

# Rank Selection Algorithm with a Cross-Correlation Coefficient for the Multistage Wiener Filter

# Junichiro Suzuki<sup>1</sup>, Yoshikazu Shoji<sup>2</sup>, Hiroyoshi Yamada<sup>3</sup>,  
Yoshio Yamaguchi<sup>3</sup>, and Masahiro Tanabe<sup>1</sup>

<sup>1</sup> Komukai Operations Radar & Sensor Systems Engineering Dept., Toshiba Corporation  
1, Komukai Toshiba-cho, Saiwai-ku, Kawasaki, Kanagawa, 212-8581 Japan,  
{junichiro.suzuki, masahiro.tanabe}@toshiba.co.jp

<sup>2</sup> Graduate School of Science and Technology, Niigata University  
8050, Ikarashi 2-no-cho, Niigata, 950-2181 Japan, shoji@wave.ie.niigata-u.ac.jp

<sup>3</sup> Faculty of Engineering, Niigata University  
8050, Ikarashi 2-no-cho, Niigata, 950-2181 Japan, {yamada, yamaguchi}@ie.niigata-u.ac.jp

## Abstract

*The multistage Wiener filter(MWF) is a reduced rank algorithm for the space-time adaptive processing(STAP). In environments with low sample support, the performance of the MWF is better than the full rank Wiener filter, but MWF stage analysis has to be operated at the optimum rank. When it is not operated at the optimum rank, it leads performance degradation. In this paper we introduce a new rank selection algorithm with a cross-correlation coefficient that can stop at the optimum rank. Since the proposed algorithm does not have to calculate the evaluation function to use a scalar coefficient of weight derivation process, the computational complexity is more advantageous than other algorithm. Here, we evaluate its performance by simulation examples, and show the effectiveness.*

## 1. INTRODUCTION

This paper discusses a result to evaluate performance of rank selection algorithm with a cross-correlation coefficient for the multistage Wiener filter(MWF). The MWF, described in [1], is an adaptive processing technique to use at low sample support environments. MWF weight calculation connects scalar Wiener filters to multistage. Stopping at the optimum stage, the MWF can significantly outperform the full rank Wiener filter.

There are two types of method to stop MWF weight calculation. One method applies diagonal loading[2] to reduce influence of a highly stage, and the calculation is run out to full stage. This method should be calculate to a full  $N$  stages, where  $N$  is the number of degrees-of-freedom(DOF), and much computational complexity is necessary. To stop MWF weight calculation at the optimum rank is the another method. It should calculate only a  $r(< N)$  stages for rank reduction method and can reduce computational complexity. Therefore we consider that a later method is very effective for MWF weight calculation.

As a method to judge the optimum rank to be it, the most common rank revealing decomposition is the eigenvalue decomposition(EVD). The EVD well uses the Akaike information criterion(AIC)[3] and the minimum description length(MDL)[4] for rank selection algorithm. As for the rest, a well known method is called the white noise gain constraint(WNGC)[5]. It has been observed that the size of the weight vector norm, grows in response to mismatch errors. Therefore thresholding  $\|w\|^2$  can be used as a rank selection algorithm. For other approaches, the method of conjugate gradients can be used to implement the MWF[5]. This is the hybrid method that uses sample mean square error(SMSE) and the Krylov power function  $\mathcal{P}_r(0)$ . Thus these rank selection algorithms are according to require auxiliary knowledge, a calculation of an evaluation function and a judgment by the threshold.

The MWF provides a stage-by-stage decomposition of the Wiener filter solution. At each stage a scalar weight is computed for the MWF which weights the contribution of that stage in removing undesired signal from the desired signal. The first stage of the MWF is responsible for removing interference most correlated with the residual undesired signal that survived all previous stages. Each successive stage acts upon a smaller residual undesired signal quantity. Hence the calculation of additional stages becomes counterproductive after reached the optimum stage. It solves this problem and makes use of a characteristic of a weight calculation of the MWF, we propose rank selection algorithm with a cross-correlation coefficient of weight derivation process. The proposed algorithm does not have to calculate the evaluation function, the computational complexity is more advantageous than other algorithm. In using proposed algorithm for MWF rank selection, MWF recursion solution can be stopped automatically at the optimum stage. And we show that it is good performance than the WNGC by simulation examples.

In this paper, we introduce the space-time signal models and the MWF algorithm in Section 2, Secondly, we describe the proposed rank selection algorithm in Section 3, and present

a simulation example in Section 4. Lastly, we concluded our result in Section 5.

## 2. BACKGROUND

### A. Space-Time Signal Models

The spatial and temporal steering vectors  $\mathbf{s}_{\bar{\theta}}$ ,  $\mathbf{s}_{\bar{f}_d}$  are defined as follows

$$\mathbf{s}_{\bar{\theta}} = [1, e^{j\bar{\theta}}, \dots, e^{j(N-1)\bar{\theta}}]^T, \quad (1)$$

$$\mathbf{s}_{\bar{f}_d} = [1, e^{j\bar{f}_d}, \dots, e^{j(M-1)\bar{f}_d}]^T, \quad (2)$$

where  $T$  is transpose,  $N$  is the number of elements,  $M$  is the number of pulses,  $\bar{\theta}$  is the normalized spatial angle,  $\bar{f}_d$  is the normalized doppler frequency. In addition the azimuth angle of the signal  $\theta$  and the doppler frequency  $f_d$  are defined as follows

$$\bar{\theta} = \frac{2\pi d}{\lambda} \sin\theta, \quad (3)$$

$$\bar{f}_d = 2\pi \frac{f_d}{f_p}, \quad (4)$$

where  $\lambda$  is the radar wavelength,  $d$  is the element spacing,  $f_p$  is the pulse repetition frequency (PRF). We combine the spatial and temporal steering vectors into the  $N \times M$  space-time steering vector  $\mathbf{s}_{\bar{\theta}, \bar{f}_d}$  for STAP as follows

$$\mathbf{s}_{\bar{\theta}, \bar{f}_d} = \mathbf{s}_{\bar{f}_d} \otimes \mathbf{s}_{\bar{\theta}}, \quad (5)$$

where  $\otimes$  is the Kronecker product.

Clutter is modeled as a zero-mean Gaussian random process. The space-time clutter covariance matrix  $\mathbf{R}_c$  is given by

$$\mathbf{R}_c = \sum_{i=1}^{N_c} \sigma_c^2(i) [\mathbf{v}_{\bar{f}_i} \mathbf{v}_{\bar{f}_i}^H] \otimes [\mathbf{v}_{\bar{\theta}_i} \mathbf{v}_{\bar{\theta}_i}^H], \quad (6)$$

where  $H$  is conjugate transpose,  $N_c$  is the number of clutter patches uniformly distributed in azimuth,  $\sigma_c^2(i)$  is the power of in the  $i^{\text{th}}$  clutter patch,  $\mathbf{v}_{\bar{\theta}_i}$  and  $\mathbf{v}_{\bar{f}_i}$  are the spatial and temporal steering vectors associated with the  $i^{\text{th}}$  clutter patch.

The space-time clutter ridge for different velocity/PRF conditions  $\beta$  can be given by

$$\beta = \frac{2v_p}{d \cdot f_p}, \quad (7)$$

where  $v_p$  is the platform velocity.

Here, in the case of a side-looking uniform linear array (ULA) with half-wavelength element spacing and  $\beta$  is an integer, the clutter covariance matrix rank holds the ‘‘Brennan’s rule [6], [7]’’;

$$\text{rank}\{\mathbf{R}_c\} \simeq \lceil N + \beta \times (M - 1) \rceil, \quad (8)$$

where  $\lceil \cdot \rceil$  is the ceiling operator (round to nearest largest integer).

Jammer is uncorrelated pulse-to-pulse, therefore jammer covariance matrix  $\mathbf{R}_j$  can be given by

$$\mathbf{R}_j = \sigma_j^2 [\mathbf{I}_M \otimes [\mathbf{v}_{\theta_j} \mathbf{v}_{\theta_j}^H]], \quad (9)$$

where  $\sigma_j^2$  is the jammer power,  $\mathbf{I}_M$  is the  $M \times M$  identity matrix,  $\mathbf{v}_{\theta_j}$  is the  $j^{\text{th}}$  jammer spatial steering vector.

Noise refers to receiver thermal noise and is modeled as a zero-mean Gaussian random process. It is uncorrelated both spatially and temporally, therefore it can be characterized by the following covariance matrix  $\mathbf{R}_n$ .

$$\mathbf{R}_n = \sigma_n^2 [\mathbf{I}_M \otimes \mathbf{I}_N], \quad (10)$$

where  $\sigma_n^2$  is the noise power,  $\mathbf{I}_N$  is the  $N \times N$  identity matrix.

The undesired signals (clutter, jammer and noise) for STAP are all assumed to be uncorrelated. Therefore, the total interference-plus-noise covariance matrix  $\mathbf{R}$  can be as the sum of the individual covariance matrices

$$\mathbf{R} = \mathbf{R}_c + \sum_{i=1}^J \mathbf{R}_j(i) + \mathbf{R}_n, \quad (11)$$

where  $J$  is the number of jammers.

The standard method of estimating the covariance matrix is by constructing the sample covariance matrix  $\mathbf{R}$  as follows

$$\mathbf{x}_i(k) = \mathbf{R}^{\frac{1}{2}} \frac{1}{\sqrt{2}} \mathbf{a}_k, \quad (12)$$

$$\mathbf{R} = \frac{1}{K} \sum_{k=1}^K \mathbf{x}_i(k) \mathbf{x}_i(k)^H, \quad (13)$$

where  $\mathbf{x}_i(k)$  is the  $k^{\text{th}}$  training sample at the  $i^{\text{th}}$  element,  $K$  is the total number of training samples and the  $N \times M$  matrix  $\mathbf{a}_k$  are generated from a complex vector, zero mean, unit variance, Gaussian distribution.

For example, a well known algorithm, the maximum signal-to-noise ratio (MSN) [8] algorithm gives the weight vector  $\mathbf{w}_{msn}$  as follows

$$\mathbf{w}_{msn} = \frac{\mathbf{R}^{-1} \mathbf{s}}{\mathbf{s}^H \mathbf{R}^{-1} \mathbf{s}}, \quad (14)$$

where  $\mathbf{s}$  is the steering vector.

### B. The Multistage Wiener Filter

The MWF is a signal dependent reduced rank processing, it is able to operate at fewer adaptive DOF than other reduced rank algorithms (e.g., principal components, the cross spectral metric) [1]. The filter structure of the MWF is shown in Fig. 1, and the recursion equations are summarized in TABLE 1, where  $E[\cdot]$  denotes ensemble average,  $\mathbf{r}_{x_i d_i}$  is the cross-correlation vectors,  $d_i(k)$  is the desired signals,  $\delta_i$  is the magnitude of cross-correlation vectors,  $\mathbf{h}_i$  is the direction of cross-correlation vectors,  $\mathbf{B}_i$  is the blocking matrices,  $\text{null}(\cdot)$  means  $\mathbf{B}_i \mathbf{h}_i = \mathbf{0}$ ,  $w_i$  is the scalar weights, and  $\epsilon_i(k)$  is the error signals.

Now the adaptive weight vector  $\mathbf{w}_{mwf}$  for the MWF solution can be represented as follows

$$\mathbf{w}_{mwf} = \mathbf{s} - w_1 \mathbf{B}_0^H \mathbf{h}_1 + w_1 w_2 \mathbf{B}_0^H \mathbf{B}_1^H \mathbf{h}_2 - w_1 w_2 w_3 \mathbf{B}_0^H \mathbf{B}_1^H \mathbf{B}_2^H \mathbf{h}_3 + \dots, \quad (15)$$

where  $\mathbf{B}_0$  is the  $\text{null}(\mathbf{s})$ .

As I mentioned previously, rank reduction is obtained by truncation at the arbitrary  $r (< N)$  stages.

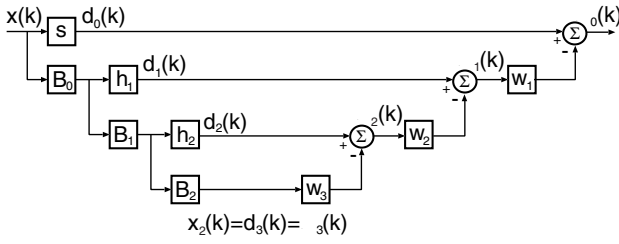


Fig. 1: The Filter Structure of the MWF for 3 Stages.

TABLE 1: RECURSION EQUATIONS FOR THE MWF.

Forward Recursion	Backward Recursion
$\mathbf{r}_{x_i d_i} = E[\mathbf{x}_i(k)d_i^*(k)]$	$\xi_N = E[ \mathbf{x}_{N-1}(k) ^2]$
$\delta_{i+1} = \sqrt{\mathbf{r}_{x_i d_i}^H \mathbf{r}_{x_i d_i}}$	$\delta_N(k) = \mathbf{r}_{x_{N-1} d_{N-1}}$
$\mathbf{h}_{i+1} = \mathbf{r}_{x_i d_i} / \delta_{i+1}$	$\xi_i = \sigma_{d_i}^2 - \delta_{i+1}^2 / \xi_{i+1}$
$\mathbf{B}_{i+1} = \text{null}(\mathbf{h}_{i+1})$	$= E[ \epsilon_i ^2]$
$d_{i+1}(k) = \mathbf{h}_{i+1}^H \mathbf{x}_i(k)$	$w_i = \delta_i / \xi_i$
$\sigma_{d_i}^2 = E[ d_i(k) ^2]$	$\epsilon_i(k) = d_i(k) - w_{i+1} \epsilon_{i+1}(k)$
$\mathbf{x}_{i+1}(k) = \mathbf{B}_{i+1} \mathbf{x}_i(k)$	
$\mathbf{x}_{N-1}(k) = \mathbf{d}_N(k) = \epsilon_N(k)$	

### 3. PROPOSED RANK SELECTION ALGORITHM

In MWF solution, a scalar coefficient  $\xi_i$  and a scalar weight  $w_i$  of a weight derivation process have a very important meaning. Recall from TABLE 1 that scalar coefficients  $\xi_i$  and a scalar weight  $w_i$  are defined as follows

$$\xi_i = \sigma_{d_i}^2 - \delta_{i+1}^2 / \xi_{i+1} = E[|\epsilon_i|^2], \quad (16)$$

$$w_i = \delta_i / \xi_i, \quad (17)$$

where (16) shows a scalar coefficient  $\xi_i$  is the expected value of the magnitude squared error at the  $i^{\text{th}}$  stage.

Because MWF stage analysis is calculated to remove undesired signal from the desired signal at each stage, a value of  $\xi_i$  which shows the magnitude of the squared error is almost zero value when reached the optimum rank. In addition, a value of coefficient  $w_i$  becomes bigger when a value of coefficient  $\xi_i$  becomes almost zero as shown in (17). Therefore, if each coefficient is able to judge by optimum threshold, scalar coefficients  $\xi_i$  and  $w_i$  can be used as rank selection criteria, however it is necessary to determine the threshold by simulation same as other algorithms.

Now we introduce a scalar coefficient  $\eta_i$  which multiplied  $\xi_i$  by  $w_i$  as follows

$$\begin{aligned} \eta_i &= \xi_i \cdot w_i, \\ &= \delta_i, \end{aligned} \quad (18)$$

where  $\eta_i$  is equal to the magnitude of cross-correlation vectors  $\delta_i$  as shown in (17).

Recall from TABLE 1 that the direction of cross-correlation vectors  $\mathbf{h}_i$ , the cross-correlation vectors  $\mathbf{r}_{x_i d_i}$ , and the desired

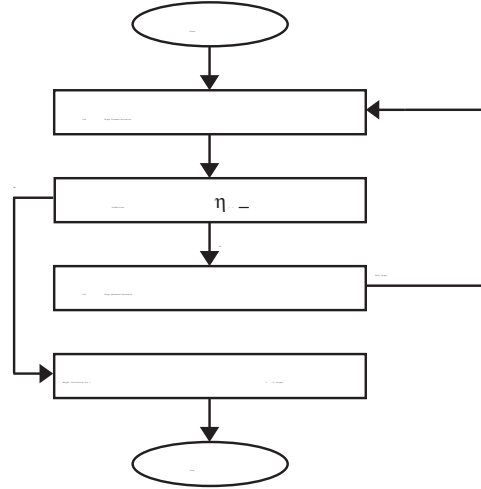


Fig. 2: The Flow Chart of Proposed Rank Selection Algorithm.

signals  $\delta_i$  are defined as follows

$$\mathbf{h}_{i+1} = \mathbf{r}_{x_i d_i} / \delta_{i+1}, \quad (19)$$

$$\mathbf{r}_{x_i d_i} = E[\mathbf{x}_i(k)d_i^*(k)], \quad (20)$$

$$d_{i+1}(k) = \mathbf{h}_{i+1}^H \mathbf{x}_i(k). \quad (21)$$

By (19), the magnitude of cross-correlation vector  $\delta_i$  is represented as follows

$$\delta_i = \mathbf{h}_i^H \mathbf{r}_{x_{i-1} d_{i-1}}. \quad (22)$$

Substituting (20) and (21) into (22) leads to

$$\begin{aligned} \delta_i &= E[\mathbf{h}_i^H \mathbf{x}_{i-1}(k)d_{i-1}^*(k)], \\ &= E[d_i(k)d_{i-1}^*(k)], \end{aligned} \quad (23)$$

where we see that  $\delta_i$  is the cross-correlation coefficient of  $i^{\text{th}}$  and  $(i-1)^{\text{th}}$  stage desired signals.

Because a cross-correlation coefficient is a statistical relationship between two data, it is thought that there is similarity for two data when it is almost 1. And so the magnitude of cross-correlation vectors  $\delta_i$  denotes the cross-correlation coefficient of the desired signals between  $i^{\text{th}}$  and  $(i-1)^{\text{th}}$  stage, the value of  $\eta_i (= \delta_i)$  becomes smaller than 1 when MWF recursion reached the optimum rank. Therefore, MWF weight solution should stop at the  $(i-1)^{\text{th}}$  stage when  $\eta_i$  (the value of  $i^{\text{th}}$  stage) is less than 1.

So we propose a condition of rank selection as follows

$$\eta_i = \delta_i \leq 1. \quad (24)$$

In addition, because a cross-correlation coefficient  $\eta_i (= \delta_i)$  is calculated by MWF forward recursion (see, TABLE 1), there is no computational complexity for proposed rank selection algorithm and be able to stop automatically at the optimum rank when (24) satisfied a condition. Here, the flow chart of proposed rank selection algorithm shows in Fig. 2.

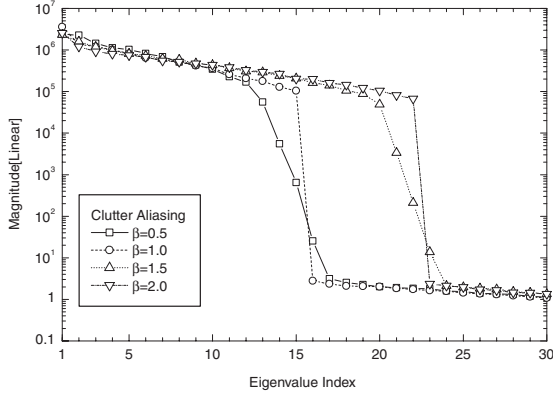


Fig. 3: Eigenvalue Index vs. Magnitude of Eigenvalue.

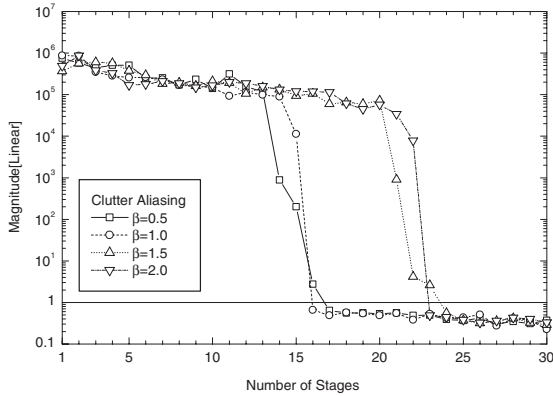


Fig. 4: Number of Stages vs. Magnitude of  $\eta_i$ .

#### 4. SIMULATION EXAMPLE

In this section, we evaluate proposed rank selection algorithm by simulation examples. We determine the magnitude of eigenvalue and  $\eta_i$  under the condition that  $N = 8$  elements ULA with half-wavelength element spacing,  $M = 8$  pulses coherent processing interval, the training data  $K = 64$  samples, noise floor is 0dB, clutter-to-noise ratio(CNR) is 30dB on each element and each pulse, and conditions of clutter aliasing  $\beta = 0.5, 1.0, 1.5$  and  $2.0$ .

From Fig. 3, it is clear that we can estimate the clutter covariance matrix rank is 16, 15, 23 and 22 because the magnitude of 17<sup>th</sup>, 16<sup>th</sup>, 23<sup>th</sup> and 22<sup>th</sup> eigenvalue are smaller than the previous stages(those are near to noise level). Fig. 4 shows the magnitude of  $\eta_i$  tends to be similar to the magnitude of eigenvalue. As in the case of the eigenvalue, we can estimate the clutter covariance matrix rank is 16, 15, 23 and 22 by the proposed algorithm.

Next, we run 1000 Monte Carlo trials for each  $\beta$  changed from 0.1 to 2.0 every 0.1 steps, and the other simulation parameters are unchanged as the previous one. Here, we

TABLE 2: OCCURRENCE OF THE PROPOSED METHOD.

$\beta$	$\Delta\text{rank}$								
	-5	-4	-3	-2	-1	0	1	2	3
0.1	0	3	258	515	224	0	0	0	0
0.2	0	0	0	0	13	987	0	0	0
0.3	0	0	3	726	255	16	0	0	0
0.4	0	0	0	0	11	989	0	0	0
0.5	0	0	0	0	6	994	0	0	0
0.6	0	0	0	616	354	30	0	0	0
0.7	0	0	0	0	1	664	335	0	0
0.8	0	0	0	0	1	999	0	0	0
0.9	0	0	0	0	5	995	0	0	0
1.0	0	0	0	0	2	998	0	0	0
1.1	0	0	0	0	2	873	125	0	0
1.2	0	0	0	0	9	795	196	0	0
1.3	0	0	0	0	0	51	949	0	0
1.4	0	0	0	0	0	1000	0	0	0
1.5	0	0	0	0	33	962	5	0	0
1.6	0	0	0	0	0	235	765	0	0
1.7	0	0	0	0	0	1000	0	0	0
1.8	0	0	0	0	288	712	0	0	0
1.9	0	0	0	122	814	64	0	0	0
2.0	0	0	0	0	0	1000	0	0	0

evaluate the following performance parameters.

$$\Delta\text{rank} = \text{rank}\{\text{best}\} - \text{rank}\{\text{prop.}\}, \quad (25)$$

$$\text{normalized SINR} = \frac{\mathbf{w}_{mvd}^H \mathbf{R} \mathbf{w}_{mvd}}{\mathbf{w}_{mwf}^H \mathbf{R} \mathbf{w}_{mwf}}, \quad (26)$$

where  $\text{rank}\{\text{best}\}$  is the rank of the best normalized signal-to-interference-plus-noise ratio(SINR),  $\text{rank}\{\text{prop.}\}$  is the rank of the proposed rank selection algorithm,  $\mathbf{w}_{mvd}$  is derived by minimizing the interference plus noise power out of the beamformer, while retaining the desired signal without distortion.

TABLE 2 and 3 show the occurrence of the proposed algorithm and the WNGC, where the threshold level is 1dB, and Fig. 5 shows a histogram of them. We consider the proposed algorithm is superior to the WNGC because the WNGC selected the  $\text{rank}\{\text{best}\}$  about 59% whereas the proposed algorithm selected about 67%. Fig. 6 shows the performance of  $\text{rank}\{\text{best}\}$ , and superior to the WNGC.

Next, Recall from TABLE 2, we see that the proposed algorithm tends to overestimate at  $\beta = 0.1, 0.3, 0.6,$  and  $1.9$ . So we think the characteristic of normalized SINR can be improved by applying to “error loading[9]” and “diagonal loading[10]”. For example, we evaluate performance that applied “error loading” to the proposed algorithm, where the loading level is 0dB. Here, the following coefficient is used instead of (17)

$$w_i = \delta_i / \{\xi_i + \sigma_{el}^2\}, \quad (27)$$

where  $\sigma_{el}^2$  is the loading level.

Fig. 7 shows the normalized SINR of the “with error loading” and “without error loading”. Clearly, the performance improves by applying error loading to proposed algorithm. Thus good performance can be realized by the proposed algorithm.

TABLE 3: OCCURRENCE OF THE WNGC.

$\beta$	$\Delta$ rank								
	-5	-4	-3	-2	-1	0	1	2	3
0.1	0	0	0	738	262	0	0	0	0
0.2	0	0	0	0	355	645	0	0	0
0.3	0	0	0	107	856	37	0	0	0
0.4	0	0	0	0	1	996	3	0	0
0.5	0	0	0	0	68	932	0	0	0
0.6	0	0	0	0	882	118	0	0	0
0.7	0	0	0	0	0	157	842	1	0
0.8	0	0	0	0	0	999	1	0	0
0.9	0	0	0	0	2	998	0	0	0
1.0	0	0	0	0	116	884	0	0	0
1.1	0	0	0	0	0	601	399	0	0
1.2	0	0	0	0	0	598	402	0	0
1.3	0	0	0	0	0	197	803	0	0
1.4	0	0	0	0	1	999	0	0	0
1.5	0	0	0	0	0	981	19	0	0
1.6	0	0	0	0	0	159	841	0	0
1.7	0	0	0	0	0	1000	0	0	0
1.8	0	0	0	0	0	499	501	0	0
1.9	0	0	0	4	905	91	0	0	0
2.0	0	0	0	0	29	971	0	0	0

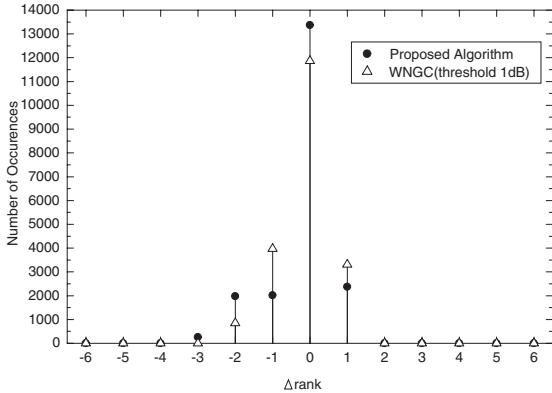


Fig. 5: Histogram of  $\Delta$ rank.

### 5. CONCLUSIONS

In this paper, we proposed a rank selection algorithm with a cross-correlation coefficient  $\eta_i$  for the MWF. The proposed algorithm has no computational complexity because  $\eta_i (= \delta_i)$  is calculated by MWF forward recursion. As example, we demonstrated characteristics of rank estimation and normalized SINR. Consequently, the proposed algorithm showed that it was good performance than the WNGC.

### REFERENCES

[1] J. S. Goldstein, I. S. Reed and L. L. Scharf, "A multistage representation of the wiener filter based on orthogonal projections," *IEEE Trans. Information Theory*, vol. IT-44, no. 7, pp. 2943-2959, Nov. 1998.  
 [2] W. F. GABRIEL, "Using spectral estimation techniques in adaptive processing antenna systems," *IEEE Trans. Antennas Propagat.*, vol. AP-34, no. 3, pp. 291-299, Mar. 1986.  
 [3] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Automatic Control*, vol. AC-19, no. 6, pp. 716-723, Dec. 1974.

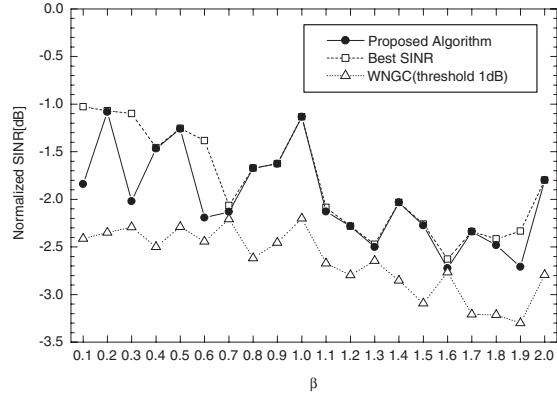


Fig. 6:  $\beta$  vs. Normalized SINR.

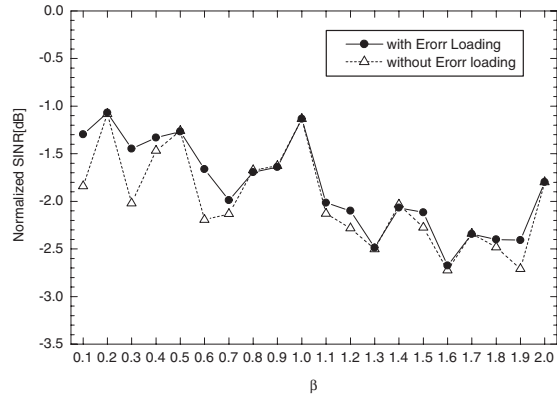


Fig. 7:  $\beta$  vs. Normalized SINR(with Error Loading and without Error Loading).

[4] J. Rissanen, "Universal coding, information, prediction, and estimation," *IEEE Trans. Information Theory*, vol. IT-30, no. 4, pp. 629-636, July. 1984.  
 [5] H. Cox, R. M. Zeskind and M. M. Owen, "Robust adaptive beamforming," *IEEE Trans. Acoust., Speech, and Signal Processing*, vol. ASSP-35, no. 10, pp. 1365-1376, Oct. 1987.  
 [6] R. Klemm, *Space-Time Adaptive Processing, principles and applications*, IEE Press, London, U.K., 2002.  
 [7] J. R. Guerci, *Space-Time Adaptive Processing for Radar*, Artech House, Norwood, MA, 2003.  
 [8] S. P. Applebaum, "Adaptive arrays," *IEEE Trans. Antennas Propagat.*, vol. AP-24, no. 5, pp. 585-598, Sept. 1976.  
 [9] J. D. Hiemstra and J. S. Goldstein, "Robust rank selection for the multistage wiener filter," in *Proc. 2002 IEEE Int. Conf. Acoust., Speech, and Signal Processing*, May 2002, vol. III, pp. 2929-2932.  
 [10] J. D. Hiemstra, M. E. Weippert, H. N. Nguyen and J. S. Goldstein, "Insertion of diagonal loading into the the multistage wiener filter," in *Proc. 2002 IEEE Sensor Array and Multichannel Signal Processing*, Aug. 2002, pp. 379-382.