Sigma Point Belief Propagation

Florian Meyer, Ondrej Hlinka, and Franz Hlawatsch

Institute of Telecommunications, Vienna University of Technology, Austria

Contribution

- The sigma point (SP) filter is an attractive method for sequential Bayesian estimation.
- We develop the **sigma point belief propagation (SPBP) algorithm** by extending the SP filter to Bayesian inference on general factor graphs.

Sigma Point Basics

- Consider a random vector $\mathbf{x} \in \mathbb{R}^J$ whose mean $\boldsymbol{\mu}_{\mathbf{x}}$ and covariance matrix $\mathbf{C}_{\mathbf{x}}$ are known, and a transformed random vector $\mathbf{y} = H(\mathbf{x})$.
- SPs and corresponding weights $\{(\mathbf{x}^{(j)}, w^{(j)})\}_{j=0}^{2J}$ are calculated from $\boldsymbol{\mu}_{\mathbf{x}}$ and $\mathbf{C}_{\mathbf{x}}$ as described in [Julier et al., 1997], and SPs $\{\mathbf{y}^{(j)}\}_{j=0}^{2J}$ are obtained as $\mathbf{y}^{(j)} = H(\mathbf{x}^{(j)})$.
- From $\{(\mathbf{x}^{(j)}, \mathbf{y}^{(j)}, w^{(j)})\}_{j=0}^{2J}$, approximations $\tilde{\boldsymbol{\mu}}_{\mathbf{y}}$, $\tilde{\mathbf{C}}_{\mathbf{y}}$, and $\tilde{\mathbf{C}}_{\mathbf{x},\mathbf{y}}$ of the mean $\boldsymbol{\mu}_{\mathbf{y}}$ and the covariance matrices $\mathbf{C}_{\mathbf{y}}$ and $\mathbf{C}_{\mathbf{xy}}$ are calculated as a weighted sample mean or as weighted sample covariances.

Bayesian Estimation with SPs

- SPs can be used for Bayesian estimation of a random vector $\mathbf{x}$ from an observed vector

$$\mathbf{z} = \mathbf{y} + \mathbf{n}\,, \quad \text{with } \mathbf{y} = H(\mathbf{x})\,.$$

Here, the noise $\mathbf{n}$ is zero-mean and statistically independent of $\mathbf{x}$ and has a known covariance matrix $\mathbf{C}_{\mathbf{n}}$.

- Bayesian estimation relies on the posterior probability density functions (pdf)

$$f(\mathbf{x}|\mathbf{z}) \propto f(\mathbf{z}|\mathbf{x}) f(\mathbf{x})\,,$$

where $f(\mathbf{z}|\mathbf{x})$ is the likelihood function and $f(\mathbf{x})$ is the prior pdf.

- A closed-form calculation of $f(\mathbf{x}|\mathbf{z})$ is usually infeasible. An exception is the case where $H(\cdot)$ is linear, i.e., $H(\mathbf{x}) = \mathbf{H}\mathbf{x}$, and $\mathbf{x}$ and $\mathbf{n}$ are Gaussian. Then $f(\mathbf{x}|\mathbf{z})$ is also Gaussian, and $\boldsymbol{\mu}_{\mathbf{x}|\mathbf{z}}$ and $\mathbf{C}_{\mathbf{x}|\mathbf{z}}$ can be calculated in closed form from $\boldsymbol{\mu}_{\mathbf{x}}$, $\boldsymbol{\mu}_{\mathbf{y}}$, $\mathbf{C}_{\mathbf{x}}$, $\mathbf{C}_{\mathbf{y}}$, and $\mathbf{C}_{\mathbf{xy}}$.
- In the nonlinear case, $\boldsymbol{\mu}_{\mathbf{x}|\mathbf{z}}$ and $\mathbf{C}_{\mathbf{x}|\mathbf{z}}$ can be approximated by means of SPs [Julier et al., 1997]. This is done by using the closed-form expressions of the linear-Gaussian case, in which $\boldsymbol{\mu}_{\mathbf{y}}$, $\mathbf{C}_{\mathbf{y}}$, and $\mathbf{C}_{\mathbf{xy}}$ are replaced by the corresponding SP approximations $\tilde{\boldsymbol{\mu}}_{\mathbf{y}}$, $\tilde{\mathbf{C}}_{\mathbf{y}}$, and $\tilde{\mathbf{C}}_{\mathbf{xy}}$.
- The sequential version of this algorithm is the SP filter.

Belief Propagation

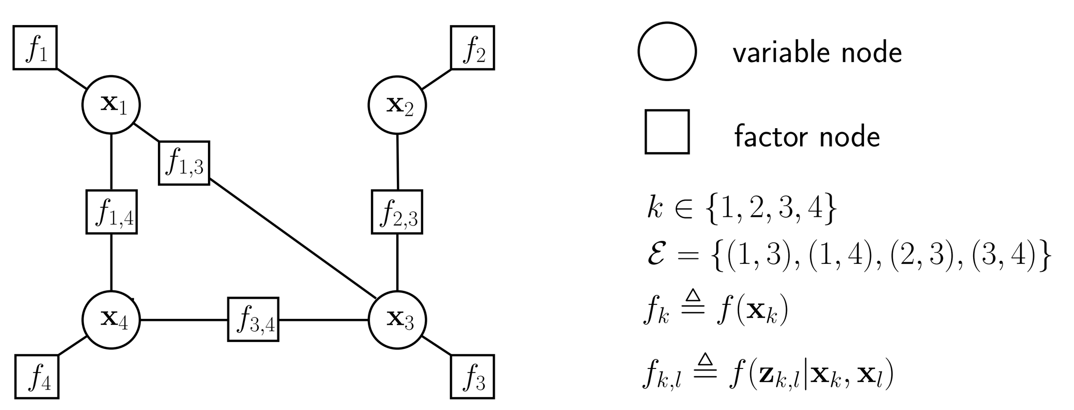
- Consider K state vectors $\mathbf{x}_k$, $k \in \{1,\ldots,K\}$, and let $\mathbf{x} = (\mathbf{x}_1^{\mathsf{T}} \cdots \mathbf{x}_K^{\mathsf{T}})^{\mathsf{T}}$ be the vector of all states. Furthermore consider noisy "pairwise" observations

$$\mathbf{z}_{k,l} = G(\mathbf{x}_k, \mathbf{x}_l) + \mathbf{v}_{k,l}\,.$$

- Assume that the posterior pdf $f(\mathbf{x}|\mathbf{z})$ factorizes as

$$f(\mathbf{x}|\mathbf{z}) \propto \left[\prod_{k=1}^{K} f(\mathbf{x}_k)\right] \prod_{(k',l)\in\mathcal{E}} f(\mathbf{z}_{k',l}|\mathbf{x}_{k'}, \mathbf{x}_l)\,.$$

- Representation by a factor graph (example):



- To obtain an estimate of $\mathbf{x}_k$, the "marginal" posterior pdf $f(\mathbf{x}_k|\mathbf{z}) = \int f(\mathbf{x}|\mathbf{z})\, d\mathbf{x}^{\sim k}$ is needed.
- An approximate marginal posterior ("belief") $b(\mathbf{x}_k) \approx f(\mathbf{x}_k|\mathbf{z})$ is obtained by executing the belief propagation (BP) message passing algorithm on the factor graph representing the joint posterior $f(\mathbf{x}|\mathbf{z})$.
- Belief of variable $\mathbf{x}_k$:

$$b(\mathbf{x}_k) \propto f(\mathbf{x}_k) \prod_{l\in\mathcal{N}_k} m_{l\to k}(\mathbf{x}_k)\,.$$

Here, $\mathcal{N}_k$ denotes the "neighbor" set of variable node $\mathbf{x}_k$, which comprises all variable nodes $\mathbf{x}_l$ with $l \neq k$ such that $(k,l) \in \mathcal{E}$.

- Message from variable node $\mathbf{x}_k$ to function node $f(\mathbf{z}_{k,l}|\mathbf{x}_k, \mathbf{x}_l)$:

$$q_{k\to l}(\mathbf{x}_k) = f(\mathbf{x}_k) \prod_{l'\in\mathcal{N}_k\setminus\{l\}} m_{l'\to k}(\mathbf{x}_k)\,.$$

- Message from function node $f(\mathbf{z}_{k,l}|\mathbf{x}_k, \mathbf{x}_l)$ to variable node $\mathbf{x}_k$:

$$m_{l\to k}(\mathbf{x}_k) = \int f(\mathbf{z}_{k,l}|\mathbf{x}_k, \mathbf{x}_l)\, q_{l\to k}(\mathbf{x}_l)\, d\mathbf{x}_l\,.$$

The Proposed SPBP Algorithm

- Equivalent formulation of the BP rules:

$$b(\mathbf{x}_k) \propto \int f(\mathbf{x}_k) \prod_{l\in\mathcal{N}_k} \left[f(\mathbf{z}_{k,l}|\mathbf{x}_k, \mathbf{x}_l)\, q_{l\to k}(\mathbf{x}_l)\, d\mathbf{x}_l\right]$$

$$q_{k\to l}(\mathbf{x}_k) = \int f(\mathbf{x}_k) \prod_{l'\in\mathcal{N}_k\setminus\{l\}} \left[f(\mathbf{z}_{k,l'}|\mathbf{x}_k, \mathbf{x}_{l'})\, q_{l'\to k}(\mathbf{x}_{l'}) d\mathbf{x}_{l'}\right].$$

- Let $\mathcal{N}_k = \{l_1, l_2, \ldots, l_{|\mathcal{N}_k|}\}$, and define the "composite" vectors $\bar{\mathbf{x}}_k \triangleq (\mathbf{x}_k^{\mathsf{T}}\, \mathbf{x}_{l_1}^{\mathsf{T}}\, \mathbf{x}_{l_2}^{\mathsf{T}} \cdots \mathbf{x}_{l_{|\mathcal{N}_k|}}^{\mathsf{T}})^{\mathsf{T}}$ ($\mathbf{x}_k$ and its neighbor states) and $\bar{\mathbf{z}}_k \triangleq (\mathbf{z}_{k,l_1}^{\mathsf{T}}\, \mathbf{z}_{k,l_2}^{\mathsf{T}} \cdots \mathbf{z}_{k,l_{|\mathcal{N}_k|}}^{\mathsf{T}})^{\mathsf{T}}$ (all observations involving $\mathbf{x}_k$).
- We can now write $b(\mathbf{x}_k)$ as

$$b(\mathbf{x}_k) \propto \int f(\bar{\mathbf{z}}_k|\bar{\mathbf{x}}_k)\, f(\bar{\mathbf{x}}_k)\, d\bar{\mathbf{x}}_k^{\sim k},$$

with the composite prior

$$f(\bar{\mathbf{x}}_k) \propto f(\mathbf{x}_k) \prod_{l\in\mathcal{N}_k} q_{l\to k}(\mathbf{x}_l)$$

and the composite likelihood function

$$f(\bar{\mathbf{z}}_k|\bar{\mathbf{x}}_k) = \prod_{l\in\mathcal{N}_k} f(\mathbf{z}_{k,l}|\mathbf{x}_k, \mathbf{x}_l)\,.$$

- Note that $f(\bar{\mathbf{z}}_k|\bar{\mathbf{x}}_k)$ corresponds to the composite observation model

$$\bar{\mathbf{z}}_k = \bar{\mathbf{y}}_k + \bar{\mathbf{v}}_k\,, \quad \text{with } \bar{\mathbf{y}}_k = H(\bar{\mathbf{x}}_k),$$

where $H(\bar{\mathbf{x}}_k) \triangleq \left((G(\mathbf{x}_k, \mathbf{x}_{l_1}))^{\mathsf{T}} \cdots (G(\mathbf{x}_k, \mathbf{x}_{l_{|\mathcal{N}_k|}}))^{\mathsf{T}}\right)^{\mathsf{T}}$ and $\bar{\mathbf{v}}_k \triangleq (\mathbf{v}_{k,l_1}^{\mathsf{T}} \cdots \mathbf{v}_{k,l_{|\mathcal{N}_k|}}^{\mathsf{T}})^{\mathsf{T}}$.

- We can finally express $b(\mathbf{x}_k)$ as the result of a marginalization:

$$b(\mathbf{x}_k) = \int b(\bar{\mathbf{x}}_k)\, d\bar{\mathbf{x}}_k^{\sim k},$$

with the composite belief

$$b(\bar{\mathbf{x}}_k) \propto f(\bar{\mathbf{z}}_k|\bar{\mathbf{x}}_k)\, f(\bar{\mathbf{x}}_k)\,.$$

- This expression of $b(\bar{\mathbf{x}}_k)$ has the same form as the expression $f(\mathbf{x}|\mathbf{z}) \propto f(\mathbf{z}|\mathbf{x}) f(\mathbf{x})$ that occurred in the earlier section "Bayesian Estimation with SPs" in the first column $\Rightarrow$ SPs can again be used for calculating an approximate mean and covariance of $b(\bar{\mathbf{x}}_k)$ and, in turn, of $b(\mathbf{x}_k)$.
- More specifically, $\tilde{\boldsymbol{\mu}}_{b(\bar{\mathbf{x}}_k)}$ and $\tilde{\mathbf{C}}_{b(\bar{\mathbf{x}}_k)}$ can be obtained by using the closed-form expressions of the linear-Gaussian case, in which $\boldsymbol{\mu}_{\bar{\mathbf{y}}_k}$, $\mathbf{C}_{\bar{\mathbf{y}}_k}$, and $\mathbf{C}_{\bar{\mathbf{x}}_k\bar{\mathbf{y}}_k}$ are replaced by the SP approximations $\tilde{\boldsymbol{\mu}}_{\bar{\mathbf{y}}_k}$, $\tilde{\mathbf{C}}_{\bar{\mathbf{y}}_k}$, and $\tilde{\mathbf{C}}_{\bar{\mathbf{x}}_k\bar{\mathbf{y}}_k}$.

SPBP algorithm

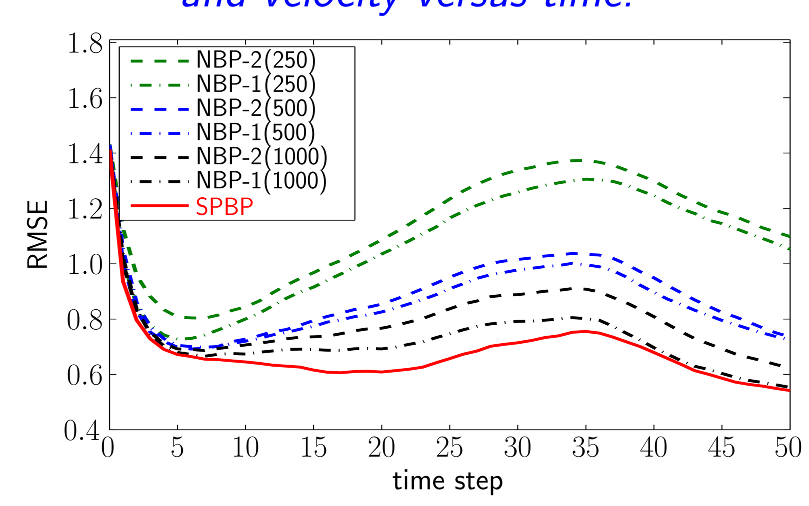
SP-based approximate calculation of the mean $\boldsymbol{\mu}_{b(\mathbf{x}_k)}$ and covariance matrix $\mathbf{C}_{b(\mathbf{x}_k)}$ of $b(\mathbf{x}_k)$:

1. Obtain the mean and covariance matrix of the composite prior $f(\bar{\mathbf{x}}_k) \propto f(\mathbf{x}_k) \prod_{l\in\mathcal{N}_k} q_{l\to k}(\mathbf{x}_l)$ as $\boldsymbol{\mu}_{\bar{\mathbf{x}}_k} = (\boldsymbol{\mu}_{\mathbf{x}_k}^{\mathsf{T}}\; \boldsymbol{\mu}_{l_1\to k}^{\mathsf{T}}\; \boldsymbol{\mu}_{l_2\to k}^{\mathsf{T}} \cdots \boldsymbol{\mu}_{l_{|\mathcal{N}_k|}\to k}^{\mathsf{T}})^{\mathsf{T}}$ and $\mathbf{C}_{\bar{\mathbf{x}}_k} = \mathrm{diag}\{\mathbf{C}_{\mathbf{x}_k}, \mathbf{C}_{l_1\to k}, \mathbf{C}_{l_2\to k}, \ldots, \mathbf{C}_{l_{|\mathcal{N}_k|}\to k}\}$, where $\boldsymbol{\mu}_{l\to k}$ and $\mathbf{C}_{l\to k}$ are the mean and covariance matrix of $q_{l\to k}(\mathbf{x}_l)$.
2. Use the "Bayesian estimation" SP scheme to calculate $\tilde{\boldsymbol{\mu}}_{b(\bar{\mathbf{x}}_k)} \approx \boldsymbol{\mu}_{b(\bar{\mathbf{x}}_k)}$ and $\tilde{\mathbf{C}}_{b(\bar{\mathbf{x}}_k)} \approx \mathbf{C}_{b(\bar{\mathbf{x}}_k)}$ representing $b(\bar{\mathbf{x}}_k)$ from $\boldsymbol{\mu}_{\bar{\mathbf{x}}_k}$ and $\mathbf{C}_{\bar{\mathbf{x}}_k}$.
3. Obtain $\tilde{\boldsymbol{\mu}}_{b(\mathbf{x}_k)}$ and $\tilde{\mathbf{C}}_{b(\mathbf{x}_k)}$ by extracting from $\tilde{\boldsymbol{\mu}}_{b(\bar{\mathbf{x}}_k)}$ and $\tilde{\mathbf{C}}_{b(\bar{\mathbf{x}}_k)}$ the elements corresponding to $\mathbf{x}_k$. (This emulates the marginalization $b(\mathbf{x}_k) = \int b(\bar{\mathbf{x}}_k)\, d\bar{\mathbf{x}}_k^{\sim k}$.)

Simulation Results

- We simulated a decentralized, cooperative, dynamic localization scenario [Wymeersch et al., 2009] using a network of three mobile sensors and two anchor sensors.
- The state $\mathbf{x}_{k,i}$ of mobile sensor $k \in \{1,2,3\}$ at time $i \in \{0,1,\ldots,50\}$ consists of the sensor's location and velocity. Each mobile sensor communicates with all other sensors, performs distance measurements relative to all other sensors, and estimates its own state.
- We compare the proposed SPBP algorithm with two nonparametric (random particle based) BP methods, referred to as NBP-1 [Ihler et al., 2005] and NBP-2 [Savic et al., 2013], both simulated with 250, 500, and 1000 particles.

Root-mean-square error (RMSE) of location and velocity versus time:



Communication and computation requirements:

	Communication between sensors [# transmitted real values]	Runtime on Xeon X5650 [seconds]
SPBP	500	0.61
NBP-1(250)	50000	1.53
NBP-2(250)	50000	2.01
NBP-1(500)	100000	5.16
NBP-2(500)	100000	7.27
NBP-1(1000)	200000	19.57
NBP-2(1000)	200000	28.10

- The proposed SPBP algorithm outperforms NBP-1 and NBP-2, and it requires significantly less communications and computations.

Conclusion

- The proposed SPBP algorithm extends the SP filter to Bayesian inference on general factor graphs.
- SPBP is well suited to certain decentralized inference problems in wireless sensor networks because of its low communication requirements.
- In a decentralized, cooperative self-localization scenario, SPBP can outperform nonparametric BP, and it requires significantly less communications and computations.

[Julier et al., 1997] S. J. Julier and J. K. Uhlmann, "A new extension of the Kalman filter to nonlinear systems," in *Proc. AeroSense-97*, 1997

[Wymeersch et al., 2009] H. Wymeersch, J. Lien, and M. Z. Win, "Cooperative localization in wireless networks," *Proc. IEEE*, 2009

[Ihler et al., 2005] A. T. Ihler, J. W. Fisher, R. L. Moses, and A. S. Willsky, "Nonparametric belief propagation for self-localization of sensor networks," *IEEE J. Sel. Areas Comm.*, 2005

[Savic et al., 2013] V. Savic and S. Zazo, "Cooperative localization in mobile networks using nonparametric variants of belief propagation," *Ad Hoc Networks*, 2013

TECHNISCHE UNIVERSITÄT WIEN Vienna University of Technology

institute of telecommunications

FWF Der Wissenschaftsfonds.