State estimation of nonlinear systems using the Unscented Kalman Filter

João Almeida

Dynamic Stochastic Filtering, Prediction, and Smoothing

Instituto Superior Técnico

July 7th, 2010

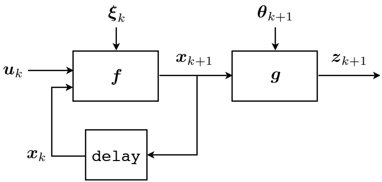
Outline

Introduction

Discrete-time nonlinear plant

- Discrete-time nonlinear plant

$$\boldsymbol{x}_{k+1} = \boldsymbol{f}(\boldsymbol{x}_k, \boldsymbol{u}_k, \boldsymbol{\xi}_k)$$
$$\boldsymbol{x}_0 \sim \mathcal{N}(\bar{\boldsymbol{x}}_0, \boldsymbol{\Sigma}_0)$$
$$\boldsymbol{z}_k = \boldsymbol{g}(\boldsymbol{x}_k, \boldsymbol{\theta}_k)$$



- For clarity, assume additive zero-mean Gaussian white-noise

$$\boldsymbol{x}_{k+1} = \boldsymbol{f}(\boldsymbol{x}_k, \boldsymbol{u}_k) + \mathbf{L}_k \boldsymbol{\xi}_k$$
$$\boldsymbol{z}_k = \boldsymbol{g}(\boldsymbol{x}_k) + \mathbf{D}_k \boldsymbol{\theta}_k$$

$\boldsymbol{\xi}_k \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Xi})$, $\boldsymbol{\theta}_k \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Theta})$, both uncorrelated

Introduction

Problem formulation

Given the **input** up to time $k-1$

$$\{\boldsymbol{u}_i : 0 \leq i \leq k-1\}$$

and the **observations** up to time k

$$\{\boldsymbol{z}_i : 0 \leq i \leq k\},$$

compute an **estimate** of $\boldsymbol{x}_k$ in a
minimum mean squared error sense

Introduction

Optimal solution

- Cannot be described by a finite number of parameters
- However, under some simplifying assumptions, it has the form

$$\begin{aligned}\hat{\boldsymbol{x}}_{k+1|k+1} &= (\text{prediction of } \boldsymbol{x}_{k+1}) \\ &\quad + \mathbf{H}_{k+1}[\boldsymbol{z}_{k+1} - (\text{prediction of } \boldsymbol{z}_{k+1})] \\ \mathbf{P}_{k+1|k+1} &= \mathbf{P}_{k+1|k} - \mathbf{H}_{k+1}\mathbf{P}_{\tilde{\boldsymbol{z}}_{k+1}}\mathbf{H}_{k+1}^{\top}\end{aligned}$$

- Optimal terms in the recursion

$$\begin{aligned}\hat{\boldsymbol{x}}_{k+1|k} &= \mathsf{E}\{\boldsymbol{f}(\hat{\boldsymbol{x}}_{k|k}, \boldsymbol{u}_k, \boldsymbol{\xi}_k)\} \\ \hat{\boldsymbol{z}}_{k+1} &= \mathsf{E}\{\boldsymbol{g}(\hat{\boldsymbol{x}}_{k+1|k}, \boldsymbol{\theta}_k)\} \\ \mathbf{H}_k &= \mathbf{P}_{\boldsymbol{x}_k\boldsymbol{z}_k}\mathbf{P}_{\tilde{\boldsymbol{z}}_k}^{-1}\end{aligned}$$

- Hard to compute in closed-form

Introduction

Optimal solution (Cont.)

The **Extended Kalman Filter** (EKF) and the **Unscented Kalman Filter** (UKF) use different methods to find an **approximate solution** for the optimal terms

Extended Kalman Filter

Review

- Optimal predictions

$$\hat{\boldsymbol{x}}_{k+1|k} = \mathsf{E}\{\boldsymbol{f}(\hat{\boldsymbol{x}}_{k|k}, \boldsymbol{u}_k, \boldsymbol{\xi}_k)\}$$
$$\hat{\boldsymbol{z}}_{k+1} = \mathsf{E}\{\boldsymbol{g}(\hat{\boldsymbol{x}}_{k+1|k}, \boldsymbol{\theta}_k)\}$$

- EKF predictions

$$\hat{\boldsymbol{x}}_{k+1|k} = \boldsymbol{f}(\hat{\boldsymbol{x}}_{k|k}, \boldsymbol{u}_k, \mathbf{0})$$
$$\hat{\boldsymbol{z}}_{k+1} = \boldsymbol{g}(\hat{\boldsymbol{x}}_{k+1|k}, \mathbf{0})$$

Extended Kalman Filter

Review (Cont.)

- Covariance propagation: apply Kalman Filter equations to the linearized system

$$\boldsymbol{x}_{k+1} \approx \hat{\mathbf{A}}_k \boldsymbol{x}_k + \mathbf{L}_k \boldsymbol{\xi}_k$$
$$\boldsymbol{z}_{k+1} \approx \hat{\mathbf{C}}_k \boldsymbol{x}_k + \mathbf{D}_k \boldsymbol{\theta}_k$$

where

$$\hat{\mathbf{A}}_k = \left. \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}} \right|_{\substack{\boldsymbol{x}=\hat{\boldsymbol{x}}_k \\ \boldsymbol{u}=\boldsymbol{u}_k}} \quad \hat{\mathbf{C}}_k = \left. \frac{\partial \boldsymbol{g}}{\partial \boldsymbol{x}} \right|_{\substack{\boldsymbol{x}=\hat{\boldsymbol{x}}_k \\ \boldsymbol{u}=\boldsymbol{u}_k}}$$

Extended Kalman Filter

Algorithm

Initialization: $\hat{\boldsymbol{x}}_{0|0} = \bar{\boldsymbol{x}}_0, \mathbf{P}_{0|0} = \boldsymbol{\Sigma}_0$
Main cycle: for $k = 0, 1, 2, \ldots$

1. Predict step:

$$\hat{\boldsymbol{x}}_{k+1|k} = \boldsymbol{f}(\hat{\boldsymbol{x}}_{k|k}, \boldsymbol{u}_k)$$
$$\mathbf{P}_{k+1|k} = \hat{\mathbf{A}}_k \mathbf{P}_{k|k} \hat{\mathbf{A}}_k^\top + \mathbf{L}_k \boldsymbol{\Xi} \mathbf{L}_k^\top$$

2. Measurement update step:

$$\hat{\boldsymbol{z}}_{k+1} = \boldsymbol{g}(\hat{\boldsymbol{x}}_{k+1|k})$$
$$\mathbf{H}_{k+1} = \mathbf{P}_{k+1|k} \hat{\mathbf{C}}_k^\top \left[\hat{\mathbf{C}}_k \mathbf{P}_{k+1|k} \hat{\mathbf{C}}_k^\top + \mathbf{D}_k \boldsymbol{\Theta} \mathbf{D}_k^\top \right]^{-1}$$
$$\hat{\boldsymbol{x}}_{k+1|k+1} = \hat{\boldsymbol{x}}_{k+1|k} + \mathbf{H}_{k+1}(\boldsymbol{z}_{k+1} - \hat{\boldsymbol{z}}_{k+1})$$
$$\mathbf{P}_{k+1|k+1} = \left[\mathbf{I} - \mathbf{H}_{k+1} \hat{\mathbf{C}}_k \right] \mathbf{P}_{k+1|k}$$

Unscented Kalman Filter

Random variables and nonlinear transformations

- Consider a random variable $\boldsymbol{x}$ and a nonlinear map

$$\boldsymbol{f} : \mathbb{R}^n \to \mathbb{R}^n, \quad \boldsymbol{x} \mapsto \boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{z}$$

- If $\boldsymbol{x}$ has mean $\bar{\boldsymbol{x}}$ and covariance $\mathbf{P}_{\boldsymbol{x}}$, what is the mean and covariance of $\boldsymbol{z}$?
- For example, EKF assumes $\bar{\boldsymbol{z}} = \boldsymbol{f}(\bar{\boldsymbol{x}})$

Unscented Kalman Filter

Transformation example

Example: $\boldsymbol{x} = (x_1, x_2) \sim \mathcal{N}(\bar{\boldsymbol{x}}, \boldsymbol{P_x})$
$(z_1, z_2) = \boldsymbol{f}(x_1, x_2) = (x_1^2 + x_1(1 - x_2), x_2(x_1 - 2))$
Mean and covariances computed by Monte Carlo methods

Figure: (x_1, x_2)

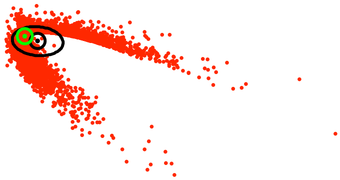


Figure: (z_1, z_2)

Unscented Kalman Filter

Unscented transformation

- Proposed by Julier, Uhlmann, and Durrant-Whyte [1995, 2000], further developments by Wan and van der Merwe [2001]
- How it works:
  1. generate a set of points whose sample mean and covariance match those of $\boldsymbol{x}$
  2. propagate them through function $\boldsymbol{f}(\cdot)$
  3. compute the sample mean and covariance of the propagated points
- Resembles a Monte Carlo method. However, the sample points are **not drawn at random**: they are **deterministically chosen**

Unscented Kalman Filter

Unscented transformation (cont.)

1. Form the set of $2L + 1$ *sigma points* (L is the state dimension)

$$\mathcal{X}_0 = \bar{\boldsymbol{x}}$$
$$\mathcal{X}_i = \bar{\boldsymbol{x}} + \left(\sqrt{(L+\lambda)\mathbf{P}_{\boldsymbol{x}}}\right)_i, \quad i = 1, \ldots, L$$
$$\mathcal{X}_i = \bar{\boldsymbol{x}} - \left(\sqrt{(L+\lambda)\mathbf{P}_{\boldsymbol{x}}}\right)_{i-L}, \quad i = L+1, \ldots, 2L$$

where $(\mathbf{X})_i$ denotes the ith column of matrix $\mathbf{X}$

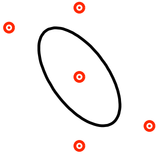
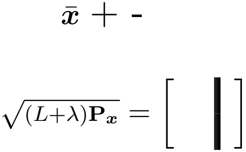


Figure: Sigma points



Unscented Kalman Filter

Unscented transformation (Cont.)

2. Transform each of the sigma points

$$\mathcal{Z}_i = \boldsymbol{f}(\mathcal{X}_i), \quad i = 0, \ldots, 2L$$

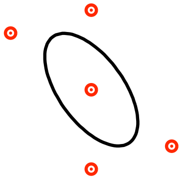


Figure: Original sigma points

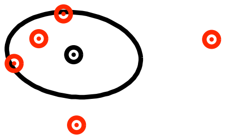


Figure: Transformed sigma points

Unscented Kalman Filter

Unscented transformation (Cont.)

3. Mean and covariance estimates for $\boldsymbol{z}$, cross-covariance of $\boldsymbol{x}$ and $\boldsymbol{z}$

$$\bar{\boldsymbol{z}} = \sum_{i=0}^{2L} W_i^{(m)} \mathcal{Z}_i \qquad \mathbf{P}_{\boldsymbol{z}} = \sum_{i=0}^{2L} W_i^{(c)} \left(\mathcal{Z}_i - \bar{\boldsymbol{z}}\right)\left(\mathcal{Z}_i - \bar{\boldsymbol{z}}\right)^\top$$

$$\mathbf{P}_{\boldsymbol{xz}} = \sum_{i=0}^{2L} W_i^{(c)} \left(\mathcal{X}_i - \bar{\boldsymbol{x}}\right)\left(\mathcal{Z}_i - \bar{\boldsymbol{z}}\right)^\top$$

where the weights are defined as

$$W_0^{(m)} = \lambda/(L+\lambda), \quad W_0^{(c)} = \lambda/(L+\lambda) + (1-\alpha^2+\beta)$$
$$W_i^{(m)} = W_i^{(c)} = 1/\{2(L+\lambda)\}, \quad i = 1, \dots, 2L$$

Unscented Kalman Filter

Unscented transformation (Example)

Example

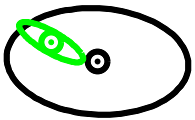


Figure: Linearization

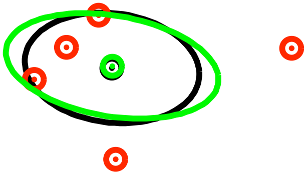


Figure: Unscented Transformation

The UT can be seen as a function with the following syntax

$$[\bar{z}, \mathbf{P}_{\boldsymbol{z}}, \mathbf{P}_{\boldsymbol{xz}}] = \mathsf{UT}\{\boldsymbol{f}(\cdot), \bar{\boldsymbol{x}}, \mathbf{P}_{\boldsymbol{x}}\}$$

Unscented Kalman Filter

Algorithm

Initialization: $\hat{\boldsymbol{x}}_{0|0} = \bar{\boldsymbol{x}}_0, \mathbf{P}_{0|0} = \boldsymbol{\Sigma}_0$
Main cycle: for $k = 0, 1, 2, \ldots$

1. Predict step:

$$[\hat{\boldsymbol{x}}_{k+1|k}, \mathbf{P}_{k+1|k}] = \mathsf{UT}\{\boldsymbol{f}(\cdot, \boldsymbol{u}_k), \hat{\boldsymbol{x}}_{k|k}, \mathbf{P}_{k|k}\}$$
$$\mathbf{P}_{k+1|k} = \mathbf{P}_{k+1|k} + \mathbf{L}_k \boldsymbol{\Xi} \mathbf{L}_k^\top$$

2. Measurement update step:

$$[\hat{\boldsymbol{z}}_{k+1}, \mathbf{P}_{\tilde{\boldsymbol{z}}}, \mathbf{P}_{\boldsymbol{x}\boldsymbol{z}}] = \mathsf{UT}\{\boldsymbol{g}(\cdot), \hat{\boldsymbol{x}}_{k+1|k}, \mathbf{P}_{k+1|k}\}$$
$$\mathbf{P}_{\tilde{\boldsymbol{z}}} = \mathbf{P}_{\tilde{\boldsymbol{z}}} + \mathbf{D}_k \boldsymbol{\Theta} \mathbf{D}_k^\top$$
$$\mathbf{H}_{k+1} = \mathbf{P}_{\boldsymbol{x}\boldsymbol{z}} \mathbf{P}_{\tilde{\boldsymbol{z}}}^{-1}$$
$$\hat{\boldsymbol{x}}_{k+1|k+1} = \hat{\boldsymbol{x}}_{k+1|k} + \mathbf{H}_{k+1}(\boldsymbol{z}_{k+1} - \hat{\boldsymbol{z}}_{k+1})$$
$$\mathbf{P}_{k+1|k+1} = \mathbf{P}_{k+1|k} - \mathbf{H}_{k+1} \mathbf{P}_{\tilde{\boldsymbol{z}}} \mathbf{H}_{k+1}$$

An illustrative example

ASC

- Simplified model of an autonomous surface craft (ASC) moving along straight lines

$$m\ddot{x}(t) + a_1\dot{x}(t)|\dot{x}(t)| + a_2\dot{x}^3(t) = b(u(t) + w(t))$$

where $w(t)$ represents an external disturbance caused by waves

- State-space representation

$$\dot{x}_1 = x_2$$
$$\dot{x}_2 = -\frac{a_1}{m}x_2|x_2| - \frac{a_2}{m}x_2^3 + \frac{b}{m}(u + w)$$

where $x_1(t) = x(t)$ and $x_2(t) = \dot{x}(t)$

An illustrative example

Waves

- Waves power spectral density

$$\Phi_{ww}(\omega) = \frac{100\omega^2}{81\omega^4 + 18\omega^2 + 1}$$

- State-space representation

$$\begin{aligned}
\dot{x}_{w1} &= x_{w2} \\
\dot{x}_{w2} &= -\frac{1}{9}x_{w1} - \frac{2}{3}x_{w2} + \frac{1}{9}\xi_w \\
w &= 10x_{w2}
\end{aligned}$$

where $\xi_w(t) \sim \mathcal{N}(0,1)$

An illustrative example

Sensors

On board the ASC, there are two sensors

- Position sensor

$$z_1 = x_1 + \theta_1$$

  where $\theta_1(t) \sim \mathcal{N}(0, \sigma_1^2)$

- Velocity sensor

$$z_2 = x_2 + \theta_2$$

  where $\theta_2 \sim \mathcal{N}(\bar{\theta}_2, \sigma_2^2)$ with $\bar{\theta}_2 \neq 0$

- θ_2 is not zero-mean. Written as the output of the LTI system

$$\dot{x}_s = 0 \qquad \theta_2 = x_s + \sigma_2 \xi_s$$

  where $x_s(0) = \bar{\theta}_2$ and $\xi_s \sim \mathcal{N}(0, 1)$

An illustrative example

Unknown parameters

- The parameters a_1, a_2, and b of the ASC are assumed unknown
- Estimates of the parameters: x_3, x_4, and x_5

$$\dot{x}_3(t) = \dot{x}_4(t) = \dot{x}_5(t) = 0$$
$$x_3(0) = a_1, x_4(0) = a_2, x_5(0) = b$$

An illustrative example

Augmented plant

- Plant's state

$$\boldsymbol{x} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 & x_5 & x_{w1} & x_{w2} & x_s \end{bmatrix}^\top$$

- Plant dynamics

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, u, \xi_w) = \begin{bmatrix} x_2 \\ -\frac{1}{m}\left(x_3 x_2 |x_2| + x_4 x_2^3 + x_5(u + 10x_{w2})\right) \\ 0 \\ 0 \\ 0 \\ x_{w2} \\ -\frac{1}{9}x_{w1} - \frac{2}{3}x_{w2} + \frac{1}{9}\xi_w \\ 0 \end{bmatrix}$$

An illustrative example

Augmented plant (Cont.)

- Output

$$\boldsymbol{z} = \begin{bmatrix} x_1 + \sigma_1 \xi_{s1} \\ x_2 + x_s + \sigma_2 \xi_{s2} \end{bmatrix}$$

where $\boldsymbol{\xi}_s(t) \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_2)$

An illustrative example

Discretization

- Continuous-time (augmented) plant is discretized using a step size of h seconds
- First-order discretization

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + h\boldsymbol{f}(\boldsymbol{x}_k, u_k, \xi_{w,k})$$

where $\boldsymbol{x}_k$ stands for $\boldsymbol{x}(kh)$

- Output

$$\boldsymbol{z}_k = \begin{bmatrix} x_{1,k} + \frac{\sigma_1}{h}\xi_{s1,k} \\ x_{2,k} + x_{s,k} + \frac{\sigma_2}{h}\xi_{s2,k} \end{bmatrix}$$

An illustrative example

Parameter and sensor configurations

Four different configurations of parameters and sensor covariances

- Parameters and their initial estimates

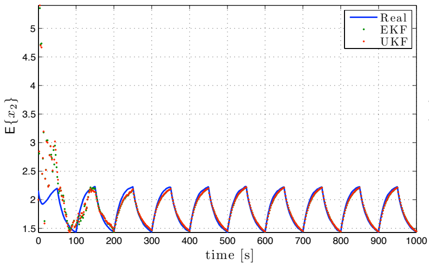
$$\mathcal{P}_1 = \{a_1 = 25, a_2 = 0, b = 1;$$
$$\hat{x}_3(0) = -5, \hat{x}_4(0) = 0, \hat{x}_5(0) = 10\}$$
$$\mathcal{P}_2 = \{a_1 = 25, a_2 = 2, b = 1;$$
$$\hat{x}_3(0) = -5, \hat{x}_4(0) = -5, \hat{x}_5(0) = 10\}$$

- Sensor covariances

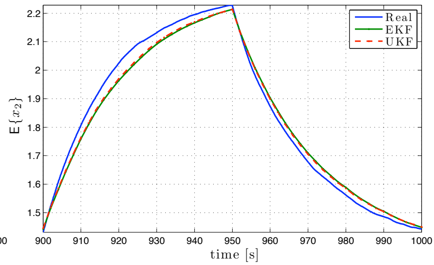
$$\mathcal{S}_1 = \{\sigma_1^2 = 1\,\mathrm{m}^2, \sigma_2^2 = 0.04\,(\mathrm{m/s})^2\}$$
$$\mathcal{S}_2 = \{\sigma_1^2 = 10\,\mathrm{m}^2, \sigma_2^2 = 0.4\,(\mathrm{m/s})^2\}$$

An illustrative example

Simulation results



(a) 0 to 1000 s



(b) 900 to 1000 s

Figure: Evolution of E$\{x_2\}$ for configuration ($\mathcal{P}_2$, $\mathcal{S}_2$)

An illustrative example

Simulation results (Cont.)

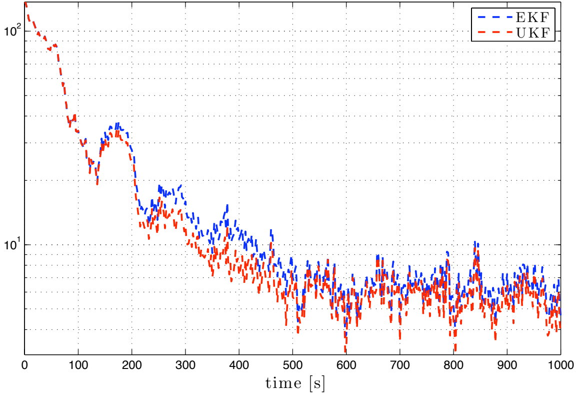


Figure: Evolution of $\mathrm{E}\{\tilde{\boldsymbol{x}}^\top \tilde{\boldsymbol{x}}\}$ (sensors: $\mathcal{S}_1$, parameters: $\mathcal{P}_2$)

An illustrative example

Simulation results (Cont.)

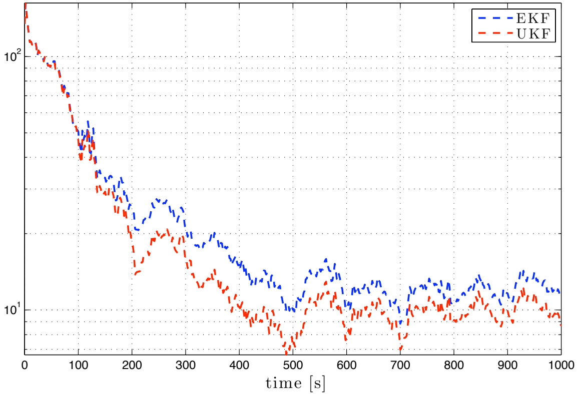


Figure: Evolution of $\mathrm{E}\{\tilde{\boldsymbol{x}}^\top \tilde{\boldsymbol{x}}\}$ (sensors: $\mathcal{S}_2$, parameters: $\mathcal{P}_2$)

An illustrative example

Simulation results (Comparison)

RMS values of both filters for different parameter and sensor configurations

Parameters	Sensors	Filter	RMS value of $\tilde{x}$
$\mathcal{P}_1$	$\mathcal{S}_1$	EKF	**9.769**
		UKF	**9.747** (0.23%)
	$\mathcal{S}_2$	EKF	**11.547**
		UKF	**11.487** (0.52%)
$\mathcal{P}_2$	$\mathcal{S}_1$	EKF	**12.022**
		UKF	**11.588** (3.61%)
	$\mathcal{S}_2$	EKF	**14.003**
		UKF	**13.071** (6.66%)

$$\mathrm{RMS} = \sqrt{\frac{1}{N}\sum_{k=0}^{N} \mathsf{E}\{\tilde{\boldsymbol{x}}_k^\top \tilde{\boldsymbol{x}}_k\}}$$

Conclusion

Summary

- State estimation of a nonlinear system perturbed by additive sources of zero-mean Gaussian white-noise
- The EKF deals with a linearized version of the system and then applies the standard KF equations
- The UKF uses the unscented transformation to propagate the mean and covariance of the state
- The performance of both filters was compared through simulation. It is shown that
    - only when more nonlinear terms are included
    - or, under severe noise

  does the gain in performance of the UKF stands out

Conclusion

Remarks

- The Unscented transformation can also be applied to prediction and smoothing
- Also available: continuous-time version, unscented particle filter
- Like the EKF, there are **no guarantees of stability or optimality**
- In my humble opinion, if you are using an EKF, try using an UKF. The worst it can happen is that you spend some time implementing it (not much, if I did my job right)

Bibliography

Simon Julier, Jeffrey Uhlmann, and Hugh F. Durrant-Whyte. A new approach for filtering nonlinear systems. In *Proc. of Am. Contr. Conf.*, pages 1628–1632, Seattle, WA, 21–23 June 1995.

Simon Julier, Jeffrey Uhlmann, and Hugh F. Durrant-Whyte. A new method for nonlinear transformation of means and covariances in filters and estimators. *IEEE Transactions on Automatic Control*, 45(3):477–482, March 2000.

Simo Särkkä. On unscented kalman filtering for state estimation of continuous-time nonlinear systems. *IEEE Transactions on Automatic Control*, 52(9):1631–1641, September 2007.

Eric A. Wan and Rudolph van der Merwe. *The Unscented Kalman Filter*, chapter 7. Kalman Filtering and Neural Networks. John Wiley and Sons, Inc., 2001.