An Extended Set-valued Kalman Filter

D. R. MORRELL
Arizona State University, USA

W. C. STIRLING
Brigham Young University, USA

Abstract

Set-valued estimation offers a way to account for imprecise knowledge of the prior distribution of a Bayesian statistical inference problem. The set-valued Kalman filter, which propagates a set of conditional means corresponding to a convex set of conditional probability distributions of the state of a linear dynamic system, is a general solution for linear Gaussian dynamic systems. In this paper, the set-valued Kalman filter is extended to the non-linear case by approximating the non-linear model with a linear model that is chosen to minimize the error between the non-linear dynamics and observation models and the linear approximation. An application is presented to illustrate and interpret the estimator results.

Keywords

imprecise probabilities, statistical inference, dynamic systems, convex sets of probability measures, set-valued estimation

1 Introduction

In this paper we address the statistical inference problem of estimating a set of time-varying parameters of a discrete-time dynamical system that is monitored with discrete-time observations of its behavior. Such a real-time estimator is called a *filter*. For example, consider an aircraft flight for which radar data are collected as functions of its kinematic parameters (position and velocity). The filtering problem is to obtain instantaneous estimates of its trajectory.

A reasonable model structure for this class of problems is for the system dynamics to be modeled as a finite-dimensional Markov process that is characterized by a stochastic difference equation of the form

$$\mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k) + \mathbf{w}_k \tag{1}$$

for $k = 0, 1, \ldots$, where the p-dimensional vector $\mathbf{x}_k$ is the *state* of the system at time k, and is the time-varying parameter set to be estimated. The p-dimensional vector function $\mathbf{f}$ is the dynamical model for the system, and the p-dimensional vector $\mathbf{w}_k$ is an uncorrelated process termed the *process noise*, with covariance matrix $\mathbf{Q}_k$. The process noise represents random disturbances to the system.

The observation model for this system is of the form

$$\mathbf{y}_k = \mathbf{h}(\mathbf{x}_k) + \mathbf{v}_k \tag{2}$$

for $k = 1, 2, \ldots$, where the q-dimensional vector function $\mathbf{h}$ models the observations as a function of the state. The q-dimensional vector $\mathbf{v}_k$ is an uncorrelated process, termed the *observation noise*, with covariance matrix $\mathbf{R}$. The observation noise represents random measurement errors.

The general filtering problem for this class of systems is to determine the conditional distribution of $\{\mathbf{x}_k, k > 0\}$, given $\{\mathbf{y}_j,\ j \leq k\}$. This problem is easily solved formally: densities are propagated forward via the Chapman-Kolmogorov equation and observations are incorporated using Bayes theorem. However, there are very few system models that lead to closed form solutions. An important special case for which the solution is well known is the linear Gaussian system with precise probability distributions. According to this model, the dynamical and observational equations are linear functions of the state, *i.e.* ,

$$\mathbf{x}_{k+1} = \mathbf{F}_k \mathbf{x}_k + \mathbf{w}_k \tag{3}$$

and

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k, \tag{4}$$

where the processes $\mathbf{w}_k$ and $\mathbf{v}_k$ and the initial state $\mathbf{x}_0$ are all assumed to be jointly normally distributed and mutually uncorrelated. For this special case, the subsequent states $\mathbf{x}_k$, being linear combinations of normally distributed random variables, are also normally distributed, and the problem is solved by directly computing the conditional expectation and covariance of the state. The stationary linear filtering problem (that is, when $\mathbf{F}_k$ and $\mathbf{H}_k$ are constant matrices) was solved by Wiener [14, 4], and the nonstationary case was solved by Kalman [5], Kalman and Bucy [7], and Kalman [6], resulting in the well-known Kalman filter.

Since the normal distribution is not preserved under non-linear transformations, it is not straightforward to compute the conditional mean and variance for non-linear systems. The set-valued estimation problem was addressed for the non-linear case by Kenney and Stirling [8], who provide an approximate solution for the propagation for a set of conditional densities of the state based upon Galerkin approximations to Kolmogorov's equations. Unfortunately, however, this latter approach, although theoretically elegant, is very computationally intensive and has not yet proven to be a practical solution. Practical non-linear estimation techniques include linearization approaches such as the extended Kalman filter [3], Monte Carlo particle filters [2], and interacting multiple-model filtering [1]. Although our approach is essentially Kalman filter based, alternative approaches to set-valued filtering are possible topics of future research. Our preliminary assessment, however, is that extending a particle filter to the imprecise case would be computationally very demanding.

Kalman filter-based approaches (both linear and extended) typically employ a precise prior distribution for the initial state. This is a strict Bayesian approach that is often assumed out of convenience. If this assumption is unwarranted, the precision attributed to the resulting state estimates will not be a realistic indication of their accuracy. Of course, if the system is observable, the influence of the initial conditions will become asymptotically negligible as more and more data are processed. But, for systems with limited data, or if accuracy assessments after a few observations are of interest, then the effect of the initial conditions will be critical.

Imprecise probability theory [13] has emerged as a way to account for ignorance as well as uncertainty in decision making. For the problem here considered, we are concerned with situations where we are unable to specify with confidence the prior distribution of $\mathbf{x}_0$. One way to approach this problem is to characterize the prior as a convex set of distributions, rather than a singleton. This convex Bayesian approach is advocated by Levi [9, 10] as a way of suspending judgment between choices when there is insufficient information to choose a single distribution. Thus, if $p_1(\mathbf{x})$ and $p_2(\mathbf{x})$ are possible prior distributions for $\mathbf{x}_0$, then so is every convex combination $\alpha p_1(\mathbf{x}) + (1-\alpha)p_2(\mathbf{x})$, where $\alpha \in [0,1]$. The filtering problem is then to propagate and update this convex set of distributions. This problem was solved for the linear case by Morrell and Stirling [12], resulting in the *set-valued Kalman filter*.

This paper presents an alternative approach to set-valued non-linear filtering. In Section 2 we review linear set-valued Kalman filtering, which we then extend to deal with non-linear systems in Section 3. Finally, we provide an example in Section 4, and we finish with a discussion in Section 5.

2 Linear Set-Valued Filtering

Consider the system dynamics and observation equations presented in (3) and (4). The set-valued Kalman filter computes a sequence of estimate sets and a corresponding sequence of estimate covariances [12]. An estimate set is denoted $X_{k|j}$, the set of estimates of the system state at time k given the observations $\mathbf{y}_1$ through $\mathbf{y}_j$, and is represented in terms of the p-dimensional vector $\mathbf{c}_{k|j}$ and the $p \times p$ matrix $\mathbf{K}_{k|j}$ as

$$X_{k|j} = \left\{\mathbf{x}\colon \left(\mathbf{x} - \mathbf{c}_{k|j}\right)^T \mathbf{S}_{k|j}^{-1} \left(\mathbf{x} - \mathbf{c}_{k|j}\right) \leq 1\right\}, \tag{5}$$

where $\mathbf{S}_{k|j} = \mathbf{K}_{k|j}\mathbf{K}_{k|j}{}^T$. The set-valued Kalman filtering equations provide a two-stage recursion for computing $\mathbf{c}_{k|j}$, $\mathbf{K}_{k|j}$, and the estimation error covariance $\mathbf{P}_{k|j}$

for $j = k-1$ (prediction between observations) and $j = k$ (updating new observations, or filtering).

Initialization: We assume that the initial state of the dynamic system is characterized by a distribution that lies in the set

$$\mathbf{X} = \left\{\mathbf{x} \sim \mathcal{N}\left(\mathbf{m}, \mathbf{P}_{0|0}\right) \colon \mathbf{m} \in X_{0|0}\right\},$$

where $\mathcal{N}\left(\mathbf{m}, \mathbf{P}_{0|0}\right)$ denotes the normal distribution with mean $\mathbf{m}$ and positive-definite covariance matrix $\mathbf{P}_{0|0}$, and $X_{0|0}$ denotes a hyper-ellipsoid defined by

$$X_{0|0} = \left\{\mathbf{x} \colon \left(\mathbf{x} - \mathbf{c}_{0|0}\right)^T \mathbf{S}_{0|0}^{-1} \left(\mathbf{x} - \mathbf{c}_{0|0}\right) \leq 1\right\}, \tag{6}$$

where $\mathbf{S}_{0|0} = \mathbf{K}_{0|0}\mathbf{K}_{0|0}{}^T$.

Prediction Step:

$$\mathbf{c}_{k|k-1} = \mathbf{F}_{k-1}\mathbf{c}_{k-1|k-1} \tag{7}$$

$$\mathbf{P}_{k|k-1} = \mathbf{F}_{k-1}\mathbf{P}_{k-1|k-1}\mathbf{F}_{k-1}{}^T + \mathbf{Q}_{k-1} \tag{8}$$

$$\mathbf{K}_{k|k-1} = \mathbf{F}_{k-1}\mathbf{K}_{k-1|k-1} \tag{9}$$

The predicted set-valued state estimate is given by

$$X_{k|k-1} = \left\{\mathbf{x} \colon \left(\mathbf{x} - \mathbf{c}_{k|k-1}\right)^T \mathbf{S}_{k|k-1}^{-1} \left(\mathbf{x} - \mathbf{c}_{k|k-1}\right) \leq 1\right\}, \tag{10}$$

with $\mathbf{S}_{k|k-1} = \mathbf{K}_{k|k-1}\mathbf{K}_{k|k-1}{}^T$.

Filter Step:

$$\mathbf{c}_{k|k} = \mathbf{c}_{k|k-1} + \mathbf{W}_k\left[\mathbf{y}_k - \mathbf{H}_k\mathbf{c}_{k|k-1}\right] \tag{11}$$

$$\mathbf{P}_{k|k} = \left[\mathbf{I} - \mathbf{W}_k\mathbf{H}_k\right]\mathbf{P}_{k|k-1} \tag{12}$$

$$\mathbf{K}_{k|k} = \left[\mathbf{I} - \mathbf{W}_k\mathbf{H}_k\right]\mathbf{K}_{k|k-1}, \tag{13}$$

where $\mathbf{W}_k$ is the Kalman gain:

$$\mathbf{W}_k = \mathbf{P}_{k|k-1}\mathbf{H}_k{}^T\left[\mathbf{H}_k\mathbf{P}_{k|k-1}\mathbf{H}_k{}^T + \mathbf{R}\right]^{-1}. \tag{14}$$

The filtered set-valued estimate is then given by

$$X_{k|k} = \left\{\mathbf{x} \colon \left(\mathbf{x} - \mathbf{c}_{k|k}\right)^T \mathbf{S}_{k|k}^{-1} \left(\mathbf{x} - \mathbf{c}_{k|k}\right) \leq 1\right\} \tag{15}$$

It is shown in [12] that, if the linear system defined by (3) and (4) is uniformly observable and controllable (*e.g.* , see [3]), then $\mathbf{K}_{k|k} \to \mathbf{0}$ as $k \to \infty$. Thus, in the limit, the set-valued estimates converge to a point, and the imprecise probability distributions converge to a precise distribution. For systems that are not uniformly observable and controllable, or if the time sequence is not infinite, then imprecision cannot be eliminated. Observability and controllability guarantee only that $\mathbf{P}_{k|k}$ will be bounded [3]. This does *not* mean, however, that the estimation error covariance $\mathbf{P}_{k|k}$ tends to zero as $k \to \infty$.

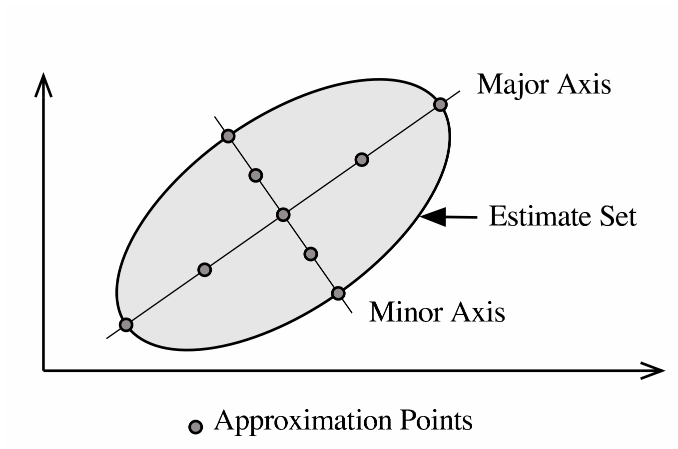

Figure 1: Approximation points for the extended set-valued Kalman filter.

3 Extension to Non-linear System Models

We desire to apply the set-valued Kalman filter to non-linear systems using a linear approximation to the system model. In an extended Kalman filter, such an approximation is made by computing a first-order Taylor series expansion of the non-linear functions about a point-valued state estimate. Unfortunately, because we have a set of state estimates, the set-valued Kalman filter cannot be extended in the same way, and we instead choose approximations that best fit the non-linear functions over the estimate set.

We propose the following approach to finding approximations of the system dynamic and observation functions over the entire estimate set. A set of *approximation points* is chosen; the parameters of affine approximations to the dynamics and observation functions are computed to minimize the weighted squared errors between the function values and approximation values evaluated at the approximation points. Our method of choosing approximation points relies on the hyper-ellipsoidal shape of the estimation sets. Figure 1 illustrates our method for a two-dimensional estimate set (i.e., $p = 2$). Specifically, we form the set of approximation points from the centroid of the estimate set, each point where an axis of the hyper-ellipse intersects the ellipse, and all points equidistant from the centroid and boundary points. Since the estimate set is a p-dimensional hyper-ellipse, the set of approximation points will require $4p+1$ elements. The set-valued Kalman filter requires (approximate) linear dynamics and observation models in which the approximations are good over the entire estimate set.

3.1 Approximating the Dynamics and Observation Functions

We choose approximations of the following form:

$$\mathbf{f}(\mathbf{x}_k) \approx \mathbf{F}_k \mathbf{x}_k + \mathbf{f}_k^0 \tag{16}$$

and

$$\mathbf{h}(\mathbf{x}_k) \approx \mathbf{H}_k \mathbf{x}_k + \mathbf{h}_k^0. \tag{17}$$

The linearizations are obtained by solving the following problems for $\mathbf{F}_k$ and $\mathbf{f}_k^0$ and for $\mathbf{H}_k$ and $\mathbf{h}_k^0$. Let $\mathbf{x}_{k|k}^{(0)}$ through $\mathbf{x}_{k|k}^{(N-1)}$ be values in $X_{k|k}$, denoted the *filtered approximation points*, and let $\mathbf{x}_{k|k-1}^{(0)}$ through $\mathbf{x}_{k|k-1}^{(N-1)}$ be values in $X_{k|k-1}$, denoted the *predicted approximation points*. Let $\mathbf{d}_k^{(n)}$ be the error between the actual dynamics function and the linear approximation evaluated at $\mathbf{x}_{k|k}^{(n)}$:

$$\mathbf{d}_k^{(n)} = \mathbf{f}\left(\mathbf{x}_{k|k}^{(n)}\right) - \mathbf{F}_k \mathbf{x}_{k|k}^{(n)} - \mathbf{f}_k^0.$$

Also, let $\mathbf{e}_k^{(n)}$ be the error between the actual observation function and the linear approximation evaluated at $\mathbf{x}_{k|k-1}^{(n)}$:

$$\mathbf{e}_k^{(n)} = \mathbf{h}\left(\mathbf{x}_{k|k-1}^{(n)}\right) - \mathbf{H}_k \mathbf{x}_{k|k-1}^{(n)} - \mathbf{h}_k^0.$$

We choose $\mathbf{F}_k$, $\mathbf{f}_k^0$ and $\mathbf{H}_k$, $\mathbf{h}_k^0$ to minimize the sums, respectively, of weighted squared dynamics and observation errors evaluated at the approximation points:

$$\mathbf{F}_k, \mathbf{f}_k^0 = \arg\min_{\mathbf{F},\mathbf{f}^0} \sum_{n=0}^{N-1} \ell_k^{(n)} \left[\mathbf{d}_k^{(n)}\right]^T \left[\mathbf{d}_k^{(n)}\right]$$

and

$$\mathbf{H}_k, \mathbf{h}_k^0 = \arg\min_{\mathbf{H},\mathbf{h}^0} \sum_{n=0}^{N-1} \ell_k^{(n)} \left[\mathbf{e}_k^{(n)}\right]^T \left[\mathbf{e}_k^{(n)}\right],$$

where $\ell_k^{(n)}$ is a weight associated with the nth approximation point.

This is a simple weighted least squares problem [11]. We define the following matrices:

$$\mathbf{L}_k = \operatorname{diag}\left(\ell_k^{(0)}, \ldots, \ell_k^{(N-1)}\right)$$

$$\mathbf{A}_k = \begin{bmatrix} \mathbf{x}_{k|k}^{(0)} & \ldots & \mathbf{x}_{k|k}^{(N-1)} \\ 1 & \ldots & 1 \end{bmatrix}, \qquad \mathbf{B}_k = \begin{bmatrix} \mathbf{x}_{k|k-1}^{(0)} & \ldots & \mathbf{x}_{k|k-1}^{(N-1)} \\ 1 & \ldots & 1 \end{bmatrix}$$

$$\mathbf{C}_k = \begin{bmatrix} \mathbf{f}^T\left(\mathbf{x}_{k|k}^{(0)}\right) \\ \vdots \\ \mathbf{f}^T\left(\mathbf{x}_{k|k}^{(N-1)}\right) \end{bmatrix}, \qquad \mathbf{D}_k = \begin{bmatrix} \mathbf{h}^T\left(\mathbf{x}_{k|k-1}^{(0)}\right) \\ \vdots \\ \mathbf{h}^T\left(\mathbf{x}_{k|k-1}^{(N-1)}\right) \end{bmatrix}.$$

The solution to the weighted least squares problem is

$$\begin{bmatrix} \mathbf{F}_k{}^T \\ \mathbf{f}_k^{0T} \end{bmatrix} = \left(\mathbf{A}_k\mathbf{L}_k\mathbf{A}_k{}^T\right)^{-1}\mathbf{A}_k\mathbf{L}_k\mathbf{C}_k$$

and

$$\begin{bmatrix} \mathbf{H}_k{}^T \\ \mathbf{h}_k^{0T} \end{bmatrix} = \left(\mathbf{B}_k\mathbf{L}_k\mathbf{B}_k{}^T\right)^{-1}\mathbf{B}_k\mathbf{L}_k\mathbf{D}_k.$$

An example of choosing approximation points is given in Section 4 in the context of a target tracking problem. Once these quantities are defined, the set-valued Kalman filter is applied to the equations

$$\mathbf{x}_{k+1} = \mathbf{F}_k\mathbf{x}_k + \mathbf{f}_k^0 + \mathbf{w}_k \tag{18}$$

and

$$\mathbf{y}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{h}_k^0 + \mathbf{v}_k. \tag{19}$$

Initialization: The extended set-valued Kalman filter is initialized in the same way the set-valued Kalman filter is initialized.

Prediction Step:

$$\mathbf{c}_{k|k-1} = \mathbf{f}(\mathbf{c}_{k-1|k-1}) \tag{20}$$

$$\mathbf{P}_{k|k-1} = \mathbf{F}_{k-1}\mathbf{P}_{k-1|k-1}\mathbf{F}_{k-1}{}^T + \mathbf{Q}_{k-1} \tag{21}$$

$$\mathbf{K}_{k|k-1} = \mathbf{F}_{k-1}\mathbf{K}_{k-1|k-1} \tag{22}$$

Filter Step:

$$\mathbf{c}_{k|k} = \mathbf{c}_{k|k-1} + \mathbf{W}_k\left[\mathbf{y}_k - \mathbf{h}(\mathbf{c}_{k|k-1})\right] \tag{23}$$

$$\mathbf{P}_{k|k} = [\mathbf{I} - \mathbf{W}_k\mathbf{H}_k]\,\mathbf{P}_{k|k-1} \tag{24}$$

$$\mathbf{K}_{k|k} = [\mathbf{I} - \mathbf{W}_k\mathbf{H}_k]\,\mathbf{K}_{k|k-1}, \tag{25}$$

where $\mathbf{W}_k$ is the Kalman gain as given by (14). The filtered set estimate is then given by (15).

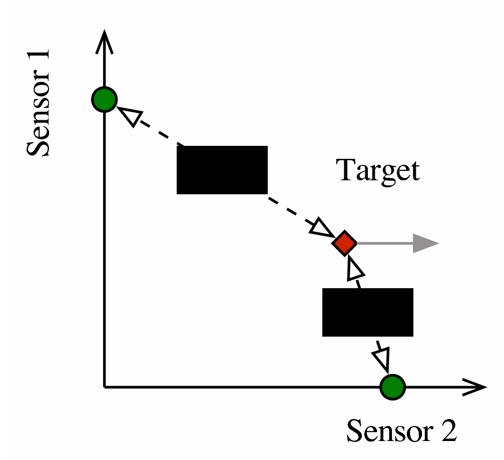

Figure 2: The tracking scenario for application of the set-valued Kalman filter. The target moves in a straight line from left to right. Sensors 1 and 2 measure their range to the target at each time.

4 Example: Target Tracking using Range Measurements

In this section, we present an example of this linearization technique for the set-valued filter. We track a moving target using measured range from one or two fixed sensors; one or both sensors may operate at any point in time. The target moves in a two-dimensional Cartesian coordinate system. Figure 2 illustrates the target motion, sensor locations, and range measurements. The set-valued filter estimates the target position and velocity in both dimensions as a function of time from the range measurements.

We use a linear model of the form (3) for the target dynamics. The target state $\mathbf{x}_k$ consists of four elements: the target position in the x and y directions, denoted $x_k(0)$ and $x_k(1)$, and the target velocity in the x and y directions, denoted $x_k(2)$ and $x_k(3)$. The system dynamics matrix is the following:

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

where Δt is the time between observations. The process noise covariance matrix is

$$\mathbf{Q}_k = \sigma^2 \begin{bmatrix} \frac{\Delta t^3}{3} & 0 & \frac{\Delta t^2}{2} & 0 \\ 0 & \frac{\Delta t^3}{3} & 0 & \frac{\Delta t^2}{2} \\ \frac{\Delta t^2}{2} & 0 & \Delta t & 0 \\ 0 & \frac{\Delta t^2}{2} & 0 & \Delta t \end{bmatrix},$$

where σ^2 is the intensity of a white continuous-time Gaussian noise process modeling the target acceleration.

For this example, we locate Sensor 1 at coordinates $(0,20)$ and Sensor 2 at coordinates $(20,0)$. The ranges from the sensors to the target at time k are denoted as $r_k(1)$ and $r_k(2)$. These ranges are computed as

$$r_k(1) = \sqrt{(x_k(0)-0)^2 + (x_k(1)-20)^2}$$

$$r_k(2) = \sqrt{(x_k(0)-20)^2 + (x_k(1)-0)^2}.$$

Since one or both sensors may be in use at any given k, the observation function $\mathbf{h}(\mathbf{x}_k)$ will be either a one- or two- dimensional vector function of the state $\mathbf{x}_k$. When only Sensor 1 is in use, $\mathbf{h}(\mathbf{x}_k) = r_k(1)$.When both sensors are in use, $\mathbf{h}(\mathbf{x}_k) = \begin{bmatrix} r_k(1) \\ r_k(2) \end{bmatrix}$.

In this problem, the system dynamics are linear; thus, we need to approximate only the observation function. We select the predicted approximation points $\mathbf{x}^{(n)}_{k|k-1}$ as follows. The observations depend only on the target position and not on its velocity, so we select approximation points to cover the range of position values in the estimate set $X_{k|k-1}$. These position values lie in an ellipse defined by the upper left sub-matrix of $\mathbf{S}_{k|k-1}$. We use the centroid of the ellipse, the four points at the intersection of the boundary of the ellipse with its axes, and the four points equidistant between the centroid and boundary points. We use weights of 1.0 for the centroid point, 0.5 for the midpoints, and 0.1 for the boundary points.

In the example scenario, the target starts at the point (10,10) with a velocity of one unit/second to the right. The time Δt between observations is 2 seconds. Only Sensor 1 provides range measurements from time $k=1$ to $k=4$; after $k=4$, both sensors provide range measurements. Figure 3 shows the set estimates of the target position for this scenario. The initial estimate set is circular. The range observations from Sensor 1 quickly reduce the size of the estimate set in the direction of the target from the sensor, but do not provide information about the target location along the perpendicular direction. When range information from Sensor 2 becomes available at time $k=5$, the set of estimates becomes much smaller, since now there is enough information in the observations to locate the target. In other words, during the first 4 time units, when the system is not fully observable, the set of estimates does not shrink in the unobservable direction, but

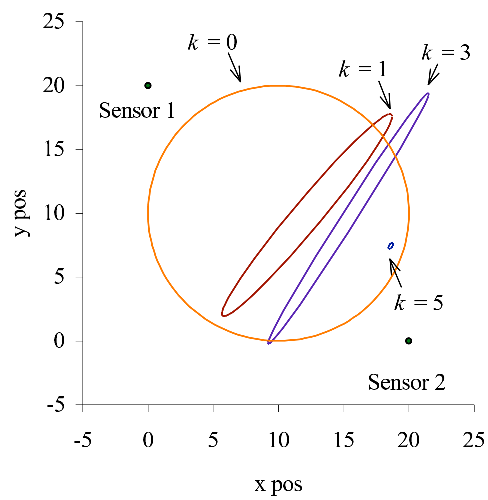

Figure 3: Estimate sets.

thereafter, the system is fully observable and the set of means shrinks in both directions.

It must be emphasized that the ellipses in Figure 3 do *not* correspond to likelihood contours (contours of constant value of probability density); rather, they define a set of position estimates, each of which has a legitimate claim to being a valid assessment of the true state of the system. If time were to increase without bound with both sets of observations available, the system would be observable and, in the limit, it would converge to a singleton representing the mean value of a unique limiting distribution (a precise probability). The covariance of this distribution, however, would converge to a steady-state, but non-zero, level, such that no increase in the accuracy of the (now point-valued) state estimates can be achieved.

5 Discussion

For time-varying estimation scenarios that are either not uniformly observable and controllable or, even if they are observable and controllable, are of such short duration that transients in the estimator dynamics do not have time to damp out, set-valued estimation provides a realistic means of accounting for imprecise knowledge of the mean of the prior distribution.

Non-linear filtering requires the propagation of the entire distribution, in contrast to the need to propagate only the first two moments with linear filtering. This accounts for the difficulty associated with non-linear estimation. The conventional extended Kalman filter is a well-accepted and practical solution for point-valued estimates, but it does not apply to the set-valued case. The extended set-valued Kalman filter provides an approximate solution to the non-linear set-valued dynamic state estimation problem that is computationally feasible. As with the conventional extended Kalman filter, however, it is not possible to prove global convergence of the extended set-valued Kalman filter.

References

[1] Blom, H. A. P., and Bar-Shalom, Y. The interacting multiple model algorithm for systems with markovian switching coefficients. *IEEE Trans. on Automatic Control AC-33*, 8 (1988), 780–783.

[2] Doucet, A. N., de Freitas, N., and Gordon, N. J., Eds. *Sequential Monte Carlo Methods in Practice*. Springer-Verlag, New York, 2001.

[3] Jazwinski, A. H. *Stochastic Processes and Filtering Theory*. Academic Press, New York, 1970.

[4] Kailath, T. *Lectures on Wiener and Kalman Filtering*. Springer-Verlag, New York, 1981.

[5] Kalman, R. E. A new approach to linear filtering and prediction problems. *Trans. ASME, Ser. D: J. Basic Eng 82* (1960), 35–45.

[6] Kalman, R. E. New methods in Wiener filtering theory. In *Proc. Symp. Appl. Random Function Theory and Probability* (New York, 1963), J. L. Bogdanoff and F. Kozin, Eds., Wiley.

[7] Kalman, R. E., and Bucy, R. S. New results in linear filtering and prediction theory. *Trans. ASME, Ser. D: J. Basic Eng 83* (1961), 95–108.

[8] Kenney, J. D., and Stirling, W. C. Nonlinear filtering of convex sets of probability distributions. *J. Stat. Plann. Inference 105* (2002), 123–137.

[9] Levi, I. *The Enterprise of Knowledge*. MIT Press, Cambridge, MA, 1980.

[10] Levi, I. Imprecision and indeterminacy in probability judgement. *Philosophy of Science 52*, 3 (1985), 390–409.

[11] Moon, T. K., and Stirling, W. C. *Mathematical Methods and Algorithms in Signal Processing*. Prentice-Hall, Upper Saddle River, NJ, 2000.

[12] Morrell, D. R., and Stirling, W. C. Set-valued filtering and smoothing. *IEEE Trans. Systems, Man, Cybernet. 21*, 1 (January/February 1991), 184–193.

[13] Walley, P. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London, 1991.

[14] Wiener, N. *The extrapolation, Interpolation and Smoothing of Stationary Time Series*. Wiley, New York, 1949.

Darryl R. Morrell is with the Department of Electrical Engineering, Arizona State University, Tempe, Arizona, USA. E-mail Morrell@asu.edu

Wynn C. Stirling is with the Department of Electrical and Computer Engineering, Brigham Young University, Provo, Utah, USA. E-mail wynn@ee.byu.edu