# Sigma-point methods

Sensor fusion & nonlinear filtering

Lars Hammarstrand

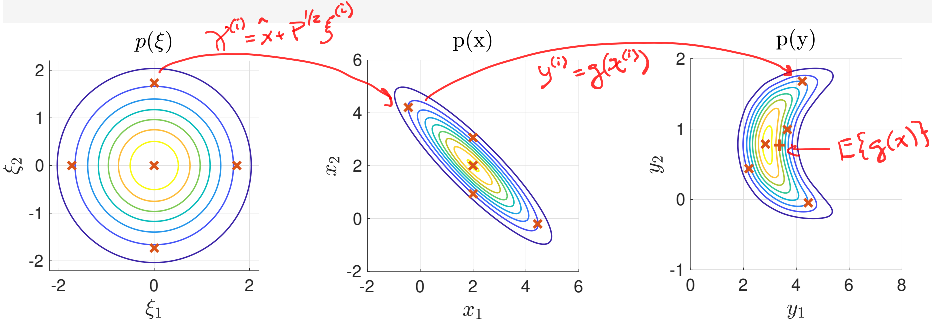
# SIGMA-POINT METHODS – INTEGRAL APPROXIMATION

## General idea

- Suppose $\mathbf{x} \sim \mathcal{N}(\hat{\mathbf{x}}, \mathbf{P})$, we can then approximate

$$\mathbb{E}\{\mathbf{g}(\mathbf{x})\} = \int \mathbf{g}(\hat{\mathbf{x}} + \mathbf{P}^{1/2}\xi)\mathcal{N}(\xi; \mathbf{0}, \mathbf{I})\, d\xi \approx \sum_{i=1}^{N} W_i \overbrace{\mathbf{g}(\underbrace{\hat{\mathbf{x}} + \mathbf{P}^{1/2}\xi^{(i)}}_{x^{(i)}})}^{y^{(i)}}$$

where $\xi^{(i)}$ are called $\sigma$-points and $W_i$ are weights.



# SIGMA-POINT METHOD IN GAUSSIAN FILTERING

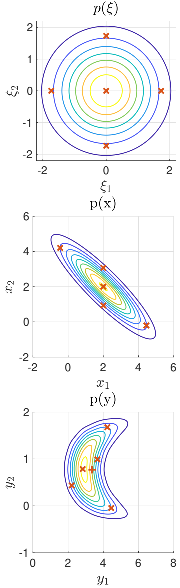
## Filtering idea

- Idea 1: *it is easier to approximate a probability distribution than it is to approximate an arbitrary nonlinear function or transformation*.
- Idea 2: if $\mathbf{y} = \mathbf{g}(\mathbf{x})$ and $\mathbf{x} \sim \mathcal{N}(\hat{\mathbf{x}}, \mathbf{P})$ we can approximate $p(\mathbf{y})$ using

$$\mathbb{E}\{\mathbf{y}\} \approx \boldsymbol{\mu}_y = \sum_{i=0}^{N} W_i \, \mathbf{g}(\mathcal{X}^{(i)})$$

$$\mathrm{Cov}\{\mathbf{y}\} \approx \sum_{i=0}^{N} W_i \overbrace{\left(\mathbf{g}(\mathcal{X}^{(i)}) - \boldsymbol{\mu}_y\right)\left(\mathbf{g}(\mathcal{X}^{(i)}) - \boldsymbol{\mu}_y\right)^T}^{\tilde{g}(x)}$$

where $\mathcal{X}^{(i)}$ are $\sigma$-points and $W_i$ the associated weights.



# THE UNSCENTED TRANSFORM (UT)

## Unscented transform (UT)

- Form a set of $2n+1$ $\sigma$-points as follows:

$$\sum_{i=0}^{2n} W_i \mathcal{X}^{(i)} = \hat{x}$$

$$\mathcal{X}^{(0)} = \hat{\mathbf{x}}$$

$$\mathcal{X}^{(i)} = \hat{\mathbf{x}} + \sqrt{\frac{n}{1-W_0}}\mathbf{P}_i^{1/2}, \quad i = 1, 2, \ldots, n,$$

$$\mathcal{X}^{(i+n)} = \hat{\mathbf{x}} - \sqrt{\frac{n}{1-W_0}}\mathbf{P}_i^{1/2}, \quad i = 1, 2, \ldots, n,$$

$$\sum_{i=0}^{2n} W_i = 1$$

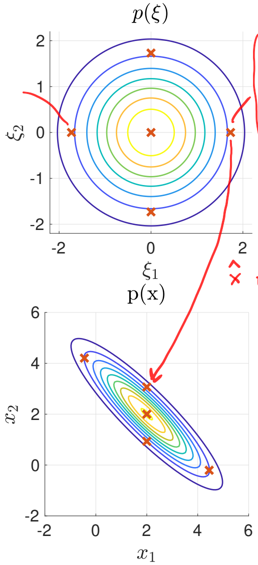
$$W_i = \frac{1-W_0}{2n}$$

$$\sum_{n=0}^{2n} W_i(\mathcal{X}^{(i)} - \hat{x})(\cdot)^T = P$$

where $\mathbf{P}_i^{1/2}$ is the $i$th column of $\mathbf{P}^{1/2}$.

- Other versions with more design parameters.
- If $\mathbf{x}$ is Gaussian, set $W_0 = 1 - n/3$.

$$\begin{bmatrix} -\sqrt{\frac{n}{1-W_0}} \\ 0 \end{bmatrix} \qquad \begin{bmatrix} \sqrt{\frac{n}{1-W_0}} \\ 0 \end{bmatrix}$$




$\hat{x} + P^{1/2}\xi$

# AN ILLUSTRATION OF THE UNSCENTED TRANSFORM

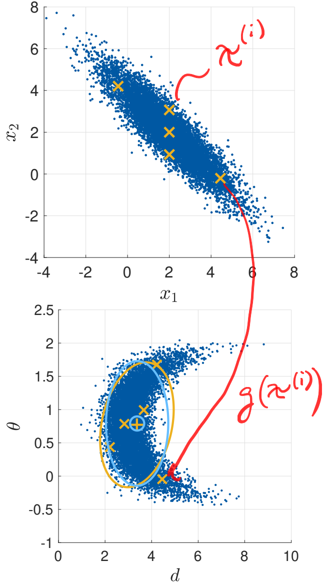
## UT for polar measurements

- Consider again the example where

$$\mathbf{y} = \mathbf{g}(\mathbf{x}) = \begin{bmatrix} \sqrt{x_1^2 + x_2^2} \\ \arctan\left(\frac{x_2}{x_1}\right) \end{bmatrix}$$

and

$$\mathbf{x} \sim \mathcal{N}\left( \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 & -1.8 \\ -1.8 & 2 \end{bmatrix} \right)$$



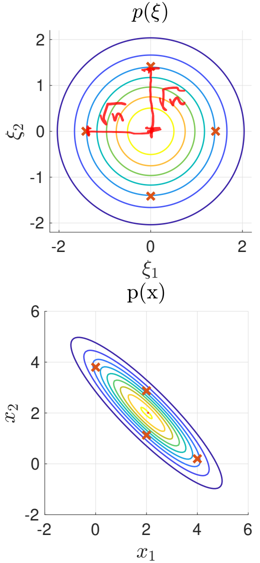
# THE CUBATURE RULE

## Cubature rule

- Forms a set of $2n$ $\sigma$-points as follows:

$$\mathcal{X}^{(i)} = \hat{\mathbf{x}} + \sqrt{n}\mathbf{P}_i^{1/2}, \qquad i = 1, 2, \dots, n,$$
$$\mathcal{X}^{(i+n)} = \hat{\mathbf{x}} - \sqrt{n}\mathbf{P}_i^{1/2}, \qquad i = 1, 2, \dots, n,$$
$$W_i = \frac{1}{2n}$$

where $\mathbf{P}_i^{1/2}$ is the $i$th column of $\mathbf{P}^{1/2}$.




Note:

- Special case of UT: $W_0 = 0$!
- No tuning parameters and no negative weights.
- Popularized in 2009.

# AN ILLUSTRATION OF THE CUBATURE RULE

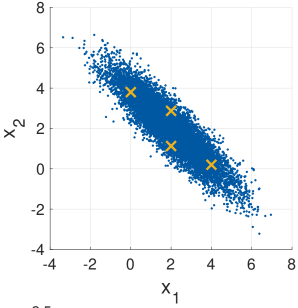
## Cubature rule for polar measurements

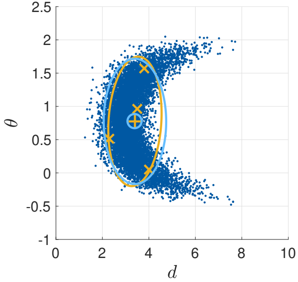
- Consider again the example where

$$\mathbf{y} = \mathbf{g}(\mathbf{x}) = \begin{bmatrix} \sqrt{x_1^2 + x_2^2} \\ \arctan\left(\frac{x_2}{x_1}\right) \end{bmatrix}$$

and

$$\mathbf{x} \sim \mathcal{N}\left(\begin{bmatrix} 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 & -1.8 \\ -1.8 & 2 \end{bmatrix}\right)$$



# AN ILLUSTRATION OF EKF

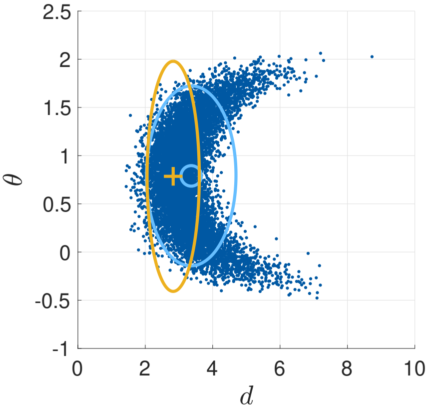
## EKF for polar measurements

- An EKF would instead handle

$$\mathbf{g}(\mathbf{x}) = \begin{bmatrix} \sqrt{x_1^2 + x_2^2} \\ \arctan\left(\frac{x_2}{x_1}\right) \end{bmatrix}$$

by linearization:

$$\mathbf{y} \approx \mathbf{g}\left(\begin{bmatrix} 2 \\ 2 \end{bmatrix}\right) + \mathbf{g}'\left(\begin{bmatrix} 2 \\ 2 \end{bmatrix}\right)\left(\mathbf{x} - \begin{bmatrix} 2 \\ 2 \end{bmatrix}\right)$$



# REMARKS ON THE UT AND THE CUBATURE RULE

Unscented transform:

- $W_0 = 1 - n/3$ is negative for $n > 3$ $\Rightarrow \text{Cov}\{\mathbf{y}\}$ can in become negative definite.
- UT variant with more design parameters but less intuitive.

Cubature rule:

- No theoretical risk of covariance becoming negative definite.
- No design parameters
- The spread of the $\sigma$-points increase with $n$, since $\mathcal{X}^{(i)} = \hat{\mathbf{x}} \pm \sqrt{n}\mathbf{P}_i^{1/2}$.

- Both the Unscented transform and the cubature rule computes the mean exactly for polynomials up to order 3, but the covariance is only exact when $\mathbf{g}(\mathbf{x})$ is of order 1.

# SELF ASSESSMENT

Check all that apply!

- The number of $\sigma$-points used in an UT grows linearly with the dimension of $\mathbf{x}$.
- The number of $\sigma$-points in UT is determined by the dimensionality of $\mathbf{y} = \mathbf{g}(\mathbf{x})$.
- We need to evaluate $\mathbf{g}(\mathbf{x})$ at $4n + 2$ different points since need $2n + 1$ points to approximate the mean and later $2n + 1$ points to approximate the covariance.

# SELF ASSESSMENT

Suppose that $n = 1$ and $p(x) = \mathcal{N}(x; 0, 1)$, the cubature rule is then simply

$$\mathbb{E}\{\mathbf{g}(x)\} \approx \mathbf{g}(-1)\frac{1}{2} + \mathbf{g}(1)\frac{1}{2}.$$

Compute the cubature rule approximation to $\mathbb{E}\{\mathbf{g}(x)\}$ when we further assume that

$$\mathbf{g}(x) = \begin{bmatrix} 1 + x + 5x^4 \\ 3x^2 + 4 \\ 2x^3 + x^5 \end{bmatrix}.$$

(You can also ask yourself which elements of $\mathbb{E}\{\mathbf{g}(x)\}$ that we have thus computed exactly.)

- $\begin{bmatrix} 3 & 3.5 & 1.5 \end{bmatrix}^T$
- $\begin{bmatrix} 6 & 7 & 3 \end{bmatrix}^T$
- $\begin{bmatrix} 7 & 7 & 3 \end{bmatrix}^T$
- $\begin{bmatrix} 6 & 7 & 0 \end{bmatrix}^T$