Investigation in extended kalman filter for indoor robot navigation

1st Yihui Chen
National University of Singapore
e1010473@u.nus.edu

***Abstract*—This document is a model and instructions for LaTeX. This and the IEEEtran.cls file define the components of your paper [title, text, heads, etc.]. *CRITICAL: Do Not Use Symbols, Special Characters, Footnotes, or Math in Paper Title or Abstract.**

***Index Terms*—component, formatting, style, styling, insert**

I. Literature review

II. Bayesian filter

A. Bayes theorem in discrete and continuous cases

Bayes theorem describes the probability of an event based on prior knowledge. Suppose that random variable X denote the event to be estimated while Y represent the observation result, x and y are their specific values respectively. Consdier Bayes rule in discrete cases and get

$$P(X = x \mid Y = y) = \frac{P(Y = y \mid X = x)P(X = x)}{P(Y = y)} \quad (1)$$

where $P(Y = y \mid X = x)$ is likelyhood that often describes the measurement accuracy of sensor, and $P(X = x)$ denotes the prior probability based on current knowledge. Using the law of total probability, $P(Y = y)$ on the denominator can be expanded as $P(Y = y) = \sum_{i=1}^{n} P(Y = y \mid X = x_i)P(X = x_i)$, which shows that its value is not related to the value of Y and thus it can be replaced by a constant η. Then Bayes theorem combines them and gives the posterior probability.

Now consider the Bayes fomula for continuous random variables

$$P(X < x \mid Y = y) = \frac{P(Y = y \mid X < x)P(X < x)}{P(Y = y)} \quad (2)$$

Since the fomula can not be directly used, the right side is written as a sum of discrete probabilities and turned into a limit form.

$$RHS = \sum_{u=-\infty}^{x} \frac{P(Y = y \mid X = u)P(X = u)}{P(Y = y)}$$
$$= \lim_{\varepsilon \to 0} \sum_{u=-\infty}^{x} \frac{P(y < Y < y + \varepsilon \mid X = u)P(u < X < u + \varepsilon)}{P(y < Y < y + \varepsilon)}$$

Let $f(\cdot)$ denotes the probability density function and apply mean value theorem, we derive

$$\begin{aligned} RHS &= \lim_{\varepsilon_1,\varepsilon_2,\varepsilon_3 \to 0} \sum_{u=-\infty}^{x} \frac{(f_{Y|X}(\xi_1 \mid u))(f_X(\xi_2)) \cdot \varepsilon_1\varepsilon_2}{f_Y(\xi_3) \cdot \varepsilon_3} \\ &= \lim_{\varepsilon \to 0} \sum_{u=-\infty}^{x} \frac{f_{Y|X}(y \mid u)f_X(u)}{f_Y(y)} \cdot \varepsilon \\ &= \int_{-\infty}^{x} \frac{f_{Y|X}(y \mid x)f_X(x)}{f_Y(y)} dx \end{aligned}$$

where $\xi_1 \in (y, y+\varepsilon_1)$, $\xi_2 \in (u, u+\varepsilon_2)$ and $\xi_3 \in (y, y+\varepsilon_3)$. Also write the left side of Eq. 2 in a form of probability density function, i.e. $LHS = \int_{-\infty}^{x} f_{X|Y}(x \mid y)dx$, we finally get the Bayes formula for pdf in continuous cases.

$$f_{X|Y}(x \mid y) = \frac{f_{Y|X}(y \mid x)f_X(x)}{f_Y(y)} \quad (3)$$

Similarily, the denominator can be replace by a constant η, where $\eta = [\int_{-\infty}^{+\infty} f_{Y|X}(y \mid x)f_X(x)dx]^{-1}$.

B. Bayesian filtering algorithm

Suppose we have established the prediction and observation equation as

$$\begin{aligned} X_k &= F(X_{k-1}) + Q_k \\ Y_k &= H(X_k) + R_k \end{aligned} \quad (4)$$

where X_k, Y_k, Q_k, R_k are all random variables, and X_0, Q_k and R_k are mutually independent. In the prediction step, pdf of prior estimation $f_k^-(x)$ is firstly calculated.

$$f_k^-(x) = \frac{d}{dx}(P(X_k < x)) = \frac{d}{dx}(\sum_{u=-\infty}^{x} P(X_k = u)) \quad (5)$$

Expand $P(X_k = u)$ by total probability law and substitute the prediction function into it, we get

$$P(X_k = u) = \sum_{v=-\infty}^{+\infty} P(X_k = u \mid X_{k-1} = v)P(X_{k-1} = v)$$
$$= \sum_{v=-\infty}^{+\infty} P(X_k - F(X_{k-1}) = u - F(v) \mid X_{k-1} = v)P(X_{k-1} = v)$$
$$= \sum_{v=-\infty}^{+\infty} P(Q_k = u - F(v))P(X_{k-1} = v)$$

Repeat what we do on RHS of Eq. 2, it becomes

$$\begin{aligned}P(X_k = u) &= \lim_{\varepsilon\to 0} \sum_{v=-\infty}^{+\infty} f_{Q_k}[u - F(v)] f_{k-1}(v)\cdot\varepsilon^2 \\ &= \lim_{\varepsilon\to 0}\int_{-\infty}^{+\infty} f_{Q_k}[u - F(v)] f_{k-1}(v) dv\cdot\varepsilon\end{aligned}$$

Substitute it into Eq. 5 and get the pdf of prior estimation

$$\begin{aligned}f_k^-(x) &= \frac{d}{dx}\left\{\textstyle\sum_{-\infty}^{x} \lim_{\varepsilon\to 0}\int_{-\infty}^{+\infty} f_{Q_k}[u - F(v)] f_{k-1}(v) dv\cdot\varepsilon\right\} \\ &= \frac{d}{dx}\left\{\textstyle\int_{-\infty}^{x}\int_{-\infty}^{+\infty} f_{Q_k}[u - F(v)] f_{k-1}(v) dv du)\right\} \\ &= \int_{-\infty}^{+\infty} f_{Q_k}[x - F(v)] f_{k-1}(v) dv\end{aligned} \tag{6}$$

After the observed value $Y_k = y_k$ is acquired by the sensor, the pdf of likelyhood is derived as

$$\begin{aligned}f_{Y_k|X_k}(y_k \mid x) &= \lim_{\varepsilon\to 0}\frac{P(y_k < Y_k < y_k + \varepsilon \mid X_k = x)}{\varepsilon} \\ &= \lim_{\varepsilon\to 0}\frac{P(y_k - H(x) < R_k < y_k - H(x) + \varepsilon)}{\varepsilon} \\ &= f_{R_k}[y_k - H(x)]\end{aligned} \tag{7}$$

Substitute Eq. 6 and Eq. 7 into Eq. 3, we get the pdf of posterior estimation:

$$f_k^+(x) = \eta_k\cdot f_{R_k}[y_k - H(x)]\cdot f_k^-(x) \tag{8}$$

where $\eta_k = (\int_{-\infty}^{+\infty} f_{R_k}[y_k - H(x)]\cdot f_k^-(x)dx)^{-1}$. With the prediction and update steps, the current state of robot can be estimated realtime by Bayes formula. The algorithm flow is given below.

Algorithm 1 Bayesian filter

INITIALIZE $f_0^+(x), Q, R$
for i=1,...,r **do**
  Predict Step $f_i^-(x) = \int_{-\infty}^{+\infty} f_{Q_i}[x - F(v)] f_{i-1}^+(v)dv$
  Update Step $f_i^+(x) = \eta_i\cdot f_{R_i}[y_i - H(x)] f_i^-(x)$
  Estimate State $\hat{x}_i^+ = \int_{-\infty}^{+\infty} x f_i^+(x)dx$
end for

III. Kalman filter

A. Kalman filter

Bayes filter estimates unknown probability density function recursively using a mathematical process model and external measurements, laying the foundation for realtime state estimation. However, it is usually hard to get analytical solutions when calculating improper integrals in the algorithm steps. Therefore, Kalman filter made some assumptions to simplify the process. Consider the prediction and observation equations in Eq. 4, suppose that both $F(\cdot)$ and $H(\cdot)$ are linear functions, and Q_k, R_k follow the normal distribution, i.e.

$$\begin{aligned}X_k &= F X_{k-1} + Q_k \\ Y_k &= H X_k + R_k \\ Q_k \sim N(0, Q)&, R_k \sim N(0, R)\end{aligned} \tag{9}$$

Then if the random variable at the last time also follows a normal distribution, i.e. $X_{k-1}^+ \sim N(\mu_{k-1}^+, \sigma_{k-1}^+)$, we can easily infer the pdf of prior probability is a normal distribution by convolution theorem.

$$\begin{aligned}&F X_{k-1}^+ \sim N(F\mu_{k-1}^+, F^2\sigma_{k-1}^+), Q_k \sim N(0, Q) \\ &\Rightarrow X_k^- \sim N(F\mu_{k-1}^+, F^2\sigma_{k-1}^+ + Q)\end{aligned} \tag{10}$$

Denote the mean and variance of X_k^- as μ_k^- and σ_k^-, the pdf of posterior probability $f_k^+(x)$ and estimated state $\hat{x}_k^+$ can be derived following the update rule. Then the random variable X_k^+ at the current time also follows the normal distribution, i.e. $X_k^+ \sim N(\mu_k^+, \sigma_k^+)$, where $\hat{x}_k^+ = \mu_k^+$ is the estimated state and they are given by

$$\begin{aligned}\mu_k^+ &= \mu_k^- + K(y_k - H\mu_k^-) \\ \sigma_k^+ &= (1 - KH)\sigma_k^- \\ K &= \frac{H\sigma_k^-}{H^2\sigma_k^- + R}\end{aligned} \tag{11}$$

Since the state variables are usually vectors, kalman filter algorithm can be generalized to a matrix form as below, where $\boldsymbol{F}$ and $\boldsymbol{H}$ are matrixes in prediction and observation equatons, and $\boldsymbol{\Sigma_k}$ is the covariance matrix.

Algorithm 2 Kalman filter

Prediction equation $\boldsymbol{X_k = F X_{k-1} + Q_k}$
Observation equation $\boldsymbol{Y_k = H X_k + R_k}$
$\boldsymbol{Q_k} \sim N(\boldsymbol{0}, \boldsymbol{Q})$, $\boldsymbol{R_k} \sim N(\boldsymbol{0}, \boldsymbol{R})$, $\boldsymbol{X_0} \sim N(\boldsymbol{\mu_0^+}, \boldsymbol{\Sigma_0^+})$
Initialize $\boldsymbol{\mu_0^+, \Sigma_0^+, Q, R}$
for k=1,...,r **do**
  $\boldsymbol{\mu_k^- = F\mu_{k-1}^+}$
  $\boldsymbol{\Sigma_k^- = F\Sigma_{k-1}^+ F}^T \boldsymbol{+ Q}$ {prediction end}
  $\boldsymbol{K = \Sigma_k^- H}^T\boldsymbol{(H\Sigma_k^- H}^T \boldsymbol{+ R)}^{-1}$ {the kalman gain}
  Obtain an observation value $\boldsymbol{y_k}$
  $\boldsymbol{\mu_k^+ = \mu_k^- + K(y_k - H\mu_k^-)}^{-1}$
  $\boldsymbol{\Sigma_k^+ = (I - KH)\Sigma_k^-}$ {update end}
  $\boldsymbol{\hat{x}_k^+ = \mu_k^+}$
end for

B. Extended Kalman filter

Prediction and observation functions are usually nonlinear in real cases, but they are simply supposed to be linear in Kalman filter, which leads to bad performance. Extended Kalman filter improve it by linearizing them by Taylor series. We still suppose that $X_{k-1}^+ \sim N(\mu_{k-1}^+, \sigma_{k-1}^+)$, and expand $F(X_{k-1}^+)$ by its first-order Taylor series about μ_{k-1}^+ as

$$F(X_{k-1}^+) \approx F(\mu_{k-1}^+) + F'(\mu_{k-1}^+)(X_{k-1}^+ - \mu_{k-1}^+) \\ \approx AX_{k-1}^+ + B \tag{12}$$

where $A = F'(\mu_{k-1}^+)$ and $B = F(\mu_{k-1}^+) - F'(\mu_{k-1}^+)\mu_{k-1}^+$. Similarly, we follow the prediction step in Bayes filter algorithm and find X_k^- also follows the normal distribution, i.e. $X_k^- \sim N(A\mu_{k-1}^+ + B, A^2\sigma_{k-1}^+ + Q)$. Substitute A and B into it and it becomes

$$X_k^- \sim N(F(\mu_{k-1}^+), A^2\sigma_{k-1}^+ + Q) \tag{13}$$

Denote the mean and variance in Eq. 13 are μ_k^- and σ_k^- respectively, we calculate the pdf of posterior probability $f_k^+(x)$ by the update step in Bayes filter. Similarily, we expand the nonlinear observation function $H(X_k)$ by Taylor series about μ_k^- as

$$H(X_k) \approx H(\mu_k^-) + H'(\mu_k^-)(X_k - \mu_k^-) \\ \approx CX_k + D \tag{14}$$

where $C = H'(\mu_k^-)$ and $D = H(\mu_k^-) - H'(\mu_k^-)\mu_k^-$. Following the update step we find the pdf of posterior probability is also in shape of normal distribution. Therefore, the random variable after update follows

$$X_k^+ \sim N(\frac{R\mu_k^- + C\sigma_k^-(y_k - D)}{R + C^2\sigma_k^-}, (1 - \frac{C^2\sigma_k^-}{R + C^2\sigma_k^-})\sigma_k^-) \\ \sim N(\mu_k^- + K(y_k - H(\mu_k^-)), (1 - KC)\sigma_k^-) \tag{15}$$

where $K = \frac{C\sigma_k^-}{R + C^2\sigma_k^-}$. Then we get our estimated state $x_k^+ = \mu_k^+$. We can also generate the extended Kalman filter into matrix case. Different from the scalar case, here $\boldsymbol{Q}$, $\boldsymbol{R}$, $\boldsymbol{\Sigma_k}$ are all covariance matrixes, and $\boldsymbol{A}_{n\times n}$, $\boldsymbol{C}_{m\times n}$ (m observers and n state variables) need to be calculated in each loop:

$$\boldsymbol{A} = \begin{pmatrix} \frac{\partial F_1}{\partial X_{k-1}^1} & \frac{\partial F_1}{\partial X_{k-1}^2} & \cdots & \frac{\partial F_1}{\partial X_{k-1}^n} \\ \frac{\partial F_2}{\partial X_{k-1}^1} & \frac{\partial F_2}{\partial X_{k-1}^2} & \cdots & \frac{\partial F_2}{\partial X_{k-1}^n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial F_n}{\partial X_{k-1}^1} & \frac{\partial F_n}{\partial X_{k-1}^2} & \cdots & \frac{\partial F_n}{\partial X_{k-1}^n} \end{pmatrix}_{|X_{k-1} = \hat{x}_{k-1}^+ = \mu_{k-1}^+} \tag{16}$$

$$\boldsymbol{C} = \begin{pmatrix} \frac{\partial H_1}{\partial X_k^1} & \frac{\partial H_1}{\partial X_k^2} & \cdots & \frac{\partial H_1}{\partial X_k^2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial H_m}{\partial X_k^1} & \frac{\partial H_m}{\partial X_k^2} & \cdots & \frac{\partial H_m}{\partial X_k^n} \end{pmatrix}_{|X_k = \hat{x}_k^- = \mu_k^-} \tag{17}$$

The full flow of EKF algorithm is shown below.

Algorithm 3 Extended Kalman filter

Prediction equation $\boldsymbol{X_k} = \boldsymbol{F}(\boldsymbol{X_{k-1}}) + \boldsymbol{Q_k}$
Observation equation $\boldsymbol{Y_k} = \boldsymbol{H}(\boldsymbol{X_k}) + \boldsymbol{R_k}$
$\boldsymbol{Q_k} \sim N(\boldsymbol{0}, \boldsymbol{Q})$, $\boldsymbol{R_k} \sim N(\boldsymbol{0}, \boldsymbol{R})$, $\boldsymbol{X_0} \sim N(\boldsymbol{\mu_0^+}, \boldsymbol{\Sigma_0^+})$
Initialize $\boldsymbol{\mu_0^+}, \boldsymbol{\Sigma_0^+}, \boldsymbol{Q}, \boldsymbol{R}$
for k=1,...,r **do**
  Calculate $\boldsymbol{A}$ using Eq. 16
  $\boldsymbol{\mu_k^-} = \boldsymbol{F}(\boldsymbol{\mu_{k-1}^+})$
  $\boldsymbol{\Sigma_k^-} = \boldsymbol{A}\boldsymbol{\Sigma_{k-1}^+}\boldsymbol{A}^T + \boldsymbol{Q}$ {prediction end}
  Calculate $\boldsymbol{C}$ using Eq. 17
  $\boldsymbol{K} = \boldsymbol{\Sigma_k^-}\boldsymbol{C}^T(\boldsymbol{C}\boldsymbol{\Sigma_k^-}\boldsymbol{C}^T + \boldsymbol{R})^{-1}$ {the kalman gain}
  Obtain an observation value $\boldsymbol{y_k}$
  $\boldsymbol{\mu_k^+} = \boldsymbol{\mu_k^-} + \boldsymbol{K}[\boldsymbol{y_k} - \boldsymbol{H}(\boldsymbol{\mu_k^-})]$
  $\boldsymbol{\Sigma_k^+} = (\boldsymbol{I} - \boldsymbol{K}\boldsymbol{C})\boldsymbol{\Sigma_k^-}$ {update end}
  $\boldsymbol{\hat{x}_k^+} = \boldsymbol{\mu_k^+}$
end for

C. Test results

To test the performance of these filters, we create a signal $x_1(t) = t^2$ and suppose the sensor noise follows the gaussian distribution, i.e. $\omega_1 \sim N(0, 0.1)$. First we use Kalman filter and establish prediction and observation equations. Since X_k can be expanded about X_{k-1} as $X_k = X_{k-1} + X_{k-1}'dt + \frac{1}{2}X_{k-1}''(dt)^2$, we choose $\begin{pmatrix} X_{k-1} & X_{k-1}' & X_{k-1}'' \end{pmatrix}^T$ and the two equations becomes

$$\begin{pmatrix} X_k \\ X_k' \\ X_k'' \end{pmatrix} = \begin{pmatrix} 1 & dt & \frac{1}{2}(dt)^2 \\ 0 & 1 & dt \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X_{k-1} \\ X_{k-1}' \\ X_{k-1}'' \end{pmatrix} + Q_k \tag{18}$$

$$Y_k = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} X_k \\ X_k' \\ X_k'' \end{pmatrix} + R_k \tag{19}$$

where $Q_k \sim N(0, Q)$, $R_k \sim N(0, R)$. The prediction for the higher derivative is considered more accurate and sensor noise is a little bit large, so here let $Q = diag\begin{Bmatrix} 1 & 0.01 & 0.0001 \end{Bmatrix}$ and $R = 20$. Also, $\boldsymbol{\mu_0^+}$ and $\boldsymbol{\Sigma_0^+}$ are initialized as $\begin{pmatrix} 0.01 & 0 & 0 \end{pmatrix}^T$ and $diag\begin{Bmatrix} 0.01 & 0.01 & 0.0001 \end{Bmatrix}$ respectively. As shown in Fig. 1, Kalman filter cut down the error between observation and real signals.

Now we add another sensor with higher measurement accuracy and modify the observation equation (Eq. 20), where we tend to trust more on the second sensor and let $R = diag\begin{Bmatrix} 3 & 5 \end{Bmatrix}$. The filtering result after sensor fusion is illustrated in Fig. 2, where the estimation error is further reduced.

$$\begin{pmatrix} Y_{k_1} \\ Y_{k_2} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} X_k \\ X_k' \\ X_k'' \end{pmatrix} + R_k \tag{20}$$

Then we choose another signal with stronger nonlinearity. The real signal $x_2(t)$ follows that $x_2(t+1) = sin(3x_2(t))$ and $x_2(0) = 0.1$. We still use the model in Eq. 18 and Eq. 19, and the simulation result is shown in Fig. 3. The limitation of

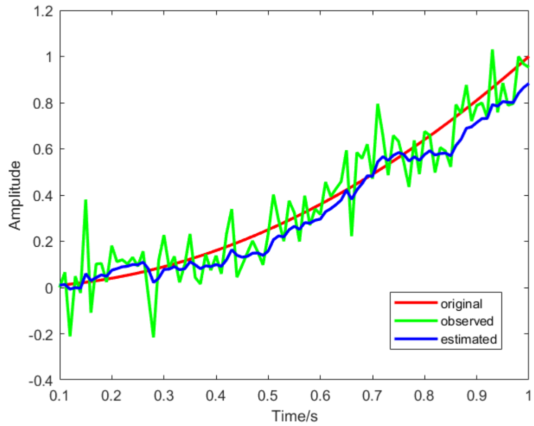

Fig. 1. Estimation for $x_1(t) = t^2$ using Kalman filter with one sensor

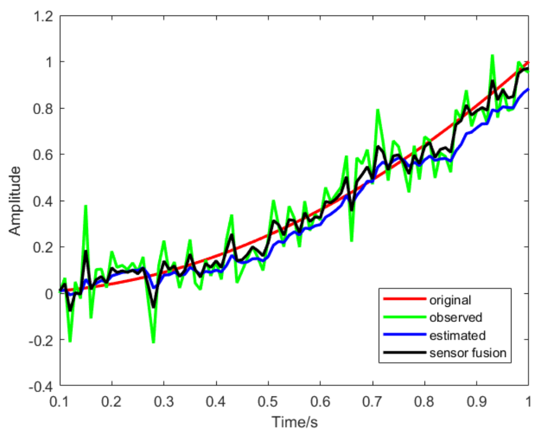

Fig. 2. Estimation for $x_1(t) = t^2$ using Kalman filter after sensor fusion

Kalman is shown that it cannot track the nonlinear signal and result in a large estimation error.

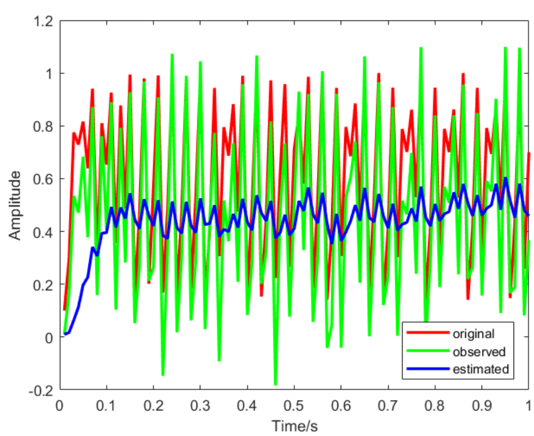

Fig. 3. Estimation for $x_2(t)$ using Kalman filter

Instead, we use the extended Kalman filter to process the same signal. The prediction and observation equation is given in Eq. 21, where $Q_k \sim N(0, 0.0001)$ and $R_k \sim N(0, 1)$. Suppose the initial state $X_0 \sim N(0.1, 0.1)$, the estimation performance is plot in Fig. 4, where tracking error is much smaller than that using Kalman filter.

$$
\begin{aligned}
X_k &= sin(3X_{k-1}) + Q_k \\
Y_k &= X_k^2 + R_k
\end{aligned} \tag{21}
$$

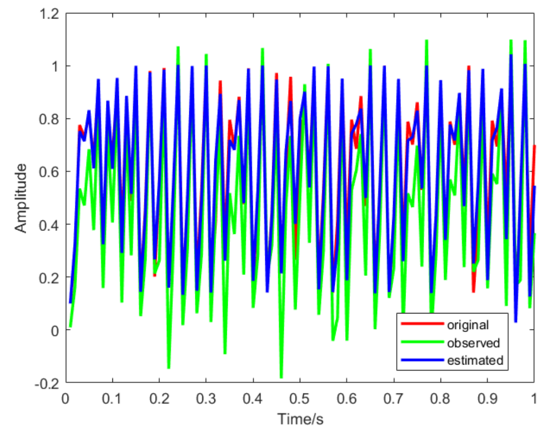

Fig. 4. Estimation for $x_2(t)$ using extended Kalman filter

IV. Particle filters

Instead, we use the extended Kalman filter to process the same signal. The prediction and observation equation is given in Eq. 21, where $Q_k \sim N(0, 0.0001)$ and $R_k \sim N(0, 1)$. Suppose the initial state $X_0 \sim N(0.1, 0.1)$, the estimation performance is plot in Fig. 4, where tracking error is much smaller than that using Kalman filter.

Acknowledgment

The preferred spelling of the word "acknowledgment" in America is without an "e" after the "g". Avoid the stilted expression "one of us (R. B. G.) thanks …". Instead, try "R. B. G. thanks…". Put sponsor acknowledgments in the unnumbered footnote on the first page.