Reducing Bias in Bayesian Shape Estimation

Florian Faion, Antonio Zea, and Uwe D. Hanebeck

Intelligent Sensor-Actuator-Systems Laboratory (ISAS) Institute for Anthropomatics and Robotics Karlsruhe Institute of Technology (KIT), Germany florian.faion@kit.edu, antonio.zea@kit.edu, uwe.hanebeck@ieee.org

Abstract—This work considers the problem of estimating the parameters of an extended object based on noisy point observations from its boundary. The intention is to explore relationships between common approaches by breaking them down into their basic assumptions within the Bayesian framework. In doing so, we find that distance-minimizing curve fitting algorithms can be modeled by using a special Spatial Distribution Model, where the source distribution is approximated by a greedy one-to-one association of points to sources on the shape boundary. Based on this insight, we explore the origin of the estimation bias, which is a well-known issue of curve fitting algorithms. Furthermore, we derive a general scheme to alleviate its effect for arbitrary shapes, as well as for non-isotropic noise. This procedure is shown to be a generalization of related special solutions.

Keywords—Bias reduction, shape fitting, Spatial Distribution Model, Linear Regression Kalman Filter, non-isotropic noise.

I. INTRODUCTION

Estimating the pose and shape of an extended object based on noisy point observations from its boundary is an essential task of many tracking algorithms. Applications can be found in fields related to robotics, industries, medicine, entertainment, automotive, surveillance, and telepresence. Relevant sensors include laser range finders, depth cameras, or radar devices.

Approaches to shape estimation can be generally distinguished in terms of the generative models they assume. Two popular classes are given by *Spatial Distribution Models* and *implicit models*, as used in curve fitting. Our intention is to explore the relationship between both classes by investigating how they associate the noisy points to their generating sources on the shape and how this association affects the estimation result.

The first class [1], [2] assumes that points are randomly drawn from a known distribution along the boundary and considers each source, in order to avoid explicit one-to-one association. This class also includes the *Random Matrices* [3], and, partially the *Random Hypersurface Models* [4]. The implicit models from the second class do not depend on the knowledge of a spatial distribution but greedily establish correspondences from noisy points to sources on the shape boundary whose distance then is minimized by the estimator. Relevant expressions include the Euclidean distance [5], Mahalanobis distance [6], radial distance [4], squared or signed distance [2], [7], up to abstract algebraic expressions for circles [9], [10] and conics [5].

As a result of our investigations, we establish the nontrivial relationship that the implicit models used in distanceminimizing approaches can be interpreted as Spatial Distribution Models with special assumptions. Furthermore, we



(a) Common estimator.

(b) Proposed estimator.

Figure 1: Common distance-minimizing estimators (a) are biased in the case of noise. This paper explores the bias and proposes an alleviation scheme (b).

consider the problem of bias in the estimated parameters in the presence of noise, which is a known drawback when using these implicit models. Analyzing the origin of this bias lets us conclude a general scheme to derive a bias-reduced estimator for arbitrary shapes that works even in the case of non-isotropic noise. This scheme is inspired by ideas from [11] and intuitively motivated from a geometric point of view. There are several related approaches that propose bias reduction for specific shapes, e.g., circles [7], [10], ellipses [5], [6], [11], and curve fitting [12]. General considerations on bias in recursive filtering are presented in [13] and [14].

Summed up, in this work we 1) show that the implicit models used in distance-minimizing approaches are a special case of Spatial Distribution Models, and 2) derive a bias-reduced estimator for arbitrary shapes that works even in the case of non-isotropic noise. The improvement achieved by this estimator is illustrated in Fig. 1 for a spherical shape. Note that wider considerations specific to tracking, such as motion models [15], or the treatment of clutter measurements [1], [16], [17] are out of the scope of this paper. Another important aspect, reserved for future work, are filled shapes where measurements may originate from the inside of the object.

II. PROBLEM STATEMENT

We consider the following specific instance of shape estimation problem.

Given: A set of noisy point measurements $\hat{\mathcal{Z}} = \{\hat{\underline{z}}_i | i = 1, \ldots, n\}$ is observed from a static object, where the points are from the domain \mathbb{R}^d and d is usually two or three. The object is characterized by a set of sources $\underline{\tilde{z}}_i \in \tilde{\mathcal{Z}}$ that



Figure 2: Problem Statement: Find the object parameters that most likely have generated the points (blue crosses). Red dots mark source points on the object.

together form its boundary. We assume that each point $\underline{\hat{z}}_i$ is an independent observation of a specific, but unknown source $\underline{\tilde{z}}_i$ according to

$$\underline{\hat{z}}_i = \underline{\tilde{z}}_i + \underline{w}_i \ . \tag{1}$$

The additive noise term \underline{w}_i is drawn from the Gaussian distribution $\mathcal{N}(\underline{0}, \mathbf{C}_{w_i})$ with known, sensor-specific covariance matrix \mathbf{C}_{w_i} . In general, the sources cannot be recovered from the points, which is known as the *association problem*. Both the generation process and association problem are illustrated in Fig. 2.

Desired: Let the parameter vector \underline{x} , also referred to as the *state*, parameterize the set of sources \tilde{Z}_x of the desired object. The estimation task now is to find the specific instance of parameters that most likely have generated the measurements \hat{Z} . Besides position and orientation, \underline{x} can contain information related to scale, aspect ratio, and any other parameter related to the appearance.

III. BAYESIAN SHAPE ESTIMATION

Usually, the relationship between object parameters \underline{x} and measurements $\hat{\mathcal{Z}}$ is expressed in terms of the *likelihood* $p(\hat{\mathcal{Z}}|\underline{x})$. Approaches without prior knowledge of the parameters \underline{x} usually calculate the instance of parameters $\underline{x}_{\text{ML}}$ that yields the *maximum likelihood*

$$\underline{x}_{\mathrm{ML}} = \operatorname*{arg\,max}_{\underline{x}} \ p(\hat{\mathcal{Z}}|\underline{x}) \ .$$

In the Bayesian paradigm, a prior distribution $p(\underline{x})$ on the parameters is incorporated in the estimator and yields a *maximum a posteriori* estimate

$$\underline{x}_{\text{MAP}} = \underset{\underline{x}}{\arg\max} \ p(\hat{\mathcal{Z}}|\underline{x}) \cdot p(\underline{x}) \ .$$

Based on Bayes' rule, we can also keep the posterior distribution of parameters

$$p(\underline{x}|\mathcal{Z}) \propto p(\mathcal{Z}|\underline{x}) \cdot p(\underline{x})$$

which allows for developing a recursive Bayesian estimator, by using the posterior distribution as new prior. Note that all these estimation techniques require the likelihood $p(\hat{z}|\underline{x})$ to be defined. This likelihood can be derived as follows.

If we assume all points $\underline{\hat{z}}_i \in \hat{\mathcal{Z}}$ to be measured independently, it allows for separately considering each measurement according to

$$p(\hat{\mathcal{Z}}|\underline{x}) = \prod_{i} p(\hat{\underline{z}}_{i}|\underline{x}) .$$
⁽²⁾



Figure 3: Different concepts to derive a likelihood $p(\underline{\hat{z}}|\underline{x})$ from the individual $p(\underline{\hat{z}}|\underline{x},s)$ from (a). An SDM (b) assumes a distribution over s. A GAM (c) exclusively selects one individual likelihood.

Thus, we only need to derive the likelihood $p(\underline{\hat{z}}_i | \underline{x})$ for every individual measurement $\underline{\hat{z}}_i$ and can use (2) for aggregation. The additive noise model from (1) can be written as the convolution

$$p(\underline{\hat{z}}|\underline{x}) = \int_{\mathbb{R}^d} p(\underline{\hat{z}}|\underline{z}) \cdot p(\underline{z}|\underline{x}) \, \mathrm{d}\underline{z} \tag{3}$$

of a sensor model $p(\underline{\hat{z}}|\underline{z})$ and a source model $p(\underline{z}|\underline{x})$. The former model specifies the distribution of measurements for a point \underline{z} and is immediately given by the Gaussian distribution $p(\underline{\hat{z}}|\underline{z}) = \mathcal{N}(\underline{\hat{z}} - \underline{z}, \mathbf{C}_w)$. The latter model specifies the distribution of points \underline{z} to be measurement sources for a given state \underline{x} . Let $s \in S$ be an index parameter that iterates through all possible sources $\underline{\tilde{z}}_{x,s} \in \underline{\tilde{Z}}_x$ for a given state \underline{x} . Then, for each s, an individual source model can be modeled by the Dirac delta distribution $p(\underline{z}|\underline{x}, s) = \delta(\underline{z} - \underline{\tilde{z}}_{x,s})$.

Taking s into account, this source model and the Gaussian sensor model can be plugged into (3). Then, applying the sifting property of the Dirac delta distribution yields the following individual likelihood for each s

$$p(\underline{\hat{z}}|\underline{x},s) \stackrel{(3)}{=} \int_{\mathbb{R}^d} p(\underline{\hat{z}}|\underline{z}) \cdot p(\underline{z}|\underline{x},s) \, \mathrm{d}\underline{z}$$
$$= \int_{\mathbb{R}^d} \mathcal{N}(\underline{\hat{z}} - \underline{z}, \mathbf{C}_w) \cdot \delta(\underline{z} - \underline{\tilde{z}}_{x,s}) \, \mathrm{d}\underline{z}$$
$$= \mathcal{N}(\underline{\hat{z}} - \underline{\tilde{z}}_{x,s}, \mathbf{C}_w) \, . \tag{4}$$

The meaning of these individual likelihoods is illustrated in Fig. 3a. For each $s \in S$, the observed point \hat{z} is treated as if it would exclusively originate from the particular source $\tilde{z}_{x,s}$. In other words, each *s* refers to an individual association hypothesis. Next, in order to deal with these hypotheses, there are two strategies: 1) Keeping all hypotheses and obtaining a set of likelihoods and posterior distributions, or 2) Fusing all hypotheses and obtaining a single likelihood and posterior distribution. In this work, we exclusively focus on the second strategy, which leads to Spatial Distribution Models.

A. Spatial Distribution Model (SDM)

A Spatial Distribution Model (SDM) assumes a distribution p(s) over all hypotheses and uses marginalization over all (4)

to derive the single likelihood

$$p(\underline{\hat{z}}|\underline{x}) = \int_{\mathcal{S}} p(\underline{\hat{z}}|\underline{x}, s) \cdot p(s) \, \mathrm{d}s$$

$$\stackrel{(4)}{=} \int_{\mathcal{S}} \mathcal{N}(\underline{\hat{z}} - \underline{\tilde{z}}_{x,s}, \mathbf{C}_w) \cdot p(s) \, \mathrm{d}s .$$
(5)

Using SDMs is widely used in fitting [2] and tracking [1], [16], [18] applications. As a short remark, modeling an ellipsoidal extended object by means of *Random Matrices* [3] is also related to the SDM, as sources are assumed to be Gaussian distributed around the object center.

An interesting special case is shown in Fig. 3b, where a uniform distribution $\mathcal{U}(S)$ over $s \in S$ is assumed, which lets us rewrite (5) to

$$p(\underline{\hat{z}}|\underline{x}) \propto \int_{\mathcal{S}} p(\underline{\hat{z}}|\underline{x}, s) \, \mathrm{d}s \qquad (6)$$
$$\stackrel{(4)}{=} \int_{\mathcal{S}} \mathcal{N}(\underline{\hat{z}} - \underline{\tilde{z}}_{x,s}, \mathbf{C}_w) \, \mathrm{d}s \; .$$

This can be imagined as integrating a Gaussian with mean \hat{z} and covariance C_w along the shape boundary. We will show in the evaluation that estimation based on an SDM is the method of choice, as long as the correct distribution p(s) is used. However, in most real-life scenarios, p(s) is not known in advance, and deriving it is a non-trivial task, as it depends on factors as the sensor to object geometry, the specific segmentation algorithm, and occlusions, among others. This raises the need for a model that depends less on the source distribution.

B. Greedy Association Model (GAM)

Let us assume $\underline{\tilde{z}}_{x,s'}$ would be the true source that has generated the measurement $\underline{\hat{z}}$. Based on this assumption, we could specify p(s) as illustrated in Fig. 3c, where all probability mass is exclusively assigned to s' according to

$$p(s) = \delta(s - s') . \tag{7}$$

Then, plugging (7) into the generic SDM from (5) lets us eliminate all other association hypotheses

$$p(\underline{\hat{z}}|\underline{x}) \stackrel{(7)}{=} \int_{\mathcal{S}} \mathcal{N}(\underline{\hat{z}} - \underline{\tilde{z}}_{x,s}, \mathbf{C}_w) \cdot \delta(s - s') \, \mathrm{d}s \qquad (8)$$
$$= \mathcal{N}(\underline{\hat{z}} - \underline{\tilde{z}}_{x,s'}, \mathbf{C}_w) \, .$$

Note that (8) imposes that $\underline{\tilde{z}}_{x,s'}$ must have generated the measurement $\underline{\hat{z}}$.

However, as the true source $\underline{\tilde{z}}_{x,s'}$ is usually unknown, an appropriate approximation has to be found. A reasonable choice would be selecting the "best" source, in the sense that it yields the highest individual likelihood

$$\underline{\tilde{z}}_{x,s'} := \underset{\underline{\tilde{z}}_{x,s} \in \tilde{\mathcal{Z}}_x}{\arg \max} \, \mathcal{N}(\underline{\hat{z}} - \underline{\tilde{z}}_{x,s}, \mathbf{C}_w) \, . \tag{9}$$

In doing so, a greedy association is established between the measurement $\hat{\underline{z}}$ and its most likely source $\tilde{\underline{z}}_{x,s'}$. Based on these considerations, we define a *Greedy Association Model* (GAM) to be an SDM, where all probability is greedily assigned to the single source $\tilde{\underline{z}}_{x,s'}$ that is selected according to (9). It is worth mentioning that the greedy association becomes increasingly more correct for decreasing noise. In the following, we show that the GAM will directly lead us to the class of implicit models used in distance-minimizing estimators.



Figure 4: Projection for isotropic and non-isotropic noise.



Figure 5: Finding the source $\underline{\tilde{z}}_{x,s'}$ for a point $\underline{\hat{z}}$ in a GAM- π .

C. Relation to Projections

As already mentioned in the previous section, using the GAM from (8) and (9) builds the theoretical basis for common distance-minimizing estimators. Let us now rearrange this model in order to illustrate this relationship. As a consequence of the Gaussian sensor model, finding the most likely source $\underline{\tilde{z}}_{x,s'}$ in (9) is equivalent to finding the closest source to the point $\underline{\hat{z}}$ in terms of the Mahalanobis distance.

Definition 1 (Projection Function)

Let $\underline{\hat{z}}$ be a point measurement with covariance \mathbf{C}_w and let \mathcal{Z}_x be a set of sources, specified by the parameter vector \underline{x} . Then, the projection π of the point $\underline{\hat{z}}$ onto $\overline{\mathcal{Z}}_x$ is defined to be the closest source $\underline{\tilde{z}}_{x,s} \in \overline{\mathcal{Z}}_x$ in terms of the Mahalanobis distance

$$\pi(\underline{\hat{z}}, \underline{x}, \mathbf{C}_w) := \operatorname*{arg\,min}_{\underline{\tilde{z}}_{x,s} \in \tilde{\mathcal{Z}}_x} \left(\underline{\hat{z}} - \underline{\tilde{z}}_{x,s}\right)^{\mathrm{T}} \mathbf{C}_w^{-1} \left(\underline{\hat{z}} - \underline{\tilde{z}}_{x,s}\right) .(10)$$

Note that in case of isotropic noise, the projection coincides with the Euclidean projection. In the case of multiple equivalent projection candidates (with equal distance to \hat{z}), one has to be chosen, e.g., randomly.

For the sake of readability, we define the abbreviation $\pi_w(\underline{\hat{z}}, \underline{x}) := \pi(\underline{\hat{z}}, \underline{x}, \mathbf{C}_w)$. An illustration of the projection is shown in Fig. 4, where the measurement $\underline{\hat{z}}$ (blue cross) is projected onto the line segment. Using the projection function we can rewrite (9) to

$$\underline{\tilde{z}}_{x,s'} = \pi_w(\underline{\hat{z}}, \underline{x}) , \qquad (11)$$

and in consequence simplify likelihood (8)

$$p(\underline{\hat{z}}|\underline{x}) \stackrel{(8)}{=} \mathcal{N}(\underline{\hat{z}} - \underline{\tilde{z}}_{x,s'}, \mathbf{C}_w)$$

$$\stackrel{(11)}{=} \mathcal{N}(\underline{\hat{z}} - \pi_w(\underline{\hat{z}}, x), \mathbf{C}_w) .$$
(12)

We will denote this model as GAM- π .

The motivation for projections is that there exist analytic expressions for many shapes and, thus, iterating through all sources can be avoided, as illustrated in Fig. 5. An estimator that uses the GAM- π (12) will minimize the component-wise difference between each measurement $\hat{\underline{z}}$ and its projection

 $\pi_w(\underline{\hat{z}}, \underline{x})$, while taking into account the specific characteristics of the noise.

D. Relation to Distances

Minimizing the component-wise difference as in the GAM- π is closely related to minimizing the distance. Thus, the expression $\underline{\hat{z}} - \pi_w(\underline{\hat{z}}, \underline{x})$ is often approximated by a function g that fulfills

$$g(\underline{x}, \underline{\hat{z}}) = 0 \Leftrightarrow \underline{\hat{z}} \in \tilde{\mathcal{Z}}_x$$

In literature, g is sometimes referred to as *implicit shape* function, as it implicitly relates points in space $\hat{\underline{z}}$ to the shape. By means of the shape function, the multivariate likelihood (12) is approximated by the univariate

$$p(\underline{\hat{z}}|\underline{x}) = \mathcal{N}(g(\underline{x}, \underline{\hat{z}}), \operatorname{Var}\{g(\underline{x}_*, \underline{\hat{z}})\}) , \qquad (13)$$

which we will denote as GAM-g. This approximation, however, also affects the covariance term by replacing it by a variance term that unfortunately depends on the true, unknown state parameters \underline{x}_* . Hence, its calculation requires the approximation of the true parameters and is discussed in Sec. V.

In a GAM-g, the function g implicitly establishes the greedy one-to-one association between \hat{z} and a point on the shape. This can be easily verified when assuming g to be the Mahalanobis distance according to

$$g(\underline{x}, \underline{\hat{z}}) = \sqrt{\left[\underline{\hat{z}} - \pi_w(\underline{\hat{z}}, \underline{x})\right]^{\mathrm{T}} \mathbf{C}_w^{-1} \left[\underline{\hat{z}} - \pi_w(\underline{\hat{z}}, \underline{x})\right]} .$$
(14)

In this case, the greedily associated source for $\underline{\hat{z}}$ is given by the Mahalanobis projection $\pi_w(\underline{\hat{z}}, \underline{x})$. For the important special case of isotropic noise, minimizing (14) is equivalent to minimizing the Euclidean distance

$$g(\underline{x}, \underline{\hat{z}}) = \|\underline{\hat{z}} - \pi(\underline{\hat{z}}, \underline{x}, \mathbf{I})\|$$

from the points to their Euclidean projection $\pi(\hat{z}, \underline{x}, \mathbf{I})$.

Besides these common shape functions, related works propose to use other expressions g with the advantage of simplified equations. For example [4] estimate a star-convex object by minimizing the squared radial distance and [10] (among others) minimize an algebraic expression for fitting circles.

IV. BIAS IN SHAPE ESTIMATION

The GAM- π , as well as the GAM-g establish a greedy association of measurements \hat{z} to the shape in order to recover the true source. This association, however, turns out to be a rough approximation in the case of noise and introduces a *bias* in the estimated parameters \underline{x}_{e} .

For an illustrative example see Fig. 6. According to a uniform distribution p(s) = U(0, 1), sources were drawn along a line segment of length 1 in 2D and then additively disturbed according to (1) by isotropic Gaussian noise with covariance $\mathbf{C}_w = \sigma_w^2 \cdot \mathbf{I}$. Then, we found the most likely source for each \hat{z} by applying (10) and calculated the histogram of these recovered sources. It can be seen that for low noise in Fig. 6a, the approximation is close to correct, while in the case of increasing noise, too much probability mass is assigned to the edges. In the following, we take a closer look at the effects of this miss-assignment on the GAM- π , and the GAM-g. Based on this investigation, we then propose a scheme to alleviate the bias introduced by the misalignment.



Figure 6: Approximation error of recovering the source distribution by the greedy association. The uniform distribution (black) along a line segment cannot be recovered (red) from measurements in the presence of noise.



Figure 7: Origin of the bias in shape estimation.

A. Bias of GAM-g

Let us first explore the bias when using a GAM-g. The estimation algorithm will find an estimate \underline{x}_e that maximizes the likelihood (13). Coincidentally, this estimate \underline{x}_e will minimize $g(\underline{x}_e, \underline{\hat{z}})$ and thus, will try to fulfill

$$\mathbf{E}\{g(\underline{x}_{\mathbf{e}},\underline{\hat{z}})\} \stackrel{!}{=} 0 \ . \tag{15}$$

However, this estimate \underline{x}_e can easily be shown to be biased in the presence of noise. The idea is to show that $E\{g(\underline{x}_*, \underline{\hat{z}})\}$ does not necessarily equal zero for the true parameters \underline{x}_* . For simplicity, let the shape function be given by (14), and $C_w = I$, such that g coincides with the Euclidean distance $\|\cdot\|$. Then,

$$E\{g(\underline{x}_*, \underline{\hat{z}})\} = E\{g(\underline{x}_*, \underline{\tilde{z}} + \underline{w})\}$$

$$\stackrel{(14)}{=} E\{\|(\underline{\tilde{z}} + \underline{w}) - \pi_w(\underline{\tilde{z}} + \underline{w}, \underline{x}_*)\|\}$$

holds for a measurement $\underline{\hat{z}}$ and its generating source $\underline{\tilde{z}}$. Now, let us conduct the following thought experiment. Fig. 7(a,d,g) show three shape boundaries. The true shape in each figure is marked in black, a source $\underline{\tilde{z}}$ is drawn as a red dot, and the isotropic uncertainty around it is schematically indicated by the filled circle. Probability mass for expected measurements $\underline{\hat{z}} = \underline{\tilde{z}} + \underline{w}$ from this source is schematically colored in blue (left of the boundary) and gray (right of the boundary). For illustration,

let blue denote the regions with negative signed distance to the shape, and gray the regions with positive signed distance, respectively. The estimated shape that fulfills (15) is depicted as dashed, orange line. In Fig. 7(c,f,i), this can be thought of finding the shape that perfectly balances measurements on the left and the right side.

Line Shape: For the line shape, the estimate coincides with the true shape. This is due to the fact that the true ratio between positive and negative distances is perfectly balanced (see Fig. 7b), such that

$$\mathbf{E}\{g(\underline{x}_*, \underline{\hat{z}})\} = 0 = \mathbf{E}\{g(\underline{x}_e, \underline{\hat{z}})\}$$

holds. In consequence, the algorithm will find the true parameters as they maximize the likelihood in (13).

Curve and Corner Shape: For the curve and corner shape, the estimate does not coincide with the true shape. This is due to the fact that the true ratio between positive and negative distances is unbalanced, such that

$$\mathbf{E}\{g(\underline{x}_*, \underline{\hat{z}})\} \neq 0 = \mathbf{E}\{g(\underline{x}_e, \underline{\hat{z}})\}$$

holds. In other words, for the true shape, there is more probability mass for points on the right side of the boundary than for points on the left side. In consequence, the algorithm will find biased parameters, as it assumes a balanced ratio.

B. Bias of GAM- π

Looking at the GAM- π will provide further insights about the bias. Similar to (15), the estimator based on (12) minimizes the differences between $\underline{\hat{z}}$ and $\pi_w(\underline{\hat{z}}, \underline{x})$. In consequence, for the true parameters, their expected values should be equal. According to (1), the first term is immediately given by

$$\mathbf{E}\{\hat{\underline{z}}\} \stackrel{(1)}{=} \underbrace{\mathbf{E}\{\underline{\tilde{z}}\}}_{=\underline{\tilde{z}}} + \underbrace{\mathbf{E}\{\underline{w}\}}_{=0} , \qquad (16)$$

and the second term by

$$\mathrm{E}\{\pi_w(\underline{\hat{z}},\underline{x}_*)\} \stackrel{(1)}{=} \mathrm{E}\{\pi_w((\underline{\tilde{z}}+\underline{w}),\underline{x}_*)\} .$$

In Fig. 8, both terms are evaluated for entire shapes, and drawn against each other. The true sources $\underline{\tilde{z}}$ are drawn in black, and the expected $\mathbb{E}\{\pi_w((\underline{\tilde{z}} + \underline{w}), \underline{x}_*)\}$ in red, respectively. This second term was calculated using random sampling of the measurement noise. An important observation is that the true sources do only coincide with the expected ones for straight parts (Fig. 8e) along the shape or for vanishingly low noise (Fig. 8a). The estimator, however, will try to find a biased estimate \underline{x}_e where $\mathbb{E}\{\pi_w((\underline{\tilde{z}} + \underline{w}), \underline{x}_e)\}$ is closer to $\underline{\tilde{z}}$. This bias will depend on the level of noise and the true usually unknown distribution of sources.

C. Bias Alleviation

In this section, a general scheme for reducing the bias in GAMs is proposed, inspired by ideas from perturbation theory [5], [11]. The idea is to slightly adjust $p(\hat{z}|\underline{x})$ in order to ensure that it takes its maximum for the true parameters. A reasonable choice for a GAM- π is to replace (12) by

$$p(\underline{\hat{z}}|\underline{x}) = \mathcal{N}(\underline{\hat{z}} - \pi_w(\underline{\hat{z}}, \underline{x}) - \mathrm{E}\{\underline{\hat{z}} - \pi_w(\underline{\hat{z}}, \underline{x}_*)\}, \mathbf{C}_w) , (17)$$

where $E\{\hat{\underline{z}} - \pi_w(\hat{\underline{z}}, \underline{x}_*)\}$ is a *correction term* that ensures that the likelihood takes its maximum at $\underline{x}_e = \underline{x}_*$. To verify this,



Figure 8: Illustration of the bias. Groundtruth sources $\underline{\tilde{z}}$ are marked in black, $E\{\pi_w((\underline{\tilde{z}} + \underline{w}), \underline{x}_*)\}$ in red.

we can use the same technique as in Sec. IV-B. The estimator based on (17) will try to minimize the differences between $\hat{\underline{z}}$ and $\pi_w(\hat{\underline{z}}, \underline{x}_e) + E\{\hat{\underline{z}} - \pi_w(\hat{\underline{z}}, \underline{x}_*)\}$. In consequence, for the true parameters, their expected values should be equal. For the first term, $E\{\hat{\underline{z}}\} = \tilde{\underline{z}}$ is already given by (16). Then plugging $\underline{x}_e = \underline{x}_*$ into the second term yields

$$\mathbf{E}\{\pi_w(\underline{\hat{z}},\underline{x}_*) + \mathbf{E}\{\underline{\hat{z}} - \pi_w(\underline{\hat{z}},\underline{x}_*)\}\} = \underbrace{\mathbf{E}\{\underline{\hat{z}}\}}_{=\underline{\tilde{z}}} \cdot \mathbf{E}\{\underline{\hat{z}}\}$$

From this, we can conclude that (17) will be ideally unbiased, as the differences will approach zero for the true parameters. However, as the correction term $E\{\hat{z} - \pi_w(\hat{z}, \underline{x}_*)\}$ depends on the true parameters \underline{x}_* , deriving it requires an appropriate approximation. We propose to use a recursive Bayesian estimator, to allow for approximating the true parameters \underline{x}_* by the mean $\underline{\mu}_x$ from the prior distribution $p(\underline{x})$, i.e., $\underline{x}_* \approx \underline{\mu}_x$. As the uncertainty of $p(\underline{x})$ usually decreases during the estimation process, the quality of the approximation increases. For this reason, the influence of $p(\underline{x})$ on the bias vanishes and, thus, it is excluded from our considerations. The correction term can be calculated according to

$$E\{\hat{\underline{z}} - \pi_w(\hat{\underline{z}}, \underline{x}_*)\} \stackrel{(1)}{=} \tilde{\underline{z}} - E\{\pi_w(\tilde{\underline{z}} + \underline{w}, \underline{x}_*)\}$$
(18)
$$\stackrel{(11)}{=} \pi_w(\hat{\underline{z}}, \underline{x}_*) - E\{\pi_w(\pi_w(\hat{\underline{z}}, \underline{x}_*) + \underline{w}, \underline{x}_*)\}$$
$$\approx \pi_w(\hat{\underline{z}}, \underline{\mu}_r) - E\{\pi_w(\pi_w(\hat{\underline{z}}, \underline{\mu}_r) + \underline{w}, \underline{\mu}_r)\}$$
.

Note that the proposed alleviation scheme can be immediately applied to the GAM-g, and turns (13) into

$$p(\underline{\hat{z}}|\underline{x}) = \mathcal{N}(g(\underline{x}, \underline{\hat{z}}) - \mathrm{E}\{g(\underline{x}_*, \underline{\hat{z}})\}, \mathrm{Var}\{g(\underline{x}_*, \underline{\hat{z}})\}) . (19)$$

The calculation of the correction term $E\{g(\underline{x}_*, \underline{\hat{z}})\}$ is analog to (18).

Example: Circle

We found that a particular estimator for circular shapes [7] is a special case of the proposed GAM-g (19) estimator with bias alleviation.

Let us assume that the circle is centered on the origin, so that the task is to estimate the radius $\underline{x} = r$. Furthermore, let the noise covariance C_w be isotropic. In [7], they implicitly use an Euclidean projection of measurements $\underline{\hat{z}}$ onto the circle according to

$$\pi_w(\underline{\hat{z}},\underline{x}) = \frac{\underline{\hat{z}}}{\|\underline{\hat{z}}\|}r$$

As shape function g, they use the squared, signed radial distance

$$g(\underline{x}, \underline{\hat{z}}) = \|\underline{\hat{z}}\|^2 - \|\pi_w(\underline{\hat{z}}, \underline{x})\|^2 = \underline{\hat{z}}^2 - r^2 .$$

Evaluating the likelihood (19) further requires the correction term $E\{g(\underline{x}_*, \underline{\hat{z}})\}$ and the variance $Var\{g(\underline{x}_*, \underline{\hat{z}})\}$. The derivation for both is given in the appendix. Finally, we arrive at the GAM-g

$$p(\underline{\hat{z}}|\underline{x}) \stackrel{(19)}{=} \mathcal{N}\left(\underline{\hat{z}}^2 - r^2 - \operatorname{Tr}(\mathbf{C}_w), \operatorname{Tr}(\mathbf{C}_w)^2 + 2\mu_r^2 \operatorname{Tr}(\mathbf{C}_w)\right) ,$$

which exactly coincides with the result from [7].

V. IMPLEMENTATION

In this section, we give some general comments and explicit instructions for implementing a Bayes update based on a GAM. In doing so, we focus on the GAM- π . However, all derivations apply to the GAM-g in a straightforward way. In the following, we refer to a single measurement $\hat{\underline{z}}$ with known noise covariance C_w . More measurements can be processed sequentially according to (2). Prior knowledge on the parameters is given by $p(\underline{x})$ with $\underline{\mu}_x$ being selected as the mean or representative mode of $p(\underline{x})$.

Implementing a GAM- π requires 1) defining a projection function $\pi_w(\underline{\hat{z}}, \underline{x})$ according to (10) for the specific shape and 2) calculating the correction term $E\{\underline{\hat{z}} - \pi_w(\underline{\hat{z}}, \underline{x}_*)\}$ from (18). Besides for simple shapes, $E\{\underline{\hat{z}} - \pi_w(\underline{\hat{z}}, \underline{x}_*)\}$ cannot be derived analytically and thus, requires an approximation by, e.g., sampling techniques. Let $\{\underline{w}_l\}_{l=1}^L$ be a representative set of samples from the noise distribution $\mathcal{N}(\underline{0}, \mathbf{C}_w)$, drawn randomly or deterministically according to [19], [20]. Then, from these samples, the correction term (18) can be calculated as the *sample mean*

$$\mathbf{E}\{\underline{\hat{z}} - \pi_w(\underline{\hat{z}}, \underline{x}_*)\} \approx \frac{1}{L} \sum_{l=1}^L \pi_w(\underline{\hat{z}}, \underline{\mu}_x) - \mathbf{E}\left\{\pi_w(\pi_w(\underline{\hat{z}}, \underline{\mu}_x) + \underline{w}_l, \underline{\mu}_x)\right\}$$

As a short remark, for the GAM-g, the mean $E\{g(\underline{x}_*, \underline{\hat{z}})\}$ and variance $Var\{g(\underline{x}_*, \underline{\hat{z}})\}$ can analogously be calculated using sampling techniques.

Linear Regression Kalman Filter: For estimation using a Linear Regression Kalman Filter (LRKF) [19], [20], a set of samples $\{\underline{x}_j\}_{j=1}^J$ from the state distribution p(x) is given. The GAM- π measurement function then is

$$h(\underline{x}_j, w, \underline{\hat{z}}) := \underline{\hat{z}} - \pi_w(\underline{\hat{z}}, \underline{x}_j) - \mathrm{E}\{\underline{\hat{z}} - \pi_w(\underline{\hat{z}}, \underline{x}_*)\} + \underline{w}$$

= $\underline{0}$, (20)

where \underline{w} is an additive noise term with $\mathcal{N}(\underline{0}, \mathbf{C}_w)$. To the estimator, zero acts as a constant *pseudo-measurement*, while



Figure 9: Simulation Setup.

 $\hat{\underline{z}}$ is modeled as a parameter. It is notable that the common biased measurement function would simply be (20) without the correction term.

VI. EVALUATION

In this section, we evaluate the proposed bias reduction technique. More specifically, we will answer three evaluation questions:

- Question 1: How does the proposed GAM- π with bias reduction compare to the well-known SDM?
- Question 2: How does the proposed GAM-π with bias reduction compare to common distance-minimizing approaches, using a GAM-g?
- Question 3: How high is the improvement of the bias reduction compared to the common GAM- π ?

For this purpose, we defined the following estimation task.

A. Estimation Task

As evaluation object, we chose the example shape in Fig. 9, known from the previous sections. From this shape, 500 measurements were generated according to the following rule: Sources were uniformly drawn along $\frac{3}{4}$ of the boundary (red) and then disturbed by additive non-isotropic noise with covariance matrix $\mathbf{C}_w = [1.5 \cdot 10^{-2}, 5 \cdot 10^{-3}; 5 \cdot 10^{-3}, 1.5 \cdot 10^{-2}]$ (as indicated in blue). Parameters to be estimated consist of the two dimensional position, the orientation angle, and the scale. Technically, the shape is modeled by a polygon with 75 vertices.

GAM- π : We implemented an LRKF based on the GAM- π measurement function (20) using 400 state samples. Note that the Mahalanobis projection takes into account the specific characteristics of the non-isotropic noise. The correction term was calculated based on 40 samples. All samples were deterministically drawn according to [19]. Further, we distinguish between two instances of this estimator: one *common* biased version and another with the *proposed* alleviation.

SDM: For the SDM, we implemented a particle filter [21] based on the likelihood (6) that uses 500 particles and performs numerical integration over s with 100 samples. Further, we distinguish between two instances of this estimator. The first "SDM correct" approach uses the *correct* distribution p(s), i.e., it assumes that sources originate uniformly from $\frac{3}{4}$ of the boundary. The second "SDM wrong" approach uses a *wrong* distribution p(s) that assumes that sources originate uniformly from the entire boundary. Note that this second SDM is

commonly used, when no further knowledge is available about the true distribution.

GAM-g: As a representative curve fitting approach, we implemented an LRKF based on a GAM-g that uses 400 state samples. This estimator was designed in the common fashion and minimizes the Euclidean distance between measurements and shape boundary. Furthermore, it does not take advantage of the bias alleviation. The sample variance was calculated based on 40 samples. All samples were deterministically drawn according to [19].

B. Implementation Details

All estimators were initialized with random values drawn from $\mathcal{N}(\underline{x}_*, 10^{-1} \cdot \mathbf{I})$. The state distribution $p(\underline{x})$ then was sequentially updated with each of the 500 measurements, in alternation with a *prediction step* using a *random walk model* in the magnitude of $\mathcal{N}(\underline{0}, 10^{-5} \cdot \mathbf{I})$, to avoid getting stuck into a local minimum.

C. Results

For the evaluation, we performed 100 runs for all estimators in the estimation tasks. In Fig. 10, the average shape as estimated by the proposed GAM- π approach is drawn against the SDM approach(Fig. 10a), and the common GAM approaches (Fig. 10b). The *root mean squared error* (RMSE) for position, orientation angle, and scale over all runs is summarized in Fig. 11 for all approaches. From this, we can answer the three evaluation questions.

Question 1: From Fig. 11, it is easy to verify that the SDM approach with correctly modeled distribution shows the best performance in all parameters, in terms of the RMSE. The proposed GAM- π approach with bias alleviation shows a comparable estimation quality as the "SDM correct". This is remarkable, as the GAM- π approach is not given any information about the source distribution. In contrast, the "SDM wrong" using the uniform distribution along the entire boundary performs significantly worse than the GAM- π approach. In sum, the GAM- π approach is the better choice, when the source distribution is not available.

Question 2: The distance-minimizing GAM-g approach is heavily biased in all parameters. This is due to the nonisotropic noise characteristics, as well as the complicated shape. Thus, compared to standard curve fitting approaches, for this scenario, the "GAM- π proposed" can reduce the RMSE about 80% in terms of scaling and about 70% in terms of orientation angle.

Question 3: The estimation performance of the "GAM- π common" is quite similar to the bias alleviated "GAM- π proposed". Especially position (Fig. 11a) and orientation (Fig. 11b) converge to the same values as the unbiased SDM. However, the common approach estimates a biased scale (Fig. 11c). In sum, for this scenario, applying the bias alleviation in the GAM- π approach can reduce the scaling bias about 50%.

VII. CONCLUSION

In this work, we presented an intuitive derivation of two commonly used approaches for Bayesian shape estimation. The first approach uses the well-known *Spatial Distribution Model* (SDM) and the second uses an implicit model known from curve fitting algorithms that minimize a distance-related expression. We found that these implicit models (GAM-g) can be derived as an SDM with special (greedy) assumptions. Based on this concept, we derived an estimator based on a *Greedy* Association Model (GAM- π) that minimizes the differences between measurements and their most likely sources on the shape and can naturally deal with non-isotropic noise. For this estimator (GAM- π), as well as for those using a GAM-g we explored the well-known parameter bias in the presence of noise. Inspired by ideas from related approaches, we then proposed a general alleviation scheme for this bias that works with non-isotropic noise, as well as with arbitrary shapes. In addition, we showed that this scheme only requires the subtraction of a correction term in a nonlinear measurement function with additive noise.

From the evaluation, it can be concluded that in the presence of non-isotropic noise, and if the true spatial distribution is unknown, the proposed estimator outperforms all standard approaches, even the SDM. In numerical terms, by using the GAM- π , we were able to reduce the orientation RMSE about 70% and the scaling RMSE about 80% compared to common distance-minimizing approaches in the considered scenario.

APPENDIX

This section includes the proof that the proposed correction term is equivalent to the one from [7] when considering a circular shape. The expectation value evaluates to

$$\begin{split} \mathbf{E}\{g(\underline{x}, \underline{\hat{z}})\} &\approx \mathbf{E}\left\{g\left(\underline{\mu}_{x}, \pi_{w}(\underline{\hat{z}}, \underline{\mu}_{x}) + \underline{w}\right)\right\} \\ &= \mathbf{E}\left\{\left(\frac{\mu_{r}}{\|\underline{\hat{z}}\|}\underline{\hat{z}} + \underline{w}\right)^{2} - \mu_{r}^{2}\right\} \\ &= \mathbf{E}\left\{\underbrace{\frac{\mu_{r}}{\|\underline{\hat{z}}\|^{2}}\underline{\hat{z}}^{2}}_{\|\underline{\hat{z}}\|^{2}} + \underbrace{\frac{2\mu_{r}}{\|\underline{\hat{z}}\|}\underline{\hat{z}}^{\mathrm{T}}\underline{w}}_{=0} + \underbrace{\underline{w}^{2}}_{=\mathrm{Tr}(\mathbf{C}_{w})} - \mu_{r}^{2}\right\} \\ &= \mathrm{Tr}(\mathbf{C}_{w}) \;, \end{split}$$

where $\underline{\mu}_x = \mu_r$ is the mean of the prior distribution $p(\underline{x})$. Analogously, the variance evaluates to

$$\begin{aligned} \operatorname{Var}\left\{g(\underline{x}_{*},\underline{\hat{z}})\right\} &\approx \operatorname{E}\left\{\left(g(\underline{\mu}_{x},\pi_{w}(\underline{\hat{z}},\underline{\mu}_{x})+\underline{w})-\operatorname{Tr}(\mathbf{C}_{w})\right)^{2}\right\} \\ &= \operatorname{E}\left\{\left(\left(\frac{\mu_{r}}{\|\underline{\hat{z}}\|}\underline{\hat{z}}+\underline{w}\right)^{2}-\mu_{r}^{2}-\operatorname{Tr}(\mathbf{C}_{w})\right)^{2}\right\} \\ &= \operatorname{E}\left\{\left(\frac{2\mu_{r}}{\|\underline{\hat{z}}\|}\underline{\hat{z}}^{\mathrm{T}}\underline{w}+\underline{w}^{2}-\operatorname{Tr}(\mathbf{C}_{w})\right)^{2}\right\} \\ &= \operatorname{E}\left\{4\mu_{r}^{2}\underline{w}^{2}+\underline{w}^{4}+\operatorname{Tr}(\mathbf{C}_{w})^{2}-2\underline{w}^{2}\operatorname{Tr}(\mathbf{C}_{w})\right\} \\ &= 2r^{2}\operatorname{Tr}(\mathbf{C}_{w})+3\operatorname{Tr}(\mathbf{C}_{w})^{2}+\operatorname{Tr}(\mathbf{C}_{w})^{2}-2\operatorname{Tr}(\mathbf{C}_{w})^{2} \\ &= \operatorname{Tr}(\mathbf{C}_{w})^{2}+2\mu_{r}^{2}\operatorname{Tr}(\mathbf{C}_{w}) \;.\end{aligned}$$

References

- K. Gilholm and D. Salmond, "Spatial Distribution Model for Tracking Extended Objects," *IEE Proceedings on Radar, Sonar and Navigation*, vol. 152, no. 5, pp. 364–371, 2005.
- [2] M. Werman and D. Keren, "A Bayesian method for fitting parametric and nonparametric models to noisy data," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 5, pp. 528–534, 2001.
- [3] M. Feldmann, D. Franken, and W. Koch, "Tracking of Extended Objects and Group Targets Using Random Matrices," *IEEE Transactions on Signal Processing*, vol. 59, no. 4, pp. 1409–1420, Apr. 2011.



Figure 10: Average shape estimate over 100 runs. The true shape is drawn in black.



Figure 11: RMSE of estimated parameters over 100 runs.

- [4] M. Baum and U. D. Hanebeck, "Shape Tracking of Extended Objects and Group Targets with Star-Convex RHMs," in *Proceedings of the* 14th International Conference on Information Fusion (Fusion 2011), Chicago, Illinois, USA, 2011.
- [5] Z. Zhang, "Parameter estimation techniques: A tutorial with application to conic fitting," *Image and vision Computing*, vol. 15, no. 1997, pp. 59–76, 1997.
- [6] M. Baum and U. D. Hanebeck, "Fitting Conics to Noisy Data Using Stochastic Linearization," in *Proceedings of the 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2011)*, San Francisco, California, USA, Sep. 2011.
- [7] M. Baum, "A novel Bayesian method for fitting a circle to noisy points," in 13th International Conference on Information Fusion (2010), 2010.
- [8] A. Fitzgibbon, M. Pilu, and R. Fisher, "Direct least square fitting of ellipses," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 5, pp. 476–480, May 1999.
- [9] D. Umbach and K. Jones, "A few methods for fitting circles to data," *IEEE Transactions on Instrumentation and Measurement*, vol. 52, no. 6, pp. 1881–1885, Dec. 2003.
- [10] A. Al-Sharadqah and N. Chernov, "Error analysis for circle fitting algorithms," *Electronic Journal of Statistics*, vol. 3, pp. 886–911, 2009.
- [11] K. Kanatani, "Statistical bias of conic fitting and renormalization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, no. 3, pp. 320–326, Mar. 1994.
- [12] D. Miller, "Reducing transformation bias in curve fitting," *The American Statistician*, vol. 38, no. 2, pp. 124–126, 1984.

- [13] B. Friedland, "Treatment of bias in recursive filtering," *IEEE Transactions on Automatic Control*, vol. 14, no. 4, pp. 359–367, Aug. 1969.
- [14] J.-P. Drécourt, H. Madsen, and D. Rosbjerg, "Bias aware Kalman filters: Comparison and improvements," *Advances in Water Resources*, vol. 29, no. 5, pp. 707–718, May 2006.
- [15] X. Li and V. Jilkov, "A survey of maneuvering target tracking: Dynamic models," in *Proc. 2000 SPIE Conf. on Signal and Data Processing*, no. April, Orlando, Florida, USA, 2000.
- [16] J. Vermaak, N. Ikoma, and S. Godsill, "Sequential Monte Carlo framework for extended object tracking," *IEE Proceedings on Radar, Sonar* and Navigation, pp. 353–363, 2005.
- [17] A. Stein and M. Werman, "Robust Statistics for Shape Fitting and Pattern Matching," in *Computer Vision and Pattern Recognition (CVPR)*, Champaign, IL, USA, 1992, pp. 540–546.
- [18] M. Baum, F. Faion, and U. D. Hanebeck, "Modeling the Target Extent with Multiplicative Noise," in *Proceedings of the 15th International Conference on Information Fusion (Fusion 2012)*, Singapore, 2012.
- [19] J. Steinbring and U. D. Hanebeck, "S2KF: The Smart Sampling Kalman Filter," in *Proceedings of the 16th International Conference* on Information Fusion (Fusion 2013), Istanbul, Turkey, 2013.
- [20] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proceedings of the IEEE*, vol. 92, no. 3, pp. 401–422, Mar. 2004.
- [21] M. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "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.