2014 American Control Conference (ACC)
June 4-6, 2014. Portland, Oregon, USA

Estimation of Saturated Data Using the Tobit Kalman Filter

Bethany Allik, Cory Miller, Michael J. Piovoso, and Ryan Zurakowski,

***Abstract*— Saturated, clipped or censored data arises in multiple engineering applications including sensors systems and image based tracking. The saturation limits of a measurement consist of an upper limit and lower limit on the measurements. When a measurement is near a saturated region or in saturated region a standard Kalman filter will be biased and unable to track the true states. In this paper, we use a novel formulation of the Kalman filter for Tobit Type 1 censored measurements. The proposed formulation, called the Tobit Kalman filter for saturated data, converges to the standard Kalman filter in the no-censoring case. A motivating example is presented to demonstrate the usefulness of an estimator for censored data.**

I. Introduction

Censoring frequently occurs in data systems often as a clipped signal, or a saturated sensor . These forms of censoring are known as Tobit model Type 1 censoring [1]. When measurements that are subjected to Tobit model Type 1 censoring are introduced into the standard Kalman filter the resulting estimates will be biased. In this paper, we present a Kalman filter capable of estimating state variables from saturated data without bias. We refer to this formulation as the Tobit Kalman filter for saturated data, where the first formulation of the Tobit Kalman filter for single sided censoring is described in [2].

The Type 1 Tobit model is shown below,

$$y_k = \begin{cases} y_k^*, & T_{low} < y_k^* < T_{high} \\ T_{low}, & y_k^* \leq T_{low} \\ T_{high}, & y_k^* \geq T_{high} \end{cases} \tag{1}$$

$T_{low} \to -\infty$ or $T_{high} \to \infty$ yields either left and right censoring, where y_k^* is the latent variable. Non-recursive estimation for Tobit models has been studied in [3]–[9]; a recursive estimator based on the Kalman filter has been recently developed in [2].

A related phenomenon is dropped or missing measurements. A formulation of the Kalman filter designed for the missing measurement case is presented in [11], [12]. This formulation reduces the Kalman filter to a predictor when the measurement is missing. This formulation relies on the assumption that missed measurements were uncorrelated with the true state value. If censored measurements are treated as dropped measurements the result will be a biased estimate of the state. Likewise, treating the censored values as true measurement in a standard Kalman filter also results in biased estimates.

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. Piovoso is with Pennsylvania State University, Malvern, PA 19355 USA email:piovoso@psu.edu

Manuscript received August 30, 2013; revised .

Tracking the latent variable in Equation 1 has received some attention in [13]. The solution in [13] is not recursive; in that the entire history of the data is needed. Other types of censoring include [14], [15] which studies a different type of censoring involved in distributed detection systems, however they were able to show improved performance by using information that data is censored.

Censoring can also be modeled as an output nonlinearity. These can be addressed using the Extended Kalman filter (EKF) or the particle filter. The sharp discontinuity at the threshold value of the censoring region creates problems for the EKF as the gradient does not exist at this discontinuity. The particle filter formulated for partially observed Gaussian state space models is presented in [16] but is much more computationally expensive than an extended Kalman filter or linear Kalman filter.

In this paper, the theory for the Tobit Kalman filter is extended from the left censored case [2] to the left and right censored case. The applications for this kind of filter include MEMS sensor tracking with saturation, visual tracking with camera frame censoring [17] and biological measurements with limit of detection saturations [18]–[22] . This new formulation provides unbiased recursive estimates of latent state variables from saturated observations. Section II introduces the problem and notation. Section III explains the Tobit model in detail. Section IV derives the Tobit Kalman filter. Section V shows the filter is equivalent to the Kalman filter when the distance between the state estimate and the threshold is large compared to the standard deviation of the measurement noise. Section VI briefly describes the computational complexity of the Tobit Kalman filter. Section VII contains some experimental results including estimation of sinusoidal data with and without disturbance.

II. Problem Formulation

To define the left and right censoring problem we consider the evolution of a scalar output state sequence as,

$$\begin{gathered} x_k = Ax_{k-1} + w_{k-1} \\ y_k^* = Cx_k + v_k \\ y_k = \begin{cases} y_k^*, & T_{low} < y_k^* < T_{high} \\ T_{low}, & y_k^* \leq T_{low} \\ T_{high}, & y_k^* \geq T_{high} \end{cases} \end{gathered} \tag{2}$$

$x_k \in \mathbb{R}^{n\times 1}$ is the state vector and y_k is the scalar measurement . The $A \in \mathbb{R}^{n\times n}$ is the state transition matrix and the $C \in \mathbb{R}^{1\times n}$ is the measurement matrix. The w_k and v_k are Gaussian random vectors with zero mean, they have covariance $Q \in \mathbb{R}^{n\times n}$ and $R = \sigma^2$, respectively. Where σ is the standard deviation of the measurement noise. The

y^* is called the latent variable and is only observable as a measurement when it lies between T_{low}, the lower threshold, and T_{high}, the upper threshold.

It is important to note that the saturation censoring we are presenting in this paper is assuming that the Gaussian noise on the measurements, or the disturbance on the system, will never cause the latent variable to jump from the region $y_k^* > T_{high}$ to $y_k^* < T_{low}$ without passing through the uncensored region. One major application when the occurrence of $y_k^* > T_{high}$ then $y_{k+n}^* < T_{low}$ without a measurement in the uncensored region would occur is in a computer vision tracking application where a target in an image frame exits the field of view on one side and reenters the field of view on the other side of the frame. This is another type of censoring that is not discussed in this paper.

III. Problem formulation in the Tobit Case

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

$$\begin{aligned}&f(y_k|x_k) = \tfrac{1}{\sigma}\phi(\tfrac{y_k - Cx_k}{\sigma})u(y_k - T_{low})u(T_{high} - y_k)\\ &+\delta(T_{high} - y_k)\Phi(\tfrac{Cx_k - T_{high}}{\sigma}) + \delta(T_{low} - y_k)\Phi(\tfrac{T - Cx_k}{\sigma})\end{aligned} \tag{3}$$

where

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

and

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

are the probability density function and the cumulative distribution function of a Gaussian random variable whose mean is Cx_k. δ is the Dirac delta function. The $u(\alpha)$ is a step function and is $u(\alpha) = 1$ when $\alpha \geq 0$ and $u(\alpha) = 0$ when $\alpha < 0$. The delta function at $T_{low} - y_k$ and $T_{high} - y_k$ is present when measurements at the censoring limit are recorded, this function is absent in a truncated model. Figure 1 shows the probability distribution of the saturated measurement.

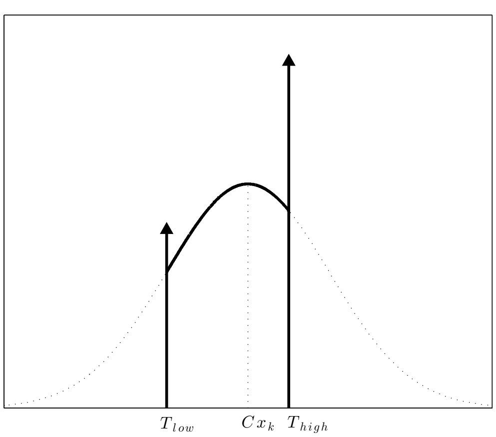

Fig. 1. Probability distribution of a saturated measurement

To calculate the expectation of the measurements we first must find the probability of a measurement being uncensored, p_{uc}, the probability of being censored from above, p_h, and the probability of being censored from below, p_l.

$$\begin{aligned} p_{uc} &= \int_{T_{low}}^{T_{high}} \phi(\tfrac{z - Cx_k}{\sigma})dz \\ &= \Phi(\tfrac{T_{high} - Cx_k}{\sigma}) - \Phi(\tfrac{T_{low} - Cx_k}{\sigma}) \end{aligned} \tag{6}$$

$$p_{low} = \int_{-\infty}^{T_{low}} \phi(\frac{z - Cx_k}{\sigma})dz = \Phi(\frac{T_{low} - Cx_k}{\sigma}) \tag{7}$$

$$p_{high} = \int_{T_{high}}^{\infty} \phi(\frac{z - Cx_k}{\sigma})dz = 1 - \Phi(\frac{T_{high} - Cx_k}{\sigma}) \tag{8}$$

The mean of the measurements when the measurements are not censored is given by:

$$\begin{aligned} &E(y_k|T_{high} > y_k > T_{low}, x_k, \sigma) \\ &= (\sigma p_{uc})^{-1} \int_{T_{low}}^{T_{high}} z\phi(\tfrac{z - Cx_k}{\sigma})dz \\ &= Cx_k - \sigma\lambda(T_{high}, T_{low}) \end{aligned} \tag{9}$$

This p_{uc} is a normalization factor in the uncensored region. This expectation differs from the true value of the latent variable by a bias of $\sigma\lambda(T_{high}, T_{low})$ where,

$$\lambda(T_{high}, T_{low}) = \frac{\phi(\frac{T_{high} - Cx_k}{\sigma}) - \phi(\frac{T_{low} - Cx_k}{\sigma})}{p_{uc}} \tag{10}$$

The expected measured value when censored measurements are included is:,

$$\begin{aligned} &E[y_k|x_{k|k-1}, \sigma] = \\ &P[T_{high} > y_k > T_{low}]E[y_k|T_{high} > y_k > T_{low}] \\ &+P[y_k < T_{low}]E[y_k|y_k < T_{low}] \\ &+P[T_{high} < y_k]E[y_k|T_{high} < y_k] \end{aligned} \tag{11}$$

In the above expectations and probability the $x_{k|k-1}, \sigma$ is dropped for notational purposes, we write $P(\alpha|x_{k|k-1}, \sigma)$ as $P(\alpha)$. The remaining two expectations are $E[y_k|y_k < T_{low}] = T_{low}$ and $E[y_k|T_{high} < y_k]] = T_{high}$

The variance of the expected measured value is derived below:

$$\begin{aligned} Var[y_k|T_{high} > y_k > T_{low}] &= E[y_k^2|T_{high} > y_k > T_{low}] \\ &-[E[y_k|T_{high} > y_k > T_{low}]]^2 \end{aligned} \tag{12}$$

With the first term being,

$$\begin{aligned} E[y_k^2|T_{high} > y_k > T_{low}] &= \sigma^{-1}\frac{1}{p_{uc}}\int_{T_{low}}^{T_{high}} z^2\phi(\tfrac{z - Cx_k}{\sigma})dz \\ &= (Cx_k)^2 + \sigma^2 - \sigma Cx_k\lambda(T_{high}, T_{low}) \\ &+ \frac{\sigma T_{low}\phi(\frac{T_h - Cx_k}{\sigma}) - T_{high}\phi(\frac{T_h - Cx_k}{\sigma})}{p_{uc}} \end{aligned} \tag{13}$$

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

IV. The Tobit Kalman filter

In this section we derive the optimal Kalman formulation for Tobit censored measurements using a linear estimator.

A. The Predict stage

The prior estimate of the state and it's probability distribution may be written

$$\mathbf{P}(\mathbf{x_{k|k-1}}) \sim \mathcal{N}(\mathbf{E}(\mathbf{x_{k|k-1}}), \mathbf{Var}(\mathbf{x_{k|k-1}})) \tag{14}$$

where $\mathbf{x_{k|k-1}} \in \mathbb{R}^{n\times 1}$ is the state estimate vector of $\mathbf{x_k}$ given all estimates and measurements up to time $k-1$. The predict equation of the state may be written as

$$\mathbf{E}(\mathbf{x_{k|k-1}}) = \mathbf{E}(\mathbf{A}\mathbf{x_{k-1|k-1}} + \mathbf{w_k}) = \mathbf{A}\mathbf{x_{k-1|k-1}} \tag{15}$$

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

$$\begin{aligned}\mathbf{cov}(\mathbf{x_{k|k-1}} - \mathbf{x_k}) &= \mathbf{cov}(\mathbf{A}\mathbf{x_{k-1|k-1}} + \mathbf{w_k} - \mathbf{A}\mathbf{x_{k-1}}) \\ &= \mathbf{A}\mathbf{cov}(\mathbf{x_{k-1|k-1}} - \mathbf{x_{k-1}})\mathbf{A^T} + \mathbf{Q} \\ &= \mathbf{A}\mathbf{\Psi_{k-1|k-1}}\mathbf{A^T} + \mathbf{Q}\end{aligned} \tag{16}$$

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

B. The Update Stage

The optimal Kalman filter must minimize the state error covariance, $\mathbf{\Psi_{k|k}}$. The update step shown below will minimize the state error covariance,

$$\mathbf{x_{k|k}} = \mathbf{x_{k|k-1}} + \mathbf{R_{xe_k}}\mathbf{R_{ee_k}^{-1}}(\mathbf{y_k} - \mathbf{E}(\mathbf{y_k})) \tag{17}$$

This is a linear estimator that minimizes the mean squared error and has been proven in [2] and many others. To obtain our estimate of $\mathbf{x_{k|k}}$ we must have the values of $\mathbf{y_k}$, $\mathbf{R_{xe_k}}$ and $\mathbf{R_{ee_k}}$ where $\mathbf{R_{xe_k}}$ is the cross covariance between the Kalman error and the state and $\mathbf{R_{ee_k}}$ is the variance of the Kalman error. The determination of $\mathbf{R_{xe_k}}$ and $\mathbf{R_{ee_k}}$ defined by,

$$\mathbf{R_{xe_k}} = \mathbf{E}((\mathbf{x_k} - \mathbf{x_{k|k-1}})(\mathbf{y_k} - \mathbf{E}(\mathbf{y_k}))^\mathbf{T}) \tag{18}$$

$$\mathbf{R_{ee_k}} = \mathbf{E}((\mathbf{y_k} - \mathbf{E}(\mathbf{y_k}))(\mathbf{y_k} - \mathbf{E}(\mathbf{y_k}))^\mathbf{T}) \tag{19}$$

is described below.

The value of $\mathbf{E}(\mathbf{y_k})$ was calculated for a scalar case for a censored value in Equation 11; in this notation $\mathbf{E}(\mathbf{y_k}) \in \mathbb{R}^{m\times 1}$ is a vector, in which each scalar component can be censored at any given time and may have different threshold limits $\mathbf{T_{high}} = [T_{high}(1), T_{high}(2), ..., T_{high}(m)]$, $\mathbf{T_{low}} = [T_{low}(1), T_{low}(2), ..., T_{low}(m)]$ with $T_{high}(l)$, $T_{low}(l)$, $y_k(l)$ representing the lth component of arrays $\mathbf{T_{high}}$, $\mathbf{T_{low}}$ and $\mathbf{y_k}$ respectively.

To find $\mathbf{K_k} = \mathbf{R_{xe_k}}\mathbf{R_{ee_k}^{-1}}$ in Equation 17 we minimize the state error covariance,

$$\begin{aligned}\mathbf{\Psi_{k|k}} &= \mathbf{cov}(\mathbf{x_k} - \mathbf{x_{k|k}}) \\ &= \mathbf{cov}(\mathbf{x_k} - \mathbf{x_{k|k-1}} - \mathbf{K_k}(\mathbf{y_k} - \mathbf{E}(\mathbf{y_k})))\end{aligned} \tag{20}$$

Three Bernoulli random variables (ζ,ξ, ν) will be introduced to model the occurrence of a censored measurements at T_{high}, T_{low} and a measurement of the latent variable when it is in the uncensored region respectively. The variable $\zeta_k(l) = 1$ when the measurement is censored at T_{high} and $\zeta_k(l) = 0$ when the measurement is not equal to the threshold value. The measurement model for the bernoulli variables are,

$$\zeta_k(l) = \begin{cases} 1, & Cx_k(l) + v_t(l) > T_{high}(l) \\ 0, & Cx_k(l) + v_t(l) \le T_{high}(l) \end{cases} \tag{21}$$

$$\xi_k(l) = \begin{cases} 1, & Cx_k(l) + v_t(l) < T_{low}(l) \\ 0, & Cx_k(l) + v_t(l) \ge T_{low}(l) \end{cases} \tag{22}$$

$$\nu_k(l) = \begin{cases} 1, & T_{low}(l) < Cx_k(l) + v_t(l) < T_{high}(l) \\ 0, & otherwise \end{cases} \tag{23}$$

At any given time step the measurement will represent the state by $Cx_k(l) + v_k(l)$ with probability $E(\nu_k(l))$. In matrix notation the Bernoulli random matrices will be diagonal $\zeta_\mathbf{k} \in \mathbb{R}^{m\times m}$, $\xi_\mathbf{k} \in \mathbb{R}^{m\times m}$,$\nu_\mathbf{k} \in \mathbb{R}^{m\times m}$so the measurements will be arriving by the following equation

$$\mathbf{y_k} = \nu_\mathbf{k}(\mathbf{C}\mathbf{x_k} + \mathbf{v_k}) + \zeta_\mathbf{k}\mathbf{T_{high}} + \xi_\mathbf{k}\mathbf{T_{low}} \tag{24}$$

Substituting into Equation 17 yields

$$\begin{aligned}\mathbf{\Psi_{k|k}} &= \mathbf{cov}(\mathbf{x_k} - \mathbf{x_{k|k}}) \\ &= \mathbf{cov}(\mathbf{x_k} - \mathbf{x_{k|k-1}} - \mathbf{K_k}(\nu_\mathbf{k}(\mathbf{C}\mathbf{x_k} + \mathbf{v_k}) \\ &\quad + \zeta_\mathbf{k}\mathbf{T_{high}} + +\xi_\mathbf{k}\mathbf{T_{low}}))\end{aligned} \tag{25}$$

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

$$\mathbf{G_k} = \nu_\mathbf{k}(\mathbf{C}\mathbf{x_k} + \mathbf{v_k}) + \zeta_\mathbf{k}\mathbf{T_{high}} + \xi_\mathbf{k}\mathbf{T_{low}} - \mathbf{E}(\mathbf{y_k}) \tag{26}$$

so the covariance of the state estimate becomes

$$\begin{aligned}&\mathbf{\Psi_{k|k}} = \mathbf{E}((\mathbf{x_k} - \mathbf{x_{k|k-1}} - \mathbf{K_k}\mathbf{G_k}) \\ &(\mathbf{x_k} - \mathbf{x_{k|k-1}} - \mathbf{K_k}\mathbf{G_k})^\mathbf{T}) \\ &= \mathbf{\Psi_{k|k-1}} - \mathbf{E}((\mathbf{x_k} - \mathbf{x_{k|k-1}})\mathbf{G_k^T})\mathbf{K_k^T} \\ &- \mathbf{K_k}\mathbf{E}(\mathbf{G_k}(\mathbf{x_k} - \mathbf{x_{k|k-1}})^\mathbf{T}) + \mathbf{K_k}\mathbf{E}(\mathbf{G_k}\mathbf{G_k^T})\mathbf{K_k^T}\end{aligned} \tag{27}$$

with

$$\mathbf{\Psi_{k|k-1}} = \mathbf{E}((\mathbf{x_k} - \mathbf{x_{k|k-1}})(\mathbf{x_k} - \mathbf{x_{k|k-1}})^\mathbf{T}) \tag{28}$$

$$\mathbf{R_{xe_k}} = \mathbf{E}((\mathbf{x_k} - \mathbf{x_{k|k-1}})\mathbf{G_k^T}) \tag{29}$$

$$\mathbf{R_{ee_k}} = \mathbf{E}(\mathbf{G_k}\mathbf{G_k^T}) \tag{30}$$

Now we need to find the values for $\mathbf{R_{xe}}$ and $\mathbf{R_{ee}}$. The function for $\mathbf{R_{xe}}$ is,

$$
\begin{aligned}
&\mathbf{R_{xe_k}} = \mathbf{E}((\mathbf{x_k} - \mathbf{x_{k|k-1}})(\nu_\mathbf{k}(\mathbf{Cx_k} + \mathbf{v_k}) \\
&+\zeta_\mathbf{k}\mathbf{T_{high}} + \xi_\mathbf{k}\mathbf{T_{low}} - \mathbf{E}(\mathbf{y_k}))^\mathbf{T}) \\
&= \mathbf{E}(\mathbf{x_k}(\mathbf{Cx_k} + \mathbf{v_k})^\mathbf{T}\nu_\mathbf{k}^\mathbf{T}) + \mathbf{E}(\mathbf{x_k}\mathbf{T_{high}^T}\zeta_\mathbf{k}^\mathbf{T}) \\
&+\mathbf{E}(\mathbf{x_k}\mathbf{T_{low}^T}\xi_\mathbf{k}^\mathbf{T}) - \mathbf{E}(\mathbf{x_k}\mathbf{E}(\mathbf{y_k})^\mathbf{T}) \\
&-\mathbf{E}(\mathbf{x_{k|k-1}}(\mathbf{Cx_k} + \mathbf{v_k})^\mathbf{T}\nu_\mathbf{k}^\mathbf{T}) - \mathbf{E}(\mathbf{x_{k|k-1}}\mathbf{T_{high}^T}\zeta_\mathbf{k}^\mathbf{T}) \\
&-\mathbf{E}(\mathbf{x_{k|k-1}}\mathbf{T_{low}^T}\xi_\mathbf{k}^\mathbf{T}) + \mathbf{E}(\mathbf{x_{k|k-1}}\mathbf{E}(\mathbf{y_k})^\mathbf{T})
\end{aligned} \tag{31}
$$

The probability of the measurement being non censored is a function of the distance between the latent measured variable and the threshold value. The expected value of $\zeta_k(l,l)$, $\xi_k(l,l)$ and $\nu_k(l,l)$ may be written as

$$E(\zeta_k(l,l)) = \Phi(\frac{Cx_k(l) - T_{high}(l)}{\sigma(l)}) \tag{32}$$

$$E(\xi_k(l,l)) = \Phi(\frac{T_{low}(l) - Cx_k(l)}{\sigma(l)}) \tag{33}$$

$$E(\nu_k(l,l)) = \Phi(\frac{T_{high}(l) - Cx_k(l)}{\sigma(l)}) - \Phi(\frac{T_{low}(l) - Cx_k(l)}{\sigma(l)}) \tag{34}$$

Where $Cx_k(l)$ is the l^{th} element of the measurement vector and $\sigma(l)$ is the variance of the noise on that element. The above equation requires knowledge of the true state value, the following assumption allows us to relax this dependence and use the estimated state value instead to obtain values for $\mathbf{R_{xe_k}}$ and $\mathbf{R_{ee_k}}$.

Assumption 1: We assume that the state prediction is a sufficiently accurate estimate of the probability of censoring from above or below:

$$
\begin{aligned}
E(\zeta_k(l,l)) &= \Phi(\tfrac{Cx_k(l)-T_{high}(l)}{\sigma(l)}) \\
&\approx \Phi(\tfrac{Cx_{k|k-1}(l)-T_{high}(l)}{\sigma(l)})
\end{aligned} \tag{35}
$$

$$
\begin{aligned}
E(\xi_k(l,l)) &= \Phi(\tfrac{T_{low}(l)-Cx_k(l)}{\sigma(l)}) \\
&\approx \Phi(\tfrac{T_{low}(l)-Cx_{k|k-1}(l)}{\sigma(l)})
\end{aligned} \tag{36}
$$

$$
\begin{aligned}
&E(\nu_k(l,l)) = \Phi(\tfrac{T_{high}(l)-Cx_k(l)}{\sigma(l)}) - \Phi(\tfrac{T_{low}(l)-Cx_k(l)}{\sigma(l)}) \\
&\approx \Phi(\tfrac{T_{high}(l)-Cx_{k|k-1}(l)}{\sigma(l)}) - \Phi(\tfrac{T_{low}(l)-Cx_{k|k-1}(l)}{\sigma(l)})
\end{aligned} \tag{37}
$$

Assumption 2: For simplicity we assume no cross-dependence in the measurements. Consequently, R is diagonal and:

$$cov(y_k(d), y_k(l)) = 0 \;\forall\; d, l \tag{38}$$

C. The Update Stage, continued

The above assumptions allows us to estimate $\zeta_\mathbf{k}$, $\xi_\mathbf{k}$ and $\nu_\mathbf{k}$ at each iteration and obtain values of $\mathbf{R_{xe}}$ and $\mathbf{R_{ee}}$ without the knowledge of $\mathbf{x_k}$. Where Assumptions 1 and 2 hold,

$$\mathbf{E}(\zeta_\mathbf{k}) = \mathbf{Diag}\begin{pmatrix} \Phi(\frac{Cx_{k|k-1}(1)-T_{high}(1)}{\sigma(1)}) \\ \Phi(\frac{Cx_{k|k-1}(2)-T_{high}(2)}{\sigma(2)}) \\ \vdots \\ \Phi(\frac{Cx_{k|k-1}(m)-T_{high}(m)}{\sigma(m)}) \end{pmatrix}. \tag{39}$$

$$\mathbf{E}(\xi_\mathbf{k}) = \mathbf{Diag}\begin{pmatrix} \Phi(\frac{T_{low}(1)-Cx_{k|k-1}(1)}{\sigma(1)}) \\ \Phi(\frac{T_{low}(2)-Cx_{k|k-1}(2)}{\sigma(2)}) \\ \vdots \\ \Phi(\frac{T_{low}(m)-Cx_{k|k-1}(m)}{\sigma(m)}) \end{pmatrix}. \tag{40}$$

$$\mathbf{E}(\nu_\mathbf{k}) = \mathbf{Diag}\begin{pmatrix} \Phi(\frac{T_{high}(1)-Cx_{k|k-1}(1)}{\sigma(1)}) \\ -\Phi(\frac{T_{low}(1)-Cx_{k|k-1}(1)}{\sigma(1)}) \\ \Phi(\frac{T_{high}(2)-Cx_{k|k-1}(2)}{\sigma(2)}) \\ -\Phi(\frac{T_{low}(2)-Cx_{k|k-1}(2)}{\sigma(2)}) \\ \vdots \\ \Phi(\frac{T_{high}(m)-Cx_{k|k-1}(m)}{\sigma(m)}) \\ -\Phi(\frac{T_{low}(m)-Cx_{k|k-1}(m)}{\sigma(m)}) \end{pmatrix} \tag{41}$$

Revisiting $\mathbf{R_{xe_k}}$, and using $\mathbf{E}(\mathbf{x_{k|k-1}}) = \mathbf{x_{k|k-1}}$, $\mathbf{E}(\mathbf{x_k}) = \mathbf{x_{k|k-1}}$ and

$$
\begin{aligned}
\mathbf{E}(\mathbf{x_k}\mathbf{x_k^T}) &= \mathbf{E}((\mathbf{x_k} - \mathbf{E}(\mathbf{x_{k|k-1}}))(\mathbf{x_k} - \mathbf{E}(\mathbf{x_{k|k-1}}))^\mathbf{T}) \\
&+\mathbf{E}(\mathbf{x_k})\mathbf{E}(\mathbf{x_k})^\mathbf{T} \\
&= \mathbf{\Psi_{k|k-1}} + \mathbf{x_{k|k-1}}\mathbf{x_{k|k-1}^T}
\end{aligned} \tag{42}
$$

$$
\begin{aligned}
&\mathbf{R_{xe_k}} = (\mathbf{\Psi_{k|k-1}} + \mathbf{x_{k|k-1}}\mathbf{x_{k|k-1}^T})\mathbf{C^T}\mathbf{E}(\nu_\mathbf{k}) \\
&+\mathbf{x_{k|k-1}}\mathbf{T_{high}^T}\mathbf{E}(\zeta_\mathbf{k}) + \mathbf{x_{k|k-1}}\mathbf{T_{low}^T}\mathbf{E}(\xi_\mathbf{k}) \\
&-\mathbf{x_{k|k-1}}\mathbf{x_{k|k-1}^T}\mathbf{C^T}\mathbf{E}(\nu_\mathbf{k})^\mathbf{T} - \mathbf{x_{k|k-1}}\mathbf{T_{high}^T}\mathbf{E}(\zeta_\mathbf{k})^\mathbf{T} \\
&-\mathbf{x_{k|k-1}}\mathbf{T_{low}^T}\xi_\mathbf{k}^\mathbf{T} \\
&= \mathbf{\Psi_{k|k-1}}\mathbf{C^T}\mathbf{E}(\nu_\mathbf{k})
\end{aligned} \tag{43}
$$

Repeat the above steps for $\mathbf{R_{xe}}$ to compute $\mathbf{R_{ee}}$

$$\mathbf{R_{ee_k}} = \mathbf{E}(\nu_\mathbf{k})\mathbf{C}\mathbf{\Psi_{k|k-1}}\mathbf{C^T}\mathbf{E}(\nu_\mathbf{k}) + \mathbf{E}(\nu_\mathbf{k}\mathbf{v_k}\mathbf{v_k^T}\nu_\mathbf{k}) \tag{44}$$

where $\mathbf{E}(\nu_\mathbf{k}\mathbf{v_k}\mathbf{v_k^T}\nu_\mathbf{k})^\mathbf{T}$ is related to the scalar Equation 12. With assumption 2, the diagonal matrix is written as:

$$\mathbf{E}(\nu_\mathbf{k}\mathbf{v_k}\mathbf{v_k^T}\nu_\mathbf{k})^\mathbf{T} = \tag{45}$$

$$\mathbf{Diag}\begin{pmatrix} Var[y_k(1)|T_{high}(1) > y_k(1) > T_{low}(1)] \\ Var[y_k(2)|T_{high}(2) > y_k(2) > T_{low}(2)] \\ \vdots \\ Var[y_k(m)|T_{high}(m) > y_k(m) > T_{low}(m)] \end{pmatrix} \tag{46}$$

where $Var[y_k(i)|T_{high}(i) > y_k(i) > T_{low}(i)]$ is calculated according to Equation 12. Substituting this optimal Kalman gain into Equation 27 yields the simplified covariance update equations:

$$\mathbf{\Psi_{k|k}} = (\mathbf{I_{m\times m}} - \mathbf{K_k}\mathbf{E}(\nu_\mathbf{k})\mathbf{C})\mathbf{\Psi_{k|k-1}} \tag{47}$$

The complete Tobit Kalman filter for saturated data is:

$$
\begin{aligned}
\mathbf{x_{k|k-1}} &= \mathbf{A}\mathbf{x_{k-1|k-1}} \\
\mathbf{\Psi_{k|k-1}} &= \mathbf{A}\mathbf{\Psi_{k-1|k-1}}\mathbf{A^T} + \mathbf{Q} \\
\mathbf{x_{k|k}} &= \mathbf{x_{k|k-1}} + \mathbf{R_{xe_k}}\mathbf{R_{ee_k}^{-1}}(\mathbf{y_k} - \mathbf{E}(\mathbf{y_k})) \\
\mathbf{\Psi_{k|k}} &= (\mathbf{I_{m\times m}} - \mathbf{R_{xe_k}}\mathbf{R_{ee_k}^{-1}}\mathbf{E}(\nu_\mathbf{k})\mathbf{C})\mathbf{\Psi_{k|k-1}}
\end{aligned}
\tag{48}
$$

where $\mathbf{R_{xe_k}}$ is given by Equation 43, $\mathbf{R_{ee_k}}$ is given by Equation 44, $\mathbf{E}(\mathbf{y_k})$ is given by Equation 11, and $\mathbf{E}(\nu_\mathbf{k})$ is given by Equation 41.

V. Equivalence to the Standard Kalman Filter

The Tobit Kalman filter will converge to the standard Kalman filter as the threshold values approach $T_{low} \to -\infty$ and $T_{high} \to \infty$,

$$
\lim_{\mathbf{T_{high}}\to\infty \cap \mathbf{T_{low}}\to-\infty}
\begin{cases}
\mathbf{\Phi}(\frac{\mathbf{T_{low}} - \mathbf{C}\mathbf{x_{k|k-1}}}{\sigma}) = [\mathbf{0}\,\mathbf{0}\ldots]^\mathbf{T} \\
\mathbf{\Phi}(\frac{\mathbf{C}\mathbf{x_{k|k-1}} - \mathbf{T_{high}}}{\sigma}) = [\mathbf{0}\,\mathbf{0}\ldots]^\mathbf{T} \\
\mathbf{E}(\mathbf{y_k}) = \mathbf{C}\mathbf{x_{k|k-1}} \\
\mathbf{R} = \sigma^\mathbf{2} \\
\mathbf{R_{xe}} = \mathbf{C}\mathbf{\Psi_{k|k-1}} \\
\mathbf{R_{ee}} = \mathbf{C}\mathbf{\Psi_{k|k-1}}\mathbf{C^T} + \mathbf{R} \\
\mathbf{\Psi_{k|k}} = (\mathbf{I_{m\times m}} - \mathbf{K_k}\mathbf{C})\mathbf{\Psi_{k|k-1}}
\end{cases}
\tag{49}
$$

so Tobit Kalman filter for left and right censoring is a generalization of a standard Kalman filter. For state estimates close to the censoring region the convergence will be directly proportional to the distance to the censoring limit and the measurement noise.

VI. Computation

The computational difference between the Kalman filter and the Tobit Kalman filter is the addition of $2 \times m$ normal PDFs and $2 \times m$ normal CDFs. The extra computations are performed only once per iteration of the estimator, and are only needed for the update stage.

VII. Experiments

In this section we will present a motivating example of tracking a sinusoidal model with a standard Kalman filter where the censored data is used as measurements, the Kalman filter for intermittent measurements and the Tobit Kalman Filter for saturated data. Tracking a sinusoid is important in ballistic problems to calculate roll rate for state estimation [23]. The state space equations we will be using are,

$$
\begin{aligned}
\mathbf{x_k} &= \mathbf{A}\mathbf{x_{k-1}} + \mathbf{B}\mathbf{u_k} \\
y_k^* &= \mathbf{C}\mathbf{x_k} \\
y_k &= \begin{cases} y_k^*, & T_{low} < y_k^* < T_{high} \\ T_{low}, & y_k^* \le T_{low} \\ T_{high}, & y_k^* \ge T_{high} \end{cases}
\end{aligned}
\tag{50}
$$

$$
\mathbf{A} = \alpha \begin{bmatrix} cos(\omega) & -sin(\omega) \\ sin(\omega) & cos(\omega) \end{bmatrix} \tag{51}
$$

$$
\mathbf{B} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{52}
$$

$$
\mathbf{C} = \begin{bmatrix} 1 & 0 \end{bmatrix} \tag{53}
$$

This experiment shows a robust tracking ability with a known model and unknown disturbance that enters the system in $\mathbf{u_k}$. In this example, $\alpha = .999$, the disturbance is normally distributed with variance of 0.05 and is uncorrelated to the measurement noise which is normally distributed with variance of 0.4. The state has $\mathbf{x_0} = [\mathbf{0}\ \mathbf{0}]^\mathbf{T}$ and covariance $\mathbf{P_0} = \mathbf{.05I_{2x2}}$ with threshold limits $T_{low} = -.5$ and $T_{high} = .5$ for all three methods.

In Figure 2 the first plot shows the tracking ability of all three estimators. The standard Kalman filter will converge to the censored value when the measurements are in saturation. The error covariance matrix is not plotted for the the standard Kalman filter but will converge quickly even though uncertainty should increase in saturation. The Kalman filter with intermittent measurements will have erratic behavior because the error covariance matrix is growing exponentially when the measurements are censored. The Tobit Kalman filter shows better tracking ability and an error covariance matrix that reflects the actual uncertainty when near or in a saturated region.

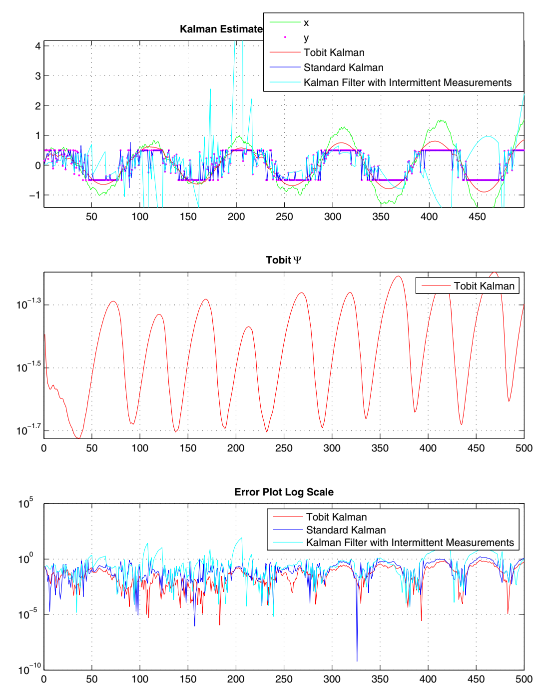

Fig. 2. Sinusoid with disturbance

In Figure 3 all initial conditions are the same except the disturbance is now set to 0. The covariance of the state error in the Tobit Kalman filter will decay to zero, along with the error. The Kalman filter with intermittent measurements still

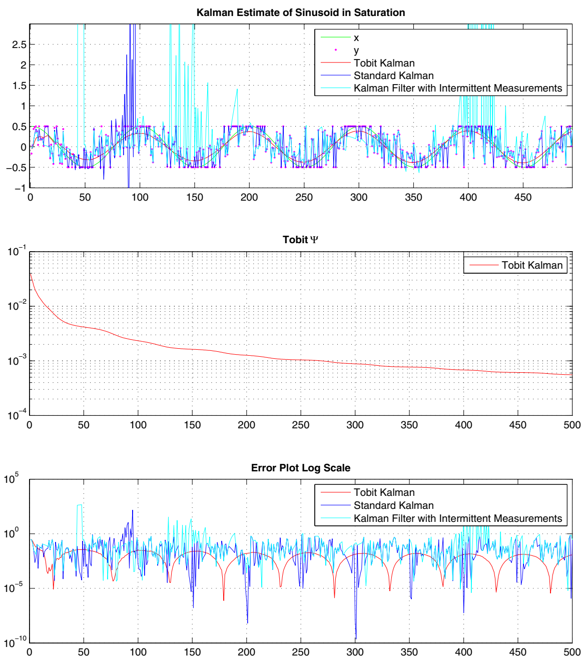

Fig. 3. Sinusoid without disturbance

shows erratic behavior because the state error covariance matrix will grow unbounded when measurments are censored. The standard Kalman filter will falsely converge to censored data when in saturation.

VIII. Conclusion

In this paper we have presented a Tobit Kalman filter for saturated measurements. The estimator provides an unbiased estimate even when a high proportion of the measurements are censored from above or below. The estimator has almost equivalent computational requirements as the standard Kalman filter and converges to the standard Kalman filter when the threshold values approach ∞ and $-\infty$.

The tracking ability of the Tobit Kalman filter in a saturated case is compared to two other methods: one in which the censored values are used as true measurements and another which treats the censored values as missing. The Tobit Kalman filter outperforms the other two estimators and gives a less reactive estimate of the error covariance matrix.

Saturated data arises naturally in a number of engineering applications. Applications for the presented formulation include biochemical measurements with limit-of-detection saturations and inexpensive sensors with saturation censoring.

References

[1] J. Tobin, "Estimation of relationships for limited dependent variables," *Econometrica: Journal of the Econometric Society*, pp. 24–36, 1958.

[2] B. Allik, C. Miller, M. Piovoso, and R. Zurakowski, "The tobit kalman filter: An estimator for censored data," *IEEE Transactions on Control Systems Technology*, p. (submitted), Sept 2013.

[3] J. Hampshire and J. Strohbehn, "Tobit maximum-likelihood estimation for stochastic time series affected by receiver saturation," *Information Theory, IEEE Transactions on*, vol. 38, no. 2, pp. 457–469, 1992.

[4] J. McDonald and R. Moffitt, "The uses of Tobit analysis," *The review of economics and statistics*, vol. 62, no. 2, pp. 318–321, 1980.

[5] T. Amemiya, "The estimation of a simultaneous-equation Tobit model," *International Economic Review*, pp. 169–181, 1979.

[6] R. Carson and Y. Sun, "The Tobit model with a nonzero threshold," *The Econometrics Journal*, vol. 10, no. 3, pp. 488–502, 2007.

[7] T. Amemiya, "Regression analysis when the dependent variable is truncated normal," *Econometrica: Journal of the Econometric Society*, pp. 997–1016, 1973.

[8] R. Smith and R. Blundell, "An exogeneity test for a simultaneous equation Tobit model with an application to labor supply," *Econometrica: Journal of the Econometric Society*, pp. 679–685, 1986.

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

[10] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. Jordan, and S. Sastry, "Kalman filtering with intermittent observations," *Automatic Control, IEEE Transactions on*, vol. 49, no. 9, pp. 1453–1464, 2004.

[11] S. Kar, B. Sinopoli, and J. Moura, "Kalman filtering with intermittent observations: Weak convergence to a stationary distribution," *Automatic Control, IEEE Transactions on*, vol. 57, no. 2, pp. 405–420, 2012.

[12] B. Ibarz-Gabardos and P. J. Zufiria, "A Kalman filter with censored data," *Intelligent Signal Processing, 2005 IEEE International Workshop on*, pp. 74–79, 2005.

[13] Y. Zheng, R. Niu, and P. K. S. S. P. W. S. . I. Varshney, "Sequential Bayesian estimation with censored data," *Statistical Signal Processing Workshop (SSP), 2012 IEEE*, 2012.

[14] C. Rago, P. Willett, and Y. Bar-Shalom, "Censoring sensors: A low-communication-rate scheme for distributed detection," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 32, no. 2, pp. 554–568, 1996.

[15] C. Andrieu and A. Doucet, "Particle filtering for partially observed gaussian state space models," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 64, no. 4, pp. 827–836, 2002.

[16] M. C, B. Allik, M. Ilg, and R. Zurakowski, "Kalman filter-based tracking of multiple similar objects from a moving camera platform," in *Decision and Control (CDC), 2012 IEEE 51st Annual Conference on*, 2012, pp. 5679–5684.

[17] 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.

[18] J. David, H. Tran, and H. T. Banks, "Receding Horizon Control of HIV," *Optim. Control Appl. Meth.*, pp. n/a–n/a, Oct. 2010.

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

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

[21] R. Luo, E. F. Cardozo, M. J. Piovoso, H. Wu, M. J. Buzón, J. Martinez-Picado, and R. Zurakowski, "Modelling HIV-1 2-LTR Dynamics Following Raltegravir Intensification," *Journal of the Royal Society Interface*, vol. 10, no. 84, p. 20130186, Apr. 2013.

[22] B. Allik, M. Ilg, and R. Zurakowski, "Ballistic roll estimation using ekf frequency tracking adaptive noise cancellation," *IEEE Transactions on Aerospace and Electronic Systems*, pp. 2546 – 2553, Oct 2013.