Maximum-likelihood DOA estimation at low SNR in Laplace-like noise

Christoph F. Mecklenbräuker
Inst. of Telecommunications
TU Wien, 1040 Wien, Austria

Peter Gerstoft[1]
Scripps Institution of Oceanography
University of California San Diego

***Abstract*—We consider the estimation of the direction of arrivals (DOAs) of plane waves hidden in additive, mutually independent, complex circularly symmetric noise at very low signal to noise ratio (SNR). The maximum-likelihood estimator (ML) for the DOAs of deterministic signals carried by plane waves hidden in noise with a Laplace-like distribution is derived. This leads to a DOA estimator based on the Least Absolute Deviation (LAD) criterion. We prove analytically that a weighted phase-only beamformer (which evaluates the scalar product between the steering vector and the complex signum function of the observed array data) is an approximation to a beamformer based on the Least Absolute Deviation (LAD) criterion. The root mean squared error of DOA estimators versus SNR is compared in a simulation study: the conventional beamformer (CBF), the weighted phase-only beamformer, and sparse Bayesian learning (SBL3). This shows show that the ML estimator and weighted phase-only beamformer are well performing DOA estimators at low SNR for additive homoscedastic and heteroscedastic Gaussian noise, as well as Laplace-like noise.**

I. Introduction

Laplace distributions for signal and noise models have been used for studying detection and estimation problems since a long time. They are often used as a model for impulsive noise which is prevalent in seismic, acoustic, and radio environments [1], [2], [3]. Multivariate Laplace probability models and their generalizations are discussed in Refs. [4], [5]. Alternatively, a heteroscedastic noise model is used in [6] which is a hierarchical probability model based on a conditional normal distribution given the noise variance parameter σ^2 and assuming a probability distribution for σ^2.

Previously, a Laplace-like probability density has been assumed for the complex *source* amplitudes of plane waves hidden in additive white Gaussian noise in [7], [8]. In this context, the Laplace-like probability density has been used to promote sparse solutions [9], [10]. Such Laplace-like probability density is similar to the multivariate Laplace probability density, but the two should not be confused.

Here, we assume a Laplace-like probability density for the additive noise and regard the unknown complex source amplitudes as deterministic unknown variables. Based on the maximum-likelihood principle, this leads to complex-valued least absolute deviation (LAD) regression. Further, we show analytically that a weighted phase-only beamformer which computes the scalar product between the steering vector for a given DOA and the complex signum function of the observed array data approximates the LAD regression at low SNR.

It is well known that LAD regression has desirable robustness properties [11]. For example, influence function for LAD regression is universally bounded which guarantees that outliers in the observation data can only have a limited influence on the estimated regression result.

Robust estimation and detection of signals hidden in additive non-Gaussian noise has been considered in e.g. [12], [13], [14]. Huber's saddlepoint method is used in [12], which allows the specification of the robust detector for nominally Laplace distributed noise and the analysis of its performance.

[1]Supported by the Office of Naval Research, Grant No N00014-18-1-2118

II. Single Source in Laplace-like Noise

The array data model for multiple measurement vectors $\boldsymbol{y}_\ell$ for a single plane wave arriving from direction of arrival (DOA) θ in additive independent identically distributed zero-mean complex-valued Laplace-like noise is

$$\boldsymbol{y}_\ell = \boldsymbol{a}(\theta) x_\ell + \boldsymbol{n}_\ell, \qquad (\ell = 1, \ldots, L) \tag{1}$$

where $\boldsymbol{a}(\theta) = [a_1(\theta), \ldots, a_N(\theta)]^T$ is the plane wave steering vector for a single wave with unknown complex source amplitude x_ℓ. The nth element is given by $a_n(\theta) = \mathrm{e}^{\mathrm{j}\frac{\omega d_n}{c}\sin\theta}$ (d_n is the distance to the reference element and c the phase speed). The complex random vector $\boldsymbol{n}_\ell$ has the Laplace-like probability density function (pdf)

$$p(\boldsymbol{n}_\ell) = \prod_{k=1}^{N} \left(\frac{\lambda}{\sqrt{2\pi}}\right)^2 \mathrm{e}^{-\lambda |n_{k\ell}|} = \frac{\lambda^{2N}}{(2\pi)^N} \mathrm{e}^{-\lambda \|\boldsymbol{n}_\ell\|_1}. \tag{2}$$

We assume the noise to be independent across all measurement vectors, $\ell = 1, \ldots, L$. Given the DOA θ and the (row-)vector of complex source amplitudes $\boldsymbol{x} = [x_1, \ldots, x_L]$, the probability density of the array data matrix $\boldsymbol{Y} = [\boldsymbol{y}_1, \ldots, \boldsymbol{y}_L]$ with typical element $y_{n\ell}$ is

$$p(\boldsymbol{Y}; \theta, \boldsymbol{x}) = \prod_{\ell=1}^{L} \prod_{n=1}^{N} \left(\frac{\lambda}{\sqrt{2\pi}}\right)^2 \mathrm{e}^{-\lambda |y_{n\ell} - a_n(\theta) x_\ell|} \tag{3}$$

$$= \prod_{\ell=1}^{L} \frac{\lambda^{2N}}{(2\pi)^N} \, \mathrm{e}^{-\lambda \|\boldsymbol{y}_\ell - \boldsymbol{a}(\theta)\, x_\ell\|_1}. \tag{4}$$

978-9-0827-9703-9/19/$31.00 ©2019 IEEE

The negative log-likelihood for the unknown parameters $\theta, \boldsymbol{x}, \lambda$ becomes

$$\begin{aligned}\mathcal{L}(\theta,\boldsymbol{x},\lambda) &= -\ln p(\boldsymbol{y};\theta,x)\\ &= \lambda\sum_{\ell=1}^{L}\sum_{n=1}^{N}\left(|y_{n\ell}-a_n(\theta)x_\ell| - N\ln\frac{\lambda^2}{2\pi}\right) \quad (5)\\ &= \lambda\sum_{\ell=1}^{L}\|\boldsymbol{y}_\ell-\boldsymbol{a}(\theta)x_\ell\|_1 - LN\ln\frac{\lambda^2}{2\pi}. \quad (6)\end{aligned}$$

Estimators based on minimizing the criterion (5) are known as Least Absolute Deviation (LAD) estimators [11]. The LAD estimator is the ML-estimator for DOA θ in Laplace-like noise.

$$\hat{\theta}_{\text{LAD}} = \underset{\boldsymbol{x}\in\mathbb{C}^L,\lambda>0,\theta\in\Theta}{\arg\min}\ \mathcal{L}(\theta,\boldsymbol{x},\lambda), \quad (7)$$

where Θ is the feasible set of DOAs.

The numerical solution to (7) is found as follows. The *complex conjugate derivative* (in the sense of the Wirtinger calculus [11, Sec. 2.1]) of L with respect to the complex source amplitude is the direction of largest change [11, Eq.(2.3)]

$$\begin{aligned}\frac{\partial\mathcal{L}}{\partial x_\ell^*} &= \lambda\sum_{n=1}^{N}\frac{\partial|y_n-a_n(\theta)x_\ell|}{\partial x_\ell^*}\\ &= -\frac{\lambda}{2}\sum_{n=1}^{N}a_n^*(\theta)\frac{y_n-a_n(\theta)x_\ell}{|y_n-a_n(\theta)x_\ell|}\\ &= -\frac{\lambda}{2}\boldsymbol{a}^H(\theta)\,\boldsymbol{W}_\ell(\theta,x_\ell)\,(\boldsymbol{y}-\boldsymbol{a}(\theta)x_\ell) \quad (8)\end{aligned}$$

where we introduced the diagonal weighting matrix

$$\boldsymbol{W}_\ell(\theta,x_\ell) = \text{diag}\left(\frac{1}{|y_{1\ell}-a_1(\theta)x_\ell|},\ldots,\frac{1}{|y_{N\ell}-a_N(\theta)x_\ell|}\right) \quad (9)$$

Equating (8) to zero shows that a stationary point in $\mathcal{L}$ satisfies

$$x_\ell = \hat{x}_{\ell,\text{LAD}}(\theta) = \frac{\boldsymbol{a}^H(\theta)\,\boldsymbol{W}_\ell(\theta,x_\ell)\,\boldsymbol{y}_\ell}{\boldsymbol{a}^H(\theta)\,\boldsymbol{W}_\ell(\theta,x_\ell)\,\boldsymbol{a}(\theta)}, \quad (10)$$

for $\ell = 1,\ldots,L$. This expression does not allow direct evaluation because the weighting matrix depends on the unknowns. In the following, we assume low signal to noise ratio (SNR),

$$|x_\ell| \ll \min_n(|a_n^* y_{n\ell}|) = \min_n(|y_{n\ell}|) \quad (11)$$

and expand $\boldsymbol{W}_\ell(\theta,\mathrm{d}x)$ for infinitesimal change $\mathrm{d}x$ from $x_\ell = 0$ and any θ as

$$\begin{aligned}\boldsymbol{W}_\ell(\theta,\mathrm{d}x) &= \boldsymbol{W}_{0\ell} + \left.\frac{\partial W_\ell}{\partial x_\ell}\right|_{x_\ell=0}\mathrm{d}x + \left.\frac{\partial W_\ell}{\partial x_\ell^*}\right|_{x_\ell^*=0}\mathrm{d}x^*\\ &= \boldsymbol{W}_{0\ell}+\frac{1}{2}\boldsymbol{y}_\ell^H\,\boldsymbol{W}_{0\ell}^3\,\boldsymbol{a}(\theta)\,\mathrm{d}x + \frac{1}{2}\boldsymbol{y}_\ell^T\,\boldsymbol{W}_{0\ell}^3\,\boldsymbol{a}(\theta)^*\,\mathrm{d}x^*\\ &= \boldsymbol{W}_{0\ell}+\text{Re}[\boldsymbol{y}_\ell^H\,\boldsymbol{W}_{0\ell}^3\,\boldsymbol{a}(\theta)\,\mathrm{d}x] \quad (12)\end{aligned}$$

where the complex derivatives with respect to x_ℓ and x_ℓ^* are again in the sense of the Wirtinger calculus and

$$\boldsymbol{W}_{0\ell} = \boldsymbol{W}_\ell(\theta,0) = \text{diag}\left(\frac{1}{|y_{1\ell}|},\ldots,\frac{1}{|y_{N\ell}|}\right) \quad (13)$$

A first approximation $\hat{x}_\ell^{(1)}$ to the ML estimate (10) is found using the first term in (12)

$$\hat{x}_\ell^{(1)} = \frac{\boldsymbol{a}^H(\theta)\,\boldsymbol{W}_{0\ell}\,\boldsymbol{y}_\ell}{\boldsymbol{a}^H(\theta)\,\boldsymbol{W}_{0\ell}\,\boldsymbol{a}(\theta)} = c_\ell\,\boldsymbol{a}^H(\theta)\,\boldsymbol{W}_{0\ell}\,\boldsymbol{y}_\ell = c_\ell\boldsymbol{a}^H(\theta)\,\tilde{\boldsymbol{y}}_\ell \quad (14)$$

$$c_\ell = \frac{1}{\boldsymbol{a}^H(\theta)\,\boldsymbol{W}_{0\ell}\,\boldsymbol{a}(\theta)} = \left(\sum_{n=1}^{N}\frac{1}{|y_{n\ell}|}\right)^{-1} \quad (15)$$

$$\tilde{\boldsymbol{y}}_\ell = \boldsymbol{W}_{0\ell}\,\boldsymbol{y}_\ell = \text{sign}(\boldsymbol{y}_\ell) = \left(\frac{y_{1\ell}}{|y_{1\ell}|},\ldots,\frac{y_{N\ell}}{|y_{N\ell}|}\right)^T, \quad (16)$$

where $\text{sign}(\cdot)$ is the complex signum function [11, Eq.(2.6)]. Note that $\boldsymbol{a}^H\tilde{\boldsymbol{y}}_\ell$ in (14) resembles a CBF, with the steering vector $\boldsymbol{a}(\theta)$ applied to complex signum function $\tilde{\boldsymbol{y}}_\ell$, cf. [6]. Thus $\boldsymbol{a}^H(\theta)\tilde{\boldsymbol{y}}_\ell$ becomes independent of the magnitude of the array data which indicates robustness against outliers in the array data. The scaling c_ℓ (15) is positive real and independent of DOA θ. The scaling does, however, depend on the array data magnitudes.

We refer to the averaged squared magnitude of (14) as the weighted phase-only beamformer. The averaging of the estimated source power is carried out across measurement vectors,

$$\begin{aligned}\tilde{B}_L(\theta) &= \frac{1}{L}\sum_{\ell=1}^{L}|\hat{x}_\ell^{(1)}|^2 = \frac{1}{L}\sum_{\ell=1}^{L}\left|c_\ell\boldsymbol{a}^H(\theta)\tilde{\boldsymbol{y}}_\ell\right|^2 \quad (17)\\ &= \boldsymbol{a}^H(\theta)\left(\frac{1}{L}\sum_{\ell=1}^{L}c_\ell^2\tilde{\boldsymbol{y}}_\ell\tilde{\boldsymbol{y}}_\ell^H\right)\boldsymbol{a}(\theta) \quad (18)\end{aligned}$$

We use the first approximation (14) to initialize an iteratively re-weighted least squares (IRWLS) algorithm for $i = 1,2,\ldots$ until convergence [11, Algorithm 3],

$$\hat{x}_\ell^{(i+1)} = \frac{\boldsymbol{a}^H(\theta)\,\boldsymbol{W}_\ell(\theta,\hat{x}_\ell^{(i)})\,\boldsymbol{y}_\ell}{\boldsymbol{a}^H(\theta)\,\boldsymbol{W}_\ell(\theta,\hat{x}_\ell^{(i)})\,\boldsymbol{a}(\theta)} \quad \underset{i\to\infty}{\longrightarrow} \quad \hat{x}_{\ell,\text{LAD}}(\theta)\ . \quad (19)$$

Substituting $\boldsymbol{x} = \hat{\boldsymbol{x}}_{\text{LAD}}(\theta)$ in (6) gives

$$\mathcal{L}(\theta,\hat{\boldsymbol{x}}_{\text{LAD}}(\theta),\lambda) = \lambda\sum_{\ell=1}^{L}\|\boldsymbol{y}_\ell-\boldsymbol{a}(\theta)\hat{x}_{\ell,\text{LAD}}(\theta)\|_1 - LN\ln\frac{\lambda^2}{2\pi}. \quad (20)$$

Here $\hat{x}_{\ell,\text{LAD}}(\theta)$ is the ℓth element of $\hat{\boldsymbol{x}}_{\text{LAD}}(\theta)$, the LAD estimate for the complex source amplitude for DOA θ obtained by IRWLS. Minimizing (20) with respect to λ gives

$$\hat{\lambda}_{\text{LAD}}(\theta) = \frac{2LN}{\sum_{\ell=1}^{L}\|\boldsymbol{y}_\ell-\boldsymbol{a}(\theta)\hat{x}_{\ell,\text{LAD}}(\theta)\|_1} \quad (21)$$

Substituting $\lambda = \hat{\lambda}_{\text{LAD}}(\theta)$ into (20) gives the concentrated likelihood function for DOA θ:

$$\mathcal{L}(\theta,\hat{\boldsymbol{x}}_{\text{LAD}}(\theta),\hat{\lambda}_{\text{LAD}}(\theta)) = 2LN\ln\sum_{\ell=1}^{L}\|\boldsymbol{y}_\ell-\boldsymbol{a}(\theta)\hat{x}_{\ell,\text{LAD}}(\theta)\|_1 + \text{const.} \quad (22)$$

The numerical solution to (7) is found by a global grid search over (22) for $\theta \in \Theta$.

III. Multiple Sources in Laplace-like Noise

The array data model for multiple measurement vectors (MMV) $\boldsymbol{y}_l$ with multiple sources in additive independent identically distributed (i.i.d.) Laplace-like noise is

$$\boldsymbol{Y} = \boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{X} + \boldsymbol{N}, \tag{23}$$

where $\boldsymbol{Y} = [\boldsymbol{y}_1, \ldots, \boldsymbol{y}_L] \in \mathbb{C}^{N\times L}$, $\boldsymbol{X} = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_L] \in \mathbb{C}^{K\times L}$, and similarly for $\boldsymbol{N}$. We assume that K plane waves are arriving in additive white Laplace-like noise and the antenna array has N elements. Here, $\boldsymbol{A}(\boldsymbol{\theta}) = (\boldsymbol{a}(\theta_1), \ldots, \boldsymbol{a}(\theta_K) \in \mathbb{C}^{N\times K}$ is the steering matrix which contains the steering vectors $\boldsymbol{a}(\theta_k)$ as columns. The DOAs of the plane waves are collected in the vector $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_K)^T$. The negative log-likelihood is

$$\begin{aligned}\mathcal{L}(\boldsymbol{\theta}, \boldsymbol{X}, \lambda) &= \lambda \sum_{\ell=1}^{L} \|\boldsymbol{y}_\ell - \boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{x}_\ell\|_1 - LN \ln \frac{\lambda^2}{2\pi}. \\ &= \lambda \|\boldsymbol{Y} - \boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{X}\|_1 - LN \ln \frac{\lambda^2}{2\pi}, \end{aligned} \tag{24}$$

where we have introduced the entry-wise l_1-norm for matrices, i.e., $\|\boldsymbol{Y}\|_1 = \sum_{n=1}^{N} \sum_{\ell=1}^{L} |y_{n\ell}|$. The LAD estimator (7) is formally defined by

$$\hat{\boldsymbol{\theta}}_{\mathrm{LAD}} = \underset{\boldsymbol{X}\in\mathbb{C}^{K\times L},\, \lambda>0,\, \boldsymbol{\theta}\in\boldsymbol{\Theta}}{\arg\min} \mathcal{L}(\boldsymbol{\theta}, \boldsymbol{X}, \lambda). \tag{25}$$

The feasible set $\boldsymbol{\Theta}$ is discrete and contains all combinations of DOAs. The numerical solution to (25) is found as follows. First, we assume that the DOA vector $\boldsymbol{\theta}$ is given and $\boldsymbol{A}(\boldsymbol{\theta})$ is fixed and known. The *complex conjugate derivative* (in the sense of the Wirtinger calculus) of L with respect to the complex source amplitude vector is the direction of largest change [11, Eq.(2.3)]

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{x}_\ell^*} = -\frac{\lambda}{2} \boldsymbol{A}^H(\theta)\, \boldsymbol{W}_\ell(\theta, \boldsymbol{x}_\ell)\, (\boldsymbol{y}_\ell - \boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{x}_\ell) \tag{26}$$

where we introduced the diagonal weighting matrix

$$\boldsymbol{W}_\ell(\boldsymbol{\theta}, \boldsymbol{x}_\ell) = \operatorname{diag}\left(\frac{1}{|\boldsymbol{e}_1^T(\boldsymbol{y}_\ell - \boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{x}_\ell)|}, \ldots, \frac{1}{|\boldsymbol{e}_N^T(\boldsymbol{y}_\ell - \boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{x}_\ell)|} \right). \tag{27}$$

Here $\boldsymbol{e}_n$ is the nth standard basis vector. Equating (26) to zero shows that a stationary point in L satisfies

$$\boldsymbol{A}^H(\boldsymbol{\theta})\boldsymbol{W}_\ell(\boldsymbol{\theta}, \boldsymbol{x}_\ell)\boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{x}_\ell = \boldsymbol{A}^H(\boldsymbol{\theta})\, \boldsymbol{W}_\ell(\boldsymbol{\theta}, \boldsymbol{x}_\ell)\, \boldsymbol{y}_\ell. \tag{28}$$

This expression does not allow direct evaluation since the weighting matrix depends on the unknowns. We now assume a low SNR and a first approximation $\hat{\boldsymbol{x}}_\ell^{(1)}$ to the ML estimate (28) is found using the first term in (12)

$$\begin{aligned}\hat{\boldsymbol{x}}_\ell^{(1)} &= \left(\boldsymbol{A}^H(\boldsymbol{\theta})\boldsymbol{W}_{0\ell}\boldsymbol{A}(\boldsymbol{\theta})\right)^{-1} \boldsymbol{A}^H(\boldsymbol{\theta})\boldsymbol{W}_{0\ell}\, \boldsymbol{y}_\ell \\ &= \left(\boldsymbol{A}^H(\boldsymbol{\theta})\boldsymbol{W}_{0\ell}\boldsymbol{A}(\boldsymbol{\theta})\right)^{-1} \boldsymbol{A}^H(\boldsymbol{\theta})\, \tilde{\boldsymbol{y}}_\ell \end{aligned} \tag{29}$$

Then the LAD estimate $\hat{\boldsymbol{X}}_{\mathrm{LAD}}(\boldsymbol{\theta})$ is computed column-by-column ($\ell = 1, \ldots, L$) by applying the IRWLS algorithm to each data measurement vector individually for $i = 1, 2, \ldots$ until convergence [11, Algorithm 3], cf. (19),

$$\begin{aligned}\hat{\boldsymbol{x}}_\ell^{(i+1)} &= \left[\boldsymbol{A}^H(\theta)\boldsymbol{W}_\ell(\boldsymbol{\theta}, \hat{\boldsymbol{x}}_\ell^{(i)})\boldsymbol{A}(\theta)\right]^{-1} \boldsymbol{A}^H(\boldsymbol{\theta})\, \boldsymbol{W}_\ell(\boldsymbol{\theta}, \hat{\boldsymbol{x}}_\ell^{(i)})\, \boldsymbol{y}_\ell \\ &\underset{i\to\infty}{\longrightarrow} \quad \hat{\boldsymbol{x}}_{\ell,\mathrm{LAD}}(\boldsymbol{\theta}) \,, \end{aligned} \tag{30}$$

$$\hat{\boldsymbol{X}}_{\mathrm{LAD}}(\boldsymbol{\theta}) = (\hat{\boldsymbol{x}}_{1,\mathrm{LAD}}(\boldsymbol{\theta}), \ldots, \hat{\boldsymbol{x}}_{L,\mathrm{LAD}}(\boldsymbol{\theta})). \tag{31}$$

After substituting $\boldsymbol{X} = \hat{\boldsymbol{X}}_{\mathrm{LAD}}(\boldsymbol{\theta})$ in (24), we get

$$\mathcal{L}(\boldsymbol{\theta}, \hat{\boldsymbol{X}}_{\mathrm{LAD}}(\boldsymbol{\theta}), \lambda) = \lambda \left\|\boldsymbol{Y} - \boldsymbol{A}(\boldsymbol{\theta})\hat{\boldsymbol{X}}_{\mathrm{LAD}}(\boldsymbol{\theta})\right\|_1 - LN \ln \frac{\lambda^2}{2\pi}. \tag{32}$$

Minimizing (32) with respect to λ gives the ML estimate

$$\hat{\lambda}_{\mathrm{LAD}}(\boldsymbol{\theta}) = 2LN \Big/ \left\|\boldsymbol{Y} - \boldsymbol{A}(\boldsymbol{\theta})\hat{\boldsymbol{X}}_{\mathrm{LAD}}(\boldsymbol{\theta})\right\|_1 \tag{33}$$

Substituting $\lambda = \hat{\lambda}_{\mathrm{LAD}}(\boldsymbol{\theta})$ into (32) gives the concentrated likelihood function for the DOA vector $\boldsymbol{\theta}$,

$$\begin{aligned}&\mathcal{L}(\boldsymbol{\theta}, \hat{\boldsymbol{X}}_{\mathrm{LAD}}(\theta), \hat{\lambda}_{\mathrm{LAD}}(\boldsymbol{\theta})) \\ &\qquad = 2LN \ln \left\|\boldsymbol{Y} - \boldsymbol{A}(\boldsymbol{\theta})\hat{\boldsymbol{X}}_{\mathrm{LAD}}(\boldsymbol{\theta})\right\|_1 + \text{const.} \end{aligned} \tag{34}$$

The numerical solution to (25) is found by a global grid search over (34) for all combinations of DOAs, $\boldsymbol{\theta} \in \boldsymbol{\Theta}$.

IV. Simulations

Performance is assessed by numerical simulations using synthetic data. For the results in Figs. 1 and 2, a single plane wave with DOA $-45°$ with additive noise is observed with a uniform linear antenna array with $N = 20$ elements and spacing $\lambda/2$. The feasible set Θ in (7) is $\{0°, 1°, \ldots, 180°\}$. Three different types of zero-mean circularly symmetric complex-valued noise $\boldsymbol{N}$ in (23) are simulated.

The first type is i.i.d. Laplace-like noise with probability density (2). The second type is i.i.d. Gaussian noise with constant parameter σ. The third type is heteroscedastic Gaussian noise, cf. [6]. The heteroscedastic noise matrix $\boldsymbol{N}$ in the model (23) is conditionally Gaussian given the parameters $\sigma_{n\ell}$ for $n = 1, \ldots, N$ and $\ell = 1, \ldots, L$ and conditionally independent across sensors and measurement vectors,

$$p(\boldsymbol{N}|\boldsymbol{\Sigma}) = \prod_{\ell=1}^{L} \prod_{k=1}^{N} \frac{1}{(2\pi\sigma_{k\ell}^2)} \mathrm{e}^{-|n_{k\ell}|^2/\sigma_{n\ell}^2}, \tag{35}$$

The parameter matrix $\boldsymbol{\Sigma} = (\sigma_{n\ell})$ is an i.i.d. random matrix,

$$\sigma_{n\ell} = \sigma\, s_{n\ell} \text{ with } s_{n\ell} = \frac{t_{n\ell}}{\sqrt{\frac{1}{NL}\sum_n \sum_\ell t_{n\ell}^2}} \tag{36}$$

$$t_{n\ell} = 10^{U_{n\ell}}, \text{ where } U_{n\ell} \text{ is uniformly on } [-1, 1]. \tag{37}$$

The deterministic parameter σ is used to define the SNR in the plots. Fig. 1 shows the performance in terms of the root mean squared error (RMSE) of several DOA estimators for a single plane wave hidden in additive noise. These RMSE results are obtained for the DOA estimators based a single measurement

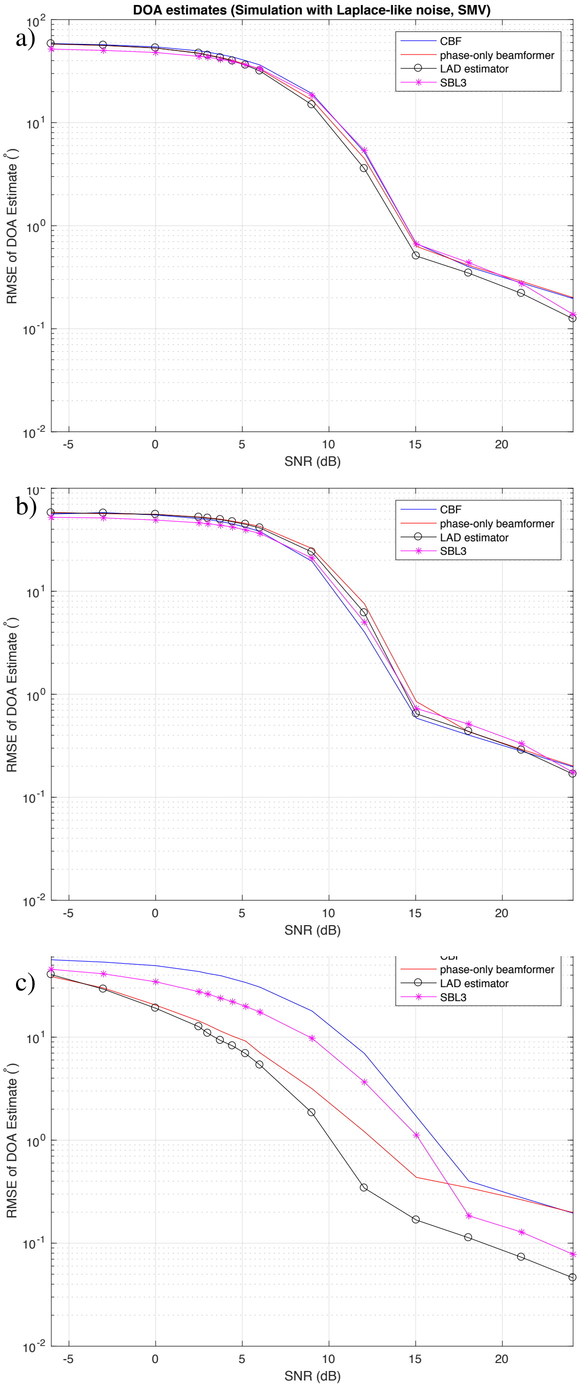

Fig. 1. RMSE of DOA estimate for a single DOA $-45°$ (20 element uniform linear array, single measurement vector, $L = 1$) with additive noise, (a) Laplace-like, (b) Gaussian, (c) heteroscedastic Gaussian.

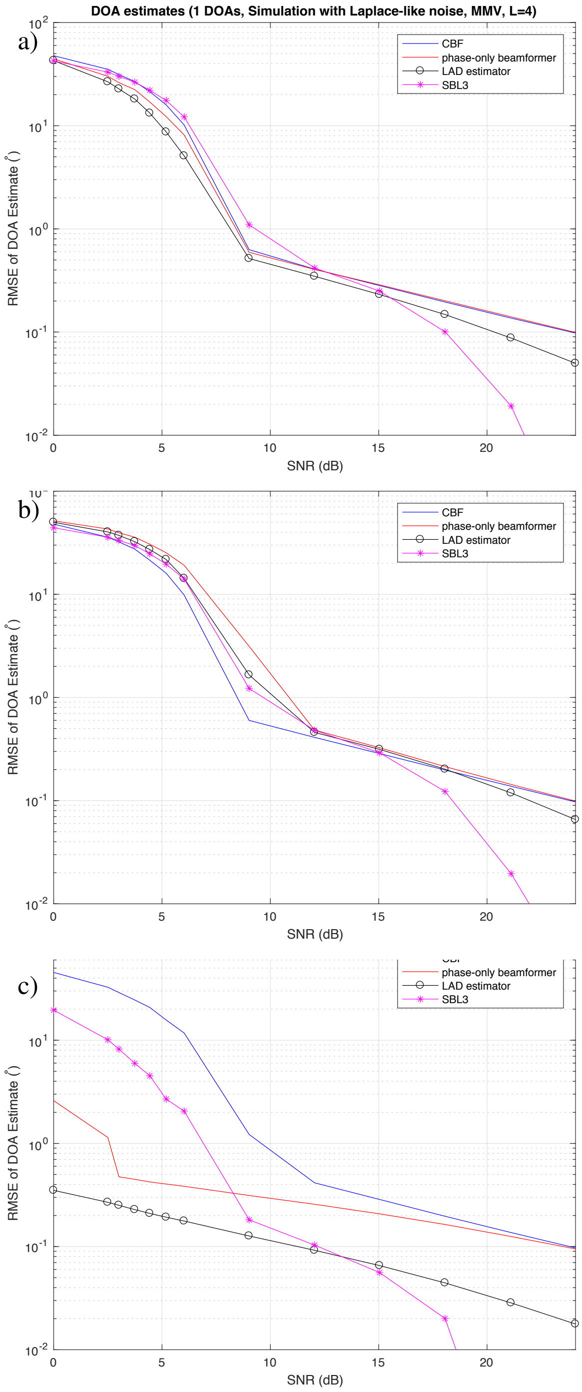

Fig. 2. RMSE of DOA estimate for a single DOA $-45°$ (20 element uniform linear array, multiple measurement vector, $L = 4$) with additive noise, (a) Laplace-like, (b) Gaussian, (c) heteroscedastic Gaussian.

vector, $L = 1$, as derived in Sec. II. Blue curve: CBF, red curve: DOA estimator which maximizes the weighted phase-only beamformer (18), black curve with '∘'-markers: LAD estimator (7), magenta curve with '*'-markers: sparse bayesian learning SBL3 [6]. The results for Laplace-like and Gaussian noise in Figs. 1ab are all very close to eachother they all reach the asymptotic regime above the threshold SNR≈ 15 dB. For heteroscedastic noise, however, the difference in estimator performance is clearly seen in Fig. 1c and the LAD-based DOA estimator has the lowest RMSE over the whole SNR range. The weighted phase-only beamformer approaches the LAD-based DOA estimator for low SNR.

Fig. 2 shows RMSE results for several DOA estimators using $L = 4$ measurement vectors as derived in Sec. II. For the CBF, the RMSE of the DOA estimate compared to single measurement vector shows a shift in SNR of 6 dB compared to Fig. 1. The corresponding shifts in SNR for the weighted phase-only beamformer (18) and LAD estimator (7) are either similar or higher than 6 dB. Figs. 1c and 2c also reveal that CBF and SBL3 suffer from heteroscedastic noise at low SNR in contrast to the weighted phase-only beamformer (18) and LAD-based DOA estimator (22). It is also seen in Fig. 2 that SBL3 significantly outperforms the other DOA estimators at high SNR. This behavior is due to the fact that the true source location is on the search grid.

Finally, we simulate three plane waves with DOAs $\boldsymbol{\theta} = [-2°, 3°, 75°]$. The minimization of (34) over a discrete grid with $1°$ spacing requires $\binom{181}{3} = 971970$ evaluations. The minimization is approximated by reducing $\boldsymbol{\Theta}$ to 50000 DOA combinations. The search is carried out over the subset of combinations with the lowest 50000 objective function values without noise. Fig. 3 shows the approximate RMSE of DOA results for the LAD-based DOA estimator (34) using $L = 4$ measurement vectors. The noise distribution is more heavy tailed in the heteroscedastic case, than Laplace-like or Gauss. Fig. 3 confirms that.

V. Conclusion

The weighted phase-only beamformer is a low-cost approximation to a beamformer based on the LAD criterion. The weighted phase-only beamformer performs well in terms of RMSE of its DOA estimate for a single DOA at low SNR for all three investigated noise types. Finally, we compare the RMSE of several different DOA estimators versus SNR in simulations: the CBF, weighted phase-only beamformer, LAD-based estimator, and sparse Bayesian learning (SBL3). The simulations indicate that the LAD-based DOA estimator and weighted phase-only beamformer are less sensitive to heteroscedastic noise than SBL3.

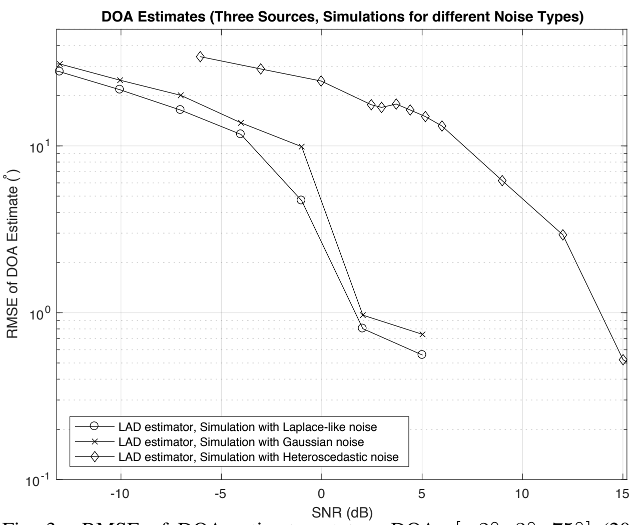

Fig. 3. RMSE of DOA estimates at true DOAs $[-2°, 3°, 75°]$ (20 element uniform linear antenna array, multiple measurement vector, $L = 4$) hidden in additive complex circularly symmetric (a) Laplace-like noise, (b) homoscedastic Gaussian noise, (c) heteroscedastic Gaussian noise.

References

[1] K. L. Blackard, T. S. Rappaport, and C. W. Bostian. Measurements and models of radio frequency noise for indoor wireless communications. *IEEE J. Sel. Areas Commun.*, 11(7):991–1001, 1993.

[2] D. Middleton and A. D. Spaulding. Elements of weak signal detection in non-Gaussian noise. *Advances in Statistical Signal Processing, Signal Detection*, 2:137–215, 1993.

[3] S. Jiang and N. C. Beaulieu. Precise ber computation for binary data detection in bandlimited white laplace noise. *IEEE Trans. Commun.*, 59(6):1570–1579, Jun. 2011.

[4] S. Kotz, T. J. Kozubowski, and K. Podgorski. *The Laplace distribution and generalizations. A revisit with applications to communications, economics, engineering, and finance*. Birkhäuser, 2001.

[5] T. Eltoft, Taesu Kim, and Te-Won Lee. On the multivariate Laplace distribution. *IEEE Sig. Proc. Lett.*, 13(5):300–303, May 2006.

[6] Peter Gerstoft, Santosh Nannuru, Christoph F. Mecklenbräuker, and Geert Leus. DOA estimation in heteroscedastic noise. *Signal Processing*, 161:63–73, 2019.

[7] C. F. Mecklenbräuker, P. Gerstoft, A. Panahi, and M. Viberg. Sequential Bayesian sparse signal reconstruction using array data. *IEEE Trans. Signal Process.*, **61**(24):6344–6354, 2013.

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

[9] Daniele Angelosante, Juan Andrés Bazerque, and Georgios B Giannakis. Online adaptive estimation of sparse signals: Where rls meets the ℓ_1-norm. *IEEE Transactions on Signal Processing*, 58(7):3436–3447, 2010.

[10] S. D. Babacan, R. Molina, and A. K. Katsaggelos. Bayesian compressive sensing using laplace priors. *IEEE Trans. Image Process.*, **19**(1):53–63, 2010.

[11] A. M. Zoubir, V. Koivunen, E. Ollila, and M. Muma. *Robust Statistics for Signal Processing*. Cambridge University Press, 2018.

[12] M. W. Thompson, D. R. Halverson, and G. L. Wise. Robust detection in nominally laplace noise. *IEEE Trans. Commun*, 42(2/3/4):1651–1660, Feb./Mar./Apr. 1994.

[13] Gencheng Guo, Mrinal Mandal, and Yindi Jing. A robust detector of known signal in non-Gaussian noise using threshold systems. *Signal Processing*, 92(11):2676–2688, 2012.

[14] E. Ollila. Multichannel sparse recovery of complex-valued signals using Huber's criterion. In 3rd *Int. Workshop on Compressed Sensing Theory and Appl. to Radar, Sonar, and Remote Sensing*, Pisa, Italy, June 2015.