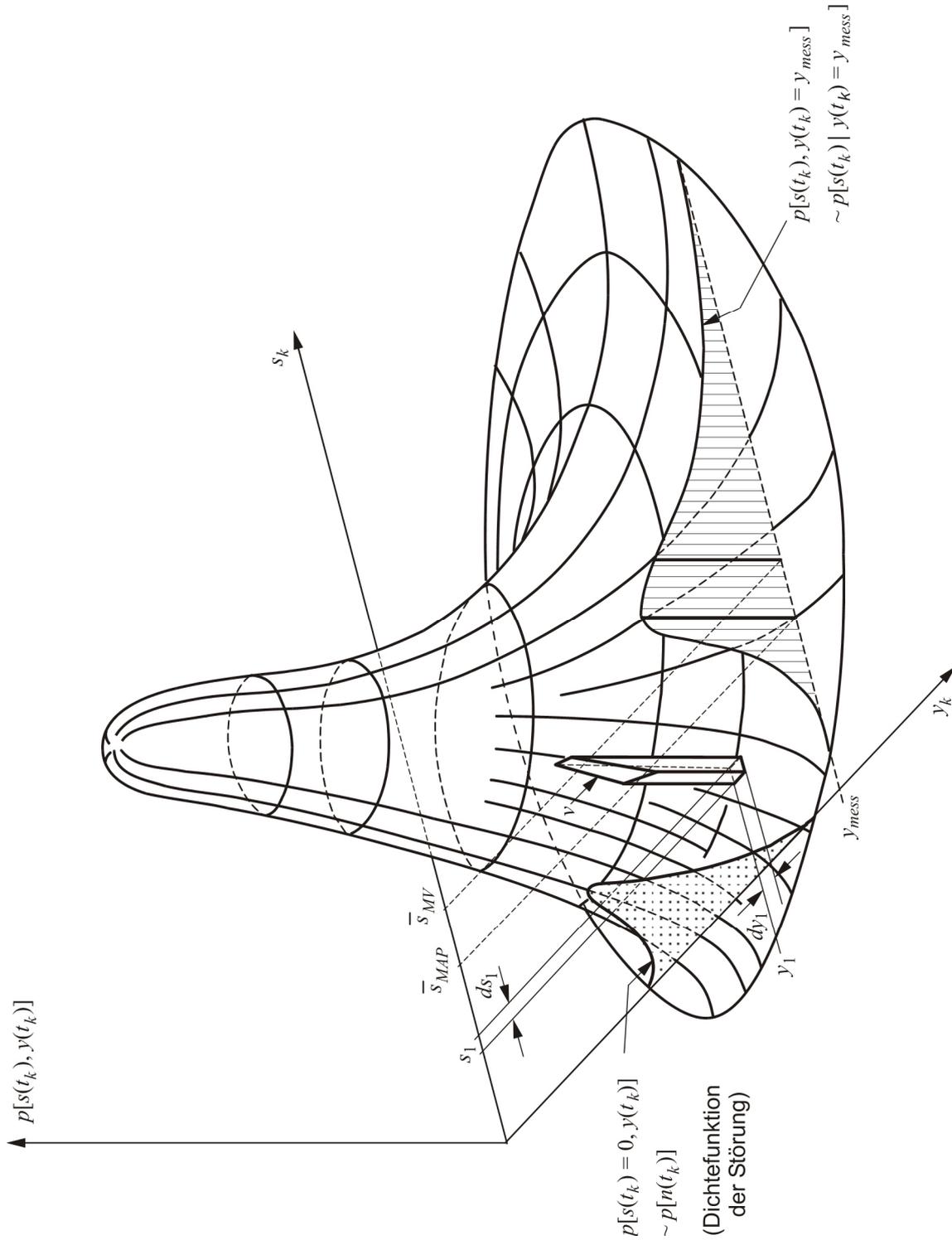
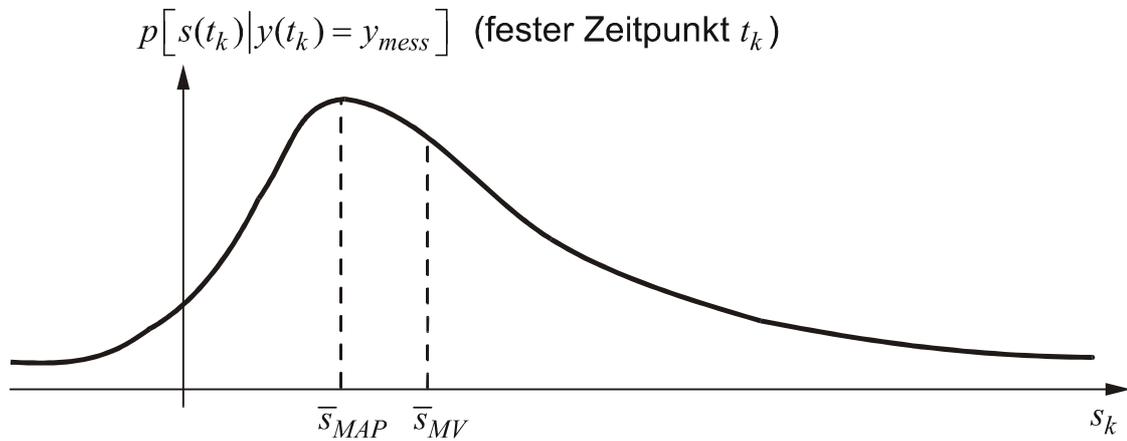


Verbunddichtefunktion und bedingte Verbunddichtefunktion

Momentaufnahme bei festem Zeitpunkt t_k



Schätzwerte



\bar{s}_{MAP} : Maximum-a-posteriori-Schätzwert
(bei Maximum der bedingten Verbunddichtefunktion)

$\bar{s}_{MV} = \hat{s}$: Minimal-Varianz-Schätzwert
(bei Flächenschwerpunkt der bedingten Verbunddichtefunktion)

Gauß-Dichtefunktionen zeichnen sich durch ihre typische Gauß-Glocken-Gestalt aus. Beispielsweise ergibt sich für den Fall $n = 2$ für die beiden zeitdiskreten Zufallsvariablen zum Zeitpunkt k folgendes Bild 1.

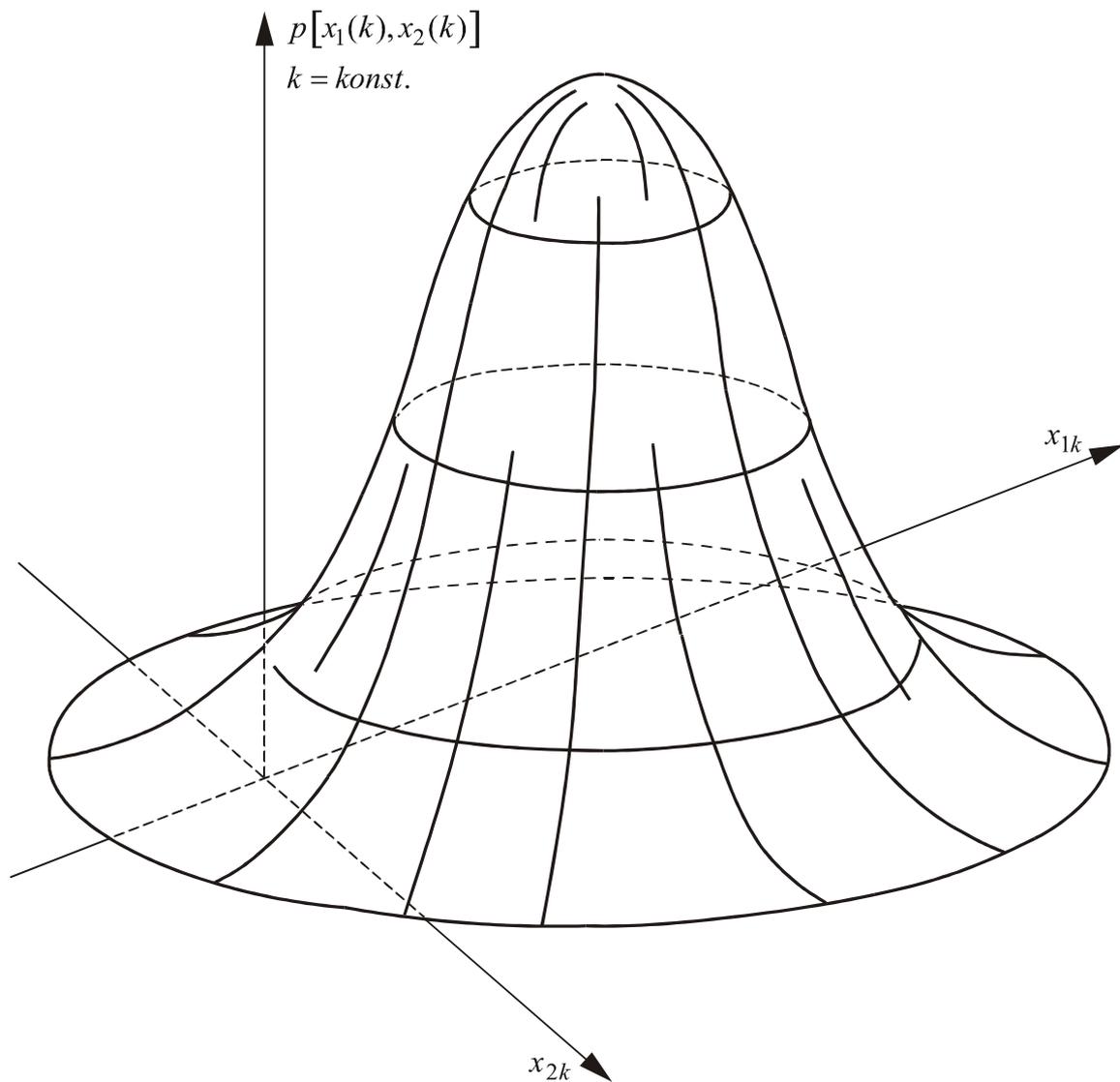


Bild 1

Beim Systemdurchgang bleibt die Gaußglockenförmige Verteilung der Zufallsvariablen erhalten, Mittelwerte und Varianzen entwickeln sich jedoch dynamisch. In Bild 2 ist dies exemplarisch für den Fall einer einzelnen Zufallsvariablen angedeutet ($n = 1$).

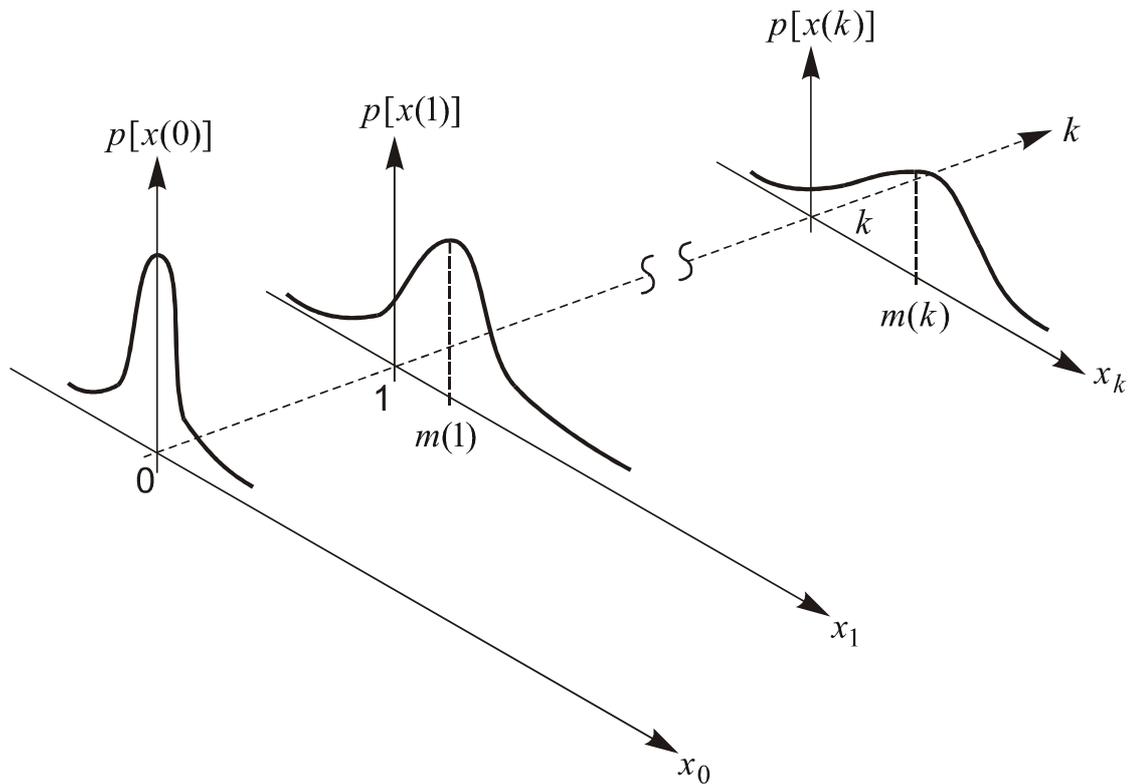


Bild 2

Aus der Least-Squares (LS)-Schätzung stochastisch gestörter Parameter sind folgende rekursive Schätzgleichungen bekannt (vgl. z.B. Vorlesung „Modellbildung und Identifikation“):

Als Messgleichung liegt das *allgemeine lineare Regressionsmodell* vor:

$$\underline{y}(k) = \underline{U}(k)\underline{\theta}(k) + \underline{\varepsilon}(k).$$

Hierbei sind:

- \underline{y} : Messgrößenvektor (Dimension $(N, 1)$)
- \underline{U} : Eingangsgrößenmatrix (Dimension (N, n))
- $\underline{\theta}$: Parametervektor (Dimension $(n, 1)$)
- $\underline{\varepsilon}$: vektorielles Messrauschen (Dimension $(N, 1)$)

Mit den Voraussetzungen

- α) als Messrauschen liegt mittelwertfreies ($E\{\underline{\varepsilon}\} = \underline{0}$) Gauß'sches Weißes Rauschen mit bekannter Kovarianzmatrix \underline{R} vor
- β) Eingangsdaten und Messrauschen werden als unkorreliert angenommen ($E\{\underline{U} \cdot \underline{\varepsilon}\} = E\{\underline{U}\} \cdot \underbrace{E\{\underline{\varepsilon}\}}_{=0} = \underline{0}$)

ergibt sich aus der Forderung nach dem kleinsten quadratischen Fehler der *Schätzabweichung*

$$\underline{\tilde{\theta}} := \underline{\theta} - \underline{\hat{\theta}}$$

folgender **rekursiver LS-Algorithmus** für den Schätzwert $\underline{\hat{\theta}}(k)$ des Parametervektors:

$$\begin{aligned}\underline{\hat{\theta}}(k) &= \underline{\hat{\theta}}(k-1) + \underline{K}(k) \left\{ \underline{y}(k) - \underline{U}(k)\underline{\hat{\theta}}(k-1) \right\} \\ \underline{K}(k) &= \underline{P}(k-1)\underline{U}^T(k) \left(\underline{U}(k)\underline{P}(k-1)\underline{U}^T(k) + \underline{R}(k) \right)^{-1} \\ &= \underline{P}(k)\underline{U}^T(k)\underline{R}^{-1}(k) \\ \underline{P}(k) &= \underline{P}(k-1) - \underline{K}(k)\underline{U}(k)\underline{P}(k-1)\end{aligned}$$

Folgende stochastische Eigenschaften des geschätzten Parameters sind hervorzugeben (ohne Herleitung):

- a) Es gilt $E\{\underline{\tilde{\theta}}\} = \underline{0}$ bzw. $E\{\underline{\hat{\theta}}\} = \underline{\theta}$. Man nennt die Schätzung daher auch **erwartungstreu** (unbiased).
- b) Wegen $\underline{P}(k) = E\{\underline{\tilde{\theta}}\underline{\tilde{\theta}}^T\}$ handelt es sich bei $\underline{P}(k)$ um die **Kovarianzmatrix des Schätzwerts**.

Kalman-Filter-Gleichungen:

- **Filtergleichungen:**

$$\hat{\underline{x}}(k|k) = \hat{\underline{x}}(k|k-1) + \underline{K}(k) \{ \underline{y}(k) - \underline{C}(k) \hat{\underline{x}}(k|k-1) \}$$

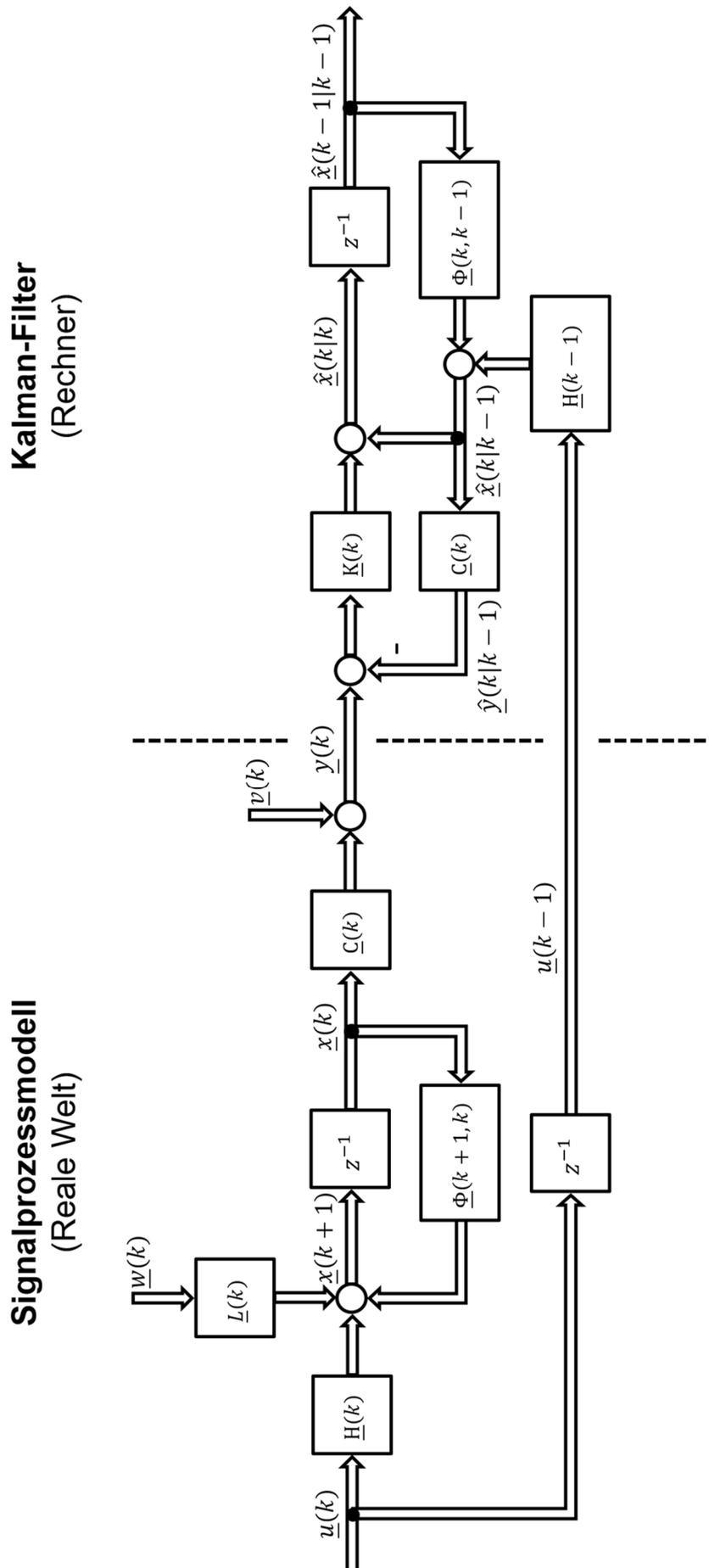
$$\underline{P}(k|k) = \underline{P}(k|k-1) - \underline{K}(k) \underline{C}(k) \underline{P}(k|k-1)$$

$$\begin{aligned} \underline{K}(k) &= \underline{P}(k|k-1) \underline{C}^T(k) [\underline{C}(k) \underline{P}(k|k-1) \underline{C}^T(k) + \underline{R}(k)]^{-1} \\ &= \underline{P}(k|k) \underline{C}^T(k) \underline{R}^{-1}(k), \quad k = 0, 1, \dots \end{aligned}$$

- **Prädiktionsgleichungen:**

$$\hat{\underline{x}}(k|k-1) = \underline{\Phi}(k|k-1) \hat{\underline{x}}(k-1|k-1) + \underline{H}(k-1) \underline{u}(k-1)$$

$$\begin{aligned} \underline{P}(k|k-1) &= \underline{\Phi}(k|k-1) \underline{P}(k-1|k-1) \underline{\Phi}^T(k|k-1) \\ &\quad + \underline{L}(k-1) \underline{Q}(k-1) \underline{L}^T(k-1), \quad k = 0, 1, \dots \end{aligned}$$



Betrachtet sei das folgende zeitinvariante Signalprozessmodell:

$$\underline{x}(k+1) = \begin{bmatrix} 1,1269 & -0,494 & 0,1129 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \underline{x}(k) + \begin{bmatrix} -0,3832 \\ 0,5919 \\ 0,5191 \end{bmatrix} u(k) + \begin{bmatrix} -0,3832 \\ 0,5919 \\ 0,5191 \end{bmatrix} w(k),$$

$$y(k) = [1 \ 0 \ 0] \underline{x}(k) + v(k).$$

Hierbei sind $w(k)$ und $v(k)$ stationäre skalare Gauß'sche Rauschprozesse mit den Varianzen $\underline{Q} = q$ bzw. $\underline{R} = r$, welche System- bzw. Messrauschen darstellen.

Zur Analyse der Arbeitsweise des Kalman-Filters werden nun drei Fälle betrachtet:

- a) Fall $q = r = 1$
- b) Fall $q = 10, r = 1$
- c) Fall $q = 1, r = 10$

Als deterministische Eingangsgröße aufgeschaltet werde in allen Fällen

$$u(t) = \sin\left(\frac{1}{10}t\right)\sigma(t),$$

als Anfangswerte der Schätzung liegen $\hat{\underline{x}}(0|0) = \underline{0}$ und $\underline{P}(0|0) = 10 \cdot \underline{I}$ vor.

- a) $q = r = 1$

In diesem Beispiel haben System- und Messrauschen die gleiche Leistung. Bild 1 zeigt die ungestörte Messgröße (grün) sowie die Messgröße unter Berücksichtigung des Systemrauschens (blau) bzw. des Messrauschens (rot).

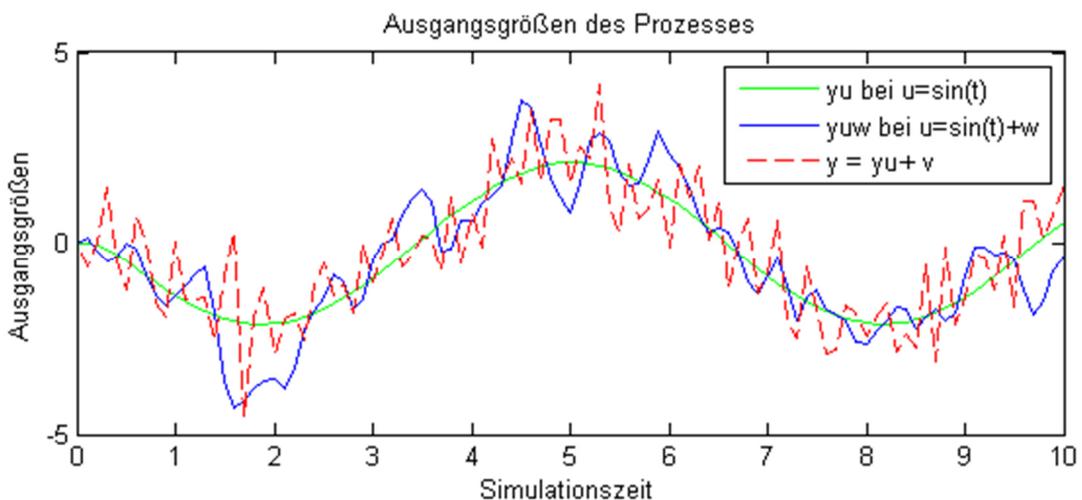


Bild 1

Durch Anwendung des Kalman-Filters erhält man aus den vorhandenen Informationen folgendes Schätzergebnis:

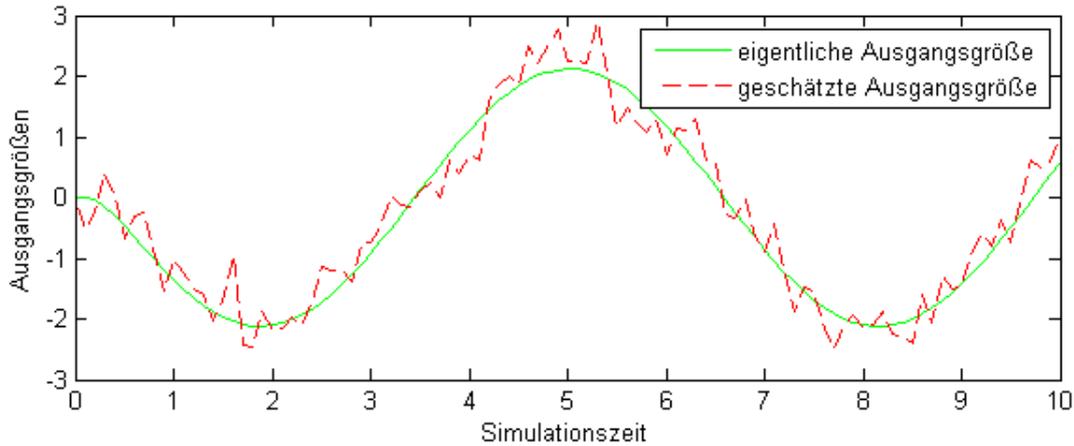


Bild 2

Aus Bild 2 wird ersichtlich, dass das ursprüngliche Ausgangssignal adäquat geschätzt werden kann. Die Varianz (und damit die Leistung) des Fehlersignals wird unter den getroffenen Annahmen minimiert.

Die folgenden Schaubilder geben abschließend die Verläufe der Varianzen der geschätzten Ausgangswerte und die Filterverstärkung $\underline{K}(k) = \underline{k}(k)$ an.

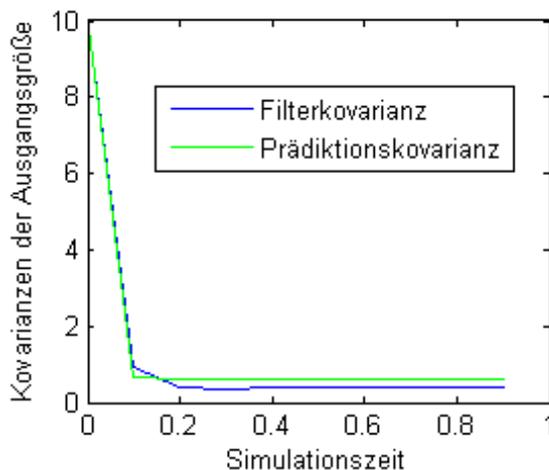


Bild 3

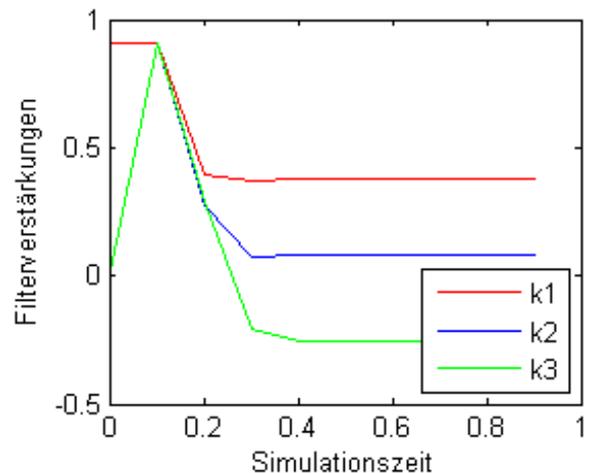


Bild 4

Aufgrund der Stationarität der Rauschprozesse und der Zeitinvarianz des Signalprozessmodells konvergieren die Verläufe sehr schnell gegen stationäre Endwerte. Es ist anzumerken, dass die Konvergenz weitgehend unabhängig vom Startwert der Schätzwertkovarianz verläuft.

b) $q = 10, r = 1$

Bild 5 stellt die Ergebnisse der Schätzung dar. Das Kalman-Filter verwendet in diesem Fall hauptsächlich die Messinformation zur Berechnung des Schätzwerts.

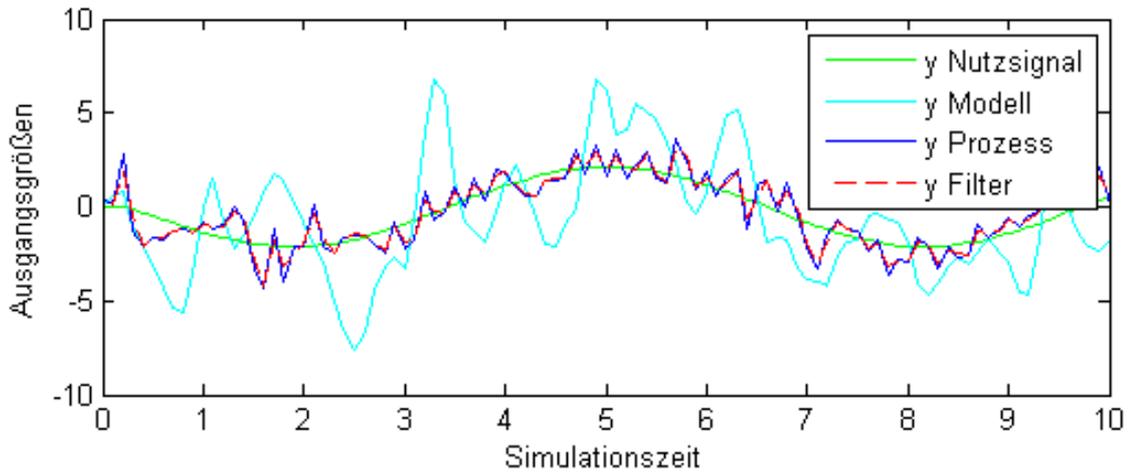


Bild 5

Aufgrund der großen Varianz des Systemrauschens ist die Schätzwert-Varianz im Prädiktionsschritt sehr groß und wird durch den Filterschritt reduziert, was in Bild 6 dargestellt ist. Die Beträge der Einträge des Kalman-Gains $\underline{k}(k)$ nehmen im Vergleich zu Fall a) kleinere Werte an (Bild 7).

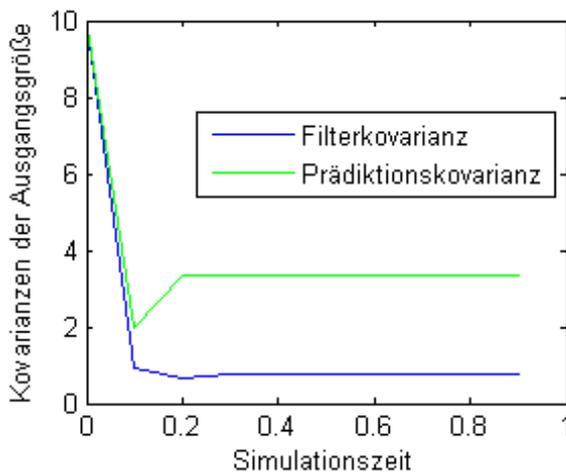


Bild 6

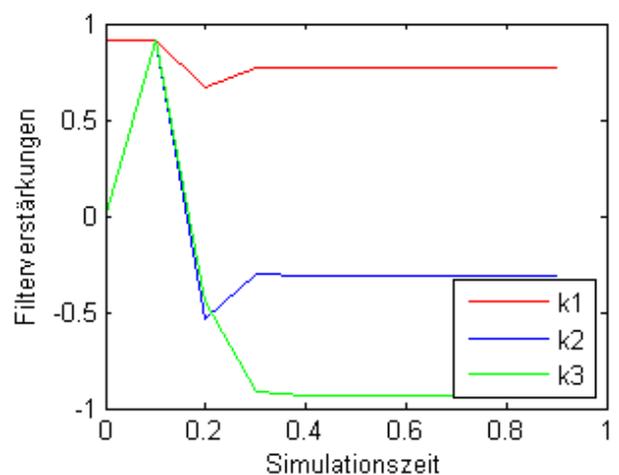


Bild 7

c) $q = 1, r = 10$

Im Gegensatz zu Fall b) basieren der Schätzwert und der daraus resultierende geschätzte Wert der Ausgangsgröße sehr viel stärker auf der Modellinformation, was in Bild 8 dargestellt ist.

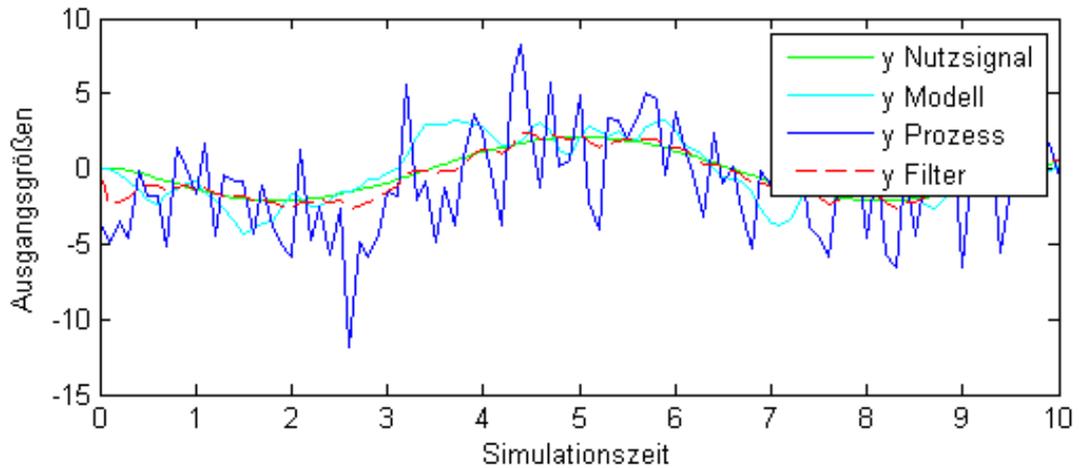


Bild 8

Aufgrund der unpräzisen Messinformation wird die Prädiktion nur relativ wenig korrigiert. Entsprechend sind die Varianzen in Prädiktions- und Filterschritt im eingeschwungenen Zustand annähernd gleich (Bild 9). Die Filterverstärkung nimmt betragskleine Werte an (Bild 10).

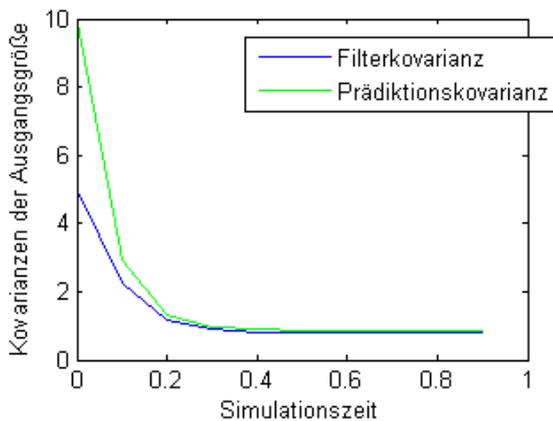


Bild 9

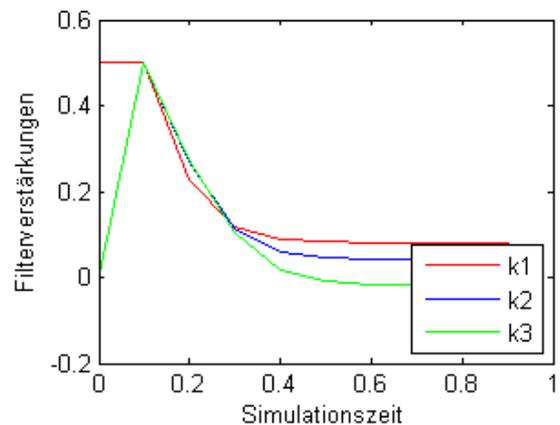


Bild 10

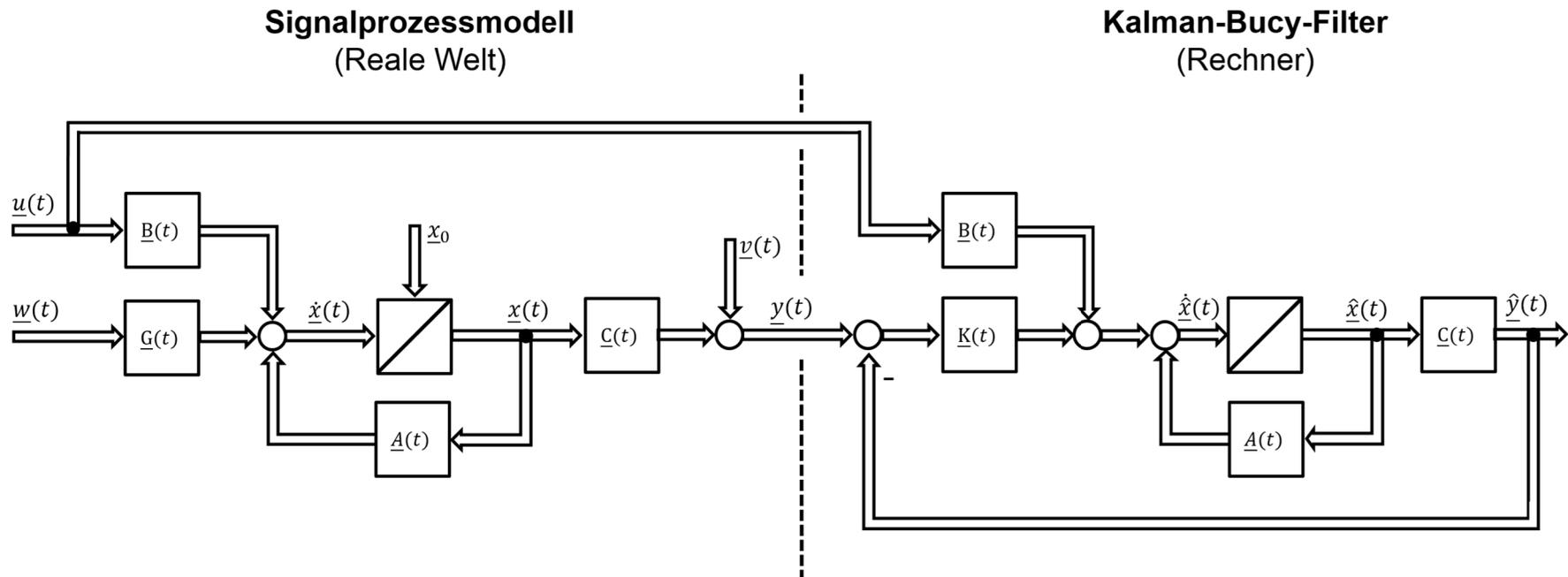
- **Gleichungen des Kalman-Bucy-Filters**

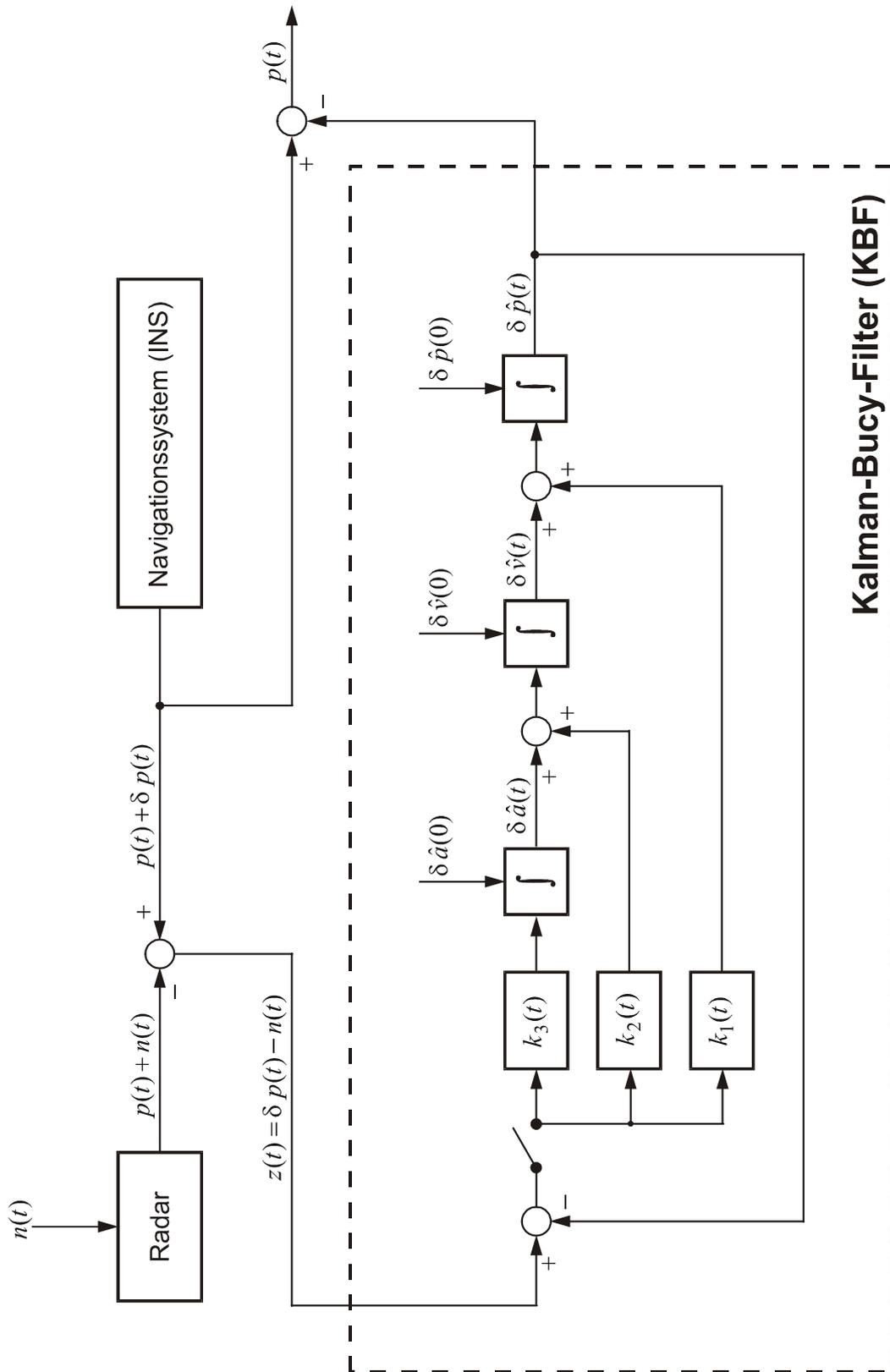
$$\dot{\hat{x}}(t) = \underline{A}(t)\hat{x}(t) + \underline{B}(t)u(t) + \underline{K}(t)\{y(t) - \underline{C}(t)\hat{x}(t)\}$$

$$\dot{\underline{P}}(t) = \underline{A}(t)\underline{P}(t) + \underline{P}(t)\underline{A}^T(t) + \underline{G}(t)\underline{Q}(t)\underline{G}^T(t) - \underline{P}(t)\underline{C}^T(t)\underline{R}^{-1}(t)\underline{C}(t)\underline{P}(t)$$

$$\underline{K}(t) = \underline{P}(t)\underline{C}^T(t)\underline{R}^{-1}(t)$$

- **Struktur des Kalman-Bucy-Filters**





Linearisiertes Kalman-Bucy-Filter:

$$(\delta \hat{\underline{x}}(t))' = \underline{A}_0(t) \delta \hat{\underline{x}}(t) + \underline{K}(t) \{ \delta \underline{y}(t) - \underline{C}_0(t) \delta \hat{\underline{x}}(t) \}$$

Anfangswerte:
 $\delta \hat{\underline{x}}(t_0) = \underline{0}, \underline{P}(t_0)$

$$\underline{K}(t) = \underline{P}(t) \underline{C}_0^T(t) \underline{R}_0^{-1}(t)$$

$$\dot{\underline{P}}(t) = \underline{A}_0(t) \underline{P}(t) + \underline{P}(t) \underline{A}_0^T(t) - \underline{P}(t) \underline{C}_0^T(t) \underline{R}_0^{-1}(t) \underline{C}_0(t) \underline{P}(t) + \underline{G}_0(t) \underline{Q}(t) \underline{G}_0^T(t)$$

$$\hat{\underline{x}}(t) = \underline{x}_0(t) + \delta \hat{\underline{x}}(t)$$

mit:

$$\underline{A}_0(t) = \left. \frac{\partial \underline{f}(\underline{x}, \underline{u}, \underline{w}, t)}{\partial \underline{x}(t)} \right|_0; \quad \underline{C}_0(t) = \left. \frac{\partial \underline{g}(\underline{x}, \underline{v}, t)}{\partial \underline{x}(t)} \right|_0; \quad \underline{G}_0(t) = \left. \frac{\partial \underline{f}(\underline{x}, \underline{u}, \underline{w}, t)}{\partial \underline{w}} \right|_0;$$

$$\underline{V}_0(t) = \left. \frac{\partial \underline{g}(\underline{x}, \underline{v}, t)}{\partial \underline{v}} \right|_0; \quad \underline{R}_0(t) = \underline{V}_0(t) \underline{R}(t) \underline{V}_0^T(t)$$

Erweitertes Kalman-Bucy-Filter:

$$\hat{\underline{x}}(t) = \underline{f}(\hat{\underline{x}}, \underline{u}, \underline{0}, t) + \underline{K}(t) \{ \underline{y}(t) - \underline{g}(\hat{\underline{x}}, \underline{0}, t) \}$$

Anfangswerte:
 $\hat{\underline{x}}(t_0) = \underline{\hat{x}}_0, \underline{P}(t_0)$

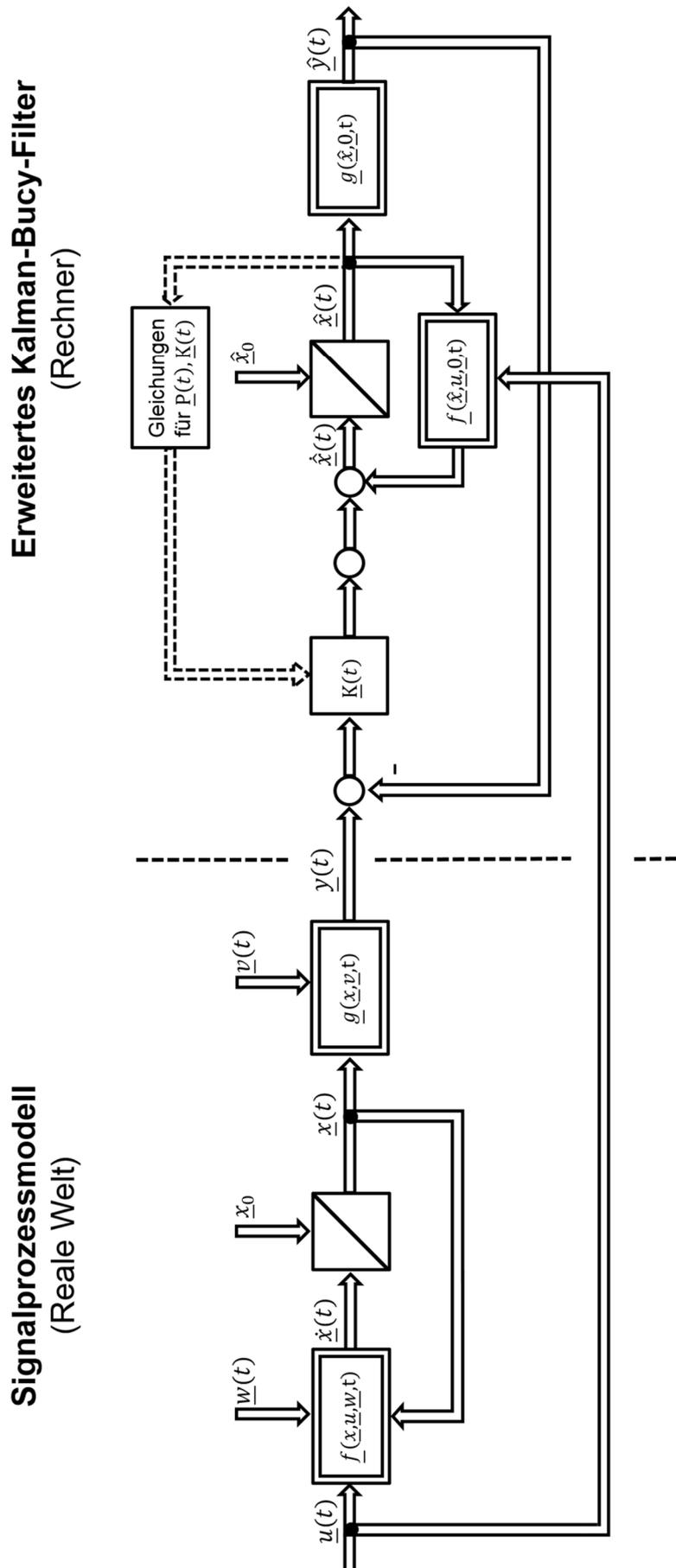
$$\underline{K}(t) = \underline{P}(t) \hat{\underline{C}}^T(t) \hat{\underline{R}}^{-1}(t)$$

$$\dot{\underline{P}}(t) = \hat{\underline{A}}(t) \underline{P}(t) + \underline{P}(t) \hat{\underline{A}}^T(t) - \underline{P}(t) \hat{\underline{C}}^T(t) \hat{\underline{R}}^{-1}(t) \hat{\underline{C}}(t) \underline{P}(t) + \hat{\underline{G}}(t) \underline{Q}(t) \hat{\underline{G}}^T(t)$$

mit:

$$\hat{\underline{A}}(t) = \left. \frac{\partial \underline{f}(\underline{x}, \underline{u}, \underline{w}, t)}{\partial \underline{x}(t)} \right|_{\wedge}; \quad \hat{\underline{C}}(t) = \left. \frac{\partial \underline{g}(\underline{x}, \underline{v}, t)}{\partial \underline{x}(t)} \right|_{\wedge}; \quad \hat{\underline{G}}(t) = \left. \frac{\partial \underline{f}(\underline{x}, \underline{u}, \underline{w}, t)}{\partial \underline{w}} \right|_{\wedge};$$

$$\hat{\underline{V}}(t) = \left. \frac{\partial \underline{g}(\underline{x}, \underline{v}, t)}{\partial \underline{v}} \right|_{\wedge}; \quad \hat{\underline{R}}(t) = \hat{\underline{V}}(t) \underline{R}(t) \hat{\underline{V}}^T(t)$$



Das Sigma-Punkt-Kalman-Filter (SPKF) geht von einem nichtlinearen Signalprozessmodell (SPM) mit n -dimensionalem Zustandsvektor und q -dimensionalen Ausgangsvektor aus:

$$\begin{aligned}\underline{x}(k+1) &= \underline{\Phi}(\underline{x}(k), \underline{u}(k), \underline{w}(k), k) \\ y(k) &= \underline{h}(\underline{x}(k), \underline{v}(k), k).\end{aligned}\tag{1}$$

Zentraler Bestandteil des SPKFs ist die sogenannte „*Unscented Transformation*“ (UT), die eine nichtlineare Transformation sogenannter *Sigma-Punkte* darstellt, wodurch sich auch der Name des Filters erklärt.

Für das SPKF muss das SPM (1) zunächst mit einem erweiterten Zustandsvektor beschrieben werden, der neben den ursprünglichen Zustandsgrößen auch die Rauschgrößen am Ein- und Ausgang des Systems enthält:

$$\underline{x}_e(k) = \begin{bmatrix} \underline{x}(k) \\ \underline{w}(k) \\ \underline{v}(k) \end{bmatrix}.$$

Schätzgröße ist dann dieser vergrößerte Zustandsvektor, dessen ersten beide Momente sich zum Beginn der Schätzung zu

$$\underline{x}_e(0) = \begin{bmatrix} \underline{m}_x(0) \\ \underline{0} \\ \underline{0} \end{bmatrix} \quad \text{und} \quad \underline{P}_e(0) = \begin{bmatrix} \underline{P}(0) & \underline{0} & \underline{0} \\ \underline{0} & \underline{Q}(0) & \underline{0} \\ \underline{0} & \underline{0} & \underline{R}(0) \end{bmatrix}\tag{2}$$

ergeben. Im Folgenden wird zur vereinfachten Schreibweise dieser vergrößerte Zustandsvektor wieder mit $\underline{x}(k)$ bezeichnet. Das SPM als Ausgangspunkt für das SPKF lautet dann

$$\underline{x}(k+1) = \underline{f}(\underline{x}(k), \underline{u}(k), k)\tag{3}$$

$$\underline{y}(k) = \underline{g}(\underline{x}(k), k).\tag{4}$$

Die Gleichungen des SPKFs gliedern sich dann wie bei den anderen Filtern gemäß der beiden Aufgaben Prädiktion und Filterung:

- **Prädiktionsgleichungen**

$$\hat{\underline{x}}(k|k-1) = \sum_{i=0}^{2n} W^{(i)} \cdot \hat{\underline{x}}^{(i)}(k)$$

$$\underline{P}(k|k-1) = \sum_{i=0}^{2n} W^{(i)} \cdot \left(\hat{\underline{x}}^{(i)}(k) - \hat{\underline{x}}(k|k-1) \right) \left(\hat{\underline{x}}^{(i)}(k) - \hat{\underline{x}}(k|k-1) \right)^T$$

(Anfangswerte: $\hat{\underline{x}}(0|-1) = \hat{\underline{x}}(0)$, $\underline{P}(0|-1) = \underline{P}(0)$, vgl. (2))

Hierbei sind definiert:

- Gewichte: $W^{(0)} = 1 - \frac{n}{\alpha}$, $W^{(i)} = \frac{1}{2\alpha}$ ($i = 1, \dots, 2n$) (Gauß-ZVn: $\alpha = 3$)
- Sigma-Punkte: $\hat{\underline{x}}^{(i)}(k) = \underline{f}(\hat{\underline{x}}^{(i)}(k-1), \underline{u}(k), k)$

Mit:

$$\hat{\underline{x}}^{(0)}(k-1) = \hat{\underline{x}}(k-1|k-1),$$

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

$$\tilde{\underline{x}}^{(i)} = \left(\sqrt{\alpha \underline{P}(k-1|k-1)} \right)_i^T \quad (i = 1, \dots, n),$$

$$\tilde{\underline{x}}^{(n+i)} = -\left(\sqrt{\alpha \underline{P}(k-1|k-1)} \right)_i^T \quad (i = 1, \dots, n),$$

wobei $\sqrt{\underline{A}} = \underline{V}_A \underline{\Lambda}_A^{1/2} \underline{V}_A^{-1}$ mit $\underline{\Lambda}_A^{1/2} = \text{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n})$

- **Filtergleichungen**

$$\underline{\hat{x}}(k|k) = \underline{\hat{x}}(k|k-1) + \underline{K}(k) \{ \underline{y}(k) - \underline{\hat{y}}(k|k-1) \}$$

$$\underline{P}(k|k) = \underline{P}(k|k-1) - \underline{K}(k) \underline{P}_{yy}(k|k-1) \underline{K}^T(k)$$

$$\underline{K}(k) = \underline{P}_{xy}(k|k-1) \underline{P}_{yy}^{-1}(k|k-1)$$

Hierbei sind:

- $\underline{\hat{y}}(k|k-1) = \sum_{i=0}^{2n} W^{(i)} \underline{\hat{y}}^{(i)}(k),$

$$\underline{\hat{y}}^{(i)}(k) = \underline{g}(\underline{\hat{x}}^{(i)}(k), k)$$

- $\underline{P}_{yy}(k|k-1) = \sum_{i=0}^{2n} W^{(i)} \left(\underline{\hat{y}}^{(i)}(k) - \underline{\hat{y}}(k|k-1) \right) \left(\underline{\hat{y}}^{(i)}(k) - \underline{\hat{y}}(k|k-1) \right)^T,$

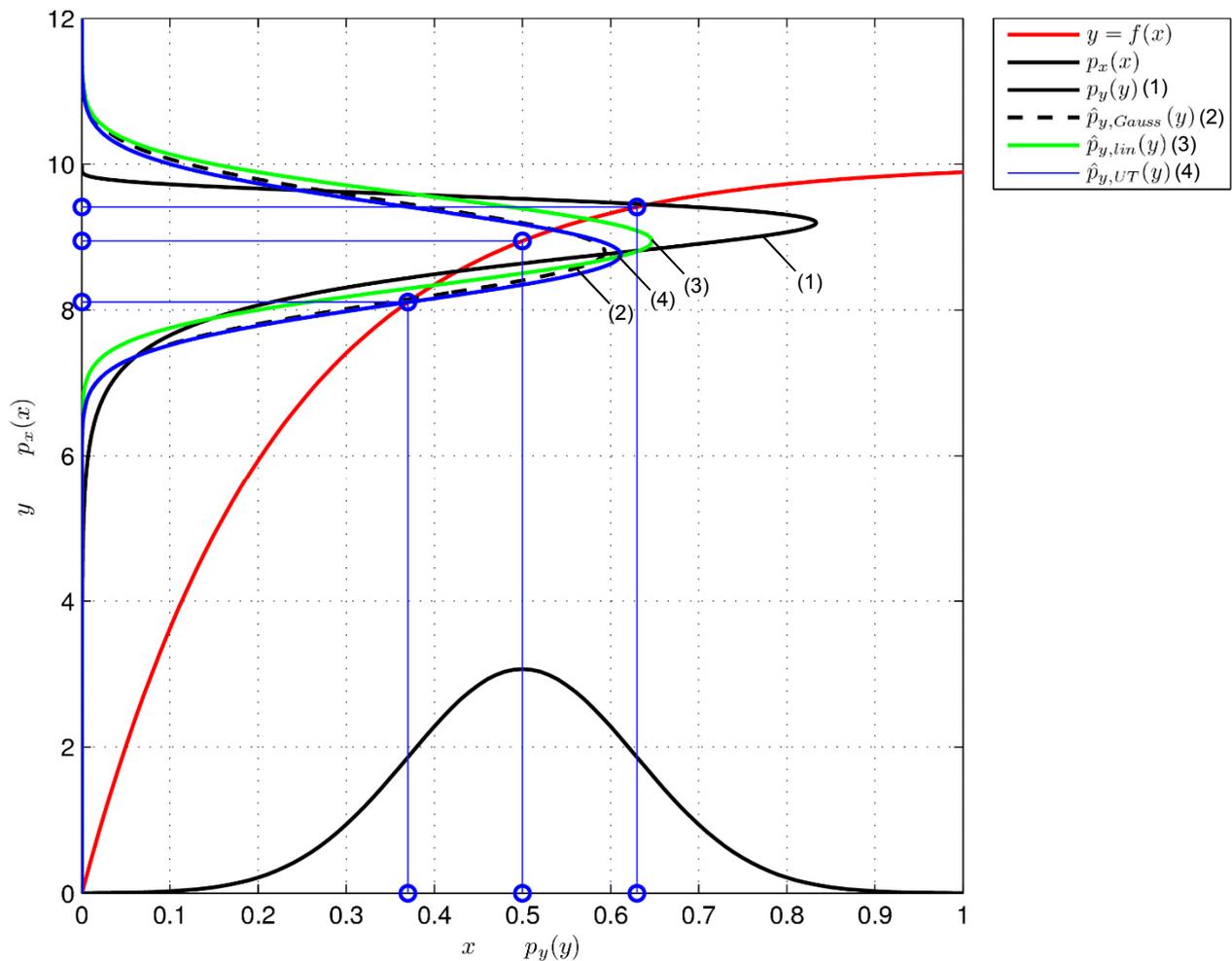
$$\underline{P}_{xy}(k|k-1) = \sum_{i=0}^{2n} W^{(i)} \left(\underline{\hat{x}}^{(i)}(k) - \underline{\hat{x}}(k|k-1) \right) \left(\underline{\hat{y}}^{(i)}(k) - \underline{\hat{y}}(k|k-1) \right)^T$$

Durch die hier angegebene Wahl der Sigma-Punkte wird bei Gauß-Zufallsvariablen eine gute Approximation von Erwartungswert und Kovarianz bis zu den Termen 4. Ordnung erreicht.

Vergleich verschiedener Dichtefunktionen

Beispielhafte Verhältnisse bei der Abbildung der Dichtefunktion $p_x(x)$ als

- $p_y(y)$ bei nichtlinearer Abbildung mit $y = f(x)$
- $\hat{p}_{y,Gau\beta}(y)$ bei Gauß-Approximation der Momente von $p_y(y)$
- $\hat{p}_{y,lin}(y)$ bei Gauß-Approximation der Momente beim EKF
- $\hat{p}_{y,UT}(y)$ bei Gauß-Approximation der Momente beim SPKF



Am Beispiel der Entfernungsschätzung sollen die Genauigkeiten bei der Momentenapproximation bei der Modelllinearisierung einerseits (Basis z.B. für das EKF) und der Unscented Transformation andererseits (Basis für das SPKF) verglichen werden.

Ausgangspunkt ist das nichtlineare Modell zur Umrechnung von Polarkoordinaten in kartesische Koordinaten bei der Entfernungsmessung gemäß Bild 1.

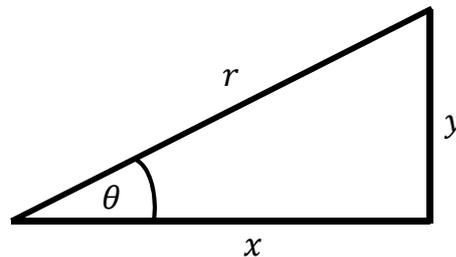


Bild 1

Es gilt der Zusammenhang

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} r \cdot \cos \theta \\ r \cdot \sin \theta \end{bmatrix}. \quad (1)$$

Nun ist die Entfernung r in der Regel relativ genau messbar ($\sigma_r \approx 2\text{cm}$), wohingegen der Peilwinkel θ größeren Schwankungen unterliegt ($\sigma_\theta \approx 15^\circ$).

Als konkretes Anschauungsbeispiel werden Simulationen aus Julier/Uhlmann: „*Unscented Filtering and Nonlinear Estimation*“ (Proc. of the IEEE: 92(3): S.401-422) betrachtet. Hierbei wurden einige hundert Datenpaare $[x, y]^T$ gemäß (1) aufgezeichnet; die zugehörigen Polarkoordinaten r und θ wurden additiv mit mittelwertfreiem Gauß'schem weißem Rauschen gestört.

3.4.2 Beispiel Entfernungsschätzung: Vergleich von Linearisierung und UT (2)

ORS 3-21

Bild 2 zeigt die Daten (oben) sowie den echten (\times) und den durch Linearisierung (\circ) erhaltenen Mittelwert. Zusätzlich sind die echten ($-$) und die durch die Linearisierung erhaltenen (\bullet) 1σ -Konturplots dargestellt. Es ist ein deutlicher Bias zu erkennen, der durch die Transformation (1) für die Mittelwerte entsteht und auf den Fehler durch Taylorreihenabbruch bei der Linearisierung zurückzuführen ist.

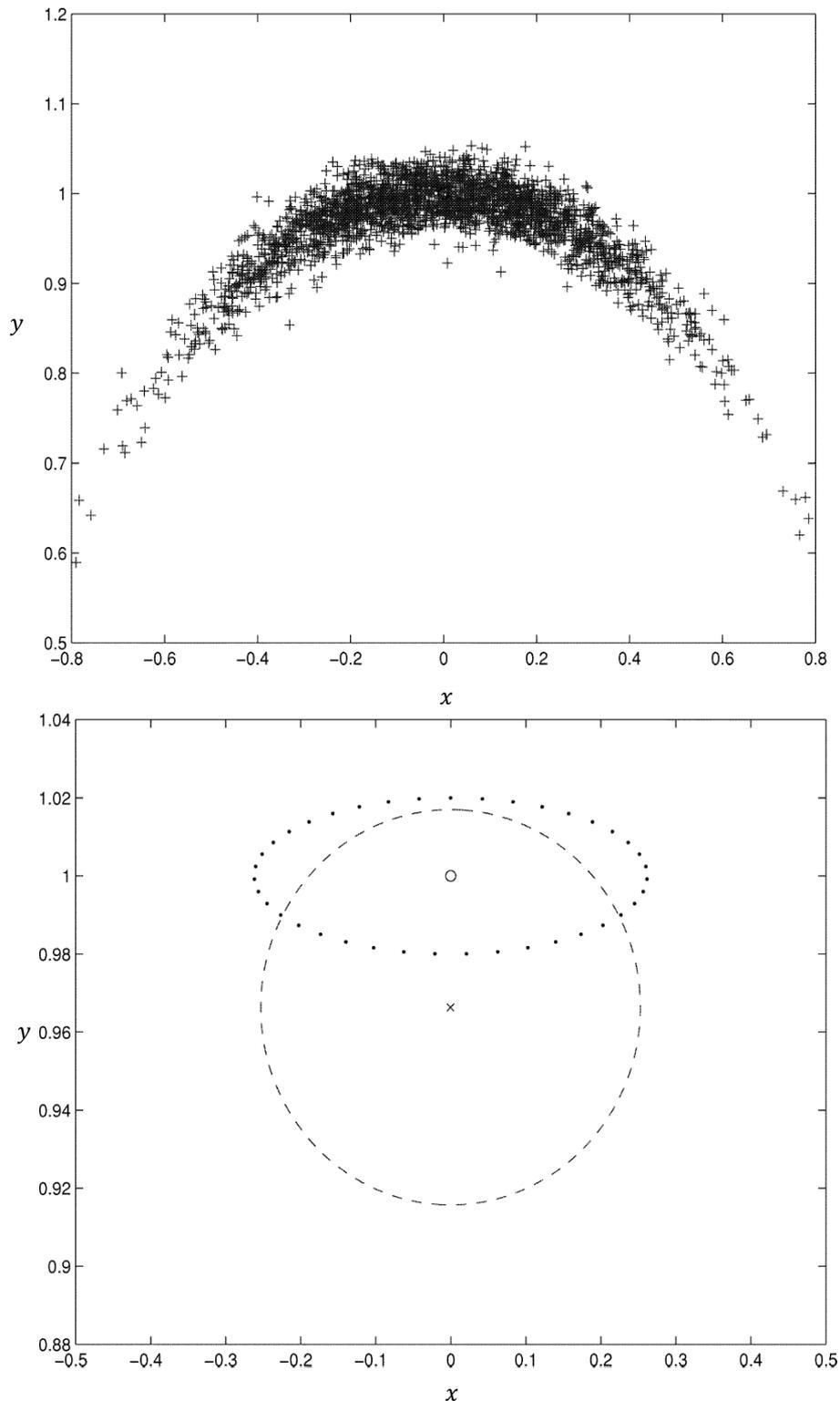


Bild 2

3.4.2 Beispiel Entfernungsschätzung: Vergleich von Linearisierung und UT (3)

ORS 3-22

Als Gegenentwurf zeigt Bild 3 im oberen Teil die Sigma-Punkte, welche mittels der nichtlinearen Abbildung (1) und der Gewichtung entsprechend der in der Vorlesung gezeigten Unscented Transformation (gemäß BB ORS 3-16 bis 3-18) den Erwartungswert (+) und den zugehörigen 1σ -Konturplot (*) liefern. Man erkennt deutlich die hohe Genauigkeit, die auf die bessere Approximation (bis zu Termen 4. Ordnung) zurückzuführen ist.

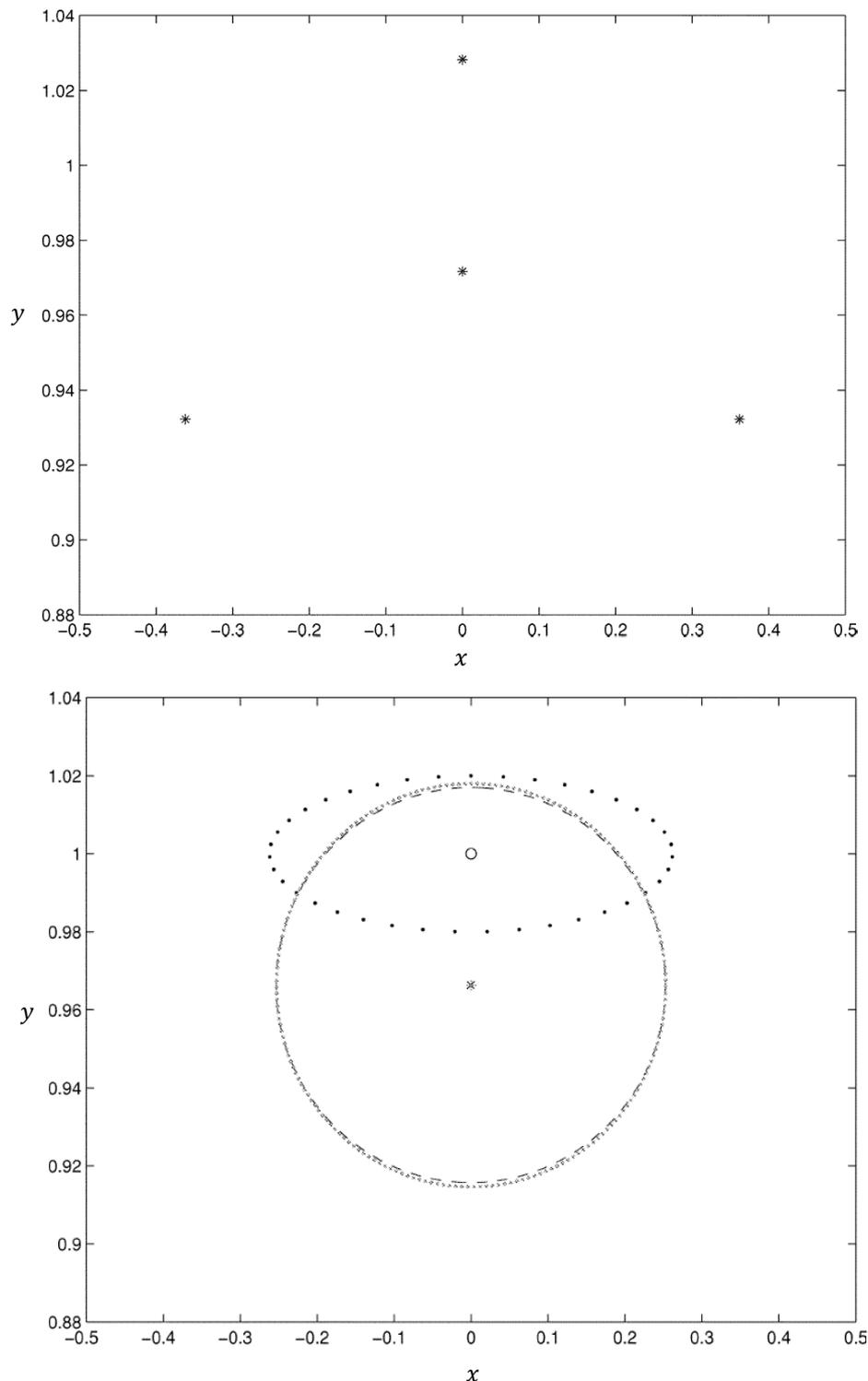


Bild 3