

## 6 Das Kalman-Filter

Ein Kalman-Filter ist ein Algorithmus zur Schätzung des Zustandes eines linearen Systems. Dazu werden Messwerte verarbeitet, die in linearer Weise mit dem Systemzustand zusammenhängen müssen. Ist der Zusammenhang zwischen Messwerten und Systemzustand oder das beobachtete System schwach nichtlinear, kann ein erweitertes Kalman-Filter (EKF) eingesetzt werden, siehe Abschnitt 6.4.

Um eine Zustandsschätzung durchführen zu können, benötigt das Kalman-Filter ein Modell des beobachteten Systems in der Form

$$\vec{x}_{k+1} = \mathbf{\Phi}_k \vec{x}_k + \mathbf{B}_k \vec{u}_k + \mathbf{G}_k \vec{w}_k \quad (6.1)$$

Gegenüber dem bereits eingeführten Systemmodell Gl. (2.71) ist ein zusätzlicher Term hinzugekommen, der das System- oder Prozessrauschen  $\vec{w}_k$  enthält. Bei dem Systemrauschen handelt es sich um mittelwertfreies, normalverteiltes, weißes Rauschen; es gilt

$$E [\vec{w}_i \vec{w}_k^T] = \begin{cases} \mathbf{Q}_k & i = k \\ \mathbf{0} & i \neq k \end{cases} \quad (6.2)$$

Das Systemrauschen dient zur Berücksichtigung von Modellunsicherheiten. Damit kann der Tatsache Rechnung getragen werden, dass das reale System durch ein Modell der Form Gl. (2.71) nur näherungsweise beschrieben wird. Desweiteren ist es vorstellbar, dass die Eingangsgrößen  $\vec{u}_k$  mit Rauschen behaftet sind: auch diese Eingangsgrößen müssen in der Regel durch Messung ermittelt werden. Das Messrauschen dieser Eingangsgrößen kann dann dem Systemrauschen zugeschlagen werden. Eine Berücksichtigung von Messwerten als Eingangsgrößen und nicht als Messwerte im eigentlichen Sinne setzt voraus, dass keine Ableitungen dieser Größen im Zustandsvektor des Systemmodells auftreten.

Der Zusammenhang zwischen Messwerten  $\tilde{y}_k$  und dem Systemzustand  $\vec{x}_k$  wird durch ein Messmodell der Form

$$\tilde{y}_k = \mathbf{H}_k \vec{x}_k + \vec{v}_k \quad (6.3)$$

beschrieben. Die Matrix  $\mathbf{H}_k$  wird als Messmatrix bezeichnet, bei  $\vec{v}_k$  handelt es sich um das Messrauschen. Das Messrauschen wird – genau wie das Systemrauschen – als weiß, mittelwertfrei und normalverteilt angenommen; es gilt

$$E [\vec{v}_i \vec{v}_k^T] = \begin{cases} \mathbf{R}_k & i = k \\ \mathbf{0} & i \neq k \end{cases} \quad (6.4)$$

Als weitere Voraussetzung wird angenommen, dass die Kreuzkorrelation von Messrauschen und Systemrauschen verschwindet:

$$E [\vec{v}_k \vec{w}_k^T] = \mathbf{0} \quad (6.5)$$

Unter den angesprochenen Voraussetzungen – lineares System- und Messmodell, weißes, normalverteiltes, mittelwertfreies System- und Messrauschen – ist das Kalman-Filter optimal, d.h. kein anderer Algorithmus liefert im Mittel eine kleinere Summe der quadrierten Schätzfehler. Neben einer Schätzung des Systemzustandes liefert das Kalman-Filter auch noch Informationen bezüglich der Güte der Schätzung in Form einer Schätzfehlerkovarianzmatrix.

Die Gleichungen des Kalman-Filters werden im Folgenden hergeleitet.

## 6.1 Kalman-Filter-Gleichungen

Rudolf Emil Kalman leitete 1960 die nach ihm benannten Kalman-Filter-Gleichungen über ein Orthogonalitätsprinzip her, siehe [63]. Neben dieser sehr mathematischen Vorgehensweise gibt es auch leichter einsichtige Wege der Herleitung dieser Gleichungen, eine Möglichkeit besteht über die Formulierung linearer Operationen mit normalverteilten Zufallsvektoren.

### 6.1.1 Herleitung über normalverteilte Zufallsvektoren

Prinzipiell können die Kalman-Filter-Gleichungen in zwei Gruppen eingeteilt werden, den Estimationsschritt und den Propagationsschritt.

#### Estimation

Im Estimationsschritt des Kalman-Filters werden die zum aktuellen Zeitpunkt vorliegenden Messwerte verarbeitet. Dadurch wird die Schätzung des Systemzustandes verbessert, entsprechend muss auch die Kovarianzmatrix des Schätzfehlers angepasst werden. Es soll folgende Nomenklatur verwendet werden: Der tatsächliche Systemzustand wird mit  $\vec{x}_k$  bezeichnet, die Schätzung des Systemzustandes mit  $\hat{\vec{x}}_k$ . Obere Indizes  $+$  und  $-$  geben an, ob die zum aktuellen Zeitpunkt vorliegenden Messwerte bereits verarbeitet wurden  $(\cdot)^+$  oder die Messwertverarbeitung noch aussteht  $(\cdot)^-$ . Im ersten Fall spricht man von einem a-posteriori-Schätzwert, bei noch ausstehender Messwertverarbeitung handelt es sich um einen a-priori-Schätzwert. Die a-priori- und a-posteriori-Kovarianzmatrizen des Schätzfehlers sind gegeben durch

$$\mathbf{P}_{xx,k}^- = E \left[ (\vec{x}_k - \hat{\vec{x}}_k^-)(\vec{x}_k - \hat{\vec{x}}_k^-)^T \right] \quad (6.6)$$

$$\mathbf{P}_{xx,k}^+ = E \left[ (\vec{x}_k - \hat{\vec{x}}_k^+)(\vec{x}_k - \hat{\vec{x}}_k^+)^T \right] . \quad (6.7)$$

Grundlage für die Herleitung des Estimationsschrittes sind die Gleichungen (5.31) und (5.32), die sich mit der gerade eingeführten Nomenklatur wie folgt schreiben lassen:

$$\hat{\vec{x}}_k^+ = \hat{\vec{x}}_k^- + \mathbf{P}_{xy,k} \mathbf{P}_{yy,k}^{-1} (\tilde{y}_k - \hat{y}_k) \quad (6.8)$$

$$\mathbf{P}_{xx,k}^+ = \mathbf{P}_{xx,k}^- - \mathbf{P}_{xy,k} \mathbf{P}_{yy,k}^{-1} \mathbf{P}_{yx,k} \quad (6.9)$$

Um anhand der vorliegenden Messwerte  $\tilde{y}_k$  aus der a-priori-Zustandsschätzung und -schätzfehlerkovarianzmatrix die entsprechenden a-posteriori-Größen berechnen zu können, müssen die noch unbekannten Größen  $\hat{y}_k$ ,  $\mathbf{P}_{xy,k}$ ,  $\mathbf{P}_{yy,k}$  und  $\mathbf{P}_{yx,k}$  ermittelt werden.

Der Erwartungswert des Messwertvektors ergibt sich unter Berücksichtigung des Messmodells zu

$$\begin{aligned}\hat{y}_k &= E[\tilde{y}_k] \\ &= E[\mathbf{H}_k \vec{x}_k + \vec{v}_k] \\ &= \mathbf{H}_k E[\vec{x}_k] + E[\vec{v}_k] \\ \hat{y}_k &= \mathbf{H}_k \hat{x}_k^- .\end{aligned}\tag{6.10}$$

Die Kovarianzmatrix  $\mathbf{P}_{yy,k}$  ist gegeben durch

$$\mathbf{P}_{yy,k} = E[(\tilde{y}_k - \hat{y}_k)(\tilde{y}_k - \hat{y}_k)^T] .\tag{6.11}$$

Durch Einsetzen des Messmodells und Gl. (6.10) erhält man

$$\begin{aligned}\mathbf{P}_{yy,k} &= E[(\mathbf{H}_k \vec{x}_k + \vec{v}_k - \mathbf{H}_k \hat{x}_k^-)(\mathbf{H}_k \vec{x}_k + \vec{v}_k - \mathbf{H}_k \hat{x}_k^-)^T] \\ &= E[\mathbf{H}_k(\vec{x}_k - \hat{x}_k^-)(\vec{x}_k - \hat{x}_k^-)^T \mathbf{H}_k^T] \\ &\quad + E[\mathbf{H}_k(\vec{x}_k - \hat{x}_k^-)\vec{v}_k^T] + E[\vec{v}_k(\vec{x}_k - \hat{x}_k^-)^T \mathbf{H}_k^T] + E[\vec{v}_k \vec{v}_k^T]\end{aligned}\tag{6.12}$$

Da das Messrauschen  $\vec{v}_k$  vor Verarbeitung der Messwerte nicht mit dem Schätzfehler korreliert sein kann, verschwinden in obiger Gleichung der zweite und der dritte Erwartungswert, es ergibt sich damit

$$\begin{aligned}\mathbf{P}_{yy,k} &= \mathbf{H}_k E[(\vec{x}_k - \hat{x}_k^-)(\vec{x}_k - \hat{x}_k^-)^T] \mathbf{H}_k^T + E[\vec{v}_k \vec{v}_k^T] \\ &= \mathbf{H}_k \mathbf{P}_{xx,k}^- \mathbf{H}_k^T + \mathbf{R}_k .\end{aligned}\tag{6.13}$$

Die Matrix  $\mathbf{P}_{xy,k}$  ergibt sich analog zu

$$\begin{aligned}\mathbf{P}_{xy,k} &= E[(\vec{x}_k - \hat{x}_k^-)(\tilde{y}_k - \hat{y}_k)^T] \\ &= E[(\vec{x}_k - \hat{x}_k^-)(\mathbf{H}_k \vec{x}_k + \vec{v}_k - \mathbf{H}_k \hat{x}_k^-)^T] \\ &= E[(\vec{x}_k - \hat{x}_k^-)(\vec{x}_k - \hat{x}_k^-)^T \mathbf{H}_k^T] + E[(\vec{x}_k - \hat{x}_k^-)\vec{v}_k^T] \\ &= E[(\vec{x}_k - \hat{x}_k^-)(\vec{x}_k - \hat{x}_k^-)^T] \mathbf{H}_k^T \\ \mathbf{P}_{xy,k} &= \mathbf{P}_{xx,k}^- \mathbf{H}_k^T\end{aligned}\tag{6.14}$$

Setzt man Gl. (6.10), (6.13) und (6.14) in Gl. (6.8) ein, erhält man mit

$$\mathbf{P}_{y,k} = \mathbf{P}_{xy,k}^T = \mathbf{H}_k \mathbf{P}_{xx,k}^- \quad (6.15)$$

den gesuchten Zusammenhang zur Berechnung des a-posteriori-Schätzwertes:

$$\hat{x}_k^+ = \hat{x}_k^- + \mathbf{P}_{xx,k}^- \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_{xx,k}^- \mathbf{H}_k^T + \mathbf{R}_k)^{-1} (\tilde{y} - \mathbf{H}_k \hat{x}_k^-) \quad (6.16)$$

Es ist üblich, folgende Abkürzung einzuführen:

$$\mathbf{K}_k = \mathbf{P}_{xx,k}^- \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_{xx,k}^- \mathbf{H}_k^T + \mathbf{R}_k)^{-1} \quad (6.17)$$

Die Matrix  $\mathbf{K}_k$  wird als Kalman-Gain-Matrix bezeichnet. Damit lassen sich der a-posteriori-Schätzwert und die a-posteriori-Kovarianzmatrix des Schätzfehlers wie folgt schreiben:

$$\hat{x}_k^+ = \hat{x}_k^- + \mathbf{K}_k (\tilde{y} - \mathbf{H}_k \hat{x}_k^-) \quad (6.18)$$

$$\mathbf{P}_{xx,k}^+ = \mathbf{P}_{xx,k}^- - \mathbf{K}_k \mathbf{H}_k \mathbf{P}_{xx,k}^- = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_{xx,k}^- \quad (6.19)$$

Der Kalman-Filter-Estimationsschritt ist folglich durch die Gleichungen (6.17)–(6.19) gegeben. Alternativ zu (6.19) wird häufig eine andere Form des Kovarianzmatrix-Updates verwendet, die sogenannte Joseph's Form. Diese ist gegeben durch

$$\mathbf{P}_{xx,k}^+ = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_{xx,k}^- (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)^T + \mathbf{K}_k \mathbf{R}_k \mathbf{K}_k^T \quad (6.20)$$

Die Gl. (6.19) und (6.20) sind mathematisch äquivalent: Ausmultiplizieren und Zusammenfassen von Gl. (6.20) liefert

$$\begin{aligned} \mathbf{P}_{xx,k}^+ &= (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_{xx,k}^- (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)^T + \mathbf{K}_k \mathbf{R}_k \mathbf{K}_k^T \\ &= \mathbf{P}_{xx,k}^- - \mathbf{K}_k \mathbf{H}_k \mathbf{P}_{xx,k}^- - \mathbf{P}_{xx,k}^- \mathbf{H}_k^T \mathbf{K}_k^T \\ &\quad + \mathbf{K}_k \mathbf{H}_k \mathbf{P}_{xx,k}^- \mathbf{H}_k^T \mathbf{K}_k^T + \mathbf{K}_k \mathbf{R}_k \mathbf{K}_k^T \\ &= (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_{xx,k}^- - \mathbf{P}_{xx,k}^- \mathbf{H}_k^T \mathbf{K}_k^T \\ &\quad + \mathbf{K}_k (\mathbf{H}_k \mathbf{P}_{xx,k}^- \mathbf{H}_k^T + \mathbf{R}_k) \mathbf{K}_k^T. \end{aligned} \quad (6.21)$$

Setzt man in den letzten Term die Kalman-Gain-Matrix nach Gl. (6.17) ein, erhält man

$$\begin{aligned} \mathbf{P}_{xx,k}^+ &= (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_{xx,k}^- - \mathbf{P}_{xx,k}^- \mathbf{H}_k^T \mathbf{K}_k^T \\ &\quad + \mathbf{P}_{xx,k}^- \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_{xx,k}^- \mathbf{H}_k^T + \mathbf{R}_k)^{-1} (\mathbf{H}_k \mathbf{P}_{xx,k}^- \mathbf{H}_k^T + \mathbf{R}_k) \mathbf{K}_k^T \\ &= (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_{xx,k}^- - \mathbf{P}_{xx,k}^- \mathbf{H}_k^T \mathbf{K}_k^T + \mathbf{P}_{xx,k}^- \mathbf{H}_k^T \mathbf{K}_k^T \\ &= (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_{xx,k}^-, \end{aligned} \quad (6.22)$$

das Kovarianzmatrix-Update nach Gl. (6.19). Die Joseph's Form des Kovarianzmatrix-Updates ist numerisch robuster und sollte daher bevorzugt werden, wenn der zusätzliche Rechenaufwand toleriert werden kann.

## Propagation

Im Propagationsschritt wird der geschätzte Systemzustand in der Zeit propagiert, d.h. aus dem a-posteriori-Schätzwert zum Zeitpunkt  $k$  wird der a-priori-Schätzwert zum nächsten Zeitpunkt, also  $k+1$ , berechnet. Dabei wächst die Unsicherheit bezüglich der Zustandsschätzung und die Schätzfehlerkovarianzmatrix muss entsprechend angepasst werden.

Die Zustandsschätzung zum nächsten Zeitpunkt ist gegeben durch

$$\hat{\vec{x}}_{k+1}^- = E[\vec{x}_{k+1}] , \quad (6.23)$$

durch Einsetzen des Systemmodells erhält man

$$\begin{aligned} \hat{\vec{x}}_{k+1}^- &= E[\Phi_k \vec{x}_k + \mathbf{B}_k \vec{u}_k + \mathbf{G}_k \vec{w}_k] \\ \hat{\vec{x}}_{k+1}^- &= \Phi_k \hat{\vec{x}}_k^+ + \mathbf{B}_k \vec{u}_k . \end{aligned} \quad (6.24)$$

Die zugehörige Schätzfehlerkovarianzmatrix ergibt sich mit Gl. (6.1) und Gl. (6.24) zu

$$\begin{aligned} \mathbf{P}_{xx,k+1}^- &= E[(\vec{x}_{k+1} - \hat{\vec{x}}_{k+1}^-)(\vec{x}_{k+1} - \hat{\vec{x}}_{k+1}^-)^T] \\ &= E[(\Phi_k \vec{x}_k + \mathbf{B}_k \vec{u}_k + \mathbf{G}_k \vec{w}_k - (\Phi_k \hat{\vec{x}}_k^+ + \mathbf{B}_k \vec{u}_k)) \\ &\quad (\Phi_k \vec{x}_k + \mathbf{B}_k \vec{u}_k + \mathbf{G}_k \vec{w}_k - (\Phi_k \hat{\vec{x}}_k^+ + \mathbf{B}_k \vec{u}_k))^T] \\ &= E[(\Phi_k(\vec{x}_k - \hat{\vec{x}}_k^+) + \mathbf{G}_k \vec{w}_k)((\vec{x}_k - \hat{\vec{x}}_k^+)^T \Phi_k^T + \vec{w}_k^T \mathbf{G}_k^T)] \\ &= E[\Phi_k(\vec{x}_k - \hat{\vec{x}}_k^+)(\vec{x}_k - \hat{\vec{x}}_k^+)^T \Phi_k^T] + E[\Phi_k(\vec{x}_k - \hat{\vec{x}}_k^+) \vec{w}_k^T \mathbf{G}_k^T] \\ &\quad + E[\mathbf{G}_k \vec{w}_k(\vec{x}_k - \hat{\vec{x}}_k^+)^T \Phi_k^T] + E[\mathbf{G}_k \vec{w}_k \vec{w}_k^T \mathbf{G}_k^T] \end{aligned} \quad (6.25)$$

Da das Systemrauschen  $\vec{w}_k$  erst auf den Systemzustand zum Zeitpunkt  $k+1$  einen Einfluss hat, siehe Gl. (6.1), ist das Systemrauschen  $\vec{w}_k$  nicht mit dem Schätzfehler zum Zeitpunkt  $k$ , also  $\hat{\vec{x}}_k^+ - \vec{x}_k$ , korreliert. Damit folgt

$$\begin{aligned} \mathbf{P}_{xx,k+1}^- &= \Phi_k E[(\vec{x}_k - \hat{\vec{x}}_k^+)(\vec{x}_k - \hat{\vec{x}}_k^+)^T] \Phi_k^T + \mathbf{G}_k E[\vec{w}_k \vec{w}_k^T] \mathbf{G}_k^T \\ &= \Phi_k \mathbf{P}_{xx,k}^+ \Phi_k^T + \mathbf{G}_k \mathbf{Q}_k \mathbf{G}_k^T . \end{aligned} \quad (6.26)$$

Die Gleichungen (6.24) und (6.26) stellen den Propagationsschritt des Kalman-Filters dar.

### 6.1.2 Herleitung über Minimierung einer Kostenfunktion

Die Herleitung der Gleichungen des Kalman-Filter-Estimationsschrittes kann auch durch Lösung einer Optimierungsaufgabe erfolgen. Dazu wird davon ausgegangen, dass

sich der a-posteriori-Schätzwert aus dem a-priori-Schätzwert und der Differenz zwischen vorliegendem Messwertvektor  $\tilde{\mathbf{y}}$  und erwartetem Messwertvektor  $\hat{\mathbf{y}}_k = \mathbf{H}_k \hat{\mathbf{x}}_k^-$  ergibt, wobei diese Differenz mit einer zunächst unbekannten Matrix  $\mathbf{K}_k$  gewichtet wird:

$$\hat{\mathbf{x}}_k^+ = \hat{\mathbf{x}}_k^- + \mathbf{K}_k \left( \tilde{\mathbf{y}} - \mathbf{H}_k \hat{\mathbf{x}}_k^- \right) \quad (6.27)$$

Die eigentliche Optimierungsaufgabe besteht nun darin, die Gewichtungsmatrix  $\mathbf{K}_k$  so zu wählen, dass die Spur der a-posteriori-Kovarianzmatrix des Schätzfehlers minimal wird. Diese Wahl der Kostenfunktion liegt nahe, da die Spur von  $\mathbf{P}_{xx,k}^+$  gerade die Summe der Varianzen der Schätzfehler ist.

Allgemein, d.h. für eine beliebig gewählte Gewichtungsmatrix  $\mathbf{K}_k$ , erhält man mit Gl. (6.27) für  $\mathbf{P}_{xx,k}^+$ :

$$\begin{aligned} \mathbf{P}_{xx,k}^+ &= E \left[ (\hat{\mathbf{x}}_k^+ - \vec{x}_k)(\hat{\mathbf{x}}_k^+ - \vec{x}_k)^T \right] \\ &= E \left[ \left( \hat{\mathbf{x}}_k^- - \vec{x}_k + \mathbf{K}_k(\tilde{\mathbf{y}} - \mathbf{H}_k \hat{\mathbf{x}}_k^-) \right) \left( \hat{\mathbf{x}}_k^- - \vec{x}_k + \mathbf{K}_k(\tilde{\mathbf{y}} - \mathbf{H}_k \hat{\mathbf{x}}_k^-) \right)^T \right] \\ &= E \left[ \left( \hat{\mathbf{x}}_k^- - \vec{x}_k + \mathbf{K}_k(\mathbf{H}_k \vec{x}_k + \vec{v}_k - \mathbf{H}_k \hat{\mathbf{x}}_k^-) \right) \right. \\ &\quad \left. \left( \hat{\mathbf{x}}_k^- - \vec{x}_k + \mathbf{K}_k(\mathbf{H}_k \vec{x}_k + \vec{v}_k - \mathbf{H}_k \hat{\mathbf{x}}_k^-) \right)^T \right] \\ &= E \left[ \left( (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)(\hat{\mathbf{x}}_k^- - \vec{x}_k) + \mathbf{K}_k \vec{v}_k \right) \right. \\ &\quad \left. \left( (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)(\hat{\mathbf{x}}_k^- - \vec{x}_k) + \mathbf{K}_k \vec{v}_k \right)^T \right] \end{aligned} \quad (6.28)$$

Da das Messrauschen  $\vec{v}_k$  nicht mit dem a-priori-Schätzfehler  $(\hat{\mathbf{x}}_k^- - \vec{x}_k)$  korreliert sein kann, ergibt sich

$$\begin{aligned} \mathbf{P}_{xx,k}^+ &= E \left[ (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)(\hat{\mathbf{x}}_k^- - \vec{x}_k)(\hat{\mathbf{x}}_k^- - \vec{x}_k)^T (\mathbf{I} - \mathbf{H}_k^T \mathbf{K}_k^T) + \mathbf{K}_k \vec{v}_k \vec{v}_k^T \mathbf{K}_k^T \right] \\ &= (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_{xx,k}^- (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)^T + \mathbf{K}_k \mathbf{R}_k \mathbf{K}_k^T \end{aligned} \quad (6.29)$$

Das ist gerade die Joseph's Form des Kovarianzmatrix-Updates, während diese offensichtlich für beliebige Gewichtsmatrizen  $\mathbf{K}_k$  gilt, hat Gl. (6.19) nur bei Berechnung der Gewichtsmatrix  $\mathbf{K}_k$  gemäß Gl. (6.17) Gültigkeit.

Gesucht ist nun das Minimum der Spur von  $\mathbf{P}_{xx,k}^+$ , wobei der beeinflussende Faktor die Gewichtsmatrix  $\mathbf{K}_k$  ist. Daher muss zur Bestimmung dieses Minimums die Spur von  $\mathbf{P}_{xx,k}^+$  nach  $\mathbf{K}_k$  abgeleitet und die Nullstelle dieser Ableitung ermittelt werden:

$$\frac{d \text{Spur}(\mathbf{P}_{xx,k}^+)}{d \mathbf{K}_k} \stackrel{!}{=} \mathbf{0} \quad (6.30)$$

Dazu werden einige Zusammenhänge aus der Matrizenrechnung benötigt:

Bei der Spur einer Matrix handelt es sich um eine skalare Größe, die Ableitung eines Skalars  $s$  nach einer Matrix  $\mathbf{A}$  ist allgemein gegeben durch

$$\frac{d s}{d \mathbf{A}} = \begin{pmatrix} \frac{d s}{d a_{11}} & \frac{d s}{d a_{12}} & \cdots \\ \frac{d s}{d a_{21}} & \frac{d s}{d a_{22}} & \cdots \\ \vdots & & \end{pmatrix}. \quad (6.31)$$

Für zwei Matrizen  $\mathbf{A}$  und  $\mathbf{B}$  gilt unter der Voraussetzung, dass es sich bei dem Produkt  $\mathbf{AB}$  um eine quadratische Matrix handelt, der Zusammenhang

$$\frac{d \text{Spur}(\mathbf{AB})}{d \mathbf{A}} = \mathbf{B}^T. \quad (6.32)$$

Desweiteren gilt für eine beliebige Matrix  $\mathbf{A}$  und eine symmetrische Matrix  $\mathbf{B}$

$$\frac{d \text{Spur}(\mathbf{ABA}^T)}{d \mathbf{A}} = 2\mathbf{AB}. \quad (6.33)$$

Da sich beim Transponieren einer Matrix die Hauptdiagonalelemente nicht ändern, folgt sofort für die Spur der Transponierten

$$\text{Spur}(\mathbf{A}) = \text{Spur}(\mathbf{A}^T). \quad (6.34)$$

Durch Ausmultiplizieren von Gl. (6.29) findet man nun

$$\mathbf{P}_{xx,k}^+ = \mathbf{P}_{xx,k}^- - \mathbf{K}_k \mathbf{H}_k \mathbf{P}_{xx,k}^- - \mathbf{P}_{xx,k}^- \mathbf{H}_k^T \mathbf{K}_k^T + \mathbf{K}_k (\mathbf{H}_k \mathbf{P}_{xx,k}^- \mathbf{H}_k^T + \mathbf{R}_k) \mathbf{K}_k^T. \quad (6.35)$$

Anhand obiger Rechenregeln lassen sich folgende Ableitungen berechnen:

$$\begin{aligned} \frac{d \text{Spur}(\mathbf{P}_{xx,k}^-)}{d \mathbf{K}_k} &= \mathbf{0} \\ \frac{d \text{Spur}(\mathbf{P}_{xx,k}^- \mathbf{H}_k^T \mathbf{K}_k^T)}{d \mathbf{K}_k} &= \frac{d \text{Spur}(\mathbf{K}_k \mathbf{H}_k \mathbf{P}_{xx,k}^-)}{d \mathbf{K}_k} = \left( \mathbf{H}_k \mathbf{P}_{xx,k}^- \right)^T \\ \frac{d \text{Spur} \left( \mathbf{K}_k (\mathbf{H}_k \mathbf{P}_{xx,k}^- \mathbf{H}_k^T + \mathbf{R}_k) \mathbf{K}_k^T \right)}{d \mathbf{K}_k} &= 2\mathbf{K}_k (\mathbf{H}_k \mathbf{P}_{xx,k}^- \mathbf{H}_k^T + \mathbf{R}_k) \end{aligned}$$

Daraus folgt mit Gl. (6.35)

$$\frac{d \text{Spur}(\mathbf{P}_{xx,k}^+)}{d \mathbf{K}_k} = -2(\mathbf{H}_k \mathbf{P}_{xx,k}^-)^T + 2\mathbf{K}_k (\mathbf{H}_k \mathbf{P}_{xx,k}^- \mathbf{H}_k^T + \mathbf{R}_k) \stackrel{!}{=} \mathbf{0}. \quad (6.36)$$

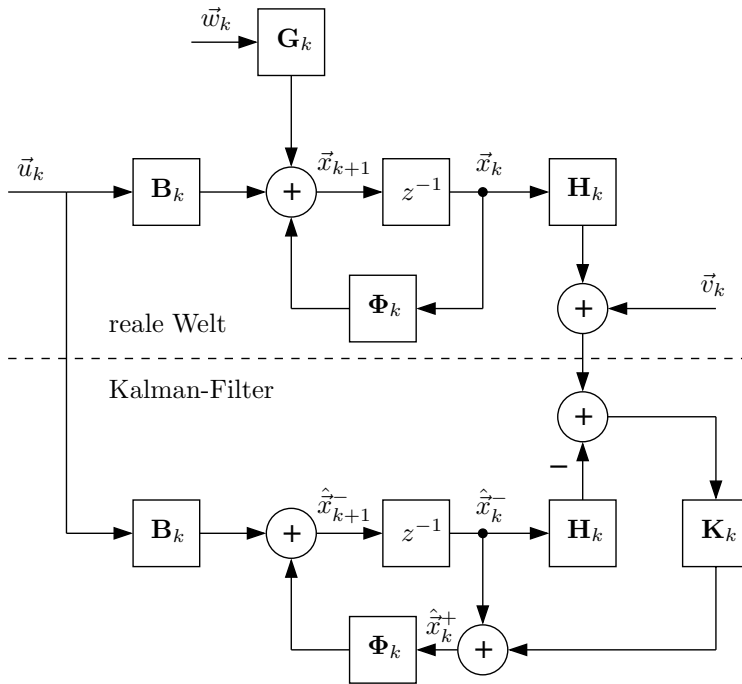


Abbildung 6.1: Kalman-Filter-Blockdiagramm.

Auflösen nach  $\mathbf{K}_k$  liefert die gesuchte Gewichtungsmatrix

$$\begin{aligned} \mathbf{K}_k(\mathbf{H}_k \mathbf{P}_{xx,k}^- \mathbf{H}_k^T + \mathbf{R}_k) &= \mathbf{P}_{xx,k}^- \mathbf{H}_k^T \\ \mathbf{K}_k &= \mathbf{P}_{xx,k}^- \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_{xx,k}^- \mathbf{H}_k^T + \mathbf{R}_k)^{-1}, \end{aligned} \quad (6.37)$$

die natürlich mit der im vorigen Abschnitt gefundenen Kalman-Gain-Matrix Gl. (6.17) übereinstimmt.

### 6.1.3 Diskussion der Filtergleichungen

Abb. 6.1 zeigt die prinzipielle Struktur eines Kalman-Filters in Form eines Blockdiagramms. Wie man anhand dieses Blockdiagramms oder den Gleichungen des Estimationsschrittes erkennt, erfolgt eine Anpassung der Zustandsschätzung anhand des vorliegenden Messwertvektors durch Rückführung der Differenz zwischen vorliegendem und erwartetem Messwertvektor. Die Grundidee dieser Vorgehensweise soll für den Fall eines idealen, nicht mit Rauschen behafteten Messwertvektors, d.h.  $\vec{v}_k = \vec{0}$ , verdeutlicht werden: Der erwartete Messwertvektor basiert auf der aktuellen Zustandsschätzung – ist



diese korrekt, werden der erwartete und der vorliegende Messwertvektor übereinstimmen und die Zustandsschätzung bleibt unverändert. Ist die aktuelle Zustandsschätzung nicht korrekt, so stimmen auch der erwartete und der vorliegende Messwertvektor nicht überein. In dieser Situation ist es plausibel, basierend auf der verbleibenden Differenz die Zustandsschätzung anzupassen.

### Beispiel Parameterschätzung

Der Einfluss der Kovarianz von Messrauschen und Schätzfehler im Rahmen der Messwertverarbeitung soll anhand eines einfachen Beispiels verdeutlicht werden. Hierzu soll das Kalman-Filter als Parameterschätzer eingesetzt werden, es soll eine konstante, skalare Größe geschätzt werden, die direkt gemessen werden kann. In diesem Szenario sind die Transitionsmatrix  $\Phi_k$  und die Messmatrix  $\mathbf{H}_k$  formal Einheitsmatrizen der Dimension  $1 \times 1$ . Da die Transitionsmatrix das Systemverhalten perfekt beschreibt und auch keine mit Rauschen behafteten Eingangsgrößen vorliegen, kann das Systemrauschen zu Null angenommen werden.

Die Kovarianzmatrizen des Messrauschens und des Schätzfehlers sind hier skalare Größen, die Kalman-Gain-Matrix kann daher wie folgt geschrieben werden:

$$\mathbf{K}_k = \frac{\mathbf{P}_{xx,k}^-}{\mathbf{P}_{xx,k}^- + \mathbf{R}_k} \quad (6.38)$$

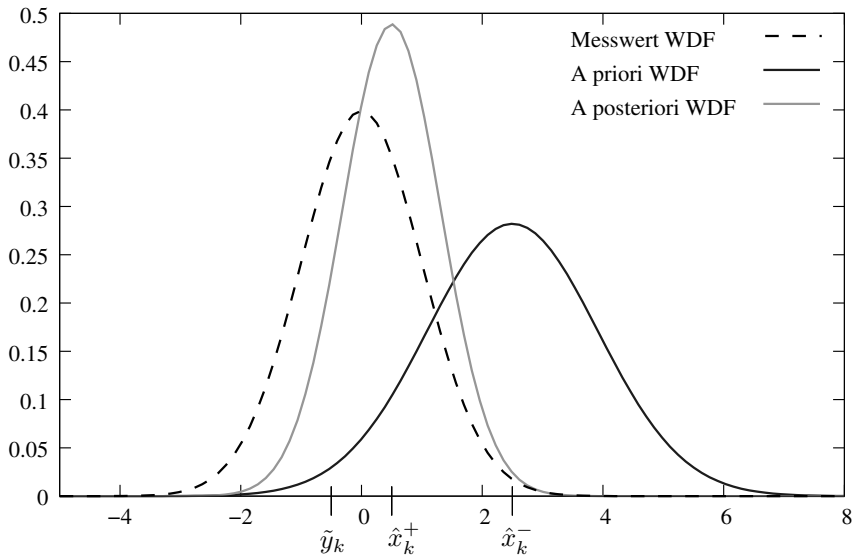
Damit vereinfachen sich die Gleichungen des Estimationsschrittes zu

$$\hat{x}_k^+ = \hat{x}_k^- + \frac{\mathbf{P}_{xx,k}^-}{\mathbf{P}_{xx,k}^- + \mathbf{R}_k} (\tilde{y}_k - \hat{x}_k^-) \quad (6.39)$$

$$\mathbf{P}_{xx,k}^+ = \mathbf{P}_{xx,k}^- \cdot \frac{\mathbf{R}_k}{\mathbf{P}_{xx,k}^- + \mathbf{R}_k} \quad (6.40)$$

Im Folgenden soll die resultierende Zustandsschätzung für einige Sonderfälle betrachtet werden:

1.  $\mathbf{R}_k = 0$ . Es liegt ein perfekter Messwert vor. Durch Einsetzen in (6.39) erkennt man, dass dieser Messwert als neue Zustandsschätzung direkt übernommen wird. Die bisherige Zustandsschätzung wird verworfen:  $\hat{x}_k^+ = \tilde{y}$ .  
Da aufgrund der Übernahme des perfekten Messwertes eine perfekte Zustandsschätzung vorliegt, liefert Gl. (6.40) eine verschwindende Schätzfehlervarianz:  $\mathbf{P}_{xx,k}^+ = 0$ .
2.  $\mathbf{R}_k \rightarrow \infty$ . Der Messwert enthält praktisch keine Information. In diesem Fall wird die Zustandsschätzung nicht verändert:  $\hat{x}_k^+ = \hat{x}_k^-$ .  
Durch Berechnung des entsprechenden Grenzwertes von Gl. (6.40) erkennt man, dass wie zu erwarten auch keine Anpassung der Schätzfehlervarianz stattfindet:  $\mathbf{P}_{xx,k}^+ = \mathbf{P}_{xx,k}^-$ .



**Abbildung 6.2:** Wahrscheinlichkeitsdichtefunktionen des Messwertes, der a-priori- und der a-posteriori-Zustandsschätzung.

3.  $\mathbf{P}_{xx,k}^- = 0$ . Die aktuelle Zustandsschätzung wird als perfekt angesehen, daher wird diese auch bei Vorliegen eines mit Rauschen behafteten Messwertes nicht mehr verändert:  $\hat{x}_k^+ = \hat{x}_k^-$ .  
Folglich resultiert für die Schätzfehlervarianz:  $\mathbf{P}_{xx,k}^+ = 0$ .
4.  $\mathbf{P}_{xx,k}^- \rightarrow \infty$ . Die bisherige Zustandsschätzung enthält praktisch keine Information. In dieser Situation ist es sinnvoll, bei Vorliegen eines Messwertes diese Zustandsschätzung zu verwerfen und diesen Messwert direkt zu übernehmen:  $\hat{x}_k^+ = \tilde{y}$ .  
Folglich ist die neue Schätzfehlervarianz durch die Varianz des Messwertes bestimmt:  $\mathbf{P}_{xx,k}^+ = \mathbf{R}_k$ .

In allen anderen Fällen ergibt sich der a-posteriori-Schätzwert als ein gewichteter Mittelwert von Messwert und a-priori-Schätzwert, wobei die Gewichtung von den vorliegenden Messwert- und Schätzfehlervarianzen abhängt. Dies ist anhand der Wahrscheinlichkeitsdichtefunktionen in Abb. 6.2 illustriert.

■

Ein großer Vorteil des Kalman-Filters ist dessen Modularität. Es ist keinesfalls notwendig, den Zeitraum zwischen zwei Messwerten mit einem einzigen Propagationsschritt zu überbrücken. Meist ist es sinnvoll, stattdessen den Propagationsschritt mit einer festen Update-Rate aufzurufen und dann zu jedem Zeitpunkt festzustellen, ob Messwerte vorliegen und ein Estimationsschritt ausgeführt werden muss. Ferner ist es nicht zwingend

notwendig, aus allen zu einem Zeitpunkt vorliegenden Messwerten einen Messwertvektor zu bilden und diesen in einem einzigen Estimationsschritt zu verarbeiten. Sind die einzelnen Messwerte nicht miteinander korreliert<sup>1</sup>, kann stattdessen jede skalare Einzelmessung im Rahmen eines Estimationsschrittes verarbeitet werden, es werden dann soviele Estimationsschritte durchgeführt, wie Messwerte zu diesem Zeitpunkt vorliegen. Dies hat zwei Vorteile: Bei der Berechnung der Kalman-Gain-Matrix muss eine Matrix invertiert werden, deren Dimension durch die Anzahl der Messwerte gegeben ist; liegen viele Messwerte zu einem Zeitpunkt vor, muss eine entsprechend große Matrix invertiert werden, was numerisch kritisch sein kann. Werden hingegen nur skalare Messwerte verarbeitet, reduziert sich die Berechnung der Inversen zu einer einfachen Division, die numerisch unkritisch ist. Desweiteren ist denkbar, dass zu verschiedenen Zeitpunkten eine unterschiedliche Anzahl von Messwerten vorliegt, z.B. wenn mehrere Sensoren mit verschiedenen Datenraten verwendet werden. Sollten alle zu einem Zeitpunkt vorliegenden Messwerte in einem einzigen Estimationsschritt verarbeitet werden, müsste die Implementierung des Filters alle bezüglich der Messwertverfügbarkeit möglichen Permutationen abdecken. In diesem Szenario wäre es sinnvoller, für jeden Sensor einen eigenen Estimationsschritt zu implementieren und diesen bei Bedarf aufzurufen.

## Eigenschaften

Neben den bisher angesprochenen Aspekten sollen im Folgenden die wesentlichen Eigenschaften des Kalman-Filters kurz angesprochen werden.

Für den Fall, dass die in Abschnitt 6 angesprochenen Voraussetzungen erfüllt sind, gilt:

- Das Kalman-Filter ist der optimale Schätzalgorithmus zur Zustandsschätzung dieses linearen Systems. Das bedeutet, dass kein anderes Filter bei dieser Aufgabenstellung bessere Ergebnisse erzielen kann.
- Das Kalman-Filter ist asymptotisch stabil: Der Einfluss der notwendigen Initialisierung der Kovarianzmatrix  $\mathbf{P}_0$  und des Zustandsvektors  $\hat{\mathbf{x}}_0^-$  verschwindet für  $t \rightarrow \infty$ .
- Die gelieferten Schätzwerte sind wirksam (engl. efficient estimate): Kein anderer Schätzer kann Schätzwerte mit einer kleineren Varianz liefern.
- Die Schätzwerte sind erwartungstreu (engl. unbiased): Die Schätzwerte stimmen im Mittel mit den tatsächlichen Werten überein.
- Wird das Kalman-Filter als Parameterschätzer eingesetzt, sind die gelieferten Schätzungen konsistent: Geht die Anzahl der zur Verfügung stehenden Messungen gegen unendlich, so gehen die Varianzen der Schätzfehler gegen null.
- Das Kalman-Filter liefert den wahrscheinlichsten Schätzwert (engl. maximum likelihood estimate): Der vom Kalman-Filter gelieferte Schätzwert ist der Mittelwert

---

<sup>1</sup>Eine separate Verarbeitung skalarer Einzelmessungen ist auch dann möglich, wenn diese miteinander in bekannter Weise korreliert sind. In diesem Fall muss vor der Verarbeitung durch eine geeignete Transformation eine Dekorrelation durchgeführt werden, siehe [38].

des Zustandsvektors  $\vec{x}_k$ . Da bei einer Gaußverteilung aber Mittelwert, Median und Modalwert zusammenfallen, ist dies auch der wahrscheinlichste Wert.

- Liegt kein Systemrauschen vor, ist das Kalman-Filter identisch zu einem rekursiven Least-Squares-Schätzer.

Ausführliche Diskussionen und weiterführende Betrachtungen des Kalman-Filters sind in [18],[21],[26],[35],[38],[85],[86],[90] oder [76] zu finden. Im Folgenden sollen lediglich einige der wesentlichsten Aspekte, die beim Einsatz eines Kalman-Filters zu berücksichtigen sind, angesprochen werden.

## 6.2 Beobachtbarkeit

Ein System heißt beobachtbar, wenn bei bekannten Eingangsgrößen der initiale Systemzustand  $\vec{x}_0$  nach endlicher Zeit anhand der zur Verfügung stehenden Messungen  $\vec{y}_0, \vec{y}_1, \dots, \vec{y}_{k-1}$  eindeutig ermittelt werden kann. Eine Aussage über die Beobachtbarkeit eines Systems kann anhand der Beobachtbarkeitsmatrix getroffen werden.

Für skalare Messwerte ist die Beobachtbarkeitsmatrix  $\mathbf{M}_b$  gegeben durch

$$\begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_{n-1} \end{pmatrix} = \underbrace{\begin{pmatrix} \mathbf{H}_0 \\ \mathbf{H}_1 \Phi_1 \\ \vdots \\ \mathbf{H}_n \Phi_{n-1} \end{pmatrix}}_{\mathbf{M}_b} \vec{x}_0, \quad (6.41)$$

wobei  $n$  die Dimension des Zustandsvektors bezeichnet. Das System ist beobachtbar, wenn die Matrix  $\mathbf{M}_b$  der Dimension  $n \times n$  den Höchstrang  $n$  besitzt, d.h. die Determinante von  $\mathbf{M}_b$  nicht verschwindet. In diesem Fall ist  $\mathbf{M}_b$  invertierbar, und das Gleichungssystem (6.41) kann nach dem Systemzustand  $\vec{x}_0$  aufgelöst werden.

Dieses Konzept soll anhand eines einfachen Beispiels verdeutlicht werden.

### Beispiel eindimensionale Bewegung

Gegeben sei das System

$$\begin{pmatrix} x \\ v \end{pmatrix}_{k+1} = \begin{pmatrix} 1 & T \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ v \end{pmatrix}_k + \begin{pmatrix} 0 \\ 1 \end{pmatrix} u_k. \quad (6.42)$$

Stehen für dieses System Positionsmessungen zur Verfügung,

$$y_k = (1 \ 0) \begin{pmatrix} x \\ v \end{pmatrix}_k, \quad (6.43)$$

ergibt sich die Beobachtbarkeitsmatrix zu

$$\mathbf{M}_b = \begin{pmatrix} 1 & 0 \\ 1 & T \end{pmatrix}. \quad (6.44)$$

Die Determinante dieser Matrix ist ungleich Null, die Matrix ist invertierbar und das System damit beobachtbar. Dies ist nicht verwunderlich, da durch Messungen der Position sowohl auf Position als auch auf Geschwindigkeit geschlossen werden kann.

Stehen jedoch anstelle von Positionsmessungen nur Geschwindigkeitsmessungen zu Verfügung,

$$y_k = (0 \ 1) \begin{pmatrix} x \\ v \end{pmatrix}_k, \quad (6.45)$$

erhält man als Beobachtbarkeitsmatrix

$$\mathbf{M}_b = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}. \quad (6.46)$$

Diese Matrix ist nicht invertierbar und das System damit unbeobachtbar. Das ist auch anschaulich unmittelbar klar, da nur durch Messung der Geschwindigkeit nicht auf eine unbekannte Anfangsposition geschlossen werden kann.

Beobachtbarkeit im Rahmen der stochastischen Filterung kann als Analogon zur Steuerbarkeit bei einem regelungstechnischen Problem gesehen werden: Steuerbarkeit besagt, dass ein System durch geeignete Eingangsgrößen in endlicher Zeit von einem beliebigen Anfangszustand nach  $\vec{x}_k = \vec{0}$  überführt werden kann.

## 6.3 Übergang kontinuierlich - diskret

Häufig ist das einem Kalman-Filter zugrundeliegende Systemmodell nicht in zeitdiskreter Form wie in Gl. (6.1) gegeben, sondern liegt im Zeitkontinuierlichen in Form von Differentialgleichungen vor:

$$\dot{\vec{x}}(t) = \mathbf{F}(t)\vec{x}(t) + \mathbf{B}(t)\vec{u}(t) + \mathbf{G}(t)\vec{w}(t) \quad (6.47)$$

Wie ein zeitdiskretes Äquivalent des deterministischen Anteils des Systemmodells Gl. (6.47) gewonnen werden kann, wurde bereits in Abschnitt 2.3 erörtert. Für das Systemrauschen gilt im Zeitkontinuierlichen

$$E [\vec{w}(t)\vec{w}(\tau)^T] = \mathbf{Q}\delta(t - \tau), \quad (6.48)$$

wobei es sich bei  $\mathbf{Q}$  um die spektrale Leistungsdichte des Systemrauschens handelt.

Es wird angenommen, dass sich der zeitdiskrete Rauschterm als Mittelwert des zeitkontinuierlichen Rauschterms innerhalb des Abtastintervalls ergibt:

$$\vec{w}_k = \frac{1}{T} \int_{t_k}^{t_{k+1}} \vec{w}(\tau) d\tau \quad (6.49)$$

Damit findet man für das Systemrauschen in analoger Weise zu Gl. (3.198) einen Zusammenhang zwischen Kovarianzmatrix im Zeitdiskreten und spektraler Leistungsdichte im

Zeitkontinuierlichen gemäß

$$\begin{aligned}
 \mathbf{Q}_k &= E [\vec{w}_k \vec{w}_k^T] = E \left[ \left( \frac{1}{T} \int_{t_k}^{t_{k+1}} \vec{w}(\tau) d\tau \right) \left( \frac{1}{T} \int_{t_k}^{t_{k+1}} \vec{w}(\tau) d\tau \right)^T \right] \\
 &= \frac{1}{T^2} \int_{t_k}^{t_{k+1}} \int_{t_k}^{t_{k+1}} E [\vec{w}(\tau) \vec{w}(\alpha)^T] d\alpha d\tau \\
 &= \frac{1}{T^2} \int_{t_k}^{t_{k+1}} \int_{t_k}^{t_{k+1}} \mathbf{Q} \cdot \delta(\tau - \alpha) d\alpha d\tau = \frac{1}{T^2} \int_{t_k}^{t_{k+1}} \mathbf{Q} d\tau = \frac{\mathbf{Q}}{T}. \quad (6.50)
 \end{aligned}$$

Bei der Berechnung der Matrix  $\mathbf{G}_k$  sind – wie bei der Eingangsmatrix, siehe Gl. 2.70 sowie Gl. 2.72 und 2.73 – verschiedene Approximationen möglich:

$$\mathbf{G}_k \approx \frac{1}{2} (\mathbf{I} + \Phi_k) \mathbf{G} T \quad (6.51)$$

$$\mathbf{G}_k \approx \Phi_k \mathbf{G} T \quad (6.52)$$

$$\mathbf{G}_k \approx \mathbf{G} T \quad (6.53)$$

Bei der Verwendung der Näherung Gl. (6.51) erhält man für das Systemrauschen im Zeitdiskreten

$$\begin{aligned}
 \mathbf{G}_k \mathbf{Q}_k \mathbf{G}_k^T &= E \left[ (\mathbf{G}_k \vec{w}_k) (\mathbf{G}_k \vec{w}_k)^T \right] \\
 &= \frac{1}{2} (\mathbf{I} + \Phi_k) \mathbf{G} T \cdot \frac{\mathbf{Q}}{T} \cdot \frac{1}{2} \mathbf{G}^T (\mathbf{I} + \Phi_k)^T T \\
 &= \frac{1}{4} (\mathbf{I} + \Phi_k) \mathbf{G} \mathbf{Q} \mathbf{G}^T (\mathbf{I} + \Phi_k)^T T, \quad (6.54)
 \end{aligned}$$

bei Verwendung von Gl. (6.52) ergibt sich entsprechend

$$\mathbf{G}_k \mathbf{Q}_k \mathbf{G}_k^T = \Phi_k \mathbf{G} \mathbf{Q} \mathbf{G}^T \Phi_k^T T, \quad (6.55)$$

und mit Gl. (6.53) erhält man

$$\mathbf{G}_k \mathbf{Q}_k \mathbf{G}_k^T = \mathbf{G} \mathbf{Q} \mathbf{G}^T T. \quad (6.56)$$

Häufig ist es nicht von entscheidender Bedeutung, welche der angegebenen Näherungen verwendet wird: Bei der Modellierung der Unsicherheit des Systemmodells als weißer,

normalverteilter Zufallsvektor ist eine gewisse Heuristik unvermeidlich, so dass bei komplexen Systemen das Systemrauschen meist eher als eine Möglichkeit des Filter-Tunings verstanden werden muss und nicht ausschließlich aufgrund mathematischer Überlegungen gewählt werden kann.

Neben der unabhängigen Diskretisierung von  $\mathbf{G}(t)$  und  $\vec{w}(t)$  ist auch eine gemeinsame Behandlung möglich. Das zeitdiskrete Äquivalent von Gl. (6.47) lautet

$$\begin{aligned} \vec{x}_{k+1} = e^{\mathbf{A}(t_{k+1}-t_k)} \vec{x}_k &+ \int_{t_k}^{t_{k+1}} e^{\mathbf{A}(t_{k+1}-\tau)} \mathbf{B} \vec{u}(\tau) d\tau \\ &+ \int_{t_k}^{t_{k+1}} e^{\mathbf{A}(t_{k+1}-\tau)} \mathbf{G} \vec{w}(\tau) d\tau . \end{aligned} \quad (6.57)$$

Nach einiger Rechnung erhält man für die Kovarianzmatrix des Systemrauschens den allgemeinen Ausdruck

$$\begin{aligned} \mathbf{G}_k \mathbf{Q}_k \mathbf{G}_k^T &= E \left[ \left( \int_{t_k}^{t_{k+1}} e^{\mathbf{A}(t_{k+1}-\tau)} \mathbf{G} \vec{w}(\tau) d\tau \right) \right. \\ &\quad \left. \left( \int_{t_k}^{t_{k+1}} e^{\mathbf{A}(t_{k+1}-\tau)} \mathbf{G} \vec{w}(\tau) d\tau \right)^T \right] \\ &= \int_0^T e^{\mathbf{A}\tau} \mathbf{G} \mathbf{Q} \mathbf{G}^T (e^{\mathbf{A}\tau})^T d\tau . \end{aligned} \quad (6.58)$$

Die Berechnung des Integrals ist in den meisten Fällen nur näherungsweise möglich.

## 6.4 Nichtlineare System- und Messmodelle

Handelt es sich beim zu beobachtenden System um ein nichtlineares System oder ist der Zusammenhang zwischen Messwerten und Zustandsvektor nichtlinear, kann bei nicht zu großen Nichtlinearitäten ein linearisiertes oder ein erweitertes Kalman-Filter eingesetzt werden.

### 6.4.1 Linearisiertes Kalman-Filter

Bei einem linearisierten Kalman-Filter wird ein zu beobachtendes, nichtlineares System

$$\dot{\vec{x}} = \vec{f}(\vec{x}) + \mathbf{B}\vec{u} + \mathbf{G}\vec{w} \quad (6.59)$$

um einen Linearisierungspunkt  $\vec{x}$  in eine Taylor-Reihe entwickelt, die nach dem linearen Glied abgebrochen wird:

$$\dot{\vec{x}} \approx \vec{f}(\vec{x}) + \left. \frac{\partial \vec{f}(\vec{x})}{\partial \vec{x}} \right|_{\vec{x}=\vec{x}} \cdot (\vec{x} - \vec{x}) + \mathbf{B}\vec{u} + \mathbf{G}\vec{w} \quad (6.60)$$

Analog findet man für den geschätzten Systemzustand

$$\begin{aligned} \dot{\hat{\vec{x}}} &= \vec{f}(\hat{\vec{x}}) + \mathbf{B}\vec{u} \\ &\approx \vec{f}(\vec{x}) + \left. \frac{\partial \vec{f}(\vec{x})}{\partial \vec{x}} \right|_{\vec{x}=\vec{x}} \cdot (\hat{\vec{x}} - \vec{x}) + \mathbf{B}\vec{u}. \end{aligned} \quad (6.61)$$

Subtrahiert man Gl. (6.60) von Gl. (6.61), erhält man

$$\dot{\hat{\vec{x}}} - \dot{\vec{x}} = \left. \frac{\partial \vec{f}(\vec{x})}{\partial \vec{x}} \right|_{\vec{x}=\vec{x}} \cdot (\hat{\vec{x}} - \vec{x}) - \mathbf{G}\vec{w}. \quad (6.62)$$

Setzt man  $\Delta\vec{x} = \hat{\vec{x}} - \vec{x}$  und wählt den geschätzten Systemzustand  $\hat{\vec{x}}$  als Linearisierungspunkt<sup>2</sup>, so ergibt sich

$$\Delta\dot{\vec{x}} = \left. \frac{\partial \vec{f}(\vec{x})}{\partial \vec{x}} \right|_{\vec{x}=\hat{\vec{x}}} \cdot \Delta\vec{x} - \mathbf{G}\vec{w}. \quad (6.63)$$

Gleichung (6.63) beschreibt näherungsweise die zeitliche Propagation der Schätzfehler durch das nichtlineare System und stellt das einem linearisierten Kalman-Filter zugrundeliegende Systemmodell dar. Die Matrix

$$\mathbf{F} = \left. \frac{\partial \vec{f}(\vec{x})}{\partial \vec{x}} \right|_{\vec{x}=\hat{\vec{x}}} \quad (6.64)$$

wird als Jacobi-Matrix der Funktion  $\vec{f}(\vec{x})$  bezeichnet.

Das linearisierte Kalman-Filter schätzt den Systemzustand nicht direkt, stattdessen wird der Fehler eines außerhalb des eigentlichen Filters gespeicherten, vermuteten Systemzustandes geschätzt. Daher ist auch die Bezeichnung error state space oder indirektes Kalman-Filter gebräuchlich. Wird anhand dieses geschätzten Fehlers der vermutete Systemzustand korrigiert, spricht man von einem closed-loop-error-state-space-Kalman-Filter.

---

<sup>2</sup>Es ist auch möglich, nicht den geschätzten Systemzustand sondern eine nominale Trajektorie als Linearisierungspunkt zu wählen.



Ein solches Filter verarbeitet auch die vorliegenden Messwerte nicht direkt, statt dessen wird die Differenz aus den vorliegenden Messwerten und den aufgrund des vermuteten Systemzustandes erwarteten Messwerten als 'Messung' verarbeitet. Das Messmodell wird hierbei in gleicher Weise wie das Systemmodell bestimmt: Der nichtlineare Zusammenhang zwischen Messwertvektor und Systemzustand

$$\tilde{\vec{y}}_k = \vec{h}_k(\vec{x}_k) + \vec{v}_k \quad (6.65)$$

wird in eine Taylor-Reihe entwickelt, wobei nach dem linearen Glied abgebrochen wird:

$$\tilde{\vec{y}}_k \approx \left. \frac{\partial \vec{h}_k(\vec{x}_k)}{\partial \vec{x}_k} \right|_{\vec{x}_k = \bar{\vec{x}}_k} \cdot (\vec{x}_k - \bar{\vec{x}}_k) + \vec{v}_k \quad (6.66)$$

Analog kann man schreiben:

$$\hat{\vec{y}}_k \approx \left. \frac{\partial \vec{h}_k(\vec{x}_k)}{\partial \vec{x}_k} \right|_{\vec{x}_k = \bar{\vec{x}}_k} \cdot (\hat{\vec{x}}_k - \bar{\vec{x}}_k) \quad (6.67)$$

Die Subtraktion dieser beiden Gleichungen liefert mit  $\Delta \tilde{\vec{y}}_k = \hat{\vec{y}}_k - \tilde{\vec{y}}_k$  bei Verwendung des vermuteten Systemzustandes als Linearisierungspunkt das Messmodell des linearisierten Kalman-Filters,

$$\Delta \tilde{\vec{y}}_k = \left. \frac{\partial \vec{h}_k(\vec{x}_k)}{\partial \vec{x}_k} \right|_{\vec{x}_k = \hat{\vec{x}}_k} \cdot \Delta \vec{x}_k - \vec{v}_k \quad (6.68)$$

Auf der Grundlage des Systemmodells (6.63) und des Messmodells (6.68) kommen die üblichen Kalman-Filter-Gleichungen zum Einsatz. Es ist jedoch darauf zu achten, dass die Propagation des vermuteten Systemzustandes durch Lösung des deterministischen Anteils der nichtlinearen Differentialgleichung (6.59) erfolgt; die Transitionsmatrix, die anhand der Jacobi-Matrix nach Gl. (6.64) und Gl. (2.68) ermittelt wird, wird nur zur Propagation der Kovarianzmatrix des Schätzfehlers benötigt. Die Berechnung des zu verarbeitenden Messwertes  $\Delta \tilde{\vec{y}}_k$  muss ebenfalls auf Grundlage des nichtlinearen Messmodells erfolgen:

$$\Delta \tilde{\vec{y}}_k = \vec{h}_k(\hat{\vec{x}}_k) - \tilde{\vec{y}}_k \quad (6.69)$$

Die durch die Jacobi-Matrix des nichtlinearen Messmodells gegebene Messmatrix

$$\mathbf{H}_k = \left. \frac{\partial \vec{h}_k(\vec{x}_k)}{\partial \vec{x}_k} \right|_{\vec{x}_k = \hat{\vec{x}}_k} \quad (6.70)$$

wird nur zur Berechnung der Kalman-Gain-Matrix und zur Anpassung der Kovarianzmatrix des Schätzfehlers verwendet.

Wird nach der Messwertverarbeitung der vermutete Systemzustand mit Hilfe der geschätzten Fehler korrigiert, muss der die Schätzfehler enthaltende Zustandsvektor des Filters zu Null gesetzt werden. Daher kann im anschließenden Propagationsschritt auf die Propagation des Filterzustandes verzichtet werden, neben dem vermuteten Systemzustand muss nur die Kovarianzmatrix der Schätzfehler propagiert werden.

## 6.4.2 Erweitertes Kalman-Filter

Das erweiterte Kalman-Filter<sup>3</sup> (EKF) unterscheidet sich von dem linearisierten Kalman-Filter dadurch, dass nicht die Fehler eines vermuteten Systemzustandes, sondern der Systemzustand selbst geschätzt wird. Man spricht daher auch von einem direkten Kalman-Filter, die Bezeichnung total-state-space-Kalman-Filter ist ebenfalls üblich.

Im Propagationsschritt wird der geschätzte Systemzustand anhand des nichtlinearen Systemmodells in der Zeit propagiert. Die Propagation der Kovarianzmatrix des Schätzfehlers erfolgt – wie beim linearisierten Kalman-Filter – unter Verwendung einer durch Linearisierung gewonnenen Transitionsmatrix. Ist das nichtlineare Systemmodell im Zeitdiskreten gegeben,

$$\vec{x}_{k+1} = \vec{f}(\vec{x}_k, \vec{u}_k) + \mathbf{G}_k \vec{w}_k, \quad (6.71)$$

erhält man die Transitionsmatrix unmittelbar durch

$$\Phi_k = \left. \frac{\partial \vec{f}(\vec{x}_k, \vec{u}_k)}{\partial \vec{x}} \right|_{\vec{x}=\hat{\vec{x}}} \quad (6.72)$$

Bei einem zeitkontinuierlichen Systemmodell wird die im vorigen Abschnitt beschriebene Vorgehensweise angewandt.

Die Messmatrix  $\mathbf{H}_k$  zur Berechnung der Kalman-Gain-Matrix und zur Anpassung der Kovarianzmatrix des Schätzfehlers ist durch Gl. (6.70) gegeben, die Prädiktion der Messwerte erfolgt auf Grundlage des nichtlinearen Messmodells Gl. (6.65).

Die Gleichungen des erweiterten Kalman-Filters lauten damit insgesamt:

$$\hat{\vec{x}}_{k+1}^- = \vec{f}(\hat{\vec{x}}_k^+, \vec{u}_k) \quad (6.73)$$

$$\mathbf{P}_{xx,k+1}^- = \Phi_k \mathbf{P}_{xx,k}^+ \Phi_k^T + \mathbf{G}_k \mathbf{Q}_k \mathbf{G}_k^T \quad (6.74)$$

$$\mathbf{K}_k = \mathbf{P}_{xx,k}^- \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_{xx,k}^- \mathbf{H}_k^T + \mathbf{R}_k)^{-1} \quad (6.75)$$

$$\hat{\vec{x}}_k^+ = \hat{\vec{x}}_k^- + \mathbf{K}_k (\tilde{y}_k - \vec{h}_k(\hat{\vec{x}}_k^-)) \quad (6.76)$$

$$\mathbf{P}_{xx,k}^+ = \mathbf{P}_{xx,k}^- - \mathbf{K}_k \mathbf{H}_k \mathbf{P}_{xx,k}^- \quad (6.77)$$

Während bei linearen Schätzproblemen mit normalverteiltem System- und Messrauschen das Kalman-Filter die optimale Lösung darstellt, sind linearisierte und erweiterte Kalman-Filter suboptimale, im allgemeinen nicht erwartungstreue Schätzalgorithmen. Sind die Nichtlinearitäten von System- und/oder Messmodell signifikant, können diese Filter unzureichende Ergebnisse liefern oder sogar divergieren, so dass andere Filtertypen wie z.B. Particle Filter eingesetzt werden müssen. In vielen Fällen ist der Einsatz eines linearisierten oder erweiterten Kalman-Filters jedoch gerechtfertigt, da der mit dem Einsatz eines 'echten' nichtlinearen Filters verbundene Aufwand in keinem Verhältnis zur Verbesserung des Schätzergebnisses stehen würde.

Eine ausführliche Diskussion von linearisiertem und erweitertem Kalman-Filter ist in [86] zu finden.

---

<sup>3</sup>engl. extended Kalman-Filter

### 6.4.3 Sigma-Point-Kalman-Filter

Die Gruppe der Sigma-Point-Kalman-Filter stellt eine Alternative zum Einsatz eines erweiterten Kalman-Filters dar. Die Grundidee der Sigma-Point-Kalman-Filter besteht darin, Mittelwert und Kovarianz eines normalverteilten Zufallsvektors durch eine Menge deterministisch gewählter Sigma-Punkte darzustellen, wobei jeder Sigma-Punkt als ein Zustandsvektor verstanden werden kann. Die Propagation des Zufallsvektors durch eine nichtlineare Funktion wird dadurch realisiert, dass die einzelnen Sigma-Punkte durch diese nichtlineare Funktion propagiert werden und anschließend Mittelwert und Varianz dieser transformierten Sigma-Punkte berechnet werden. Die Umsetzung dieser Idee in einen Filteralgorithmus wird im Folgenden aufgezeigt.

Das nichtlineare Systemmodell sei gegeben durch

$$\vec{x}_k = \vec{f}(\vec{x}_{k-1}, \vec{u}_{k-1}, \vec{w}_k), \quad (6.78)$$

das Messmodell liege in der Form

$$\vec{y}_k = \vec{h}(\vec{x}_k, \vec{v}_k) \quad (6.79)$$

vor. Im Rahmen dieser Formulierung besteht die Möglichkeit, dass das System- und Messrauschen nichtlinear in das entsprechende Modell eingeht, bei einem erweiterten Kalman-Filter müsste in diesem Fall erst entsprechend linearisiert werden.

Üblicherweise wird zunächst ein erweiterter Zustandsvektor konstruiert, der das System- und das Messrauschen beinhaltet:

$$\hat{\vec{x}}_k^a = E[\vec{x}_k^a] = (\vec{x}_k^T \ \vec{w}_k^T \ \vec{v}_k^T)^T \quad (6.80)$$

Die Kovarianzmatrix des Schätzfehlers dieses erweiterten Zustandsvektors ist gegeben durch

$$\mathbf{P}_{xx,k}^a = E[(\hat{\vec{x}}_k^a - \vec{x}_k^a)(\hat{\vec{x}}_k^a - \vec{x}_k^a)^T] = \begin{pmatrix} \mathbf{P}_{xx,k} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{R}_k \end{pmatrix}. \quad (6.81)$$

Zur Repräsentation von Mittelwert und Kovarianz werden  $2L+1$  Sigma-Punkte<sup>4</sup> gemäß

$$\vec{\chi}_{0,k}^a = \hat{\vec{x}}_k^a \quad (6.82)$$

$$\vec{\chi}_{i,k}^a = \hat{\vec{x}}_k^a + \zeta \sqrt{\mathbf{P}_{ki}^a}, \quad i = 1 \dots L \quad (6.83)$$

$$\vec{\chi}_{i,k}^a = \hat{\vec{x}}_k^a - \zeta \sqrt{\mathbf{P}_{ki}^a}, \quad i = L+1 \dots 2L \quad (6.84)$$

gewählt, wobei  $L$  die Anzahl der Komponenten des erweiterten Zustandsvektors bezeichnet. Jeder Sigma-Punkt ist folglich ein Vektor der Dimension  $L \times 1$ ,  $\zeta$  ist ein

<sup>4</sup>Es existieren auch Sigma-Point-Kalman-Filter-Formulierungen, die mit weniger Sigma-Punkten auskommen.

Skalierungsparameter der bestimmt, in welchem Abstand sich die Sigma-Punkte vom Mittelwert befinden. Die Wurzel der Matrix  $\mathbf{P}_{xx,k}^a$ ,

$$\sqrt{\mathbf{P}_{xx,k}^a} = \begin{pmatrix} \sqrt{\mathbf{P}_{xx,k}^a} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \sqrt{\mathbf{Q}_k} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \sqrt{\mathbf{R}_k} \end{pmatrix} \quad (6.85)$$

kann mit Hilfe einer Cholesky-Zerlegung berechnet werden. Wird diese Wurzel so berechnet, dass

$$\mathbf{P}_{xx,k}^a = \sqrt{\mathbf{P}_{xx,k}^a}^T \sqrt{\mathbf{P}_{xx,k}^a} \quad (6.86)$$

gilt, dann bezeichnet  $\sqrt{\mathbf{P}_{xx,k}^a}$  einen Vektor der Dimension  $L \times 1$ , dessen Komponenten durch die  $i$ -te Zeile der Matrix  $\sqrt{\mathbf{P}_{xx,k}^a}$  gegeben sind.

Im Prädiktionsschritt des Sigma-Point-Kalman-Filters werden diese Sigma-Punkte durch das nichtlineare Systemmodell propagiert:

$$\vec{\chi}_{i,k}^{a,-} = \vec{f}(\vec{\chi}_{i,k-1}^{a,+}, \vec{u}_{k-1}) \quad (6.87)$$

Anschließend werden anhand der transformierten Sigma-Punkte der propagierte Mittelwert und die propagierte Kovarianzmatrix berechnet:

$$\hat{\vec{x}}_k^{a,-} = \sum_{i=0}^{2L} w_i^m \vec{\chi}_{i,k}^{a,-} \quad (6.88)$$

$$\mathbf{P}_{xx,k}^{a,-} = \sum_{i=0}^{2L} \sum_{j=0}^{2L} w_{ij}^c (\vec{\chi}_{i,k}^{a,-} - \hat{\vec{x}}_k^{a,-})(\vec{\chi}_{j,k}^{a,-} - \hat{\vec{x}}_k^{a,-})^T \quad (6.89)$$

Bei  $w_i^m$  und  $w_{ij}^c$  handelt es sich erneut um Gewichtungsfaktoren.

Im Messschritt des Sigma-Point-Kalman-Filters werden die Sigma-Punkte durch das nichtlineare Messmodell propagiert:

$$\vec{\Upsilon}_{i,k} = \vec{h}(\vec{\chi}_{i,k}^{a,-}) \quad (6.90)$$

Basierend auf dieser Menge prädizierter Messwerte können der Mittelwert, die Kovarianz und die Kreuzkorrelation berechnet werden:

$$\hat{\vec{y}}_k = \sum_{i=0}^{2L} w_i^m \vec{\Upsilon}_{i,k} \quad (6.91)$$

$$\mathbf{P}_{yy,k} = \sum_{i=0}^{2L} \sum_{j=0}^{2L} w_{ij}^c (\vec{\Upsilon}_{i,k} - \hat{\vec{y}}_k)(\vec{\Upsilon}_{j,k} - \hat{\vec{y}}_k)^T \quad (6.92)$$

$$\mathbf{P}_{xy,k} = \sum_{i=0}^{2L} \sum_{j=0}^{2L} w_{ij}^c (\vec{\chi}_{i,k}^{a,-} - \hat{\vec{x}}_k^{a,-})(\vec{\Upsilon}_{j,k} - \hat{\vec{y}}_k)^T \quad (6.93)$$

Auf Grundlage dieser Größen ergeben sich die a-posteriori-Zustandsschätzung und Schätzfehlerkovarianz zu

$$\mathbf{K}_k = \mathbf{P}_{xy,k} \mathbf{P}_{yy,k}^{-1} \quad (6.94)$$

$$\hat{x}_k^{a,+} = \hat{x}_k^{a,-} + \mathbf{K}_k (\tilde{y}_k - \hat{y}_k) \quad (6.95)$$

$$\mathbf{P}_{xx,k}^+ = \mathbf{P}_{xx,k}^- - \mathbf{K}_k \mathbf{P}_{yy,k} \mathbf{K}_k^T. \quad (6.96)$$

Das Kovarianzmatrix-Update Gl. (6.96) erscheint auf den ersten Blick etwas ungewöhnlich, hierbei handelt es sich aber lediglich um eine äquivalente Schreibweise zu Gl. (5.32):

$$\begin{aligned} \mathbf{P}_{xx,k}^+ &= \mathbf{P}_{xx,k}^- - \mathbf{P}_{xy,k} \mathbf{P}_{yy,k}^{-1} \mathbf{P}_{xy,k}^T \\ &= \mathbf{P}_{xx,k}^- - \mathbf{P}_{xy,k} \mathbf{P}_{yy,k}^{-1} \mathbf{P}_{yy,k} \mathbf{P}_{yy,k}^{-1} \mathbf{P}_{xy,k}^T \\ \mathbf{P}_{xx,k}^+ &= \mathbf{P}_{xx,k}^- - \mathbf{K}_k \mathbf{P}_{yy,k} \mathbf{K}_k^T \end{aligned} \quad (6.97)$$

Gehen System- und Messrauschen linear in die entsprechenden Modelle ein, kann auf eine Erweiterung des Zustandsvektors verzichtet werden: In diesem Fall kann im Prädiktionsschritt die Kovarianzmatrix des Systemrauschens einfach additiv auf der rechten Seite von Gl. (6.89) berücksichtigt werden, das Messrauschen geht durch Addition von  $\mathbf{R}_k$  auf der rechten Seite von Gl. (6.92) ein. Dadurch kann der Rechenaufwand des Filters unter Umständen deutlich verringert werden.

Bisher wurde noch nicht näher auf den Skalierungsparameter  $\zeta$  und die Gewichtungsfaktoren  $w_i^m$  und  $w_{ij}^c$  eingegangen. Diese müssen bestimmten Bedingungen genügen, damit ein sinnvoller Filteralgorithmus resultiert; ein sinnvoller Mittelwert kann sich beispielsweise nur ergeben, wenn

$$\sum_{i=0}^{2L} w_i^m = 1 \quad (6.98)$$

gilt. Die konkrete Wahl dieser Parameter bestimmt den Typ des Sigma-Point-Kalman-Filters, die bekanntesten Vertreter sind der unscented Kalman-Filter und der central difference Kalman-Filter.

Eine einfache Möglichkeit der Parameterwahl, die einen Sonderfall des unscented Kalman-Filters darstellt, ist gegeben durch

$$w_i^m = \begin{cases} w_0 & i = 0 \\ \frac{1-w_0}{2L} & i > 0 \end{cases} \quad (6.99)$$

$$w_{ij}^c = \begin{cases} 0 & i \neq j \\ \frac{1-w_0}{2L} & i = j \end{cases} \quad (6.100)$$

$$\zeta = \sqrt{\frac{L}{1-w_0}}. \quad (6.101)$$

Es lässt sich leicht zeigen, dass mit dieser Wahl der Parameter eine korrekte Repräsentation von Mittelwert und Kovarianzmatrix gegeben ist. Durch Einsetzen erhält man

$$\begin{aligned}
 \sum_{i=0}^{2L} w_i^m \vec{\chi}_{i,k}^a &= w_0 \hat{x}_k^a + \sum_{i=1}^L \frac{1-w_0}{2L} \left( \hat{x}_k^a + \sqrt{\frac{L}{1-w_0}} \sqrt{\mathbf{P}_{ki}^a} \right) \\
 &\quad + \sum_{i=1}^L \frac{1-w_0}{2L} \left( \hat{x}_k^a - \sqrt{\frac{L}{1-w_0}} \sqrt{\mathbf{P}_{ki}^a} \right) \\
 &= w_0 \hat{x}_k^a + \frac{1-w_0}{2} \hat{x}_k^a + \frac{1-w_0}{2} \hat{x}_k^a \\
 &= \hat{x}_k^a.
 \end{aligned} \tag{6.102}$$

Für die Kovarianz ergibt sich

$$\begin{aligned}
 \sum_{i=0}^{2L} \sum_{j=0}^{2L} w_{ij}^c (\vec{\chi}_i^a - \hat{x}_k^a) (\vec{\chi}_j^a - \hat{x}_k^a)^T &= \frac{1-w_0}{2L} (\hat{x}_k^a - \hat{x}_k^a) (\hat{x}_k^a - \hat{x}_k^a)^T \\
 &\quad + \sum_{i=1}^L \frac{1-w_0}{2L} \left( \sqrt{\frac{L}{1-w_0}} \sqrt{\mathbf{P}_{ki}^a} \right) \left( \sqrt{\frac{L}{1-w_0}} \sqrt{\mathbf{P}_{ki}^a} \right)^T \\
 &\quad + \sum_{i=1}^L \frac{1-w_0}{2L} \left( -\sqrt{\frac{L}{1-w_0}} \sqrt{\mathbf{P}_{ki}^a} \right) \left( -\sqrt{\frac{L}{1-w_0}} \sqrt{\mathbf{P}_{ki}^a} \right)^T \\
 &= \frac{1}{2} \sum_{i=1}^L \left( \sqrt{\mathbf{P}_{ki}^a} \right) \left( \sqrt{\mathbf{P}_{ki}^a} \right)^T + \frac{1}{2} \sum_{i=1}^L \left( \sqrt{\mathbf{P}_{ki}^a} \right) \left( \sqrt{\mathbf{P}_{ki}^a} \right)^T \\
 &= \sum_{i=1}^L \left( \sqrt{\mathbf{P}_{ki}^a} \right) \left( \sqrt{\mathbf{P}_{ki}^a} \right)^T = \mathbf{P}_k^a.
 \end{aligned} \tag{6.103}$$

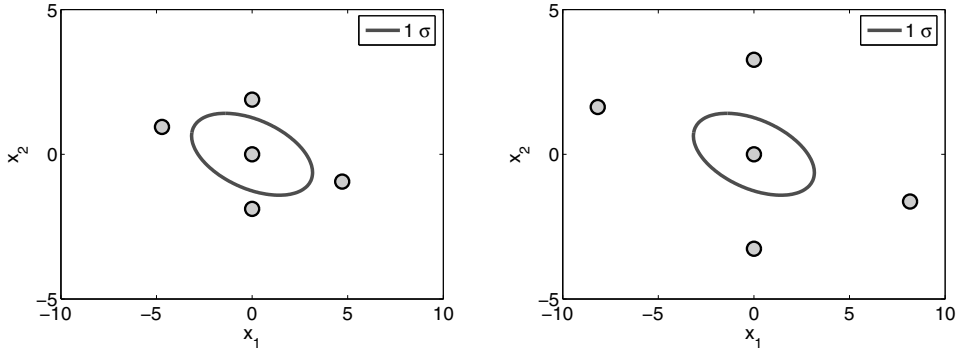
Je größer bei dieser Filterformulierung der Parameter  $w_0 < 1$  gewählt wird, desto weiter entfernen sich die Sigma-Punkte vom Mittelwert. Abb. 6.3 illustriert für einen zweidimensionalen Zufallsvektor die Lage der Sigma-Punkte sowie die Ellipse, die den  $(1 - \sigma)$ -Bereich markiert.

### Theoretischer Vergleich

Erweitertes Kalman-Filter und Sigma-Point Kalman-Filter verfolgen unterschiedliche Ansätze, um die Propagation von Mittelwert und Kovarianz eines Zufallsvektors durch eine nichtlineare Funktion zu approximieren. Am Beispiel des Mittelwertes sollen die Fehler dieser Approximationen näher betrachtet werden. Um diese Fehler berechnen zu können ist es zunächst notwendig, eine Referenzlösung zu bestimmen.

Die Taylor-Reihenentwicklung einer stetigen, nichtlinearen Funktion  $\vec{f}(\vec{x})$  ist gegeben durch

$$\vec{f}(\hat{x} + \Delta \vec{x}) = \vec{f}(\hat{x}) + D_{\Delta \vec{x}} \vec{f} + \frac{D_{\Delta \vec{x}}^2 \vec{f}}{2!} + \frac{D_{\Delta \vec{x}}^3 \vec{f}}{3!} + \frac{D_{\Delta \vec{x}}^4 \vec{f}}{4!} + \dots \tag{6.104}$$



**Abbildung 6.3:** Lage der Sigma-Punkte für  $w_0 = 0.1$  (links) und  $w_0 = 0.7$  (rechts) für einen zweidimensionalen Zufallsvektor.

Der Operator  $D_{\Delta\vec{x}}\vec{f}$  lässt sich mit Hilfe des Nabla-Operators

$$\nabla = \left( \frac{\partial}{\partial x_1} \quad \frac{\partial}{\partial x_2} \quad \dots \quad \frac{\partial}{\partial x_L} \right)^T \quad (6.105)$$

darstellen als

$$D_{\Delta\vec{x}}\vec{f} = \left( \Delta\vec{x}^T \nabla \right) \vec{f}(\vec{x}) \Big|_{\vec{x}=\hat{\vec{x}}} , \quad (6.106)$$

durch Ausmultiplizieren erhält man

$$D_{\Delta\vec{x}}\vec{f} = \sum_{i=1}^N \Delta x_i \frac{\partial}{\partial x_i} \vec{f}(\vec{x}) \Big|_{\vec{x}=\hat{\vec{x}}} . \quad (6.107)$$

Für die Reihenterme höherer Ordnung ergibt sich analog

$$\frac{D_{\Delta\vec{x}}^i \vec{f}}{i!} = \frac{1}{i!} \left( \sum_{j=1}^N \Delta x_j \frac{\partial}{\partial x_j} \right)^i \vec{f}(\vec{x}) \Big|_{\vec{x}=\hat{\vec{x}}} . \quad (6.108)$$

Um den Mittelwert eines Zufallsvektors  $\vec{x}$  nach der Propagation durch die nichtlineare Funktion berechnen zu können, muss der Erwartungswert der Reihenentwicklung berechnet werden. Der transformierte Zufallsvektor sei im Folgenden mit  $\vec{y} = \vec{f}(\vec{x})$  bezeichnet, dessen Erwartungswert ergibt sich zu

$$\begin{aligned} \hat{\vec{y}} &= E \left[ \vec{f}(\hat{\vec{x}} + \Delta\vec{x}) \right] \\ &= \vec{f}(\hat{\vec{x}}) + E \left[ D_{\Delta\vec{x}}\vec{f} + \frac{D_{\Delta\vec{x}}^2 \vec{f}}{2!} + \frac{D_{\Delta\vec{x}}^3 \vec{f}}{3!} + \frac{D_{\Delta\vec{x}}^4 \vec{f}}{4!} + \dots \right] . \end{aligned} \quad (6.109)$$

Da die Dichte  $\Delta\vec{x}$  als gaußförmig angenommen ist und per Definition mittelwertfrei ist, verschwinden alle Momente ungerader Ordnung: Für eine symmetrische Dichtefunktion  $p_x(x)$  einer mittelwertfreien Zufallsvariablen  $x$  gilt

$$p_x(x) = p_x(-x). \quad (6.110)$$

die Momente ungerader Ordnung ergeben sich mit diesen Voraussetzungen zu

$$\begin{aligned} E[x^{2n+1}] &= \int_{-\infty}^{\infty} x^{2n+1} p_x(x) dx \\ &= \int_0^{\infty} x^{2n+1} p_x(x) dx + \int_{-\infty}^0 x^{2n+1} p_x(x) dx \\ &= \int_0^{\infty} x^{2n+1} p_x(x) dx - \int_0^{\infty} x^{2n+1} p_x(x) dx. \end{aligned} \quad (6.111)$$

Durch Substitution von  $x = -u$  und folglich  $dx = -du$  im zweiten Integralterm erhält man schließlich

$$\begin{aligned} E[x^{2n+1}] &= \int_0^{\infty} x^{2n+1} p_x(x) dx - \int_0^{\infty} (-u)^{2n+1} p_x(-u) (-1) du \\ &= \int_0^{\infty} x^{2n+1} p_x(x) dx - \int_0^{\infty} (-1)^{2n+2} u^{2n+1} p_x(u) du \\ &= \int_0^{\infty} x^{2n+1} p_x(x) dx - \int_0^{\infty} u^{2n+1} p_x(u) du \\ &= 0. \end{aligned} \quad (6.112)$$

Damit verschwinden alle Reihenglieder ungerader Ordnung in Gl. (6.109) und man kann schreiben:

$$\hat{\vec{y}} = \vec{f}(\hat{\vec{x}}) + E \left[ \frac{D_{\Delta\vec{x}}^2 \vec{f}}{2!} + \frac{D_{\Delta\vec{x}}^4 \vec{f}}{4!} + \dots \right] \quad (6.113)$$

Im Folgenden soll die Taylor-Reihe nur noch bis zum Glied zweiter Ordnung detailliert betrachtet werden, da die Gleichungen für die Glieder höherer Ordnung sehr unübersichtlich werden. Mit

$$\frac{D_{\Delta\vec{x}}^2 \vec{f}}{2!} = \frac{D_{\Delta\vec{x}} (D_{\Delta\vec{x}} \vec{f})}{2!} = \left( \frac{\Delta\vec{x}^T \nabla \Delta\vec{x}^T \nabla}{2!} \right) \vec{f} = \left( \frac{\nabla^T \Delta\vec{x} \Delta\vec{x}^T \nabla}{2!} \right) \vec{f} \quad (6.114)$$

und  $E[\Delta\vec{x} \Delta\vec{x}^T] = \mathbf{P}^a$  erhält man für den Mittelwert des transformierten Zufallsvektors

$$\hat{\vec{y}} = \vec{f}(\hat{\vec{x}}) + \left( \frac{\nabla^T \mathbf{P}^a \nabla}{2!} \right) \vec{f} + \dots \quad (6.115)$$



Dieser Ausdruck kann mit den von einem erweiterten Kalman-Filter und von einem Sigma-Point-Kalman-Filter gelieferten Approximationen verglichen werden.

### Transformation des Mittelwertes beim erweiterten Kalman-Filter

Der transformierte Mittelwert ergibt sich bei Verwendung eines erweiterten Kalman-Filters zu

$$\hat{\vec{y}}_{EKF} = \vec{f}(\hat{\vec{x}}) . \quad (6.116)$$

Offensichtlich geht die Varianz der Zustandsschätzung nicht in die Propagation des Mittelwertes ein. Durch Vergleich mit dem exakten Ergebnis Gl. (6.115) erkennt man, dass die beim erweiterten Kalman-Filter verwendete Approximation nur bis zur ersten Ordnung genau ist<sup>5</sup>.

### Transformation des Mittelwertes beim Sigma-Point-Kalman-Filter

Die Berechnung des transformierten Mittelwertes erfolgt beim Sigma-Point-Kalman-Filter anhand der transformierten Sigma-Punkte, die mit der Kurzschreibweise  $\vec{\sigma}_i = \zeta\sqrt{\mathbf{P}^a}_i$  wie folgt dargestellt werden können:

$$\vec{\Upsilon}_i = \vec{f}(\vec{\chi}_i) = \vec{f}(\hat{\vec{x}} + \vec{\sigma}_i) \quad (6.117)$$

Durch eine Taylor-Reihenentwicklung der nichtlinearen Funktion um den Entwicklungspunkt  $\hat{\vec{x}}$  erhält man

$$\vec{\Upsilon}_i = \vec{f}(\hat{\vec{x}}) + D_{\vec{\sigma}_i} \vec{f} + \frac{D_{\vec{\sigma}_i}^2 \vec{f}}{2!} + \frac{D_{\vec{\sigma}_i}^3 \vec{f}}{3!} + \frac{D_{\vec{\sigma}_i}^4 \vec{f}}{4!} + \dots . \quad (6.118)$$

Bei einer Wahl der Gewichtungsfaktoren nach Gl. (6.99) ergibt sich folgender transformierter Mittelwert:

$$\begin{aligned} \hat{\vec{y}}_{SPKF} &= w_0 \vec{\Upsilon}_0 + \frac{1 - w_0}{2L} \sum_{i=1}^{2L} \vec{\Upsilon}_i \\ &= w_0 \vec{f}(\hat{\vec{x}}) + \frac{1 - w_0}{2L} \sum_{i=1}^{2L} \left( \vec{f}(\hat{\vec{x}}) + D_{\vec{\sigma}_i} \vec{f} + \frac{D_{\vec{\sigma}_i}^2 \vec{f}}{2!} + \frac{D_{\vec{\sigma}_i}^3 \vec{f}}{3!} + \dots \right) \end{aligned} \quad (6.119)$$

Da die Sigma-Punkte symmetrisch um den Mittelwert  $\hat{\vec{x}}$  verteilt sind, entfällt in der Summe auf der rechten Seite von Gl. (6.119) der Term erster Ordnung. Die höheren Terme ungerader Ordnung entfallen ebenso, da diese von den entsprechenden Momenten ungerader Ordnung abhängen, die bei symmetrischen Dichtefunktionen verschwinden. Damit erhält man

$$\hat{\vec{y}}_{SPKF} = \vec{f}(\hat{\vec{x}}) + \frac{1 - w_0}{2L} \sum_{i=1}^{2L} \left( \frac{D_{\vec{\sigma}_i}^2 \vec{f}}{2!} + \frac{D_{\vec{\sigma}_i}^4 \vec{f}}{4!} + \dots \right) . \quad (6.120)$$

<sup>5</sup>Es können durch Linearisierung auch Filter höherer Ordnung gefunden werden, siehe z.B. [35].

Aus Gründen der Übersichtlichkeit sollen nur Terme bis zur zweiten Ordnung weiter betrachtet werden. Mit Gl. (6.114) ergibt sich

$$\begin{aligned}
 \hat{y}_{SPKF} &= \vec{f}(\hat{x}) + \frac{1-w_0}{2L} \sum_{i=1}^{2L} \left( \frac{\nabla^T \vec{\sigma}_i \vec{\sigma}_i^T \nabla}{2!} \right) \vec{f} + \dots \\
 &= \vec{f}(\hat{x}) + \frac{1-w_0}{2L} \left( \frac{\nabla^T \sum_{i=1}^{2L} \vec{\sigma}_i \vec{\sigma}_i^T \nabla}{2!} \right) \vec{f} + \dots \\
 &= \vec{f}(\hat{x}) + \frac{1-w_0}{L} \left( \frac{\nabla^T \sum_{i=1}^L \vec{\sigma}_i \vec{\sigma}_i^T \nabla}{2!} \right) \vec{f} + \dots .
 \end{aligned} \tag{6.121}$$

Hierbei wurde ausgenutzt, dass aufgrund der Symmetrie der Sigma-Punkte  $\vec{\sigma}_i = -\vec{\sigma}_{L+i}$  gilt und damit  $\vec{\sigma}_i \vec{\sigma}_i^T = \vec{\sigma}_{L+i} \vec{\sigma}_{L+i}^T$ . Da  $\vec{\sigma}_i = \zeta \sqrt{\mathbf{P}^a}_i$  einen Vektor der Dimension  $L \times 1$  bezeichnet, dessen Komponenten sich aus der skalierten  $i$ -te Zeile der Matrix  $\sqrt{\mathbf{P}^a}$  ergeben, kann man schreiben:

$$\sum_{i=1}^L \vec{\sigma}_i \vec{\sigma}_i^T = \zeta \sqrt{\mathbf{P}^a}^T \zeta \sqrt{\mathbf{P}^a} = \frac{L}{1-w_0} \mathbf{P}^a \tag{6.122}$$

Insgesamt erhält man so für den transformierten Mittelwert bei Verwendung eines Sigma-Point-Kalman-Filters

$$\begin{aligned}
 \hat{y}_{SPKF} &= \vec{f}(\hat{x}) + \frac{1-w_0}{L} \left( \frac{\nabla^T \sum_{i=1}^L \vec{\sigma}_i \vec{\sigma}_i^T \nabla}{2!} \right) \vec{f} + \dots \\
 &= \vec{f}(\hat{x}) + \left( \frac{\nabla^T \mathbf{P}^a \nabla}{2!} \right) \vec{f} + \dots .
 \end{aligned} \tag{6.123}$$

Durch Vergleich mit Gl. (6.115) erkennt man, dass dieses Ergebnis bis zum Term zweiter Ordnung mit dem exakten Ergebnis übereinstimmt, und zwar unabhängig von der Wahl des Parameters  $w_0$ . Da der Term dritter Ordnung sowohl beim Sigma-Point-Kalman-Filter als auch bei der exakten Berechnung verschwindet, spricht man bezüglich des transformierten Mittelwertes von einer Genauigkeit dritter Ordnung. Werden wie z.B. beim unscented Filter die Gewichtungsfaktoren nach einer etwas komplexeren Vorschrift als Gl. (6.99)–(6.101) gewählt, kann durch eine geschickte Festlegung der Gewichtungsfaktoren zusätzlich versucht werden, den Term vierter Ordnung anzunähern, siehe [61]. Wie gut dies gelingt, hängt von der vorliegenden Nichtlinearität ab.

Für die Transformation der Kovarianz lässt sich in ähnlicher Art und Weise für den Mittelwert zeigen, dass ein erweitertes Kalman-Filter bis zur dritten Ordnung, ein Sigma-Point-Kalman-Filter bis zur fünften Ordnung korrekte Approximationen liefert.

Weiterführende Betrachtungen zur Klasse der Sigma-Point-Kalman-Filter sind in [59] zu finden.

### Fazit

Ein Sigma-Point-Kalman-Filter ist in der Lage, Mittelwert und Kovarianz eines Zufallsvektors bei nichtlinearen Transformationen genauer zu erfassen als ein erweitertes Kalman-Filter. Diesem Vorteil steht je nach Applikation ein unter Umständen deutlich größerer Rechenaufwand gegenüber. Ob ein Sigma-Point-Kalman-Filter einem erweiterten Kalman-Filter vorzuziehen ist hängt davon ab, ob in dem betrachteten Szenario die Terme höherer Ordnung der nichtlinearen Modelle einen signifikanten Einfluss haben oder nicht. Vereinfacht kann man sagen, je nichtlinearer sich die Filteraufgabe gestaltet, desto eher kann ein Sigma-Point-Kalman-Filter vorteilhaft sein. Bei linearen System- und Messmodellen liefern Kalman-Filter und Sigma-Point-Kalman-Filter identische Ergebnisse.

Da am Ende jedes Propagationsschrittes und jedes Messschrittes bei einem erweiterten Kalman-Filter genau wie bei einem Sigma-Point-Kalman-Filter die Wahrscheinlichkeitsdichtefunktion der Zustandsschätzung durch Mittelwert und Varianz repräsentiert wird, lassen sich Sigma-Point-Kalman-Filter und erweitertes Kalman-Filter problemlos kombinieren: So kann z.B. bei einem linearen Systemmodell aber signifikant nichtlinearen Messmodell der Propagationsschritt eines Kalman-Filters und der Messschritt eines Sigma-Point-Kalman Filters verwendet werden.

## 6.4.4 Kalman-Filter 2. Ordnung

Bei nichtlinearen Messgleichungen kann die Leistungsfähigkeit eines erweiterten Kalman-Filters in manchen Situationen durch eine einfache Modifikation gesteigert werden: Der Messwert wird wie bei einem EKF üblich verarbeitet, die Kovarianzmatrix des Schätzfehlers wird jedoch noch nicht angepasst. Stattdessen wird basierend auf der verbesserten Zustandsschätzung erneut linearisiert, und so eine verbesserte Messmatrix bestimmt. Nun wird der Messwert unter Verwendung dieser verbesserten Messmatrix erneut verarbeitet, um eine erneut verbesserte Zustandsschätzung zu erhalten. Diese Iteration kann einige Male durchgeführt werden, bis sich die Zustandsschätzung nicht mehr ändert. Ist dies geschehen, schließt die Anpassung der Kovarianzmatrix des Schätzfehlers den Messschritt ab. Ein solches erweitertes Kalman-Filter wird als iterated extended Kalman-Filter (IEKF)<sup>6</sup> bezeichnet.

Natürlich kann auch ein erweitertes Kalman-Filter hergeleitet werden, indem die Taylorreihenentwicklung des nichtlinearen System- und Messmodells nicht nach dem linearen Glied, sondern erst nach dem quadratischen Glied abgebrochen wird. Dieses Filter wird manchmal als Kalman-Filter 2. Ordnung bezeichnet.

Gegeben sei das nichtlineare Systemmodell

$$\vec{x}_{k+1} = \vec{f}(\vec{x}_k) + \vec{w}_k \quad (6.124)$$

wobei  $\vec{w}_k$  mittelwertfreies, weißes, normalverteiltes Rauschen mit der Kovarianzmatrix  $E[\vec{w}_k \vec{w}_k^T] = \mathbf{Q}_k$  ist. Der Zustandsvektor besitzt die Dimension  $n \times 1$ .

---

<sup>6</sup>Es wird zwischen lokalem IEKF und globalem IEKF unterschieden, bei der beschriebenen Vorgehensweise handelt es sich um ein lokales IEKF.

Der Propagationsschritt des Kalman-Filters 2. Ordnung lautet dann [73]:

$$\hat{\vec{x}}_{k+1}^- = \vec{f}_k(\hat{\vec{x}}_k^+) + \frac{1}{2} \sum_{p=1}^n \sum_{q=1}^n \frac{\partial^2 \vec{f}(\vec{x}_k)}{\partial x_p \partial x_q} \bigg|_{\vec{x}_k = \hat{\vec{x}}_k^+} \cdot P_{pq} \quad (6.125)$$

$$\mathbf{P}_{k+1}^- = \mathbf{\Phi}_k \mathbf{P}_k^+ \mathbf{\Phi}_k^T + \mathbf{A}_k + \mathbf{Q}_k \quad (6.126)$$

$$\mathbf{\Phi}_k = \frac{\partial \vec{f}(\vec{x}_k)}{\partial \vec{x}_k} \bigg|_{\vec{x}_k = \hat{\vec{x}}_k^+} \quad (6.127)$$

$$\mathbf{A}_k = \frac{1}{4} \sum_{p=1}^n \sum_{q=1}^n \sum_{r=1}^n \sum_{s=1}^n \frac{\partial^2 \vec{f}(\vec{x}_k)}{\partial x_p \partial x_q} \frac{\partial^2 \vec{f}^T(\vec{x}_k)}{\partial x_r \partial x_s} \bigg|_{\vec{x}_k = \hat{\vec{x}}_k^+} \cdot (P_{qr} P_{ps} + P_{qs} P_{pr}) \quad (6.128)$$

Hierbei bezeichnet  $x_p$  die  $p$ -te Komponente des Zustandsvektors  $\vec{x}_k$ ,  $P_{pq}$  ist der Eintrag in der  $p$ -ten Zeile und  $q$ -ten Spalte der Kovarianzmatrix der Zustandsschätzung,  $\mathbf{P}_k$ .

Unter Annahme des Messmodells Gl. (6.65) ist der Messschritt des Filters 2. Ordnung gegeben durch:

$$\mathbf{K}_k = \mathbf{P}_k^- \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^T + \mathbf{S}_k + \mathbf{R}_k)^{-1} \quad (6.129)$$

$$\hat{\vec{x}}_k^+ = \hat{\vec{x}}_k^- + \mathbf{K}_k \left( \vec{y}_k - \vec{h}_k(\hat{\vec{x}}_k^-) - \frac{1}{2} \sum_{p=1}^n \sum_{q=1}^n \frac{\partial^2 \vec{h}_k(\vec{x}_k)}{\partial x_p \partial x_q} \bigg|_{\vec{x}_k = \hat{\vec{x}}_k^-} \cdot P_{pq} \right) \quad (6.130)$$

$$\mathbf{P}_k^+ = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_k^- \quad (6.131)$$

$$\mathbf{H}_k = \frac{\partial \vec{h}_k(\vec{x}_k)}{\partial \vec{x}_k} \bigg|_{\vec{x}_k = \hat{\vec{x}}_k^-} \quad (6.132)$$

$$\mathbf{S}_k = \frac{1}{4} \sum_{p=1}^n \sum_{q=1}^n \sum_{r=1}^n \sum_{s=1}^n \frac{\partial^2 \vec{h}_k(\vec{x}_k)}{\partial x_p \partial x_q} \frac{\partial^2 \vec{h}_k^T(\vec{x}_k)}{\partial x_r \partial x_s} \bigg|_{\vec{x}_k = \hat{\vec{x}}_k^-} \cdot (P_{qr} P_{ps} + P_{qs} P_{pr}) \quad (6.133)$$

Das Kalman-Filter zeigt ein ähnliches Verhalten wie das Sigma-Point-Kalman-Filter, die Berechnung der benötigten 2. Ableitungen kann jedoch umständlich und fehlerträchtig sein. Andererseits kann durch Berechnung einiger zusätzlicher Korrekturterme ein bereits bestehendes EKF in ein Filter 2. Ordnung umgewandelt werden. Ist von vorneherein klar, dass ein Filter 2. Ordnung benötigt wird, ist der Entwurf eines Sigma-Point-Kalman-Filters meist weniger aufwändig und daher vorzuziehen.

## 6.5 Filterung bei zeitkorreliertem Rauschen

Bei vielen praktischen Anwendungen ist die bisher getroffene Voraussetzung, dass das Mess- und Systemrauschen als weiß angenommen werden kann, nicht erfüllt. Ein Beispiel

hierfür wären Inertialsensoren, die Vibrationen ausgesetzt sind. Sind die Inertialsensoren nicht in der Lage, die den Vibrationen zugrundeliegenden oszillatorischen Rotationen und Translationen aufzulösen, erscheinen diese als zusätzliches Rauschen in den Inertialsensordaten. Diese Vibrationen lassen sich nicht als weißes Rauschen modellieren, meist handelt es sich um Schwingungen mit bestimmten Eigenfrequenzen. Damit ist das Leistungsdichtespektrum nicht konstant, folglich liegen Zeitkorrelationen vor. Werden diese Zeitkorrelationen nicht im Filteralgorithmus berücksichtigt, kann das zu schlechten oder sogar unbrauchbaren Filterergebnissen führen. Darüber hinaus spiegelt die vom Filter gelieferte Kovarianzmatrix der Schätzfehler nicht die tatsächliche Güte der Zustandsschätzung wieder. Eine überoptimistische Einschätzung der Güte der Zustandsschätzung führt zu einer geringeren Gewichtung der vorliegenden Messwerte, was sogar die Divergenz des Filters verursachen kann. Eine Möglichkeit zur Berücksichtigung von Zeitkorrelationen ist also nicht nur für die Genauigkeit der Zustandsschätzung, sondern auch zur Sicherstellung deren Zuverlässigkeit von Interesse.

Im Folgenden sollen daher die gängigen Verfahren zur Berücksichtigung von Zeitkorrelationen bei Filterproblemen vorgestellt werden. Zur besseren Übersichtlichkeit wird ein Systemmodell ohne Eingangsgrößen betrachtet,

$$\vec{x}_{k+1} = \Phi_k \vec{x}_k + \mathbf{G}_k \vec{w}_k . \quad (6.134)$$

Das zeitkorrelierte Systemrauschen wird durch einen Rauschprozess der Form

$$\vec{w}_k = \mathbf{C} \vec{w}_{k-1} + \vec{\eta}_k \quad (6.135)$$

beschrieben. Mit einem solchen Modell lassen sich bei geeigneter Wahl des Zustandsvektors  $\vec{w}_k$ , der Gewichtungsmatrix  $\mathbf{C}$  und dem Vektor  $\vec{\eta}_k$  auch Rauschprozesse höherer Ordnung darstellen. Das in das Messmodell

$$\tilde{y}_k = \mathbf{H}_k \vec{x}_k + \vec{v}_k \quad (6.136)$$

eingehende Messrauschen soll in analoger Weise durch das Rauschprozessmodell

$$\vec{v}_k = \mathbf{D} \vec{v}_{k-1} + \vec{\mu}_k \quad (6.137)$$

beschrieben werden.

### 6.5.1 Erweiterung des Zustandsvektors

Der übliche Ansatz zur Berücksichtigung von zeitkorreliertem Rauschen besteht in der Erweiterung des Systemmodells um die Zustandsraummodelle dieser zeitkorrelierten Rauschprozesse. In dem hier betrachteten Fall ergibt sich das erweiterte Systemmodell zu

$$\begin{pmatrix} \vec{x} \\ \vec{w} \\ \vec{v} \end{pmatrix}_{k+1} = \begin{pmatrix} \Phi_k & \mathbf{G}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{C} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{D} \end{pmatrix} \begin{pmatrix} \vec{x} \\ \vec{w} \\ \vec{v} \end{pmatrix}_k + \begin{pmatrix} \mathbf{0} \\ \vec{\eta} \\ \vec{\mu} \end{pmatrix}_k . \quad (6.138)$$

Das Systemrauschen ist durch das weiße Rauschen gegeben, das die zeitkorrelierten Rauschprozesse treibt. Zur Verarbeitung der Messwerte wird die Messmatrix ebenfalls erweitert, man erhält

$$\tilde{\mathbf{y}}_k = (\mathbf{H}_k \ \mathbf{0} \ \mathbf{I}) \begin{pmatrix} \vec{x} \\ \vec{w} \\ \vec{v} \end{pmatrix}_k \quad (6.139)$$

Der normalerweise übliche, additive Messrauschterm verschwindet in diesem Messmodell, da das zeitkorrelierte Messrauschen im Zustandsvektor enthalten ist. Offensichtlich wäre es sehr einfach möglich, noch zusätzlich einen weißen Rauschterm für das Messrauschen einzuführen, so dass insgesamt das Messrauschen als Summe des zeitkorrelierten Rauschprozesses und dieses weißen Rauschterms modelliert werden könnte.

Da durch diese Erweiterung des Zustandsvektors formal wieder weißes Mess- und Systemrauschen vorliegt, können die üblichen Kalman-Filter-Gleichungen Anwendung finden.

Werden zeitkorrelierte Rauschprozesse höherer Ordnung zur Modellierung des System- und Messrauschens benötigt, müssen die Matrizen, die den Zusammenhang zwischen Zustandsvektor bzw. Messwertvektor und dem jeweiligen zeitkorrelierten Rauschprozess beschreiben, entsprechend angepasst werden.

Der Nachteil einer Erweiterung des Zustandsvektors besteht, je nachdem wieviele Rauschprozesse welcher Ordnung berücksichtigt werden müssen, in einer unter Umständen drastischen Erhöhung des Rechenaufwandes. Zusätzlich können numerische Probleme auftreten. Häufig sind auch die zusätzlichen Zustände mit den zu Verfügung stehenden Messungen kaum beobachtbar<sup>7</sup>, die Schätzung des Rauschprozesszustandes ist unbrauchbar. Entscheidend bei der Erweiterung des Zustandsvektors ist jedoch nicht der geschätzte Rauschprozesszustand, sondern der Einfluss dieser Zustände auf die Propagation und das Update der Kovarianzmatrix des Schätzfehlers. Dieser Einfluss lässt sich durch die häufig zu beobachtende Praxis einer einfachen Erhöhung eines als weiß angenommenen Mess- und Systemrauschens nicht nachbilden. Trotz der angesprochenen Nachteile stellt die Erweiterung des Zustandsvektors, die in der englischsprachigen Literatur als *state augmentation* oder *shaping filter approach* zu finden ist, ein leistungsfähiges und universell einsetzbares Verfahren zur Berücksichtigung von zeitkorreliertem Rauschen dar.

## 6.5.2 Messwertdifferenzen

Ein als *measurement differencing* bekanntes Verfahren erlaubt es, auf eine Erweiterung des Zustandsvektors zu verzichten. Die Idee dieses Verfahrens besteht darin, durch einen gewichteten Mittelwert zweier aufeinanderfolgender, mit zeitkorreliertem Rauschen behafteter Messwerte ein neue Messgröße zu konstruieren, deren Messrauschen weiß ist.

<sup>7</sup>In der englischsprachigen Literatur findet man für diese Zustände daher die Bezeichnung 'crap states'

Diese neue Messgröße  $\tilde{z}_k$  ist gegeben durch

$$\tilde{z}_k = \tilde{y}_{k+1} - \mathbf{D}\tilde{y}_k . \quad (6.140)$$

Durch Einsetzen des Messmodells erhält man

$$\tilde{z}_k = \mathbf{H}_{k+1}\vec{x}_{k+1} + \vec{v}_{k+1} - \mathbf{D}(\mathbf{H}_k\vec{x}_k + \vec{v}_k) , \quad (6.141)$$

mit Hilfe des Systemmodells ergibt sich schließlich

$$\begin{aligned} \tilde{z}_k &= \mathbf{H}_{k+1}(\Phi_k\vec{x}_k + \mathbf{G}_k\vec{w}_k) + \vec{v}_{k+1} - \mathbf{D}(\mathbf{H}_k\vec{x}_k + \vec{v}_k) \\ &= (\mathbf{H}_{k+1}\Phi_k - \mathbf{D}\mathbf{H}_k)\vec{x}_k + (\mathbf{H}_{k+1}\mathbf{G}_k\vec{w}_k + \mathbf{D}\vec{v}_k + \vec{\mu}_{k+1} - \mathbf{D}\vec{v}_k) \\ \tilde{z}_k &= (\mathbf{H}_{k+1}\Phi_k - \mathbf{D}\mathbf{H}_k)\vec{x}_k + (\mathbf{H}_{k+1}\mathbf{G}_k\vec{w}_k + \vec{\mu}_{k+1}) . \end{aligned} \quad (6.142)$$

Unter der Voraussetzung, dass es sich beim Systemrauschen  $\vec{w}_k$  um weißes Rauschen handelt, ist das in der Größe  $\tilde{z}_k$  enthaltene Rauschen als Summe des Systemrauschens und des Rauschterms, das den zeitkorrelierten Messrauschprozess treibt, ebenfalls ein weißes Rauschen. Allerdings ist dieser Rauschterm mit dem Systemrauschen korreliert, diese Korrelation lässt sich jedoch berechnen und mit einem sog. generalized Kalman-Filter berücksichtigen. Details zu dieser Vorgehensweise sind in [19] zu finden.

Neben einem unbedeutenden Initialisierungsproblem ist ein Nachteil dieses Verfahrens, dass es sich nur schwer auf zeitkorrelierte Rauschprozesse höherer Ordnung erweitern lässt, da die resultierenden Gleichungen sehr unhandlich werden. Der Hauptnachteil ist jedoch, dass diese Vorgehensweise auf zeitkorreliertes Messrauschen beschränkt ist, zeitkorreliertes Systemrauschen lässt sich nicht berücksichtigen.

Ein alternatives Verfahren zur Berücksichtigung sowohl von zeitkorreliertem Mess- als auch von zeitkorreliertem Systemrauschen, das ohne eine Erweiterung des Zustandsvektors auskommt, wird in Abschnitt 9.3 vorgestellt.

## 6.6 Covariance Intersection

Der bisher vorgestellte Standard-Kalman-Filter-Algorithmus setzt voraus, dass das Messrauschen nicht mit dem zu schätzenden Systemzustand korreliert ist:

$$E[(\vec{x}_k - \hat{x}_k)\vec{v}_k^T] = \mathbf{0} \quad (6.143)$$

Ist diese Voraussetzung verletzt, ist der bisher vorgestellte Kalman-Filter suboptimal. Das kann – je nach Schätzproblem – irrelevant sein, aber auch von einer Verschlechterung der Performance bis zur Divergenz des Filters führen.

Die Ursachen für diese Kreuzkorrelationen können unterschiedlich sein. Eine mögliche Ursache ist zeitkorreliertes Messrauschen. Der Messrauschterm  $\vec{v}_k$  geht in die Zustandsschätzung  $\hat{x}_k^+$  ein. Liegen nun Zeitkorrelationen vor, so sind die Rauschterme  $\vec{v}_k$  und

$\vec{v}_{k+1}$  nicht unabhängig voneinander. Über den Propagationsschritt des Filters beeinflusst  $\vec{v}_k$  aber auch die a-priori-Zustandsschätzung einen Zeitschritt später,  $\vec{x}_{k+1}^-$ . Daher können nun  $\vec{x}_{k+1}^-$  und  $\vec{v}_{k+1}$  nicht unabhängig voneinander sein, auch wenn  $\vec{x}_k^-$  und  $\vec{v}_k$  noch unkorreliert waren, siehe hierzu auch Abschnitt 9.3.

Korrelationen zwischen Messwerten und Zustandsschätzung treten auch in SLAM<sup>8</sup>-Szenarien auf. Als Beispiel soll hier die Navigation eines Roboters in unbekanntem Gelände angeführt werden. Entdeckt dieser Roboter mit einem Laserscanner eine neue Landmarke, so wird anhand der aktuellen Roboterposition die Position dieser Landmarke geschätzt. Anschließend kann relativ zu dieser Landmarke navigiert werden. Der Fehler der geschätzten Roboterposition geht damit aber auch in die Positionsschätzung der Landmarke ein, und folglich in jede Abstandsmessung, die anschließend zu Navigationszwecken zu dieser Landmarke gemacht wird.

In dezentralen Datenfusionsproblemen sind Korrelationen ebenfalls praktisch unvermeidlich. Ein Beispiel hierfür wäre ein Sensornetzwerk, bei dem jeder mit einem Sensor ausgestattete Knoten seine Messwerte mit jedem anderen Knoten teilt. Angenommen, Knoten A verteilt eine Messung und die zugehörige Varianz an Knoten B und Knoten C. Würde nun ein Knoten D diese Messungen von Knoten B und Knoten C als unabhängig verarbeiten, so würde eine deutlich zu geringe Varianz der Zustandsschätzung resultieren – schließlich enthält die von Knoten C zur Verfügung gestellte Messung gegenüber der von Knoten B zur Verfügung gestellten Messung keine zusätzliche Information, es handelt sich ja nur um die eine, ursprünglich von Knoten A stammende Messung. Auch wenn in diesem einfachen Szenario eine Mehrfachverarbeitung von Messwerten sehr einfach verhindert werden könnte, ist dies bei großen, komplexen Netzwerken schwierig.

Eine zusätzliche Schwierigkeit besteht nun meist darin, dass Kreuzkorrelationen zwar vorliegen, diese aber unbekannt sind. In diesen Fällen kann anstelle eines Kalman-Filters der Covariance-Intersection-Algorithmus (CI) verwendet werden [60][128]. In den meisten Beschreibungen zu CI wird das Problem betrachtet, zwei Zufallsvektoren gleicher Dimension zu fusionieren. Dieses Problem kann als eine Untermenge des in den vorangegangenen Abschnitten betrachteten Schätzproblems verstanden werden: Betrachtet man ausschließlich den Messschritt und fordert  $\mathbf{H}_k = \mathbf{I}$ , so resultiert das angesprochene Schätzproblem, wobei einer der Zufallsvektoren die a-priori-Zustandsschätzung, der andere der Messwertvektor ist. Im Folgenden soll CI angepasst auf das schon beim Kalman-Filter verwendete, allgemeinere Messmodell beschrieben werden. Durch entsprechende Vereinfachungen kann dann zu den üblicherweise in der Literatur zu findenden CI-Gleichungen übergegangen werden.

Bevor auf die Filterung bei unbekannten Kreuzkorrelationen eingegangen wird, soll jedoch der Fall betrachtet werden, dass bekannte Kreuzkorrelationen zwischen Messwerten und Zustandsschätzung vorliegen.

### 6.6.1 Bekannte Kreuzkorrelationen

Die Gleichungen (6.8) und (6.9) stellen eine allgemeingültige Vorschrift dar, wie bei gaußverteilten Zufallsvektoren aus einem a-priori-Schätzwert  $\hat{\vec{x}}_k^-$  der a-posteriori-

<sup>8</sup>Simultaneous Location and Map Building



Schätzwert  $\hat{x}_k^+$  berechnet werden kann:

$$\begin{aligned}\hat{x}_k^+ &= \hat{x}_k^- + \mathbf{P}_{xy,k} \mathbf{P}_{yy,k}^{-1} (\tilde{y}_k - \hat{y}_k) \\ \mathbf{P}_{xx,k}^+ &= \mathbf{P}_{xx,k}^- - \mathbf{P}_{xy,k} \mathbf{P}_{yy,k}^{-1} \mathbf{P}_{yx,k}\end{aligned}$$

Zur Herleitung des Kalman-Filter-Messschrittes wurde vom Messmodell Gl. (6.3),

$$\tilde{y}_k = \mathbf{H}_k \vec{x}_k + \vec{v}_k ,$$

ausgegangen. Dabei wurde angenommen, dass die Korrelation von Schätzfehler und Messfehler,  $\mathbf{P}_{xv,k}$ , verschwindet. Es ist jedoch auch möglich, ohne diese Annahme einen allgemeineren Messschritt herzuleiten, siehe hierzu Abschnitt 9.3. In diesem Fall erhält man

$$\mathbf{P}_{xy,k} = \mathbf{P}_{xx,k}^- \mathbf{H}_k^T + \mathbf{P}_{xv,k} \quad (6.144)$$

$$\mathbf{P}_{yy,k} = \mathbf{H}_k \mathbf{P}_{xx,k}^- \mathbf{H}_k^T + \mathbf{H}_k \mathbf{P}_{xv,k} + \mathbf{P}_{vx,k} \mathbf{H}_k^T + \mathbf{R}_k \quad (6.145)$$

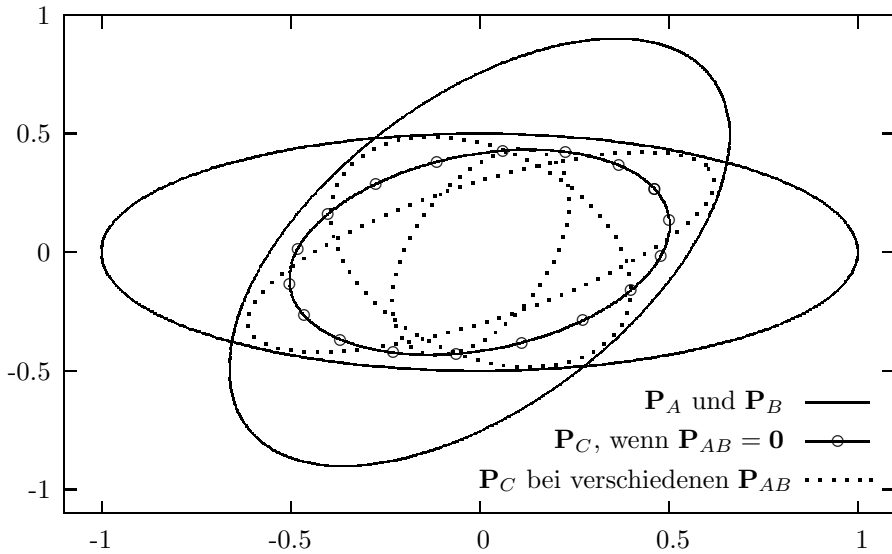
und mit (6.8) und (6.9) unter Verwendung der Abkürzung

$$\mathbf{K}_k = \left( \mathbf{P}_{xx,k}^- \mathbf{H}_k^T + \mathbf{P}_{xv,k}^- \right) \left( \mathbf{H}_k \mathbf{P}_{xx,k}^- \mathbf{H}_k^T + \mathbf{H}_k \mathbf{P}_{xv,k}^- + \mathbf{P}_{vx,k}^- \mathbf{H}_k^T + \mathbf{P}_{vv,k} \right)^{-1}$$

schließlich

$$\begin{aligned}\hat{x}_k^+ &= \hat{x}_k^- + \mathbf{K}_k \left( \tilde{y}_k - \mathbf{H}_k \hat{x}_k^- \right) \\ \mathbf{P}_{xx,k}^+ &= \mathbf{P}_{xx,k}^- - \mathbf{K}_k \mathbf{H}_k \left( \mathbf{P}_{xx,k}^- + \mathbf{P}_{vx,k}^- \right) .\end{aligned}$$

Um den Einfluss von Kreuzkorrelationen zu verdeutlichen, wird die Fusion zweier zweidimensionaler Zufallsvektoren  $\vec{a}$  und  $\vec{b}$  betrachtet, die die Kovarianzmatrizen  $\mathbf{P}_A$  und  $\mathbf{P}_B$  besitzen. Das Fusionsergebnis ist ein Vektor  $\vec{c}$  mit der Kovarianzmatrix  $\mathbf{P}_C$ . Die 1-Sigma-Kovarianzellipsen der Vektoren  $\vec{a}$  und  $\vec{b}$ , die sich anhand von  $\mathbf{P}_A$  und  $\mathbf{P}_B$  bestimmen lassen, sind in Abb. 6.4 dargestellt. Zusätzlich sind in Abb. 6.4 auch noch die aus  $\mathbf{P}_C$  bestimmte Kovarianzellipse des Fusionsergebnisses für unterschiedliche Szenarien zu sehen: Zum Einen, wenn  $\vec{a}$  und  $\vec{b}$  unkorreliert sind ( $\mathbf{P}_{AB} = \mathbf{0}$ ); zum Anderen, wenn  $\vec{a}$  und  $\vec{b}$  auf unterschiedliche Weise korreliert sind. Man erkennt, dass bei vorliegenden Korrelationen die Unsicherheit in einer von der Korrelation abhängigen Vorzugsrichtung geringer ausfällt als im unkorrelierten Fall, senkrecht dazu ergibt sich jedoch eine größere Unsicherheit. Würden nun bei der Fusion der Zufallsvektoren die Kreuzkorrelationen ignoriert, würde in dieser Richtung eine zu geringe Varianz des Fusionsergebnisses berechnet. Ein Filter würde dadurch in diesem Szenario mit der Zeit über-optimistisch,



**Abbildung 6.4:** Fusion zweier Zufallsvektoren bei unterschiedlichen Kreuzkorrelationen.

was die Güte der Zustandsschätzung anbelangt – die Folge ist eine zu geringe Gewichtung der Messwerte, das Filter hält zu stark an der eigenen Zustandsschätzung fest, was bis zur Divergenz des Filters führen kann. Sind die Kreuzkorrelationen anders als in diesem Beispiel unbekannt, muss dennoch sichergestellt werden, dass keine zu geringe Schätzfehlerkovarianz auftritt, die zu dieser Entkopplung von Filter und realer Welt führen kann.

Es lässt sich zeigen, dass für alle möglichen Kreuzkorrelationen die Kovarianzellipse des Fusionsergebnisses  $\vec{c}$  innerhalb der Schnittmenge der Kovarianzellipsen von  $\vec{a}$  und  $\vec{b}$  liegt. Die Grundidee von CI besteht darin, bei unbekannten Kreuzkorrelationen die Kovarianzellipse von  $\vec{c}$  so zu wählen, dass sie diese Schnittmenge einschließt. Ist das gegeben, spricht man von einer konservativen Schätzung.

## 6.6.2 Unbekannte Kreuzkorrelationen

Um eine konservative Schätzung zu erhalten, werden in [48] folgende Vorüberlegungen angestellt: Wenn  $\mathbf{C}$  eine positiv semi-definite, symmetrische Matrix ist mit

$$\mathbf{C} = \begin{pmatrix} \mathbf{P}_{xx} & \mathbf{P}_{xv} \\ \mathbf{P}_{vx} & \mathbf{P}_{vv} \end{pmatrix}, \quad (6.146)$$

dann gilt

$$\mathbf{C}_\alpha \geq \mathbf{C} \quad (6.147)$$

mit

$$\mathbf{C}_\alpha = \begin{pmatrix} \frac{1}{1-\alpha} \mathbf{P}_{xx} & 0 \\ 0 & \frac{1}{\alpha} \mathbf{P}_{vv} \end{pmatrix} \quad (6.148)$$

wobei  $0 \leq \alpha \leq 1$  gilt. Ein Beweis dieser Tatsache ist ebenfalls in [48] zu finden. Gleichung (6.147) ist dabei so zu verstehen, dass die Matrix  $\mathbf{C}_\alpha - \mathbf{C}$  eine positiv semi-definite Matrix ist.

Soll anhand von (6.8) und (6.9) ein Messschritt hergeleitet werden und ist  $\mathbf{P}_{xv,k}$  unbekannt, so erhält man eine konservative Schätzung wenn man in Gl. (6.144)–(6.145) nicht mit den Einträgen der Matrix  $\mathbf{C}$ , sondern mit den Einträgen von  $\mathbf{C}_\alpha$  arbeitet. Auf diese Weise ergibt sich

$$\mathbf{P}_{xy,k} = \frac{1}{1-\alpha} \mathbf{P}_{xx,k}^- \mathbf{H}_k^T \quad (6.149)$$

$$\mathbf{P}_{yy,k} = \frac{1}{1-\alpha} \mathbf{H}_k \mathbf{P}_{xx,k}^- \mathbf{H}_k^T + \frac{1}{\alpha} \mathbf{R}_k \quad (6.150)$$

und folglich mit der Abkürzung

$$\begin{aligned} \mathbf{K}_{\alpha,k} &= \mathbf{P}_{xy,k} \mathbf{P}_{yy,k}^{-1} \\ &= \frac{1}{1-\alpha} \mathbf{P}_{xx,k}^- \mathbf{H}_k^T \left( \frac{1}{1-\alpha} \mathbf{H}_k \mathbf{P}_{xx,k}^- \mathbf{H}_k^T + \frac{1}{\alpha} \mathbf{R}_k \right)^{-1} \\ &= \frac{1}{1-\alpha} \mathbf{P}_{xx,k}^- \mathbf{H}_k^T \left( \frac{\alpha}{\alpha(1-\alpha)} \mathbf{H}_k \mathbf{P}_{xx,k}^- \mathbf{H}_k^T + \frac{1-\alpha}{\alpha(1-\alpha)} \mathbf{R}_k \right)^{-1} \\ \mathbf{K}_{\alpha,k} &= \alpha \mathbf{P}_{xx,k}^- \mathbf{H}_k^T \left( \alpha \mathbf{H}_k \mathbf{P}_{xx,k}^- \mathbf{H}_k^T + (1-\alpha) \mathbf{R}_k \right)^{-1} \end{aligned} \quad (6.151)$$

schließlich

$$\hat{\tilde{x}}_k^+ = \hat{\tilde{x}}_k^- - \mathbf{K}_{\alpha,k} \left( \mathbf{H}_k \hat{\tilde{x}}_k^- - \tilde{y}_k \right) \quad (6.152)$$

$$\mathbf{P}_{xx,k}^+ = \frac{1}{1-\alpha} \left( \mathbf{P}_{xx,k}^- - \mathbf{K}_{\alpha,k} \mathbf{H}_k \mathbf{P}_{xx,k}^- \right) . \quad (6.153)$$

Die Gleichungen (6.151)–(6.153) stellen den CI-Messschritt dar.

In der Literatur wird nun bei der Beschreibung von CI üblicherweise nicht nur wie eingangs erwähnt eine Untermenge ( $\mathbf{H}_k = \mathbf{I}$ ) des hier betrachteten Schätzproblems betrachtet, meist werden die CI-Gleichungen in Invers-Kovarianzform dargestellt.

Die Herleitung der Invers-Kovarianzform von Gl. (6.151)–(6.153) soll im Folgenden skizziert werden, wobei zur besseren Übersichtlichkeit kurz  $\mathbf{P}_{xx,k} = \mathbf{P}_k$  geschrieben wird. Einsetzen von  $\mathbf{K}_{\alpha,k}$  in Gl. (6.153) liefert

$$\mathbf{P}_k^+ = \frac{1}{1-\alpha} \left( \mathbf{P}_k^- - \alpha \mathbf{P}_k^- \mathbf{H}_k^T \left( \alpha \mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^T + (1-\alpha) \mathbf{R}_k \right)^{-1} \mathbf{H}_k \mathbf{P}_k^- \right) . \quad (6.154)$$

Mit den Substitutionen  $\mathbf{A} = (1 - \alpha)\mathbf{R}_k$ ,  $\mathbf{B} = \mathbf{I}$ ,  $\mathbf{C}^T = \alpha\mathbf{H}_k\mathbf{P}_k^-\mathbf{H}_k^T$  folgt aus Gl. (A.3)

$$\mathbf{P}_k^+ = \frac{1}{1 - \alpha} \left[ \mathbf{P}_k^- - \alpha\mathbf{P}_k^-\mathbf{H}_k^T \frac{1}{1 - \alpha} \mathbf{R}_k^{-1} \cdot \left( \mathbf{I} + \alpha\mathbf{H}_k\mathbf{P}_k^-\mathbf{H}_k^T \frac{1}{1 - \alpha} \mathbf{R}_k^{-1} \right)^{-1} \mathbf{H}_k\mathbf{P}_k^- \right]. \quad (6.155)$$

Mit  $\mathbf{A}^{-1} = \mathbf{P}_k^-$ ,  $\mathbf{B} = \frac{\alpha}{1 - \alpha}\mathbf{H}_k^T\mathbf{R}_k^{-1}$ ,  $\mathbf{C}^T = \mathbf{H}_k$  und Gl. (A.2) erhält man

$$\begin{aligned} \mathbf{P}_k^+ &= \frac{1}{1 - \alpha} \left( (\mathbf{P}_k^-)^{-1} + \frac{\alpha}{1 - \alpha} \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{H}_k \right)^{-1} \\ \mathbf{P}_k^+ &= \frac{1}{1 - \alpha} \left( \frac{1 - \alpha}{1 - \alpha} (\mathbf{P}_k^-)^{-1} + \frac{\alpha}{1 - \alpha} \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{H}_k \right)^{-1} \\ \mathbf{P}_k^+ &= \left( (1 - \alpha)(\mathbf{P}_k^-)^{-1} + \alpha\mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{H}_k \right)^{-1} \\ (\mathbf{P}_k^+)^{-1} &= (1 - \alpha)(\mathbf{P}_k^-)^{-1} + \alpha\mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{H}_k \end{aligned} \quad (6.156)$$

Die Gain-Matrix lässt sich nun ebenfalls umschreiben:

$$\begin{aligned} \mathbf{K}_{\alpha,k} &= \alpha\mathbf{P}_k^-\mathbf{H}_k^T (\alpha\mathbf{H}_k\mathbf{P}_k^-\mathbf{H}_k^T + (1 - \alpha)\mathbf{R}_k)^{-1} \\ &= \alpha (\mathbf{P}_k^+(\mathbf{P}_k^+)^{-1}) \mathbf{P}_k^-\mathbf{H}_k^T \left( \frac{1}{1 - \alpha} \mathbf{R}_k^{-1} (1 - \alpha)\mathbf{R}_k \right) \\ &\quad \cdot (\alpha\mathbf{H}_k\mathbf{P}_k^-\mathbf{H}_k^T + (1 - \alpha)\mathbf{R}_k)^{-1} \\ &= \alpha (\mathbf{P}_k^+(\mathbf{P}_k^+)^{-1}) \mathbf{P}_k^-\mathbf{H}_k^T \left( \frac{1}{1 - \alpha} \mathbf{R}_k^{-1} (1 - \alpha)\mathbf{R}_k \right) \\ &\quad \cdot \frac{1}{1 - \alpha} \mathbf{R}_k^{-1} \left( \mathbf{I} + \alpha\mathbf{H}_k\mathbf{P}_k^-\mathbf{H}_k^T \frac{1}{1 - \alpha} \mathbf{R}_k^{-1} \right)^{-1} \end{aligned} \quad (6.157)$$

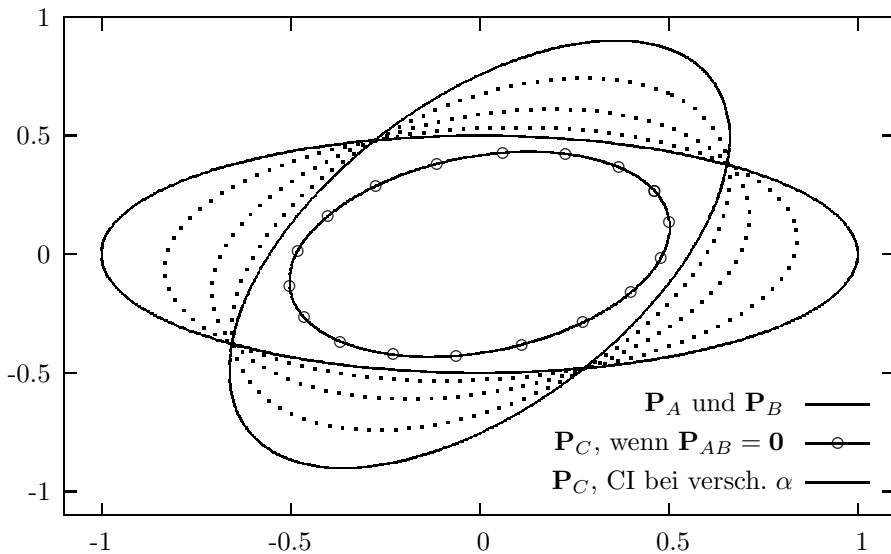
Einsetzen von Gl. (6.156) liefert

$$\begin{aligned}
\mathbf{K}_{\alpha,k} &= \alpha \mathbf{P}_k^+ \left( (1-\alpha)(\mathbf{P}_k^-)^{-1} + \alpha \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{H}_k \right) \mathbf{P}_k^- \mathbf{H}_k^T \\
&\quad \cdot \frac{1}{1-\alpha} \mathbf{R}_k^{-1} \left( \mathbf{I} + \alpha \mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^T \frac{1}{1-\alpha} \mathbf{R}_k^{-1} \right)^{-1} \\
&= \alpha \mathbf{P}_k^+ \left( \mathbf{I} + \alpha \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{H}_k \frac{1}{1-\alpha} \mathbf{P}_k^- \right) (1-\alpha)(\mathbf{P}_k^-)^{-1} \mathbf{P}_k^- \mathbf{H}_k^T \\
&\quad \cdot \frac{1}{1-\alpha} \mathbf{R}_k^{-1} \left( \mathbf{I} + \alpha \mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^T \frac{1}{1-\alpha} \mathbf{R}_k^{-1} \right)^{-1} \\
&= \alpha \mathbf{P}_k^+ \left( \mathbf{I} + \alpha \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{H}_k \frac{1}{1-\alpha} \mathbf{P}_k^- \right) \mathbf{H}_k^T \mathbf{R}_k^{-1} \\
&\quad \cdot \left( \mathbf{I} + \alpha \mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^T \frac{1}{1-\alpha} \mathbf{R}_k^{-1} \right)^{-1} \\
&= \alpha \mathbf{P}_k^+ \left( \mathbf{H}_k^T \mathbf{R}_k^{-1} + \alpha \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{H}_k \frac{1}{1-\alpha} \mathbf{P}_k^- \mathbf{H}_k^T \mathbf{R}_k^{-1} \right) \\
&\quad \cdot \left( \mathbf{I} + \alpha \mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^T \frac{1}{1-\alpha} \mathbf{R}_k^{-1} \right)^{-1} \\
&= \alpha \mathbf{P}_k^+ \mathbf{H}_k^T \mathbf{R}_k^{-1} \left( \mathbf{I} + \alpha \mathbf{H}_k \frac{1}{1-\alpha} \mathbf{P}_k^- \mathbf{H}_k^T \mathbf{R}_k^{-1} \right) \\
&\quad \cdot \left( \mathbf{I} + \alpha \mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^T \frac{1}{1-\alpha} \mathbf{R}_k^{-1} \right)^{-1} \\
\mathbf{K}_{\alpha,k} &= \alpha \mathbf{P}_k^+ \mathbf{H}_k^T \mathbf{R}_k^{-1} .
\end{aligned} \tag{6.158}$$

Multiplikation von Gl. (6.152) von links mit  $(\mathbf{P}_k^+)^{-1}$  ergibt

$$\begin{aligned}
(\mathbf{P}_k^+)^{-1} \hat{\tilde{x}}_k^+ &= (\mathbf{P}_k^+)^{-1} \hat{\tilde{x}}_k^- - (\mathbf{P}_k^+)^{-1} \mathbf{K}_{\alpha,k} \left( \mathbf{H}_k \hat{\tilde{x}}_k^- - \tilde{y}_k \right) \\
&= (\mathbf{P}_k^+)^{-1} \hat{\tilde{x}}_k^- - (\mathbf{P}_k^+)^{-1} \alpha \mathbf{P}_k^+ \mathbf{H}_k^T \mathbf{R}_k^{-1} \left( \mathbf{H}_k \hat{\tilde{x}}_k^- - \tilde{y}_k \right) \\
&= (\mathbf{P}_k^+)^{-1} \hat{\tilde{x}}_k^- - \alpha \mathbf{H}_k^T \mathbf{R}_k^{-1} \left( \mathbf{H}_k \hat{\tilde{x}}_k^- - \tilde{y}_k \right) \\
&= ((\mathbf{P}_k^+)^{-1} - \alpha \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{H}_k) \hat{\tilde{x}}_k^- + \alpha \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{H}_k \tilde{y}_k \\
(\mathbf{P}_k^+)^{-1} \hat{\tilde{x}}_k^+ &= (1-\alpha)(\mathbf{P}_k^-)^{-1} \hat{\tilde{x}}_k^- + \alpha \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{H}_k \tilde{y}_k \\
\hat{\tilde{x}}_k^+ &= \mathbf{P}_k^+ \left( (1-\alpha)(\mathbf{P}_k^-)^{-1} \hat{\tilde{x}}_k^- + \alpha \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{H}_k \tilde{y}_k \right)
\end{aligned} \tag{6.159}$$

Die Invers-Kovarianzform der Gleichungen (6.151)–(6.153) ist durch (6.156) und (6.159) gegeben.



**Abbildung 6.5:** Fusion zweier in unbekannter Weise korrelierter Zufallsvektoren mit dem CI-Algorithmus für unterschiedliche Parameter  $\alpha$ .

In den CI-Gleichungen tritt nun ein hier mit  $\alpha$  bezeichneter Parameter auf. Abb. 6.5 zeigt nun für das schon im vorigen Abschnitt betrachtete Schätzproblem die Kovarianzellipse des Fusionsergebnisses für unterschiedliche Parameter  $\alpha$ . Der Parameter  $\alpha$  kann als ein Scharparameter für die Kovarianzellipse des Fusionsergebnisses gesehen werden; es lässt sich leicht überlegen, dass die bei der Fusion resultierende Kovarianzmatrix für  $\alpha = 0$  gleich der Kovarianzmatrix des einen Zufallsvektors und für  $\alpha = 1$  gleich der Kovarianzmatrix des anderen Zufallsvektors ist. Bei dieser Parameterwahl wäre der gefundene Schätzer zwar sicherlich konservativ, die bei der Fusion gewünschte Verringerung der Schätzfehlervarianz würde jedoch ausbleiben. Der Parameter  $\alpha$  sollte so gewählt werden, dass die Spur der a-posteriori-Kovarianzmatrix minimal wird – dieses Optimierungsproblem kann numerisch behandelt werden<sup>9</sup>, in [48] werden aber auch analytische Lösungen für  $\alpha$  sowohl bei skalaren als auch bei vektorwertigen Messungen angegeben.

Abschließend sei noch darauf hingewiesen, dass eine Vielzahl von Erweiterungen und Verfeinerungen des CI-Algorithmus existieren. So wird z.B. in [49] aufgezeigt, wie bei unbekannten aber beschränkten Korrelationen eine kleinere Kovarianzmatrix des Fusionsergebnisses erzielt werden kann, als dies mit dem hier vorgestellten CI-Algorithmus möglich wäre.

<sup>9</sup>Im einfachsten Fall durch Probieren verschiedener Parameter  $\alpha$ .

## 6.7 Adaptive Filterung

Allgemein gesagt ist ein adaptives Filter in der Lage, sich an aktuell vorliegende Bedingungen anzupassen – das kann notwendig sein, weil diese entweder im voraus nicht bekannt sind, oder sich während des Betriebes des Filters ändern können. In Abschnitt 8 zur GPS/INS-Integration werden beispielsweise Filter entwickelt, die im vorhinein unbekannte Parameter, die Biase der Inertialsensoren, schätzen und dadurch ihre Performance verbessern. So gesehen könnte man diese Filter auch als adaptive Filter bezeichnen. Auf diese Art der Adaptivität, die Schätzung unbekannter Parameter durch Erweiterung des Zustandsvektors des Filters, soll hier jedoch nicht weiter eingegangen werden. Ein Grundproblem beim Einsatz eines Kalman-Filters lässt sich auf diese Weise nämlich nicht lösen: Um ein gut funktionierendes Filter zu erhalten, müssen die Kovarianzmatrix des Systemrauschens,  $\mathbf{Q}_k$ , und die Kovarianzmatrix des Messrauschens,  $\mathbf{R}_k$ , geeignet gewählt werden. Die Ansätze zur Lösung dieses Problems lassen sich grob in drei Klassen einteilen.

Zum Einen kann anhand der Sensordaten versucht werden, die unbekannten Kovarianzmatrizen zu schätzen – das in Abschnitt 9.5 vorgestellte Verfahren ist ein Beispiel hierfür. Dabei wurde ausgenutzt, dass bei dem betrachteten Problem das Systemrauschen durch das die Inertialsensordaten behaftende Rauschen gegeben ist, da diese als bekannte Eingangsgrößen modelliert wurden.

Ein anderer Ansatz besteht in der Analyse der Innovation des Filters, also der Differenz zwischen prädizierten und vorliegenden Messwerten, siehe hierzu beispielsweise [92],[88].

Schließlich ist die Gruppe der Multiple-Model-Algorithmen zu nennen; hierbei wird eine Bank von Filtern betrieben, wobei jeder der elementaren Filter unter einer anderen Annahme entworfen wird – diese unterschiedlichen Annahmen können sich auf die Kovarianzmatrizen des System- und/oder Messrauschens beziehen, es können aber auch verschiedene Systemmodelle angenommen werden<sup>10</sup>. Schließlich wird durch Analyse der Innovationen der elementaren Filter eine Gewichtung bestimmt, die festlegt, wie die Schätzungen der elementaren Filter zum endgültigen Schätzergebnis beitragen. Der wichtigste Vertreter dieser Gruppe ist der Interacting Multiple Model (IMM) Filter, der im Folgenden beschrieben werden soll.

### 6.7.1 Interacting Multiple Model Filter: Problemformulierung

Ein IMM hat das Ziel, den Zustand des Systems

$$\vec{x}_k = \Phi_k(s_k)\vec{x}_{k-1} + \mathbf{G}_k(s_k)\vec{w}_k \quad (6.160)$$

zu schätzen. Dabei stehen Messungen zur Verfügung, deren Beziehung zum Systemzustand durch das Messmodell

$$\vec{y}_k = \mathbf{H}_k(s_k)\vec{x}_k + \mathbf{C}_k(s_k)\vec{v}_k \quad (6.161)$$

<sup>10</sup>Vor allem im Bereich Target Tracking kommen häufig Multiple-Model-Ansätze mit verschiedenen Systemmodellen zum Einsatz, z.B. für Geradeausflug und Kurvenflug des Ziels.

beschrieben wird. Bei diesem System- und Messmodell sind die Transitionsmatrix, die Messmatrix sowie die Einwirkung von System- und Messrauschen in Abhängigkeit eines Parameters  $s_k$  formuliert. Bei  $s_k$  handelt es sich um eine Markov-Kette, die Werte zwischen 1 und  $m$  annehmen kann. Die Markov-Kette wird durch eine Übergangswahrscheinlichkeitsmatrix  $\mathbf{M}$  beschrieben, wobei das Element  $M_{ij}$  die Wahrscheinlichkeit angibt, dass sich beim Übergang vom Zeitschritt  $k-1$  zum Zeitschritt  $k$  der Zustand der Markov-Kette von  $s_{k-1} = i$  zu  $s_k = j$  ändert. Diese Markov-Kette wirkt wie ein Schalter, der zwischen verschiedenen Transitionsmatrizen, Messmatrizen usw. hin- und herschaltet und so das Systemverhalten beeinflusst. Es handelt sich folglich um ein nichtlineares System, da der  $\vec{x}_k$  nicht linear von  $\vec{x}_{k-1}, s_k$  abhängt<sup>11</sup>.

Ein optimaler Schätzalgorithmus für ein solches System würde den bedingten Erwartungswert

$$E[\vec{x}_k | \mathbf{Y}_k] = \sum_{q=1}^{m^k} E[\vec{x}_k | S_q, \mathbf{Y}_k] P(S_q | \mathbf{Y}_k) \quad (6.162)$$

berechnen, wobei  $\mathbf{Y}_k$  für alle bis zum Zeitschritt  $k$  aufgetretenen Messwerte  $\mathbf{Y}_k = \tilde{y}_k, \tilde{y}_{k-1}, \dots, \tilde{y}_0$  steht.  $S_q$  bezeichnet die  $q$ -te Sequenz von  $m^k$  möglichen Sequenzen der Zustände der Markov-Kette,  $P(S_q | \mathbf{Y}_k)$  ist die Wahrscheinlichkeit dafür, dass die  $q$ -te Sequenz der tatsächlich vorliegenden Sequenz von Zuständen der Markov-Kette entspricht. Offensichtlich wächst die Anzahl der möglichen Sequenzen exponentiell mit der Zeit, so dass ein praktisch implementierbarer Filter nicht in der Lage ist, alle möglichen Sequenzen für alle Zeiten zu betrachten.

Die Grundidee des IMM besteht nun darin, Gl. (6.162) anhand von

$$E[\vec{x}_k | \mathbf{Y}_k] = \sum_{j=1}^m E[\vec{x}_k | s_k = j, \mathbf{Y}_k] P(s_k = j | \mathbf{Y}_k) \quad (6.163)$$

zu approximieren<sup>12</sup>. Dazu laufen  $m$  elementare Filter parallel, wobei jeder dieser Filter für einen bestimmten Zustand  $s_k$  der Markov-Kette optimal ist, d.h. sein System- und Messmodell für ein bestimmtes  $s_k$  gerade mit Gl. (6.160) und (6.161) übereinstimmt. Diese elementaren Filter werden auch als Modelle bezeichnet. Gleichzeitig wird die Wahrscheinlichkeit der einzelnen Modelle geschätzt, was einer Schätzung des Zustandes der Markov-Kette entspricht.

## 6.7.2 Herleitung der IMM-Filtergleichungen

Zur Herleitung des IMM werden folgende Schreibweisen eingeführt, um eine kompakte Darstellung zu ermöglichen:

<sup>11</sup>Ist  $s_k$  bekannt, liegt ein lineares System vor.

<sup>12</sup>Damit ist sofort klar, dass es sich beim IMM um einen suboptimalen Filteralgorithmus handelt.



$$\mu_{k-1}^{j_k} = P(s_k = j | \mathbf{Y}_{k-1})$$

Bedingte Wahrscheinlichkeit unter der Voraussetzung aller bis zum Zeitpunkt  $k - 1$  eingetroffenen Messungen, dass das Modell  $j$  im Zeitschritt  $k$  korrekt ist.

$$\mu_{k-1}^{j_k | i_{k-1}} = P(s_k = j | s_{k-1} = i, \mathbf{Y}_{k-1})$$

Bedingte Wahrscheinlichkeit unter der Voraussetzung aller bis zum Zeitpunkt  $k - 1$  eingetroffenen Messungen, dass das Modell  $j$  im Zeitschritt  $k$  korrekt ist, wenn das Modell  $i$  im Zeitschritt  $k - 1$  korrekt war.

$$\mu_{k-1}^{j_k | i_{k-1}} = M_{ij}$$

Element der Übergangswahrscheinlichkeitsmatrix

$$p(\vec{x}_{k-1} | s_k = i, \mathbf{Y}_{k-1})$$

Wahrscheinlichkeitsdichte des Systemzustands, basierend auf Modell  $i$  und allen bis zum Zeitpunkt  $k - 1$  vorliegenden Messwerten. Hierbei handelt es sich um die WDF des elementaren Filters  $i$ .

Die a-posteriori-Zustandsschätzung des elementaren Filters  $i$  wird mit  $\hat{x}_{i,k}^+$  bezeichnet, die zugehörige Kovarianzmatrix mit  $\mathbf{P}_{i,k}^+$ . Die Zustandsschätzung der Filterbank basierend auf allen elementaren Filtern wird mit  $\hat{x}_k^+$ , die zugehörige Kovarianz mit  $\hat{\mathbf{P}}_k^+$  bezeichnet. Alle anderen im folgenden verwendeten Bezeichnungen können anhand dieser Definitionen erschlossen werden.

Der Ablauf des IMM-Algorithmus lässt sich in drei Schritte unterteilen, das Mischen, den Propagationsschritt und den Messschritt. Die nachfolgende Herleitung dieser Schritte orientiert sich an [15].

### Mischen

Der erste Schritt des IMM-Algorithmus ist das Mischen. Die Markov-Kette propagiert in der Zeit, daher müssen auch die Schätzungen des Zustands der Markov-Kette und des Systemzustandes angepasst werden. Dazu müssen folgende Übergänge ermittelt werden:

$$\mu_{k-1}^{j_{k-1}} \rightarrow \mu_{k-1}^{i_k}, \quad \forall i, j \quad (6.164)$$

$$p(\vec{x}_{k-1} | s_{k-1} = j, \mathbf{Y}_{k-1}) \rightarrow p(\vec{x}_{k-1} | s_k = i, \mathbf{Y}_{k-1}), \quad \forall i, j \quad (6.165)$$

Die Chapman-Kolmogorov-Gl. (7.7) angewandt auf eine Markov-Kette liefert

$$\mu_{k-1}^{i_k} = \sum_{j=1}^m \mu_{k-1}^{i_k | j_{k-1}} \mu_{k-1}^{j_{k-1}} = \sum_{j=1}^m M_{ij} \mu_{k-1}^{j_{k-1}}, \quad (6.166)$$

der gesuchte Zusammenhang Gl. (6.164) ist damit gefunden.

Mit der Bayes'schen Regel folgt ausserdem

$$\mu_{k-1}^{j_{k-1} | i_k} \mu_{k-1}^{i_k} = \mu_{k-1}^{i_k | j_{k-1}} \mu_{k-1}^{j_{k-1}} = M_{ij} \mu_{k-1}^{j_{k-1}} \quad (6.167)$$

und damit

$$\mu_{k-1}^{j_{k-1}|i_k} = \frac{M_{ij} \mu_{k-1}^{j_{k-1}}}{\mu_{k-1}^{i_k}} = \frac{M_{ij} \mu_{k-1}^{j_{k-1}}}{\sum_{j=1}^m M_{ij} \mu_{k-1}^{j_{k-1}}} . \quad (6.168)$$

Die gesuchte Dichte Gl. (6.165) lässt sich darstellen als

$$p(\vec{x}_{k-1}|s_k = i, \mathbf{Y}_{k-1}) = \sum_{j=1}^m p(\vec{x}_{k-1}|s_k = i, s_{k-1} = j, \mathbf{Y}_{k-1}) \mu_{k-1}^{j_{k-1}|i_k} . \quad (6.169)$$

Mit Hilfe der Bayes'sche Regel kann man schreiben

$$\begin{aligned} p(\vec{x}_{k-1}|s_k = i, s_{k-1} = j, \mathbf{Y}_{k-1}) P(s_k = i|s_{k-1} = j, \mathbf{Y}_{k-1}) \\ = P(s_k = i|\vec{x}_{k-1}, s_{k-1} = j, \mathbf{Y}_{k-1}) p(\vec{x}_{k-1}|s_{k-1} = j, \mathbf{Y}_{k-1}) \end{aligned} \quad (6.170)$$

und schließlich

$$\begin{aligned} p(\vec{x}_{k-1}|s_k = i, s_{k-1} = j, \mathbf{Y}_{k-1}) \\ = \frac{P(s_k = i|\vec{x}_{k-1}, s_{k-1} = j, \mathbf{Y}_{k-1})}{P(s_k = i|s_{k-1} = j, \mathbf{Y}_{k-1})} p(\vec{x}_{k-1}|s_{k-1} = j, \mathbf{Y}_{k-1}) . \end{aligned} \quad (6.171)$$

Setzt man  $s_{k-1} = j$  voraus, so beeinflusst das die Wahrscheinlichkeit für  $s_k = i$ . In diesem Fall hat jedoch die zusätzliche Voraussetzung  $\vec{x}_{k-1}$  keinen Einfluss auf die Wahrscheinlichkeit für  $s_k = i$ . Aufgrund dieser Überlegungen folgt

$$P(s_k = i|\vec{x}_{k-1}, s_{k-1} = j, \mathbf{Y}_{k-1}) = P(s_k = i|s_{k-1} = j, \mathbf{Y}_{k-1}) , \quad (6.172)$$

Einsetzen in Gl. (6.171) liefert

$$\begin{aligned} p(\vec{x}_{k-1}|s_k = i, s_{k-1} = j, \mathbf{Y}_{k-1}) \\ = \frac{P(s_k = i|s_{k-1} = j, \mathbf{Y}_{k-1})}{P(s_k = i|s_{k-1} = j, \mathbf{Y}_{k-1})} p(\vec{x}_{k-1}|s_{k-1} = j, \mathbf{Y}_{k-1}) \\ = p(\vec{x}_{k-1}|s_{k-1} = j, \mathbf{Y}_{k-1}) . \end{aligned} \quad (6.173)$$

Damit kann Gl. (6.169) vereinfacht werden:

$$p(\vec{x}_{k-1}|s_k = i, \mathbf{Y}_{k-1}) = \sum_{j=1}^m p(\vec{x}_{k-1}|s_{k-1} = j, \mathbf{Y}_{k-1}) \mu_{k-1}^{j_{k-1}|i_k} \quad (6.174)$$

Die Wahrscheinlichkeitsdichte des Systemzustandes  $\vec{x}_{k-1}$  unter Voraussetzung des Modells  $i$  ist also durch eine gewichtete Summe von  $m$  Gaußverteilungen gegeben. Diese

Dichte muss nun durch eine einzelne Gaußverteilung approximiert werden, um mit dieser für den auf das Mischen folgenden Propagationsschritt den elementaren Filter  $i$  initialisieren zu können.

Gesucht sind also Erwartungswert und Kovarianzmatrix einer Summe von Gaußverteilungen. Zur Herleitung wird allgemein ein Zufallsvektor  $\vec{z}$  betrachtet, dessen WDF  $p(\vec{z})$  durch eine gewichtete Summe von Gaußverteilungen  $p_i(\vec{z})$

$$p(\vec{z}) = \sum_{i=1}^m \alpha_i p_i(\vec{z}) \quad (6.175)$$

gegeben ist, wobei

$$\sum_{i=1}^m \alpha_i = 1 \quad (6.176)$$

und

$$p_i(\vec{z}) = \frac{1}{\sqrt{(2\pi)^n |\mathbf{P}_i|}} e^{-\frac{1}{2}(\vec{z} - \hat{\vec{z}}_i)^T \mathbf{P}_i^{-1} (\vec{z} - \hat{\vec{z}}_i)} \quad (6.177)$$

gilt,  $n$  ist die Anzahl der Komponenten des Zufallsvektors  $\vec{z}$ .

Der Erwartungswert von  $\vec{z}$  ergibt sich zu

$$E[\vec{z}] = \int_{-\infty}^{\infty} \vec{z} \sum_{i=1}^m \alpha_i p_i(\vec{z}) d\vec{z} = \sum_{i=1}^m \alpha_i \int_{-\infty}^{\infty} \vec{z} p_i(\vec{z}) d\vec{z} = \sum_{i=1}^m \alpha_i \hat{\vec{z}}_i. \quad (6.178)$$

Die Kovarianzmatrix des Zufallsvektors  $\vec{z}$  lässt sich darstellen als

$$\begin{aligned} E[(\vec{z} - E[\vec{z}])(\vec{z} - E[\vec{z}])^T] &= E\left[\vec{z}\vec{z}^T - \vec{z}\hat{\vec{z}}^T - \hat{\vec{z}}\vec{z}^T + \hat{\vec{z}}\hat{\vec{z}}^T\right] \\ &= E[\vec{z}\vec{z}^T] - E[\vec{z}]\hat{\vec{z}}^T - \hat{\vec{z}}E[\vec{z}^T] + \hat{\vec{z}}\hat{\vec{z}}^T \\ &= E[\vec{z}\vec{z}^T] - \hat{\vec{z}}\hat{\vec{z}}^T \\ &= \int_{-\infty}^{\infty} \vec{z}\vec{z}^T \sum_{i=1}^m \alpha_i p_i(\vec{z}) d\vec{z} - \hat{\vec{z}}\hat{\vec{z}}^T \\ &= \sum_{i=1}^m \alpha_i \int_{-\infty}^{\infty} \vec{z}\vec{z}^T p_i(\vec{z}) d\vec{z} - \hat{\vec{z}}\hat{\vec{z}}^T, \end{aligned} \quad (6.179)$$

mit

$$\mathbf{P}_i = E[(\vec{z} - \hat{\vec{z}}_i)(\vec{z} - \hat{\vec{z}}_i)^T] = E[\vec{z}\vec{z}^T] - \hat{\vec{z}}_i \hat{\vec{z}}_i^T \quad (6.180)$$

folgt

$$E[\vec{z}\vec{z}^T] = \mathbf{P}_i + \hat{\vec{z}}_i \hat{\vec{z}}_i^T \quad (6.181)$$

und damit

$$\begin{aligned} E[(\vec{z} - E[\vec{z}])(\vec{z} - E[\vec{z}])^T] &= \sum_{i=1}^m \alpha_i \left( \mathbf{P}_i + \hat{\vec{z}}_i \hat{\vec{z}}_i^T \right) - \hat{\vec{z}} \hat{\vec{z}}^T \\ &= \sum_{i=1}^m \alpha_i \mathbf{P}_i + \left( \sum_{i=1}^m \alpha_i \hat{\vec{z}}_i \hat{\vec{z}}_i^T \right) - \hat{\vec{z}} \hat{\vec{z}}^T. \end{aligned} \quad (6.182)$$

Gl. (6.182) kann weiter vereinfacht werden. Mit

$$\begin{aligned} \left( \sum_{i=1}^m \alpha_i \hat{\vec{z}}_i \hat{\vec{z}}_i^T \right) - \hat{\vec{z}} \hat{\vec{z}}^T &= \left( \sum_{i=1}^m \alpha_i \hat{\vec{z}}_i \hat{\vec{z}}_i^T \right) - \hat{\vec{z}} \hat{\vec{z}}^T - \hat{\vec{z}} \hat{\vec{z}}^T + \hat{\vec{z}} \hat{\vec{z}}^T \\ &= \left( \sum_{i=1}^m \alpha_i \hat{\vec{z}}_i \hat{\vec{z}}_i^T \right) - \left( \sum_{i=1}^m \alpha_i \hat{\vec{z}}_i \right) \hat{\vec{z}}^T \\ &\quad - \hat{\vec{z}} \left( \sum_{i=1}^m \alpha_i \hat{\vec{z}}_i \right)^T + \hat{\vec{z}} \hat{\vec{z}}^T \sum_{i=1}^m \alpha_i \\ &= \sum_{i=1}^m \alpha_i \left( \hat{\vec{z}}_i \hat{\vec{z}}_i^T - \hat{\vec{z}}_i \hat{\vec{z}}^T - \hat{\vec{z}} \hat{\vec{z}}_i^T + \hat{\vec{z}} \hat{\vec{z}}^T \right) \\ &= \sum_{i=1}^m \alpha_i (\hat{\vec{z}}_i - \hat{\vec{z}})(\hat{\vec{z}}_i - \hat{\vec{z}})^T \end{aligned} \quad (6.183)$$

erhält man schließlich

$$\begin{aligned} E[(\vec{z} - E[\vec{z}])(\vec{z} - E[\vec{z}])^T] &= \sum_{i=1}^m \alpha_i \mathbf{P}_i + \sum_{i=1}^m \alpha_i (\hat{\vec{z}}_i - \hat{\vec{z}})(\hat{\vec{z}}_i - \hat{\vec{z}})^T \\ &= \sum_{i=1}^m \alpha_i \left( \mathbf{P}_i + (\hat{\vec{z}}_i - \hat{\vec{z}})(\hat{\vec{z}}_i - \hat{\vec{z}})^T \right). \end{aligned} \quad (6.184)$$

Anhand dieser allgemeingültigen Formeln für Erwartungswert und Kovarianz einer Summe von Gaußverteilungen können die elementaren Filter wie folgt initialisiert werden:

$$\vec{x}_{i,k-1}^{0,+} = \sum_{j=1}^m \mu_{k-1}^{j,k-1|i_k} \vec{x}_{j,k-1}^+ \quad (6.185)$$

$$\mathbf{P}_{i,k-1}^{0,+} = \sum_{j=1}^m \mu_{k-1}^{j,k-1|i_k} \left( \mathbf{P}_{j,k-1}^+ + (\hat{\vec{x}}_{j,k-1}^+ - \vec{x}_{i,k-1}^{0,+})(\hat{\vec{x}}_{j,k-1}^+ - \vec{x}_{i,k-1}^{0,+})^T \right) \quad (6.186)$$

### Propagationsschritt

Im Propagationsschritt werden die Wahrscheinlichkeitsdichten der einzelnen Modelle in der Zeit propagiert, es wird also der Übergang

$$p(\vec{x}_{k-1}|s_k = i, \mathbf{Y}_{k-1}) \rightarrow p(\vec{x}_k|s_k = i, \mathbf{Y}_{k-1}), \quad \forall i \quad (6.187)$$

durchgeführt. Realisiert wird dieser Übergang, indem jeder elementare Filter einen gewöhnlichen Propagationsschritt durchführt, dies stellt sich wie folgt dar:

$$\vec{x}_{i,k-1}^{0,+} \rightarrow \vec{x}_{i,k}^- \quad (6.188)$$

$$\mathbf{P}_{i,k-1}^{0,+} \rightarrow \mathbf{P}_{i,k}^- \quad (6.189)$$

Die Modellwahrscheinlichkeiten sind davon nicht betroffen.

### Messschritt

Im Messschritt werden die vorhandenen Messwerte von jedem elementaren Filter verarbeitet. Das Ergebnis davon ist der Übergang

$$p(\vec{x}_k|s_k = i, \mathbf{Y}_{k-1}) \rightarrow p(\vec{x}_k|s_k = i, \mathbf{Y}_k), \quad \forall i \quad (6.190)$$

bzw.

$$\vec{x}_{i,k}^- \rightarrow \vec{x}_{i,k}^+ \quad (6.191)$$

$$\mathbf{P}_{i,k}^- \rightarrow \mathbf{P}_{i,k}^+ . \quad (6.192)$$

Zusätzlich müssen die Modellwahrscheinlichkeiten angepasst werden:

$$\mu_{k-1}^{i_k} \rightarrow \mu_k^{i_k}, \quad \forall i \quad (6.193)$$

Mit Hilfe der Bayes'schen Regel erhält man

$$P(s_k = i|\tilde{y}_k, \mathbf{Y}_{k-1}, \vec{x}_{i,k}^-) = \frac{p(\tilde{y}_k|s_k = i, \mathbf{Y}_{k-1}, \vec{x}_{i,k}^-)P(s_k = i|\mathbf{Y}_{k-1}, \vec{x}_{i,k}^-)}{p(\tilde{y}_k|\mathbf{Y}_{k-1}, \vec{x}_{i,k}^-)} . \quad (6.194)$$

Da alle Messwerte bis zum Zeitpunkt  $k-1$  bereits in  $\vec{x}_{i,k}^-$  berücksichtigt sind, kann man schreiben:

$$p(\tilde{y}_k|s_k = i, \mathbf{Y}_{k-1}, \vec{x}_{i,k}^-) = p(\tilde{y}_k|\vec{x}_{i,k}^-) \quad (6.195)$$

Damit ergibt sich aus Gl. (6.194)

$$\mu_k^{i_k} = \frac{p(\tilde{y}_k|\vec{x}_{i,k}^-)\mu_{k-1}^{i_k}}{\sum_i p(\tilde{y}_k|\vec{x}_{i,k}^-)\mu_{k-1}^{i_k}} . \quad (6.196)$$

Der Nenner in Gl. (6.196) dient dabei zur Normierung der Modellwahrscheinlichkeiten, so dass deren Summe eins ergibt. Um Gl. (6.196) auswerten zu können, muss  $p(\tilde{y}_k | \vec{x}_{i,k}^-)$  ermittelt werden, die Wahrscheinlichkeitsdichte von  $\tilde{y}_k$  unter Voraussetzung des Zustands  $\vec{x}_{i,k}^-$  des elementaren Filters  $i$ . Normalerweise spricht man bei  $p(\tilde{y}_k | \vec{x}_{i,k}^-)$  nicht von einer Wahrscheinlichkeitsdichte, sondern von einer Likelihood-Funktion. Diese hat die gleiche Form wie eine Wahrscheinlichkeitsdichte und beschreibt, wie 'gut' eine tatsächlich vorliegende Realisation, hier  $\tilde{y}_k$ , zu einer Wahrscheinlichkeitsdichte, hier die von  $\vec{x}_{i,k}^-$ , 'passt'. Damit erschließt sich eine intuitive Interpretation von Gl. (6.196): Passt der vorliegende Messwert gut zur WDF von Filter  $i$ , so wird die zugehörige Modellwahrscheinlichkeit vergrößert. War das Auftreten des Messwertes eher unwahrscheinlich wenn man die WDF von Filter  $i$  zugrunde legt, wird die zugehörige Modellwahrscheinlichkeit entsprechend verringert. Der Filter, der auf dem zum tatsächlichen Zustand der Markov-Kette gehörenden Modell basiert, wird besser in der Lage sein, Messwerte vorherzusagen als ein Filter, dessen System- und/oder Messmodell sich vom tatsächlichen System-/Messmodell unterscheidet.

Die Likelihood-Funktion  $p(\tilde{y}_k | \vec{x}_{i,k}^-)$  ist gegeben durch

$$p(\tilde{y}_k | \vec{x}_{i,k}^-) = \frac{1}{\sqrt{(2\pi)^l |\mathbf{S}_{i,k}|}} e^{-\frac{1}{2} \vec{r}_{i,k}^T \mathbf{S}_{i,k}^{-1} \vec{r}_{i,k}} \quad (6.197)$$

mit

$$\mathbf{S}_{i,k} = \text{cov}(\tilde{y}_k | \vec{x}_{i,k}^-) = \mathbf{H}_{i,k} \mathbf{P}_{i,k}^- \mathbf{H}_{i,k}^T + \mathbf{R}_{i,k} \quad (6.198)$$

$$\mathbf{R}_{i,k} = \text{cov}(\mathbf{C}_k(s_k = i) \vec{v}_k) \quad (6.199)$$

$$\vec{r}_{i,k} = \tilde{y}_k - E[\tilde{y}_k | \vec{x}_{i,k}^-] = \tilde{y}_k - \mathbf{H}_{i,k} \vec{x}_{i,k}^- , \quad (6.200)$$

die Anzahl der Komponenten von  $\tilde{y}_k$  ist mit  $l$  bezeichnet.

Die Zustandsschätzung der Filterbank ist durch Erwartungswert und Kovarianz der mit den Modellwahrscheinlichkeiten gewichteten Summe der Wahrscheinlichkeitsdichten der elementaren Filter gegeben:

$$\hat{x}_k^+ = \sum_{i=1}^m \mu_k^{i_k} \vec{x}_{i,k}^+ \quad (6.201)$$

$$\hat{\mathbf{P}}_k^+ = \sum_{i=1}^m \mu_k^{i_k} \left( \mathbf{P}_{i,k}^+ + (\vec{x}_{i,k}^+ - \hat{x}_k^+) (\vec{x}_{i,k}^+ - \hat{x}_k^+)^T \right) \quad (6.202)$$

Damit ist ein Zyklus des IMM abgeschlossen, die nach diesem Messschritt zur Verfügung stehenden Größen dienen als Eingangsgrößen für das nächste Mischen.

Beim IMM wird davon ausgegangen, dass sich die Zustände der Markov-Kette mit der Zeit ändern können. Geht man von einem unbekannten, aber konstanten Zustand aus, so reduziert sich der IMM zum Multiple Model Adaptive Estimator (MMAE). Der MMAE identifiziert das korrekte Modell, dessen Wahrscheinlichkeit – ein entsprechendes

Szenario mit aussagekräftigen Messwerten vorausgesetzt – sich dann ziemlich schnell gegen eins bewegt. Dieses Verhalten kann bei dem hier vorgestellten IMM durch eine entsprechende Wahl der Übergangswahrscheinlichkeitsmatrix  $M$  erzwungen werden.

Mit wachsender Anzahl von Modellen sinkt im Allgemeinen die Effektivität eines IMM: Es wird zunehmend Rechenleistung für Modelle mit geringer Modellwahrscheinlichkeit verschwendet, die wenig zum Gesamtergebnis beitragen – der dennoch vorhandene Beitrag dieser Modelle bedeutet zudem meist eine Verschlechterung. Dem kann begegnet werden, indem online entschieden wird, welche und wieviele Modelle verwendet werden. Diese als Variable-Structure IMM (VSIMM) bekannten Algorithmen sollen hier jedoch nicht weiter betrachtet werden.

### IMM-Algorithmus

Zur besseren Übersichtlichkeit werden hier die Gleichungen des IMM nochmals zusammengestellt:

#### Mischen

$$\begin{aligned}\mu_{k-1}^{j_{k-1}|i_k} &= \frac{M_{ij} \mu_{k-1}^{j_{k-1}}}{\sum_{j=1}^m M_{ij} \mu_{k-1}^{j_{k-1}}} \\ \vec{x}_{i,k-1}^{0,+} &= \sum_{j=1}^m \mu_{k-1}^{j_{k-1}|i_k} \vec{x}_{j,k-1}^+ \\ \mathbf{P}_{i,k-1}^{0,+} &= \sum_{j=1}^m \mu_{k-1}^{j_{k-1}|i_k} \left( \mathbf{P}_{j,k-1}^+ + (\hat{x}_{j,k-1}^+ - \vec{x}_{i,k-1}^{0,+})(\hat{x}_{j,k-1}^+ - \vec{x}_{i,k-1}^{0,+})^T \right)\end{aligned}$$

### Propagationsschritt

Durchführen gewöhnlicher Propagationsschritte für jeden elementaren Filter  $i$ , nachdem diese mit  $\vec{x}_{i,k-1}^{0,+}$ ,  $\mathbf{P}_{i,k-1}^{0,+}$  initialisiert wurden:

$$\begin{aligned}\vec{x}_{i,k-1}^{0,+} &\rightarrow \vec{x}_{i,k}^- \\ \mathbf{P}_{i,k-1}^{0,+} &\rightarrow \mathbf{P}_{i,k}^-\end{aligned}$$

### Messschritt

Durchführen gewöhnlicher Messschritte für jeden elementaren Filter  $i$ :

$$\begin{aligned}\vec{x}_{i,k}^- &\rightarrow \vec{x}_{i,k}^+ \\ \mathbf{P}_{i,k}^- &\rightarrow \mathbf{P}_{i,k}^+ .\end{aligned}$$

Anpassung der Modellwahrscheinlichkeiten:

$$\begin{aligned}\mu_k^{i_k} &= \frac{p(\tilde{y}_k | \vec{x}_{i,k}^-) \mu_{k-1}^{i_k}}{\sum_i p(\tilde{y}_k | \vec{x}_{i,k}^-) \mu_{k-1}^{i_k}} \\ p(\tilde{y}_k | \vec{x}_{i,k}^-) &= \frac{1}{\sqrt{(2\pi)^l |\mathbf{S}_{i,k}|}} e^{-\frac{1}{2} \vec{r}_{i,k}^T \mathbf{S}_{i,k}^{-1} \vec{r}_{i,k}} \\ \mathbf{S}_{i,k} &= \mathbf{H}_{i,k} \mathbf{P}_{i,k}^- \mathbf{H}_{i,k}^T + \mathbf{R}_{i,k} \\ \vec{r}_{i,k} &= \tilde{y}_k - \mathbf{H}_{i,k} \vec{x}_{i,k}^-\end{aligned}$$

Berechnung der Ausgangsgrößen:

$$\begin{aligned}\hat{\vec{x}}_k^+ &= \sum_{i=1}^m \mu_k^{i_k} \vec{x}_{i,k}^+ \\ \hat{\mathbf{P}}_k^+ &= \sum_{i=1}^m \mu_k^{i_k} \left( \mathbf{P}_{i,k}^+ + (\vec{x}_{i,k}^+ - \hat{\vec{x}}_k^+) (\vec{x}_{i,k}^+ - \hat{\vec{x}}_k^+)^T \right)\end{aligned}$$