Experimental Evaluation of Estimator Mean Square Error Curve for Cognitive Tracking Radar

Michał Meller

Department of Automatic Control, Faculty of Electronics, Telecommunications and Computer Science,
Gdańsk University of Technology, ul. Narutowicza 11/12, 80-233 Gdańsk, POLAND
PIT-RADWAR S.A., ul. Hallera 233A, 80-502 Gdańsk, POLAND
e-mail: michal.meller@eti.pg.gda.pl, michal.meller@pitradwar.com

***Abstract*—To make decisions, cognitive radar must rely on predictions of its own performance. In the literature, these predictions are usually based on some form of Cramér-Rao lower bound. This approach is scientifically sound, but it also brings a possibility of the cognitive controller overestimating radar performance. It therefore makes sense to back theoretical predictions with careful experiments which will verify their applicability. Using a simple direction of arrival estimator as an example we discuss how an experimental performance curve can be obtained. We also propose an extended approach, which employs experimental data to estimate parameters of a performance curve which includes floor and threshold effects.**

***Index Terms*—cognitive radar, performance bounds, direction of arrival estimation**

I. Introduction

Cognitive radar, introduced by Guerci [1] and Haykin [2], is a novel approach to radar management which emphasizes information flow from the receiver to the transmitter via feedback. The concept was refined and formalized by Bell and coworkers in [3]. The feedback information, in the form of posterior probability density function of target location, is employed to make decisions about future radar actions (Fig. 1). To make these decisions, rather than relying on a set of heuristic rules, the cognitive controller addresses a certain optimization problem, where the cost function is set so as to reflect goals specific to radar mission and mode of operation and the optimization is carried out subject to a set of constraints.

Thus, at the very heart of cognitive radar lies the issue of predicting system performance accurately and reliably. If predictions are pessimistic, the radar will waste resources unnecessarily. On the other hand, optimistic predictions are equally, if not more, unwanted, because relying on them can result in the radar failing to meet its goals.

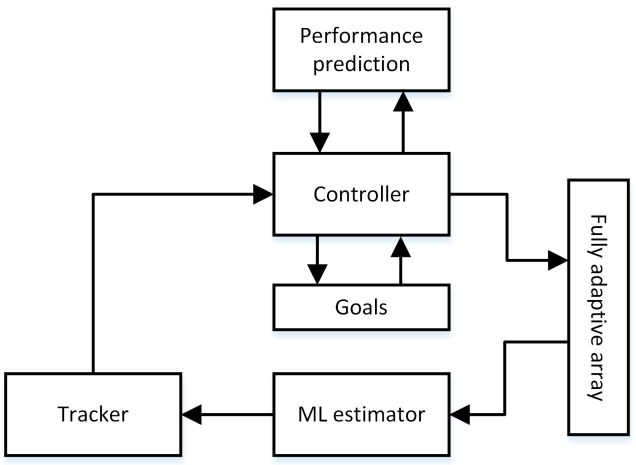

Figure 1. Block diagram of cognitive radar system.

The dominant approach to predicting cognitive radar performance in tracking applications seems to be based on analyzing Cramér-Rao lower bound and its relatives. For instance, both Haykin and Bell employed Bayesian bounds [4] in their solution of tracking problem [2], [3], while in [5] modified Miller-Chang bound was used. Such an approach is both elegant and scientifically sound, because analysis of many estimators shows that they are able to reach these bounds in medium and high signal to noise (SNR) range [4].

One must not forget however, that these bounds are *absolute*, i.e. no system can exceed (reach below) performance limitations set by these bounds. In fact, actual systems always perform worse. In classical radar design this discrepancy is well recognized and accounted for in terms of so-called SNR loss [6].

It follows that during development of cognitive radar it would make sense to evaluate its actual performance experimentally and employ realistic performance curves in the predictor. Such a curve would also be of interest for traditional radar design, because it could be used to assess quality of its subcomponents, including both hardware and software.

The main contribution of this paper is a procedure for obtaining estimator mean square error curves using experimental data. The basic variant of the method, drawn from analysis of general behavior of Cramér-Rao lower bound, is introduced in Section II. Extended version of the method, discussed in Section III, allows one estimate parametric curves which exhibit Cramér-Rao-like behavior but include realistic features, such as threshold effect [4] and floor. Some illustrative results are presented in Section IV. Section V concludes.

II. Performance prediction

A. Theory-based approach

To keep our discussion focused, we will restrict ourselves to a specific application considered in [5]. The problem discussed in [5] involves making predictions of accuracy of maximum likelihood (ML) direction of arrival estimator.

The echo signal received by the radar can be modeled as

$$\boldsymbol{y} = re^{j\phi}\boldsymbol{a}(\alpha) + \boldsymbol{v}\,, \tag{1}$$

where $r = \sigma$ denotes echo amplitude, ϕ is echo phase and α is target angle. The symbol $\boldsymbol{v}$ denotes measurement noise, assumed to be complex circular Gaussian distributed with

covariance matrix $\sigma_v^2\boldsymbol{I}$, where $\boldsymbol{I}$ denotes eye matrix with size corresponding to size of $\boldsymbol{v}$ and $\boldsymbol{y}$.

The maximum likelihood estimator is obtained by maximizing the log-likelihood function

$$\hat{\alpha} = \arg\max_{\alpha,A}\ l(\boldsymbol{y},\alpha,A)$$
$$l(\boldsymbol{y},\alpha,A) = C - \frac{[\boldsymbol{y}-A\boldsymbol{a}(\alpha)]^{\mathrm{H}}[\boldsymbol{y}-A\boldsymbol{a}(\alpha)]}{\sigma_v^2}\ ,$$

where $A = re^{j\phi}$ is a complex amplitude and C denotes a constant term which is of no importance. The so-called compressed likelihood is obtained by substituting the optimal choice of amplitude

$$A_o(\alpha) = \boldsymbol{a}^{\#}(\alpha)\boldsymbol{y}\ ,$$

where $\boldsymbol{a}^{\#}(\alpha) = [\boldsymbol{a}^{\mathrm{H}}(\alpha)\boldsymbol{a}(\alpha)]^{-1}\boldsymbol{a}^{\mathrm{H}}(\alpha)$ is pseudoinverse of $\boldsymbol{a}(\alpha)$. This leads to the following formula for ML estimator [4]

$$\hat{\alpha} = \arg\min_{\alpha}\ \boldsymbol{y}^{\mathrm{H}}\boldsymbol{Q}(\alpha)\boldsymbol{y}\ , \tag{2}$$

where

$$\boldsymbol{Q}(\alpha) = \boldsymbol{I} - \frac{\boldsymbol{a}(\alpha)\boldsymbol{a}^{\mathrm{H}}(\alpha)}{\boldsymbol{a}^{\mathrm{H}}(\alpha)\boldsymbol{a}(\alpha)}$$

is the noise subspace projection matrix.

The performance of ML angle estimator is well predicted by Cramér-Rao lower bound, which takes the form

$$\mathbf{CRB} = \mathbf{FIM}(\boldsymbol{\theta})^{-1}\ , \tag{3}$$

where $\boldsymbol{\theta} = [\alpha\ r\ \phi]^{\mathrm{T}}$ is the vector of parameters and

$$\mathbf{FIM}(\boldsymbol{\theta}) = \mathrm{E}\left[\left(\frac{\partial l}{\partial \boldsymbol{\theta}}\right)\left(\frac{\partial l}{\partial \boldsymbol{\theta}}\right)^{\mathrm{T}}\right]\ ,$$

denotes Fisher information matrix.

Under adopted Gaussian assumptions the FIM can be found using textbook approach [4]. Furthermore, the Schur complement technique can be employed to get rid of nuisance parameters r, ϕ. Doing so allows one to obtain a closed form solution for the top left element of inverse FIM. It takes the form [5]

$$[\mathbf{FIM}(\boldsymbol{\theta})^{-1}]_{1,1} = \frac{\sigma_v^2}{2r^2}\left[\left\|\frac{\partial \boldsymbol{a}(\alpha)}{\partial\alpha}\right\|^2 - \left(\frac{\partial \boldsymbol{a}(\alpha)}{\partial\alpha}\right)^{\mathrm{H}}\frac{\boldsymbol{a}(\alpha)\boldsymbol{a}^{\mathrm{H}}(\alpha)}{\|\boldsymbol{a}(\alpha)\|^2}\frac{\partial \boldsymbol{a}(\alpha)}{\partial\alpha}\right]^{-1}\ , \tag{4}$$

where $\|\boldsymbol{x}\|^2 = \boldsymbol{x}^{\mathrm{H}}\boldsymbol{x}$.

Remark: In practice it may occur that measurement noise variance σ_v^2 is unknown. In such case it can be estimated using the following formula

$$\hat{\sigma}_v^2 = \frac{2}{2K-3}\boldsymbol{y}^{\mathrm{H}}\boldsymbol{Q}(\alpha)\boldsymbol{y}\ , \tag{5}$$

where K denotes the size of vectors $\boldsymbol{y}$, $\boldsymbol{v}$ and the normalizing factor $2/(2K-3)$ is intended to reduce estimator bias. Finally, note that (4) stays valid when σ_v^2 is unknown and must be estimated.

B. Performance curve estimation

The actual, rather than theoretical, system performance can be established by observing a cooperating target equipped with GPS receiver. Using such 'ground truth' one can compute radar angle estimation errors easily. However, due to fluctuations of target's radar cross section, such data needs additional processing before a realistic performance curve can be obtained. In this section we line out such a procedure.

First, we assign an estimate of signal to noise ratio to each measurement. Such an estimate may be formed as a ratio of estimated signal and noise powers

$$\hat{SNR} = \frac{|\boldsymbol{a}^{\#}(\hat{\alpha})\boldsymbol{y}|^2}{\hat{\sigma}_v^2}\ . \tag{6}$$

To obtain experimental performance curve we will exploit the dominant feature of Cramér-Rao lower bound (4), i.e. its dependence on inverse SNR. We therefore adopt the following model of estimator variance

$$\mathrm{E}[(\hat{\alpha}-\alpha)^2|SNR] = \frac{k_1}{SNR}\ , \tag{7}$$

where k_1 is a model coefficient which will be estimated from data.

To describe the distribution of estimation errors for a particular value of signal to noise ratio we will use generalized normal family [7]

$$p(x) = \frac{\beta}{2\alpha\Gamma(1/\beta)}e^{-(|x|/\alpha)^{\beta}}\ , \tag{8}$$

where $\Gamma(\cdot)$ denotes the Gamma function.

The generalized normal family is parametrized by two parameters: scale $\alpha > 0$ and shape $\beta > 0$. It includes standard Gaussian distribution ($\beta = 2$) as well as e.g. Laplace distribution ($\beta = 1$). This ability to model heavy tailed distributions is a desirable feature, because it will allow us to cope with outliers without relying on heuristics. Another interesting feature of the family is that, in the limiting case $\beta \to \infty$, the density $p(x)$ converges to uniform distribution on the interval $(-\alpha, \alpha)$.

A variance of a random variable X with generalized Gaussian density equals

$$\mathrm{E}[X^2] = \frac{\alpha^2\Gamma(3/\beta)}{\Gamma(1/\beta)}\ . \tag{9}$$

Combining (7) with (9) leads to

$$\alpha = \sqrt{\frac{\Gamma(1/\beta)}{\Gamma(3/\beta)}\frac{k_1}{SNR}}$$

which can be substituted into (8) to obtain SNR-dependent distribution of errors. Replacing unknown value of SNR with its estimate $\hat{SNR}$ leads us to the following procedure for estimating the performance curve: given N data points (values of measurement error) $\Delta\alpha_1, \Delta\alpha_2, \ldots, \Delta\alpha_N$ with assigned

estimates of signal to noise ratio $S\hat{N}R_n$, $n = 1, 2, \ldots, N$, maximize the following semi-loglikelihood function

$$J(e_1, \ldots, e_N, S\hat{N}R_1, \ldots, S\hat{N}R_N, k_1, \beta) = \sum_{n=1}^{N} \log\left[\frac{\beta}{2\alpha_n \Gamma(1/\beta)} e^{-(|e_n|/\hat{\alpha}_n)^\beta}\right] = \sum_{n=1}^{N} \left[\log(\beta) - \log(2\hat{\alpha}_n) - \log[\Gamma(1/\beta)] - (|\Delta\alpha_n|/\hat{\alpha}_n)^\beta\right] , \tag{10}$$

where

$$\hat{\alpha}_n = \sqrt{\frac{\Gamma(1/\beta)}{\Gamma(3/\beta)} \frac{k_1}{S\hat{N}R_n}} .$$

Optimization should be carried out with respect to variables k_1, β and subject to constraints $k_1 > 0$, $\beta > 0$.

In the next section we will discuss some extensions of the above proposed method. However, it would make sense to gain trust in the basic variant first. To this end we will employ simulation-based approach. This will allow us to have full knowledge of system performance, i.e. we will be able to compare estimated performance curve with actual one.

The simulation employs a 12-element linear array whose elements are spaced at half wavelength. The array manifold equals

$$\boldsymbol{a}(\alpha) = \begin{bmatrix} 1 & e^{j\alpha\pi/2} & e^{j2\alpha\pi/2} & \ldots & e^{j(K-1)\alpha\pi/2} \end{bmatrix}^{\mathrm{T}}$$

where $K = 12$. Note that $\alpha \in [-1, 1]$ is defined as *sine* of angle at which the signal impinges the array.

We generated $N = 500$ random realizations of measurements $\boldsymbol{y}_n$. In each case the signal to noise ratio was randomly generated with uniform distribution from the interval [20 dB, 40 dB] (note that the term "uniform" applies to logarithmic scale) and α_n was drawn uniformly from the interval $[-0.5,\ 0.5]$. We then estimated α_n, noise variance σ_v^2 and signal to noise $S\hat{N}R_n$ using (2), (5) and (6), respectively. Optimization was carried out using Nedler-Mead simplex search algorithm [8] which was found to be superior to e.g. Broyden–Fletcher–Goldfarb–Shanno (commonly abbreviated BFGS) algorithm [9] in the sense that it exhibited better capability to avoid being trapped in local minimums.

Under the adopted range of signal to noise ratio the accuracy of ML estimator reaches the Cramér-Rao lower bound (4). Therefore we have a very clear point of reference for the experimental curve obtained via (10). Fig. 2 overlays the data, i.e. the pairs $(\Delta\alpha_n^2,\ S\hat{N}R_n)$, Cramér-Rao lower bound and the estimated curve. Observe that, due to the fact that signal to noise ratio is unknown and estimated, the range of $S\hat{N}R$ exceeds the actual interval [20 dB, 40 dB]. Despite this obvious discrepancy of true and estimated SNRs the estimated mean square error curve overlaps with the true one, which confirms that the basic method (10) works.

III. Extensions

We now shift focus to extensions of our basic scheme (10). Specifically, we will discuss modeling floor and threshold effects. The floor effect occurs at high signal to noise ratio, when system performance is dominated by array calibration errors, rather than noise. This causes the accuracy of the estimator to saturate at some value, rather than continue to decrease with increasing SNR, as (4) would suggest. Threshold

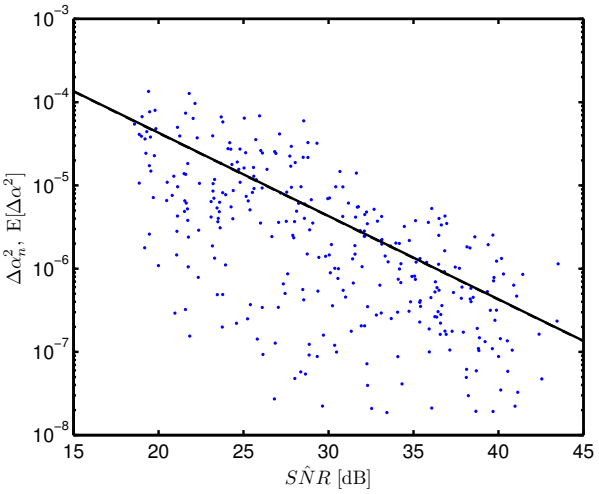

Figure 2. Comparison of data points (blue dots), Cramér-Rao lower bound (dashed line) and estimated mean square error curves (solid line). The solid line overlaps the dashed one.

effect, on the other hand, takes place at low signal to noise ratio, when the signal is buried in noise so deeply that accuracy of the estimator collapses and the mean square errors deviate from Cramér-Rao lower bound sharply.

These effects restrict applicability of Cramér-Rao bound and the model (7) to a certain range of signal to noise ratio. It is therefore important to be able to recognize when these effects take place. In the following two sections we discuss appropriate extensions of (10).

A. Floor effect

In order to include floor effect it is sufficient to extend our basic model (7) with constant term k_2

$$\mathrm{E}[(\hat{\alpha} - \alpha)^2 | SNR] = \frac{k_1}{SNR} + k_2 . \tag{11}$$

It leads to a straghtforward modification of (10), where the original formula for $\hat{\alpha}_n$ should be replaced with the following one

$$\hat{\alpha}_n = \sqrt{\frac{\Gamma(1/\beta)}{\Gamma(3/\beta)} \left[\frac{k_1}{S\hat{N}R_n} + k_2\right]} . \tag{12}$$

Note however, that when dealing with floor, caution is necessary when estimating noise level. The effect of calibration errors can be modeled by modifying (1) as follows

$$\boldsymbol{y} = re^{j\phi}\left[\boldsymbol{a}(\alpha) + \Delta\boldsymbol{a}\right] + \boldsymbol{v} , \tag{13}$$

where $\Delta\boldsymbol{a}$ is a vector of random calibration errors. Note that the quantity $re^{j\phi}\Delta\boldsymbol{a}$ can be interpreted as additional noise, whose power is in linear relationship with power of the useful signal $re^{j\phi}\boldsymbol{a}(\alpha)$. For this reason formula (5) may lead to erroneous estimation of noise variance when the signal level is high. To avoid this problem σ_v^2 should be estimated under small signal conditions, e.g. in a separate experiment or by discarding portion of data where the estimated signal amplitude is high.

Validation of the formula (12) can be performed using a modified version of simulation experiment described in section II. Calibration errors were introduced using (13), where $\Delta\boldsymbol{a}$ was zero mean circular complex Gaussian vector with covariance matrix $0.01\boldsymbol{I}$. Additionally, in order to better expose the

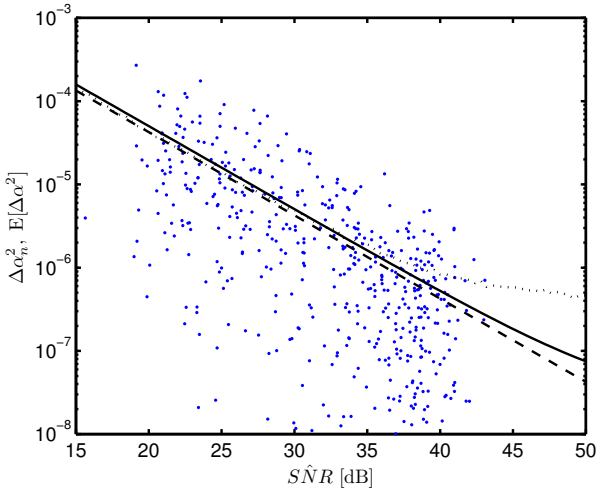

Figure 3. Comparison of data points (blue dots), Cramér-Rao lower bound (dashed line), actual (dotted line) and estimated (solid line) mean square error curves in presence of array calibration errors without accounting for apparent increase of noise power for high SNR.

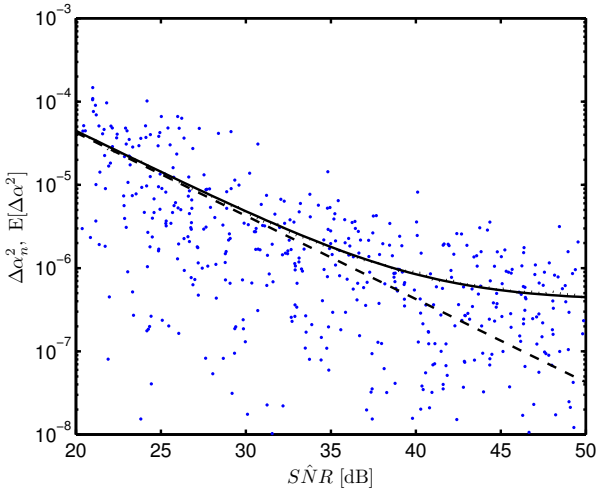

Figure 4. Comparison of data points (blue dots), Cramér-Rao lower bound (dashed line), actual (dotted line) and estimated (solid line) mean square error curves in presence of array calibration errors when noise level is estimated properly.

floor effect, the upper range of signal to noise was increased to 50 dB.

Introduction of calibration errors into the simulation means that Cramér-Rao lower bound should be supplemented with actual performance curve of the estimator. Such curve was obtained by running 200000 number of Monte-Carlo simulations for 100 values of SNR (2000 runs per SNR value). The values of SNR were distributed uniformly on the log scale in the interval [20 dB, 50 dB].

Fig. 3 shows simulation results obtained using (12) when the noise level was estimated using a faulty procedure, which did not account for influence of calibration errors on (5). Note that the estimated values of SNR saturate at about 40 dB. The outcome is optimistic performance curve, which barely exhibits any floor. Fig. 4, on the other hand, shows results obtained using (12) with properly estimated noise level. This time the experimental curve almost coincides with the actual one.

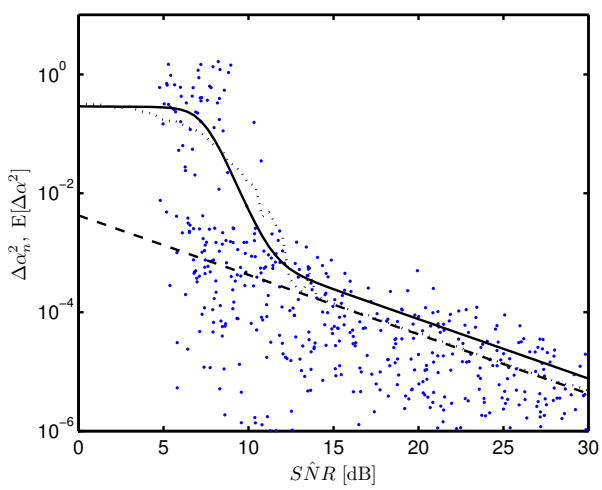

Figure 5. Comparison of data points (blue dots), Cramér-Rao lower bound (dashed line), actual (dotted line) and estimated (solid line) mean square error curves with threshold effect in place.

B. Threshold effect

Modeling threshold effect is more difficult than modeling the floor. The following formula is proposed for this purpose

$$\mathrm{E}[(\hat{\alpha}-\alpha)^2|SNR] = \frac{k_1}{SNR} + k_2 S(SNR)\ , \qquad (14)$$

where

$$S(SNR) = \frac{1-\tanh\left[k_3 \log\frac{SNR}{k_4}\right]}{2}$$

is a sigmoid-type term which will increase mean square error curve for signal to noise smaller than, approximately, k_4. The coefficient k_3 governs the "steepness" of $S(SNR)$ and, finally, k_2 is intended to correspond to estimator mean square error when the threshold effect is fully developed.

The resulting algorithm is another straightforward modification of (10), i.e. with α_n replaced by

$$\hat{\alpha}_n = \sqrt{\frac{\Gamma(1/\beta)}{\Gamma(3/\beta)}\left[\frac{k_1}{\hat{SNR}_n} + k_2 S(\hat{SNR}_n)\right]}\ . \qquad (15)$$

Once again the method will be verified using simulations. The overall scenario is the same as in the preceding subsection, but this time the SNR was lowered to range between 5 dB and 30 dB. Fig. 5 depicts the results. Observe that some bias has developed in k_1, as the experimental curve does not overlap with Cramér-Rao lower bound for medium SNR. However, the threshold is modeled reasonably well and the model certainly provides a good hint about the range of SNR when Cramér-Rao lower bound can no longer be trusted.

Although the cause of bias of k_1 is not fully understood, there is a very simple way to reduce it. Recall that the model (14) is an extension of (7). In particular, the role of coefficient k_1 is the same in both models. The curve obtained using (15) can be treated as preliminary, and the coefficient k_1 can be reestimated using (10) and a truncated data set, which excludes all points with $\hat{SNR}$ such that $S(\hat{SNR})$ deviates from zero farther than some tolerance level ε. Typical result of such two stage procedure, with tolerance level set to $\varepsilon = 0.001$, is shown in Fig. 6.

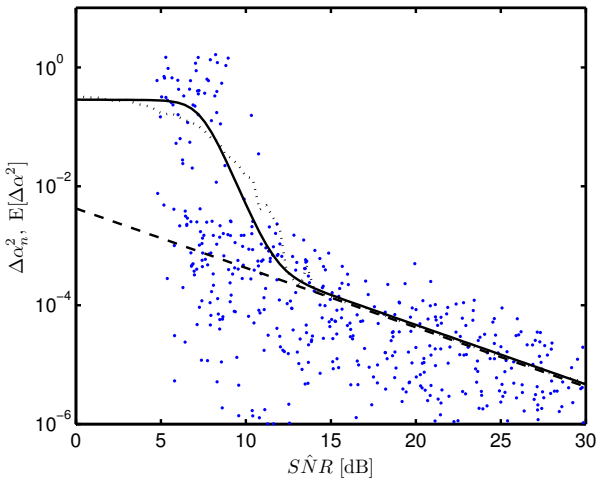

Figure 6. Comparison of data points (blue dots), Cramér-Rao lower bound (dashed line), actual (dotted line) and estimated (solid line) mean square error curves with threshold effect in place. The estimated mean square error curve was obtained using the two-step estimation procedure.

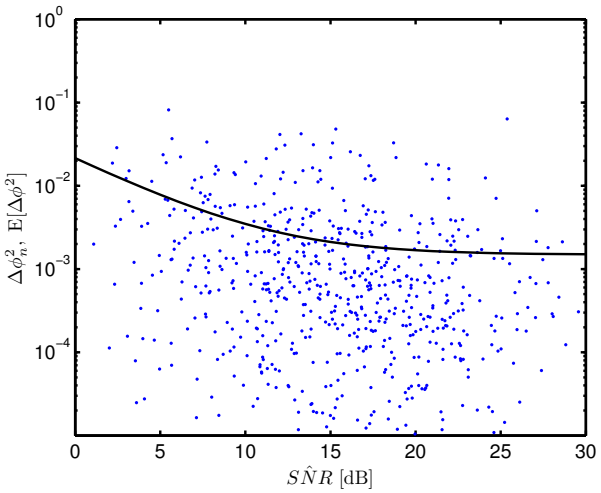

Figure 7. Comparison of data points (blue dots) and estimated (solid line) mean square error curve obtained for a real-world system. All values are normalized with respect to squared beamwidth.

IV. Real-world example

To illustrate the proposed method with some real world results we perform a case study of an experimental radar system. The radar employs rotating array with a rather wide illumination/receive beam. While this has an adverse effect of lowering antenna gain, it aids detection and classification by enabling very long coherent processing.

We study accuracy of DOA estimates of the system. The data consists of 768 measurements of raw signals in the coherent processing interval, as well as estimated and true DOA. The signals were used to establish noise level of the system and the model (11) was fitted to results. The resulting estimate of mean square error behavior is shown in Fig. 7. We see that the system performance saturates at rather low SNR, which stems from mediocre array calibration. On the other hand, overall system performance is quite satisfactory as it reaches accuracy better than one tenth of a beamwidth at SNR as low as 5 dB.

V. Conclusions

It was shown how mean square error curves can be estimated using results of real world trials. The proposed procedure is based on fitting a generalized normal model, constructed in such a way so that its variance depends on signal to noise ratio. The extensions of the basic procedure allow the model to include floor and threshold effects.

In case of current generation of radar systems the primary role for experimental curves is system evaluation: they can be compared with Cramér-Rao lower bound to judge the overall quality, and room for improvement, of the radar system. However, it can be expected that their importance will grow in the future, due to anticipated proliferation of cognition-inspired radars, where they will be a key component enabling accurate prediction and optimization of radar resource management.

References

[1] J. Guerci, *Cognitive Radar: The Knowledge-aided Fully Adaptive Approach*. Artech House, 2010.

[2] S. Haykin, *Cognitive Dynamic Systems: Perception-action Cycle, Radar and Radio*. Cambridge University Press, 2012.

[3] K. L. Bell, C. J. Baker, G. E. Smith, J. T. Johnson, and M. Rangaswamy, "Cognitive radar framework for target detection and tracking," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 8, pp. 1427–1439, 2015.

[4] H. L. Van Trees, K. L. Bell, and Z. Tian, *Detection, Estimation, and Modulation Theory Part I*. New York: John Wiley & Sons, 2013.

[5] M. Meller, "Radar time budget optimization subject to angle accuracy constraint via cognitive approach," in *Proc. 2017 International Radar Symposium (IRS 2017)*, Prague, Czech Republic, 2017.

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

[7] S. Nadarajah, "A generalized normal distribution," *Journal of Applied Statistics*, vol. 32, no. 7, pp. 685–694, 2005.

[8] J. Nelder and R. Mead, "A simplex method for function minimization," *The Computer Journal*, vol. 7, no. 4, p. 308, 1965.

[9] R. Fletcher, *Practical methods of optimization*. New York: John Wiley & Sons, 1987.