The Tobit Kalman Filter: An Estimator for Censored Measurements

Bethany Allik, *Member, IEEE*, Cory Miller, Michael J. Piovoso, *Life Senior Member, IEEE*, and Ryan Zurakowski, *Senior Member, IEEE*

***Abstract*— Tobit model censored data arise in multiple engineering applications through saturating sensors, limit-of-detection effects, and image frame effects. In this brief, we introduce a novel formulation of the Kalman filter for Tobit Type 1 censored measurements. Our proposed formulation, called the Tobit Kalman filter, is identical to the standard Kalman filter in the no-censoring case. At or near the censored region, the Tobit Kalman filter utilizes a local approximation of the probability of censoring in order to provide a fully recursive estimate of the state and state error covariance. The additional computational burden of the method compared with the standard Kalman filter is limited to the calculation of m normal probability density functions and m normal cumulative density functions per update, where m is the number of measurements.**

***Index Terms*— Censored data, Kalman filtering, output nonlinearity, recursive estimation, Tobit model.**

I. Introduction

The Kalman filter [1] has become ubiquitous in tracking and estimation. Many estimation applications, especially those using low cost commercial off-the-shelf sensors, are subject to a specific type of measurement nonlinearity called censoring. Censoring frequently takes the form of limit-of-detection, occlusion region, or sensor saturation. All these forms of censoring are known as Tobit Type 1 censoring [2]. In this brief, we present the first fully recursive formulation of the Kalman filter for estimating the state variables from censored data. In order to develop such a recursive estimator, we rely on two novel formulations and one local approximation (detailed in Section IV, Assumption 1). We introduce a new definition of the measurement innovation, computed as the difference between the measured value and the expected measurement value conditioned on the state estimate. We introduce a binary random variable to track the probability of a measurement being censored, conditioned on the state. We then employ the local approximation that this probability of measurement censoring can be adequately estimated using the state estimate rather than the state. We then compute the optimal (subject to the local approximation) recursive update algorithm for the state and state covariance estimates for Tobit model censored measurements. We refer to this formulation as the Tobit Kalman filter.

Manuscript received November 7, 2014; revised April 17, 2015; accepted May 3, 2015. Date of publication June 1, 2015; date of current version December 21, 2015. Manuscript received in final form May 8, 2015. This work was supported by the U.S. Army Research Office under Grant W911NF-10-1-0386. Recommended by Associate Editor A. Chiuso.

B. Allik, C. Miller, and R. Zurakowski are with the Department of Electrical and Computer Engineering, University of Delaware, Newark, DE 19711 USA (e-mail: ballik@udel.edu; cimiller@udel.edu; ryanz@udel.edu).

M. J. Piovoso is with Pennsylvania State University, Malvern, PA 19355 USA (e-mail: piovoso@psu.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCST.2015.2432155

Despite many obvious examples of censored data in estimation and tracking, the Tobit model has not received much attention in either the field of signal processing or control theory. It has been widely used, however, in medicine and economics. The model was first proposed in [2] as a regression model for household expenditure, where the dependent variable could not be observed below a certain limit. The standard Tobit model formulation is

$$y_t = \begin{cases} \beta x_t + u_t, & \beta x_t + u_t > \tau \\ \tau, & \beta x_t + u_t \leq \tau \end{cases} \tag{1}$$

where $\beta \in \mathbb{R}^{1\times n}$ is a vector of constants, $x_t \in \mathbb{R}^{n\times 1}$ is the input vector at time t, y_t is a scalar output, and u_t is a Gaussian random number with zero mean and variance σ_u^2. The use of ordinary least squares to estimate the β or σ_u from the output would be inconsistent because the entire population of the dependent variable is not being observed.

Many methods have been devised to solve for the parameters in the Tobit model, including Tobin's original maximum-likelihood estimator [3]–[9]. An analysis of the method and the consistency of the estimates is presented in [7]. Most of these methods require a knowledge of the entire measurement history; a recursive estimator such as the Kalman filter has been lacking for this type of censored measurement nonlinearity.

The Tobit model is a special case of a Wiener model [10]. A recursive identification method for a Weiner model nonlinearity is presented in [11]. However, when the output is censored, the identification update is halted because the input has driven the output to a slope zero region. The method is therefore inappropriate for censored data systems.

A similar measurement nonlinearity that has been previously considered is the problem of intermittent measurement. A formulation of the Kalman filter designed for the intermittent measurement case is presented in [12] and [13]. This formulation reduces to a Kalman filter with no measurement update when the measurements are missing. The estimator in [12] provides the minimum state error variance filter, given all past observations and arrival sequences.

The formulation for intermittent measurements relies on the assumption that missed measurements were uncorrelated with the state value. Censored measurements, however, are correlated with the state values. If the Kalman filter for intermittent measurements is used on a censored data system, the resulting estimates of the state will be biased.

Another attempt to solve the censored measurement problem was presented in [14]. This formulation treats the censored and noncensored measurements differently, and has a formulation that is not recursive (it requires knowledge of the entire history of censored measurements); recursion is a major motivation for using a Kalman filter. A different kind of censoring in distributed detection systems has been studied in [15] and [16]. The censored data are either sent or not sent to a fusion center based on its informativeness. This is a different type of censoring from what we consider in this brief.

One difficulty in using a Kalman filter for censored measurements is that the measurement noise is non-Gaussian near the censoring region. Suppose the state variable is a constant near the censored region, noise on the measurements will cause some of the measurements to be censored. The standard Kalman filter will produce a biased estimate of the state because of the assumption that the noise is zero mean Gaussian. Previous research has also designed estimators using the likelihood function that allow the Gaussian observation noise to be state dependent. In [17], an iterative Kalman filter was created to solve the nonlinear least square problem using the likelihood function. In this brief, the censored measurement case can be estimated using a linear Kalman filter.

Censoring can be generalized as an output nonlinearity, and general output nonlinearities can be addressed using either the extended Kalman filter (EKF), unscented Kalman filter (UKF), or the particle filter. A comparison of the EKF and UKF for censored data is presented in [18]. One example of the particle filter formulated for partially observed Gaussian state-space models is presented in [19]. Particle filters are much more computationally expensive than the Kalman filter because they require the use of a weighted set of samples called particles to generate the posteriori distribution $p(x_k|y_{1:k})$. The method described in this brief avoids the use of numerical approximation methods such as the particle filter by directly computing the relevant posteriori distributions from the censored data measurement model. The resulting Tobit Kalman filter has a similar computational burden to the standard Kalman filter, making it practical in computation-limited environments such as embedded systems.

In this brief, we introduce a novel formulation of the Kalman filter for censored data applications. Many applications for this kind of filter include MEMS sensor tracking with saturation, visual tracking with camera frame censoring [20], and biological measurements with a limit of detection saturation [21]–[25]. This new formulation provides recursive estimates of latent state variables from censored observations.

This brief is divided into seven sections. Section II defines the problem and notation. Section III explains the Tobit model in detail. Section IV derives the Tobit Kalman filter. The predict stage of the Tobit Kalman filter remains unaltered from the Kalman filter while the update stage is rederived based on two assumptions. Section V shows that our filter is equivalent to the standard Kalman filter when the distance between the state estimate and the censoring threshold is large relative to the standard deviation of the measurement noise. Section VI contains the simulation results including the estimation of a constant near a censoring limit, estimating a Brownian motion model and a sinusoidal model.

II. Problem Formulation

Consider the evolution of a scalar output state sequence

$$
\begin{aligned}
x_k &= Ax_{k-1} + w_{k-1} \\
y_k^* &= Cx_k + v_k \\
y_k &= \begin{cases} y_k^*, & y_k^* > \tau \\ \tau, & y_k^* \le \tau \end{cases}
\end{aligned} \tag{2}
$$

where $x_k \in \mathbb{R}^{n\times 1}$ is the state vector, y_k is the scalar measurement, $A \in \mathbb{R}^{n\times n}$ is the state transition matrix, and $C \in \mathbb{R}^{1\times n}$ is the measurement vector. w_k and v_k are the Gaussian random vectors with zero mean and covariance $Q \in \mathbb{R}^{n\times n}$ and $R = \sigma^2$, respectively, where σ is the standard deviation of the measurement noise. The Kalman filter is optimal in the Gaussian sense; however, when the noise distribution on y_k is a censored Gaussian, the filter is suboptimal, and since the noise is correlated to the state value, there is a violation of the assumptions of the Kalman filter. The closer the state is to the threshold value the more censored the Gaussian distribution on y_k becomes.

III. Tobit Regression

Using (2), we define y_k as the censored observation and y_k^* as the latent variable. The probability distribution of a censored variable with a normally distributed noise is

$$
\begin{aligned}
f(y_k|x_k) = &\frac{1}{\sigma}\phi\left(\frac{y_k - Cx_k}{\sigma}\right)u(y_k - \tau) \\
&+ \delta(\tau - y_k)\Phi\left(\frac{\tau - Cx_k}{\sigma}\right)
\end{aligned} \tag{3}
$$

where

$$
\phi\left(\frac{y_k - Cx_k}{\sigma}\right) = \frac{1}{\sqrt{2\pi}}e^{-\frac{(y_k - Cx_k)^2}{2\sigma^2}} \tag{4}
$$

and

$$
\Phi\left(\frac{y_k - Cx_k}{\sigma}\right) = \int_{-\infty}^{y_k} \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(z_k - Cx_k)^2}{2\sigma^2}}dz_k \tag{5}
$$

are the probability density function (pdf) and the cumulative distribution function (cdf) of a Gaussian random variable whose mean is Cx_k. $\delta(\cdot)$ is the Dirac delta function and $u(\cdot)$ is the unit step function.

The likelihood function for the standard Tobit model is

$$
L = \prod_{y_k^* \le \tau}\left[1 - \Phi\left(\frac{Cx_k - \tau}{\sigma}\right)\right]\prod_{y_k^* \ge \tau}\sigma^{-1}\phi\left(\frac{y_k - Cx_k}{\sigma}\right) \tag{6}
$$

as formulated in [2].

The expected value of the measurements when uncensored is given by

$$
\begin{aligned}
E(y_k|y_k > \tau, x_k, \sigma) &= \sigma^{-1}\int_{\tau}^{+\infty} z\frac{\phi\left(\frac{z - Cx_k}{\sigma}\right)}{1 - \Phi\left(\frac{\tau - Cx_k}{\sigma}\right)}dz \\
&= Cx_k + \sigma\lambda((\tau - Cx_k)/\sigma)
\end{aligned} \tag{7}
$$

this differs from the true value of the latent variable by a bias of $\sigma\lambda((\tau - Cx_k)/\sigma)$, where $\lambda(\alpha) = (\phi(\alpha)/[1-\Phi(\alpha)])$ is the inverse Mills ratio [7].

The expected measured value when both uncensored and censored measurements are included is

$$\begin{aligned} E[y_k|x_k,\sigma] &= P[y_k > \tau|x_k,\sigma]E[y_k|y_k > \tau, x_k,\sigma] \\ &\quad + P[y_k = \tau|x_k,\sigma]E[y_k|y_k = \tau, x_k,\sigma] \\ &= \Phi\left(\frac{Cx_k - \tau}{\sigma}\right)\left[Cx_k + \sigma\lambda\left(\frac{\tau - Cx_k}{\sigma}\right)\right] \\ &\quad + \Phi\left(\frac{\tau - Cx_k}{\sigma}\right)\tau. \end{aligned} \tag{8}$$

The variance of the expected measured value is derived in [26] and can be written as

$$\begin{aligned} &\mathrm{Var}[y_k|y_k > \tau, x_k,\sigma] \\ &\quad = E\left[y_k^2|y_k > \tau, x_k,\sigma\right] - [E[y_k|y_k > \tau, x_k,\sigma]]^2 \end{aligned} \tag{9}$$

$$\begin{aligned} &E\left[y_k^2|y_k > \tau, x_k,\sigma\right] \\ &\quad = \sigma^{-1}\frac{1}{1-\Phi\left(\frac{\tau - Cx_k}{\sigma}\right)}\int_\tau^{+\infty} z^2\phi\left(\frac{z - Cx_k}{\sigma}\right)dz \end{aligned} \tag{10}$$

so

$$\mathrm{Var}[y_k|y_k > \tau, x_k,\sigma] = \sigma^2\left[1 - \eth\left(\frac{\tau - Cx_k}{\sigma}\right)\right] \tag{11}$$

where

$$\begin{aligned} &\eth\left(\frac{\tau - Cx_k}{\sigma}\right) \\ &\quad = \lambda\left(\frac{\tau - Cx_k}{\sigma}\right)\left[\lambda\left(\frac{\tau - Cx_k}{\sigma}\right) - \left(\frac{\tau - Cx_k}{\sigma}\right)\right]. \end{aligned} \tag{12}$$

Note that $\mathrm{Var}[y_k|x_k,\sigma] = \mathrm{Var}[y_k|y_k > \tau, x_k,\sigma]$ since $\mathrm{Var}[y_k|y_k < \tau, x_k,\sigma] = 0$.

IV. Tobit Kalman Filter

In this section, we derive the Kalman formulation for Tobit censored measurements. The derivation is similar to the derivation for the standard Kalman filter; however, the censoring results in new definitions for the measurement residual, and consequently for the optimal Kalman gain and the estimated state covariance. Following is the notation for our model with the state vector $\mathbf{x_k} \in \mathbb{R}^{n\times 1}$, and the measurement vector $\mathbf{y_k} \in \mathbb{R}^{m\times 1}$:

$$\begin{aligned} \mathbf{x}_k &= \mathbf{Ax_{k-1}} + \mathbf{w_{k-1}} \\ \mathbf{y_k^*} &= \mathbf{Cx_k} + \mathbf{v_k} \\ \mathbf{y_k} &= \begin{cases} \mathbf{y_k^*}, & \mathbf{y_k^*} > \mathcal{T} \\ \mathcal{T}, & \mathbf{y_k^*} \le \mathcal{T}. \end{cases} \end{aligned} \tag{13}$$

The matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$ is the state transition matrix, $\mathbf{C} \in \mathbb{R}^{m\times n}$ is the measurement model, and $\mathcal{T} \in \mathbb{R}^{m\times 1}$ is the vector of threshold values. The noises $\mathbf{w_k}$ and $\mathbf{v_k}$ are the zero mean white Gaussian noise with covariance matrices $\mathbf{Q} \in \mathbb{R}^{n\times n}$ and $\mathbf{R} \in \mathbb{R}^{m\times m}$, respectively.

A. Predict Stage

The predict equation of the state may be written as

$$\mathbf{E}(\mathbf{x_k}|\mathbf{y_{1:k-1}}) = \mathbf{E}(\mathbf{Ax_{k-1}} + \mathbf{w_k}) = \mathbf{Ax_{k-1|k-1}} \tag{14}$$

where $\mathbf{x_{k-1|k-1}}$ is the estimate of $\mathbf{x_{k-1}}$, given all estimates and measurements up to time $k-1$. The state error covariance given measurements and state information up to time $k-1$ may be written as

$$\begin{aligned} \mathbf{cov}(\mathbf{x_k} - \mathbf{x_{k|k-1}}) &= \mathbf{cov}(\mathbf{Ax_{k-1}} + \mathbf{w_k} - \mathbf{Ax_{k-1|k-1}}) \\ &= \mathbf{A}\mathbf{cov}(\mathbf{x_{k-1}} - \mathbf{x_{k-1|k-1}})\mathbf{A^T} + \mathbf{Q} \\ &= \mathbf{A}\Psi_{\mathbf{k-1|k-1}}\mathbf{A^T} + \mathbf{Q} \end{aligned} \tag{15}$$

where $\mathbf{Q}$ is the model covariance matrix and $\Psi_{\mathbf{k-1|k-1}}$ is the previous *a posteriori* estimate of the state error covariance.

B. Update Stage

The optimal Kalman filter must minimize the state error covariance, $\Psi_{\mathbf{k|k}}$. The update stage corrects the state estimate using the current measurements. The update step will reduce the state error covariance, whereas the predict step will result in a widening of the state error covariance. The Kalman correction step to obtain the current estimate, given all observations up to time k, may be written as

$$\mathbf{x_{k|k}} = \mathbf{x_{k|k-1}} + \mathbf{K_k}(\mathbf{y_k} - \mathbf{E}(\mathbf{y_k}|\mathbf{x_{k|k-1}})). \tag{16}$$

The value of $E(y_k)$ was calculated for a scalar case for a censored value in (8); in this notation, $\mathbf{E}(\mathbf{y_k}|\mathbf{x_{k|k-1}}) \in \mathbb{R}^{m\times 1}$ is a vector, denoted as $\mathbf{E}(\mathbf{y_k})$ for the rest of this section. Each scalar component of $\mathbf{E}(\mathbf{y_k})$ can be censored at any given time and will have different threshold limits $\mathcal{T} = [\tau(1), \tau(2), \ldots, \tau(m)]$ with $\tau(l)$ and $y_k(l)$ representing the lth component of arrays $\mathcal{T}$ and $\mathbf{y_k}$, respectively.

To find $\mathbf{K_k}$ in (16), we minimize the state error covariance

$$\begin{aligned} \Psi_{\mathbf{k|k}} &= \mathbf{cov}(\mathbf{x_k} - \mathbf{x_{k|k}}) \\ &= \mathbf{cov}(\mathbf{x_k} - \mathbf{x_{k|k-1}} - \mathbf{K_k}(\mathbf{y_k} - \mathbf{E}(\mathbf{y_k}))). \end{aligned} \tag{17}$$

A Bernoulli random variable is introduced to model the occurrence of a censored measurement versus an actual measurement. The variable $p_k(l) = 1$ when the measurement is not censored and $p_k(l) = 0$ when the measurement is equal to the threshold value. The measurement model can be written as

$$p_k(l) = \begin{cases} 1, & Cx_k(l) + \upsilon_k(l) > \tau(l) \\ 0, & Cx_k(l) + \upsilon_k(l) \le \tau(l). \end{cases} \tag{18}$$

At any given time step, the measurement will represent the state by $Cx_k(l) + \upsilon_k(l)$ with probability $E(p_k(l))$. In matrix notation, the Bernoulli random matrix will be diagonal $\mathbf{p_k} \in \mathbb{R}^{m\times m}$ so the measurements will be arrived by the following equation:

$$\mathbf{y_k} = \mathbf{p_k}(\mathbf{Cx_k} + \mathbf{v_k}) + (\mathbf{I_{m\times m}} - \mathbf{p_k})\mathcal{T} \tag{19}$$

where $\mathbf{I_{m\times m}}$ is the identity matrix. Substituting into (16) yields

$$\begin{aligned} \Psi_{\mathbf{k|k}} &= \mathbf{cov}(\mathbf{x_k} - \mathbf{x_{k|k}}) \\ &= \mathbf{cov}\left(\begin{matrix} \mathbf{x_k} - \mathbf{x_{k|k-1}} \\ -\mathbf{K_k}\left(\begin{matrix} \mathbf{p_k}(\mathbf{Cx_k} + \mathbf{v_k}) \\ + (\mathbf{I_{m\times m}} - \mathbf{p_k})\mathcal{T} \\ -\mathbf{E}(\mathbf{y_k}) \end{matrix}\right) \end{matrix}\right). \end{aligned} \tag{20}$$

To simplify the notation in the derivation, we set the Kalman error to

$$\mathbf{\tilde{y}_k} = \mathbf{p_k}(\mathbf{Cx_k} + \mathbf{v_k}) + (\mathbf{I_{m\times m}} - \mathbf{p_k})\mathcal{T} - \mathbf{E}(\mathbf{y_k}) \tag{21}$$

so the covariance of the state estimate becomes

$$\begin{aligned}\Psi_{\mathbf{k|k}} &= \mathbf{E}\big((\mathbf{x_k} - \mathbf{x_{k|k-1}} - \mathbf{K_k\tilde{y}_k})(\mathbf{x_k} - \mathbf{x_{k|k-1}} - \mathbf{K_k\tilde{y}_k})^{\mathbf{T}}\big)\\ &= \Psi_{\mathbf{k|k-1}} - \mathbf{E}\big((\mathbf{x_k} - \mathbf{x_{k|k-1}})\mathbf{\tilde{y}_k^T}\big)\mathbf{K_k^T}\\ &\quad - \mathbf{K_k E}\big(\mathbf{\tilde{y}_k}(\mathbf{x_k} - \mathbf{x_{k|k-1}})^{\mathbf{T}}\big) + \mathbf{K_k E}\big(\mathbf{\tilde{y}_k\tilde{y}_k^T}\big)\mathbf{K_k^T}\end{aligned} \tag{22}$$

with

$$\Psi_{\mathbf{k}}|_{\mathbf{k-1}} = \mathbf{E}\big((\mathbf{x_k} - \mathbf{x_{k|k-1}})(\mathbf{x_k} - \mathbf{x_{k|k-1}})^{\mathbf{T}}\big) \tag{23}$$

$$\mathbf{R_{\tilde{x}\tilde{y}_k}} = \mathbf{E}\big((\mathbf{x_k} - \mathbf{x_{k|k-1}})\mathbf{\tilde{y}_k^T}\big) \tag{24}$$

$$\mathbf{R_{\tilde{y}\tilde{y}_k}} = \mathbf{E}\big(\mathbf{\tilde{y}_k\tilde{y}_k^T}\big) \tag{25}$$

where the Kalman gain is $\mathbf{K_k} = \mathbf{R_{\tilde{x}\tilde{y}_k}}\mathbf{R_{\tilde{y}\tilde{y}_k}^{-1}}$.

In a standard linear Kalman filter, the values of $\mathbf{R_{\tilde{x}\tilde{y}}}$ and $\mathbf{R_{\tilde{y}\tilde{y}}}$ are the functions of Ψ, $\mathbf{H}$, and $\mathbf{R}$. Because our measurements are not linearly related to the state vector in or around a censored region, we must explicitly find the values for $\mathbf{R_{\tilde{x}\tilde{y}}}$ and $\mathbf{R_{\tilde{y}\tilde{y}}}$. The function for $\mathbf{R_{\tilde{x}\tilde{y}}}$ is

$$\begin{aligned}\mathbf{R_{\tilde{x}\tilde{y}_k}} &= \mathbf{E}\left((\mathbf{x_k} - \mathbf{x_{k|k-1}})\begin{pmatrix}(\mathbf{Cx_k} + \mathbf{v_k})^{\mathbf{T}}\mathbf{p_k}\\ +\mathcal{T}^{\mathbf{T}}(\mathbf{I_{m\times m}} - \mathbf{p_k})\\ -\mathbf{E}(\mathbf{y_k})^{\mathbf{T}}\end{pmatrix}\right)\\ &= \mathbf{E}\big(\mathbf{x_k x_k^T C^T p_k}\big) + \mathbf{E}\big(\mathbf{x_k v_k^T p_k}\big)\\ &\quad + \mathbf{E}\big(\mathbf{x_k}\mathcal{T}^{\mathbf{T}}(\mathbf{I_{m\times m}} - \mathbf{p_k})\big) - \mathbf{E}\big(\mathbf{x_k E}(\mathbf{y_k})^{\mathbf{T}}\big)\\ &\quad - \mathbf{E}\big(\mathbf{x_{k|k-1}x_k^T C^T p_k}\big) - \mathbf{E}\big(\mathbf{x_{k|k-1}v_k^T p_k}\big)\\ &\quad - \mathbf{E}\big(\mathbf{x_{k|k-1}}\mathcal{T}^{\mathbf{T}}(\mathbf{I_{m\times m}} - \mathbf{p_k})\big)\\ &\quad + \mathbf{E}(\mathbf{x_{k|k-1}})\mathbf{E}(\mathbf{y_k})^{\mathbf{T}}.\end{aligned} \tag{26}$$

The probability of the measurement being uncensored is a function of the distance between the latent measured variable and the threshold value. The expected value of $p_k(l,l)$ may be written as

$$E(p_k(l,l)) = \Phi\left(\frac{Cx_k(l) - \tau(l)}{\sigma(l)}\right) \tag{27}$$

where $Cx_k(l)$ is the lth element of the measurement vector and $\sigma(l)$ is the variance of the noise on that element. In principle, this requires knowledge of the true state value. The following assumption allows us to relax this dependence and use the estimated state value instead.

Assumption 1: For small state estimation errors, the state prediction provides a reasonably accurate estimate of the probability of censoring. For the purposes of this derivation, we will use the approximation

$$E(p_k(l,l)) = \Phi\left(\frac{Cx_k(l) - \tau(l)}{\sigma(l)}\right) \approx \Phi\left(\frac{Cx_{k|k-1}(l) - \tau(l)}{\sigma(l)}\right). \tag{28}$$

To the extent that Assumption 1 holds true, $\mathbf{x_k}$ and $\mathbf{p_k}$ are independent, resulting in the following approximations:

$$\begin{aligned}\mathbf{E}(\mathbf{x_k p_k}|\mathbf{y_{1:k-1}}) &\approx \mathbf{E}(\mathbf{x_k}|\mathbf{y_{1:k-1}})\mathbf{E}(\mathbf{p_k}|\mathbf{y_{1:k-1}})\\ \mathbf{E}\big(\mathbf{x_k v_k^T p_k}|\mathbf{y_{1:k-1}}\big) &\approx \mathbf{E}(\mathbf{x_k}|\mathbf{y_{1:k-1}})\mathbf{E}\big(\mathbf{v_k^T p_k}|\mathbf{y_{1:k-1}}\big)\end{aligned} \tag{29}$$

where $\mathbf{E}(\mathbf{x_k}|\mathbf{y_{1:k-1}}) = \mathbf{x_{k|k-1}}$.

Assumption 2: In most applications, the $\mathbf{R}$ matrix is diagonal, meaning the measurement noise is independent amongst measurements. Because of the commonality of $\mathbf{R}$ being diagonal, we will restrict our derivation to this case

$$\mathrm{cov}(y_k(d), y_k(l)) = 0 \quad \forall\, d \neq l. \tag{30}$$

Extending the derivation to a model with off-diagonal $\mathbf{R}$ elements is straightforward, but notationally cumbersome.

C. Update Stage, Continued

The above assumptions allow us to estimate $\mathbf{p_k}$ at each iteration and obtain the values of $\mathbf{R_{\tilde{x}\tilde{y}}}$ and $\mathbf{R_{ee}}$ without the knowledge of $\mathbf{x_k}$. When Assumptions 1 and 2 hold

$$\mathbf{E}(\mathbf{p_k}) = \mathbf{Diag}\begin{pmatrix}\Phi\left(\frac{Cx_{k|k-1}(1) - \tau(1)}{\sigma(1)}\right)\\ \Phi\left(\frac{Cx_{k|k-1}(2) - \tau(2)}{\sigma(2)}\right)\\ \vdots\\ \Phi\left(\frac{Cx_{k|k-1}(m) - \tau(m)}{\sigma(m)}\right)\end{pmatrix}. \tag{31}$$

Revisiting $\mathbf{R_{\tilde{x}\tilde{y}_k}}$, and using Assumption 1

$$\begin{aligned}\mathbf{E}\big(\mathbf{x_k x_k^T}\big) &= \mathbf{E}\big((\mathbf{x_k} - \mathbf{E}(\mathbf{x_{k|k-1}}))(\mathbf{x_k} - \mathbf{E}(\mathbf{x_{k|k-1}}))^{\mathbf{T}}\big)\\ &\quad + \mathbf{E}(\mathbf{x_k})\mathbf{E}(\mathbf{x_k})^{\mathbf{T}}\\ &= \Psi_{\mathbf{k|k-1}} + \mathbf{x_{k|k-1}x_{k|k-1}^T}.\end{aligned} \tag{32}$$

The value of $\mathbf{R_{\tilde{x}\tilde{y}_k}}$ is

$$\begin{aligned}\mathbf{R_{\tilde{x}\tilde{y}_k}} &= \big(\Psi_{\mathbf{k|k-1}} + \mathbf{x_{k|k-1}x_{k|k-1}^T}\big)\mathbf{C^T E}(\mathbf{p_k})\\ &\quad + \mathbf{x_{k|k-1}}\mathcal{T}^{\mathbf{T}}(\mathbf{I_{m\times m}} - \mathbf{E}(\mathbf{p_k})) - \mathbf{x_{k|k-1}E}(\mathbf{y_k})^{\mathbf{T}}\\ &\quad - \mathbf{x_{k|k-1}x_{k|k-1}^T C^T E}(\mathbf{p_k})\\ &\quad + \mathbf{x_{k|k-1}}\mathcal{T}^{\mathbf{T}}\big(\mathbf{I_{m\times m}} - \mathbf{E}(\mathbf{p_k})\big) + \mathbf{x_{k|k-1}E}(\mathbf{y_k})^{\mathbf{T}}\big)\\ &= \Psi_{\mathbf{k|k-1}}\mathbf{C^T E}(\mathbf{p_k}).\end{aligned} \tag{33}$$

Repeat the above steps, using the assumptions along with the definition of $E(y_k)$ to compute $\mathbf{R_{\tilde{y}\tilde{y}}}$

$$\begin{aligned}\mathbf{R_{\tilde{y}\tilde{y}_k}} &= \mathbf{E}(\mathbf{p_k})\mathbf{C}\Psi_{\mathbf{k|k-1}}\mathbf{C^T E}(\mathbf{p_k})\\ &\quad + \mathbf{E}\big(\mathbf{p_k}(\mathbf{v_k} - \mathbf{\tilde{v}_k})(\mathbf{v_k} - \mathbf{\tilde{v}_k})^{\mathbf{T}}\mathbf{p_k}\big)\end{aligned} \tag{34}$$

where $\mathbf{\tilde{v}_k} = [\sigma(1)\lambda((\tau(1) - Cx_{k|k-1})/\sigma(1)), \sigma(2)\lambda((\tau(2) - Cx_{k|k-1})/\sigma(1)), \ldots]^T$ is a constant equal to the bias censoring induced on the measurement noise and $\mathbf{E}(\mathbf{p_k}(\mathbf{v_k} - \mathbf{\tilde{v}_k})(\mathbf{v_k} - \mathbf{\tilde{v}_k})^{\mathbf{T}}\mathbf{p_k})^{\mathbf{T}}$ is related to the scalar equation (11). If Assumption 2 holds, $\mathbf{E}(\mathbf{p_k}(\mathbf{v_k} - \mathbf{\tilde{v}_k})(\mathbf{v_k} - \mathbf{\tilde{v}_k})^{\mathbf{T}}\mathbf{p_k})$ is a diagonal matrix written as

$$\begin{aligned}&\mathbf{E}\big(\mathbf{p_k}(\mathbf{v_k} - \mathbf{\tilde{v}_k})(\mathbf{v_k} - \mathbf{\tilde{v}_k})^{\mathbf{T}}\mathbf{p_k}\big)^{\mathbf{T}}\\ &\quad = \mathbf{Diag}\begin{pmatrix}\mathrm{Var}[y_k(1)|x_{k|k-1}(1), \sigma(1)]\\ \mathrm{Var}[y_k(2)|x_{k|k-1}(2), \sigma(2)]\\ \vdots\\ \mathrm{Var}[y_k(m)|x_{k|k-1}(m), \sigma(m)]\end{pmatrix}\end{aligned} \tag{35}$$

where $\mathrm{Var}[y_k(i)|x_{k|k-1}(i), \sigma(i)]$ is calculated according to (11). Substituting this optimal Kalman gain into (22) yields the simplified covariance update equations

$$\Psi_{\mathbf{k|k}} = (\mathbf{I_{m\times m}} - \mathbf{E}(\mathbf{p_k})\mathbf{K_k C})\Psi_{\mathbf{k|k-1}}. \tag{36}$$

The complete Tobit Kalman filter is

$$
\begin{aligned}
\mathbf{x_{k|k-1}} &= \mathbf{A}\mathbf{x_{k-1|k-1}}\\
\Psi_{\mathbf{k|k-1}} &= \mathbf{A}\Psi_{\mathbf{k-1|k-1}}\mathbf{A^T} + \mathbf{Q}\\
\mathbf{x_{k|k}} &= \mathbf{x_{k|k-1}} + \mathbf{R}_{\tilde{\mathbf{x}}\tilde{\mathbf{y}}_\mathbf{k}}\mathbf{R}^{-1}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}_\mathbf{k}}(\mathbf{y_k} - \mathbf{E}(\mathbf{y_k}))\\
\Psi_{\mathbf{k|k}} &= \left(\mathbf{I_{m\times m}} - \mathbf{E}(\mathbf{p_k})\mathbf{R}_{\tilde{\mathbf{x}}\tilde{\mathbf{y}}_\mathbf{k}}\mathbf{R}^{-1}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}_\mathbf{k}}\mathbf{C}\right)\Psi_{\mathbf{k|k-1}}
\end{aligned} \tag{37}
$$

where $\mathbf{R}_{\tilde{\mathbf{x}}\tilde{\mathbf{y}}_\mathbf{k}}$ is given by (33), $\mathbf{R}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}_\mathbf{k}}$ is given by (34), $\mathbf{E}(\mathbf{y_k})$ is given by (8), and $\mathbf{E}(\mathbf{p_k})$ is given by (31).

V. Equivalence to the Standard Kalman Filter

The Tobit Kalman filter will converge to the standard Kalman filter when the state value is far away from the censoring region

$$
\lim_{\frac{\mathbf{x_{k|k-1}}-\mathcal{T}}{\sigma}\to\infty}
\begin{cases}
\mathbf{E}(\mathbf{p_k}) = \mathbf{I_{m\times m}}\\
\mathbf{E}(\mathbf{y_k}) = \mathbf{C}\mathbf{x_{k|k-1}}\\
\mathbf{R} = \sigma^2\\
\mathbf{R}_{\tilde{\mathbf{x}}\tilde{\mathbf{y}}} = \mathbf{C}\Psi_{\mathbf{k|k-1}}\\
\mathbf{R}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}} = \mathbf{C}\Psi_{\mathbf{k|k-1}}\mathbf{C^T} + \mathbf{R}\\
\Psi_{\mathbf{k|k}} = (\mathbf{I_{m\times m}} - \mathbf{K_k}\mathbf{C})\Psi_{\mathbf{k|k-1}}
\end{cases} \tag{38}
$$

so this formulation is a generalization of a standard Kalman filter.

We have derived a Kalman filter for censored data making an assumption on the predictability of the amount of censorship. It is important to note that this formulation only requires the extra complexity of computing m normal pdfs and m normal cdfs at each iteration. These extra computations need to be performed only once per iteration, after the predict stage.

VI. Simulations

In this section, the results from two simulations show the value of the Tobit Kalman filter in comparison with the standard Kalman filter, the Kalman filter for intermittent measurements, and the particle filter. The first simulations are estimating the constant value near a censoring limit and show that the Tobit Kalman filter is unbiased. The next simulation shows a sinusoidal motion model that has disturbances as well as additive noise.

The Tobit Kalman filter will be compared with the Kalman filter with intermittent measurements which is outlined in [13]. This filter operates as a Kalman filter until a missing measurement occurs, then it will only predict the current state and covariance matrix and not update the filter, so $\mathbf{x_{k|k}} = \mathbf{x_{k|k-1}}$ and $\Psi_{\mathbf{k|k}} = \Psi_{\mathbf{k|k-1}}$. This filter treats the censored measurements as missing, instead of censored. The results are also compared with the standard Kalman filter (that treats the censored values as regular measurements) and the particle filter. Particle filters are suboptimal, approaching optimality as the number of particles increases. Many types of particle filter have been developed [27]; in this comparison, we will use the sequential importance resampling, with resampling if the effective number of particles is less than 50% of the particles used. We will use a simple resampling strategy, the systematic resampling approach in [28]. It is possible that a more specialized resampling approach, such as those described in [29], could result in higher performance. The conditional probability of a measurement is given in (3), and will be used in the simulations for the weight update equation.

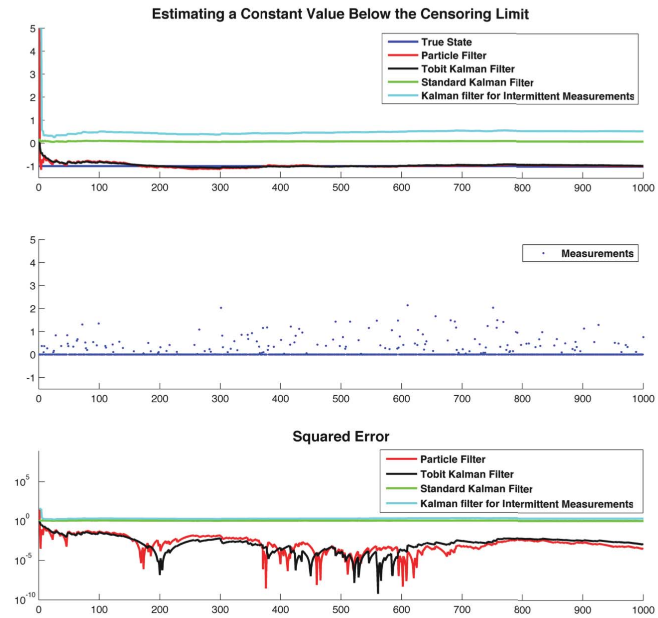

Fig. 1. 1-D example: estimation of a constant value one standard deviation of the measurement noise below the censoring limit.

A. Constant Value Estimation

In this example, we will estimate a constant value near a censoring region. In Fig. 1, we have a constant value at -1 with a noise of $\sigma = 1$ and a censoring limit of $\tau = 0$. The initial conditions are $x_0 = 5$ and $\Psi_0 = 25$, with $Q = 10^{-10}$. The particle filter is represented with 100 particles in this example.

As shown in Fig. 1, the Tobit Kalman filter converges to the true value along with the particle filter. All other methods are biased with estimates produced by the standard Kalman and Kalman filter for intermittent measurements remaining in the uncensored region. This simulation shows that the Tobit Kalman filter is a deterministic and computationally efficient alternative to the particle filter.

B. Oscillator

The following example has dynamics of the form of (13) with the state-space matrices:

$$
A = \alpha\begin{bmatrix}\cos(\omega) & -\sin(\omega)\\ \sin(\omega) & \cos(\omega)\end{bmatrix} \tag{39}
$$

$$
C = [1 \quad 0]. \tag{40}
$$

This example is motivated by the problem of estimating ballistic roll rates from censored magnetometer data, explored in [30]. This simulation shows a robust tracking ability with a known model and unknown disturbance that enters the system through $\mathbf{w_k}$. In this example, $\alpha = 1$, the disturbance is normally distributed with a standard deviation of .05, which is uncorrelated to the measurement noise, which is normally distributed with a variance of $\sigma = 1$. The initial conditions are

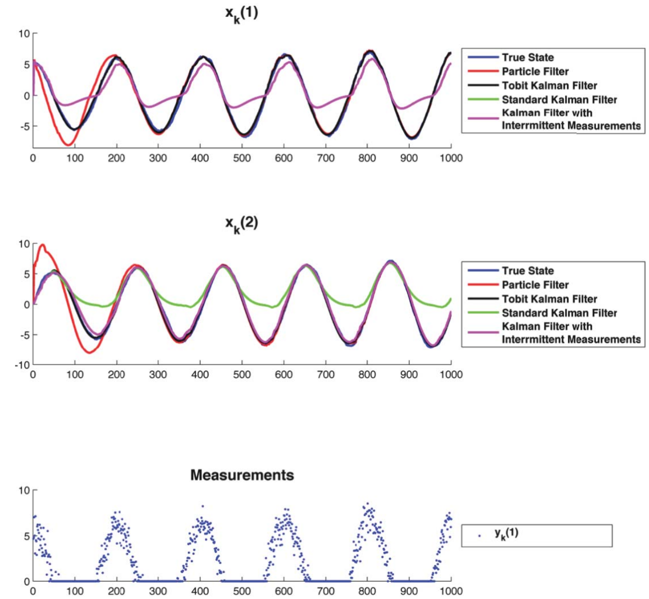

Fig. 2. 2-D example: sinusoidal model with one censored output state.

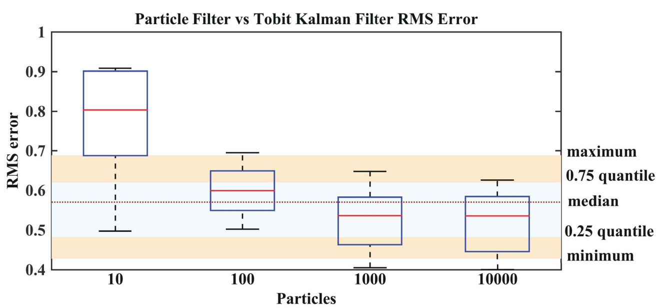

Fig. 3. RMS error and speed comparison of the Tobit Kalman filter and the particle filter. The Tobit Kalman maximum, minimum, and 0.25 and 0.75 quantiles are plotted as color regions, with the dashed line representing the median. The particle filter is denoted by box and whisker plots, representing the same values as the color regions.

$\mathbf{x_0} = [\mathbf{5\ 0}]^{\mathbf{T}}$ and $\Psi_0 = \mathbf{I}_{2\times 2}$ and the frequency is $\omega = .005\ 2\pi$ with a sampling period of $T = 1$. This model has small process noise that makes the particle filter with 100 particles have a long convergence time. Increasing the process noise of the measurements will increase the performance of the particle filter; however, in smooth systems, such as the one in Fig. 2 artificially increasing the process noise of the particle filter will cause a large variance in the estimate.

Fig. 2 shows that when the measurements are being censored, the output of the Tobit Kalman filter closely tracks the actual values while the Kalman filter for the intermittent measurements and standard Kalman filter method are unable to track as they place heavy emphasis on stray noncensored data. In this example, the rms error of the particle filter with the initial 200 points ignored is 0.3792 and the rms error of the Tobit Kalman filter is 0.3480.

The particle filter is the closest to performing the same as the Tobit Kalman filter; however, the Tobit Kalman filter far outperforms the particle filter in speed. In Fig. 3, the sinusoidal example with an average rms error is plotted with 10–10 000 particles. Each simulation is run with 50 sets of simulated data. The Tobit Kalman maximum, minimum, and 0.25 and 0.75 quantiles are plotted as color regions in Fig. 3, with the dashed line representing the median. The particle filter is denoted with box and whisker plots, representing the same values as the color regions. The process noise is higher than the previous example, having a standard deviation of 0.1 and the measurement noise is $\sigma = 1$. The initial condition for the created data and the estimator is $x = [5\ 0]^T$. The computational time for the Tobit Kalman filter was 0.18 s; the computational time for the particle filter increased linearly with the number of particles from 1.2 s for 10 particles to 865 s for 10 000 particles. The Tobit Kalman filter outperforms the particle filter in rms error until a certain number of particles are reached, and in computational speed regardless of the number of particles. This point of intersection of the rms error is heavily dependent on the model. Also, since the particle filter is a stochastic technique, there is a possibility that even with a high number of particles the Tobit Kalman filter will outperform the particle filter.

VII. Conclusion

We present a novel adaptation of the Kalman filter for Tobit Type 1 censored measurements, which we call the Tobit Kalman filter. The resulting formulation provides accurate estimates even when a high proportion of the measurements are censored. In the disturbance-free condition, the estimate converges to the correct value. The behavior of the Tobit Kalman filter is compared with those of three other methods: one in which the censored values are used as true measurements, another which treats the censored data points as missing, and the particle filter. The Tobit Kalman filter consistently outperforms two computationally equivalent filters in situations with moderate censoring. The Tobit Kalman filter often performs better than the particle filter with much less computational requirements.

Censored data arise naturally in a number of engineering applications. Applications for the presented formulation include biochemical measurements with limit-of-detection saturation, inexpensive sensors with saturation censoring, visual tracking with camera frame censoring, and line-of-sight tracking with occlusion.

References

[1] R. E. Kalman, "A new approach to linear filtering and prediction problems," *J. Fluids Eng.*, vol. 82, no. 1, pp. 35–45, 1960.
[2] J. Tobin, "Estimation of relationships for limited dependent variables," *Econometrica*, vol. 26, no. 1, pp. 24–36, 1958.
[3] J. B. Hampshire, II, and J. W. Strohbehn, "Tobit maximum-likelihood estimation for stochastic time series affected by receiver saturation," *IEEE Trans. Inf. Theory*, vol. 38, no. 2, pp. 457–469, Mar. 1992.
[4] J. F. McDonald and R. A. Moffitt, "The uses of Tobit analysis," *Rev. Econ. Statist.*, vol. 62, no. 2, pp. 318–321, 1980.
[5] T. Amemiya, "The estimation of a simultaneous-equation Tobit model," *Int. Econ. Rev.*, vol. 20, no. 1, pp. 169–181, 1979.
[6] R. T. Carson and Y. Sun, "The Tobit model with a non-zero threshold," *Econ. J.*, vol. 10, no. 3, pp. 488–502, Nov. 2007.
[7] T. Amemiya, "Regression analysis when the dependent variable is truncated normal," *Econometrica*, vol. 41, no. 6, pp. 997–1016, Nov. 1973.
[8] R. J. Smith and R. W. Blundell, "An exogeneity test for a simultaneous equation Tobit model with an application to labor supply," *Econometrica*, vol. 54, no. 3, pp. 679–685, May 1986.

[9] W. Schnedler, "Likelihood estimation for censored random vectors," *Econ. Rev.*, vol. 24, no. 2, pp. 195–217, 2005.

[10] F. Giri and E.-W. Bai, Eds., *Block-Oriented Nonlinear System Identification* (Lecture Notes in Control and Information Sciences), vol. 404. London, U.K.: Springer-Verlag, 2010.

[11] T. Wigren, "Convergence analysis of recursive identification algorithms based on the nonlinear Wiener model," *IEEE Trans. Autom. Control*, vol. 39, no. 11, pp. 2191–2206, Nov. 1994.

[12] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S. S. Sastry, "Kalman filtering with intermittent observations," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1453–1464, Sep. 2004.

[13] S. Kar, B. Sinopoli, and J. M. F. Moura, "Kalman filtering with intermittent observations: Weak convergence to a stationary distribution," *IEEE Trans. Autom. Control*, vol. 57, no. 2, pp. 405–420, Feb. 2012.

[14] B. Ibarz-Gabardos and P. J. Zufiria, "A Kalman filter with censored data," in *Proc. IEEE Int. Workshop Intell. Signal Process.*, Sep. 2005, pp. 74–79.

[15] Y. Zheng, R. Niu, and P. K. Varshney, "Sequential Bayesian estimation with censored data," in *Proc. IEEE Statist. Signal Process. Workshop (SSP)*, Aug. 2012, pp. 57–60.

[16] C. Rago, P. Willett, and Y. Bar-Shalom, "Censoring sensors: A low-communication-rate scheme for distributed detection," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 32, no. 2, pp. 554–568, Apr. 1996.

[17] D. Spinello and D. J. Stilwell, "Nonlinear estimation with state-dependent Gaussian observation noise," *IEEE Trans. Autom. Control*, vol. 55, no. 6, pp. 1358–1366, Jun. 2010.

[18] B. Allik, C. Miller, M. J. Piovoso, and R. Zurakowski, "Nonlinear estimators for censored data: A comparison of the EKF, the UKF and the Tobit Kalman filter," in *Proc. Amer. Control Conf.*, 2015.

[19] C. Andrieu and A. Doucet, "Particle filtering for partially observed Gaussian state space models," *J. Roy. Statist. Soc., B (Statist. Methodol.)*, vol. 64, no. 4, pp. 827–836, Oct. 2002.

[20] C. Miller, B. Allik, M. Ilg, and R. Zurakowski, "Kalman filter-based tracking of multiple similar objects from a moving camera platform," in *Proc. 51st IEEE Conf. Decision Control*, Dec. 2012, pp. 5679–5684.

[21] B. M. Adams, H. T. Banks, M. Davidian, and E. S. Rosenberg, "Estimation and prediction with HIV-treatment interruption data," *Bull. Math. Biol.*, vol. 69, no. 2, pp. 563–584, Feb. 2007.

[22] J. David, H. Tran, and H. T. Banks, "Receding horizon control of HIV," *Optim. Control Appl. Methods*, vol. 32, no. 6, pp. 681–699, Nov./Dec. 2010.

[23] Y. Huang, H. Wu, and E. P. Acosta, "Hierarchical Bayesian inference for HIV dynamic differential equation models incorporating multiple treatment factors," *Biometrical J.*, vol. 52, no. 4, pp. 470–486, Aug. 2010.

[24] H. Putter, S. H. Heisterkamp, J. M. A. Lange, and F. de Wolf, "A Bayesian approach to parameter estimation in HIV dynamical models," *Statist. Med.*, vol. 21, no. 15, pp. 2199–2214, Aug. 2002.

[25] R. Luo *et al.*, "Modelling HIV-1 2-LTR dynamics following raltegravir intensification," *J. Roy. Soc. Interf.*, vol. 10, no. 84, p. 20130186, Jul. 2013.

[26] G. Calzolari and G. Fiorentini, "Estimating variances and covariances in a censored regression model," *Statistica*, vol. 53, no. 3, pp. 323–339, 1993.

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

[28] R. Douc and O. Cappe, "Comparison of resampling schemes for particle filtering," in *Proc. 4th Int. Symp. Image Signal Process. Anal.*, Sep. 2005, pp. 64–69.

[29] A. Doucet, S. Godsill, and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering," *Statist. Comput.*, vol. 10, no. 3, pp. 197–208, Jul. 2000.

[30] B. Allik, M. Ilg, and R. Zurakowski, "Ballistic roll estimation using EKF frequency tracking and adaptive noise cancellation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 49, no. 4, pp. 2546–2553, Oct. 2013.