

New Estimators for Mixed Stochastic and Set Theoretic Uncertainty Models: The General Case

Uwe D. Hanebeck

Institute of Automatic Control Engineering
Technische Universität München
80290 München, Germany
Uwe.Hanebeck@ieee.org

Joachim Horn

Siemens AG, Corporate Technology
Information and Communications
81730 München, Germany
Joachim.Horn@ieee.org

Abstract

New filters are derived for estimating the n -dimensional state of a linear dynamic system based on uncertain m -dimensional observations, which suffer from two types of uncertainties simultaneously. The first uncertainty is a stochastic process with given distribution. The second uncertainty is only known to be bounded, the exact underlying distribution is unknown. The new estimators combine set theoretic and stochastic estimation in a rigorous manner and provide a continuous transition between the two classical estimation concepts. They converge to a set theoretic estimator, when the stochastic error goes to zero, and to a Kalman filter, when the bounded error vanishes. In the mixed noise case, solution sets are provided that are uncertain in a stochastic sense.

1 Introduction

In many technical systems the internal state, which is for example required for control purposes, is not directly observable and has to be reconstructed on the basis of uncertain measurements of the system output. In most cases, a stochastic uncertainty description is chosen and the state is estimated by means of a Kalman filter or one of its variations [1].

In many cases, however, uncertainties arise, for example from unmodeled dynamics or unmodeled nonlinearities, which cannot satisfactorily be described as stochastic signals with known distribution. In addition, correlated noise terms or systematic errors may be present. For these types of uncertainties, Kalman filter estimates tend to be overoptimistic [11], i.e., the covariance is underestimated. Several heuristics have been suggested for coping with this problem, which of course do not provide optimal estimators.

In some situations, bounds for these uncertainties can be provided. In that case, set theoretic estimation can be applied [13]. However, when additional uncorrelated noise is present, the error bounds become unnecessarily conservative.

In [2, 5], a basic concept for estimation in the presence of both bounded and stochastic uncertainties has been introduced. The proposed algorithm for the case of a scalar state is exact, but computationally complex. In [3, 4], an approximate solution for the case of a scalar state has been derived, that is computationally attractive. In addition, a generalization towards arbitrary dimensional states and observations of the same dimension has been proposed in [6]. Furthermore, the case of scalar measurements and arbitrary dimensional states has been treated in [8, 9, 10].

This paper is concerned with updating the estimate of an n -dimensional state based on m -dimensional observations. For this most general case, a new, approximate solution is derived, that is computationally attractive. Nevertheless, it combines both set theoretic and stochastic estimation in a rigorous manner. It bridges the gap between both estimation schemes, because a set theoretic estimator is obtained, when the stochastic error goes to zero, and a Kalman filter is obtained, when the bounded error vanishes. When both types of uncertainty are present, the new estimator provides solution sets that are uncertain in a stochastic sense. The propagation of estimates suffering from both uncertainties through a dynamic system has been discussed in [7].

In Section 2, a system model suffering from mixed stochastic and set theoretic uncertainties is introduced. The basic concept for solving the state estimation problem in this context is described in Section 3. In Section 4, a useful estimator is derived on the basis of this concept. The proposed estimator is then applied to a simple synthetic example in Section 5.

2 Problem Formulation

The key point of this paper is the use of a generalized uncertainty model unifying stochastic and set theoretic modeling. This allows the treatment of systems corrupted by both bounded and stochastic uncertainties simultaneously. Hence, the model is well-suited

for, but not limited to, the combination of deterministic/systematic errors and random noise.

To be specific, we consider a linear measurement equation given by

$$\hat{\mathbf{y}} = \mathbf{H}\mathbf{x} + \mathbf{e}_y + \mathbf{c}_y$$

with m -dimensional observation vector $\hat{\mathbf{y}}$, n -dimensional state vector \mathbf{x} , and additive uncertainties \mathbf{e}_y , \mathbf{c}_y . Furthermore, there exists a prior estimate $\hat{\mathbf{x}}_p$ of the state vector. $\hat{\mathbf{x}}_p$ also suffers from additive uncertainties \mathbf{e}_p , \mathbf{c}_p according to

$$\hat{\mathbf{x}}_p = \mathbf{x} + \mathbf{e}_p + \mathbf{c}_p .$$

The corresponding additive uncertainties are of different type:

- 1) \mathbf{e}_p , \mathbf{e}_y are uncertainties where the only prior knowledge is their boundedness, which is expressed by

$$\mathbf{e}_p^T \mathbf{E}_p^{-1} \mathbf{e}_p \leq 1, \quad \mathbf{e}_y^T \mathbf{E}_y^{-1} \mathbf{e}_y \leq 1 .$$

- 2) \mathbf{c}_p , \mathbf{c}_y are Gaussian random variables

$$\mathbf{c}_p \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_p), \quad \mathbf{c}_y \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_y)$$

which are assumed to be uncorrelated.

3 The Basic Estimation Concept

For deriving an appropriate state estimator, we define

$$\begin{aligned} \bar{\mathbf{x}}_p &= \hat{\mathbf{x}}_p - \mathbf{c}_p, \\ \bar{\mathbf{y}} &= \hat{\mathbf{y}} - \mathbf{c}_y . \end{aligned}$$

Since there is no prior information about the remaining uncertainties \mathbf{e}_p , \mathbf{e}_y besides their boundedness, we make the worst case assumption that \mathbf{e}_p , \mathbf{e}_y are fully correlated. Hence, a set theoretic estimator is appropriate for fusing $\bar{\mathbf{y}}$ and $\bar{\mathbf{x}}_p$. The fusion result is then given by the set

$$\mathcal{X}_s = \{ \underline{\mathbf{x}}_s : (\underline{\mathbf{x}}_s - \bar{\mathbf{x}}_s)^T \mathbf{E}_s^{-1} (\underline{\mathbf{x}}_s - \bar{\mathbf{x}}_s) \leq 1 \}$$

with

$$\begin{aligned} \bar{\mathbf{x}}_s &= \bar{\mathbf{x}}_p + \lambda \mathbf{E}_p \mathbf{H}^T (\mathbf{E}_y + \lambda \mathbf{H} \mathbf{E}_p \mathbf{H}^T)^{-1} \bar{\mathbf{y}}, \\ \eta &= \bar{\mathbf{y}} - \mathbf{H} \bar{\mathbf{x}}_p \end{aligned} \quad (1)$$

and

$$\mathbf{E}_s = d \mathbf{P}_s, \quad (2)$$

where d is given by

$$d = 1 + \lambda - \lambda \bar{\eta}^T (\mathbf{E}_y + \lambda \mathbf{H} \mathbf{E}_p \mathbf{H}^T)^{-1} \bar{\eta}$$

and \mathbf{P}_s is given by

$$\mathbf{P}_s = \mathbf{E}_p - \lambda \mathbf{E}_p \mathbf{H}^T (\mathbf{E}_y + \lambda \mathbf{H} \mathbf{E}_p \mathbf{H}^T)^{-1} \mathbf{H} \mathbf{E}_p .$$

The appropriate selection of the parameter $\lambda \in [0, \infty)$ will be discussed later. (1) can be rewritten as

$$\bar{\mathbf{x}}_s = \mathbf{W}_x \bar{\mathbf{x}}_p + \mathbf{W}_y \bar{\mathbf{y}}$$

with

$$\mathbf{W}_x = \mathbf{I} - \lambda \mathbf{E}_p \mathbf{H}^T (\mathbf{E}_y + \lambda \mathbf{H} \mathbf{E}_p \mathbf{H}^T)^{-1} \mathbf{H},$$

$$\mathbf{W}_y = \lambda \mathbf{E}_p \mathbf{H}^T (\mathbf{E}_y + \lambda \mathbf{H} \mathbf{E}_p \mathbf{H}^T)^{-1}$$

and

$$\mathbf{W}_x + \mathbf{W}_y \mathbf{H} = \mathbf{I} .$$

However, $\bar{\mathbf{x}}_p$ and $\bar{\mathbf{y}}$ cannot be measured directly. Only their noisy counterparts given by $\hat{\mathbf{x}}_p = \bar{\mathbf{x}}_p + \mathbf{c}_p$ and $\hat{\mathbf{y}} = \bar{\mathbf{y}} + \mathbf{c}_y$ are available. Hence, the midpoint of the ellipsoidal set \mathcal{X}_s is a random variable denoted by $\underline{\mathbf{X}}_s$. It is defined only when the difference $\bar{\mathbf{y}} - \mathbf{H} \bar{\mathbf{x}}_p$ is bounded by the Minkowski sum $\tilde{\mathcal{B}}$ of \mathbf{E}_y and $\mathbf{H} \mathbf{E}_p \mathbf{H}^T$, which is not an ellipsoidal set.

To simplify the following derivations, we note that the set theoretic uncertainty \mathbf{E}_s given by (2) depends on $\bar{\mathbf{y}}$ and $\bar{\mathbf{x}}_p$. Setting $\bar{\eta} = \mathbf{0}$ leads to $d = 1 + \lambda$ and is equivalent to bounding \mathbf{E}_s from above. The resulting \mathbf{E}_s is then given by

$$\begin{aligned} \mathbf{E}_s &= (1 + \lambda) \mathbf{E}_p \\ &\quad - \lambda(1 + \lambda) \mathbf{E}_p \mathbf{H}^T (\mathbf{E}_y + \lambda \mathbf{H} \mathbf{E}_p \mathbf{H}^T)^{-1} \mathbf{H} \mathbf{E}_p . \end{aligned} \quad (3)$$

Since the simplified \mathbf{E}_s in (3) does not depend on the actual values, it is not a random variable.

4 Derivation of the New Estimator

Since the set $\tilde{\mathcal{B}}$ is not an ellipsoid, a simplification is obtained by using a bounding ellipsoid according to

$$\mathcal{B} = \{ \underline{\mathbf{z}} : \underline{\mathbf{z}}^T \mathbf{B}^{-1} \underline{\mathbf{z}} \leq 1 \}$$

with

$$\bar{\mathbf{y}} - \mathbf{H} \bar{\mathbf{x}}_p \in \tilde{\mathcal{B}} \subset \mathcal{B} .$$

A parametrized family of bounding ellipsoid is obtained by

$$\mathbf{B} = \frac{1}{0.5 - \kappa} \mathbf{E}_y + \frac{1}{0.5 + \kappa} \mathbf{H} \mathbf{E}_p \mathbf{H}^T$$

for $\kappa \in (-0.5, 0.5)$. κ is selected in such a way, that the volume of the resulting bounding ellipsoid for the Minkowski sum is minimized. Thus, $\underline{\mathbf{X}}_s$ is a random variable approximately given by

$$\underline{\mathbf{X}}_s = \begin{cases} \mathbf{W}_x \underline{\mathbf{X}}_p + \mathbf{W}_y \underline{\mathbf{Y}} & \text{for } \bar{\mathbf{y}} - \mathbf{H} \bar{\mathbf{x}}_p \in \mathcal{B} \\ \text{undefined} & \text{elsewhere} \end{cases}$$

Hence, the desired density is given by [12]

$$f_{\underline{x}_s}(\underline{x}_s) = \frac{1}{|\mathbf{W}_x|} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{py}(\mathbf{W}_x^{-1}(\underline{x}_s - \mathbf{W}_y \underline{y}), \underline{y}) d\underline{y}, \quad (4)$$

where the weighting matrix \mathbf{W}_x is assumed to be regular, and $f_{py}(\underline{x}_p, \underline{y})$ is defined by

$$f_{py}(\underline{x}_p, \underline{y}) = \begin{cases} c_{py} f_p(\underline{x}_p) f_y(\underline{y}) & \text{for } \underline{y} - \mathbf{H}\underline{x}_p \in \mathcal{B} \\ 0 & \text{elsewhere} \end{cases} \quad (5)$$

with normalizing constant c_{py} . By means of an indicator function

$$I(\underline{x}_p, \underline{y}) = \begin{cases} 1 & \text{for } \underline{y} - \mathbf{H}\underline{x}_p \in \mathcal{B} \\ 0 & \text{elsewhere} \end{cases},$$

(5) can be compactly written as

$$f_{py}(\underline{x}_p, \underline{y}) = c_{py} f_p(\underline{x}_p) f_y(\underline{y}) I(\underline{x}_p, \underline{y}).$$

The key idea to finding an approximate solution for the probability density function is to approximate the indicator function by a weighted sum of Gaussians according to

$$I(\underline{x}_p, \underline{y}) \approx \sum_{i=1}^L \exp \left\{ -\frac{1}{2} (\underline{y} - \mathbf{H}\underline{x}_p - \underline{m}_g^i)^T \mathbf{C}_g^{-1} (\underline{y} - \mathbf{H}\underline{x}_p - \underline{m}_g^i) \right\},$$

with \underline{m}_g^i and symmetric, positive definite matrix \mathbf{C}_g appropriately chosen.

Based on this approximation of the bound, the exact density $f_{\underline{x}_s}$ in (4) can be approximated by a sum of simple densities. For that purpose, we first consider one term of the sum which gives

$$f_{\underline{x}_s}^i(\underline{x}_s) = c_1 \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp \left\{ -\frac{1}{2} \left[(\underline{x}_p - \hat{\underline{x}}_p)^T \mathbf{C}_p^{-1} (\underline{x}_p - \hat{\underline{x}}_p) + (\underline{y} - \hat{\underline{y}})^T \mathbf{C}_y^{-1} (\underline{y} - \hat{\underline{y}}) + (\underline{y} - \mathbf{H}\underline{x}_p - \underline{m}_g^i)^T \mathbf{C}_g^{-1} (\underline{y} - \mathbf{H}\underline{x}_p - \underline{m}_g^i) \right] \right\} d\underline{y}$$

for $i = 1, \dots, L$, $\underline{x}_p = \mathbf{W}_x^{-1}[\underline{x}_s - \mathbf{W}_y \underline{y}]$, and normalizing constant c_1 . A tedious calculation reveals that this approximation can be simplified to

$$f_{\underline{x}_s}^i(\underline{x}_s) = g_i c_2 \exp \left\{ -\frac{1}{2} (\underline{x}_s - \hat{\underline{x}}_s^i)^T (\mathbf{C}_s^i)^{-1} (\underline{x}_s - \hat{\underline{x}}_s^i) \right\}$$

with weighting factors

$$g_i = \exp \left\{ -\frac{1}{2} (\hat{\underline{y}} - \mathbf{H}\hat{\underline{x}}_p - \underline{m}_g^i)^T (\mathbf{H}\mathbf{C}_p\mathbf{H}^T + \mathbf{C}_y + \mathbf{C}_g)^{-1} (\hat{\underline{y}} - \mathbf{H}\hat{\underline{x}}_p - \underline{m}_g^i) \right\},$$

normalizing constant c_2 , and individual means

$$\hat{\underline{x}}_s^i = \mathbf{W}_x \hat{\underline{x}}_p + \mathbf{W}_y \hat{\underline{y}} + (\mathbf{W}_x \mathbf{C}_p \mathbf{H}^T - \mathbf{W}_y \mathbf{C}_y) (\mathbf{H}\mathbf{C}_p\mathbf{H}^T + \mathbf{C}_y + \mathbf{C}_g)^{-1} (\hat{\underline{y}} - \mathbf{H}\hat{\underline{x}}_p - \underline{m}_g^i)$$

for $i = 1, \dots, L$. The covariance matrices are the same for each term in the sum and given by

$$\mathbf{C}_s^i = \mathbf{W}_x \mathbf{C}_p \mathbf{W}_x^T + \mathbf{W}_y \mathbf{C}_y \mathbf{W}_y^T - (\mathbf{W}_x \mathbf{C}_p \mathbf{H}^T - \mathbf{W}_y \mathbf{C}_y) (\mathbf{H}\mathbf{C}_p\mathbf{H}^T + \mathbf{C}_y + \mathbf{C}_g)^{-1} (\mathbf{W}_x \mathbf{C}_p \mathbf{H}^T - \mathbf{W}_y \mathbf{C}_y)^T.$$

The approximate solution for the density $f_{\underline{x}_s}$ is then

$$f_{\underline{x}_s}(\underline{x}_s) \approx \sum_{i=1}^L f_{\underline{x}_s}^i(\underline{x}_s),$$

which is a weighted sum of Gaussian densities, where the weighting factors g_i are themselves values of a Gaussian function.

An approximate expression for the expected value $\hat{\underline{x}}_s = E[\underline{X}_s]$ of the random variable \underline{X}_s is then given by

$$\hat{\underline{x}}_s \approx \left(\sum_{i=1}^L g_i \hat{\underline{x}}_s^i \right) / \left(\sum_{i=1}^L g_i \right), \quad (6)$$

an approximate expression for the covariance by

$$\mathbf{C}_s \approx \frac{\sum_{i=1}^L g_i \{ \mathbf{C}_s^i + \hat{\underline{x}}_s^i (\hat{\underline{x}}_s^i)^T \}}{\sum_{i=1}^L g_i} - \hat{\underline{x}}_s \hat{\underline{x}}_s^T. \quad (7)$$

To summarize the results: The uncertainty of the fusion result is given by a bounded uncertainty and a weighted sum of Gaussian densities. To end up with a second order estimator, the weighted sum of Gaussian densities is approximated by mean and covariance.

The new estimator unifies Kalman filtering and set theoretic estimation: A Kalman filter is approached, when the bounded error vanishes. On the other hand, a set theoretic estimator is attained, when the stochastic error goes to zero. When both types of uncertainty are present simultaneously, the new estimator provides solution sets that are uncertain in a stochastic sense.

5 Simulation Example

Consider two measurement equations according to

$$\underline{y}_1^k = \mathbf{H}_1 \underline{x} + \underline{e}_1^k + \underline{c}_1^k,$$

$$\underline{y}_2^k = \mathbf{H}_2 \underline{x} + \underline{e}_2^k + \underline{c}_2^k$$

with

$$\mathbf{H}_1 = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{H}_2 = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix}.$$

$\underline{e}_1^k, \underline{e}_2^k$ are unknown errors with bounds given by

$$\mathbf{E}_1 = \begin{bmatrix} 36900 & 24300 \\ 24300 & 16200 \end{bmatrix}, \quad \mathbf{E}_2 = \begin{bmatrix} 1200 & 0 \\ 0 & 1200 \end{bmatrix}.$$

In this simulation, $\underline{e}_1^k, \underline{e}_2^k$ are constant values given by $\underline{e}_1^k = [20 \ 10]^T$, $\underline{e}_2^k = [5 \ 25]^T$. $\underline{c}_1^k, \underline{c}_2^k$ are samples from independent, zero mean white Gaussian random processes with covariance matrices

$$\mathbf{C}_1 = \begin{bmatrix} 100^2 & 0 \\ 0 & 100^2 \end{bmatrix}, \quad \mathbf{C}_2 = \begin{bmatrix} 100^2 & 0 \\ 0 & 100^2 \end{bmatrix}.$$

No prior knowledge is given.

In order to apply the standard approach for estimating the state $\underline{x} = [x_1 \ x_2]^T$, the Kalman filter, the bounded uncertainties are interpreted as additional white Gaussian noise resulting in total covariances $\mathbf{E}_1 + \mathbf{C}_1$ and $\mathbf{E}_2 + \mathbf{C}_2$. The evolution of the resulting confidence set over time is depicted in Figure 1. The confidence set has been calculated based on 9 times the covariance matrix centered at $\hat{\underline{x}}_s^k$. The true state $\underline{x} = [100 \ 100]^T$ is marked by a dot. **Note:** The confidence set for $k \rightarrow \infty$ does not contain the true state.

For estimating the state \underline{x} by using the proposed estimator, the formulae for $\hat{\underline{x}}_s^k$ in (6), \mathbf{E}_s^k in (3), and \mathbf{C}_s^k in (7) are used for both measurement equations at time k . The parameter λ_k is chosen such that $\det(\mathbf{E}_s^k + \mathbf{C}_s^k)$ is minimized. Figure 2 depicts how the resulting estimate evolves over time. Here, the confidence set is given by the Minkowski sum of \mathbf{E}_s^k and $9\mathbf{C}_s^k$ centered at $\hat{\underline{x}}_s^k$. The optimal estimate for an infinite number of measurements would be given by the set resulting from intersecting the two ellipses representing the sets of states that would be obtained by inverting the measurement equations when no stochastic uncertainty is present. **Note:** The confidence set for $k \rightarrow \infty$ bounds the exact set from above and contains the true state.

6 Conclusions

Many applications require estimating the state of a linear system from uncertain observations, where the uncertainties are additively composed of both 1) noise with known distribution and 2) noise with known bounds. The new estimator then provides a rigorous framework for solving these problems efficiently.

This paper focused on the measurement update, i.e., on updating the estimate of an n -dimensional state based on given m -dimensional observations. The time update, i.e., propagating the state estimate through a dynamic system, is discussed in [7].

References

- [1] B. D. O. Anderson, J. B. Moore, *Optimal Filtering*, Prentice-Hall, 1979.
- [2] U. D. Hanebeck, J. Horn, G. Schmidt, "On Combining Set Theoretic and Bayesian Estimation", *Proceedings of the 1996 IEEE International Conference on Robotics and Automation (ICRA '96)*, pp. 3081-3086, 1996.
- [3] U. D. Hanebeck, J. Horn, "A New State Estimator for a Mixed Stochastic and Set Theoretic Uncertainty Model", *Proceedings of SPIE Vol. 3720, AeroSense Symposium*, pp. 336-344, 1999.
- [4] U. D. Hanebeck, J. Horn, "A New Estimator for Mixed Stochastic and Set Theoretic Uncertainty Models Applied to Mobile Robot Localization", *Proceedings of the 1999 IEEE International Conference on Robotics and Automation (ICRA '99)*, pp. 1335-1340, 1999.
- [5] U. D. Hanebeck, J. Horn, G. Schmidt, "On Combining Statistical and Set-Theoretic Estimation", *Automatica*, Vol. 35, No. 6, 1999, pp. 1101-1109.
- [6] U. D. Hanebeck, J. Horn, "New Estimators for Mixed Stochastic and Set Theoretic Uncertainty Models: The Vector Case", *Proceedings of the Fifth European Control Conference (ECC'99)*, 1999.
- [7] U. D. Hanebeck, J. Horn, "New Results for State Estimation in the Presence of Mixed Stochastic and Set Theoretic Uncertainties", *Proceedings of the 1999 IEEE Systems, Man, and Cybernetics Conference (SMC'99)*, Vol. 5, pp. 86-91, 1999.
- [8] U. D. Hanebeck, J. Horn, "State Estimation Based on Observations Simultaneously Corrupted by Random Noise with Known Distribution and Uncertainties with Known Bounds", *1999 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'99)*, pp. 665-670, 1999.
- [9] U. D. Hanebeck, J. Horn, "New Estimators for Mixed Stochastic and Set Theoretic Uncertainty Models: The Scalar Measurement Case", *Proceedings of the 1999 IEEE Conference on Decision and Control (CDC'99)*, pp. 1934-1939, 1999.
- [10] U. D. Hanebeck, J. Horn, "Fusing Information Simultaneously Corrupted by Uncertainties with Known Bounds and Random Noise with Known Distribution", *Information Fusion*, Vol. 1, No. 1, pp. 55-63, 2000.
- [11] S. J. Julier, J. K. Uhlmann, "A Non-Divergent Estimation Algorithm in the Presence of Unknown Correlations", *Proceedings of the 1997 American Control Conference (ACC'97)*, pp. 2369-2373, 1997.
- [12] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, Second Edition, McGraw-Hill Book Company, 1984.
- [13] F. C. Schweppe, *Uncertain Dynamic Systems*, Prentice-Hall, 1973.

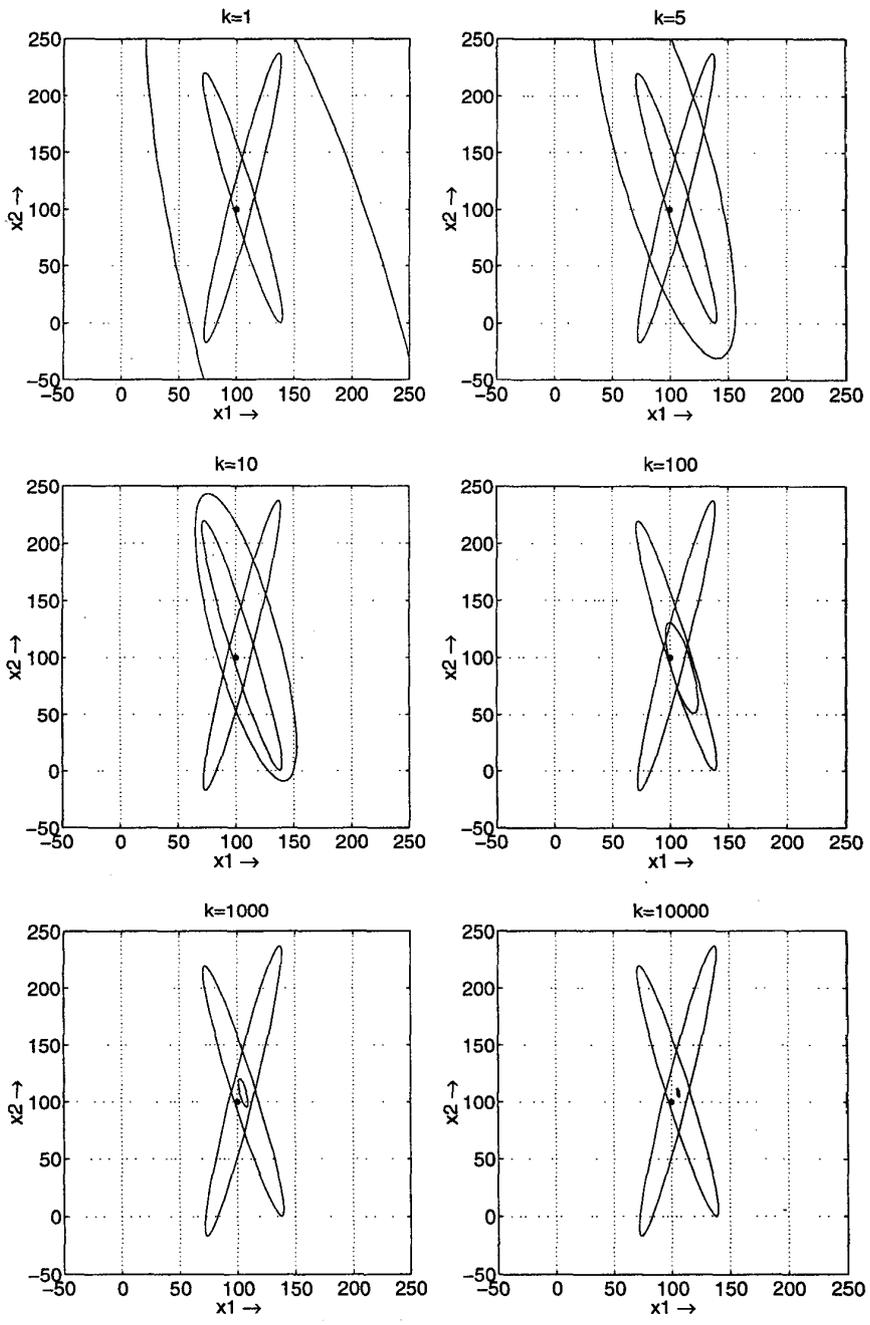


Figure 1: Results of applying the Kalman filter.

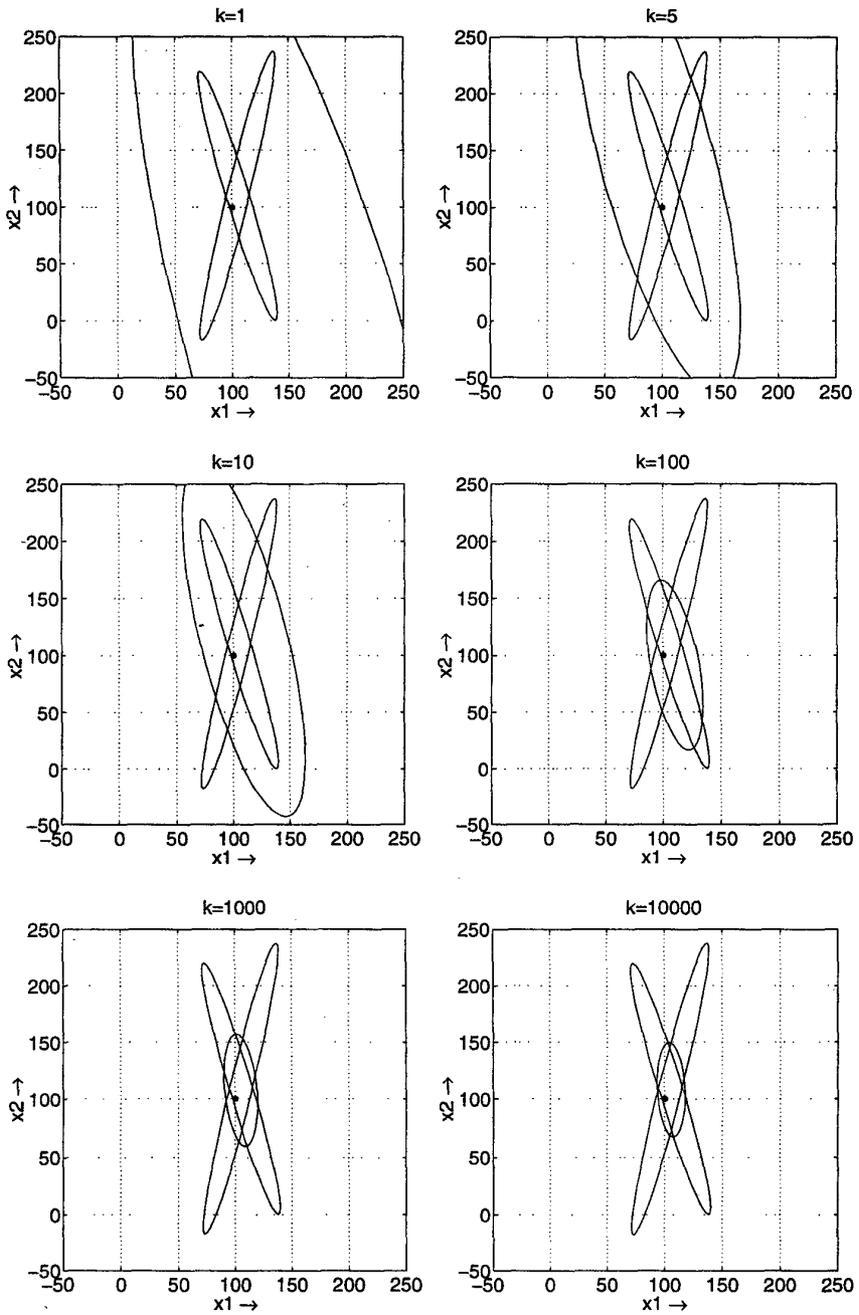


Figure 2: Results of applying the new estimator.