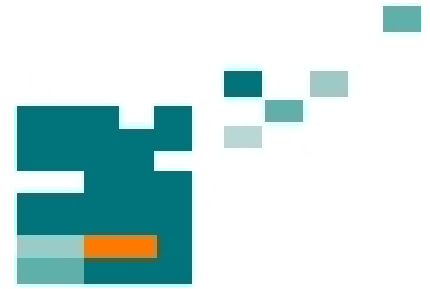**54. IWK**

Internationales Wissenschaftliches Kolloquium
International Scientific Colloquium

Information Technology and Electrical Engineering - Devices and Systems, Materials and Technologies for the Future

Faculty of Electrical Engineering and Information Technology

Startseite / Index:
http://www.db-thueringen.de/servlets/DocumentServlet?id=14089

TECHNISCHE UNIVERSITÄT ILMENAU

DETERMINISTIC PREWHITENING TO IMPROVE
SUBSPACE BASED PARAMETER ESTIMATION TECHNIQUES
IN SEVERELY COLORED NOISE ENVIRONMENTS

João Paulo C. L. da Costa[1], Florian Roemer, and Martin Haardt

Ilmenau University of Technology
Communications Research Laboratory
P.O. Box 100565, D-98684 Ilmenau, Germany
{joaopaulo.dacosta,florian.roemer,martin.haardt}@tu-ilmenau.de

ABSTRACT

Colored noise is encountered in a variety of signal processing applications. For such applications the prewhitening step becomes essential, since parameter estimation without prewhitening can be severely degraded.

Traditionally stochastic prewhitening techniques transform the colored noise into white noise keeping the SNR constant. In this paper, we propose a deterministic approach for subspace prewhitening, where we remove the correlation, which increases the SNR. Consequently, in high noise correlation scenarios, where the subspace is prewhitened by our deterministic approach, there is a significant improvement in the parameter estimation accuracy. The proposed deterministic prewhitening requires knowledge of the noise correlation. Therefore, we also propose solutions to estimate the correlation coefficients.

Index Terms— prewhitening, array signal processing, noise correlation estimation, colored noise.

1. INTRODUCTION

In practical applications using sensor arrays the assumption that the noise of the sensors is uncorrelated may be not valid. For example, underwater noise components of a sonar system are in general spatially correlated [1]. Therefore, if no prewhitening step is applied, a severe degradation of the performance is observed.

Typically the prewhitening approaches require the estimation of the noise covariance matrix R_{ww}, which is performed by collecting measurement samples in the absence of signal components. For example, in speech processing applications, the noise can be recorded in speechless frames [2]. The level of noise correlation (ρ) depends on the specific application. For example, in [3] and [4], ρ assumes values up to 0.99. For other applications, the correlation can assume smaller values.

In the stochastic prewhitening approaches of the literature [2, 5, 6], the data samples are multiplied by some prewhitening matrix, L^{-1}, which transforms the correlated noise into white noise. On the other hand, in our proposed deterministic approach, one sensor is used as the reference, and then, the correlated part of the noise is removed. In order to apply the deterministic approach, the correlation coefficients should be estimated in terms of their amplitudes and phases. In this paper, we also propose techniques to estimate these correlation parameters. We compare stochastic and deterministic prewhitening in computer simulations and demonstrate the improved performance of the deterministic approach. Here we restrict the application of the proposed deterministic prewhitening to ESPRIT-type algorithms. Nevertheless our technique can be applied together with other subspace based parameter estimation techniques, like MUSIC, Root MUSIC, or RARE.

The remainder of this paper is organized as follows. After reviewing the notation in Section 2, the colored noise model is presented in Section 3. Then, the estimation of spatial frequencies with a uniform linear array is presented as a data model example. In Section 4, we propose the new deterministic approach, which preserves the shift invariance structure and requires the estimation of the correlation coefficients. In Section 5, we propose methods to estimate the correlation coefficients for the noise model described in Section 3. Simulations results comparing the different prewhitening schemes are presented for the DOA estimation problem in Section 6. In Section 7, conclusions are drawn.

2. NOTATION

In order to facilitate the distinction between scalars and matrices, the following notation

---

is used: Scalars are denoted as italic letters $(a, b, \ldots, A, B, \ldots, \alpha, \beta, \ldots)$, column vectors as lower-case bold-face letters $(\boldsymbol{a}, \boldsymbol{b}, \ldots)$ and matrices as bold-face capitals $(\boldsymbol{A}, \boldsymbol{B}, \ldots)$. Lower-order parts are consistently named: the (i, j)-element of the matrix $\boldsymbol{A}$, is denoted as $a_{i,j}$.

We use the superscripts $^\mathrm{T}, ^\mathrm{H}, ^{-1}, ^{+}$, and * for transposition, Hermitian transposition, matrix inversion, the Moore-Penrose pseudo inverse of matrices, and complex conjugation, respectively.

3. DATA MODEL

As an example for the prewhitening schemes discussed in this paper, we consider a superposition of d planar wavefronts received by a uniform linear array with M sensors at N subsequent time instants. The measurement samples are given by

$$x_m(n) = \sum_{i=1}^{d} s_i(n) \cdot e^{j \cdot (m-1) \cdot \mu_i} + w_m^{(\mathrm{c})}(n), \qquad (1)$$

where $m = 1, 2, \ldots, M$, $n = 1, 2, \ldots, N$, $s_i(n)$, whose variance is σ_s^2, denotes the complex amplitude of the i-th exponential at time instant n, μ_i symbolizes the spatial frequency of the i-th exponential, and $w_m^{(\mathrm{c})}(n)$, whose variance is σ_w^2, models the additive spatially correlated noise component inherent in the measurement process. In the context of array signal processing, each of the exponentials represents one planar wavefront.

In matrix form, we can represent (1) in the following way

$$\boldsymbol{X} = \boldsymbol{A} \cdot \boldsymbol{S} + \boldsymbol{W}^{(\mathrm{c})}, \qquad (2)$$

where $\boldsymbol{A} \in \mathbb{C}^{M \times d}$ contains the steering vectors $\boldsymbol{a}_i \in \mathbb{C}^{M \times 1}$ for each of the d sources, $\boldsymbol{S} \in \mathbb{C}^{d \times N}$ contains the symbols $s_i(n)$, and $\boldsymbol{X}$ is corrupted by some spatially correlated noise matrix $\boldsymbol{W}^{(\mathrm{c})} \in \mathbb{C}^{M \times N}$. We can model the noise $\boldsymbol{W}^{(\mathrm{c})}$ as $\boldsymbol{W}^{(\mathrm{c})} = \boldsymbol{L} \cdot \boldsymbol{W}$, where $\boldsymbol{L}$ correlates the white noise matrix $\boldsymbol{W}$. The noise elements $w_m(n)$ of $\boldsymbol{W}$ are modeled as ZMCSCG (zero-mean circularly-symmetric complex Gaussian) random variables.

In practical applications, the model order must first be estimated. Model order estimation techniques that are suitable for this scenario are, for example, ESTimation ERror (ESTER) [7] and Subspace-based Automatic Model Order Selection (SAMOS) [8], since both are based on the shift invariance equation and, therefore, are compatible to the subspace deterministic prewhitening technique proposed here.

The covariance matrix of the data model (2) is given by

$$\begin{aligned} \boldsymbol{R}_{xx} &= \frac{1}{N} \cdot \mathrm{E}\{\boldsymbol{X} \cdot \boldsymbol{X}^{\mathrm{H}}\} \qquad (3) \\ &= \boldsymbol{A} \cdot \boldsymbol{R}_{\mathrm{ss}} \cdot \boldsymbol{A}^{\mathrm{H}} + \sigma_w^2 \cdot \boldsymbol{R}_{ww}, \end{aligned}$$

where $\boldsymbol{R}_{\mathrm{ss}}$ is the signal covariance matrix and $\boldsymbol{R}_{ww}$ is the noise covariance matrix, such that $\mathrm{tr}(\boldsymbol{R}_{ww}) = M$. In practice, $\boldsymbol{R}_{xx}$ can be estimated from a finite set of realizations via

$$\boldsymbol{R}_{xx} \approx \frac{1}{N} \cdot \boldsymbol{X} \cdot \boldsymbol{X}^{\mathrm{H}}. \qquad (4)$$

In the absence of signals, $\boldsymbol{R}_{ww}$ can also be estimated by using (4).

To demonstrate the estimation of the noise correlation coefficients, we consider the following specific correlation model [9]

$$w_{m+1}^{(\mathrm{c})}(n) = \rho \cdot w_m(n) + \sqrt{1 - |\rho|^2} \cdot w_{m+1}(n), \quad (5)$$

where $m = 1, 2, \ldots, M - 1$ and indicates the sensor position. Here, $\rho \in \mathbb{C}$ represents the noise correlation coefficient between the sensors m and $m + 1$, such that $0 \leq |\rho| \leq 1$.

Using this correlation model, the noise covariance matrix for $M = 3$ is given by

$$\boldsymbol{R}_{ww} = \begin{bmatrix} 1 & \rho^* & (\rho^*)^2 \\ \rho & 1 & \rho^* \\ \rho^2 & \rho & 1 \end{bmatrix}. \qquad (6)$$

4. SUBSPACE PREWHITENING APPROACHES

In the literature, we find three main stochastic approaches for the subspace prewhitening: the traditional one based on the Cholesky factorization of the noise covariance matrix [2], the GEVD approach [5, 6], and the GSVD approach [2, 5, 6].

In contrast to the stochastic approaches, a linear preprocessing with a deterministic matrix $\boldsymbol{D}$ is applied to remove the noise correlation between two adjacent sensors m and $m + 1$ in our proposed deterministic approach. First we consider the case where no signal component is present. Let us assume the noise covariance model in (6) to derive a preprocessing matrix $\boldsymbol{D}$. To this end, each element of the prewhitened noise matrix $\boldsymbol{V} = \boldsymbol{D} \cdot \boldsymbol{W}^{(\mathrm{c})}$ is transformed as

$$v_m(n) = w_{m+1}(n) \cdot \sqrt{1 - |\rho|^2}. \qquad (7)$$

Note that the greater $|\rho|$, the smaller is the variance of the deterministic prewhitened noise samples $v_m(n)$. Note that the correlated noise $w_{m+1}^{(\mathrm{c})}(n)$ in (5) is composed of a white noise component correlated with the noise at the previous sensor $w_m(n)$ and another uncorrelated white component $w_{m+1}(n)$. Since the white noise component of $w_m(n)$ is known from the previous sensor m, we can use this fact to obtain (7).

In order to decorrelate the noise as in (7), we propose a deterministic prewhitening matrix $\boldsymbol{D} = \boldsymbol{J}_2 - \hat{\rho} \cdot \boldsymbol{J}_1$, where $\boldsymbol{J}_2 \in \mathbb{R}^{M-1 \times M}$ and $\boldsymbol{J}_1 \in \mathbb{R}^{M-1 \times M}$ are the selection matrices for the last $M - 1$ sensors and for the first $M - 1$ sensors, respectively, and $\hat{\rho}$ is an estimate of ρ. In this section we consider that $\hat{\rho} = \rho$, and in Section 5, we propose ways of calculating $\hat{\rho}$.

Next, we consider the presence of signal components and the prewhitening matrix D is applied in the following fashion

$$Y = D \cdot X = \tilde{A} \cdot S + V, \qquad (8)$$

where the i-th column of $\tilde{A}$ is of the form[1]

$$\tilde{a}_i = [1, e^{j\mu_i}, \ldots, e^{(M-2) \cdot j\mu_i}]^T \cdot (e^{j\mu_i} - \rho). \quad (9)$$

The structure of $\tilde{a}_i$ can be derived by observing one sample of $y_m(n)$

$$y_m(n) = x_{m+1}(n) - \rho \cdot x_m(n) \qquad (10)$$
$$= \sum_{i=1}^{d} [e^{j \cdot m \cdot \mu_i} - \rho \cdot e^{j \cdot (m-1) \cdot \mu_i}] \cdot s_i(n) + v_m(n)$$
$$= \sum_{i=1}^{d} [e^{j \cdot (m-1) \cdot \mu_i}] \cdot (e^{j\mu_i} - \rho) \cdot s_i(n) + v_m(n),$$

where $1 \le m \le M - 1$.

One important property of $\tilde{A}$ is that it has a Vandermonde structure and therefore the spatial frequencies can be obtained from the transformed measurement matrix Y in the same way as from X. Consequently,

$$\tilde{J}_2 \cdot \tilde{A} = \tilde{J}_1 \cdot \tilde{A} \cdot \Phi, \qquad (11)$$

where Φ is a diagonal matrix with the spatial frequencies $e^{j\mu_i}$, and $\tilde{J}_1$ and $\tilde{J}_2$ are now of size $(M - 2) \times (M - 1)$.

The SVD of Y is given by $U_y \cdot \Sigma \cdot P^H$. We define the matrix $U_s \in \mathbb{C}^{(M-1) \times d}$ as the first d columns vectors of U_y. Then, there is a certain $T \in \mathbb{C}^{d \times d}$, such that $\tilde{A} = U_s \cdot T^{-1}$. Therefore, we can rewrite (11) as

$$\tilde{J}_2 \cdot U_s = \tilde{J}_1 \cdot U_s \cdot \Psi, \qquad (12)$$

where $\Psi = T^{-1} \cdot \Phi \cdot T$. Note that Ψ and Φ share the same set of eigenvalues.

The prewhitened noise power P_{pwt} is given by

$$P_{\text{pwt}} = \text{E}\{v_m(n) \cdot v_m^*(n)\} = \sigma_w^2 \cdot (1 - |\rho|^2), \quad (13)$$

where $P_w = \text{E}\{w_m(n) \cdot w_m^*(n)\} = \sigma_w^2$.

Since $\frac{P_{\text{pwt}}}{P_w} = 1 - |\rho|^2$, we have $P_{\text{pwt}} < P_w$ for $|\rho|^2 > 0$, which means that this approach always gives a better SNR, for $|\rho| > 0$ than in the white noise scenario. It is possible to observe this behavior by considering the SNR

$$\text{SNR}_{\text{pwt}} = 10 \cdot \log_{10} \left[\frac{\sigma_s^2 \cdot |e^{j\mu_i} - \rho|^2}{\sigma_w^2 \cdot (1 - |\rho|^2)} \right]. \qquad (14)$$

If the correlation $|\rho|$ is close to 1, it implies that SNR_{pwt} approaches infinity. Therefore, the higher the correlation, the better the parameter estimation for the deterministic approach, and this is shown in the simulations

---

[1]Note that for the case that ρ is close to $e^{j\mu_i}$ in magnitude and phase, then σ_s^2, the signal power for this particular μ_i is reduced.

in Section 6. Note that for such a gain, it is necessary that the correlation ρ should not be in phase to the signal.

The drawback of the deterministic approach is that the array aperture is reduced from M to $M - 1$ sensors. This leads to a minor performance degradation, which becomes visible for small M and low correlations.

In summary, the objective of the *stochastic prewhitening* approaches is given a certain colored noise matrix $W^{(\text{c})}$ and an estimate of the stochastic prewhitening matrix L^{-1}, to prewhiten the noise in (2), such that the elements of the prewhitened noise $W = L^{-1} \cdot W^{(\text{c})}$ have zero mean and variance σ_w^2, where σ_w^2 denotes the noise power. In our *deterministic approach*, we propose to use D, instead of the prewhitening matrix L^{-1}, such that $V = D \cdot W^{(\text{c})}$, and the elements of V have zero mean and variance $(1 - |\rho|^2) \cdot \sigma_w^2$, where $(1 - |\rho|^2) \cdot \sigma_w^2$ denotes the noise power after applying the proposed deterministic prewhitening.

5. ESTIMATION OF THE CORRELATION COEFFICIENT FOR THE DETERMINISTIC APPROACH

Since for the deterministic approach in Section 4, the estimation of ρ is necessary, we propose different ways of performing this estimation from measurements taken in the absence of signal components. First we represent ρ as a function of its phase and magnitude, such that $\rho = |\rho| \cdot e^{j \cdot \phi}$.

Let us first take the sample estimate to obtain the phase and magnitude of ρ

$$\hat{\rho} = \frac{\sum\limits_{n=1}^{N} w_{m+1}^{(\text{c})}(n) \cdot [w_m^{(\text{c})}(n)]^*}{\sum\limits_{n=1}^{N} w_m^{(\text{c})}(n) \cdot [w_m^{(\text{c})}(n)]^*}, \qquad (15)$$

where $m = 1, ..., M - 1$. Note that the enumerator in (15) is N times the element of $\hat{R}_{ww}$ in row $m + 1$ and column m. The sample estimate is applicable to arbitrary colored noise models. Next we show that for the considered specific correlated noise model, it is possible to improve the estimation of the correlation considerably. To this end, we propose two other techniques: the ESPRIT based phase estimation and the magnitude estimation of ρ.

For $M = 3$, the colored noise samples can be written as

$$W^{(\text{c})} = \begin{bmatrix} 1 & 0 & 0 \\ \rho & \sqrt{1 - |\rho|^2} & 0 \\ \rho^2 & \rho \cdot \sqrt{1 - |\rho|^2} & \sqrt{1 - |\rho|^2} \end{bmatrix} \cdot W. \quad (16)$$

Note that this linear transformation has a specific structure that can effectively be exploited to estimate ρ.

For instance, the first column of the transformation matrix, which has the strongest power, has a Vandermonde structure with rate equal to ρ. Therefore, an

ESPRIT based approach can be applied for the estimation of the phase of ρ, as shown in the following shift invariance equation

$$\boldsymbol{J}_2 \cdot \boldsymbol{u}_1^{(c)} = \boldsymbol{J}_1 \cdot \boldsymbol{u}_1^{(c)} \cdot \hat{\rho}, \tag{17}$$

where the SVD of $\boldsymbol{W}^{(c)}$ is given by $\boldsymbol{U}^{(c)} \cdot \boldsymbol{\Sigma}^{(c)} \cdot (\boldsymbol{P}^{(c)})^{\mathrm{H}}$ and $\boldsymbol{u}_1^{(c)}$ is the first column of the matrix $\boldsymbol{U}^{(c)}$. The estimated phase shift $e^{j \cdot \hat{\phi}}$ is given by $\frac{\hat{\rho}}{|\hat{\rho}|}$.

In Figure 1, we compare the sample estimate in (15) with the ESPRIT based approach in (17) for the estimation of $e^{j\phi}$. We note that the performance of the ESPRIT based approach in (17) is far better than the performance of the sample estimate in (15).

Since we have already estimated the phase shift between two consecutive sensors, this information can be applied in order to phase align the correlated noise samples. Therefore, we consider the case that $\rho \in \mathbb{R}$, and the outputs of two consecutive sensors are given by

$$
\begin{aligned}
x_m(n) &= w_m^{(c)}(n) \\
x_{m+1}(n) &= w_{m+1}(n) \cdot \sqrt{1 - \rho^2} + w_m^{(c)}(n) \cdot \rho.
\end{aligned}
\tag{18}
$$

The noise power can be estimated by

$$\hat{\sigma}_w^2 = \frac{1}{N} \cdot \sum_{n=1}^{N} x_m^2(n) = \frac{1}{N} \cdot \sum_{n=1}^{N} [w_m^{(c)}(n)]^2. \tag{19}$$

Given the noise model, the following expression can be derived

$$\mathrm{E}\{[x_{m+1}(n) - x_m(n)]^2\} = 2 \cdot \sigma_w^2 \cdot (1 - \rho). \tag{20}$$

Since σ_w^2 was estimated in (19) and using the expression in (20), the magnitude estimation of ρ is calculated according to

$$\hat{\rho} = 1 - \frac{1}{2 \cdot N \cdot \hat{\sigma}_w^2} \cdot \sum_{n=1}^{N} [x_{m+1}(n) - x_m(n)]^2 \tag{21}$$

In Figure 2, we compare the performance of the magnitude estimation in (21) with the sample estimate in (15), and the magnitude estimation outperforms significantly the estimation by the sample estimate. Note that both approaches have a better estimation accuracy when the number of samples is increased.

6. SIMULATION RESULTS

In this section we present simulations results, with which we compare the proposed deterministic method to the stochastic approaches. We consider the data symbols $s_i(n)$ as being ZMCSCG distributed. The performance comparison is based on the spatial frequency estimation with standard ESPRIT (SE) [5] and for each realization, the spatial frequency for each source is chosen randomly from a uniform distribution in the interval from $-\frac{\pi}{2}$ to $\frac{\pi}{2}$. In addition, we assume that ρ is
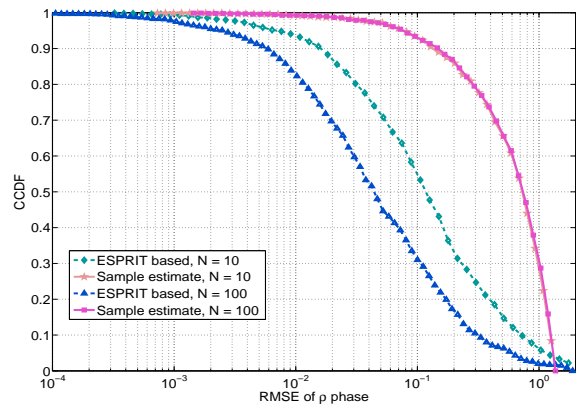
Fig. 1. Complementary cumulative distribution function (CCDF) of the root mean square error (RMSE) of $\frac{\hat{\rho}}{|\hat{\rho}|}$ considering a only noise system with $M = 11$ sensors, and $N = 10$ and $N = 100$ snapshots.

known, since we are interested in evaluating the different prewhitening schemes and not the estimation of ρ, which may vary, for example, for a different number of snapshots according to Section 5. The notation in the legends is the following: SE Color stands for the estimation without using any prewhitening, SE DET for the deterministic approach proposed here, SE CP for the classical prewhitening in [2], SE GEVD for the prewhitening scheme in [5, 6], and SE GSVD for the prewhitening in [2, 5, 6].
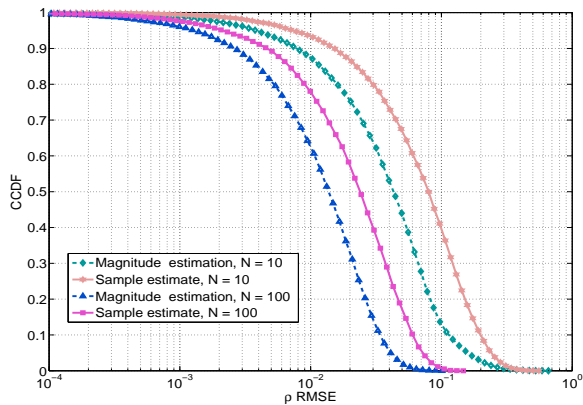
In Figure 3, the RMSE of the spatial frequencies is plotted versus the SNR. For this scenario, $\rho = 0.7$ and all the prewhitening techniques outperform SE Color. Note that the SE DET outperforms all the other prewhitening techniques.

In order to observe the performance as a function of ρ, we fix the SNR to 1 dB in Figure 4. Note that a considerable improvement by using all types of prewhitening is only observed for $\rho > 0.3$. Also note that the stochastic prewhitening schemes tend to keep the noise power σ_w^2 constant for all values of ρ. On the other hand, for the deterministic approach, the greater the correlation, the greater the gain obtained, which is expected according to (14). Therefore, in Figure 4, the deterministic prewhitening outperforms significantly the stochastic approaches for $\rho > 0.7$, and only slightly for $0.4 \leq \rho \leq 0.7$. As a drawback, we note that for $\rho \leq 0.4$, the stochastic prewhitening techniques slightly outperforms SE DET, and for $\rho < 0.3$, the estimation without prewhitening also slightly outperforms SE DET. This phenomenon is due to the aperture reduction already mentioned in Section 4.

7. CONCLUSIONS

In this paper, we propose a deterministic prewhitening technique, which outperforms the prewhitening techniques presented in the literature in case of high noise correlation. Observe that in general we have three cases:
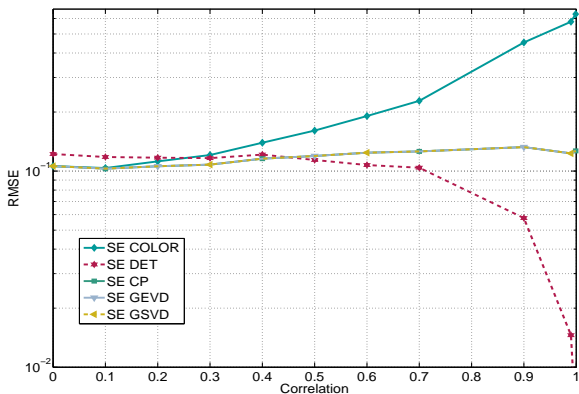
Fig. 2. Complementary cumulative distribution function (CCDF) of the root mean square error (RMSE) of $|\hat{\rho}|$ considering a only noise system with $M = 2$ sensors, and $N = 10$ and $N = 100$ snapshots.
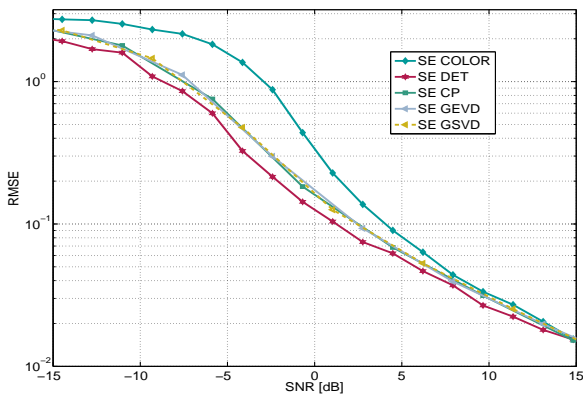
Fig. 3. RMSE of the spatial frequencies vs. SNR considering a system with $N = 10$ snapshots, with $M = 10$ sensors and with $\rho = 0.7$. $d = 3$ sources are present.

Fig. 4. RMSE of the spatial frequencies vs. correlation factor ρ considering a system with $N = 10$ snapshots and with $M = 10$ sensors. $d = 3$ sources are present. It was fixed SNR = 1 dB.

First there is the case of a small noise correlation, when the prewhitening step for the simulated scenario does not give a significant improvement. For an intermediate level of noise correlation, the stochastic prewhitening slightly outperforms the deterministic. Finally, for a high noise correlation, the proposed deterministic approach outperforms significantly the stochastic approaches.

Moreover we propose an ESPRIT based phase estimation together with the proposed magnitude estimation to obtain the correlation coefficients. Therefore, depending on the estimated level of correlation, it is possible to switch between no prewhitening, the deterministic and stochastic prewhitening.

8. REFERENCES

[1] Q. T. Zhang and K. M. Wong, "Information theoretic criteria for the determination of the number of signals in spatially correlated noise," *IEEE Transactions on Signal Processing*, vol. 41, pp. 1652–1662, Apr. 1993.

[2] P. C. Hansen and S. H. Jensen, "Prewhitening for rank-deficient noise in subspace methods for noise reduction," *IEEE Trans. Signal Processing*, vol. 53, pp. 3718–3726, Oct. 2005.

[3] L. Cao and D. J. Wu, "Noise-induced transport in a periodic system driven by gaussian white noises with intensive cross-correlation," *Physics Letters A*, vol. 291, pp. 371–375, Dec. 2001.

[4] T. Liu and S. Gazor, "Adaptive MLSD receiver employing noise correlation," *IEEE Proc.-Commun.*, vol. 53, pp. 719–724, Oct. 2006.

[5] R. Roy and T. Kailath, "ESPRIT - Estimation of signal parameters via rotational invariance techniques," in *Signal Processing Part II: Control Theory and Applications*, L. Auslander, F. A. Grünbaum, J. W. Helton, T. Kailath, P. Khargonekar, and S. Mitter, Eds. 1990, pp. 369–411, Springer-Verlag.

[6] M. Haardt, R. S. Thomä, and A. Richter, "Multidimensional high-resolution parameter estimation with applications to channel sounding," in *High-Resolution and Robust Signal Processing*, Y. Hua, A. Gershman, and Q. Chen, Eds. 2004, pp. 255–338, Marcel Dekker, New York, NY, Chapter 5.

[7] R. Badeau, B. David, and G. Richard, "Selecting the modeling order for the esprit high resolution method: an alternative approach," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2004)*, Montreal, Canada, May 2004.

[8] J.-M. Papy, L. De Lathauwer, and S. Van Huffel, "A shift invariance-based order-selection technique

for exponential data modelling," *IEEE Signal Processing Letters*, vol. 14, pp. 473–476, July 2007.

[9] Y. H. Chen and Y. S. Lin, "DOA estimation by fourth order cumulants in unknown noise environments," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 1993)*, 1993, vol. 4, pp. 296–299.