DOA ESTIMATION IN STRUCTURED PHASE-NOISY ENVIRONMENTS

Angélique Drémeau[⊛]*, *Cédric Herzet*[⁎]

[⊛] ENSTA Bretagne and Lab-STICC UMR 6285, Brest, F-29200, France
[⁎] INRIA Centre Rennes-Bretagne Atlantique, Campus universitaire de Beaulieu, F-35000 Rennes, France

ABSTRACT

In this paper we focus on the problem of estimating the directions of arrival (DOA) of a set of incident plane waves. Unlike many previous works, which assume that the received observations are only affected by additive noise, we consider the setup where some phase noise also corrupts the data (as for example observed in atmospheric sound propagation or underwater acoustics). We propose a new methodology to solve this problem in a Bayesian framework by resorting to a variational mean-field approximation. Our simulation results illustrate the benefits of carefully accounting for the phase noise in the DOA estimation process.

Index Terms— Direction-of-arrival estimation, phase noise, variational Bayesian approximation, mean-field approximation

1. INTRODUCTION

Estimating the directions of arrival (DOA) of propagating plane waves is at the heart of many applicative domains including sonar, radar and mobile telecommunications. Among the rich literature dealing with this problem, the most popular method is probably conventional beamforming [1] which can roughly be interpreted as a least-square estimator. *Per se*, this approach suffers from a lower limit on the resolution achievable in the DOA estimation process (conditioned by the length of the sensor array). To overcome this issue, so-called "high-resolution" techniques, taking advantage of more prior information on (the number and the nature of) the sources, have been proposed in the literature (see *e.g.,* [2, 3, 4, 5]). In [2], the authors introduced the well-known MUSIC algorithm which benefits from the knowledge of the number of the sources and rely on the assumption that the noise and the signal of interest live in perfectly separable subspaces. More recently, a "compressive" beamforming approach was proposed in [3, 4, 5], where a sparse prior was exploited to address the DOA estimation problem.

The contributions mentioned above assume that the incident plane waves are only corrupted by some additive noise. Unfortunately, when the waves travel through highly fluctuating media, as in the case of *e.g.,* atmospheric sound propagation [6] or underwater acoustics [7], this model does no longer describe accurately the physics underlying the propagation process. In such cases, a multiplicative phase noise typically corrupts the collected signal, making the corresponding DOA estimation problem quite challenging. We address this problem in the present paper.

Our approach is inspired from the recent standard "high-resolution" DOA methods [3, 4, 5] and some phase retrieval algorithms presented in [8, 9]. More specifically, we model the received signal as a sparse combination of elementary signals (taken from a redundant dictionary) and assume that the latter is corrupted by both additive and phase noise. Our methodology is grounded on a probabilistic Bayesian framework and relies on a variational mean-field approximation. In particular, we show how to nicely incorporate fine noise-phase models in this framework, extending in this respect, the approaches proposed in [8, 9].

2. PROBLEM STATEMENT

Our derivations are based on the following formulation of the DOA estimation problem: we consider an antenna composed of N sensors and assume that the collected observation vector $\mathbf{y} \in \mathbb{C}^N$ can be expressed as

$$\mathbf{y} = \mathbf{PDz} + \boldsymbol{\omega}, \tag{1}$$

where $\boldsymbol{\omega} \in \mathbb{C}^N$ and $\mathbf{P} = \mathrm{diag}(\{e^{j\theta_n}\}_{n=1}^N) \in \mathbb{C}^{N\times N}$ play respectively the role of an additive and a multiplicative phase noise. Matrix $\mathbf{D} = [\mathbf{d}_1 \dots \mathbf{d}_M] \in \mathbb{C}^{N\times M}$ is made up of the steering vectors $\mathbf{d}_i \triangleq [e^{j\frac{2\pi}{\lambda}\Delta\sin(\phi_i)} \dots e^{j\frac{2\pi}{\lambda}\Delta N\sin(\phi_i)}]^T$, where ϕ_i's are some potential angles of arrival, Δ is the distance between two adjacent sensors, and λ is the wavelength of the propagation waves.

With this formulation, assuming that $\mathbf{y}$ results from the combination of a few waves arriving from different angles ϕ_i's, the DOA estimation problem is basically equivalent to identifying the positions of the nonzero coefficients in $\mathbf{z}$ (since each column of $\mathbf{D}$ corresponds to a particular angle of arrival). In the standard DOA estimation problem, $\mathbf{P}$ is assumed to be known with $\mathbf{P} = \mathbf{I}_N$, where $\mathbf{I}_N$ is the $N \times N$ identity matrix. In this case, the model connecting the unknown vector $\mathbf{z}$ to the measurements $\mathbf{y}$ is linear; finding the

*This work has been supported by the DGA/MRIS.

position of the nonzero elements in $\mathbf{z}$ can then be carried out with standard sparse-representation algorithms, see *e.g.,* [10].

In this paper, we consider the more complex case where $\mathbf{P} = \mathrm{diag}(\{e^{j\theta_n}\}_{n=1}^N)$ and the θ_n's are unknown. More specifically, we assume that the phases θ_n's obey the following Markov model:

$$p(\boldsymbol{\theta}) = \prod_{n=2}^{N} p(\theta_n|\theta_{n-1})\, p(\theta_1), \tag{2}$$

with $p(\theta_n|\theta_{n-1}) = \mathcal{N}(a\,\theta_{n-1}, \sigma_\theta^2)$, $\forall n \in \{2, \dots, M\}$, $a \in \mathbb{R}_+$, and $p(\theta_1) = \mathcal{N}(0, \sigma_1^2)$. From a practical point of view, this model allows us to describe spatial fluctuations of the propagation medium all along the antenna; the strength of the fluctuations is related to the value of parameter σ_θ^2.

As noted in the introduction, assuming that $\mathbf{P}$ is unknown renders the DOA estimation much more difficult since it introduces uncertainties in the observation model. Before proceeding to the presentation of the proposed methodology to address this issue, we draw some connections with other applicative fields:

- Our work relates first to the phase retrieval problem (*e.g.,* [11]) where the phase information of the observations is completely missing: only intensities or amplitudes are acquired. Formally, both problems share - explicitely or not - the same observation model (1) but differ in the prior distribution they enforce on the phases $\boldsymbol{\theta}$, the absence of phase information being modeled by a non-informative prior, such as a uniform law (see [8, 9]).

- Then we note that, in top of being relevant for the DOA estimation problem with fluctuating media, this model is also of interest in the domain of digital communications where it can be used to characterize the transmission of complex modulation symbols over a channel affected by carrier phase noise, see *e.g.,* [12].

3. BAYESIAN FORMULATION OF THE PROBLEM

We address the problem of estimating $\mathbf{z}$ from $\mathbf{y}$ when the realizations of $\boldsymbol{\omega}$ and $\boldsymbol{\theta}$ are unknown. To that end, we first place this problem into a Bayesian framework by defining suitable additional prior distributions on the unknown quantities.

To account for the sparsity of $\mathbf{z}$, we suppose that, $\forall i \in \{1, \dots, M\}$, $z_i = x_i \cdot s_i$ with

$$p(x_i) = \mathcal{CN}(0, \sigma_x^2), \quad \text{and} \quad p(s_i) = \mathrm{Ber}(p_i). \tag{3}$$

This so-called Bernoulli-Gaussian model has been now largely used in the literature to model sparse priors (see *e.g.,* [13]). Next, we assume that $p(\boldsymbol{\omega})$ is a zero-mean Gaussian distribution with variance σ^2. This hypothesis is typically justified by the central limit theorem under the assumption that the additive noise corrupting the data results from the aggregation of a large number of random parasitic contributions. Finally, the probabilistic model describing the behavior of $\boldsymbol{\theta}$ is given by (2). As mentioned in the previous section, this simple model accounts for local variations of the propagation medium along the sensor network.

Based on this probabilistic model, we propose to look for the solution of the following Minimum Mean Square Error (MMSE) problem

$$\hat{\mathbf{z}} = \operatorname*{arg\,min}_{\tilde{\mathbf{z}}} \mathrm{E}_{\mathbf{z}|\mathbf{y}} \left[\|\mathbf{z} - \tilde{\mathbf{z}}\|_2^2 \right], \tag{4}$$

relying on the marginal posterior distribution $p(\mathbf{z}|\mathbf{y}) = \int_{\boldsymbol{\theta}} p(\mathbf{z}, \boldsymbol{\theta}|\mathbf{y})\; d\boldsymbol{\theta}$. The computation of this marginalization being an intractable problem, we propose hereafter a practical procedure based on a mean-field approximation of the joint distribution $p(\mathbf{z}, \boldsymbol{\theta}|\mathbf{y})$ to approach the solution of (4).

4. THE PROPOSED PROCEDURE

Mean-field approximations aim at approximating a posterior joint distribution by another one constrained to have a “simple” factorization while minimizing some distance with the targeted distribution. In the following, we will look for an approximation of $p(\mathbf{z}, \boldsymbol{\theta}|\mathbf{y})$, say $\hat{q}(\mathbf{z}, \boldsymbol{\theta})$, obeying the following factorization $\hat{q}(\mathbf{z}, \boldsymbol{\theta}) = \hat{q}(\boldsymbol{\theta}) \prod_{i=1}^M \hat{q}(z_i)$. With this approximation of $p(\mathbf{z}, \boldsymbol{\theta}|\mathbf{y})$, the evaluation of the marginal with respect to $\mathbf{z}$ is simplified since $\int_{\boldsymbol{\theta}} \hat{q}(\mathbf{z}, \boldsymbol{\theta}) = \prod_i \hat{q}(z_i)$ and the solution of the MMSE problem (4) can be approximated component-wise as

$$\hat{z}_i \simeq \int_{z_i} z_i\; \hat{q}(z_i)\, dz_i.$$

A well-known algorithm to find a mean-field approximation of a target distribution is the so-called “Variational Bayes Expectation-Maximization” (VBEM) algorithm (see *e.g.,* [14]). Particularized to our problem, this procedure searches for a local minimum of the following optimization problem:

$$\hat{q}(\mathbf{z}, \boldsymbol{\theta}) = \operatorname*{arg\,min}_{q} \int_{\mathbf{z}, \boldsymbol{\theta}} q(\mathbf{z}, \boldsymbol{\theta}) \log\left(\frac{q(\mathbf{z}, \boldsymbol{\theta})}{p(\mathbf{z}, \boldsymbol{\theta}|\mathbf{y})}\right) d\mathbf{z}\, d\boldsymbol{\theta}$$
$$\text{subject to } q(\mathbf{z}, \boldsymbol{\theta}) = q(\boldsymbol{\theta}) \prod_{i=1}^{M} q(z_i),$$

by sequentially minimizing the cost function with respect to $q(\boldsymbol{\theta})$ and $q(z_i)$, $\forall i \in \{1, \dots, M\}$.

We show hereafter that the sequence of distributions generated by the VBEM algorithm, say $\{q^{(k)}(\boldsymbol{\theta}), \{q^{(k)}(z_i)\}_i\}_k$, admit simple parametric expressions. We detail these expressions (at a given iteration of the procedure) in the following subsections. Due to space limitation, we do not provide the technical derivations but refer the reader to our companion report [15] for more details. For the sake of clarity, we also omit the iteration index $^{(k)}$ in the notations.

4.1. Update of $q(\boldsymbol{\theta})$

Under the condition[1] of small σ_θ^2, we can express the estimate of $q(\boldsymbol{\theta})$ as a Gaussian distribution:

$$
\begin{aligned}
& q(\boldsymbol{\theta}) = \mathcal{N}(\mathbf{m}_\theta, \Sigma_\theta), \\
\text{where} \quad & \Sigma_\theta^{-1} = \Lambda_\theta^{-1} + \operatorname{diag}\left(\frac{2}{\sigma^2}|\boldsymbol{\eta}|\right), \\
& \mathbf{m}_\theta = \Sigma_\theta \left(\operatorname{diag}\left(\frac{2}{\sigma^2}|\boldsymbol{\eta}|\right) \arg(\boldsymbol{\eta}) \right),
\end{aligned}
$$

and Λ_θ^{-1} is the precision matrix attached to the prior distribution on $\boldsymbol{\theta}$ (2), *i.e.,*

$$
\Lambda_\theta^{-1} = \begin{pmatrix} \frac{1}{\sigma_1^2} + \frac{a^2}{\sigma_\theta^2} & -\frac{a}{\sigma_\theta^2} & 0 & 0 \\ -\frac{a}{\sigma_\theta^2} & \frac{1+a^2}{\sigma_\theta^2} & \ddots & 0 \\ 0 & \ddots & \ddots & -\frac{a}{\sigma_\theta^2} \\ 0 & 0 & -\frac{a}{\sigma_\theta^2} & \frac{1}{\sigma_\theta^2} \end{pmatrix}. \quad (5)
$$

Vector $\boldsymbol{\eta}$ is defined as $\boldsymbol{\eta} \triangleq [\eta_1, \dots, \eta_N]^T$, with

$$\eta_n = y_n \sum_i \langle z_i \rangle^* d_{ni}^*,$$

where d_{ni}^* is the conjugate of the nth element of $\mathbf{d}_i$, and $\langle z_i \rangle$ is defined using the current estimate of $q(z_i)$ as

$$\langle z_i \rangle \triangleq \int_{z_i} z_i \; q(z_i) \, dz_i.$$

Note that the distribution $q(\boldsymbol{\theta})$ being Gaussian, the marginals $q(\theta_n)$ come straightforwardly as

$$q(\theta_n) = \mathcal{N}(m_{\theta_n}, \Sigma_{\theta_n}),$$

where m_{θ_n} (resp. Σ_{θ_n}) is the nth element in $\mathbf{m}_\theta$ (resp. in the diagonal of Σ_θ). In practice, estimating these parameters can be efficiently implemented through a Kalman smoother [16] because of the particular structure of the precision matrix (5).

4.2. Update of $q(z_i)$

The iterates of the $q(z_i)$'s take the form of a mixture of two Gaussian distributions and only depend on the marginal distributions $q(\theta_n)$ of $q(\boldsymbol{\theta})$. More precisely, given that

$$q(z_i) = q(x_i, s_i) = q(x_i|s_i) \, q(s_i),$$

we obtain

$$q(x_i|s_i) = \mathcal{CN}(m_{x_i}(s_i), \Sigma_{x_i}(s_i)), \quad (6)$$

$$q(s_i) \propto \sqrt{\Sigma_{x_i}(s_i)} \exp\left(\frac{m_{x_i}(s_i)^* m_{x_i}(s_i)}{\Sigma_{x_i}(s_i)} \right) p(s_i), \quad (7)$$

where $\propto$ and $.^H$ means respectively proportionality and transpose complex conjugate,

$$\Sigma_{x_i}(s_i) = \frac{\sigma^2 \sigma_x^2}{\sigma^2 + s_i \sigma_x^2 \mathbf{d}_i^H \mathbf{d}_i}, \quad (8)$$

$$m_{x_i}(s_i) = s_i \frac{\sigma_x^2}{\sigma^2 + s_i \sigma_x^2 \mathbf{d}_i^H \mathbf{d}_i} \mathbf{d}_i^H \langle \mathbf{r}_i \rangle, \quad (9)$$

$$\langle \mathbf{r}_i \rangle = \bar{\mathbf{y}} - \sum_{k \neq i} q(s_k = 1) \, m_{x_k}(s_k = 1) \, \mathbf{d}_k, \quad (10)$$

$$\bar{\mathbf{y}} = \left[y_n e^{-j m_{\theta_n}} \, \frac{I_1(1/\Sigma_{\theta_n})}{I_0(1/\Sigma_{\theta_n})} \right]_{n=\{1 \dots M\}}, \quad (11)$$

and I_0 (resp. I_1) stands for the modified Bessel of the first kind of order 0 (resp. 1).

We note that the update equation (6)-(11) share some connections with the phase retrieval algorithm presented in [9]. The latter, relying also on a VBEM algorithm, differs from the proposed procedure in the definition of the prior distributions (2) and (3), respectively replaced by a uniform and a Gaussian distributions. In practice, both procedures share a similar structure. Leaving out the choice made here of a sparse-enforcing prior on $\mathbf{z}$, the main difference lies in the "reconstructed" phases m_{θ_n} in (11): while their definition relies here on the parameters of the Markov chain through the precision matrix (5), they only depend on the observations in [9] where a non-informative prior is considered.

4.3. Noise estimation

As emphasized in [13], the estimation of model parameters can easily be embedded within the VBEM procedure. Among them, the noise variance is of particular interest. Measure of the (mean) discrepancies between the observations and the assumed model, its iterative estimation usually helps the convergence of the algorithm to a proper local minimum, as observed in [13]. Particularized to model (1)-(3), this leads to

$$
\begin{aligned}
\hat{\sigma}^2 &= \operatorname*{arg\,max}_{\sigma^2} \int_{\mathbf{z}} q(\mathbf{z}, \boldsymbol{\theta}) \log p(\mathbf{z}, \boldsymbol{\theta}, \mathbf{y}; \sigma^2) \, d\mathbf{z} \, d\boldsymbol{\theta}, \\
&= \frac{1}{N} \Bigg(\mathbf{y}^H \mathbf{y} - 2 \sum_i \Re\Big[q(s_i{=}1) \, m_{x_i}(s_i{=}1) \bar{\mathbf{y}}^H \mathbf{d}_i \Big] \\
&+ \sum_i \sum_{k \neq i} q(s_i{=}1) \, q(s_k{=}1) \, m_{x_i}(s_i{=}1)^* \, m_{x_k}(s_k{=}1) \, \mathbf{d}_i^H \mathbf{d}_k \\
&+ \sum_i q(s_i{=}1) \Big(\Sigma_{x_i}(s_i{=}1) + \big| m_{x_i}(s_i{=}1) \big|^2 \Big) \mathbf{d}_i^H \mathbf{d}_i \Bigg).
\end{aligned}
$$

In the following, we will refer to the proposed procedure as "paVBEM" for "phase-aware VBEM algorithm".

5. EXPERIMENTS

In this section, we assess numerically the effectiveness of the proposed approach. We consider the problem of identifying

[1] We justify this condition in the technical report [15].

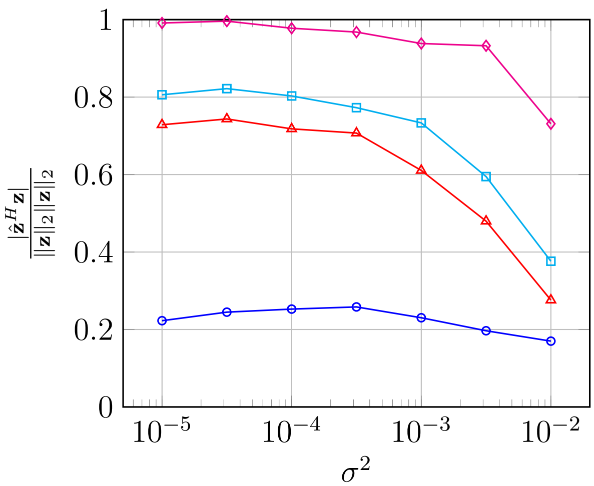

Fig. 1. Evolution of the (averaged) normalized correlation as a function of the variance σ^2 when $K = 2$.

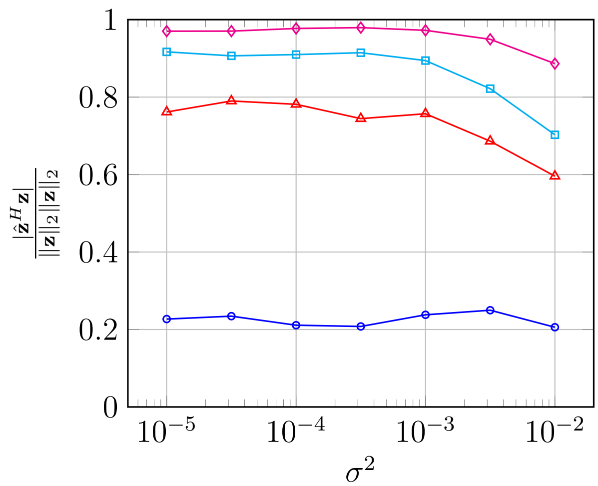

Fig. 2. Evolution of the (averaged) normalized correlation as a function of the variance σ^2 when $K = 5$.

the directions of arrival of K plane waves from $N = 256$ observations (K will be specified later on). We assume that the angles of the K incident waves can be written as $\phi_k = -\frac{\pi}{2} + i_k \frac{\pi}{50}$ with $i_k \in [1, 50]$, $\forall k \in \{1, \ldots, K\}$. The set of angles $\{\phi_i = -\pi + i\frac{\pi}{50}\}_{i\in\{1,\ldots,50\}}$ together with the choice of the parameter $\Delta/\lambda = 4$ define the columns of the dictionary $\mathbf{D}$ (see section 2). We set the following parameters for the phase Markov model (2): $\sigma_1^2 = 10^6$, $\sigma_\theta^2 = 1$ and $a = 0.8$. This corresponds to the situation where one has a large uncertainty on the initial value of the phase noise but connections exist between the phase noise on adjacent sensors.

As well as for the classic phase retrieval problem, the solution of (4) can only be found up to a global phase. Hence, computing the mean square distance between the ground truth $\mathbf{z}$ and its reconstruction $\hat{\mathbf{z}}$ does not constitute a relevant figure of merit. Instead, we consider their normalized correlation, $\frac{|\mathbf{z}^H\hat{\mathbf{z}}|}{\|\mathbf{z}\|_2\|\hat{\mathbf{z}}\|_2}$. This quantity is averaged over 50 realizations for each point of simulation. We compare the performance of the following algorithms: *i)* the conventional beamforming introduced in [1] (—○—); *ii)* the so-called *prVBEM* algorithm proposed in [9] as a solution to the phase retrieval problem (—△—); *iii)* the *paVBEM* procedure described in section 4 (—◇—); *iv)* a relaxed version of *paVBEM* in which the sparsity of $\mathbf{z}$ is not exploited but replaced by a Gaussian prior (—□—).

The performance of these procedures are illustrated in Fig. 1 and 2 as a function of the noise variance σ^2 for $K = 2$ and $K = 5$, respectively. We see that the beamforming algorithm, which was originally proposed to solve the DOA estimation problem in the standard linear setup ($\mathbf{P} = \mathbf{I}_N$), fails to cope with the presence of fluctuations in the phase $\boldsymbol{\theta}$. The three other algorithms achieve different levels of performance, depending on the power of the additive noise and the number of incident waves. We note that all these procedures are derived from a similar optimization procedure but consider different degrees of knowledge on $\boldsymbol{\theta}$ and $\mathbf{z}$. In [9], the authors assume that the phase is uniformly distributed and the sparse nature of $\mathbf{z}$ is ignored; in the relaxed version of the proposed procedure (see [15]) the phase model (2) is exploited but the sparsity of $\mathbf{z}$ is not taken into account; finally, as explained previously, the methodology presented in this paper integrates both the phase model (2) and the sparsity of $\mathbf{z}$ in the estimation process. We see from Fig. 1 and 2 that the performance of these algorithms directly relates to the level of information they exploit: the proposed methodology outperforms its relaxed counterpart which, in turn, leads to better performance than the procedure proposed in [9].

We also notice that the procedures achieve better performance when $K = 5$ than $K = 2$. This counter-intuitive behavior is typical for phase-retrieval problems. In fact, it is easy to see that, when no phase information is available (*i.e.,* θ_n is uniform on $[0, 2\pi]$), only the modulus of y_n provides some information on $\mathbf{z}$. In such a case, the worst situation occurs when $K = 1$ since all the elements of $\mathbf{D}$ has the same modulus (equal to 1) and the observations thus provide no information on $\mathbf{z}$. The setup $K = 2$ is close to this worst case, hence explaining the observed behavior.

6. CONCLUSION

We have presented a novel algorithm able to estimate DOA in environments corrupted by phase noise. Our approach relies on a mean-field approximation and exploits two types of priors: on the DOA through a sparse-enforcing distribution and on the phase noise through a Markov model. Our experiments have confronted the proposed approach to conventional beamforming and similar variational approaches handicapped by non-informative priors. In this regard, its good performance tends to prove a successful inclusion of the priors. Future work will include further assessment in underwater acoustics.

7. REFERENCES

[1] D. H. Johnson and D. E. Dudgeon, *Array signal processing: concepts and techniques*, Englewood Cliffs, NJ, 1993.

[2] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propagation*, vol. AP-34, pp. 276–280, 1986.

[3] W. Mantzel, J. Romberg, and K. Sabra, "Compressive matched-field processing," *Journal of Acoustical Society of America*, vol. 132, no. 1, pp. 90–102, 2012.

[4] A. Xenaki, P. Gerstoft, and K. Mosegaard, "Compressive beamforming," *Journal of Acoustical Society of America*, vol. 136, no. 1, pp. 260–271, 2014.

[5] S. Fortunati, R. Grasso, F. Gini, M. S. Greco, and K. Lepage, "Single snapshot DOA estimation by using compressed sensing," *EURASIP Journal on Advances in Signal Processing*, vol. 120, 2014.

[6] S. Cheinet, L. Ehrhardt, D. Juvé, and P. Blanc-Benon, "Unified modeling of turbulence effects on sound propagation," *Journal of Acoustical Society of America*, vol. 132, no. 4, pp. 2198–2209, 2012.

[7] R. Dashen, S.M. Flatté, W.H. Munk, K.M. Watson, and F. Zachariasen, *Sound Transmission Through a Fluctuating Ocean*, Cambridge Monographs on Mechanics. Cambridge University Press, 2010.

[8] P. Schniter and S. Rangan, "Compressive phase retrieval via generalized approximate message passing," in *Communication, Control, and Computing (Allerton)*, October 2012.

[9] A. Drémeau and F. Krzakala, "Phase recovery from a bayesian point of view: the variational approach," in *Proc. IEEE Int'l Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Brisbane, Australia, April 2015, pp. 3661–3665.

[10] S. Foucart and H. Rauhut, *A mathematical introduction to compressive sensing.*, Applied and Numerical Harmonic Analysis. Birkhaüser, 2013.

[11] E. J. Candès, T. Strohmer, and V. Voroninski, "Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming," *Communications on Pure and Applied Mathematics*, vol. 66, no. 8, pp. 1241–1274, 2013.

[12] G. Colavolpe, A. Barbieri, and G. Caire, "Algorithms for iterative decoding in the presence of strong phase noise," *IEEE Journal on selected areas in communications*, vol. 23, no. 9, pp. 1748 – 1757, 2005.

[13] A. Drémeau, C. Herzet, and L. Daudet, "Boltzmann machine and mean-field approximation for structured sparse decompositions," *IEEE Trans. On Signal Processing*, vol. 60, no. 7, pp. 3425–3438, July 2012.

[14] M. J. Beal and Z. Ghahramani, "The variational bayesian EM algorithm for incomplete data: with application to scoring graphical model structures," *Bayesian Statistics*, vol. 7, pp. 453–463, 2003.

[15] A. Drémeau and C. Herzet, "DOA estimation in structured phase-noisy environments: technical report," Tech. Rep., available at http://arxiv.org:1609.03503, September 2016.

[16] J. M. Mendel, *Lessons in Estimation Theory for Signal Processing Communications and Control*, Prentice Hall Signal Processing Series, Englewood Cliffs, NJ, 1995.