Discrete-Time Kalman filter under incorrect noise covariances

Conference Paper · July 1995

DOI: 10.1109/ACC.1995.520929 · Source: IEEE Xplore

1 author:

Samer Saab
Lebanese American University

Proceedings of the
American Control Conference
Seattle, Washington • June 1995

Discrete-Time Kalman Filter under Incorrect Noise Covariances

Samer S. Saab
Union Switch & Signal
Pittsburgh, PA 15237
saab@switch.com

Abstract

The optimum filtering results of Kalman filtering for linear dynamic systems require an exact knowledge of the process noise covariance matrix Q, the measurement noise covariance matrix R and the initial error covariance matrix P_0. In a number of practical solutions, Q, R *and* P_0, are either unknown or are known only approximately. In this paper the sensitivity due to class of errors in the statistical modeling employing a Kalman Filter is discussed. In particular, we present a special case where it is shown that Kalman filter gains can be insensitive to scaling of covariance matrices. Some basic results are derived to describe the mutual relations among the three covariance matrices (actual and perturbed covariance matrices), their respective Kalman gain K_k and the error covariance matrices P_k. Experimental results using a tactical grade IMU are presented to illustrate the theoretical results.

1. Introduction

Consider a linear discrete-time invariant system, e.g. describing an error model, with state-space description

$$x_{k+1} = \Phi_k \delta x_k + w_k \qquad \text{(EQ 1)}$$

and the measurement equation as follows

$$z_k = H_k \delta x_k + v_k \qquad \text{(EQ 2)}$$

Let $\{w_k\}$ and $\{v_k\}$ be sequences of zero-mean Gaussian white noise such that $\mathrm{Var}(w_k) = Q$ is positive semi-definite matrix, $\mathrm{Var}(v_k) = R$ is positive definite matrix and $E(w_k v_l^T) = 0$ for all k and l. The initial error state δx_0 is also assumed independent of w_k and v_k in the sense that $E(\delta x_0 w_k^T) = 0$ and $E(\delta x_0 v_k^T) = 0$ for all k, and the initial error covariance matrix $P_0 = \mathrm{var}(\delta x_0)$.

Problem: Assume that Q, R *and* P_0 are perturbed, how would such perturbations affect the optimality of the filter?

Several results which deal with the deviation from the basic assumptions that guarantee optimality, for example non-Gaussian models of the errors, are presented in the literature, such as in [3-6] where the robustification of the filter is considered. There are some on-line identification schemes which identify Q and R from the innovation sequence, but their assumptions are rather restrictive and are not applicable for general systems [7]. Other stability considerations were presented in [10-11], where it is shown that incorrect values of the noise covariances *can* cause the filter to diverge; i.e., variance of a linear combination of the estimation error becomes unbounded. However, in [11], it has been shown that if the system is detectable, then the filter divergence will never occur. On the other hand, none of the work has addressed situations where the optimality of a Kalman filter is insensitive to statistical modeling errors. In this paper, we present a special case where the optimality of the filter is not destroyed.

Motivation:

The results of this paper constrain the uncertainties to be a scalar multiple of the ideal model; e.g., for the process noise covariance matrix, we have $Q' = \alpha Q$. This constraint appears to be quite restrictive. In addition, in Theorem 1, we have restricted $R = 0$, and $\bar{P}_0 = \alpha P_0 \geq 0$. Does this address a real world problem? Mathematically, we will never have ideal measurements, thus restricting $R = 0$ is impossible. On the other hand, if such constraints are weighted and approximated, then the results can be applied to a real problem. For example, a planar positioning system for transit vehicles employing a tachometer and one low grade yaw gyro, or an odometer system for trains, which employs one forward accelerometer and a low grade pitch gyro. One way to stabilize these systems from drifting, is to place passive transponders along the track, say one transponder every five miles, surveyed accurately with position and/or distance information. To correct the angle drifts from distance, speed or position information, one needs an observer such as an event-driven Kalman filter. Since the period between updates is in the order of one minute, then the drift due to the gyro errors is much larger than the measurement errors and initial errors. Again, the process noise covariance matrix is dominated by the gyro errors. When designing the error model for the gyro, then at least, the model will be off by a scale factor, say α. Therefore, if we set $\bar{P}_0 = \alpha P_0 \geq 0$ and/or $\bar{R} = \alpha R \geq 0$ and the accelerometer error model is scaled by α, say for $0 < \alpha < 10$, then such errors in modeling will not contribute to any significant change in the Kalman gain. In summary, if the error model is dominated by one error source, and the error model is not scaled correctly, then Theorem 1 implies that the optimality of the filter is insensitive to such scaling errors.

Assume that a LTI system is completely observable and controllable, then the steady state estimation using a Kalman filter will be independent of the initial error covariance matrix P_0. In this case, the choice of P_0 does not change the steady state of error covariance matrix P. On the other hand, if a system is not detectable, *and* not stabilizable, then one can *not* conclude that the steady state estimation is totally independent of the choice of P_0. In this paper, it is shown, for a particular case, that the steady state estimation is independent of the choice of the initial error covariance matrix even if the system is not completely observable. An application of the last statement is presented in Theorem 3. Another motive of this work came from error in experiments done in our labs, where the process noise covariance matrix Q was badly scaled, larger than its nominal value. The error covariance matrices generated by the algorithm showed errors which are consistent with the perturbed Q, but larger than the true errors of the state estimate. Theorem 1 in the next section solves this problem and in Section 3 experimental results verify it.

Moreover, it is shown for other cases where the state estimate may still be "optimum" whenever Q, R *and* P_0 are perturbed. The linear time invariance assumption forces the error covariance and Kalman gain updates to be independent of measurements and previous state updates. Therefore, if it can be shown that the Kalman gain remains unchanged under perturbations of these covariances, then all updates will remain unchanged, hence, optimum.

2. Main Results

The iterative algorithm of the error covariance matrices P_k^+ (= P_k), and P_k^-, and the Kalman gain K_k depend on the initial matrix P_0^+ (= P_0), and the process and measurement noise covariance matrices Q and R (sampled). Therefore, one should consider the change in P_k, P_k^-, and K_k due to the change in Q and R. The iterative algorithms of these matrices [2] are:

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

$$K_k = P_k^- H^T\left(HP_k^- H^T + R\right)^{-1} \qquad \text{(EQ 3)}$$

$$P_k = \left(I - K_k H\right)P_k^-$$

where Φ is the system matrix in the discrete domain, and H is the measurement matrix which are assumed to be constant since the system considered is LTI. Note that such algorithms are independent of the measurements and they are generated using the covariances matrices P_0, Q, R, system and measurement matrices. The analysis of the error propagation in P and/or K due to change in Q and R, which depend on such algorithms, tend to be messy even if in single output systems. For this reason, we will approach this problem using some practical approximations and assumptions on P_0, Q, and R.

Let A represent the error model state matrix, and B process noise coupling matrix, and Q_c the covariance matrix in the continuous domain. If we let Q represent the covariance matrix of the *sampled* process, then Q is given by [1] (page 82)

$$Q = \int_0^T e^{At} B Q_c B^T e^{A^T t} dt \qquad \text{(EQ 4)}$$

Note that if Q_c is positive definite and diagonal, then Q is always positive semi-definite and *not* necessarily diagonal, and note that if $Q_c' = \alpha Q_c$, then $Q' = \alpha Q$. In most applications, Q_c is symmetric and positive definite. If B is nonsingular (square) matrix and H is full row rank matrices, then Q is positive definite, and hence HQH^T is symmetric positive definite matrix ($\Rightarrow$ nonsingular). In general, the assumption that HQH^T is symmetric positive definite matrix is less restrictive than B being nonsingular.

In the event driven or multi-rate Kalman Filter the corresponding sampling period T can be large; e.g., for aiding a positioning system, one can use Map Matching updates, Differential Global Positioning System, zero-velocity updates[1] (ZUPTs) [8] or zero-sideslip measurements [9], etc. which can require a larger T. Clearly, since Q_c is positive definite, then as T increases, $\|Q\|$ increases. Note that for the zero-sideslip or zero-velocity updates the respective measurement noise covariance matrices can be modeled with arbitrary small variances. Therefore, in this case, we may approximate $HP_k^- H^T + R \approx HP_k^- H^T$; i.e., $R \cong 0$. The same is analogous to the size of P_0 or $\Phi P_0 \Phi^T$ with respect to Q; i.e., approximating $\Phi P_0 \Phi^T \cong 0$ whenever Q is sufficiently large.

In the following we summarize the results showing the effects of the change in the statistics of the initialization combined with the change in the covariance modeling of the process noise, measurement noise, and process and measurement noise. Consider the following assumptions:

(A1) HQH^T is positive definite matrix.

(A2) $R = 0$.

(A3) $\tilde{P}_0 = \alpha P_0 \geq 0$, where $\alpha \neq 0$, and P_0 can be equal to zero matrix.

Notice that (A1-A3) imply that $HP_k^- H^T$ is nonsingular since $H\Phi P_{k-1}\Phi^T H^T \geq 0, \forall k$.

Theorem 1: Let $\tilde{Q} = \alpha Q$, where $|\alpha| > 0$. If assumptions (A1)-(A3) are satisfied, then the Kalman algorithm will generate equal Kalman gains for all non zero α; i.e., the Kalman gain will be independent of the size of Q. Moreover, $\tilde{P}_k^- = \alpha P_k^-$ and $\tilde{P}_k = \alpha P_k$.

Proof.

The proof is simple and done by induction. Let $\tilde{P}_k^-$, $\tilde{P}_k$, $\tilde{K}_k$ correspond to $\tilde{P}_0$, $\tilde{Q}$ and P_k, P_k, K_k correspond to Q, P_0.

Thus $\tilde{P}_1^- = \alpha\left(\Phi P_0 \Phi^t + Q\right) = \alpha P_1^-$, next we find their corresponding gains

$$\tilde{K}_1 = \alpha P_1^- H^T\left(\alpha H P_1^- H^T\right)^{-1} =$$

$$P_1^- H^T\left(HP_1^- H^T\right)^{-1} = K_1 \qquad \text{(EQ 5)}$$

Thus, $P_1 = \left(I - K_1 H\right)P_1^-$ and

$\tilde{P}_1 = \alpha\left(I - K_1 H\right)P_1^- = \alpha P_1$.

We assume that the results are satisfied for k - 1 iterate, and prove that the results will be satisfied for k. It follows that

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

---

1. Such updates are used to partially calibrate and align an inertial navigation system whenever the vehicle is at rest. In this case, the measurements will be zero (velocity vector).

$$\tilde{P}_k^- = \Phi\tilde{P}_{k-1}\Phi^T + \tilde{Q} = \alpha\left(\Phi P_{k-1}\Phi^T + Q\right) = \alpha P_k^- \quad \text{(EQ 6)}$$

Now we are able to compute the Kalman gains,

$$K_k = P_k^- H^T\left(HP_k^- H^T\right)^{-1} \quad \text{(EQ 7)}$$

$$\tilde{K}_k = \alpha P_k^- H^T\left(\alpha HP_k^- H^T\right)^{-1} = K_k \quad \text{(EQ 8)}$$

Thus, the updated error covariance can be computed

$$P_k = \left(I - K_k H\right)P_k^- \quad \text{(EQ 9)}$$

$$\tilde{P}_k = \left(I - K_k H\right)\tilde{P}_k^- = \alpha P_k \quad \text{(EQ 10)}$$

Therefore, these relations are true for all k. Q.E.D.

Theorem 1 implies that, for one case where no initial errors $P_0 = 0$, no measurement errors $R = 0$, and whenever the scale of the covariance matrix of the process noise is *off*, then Kalman gain remains unchanged from its nominal value; i.e., the state estimate remains unchanged or "optimal". On the other hand, the error covariance matrices generated by the algorithm predict errors in the state estimate consistent with αQ and *not* the generated state estimates.

Theorem 2: If $\tilde{P}_0 = \alpha P_0$, $\tilde{Q} = \alpha Q$ and $\tilde{R} = \alpha R$, where $|\alpha| > 0$, then the Kalman algorithm will generate equal Kalman gains for all non zero α; i.e., the Kalman gain will be independent of the size of Q and R. Moreover, $\tilde{P}_k^- = \alpha P_k^-$ and $\tilde{P}_k = \alpha P_k$.

Proof.

The proof is done by induction (similar steps to the proof of Theorem 1). Let $\tilde{P}_k^-$, $\tilde{P}_k$, $\tilde{K}_k$ corresponds $\tilde{P}_0$, $\tilde{Q}$ & $\tilde{R}$, and P_k, P_k, K_k correspond to P_0, Q *and* R. Thus

$$\tilde{P}_1^- = \Phi\tilde{P}_0\Phi^T + \tilde{Q} = \alpha\left(\Phi P_0\Phi^T + Q\right) = \alpha P_1^- \quad \text{(EQ 11)}$$

Next, we find their corresponding gains

$$\tilde{K}_1 = \alpha P_1^- H^T\left(\alpha HP_1^- H^T + \alpha R\right)^{-1} =$$
$$H^T P_1^- \left(HP_1^- H^T + R\right)^{-1} = K_1 \quad \text{(EQ 12)}$$

Thus, $P_1 = \left(I - K_1 H\right)P_1^-$ and
$\tilde{P}_1 = \alpha\left(I - K_1 H\right)P_1^- = \alpha P_1$.

We assume that the results are satisfied for k - 1 iterate, and prove that the results will be satisfied for k. It follows

$$P_k^- = \Phi P_{k-1}\Phi^T + Q \quad \text{(EQ 13)}$$

$$\tilde{P}_k^- = \Phi\tilde{P}_{k-1}\Phi^T + \tilde{Q} = \alpha\left(\Phi P_{k-1}\Phi^T + Q\right) = \alpha P_k^-$$

Now we are able to compute the Kalman gains,

$$K_k = P_k^- H^T\left(HP_k^- H^T + R\right)^{-1} \quad \text{(EQ 14)}$$

$$\tilde{K}_k = \alpha P_k^- H^T\left(\alpha HP_k^- H^T + \alpha R\right)^{-1} = K_k \quad \text{(EQ 15)}$$

Thus, the updated error covariance can be computed

$$P_k = \left(I - K_k H\right)P_k^- \quad \text{(EQ 16)}$$

$$\tilde{P}_k = \left(I - K_k H\right)\tilde{P}_k^- = \alpha P_k \quad \text{(EQ 17)}$$

Therefore, these relations are true for all k. Q.E.D.

Theorem 3: If $\tilde{P}_0 = \alpha P_0$, $\tilde{R} = \alpha R > 0$, and $Q = 0$, where $|\alpha| > 0$, then the Kalman algorithm will generate equal Kalman gains for all non zero α; i.e., the Kalman gain will be independent of the size of R. Moreover, $\tilde{P}_k^- = \alpha P_k^-$ and $\tilde{P}_k = \alpha P_k$.

The proof of Theorem 3 is similar to the proves of Theorems 1 and 2, thus it can be omitted.

3. Experimental Applications

In this experiment, zero-velocity updates (ZUPTs) are used to partially align (pitch and roll) an INS presented in [8] for terranean vehicle application, and partially calibrate a six degree-of-freedom tactical grade Inertial Measurement Unit (IMU), three gyros and three accelerometers. The angular rates are sensed using subminiature quartz gyros, and the linear accelerations are sensed using servo accelerometers. The IMU is supplied by Systron Donner "MotionPak". The "earth-surface" frame navigation equations are derived based on the following assumptions:

(1) Neglect the earth angular rate.

(2) A constant gravity field applied.

(3) The inertial coordinate frame is chosen as a frame attached at the surface of earth. Also the pointing east-north-up frame is orientation of the earth-surface frame navigator.

The following is the dynamics of the earth-surface frame navigator

$$\dot{R}^I = V^I$$
$$\dot{V}^I = C_B^I \cdot A_{IB}^B - g^I$$
$$\dot{C}_B^I = C_B^I \cdot \Omega_{IB}^B$$

where

$$A_{IB}^B = \begin{bmatrix} a_x^b \\ a_y^b \\ a_z^b \end{bmatrix}, \Omega_{IB}^B = \begin{bmatrix} 0 & -g_z^b & g_y^b \\ g_z^b & 0 & -g_x^b \\ -g_y^b & g_x^b & 0 \end{bmatrix}, g^I = \begin{bmatrix} 0 \\ 0 \\ -g \end{bmatrix}$$

and

$$C_I^B = \begin{bmatrix} \cos\theta\cos\psi & \cos\theta\sin\psi & -\sin\theta \\ -\cos\phi\sin\psi + \sin\phi\sin\theta\cos\psi & \cos\psi\cos\phi + \sin\phi\sin\theta\sin\psi & \sin\phi\cos\theta \\ \sin\phi\sin\psi + \cos\phi\sin\theta\cos\psi & -\sin\phi\cos\psi + \cos\phi\sin\theta\sin\psi & \cos\phi\cos\theta \end{bmatrix}$$

and a^b and g^b are the accelerometer and gyro readings in the body frame respectively. The transformation matrix is initialized using the initial conditions of the roll (ϕ), pitch (θ) and yaw (ψ) angles of IMU frame with respect to the inertial reference frame, i.e., we rotate around the x-axis (ϕ), then rotate around the y-axis (θ), and finally around the z-axis (ψ).

An error model of the overall system (INS and IMU) with a system transformation which decouples the observable modes, are based on physical insight, from the unobservable modes are presented in details in [8]. The (multi-rate Kalman filter) recursive algorithm has been applied on the transformed error model using ZUPTs to estimate the tilt errors and the IMU constant biases. The output of the IMU is sampled at 20 Hz, Euler's method is used for integration. The data was collected for 300 seconds and fed to two Kalman filters, one with process noise covariance matrix based on IMU specifications "nominal" Q and the other filter uses $10*Q$ ($\alpha = 10$). The ZUPTs were triggered every 30 seconds, which corresponds to $\|Q\|_2 = 5.17e{+}3$ and the measurement noise standard deviation was set to 1.e-4 m/s which implies that $\|R\|_2 = 1.e{-}8$. The experimental results agreed closely to the results given in Theorem 1. Figures 1-6 show the error velocity outputs of the navigator, the expected velocity error standard deviations (VESS) estimated by the filter previous to each update (square root of the diagonal entries of P_k^-), and the velocity errors of the navigator previous to each update. Computed and estimated standard deviations of the velocity errors previous to ZUPTs are given in Table 1.

Table 1. Velocity error standard deviations (m/s)

VESS	true-x	est.-x	true-y	est.-y	true-z	est.-z
Q	0.443	0.515	1.968	0.708	0.020	0.007
$10Q$	0.447	5.154	1.984	7.089	0.020	0.076

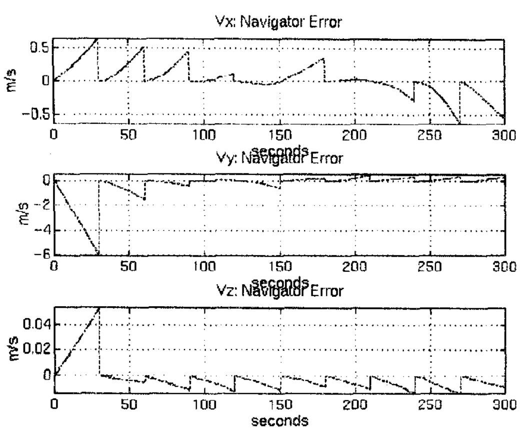

Figure 1. Continuous velocity errors using "nominal" Q

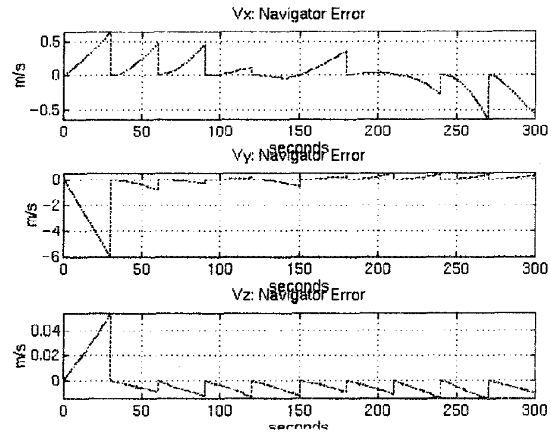

Figure 2. Continuous velocity errors using $10Q$

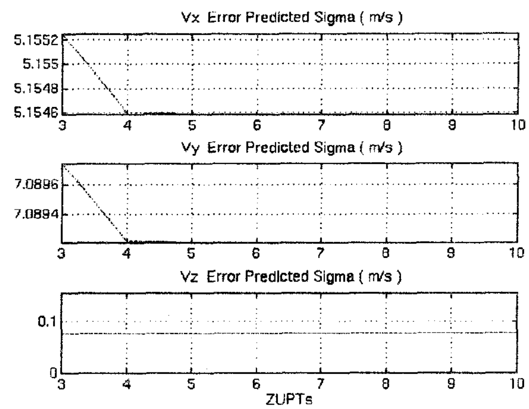

Figure 3. VESS predicted previous to ZUPTs using "nominal" Q

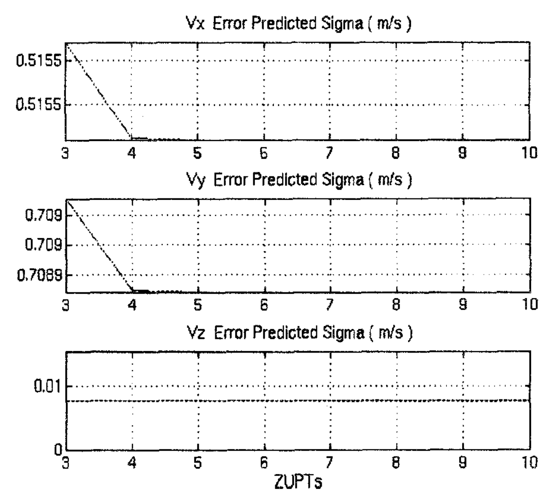

Figure 4. VESS predicted previous to ZUPTs using $10Q$

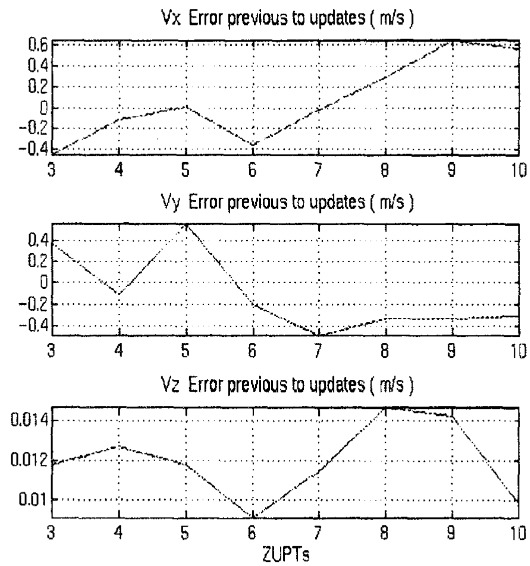

Figure 5. Velocity errors previous to ZUPTs using "nominal" Q

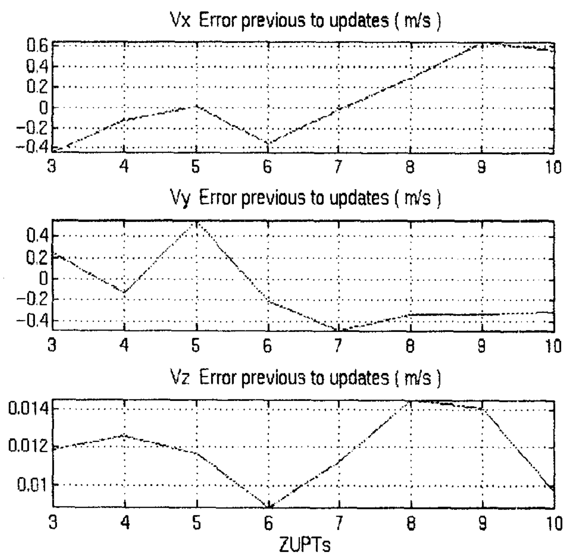

Figure 6. Velocity errors previous to ZUPTs using $10Q$

4. Conclusion

It was shown that Kalman filter gains can be insensitive to scaling of covariance matrices; i.e., the state estimate remains unchanged or "optimal" under incorrect noise covariances for a special case. Moreover, the error covariance matrices generated by the algorithm predict errors in the state estimate consistent with the scaled covariance matrices and *not* the generated state estimates. An integrated six degree-of-freedom inertial measurement unit was used coupled with a ZUPT Kalman filter which illustrated the theoretical results.

References

[1] F. Lewis, Optimal Estimation with an Introduction to Stochastic Control Theory, New York: Wiley.

[2] R. G. Brown and P. Y. C. Hwang, Introduction to Random Signals and Applied Kalman Filtering, New York: Wiley, 1992

[3] P. W. McBurney, "Robust Approach to reliable real-time Kalman filtering," *IEEE PLANS, Position Location and Navigation Symposium*, pp 549-556, Piscataway, NJ, 1990.

[4] S. Kawase and N. Yanagihara, "Robustness of the estimation with the Kalman filter," *IEEE Int. Symposium on Information Theory*, p 54, Piscataway, NJ, 1988.

[5] B. D. Kovacevic, S. S. Stankovic, and Z. M. Durovic, " Approaches to Robust Kalman Filtering," *Publications of the Faculty of Electrical Engineering, University of Belgrade: Automatic Control*, No. 1, pp 45-59, Belgrade Univ Yugoslavia, 1991.

[6] T. Cipra and R. Romera, " Robust Kalman Filter and its Application in Time Series Analysis, *Kybernetika*, Vol. 27, No. 6, pp 481-494, Prague, Czechoslovakia, 1991.

[7] K. M. Raman, "On the Identification of Variances and Adaptive Kalman Filtering," *IEEE Transactions on Automatic Control*, vol. AC-15, pp 175-184, April 1970.

[8] Samer S. Saab, and Kristjan T. Gunnarsson, "Automatic Alignment and Calibration of an Inertial Navigation System, "*IEEE Position Location and Navigation Symposium*, pp 845-852, April 1994.

[9] B. A. Kriegsman, and K. B. Mahar,"Gravity-Model Errors in Mobile Inertial Navigation Systems," *Journal of Guidance*, Vol 9, No. 3, pp 312-318, June 1986.

[10] S. Sangsuk-Iam, and T.E. Bullock, "Analysis of Discrete-Time Kalman Filtering under Incorrect Noise Covariances," *IEEE Transactions on Automatic Control*, vol. AC-35, pp 1304-1309, December 1990.

[11] S. Sungsuk-Iam, "Divergence of the Discrete-Time Kalman Filter under Incorrect Noise Covariances for Linear Periodic Systems," *Proceedings of the American Control Conference*, pp 1190-1194, Baltimore, Maryland, June 1994.