7. Nonlinear Estimation

7.1 The Extended Kalman Filter (EKF)

Consider a nonlinear, time-invariant state model of the form

$$\underline{\mathrm{x}}(n+1) = \underline{\phi}(\underline{\mathrm{x}}(n)) + \Gamma \underline{\mathrm{w}}(n), \quad \underline{\mathrm{w}}(n) \sim N(\underline{0}, Q(n)) \tag{7.1-1}$$

$$\underline{\mathrm{z}}(n) = \underline{\gamma}(\underline{\mathrm{x}}(n)) + \underline{\mathrm{v}}(n), \quad \underline{\mathrm{v}}(n) \sim N(\underline{0}, R(n)) \tag{7.1-2}$$

$$E\left\{\underline{\mathrm{w}}(n)\underline{\mathrm{v}}^{\mathrm{T}}(n)\right\} = 0,\ E\left\{\underline{\mathrm{w}}(n)\underline{\mathrm{x}}^{\mathrm{T}}(0)\right\} = 0,\ E\left\{\underline{\mathrm{v}}(n)\underline{\mathrm{x}}^{\mathrm{T}}(0)\right\} = 0.$$

Assume that $\underline{\phi}(\cdot)$ and $\underline{\gamma}(\cdot)$ are sufficiently smooth in $\underline{\mathrm{x}}$ so that each has a valid Taylor series expansion. Given a realization $\hat{\underline{\mathrm{x}}}(n)$, expand $\underline{\phi}(\cdot)$ into a Taylor series about $\hat{\underline{\mathrm{x}}}(n)$

$$\begin{aligned}\underline{\phi}(\underline{\mathrm{x}}(n)) &= \underline{\phi}(\hat{\underline{\mathrm{x}}}(n)) + J_{\underline{\phi}}(\hat{\underline{\mathrm{x}}}(n))\left[\underline{\mathrm{x}}(n) - \hat{\underline{\mathrm{x}}}(n)\right] + \cdots \\ &\triangleq \underline{\phi}(\hat{\underline{\mathrm{x}}}(n)) + \Phi(n)\left[\underline{\mathrm{x}}(n) - \hat{\underline{\mathrm{x}}}(n)\right] + \cdots\end{aligned} \tag{7.1-3}$$

where $J_{\underline{\phi}}(\underline{x})$ is the Jacobian of $\underline{\phi}(\cdot)$ evaluated at $\underline{x}$ such that

$$J_{\underline{\phi}}(\underline{x}) = \frac{\partial \underline{\phi}(\underline{x})}{\partial \underline{x}} = \begin{bmatrix} \frac{\partial \phi_1(\underline{x})}{\partial x_1} & \frac{\partial \phi_1(\underline{x})}{\partial x_2} & \cdots & \frac{\partial \phi_1(\underline{x})}{\partial x_n} \\ \frac{\partial \phi_2(\underline{x})}{\partial x_1} & \frac{\partial \phi_2(\underline{x})}{\partial x_2} & \cdots & \frac{\partial \phi_2(\underline{x})}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \phi_m(\underline{x})}{\partial x_1} & \frac{\partial \phi_m(\underline{x})}{\partial x_2} & \cdots & \frac{\partial \phi_m(\underline{x})}{\partial x_n} \end{bmatrix}. \tag{7.1-4}$$

Likewise, we expand $\underline{\gamma}(\cdot)$ about the realization $\hat{\underline{\mathrm{x}}}^-(n)$

$$\begin{aligned} \underline{\gamma}(\underline{\mathrm{x}}(n)) &= \underline{\gamma}(\hat{\underline{\mathrm{x}}}^-(n)) + J_{\underline{\gamma}}(\hat{\underline{\mathrm{x}}}^-(n))\left[\underline{\mathrm{x}}(n) - \hat{\underline{\mathrm{x}}}^-(n)\right] + \cdots \\ &\triangleq \underline{\gamma}(\hat{\underline{\mathrm{x}}}^-(n)) + H(n)\left[\underline{\mathrm{x}}(n) - \hat{\underline{\mathrm{x}}}^-(n)\right] + \cdots \end{aligned} \tag{7.1-5}$$

Keeping only the first two terms in the expansions of Eqs. (7.1-3) and (7.1-5), we have a linearized signal/measurement model

$$\underline{\mathrm{x}}(n+1) = \underline{\phi}(\hat{\underline{\mathrm{x}}}(n)) + \Phi(n)\left[\underline{\mathrm{x}}(n) - \hat{\underline{\mathrm{x}}}(n)\right] + \Gamma \underline{\mathrm{w}}(n) \tag{7.1-6}$$

$$\underline{\mathrm{z}}(n) = \underline{\gamma}(\hat{\underline{\mathrm{x}}}^-(n)) + H(n)\left[\underline{\mathrm{x}}(n) - \hat{\underline{\mathrm{x}}}^-(n)\right] + \underline{\mathrm{v}}(n)\,. \tag{7.1-7}$$

Time Propagation – Assume that $\underline{\hat{x}}(n-1)$ is unbiased and seek the a priori estimate $\underline{\hat{x}}^{-}(n)$. We want $\underline{\hat{x}}^{-}(n)$ to be unbiased, so it must satisfy

$$E\left\{\left[\underline{x}(n)-\underline{\hat{x}}^{-}(n)\right]\middle|Z\right\}=0,$$

which yields to

$$\begin{aligned}\underline{\hat{x}}^{-}(n)&=E\left\{\underline{x}(n)\middle|Z\right\}\\&=E\left\{\underline{\phi}(\underline{\hat{x}}(n-1))+\Phi(n-1)\left[\underline{x}(n-1)-\underline{\hat{x}}(n-1)\right]+\Gamma\underline{w}(n-1)\middle|Z\right\}\\&=E\left\{\underline{\phi}(\underline{\hat{x}}(n-1))+\Phi(n-1)\underline{\tilde{x}}(n-1)+\Gamma\underline{w}(n-1)\middle|Z\right\}\\&=\underline{\phi}(\underline{\hat{x}}(n-1)).\end{aligned} \tag{7.1-8}$$

Now, seek the time propagation equation for the a priori conditional error covariance $P^{-}(n)$

$$\begin{aligned}P^{-}(n)&=Cov\left\{\left[\underline{x}(n)-\underline{\hat{x}}^{-}(n)\right]\middle|Z\right\}\\&=Cov\left\{\left[\underline{\phi}(\underline{\hat{x}}(n-1))+\Phi(n-1)\underline{\tilde{x}}(n-1)+\Gamma\underline{w}(n-1)-\underline{\hat{x}}^{-}(n)\right]\middle|Z\right\}\\&=\Phi(n-1)P(n-1)\Phi^{T}(n-1)+\Gamma Q(n-1)\Gamma^{T}.\end{aligned} \tag{7.1-9}$$

Measurement Update – Assume that $\underline{\hat{x}}^-(n)$ is unbiased and the error $\underline{\tilde{x}}^-(n)$ is orthogonal to the measurements $\underline{z}(1),\cdots,\underline{z}(n-1)$. (Refer to Eq. (7.1-8).) Given the new measurement $\underline{z}(n)$, we seek the minimum mean square error estimate $\underline{\hat{x}}(n)$. Assume that $\underline{\hat{x}}(n)$ has the form

$$\underline{\hat{x}}(n) = \underline{b}(n) + K(n)\underline{z}(n). \qquad (7.1\text{-}10)$$

Since $\underline{\hat{x}}(n)$ should be unbiased, we have

$$E\left\{\left[\underline{x}(n) - \underline{\hat{x}}(n)\right] \middle| Z\right\} = 0.$$

Substitute Eqs. (7.1-7) and (7.1-10) into this expression to obtain

$$E\left\{\left[\underline{x}(n) - \underline{b}(n) - K(n)\left(\underline{\gamma}(\underline{\hat{x}}^-(n)) + H(n)\underline{\tilde{x}}^-(n) + \underline{v}(n)\right)\right] \middle| Z\right\} = 0.$$

Solve this equation for $\underline{b}(n)$

$$\begin{aligned}
\underline{b}(n) &= -K(n)\underline{\gamma}(\underline{\hat{x}}^-(n)) - K(n)H(n)E\left\{\underline{\tilde{x}}^-(n)\middle|Z\right\} + E\left\{\underline{x}(n)\middle|Z\right\} - K(n)E\left\{\underline{v}(n)\middle|Z\right\} \\
&= -K(n)\underline{\gamma}(\underline{\hat{x}}^-(n)) + E\left\{\underline{x}(n)\middle|Z\right\} \\
&= -K(n)\underline{\gamma}(\underline{\hat{x}}^-(n)) + \underline{\hat{x}}^-(n). \text{ (according to Eq. (7.1-8).)}
\end{aligned}$$

Therefore, Eq. (7.1-10) becomes

$$\underline{\hat{\mathrm{x}}}(n) = \underline{\hat{\mathrm{x}}}^-(n) + K(n)\left[\underline{\mathrm{z}}(n) - \underline{\gamma}(\underline{\hat{\mathrm{x}}}^-(n))\right]. \tag{7.1-11}$$

Now, we want to find $K(n)$. By the orthogonality principle, the following equations hold

$$\begin{aligned}
&E\left\{\left[\underline{\mathrm{x}}(n) - \underline{\hat{\mathrm{x}}}(n)\right]\underline{\mathrm{z}}^{\mathrm{T}}(i)\middle|Z\right\} \\
&= E\left\{\left[\underline{\mathrm{x}}(n) - \underline{\hat{\mathrm{x}}}^-(n) - K(n)\left[\underline{\mathrm{z}}(n) - \underline{\gamma}(\underline{\hat{\mathrm{x}}}^-(n))\right]\right]\underline{\mathrm{z}}^{\mathrm{T}}(i)\middle|Z\right\} \\
&= E\left\{\left[\underline{\tilde{\mathrm{x}}}^-(n) - K(n)\left[H(n)\underline{\tilde{\mathrm{x}}}^-(n) + \underline{\mathrm{v}}(n)\right]\right]\underline{\mathrm{z}}^{\mathrm{T}}(i)\middle|Z\right\} \\
&= 0, \quad i = 1, 2, \cdots, n.
\end{aligned} \tag{7.1-12}$$

We have assumed that, for $i = 1, 2, \cdots, n-1$,

$$E\left\{\underline{\tilde{\mathrm{x}}}^-(n)\underline{\mathrm{z}}^{\mathrm{T}}(i)\middle|Z\right\} = 0,$$

and $\underline{\mathrm{v}}(n)$ is independent of the measurements, $\underline{\mathrm{z}}(i), i = 1, 2, \cdots, n-1$. Thus, Eq. (7.1-12) is already satisfied for $i = 1, 2, \cdots, n-1$. We need only to consider the remaining case, i.e., $i = n$

$$E\left\{\left[\underline{\tilde{\mathrm{x}}}^-(n)-K(n)\left[H(n)\underline{\tilde{\mathrm{x}}}^-(n)+\underline{\mathrm{v}}(n)\right]\right]\underline{\mathrm{z}}^{\mathrm{T}}(i)\middle|Z\right\}$$
$$=E\left\{\left[\underline{\tilde{\mathrm{x}}}^-(n)-K(n)\left[H(n)\underline{\tilde{\mathrm{x}}}^-(n)+\underline{\mathrm{v}}(n)\right]\right]\left[\underline{\gamma}(\underline{\hat{\mathrm{x}}}^-(n))+H(n)\left[\underline{\mathrm{x}}(n)-\underline{\hat{\mathrm{x}}}^-(n)\right]+\underline{\mathrm{v}}(n)\right]^T\middle|Z\right\}$$
$$=P^-(n)H^T(n)-K(n)H(n)P^-(n)H^T(n)-K(n)R(n)$$
$$=0.$$

Solving this equation for $K(n)$, we have the Kalman gain for EKF

$$K(n)=P^-(n)H^T(n)\left[H(n)P^-(n)H^T(n)+R(n)\right]^{-1}. \qquad (7.1\text{-}13)$$

Finally, we need an expression for $P(n)$, the a posteriori conditional error covariance matrix

$$P(n)=Cov\left\{\underline{\tilde{\mathrm{x}}}(n)\middle|\mathrm{Z}=Z\right\}=Cov\left\{\left[\underline{\mathrm{x}}(n)-\underline{\hat{\mathrm{x}}}(n)\right]\middle|Z\right\}$$
$$=Cov\left\{\left[\underline{\mathrm{x}}(n)-\underline{\hat{\mathrm{x}}}^-(n)-K(n)\left[H(n)\underline{\tilde{\mathrm{x}}}^-(n)+\underline{\mathrm{v}}(n)\right]\right]\middle|Z\right\}$$
$$=P^-(n)-P^-(n)H^T(n)K^T(n)-K(n)H(n)P^-(n) \qquad (7.1\text{-}14)$$
$$+K(n)\left[H(n)P^-(n)H^T(n)+R(n)\right]K^T(n)$$
$$=P^-(n)-K(n)H(n)P^-(n).$$

Summary of EKF

System Model Measurement Model	$\underline{x}(n+1)=\underline{\phi}(\underline{x}(n))+\Gamma\underline{w}(n), \quad \underline{w}(n)\sim N(\underline{0},Q(n))$ $\underline{z}(n)=\underline{\gamma}(\underline{x}(n))+\underline{v}(n), \quad \underline{v}(n)\sim N(\underline{0},R(n))$
Initial Conditions Other Assumptions	$E[\underline{x}(0)]=\hat{\underline{x}}_0, E[(\underline{x}(0)-\hat{\underline{x}}_0)(\underline{x}(0)-\hat{\underline{x}}_0)^T]=P_0$ $E\left\{\underline{w}(n)\underline{v}^{\mathrm{T}}(n)\right\}=0,\ E\left\{\underline{w}(n)\underline{x}^{\mathrm{T}}(0)\right\}=0,\ E\left\{\underline{v}(n)\underline{x}^{\mathrm{T}}(0)\right\}=0$
Time Propagations	$\hat{\underline{x}}^-(n)=\underline{\phi}(\hat{\underline{x}}(n-1))$ $P^-(n)=\Phi(n-1)P(n-1)\Phi^T(n-1)+\Gamma Q(n-1)\Gamma^T$ $\Phi(n-1)=J_{\underline{\phi}}(\hat{\underline{x}}(n-1))$
Measurement Updates Kalman Gain Matrix	$\hat{\underline{x}}(n)=\hat{\underline{x}}^-(n)+K(n)\left[\underline{z}(n)-\underline{\gamma}(\hat{\underline{x}}^-(n))\right]$ $P(n)=P^-(n)-K(n)H(n)P^-(n)$ $K(n)=P^-(n)H^T(n)\left[H(n)P^-(n)H^T(n)+R(n)\right]^{-1}$ $H(n)=J_{\underline{\gamma}}(\hat{\underline{x}}^-(n))$

7.2 Frequency Demodulation

The frequency modulation (FM) is a well-known technique for transmitting analog waveforms. A continuous-time message $\mathrm{m}_c(t)$ modulates the angle of a sinusoidal carrier $\mathrm{c}_c(t)$. Let A_0 denote the carrier amplitude and Ω_0 the carrier frequency.

The carrier and message are related by

$$\mathrm{c}_c(t) = A_0 \cos\left(\Omega_0 t + \beta_0 \int_0^t \mathrm{m}_c(\tau) d\tau\right) \triangleq A_0 \cos\left(\Omega_0 t + \theta_c(t)\right), \qquad (7.2\text{-}1)$$

where β_0 is the modulation index. The problem of interest is to estimate the message $\mathrm{m}_c(t)$ from noisy measurements of the form $\mathrm{c}_c(t) + \mathrm{v}_c(t), \mathrm{v}_c(t) \sim N(0, R_c)$. This process is known as frequency demodulation.

We want to solve the frequency demodulation problem by applying the EKF. The message is a bandlimited signal in the frequency range $-\Omega_m < \Omega < \Omega_m$. We therefore model $\mathrm{m}_c(t)$ as the output of a lowpass filter that has cutoff frequency Ω_m and is excited by white noise $\mathrm{w}_c(t)$ with variance $\sigma_{w_c}^2$. We employ a first-order Butterworth filter. Then the Laplace transform $H_c(s)$ representation of the filter is

$$H_c(s) = \frac{M_c(s)}{W_c(s)} = \frac{\Omega_m}{s + \Omega_m}.$$

As a result, the differential equation relating $\mathrm{m}_c(t)$ and $\mathrm{w}_c(t)$ is

$$\dot{\mathrm{m}}_c(t) = -\Omega_m \mathrm{m}_c(t) + \Omega_m \mathrm{w}_c(t).$$

From Eq. (7.2-1), the derivative of $\theta_c(t)$ is

$$\dot{\theta}_c(t) = \beta_0 \mathrm{m}_c(t).$$

Defining a continuous-time state vector by $\underline{\mathrm{x}}_c(t) = \begin{bmatrix} \mathrm{m}_c(t) & \theta_c(t) \end{bmatrix}^T$, we obtain the following continuous-time state model

$$\dot{\underline{\mathrm{x}}}_c(t) = \begin{bmatrix} -\Omega_m & 0 \\ \beta_0 & 0 \end{bmatrix} \underline{\mathrm{x}}_c(t) + \begin{bmatrix} \Omega_m \\ 0 \end{bmatrix} \mathrm{w}_c(t) \tag{7.2-2}$$

$$\begin{aligned} \mathrm{z}_c(t) &= A_0 \cos\left(\Omega_0 t + \begin{bmatrix} 0 & 1 \end{bmatrix} \underline{\mathrm{x}}_c(t)\right) + \mathrm{v}_c(t) \\ &\triangleq \gamma_c\left(\underline{\mathrm{x}}_c(t)\right) + \mathrm{v}_c(t). \end{aligned} \tag{7.2-3}$$

To apply the EKF, discretize the continuous-time state model with a sampling period T. When we discretize the system, the discrete-time state vector is

$$\underline{\mathrm{x}}(n) = \underline{\mathrm{x}}_c(nT) = \begin{bmatrix} \mathrm{m}(n) & \theta(n) \end{bmatrix}^T = \begin{bmatrix} \mathrm{m}_c(nT) & \theta_c(nT) \end{bmatrix}^T . \tag{7.2-4}$$

To discretize Eq. (7.2-2), we use the Laplace transform relationship

$$\begin{aligned}
\Phi &= \mathfrak{L}^{-1}\left\{\left(sI - \begin{bmatrix} -\Omega_m & 0 \\ \beta_0 & 0 \end{bmatrix}\right)^{-1}\right\}_{t=T} = \mathfrak{L}^{-1}\left\{\begin{bmatrix} \dfrac{1}{s+\Omega_m} & 0 \\ \dfrac{\beta_0}{s\left(s+\Omega_m\right)} & \dfrac{1}{\mathrm{s}} \end{bmatrix}\right\}_{t=T} \\
&= \begin{bmatrix} e^{-\Omega_m T} & 0 \\ \dfrac{\beta_0}{\Omega_m}\left(1-e^{-\Omega_m T}\right) & 1 \end{bmatrix} \triangleq e^{AT}, \\
Q &= \sigma_{w_c}^2 \int_0^T e^{A\tau} B B^T e^{A^T \tau} d\tau \\
&= \sigma_{w_c}^2 \int_0^T \begin{bmatrix} {\Omega_m}^2 e^{-2\Omega_m \tau} & \beta_0 \Omega_m \left(e^{-\Omega_m \tau} - e^{-2\Omega_m \tau}\right) \\ \beta_0 \Omega_m \left(e^{-\Omega_m \tau} - e^{-2\Omega_m \tau}\right) & {\beta_0}^2 \left(1-e^{-\Omega_m \tau}\right)^2 \end{bmatrix} d\tau
\end{aligned} \tag{7.2-5}$$

$$
= \begin{bmatrix} \frac{\sigma_{w_c}^2 \Omega_m}{2}\left(1-e^{-2\Omega_m T}\right) & \frac{\sigma_{w_c}^2 \beta_0}{2}\left(1-2e^{-\Omega_m T}+e^{-2\Omega_m T}\right) \\ \frac{\sigma_{w_c}^2 \beta_0}{2}\left(1-2e^{-\Omega_m T}+e^{-2\Omega_m T}\right) & \frac{\sigma_{w_c}^2 {\beta_0}^2}{2\Omega_m}\left(-3+2\Omega_m T+4e^{-\Omega_m T}-e^{-2\Omega_m T}\right) \end{bmatrix}. \tag{7.2-6}
$$

So the state model is

$$
\underline{\mathrm{x}}(n+1) = \begin{bmatrix} e^{-\Omega_m T} & 0 \\ \frac{\beta_0}{\Omega_m}\left(1-e^{-\Omega_m T}\right) & 1 \end{bmatrix} \underline{\mathrm{x}}(n) + \underline{\mathrm{w}}(n),\ \underline{\mathrm{w}}(n) \sim N(\underline{0}, Q). \tag{7.2-7}
$$

The carrier signal is then

$$
\mathrm{c}(n) = \gamma(\underline{\mathrm{x}}(n)) = A_0 \cos\left(\Omega_0 nT + \theta(n)\right) = A_0 \cos\left(\Omega_0 nT + \begin{bmatrix} 0 & 1 \end{bmatrix} \underline{\mathrm{x}}(n)\right). \tag{7.2-8}
$$

Finally, the output equation becomes

$$
z(n) = \gamma(\underline{\mathrm{x}}(n)) + \mathrm{v}(n),\ \mathrm{v}(n) \sim \left(0, R = \frac{{\sigma_{\mathrm{v}_c}}^2}{\mathrm{T}}\right). \tag{7.2-9}
$$

Note that the state and input dynamics of Eq. (7.2-7) are linear, but the output Eq. (7.2-9) is nonlinear. Hence we need to linearize Eq. (7.2-9). Applying the EKF approximation to Eq. (7.2-8) or Eq. (7.2-1), we obtain

$$\begin{aligned} H(n) = J_{\underline{\gamma}}(\hat{\underline{\mathrm{x}}}^-(n)) &= \begin{bmatrix} 0 & -A_0 \sin\left(\Omega_0 nT + \hat{\theta}^-(n)\right) \end{bmatrix} \\ &= \begin{bmatrix} 0 & -A_0 \sin\left(\Omega_0 nT + \begin{bmatrix} 0 & 1 \end{bmatrix} \hat{\underline{\mathrm{x}}}^-(n)\right) \end{bmatrix}. \end{aligned}$$

The a posteriori state estimate Eq. (7.1-11) becomes

$$\begin{aligned} \hat{\underline{\mathrm{x}}}(n) &= \hat{\underline{\mathrm{x}}}^-(n) + K(n)\left[\underline{\mathrm{z}}(n) - A_0 \cos\left(\Omega_0 nT + \begin{bmatrix} 0 & 1 \end{bmatrix} \hat{\underline{\mathrm{x}}}^-(n)\right) \right] \\ &= \hat{\underline{\mathrm{x}}}^-(n) + K(n)\left[\underline{\mathrm{z}}(n) - A_0 \cos\left(\Omega_0 nT + \hat{\theta}^-(n)\right) \right] \end{aligned}$$

with

$$\begin{aligned} &\hat{\underline{\mathrm{x}}}^-(n) = \Phi \, \hat{\underline{\mathrm{x}}}(n-1) \\ &P^-(n) = \Phi \, P(n-1)\Phi^T + Q \\ &K(n) = P^-(n) H^T(n)\left[H(n) P^-(n) H^T(n) + R \right]^{-1} \\ &P(n) = P^-(n) - K(n) H(n) P^-(n). \end{aligned}$$

The demodulated message is then

$$\hat{\mathrm{m}}(n) = \begin{bmatrix} 1 & 0 \end{bmatrix} \hat{\underline{\mathrm{x}}}(n) .$$

With no a priori information, we may initialize the EKF with

$$\hat{\underline{\mathrm{x}}}(0) = \begin{bmatrix} 0 & 0 \end{bmatrix}^T$$

$$P(0) = \alpha I, \quad \alpha > 0 .$$

Example 7.1 Frequency Demodulation

$$A_0 = \text{carrier amplidude} = 1$$

$$f_0 = \text{carrier frequency} = 100 \text{ MHz}$$

$$f_m = \text{message bandwidth} = 15 \text{ KHz}$$

$$\beta_0 = \text{frequency modulation index} = 5$$

$$f_s = \text{sampling frequency} = 250 \text{ MHz}$$

$$T = \text{sampling period} = 1/f_s = 9 \times 10^{-9} = 9 \text{ ns}$$

$$\sigma_{wc}^2 = \text{variance of process noise} = 0.01$$

$$\sigma_{vc}^2 = \text{variance of measurement noise} = 4 \times 10^{-12} \text{ (after discretization, } \sigma_v^2 = 0.001\text{)}$$

The carrier signal-to-noise ratio is

$$SNR_c = 10 \log_{10}\left(\frac{A_0^2}{2\sigma_v^2}\right) \approx 27 \text{ dB}.$$

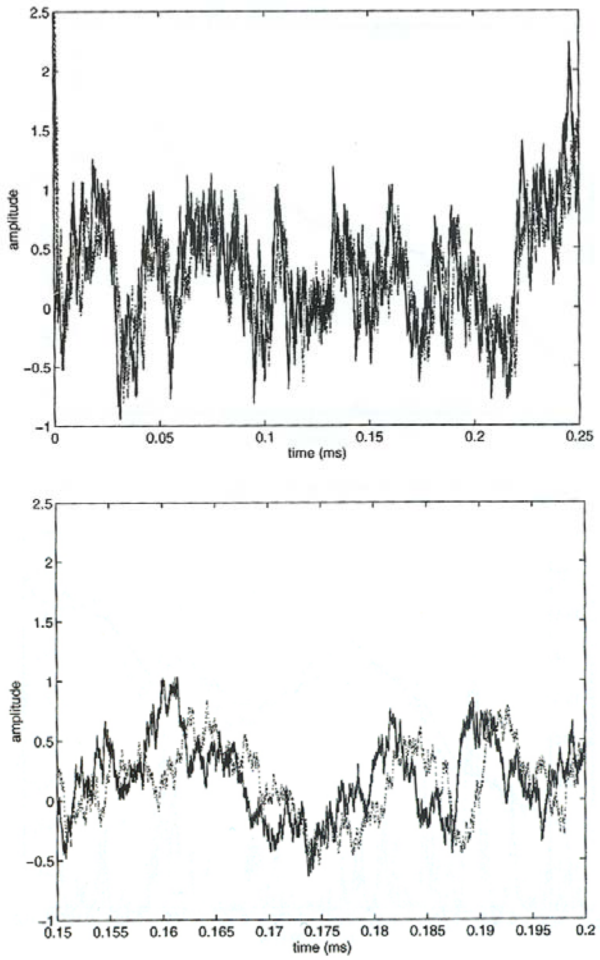

Figure 8.1. Frequency demodulation results for Example 8.1. **Top:** Actual message $m(n)$ (solid curve) and EKF estimate of message $\hat{m}(n)$ (dotted curve). **Bottom:** Detail of message.

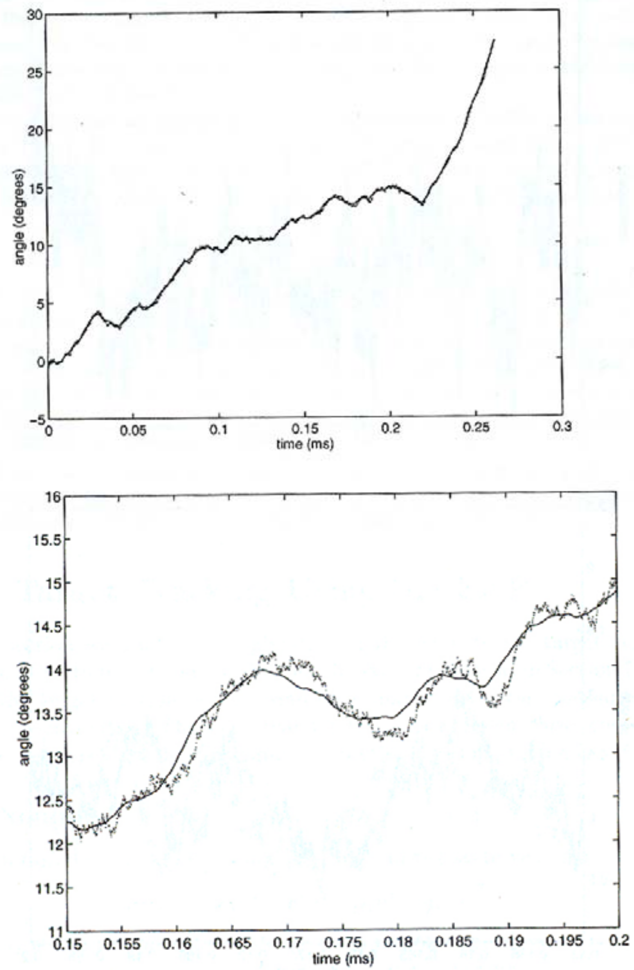

Figure 8.2. Frequency demodulation results for Example 8.1. **Top:** Actual angle $\theta(n)$ (solid curve) and EKF estimate $\hat{\theta}(n)$ (dotted curve). **Bottom:** Detail of angle and estimate.

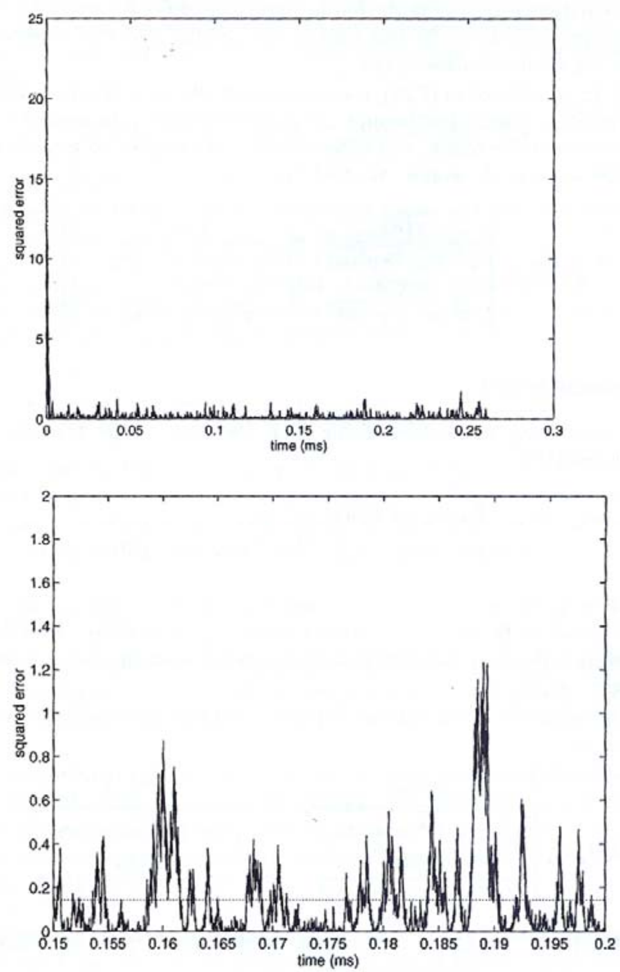

Figure 8.3. Frequency demodulation results for Example 8.1, cont. **Top:** True MSE (solid curve) and EKF estimate of the MSE ($\mathrm{tr}\,(P(n))$, dotted curve). **Bottom:** Detail of MSE.

7.3 Unscented Kalman Filter

Unscented Transformations

Suppose that $\underline{x}$ is an $n \times 1$ random vector that is transformed by a nonlinear function $\underline{y} = \underline{h}(\underline{x})$. Further assumed is that the pdf of $\underline{x}$ is symmetric around its mean and $\underline{h}(\underline{x})$ is smooth. Choose $2n$ sigma points $\underline{x}^{(i)}$ as follows:

$$\begin{aligned}
\underline{x}^{(i)} &= \overline{\underline{x}} + \underline{\tilde{x}}^{(i)} \quad i = 1, \cdots, 2n \\
\underline{\tilde{x}}^{(i)} &= \left(\sqrt{nP}\right)_i^T \quad i = 1, \cdots, n \\
\underline{\tilde{x}}^{(n+i)} &= -\left(\sqrt{nP}\right)_i^T \quad i = 1, \cdots, n
\end{aligned} \tag{7.3-1}$$

where

$\sqrt{nP}$ = the matrix square root of nP such that $\left(\sqrt{nP}\right)^T \sqrt{nP} = nP$

$\left(\sqrt{nP}\right)_i$ = the ith row of $\sqrt{nP}$.

Mean approximation – Suppose that we have a vector $\underline{x}$ with a known mean $\overline{\underline{x}}$ and covariance P, a nonlinear function $\underline{y} = \underline{h}(\underline{x})$, and we want to appproximate the mean of $\underline{y}$. Let the transformed sigma points be computed by

$$\underline{y}^{(i)} = \underline{h}(\underline{x}^{(i)}) \quad i = 1, \cdots, 2n. \tag{7.3-2}$$

Suppose that the approximated mean of $\underline{y}$ denoted as $\overline{\underline{y}}_u$ is computed as follows

$$\begin{aligned}
\overline{\underline{y}}_u &= \sum_{i=1}^{2n} W^{(i)} \underline{y}^{(i)} \\
&= \frac{1}{2n} \sum_{i=1}^{2n} \underline{y}^{(i)} \quad i = 1, \cdots, 2n \\
&= \frac{1}{2n} \sum_{i=1}^{2n} \left(\underline{h}\left(\overline{\underline{x}}\right) + D_{\underline{\tilde{x}}^{(i)}} \underline{h} + \frac{1}{2!} D^2_{\underline{\tilde{x}}^{(i)}} \underline{h} + \cdots \right) \\
&= \underline{h}\left(\overline{\underline{x}}\right) + \frac{1}{2n} \sum_{i=1}^{2n} \left(D_{\underline{\tilde{x}}^{(i)}} \underline{h} + \frac{1}{2!} D^2_{\underline{\tilde{x}}^{(i)}} \underline{h} + \cdots \right)
\end{aligned} \tag{7.3-3}$$

where $D^k_{\underline{\tilde{x}}^{(j)}} \underline{h} = \left(\sum_{i=1}^{n} \tilde{x}_i^{(j)} \frac{\partial}{\partial x_i} \right)^k \underline{h}(\underline{x}) \Big|_{\underline{x} = \overline{\underline{x}}}$ and $\underline{\tilde{x}} = \underline{x} - \overline{\underline{x}}$.

Notice that since $\underline{\tilde{x}}^{(j)} = -\underline{\tilde{x}}^{(n+j)}$ $(j = 1, \cdots, n)$ according to Eq. (7.3-1), for any odd power term, we have

$$\begin{aligned}
\sum_{j=1}^{2n} D_{\underline{\tilde{x}}^{(j)}}^{2k+1} \underline{h} &= \sum_{j=1}^{2n} \left[\left(\sum_{i=1}^{n} \tilde{x}_i^{(j)} \frac{\partial}{\partial x_i} \right)^{2k+1} \underline{h}(\underline{x}) \Big|_{\underline{x}=\overline{\underline{x}}} \right] \\
&= \sum_{j=1}^{2n} \left[\sum_{i=1}^{n} \left(\tilde{x}_i^{(j)} \right)^{2k+1} \frac{\partial^{2k+1}}{\partial x_i^{2k+1}} \underline{h}(\underline{x}) \Big|_{\underline{x}=\overline{\underline{x}}} \right] \\
&= \sum_{i=1}^{n} \left[\sum_{j=1}^{2n} \left(\tilde{x}_i^{(j)} \right)^{2k+1} \frac{\partial^{2k+1}}{\partial x_i^{2k+1}} \underline{h}(\underline{x}) \Big|_{\underline{x}=\overline{\underline{x}}} \right] \\
&= 0.
\end{aligned} \qquad (7.3\text{-}4)$$

Eq. (7.3-3) can be expressed

$$\begin{aligned}
\underline{\overline{y}}_u &= \underline{h}\left(\overline{\underline{x}}\right) + \frac{1}{2n} \sum_{i=1}^{2n} \left(\frac{1}{2!} D_{\underline{\tilde{x}}^{(i)}}^{2} \underline{h} + \frac{1}{4!} D_{\underline{\tilde{x}}^{(i)}}^{4} \underline{h} + \cdots \right) \\
&= \underline{h}\left(\overline{\underline{x}}\right) + \frac{1}{2n} \sum_{i=1}^{2n} \frac{1}{2!} D_{\underline{\tilde{x}}^{(i)}}^{2} \underline{h} + \frac{1}{2n} \sum_{i=1}^{2n} \left(\frac{1}{4!} D_{\underline{\tilde{x}}^{(i)}}^{4} \underline{h} + \frac{1}{6!} D_{\underline{\tilde{x}}^{(i)}}^{6} \underline{h} + \cdots \right).
\end{aligned} \qquad (7.3\text{-}5)$$

Look at the second term on the right side of Eq. (7.3-5)

$$\begin{aligned}
\frac{1}{2n}\sum_{i=1}^{2n}\frac{1}{2!}D^2_{\underline{\tilde{x}}^{(i)}}\underline{h} &= \frac{1}{2n}\sum_{k=1}^{2n}\frac{1}{2!}\left(\sum_{i=1}^{n}\tilde{x}_i^{(k)}\frac{\partial}{\partial x_i}\right)^2 \underline{h}(\underline{x})\Bigg|_{\underline{x}=\underline{\bar{x}}} \\
&= \frac{1}{4n}\sum_{k=1}^{2n}\sum_{i,j=1}^{n}\tilde{x}_i^{(k)}\tilde{x}_j^{(k)}\frac{\partial^2}{\partial x_i \partial x_j}\underline{h}(\underline{x})\Bigg|_{\underline{x}=\underline{\bar{x}}} \\
&= \frac{1}{4n}\sum_{i,j=1}^{n}\sum_{k=1}^{2n}\tilde{x}_i^{(k)}\tilde{x}_j^{(k)}\frac{\partial^2}{\partial x_i \partial x_j}\underline{h}(\underline{x})\Bigg|_{\underline{x}=\underline{\bar{x}}} \\
&= \frac{1}{2n}\sum_{i,j=1}^{n}\sum_{k=1}^{n}\tilde{x}_i^{(k)}\tilde{x}_j^{(k)}\frac{\partial^2}{\partial x_i \partial x_j}\underline{h}(\underline{x})\Bigg|_{\underline{x}=\underline{\bar{x}}}
\end{aligned} \qquad (7.3\text{-}6)$$

where we have again used the fact from Eq. (7.3-1) that $\underline{\tilde{x}}^{(k)} = -\underline{\tilde{x}}^{(k+n)}$ $(k = 1,\cdots,n)$. Substitute for $\tilde{x}_i^{(k)}$ and $\tilde{x}_j^{(k)}$ from Eq. (7.3-1) in Eq. (7.3-6) to obtain

$$\frac{1}{2n}\sum_{i,j=1}^{n}\sum_{k=1}^{n}\tilde{x}_i^{(k)}\tilde{x}_j^{(k)}\frac{\partial^2 \underline{h}(\underline{x})}{\partial x_i \partial x_j}\Bigg|_{\underline{x}=\underline{\bar{x}}} = \frac{1}{2n}\sum_{i,j=1}^{n}\sum_{k=1}^{n}\left(\sqrt{nP}\right)_{ki}\left(\sqrt{nP}\right)_{kj}\frac{\partial^2 \underline{h}(\underline{x})}{\partial x_i \partial x_j}\Bigg|_{\underline{x}=\underline{\bar{x}}}$$

$$
\begin{aligned}
&= \frac{1}{2n}\sum_{i,j=1}^{n} nP_{ij}\, \frac{\partial^2 \underline{h}(\underline{x})}{\partial x_i \partial x_j}\bigg|_{\underline{x}=\underline{\bar{x}}} \qquad \left(\left(\sqrt{nP}\right)^T \sqrt{nP} = nP \text{ or } \sum_{k=1}^{n} \left(\sqrt{nP}\right)_{ik} \left(\sqrt{nP}\right)_{kj} = nP_{ij} \right) \\
&= \frac{1}{2}\sum_{i,j=1}^{n} P_{ij}\, \frac{\partial^2 \underline{h}(\underline{x})}{\partial x_i \partial x_j}\bigg|_{\underline{x}=\underline{\bar{x}}} .
\end{aligned}
\tag{7.3-7}
$$

Eq. (7.3-5) can therefore be written as

$$
\underline{\bar{y}}_u = \underline{h}\left(\underline{\bar{x}}\right) + \frac{1}{2}\sum_{i,j=1}^{n} P_{ij}\, \frac{\partial^2 \underline{h}(\underline{x})}{\partial x_i \partial x_j}\bigg|_{\underline{x}=\underline{\bar{x}}} + \frac{1}{2n}\sum_{i=1}^{2n}\left(\frac{1}{4!} D^4_{\underline{\tilde{x}}^{(i)}} \underline{h} + \frac{1}{6!} D^6_{\underline{\tilde{x}}^{(i)}} \underline{h} + \cdots \right). \tag{7.3-8}
$$

Now, note that the true mean of $\underline{y}$ can be written by

$$
\begin{aligned}
\underline{\bar{y}} &= E\left\{\underline{h}(\underline{x})\right\} \\
&= E\left\{ \underline{h}(\underline{\bar{x}}) + D_{\underline{\tilde{x}}}\underline{h} + \frac{1}{2!} D^2_{\underline{\tilde{x}}}\underline{h} + \frac{1}{3!} D^3_{\underline{\tilde{x}}}\underline{h} + \cdots \right\} \\
&= \underline{h}(\underline{\bar{x}}) + E\left\{ D_{\underline{\tilde{x}}}\underline{h} + \frac{1}{2!} D^2_{\underline{\tilde{x}}}\underline{h} + \frac{1}{3!} D^3_{\underline{\tilde{x}}}\underline{h} + \cdots \right\}
\end{aligned}
$$

$$
\begin{aligned}
&= \underline{h}(\underline{\overline{x}}) + E\left\{\frac{1}{2!}D^2_{\underline{\tilde{x}}}\underline{h} + \frac{1}{4!}D^4_{\underline{\tilde{x}}}\underline{h} + \cdots\right\} \\
&= \underline{h}(\underline{\overline{x}}) + \frac{1}{2!}E\left\{D^2_{\underline{\tilde{x}}}\underline{h}\right\} + \frac{1}{4!}E\left\{D^4_{\underline{\tilde{x}}}\underline{h}\right\} + \cdots
\end{aligned}
\qquad (7.3\text{-}9)
$$

since $E\left\{D_{\underline{\tilde{x}}}\underline{h}\right\} = E\left\{D^3_{\underline{\tilde{x}}}\underline{h}\right\} = \cdots = 0$. The second term on the right side of Eq. (7.3-9) can be written

$$
\begin{aligned}
\frac{1}{2!}E\left\{D^2_{\underline{\tilde{x}}}\underline{h}\right\} &= \frac{1}{2!}E\left\{\left(\sum_{i=1}^{n}\tilde{x}_i\frac{\partial}{\partial x_i}\right)^2 \underline{h}(\underline{x})\Bigg|_{\underline{x}=\underline{\overline{x}}}\right\} \\
&= \frac{1}{2!}E\left\{\sum_{i,j=1}^{n}\tilde{x}_i\tilde{x}_j\frac{\partial^2\underline{h}}{\partial x_i\partial x_j}\Bigg|_{\underline{x}=\underline{\overline{x}}}\right\} \qquad (7.3\text{-}10) \\
&= \frac{1}{2!}\sum_{i,j=1}^{n}P_{ij}\frac{\partial^2\underline{h}}{\partial x_i\partial x_j}\Bigg|_{\underline{x}=\underline{\overline{x}}} .
\end{aligned}
$$

We therefore see that $\underline{\overline{y}}$ can be written from Eq. (7.3-9) as

$$\underline{\bar{y}} = \underline{h}(\underline{\bar{x}}) + \frac{1}{2!}\sum_{i,j=1}^{n} P_{ij} \left.\frac{\partial^2 \underline{h}}{\partial x_i \partial x_j}\right|_{\underline{x}=\underline{\bar{x}}} + \frac{1}{4!} E\left\{ D_{\underline{\tilde{x}}}^4 \underline{h} \right\} + \frac{1}{6!} E\left\{ D_{\underline{\tilde{x}}}^6 \underline{h} \right\} + \cdots. \tag{7.3-11}$$

Comparing this with Eq. (7.3-8), we see that $\underline{\bar{y}}_u$ (the approximated mean of $\underline{y}$) matches the true mean of $\underline{y}$ correctly up to the third order.

Covariance approximation - Suppose that we have a vector $\underline{x}$ with a known mean $\underline{\bar{x}}$ and covariance P, a nonlinear function $\underline{y} = \underline{h}(\underline{x})$, and we want to appproximate the covariance of $\underline{y}$. Denote the approximation as P_u and propose the following equation

$$\begin{aligned} P_u &= \sum_{i=1}^{2n} W^{(i)} \left(\underline{y}^{(i)} - \underline{\bar{y}}_u \right)\left(\underline{y}^{(i)} - \underline{\bar{y}}_u \right)^T \\ &= \frac{1}{2n}\sum_{i=1}^{2n} \left(\underline{y}^{(i)} - \underline{\bar{y}}_u \right)\left(\underline{y}^{(i)} - \underline{\bar{y}}_u \right)^T. \end{aligned} \tag{7.3-12}$$

Expanding Eq. (7.3-12) using the Taylor series and Eq. (7.3-5) gives

$$P_u = \frac{1}{2n}\sum_{i=1}^{2n} \left(\underline{h}\left(\underline{x}^{(i)} \right) - \underline{\bar{y}}_u \right)\left(\underline{h}\left(\underline{x}^{(i)} \right) - \underline{\bar{y}}_u \right)^T$$

$$
\begin{aligned}
&= \frac{1}{2n}\sum_{i=1}^{2n}\Bigg[\underline{h}(\overline{x}) + D_{\underline{\tilde{x}}^{(i)}}\underline{h} + \frac{1}{2!}D^2_{\underline{\tilde{x}}^{(i)}}\underline{h} + \frac{1}{3!}D^3_{\underline{\tilde{x}}^{(i)}}\underline{h} + \cdots \\
&\qquad - \underline{h}\left(\overline{x}\right) - \frac{1}{2n}\sum_{j=1}^{2n}\left(\frac{1}{2!}D^2_{\underline{\tilde{x}}^{(j)}}\underline{h} + \frac{1}{4!}D^4_{\underline{\tilde{x}}^{(j)}}\underline{h} + \cdots\right)\Bigg]\left[\cdots\right]^T \\
&= \frac{1}{2n}\sum_{i=1}^{2n}\Bigg\{\left(D_{\underline{\tilde{x}}^{(i)}}\underline{h}\right)\left(D_{\underline{\tilde{x}}^{(i)}}\underline{h}\right)^T + \underbrace{\left[\left(\frac{1}{2}D_{\underline{\tilde{x}}^{(i)}}\underline{h}\right)\left(D^2_{\underline{\tilde{x}}^{(i)}}\underline{h}\right)^T\right]}_{0} + \underbrace{\left[\cdots\right]^T}_{0} + \frac{1}{4}\left(D^2_{\underline{\tilde{x}}^{(i)}}\underline{h}\right)\left(D^2_{\underline{\tilde{x}}^{(i)}}\underline{h}\right)^T \\
&\qquad - \underbrace{\left[D_{\underline{\tilde{x}}^{(i)}}\underline{h}\left(\frac{1}{2n}\sum_j\frac{1}{2}D^2_{\underline{\tilde{x}}^{(i)}}\underline{h}\right)^T\right]}_{0} - \underbrace{\left[\cdots\right]^T}_{0} + \frac{1}{4n^2}\left(\sum_j\frac{1}{2}D^2_{\underline{\tilde{x}}^{(i)}}\underline{h}\right)\left(\cdots\right)^T \\
&\qquad - \left[\frac{1}{4n}D^2_{\underline{\tilde{x}}^{(i)}}\underline{h}\left(\sum_j D^2_{\underline{\tilde{x}}^{(i)}}\underline{h}\right)^T\right] - \left[\cdots\right]^T + \left[D_{\underline{\tilde{x}}^{(i)}}\underline{h}\left(\frac{1}{3!}D^3_{\underline{\tilde{x}}^{(i)}}\underline{h}\right)^T\right] + \left[\cdots\right]^T\Bigg\} \qquad (7.3\text{-}13) \\
&= \frac{1}{2n}\sum_{i=1}^{2n}\left(D_{\underline{\tilde{x}}^{(i)}}\underline{h}\right)\left(D_{\underline{\tilde{x}}^{(i)}}\underline{h}\right)^T + HOT.
\end{aligned}
$$

Some of the terms in the above equation are zero as noted because $\underline{\tilde{x}}^{(i)} = -\underline{\tilde{x}}^{(n+i)}\ (i = 1,\cdots,n)$.

Expanding Eq. (7.3-13) gives

$$P_u = \frac{1}{2n}\sum_{i=1}^{2n}\sum_{j,k=1}^{n}\left(\frac{\partial \underline{h}(\underline{x})}{\partial x_j}\tilde{x}_j^{(i)}\right)\left(\frac{\partial \underline{h}(\underline{x})}{\partial x_k}\tilde{x}_k^{(i)}\right)^T + HOT\,. \qquad (7.3\text{-}14)$$

Recall that $\tilde{x}_j^{(i)} = -\tilde{x}_j^{(i+n)}$ and $\tilde{x}_k^{(i)} = -\tilde{x}_k^{(i+n)}$ for $i = 1,\cdots,n$. The covariance approximation becomes

$$\begin{aligned} P_u &= \frac{1}{n}\sum_{i=1}^{n}\sum_{j,k=1}^{n}\left(\frac{\partial \underline{h}(\underline{x})}{\partial x_j}\tilde{x}_j^{(i)}\right)\left(\frac{\partial \underline{h}(\underline{x})}{\partial x_k}\tilde{x}_k^{(i)}\right)^T + HOT \\ &= \sum_{j,k=1}^{n}\frac{\partial \underline{h}(\underline{x})}{\partial x_j}P_{jk}\left(\frac{\partial \underline{h}(\underline{x})}{\partial x_k}\right)^T + HOT \quad \left(\text{Applied is } \underline{\tilde{x}}^{(i)} = \left(\sqrt{nP}\right)_i^T.\right) \qquad (7.3\text{-}15) \\ &= HPH^T + HOT. \end{aligned}$$

Now, note that the true covariance of $\underline{y}$ can be written by

$$P_{\underline{y}} = E\left\{\left(\underline{y} - \underline{\bar{y}}\right)\left(\underline{y} - \underline{\bar{y}}\right)^T\right\}$$

$$
= E\left\{\left(\left[\underline{h}(\underline{\bar{x}}) + D_{\underline{\tilde{x}}}\underline{h} + \frac{1}{2!}D_{\underline{\tilde{x}}}^2\underline{h} + \cdots\right] - \left[\underline{h}(\underline{\bar{x}}) + \frac{1}{2!}E\left\{D_{\underline{\tilde{x}}}^2\underline{h}\right\} + \frac{1}{4!}E\left\{D_{\underline{\tilde{x}}}^4\underline{h}\right\} + \cdots\right]\right)(\cdot)^T\right\}
$$

$$
= E\left\{D_{\underline{\tilde{x}}}\underline{h}\left(D_{\underline{\tilde{x}}}\underline{h}\right)^T\right\} + E\left\{\frac{D_{\underline{\tilde{x}}}\underline{h}\left(D_{\underline{\tilde{x}}}^3\underline{h}\right)^T}{3!} + \frac{D_{\underline{\tilde{x}}}^2\underline{h}\left(D_{\underline{\tilde{x}}}^2\underline{h}\right)^T}{2!2!} + \frac{D_{\underline{\tilde{x}}}^3\underline{h}\left(D_{\underline{\tilde{x}}}\underline{h}\right)^T}{3!}\right\} \tag{7.3-16}
$$

$$
+ E\left\{\frac{D_{\underline{\tilde{x}}}^2\underline{h}}{2!}\right\}E\left\{\left(\frac{D_{\underline{\tilde{x}}}^2\underline{h}}{2!}\right)^T\right\} + \cdots.
$$

The first term on the right side of the above equation can be written as

$$
E\left\{D_{\underline{\tilde{x}}}\underline{h}\left(D_{\underline{\tilde{x}}}\underline{h}\right)^T\right\} = E\left\{\left(\sum_{i=1}^{n}\left.\frac{\partial \underline{h}(\underline{x})}{\partial x_i}\right|_{\underline{x}=\underline{\bar{x}}} \cdot \tilde{x}_i\right)\left(\sum_{i=1}^{n}\left.\frac{\partial \underline{h}(\underline{x})}{\partial x_i}\right|_{\underline{x}=\underline{\bar{x}}} \cdot \tilde{x}_i\right)^T\right\} \tag{7.3-17}
$$

$$
= HPH^T.
$$

Substitute Eq. (7.3-17) into Eq. (7.3-16) to obtain

$$\begin{aligned} P_{\underline{y}} &= HPH^T + E\left\{ \frac{D_{\tilde{\underline{x}}}\underline{h}\left(D_{\tilde{\underline{x}}}^3\underline{h}\right)^T}{3!} + \frac{D_{\tilde{\underline{x}}}^2\underline{h}\left(D_{\tilde{\underline{x}}}^2\underline{h}\right)^T}{2!2!} + \frac{D_{\tilde{\underline{x}}}^3\underline{h}\left(D_{\tilde{\underline{x}}}\underline{h}\right)^T}{3!} \right\} \\ &\quad + E\left\{\frac{D_{\tilde{\underline{x}}}^2\underline{h}}{2!}\right\} E\left\{\left(\frac{D_{\tilde{\underline{x}}}^2\underline{h}}{2!}\right)^T\right\} + \cdots. \end{aligned} \tag{7.3-18}$$

Comparing this with Eq. (7.3-15), we see that P_u (the approximated covariance of $\underline{y}$) matches the true covariance of $\underline{y}$ correctly up to the third order.

The Unscented Kalman Filter

Based on the unscented transformations, the unscented Kalman filter (UKF) algorithm can be constructed.

1. We have an *n*-state discrete-time nonlinear system given by

$$\underline{x}_{k+1} = \underline{f}\left(\underline{x}_k, \underline{u}_k\right) + \underline{w}_k, \quad \underline{w}_k \sim \left(\underline{0}, Q_k\right)$$

$$\underline{z}_k = \underline{h}\left(\underline{x}_k\right) + \underline{v}_k, \quad \underline{v}_k \sim \left(\underline{0}, R_k\right). \tag{7.3-19}$$

2. The UKF is initialized as follows

$$\begin{aligned} \hat{\underline{x}}_0^+ &= E\left\{\underline{x}_0\right\} \\ P_0^+ &= E\left\{\left(\underline{x}_0 - \hat{\underline{x}}_0^+\right)\left(\underline{x}_0 - \hat{\underline{x}}_0^+\right)^T\right\}. \end{aligned} \tag{7.3-20}$$

3. The following time update equations are used to propagate the state estimate and covariance from one measurement time to the next.
   (a) To propagate from time step $(k-1)$ to k, first choose sigma points $\underline{x}_{k-1}^{(i)}$ as specified in Eq. (8.3-1) with appropriate changes

$$\begin{aligned} \hat{\underline{x}}_{k-1}^{(i)} &= \hat{\underline{x}}_{k-1}^+ + \tilde{\underline{x}}^{(i)} \quad i = 1, \cdots, 2n \\ \tilde{\underline{x}}^{(i)} &= \left(\sqrt{nP_{k-1}^+}\right)_i^T \quad i = 1, \cdots, n \\ \tilde{\underline{x}}^{(n+i)} &= -\left(\sqrt{nP_{k-1}^+}\right)_i^T \quad i = 1, \cdots, n. \end{aligned} \tag{7.3-21}$$

   (b) Use the known nonlinear system equation $\underline{f}(\cdot)$ to transform the sigma points into $\underline{x}_k^{(i)}$

vectors as shown in Eq. (7.3-2) with appropriate changes

$$\underline{x}_k^{(i)} = \underline{f}\left(\underline{x}_{k-1}^{(i)}, \underline{u}_k\right). \tag{7.3-22}$$

(c) Combine the $\underline{x}_k^{(i)}$ vectors to obtain the a priori state estimate at time k based on Eq. (7.3-3)

$$\hat{\underline{x}}_k^- = \frac{1}{2n}\sum_{i=1}^{2n}\hat{\underline{x}}_k^{(i)}. \tag{7.3-23}$$

(d) Estimate the a priori error covariance as shown in Eq. (7.3-12). However, we should add Q_{k-1} to the end of the equation to take the process noise into account

$$P_k^- = \frac{1}{2n}\sum_{i=1}^{2n}\left(\hat{\underline{x}}_k^{(i)} - \hat{\underline{x}}_k^-\right)\left(\hat{\underline{x}}_k^{(i)} - \hat{\underline{x}}_k^-\right)^T + Q_{k-1}. \tag{7.3-24}$$

4. Now that the time update equations are done, we implement the measurement update the equations.
   (a) Choose sigma points $\underline{x}_k^{(i)}$ as specified in Eq. (7.3-1) with appropriate changes

$$\underline{\hat{x}}_k^{(i)} = \underline{\hat{x}}_k^- + \underline{\tilde{x}}^{(i)} \quad i = 1, \cdots, 2n$$

$$\underline{\tilde{x}}^{(i)} = \left(\sqrt{nP_k^-}\right)_i^T \quad i = 1, \cdots, n \tag{7.3-25}$$

$$\underline{\tilde{x}}^{(n+i)} = -\left(\sqrt{nP_k^-}\right)_i^T \quad i = 1, \cdots, n.$$

(b) Use the known nonlinear measurement equation $\underline{h}(\cdot)$ to transform the sigma points into $\underline{\hat{z}}_k^{(i)}$ vectors (predicted measurements) as shown in Eq. (7.3-2)

$$\underline{\hat{z}}_k^{(i)} = \underline{h}\left(\underline{\hat{x}}_k^{(i)}\right). \tag{7.3-26}$$

(c) Combine the $\underline{\hat{z}}_k^{(i)}$ vectors to obtain the predicted measurement at time k based on Eq. (8.3-3)

$$\underline{\hat{z}}_k = \frac{1}{2n}\sum_{i=1}^{2n} \underline{\hat{z}}_k^{(i)}. \tag{7.3-27}$$

(d) Estimate the covariance of the predicted measurement as shown in Eq. (7.3-12). However, we should add R_k to the end of the equation to take the measurement noise into account

$$P_{\underline{z}} = \frac{1}{2n}\sum_{i=1}^{2n}\left(\hat{\underline{z}}_k^{(i)} - \hat{\underline{z}}_k\right)\left(\hat{\underline{z}}_k^{(i)} - \hat{\underline{z}}_k\right)^T + R_k \,. \tag{7.3-28}$$

(e) Estimate the cross covariance between $\hat{\underline{x}}_k^-$ and $\hat{\underline{z}}_k$ based on Eq. (7.3-12)

$$P_{\underline{xz}} = \frac{1}{2n}\sum_{i=1}^{2n}\left(\hat{\underline{x}}_k^{(i)} - \hat{\underline{x}}_k^-\right)\left(\hat{\underline{z}}_k^{(i)} - \hat{\underline{z}}_k\right)^T \,. \tag{7.3-29}$$

(f) The measurement update of the state estimate can be performed using the normal Kalman filter equations

$$\begin{aligned} K_k &= P_{\underline{xz}} P_{\underline{z}}^{-1} \\ \hat{\underline{x}}_k^+ &= \hat{\underline{x}}_k^- + K_k\left(\underline{z}_k - \hat{\underline{z}}_k\right) \\ P_k^+ &= P_k^- - K_k P_{\underline{z}} K_k^T \,. \end{aligned} \tag{7.3-30}$$

This completes the UKF algorithm which has a similar form with the EKF. Note that the EKF is based on linearization while the UKF is based on the unscented transformations which are more accurate than linearization for propagating means and covariances.

7.4 Particle Filter

Monte Carlo Integration

The Monte Carlo (MC) estimate of integral

$$I = \int f(x) p(x) dx \qquad (7.4\text{-}1)$$

is the sample mean

$$I_N = \frac{1}{N} \sum_{i=1}^{N} f\left(x^i\right) \qquad (7.4\text{-}2)$$

where $f(x)$ is an arbitrary function of x and $p(x)$ is the probability density. If the samples x^i are independent then I_N is an unbiased estimate and according to the law of large numbers I_N will almost surely converge to I. If the variance of $f(x)$,

$$\sigma^2 = \int \left(f(x) - I \right)^2 p(x) dx$$

is finite, then the central limit theorem holds and the estimation error converges in distribution

$$\lim_{N\to\infty}\sqrt{N}\left(I_N - I\right) \sim N\left(0,\sigma^2\right).$$

Importance Sampling

Suppose that we do not exactly know $p(x)$ and we can only generate samples from a density $\pi(x)$ which is similar to $p(x)$. Employing $\pi(x)$ Eq. (7.4-1) can be written

$$I = \int f(x)p(x)dx = \int f(x)\frac{p(x)}{\pi(x)}\pi(x)dx \tag{7.4-3}$$

A Monte Carlo estimate of I is computed by generating $N \gg 1$ independent samples $\left\{x^i; i = 1,\cdots,N\right\}$ distributed according to $\pi(x)$ and forming the weighted sum

$$I_N = \frac{1}{N}\sum_{i=1}^{N} f\left(x^i\right)\tilde{q}\left(x^i\right) \tag{7.4-4}$$

where

$$\tilde{q}\left(x^i\right) = \frac{p\left(x^i\right)}{\pi\left(x^i\right)}. \tag{7.4-5}$$

Normalizing $\tilde{q}\left(x^i\right)$, Eq. (7.4-4) can be expressed

$$I_N = \frac{\frac{1}{N}\sum_{i=1}^{N} f\left(x^i\right)\tilde{q}\left(x^i\right)}{\frac{1}{N}\sum_{j=1}^{N}\tilde{q}\left(x^j\right)} = \sum_{i=1}^{N} f\left(x^i\right)q\left(x^i\right) \tag{7.4-6}$$

where the normalized weights $q\left(x^i\right)$ are given by

$$q\left(x^i\right) = \frac{\tilde{q}\left(x^i\right)}{\sum_{j=1}^{N}\tilde{q}\left(x^j\right)}. \tag{7.4-7}$$

In the above, $\pi(x)$ is called the importance density, $\tilde{q}\left(x^i\right)$ the importance weights, and $q\left(x^i\right)$ the normalized importance weights.

Sequential Importance Sampling

Let $X_k = \left\{x_j, j = 0, \cdots, k\right\}$ represent the sequence of all target states up to time *k*. The joint posterior density at time k is denoted by $p\left(X_k \middle| Z_k\right)$, and its marginal is $p\left(x_k \middle| Z_k\right)$. Let $\left\{X_k^i, q_k^i\right\}_{i=1}^N$ denote a random measure that characterizes the joint posterior $p\left(X_k \middle| Z_k\right)$, where $\left\{X_k^i, i=1,\cdots,N\right\}$ is a set of support points with associated weights $\left\{q_k^i, i=1,\cdots,N\right\}$. Then, the joint posterior density at k can be approximated as follows,

$$p\left(X_k \middle| Z_k\right) \approx \sum_{i=1}^{N} q_k^i \delta\left(X_k - X_k^i\right). \tag{7.4-8}$$

We therefore have a discrete weighted approximation of the true posterior, $p\left(X_k \middle| Z_k\right)$. The normalized weights q_k^i are chosen using the principle of importance sampling described earlier. If the samples X_k^i were drawn from an importance density $\pi\left(X_k \middle| Z_k\right)$, then according to Eq. (7.4-5)

$$q_k^i \propto \frac{p\left(X_k^i \middle| Z_k\right)}{\pi\left(X_k^i \middle| Z_k\right)}. \tag{7.4-9}$$

Suppose at time step $k-1$ we have samples constituting an approximation to $p\left(X_{k-1} \middle| Z_{k-1}\right)$. With the reception of measurement z_k at time k, we wish to approximate $p\left(X_k \middle| Z_k\right)$ with a new set of samples. If the importance density is chosen to factorize such that

$$\pi\left(X_k \middle| Z_k\right) = \pi\left(x_k, X_{k-1} \middle| Z_k\right) = \pi\left(x_k \middle| X_{k-1}, Z_k\right)\pi\left(X_{k-1} \middle| Z_{k-1}\right) \tag{7.4-10}$$

Then one can obtain samples $X_k^i \sim \pi\left(X_k \middle| Z_k\right)$ by augmenting each of the existing samples $X_{k-1}^i \sim \pi\left(X_{k-1} \middle| Z_{k-1}\right)$ with the new state $x_k^i \sim \pi\left(x_k \middle| X_{k-1}, Z_k\right)$. To derive the weight update equation, the pdf $p\left(X_k \middle| Z_k\right)$ is first expressed using the Bayes' rule

$$\begin{aligned}
p\left(X_k \middle| Z_k\right) &= \frac{p\left(Z_k \middle| X_k\right) p\left(X_k\right)}{p\left(Z_k\right)} = \frac{p\left(z_k, Z_{k-1} \middle| X_k\right) p\left(X_k\right)}{p\left(Z_k\right)} \\
&= \frac{p\left(z_k \middle| X_k, Z_{k-1}\right) p\left(Z_{k-1} \middle| X_k\right) p\left(X_k\right)}{p\left(Z_k\right)} \\
&= \frac{p\left(z_k \middle| X_k, Z_{k-1}\right) p\left(X_k \middle| Z_{k-1}\right) p\left(Z_{k-1}\right) p\left(X_k\right)}{p\left(Z_k\right) p\left(X_k\right)} \\
&= \frac{p\left(z_k \middle| X_k, Z_{k-1}\right) p\left(X_k \middle| Z_{k-1}\right) p\left(Z_{k-1}\right)}{p\left(z_k \middle| Z_{k-1}\right) p\left(Z_{k-1}\right)}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{p\left(z_k \middle| X_k, Z_{k-1}\right) p\left(x_k \middle| X_{k-1}, Z_{k-1}\right) p\left(X_{k-1} \middle| Z_{k-1}\right)}{p\left(z_k \middle| Z_{k-1}\right)} \\
&= \frac{p\left(z_k \middle| x_k\right) p\left(x_k \middle| x_{k-1}\right) p\left(X_{k-1} \middle| Z_{k-1}\right)}{p\left(z_k \middle| Z_{k-1}\right)} && (7.4\text{-}11) \\
&\propto p\left(z_k \middle| x_k\right) p\left(x_k \middle| x_{k-1}\right) p\left(X_{k-1} \middle| Z_{k-1}\right) && (7.4\text{-}12)
\end{aligned}
$$

By substituting Eqs. (7.4-10) and (7.4-12) into Eq. (7.4-9), the weight update equation can then be shown to be

$$
\begin{aligned}
q_k^i &\propto \frac{p\left(z_k^i \middle| x_k^i\right) p\left(x_k^i \middle| x_{k-1}^i\right) p\left(X_{k-1}^i \middle| Z_{k-1}\right)}{\pi\left(x_k^i \middle| X_{k-1}^i, Z_k\right) \pi\left(X_{k-1}^i \middle| Z_{k-1}\right)} \\
&= q_{k-1}^i \frac{p\left(z_k^i \middle| x_k^i\right) p\left(x_k^i \middle| x_{k-1}^i\right)}{\pi\left(x_k^i \middle| X_{k-1}^i, Z_k\right)}.
\end{aligned}
\qquad (7.4\text{-}13)
$$

Furthermore, if $\pi\left(x_k \middle| X_{k-1}, Z_k\right) = \pi\left(x_k \middle| x_{k-1}, z_k\right)$, i.e., Markov process, then the importance density becomes only dependent on the x_{k-1} and z_k. The modified weight is then

$$q_k^i \propto q_{k-1}^i \frac{p\left(z_k^i \middle| x_k^i\right) p\left(x_k^i \middle| x_{k-1}^i\right)}{\pi\left(x_k^i \middle| x_{k-1}^i, z_k\right)} \tag{7.4-14}$$

and the posterior filtered density $p\left(x_k \middle| Z_k\right)$ can be approximated as

$$p\left(x_k \middle| Z_k\right) \approx \sum_{i=1}^{N} q_k^i \delta\left(x_k - x_k^i\right). \tag{7.4-15}$$

It can be shown that as $N \to \infty$, the approximation, Eq. (7.4-14), approaches the true posterior density $p\left(x_k \middle| Z_k\right)$.

Particle Filtering

Consider a nonlinear system described by the equations

$$\begin{aligned} x_{k+1} &= f_k\left(x_k, w_k\right) \\ z_k &= h_k\left(x_k, v_k\right) \end{aligned} \tag{7.4-16}$$

where k is the time index, x_k is the state, w_k is the process noise, z_k is the measurement, and v_k

is the measurement noise. w_k and v_k are assumed to be independent and white with known pdf's. The particle filtering is to numerically implement a Baysian estimation using the sequential importance sampling.

At the beginning of the estimation problem, we randomly generate a given number N state vectors based on the initial pdf $p\left(x_0\right)$ which is assumed to be known. These state vectors are called particles and are denoted as $x_0^i(+)\ (i=1,\cdots,N)$. At each time step $k=1,2,\cdots,$ we propagate the particles to the next time step using the process equation $f(\cdot)$

$$x_k^i(-)= f_{k-1}\left(x_{k-1}^i(+), w_{k-1}^i\right)\ (i=1,\cdots,N) \qquad (7.4\text{-}17)$$

where each w_{k-1}^i noise vector is randomly generated on the basis of the known pdf of w_{k-1}.

Now, we want to talk about obtaining the normalized weights q_k^i. Referring to Eq. (7.4-14), compute q_k^i from evaluating $p\left(z_k^i \middle| x_k^i(-)\right)$. Thus it may be called as the relative likelihood. For example, if $z_k = h\left(x_k\right)+v_k$ and $v_k \sim N(0,R)$, then for a specific measurement z^*,

$$q_k^i = P\left[\left(z_k = z^*\right)\middle|\left(x_k = x_k^i(-)\right)\right] = P\left[v_k = z^* - h\left(x_k^i(-)\right)\right]$$

$$\sim \frac{1}{(2\pi)^{\frac{m}{2}}\left|R\right|^{\frac{1}{2}}}\exp\left(\frac{-\left[z^* - h\left(x_k^i(-)\right)\right]^T R^{-1}\left[z^* - h\left(x_k^i(-)\right)\right]}{2}\right). \tag{7.4-18}$$

Next we resample the particles from the computed weights. That is, we compute a brand new set of particles $x_k^i(+)$ that are randomly generated on the basis of q_k^i. One straightforward way to obtain $x_k^i(+)$ is as follows:

1. Generate a random number *r* that is uniformly distributed on $[0,1]$.
2. Accumulate q_k^i into a sum, one at a time, until the accumulated sum is greater than *r*. That is, $\sum_{m=1}^{j-1} q_k^m < r$ but $\sum_{m=1}^{j} q_k^m \geq r$. The new particle $x_k^i(+)$ is then set equal to the old particle $x_k^j(-)$, i.e.,

$$x_k^i(+) = x_k^j(-) \text{ with probability } q_j \quad (i, j = 1, \cdots, N). \tag{7.4-19}$$

This is illustrated in Fig. 7.4-1.

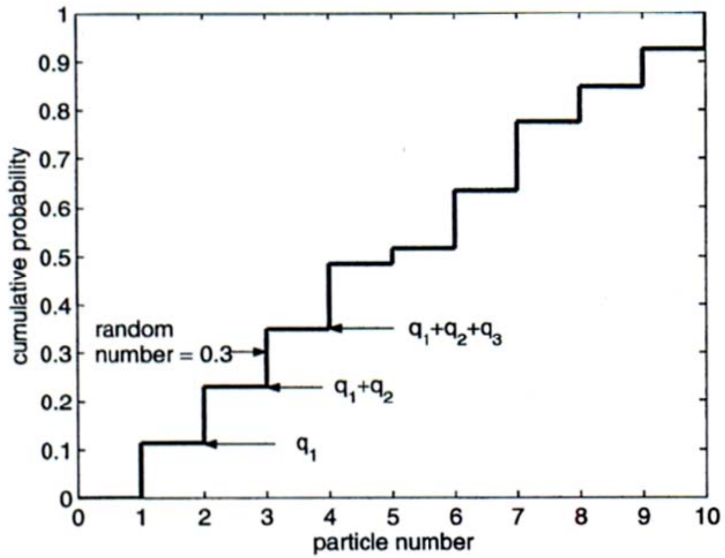

Figure 7.4-1 Illustration of resampling in the particle filter.

For example, if a random number $r = 0.3$ is generated from a distribution that is uniform on [0,1], the smallest vajue of j for which $\sum_{m=1}^{j} q_k^m \geq r$ is $j = 3$. Therefore the resampled particle is set equal to $x_k^3(-)$.

The Particle Filter

1. The system and measurement equations are given as follows

$$\begin{aligned} x_{k+1} &= f_k\left(x_k, w_k\right) \\ z_k &= h_k\left(x_k, v_k\right) \end{aligned} \qquad (7.4\text{-}20)$$

   where w_k and v_k are assumed to be independent and white with known pdf's.

2. Randomly generate N initial particles on the basis of the pdf $p\left(x_0\right)$ and denote them as $x_0^i(+)\ (i=1,\cdots,N)$.

3. For $k=1,2,\cdots,$ do the following.

   (a) Perform the time propagation step to obtain a priori particles $x_k^i(-)$

$$x_k^i(-) = f_{k-1}\left(x_{k-1}^i(+), w_{k-1}^i\right)\ (i=1,\cdots,N). \qquad (7.4\text{-}21)$$

   (b) Compute the relative likelihood q_k^i of each particle $x_k^i(-)$ conditioned on the measurement z_k. This is done by evaluating the pdf $p\left(z_k^i \middle| x_k^i(-)\right)$ on the basis of the nonlinear measurement equation and the pdf of the measurement noise.

(c) Scale the relative likelihoods obtained in the previous step as follows

$$q_k^i = \frac{q_k^i}{\sum_{j=1}^{N} q_k^j}$$

Now the sum of all the likelihoods is equal to one.

(d) Generate a set of a posteriori particles $x_k^i(+)$ on the basis of the relative likelihoods q_k^i. This is called the resampling step (for example, see Fig. 7.4-1)

(e) Using the set of particles $x_k^i(+)$, we can compute the mean and covariance,

$$
\begin{aligned}
&E\left\{x_k \middle| z_k\right\} = \frac{1}{N}\sum_{i=1}^{N} x_k^i(+) \\
&Cov\left\{x_k \middle| z_k\right\} = \frac{1}{N-1}\sum_{i=1}^{N}\left(x_k^i(+) - E\left\{x_k \middle| z_k\right\}\right)\left(x_k^i(+) - E\left\{x_k \middle| z_k\right\}\right)^T.
\end{aligned}
\tag{7.4-22}
$$

The Extended Kalman Particle Filter

1. The system and measurement equations are given as follows

$$
\begin{aligned}
x_{k+1} &= f_k\left(x_k, w_k\right) \\
z_k &= h_k\left(x_k, v_k\right)
\end{aligned}
\tag{7.4-20}
$$

   where w_k and v_k are assumed to be independent and white with known pdf's.

2. Randomly generate *N* initial particles on the basis of the pdf $p\left(x_0\right)$ and denote them as $x_0^i(+)$ and their covariances $P_0^i(+) = P_0(+)\ (i = 1, \cdots, N)$.

3. For $k = 1, 2, \cdots,$ do the following.

   (a) Perform the time propagation step to obtain a priori particles $x_k^i(-)$ and covariances

$P_k^i(-)$ using

$$x_k^i(-) = f_{k-1}\left(x_{k-1}^i(+), w_{k-1}^i\right)$$

$$P_k^i(-) = F_{k-1}^i P_{k-1}^i(+)\left(F_{k-1}^i\right)^T + Q_{k-1} \qquad (7.4\text{-}23)$$

$$F_{k-1}^i = \left.\frac{\partial f}{\partial x}\right|_{x=x_{k-1}^i(+)} \qquad (i = 1, \cdots, N)$$

where each w_{k-1}^i noise vector is randomly generated on the basis of the known pdf of w_{k-1}.

(b) Update the a priori particles and covariances to obtain a posteriori particles and covariances

$$H_k^i = \left.\frac{\partial h}{\partial x}\right|_{x=x_k^i(-)}$$

$$K_k^i = P_k^i(-)\left(H_k^i\right)^T\left(H_k^i P_k^i(-)\left(H_k^i\right)^T + R_k\right)^{-1}$$

$$x_k^i(+) = x_k^i(-) + K_k^i\left[z_k - h\left(x_k^i(-)\right)\right] \qquad (7.4\text{-}24)$$

$$P_k^i(+) = \left(I - K_k^i H_k^i\right) P_k^i(-).$$

(c) Compute the relative likelihood q_k^i of each particle $x_k^i(-)$ conditioned on the measurement z_k. This is done by evaluating the pdf $p\left(z_k^i \middle| x_k^i(-)\right)$ on the basis of the nonlinear measurement equation and the pdf fo the measurement noise.

(d) Scale the relative likelihoods obtained in the previous step as follows

$$q_k^i = \frac{q_k^i}{\sum_{j=1}^{N} q_k^j}$$

Now the sum of all the likelihoods is equal to one.

(e) Refine the set of a posteriori particles $x_k^i(+)$ and covariances $P_k^i(+)$ on the basis of the relative likelihoods q_k^i. This is the resampling step.

(f) Now we have a set of a posteriori particles $x_k^i(+)$ and covariances $P_k^i(+)$. We can compute any desired statistical measure of this set of particles.