summation in (5.10) is carried over the set \( \Omega \) of eigenvectors \( \psi_i \) associated with nonzero eigenvalues. From (5.10) and the nonnegativity of \( M \) we deduce that
\[
w_i^T M w_i = 0, \quad \forall w_i \in \Omega
\]
and that the whole set \( \Omega \) of eigenvectors belongs to the null space of \( M \). Hence (5.9) implies that
\[
\hat{R} M = \sum_{i=1}^p \psi_i w_i^T M = O
\]
where \( O \) is a null matrix of order \( K \times K \). It follows from (5.7)-(5.12) that
\[
\text{var} \{ \hat{f}_{\text{QNAR}} \} = \text{var} \{ \hat{f}_{\text{GAR}} \}.
\]
However, if \( \hat{R} \) is positive definite all the \( \psi_i \)'s are positive and the set \( \Omega \) in (5.11) constitutes a complete orthonormal basis. Hence, due to (5.11) the matrix \( M \) becomes a null matrix. Thus \( f_{\text{QNAR}} \) exists only when \( \hat{R} \) is nonnegative definite and condition (5.11) is valid. Assume that the matrix \( \hat{R} \) has \( q < K \) positive eigenvalues \( \psi_1, \psi_2, \ldots, \psi_q \) and that their associated eigenvectors are normalized to unit norm. It can be shown that \( f_{\text{QNAR}} \) is expressed as
\[
f_{\text{QNAR}}(X) = (K - q)^{-1} X^T \left[ I - \sum_{i=1}^q \psi_i \psi_i^T \right] X
\]
and its variance is given by
\[
\text{var} \{ \hat{f}_{\text{QNAR}} \} = (K - q)^{-1} \left[ \text{tr}^2 \{ \Gamma \} + \text{tr} \{ \Gamma \} \right].
\]
Clearly, the variance increases as the number of positive eigenvalues of \( \hat{R} \) increases. Evidently, \( q \) is a measure of the received signal bandwidth and also represents the dimension of the signal subspace. For \( f_{\text{QNAR}} \) to be invertible it is necessary that \( K \) be greater than or equal to \( q + N \).

It has been pointed out that if \( \hat{R} \) is positive definite, \( f_{\text{QNAR}} \) does not exist. In this case we define a quasi-NAR covariance estimator in a manner very similar to the QNAR power estimator defined in [7]. The QNAR covariance estimator is the member of the class of matrices (2.1) that simultaneously minimizes the variance in absence of signal and the sum of element biases in presence of signal subject to the usual constraint of being unbiased in absence of signal. Thus the QNAR covariance estimation problem is a biobjective optimization problem [9] formulated mathematically as
\[
\min_{\{M^2\} \geq 1} \left\{ \text{tr} \{ \Gamma \} + \text{tr} \{ \Gamma \} \right\}, \quad N^2 \xi \text{tr} \{ \hat{R}, M \}.
\]
Following the same method of solution outlined in [7], we finally get
\[
f_{\text{QNAR}}(X) = (K - \psi_{\text{max}} - 1)^{-1} X^T \left[ I - \psi_{\text{max}}^{-1} \hat{R} \right] X
\]
and
\[
E_{\{f_{\text{QNAR}} \}} = \Gamma + \frac{1 - \psi_{\text{max}}^{-1}}{K - \psi_{\text{max}}^{-1}} \text{tr} \{ \hat{R} \} \mu \mu^T
\]
\[
\text{var} \{ \hat{f}_{\text{QNAR}} \} = \text{tr} \left\{ \left[ I - \psi_{\text{max}}^{-1} \hat{R} \right] \right\} \text{tr} \{ \Gamma \} + \text{tr} \{ \Gamma \}
\]
\[
+ 2 \xi \text{tr} \left\{ \hat{R}, \left[ I - \psi_{\text{max}}^{-1} \hat{R} \right] \right\} \left[ N \text{tr} \{ \Gamma \} + u^T \Gamma u \right].
\]
It can easily be verified that both the bias and the variance of \( f_{\text{QNAR}} \) increase with the decrease of the eigenvalue spread of \( \hat{R} \); i.e., with an increase of the received signal frequency bandwidth. In the limiting case, \( \hat{R} = K^{-1} \). \( f_{\text{QNAR}} \) does not exist. The invertibility of \( f_{\text{QNAR}} \) necessitates that \( K \) be not less than \( N + 1 \).

Finally, it is noteworthy that the existence of \( f_{\text{QNAR}} \) implies the existence of \( f_{\text{GAR}} \). A comparison of the two estimators on a total mean-square error basis assuming signal present indicates the preference of \( f_{\text{QNAR}} \) at weak SNR and of \( f_{\text{GAR}} \) at large SNR. This can easily be proved using (5.15), (5.16), and (5.19).

REFERENCES


On the Number of Signals Whose Directions Can Be Estimated by an Array

Benjamin Friedlander and Anthony J. Weiss

Abstract—We investigate the dependence of direction finding accuracy and sensitivity on the number of signals and their separation. Using the Cramer–Rao lower bound on the direction-of-arrival estimates, we show that accuracy degrades rapidly with the number of signals, if the signals are separated by less than a beamwidth. Using results on the sensitivity of the direction estimates to calibration errors, we show that the sensitivity to calibration errors increases rapidly with the number of signals, if the signals are separated by less than a beamwidth.

Manuscript received August 24, 1990; revised December 20, 1990. This work was supported by the Army Research Office under Contract DAAL00-89-C-0007 sponsored by the U.S. Army Communications Electronics Command, Center for Signal Warfare.

B. Friedlander is with Signal Processing Technology, Ltd., Palo Alto, CA 94303 and with the University of California, Davis, CA 95616.

A. J. Weiss is with the Department of Electrical Engineering, Tel Aviv University, Tel Aviv, 59978, Israel.

IEEE Log Number 9144727.
I. INTRODUCTION

The problem of estimating the directions of arrival (DOA’s) of narrow-band signals received by an array of sensors has received considerable attention in recent years. An important related question is: how many sources can be uniquely resolved by a given array? This problem was addressed by several authors [1]–[3]. One of the conclusions from their work is that an N-element array can uniquely resolve \( N - 1 \) uncorrelated signals. (We are interested here only in the uncorrelated case. We note, however, that results are available for the correlated case as well, see, e.g., [2].)

The purpose of this note is to point out that the accuracy of the DOA estimates for closely spaced signals (i.e., signals spaced less than a beamwidth apart) degrades rapidly as the number of signals increases. Similarly, the sensitivity to calibration errors increases rapidly with the number of signals. Thus, it may not be practically possible to resolve more than a fraction of \( N - 1 \) theoretically resolvable signals, unless the signals are relatively well separated.

In the next section we study the Cramer–Rao lower bound on the variance of the estimates of the signal directions of arrival. This bound indicates the best achievable direction finding accuracy. By computing the bound for a few cases we get some insight into the dependence of the direction estimation accuracy on the number and separation of the signals.

We also study the sensitivity of the DOA estimates to calibration errors, using results we developed in [5], [6]. We define a sensitivity parameter \( \sigma_{\text{DOA}} \), which indicates the direction error caused by errors in the calibration of the array. In other words, we assume that the true array manifold differs from the array manifold used by the direction finding system. In particular, we considered errors due to gain and phase miscalibration, and errors due to imprecise knowledge of the array element locations. In the next section we evaluate only the phase sensitivity parameter. This parameter is the ratio of the direction errors induced by phase miscalibration, to the root mean square of the phase errors (see [5] for a more precise definition). By computing this sensitivity parameter for a few cases we get some insight into the dependence of the direction estimation errors due to miscalibration, on the number and separation of the signals.

II. SOME EXAMPLES

The following examples depict the standard deviation of the DOA estimates computed using the Cramer–Rao lower bound presented in [4] for a case of a linear uniformly spaced array (this bound is computed under the assumption that neither the signal covariance matrix nor the noise variance are known). In all cases the sources are assumed to have equal power, and the signal-to-noise ratio for each source was 10 dB. We have assumed arbitrarily that the DOA estimates are based on 100 snapshots. The results can be easily scaled for any other number of snapshots. For \( K \) snapshots the standard deviations in these figures should be multiplied by \( \sqrt{100/K} \). (Thus, it is best to read these figures in a relative sense, i.e., to look at the ratios between the values corresponding to the different curves in a given figure, since these ratios are independent of \( K \).) Each curve depicts the result for a different spacing \( \Delta \theta \) of the signals. We considered spacings of 1.25, 1.00, 0.75, 0.50, and 0.25 BW, where BW denotes the approximate beamwidth of the array, defined by \( BW = \lambda/d \), where \( \lambda \) is the wavelength, and \( d \) is the array aperture. The signal DOA’s are \( 0^\circ, \pm \Delta \theta, \pm 2\Delta \theta, \cdots \). The figures depict the standard deviation of the direction estimate for the signal at 0° only.

Examination of Figs. 1–3 reveals that the estimation error depends very weakly on the number of signals as long as the signal
Fig. 4. The ratio of standard deviation of the DOA estimation errors to the standard deviation for one signal, versus the number of signals, for different signal separations. Linear uniformly spaced array with 16 elements, \( \lambda/2 \) spacing. SNR = 10 dB, 100 snapshots, BW = 7.6°.

Fig. 5. The phase sensitivity parameter \( \sigma_{\text{DOA}} \) versus the number of signals, for different signal separations. Linear uniformly spaced array with 8 elements, \( \lambda/2 \) spacing.

Fig. 6. The phase sensitivity parameter \( \sigma_{\text{DOA}} \) versus the number of signals, for different signal separations. Linear uniformly spaced array with 16 elements, \( \lambda/2 \) spacing.

Fig. 7. The phase sensitivity parameter \( \sigma_{\text{DOA}} \) versus the number of signals, for different signal separations. Circular uniformly spaced array with 16 elements, \( \lambda/2 \) spacing.

Fig. 8. The phase sensitivity parameter \( \sigma_{\text{DOA}} \) versus the number of signals, for different signal separations. Linear uniformly spaced array with 16 elements, \( \lambda/4 \) spacing.

separation is not smaller than the beamwidth. However, for signal separations of half a beamwidth or less, the error increases very rapidly with the number of signals. To see this more clearly, we present in Fig. 4 the results of Fig. 3, normalized by the standard deviation of the single signal case. In other words, the curves in Fig. 4 depict the increase in the standard deviation of the estimation error relative to the case when only one signal is present. The corresponding figures for arrays with 8 and 32 elements are very similar and will not be shown.

Note that even for a modest number of signals, say 4, the standard deviation increases by more than two orders of magnitude compared to the case of one signal (or for 4 widely separated signals), for a quarter-beamwidth separation. In the case of 5 signals the increase is by four orders of magnitude.

Figs. 5-8 depict the phase sensitivity parameter \( \sigma_{\text{DOA}} \) as a function of the number of signals, for the same signal spacings considered before, computed using the equations in [5]. We note that the sensitivity parameter depends very weakly on the number of signals for signal separations on the order of a beamwidth or greater. However, for signal separation smaller than half a beamwidth, sensitivity increases rapidly with the number of signals. For example, the sensitivity for 5 signals separated by a quarter of a beamwidth, is a thousand times greater than for the one signal case (or for 5 widely separated signals).

These examples illustrate the inherent ill conditioning of the di-
Two Convergence Theorems on Deterministic Properties of Median Filters

Zi-Jun Gan and Ming Mao

Abstract—In this correspondence, two convergence theorems on deterministic properties of median filters are obtained. The results can be used to analyze the root and recurrent signals of median filters, and to estimate the convergence rates. The concept of convergence point introduced here may be a useful new tool for analyzing the median filters.

I. INTRODUCTION

Median filters can smooth noisy signals efficiently while maintaining the signal edges. Since median filtering is a nonlinear operation, the analysis techniques which are used in linear operation do not apply.

Root signals are signals that are invariant to filtering. They indicate which kinds of signal pass the filter unaltered; in other words, they have a similar role to that of the passband of a linear filter. The analysis of the convergence behavior of median filters is an important part of median filter theory. It is an efficient way to analyze the structures of root signals and the properties of the output signals of median filters. Some works related to convergence to roots have been reported by several authors, e.g., [1]–[6].

Although any actual obtained signal is of finite length, it is necessary to understand the nature of infinitely long signals to explain some phenomenon in the finite case. For example, a finite signal may contain a section of an infinitely long oscillatory root signal. If the section is very long, the number of filtering terms needed to remove it must be prohibitively large.

In this correspondence, a new concept called convergence point is introduced. Using this concept, two convergence theorems for arbitrary signals are obtained. The number of convergence points in a signal can be regarded as a measure of stability of the signal to the operation of median filters.

II. DEFINITIONS

For a window of length $2k + 1$, let $\{x(n)\}$ and $\{y(n)\}$ be the input and output of the median filter, respectively. Then $y(n) = \text{median}(x(n - k), \ldots, x(n), \ldots, x(n + k))$.

Definition 1: For the median filter of length $N = 2k + 1$, let $t \in \{1, k + 1\}$. Then $i$ is called a $t$-CP ($t$-class convergence point) of $\{x(n)\}$ if the following relations are satisfied:

1: $x(i + s - t + 1) = \cdots = x(i) = \cdots = x(i + s)$

2: $x(i + s - t + 1) \leq x(i + s - t), \ldots, x(i + s - k)$ and $x(i + s) = x(i + s + 1), \ldots, x(i + s + 1 + k - t)$

or 2: $x(i + s - t + 1) \leq x(i + s - t), \ldots, x(i + s - k)$ and $x(i + s) = x(i + s + 1), \ldots, x(i + s + 1 + k - t)$

where $s \in \{0, t - 1\}$.

Definition 2: $i$ is called a GP (good point) if $i$ is a 1-CP or $(k + 1)$-CP. Otherwise, $i$ is called a BP (bad point). See Fig. 1 as an example to explain definitions 1 and 2.

III. THEOREMS

For a window of length $2k + 1$, let $\{x(n)\}$ and $\{y(n)\}$ be the input and output of the median filter, respectively.

Theorem 1: If $i$ is a GP of $\{x(n)\}$, then $i - 1$, $i$, and $i + 1$ are GPs of $\{y(n)\}$.

Proof: It is sufficient to prove that $i$ and $i + 1$ are GPs of $\{y(n)\}$.

Case 1: $i$ is a 1-CP of $\{x(n)\}$. Suppose $x(i - k), \ldots, x(i - 1) \leq x(i) \leq x(i + 1), \ldots, x(i + k)$. It is obvious that $i$ is a 1-CP of $\{y(n)\}$, i.e., $y(i - k), \ldots, y(i - 1) \leq y(i) \leq y(i + 1), \ldots, y(i + k)$. Notice that $y(i + 1) = \text{median}(x(i - k + 1), \ldots, x(i + k + 1))$ and $x(i - k - 1) \leq x(i) \leq x(i + 1), \ldots, x(i + k + 1)$.

One can see that all of $(i + 1), x(i + 2), \ldots, x(i + k)$ and one of $i$ and $x(i + k + 1)$ must be $\geq y(i + 1)$. For any $r \in \{2, k\}$, let the center point in the filter window be the point $i + r$, then $x(i), \ldots, x(i + k + 1)$ must all be in the window. Thus there are at least $k + 1$ points which are $\geq y(i + 1)$ in the window. Hence $y(i + 2), \ldots, y(i + k)$ must all be $\geq y(i + 1)$. Following the above observations, if $y(i + k + 1) \geq y(i + 1)$, $i + 1$ is a 1-CP. Otherwise, if $y(i + k + 1) < y(i + 1)$, it is clear that all of $x(i + k + 1), \ldots, x(i + k + 2)$ and $x(i) \leq y(i + 1)$. Thus, $y(i) = y(i + 1) = \cdots = y(i + k) = x(i)$, and $i + 1$ is a $(k + 1)$-CP of $\{y(n)\}$.

Case 2: $i$ is a $(k + 1)$-CP of $\{x(n)\}$, i.e., $x(i - k - s) = \cdots = x(i + s), s \in \{0, k\}$. If $s \neq 0$, it is trivial, because $y(i - k - s) = \cdots = y(i - k + 1) = \cdots = y(i + s)$.

This leaves the case $s = 0$, i.e., $x(i - k) = \cdots = x(i)$. Then, if $y(i + 1) = y(i)$, it is obvious that $i + 1$ is a $(k + 1)$-CP of $\{y(n)\}$.

Otherwise, if $y(i + 1) \neq y(i)$, there is no loss in generality to assume that $y(i + 1) > y(i) = x(i)$. We can see that all of $x(i)$, $x(i + 1), \ldots, x(i + k)$, $\ldots, x(i + k + 1)$, $\ldots, x(i + k + 2)$ fall in the window.