

## 2. Probabilistische Methoden

## 2. Probabilistische Methoden

### 2.1 Stochastische Grundbegriffe

2.1.1 Definition von Wahrscheinlichkeit

2.1.2 Interpretation von Wahrscheinlichkeit

2.1.3 Bedingte Wahrscheinlichkeit und Unabhängigkeit

2.1.4 Kenngrößen von Verteilungen

### 2.2 Klassische Statistik

2.2.1 Schätzung von Zufallsgrößen

2.2.2 Beobachtungsunsicherheit

2.2.3 Konzentration und Dispersion

2.2.4 Zentrale Tendenz

### 2.3 Schätzung bei mehreren Beobachtungen

2.3.1 Schätzung einer statischen Zielgröße

2.3.2 Schätzung einer mehrdimensionalen stat. Zielgröße

2.3.3 Kalman-Filter

2.3.4 Extended Kalman-Filter

## 2. Probabilistische Methoden (Forts.)

### 2.4 Guide to the Expression of Uncertainty in Measurement (GUM)

#### 2.4.1 Grundbegriffe

#### 2.4.2 Ermittlung der Messunsicherheit gemäß GUM

#### 2.4.3 Vollständige Beschreibung von Zufallsgrößen durch PDFen

### 2.5 Bayes'sche Fusion

#### 2.5.1 Bayes'sche Methodik

#### 2.5.2 Unabhängigkeitsannahmen

#### 2.5.3 Bayes'sche Fusion

#### 2.5.4 Logarithmierung

#### 2.5.5 Marginalisierung

#### 2.5.6 Bewertung

### 2.6 Tracking

#### 2.6.1 Bayes'sches Tracking

#### 2.6.2 Tracking-Problematik

#### 2.6.3 Vorselektion

#### 2.6.4 Hypothesen-Tracking

## 2.1 Stochastische Grundbegriffe

### 2.1.1 Definition von Wahrscheinlichkeit

#### Empirisch: (Ideales) Zufallsexperiment:

- Das Experiment wird unter **genau festgelegten Bedingungen** ausgeführt
- Die **Menge der möglichen Elementarergebnisse** ist vor der Durchführung bekannt
- Das Experiment kann (zumindest prinzipiell) **beliebig oft** unter den gleichen Bedingungen **wiederholt** werden
- Führt zur **frequentistischen Interpretation** von Wahrscheinlichkeit

#### Beispiele:

- Werfen einer fairen Münze
- Würfeln
- Lotto

#### Bezeichnungen:

- **Grundraum**  $\Omega$ : (Endliche) Menge der möglichen Elementarergebnisse
- **Ereignis**: Menge  $A \subseteq \Omega$
- **Ereignisraum**: Menge aller Ereignisse  $\mathcal{P}(\Omega)$  (Potenzmenge von  $\Omega$ )

## 2.1.1 Definition von Wahrscheinlichkeit

### Axiomensystem von Kolmogorov:

Ein **endlicher Wahrscheinlichkeitsraum** ist ein Tupel  $(\Omega, \text{Pr})$  mit

- Grundraum  $\Omega$  ist nichtleere, endliche Menge
- Wahrscheinlichkeit  $\text{Pr}$  ist reellwertige Funktion auf der Potenzmenge  $\mathcal{P}(\Omega)$  mit
  - $\text{Pr}(A) \geq 0$  für  $A \subseteq \Omega$  (Nichtnegativität)
  - $\text{Pr}(\Omega) = 1$  (Normiertheit)
  - $\text{Pr}\left(\bigcup_i A_i\right) = \sum_i \text{Pr}(A_i)$  (Additivität)  
für je zwei disjunkte Ereignisse  $A_i \cap A_j = \emptyset$

Beispiel:  $\Omega = \{\text{„Kopf“}, \text{„Zahl“}\} = \{K, Z\}$  (mögliche Elementarergebnisse)

Ereignisraum:  $\mathcal{P}(\Omega) = \{\emptyset, \{K\}, \{Z\}, \{K, Z\}\}$

Axiome  $\Rightarrow \text{Pr}(\emptyset) = 0, \quad \text{Pr}(K) = 1 - \text{Pr}(Z), \quad \text{Pr}(\Omega) = 1$

Zusatzwissen:  $\text{Pr}(K) = \text{Pr}(Z) = 0,5$  (faire Münze);

$\text{Pr}(K), \text{Pr}(Z) \neq 0,5$  (gezinkte Münze)

## 2.1.2 Interpretation von Wahrscheinlichkeit

Mit den Axiomen verträgliche Interpretationen:

### Axiome nach Kolmogorov

#### Empirisch (Frequentistisch)

- Wahrscheinlichkeiten: Grenzwerte relativer Häufigkeiten, wenn Zufallsexperiment unendlich oft wiederholt wird.
- Unbedingter, objektiver Anspruch
- Beschreibung von Massenerscheinungen
- (Klassische) Statistik

#### Degree-of-Belief (Grad des Dafürhaltens)

- Wahrscheinlichkeiten: Verkörperung auch von subjektiver Unsicherheit bzw. des Wissensstands
- Kein unbedingter Anspruch
- Situationsabhängig
- Allgemeinere Interpretation als empirische Deutung

#### Objektiv

- Verkörperung von Fakten
- Modellierung von Nichtwissen (z.B. Prinzip der Maximalen Entropie)

#### Subjektiv

- Verkörperung des Wissens, der Überzeugung eines Individuums
- Z.B. fiktive Wettsituation

Interpretation

## 2.1.2 Interpretation von Wahrscheinlichkeit

### Frequentistische Interpretation:

Zufallsexperiment  $n$ -mal wiederholen

$a_i$  Ereignis bei der  $i$ -ten Durchführung ( $i = 1, \dots, n$ ),  $\mathbf{a} := (a_1, \dots, a_n)$

Relative Häufigkeit des Ereignisses  $A$ :  $r_{n,\mathbf{a}}(A) := \frac{1}{n} \left| \{j \mid a_j \in A, j \in \{1, \dots, n\}\} \right|$

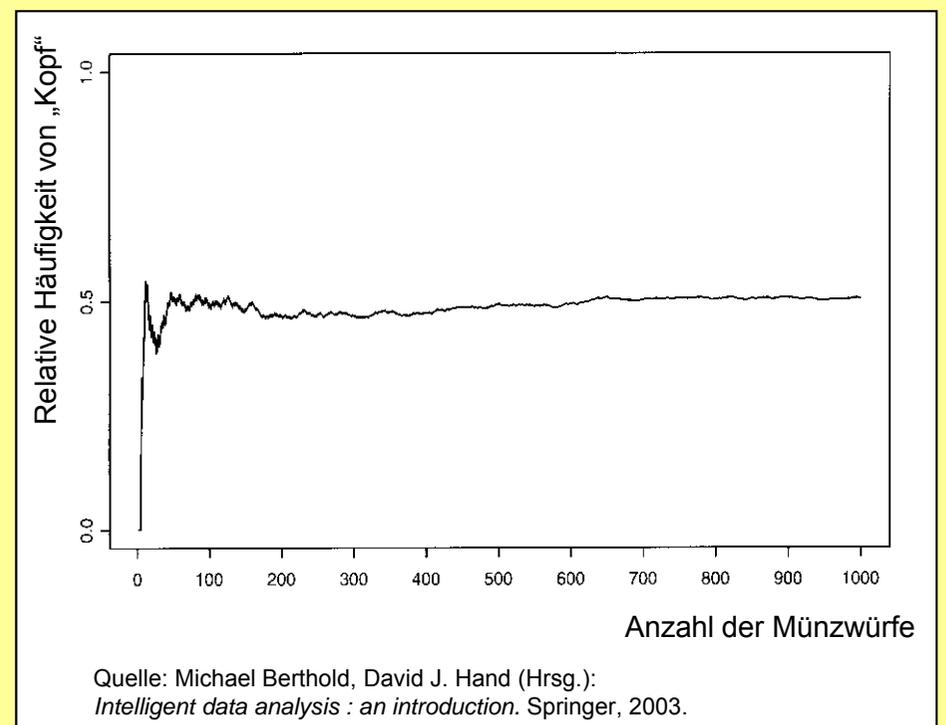
→ Empirisches Gesetz über die Stabilisierung relativer Häufigkeiten

### Bsp.: Werfen einer Münze

Relative Häufigkeit von „Kopf“ in einer Folge von 1000 Münzwürfen lässt Konvergenz annehmen:

$$\lim_{n \rightarrow \infty} r_{n,\mathbf{a}}(\text{Kopf}) = \lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n \delta_{a_i}^{\text{Kopf}}}{n} \rightarrow \Pr(\text{Kopf})$$

$$\text{mit } \delta_{a_i}^{\text{Kopf}} = \begin{cases} 1 & \text{falls } a_i = \text{Kopf} \\ 0 & \text{sonst} \end{cases}$$



## 2.1.2 Interpretation von Wahrscheinlichkeit

Bsp.: „Die Wahrscheinlichkeit einer 1 bei einem fairen Würfel ist gleich  $1/6$ .“

- Physikalische Realität, „nachprüfbar“ durch wiederholtes Experiment
- Entspricht frequentistischer Interpretation

Beispiel: Mehrheitsverhältnisse bei politischen Wahlen

- Zählvorgang, aber kein wiederholbares Experiment
- Interpretation im frequentistischen Sinn unzulässig

Bsp.: „Die Wahrscheinlichkeit für Leben auf dem Jupiter ist nach heutigem Kenntnisstand gleich  $1/10000$ .“

- Kein wiederholbares Experiment
- Grad der Überzeugung aufgrund verfügbaren Wissens (z.B. Beobachtungen, Modellrechnungen)
- Degree-of-Belief-Interpretation (objektiv)

Bsp.: „Die Wahrscheinlichkeit für Leben in anderen Galaxien ist gleich  $1/2$ .“

- Vollkommene Unwissenheit über Sachverhalt
- Degree-of-Belief-Interpretation (subjektiv)

## 2.1.3 Bedingte Wahrscheinlichkeit und Unabhängigkeit

$A, B, A_1, \dots, A_n$  Ereignisse auf einem endl. Wahrscheinlichkeitsraum  $(\Omega, \Pr)$

- **Bedingte Wahrscheinlichkeit:**

$$\Pr(A | B) := \frac{\Pr(A \cap B)}{\Pr(B)}$$

$\Pr(\cdot | B)$  : Wahrscheinlichkeitsverteilung für festes  $B$

- $A, B$  sind **unabhängig**, wenn gilt

$$\Pr(A, B) := \Pr(A \cap B) = \Pr(A)\Pr(B)$$

dann:  $\Pr(A | B) = \Pr(A)$

$$\Pr(B | A) = \Pr(B)$$

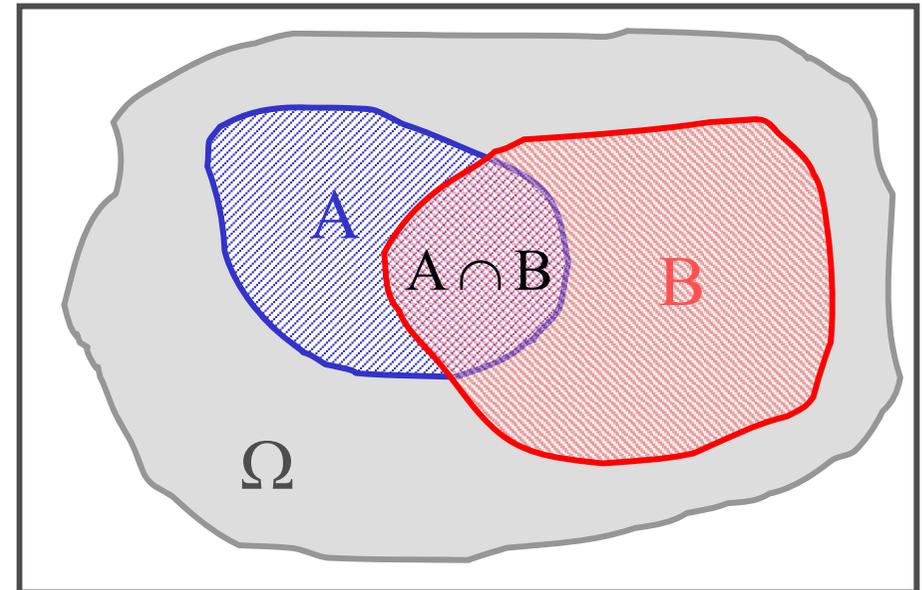
- **Multiplikationsformel:**

$$\Pr(A_1, A_2) = \Pr(A_1)\Pr(A_2 | A_1)$$

$$\Pr(A_1, \dots, A_n) = \Pr(A_1)\Pr(A_2 | A_1)\Pr(A_3 | A_1, A_2) \dots \Pr(A_n | A_1, \dots, A_{n-1})$$

- **Formel von Bayes:**

$$\Pr(A | B) = \frac{\Pr(B | A)\Pr(A)}{\Pr(B)}$$

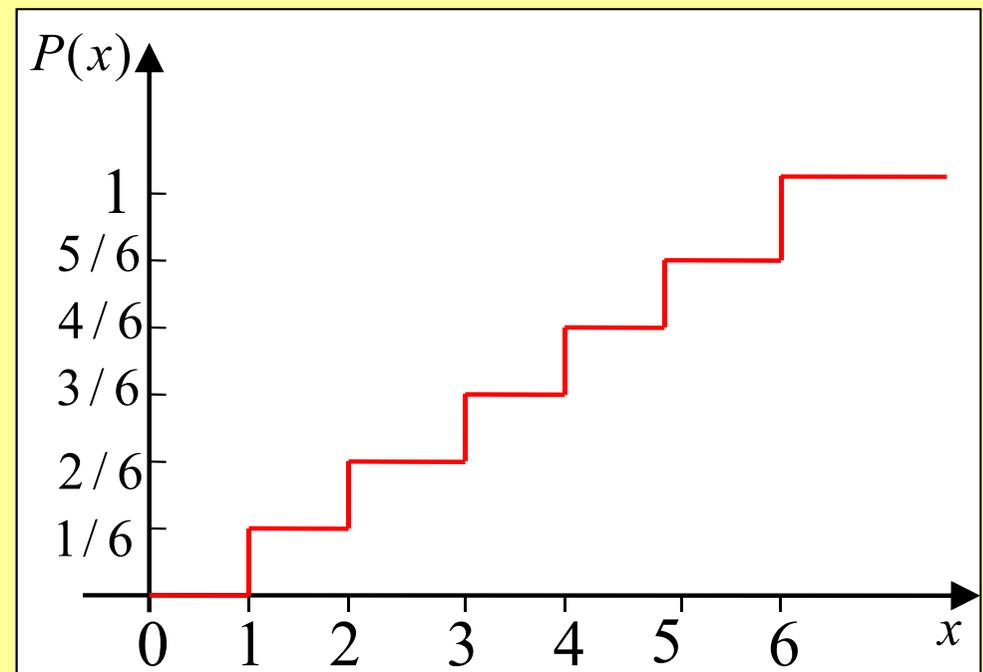
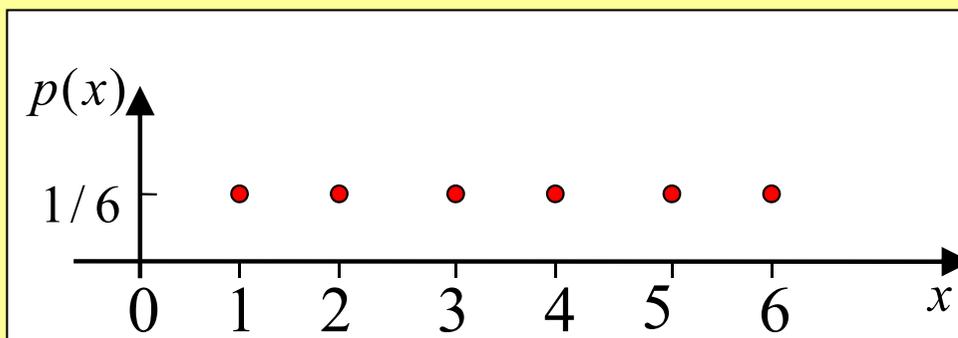


## 2.1.4 Kenngrößen von Verteilungen

- $X: \Omega \rightarrow \mathbb{R}$  (reellwertige) **Zufallsvariable** für **diskreten Grundraum**  $\Omega$   
 $x$ : Realisierung von  $X$
- **Wahrscheinlichkeitsfunktion**:  $p_i := \Pr(X = x_i)$  probability mass function  
„Dichte“ (bezüglich des Zählmaßes)
- **Verteilungsfunktion**:  $P(x_i) := \sum_{j: x_j \leq x_i} p_j$  Kumulative Verteilung  
Summation

### Beispiel:

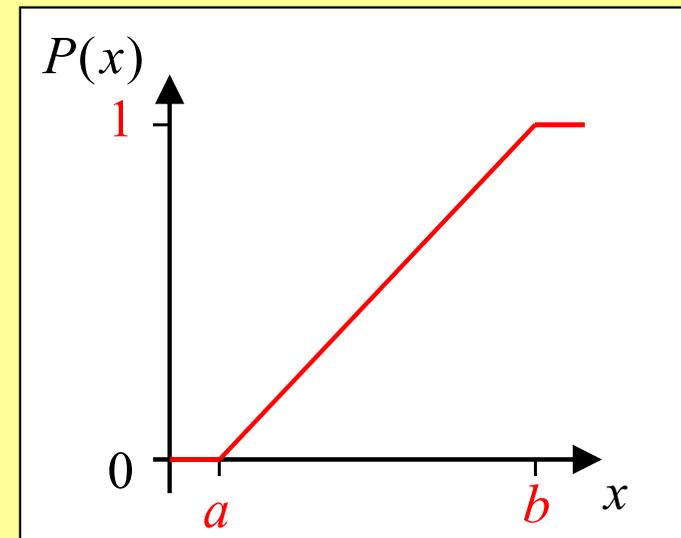
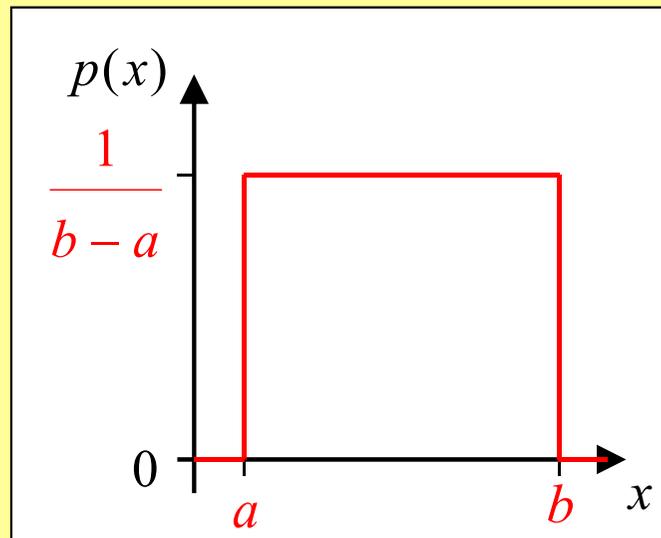
Gleichverteilung auf  $\Omega = \{1, \dots, 6\}$



## 2.1.4 Kenngrößen von Verteilungen

- $X: \Omega \rightarrow \mathbb{R}$  **Zufallsvariable** für **kontinuierlichen Grundraum**  $\Omega$   
 $x$ : Realisierung von  $X$ , aber:  $\Pr(X = x) = 0$
- **Verteilungsfunktion**:  $P(x_i) := \Pr(x \leq x_i)$  Kumulative Verteilung, CDF  
cumulative distribution function  
(Siebvorgang)
- **Wahrscheinlichkeitsdichte**:  $p(x) := \frac{dP(x)}{dx}$  Verteilungsdichte, PDF  
probability density function  
(Integral statt Summe)  
$$P(x_i) = \int_{-\infty}^{x_i} p(x) dx$$

**Beispiel:**  
Gleichverteilung  
auf  $\Omega = [a, b]$



## 2.1.4 Kenngrößen von Verteilungen

- **Erwartungswert:**

stetiger Fall: 
$$E\{X\} = \int_{-\infty}^{\infty} x p(x) dx =: \mu$$

diskreter Fall: 
$$E\{X\} = \sum_i x_i p_i =: \mu$$

Existiert, falls das ihn definierende Integral bzw. die ihn definierende Summe absolut konvergiert, d.h.

$$E\{|X|\} = \int_{-\infty}^{\infty} |x| p(x) dx < \infty \quad \text{bzw.} \quad E\{|X|\} = \sum_i |x_i| p_i < \infty$$

- Erwartungswertbildung ist eine lineare Operation, d.h.

$$E\{aX + b\} = a E\{X\} + b, \quad a, b \in \mathbb{R}$$

Für beliebige  $X_1, \dots, X_N$  gilt:

$$E\left\{\sum_{i=1}^N X_i\right\} = \sum_{i=1}^N E\{X_i\}$$

- Unkorrelierte  $X_1, X_2$  (s.u.):  $E\{X_1 \cdot X_2\} = E\{X_1\} \cdot E\{X_2\}$

## 2.1.4 Kenngrößen von Verteilungen

- **Momente:**  $\alpha_k := E\{X^k\}, \quad k \in \mathbb{N}_0$   
Spezialfälle:  $\alpha_0 = 1$   
 $\alpha_1 = E\{X\} = \mu$  Erwartungswert von  $X$

**Zentralmomente:**  $\beta_k := E\left\{\left(X - E\{X\}\right)^k\right\}, \quad k \in \mathbb{N}_0$

Spezialfall:  $\beta_2 = E\left\{\left(X - E\{X\}\right)^2\right\} =: \text{Var}\{X\}$  **Varianz**  
 $=: \sigma_x^2$   
 $= \alpha_2 - \alpha_1^2$

## 2.1.4 Kenngrößen von Verteilungen

### ■ Varianz:

- Varianzbildung ist *keine* lineare Operation:

$$\text{Var}\{aX + b\} = a^2 \text{Var}\{X\}, \quad a, b \in \mathbb{R}$$

- **Standardabweichung:**

$$\sigma_x := \sqrt{\beta_2} = \sqrt{\text{E}\{(X - \text{E}\{X\})^2\}}$$

- Für *beliebige* Zufallsvariable  $X_1, X_2, \dots, X_N$  gilt:

$$\text{Var}\left\{\sum_{i=1}^N X_i\right\} = \sum_{i=1}^N \text{Var}\{X_i\} + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N \text{Cov}\{X_i, X_j\}$$

**Kovarianz:**  $\text{Cov}\{X_i, X_j\} := \text{E}\{(X_i - \text{E}\{X_i\}) \cdot (X_j - \text{E}\{X_j\})\}$

- Falls  $\text{Cov}\{X_i, X_j\} = 0$ :  $\text{E}\{X_i \cdot X_j\} = \text{E}\{X_i\} \cdot \text{E}\{X_j\}$   
Zufallsvariablen  $X_1, X_2, \dots, X_N$  sind *unkorreliert*.

$$\text{Var}\left\{\sum_{i=1}^N X_i\right\} = \sum_{i=1}^N \text{Var}\{X_i\}$$

Aus Unabhängigkeit von  $X_1, X_2$ :  $p(x_1, x_2) = p(x_1) \cdot p(x_2)$   
folgt Unkorreliertheit, Umkehrung gilt nicht allgemein

## 2.1 Stochastische Grundlagen

---

Zusammenfassung:

- Unsicherheitsbehaftete Information lässt sich probabilistisch in Form von **Wahrscheinlichkeiten** beschreiben, die dem **Axiomensystem von Kolmogorov** genügen
- Solche Wahrscheinlichkeiten lassen sich auf unterschiedliche Arten gewinnen und interpretieren:
  - **Frequentistische Interpretation**  
Wiederholbares Zufallsexperiment
  - **Degree-of-Belief-Interpretation**  
Repräsentation von Wissen/Unwissen
- Mit Wahrscheinlichkeiten lässt sich rechnen, z.B. Bayes'sche Formel
- Wahrscheinlichkeitsverteilungen lassen sich für **diskrete und kontinuierliche Zufallsvariablen** angeben
- Einfache Kenngrößen lassen sich aus den **Momenten** der Zufallsvariablen gewinnen

## 2.2 Klassische Statistik

---

### 2.2.1 Schätzung von Zufallsgrößen

Problem: Beobachtungen sind zufallsbehaftet (gestört)

Ziel: Extraktion der ungestörten, eigentlich gewünschten Nutzinformation

#### ▪ **Verteilungsannahme:**

- Daten  $x$ : Realisierung einer Zufallsvariablen  $X$
- Wertebereich von  $X$ : Stichprobenraum  $\mathcal{X}$  (oft:  $\mathcal{X} = \mathbb{R}, \mathbb{R}^n$ )
- Annahme über Verteilung von  $X$  (kumulative Verteilung  $P$  bzw. Verteilungsdichte  $p$ ) aus Rahmenbedingungen/Vorwissen
- $(P_\theta \mid \theta \in \Theta)$  Klasse von parametrischen Wahrscheinlichkeitsmaßen mit Parameterraum  $\Theta$
- Ziel: Schätzung von  $\theta$

#### ▪ **Statistisches Modell:** $(\mathcal{X}, \{P_\theta \mid \theta \in \Theta\})$

## 2.2.1 Schätzung von Zufallsgrößen

- **Schätzer:** Abbildung  $\hat{\theta} : \mathcal{X} \rightarrow \hat{\Theta}$ ,  $\hat{\Theta} \supset \Theta$   
( $\hat{\Theta} \supset \Theta$  aus mathematischen Gründen erlaubt)
  - Beobachtungen:  $n$ -dimensionaler Zufallsvektor  $\mathbf{X} = (X_1, \dots, X_n)$ 
    - $\mathbf{X} = (X_1, \dots, X_n)$  heißt **Stichprobe vom Umfang  $n$**
    - $X_1, \dots, X_n$  unabhängig bezüglich  $P_\theta$  (für jedes  $\theta \in \Theta$ ), reellwertig
    - $\mathcal{X}_n \subseteq \mathbb{R}^n$  Stichprobenraum für  $\mathbf{X} = (X_1, \dots, X_n)$
  - $\hat{\theta}_n : \mathcal{X}_n \rightarrow \hat{\Theta}$  : Schätzer für  $\theta$ , der sich auf  $\mathbf{X}$  stützt ( $n \in \mathbb{N}$ )
  - $(\hat{\theta}_n)_{n \in \mathbb{N}}$  : **Schätzfolge**
- **Anforderungen an Schätzer/Schätzfolge:**
  - **Erwartungstreue:** Schätzer  $\hat{\theta}$  soll im Mittel den wahren Wert  $\theta$  liefern, d.h. für metrisch skalierte Merkmale:

$$E_\theta \{ \hat{\theta} \} = \theta$$

sonst Bias:  $b_\theta = E_\theta \{ \hat{\theta} \} - \theta$

## 2.2.1 Schätzung von Zufallsgrößen

---

- **Asymptotische Erwartungstreue** von Schätzfolgen:

$$\lim_{n \rightarrow \infty} E_{\theta} \{ \hat{\theta}_n \} \stackrel{!}{=} \theta$$

- **Konsistenz** von Schätzfolgen: Schätzung soll sich mit wachsender Anzahl von Messungen immer mehr dem wahren Wert nähern:

$$\lim_{n \rightarrow \infty} \Pr \left( \left| \hat{\theta}_n - \theta \right| < \varepsilon \right) \stackrel{!}{=} 1 \quad \text{für jedes } \varepsilon > 0$$

- **Wirksamkeit/Effizienz**: Ein Schätzer ist wirksam, wenn es keinen anderen mit kleinerer Varianz gibt.  
Der Schätzer  $\hat{\theta}^*$  aus allen erwartungstreuen Schätzern, der die kleinste Varianz hat, wird als effizientester oder wirksamster Schätzer bezeichnet.

## 2.2.2 Beobachtungsunsicherheit

---

- Beobachtungen eines Sensorsystems:  
Unsicherheitsbehaftete Informationen
- Vollständige Beschreibung: Verteilung der Beobachtungen
- Vereinfachung: **Zentrale Tendenz** und **Konzentration/Dispersion (Streuung)** vieler Beobachtungen
- Für eine Beobachtung: Angabe des beobachteten Werts und einer Beobachtungsunsicherheit
- **Beobachtungsunsicherheit** als Kennzeichnung des Vertrauens in die Beobachtung (Degree-of-Belief-Interpretation)
- Grundlagen für Aussagen zur Beobachtungsunsicherheit beruhen auf
  - + Statistische Analyse: Häufigkeitsverteilung  
(Erfahrung, häufige Wiederholung der Beobachtung)
  - + Art des Zustandekommens  
(z.B. Messverfahren, Umgebungsbedingungen)
- **Spezifikationsmöglichkeiten** für zentrale Tendenz und Konzentration/Dispersion **hängen von der Skalierung der Merkmale ab:**  
Nominale, ordinale, metrische Skalierung

## 2.2.2 Beobachtungsunsicherheit

### Beschreibung der Beobachtung durch statistische Analyse:

- Mehrere Beobachtungen desselben Sachverhalts durch denselben Sensor: Beispiel Beobachtungen einer metrischen Größe  $\mathbf{X} = \{x_1, x_2, \dots, x_N\}$

#### ▪ Zentrale Tendenz

- Synonym: Lagemaß  $l\{\mathbf{X}\} = l\{x_1, x_2, \dots, x_N\}$
- Kennzeichnet die „grobe Lage“ der Beobachtung auf der Skala des Merkmals
- Forderung: Bei Verschiebung der Beobachtungen Mitverschiebung der zentralen Tendenz:

$$l\{x_1 + a, x_2 + a, \dots, x_N + a\} = l\{x_1, x_2, \dots, x_N\} + a$$

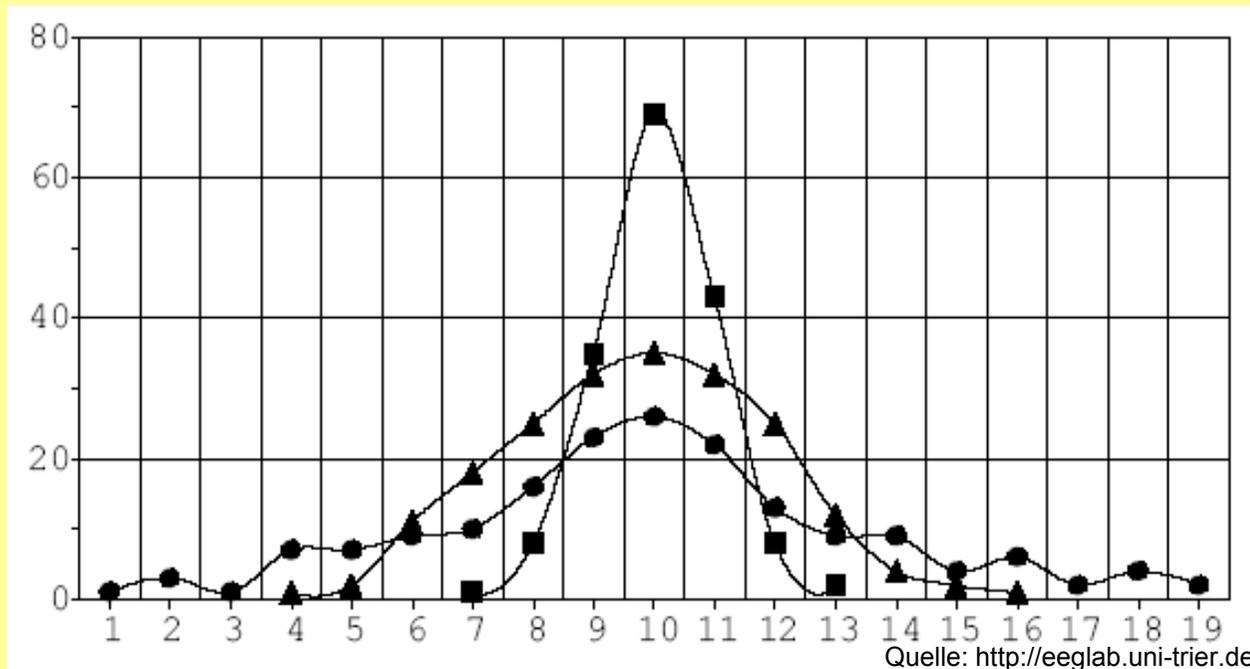
#### ▪ Dispersionsmaß

- Synonym: Streuungsmaß  $d\{\mathbf{X}\} = d\{x_1, x_2, \dots, x_N\}$
- Forderung: Keine Beeinflussung durch Verschiebung der Beobachtungen:

$$d\{x_1 + a, x_2 + a, \dots, x_N + a\} = d\{x_1, x_2, \dots, x_N\}$$

## 2.2.2 Beobachtungsunsicherheit

### Beispiel: Dispersion zur Charakterisierung von Verteilungen



- Die drei Verteilungen weisen trotz derselben **zentralen Tendenz** (Mittelwert) unterschiedliche Formen auf:  
Unterschiedliche **Dispersion**

## 2.2.3 Konzentration und Dispersion

### Entropie:

- Entropie der Beobachtungen  $\mathbf{X} = \{x_1, \dots, x_N\}$  mit  $J$  Merkmalsausprägungen  $m_j$  und den relativen Häufigkeiten

$$r(m_j) = \frac{1}{N} \left| \left\{ i \mid m_j = x_i, i = 1, \dots, N \right\} \right| \quad (j \in \{1, \dots, J\})$$

$$H\{\mathbf{X}\} = -\sum_{j=1}^J r(m_j) \log_2 r(m_j) \quad (0 \log_2(0) = 0 \text{ durch Grenzübergang})$$

- Geeignet als Konzentrationsmaß bei **nominalskalierten Merkmalen**

Beispiel: Die Elemente der Beobachtungen  $\mathbf{X}$  können zwei nominale Werte annehmen, die mit A und B bezeichnet werden.

- Keine Unsicherheit**, falls nur eines der Ereignisse  $\{x_i = A\}$  bzw.  $\{x_i = B\}$  eintritt:  $H\{\mathbf{X}\} = 0$
- Maximale Unsicherheit**, wenn die Ereignisse  $\{x_i = A\}$  und  $\{x_i = B\}$  mit gleicher relativer Häufigkeit auftreten:  $r(A) = r(B) = 1/2$

$$H\{\mathbf{X}\} = -2 \cdot \left( \frac{1}{2} \log_2 \frac{1}{2} \right) = 1 \rightarrow \text{Maximum}$$

## 2.2.3 Konzentration und Dispersion

**Spannweite:** Differenz zwischen maximalem und minimalem Wert

$$R\{\mathbf{X}\} = \max(\mathbf{X}) - \min(\mathbf{X})$$

- Hängt nur von den extremalen Beobachtungen ab
- Indirekt abhängig von der Anzahl der Beobachtungen: Kommt ein neuer Messwert hinzu, kann die Spannweite größer werden.
- Nicht erwartungstreuer Schätzer für  $\sigma_x$  :  $E\{R\} = \sigma_x \cdot d(N)$   
 $d(N)$ : Korrekturfaktor

Erwartungstreuer Schätzer für  $\sigma_x$  :  $\hat{\sigma}_x = \frac{R}{d(N)}$

- Große Ausreißerempfindlichkeit

## 2.2.3 Konzentration und Dispersion

### Spannweite: Differenz zwischen maximalem und minimalem Wert

- Oft Verwendung als Dispersionsmaß bei ordinalskalierten Merkmalen
- Aber: Bestimmung einer Differenz von Werten erforderlich!  
Daher bei ordinalskalierten Merkmalen besser Entropie als Konzentrationsmaß verwenden

#### Beispiel: Ordinalskala

Klausurnote	Häufigkeit
1 (sehr gut)	1
2 (gut)	3
3 (befriedigend)	2
4 (ausreichend)	2
5 (mangelhaft)	1
6 (ungenügend)	0

**Spannweite: 4**

Kommt noch eine Prüfung hinzu, so verändert sich die Spannweite nur, wenn die Note 6 auftritt.

## 2.2.3 Konzentration und Dispersion

**Standardabweichung:** Die (empirische) Standardabweichung der Beobachtungen  $\mathcal{S}$  ist definiert durch

$$s_x \{ \mathbf{X} \} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2}$$

- Quadrat der Standardabweichung: Empirische Varianz  $s_x^2$
- $s_x^2$  : Erwartungstreuer Schätzer für Varianz der Verteilung:

$$E\{s_x^2\} = \sigma_x^2 \quad \text{aber: } E\{s_x\} \neq \sigma_x$$

- Interpretation: Durchschnittliche Abweichung der Beobachtungen von der zentralen Tendenz
- Verwendung als Dispersionsmaß bei **metrisch skalierten Merkmalen**

<u>Beispiel:</u>	Gemessene Temperatur	Häufigkeit	Arithmetisches Mittel: ca. 14,42°C
<b>Intervallskala</b>	12°C	1	<b>Standardabweichung:</b> ca. <b>1,27°C</b>
	13°C	0	
	14°C	2	
	15°C	3	
	16°C	1	

## 2.2.4 Zentrale Tendenz

**Modus (Modalwert):** Häufigster Wert der Beobachtungen  $S = \{x_1, x_2, \dots, x_N\}$

- Wert, der am häufigsten vorkommt, wenn man aus der Gesamtheit der  $N$  Beobachtungen  $X = \{x_i\}$ ,  $x_i \in \{m_1, m_2, \dots, m_J\}$ ,  $i = 1, 2, \dots, N$  zufällig einen Wert auswählt:

$$\text{Modus}\{X\} = m_{j^*} \quad \text{mit} \quad j^* = \arg \max_j \left| \left\{ i \mid m_j = x_i, i \in \{1, \dots, N\} \right\} \right| \quad (j \in \{1, \dots, J\})$$

- Der Modus muss nicht eindeutig sein (z.B. bei **bimodalen** Verteilungen)
- Aussagekräftig, wenn viele Beobachtungen den Modus annehmen
- Bei metrischer Skalierung, kontinuierlicher Wertebereich: Einteilung der Beobachtungen in Klassen, dann Bestimmung der **modalen Klasse**.
- Erfordert lediglich **nominalskalierte Merkmale**

### Beispiele:

<u>Erkannte Farben</u>		<u>Gemessene Temperatur</u>	
	<u>Häufigkeit</u>		<u>Häufigkeit</u>
		12°C	1
Gelb	7	13°C	0
<b>Rot</b>	20	14°C	2
Grün	9	<b>15°C</b>	3
		16°C	1

**Nominalskala**

**Intervallskala**

## 2.2.4 Zentrale Tendenz

### Median (Medianwert):

- Falls die kumulierte Verteilungsfunktion  $P$  bekannt ist:  
Wert, für den  $P$  den Wert  $\frac{1}{2}$  annimmt (0,5-Quantil)

- Geordnete Beobachtungen:

$$\mathbf{X}_{( )} = \left\{ x_{(1)}, x_{(2)}, \dots, x_{(N)} \right\} \quad \text{Median} \{ \mathbf{X} \} = \begin{cases} x_{((N+1)/2)}, & N \text{ ungerade} \\ \frac{x_{(N/2)} + x_{(N/2+1)}}{2}, & N \text{ gerade} \end{cases}$$

- Wert, von dem alle übrigen Beobachtungen im Durchschnitt am wenigsten abweichen, so dass die mittlere absolute Abweichung minimal wird:

$$\frac{1}{N} \sum_{j=1}^N |x_j - \text{Median} \{ \mathbf{X} \} | \rightarrow \text{Minimum}$$

- Zentraler Wert: Ober- und unterhalb liegen gleich viele Beobachtungen
- Geringe Ausreißerempfindlichkeit

## 2.2.4 Zentrale Tendenz

### Median (Medianwert):

- Erfordert mindestens **ordinalskalierte Merkmale**  
(für  $N$  gerade eigentlich mindestens intervallskalierte Merkmale)
- Nur dann erwartungstreuer Schätzer für den Erwartungswert  $\mu$ , wenn Verteilung symmetrisch zu  $\mu$
- Bei metrischer Skala: Varianz:  $\text{Var}\{\text{Median}\{\mathbf{X}\}\} = \frac{\sigma^2}{N} \cdot \frac{\pi}{2}$

Beispiel: **Ordinalskala**  $N = 9$  (ungerade)

Klausurnote	Häufigkeit
1 (sehr gut)	1
2 (gut)	3
<b>3 (befriedigend)</b>	2
4 (ausreichend)	2
5 (mangelhaft)	1
6 (ungenügend)	0

$$\mathbf{X}_0 = \{1, 2, 2, 2, \mathbf{3}, 3, 4, 4, 5\}$$
$$\Rightarrow \text{Median}\{\mathbf{X}\} = 3$$

## 2.2.4 Zentrale Tendenz

**Arithmetisches Mittel** der Beobachtungen  $\mathbf{X} = \{x_1, x_2, \dots, x_N\}$

$$\bar{x} \{ \mathbf{X} \} = \frac{1}{N} \sum_{i=1}^N x_i$$

- Erwartungstreuer Schätzer für  $\mu$  :  $E\{\bar{x}\} = \mu$  (s.u.)
- Schätzer mit dem minimalen mittleren Abweichungsquadrat:

$$\frac{1}{N} \sum_{i=1}^N (x_i - \hat{x})^2 \rightarrow \text{Minimum}$$

Notwendige Bedingung für Optimum:

$$\begin{aligned} \frac{\partial}{\partial \hat{x}} \left( \frac{1}{N} \sum_{i=1}^N (x_i - \hat{x})^2 \right) &= -\frac{2}{N} \sum_{i=1}^N (x_i - \hat{x}) \\ &= -2 \cdot \left( \frac{1}{N} \sum_{i=1}^N x_i - \hat{x} \right) \stackrel{!}{=} 0 \quad \Rightarrow \hat{x} = \bar{x} \end{aligned}$$

## 2.2.4 Zentrale Tendenz

### Arithmetisches Mittel

- Alle Beobachtungen gehen gleich gewichtet in das Ergebnis ein
- Geringe Robustheit (großer Einfluss von Ausreißern)
- Mindestens **intervallskalierte Merkmale** erforderlich

#### Beispiel:

	Gemessene Temperatur	Häufigkeit
<b>Intervallskala</b>	12°C	1
	13°C	0
	14°C	2
	15°C	3
	16°C	1

#### **Arithmetisches Mittel:**

$$\bar{x} \approx 14,42^\circ\text{C}$$

Vergleich: Modus 15°C  
Median 15°C

Fraglich: Beobachtung 12°C  
Ausreißer??

## 2.2.4 Zentrale Tendenz

### Arithmetisches Mittel

- **Erwartungstreue:**

$\bar{x}$  ist erwartungstreuer Schätzer für  $\mu$ , da

$$E\{\bar{x}(X_1, \dots, X_N)\} = E\left\{\frac{1}{N} \sum_{i=1}^N x_i\right\} = \frac{1}{N} \sum_{i=1}^N E\{x_i\} = \frac{1}{N} N \cdot \mu = \mu$$

- **Konsistenz:**

Falls Messungen  $x_i$  paarweise unkorreliert, d.h.  $E\{X_i \cdot X_j\} = E\{X_i\} \cdot E\{X_j\}$ :

$$\text{Var}\{\bar{x}\} = E\{(\bar{x} - \mu)^2\} = \sigma_{\bar{x}}^2 = \frac{\sigma^2}{N} \xrightarrow{N \rightarrow \infty} 0$$

$$\Rightarrow \Pr(|\bar{x}(X_1, \dots, X_N) - \mu| \geq \varepsilon) \xrightarrow{N \rightarrow \infty} 0 \quad \varepsilon > 0$$

- **Wirksamkeit:**

Es gibt keinen Schätzer für  $\mu$  mit kleinerer Varianz, also ist  $\bar{x}$  ein wirksamer Schätzer (ohne Beweis).

## 2.2.4 Zentrale Tendenz

**$\alpha$ -getrimmtes Mittel** der geordneten Beobachtungen  $\mathbf{X}_{( )} = \{x_{(1)}, x_{(2)}, \dots, x_{(N)}\}$

$$x_{t,\alpha} = \frac{1}{N-2k} \sum_{i=k+1}^{N-k} x_{(i)} = \frac{1}{N-2k} (x_{(k+1)} + \dots + x_{(N-k)})$$

mit  $0 < \alpha < \frac{1}{2}$ ,  $k := \lfloor N\alpha \rfloor$  größte ganze Zahl  $\leq N\alpha$

- Mindestens **intervallskalierte Merkmale** erforderlich
- Reduzierte Ausreißerempfindlichkeit:  
Keine Berücksichtigung der  $k$  kleinsten und  $k$  größten Beobachtungen
- Für  $\alpha \rightarrow 0$  : Arithmetisches Mittel
- Für  $\alpha \rightarrow 1/2$  : Median
- Problematisch: Wahl des Parameters  $\alpha$

<u>Beispiel:</u>	Gemessene Temperatur	Häufigkeit	gewählt: $\alpha = \frac{1}{7}$ , $k = 1$ <b><math>\alpha</math>-getrimmtes Mittel:</b> $x_{t,\alpha} = 14,6^\circ\text{C}$ Vergleich: Modus $15^\circ\text{C}$ Median $15^\circ\text{C}$ arithm. M. $14,42^\circ\text{C}$
<b>Intervallskala</b>	<b>12°C</b>	<b>1</b>	
	13°C	0	
	14°C	2	
	15°C	3	
	<b>16°C</b>	<b>1</b>	

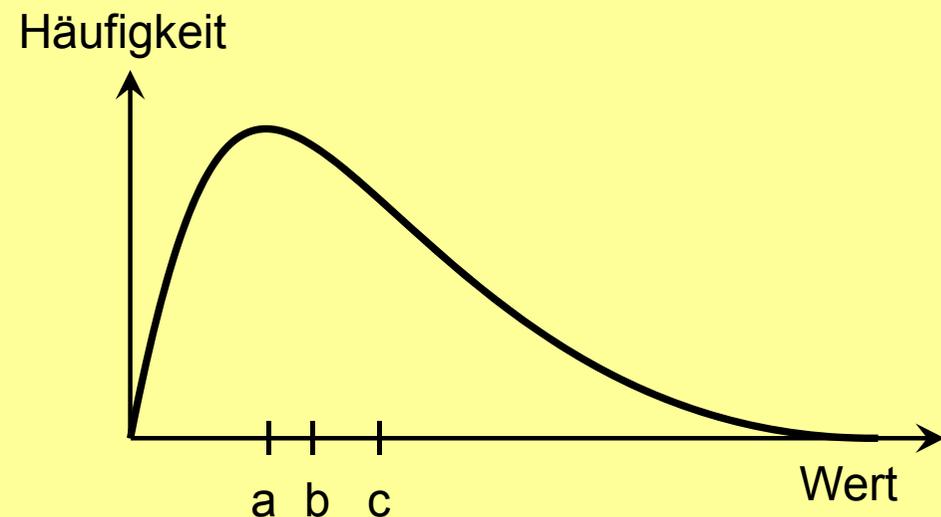
## 2.2.4 Zentrale Tendenz

### Lage der zentralen Tendenz (bei metrischer Skalierung)

- In einer **linkssteilen Verteilung** ist die Reihenfolge (links nach rechts) **Modus – Median – arithmetisches Mittel**.
- In einer **rechtssteilen Verteilung** ist die Reihenfolge **arithmetisches Mittel – Median – Modus**.
- **Symmetrische Verteilung:** Modus, Median und arithmetisches Mittel fallen zusammen.
- Falls Verteilung sehr unsymmetrisch: Aussagekraft von zentraler Tendenz (und ggf. Dispersionsmaß) fraglich

### Beispiel: **Linkssteile Verteilung**

a: Modus  
b: Median  
c: arithm. Mittel



## 2.2.4 Zentrale Tendenz

---

### Kriterien zur Wahl der für die zentrale Tendenz verwendeten Größe

- **Erwartungstreue, Konsistenz, Wirksamkeit** (s.o.)
- **Skalierung**
  - Modus: einzige mögliche Größe bei Nominalskala
  - Median: i.a. beste Wahl bei Ordinalskala
  - Arithmetisches Mittel,  $\alpha$ -getrimmtes Mittel:  
i.a. beste Wahl bei metrisch skalierten Beobachtungen
- **Vorliegende Verteilung**
  - Modus: identifiziert die am häufigsten vorkommende Beobachtung; diese kann aber auch am äußeren Ende der Skala liegen
  - Median,  $\alpha$ -getrimmtes Mittel: Geringe Beeinflussung durch stark vom Rest abweichende Werte (Ausreißer)
  - Arithmetisches Mittel: Berücksichtigung aller Beobachtungen; hohe Empfindlichkeit gegen stark vom Rest abweichende Beobachtungen

## 2.2.2-2.2.4 Beobachtungsunsicherheit – Zusammenfassung

---

- Die Unsicherheit einer Information, die in mehreren Beobachtungen enthalten ist, lässt sich mittels einer **statistischen Analyse** angeben (weitere Spezifikationsmöglichkeit: siehe „Guide to the Expression of Uncertainty in Measurement“, später)
- Grundlage der statistischen Analyse: **Häufigkeitsverteilung**
- Das Ergebnis von Beobachtungen lässt sich vereinfacht mittels einer **zentralen Tendenz** und der **Konzentration/Dispersion** beschreiben
- Je nach Skalenqualität der Beobachtungsgröße kommen **unterschiedliche Größen** in Frage

## 2.3 Schätzung bei mehreren Beobachtungen

### 2.3.1 Schätzung einer statischen Zielgröße

Aufgabe: Schätzung einer **mindestens intervallskalierten Größe**,  
z.B. der **statischen Position  $x$  (eindimensional)** eines Objekts

Verfügbar: **Beobachtungen  $\hat{z}_i$**  für Zielgröße  $x$  (z.B. mittels GPS)  
mit **Messunsicherheiten  $\mu_i, i = 1, \dots, n$**

$$\hat{z}_i = x + \mu_i$$

$\mu_i$ : Realisation einer Zufallsvariablen  $U_i$   
mit

$$E\{U_i\} = 0$$

$$E\{U_i U_j\} = \sigma_i^2 \delta(i - j)$$

d.h. **unkorrelierte Beobachtungen**

## 2.3.1 Schätzung einer statischen Zielgröße

- **Eine Beobachtung:**

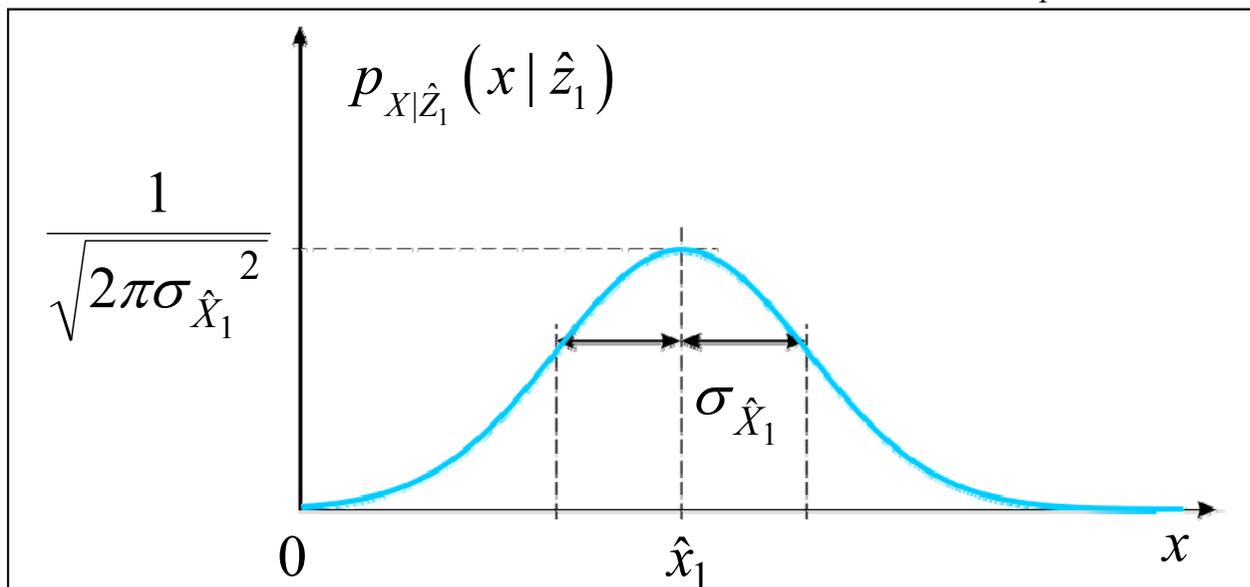
- **Gauß'sches Beobachtungsmodell:**

Normalverteilte Abweichung  $\mu_1$ :  $p(\mu_1) \sim N(0, \sigma_1^2)$

$$p_{\hat{z}_1|X}(\hat{z}_1 | x) = \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{(\hat{z}_1 - x)^2}{2\sigma_1^2}}$$

- Beste Schätzung der Position:  $\hat{x}_1 = \hat{z}_1$

- Varianz der Schätzung:  $E\left\{\left(\hat{X}_1 - E\{\hat{X}_1\}\right)^2\right\} = \sigma_{\hat{X}_1}^2 = \sigma_1^2$



## 2.3.1 Schätzung einer statischen Zielgröße

- Rechtfertigung der Normalverteilungsannahme:

### Zentraler Grenzwertsatz:

$\bar{X}_N$  Summe von unabhängigen, identisch verteilten Zufallsvariablen  $X_i$   
(i.i.d., independent and identically distributed)  
Verteilung der  $X_i$  darf ansonsten beliebig sein

$$\bar{X}_N := \frac{1}{N} \sum_{i=1}^N X_i$$

Standardisierung: 
$$Y_N := \sqrt{N} \frac{\bar{X}_N - \mu}{\sigma}$$

Dann gilt  $E\{Y_N\} = 0, \quad \text{Var}\{Y_N\} = 1$

und

$$\Pr(Y_N \leq y) \xrightarrow{N \rightarrow \infty} \Phi(y)$$

mit  $\Phi(y)$  : Verteilungsfunktion  
der Standardnormalverteilung

d.h.  $Y_N$  ist approximativ  $N(0,1)$  verteilt

## 2.3.1 Schätzung einer statischen Zielgröße

- **$n$  Beobachtungen**  $\hat{z}_1, \dots, \hat{z}_n$ :
  - **Weighted-Least-Squares- (WLS-) Schätzung** für  $n$  Beobachtungen
  - **Gauß-Markov-Beobachtungsmodell**

$$\begin{pmatrix} \hat{z}_1 \\ \vdots \\ \hat{z}_n \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \cdot x + \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} \Leftrightarrow \hat{\mathbf{z}} = \mathbf{C} \cdot x + \boldsymbol{\mu} \quad \mathbf{U} = \begin{pmatrix} U_1 \\ \vdots \\ U_n \end{pmatrix} \quad \mathbb{E}\{\mathbf{U}\} = \mathbf{0}$$

$\mathbf{C}$  : Beobachtungsmatrix, hier z.B.  $\mathbf{C} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$

$$\mathbf{R} = \mathbb{E}\{\mathbf{U}\mathbf{U}^T\} = \begin{pmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_n^2 \end{pmatrix} \quad \begin{array}{l} \text{Kovarianzmatrix} \\ \text{Unkorrelierte Beobachtungen} \end{array}$$

## 2.3.1 Schätzung einer statischen Zielgröße

- **$n$  Beobachtungen**  $\hat{z}_1, \dots, \hat{z}_n$ :

- Beste Schätzung der Position:  $\hat{x}_n = f(\hat{z}_1, \hat{z}_2, \dots, \hat{z}_n) = f(\hat{\mathbf{z}})$

Gewichtete quadratische Abweichungen:

$$S = \sum_{i=1}^n \frac{1}{\sigma_i^2} r_i^2 = \text{Spur} \{ \mathbf{r}^T \mathbf{R}^{-1} \mathbf{r} \} \rightarrow \text{Minimum} \quad \text{mit Residuen } \mathbf{r} = \hat{\mathbf{z}} - \mathbf{C}\mathbf{x}$$

Minimierung von  $S$  (siehe nächste Folie)

führt auf **Weighted-Least-Squares- (WLS-) Schätzung**:

$$\begin{aligned} \hat{x}_n &= (\mathbf{C}^T \mathbf{R}^{-1} \mathbf{C})^{-1} \mathbf{C}^T \mathbf{R}^{-1} \hat{\mathbf{z}} && \text{Gewichtetes arithmetisches Mittel} \\ &= \left( \sum_{i=1}^n \frac{1}{\sigma_i^2} \right)^{-1} \sum_{i=1}^n \frac{\hat{z}_i}{\sigma_i^2} && \text{(Aitken 1935)} \end{aligned}$$

- Varianz der Schätzung (siehe übernächste Folie):

$$\begin{aligned} \sigma_{\hat{X}_n}^2 &= \text{E} \left\{ \left( \hat{X}_n - \text{E} \{ \hat{X}_n \} \right)^2 \right\} = (\mathbf{C}^T \mathbf{R}^{-1} \mathbf{C})^{-1} && \text{Harmonisches Mittel der} \\ &= \left( \sum_{i=1}^n \frac{1}{\sigma_i^2} \right)^{-1} && \text{Varianzen} \end{aligned}$$

## 2.3.1 Schätzung einer statischen Zielgröße

- **$n$  Beobachtungen**  $\hat{z}_1, \dots, \hat{z}_n$ :

- Minimierung von  $S$  (hier allgemein für vektorielles  $\mathbf{x}_n$ ):

$$\begin{aligned} \mathbf{r}^T \mathbf{R}^{-1} \mathbf{r} &= \left( \hat{\mathbf{z}}^T - \mathbf{x}_n^T \mathbf{C}^T \right) \mathbf{R}^{-1} \left( \hat{\mathbf{z}} - \mathbf{C} \mathbf{x}_n \right) \\ &= \hat{\mathbf{z}}^T \mathbf{R}^{-1} \hat{\mathbf{z}} - \hat{\mathbf{z}}^T \mathbf{R}^{-1} \mathbf{C} \mathbf{x}_n - \mathbf{x}_n^T \mathbf{C}^T \mathbf{R}^{-1} \hat{\mathbf{z}} + \mathbf{x}_n^T \mathbf{C}^T \mathbf{R}^{-1} \mathbf{C} \mathbf{x}_n \end{aligned}$$

$$\begin{aligned} \frac{\partial \text{Spur}(\mathbf{r}^T \mathbf{R}^{-1} \mathbf{r})}{\partial \mathbf{x}_n} &= -\left( \hat{\mathbf{z}}^T \mathbf{R}^{-1} \mathbf{C} \right)^T - \mathbf{C}^T \mathbf{R}^{-1} \hat{\mathbf{z}} + \left( \mathbf{C}^T \mathbf{R}^{-1} \mathbf{C} + \left( \mathbf{C}^T \mathbf{R}^{-1} \mathbf{C} \right)^T \right) \mathbf{x}_n \\ &= -2\mathbf{C}^T \mathbf{R}^{-1} \hat{\mathbf{z}} + 2\mathbf{C}^T \mathbf{R}^{-1} \mathbf{C} \mathbf{x}_n \stackrel{!}{=} 0 \end{aligned}$$

$$\Rightarrow \mathbf{C}^T \mathbf{R}^{-1} \mathbf{C} \hat{\mathbf{x}}_n = \mathbf{C}^T \mathbf{R}^{-1} \hat{\mathbf{z}} \quad \left( \mathbf{R}^{-1} \right)^T = \mathbf{R}^{-1} \quad (\mathbf{R} \text{ symmetrisch})$$

$$\hat{\mathbf{x}}_n = \left( \mathbf{C}^T \mathbf{R}^{-1} \mathbf{C} \right)^{-1} \mathbf{C}^T \mathbf{R}^{-1} \hat{\mathbf{z}}$$

Einige Regeln der Matrizenrechnung:  $(\mathbf{ABC})^T = \mathbf{C}^T \mathbf{B}^T \mathbf{A}^T$ ;  $(\mathbf{A}^{-1})^T = (\mathbf{A}^T)^{-1}$

$$\frac{\partial \text{Spur}(\mathbf{Ax})}{\partial \mathbf{x}} = \mathbf{A}^T; \quad \frac{\partial \text{Spur}(\mathbf{x}^T \mathbf{A})}{\partial \mathbf{x}} = \mathbf{A}; \quad \frac{\partial \text{Spur}(\mathbf{x}^T \mathbf{Ax})}{\partial \mathbf{x}} = (\mathbf{A} + \mathbf{A}^T) \mathbf{x}$$

## 2.3.1 Schätzung einer statischen Zielgröße

- **$n$  Beobachtungen**  $\hat{z}_1, \dots, \hat{z}_n$ :
  - Varianz der Schätzung  $\hat{x}_n$  (hier allgemein für vektorielles  $x_n$ ):

$$\hat{X}_n = (C^T R^{-1} C)^{-1} C^T R^{-1} \hat{Z} = (C^T R^{-1} C)^{-1} C^T R^{-1} (C\hat{x} + U)$$

$$E\{\hat{X}_n\} = (C^T R^{-1} C)^{-1} C^T R^{-1} C\hat{x}$$

$$\begin{aligned} E\left\{(\hat{X}_n - E\{\hat{X}_n\})(\hat{X}_n - E\{\hat{X}_n\})^T\right\} \\ = E\left\{(C^T R^{-1} C)^{-1} C^T R^{-1} \underbrace{UU^T}_{E\{UU^T\}=R} (R^{-1})^T C \left((C^T R^{-1} C)^{-1}\right)^T\right\} = (C^T R^{-1} C)^{-1} \\ \underbrace{\mathbf{I}}_{(C^T R^{-1} C)^T} \end{aligned}$$

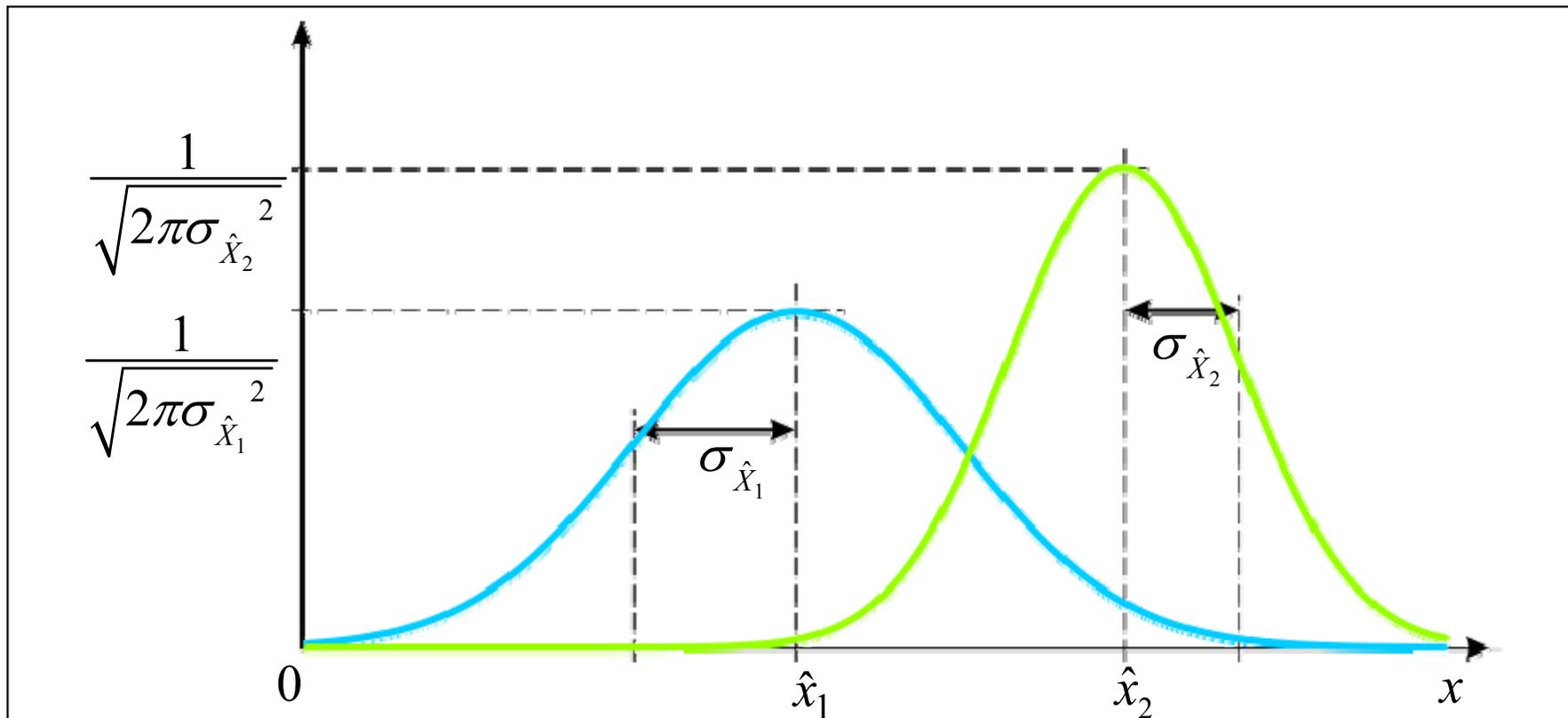
## 2.3.1 Schätzung einer statischen Zielgröße

- Beispiel  $n = 2$ :

Beobachtungen:

$$p_{\hat{z}_1|X}(\hat{z}_1 | x) = \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{(\hat{z}_1 - x)^2}{2\sigma_1^2}}, \quad p_{\hat{z}_2|X}(\hat{z}_2 | x) = \frac{1}{\sqrt{2\pi\sigma_2^2}} e^{-\frac{(\hat{z}_2 - x)^2}{2\sigma_2^2}}$$

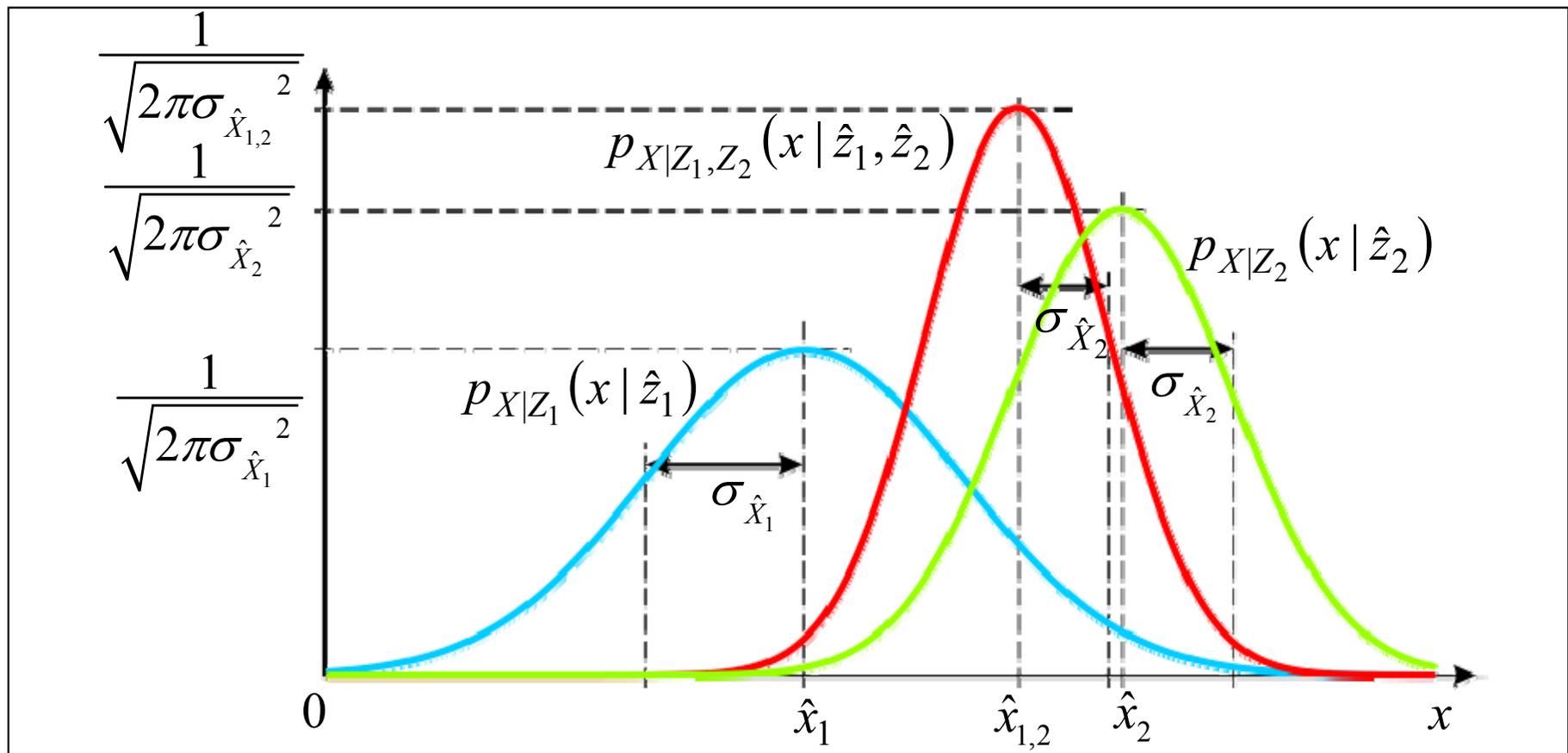
Schätzungen aus Einzelbeobachtung:  $\hat{x}_1 = \hat{z}_1, \quad \sigma_{\hat{x}_1}^2 = \sigma_1^2$   
 $\hat{x}_2 = \hat{z}_2, \quad \sigma_{\hat{x}_2}^2 = \sigma_2^2$



## 2.3.1 Schätzung einer statischen Zielgröße

- Beispiel  $n = 2$ :

Schätzung durch Fusion: 
$$p_{X|\hat{z}_1, \hat{z}_2}(x | \hat{z}_1, \hat{z}_2) = \frac{1}{\sqrt{2\pi\sigma_{\hat{x}_{1,2}}^2}} e^{-\frac{(x-\hat{x}_{1,2})^2}{2\sigma_{\hat{x}_{1,2}}^2}}$$



Vorsicht: Auf dieser Folie bezieht sich der Index 1,2 von  $x, X$  etc. auf die Fusion der beiden Beobachtungen, der Index 2 allein bezieht sich auf Beobachtung 2. Im Folgenden wird mit dem Index  $n$  meist das Fusionsresultat unter Verwendung der Beobachtungen bis einschließlich Beobachtung  $n$  bezeichnet.

## 2.3.1 Schätzung einer statischen Zielgröße

- Beispiel  $n = 2$ :

Fusion durch Multiplikation der Verteilungen  
bzgl. der Einzelbeobachtungen:

$$p_{X|\hat{Z}_1, \hat{Z}_2}(x | \hat{z}_1, \hat{z}_2) \propto p_{\hat{Z}_1|X}(\hat{z}_1 | x) \cdot p_{\hat{Z}_2|X}(\hat{z}_2 | x) = \frac{1}{\sqrt{2\pi\sigma_{\hat{X}_{1,2}}^2}} e^{-\frac{(x-\hat{x}_{1,2})^2}{2\sigma_{\hat{X}_{1,2}}^2}}$$

mit  $\hat{x}_{1,2} = \left( \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \right)^{-1} \cdot \left( \frac{\hat{z}_1}{\sigma_1^2} + \frac{\hat{z}_2}{\sigma_2^2} \right)$

$$\sigma_{\hat{X}_{1,2}}^2 = \left( \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \right)^{-1}$$

Weighted-Least-Squares-  
(WLS-) Schätzung

## 2.3.1 Schätzung einer statischen Zielgröße

- **$n$  Beobachtungen**  $\hat{z}_1, \dots, \hat{z}_n$ :

- Sonderfall: Identische Beobachtungsunsicherheiten  $\sigma_i = \sigma$ :

$$\hat{x}_n = \frac{1}{n} \sum_{i=1}^n \hat{z}_i$$

Mittlere Abweichung der Beobachtungen:  $s_{\hat{z}} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (\hat{z}_i - \hat{x}_n)^2}$

$$\mathbb{E} \left\{ \left( \hat{X}_n - \mathbb{E} \{ \hat{X}_n \} \right)^2 \right\} = \sigma_{\hat{X}_n}^2 = \frac{1}{n} \sigma^2 = \frac{1}{n} \mathbb{E} \left\{ s_{\hat{z}}^2 \right\}$$

## 2.3.1 Schätzung einer statischen Zielgröße

### Beispiel:

- Ortsfeste Kamera, starkes additives Gauß'sches Rauschen:  $\sigma_{\text{Sensor}} = 20$



Rauschfreies Bild



Stark verrauschtes Bild

## 2.3.1 Schätzung einer statischen Zielgröße

- Rauschunterdrückung durch mehrfache Aufnahme und Mittelung
- Annahme: Unkorreliertes Rauschen



Bildserie  $N = 4$ ,  $\sigma_{\text{sensor}} = 20$



Fusionsergebnis  $\sigma_{\text{Ergebnis}} = \frac{\sigma_{\text{Sensor}}}{\sqrt{N}} = 10$

## 2.3.1 Schätzung einer statischen Zielgröße

- Modellierung der Bilderfassung:

Beobachtete Bildintensität:  $g_i(\mathbf{x}) = d(\mathbf{x}) + r_i(\mathbf{x}), i = 1, \dots, N$

Realisierung einer Zufallsgröße:  $G_i(\mathbf{x}) = d(\mathbf{x}) + R_i(\mathbf{x}), i = 1, \dots, N$

- Nutzinformation:  $d(\mathbf{x})$  deterministisches Signal  
repräsentiert Information über Szene

- Sensorrauschen (z.B. thermisch):  $r_i(\mathbf{x}), E\{R_i(\mathbf{x})\} = 0,$

$$E\{R_i(\mathbf{x})R_j(\mathbf{x})\} = \delta_i^j \sigma_R^2 \text{ unkorreliert}$$

- Statistik einer Aufnahme:  $E\{G_i(\mathbf{x})\} = E\{d(\mathbf{x}) + R_i(\mathbf{x})\} = d(\mathbf{x}),$

$$\begin{aligned} \sigma_G^2 &= \text{Var}\{G_i(\mathbf{x})\} = E\left\{\left(G_i(\mathbf{x}) - E\{G_i(\mathbf{x})\}\right)^2\right\} \\ &= E\left\{\left(R_i(\mathbf{x})\right)^2\right\} = \sigma_R^2 \end{aligned}$$

- Mittelwertbildung über  $N$  Aufnahmen:

$$m(\mathbf{x}) = \frac{1}{N} \sum_{i=1}^N g_i(\mathbf{x}) = d(\mathbf{x}) + \frac{1}{N} \sum_{i=1}^N r_i(\mathbf{x})$$

## 2.3.1 Schätzung einer statischen Zielgröße

- Statistik des Mittelwerts  $m(\mathbf{x})$  von  $N$  Aufnahmen:

$$E\{M(\mathbf{x})\} = d(\mathbf{x}) + \frac{1}{N} \sum_{i=1}^N E\{R_i(\mathbf{x})\} = d(\mathbf{x}) \text{ (wie zuvor)}$$

$$\begin{aligned} \sigma_M^2 &= \text{Var}\{M(\mathbf{x})\} = E\left\{\left(M(\mathbf{x}) - E\{M(\mathbf{x})\}\right)^2\right\} \\ &= E\left\{\left(\frac{1}{N} \sum_{i=1}^N R_i(\mathbf{x})\right)^2\right\} = \frac{1}{N^2} E\left\{\sum_{i=1}^N R_i^2(\mathbf{x}) + 2 \underbrace{\sum_{i=1}^N \sum_{\substack{j=1 \\ i \neq j}}^N R_i(\mathbf{x}) R_j(\mathbf{x})}_{=0}\right\} \\ &= \frac{1}{N} \sigma_R^2 \end{aligned}$$

→ Varianz des Mittelwerts  $m(\mathbf{x})$  um Faktor  $N$  reduziert  
im Vergleich zur Einzelaufnahme  $g_i(\mathbf{x})$   
siehe oben!

## 2.3.1 Schätzung einer statischen Zielgröße

- **Laufende Nachführung** der Schätzung nach jeder Beobachtung  $\hat{z}_n$

hier für

$$\hat{z} = \mathbf{C} \cdot x + \boldsymbol{\mu}, \quad \mathbf{C} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

- Ausgangspunkt: WLS-Schätzung:

$$\begin{aligned} \hat{x}_n &= \left( \sum_{i=1}^n \frac{1}{\sigma_i^2} \right)^{-1} \sum_{i=1}^n \frac{\hat{z}_i}{\sigma_i^2} = \frac{1}{\sum_{i=1}^{n-1} \frac{1}{\sigma_i^2} + \frac{1}{\sigma_n^2}} \left( \sum_{i=1}^{n-1} \frac{\hat{z}_i}{\sigma_i^2} + \frac{\hat{z}_n}{\sigma_n^2} \right) \\ &= \frac{\sum_{i=1}^{n-1} \frac{1}{\sigma_i^2}}{\sum_{i=1}^{n-1} \frac{1}{\sigma_i^2} + \frac{1}{\sigma_n^2}} \left( \underbrace{\left( \sum_{i=1}^{n-1} \frac{1}{\sigma_i^2} \right)^{-1} \sum_{i=1}^{n-1} \frac{\hat{z}_i}{\sigma_i^2}}_{\hat{x}_{n-1}} + \left( \sum_{i=1}^{n-1} \frac{1}{\sigma_i^2} \right)^{-1} \frac{\hat{z}_n}{\sigma_n^2} \right) \\ &= \frac{\sigma_{\hat{x}_{n-1}}^{-2}}{\sigma_{\hat{x}_{n-1}}^{-2} + \sigma_n^{-2}} \left( \hat{x}_{n-1} + \frac{\sigma_{\hat{x}_{n-1}}^2}{\sigma_n^2} \hat{z}_n \right) = \underbrace{\frac{\sigma_{\hat{x}_{n-1}}^2}{\sigma_{\hat{x}_{n-1}}^2 + \sigma_n^2}}_{:=K_n} \left( \frac{\sigma_n^2}{\sigma_{\hat{x}_{n-1}}^2} \hat{x}_{n-1} + \hat{z}_n \right) \\ &= K_n \left( \frac{\sigma_n^2}{\sigma_{\hat{x}_{n-1}}^2} \hat{x}_{n-1} + \hat{z}_n \right) = (1 - K_n) \hat{x}_{n-1} + K_n \hat{z}_n = \hat{x}_{n-1} + K_n (\hat{z}_n - \hat{x}_{n-1}) \end{aligned}$$

## 2.3.1 Schätzung einer statischen Zielgröße

- **Laufende Nachführung** der Schätzung nach jeder Beobachtung  $\hat{z}_n$

- Schätzwert:

$$\hat{x}_n = \hat{x}_{n-1} + K_n (\hat{z}_n - \hat{x}_{n-1}), \quad K_n = \frac{\sigma_{\hat{X}_{n-1}}^2}{\sigma_{\hat{X}_{n-1}}^2 + \sigma_n^2}, \quad \hat{x}_1 = \hat{z}_1$$

- **Varianz der Schätzung:**

$$\frac{1}{\sigma_{\hat{X}_n}^2} = \sum_{i=1}^n \frac{1}{\sigma_i^2} = \underbrace{\sum_{i=1}^{n-1} \frac{1}{\sigma_i^2}}_{=\sigma_{\hat{X}_{n-1}}^{-2}} + \frac{1}{\sigma_n^2} = \left( \left( \sigma_{\hat{X}_{n-1}}^2 \right)^{-1} + \left( \sigma_n^2 \right)^{-1} \right)^{-1}, \quad \sigma_{\hat{X}_1}^2 = \sigma_1^2$$

- Vorteil: Konstanter Rechenaufwand für  $n > 1$
- Jedoch beschränkt auf **statische Messgröße**

## 2.3.2 Schätzung einer mehrdimensionalen statischen Zielgröße

Aufgabe: Schätzung einer mehrdimensionalen, mindestens intervallskalierten Größe,  
z.B. der statischen 2D-Position  $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$  eines Objekts

Verfügbar: Beobachtungen  $\hat{\mathbf{z}}_i = \begin{pmatrix} \hat{z}_{1,i} \\ \hat{z}_{2,i} \end{pmatrix}$  (z.B. durch GPS-Gerät)  
 $\hat{\mathbf{z}}_i = \mathbf{x} + \boldsymbol{\mu}_i$

mit Beobachtungsunsicherheiten  $\boldsymbol{\mu}_i = \begin{pmatrix} \mu_{1,i} \\ \mu_{2,i} \end{pmatrix}$

$\boldsymbol{\mu}_i$ : Realisation stochastischer Prozesse  $\mathbf{U}_i = \begin{pmatrix} U_{1,i} \\ U_{2,i} \end{pmatrix}$   
mit  $E\{\mathbf{U}_i\} = \mathbf{0}$

$$E\{\mathbf{U}_i \mathbf{U}_j^T\} = \delta(i - j) \cdot \mathbf{R}_i$$

Kovarianzmatrix beschreibt Unsicherheit der Sensordaten:

$$\mathbf{R}_i = \begin{pmatrix} \text{Var}\{U_{1,i}\} & \text{Cov}\{U_{1,i}, U_{2,i}\} \\ \text{Cov}\{U_{1,i}, U_{2,i}\} & \text{Var}\{U_{2,i}\} \end{pmatrix}$$

Unsicherheitsellipsen, aufgespannt durch die Eigenvektoren von  $\mathbf{R}_i$

## 2.3.2 Schätzung einer mehrdimensionalen statischen Zielgröße

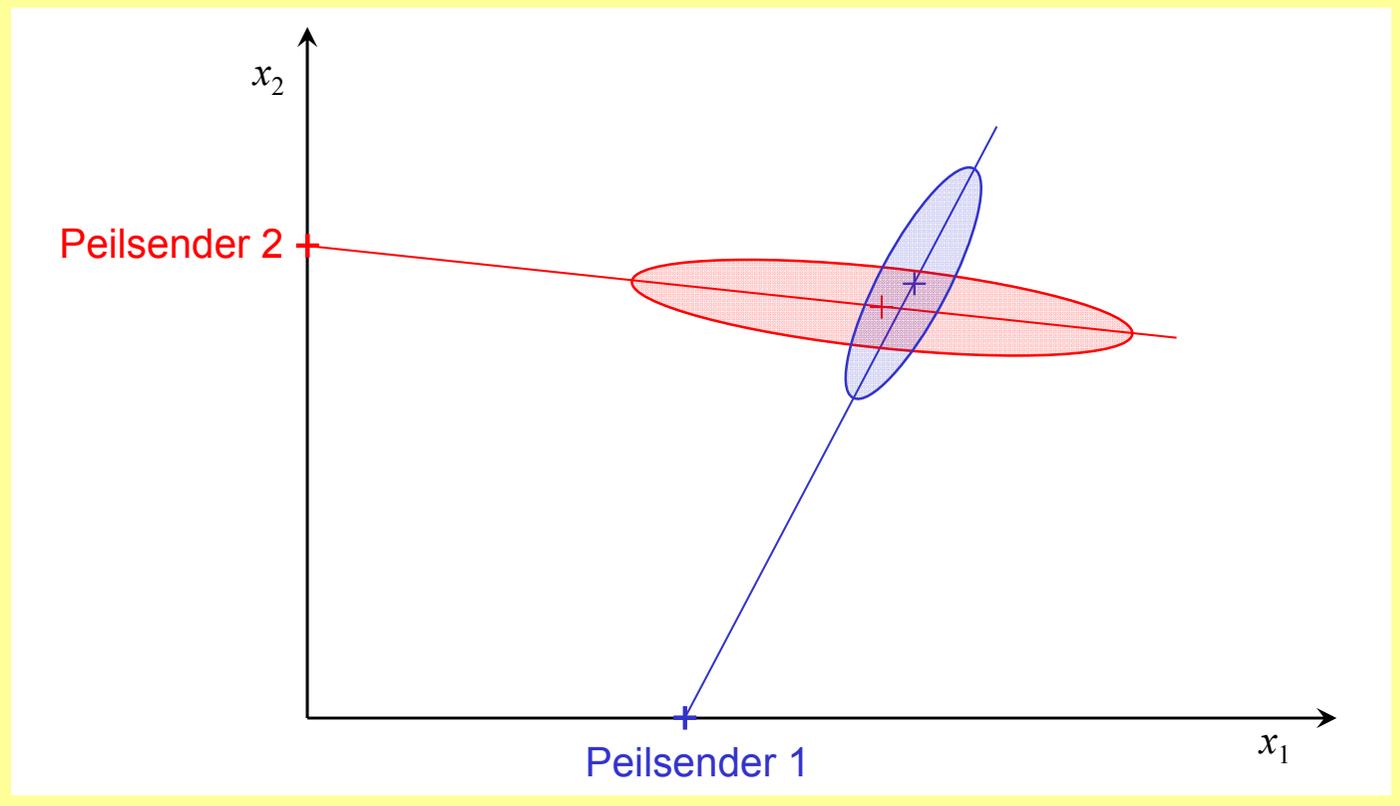
- Beste Schätzung der 2D-Position (gewichtetes arithmetisches Mittel) bei  $n$  Beobachtungen:

$$\hat{\mathbf{x}}_n = \mathbf{P}_n \sum_{i=1}^n \mathbf{R}_i^{-1} \hat{\mathbf{z}}_i$$

- Unsicherheit der Schätzung (harmonisches Mittel):  $\mathbf{P}_n := \left( \sum_{i=1}^n \mathbf{R}_i^{-1} \right)^{-1}$

Beispiel:

Radarpeilungen



## 2.3.3 Kalman-Filter

---

Voraussetzung für bisherige Schätzverfahren:

- Statische Zielgröße

Jetzt:

- Zeitlich veränderliche Zielgröße
- Laufende Nachführung des Schätzwerts
- Verwendung von verfügbarem Vorwissen (für die Initialisierung)

### Kalman-Filter:

- Spezielle Realisierung eines LS-Schätzers
- Beschränkt auf lineare dynamische Systeme
- Zustandsbeobachter:  
Zielgrößen sind nicht direkt beobachtbare Zustände

Rudolf Emil Kálmán (geb. 1930 in Ungarn)

- Miterfinder des Kalman-Filters ca. 1960
- Anfangs eher skeptische Akzeptanz in der Wissenschaft
- Einsatz im Apollo-Programm

## 2.3.3 Kalman-Filter – Dynamisches Modell

### Schätzung einer dynamischen Größe

Bsp: Objekt aus dem vorigen Beispiel bewegt sich.

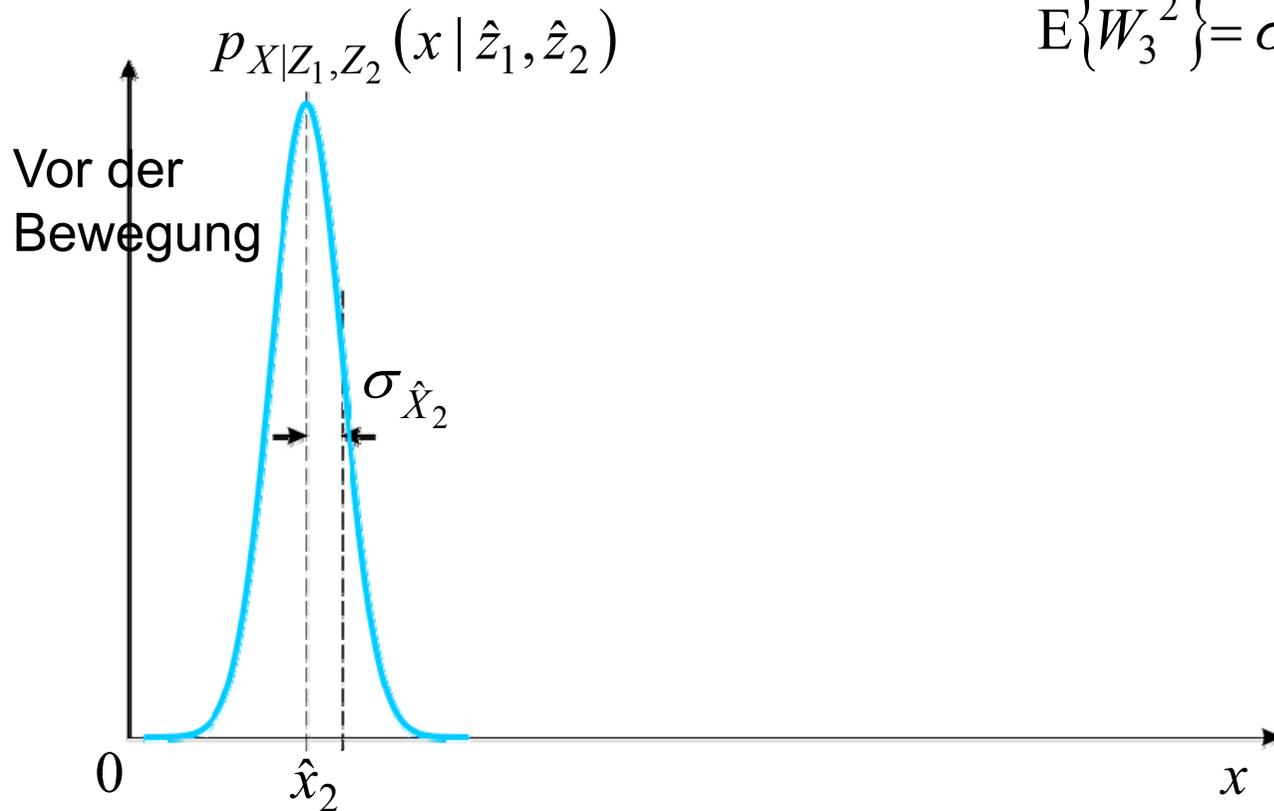
Dynamisches Modell: Konstante Geschwindigkeit  $u_3$

für die Dauer  $\Delta t$

Unsicherheit der Geschwindigkeit  $u_3$ :  $w_3$

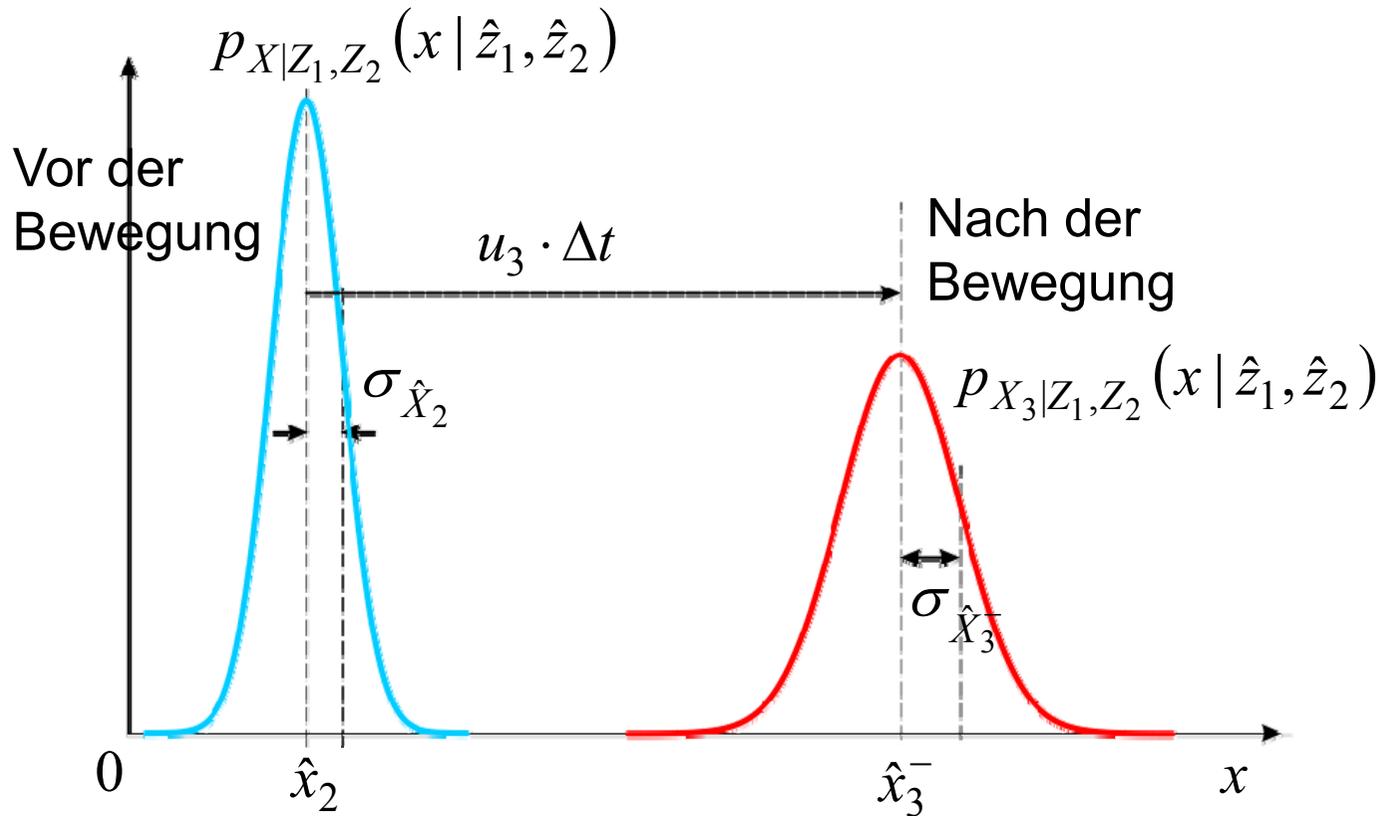
$$E\{W_3\} = 0$$

$$E\{W_3^2\} = \sigma_{W_3}^2$$



## 2.3.3 Kalman-Filter – Prädiktionsschritt

Propagation der Verteilung (**Prädiktion**):  $\hat{x}_3^- = \hat{x}_2 + u_3 \Delta t$



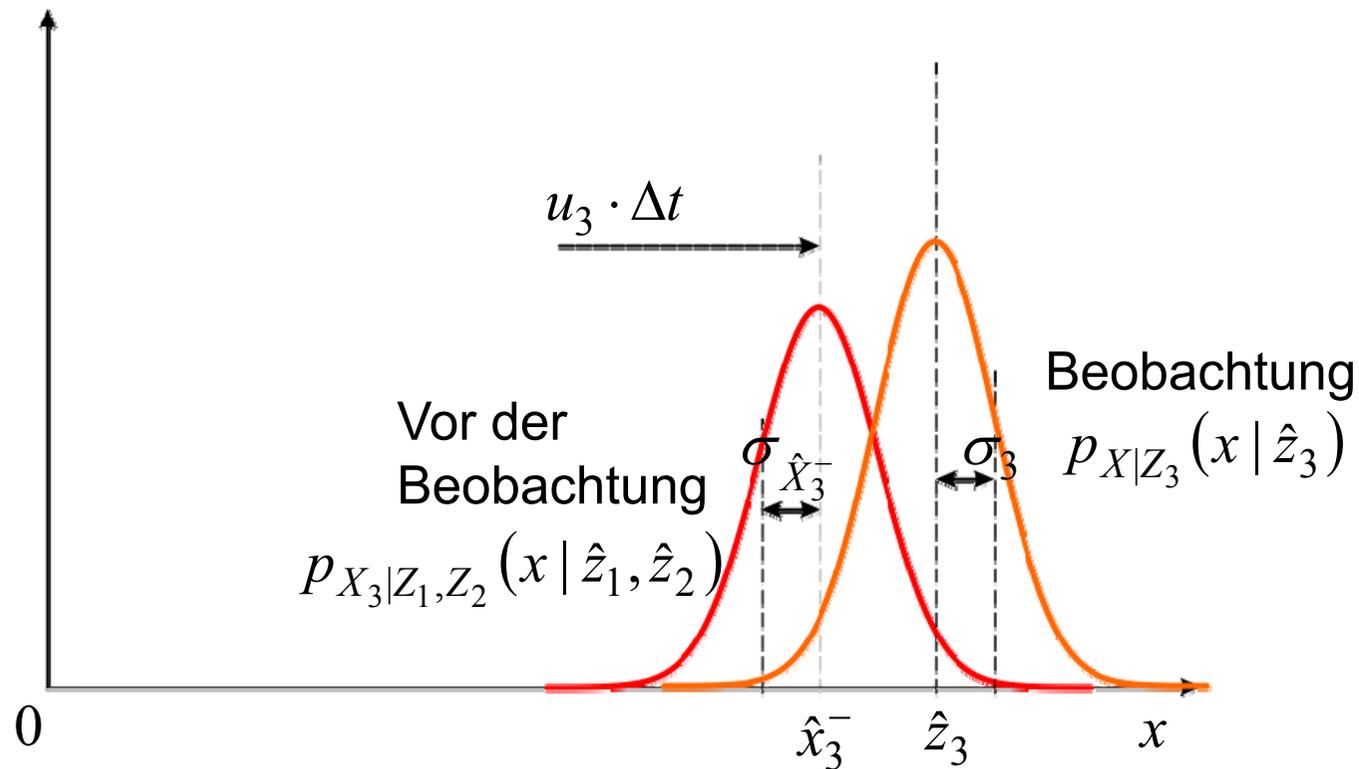
Durch Unsicherheit der Geschwindigkeit steigt Varianz der Prädiktion:

$$\sigma_{\hat{x}_3^-}^2 = E\left\{\left(\hat{x}_3^- - E\{\hat{x}_3^-\}\right)^2\right\} = \sigma_{\hat{x}_2}^2 + \sigma_{w_3}^2 (\Delta t)^2 > \sigma_{\hat{x}_2}^2$$

(unkorrelierte Unsicherheiten von Beobachtung und Bewegung)

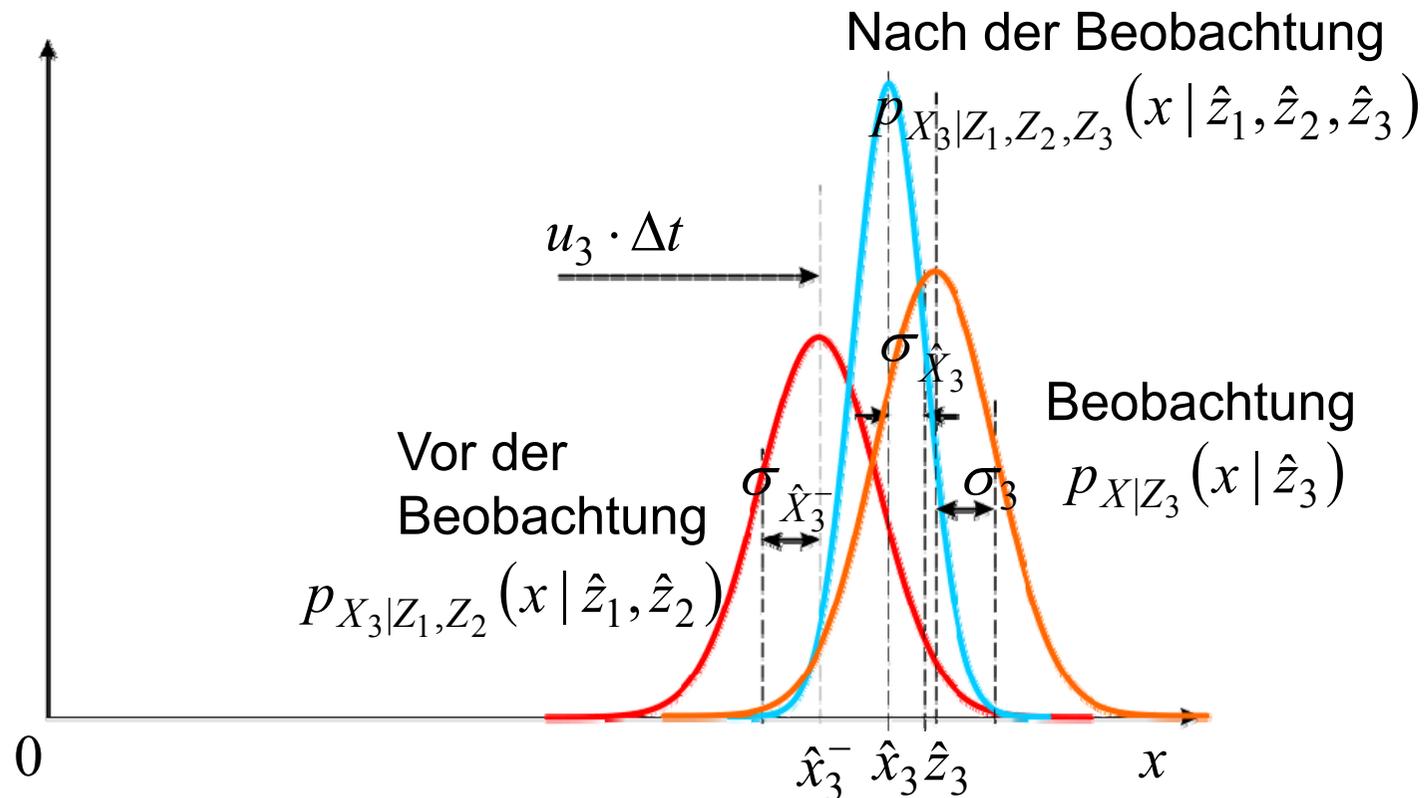
## 2.3.3 Kalman-Filter – Innovationsschritt

Neue Beobachtung nach der Bewegung:  $\hat{z}_3$  mit Unsicherheit  $\sigma_3$   
(**Innovation**)



## 2.3.3 Kalman-Filter – Innovationsschritt

Neue Beobachtung nach der Bewegung:  $\hat{z}_3$  mit Unsicherheit  $\sigma_3$



Verbesserung der Verteilung im Innovationsschritt:  
Durch die zusätzliche Beobachtung sinkt die Varianz.

## 2.3.3 Kalman-Filter – Zustandsraumbeschreibung

Zustandsraumbeschreibung von  
**linearen verschiebungsinvarianten (zeitdiskreten) Systemen:**  
(Linear Shift Invariant – LSI)

$$\mathbf{x}(n+1) = \mathbf{A}\mathbf{x}(n) + \mathbf{B}\mathbf{u}(n) + \mathbf{G}\mathbf{v}(n) \quad \text{Zustandsgleichung}$$

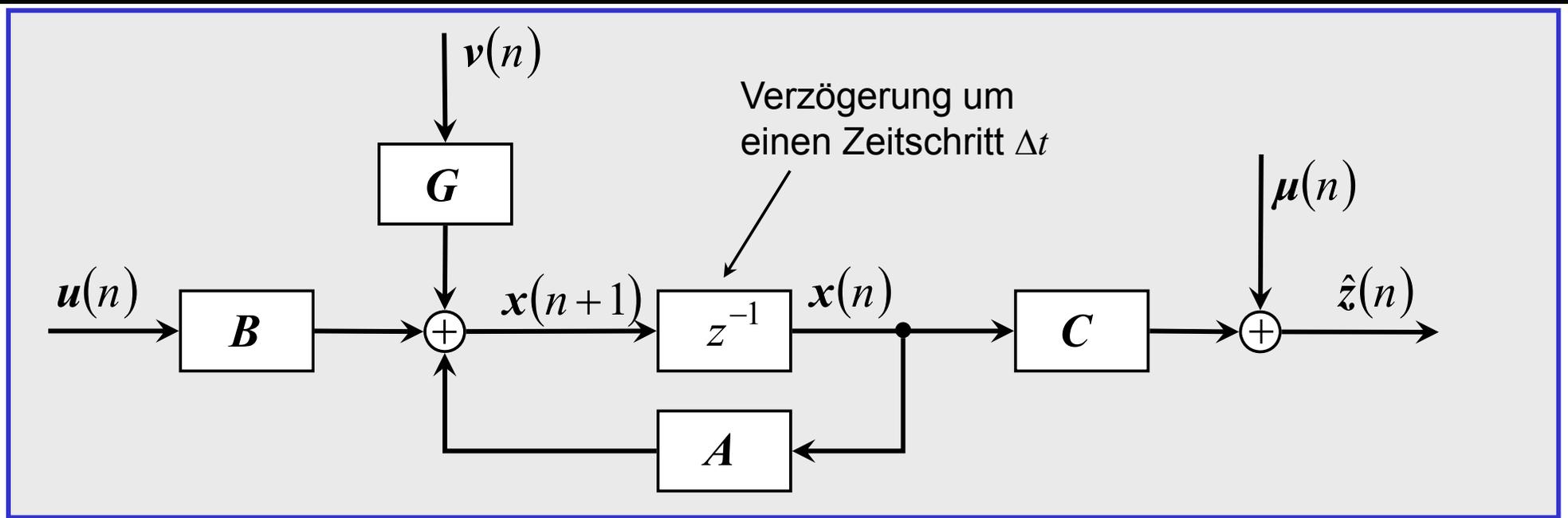
$$\hat{\mathbf{z}}(n) = \mathbf{C}\mathbf{x}(n) + \boldsymbol{\mu}(n) \quad \text{Beobachtungsgleichung}$$

$\mathbf{u}(n)$	deterministisches Eingangsvektor (Dim. $q \times 1$ )	$\mathbf{A}$	Systemmatrix ( $p \times p$ )
$\hat{\mathbf{z}}(n)$	Ausgangsvektor = Beobachtungsvektor ( $r \times 1$ )	$\mathbf{B}$	Steuermatrix ( $p \times q$ )
$\mathbf{x}(n)$	Zustandsvektor ( $p \times 1$ )	$\mathbf{C}$	Beobachtungsmatrix ( $r \times p$ )
$\mathbf{v}(n)$	System-Rauschvektor ( $s \times 1$ )	$\mathbf{G}$	Rauschmatrix ( $p \times s$ )
$\boldsymbol{\mu}(n)$	Beobachtungs-Rauschvektor ( $r \times 1$ )		

Anfangszustand vor der Beobachtung bei  $n = 0$ :

$$\mathbb{E}\{\mathbf{X}(0)\} = \mathbf{x}_0; \quad \text{Cov}\{\mathbf{X}(0)\} = \mathbb{E}\{\mathbf{X}(0)\mathbf{X}^T(0)\} = \mathbf{P}_0$$

## 2.3.3 Kalman-Filter – Blockschaltbild, Störungen



Zustand bei  $n+1$  hängt nur vom Zustand bei  $n$  und von  $v(n)$  ab  
 Beobachtung bei  $n$  hängt nur vom Zustand bei  $n$  und von  $\mu(n)$  ab

Stochastische Prozesse:  $V(i)$ ,  $U(i)$  mit den Realisierungen  $v(i)$ ,  $\mu(i)$   
 Gauß-Markov-Modell, modelliert durch

$$\begin{aligned} E\{V(i)\} &= \mathbf{0}; & E\{U(i)\} &= \mathbf{0}; \\ E\{V(i)V^T(j)\} &= \mathbf{Q}\delta(i-j); & E\{U(i)U^T(j)\} &= \mathbf{R}\delta(i-j) \end{aligned}$$

Stochastische Prozesse sind paarweise unkorreliert:  $E\{U(i)V^T(j)\} = \mathbf{0}$

## 2.3.3 Kalman-Filter – Schätzkriterium

Ziel: LS-Schätzung des jeweils aktuellen Zustands  $\mathbf{x}(n)$  bei gegebener Eingangs- und Ausgangsfolge  $\mathbf{u}(1), \dots, \mathbf{u}(n-1), \hat{\mathbf{z}}(1), \dots, \hat{\mathbf{z}}(n-1)$  sowie gegebenem Erwartungswert  $\mathbf{x}_0$  und gegebener Kovarianz  $\mathbf{P}_0$  des Anfangszustands

Schätzkriterium:

$$J(\hat{\mathbf{x}}(n)) = \underbrace{\sum_{k=1}^n \hat{\mathbf{v}}^T(k) \mathbf{Q}^{-1} \hat{\mathbf{v}}(k)}_{\text{LS-Systemfehler}} + \underbrace{\sum_{k=1}^n \hat{\boldsymbol{\mu}}^T(k) \mathbf{R}^{-1} \hat{\boldsymbol{\mu}}(k)}_{\text{LS-Beobachtungsfehler}} + \underbrace{(\hat{\mathbf{x}}(0) - \hat{\mathbf{x}}_0^-)^T \mathbf{P}_0^{-1} (\hat{\mathbf{x}}(0) - \hat{\mathbf{x}}_0^-)}_{\text{LS-Anfangswertfehler}}$$

→ Minimum

Ergibt linearen erwartungstreuen Schätzer kleinster Varianz.

## 2.3.3 Kalman-Filter – Optimaler Beobachter

---

Das LS-Schätzkriterium lässt sich in zwei Schritte zerlegen:

Prädiktionsschritt: Beste Schätzung des Systemzustands zum Zeitpunkt  $n+1$  ohne neue Beobachtung aus  $\mathbf{u}(1), \dots, \mathbf{u}(n), \hat{\mathbf{z}}(1), \dots, \hat{\mathbf{z}}(n)$  :

$$\hat{\mathbf{x}}^-(n+1) = \mathbf{A}\hat{\mathbf{x}}(n) + \mathbf{B}\mathbf{u}(n)$$

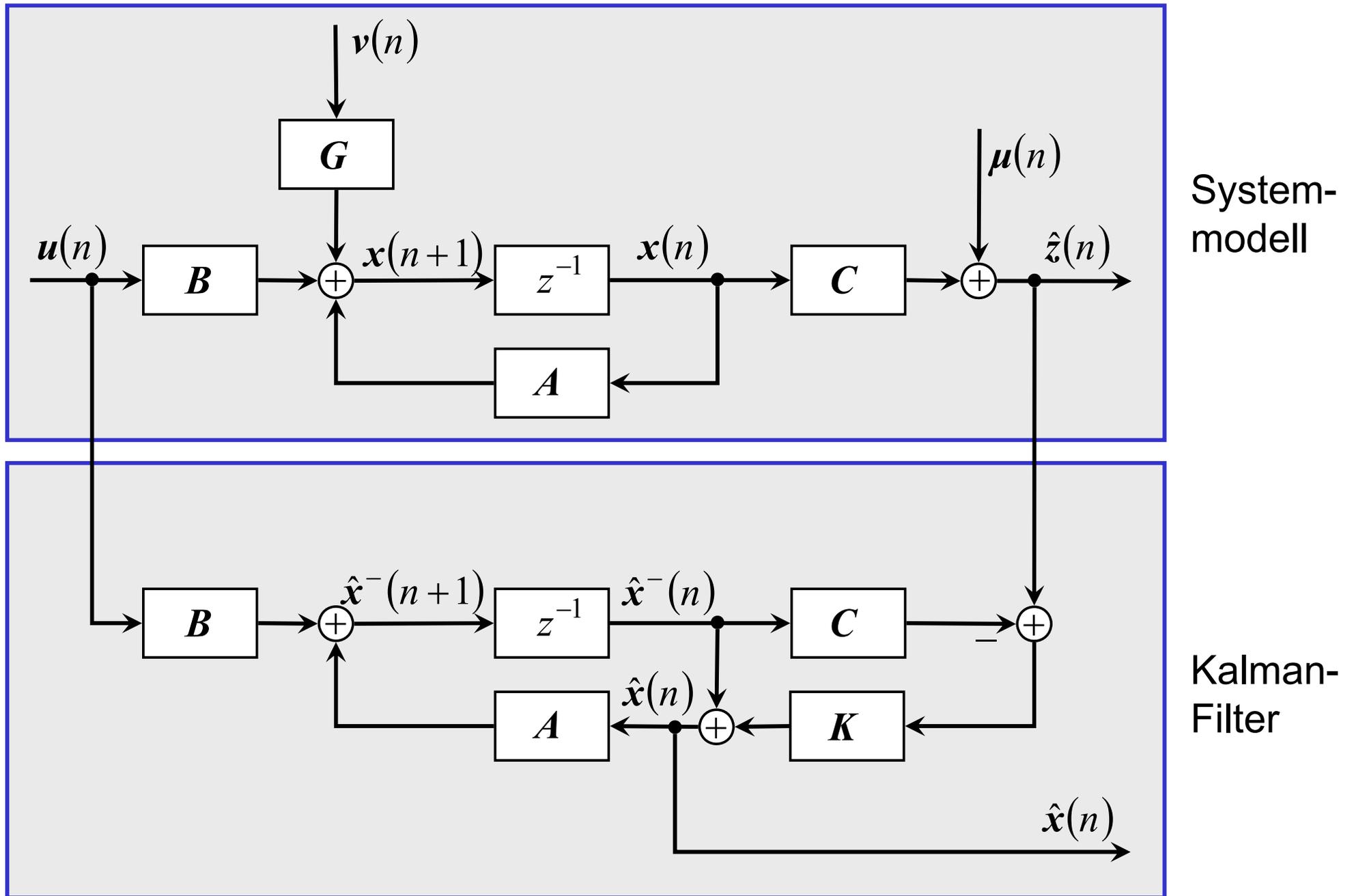
Innovationsschritt: Setze  $n := n+1$ , Neue Beobachtung:  $\hat{\mathbf{z}}(n)$   
Beste Schätzung des Systemzustands unter Berücksichtigung der neuen Beobachtung:

$$\hat{\mathbf{x}}(n) = \hat{\mathbf{x}}^-(n) + \mathbf{K}(n)(\hat{\mathbf{z}}(n) - \mathbf{C}\hat{\mathbf{x}}^-(n))$$

$\mathbf{K}(n)$ : Verstärkungsmatrix (Kalman gain)

Als Anfangsschätzung wird gewählt:  $\hat{\mathbf{x}}(0) = \mathbf{x}_0$

## 2.3.3 Kalman-Filter – Blockschaltbild



## 2.3.3 Kalman-Filter – Prädiktionsschritt

Prädiktion: Beste Schätzung des Systemzustands zum Zeitpunkt  $n+1$  ohne neue Beobachtung aus  $\mathbf{u}(1), \dots, \mathbf{u}(n)$ ,  $\hat{\mathbf{x}}(1), \dots, \hat{\mathbf{x}}(n)$  :

$$\hat{\mathbf{x}}^-(n+1) = \mathbf{A}\hat{\mathbf{x}}(n) + \mathbf{B}\mathbf{u}(n)$$

Prädiktionsfehler:  $\hat{\mathbf{e}}^-(n+1) = \mathbf{x}(n+1) - \hat{\mathbf{x}}^-(n+1) = \mathbf{A}(\underbrace{\mathbf{x}(n) - \hat{\mathbf{x}}(n)}_{\hat{\mathbf{e}}(n)}) + \mathbf{G}\mathbf{v}(n)$

### Prädiktionsschritt:

A-priori-Schätzwert:  $\hat{\mathbf{x}}^-(n+1) = \mathbf{A}\hat{\mathbf{x}}(n) + \mathbf{B}\mathbf{u}(n)$

Kovarianz des Prädiktionsfehlers:

$$\mathbf{P}^-(n+1) = \mathbf{A}\mathbf{P}(n)\mathbf{A}^T + \mathbf{G}\mathbf{Q}\mathbf{G}^T$$

Die im Schritt  $n$  bestehende Unsicherheit wird propagiert ( $\mathbf{A}\mathbf{P}(n)\mathbf{A}^T$ ), dazu kommt das neue Systemrauschen hinzu ( $\mathbf{G}\mathbf{Q}\mathbf{G}^T$ ).

## 2.3.3 Kalman-Filter – Innovationsschritt

Innovation: Korrektur des Zustands proportional zum aktuellen Beobachtungsfehler  $\hat{z}(n) - C\hat{x}^-(n)$

$$n := n + 1$$

$\mathbf{K}(n)$ : Verstärkungsmatrix (Kalman gain)

$$\hat{\mathbf{x}}(n) = \hat{\mathbf{x}}^-(n) + \mathbf{K}(n)(\hat{z}(n) - C\hat{\mathbf{x}}^-(n))$$

Innovationsfehler:  $\hat{\mathbf{e}}(n) = \mathbf{x}(n) - \hat{\mathbf{x}}(n)$

### Innovationsschritt:

A-posteriori-Schätzwert:  $\hat{\mathbf{x}}(n) = \hat{\mathbf{x}}^-(n) + \mathbf{K}(n)(\hat{z}(n) - C\hat{\mathbf{x}}^-(n))$

Kovarianz des Innovationsfehlers:

$$\mathbf{P}(n) = (\mathbf{I} - \mathbf{K}(n)\mathbf{C})\mathbf{P}^-(n)$$

Unsicherheit der Prädiktion ( $\mathbf{P}^-(n)$ ) wird durch die Beobachtung reduziert ( $\mathbf{K}(n)\mathbf{C}\mathbf{P}^-(n)$ ).

Verstärkungsmatrix:  $\mathbf{K}(n) = \mathbf{P}^-(n)\mathbf{C}^T (\mathbf{C}\mathbf{P}^-(n)\mathbf{C}^T + \mathbf{R})^{-1}$

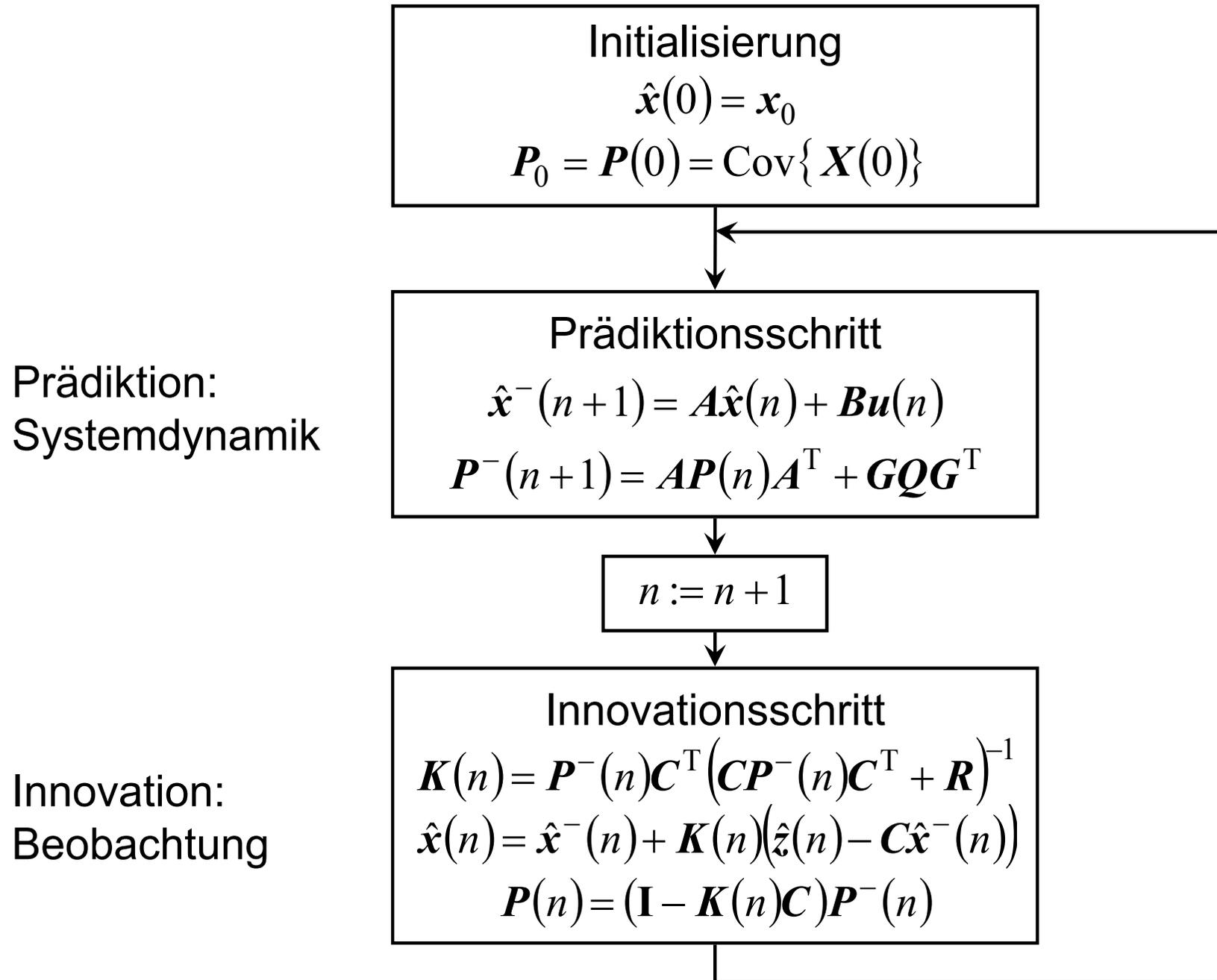
## 2.3.3 Kalman-Filter – Zusammenfassung

---

### Kalman-Filter:

- **Zeitlich veränderliche** Zielgröße zulässig
- Beobachtungen liegen **sequentiell** vor
- **Iterative Struktur**: Zielgröße wird **laufend geschätzt**, jede neue Beobachtung verbessert die Schätzung bzw. führt die Schätzung der veränderlichen Zielgröße nach
- **Least-Squares-(LS-)Schätzer**:  
Minimierung der mittleren quadratischen Abweichung
- Nur **lineare Zustands- und Beobachtungsgleichungen** zulässig

## 2.3.3 Kalman-Filter – Zusammenfassung



## 2.3.3 Kalman-Filter – Beispiel

---

### Bewegungsschätzung für Fahrzeuge

- Gesucht: Eindimensionaler Bewegungszustand eines Fahrzeugs:
  - Position
  - Geschwindigkeit
  - Beschleunigung
  
- Verfügbare Beobachtungen:
  - Positionsmesswerte (z.B. Laserentfernungsmesser, Odometer)
  - Geschwindigkeitsmesswerte
  - Beschleunigungsmesswerte

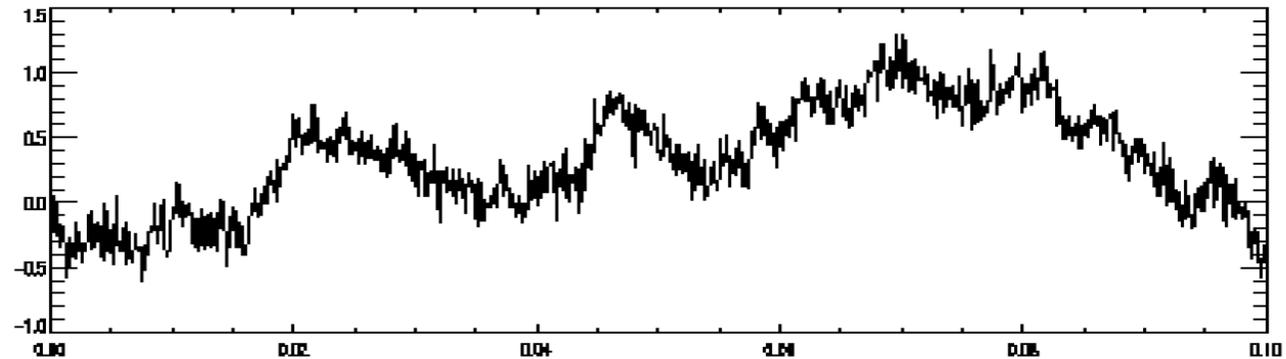
#### Problematisch:

- Unsicherheitsbehaftete Messwerte (z.B. Schlupf)
- Evtl. auftretender Sensorausfall muss erkannt werden

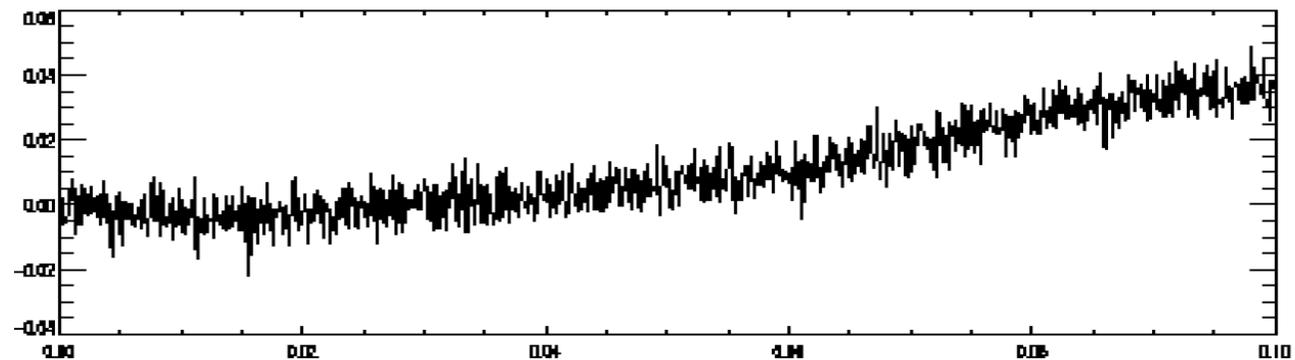
## 2.3.3 Kalman-Filter – Beispiel

### Messwerte

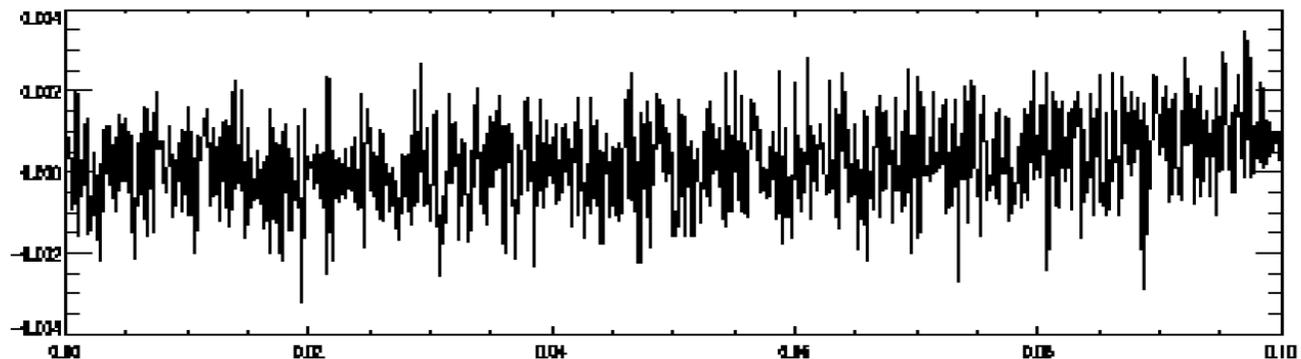
Beschleunigung



Geschwindigkeit

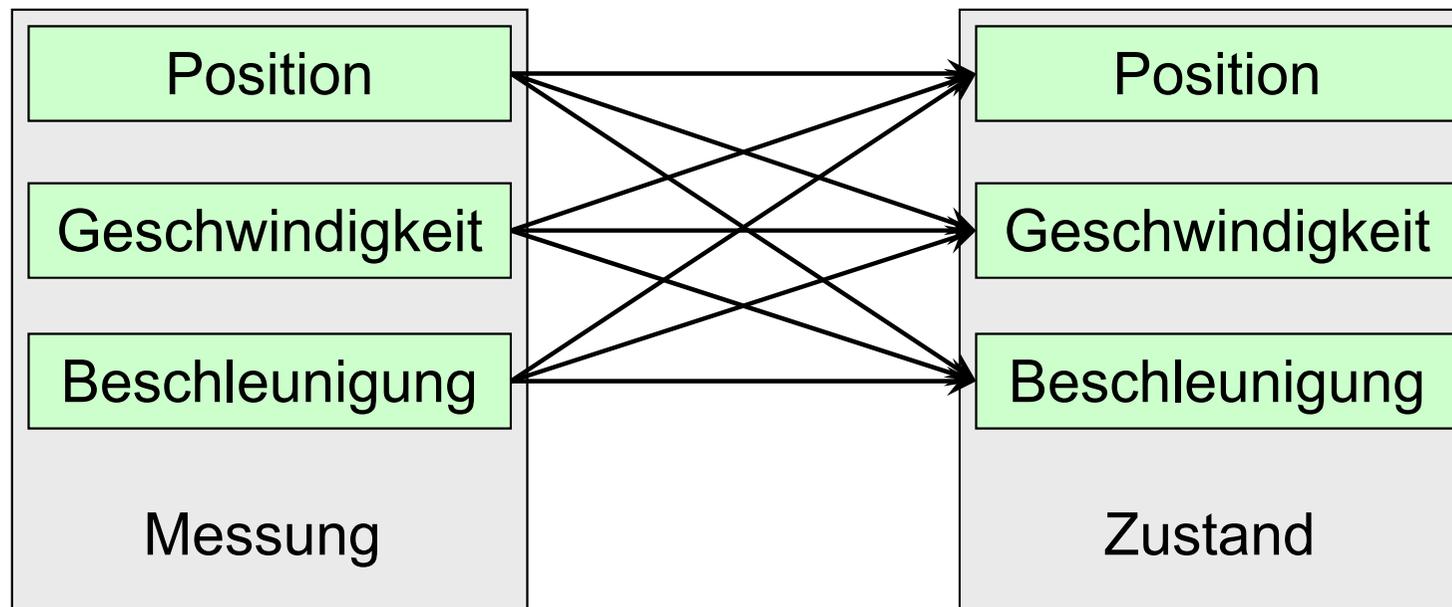


Position



## 2.3.3 Kalman-Filter – Beispiel

- Messwerte sind direkte Beobachtungen der entsprechenden Zustandsgrößen, aber gestört
- Zustandsgrößen lassen sich umrechnen (bis auf Integrationskonstanten):



- Kalman-Filter: Schätzung der zeitvarianten Zustände aus den vorhandenen Beobachtungen (Sensordatenfusion)
- Höhere Zuverlässigkeit der Schätzung des Systemzustands gegenüber unabhängigen Schätzungen der Zustandskomponenten

## 2.3.4 Extended Kalman-Filter – Zustandsraumbeschreibung

Einschränkung bei Standard-Kalman-Filter:

**Lineare** Zustands- und Beobachtungsgleichungen

→ Linearisierung des Zustandsvektors und der Kovarianzmatrix um den aktuellen Arbeitspunkt:

**Extended Kalman-Filter (EKF)**

$$\mathbf{x}(n+1) = \mathbf{f}(\mathbf{x}(n), \mathbf{u}(n), \mathbf{v}(n))$$

Zustandsgleichung

$$\hat{\mathbf{z}}(n) = \mathbf{h}(\mathbf{x}(n), \boldsymbol{\mu}(n))$$

Beobachtungsgleichung

$\mathbf{f}(\cdot, \cdot, \cdot)$ ,  $\mathbf{h}(\cdot, \cdot)$  : nichtlineare, vektorwertige, zumindest um die Arbeitspunkte ableitbare Funktionen

## 2.3.4 Extended Kalman-Filter – Arbeitspunkt

Bestimmung des aktuellen **Arbeitspunktes**:

Näherung:  $\mathbf{v}(n) = \mathbf{0}$ ,  $\boldsymbol{\mu}(n) = \mathbf{0}$  (mit  $E\{\mathbf{V}(i)\} = \mathbf{0}$ ,  $E\{\mathbf{U}(i)\} = \mathbf{0}$  )

$$\tilde{\mathbf{x}}(n+1) = \mathbf{f}(\mathbf{x}(n), \mathbf{u}(n), \mathbf{0}) \quad \text{Zustandsgleichung}$$

$$\hat{\mathbf{z}}(n) = \mathbf{h}(\tilde{\mathbf{x}}(n), \mathbf{0}) \quad \text{Beobachtungsgleichung}$$

**Problematisch:** Nichtlineare Funktionen  $\mathbf{f}(\dots)$ ,  $\mathbf{h}(\dots)$  verändern Gestalt der PDFen der stochastischen Prozesse  $\mathbf{V}(i)$ ,  $\mathbf{U}(i)$ .

- Kovarianzmatrizen  $E\{\mathbf{V}(i)\mathbf{V}^T(i)\}$ ,  $E\{\mathbf{U}(i)\mathbf{U}^T(i)\}$  evtl. nicht mehr aussagekräftig genug
- keine LS-Schätzung mehr erzielbar

## 2.3.4 Extended Kalman-Filter – Linearisierung

**Linearisierung** der Zustandsraumbeschreibung um den Arbeitspunkt:  
(hier  $\mathbf{u}(n)$  vernachlässigt)

$$\mathbf{x}(n+1) \approx \tilde{\mathbf{x}}(n) + \mathbf{A}(n)(\mathbf{x}(n) - \hat{\mathbf{x}}(n)) + \mathbf{G}(n)\mathbf{v}(n) \quad \text{Zustandsgleichung}$$

$$\hat{\mathbf{z}}(n) \approx \hat{\tilde{\mathbf{z}}}(n) + \mathbf{C}(n)(\mathbf{x}(n) - \tilde{\mathbf{x}}(n)) + \mathbf{V}(n)\boldsymbol{\mu}(n) \quad \text{Beobachtungsgleichung}$$

$\hat{\mathbf{z}}(n)$  Ausgangsvektor = Beobachtungsvektor ( $r \times 1$ )

$\mathbf{x}(n)$  Zustandsvektor ( $p \times 1$ )

$\tilde{\mathbf{x}}(n), \hat{\tilde{\mathbf{z}}}(n)$  Arbeitspunkte von  $\mathbf{x}(n), \hat{\mathbf{z}}(n)$ , s. oben

$\hat{\mathbf{x}}(n)$  A-posteriori-Schätzwert des Zeitschritts  $n$

$\mathbf{v}(n)$  Systemrauschen ( $s \times 1$ )

$\boldsymbol{\mu}(n)$  Beobachtungsrauschen ( $r \times 1$ )

## 2.3.4 Extended Kalman-Filter – Linearisierung

$$\mathbf{x}(n+1) \approx \tilde{\mathbf{x}}(n) + \mathbf{A}(n)(\mathbf{x}(n) - \hat{\mathbf{x}}(n)) + \mathbf{G}(n)\mathbf{v}(n) \quad \text{Zustandsgleichung}$$

$$\hat{\mathbf{z}}(n) \approx \tilde{\mathbf{z}}(n) + \mathbf{C}(n)(\mathbf{x}(n) - \tilde{\mathbf{x}}(n)) + \mathbf{V}(n)\boldsymbol{\mu}(n) \quad \text{Beobachtungsgleichung}$$

Jacobi-Matrizen:

$$\mathbf{A}(n) : \text{Systemmatrix} \quad \frac{\partial \mathbf{f}(\hat{\mathbf{x}}(n), \mathbf{u}(n), \mathbf{0})}{\partial \mathbf{x}(n)} \quad A_{i,j}(n) = \frac{\partial f_i(\hat{\mathbf{x}}(n), \mathbf{u}(n), \mathbf{0})}{\partial x_j(n)}$$

$$\mathbf{C}(n) : \text{Beobachtungsmatrix} \quad \frac{\partial \mathbf{h}(\tilde{\mathbf{x}}(n), \mathbf{0})}{\partial \mathbf{x}(n)} \quad C_{i,j}(n) = \frac{\partial h_i(\tilde{\mathbf{x}}(n), \mathbf{0})}{\partial x_j(n)}$$

$$\mathbf{G}(n) : \text{System-Rauschmatrix} \quad \frac{\partial \mathbf{f}(\hat{\mathbf{x}}(n), \mathbf{u}(n), \mathbf{0})}{\partial \mathbf{v}(n)} \quad G_{i,j}(n) = \frac{\partial f_i(\hat{\mathbf{x}}(n), \mathbf{u}(n), \mathbf{0})}{\partial v_j(n)}$$

$$\mathbf{V}(n) : \text{Beobacht.-Rauschmatrix} \quad \frac{\partial \mathbf{h}(\tilde{\mathbf{x}}(n), \mathbf{0})}{\partial \boldsymbol{\mu}(n)} \quad V_{i,j}(n) = \frac{\partial h_i(\tilde{\mathbf{x}}(n), \mathbf{0})}{\partial \mu_j(n)}$$

## 2.3.4 Extended Kalman-Filter – Prädiktionsschritt

Prädiktion: Beste Schätzung des Systemzustands zum Zeitpunkt  $n+1$  ohne neue Beobachtung aus  $\mathbf{u}(1), \dots, \mathbf{u}(n)$ ,  $\hat{\mathbf{x}}(1), \dots, \hat{\mathbf{x}}(n)$ :

$$\hat{\mathbf{x}}^-(n+1) = \mathbf{f}(\hat{\mathbf{x}}(n), \mathbf{u}(n), \mathbf{0})$$

Prädiktionsfehler: (Umdefinition im Vergleich zum linearen Kalman-Filter)

$$\hat{\boldsymbol{\varepsilon}}^-(n+1) = \mathbf{x}(n+1) - \tilde{\mathbf{x}}(n+1) \quad (\mathbf{x}(n+1) \text{ unbekannt})$$

### **Prädiktionsschritt:**

A-priori-Schätzwert:  $\hat{\mathbf{x}}^-(n+1) = \mathbf{f}(\hat{\mathbf{x}}(n), \mathbf{u}(n), \mathbf{0})$

Kovarianz des Prädiktionsfehlers:

$$\mathbf{P}^-(n+1) = \mathbf{A}(n)\mathbf{P}(n)\mathbf{A}(n)^T + \mathbf{G}(n)\mathbf{Q}(n)\mathbf{G}(n)^T$$

Die im Schritt  $n$  bestehende Unsicherheit wird propagiert ( $\mathbf{A}(n)\mathbf{P}(n)\mathbf{A}(n)^T$ ), dazu kommt das neue Systemrauschen hinzu ( $\mathbf{G}(n)\mathbf{Q}(n)\mathbf{G}(n)^T$ ).

## 2.3.4 Extended Kalman-Filter – Innovationsschritt

Innovation: Korrektur des Zustands proportional zum aktuellen Beobachtungsfehler  $\hat{z}(n) - \mathbf{h}(\tilde{\mathbf{x}}(n), \mathbf{0})$

$$n := n + 1$$

$\mathbf{K}(n)$ : Verstärkungsmatrix (Kalman gain)

$$\hat{\mathbf{x}}(n) = \hat{\mathbf{x}}^-(n) + \mathbf{K}(n)(\hat{z}(n) - \mathbf{h}(\tilde{\mathbf{x}}(n), \mathbf{0})) \quad \tilde{\mathbf{x}}(n) = \hat{\mathbf{x}}^-(n)$$

$\mathbf{P}(n)$  und  $\mathbf{K}(n)$  analog zu linearem Kalman-Filter

### Innovationsschritt:

A-posteriori-Schätzwert:  $\hat{\mathbf{x}}(n) = \hat{\mathbf{x}}^-(n) + \mathbf{K}(n)(\hat{z}(n) - \mathbf{h}(\tilde{\mathbf{x}}(n), \mathbf{0}))$

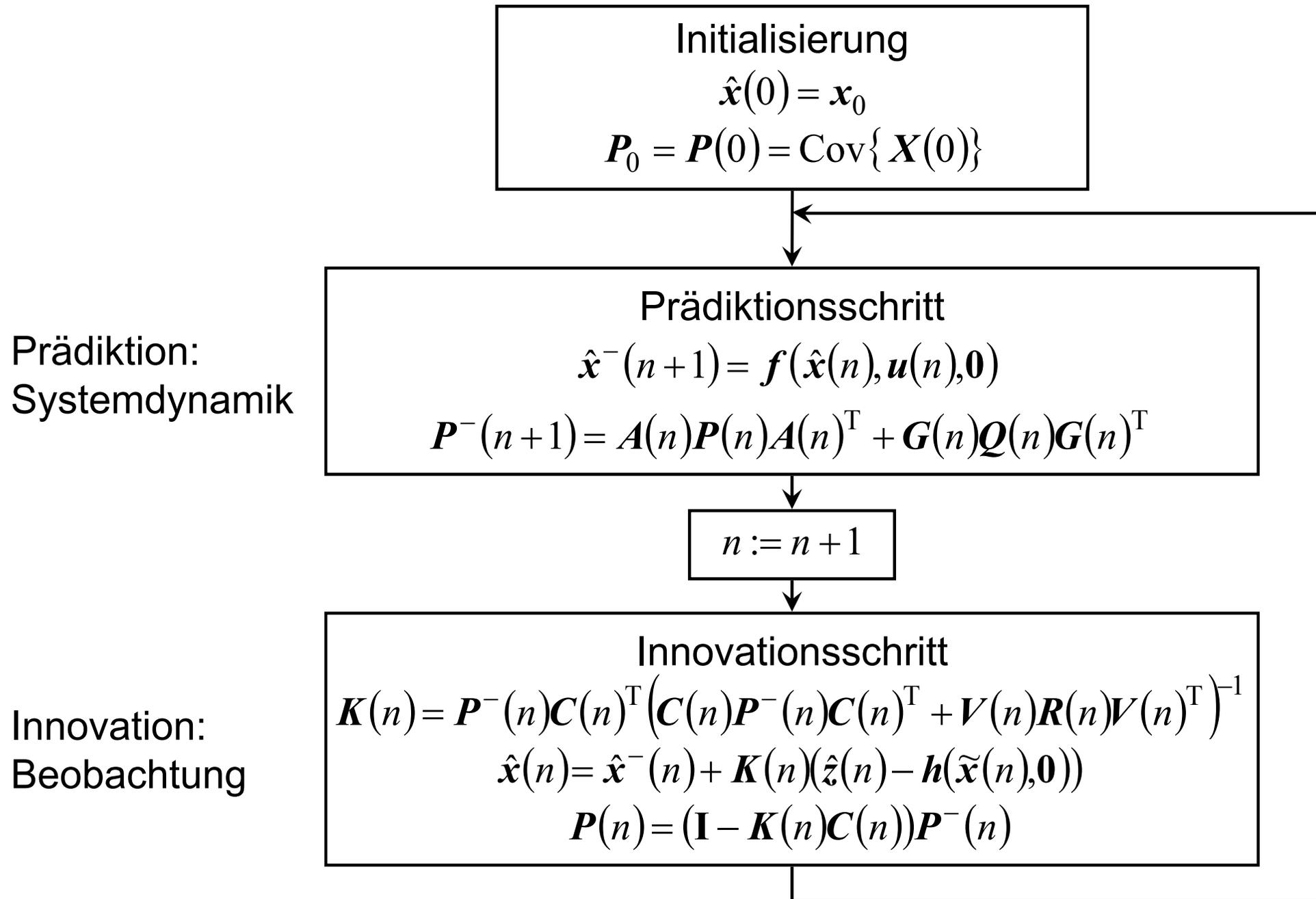
Kovarianz des Innovationsfehlers:

$$\mathbf{P}(n) = (\mathbf{I} - \mathbf{K}(n)\mathbf{C}(n))\mathbf{P}^-(n)$$

Unsicherheit der Prädiktion ( $\mathbf{P}^-(n)$ ) wird durch die Beobachtung reduziert ( $\mathbf{K}(n)\mathbf{C}\mathbf{P}^-(n)$ ).

Verstärkungsmatrix:  $\mathbf{K}(n) = \mathbf{P}^-(n)\mathbf{C}(n)^T \left( \mathbf{C}(n)\mathbf{P}^-(n)\mathbf{C}(n)^T + \mathbf{V}(n)\mathbf{R}(n)\mathbf{V}(n)^T \right)^{-1}$

## 2.3.4 Extended Kalman-Filter – Zusammenfassung



## 2.3.4 Extended Kalman-Filter – Zusammenfassung

---

### Extended Kalman-Filter:

- **Nichtlineare Abhängigkeiten** in Zustands- und Beobachtungsgleichung
- **Jacobi-Matrizen** zur Linearisierung der System- und Beobachtungsgleichungen um den Arbeitspunkt

### Probleme des EKF:

- Kein optimaler Schätzer
- Divergenz bei schlechter Anfangsschätzung möglich
- Nichtlinearitäten werden nur durch Ableitungen erster Ordnung angenähert (Truncation error)
- „Kleine“ Zeitschritte erforderlich
- Linearisierung um Schätzwert, nicht wahren Wert (Base point error)

## 2.4 Guide to the Expression of Uncertainty in Measurement (GUM)

---

- Internationale Norm (1993)
- **Ziel:** Gewährleistung der Vergleichbarkeit von Messergebnissen miteinander bzw. mit Referenzwerten durch **einheitliche, international vereinbarte und anerkannte Grundsätze sowie universelle Vorgehensweisen** zur Bestimmung der Messunsicherheit
- Konsistente und übertragbare Spezifikation der Messunsicherheit
- Probabilistische Beschreibung der Messunsicherheit
- Beschränkt auf Merkmale mit **metrischer Skalierung**
- Hier: Vorgehensweise zur Gewinnung von zu fusionierender Information

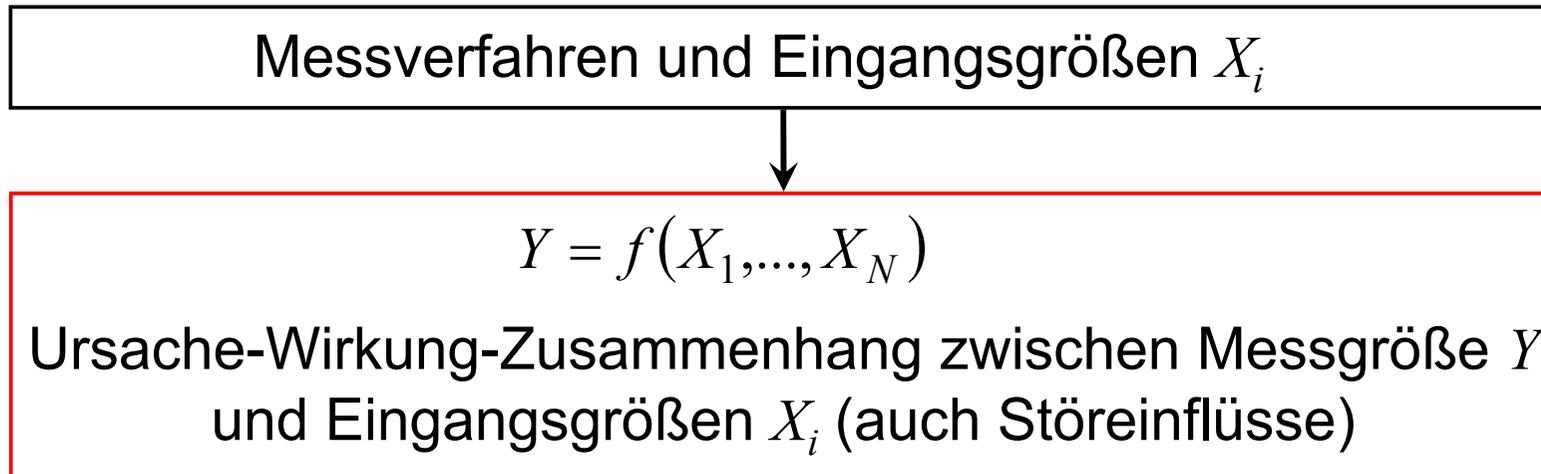
## 2.4.1 GUM: Grundbegriffe

---

- **Messung**
  - Ermittlung des Wertes der zu messenden Größe  $Y$
  - Spezifikation von Messgröße, -methode, -verfahren, -bedingungen
  
- **Messergebnis**
  - Bei Messung erhaltene Schätzung des (wahren) Wertes der Messgröße
  - Nur mit Angabe der Messunsicherheit vollständig/vergleichbar
  
- **Messunsicherheit**
  - Spiegelt die unvollständige Kenntnis des (wahren) Werts der Messgröße wider
  - Parameter zur Kennzeichnung der Dispersion, der der Messgröße zugeordnet werden könnte
  - Beispiel: Standardabweichung

## 2.4.2 Ermittlung der Messunsicherheit gemäß GUM

(1) Formulierung der Modellgleichung  $Y = f(X_1, \dots, X_N)$  der Messung:



- Eingangsgrößen  $X_i$  können ihrerseits nach GUM bestimmte Messgrößen sein
- Berücksichtigung aller zur Unsicherheit von  $Y$  beitragenden Größen
- $f$  kann kompliziert und nicht explizit formulierbar sein
- $f$  kann experimentell ermittelt werden oder als Algorithmus vorliegen

## 2.4.2 Ermittlung der Messunsicherheit gemäß GUM

- (2) (a) Ermittlung der Schätzwerte  $\hat{x}_i$  der Eingangsgrößen  $X_i$  ( $i = 1, \dots, N$ )
- (b) Ermittlung der Standardunsicherheiten  $u_i$  für jeden Schätzwert  $\hat{x}_i$  (Methode A oder B, hängt davon ab, wie die  $\hat{x}_i$  gewonnen wurden und wie ihre Unsicherheiten interpretiert werden können)

### Ermittlungsmethode A der Standardunsicherheit

- Eingangsschätzwert  $\hat{x}_i$ , gewonnen durch **statistische Analyse** von Reihen von Beobachtungen unter gleichen Messbedingungen
- **Frequentistische Wahrscheinlichkeitsinterpretation**
- Kennwerte aus **Häufigkeitsverteilung** ableiten
- Erfordert **ausreichende Anzahl von unabhängigen Beobachtungen**
- **Eingangsgröße  $X_i$**  aus  $n$  unabhängigen Beobachtungen  $x_{i,k}$  schätzen:
  - I.  $\hat{x}_i := \bar{x}_i = \frac{1}{n} \sum_{k=1}^n x_{i,k}$  (Arithmetisches Mittel)

als Schätzwert für den „wahren“ Wert  $E\{X_i\} = \mu_i$

## 2.4.2 Ermittlung der Messunsicherheit gemäß GUM

### II. Schätzung der Varianz der Verteilung von $X_i$ :

$$s^2(x_i) = \frac{1}{n-1} \sum_{k=1}^n (x_{i,k} - \bar{x}_i)^2 \quad \text{Empirische Varianz der Beobachtungen}$$

$$s(x_i) = \sqrt{s^2(x_i)} \quad \text{Empirische Standardabweichung der Beobachtungen}$$

beschreibt die Streuung der Beobachtungswerte  $x_{i,k}$  um  $\bar{x}_i$

### III. Schätzung der Varianz des Mittelwerts:

Quantitative Aussage, wie gut  $\bar{x}_i$  den „wahren“ Wert  $E\{X_i\}$  schätzt

$$s^2(\bar{x}_i) = \frac{s^2(x_i)}{n} \quad \text{Empirische Varianz des Mittelwertes}$$

$$u_i := s(\bar{x}_i) = \sqrt{\frac{s^2(x_i)}{n}} \quad \text{Empirische Standardabweichung des Mittelwertes = Standardunsicherheit}$$

### Ermittlungsmethode A: Zusammenfassung

$\hat{x}_i := \bar{x}_i$  Eingangsschätzwert

$u_i := s(\bar{x}_i)$  Standardunsicherheit

## 2.4.2 Ermittlung der Messunsicherheit gemäß GUM

---

**Ermittlungsmethode B** der Standardunsicherheit:

- Eingangsschätzwert  $\hat{x}_i$ , der *nicht* durch statistische Analyse von Messwerten gewonnen wurde
- Degree-of-Belief-Interpretation von Wahrscheinlichkeiten
- Ermittlung der Standardunsicherheit  $u_i$  durch **Beurteilung aller verfügbaren Informationen** über die Streuung von  $X_i$ , z.B.
  - Daten aus früheren Messungen
  - Herstellerangaben, Handbücher
  - Erfahrungen/Wissen über Messgeräte und Materialien
  - Referenzdaten, Daten von Kalibrierscheinen

## 2.4.2 Ermittlung der Messunsicherheit gemäß GUM

### Ermittlungsmethode B der Standardunsicherheit

#### Beispiel 1:

Schätzwert  $\hat{x}_i$  aus Kalibrierzertifikat,  
Unsicherheit wird dort als Mehrfaches der Standardabweichung angegeben  
→  $u_i$  durch Teilen des angegebenen Wertes durch den Multiplikationsfaktor

#### Beispiel 2:

„Es besteht eine Chance von zwei Dritteln, dass der Wert der Eingangsgröße im Bereich  $a_-$  bis  $a_+$  liegt.  $X_i$  ist näherungsweise normalverteilt.“

$$\rightarrow \hat{x}_i = \frac{a_+ + a_-}{2} \quad u_i = \frac{a_+ - a_-}{2}$$

$[\mu - \sigma, \mu + \sigma]$  schließt ca. 68,3%  $\approx 2/3$  der  $N(\mu, \sigma^2)$ -Verteilung ein

## 2.4.2 Ermittlung der Messunsicherheit gemäß GUM

### Ermittlungsmethode B der Standardunsicherheit

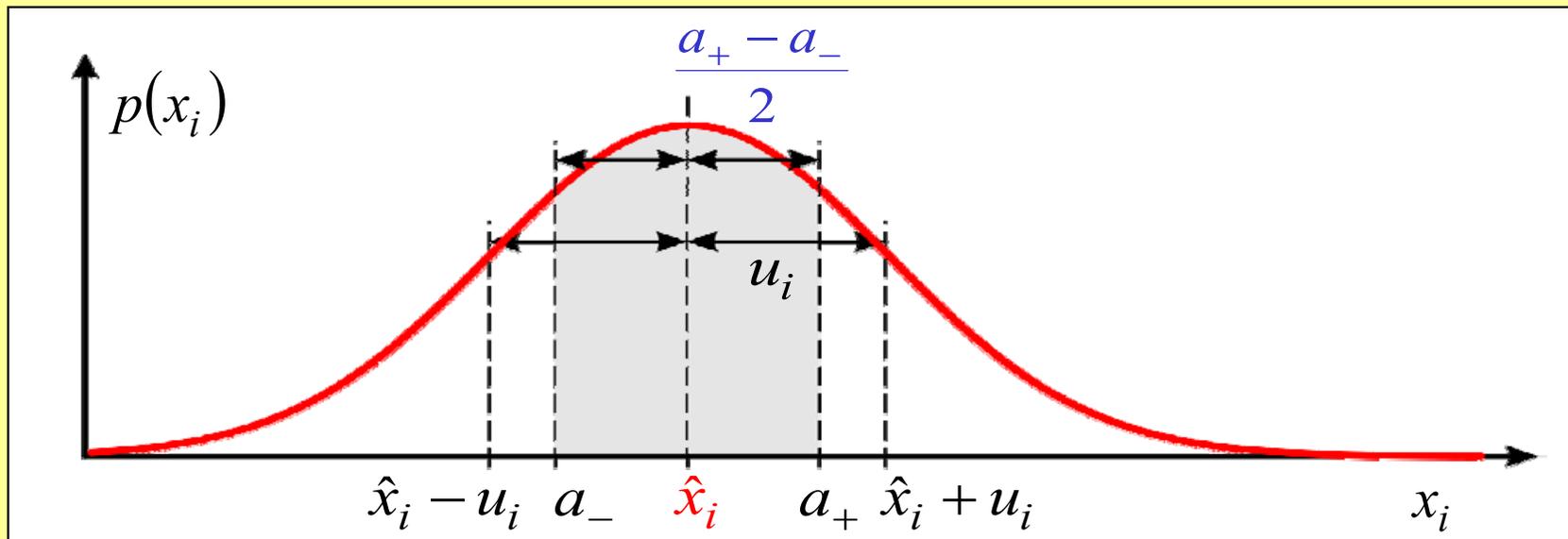
#### Beispiel 3:

„Es besteht eine 50%ige Chance, dass der Wert der Eingangsgröße im Bereich  $a_-$  bis  $a_+$  liegt.  $X_i$  ist näherungsweise normalverteilt.“

→ Schätzwert von  $X_i$ :  $\hat{x}_i = \frac{a_+ + a_-}{2}$ , da Symmetrie der Normalverteilung

Standardabweichung:  $u_i = 1,48 \cdot \frac{a_+ - a_-}{2}$ , denn

$\left[ \mu - \frac{\sigma}{1,48}, \mu + \frac{\sigma}{1,48} \right]$  schließt ca. 50% der  $N(\mu, \sigma^2)$ -Verteilung ein



## 2.4.2 Ermittlung der Messunsicherheit gemäß GUM

### Ermittlungsmethode B der Standardunsicherheit

#### Beispiel 4:

„Die Wahrscheinlichkeit, dass der Wert der Eingangsgröße im Bereich  $[a, b]$  liegt, ist praktisch gleich Eins, und die Wahrscheinlichkeit, dass  $X_i$  außerhalb dieses Bereiches liegt, ist im Wesentlichen Null.“

Weitere Kenntnisse über die Verteilung von  $X_i$  liegen nicht vor.

→ Annahme einer Gleichverteilung für  $X_i$  innerhalb des Bereichs  $[a, b]$ :

$$p(x_i) = \begin{cases} \frac{1}{b-a} & x_i \in [a, b] \\ 0 & \text{sonst} \end{cases} = \frac{1}{b-a} \operatorname{rect}\left(\frac{x_i - \frac{a+b}{2}}{b-a}\right)$$

Schätzwert von  $X_i$ :  $\hat{x}_i = \frac{a+b}{2}$ ,

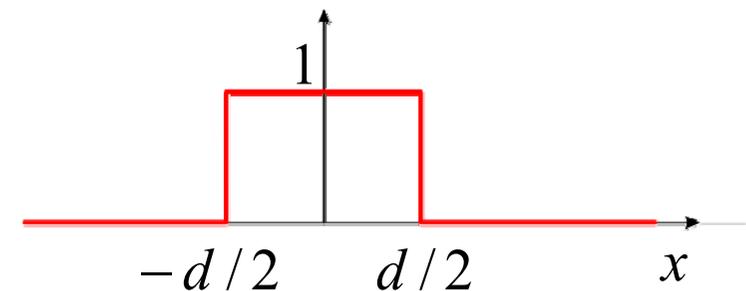
da Symmetrie der Gleichverteilung

Standardabweichung:  $u_i = \sqrt{\operatorname{Var}\{X_i\}} = \frac{b-a}{\sqrt{12}}$ ,

$$\text{da } \operatorname{Var}\{X_i\} = \int_a^b (x_i - E\{X_i\})^2 p(x_i) dx_i = \frac{(b-a)^2}{12}$$

Rechteck-Funktion:

$$\operatorname{rect}\left(\frac{x}{d}\right) := \begin{cases} 1 & x \in \left[-\frac{d}{2}, \frac{d}{2}\right] \\ 0 & \text{sonst} \end{cases}$$



## 2.4.2 Ermittlung der Messunsicherheit gemäß GUM

### (3) Berechnung des Schätzwertes $\hat{y}$ für die Messgröße $Y$

- Standardmethode für (hinreichend) lineare Modellgleichungen:

$$\hat{y} := f(\hat{x}_1, \dots, \hat{x}_N) \quad (*)$$

- Nicht (hinreichend) lineare Modellgleichung:

- $M$  unabhängige Beobachtungen aller  $N$  Eingangsgrößen  $X_i$
- $\hat{y}$  als arithmetisches Mittel von  $M$  unabhängigen Ermittlungen von  $Y$ :

$$\hat{y} := \bar{y} = \frac{1}{M} \sum_{k=1}^M y_k = \frac{1}{M} \sum_{k=1}^M f(x_{1,k}, \dots, x_{N,k}) \quad (**)$$

- Wenn  $f$  lineare Funktion: (\*), (\*\*) sind identisch

## 2.4.2 Ermittlung der Messunsicherheit gemäß GUM

### (4) Ermittlung der kombinierten Standardunsicherheit $u_c^y$ :

- Standardunsicherheit des Schätzwertes  $\hat{y}$  von  $Y$  durch Kombination der Standardunsicherheiten  $u_i$ ,  $i = 1, \dots, N$

### Unsicherheitsfortpflanzungsgesetz:

$$Y = f(X_1, \dots, X_N), \quad E\{X_i\} = \mu_i, \quad \mu^y := f(\mu_1, \dots, \mu_N)$$

$$\begin{aligned} Y - \mu^y &= f(X_1, \dots, X_N) - f(\mu_1, \dots, \mu_N) \\ &= f(\mu_1 + \Delta_1, \dots, \mu_N + \Delta_N) - f(\mu_1, \dots, \mu_N) \quad \text{mit } \Delta_i = X_i - \mu_i \end{aligned}$$

$$\begin{aligned} &\approx f(\mu_1, \dots, \mu_N) + \sum_{i=1}^N \frac{\partial f}{\partial X_i}(\mu_i) \cdot \Delta_i - f(\mu_1, \dots, \mu_N) \end{aligned}$$

### Taylorentwicklung

$$= \sum_{i=1}^N \frac{\partial f}{\partial X_i}(\mu_i) \cdot (X_i - \mu_i)$$

- Wenn  $f$  nicht hinreichend linear:  
Einbeziehung von Gliedern höherer Ordnung in der Taylorentwicklung

## 2.4.2 Ermittlung der Messunsicherheit gemäß GUM

### Unsicherheitsfortpflanzungsgesetz

$$\begin{aligned}(Y - \mu^y)^2 &= \left( \sum_{i=1}^N \frac{\partial f}{\partial X_i}(\mu_i) \cdot (X_i - \mu_i) \right)^2 \\ &= \sum_{i=1}^N \left( \frac{\partial f}{\partial X_i}(\mu_i) \right)^2 \cdot (X_i - \mu_i)^2 \\ &\quad + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N \frac{\partial f}{\partial X_i}(\mu_i) \frac{\partial f}{\partial X_j}(\mu_j) \cdot (X_i - \mu_i)(X_j - \mu_j) \\ \underbrace{\mathbb{E}\left\{(Y - \mu^y)^2\right\}}_{\text{Var}\{Y\}} &= \sum_{i=1}^N \left( \frac{\partial f}{\partial X_i}(\mu_i) \right)^2 \cdot \underbrace{\mathbb{E}\left\{(X_i - \mu_i)^2\right\}}_{\text{Var}\{X_i\}} \\ &\quad + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N \frac{\partial f}{\partial X_i}(\mu_i) \frac{\partial f}{\partial X_j}(\mu_j) \cdot \underbrace{\mathbb{E}\left\{(X_i - \mu_i)(X_j - \mu_j)\right\}}_{\text{Cov}\{X_i, X_j\}}\end{aligned}$$

## 2.4.2 Ermittlung der Messunsicherheit gemäß GUM

### Unsicherheitsfortpflanzungsgesetz

$$\begin{aligned}\text{Var}\{Y\} &= \sum_{i=1}^N \left( \frac{\partial f}{\partial X_i}(\mu_i) \right)^2 \text{Var}\{X_i\} + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N \frac{\partial f}{\partial X_i}(\mu_i) \frac{\partial f}{\partial X_j}(\mu_j) \text{Cov}\{X_i, X_j\} \\ &= \sum_{i=1}^N \left( \frac{\partial f}{\partial X_i}(\mu_i) \right)^2 \text{Var}\{X_i\} \\ &\quad + 2 \underbrace{\sum_{i=1}^{N-1} \sum_{j=i+1}^N \frac{\partial f}{\partial X_i}(\mu_i) \frac{\partial f}{\partial X_j}(\mu_j) \sqrt{\text{Var}\{X_i\} \text{Var}\{X_j\}} r(X_i, X_j)}_{= 0, \text{ falls } X_i, X_j \text{ unkorreliert}}\end{aligned}$$

mit  $r(X_i, X_j) = \frac{\text{Cov}\{X_i, X_j\}}{\sqrt{\text{Var}\{X_i\} \text{Var}\{X_j\}}} \in [-1, +1]$  : Korrelationskoeffizient

## 2.4.2 Ermittlung der Messunsicherheit gemäß GUM

### (4) Kombinierte Standardunsicherheit $u_c^y$

- Für **unkorrel.** Eingangsgrößen gilt näherungsw. ( $\hat{x}_i \approx \mu_i, u_i^2 \approx \text{Var}\{X_i\}$ )

$$(u_c^y)^2 = \sum_{i=1}^N \left( \frac{\partial f}{\partial X_i}(\hat{x}_i) \right)^2 u_i^2 =: \sum_{i=1}^N (u_i^y)^2, \text{ wobei } u_i^y = |c_i| u_i, c_i = \frac{\partial f}{\partial X_i}(\hat{x}_i)$$

- Für **korrelierte** Eingangsgrößen gilt näherungsweise:

$$\begin{aligned} (u_c^y)^2 &= \sum_{i=1}^N \left( \frac{\partial f}{\partial X_i}(\hat{x}_i) \right)^2 u_i^2 + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N \frac{\partial f}{\partial X_i}(\hat{x}_i) \frac{\partial f}{\partial X_j}(\hat{x}_j) u_{i,j} \\ &= \sum_{i=1}^N \left( \frac{\partial f}{\partial X_i}(\hat{x}_i) \right)^2 u_i^2 + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N \frac{\partial f}{\partial X_i}(\hat{x}_i) \frac{\partial f}{\partial X_j}(\hat{x}_j) u_i u_j r_{i,j} \\ &= \sum_{i=1}^N c_i^2 u_i^2 + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N c_i c_j u_i u_j r_{i,j} \end{aligned}$$

wobei  $u_{i,j} = u_{j,i}$ : geschätzte Kovarianz zwischen  $X_i$  und  $X_j$

$r_{i,j} = r_{j,i} = \frac{u_{i,j}}{u_i u_j} \in [-1, +1]$ : Schätzung des Korrelationskoeffizienten

## 2.4.2 Ermittlung der Messunsicherheit gemäß GUM

### (5) Falls erforderlich: Spezifizierung einer erweiterten Unsicherheit $U$

- Falls Angabe der Standardunsicherheit nicht ausreichend ist, z.B. wegen Gesundheits-, Sicherheitsaspekten
- Angabe eines Bereiches  $[\hat{y} - U, \hat{y} + U]$ , von dem erwartet werden kann, dass er einen „großen Anteil“ der Verteilung der Werte umfasst, die der Messgröße  $Y$  sinnvollerweise zugeordnet werden können:

$$U := k \cdot u_c^y \text{ erweiterte Unsicherheit}$$

- $k$ : **Erweiterungsfaktor**, abhängig vom geforderten „Grad des Vertrauens“  $p$ , wird i.d.R. zwischen zwei und drei gewählt
- Zuordnung eines Erweiterungsfaktors  $k$  an einen Grad des Vertrauens  $p$  nur möglich, wenn Wahrscheinlichkeitsverteilung bekannt  
Bsp.: Normalverteilung:  $k = 2 \Leftrightarrow p = 95,5\%$ ,  $k = 3 \Leftrightarrow p = 99,7\%$ ,

## 2.4.2 Ermittlung der Messunsicherheit gemäß GUM

Beispiel: Verteilung der Ausgangsgröße bekannt

- Normalverteilung:

$$p(y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-\mu)^2}{2\sigma^2}}$$

- Grad des Vertrauens:

$$p = 95,4\%$$

$$\Rightarrow k = 2, U \approx 2\sigma$$

$$\Pr(\mu - 2\sigma \leq Y \leq \mu + 2\sigma) \approx 95,4\%$$

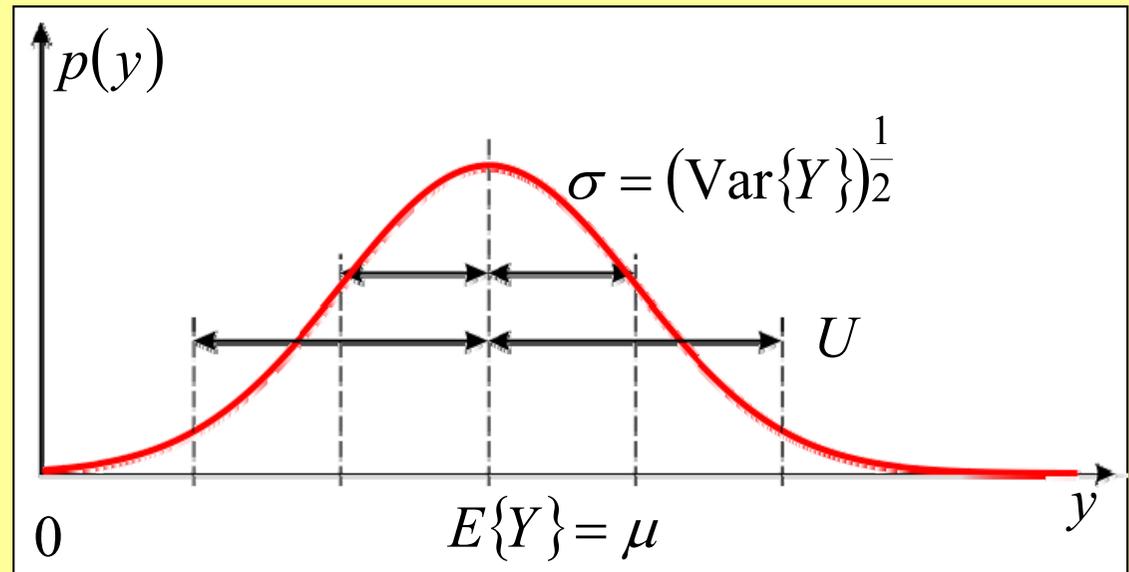
GUM-konforme Angabe des Messergebnisses:  $Y = \mu \pm U$  ( $k = 2$ )

- Genauso: Grad des Vertrauens:  $p = 99,7\%$

$$\Rightarrow k = 3, U \approx 3\sigma$$

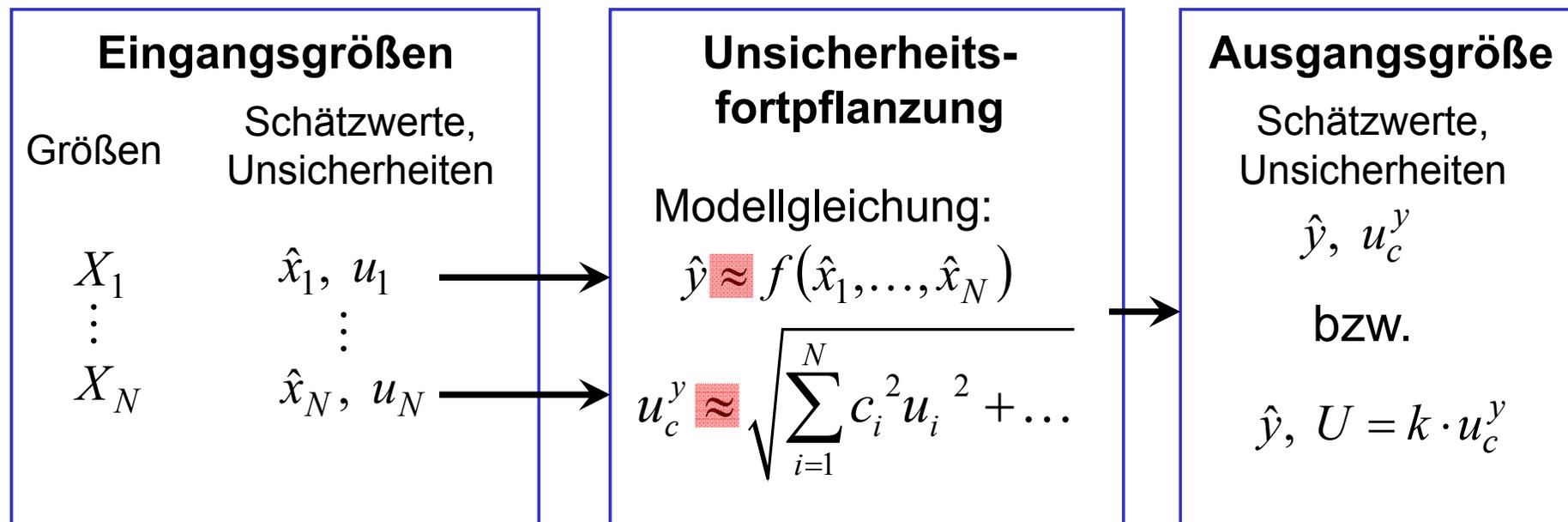
$$\Pr(\mu - 3\sigma \leq Y \leq \mu + 3\sigma) \approx 99,7\%$$

GUM-konforme Angabe des Messergebnisses:  $Y = \mu \pm U$  ( $k = 3$ )



## 2.4.2 Ermittlung der Messunsicherheit gemäß GUM – Zusammenfassung

- Vereinfachte probabilistische Betrachtung der Messung:
  - Beschreibung aller Eingangsgrößen  $X_i$ ,  $i = 1, \dots, N$  als Zufallsgrößen mittels Schätzwerten  $\hat{x}_1$  und Unsicherheiten  $u_i$
  - Modellgleichung  $Y = f(X_1, \dots, X_N)$  beschreibt Ursache-Wirkung-Zusammenhang zw. der Messgröße  $Y$  und den Eingangsgrößen  $X_i$
  - Messgröße  $Y$ : abgeleitete **Zufallsvariable**, deren Wert ebenfalls durch einen Schätzwert  $\hat{y}$  und eine Unsicherheit  $u_c^y$  beschrieben wird



## 2.4.3 Vollständige Beschreibung von Zufallsgrößen durch PDFen

---

Falls näherungsweise Angabe von Schätzwert und Unsicherheit nach GUM nicht ausreichend ist:

**Vollständige probabilistische Beschreibung** der Eingangsgrößen  $X_i$  und der Ausgangsgröße  $Y$  durch Wahrscheinlichkeitsdichtefunktionen (PDFen)

$$p_{X_i}(\xi), i = 1, \dots, N, \quad p_Y(\eta)$$

- Möglich bei **verfügbarer (objektiver) Information über die PDFen der Eingangsgrößen  $X_i$** , z.B. durch
  - Direkte (frequentistische) Beobachtungen
  - Werte aus Kalibrier-/Prüfscheinen, Herstellerangaben, Literaturwerte
  - Prinzip der Maximalen Entropie (PME, s.u.),  
d.h. Wahl derjenigen WV, die in Erfüllung gegebener Bedingungen die Entropie maximiert

## 2.4.3 Vollständige Beschreibung von Zufallsgrößen durch PDFen

Wahrscheinlichkeitsdichtefunktion der Ausgangsgröße  $Y$   
aus der Modellgleichung  $Y = f(X_1, \dots, X_N)$ :

$$p_Y(\eta) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \underbrace{\delta(\eta - f(\xi_1, \dots, \xi_N))}_{\text{Delta-Distribution:}} \cdot \underbrace{p_{X_1, \dots, X_N}(\xi_1, \dots, \xi_N)}_{\text{Verbundwahrscheinlichkeitsdichte}} d\xi_1 \cdots d\xi_N$$
$$\delta(x \neq 0) := 0, \quad \int_{-\infty}^{\infty} \delta(x) dx = 1$$

### Schwierigkeiten

- Verbundwahrscheinlichkeitsdichte  $p_{X_1, \dots, X_N}(\xi_1, \dots, \xi_N)$
- Sind die  $X_i$  unabhängig, unkorreliert?  
Falls  $X_i$  unabhängig:  $p_{X_1, \dots, X_N}(\xi_1, \dots, \xi_N) = p_{X_1}(\xi_1) \cdots p_{X_N}(\xi_N)$
- Rechenaufwand!

## 2.5 Bayes'sche Fusion

### 2.5.1 Bayes'sche Methodik

Zielgrößen (gesucht)

$$\mathbf{x} = (x_1, \dots, x_N) \in \mathbf{X}$$

$$\mathbf{X} = \mathbf{X}_1 \times \dots \times \mathbf{X}_N, N \in \mathbb{N}$$

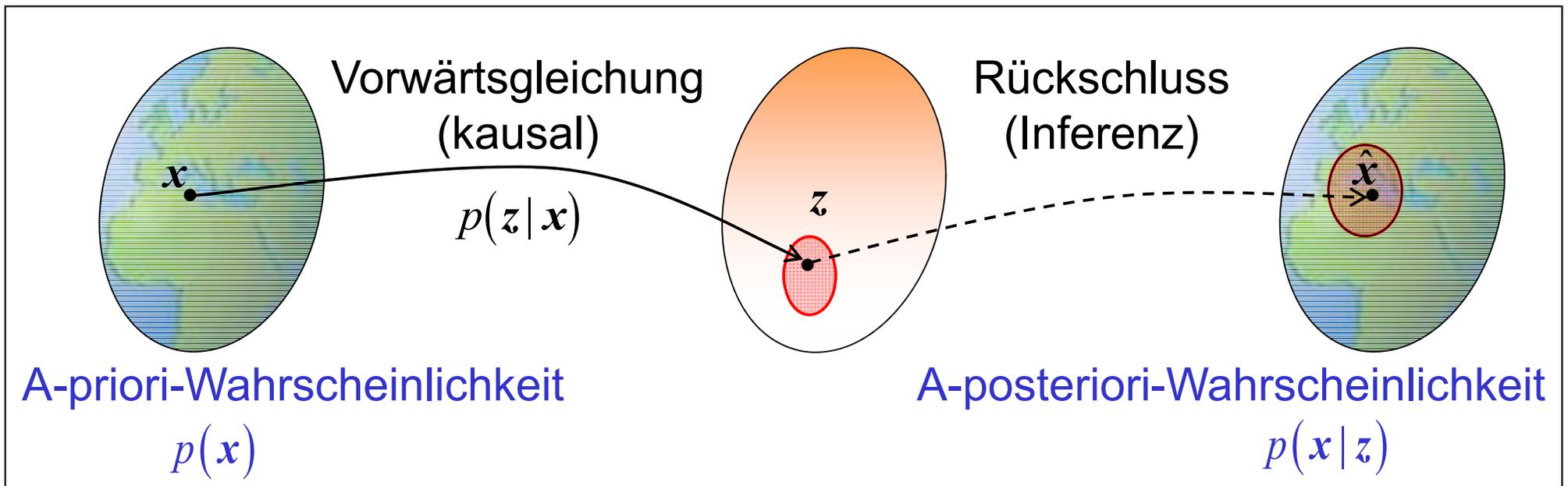
Daten (beobachtbar)

$$\mathbf{z} = (z_1, \dots, z_S) \in \mathbf{Z}$$

$$\mathbf{Z} = \mathbf{Z}_1 \times \dots \times \mathbf{Z}_S, S \in \mathbb{N}$$

Zielgrößen (Schätzung)

$$\hat{\mathbf{x}} \in \mathbf{X}$$



**Basis für  
Bayes'sche  
Fusion:**

$$p(\mathbf{x}|\mathbf{z}) = \frac{p(\mathbf{z}|\mathbf{x})p(\mathbf{x})}{p(\mathbf{z})} = \underbrace{\frac{p(\mathbf{z}|\mathbf{x})p(\mathbf{x})}{\sum_{\mathbf{x}} \int p(\mathbf{z}|\mathbf{x})p(\mathbf{x}) d\mathbf{x}}}_{\text{Normierungskonstante}} \propto p(\mathbf{z}|\mathbf{x})p(\mathbf{x})$$

Satz von Bayes

## 2.5.1 Bayes'sche Methodik

- Bestimmung der **A-priori-Wahrscheinlichkeit**  $p(\mathbf{x})$

### Beispiele:

- Zielgrößen erfüllen gewisse **Glattheitsanforderungen**  
(siehe später: Energiefunktionale, Bayes'sche Interpretation mittels Gibbs'scher Dichten)
- **Naturgesetze**  
(z.B. geometrische Eigenschaften der Zielgröße: Größe von Fahrzeugen bei gegebenem Abstand und Typ)
- **Systemmodell**:  $p(\mathbf{x}(n) | \mathbf{x}(n-1))$   
„Wahre“ A-priori-Wahrscheinlichkeit:  $p(\mathbf{x}(1) | \mathbf{x}(0)) = p(\mathbf{x}(1))$  ,  
da  $\mathbf{x}(0)$  fest vorgegeben
- **Empirisches Wissen**  
(z.B. Auftretenswahrscheinlichkeiten von Fahrzeugen im Gebirge)
- **Naiver Ansatz**: Gleichverteilung auf  $X$  annehmen  
(entspricht PME bei fehlendem Wissen über  $X$ )

## 2.5.1 Bayes'sche Methodik

- Prinzip der Maximalen Entropie:

$$\text{Entropie: } H(\mathbf{x}) = -\sum_{\mathbf{x}} \text{Pr}(\mathbf{x}) \log_2 \text{Pr}(\mathbf{x}) = \text{E}\{-\log_2 \text{Pr}(\mathbf{x})\} \quad (\mathbf{x} \text{ diskret})$$

$$h(\mathbf{x}) = -\int p(\mathbf{x}) \log_2 p(\mathbf{x}) \, d\mathbf{x} = \text{E}\{-\log_2 p(\mathbf{x})\} \quad (\mathbf{x} \text{ kontinuierlich})$$

Vorsicht:  $h$  resultiert nicht als Grenzwert aus  $H$  bei einer immer feiner werdenden Diskretisierung (siehe Übung), d.h.  $H$  und  $h$  sind zwei verschiedene Größen!

- Gegebene Fakten: Nebenbedingungen
- Bestimmung derjenigen Wahrscheinlichkeitsverteilung  $p_{\text{ME}}(\mathbf{x})$ , die unter Einhaltung der Nebenbedingungen maximale Entropie (d.h. maximale Unsicherheit) aufweist:

$$p_{\text{ME}}(\mathbf{x}) := \arg \max_{q(\mathbf{x}) \in \Pi} \text{E}\{-\log_2 q(\mathbf{x})\}$$

$$\Pi := \{q(\mathbf{x}) \mid q(\mathbf{x}) \text{ ist WV auf } \mathbf{X} \wedge q(\mathbf{x}) \text{ ist mit den Fakten verträglich}\}$$

- $p_{\text{ME}}(\mathbf{x})$  verkörpert die Fakten in wahrscheinlichkeitstheoretischer Notation; die Fakten sind hieraus rekonstruierbar.

## 2.5.1 Bayes'sche Methodik

- Prinzip der Maximalen Entropie:

(1) Nebenbedingungen als Erwartungswerte:  $E\{k_l(\mathbf{x})\} = \kappa_l, l = 1, \dots, L$

Isoperimetrisches Variationsproblem für  $p(\mathbf{x}) \geq 0$ :

$$h(\mathbf{x}) = -\int p(\mathbf{x}) \log_2 p(\mathbf{x}) \, d\mathbf{x} \stackrel{!}{=} \max \Leftrightarrow -\int p(\mathbf{x}) \ln p(\mathbf{x}) \, d\mathbf{x} \stackrel{!}{=} \max$$

$$\text{mit Nebenbedingungen: } \int k_l(\mathbf{x}) p(\mathbf{x}) \, d\mathbf{x} = \kappa_l$$

$$\int p(\mathbf{x}) \, d\mathbf{x} = 1$$

Lagrange-Ansatz:

$$J(p(\mathbf{x})) := -\int p(\mathbf{x}) \ln p(\mathbf{x}) \, d\mathbf{x} + \lambda_0 \int p(\mathbf{x}) \, d\mathbf{x} + \sum_{l=1}^L \lambda_l \int k_l(\mathbf{x}) p(\mathbf{x}) \, d\mathbf{x}$$

$$\frac{\partial J(p(\mathbf{x}))}{\partial p(\mathbf{x})} = -\ln p(\mathbf{x}) - 1 + \lambda_0 + \sum_{l=1}^L \lambda_l k_l(\mathbf{x}) \stackrel{!}{=} 0$$

$$\Rightarrow p_{\text{ME}}(\mathbf{x}) = \exp \left\{ \lambda_0 - 1 + \sum_{l=1}^L \lambda_l k_l(\mathbf{x}) \right\}$$

$$\text{d.h. } p_{\text{ME}}(\mathbf{x}) = \tilde{\lambda}_0 \cdot \exp \left( \sum_{l=1}^L \lambda_l k_l(\mathbf{x}) \right), \quad \lambda_l \text{ aus Erwartungswerten}$$

## 2.5.1 Bayes'sche Methodik

- Prinzip der Maximalen Entropie:

(2) Nebenbedingungen als Gleichungen oder Ungleichungen, d.h. Beschränkung des Grundraums  $\Omega$  auf die Lösungsmenge  $X'$  der Nebenbedingungen:  $f(\mathbf{x}) \begin{matrix} \geq \\ \leq \end{matrix} 0$

$$\text{Indikatorfunktion: } I(\mathbf{x}) := \begin{cases} 1 & \mathbf{x} \in X' \\ 0 & \mathbf{x} \notin X' \end{cases}$$

$$\text{somit: } p_{ME}(\mathbf{x}) = \tilde{\lambda}_0 \cdot I(\mathbf{x}) \cdot \exp\left(\sum_{l=1}^L \lambda_l k_l(\mathbf{x})\right)$$

Beispiel: Bekannt:  $x \in [a, b]$

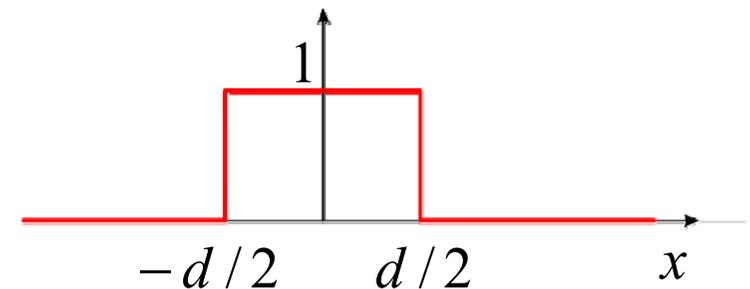
$$p_{ME}(x) = \frac{1}{b-a} \text{rect}\left(\frac{x - \frac{a+b}{2}}{b-a}\right)$$

Z.B. bei

- vorgegebenen Toleranzen
- bekannten Fehlergrenzen
- Quantisierungsfehlern

Rechteck-Funktion:

$$\text{rect}\left(\frac{x}{d}\right) := \begin{cases} 1 & x \in \left[-\frac{d}{2}, \frac{d}{2}\right] \\ 0 & \text{sonst} \end{cases}$$



## 2.5.1 Bayes'sche Methodik

---

- **Bedingte Wahrscheinlichkeit**  $p(z | x)$ :

Wahrscheinlichkeit, dass die Daten  $z$  beobachtet werden, wenn die Zielgröße  $x$  vorliegt.

- Sensormodell, z.B. Gauß'sches Beobachtungsmodell, Modellierung von Beobachtungsunsicherheiten
- Ausgabe des Sensorsystems in Abhängigkeit von der phys. Realität
- Für jede Realisierung von  $x$ : Konstruktion einer WV über  $z$
- „Familie“ von WV für variiertes  $z$  mit Parameter  $x$
- Interpretation für festes  $z$ : Likelihood-Funktion für  $x : \Lambda(x)$
- **Vorsicht:**  $p(z | x)$  ist nur WV für festes  $x$ , nicht für festes  $z$  !  
(Normiertheitsbedingung von Kolmogorov ist meist verletzt)

## 2.5.1 Bayes'sche Methodik

### Beispiel: Kamerabild eines Fahrzeugs

- $\mathbf{x} = (\text{Position, Orientierung, Typ})^T$
- $\mathbf{z} = \text{Kamerabild}$
- Modellierung der Abbildung: CAD-Modell, Beleuchtungsmodellierung, Reflexionsmodellierung, Raytracing
- Bei idealer Beobachtung (festgelegte Umgebungsbedingungen, keine Beobachtungsunsicherheit etc.):  $\mathbf{z} = \mathbf{f}(\mathbf{x})$  unrealistisch!
- Unsicherheiten der Modellierung und Beobachtung: Umgebungslicht, Anbauten, Verdeckungen, Verschmutzungen, Beschädigungen, Wetter (Nebel, Regen), ...  
→ Modellierung als WV  $p(\mathbf{z} | \mathbf{x})$



Bundesarchiv, 8 146 Bild-F038805-0013  
Foto: Schaack, Lothar | 25. Januar 1973

Quelle: de.wikipedia.org

## 2.5.1 Bayes'sche Methodik

---

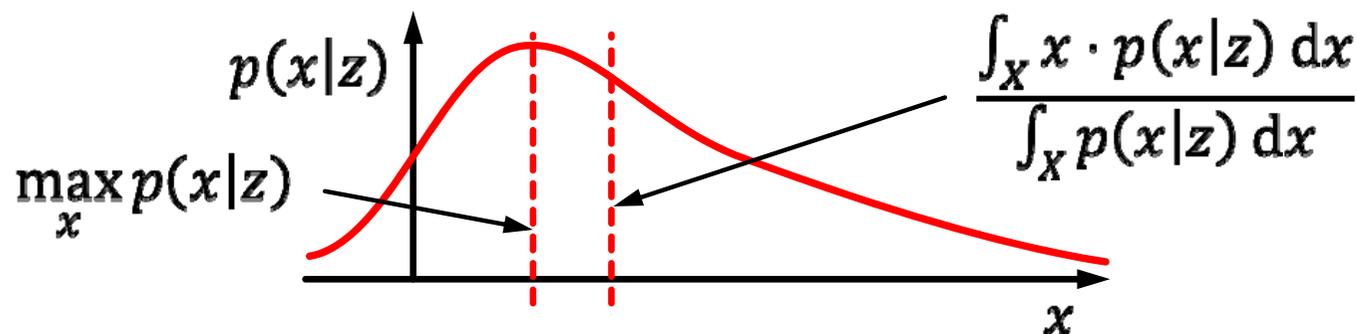
- **A-posteriori-Wahrscheinlichkeitsverteilung**  $p(x | z)$  der Zielgröße  $x$  verkörpert umfassend die bis dahin zusammengefassten Daten und Informationen.
  - ⇒ A-posteriori-WV muss von Sensorsystemen und Fusionssystemen kommuniziert werden
  - ⇒ A-posteriori-WV als „natürliche Eingangsgröße“ für Verfahren der Entscheidungstheorie

## 2.5.1 Bayes'sche Methodik

**Satz von Bayes:** 
$$p(\mathbf{x}|\mathbf{z}) = \frac{p(\mathbf{z}|\mathbf{x})p(\mathbf{x})}{p(\mathbf{z})}$$

Rolle des Nenners:  $p(\mathbf{z})$ :

- Wahrscheinlichkeitsverteilung für das Auftreten der Beobachtungen
- Unabhängig von  $\mathbf{x}$
- Normierungskonstante, damit  $\int_{\mathbf{X}} p(\mathbf{x}|\mathbf{z}) d\mathbf{x} = 1$
- Bestimmung von  $p(\mathbf{z})$ :  $p(\mathbf{z}) = \int_{\mathbf{X}} p(\mathbf{x}, \mathbf{z}) d\mathbf{x} = \int_{\mathbf{X}} p(\mathbf{z}|\mathbf{x})p(\mathbf{x}) d\mathbf{x}$ ,  
d.h. Integration des Zählers der Bayes'schen Formel über alle  $\mathbf{x}$
- Vereinfachung:  $p(\mathbf{x}|\mathbf{z}) \propto p(\mathbf{z}|\mathbf{x})p(\mathbf{x})$   
möglich, wenn Wert der Wahrscheinlichkeit  $p(\mathbf{x}|\mathbf{z})$  nicht erforderlich ist,  
z.B. für Bestimmung des Maximums oder des Schwerpunkts von  $p(\mathbf{x}|\mathbf{z})$



## 2.5.1 Bayes'sche Methodik

Übergang der A-posteriori-WV auf „scharfe“ Werte:  
Schätzung von  $\mathbf{x}$ :

- **Maximum-Likelihood-Schätzung:**  $\hat{\mathbf{x}}_{\text{MLE}} = \arg \max_{\mathbf{x}} p(\mathbf{z} | \mathbf{x})$
- **Maximum-A-Posteriori-Schätzung:**  $\hat{\mathbf{x}}_{\text{MAP}} = \arg \max_{\mathbf{x}} p(\mathbf{x} | \mathbf{z})$

$$\begin{aligned} \text{Für } p(\mathbf{x}) = \text{const.} : \quad \hat{\mathbf{x}}_{\text{MAP}} &= \arg \max_{\mathbf{x}} p(\mathbf{x} | \mathbf{z}) \\ &= \arg \max_{\mathbf{x}} \left[ p(\mathbf{z} | \mathbf{x}) p(\mathbf{x}) \right] \\ &= \arg \max_{\mathbf{x}} \left[ p(\mathbf{z} | \mathbf{x}) \right] = \hat{\mathbf{x}}_{\text{MLE}} \end{aligned}$$

- **Minimum-Mean-Square-Error-Schätzung:**

$$J(\mathbf{x}) = \int_{\mathbf{X}} (\mathbf{x} - \xi)^T \mathbf{C} (\mathbf{x} - \xi) p(\xi | \mathbf{z}) \, d\xi \rightarrow \text{Minimum}$$

$\mathbf{C}$  : positiv semidefinite Gewichtungsmatrix (Kostenmatrix)

$$\hat{\mathbf{x}}_{\text{MMSE}} = \arg \min_{\mathbf{x}} J(\mathbf{x}) = \int_{\mathbf{X}} \xi p(\xi | \mathbf{z}) \, d\xi \quad (\text{unabhängig von } \mathbf{C})$$

Erwartungswert der A-posteriori-Verteilung

## 2.5.1 Bayes'sche Methodik

Beispiel: Kontinuierliche Zielgröße  $x$  (z.B. Abstand),  
eine Beobachtung  $z$ :

Gauß'sches Beobachtungsmodell:

$$p_{z|x}(z|x) = \frac{1}{\sqrt{2\pi\sigma_z^2}} e^{-\frac{(z-x)^2}{2\sigma_z^2}} \quad \text{Funktion von } z \text{ und } x$$

Konstruktion des Modells: Zielgröße festhalten:  $x = x_0$ , dann  
WV als Funktion von  $z$  bestimmen

Verwendung des Modells: Beobachtung machen:  $z = z_0$ , dann  
als Funktion von  $x$  auswerten:  
Likelihood-Funktion

A-priori-Wahrscheinlichkeit:

$$p_x(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} e^{-\frac{(x-x_p)^2}{2\sigma_x^2}}$$

$x_p, \sigma_x$  : Bekannte Eigenschaften der Zielgröße

## 2.5.1 Bayes'sche Methodik

Beispiel: (Fortsetzung)

A-posteriori-Wahrscheinlichkeit:

Beobachtung  $z_0$

$$\begin{aligned} p_{X|Z}(x|z) &= C p_{Z|X}(z|x) p_X(x) = C \frac{1}{\sqrt{2\pi\sigma_Z^2}} e^{-\frac{(x-z_0)^2}{2\sigma_Z^2}} \frac{1}{\sqrt{2\pi\sigma_X^2}} e^{-\frac{(x-x_p)^2}{2\sigma_X^2}} \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\hat{x})^2}{2\sigma^2}} \end{aligned}$$

$$\hat{x} = \frac{\sigma_X^2}{\sigma_X^2 + \sigma_Z^2} z_0 + \frac{\sigma_Z^2}{\sigma_X^2 + \sigma_Z^2} x_p$$

$$\sigma^2 = \mathbb{E}\{\hat{X}^2\} = \frac{\sigma_Z^2 \sigma_X^2}{\sigma_Z^2 + \sigma_X^2} = \left( \frac{1}{\sigma_Z^2} + \frac{1}{\sigma_X^2} \right)^{-1}$$

## 2.5.2 Unabhängigkeitsannahmen

### Unabhängigkeit und bedingte Unabhängigkeit: (Zufallsvariablen $x, y, z$ )

- Annahme: Kenntnis von  $y$  liefert keine Information über  $x$ :

$$p(\mathbf{x} | \mathbf{y}) = p(\mathbf{x})$$

$x$  und  $y$  sind **stochastisch unabhängig**:

$$p(\mathbf{x}, \mathbf{y}) = p(\mathbf{x} | \mathbf{y})p(\mathbf{y}) = p(\mathbf{x})p(\mathbf{y})$$

- Annahme: Kenntnis von  $z$  macht  $x$  und  $y$  stochastisch unabhängig:

$$p(\mathbf{x} | \mathbf{y}, \mathbf{z}) = p(\mathbf{x} | \mathbf{z})$$

$x$  und  $y$  sind **bedingt stochastisch unabhängig** (gegeben  $z$ ):

$$p(\mathbf{x}, \mathbf{y} | \mathbf{z}) = p(\mathbf{x} | \mathbf{y}, \mathbf{z})p(\mathbf{y} | \mathbf{z}) = p(\mathbf{x} | \mathbf{z})p(\mathbf{y} | \mathbf{z})$$

$z$  enthält indirekt den gesamten Informationsbeitrag von  $y$  an  $x$   
(und umgekehrt)

## 2.5.2 Unabhängigkeitsannahmen

**Bedingte stochastische Unabhängigkeit:**  
Grundlage für viele Verfahren zur Informationsfusion  
**Naive Bayes'sche Fusion**

- Bedingt unabhängige Daten („independent feature model“):

$z_1$  und  $z_2$ : Daten bezüglich derselben Zielgröße  $x$

→  $z_1$  und  $z_2$  sind nicht stochastisch unabhängig, da sie sich auf dieselbe Zielgröße beziehen:  $p(z_1, z_2) \neq p(z_1)p(z_2)$

**Annahme der bedingten stochastischen Unabhängigkeit:**

$z_1$  und  $z_2$  sind stochastisch unabhängig bei festem  $x$  :

$$p(z_1 | z_2, \mathbf{x}) = p(z_1 | \mathbf{x}), \quad p(z_1, z_2 | \mathbf{x}) = p(z_1 | \mathbf{x}) \cdot p(z_2 | \mathbf{x})$$

$$\Rightarrow p(\mathbf{x} | z_1, z_2) \propto p(z_1, z_2 | \mathbf{x}) p(\mathbf{x}) = p(z_1 | \mathbf{x}) p(z_2 | \mathbf{x}) p(\mathbf{x})$$

## 2.5.2 Unabhängigkeitsannahmen

### Beispiel: Kriminalistische Untersuchung

(nach einem wahren Fall: People vs. Collins, L.A. 1964)

Raubüberfall, detaillierte Täterbeschreibung eines Paares  
(Beobachtungen  $z_i$ ):

- Mann mit Oberlippenbart
- Schwarzer Mann mit Kinnbart
- Frau mit Pferdeschwanzfrisur
- Blonde Frau
- Paar aus unterschiedlichen Rassen
- Gelbes Fluchtfahrzeug

Passende Tatverdächtige gefunden

Wie hoch ist die Wahrscheinlichkeit, dass dieses Paar das  
gesuchte ist?

Gesucht:  $\Pr(x = w \mid z)$

Quelle: K. Korb: Bayesian artificial intelligence. Chapman & Hall, 2004.

## 2.5.2 Unabhängigkeitsannahmen

### Beispiel: (Fortsetzung)

#### Probabilistische Überlegung:

- $\Pr(z_1) = \Pr(\text{Mann mit Oberlippenbart}) = 1/4$
- $\Pr(z_2) = \Pr(\text{Schwarzer Mann mit Kinnbart}) = 1/10$
- $\Pr(z_3) = \Pr(\text{Frau mit Pferdeschwanzfrisur}) = 1/10$
- $\Pr(z_4) = \Pr(\text{Blonde Frau}) = 1/3$
- $\Pr(z_5) = \Pr(\text{Paar aus unterschiedlichen Rassen}) = 1/1000$
- $\Pr(z_6) = \Pr(\text{Gelbes Fluchtfahrzeug}) = 1/10$

Annahme des Sachverständigen: Unabhängige Eigenschaften

$$\Pr(\mathbf{z} \mid x = f) = \prod_i \Pr(z_i \mid x = f) = \frac{1}{12.000.000} \rightarrow \text{Verurteilung!}$$

**Fehler 1: Keine unabhängigen Eigenschaften!**

Bessere Annahmen:  $\Pr(\mathbf{z} \mid x = f) = \prod_{i=2,3,4,6} \Pr(z_i \mid x = f) = \frac{1}{3.000}$

- $z_2$  impliziert meist  $z_1$
- $z_2, z_3, z_4$  implizieren  $z_5$

**Fehler 2:**  $\Pr(x = w \mid \mathbf{z}) \neq 1 - \Pr(\mathbf{z} \mid x = f)$

Quelle: K. Korb: Bayesian artificial intelligence. Chapman & Hall, 2004.

## 2.5.2 Unabhängigkeitsannahmen

Beispiel: (Fortsetzung)

Vollständige Anwendung der Bayes'schen Methodik:

$$\begin{aligned}\Pr(x = w | z) &= \frac{\Pr(z | x = w) \Pr(x = w)}{\Pr(z)} \\ &= \frac{\Pr(z | x = w) \Pr(x = w)}{\Pr(z | x = w) \Pr(x = w) + \Pr(z | x = f) \Pr(x = f)}\end{aligned}$$

Vereinfachende Annahmen:  $\Pr(z | x = w) = 1$   
(d.h. die Zeugen sagen die Wahrheit)

A-priori-WV (Gleichverteilung für alle Paare  
im Großraum L.A.):  $\Pr(x = w) = \frac{1}{1.625.000}$

$$\Pr(x = w | z) = \frac{1 \cdot \frac{1}{1.625.000}}{1 \cdot \frac{1}{1.625.000} + \frac{1}{3.000} \left(1 - \frac{1}{1.625.000}\right)} \approx 0,002$$

D.h. zufällig gewählte Verdächtige, auf die die Beobachtungen  
passen, sind zu 99,8% unschuldig!

Quelle: K. Korb: Bayesian artificial intelligence. Chapman & Hall, 2004.

## 2.5.2 Unabhängigkeitsannahmen

Beispiel: Spam-Klassifikation (vgl. Übung 1)

$S$  : Mail ist Spam

$W_i$  : Wort  $i$  tritt in der Mail auf,  $i = 1, \dots, N$

$$\text{Bayes: } \Pr(S | W_1, \dots, W_N) = \frac{\Pr(W_1, \dots, W_N | S) \Pr(S)}{\Pr(W_1, \dots, W_N)}$$

Annahme bedingter stochastischer Unabhängigkeit des Auftretens der Wörter  $W_1, \dots, W_N$  in einer Spam-Mail:

$$\Pr(W_1, \dots, W_N | S) = \Pr(W_1 | S) \cdot \dots \cdot \Pr(W_N | S)$$

Problem:  $\Pr(W_1, \dots, W_N)$  unbekannt

→ Likelihood-Verhältnis:

$$\begin{aligned} \frac{\Pr(S | W_1, \dots, W_N)}{\Pr(\bar{S} | W_1, \dots, W_N)} &= \frac{\Pr(W_1, \dots, W_N | S) \Pr(S)}{\Pr(W_1, \dots, W_N | \bar{S}) \Pr(\bar{S})} \\ &= \frac{\Pr(S)}{\Pr(\bar{S})} \prod_{i=1}^N \frac{\Pr(W_i | S)}{\Pr(W_i | \bar{S})} \end{aligned}$$

## 2.5.3 Bayes'sche Fusion

Satz von Bayes für  $L$  Datenquellen:

$$p(\mathbf{x} | \mathbf{z}_1, \dots, \mathbf{z}_L) = \frac{p(\mathbf{z}_1, \dots, \mathbf{z}_L | \mathbf{x}) p(\mathbf{x})}{p(\mathbf{z}_1, \dots, \mathbf{z}_L)}$$

Annahme der bedingten stochastischen Unabhängigkeit:

$\mathbf{z}_i$  sind stochastisch unabhängig bei festem  $\mathbf{x}$  :

$$p(\mathbf{z}_1, \dots, \mathbf{z}_L | \mathbf{x}) = \prod_{i=1}^L p(\mathbf{z}_i | \mathbf{x})$$

Dann gilt:

$$p(\mathbf{x} | \mathbf{z}_1, \dots, \mathbf{z}_L) = \frac{\prod_{i=1}^L p(\mathbf{z}_i | \mathbf{x}) p(\mathbf{x})}{p(\mathbf{z}_1, \dots, \mathbf{z}_L)} \propto \underbrace{\prod_{i=1}^L p(\mathbf{z}_i | \mathbf{x})}_{\text{Likelihood-Produkt:}} \times \underbrace{p(\mathbf{x})}_{\text{A-priori-WV:}}$$

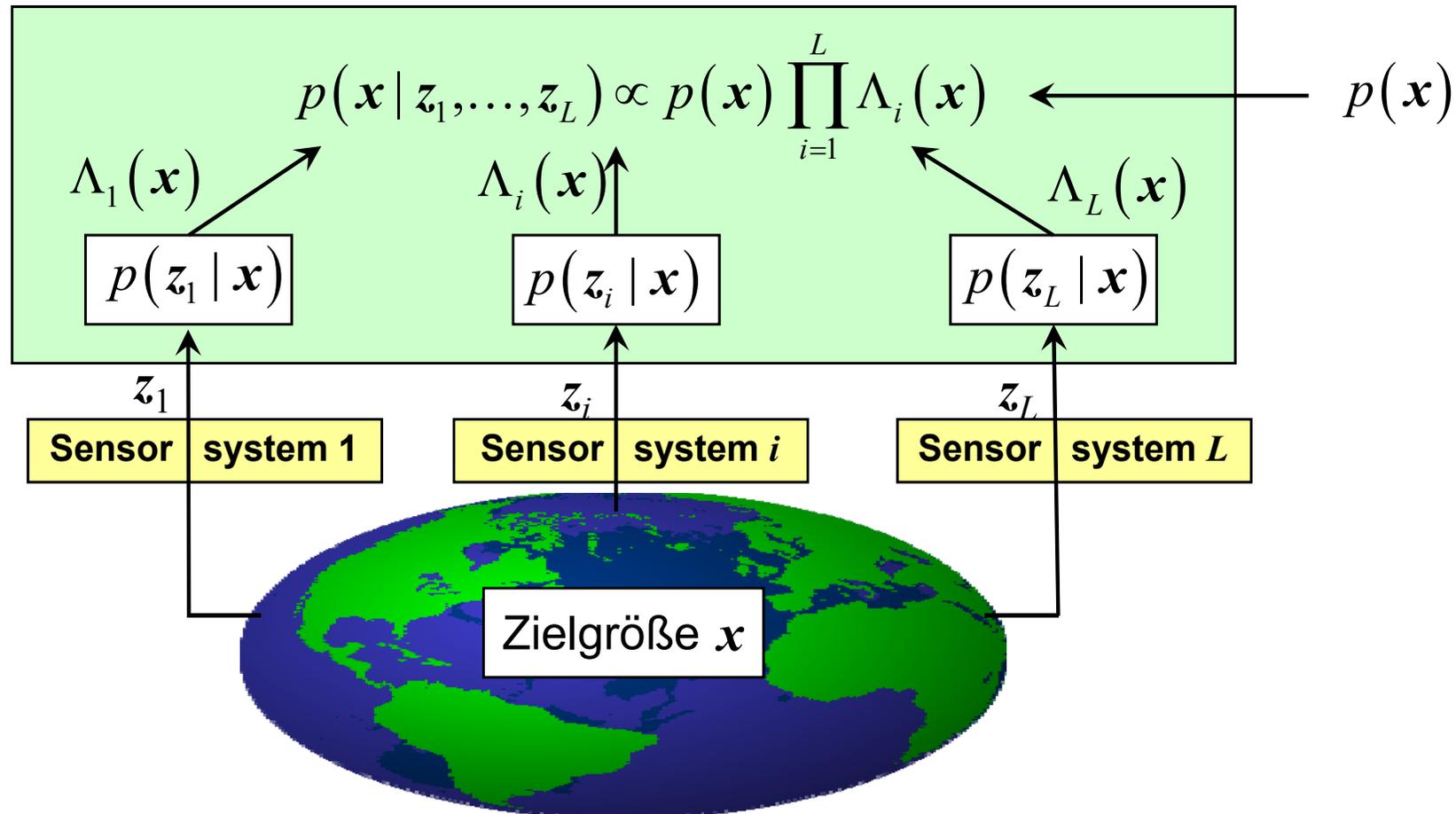
Schnittstelle für  
physikalisch modellierte,  
sensorielle Information

Schnittstelle für nicht  
sensorielle Information

## 2.5.3 Bayes'sche Fusion

### Realisierung der Fusion (zentralisiert):

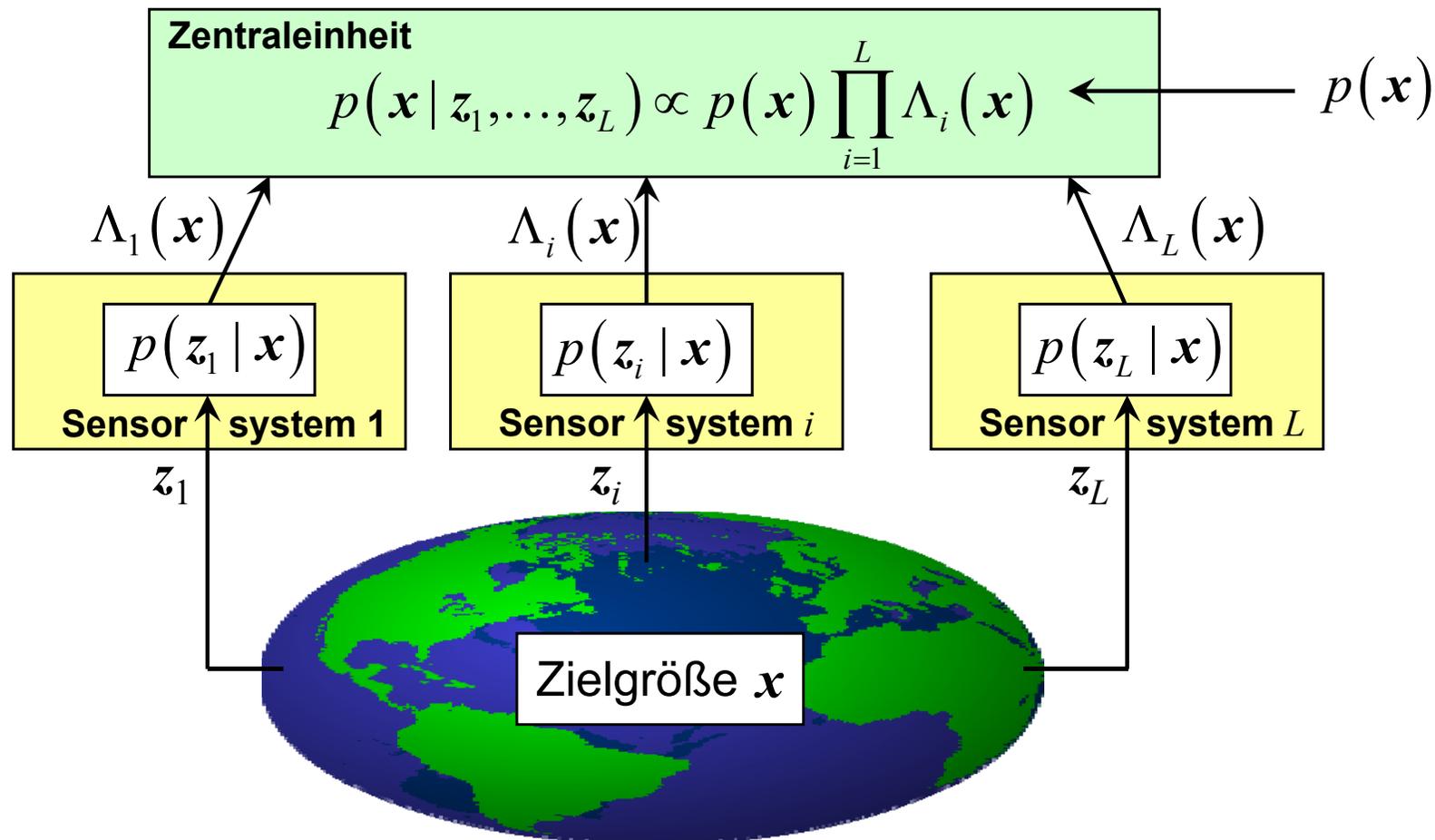
- Bedingte Wahrscheinlichkeiten  $p(z_i | x)$  werden **vorab berechnet** und gespeichert als Funktionen von  $z_i$  und  $x$
- Beobachtung  $z_i$ : Konstruktion von Likelihood-Funktionen  $\Lambda_i(x)$
- **Kommunikation der Beobachtungen**  $z_i$  an die Zentraleinheit



## 2.5.3 Bayes'sche Fusion

### Realisierung der Fusion (zentralisiert):

- Alternativ: Lokale Konstruktion von Likelihood-Funktionen  $\Lambda_i(\mathbf{x})$  in den Sensorsystemen
- Kommunikation der Likelihood-Funktionen  $\Lambda_i(\mathbf{x})$  an die Zentraleinheit
- Vorteil: Reduzierter Kommunikationsaufwand



## 2.5.3 Bayes'sche Fusion

### Realisierung der Fusion (verteilt):

- Lokale Konstruktion der A-posteriori-WV in den Sensorsystemen:

$$p(\mathbf{x} | \mathbf{z}_i) \propto p(\mathbf{z}_i | \mathbf{x}) p(\mathbf{x})$$

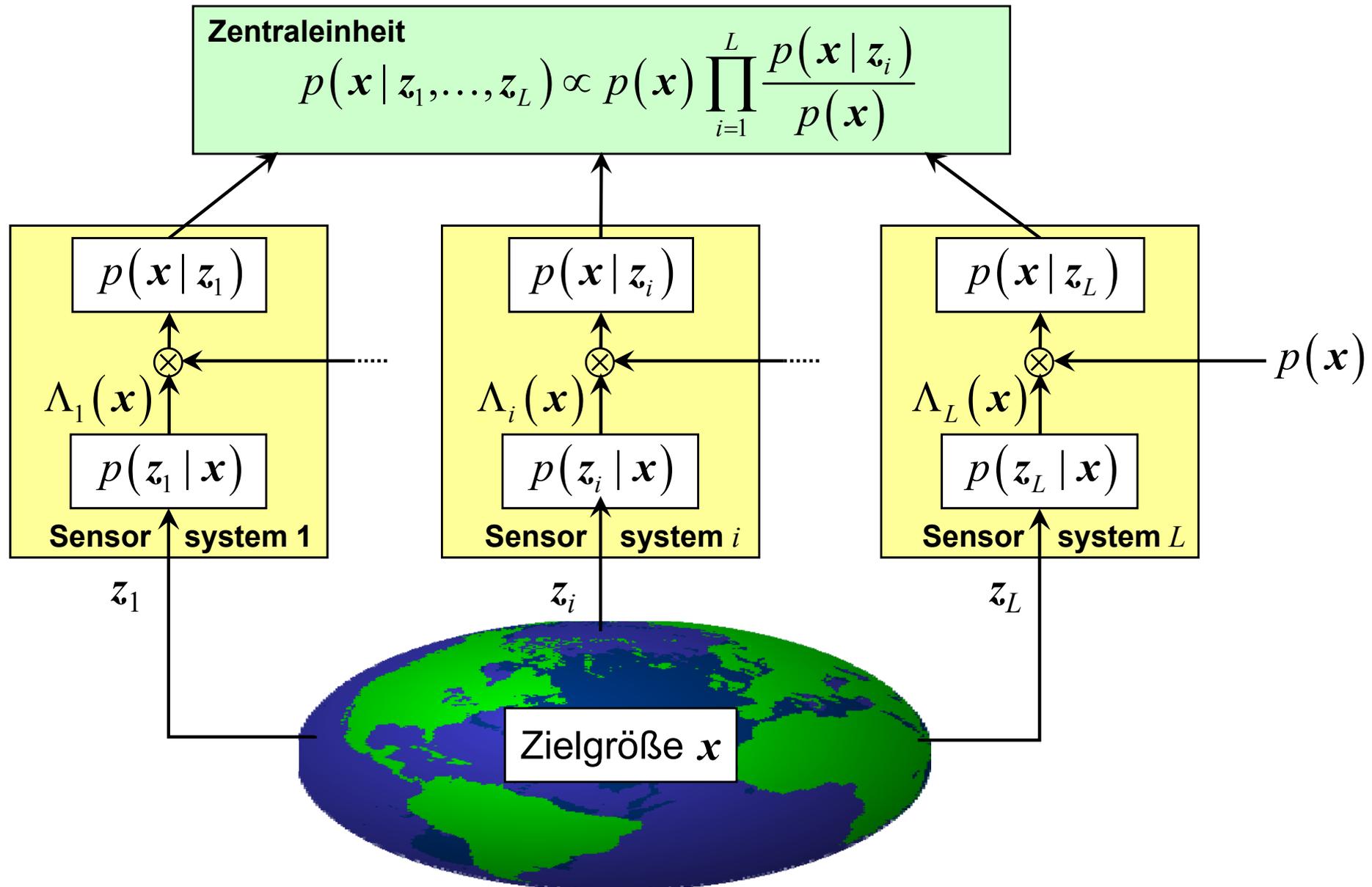
- Kommunikation der A-posteriori-WV  $p(\mathbf{x} | \mathbf{z}_i)$  an die Zentraleinheit

$$\begin{aligned} p(\mathbf{x} | \mathbf{z}_1, \dots, \mathbf{z}_L) &\propto p(\mathbf{x}) \prod_{i=1}^L \Lambda_i(\mathbf{x}) \\ &\propto p(\mathbf{x}) \prod_{i=1}^L \frac{p(\mathbf{x} | \mathbf{z}_i)}{p(\mathbf{x})} \end{aligned}$$

- Vergleich zur zentralisierten Fusion:
  - + Geringerer Kommunikationsbedarf (ähnlich wie Likelihood-Funktionen)
  - + Priorisierbarkeit von Informationen im Kommunikationskanal  
entsprechend lokaler A-posteriori-WV
  - Ggf. Informationsreduktion durch Priorisierung

## 2.5.3 Bayes'sche Fusion

### Realisierung der Fusion (verteilt):



## 2.5.3 Bayes'sche Fusion

### Rekursive Bayes'sche Fusion

- Annahme: Nacheinander vorliegende Beobachtungen  $z(n)$ , Zeitpunkt  $n$   
Zum Zeitpunkt  $n$  bekannte Zielgröße:  $x(n) \quad n > 0$
- Menge aller zum Zeitpunkt  $n$  vorliegenden Beobachtungen:  
$$\mathbf{Z}(n) := \{z(n), \mathbf{Z}(n-1)\} \quad n > 1$$
- Annahme: Beobachtungen  $z(n)$  sind bedingt stochastisch unabhängig:  
$$p(\mathbf{Z}(n) | x(n)) = p(z(n) | x(n)) \cdot p(\mathbf{Z}(n-1) | x(n)) \quad n > 1$$
  
aber: 
$$p(\mathbf{Z}(n)) = p(z(n), \mathbf{Z}(n-1)) = p(z(n) | \mathbf{Z}(n-1)) \cdot p(\mathbf{Z}(n-1)) \quad n > 1 \quad (*)$$
  
$$\neq p(z(n)) \cdot p(\mathbf{Z}(n-1))$$

- Rekursionsvorschrift:  
$$p(x(n) | \mathbf{Z}(n)) = p(x(n-1) | \mathbf{Z}(n-1)) \frac{p(z(n) | x(n))}{p(z(n) | \mathbf{Z}(n-1))}$$

- Berechnen und speichern nur von  $p(x(n) | \mathbf{Z}(n))$ :  
Enthält alle bis zum Zeitschritt  $n$  verfügbare Information
- Neue Messung  $n + 1$ :  $p(x(n) | \mathbf{Z}(n))$  wird zur neuen A-priori-WV  
$$n \rightarrow n - 1$$

## 2.5.3 Bayes'sche Fusion

### Rekursive Bayes'sche Fusion

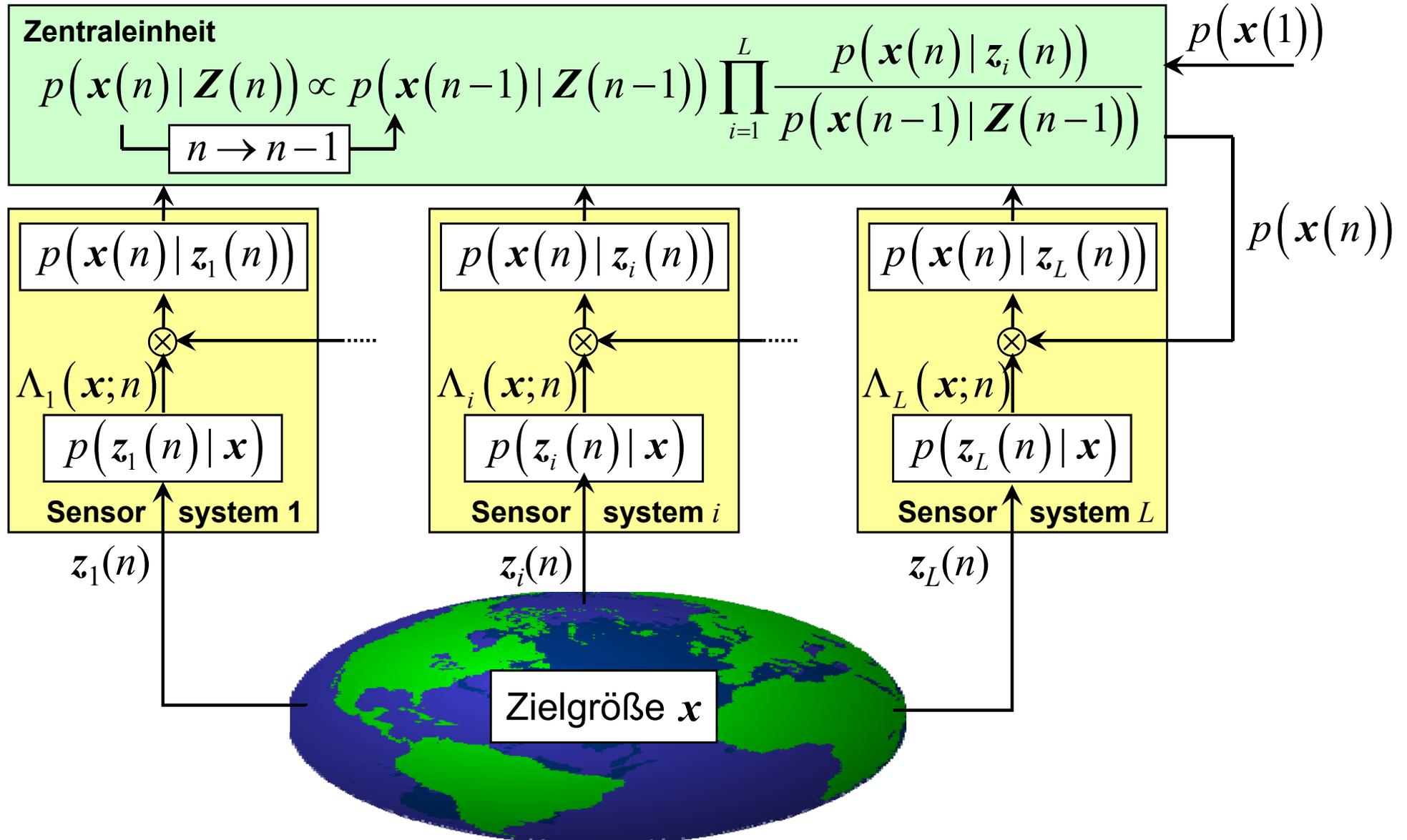
- Abkürzung:  $\mathbf{Z}(n) := \{\mathbf{Z}_1(n), \dots, \mathbf{Z}_L(n)\}$
- Durchführung der Fusion  
(wie oben, Annahme der bedingten stochastischen Unabhängigkeit der Beobachtungen):

$$p(\mathbf{x}(n) | \mathbf{Z}(n)) \propto p(\mathbf{x}(n-1) | \mathbf{Z}(n-1)) \prod_{i=1}^L p(\mathbf{z}_i(n) | \mathbf{x}(n)) \quad (\text{zentralisiert})$$
$$\propto p(\mathbf{x}(n-1) | \mathbf{Z}(n-1)) \prod_{i=1}^L \frac{p(\mathbf{x}(n) | \mathbf{z}_i(n))}{p(\mathbf{x}(n-1) | \mathbf{Z}(n-1))} \quad (\text{verteilt})$$

## 2.5.3 Bayes'sche Fusion

### Rekursive Bayes'sche Fusion (verteilt)

Bestimmung von  $x$ , Beobachtung zum Zeitpunkt  $n$ :  $z_i(n)$



## 2.5.3 Bayes'sche Fusion

### Fusion bei Daten über unterschiedliche unabhängige Zielgrößen

Annahme:  $z_1$  sind Daten bzgl. der Zielgröße  $x_1$   
 $z_2$  sind Daten bzgl. der Zielgröße  $x_2$

Sensormodelle:  $x_1 \rightarrow z_1 : p(z_1 | x_1)$   
 $x_2 \rightarrow z_2 : p(z_2 | x_2)$

Fusion:

$$\begin{aligned} p(x_1, x_2 | z_1, z_2) &\propto p(z_1, z_2 | x_1, x_2) p(x_1, x_2) \\ &= p(z_1 | z_2, x_1, x_2) p(z_2 | x_1, x_2) p(x_1, x_2) \end{aligned}$$

Annahme:  $z_1$  und  $z_2$  sind bedingt stochastisch unabhängig  
bei festen  $x_1$  und  $x_2$

$$= p(z_1 | x_1, x_2) p(z_2 | x_1, x_2) p(x_1, x_2)$$

Annahme:  $z_1$  und  $x_2$  bzw.  $z_2$  und  $x_1$  sind bedingt stochastisch  
unabhängig bei festen  $x_1$  bzw.  $x_2$

$$= p(z_1 | x_1) p(z_2 | x_2) p(x_1, x_2)$$

## 2.5.4 Logarithmierung

### Verwendung von logarithmierten WV

- Logarithmen der WV:

$$l(\mathbf{x}) := \log(\Pr(\mathbf{x}))$$

$$l(\mathbf{z} | \mathbf{x}) := \log(\Pr(\mathbf{z} | \mathbf{x}))$$

- Negativität:  $\Pr(\mathbf{x}) \leq 1 \Rightarrow l(\mathbf{x}) \leq 0$

- Satz von Bayes:

$$\Pr(\mathbf{x} | \mathbf{z}) = \frac{\Pr(\mathbf{z} | \mathbf{x})\Pr(\mathbf{x})}{\Pr(\mathbf{z})} \Rightarrow l(\mathbf{x} | \mathbf{z}) = l(\mathbf{z} | \mathbf{x}) + l(\mathbf{x}) - l(\mathbf{z})$$

- Vorteil: Additionen statt Multiplikationen (schnellere Berechenbarkeit)

- Nachteile:

- Logarithmierung erforderlich, daher nur vorteilhaft bei
  - Mehrfacher Verwendung (z.B.  $l(\mathbf{x})$ )
  - Lookup-Tabellen für Verteilungen (z.B.  $l(\mathbf{z} | \mathbf{x})$ )
- Normierung nicht unmittelbar möglich

## 2.5.5 Marginalisierung

„**Herausintegrieren**“ uninteressanter Zielgrößen: **Fokussierung**

Falls nur eine von mehreren Zielgrößen interessiert:

$$p(x_k | \mathbf{z}) = \int_{x_l, l \in \{1, \dots, N\} \setminus \{k\}} p(x_1, \dots, x_k, \dots, x_N | \mathbf{z}) dx_l$$

bzw.

$$\Pr(x_k | \mathbf{z}) = \sum_{x_l, l \in \{1, \dots, N\} \setminus \{k\}} \Pr(x_1, \dots, x_k, \dots, x_N | \mathbf{z})$$

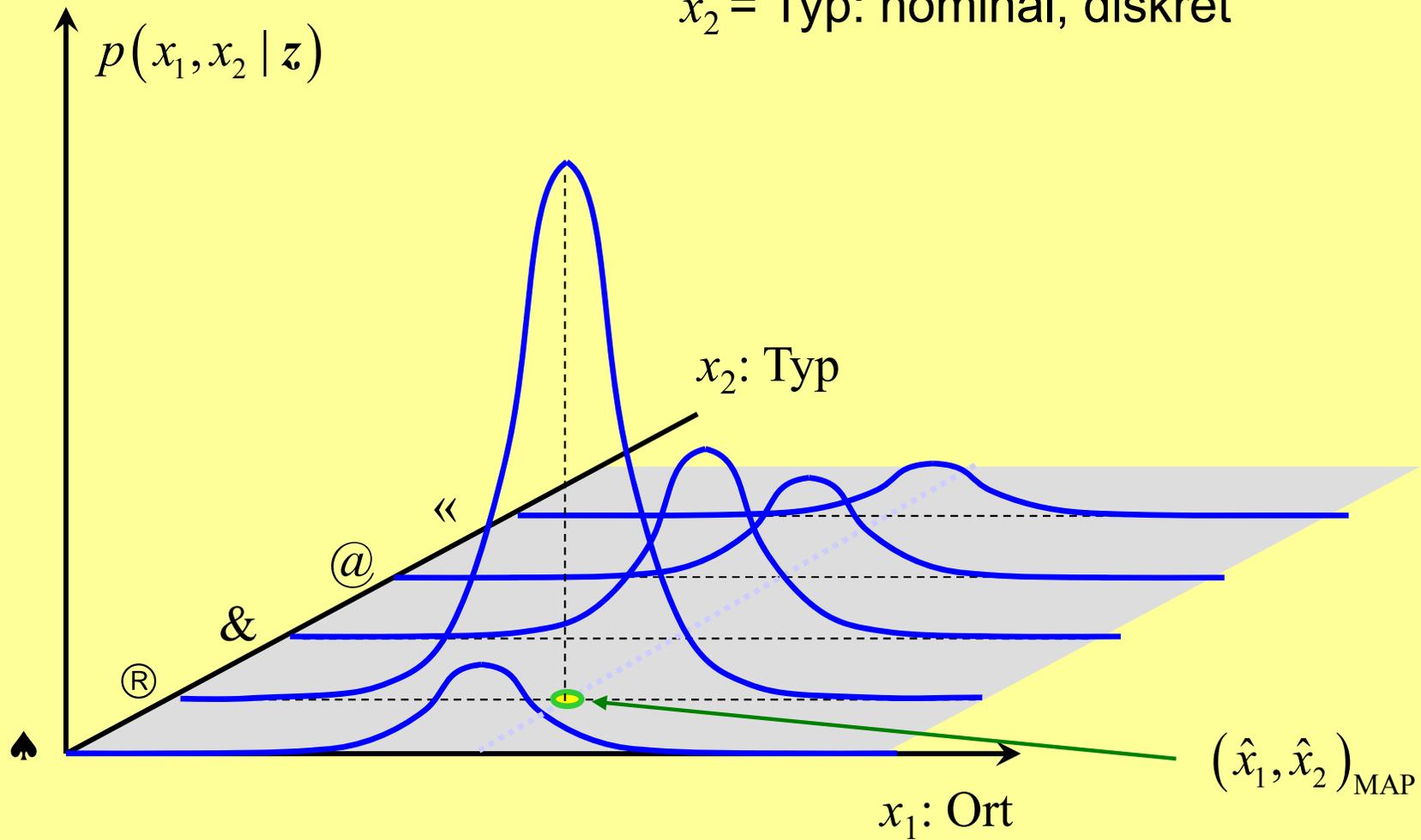
**Marginalisierung** auf den „Rand“ der Verbundverteilung

## 2.5.5 Marginalisierung

### Beispiel: Gemischte Größen

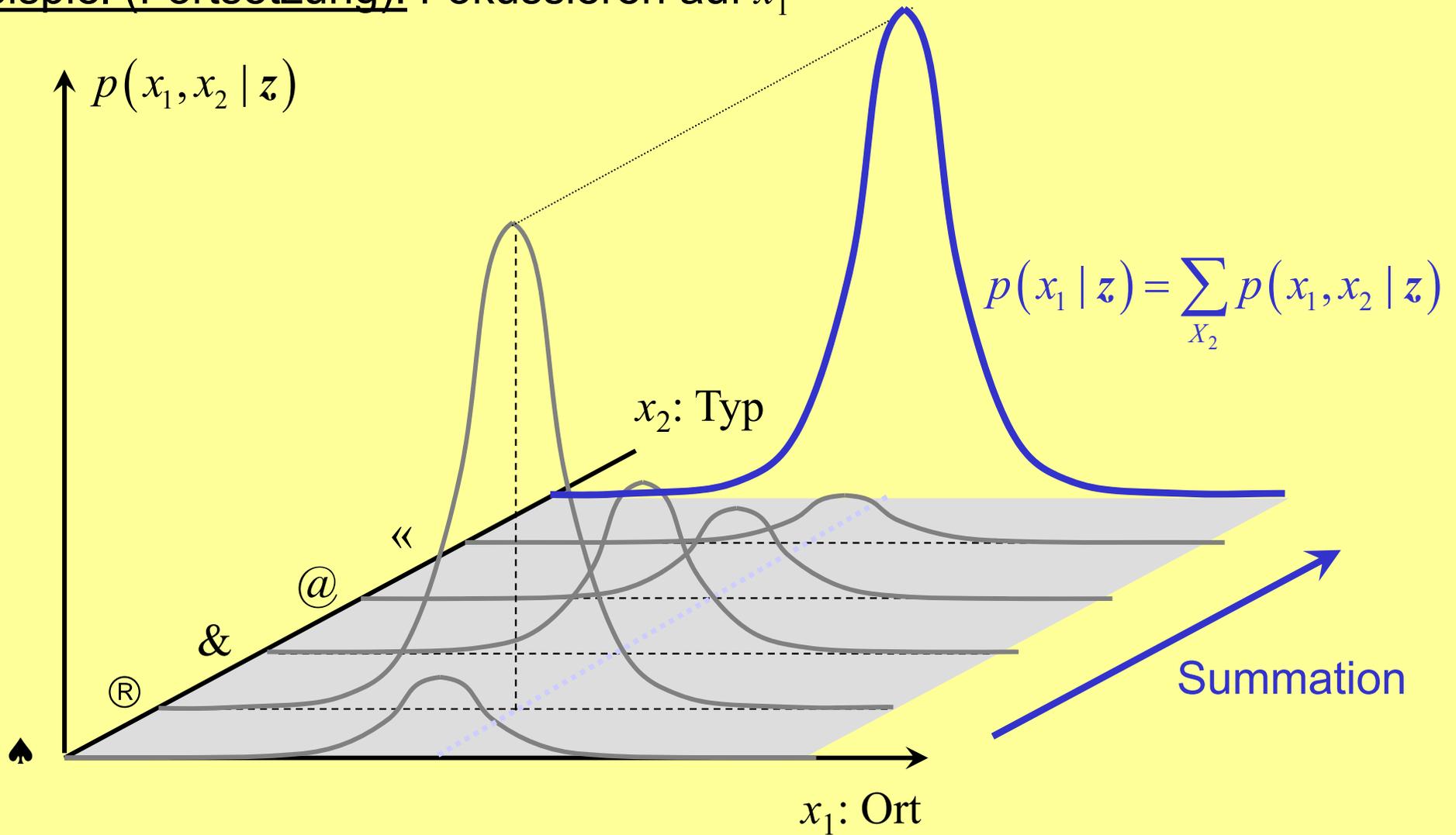
$x_1 = \text{Ort: kardinal, kontinuierlich}$

$x_2 = \text{Typ: nominal, diskret}$



## 2.5.5 Marginalisierung

Beispiel (Fortsetzung): Fokussieren auf  $x_1$



$(p(x_1 | z))$  nicht maßstäblich



## 2.5.6 Bewertung der Bayes'schen Methodik

---

- Vorteile:**
- + Erfüllt Forderungen nach Transformation, Fusion, Fokussierung
  - + Theoretisch gesehen verlustfreie Fusionsmethodik (Resultat: vollständige A-posteriori-Verteilung)
  - + Simultane Handhabung unterschiedlich skalierten Größen
  - + Kein Problem mit unterschiedlichen Abstraktionsniveaus

## 2.5.6 Bewertung der Bayes'schen Methodik

- Probleme:** - Hoher Rechenaufwand zur Bestimmung der Verbundwahrscheinlichkeiten und Marginalisierung  
→ Bei hochdimensionalen Räumen  $X$  bzw.  $Z$  unbrauchbar

### Beispiel:

- Detektion eines Objekts mit  $n$  Sensoren:  $\mathbf{z} = \{z_1, z_2, \dots, z_n\}$ ,  $\mathbf{Z} = \{0, 1\}^n$
- Konstruktion der bedingten Wahrscheinlichkeit  $\Pr(\mathbf{z} | x)$  erfordert  $2^n$  Werte (z.B.  $n = 6$ : 64) durch Beobachtung oder Simulation

### Beispiel:

- Zielgröße mit  $m$  Komponenten:  $\mathbf{x} = \{x_1, x_2, \dots, x_m\}$ ,  $\mathbf{X} = \mathbf{X}_1 \times \dots \times \mathbf{X}_m$
- Zur Bestimmung des Ergebnisses für eine Komponente  $z_i$  Summation bzw. Integration über  $m - 1$  Komponenten, d.h. Dimensionen, erforderlich:

$$p(x_k | \mathbf{z}) = \sum \int_{\underbrace{X_i, i \in \{1, \dots, m\} \setminus \{k\}}}} p(x_1, \dots, x_k, \dots, x_m | \mathbf{z})$$

## 2.5.6 Bewertung der Bayes'schen Methodik

### Abhilfe:

- Vereinfachungen (z.B. Annahme von bedingter stochastischer Unabhängigkeit)

### Beispiel (siehe oben):

- Detektion eines Objekts mit  $n$  Sensoren:  $\mathbf{z} = \{z_1, z_2, \dots, z_n\}$ ,  $\mathbf{Z} = \{0, 1\}^n$
- Annahme d. stochast. Unabhängigkeit:  $\Pr(\mathbf{z} | x) = \Pr(z_1 | x) \cdot \dots \cdot \Pr(z_n | x)$   
Nur noch  $2n$  Werte erforderlich (statt  $2^n$  Werte)

- Hierarchische Vorgehensweise
- Lokale Ansätze

## 2.6 Tracking

---

**Tracking: Nachführen der Zustandsschätzung** eines oder mehrerer dynamischer Objekte

- Objektcharakterisierung: Zustandsgrößen bzw. -vektoren (z.B. Objektlabel, -position, -geschwindigkeit, -beschleunigung)
- Track: Folge der Zustände eines Objekts
- Zielverfolgung (Target Tracking, oft militärischer Kontext)
- Verfolgung von sich im Raum bewegenden Objekten durch Beobachtungen zu diskreten Zeitpunkten  $\{t_0, t_1, t_2, \dots\}$
- Identifikation zusammengehörender Objektbeobachtungen über der Zeit (Assoziation, ggf. mit Vorauswahl: Gating)

## 2.6.1 Bayes'sches Tracking

### Im Bayes'schen Sinne:

- Iteratives Nachführen (Updaten) bedingter Verteilungen  $p(x_m | Z^n)$ :

$t_m > t_n \rightarrow$  Prädiktion

$t_m = t_n \rightarrow$  Innovation

$t_m < t_n \rightarrow$  Retrodiktion

- $x_m$ : Zustand(svektor) der zu verfolgenden Objekte zum Zeitpunkt  $t_m$
- $Z^n$ : bis einschließlich zum Zeitpunkt  $t_n$  verfügbare (Sensor-)daten (inkl. verfügbares A-priori-Wissen)  $Z^n = \{z_0, z_1, \dots, z_n\}$
- $p(x_m | Z^n)$ : verkörpert das gesamte verfügbare Wissen über  $x_m$
- Objekt wird durch seine WV über den Zustandsvektor charakterisiert, keine „harte“ Entscheidung (Soft Decision Prinzip)

## 2.6.1 Bayes'sches Tracking

Berechnung der Verteilungen  $p(x_m | Z^n)$ :

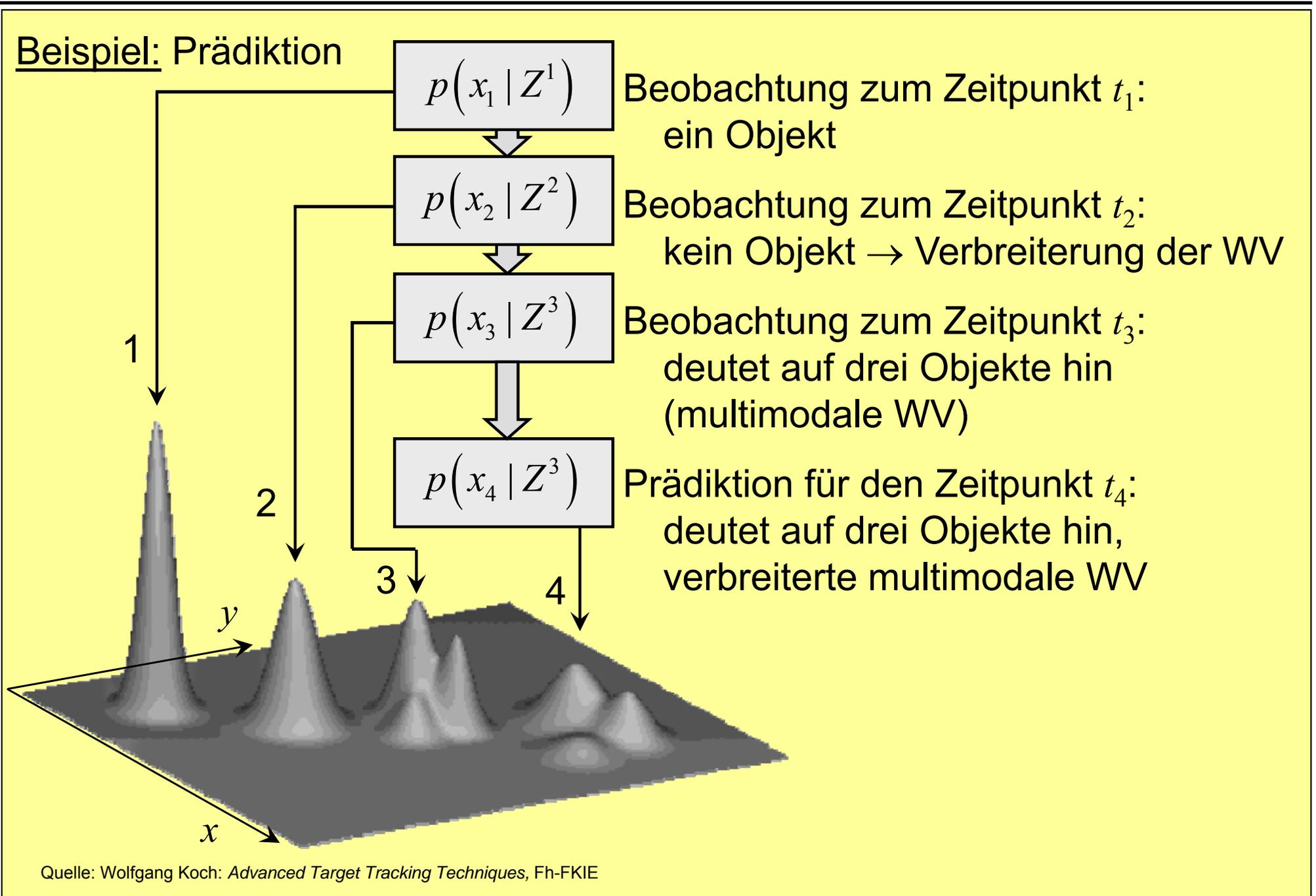
- Initialisierung durch A-priori-WV
- Iterative Schemata:

**Prädiktion**  $t_m > t_n$   $p(x_{k-1} | Z^{k-1})$   $\xrightarrow[\text{Systemunsicherheit}]{\text{Dynamisches Modell}}$   $p(x_k | Z^{k-1})$

**Innovation**  $t_m = t_n$   $p(x_k | Z^{k-1})$   $\xrightarrow[\text{Beobachtungsunsicherheit}]{\text{Aktuelle Sensordaten}}$   $p(x_k | Z^k)$

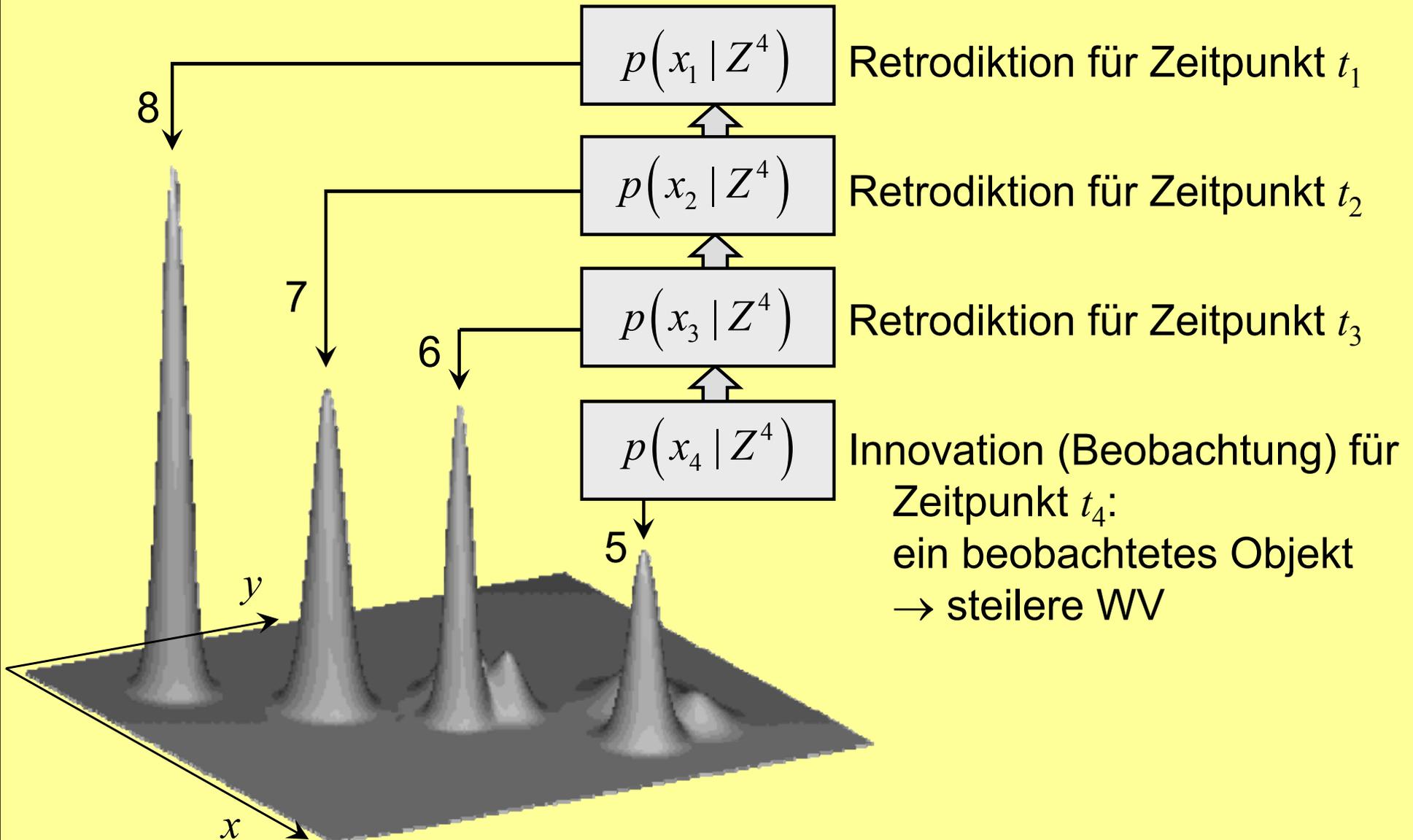
**Retrodiktion**  $t_m < t_n$   $p(x_{k-1} | Z^k)$   $\xleftarrow{\text{Dynamisches Modell}}$   $p(x_k | Z^k)$

## 2.6.1 Bayes'sches Tracking



## 2.6.1 Bayes'sches Tracking

### Beispiel: Innovation und Retrodiktion

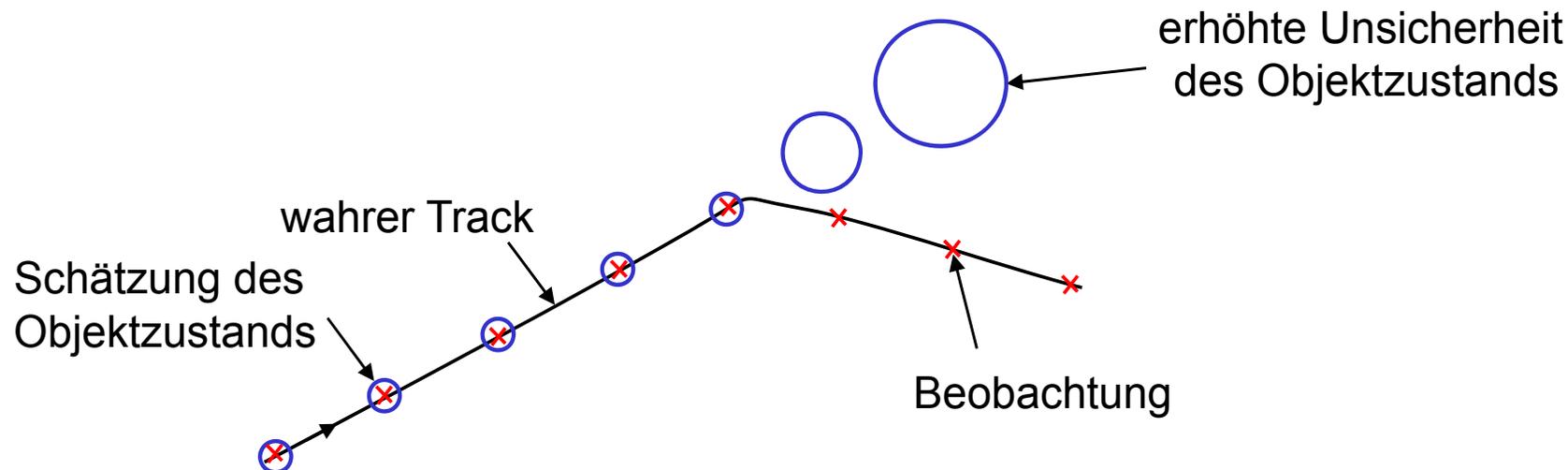


Quelle: Wolfgang Koch: *Advanced Target Tracking Techniques*, Fh-FKIE

## 2.6.2 Tracking-Problematik

### Konfliktsituation – Fehlertypen

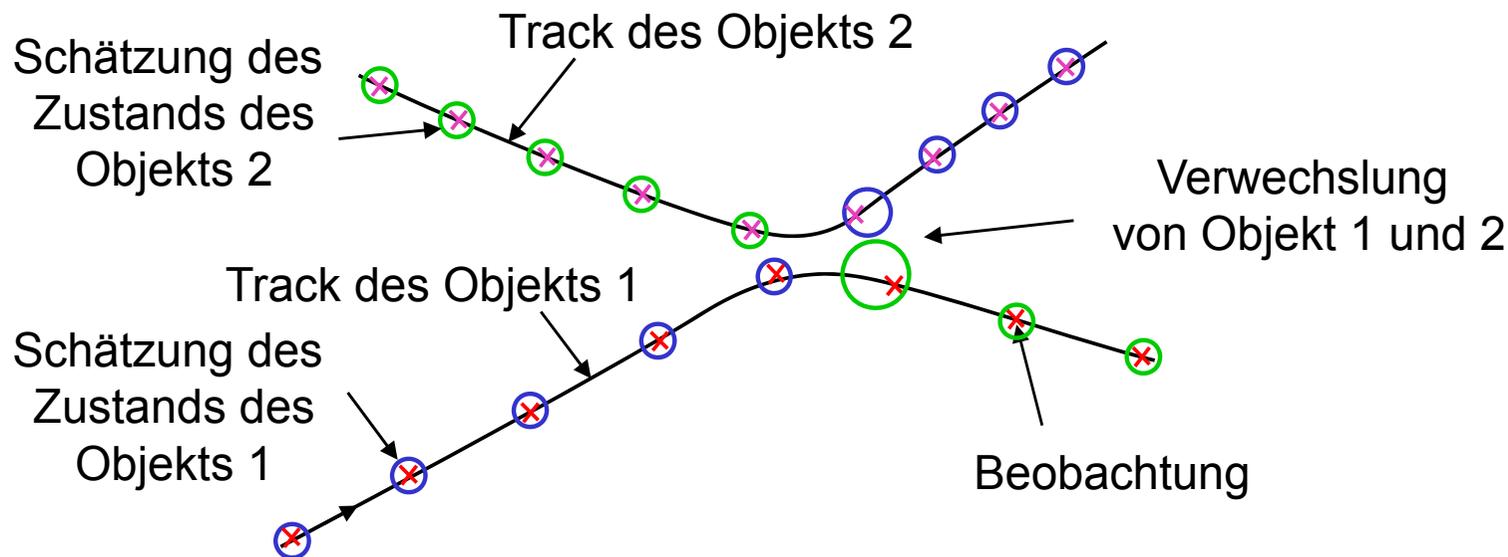
- Grundproblem: Zuordnung von Beobachtungen zu Objekten  
Assoziation Daten – Objekte
- **Fehlertyp 1:** Obwohl eine Beobachtung zu einem Objekt gehört, findet keine Zuordnung statt
  - Ursachen: Beobachtungsunsicherheit, Dynamik des Objekts
  - Folgen:
    - Erhöhung der Unsicherheit des Objektzustands
    - Evtl. Initialisierung eines neuen Tracks



## 2.6.2 Tracking-Problematik

### Konfliktsituation – Fehlertypen

- **Fehlertyp 2:** Eine Beobachtung wird einem Objekt zugeordnet, zu dem sie nicht gehört (Verwechslung)
  - Ursachen: Beobachtungsunsicherheit, Dynamik des Objekts, mehrere benachbarte Objekte
  - Folge: keine korrekten Tracks der Objekte
- Fehlertypen 1 und 2 können sich gegenseitig erzeugen und verstärken



## 2.6.2 Tracking-Problematik

---

### Distanz

- Quantifizierung der Abstände zwischen **prädiziertem Objektzustand (Hypothese)** und **Beobachtungen (Innovation)** durch einen geeigneten **Distanzbegriff**
- Distanz kann **alle Komponenten des Objektzustands** umfassen, nicht nur geometrische Position
- Falls Abstand zwischen Beobachtung und Prädiktion zu groß: **Ausreißer, Geisterziele** oder **neue (interessierende) Objekte**
- Festlegung von **Suchbereichen (Gates)** für die Zuordnung einer Beobachtung zur Prädiktion (z.B. über einen Schwellwert)
- Keine Berücksichtigung von Beobachtungen außerhalb des Suchbereichs → **Reduktion der Komplexität, Erhöhung der Robustheit**
- Identifikation eines neuen Objekts: **neuer Track**
- Evtl. Berücksichtigung weiterer Objekteigenschaften zur Identifikation von Objekten

## 2.6.3 Vorselektion

### Festlegung von Suchbereichen (Gates)

- 1. Möglichkeit:

Suchbereiche **getrennt für jede Komponente** des Zustandsvektors

- **Achsparallele Suchbereiche** (im zweidimensionalen kartesischen Koordinatensystem rechteckige Suchbereiche)

- Abstandsbegriff: 
$$\text{dist}(\hat{\mathbf{x}}, \mathbf{x}) = \max_i \left( \frac{|\hat{x}_i - x_i|}{\sqrt{P_{i,i}}} \right)$$

$P$ : Kovarianzmatrix des Zustandsvektors nach der Innovation  
(Indizes: Vektor-/Matrixkomponenten)

- Keine Berücksichtigung von Kovarianzen  $P_{ij}, i \neq j$ ,  
d.h. wechselseitige Unsicherheiten bzgl. der Komponenten des Zustandsvektors werden ignoriert
- Suchbereich:  $\text{dist}(\hat{\mathbf{x}}, \mathbf{x}) \leq k$   
z.B.  $k \geq 3$  sorgt bei Normalverteilung der Komponenten des Zustandsvektors für Lage von 99,7% der Beobachtungen innerhalb des Suchbereichs

## 2.6.3 Vorselektion

### Festlegung von Suchbereichen (Gates)

- 2. Möglichkeit:

#### Ellipsoidische Suchbereiche

- Ergeben sich bei Berücksichtigung der Kovarianzmatrix
- Unsicherheitsbereich wird nicht von den Koordinatenachsen (der Komponenten des Zustandsvektors), sondern von den **Eigenvektoren der Kovarianzmatrix des Zustandsvektors** nach dem Innovationsschritt aufgespannt

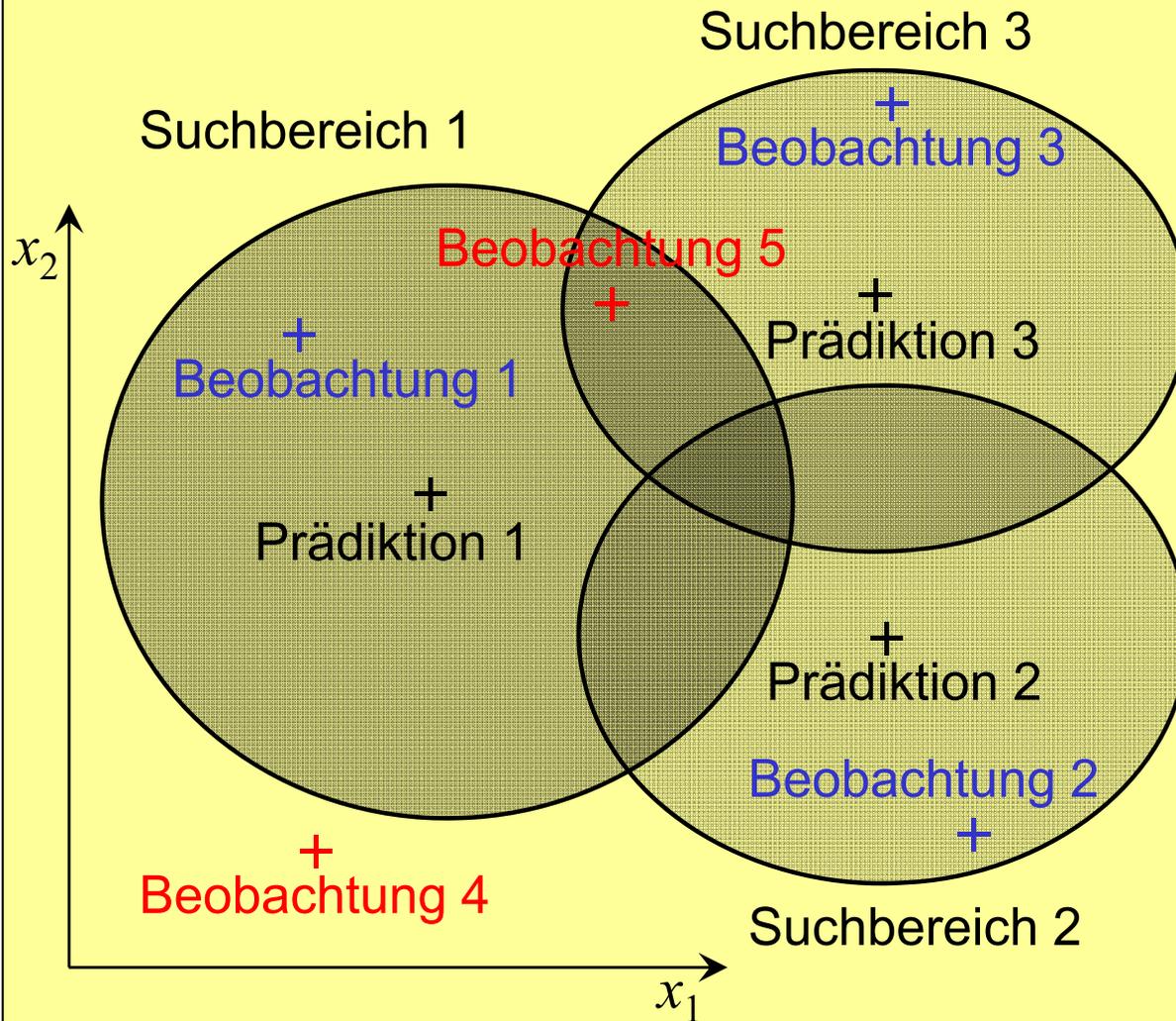
- Abstandsbegriff: z.B. Mahalanobis-Distanz:

$$\text{dist}(\hat{\mathbf{x}}, \mathbf{x}) = \sqrt{(\hat{\mathbf{x}} - \mathbf{x})^T P^{-1} (\hat{\mathbf{x}} - \mathbf{x})}$$

- Invariant bei Skalierung, Translation
- Punkte gleichen Abstands im Zweidimensionalen: Ellipsen
- für  $P = I$  (unkorrelierte Komponenten, gleiche Unsicherheiten 1): Euklidischer Abstand

## 2.6.3 Vorselektion

### Beispiel: Suchbereiche und Konfliktsituationen



- Beobachtungen 1 bis 3: Eindeutig einem Suchbereich zuordenbar
- Beobachtung 4: Neues (interessierendes) Objekt oder Fehler vom Typ 1
- Beobachtung 5: Fehlzuordnung möglich (Fehlertyp 2), kann evtl. in weiteren (Zeit-) Schritten aufgelöst werden (Multi-Hypothesen-Tracking, s.u.)

Quelle: Dirk Stüker: *Heterogene Sensordatenfusion zur robusten Objektverfolgung im automobilen Straßenverkehr*, Dissertation. Carl von Ossietzky-Universität Oldenburg, 2004

## 2.6.4 Hypothesen-Tracking

---

- Bisherige Vorselektion zur Auflösung von Zuordnungskonflikten: nur ein Objekt und ein Zeitschritt wird berücksichtigt
- Falls immer noch Zuordnungskonflikte existieren: **Berücksichtigung aller Objekte und der Objekt-Historie** zur Auflösung der Zuordnungsproblematik
- Hypothese: mögliche Zuordnung von Beobachtungen zu Objekten
- Unterscheidung in Einfach- und Multi-Hypothesen-Tracking

## 2.6.4 Hypothesen-Tracking

---

- **Einfach-Hypothesen-Tracking (EHT):**
  - Zuordnung aller Beobachtungen zu den Objekten erfolgt in jedem Iterationsschritt neu
  - Zusammenfassung aller während einer Iteration betrachteten Hypothesen zu einer Gesamthypothese
  - Prüfung der Gesamthypothese
  - Keine Berücksichtigung der Historie: „Zero-Scan-Algorithmen“
  
- **Multi-Hypothesen-Tracking (MHT):**
  - Bei Zuordnungskonflikten während einer Iteration: Verfolgung von Hypothesen über mehrere Iterationen
  - Konflikte werden während der folgenden Iterationen aufgelöst: Verzögerte Entscheidung
  - Höhere Komplexität als Einfach-Hypothesen-Tracking
  - „Multiple-Scan-Algorithmen“

## 2.6.4 Hypothesen-Tracking

---

### Multi-Hypothesen-Tracking

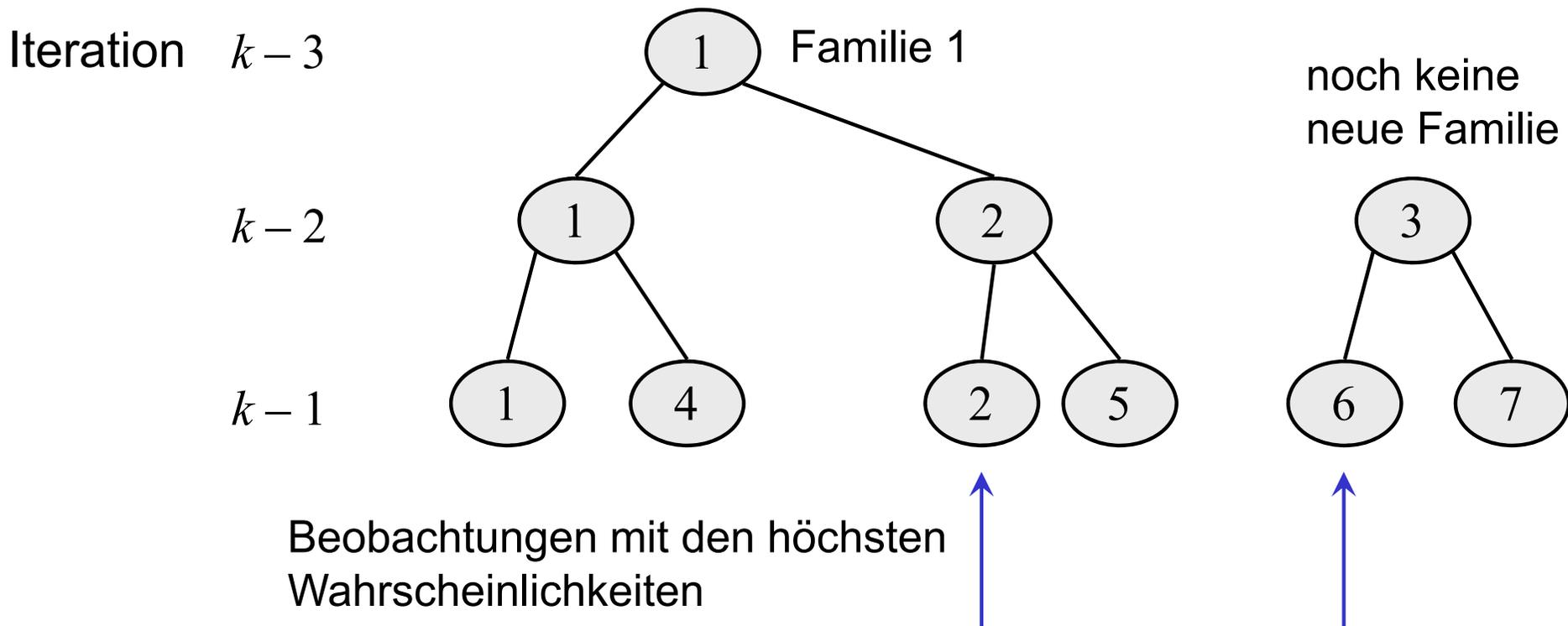
- Aufnahme **aller Assoziationshypothesen** in einen Hypothesenbaum
- Entscheidung über zutreffende Hypothese erfolgt erst dann, wenn hinreichende Evidenz vorliegt (z.B. nach 5 Zeitschritten)
- Problematisch: mit Anzahl der zu berücksichtigenden Iterationen **exponentiell wachsender Hypothesenbaum**  
→ Verfolgung aller Hypothesen über längere Zeit oft nicht möglich
- Mögliche Abhilfen:
  - Entfernung von Hypothesen mit Wahrscheinlichkeiten unterhalb eines Schwellwertes
  - Kombination von Hypothesen mit ähnlichen Zuordnungen
  - Entfernung von Zweigen, die nicht zur Hypothese mit den aktuell höchsten Wahrscheinlichkeiten führen (siehe nächste Folie)

## 2.6.4 Hypothesen-Tracking

### Multi-Hypothesen-Tracking

Entfernung von Zweigen aus dem Hypothesenbaum

hier: Entscheidungen werden über 2 Iterationen hinweg getroffen

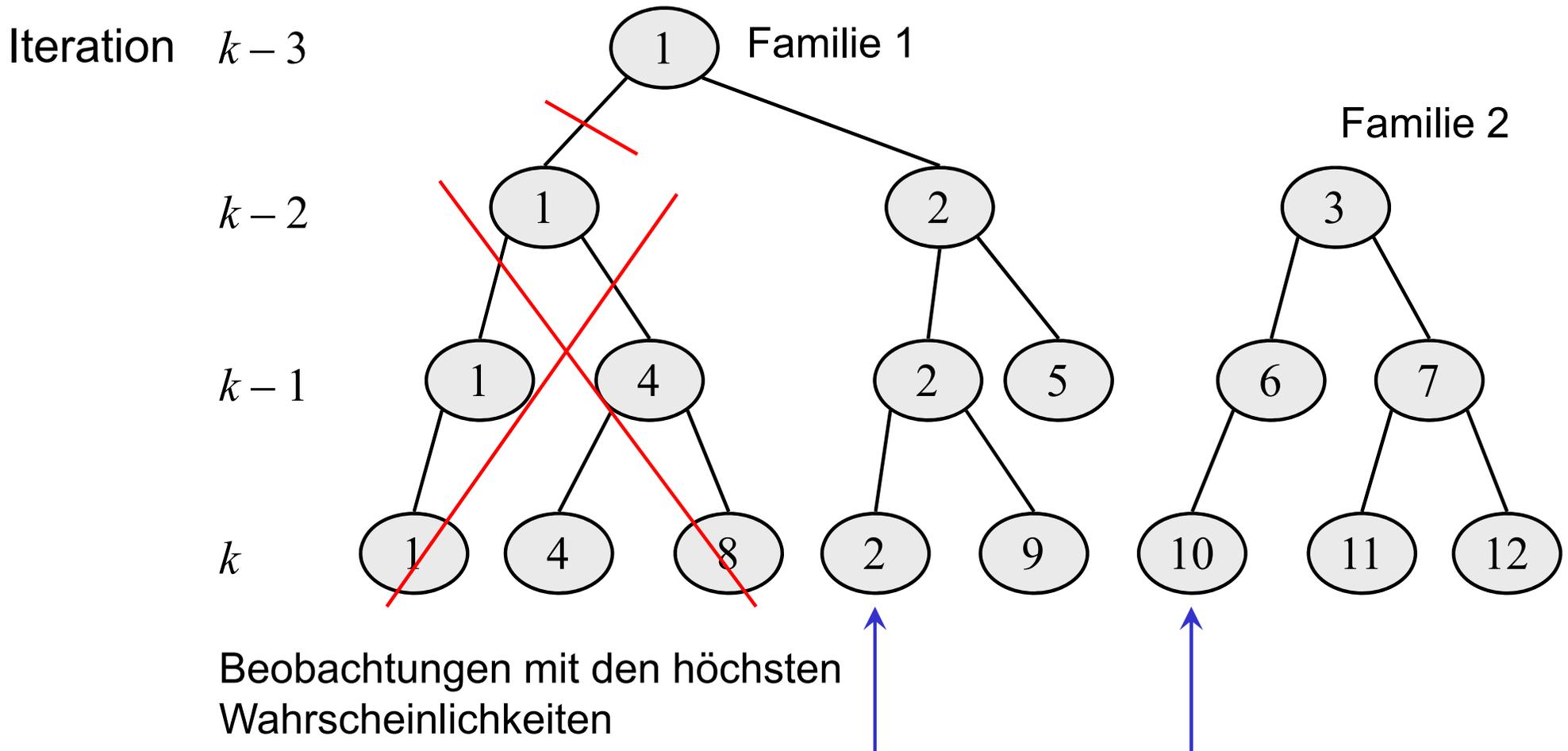


Quelle: S. S. Blackman: *Multiple hypothesis tracking for multiple target tracking*. Aerospace and Electronic Systems Magazine, IEEE, Vol. 19, Issue 1, Part 2, Jan 2004, S. 5-18

## 2.6.4 Hypothesen-Tracking

### Multi-Hypothesen-Tracking

Entfernung von Zweigen aus dem Hypothesenbaum  
neue Beobachtungen, aktualisierter Baum



Quelle: S. S. Blackman: *Multiple hypothesis tracking for multiple target tracking*. Aerospace and Electronic Systems Magazine, IEEE, Vol. 19, Issue 1, Part 2, Jan 2004, S. 5-18

- Berger, James O.: *Statistical Decision Theory and Bayesian Analysis*. 2nd edition, Springer, 2006.
- Bernardo, José M.: *Bayesian theory*. Wiley, 2000.
- Berthold, Michael; Hand, David J. (Hrsg.): *Intelligent data analysis: an introduction*. Springer, 2003.
- Blackman, Samuel S.: *Multiple hypothesis tracking for multiple target tracking*. Aerospace and Electronic Systems Magazine, IEEE, Vol. 19, Issue 1, Part 2, Jan 2004, S. 5-18.
- Hall, David L.: *Mathematical Techniques in Multisensor Data Fusion*. 2. Aufl., Artech House, 2004.
- Henze, Norbert: *Stochastik I : Einführung in die Wahrscheinlichkeitstheorie und Statistik*. Skriptum, Universität Karlsruhe, 1995.
- Henze, Norbert: *Stochastik für Einsteiger*. Vieweg, 1999.
- Klein, Lawrence A.: *Sensor and Data Fusion – A Tool for Information Assessment and Decision Making*. SPIE, 2004.

- Korb, Kevin: *Bayesian artificial intelligence*. Chapman & Hall, 2004.
- Stüker, Dirk: *Heterogene Sensordatenfusion zur robusten Objektverfolgung im automobilen Straßenverkehr*. Dissertation. Carl von Ossietzky-Universität Oldenburg, 2004.
- Welch, Greg; Bishop, Gary: <http://www.cs.unc.edu/~welch/kalman/>