ON REDUCED-RANK APPROACHES TO MATRIX WIENER FILTERS IN MIMO SYSTEMS

Guido Dietl and Wolfgang Utschick

Institute for Circuit Theory and Signal Processing
Munich University of Technology, 80333 Munich, Germany
{gudi,wout}@nws.ei.tum.de

ABSTRACT

Reduced-rank processing is a well-known strategy in order to reduce computational complexity and enhance the performance in case of low sample support. In this paper, we use the eigenspace based *Principal Component* (PC) and *Cross-Spectral* (CS) method for rank-reduction of a *matrix Wiener Filter* (WF) which estimates a signal vector instead of a scalar by minimizing the *mean square error.*

Finally, we apply the resulting filters to a frequency-flat *Multi-Input Multi-Output* (MIMO) transmission channel. Although the matrix PC algorithm is computational cheaper than the matrix CS algorithm, we show analytically that the two methods are equal if we assume i.i.d. transmit symbols and uncorrelated white Gaussian noise. Simulation results show additionally that the matrix *Multi-Stage WF* (MSWF) which approximates the WF in a Krylov subspace is partially outperformed in the considered MIMO case.

1. INTRODUCTION

The *Wiener Filter* (WF) [1] recovers an unknown signal vector from an observation vector by minimizing the *Mean Square Error* (MSE) and exploiting only second order statistics. The derivation ends up in solving the *Wiener-Hopf* equation which is computational intense for observations of high dimensionality. Since reduced-rank methods approximate the WF in a subspace, they reduce computational complexity and enhance the robustness against estimation errors of statistics due to low sample support.

In this paper, we consider the approximation of the WF in subspaces spanned by eigenvectors of the covariance matrix of the observation vector. The *Principal Component* (PC) method [2] chooses the eigenvectors corresponding to the largest eigenvalues. Nevertheless, in the case of a mixture of signals, there is no distinction between the signal of interest and the interference signal. Hence, the performance of the algorithm degrades dramatically if the interference power is larger than the power of the unknown signal. More recently, Goldstein et al. [3] introduced the *Cross-Spectral* metric as a selection criterion for the eigenvectors spanning the eigenspace. Compared to the PC method, the CS method is more robust against strong interference because it considers additionally the cross-correlation between the signal and the observation vector. Moreover, it minimizes the MSE over all eigenspace based methods with equal rank.

Our contribution is to derive a reduced-rank *matrix WF* based on the PC and CS method, respectively, and to show that the two eigenspace based methods are identical if we apply them to a frequency-flat *Multi-Input Multi-Output* (MIMO) system where the transmitted symbols are i.i.d. and the noise is uncorrelated white Gaussian. Moreover, the comparison with the matrix version of the *Multi-Stage WF* (MSWF), a reduced-rank WF developed by Goldstein et al. [4, 5] which is based on the Krylov subspace of the covariance matrix of the observation and the cross-correlation matrix between the observation and the signal vector, shows that the eigenspace based methods exceed the performance of the matrix MSWF in certain *Signal-to-Noise Ratio* (SNR) regions and have the possibility of a more flexible rank selection.

The next section briefly reviews the WF. In Section 3, we derive a matrix version of the PC and CS algorithm and finally apply them to a frequency-flat MIMO system in Section 4. Throughout the paper, the covariance matrix of a random vector $\mathbf{u}$ is $\boldsymbol{R}_{\mathsf{u}} = \mathrm{E}\left\{\mathbf{u}\mathbf{u}^{\mathrm{H}}\right\}$ and the cross-correlation matrix between the vectors $\mathbf{u}$ and $\mathbf{v}$ is $\boldsymbol{R}_{\mathsf{u,v}} = \mathrm{E}\left\{\mathbf{u}\mathbf{v}^{\mathrm{H}}\right\}$ where '$\mathrm{E}\{\cdot\}$' denotes expectation and '$(\cdot)^{\mathrm{H}}$' Hermitian.

2. WIENER FILTER

The WF [1] $\boldsymbol{W} \in \mathbb{C}^{M\times N}$ estimates the unknown signal vector $\mathbf{x}[n] \in \mathbb{C}^{M}$ from the observation vector $\mathbf{y}[n] \in \mathbb{C}^{N}$, $N \geq M$, by minimizing the MSE $\xi(\mathcal{W})$ between $\mathbf{x}[n]$ and its estimate $\hat{\mathbf{x}}[n] = \mathcal{W}\mathbf{y}[n]$, i. e.

$$\boldsymbol{W} = \arg\min_{\mathcal{W}} \xi(\mathcal{W}), \quad \text{with} \tag{1}$$

$$\xi(\mathcal{W}) = \mathrm{tr}\left\{\boldsymbol{R}_{\mathsf{x}} - 2\,\mathrm{Re}\left\{\mathcal{W}\boldsymbol{R}_{\mathsf{y,x}}\right\} + \mathcal{W}\boldsymbol{R}_{\mathsf{y}}\mathcal{W}^{\mathrm{H}}\right\}, \tag{2}$$

where the operation '$\mathrm{tr}\{\cdot\}$' denotes the trace of a matrix and '$\mathrm{Re}\{\cdot\}$' the real part of a complex variable.

The optimization leads to the Wiener-Hopf equation whose solution, the WF

$$W = R_{y,x}^{H} R_y^{-1}, \quad (3)$$

achieves the *Minimum Mean Square Error* (MMSE)

$$\xi(W) = \mathrm{tr}\left\{R_x - R_{y,x}^{H} R_y^{-1} R_{y,x}\right\}. \quad (4)$$

Note that we consider a *matrix WF* which estimates a M-dimensional signal $\mathbf{x}[n]$. If $M = 1$, the cross-correlation matrix $R_{y,x}$ shrinks to a cross-correlation vector and we end up in the *vector WF* which produces a scalar output. Considering Equation (3), it is obvious that estimating each element of $\mathbf{x}[n]$ with M parallel vector WFs produces the same estimate $\hat{\mathbf{x}}[n]$ as a matrix WF since the covariance matrix R_y is the same and the cross-correlation vectors of the vector WFs are the columns of the cross-correlation matrix $R_{y,x}$.

3. REDUCED-RANK MATRIX WIENER FILTERS

The basic idea of reduced-rank equalization is to prefilter the N-dimensional observation vector $\mathbf{y}[n]$ by the matrix $T^{(D)} \in \mathbb{C}^{N\times D}$, $D < N$, and to get the estimate $\hat{\mathbf{x}}^{(D)}[n] \in \mathbb{C}^{M}$ by applying a *reduced-dimension matrix WF*[1] $W_{\mathrm{rd}}^{(D)} \in \mathbb{C}^{M\times D}$ to the transformed observation vector $\mathbf{y}^{(D)}[n] = T^{(D),\mathrm{H}}\mathbf{y}[n] \in \mathbb{C}^{D}$, i. e. $\hat{\mathbf{x}}^{(D)}[n] = W_{\mathrm{rd}}^{(D)} T^{(D),\mathrm{H}}\mathbf{y}[n]$. Thus, the reduced-dimension matrix WF is the solution of the optimization

$$W_{\mathrm{rd}}^{(D)} = \arg\min_{\mathcal{W}_{\mathrm{rd}}^{(D)}} \xi\left(\mathcal{W}_{\mathrm{rd}}^{(D)} T^{(D),\mathrm{H}}\right). \quad (5)$$

The *reduced-rank matrix WF*[2] is the combination of the prefilter matrix and the reduced-dimension matrix WF, i. e.

$$W^{(D)} = W_{\mathrm{rd}}^{(D)} T^{(D),\mathrm{H}} \in \mathbb{C}^{M\times N}, \quad (6)$$

$$W_{\mathrm{rd}}^{(D)} = R_{y,x}^{H} T^{(D)} \left(T^{(D),\mathrm{H}} R_y T^{(D)}\right)^{-1}, \quad (7)$$

and achieves the MMSE

$$\xi\left(W^{(D)}\right) = \mathrm{tr}\left\{R_x - R_{y^{(D)},x}^{H} R_{y^{(D)}}^{-1} R_{y^{(D)},x}\right\}. \quad (8)$$

Note that in general, $\hat{\mathbf{x}}^{(D)}[n]$ is unequal to the estimate $\hat{\mathbf{x}}[n]$ since the reduced-rank matrix WF $W^{(D)}$ is only an approximation of the matrix WF W in the subspace spanned by the columns of the prefilter matrix $T^{(D)}$.

In this paper, we restrict ourselves on the approximation of the matrix WF in eigenspaces, i. e. the prefilter matrix $T^{(D)}$ is composed by D eigenvectors of the covariance matrix R_y. In the sequel, we assume the eigenvalue decomposition $R_y = Q\Lambda Q^{\mathrm{H}}$ and consider two reduced-rank methods with a different choice of eigenvectors.

[1]The notation *reduced-dimension matrix WF* denotes a matrix WF whose rows are reduced-dimension vector WFs.

[2]Again, the *reduced-rank matrix WF* is a matrix whose rows are reduced-rank vector WFs.

3.1. Principal Component Method

The PC method [2] chooses the eigenspace such that the sum of signal powers in the eigenmodes of the covariance matrix R_y of the observation is maximized.

Consider the unitary modal matrix $Q = [q_1, \dots, q_N]$ corresponding to the diagonal matrix of eigenvalues $\Lambda = \mathbf{diag}\{\lambda_1, \dots, \lambda_N\}$ where $\lambda_1 \geq \dots \geq \lambda_N \geq 0$. Then, the PC prefilter matrix $T_{\mathrm{PC}}^{(D)}$ is composed by the eigenvectors corresponding to the largest eigenvalues, i. e. $T_{\mathrm{PC}}^{(D)} = [q_1, \dots, q_D]$. Thus, the MMSE may be written as

$$\xi\left(W_{\mathrm{PC}}^{(D)}\right) = \mathrm{tr}\{R_x\} - \sum_{i=1}^{D} \lambda_i^{-1} \|R_{y,x}^{H} q_i\|_2^2. \quad (9)$$

Since only the N-dimensional eigenvectors corresponding to the D principal eigenvalues must be computed and the inversion of the diagonal matrix $R_{y^{(D)},\mathrm{PC}} = T_{\mathrm{PC}}^{(D),\mathrm{H}} R_y T_{\mathrm{PC}}^{(D)}$ is computational negligible, the PC method has a computational complexity of $O(DN^2)$.

Again, the parallel application of reduced-rank vector WFs based on the PC method yields the same estimate as the application of the reduced-rank matrix WF $W_{\mathrm{PC}}^{(D)}$ because the prefilter matrix $T_{\mathrm{PC}}^{(D)}$ does not depend on cross-correlation information and therefore, the term $T_{\mathrm{PC}}^{(D)} R_{y^{(D)},\mathrm{PC}}^{-1} T_{\mathrm{PC}}^{(D),\mathrm{H}}$ is identical for the equalization of each element of $\mathbf{x}[n]$ (cf. Equation 6 and 7).

3.2. Cross-Spectral Method

Compared to the PC method, the CS method [3] chooses the D eigenvectors which achieve the minimum MSE of all eigenspace based reduced-rank matrix WFs with the same rank D, i. e.

$$\xi\left(W_{\mathrm{CS}}^{(D)}\right) = \mathrm{tr}\{R_x\} - \sum_{i\in\mathbb{M}} \lambda_i^{-1} \|R_{y,x}^{H} q_i\|_2^2, \quad (10)$$

where the set

$$\mathbb{M} = \arg\max_{\substack{\mathcal{M}\subset\{1,2,\dots,N\}\\ |\mathcal{M}|=D}} \sum_{i\in\mathcal{M}} \lambda_i^{-1} \|R_{y,x}^{H} q_i\|_2^2 \quad (11)$$

includes the indices of the eigenvectors with the largest CS metrics $\alpha_i = \lambda_i^{-1}\|R_{y,x}^{H} q_i\|_2^2$. $|\mathcal{M}|$ denotes the power of the set $\mathcal{M}$. The CS prefilter matrix $T_{\mathrm{CS}}^{(D)}$ is finally composed by the eigenvectors q_i, $i \in \mathbb{M}$.

Note that the complexity of the CS method is equal to the one of the WF, i. e. $O(N^3)$ due to the inversion of the $N \times N$ covariance matrix, because all eigenvectors have to be computed before determining the CS metrics. It remains the superiority of reduced-rank processing compared to the full approach in case of estimation errors of statistics due to low sample support [6].

Here, the prefilter matrix $T_{CS}^{(D)}$ depends on the cross-correlation information and hence, the term $T_{CS}^{(D)} R_{y^{(D)},CS}^{-1} T_{CS}^{(D),H}$ is different for the estimation of each element of $\mathbf{x}[n]$. Thus, the reduced-rank matrix WF based on the CS method is different from the parallel application of M reduced-rank vector WFs (cf. Figure 1).

4. APPLICATION TO A MULTI-INPUT MULTI-OUTPUT TRANSMISSION CHANNEL

4.1. Channel Model

The MIMO transmission channel with M inputs and N outputs is described by the channel matrix $H \in \mathbb{C}^{N\times M}$. The received signal vector

$$\mathbf{y}[n] = H\mathbf{x}[n] + \mathbf{n}[n] \in \mathbb{C}^N, \quad (12)$$

is perturbed by additive white Gaussian noise $\mathbf{n}[n]$ with the complex normal distribution $\mathcal{N}_c(\mathbf{0}_{N\times 1}, \sigma_n^2 \mathbf{1}_N)$ where $\mathbf{0}_{N\times 1}$ denotes the $N \times 1$ zero matrix and $\mathbf{1}_N$ the $N \times N$ identity matrix. The transmit signal vector $\mathbf{x}[n] \in \mathbb{C}^M$ at time index n is composed by zero-mean i.i.d. symbols with variance σ_x^2. Note that the total transmit power is $P_{Tx} = \mathrm{tr}\{R_x\} = \sigma_x^2 M$.

We consider a MIMO scenario [7] with flat fading where the channels H are realizations of the random variable

$$\mathbf{H} = U_{Rx}\Sigma_{Rx}\mathbf{Z}\Sigma_{Tx}V_{Tx}^H. \quad (13)$$

The matrix $\mathbf{Z} \in \mathbb{C}^{N\times M}$ has i.i.d. random entries with $\mathcal{N}_c(0, \sigma_z^2)$. The columns of the unitary matrix $U_{Rx} \in \mathbb{C}^{N\times N}$ are eigenmodes of the correlations at the receiver, i.e. $\mathrm{E}\{\mathbf{H}\mathbf{H}^H\} = \sigma_z^2 \,\mathrm{tr}\{\Sigma_{Tx}^2\} U_{Rx}\Sigma_{Rx}^2 U_{Rx}^H$ with the power distribution described by the diagonal matrix of eigenvalues $\Sigma_{Rx}^2 \in \mathbb{R}_0^{+,N\times N}$. Analogous, the unitary matrix $V_{Tx} \in \mathbb{C}^{M\times M}$ and the diagonal matrix $\Sigma_{Tx}^2 \in \mathbb{R}_0^{+,M\times M}$ from the eigenvalue decomposition of $\mathrm{E}\{\mathbf{H}^H\mathbf{H}\} = \sigma_z^2 \,\mathrm{tr}\{\Sigma_{Rx}^2\} V_{Tx}\Sigma_{Tx}^2 V_{Tx}^H$ depict the correlations at the transmitter. Note that the channel amplification concerning the variance σ_x^2 is given by $\mathrm{tr}\{\mathrm{E}\{\mathbf{H}\mathbf{H}^H\}\} = \mathrm{tr}\{\mathrm{E}\{\mathbf{H}^H\mathbf{H}\}\} = \sigma_z^2 \,\mathrm{tr}\{\Sigma_{Rx}^2\}\,\mathrm{tr}\{\Sigma_{Tx}^2\}$.

4.2. Relationship between Matrix PC and CS Method

In this subsection, we prove that the matrix PC and CS algorithm are equal since they choose exactly the same eigenspace for the approximation of the WF assuming the MIMO channel model defined in Subsection 4.1. First, consider the eigenvalue decomposition of the covariance matrix

$$R_y = \sigma_x^2 HH^H + \sigma_n^2 \mathbf{1}_N = Q\Lambda Q^H, \quad (14)$$

with Q and Λ from Subsection 3.1. Additionally, assume the *Singular Value Decomposition* (SVD) of

$$R_{y,x} = \sigma_x^2 H = U\Sigma V^H, \quad (15)$$

with the unitary matrices $U \in \mathbb{C}^{N\times N}$ and $V \in \mathbb{C}^{M\times M}$, and the rank $R = \mathrm{rank}\{H\} \leq M$ matrix

$$\Sigma = \begin{bmatrix} \mathrm{diag}\{\sigma_1, \ldots, \sigma_M\} \\ \mathbf{0}_{(N-M)\times M} \end{bmatrix} \in \mathbb{R}_0^{+,N\times M}, \quad (16)$$

composed by the singular values $\sigma_1 \geq \ldots \geq \sigma_R > \sigma_{R+1} = \ldots = \sigma_M = 0$.

It can easily be seen that $Q = U$ and $\Lambda = \sigma_x^{-2}\Sigma\Sigma^T + \sigma_n^2 \mathbf{1}_N$ where '$(\cdot)^T$' denotes transpose. Consequently, with Equation (15) and the characteristics of the left and right singular vectors, the CS metric for $i \in \{1, \ldots N\}$ gets

$$\alpha_i = \begin{cases} \lambda_i^{-1}\sigma_i^2 = \dfrac{\sigma_x^2}{1 + \sigma_x^2\sigma_n^2/\sigma_i^2}, & i \leq R, \\ 0, & R < i \leq N. \end{cases} \quad (17)$$

Due to the descending order of the singular values, it follows $\alpha_1 \geq \ldots \geq \alpha_R > \alpha_{R+1} = \ldots = \alpha_N = 0$ and the set of indices reads as $\mathbb{M} = \{1, \ldots, D\}$ if we recall Equation (11). Thus, the CS method chooses the same set of eigenvectors $\{q_1, \ldots, q_D\}$ as those selected by the PC method.

4.3. Simulation Results

We consider the transmission of QPSK symbols with variance $\sigma_x^2 = 1$ over a channel with $M = 4$ and $N = 8$. We choose a strongly correlated MIMO channel where $\Sigma_{Rx}^2 = \mathrm{diag}\{0.7, 0.1, 0.1, 0.05, 0.02, 0.01, 0.01, 0.01\}$ and $\Sigma_{Tx}^2 = \mathrm{diag}\{0.8, 0.1, 0.05, 0.05\}$. The unitary matrices U_{Rx} and V_{Tx} are chosen arbitrarily for every realization of the channel $\mathbf{H}$. Note that the channel amplification is comparable to the one of an uncorrelated MIMO channel where the zero-mean i.i.d. complex normal distributed random entries have variance one if we set $\sigma_z^2 = 32 = NM$ because $\mathrm{tr}\{\Sigma_{Rx}^2\} = \mathrm{tr}\{\Sigma_{Tx}^2\} = 1$. Additionally, we assume perfect estimation of second order statistics at the receiver.

Figure 1 shows the uncoded *Bit Error Rate* (BER) over the SNR $10\lg(P_{Tx}/\sigma_n^2)$ in dB for the proposed equalizers. It can be seen that all rank 4 approximations of the matrix WF perform equal to the optimum linear filter since $R = 4$. If second order statistics are estimated imperfectly, the reduced-rank methods may even be better than the matrix WF as shown by Goldstein et al. [6]. Contrary to the matrix MSWF whose rank is restricted to $D = 4$ in the given scenario since it reduces to a matched filter followed by a quadratic WF [5], the matrix PC and CS method may have ranks $D \in \{1, \ldots, 8\}$. Therefore, the rank can even be smaller than the dimension of the signal vector. The matrix PC and CS algorithm with rank $D = 3 < R$ saturate for high SNR values but produce identical BERs as derived in Section 4.2 although the matrix PC method is much less computational intense. Nevertheless, the vector CS method is much closer to the optimum solution while having only

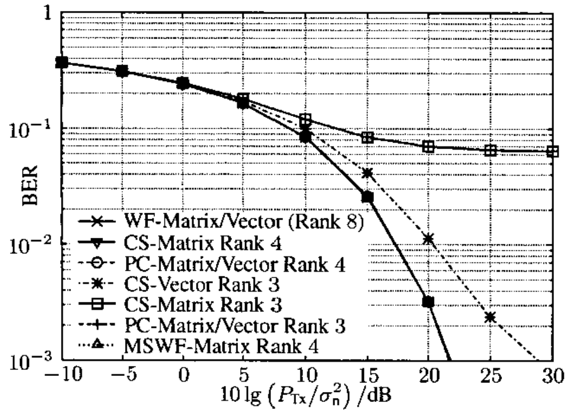

Fig. 1. Uncoded BER comparison of different reduced-rank equalizers to the WF in a flat MIMO scenario

a slightly more computational complexity since the SVD is performed only once whereas the CS metrics have to be computed for every element of the signal vector.

Now we consider the MIMO scenario described above with an additional weak interferer. In this case, the received signal vector may be written as

$$\mathbf{y}[n] = \boldsymbol{H}\mathbf{x}[n] + \boldsymbol{H}_{\mathrm{I}}\mathbf{x}_{\mathrm{I}}[n] + \mathbf{n}[n], \tag{18}$$

where the matrix $\boldsymbol{H}_{\mathrm{I}}$ is also a realization of $\mathbf{H}$. We assume the variance of the elements of the interference signal vector $\mathbf{x}_{\mathrm{I}}[n]$ to be $\sigma_{x_{\mathrm{I}}}^2 = 1/3$.

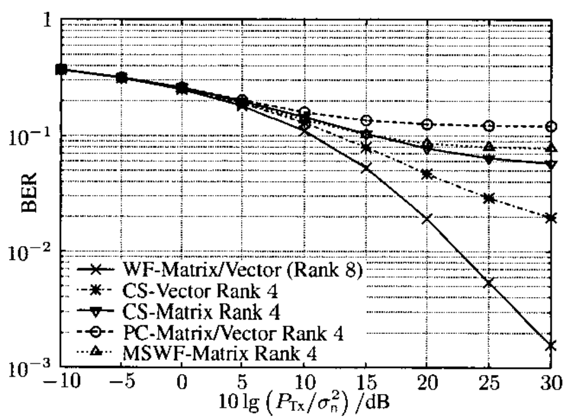

Fig. 2. Uncoded BER comparison of different reduced-rank equalizers to the WF in an interfered flat MIMO scenario

The simulation results are plotted in Figure 2 where only the signal vector $\mathbf{x}[n]$ of interest is detected (*single-user detection*). In this case the matrix PC and CS algorithm are no longer identical. This would change if we detect the user of interest and the interferer together (*multi-user detection*) because we end up in the case similar to a single-user MIMO scenario where the channel matrix has the dimension $N \times 2M$ and the signal vector $\left[\mathbf{x}^{\mathrm{T}}[n], \mathbf{x}_{\mathrm{I}}^{\mathrm{T}}[n]\right]^{\mathrm{T}} \in \mathbb{C}^{2M}$. Finally, the rank 4 approximations of the matrix WF are no longer equal to the optimum filter since the rank of the channel $R > 4$ due to the additional interferer. Note that the matrix MSWF performs even worse than the matrix CS method for SNR values larger than approximately 15 dB but it has also a smaller computational complexity.

5. CONCLUSIONS

In this paper, we extended two eigenspace based methods, the PC and CS method, for rank-reduction of a matrix WF which estimates an unknown signal vector instead of a scalar. Moreover, we showed that the PC and CS algorithms are exactly the same if we assume a flat MIMO channel with i.i.d. symbols and uncorrelated white Gaussian noise. Simulation results demonstrated the ability of the matrix CS algorithm to partially outperform the matrix MSWF in this special scenario. Additionally, the eigenspace based methods have more freedom in rank selection.

6. REFERENCES

[1] L. L. Scharf, *Statistical Signal Processing*, Addison-Wesley, 1991.

[2] H. Hotelling, "Analysis of a Complex of Statistical Variables into Principal Components," *Journal of Educational Psychology*, vol. 24, no. 6/7, pp. 417–441, 498–520, September/October 1933.

[3] J. S. Goldstein and I. S. Reed, "Reduced-Rank Adaptive Filtering," *IEEE Transactions on Signal Processing*, vol. 45, no. 2, pp. 492–496, February 1997.

[4] J. S. Goldstein, I. S. Reed, D. E. Dudgeon, and J. R. Guerci, "A Multistage Matrix Wiener Filter for Subspace Detection," in *Proc. IT Workshop on Detection, Estimation and Classification and Imaging*, February 1999.

[5] P. Cifuentes, W. L. Myrick, S. Sud, J. S. Goldstein, and M. D. Zoltowski, "Reduced Rank Matrix Multistage Wiener Filter with Applications in MMSE Joint Multiuser Detection for DS-CDMA," in *Proc. ICASSP 2002*, May 2002, vol. 3, pp. 2605–2608.

[6] J. S. Goldstein and I. S. Reed, "Theory of Partially Adaptive Radar," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 33, no. 4, pp. 1309–1325, October 1997.

[7] J. P. Kermoal, L. Schumacher, K. I. Pedersen, P. E. Mogensen, and F. Frederiksen, "A Stochastic MIMO Radio Channel Model With Experimental Validation," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 6, pp. 1211–1226, August 2002.