State Estimation with Point and Set Measurements*

Zhansheng Duan **X. Rong Li** **Vesselin P. Jilkov**
Department of Electrical Engineering
University of New Orleans
New Orleans, LA 70148, U.S.A.
{zduan,xli,vjilkov}@uno.edu

Abstract – *Numerous state estimation problems (e.g., under linear or nonlinear inequality constraints, with quantized measurements) can be formulated as those with point and set measurements. Inspired by the estimation with quantized measurements developed by Curry [1], under a Gaussian assumption, the minimum mean-squared error (MMSE) filtering with point measurements and set measurements of any shape is proposed by discretizing continuous set measurements. Possible ways to relax the Gaussian assumption and to discretize the involved Gaussian and truncated Gaussian distributions are discussed. Through an inequality constrained state estimation example, it is shown that under a certain condition, the update by inequality constraints as set measurements is redundant, otherwise the update is necessary and helpful. Supporting numerical examples are provided.*

Keywords: State estimation, point measurement, set measurement, inequality constraint, quantized measurement, nonlinear filtering.

1 Introduction

In this paper, a *point measurement* is one that is a single point in the measurement space, while a *set measurement* is a subset of the measurement space.

Set measurements are available in numerous cases, for example, state estimation under linear or nonlinear inequality constraints, with quantized measurements. These cases of state estimation are much harder than those with point measurements.

Results about inequality constrained estimation are available [2, 3, 4, 5]. The main idea is to extend the existing results in equality constrained estimation [6] to inequality constrained problem by utilizing classical optimization techniques (e.g., the active set method [2, 3, 4] and the interior point method [5]). If we treat the inequality constraints on the state as another piece of observation information, then we do have set measurement in this case.

State estimation with quantized measurement is an old research topic [1], and it has attracted recent interests [7, 8, 9] with new developments in hybrid estimation, nonlinear filtering and numerical computation. For instance, given the quantization rules, [7] applied the first-order generalized pseudo-Bayesian (GPB1) idea in maneuvering target tracking to improve the performance of Sheppard's method; [9] applied particle filtering; and a numerically efficient implementation of the state estimator with quantized measurement in [1] was proposed in [8]. In general, with the help of a quantization rule, the only thing we can infer from a quantized measurement about the measurement before quantization is that it belongs to some subset of the measurement space.

This paper is an extension of our recent work in [7, 8] to a more general framework—state estimation with point and set measurements. In this paper, inspired by the estimation with quantized measurements developed by Curry [1], under a Gaussian assumption, the MMSE filtering with point measurements and set measurements of any shape is proposed by discretizing continuous set measurements. Possible ways to relax the Gaussian assumption and to discretize the involved Gaussian and truncated Gaussian distributions are also discussed. Supporting numerical examples are provided.

The paper is organized as follows. Sec. 2 formulates the problem. Sec. 3 presents the MMSE filter with point and set measurements. Sec. 4 discusses a possible way to discretize the involved Gaussian and truncated Gaussian distributions. Sec. 5 provides supporting numerical examples. Sec. 6 gives conclusions.

2 Problem formulation

Consider the following generic linear dynamic system

$$x_k = F_{k-1}x_{k-1} + G_{k-1}w_{k-1}$$

with zero-mean white Gaussian noise w_k with $\text{cov}(w_k) = Q_k \geq 0$ and $x_k \in \mathbb{R}^n$, $x_0 \sim \mathcal{N}(\bar{x}_0, P_0)$.

*Research supported in part by NSFC grant 60602026, ARO through grant W911NF-08-1-0409, ONR-DEPSCoR through grant N00014-09-1-1169, NAVO through Contract # N62306-09-P-3S01, and the Louisiana BoR through Grant LEQSF(2009-12)-RD-A-25.

Assuming that two types of measurements of the state are available. The first is the conventional point measurement

$$z_k^{(1)} = H_k^{(1)} x_k + v_k^{(1)} \tag{1}$$

with zero-mean white Gaussian noise $v_k^{(1)}$ having $\mathrm{cov}(v_k^{(1)}) = R_k^{(1)} > 0$ and $z_k^{(1)} \in \mathbb{R}^{m_1}$. $\langle w_k \rangle$, $\langle v_k^{(1)} \rangle$ and x_0 are independent of each other.

The second type is the set measurement:

$$Y_k = \{z_k^{(2)} \in \mathcal{Z}_k\},\ \mathcal{Z}_k \subseteq \mathbb{R}^{m_2}$$

where $z_k^{(2)}$ is a linear or nonlinear function of x_k which may or may not be driven by random noise, $\mathcal{Z}_k$ is a subset of the measurement space $\mathbb{R}^{m_2}$.

Remark: In the extreme case, $\mathcal{Z}_k$ can be the empty set $\emptyset$, singleton or even $\Omega = \mathbb{R}^{m_2}$ (the measurement space). Both $\mathcal{Z}_k = \emptyset$ and $\mathcal{Z}_k = \Omega$ are trivial, since they do not provide any useful information normally and $z_k^{(2)}$ can be ignored. If $\mathcal{Z}_k$ is a singleton, $z_k^{(2)}$ reduces to a point measurement.

In this paper, we estimate the state as best as we can in the MMSE sense based on point measurement $z_k^{(1)}$ and set measurement Y_k, that is,

$$\hat{x}_{k|k}^{\text{MMSE}} = \arg\min_{\hat{x}_{k|k}} E[\tilde{x}_{k|k}\tilde{x}'_{k|k} | \{z_t^{(1)}, Y_t\}_{t=1}^k]$$

where $\tilde{x}_{k|k} = x_k - \hat{x}_{k|k}$.

3 MMSE filtering with point and set measurements

State estimation with point and set measurements is essentially a nonlinear filtering problem. Although nonlinearity destroys Gaussianality in general, to simplify derivation, we will still make Gaussian approximation to the updated state estimate at each time step. A similar idea can be found in the Gaussian filter developed for general nonlinear filtering problems in [10].

3.1 General form of MMSE filter

As discussed above, a quantized measurement is a special type of set measurement. Inspired by the ideas of state estimation with quantized measurements in [1], we extend them to develop an approximate MMSE filtering algorithm with point and set measurements.

Define

$$z_k = \{z_k^{(1)}, Y_k\},\ z^k = \{z_1, z_2, \cdots, z_k\}$$

It is well known that the MMSE filter is

$$\hat{x}_{k|k} = E[x_k | z^k] \tag{2}$$

Using the total expectation theorem, we have

$$\begin{aligned}
\hat{x}_{k|k} &= E[x_k|z^k] = E[E[x_k|z_k^{(2)}, z^k]|z^k] \\
&= \int_{\mathcal{Z}_k} E[x_k|z_k^{(2)}, z^k] p(z_k^{(2)}|z^k) dz_k^{(2)} \\
&= \int_{\mathcal{Z}_k} E[x_k|z_k^{(2)}, z^{k-1}, z_k^{(1)}] p(z_k^{(2)}|z^k) dz_k^{(2)}
\end{aligned} \tag{3}$$

with

$$\begin{aligned}
P_{k|k} &= \text{MSE}(\hat{x}_{k|k}) = E[\tilde{x}_{k|k}\tilde{x}'_{k|k}|z^k] \\
&= E[E[\tilde{x}_{k|k}\tilde{x}'_{k|k}|z_k^{(2)}, z^k]|z^k] \\
&= \int_{\mathcal{Z}_k} E[\tilde{x}_{k|k}\tilde{x}'_{k|k}|z_k^{(2)}, z^k] p(z_k^{(2)}|z^k) dz_k^{(2)} \\
&= \int_{\mathcal{Z}_k} E[\tilde{x}_{k|k}\tilde{x}'_{k|k}|z_k^{(2)}, z^{k-1}, z_k^{(1)}] p(z_k^{(2)}|z^k) dz_k^{(2)}
\end{aligned} \tag{4}$$

That is, $\hat{x}_{k|k}$ and $P_{k|k}$ are the means of $E[x_k|z_k^{(2)}, z^{k-1}, z_k^{(1)}]$ and $E[\tilde{x}_{k|k}\tilde{x}'_{k|k}|z_k^{(2)}, z^{k-1}, z_k^{(1)}]$ (functions of $z_k^{(2)}$ given z^k) with respect to the distribution $p(z_k^{(2)}|z^k)$.

From the definition of z^k, it follows that

$$p(z_k^{(2)}|z^k) = \frac{1}{c_z} p(z_k^{(2)}|z^{k-1}, z_k^{(1)}) I_{\mathcal{Z}_k}(z_k^{(2)})$$

where

$$I_{\mathcal{Z}_k}(z_k^{(2)}) = \begin{cases} 1 & \text{if } z_k^{(2)} \in \mathcal{Z}_k \\ 0 & \text{otherwise} \end{cases}$$

$$c_z = \int_{\mathcal{Z}_k} p(z_k^{(2)}|z^{k-1}, z_k^{(1)}) dz_k^{(2)}$$

That is, $p(z_k^{(2)}|z^k)$ is a truncated version of $p(z_k^{(2)}|z^{k-1}, z_k^{(1)})$.

In general, it is hard to solve (3) and (4) analytically except in some special cases due to several difficulties. First, it is hard to obtain the truncated distribution $p(z_k^{(2)}|z^k)$ exactly. Second, it is also hard to find an analytical form of $E[x_k|z_k^{(2)}, z^{k-1}, z_k^{(1)}]$ and $E[\tilde{x}_{k|k}\tilde{x}'_{k|k}|z_k^{(2)}, z^{k-1}, z_k^{(1)}]$ if $z_k^{(2)}$ is a nonlinear function of x_k. Third, even if we have an analytical form of $p(z_k^{(2)}|z^k)$, $E[x_k|z_k^{(2)}, z^{k-1}, z_k^{(1)}]$ and $E[\tilde{x}_{k|k}\tilde{x}'_{k|k}|z_k^{(2)}, z^{k-1}, z_k^{(1)}]$, to evaluate the involved integrals in a closed form is not easy and only numerical quadrature can be done in practice.

Given

$$\begin{aligned}
\hat{x}_{k-1|k-1} &= E[x_{k-1}|z^{k-1}] \\
P_{k-1|k-1} &= \text{MSE}(\hat{x}_{k-1|k-1}|z^{k-1})
\end{aligned}$$

assume that

$$p(x_{k-1}|z^{k-1}) = \mathcal{N}(x_{k-1}; \hat{x}_{k-1|k-1}, P_{k-1|k-1}) \tag{5}$$

Then the state at time k can be estimated as follows

$$\begin{aligned}
\hat{x}_{k|k-1} &= E[x_k|z^{k-1}] = F_{k-1}\hat{x}_{k-1|k-1} && (6)\\
P_{k|k-1} &= \text{MSE}(\hat{x}_{k|k-1}|z^{k-1})\\
&= F_{k-1}P_{k-1|k-1}F'_{k-1} + G_{k-1}Q_{k-1}G'_{k-1} && (7)\\
\hat{x}^{(1)}_{k|k} &= E[x_k|z^{k-1}, z^{(1)}_k]\\
&= \hat{x}_{k|k-1} + K^{(1)}_k(z^{(1)}_k - H^{(1)}_k\hat{x}_{k|k-1}) && (8)\\
P^{(1)}_{k|k} &= \text{MSE}(\hat{x}^{(1)}_{k|k}|z^{k-1}, z^{(1)}_k)\\
&= P_{k|k-1} - K^{(1)}_k S^{(1)}_k (K^{(1)}_k)' && (9)\\
K^{(1)}_k &= P_{k|k-1}(H^{(1)}_k)'(S^{(1)}_k)^{-1}\\
S^{(1)}_k &= H^{(1)}_k P_{k|k-1}(H^{(1)}_k)' + R^{(1)}_k
\end{aligned}$$

This is just an application of the Kalman filter.

Depending on whether $z^{(2)}_k$ is a linear or nonlinear function of x_k, we discuss next how to solve (3) and (4) approximately in an efficient way.

3.1.1 Linear case

When $z^{(2)}_k$ takes the linear form

$$z^{(2)}_k = H^{(2)}_k x_k + v^{(2)}_k \tag{10}$$

where $v^{(2)}_k$ is zero-mean white Gaussian noise with $\text{cov}(v^{(2)}_k) = R^{(2)}_k \geq 0$ and uncorrelated with $\langle w_k \rangle$, $\langle v^{(1)}_k \rangle$ and x_0, it follows from the Kalman filter that

$$\begin{aligned}
\hat{x}^*_{k|k} &= E[x_k|z^{k-1}, z^{(1)}_k, z^{(2)}_k] = \hat{x}^{(1)}_{k|k} + K_k(z^{(2)}_k - \hat{z}^{(2)}_{k|k})\\
P^*_{k|k} &= \text{MSE}(\hat{x}^*_{k|k}|z^{k-1}, z^{(1)}_k, z^{(2)}_k) = P^{(1)}_{k|k} - K_k S^{(2)}_k K'_k
\end{aligned}$$

where

$$\begin{aligned}
\hat{z}^{(2)}_{k|k} &= E[z^{(2)}_k|z^{k-1}, z^{(1)}_k] = H^{(2)}_k \hat{x}^{(1)}_{k|k}\\
K_k &= P^{(1)}_{k|k}(H^{(2)}_k)'(S^{(2)}_k)^{+}\\
S^{(2)}_k &= \text{MSE}(\hat{z}^{(2)}_{k|k}|z^{k-1}, z^{(1)}_k) = H^{(2)}_k P^{(1)}_{k|k}(H^{(2)}_k)' + R^{(2)}_k
\end{aligned}$$

Note that $\hat{x}^{(1)}_{k|k}$, K_k and $H^{(2)}_k$ have nothing to do with $z^{(2)}_k$. Substituting $\hat{x}^*_{k|k}$ into Eq. (3) yields

$$\hat{x}_{k|k} = \hat{x}^{(1)}_{k|k} + K_k(E[z^{(2)}_k|z^k] - \hat{z}^{(2)}_{k|k}) \tag{11}$$

Since

$$\begin{aligned}
\tilde{x}_{k|k} &= x_k - \hat{x}_{k|k} = x_k - \hat{x}^*_{k|k} + \hat{x}^*_{k|k} - \hat{x}_{k|k}\\
&= x_k - \hat{x}^*_{k|k} + K_k(z^{(2)}_k - E[z^{(2)}_k|z^k])
\end{aligned}$$

we have

$$\begin{aligned}
&E[\tilde{x}_{k|k}\tilde{x}'_{k|k}|z^{(2)}_k, z^{k-1}, z^{(1)}_k]\\
&= P^*_{k|k} + K_k(z^{(2)}_k - E[z^{(2)}_k|z^k])(z^{(2)}_k - E[z^{(2)}_k|z^k])'K'_k
\end{aligned}$$

Note that $P^*_{k|k}$ and K_k have nothing to do with $z^{(2)}_k$. Substituting this into Eq. (4) yields

$$P_{k|k} = P^*_{k|k} + K_k\text{cov}(z^{(2)}_k|z^k)K'_k \tag{12}$$

These are the formulas for state estimation with quantized measurements in [1, 8], except that now we have both point measurement $z^{(1)}_k$ and set measurement Y_k.

It follows from the Gaussian assumption (5) that $p(z^{(2)}_k|z^k)$ is a truncated Gaussian distribution

$$p(z^{(2)}_k|z^k) = \frac{1}{c_z}\mathcal{N}(z^{(2)}_k; \hat{z}^{(2)}_{k|k}, S^{(2)}_k) I_{\mathcal{Z}_k}(z^{(2)}_k)$$

where $c_z = \int_{\mathcal{Z}_k} \mathcal{N}(z^{(2)}_k; \hat{z}^{(2)}_{k|k}, S^{(2)}_k) dz^{(2)}_k$.

Clearly, the key to this MMSE filtering is to compute the mean and covariance matrix of the truncated Gaussian distribution $p(z^{(2)}_k|z^k)$. For the scalar case (i.e., $m_2 = 1$), they can be obtained analytically [8] in terms of the error function $\Phi(\cdot)$ and the density function of the standard Gaussian distribution. For the multi-dimensional case (i.e., $m_2 > 1$), their computation usually relies on multi-dimensional integration over the truncation region without a closed form and only numerical quadrature can be done in practice.

There are several difficulties associated with the numerical quadratures for $E[z^{(2)}_k|z^k]$ and $\text{cov}(z^{(2)}_k|z^k)$. For example, the Gaussian distribution $\mathcal{N}(z^{(2)}_k; \hat{z}^{(2)}_{k|k}, S^{(2)}_k)$ may be singular[1] in some cases due to, e.g., $|R^{(2)}_k| = 0$. Such a density function is not defined, let alone the quadratures based on it. Even if $\mathcal{N}(z^{(2)}_k; \hat{z}^{(2)}_{k|k}, S^{(2)}_k)$ is nonsingular, the following two aspects may still complicate the quadratures in general. First, the curse of dimensionality may occur. Second, the numerical quadrature may be over a truncation region not of a regular shape. Recall that in the problem formulation, the truncation region is only required to be a subset of the m_2-dimensional measurement space, so it can be of any shape depending on the problem.

We give two examples to illustrate how irregular the shapes of the truncation region can be for the case $m_2 = 3$. First, the domain of integration can be unbounded in a three-dimensional space like

$$-\infty < z^{(2)}_{k,1} \leq b_1,\ a_2 \leq z^{(2)}_{k,2} < b_2,\ a_3 < z^{(2)}_{k,3} < +\infty$$

Second, the domain of integration may be an ellipsoid [11] like

$$\sum\nolimits_{i=1}^{3}((z^{(2)}_{k,i} - a_i)/b_i)^2 = 1$$

where $z^{(2)}_{k,i}$ is the i-th component of the vector $z^{(2)}_k$, and a_i and b_i are constants.

If the quadratures used to evaluate $E[z^{(2)}_k|z^k]$ and $\text{cov}(z^{(2)}_k|z^k)$ are well defined, we may use a mature numerical quadrature rule like the Newton–Cotes rule, which

[1]The Gaussian distribution $\mathcal{N}(z^{(2)}_k; \hat{z}^{(2)}_{k|k}, S^{(2)}_k)$ with $|S^{(2)}_k| = 0$ is called a singular Gaussian.

includes the rectangle rule, trapezoidal rule and Simpson's rule as special cases, and the Clenshaw–Curtis rule, but they are mainly for integrations over a one-dimensional bounded region and can not be easily extended to other cases. Efficient numerical quadrature rules like the Gaussian quadrature can indeed handle integration over unbounded domain, but they are still mainly for the one-dimensional case. There do exist exceptions, for instance, the Gaussian-Hermite quadrature can be applied in the high dimensional case because the domain of integration is from $-\infty$ to $+\infty$, which is kept the same after changing variables through decoupling to the original ones. The other Gaussian quadratures are not so lucky because after decoupling of the original variables, the domains of the new variables may be strongly coupled, which will complicate the quadrature greatly and the existing Gaussian quadrature rules can not be applied. To handle multi-dimensional integration, Monte Carlo and sparse grid methods are usually suggested. The sparse grid method is based on a one-dimensional quadrature rule, but performs a more sophisticated combination of univariate results. It is mainly for integrations with a hyper-rectangular domain. Monte Carlo integration may yield better accuracy for the same number of function evaluations than repeated integrations using one-dimensional methods, but the computational complexity is usually too high to be affordable online.

Next, a new way to obtain $E[z_k^{(2)}|z^k]$ and $\text{cov}(z_k^{(2)}|z^k)$ without resorting to any conventional quadrature rules is discussed, which is more general in that it can be applied to set measurements of any shape or dimension regardless if the truncated Gaussian distribution involved is singular or not.

Suppose that the discrete random vector $z_{k,td}^{(2)}$ approximates a continuous one with the conditional distribution $p(z_k^{(2)}|z^{k-1}, z_k^{(1)}, Y_k)$:

$$P\{z_{k,td}^{(2)} = z_k^{(2,i)}\} = \mu_k^{(i)},\ i = 1, 2, \cdots, M$$

where

$$\mu_k^{(i)} \geq 0,\ \sum\nolimits_{i=1}^{M} \mu_k^{(i)} = 1$$

Then it follows that

$$\hat{z}_{k|k} = E[z_k^{(2)}|z^{k-1}, z_k^{(1)}, Y_k] \approx E[z_{k,td}^{(2)}] = \sum\nolimits_{i=1}^{M} \mu_k^{(i)} z_k^{(2,i)}$$

$$\begin{aligned}
&\text{cov}(z_k^{(2)}|z^{k-1}, z_k^{(1)}, Y_k) \\
&\approx \text{cov}(z_{k,td}^{(2)}) = \sum\nolimits_{i=1}^{M} \mu_k^{(i)} (z_k^{(2,i)} - \hat{z}_{k|k})(z_k^{(2,i)} - \hat{z}_{k|k})' \\
&= \sum\nolimits_{i=1}^{M} \mu_k^{(i)} z_k^{(2,i)} (z_k^{(2,i)})' - \hat{z}_{k|k} \hat{z}_{k|k}'
\end{aligned}$$

How to discretize a truncated Gaussian distribution will be discussed in detail later.

3.1.2 Nonlinear case

If $z_k^{(2)}$ takes a nonlinear form $z_k^{(2)} = h_k(x_k, v_k^{(2)})$, even under the Gaussian assumption (5), we do not have $p(z_k^{(2)}|z^{k-1}, z_k^{(1)}) = \mathcal{N}(z_k^{(2)}; \hat{z}_{k|k}^{(2)}, S_k^{(2)})$ and it is usually hard to obtain the exact distribution $p(z_k^{(2)}|z^{k-1}, z_k^{(1)})$, let alone its truncated version $p(z_k^{(2)}|z^k)$. Thus the MMSE filter with a nonlinear $z_k^{(2)}$ is much harder to obtain than the one with a linear $z_k^{(2)}$.

Under the Gaussian assumption (5), we have $p(x_k|z^{k-1}, z_k^{(1)}) = \mathcal{N}(x_k; \hat{x}_{k|k}^{(1)}, P_{k|k}^{(1)})$. Suppose that there exists optimal discrete approximations to $p(x_k|z^{k-1}, z_k^{(1)})$ and $p(v_k^{(2)}) = \mathcal{N}(v_k^{(2)}; 0, R_k^{(2)})$:

$$\begin{aligned}
P\{x_{k,d} = x_k^{(i)}\} &= \alpha_k^{(i)},\ i = 1, 2, \cdots, M_x \\
P\{v_{k,d}^{(2)} = v_k^{(2,j)}\} &= \beta_k^{(j)},\ j = 1, 2, \cdots, M_v
\end{aligned}$$

where

$$\alpha_k^{(i)} \geq 0,\ \sum\nolimits_{i=1}^{M_x} \alpha_k^{(i)} = 1,\ \beta_k^{(j)} \geq 0,\ \sum\nolimits_{j=1}^{M_v} \beta_k^{(j)} = 1$$

Then a close discrete approximation to $p(z_k^{(2)}|z^{k-1}, z_k^{(1)})$ is

$$P\{z_{k,d}^{(2)} = h_k(x_k^{(i)}, v_k^{(2,j)})\} = \gamma_k^{(i,j)} = \alpha_k^{(i)} \beta_k^{(j)}$$

where $i = 1, 2, \cdots, M_x$, $j = 1, 2, \cdots, M_v$.

Selecting only those $h_k(x_k^{(i)}, v_k^{(2,j)}) \in \mathcal{Z}_k$ and normalizing their corresponding probability masses yield a close discrete approximation to the truncated distribution $p(z_k^{(2)}|z^k)$ having the probability mass function (pmf)

$$P\{z_{k,td}^{(2)} = z_k^{(2,l)}\} = \zeta_k^{(l)},\ l = 1, 2, \cdots, L$$

with

$$\zeta_k^{(l)} \geq 0,\ \sum\nolimits_{l=1}^{L} \zeta_k^{(l)} = 1$$

Then we can approximate $\hat{x}_{k|k}$ and $P_{k|k}$ by

$$\hat{x}_{k|k} \approx \sum\nolimits_{l=1}^{L} \hat{x}_{k|k}^{(l*)} \zeta_k^{(l)} \tag{13}$$

$$\begin{aligned}
&P_{k|k} \\
&\approx \sum\nolimits_{l=1}^{L} E[\tilde{x}_{k|k} \tilde{x}_{k|k}' | z_{k,td}^{(2)} = z_k^{(2,l)}, z^{k-1}, z_k^{(1)}] \zeta_k^{(l)} \\
&= \sum\nolimits_{l=1}^{L} \zeta_k^{(l)} E[\tilde{x}_{k|k}^{(l*)} (\tilde{x}_{k|k}^{(l*)})' | z_{k,td}^{(2)} = z_k^{(2,l)}, z^{k-1}, z_k^{(1)}] \\
&+ \sum\nolimits_{l=1}^{L} \zeta_k^{(l)} (\hat{x}_{k|k}^{(l*)} - \hat{x}_{k|k})(\hat{x}_{k|k}^{(l*)} - \hat{x}_{k|k})' \\
&= \sum\nolimits_{l=1}^{L} [P_{k|k}^{(l*)} + (\hat{x}_{k|k}^{(l*)} - \hat{x}_{k|k})(\hat{x}_{k|k}^{(l*)} - \hat{x}_{k|k})'] \zeta_k^{(l)} \\
&= \sum\nolimits_{l=1}^{L} \zeta_k^{(l)} P_{k|k}^{(l*)} + \sum\nolimits_{l=1}^{L} \zeta_k^{(l)} \hat{x}_{k|k}^{(l*)} (\hat{x}_{k|k}^{(l*)})' - \hat{x}_{k|k} \hat{x}_{k|k}'
\end{aligned} \tag{14}$$

where

$$\begin{aligned}
&\tilde{x}_{k|k} = x_k - \hat{x}_{k|k},\ \tilde{x}_{k|k}^{(l*)} = x_k - \hat{x}_{k|k}^{(l*)} \\
&\hat{x}_{k|k}^{(l*)} = E[x_k | z_{k,td}^{(2)} = z_k^{(2,l)}, z^{k-1}, z_k^{(1)}] \\
&P_{k|k}^{(l*)} = \text{MSE}(\hat{x}_{k|k}^{(l*)})
\end{aligned}$$

Note that $\hat{x}_{k|k}^{(l*)}$ is a standard nonlinear filter. Existing nonlinear filtering methods for point estimation can be applied here. As an illustration, only the general form of linear minimum mean-squared error (LMMSE) estimator[2] is listed here:

$$\hat{x}_{k|k}^{(l*)} = E[x_k|z_{k,td}^{(2)} = z_k^{(2,l)}, z^{k-1}, z_k^{(1)}] = \hat{x}_{k|k}^{(1)} + K_k^{(2,l)} \tilde{z}_{k|k}^{(2,l)}$$
$$P_{k|k}^{(l*)} = P_{k|k}^{(1)} - K_k^{(2,l)} S_k^{(2,l)} (K_k^{(2,l)})'$$

where

$$K_k^{(2,l)} = \text{cov}(\tilde{x}_{k|k}^{(1)}, \tilde{z}_{k|k}^{(2,l)})(S_k^{(2,l)})^+, \ S_k^{(2,l)} = \text{cov}(\tilde{z}_{k|k}^{(2,l)})$$
$$\tilde{z}_k^{(2,l)} = z_k^{(2,l)} - E^*[z_k^{(2)}|z^{k-1}, z_k^{(1)}]$$

Remark: Due to the nonlinearity of $z_k^{(2)}$, we do not have an elegant analytical form for $E^*[z_k^{(2)}|z^{k-1}, z_k^{(1)}]$, $\text{cov}(\tilde{x}_{k|k}^{(1)}, \tilde{z}_{k|k}^{(2,l)})$ and $S_k^{(2,l)}$ in general, but they can be approximated by the extended Kalman filter, unscented filtering [12], DD2 [13], Gaussian Hermite filter [10] or even from the original definition of LMMSE estimation as in [14].

Remark: If $\text{cov}(\tilde{x}_{k|k}^{(1)}, \tilde{z}_{k|k}^{(2,l)})$ and $S_k^{(2,l)}$ do not depend on $z_k^{(2,l)}$, (13) and (14) will be equivalent to (11) and (12). In this case, (11) and (12) are more preferable due to their efficiency in implementation.

3.2 Relaxation of Gaussian assumption

Clearly, the key to the above derivation of the approximate MMSE filter is the Gaussian assumption (5). Due to the involvement of set measurements, this assumption is not valid in general. An idea is to use a Gaussian mixture to approximate $p\left(x_{k-1}|z^{k-1}\right)$ after each cycle of filtering as

$$p\left(x_{k-1}|z^{k-1}\right) = \sum\nolimits_{i=1}^{m} \beta_{k-1}^{(i)} \mathcal{N}(x_{k-1}; \hat{x}_{k-1|k-1}^{(i)}, P_{k-1|k-1}^{(i)})$$

This is supported by the fact that any distribution can be approximated by a Gaussian mixture distribution as closely as desired [10].

For the next filtering cycle, the approximate MMSE filter is applied to each component to obtain the corresponding updated estimate $\mathcal{N}(x_k; \hat{x}_{k|k}^{(i)}, P_{k|k}^{(i)})$ autonomously. The weight $\beta_k^{(i)}$ of each component in the Gaussian mixture can be updated as follows when $z_k^{(2)}$ takes the linear form (10)

$$\begin{aligned}\tilde{\beta}_k^{(i)} &= \beta_{k-1}^{(i)} p(z_k|\hat{x}_{k|k-1}^{(i)}) \\ &= \beta_{k-1}^{(i)} \mathcal{N}(z_k^{(1)}; \hat{z}_{k|k-1}^{(1,i)}, S_k^{(1,i)}) \\ &\cdot P\{z_k^{(2)} \in \mathcal{Z}_k | z_k^{(2)} \sim \mathcal{N}(\hat{z}_{k|k-1}^{(2,i)}, S_k^{(2,i)})\} \\ \beta_k^{(i)} &= \tilde{\beta}_k^{(i)} / \sum\nolimits_{j=1}^{m} \tilde{\beta}_k^{(j)}, \ i = 1, 2, \cdots, m\end{aligned}$$

where

$$\begin{aligned}\hat{x}_{k|k-1}^{(i)} &= F_{k-1} \hat{x}_{k-1|k-1}^{(i)} \\ P_{k|k-1}^{(i)} &= F_{k-1} P_{k-1|k-1}^{(i)} F_{k-1}' + G_{k-1} Q_{k-1} G_{k-1}' \\ \hat{z}_{k|k-1}^{(j,i)} &= H_k^{(j)} \hat{x}_{k|k-1}^{(i)} \\ S_k^{(j,i)} &= H_k^{(j)} P_{k|k-1}^{(i)} (H_k^{(j)})' + R_k^{(j)}, \ j = 1, 2\end{aligned}$$

Remark: If a set measurement is available, the computation of the likelihood $P\{z_k^{(2)} \in \mathcal{Z}_k | z_k^{(2)} \sim \mathcal{N}(\hat{z}_{k|k-1}^{(2,i)}, S_k^{(2,i)})\}$ above is a headache. It is suggested to use our proposed discretization method to get this probability as

$$\begin{aligned}P\{z_k^{(2)} \in \mathcal{Z}_k | z_k^{(2)} \sim \mathcal{N}(\hat{z}_{k|k-1}^{(2,i)}, S_k^{(2,i)})\} \\ = \sum\nolimits_{l=1}^{L} \mu_k^{(i,l)} I_{\mathcal{Z}_k}(y_k^{(i,l)})\end{aligned}$$

where $\{\mu_k^{(i,1)}, \mu_k^{(i,2)}, \cdots, \mu_k^{(i,L)}\}$ is the set of probability masses of a discrete random vector $y_k^{(i)}$ over the space $\{y_k^{(i,1)}, y_k^{(i,2)}, \cdots, y_k^{(i,L)}\}$, which is an approximation of the Gaussian distribution $\mathcal{N}(z_k^{(2)}; \hat{z}_{k|k-1}^{(2,i)}, S_k^{(2,i)})$.

Remark: The other forms of weight update in [10], derived under different optimality criteria, can also be applied here similarly.

4 Discretization of Gaussian-related distribution

In summary, we met two types of discretization problems above. One is for a Gaussian distribution and the other for a truncated Gaussian distribution. Next, we discuss how to find a close discretization for each type.

From [15], it is known that a continuous random variable can be approximated by a discrete random variable as closely as desired, as described by the following lemma.

Lemma 1 [15]. Given a distance metric tolerance ϵ and a scalar continuous random variable x with range $\mathcal{X}$ and cdf $F_x(t)$, the scalar discrete random variable y characterized by range $\mathcal{Y}^*$ (with minimum cardinality $L = \lceil 1/2\epsilon \rceil =$ smallest integer not smaller than $1/2\epsilon$) and pmf μ^*, which is closest to x in the sense of minimum distribution mismatch, is given by

$$\begin{aligned}\mathcal{Y}^* &= \{y^{(1)}, y^{(2)}, \cdots, y^{(L)}\} \\ \mu^* &= \{\mu^{(1)}, \mu^{(2)}, \cdots, \mu^{(L)}\} \\ y^{(i)} &= \arg_{t \in \mathcal{X}} [F_x(t) = (i - 1/2)/L], \ i = 1, 2, \cdots, L \\ \mu^{(i)} &= P\{y = y^{(i)} | y \in \mathcal{Y}^*\} = 1/L\end{aligned}$$

Unfortunately, we do not have such a nice result for the vector case. Now let us discuss how to discretize a general singular or nonsingular multivariate Gaussian distribution.

[2] In this paper, $E^*[x|z]$ denotes the LMMSE estimator of x with respect to z.

Given an m-dimensional Gaussian distribution $p(z) = \mathcal{N}(z; \mu_z, R_z)$, it follows from the singular value decomposition that there must exist a unitary matrix U such that

$$UR_zU' = \begin{bmatrix} R_z^{(\#)} & 0_{r\times(m-r)} \\ 0_{(m-r)\times r} & 0_{(m-r)\times(m-r)} \end{bmatrix}$$

where $R_z^{(\#)}$ is an $r \times r$ diagonal matrix with $R_z^{(\#)} > 0$ and $r = \text{rank}(R_z)$.

Let

$$a = [((R_z^{(\#)})^{1/2}b)', 0_{1\times(m-r)}]' \tag{15}$$

where $b \in \mathbb{R}^r$ and $b \sim \mathcal{N}(0_{r\times 1}, I_{r\times r})$. It can then be easily verified that

$$U'a + \mu_z \sim \mathcal{N}(\mu_z, R_z) \tag{16}$$

The discretization of a Gaussian distribution can then be summarized as follows.

Step 1: Following Lemma 1, obtain the optimal discrete approximation to the standard Gaussian distribution and get L mass points in the one-dimensional space, each of which has a probability mass $1/L$. Construct r copies of these L mass points with respect to r components of b and then obtain the L^r mass points in an r-dimensional space simply through combining points in the r copies, which are a close approximation to the distribution of b, each having probability mass $1/L^r$.

Step 2: Transform all L^r mass points through Eqs. (15) and (16) to get a close discrete approximation of the Gaussian distribution $p(z) = \mathcal{N}(z; \mu_z, R_z)$.

Remark: As L increases, the approximation accuracy becomes better.

Remark: Since our set measurement formulation includes the linear inequality constraint as a special case, it follows that $S_k^{(2)} > 0$ may not always be true. In this case, the untruncated distribution $p(z_k^{(2)}|z^{k-1}, z_k^{(1)})$ may be a singular Gaussian distribution. That is exactly why we have the block forms above.

Remark: In essence, this discretization is a *quasi Monte Carlo* method and it samples deterministically rather than randomly as in the Monte Carlo method.

In order to discretize a truncated Gaussian distribution, we can discretize the untruncated Gaussian distribution first, and then reject all those mass points located outside the truncation region $\mathcal{Z}$ and renormalize the probability masses of all those mass points within the truncation region to sum up to the unity. This will be a close discrete approximation to the truncated Gaussian distribution. Note, however, that this *rejection method* is not efficient, although it works in general. An efficient way was proposed in [16].

5 Illustrative examples

Under a certain condition [17], the update by set constraints as set measurements is redundant. But if there exists model mismatch, the update will improve performance in general. We now verify these findings through numerical examples.

Consider the following dynamic system, which describes the motion of an on-road vehicle [18]:

$$x_k = F_{k-1}x_{k-1} + G_{k-1}u_{k-1} + w_{k-1} \tag{17}$$

where

$$\begin{aligned}
x_k &= [\ \mathsf{x}_k \quad \mathsf{y}_k \quad \dot{\mathsf{x}}_k \quad \dot{\mathsf{y}}_k\ ]',\ w_k \sim \mathcal{N}(0, Q_k) \\
x_0 &\sim \mathcal{N}(\bar{x}_0, P_0),\ \bar{x}_0 = [\ 0 \quad 0 \quad 10\sqrt{3} \quad 10\ ]' \\
F_k &= \begin{bmatrix} 1 & 0 & T & 0 \\ 0 & 1 & 0 & T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},\ G_k = \begin{bmatrix} 0 \\ 0 \\ T\sin\theta \\ T\cos\theta \end{bmatrix} \\
Q_k &= \begin{bmatrix} 30 & 10\sqrt{3} & 0 & 0 \\ 10\sqrt{3} & 10 & 0 & 0 \\ 0 & 0 & 10 & 10\sqrt{3}/3 \\ 0 & 0 & 10\sqrt{3}/3 & 10/3 \end{bmatrix} \\
P_0 &= N_k \text{diag}\{400, 400, 10, 10\} N_k \\
N_k &= I_4 - (H_k^{(2)})'(H_k^{(2)}(H_k^{(2)})')^{-1}H_k^{(2)} \\
H_k^{(2)} &= [\ 0 \quad 0 \quad 1 \quad -\tan\theta\ ],\ \theta = \pi/3,\ T = 2 \\
u_k &= \begin{cases} 1, \text{ if } k \text{ is odd} \\ -1, \text{ if } k \text{ is even} \end{cases}
\end{aligned}$$

The state satisfies the following linear equality constraint [18]

$$H_k^{(2)}x_k = 0 \tag{18}$$

This is because the angle between the y axis and the road (treated as a straight line without width) is θ. For the filter, it only knows the following linear inequality constraint

$$-1 \le H_k^{(2)}x_k \le 1 \tag{19}$$

The measurement is described by

$$z_k^{(1)} = H_k^{(1)}x_k + v_k^{(1)} \tag{20}$$

where

$$\begin{aligned}
v_k^{(1)} &\sim \mathcal{N}(0, R_k^{(1)}),\ R_k^{(1)} = \text{diag}(400, 400, 10) \\
H_k^{(1)} &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}
\end{aligned}$$

We want to estimate the state in the MMSE sense based on the measurement $z_k^{(1)}$ and linear inequality constraint (19). In the following, all estimators were initialized with $P_{0|0} = P_0$. A mismatched Q_k:

$$Q_k^{\text{mis}} = \begin{bmatrix} 30 & 10\sqrt{3} & 0 & 0 \\ 10\sqrt{3} & 10 & 0 & 0 \\ 0 & 0 & 13.25 & 0.1443 \\ 0 & 0 & 0.1443 & 13.0833 \end{bmatrix}$$

is used in place of Q_k in some estimators, possibly along with a mis-specified initial estimate

$$\hat{x}_{0|0}^{\text{mis}} = \bar{x}_0 + [\ 0 \quad 0 \quad 6 \quad 12\ ]'$$

Table 1: Estimators used in inequality constrained estimation example

name	explanation
KF	Q_k, $\hat{x}_{0\|0} = \bar{x}_0$, updated only by $z_k^{(1)}$
IEC	true Q_k, $\hat{x}_{0\|0} = \bar{x}_0$, updated by $z_k^{(1)}$ first and then by (19)
PM-Q	Q_k^{mis}, $\hat{x}_{0\|0} = \bar{x}_0$, updated only by $z_k^{(1)}$
IEC-Q	Q_k^{mis}, $\hat{x}_{0\|0} = \bar{x}_0$, updated by $z_k^{(1)}$ first and then by (19)
PM-0	Q_k, $\hat{x}_{0\|0} = \hat{x}_{0\|0}^{\text{mis}}$, updated only by $z_k^{(1)}$
IEC-0	Q_k, $\hat{x}_{0\|0} = \hat{x}_{0\|0}^{\text{mis}}$, updated by $z_k^{(1)}$ first and then by (19)

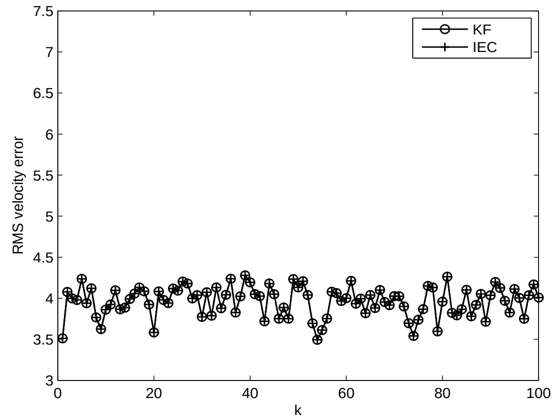

Figure 2: RMS velocity error comparison with correct Q and $\hat{x}_{0|0} = \bar{x}_0$. Note that KF overlaps with IEC.

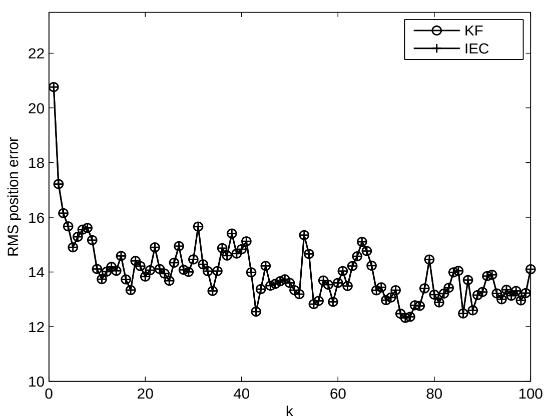

Figure 1: RMS position error comparison with correct Q and $\hat{x}_{0|0} = \bar{x}_0$. Note that KF overlaps with IEC.

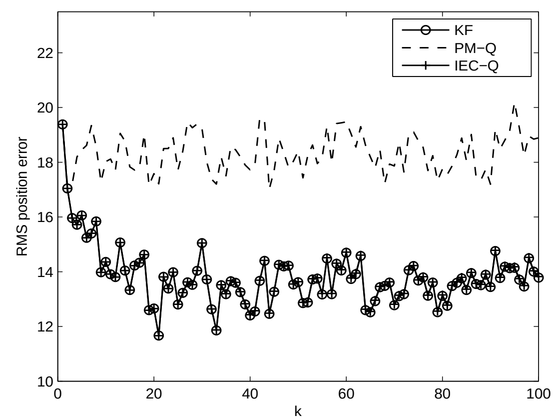

Figure 3: RMS position error comparison with Q mismatch. Note that KF and IEC-Q have no noticeable difference.

Also, all results were averaged over 200 Monte Carlo runs.

Figs. 1 through 6 show comparison results of estimators in Table 1.

Since the noisy point measurement $z_k^{(1)}$ is linear in x_k, under the given knowledge and with correct Q and $\hat{x}_{0|0}$, the KF updated only by noisy measurement is optimal. It turns out that the update by the linear inequality constraint can be skipped. From Figs. 1 through 6, it can be seen that the best performance is indeed achieved by the KF.

From Figs. 3 and 4, it can be seen that if there is Q mismatch and linear inequality constraint is not taken into account in the update step, the PM-Q filter has much worse performance than the KF. But if there is Q mismatch and the linear inequality constraint is fully accounted for in the update step, the performance of the IEC-Q filter is almost the same as that of the KF. From Figs. 5 and 6, it can be seen that if $\hat{x}_{0|0}$ is mis-specified and the linear inequality constraint is not taken into account in the update step, the PM-0 estimator has the worst performance and its position RMSE diverges. But if $\hat{x}_{0|0}$ is mis-specified and the linear inequality constraint is fully taken into account in the update step, except during the initial short transient, IEC-0 and KF have comparable position RMSE and almost identical velocity RMSE. All these verify our statements about the contribution of set constraints in practical constrained estimation problems. We should take the set constraint into account when there exists model mismatch.

6 Conclusions

This paper targets the state estimation problem with point and set measurements. The main difficulty for solving this type of problem comes from the uncertainty associated with set measurements. The key idea proposed in this paper is to approximate a continuous set measurement by discrete point measurements under a Gaussian assumption. Possible ways to relax the Gaussian assumption and to discretize the involved Gaussian and truncated Gaussian distributions are discussed. Numerical examples show that for state estimation subject to inequality constraints, under a certain con-

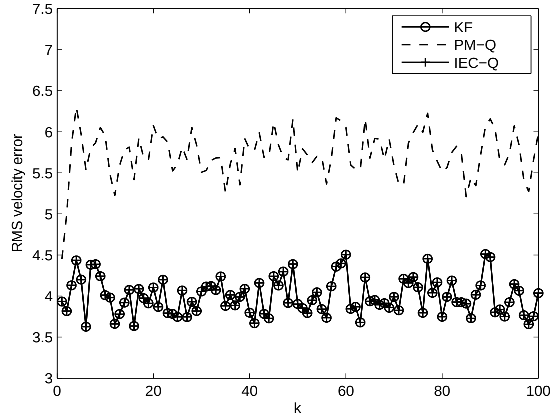

Figure 4: RMS velocity error comparison with Q mismatch. Note that KF and IEC-Q have no noticeable difference.

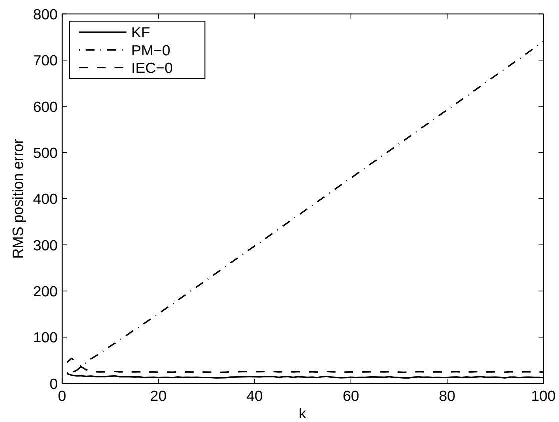

Figure 5: RMS position error comparison with mis-specified $\hat{x}_{0|0}$.

dition, the update by inequality constraints as set measurements is redundant, otherwise the update is necessary and helpful.

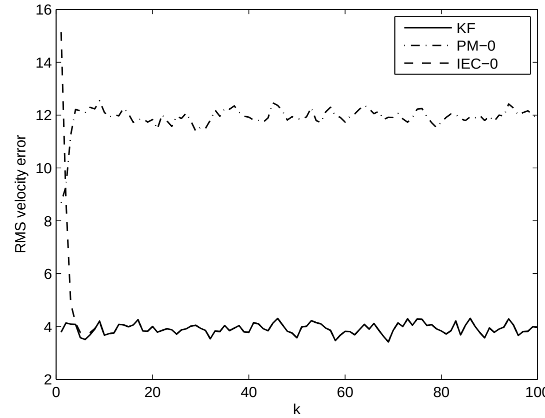

Figure 6: RMS velocity error comparison with mis-specified $\hat{x}_{0|0}$.

References

[1] R. E. Curry, *Estimation and Control with Quantized Measurements*. MIT Press, February 1970.

[2] D. Simon and D. L. Simon, "Aircraft turbofan engine health estimation using constrained kalman filtering," *ASME Journal of Engineering for Gas Turbines and Power*, vol. 127, no. 2, pp. 323–328, April 2005.

[3] D. Simon, "A game theory approach to constrained minimax state estimation," *IEEE Transactions on Signal Processing*, vol. 54, no. 2, pp. 405–412, February 2006.

[4] D. Simon and D. L. Simon, "Kalman filtering with inequality constraints for turbofan engine health estimation," *IEE Proceedings Control Theory and Applications*, vol. 153, no. 3, pp. 371–378, May 2006.

[5] B. M. Bell, J. V. Burke, and G. Pillonetto, "An inequality constrained nonlinear Kalman-Bucy smoother by interior point likelihood maximization," *Automatica*, vol. 45, no. 1, pp. 25–33, January 2009.

[6] D. Simon and T. L. Chia, "Kalman filtering with state equality constraints," *IEEE Transactions on Aerospace and Elctronic Systems*, vol. 38, no. 1, pp. 128–136, January 2002.

[7] V. P. Jilkov, Z. S. Duan, and X. R. Li, "State estimation with quantized measurements: Multiple model approach," in *Proceedings of 2007 International Colloquium on Information Fusion*, Xi'an, China, August 2007, pp. 1–7.

[8] Z. S. Duan, V. P. Jilkov, and X. R. Li, "State estimation with quantized measurements: Approximate MMSE approach," in *Proceedings of the 11th International Conference on Information Fusion*, Cologne, Germany, July 2008, pp. 1067–1072.

[9] R. Karlsson and F. Gustafsson, "Particle filtering for quantized sensor information," in *Proceedings of the 13th European Signal Processing Conference*, Antalya, Turkey, September 2005.

[10] K. Ito and K. Q. Xiong, "Gaussian filters for nonlinear filtering problems," *IEEE Transactions on Automatic Control*, vol. 45, no. 5, pp. 910–927, May 2000.

[11] F. W. Yang and Y. M. Li, "Set-membership filtering with state constraints," *IEEE Transactions on Aerospace and Elctronic Systems*, vol. 45, no. 4, pp. 1619–1629, October 2009.

[12] S. Julier, J. Uhlmann, and H. F. Durrant-Whyte, "A new method for nonlinear transformation of means and covariances in filters and estimators," *IEEE Transactions on Automatic Control*, vol. 45, no. 3, pp. 477–482, March 2000.

[13] M. Norgaard, N. K. Poulsen, and O. Ravn, "New developments in state estimation for nonlinear systems," *Automatica*, vol. 36, no. 11, pp. 1627–1638, November 2000.

[14] Z. L. Zhao, X. R. Li, and V. P. Jilkov, "Best linear unbiased filtering with nonlinear measurements for target tracking," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 40, no. 4, pp. 1324–1336, October 2004.

[15] X. R. Li, Z. L. Zhao, and X. B. Li, "General model-set design methods for multiple-model approach," *IEEE Transactions on Automatic Control*, vol. 50, no. 9, pp. 1260–1276, September 2005.

[16] J. H. Kotecha and P. M. Djuric, "Gibbs sampling approach for generation of truncated multivariate gaussian random variables," in *Proceedings of the 1999 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 3, Phoenix, AZ, USA, March 1999, pp. 1757–1760.

[17] Z. S. Duan, X. R. Li, and V. P. Jilkov, "State estimation with point and set measurements," *to be submitted for journal publication*, 2010.

[18] S. Ko and R. R. Bitmead, "State estimation for linear systems with state equality constraints," *Automatica*, vol. 43, no. 8, pp. 1363–1368, August 2007.