Postprint of: Meller M., On Bayesian Tracking and Prediction of Radar Cross Section, IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONIC SYSTEMS, Vol. 55, iss. 4 (2018), pp. 1756-1768, DOI: 10.1109/TAES.2018.2875572

On Bayesian Tracking and Prediction of Radar Cross Section

Michał Meller

***Abstract*—We consider the problem of Bayesian tracking of radar cross section. The adopted observation model employs the gamma family, which covers all Swerling cases in a unified framework. State dynamics are modeled using a nonstationary autoregressive gamma process. The principal component of the proposed solution is a nontrivial gamma approximation, applied during the time update recursion. The superior performance of the proposed approach is confirmed using simulations and a real-world dataset.**

***Index Terms*—radar cross section, Bayesian tracking, autoregressive gamma process**

I. Introduction

To characterize a target succinctly, a radar engineer may choose to specify its radar cross section (RCS). Radar cross section of the target quantifies its magnitude. It is defined as a cross-sectional area of a perfectly reflecting sphere that produces the echo of the same strength as the actual target [1]–[3]. The concept of RCS is essential to system analysis and design because radar performance figures, such as the detection range or accuracy, are meaningless if they are not accompanied by a specification of a "nominal" target of interest. Typical values of RCS employed in system analysis are in the range of 1-2 m^2 for "fighter sized" targets, 10-100 m^2 for "bomber sized" targets, and 0.0001-0.01 m^2 for the so-called low observable (LO) targets, such as stealth airplanes, missiles or projectiles, among others [1].

However, the benefits of knowing RCS figures are not limited to system analysis and design. Improved performance has been demonstrated in the areas such as discrimination between targets and clutter [4], [5], multi-target tracking [6]–[9], track-to-truth assignment [10], non-cooperative target recognition [11]–[14], and resource management [15]–[17].

Resource management is, in fact, the application that motivated this paper. A multifunction radar must share its resources between multiple tasks, such as the search and active tracking of multiple targets, among others. Knowledge of the tracked targets' RCS allows one to schedule the optimal, in the sense of minimizing the time/power budget allocation, track update looks, which will increase the number of targets that the radar can handle before it saturates. The difficulty of implementing such an advanced radar management scheme lies in the fact that the RCS of a target is often unknown. To mitigate this obstacle, one has two options. First, one can employ a database containing values of RCS for various target classes, obtained using measurements [18]–[20] or computational methods [21]–[24]. However, no matter how extensive, such a database will likely need updating with new target classes. Additionally, this approach requires one to know the target class accurately, which is difficult in military applications. The second option, which is free of such limitations, is to equip the radar with an RCS estimator and entrust it with the task of assessing the target's RCS on-line, i.e., by analyzing the contents of the actual echo signal.

The design of the RCS estimator should take into account the properties of the RCS specific to the radar application. When one measures the RCS on a range, it natural to treat it as a deterministic quantity, whose value is a function of an aspect angle. However, when the radar is observing a target, the aspect angle varies, which causes fluctuations of the radar echo strength. Since the fluctuations appear to be random, it makes more sense to employ the stochastic modeling tools than to rely on the deterministic modeling techniques [3]. A typical stochastic model of RCS specifies a value of the average RCS and a probability density family that describes the distribution of the fluctuations. This class includes, among others, the Swerling I-IV models [1], [25], which belong to the broader family of χ^2 models [26].

It can be argued, however, that the models of this form, while adequate for system analysis and design, are too simplistic for the resource management. It is well known that radar targets can exhibit different statistical properties when viewed from different sectors [10]. For example, the designs of modern fighter planes include the features that reduce their backscatter in the frontal area. To avoid feeding the resource manager with outdated information, the RCS estimator should work out the *local* estimates of the parameters of the stochastic RCS model, i.e., it should be a *tracker* with limited memory.

The final requirement concerns the ability to deliver predictive distributions of the future RCS. Such distributions are necessary for accurate decision-making because of the random behavior of RCS in radar. To enable generation of the predictive distributions, one can design the tracker using the Bayesian framework. In this approach, the parameters are treated as random variables and assigned a priori distributions. Application of Bayes' formula allows one to work out the a posteriori distribution of the parameters [27], which is a complete description of one's current state of knowledge about the tracked process. To form the predictive distribution, one should propagate the a posteriori distribution into the future and compound it with the model distribution. Note that the straightforward approach of substituting point estimates into the model distribution can result in a degraded performance

The author is with the Department of Automatic Control, Gdańsk University of Technology, Faculty of Electronics, Telecommunications and Computer Science, Narutowicza 11/12, 80-233 Gdańsk, Poland and PIT-RADWAR S.A., Poligonowa 30, 04-051 Warsaw, Poland. e-mail: michal.meller@eti.pg.edu.pl

because of long sequences of data required to obtain accurate estimates of the average RCS [5].

The problem of on-line RCS estimation is typically considered in the context of target tracking. Unfortunately, the existing solutions fail to satisfy at least one out of three of the above requirements. The estimators proposed in [7], [8] are parts of larger probability hypothesis density (PHD) filtering schemes. These algorithms employ stochastic models of RCS, whose parameters are assumed to be constant. Additionally, even though these trackers include the formulas that allow one to compute the estimates of the average RCS and their uncertainties, the resultant predictive pdfs were not presented. Another solution was proposed in [10]. It is a heuristic design which employs the alpha filtering approach to smooth out the fluctuations of the individual measurements. This solution, owing to its limited memory, is capable of tracking the time-varying parameters of the model. However, like in the previous case, the issue of computing the predictive distribution was not addressed. Finally, in [15] the authors described the cognitive resource management scheme which includes the estimation of the SNR, rather than RCS. Their solution relies on extracting the estimate of the SNR from the ambiguity function and treating it as valid for the next decision cycle. Neither the issue of uncertainty nor the random behavior of RCS was recognized in this scheme.

This paper proposes a novel Bayesian tracker which satisfies all of the above requirements. Derivation of the algorithm employs a nongaussian model that combines the observation process and the underlying, hidden, state process. The model is constructed such that, using a nontrivial gamma approximation, one can reach a remarkably simple final form of the tracker. Interestingly, a closer inspection of our algorithm shows that one may regard the existing methods as approximations of the Bayesian solution proposed here.

The paper is organized as follows. Section II specifies the problem assumptions and presents the derivation of the tracking algorithm. Section III presents the extensions of the basic scheme. Section IV validates the proposed tracker using computer simulations and real-world data. Section V concludes the paper.

II. Bayesian filtering of a hidden autoregressive gamma process

A. A quick review of Bayesian filtering

We will begin with a quick review of a general nonlinear Bayesian filtering problem. This material will serve as the reference when we derive the tracker.

Denote by $\boldsymbol{x}_n \in X$ ($n = 0, 1, \dots$ is a discrete, dimensionless time) a state vector, i.e., a p-dimensional Markov process with state transition probability density $p(\boldsymbol{x}_n|\boldsymbol{x}_{n-1})$ and the initial distribution of the state $p(\boldsymbol{x}_0)$. The state is assumed to be *hidden*, i.e., not observed directly. The information about the state is conveyed via an observation process $\boldsymbol{y}_n$, which has the following property

$$p(\boldsymbol{y}_n|\boldsymbol{x}_0, \boldsymbol{x}_1, \dots, \boldsymbol{x}_n, \dots) = p(\boldsymbol{y}_n|\boldsymbol{x}_n) \ ,$$

where $p(\boldsymbol{y}_n|\boldsymbol{x}_n)$ is a known probability density function.

The Bayesian filtering problem is to find the a posteriori density $p(\boldsymbol{x}_n|\mathcal{Y}_n)$, where $\mathcal{Y}_n = \{\boldsymbol{y}_0, \boldsymbol{y}_1, \dots, \boldsymbol{y}_n\}$ denotes the observation history available at the time instant n. Its solution is a generic recursion, consisting of two steps [27]

1) Time update

$$p(\boldsymbol{x}_n|\mathcal{Y}_{n-1}) = \int_X p(\boldsymbol{x}_n|\boldsymbol{x}_{n-1}) p(\boldsymbol{x}_{n-1}|\mathcal{Y}_{n-1}) \mathrm{d}\boldsymbol{x}_{n-1} \ . \quad (1)$$

2) Measurement update

$$p(\boldsymbol{x}_n|\mathcal{Y}_n) = c\, p(\boldsymbol{y}_n|\boldsymbol{x}_n) p(\boldsymbol{x}_n|\mathcal{Y}_{n-1}) \ , \quad (2)$$

where

$$c = \left[\int_X p(\boldsymbol{y}_n|\boldsymbol{x}_n) p(\boldsymbol{x}_n|\mathcal{Y}_{n-1}) \mathrm{d}\boldsymbol{x}_n \right]^{-1}$$

is a normalizing constant which guarantees that

$$\int_X p(\boldsymbol{x}_n|\mathcal{Y}_n) \mathrm{d}\boldsymbol{x}_n = 1 \ .$$

The simplicity of equations (1)-(2) can be deceiving. The main difficulty arising in the Bayesian filtering is that the solution is very often analytically or computationally intractable. Faced with this problem, one may resort to approximations, such as Grid-based or Monte Carlo methods, that often produce good results, albeit at the expense of a high computational cost [27]–[30].

The Bayesian filtering problem can be solved using the analytical approach if certain regularity conditions are satisfied. Let $p(\boldsymbol{y}_n|\boldsymbol{x}_n)$ belong to a family $\mathcal{F}$, where $\boldsymbol{x}_n$ enters as a parameter

$$p(\boldsymbol{y}_n|\boldsymbol{x}_n) = \mathcal{F}(\boldsymbol{y}_n; \boldsymbol{x}_n) \ .$$

Suppose that $p(\boldsymbol{x}_n|\mathcal{Y}_{n-1})$ happens to be the so-called conjugate prior distribution of $p(\boldsymbol{y}_n|\boldsymbol{x}_n)$,

$$p(\boldsymbol{x}_n|\mathcal{Y}_{n-1}) = \mathcal{F}'(\boldsymbol{x}_n; \boldsymbol{\theta}_{n|n-1}) \ ,$$

where $\mathcal{F}'(\cdot)$ denotes the conjugate prior of $\mathcal{F}(\cdot)$ and $\boldsymbol{\theta}_{n|n-1}$ is the vector of the so-called hyperparameters. Then the a posteriori distribution also belongs to the family $\mathcal{F}'$

$$p(\boldsymbol{x}_n|\mathcal{Y}_n) = \mathcal{F}'(\boldsymbol{x}_n; \boldsymbol{\theta}_{n|n}) \ ,$$

where the a posteriori vector of the hyperparameters $\boldsymbol{\theta}_{n|n}$ is a function of $\boldsymbol{\theta}_{n|n-1}$ and $\boldsymbol{y}_n$ – see [31] for more details and an exhaustive number of examples of the conjugate prior pairs.

Note that, the conjugate prior property can be employed repeatedly, i.e., in each iteration of the recursion, only when $p(\boldsymbol{x}_n|\boldsymbol{x}_{n-1})$ satisfies the following condition

$$p(\boldsymbol{x}_{n+1}|\mathcal{Y}_n) = \mathcal{F}'(\boldsymbol{x}_{n+1}, \boldsymbol{\theta}_{n+1|n}) \ . \quad (3)$$

This situation occurs, among others, when the probabilities $p(\boldsymbol{x}_n|\boldsymbol{x}_{n-1})$, $p(\boldsymbol{y}_n|\boldsymbol{x}_n)$ stem from the linear Gaussian model

$$\begin{aligned} \boldsymbol{x}_{n+1} &= \boldsymbol{A}_n \boldsymbol{x}_n + \boldsymbol{w}_n \\ \boldsymbol{y}_n &= \boldsymbol{C}_n \boldsymbol{x}_n + \boldsymbol{v}_n \ , \end{aligned}$$

where $\{\boldsymbol{w}_n\}$, $\{\boldsymbol{v}_n\}$ are mutually independent zero-mean Gaussian white noise processes and $p(\boldsymbol{x}_0)$ is a known Gaussian

density. In such case, the Bayesian filter takes the form of the celebrated Kalman filter [27], [32].

When the condition specified by Eq. (3) does not take place, it might be possible to approximate $p(\boldsymbol{x}_{n+1}|\mathcal{Y}_n)$ with a suitably chosen distribution $\mathcal{F}'(\boldsymbol{x}_{n+1}, \boldsymbol{\theta}_{n+1|n})$. Approximations of this kind are the cornerstones of the so-called nonlinear Kalman filters, such as the unscented Kalman filter [33] and the cubature Kalman filter [34]. Note that, even though the nonlinear Kalman filters are no longer optimal, they perform more than adequately in a broad spectrum of applications [35]–[38]. We will refrain from discussing the details of these algorithms any further – the purpose of mentioning this approach stems from the fact the proposed tracker will also exploit a favorable approximation of $p(\boldsymbol{x}_{n+1}|\mathcal{Y}_n)$, although not the Gaussian one.

B. Assumptions

Following the above short discussion, we can now introduce the process model which become a foundation for our tracker. Denote by $\gamma(x; \alpha, \beta)$ the probability density function of the gamma distribution with a shape parameter α and a rate parameter β, $\alpha > 0$, $\beta > 0$

$$\gamma(x; \alpha, \beta) = \begin{cases} 0 & \text{for } x < 0 \\ \dfrac{\beta^\alpha}{\Gamma(\alpha)} x_0^{\alpha-1} e^{-\beta x} & \text{for } x \geq 0 \end{cases} \tag{4}$$

Suppose that the radar's plot extractor can assess the target's RCS during a dwell. Such a functionality can be implemented by a straightforward scaling of the target's echo power with a factor obtainable from the radar equation [7]. The measurements of the RCS obtained from the plot extractor will play the role of the (scalar) sequence of observations $\{y_n\}$. We will assume that

(**A1**) The conditional distribution of the observations is the gamma with a known shape parameter $a > 0$ and an unknown rate parameter x_n

$$p(y_n|x_n) = \gamma(y_n; a, x_n) = \begin{cases} 0 & \text{for } y_n < 0 \\ \dfrac{x_n^a}{\Gamma(a)} y_0^{a-1} e^{-y_n x_n} & \text{for } y_n \geq 0 \end{cases} \tag{5}$$

Assumption (A1) can be regarded as a generalization of the Swerling family [25] in two ways. First, with a proper choice of a, the density in Eq. (5) can cover all Swerling cases. For instance, the Swerling-I case corresponds to $a = 1$, while $a = 2$ yields the Swerling-III case. The extended Swerling cases, introduced in [26], can be obtained as well. Second, the assumption introduces the state variable x_n, whose role is to explain differences in behavior of y_n over different aspect angles. Observe that the conditional mean of y_n, which corresponds to the average RCS, is inversely proportional to x_n

$$\mathrm{E}[y_n|x_n] = \frac{a}{x_n} = a x_n^{-1} \,. \tag{6}$$

By allowing x_n to vary, one can model targets whose statistical properties are time-dependent, rather than fixed. Since, in such a case, the quantity on the right-hand side of Eq. (6) characterizes the local properties of the observation process, we will further refer to it as the local average RCS.

Finally, note that the form of (A1) implies that any influence of the receiver noise on y_n is neglected. Arguably, this is not a severe limitation. Compared with fluctuations of RCS, the effects of the noise are rather small, especially if one considers the problem of predicting the *future* values of y_n. On the other hand, the benefits of neglecting the noise are quite substantial because the construction of the tracker becomes considerably simplified – see, e.g., [7] for an example of the RCS estimator that employs similar simplifications.

The next two assumptions specify the dynamics of the state variable and its initial probability distribution.

(**A2**) The state sequence $\{x_n\}$ forms a scalar Markov chain with state transition probability density $p(x_n|x_{n-1})$ given by the following limit [39]

$$p(x_n|x_{n-1}) = \lim_{\delta \to 0} p_\delta(x_n|x_{n-1})$$

$$p_\delta(x_n|x_{n-1}) = \begin{cases} 0 & \text{for } x_n < 0 \\ \exp\left(-\dfrac{x_n}{c}\right) \displaystyle\sum_{k=0}^{\infty} \frac{e^{-b x_{n-1}} b^k x_{n-1}^k}{c^{\delta+k}\Gamma(\delta+k)} \frac{x_n^{\delta+k-1}}{k!} & \text{for } x_n \geq 0 \end{cases} \tag{7}$$

where $c > 0$ is a small constant, $c \approx 0$, and $b = 1/c$.

(**A3**) The initial state x_0 is a gamma-distributed random variable with known shape and rate parameters, $\alpha_0 > 0$, $\beta_0 > 0$

$$p(x_0) = \gamma(x_0; \alpha_0, \beta_0) = \begin{cases} 0 & \text{for } x_0 < 0 \\ \dfrac{\beta_0^{\alpha_0}}{\Gamma(\alpha_0)} x_0^{\alpha_0-1} e^{-\beta_0 x_0} & \text{for } x_0 \geq 0 \end{cases} \tag{8}$$

Following [39], the state is modeled using the limiting case ($\delta \to 0$) of the autoregressive gamma process [c.f. (A2)-(A3)]. Although the form of (7) is unusually complex, it is used here deliberately to place the paper in a broader context of nongaussian autoregressive modeling [40], [41]. In practice, the two following results are more useful. First, one can generate trajectories of x_n corresponding to (A2) using the following recursion, which is the special case of a more generic formula presented in [40], [41], obtained for $\delta \to 0$,

$$x_n = \sum_{i=1}^{N(x_{n-1})} W_i \,, \tag{9}$$

where W_i are exponential i.i.d. random variables with parameter $1/c$ and $N(x)$ is a Poisson distributed random variable with parameter x/c. Second, Laplace transform of $p(x_n|x_{n-1})$ has a remarkably simple form (see Appendix A for the derivations of the results presented in this part of our discussion)

$$\begin{aligned} \mathcal{L}\left(p(x_n|x_{n-1})\right) &= \int_0^\infty p(x_n|x_{n-1}) e^{-s x_n} \mathrm{d}x_n \\ &= \exp\left(-\frac{s x_{n-1}}{sc+1}\right) \,. \end{aligned} \tag{10}$$

We will employ (10), rather than (7), throughout most of the paper.

It is clear from (9) that the adopted model is not backed by any actual physics. However, the model in this form allows one to perform a tractable analysis of the problem without sacrificing the essential features of the dynamic model of RCS. Observe that, under $\delta \to 0$ and $b = 1/c$, the process $\{x_n\}$ has the martingale property

$$\mathrm{E}\left[x_n | x_n - 1\right] = x_{n-1} \tag{11}$$

and can drift unbounded. The degree of nonstationarity of x_n is governed by the parameter c, as can be seen from the expression for the conditional variance

$$\mathrm{V}\left[x_n | x_{n-1}\right] = 2cx_{n-1} \ . \tag{12}$$

It follows that, for small values of c, the state variable will drift slowly. Since any changes in x_n immediately affect the local average RCS [c.f. (6)], the proposed model offers enough flexibility to explain the behavior of RCS, at least in a local timeframe.

C. The tracker – the time update step

Since the statistical properties of y_n are governed by the state, we will design the tracker to estimate x_n. Such an approach will later allow us to construct a forecasting mechanism to predict the future values of y_n.

Suppose that $p(x_{n-1}|\mathcal{Y}_{n-1})$ corresponds to the gamma pdf with the shape α_{n-1} and the rate β_{n-1}

$$p(x_{n-1}|\mathcal{Y}_{n-1}) = \gamma(x_{n-1}; \alpha_{n-1}, \beta_{n-1}) \ . \tag{13}$$

Using (1), (4) and (10) one can arrive at

$$\begin{aligned} p(x_n|\mathcal{Y}_{n-1}) &= \int_0^\infty \mathrm{d}x_{n-1} p(x_n|x_{n-1}) p(x_{n-1}|\mathcal{Y}_{n-1}) \\ &= \int_0^\infty \mathrm{d}x_{n-1} \int_{\sigma-j\infty}^{\sigma+j\infty} \mathrm{d}s \exp\left(-\frac{sx_{n-1}}{sc+1}\right) e^{sx_n} p(x_{n-1}|\mathcal{Y}_{n-1}) \ . \end{aligned}$$

Reversing the order of the integration yields the formula that can be identified as the inverse Laplace transform

$$p(x_n|\mathcal{Y}_{n-1}) = \int_{\sigma-j\infty}^{\sigma+j\infty} K(s) e^{sx_n} \mathrm{d}s = \mathcal{L}^{-1}(K(s)) \ , \tag{14}$$

where

$$K(s) = \int_0^\infty \exp\left(-\frac{sx_{n-1}}{sc+1}\right) p(x_{n-1}|\mathcal{Y}_{n-1}) \mathrm{d}x_{n-1} \ .$$

Substitution of Eq. (13) into $K(s)$ eventually leads to

$$K(s) = \left[\frac{sc+1}{s(\beta_{n-1}^{-1} + c) + 1}\right]^{\alpha_{n-1}} \ . \tag{15}$$

Fig. 1 shows a comparison of $p(x_{n-1}|\mathcal{Y}_{n-1})$ with the corresponding $p(x_n|\mathcal{Y}_{n-1})$ for $c = 0.05$, $\alpha_{n-1} = 10$ and three choices of β_{n-1}. All plots of $p(x_n|\mathcal{Y}_{n-1})$ were obtained

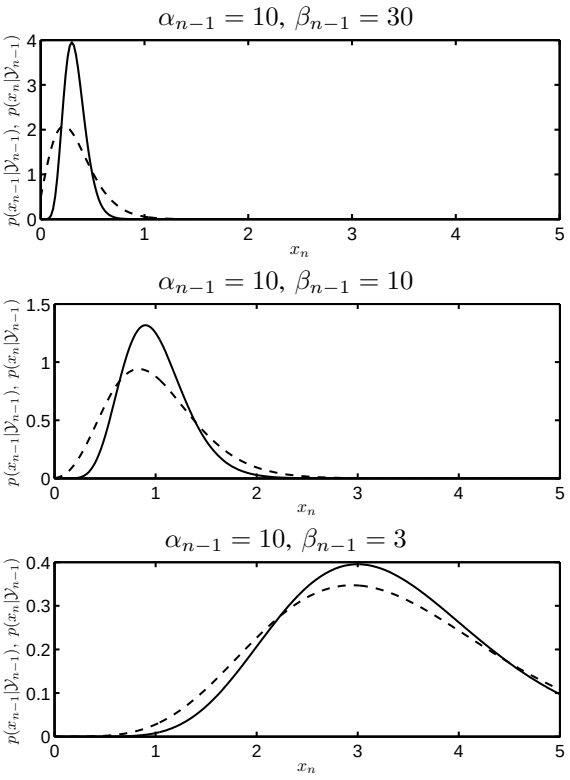

Figure 1. Comparison of $p(x_{n-1}|\mathcal{Y}_{n-1})$ (solid line) and $p(x_n|\mathcal{Y}_{n-1})$ (dashed line) for $c = 0.05$ and three choices of α_{n-1}, β_{n-1}.

by evaluating the inverse Laplace transform of the formula (15). Observe that, even though the predictive distribution of x_n widens, its overall shape remains similar to the gamma distribution, which presents an opportunity to approximate $p(x_n|\mathcal{Y}_{n-1})$ using the gamma pdf.

Indeed, the $\mathcal{L}$-transform of $\gamma(x_n; \tilde{\alpha}_n, \tilde{\beta}_n)$ reads

$$K_\gamma(s) = \frac{1}{\left(s\tilde{\beta}_n^{-1} + 1\right)^{\tilde{\alpha}_n}} \ ,$$

and differs from $K(s)$ primarily by the absence of zeros. Since the zeros of $K(s)$ are located in the high-frequency range, their influence on the overall shape of $p(x_n|\mathcal{Y}_{n-1})$ is small.

One can obtain a simple gamma approximation of $p(x_n|\mathcal{Y}_{n-1})$ using the moment matching technique. We shall equalize the first two moments of $p(x_n|\mathcal{Y}_{n-1})$ and its approximation, i.e., we shall request that

$$\begin{aligned} -\frac{\mathrm{d}}{\mathrm{d}s} K_\gamma(s)\Big|_{s\to 0} &= -\frac{\mathrm{d}}{\mathrm{d}s} K(s)\Big|_{s\to 0} \\ \frac{\mathrm{d}^2}{\mathrm{d}s^2} K_\gamma(s)\Big|_{s\to 0} &= \frac{\mathrm{d}^2}{\mathrm{d}s^2} K(s)\Big|_{s\to 0} \ . \end{aligned} \tag{16}$$

The left-hand side terms of the above system of equations are equal to $\tilde{\alpha}_n/\tilde{\beta}_n$ and $\tilde{\alpha}_n(\tilde{\alpha}_n+1)/\tilde{\beta}_n^2$, respectively. The terms on the right-hand side read (see Appendix B for the derivation)

$$\begin{aligned} -\frac{\mathrm{d}}{\mathrm{d}s} K(s)\Big|_{s\to 0} &= \frac{\alpha_{n-1}}{\beta_{n-1}} \\ \frac{\mathrm{d}^2}{\mathrm{d}s^2} K(s)\Big|_{s\to 0} &= \frac{\alpha_{n-1}}{\beta_{n-1}} \left(\frac{\alpha_{n-1}+1}{\beta_{n-1}} + 2c\right) \ . \end{aligned} \tag{17}$$

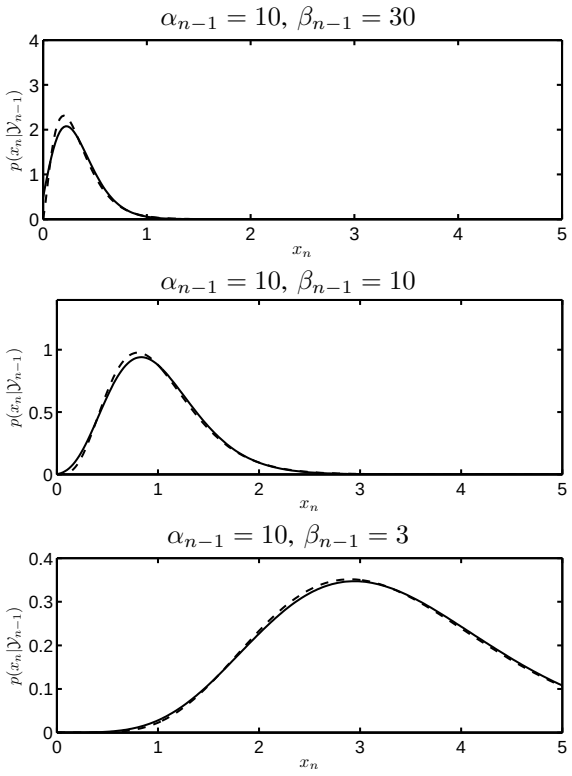

Figure 2. Comparison of $p(x_n|\mathcal{Y}_{n-1})$ (solid line) and its approximation using the gamma density $\gamma(x_n; \tilde{\alpha}_n, \tilde{\beta}_n)$ (dashed line) for $c = 0.05$ and three choices of α_{n-1}, β_{n-1}.

Combining the above partial results allows one to arrive at the following formulas for the parameters $\tilde{\alpha}_n$, $\tilde{\beta}_n$ of the approximation $\gamma(x_n; \tilde{\alpha}_n, \tilde{\beta}_n)$

$$\tilde{\alpha}_n = \frac{\alpha_{n-1}}{1 + 2c\beta_{n-1}} \qquad \tilde{\beta}_n = \frac{\beta_{n-1}}{1 + 2c\beta_{n-1}} \ . \tag{18}$$

Fig. 2 shows a comparison of $p(x_n|\mathcal{Y}_{n-1})$ and its approximation (18) for $c = 0.05$ and the same three choices of α_{n-1}, β_{n-1} as in Fig. 1. Even though the plots presented in Fig. 2 cannot be regarded as a proof that (18) works, they show a good agreement of both densities for a fairly wide range of β_{n-1}. Additional tests show that the quality of (18) decreases with increasing c, but stays reasonable for c as large as 0.1. Note that $c = 0.05$ already corresponds to quite fast changes of x_n [c.f. (12) and Fig. 1],which means that larger values of c are unlikely to be of practical importance.

Remark: A "quick and dirty" approximation of $K(s)$ can be obtained by neglecting the zeros of (15) and modifying its exponent so as to match the means of $K(s)$ and $K_\gamma(s)$. This procedure yields the rule whose form is almost the same as (18), except that the coefficient 2 in the denominator is replaced with 1. The accuracy of the resulting approximation is substantially worse than (18), however.

D. The tracker – the measurement update step

We will now derive the measurement update rule. To this end, we will assume that the approximation introduced in the previous subsection is sufficiently accurate, $p(x_n|\mathcal{Y}n-1) \simeq \gamma(x_m; \tilde{\alpha}_n, \tilde{\beta}_n)$.

Under (A1), one is allowed to take advantage of the conjugate prior property of the gamma family, summarized in the following lemma [31]

Lemma 1. *Let V be a gamma distributed random variable with a known shape $a > 0$ and an unknown rate b. Assume that the a priori distribution of b is also a gamma with a known shape $\tilde{\alpha} > 0$ and rate $\tilde{\beta} > 0$. The a posteriori distribution of b, $p(b|v)$ is the gamma with the shape $\alpha = \tilde{\alpha} + a$ and the rate $\beta = \tilde{\beta} + v$, where v denotes the observed realization of V.*

A straightforward application of Lemma 1 leads to the conclusion that the conditional distribution $p(x_n|\mathcal{Y}n)$ is $\gamma(x_n; \alpha_n, \beta_n)$, whose shape and rate parameters take the following form

$$\alpha_n = \tilde{\alpha}_n + a \qquad \beta_n = \tilde{\beta}_n + y_n \ . \tag{19}$$

Note that the proposed update rules, stated in equations (18) and (19), can be employed perpetually, i.e., they form a recursion that can be executed for every n.

E. A summary of the algorithm

The proposed tracker is summarized in Table I. Since the table is meant to serve as a quick reference, it includes interpretations of all quantities appearing in the equations.

The pair of measurement update equations is responsible for factoring the new observation y_n into the Bayesian state estimate, i.e., into the a posteriori density $p(x_n|\mathcal{Y}_n)$. The form of these equations suggests that α_n/a can be interpreted as the effective number of observations, while β_n – as the effective sum of observations. The minimum mean square error (MMSE) estimate of the state, which reads,

$$\mathrm{E}\left[x_n|\mathcal{Y}_n\right] = \frac{\alpha_n}{\beta_n} \tag{20}$$

is proportional to the inverse of the effective sum of observations.

The pair of time update equations implements the mechanism of discarding the old data by applying a *variable* forgetting factor $1/(1 + c\beta_{n-1}^{-1})$ on α_{n-1} and β_{n-1}. The fact that the forgetting factor depends on the current estimate of the state stays in agreement with the behavior of the state process [recall Eq. (12)].

Since the a posteriori distribution of the state is the gamma with the shape and rate equal to α_n and $a\beta_n$, respectively, the a posteriori distribution of the local average radar cross section is the *inverse gamma*, with the shape α_n and the *scale* $a\beta_n$ [c.f. (6)]. It follows that the MMSE estimate of the local average cross section reads

$$a\mathrm{E}\left[x_n^{-1}|\mathcal{Y}_n\right] = a\int_0^\infty x_n^{-1} p(x_n|\mathcal{Y}_n)\mathrm{d}x_n = a\frac{\beta_n}{(\alpha_n - 1)} \ , \tag{21}$$

where it must hold that $\alpha_n > 1$. Observe that, when the effective number of observations is large, Eq. (21) agrees with the more straightforward estimate

$$a\mathrm{E}\left[x_n^{-1}|\mathcal{Y}_n\right] \approx a\frac{\beta_n}{\alpha_n} \ . \tag{22}$$

Prerequisites
Observation model: shape parameter a
$y_n \sim \gamma(a, x_n)$
State dynamics model: nonstationarity coefficient c
Initial distribution of the state: shape and rate parameters α_0, β_0
$x_0 \sim \gamma(\alpha_0, \beta_0)$
Interpretation of quantities
Observations y_n are RCS values obtained from the plot extractor
State variable x_n is proportional to the inverse of the local average RCS
$\mathrm{E}[y_n \vert x_n] = a x_n^{-1}$
Tracker variables $\tilde{\alpha}_n$, $\tilde{\beta}_n$ parametrize the a priori pdf of the state
$p(x_n \vert \mathcal{Y}_{n-1}) \simeq \gamma(x_n; \tilde{\alpha}_n, \tilde{\beta}_n)$
Tracker variables α_n, β_n parametrize the a posteriori pdf of the state
$p(x_n \vert \mathcal{Y}_n) \simeq \gamma(x_n; \alpha_n, \beta_n)$
For $n = 1, 2, \ldots$ execute
Time update
$\tilde{\alpha}_n = \dfrac{\alpha_{n-1}}{1 + 2c\beta_{n-1}}$
$\tilde{\beta}_n = \dfrac{\beta_{n-1}}{1 + 2c\beta_{n-1}}$
Measurement update
$\alpha_n = \tilde{\alpha}_n + a$
$\beta_n = \tilde{\beta}_n + y_n$
MMSE estimate of the local average RCS
$a \dfrac{\beta_n}{\alpha_n - 1}$

Table I
SUMMARY OF THE PROPOSED TRACKER.

The advantage of the formula (21) over (22) lies in the fact that it accounts for the extra uncertainty caused by the tail behavior of $p(x_n|\mathcal{Y}_n)$, which can be a significant factor for a small number of observations or fast changes of the state.

F. Initialization of the scheme

Assumption (A3) can be relaxed to include useful uninformative (improper) priors. The uniform prior

$$p(x_0) \propto 1, \qquad x_0 > 0$$

can be recognized as the limiting case of the gamma distribution, $\alpha_0 \to 1$, $\beta_0 \to 0$.

The Jeffreys prior, i.e., the prior which is proportional to the square root of the Fisher information, takes the form

$$p(x_0) \propto \frac{\sqrt{a}}{x_0}\,, \qquad x_0 > 0$$

which corresponds to the initialization using $\alpha_0 = \beta_0 = 0$.

G. A qualitative comparison with existing approaches

We will now compare the proposed tracker with three existing solutions, presented in [8], [10], and explain the advantages of our approach.

The tracker proposed in [8] is a part of a larger PHD filter, which precludes direct comparison with our solution. However, under the assumption that the probability of detection equals 1, the Swerling-I tracker proposed in [8] reduces to the following pair of equations

$$\begin{aligned} \hat{\sigma}_n &= \frac{\mu_{n-1} - 1}{\mu_{n-1}} \hat{\sigma}_{n-1} + \frac{y_n}{\mu_{n-1}} \\ \mu_n &= \mu_{n-1} + 1\,, \end{aligned} \tag{23}$$

where $\hat{\sigma}_n$ denotes the estimate of target average RCS, assumed to be constant by this algorithm. One can show that (23) is equivalent to the proposed tracker if the nonstationarity coefficient c is assumed to be zero. Indeed, multiplying both sides of the first recursion of (23) with μ_{n-1} yields

$$\mu_{n-1}\hat{\sigma}_n = (\mu_{n-1} - 1)\hat{\sigma}_{n-1} + y_n\,,$$

which, using the second recursion in (23), can be rewritten in the following form

$$(\mu_n - 1)\hat{\sigma}_n = (\mu_{n-1} - 1)\hat{\sigma}_{n-1} + y_n\,.$$

Set $a = 1$, $\alpha_n = \mu_n$ and $\beta_n = (\mu_n - 1)\hat{\sigma}_n$ and the equivalence of (23) and the proposed algorithm becomes apparent. Moreover, the second tracker proposed in [8], i.e., the Swerling-III tracker, is also equivalent to the proposed one ($a = 2$, $c = 0$), which can be shown using a similar argument.

Note that, unlike (23), our solution is free from the assumption that the (local) average RCS is constant, which is a considerable advantage. As argued in the introduction, the observation process y_n might be nonstationary when a maneuvering target is observed. In this situation, one should adjust the memory length of the tracker to achieve a balance between its bias and variance. Generally, increasing the tracker's memory will increase the bias, but reduce the variance, and vice versa. However, the algorithm (23) pays equal attention to the observations from the distant and the near past, i.e., it has an infinite memory, which means that the estimates $\hat{\sigma}_n$ will suffer from a substantial bias. The proposed solution, on the other hand, includes a simple, but efficient, mechanism for discarding the old data and does not suffer from such a problem.

The paper [10] also presents two trackers. The first tracker takes the form of the so-called alpha filter, i.e., a fading memory filter based on exponential forgetting. When written using our notation, this algorithm takes the following form

$$\hat{\sigma}_n = (1 - \lambda)\hat{\sigma}_{n-1} + \lambda y_n\,, \tag{24}$$

where $\hat{\sigma}_n$ denotes the estimate of the local average RCS and $0 < \lambda < 1$ is the filter gain. Comparing (24) and our solution, it straightforward to see that (24) can be regarded as a simplified version of our approach, where the time-varying forgetting factor is replaced with a constant one and the MMSE estimate of the local average RCS (21) is replaced with the formula (22).

The second tracker proposed in [10] is a simple median of last N observations, where N was chosen to be 10. Note that the median is a questionable estimator of the average RCS because it will treat spikes of y_n as outliers and reject them.

The spikes are, however, typical to the behavior of RCS. As a consequence, the median estimator can exhibit a significant downward bias.

III. Extensions

A. Forecasting the future RCS

As argued in the introduction, a statistical characterization of y_{n+1}, conditioned on the available observation history $\mathcal{Y}_n$, may be valuable for radar resource management. A complete description of one's knowledge about the behavior of y_{n+1} is encoded in the conditional density $p(y_{n+1}|\mathcal{Y}_n)$

$$p(y_{n+1}|\mathcal{Y}_n) = \int_0^\infty p(y_{n+1}|x_{n+1})p(x_{n+1}|\mathcal{Y}_n)\mathrm{d}x_n \ .$$

Substituting (8) and approximating $p(x_{n+1}|\mathcal{Y}_n)$ with $\gamma(x_{n+1}; \tilde{\alpha}_{n+1}, \tilde{\beta}_{n+1})$ leads to [42], [43]

$$p(y_{n+1}|\mathcal{Y}_n) = \beta'(y_{n+1}; a, \tilde{\alpha}_{n+1}, 1, \tilde{\beta}_{n+1}) \ , \qquad (25)$$

where

$$\beta'(x; \alpha, \beta, p, q) = \frac{p\left(\frac{x}{q}\right)^{\alpha p-1}\left[1+\left(\frac{x}{q}\right)^{p}\right]^{-\alpha-\beta}}{q\mathrm{B}(\alpha,\beta)} \qquad (26)$$

denotes the generalized beta prime distribution with parameters $\alpha > 0$, $\beta > 0$, $p > 0$, $q > 0$ and $\mathrm{B}(x, y)$ denotes the beta function [the distribution resulting from (25) is commonly called the compound gamma].

If needed, the conditional moments of (25) can be computed using the following lemma [42]

Lemma 2. *Let X be a compound gamma distributed random variable, $X \sim \beta'(x; \alpha, \beta, 1, q)$. For $-\alpha < t < \beta$ it holds that*

$$\mathrm{E}[X^t] = q^t \frac{\mathrm{B}(\alpha+t, \beta-t)}{\mathrm{B}(\alpha,\beta)} \ . \qquad (27)$$

Combining (25) with (27) yields

$$\mathrm{E}\left[y_n^t|\mathcal{Y}_{n-1}\right] = \tilde{\beta}_n^t \frac{\mathrm{B}(a+t, \tilde{\alpha}_n - t)}{\mathrm{B}(a, \tilde{\alpha}_n)} \ . \qquad (28)$$

In particular, for $t = 1$ the above formula simplifies to

$$\mathrm{E}\left[y_n|\mathcal{Y}_{n-1}\right] = a\frac{\tilde{\beta}_n}{(\tilde{\alpha}_n - 1)} \quad \text{for } \tilde{\alpha}_n > 1$$

and for $t = -1$ to

$$\mathrm{E}\left[y_n^{-1}|\mathcal{Y}_{n-1}\right] = \frac{\tilde{\alpha}_n}{\tilde{\beta}_n(a-1)} \quad \text{for } a > 1 \ .$$

B. The Interacting Multiple Model solution

The proposed tracking algorithm requires one to specify the parameters a and c. It is, however, unrealistic to assume that these quantities can be known precisely because any of them, c in particular, can be time varying.

The uncertainty of model parameters can be addressed using the interacting multiple model (IMM) approach [44]. Denote by $\bar{x}_n$ the normalized state variable,

$$\bar{x}_n = \frac{x_n}{a} \ .$$

State interaction
For $j = 1, 2, \ldots, K$
$$\psi_{n-1,j} = \sum_{i=1}^{K} \mu_{n-1,i}\, p_{ij}$$
$$\mu_{n-1,i|j} = \frac{\mu_{n-1,i}\, p_{ij}}{\psi_{n-1,j}}, \quad i = 1, 2, \ldots, K$$
$$\bar{m}^0_{n-1,j} = \sum_{i=1}^{K} \mu_{n-1,i|j} \frac{\bar{\alpha}_{n-1,i}}{\bar{\beta}_{n-1,i}}$$
$$\bar{v}^0_{n-1,j} = \sum_{i=1}^{K} \mu_{n-1,i|j}\left[\frac{\bar{\alpha}_{n-1,i}}{\bar{\beta}^2_{n-1,i}} + \left(\frac{\bar{\alpha}_{n-1,i}}{\bar{\beta}_{n-1,i}} - \bar{m}^0_{n-1,j}\right)^2\right]$$
$$\bar{\alpha}^0_{n-1,j} = \frac{\left(\bar{m}^0_{n-1,j}\right)^2}{\bar{v}^0_{n-1,j}} \qquad \bar{\beta}^0_{n-1,j} = \frac{\bar{m}^0_{n-1,j}}{\bar{v}^0_{n-1,j}}$$
Filter propagation
For $j = 1, 2, \ldots, K$
$$\tilde{\bar{\alpha}}_{n,j} = \frac{\bar{\alpha}^0_{n-1,j}}{1 + c_j\bar{\beta}^0_{n-1,j}/a_j} \qquad \tilde{\bar{\beta}}_{n,j} = \frac{\bar{\beta}^0_{n-1,j}}{1 + c_j\bar{\beta}^0_{n-1,j}/a_j}$$
$$\bar{\alpha}_{n,j} = \tilde{\bar{\alpha}}_{n,j} + a_j \qquad \bar{\beta}_{n,j} = \tilde{\bar{\beta}}_{n,j} + a_j y_n$$
Model likelihood update
For $j = 1, 2, \ldots, K$
$$\mu_{n,j} \propto \psi_{n-1,j}\beta'(y_n; a_j, \tilde{\bar{\alpha}}_{n,j}, 1, \tilde{\bar{\beta}}_{n,j}/a_j)$$
Mixing of local average RCS estimates
$$\sum_{j=1}^{K} \mu_{n,j} \frac{\bar{\beta}_{n,j}}{\bar{\alpha}_{n,j} - 1}$$

Table II
The proposed IMM solution.

The normalization facilitates implementation of the IMM scheme, because estimates of the normalized state, yielded by filters with different assumed values of the parameter a, are directly comparable. Under such a transformation of variables, it holds that $\mathrm{E}[y_n|\bar{x}_n] = \bar{x}_n^{-1}$ and $p(\bar{x}_n|\mathcal{Y}_{n-1}) \simeq \gamma(\bar{x}_n; \tilde{\bar{\alpha}}_n, \tilde{\bar{\beta}}_n)$, $p(\bar{x}_n|\mathcal{Y}_n) \simeq \gamma(\bar{x}_n; \bar{\alpha}_n, \bar{\beta}_n)$, where

$$\tilde{\bar{\alpha}}_n = \frac{\bar{\alpha}_{n-1}}{1 + c\bar{\beta}_{n-1}/a} \qquad \tilde{\bar{\beta}}_n = \frac{\bar{\beta}_{n-1}}{1 + c\bar{\beta}_{n-1}/a}$$
$$\bar{\alpha}_n = \tilde{\bar{\alpha}}_n + a \qquad \bar{\beta}_n = \tilde{\bar{\beta}}_n + a y_n \ .$$

Suppose that the behavior of the state and the observations can be governed by one of K models of the form (A1)-(A3), with parameters $(a_k,\ c_k),\ k = 1, 2, \ldots, K$. Let p_{ij}, $i = 1, 2, \ldots, K$, $j = 1, 2, \ldots, K$ denote the probability of a switch from the model i to the model j. Furthermore, assume that, when the switch occurs, the state variable x_n undergoes a renormalization by the ratio of the new and the old value of the parameter a, which ensures the normalized state $\bar{x}_n$ remains unaffected by the model transition. The IMM tracker resulting from these assumptions is summarized in Table II. For this algorithm, the predictive distribution of observations is a mixture of the predictive distributions corresponding to

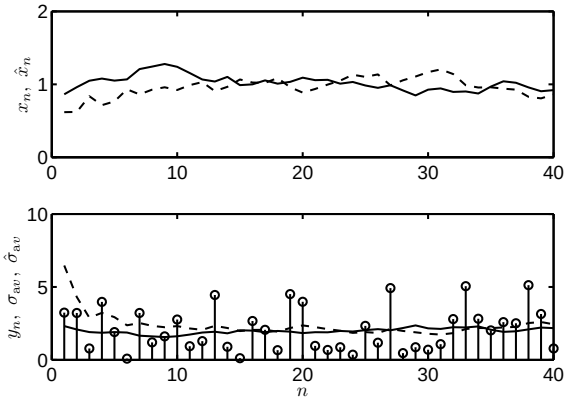

Figure 3. Typical behavior of the hidden autoregressive gamma process ($a = 2$, $c = 0.002$, $x_0 = 1$) and corresponding tracking results. Top plot: x_n (solid line) and its MMSE estimate (dashed line). Bottom plot: observations y_n (lollipops), local average RCS (solid line), and its MMSE estimate (dashed line).

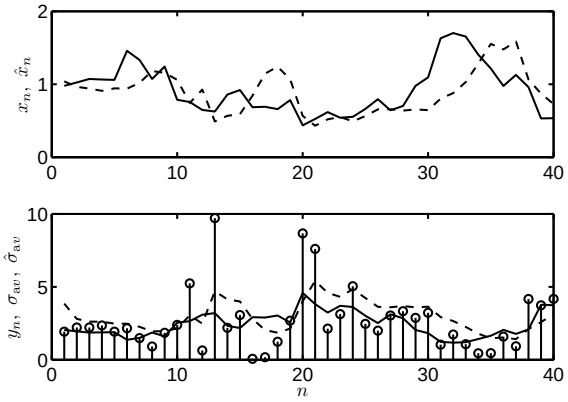

Figure 4. Typical behavior of the hidden autoregressive gamma process ($a = 2$, $c = 0.02$, $x_0 = 1$) and corresponding tracking results. Top plot: x_n (solid line) and its MMSE estimate (dashed line). Bottom plot: observations y_n (lollipops), local average RCS (solid line) and its MMSE estimate (dashed line).

each model in the model set with weights $\psi_{n,j}$

$$p(y_{n+1}|\mathcal{Y}_n) = \sum_{j=1}^{K} \psi_{n,j}\beta'(y_{n+1}; a_j, \tilde{\tilde{\alpha}}_{n+1,j}, 1, \tilde{\tilde{\beta}}_{n+1,j}/a_j) \ . \tag{29}$$

IV. Simulated and real world results

A. A basic example

Fig. 3 shows a typical realization of (x_n, y_n), obtained using (9) for $x_0 = 1$, $a = 2$ and $c = 0.002$. We also plot the MMSE estimates of x_n and of the local average RCS, computed using (20) and (21), respectively. The tracker, which employed the correct values of a and c, was initialized using the Jeffreys prior and exhibits good accuracy through the entire simulation.

Fig. 4 shows the results of a similar simulation experiment, differing in the value of c, which was increased to 0.02. Observe qualitative differences between the two presented choices of c – for $c = 0.02$ the changes in the local average RCS are much faster, which makes tracking considerably more difficult.

B. A quantitative comparison with existing approaches

We investigated how the proposed approach compares against existing solutions in quantitative terms. The following algorithms were compared:

- The proposed tracker.
- Algorithm (23).
- Algorithm (24), where the parameter λ was set to 0.1.
- The median of 10 observations.
- The Extended Kalman Filter constructed using assumptions (A1)-(A3), which takes the form

$$\begin{aligned} p_{n|n-1} &= f_n^2 p_{n-1|n-1} + q_n \\ \hat{x}_{n|n-1} &= \hat{x}_{n-1|n-1} \\ e_n &= y_n - a/\hat{x}_{n|n-1} \\ k_n &= p_{n|n-1}h_n/(h_n^2 p_{n|n-1} + r_n) \\ \hat{x}_{n|n} &= \hat{x}_{n|n-1} + k_n e_n \\ p_{n|n} &= (1 - k_n h_n)p_{n|n-1} \ , \end{aligned} \tag{30}$$

where $f_n = 1$ [c.f. (11)], $q_n = 2c\hat{x}_{n-1|n-1}$ [c.f. (12)], $h_n = -a/\hat{x}_{n|n-1}^2$, $r_n = a/\hat{x}_{n|n-1}^2$ [c.f. (A1)]. The local average RCS was estimated using the formula $a/x_{n|n}$ [c.f. (6)].

All algorithms were compared using 1000 realizations of (x_n, y_n), generated in the following way: for each realization x_0 was drawn from the gamma distribution with shape and rate equal to 20. The remaining trajectory of x_n, $1 \le n \le 100$, was generated using (9). The observation process was governed by assumption (A1) with $a = 1$.

We compared tracking accuracy and quality of forecasts generated by each algorithm. The accuracy was scored by accumulating the squared errors of local average RCS estimates. To eliminate the influence of the initial conditions on the results, we discarded twenty initial samples from the accumulation. The quality of forecasts was measured by accumulating predictive log-likelihoods of the observations in the same interval. The predictive likelihoods were obtained by substituting the actual observed values of y_{n+1} into the predictive distributions $p(y_{n+1}|\mathcal{Y}_n)$ generated by each of the compared trackers [45], [46]. For the proposed algorithm we employed Eq. (25). In case of the EKF, $p(y_{n+1}|\mathcal{Y}_n)$ takes the form of the Gaussian distribution with mean $a/\hat{x}_{n|n-1}$ and variance s_n. The remaining algorithms were not fitted with the forecasting mechanism and were modified to include one.

Owing to the discussion presented in Section II.G, extending algorithms (23) and (24) was straightforward. For (23), one can use $\beta'(y_{n+1}; a, \mu_n, 1, (\mu_n - 1)\hat{\sigma}_n)$. In case of (24), one can add the second recursion

$$\mu_n = (1 - \lambda)\mu_{n-1} + \lambda$$

and employ $\beta'(y_{n+1}; a, \mu_n, 1, \mu_n\hat{\sigma}_n)$ as the predictive distribution.

Finally, for the median filter, the inverse of its output served as the estimate of the state and was substituted into (5), which neglects the uncertainty of such an estimate.

Table III shows the comparison of mean accumulated squared local average RCS estimation errors yielded by all

c	Proposed	Alg. (23)	Alg. (24)	Median	EKF
0.0000	**1.34**	1.34	4.95	15.6	1.41
0.0001	**1.96**	2.04	5.19	15.9	2.10
0.0002	**2.57**	2.92	5.53	16.6	2.65
0.0005	**4.53**	6.71	6.93	20.3	5.43
0.0010	**10.2**	20.9	11.1	29.4	—
0.0020	**34.1**	86.2	35.1	65.9	—

Table III
COMPARISON OF MEAN ACCUMULATED SQUARED LOCAL AVERAGE RCS ESTIMATION ERRORS YIELDED BY THE PROPOSED TRACKER, ALGORITHMS (23), (24), THE MEDIAN FILTER, AND THE EXTENDED KALMAN FILTER, FOR SEVERAL VALUES OF c.

c	Proposed	Alg. (23)	Alg. (24)	Median	EKF
0.0000	**-82.9**	-82.9	-84.5	-88.1	-121
0.0001	**-83.1**	-83.1	-84.5	-88.5	-122
0.0002	**-84.0**	-84.1	-85.3	-88.9	-124
0.0005	**-86.6**	-87.0	-87.6	-91.7	-130
0.0010	**-88.9**	-90.2	-89.5	-93.8	—
0.0020	**-94.0**	-97.6	-94.3	-98.3	—

Table IV
COMPARISON OF MEAN ACCUMULATED PREDICTIVE LOG-LIKELIHOODS OBTAINED WITH THE PROPOSED TRACKER, ALGORITHMS (23), (24), THE MEDIAN FILTER, AND THE EXTENDED KALMAN FILTER, FOR SEVERAL VALUES OF c.

algorithms for several values of c. Table IV shows the corresponding mean accumulated predictive log-likelihood scores.

In all cases, the proposed approach outperforms the other ones. As could be expected, algorithm (23) works well only for slow changes of the process parameters, i.e., for small values of c. Algorithm (24), on the other hand, starts poorly but gradually catches up with increasing c. The median filter does very poorly at estimating the local average RCS and not too well in generating forecasts. The EKF works reasonably well at tracking until $c = 0.001$, when its performance breaks down. However, it consistently fails at forecasting because the Gaussian distribution matches the behavior of y_n poorly. We verified that one can obtain considerably better results by forming the predictive distribution using $\hat{x}_{n|n-1}$ and (5), but such a solution still lags behind the proposed approach.

C. The IMM tracker – a basic example

To demonstrate the advantages of the IMM estimator, we will use the following scenario. The behavior of x_n is governed by (A2) with $c = 0.002$ at all times except $n = 20$ when

$$x_{20} = \frac{x_{19}}{3}$$

is enforced. The initial condition is set to $x_0 = 2$. The observations are governed by (8), where $a = 2$ at all times. Under such setup, the behavior of the state variable at $n = 20$ corresponds to an abrupt, three-fold, increase of the local average RCS.

The IMM structure was built around four models, specified by the following values of a and c: $(a_1, c_1) = (1, 0.002)$, $(a_2, c_2) = (1, 0.02)$, $(a_3, c_3) = (2, 0.002)$ and $(a_4, c_4) = (2, 0.02)$. The probabilities of the model switch p_{ij} are set to 0.025 ($i \neq j$) and 0.925 ($i = j$), respectively.

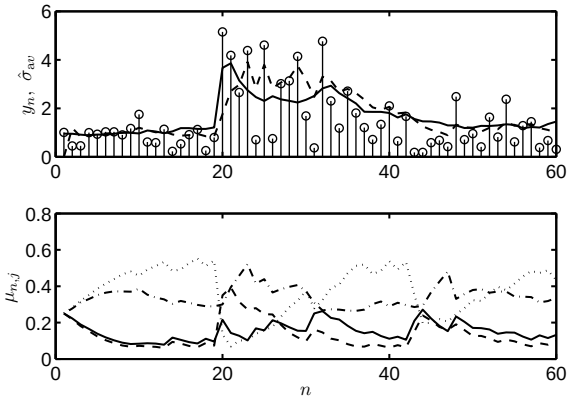

Figure 5. The behavior of the IMM estimator under an abrupt change of the local average RCS. Top plot – observations, true values of the local average RCS (solid line) and their IMM estimates (dashed line). Bottom plot – model likelihoods: solid line $a = 1$, $c = 0.002$; dashed line $a = 1$, $c = 0.02$; dotted line $a = 2$, $c = 0.002$; dash-dotted line $a = 2$, $c = 0.02$.

Fig. 5 shows the behavior of the IMM estimator observed for this scenario. Inspection of the model likelihoods around $n = 20$ confirms their proper reaction to the abrupt change. Furthermore, the IMM structure achieves better accuracy than any of the four constituting trackers alone. In the simulation shown, the accumulated squared RCS estimation error yielded by the IMM solution was equal to 15.6, against 19.4, 55.3, 20.5 and 18.1 obtained using (a_1, c_1), (a_2, c_2), (a_3, c_3) and (a_4, c_4), respectively.

D. The IMM tracker applied to a real-world dataset

Fig. 6 shows the behavior of the proposed IMM tracker for real-world data, collected by observing a cooperative target. For simplicity, all settings of the tracker were left unchanged from the last simulation example.

We will start with the qualitative inspection of the data and the behavior of the tracker. Consider the fragment between $n_1 = 20$ and $n_2 = 35$, for example. The average RCS during this period varies rather slowly, and the distribution of the observations seems closer to Swerling-III than to Swerling-I. The IMM scheme recognizes this situation correctly – the likelihoods of the two models with $a = 2$ are larger than for the models with $a = 1$. Near the end of that fragment, the likelihood of the model with $a = 2$ and $c = 0.002$ exceeds 0.6. Observe, however, that when the RCS decreases sharply between $n_2 = 35$ and $n_3 = 50$, the likelihood of the model with $a = 2$ and $c = 0.002$ decreases instantly.

Another interesting part occurs between $n_4 = 80$ and $n_5 = 120$, where the observations become very spiky. Observe a nearly instantaneous reaction of the model likelihoods to this change of behavior.

Since the true values of the local average RCS are unknown for real-world datasets, we can quantify only the quality of the forecasts generated by the algorithm. As in Section IV.B, the assessment can be based on the cumulative predictive log-likelihood scores. For the data shown, the IMM algorithm's score was -221, while algorithms (23) and (24) could only reach -263 and -226, respectively. These values confirm that

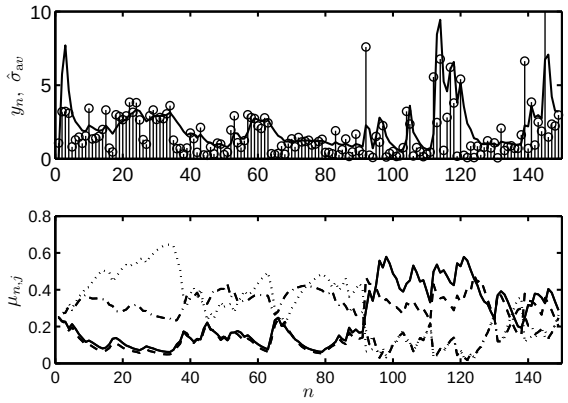

Figure 6. The behavior of the IMM estimator for real-world measurements. Top plot – observations and estimates of the local average RCS. Bottom plot – model likelihoods: solid line $a = 1$, $c = 0.002$; dashed line $a = 1$, $c = 0.02$; dotted line $a = 2$, $c = 0.002$; dash-dotted line $a = 2$, $c = 0.02$.

the forecasts generated using the proposed approach are superior not only in the simulations but also for the real-world data.

V. Conclusions

An approximate Bayesian tracker of radar cross section was proposed. The scheme exploits the conjugate prior property of the gamma family by modeling the state using an autoregressive gamma process and applying a nontrivial gamma approximation during the time update step. The analysis of the tracker allowed us to formulate noteworthy observations about the existing RCS tracking solutions. Finally, to cope with the unknown process parameters, the interacting multiple model extension of the basic scheme was proposed. The behavior of all proposed solutions was verified using simulations and a real-world dataset.

References

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

[2] Skolnik, M., *Radar Handbook*, 3rd ed. McGraw-Hill, 2008.

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

[4] Lerro, D. and Bar-Shalom, Y., "Automated tracking with target amplitude information," in *Proc. American Control Conference*, San Diego, CA, 1990, pp. 2875—-2880.

[5] Clark, D., Ristić, B., Vo, B.-N., and Vo, B., "Bayesian multi-object filtering with amplitude feature likelihood for unknown object SNR," *IEEE Transactions on Signal Processing*, vol. 58, no. 1, pp. 26–37, 2010.

[6] van Keuk, G., "Multihypothesis tracking using incoherent signal-strength information," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 32, no. 3, pp. 1164—-1170, 1996.

[7] Mertens, M. and Ulmke, M., "Ground target tracking with RCS estimation utilizing probability hypothesis density filters," in *Proceedings of the 16th International Conference on Information Fusion*, 2013, pp. 2145–2152.

[8] Mertens, M., Ulmke, M., and Koch, W., "Ground target tracking with RCS estimation based on signal strength measurements," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 52, no. 1, pp. 205–220, 2016.

[9] Ehrman, L. M. and Blair, W. D., "Comparison of methods for using target amplitude to improve measurement-to-track association in multi-target tracking," in *Proc. 9th Intenrnational Conference on Information Fusion*, Florence, Italy, 2006.

[10] Dale Blair, W., Kramer, J., and Miceli, P., "Use of radar cross section in track-to-truth assignment for assessment of multiple target tracking algorithms," in *2014 IEEE Aerospace Conference*, 2014, pp. 1–8.

[11] Pisane, J., "Automatic target recognition using passive bistatic radar signals," Ph.D. dissertation, University of Liége, 2013.

[12] Lee, Y., Choos, H., Kim, S., and Kim, H., "RCS based target recognition with real FMCW radar implementation," *Microwave and Optical Technology Letters*, vol. 58, no. 7, pp. 1745–1750, 2016.

[13] Ehrman, L. and Lanterman, A., "A robust algorithm for automatic target recognition using passive radar," in *Proc. thirty-sixth southeastern symposium on system theory*, 2004, pp. 102–106.

[14] Jiang, H., Xu, L., and Zhan, K., "Joint tracking and classification based on aerodynamic model and radar cross section," *Pattern recognition*, vol. 47, no. 9, pp. 3096–3105, 2014.

[15] Bell, K. L., Johnson, J. T., Smith, G. E., Baker, C. J., and Rangaswamy, M., "Cognitive radar for target tracking using a software defined radar system," in *2015 IEEE Radar Conference (RadarCon)*, 2015, pp. 1394–1399.

[16] Meller, M., "Radar time budget optimization subject to angle accuracy constraint via cognitive approach," in *Proc. 18th International Radar Symposium (IRS 2017)*, Prague, Czech Republic, 2017.

[17] Han, Q., Pan, M., and Liang, Z., "Joint power and beam allocation of opportunistic array radar for multiple target tracking in clutter," *Digital Signal Processing*, vol. 78, pp. 136–151, 2018.

[18] Ochodnicky, J., Matousek, Z., Sostronek, M., and Hykel, A., "Radar cross section measurement by subscale models," in *2008 International Radar Symposium*, Wroclaw, Poland, 2008.

[19] Knott, E. F., *Radar Cross Section Measurements*. New York: Van Nostrand Reinhold, 1993.

[20] Dybdal, R. B., "Radar cross section measurements," *Proceedings of the IEEE*, vol. 75, no. 4, pp. 498–516, 1987.

[21] Crispin, J. W. and Maffett, A. L., "Radar cross-section estimation for simple shapes," *Proceedings of the IEEE*, vol. 53, no. 8, pp. 833–848, 1965.

[22] ——, "Radar cross-section estimation for complex shapes," *Proceedings of the IEEE*, vol. 53, no. 8, pp. 972–982, 1965.

[23] White, R. G., "Simulated annealing algorithm for radar cross-section estimation and segmentation," *Proc. SPIE*, vol. 2243, pp. 231–240, 1994.

[24] Crispin, J. J., *Methods of radar cross-section analysis*. Elsevier, 2013.

[25] Swerling, P., "Probability of detection for fluctuating targets," *IRE Transactions on Information Theory*, vol. 6, no. 2, pp. 269–308, 1960.

[26] ——, "Radar probability of detection for some additional fluctuating target cases," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 33, no. 2, pp. 698–709, 1997.

[27] van Trees, H. L., Bell, K. L., and Tian, Z., *Detection, Estimation and Modulation Theory, Part I*. John Wiley & Sons, 2011.

[28] Stone, L. D., *Handbook of Multisensor Data Fusion*. Boca Raton, FL: CRC Press, 2001, chapter A Bayesian approach to multiple-target tracking.

[29] Arulampalam, M. S., Maskell, S., Gordon, N., and Clapp, T., "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 174–188, 2002.

[30] Doucet, A. and Johansen, A. M., "A tutorial on particle filtering and smoothing: Fifteen years later," Institute of Statistical Mathematics, Japan and Department of Statistics, University of Warwick, Tech. Rep., 2008.

[31] Fink, D., "A compendium of conjugate priors," Environmental Statistics Group, Department of Biology, Montana State University, Tech. Rep., 1997.

[32] Anderson, B. D. O. and Moore, J. B., *Optimal Filtering*. Englewood Cliffs, New Jersey: Prentice–Hall, 1979.

[33] Julier, S. J. and Uhlmann, J. K., "Unscented filtering and nonlinear estimation," *Proceedings of the IEEE*, vol. 92, no. 3, pp. 401–422, 2004.

[34] Arasaratnam, I. and Haykin, S., "Cubature Kalman filters," *IEEE Transactions on Automatic Control*, vol. 54, no. 6, pp. 1254–1269, 2009.

[35] Farina, A., Benvenuti, D., and Ristic, B., "Estimation accuracy of a landing point of a ballistic target," in *Proceedings of the Fifth International Conference on Information Fusion. (FUSION 2002)*, vol. 1, 2002, pp. 2–9 vol.1.

[36] Novanda, H., Regulski, P., González-Longatt, F. M., and Terzija, V., "Unscented Kalman filter for frequency and amplitude estimation," in *2011 IEEE Trondheim PowerTech*, 2011, pp. 1–6.

[37] Dong, Y., Li, Y., Xiao, M., and Lai, M., "Unscented Kalman filter for time varying spectral analysis of earthquake ground motions," *Applied Mathematical Modelling*, vol. 33, no. 1, pp. 398 – 412, 2009.

[38] Zhang, D., Deng, Z., Wang, B., and Fu, M., "The application of Square-Root Cubature Kalman Filter in the SINS/CNS integrated navigation system," in *2016 IEEE Chinese Guidance, Navigation and Control Conference (CGNCC)*, 2016, pp. 2331–2335.

[39] Gourieroux, C. and Jasiak, J., "Autoregressive gamma processes," *Journal of Forecasting*, vol. 25, no. 2, pp. 129–152, 2006.

[40] Sim, C. H., "First order autoregressive models for gamma and exponential processess," *Journal of Applied Probability*, vol. 27, pp. 325–332, 1990.

[41] ——, "Modelling non-normal first-order autoregressive time series," *Journal of Forecasting*, vol. 13, no. 4, pp. 369–381, 1994.

[42] Dubey, S. D., "Compound gamma, beta and F distributions," *Metrika*, vol. 16, no. 1, pp. 27–31, 1970.

[43] Johnson, N. L., Kotz, S., and Balakrishnan, N., *Continuous Univariate Distributions, Volume 1*, 2nd ed. John Wiley & Sons, 1994.

[44] Bar-Shalom, Y., Chang, K. C., and Blom, H. A. P., "Tracking a maneuvering target using input estimation versus the interacting multiple model algorithm," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 25, no. 2, pp. 296–300, 1989.

[45] Dawid, A. P., "Present position and potential developments: some personal views. Statistical theory. The prequential approach," *Journal of the Royal Statistical Society: Series A*, vol. 147, no. 2, pp. 278–292, 1984.

[46] ——, "Prequential data analysis," in *Current Issues in Statistical Inference: Essays in Honor of D. Basu*, 1992, p. 113–126.

APPENDIX A
DERIVATION OF (10)-(12)

We start by deriving $\mathcal{L}(p_\delta(x_n|x_{n-1}))$, the Laplace transform of $p_\delta(x_n|x_{n-1})$. Applying $\mathcal{L}$-transform to (7) yields

$$\mathcal{L}(p_\delta(x_n|x_{n-1})) = e^{-bx_{n-1}} \sum_{k=0}^{\infty} \frac{b^k x_{n-1}^k}{c^{\delta+k}\Gamma(\delta+k)} \frac{L_k(s)}{k!} , \quad (31)$$

where

$$L_k(s) = \mathcal{L}\left[\exp\left(-\frac{x_n}{c}\right) x_n^{\delta+k-1}\right] .$$

Using the substitution $y_n = (s+1/c)x_n$ yields, after some straightforward manipulations,

$$\begin{aligned} L_k(s) &= \int_0^\infty \exp\left(-\frac{x_n}{c}\right) x_n^{\delta+k-1} e^{-sx_n} \mathrm{d}x_n \\ &= \frac{c^{\delta+k}}{(sc+1)^{\delta+k}} \Gamma(\delta+k) . \end{aligned} \quad (32)$$

Substituting $L_k(s)$ into (31), after canceling and rearranging the remaining common terms, leads to the following formula

$$\mathcal{L}(p_\delta(x_n|x_{n-1})) = \frac{e^{-bx_{n-1}}}{(sc+1)^\delta} \sum_{k=0}^{\infty} \frac{b^k x_{n-1}^k}{(sc+1)^k k!} , \quad (33)$$

where the sum is immediately recognizable as the exponent function

$$\begin{aligned} \mathcal{L}(p_\delta(x_n|x_{n-1})) &= \frac{e^{-bx_{n-1}}}{(sc+1)^\delta} \exp\left(\frac{bx_{n-1}}{sc+1}\right) \\ &= \frac{1}{(sc+1)^\delta} \exp\left(-\frac{scbx_{n-1}}{sc+1}\right) . \end{aligned}$$

Equation (10) is obtained with $\delta \to 0$ and $cb = 1$.

The moments of $p(x_n|x_{n-1})$ can be computed using the following property

$$\mathcal{L}(p(x_n|x_{n-1})) = \mathrm{E}\left[e^{-sx_n}|x_{n-1}\right] .$$

It follows immediately that

$$\mathrm{E}[x_n^k|x_{n-1}] = (-1)^k \frac{\mathrm{d}^k}{\mathrm{d}s^k} \exp\left(-\frac{sx_{n-1}}{sc+1}\right)\Big|_{s\to 0} .$$

Equation (11) stems directly from the above equality ($k = 1$). Reaching Eq. (12) requires one to combine the results obtained for $k = 1$ and $k = 2$.

APPENDIX B
DERIVATION OF (17)

Let

$$k(s) = \frac{sc+1}{s(c+\beta_{n-1}^{-1})+1} .$$

Since $K(s) = [k(s)]^{\alpha_{n-1}}$, it holds that

$$\frac{\mathrm{d}}{\mathrm{d}s}K(s) = \alpha_{n-1}[k(s)]^{\alpha_{n-1}-1} \frac{\mathrm{d}}{\mathrm{d}s}k(s) , \quad (34)$$

where

$$\frac{\mathrm{d}}{\mathrm{d}s}k(s) = \frac{-\beta_{n-1}^{-1}}{[s(c+\beta_{n-1}^{-1})+1]^2} .$$

Substituting $s = 0$ into $k(s)$ and its derivative, one obtains the following two equalities

$$k(0) = 1 \qquad \frac{\mathrm{d}}{\mathrm{d}s}k(s)\Big|_{s\to 0} = -\frac{1}{\beta_{n-1}} \quad (35)$$

which, when combined with Eq. (34), lead to

$$-\frac{\mathrm{d}}{\mathrm{d}s}K(s)\Big|_{s\to 0} = \frac{\alpha_{n-1}}{\beta_{n-1}} .$$

Note that this result could be alternatively derived by invoking (11) and evaluating its expectation over x_{n-1}.

The second derivative of $K(s)$ reads

$$\begin{aligned} \frac{\mathrm{d}^2}{\mathrm{d}s^2}K(s) = -\frac{\alpha_{n-1}}{\beta_{n-1}} \Bigg\{ & \frac{(\alpha_{n-1}-1)[k(s)]^{\alpha_{n-1}-2}}{[s(c+\beta_{n-1}^{-1})+1]^2} \frac{\mathrm{d}}{\mathrm{d}s}k(s) \\ & - 2[k(s)]^{\alpha_{n-1}} \frac{c+\beta_{n-1}^{-1}}{[s(c+\beta_{n-1}^{-1})+1]^3} \Bigg\} . \end{aligned}$$

Evaluating $\mathrm{d}^2K(s)/\mathrm{d}s^2$ for $s \to 0$, substituting (35), and performing some straightforward simplifications allows one to obtain the second formula in Eq. (17).

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 architect of digital systems. In 2018, in recognition of his outstanding industrial achievements, he was awarded 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.

He is a member of the NATO SET-227 Research Task Group on "Cognitive Radar".