

# Distribuovaná kalmanovská filtrace při neznámém prostorově heterogenním šumu

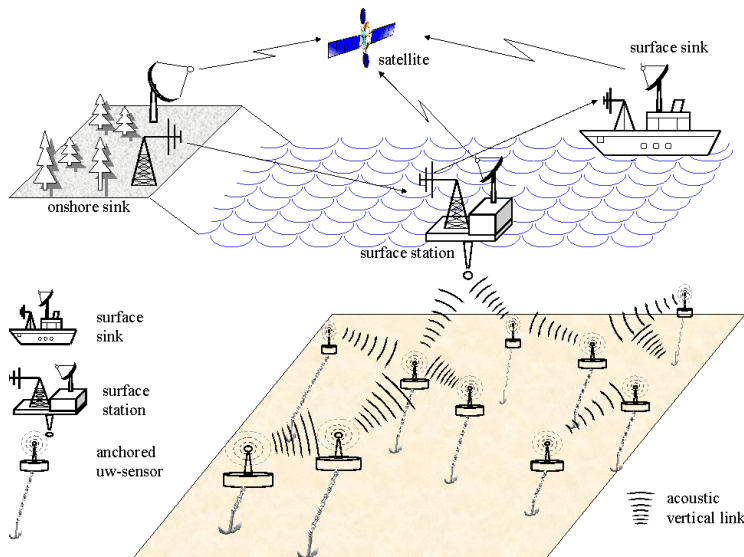
Daniel Hnyk

Matematická informatika  
Katedra matematiky

Fakulta jaderná a fyzikálně inženýrská  
České vysoké učení technické v Praze

Obhajoba diplomové práce (2018)

## Co je to difúzní síť?



## Cíl práce

### Návrh algoritmu...

...k distribuované kalmanovské filtraci v difuzní síti, kde každý uzel sítě může potenciálně mít jinou kovarianci šumu na stavovém procesu i na měřeních a tato kovariance je navíc neznámá.

### Řešení

Distribuovaný adaptivní kalmanův filtr za použití variačních Bayesovských metod.

## Cíl práce

### Návrh algoritmu...

...k distribuované kalmanovské filtraci v difuzní síti, kde každý uzel sítě může potenciálně mít jinou kovarianci šumu na stavovém procesu i na měřeních a tato kovariance je navíc neznámá.

### Řešení

Distribuovaný adaptivní kalmanův filtr za použití variačních Bayesovských metod.

# Kalmanův filtr

- používá se pro zpřesnění předpovědí modelu na základě měření
- distributivní - více propojených agentů provádějící filtraci
- optimální lineární filtr (v MSE) pokud:
  - model přesně reflektuje skutečnost
  - šum je Gaussovský bílý
  - kovariance šumu jsou přesně známy
- tato práce se zabývá rozšířením kdy 3. bod není splněn

# Bayesovská pravděpodobnost

- interpretace pravděpodobnosti
- inference pro model  $M = \{P_{\theta} | \theta \in \Theta\}$
- chceme získat aposteriorní pravděpodobnost dle

$$p(\mathbf{Z}|\mathbf{X}) = p(\mathbf{Z}) \cdot \frac{p(\mathbf{X}|\mathbf{Z})}{p(\mathbf{X})}$$

- přímý výpočet obvykle není možný, aproximace například faktorizovatelnými distribucemi

$$q(\mathbf{Z}) = \prod_{i=1}^M q_i(\mathbf{Z})$$

## Optimalizační problém

- vzdálenost měříme pomocí Kullback-Leiblerovy divergence

$$\text{KL}(q||p) = - \int q(\mathbf{Z}) \ln \frac{p(\mathbf{Z}|\mathbf{X})}{q(\mathbf{Z})} d\mathbf{Z}$$

- optimalizační problém

$$q(\mathbf{Z}) = \underset{q(\mathbf{Z})\text{-fakt.}}{\text{argmin}} \text{KL}(q(\mathbf{Z})||p(\mathbf{Z}|\mathbf{X}))$$

- stále nejde přímo, ekvivalentní maximalizaci

$$\mathcal{L}(q) = \int q(\mathbf{Z}) \ln \frac{p(\mathbf{X}, \mathbf{Z})}{q(\mathbf{Z})} d\mathbf{Z}$$

## Coordinate ascent mean-field variational

- optimální řešení problému

$$q_i^*(\mathbf{Z}_i) \propto \exp\{\mathbb{E}_{-i}[\ln p(\mathbf{Z}_i, \mathbf{Z}_{\{-i\}}, \mathbf{X})]\}$$

- v praxi se postupuje iterativně
- vhodná volba apriorních distribucí zjednodušuje update
- elegantní reprezentace použije-li se *forma exponenciální rodiny*



## Pravděpodobností model

- odhady pravděpodobnostně odpovídají

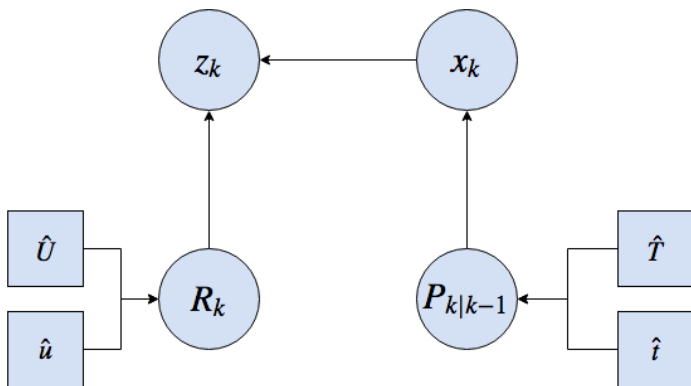
$$p(\mathbf{z}_k | \mathbf{x}_k) = N(\mathbf{z}_k; \mathbf{H}_k \mathbf{x}_k, \mathbf{R}_k)$$

$$p(\mathbf{x}_k | \mathbf{x}_{k-1}) = N(\mathbf{x}_k; F_{k|k-1} \mathbf{x}_{k-1}, \mathbf{Q}_k).$$

- cílem je odhadnout kovariační matice šumů:  $\mathbf{R}_k, \mathbf{Q}_k$
- konjugovaná apriorná jsou inverzní Wishartovy distribuce
- pro CAVI předpokládáme, že distribuce je faktorizovatelná

$$\begin{aligned} p(\boldsymbol{\Xi}, \mathbf{z}_{1:k}) &= N(\mathbf{z}_k; \mathbf{H}_k \mathbf{x}_k, \mathbf{R}_k) N(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1}) \\ &\times \text{IW}(\mathbf{P}_{k|k-1}; \hat{t}_{k|k-1}, \hat{\mathbf{T}}_{k|k-1}) \text{IW}(\mathbf{R}_k; \hat{u}_{k|k-1}, \hat{\mathbf{U}}_{k|k-1}) \\ &\times p(\mathbf{z}_{1:k-1}) \end{aligned}$$

# Grafický pravděpodobnostní model



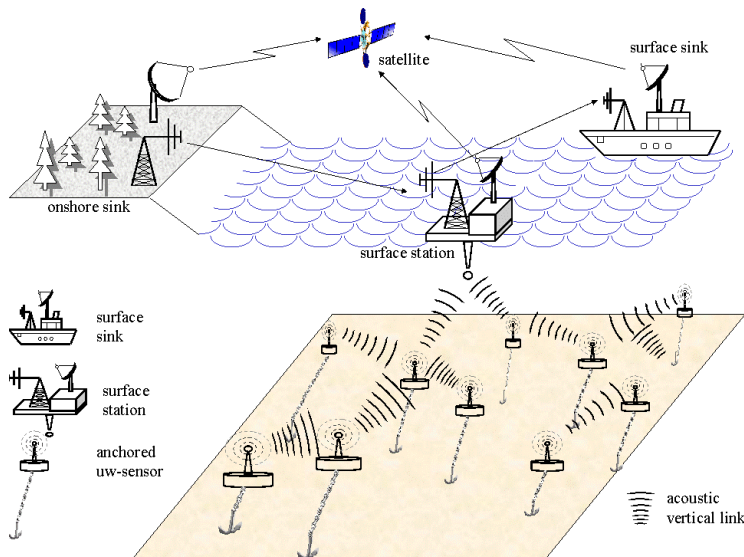
## Difúze parametrů

- více agentů provádí filtraci lokálně
- vzájemně propojeni dle dané topografie
- po  $n$  krocích si nějakým způsobem vymění informace o odhadech

$$\Xi_*^{i,t} = g(\Xi^{i,t}, \Xi^{1,t}, \dots, \Xi^{d,t})$$

- snaha o zlepšení lokálního odhadu *zprůměrováním* s ostatními
- je také Kullback-Leibler optimální

# Co je to difúzní síť?



## Bayesovský update

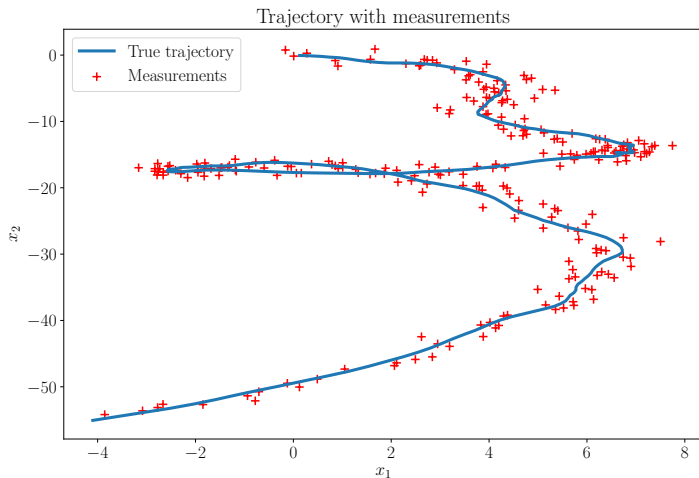
využitím formy exponenciální rodiny se celý update zjednodušuje na

$$\begin{aligned}\Xi_{i+1}^n &\leftarrow \Xi^{n-1} - \frac{1}{2} \begin{bmatrix} \mathbf{P}_{k|k}^i + (\hat{\mathbf{x}}_{k|k}^i - \hat{\mathbf{x}}_{k|k-1})(\hat{\mathbf{x}}_{k|k}^i - \hat{\mathbf{x}}_{k|k-1})^T \\ 1 \end{bmatrix} \\ \Omega_{i+1}^n &\leftarrow \Omega^{n-1} + \begin{bmatrix} \mathbf{H}_k \mathbf{P}_{k|k}^i \mathbf{H}_k^T + (\mathbf{z}_k - \hat{\mathbf{x}}_{k|k}^i)(\mathbf{z}_k - \hat{\mathbf{x}}_{k|k}^i)^T \\ -\frac{1}{2} \end{bmatrix}\end{aligned}$$

## Shrnutí algoritmu

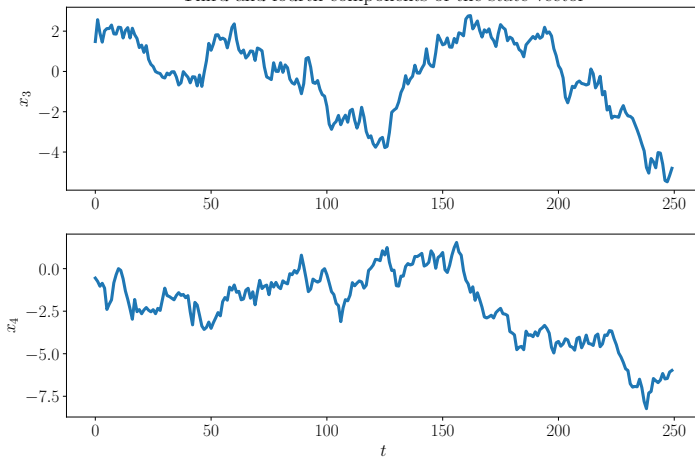
1. **in:**  $\hat{\mathbf{x}}_{k-1|k-1}, \mathbf{P}_{k-1|k-1}, \hat{u}_{k-1|k-1}, \hat{\mathbf{U}}_{k-1|k-1}, \mathbf{F}, \mathbf{H}, \mathbf{z}_k, \tilde{\mathbf{Q}}_{k-1}, \tau, \rho, N$
2. predikce kalmanova filtru
3. inicializace hyperparametrů  $\Xi_{\text{init}}, \Omega_{\text{init}}, \hat{\mathbf{x}}_{k|k}^{(0)} = \hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k}^{(0)} = \tilde{\mathbf{P}}_k$
4. **for**  $i$  in  $N$  **do**:
  - 4.1 updatuj  $\Omega^i$  a spočti z něj  $\mathbb{E}[\mathbf{R}_k]$
  - 4.2 updatuj  $\Xi^i$  a spočti z něj  $\mathbb{E}[\mathbf{P}_{k|k-1}]$
  - 4.3 kalmanova korekce na základě  $\hat{\mathbf{x}}_{k|k}^{i+1}, \mathbf{P}_{k|k-1}^{i+1}$
5. **end for**
6. difúze:  $\Xi_*^{i,t} = g(\Xi^{i,t}, \Xi^{1,t}, \dots, \Xi^{d,t}), \Omega_*^{i,t} = g(\Omega^{i,t}, \Omega^{1,t}, \dots, \Omega^{d,t})$
7. **out:**  $\hat{\mathbf{x}}_{k|k}^{i+1}, \mathbf{P}_{k|k-1}^{i+1}, \Xi_*^{i,t}, \Omega_*^{i,t}$

# Testovací úloha



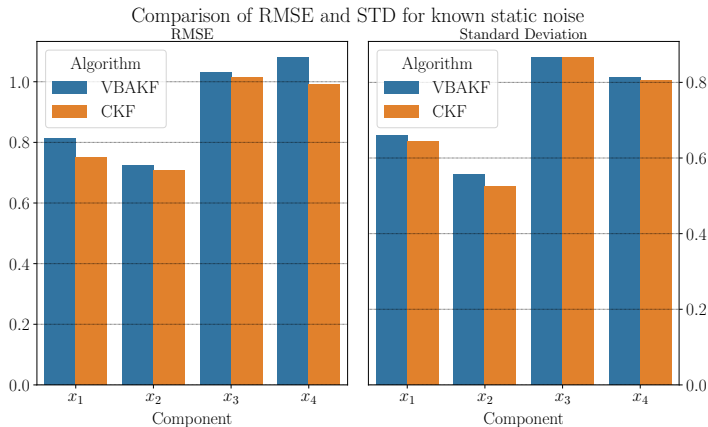
# Testovací úloha

Third and fourth components of the state vector

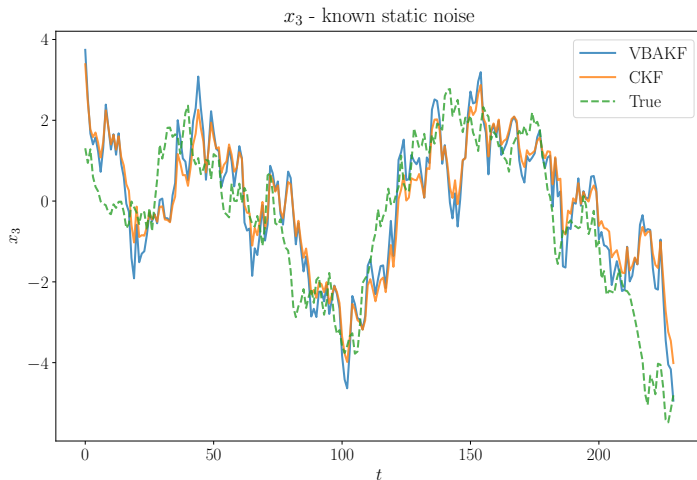




# Statický známý šum - srovnání

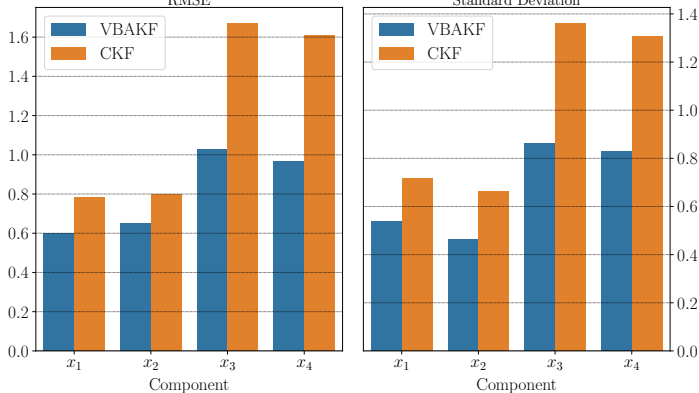


# Statický známý šum - 3 komponenta

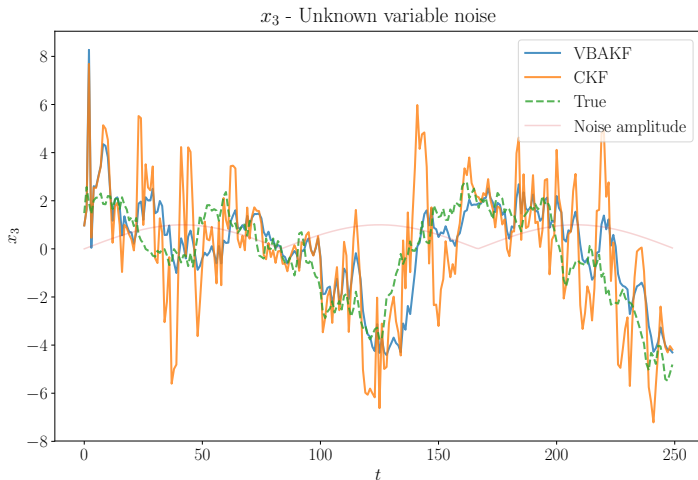


# Proměnlivý šum - srovnání

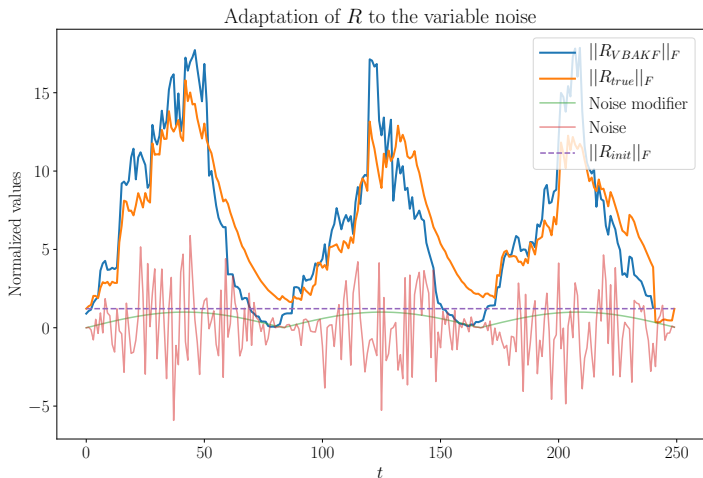
Comparison of RMSE and STD for unknown variable noise



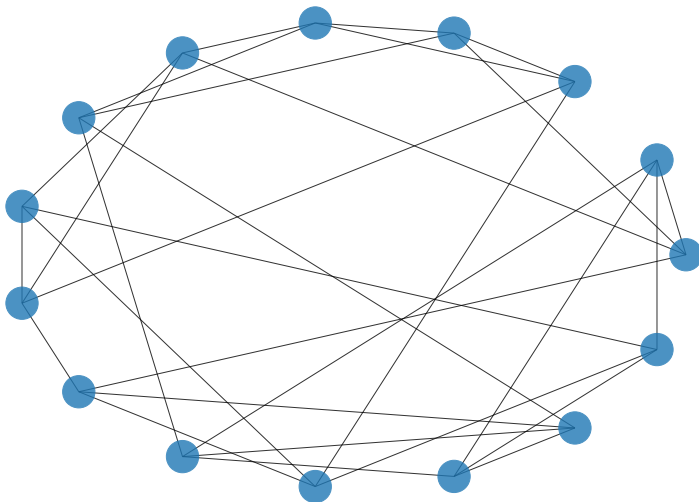
# Proměnlivý šum - 3 komponenta



# Proměnlivý šum - kovariance

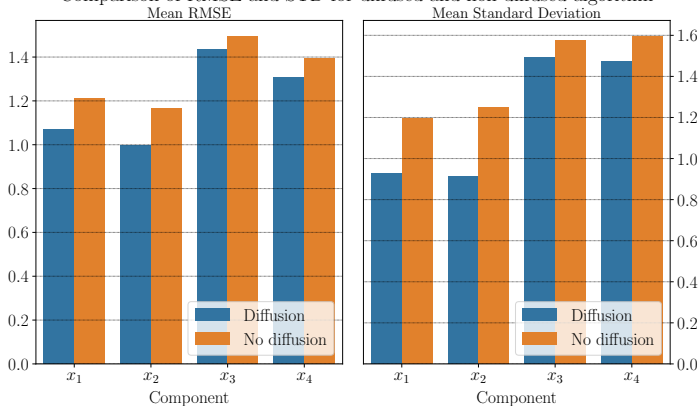


## Distribuovaná verze - topologie



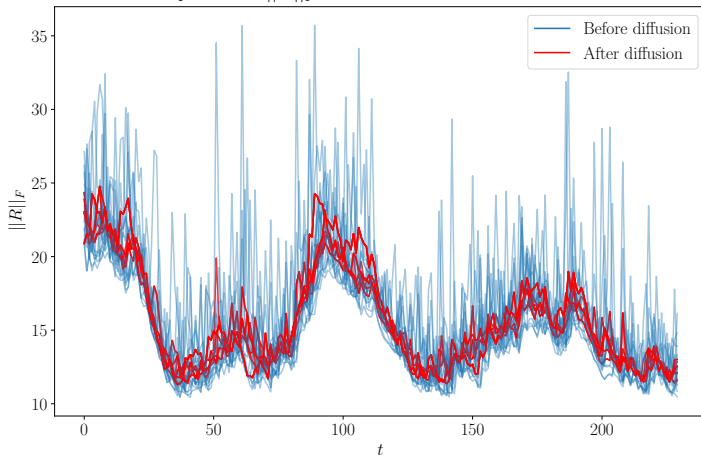
## Proměnlivý šum - srovnání

Comparison of RMSE and STD for diffused and non-diffused algorithm



# Proměnlivý šum - srovnání kovariancí

Comparison of  $\|R\|_F$  before and after diffusion - 15 nodes





## Shrnutí

- navržený algoritmus má v případech neznámého šumu lepší výkon než základní verze filtru
- vede si zanedbatelně hůře v případě kdy je kovariance známá
- difúze přináší vyšší přesnost, nižší varianci a obranu proti divergenci
- VBAKF vyžaduje nastavení dalších 2 („tunning“) parametrů

## Budoucí práce

- jaký vliv má topologie sítě a komunikace
- srovnání různých strategií difúzní funkce
- citlivost na jiné typy šumů než Gaussovský
- konvergenční vlastnosti

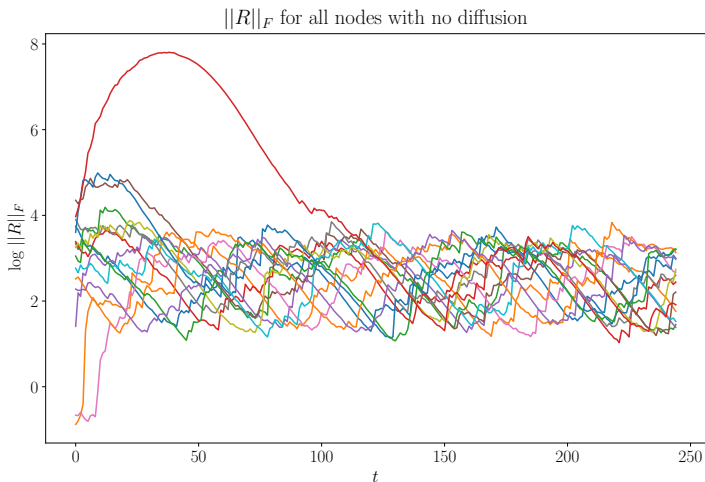
## Faktické poznámky

- úvodní kapitola je o interpretacích, ne jen o srovnání
- *2D tracking problem* – ustálený pojem
- difúzní síť a difúze parametrů – uvedené minimum využité v práci
- konvergence
  - důkaz v [11] citovaný v úvodu kapitoly
  - lokální versus globální: CAVI podobné vlastnosti jako gradient descent

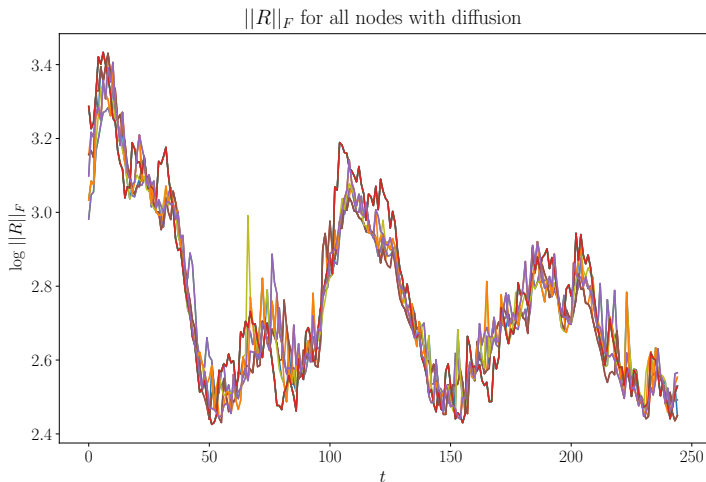
# Sazba

- $T$  pro transpozici
- značení pro distribuci:  $D(\alpha)$  versus  $D(x; \alpha)$
- značení vektorů a skalárů

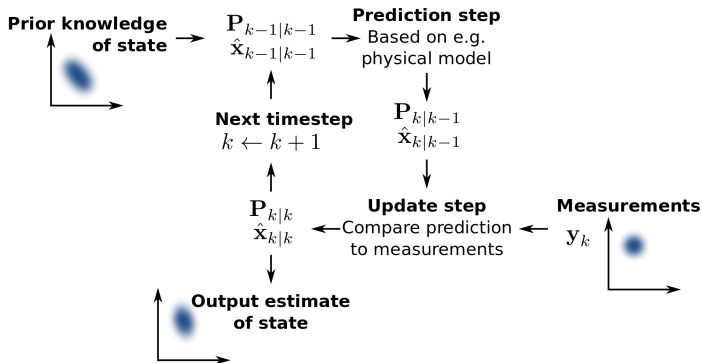
# Proměnlivý šum - kovariance bez difúze



# Proměnlivý šum - kovariance s difúzí



## Kalmanův filtr vizuálně



## Lower variational boundary

