SPARSE BAYESIAN LEARNING FOR DOA ESTIMATION OF CORRELATED SOURCES

Christoph F. Mecklenbräuker[†], *Peter Gerstoft*[⋆], *and Geert Leus*[+]

[†]Inst. of Telecommunications, TU Wien, 1040 Wien, Austria
[⋆]University of California San Diego
[+]Dept. of Electrical Eng., Math. and Comp. Science, Delft Univ. of Technology

ABSTRACT

Direction of arrival (DOA) estimation from array observations in a noisy environment is discussed. The source amplitudes are assumed to be correlated zero-mean complex Gaussian distributed with unknown covariance matrix. The DOAs and covariance parameters of plane waves are estimated from multi-snapshot sensor array data using sparse Bayesian learning (SBL). The performance of SBL is evaluated in terms of the fidelity of the reconstructed coherency matrix of the estimated plane waves.

1. INTRODUCTION

When the sources are weak and closely spaced, parametric methods are needed for high-resolution DOA estimation. This is demonstrated for uncorrelated sources and the application of multiple measurement vector (MMV, or multi-snapshot) compressive beamforming [1, 2, 3, 4]. The MMV problem was previously solved using the sparse Bayesian learning (SBL) framework [3, 5, 6] with the maximum-a-posteriori (MAP) estimate for DOA reconstruction.

Here, we allow the sources to be correlated. Thus, we assume the source signals to jointly follow a zero-mean multivariate complex normal distribution with unknown covariance parameters. The noise across sensors and snapshots also follows a zero-mean multivariate normal distribution with unknown variance. These assumptions lead to a Gaussian likelihood function.

The corresponding posterior distribution is also Gaussian and already developed SBL approaches solve this well when the sources are uncorrelated. We base our present development on our fast SBL method [5, 6] which we augment to estimate the signal covariance parameters and noise variance. Standard techniques are based on minimization-majorization [7] and expectation maximization (EM) [3, 8, 9, 10, 11, 12, 13], though not all estimates work well. Instead, we estimate the unknown (co-)variances using stochastic maximum likelihood [14, 15, 16].

Supported by the Office of Naval Research, Grant Nos. N00014-1110439.

1.1. Noisy sensor array observation model

For the lth observation snapshot, we assume the linear model

$$\mathbf{y}_l = \mathbf{A}\mathbf{x}_l + \mathbf{n}_l, \tag{1}$$

where the dictionary $\mathbf{A} \in \mathbb{C}^{N\times M}$ is constant and known, and the source vector $\mathbf{x}_l \in \mathbb{C}^M$ contains the physical information of interest. Further, $\mathbf{n}_l \in \mathbb{C}^N$ is additive zero-mean circularly symmetric complex Gaussian noise, $\mathbf{n}_l \sim \mathcal{CN}(\mathbf{n}_l; \mathbf{0}, \boldsymbol{\Sigma}_\mathbf{n})$. Due to the circular symmetry of the noise the phase is uniformly distributed.

We specialize to diagonal noise covariance matrices, parameterized as

$$\boldsymbol{\Sigma}_\mathbf{n} = \sigma^2\mathbf{I} \tag{2}$$

Let $\mathbf{X}{=}[\mathbf{x}_1,\ldots,\mathbf{x}_L]{\in}\mathbb{C}^{M\times L}$ be the complex source amplitudes, $x_{ml} = [\mathbf{X}]_{m,l} = [\mathbf{x}_l]_m$ with $m \in \{1,\cdots,M\}$ and $l \in \{1,\cdots,L\}$, at M DOAs (e.g., $\theta_m = -90^\circ + \frac{m-1}{M}180^\circ$) and L snapshots for a frequency ω. We observe narrowband waves on N sensors for L snapshots $\mathbf{Y} = [\mathbf{y}_1,\ldots,\mathbf{y}_L] \in \mathbb{C}^{N\times L}$. A linear regression model relates the array data $\mathbf{Y}$ to the source amplitudes $\mathbf{X}$ as

$$\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{N}. \tag{3}$$

The dictionary $\mathbf{A}{=}[\mathbf{a}_1,...,\mathbf{a}_M]{\in}\mathbb{C}^{N\times M}$ contains the array steering vectors for all hypothetical DOAs as columns, with the (n,m)th element given by $\mathrm{e}^{-\mathrm{j}\frac{\omega d_n}{c}\sin\theta_m}$ (d_n is the distance to the reference element and c the sound speed).

We assume $M{>}N$ and thus (3) is underdetermined. In the presence of only few stationary sources, the source vector $\mathbf{x}_l$ is K-sparse with $K{\ll}M$. We define the lth active set

$$\mathcal{M}_l = \{m \in \mathbb{N} | x_{ml} \neq 0\}, \tag{4}$$

and assume $\mathcal{M}_l{=}\mathcal{M}{=}\{m_1,...,m_K\}$ is constant across all snapshots l. Also, we define $\mathbf{A}_\mathcal{M}{\in}\mathbb{C}^{N\times K}$ which contains only the K "active" columns of $\mathbf{A}$. In the following, $\|\cdot\|_p$ denotes the vector p-norm and $\|\cdot\|_\mathcal{F}$ the matrix Frobenius norm.

Similar to other DOA estimators, K can be estimated by model order selection criteria or by examining the angular spectrum. The parameter K is required only for the noise power in the SBL algorithm. An inaccurate estimate influences the algorithm's convergence.

1.2. Prior on the sources

We assume that the complex source amplitudes x_{ml} are independent both across snapshots. They follow a zero-mean circularly symmetric complex Gaussian distribution with covariance $\mathbf{\Gamma}$,

$$p(\mathbf{X};\mathbf{\Gamma}) = \prod_{l=1}^{L} p(\mathbf{x}_l;\mathbf{\Gamma}) = \prod_{l=1}^{L} \mathcal{CN}(\mathbf{x}_l;\mathbf{0},\mathbf{\Gamma}), \tag{5}$$

i.e., the source vector $\mathbf{x}_l$ at each snapshot $l\in\{1,\cdots,L\}$ is multivariate Gaussian with potentially singular covariance matrix,

$$\mathbf{\Gamma} = \mathsf{E}[\mathbf{x}_l\mathbf{x}_l^H] = \begin{pmatrix} \Gamma_{11} & \cdots & \Gamma_{1M} \\ \vdots & \ddots & \vdots \\ \Gamma_{M1} & \cdots & \Gamma_{MM} \end{pmatrix}, \tag{6}$$

as $\mathrm{rank}(\mathbf{\Gamma})=\mathrm{card}(\mathcal{M})=K\leq M$ (typically $K \ll M$). Note that the diagonal elements of $\mathbf{\Gamma}$ represent source powers and thus $\Gamma_{mm} \geq 0$ for all $1 \leq m \leq M$. When the variance $\Gamma_{mm}=0$, then $x_{ml}=0$ with probability 1. The sparsity of the model is thus controlled by the diagonal elements of $\mathbf{\Gamma}$, and the active set $\mathcal{M}$ is equivalently

$$\mathcal{M} = \{m \in \mathbb{N} | \Gamma_{mm} > 0\}\,. \tag{7}$$

The SBL algorithm estimates $\mathbf{\Gamma}$ rather than the complex source amplitudes $\mathbf{X}$ yielding a significant reduction of the degrees of freedom.

1.3. Stochastic likelihood

We here derive the well-known stochastic maximum likelihood function [17, 18, 19]. Given the linear model (3) with Gaussian source (5) and noise (2) the array data $\mathbf{Y}$ is Gaussian with covariance $\mathbf{\Sigma_y}$ given by

$$\mathbf{\Sigma_y} = \mathsf{E}[\mathbf{y}_l\mathbf{y}_l^H] = \mathbf{\Sigma_n} + \mathbf{A}\mathbf{\Gamma}\mathbf{A}^H \tag{8}$$

The probability density function of $\mathbf{Y}$ is thus given by

$$p(\mathbf{Y}) = \prod_{l=1}^{L} \mathcal{CN}(\mathbf{y}_l;\mathbf{0},\mathbf{\Sigma_y}) = \prod_{l=1}^{L} \frac{\mathrm{e}^{-\mathbf{y}_l^H\mathbf{\Sigma_y}^{-1}\,\mathbf{y}_l}}{\pi^N \det \mathbf{\Sigma_y}}\,. \tag{9}$$

The L-snapshot log-likelihood for estimating $\mathbf{\Gamma}$ and $\mathbf{\Sigma_n}$ is

$$\log p(\mathbf{Y};\mathbf{\Gamma},\mathbf{\Sigma_n}) = -\sum_{l=1}^{L} \left(\mathbf{y}_l^H\mathbf{\Sigma_y}^{-1}\mathbf{y}_l + \log\det\mathbf{\Sigma_y}\right) \tag{10}$$

where we have neglected irrelevant constants. This likelihood function is identical to the Type II likelihood function (evidence) in standard SBL [9, 8, 5] which is obtained by treating $\mathbf{\Gamma}$ as a hyperparameter. The Type II likelihood is obtained by integrating the likelihood function over the complex source amplitudes, cf. (29) in [5]. The stochastic maximum likelihood approach is used here as it is more direct.

The parameter estimates $\hat{\mathbf{\Gamma}}$ and $\widehat{\mathbf{\Sigma_n}}$ are obtained by maximizing the likelihood, leading to

$$(\hat{\mathbf{\Gamma}},\widehat{\mathbf{\Sigma_n}}) = \underset{\mathbf{\Gamma}\succeq\mathbf{0},\ \mathbf{\Sigma_n}\succ\mathbf{0}}{\arg\max}\ \log p(\mathbf{Y};\mathbf{\Gamma},\mathbf{\Sigma_n}), \tag{11}$$

where $\mathbf{\Gamma}\succeq\mathbf{0}$ restricts the maximization over $\mathbf{\Gamma}$ to the domain of positive semi-definite Hermitian matrices and $\mathbf{\Sigma_n}\succ\mathbf{0}$ ensures that $\sigma^2 > 0$ in (2). The likelihood function (10) is similar to the one derived for LIKES [7]. If $\mathbf{\Gamma}$ and $\mathbf{\Sigma_n}$ are known, then the MAP estimate is the posterior mean $\hat{\mathbf{x}}_l^{\mathrm{MAP}}$ and covariance $\mathbf{\Sigma}_{\mathbf{x}_l}$ [20, 5],

$$\hat{\mathbf{x}}_l^{\mathrm{MAP}} = \mathbf{\Gamma}\mathbf{A}^H\mathbf{\Sigma_y}^{-1}\mathbf{y}_l, \tag{12}$$

$$\mathbf{\Sigma_x} = \left(\mathbf{A}^H\mathbf{\Sigma_n}^{-1}\mathbf{A} + \mathbf{\Gamma}^{-1}\right)^{-1}. \tag{13}$$

The diagonal elements of $\mathbf{\Gamma}$ control the sparsity of $\hat{\mathbf{x}}_l^{\mathrm{MAP}}$ as for $\Gamma_{mm} = 0$ the corresponding mth element of $\hat{\mathbf{x}}_l^{\mathrm{MAP}}$ becomes 0.

It is well-known that the maximum-likelihood DOA estimate for a single source in additive white Gaussian noise coincides with the peak finder in the conventional beamformer. In this case, there is also no difference between maximum-likelihood Type I and II estimates for DOA. Here, this means that the solution to (11) coincides with the peak finder in the conventional beamformer for the special case $K = 1$ formulated as the angular power spectrum at DOA θ_m,

$$P_{\text{CBF}}(\theta_m) = \mathbf{a}_m^H\mathbf{Y}\mathbf{Y}^H\mathbf{a}_m/L = \mathbf{a}_m^H\mathbf{S_y}\mathbf{a}_m\,. \tag{14}$$

Here, the sample covariance matrix is defined as

$$\mathbf{S_y} = \mathbf{Y}\mathbf{Y}^H/\,L. \tag{15}$$

2. SOURCE COVARIANCE MATRIX ESTIMATION

For solving (11), the source signal covariance matrix $\mathbf{\Gamma}$ is estimated in two steps. In the first step, the active set $\mathcal{M}$ is estimated iteratively. In the second step, $\mathbf{\Gamma}$ and σ^2 are estimated by assuming $\mathcal{M}$ known. After detailing the estimation procedure, the algorithm is summarized in Table 1.

2.1. Estimation of the active set

We allow correlated sources and thus a *full* $K\times K$ source covariance $\mathbf{\Gamma}_{\mathcal{M}}\succ\mathbf{0}$ and derive (10) with respect to all elements $\Gamma_{mm'}$ with pairs of active indices, cf. [17]. Using

$$\frac{\partial\mathbf{\Sigma_y}^{-1}}{\partial\Gamma_{mm'}} = -\mathbf{\Sigma_y}^{-1}\frac{\partial\mathbf{\Sigma_y}}{\partial\Gamma_{mm'}}\mathbf{\Sigma_y}^{-1} = -\mathbf{\Sigma_y}^{-1}\mathbf{a}_m\mathbf{a}_{m'}^H\mathbf{\Sigma_y}^{-1}, \tag{16}$$

$$\frac{\partial\log\det(\mathbf{\Sigma_y})}{\partial\Gamma_{mm'}} = \mathrm{tr}\left(\mathbf{\Sigma_y}^{-1}\frac{\partial\mathbf{\Sigma_y}}{\partial\Gamma_{mm'}}\right) = \mathbf{a}_{m'}^H\mathbf{\Sigma_y}^{-1}\mathbf{a}_m, \tag{17}$$

the derivative of (10) is formulated as

$$\frac{\partial\log p(\mathbf{Y};\mathbf{\Gamma},\mathbf{V_N})}{\partial\Gamma_{mm'}} = \sum_{l=1}^{L}\mathbf{a}_{m'}^H(\mathbf{\Sigma_y}^{-1}\mathbf{y}_l\mathbf{y}_l^H\mathbf{\Sigma_y}^{-1} - \mathbf{\Sigma_y}^{-1})\mathbf{a}_m \tag{18}$$

The solution to (11) must satisfy the necessary condition $\frac{\partial \log p(\mathbf{Y};\mathbf{\Gamma},\mathbf{V_N})}{\partial \Gamma_{mm'}} = 0$ for all pairs of active indices, $(m, m') \in \mathcal{M} \times \mathcal{M}$.

First, we use (18) for the diagonal elements of $\mathbf{\Gamma}$ only (i.e. $m = m'$) by ignoring the off-diagonal elements of $\mathbf{\Gamma}$. The diagonal $\mathbf{\Gamma}$ approximation is found iteratively by assuming Γ_{mm}^{old} and $\mathbf{\Sigma_y}$ given (from previous iterations or initialization), we obtain the following fixed point iteration [21] for Γ_{mm}:

$$\Gamma_{mm}^{\text{new}} = \Gamma_{mm}^{\text{old}} \left(\frac{\mathbf{a}_m^H \mathbf{U} \mathbf{a}_m}{\mathbf{a}_m^H \mathbf{V} \mathbf{a}_m} \right)^b, \text{ where} \tag{19}$$

$$\mathbf{U} = \sum_{l=1}^{L} \mathbf{\Sigma_y}^{-1} \mathbf{y}_l \mathbf{y}_l^H \mathbf{\Sigma_y}^{-1}, \quad \mathbf{V} = \sum_{l=1}^{L} \mathbf{\Sigma_y}^{-1}. \tag{20}$$

The choice of $b = 1$ gives the update equation used in [20, 3, 6] and $b = 0.5$ gives the update equation used in [5]. Here, we use $b = 0.5$.

The converged $\Gamma_{mm}^{\text{new}} = \Gamma_{mm}^{\text{old}}$ are solely used to estimate the active set $\mathcal{M}$ by thresholding according to (7). The Γ_{mm}^{new} are ignored thereafter.

2.2. Signal Covariance and Noise Variance Estimates

In this section we estimate all elements of $\mathbf{\Gamma}$ and the noise variance σ^2. We assume that all index pairs (m, m') for non-zero $\Gamma_{mm'}$ are known.

When $\mathbf{\Sigma_n} = \sigma^2 \mathbf{I}_N$ with $\mathbf{I}_N$ the identity matrix of size N, stochastic maximum likelihood [11, 14, 16] provides an asymptotically efficient estimate of σ^2 if the set of active DOAs $\mathcal{M}$ is known.

Let $\mathbf{\Gamma}_\mathcal{M}$ be the covariance matrix of the K active sources estimated above with corresponding active steering matrix $\mathbf{A}_\mathcal{M}$ which maximizes (10). The corresponding data covariance matrix is

$$\mathbf{\Sigma_y} = \sigma^2 \mathbf{I}_N + \mathbf{A}_\mathcal{M} \mathbf{\Gamma}_\mathcal{M} \mathbf{A}_\mathcal{M}^H. \tag{21}$$

Note that,the data covariance matrices (8) and (21) are identical. Following [15], we continue from (18),

$$\frac{\partial \log p(\mathbf{Y};\boldsymbol{\gamma},\mathbf{\Sigma_n})}{L\ \partial \Gamma_{mm'}} = \mathbf{a}_m^H \left(\mathbf{\Sigma_y}^{-1} \mathbf{S_y} \mathbf{\Sigma_y}^{-1} - \mathbf{\Sigma_y}^{-1} \right) \mathbf{a}_{m'} \tag{22}$$

$$= \mathbf{a}_m^H \mathbf{\Sigma_y}^{-1} \left(\mathbf{S_y} - \mathbf{\Sigma}_{\mathbf{y}_l} \right) \mathbf{\Sigma_y}^{-1} \mathbf{a}_{m'} = 0, \tag{23}$$

for all pairs of active sources $(m, m' \in \mathcal{M})$. Now we use $\text{range}(\mathbf{\Sigma_y}^{-1} \mathbf{A}_\mathcal{M}) = \text{range}(\mathbf{A}_\mathcal{M})$ and Eq.(23) simplifies to

$$\mathbf{a}_m^H \left(\mathbf{S_y} - \mathbf{\Sigma_y} \right) \mathbf{a}_{m'} = 0, \quad \forall m, m' \in \mathcal{M}\ . \tag{24}$$

This is Jaffer's condition ([15]:Eq.(6)), i.e.,

$$\mathbf{A}_\mathcal{M}^H \left(\mathbf{S_y} - \mathbf{\Sigma_y} \right) \mathbf{A}_\mathcal{M} = \mathbf{0}, \tag{25}$$

which we enforce at the optimal solution $(\mathbf{\Gamma}_\mathcal{M}, \sigma^2)$. Jaffer's condition follows from allowing arbitrary correlations among the source signals, i.e. when the Γ matrix is not restricted to be diagonal. Substituting (21) into (25) gives

$$\mathbf{A}_\mathcal{M}^H \left(\mathbf{S_y} - \sigma^2 \mathbf{I}_N \right) \mathbf{A}_\mathcal{M} = \mathbf{A}_\mathcal{M}^H \mathbf{A}_\mathcal{M} \mathbf{\Gamma}_\mathcal{M} \mathbf{A}_\mathcal{M}^H \mathbf{A}_\mathcal{M}. \tag{26}$$

This suggests the signal covariance estimate

$$\hat{\mathbf{\Gamma}}_\mathcal{M} = \mathbf{A}_\mathcal{M}^{+H} \left(\mathbf{S_y} - \sigma^2 \mathbf{I}_N \right) \mathbf{A}_\mathcal{M}^+, \tag{27}$$

where $\mathbf{A}_\mathcal{M}^+ = (\mathbf{A}_\mathcal{M}^H \mathbf{A}_\mathcal{M})^{-1} \mathbf{A}_\mathcal{M}^H$. Let us then define the projection matrix onto the subspace spanned by the active steering vectors

$$\mathbf{P} = \mathbf{A}_\mathcal{M} \mathbf{A}_\mathcal{M}^+ = \mathbf{A}_\mathcal{M} (\mathbf{A}_\mathcal{M}^H \mathbf{A}_\mathcal{M})^{-1} \mathbf{A}_\mathcal{M}^H = \mathbf{P}^H = \mathbf{P}^2. \tag{28}$$

Left-multiplying (26) with $\mathbf{A}_\mathcal{M}^{+H} = \mathbf{A}_\mathcal{M} (\mathbf{A}_\mathcal{M}^H \mathbf{A}_\mathcal{M})^{-1}$ and right-multiplying it with $\mathbf{A}_\mathcal{M}^+ = (\mathbf{A}_\mathcal{M}^H \mathbf{A}_\mathcal{M})^{-1} \mathbf{A}_\mathcal{M}^H$, we obtain

$$\mathbf{P} \mathbf{S_y} \mathbf{P}^H - \sigma^2 \mathbf{P} \mathbf{P}^H = \mathbf{P} \mathbf{A}_\mathcal{M} \mathbf{\Gamma}_\mathcal{M} \mathbf{A}_\mathcal{M}^H \mathbf{P}^H = \mathbf{A}_\mathcal{M} \mathbf{\Gamma}_\mathcal{M} \mathbf{A}_\mathcal{M}^H$$
$$= \mathbf{\Sigma}_{\mathbf{y}_l} - \sigma^2 \mathbf{I}_N. \tag{29}$$

Evaluating the trace, using $\text{tr}(\mathbf{P}\mathbf{P}^H) = K$ and $\text{tr}(\mathbf{P}\mathbf{S_y}\mathbf{P}^H) = \text{tr}(\mathbf{P}\mathbf{S_y})$, gives

$$\sigma^2 = \frac{\text{tr}[(\mathbf{S_y} - \mathbf{P}\mathbf{S_y}] + \epsilon}{N - K} \approx \frac{\text{tr}[(\mathbf{I}_N - \mathbf{P})\mathbf{S_y}]}{N - K} = \hat{\sigma}^2, \tag{30}$$

where we have defined $\epsilon = \text{tr}[\mathbf{\Sigma}_{\mathbf{y}_l} - \mathbf{S_y}]$.

The above approximation motivates the noise power estimate which is *error-free* if $\text{tr}[\mathbf{\Sigma_y}] = \text{tr}[\mathbf{S_y}]$, *unbiased* because $\text{E}[\epsilon] = 0$, *consistent* since also its variance tends to zero for $L \to \infty$ [22], and *asymptotically efficient* as it approaches the CRLB for $L \to \infty$ [23]. Note that, the estimate (30) is valid for any number of snapshots, even for just one snapshot.

3. EXAMPLES

We carry out simulations to assess the performance of the algorithm in Table 1. We consider a scenario with a uniform linear sensor array with $N = 20$ elements and half wavelength spacing. Three plane waves are arriving ($K = 3$). Two of the plane waves are closely spaced arriving approximately from broadside and the third wave arrives approximately from endfire direction. The corresponding directions of arrival are $\theta_1 = -3°$, $\theta_2 = 2°$, and $\theta_3 = 75°$. The associated three complex-valued amplitudes are jointly Gaussian with zero mean and covariance matrix

$$\mathbf{\Gamma}_\mathcal{M} = \begin{pmatrix} 16 & \Gamma_{12} & \Gamma_{13} \\ \Gamma_{12} & 169 & \Gamma_{23} \\ \Gamma_{13} & \Gamma_{23} & 100 \end{pmatrix}. \tag{31}$$

We investigate two extreme cases in the simulations: uncorrelated plane waves (modeled by $\Gamma_{mm'} = 0$ for $m \neq m'$) and fully correlated plane waves (modeled by $\Gamma_{mm'} = \sqrt{\Gamma_{mm}\Gamma_{m'm'}}$ for $m \neq m'$).

0	Initialize: $\mathbf{\Gamma}^{\text{new}} = \text{diag}[\mathbf{A}^H \mathbf{S_y} \mathbf{A}]$
	$\mathbf{\Sigma}_{\mathbf{n}}^{\text{new}} = \hat{\sigma}^2 \mathbf{I}$ using noise power estimate (30)
	$\epsilon_{\text{min}} = 0.001, \epsilon = 2\epsilon_{\text{min}}, j = 0, j_{\text{max}} = 100$
1	while $(\epsilon > \epsilon_{\text{min}})$ and $(j < j_{\text{max}})$
2	$\mathbf{\Gamma}^{\text{old}} = \mathbf{\Gamma}^{\text{new}},\ \mathbf{\Gamma} = \mathbf{\Gamma}^{\text{old}},\ \mathbf{\Sigma}_{\mathbf{n}}^{\text{old}} = \mathbf{\Sigma}_{\mathbf{n}}^{\text{new}}$
3	$\mathbf{\Sigma_y} = \mathbf{\Sigma}_{\mathbf{n}}^{\text{old}} + \mathbf{A \Gamma A}^H$ (8)
4	$\Gamma_{mm}^{\text{new}} =$ calculate update based on Γ_{mm}^{old} (19)
5	$\mathcal{M} = \{m \in \mathbb{N} \mid K$ largest peaks in $\Gamma_{mm}^{\text{new}}\}$
	$= \{m_1, \ldots, m_K\}$ (7)
6	$\mathbf{A}_{\mathcal{M}} = (a_{m_1}, \ldots, a_{m_K}),\ \mathbf{P} = \mathbf{A}_{\mathcal{M}} \mathbf{A}_{\mathcal{M}}^+$
7	$\hat{\sigma}^2 =$ use noise power estimate (30)
8	$\epsilon = \lVert \boldsymbol{\gamma}^{\text{new}} - \boldsymbol{\gamma}^{\text{old}} \rVert_1 / \lVert \boldsymbol{\gamma}^{\text{old}} \rVert_1,\ \ j = j + 1$
9	$\hat{\mathbf{\Gamma}}_{\mathcal{M}} = \mathbf{A}_{\mathcal{M}}^+ (\mathbf{S_y} - \hat{\sigma}^2 \mathbf{I}_N) \mathbf{A}_{\mathcal{M}}^{+H}$ (27)
10	Output: $\mathcal{M}, \hat{\mathbf{\Gamma}}_{\mathcal{M}}, \hat{\sigma}^2$

Table 1. SBL Algorithm for correlated sources.

The dictionary $\mathbf{A}$ defined in (3) consists of $M = 360$ plane wave steering vectors sampling all bearings in $[0°, 179.5°]$ in steps of $0.5°$.

Figure 1 shows the reconstruction accuracy as evaluated from the true and estimated coherence matrices with elements

$$C_{mm'} = \Gamma_{mm'} / \sqrt{\Gamma_{mm} \Gamma_{m'm'}} \tag{32}$$

for the three waves. The three left images show the case of three uncorrelated waves. Here, the true coherency matrix is $\mathbf{C} = \mathbf{I}_3$. The reconstructed coherency matrices in very high SNR= 100 dB (middle) and very low SNR= -10 dB (bottom) are shown left. The three right images show the case of three fully coherent waves. The true coherency matrix $\mathbf{C}$ with $C_{mm'} = 1\ \forall m, m'$ is shown (top right). The reconstructed coherency matrices in very high SNR= 100 dB (middle) and very low SNR= -10 dB (bottom) are shown right. The SBL estimate computed by the algorithm in Table 1 is very close to the result computed by the method in Ref. [4]. The final paper will discuss the SBL reconstruction accuracy versus SNR and compare the performance with Ref. [4].

4. CONCLUSION

We develop a Sparse Bayesian Learning (SBL) approach to estimate plane wave directions of arrival, their correlations, and the background noise variance from multi-snapshot sensor array data. The plane wave amplitudes are assumed to be correlated zero-mean complex Gaussian distributed with unknown covariance matrix, inspiring a stochastic maximum likelihood approach. The performance of SBL is evaluated in terms of the fidelity of the reconstructed coherency matrix of the estimated plane waves. Simulations indicate that SBL performs well in this setting.

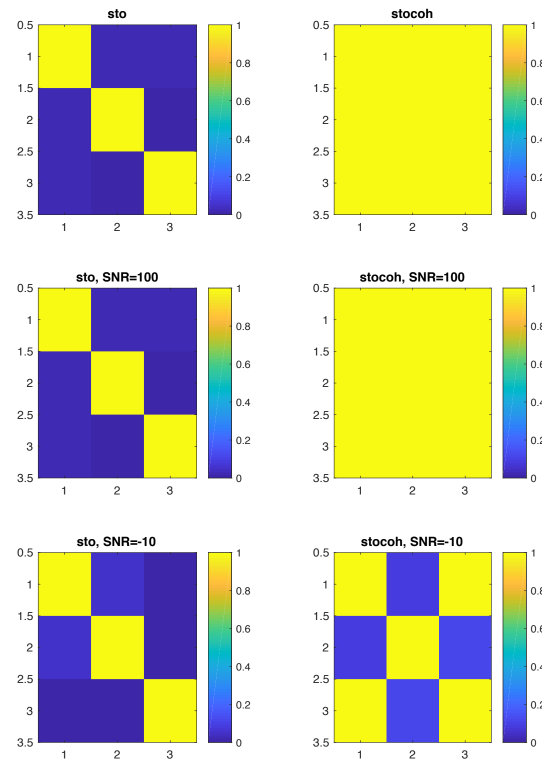

Fig. 1. Coherence matrices for three correlated sources (right) and uncorrelated sources (left). Top row: True coherence matrix. Middle row: Estimated coherence matrix at SNR= 100 dB. Bottom row: SNR= -10 dB.

5. REFERENCES

[1] D. Malioutov, M. Çetin, and A. S. Willsky. A sparse signal reconstruction perspective for source localization with sensor arrays. *IEEE Trans. Signal Process.*, **53**(8):3010–3022, 2005.

[2] A. Xenaki, P. Gerstoft, and K. Mosegaard. Compressive beamforming. *J. Acoust. Soc. Am.*, **136**(1):260–271, 2014.

[3] D. P. Wipf and B.D. Rao. An empirical Bayesian strategy for solving the simultaneous sparse approximation problem. *IEEE Trans. Signal Proc.*, **55**(7):3704–3716, 2007.

[4] P. Gerstoft, A. Xenaki, and C.F. Mecklenbräuker. Multiple and single snapshot compressive beamforming. *J. Acoust. Soc. Am.*, 138(4):2003–2014, 2015.

[5] P. Gerstoft, C. F. Mecklenbräuker, A. Xenaki, and S. Nannuru. Multisnapshot sparse Bayesian learning for DOA. *IEEE Signal Proc. Lett.*, 23(10):1469–1473, 2016.

[6] S. Nannuru, K.L Gemba, P. Gerstoft, W.S. Hodgkiss, and C.F. Mecklenbräuker. Sparse Bayesian learning with uncertainty models and multiple dictionaries. *arXiv:1704.00436*, 2017.

[7] P. Stoica and P. Babu. SPICE and LIKES: Two hyperparameter-free methods for sparse-parameter estimation. *Signal Proc.*, 92(7):1580–1590, 2012.

[8] D. P. Wipf and S. Nagarajan. Beamforming using the relevance vector machine. In *Proc. 24th Int. Conf. Machine Learning*, New York, NY, USA, 2007.

[9] D. P. Wipf and B.D. Rao. Sparse Bayesian learning for basis selection. *IEEE Trans. Signal Proc*, 52(8):2153–2164, 2004.

[10] Z. Zhang and B. D Rao. Sparse signal recovery with temporally correlated source vectors using sparse Bayesian learning. *IEEE J Sel. Topics Signal Proc.,*, 5(5):912–926, 2011.

[11] Z.-M. Liu, Z.-T. Huang, and Y.-Y. Zhou. An efficient maximum likelihood method for direction-of-arrival estimation via sparse Bayesian learning. *IEEE Trans. Wireless Comm.*, 11(10):1–11, Oct. 2012.

[12] JA. Zhang, Z. Chen, P. Cheng, and X. Huang. Multiple-measurement vector based implementation for single-measurement vector sparse Bayesian learning with reduced complexity. *Signal Proc.*, 118:153–158, 2016.

[13] R. Giri and B. Rao. Type I and type II Bayesian methods for sparse signal recovery using scale mixtures. *IEEE Trans Signal Proc. Signal Proc.*, 64(13):3418–3428, July 2016.

[14] J.F. Böhme. Source-parameter estimation by approximate maximum likelihood and nonlinear regression. *IEEE J. Oceanic Eng.*, 10(3):206–212, 1985.

[15] A.G. Jaffer. Maximum likelihood direction finding of stochastic sources: A separable solution. In *IEEE Int. Conf. on Acoust., Speech, and Sig. Proc. (ICASSP-88)*, volume 5, pages 2893–2896, 1988.

[16] P. Stoica and A. Nehorai. On the concentrated stochastic likelihood function in array processing. *Circuits Syst. Signal Proc.*, 14(5):669–674, 1995.

[17] Johann F Böhme. Estimation of spectral parameters of correlated signals in wavefields. *Signal Proc.*, 11(4):329–337, 1986.

[18] H. Krim and M. Viberg. Two decades of array signal processing research: the parametric approach. *IEEE Signal Proc. Mag.*, 13(4):67–94, 1996.

[19] P. Stoica, B. Ottersten, M. Viberg, and R.L. Moses. Maximum likelihood array processing for stochastic coherent sources. *IEEE Trans. Signal Proc.*, 44(1):96–105, 1996.

[20] M. E. Tipping. Sparse Bayesian learning and the relevance vector machine. *J. Machine Learning Research*, 1:211–244, 2001.

[21] R.L. Burden, J.D. Faires, and A.M. Burden. *Numerical analysis*. Cengage Learning, 2016. (chapter 2.2).

[22] D. Kraus. *Approximative Maximum-Likelihood-Schätzung und verwandte Verfahren zur Ortung und Signalschätzung mit Sensorgruppen*. Shaker Verlag, Aachen, Germany, 1993.

[23] G. Bienvenu and L. Kopp. Optimality of high resolution array processing using the eigensystem approach. *IEEE Trans on Acous., Speech, Signal Proc.*, 31(5):1235–1248, Oct 1983.