AN AMPLITUDE AND COVARIANCE MATRIX ESTIMATOR FOR SIGNALS IN COLORED GAUSSIAN NOISE

Jesper Højvang Jensen, Mads Græsbøll Christensen, and Søren Holdt Jensen

Deptartment of Electronic Systems
Aalborg University, Denmark
{jhj,mgc,shj}@kom.aau.dk.

ABSTRACT

We show that by considering the estimation of the amplitudes of sinusoids in colored, Gaussian noise as a joint amplitude and noise covariance matrix estimation problem, we obtain an iterative estimator that has the Capon spectral estimator as a special case. The estimator is also closely related to the amplitude and phase estimator (APES). In experiments, the proposed joint estimator in most cases outperforms Capon and APES.

1. INTRODUCTION

The estimation of the amplitudes of sinusoids in colored, Gaussian noise has a long history with applications such as radar imaging and audio coding (see e.g. [1] and the references therein). The simplest approach is to ignore the coloring of the noise and use methods such as least squares that assume white noise. A refinement is e.g. the Capon spectral estimator that uses the signal covariance matrix as an estimate of the noise covariance [2] (see also [3, 4]). An evolution of this is the amplitude and phase estimator (APES), which uses a cheap amplitude estimate to obtain a refined noise covariance estimate [4,5]. A related approach for audio signals can be found in [6]. The single-sinusoid APES algorithm was originally derived as an approximation to the exact joint maximum likelihood amplitude and noise covariance matrix estimator [5], similar to what we propose here for multiple sinusoids. However, when the multiple sinusoids APES algorithm was derived in [4], the noise covariance matrix was estimated prior to the amplitudes, not jointly.

In this paper, we do not consider the estimation of the noise covariance matrix and the sinusoid amplitudes as two separate tasks, but rather estimate them jointly. The resulting estimator is indeed closely related to both the Capon amplitude estimator and APES. While the joint estimator is computationally more demanding than the former two, it has the advantage that it avoids ad hoc noise covariance estimates.

In Section 2, we derive the joint noise covariance and sinusoid amplitude estimator, in Section 3 we evaluate it, and in Section 4 we conclude on the proposed approach.

This research was supported by the Intelligent Sound project, Danish Technical Research Council grant no. 26–04–0092, and the Parametric Audio Processing project, Danish Research Council for Technology and Production Sciences grant no. 274–06–0521.

2. MAXIMUM LIKELIHOOD PARAMETER ESTIMATION

We are concerned with the following complex-valued signal model:

$$x(n) = e(n) + \sum_{l=1}^{L} \alpha_l \exp(i\omega_l n) \text{ for } n \in \{0, 1, \ldots, N-1\}, \quad (1)$$

where $x(n)$ is the observed signal, ω_l are the known frequencies of the complex sinusoids, α_l are the unknown, complex amplitudes, and $e(n)$ is complex, colored, zero-mean Gaussian noise. The assumption of zero-mean noise is without loss of generality, since a non-zero mean is equivalent to an additional sinusoid with a frequency of zero. We define the vectors

$$\boldsymbol{x}(n) = \begin{bmatrix} x(n) & \cdots & x(n+M-1) \end{bmatrix}^{\mathrm{T}}, \quad (2)$$

$$\boldsymbol{e}(n) = \begin{bmatrix} e(n) & \cdots & e(n+M-1) \end{bmatrix}^{\mathrm{T}}, \quad (3)$$

$$\boldsymbol{s}_l(n) = \begin{bmatrix} \exp(i\omega_l n) & \cdots & \exp(i\omega_l (n+M-1)) \end{bmatrix}^{\mathrm{T}}, \quad (4)$$

$$\boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 & \cdots & \alpha_L \end{bmatrix}^{\mathrm{T}}, \quad (5)$$

and the matrix

$$\boldsymbol{S}(n) = \begin{bmatrix} \boldsymbol{s}_1(n) & \cdots & \boldsymbol{s}_L(n) \end{bmatrix}. \quad (6)$$

Letting $G = N - M + 1$ be the number of observed vectors, (1) is equivalent to

$$\boldsymbol{x}(n) = \boldsymbol{S}(n)\boldsymbol{\alpha} + \boldsymbol{e}(n) \text{ for } n \in \{0, 1, \ldots, G-1\}. \quad (7)$$

We will assume that the noise vectors $\boldsymbol{e}(n)$ are independent and Gaussian with $M \times M$ covariance matrix $\boldsymbol{Q} = \mathrm{E}[\boldsymbol{e}(n)\boldsymbol{e}(n)^{\mathrm{H}}]$. In this case, the maximum likelihood estimates of $\boldsymbol{\alpha}$ and $\boldsymbol{Q}$, denoted by $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{Q}}$, are given by

$$[\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{Q}}] = \underset{\boldsymbol{\alpha}, \boldsymbol{Q}}{\arg\max}\, J(\boldsymbol{\alpha}, \boldsymbol{Q}) \quad (8)$$

where

$$J(\boldsymbol{\alpha}, \boldsymbol{Q}) = \frac{1}{G} \sum_{n=0}^{G-1} \log p(x(n)) \quad (9)$$

$$= \frac{1}{G} \sum_{n=0}^{G-1} \log \Big[\frac{1}{\pi^M |\boldsymbol{Q}|} \exp(-\boldsymbol{e}(n)^{\mathrm{H}} \boldsymbol{Q}^{-1} \boldsymbol{e}(n)) \Big], \quad (10)$$

with $\boldsymbol{e}(n) = \boldsymbol{x}(n) - \boldsymbol{S}(n)\boldsymbol{\alpha}$.

We find $\hat{\alpha},\hat{\boldsymbol{Q}}$ in a two-step process. We first find the maximum likelihood estimate of $\hat{\boldsymbol{Q}}$ given the amplitudes, i.e.,

$$\hat{\boldsymbol{Q}}(\boldsymbol{\alpha}) = \arg\max_{\boldsymbol{Q}} J(\boldsymbol{\alpha},\boldsymbol{Q}). \tag{11}$$

Next, inserting $\hat{\boldsymbol{Q}}(\boldsymbol{\alpha})$ in J, we find the $\boldsymbol{\alpha}$ that maximizes J with $\hat{\boldsymbol{Q}}(\boldsymbol{\alpha})$ inserted for $\boldsymbol{Q}$:

$$\hat{\boldsymbol{\alpha}} = \arg\max_{\boldsymbol{\alpha}} J(\boldsymbol{\alpha},\hat{\boldsymbol{Q}}(\boldsymbol{\alpha})). \tag{12}$$

If all values exist and are unique, $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{Q}}(\hat{\boldsymbol{\alpha}})$ found in this way are identical to $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{Q}}$ found by directly maximizing (8) (this can be proven by assuming the opposite and show that this leads to contradictions). Using this procedure, we show that maximizing the log-likelihood of the observed data is equivalent to minimizing the determinant of the estimated noise covariance matrix, before we derive an iterative estimator that finds this minimum.

It is well-known that the maximum likelihood estimate of the noise covariance matrix given the amplitudes is given by [1]

$$\hat{\boldsymbol{Q}}(\boldsymbol{\alpha}) = \frac{1}{G}\sum_{n=0}^{G-1}\boldsymbol{e}(n)\boldsymbol{e}(n)^{\mathrm{H}}, \tag{13}$$

still with $\boldsymbol{e}(n) = \boldsymbol{x}(n) - \boldsymbol{S}(n)\boldsymbol{\alpha}$.

To see that maximizing the log-likelihood is equivalent to minimizing the determinant of the estimated noise covariance matrix (as shown by e.g. [7] and [5]), we first insert (13) in (10):

$$J(\boldsymbol{\alpha},\hat{\boldsymbol{Q}}(\boldsymbol{\alpha})) = \frac{1}{G}\sum_{n=0}^{G-1}\log\left[\frac{1}{\pi^M|\hat{\boldsymbol{Q}}(\boldsymbol{\alpha})|}\mathrm{e}^{-\boldsymbol{e}(n)^{\mathrm{H}}\hat{\boldsymbol{Q}}(\boldsymbol{\alpha})^{-1}\boldsymbol{e}(n)}\right] \tag{14}$$

$$= -M\log\pi - \log|\hat{\boldsymbol{Q}}(\boldsymbol{\alpha})| - \frac{1}{G}\sum_{n=0}^{G-1}\boldsymbol{e}(n)^{\mathrm{H}}\hat{\boldsymbol{Q}}(\boldsymbol{\alpha})^{-1}\boldsymbol{e}(n). \tag{15}$$

Traditionally in maximum likelihood amplitude estimation, the first two terms of (15) are considered constant, while the last term is maximized. However, when $\hat{\boldsymbol{Q}}$ is given by (13) instead of being independent of the estimated amplitudes, we can show that the last term vanishes, leaving only the second term to be optimized. Inserting (13) in the last term, we obtain

$$\begin{aligned}\frac{1}{G}\sum_{n=0}^{G-1}\boldsymbol{e}(n)^{\mathrm{H}}\hat{\boldsymbol{Q}}(\boldsymbol{\alpha})^{-1}\boldsymbol{e}(n) &= \mathrm{tr}\Big[\hat{\boldsymbol{Q}}(\boldsymbol{\alpha})^{-1}\frac{1}{G}\sum_{n=0}^{G-1}\boldsymbol{e}(n)\boldsymbol{e}(n)^{\mathrm{H}}\Big] \\ &= \mathrm{tr}[\hat{\boldsymbol{Q}}(\boldsymbol{\alpha})^{-1}\hat{\boldsymbol{Q}}(\boldsymbol{\alpha})] = M.\end{aligned} \tag{16}$$

In the above, $\mathrm{tr}[\cdot]$ denotes the trace operator. Thus, $J(\boldsymbol{\alpha},\hat{\boldsymbol{Q}}(\boldsymbol{\alpha}))$ is given by

$$J(\boldsymbol{\alpha},\hat{\boldsymbol{Q}}(\boldsymbol{\alpha})) = -M\log\pi - \log|\hat{\boldsymbol{Q}}(\boldsymbol{\alpha})| - M. \tag{17}$$

The maximum likelihood estimates $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{Q}}$ are therefore simply given by

$$\hat{\boldsymbol{\alpha}} = \arg\min_{\boldsymbol{\alpha}} J(\boldsymbol{\alpha},\hat{\boldsymbol{Q}}(\boldsymbol{\alpha})) \tag{18}$$

$$= \arg\min_{\boldsymbol{\alpha}} \log|\hat{\boldsymbol{Q}}(\boldsymbol{\alpha})| \tag{19}$$

and

$$\hat{\boldsymbol{Q}} = \hat{\boldsymbol{Q}}(\hat{\boldsymbol{\alpha}}). \tag{20}$$

For $\boldsymbol{S}(n)$ being a single sinusoid, a closed form expression for $\hat{\alpha}$ was derived in [5]. To compute $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{Q}}$ in the general case, we find the derivative of $\log|\hat{\boldsymbol{Q}}(\boldsymbol{\alpha})|$ with respect to $\boldsymbol{\alpha}$ and find where it equals zero. First, we use the chain rule to see that

$$\frac{d\log|\hat{\boldsymbol{Q}}(\boldsymbol{\alpha})|}{d\alpha_l} = \mathrm{tr}\left(\frac{\partial\log|\hat{\boldsymbol{Q}}(\boldsymbol{\alpha})|}{\partial\hat{\boldsymbol{Q}}(\boldsymbol{\alpha})^{\mathrm{T}}}\frac{\partial\hat{\boldsymbol{Q}}(\boldsymbol{\alpha})}{\partial\alpha_l}\right). \tag{21}$$

Using the relation $\frac{\partial\log|\hat{\boldsymbol{Q}}(\boldsymbol{\alpha})|}{\partial\hat{\boldsymbol{Q}}(\boldsymbol{\alpha})^{\mathrm{T}}} = \hat{\boldsymbol{Q}}(\boldsymbol{\alpha})^{-1}$ (see [8, 9]), we first compute the partial derivative with respect to a single amplitude:

$$\frac{d\log|\hat{\boldsymbol{Q}}(\boldsymbol{\alpha})|}{d\alpha_l} = \mathrm{tr}\left(\hat{\boldsymbol{Q}}(\boldsymbol{\alpha})^{-1}\frac{\partial\hat{\boldsymbol{Q}}(\boldsymbol{\alpha})}{\partial\alpha_l}\right) \tag{22}$$

$$= \mathrm{tr}\left(\hat{\boldsymbol{Q}}(\boldsymbol{\alpha})^{-1}\frac{\partial}{\partial\alpha_l}\frac{1}{G}\sum_{n=0}^{G-1}(\boldsymbol{x}(n)-\boldsymbol{S}(n)\boldsymbol{\alpha})(\boldsymbol{x}(n)-\boldsymbol{S}(n)\boldsymbol{\alpha})^{\mathrm{H}}\right) \tag{23}$$

$$= -\frac{1}{G}\mathrm{tr}\left(\hat{\boldsymbol{Q}}(\boldsymbol{\alpha})^{-1}\sum_{n=0}^{G-1}\boldsymbol{s}_l(n)(\boldsymbol{x}(n)-\boldsymbol{S}(n)\boldsymbol{\alpha})^{\mathrm{H}}\right) \tag{24}$$

$$= -\frac{1}{G}\mathrm{tr}\left(\sum_{n=0}^{G-1}(\boldsymbol{x}(n)-\boldsymbol{S}(n)\boldsymbol{\alpha})^{\mathrm{H}}\hat{\boldsymbol{Q}}(\boldsymbol{\alpha})^{-1}\boldsymbol{s}_l(n)\right) \tag{25}$$

$$= -\frac{1}{G}\sum_{n=0}^{G-1}(\boldsymbol{x}(n)-\boldsymbol{S}(n)\boldsymbol{\alpha})^{\mathrm{H}}\hat{\boldsymbol{Q}}(\boldsymbol{\alpha})^{-1}\boldsymbol{s}_l(n). \tag{26}$$

From this we see that the derivative with respect to $\boldsymbol{\alpha}$ is given by

$$\frac{d\log|\hat{\boldsymbol{Q}}(\boldsymbol{\alpha})|}{d\boldsymbol{\alpha}^{\mathrm{T}}} = \frac{1}{G}\sum_{n=0}^{G-1}(\boldsymbol{x}(n)-\boldsymbol{S}(n)\boldsymbol{\alpha})^{\mathrm{H}}\hat{\boldsymbol{Q}}(\boldsymbol{\alpha})^{-1}\boldsymbol{S}(n). \tag{27}$$

Setting this equal to zero, we obtain

$$\hat{\boldsymbol{\alpha}} = \left[\sum_{n=0}^{G-1}\boldsymbol{S}(n)^{\mathrm{H}}\hat{\boldsymbol{Q}}(\hat{\boldsymbol{\alpha}})^{-1}\boldsymbol{S}(n)\right]^{-1}\left[\sum_{n=0}^{G-1}\boldsymbol{S}(n)^{\mathrm{H}}\hat{\boldsymbol{Q}}(\hat{\boldsymbol{\alpha}})^{-1}\boldsymbol{x}(n)\right]. \tag{28}$$

However, this is not a closed form solution, since the $\hat{\boldsymbol{Q}}(\hat{\boldsymbol{\alpha}})$ term also depends on $\hat{\boldsymbol{\alpha}}$. Nevertheless, we can choose an initial guess for $\hat{\boldsymbol{\alpha}}$, e.g. $\hat{\boldsymbol{\alpha}}_0 = \boldsymbol{0}$, and cyclically compute refined covariance matrix and amplitude estimates:

$$\hat{\boldsymbol{Q}}(\hat{\boldsymbol{\alpha}}_k) = \frac{1}{G}\sum_{n=0}^{G-1}(\boldsymbol{x}(n)-\boldsymbol{S}(n)\hat{\boldsymbol{\alpha}}_k)(\boldsymbol{x}(n)-\boldsymbol{S}(n)\hat{\boldsymbol{\alpha}}_k)^{\mathrm{H}} \tag{29}$$

and

$$\hat{\boldsymbol{\alpha}}_{k+1} = \left[\sum_{n=0}^{G-1}\boldsymbol{S}(n)^{\mathrm{H}}\hat{\boldsymbol{Q}}(\hat{\boldsymbol{\alpha}}_k)^{-1}\boldsymbol{S}(n)\right]^{-1}\left[\sum_{n=0}^{G-1}\boldsymbol{S}(n)^{\mathrm{H}}\hat{\boldsymbol{Q}}(\hat{\boldsymbol{\alpha}}_k)^{-1}\boldsymbol{x}(n)\right]. \tag{30}$$

Note that the maximum likelihood amplitude estimator for sinusoids in additive Gaussian noise with known covariance matrix is identical to (30) with $\hat{\boldsymbol{Q}}(\boldsymbol{\alpha}_k)$ replaced by the true covariance matrix. This is surprising, since the latter maximizes the last term of (15), while the joint estimator maximizes the second term. Since (29) is the maximum likelihood estimate of the covariance matrix given the amplitudes, we cyclically compute the maximum likelihood estimate of one parameter using the current estimate of the other parameter. This guarantees that the log-likelihood of each refined estimate never decreases, whereby convergence is guaranteed. It is an open question, though, whether convergence is always to the global optimum or if it is only to local optima. In practice, we have observed that when using the zero vector as the initial estimate of $\boldsymbol{\alpha}$, the iterations converge in very few iterations.

Using the zero vector as the initial estimate, the first estimate of $\boldsymbol{Q}$ computed from (29) becomes

$$\hat{\boldsymbol{Q}}(\boldsymbol{0}) = \frac{1}{G}\sum_{n=0}^{G-1} \boldsymbol{x}(n)\boldsymbol{x}(n)^{\mathrm{H}}, \tag{31}$$

and the first non-trivial amplitude estimate becomes

$$\hat{\boldsymbol{\alpha}}_1 = \left[\sum_{n=0}^{G-1} \boldsymbol{S}(n)^{\mathrm{H}}\hat{\boldsymbol{Q}}(\boldsymbol{0})^{-1}\boldsymbol{S}(n)\right]^{-1}\left[\sum_{n=0}^{G-1} \boldsymbol{S}(n)^{\mathrm{H}}\hat{\boldsymbol{Q}}(\boldsymbol{0})^{-1}\boldsymbol{x}(n)\right], \tag{32}$$

which is nothing but the Capon amplitude estimator. The Capon estimator is thus the special case of the proposed joint estimator where we use the zero vector as the initial amplitude estimate and stop after a single iteration.

The APES algorithm uses a cheap estimate of the sinusoids' amplitudes to obtain a refined noise covariance estimate. For the multiple sinusoid version in [4], the refined noise covariance matrix is given by

$$\hat{\boldsymbol{Q}} = \frac{1}{G}\sum_{n=0}^{G-1} \boldsymbol{x}(n)\boldsymbol{x}(n)^{\mathrm{H}} - \sum_{l=1}^{L} \boldsymbol{g}_l\boldsymbol{g}_l^{\mathrm{H}}, \tag{33}$$

where

$$\boldsymbol{g}_l = \frac{1}{G}\sum_{n=0}^{G-1} \boldsymbol{x}(n)\exp(-i\omega_l n). \tag{34}$$

This has some resemblance to the second iteration of the proposed estimator, where the first amplitude estimate is used to obtain a refined noise covariance estimate. However, since no further iterations are performed with the APES algorithm, it does not converge to the maximum likelihood estimate.

Note that in showing that maximizing the log-likelihood is equivalent to minimizing the estimated noise covariance matrix, we only use that the same error signal estimate is used for calculating the noise covariance estimate and for computing the log-likelihood. Hence, this property holds for all signals in additive Gaussian noise, whether it is a linear or a nonlinear combination of deterministic signals, and it also holds if we use the forward-backward estimate for both the covariance matrix and the log-likelihood. In deriving the iterative estimator in (29) and (30), we restrict ourselves to linear combinations of signals, although they do not need to be sinusoids. Furthermore, since the function we end up optimizing in (19) is the determinant of the noise covariance matrix, we can analyze its properties using asymptotic eigenvalue properties. In particular, if the covariance matrix

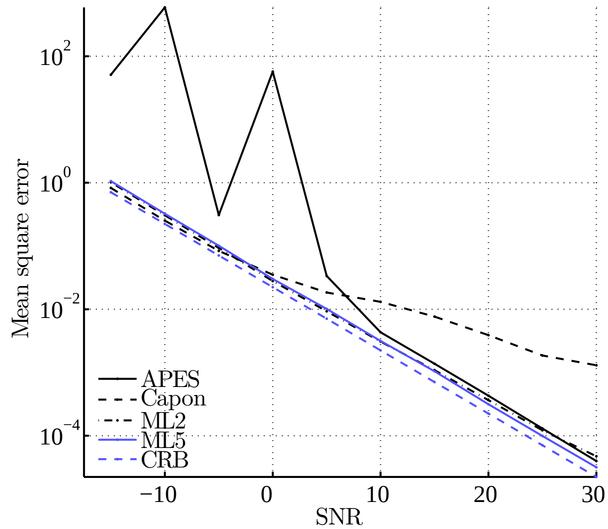

Figure 1: The mean square estimation error of α_1 in (35), located at $f_1 = 0.1$ with f_2 interfering at 0.11. ML2 and ML5 are the proposed estimator with two and five iterations, respectively, and CRB is the Cramér-Rao lower bound.

is Toeplitz and obeys certain regularity properties, a special case of Szegő's theorem asymptotically relates the log determinant to the logarithm of the Fourier spectrum of the autocorrelation function [10].

3. EVALUATION

We have tested the proposed estimator in the same setup as used in [4], where a sum of three complex sinusoids are buried in autoregressive Gaussian noise. The observed signal is given by

$$x(n) = e(n) + \sum_{l=1}^{3} \alpha_l \exp(i2\pi f_l n) \tag{35}$$

with $\alpha_1 = \exp(i\pi/4)$, $\alpha_2 = \exp(i\pi/3)$, $\alpha_3 = \exp(i\pi/4)$, and $f_1 = 0.1$, $f_2 = 0.11$, and $f_3 = 0.3$. The colored noise $e(n)$ is given by

$$e(n) = 0.99e(n-1) + v(n), \tag{36}$$

where $v(n)$ is white, Gaussian noise. The observation length N is 32, and the dimension of the covariance matrix, M, is 8. The mean square error for the first and third sinusoid are shown in Figure 1 and 2, respectively, for different signal to noise ratios. The mean square error of the first sinusoid, located at frequency 0.1 with a neighboring sinusoid at 0.11, shows how well the different estimators handle spectral leakage from sinusoids at neighboring frequencies. The mean square errors of the third sinusoid, located at 0.3, are indicative of performance when no other sinusoids are close. Consequently, the mean square errors in Figure 2 are much lower than in Figure 1. The mean square errors are averaged over 1000 realizations. We have compared APES and Capon to the proposed estimator with two and five iterations of (29) and (30).

In Figure 1, we see that in the case of an interfering neighboring sinusoid, the proposed estimator has uniformly good performance at all SNR and is consistently close to the

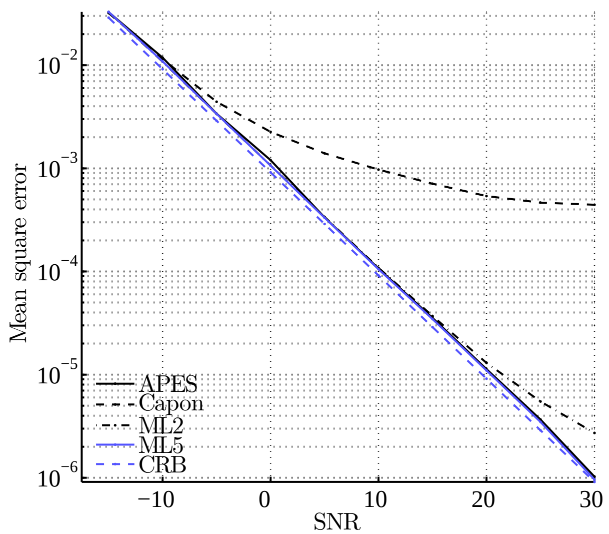

Figure 2: The mean square estimation error of α_3 in (35), located at $f_3 = 0.3$ with the closest interfering sinusoid located at 0.11.

Cramér-Rao lower bound. At low SNR, additional iterations decrease performance slightly, and the Capon amplitude estimator (which corresponds to a single iteration) has best performance. At higher SNR, the proposed estimator with five iterations performs best. In Figure 2, the estimation errors are much lower due to the less dominant interference. APES and the proposed estimator with five iterations seem to have nearly identical performance close to the Cramér-Rao bound, while two iterations and the Capon estimator sees a performance decrease at high SNR.

4. CONCLUSION

We have shown that maximizing the log-likelihood of a signal in colored, Gaussian noise is equivalent to minimizing the determinant of the estimated noise covariance matrix, and we have derived an iterative algorithm to find the optimum for a sum of sinusoids with unknown amplitudes. The derived algorithm has the Capon amplitude estimator as a special case, and experimentally, the new estimator shows consistently good performance. The proposed estimator has interesting theoretical implications, since it demonstrates that sinusoidal amplitude estimation in colored noise can elegantly be cast as a joint amplitude and noise covariance matrix estimation problem, instead of using ad hoc noise covariance estimates, and because it allows the use of asymptotic determinant properties such as Szegő's theorem for the analysis of maximum likelihood estimators. The latter may also be useful for deriving computationally cheap, asymptotically efficient estimators.

REFERENCES

[1] P. Stoica and R. Moses, *Spectral Analysis of Signals*. Prentice Hall, 2005.

[2] J. Capon, "High-resolution frequency-wavenumber spectrum analysis," *Proc. IEEE*, vol. 57, no. 8, pp. 1408–1418, 1969.

[3] M. A. Lagunas, M. E. Santamaria, A. Gasull, and A. Moreno, "Maximum likelihood filters in spectral estimation problems," *Elsevier Signal Processing*, vol. 10(1), pp. 19–34, 1986.

[4] P. Stoica, H. Li, and J. Li, "Amplitude estimation of sinusoidal signals: survey, new results, and an application," *IEEE Trans. Signal Processing*, vol. 48, no. 2, pp. 338–352, 2000.

[5] J. Li and P. Stoica, "An adaptive filtering approach to spectral estimation and sar imaging," *IEEE Trans. Signal Processing*, vol. 44, no. 6, pp. 1469–1484, 1996.

[6] M. G. Christensen and S. H. Jensen, "Variable order harmonic sinusoidal parameter estimation for speech and audio signals," in *Proc. Fortieth Asilomar Conference on Signals, Systems and Computers ACSSC '06*, 2006, pp. 1126–1130.

[7] G. Bienvenu and L. Kopp, "Optimality of high resolution array processing using the eigensystem approach," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 31, no. 5, pp. 1235–1248, 1983.

[8] A. Hjorungnes and D. Gesbert, "Complex-valued matrix differentiation: Techniques and key results," *IEEE Trans. Signal Processing*, vol. 55, no. 6, pp. 2740–2746, 2007.

[9] T. P. Minka, "Old and new matrix algebra useful for statistics," Tech. Rep., 2000.

[10] R. M. Gray, "Toeplitz and circulant matrices: A review," *Foundations and Trends in Communications and Information Theory*, vol. 2, pp. 155–239, 2006.