2010 IEEE International Conference on Multisensor Fusion and
Integration for Intelligent Systems
University of Utah, Salt Lake City, UT, USA, Sept. 5-7, 2010

Semi-Analytic Stochastic Linearization for Range-Based Pose Tracking

Frederik Beutler, Marco F. Huber, and Uwe D. Hanebeck

Abstract— **In range-based pose tracking, the translation and rotation of an object with respect to a global coordinate system has to be estimated. The ranges are measured between the target and the global frame. In this paper, an intelligent decomposition is introduced in order to reduce the computational effort for pose tracking. Usually, decomposition procedures only exploit conditionally linear models. In this paper, this principle is generalized to conditionally integrable substructures and applied to pose tracking. Due to a modified measurement equation, parts of the problem can even be solved analytically.**

I. Introduction

In many applications, the estimation of an object pose, i.e., the translation and the rotation, is essential, e.g., in large scale telepresence (Fig. 1) [1]. In [1], the pose of a human has to be tracked in order to steer the teleoperator. For tracking the user's pose, several emitters are located at known positions in a global coordinate system. They are emitting signals that are received by several sensors attached to the target frame. Based on the emitted and received signals, ranges between emitters and sensors can be determined. Due to disturbances and a nonlinear measurement equation, which describes the relationhip between measured ranges and pose, an exact estimator cannot be applied and so approximative estimators have to be used. Algorithms for estimating the pose based on range measurements are closed-form solutions [2], gradient descent algorithms [1], or state estimators [3]. Popular state estimators rely on the Gaussian assumption [4], [5], [6], e.g., the Unscented Kalman Filter, where all involved random variables are described by mean and covariance. Furthermore, the random variable for the measurement and for the state are assumed to be jointly Gaussian distributed. In order to calculate the mean and covariance, sample-based approaches are used. For an efficient implementation, the structure of the measurement and system equation can be exploited and so the number of sample points can be reduced. For conditionally linear substructures, the reader is referred to [3] and [7]. In order to reduce the number of sample points, the decomposition in conditionally linear substructures can be generalized to conditionally integrable substructures. In this case, nonlinear parts of the problem can be solved in closed form.

F. Beutler and U. D. Hanebeck are with the Intelligent Sensor-Actuator-Systems Laboratory (ISAS), Institute for Anthropomatics, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany. beutler@ieee.org, uwe.hanebeck@ieee.org.

Marco F. Huber is with the Variable Image Acquisition and Processing Research Group, Fraunhofer Institute of Optronics, System Technologies and Image Exploitation IOSB, Karlsruhe, Germany. marco.huber@ieee.org.

Fig. 1. Person using large scale telepresence.

Compared to the previous approach in [3], where the density of the translation and rotation has to be approximated for the filter step, merely the density of the rotation is processed approximately in the proposed approach, while the remaining part can be calculated in closed form. In doing so, the number of sample points can be decreased.

The structure of the paper is as follows. In the problem formulation in Sec. II, the measurement (see Sec. II-A) and system equation (see Sec. II-B) for pose tracking are described. The proposed approach makes use of a state estimator, the generalized Gaussian assumed density filter, which is described in Sec. II-C. This filter consists of a prediction step (Sec. II-C.1) and a filter step (Sec. II-C.2). In the filter step, the assumption that measurements and the state are jointly Gaussian distributed is applied. The proposed approach for pose tracking is shown in Sec. III. First, the measurement equation is modified in Sec. III-A. Based on this modified measurement equation, parts of the problem can be solved in closed form, where first the decomposition is explained in Sec. III-B. This decomposition is then used in order to calculate the mean (Sec. III-C), the covariance (Sec. III-D), and the cross-covariance (Sec. III-E), which is required for the filter step (see Sec. II-C.2). A short wrap-up of the filter step is shown in Sec. III-F. The proposed approach is compared to the standard decomposition via simulations in Sec. IV. Finally, the paper ends with conclusions.

II. Problem Formulation

In pose tracking, the translation $\underline{t}_k$ and the rotation $\underline{r}_k$ of an extended object have to be estimated. Due to the fact that the object is in motion, it is essential to consider the dynamic behavior of the object by means of the translation velocity $\dot{\underline{t}}_k$ and angular velocity $\underline{w}_k$. In order to estimate the state $\underline{x}_k$ of the object

$$\underline{x}_k = \begin{bmatrix} \underline{t}_k^{\mathrm{T}} & \dot{\underline{t}}_k^{\mathrm{T}} & \underline{r}_k^{\mathrm{T}} & \underline{w}_k^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}} ,$$

a state estimator is used.

A. Measurement Equation

The measurement equation describes the relationship between the measurement and the state. In range-based pose estimation, the measured ranges depend on the known landmarks located on the extended target, known landmarks in a static global coordinate system, and the unknown translation and rotation of the object with respect to the global coordinate system. The relationship between measured ranges and the translation and rotation is given by

$$\boldsymbol{d}_{i,j,k} = ||\underline{L}_j - \mathbf{D}(\underline{r}_k) \cdot \underline{M}_i - \underline{t}_k - \underline{v}_{i,j,k}||_2 , \qquad (1)$$

where $\mathbf{D}(\cdot)$ is the rotation matrix parametrized by the vector $\underline{r}_k$, which describes the rotation. $\underline{L}_j$ is the position of the jth landmark with respect to the global coordinate system and $\underline{M}_i$ is the position of the ith landmark with respect to the target coordinate system. $\underline{v}_{i,j,k}$ is the measurement noise between landmark $\underline{L}_j$ and landmark $\underline{M}_i$. $\boldsymbol{d}_{i,j,k}$ is the measured range between these two landmarks.

The rotation matrix can be parameterized by quaternions, Euler angles, roll-pitch-yaw, or a rotation vector [8]. In this paper, the parameterization of the rotation matrix is based on the rotation vector [3] according to

$$\begin{aligned} \mathbf{D}(\underline{r}_k) = \mathbf{I} &+ \frac{\sin(\|\underline{r}_k\|)}{\|\underline{r}_k\|} \cdot \mathbf{H}(\underline{r}_k) \\ &+ \frac{1-\cos(\|\underline{r}_k\|)}{\|\underline{r}_k\|^2} \cdot \mathbf{H}(\underline{r}_k) \cdot \mathbf{H}(\underline{r}_k) , \end{aligned}$$

where $\mathbf{I}$ is the identity matrix and $\mathbf{H}$ is a skew-symmetric matrix given by

$$\mathbf{H}(\underline{r}_k) = \begin{bmatrix} 0 & -\boldsymbol{r}_z & \boldsymbol{r}_y \\ \boldsymbol{r}_z & 0 & -\boldsymbol{r}_x \\ -\boldsymbol{r}_y & \boldsymbol{r}_x & 0 \end{bmatrix}$$

for the three-dimensional space.

B. System equation

The system equation describes how the state evolves over time. In state estimation, the system equation is used to predict the state at the next time step when a new measurement is taken. In the considered example, two separate motion models are assumed, one for the translation and the second for the rotation. The discrete-time motion model for the translation is given by a linear equation according to

$$\begin{bmatrix} \underline{t}_{k+1} \\ \dot{\underline{t}}_{k+1} \end{bmatrix} = \mathbf{A} \cdot \begin{bmatrix} \underline{t}_k \\ \dot{\underline{t}}_k \end{bmatrix} + \underline{\omega}_k^t ,$$

where the process noise $\underline{\omega}_k^t$ is assumed to be Gaussian distributed with covariance $\mathbf{Q}^t$ [9]. The matrix $\mathbf{A}$ and the covariance $\mathbf{Q}^t$ are given by

$$\mathbf{A} = \begin{bmatrix} \mathbf{I} & \Delta_t \cdot \mathbf{I} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} , \; \mathbf{Q}^t = \begin{bmatrix} \frac{\Delta_t^3}{3} \mathbf{Q}_c^t & \frac{\Delta_t^2}{2} \mathbf{Q}_c^t \\ \frac{\Delta_t^2}{2} \mathbf{Q}_c^t & \Delta_t \cdot \mathbf{Q}_c^t \end{bmatrix} ,$$

where Δ_t is the sampling time and $\mathbf{Q}_c^t$ is the covariance of the process noise from the continuous-time system model $\mathbf{Q}_c^t = \mathrm{diag}(\begin{bmatrix} Q_{c,x}^t & Q_{c,y}^t & Q_{c,z}^t \end{bmatrix})$.

The evolution of the rotation vector over time is described by a nonlinear equation [10], [11]

$$\begin{aligned} \underline{r}_{k+1} &= \underline{r}_k + \\ &\underbrace{\Delta_t \cdot (\mathbf{I} + 0.5\mathbf{H}(\underline{r}_k) + a(\underline{r}_k) \cdot \mathbf{H}(\underline{r}_k) \cdot \mathbf{H}(\underline{r}_k))}_{\Lambda(\underline{r}_k)} \cdot \underline{w}_k , \end{aligned}$$

with

$$a(\underline{r}_k) = \frac{1 - 0.5||\underline{r}_k||}{||\underline{r}_k||^2} \cdot \cot(\tfrac{||\underline{r}_k||}{2}) .$$

The system model for the angular velocity $\underline{w}_k$ is assumed to be

$$\underline{w}_{k+1} = \underline{w}_k + \underline{\omega}_k^w ,$$

where $\underline{\omega}_k^w$ is the process noise that affects the angular velocity. The process noise has zero mean and is Gaussian distributed with covariance $\mathbf{Q}^w$.

The resulting system model for the pose tracking scenario can be written as

$$\underline{x}_{k+1} = \begin{bmatrix} \mathbf{A} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \Lambda(\underline{r}_k) \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix} \cdot \underline{x}_k + \underline{\omega}_k , \qquad (2)$$

where the covariance of the process noise $\underline{\omega}_k$ contains the covariance of the process noises from the translation model $\mathbf{Q}^t$ and the rotation model $\mathbf{Q}^w$

$$\mathbf{Q} = \begin{bmatrix} \mathbf{Q}^t & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{Q}^w \end{bmatrix} .$$

Furthermore, the rotation vector has to be bounded. If the norm of the rotation vector $||\underline{r}_k|| > \pi$, then the transformed rotation vector is given by $\underline{r}_k^{new} = \underline{r}_k \cdot (1 - \frac{2\pi}{||\underline{r}_k||})$.

For the rest of the paper, the time index k is omitted.

C. Recursive State Estimation

In state estimation, filtering and prediction are performed recursively. The two steps depend on the measurement equation (1) and the system equation (2), respectively. Furthermore, it is assumed that the state $\underline{x}$ can be described by the mean $\underline{\mu}^x$ and the covariance $\mathbf{C}^x$ according to

$$\underline{\mu}^x = \begin{bmatrix} \underline{\mu}^t \\ \underline{\mu}^{\dot{t}} \\ \underline{\mu}^r \\ \underline{\mu}^w \end{bmatrix} , \; \mathbf{C}^x = \begin{bmatrix} \mathbf{C}^t & \mathbf{C}^{t,\dot{t}} & \mathbf{C}^{t,r} & \mathbf{C}^{t,w} \\ \mathbf{C}^{\dot{t},t} & \mathbf{C}^{\dot{t}} & \mathbf{C}^{\dot{t},r} & \mathbf{C}^{\dot{t},w} \\ \mathbf{C}^{r,t} & \mathbf{C}^{r,\dot{t}} & \mathbf{C}^r & \mathbf{C}^{r,w} \\ \mathbf{C}^{w,t} & \mathbf{C}^{w,\dot{t}} & \mathbf{C}^{w,r} & \mathbf{C}^w \end{bmatrix} .$$

1) Prediction Step: In the prediction step, the estimated mean $\underline{\mu}^{x,e}$ and covariance $\mathbf{C}^{x,e}$ of the previous filter step as well as the probabilistic model according to the system equation (2) are used in order to determine the predicted mean $\underline{\mu}^{x,p}$ and covariance $\mathbf{C}^{x,p}$. Due to the fact that the system model is conditionally linear, i.e., if the rotation vector is set to a fixed value, the system model becomes linear and the prediction for each value can be performed by using the Kalman predictor equation. Based on the predicted quantities for each fixed value, the predicted mean and covariance are calculated [3].

2) Filter Step: In the filter step, the state is updated based on the actual measurement $\underline{\hat{z}}$ by using Bayesian inference. Due to the nonlinear measurement equation (1) the filter step cannot be solved in closed form. Hence, due to the assumption that the joint density of the measurement process and the state is Gaussian distributed, the estimated mean $\underline{\mu}^{x,e}$ and covariance $\mathbf{C}^{x,e}$ can be efficiently calculated by using

$$\begin{aligned} \underline{\mu}^{x,e} &= \underline{\mu}^{x,p} + \mathbf{C}^{x,z} \cdot (\mathbf{C}^z)^{-1} (\underline{\hat{z}} - \underline{\mu}^z) \ , \\ \mathbf{C}^{x,e} &= \mathbf{C}^{x,p} - \mathbf{C}^{x,z} \cdot (\mathbf{C}^z)^{-1} \cdot \mathbf{C}^{z,x} \ , \end{aligned} \tag{3}$$

where $\mathbf{C}^{x,z}$ is the cross-covariance between the state and the measurement, $\mathbf{C}^z$ is the covariance, and $\underline{\mu}^z$ the mean of the measurement process $\underline{\boldsymbol{z}}$. These quantities depend on the considered nonlinear function $\underline{h}(\underline{\boldsymbol{x}}, \underline{\boldsymbol{v}})$ and the involved densities for the state $\underline{\boldsymbol{x}}$ and the noise $\underline{\boldsymbol{v}}$ and are given by

$$\begin{aligned} \underline{\mu}^z &= \mathrm{E}_{\underline{\boldsymbol{x}},\underline{\boldsymbol{v}}}\{\underline{h}(\underline{\boldsymbol{x}}, \underline{\boldsymbol{v}})\} \ , \\ \mathbf{C}^z &= \mathrm{E}_{\underline{\boldsymbol{x}},\underline{\boldsymbol{v}}}\{(\underline{h}(\underline{\boldsymbol{x}}, \underline{\boldsymbol{v}}) - \underline{\mu}^z) \cdot (\underline{h}(\underline{\boldsymbol{x}}, \underline{\boldsymbol{v}}) - \underline{\mu}^z)^{\mathrm{T}}\} \ , \\ \mathbf{C}^{x,z} &= \mathrm{E}_{\underline{\boldsymbol{x}},\underline{\boldsymbol{v}}}\{(\underline{\boldsymbol{x}} - \underline{\mu}^{x,p}) \cdot (\underline{h}(\underline{\boldsymbol{x}}, \underline{\boldsymbol{v}}) - \underline{\mu}^z)^{\mathrm{T}}\} \ . \end{aligned} \tag{4}$$

In general, these quantities cannot be calculated in closed form. However, some classes of nonlinear functions lead to analytic expressions [12]. In the following, 1. the measurement equation is modified to a polynomial function and 2. the density of the state is decomposed in order to solve parts of the problem in closed form.

For the rest of the paper, the indices for the predicted and estimated state are omitted.

III. Semi-Analytic Linearization for Solving the Filter Step

A. Modified Measurement Equation

In order to achieve that a part of the filter step can be solved in closed form, the measurement equation (1) is squared, which results in

$$\begin{aligned} (\boldsymbol{d}_{i,j})^2 &= h_{i,j}(\underline{\boldsymbol{x}}, \underline{\boldsymbol{v}}) \\ &= (\underline{g}_{i,j}(\underline{\boldsymbol{r}}) - \underline{\boldsymbol{t}} - \underline{\boldsymbol{v}}_{i,j})^{\mathrm{T}} \cdot (\underline{g}_{i,j}(\underline{\boldsymbol{r}}) - \underline{\boldsymbol{t}} - \underline{\boldsymbol{v}}_{i,j}) \ , \end{aligned} \tag{5}$$

with

$$\underline{g}_{i,j}(\underline{\boldsymbol{r}}) = \underline{L}_j - \mathbf{D}(\underline{\boldsymbol{r}}) \cdot \underline{M}_i \ . \tag{6}$$

Furthermore, it is assumed that squared ranges $\underline{\boldsymbol{z}}$ are measured according to

$$\boldsymbol{z}_{i,j} := (\boldsymbol{d}_{i,j})^2 \ ,$$

i.e., if a measurement $\hat{d}_{i,j}$ is taken, this measurement is mapped to the new measurement $\hat{z}_{i,j}$ by using

$$\hat{z}_{i,j} = (\hat{d}_{i,j})^2 \ . \tag{7}$$

For all measured ranges from $i = 1, \ldots, N$ and $j = 1, \ldots, M$, the nonlinear measurement equation is given by

$$\underline{\boldsymbol{z}} = \underline{h}(\underline{\boldsymbol{t}}, \underline{\boldsymbol{v}}, \underline{g}(\underline{\boldsymbol{r}})) \ ,$$

where N is the number of landmarks with respect to the target frame and M is the number of landmarks with respect to the global coordinate system. $\underline{\boldsymbol{v}}$ is the measurement noise, which has zero mean and is Gaussian distributed with covariance $\mathbf{C}^v$ given by

$$\mathbf{C}^v = \begin{bmatrix} \mathbf{C}_1^v & \ldots & \mathbf{C}_{1,j}^v & \ldots & \mathbf{C}_{1,N\cdot M}^v \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{C}_{i,1}^v & \ldots & \mathbf{C}_{i,j}^v & \ldots & \mathbf{C}_{i,N\cdot M}^v \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{C}_{N\cdot M,1}^v & \ldots & \mathbf{C}_{N\cdot M,j}^v & \ldots & \mathbf{C}_{N\cdot M}^v \end{bmatrix} ,$$

where the submatrix $\mathbf{C}_{i,j}^v$ describes the covariance between the landmarks from the different coordinate systems.

B. Decomposition

In order to decompose the problem into parts that can be solved in closed form, the density of the state $f(\underline{x})$ can be written as

$$f(\underline{x}) = f(\underline{t}, \underline{\dot{t}}, \underline{w} | \underline{r}) \cdot f(\underline{r}) \ . \tag{8}$$

Furthermore, the density for the rotation $f(\underline{r})$ is approximated by a sample-based representation according to

$$f(\underline{r}) \approx \sum_{u=1}^{L} w_u \cdot \delta(\underline{r} - \underline{\mu}_u) \ , \tag{9}$$

where L is the number of sample points, $\underline{\mu}_u$ the sample positions, w_u the sample weights, and $\delta(\cdot)$ the Dirac delta distribution. For determining the sample points $\underline{\mu}_u$, several sampling schemes such as the UKF [4] or [5], [6] can be applied.

Using (9) in (8), for a single sample point $\underline{\mu}_u$ the condition density $f(\underline{x}^a | \underline{r})$ with

$$\underline{x}^a = \begin{bmatrix} \underline{t}^{\mathrm{T}} & \underline{\dot{t}}^{\mathrm{T}} & \underline{w}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}} ,$$

can be written as $f(\underline{x}^a | \underline{\mu}_u) = \mathcal{N}(\underline{x}^a - \underline{\mu}_u^a, \mathbf{C}^a)$ with mean and covariance

$$\begin{aligned} \underline{\mu}_u^a &= \begin{bmatrix} \underline{\mu}_u^t \\ \underline{\mu}_u^{\dot{t}} \\ \underline{\mu}_u^w \end{bmatrix} = \begin{bmatrix} \underline{\mu}^t \\ \underline{\mu}^{\dot{t}} \\ \underline{\mu}^w \end{bmatrix} + \begin{bmatrix} \mathbf{C}^{t,r} \\ \mathbf{C}^{\dot{t},r} \\ \mathbf{C}^{w,r} \end{bmatrix} \cdot (\mathbf{C}^r)^{-1} \cdot (\underline{\mu}_u - \underline{\mu}^r) \ , \\ \mathbf{C}^a &= \begin{bmatrix} \mathbf{C}_t^a & \mathbf{C}_{t,\dot{t}}^a & \mathbf{C}_{t,w}^a \\ \mathbf{C}_{\dot{t},t}^a & \mathbf{C}_{\dot{t}}^a & \mathbf{C}_{\dot{t},w}^a \\ \mathbf{C}_{w,t}^a & \mathbf{C}_{w,\dot{t}}^a & \mathbf{C}_w^a \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{C}^t & \mathbf{C}^{t,\dot{t}} & \mathbf{C}^{t,w} \\ \mathbf{C}^{\dot{t},t} & \mathbf{C}^{\dot{t}} & \mathbf{C}^{\dot{t},w} \\ \mathbf{C}^{w,t} & \mathbf{C}^{w,\dot{t}} & \mathbf{C}^w \end{bmatrix} - \begin{bmatrix} \mathbf{C}^{t,r} \\ \mathbf{C}^{\dot{t},r} \\ \mathbf{C}^{w,r} \end{bmatrix} (\mathbf{C}^r)^{-1} \begin{bmatrix} \mathbf{C}^{t,r} \\ \mathbf{C}^{\dot{t},r} \\ \mathbf{C}^{w,r} \end{bmatrix}^{\mathrm{T}} . \end{aligned} \tag{10}$$

C. Mean

In order to calculate the required mean $\underline{\mu}^z$ in (4), (8) and (9) are used to obtain

$$\begin{aligned}\underline{\mu}^z &= \mathrm{E}_{\underline{\boldsymbol{x}},\underline{\boldsymbol{v}}}\{\underline{h}(\underline{\boldsymbol{t}},\underline{\boldsymbol{v}},\underline{g}(\underline{\boldsymbol{r}}))\} \\ &\approx \sum_{u=1}^{L} w_u \iint \underline{h}(\underline{t},\underline{v},\underline{g}(\underline{r})) f(\underline{x}^a|\underline{\mu}_u)\delta(\underline{r}-\underline{\mu}_u) f(\underline{v})\,\mathrm{d}\underline{v}\,\mathrm{d}\underline{x}\ .\end{aligned}$$

Due to the sifting property of the Dirac delta distribution, the integral for the rotation variable can be solved. Accordingly, the nonlinear function in (5) then only depends on the translation. The variable for the velocity (translational and angular velocity) can be marginalized. The predicted mean is then given by

$$\begin{aligned}\underline{\mu}^z &\approx \sum_{u=1}^{L} w_u \iint \underline{h}(\underline{t},\underline{v},\underline{g}(\underline{\mu}_u)) \cdot f(\underline{x}^a|\underline{\mu}_u) \cdot f(\underline{v})\,\mathrm{d}\underline{v}\,\mathrm{d}\underline{x}^a \\ &= \sum_{u=1}^{L} w_u \iint \underline{h}(\underline{t},\underline{v},\underline{g}(\underline{\mu}_u)) \cdot f(\underline{t}|\underline{\mu}_u) \cdot f(\underline{v})\,\mathrm{d}\underline{v}\,\mathrm{d}\underline{t} \\ &= \sum_{u=1}^{L} w_u \cdot \mathrm{E}_{\underline{\boldsymbol{t}},\underline{\boldsymbol{v}}}\{\underline{h}(\underline{\boldsymbol{t}},\underline{\boldsymbol{v}},\underline{g}(\underline{\mu}_u))\}\ . \qquad (11)\end{aligned}$$

The function $\underline{h}(\cdot)$ contains all combinations between the landmarks in the target and global coordinate system and is given by

$$\begin{aligned}\underline{h}(\underline{\boldsymbol{t}},\underline{\boldsymbol{v}},\underline{g}(\underline{\mu}_u)) = \mathbf{K}^{\mathrm{T}} \cdot ((&\underline{g}(\underline{\mu}_u) - \underline{1}_{N\cdot M} \otimes \underline{\boldsymbol{t}} - \underline{\boldsymbol{v}}) \circ \\ &(\underline{g}(\underline{\mu}_u) - \underline{1}_{N\cdot M} \otimes \underline{\boldsymbol{t}} - \underline{\boldsymbol{v}}))\ ,\end{aligned}$$

where $\otimes$ is the Kronecker product, $\circ$ the element-wise product, $\underline{1}$ is the one-vector, D is the dimension of $\underline{\boldsymbol{t}}$, and $\mathbf{K} = \mathbf{I}_{N\cdot M} \otimes \underline{1}_D$. The vector $\underline{g}(\underline{\mu}_u)$, which depends on the sample points $\underline{\mu}_u$, consists of the entries from (6) and is given by

$$\underline{g}(\underline{\mu}_u) = \begin{bmatrix} \underline{g}_{1,1}(\underline{\mu}_u)^{\mathrm{T}} & \cdots & \underline{g}_{N,1}(\underline{\mu}_u)^{\mathrm{T}} & \cdots & \underline{g}_{N,M}(\underline{\mu}_u)^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}} .$$

For calculating the mean $\underline{\mu}^z$, a new variable $\underline{\boldsymbol{f}}_u$ is introduced, which is defined by

$$\underline{\boldsymbol{f}}_u := \underline{g}(\underline{\mu}_u) - \underline{1}_{N\cdot M} \otimes \underline{\boldsymbol{t}} - \underline{\boldsymbol{v}}\ .$$

The density of the variable $\underline{\boldsymbol{f}}_u$ is Gaussian distributed due to the linear relation. The mean and covariance are given by

$$\begin{aligned}\underline{\mu}_u^f &= \underline{g}(\underline{\mu}_u) - \underline{1}_{N\cdot M} \otimes \underline{\mu}_u^t\ , \qquad (12) \\ \mathbf{C}^f &= \mathbf{C}^v + (\underline{1}_{N\cdot M} \cdot (\underline{1}_{N\cdot M})^{\mathrm{T}}) \otimes \mathbf{C}_t^a\ ,\end{aligned}$$

where the quantities $\underline{\mu}_u^t$ and $\mathbf{C}_t^a$ are stemmed from $\underline{\mu}_u^a$ and $\mathbf{C}^a$ with respect to the translation (10). Based on this new random variable, the expected value (11) for a fixed value for $\underline{\mu}_u$ is calculated as

$$\begin{aligned}\mathrm{E}_{\underline{\boldsymbol{t}},\underline{\boldsymbol{v}}}\{\underline{h}(\underline{\boldsymbol{t}},\underline{\boldsymbol{v}},\underline{g}(\underline{\mu}_u))\} &= \mathrm{E}_{\underline{\boldsymbol{t}},\underline{\boldsymbol{v}}}\{\underline{h}(\cdot)\} = \mathrm{E}_{\underline{\boldsymbol{f}}_u}\{\mathbf{K}^{\mathrm{T}} \cdot (\underline{\boldsymbol{f}}_u \circ \underline{\boldsymbol{f}}_u)\} \qquad (13) \\ &= \mathbf{K}^{\mathrm{T}} \cdot (\underline{\mu}_u^f \circ \underline{\mu}_u^f + \mathrm{diag}(\mathbf{C}^f))\ .\end{aligned}$$

In order to determine the mean $\underline{\mu}^z$, the expected value for every fixed value $\underline{\mu}_u$ $u = 1, \ldots, L$ has to be calculated. The mean $\underline{\mu}^z$ is then given by

$$\underline{\mu}^z = \sum_{u=1}^{L} w_u \cdot \mathbf{K}^{\mathrm{T}} \cdot (\underline{\mu}_u^f \circ \underline{\mu}_u^f + \mathrm{diag}(\mathbf{C}^f))\ . \qquad (14)$$

D. Covariance

Similar to the calculation of the mean, the covariance is calculated according to

$$\begin{aligned}\mathbf{C}^z = \sum_{u=1}^{L} w_u \cdot \mathrm{E}_{\underline{\boldsymbol{t}},\underline{\boldsymbol{v}}}\{&(\underline{h}(\underline{\boldsymbol{t}},\underline{\boldsymbol{v}},\underline{g}(\underline{\mu}_u)) - \underline{\mu}^z) \\ &\cdot (\underline{h}(\underline{\boldsymbol{t}},\underline{\boldsymbol{v}},\underline{g}(\underline{\mu}_u)) - \underline{\mu}^z)^{\mathrm{T}}\}\ . \qquad (15)\end{aligned}$$

Due to the linearity of the expectation operator, each summand of the product

$$\mathrm{E}_{\underline{\boldsymbol{t}},\underline{\boldsymbol{v}}}\{(\underline{h}(\cdot) - \underline{\mu}^z) \cdot (\underline{h}(\cdot) - \underline{\mu}^z)^{\mathrm{T}}\}$$

can be calculated separately. The product $\mathrm{E}_{\underline{\boldsymbol{t}},\underline{\boldsymbol{v}}}\{\underline{h}(\cdot) \cdot \underline{h}(\cdot)^{\mathrm{T}}\}$ for the variable $\underline{\boldsymbol{f}}_u$ is given by

$$\begin{aligned}&\mathrm{E}_{\underline{\boldsymbol{t}},\underline{\boldsymbol{v}}}\{\underline{h}(\cdot) \cdot \underline{h}(\cdot)^{\mathrm{T}}\} \\ &\quad = \mathrm{E}_{\underline{\boldsymbol{f}}_u}\{\mathbf{K}^{\mathrm{T}} \cdot (\underline{\boldsymbol{f}}_u \circ \underline{\boldsymbol{f}}_u) \cdot (\mathbf{K}^{\mathrm{T}} \cdot (\underline{\boldsymbol{f}}_u \circ \underline{\boldsymbol{f}}_u))^{\mathrm{T}}\}\ . \qquad (16)\end{aligned}$$

The matrix $\mathbf{K}$ can be pulled out of the integral, which results in

$$\mathrm{E}_{\underline{\boldsymbol{f}}_u}\{\underline{h}(\cdot) \cdot \underline{h}(\cdot)^{\mathrm{T}}\} = \mathbf{K}^{\mathrm{T}} \cdot \mathrm{E}_{\underline{\boldsymbol{f}}_u}\{(\underline{\boldsymbol{f}}_u \circ \underline{\boldsymbol{f}}_u) \cdot (\underline{\boldsymbol{f}}_u^{\mathrm{T}} \circ \underline{\boldsymbol{f}}_u^{\mathrm{T}})\} \cdot \mathbf{K}\ .$$

Furthermore, the Cartesian product of the squared variable $\underline{\boldsymbol{f}}_u$ is written as the squared Cartesian product of the variable $\underline{\boldsymbol{f}}_u$, which is given by

$$\mathrm{E}_{\underline{\boldsymbol{f}}_u}\{\underline{h}(\cdot) \cdot \underline{h}(\cdot)^{\mathrm{T}}\} = \mathbf{K}^{\mathrm{T}} \cdot \mathrm{E}_{\underline{\boldsymbol{f}}_u}\{(\underline{\boldsymbol{f}}_u \cdot \underline{\boldsymbol{f}}_u^{\mathrm{T}}) \circ (\underline{\boldsymbol{f}}_u \cdot \underline{\boldsymbol{f}}_u^{\mathrm{T}})\} \cdot \mathbf{K}\ .$$

The expected value of the matrix $\mathrm{E}_{\underline{\boldsymbol{f}}_u}\{(\underline{\boldsymbol{f}}_u \cdot \underline{\boldsymbol{f}}_u^{\mathrm{T}}) \circ (\underline{\boldsymbol{f}}_u \cdot \underline{\boldsymbol{f}}_u^{\mathrm{T}})\}$ can be calculated separately. Each entry of the matrix corresponds to $\mathrm{E}_{\underline{\boldsymbol{f}}_u}\{(\boldsymbol{f}_{u,i})^2 \cdot (\boldsymbol{f}_{u,j})^2\}$, which is a fourth-order non-central moment according to

$$\begin{aligned}&\mathrm{E}_{\boldsymbol{f}_{u,i},\boldsymbol{f}_{u,j}}\{(\boldsymbol{f}_{u,i})^2 \cdot (\boldsymbol{f}_{u,j})^2\} = \\ &\quad \left(\left(\mu_{u,i}^f\right)^2 + C_{i,i}^f\right) \cdot \left(\left(\mu_{u,j}^f\right)^2 + C_{j,j}^f\right) \\ &\quad + 4 \cdot \mu_{u,i}^f \cdot \mu_{u,j}^f \cdot C_{i,j}^f + 2 \cdot \left(C_{i,j}^f\right)^2\ ,\end{aligned}$$

where $\mu_{u,i}^f$ is the ith entry of the vector $\underline{\mu}_u^f$ and $C_{i,j}^f$ is the entry (i,j) of the matrix $\mathbf{C}^f$ for $i = 1, \ldots, N \cdot M \cdot D$ and $j = 1, \ldots, N \cdot M \cdot D$. According to this, the expected value of the matrix in (16) is given by

$$\begin{aligned}&\mathrm{E}_{\underline{\boldsymbol{t}},\underline{\boldsymbol{v}}}\{\underline{h}(\cdot) \cdot \underline{h}(\cdot)^{\mathrm{T}}\} = \mathrm{E}_{\underline{\boldsymbol{t}},\underline{\boldsymbol{v}}}\{\underline{h}(\cdot)\} \cdot \mathrm{E}_{\underline{\boldsymbol{t}},\underline{\boldsymbol{v}}}\{\underline{h}(\cdot)\}^{\mathrm{T}} + \qquad (17) \\ &\quad \mathbf{K}^{\mathrm{T}} \cdot \left(4 \cdot \left(\underline{\mu}_u^f \cdot \left(\underline{\mu}_u^f\right)^{\mathrm{T}}\right) \circ \mathbf{C}^f + 2 \cdot \mathbf{C}^f \circ \mathbf{C}^f\right) \cdot \mathbf{K}\ ,\end{aligned}$$

where the expected value for $\mathrm{E}_{\underline{\boldsymbol{t}},\underline{\boldsymbol{v}}}\{\underline{h}(\cdot)\}$ is given in (13).

E. Cross-Covariance

The cross-covariance is separated into the sub-cross-covariance matrices according to

$$\mathbf{C}^{x,z} = \tag{18}$$
$$\sum_{u=1}^{L} w_u \begin{bmatrix} \mathrm{E}_{\underline{t},\underline{v}}\{(\underline{t} - \underline{\mu}^t) \cdot (\underline{h}(\underline{t}, \underline{v}, \underline{g}(\underline{\mu}_u)) - \underline{\mu}^z)^\mathrm{T}\} \\ \mathrm{E}_{\underline{t},\underline{\dot{t}},\underline{v}}\{(\underline{\dot{t}} - \underline{\mu}^{\dot{t}}) \cdot (\underline{h}(\underline{t}, \underline{v}, \underline{g}(\underline{\mu}_u)) - \underline{\mu}^z)^\mathrm{T}\} \\ \mathrm{E}_{\underline{t},\underline{v}}\{(\underline{\mu}_u - \underline{\mu}^r) \cdot (\underline{h}(\underline{t}, \underline{v}, \underline{g}(\underline{\mu}_u)) - \underline{\mu}^z)^\mathrm{T}\} \\ \mathrm{E}_{\underline{t},\underline{w},\underline{v}}\{(\underline{w} - \underline{\mu}^w) \cdot (\underline{h}(\underline{t}, \underline{v}, \underline{g}(\underline{\mu}_u)) - \underline{\mu}^z)^\mathrm{T}\} \end{bmatrix} .$$

If each entry of the product is expanded, the remaining unknown expectation values are given by

$$\mathrm{E}_{\underline{t},\underline{v}}\{\underline{t} \cdot \underline{h}(\cdot)^\mathrm{T}\} = -2\mathbf{C}_t^a \cdot \mathbf{S}_u + \underline{\mu}_u^t \cdot \mathrm{E}_{\underline{t},\underline{v}}\{\underline{h}(\cdot)\}^\mathrm{T} , \tag{19}$$
$$\mathrm{E}_{\underline{t},\underline{\dot{t}},\underline{v}}\{\underline{\dot{t}} \cdot \underline{h}(\cdot)^\mathrm{T}\} = -2\mathbf{C}_{\dot{t},t}^a \cdot \mathbf{S}_u + \underline{\mu}_u^{\dot{t}} \cdot \mathrm{E}_{\underline{t},\underline{v}}\{\underline{h}(\cdot)\}^\mathrm{T} ,$$
$$\mathrm{E}_{\underline{t},\underline{w},\underline{v}}\{\underline{w} \cdot \underline{h}(\cdot)^\mathrm{T}\} = -2\mathbf{C}_{w,t}^a \cdot \mathbf{S}_u + \underline{\mu}_u^w \cdot \mathrm{E}_{\underline{t},\underline{v}}\{\underline{h}(\cdot)\}^\mathrm{T} ,$$

with

$$\mathbf{S}_u = \begin{bmatrix} \underline{g}_{1,1}(\underline{\mu}_u) & \cdots & \underline{g}_{N,1}(\underline{\mu}_u) \cdots & \underline{g}_{N,M}(\underline{\mu}_u) \end{bmatrix} - \underline{1}_{N \cdot M}^\mathrm{T} \otimes \underline{\mu}_u^t .$$

The quantities $\underline{\mu}_u^{\dot{t}}$, $\mathbf{C}_{\dot{t},t}^a$, $\underline{\mu}_u^w$, and $\mathbf{C}_{w,t}^a$ are the entries of $\underline{\mu}^a$ and $\mathbf{C}^a$ from (10) with respect to the translation and angular velocity, respectively.

F. Wrap-Up

In the following, a short wrap-up for the filter is provided:

1) Determine the sample points $\underline{\mu}_u$ with weights w_u of the density of the rotation $f(\underline{r})$, where $u = 1, \ldots, L$.
2) Calculate the conditional means $\underline{\mu}_u^t$, $\underline{\mu}_u^{\dot{t}}$, and $\underline{\mu}_u^w$, the conditional covariance $\mathbf{C}_t^a$, and conditional cross-covariances $\mathbf{C}_{\dot{t},t}^a$ and $\mathbf{C}_{w,t}^a$ for all sample points (see (10)).
3) Calculate the mean $\underline{\mu}_u^f$ and covariance $\mathbf{C}^f$ of the random variable $\underline{\boldsymbol{f}}_u$ for all sample points (see (12)).
4) Calculate the expected values $\mathrm{E}_{\underline{t},\underline{v}}\{\underline{h}(\cdot)\}$ (see (13)), $\mathrm{E}_{\underline{t},\underline{v}}\{\underline{h}(\cdot) \cdot \underline{h}(\cdot)^\mathrm{T}\}$ (see (17)), $\mathrm{E}_{\underline{t},\underline{v}}\{\underline{t} \cdot \underline{h}(\cdot)^\mathrm{T}\}$, $\mathrm{E}_{\underline{t},\underline{\dot{t}},\underline{v}}\{\underline{\dot{t}} \cdot \underline{h}(\cdot)^\mathrm{T}\}$, and $\mathrm{E}_{\underline{t},\underline{w},\underline{v}}\{\underline{w} \cdot \underline{h}(\cdot)^\mathrm{T}\}$ (see (19)) for all sample points.
5) Combine the results of the expected values for determining the mean $\underline{\mu}^z$ (see (14)), the covariance $\mathbf{C}^z$ (see (15)), and the cross-covariance $\mathbf{C}^{x,z}$ (see (18)).
6) Square the measured ranges $\underline{\hat{d}}$ (see (7)).
7) Perform the filter step in order to calculate the estimated mean and covariance (see (3)).

G. Computational Complexity

The determination of the sample points has the highest computational cost for the calculating the required mean $\underline{\mu}^z$, covariance $\mathbf{C}^z$, and cross-covariance $\mathbf{C}^{x,z}$. For the proposed approach, the sample points only have to be calculated for the rotation. In this case, the computational complexity is in $\mathrm{O}(R^3)$, where R is the dimension of the rotation vector. If the decomposition lies only in separation according to directly and indirectly observed parts, which is explained in [3], the computational complexity is in $\mathrm{O}\left((R + D + N \cdot M)^3\right)$, because the translation and the noise is in the nonlinearity, which then has to be sampled, too.

IV. Simulation Results

In the simulation, a two-dimensional coordinate system is considered. Four emitters are located at the positions

$$\begin{bmatrix} -2 & -2 \\ -2 & 2 \\ 2 & -2 \\ 2 & 2 \end{bmatrix} \mathrm{m}$$

with respect to the global coordinate system. Furthermore, four sensors are placed on the target frame

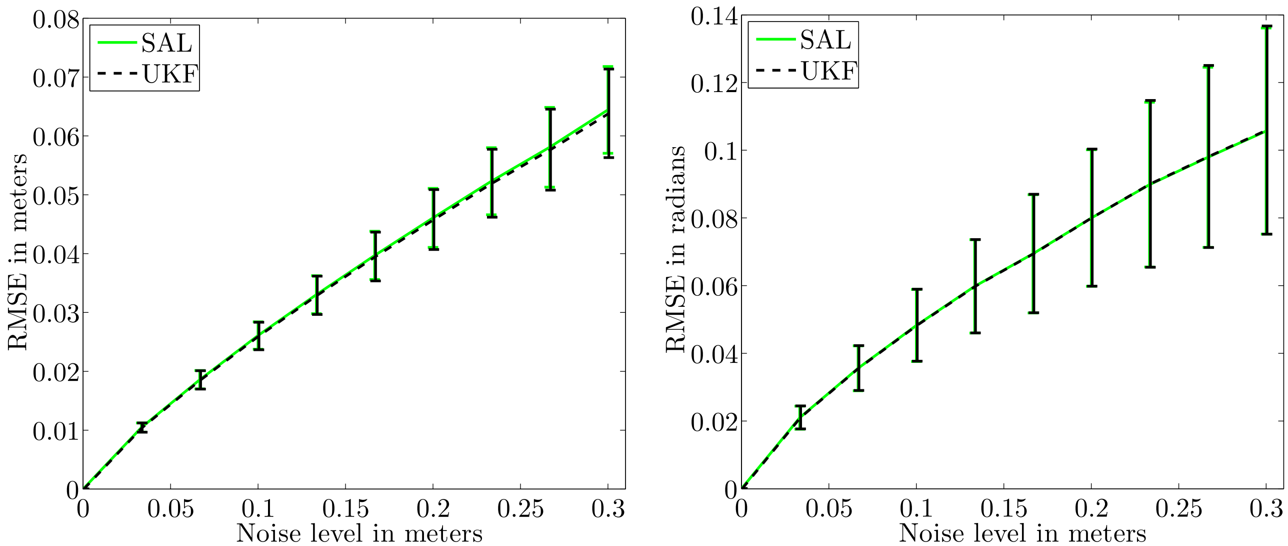
$$\begin{bmatrix} -0.2 & -0.2 \\ -0.2 & 0.2 \\ 0.2 & -0.2 \\ 0.2 & 0.2 \end{bmatrix} \mathrm{m} .$$

At different noise levels ranging from $[0.000001, \ldots, 0.3]$ meters, 1000 random trajectories are generated, where the sampling time was 0.1 seconds. The noise process is assumed as isotropic. The measured ranges were generated with (1).

In the simulation, the proposed approach (SAL) is compared to a standard estimator. As a standard estimator, the Unscented Kalman Filter (UKF) was used, where the decomposition lies in separation according to directly and indirectly observed parts, which is explained in [3]. In this case, the density for the translation and the rotation has to be approximated with samples. Furthermore, due to the fact that the measurement noise is mapped through the nonlinear transformation, it has to be approximated with samples, too. For the UKF, 71 sample points are used to approximate the density of the translation, rotation, and the measurement noise. On the other hand, the proposed approach only has to approximate the rotation by sample points, where for determining the sample points the approach presented in [6] is used. Here, 5 sample points are used to approximate the density for the rotation.

Furthermore, the system equation in (2) for a two-dimensional coordinate system becomes linear and so the prediction step can be solved by using the Kalman prediction step. This is exploited for both estimators. The covariance of the continuous process noise for the velocities (translation and angle) is set to $\mathbf{Q}_c^t = \mathrm{diag}(\begin{bmatrix} 0.1 & 0.1 \end{bmatrix})$ and $Q_c^w = 0.1$, respectively. The initial covariance for the translation is $\mathbf{C}_0^t = \mathrm{diag}(\begin{bmatrix} 10 & 10 \end{bmatrix})$, for the translation velocity $\mathbf{C}_0^{\dot{t}} = \mathrm{diag}(\begin{bmatrix} 10 & 10 \end{bmatrix})$, for the angle $C_0^r = 0.001$, and for the angular velocity $C_0^w = 0.0001$. The initial mean is initialized with zeros.

In Fig. 2, the average and the standard deviation of the Root Mean Square Error (RMSE) over 1000 trajectories for each noise level are plotted. The performance of the two estimators is nearly equal. Regarding the computational effort, the proposed approach only has to determine sample points for one dimension, which can be implemented very efficiently. On the other hand, the UKF calculates a matrix root for the covariance of the noise and the state (translation and rotation), which has a high computational effort considering the size of the combined covariance matrix, which is



(a) Simulation results for the RMSE of the translation.

(b) Simulation results for the RMSE of the rotation.

Fig. 2. Simulation results for the two estimators SAL (Semi-Analytic Linearization) and UKF. The average and the standard deviation of the RMSE for different noise level.

35×35. In the simulation, the proposed approach is three times faster than the standard approach.

V. Conclusions

In this paper, a state estimator for range-based pose tracking is presented, which relies on an intelligent decomposition of the proposed problem. This state estimator exploits the Gaussian assumption and makes use of a modified measurement equation, where the measurement noise process is inside of the nonlinearity. Due to the modified measurement equation, the filter step is separated into an analytically integrable and an approximative part. In the approximative part, the density for the rotation is represented by samples. For every sample point, the required moments for the filter step are then calculated by analytic moment calculation. In doing so, the computational demand for the approximation is drastically reduced compared to a standard decomposition, which relies on conditionally linear substructures.

The new approach was evaluated in a two-dimensional simulation and compared to the standard approach. Regarding the RMSE, the performance of the two estimators is similar. However, in the two-dimensional simulation example, only a one-dimensional density of the rotation has to be approximated for the proposed approach, which is feasible for an embedded system, compared to the standard estimator, where the matrix root of a large covariance matrix has to be calculated. In summary, if the decomposition in integrable substructure is exploited, the number of sample points can be drastically reduced.

References

[1] P. Rößler, F. Beutler, U. D. Hanebeck, and N. Nitzsche, "Motion Compression Applied to Guidance of a Mobile Teleoperator," in *Proceedings of the 2005 IEEE International Conference on Intelligent Robots and Systems (IROS 2005)*, Edmonton, Canada, Aug. 2005, pp. 2495–2500.

[2] F. Beutler and U. D. Hanebeck, "Closed-Form Range-Based Posture Estimation Based on Decoupling Translation and Orientation," in *Proceedings of the 2005 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2005)*, vol. 4, Philadelphia, Pennsylvania, Mar. 2005, pp. 989–992.

[3] F. Beutler, M. F. Huber, and U. D. Hanebeck, "Gaussian Filtering using State Decomposition Methods," in *Proceedings of the 12th International Conference on Information Fusion (Fusion 2009)*, Seattle, Washington, July 2009.

[4] S. J. Julier and J. K. Uhlmann, "Unscented Filtering and Nonlinear Estimation," *Proceedings of the IEEE*, vol. 92, no. 3, pp. 401–422, 2004.

[5] K. Ito and K. Xiong, "Gaussian Filters for Nonlinear Filtering Problems," *IEEE Transactions on Automatic Control*, vol. 45, no. 5, pp. 910–927, May 2000.

[6] M. F. Huber and U. D. Hanebeck, "Gaussian Filter based on Deterministic Sampling for High Quality Nonlinear Estimation," in *Proceedings of the 17th IFAC World Congress (IFAC 2008)*, vol. 17, no. 2, Seoul, Korea, July 2008.

[7] M. Morelande and B. Moran, "An Unscented Transformation for Conditionally Linear Models," *IEEE International Conference on Acoustics, Speech and Signal Processing, 2007 (ICASSP 2007)*, vol. 3, pp. 1417–1420, April 2007.

[8] N. Ayache, *Artifical Vision for Mobile Robots: Stereo Vision and Multisensory Perception*. Massachusetts Institute of Technology, 1991.

[9] G. Welch, B. D. Allen, A. Ilie, and G. Bishop, "Measurement Sample Time Optimization for Human Motion Tracking/Capture Systems," in *Proceedings of Trends and Issues in Tracking for Virtual Environments, Workshop at the IEEE Virtual Reality 2007 Conference*, 2007.

[10] J. E. Bortz, "A New Mathematical Formulation for Strapdown Inertial Navigation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-7, no. 1, pp. 61–66, Januar 1971.

[11] F. Beutler, M. F. Huber, and U. D. Hanebeck, "Instantaneous Pose Estimation using Rotation Vectors," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2009) in Taipei, Taiwan*, April 2009, pp. 3413–3416.

[12] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes*. McGraw-Hill Science/Engineering/Math; 4. Edition, 2002.