This paper is a postprint of a paper submitted to and accepted for publication in **IET Radar Sonar and Navigation** and is subject to Institution of Engineering and Technology Copyright. The copy of record is available at the IET Digital Library

Postprint of: Meller M., Stawiarski K., Soft-decision schemes for radar estimation of elevation at low grazing angles, IET Radar Sonar and Navigation., Vol. 13, iss. 11 (2019), pp.1934-1942, DOI: 10.1049/iet-rsn.2019.0042

Soft-decision schemes for radar estimation of elevation at low grazing angles

Michał Meller, Kamil Stawiarski

***Abstract*—In modern radars, the problem of estimating elevation angle at low grazing angles is typically solved using superresolution techniques. These techniques often require one to provide an estimate of the number of waveforms impinging the array, which one can accomplish using model selection techniques. In this paper, we investigate the performance of an alternative approach, based on the Bayesian-like model averaging. The Bayesian approach exploits the fact that the parameters of the model related to multipath signals are nuisance ones, which allows one to avoid the estimation of the number of waveforms and improves estimation performance. The method is introduced for the classical conditional maximum likelihood estimator and extended to its, recently proposed, robustified version. We find, however, that the robustified estimator includes its own soft-decision mechanism and benefits from the averaging only for low levels of model uncertainty.**

I. Introduction

Estimation of elevation at low grazing angles is one of long-standing challenges in radar signal processing. What makes this problem difficult is the phenomenon of multipath propagation Barton [1974]. The term multipath propagation refers to the situation when the waveform experiences one or more reflections which result in multiple spurious wavefronts arriving at the receiver, along with the direct one.

In radar, the predominant source of the multipath is the reflection of the target's backscatter from the earth surface. At medium and high grazing angles, one can reject the spurious wavefront(s) using the spatial filtering (that is, beamforming), which makes their influence on the estimation accuracy negligible. However, as the target approaches the horizon, the spatial proximity of the desired (direct) signal and the unwanted one (multipath) reduces the effectiveness of the spatial filtering and the difficulty of the estimation problem increases. Eventually, the direct signal and the multipath both enter the array mainlobe. This situation, if not recognized in the design of the estimator, will result in a substantial estimation bias.

Multiple techniques were proposed to address this problem. Early solutions included the off-boresight tracking Barton [1974], the complex-valued monopulse method Sherman [1971], asymmetric beampatterns Barton [1974] or nulling out the reflections White [1974]. Modern approaches typically employ the so-called superresolution spectrum estimation techniques Bonacci et al. [2015], Nickel [2013], van Trees et al. [2002], Vincent et al. [2014].

The authors are with the Department of Automatic Control, Gdańsk University of Technology, Faculty of Electronics, Telecommunications and Computer Science, Narutowicza 11/12, 80-233 Gdańsk, Poland. Kamil Stawiarski also works for PIT-RADWAR S.A., ul. Poligonowa 30, 04-051 Warsaw, Poland. e-mails: michal.meller@eti.pg.edu.pl, kamil.stawiarski@pitradwar.com

Speaking succinctly, one can divide the superresolution methods into two groups. Nonparametric methods usually attempt to work out an estimate of the spatial spectrum, whose peaks are interpreted as DoAs of the signal sources. Nonparametric methods include, among others, the Capon estimator Capon [1969], MUSIC, and ESPRIT methods Roy and Kailath [1989], Schimidt [1986] or the minimum norm approach Reddy et al. [2015]. Note that nonparametric methods are often not suited to the specifics of DoA in radar, because of their requirements on the minimum number of observations, source independence or array properties.

In the case of the parametric approach, the direction of the target is found by fitting a parametric model to available observations. Parametric methods are usually embedded in the maximum likelihood (ML) framework Jaffer [1988], Nickel [2013] and, to preserve much-desired model parsimony, employ the flat-earth, specular reflection assumption Meeks [1982].

The advantages of the parametric approach that are important to the application in radar include, among others, modest computational complexity, ability to work with a single observation (i.e., the monopulse capability) and satisfactory performance, in most cases. However, the methods belonging to this class can be sensitive to the modeling uncertainty Nickel [2013] – an unknown mismatch between the assumed and the actual steering vectors. Modeling uncertainty can be caused by multiple factors, such as poor array calibration, model simplifications, or the presence of diffuse reflections. When the mismatch reaches a certain level, the accuracy of the maximum likelihood estimator breaks down, which means that gross estimation errors start occur increasingly often. To improve the behavior of the maximum likelihood estimator under modeling errors, we recently proposed its robustified variant Meller and Stawiarski, which employs the minimax principle. Compared to the conventional approach, the application of the modified estimator reduces the number of occurrences of estimation outliers considerably.

Another factor that can degrade the performance of the parametric approach is the wrong assessment of the number of sources. The consequences of using a model that assumes a wrong number of sources depend on the kind of mistake one made when choosing the model structure. If the assumed number of sources is too small, the model lacks the capacity to accommodate all waveforms arriving at the array. This situation, which is often referred to as underfitting, usually results in biased estimates of model parameters, including the estimates of source DoAs, in particular. When the assumed number of sources is too large, the estimator is unbiased, but two other adverse effects take place. First, the variance of DoA estimates increases. Second, the interpretation of origins of the superfluous "sources" becomes difficult Nickel [2013], Wirth [2013].

In this paper, we study the solutions that allows one to mitigate this difficulty. First, we exploit the fact that, in the particular situation of estimating the elevation, one can treat a large subset of model parameters as nuisance parameters. This observation allows us to develop a Bayesian-like collaborative approach based on the model averaging technique Longford [2012], Wit et al. [2012]. We develop the proposed approach for the deterministic (conditional) maximum likelihood estimator. We also study a, recently proposed, robustified maximum-likelihood like estimator Meller and Stawiarski, and investigate the possibility of improving its performance using the model averaging approach. We find that, however, that the robustified estimator includes another soft-decision mechanism, which reduces the benefits of the averaging.

The organization of the paper is the following one: In Section 2, we introduce the collaborative approach in the context of the maximum likelihood method. In Section 3, we extend the method to fit the robustified estimator and analyze the resultant solution. Section 4 presents results of computer simulations based on an extended model that included diffuse scattering. Section 5 demonstrates the application of all algorithms to data collected from a real-world radar system. Section 6 concludes the paper.

II. Collaborative approach – maximum likelihood estimator

A. Deterministic maximum likelihood estimator revisited

Denote by $\boldsymbol{y}_n$, $n = 1, 2, \ldots, N$, the available snapshots of M-variate data, assumed to be generated from the model

$$\boldsymbol{y}_n = \boldsymbol{\Psi}_{n|k}(\boldsymbol{\Phi}_k)\boldsymbol{\theta}_{n|k} + \boldsymbol{v}_{n|k} \ , \tag{1}$$

where $k \geq 1$ denotes the number of sources,

$$\boldsymbol{\Phi}_k = \begin{bmatrix} \varphi_{1|k} & \varphi_{2|k} & \ldots & \varphi_{k|k} \end{bmatrix}$$

is the vector of the source DoAs, and

$$\boldsymbol{\Psi}_{n|k}(\boldsymbol{\Phi}_k) = \begin{bmatrix} \boldsymbol{a}_n(\varphi_{1|k}) & \boldsymbol{a}_n(\varphi_{2|k}) & \ldots & \boldsymbol{a}_n(\varphi_{k|k}) \end{bmatrix}$$

is the matrix of the source steering vectors at the n-th snapshot. Similarly,

$$\boldsymbol{\theta}_{n|k} = \begin{bmatrix} A_{1,n|k} & A_{2,n|k} & \ldots & A_{k,n|k} \end{bmatrix}^{\mathrm{T}}$$

is the vector of complex-valued source amplitudes at the n-th snapshot. The sequence $\{\boldsymbol{v}_{n|k}\}$ is assumed to form a zero mean i.i.d. multivariate circular complex Gaussian white noise with unknown variance σ_k^2, $\boldsymbol{v}_{n|k} \sim \mathcal{CN}(0, \sigma_k^2 \boldsymbol{I})$. Moreover, we will treat all unknown model parameters, i.e., the angles, the amplitudes, and the noise variance, as deterministic quantities.

Observe that model (1) assumes that the source steering vectors and amplitudes are different at each snapshot. This feature, which differs from the typical assumption of fixed steering vectors and source amplitudes, allows us to increase the model's flexibility. The situations where a model like (1) is advantageous include, among others, the cases of varying the radar's frequency or the length of its coherent processing interval. The drawback of (1) is that the number of model parameters increases.

Under (1), the likelihood function of the data reads van Trees et al. [2013]

$$L_k(\boldsymbol{\Phi}_k, \boldsymbol{\theta}_{1|k}, \ldots, \boldsymbol{\theta}_{n|k}, \sigma_k^2) = \left(\frac{1}{\pi\sigma_k^2}\right)^{NM} \prod_{n=1}^{N} \exp\left(-\frac{\|\boldsymbol{y}_n - \boldsymbol{\Psi}_{n|k}(\boldsymbol{\Phi}_k)\boldsymbol{\theta}_{n|k}\|^2}{\sigma_k^2}\right) \ . \tag{2}$$

It is well known that, for any values of the angles and the variance, one can obtain the optimal, in the sense of maximizing the likelihood function, estimates of the source amplitudes using the following formula van Trees et al. [2013]

$$\hat{\boldsymbol{\theta}}_{n|k}(\boldsymbol{\Phi}_k, \sigma_k^2) = \hat{\boldsymbol{\theta}}_{n|k}(\boldsymbol{\Phi}_k) = \boldsymbol{P}_{n|k}(\boldsymbol{\Phi}_k)\boldsymbol{y}_n \ , \tag{3}$$

where

$$\boldsymbol{P}_{n|k}(\boldsymbol{\Phi}_k) = \boldsymbol{\Psi}_{n|k}(\boldsymbol{\Phi}_k)\left[\boldsymbol{\Psi}_{n|k}^{\mathrm{H}}(\boldsymbol{\Phi}_k)\boldsymbol{\Psi}_{n|k}(\boldsymbol{\Phi}_k)\right]^{-1}\boldsymbol{\Psi}_{n|k}^{\mathrm{H}}(\boldsymbol{\Phi}_k)$$

denotes the matrix projecting the data onto the signal subspace, i.e., the subspace spanned by the columns of $\boldsymbol{\Psi}_{n|k}(\boldsymbol{\Phi}_k)$.

Substituting (3) into (2) allows one to arrive at the, so-called, deterministic maximum likelihood estimator of the source DoAs

$$\hat{\boldsymbol{\Phi}}_k = \arg\min_{\boldsymbol{\Phi}_k} \sum_{n=1}^{N} \boldsymbol{y}_n^{\mathrm{H}}\boldsymbol{Q}_{n|k}(\boldsymbol{\Phi}_k)\boldsymbol{y}_n \ , \tag{4}$$

where

$$\boldsymbol{Q}_{n|k}(\boldsymbol{\Phi}_k) = \boldsymbol{I} - \boldsymbol{P}_{n|k}(\boldsymbol{\Phi}_k)$$

is the matrix projecting the data onto the noise subspace, and $\boldsymbol{I}$ denotes the $k \times k$ identity matrix. The corresponding maximum likelihood estimate of the noise variance σ_k^2 reads

$$\hat{\sigma}_k^2 = \frac{1}{NM}\sum_{n=1}^{N} \boldsymbol{y}_n^{\mathrm{H}}\boldsymbol{Q}_{n|k}(\hat{\boldsymbol{\Phi}}_k)\boldsymbol{y}_n \ . \tag{5}$$

Let us succinctly summarize how one can fit the estimator (4)-(5) to aid solving the problem of estimating the elevation at low grazing angles. Under the simplified flat-earth specular multipath model, depicted in Fig. 1, the multipath can be regarded as originating from an apparent source located below the horizon Barton [1974]. It follows that one should fit the model that assumes $k = 2$ sources. In this case, one may arbitrarily assign the first source to the target and the second source to the multipath. Based on such labeling, one may restrict $\hat{\varphi}_{1|2}$ to the positive values, $\hat{\varphi}_{1|2} > 0$, and $\hat{\varphi}_{2|2}$ to the negative values, $\hat{\varphi}_{2|2} < 0$, which will make it easier to interpret the estimation results.

B. Competitive solution

A significant problem with the flat-earth specular multipath model lies in the fact that it is overly simplistic and does not account for a variety of scattering phenomena Barton [1974].

On rare occasions, there may be more than one specular reflections present, in which case the number of sources in the model should be increased, i.e., one should assume $k > 2$. Since, in this case, the variables $\hat{\varphi}_{2|k}, \ldots, \hat{\varphi}_{k|k}$ can

MOST WIEDZY Downloaded from mostwiedzy.pl

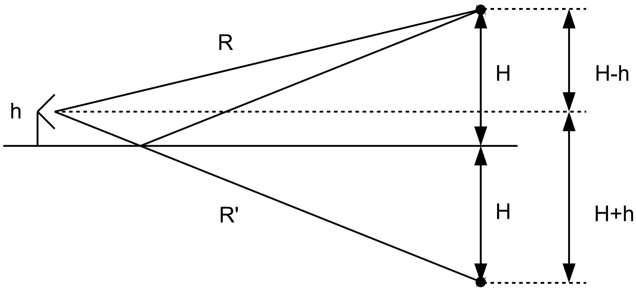

Figure 1. Flat-earth model of specular multipath.

be interpreted as corresponding to the multipath sources, one may constrain them to the range of negative numbers.

More often, however, the specular reflection might be weak or absent (e.g., due to strong attenuation by vegetation), or the specular reflection assumption might be inadequate (e.g., due to the dominance of the diffuse scattering). In any of these situations, using the parsimonious model with $k = 1$ might lead to better estimation accuracy than keeping the number of sources at $k = 2$ – as a general rule, the number of parameters in the model should be kept at a minimum level Claeskens and Hjort [2008].

Unfortunately, the "right" model structure is typically unknown a priori. The standard way to cope with this difficulty involves simultaneously running several estimators of the form (4)-(5), each assuming a model $\mathcal{M}_k$ with k sources, $k \in \{1, 2, \ldots, K\}$. Based on the outputs of all K estimators, one may attempt to determine the optimal choice of k using a suitably designed criterion. To this end, one can use several model selection approaches, which include the hypothesis testing Nickel [1998], Wirth [2013], methods developed using the random matrix theory Kritchman and Nadler [2009], Rissanen's minimum description length principle Rissanen [1978], or one of so-called information criteria Claeskens and Hjort [2008]. We will discuss the last approach in more details.

The most widely adopted information criteria are the Akaike's information statistic (also known as Akaike information criterion) Akaike [1974]

$$\mathrm{AIC}_k = -2\log L(\hat{\boldsymbol{\Phi}}_k, \hat{\sigma}_k^2) + 2n_\mathrm{p}(k) \tag{6}$$

and the Schwarz criterion (often referred to as the Bayesian information criterion) Schwarz [1978]

$$\mathrm{BIC}_k = -2\log L(\hat{\boldsymbol{\Phi}}_k, \hat{\sigma}_k^2) + \log(n_\mathrm{o})n_\mathrm{p}(k)\ . \tag{7}$$

In the above formulas, $l_k = \log L(\hat{\boldsymbol{\Phi}}_k, \hat{\boldsymbol{\theta}}_{1|k}, \ldots, \hat{\boldsymbol{\theta}}_{n|k}, \hat{\sigma}_k^2)$ denotes the log-likelihood of the maximum-likelihood estimate (4)-(5) obtained for the model $\mathcal{M}_k$, equal to

$$\log L(\hat{\boldsymbol{\Phi}}_k, \hat{\boldsymbol{\theta}}_{1|k}, \ldots, \hat{\boldsymbol{\theta}}_{n|k}, \hat{\sigma}_k^2) = -NM\left[\log \hat{\sigma}_k^2 + 1 + \log \pi\right]\ , \tag{8}$$

$n_\mathrm{p}(k)$ is the number of real-valued parameters of the model, and $n_\mathrm{o} = 2NM$ is the number of real valued observations. Note that one should not confuse the number of model parameters with the number of sources[1]. The parameters in the model $\mathcal{M}_k$ are the source DoAs $\phi_{1|k}, \phi_{2|k}, \ldots, \phi_{k|k}$, the amplitudes $A_{1,n|k}, A_{2,n|k}, \ldots, A_{k,n|k}$, $n = 1, 2, \ldots, N$, and the variance σ_k. Moreover, since all amplitudes are complex quantities, they should be treated as two real parameters, and counted twice.

[1]To avoid the ambiguity with the traditional notation, which counts the number of model parameters using the symbol k, we denote the number of model parameters as $n_\mathrm{p}(k)$.

Bank of ML estimators

For $k = 1, 2, \ldots, K$

$$\hat{\boldsymbol{\Phi}}_k = \arg\min_{\boldsymbol{\Phi}_k} \sum_{n=1}^{N} \boldsymbol{y}_n^\mathrm{H} \boldsymbol{Q}_{n|k}(\boldsymbol{\Phi}_k)\boldsymbol{y}_n$$

$$\hat{\sigma}_k^2 = \frac{1}{NM}\sum_{n=1}^{N} \boldsymbol{y}_n^\mathrm{H} \boldsymbol{Q}_{n|k}(\hat{\boldsymbol{\Phi}}_k)\boldsymbol{y}_n$$

$$l_k = -NM\left[\log \hat{\sigma}_k^2 + 1 + \log \pi\right]$$

Model selection

$$k_\mathrm{o} = \arg\min_{k=1,2,\ldots,K} \mathrm{IC}_k$$

where

$$\mathrm{IC}_k = \begin{cases} -2l_k + 2n_\mathrm{p}(k) & \text{for AIC} \\ -2l_k + \log(n_\mathrm{o})n_\mathrm{p}(k) & \text{for BIC} \\ -2l_k + 2n_\mathrm{p}(k) + 2\frac{n_\mathrm{p}(k)[n_\mathrm{p}(k)+1]}{n_\mathrm{o}-n_\mathrm{p}(k)-1} & \text{for AICc} \end{cases}$$

Final estimate of target DoA

$$\hat{\varphi}_{1|K} = \hat{\phi}_{1|k_\mathrm{o}}$$

Table I

SUMMARY OF COMPETITIVE MAXIMUM LIKELIHOOD ADAPTIVE FAMILY OF ADAPTIVE ESTIMATORS. DEPENDING ON VARIANT, DECISION STATISTIC IC_k SHOULD BE EITHER AIC, BIC, OR AICc.

Many authors also recommend the so-called corrected AIC criterion, which takes the form Claeskens and Hjort [2008], Hurvich et al. [1998]

$$\mathrm{AICc}_k = \mathrm{AIC}_k + 2\frac{n_\mathrm{p}(k)[n_\mathrm{p}(k)+1]}{n_\mathrm{o} - n_\mathrm{p}(k) - 1}\ , \tag{9}$$

i.e., it penalizes models complexity stronger than AIC, particularly for a small number of observations.

The resultant family of adaptive estimators, further referred to as the competitive ML approach, is summarized in Table I. Note that, since in each model all DoAs except the first one are interpreted as multipath, they are discarded irrespective of the choice of the model, and only the first DoA is included in the final estimate.

C. *Collaborative solution*

The competitive approach discussed in the preceding subsection may be criticized for neglecting two facts. First, that there is an unavoidable element of uncertainty in the model selection process. Model selection decision statistics, such as AIC, AICc, or BIC, are random variables, which means that erroneous choices are bound to occur from time to time. The presence of such mistakes inflates the mean squared error of the estimator Claeskens and Hjort [2008]. Second, in estimating the elevation at low grazing angles, there is really no need to establish the number of sources, because only the first source corresponds to the actual target. It follows that one can "get away" with a wrong choice of k, as long as the accuracy of the target DoA estimate does not suffer.

MOST WIEDZY Downloaded from mostwiedzy.pl

Bank of ML estimators
For $k = 1, 2, \ldots, K$

$$\hat{\boldsymbol{\Phi}}_k = \arg\min_{\boldsymbol{\Phi}_k} \sum_{n=1}^{N} \boldsymbol{y}_n^{\mathrm{H}} \boldsymbol{Q}_{n|k}(\boldsymbol{\Phi}_k)\boldsymbol{y}_n$$

$$\hat{\sigma}_k^2 = \frac{1}{NM} \sum_{n=1}^{N} \boldsymbol{y}_n^{\mathrm{H}} \boldsymbol{Q}_{n|k}(\hat{\boldsymbol{\Phi}}_k)\boldsymbol{y}_n$$

$$l_k = -NM\left[\log \hat{\sigma}_k^2 + 1 + \log \pi\right]$$

Final DoA estimate using averaging

$$\bar{\varphi}_{1|K} = \sum_{k=1}^{K} \mu_k \hat{\varphi}_{1|k}$$

where

$$\mu_k = \frac{\exp\left(-\frac{1}{2}\mathrm{IC}_k\right)}{\sum_{k=1}^{K} \exp\left(-\frac{1}{2}\mathrm{IC}_k\right)}$$

$$\mathrm{IC}_k = \begin{cases} -2l_k + 2n_{\mathrm{p}}(k) & \text{for} \quad \mathrm{AIC} \\ -2l_k + \log(n_{\mathrm{o}})n_{\mathrm{p}}(k) & \text{for} \quad \mathrm{BIC} \\ -2l_k + 2n_{\mathrm{p}}(k) + 2\frac{n_{\mathrm{p}}(k)[n_{\mathrm{p}}(k)+1]}{n_{\mathrm{o}}-n_{\mathrm{p}}(k)-1} & \text{for} \quad \mathrm{AICc} \end{cases}$$

Table II
SUMMARY OF PROPOSED COLLABORATIVE MAXIMUM LIKELIHOOD ADAPTIVE ESTIMATION APPROACH. DEPENDING ON VARIANT, IC_k REFERS TO EITHER AIC, BIC, OR AICc.

One may avoid making the decision about the number sources using the Bayesian framework. In this approach, the parameters are treated as random variables and assigned prior distributions. Consider one more time the bank of K estimators of the form (4)-(5), each delivering its own estimate of target DoA $\hat{\varphi}_{1|k}$ under the assumption that the model $\mathcal{M}_k$, $k = 1, 2, \ldots, K$ is correct. The Bayesian estimate of $\bar{\varphi}_{1|K}$ can be obtained from the following convex combination van Trees et al. [2013]

$$\bar{\varphi}_{1|K} = \sum_{k=1}^{K} \hat{\varphi}_{1|k}\mu_k \ , \tag{10}$$

where $\mu_k = p(\mathcal{M}_k|\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_N)$, $k = 1, 2, \ldots, K$, denotes the a posteriori probability of the model $\mathcal{M}_k$.

The difficulty in implementing Eq. (10) stems from the fact that computation of the a posteriori probabilities is computationally exhaustive. Akaike has shown that, under uniform priors, one can approximate μ_k using the following formula Akaike [1978]

$$\mu_k \propto \exp\left(-\frac{1}{2}\mathrm{AIC}_k\right) \ , \tag{11}$$

where AIC_k is defined in (6).

One may obtain similar approximations using AICc

$$\mu_k \propto \exp\left(-\frac{1}{2}\mathrm{AICc}_k\right) \tag{12}$$

or BIC

$$\mu_k \propto \exp\left(-\frac{1}{2}\mathrm{BIC}_k\right) \ . \tag{13}$$

The resultant three schemes, which avoid the estimation of k using the soft-decision approach, will be referred to as collaborative maximum likelihood estimators. They are summarized in Table II.

D. Computer simulations

While the collaborative approach can be, rather safely, expected to overperform the competitive one, the choice between AIC, AICc, and BIC is not an obvious one. First, both AIC and BIC are large sample approximations, so it is unclear how they will perform with small sample, which is typical situation in radar. Second, even though theses criteria seem to differ only in the term that penalizes the model complexity, they are, in fact, based on different principles and assumptions, and therefore exhibit different properties. BIC employs the assumption that the true model lies in the considered model set $\{\mathcal{M}_1, \mathcal{M}_2, \ldots, \mathcal{M}_K\}$, which can be regarded somewhat unrealistic. AIC, on the other hand, is free of such restriction – this statistic attempts to measure the proximity to an *unknown* truth Akaike [1974], Wit et al. [2012]. Unfortunately, this feature makes AIC an inconsistent estimator of k – even when the actual mechanism that generated the observations happens to be included in the model set, the probability of selecting $\mathcal{M}_k$ does not approach one as the number of observations increases to infinity. This behavior, however, is not necessarily a problem. While AIC is well known to overestimate k frequently, it is an efficient criterion, which means that it tends to minimize the mean square error – see Claeskens and Hjort [2008] for more details. Moreover, since in the collaborative approach we completely avoid selecting k, the inconsistency of AIC might become even less problematic.

One can investigate the properties of the collaborative averaging schemes using real-world datasets or carefully designed simulation experiments. Since generating synthetic datasets is considerably cheaper than collecting real-world ones, and it allows one to reach more general conclusions, the computer simulations are the preferred approach. Concordantly, we performed the several Monte Carlo experiments, where we simulated radar systems equipped with standard uniform linear arrays of lengths $M \in \{4, 8, 16\}$ and using snapshot sizes $N \in \{1, 2, 3\}$. For each pair $(N,\ M)$, we generated a dataset consisting of 5000 realizations of model (1) using the following procedure. For each data sample generated, the number of sources of was established first, by drawing randomly from the set $\{1, 2\}$, with equal probabilities of each choice. If $k = 1$ was drawn, which corresponds to the absence of the multipath, the source magnitudes were set to 10 for all snapshots, $|A_{1,n|k}| = 10$, $n = 1, 2, \ldots, N$. The source phases were generated as realizations of independent random variables, uniformly distributed in $[0, 2\pi]$. For $k = 2$, which corresponds to the presence of the multipath, the complex amplitudes of the direct signal were chosen in the same way as for $k = 1$. To ensure richness of the dataset, the complex amplitudes of the specular reflection were obtained as $A_{2,n|k} = \alpha e^{j\beta} A_{1,n|k}$, where α is a random variable, uniformly distributed in the interval [1, 10], and β is another random variable, uniformly distributed in the interval $[0, 2\pi]$. For the same reasons, the source angle φ_1 was drawn randomly from the interval 2° - 5°. The angle φ_2 was computed from the flat earth model, assuming antenna height $h = 4$ meters and target range $R = 5$ kilometers. In all cases, the variance of the measurement noise was equal to $\sigma_v^2 = 1$.

MOST WIEDZY Downloaded from mostwiedzy.pl

N	M	Nonadaptive $k=1$	Nonadaptive $k=2$	Competitive AIC	Competitive AICc	Competitive BIC	Collaborative AIC	Collaborative AICc	Collaborative BIC	Clairvoyant estimator
1	4	6.37	6.76	5.41	6.37	5.37	4.76	6.37	**4.72**	3.91
	8	2.780	0.464	0.386	0.338	0.360	0.358	**0.317**	0.338	0.290
	16	0.5892	0.0105	0.0095	0.0090	0.0088	0.0091	0.0088	**0.0087**	0.0084
2	4	6.71	4.64	3.55	4.56	3.14	3.20	4.25	**2.84**	2.51
	8	2.707	0.228	0.183	0.172	0.171	0.174	0.165	**0.164**	0.146
	16	0.6293	0.0051	0.0043	0.0042	0.0041	0.0042	0.0041	**0.0040**	0.0040
3	4	6.18	3.71	2.79	3.11	2.30	2.53	2.95	**2.13**	1.97
	8	2.830	0.164	0.132	0.129	0.131	0.127	**0.124**	0.127	0.112
	16	0.6219	0.0040	0.0034	0.0033	0.0032	0.0034	0.0033	**0.0032**	0.0032

Table III

COMPARISON OF MEAN SQUARED ESTIMATION ERRORS YIELDED BY NONADAPTIVE ESTIMATORS ASSUMING $k=1$ AND $k=2$, THREE COMPETITIVE MAXIMUM LIKELIHOOD SCHEMES BASED ON AIC, AIC$_c$, AND BIC CRITERIA, THREE CORRESPONDING COLLABORATIVE MAXIMUM LIKELIHOOD SCHEMES BASED ON SAME CRITERIA, AND CLAIRVOYANT ESTIMATOR, OBTAINED FOR THREE STANDARD UNIFORM ARRAYS, USING 5000 MONTE CARLO TRIALS FOR EACH CONFIGURATION. UNITS ARE DEGREES SQUARED.

We compared the accuracy of nonadaptive estimators employing the single-source model ($k = 1$) and the two-source model ($k = 2$), three adaptive estimators employing the competitive model selection approach, and three adaptive estimators based on the collaborative approach. Additionally, to establish baseline performance, we implemented a "clairvoyant" estimator, i.e., the estimator that has the unfair advantage of knowing the true value of k.

All estimators employing the two-source model took advantage of the fact that $\phi_1 \approx -\phi_2$ (c.f. Fig. 1), by forcing that

$$\hat{\varphi}_{2|2} = -\hat{\varphi}_{1|2} \ . \tag{14}$$

This substitution reduces the estimator's computational complexity greatly, because the underlying minimization of a function of two variables ($\varphi_{1|2}$, $\varphi_{2|2}$) is replaced with a minimization of a function of one variable. Additionally, this trick reduces the number of parameters in the model by one, which means that

$$n_{\mathrm{p}}(k) = 2Nk + 2$$
$$n_{\mathrm{o}} = 2MN \ ,$$

where the formula for $n_{\mathrm{p}}(k)$ stems from the fact that one is estimating Nk complex amplitudes ($A_{1,n|k}, A_{2,n|k}, \ldots, A_{k,n|k}$, $n = 1, 2, \ldots, N$), one angle ($\varphi_{1|k}$) and one noise variance parameter (σ_k^2).

The results, in the form of mean squared errors of the elevation estimates, are summarized in Table III. Not unexpectedly, the nonadaptive estimator that assumes $k = 1$ performs poorly for all values of N and M. Such high values of estimation errors are caused by the fact that half of the dataset contains the specular multipath, which cannot be appropriately captured by the model that assumes one source. Observe, however, that the performance of the nonadaptive estimator that employs the model with $k = 2$ sources, while significantly better, can be improved further using any of the adaptive approaches considered here. Moreover, the results obtained using the collaborative approach are uniformly better than those obtained using the competitive approach, regardless of the information criterion employed. It is also revealed that the beneficial effects of model averaging are more apparent for the smaller values of M, and that the best results are obtained using BIC and AICc criteria, except $M = 4$ where AICc works poorly. Finally, observe that the performance of the best collaborative estimators is rather impressively close (5%-10% loss) to the performance of the clairvoyant solution.

III. COLLABORATIVE APPROACH – ROBUSTIFIED MAXIMUM LIKELIHOOD ESTIMATOR

A. *Robustified maximum likelihood estimator*

We now switch our attention from the conventional maximum likelihood estimator, given by Eq. (4)-(5), to its robustified version, proposed in Meller and Stawiarski. The introduction of the robustified estimator was motivated by the observation that, when implemented in an actual radar system, the conventional solution exhibits an instability that results in occasional gross errors. The analysis of the phenomenon resulted in the hypothesis that this behavior is related to the modeling uncertainty, i.e., an unknown discrepancy between the assumed (nominal) steering vectors and the actual ones. Consequently, the robustified estimator includes the assumption that the steering vectors, i.e., the columns of $\boldsymbol{\Psi}_{n|k}(\boldsymbol{\Phi}_k)$ in (1), are subject to unknown deterministic perturbations with bounded norms.

The robustified estimator deals with this situation by employing a minimax-type extension of formula (4), which takes the form

$$\tilde{\boldsymbol{\Phi}}_k = \arg\min_{\boldsymbol{\Phi}_k} \sum_{n=1}^{N} J_{n|k}^2(\boldsymbol{\Phi}_k) \ , \tag{15}$$

where

$$J_{n|k}(\boldsymbol{\Phi}_k) = \min_{\boldsymbol{\theta}_{n|k}} \max_{\tilde{\boldsymbol{\Psi}}_{n|k}} \|\boldsymbol{y}_n - \tilde{\boldsymbol{\Psi}}_{n|k}\boldsymbol{\theta}_{n|k}\|$$
$$\tilde{\boldsymbol{\Psi}}_{n|k} = \left[\tilde{\boldsymbol{a}}_{1,n|k} \ \tilde{\boldsymbol{a}}_{2,n|k} \ \ldots \ \tilde{\boldsymbol{a}}_{k,n|k}\right] \tag{16}$$

and the "max" subproblem is subject to the constraints

$$\|\tilde{\boldsymbol{a}}_{j,n|k} - \boldsymbol{a}_n(\varphi_{j|k})\|^2 \le \epsilon_j^2$$
$$j = 1, 2, \ldots, K$$
$$n = 1, 2, \ldots, N \ , \tag{17}$$

where ϵ_j, $j = 1, 2, \ldots, K$, denote the perturbation bounds. Note that, under $\epsilon_j = 0 \ \forall j$, estimators (4) and (15) are equivalent.

The fact that one can assign the bounds of perturbations individually to each source is an important feature of the

MOST WIEDZY Downloaded from mostwiedzy.pl

proposed robustification approach. In the classical-style solution, one would lump all uncertainty into a single matrix perturbation with bounded norm (typically, the largest singular value) – see, e.g., Ghaoui and Lebret [1997]

$$\|\tilde{\mathbf{\Psi}}_{n|k}(\Phi_k) - \mathbf{\Psi}_{n|k}(\Phi_k)\|^2 \le \epsilon^2 \ .$$

Our approach is motivated by the fact that, for a sufficiently well-calibrated array, one may expect that the steering vectors that correspond to the direct signal, i.e., $\tilde{\boldsymbol{a}}_n(\varphi_{1|k})$, $n = 1, 2, \ldots, N$, exhibit a significantly smaller distortion than the steering vectors that correspond to the multipath. It is straightforward to express this a priori knowledge using (17), by setting ϵ_1 to a much smaller value than remaining perturbation bounds.

The minimax problem appearing in Eq. (16) is convex. Following Meller and Stawiarski, it can be solved efficiently using the following recursion

$$\begin{aligned} &\text{For } i = 0, 1, \ldots \\ &\tilde{\boldsymbol{\theta}}_{n,i+1|k} = \left[\mathbf{\Psi}^{\mathrm{H}}_{n|k}(\boldsymbol{\Phi}_k)\tilde{\mathbf{\Psi}}_{n|k}(\boldsymbol{\Phi}_k) + \mathbf{\Lambda}_{i,n|k}\right]^{-1} \mathbf{\Psi}^{\mathrm{H}}_{n|k}(\boldsymbol{\Phi}_k)\boldsymbol{y}_n \\ &v_{n,i+1|k} = \|\boldsymbol{y}_n - \mathbf{\Psi}_{n|k}(\boldsymbol{\Phi}_k)\tilde{\boldsymbol{\theta}}_{n,i|k}\| \\ &\mathbf{\Lambda}_{i+1,n|k} = \mathrm{diag}\left(\epsilon_1 \frac{v_{n,i+1|k}}{|\tilde{A}_{1,n,i+1|k}|}, \ldots, \epsilon_k \frac{v_{n,i+1|k}}{|\tilde{A}_{k,n,i+1|k}|}\right) \ , \end{aligned} \tag{18}$$

where

$$\tilde{\boldsymbol{\theta}}_{n,i+1|k} = \left[\ \tilde{A}_{1,n,i+1|k} \quad \tilde{A}_{2,n,i+1|k} \quad \ldots \quad \tilde{A}_{k,n,i+1|k}\ \right]^{\mathrm{T}}$$

denotes the estimates of the source amplitudes for the n-th snapshot in the $i+1$-th iteration of the algorithm. One may obtain the estimates of $J^2_{n|k}(\boldsymbol{\Phi}_k)$ corresponding to each iteration using

$$J^2_{n,i+1|k}(\boldsymbol{\Phi}_k) = \left(v_{n,i+1|k} + \sum_{j=1}^{K} \epsilon_j \left|\tilde{A}_{j,n,I_{\max}|k}\right|\right)^2 \ . \tag{19}$$

The recursion can be initialized using $v_{n,0|k} = 0$, $\mathbf{\Lambda}_{0,n|k} = \mathbf{0}$, $\forall n$, and should be iterated until its convergence. In most cases, three to five iterations are sufficient to reach the optimum with practically required accuracy. Such a low number of iterations means that the resultant increase of the computational cost over the conventional estimator (4)-(5) is modest (2-3 times), provided that one implements the recursion (18) with appropriate attention to optimization opportunities Meller and Stawiarski.

Table IV presents the detailed summary of the robustified estimator.

B. *Collaborative solution*

In light of the advantages of the collaborative approach demonstrated in Section 2, it is natural to attempt to extend the robustified estimator using this technique. To obtain a collaborative solution, we will employ the approach similar to the one adopted for the maximum-likelihood estimator, i.e., we will combine several estimates $\tilde{\varphi}_{1|k}$, $k = 1, 2, \ldots, K$. Note, however, that the robustified estimator is no longer the maximum-likelihood estimator, which means that the AIC/BIC frameworks need certain adjustments.

Top-level DoA search

$$\tilde{\boldsymbol{\Phi}}_k = \arg\min_{\boldsymbol{\Phi}_k} \sum_{n=1}^{N} J^2_{n|k}(\boldsymbol{\Phi}_k) \ ,$$

Lower-level iterative minimax solving procedure

For $n = 1, 2, \ldots, N$ execute:

1. Set $v_{n,0|k} = 0$, $\mathbf{\Lambda}_{0,n|k} = \mathbf{0}$
2. For $i = 0, 1, 2, .., I_{\max} - 1$

$$\begin{aligned} &\tilde{\boldsymbol{\theta}}_{n,i+1|k} = \left[\mathbf{\Psi}^{\mathrm{H}}_{n|k}(\boldsymbol{\Phi}_k)\tilde{\mathbf{\Psi}}_{n|k}(\boldsymbol{\Phi}_k) + \mathbf{\Lambda}_{i,n|k}\right]^{-1} \mathbf{\Psi}^{\mathrm{H}}_{n|k}(\boldsymbol{\Phi}_k)\boldsymbol{y}_n \\ &v_{n,i+1|k} = \|\boldsymbol{y}_n - \mathbf{\Psi}_{n|k}(\boldsymbol{\Phi}_k)\tilde{\boldsymbol{\theta}}_{n,i|k}\| \\ &\mathbf{\Lambda}_{i+1,n|k} = \mathrm{diag}\left(\epsilon_1 \frac{v_{n,i+1|k}}{|\tilde{A}_{1,n,i+1}|}, \ldots, \epsilon_k \frac{v_{n,i+1|k}}{|\tilde{A}_{k,n,i+1}|}\right) \end{aligned}$$

3. Set

$$J^2_{n|k}(\boldsymbol{\Phi}_k) = \left(v_{n,I_{\max}|k} + \sum_{j=1}^{K} \epsilon_j \left|\tilde{A}_{j,n,I_{\max}|k}\right|\right)^2$$

Table IV
SUMMARY OF ROBUSTIFIED ESTIMATOR PROPOSED IN MELLER AND STAWIARSKI.

We will consider two approaches, summarized below.

Scheme 1: In the first proposed approach, we will employ the robustified estimator to find the DoA of the echo, and then fall back to the conventional approach to find the likelihood of the model. That is, we use (15) to find the DoA estimate $\tilde{\boldsymbol{\Phi}}_k$, plug this estimate into (5) and (8) to obtain the likelihood, and eventually, compute AIC/AICc/BIC which may be used for model selection or averaging. The approach, in the collaborative variant, is summarized in Table V. The rationale behind such a scheme lies in the fact that, in the majority of cases, the behaviors of the conventional and the robustified estimators are practically identical. Consequently, most of the time, the corresponding averaged estimates will be computed using a reasonable set of weights. In the remaining, relatively small, fraction of events, when the the robustified estimator yields significantly different estimates of DoA than the conventional approach, we "place our bets" that the likelihoods and weights obtained using this procedure will not be corrupted to the point that the scheme breaks down.

Scheme 2: The second approach employs the fact that estimator (15) belongs to the class of M-estimators. For such estimators, a suitable generalization of AIC was proposed by Ronchetti in Rochetti [1997]. When applied to (15), Ronchetti's robust AIC takes the form

$$\mathrm{AICR} = 2V_k(\tilde{\boldsymbol{\Phi}}_k) + 2\mathrm{tr}(\tilde{\boldsymbol{J}}_k^{-1}\tilde{\boldsymbol{K}}_k) \ , \tag{20}$$

where

$$V_k(\boldsymbol{\Phi}_k) = \sum_{n=1}^{N} \left(\|\boldsymbol{y}_n - \mathbf{\Psi}_{n|k}(\boldsymbol{\Phi}_k)\tilde{\boldsymbol{\theta}}_{n,I_{\max}|k}\| + \sum_{j=1}^{K} \epsilon_j \left|\tilde{A}_{j,n,I_{\max}|k}\right|\right)^2$$

$\mathrm{tr}(\cdot)$ denotes the matrix trace, and $\tilde{\boldsymbol{J}}_k$ and $\tilde{\boldsymbol{K}}_k$ are estimators

MOST WIEDZY Downloaded from mostwiedzy.pl

DoA estimation using robustified algorithm

For $k = 1, 2, \ldots, K$

1. Find $\tilde{\boldsymbol{\Phi}}_k$ using algorithm from table IV.
2. Compute

$$\tilde{\sigma}_k^2 = \frac{1}{NM} \sum_{n=1}^{N} \boldsymbol{y}_n^{\mathrm{H}} \boldsymbol{Q}_{n|k}(\tilde{\boldsymbol{\Phi}}_k) \boldsymbol{y}_n$$
$$\tilde{l}_k = -NM \left[\log \tilde{\sigma}_k^2 + 1 + \log \pi\right]$$

Final DoA estimate using averaging

$$\bar{\tilde{\varphi}}_{1|K} = \sum_{k=1}^{K} \tilde{\mu}_k \tilde{\varphi}_{1|k}$$

where

$$\tilde{\mu}_k = \frac{\exp\left(-\frac{1}{2}\tilde{\mathrm{I}}\mathrm{C}_k\right)}{\sum_{k=1}^{K} \exp\left(-\frac{1}{2}\tilde{\mathrm{I}}\mathrm{C}_k\right)}$$

$$\tilde{\mathrm{I}}\mathrm{C}_k = \begin{cases} -2\tilde{l}_k + 2n_{\mathrm{p}}(k) & \text{for} \quad \mathrm{AIC} \\ -2\tilde{l}_k + \log(n_{\mathrm{o}})n_{\mathrm{p}}(k) & \text{for} \quad \mathrm{BIC} \\ -2\tilde{l}_k + 2n_{\mathrm{p}}(k) + 2\frac{n_{\mathrm{p}}(k)[n_{\mathrm{p}}(k)+1]}{n_{\mathrm{o}}-n_{\mathrm{p}}(k)-1} & \text{for} \quad \mathrm{AICc} \end{cases}$$

Table V

SUMMARY OF THE COLLABORATIVE ROBUSTIFIED ESTIMATOR EMPLOYING AIC/AICC/BIC.

of the matrices

$$\boldsymbol{J}_k = \mathrm{E}\left[-\frac{\partial^2 V_k(\boldsymbol{\Phi}_k)}{\partial \boldsymbol{\Phi}_k \partial \boldsymbol{\Phi}_k^{\mathrm{T}}}\right]\Bigg|_{\boldsymbol{\Phi}_k = \tilde{\boldsymbol{\Phi}}_k}$$
$$\boldsymbol{K}_k = \mathrm{E}\left[\frac{\partial V_k(\boldsymbol{\Phi}_k)}{\partial \boldsymbol{\Phi}_k} \frac{\partial V_k(\boldsymbol{\Phi}_k)}{\partial \boldsymbol{\Phi}_k^{\mathrm{T}}}\right]\Bigg|_{\boldsymbol{\Phi}_k = \tilde{\boldsymbol{\Phi}}_k}.$$

Following the standard practice, we propose to compute the estimates $\tilde{\boldsymbol{J}}_k$, $\tilde{\boldsymbol{K}}_k$ using the following formulas Claeskens and Hjort [2008]

$$\tilde{\boldsymbol{J}}_k = -\sum_{n=1}^{N} \frac{\partial^2 V_k(\boldsymbol{\Phi}_k))}{\partial \boldsymbol{\Phi}_k \partial \boldsymbol{\Phi}_k^{\mathrm{T}}}\Bigg|_{\boldsymbol{\Phi}_k = \tilde{\boldsymbol{\Phi}}_k}$$
$$\tilde{\boldsymbol{K}}_k = \sum_{n=1}^{N} \frac{\partial V_k(\boldsymbol{\Phi}_k)}{\partial \boldsymbol{\Phi}_k} \frac{\partial V_k(\boldsymbol{\Phi}_k)}{\partial \boldsymbol{\Phi}_k^{\mathrm{T}}}\Bigg|_{\boldsymbol{\Phi}_k = \tilde{\boldsymbol{\Phi}}_k}, \quad (21)$$

and to compute the derivatives using the numerical approach.

One may use Ronchetti's AIC to compute averaging weights by a straightforward modification of (11). The resultant collaborative scheme is summarized in Table VI.

Remark: : In principle, $\boldsymbol{J}_k$ and $\boldsymbol{K}_k$ above should include the derivatives over the real and the imaginary parts of the source amplitude vector $\tilde{\boldsymbol{\theta}}_{n,I_{\max}|k}$. However, by construction of the robustified estimator, the sample partial derivatives of $V_k(\boldsymbol{\Phi})$ over the source amplitudes are zero at the optimum, which means that one cannot estimated their ensemble means reliably. Unfortunately, this property means that the penalty computed using (21) is likely to be smaller than its correct amount, which may affect the performance of this approach adversely.

C. Computer simulations

We compared the performance of the nonadaptive robustified estimators assuming $k = 1$ and $k = 2$ and their adaptive – competitive and collaborative – counterparts by repeating the experiment performed in Section 2 (we used exactly the same data). Table VII shows the mean squared estimation errors of all algorithms obtained under small uncertainty of both model steering vectors ($\epsilon_1 = 0.01\|a_1(\phi_1)\|$ and $\epsilon_2 = 0.01\|a_2(\phi_2)\|$).

DoA estimation using robustified algorithm

For $k = 1, 2, \ldots, K$

1. Find $\tilde{\boldsymbol{\Phi}}_k$ using algorithm from table IV.
2. Compute AICR

$$\mathrm{AICR} = 2V_k(\tilde{\boldsymbol{\Phi}}_k) + 2\mathrm{tr}(\tilde{\boldsymbol{J}}_k^{-1}\tilde{\boldsymbol{K}}_k)$$

where

$$\tilde{\boldsymbol{J}}_k = -\sum_{n=1}^{N} \frac{\partial^2 V_k(\boldsymbol{\Phi}_k))}{\partial \boldsymbol{\Phi}_k \partial \boldsymbol{\Phi}_k^{\mathrm{T}}}\Bigg|_{\boldsymbol{\Phi}_k = \tilde{\boldsymbol{\Phi}}_k}$$
$$\tilde{\boldsymbol{K}}_k = \sum_{n=1}^{N} \frac{\partial V_k(\boldsymbol{\Phi}_k)}{\partial \boldsymbol{\Phi}_k} \frac{\partial V_k(\boldsymbol{\Phi}_k)}{\partial \boldsymbol{\Phi}_k^{\mathrm{T}}}\Bigg|_{\boldsymbol{\Phi}_k = \tilde{\boldsymbol{\Phi}}_k}$$
$$V_k(\boldsymbol{\Phi}_k) =$$
$$\sum_{n=1}^{N} \left(\|\boldsymbol{y}_n - \boldsymbol{\Psi}_{n|k}(\boldsymbol{\Phi}_k)\tilde{\boldsymbol{\theta}}_{n,I_{\max}|k}\| + \sum_{j=1}^{K} \epsilon_j \left|\tilde{A}_{j,n,I_{\max}|k}\right| \right)^2$$

Final DoA estimate using averaging

$$\bar{\tilde{\varphi}}_{1|K} = \sum_{k=1}^{K} \tilde{\mu}_k \tilde{\varphi}_{1|k}$$

where

$$\tilde{\mu}_k = \frac{\exp\left(-\frac{1}{2}\tilde{\mathrm{A}}\mathrm{ICR}_k\right)}{\sum_{k=1}^{K} \exp\left(-\frac{1}{2}\tilde{\mathrm{A}}\mathrm{ICR}_k\right)}$$

Table VI

SUMMARY OF THE COLLABORATIVE ROBUSTIFIED ESTIMATOR EMPLOYING AICR.

Observe that the results follow the pattern that occurred in Table III. The nonadaptive estimators are the poorest ones, and the collaborative approach works best. AICc and BIC demonstrate the best performance, and are closely followed by AIC. AICR works poorly, which is not unexpected given the previously discussed difficulties in estimating the penalty term reliably. What is quite remarkable, however, is the fact that the application of the basic (nonadaptive) robustified estimator improves the accuracy of the estimates over the conventional maximum likelihood approach considerably, which becomes apparent when one compares the fourth columns of Tables III and V.

This, rather unexpected, effect is actually a combination of two factors, which are related to how the robustified estimator reacts to different kind modeling errors. Table VIII shows a breakdown of mean squared errors achieved by algorithms (4) and (15), both using $k = 2$ sources, for the subsets of data where the multipath was absent and present, respectively.

Observe that the robustified estimator works better than the maximum likelihood approach for both subsets. To explain why, let us first recall that, while all estimators employed the simplification (14), the data was generated from a more elaborate model that accounted for the nonzero antenna height. That is, for the subset of data that includes multipath, the model used in the estimator is slightly erroneous, which hurts the accuracy of the maximum likelihood solution. The robustified estimator, on the other hand, proves much more robust to such

MOST WIEDZY Downloaded from mostwiedzy.pl

N	M	Nonadaptive $k=1$	Nonadaptive $k=2$	Competitive AIC	Competitive AICc	Competitive BIC	Competitive AICR	Collaborative AIC	Collaborative AICc	Collaborative BIC	Collaborative AICR	Clairvoyant estimator
1	4	6.37	3.56	3.34	6.37	3.34	3.58	3.01	6.37	**3.00**	3.12	2.40
	8	2.780	0.250	0.220	0.218	0.214	0.253	0.208	0.205	**0.202**	0.228	0.174
	16	0.5892	0.0097	0.0089	0.0085	0.0084	0.0097	0.0086	0.0083	**0.0082**	0.0092	0.0079
2	4	6.71	2.44	2.22	4.71	2.11	2.44	2.05	4.35	**1.99**	2.30	1.71
	8	2.707	0.134	0.113	0.119	0.115	0.134	**0.109**	0.113	0.110	0.126	0.096
	16	0.6293	0.0049	0.0042	0.0041	0.0040	0.0049	0.0042	0.0041	**0.0040**	0.0047	0.0040
3	4	6.18	1.83	1.66	2.91	1.63	1.83	1.55	2.79	**1.52**	1.78	1.26
	8	2.830	0.104	0.088	0.094	0.098	0.104	**0.086**	0.091	0.094	0.102	0.078
	16	0.6219	0.0036	0.0031	0.0030	**0.0030**	0.0036	0.0031	0.0030	0.0030	0.0035	0.0030

Table VII

COMPARISON OF MEAN SQUARED ESTIMATION ERRORS YIELDED BY NONADAPTIVE ROBUSTIFIED ESTIMATORS ASSUMING $k=1$ AND $k=2$ ($\epsilon_1 = 0.01\|a_1(\phi_1)\|$, $\epsilon_2 = 0.01\|a_2(\phi_2)\|$), COMPETITIVE AND COLLABORATIVE VERSIONS OF ROBUSTIFIED ESTIMATOR, AND CLAIRVOYANT ESTIMATOR, OBTAINED FOR THREE STANDARD UNIFORM ARRAYS, USING 5000 MONTE CARLO TRIALS FOR EACH CONFIGURATION. UNITS ARE DEGREES SQUARED.

N	M	Multipath absent ML	Multipath absent Robustified	Multipath present ML	Multipath present Robustified
1	4	6.39	**2.75**	7.07	**4.17**
	8	0.401	**0.195**	0.536	**0.270**
	16	0.0092	**0.0085**	0.0120	**0.0101**
2	4	4.50	**1.62**	5.02	**3.08**
	8	0.180	**0.095**	0.262	**0.170**
	16	0.0046	**0.0042**	0.0068	**0.0060**
3	4	3.50	**1.17**	3.98	**2.57**
	8	0.121	**0.066**	0.211	**0.128**
	16	0.0031	**0.0029**	0.0046	**0.0042**

Table VIII

COMPARISON OF MEAN SQUARED ESTIMATION ERRORS OF ALGORITHMS (4) AND (15), BOTH USING THE MODEL THAT ASSUMES $k=2$ SOURCES, UNDER THE ABSENCE AND THE PRESENCE OF THE MULTIPATH FOR $\epsilon_1 = 0.01\|a_1(\phi_1)\|$, $\epsilon_2 = 0.01\|a_2(\phi_2)\|$. THE UNITS ARE DEGREES SQUARED.

modeling errors. Second, the robustified estimator includes its own soft-decision mechanism that can recognize the absence of multipath efficiently. The mechanism is the consequence of the form of loading matrix $\mathbf{\Lambda}_{i,n|k}$ in recursion (18). If there is no multipath, the elements of the loading matrix $\mathbf{\Lambda}_{i,n|k}$ corresponding to the absent specular reflections will tend to inflate in consecutive iterations (18). In this way, the estimation of these components will softly “shut down”.

The strength of this mechanism increases with $\epsilon_2, \epsilon_3, \ldots$ until it makes the application of the collaborative extensions practically irrelevant – see Table IX for the results of an additional simulation experiment, this time performed with ϵ_2 increased to $0.1\|a_2(\phi_2)\|$. Since, in our experience Meller and Stawiarski, such levels of ϵ_2 work well with real-world data, we recommend to use the robustified estimator without any extensions, unless one decides to employ smaller levels of uncertainty.

IV. ADDITIONAL SIMULATION EXPERIMENTS

To gain more trust in our preliminary conclusions, we repeated our simulation experiments using a more realistic model, adopted from Barton [1974], which includes both the specular and the diffuse components of the multipath. In this model, the main parameters influencing the amount of both components are the surface roughness factor, defined as the ratio of the average height variation of the ground to the wavelength, σ_h/λ, the maximum ground surface slope β_0, and the vegetation attenuation factor ρ_v.

As previously, the simulated target was placed 5 km away from the radar, the targets's elevation angle was drawn randomly from the interval 2°-5°, and the SNR of its direct echo signal was fixed at 20 dB. The ground roughness factor was drawn randomly from the interval $[1, 20]$, where 1 corresponds to smooth ground, in which case the primary component of the multipath is the specular reflection, and 20 – to rough ground, when the diffuse scattering is the dominant one and the specular component hardly exists. The vegetation attenuation factor was also randomized, and drawn uniformly from the interval $[6, 30]$ in the dB scale. Finally, following Barton [1974], we set $\beta_0 = 0.1$.

We simulated a system using one snapshot, $N = 1$, and 16-element standard uniform array elevated to height $h_r = 4$ meters. To reduce the amount of of the diffuse component in the data, we preprocessed the raw array response using a beamformer and a whitening transformation.

The beamformer synthesized four beams: two sum beams, placed at 4° and 8°, and two corresponding difference beams. The sum beams employed the Chebyshev taper with -30 dB sidelobes, which resulted in the 3 dB beamwidth of approximately 8°. The difference beams employed the Bayliss taper, also with sidelobes at -30 dB.

The role of the whitening transformation was to decorrelate the measurement noise in the beamformer output, so as to make the data concordant with (1) which assumes uncorrelated components in the noise vector $\boldsymbol{v}_{n|k}$. Let $\boldsymbol{x}$, $\boldsymbol{W} = [\boldsymbol{w}_1\ \boldsymbol{w}_2\ \boldsymbol{w}_3\ \boldsymbol{w}_4]$ and $\boldsymbol{b} = \boldsymbol{W}^\mathrm{H}\boldsymbol{x}$ denote the beamformer input (unprocessed array output), the matrix of beamformer weights, and the beamformer output, respectively. The whitening transformation reads

$$\boldsymbol{y} = (\boldsymbol{W}^\mathrm{H}\boldsymbol{W})^{-1/2}\boldsymbol{b}\ ,$$

where $(\boldsymbol{W}^\mathrm{H}\boldsymbol{W})^{-1/2}$ denotes the matrix square root of $(\boldsymbol{W}^\mathrm{H}\boldsymbol{W})^{-1}$. The size of the observation vector resultant from the combination of beamforming and whitening is $M = 4$.

We compared the performance of the following algorithms: maximum likelihood estimators assuming $k = 1$ and $k = 2$, collaborative maximum likelihood estimators employing AIC and BIC (AICc was not included due to its poor performance for small M), and the vanilla robustified estimator with

MOST WIEDZY Downloaded from mostwiedzy.pl

$k = 2$, $\epsilon_1 = 0.01$ and $\epsilon_2 = 0.1$. The results are shown in Table X. Observe that, the application of the collaborative estimation brought an improvement over the conventional maximum likelihood estimator, but a smaller one than in our previous experiments. Moreover, there is no significant difference between AIC and BIC. On the other hand, the robustified estimator outperformed the remaining approaches by a rather safe margin (17% improvement over the ML estimator), which suggest that it is the best solution among the soft-decision approaches discussed in the paper.

V. Results from a real-world radar system

In this section, we illustrate the application of the soft-decision algorithms to the processing of data obtained from a real-world radar system. The data, which is the same as used in Meller and Stawiarski, was collected at an airport in northern Poland using a C-band radar. Similar to the simulated system is the previous section, this radar also employs the beamspace processing van Trees et al. [2002], and synthesizes two pairs of low-sidelobe sum-difference beams ($M = 4$), which were treated with the whitening transformation prior to the estimation.

The data was collected by observing a cooperative target, which was flying at a constant height at ranges from eight to thirteen kilometers. Overall, the target was illuminated by the radar beam 241 times. However, to avoid ambiguities in interpreting the results, we decided to remove one point of data that causes the maximum likelihood estimator particularly large problems (see Meller and Stawiarski for more details), which reduces the size of the dataset to 240 scans.

During each such "scan" the radar transmitted three bursts of pulses. Echo signals collected from of each burst were processed using a bank of Doppler filters, so that three ($N = 3$) snapshots were collected per scan.

We compared the behavior of three solutions: the conventional maximum likelihood estimator using the model with two sources, the collaborative solution employing the combination of the conventional maximum likelihood estimator (4) and weighting using AIC (13), and the robustified estimator (16). Fig. 2 shows the comparison of the estimates of target elevation obtained using the compared methods. Note that the results were normalized by the 3 dB beamwidth of the radar, which is in the order of a few degrees.

On the qualitative level, the estimates obtained using all approaches are similar. The quantitative results, summarized in Table XI, agree with our previous simulations – the nonadaptive estimator has the worst performance, the application of the averaging technique improves the results, but only slightly, and, once again, best performance was obtained using the robustified estimator.

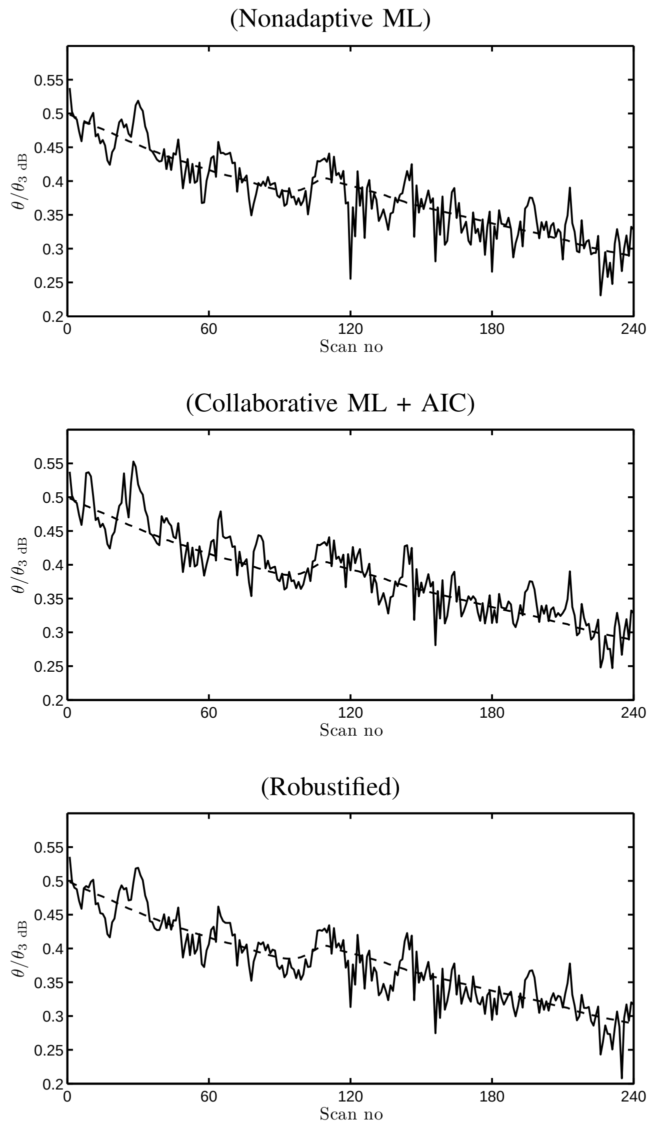

Figure 2. Comparison of estimates of target elevation, normalized by 3 dB beamwidth of the radar, obtained using three approaches. The dashed lines shows the true elevation of the target.

VI. Conclusions

In this paper, we studied the application of novel estimators that involve soft-decision mechanisms, to the problem of estimating the elevation under specular multipath. In the context of the maximum likelihood framwork, we developed such schemes by averaging the partial estimates obtained using the models with different number of sources. We also studied the, recently introduced, robustified estimator, and found that it includes a soft-decision mechanism of its own kind. The behavior of the proposed algorithms was verified using several computer simulations and a real-world dataset. While all proposed solutions improve the estimation accuracy over the basic maximum likelihood estimator, best results were obtained using the robustified solution.

References

H. Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19:716–723, 1974.

H. Akaike. A Bayesisan analysis of the minimum AIC procedure. *Annals of the Institute of Statistical Mathematics*, 30:9–14, 1978.

D. Barton. Low-angle radar tracking. *Proceeding of the IEEE*, 62(6):687–704, 1974.

MOST WIEDZY Downloaded from mostwiedzy.pl

N	M	Nonadaptive $k=1$	Nonadaptive $k=2$	Competitive AIC	Competitive AICc	Competitive BIC	Competitive AICR	Collaborative AIC	Collaborative AICc	Collaborative BIC	Collaborative AICR	Clairvoyant estimator
1	4	6.35	3.61	3.66	6.35	3.66	3.75	3.58	6.35	3.58	**3.52**	3.21
	8	2.780	0.331	0.329	0.334	0.328	0.374	0.327	0.333	**0.327**	0.366	0.316
	16	0.5892	0.0105	0.0103	0.0102	0.0101	0.0105	0.0102	0.0102	**0.0101**	0.0103	0.0100
2	4	6.71	3.05	3.06	5.26	3.08	3.19	**3.05**	5.06	3.07	3.16	2.94
	8	2.707	0.251	0.249	0.255	0.254	0.287	**0.249**	0.254	0.252	0.284	0.245
	16	0.6293	0.0068	0.0067	0.0067	0.0067	0.0068	0.0067	0.0066	**0.0066**	0.0068	0.0066
3	4	6.18	**2.58**	2.59	3.68	2.63	2.74	2.58	3.63	2.62	2.72	2.51
	8	2.830	0.249	**0.248**	0.253	0.257	0.285	0.248	0.253	0.255	0.284	0.246
	16	0.6219	0.0058	0.0057	0.0057	**0.0057**	0.0058	0.0057	0.0057	0.0057	0.0058	0.0057

Table IX
COMPARISON OF MEAN SQUARED ESTIMATION ERRORS YIELDED BY NONADAPTIVE ROBUSTIFIED ESTIMATORS ASSUMING $k=1$ AND $k=2$ ($\epsilon_1 = 0.01\|a_1(\phi_1)\|$, $\epsilon_2 = 0.1\|a_2(\phi_2)\|$), COMPETITIVE AND COLLABORATIVE VERSIONS OF OF ROBUSTIFIED ESTIMATOR, AND CLAIRVOYANT ESTIMATOR, OBTAINED FOR THREE STANDARD UNIFORM ARRAYS, USING 5000 MONTE CARLO TRIALS FOR EACH CONFIGURATION. UNITS ARE DEGREES SQUARED.

Estimator	MSE
Nonadaptive ML (k=1)	0.0815
Nonadaptive ML (k=2)	0.0481
Collaborative ML + AIC	0.0470
Collaborative ML + BIC	0.0470
Robustified (k=2)	$\mathbf{0.0400}$

Table X
COMPARISON OF MEAN SQUARED ESTIMATION ERRORS YIELDED BY SEVERAL SOFT-DECISION ESTIMATORS IN AN EXTENDED SIMULATION EXPERIMENT THAT INCLUDES DIFFUSE MULTIPATH. UNITS ARE DEGREES SQUARED.

Estimator	Normalized MSE
Nonadaptive ML (k=2)	$8.06 \cdot 10^{-4}$
Collaborative ML + AIC	$8.01 \cdot 10^{-4}$
Robustified (k=2)	$\mathbf{7.04 \cdot 10^{-4}}$

Table XI
COMPARISON OF MEAN SQUARE ERRORS OF NORMALIZED ESTIMATES OF ELEVATION OBTAINED USING THREE SOFT-DECISION ESTIMATORS FOR A REAL-WORLD DATASET.

D. Bonacci, F. Vincent, and B. Gigleux. Robust DoA estimation in case of multipath environment for a sense and avoid airborne radar. *IET Radar, Sonar & Navigation*, 11 (5):797–801, 2015.

J. Capon. High resolution frequency-wavenumber spectrum analysis. *Proceeding of the IEEE*, 57:1408–1418, 1969.

G. Claeskens and N. L. Hjort. *Model Selection and Model Averaging*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2008.

L. E. Ghaoui and H. Lebret. Robust solutions to least-squares problems with uncertain data. *SIAM Journal on Matrix Analysis and Applications*, 18(4):1035–1064, 1997.

C. M. Hurvich, J. S. Simonoff, and C.-L. Tsai. Smoothing parameter selection in nonparametric regression using an improved Akaike information criterion. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 60 (2):271–293, 1998.

A. G. Jaffer. Maximum likelihood direction finding of stochastic sources: a separable solution. In *Proc. 1988 International Conference on Acoustics, Speech, and Signal Processing (ICASSP-88)*, pages 2893–2896, 1988.

S. Kritchman and B. Nadler. Non-parametric detection of the number of signals: hypothesis testing and random matrix theory. *IEEE Transactions on Signal Processing*, 57(10): 3930–3941, 2009.

N. T. Longford. 'Which model?' is the wrong question. *Statistica Neerlandica*, 66(3), 2012.

M. L. Meeks. *Radar Propagation at Low Altitudes*. Artech House, 1982.

M. Meller and K. Stawiarski. Robust estimators of elevation using the specular multipath model. Technical report WETI/KSA/2018, available at https://pg.edu.pl/documents/1094802/fee9a155-d2d8-4ec4-831e-7bbf78e4b131.

U. Nickel. Determination of the dimension of the signal subspace for small sample size. In *Proceedings of the IASTED international conference on Signal Processing and Communication Systems*, pages 119–122. IASTED/Acta Press, 1998.

U. Nickel. Array processing for radar: Achievements and challenges. *International Journal of Antennas and Propagation*, 2013.

V. V. Reddy, M. Mubeen, and B. P. Ng. Reduced-complexity super-resolution DOA estimation with unknown number of sources. *IEEE Signal Processing Letters*, 22(6):772–776, 2015.

J. Rissanen. Modeling by shortest data description. *Automatica*, 14(5):465–471, 1978.

E. Rochetti. Robustness aspects of model choice. *Statistica Sinica*, 7:327–338, 1997.

R. Roy and T. Kailath. ESPRIT – Estimation of Signal Parameters via Rotational Invariance Techniques. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 37(7):984–995, 1989.

R. Schimidt. Multiple emiter location and signal parameter estimation. *IEEE Transactions on Antennas and Propagation*, 34(3):276–280, 1986.

G. Schwarz. Estimating the dimension of a model. *The Annals of Statistics*, 6(2):461–464, 03 1978.

S. M. Sherman. Complex indicated angles applied to unresolved radar targets and multipath. *IEEE Transactions on Aerospace and Electronic Systems*, AES-7(1):160–170, 1971.

H. L. van Trees, K. L. Bell, and Z. Tian. *Detection, Estimation and Modulation Theory, Part IV: Optimum Array Processing*. John Wiley & Sons, 2002.

MOST WIEDZY Downloaded from mostwiedzy.pl

H. L. van Trees, K. L. Bell, and Z. Tian. *Detection, Estimation and Modulation Theory, Part I: Detection, Estimation, and Filtering Theory*. John Wiley & Sons, 2013.

F. Vincent, O. Besson, and E. Chaumette. Approximate maximum likelihood estimation of two closely spaced sources. *Signal Processing*, 97(Supplement C):83 – 90, 2014.

W. D. White. Low-angle radar tracking in the presence of multipath. *IEEE Transactions on Aerospace and Electronic Systems*, AES-10(6):835–852, 1974.

W.-D. Wirth. *Radar Techniques Using Array Antennas*. Radar, Sonar & Navigation. Institution of Engineering and Technology, 2013.

E. Wit, E. v. d. Heuvel, and J.-W. Romeijn. 'All models are wrong...': an introduction to model uncertainty. *Statistica Neerlandica*, 66(3), 2012.

MOST WIEDZY Downloaded from mostwiedzy.pl