APPENDIX D

Additional Topics on State Estimation

Paul D. Groves

This appendix comprises additional topics on state estimation, building on Chapter 3. Section D.1 describes least-squares estimation. Section D.2 summarizes the Schmidt-Kalman filter. Section D.3 then provides further details of the particle filter, covering resampling, importance sampling, and hybrid filtering.

D.1 Least-Squares Estimation

Least-squares estimation forms the basis of single-epoch positioning in GNSS and other position-fixing systems. This section derives the unweighted and weighted least-squares estimation equations together with the residuals and their covariance matrix. It also summarizes the iterated least-squares algorithm. Kalman filter-based notation and sign conventions are used.

Consider the linear equation

$$\delta\mathbf{z}^- = \mathbf{H}\delta\mathbf{x} + \delta\mathbf{z}^+ , \tag{D.1}$$

where $\delta\mathbf{z}^-$ is the measurement innovation, defined as

$$\delta\mathbf{z}^- = \mathbf{z} - \mathbf{h}\left(\hat{\mathbf{x}}^-\right), \tag{D.2}$$

$\delta\mathbf{z}^+$ is the measurement residual, defined as

$$\delta\mathbf{z}^+ = \mathbf{z} - \mathbf{h}\left(\hat{\mathbf{x}}^+\right), \tag{D.3}$$

$\delta\mathbf{x}$ is the state vector innovation, defined as

$$\delta\mathbf{x} = \hat{\mathbf{x}}^+ - \hat{\mathbf{x}}^- , \tag{D.4}$$

and $\mathbf{H}$ is the measurement, observation or design matrix, which is the same as the geometry matrix for positioning problems. This is defined as

$$\mathbf{H} = \frac{d\mathbf{h}}{d\mathbf{x}} = \frac{d\mathbf{z}}{d\mathbf{x}} , \tag{D.5}$$

where $\mathbf{z}$ is the measurement vector, which is known; $\hat{\mathbf{x}}^-$ is the a priori state vector estimate, also known; $\mathbf{h}$ is the measurement function, which is a known model; and $\hat{\mathbf{x}}^+$ is the a posteriori state vector estimate, which is unknown and to be determined.

The aim is to determine the value of $\hat{\mathbf{x}}^+$ that is most consistent with the measurement vector. This is found by minimizing the sum of squares of the residuals. No assumptions about the statistical distributions of the states or measurements are made. Thus

$$\frac{\partial}{\partial \delta\mathbf{x}}\left(\delta\mathbf{z}^{+\mathrm{T}}\delta\mathbf{z}^{+}\right) = 0\,. \tag{D.6}$$

Substituting in (D.1) and expanding,

$$\frac{\partial}{\partial \delta\mathbf{x}}\left(\delta\mathbf{x}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{H}\delta\mathbf{x} - \delta\mathbf{x}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\delta\mathbf{z}^{-} - \delta\mathbf{z}^{-\mathrm{T}}\mathbf{H}\delta\mathbf{x} + \delta\mathbf{z}^{-\mathrm{T}}\delta\mathbf{z}^{-}\right) = 0\,. \tag{D.7}$$

Applying (A.43) and (A.44) from Appendix A, also on CD,

$$2\delta\mathbf{x}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{H} - 2\delta\mathbf{z}^{-\mathrm{T}}\mathbf{H} = 0\,. \tag{D.8}$$

Rearranging and postmultiplying by $\left(\mathbf{H}^{\mathrm{T}}\mathbf{H}\right)^{-1}$,

$$\delta\mathbf{x}^{\mathrm{T}} = \delta\mathbf{z}^{-\mathrm{T}}\mathbf{H}\left(\mathbf{H}^{\mathrm{T}}\mathbf{H}\right)^{-1}. \tag{D.9}$$

Finally, performing a matrix transpose,

$$\delta\mathbf{x} = \left(\mathbf{H}^{\mathrm{T}}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathrm{T}}\delta\mathbf{z}^{-}\,. \tag{D.10}$$

The remainder of the section describes iterated least squares, weighted least squares, and the computation of the residuals and their covariance.

D.1.1 Iterated Least Squares

Where the measurement matrix, **H**, is independent of **x**, the problem is a linear function of **x**, so the above solution may be considered exact. Where **H**, is a function of **x**, the problem is a linearized approximation to a nonlinear problem and **H** is approximate as it is calculated using the a priori estimate of the states, $\hat{\mathbf{x}}^-$, instead of the true values, which are unknown. However, a better approximation of **H** may be calculated using the a posteriori estimate, $\hat{\mathbf{x}}^+$, once available.

Therefore, an iterated least squares (ILS) algorithm may obtain a more accurate solution of **x** by setting $\hat{\mathbf{x}}^-$ to $\hat{\mathbf{x}}^+$, recalculating **H** and $\delta\mathbf{z}^-$, and repeating the least-squares estimation process. Provided the problem converges, $\hat{\mathbf{x}}^+ - \hat{\mathbf{x}}^-$ will become much smaller on each iteration. Once each element of $\hat{\mathbf{x}}^+ - \hat{\mathbf{x}}^-$ is smaller than the precision required, the solution may be considered converged and the ILS algorithm ends.

D.1.2 Weighted Least Squares

In a weighted least-squares problem, some measurements are considered more important than others, typically because they are considered more accurate. Minimization of the residuals of these more important measurements is thus prioritized by minimizing a weighted sum of the squares of the residuals:

$$\frac{\partial}{\partial \delta\mathbf{x}}\left(\delta\mathbf{z}^{+\mathrm{T}}\mathbf{W}\delta\mathbf{z}^{+}\right) = 0\,, \tag{D.11}$$

where **W** is the weighting matrix. Substituting in (D.1) and expanding,

$$\frac{\partial}{\partial \delta\mathbf{x}}\left(\delta\mathbf{x}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{W}\mathbf{H}\delta\mathbf{x} - \delta\mathbf{x}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{W}\delta\mathbf{z}^{-} - \delta\mathbf{z}^{-\mathrm{T}}\mathbf{W}\mathbf{H}\delta\mathbf{x} + \delta\mathbf{z}^{-\mathrm{T}}\mathbf{W}\delta\mathbf{z}^{-}\right) = 0\,. \tag{D.12}$$

Applying (A.43) and (A.44) from Appendix A, also on CD,

$$2\delta\mathbf{x}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{W}\mathbf{H} - 2\delta\mathbf{z}^{-\mathrm{T}}\mathbf{W}\mathbf{H} = 0\,. \tag{D.13}$$

Rearranging and postmultiplying by $\left(\mathbf{H}^{\mathrm{T}}\mathbf{W}\mathbf{H}\right)^{-1}$,

$$\delta\mathbf{x}^{\mathrm{T}} = \delta\mathbf{z}^{-\mathrm{T}}\mathbf{W}\mathbf{H}\left(\mathbf{H}^{\mathrm{T}}\mathbf{W}\mathbf{H}\right)^{-1}. \qquad \text{(D.14)}$$

Finally, performing a matrix transpose,

$$\delta\mathbf{x} = \left(\mathbf{H}^{\mathrm{T}}\mathbf{W}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{W}\delta\mathbf{z}^{-}. \qquad \text{(D.15)}$$

The inverse of the measurement error covariance matrix, $\mathbf{C}_{\mathbf{z}}$, maybe used as the weighting matrix. Thus,

$$\delta\mathbf{x} = \left(\mathbf{H}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\delta\mathbf{z}^{-}. \qquad \text{(D.16)}$$

A weighted least-squares algorithm may also be iterated where appropriate.

D.1.3 Least Squares Residuals and their Covariance

From (D.1) and (D.16), the weighted least squares residuals are

$$\begin{aligned}\delta\mathbf{z}^{+} &= \delta\mathbf{z}^{-} - \mathbf{H}\delta\mathbf{x} \\ &= \left[\mathbf{I} - \mathbf{H}\left(\mathbf{H}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\right]\delta\mathbf{z}^{-}\end{aligned}, \qquad \text{(D.17)}$$

The covariance of the residuals is therefore

$$\begin{aligned}\mathbf{C}_{\delta\mathbf{z}}^{+} &= \mathrm{E}\left(\delta\mathbf{z}^{+}\delta\mathbf{z}^{+\mathrm{T}}\right) \\ &= \left[\mathbf{I} - \mathbf{H}\left(\mathbf{H}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\right]\mathrm{E}\left(\delta\mathbf{z}^{-}\delta\mathbf{z}^{-\mathrm{T}}\right)\left[\mathbf{I} - \mathbf{H}\left(\mathbf{H}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\right]^{\mathrm{T}}\end{aligned}, \qquad \text{(D.18)}$$

In least-squares estimation, the measurement error covariance, $\mathbf{C}_{\mathbf{z}}$, is also the covariance of the measurement innovation because the a priori state vector estimate, $\hat{\mathbf{x}}^{-}$, is treated as a known quantity. Therefore

$$\begin{aligned}\mathbf{C}_{\delta\mathbf{z}}^{+} &= \left[\mathbf{I} - \mathbf{H}\left(\mathbf{H}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\right]\mathbf{C}_{\mathbf{z}}\left[\mathbf{I} - \mathbf{H}\left(\mathbf{H}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\right]^{\mathrm{T}} \\ &= \mathbf{C}_{\mathbf{z}} - 2\mathbf{H}\left(\mathbf{H}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathrm{T}} + \mathbf{H}\left(\mathbf{H}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{H}\left(\mathbf{H}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathrm{T}} \\ &= \mathbf{C}_{\mathbf{z}} - \mathbf{H}\left(\mathbf{H}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathrm{T}}\end{aligned}. \qquad \text{(D.19)}$$

D.2 Schmidt-Kalman Filter

The Schmidt-Kalman filter with uncertain parameters is used to model time-correlated system and measurement noise parameters without estimating them as states [1].

The covariance matrix of the unestimated parameters that account for time-correlated system and measurement noise is **W**. These parameters are sometimes known as consider states. The correlation matrix, **U**, models the correlation between the unestimated parameters and the states. The transition matrix and system noise covariance matrix for the correlated noise parameters are, respectively, $\mathbf{\Phi}_U$ and $\mathbf{Q}_U$. An additional transition matrix, $\mathbf{\Psi}$, models how the states vary with the unestimated parameters. Consequently, the basic error covariance system propagation equation, (3.15), is replaced by

$$\begin{pmatrix}\mathbf{P}_k^{-} & \mathbf{U}_k^{-} \\ \mathbf{U}_k^{-\mathrm{T}} & \mathbf{W}_k^{-}\end{pmatrix} = \begin{pmatrix}\mathbf{\Phi}_{k-1} & \mathbf{\Psi}_{k\text{-}1} \\ 0 & \mathbf{\Phi}_{U,k-1}\end{pmatrix}\begin{pmatrix}\mathbf{P}_{k-1}^{+} & \mathbf{U}_{k-1}^{+} \\ \mathbf{U}_{k-1}^{+\ \mathrm{T}} & \mathbf{W}_{k-1}^{+}\end{pmatrix}\begin{pmatrix}\mathbf{\Phi}_{k-1}^{\mathrm{T}} & 0 \\ \mathbf{\Psi}_{k-1}^{\mathrm{T}} & \mathbf{\Phi}_{U,k-1}^{\mathrm{T}}\end{pmatrix} + \begin{pmatrix}\mathbf{Q}_{k-1} & 0 \\ 0 & \mathbf{Q}_{U,k-1}\end{pmatrix}. \qquad \text{(D.20)}$$

This separates into different propagation equations for the error covariance, correlated noise covariance, and correlation matrices, noting that the state propagation equation, (3.14), is unchanged from the standard Kalman filter:

$$\mathbf{P}_k^{-} = \mathbf{\Phi}_{k-1}\mathbf{P}_{k-1}^{+}\mathbf{\Phi}_{k-1}^{\mathrm{T}} + \mathbf{\Phi}_{k-1}\mathbf{U}_{k-1}^{+}\mathbf{\Psi}_{k-1}^{\mathrm{T}} + \mathbf{\Psi}_{k\text{-}1}\mathbf{U}_{k-1}^{+\ \mathrm{T}}\mathbf{\Phi}_{k-1}^{\mathrm{T}} + \mathbf{\Psi}_{k\text{-}1}\mathbf{W}_{k-1}^{+}\mathbf{\Psi}_{k-1}^{\mathrm{T}} + \mathbf{Q}_{k-1}. \qquad \text{(D.21)}$$

$$\mathbf{W}_k^- = \mathbf{\Phi}_{U,k-1}\mathbf{W}_{k-1}^+\mathbf{\Phi}_{U,k-1}^{\mathrm{T}} + \mathbf{Q}_{U,k-1}\,. \tag{D.22}$$

$$\mathbf{U}_k^- = \mathbf{\Phi}_{k-1}\mathbf{U}_{k-1}^+\mathbf{\Phi}_{U,k-1}^{\mathrm{T}} + \mathbf{\Psi}_{k\text{-}1}\mathbf{W}_{k-1}^+\mathbf{\Phi}_{U,k-1}^{\mathrm{T}}\,. \tag{D.23}$$

Defining **J** as the matrix coupling the unestimated parameters to the measurement vector, corresponding to the **H** matrix for the estimated states, the Kalman gain becomes

$$\mathbf{K}_k = \left(\mathbf{P}_k^-\mathbf{H}_k^{\mathrm{T}} + \mathbf{U}_k^-\mathbf{J}_k^{\mathrm{T}}\right)\left(\mathbf{H}_k\mathbf{P}_k^-\mathbf{H}_k^{\mathrm{T}} + \mathbf{H}_k\mathbf{U}_k^-\mathbf{J}_k^{\mathrm{T}} + \mathbf{J}_k\mathbf{U}_k^{-\mathrm{T}}\mathbf{H}_k^{\mathrm{T}} + \mathbf{J}_k\mathbf{W}_k^-\mathbf{J}_k^{\mathrm{T}} + \mathbf{R}_k\right)^{-1}. \tag{D.24}$$

Apart from the redefinition of the Kalman gain, the state update equation, (3.24), is unchanged. The error covariance, correlated noise, and correlation matrices update as follows:

$$\mathbf{P}_k^+ = \left(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k\right)\mathbf{P}_k^- - \mathbf{K}_k\mathbf{J}_k\mathbf{U}_k^{-\mathrm{T}}, \tag{D.25}$$

$$\mathbf{W}_k^+ = \mathbf{W}_k^-, \tag{D.26}$$

$$\mathbf{U}_k^+ = \left(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k\right)\mathbf{U}_k^- - \mathbf{K}_k\mathbf{J}_k\mathbf{W}_k^-. \tag{D.27}$$

Where none of the unestimated parameters are directly correlated with the measurements, there is only time-correlated system noise and $\mathbf{J} = 0$. Similarly, if none of the states vary as a function of the unestimated parameters, there is only time-correlated measurement noise and $\mathbf{\Psi} = 0$. Note that an unestimated parameter can impact both the system and measurement models in the same manner that a state can.

As the system propagation and measurement update of the error covariance matrix dominates the Kalman filter's processor load, the Schmidt-Kalman filter typically uses at least a quarter as much processing power for the time-correlated noise sources as a Kalman filter that estimates them as states. This goes up to at least half the Kalman filter's load where all of the parameters are correlated with the measurements. The processor load can be reduced by neglecting parts of the correlation matrix, **U**. However, great caution must then be taken to avoid filter instability.

D.3 Particle Filtering Topics

This section provides further details of the particle filter, building on Section 3.5. Resampling, importance sampling, and hybrid filtering are discussed.

D.3.1 Resampling

The aim of a particle resampling algorithm is to produce a set of output particles with a more even probability distribution between particles than the set of input particles, but without introducing biases. Although resampling may be entirely rule based, most approaches incorporate a random element.

Most resampling methods use the cumulative distribution function to match output particles to input particles, attributing equal probability to the output particles [2, 3]. Each output particle is allocated a CDF between 0 and 1. In random resampling, these are allocated randomly by sampling from a uniform distribution. In systematic resampling, a single random number, u, is drawn from a uniform distribution between 0 and 1. The CDFs are then allocated using

$$F_j = \frac{u + j - 1}{N}, \tag{D.28}$$

where N is the number of particles. In stratified resampling, each CDF, F_j, is allocated a value between $(j - 1) / N$ and j / N, selected by sampling from a uniform distribution. In all cases, each output particle, denoted by j, is then set to the input particle, denoted by i, that obeys

$$\sum_{r=1}^{i} p_{\mathbf{X}}^{(r)} < F_j \leq \sum_{r=1}^{i+1} p_{\mathbf{X}}^{(r)} , \tag{D.29}$$

where $p_{\mathbf{X}}^{(r)}$ is the probability of the r^{th} input particle. Figure D.1 illustrates this for resampling five particles.

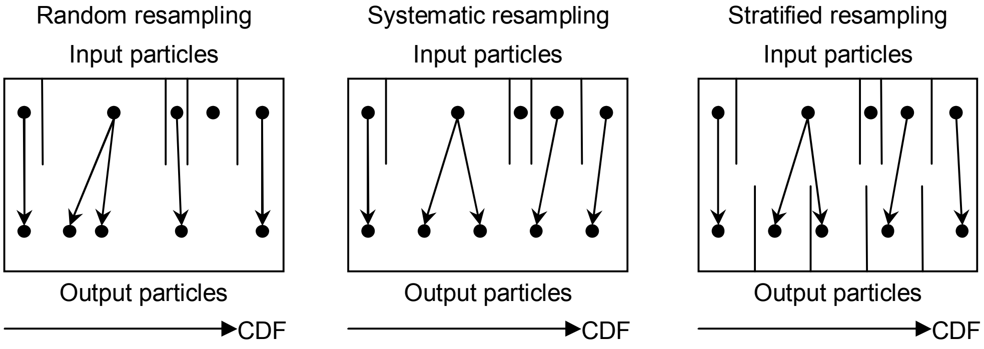

Figure D.1 CDF-based particle resampling.

Following resampling, some of the particles will be identical. The addition of system noise during the next system propagation phase will cause these particles to diverge. However, if the system noise is not large enough, there will be too many particles in some regions of the state-estimate probability distribution and insufficient particles in other regions. This is known as sample impoverishment or depletion. One remedy is simply to increase the modeled system noise. Alternatively, an additional roughening step may be added to the resampling process.

Roughening adds a noise term, $\Delta\mathbf{x}_k^{(i)}$, to each particle immediately after resampling. The noise comprises vector samples from a zero-mean distribution, usually Gaussian, with standard deviation $K\mathbf{m}_k N^{-1/n}$, where K is a tuning parameter of order 0.2, $\mathbf{m}_k$ is a vector comprising the maximum difference between state estimates across the particle set at the current epoch (determined independently for each state), and n is the number of states [2, 4].

D.3.2 Importance Sampling

Generating random samples from an arbitrary distribution is not straightforward. Importance sampling circumvents this by initially drawing a set of random samples from a distribution similar to the target that is easier to sample from. This sampling distribution may be a uniform distribution, a Gaussian distribution, a sum of Gaussian distributions, or some other distribution. Each sample, $\mathbf{x}^{(i)}$, is then given a likelihood of

$$\Lambda_{\mathbf{X}}^{(i)} = \frac{f_{\mathbf{X}}\left(\mathbf{x}^{(i)}\right)}{f_{\mathbf{S}}\left(\mathbf{x}^{(i)}\right)}, \tag{D.30}$$

where $f_{\mathbf{X}}$ is the PDF of the target distribution and $f_{\mathbf{S}}$ is the PDF of the sampling distribution. Particle probabilities may be obtained by dividing each likelihood by the sum of likelihoods across all particles. If N_{eff} is too small or equal-probability particles are desired, importance sampling may be followed by resampling as described in Section D.3.1.

D.3.3 Hybrid Filtering

Particle filtering and Kalman filtering may be combined in a number of different ways. Here, the extended particle filter (EPF), unscented particle filter (UPF), and Rao-Blackwellized particle filter (RBPF) are discussed. However, there are other methods.

In an extended particle filter, each particle is an EKF. Similarly, in an unscented particle filter, each particle is a UKF. The system propagation and measurement update of each particle is thus performed in the same way as for an EKF (Section 3.4.1) or UKF (Section 3.4.2). This requires considerably more processing capacity per particle than a basic particle filter. However, as each particle occupies a region of the state estimate probability distribution, instead of a point, fewer particles should be required.

The likelihood of each particle in an EPF or UPF is based on an estimate of the measurement vector obtained from the state vector estimate after the EKF/UKF measurement update. Thus,

$$\begin{aligned}\Lambda_{\mathbf{X},k}^{(i)} &= p_{\mathbf{X},k}^{-(i)} \tilde{f}_{\mathbf{Z},k}\left(\hat{\mathbf{z}}_k^{+(i)}\right) \\ &= p_{\mathbf{X},k}^{-(i)} \tilde{f}_{\mathbf{Z},k}\left[\mathbf{h}\left(\hat{\mathbf{x}}_k^{+(i)}\right)\right].\end{aligned} \tag{D.31}$$

Particle probabilities are calculated using (3.129) and resampling is performed in the same way as for a standard particle filter.

Because each particle is updated as an EKF or UKF, the state estimates of the different particles will converge, causing sample impoverishment. Therefore, noise must be added to separate the particles. There are two approaches. The first is to implement the particle filter version of the state vector propagation, whereby system noise is added to the state estimates [2]. Additional system noise may be required to achieve sufficient particle separation. The state estimation error covariance matrix is propagated using the EKF or UKF method.

The second way of separating the particles is to add a randomization step between the EKF/UKF measurement update and the likelihood determination. This resamples the state vector estimate from a multivariate Gaussian distribution of mean $\hat{\mathbf{x}}_k^{+(i)}$ and covariance $\mathbf{P}_k^{+(i)}$.

In a Rao-Blackwellized particle filter, particles are only distributed through the space of some of the states, with the remaining states determined using Kalman filter-based estimation. This enables fewer particles to be used. It is suited to applications where the system model is predominantly linear and Gaussian, but the measurement model is not.

The state vector is partitioned using

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}^P \\ \mathbf{x}^K \end{pmatrix}, \tag{D.32}$$

where $\mathbf{x}^P$ comprises the state space over which the particles are distributed and $\mathbf{x}^K$ comprises the remaining states, to which a Kalman filter system model is applied . In determining which partition to place each state, state propagation of $\mathbf{x}^P$, may depend on $\mathbf{x}^K$, but not vice versa.

The measurement update and resampling phases of a RBPF are the same as those of a standard particle filter. However, the system propagation phase is different. The system model may be represented as [5]

$$\begin{pmatrix} \mathbf{x}_k^P \\ \mathbf{x}_k^K \end{pmatrix} = \begin{pmatrix} \boldsymbol{\varphi}_{k-1}^P\left(\mathbf{x}_{k-1}^P\right) \\ \mathbf{0} \end{pmatrix} + \begin{pmatrix} \mathbf{\Phi}_{k-1}^{PK} \\ \mathbf{\Phi}_{k-1}^K \end{pmatrix} \mathbf{x}_{k-1}^K + \begin{pmatrix} \mathbf{\Gamma}_{k-1}^P \\ \mathbf{\Gamma}_{k-1}^K \end{pmatrix} \mathbf{w}_{s,k-1}, \tag{D.33}$$

where $\boldsymbol{\varphi}_{k-1}^P$ is the transition function for the particle filter states, $\mathbf{\Phi}_{k-1}^K$ is the transition matrix for the Kalman filter states, $\mathbf{\Phi}_{k-1}^{PK}$ is the transition matrix for the propagation of the Kalman filter states into the particle filter states, $\mathbf{\Gamma}_{k-1}^P$ is the discrete system noise distribution matrix for the particle filter states, $\mathbf{\Gamma}_{k-1}^K$ is the discrete system noise distribution matrix for the Kalman filter states, and $\mathbf{w}_{s,k-1}$ is the discrete system noise vector. A state estimation error covariance matrix, $\mathbf{P}^K$, is maintained for the Kalman filter states only, this is common to all particles.

The first step of the system propagation is to determine a Kalman gain matrix relating the Kalman filter states (one per row), to the particle filter states (one per column). This is common to all particles and is given by

$$\mathbf{K}_k^K = \mathbf{P}_{k-1}^K \mathbf{\Phi}_{k-1}^{PK\,\mathrm{T}} \left(\mathbf{\Phi}_{k-1}^{PK} \mathbf{P}_{k-1}^K \mathbf{\Phi}_{k-1}^{PK\,\mathrm{T}} + \mathbf{\Gamma}_{k-1}^P \mathbf{W}_{k-1} \mathbf{\Gamma}_{k-1}^{P\,\mathrm{T}}\right)^{-1}, \tag{D.34}$$

where $\mathbf{W}_{k-1}$ is the covariance of the discrete system noise, i.e.,

$$\mathbf{W}_{k-1} = \mathrm{E}\left(\mathbf{w}_{s,k-1}\mathbf{w}_{s,k-1}^{\mathrm{T}}\right). \tag{D.35}$$

The error covariance matrix of the Kalman filter states is propagated using

$$\mathbf{P}_k^K = \left[\mathbf{\Phi}_{k-1}^K - \mathbf{\Gamma}_{k-1}^K\left(\mathbf{\Gamma}_{k-1}^{P\ \mathrm{T}}\mathbf{\Gamma}_{k-1}^P\right)^{-1}\mathbf{\Gamma}_{k-1}^{P\ \mathrm{T}}\mathbf{\Phi}_{k-1}^{PK}\right]\left(\mathbf{P}_{k-1}^K - \mathbf{K}_k^K\mathbf{\Phi}_{k-1}^{PK}\mathbf{P}_{k-1}^K\right)\left[\mathbf{\Phi}_{k-1}^K - \mathbf{\Gamma}_{k-1}^K\left(\mathbf{\Gamma}_{k-1}^{P\ \mathrm{T}}\mathbf{\Gamma}_{k-1}^P\right)^{-1}\mathbf{\Gamma}_{k-1}^{P\ \mathrm{T}}\mathbf{\Phi}_{k-1}^{PK}\right]^{\mathrm{T}}, \tag{D.36}$$

State propagation is on a particle-by-particle basis. The particle filter states are propagated using

$$\hat{\mathbf{x}}_k^{(i)P} = \boldsymbol{\varphi}_{k-1}^P\left(\hat{\mathbf{x}}_{k-1}^{(i)P}\right) + \mathbf{\Phi}_{k-1}^{PK}\hat{\mathbf{x}}_{k-1}^{(i)K} + \Delta\mathbf{x}_{k-1}^{(i)P}, \tag{D.37}$$

where $\Delta\mathbf{x}_{k-1}^{(i)P}$ is randomly sampled from a zero-mean multivariate Gaussian distribution with covariance $\mathbf{\Phi}_{k-1}^{PK}\mathbf{P}_{k-1}^K\mathbf{\Phi}_{k-1}^{PK\,\mathrm{T}} + \mathbf{\Gamma}_{k-1}^P\mathbf{W}_{k-1}\mathbf{\Gamma}_{k-1}^{P\ \mathrm{T}}$.

The remaining states are then propagated using

$$\hat{\mathbf{x}}_k^{(i)K} = \mathbf{\Phi}_{k-1}^K\left[\hat{\mathbf{x}}_{k-1}^{(i)K} + \mathbf{K}_k^K\left(\hat{\mathbf{x}}_k^{(i)P} - \hat{\mathbf{x}}_{k-1}^{(i)P} - \mathbf{\Phi}_{k-1}^{PK}\hat{\mathbf{x}}_{k-1}^{(i)K}\right)\right] + \mathbf{\Gamma}_{k-1}^K\left(\mathbf{\Gamma}_{k-1}^{P\ \mathrm{T}}\mathbf{\Gamma}_{k-1}^P\right)^{-1}\mathbf{\Gamma}_{k-1}^{P\ \mathrm{T}}\left(\hat{\mathbf{x}}_k^{(i)P} - \hat{\mathbf{x}}_{k-1}^{(i)P}\right). \tag{D.38}$$

To initialize a RBPF, the particle filter states for each particle are generated by sampling randomly from the initial distribution of the states. For the Kalman filter states, a single set of initial values is used in all particles. An initial covariance, $\mathbf{P}_0^K$, is also required.

References

[1] Schmidt, S. F., “Application of State Space Methods to Navigation Problems,” In *Advanced in Control Systems*, Vol. 3, C. T. Leondes (ed.), New York: Academic Press, 1966.

[2] Simon, D., *Optimal State Estimation*, New York: Wiley, 2006.

[3] Doucet, A., and A. M. Johansen, “A Tutorial on Particle Filtering and Smoothing: Fifteen years later,” In *Oxford Handbook of Nonlinear Filtering*, C. Crisan and B. Rozovsky (eds.), Oxford, UK: OUP, 2011, pp. 656–704.

[4] Gordon, N. J., D. J. Salmond, and A. F. M. Smith, “A Novel Approach to Nonlinear/Non-Gaussian Bayesian State Estimation,” *Proc. IEE Radar Signal Process.*, Vol. 140, 1993, pp. 107–113.

[5] Gustafsson, F. et al., “Particle Filters for Positioning, Navigation and Tracking,” *IEEE Trans. on Signal Processing*, Vol. 50, 2002, pp. 425–437.