M. Meller and K. Stawiarski, "Robustified estimators of radar elevation angle using a specular multipath model," in IEEE Transactions on Aerospace and Electronic Systems. DOI: 10.1109/TAES.2019.2927906

Robustified estimators of radar elevation angle using a specular multipath model

Michał Meller, Kamil Stawiarski

***Abstract*—We consider the problem of estimating the elevation angle in the presence of multipath. The proposed method belongs to the class of maximum likelihood-like estimators and employs a modified specular reflection model that accounts for the uncertainty of the steering vector by assuming that they are subject to unknown deterministic perturbations with bounded norms. The analysis, performed using convex optimization methods, allows us to obtain computationally efficient implementations of the approach. Real-world results and computer simulations confirm the improved behavior of the proposed robustified estimators.**

***Index Terms*—radar, multipath, low-angle estimation, robust estimation**

I. Introduction

MODERN radars usually belong to the class of three dimensional (3D) systems, which means they are capable of determining the range, azimuth angle, and elevation angle coordinates of detected targets. The ability to estimate the elevation angle differentiates 3D radars from the older generation of two dimensional systems, which were able to measure the range and azimuth only [1]. In 3D systems, the estimates of target elevation angle are typically obtained using the conventional monopulse processing. This technique, despite its age, is still very much widespread because it is computationally cheap and provides a good accuracy over a broad envelope of conditions.

The principle of the monopulse method can be succinctly summarized as follows [2]–[4]. The antenna subsystem of the monopulse radar synthesizes two patterns, called the sum and the difference pattern, respectively. The sum pattern is used to detect targets and to provide a coarse estimate of their direction, whose accuracy is comparable to the width of the sum beam. In order to precisely localize the target inside the beam, the difference pattern, which has a null at the peak of the sum pattern, is employed. The monopulse estimator relies on the fact that, assuming that the target indeed resides inside the mainlobe of the sum beam, the so-called monopulse ratio – the ratio of the difference and the sum signals – is a function of the target's deviation from the center of the beam [2], [4].

A considerable weakness of the monopulse method is its poor performance at estimating the elevation angle of low-altitude targets [5]. The poor accuracy of the monopulse stems from the fact that the method is not designed to cope with the situations when there are multiple closely spaced waveforms arriving at the array. Unfortunately, due to the phenomenon called multipath, precisely this situation arises when the radar is observing the low flying targets. The term multipath refers to the situation when the target's backscatter reflects from the earth surface and reaches the radar array from multiple directions. For the low flying targets, the direct and the reflected waveforms are indeed very close to each other, which causes the monopulse estimator to fail.

Multiple methods have been proposed to solve this problem. Early approaches included the off-boresight tracking, frequency diversity/hopping [6], [7] or the complex indicated angle method [8]. Unlike the standard monopulse, which employs only the real part of the monopulse ratio, the complex indicated angle method utilizes the full complex value of the monopulse ratio. In theory, such processing enables one to compensate the influence of the multipath. Unfortunately, the method is based on a restrictive set of assumptions about the nature of the multipath and fails when these conditions do not take place [6]. Despite its flaws, the complex angle method has left its footprint in the practice of using both components of the monopulse ratio to detect the presence of multiple waveforms within a single beamwidth [4].

Modern solutions typically employ one of many so-called "superresolution" methods. The adjective "superresolution" comes from the fact that, in principle, these methods can resolve multiple sources, separated by less than one beamwidth.

Superresolution methods generally belong to one of two groups, called nonparametric methods and parametric methods, respectively.

Nonparametric methods, which include, among others, the classical Capon method [9], the MUSIC algorithm [10], and the minimum norm method [11], [12], attempt to work out nonparametric estimates of the spatial spectrum. In principle, by scanning the spectral estimate for peaks, one may identify the directions of the waveforms arriving at the array. Unfortunately, many nonparametric methods are poorly suited to radar because of their two characteristics. First, they usually require multiple snapshots of data to form a full-rank covariance matrix. However, in radar, it is not uncommon (in fact it is typical) to work with only one snapshot. Second, these methods often assume that the sources are independent, which is not true in radar, where, excluding jammers, all sources reflect the same waveform and therefore are correlated. To mitigate both problems, one may employ the spatial smoothing technique [13]–[15]. Unfortunately, the application of this method reduces the array resolution, which is very undesirable and may be impractical for small arrays.

The second group of methods, i.e., parametric methods, consists of the model-based techniques. In this approach, one attempts to fit a parametric model, that typically includes the

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. e-mails: michal.meller@eti.pg.edu.pl, kamil.stawiarski@pitradwar.com

unknown source angles, to available data. Parametric methods usually employ the maximum likelihood principle and Gaussian models [16]. Gaussian models may come in two flavors, which gives rise to two types of Gaussian maximum likelihood estimators. If the amplitudes of the waveforms are treated as unknown deterministic variables and included into the set of the model parameters, the resultant estimator is referred to as the deterministic, or conditional, maximum likelihood estimator. Alternatively, one may treat the waveform amplitudes as independent realizations of zero-mean Gaussian-distributed random variables and parametrize only their variances. In such a case, the resultant estimator is called the stochastic, or the unconditional, maximum likelihood estimator [16]–[18].

A considerable advantage of the maximum likelihood approach lies in the fact that it does not require the assumption that the sources are uncorrelated and can work with one snapshot. However, the direct application of the approach results in the estimators that are computationally complex because of the need to solve a multidimensional, nonconvex optimization problem. To overcome this obstacle, multiple iterative and approximate approaches have been proposed, such as the alternating projection algorithm [19], the EM algorithm [20] or the IQML algorithm [21]. Another low computational cost solution, which employed three specially arranged beams, was proposed in [22]. More recently, approximate versions of the conditional and the unconditional maximum likelihood estimators were studied in [23] and [24]. Successful applications of the maximum likelihood approach to the problem of estimating the elevation angle at low angles can be found in [5], [18], [25], among others.

In [18] Nickel pointed out that, despite their differences, none of the superresolution methods stands out concerning the estimation accuracy when applied to real-world datasets. He attributed this behavior to the sensitivity of these methods to the inaccuracies of the underlying model. On these grounds, one could conclude that one of the primary factors determining the choice of the estimation method should be the computational complexity, which makes the conditional maximum likelihood estimator particularly attractive for application in radar. However, we observed that, under model uncertainty, the conditional maximum likelihood estimator occasionally fails, in the sense that it delivers grossly erroneous estimates. Such instability of the estimator is highly undesirable, not only because of the resultant increase in the mean squared estimation error. The "observed" behavior of the target may be interpreted as a rapid maneuver, which may lead to erroneous predictions of its position, and, possibly, to a loss of track. Additionally, the radar scheduler may react to the apparent maneuver by increasing the track update rate, which will waste the ever-scarce resources of the radar.

The contributions of the paper are the following. First, we show that one can improve the behavior of the conditional maximum likelihood estimator using the minimax approach. Second, studying the proposed minimax estimator using convex optimization methods allows us to reach a computationally attractive solution, whose complexity remains comparable to the conventional approach. The low cost of the proposed estimator facilitates its implementation and makes it an attractive alternative to the standard conditional maximum likelihood method. Third, we demonstrate the behavior of the conventional and the proposed approach using two real-world datasets and realistic simulations.

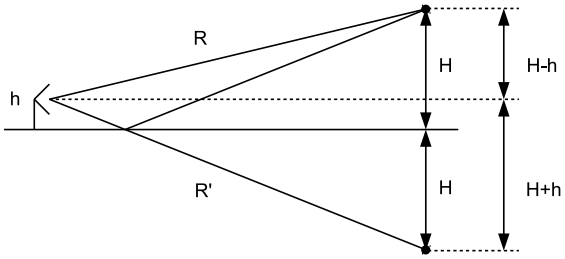


Figure 1. The flat earth model of the specular multipath.

The paper is organized as follows: Section II explains the motivating factors for this paper. We review the specular multipath model and two estimators based on the maximum likelihood principle. Using two real-world datasets, we demonstrate that these estimators can exhibit gross estimation errors. Sections III and IV form the main part of the paper. In Section III, the proposed robustified maximum likelihood estimator is derived. Its behavior is verified using the real-world datasets and simulations in Section IV. The conclusions are summarized in Section V.

II. Conventional Maximum Likelihood Estimation of Elevation Angle

Like many authors before us, we begin by reviewing a basic model of the specular multipath. Next, we consider the deterministic maximum likelihood direction-of-arrival (DoA) estimator. Using the geometric properties of the specular multipath model, we propose a quick, i.e., a computationally cheap, estimator of elevation. Note that, this solution will only serve a preliminary role and will be improved in Section III.

A. Flat earth model of specular multipath

Qualitative and, to a large extent, quantitative effects of the multipath can be explained using a simple model of the specular reflection. In this model, depicted in Fig. 1, the ground is treated as a flat, reflective surface, and the laws of optics are used to study the wave propagation [2], [26].

Suppose that the radar antenna is elevated to a height h and that the target altitude above ground is H, $H \gg h$. Denote by R, $R \gg H$, the target range, i.e., the distance traveled by the direct echo. The target's elevation angle, as seen from the array, is

$$\alpha_{\mathrm{d}} = \frac{H-h}{R} \,. \tag{1}$$

Following the laws of optics, one may regard the reflected waveform may as originating from the mirror source, located at the depth H below the ground surface, i.e., at the elevation angle

$$\alpha_{\mathrm{m}} = -\frac{H+h}{R'} \,, \tag{2}$$

where [26]

$$R' \simeq R + \frac{2hH}{R}$$

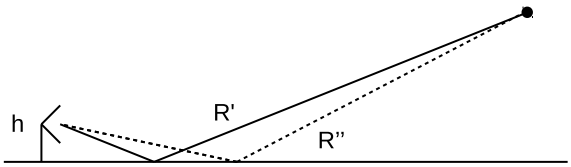


Figure 2. The definition of the quantity $R''(\alpha)$.

denotes the total distance traveled by the reflected waveform.

For low elevation angles, the amplitude of the reflected signal may be close to that of the direct one. Additionally, under this basic model, the phase of the reflected signal exhibits a delay equal to [2], [26]

$$\Delta\varphi \cong \frac{2\pi}{\lambda}\frac{2hH}{R} + \pi \ , \tag{3}$$

where the first term corresponds to the difference in the lengths of the two propagation paths, and π is an additional shift, caused by the reflection from the ground.

The output of the radar array is a noise-corrupted combination of its responses to the direct and the mirror waveforms

$$\boldsymbol{y} = A_{\mathrm{d}}\boldsymbol{a}(\alpha_{\mathrm{d}}) + A_{\mathrm{m}}\boldsymbol{a}(\alpha_{\mathrm{m}}) + \boldsymbol{v} \ , \tag{4}$$

where A_{d}, A_{m} denote the direct and the mirror signal complex "amplitudes", respectively, $\boldsymbol{a}(\alpha)$ is the array response vector in the direction α, and $\boldsymbol{v}$ is a zero mean, complex circular Gaussian white noise with a known covariance matrix $\boldsymbol{R} = \sigma_v^2\boldsymbol{I}$, $\boldsymbol{v} \sim \mathcal{CN}(\boldsymbol{0}, \sigma_v^2\boldsymbol{I})$.

Remark 1: The specular reflection model is valid provided that the ground surface is sufficiently smooth. According to [6], the following condition

$$\delta \le \frac{1}{8}\frac{\lambda}{\sin\alpha} \ , \tag{5}$$

where δ denotes the difference in the ground height, must hold in the first Fresnel zone around the specular-reflected ray, i.e., in the zone where

$$R''(\alpha) - R' \le \lambda/2$$

and $R''(\alpha)$ denotes the distance traveled by the hypothetical diffuse reflection ray arriving from the direction α – see Fig. 2 for the graphical explanation.

Remark 2: The noise covariance matrix does not, in general, adopt the diagonal structure assumed above. Consider, for example, the beamspace processing, which we will employ in our examples. The beamspace approach facilitates computations by reducing the dimensionality of the data [15]. In such case, the vector $\boldsymbol{y}$ might correspond to a fan of beams, covering a sector of elevation of interest. Assuming that the beams were digitally formed using weight vectors $\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_K$, the noise covariance matrix takes the form

$$\boldsymbol{R} = \sigma_v^2\boldsymbol{W}^{\mathrm{H}}\boldsymbol{W} \ ,$$

where σ_v^2 is the receiver noise variance, and

$$\boldsymbol{W} = [\boldsymbol{w}_1 \ \boldsymbol{w}_2 \ \ldots \ \boldsymbol{w}_K] \ ,$$

denotes the matrix of beamformer weights, assumed to have a full column rank.

It is well known that one may handle a case like this using the whitening transformation [16]. Let $\boldsymbol{X}^{1/2}$ denote the square root of a conjugate symmetric positive definite matrix $\boldsymbol{X}$, i.e., any matrix satisfying $\boldsymbol{X}^{1/2}(\boldsymbol{X}^{1/2})^{\mathrm{H}} = \boldsymbol{X}$. Set $\boldsymbol{R}^{-1/2} = (\boldsymbol{R}^{-1})^{1/2}$. Preprocessing the data using the following linear transformation

$$\boldsymbol{y} \leftarrow \boldsymbol{R}^{-1/2}\boldsymbol{y} \tag{6}$$

will make it concordant with (4).

B. Maximum likelihood estimator of elevation

Suppose that N independent L-variate observations of the target echo, $\boldsymbol{y}_n$, $n = 1, 2, \ldots, N$, are available, e.g., from several consecutive coherent processing intervals (CPIs). To enable greater flexibility, such as allowing the antenna rotation, varying the CPI length or employing the frequency agility, we shall treat the amplitudes of the direct and the mirror signal in each observation as different deterministic variables, i.e., we adopt the following model

$$\boldsymbol{y}_n = A_{\mathrm{d},n}\boldsymbol{a}_n(\alpha_{\mathrm{d}}) + A_{\mathrm{m},n}\boldsymbol{a}_n(\alpha_{\mathrm{m}}) + \boldsymbol{v}_n \ , \tag{7}$$

where the noise $\boldsymbol{v}_n$ is assumed to form an i.i.d. sequence of mutually independent random vectors.

Generally, the vector $\boldsymbol{y}_n$ is the preprocessed array response at the range of the target in the n-th CPI. The preprocessing may include, among others, beamforming, Doppler filtering, and a CFAR detector to separate the target from the clutter.

Let $\boldsymbol{\alpha} = [\alpha_{\mathrm{d}} \ \alpha_{\mathrm{m}}]^{\mathrm{T}}$. The application of standard results from the Gaussian maximum likelihood estimation theory [16] leads to the following conditional maximum likelihood estimator, which one can recognize as a nonlinear least squares fitting problem

$$\hat{\boldsymbol{\alpha}} = \arg\min_{\boldsymbol{\alpha}} \ \sum_{n=1}^{N} \left\| \boldsymbol{y}_n - \boldsymbol{\Psi}_n(\boldsymbol{\alpha})\hat{\boldsymbol{\theta}}_n(\boldsymbol{\alpha}) \right\|^2 \ , \tag{8}$$

where

$$\boldsymbol{\Psi}_n(\boldsymbol{\alpha}) = [\boldsymbol{a}_n(\alpha_{\mathrm{d}}) \ \boldsymbol{a}_n(\alpha_{\mathrm{m}})]$$
$$\hat{\boldsymbol{\theta}}_n(\boldsymbol{\alpha}) = \left[\boldsymbol{\Psi}_n^{\mathrm{H}}(\boldsymbol{\alpha})\boldsymbol{\Psi}_n(\boldsymbol{\alpha})\right]^{-1}\boldsymbol{\Psi}_n^{\mathrm{H}}(\boldsymbol{\alpha})\boldsymbol{y}_n$$

denote the regression matrix and the maximum likelihood estimate of the n-th vector of complex amplitudes, $\boldsymbol{\theta}_n = [A_{\mathrm{d},n} \ A_{\mathrm{m},n}]^{\mathrm{T}}$, respectively.

C. Quick estimator of elevation

Implementation of the estimator (8) requires one to perform a minimization over two variables, α_{d} and α_{m}, which is computationally expensive and may limit the method's attractiveness, especially under tight constraints of real-time systems.

To reduce the computational complexity, one can exploit the fact that, when $h \ll H$, it holds that [c.f. (1), (2)]

$$\alpha_{\mathrm{m}} \cong -\alpha_{\mathrm{d}} \ . \tag{9}$$

Provided that the 3 dB beamwidth of the array is sufficiently large, one can replace $\boldsymbol{a}_n(\alpha_\mathrm{m})$ in (7) with $\boldsymbol{a}_n(-\alpha_\mathrm{d})$. The resulting model

$$\boldsymbol{y}_n \cong A_{\mathrm{d},n}\boldsymbol{a}_n(\alpha_\mathrm{d}) + A_{\mathrm{m},n}\boldsymbol{a}_n(-\alpha_\mathrm{d}) + \boldsymbol{v}_n \ , \tag{10}$$

leads to the following "quick" estimator (see, e.g., [5], [27], [28] for other solutions employing this approximation)

$$\hat{\alpha}_\mathrm{d} = \arg\min_{\alpha_\mathrm{d}} \sum_{n=1}^{N} \left\| \boldsymbol{y}_n - \boldsymbol{\Psi}_n(\alpha_\mathrm{d}, -\alpha_\mathrm{d})\hat{\boldsymbol{\theta}}_n(\alpha_\mathrm{d}, -\alpha_\mathrm{d}) \right\|^2$$
$$\hat{\boldsymbol{\theta}}_n(\boldsymbol{\alpha}) = \left[\boldsymbol{\Psi}_n^\mathrm{H}(\boldsymbol{\alpha})\boldsymbol{\Psi}_n(\boldsymbol{\alpha})\right]^{-1}\boldsymbol{\Psi}_n^\mathrm{H}(\boldsymbol{\alpha})\boldsymbol{y}_n \ , \tag{11}$$

which is considerably less costly than (8) because it requires one to perform the search over only one variable.

Estimator (11) is particularly attractive when it is implemented with the aid of the beamspace processing. As we shall demonstrate in the next subsection, under such setup, the quick approach – although not entirely free of problems – can be very competitive in terms of estimation accuracy. Note however that, in the context of this work, estimator (11), serves a preliminary role and will undergo an enhancement to improve its properties.

D. Real-world behavior of maximum likelihood estimators – a case study

To motivate the remaining part of the paper, we will show two examples of the real-world behavior of the maximum likelihood estimator (8) and its quick counterpart (11).

A cooperative air-breathing target was observed using a C-band radar ($\lambda \approx 5$ cm) deployed at an airport in northern Poland. The terrain at the airport is level and, in the direction of the observed target, covered by a grass until at least 1.5km away from the radar. Overall, it is reasonable to regard the condition (5) as generally satisfied at this location.

The target, which was flying at a constant height, was observed in the track-while-scan mode. At the start of the session, the range of the target was, approximately, 7500 m. Excluding a short loop imposed by the air traffic control, the target moved away from the radar. The collection of data ended when the range of the target reached 13500 m. During this time, more than 200 scans (dwells) of the radar beam over the target occurred.

The preprocessing chain in the radar consists of a beamformer, a bank of MTD filters, a CFAR detector, and the whitening transformation (6). The beamformer forms two pairs of the low-sidelobe sum-difference patterns (Fig. 3), i.e., the size of the observation vector $\boldsymbol{y}_n$, $n = 1,\ 2,\ \ldots,\ N$ is $L = 4$. Moreover, the radar's scan program was adjusted such that, for each scan, the target was illuminated for $N = 3$ consecutive CPIs.

Fig. 4 compares the true values of normalized target elevation angle, defined as the ratio of the elevation angle to the 3 dB beamwidth of the system (whose value is in the lower end of a few degrees), with their estimates, obtained using the monopulse method (the monopulse ratio was computed from the lower sum-difference pair from the snapshot with the highest power of the sum signal), the maximum likelihood estimator (8) and the quick maximum likelihood estimator (11).

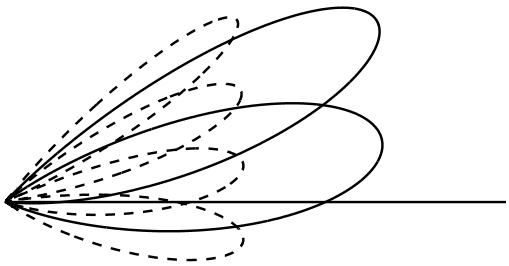


Figure 3. A schematic representation of the arrangement of the beams employed by the radar system discussed in the case study. Solid lines – sum beams. Dashed lines – difference beams. Note that the Y axis is expanded greatly.

The results yielded by the monopulse estimator confirm that there is a good deal of multipath, especially in the first half of the dataset, when the errors are quite large. The observed behavior of estimator (8) is very erratic, which can be explained by the unfavorable relation of the number of unknowns to the number of observations. In the system considered, estimator (8) must work out the values of following parameters: two angles, α_d and α_m, $2N$ complex amplitudes $A_{\mathrm{d},n}$, $A_{\mathrm{m},n}$, $n = 1, 2, \ldots, N$ and, implicitly, the noise variance σ_v^2. The total count of real-valued quantities to estimate is $3 + 4N$, or 15 for $N = 3$. On the other hand, only N 4-element complex valued (recall that $L = 4$ because the preprocessing includes a beamformer), snapshots are available, which accumulates to $8N = 24$ real-valued observables for $N = 3$. Such a ratio of the number of unknowns to the available data, coupled with the nonlinearity of the parametric model, results in the phenomenon known as overfitting [16]. The overfitting occurs when there is insufficient data to support the complexity of the model. In such a case the model's goodness of fit is seemingly very high because the residual error in (8) is minimal. However, the reduction of the residual error does not translate to the improvement of the actual accuracy of the estimator – the estimation errors increase [16].

The overfitting can be avoided by increasing the amount of the available data or by using simpler models. One can increase the amount of available data in two ways: by increasing the number of snapshots N, or by increasing the dimension of the vector $\boldsymbol{y}_n$. However, both solutions were found unacceptable for this system either due to the increase in the system's update period beyond the limit or because of the increase in the computational complexity.

For the system in question, the application of the simplified estimator (11) is a better solution than increasing N or the number of processed beams. The reduced dimensionality of the simplified model, achieved with the substitution (9), pays twofold. First, it reduces computational costs because the simplified estimator requires a one-dimensional search [c.f. (11)]. Second, it eliminates the negative effects of overfitting.

Unfortunately, the simplified estimator is not entirely free of problems. The estimate of the target's elevation at scan 217

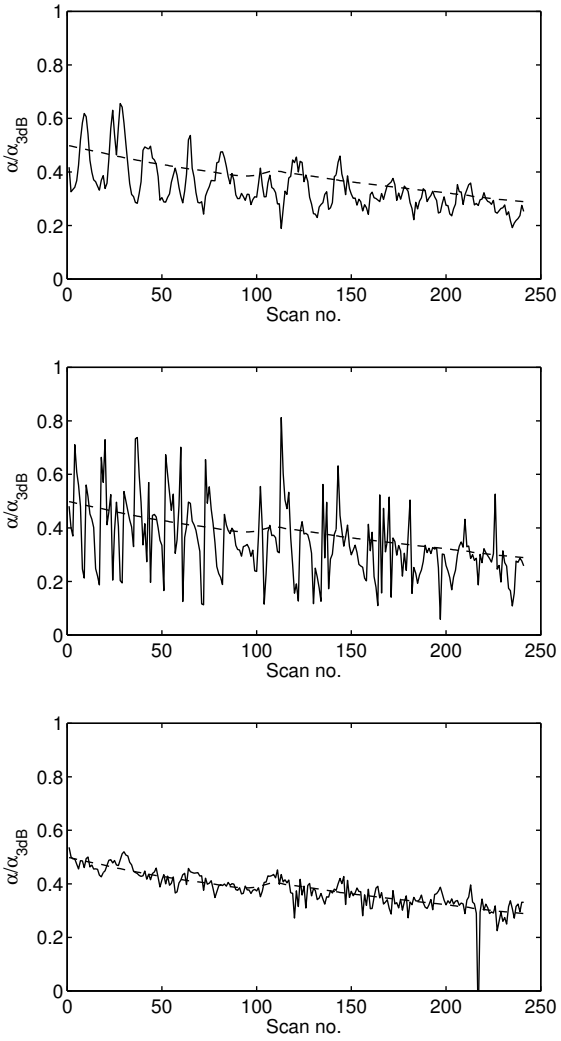


Figure 4. The comparison of the true values of the target's normalized elevation angle (dashed line) with their estimates obtained using the monopulse method (top plot), the maximum likelihood method (middle plot) and the quick maximum likelihood method (bottom plot).

is erroneous. We have witnessed such gross errors occurring not only at this site, but in other places as well. Moreover, we established that they unlikely to be related to the overfitting – Fig. 5 shows another instance of this problem, which occurred despite the fact that the size of the observation vector was increased to 6 elements by including an additional sum-difference pair. We also note that the outliers are not always as severe as shown in Figs. 4 and 5, i.e., the estimated elevation angle is not necessarily that close to zero, which precludes the application of simple rules to filter the errors out.

Even if such erroneous estimates seem relatively isolated, they can have rather undesirable consequences when the presence of a tracker is taken into account. If the erroneous estimate is assigned to an existing track, the accuracy of the track will degrade significantly, which may cause the loss of track.

Additional problems may arise if one considers a multifunction radar system. In the simplest case, a multifunction radar divides its time budget between two tasks: search, where the

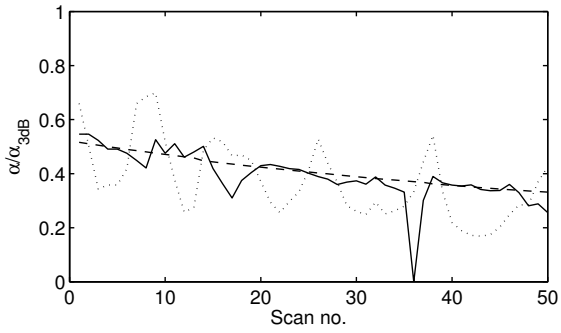


Figure 5. The comparison of the true values of the target's normalized elevation angle (dashed line) with their estimates obtained using the monopulse method (dotted line) and the quick maximum likelihood method (solid line) – the second dataset.

radar scans the surrounding space seeking new targets, and track, where the radar updates the targets that were previously detected and are included in the track file [3]. To improve their resource management, multifunction radars may employ various techniques, such as the variable update rate and the sequential detection [2]. If a tracker in such a radar interprets the change in the target's elevation angle as resulting from an abrupt vertical maneuver, the track update rate may be increased, which will reduce the time budget available for the search function and for maintaining other tracks. If, on the other hand, the erroneous plot is treated as originating from a possible new target, additional confirmation dwells may be scheduled, which will also affect the resources available for other tasks.

In the next section, we will elaborate on the possible reasons for the observed estimator instability. We will also propose a robustified version of the quick estimator, which eliminates this undesired effect.

Remark: Originally, the theoretical methods presented in this paper were developed to "cure" a specific problem occurring in a specific radar system. Admittedly, we lack experimental material to confirm that the instability of the quick estimator can occur in different bands or with different array configurations, but our computer simulations, reported in Section IV, suggest that the occurrence of the outliers is a generic behavior of estimator (11). At the very least, it does not seem unrealistic to suspect that the observed effects are not limited to the C-band and can also occur in the neighboring S and X bands.

III. Robustification of Quick Maximum Likelihood Estimator

To justify the adopted approach to eliminating the outliers, let us briefly speculate about the cause of their occurrence. It is well known that maximum likelihood estimators can be quite sensitive to the accuracy of the underlying model [18]. Algorithm (11) is no exception to this rule – when $\alpha_{\mathrm{d}} \approx 0$, the matrix $\boldsymbol{\Psi}_n^{\mathrm{H}}(\boldsymbol{\alpha})\boldsymbol{\Psi}_n(\boldsymbol{\alpha})$ becomes ill-conditioned. This observation may suggest that the observed gross errors are related to the modeling uncertainty – an unknown discrepancy between the simplified specular reflection model (10) and the

actual (true) pattern of wave propagation. In case of the quick estimator, the sources of model uncertainty include, among others:

- Array calibration errors, caused by finite manufacturing tolerances and differences between receive paths of the array. Calibration errors, though unavoidable in all real systems, are usually small, when compared with the next three sources of uncertainty.
- Approximation (9), which allows one to considerably reduce the model complexity, at the expense of making the assumed steering vector of the mirror signal different from the true one, even under ideal conditions.
- The implicit assumption about no curvature of the ground, which makes the flat earth model prone to factors such as imprecise leveling of the radar platform or a local slope of the terrain.
- When the ground surface is rough, i.e., when the condition (5) does not hold, the diffuse multipath may appear and introduce additional uncertainty [6]. Even though the terrain at the airport area is generally flat, one cannot fully exclude the possibility of the presence of diffuse multipath.

Computer simulations, reported in Section IV.B confirm the last hypothesis as particularly likely.

A. Problem formulation

We will now revisit estimator (11) to make it more robust against the model uncertainty. Consider modifying (11) using the minimax principle, i.e., consider the estimator of the form

$$\hat{\alpha}_{\mathrm{d}} = \arg\min_{\alpha_{\mathrm{d}}} \sum_{n=1}^{N} J_n^2(\alpha_{\mathrm{d}}, -\alpha_{\mathrm{d}}) \ , \tag{12}$$

where

$$\begin{aligned} J_n(\alpha_{\mathrm{d}}, \alpha_{\mathrm{m}}) &= \min_{A_{\mathrm{d}},\ A_{\mathrm{m}}} \max_{\substack{\tilde{\boldsymbol{a}}_{\mathrm{d},n} \\ \tilde{\boldsymbol{a}}_{\mathrm{m},n}}} \left\| \boldsymbol{y}_n - \tilde{\boldsymbol{\Psi}}_n \boldsymbol{\theta}_n \right\| \\ \tilde{\boldsymbol{\Psi}}_n &= [\tilde{\boldsymbol{a}}_{\mathrm{d},n} \ \tilde{\boldsymbol{a}}_{\mathrm{m},n}] \\ \boldsymbol{\theta}_n &= [A_{\mathrm{d},n} \ A_{\mathrm{m},n}]^{\mathrm{T}} \end{aligned} \tag{13}$$

$\tilde{\boldsymbol{a}}_{\mathrm{d},n}$, $\tilde{\boldsymbol{a}}_{\mathrm{m},n}$ are the perturbed steering vectors, and the "max" part is carried out subject to the constraints

$$\begin{aligned} \|\tilde{\boldsymbol{a}}_{\mathrm{d},n} - \boldsymbol{a}_n(\alpha_{\mathrm{d}})\|^2 &\le \varepsilon_{\mathrm{d}}^2 \\ \|\tilde{\boldsymbol{a}}_{\mathrm{m},n} - \boldsymbol{a}_n(\alpha_{\mathrm{m}})\|^2 &\le \varepsilon_{\mathrm{m}}^2 \ . \end{aligned} \tag{14}$$

The internal minimax fitting, which is the backbone of the proposed extension, is a variation of the robust least squares problem, discussed in [29], [30]. Note that, although similar in appearance, our formulation differs from the ones made in [29], [30]. Both [29] and [30] employed the technique of lumping the uncertainty into a single perturbation matrix,

$$\Delta\tilde{\Psi}_n(\alpha_{\mathrm{d}}, \alpha_{\mathrm{m}}) = \tilde{\boldsymbol{\Psi}}_n - \boldsymbol{\Psi}_n(\alpha_{\mathrm{d}}, \alpha_{\mathrm{m}}) \ ,$$

whose l_2-induced norm, also known as the spectral norm, i.e., its largest singular value, was bounded

$$\left\| \Delta\tilde{\Psi}_n(\alpha_{\mathrm{d}}, \alpha_{\mathrm{m}}) \right\|_2^2 \le \varepsilon^2 \ .$$

The disadvantage of the lumping approach lies in the fact that such unstructered perturbations do not reflect the actual properties of the problem. Indeed, the perturbations of the vector $\boldsymbol{a}(\alpha_{\mathrm{d}})$ are primarily related to quality of array calibration and should be regarded as smaller than the perturbations of the mirror steering vector, $\boldsymbol{a}(\alpha_{\mathrm{m}})$. Using our approach, one can express this specification by setting

$$\varepsilon_{\mathrm{d}}^2 < \varepsilon_{\mathrm{m}}^2 \ .$$

Under such setup, one can expect that the proposed estimator should pursue the direct signal more aggressively than the standard robust least squares method and, at the same time, exhibit more caution regarding the mirror signal.

B. Simplification of the minimax part

In the next two subsections we will focus exclusively on the minimax problem. Let us then switch, temporarily, to a simplified notation. We will drop the subscript n and employ the following shorthands

$$\begin{aligned} J &= J(\alpha_{\mathrm{d}}, \alpha_{\mathrm{m}}) \\ \boldsymbol{y} &= \boldsymbol{y}_n \\ \boldsymbol{a}_{\mathrm{d}} &= \boldsymbol{a}_n(\alpha_{\mathrm{d}}) \\ \boldsymbol{a}_{\mathrm{m}} &= \boldsymbol{a}_n(\alpha_{\mathrm{m}}) \\ \tilde{\boldsymbol{a}}_{\mathrm{d}} &= \tilde{\boldsymbol{a}}_{\mathrm{d},n} \\ \tilde{\boldsymbol{a}}_{\mathrm{m}} &= \tilde{\boldsymbol{a}}_{\mathrm{m},n} \\ \boldsymbol{\Psi} &= \boldsymbol{\Psi}_n(\alpha_{\mathrm{d}}, \alpha_{\mathrm{m}}) \ . \end{aligned}$$

The maximization task in (13) can be replaced with its closed form solution [29], [30]. Using the triangle inequality, one obtains

$$\begin{aligned} &\|\boldsymbol{y} - \tilde{\boldsymbol{a}}_{\mathrm{d}} A_{\mathrm{d}} - \tilde{\boldsymbol{a}}_{\mathrm{m}} A_{\mathrm{m}}\| \\ &\quad \le \nu(A_{\mathrm{d}}, A_{\mathrm{m}}) + \|\Delta\tilde{\boldsymbol{a}}_{\mathrm{d}} A_{\mathrm{d}}\| + \|\Delta\tilde{\boldsymbol{a}}_{\mathrm{m}} A_{\mathrm{m}}\| \\ &\quad = \nu(A_{\mathrm{d}}, A_{\mathrm{m}}) + \|\Delta\tilde{\boldsymbol{a}}_{\mathrm{d}}\| |A_{\mathrm{d}}| + \|\Delta\tilde{\boldsymbol{a}}_{\mathrm{m}}\| |A_{\mathrm{m}}| \ , \end{aligned} \tag{15}$$

where

$$\nu(A_{\mathrm{d}}, A_{\mathrm{m}}) = \|\boldsymbol{y} - \boldsymbol{a}_{\mathrm{d}} A_{\mathrm{d}} - \boldsymbol{a}_{\mathrm{m}} A_{\mathrm{m}}\| \tag{16}$$

and

$$\begin{aligned} \Delta\tilde{\boldsymbol{a}}_{\mathrm{d}} &= \tilde{\boldsymbol{a}}_{\mathrm{d}} - \boldsymbol{a}_{\mathrm{d}} \\ \Delta\tilde{\boldsymbol{a}}_{\mathrm{m}} &= \tilde{\boldsymbol{a}}_{\mathrm{m}} - \boldsymbol{a}_{\mathrm{m}} \end{aligned}$$

denote the steering vectors' perturbations.

The inequality in (15) becomes an equality when

$$\begin{aligned} \Delta\tilde{\boldsymbol{a}}_{\mathrm{d}} &= -c_{\mathrm{d}} A_{\mathrm{d}}^* [\boldsymbol{y} - \boldsymbol{a}_{\mathrm{d}} A_{\mathrm{d}} - \boldsymbol{a}_{\mathrm{m}} A_{\mathrm{m}}] \\ \Delta\tilde{\boldsymbol{a}}_{\mathrm{m}} &= -c_{\mathrm{m}} A_{\mathrm{m}}^* [\boldsymbol{y} - \boldsymbol{a}_{\mathrm{d}} A_{\mathrm{d}} - \boldsymbol{a}_{\mathrm{m}} A_{\mathrm{m}}] \ , \end{aligned} \tag{17}$$

where $c_{\mathrm{d}} \ge 0$, $c_{\mathrm{m}} \ge 0$ are arbitrary nonnegative factors and z^* denotes the complex conjugate of z. Choosing c_{d} and c_{m} so as to make both constraints in (14) active (i.e., met with equality) yields

$$\|\boldsymbol{y} - \tilde{\boldsymbol{a}}_{\mathrm{d}} A_{\mathrm{d}} - \tilde{\boldsymbol{a}}_{\mathrm{m}} A_{\mathrm{m}}\| = \nu(A_{\mathrm{d}}, A_{\mathrm{m}}) + \varepsilon_{\mathrm{d}} |A_{\mathrm{d}}| + \varepsilon_{\mathrm{m}} |A_{\mathrm{m}}|$$

which, when substituted into eq. (13), simplifies it to the following unconstrained minimization problem

$$J = \min_{A_{\mathrm{d}},\ A_{\mathrm{m}}} \nu(A_{\mathrm{d}}, A_{\mathrm{m}}) + \varepsilon_{\mathrm{d}} |A_{\mathrm{d}}| + \varepsilon_{\mathrm{m}} |A_{\mathrm{m}}| \ . \tag{18}$$

C. Analysis using duality theory

Observe that (18) is a convex optimization problem of two complex variables (i.e., four real variables). The objective function is, however, not differentiable when any of the following conditions takes place: $\nu(A_\mathrm{d}, A_\mathrm{m}) = 0$, $A_\mathrm{d} = 0$, $A_\mathrm{m} = 0$.

To gain more insight and, eventually, reach a simpler solution, we will reformulate (18) as the second order cone program [31]

$$\begin{aligned} \min\ & v + \varepsilon_\mathrm{d} s_\mathrm{d} + \varepsilon_\mathrm{m} s_\mathrm{m} \\ & \text{s.t.} \\ & \|\boldsymbol{y} - \boldsymbol{a}_\mathrm{d} A_\mathrm{d} - \boldsymbol{a}_\mathrm{m} A_\mathrm{m}\| \le v \\ & |A_\mathrm{d}| \le s_\mathrm{d} \\ & |A_\mathrm{m}| \le s_\mathrm{m} \end{aligned} \tag{19}$$

and employ the duality theory [32]. Technical details of our analysis are similar to the approach used in [29], although, due to differences in the problem formulation, not identical.

The dual of (19) reads [31]

$$\begin{aligned} \max\ & \boldsymbol{y}^\mathrm{H} \boldsymbol{u}_1 \\ & \text{s.t.} \\ & \|\boldsymbol{u}_1\| \le 1 \\ & |u_2| \le \varepsilon_\mathrm{d} \\ & |u_3| \le \varepsilon_\mathrm{m} \\ & \boldsymbol{a}_\mathrm{d}^\mathrm{H} \boldsymbol{u}_1 = u_2 \\ & \boldsymbol{a}_\mathrm{m}^\mathrm{H} \boldsymbol{u}_1 = u_3 \ . \end{aligned} \tag{20}$$

The first result explains the relationship between solutions of the primal and the dual problem.

Lemma 1: Denote by v_o, $s_\mathrm{d,o}$, $s_\mathrm{m,o}$, $A_\mathrm{d,o}$, $A_\mathrm{m,o}$ the solution of the primal problem. Similarly, let $\boldsymbol{u}_\mathrm{1,o}$, $u_\mathrm{2,o}$, $u_\mathrm{3,o}$ denote the solution of the dual problem. The following hold true

$$\begin{aligned} v_\mathrm{o} &= \|\boldsymbol{y} - \boldsymbol{a}_\mathrm{d} A_\mathrm{d,o} - \boldsymbol{a}_\mathrm{m} A_\mathrm{m,o}\| \\ s_\mathrm{d,o} &= |A_\mathrm{d,o}| \\ s_\mathrm{m,o} &= |A_\mathrm{m,o}| \\ \boldsymbol{u}_\mathrm{1,o} &= \frac{\boldsymbol{y} - \boldsymbol{a}_\mathrm{d} A_\mathrm{d,o} - \boldsymbol{a}_\mathrm{m} A_\mathrm{m,o}}{v_\mathrm{o}} \\ u_\mathrm{2,o} &= \varepsilon_\mathrm{d} \frac{A_\mathrm{d,o}}{|A_\mathrm{d,o}|} = \varepsilon_\mathrm{d} \frac{A_\mathrm{d,o}}{s_\mathrm{d,o}} \\ u_\mathrm{3,o} &= \varepsilon_\mathrm{m} \frac{A_\mathrm{m,o}}{|A_\mathrm{m,o}|} = \varepsilon_\mathrm{m} \frac{A_\mathrm{m,o}}{s_\mathrm{m,o}} \ . \end{aligned} \tag{21}$$

Proof: See Appendix A.

Using Lemma 1, one can show that the solution of the primal problem can be obtained by a nontrivial diagonal loading.

Lemma 2: The optimal values of complex amplitudes, $A_\mathrm{d,o}$, $A_\mathrm{m,o}$ satisfy the equation

$$[A_\mathrm{d,o}\ A_\mathrm{m,o}]^\mathrm{T} = \left[\boldsymbol{\Psi}^\mathrm{H}\boldsymbol{\Psi} + \boldsymbol{\Lambda}\right]^{-1} \boldsymbol{\Psi}^\mathrm{H}\boldsymbol{y} \ , \tag{22}$$

where

$$\boldsymbol{\Lambda} = \begin{bmatrix} \varepsilon_\mathrm{d} \frac{v_\mathrm{o}}{|A_\mathrm{d,o}|} & 0 \\ 0 & \varepsilon_\mathrm{m} \frac{v_\mathrm{o}}{|A_\mathrm{m,o}|} \end{bmatrix} . \tag{23}$$

Proof: See Appendix B. □

In principle, one could use Lemma 2 to solve the primal problem by finding the solution of eq. (22). This procedure is, however, not recommended. The unknown complex-valued quantities $A_\mathrm{d,o}$, $A_\mathrm{,o}$ appear in both sides of the equation, including v_o, which must satisfy the first equality in (21). It follows that, while using a generic nonlinear solver is certainly an option, it can hardly be expected to lead to a computationally attractive solution due to the substantial nonlinearity of equation (22). Proposition 1 points to a different, more efficient, approach

Proposition 1: The square of the primal objective can be found by minimizing the following, convex, function of two real variables

$$f(x_\mathrm{d}, x_\mathrm{m}) = \frac{\|\boldsymbol{y}\|^2 - \boldsymbol{y}^\mathrm{H}\boldsymbol{\Psi}\left[\boldsymbol{\Psi}^\mathrm{H}\boldsymbol{\Psi} + \boldsymbol{\Gamma}\right]^{-1}\boldsymbol{\Psi}^\mathrm{H}\boldsymbol{y}}{1 - x_\mathrm{d} - x_\mathrm{m}} \ , \tag{24}$$

where $x_\mathrm{d} > 0$, $x_\mathrm{m} > 0$, $1 - x_\mathrm{d} - x_\mathrm{m} > 0$ and

$$\boldsymbol{\Gamma} = \begin{bmatrix} \varepsilon_\mathrm{d}^2 \frac{1 - x_\mathrm{d} - x_\mathrm{m}}{x_\mathrm{d}} & 0 \\ 0 & \varepsilon_\mathrm{m}^2 \frac{1 - x_\mathrm{d} - x_\mathrm{m}}{x_\mathrm{m}} \end{bmatrix} . \tag{25}$$

Proof: See Appendix C.

D. The robustified estimator

Let us now look back and review the three, introduced so far, ways to solve the minimax problem (13).

The first option, summarized in eq. (18), involves unconstrained minimization of a convex function. While convexity certainly plays to one's advantage, the implementation using (18) is discouraged, primarily due to the fact that (18) is a function of four real variables (two complex variables). Additionally, (18) is not differentiable in every point of its domain, which may cause problems for the optimizers that require the function gradient.

The approach based on the second order cone programming (19) is even more computationally demanding. Its significance stems from the fact that the analysis of (19) and its dual (20) sheds a lot of light on the properties of minimax problem.

In Proposition 1, we showed that one may solve the minimax problem may in a computationally efficient manner by minimizing a convex and differentiable function of two real variables (24). Indeed, this is the recommended approach. The resulting, robustified version of the quick estimator is summarized below

$$\begin{aligned} \hat{\alpha}_\mathrm{d} &= \arg\min_{\alpha_\mathrm{d}} \sum_{n=1}^{N} J_n^2(\alpha_\mathrm{d}, -\alpha_\mathrm{d}) \\ J_n^2(\alpha_\mathrm{d}, \alpha_\mathrm{m}) &= \min_{x_\mathrm{d}, x_\mathrm{m}} f(x_\mathrm{d}, x_\mathrm{m}; \boldsymbol{y}_n, \alpha_\mathrm{d}, \alpha_\mathrm{m}) \ , \end{aligned} \tag{26}$$

where

$$f(x_\mathrm{d}, x_\mathrm{m}; \boldsymbol{y}_n, \alpha_\mathrm{d}, \alpha_\mathrm{m}) = \frac{\|\boldsymbol{y}_n\|^2 - \boldsymbol{y}_n^\mathrm{H}\boldsymbol{X}_n(\alpha_\mathrm{d}, \alpha_\mathrm{m}, x_\mathrm{d}, x_\mathrm{m})\boldsymbol{y}_n}{1 - x_\mathrm{d} - x_\mathrm{m}} \tag{27}$$

and

$$\boldsymbol{X}_n(\alpha_\mathrm{d},\alpha_\mathrm{m},x_\mathrm{d},x_\mathrm{m}) = \boldsymbol{\Psi}_n(\alpha_\mathrm{d},\alpha_\mathrm{m})\left[\boldsymbol{\Psi}_n(\alpha_\mathrm{d},\alpha_\mathrm{m})^\mathrm{H}\boldsymbol{\Psi}_n(\alpha_\mathrm{d},\alpha_\mathrm{m})+\boldsymbol{\Gamma}\right]^{-1}\boldsymbol{\Psi}_n^\mathrm{H}(\alpha_\mathrm{d},\alpha_\mathrm{m})$$
$$\boldsymbol{\Gamma} = \begin{bmatrix} \varepsilon_\mathrm{d}^2\frac{1-x_\mathrm{d}-x_\mathrm{m}}{x_\mathrm{d}} & 0 \\ 0 & \varepsilon_\mathrm{m}^2\frac{1-x_\mathrm{d}-x_\mathrm{m}}{x_\mathrm{m}} \end{bmatrix}. \quad (28)$$

Note that one can accelerate the minimization of $f(x_\mathrm{d},x_\mathrm{m};\boldsymbol{y}_n,\alpha_\mathrm{d},\alpha_\mathrm{m})$ by precomputing the terms independent of the variables x_d, x_m, i.e., $\|\boldsymbol{y}_n\|^2$, $\boldsymbol{\Psi}_n^\mathrm{H}(\alpha_\mathrm{d},\alpha_\mathrm{m})\boldsymbol{y}_n$ and $\boldsymbol{\Psi}_n^\mathrm{H}(\alpha_\mathrm{d},\alpha_\mathrm{m})\boldsymbol{\Psi}_n(\alpha_\mathrm{d},\alpha_\mathrm{m})$. Moreover, since the matrix under inversion is very small (2×2), it is best to simply hard-code the inverse, rather than employ elaborate matrix decompositions.

E. Lightweight alternatives

In this subsection we will discuss two, computationally attractive, alternatives to (26)-(28).

The first opportunity arises when one regards uncertainty of the steering vector $\tilde{\boldsymbol{a}}(\alpha_\mathrm{d})$ as negligible. It is summarized with the following proposition

Proposition 2: Consider the program (19) and set $\varepsilon_\mathrm{d}=0$. The square of the primal objective can be found by minimizing the following, convex function of one real variable

$$g(x) = \frac{\|\boldsymbol{y}\|^2 - \boldsymbol{y}^\mathrm{H}\boldsymbol{\Psi}\left[\boldsymbol{\Psi}^\mathrm{H}\boldsymbol{\Psi}+\boldsymbol{\Gamma}_g\right]^{-1}\boldsymbol{\Psi}^\mathrm{H}\boldsymbol{y}}{1-x} \quad (29)$$

where $0<x<1$ and

$$\boldsymbol{\Gamma}_g = \begin{bmatrix} 0 & 0 \\ 0 & \varepsilon_\mathrm{m}^2\frac{1-x}{x} \end{bmatrix}. \quad (30)$$

The resulting modifications to (26)-(28) are straightforward.

Reviewing Lemma 2 yields another option. Equations (22)-(23) suggest that it might be possible to find the optimal values of the complex amplitudes using the following iterative procedure

$$\begin{aligned} &\text{For } k=0,1,\ldots,K-1 \\ &[A_{\mathrm{d},k+1}\ A_{\mathrm{m},k+1}]^\mathrm{T} = \left[\boldsymbol{\Psi}^\mathrm{H}\boldsymbol{\Psi}+\boldsymbol{\Lambda}_k\right]^{-1}\boldsymbol{\Psi}^\mathrm{H}\boldsymbol{y} \\ &v_{k+1} = \|\boldsymbol{y}-\boldsymbol{a}_\mathrm{d}A_{\mathrm{d},k+1}-\boldsymbol{a}_\mathrm{m}A_{\mathrm{m},k+1}\| \\ &\boldsymbol{\Lambda}_{k+1} = \begin{bmatrix} \varepsilon_\mathrm{d}\frac{v_{k+1}}{|A_{\mathrm{d},k+1}|} & 0 \\ 0 & \varepsilon_\mathrm{m}\frac{v_{k+1}}{|A_{\mathrm{m},k+1}|} \end{bmatrix}. \end{aligned} \quad (31)$$

Note that (31) is, essentially, a variant of Miller's algorithm [33]. For simplicity, it can be started with $v_0=0$, $\boldsymbol{\Lambda}_0=\boldsymbol{0}$, which will yield the (penalized) Least Squares solution for $k=0$. Although this procedure is not guaranteed to converge to the optimal solution (for instance, observe that it will stall at the Least Squares solution whenever $\boldsymbol{y}$ happens to belong to the column span of $\boldsymbol{\Psi}$), we observed that it is extremely reliable in practice. Typically, iterating (31) three to five times is sufficient for the convergence to take place (11).

To obtain the resultant value of the cost function, one should substitute the amplitudes obtained in the final iteration into the following formula [c.f. (18)]

$$J^2 \simeq \left[v_K+\varepsilon_\mathrm{d}|A_{\mathrm{d,K}}|+\varepsilon_\mathrm{m}|A_{\mathrm{m,K}}|\right]^2.$$

The modifications of algorithm (26)-(28) needed to accommodate this approach are straightforward.

Despite the need to perform several iterations, the computational complexity of the estimator (31) is not far from the baseline quick estimator (11). One may asses the number of operations needed by the conventional approach by breaking it into elementary steps: computing $\boldsymbol{\Psi}^\mathrm{H}\boldsymbol{\Psi}$ ($4L$ complex multiplications, $4L-4$ complex additions), inverting $\boldsymbol{\Psi}^\mathrm{H}\boldsymbol{\Psi}$, which is a 2×2 matrix (6 complex multiplications, one complex division, one complex addition), computing $\boldsymbol{\Psi}^\mathrm{H}\boldsymbol{y}$ ($2L$ complex multiplications, $2L-2$ complex additions), computing $(\boldsymbol{\Psi}^\mathrm{H}\boldsymbol{\Psi})^{-1}\boldsymbol{\Psi}^\mathrm{H}\boldsymbol{y}$ (4 complex multiplications, 2 complex additions), and computing the squared norm of the residue ($3L$ complex multiplications, $3L-1$ complex additions). The total number of operations adds up to $9L+10$ complex multiplications, 1 complex division, and $9L-4$ complex additions.

Computational complexity of the iterative estimator scales nonlinearly with the number of iterations K. The cost of the first iteration ($k=0$) is almost the same as of the conventional estimator, and adds up to $9L+10$ complex multiplications, 1 complex division, $9L-4$ complex additions, and (additionally) one real square root. The remaining iterations, however, require only $3L+12$ complex multiplications, $3L+2$ complex additions, 1 complex division, 2 real multiplications, two real divisions, and three real square roots, because one does not need to compute $\boldsymbol{\Psi}^\mathrm{H}\boldsymbol{\Psi}$ and $\boldsymbol{\Psi}^\mathrm{H}\boldsymbol{y}$ again. The computation of J^2 can be implemented using 2 complex and 3 real multiplications, 2 real additions, and two real square roots. Given that modern CPUs typically execute multiplications, additions, divisions, and even square roots with comparable speed, and taking into account additional benefits from the cache hits at the second and later iterations, it is reasonable to estimate the cost of the second and subsequent iterations as one third of the first one. Overall, if we assume $K=5$ iterations, we can expect that the robustified estimator will require about $(1+(K-1)/3)=2.33$ times more time to execute that the conventional estimator.

F. Choice of design parameters

The parameters ε_d and ε_m specify the level of one's "trust" that the nominal steering vectors $\boldsymbol{a}_\mathrm{d}$, $\boldsymbol{a}_\mathrm{m}$ are accurate. Since the diffuse reflection is likely the primary source of the modeling errors, a general advice is to set $\varepsilon_\mathrm{m}>\varepsilon_\mathrm{d}$. Moreover, if ε_d or ε_m are set too close to zero, their influence on the behavior of the estimator will be marginal. On the other hand, if the uncertainty is set too high, the robustified estimator will become overly cautious. In particular, once $\varepsilon_\mathrm{m}>\|\boldsymbol{a}_\mathrm{m}\|$ ($\varepsilon_\mathrm{d}>\|\boldsymbol{a}_\mathrm{d}\|$) the optimal value of the amplitude $A_{\mathrm{d},n}$ ($A_{\mathrm{m},n}$) is zero, which means that the estimator looses the ability to estimate the corresponding component of the model. Based on our experience with the estimator, we recommended to set ε_d to a small value (possibly zero) and ε_m in the range of 5%-20% of the norm of $\boldsymbol{a}_\mathrm{m}$.

IV. Results

A. The case study revisited

The real-world datasets used in our case study (Figs. 4, 5) were processed using the proposed algorithms, i.e., algorithms

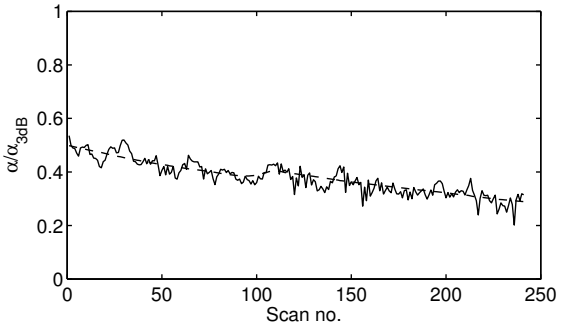


Figure 6. The comparison of the true values of the target elevation angle (dashed line) with its estimates yielded by the proposed robustified estimator (26)-(28) (solid line).

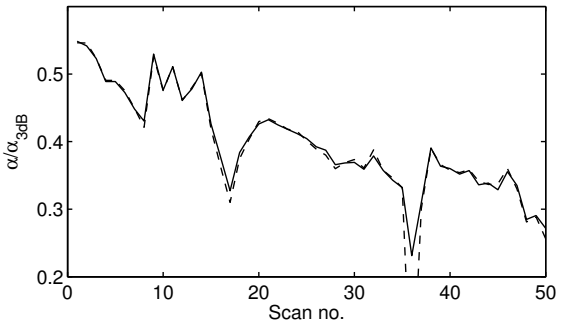


Figure 7. The comparison of the estimates yielded by algorithms (11) (dashed line) and (31) (solid line) for the second dataset. Note that, compared with Fig. 5, the vertical scale was changed to improve the plot's readability.

(26)-(28), (29), and (31). The adopted levels of uncertainty were $\epsilon_{\rm d} = 0.0125$ and $\epsilon_{\rm m} = 0.5$, respectively. Since, for the datasets in question, it holds that the norm $\|\boldsymbol{a}(\alpha)\|$ stays close to 4 for all angles of interest (recall that the radar employs the beamspace processing, which means that the norm of the steering vector is a function of angle), the adopted value of $\epsilon_{\rm m}$ corresponds to, approximately, 12.5% relative uncertainty of the mirror steering vector. Finally, for the iterative algorithm (31), we employed $K = 5$ iterations.

Fig. 6 shows the results obtained using algorithm (26)-(28). The results obtained using the other two algorithms were nearly identical, and therefore are not shown. Observe that the outlier that appeared when the conventional algorithm (11) was used, is now gone. The same effect can be observed for the second dataset, which is shown in Fig. 7. Moreover, Fig. 7 shows that the influence of the proposed robustification is practically limited to preventing the outliers from occurring – when the conventional estimator (11) works correctly, the robustified approach yields very similar estimates, which was precisely our goal.

Table I shows the values of mean squared errors of normalized elevation angle estimates, obtained using all algorithms discussed so far, for the two real-world datasets. Moreover, to make our case study more comprehensive, we also included the stochastic maximum likelihood estimator [17], [18]. For both datasets, the best results were obtained using algorithms (26)-(28) and (31). Note that, the equal performance of both algorithms is not a coincidence because the iterative procedure employed by estimator (31) is intended to converge to the solution of (26)-(28).

Summarizing this part of the discussion, we recommend using the iterative solution, because it has the smallest computational complexity. For $K = 5$, which we regard as sufficient, our MATLAB implementation of the algorithm is only about 2.5 times slower than the quick estimator (11), which agrees well with the estimate of the computational complexity presented in Section III.E.

Algorithm	Dataset 1	Dataset 2
Monopulse	$6.10 \cdot 10^{-3}$ (0%)	$1.43 \cdot 10^{-2}$ (0%)
ML (8)	$1.59 \cdot 10^{-2}$ (-161%)	$4.07 \cdot 10^{-2}$ (-184%)
Quick ML (11)	$1.54 \cdot 10^{-3}$ (75%)	$4.12 \cdot 10^{-3}$ (71%)
Algorithm (26)-(28)	$7.35 \cdot 10^{-4}$ (88%)	$1.16 \cdot 10^{-3}$ (92%)
Algorithm (29)	$7.37 \cdot 10^{-4}$ (88%)	$1.17 \cdot 10^{-3}$ (92%)
Algorithm (31)	$7.35 \cdot 10^{-4}$ (88%)	$1.16 \cdot 10^{-3}$ (92%)
Stochastic ML	$9.08 \cdot 10^{-4}$ (85%)	$1.54 \cdot 10^{-3}$ (90%)

Table I
THE COMPARISON OF THE MEAN SQUARED ERRORS OF THE NORMALIZED ELEVATION ESTIMATES YIELDED BY SEVERAL ALGORITHMS FOR THE TWO REAL-WORLD DATASETS. THE VALUES IN THE PARENTHESES SHOW THE RELATIVE REDUCTION OF THE MSE OVER THE MONOPULSE METHOD.

B. Results of computer simulations

To investigate the properties of the proposed method further, one may employ computer simulations. To this end, we used a realistic model that includes the specular and the diffuse reflection, adopted from [6]. In this model, the ground surface is treated as a "glistering surface", and the main factors influencing the amounts of the specular and the diffuse scattering are the grazing angle, the surface roughness parameter, defined as the ratio of the RMS surface height variation σ_h to the wavelength λ, σ_h/λ, and the maximum surface facet slope β_0.

Following [6], we assumed that $\beta_0 = 0.1$ radians and a simulated a system that employs the 16-element standard (half-wavelength element spacing) uniform linear array, elevated to the height of $h = 5$ meters, whose boresight was parallel to the ground. The direct signal originated from a simulated target, placed at range $R = 10$ km. The signal to noise ratio of the direct signal, i.e., the ratio $|A_{\rm d}|^2/\sigma_v^2$, was set to 15 dB.

The specular reflection was inserted at the elevation given by eq. (2), with its amplitude computed using [6, eq. (5)]. To simulate the diffuse multipath, we introduced additional reflections every 0.25°. The phases of these reflections were random in the interval $[0, 2\pi)$, and their magnitudes were obtained from the power distribution of the diffuse scattering, computed using [6, eq. (13)-(17)]. Finally, we modeled the presence of a vegetation by introducing 6 dB attenuation of all reflections.

We assumed that only one snapshot is available ($N = 1$), and compared four algorithms: the monopulse estimator, the quick conditional maximum likelihood estimator (11), the proposed iterative estimator (31), and the Capon method.

In the case of the first three methods, the data was preprocessed using a beamformer – the simulated system synthesized two pairs of sum-difference beams (i.e., $L = 4$) using full 16-

element aperture. The first sum-difference pair was positioned at 4 degrees above the horizon, and the second pair – at 8 degrees. The sum beams employed the Chebyshev taper, and the difference beams employed the Bayliss taper, both with sidelobes at -30 dB. The resultant 3dB beamwidth of the sum beams is about 8°. Moreover, after some preliminary tuning of the iterative estimator, we decided to use ϵ_{d} and ϵ_{m} that correspond to 2.5% and 10% relative uncertainty, respectively, and $K = 5$ iterations.

In the case of the Capon method, we used the forward-backward smoothing technique [15] to decorrelate the direct and the mirror signals, and to form a nonsingular correlation matrix. The forward-backward smoothing method is based on dividing the array into a number of smaller overlapping subarrays, from which independent snapshots are obtained [15]. To keep the computing time of all estimators at a comparable level, and to satisfy the well-known rule of thumb that says that the number of independent snapshots used to form the covariance matrix should be at least twice the array size, we used 6-element subarrays.

Since the benefits of the proposed estimators are more of qualitative, rather than quantitative, nature, we compare the histograms of the estimates of the target elevation angle obtained using each method. Fig 8 shows such histograms, computed from 10000 Monte Carlo trials, for the target elevation equal to 2° (1/4 beamwidth), and for three choices of the surface roughness factor, $\sigma_h/\lambda \in \{0,\ 1,\ 5\}$. The first choice corresponds to the perfectly smooth ground, in which case only the specular reflection is present. The second choice represents a "realistically" smooth ground, where the specular reflection dominates, but small amounts of the diffuse multipath can be expected. Finally, the surface roughness factor equal to 5 corresponds to a moderately rough ground, when there is a substantial amount of the diffuse multipath, and the specular reflection occurs only for targets at elevation angles below 3°.

In the case of the perfectly smooth ground, the monopulse method and the Capon method both exhibit a significant bias, while the quick estimator and the proposed method can resolve the direct and the mirror signals properly. When the surface roughness factor is increased to one, the spread of all histograms increases due to the presence of the diffuse multipath. More importantly, however, a peak emerges at zero elevation angle for the quick maximum likelihood estimator, which shows that about 3% of the estimates are outliers similar to those observed in the real system. The robustified estimator does not exhibit such behavior. When the surface roughness factor is increased even further, the fraction of outliers grows to about 7%, while the proposed approach continues to deliver stable estimates. In the case of the monopulse and the Capon methods, increasing the roughness factor results in the reduction in their bias, particularly for $\sigma_h/\lambda = 5$. However, the histograms are still wider than in the case of the proposed method, which means that the MSE of these estimators is larger.

We repeated this simulation experiment for several additional values of the target elevation angle $\alpha_{\mathrm{d}} \in [1^\circ,\ 4^\circ]$, $\sigma_h/\lambda \in \{1, 5, 20\}$, and estimated the probability of estimator

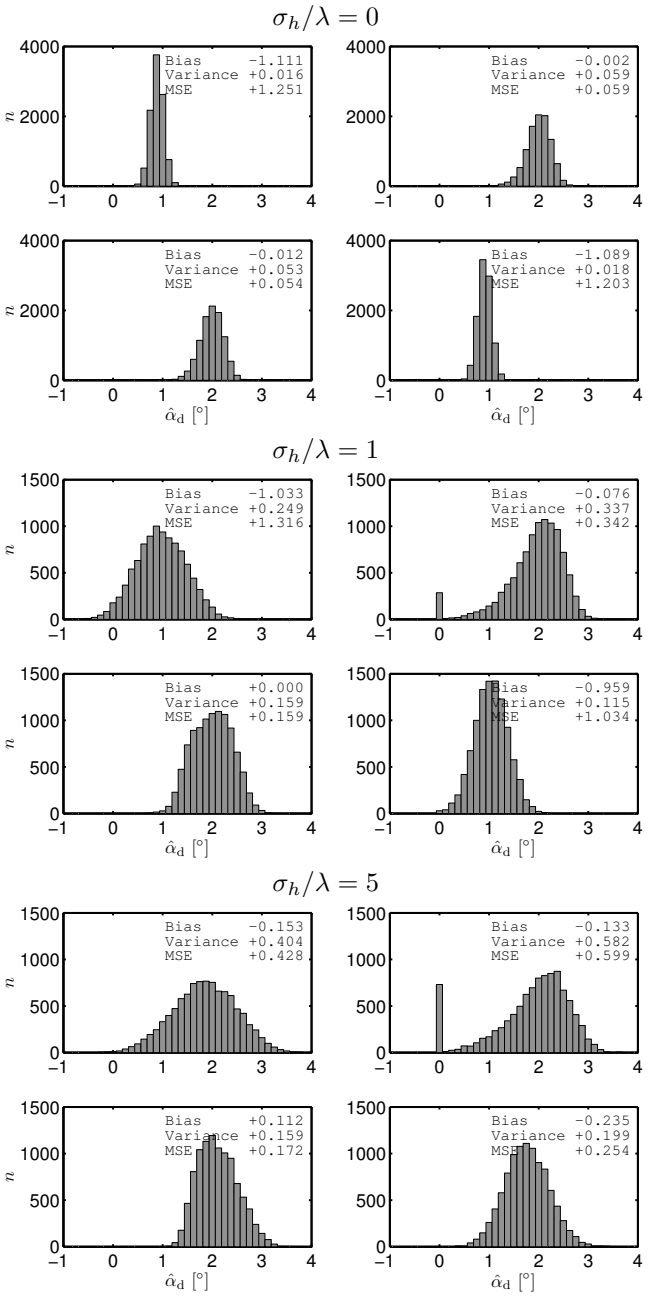


Figure 8. The comparison of the histograms of the estimates of the elevation angle obtained using the monopulse method (top-left), the quick maximum likelihood estimator (top-right), the proposed iterative algorithm (bottom left), and the Capon method (bottom right) for a target at 2° and three values of the surface roughness factor σ_h/λ.

failure by counting the number of estimates that fell below 0.1°. The results, depicted in Fig. 9, show that, in case of the nonrobustified estimator, the outliers can exceed 30% at 1° elevation angle (1/8 beamwidth), which is a very high level indeed. Increasing the elevation angle causes a gradual improvement in the estimator behavior, until it becomes practically error-free at 3°. In case of the robustified estimator, we observed that it completely avoids the outliers for elevation angles above 1.25°. For smaller angles, the contribution of outliers was marginal (below 0.2% at 1°), which confirms that the proposed technique improves the stability of the

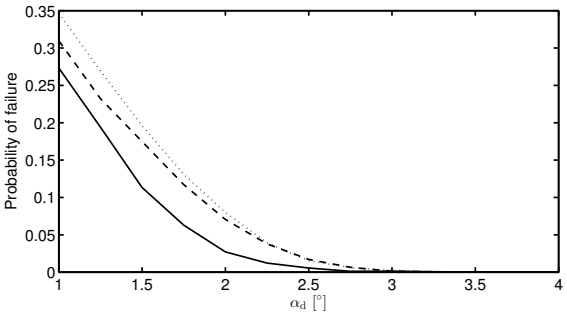


Figure 9. The probability of the conventional estimator failing as a function of the target elevation angle for roughness factor equal to 1 (solid line), 5 (dashed line), and 20 (dotted line).

estimation greatly. Finally, the monopulse and the Capon method exhibited the outlier-free behavior, but their MSE was consistently greater than for the proposed method.

V. Conclusions

We proposed a novel robustification of the conditional maximum-likelihood estimator. In the proposed solution, the sensitivity to the modeling errors is reduced by treating the steering vectors as subject to unknown bounded perturbations. Combined with the specular reflection model, the proposed robustification technique yields a computationally attractive estimator of the elevation angle. The results, obtained using real-world dataset and simulations, confirm that the proposed solution exhibits improved behavior under the presence of the modeling uncertainty.

Appendix A
Proof of Lemma 1

The first three relationships in (21) are trivial, given the form of the primal problem. To prove the remaining three, first note that the strong duality holds [32], i.e., the primal and the dual objectives are equal

$$v_{\rm o} + \varepsilon_{\rm d} s_{\rm d,o} + \varepsilon_{\rm m} s_{\rm m,o} = \boldsymbol{y}^{\rm H} \boldsymbol{u}_{1,{\rm o}} \; .$$

Rearranging the right-hand side as

$$\begin{aligned} \boldsymbol{y}^{\rm H} \boldsymbol{u}_{1,{\rm o}} = &\left[\boldsymbol{y} - \boldsymbol{a}_{\rm d} A_{\rm d,o} - \boldsymbol{a}_{\rm m} A_{\rm m,o}\right]^{\rm H} \boldsymbol{u}_{1,{\rm o}} \\ &+ A_{\rm d,o}^{*} \boldsymbol{a}_{\rm d}^{\rm H} \boldsymbol{u}_{1,{\rm o}} + A_{\rm m,o}^{*} \boldsymbol{a}_{\rm m}^{\rm H} \boldsymbol{u}_{1,{\rm o}} \; , \end{aligned}$$

and substituting the last two constraints from (20) lead to

$$\begin{aligned} v_{\rm o} + \varepsilon_{\rm d} s_{\rm d,o} + \varepsilon_{\rm m} s_{\rm m,o} = &\left[\boldsymbol{y} - \boldsymbol{a}_{\rm d} A_{\rm d,o} - \boldsymbol{a}_{\rm m} A_{\rm m,o}\right]^{\rm H} \boldsymbol{u}_{1,{\rm o}} \\ &+ A_{\rm d,o}^{*} u_{2,{\rm o}} + A_{\rm m,o}^{*} u_{3,{\rm o}} \; . \end{aligned}$$

Recall that the dual objective, i.e., the right hand side, is obtained by performing the maximization subject to the constraints in (20). Taking this into account yields

$$\begin{aligned} \boldsymbol{u}_{1,{\rm o}} &= \frac{\boldsymbol{y} - \boldsymbol{a}_{\rm d} A_{\rm d,o} - \boldsymbol{a}_{\rm m} A_{\rm m,o}}{\left\| \boldsymbol{y} - \boldsymbol{a}_{\rm d} A_{\rm d,o} - \boldsymbol{a}_{\rm m} A_{\rm m,o} \right\|} \\ &= \frac{\boldsymbol{y} - \boldsymbol{a}_{\rm d} A_{\rm d,o} - \boldsymbol{a}_{\rm m} A_{\rm m,o}}{v_{\rm o}} \\ u_{2,{\rm o}} &= \varepsilon_{\rm d} \frac{A_{\rm d,o}}{|A_{\rm d,o}|} = \varepsilon_{\rm d} \frac{A_{\rm d,o}}{|s_{\rm d,o}|} \\ u_{3,{\rm o}} &= \varepsilon_{\rm m} \frac{A_{\rm m,o}}{|A_{\rm m,o}|} = \varepsilon_{\rm m} \frac{A_{\rm m,o}}{|s_{\rm m,o}|} \; , \end{aligned} \tag{32}$$

which completes the proof. □

Appendix B
Proof of Lemma 2

Premultiplying both sides of $\boldsymbol{u}_{1,{\rm o}}$ from (21) with $v_{\rm o} \boldsymbol{a}_{\rm d}^{\rm H}$ yields [c.f. (20)]

$$u_{2,{\rm o}} v_{\rm o} = \boldsymbol{a}_{\rm d}^{\rm H} \boldsymbol{y} - \boldsymbol{a}_{\rm d}^{\rm H} \boldsymbol{a}_{\rm d} A_{\rm d,o} - \boldsymbol{a}_{\rm d}^{\rm H} \boldsymbol{a}_{\rm m} A_{\rm m,o} \; .$$

Substituting $u_{2,{\rm o}}$ from (21), after some minor rearranging, leads to

$$\boldsymbol{a}_{\rm d}^{\rm H} \boldsymbol{a}_{\rm d} A_{\rm d,o} + \varepsilon_{\rm d} \frac{A_{\rm d,o}}{s_{\rm d,o}} v_{\rm o} + \boldsymbol{a}_{\rm d}^{\rm H} \boldsymbol{a}_{\rm m} A_{\rm m,o} = \boldsymbol{a}_{\rm d}^{\rm H} \boldsymbol{y} \; . \tag{33}$$

Repeating the reasoning, but starting with pre-multiplication by $v_{\rm o} \boldsymbol{a}_{\rm m}^{\rm H}$, results in

$$\boldsymbol{a}_{\rm m}^{\rm H} \boldsymbol{a}_{\rm d} A_{\rm d,o} + \boldsymbol{a}_{\rm m}^{\rm H} \boldsymbol{a}_{\rm m} A_{\rm m,o} + \varepsilon_{\rm m} \frac{A_{\rm m,o}}{s_{\rm m,o}} v_{\rm o} = \boldsymbol{a}_{\rm m}^{\rm H} \boldsymbol{y} \; . \tag{34}$$

To obtain (22)-(23), first rewrite equations (33) and (34) using the matrix notation

$$\begin{bmatrix} \boldsymbol{a}_{\rm d}^{\rm H} \boldsymbol{a}_{\rm d} + \frac{\epsilon_{\rm d}}{s_{\rm d,o}} v_{\rm o} & \boldsymbol{a}_{\rm d}^{\rm H} \boldsymbol{a}_{\rm m} \\ \boldsymbol{a}_{\rm m}^{\rm H} \boldsymbol{a}_{\rm d} & \boldsymbol{a}_{\rm m}^{\rm H} \boldsymbol{a}_{\rm m} + \frac{\epsilon_{\rm m}}{s_{\rm m,o}} v_{\rm o} \end{bmatrix} \begin{bmatrix} A_{\rm d,o} \\ A_{\rm m,o} \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}_{\rm d}^{\rm H} \boldsymbol{y} \\ \boldsymbol{a}_{\rm m}^{\rm H} \boldsymbol{y} \end{bmatrix} \; .$$

It is straightforward to verify that the matrix on the left equals $\boldsymbol{\Psi}^{\rm H} \boldsymbol{\Psi} + \boldsymbol{\Lambda}$, where $\boldsymbol{\Lambda}$ is defined in (23), and that the vector on the right equals $\boldsymbol{\Psi}^{\rm H} \boldsymbol{y}$. These observations lead to the following formula

$$\left[\boldsymbol{\Psi}^{\rm H} \boldsymbol{\Psi} + \boldsymbol{\Lambda}\right] \left[A_{\rm d,o} \; A_{\rm m,o}\right]^{\rm T} = \boldsymbol{\Psi}^{\rm H} \boldsymbol{y} \; ,$$

which is just one, trivial step from (22)-(23). □

Appendix C
Proof of Proposition 1

Combining the first equation from (20) with the fourth from (21) yields

$$J = \boldsymbol{y}^{\rm H} \boldsymbol{u}_{1,{\rm o}} = \boldsymbol{y}^{\rm H} \frac{\boldsymbol{y} - \boldsymbol{a}_{\rm d} A_{\rm d,o} - \boldsymbol{a}_{\rm m} A_{\rm m,o}}{v_{\rm o}} \; . \tag{35}$$

Observe that $v_{\rm o}$ can be expressed as [c.f. (19)]

$$v_{\rm o} = J - \varepsilon_{\rm d} s_{\rm d,o} - \varepsilon_{\rm m} s_{\rm m,o} \tag{36}$$

and replace $v_{\rm o}$ in (35) with (36) to obtain

$$J = \boldsymbol{y}^{\rm H} \frac{\boldsymbol{y} - \boldsymbol{a}_{\rm d} A_{\rm d,o} - \boldsymbol{a}_{\rm m} A_{\rm m,o}}{J - \varepsilon_{\rm d} s_{\rm d,o} - \varepsilon_{\rm m} s_{\rm m,o}} \; .$$

Now substitute (22), which leads to

$$J = \frac{\|\boldsymbol{y}\|^2 - \boldsymbol{y}^{\mathrm{H}}\boldsymbol{\Psi}(\boldsymbol{\Psi}^{\mathrm{H}}\boldsymbol{\Psi} + \boldsymbol{\Lambda})^{-1}\boldsymbol{\Psi}^{\mathrm{H}}\boldsymbol{y}}{J - \varepsilon_{\mathrm{d}} s_{\mathrm{d,o}} - \varepsilon_{\mathrm{m}} s_{\mathrm{m,o}}} \,. \tag{37}$$

Moreover, using (36) in (23) yields the following alternative expression for $\boldsymbol{\Lambda}$ [c.f. (21)]

$$\boldsymbol{\Lambda} = \begin{bmatrix} \varepsilon_{\mathrm{d}} \frac{J - \varepsilon_{\mathrm{d}} s_{\mathrm{d,o}} - \varepsilon_{\mathrm{m}} s_{\mathrm{m,o}}}{s_{\mathrm{d,o}}} & 0 \\ 0 & \varepsilon_{\mathrm{m}} \frac{J - \varepsilon_{\mathrm{d}} s_{\mathrm{d,o}} - \varepsilon_{\mathrm{m}} s_{\mathrm{m,o}}}{s_{\mathrm{m,o}}} \end{bmatrix} .$$

Introducing two new variables

$$x_{\mathrm{d}} = \epsilon_{\mathrm{d}} \frac{s_{\mathrm{d,o}}}{J} \qquad x_{\mathrm{m}} = \epsilon_{\mathrm{m}} \frac{s_{\mathrm{m,o}}}{J}$$

one may observe that $\boldsymbol{\Lambda} = \boldsymbol{\Gamma}$. Finally, multiplying both sides of (37) by J leads to

$$J^2 = f(x_{\mathrm{d}}, x_{\mathrm{m}}) = \frac{\|\boldsymbol{y}\|^2 - \boldsymbol{y}^{\mathrm{H}}\boldsymbol{\Psi}\left[\boldsymbol{\Psi}^{\mathrm{H}}\boldsymbol{\Psi} + \boldsymbol{\Gamma}\right]^{-1}\boldsymbol{\Psi}^{\mathrm{H}}\boldsymbol{y}}{1 - x_{\mathrm{d}} - x_{\mathrm{m}}} \,, \tag{38}$$

i.e., the minimization (24) allows one to find the square of the primal objective.

To show that (24) is a convex function, one can employ the matrix inversion lemma

$$\boldsymbol{I} - \boldsymbol{\Psi}(\boldsymbol{\Psi}^{\mathrm{H}}\boldsymbol{\Psi} + \boldsymbol{\Lambda})^{-1}\boldsymbol{\Psi} = (\boldsymbol{I} + \boldsymbol{\Psi}\boldsymbol{\Lambda}^{-1}\boldsymbol{\Psi}^{\mathrm{H}})^{-1} \,,$$

where $\boldsymbol{I}$ denotes the eye matrix of the appropriate dimension. After some minor manipulations, one arrives at

$$f(x_{\mathrm{d}}, x_{\mathrm{m}}) = \boldsymbol{y}^{\mathrm{H}}\left[(1 - x_{\mathrm{d}} - x_{\mathrm{m}})\,\boldsymbol{I} + \boldsymbol{\Psi}\boldsymbol{\Delta}\boldsymbol{\Psi}^{\mathrm{H}}\right]^{-1}\boldsymbol{y} \,, \tag{39}$$

where

$$\boldsymbol{\Delta} = \begin{bmatrix} \frac{x_{\mathrm{d}}}{\varepsilon_{\mathrm{d}}^2} & 0 \\ 0 & \frac{x_{\mathrm{m}}}{\varepsilon_{\mathrm{m}}^2} \end{bmatrix} ,$$

which shows that, for $x_{\mathrm{d}} \geq 0$, $x_{\mathrm{m}} \geq 0$, $1 - x_{\mathrm{d}} - x_{\mathrm{m}} \geq 0$, $f(x_{\mathrm{d}}, x_{\mathrm{m}})$ is a composition of convex mappings, i.e., it is a convex function [32]. □

REFERENCES

[1] M. Skolnik, *Radar Handbook*. McGraw-Hill, 2008.

[2] D. K. Barton, *Radar System Analysis and Modeling*. Artech House, Inc., 2005.

[3] M. Skolnik, *Introduction to Radar Systems*. McGraw-Hill, 2002.

[4] U. Nickel, "Overview of generalized monopulse estimation," *IEEE Aerospace and Electronic Systems Magazine*, vol. 21, no. 6, pp. 27–56, 2006.

[5] W. D. White, "Low-angle radar tracking in the presence of multipath," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-10, no. 6, pp. 835–852, 1974.

[6] D. K. Barton, "Low-angle radar tracking," *Proceedings of the IEEE*, vol. 62, no. 6, pp. 687–704, 1974.

[7] V. Mangulis, "Frequency diversity in low-angle radar tracking," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-17, no. 1, pp. 149–153, 1981.

[8] S. M. Sherman, "Complex indicated angles applied to unresolved radar targets and multipath," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-7, no. 1, pp. 160–170, 1971.

[9] J. Capon, "High resolution frequency-wavenumber spectrum analysis," *Proceedings of the IEEE*, vol. 57, pp. 1408–1418, 1969.

[10] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276–280, 1986.

[11] V. V. Reddy, M. Mubeen, and B. P. Ng, "Reduced-complexity super-resolution DOA estimation with unknown number of sources," *IEEE Signal Processing Letters*, vol. 22, no. 6, pp. 772–776, 2015.

[12] C. Niu, Y. Zhang, and J. Guo, "Low angle estimation in diffuse multipath environment by time-reversal minimum-norm-like technique," *IET Radar, Sonar Navigation*, vol. 11, no. 10, pp. 1483–1487, 2017.

[13] T.-J. Shan, M. Wax, and T. Kailath, "On spatial smoothing for direction-of-arrival estimation of coherent signals," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 33, no. 4, pp. 806–811, 1985.

[14] K. G. A. Thakre and M. Haardt, "Single snapshot spatial smoothing with improved effective array aperture," *IEEE Signal Processing Letters*, vol. 16, no. 6, pp. 505–508, 2009.

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

[16] ——, *Detection, Estimation and Modulation Theory, Part I: Detection, Estimation, and Filtering Theory*. John Wiley & Sons, 2013.

[17] 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)*, 1988, pp. 2893–2896.

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

[19] I. Ziskind and M. Wax, "Maximum likelihood localization of multiple sources by alternating projection," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 36, no. 10, pp. 1553–1560, 1988.

[20] A. J. Weiss, A. S. Willsky, and B. C. Levy, "Maximum likelihood array processing for the estimation of superimposed signals," *Proceedings of the IEEE*, vol. 76, no. 2, pp. 203–204, 1988.

[21] Y. Bresler and A. Makovski, "Exact maximum likelihood parameter estimation of superimposed exponential signals in noise," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 34, no. 5, pp. 307–310, 1986.

[22] M. D. Zoltowski and T. S. Lee, "Maximum likelihood based sensor array signal processing in the beamspace domain for low angle radar tracking," *IEEE Transactions on Signal Processing*, vol. 39, no. 3, pp. 656–671, 1991.

[23] F. Vincent, O. Besson, and E. Chaumette, "Approximate maximum likelihood estimation of two closely spaced sources," *Signal Processing*, vol. 97, no. Supplement C, pp. 83–90, 2014.

[24] ——, "Approximate unconditional maximum likelihood direction of arrival estimation for two closely spaced targets," *IEEE Signal Processing Letters*, vol. 22, no. 1, pp. 86–89, 2015.

[25] 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*, vol. 11, no. 5, pp. 797–801, 2015.

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

[27] D. Park, E. Yang, S. Ahn, and J. Chun, "Adaptive beamforming for low-angle target tracking under multipath interference," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 50, no. 4, pp. 2564–2577, 2014.

[28] M. A. Sebt, A. Sheikhi, and M. M. Nayebi, "Robust low-angle estimation by an array radar," *IET Radar, Sonar Navigation*, vol. 4, no. 6, pp. 780–790, 2010.

[29] L. E. Ghaoui and H. Lebret, "Robust solutions to least-squares problems with uncertain data," *SIAM Journal on Matrix Analysis and Applications*, vol. 18, no. 4, pp. 1035–1064, 1997.

[30] S. Chandrasekaran, G. H. Golub, M. Gu, and A. H. Sayed, "Parameter estimation in the presence of bounded data uncertainties," *SIAM Journal on Matrix Analysis and Applications*, vol. 19, no. 1, pp. 235–252, 1998.

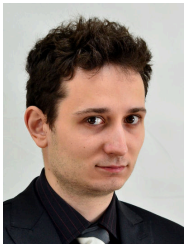
[31] M. S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret, "Applications of second-order cone programming," *Linear Algebra and its Applications*, vol. 284, pp. 193–228, 1998.

[32] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[33] K. Miller, "Least squares methods for ill-posed problems with a prescribed bound," *SIAM Journal on Mathematical Analysis*, vol. 1, no. 1, pp. 52–74, 1970.

Michał Meller was born in Suwałki, Poland in 1983. He received the M.Sc. and Ph.D. degrees in automatic control from the Gdańsk University of Technology in 2007 and 2010, respectively, and the Dr.Hab. (D.Sc.) degree from the Warsaw University of Technology in 2017. He holds a position of an associate professor at the Department of Automatic Control, Faculty of Electronics, Telecommunications and Computer Science, Gdańsk University of Technology. He also works for PIT-RADWAR S.A., currently as an analyst of digital systems. In 2018, in recognition of his outstanding industrial achievements, he received the bronze medal "for the contributions to the defense of the country."

His research interests include adaptive signal processing, adaptive control, statistical signal processing, estimation, and system identification. He has extensive experience in signal processing for radar applications, including noise, FMCW, and AESA pulse-Doppler radars.



Kamil Stawiarski received the M.Sc. degree in automatic control from the Gdańsk University of Technology in 2015, where he is currently pursuing Ph.D. degree. Since 2015, he has been working as a software developer at PIT-RADWAR S.A. His research interests are in the areas of signal processing, especially detection and estimation.