BAYESIAN INFERENCE FOR MULTI-LINE SPECTRA IN LINEAR SENSOR ARRAY

Viet Hung Tran[⋆] *Wenwu Wang*[⋆] *Yuhui Luo*[†] *Jonathon Chambers*[⋆⋆]

[⋆] CVSSP, University of Surrey, U.K.
[†] National Physical Laboratory (NPL), U.K.
[⋆⋆] Newcastle University, U.K.

ABSTRACT

For a linear sensor array, using line spectra is a common technique for estimating directions of arrival (DOA) of single-tone sources. Yet, very few papers consider multi-tone sources. For the first time, we provide the optimal Bayesian inference for multi-line spectra, i.e. a superposition of line spectra, and estimate the DOAs of the multi-tone sources. For tractable computation via fast Fourier transform, we apply a grid-based method, in which source's tones and sensor's array measure are both uncorrelated. Exploiting this method, we interpret the superposition of sensor's data as a complex Gaussian mixture of multi-tone signals. We then estimate DOA via conjugate Von-Mises, also known as circular Gaussian distribution. Our simulation shows that the multi-tone method is superior to traditional single-tone method for detecting multi-tone source's frequencies, particularly for the sources with overlapping frequencies. The posterior DOA's resolution can be tuned via Von-Mises' parameter *a priori*, which enhances the sparsity of DOA's estimation.

***Index Terms*—** DOA, Von-Mises distribution, multi-tone sources, sensor array, LASSO

I. INTRODUCTION

For a linear sensor array, it is well-known that a far distant source with different directions of arrival (DOA) oscillates the steering array's measure with different angular frequencies [1], for which the line spectrum is a popular estimation method [2], [3]. The superposition of the array's spatial frequencies, however, becomes challenging for estimation, particularly when the number of sources is unknown.

For a small number of sources, the array's line spectra are sparse and can be estimated effectively via sparsity-motivated ℓ_1-normed techniques like atomic norm [1], [2] and LASSO [4], [5]. The near optimal bound for the atomic norm approach was also given in [6]. When the number of sources is high, the array's superposition was then modeled as a mixture of all potential mono-tone sources, whose appearance is given by a boolean indication vector. The problem is then to estimate DOA in tandem with the indication vector.

Given noisy data, Bayesian inference is an optimal method for estimating model's parameters in terms of least averaged minimum-risk, of which the Mean Square Error (MSE) is a special case [7], [8]. However, the Bayesian indicator's inference is intractable, since the number of indication vector's possible values grows exponentially with the number of sources. Thus the iterative Variational Bayes approximation [9], which is popular in mixture context, was proposed for the line spectra problem [3], [10]. In the sparse context, these Bayesian indicator techniques are shown equivalent to the ℓ_1-normed LASSO techniques, if the prior is a Laplacian distribution [4], [10], [11].

From the works on DOA, we recognize that very few papers consider the case of multi-tone sources, even though such multi-tone sources appear frequently in practice. All narrow band and frequency-overlapping sources are examples of this case. In this paper, we extend the above indication vector to an indication matrix, which, in contrast to the above works, allows us to detect a source with more than one frequency component at each DOA. For fast computation without losing the optimality of the Bayesian method, we will exploit the uncorrelated property of frequency components via the fast Fourier transform (FFT) method. This special property allows the factorization of DOA's posterior distributions and, hence, yields linear computational complexity. To our knowledge, this is the first time that the exact posterior distribution can be derived for DOA of multi-tone sources. Instead of Laplace, we also enhance the sparsity via the Von-Mises prior distributions, which is conjugate to directional angle of additive white Gaussian noisy observation. Hence, our novel approach with indication matrix can be regarded as a generalization of LASSO method for the DOA problem.

II. ARRAY DATA MODEL

At time $t \in \{1, 2, \ldots, N\}$, let $\boldsymbol{x}_t \triangleq [x_{1,t}, \ldots, x_{D,t}]^T \in \mathbb{C}^D$ and $\boldsymbol{s}_t \triangleq [s_{1,t}, \ldots, s_{M,t}]^T \in \mathbb{C}^M$ be a complex-valued snapshot of D sensors and M known tones respectively, with $s_{m,t} \triangleq \varsigma_m e^{j\gamma_m i}$, $i \in \{1, 2, \ldots, D\}$, where $\varsigma_m \in \mathbb{C}$

This work was supported by the Engineering and Physical Sciences Research Council (EPSRC) Grant number EP/K014307/2 and the MOD University Defence Research Collaboration in Signal Processing. Emails: v.tran@surrey.ac.uk, w.wang@surrey.ac.uk, yuhui.luo@npl.co.uk, jonathon.chambers@newcastle.ac.uk.

$$\underbrace{\begin{bmatrix} \overbrace{x_{1,1}}^{\boldsymbol{x}_1} & \cdots & \overbrace{x_{1,N}}^{\boldsymbol{x}_N} \\ \vdots & & \vdots \\ x_{D,1} & \cdots & x_{D,N} \end{bmatrix}}_{\boldsymbol{X}_{D\times N}} = \underbrace{\begin{bmatrix} \overbrace{e^{j\omega_1}}^{\boldsymbol{a}_1(\omega_1)} & \cdots & \overbrace{e^{j\omega_K}}^{\boldsymbol{a}_K(\omega_K)} \\ \vdots & & \vdots \\ e^{j\omega_1 D} & \cdots & e^{j\omega_K D} \end{bmatrix}}_{\boldsymbol{A}_{D\times K}} \underbrace{\begin{bmatrix} \overbrace{l_{1,1}}^{\boldsymbol{l}_1} & \cdots & \overbrace{l_{1,M}}^{\boldsymbol{l}_M} \\ \vdots & & \vdots \\ l_{K,1} & \cdots & l_{K,M} \end{bmatrix}}_{\boldsymbol{L}_{K\times M}} \underbrace{\begin{bmatrix} \overbrace{s_{1,1}}^{\boldsymbol{s}_1} & \cdots & \overbrace{s_{1,N}}^{\boldsymbol{s}_N} \\ \vdots & & \vdots \\ s_{M,1} & \cdots & s_{M,N} \end{bmatrix}}_{\boldsymbol{S}_{M\times N}} + \boldsymbol{Z}_{D\times N}, \quad (1)$$

is complex amplitude and $\gamma_m \in [0, \pi)$ is temporal angular frequency, $m \in \{1, 2, \ldots, M\}$. Without loss of generality, we assume $\varsigma_m = 1$, $\forall m$. In order to apply the fast Fourier transform (FFT), we assume all frequency tones fall into discrete Fourier transform (DFT) bins $\frac{2\pi}{N}$, i.e. $\gamma_m \in \left\{0, \frac{2\pi}{N} \ldots, (N-1)\frac{2\pi}{N}\right\}$.

Let $\boldsymbol{X} \triangleq [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N] \in \mathbb{C}^{D\times N}$ and $\boldsymbol{S} \triangleq [\boldsymbol{s}_1, \boldsymbol{s}_2, \ldots, \boldsymbol{s}_N] \in \mathbb{C}^{M\times N}$ denote the matrix of sensor's output and frequency components over N time points, respectively. Let us also call $\boldsymbol{A} \triangleq [\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_K] \in \mathbb{C}^{D\times K}$ a steering array matrix, whose $\{i, k\}$-element is $a_{i,k} \triangleq e^{j2\varrho_i\omega_k} = e^{j\omega_k i}$, $i \in \{1, 2, \ldots, D\}$, with radius $\varrho_i \triangleq \frac{\lambda}{2} i \in \mathbb{R}$ denoting positions of D sensors spaced at half the unit wavelength $\lambda = 1$ and spatial angular frequencies $\omega_k \triangleq \pi \cos\phi_k \in [0, 2\pi)$ corresponding to upper half-space arrival angle $\phi_k \in [0, \pi)$, $k \in \{1, 2, \ldots, K\}$.

Let us now define a boolean incidence matrix $\boldsymbol{L} = [\boldsymbol{l}_1, \boldsymbol{l}_2, \ldots, \boldsymbol{l}_M] \in \mathbb{I}^{K\times M}$, whose $\{k, m\}$-element $l_{k,m} \in \mathbb{I} \triangleq \{0, 1\}$ links the mth frequency component to the arrival angle of the kth source. Our data model for the linear sensor array is then written in matrix form in (1), in which $\boldsymbol{Z}$ is a $D \times N$ matrix of complex additive white Gaussian noise (AWGN) with power σ^2. Element-wise, we can rewrite our model (1) as follows:

$$x_{i,t} = \sum_{k=1}^{K}\sum_{m=1}^{M} l_{k,m} e^{j\omega_k i} s_{m,t} + z_{i,t}, \qquad (2)$$

which is a linear mixture of multi-tone sources. The DOA problem is now equivalent to an estimation problem of spatial frequencies $\boldsymbol{\omega} \triangleq [\omega_1, \omega_2, \ldots, \omega_K]^T$ and their associated incidence matrix $\boldsymbol{L}$. In the literature, the incidence matrix $\boldsymbol{L}$ in (1) was reduced to a diagonal matrix of boolean values [3], [10], which corresponds to the case of single-tone sources.

III. BAYESIAN MODEL

From matrix (1) and element form (2), the observation model is a complex Gaussian distribution $f(\boldsymbol{X}|\boldsymbol{\omega}, \boldsymbol{L}) = \prod_{t=1}^{N}\prod_{i=1}^{D} \mathcal{CN}_{x_{i,t}}(\sum_{k=1}^{K}\sum_{m=1}^{M} l_{k,m} a_{i,k} s_{m,t}, \sigma^2) = \prod_{t=1}^{N} \mathcal{CN}_{\boldsymbol{x}_t}(\boldsymbol{A}\boldsymbol{L}\boldsymbol{s}_t, \sigma^2 \mathbf{I}_D)$, in which $\mathbf{I}_D$ is a $D \times D$ identity matrix, as follows:

$$f(\boldsymbol{X}|\boldsymbol{\omega}, \boldsymbol{L}) = \frac{1}{(\pi\sigma^2)^{ND}} \exp\left(-\frac{\sum_{t=1}^{N} |\boldsymbol{x}_t - \boldsymbol{A}\boldsymbol{L}\boldsymbol{s}_t|^2}{\sigma^2}\right). \qquad (3)$$

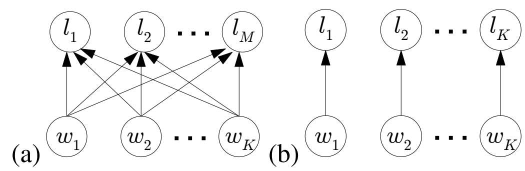

Fig. 1. Directed acyclic graph (DAG) for posterior distribution of indicators and DOAs in case of (a) multi-tone sources and (b) single-tone sources. All the tones are assumed uncorrelated.

Let $\widehat{\boldsymbol{X}} \triangleq (\boldsymbol{S}\boldsymbol{X}^*)^T = \sum_{t=1}^{N} (\boldsymbol{s}_t \boldsymbol{x}_t^*)^T$ be the DFT of the array data and $\widehat{\boldsymbol{\Sigma}} \triangleq \boldsymbol{S}\boldsymbol{S}^* = \sum_{t=1}^{N} [\boldsymbol{s}_t \boldsymbol{s}_t^*]^T$ be the covariance matrix of tone components. For FFT, we assume that all source's tones are set at DFT bins and, hence, $\widehat{\boldsymbol{\Sigma}}$ is a diagonal matrix, i.e. $\widehat{\boldsymbol{\Sigma}} = N\mathbf{I}_M$. Then, in (3), we have:

$$\begin{aligned} \sum_{t=1}^{N} |\boldsymbol{x}_t - \boldsymbol{A}\boldsymbol{L}\boldsymbol{s}_t|^2 &= \sum_{t=1}^{N} (|\boldsymbol{x}_t|^2 - 2\mathrm{Re}\{\boldsymbol{x}_t^* \boldsymbol{A}\boldsymbol{L}\boldsymbol{s}_t\} + |\boldsymbol{A}\boldsymbol{L}\boldsymbol{s}_t|^2), \\ \sum_{t=1}^{N} \boldsymbol{x}_t^* \boldsymbol{A}\boldsymbol{L}\boldsymbol{s}_t &= \mathbf{1}_D^T \widehat{\boldsymbol{X}} \circ (\boldsymbol{A}\boldsymbol{L}) \mathbf{1}_M = \sum_{k=1}^{K}\sum_{m=1}^{M} \widehat{\boldsymbol{x}}_m^T \boldsymbol{a}_k l_{k,m}, \\ \sum_{t=1}^{N} |\boldsymbol{A}\boldsymbol{L}\boldsymbol{s}_t|^2 &= \mathbf{1}_M^T \widehat{\boldsymbol{\Sigma}} \circ \left(\boldsymbol{L}^T \boldsymbol{A}^* \boldsymbol{A}\boldsymbol{L}\right) \mathbf{1}_M = ND \sum_{k=1}^{K}\sum_{m=1}^{M} l_{k,m}, \end{aligned} \qquad (4)$$

in which $*$, T, $\circ$, $\widehat{\boldsymbol{x}}_m$ and $\mathbf{1}_M$ denote conjugate transpose, matrix transpose, Hadamard product, mth column of $\widehat{\boldsymbol{X}}$ and unit column vector of length M with all elements 1, respectively. Note that, in (4), we have $\boldsymbol{A}^*\boldsymbol{A} = D\mathbf{I}_K$ for the case of uncorrelated sensor's steering matrix. Our assumption on uncorrelated tones and steering matrix is not restrictive in practice, since we can always sample more data and increase the number of sensors in order to increase the resolution of DFT bins in FFT, respectively. For single-tone sources, the matrix $\boldsymbol{L}$ is diagonal and, hence, the observations (3-4) and posteriors (5) are completely factorized, as shown in Fig. 1.

III-A. Non-informative prior

Let us firstly assume non-informative, i.e. uniform, priors $f(\boldsymbol{\omega}, \boldsymbol{L}) = f(\boldsymbol{L})f(\boldsymbol{\omega})$, in which the links $\boldsymbol{L}$ follow Bernoulli distributions with equal probabilities and all angular frequencies ω_k are uniform *a priori* over $[0, 2\pi)$. By Bayes' rule, substituting the uncorrelated forms (4) to (3), we can factorize the joint posterior $f(\boldsymbol{\omega}, \boldsymbol{L}|\boldsymbol{X}) \propto f(\boldsymbol{X}|\boldsymbol{\omega}, \boldsymbol{L})$, as follows:

$$f(\boldsymbol{\omega}, \boldsymbol{L}|\boldsymbol{X}) \propto \prod_{k=1}^{K}\prod_{m=1}^{M}\left(\prod_{i=1}^{D}\prod_{t=1}^{N}\mathcal{CN}_{x_{i,t}}(l_{k,m}a_{i,k}s_{m,t}, \sigma^2)\right)$$
$$= \prod_{k=1}^{K}\prod_{m=1}^{M} f(\omega_k, l_{k,m}|\boldsymbol{X}), \quad (5)$$

which resembles the element-wise form in (2). Also, from (3-4) and (5), we can derive the following equivalent forms:

$$f(\omega_k, l_{k,m}|\boldsymbol{X}) \propto \frac{1}{\pi\sigma^2}\exp\left(-\frac{||\boldsymbol{X} - l_{k,m}\boldsymbol{a}_k(\omega_k)\boldsymbol{s}_{m,:}||^2}{\sigma^2}\right)$$
$$= \frac{1}{\varrho_{k,m}}\exp\left(\frac{2\mathrm{Re}\{\widehat{\boldsymbol{x}}_m^T\boldsymbol{a}_k(\omega_k)\} - ND}{\sigma^2} l_{k,m}\right)$$
$$= \frac{\xi_k}{\varrho_{k,m}}\frac{\prod_{i=1}^{D}\mathcal{VM}_{\omega_k i}\left(\kappa_{k,m}l_{k,m}\right)}{\exp\left(\frac{ND}{\sigma^2}l_{k,m}\right)}, \quad (6)$$

where the $||\cdot||^2$ operator is the squared ℓ_2-norm, $\boldsymbol{s}_{m,:}$ is the mth row vector of $\boldsymbol{S}$, $\varrho_{k,m}$ is a normalizing constant, $\xi_k = \prod_{m=1}^{M} 2\pi I_0(|\kappa_{m,k}|)$, $I_q(\cdot)$ is the modified Bessel function of the first kind with order q, the parameter of the Von-Mises distribution is $\kappa_{k,m} \triangleq \frac{2}{\sigma^2}\widehat{x}^*_{k,m}$ [12] and $\widehat{x}_{i,m}$ is the $\{i,m\}$th-element of the FFT matrix $\widehat{\boldsymbol{X}}$. Then, from (5-6), we can derive feasibly the marginal posteriors $f(\omega_k|\boldsymbol{X}) = \sum_{\boldsymbol{l}_{k,:}} f(\omega_k, \boldsymbol{l}_{k,:}|\boldsymbol{X}) = \prod_{m=1}^{M}\sum_{l_{k,m}} f(\omega_k, l_{k,m}|\boldsymbol{X})$ and $f(l_{k,m}|\boldsymbol{X}) = \int_0^{2\pi} f(\omega_k, l_{k,m}|\boldsymbol{X})d\omega_k$, where $\boldsymbol{l}_{k,:} \triangleq [l_{k,1}, \ldots, l_{k,M}]$, as follows:

$$f(\omega_k|\boldsymbol{X}) = \frac{1}{\varrho_{k,m}}\prod_{m=1}^{M}\left(1 + \varrho_{k,m}f(\omega_k, l_{k,m} = 1|\boldsymbol{X})\right),$$
$$f(l_{k,m}|\boldsymbol{X}) = \mathrm{Ber}(p_{k,m}),\; p_{k,m} \triangleq \frac{\chi_{k,m}}{\chi_{k,m}+1}, \quad (7)$$

where $\chi_{k,m} = \int_0^{2\pi}\frac{1}{2\pi}\varrho_{k,m}f(\omega_k, l_{k,m} = 1|\boldsymbol{X})d\omega_k$. Since the i-fold wrapped Von-Mises distribution in (6) is not given in closed-form, it can be evaluated via a grid-based quantization method [13]. For closed-form solution, we can approximate (6) by a mixture of Von-Mises distributions at periodic locations. Since wrapped Von-Mises distribution is periodic over circle $[0, 2\pi)$, this approximation yields high accuracy, as shown in [3].

In observation model (2-3), array data $\boldsymbol{X}$ are superposition of KM components, each of which is indicated by $l_{k,m}$ and a bi-linear product of the steering array $\boldsymbol{a}_k(\omega_k)$ and the mth signal's tone $\boldsymbol{s}_{m,:}$. Then, the posterior probability (6) of DOA's ω_k and each component's indicator $l_{k,m}$ is inversely proportional to the Euclidean distance between $\boldsymbol{X}$ and each of the components $\boldsymbol{a}_k(\omega_k)\boldsymbol{s}_{m,:}$. This appearance probability is highest when ω_k and $\boldsymbol{s}_{m,:}$ are closest to the true value of DOA and signal's tones, respectively.

Nevertheless, the non-informative prior on ω_k yields ambiguous and possibly overlapping DOA's values in (6).

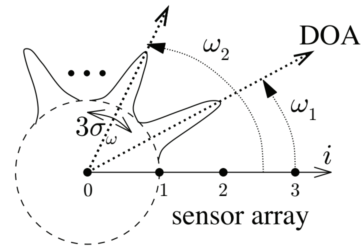

Fig. 2. Von-Mises prior distributions, also known as circular Normal distribution, for DOA's angular frequency ω. Mean prior values of ω are spanned equally over $[0, 2\pi)$. Each Von-Mises is mostly contained within three standard deviation $3\sigma_\omega$.

The reason is that our observation model (2-3) imposes no difference between $\omega_1, \ldots, \omega_K$. In order to avoid this ambiguity, each ω_k should have distinct properties, which can be imposed via conjugate priors, as shown below. Note that, in the case of a single-tone, we can simply set $\boldsymbol{L}$ as a diagonal matrix in above formulae. Hence, the single-tone method does not suffer from this ambiguity of DOA, since each DOA is associated with only one distinct tone at all time.

III-B. Conjugate prior

Let us now consider an informative prior $f(\boldsymbol{\omega}, \boldsymbol{L}) = f(\boldsymbol{\omega})f(\boldsymbol{L})$, in which the links $\boldsymbol{L}$ are uniform *a priori* like above. For angular frequency $\omega_k \in [0, 2\pi)$, we set $f(\boldsymbol{\omega}) = \prod_{k=1}^{K}\mathcal{VM}_{\omega_k}(\beta_k)$, which is conjugate to Von-Mises distributions in (6) [3], [14]. The concentrating parameters $\boldsymbol{\beta} \triangleq [\beta_1, \beta_2, \ldots, \beta_K]^T \in \mathbb{C}^K$ can be tuned such that, for each ω_k, the mean $\mathbb{E}(\omega_k) = \angle\beta_k = \frac{2\pi}{K}k$ and variance $\sigma^2_\omega \triangleq \mathrm{var}(\omega_k) = 1 - \frac{I_1(|\beta_k|)}{I_0(|\beta_k|)} \approx \frac{1}{|\beta_k|}$, $k \in \{1, 2, \ldots, K\}$, are spanned equally over $[0, 2\pi)$ *a priori*, as illustrated in (2). With this conjugate prior, there is an *a priori* knowledge β_k in posterior of each ω_k, i.e.:

$$f(\omega_k|\boldsymbol{X}, \beta_k) \propto f(\omega_k|\boldsymbol{X})\mathcal{VM}_{\omega_k}(\beta_k).$$

The angle's estimation $\hat{\phi}_k = \arccos\frac{\mathbb{E}(\omega_k)}{\pi}$ can be estimated via posterior mean of ω_k, which differs from non-informative case (7) by a concentrating variance $\sigma^2_\omega \approx \frac{1}{|\beta_k|}$ around offset means $\angle\beta_k$, $k \in \{1, 2, \ldots, K\}$. In this paper, we set $\sigma_\omega = \frac{1}{3K}\sqrt{\frac{\sigma^2}{N}}$, which is enough to separate K Von-Mises distributions equally over K parts of a circle, as illustrated in Fig. 2. The ratio $\frac{\sigma^2}{N}$ is the inverse of signal-to-noise ratio (SNR), where N is the sum of tone's power unit over N snapshots, as shown in (4). Hence, the higher the SNR, the smaller is the standard deviation σ_ω and the Von-Mises prior is closer to Gaussian distribution around the signal's value. Reversely, the lower the SNR, the higher is σ_ω and Von-Mises prior is closer to uniform distribution [12]. In the literature, this ratio $\frac{\sigma^2}{N}$ also resembles the near optimal bound via sparse atomic norm optimization [2], [6]. In practice, the number K should be set as high as possible, so that there is at least one ω_k close enough to the true DOA's value.

The probability for indicators in (7) now becomes $\chi_{k,m} = \int_0^{2\pi} \mathcal{VM}_{\omega_k}(\beta_k)\varrho_{k,m} f(\omega_k, l_{k,m} = 1|\boldsymbol{X})d\omega_k$, in which the uniform distribution $\frac{1}{2\pi}$ in (7) is now replaced by a more locally concentrated density $\mathcal{VM}_{\omega_k}(\beta_k)$. Generally, we can also replace this Von-Mises prior by other locally concentrated priors, e.g. truncated uniform or Laplacian density in LASSO [4], [10].

For estimating the number of sources, let us define DOA's boolean vector $\boldsymbol{b}(\boldsymbol{L}) \triangleq [b_1, b_2, \ldots, b_K]^T \in \mathbb{I}^K$, with $b_k \triangleq b_k(\boldsymbol{l}_{k,:}) = \delta[\sum_{m=1}^M l_{k,m} \neq 0]$ and $\delta[\cdot]$ denoting Kronecker delta function. Using this technique, the DOA's ω_k is active only if the number of active tones at the kth DOA is not null, i.e. $b_k \neq 0$, with Bernoulli probability $f(b_k) = \text{Ber}(q_k)$ and $q_k = 1 - f(b_k = 0) = 1 - \prod_{m=1}^M p_{k,m}$. Since the total number B of active sources is a summation of Bernoulli variables $B \triangleq \sum_{k=1}^K b_k$, it follows a Poisson binomial distribution, i.e. $\mathbb{E}[B] = \sum_{k=1}^K q_k$ and $\text{var}[B] = \sum_{k=1}^K q_k(1 - q_k)$. Hence, this case is called sparse if we have $B \ll K$.

IV. SIMULATIONS

For simulation, we will consider the case of three sources, of which the frequency components are: γ_0 and $2\gamma_0$ for the first source at angle $\phi_1 = \frac{\pi}{8}$ (rad); $3\gamma_0$ for the second source at angle $\phi_2 = \frac{\pi}{4}$ (rad); γ_0 and $3\gamma_0$ for the third source at angle $\phi_3 = \frac{\pi}{3}$ (rad). By this way, the first and second source each has one overlapping frequency with the third source, while being not overlapped with each other. The frequency base $\gamma_0 \triangleq \frac{\pi}{4}$ (rad) is set at $\frac{N}{8}$ DFT bins. Each DFT bin is a quantization bin with the width $\frac{2\pi}{N}$ (rad/sample). Here we set $N = 2^{10}$ time samples. We assume the maximum number of sources and their tones are $K = M = 8$. The potential tone values are $\frac{m}{2}\gamma_0$, $m = 1, \ldots, 8$, corresponding to x-axis in Fig. 3. The number of sensors is $D = 20$. The varying signal-to-noise ratio is SNR $=\frac{1}{\sigma^2}$, where σ^2 is noise power. Since the uniform prior would cause DOA's ambiguity, we will apply the conjugate prior method in our simulations.

The superiority of multi-tone to single-tone method is illustrated in Fig. 3 and 4. Fig. 3 shows the posterior probabilities in (5) and mean estimations $\mathbb{E}(\omega_k)$ of DOA's angular frequency $\frac{\omega}{2\pi} \in [0, 1)$. The single-tone method, though being fairly accurate, can only capture one DOA for each tone. In contrast, the multi-tone method successfully detects all the sources, even with overlapping tones.

From this posterior distribution, Fig. 4 shows the MSE value $\frac{1}{ND}\mathbb{E}||\widehat{\boldsymbol{ALS}} - \boldsymbol{ALS}||^2$ of estimated means versus the level of noise, measured in SNR, in the sensor's data. At low SNR, the noise dominates the signal and, hence, both methods have the same low performance. When SNR increases, the multi-tone method is able to retrieve more tones of the sources and, hence, is significantly more accurate than the single-tone method.

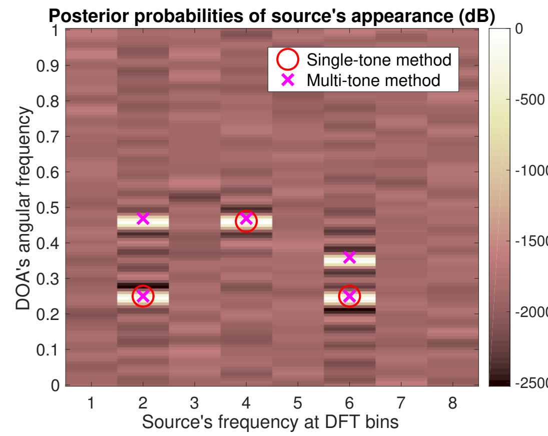

Fig. 3. Mean estimation of source's appearance and DOA via the single-tone and multi-tone method, SNR = -20 (dB)

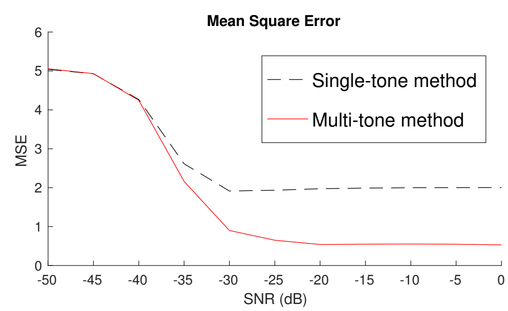

Fig. 4. Mean square error for sensor array data with 10^4 Monte Carlo runs.

V. CONCLUSION

In this paper, a full Bayesian multi-line spectra method was proposed for estimating the DOAs of multi-tone sources. In contrast to traditional single-line sprectra method, whose assumption is based on single-tone source, the multi-tone approach can detect all DOAs and the source's frequencies, even with overlapping frequencies. Owing to exact Bayesian inference, these estimations are optimal in term of Mean Square Error, as illustrated in our simulations. Also, this Bayesian method yields the optimal estimation for the number of active sources in the area, which is often unknown in practice. The computational complexity is only sub-linear, owing to applicability of fast Fourier transform.

VI. REFERENCES

[1] Z. Yang and L. Xie, "Enhancing sparsity and resolution via reweighted atomic norm minimization," *IEEE Transactions on Signal Processing*, vol. 64, no. 4, pp. 995–1006, Feb. 2016.

[2] B. N. Bhaskar, G. Tang, and B. Recht, "Atomic norm denoising with applications to line spectral estimation," *IEEE Transactions on Signal Processing*, vol. 61, no. 23, pp. 5987–5999, Dec. 2013.
[3] M.-A. Badiu, T. L. Hansen, and B. H. Fleury, "Variational Bayesian inference of line spectra," *IEEE Transactions on Signal Processing*, vol. 65, no. 9, pp. 2247–2261, May 2017.
[4] Z. Zhang, S. Wang, D. Liu, and M. I. Jordan, "EP-GIG priors and applications in Bayesian sparse learning," *J. Mach. Learn. Res.*, vol. 13, pp. 2031–2061, Jun. 2012.
[5] C. F. Mecklenbrauker, P. Gerstoft, A. Panahi, and M. Viberg, "Sequential Bayesian sparse signal reconstruction using array data," *IEEE Transactions on Signal Processing*, vol. 61, no. 24, pp. 6344–6354, Dec. 2013.
[6] G. Tang, B. N. Bhaskar, and B. Recht, "Near minimax line spectral estimation," *IEEE Transactions on Information Theory*, vol. 61, no. 1, pp. 499–512, Jan. 2015.
[7] J. M. Bernardo and A. F. M. Smith, *Bayesian Theory*. John Wiley & Sons Canada, 2006.
[8] V. H. Tran, "Variational Bayes inference in digital receivers," Ph.D. dissertation, Trinity College Dublin, 2014.
[9] V. Smidl and A. Quinn, *The Variational Bayes Method in Signal Processing*. Springer, 2006.
[10] K. E. Themelis, A. A. Rontogiannis, and K. D. Koutroumbas, "A Variational Bayes framework for sparse adaptive estimation," *IEEE Transactions on Signal Processing*, vol. 62, no. 18, pp. 4723–4736, Sep. 2014.
[11] H. J. Kappen and V. Gómez, "The variational garrote," *Machine Learning*, vol. 96, no. 3, pp. 269–294, Sep. 2014.
[12] K. V. Mardia and P. E. Jupp, *Directional Statistics*. John Wiley & Sons, 2008.
[13] 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 Transactions on Wireless Communications*, vol. 11, no. 10, pp. 3607–3617, Oct. 2012.
[14] A. Quinn, J.-P. Barbot, and P. Larzabal, "The Bayesian inference of phase," *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 4276–4279, May 2011.