Partial Likelihood for Unbiased Extended Object Tracking

Florian Faion, **Antonio Zea**, **Marcus Baum**, 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, marcus.baum@kit.edu, uwe.hanebeck@ieee.org

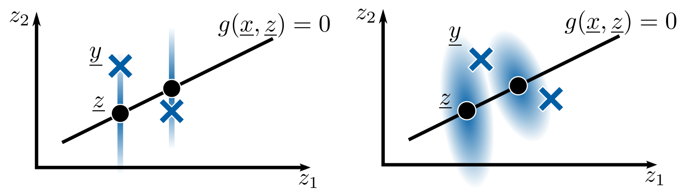
***Abstract*—An extended object gives rise to several measurements that originate from unknown measurement sources on the object. In this paper, we consider the tracking and parameter estimation of extended objects that are modeled as a curve in 2D such as a circle or an ellipse. A standard model for such extended objects is to assume that the unknown measurement sources are uniformly distributed on the curve. We argue that the uniform distribution may not be the best choice in scenarios where the true distribution of the measurements significantly differs from a uniform distribution. Based on results from curve fitting and errors-in-variables models, we develop a partial likelihood that ignores the distribution of measurement sources and can be shown to outperform the likelihood for a uniform distribution in these scenarios. If the true measurement sources are in fact uniformly distributed, our new likelihood results in a slightly slower convergence but has the same asymptotic behavior.**

I. Introduction

Tracking an extended object based on noisy measurements of its boundary is an instance of errors-in-variables (EIV) problems, as we do not know their originating measurement sources on the boundary. Besides, EIV occur in many other tasks, such as estimating the parameters of a geometric curve [1], the fundamental matrix [2], or even the optical flow [3].

To explain the basic problem structure, we consider the familiar example of fitting a line to noisy data, where we want to estimate slope and intercept parameters, encoded in the vector $\underline{x}$, of a linear constraint from noisy points $\underline{y}$. In classical regression, as illustrated in Fig. 1a, the abscissa can be measured exactly, while the ordinate is subject to noise. The estimator would adjust the parameters, so that the error between line and measurements along the ordinate is minimized. In doing so, the measurements are correctly associated to their generating sources $\underline{z}$ on the line.

In contrast, in EIV problems, all dimensions of a measurement $\underline{y}$ are subject to noise, as illustrated in Fig. 1b. This makes the association of a measurement to its source on the line ambiguous. In consequence, we have to make assumptions on the measurement source $\underline{z}$ for each $\underline{y}$ to design an estimator. To make things worse, each additional measurement introduces another unknown measurement source that, in turn, requires an additional assumption. In statistics literature, this association problem is sometimes referred to as the *Neyman-Scott problem* [4] and $\underline{z}$ is called a *nuisance parameter* [5]. This work is about dealing with the association problem in the context of extended object tracking.



(a) Classical regression. (b) EIV regression.

Figure 1: In classical regression (a), each measurement $\underline{y}$ can be exactly associated to its originating source $\underline{z}$, where in EIV regression (b), this is not possible.

A. Problem Statement

We consider objects, whose boundary can be modeled as a geometric constraint such as an ellipse in 2D. Then, given a set of noisy *measurements* $\underline{y}_1, \ldots, \underline{y}_n$ of the object boundary, the task is to find the *state* parameter vector $\underline{x}$ that best fits the constraint to the noisy data according to the following implicit model. Each measurement $\underline{y}_i$ is a noisy observation of a true value $\underline{z}_i$, denoted as the *measurement source*, which is distorted by additive noise $\underline{v}_i \sim p(\underline{v}_i)$ according to

$$\underline{y}_i = \underline{z}_i + \underline{v}_i \ . \tag{1}$$

The measurement source in turn lies on the object boundary and fulfills the implicit and typically nonlinear relationship

$$g(\underline{x}, \underline{z}_i) = 0 \ . \tag{2}$$

For convenience, we will write $g_x(\underline{z}_i) := g(\underline{x}, \underline{z}_i)$. Note that due to the noise, the measurements $\underline{y}_1, \ldots, \underline{y}_n$ themselves generally do not fulfill the constraint and thus, measurements and state are only related via the unknown measurement sources. The challenging aspect is that in order to infer the state parameters $\underline{x}$, we need to additionally make assumptions about the measurement sources. In addition, as we consider extended objects, we also have a dynamic component, which means that i) packages of measurements sequentially arrive over time and ii) state parameters may change between the measurements. However, for our derivations, it is sufficient to consider a single time instance of this process, such that we can omit all potential time indices.

B. Related Work

In the context of extended object tracking, a straightforward way to deal with the association problem is the Spatial Distribution Model (SDM), where each point on the object boundary is assigned a probability to be measured [6], [7].

Usually, the distribution is assumed to be either uniform or is extracted from the sensor to object geometry [8], [9]. However, due to occlusions and unpredictable factors such as sensor artifacts, assuming a distribution of measurement sources is not always reasonable. Furthermore, SDMs involve computationally demanding integrals. An alternative approach to address the association problem is greedily associating a measurement to a specific source on the object. In doing so, some sort of distance-related expression is minimized by the estimator as shown in [10]. These Greedy Association Models (GAMs) tend to be biased in the presence of noise. Other approaches [11], [12] consider constraints in the form of $g_x(\underline{z}_i) \in [a, b]$ which occur, e.g., when a region is to be tracked in a 2D image. Recently, there has also been great effort in studying errors-in-variables problems in a more general scope, but mostly focused on static problems [13–15].

C. Contribution

Our main contribution is a partial likelihood for extended objects, which is partial in the sense that it puts ignorance on the association problem instead of incorporating a potentially improper heuristic. Specifically, we propose a re-parametrization of the measurements that decouples their encoded information into "how well" they fit to the object boundary and "where" on the boundary they are related to. Ignoring the second type of information we develop an estimator that

1) can handle occlusions and does not rely on a probability distribution along the boundary,
2) can naturally deal with anisotropic measurement noise,
3) is unbiased by design,
4) and can be implemented using common recursive Bayesian estimation techniques such as, e.g., a nonlinear Kalman filter.

Specifically, we derive a measurement equation with additive noise and propose a sampling-based approach to evaluate the likelihood.

D. Outline

In Sec. II, we briefly discuss the considered estimation techniques and then, in Sec. III, we derive a general likelihood for extended objects. Sec. IV summarizes two traditional models to evaluate this likelihood before we introduce the new model and propose a sampling-based approach for its implementation in Sec. V. Finally, in Sec. VII we evaluate the new model in a recursive ellipse-estimation task and draw conclusions in Sec. VIII.

II. Recursive Bayesian Estimation

In extended object tracking, the relationship between object parameters $\underline{x}$ and measurements $\underline{y}_1, \ldots, \underline{y}_n$ is usually expressed in terms of a *likelihood* $p(\underline{y}_1, \ldots, \underline{y}_n|\underline{x})$. The likelihood then can be used to update a prior distribution $p(\underline{x})$ on the parameters by applying Bayes' rule

$$p(\underline{x}|\underline{y}_1, \ldots, \underline{y}_n) \propto p(\underline{y}_1, \ldots, \underline{y}_n|\underline{x}) \cdot p(\underline{x}) \ . \tag{3}$$

This allows for developing a recursive Bayesian estimator by using the posterior distribution as the prior for the next update. A recursive Bayesian estimator [16], in turn, can

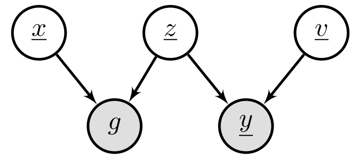


Figure 2: Probabilistic graphical model for EIV.

be employed to design a tracking algorithm by alternating between two steps. First, the *prediction step* lets the distribution $p(\underline{x})$ evolve over time according to a system model. Second, the *measurement update step* incorporates new measurements $\underline{y}_1, \ldots, \underline{y}_n$ according to (3). Next, we show how a likelihood for extended objects can be derived.

III. A Likelihood for Extended Objects

We want to derive a likelihood $p(\underline{y}_1, \ldots, \underline{y}_n|\underline{x})$ for extended objects. Specifically, given a set of measurements $\underline{y}_1, \ldots, \underline{y}_n$ obtained according to (1), where the noise terms $\underline{v}_i$ are assumed to be mutually independent and Gaussian distributed according to $\underline{v}_i \sim \mathcal{N}(\underline{0}, \mathbf{C}_{v,i})$ with known, but not necessarily identical, covariance matrices. In addition, it is known that the state parameters $\underline{x}$ and the sources $\underline{z}_i$ are related through the nonlinear constraint (2) that describes the object boundary.

The independence between the measurements lets us factorize the likelihood as

$$p(\underline{y}_1, \ldots, \underline{y}_n|\underline{x}) = \prod_{i=1}^{n} p(\underline{y}_i|\underline{x}) \ . \tag{4}$$

As a result of this factorization, we can exclusively consider the likelihood $p(\underline{y}_i|\underline{x})$ for a single measurement $\underline{y}_i$, which lets us drop the index i for readability. Nevertheless, defining the likelihood $p(\underline{y}|\underline{x})$ is not trivial, as measurement $\underline{y}$ and state $\underline{x}$ are only connected via the unknown source $\underline{z}$ through the constraint g_x.

A. Probabilistic Graphical Model

For a more intuitive treatment, we encode the dependencies between all involved variables visually using a *Probabilistic Graphical Model*, as seen in Fig. 2. In this model, $\underline{y}$, g are observable variables, while $\underline{x}$, $\underline{z}$, $\underline{v}$ are latent variables that are not directly accessible. Specifically, g represents the constant *pseudo-measurement* 0, which arises from the relationship $g_x(\underline{z}) = g := 0$ as seen in (2). From the dependency structure in Fig. 2 we obtain the joint probability distribution $p(\underline{x}, \underline{z}, \underline{v}, g, \underline{y})$, which leads to an intermediate likelihood $p(\underline{y}, g|\underline{x})$ by marginalizing out $\underline{z}, \underline{v}$ and dividing by $p(\underline{x})$, in the form of

$$p(\underline{y}, g|\underline{x}) = \int_{\mathbb{R}^d} \underbrace{\int_{\mathbb{R}^d} p(\underline{y}|\underline{z}, \underline{v}) \cdot p(\underline{v}) \ \mathrm{d}\underline{v}}_{p(\underline{y}|\underline{z})} \cdot p(g|\underline{x}, \underline{z}) \cdot p(\underline{z}) \ \mathrm{d}\underline{z} \ . \tag{5}$$

The inner integral is given by the additive noise model (1) as

$$p(\underline{y}|\underline{z}) = \mathcal{N}(\underline{y}; \underline{z}, \mathbf{C}_v) \ . \tag{6}$$

For evaluation of the remainder of (5), we have to define $p(g|\underline{x}, \underline{z})$ and $p(\underline{z})$. The first term $p(g|\underline{x}, \underline{z})$ is the probability

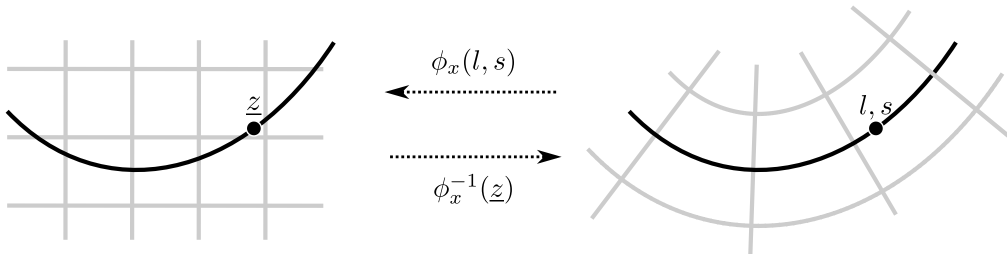


Figure 3: Sketch of the Cartesian (left) and the constraint-induced (right) parametrization. For $l = 0$, the s-component lets us iterate in the constraint (black curve).

that $\underline{x}$ and $\underline{z}$ are bound through the constraint in (2) which is 1 if $g_x(\underline{z}) = 0$ holds and 0 else. The second term $p(\underline{z})$ is the probability distribution over the domain that rates how likely a point $\underline{z}$ in space generates a measurement. Technically, to evaluate the likelihood (5), we must integrate a Gaussian, centered on the measurement, over the object boundary, where each point is weighted by its probability to be the measurement source. However, parameterizing the boundary is not straightforward in Cartesian coordinates. Thus, we will instead change $\underline{z}$ to a more convenient parameterization.

B. Constraint-induced Parametrization

Specifically, we want to represent a point $\underline{z}$ in a coordinate system in the form of $[l, \underline{s}]^{\mathrm{T}} \in L \times S \subset \mathbb{R}^d$, where the scalar $l := g_x(\underline{z})$ specifies how well $\underline{z} \in \mathbb{R}^d$ fulfills the constraint, and the vector $\underline{s}$ determines the remaining degrees of freedom. This can be visualized in Fig. 3 (right). For $l = 0$ (black line), it follows that s iterates in all points $\underline{z}$ that satisfy the constraint (2). To make the following considerations easier to follow, we focus on measurements and sources in $\mathbb{R}^2$, i.e., where both l and s are scalars. Thus, we introduce the transformation function $\phi_x : L \times S \to \mathbb{R}^2$, in the form

$$\phi_x(l, s) = \underline{z} \ , \tag{7}$$

which is bijective almost everywhere and maps parameters l, s to their corresponding Cartesian coordinates. To demonstrate this concept, let us consider a circular constraint.

Example 1 (Circle)
A circle allows for a comfortable treatment in terms of a polar representation. Mathematically, the constraint can be defined as $g_x(\underline{z}) = \|\underline{z}\| - r$. Then, the parametrization for a circle with radius $\underline{x} = r$ that is centered on the origin yields

$$\phi_x(l, s) = (r + l) \cdot \begin{bmatrix} \cos(s) \\ \sin(s) \end{bmatrix} \ , \tag{8}$$

where l is the signed distance in $L = (-r, \infty)$ and s is an angle in $S = [0, 2\pi]$. It can be seen that for $l = 0$, the parameter s lets us iterate through all points $\phi_x(0, s)$ on the circle. □

C. Changing Variables

In order to change the variable $\underline{z}$ in the integral (5) from Cartesian coordinates to the constraint-induced parametrization $[l, s]^{\mathrm{T}}$, we can apply the following lemma.

Lemma 1 (Change of Variables)
Let $p(\underline{a})$ and $p(\underline{b})$ be probability distributions where $\underline{a}$, $\underline{b}$ describe equal events in a different parametrization and A, B, represent identical regions in their respective parametrization. Then, it holds that

$$\int_A p(\underline{a}) \, \mathrm{d}\underline{a} = \int_B p(\underline{b}) \, \mathrm{d}\underline{b} \ . \tag{9}$$

□

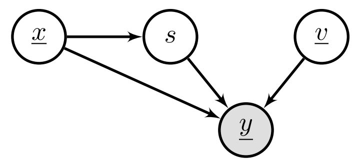


Figure 4: Probabilistic Graphical Model for the SDM.

With (9), we can rewrite the integral in (5) as

$$p(\underline{y}|\underline{x}) = \int_S \underbrace{\mathcal{N}(\underline{y}; \phi_x(0, s), \mathbf{C}_v)}_{=p(\underline{y}|\phi_x(0,s))} \cdot p(s|\underline{x}) \, \mathrm{d}s \ , \tag{10}$$

which is the likelihood of the Spatial Distribution Model [6], and well-known in extended object tracking. Its probabilistic graphical model is shown in Fig. 4. At this point, we are left with the task of specifying $p(s|\underline{x})$, which is an instance of the association problem where s acts as the *nuisance parameter*.

IV. Two Traditional Models

In the following, we discuss two approaches to address the association problem for extended objects.

A. Spatial Distribution Model (SDM)

The Spatial Distribution Model [6], [7] is a widely used model in extended object tracking and requires that $p(s|\underline{x})$ is known in the likelihood (10). Technically, this distribution specifies how likely it is that a point $\phi_x(0, s)$ in the constraint is measured, and is generally modeled as a uniform distribution. Recent approaches [8], [9] also incorporate heuristics, e.g., when the sensor observes one side of the object, they model $p(s|\underline{x})$ as a uniform distribution over the visible part only. When incorporating the correct distribution, it was shown in [6] that the SDM yields an unbiased estimator. However, wrong assumptions about $p(s|\underline{x})$ generally cause biased estimates [10].

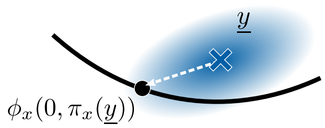


Figure 5: Finding the most likely source on the constraint for a given measurement.

B. Greedy Association Model (GAM)

Another widely-used model that we will denote as the Greedy Association Model, does not require any knowledge about $p(s|\underline{x})$. Instead, it uses the measurement $\underline{y}$ to derive a greedy estimate of its source and, in doing so, the estimator will minimize some sort of distance. For a measurement $\underline{y}$, a maximum likelihood estimate for s can be calculated by

$$\pi_x(\underline{y}) := \arg\max_{s\in\mathcal{S}} \mathcal{N}(\underline{y}; \phi_x(0, s), \mathbf{C}_v) , \tag{11}$$

where $\pi_x(\underline{y})$ refers to the most likely source $\phi_x(0, \pi_x(\underline{y}))$. See Fig. 5 for an illustration. Based on (11), we can approximate the spatial distribution by the Dirac δ-distribution

$$p(s|\underline{x}, \underline{y}) = \delta(s - \pi_x(\underline{y})) , \tag{12}$$

which now additionally depends on the measurement $\underline{y}$. Then, by plugging (12) into (10) for $p(s|\underline{x})$ we obtain

$$\begin{aligned} p(\underline{y}|\underline{x}, \underline{y}) &= \int_{\mathcal{S}} \mathcal{N}(\underline{y} - \phi_x(0, s); \underline{0}, \mathbf{C}_v) \cdot \delta\left(s - \pi_x(\underline{y})\right) \, \mathrm{d}s \\ &= \mathcal{N}(\underline{y}; \phi_x(0, \pi_x(\underline{y})), \mathbf{C}_v) , \end{aligned} \tag{13}$$

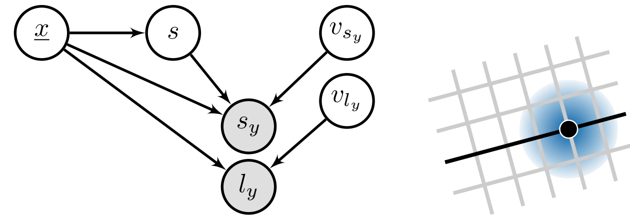
which is a corrupt expression in the sense that it has a cyclic dependency between measurement and its originating source. Thus, we intentionally abuse the notation $p(\underline{y}|\underline{x}, \underline{y})$ to emphasize this issue. It is well-known that (13) produces biased estimates for nonlinear constraints in the presence of noise. Effort has been made to understand this bias and to re-engineer it in order to reduce its effect [10], [15].

V. Partial Likelihood for Extended Objects

In this section, we propose a mathematically sound likelihood for extended objects based on the concept of *partial likelihood* [17], [18]. This likelihood yields an unbiased estimator while neither requiring assumptions about the measurement sources, nor requiring artificial re-engineering.

A. Key Idea

Let us now apply the constraint-induced parametrization as in Sec. III-B to the measurements $\phi_x(l_y, s_y) = \underline{y}$, too. Then l_y and s_y encode two types of measurement information, where the first is related to "how well" the measurement fits to the object boundary and the second refers to "where" on the boundary it is related to. Actually, designing the likelihood $p(\underline{y}|\underline{x})$ is difficult as we need a heuristic for the second type of information, e.g., spatial distribution (SDM) or greedy association (GAM). The key idea of our approach is using only the l_y information of a measurement for the Bayes update while putting ignorance on the critical s_y information.



(a) Desired Probabilistic Graphical Model. (b) ideal

Figure 6: Ideal case, where the constraint-induced parametrization makes the l_y and s_y dimensions (gray grid) independent.

Mathematically, this refers to the statistical concept of partial likelihood, which is an approximation of the full likelihood

$$\begin{aligned} p(l_y, s_y|\underline{x}) &= p(l_y|\underline{x}, s_y) \cdot p(s_y|\underline{x}) \\ &\approx p(l_y|\underline{x}, s_y) . \end{aligned} \tag{14}$$

By dropping $p(s_y|\underline{x})$ we put ignorance on the association heuristics. Of course, we have to trade this ignorance for the amount of measurement information encoded in s_y. The effect of this trade-off will be discussed in Sec. VII.

In addition, as the partial likelihood $p(l_y|\underline{x}, s_y)$ still relies on s_y to some degree, l_y preferably should become independent of the critical measurement dimension

$$p(l_y|\underline{x}, s_y) = p(l_y|\underline{x}) . \tag{15}$$

Hence, our tasks are i) finding a parametrization $\phi_x(l_y, s_y)$ for the measurement $\underline{y}$, such that l_y and s_y are mutually independent and ii) deriving the partial likelihood $p(l_y|\underline{x})$.

B. Finding the Parameterization

Measurements $\underline{y}$ of a source $\phi_x(0, s)$ on the constraint follow the additive noise model from (1) according to

$$\underline{y} = \phi_x(0, s) + \underline{v} , \tag{16}$$

When expressing these measurements in constraint-induced parametrization $\phi_x(l_y, s_y) = \underline{y}$, their l_y and s_y component are statistically independent if the measurement noise $\underline{v}$ is independent in their parametrization. The following example explains what that means.

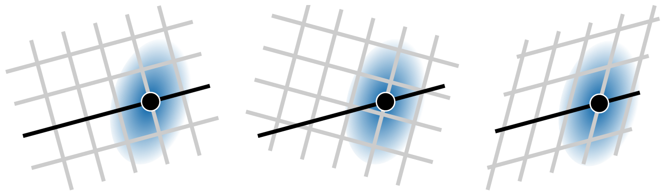
Example 2 (Linear Constraint)
In Fig. 6b, a linear constraint is shown together with its constraint-induced parametrization and a specific source. Then, measurements, which originate from this source according to isotropic Gaussian noise with $\mathbf{C}_v = \sigma^2 \cdot \mathbf{I}$ are independent in their l_y and s_y component and we can write

$$p(l_y|\underline{x}) = \mathcal{N}(l_y; 0, \sigma^2) \tag{17}$$

with $l_y = g_x(\underline{y})$, and

$$p(s_y|\underline{x}) = \int_S \mathcal{N}(s_y - s; 0, \sigma^2) \cdot p(s|\underline{x}) \, \mathrm{d}s . \tag{18}$$

Note that all information about s is isolated in the second term (18) *Thus, by using the partial likelihood* (17)*, we do not need to explicitly model s anymore. For this ideal case, the probabilistic graphical model from Fig. 4 changes to the one in Fig. 6a. As a remark, maximizing the partial likelihood* (18)



(a) constraint-induced (b) noise-induced (c) compromise

Figure 7: Competing designg properties for the desired parametrization.

corresponds to orthogonal least squares as the distance along the normals to the constraint is minimized. [19] □

While for the case of a linear constraint (and isotropic noise), full independence could be achieved, for a nonlinear constraint (and/or isotropic noise), we find us in a dilemma. On the one hand, the constraint-induced parametrization dictates the coordinate axes to lie parallel and normal to the constraint, as indicated in Fig. 7a. On the other hand, the noise will only be independent for the case that the axes coincide with the *principal components* of the Gaussian measurement covariance $\mathbf{C}_v$ (see Fig. 7b). However, these noise-induced axes no longer have a meaningful interpretation of l_y and s_y in terms of "how well" and "where" measurements are related to the boundary.

As a compromise, we propose to sacrifice orthogonality of the coordinate system, in order to come up with the parametrization in Fig. 7c, which is still meaningful as the one in Fig. 7a yet less correlated as the one in Fig. 7b. Technically, the proposed parametrization is based on the signed Mahalanobis distance to the constraint

$$g_x(\underline{y}) := \pm \|\underline{e}^{\mathrm{T}} \mathbf{C}_v^{-1} \underline{e}\|^{\frac{1}{2}} \ , \tag{19}$$

with $\underline{e} = \underline{y} - \phi_x(0, \pi_x(\underline{y}))$ and the sign indicates on which side of the constraint the measurement lies. Note that $\pi_x(\underline{y})$, as defined in (11), specifies the most likely point $\phi_x(0, \pi_x(\underline{y}))$ on the constraint for a measurement $\underline{y}$, which essentially is the point with the smallest Mahalanobis distance. Then, we define the desired parametrization in Fig. 7c as

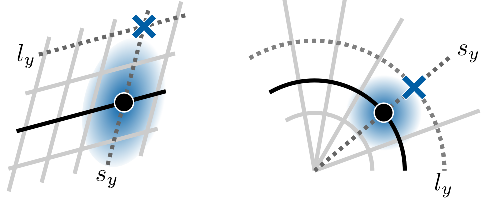
$$\begin{bmatrix} l_y \\ s_y \end{bmatrix} = \begin{bmatrix} g_x(\underline{y}) \\ \pi_x(\underline{y}) \end{bmatrix} \ . \tag{20}$$

That is, all points in the lines parallel to the constraint have equal Mahalanobis distance l_y and all points in the lines skewed to the constraint correspond to equal points $\phi_x(0, s_y)$ on the constraint. Fig. 8 shows this for several examples where the measurement covariances $\mathbf{C}_v$ are indicated as shaded blue ellipses/circles.

Note that, due to the skewness of the coordinate system and potential nonlinearities in the constraint, the proposed parametrization (20) still retains a degree of correlation between l_y and s_y that must be considered in the partial likelihood.

C. Deriving the Partial Likelihood

The remaining correlation between l_y and s_y in the proposed parametrization causes that $p(l_y|\underline{x}, s_y)$ generally cannot be simplified to $p(l_y|\underline{x})$. In consequence, taking the s_y component of a measurement as a given parameter refers to the point $\phi_x(0, s_y)$ on the constraint, marked as black circles in Fig. 8.



(a) anisotropic noise (b) nonlinear constraint

Figure 8: Proposed parametrization applied to anisotropic noise (a) and nonlinear constraint (b).

Our next step is to investigate how likely it is to measure l_y for measurements $\underline{\tilde{y}}$ that are produced by $\phi_x(0, s_y)$. For this purpose, we again apply the additive noise model from (1) $\underline{\tilde{y}} = \phi_x(0, s_y) + \underline{v}$, which refers to

$$p(\underline{\tilde{y}}|\underline{x}, s_y) = \mathcal{N}\left(\underline{\tilde{y}}; \phi_x(0, s_y), \mathbf{C}_v\right) \ . \tag{21}$$

However, within the proposed parametrization $\underline{\tilde{y}} = \phi_x(\tilde{l}_y, \tilde{s}_y)$, we are only interested in the probability of measurements fulfilling $\tilde{l}_y = l_y$. In Fig. 8 this corresponds the probability of measurements lying in the dashed l_y line.

Deriving this probability requires integrating $p(\underline{\tilde{y}}|\underline{x}, s_y)$ over the l_y-line. After applying a change of variables from $\underline{\tilde{y}}$ to $\tilde{l}_y$ and $\tilde{s}_y$ as explained in Lemma 1, we obtain

$$\begin{aligned} & p(\tilde{l}_y, \tilde{s}_y|\underline{x}, s_y) \\ & = p(\underline{\tilde{y}}|\underline{x}, s_y) \cdot \left|\det\left(\mathbf{J}_{\phi_x}(\tilde{l}_y, \tilde{s}_y)\right)\right| \\ & = \mathcal{N}\left(\phi_x(\tilde{l}_y, \tilde{s}_y); \phi_x(0, s_y), \mathbf{C}_v\right) \cdot \left|\det\left(\mathbf{J}_{\phi_x}(\tilde{l}_y, \tilde{s}_y)\right)\right| \ , \end{aligned} \tag{22}$$

where $\mathbf{J}_{\phi_x}(\tilde{l}_y, \tilde{s}_y)$ is the Jacobian matrix of the transformation ϕ_x according to

$$\mathbf{J}_{\phi_x}(\tilde{l}_y, \tilde{s}_y) = \frac{\mathrm{d}\underline{\tilde{y}}}{\mathrm{d}\tilde{l}_y \ \mathrm{d}\tilde{s}_y} = \frac{\mathrm{d}\phi_x(\tilde{l}_y, \tilde{s}_y)}{\mathrm{d}\tilde{l}_y \ \mathrm{d}\tilde{s}_y} \ , \tag{23}$$

and takes into account the skewness of the proposed coordinate system. Finally, integrating $\tilde{s}_y$ out of (22), we arrive at the desired partial likelihood that defines the new model.

Definition 1 (Partial Likelihood for Extended Objects) *For a given state $\underline{x}$ and a given measurement $\underline{y} = \phi_x(l_y, s_y)$ parametrized in terms of its Mahalanobis distance (20) to the object boundary with respect to the noise covariance matrix $\mathbf{C}_v$, the partial likelihood for l_y is*

$$\begin{aligned} & p(l_y|\underline{x}, s_y) = \\ & \int_S \mathcal{N}\left(\phi_x(l_y, \tilde{s}_y); \phi_x(0, s_y), \mathbf{C}_v\right) \cdot \left|\det\left(\mathbf{J}_{\phi_x}(l_y, \tilde{s}_y)\right)\right| \mathrm{d}\tilde{s}_y \ . \end{aligned} \tag{24}$$

□

Conceptionally, partial likelihood $p(l_y|\underline{x}, s_y)$ is somewhat related to the traditional GAM $p(\underline{y}|\underline{x}, \underline{y})$ from (13). However, while the traditional GAM derives the likelihood of a measurement producing itself (roughly speaking), the new model derives the likelihood of one dimension of a measurement producing another dimension. In order to demonstrate the new model let us look again at the circle from Example 1.

Example 3 (Circle)

For this example, let us assume isotropic measurement noise with $\mathbf{C}_v = \sigma^2 \cdot \mathbf{I}$. Then, $\phi_x(l_y, s_y)$ can be specified by (8). The determinant of the Jacobian of $\phi_x(l_y, s_y)$ evaluates to

$$\det\left(\mathbf{J}_\phi(l_y, s_y)\right) = \det\left(\begin{bmatrix} \cos(s_y) & -(r+l_y)\cdot\sin(s_y) \\ \sin(s_y) & (r+l_y)\cdot\cos(s_y)\end{bmatrix}\right) = r + l_y \ . \qquad (25)$$

Then, plugging (8) *and* (25) *into* (24) *yields*

$$p(l_y|\underline{x}, s_y) = \qquad (26)$$
$$\int_S \mathcal{N}\left((r+l_y)\begin{bmatrix}\cos(\tilde{s}_y)\\ \sin(\tilde{s}_y)\end{bmatrix}; r\begin{bmatrix}\cos(s_y)\\ \sin(s_y)\end{bmatrix}, \mathbf{C}_v\right)(r+l_y)\,\mathrm{d}\tilde{s}_y \ .$$

It is interesting to note that this likelihood actually is the line integral along a circle with radius $r + l_y$ over an isotropic Gaussian centered on a circle with radius r at angle s_y. Due to the isotropic character of the noise, this integral is independent of the specific instance of s_y, which allows us to set $s_y = 0$ and to obtain $p(l_y|\underline{x})$. □

However, analytic solutions of the integral in (24) generally cannot be found, not even for the circle and the case of isotropic noise (26), which raises the need for approximation techniques.

VI. Implementation

In this section, we show how to implement the new model. Specifically, we derive i) a measurement equation with non-additive noise, and ii) its approximation as a measurement equation with additive noise, and iii) its explicit sampling-based representation as a likelihood.

A. Measurement Equation with Non-additive Noise

Roughly speaking, given a measurement $\underline{y} = \phi_x(l_y, s_y)$ in the proposed parametrization (20), we want to know the probability of the point $\phi_x(0, s_y)$ on the object boundary to produce a measurement $\tilde{\underline{y}} = \phi_x(0, s_y) + \underline{v}$ with $g_x(\tilde{\underline{y}}) = l_y$. This equation, in turn, can be rearranged to a generative measurement model with non-additive noise in classical notation

$$\begin{aligned} 0 = h(\underline{x}, \underline{v}, \underline{y}) &= l_y - g_x(\tilde{\underline{y}}) \\ &= g_x(\underline{y}) - g_x(\phi_x(0, \pi_x(\underline{y})) + \underline{v}) \ , \end{aligned} \qquad (27)$$

where $\underline{x}$ is the state, $\underline{v}$ is the non-additive Gaussian noise, $\underline{y}$ acts as a model parameter, and 0 is a *pseudo-measurement*.

B. Measurement Equation with Additive Noise

Assume we would know the probability distribution of $g_x(\tilde{\underline{y}})$ in (27). Then, we could define a Gaussian noise variable $w \sim \mathcal{N}(\mathrm{E}\{g_x(\tilde{\underline{y}})\}, \mathrm{Var}\{g_x(\tilde{\underline{y}})\})$ which, in turn, would let us define a generative measurement model with additive noise in classical notation

$$\begin{aligned} 0 = h(\underline{x}, w, \underline{y}) &= l_y - w \\ &= g_x(\underline{y}) - w \ , \end{aligned} \qquad (28)$$

where $\underline{x}$ is the state, w is the additive Gaussian noise, $\underline{y}$ acts as a model parameter, and 0 is a *pseudo-measurement*. Note that w actually depends on the state parameters $\underline{x}$. Thus, a point estimate for $\underline{x}$ is to be used in order to derive its moments. Next, we propose a sampling-based approach for this task.

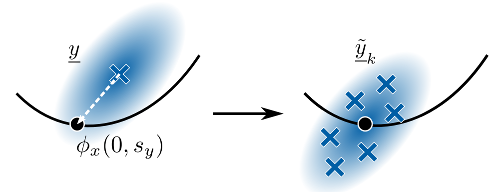


Figure 9: Sampling-based approximation of the new model.

C. Sampling-based Gaussian Approximation

The measurement equation with additive Gaussian noise from (28) immediately refers to the likelihood

$$p(l_y|\underline{x}, s_y) \approx \mathcal{N}\left(g_x(\underline{y}); \mathrm{E}\{g_x(\tilde{\underline{y}})\}, \mathrm{Var}\{g_x(\tilde{\underline{y}})\}\right) \ . \qquad (29)$$

We now propose an approach to derive the mean $\mathrm{E}\{g_x(\tilde{\underline{y}})\}$ and variance $\mathrm{Var}\{g_x(\tilde{\underline{y}})\}$ based on sampling. The idea, as shown in Fig. 9, is to simulate K measurements $\tilde{\underline{y}}$ of the point $\phi_x(0, s_y)$ and then derive the sample mean and variance of $g_x(\tilde{\underline{y}})$. This approach boils down to four steps.

1) calculate $s_y = \pi_x(\underline{y})$ using (11)
2) draw K samples $\{\underline{v}_1, \ldots, \underline{v}_K\}$ from $\mathcal{N}(\underline{v}; \underline{0}, \mathbf{C}_v)$
3) generate K measurement samples in the form of $\tilde{\underline{y}}_k = \phi_x(0, s_y) + \underline{v}_k$
4) calculate sample moments according to

$$\mathrm{E}\{g_x(\tilde{\underline{y}})\} \approx \frac{1}{K}\sum_{k=1}^{K} g_x(\tilde{\underline{y}}_k) \ , \qquad (30)$$
$$\mathrm{Var}\{g_x(\tilde{\underline{y}})\} \approx \frac{1}{K}\sum_{k=1}^{K}\left(g_x(\tilde{\underline{y}}_k)^2 - \mathrm{E}\{g_x(\tilde{\underline{y}})\}^2\right) .$$

As a remark, we recommend drawing the noise samples deterministically (e.g., by using [20]) as this allows for reproducible results while keeping the numbers of samples low (we used only $K = 5$ samples for 2D measurements).

VII. Evaluation

Let us now study the estimation quality of the new model. For this purpose, we i) show that using the model theoretically yields an unbiased estimator and ii) evaluate its performance in two recursive ellipse estimation experiments against the common models that either assume a uniform distribution or perform a greedy association.

A. Unbiasedness

Godambe and Thompson [5] defined that an *estimating equation* $g_x^*(\underline{y})$ is unbiased, if

$$\mathrm{E}\{g_x^*(\underline{y})\} = 0 \qquad (31)$$

holds for the true parameters $\underline{x}$. This definition was motivated by the fact that the likelihood $\mathcal{N}(g_x^*(\underline{y}); 0, \sigma^2)$ then has a maximum for the true parameters. In addition, they found that an estimating equation $g_x(\underline{y})$ that includes a nuisance parameter s can be modified according to $g_x^*(\underline{y}) = g_x(\underline{y}) - \mathrm{E}\{g_x(\tilde{\underline{y}})\}$, where the nuisance parameter is set to its maximum likelihood estimate, in order to make it unbiased in the sense of (31).

With this in mind, (29) can be interpreted as containing the unbiased estimating equation

$$\begin{aligned} g_x^*(\underline{y}) &= g_x(\underline{y}) - \mathrm{E}\{g_x(\tilde{\underline{y}})\} \\ &= g_x(\underline{y}) - \mathrm{E}\{g_x(\phi_x(0, \pi_x(\underline{y})) + \underline{v})\} \ , \end{aligned} \quad (32)$$

where the nuisance parameter s also has been set to its maximum likelihood estimate $\pi_x(\underline{y})$. In consequence, $g_x^*(\underline{y})$ is theoretically unbiased according to (31).

B. Ellipse Experiment

We consider recursively estimating a static ellipse $\underline{x} = [\alpha, \underline{t}^{\mathrm{T}}, a, b]^{\mathrm{T}}$ with angle $\alpha = \frac{\pi}{8}$, center $\underline{t}^{\mathrm{T}} = [\frac{1}{10}, \frac{4}{10}]$, and axes $a = 2$, and $b = 1$ based on 750 noisy point measurements of its boundary, which arrive sequentially in packages of five measurements. In a first experiment (E1), the measurement sources were uniformly drawn from the true ellipse and distorted by Gaussian noise with $\mathbf{C}_v = \mathrm{diag}(.2, .02)$. In a second experiment (E2), the sources were uniformly drawn from only a fraction of $\frac{2}{3}$ and distorted by Gaussian noise with $\mathbf{C}_v = \mathrm{diag}(.1, .01)$. We implemented an estimator based on our model against estimators based on the commonly used models (in the presence of anisotropic noise) from Sec. IV. Even though these experiments could be easily extended to a dynamic tracking task by moving the ellipse between the measurements, we decided to leave this component out, in order to fully concentrate on the measurement updates.

Modeling an Ellipse: Ignoring its position and orientation, an ellipse can be characterized by its semi-major axes a, and b. Each source on the ellipse then can be reached according to $\phi_x(0, s) = [a \cdot \cos(s), b \cdot \sin(s)]^{\mathrm{T}}$, where s is an angle from the interval $[0, 2\pi]$.

1) Spatial Distribution Model: For reference, we set up an SDM (10), where $p(s|\underline{x})$ was modeled by a uniform distribution over the entire ellipse boundary. Note that this uniform model is indeed true for E1, but not for E2. The integral was numerically evaluated, using a polygon approximation [10] for the ellipse with 72 vertices.

2) Proposed Model: For the proposed model (29), both components $l_y = g_x(\underline{y})$, and $s_y = \pi_x(\underline{y})$ of the proposed parametrization were calculated according to (11) and (19). However, as both equations require finding the most likely point on an ellipse, we also used a polygon approximation for speed-up. Mean $\mathrm{E}\{g_x(\tilde{y})\}$, and variance $\mathrm{Var}\{g_x(\tilde{y})\}$ were calculated using the sampling-based approach from (30) with 5 deterministic samples [20].

3) Traditional GAM: Finally, a GAM (13) was set up as a representative model for distance minimizing approaches.

Estimators: The measurement update step for the proposed model and the traditional GAM were implemented using a common *Unscented Kalman Filter* (UKF) [20] while the SDM required a more advanced likelihood-based filter, where we decided for a *Progressive Gaussian Filter* (PGF) [21]. Between the measurement update steps, we added prediction steps which inflate the state covariance according to a *random walk model*, in order to recover from local minima. The inflating covariance matrix was set to $\gamma \cdot \mathbf{I}$, where the factor γ logarithmically decreased from 10^{-2} to 10^{-6} from the first to the last prediction step. The ellipse state $\underline{x}$ was initialized as a circle from the first five measurements. The center $\underline{t}$ was set to their mean and a, b were both set to the maximum distance from the measurements to the center.

Results: Fig. 10 visualizes the ellipses for the average parameters over 100 runs for both experiments. These ellipses are representative, as their standard deviation through all runs and parameters was in a magnitude of 10^{-2}.

As can be seen, the new approach converges to the true ellipse in both experiments with an average systematic error in the magnitude of 10^{-2} through all parameters. In contrast, the SDM approach finds the true ellipse in E1, yet in E2 it produces a systematic error for position and semi-major axis in the magnitude of 10^{-1}. Conversely, the GAM approach manages to find the ellipse in E2, but in E1 it produces a systematic error for the semi-major axis also in the magnitude of 10^{-1}.

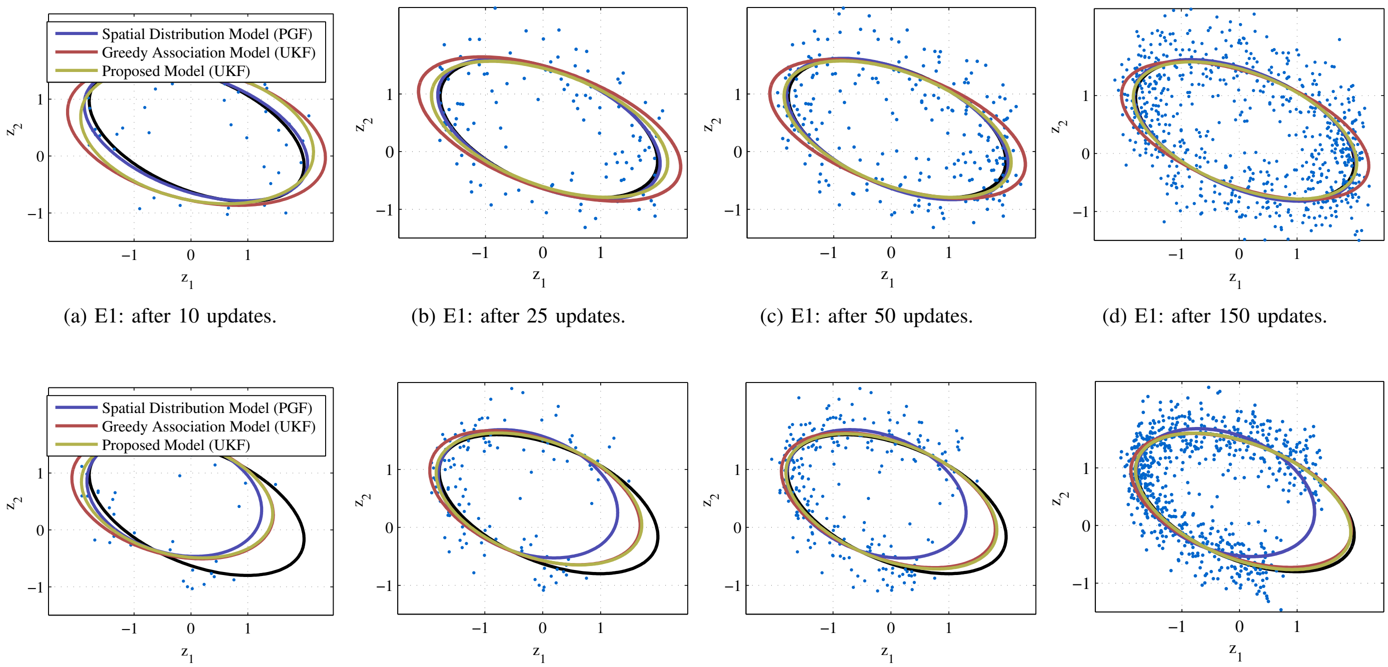
These behaviors have an intuitive interpretation: The SDM bias boils down to the correct (E1) and incorrect (E2) assumptions about $p(s|\underline{x})$, and the GAM bias is a result of its corrupt likelihood (see Sec. IV). As our approach neither requires assumptions on $p(s|\underline{x})$, nor builds upon a corrupt probabilistic model, it is not affected by these systematic errors.

However, in E1, the proposed approach has a slightly slower convergence compared to the SDM approach. We suspect this issue to be the price of ignoring the source distribution. In numerical terms, where the SDM approach finds the true ellipse right after ten update steps (Fig. 10a), our approach takes ~ 25 updates until convergence. Even though this drawback practically should be of minor importance, it must be kept in mind when designing the initialization procedure.

VIII. Conclusions

In this work, we proposed a partial likelihood for extended objects that can be used in situations when there is no knowledge about the distribution of the measurement sources. For this purpose, we developed a re-parametrization of the measurements to decouple their encoded information into "how well" they fit to the object boundary and "where" on the boundary they are related to. Then, ignoring the second type of information allowed us to design a partial likelihood that does not require incorporating a probability distribution on the measurement sources. The resulting model was shown to be theoretically unbiased and to outperform the state-of-the-art in Monte Carlo experiments.

Specifically, we considered the common task of recursively estimating the parameters of an ellipse based on sequentially arriving measurements with anisotropic Gaussian noise characteristics. We observed that our approach converges to the true ellipse, regardless of whether measurements originate only from a part of the ellipse boundary or whether they were affected by high anisotropic noise. In contrast, the common approaches of i) using a spatial distribution and ii) minimizing a distance to the boundary were either biased when applied to a non-uniform source distribution or biased in the case of high noise. In numerical terms, the systematic error in the estimated parameters could be reduced by a full order of magnitude compared to both traditional approaches.



(a) E1: after 10 updates. (b) E1: after 25 updates. (c) E1: after 50 updates. (d) E1: after 150 updates.

(e) E2: after 10 updates. (f) E2: after 25 updates. (g) E2: after 50 updates. (h) E2: after 150 updates.

Figure 10: Results of the ellipse experiments over 100 runs. While in experiment E1, measurements originate from the entire ellipse boundary, in E2, measurements originate only from $\frac{2}{3}$ of the boundary. Note the anisotropic character of the noise.

References

[1] 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.

[2] K. Kanatani, "Overviews of Optimization Techniques for Geometric Estimation," *Memoirs of the Faculty of Engineering, Okayama University*, vol. 47, no. January, pp. 1–18, 2013.

[3] N. Ohta, "Motion parameter estimation from optical flow without nuisance parameters," *Third IEEE Workshop on Statistical and Computational Theories of Vision*, 2003.

[4] J. Neyman and E. Scott, "Consistent Estimates Based on Partially Consistent Observations," *Econometrica: Journal of the Econometric Society*, vol. 16, no. 1, pp. 1–32, 1948.

[5] V. Godambe and M. Thompson, "Estimating Equations in the Presence of a Nuisance Parameter," *The Annals of Statistics*, vol. 2, no. 3, pp. 568–571, 1974.

[6] 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.

[7] 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.

[8] K. Granstrom, "Tracking Rectangular and Elliptical Extended Targets Using Laser Measurements," in *Proceedings of the 15th International Conference on Information Fusion (FUSION)*, Edinburgh, United Kingdom, 2011.

[9] N. Petrov, L. Mihaylova, A. Gning, and D. Angelova, "A Novel Sequential Monte Carlo Approach for Extended Object Tracking Based on Border Parameterisation," in *Proceedings of the 15th International Conference on Information Fusion (FUSION)*, Edinburgh, United Kingdom, 2011.

[10] F. Faion, A. Zea, and U. D. Hanebeck, "Reducing Bias in Bayesian Shape Estimation," in *IEEE 17th International Conference on Information Fusion (FUSION)*. Salamanca, Spain: IEEE, 2014, pp. 1–8.

[11] 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)*, Chicago, Illinois, USA, 2011, pp. 1–8.

[12] 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.

[13] B. C. Matei and P. Meer, "Estimation of nonlinear errors-in-variables models for computer vision applications." *IEEE transactions on pattern analysis and machine intelligence*, vol. 28, no. 10, pp. 1537–52, Oct. 2006.

[14] D. Gawley, A. van den Hengel, M. Brooks, and W. Chojnacki, "On the Fitting of Surfaces to Data with Covariances," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1294–1303, 2000.

[15] T. Okatani and K. Deguchi, "On Bias Correction for Geometric Parameter Estimation in Computer Vision," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 959–966, 2009.

[16] Z. Chen, "Bayesian filtering: From Kalman filters to particle filters, and beyond," *Statistics*, 2003.

[17] D. Cox, "Partial likelihood," *Biometrika*, 1975.

[18] W. Wong, "Theory of partial likelihood," *The Annals of Statistics*, vol. 14, no. 1, pp. 88–123, 1986.

[19] W. E. Deming, "The application of least squares," *Philosophical Magazine*, vol. 7, pp. 146–158, 1931.

[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] J. Steinbring and U. D. Hanebeck, "Progressive Gaussian Filtering Using Explicit Likelihoods," in *Proceedings of the 17th International Conference on Information Fusion (Fusion 2014)*, Salamanca, Spain, 2014.