

ECE276A: Sensing & Estimation in Robotics

Lecture 10: Extended and Unscented Kalman Filter

Nikolay Atanasov
natanasov@ucsd.edu

UC San Diego
JACOBS SCHOOL OF ENGINEERING
Electrical and Computer Engineering

Outline

Nonlinear Kalman Filter

Extended Kalman Filter

Unscented Kalman Filter

Comparison of EKF and UKF

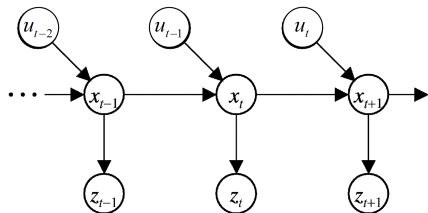
Bayes Filter

► **Motion model:**

$$\mathbf{x}_{t+1} = f(\mathbf{x}_t, \mathbf{u}_t, \mathbf{w}_t) \sim p_f(\cdot | \mathbf{x}_t, \mathbf{u}_t)$$

► **Observation model:**

$$\mathbf{z}_t = h(\mathbf{x}_t, \mathbf{v}_t) \sim p_h(\cdot | \mathbf{x}_t)$$



► **Prior:** $p_{t|t}(\mathbf{x}_t) := p(\mathbf{x}_t | \mathbf{z}_{0:t}, \mathbf{u}_{0:t-1})$

► **Prediction:** $p_{t+1|t}(\mathbf{x}) = \int p_f(\mathbf{x} | \mathbf{s}, \mathbf{u}_t) p_{t|t}(\mathbf{s}) d\mathbf{s}$

► **Update:** $p_{t+1|t+1}(\mathbf{x}) = \frac{p_h(\mathbf{z}_{t+1} | \mathbf{x}) p_{t+1|t}(\mathbf{x})}{p(\mathbf{z}_{t+1} | \mathbf{z}_{0:t}, \mathbf{u}_{0:t})} = \frac{p_h(\mathbf{z}_{t+1} | \mathbf{x}) p_{t+1|t}(\mathbf{x})}{\int p_h(\mathbf{z}_{t+1} | \mathbf{s}) p_{t+1|t}(\mathbf{s}) d\mathbf{s}}$

Kalman Filter

- ▶ **Kalman filter:** a Bayes filter with additional assumptions:
 - ▶ The prior pdf $p_{t|t}$ is Gaussian
 - ▶ The motion model is linear in the state \mathbf{x}_t with Gaussian noise \mathbf{w}_t
 - ▶ The observation model is linear in the state \mathbf{x}_t with Gaussian noise \mathbf{v}_t
 - ▶ The motion noise \mathbf{w}_t and observation noise \mathbf{v}_t are independent of each other, of the state \mathbf{x}_t , and across time

- ▶ **Prior:** $\mathbf{x}_t \mid \mathbf{z}_{0:t}, \mathbf{u}_{0:t-1} \sim \mathcal{N}(\boldsymbol{\mu}_{t|t}, \boldsymbol{\Sigma}_{t|t})$

- ▶ **Motion model:**

$$\mathbf{x}_{t+1} = f(\mathbf{x}_t, \mathbf{u}_t, \mathbf{w}_t) := F\mathbf{x}_t + G\mathbf{u}_t + \mathbf{w}_t, \quad \mathbf{w}_t \sim \mathcal{N}(0, W)$$

$$\mathbf{x}_{t+1} \mid \mathbf{x}_t, \mathbf{u}_t \sim \mathcal{N}(F\mathbf{x}_t + G\mathbf{u}_t, W), \quad F \in \mathbb{R}^{d_x \times d_x}, G \in \mathbb{R}^{d_x \times d_u}, W \in \mathbb{R}^{d_x \times d_x}$$

- ▶ **Observation model:**

$$\mathbf{z}_t = h(\mathbf{x}_t, \mathbf{v}_t) := H\mathbf{x}_t + \mathbf{v}_t, \quad \mathbf{v}_t \sim \mathcal{N}(0, V)$$

$$\mathbf{z}_t \mid \mathbf{x}_t \sim \mathcal{N}(H\mathbf{x}_t, V), \quad H \in \mathbb{R}^{d_z \times d_x}, V \in \mathbb{R}^{d_z \times d_z}$$

Nonlinear Kalman Filter

- ▶ **Nonlinear Kalman filter:** a Bayes filter with additional assumptions:
 - ▶ The prior pdf $p_{t|t}$ is Gaussian
 - ▶ The motion model is ~~linear in the state \mathbf{x}_t~~ with Gaussian noise \mathbf{w}_t
 - ▶ The observation model is ~~linear in the state \mathbf{x}_t~~ with Gaussian noise \mathbf{v}_t
 - ▶ The motion noise \mathbf{w}_t and observation noise \mathbf{v}_t are independent of each other, of the state \mathbf{x}_t , and across time
 - ▶ The predicted and updated pdfs are **forced to be Gaussian via approximation**
- ▶ **Prior:** $\mathbf{x}_t \mid \mathbf{z}_{0:t}, \mathbf{u}_{0:t-1} \sim \mathcal{N}(\boldsymbol{\mu}_{t|t}, \boldsymbol{\Sigma}_{t|t})$
- ▶ **Motion model:** $\mathbf{x}_{t+1} = f(\mathbf{x}_t, \mathbf{u}_t, \mathbf{w}_t), \quad \mathbf{w}_t \sim \mathcal{N}(0, W)$
- ▶ **Observation model:** $\mathbf{z}_t = h(\mathbf{x}_t, \mathbf{v}_t), \quad \mathbf{v}_t \sim \mathcal{N}(0, V)$
- ▶ **Challenge:** the predicted and updated pdfs are not Gaussian and can no longer be evaluated in closed form
- ▶ **Moment matching:** we can force the predicted and updated pdfs to be Gaussian by evaluating their first and second moments and approximating the pdfs using Gaussians with the same moments

Moment Matching

- ▶ Let $\mathbf{y} = f(\mathbf{x})$ be a nonlinear transformation of $\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$
- ▶ The mean and (co)variance of \mathbf{y} are:

$$\mathbf{m} := \mathbb{E}[\mathbf{y}] = \int f(\mathbf{x})\phi(\mathbf{x}; \boldsymbol{\mu}, \Sigma)d\mathbf{x}$$

$$\begin{aligned} S &:= \mathbb{E} \left[(\mathbf{y} - \mathbb{E}[\mathbf{y}]) (\mathbf{y} - \mathbb{E}[\mathbf{y}])^\top \right] = \mathbb{E} [\mathbf{y}\mathbf{y}^\top] - \mathbb{E}[\mathbf{y}]\mathbb{E}[\mathbf{y}]^\top \\ &= \int f(\mathbf{x})f(\mathbf{x})^\top \phi(\mathbf{x}; \boldsymbol{\mu}, \Sigma)d\mathbf{x} - \mathbf{m}\mathbf{m}^\top \end{aligned}$$

- ▶ The covariance of \mathbf{x} and \mathbf{y} is:

$$C := \mathbb{E} \left[(\mathbf{x} - \boldsymbol{\mu}) (\mathbf{y} - \mathbb{E}[\mathbf{y}])^\top \right] = \int \mathbf{x}f(\mathbf{x})^\top \phi(\mathbf{x}; \boldsymbol{\mu}, \Sigma)d\mathbf{x} - \boldsymbol{\mu}\mathbf{m}^\top$$

- ▶ The joint distribution of \mathbf{x} and \mathbf{y} can be approximated by a Gaussian:

$$\begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} \boldsymbol{\mu} \\ \mathbf{m} \end{pmatrix}, \begin{bmatrix} \Sigma & C \\ C^\top & S \end{bmatrix} \right)$$

- ▶ The approximate distribution of \mathbf{x} conditioned on \mathbf{y} is:

$$\mathbf{x} \mid \mathbf{y} \sim \mathcal{N} (\boldsymbol{\mu} + CS^{-1}(\mathbf{y} - \mathbf{m}), \Sigma - CS^{-1}C^\top)$$

Nonlinear Kalman Filter Prediction

- ▶ **Prior:** $\mathbf{x}_t \mid \mathbf{z}_{0:t}, \mathbf{u}_{0:t-1} \sim \mathcal{N}(\boldsymbol{\mu}_{t|t}, \boldsymbol{\Sigma}_{t|t})$
- ▶ **Motion model:** $\mathbf{x}_{t+1} = f(\mathbf{x}_t, \mathbf{u}_t, \mathbf{w}_t), \quad \mathbf{w}_t \sim \mathcal{N}(0, W)$
- ▶ **Prediction step:** force a Gaussian predicted pdf via moment matching:

$$\mathbf{x}_{t+1} \mid \mathbf{z}_{0:t}, \mathbf{u}_{0:t} \sim \mathcal{N}(\boldsymbol{\mu}_{t+1|t}, \boldsymbol{\Sigma}_{t+1|t})$$

$$\boldsymbol{\mu}_{t+1|t} = \mathbb{E}[\mathbf{x}_{t+1} \mid \mathbf{z}_{0:t}, \mathbf{u}_{0:t}] = \int \int f(\mathbf{x}, \mathbf{u}_t, \mathbf{w}) \phi(\mathbf{x}; \boldsymbol{\mu}_{t|t}, \boldsymbol{\Sigma}_{t|t}) \phi(\mathbf{w}; 0, W) d\mathbf{x} d\mathbf{w}$$

$$\begin{aligned} \boldsymbol{\Sigma}_{t+1|t} &= \mathbb{E} \left[\left(\mathbf{x}_{t+1} - \boldsymbol{\mu}_{t+1|t} \right) \left(\mathbf{x}_{t+1} - \boldsymbol{\mu}_{t+1|t} \right)^\top \mid \mathbf{z}_{0:t}, \mathbf{u}_{0:t} \right] \\ &= \mathbb{E} \left[\mathbf{x}_{t+1} \mathbf{x}_{t+1}^\top \mid \mathbf{z}_{0:t}, \mathbf{u}_{0:t} \right] - \boldsymbol{\mu}_{t+1|t} \boldsymbol{\mu}_{t+1|t}^\top \\ &= \int \int f(\mathbf{x}, \mathbf{u}_t, \mathbf{w}) f(\mathbf{x}, \mathbf{u}_t, \mathbf{w})^\top \phi(\mathbf{x}; \boldsymbol{\mu}_{t|t}, \boldsymbol{\Sigma}_{t|t}) \phi(\mathbf{w}; 0, W) d\mathbf{x} d\mathbf{w} - \boldsymbol{\mu}_{t+1|t} \boldsymbol{\mu}_{t+1|t}^\top \end{aligned}$$

Nonlinear Kalman Filter Update

- ▶ **Prior:** $\mathbf{x}_{t+1} \mid \mathbf{z}_{0:t}, \mathbf{u}_{0:t} \sim \mathcal{N}(\boldsymbol{\mu}_{t+1|t}, \boldsymbol{\Sigma}_{t+1|t})$
- ▶ **Observation model:** $\mathbf{z}_{t+1} = h(\mathbf{x}_{t+1}, \mathbf{v}_{t+1}), \quad \mathbf{v}_{t+1} \sim \mathcal{N}(0, V)$
- ▶ **Update step:** the Gaussian distribution which approximates the joint of \mathbf{x}_{t+1} and \mathbf{z}_{t+1} conditioned on $\mathbf{z}_{0:t}, \mathbf{u}_{0:t}$ via moment matching is:

$$\begin{pmatrix} \mathbf{x}_{t+1} \mid \mathbf{z}_{0:t}, \mathbf{u}_{0:t} \\ \mathbf{z}_{t+1} \mid \mathbf{z}_{0:t}, \mathbf{u}_{0:t} \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} \boldsymbol{\mu}_{t+1|t} \\ \mathbf{m}_{t+1|t} \end{pmatrix}, \begin{bmatrix} \boldsymbol{\Sigma}_{t+1|t} & \mathbf{C}_{t+1|t} \\ \mathbf{C}_{t+1|t}^\top & S_{t+1|t} \end{bmatrix} \right)$$

$$\mathbf{m}_{t+1|t} = \int \int h(\mathbf{x}, \mathbf{v}) \phi(\mathbf{x}; \boldsymbol{\mu}_{t+1|t}, \boldsymbol{\Sigma}_{t+1|t}) \phi(\mathbf{v}; 0, V) d\mathbf{x} d\mathbf{v}$$

$$S_{t+1|t} = \int \int (h(\mathbf{x}, \mathbf{v}) - \mathbf{m}_{t+1|t})(h(\mathbf{x}, \mathbf{v}) - \mathbf{m}_{t+1|t})^\top \phi(\mathbf{x}; \boldsymbol{\mu}_{t+1|t}, \boldsymbol{\Sigma}_{t+1|t}) \phi(\mathbf{v}; 0, V) d\mathbf{x} d\mathbf{v}$$

$$\mathbf{C}_{t+1|t} = \int \int (\mathbf{x} - \boldsymbol{\mu}_{t+1|t})(h(\mathbf{x}, \mathbf{v}) - \mathbf{m}_{t+1|t})^\top \phi(\mathbf{x}; \boldsymbol{\mu}_{t+1|t}, \boldsymbol{\Sigma}_{t+1|t}) \phi(\mathbf{v}; 0, V) d\mathbf{x} d\mathbf{v}$$

- ▶ The conditional Gaussian distribution of $\mathbf{x}_{t+1} \mid \mathbf{z}_{t+1}, \mathbf{z}_{0:t}, \mathbf{u}_{0:t}$ is then:

$$\boldsymbol{\mu}_{t+1|t+1} = \boldsymbol{\mu}_{t+1|t} + \mathbf{K}_{t+1|t}(\mathbf{z}_{t+1} - \mathbf{m}_{t+1|t})$$

$$\boldsymbol{\Sigma}_{t+1|t+1} = \boldsymbol{\Sigma}_{t+1|t} - \mathbf{K}_{t+1|t} S_{t+1|t} \mathbf{K}_{t+1|t}^\top$$

$$\mathbf{K}_{t+1|t} = \mathbf{C}_{t+1|t} S_{t+1|t}^{-1}$$

Extended and Unscented Kalman Filters

- ▶ Implementing a nonlinear KF requires approximating the mean $\boldsymbol{\mu}_{t+1|t}$ and covariance $\boldsymbol{\Sigma}_{t+1|t}$ integrals in the prediction step as well as the measurement mean $\mathbf{m}_{t+1|t}$, measurement covariance $\mathbf{S}_{t+1|t}$, and state-measurement correlation $\mathbf{C}_{t+1|t}$ integrals in the update step
- ▶ The **EKF** and **UKF** use different methods to approximate these five integrals
- ▶ The **EKF** uses a first-order Taylor series approximation to the motion and observation models around the state and noise means:

$$f(\mathbf{x}_t, \mathbf{u}_t, \mathbf{w}_t) \approx f(\boldsymbol{\mu}_{t|t}, \mathbf{u}_t, \mathbf{0}) + \left[\frac{df}{d\mathbf{x}}(\boldsymbol{\mu}_{t|t}, \mathbf{u}_t, \mathbf{0}) \right] (\mathbf{x}_t - \boldsymbol{\mu}_{t|t}) + \left[\frac{df}{d\mathbf{w}}(\boldsymbol{\mu}_{t|t}, \mathbf{u}_t, \mathbf{0}) \right] (\mathbf{w}_t - \mathbf{0})$$

$$h(\mathbf{x}_{t+1}, \mathbf{v}_{t+1}) \approx h(\boldsymbol{\mu}_{t+1|t}, \mathbf{0}) + \left[\frac{dh}{d\mathbf{x}}(\boldsymbol{\mu}_{t+1|t}, \mathbf{0}) \right] (\mathbf{x}_{t+1} - \boldsymbol{\mu}_{t+1|t}) + \left[\frac{dh}{d\mathbf{v}}(\boldsymbol{\mu}_{t+1|t}, \mathbf{0}) \right] (\mathbf{v}_{t+1} - \mathbf{0})$$

- ▶ The **UKF** uses a finite set of **sigma points** to approximate the prior Gaussian pdfs and convert the integrals to sums. This resembles Monte Carlo approximation but the sigma points are selected **deterministically**.

Outline

Nonlinear Kalman Filter

Extended Kalman Filter

Unscented Kalman Filter

Comparison of EKF and UKF

Extended Kalman Filter Prediction

- ▶ Let $F_t := \frac{df}{d\mathbf{x}}(\boldsymbol{\mu}_{t|t}, \mathbf{u}_t, \mathbf{0})$ and $Q_t := \frac{df}{d\mathbf{w}}(\boldsymbol{\mu}_{t|t}, \mathbf{u}_t, \mathbf{0})$ so that:

$$f(\mathbf{x}_t, \mathbf{u}_t, \mathbf{w}_t) \approx f(\boldsymbol{\mu}_{t|t}, \mathbf{u}_t, \mathbf{0}) + F_t(\mathbf{x}_t - \boldsymbol{\mu}_{t|t}) + Q_t\mathbf{w}_t$$

- ▶ **Predicted mean:**

$$\begin{aligned}\boldsymbol{\mu}_{t+1|t} &\approx \int \int \left(f(\boldsymbol{\mu}_{t|t}, \mathbf{u}_t, \mathbf{0}) + F_t(\mathbf{x} - \boldsymbol{\mu}_{t|t}) + Q_t\mathbf{w} \right) \phi(\mathbf{x}; \boldsymbol{\mu}_{t|t}, \Sigma_{t|t}) \phi(\mathbf{w}; 0, W) d\mathbf{x} d\mathbf{w} \\ &= f(\boldsymbol{\mu}_{t|t}, \mathbf{u}_t, \mathbf{0}) + F_t \left(\int \mathbf{x} \phi(\mathbf{x}; \boldsymbol{\mu}_{t|t}, \Sigma_{t|t}) d\mathbf{x} - \boldsymbol{\mu}_{t|t} \right) + Q_t \int \mathbf{w} \phi(\mathbf{w}; 0, W) d\mathbf{w} \\ &= f(\boldsymbol{\mu}_{t|t}, \mathbf{u}_t, \mathbf{0})\end{aligned}$$

Extended Kalman Filter Prediction

- ▶ Let $F_t := \frac{df}{d\mathbf{x}}(\boldsymbol{\mu}_{t|t}, \mathbf{u}_t, \mathbf{0})$ and $Q_t := \frac{df}{d\mathbf{w}}(\boldsymbol{\mu}_{t|t}, \mathbf{u}_t, \mathbf{0})$ so that:

$$f(\mathbf{x}_t, \mathbf{u}_t, \mathbf{w}_t) \approx f(\boldsymbol{\mu}_{t|t}, \mathbf{u}_t, \mathbf{0}) + F_t(\mathbf{x}_t - \boldsymbol{\mu}_{t|t}) + Q_t\mathbf{w}_t$$

- ▶ **Predicted covariance:**

$$\begin{aligned}\Sigma_{t+1|t} &\approx \iint \left(f(\boldsymbol{\mu}_{t|t}, \mathbf{u}_t, \mathbf{0}) + F_t(\mathbf{x} - \boldsymbol{\mu}_{t|t}) + Q_t\mathbf{w} \right) \left(f(\boldsymbol{\mu}_{t|t}, \mathbf{u}_t, \mathbf{0}) + F_t(\mathbf{x} - \boldsymbol{\mu}_{t|t}) + Q_t\mathbf{w} \right)^\top \\ &\quad \phi(\mathbf{x}; \boldsymbol{\mu}_{t|t}, \Sigma_{t|t}) \phi(\mathbf{w}; 0, W) d\mathbf{x} d\mathbf{w} - \boldsymbol{\mu}_{t+1|t} \boldsymbol{\mu}_{t+1|t}^\top \\ &= f(\boldsymbol{\mu}_{t|t}, \mathbf{u}_t, \mathbf{0}) \left(\int (\mathbf{x} - \boldsymbol{\mu}_{t|t})^\top \phi(\mathbf{x}; \boldsymbol{\mu}_{t|t}, \Sigma_{t|t}) d\mathbf{x} \right) F_t^\top \\ &\quad + F_t \left(\int (\mathbf{x} - \boldsymbol{\mu}_{t|t}) \phi(\mathbf{x}; \boldsymbol{\mu}_{t|t}, \Sigma_{t|t}) d\mathbf{x} \right) f(\boldsymbol{\mu}_{t|t}, \mathbf{u}_t, \mathbf{0})^\top \\ &\quad + F_t \left(\int (\mathbf{x} - \boldsymbol{\mu}_{t|t})(\mathbf{x} - \boldsymbol{\mu}_{t|t})^\top \phi(\mathbf{x}; \boldsymbol{\mu}_{t|t}, \Sigma_{t|t}) d\mathbf{x} \right) F_t^\top \\ &\quad + Q_t \left(\int \mathbf{w}\mathbf{w}^\top \phi(\mathbf{w}; 0, W) d\mathbf{w} \right) Q_t^\top \\ &= F_t \Sigma_{t|t} F_t^\top + Q_t W Q_t^\top\end{aligned}$$

Extended Kalman Filter Update

- ▶ Let $H_{t+1} := \frac{dh}{dx}(\boldsymbol{\mu}_{t+1|t}, \mathbf{0})$ and $R_{t+1} := \frac{dh}{dv}(\boldsymbol{\mu}_{t+1|t}, \mathbf{0})$ so that:

$$h(\mathbf{x}_{t+1}, \mathbf{v}_{t+1}) \approx h(\boldsymbol{\mu}_{t+1|t}, \mathbf{0}) + H_{t+1}(\mathbf{x}_{t+1} - \boldsymbol{\mu}_{t+1|t}) + R_{t+1}\mathbf{v}_{t+1}$$

- ▶ The joint distribution of \mathbf{x}_{t+1} and \mathbf{z}_{t+1} can be computed in closed form:

$$\mathbf{m}_{t+1|t} = \int \int h(\mathbf{x}, \mathbf{v}) \phi(\mathbf{x}; \boldsymbol{\mu}_{t+1|t}, \Sigma_{t+1|t}) \phi(\mathbf{v}; 0, V) d\mathbf{x} d\mathbf{v} \approx h(\boldsymbol{\mu}_{t+1|t}, \mathbf{0})$$

$$\begin{aligned} S_{t+1|t} &= \int \int (h(\mathbf{x}, \mathbf{v}) - \mathbf{m}_{t+1|t})(h(\mathbf{x}, \mathbf{v}) - \mathbf{m}_{t+1|t})^\top \phi(\mathbf{x}; \boldsymbol{\mu}_{t+1|t}, \Sigma_{t+1|t}) \phi(\mathbf{v}; 0, V) d\mathbf{x} d\mathbf{v} \\ &\approx H_{t+1} \Sigma_{t+1|t} H_{t+1}^\top + R_{t+1} V R_{t+1}^\top \end{aligned}$$

$$\begin{aligned} C_{t+1|t} &= \int \int (\mathbf{x} - \boldsymbol{\mu}_{t+1|t})(h(\mathbf{x}, \mathbf{v}) - \mathbf{m}_{t+1|t})^\top \phi(\mathbf{x}; \boldsymbol{\mu}_{t+1|t}, \Sigma_{t+1|t}) \phi(\mathbf{v}; 0, V) d\mathbf{x} d\mathbf{v} \\ &\approx \Sigma_{t+1|t} H_{t+1}^\top \end{aligned}$$

- ▶ The conditional Gaussian distribution of $\mathbf{x}_{t+1} | \mathbf{z}_{t+1}$ is then:

$$\boldsymbol{\mu}_{t+1|t+1} = \boldsymbol{\mu}_{t+1|t} + K_{t+1|t}(\mathbf{z}_{t+1} - \mathbf{m}_{t+1|t})$$

$$\Sigma_{t+1|t+1} = \Sigma_{t+1|t} - K_{t+1|t} S_{t+1|t} K_{t+1|t}^\top$$

$$K_{t+1|t} = C_{t+1|t} S_{t+1|t}^{-1}$$

Extended Kalman Filter

Prior: $\mathbf{x}_t \mid \mathbf{z}_{0:t}, \mathbf{u}_{0:t-1} \sim \mathcal{N}(\boldsymbol{\mu}_{t|t}, \boldsymbol{\Sigma}_{t|t})$

Motion model: $\mathbf{x}_{t+1} = f(\mathbf{x}_t, \mathbf{u}_t, \mathbf{w}_t), \quad \mathbf{w}_t \sim \mathcal{N}(\mathbf{0}, W)$
 $F_t := \frac{df}{d\mathbf{x}}(\boldsymbol{\mu}_{t|t}, \mathbf{u}_t, \mathbf{0}), \quad Q_t := \frac{df}{d\mathbf{w}}(\boldsymbol{\mu}_{t|t}, \mathbf{u}_t, \mathbf{0})$

Observation model: $\mathbf{z}_t = h(\mathbf{x}_t, \mathbf{v}_t), \quad \mathbf{v}_t \sim \mathcal{N}(\mathbf{0}, V)$
 $H_t := \frac{dh}{d\mathbf{x}}(\boldsymbol{\mu}_{t|t-1}, \mathbf{0}), \quad R_t := \frac{dh}{d\mathbf{v}}(\boldsymbol{\mu}_{t|t-1}, \mathbf{0})$

Prediction: $\boldsymbol{\mu}_{t+1|t} = f(\boldsymbol{\mu}_{t|t}, \mathbf{u}_t, \mathbf{0})$
 $\boldsymbol{\Sigma}_{t+1|t} = F_t \boldsymbol{\Sigma}_{t|t} F_t^\top + Q_t W Q_t^\top$

Update: $\boldsymbol{\mu}_{t+1|t+1} = \boldsymbol{\mu}_{t+1|t} + K_{t+1|t}(z_{t+1} - h(\boldsymbol{\mu}_{t+1|t}, \mathbf{0}))$
 $\boldsymbol{\Sigma}_{t+1|t+1} = (I - K_{t+1|t} H_{t+1}) \boldsymbol{\Sigma}_{t+1|t}$

Kalman gain: $K_{t+1|t} := \boldsymbol{\Sigma}_{t+1|t} H_{t+1}^\top (H_{t+1} \boldsymbol{\Sigma}_{t+1|t} H_{t+1}^\top + R_{t+1} V R_{t+1}^\top)^{-1}$

Outline

Nonlinear Kalman Filter

Extended Kalman Filter

Unscented Kalman Filter

Comparison of EKF and UKF

Unscented Transform

- ▶ The **unscented transform** (Julier et al., 1995, 2000) is a numerical method for approximating the joint distribution of a Gaussian random variable $\mathbf{x} \in \mathbb{R}^d$ and a nonlinear function f of it:

$$\mathbf{y} = f(\mathbf{x}), \quad \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} \boldsymbol{\mu} \\ \mathbf{m} \end{pmatrix}, \begin{bmatrix} \boldsymbol{\Sigma} & \mathbf{C} \\ \mathbf{C}^\top & \mathbf{S} \end{bmatrix} \right)$$

- ▶ Choose $2d + 1$ **sigma points** using columns $[\sqrt{\boldsymbol{\Sigma}}]_i$ of the square root $\sqrt{\boldsymbol{\Sigma}}$ of the covariance matrix $\boldsymbol{\Sigma} = \sqrt{\boldsymbol{\Sigma}}\sqrt{\boldsymbol{\Sigma}}^\top$:

$$\mathbf{x}^{(0)} = \boldsymbol{\mu}, \quad \mathbf{x}^{(i)} = \boldsymbol{\mu} \pm \alpha\sqrt{d+k} [\sqrt{\boldsymbol{\Sigma}}]_i, \quad i = 1, \dots, d$$

- ▶ $\sqrt{\boldsymbol{\Sigma}}$ is lower-triangular and can be obtained via **Cholesky factorization**
- ▶ Parameters $\alpha \in (0, 1]$ and $k > -d$ determine the spread of the sigma points
- ▶ The sigma points capture the shape of the original distribution of \mathbf{x}

Unscented Transform

- ▶ Each sigma point $\mathbf{x}^{(i)}$ is associated with mean weight $v^{(i)}$ and covariance weight $w^{(i)}$:
 - ▶ Choose $v^{(0)} = 1 - \frac{d}{\alpha^2(d+k)} < 1$ and $w^{(0)} \geq v^{(0)}$
 - ▶ Let $v^{(i)} = w^{(i)} = \frac{1-v^{(0)}}{2d}$ for $i = 1, \dots, 2d$
 - ▶ Let $\mathbf{x}^{(0)} = \boldsymbol{\mu}$ and $\mathbf{x}^{(i)} = \boldsymbol{\mu} \pm \sqrt{\frac{d}{1-v^{(0)}}} \left[\sqrt{\boldsymbol{\Sigma}} \right]_i$ for $i = 1, \dots, d$
- ▶ The weighted sigma points are used to approximate the integrals that determine the mean and covariance of $\mathbf{y} = f(\mathbf{x})$:

$$\mathbb{E}[\mathbf{y}] \approx \mathbf{m} = \sum_{i=0}^{2d} v^{(i)} f(\mathbf{x}^{(i)})$$

$$\text{Cov}[\mathbf{y}, \mathbf{y}] \approx S = \sum_{i=0}^{2d} w^{(i)} \left(f(\mathbf{x}^{(i)}) - \mathbf{m} \right) \left(f(\mathbf{x}^{(i)}) - \mathbf{m} \right)^\top$$

$$\text{Cov}[\mathbf{x}, \mathbf{y}] \approx C = \sum_{i=0}^{2d} w^{(i)} \left(\mathbf{x}^{(i)} - \boldsymbol{\mu} \right) \left(f(\mathbf{x}^{(i)}) - \mathbf{m} \right)^\top$$

Unscented Kalman Filter Prediction

- ▶ **Prior:** $\mathbf{x}_t \mid \mathbf{z}_{0:t}, \mathbf{u}_{0:t-1} \sim \mathcal{N}(\boldsymbol{\mu}_{t|t}, \boldsymbol{\Sigma}_{t|t})$
- ▶ **Motion model:** $\mathbf{x}_{t+1} = f(\mathbf{x}_t, \mathbf{u}_t, \mathbf{w}_t), \quad \mathbf{w}_t \sim \mathcal{N}(0, W)$
- ▶ **Sigma point weights:**

$$v^{(0)} < 1, \quad w^{(0)} \geq v^{(0)}, \quad v^{(i)} = w^{(i)} = \frac{1 - v^{(0)}}{2(d_x + d_w)}, \quad i = 1, \dots, 2(d_x + d_w)$$

- ▶ **Sigma points:**

$$\begin{pmatrix} \mathbf{x}_{t|t}^{(0)} \\ \mathbf{w}^{(0)} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\mu}_{t|t} \\ 0 \end{pmatrix}, \quad \begin{pmatrix} \mathbf{x}_{t|t}^{(i)} \\ \mathbf{w}^{(i)} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\mu}_{t|t} \\ 0 \end{pmatrix} \pm \sqrt{\frac{(d_x + d_w)}{1 - v^{(0)}}} \begin{bmatrix} \sqrt{\boldsymbol{\Sigma}_{t|t}} & 0 \\ 0 & \sqrt{W} \end{bmatrix}_i$$

- ▶ **UKF prediction step:**

$$\boldsymbol{\mu}_{t+1|t} = \sum_{i=0}^{2(d_x+d_w)} v^{(i)} f(\mathbf{x}_{t|t}^{(i)}, \mathbf{u}_t, \mathbf{w}^{(i)})$$

$$\boldsymbol{\Sigma}_{t+1|t} = \sum_{i=0}^{2(d_x+d_w)} w^{(i)} \left(f(\mathbf{x}_{t|t}^{(i)}, \mathbf{u}_t, \mathbf{w}^{(i)}) - \boldsymbol{\mu}_{t+1|t} \right) \left(f(\mathbf{x}_{t|t}^{(i)}, \mathbf{u}_t, \mathbf{w}^{(i)}) - \boldsymbol{\mu}_{t+1|t} \right)^\top$$

Unscented Kalman Filter Update

- ▶ **Prior:** $\mathbf{x}_{t+1} \mid \mathbf{z}_{0:t}, \mathbf{u}_{0:t} \sim \mathcal{N}(\boldsymbol{\mu}_{t+1|t}, \boldsymbol{\Sigma}_{t+1|t})$
- ▶ **Observation model:** $\mathbf{z}_{t+1} = h(\mathbf{x}_{t+1}, \mathbf{v}_{t+1}), \quad \mathbf{v}_{t+1} \sim \mathcal{N}(0, V)$
- ▶ **Sigma points:**

$$\begin{pmatrix} \mathbf{x}_{t+1|t}^{(0)} \\ \mathbf{v}^{(0)} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\mu}_{t+1|t} \\ 0 \end{pmatrix}, \quad \begin{pmatrix} \mathbf{x}_{t+1|t}^{(i)} \\ \mathbf{v}^{(i)} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\mu}_{t+1|t} \\ 0 \end{pmatrix} \pm \sqrt{\frac{(d_x + d_v)}{1 - \nu^{(0)}}} \begin{bmatrix} \sqrt{\boldsymbol{\Sigma}_{t+1|t}} & 0 \\ 0 & \sqrt{V} \end{bmatrix}_i$$

- ▶ **UKF update step:**
$$\boldsymbol{\mu}_{t+1|t+1} = \boldsymbol{\mu}_{t+1|t} + K_{t+1|t}(\mathbf{z}_{t+1} - \mathbf{m}_{t+1|t})$$
$$\boldsymbol{\Sigma}_{t+1|t+1} = \boldsymbol{\Sigma}_{t+1|t} - K_{t+1|t} S_{t+1|t} K_{t+1|t}^\top$$
- ▶ **Kalman gain:**
$$K_{t+1|t} = C_{t+1|t} S_{t+1|t}^{-1}$$

$$\mathbf{m}_{t+1|t} = \sum_{i=0}^{2(d_x+d_v)} w^{(i)} h(\mathbf{x}_{t+1|t}^{(i)}, \mathbf{v}^{(i)})$$

$$S_{t+1|t} = \sum_{i=0}^{2(d_x+d_v)} w^{(i)} \left(h(\mathbf{x}_{t+1|t}^{(i)}, \mathbf{v}^{(i)}) - \mathbf{m}_{t+1|t} \right) \left(h(\mathbf{x}_{t+1|t}^{(i)}, \mathbf{v}^{(i)}) - \mathbf{m}_{t+1|t} \right)^\top$$

$$C_{t+1|t} = \sum_{i=0}^{2(d_x+d_v)} w^{(i)} \left(\mathbf{x}_{t+1|t}^{(i)} - \boldsymbol{\mu}_{t+1|t} \right) \left(h(\mathbf{x}_{t+1|t}^{(i)}, \mathbf{v}^{(i)}) - \mathbf{m}_{t+1|t} \right)^\top$$

Unscented Kalman Filter with Additive Noise

Prior $\mathbf{x}_{t+1} \mid \mathbf{z}_{0:t}, \mathbf{u}_{0:t} \sim \mathcal{N}(\boldsymbol{\mu}_{t+1|t}, \boldsymbol{\Sigma}_{t+1|t})$

Motion model $\mathbf{x}_{t+1} = f(\mathbf{x}_t, \mathbf{u}_t) + \mathbf{w}_t, \quad \mathbf{w}_t \sim \mathcal{N}(0, W)$

Obs. model $\mathbf{z}_{t+1} = h(\mathbf{x}_{t+1}) + \mathbf{v}_{t+1}, \quad \mathbf{v}_{t+1} \sim \mathcal{N}(0, V)$

Predict
$$\boldsymbol{\mu}_{t+1|t} = \sum_{i=0}^{2d_x} v^{(i)} f(\mathbf{x}_{t|t}^{(i)}, \mathbf{u}_t), \quad \mathbf{x}_{t|t}^{(0)} = \boldsymbol{\mu}_{t|t}, \quad \mathbf{x}_{t|t}^{(i)} = \boldsymbol{\mu}_{t|t} \pm \sqrt{\frac{d_x}{1-v^{(0)}}} \left[\sqrt{\boldsymbol{\Sigma}_{t|t}} \right]_i$$

$$\boldsymbol{\Sigma}_{t+1|t} = \sum_{i=0}^{2d_x} w^{(i)} \left(f(\mathbf{x}_{t|t}^{(i)}, \mathbf{u}_t) - \boldsymbol{\mu}_{t+1|t} \right) \left(f(\mathbf{x}_{t|t}^{(i)}, \mathbf{u}_t) - \boldsymbol{\mu}_{t+1|t} \right)^\top + W$$

Update
$$\boldsymbol{\mu}_{t+1|t+1} = \boldsymbol{\mu}_{t+1|t} + K_{t+1|t}(\mathbf{z}_{t+1} - \mathbf{m}_{t+1|t})$$

$$\boldsymbol{\Sigma}_{t+1|t+1} = \boldsymbol{\Sigma}_{t+1|t} - K_{t+1|t} S_{t+1|t} K_{t+1|t}^\top$$

Kalman gain
$$K_{t+1|t} = C_{t+1|t} S_{t+1|t}^{-1}$$

$$\mathbf{m}_{t+1|t} = \sum_{i=0}^{2d_x} v^{(i)} h(\mathbf{x}_{t+1|t}^{(i)}), \quad \mathbf{x}_{t+1|t}^{(0)} = \boldsymbol{\mu}_{t+1|t}, \quad \mathbf{x}_{t+1|t}^{(i)} = \boldsymbol{\mu}_{t+1|t} \pm \sqrt{\frac{d_x}{1-v^{(0)}}} \left[\sqrt{\boldsymbol{\Sigma}_{t+1|t}} \right]_i$$

$$S_{t+1|t} = \sum_{i=0}^{2d_x} w^{(i)} \left(h(\mathbf{x}_{t+1|t}^{(i)}) - \mathbf{m}_{t+1|t} \right) \left(h(\mathbf{x}_{t+1|t}^{(i)}) - \mathbf{m}_{t+1|t} \right)^\top + V$$

$$C_{t+1|t} = \sum_{i=0}^{2d_x} w^{(i)} \left(\mathbf{x}_{t+1|t}^{(i)} - \boldsymbol{\mu}_{t+1|t} \right) \left(h(\mathbf{x}_{t+1|t}^{(i)}) - \mathbf{m}_{t+1|t} \right)^\top$$

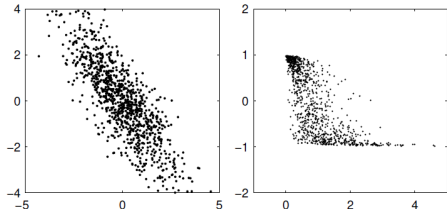
Outline

Nonlinear Kalman Filter

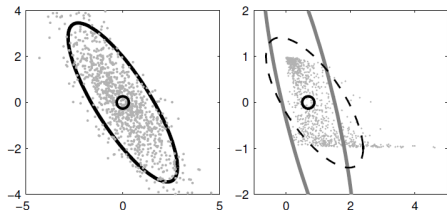
Extended Kalman Filter

Unscented Kalman Filter

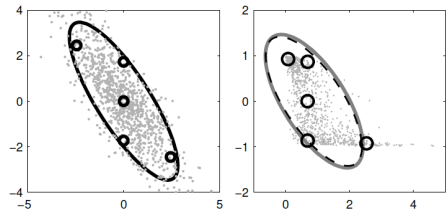
Comparison of EKF and UKF



Applying a nonlinear transformation to the random variable on the left results in a new random variable (right)



EKF: Linearization-based approximation to the transformation formed by calculating the curvature at the mean (solid) and true covariance (dashed).



UKF: Unscented transform approximation to the transformation formed by propagating sigma points through the nonlinearity and estimating the covariance from them (solid) and true covariance (dashed).

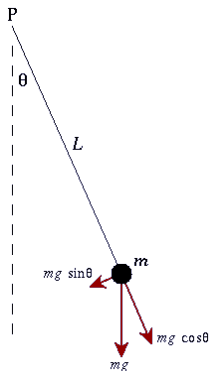
Noisy Pendulum Tracking

- ▶ Consider a pendulum consisting of mass m hanging by string of length L from pivot point P
- ▶ The differential equation of motion for the pendulum is obtained using Newton's second law for rotational systems which relates the net external torque τ (position \times force) to the product of the moment of inertia $I = mL^2$ and the angular acceleration $\ddot{\theta}(t)$:

$$\tau = -mgL \sin \theta(t) = mL^2 \ddot{\theta}(t) \quad \Rightarrow \quad \ddot{\theta}(t) = -\frac{g}{L} \sin \theta(t) + \underbrace{w(t)}_{\text{noise} \sim \mathcal{N}(0, q)}$$

- ▶ The model can be converted into a state-space motion model with state $\mathbf{x}(t) := (\theta(t), \omega(t))^T$, where $\omega(t) := \dot{\theta}(t)$ as follows:

$$\frac{d}{dt} \begin{pmatrix} \theta(t) \\ \omega(t) \end{pmatrix} = \begin{bmatrix} \omega(t) \\ -\frac{g}{L} \sin(\theta(t)) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} w(t)$$



Discrete-time Model

- ▶ **Motion model:** Euler discretization of the pendulum state-space model with sampling period τ leads to:

$$\mathbf{x}_{t+1} = \begin{pmatrix} \theta_{t+1} \\ \omega_{t+1} \end{pmatrix} = \underbrace{\begin{bmatrix} \theta_t + \tau\omega_t \\ \omega_t - \tau\frac{g}{L}\sin\theta_t \end{bmatrix}}_{f(\mathbf{x}_t, \mathbf{w}_t)} + \mathbf{w}_t, \quad \mathbf{w}_t \sim \mathcal{N}\left(\mathbf{0}, \underbrace{q \begin{bmatrix} \frac{\tau^3}{3} & \frac{\tau^2}{2} \\ \frac{\tau^2}{2} & \tau \end{bmatrix}}_W\right)$$

- ▶ **Observation model:** estimate the angle θ_t and angular velocity ω_t of the pendulum using measurements of its deviation from rest position:

$$z_t = \underbrace{L \sin(\theta_t)}_{h(\mathbf{x}_t, v_t)} + v_t, \quad v_t \sim \mathcal{N}(0, V)$$

Extended Kalman Filter Prediction Step

► **Prior:** $\mathbf{x}_t \mid \mathbf{z}_{0:t} \sim \mathcal{N}(\boldsymbol{\mu}_{t|t}, \boldsymbol{\Sigma}_{t|t})$ with $\boldsymbol{\mu}_{t|t} = \begin{bmatrix} \mu_{t|t}^\theta \\ \mu_{t|t}^\omega \end{bmatrix}$

► **Motion model Jacobian:**

$$F_t := \begin{bmatrix} 1 & \tau \\ -\tau \frac{g}{L} \cos \mu_{t|t}^\theta & 1 \end{bmatrix} \quad Q_t := I$$

► **EKF prediction step:**

$$\boldsymbol{\mu}_{t+1|t} = f(\boldsymbol{\mu}_{t|t}, \mathbf{0}) = \begin{bmatrix} \mu_{t|t}^\theta + \tau \mu_{t|t}^\omega \\ \mu_{t|t}^\omega - \tau \frac{g}{L} \sin \mu_{t|t}^\theta \end{bmatrix}$$
$$\boldsymbol{\Sigma}_{t+1|t} = F_t \boldsymbol{\Sigma}_{t|t} F_t^\top + Q_t W Q_t^\top$$

Extended Kalman Filter Update Step

- ▶ **Prediction:** $\mathbf{x}_{t+1} \mid \mathbf{z}_{0:t} \sim \mathcal{N}(\boldsymbol{\mu}_{t+1|t}, \boldsymbol{\Sigma}_{t+1|t})$ with $\boldsymbol{\mu}_{t+1|t} = \begin{bmatrix} \mu_{t+1|t}^\theta \\ \mu_{t+1|t}^\omega \end{bmatrix}$

- ▶ **Observation model Jacobian:**

$$H_{t+1} := \begin{bmatrix} L \cos \mu_{t+1|t}^\theta & 0 \end{bmatrix} \quad R_{t+1} := I$$

- ▶ **Innovation:** $r_{t+1|t} = z_{t+1} - L \sin(\mu_{t+1|t}^\theta)$

- ▶ **Measurement/innovation covariance:** $S_{t+1|t} = H_{t+1} \boldsymbol{\Sigma}_{t+1|t} H_{t+1}^\top + V$

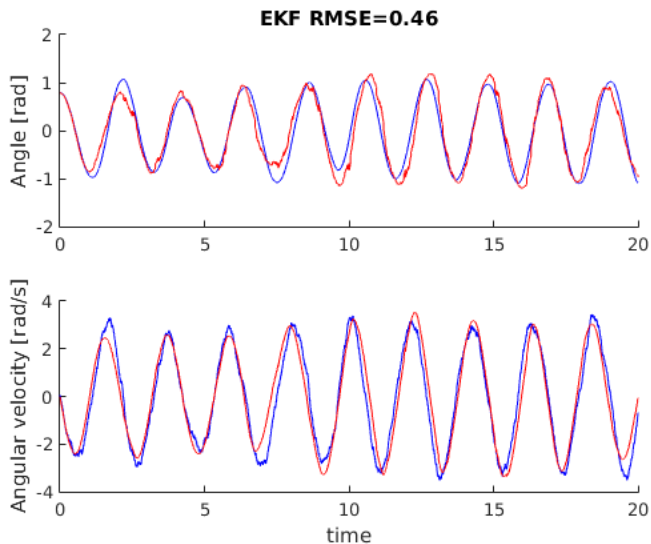
- ▶ **State-measurement cross-covariance:** $\boldsymbol{\Sigma}_{t+1|t} H_{t+1}^\top$

- ▶ **Kalman gain:** $K_{t+1|t} = \boldsymbol{\Sigma}_{t+1|t} H_{t+1}^\top S_{t+1|t}^{-1}$

- ▶ **EKF update step:**
$$\boldsymbol{\mu}_{t+1|t+1} = \boldsymbol{\mu}_{t+1|t} + K_{t+1|t} r_{t+1|t}$$
$$\boldsymbol{\Sigma}_{t+1|t+1} = \boldsymbol{\Sigma}_{t+1|t} - K_{t+1|t} H_{t+1} \boldsymbol{\Sigma}_{t+1|t}$$

EKF Performance

- ▶ $\tau = 0.001$, $q = 0.3$, $g = 9.81$, $L = 1$, $V = 0.64$
- ▶ Prediction at 1000 Hz, update at 20 Hz



Unscented Kalman Filter Prediction Step

► **Prior:** $\mathbf{x}_t \mid z_{0:t} \sim \mathcal{N}(\boldsymbol{\mu}_{t|t}, \boldsymbol{\Sigma}_{t|t})$

► **Sigma points:**

$$v^{(0)} < 1, \quad w^{(0)} \geq v^{(0)}, \quad v^{(i)} = w^{(i)} = \frac{1 - v^{(0)}}{2d_x}, \quad i = 1, \dots, 2d_x$$

$$\mathbf{x}_{t|t}^{(0)} = \boldsymbol{\mu}_{t|t}, \quad \mathbf{x}_{t|t}^{(i)} = \boldsymbol{\mu}_{t|t} \pm \sqrt{\frac{d_x}{1 - v^{(0)}}} \left[\sqrt{\boldsymbol{\Sigma}_{t|t}} \right]_i, \quad i = 1, \dots, 2d_x$$

► **UKF prediction step:**

$$\boldsymbol{\mu}_{t+1|t} = \sum_{i=0}^{2d_x} v^{(i)} f(\mathbf{x}_{t|t}^{(i)}, \mathbf{0})$$

$$\boldsymbol{\Sigma}_{t+1|t} = \sum_{i=0}^{2d_x} w^{(i)} \left(f(\mathbf{x}_{t|t}^{(i)}, \mathbf{0}) - \boldsymbol{\mu}_{t+1|t} \right) \left(f(\mathbf{x}_{t|t}^{(i)}, \mathbf{0}) - \boldsymbol{\mu}_{t+1|t} \right)^\top + W$$

Unscented Kalman Filter Update Step

- ▶ **Sigma points:**

$$\mathbf{x}_{t+1|t}^{(0)} = \boldsymbol{\mu}_{t+1|t}, \quad \mathbf{x}_{t+1|t}^{(i)} = \boldsymbol{\mu}_{t+1|t} \pm \sqrt{\frac{d_x}{1 - \nu^{(0)}}} \left[\sqrt{\boldsymbol{\Sigma}_{t+1|t}} \right]_i, \quad i = 1, \dots, 2d_x$$

- ▶ **Expected measurement:** $m_{t+1|t} = \sum_{i=0}^{2d_x} \nu^{(i)} h(\mathbf{x}_{t+1|t}^{(i)}, 0)$

- ▶ **Innovation:** $r_{t+1|t} = z_{t+1} - m_{t+1|t}$

- ▶ **Measurement/innovation covariance:**

$$S_{t+1|t} = \sum_{i=0}^{2d_x} w^{(i)} \left(h(\mathbf{x}_{t+1|t}^{(i)}, 0) - m_{t+1|t} \right) \left(h(\mathbf{x}_{t+1|t}^{(i)}, 0) - m_{t+1|t} \right)^\top + V$$

- ▶ **State-measurement cross-covariance:**

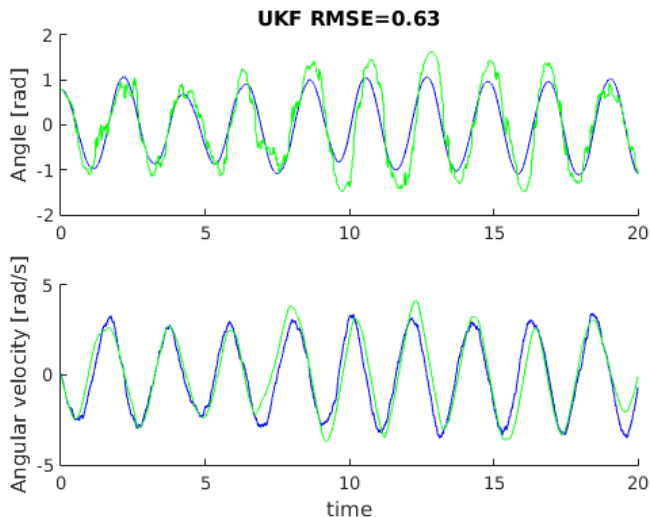
$$C_{t+1|t} = \sum_{i=0}^{2d_x} w^{(i)} \left(\mathbf{x}_{t+1|t}^{(i)} - \boldsymbol{\mu}_{t+1|t} \right) \left(h(\mathbf{x}_{t+1|t}^{(i)}, 0) - m_{t+1|t} \right)^\top$$

- ▶ **Kalman gain:** $K_{t+1|t} = C_{t+1|t} S_{t+1|t}^{-1}$

- ▶ **UKF update step:**
$$\begin{aligned} \boldsymbol{\mu}_{t+1|t+1} &= \boldsymbol{\mu}_{t+1|t} + K_{t+1|t} r_{t+1|t} \\ \boldsymbol{\Sigma}_{t+1|t+1} &= \boldsymbol{\Sigma}_{t+1|t} - K_{t+1|t} S_{t+1|t} K_{t+1|t}^\top \end{aligned}$$

UKF Performance

- ▶ $\tau = 0.001$, $q = 0.3$, $g = 9.81$, $L = 1$, $V = 0.64$
- ▶ Prediction at 1000 Hz, update at 20 Hz



UKF vs EKF Predicted Covariance

► Prior: $\mathcal{N}\left(\left(\frac{\pi}{4}\right), \begin{bmatrix} 2 & -0.3 \\ -0.3 & 0.5 \end{bmatrix}\right)$

► One prediction step with parameters $\tau = 1$, $g = 9.81$, $L = 1$

