RIDGE REGRESSION SIGNAL PROCESSING

Mark R. Kuhl
Ohio University
Athens, Ohio

SUMMARY

The introduction of the Global Positioning System (GPS) into the National Airspace System (NAS) necessitates the development of Receiver Autonomous Integrity Monitoring (RAIM) techniques as identified by RTCA Special Committee 159. In addition, it is anticipated that future multisensor positioning systems will also utilize RAIM schemes to guarantee a high level of integrity, as evidenced by current FAA programs involving hybrid positioning systems such as GPS/LORAN-C.

In order to guarantee a certain level of integrity, a thorough understanding of modern estimation techniques applied to navigational problems is required. In this paper, the extended Kalman filter (EKF) is derived and analyzed under poor geometry conditions. It was found that the performance of the EKF is difficult to predict, since the EKF is designed for a Gaussian environment. A novel approach is implemented which incorporates ridge regression to explain the behavior of an EKF in the presence of dynamics under poor geometry conditions.

The basic principles of ridge regression theory are presented, followed by the derivation of a linearized recursive ridge estimator. Computer simulations are performed to confirm the underlying theory and to provide a comparative analysis of the EKF and the recursive ridge estimator.

BACKGROUND OF RIDGE REGRESSION

Ridge regression signal processing is introduced in order to counter the effects of poor geometry on position estimation. Robert Kelly first applied ridge regression to navigational problems in 1988 (ref. 1). Many other papers have followed since then explaining the further developments of ridge regression (refs. 1-6). The basic idea behind ridge regression is presented below.

For navigation, the appropriate error criterion is the mean square error (MSE). It expresses the deviation of the vehicle with respect to its intended path. The MSE consists of the sum of two error components; MSE = Variance + Bias^2. For ordinary least squares (OLS) estimation, the position solution is unbiased; therefore, the MSE is only made up of the variance term. (Note that the Kalman filter is a recursive form of the OLS estimator.) Under poor geometry conditions, the variance of the position errors inflates significantly, giving an inaccurate position estimate. The ridge estimator takes advantage of an extra degree of freedom in the MSE which is not used by the OLS estimator - the bias term. In effect, small biases induced by the ridge estimator decrease the variance term such that the overall MSE is

smaller than the MSE obtained from an unbiased estimator, as illustrated in figure 1.

The linear model for a system with an unknown n x 1 measurement bias vector, δB, and a measurement noise vector, $\underline{e}$, is given by

$$\underline{Y} = H\underline{\beta} + \delta B + \underline{e} \tag{1}$$

where $\underline{Y}$ is the n x 1 range measurement vector, $\underline{\beta}$ is the p x 1 unknown system state vector (or parameter vector), and H is the n x p predictor (or design) matrix which relates the range measurement vector to the system state vector. Also, the measurement noise is uncorrelated; $\text{cov}(\underline{e} + \delta B) = [\underline{e}\,\underline{e}^T] = \sigma^2 I$, where I is the n x n identity matrix.

The OLS estimate of equation 1 is

$$\hat{\underline{\beta}}_{OLS} = (H^T H)^{-1} H^T \underline{Y} \tag{2}$$

The corresponding ridge estimate of equation 1 is

$$\hat{\underline{\beta}}_R = (H^T H + P_R)^{-1} H^T \underline{Y} \tag{3}$$

where P_R is the ridge parameter matrix (which is formed from the chosen ridge biasing parameters, k_i). When P_R consists of zeros, the ridge estimator reduces to the OLS estimator. Adding a non-zero ridge parameter matrix to $H^T H$ purposely upsets the balance between the first and second moment components of $\hat{\underline{\beta}}$, thereby introducing a bias.

For a complete discussion on the comparative analysis of the ridge and OLS estimators, see reference 6.

THE EXTENDED KALMAN FILTER (EKF)

From reference 6, figure 2 summarizes how the normal Kalman equations relate to each other to form the Kalman filter. Since the equations that relate the measurements to the state vector are usually nonlinear (i.e. the H matrix is nonlinear), an extended Kalman filter (EKF) is needed. Therefore, a linearization procedure is performed when deriving the Kalman filter equations, see reference 7.

After the linearization procedure, the recursive update equation for the extended Kalman filter is given by

$$\hat{\underline{\beta}}_k = \hat{\underline{\beta}}_{k/k-1} + (P^{-1}_{k/k-1} + H_k^T R^{-1} H_k)^{-1} (H_k^T R^{-1} \delta \underline{Y}_k) \tag{4}$$

where $\delta \underline{Y}_k$ is the difference between the actual measurements and the predicted measurements (the innovations) as shown below

$$\delta \underline{Y}_k = \underline{Y}_k - g_k(\hat{\underline{\beta}}_{k/k-1}) \tag{5}$$

where g_k is a nonlinear vector-valued function which relates the system state to the measurements.

An optimum unbiased estimator arises when both the model and the estimator match the process which generates the data. Kalman filter optimization techniques include the selection of the initial error covariance matrix, P_0, and the system covariance matrix, Q, based upon past experience or by adaptively tuning the filter until its innovations (residuals) become white (i.e. zero mean and random). However, the selection of the proper Q matrix is usually not a very easy task (ref. 8).

For example, the system covariance matrix Q is often set artificially high such that the Kalman filter can track the vehicle when it encounters dynamics such as turning-induced accelerations. Therefore, the Kalman filter allows more noise in the solution during periods of low dynamics. Although some methods exist for selecting Q adaptively (ref. 9), these are stochastic in nature.

Another problem arises when the Kalman filter is subjected to a poor geometry condition. In the case of inaccurate P_0 and Q matrices (a mismatched filter), the filter may become biased. The performance of a biased Kalman filter is not readily understood, as the Kalman filter is optimal and defined in a Gaussian environment only. Therefore, the performance of a Kalman filter for a deterministic maneuver in a poor geometry condition cannot be predicted from the regular Kalman filter equations. Furthermore, poor geometry effects cannot be minimized by observing the innovations in the Kalman filter's update equation (see equation 4), because these effects only appear in the estimator $\hat{\underline{\beta}}$.

In applying the Kalman filter to navigation, the MSE is the appropriate error criterion to be minimized. The MSE of a mismatched Kalman filter is not necessarily the smallest obtainable. Recall that the MSE is the sum of two components: the variance term and bias term squared. Since the Kalman filter is developed from the OLS estimator, it is inherently an unbiased estimator. On the other hand, a biased estimator, such as the ridge recursive filter presented in the next section, is purposely not matched to the process that generates the data in order to achieve a smaller MSE.

THE RIDGE/EKF FILTER

Extending the work presented in reference 5, the linearized state update equation for the extended Ridge/EKF is obtained as

$$\hat{\underline{\beta}}_{Rk} = \hat{\underline{\beta}}_{Rk/k-1} + (P^{-1}_{k/k-1} + P_{Rk} + H_k^T R^{-1} H_k)^{-1} (H_k^T R^{-1} \delta \underline{Y}_k) \tag{6}$$

where P_{Rk} has been added to the $(P^{-1}_{k/k-1} + H_k^T R^{-1} H_k)$ term in order to reduce the effects of poor geometry. Also, the update equation for the error covariance matrix is given by

$$P_k = (P^{-1}_{k/k-1} + P_{Rk} + H_k^T R^{-1} H_k)^{-1}. \tag{7}$$

Two cases may now be defined in which ridge regression can explain the behavior of the Kalman filter. First, in the absence of system dynamics, the ridge parameter matrix, P_{Rk} is functionally equivalent to the inverse of the state error covariance matrix, $P^{-1}_{k/k-1}$. Therefore, the EKF has similar convergence properties when the model incorrectly represents the system and $P_{k/k-1}$ is small. (See reference 6 for a discussion on the convergence properties of estimators.) Second, if dynamics exist, P_{Rk} is related to both $P^{-1}_{k/k-1}$ and the system error

covariance matrix, Q, given by the following relation

$$(\Phi P_k \Phi^T + Q)^{-1} \sim (\Phi P_k \Phi^T + Q)^{-1} + P_{Rk} \tag{8}$$

where, in this case, the $\sim$ symbol means "relating to". Again, the EKF has similar convergence properties when the model incorrectly represents the system and $(\Phi P_k \Phi^T + Q)^{-1}$ is small.

Usually P_k is small, which means that the $P_{k/k-1}^{-1}$ term in equation 6 is large; therefore, there is no geometry problem. Note that P_k varies as the Kalman filter is updated, but normally a constant Q is added which "limits" it (i.e. puts uncertainty in the model). As seen in the above equation, P_k will be large when Q is large. Q is chosen large for dynamic situations (i.e. turn-induced accelerations). When a poor geometry condition exists in addition to the dynamic situation, P_{Rk} may be added to counteract the large Q. Therefore, a proper P_{Rk} can be chosen to incorporate ridge regression into the Kalman filter.

The key idea in developing the ridge recursive filter is the following: Each step in the recursive process is viewed as a new prior linear model wherein the last estimate $\hat{\beta}_{Rk/k-1}$ is the prior equation for the next iteration. The ridge solution is recomputed at each step using selection rules, described in reference 6, to determine a proper P_{Rk}.

SIMULATION RESULTS

This section provides a simulation to evaluate the behavior of an extended Kalman filter (EKF) when dynamics exist and a poor geometry condition occurs. Initial results show how ridge regression can explain the behavior of a mismatched Kalman filter.

The EKF is implemented in the following simulation scenario: An aircraft is moving along a constant velocity flight path making range measurements to two Distance Measuring Equipment (DME) stations. At first (t = 0 seconds), the station locations constitute a good geometry situation with respect to the aircraft position. At t = 50 seconds, the aircraft "enters a poor geometry condition". This happens when the aircraft switches from making range measurements to DME stations 1 and 2 to making range measurements to DME stations 1 and 3. After some time in this poor geometry condition, the aircraft turns 30° to the right (t = 150 seconds) and continues in this situation until the end of the simulation run (t = 300 seconds). The numbers for the scenario set-up are given below:

DME station 1 location, $[X_1 \ Y_1]$ = [10.0e05 0] feet
DME station 2 location, $[X_2 \ Y_2]$ = [85.0e05 75.0e05] feet
DME station 3 location, $[X_3 \ Y_3]$ = [160.0e05 0] feet

initial aircraft location, $[X_{TR} \ Y_{TR}]$ = [85.0e05 -1.0e05] feet
initial aircraft velocity, $[V_x \ V_y]$ = [100 0] feet/sec
aircraft velocity after turn, $[V_x \ V_y]$ = [86.6 50] feet/sec
measurement noise, σ = 30 feet
measurement bias, $\underline{\delta B}$ = [0 0] feet
aircraft initial position offset, $[\delta X \ \delta Y]$ = [100 100] feet

70

The aircraft makes 16 measurements to each DME station per second and updates its position estimate once every 16th of a second (i.e. recursive estimation).

The initialization of the tuned Kalman filter is as follows: The initial estimated system state vector and its associated error covariance matrix are

$$\hat{\underline{B}}_0 = \begin{bmatrix} \hat{X} \\ \hat{Y} \\ \hat{V}_X \\ \hat{V}_Y \end{bmatrix} = \begin{bmatrix} 85.0e05 - 100 \\ -1.0e05 - 100 \\ 0 \\ 0 \end{bmatrix} \qquad P_0 = \begin{bmatrix} 100^2 & 0 & 0 & 0 \\ 0 & 100^2 & 0 & 0 \\ 0 & 0 & 100^2 & 0 \\ 0 & 0 & 0 & 100^2 \end{bmatrix}$$

The system process covariance matrix is

$$Q = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0.1 & 0 \\ 0 & 0 & 0 & 0.1 \end{bmatrix}$$

And the measurement noise covariance matrix is

$$R = \begin{bmatrix} 30^2 & 0 \\ 0 & 30^2 \end{bmatrix}$$

The results of the tuned Kalman filter for the above scenario are portrayed in figure 3 where the estimated flight path (each estimate is represented by an "o") is compared to the true flight path (represented by a straight line). Note that there is some lag in the Kalman filter estimates when the aircraft starts to turn. The square root of the MSE is 134 feet. This result shows that the Kalman filter can be properly tuned to handle dynamics when conditions of poor geometry arise.

Normally, the Q matrix in the Kalman filter is set artificially high such that the filter can follow a vehicle through periods when dynamics exist. In order to see how ridge regression can reduce the square root of the MSE, implement the same simulation set-up provided above with a much higher Q matrix.

$$Q = \begin{bmatrix} 500 & 0 & 0 & 0 \\ 0 & 500 & 0 & 0 \\ 0 & 0 & 500 & 0 \\ 0 & 0 & 0 & 500 \end{bmatrix}$$

The estimated and true aircraft flight paths are shown in figure 4. The square root of the MSE becomes 304 feet.

Now, add a proper ridge parameter matrix, P_R, to counteract the effects of poor geometry. In effect, adding P_R is similar to making Q smaller. Figure 5 depicts the resulting estimated and true aircraft flight paths. The square root of the MSE reduces to 142 feet. This shows that a properly selected P_R matrix can give results that are comparable to the tuned Kalman filter described above. Efforts are continuing to address the performance of the extended Ridge/EKF in detail.

CONCLUSIONS

Ridge regression signal processing is introduced to explain the behavior of a mismatched Kalman filter. This paper provides an initial comparison of the extended Kalman filter (EKF) and the recursive ridge filter (the Ridge/EKF). It is shown that a proper ridge parameter matrix P_R, may be chosen to counteract poor geometry conditions that may arise when estimating a position.

REFERENCES

1. Kelly, R. J.: Reducing Geometric Dilution of Precision Using Ridge Regression Signal Processing. Plans '88, Orlando, FL, November 29 through December 2, 1988.

2. Kelly, R. J.: OMEGA Accuracy Improvement Using Ridge Regression. OMEGA: International and Economic, Proceedings of the Fourteenth Annual Meeting, Long Beach, CA, 1989.

3. Kelly, R. J.: Global Positioning System Accuracy Improvement Using Ridge Regression. AIAA Guidance, Navigation and Control Conference, Boston, MA, August 14-16, 1989.

4. Kelly, R. J., van Graas, F., and Kuhl, M. R.: Improved Effectiveness of GPS RAIM through Ridge Regression Signal Processing. Proceedings of the 2nd International Satellite Division Meeting of the ION, Colorado Springs, CO, September 27-29, 1989.

5. Kelly, R. J.: Reducing Geometric Dilution of Precision Using Ridge Regression. IEEE Transactions on Aerospace and Electronic Systems, Vol. 26, No. 1, January 1990.

6. Kuhl, M. R.: Ridge Regression Signal Processing Applied to Multisensor Position Fixing. Master's Thesis, Ohio University, Department of Electrical and Computer Engineering, Athens, OH, June 1990.

7. Ewing, P. L.: The Design and Implementation of Tracking and Filtering Algorithms for an Aircraft Beacon Collision Warning System. Master's Thesis, Ohio University, Department of Electrical and Computer Engineering, Athens, OH, March 1989.

8. Brown, R. G.: Introduction to Random Signal Analysis and Kalman Filtering. John Wiley and Sons, New York, 1983.

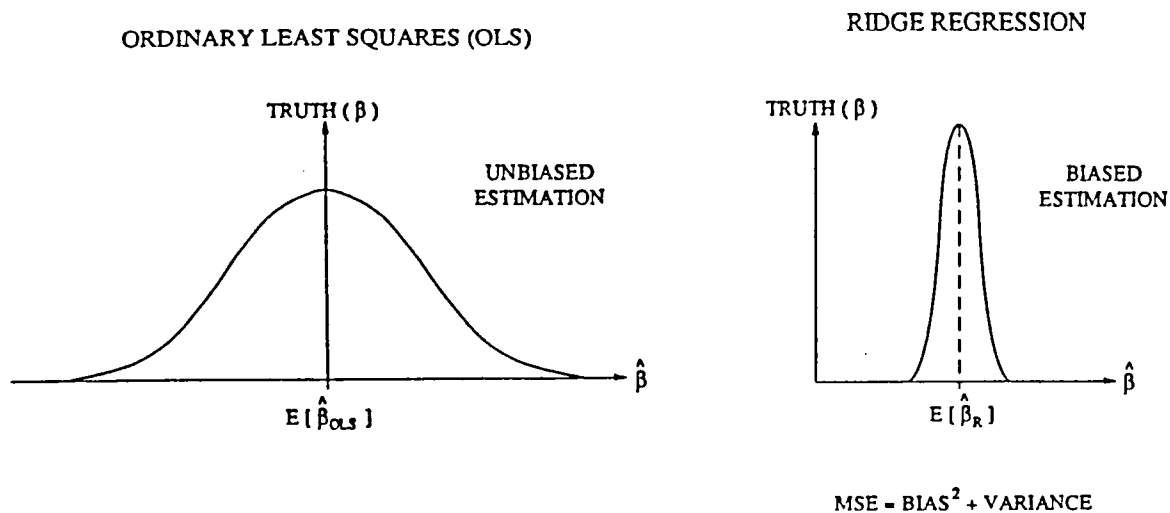9. Sorenson, H. W.: Kalman Filtering: Theory and Applications. IEEE Press, New York, 1985.

ORDINARY LEAST SQUARES (OLS)   RIDGE REGRESSION

TRUTH (β)   TRUTH (β)

UNBIASED   BIASED
ESTIMATION   ESTIMATION

$\hat{\beta}$ $\hat{\beta}$

$E[\hat{\beta}_{OLS}]$ $E[\hat{\beta}_R]$

$$MSE = BIAS^2 + VARIANCE$$

Figure 1. Comparison of unbiased and biased solution distributions and definition of the Mean Squared Error (MSE).

Enter the *a priori* estimate $\hat{\beta}_0$
and its error covariance P_0

Compute Kalman gain:

$$K_k = P_{k/k-1}H_k^T(H_kP_{k/k-1}H_k^T + R)^{-1}$$

Project Ahead:

$$\hat{\beta}_{k+1/k} = \Phi\hat{\beta}_k$$

$$P_{k+1/k} = \Phi P_k\Phi^T + Q$$

Update estimate with measurement Y_k:

$$\hat{\beta}_k = \hat{\beta}_{k/k-1} + K_k(Y_k - H_k\hat{\beta}_{k/k-1})$$

Compute error covariance for updated estimate:
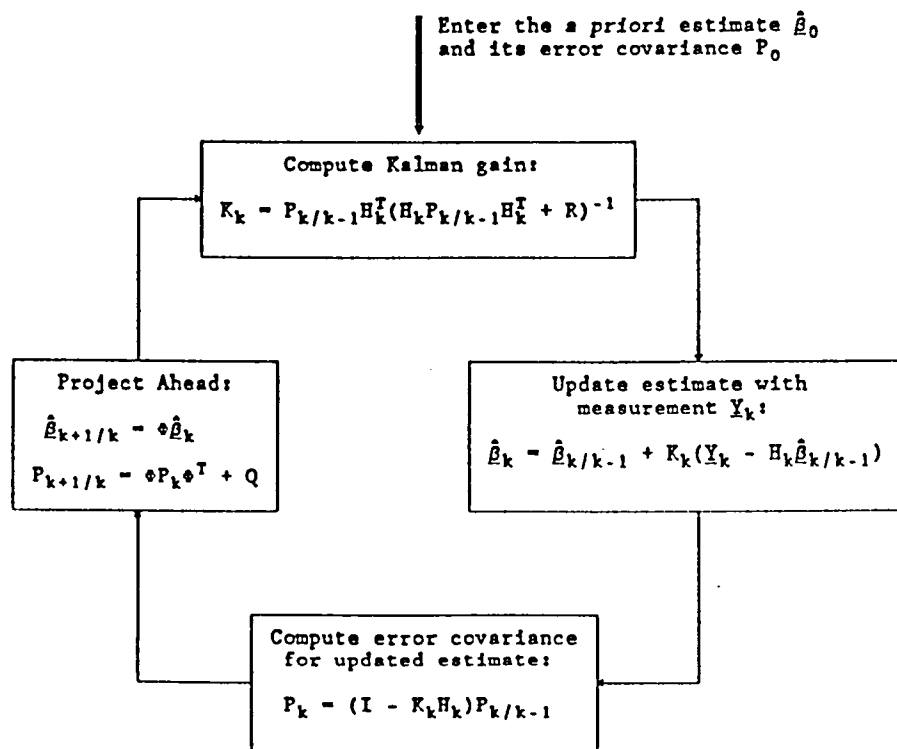
$$P_k = (I - K_kH_k)P_{k/k-1}$$

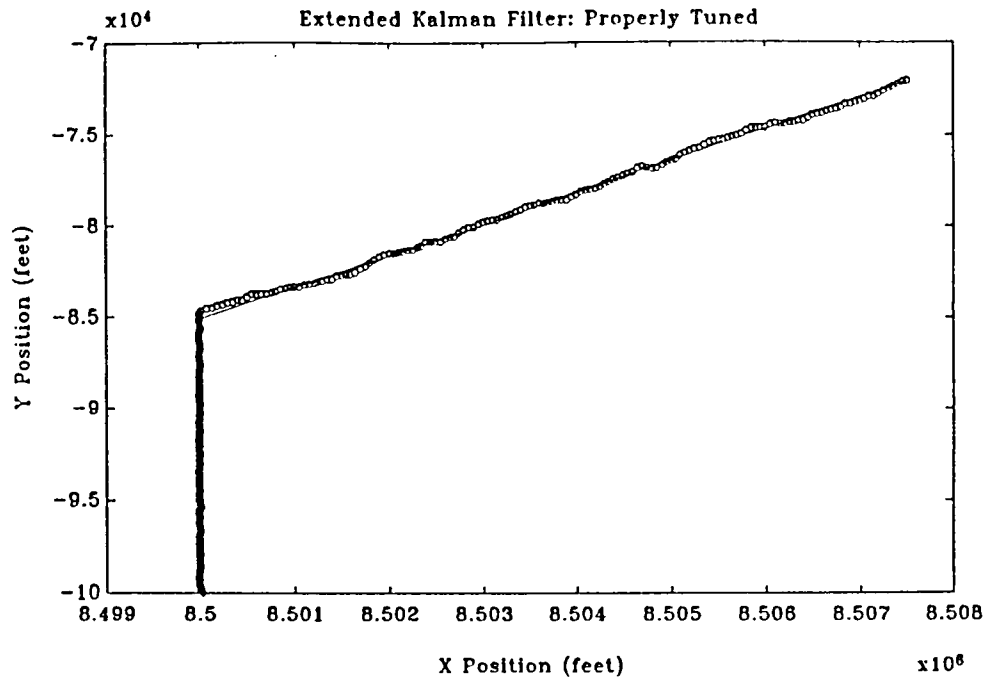Figure 2. Kalman filter equations for a linear system.

Figure 3. Simulation results for a properly tuned extended Kalman filter (small system covariance).
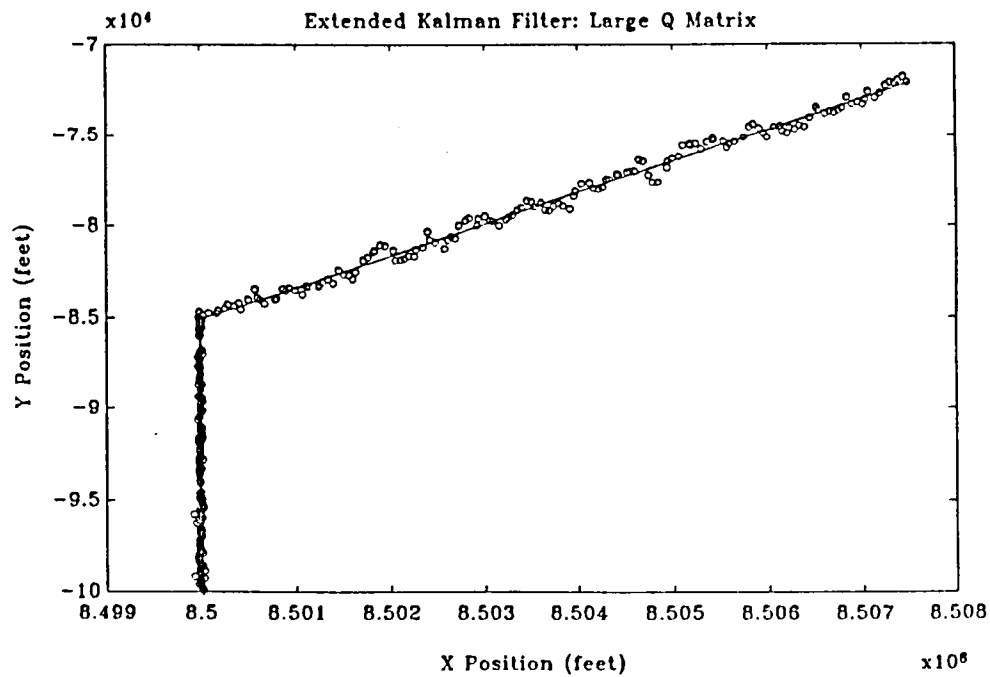

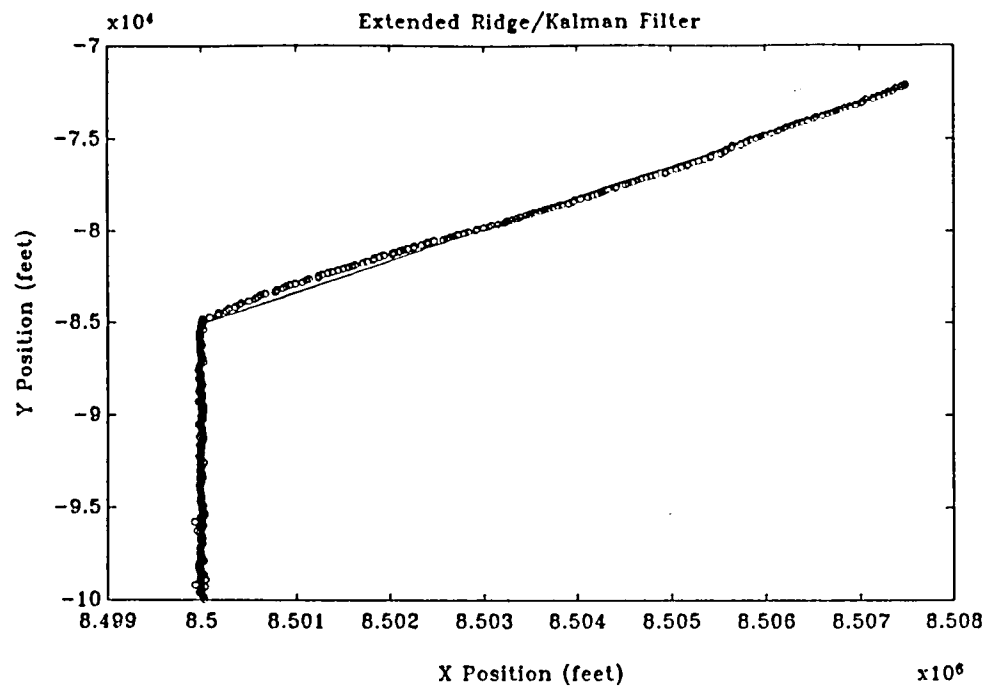Figure 4. Simulation results for an extended Kalman filter allowing for dynamics (large system covariance).

Figure 5. Simulation results for a Ridge/EKF filter.