# Lineare optimale Regelung und Schätzung mittels Message Passing auf Faktorgraphen

Christian Hoffmann * Philipp Rostalski *

* Institut für Medizinische Elektrotechnik
Universität zu Lübeck
Ratzeburger Allee 160
+49 451 3101 6200
Philipp.Rostalski@uni-luebeck.de
Christian.Hoffmann@uni-luebeck.de

**Schlüsselwörter:** Message Passing, Probabilistische Modelle, LQG Regelung, Faktorgraphen, Kalman Filterung, Expectation Maximization, Iterativ Lernende Regelung

Faktorgraphen bilden eine Klasse probabilistischer graphischer Modelle, welche die Faktorisierung von Wahrscheinlichkeitsdichteverteilungen als bipartite Graphen abbilden [1, 2]. Eine solche Zerlegung kann genutzt werden, um Aussagen bedingter Unabhängigkeiten darzustellen und probabilistische Inferenzprobleme zu lösen. Dazu werden häufig sogenannte Max-Product und Sum-Product Algorithmen [3] verwendet, um Nachrichten auf den Kanten des Graphen zu berechnen, die eine effiziente Zusammenfassung aller vorangegangenen Nachrichten darstellen. So kann die Struktur des Graphen ausgenutzt werden, um Inferenzprobleme mittels der Übermittlung von Nachrichten zwischen den Knoten des Graphen effizient zu lösen bzw. zu approximieren.

Weitere bekannte graphische Darstellungen wie Bayes'sche Netze, Hidden Markov Modelle und Markov Random Fields sind Spezialfälle von Faktorgraphen [3]. Folglich lassen sich in diesem Framework viele Algorithmen aus sehr unterschiedlichen Forschungsfeldern wie der Statistik, der Kodierungstheorie, dem Maschinellen Lernen und der Signalverarbeitung darstellen [4]. Viele dieser Algorithmen – darunter der Viterbi Algorithmus, die iterative Dekodierung von Turbo Codes, Low-Density Parity-Check Codes, Expectation Maximization, sowie das Kalman Filter – sind wegweisend für ihre Forschungsfelder. Eine Hauptmotivation der Forschung an Faktorgraphen stellt daher das Potenzial dar, viele vermeintlich heterogene Algorithmen in einem vereinheitlichenden graphischen Modell kombinieren zu können, um so ganzheitliche Lösungen für komplexe Probleme zu finden. In einer Faktorgraphen-Darstellung bieten Algorithmen zudem meist einen intuitiven Zugang, weshalb das Framework auch aus didaktischen Gesichtspunkten interessant ist.

Der Vortrag zeigt daher zunächst, wie das Kalman Filter mittels Faktorgraphen dargestellt werden kann [1]. Der entsprechende Faktorgraph ist in Abbildung 1 dargestellt. Kombiniert mit der Darstellung des Expectation Maximization Algorithmus' mittels Faktorgraphen [5] lässt sich das Kalman Filter zu einem Ansatz für die Identifikation nicht-linearer Systeme in Form affin linear parameter-veränderlicher Darstellungen [6] erweitern. Die graphische Formulierung macht dabei eine rekursive Implementierung des Algorithmus' offensichtlich, die es erlaubt sowohl bekannte Parameterabhängigkeiten zu berücksichtigen, als auch langsamere unbekannte Veränderungen von Parametern zu verfolgen [7].

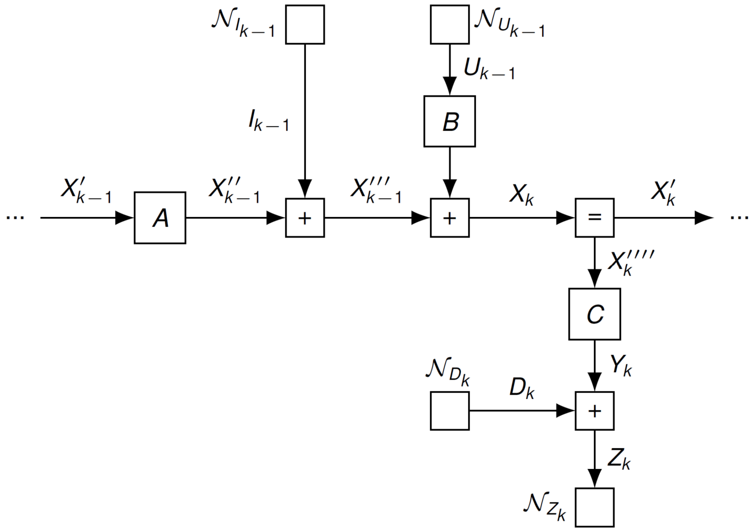



Abbildung 1 - Darstellung des Kalman Filters/Smoothers als Faktorgraph.

Weiterhin zeigt dieser Vortrag auf, inwiefern graphische Modelle in der Form von Faktorgraphen auch im Zusammenhang mit der Regelung nützlich sein können. So wird eine stochastische Interpretation optimaler Regelung [8] mittels Faktorgraphen präsentiert, auf Basis derer eine Reihe klassischer Ergebnisse hergeleitet werden können. Unter anderem, erlaubt es der probabilistische Ansatz hier zudem, die Varianz des durch einen Kalman Filter geschätzten Zustandes in der Regelung zu berücksichtigen. Weiterhin bietet eine einfache Reinterpretation der Faktorgraphen-Darstellung die Möglichkeit zur Implementierung iterativ lernender Regelungs-konzepte. Beispielhaft werden zudem Erweiterungen auf Basis einer Faktorgraphen-Darstellung des Expectation Maximization Algorithmus‘ erläutert, die es erlauben, die Sequenz optimaler Stellgrößen zu regularisieren [9].

Der zugrundeliegende Message Passing Ansatz verspricht neben zahlreichen Optionen zur Kombination von Algorithmen vor allem die Möglichkeit, Algorithmen auf natürliche Weise in separate Recheneinheiten zu modularisieren, um so neue Ansätze für Algorithmen zur verteilten Schätzung und Regelungen zu finden.

**Literatur:**


[1] H.-A. Loeliger, “An Introduction to factor graphs,” (en), *IEEE Signal Process. Mag.*, vol. 21, no. 1, pp. 28–41, 2004.

[2] G. D. Forney, “Codes on graphs: Normal realizations,” (en), *IEEE Trans. Inform. Theory*, vol. 47, no. 2, pp. 520–548, 2001.

[3] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, “Factor graphs and the sum-product algorithm,” (en), *IEEE Trans. Inform. Theory*, vol. 47, no. 2, pp. 498–519, 2001.

[4] H.-A. Loeliger, “An Introduction to Factor Graphs,” 2008.

[5] J. Dauwels, A. W. Eckford, S. Korl, and H.-A. Loeliger, “Expectation Maximization as Message Passing - Part I: Principles and Gaussian Messages,” *arXiv,* arXiv:0910.2832, 2009.

[6] F. Wu, “Control of Linear Parameter Varying Systems,” Ph.D. Thesis, Mechanical Engineering, University of Berkeley, California, USA, 1995.

[7] C. Hoffmann, A. Isler, and P. Rostalski, “A Factor Graph Approach to Parameter Identification for Affine LPV Systems,” *American Control Conference*, 2017 (submitted).

[8] M. Kárný, “Towards fully probabilistic control design,” *Automatica*, vol. 32, no. 12, pp. 1719–1722, 1996.

[9] C. Hoffmann and P. Rostalski, “Linear Optimal Control on Factor Graphs - A Message Passing Perspective,” in *20th IFAC World Congress*, 2017 (submitted).

# Lineare optimale Regelung und Schätzung mittels Message Passing auf Faktorgraphen

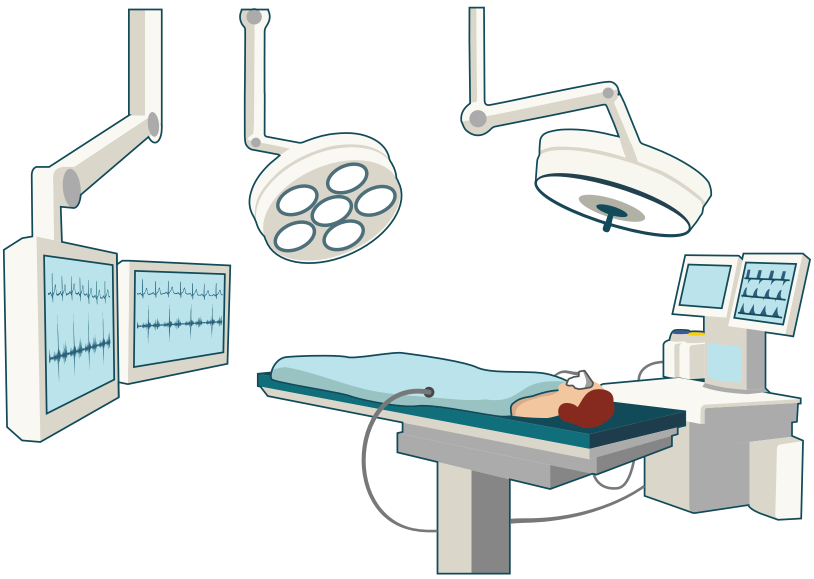
Christian Hoffmann

Inst. für Medizinische Elektrotechnik

23. Februar 2017

---

## Anforderungen moderner Medizinischer Assistenzsysteme



- Zusammenwirken verteilter Algorithmen
- Optimale Steuerung und Regelung
- Signalverarbeitung auf Messdaten
- Kommunikation zwischen verteilten intelligenten Teilsystemen

# Inhalt







1. Faktorgraphen und Gaußsches Message Passing

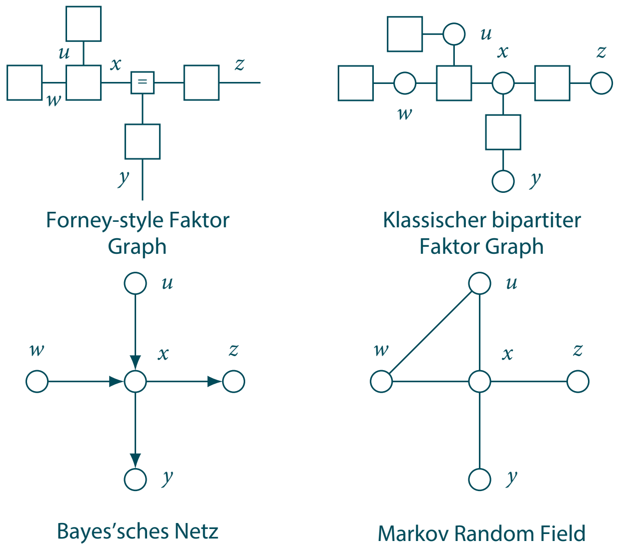
2. Message Passing für Zustands- und Parameterschätzung

3. Regelung mittels Faktorgraphen

4. Zusammenfassung

C. Hoffmann (Inst. für Medizinische Elektrotechnik) Regelung/Schätzung auf Faktorgraphen 23. Februar 2017 3 / 22

# Graphische probabilistische Modelle
## Darstellung bedingter Unabhängigkeit (Loeliger 2004)




### Visualisierung der Faktorisierung von

$$f(u,w,x,y,z) = f_1(u)f_2(w)f_3(x \mid y,z)f_4(y \mid x)f_5(z \mid x)$$

### Vorteile von Forney-style Faktorgraphen

- Geeignet für hierarchische Modellierung
- Einfachste Formulierung der Summary-Product Message Passing Regeln





# Message Passing auf Faktorgraphen
## Ein einheitlicher Ansatz für eine Vielzahl praktischer Algorithmen

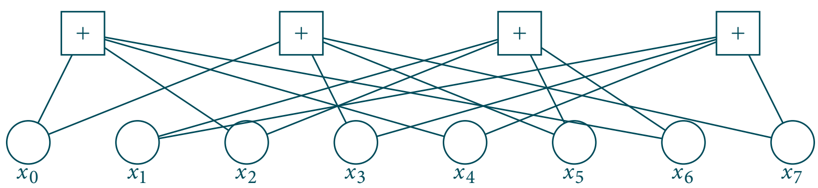



Abbildung: Faktor Graph eines LDPC Codes (Kschischang, Frey und Loeliger 2001)

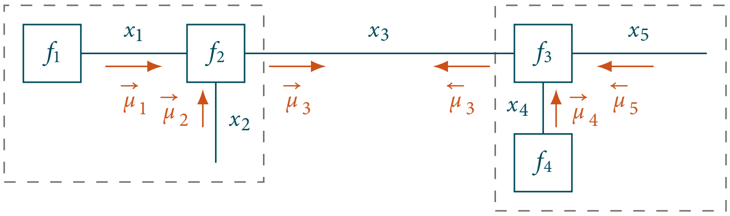
### Viele Problemklassen sind als statistische Inferenzprobleme interpretierbar

- Filterung und Smoothing (z.B. der Kalman oder Modified Bryson–Frazier Smoother (Kalman 1960))
- Fehlerkorrigierte Dekodierung (z.B. der Viterbi Algorithmus (Viterbi 1967))
- Iterative Dekodierung (z.B. Turbo codes, Low-density parity-check codes, ... (Tanner 1981))
- Parameterschätzung (z.B. Expectation Maximization (Dempster, Laird und Rubin 1977))
- Regelung (z.B. Linear Quadratic Gaussian Control(Kwakernaak und Sivan 1972))

# Berechnung von Randwahrscheinlichkeiten

## Idee der Summary-Product Messages

Ziel: Berechnung von $\bar{f}_k(x_k) = \underset{\substack{x_1,\ldots,x_n \\ \text{außer } x_k}}{\iint} f(x_1,\ldots,x_n)\,\mathrm{d}x_1\cdots\mathrm{d}x_n$




Berechnung mittels Sum-Product Algorithmus:

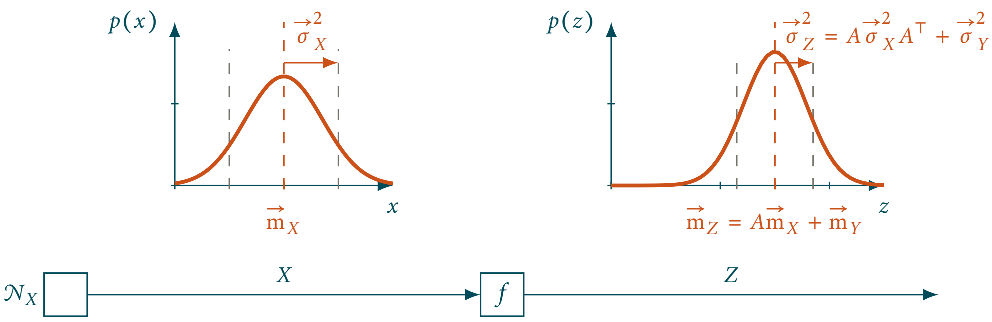
$$\bar{f}(x_3) = \underbrace{\iint_{x_1,x_2} \overbrace{f_1(x_1)}^{\vec{\mu}_1(x_1)} f_2(x_1,x_2,x_3) \overset{\vec{\mu}_2}{\overset{\downarrow}{1}} \,\mathrm{d}x_1\mathrm{d}x_2}_{\vec{\mu}_3(x_3)} \underbrace{\iint_{x_4,x_5} f_3(x_3,x_4,x_5) \overbrace{f_4(x_4)}^{\vec{\mu}_4(x_4)} \overset{\vec{\mu}_5}{\overset{\downarrow}{1}} \,\mathrm{d}x_4\mathrm{d}x_5}_{\overleftarrow{\mu}_3(x_3)}$$





# Gaußsches Message Passing

## Message Passing mit deterministischen Funktionen

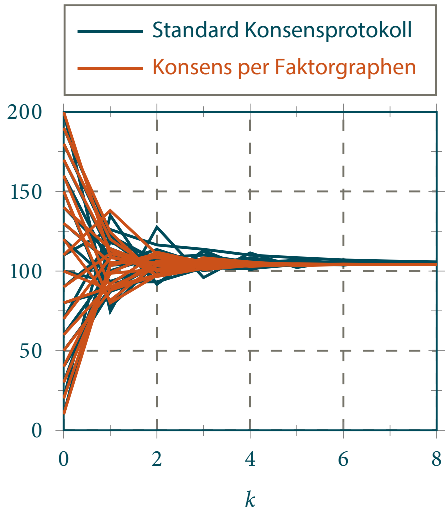



$$f(x) = \delta(z - Ax + y) \quad \Longleftrightarrow \quad z = Ax + y$$

- Tabellierung von Message Passing Regeln für Verteilungsparameter

# Faktorgraphen mit Zyklen

## Beispiel: Konsens-Algorithmen via Loopy Belief Propagation (Moallemi und van Roy 2006)




### Alternative: «Consensus Propagation»

- Spezialfall des Sum-Product Algorithmus
- Kann auch asynchron ausgeführt werden
- Auf regulären Graphen:
  - Schnellere Konvergenz als gewöhnliche, lineare und synchrone Konsensprotokolle





# LPV-Systeme und Gaußsches Message Passing Framework

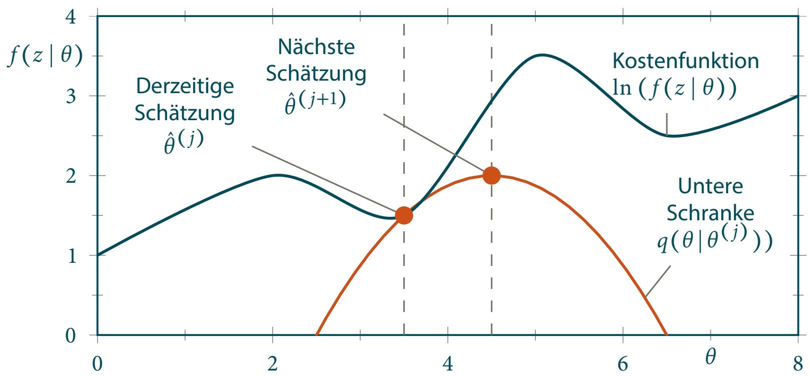
- Allgemeine Form eines linear parameter-veränderlichen (LPV) Systems:

$$\mathcal{G}_\Delta : \begin{cases} \begin{bmatrix} x_{k+1} \\ \hdashline y_k \end{bmatrix} = \left[\begin{array}{c:c} \mathcal{A}(\Delta_k) & \mathcal{B}_u(\Delta_k) \\ \hdashline \mathcal{C}_y(\Delta_k) & \mathcal{D}_{yu}(\Delta_k) \end{array}\right] \begin{bmatrix} x_k \\ \hdashline u_k \end{bmatrix} \\ \Delta_k \in \mathbf{\Delta}, \end{cases}$$

| | |
|---|---|
| Lineare Struktur der Modelle: | Erhaltung der Gaußschen Form |
| Extended Kalman Filter: | LPV-Darstellung |
| Parameterschätzung: | Strukturierte Vorannahmen |

# Expectation Maximization Algorithmus

**(Dempster, Laird und Rubin 1977)**

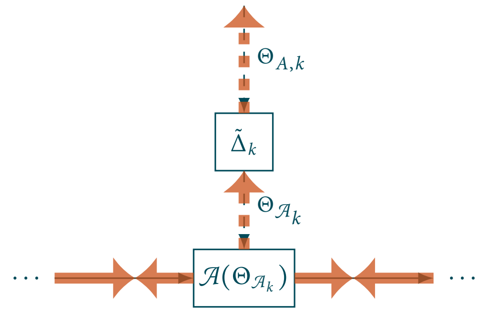



Optimierungsziel: Finde $\hat{\theta} \triangleq \underset{\theta}{\operatorname{argmax}}\, f(z \mid \theta) = \underset{\theta}{\operatorname{argmax}}\, \ln(f(z \mid \theta))$

| | | |
|---|---|---|
| Expectation-Schritt: | $Q\left(\theta \vert \hat{\theta}^{(j)}\right) = \mathrm{E}_{p_X}\left[\ln\left(f(X, Z \vert \theta)\right)\right]$ | «Berechung der Verteilung latenter Variablen als Funktion von $\theta$, gegeben $\hat{\theta}^{(j)}$ » |
| Maximization-Schritt: | $\hat{\theta}^{(j+1)} = \underset{\theta}{\operatorname{argmax}}\, Q\left(\theta \vert \hat{\theta}^{(j)}\right)$ | «Verbesserung der Modellparameter » |

# EM-basierte Schätzung von LPV-Systemparametern

**(Hoffmann, Isler und Rostalski 2017)**




- Expectation Maximization auf Faktorgraphen für lineare Systeme (Dauwels, Korl und Loeliger 2005)
- Vektorisierung der Faktor-Matrix in einen Parametervektor: $\overrightarrow{m}_\Theta \triangleq \mathrm{rvect}\,(A)$
- Schätzung des Faktors erfordert Marginalisierung der anliegenden Variablen $X$ und $Y$.
- Hochwärtige Gaußsche EM-Message Update Regeln:

$$\overleftarrow{W}_\Theta = W_Y \otimes \left(V_X + m_X m_X^\top\right), \qquad \overleftarrow{\xi}_\Theta = (W_Y \otimes I)\,\mathrm{cvect}\left(V_{XY^\top} + m_X m_Y^\top\right)$$

mit

$$V_{XY^\top} = \overrightarrow{V}_X A^\top \left(\overrightarrow{V}_Y + \overleftarrow{V}_Y\right)^{-1} \overleftarrow{V}_Y.$$

- Verwendung von $\mathrm{cvect}(MN) = (N^\top \otimes I)\mathrm{cvect}(M)$
- Faktorisierung:

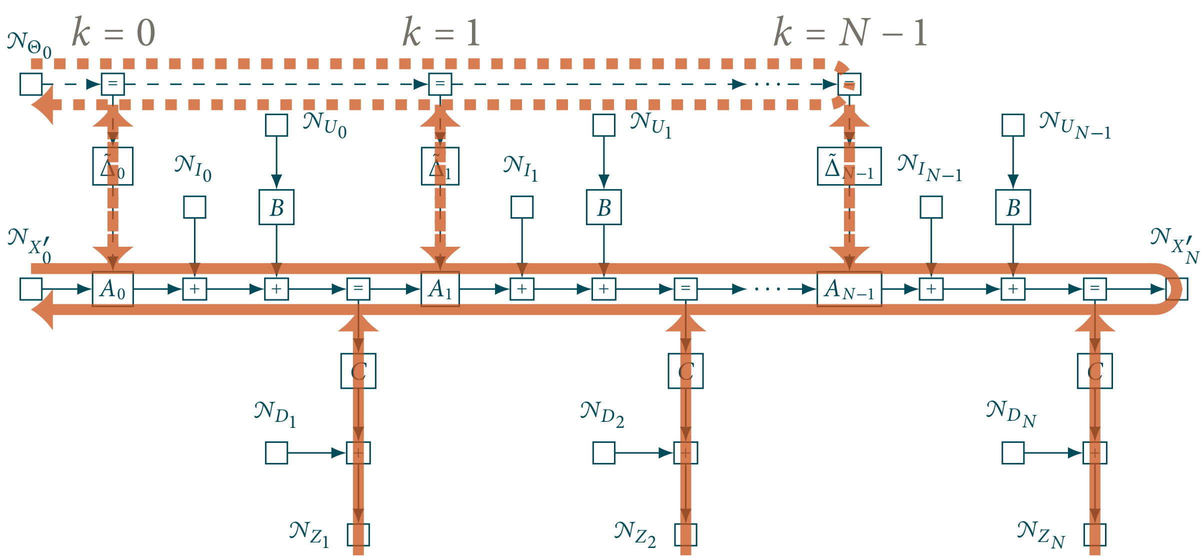
$$\theta_{\mathcal{A}_k} = \mathrm{cvect}\,(\mathcal{A}\,(\Delta_k)) = \mathrm{cvect}\left(\begin{bmatrix}A & B_\Delta\end{bmatrix}\begin{bmatrix}I \\ \Delta_k C_\Delta\end{bmatrix}\right) = \underbrace{\left[\begin{bmatrix}I \\ \Delta_k C_\Delta\end{bmatrix}^\top \otimes I_{n_x}\right]}_{\tilde{\Delta}_k} \underbrace{\mathrm{cvect}\left(\begin{bmatrix}A & B_\Delta\end{bmatrix}\right)}_{\theta_A}$$





# EM-basierte Schätzung affiner LPV-Systeme

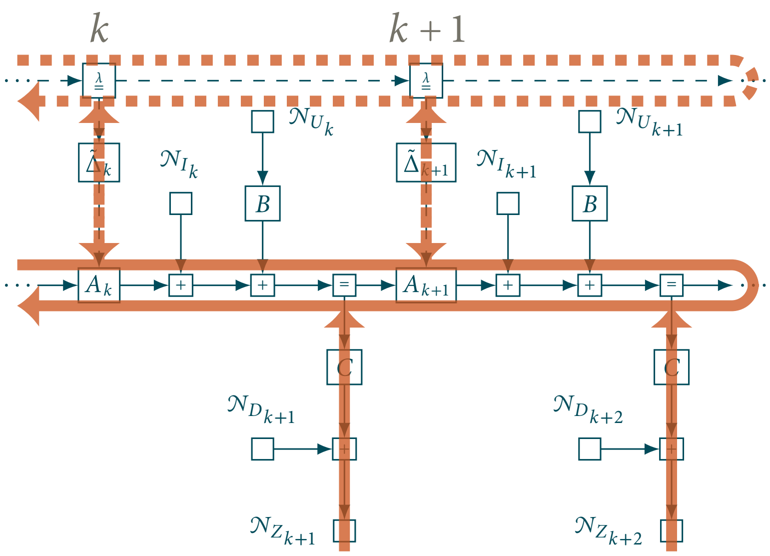
## Message Passing Schedules (Hoffmann, Isler und Rostalski 2017)




- Offline Message Passing Schedule zur Identifikation des Systems

# EM-basierte Schätzung affiner LPV-Systeme

## Message Passing Schedules (Hoffmann, Isler und Rostalski 2017)



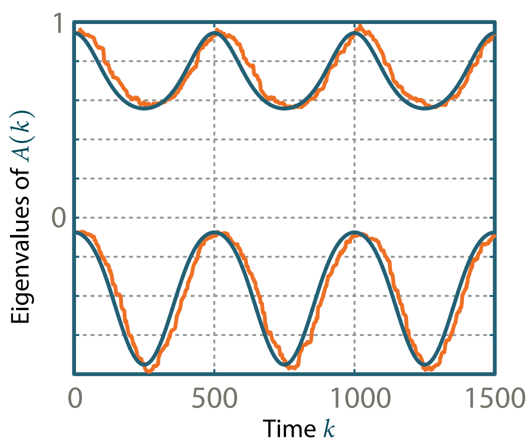

- Online Message Passing Schedule zur rekursiven Identifikation des Systems
- «Forgetting Factor» $\lambda$ degradiert vergangene Informationen





# EM-basierte Identifikation affiner LPV-Systeme

## Numerisches Beispiel - Rekursive Schätzung (Hoffmann, Isler und Rostalski 2017)



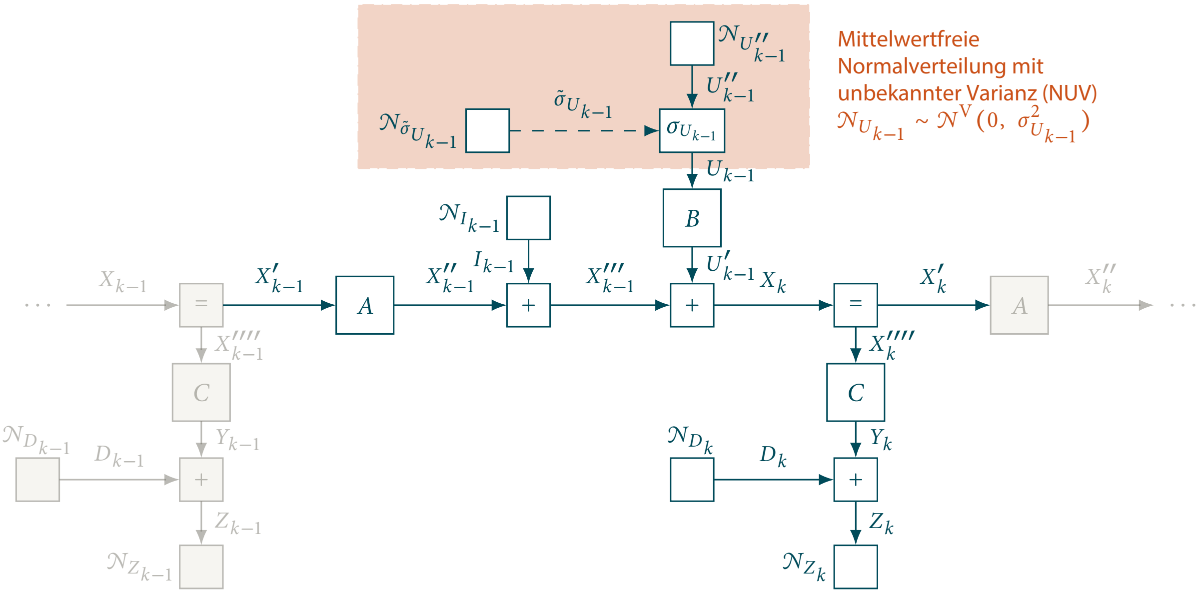

- Bilineare Systemdynamik
- Zusätzlich langsam zeit-variable nominelle Systemmatrix

$$A(k) = \begin{bmatrix} 0.5 \cdot \cos\left(\frac{k\pi}{250}\right) & 1.1 \\ 0.2 & 0.3 \end{bmatrix}$$

- Eigenwerte der Systemmatrix $A$
  - Wahre Werte (——)
  - Geschätzte Werte (——)

# Regularisierte Eingangsschätzung mittels Faktor Graphen

## Schätzung unbekannter Eingangsvarianz (Loeliger u. a. 2016)



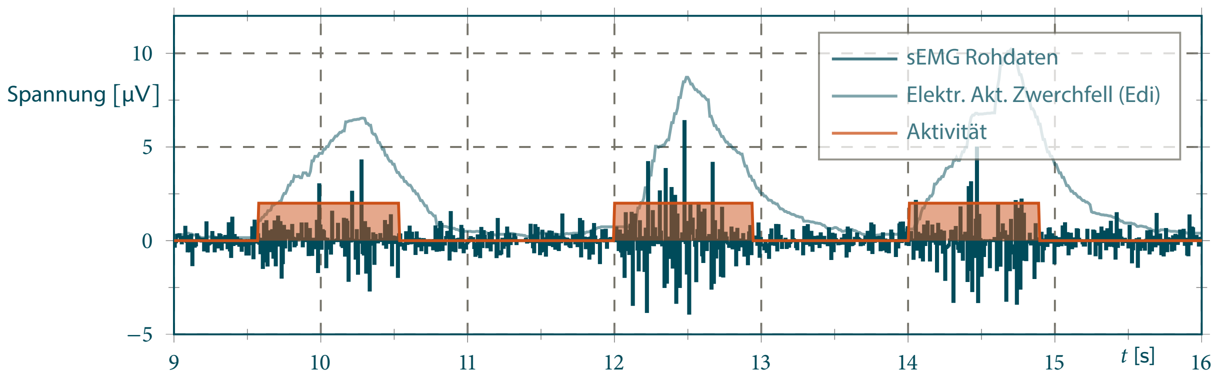

- Nur wenn die Abweichungen zwischen Messwert und prädiziertem Zustand nicht mehr durch Rauschterme zu erklären sind, wird eine Varianz $\sigma^2_{U_{k-1}} \neq 0$ geschätzt.





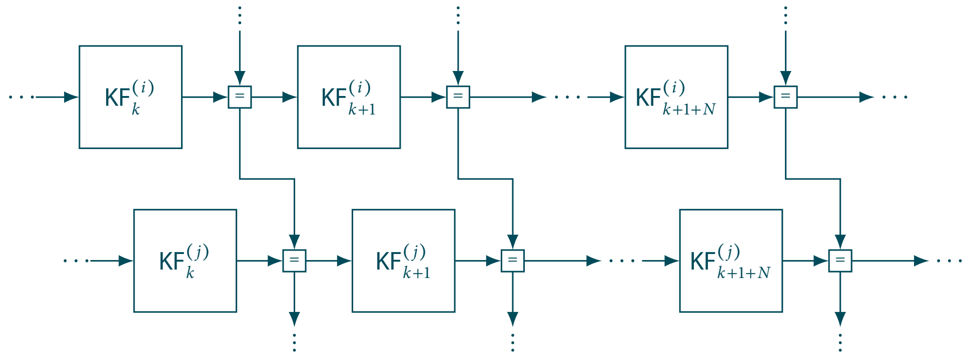
# Faktorgraphen für die Signalverarbeitung

## Beispiel: Change Point Detection auf sEMG-Signalen




- Schätzen der Eingangsvarianz zur Detektion von Aktivität
- Daten aufgenommen von Giacomo Bellani, University of Milan-Bicocca, Monza, Italien

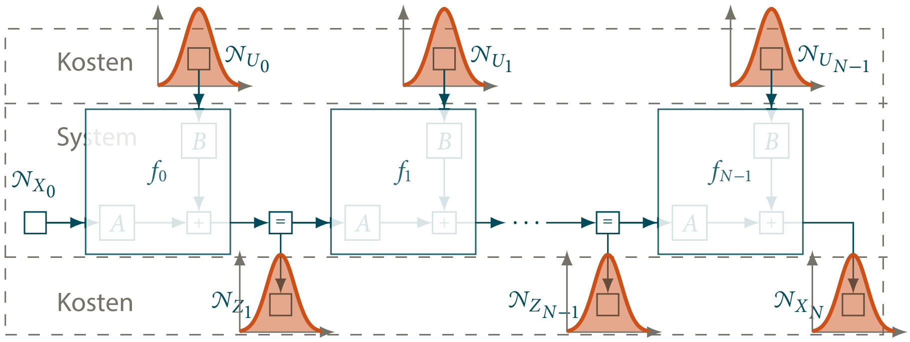
# Verteilte und «Mix & Match» Algorithmen




- Verteilte Implementierung durch «Schnitt» an beliebiger Kante
- Kombination geeigneter Algorithmen im Message Passing Framework:
  - Kalman Filterung oder Lineare Quadratische Regelung
  - Expectation Maximization (Dempster, Laird und Rubin 1977)
  - Expectation Propagation (Minka 2001)

# Lineare optimale Regelung auf Faktorgraphen

## Dynamische Programmierung mittels Message Passing (Hoffmann und Rostalski 2017)




$$f_{k-1}(\cdot) = \delta(x_k - Ax'_{k-1} - Bu_{k-1})$$

- Interpretation des Max-Product Algorithmus zur Optimierung einer quadratischen Kostenfunktion:
  - Form der Gaußschen Wahrscheinlichkeitsdichte führt zu quadratischem Kostenterm
- Interpretation der A-Priori Wahrscheinlichkeiten:
  - Gewichte entsprechen Bestrafung von Abweichungen vom angenommenen Arbeitspunkt
  - Gewichte bestimmen die gewünschte Konfidenz eines Arbeitspunktes

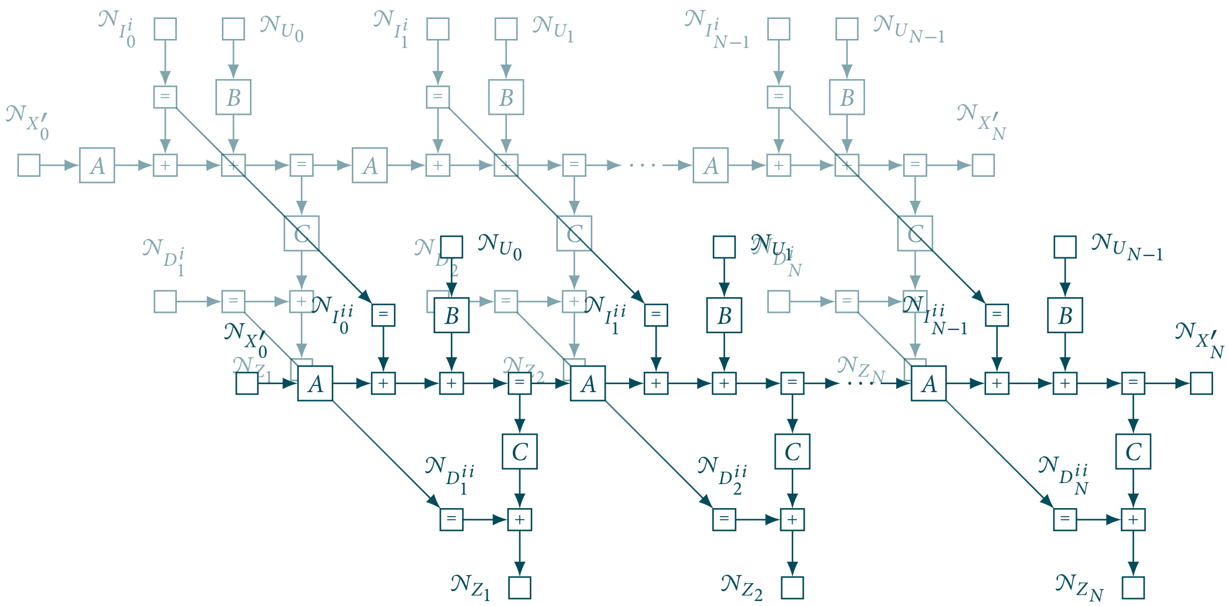
$$\bar{\mu}_{U_{k-1}}(u_{k-1}) = \vec{\mu}_{U_{k-1}}(u_{k-1}) \iint_{-\infty}^{\infty} f_{k-1}(\cdot)\,\delta(x'_{k-1} - \vec{\mathrm{m}}_{X_{k-1}})\exp\left\{-\frac{1}{2}x'^{\mathsf{T}}_{k-1}Qx'_{k-1}\right\}$$
$$\times \exp\left\{-V_k(x_k)\right\}\,\mathrm{d}x'_{k-1}\mathrm{d}x_k$$
$$= \exp\left\{-\tfrac{1}{2}\left(u^{\mathsf{T}}_{k-1}Ru_{k-1} - \vec{\mathrm{m}}^{\mathsf{T}}_{X_{k-1}}Q\vec{\mathrm{m}}_{X_{k-1}}\right) - V_k(A\vec{\mathrm{m}}_{X_{k-1}} + Bu_{k-1})\right\}$$





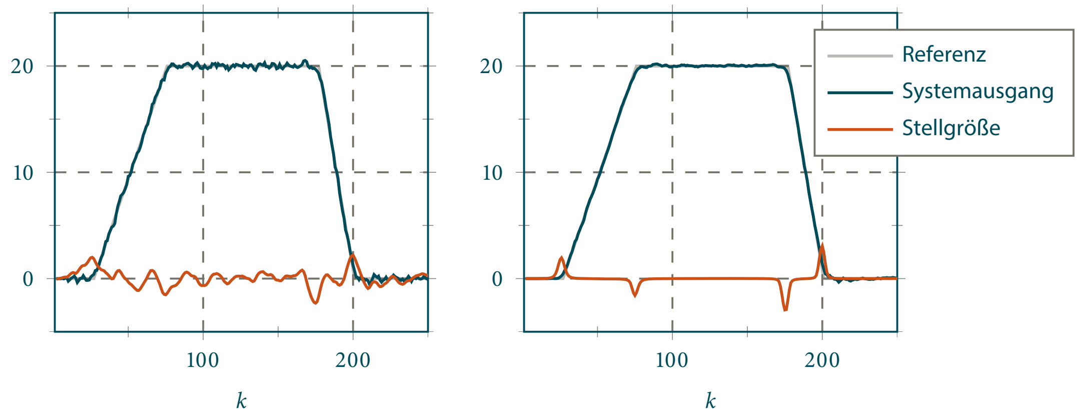
# Iterativ lernende lineare optimale Regelung auf Faktorgraphen

## Einfache Interpretation mittels Faktorgraphen (Hoffmann und Rostalski 2017)



# Regularisierte optimale Regelung mittels Faktor Graphen

## Beispiel: Regularisierte Stellgrößen in der Regelung (Hoffmann und Rostalski 2017)




(a) Linear Optimale Regelung

(b) Regularisierte Lineare Optimale Regelung





1. Faktorgraphen und Gaußsches Message Passing
2. Message Passing für Zustands- und Parameterschätzung
3. Regelung mittels Faktorgraphen
4. Zusammenfassung

# Zusammenfassung

Vielfältige Anwendungsmöglichkeiten für probabilistische graphische Modelle

Effiziente Algorithmen durch explizite Ausnutzung der Problemstruktur

Verteilte Implementierung durch Message Passing Ansatz

Einfache Herleitung neuer Algorithmen im Stile des «Mix & Match»





# Literatur I

Hoffmann, Christian, Andreas Isler und Philipp Rostalski (2017). «A Factor Graph Approach to Parameter Identification for Affine LPV Systems». In: *Proc. Amer. Contr. Conf.* accepted.

Hoffmann, Christian und Philipp Rostalski (2017). «Linear Optimal Control on Factor Graphs - A Message Passing Perspective». In: *Proc. 20th IFAC World Congress*. submitted.

Loeliger, Hans-Andrea u. a. (2016). «On Sparsity by NUV-EM, Gaussian Message Passing, and Kalman Smoothing». In: *arXiv*.

Moallemi, Ciamac C. und Benjamin van Roy (2006). «Consensus Propagation». In: *IEEE Trans. Inform. Theory* 52.11. IEEE Transactions on Information Theory, 2006, 52(11): 4753-4766 journal version, S. 4753–4766. issn: 00189448. doi: 10.1109/TIT.2006.883539.

Dauwels, J., S. Korl und Hans-Andrea Loeliger (2005). «Expectation maximization as message passing». In: *Proc. Int. Symp. Inf. Theory*, S. 583–586. isbn: 0-7803-9151-9. doi: 10.1109/ISIT.2005.1523402.

Loeliger, Hans-Andrea (2004). «An Introduction to factor graphs». In: *IEEE Signal Process. Mag.* 21.1, S. 28–41. issn: 1053-5888. doi: 10.1109/MSP.2004.1267047.

Kschischang, F. R., B. J. Frey und Hans-Andrea Loeliger (2001). «Factor graphs and the sum-product algorithm». In: *IEEE Trans. Inform. Theory* 47.2, S. 498–519. issn: 00189448. doi: 10.1109/18.910572.

Minka, Thomas (2001). «A Family of Algorithms for Approximate Bayesian Inference». Ph.D. Thesis. Massachusetts Institute of Technology.

Tanner, R. (1981). «A recursive approach to low complexity codes». In: *IEEE Trans. Inform. Theory* 27.5, S. 533–547. issn: 00189448. doi: 10.1109/TIT.1981.1056404.

Dempster, A. P., N. M. Laird und D. B. Rubin (1977). «Maximum Likelihood from Incomplete Data via the EM Algorithm». In: *J. Royal Stat. Soc. Series B* 39.1, S. 1–38.

Kwakernaak, Huibert und Raphael Sivan (1972). *Linear optimal control systems*. New York und Chichester: Wiley-Interscience. isbn: 9780471511106.

Viterbi, A. (1967). «Error bounds for convolutional codes and an asymptotically optimum decoding algorithm». In: *IEEE Trans. Inform. Theory* 13.2, S. 260–269. issn: 00189448. doi: 10.1109/TIT.1967.1054010.

Kalman, R. E. (1960). «A New Approach to Linear Filtering and Prediction Problems». In: *J. Basic Engineering* 82.1, S. 35. issn: 00219223. doi: 10.1115/1.3662552.