Sabine Horvath* and Hans Neuner

System identification of a robot arm with extended Kalman filter and artificial neural networks

https://doi.org/10.1515/jag-2018-0045
Received November 14, 2018; accepted January 16, 2019

Abstract: The development of an algorithm to describe a dynamic system and to predict its future behaviour in further consequence is the aim of the present study. Non parametric models provide a general description of object dynamics and artificial neural networks (ANN), which are a very flexible and universal learning method, belong to it. However, the standard estimation procedures for ANN like Levenberg-Marquardt (LM) do not consider that data is observed and consequently is uncertain. The combination with the extended Kalman filter (EKF) enables the consideration of the uncertainty in the estimation process. The analogies between EKF and LM are discussed and thereon the advantages of the EKF are outlined.

The integration of ANN into EKF will be evaluated on an industrial robot arm. At first, a simplified model is determined; the ANN describes the robot position deviations as a function of the joint encoder values. The robot reference positions are measured by a laser tracker. In order to compare it with the robot outputs, the observations need to be transformed to the robot frame and the offset between the end-effector and the robot flange has to be determined. A method to estimate both parameters simultaneously is developed and the results are verified on basis of simulated data.

This paper comprises two novel approaches. First, uncertainty is considered in the ANN estimation on basis of the combination with the EKF. Considering the full covariance matrix of the robot deviations leads to a better prediction of the robot's behaviour. Second, an integrated transformation and lever arm determination is introduced and the robot's repeatability presents the limiting factor of the achievable parameter uncertainty.

Keywords: ANN, EKF, Robot Arm, System Identification, Laser Tracker

***Corresponding author: Sabine Horvath,** Department of Geodesy and Geoinformation, Research Division Engineering Geodesy, TU Wien, Gusshausstr. 27-29 / E120-5, 1040 Vienna, Austria, e-mail: sabine.horvath@tuwien.ac.at
Hans Neuner, Department of Geodesy and Geoinformation, Research Division Engineering Geodesy, TU Wien, Gusshausstr. 27-29 / E120-5, 1040 Vienna, Austria, e-mail: hans.neuner@tuwien.ac.at

1 Introduction

If physical or mathematical-statistical models do not succeed to describe a complex system appropriately or in a satisfactory manner, a behavioural approach is preferable. Estimating model parameters of an unknown input-output relation from observed data, considering available a priori information is the aim of learning. Once such dependencies are estimated, future system outputs can be predicted on basis of given input data. Artificial Neural Networks (ANN) are a very flexible type of approximating functions. It is one preferred method in many different research fields to accomplish regression as well as classification tasks. In Geodesy, ANN have been applied for the non-linear propagation of quality [13] and the step length determination of a dead reckoning navigation on basis of a combination with fuzzy logic [20]. Moreover, it has been used for analysing inclination measurements [1], for the system identification of a lock [15] and for modelling the impact of weather on a bridge [7]. These three publications among others have in common that spatial information in form of coordinates, determined from measurements build the output of the ANN. In this contribution, the output of the ANN bases as well on measurements. However, in the standard ANN no uncertainties of input nor of output quantities are considered. This additional information can prevent the model to correspond too closely to the data and already represent measurement noise. The estimation of the ANN in the extended Kalman filter (EKF) enables the integration of stochastic information of the output. Another advantage of the EKF is the faster convergence in comparison to standard optimisation methods [6]. In this paper the combination of ANN and EKF is compared to the standard for ANN parameter estimation, the Levenberg-Marquardt (LM) method [14]. Their equivalence under specific assumptions is shown. The influence of the stochastic information on the system identification is evaluated and heuristics of standard optimisation methods are replaced by statistical measures.

The system to be identified is an industrial robot arm. The end-effector position of the robot arm is measured by a laser tracker and the deviation to the robot's position is used as an output in the ANN. In order to produce meaningful absolute robot position deviations, the determination of an optimal transformation between the laser

Brought to you by | TU Wien Bibliothek
Authenticated
Download Date | 3/26/19 2:46 PM

tracker and the robot arm frame as well as of the offset between the end-effector and the robot flange is necessary. To keep complexity of the ANN low and due to the reason that the zero positions of the encoders are the main error source of absolute position errors [2], the ANN input is only based on the encoder values. Thus, the ANN describes the dependency of the robot position deviations to the joint encoder values.

Classic parametric calibration methods are treated in [21]. Laser tracker measurements for calibrating an ABB robot are used in [17]. In addition, [23] determines the full robot pose by laser tracker measurements. In Geodesy, the positioning and synchronization of industrial robots has been discussed in [8]. [4] calibrates a metrological arm with a length gage and empirically figures out the optimal measurement configuration.

ANN are also a known method in robotics. They are often used for solving the inverse kinematics of a robot arm e. g. [24]. In [16] EKF and ANN are integrated to calibrate a robot arm. The geometric robot parameters are first modelled by an EKF approach and subsequently an ANN is applied on the position residuals to compensate non geometric error sources that are not captured by the EKF. Different to [16], in this contribution ANN and EKF build one model; EKF is mainly serving as an optimisation method.

The system identification of a robot arm is one possible application for the developed algorithm and has been chosen due to the robot's well understood behaviour. Besides the use as a correction tool, it can also serve for robot arm control. A typical geodetic application is to monitor and model the deformation of an object – e. g. landslide, dam – in order to warn in case of hazard. The non parametric modelling and universal approximation by ANNs enable the consideration of various influencing measures and therefore an appropriate prediction can be expected.

The two main aspects of the present study are to evaluate the potential of the chosen learning method in comparison to standard approaches and the derivation of the robot arm position deviations. The article is structured as follows: After a theoretical introduction in ANN and EKF in Section 2, the measurements and the preprocessing steps are described in Section 3. The transformation results are verified on basis of a simulation. Thereby, possible extensions of the functional model are discussed to reach sufficiently accurate transformation parameters as well as to ensure the separability of these parameters. In Section 4, the simplified system robot arm is identified. This includes the model selection task, the empirical confirmation of the equivalence of LM and EKF and the evaluation of the influence of stochastic information on the combined approach. The results are analysed and a final conclusion is drawn in Section 5.

2 Artificial neural networks and extended Kalman filter

Learning belongs to the class of data-driven methods. It consists of two steps: First, the parameters of a chosen model are estimated based on labelled pairs of input and output data in a training step. Second, the trained model is applied on new input data to test the prediction performance. Consequently, an empirical model is generated. This conforms to a general model in case of infinite training data sets. The higher the complexity of the relation between input and output measures, the more data is needed. The amount of available training data determines the generalisation capability and the reachable predictive accuracy of the trained model. A learning approach consists of an approximating function and an optimisation method [3]. The performance of ANN as an approximating function strongly depends on the chosen optimisation method. Therefore, this contribution is focused on the optimisation and the possibilities offered by the chosen method.

2.1 ANN

Artificial neural networks are a universal and flexible deployable approximating function. A brief description of the functionality of ANNs is given. For an exhaustive introduction see [9]. The mathematical notation of an ANN with one processing layer m is shown in equation (1). The function f_n includes K free weights $\boldsymbol{w} = [w_k]$, which need to be estimated. On basis of inputs $x_l(t)$ with $l = 1, \ldots, L$ and initial values of the weights $\boldsymbol{w}_0 = [\boldsymbol{w}_{ml,0}; \boldsymbol{w}_{nm,0}]$ the output $\hat{y}_n(t)$ with $n = 1, \ldots, N$ is computed. The activation functions φ are the basis functions of neural networks (see [5]). t is the specific index for the training data sample. The computed output is compared to the observed one $y_n(t)$ and the error $e_n(t)$ is back propagated through the network to update the weights $\hat{\boldsymbol{w}}_i$ per iteration i in a way that the total error ε_i (2) is minimised.

$$\begin{aligned} e_{n,i}(t) &= y_n(t) - \hat{y}_{n,i}(t) = y_n(t) - f_n(\boldsymbol{w}_i, \boldsymbol{x}(t)) \\ &= y_n(t) - \varphi^{(n)}\left(\sum_m w_{nm,i}\varphi^{(m)}\left(\sum_l w_{ml,i}x_l(t)\right)\right) \end{aligned} \tag{1}$$

$$\varepsilon_i(t) = \frac{1}{2}\sum_n e_{n,i}^2(t) \tag{2}$$

Brought to you by | TU Wien Bibliothek
Authenticated
Download Date | 3/26/19 2:46 PM

The back propagation corresponds to a gradient descent estimation method. Such linear methods are advantageous due to their permanent progress towards a local minimum, but offer a slow convergence behaviour. Quasi Newton optimisation methods approximate the second derivatives of the Hessian matrix and hence lead to faster convergence close to a minimum. The Levenberg-Marquardt (LM) optimisation combines the advantages of linear estimation methods and Quasi-Newton optimisation methods [14]. The first derivatives of the ANN (1) are contained in the Jacobian $\boldsymbol{J}$ (3). By extending the quadratic approximation term $\boldsymbol{J}^T\boldsymbol{J}$ with a term $\lambda\boldsymbol{I}$ the advantages of both methods are usable. The learning rate λ enables a change between the gradient descent (large λ) and the Newton (small λ) method. The optimisation of ANNs by LM or any other standard second order method is usually applied in a batch mode, i. e. few to all sets of labelled input-output data (t_b) cause an updating of the weights.

$$\boldsymbol{w}_{i+1} = \boldsymbol{w}_i - \left(\boldsymbol{J}^T\boldsymbol{J} + \lambda\boldsymbol{I}\right)^{-1}\boldsymbol{J}^T\boldsymbol{e} \tag{3}$$

2.2 EKF

The Kalman filter is a linear recursive data processing algorithm, which combines a physical problem description with observations related to some parameters. If no linear relation is given, the function can be linearised and needs to be estimated iteratively, the so called extended Kalman filter. The methodological description of the combination of EKF and ANN is given by [22]. The approach bases on a static system equation (4). The ANN is considered in the observation equation (5), which relates the ANN's function $\boldsymbol{f}$ to the observation noise $\boldsymbol{o}(t_b)$ and the observations $\boldsymbol{y}(t_b)$. The first corresponds to the ANNs output $\hat{\boldsymbol{y}}(t_b)$ as shown in Section 2.1. The process noise $\boldsymbol{p}$ and the observation noise $\boldsymbol{o}$ are assumed to be Gaussian with zero mean and the appropriate covariance $\boldsymbol{\Sigma}$. A transition to cofactor matrices $\boldsymbol{Q}$ is fulfilled on basis of the common variance factor of unit weight σ_0^2 ($\boldsymbol{\Sigma} = \sigma_0^2\boldsymbol{Q}$).

$$\bar{\boldsymbol{w}}_{i+1} = \boldsymbol{I}\hat{\boldsymbol{w}}_i + \boldsymbol{p}, \quad \boldsymbol{p} \sim \mathcal{N}(0, \boldsymbol{\Sigma}_{pp}) \tag{4}$$

$$\boldsymbol{y}(t_b) - \boldsymbol{o}(t_b) = \boldsymbol{f}(\hat{\boldsymbol{w}}_{i+1}, \boldsymbol{X}(t_b)), \quad \boldsymbol{o} \sim \mathcal{N}(0, \boldsymbol{\Sigma}_{yy}) \tag{5}$$

The ANN expressed in a linear form is: $\boldsymbol{f}(\bar{\boldsymbol{w}}_{i+1}, \boldsymbol{X}(t_b)) = \boldsymbol{A}_{i+1}\bar{\boldsymbol{w}}_{i+1}$. The Jacobian matrix $\boldsymbol{A}_{i+1}$ describes the linear relation between the ANN's output and the weights.

Due to the static system the predicted weights correspond to the preliminary estimated one (4). However, the cofactor matrix of the predicted weights $\boldsymbol{Q}_{\bar{w}\bar{w},i+1}$ needs to be computed because of the consideration of the process cofactor matrix $\boldsymbol{Q}_{pp}$ (6).

The updating of the weight estimate is accomplished on basis of the predicted weights $\bar{\boldsymbol{w}}_{i+1}$ and by $\boldsymbol{K}_{i+1}$ weighted innovations $\boldsymbol{e}_{i+1}$ for each batch t_b, see equation (7). The innovations are computed on basis of the non-linear function $\boldsymbol{f}$ of the ANN and corresponds to the error in the ANN estimation (1). The Kalman gain $\boldsymbol{K}_{i+1}$ given in equation (8) describes the influence of the observations on the weight estimation. If the elements of $\boldsymbol{Q}_{yy}$ are small, i. e. the observations are precise, the Kalman gain $\boldsymbol{K}_{i+1}$ becomes large and consequently the innovation $\boldsymbol{e}$ strongly contributes to the weight update. As a result of the variance propagation also a cofactor matrix of the estimated weights $\boldsymbol{Q}_{\hat{w}\hat{w},i+1}$ is available (9). These will be examined in the next section.

The estimation is carried out as follows: The weights $\boldsymbol{w}_i$ are updated for each batch t_b. The dimension of the Jacobian matrix $\boldsymbol{A}$ comprises the number of training samples belonging to one batch times the number of ANN outputs and the number of weights ($N_b \times K = (T_b \cdot N) \times K$). After updating all batches ($I = \frac{T_b}{T}$), a new iteration cycle *iter* begins. The estimation terminates if the defined number of iterations (*Iter*) has been reached.

$$\boldsymbol{Q}_{\bar{w}\bar{w},i+1} = \boldsymbol{I}\boldsymbol{Q}_{\hat{w}\hat{w},i}\boldsymbol{I}^T + \boldsymbol{Q}_{pp} \tag{6}$$

$$\hat{\boldsymbol{w}}_{i+1} = \bar{\boldsymbol{w}}_{i+1} + \boldsymbol{K}_{i+1}\Big(\boldsymbol{y}(t_b) - \boldsymbol{f}\left(\bar{\boldsymbol{w}}_{i+1}, \boldsymbol{X}(t_b)\right)\Big) \tag{7}$$

$$\boldsymbol{K}_{i+1} = \boldsymbol{Q}_{\bar{w}\bar{w},i+1}\boldsymbol{A}_{i+1}^T\left(\boldsymbol{Q}_{yy,i+1} + \boldsymbol{A}_{i+1}\boldsymbol{Q}_{\bar{w}\bar{w},i+1}\boldsymbol{A}_{i+1}^T\right)^{-1} \tag{8}$$

$$\boldsymbol{Q}_{\hat{w}\hat{w},i+1} = \boldsymbol{Q}_{\bar{w}\bar{w},i+1} - \boldsymbol{K}_{i+1}\boldsymbol{A}_{i+1}\boldsymbol{Q}_{\bar{w}\bar{w},i+1} \tag{9}$$

2.3 Analogies between LM and EKF

Comparing the updating rules of LM (3) and EKF (7)–(8) a similarity is observable. The equivalence of the two methods is derived on basis of simplifications and rearrangements of both equations. If the updating of LM is expanded as presented in (10), a similar structure as the EKF algorithm (7)–(8) is reached. The invertibility of the resulting term in the brackets of (10) is guaranteed on basis of the learning rate λ, which serves as well as a regularisation term. It prevents the term to become zero and hence full rank is achieved.

$$\begin{aligned}
&\underset{K\times N_b N_b\times K}{(\boldsymbol{J}^T}\underset{}{\boldsymbol{J}} + \lambda\underset{K\times K}{\boldsymbol{I}})^{-1}\underset{K\times N_b N_b\times 1}{\boldsymbol{J}^T\boldsymbol{e}}\\
&= (\boldsymbol{J}^T\boldsymbol{J} + \lambda\boldsymbol{I})^{-1}\boldsymbol{J}^T(\boldsymbol{J}\boldsymbol{J}^T + \lambda\boldsymbol{I})(\boldsymbol{J}\boldsymbol{J}^T + \lambda\boldsymbol{I})^{-1}\boldsymbol{e}\\
&= (\boldsymbol{J}^T\boldsymbol{J} + \lambda\boldsymbol{I})^{-1}(\boldsymbol{J}^T\boldsymbol{J} + \lambda\boldsymbol{I})\boldsymbol{J}^T(\boldsymbol{J}\boldsymbol{J}^T + \lambda\boldsymbol{I})^{-1}\boldsymbol{e}\\
&= \underset{K\times N_b}{\boldsymbol{J}^T}\,\underset{N_b\times K K\times N_b}{(\boldsymbol{J}\,\boldsymbol{J}^T}\,+\lambda\underset{N_b\times N_b}{\boldsymbol{I}})^{-1}\underset{N_b\times 1}{\boldsymbol{e}}
\end{aligned} \tag{10}$$

Necessary simplifications of the EKF updating to obtain equivalence between (7) and (10) are $\boldsymbol{Q}_{\bar{w}\bar{w},i} = \boldsymbol{I}$, $\boldsymbol{Q}_{yy,i} = \lambda\boldsymbol{I}$

Brought to you by | TU Wien Bibliothek
Authenticated
Download Date | 3/26/19 2:46 PM

and no updating of the weights cofactor matrix (6). Due to the reason, that $\boldsymbol{A}$ contains the derivatives of the observations $\left(\frac{\partial \hat{y}_n(t)}{\partial \overline{w}_k}\right)$ and $\boldsymbol{J}$ the derivatives of the error $\left(\frac{\partial e_n(t)}{\partial w_k}\right)$, a difference in the algebraic sign is present, as can be noted from (1). The transition from $\boldsymbol{A}$ to $\boldsymbol{J}$ leads to the change of the sign and the EKF updating becomes:

$$\hat{\boldsymbol{w}} = \overline{\boldsymbol{w}} - \boldsymbol{J}^T(\lambda \boldsymbol{I} + \boldsymbol{J}\boldsymbol{J}^T)^{-1}\boldsymbol{e}.$$

This indicates that in typical training of neural networks it is not accounted for stochastic information of the output data nor for the propagation of covariances between estimated weights in one epoch to subsequent ones.

The variation of the learning rate λ in LM controls the region, where the assumption of the model's linearity is admissible. If the criterion of a decreasing total error in subsequent iterations is not fulfilled, λ is increased to change to gradient descent optimisation. In case of fulfilment, λ will be decreased. Thus, a faster convergence can be achieved. A similar variation of the region can be obtained in the Kalman filter based estimation procedure on basis of the observation variance factor $q_{yy,i}$. The variance factor $q_{yy,i}$ is conform to an increment, considering $\boldsymbol{\Sigma}_{yy} = \sigma_0^2 q_{yy}\widetilde{\boldsymbol{Q}}_{yy}$. However, the adaption of the variance factor q_{yy} bases on a compatibility test. Thereby, the consistency of the functional and stochastic model is proven. The test value is given by:

$$T_{i+1} = \frac{\boldsymbol{e}_{i+1}^T \boldsymbol{Q}_{ee,i+1}^{-1} \boldsymbol{e}_{i+1}}{\sigma_0^2} \tag{11}$$

$$\boldsymbol{Q}_{ee,i+1} = q_{yy,i}\widetilde{\boldsymbol{Q}}_{yy,i+1} + \boldsymbol{A}_{i+1}\boldsymbol{Q}_{\bar{w}\bar{w},i+1}\boldsymbol{A}_{i+1}^T \tag{12}$$

and consists of the normalized and decorrelated innovation and the a priori variance factor σ_0^2. The test value follows a χ^2-distribution with $n_{y,i}$ degrees of freedom and is computed in each iteration step i. If the test value T does not lie in the corresponding quantiles of the χ^2-distribution (a two sided test is assumed), the stochastic and functional model need to be adapted. In case of ANN the model selection task is of major importance and quite challenging with regard to the generalisation capability, see [3] or [15]. However, an adequate chosen functional model is assumed and therefore we focus on the stochastic model. When the test value exceeds the upper limit of the confidence interval, one firstly seeks for outliers in the observations by applying an individual test for each innovation value. Instead of an elimination, a down-weighting of the observations is accomplished here. Increasing the observation noise has the same effect on the test outcome as decreasing the process noise. On the other hand, adapting the observation noise by $q_{yy,i}$ is quite similar to the adaption of λ in LM, when equal weighted and uncorrelated observations are considered. However, λ is a functional regularisation parameter and problem dependent defined by heuristics. $q_{yy,i}$ is derived from the properties of the used sensor and the variation is controlled by the quantiles of the χ^2-distribution.

In equation (13) the Kalman gain is derived for the one-dimensional case with a direct observed weight in order to analyse the role of the process noise. The higher the process noise σ_{pp}^2, the higher the gain and implicitly the weight for the innovation in the updating equation (7).

$$\boldsymbol{K}_i = \frac{1}{1 + \left(\frac{\sigma_{yy,i+1}^2}{\sigma_{\overline{w}\overline{w},i+1}^2}\right)} = \frac{1}{1 + \left(\frac{\sigma_{yy,i+1}^2}{\sigma_{\hat{w}\hat{w},i}^2 + \sigma_{pp,i+1}^2}\right)} \tag{13}$$

At the beginning of the estimation the weight noise $\sigma_{\bar{w}\bar{w}}^2$ is quite high in comparison to the observation and process noise. Therefore, the compatibility test will first be determined after 10 iteration steps. During the filter progress the uncertainty of the weights decreases and the interplay between process and observation noise becomes more important. The interplay is mainly steered by varying the observation noise because of demonstrating the analogy of LM and EKF.

An additional variation of the process noise influences the convergence behaviour of the filter positively and it leads to superior results in comparison to non adapting process noise according to [19]. In this paper, the variation is accomplished by the process variance factor q_{pp}, an additional increment of the covariance matrix $\boldsymbol{\Sigma}_{pp} = \sigma_0^2 q_{pp}\widetilde{\boldsymbol{Q}}_{pp}$. It will be evaluated in Section 4.3.

The weight noise $\sigma_{\hat{w}\hat{w}}^2$ provided by the filter solution enables a testing of the weights on their significance. The formulated null hypothesis postulates that the weight is equal to zero and leads to the test in Equation (14). This information can be helpful in the model selection task and will be evaluated as well in Section 4.3.

$$P\left(\frac{|\hat{w}|}{\sigma_{\hat{w}\hat{w}}} \leq t_{T-K,1-\frac{\alpha}{2}}\right) = 1 - \alpha \tag{14}$$

The combination of ANN and EKF presents some advantages: The updating of the weight cofactor matrix $\boldsymbol{Q}_{\hat{w}\hat{w}}$ is one characteristic of the EKF. A reasonable initialisation of the observation cofactor matrix $\boldsymbol{Q}_{yy}$ is possible due to the combination of EKF and ANN. The compatibility test is an alternative to the constraint of a decreasing total error. The adaption of the observation variance factor q_{yy} is accomplished according to the learning rate λ. The cofactor matrix of the estimated weights $\boldsymbol{Q}_{\hat{w}\hat{w}}$ enables a testing on the significance of the weights.

Brought to you by | TU Wien Bibliothek
Authenticated
Download Date | 3/26/19 2:46 PM

3 Experiment

The developed method including the analogy between EKF and ANN will be evaluated on the system robot arm. The position deviations of the robot arm shall be modelled as a function of the joint encoder values in order to be used as a correction tool during robot's operation. Therefore, the ANN should only reflect the positioning deviations of the robot. Considerable effort needs to be put on the reference frame transformation between measurement device and robot arm as well as the estimation of lever arm components between the end-effector and the robot flange. A possible approach to this task is discussed in the following.

3.1 Concept

Due to a typical repeatability of such an industrial robot arm of 0.1 mm, the chosen measurement method consists of a laser tracker (for specifications see Table 5) in combination with a probe, which enables a 6-DOF pose measurement. The probe is rigidly coupled with the robot arm flange. The setting for the measurement is illustrated in Figure 1. Four frames are involved and these are related by four transformations (see Table 1).

The notation for transformations is defined in equation (15) – an arbitrary point $\boldsymbol{p}_i$ in frame (S) is transformed with respect to the frame (T). The translational part consists of the three coordinate components ${}^{T}\boldsymbol{t}_S = [t_x, t_y, t_z]$ and the rotation matrix ${}^{T}\boldsymbol{R}_S$ is defined on basis of angles e. g. $\boldsymbol{\alpha} = (\alpha, \beta, \gamma)^T$ (see Table 1).

$$ {}^{T}\boldsymbol{p}_i = {}^{T}\boldsymbol{R}_S \, {}^{S}\boldsymbol{p}_i + {}^{T}\boldsymbol{t}_S \tag{15} $$

Due to the consideration of only robot position deviations $\boldsymbol{\delta}_R$ in the ANN, the orientation of the robot flange

Table 1: Four transformations (Orientations in light grey not considered in the model of equation (16)).

Symbol	Description
${}^{R}\boldsymbol{T}_{LT}$	Transformation ($\boldsymbol{T}$) representing the laser tracker frame (LT) with respect to the robot base frame (R): ${}^{R}\boldsymbol{R}_{LT} = \boldsymbol{R}(\omega, \phi, \kappa) = \boldsymbol{R}_x(\omega)\boldsymbol{R}_y(\phi)\boldsymbol{R}_z(\kappa)$, ${}^{R}\boldsymbol{t}_{LT}$
${}^{LT}\boldsymbol{T}_P$	Transformation representing the probe frame (P) with respect to the laser tracker frame: ${}^{LT}\boldsymbol{R}_P = \boldsymbol{R}(\alpha, \beta, \gamma) = \boldsymbol{R}_z(\gamma)\boldsymbol{R}_y(\beta)\boldsymbol{R}_x(\alpha)$, ${}^{LT}\boldsymbol{t}_P$
${}^{P}\boldsymbol{T}_{RF}$	Transformation representing the robot flange frame (RF) with respect to the probe frame: ${}^{P}\boldsymbol{R}_{RF}$, ${}^{P}\boldsymbol{t}_{RF}$
${}^{R}\boldsymbol{T}_{RF}$	Orientation and position of the robot flange frame with respect to the robot base frame: ${}^{R}\boldsymbol{R}_{RF}$, ${}^{R}\boldsymbol{t}_{RF}$

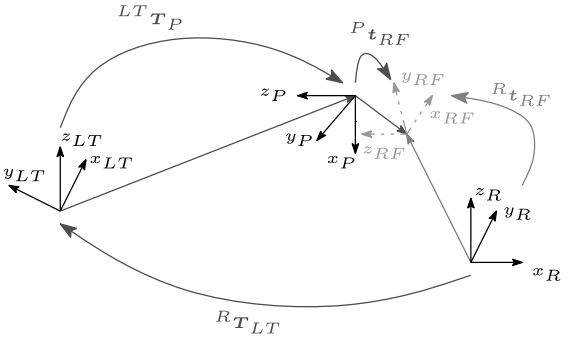

Figure 1: Sketch of involved transformations (Robot flange frame (light grey) only relevant for simulation).

frame ${}^{R}\boldsymbol{R}_{RF}$ is neglected. Consequently, the orientation of the robot flange with respect to the probe ${}^{P}\boldsymbol{R}_{RF}$ is not required and only the translational part ${}^{P}\boldsymbol{t}_{RF}$, the so called lever arm, needs to be determined. The function is given in equation (16) and follows the sketch in Figure 1. The robot flange position with respect to the robot base frame ${}^{R}\boldsymbol{t}_{RF}$ can be derived as well, if the lever arm ${}^{P}\boldsymbol{t}_{RF}$ is first transformed with respect to the laser tracker frame ${}^{LT}\boldsymbol{T}_P$ and in a second step with respect to the robot frame ${}^{R}\boldsymbol{T}_{LT}$.

$$ {}^{R}\boldsymbol{t}_{RF} = {}^{R}\boldsymbol{R}_{LT}\left({}^{LT}\boldsymbol{R}_P \, {}^{P}\boldsymbol{t}_{RF} + {}^{LT}\boldsymbol{t}_P\right) + {}^{R}\boldsymbol{t}_{LT} \tag{16} $$

The measures in (16) can be split up in unknowns $\boldsymbol{x}$ and observations $\boldsymbol{l}$ and are summarised in equation (17). The observations consist of the probe pose ${}^{LT}\boldsymbol{T}_P$ and the robot flange position ${}^{R}\boldsymbol{t}_{RF}$. The transformation of the laser tracker frame with respect to the robot base frame ${}^{R}\boldsymbol{T}_{LT}$ as well as the leverarm ${}^{P}\boldsymbol{t}_{RF}$ are the unknown measures.

In a first evaluation step, the lever arm has been determined by keeping a constant robot position while varying the robot orientation. Thereby, the probe reference points describe a sphere and the difference vector to the sphere centre gives the lever arm in the appropriate probe frame. On basis of the determined lever arm ${}^{P}\boldsymbol{t}_{RF}$, the transformation parameters ${}^{R}\boldsymbol{T}_{LT}$ have been estimated. However, the weakness of the approach is the lack of a persistent variance propagation of the observations and consequently of a reliable statement of the reached lever arm uncertainty. Therefore, the aim is to determine both unknown measures together in one adjustment on basis of equation (16).

$$ \boldsymbol{x} = \begin{bmatrix} \boldsymbol{\omega} \\ {}^{R}\boldsymbol{t}_{LT} \\ {}^{P}\boldsymbol{t}_{RF} \end{bmatrix} \qquad \boldsymbol{l} = \begin{bmatrix} \boldsymbol{\alpha}_1 \\ {}^{LT}\boldsymbol{t}_{P,1} \\ {}^{R}\boldsymbol{t}_{RF,1} \\ \vdots \\ \boldsymbol{\alpha}_N \\ {}^{LT}\boldsymbol{t}_{P,N} \\ {}^{R}\boldsymbol{t}_{RF,N} \end{bmatrix} \tag{17} $$

Brought to you by | TU Wien Bibliothek
Authenticated
Download Date | 3/26/19 2:46 PM

$$\boldsymbol{\Sigma}_{ll} = \begin{bmatrix} \boldsymbol{\Sigma}_{\boldsymbol{\alpha}_1} & & & & \\ & \boldsymbol{\Sigma}_{^{LT}\boldsymbol{t}_{P,1}} & & & \\ & & \boldsymbol{\Sigma}_{^{R}\boldsymbol{t}_{RF,1}} & & \\ & & & \ddots & \\ & & & & \boldsymbol{\Sigma}_{^{R}\boldsymbol{t}_{RF,N}} \end{bmatrix} \quad (18)$$

Due to the reason that the observations $\boldsymbol{l}$ are an implicit function of the parameters $\boldsymbol{x}$ (19), the least squares adjustment with condition equation – the Gauss-Helmert model (GHM) – needs to be applied, which is treated in the next section.

$$\boldsymbol{f}(\boldsymbol{l} + \boldsymbol{v}, \hat{\boldsymbol{x}}) = 0 \quad (19)$$

3.2 Estimation

The non-linear condition for the estimation in the Gauss-Helmert model follows from (16) into (20). The condition $\boldsymbol{f}_1$ is equal to zero in case of estimated parameters, which are defined by a hat, and the observations corrected by the residuals $\boldsymbol{v}$. The corresponding stochastic model $\boldsymbol{\Sigma}_{ll}$ is defined in (18), whereby the sub matrix $\boldsymbol{\Sigma}_{^{LT}\boldsymbol{t}_{P,i}}$ is a full covariance matrix originating from the variance propagated polar measures of the laser tracker. The sub matrices $\boldsymbol{\Sigma}_{\boldsymbol{\alpha},i}$ and $\boldsymbol{\Sigma}_{^{R}\boldsymbol{t}_{RF,i}}$ have diagonal form with equal weighted components. A transition to the cofactor matrix $\boldsymbol{Q}_{ll}$ is fulfilled by considering the variance factor of unit weight ($\boldsymbol{\Sigma}_{ll} = \sigma_0^2 \boldsymbol{Q}_{ll}$).

$$\begin{aligned} \boldsymbol{f}_1 : {}^{R}\hat{\boldsymbol{R}}_{LT} \left({}^{LT}\boldsymbol{R}_P + \boldsymbol{v}_{^{LT}\boldsymbol{R}_P} \right) {}^{P}\hat{\boldsymbol{t}}_{RF} \\ + {}^{R}\hat{\boldsymbol{R}}_{LT} \left({}^{LT}\boldsymbol{t}_P + \boldsymbol{v}_{^{LT}\boldsymbol{t}_P} \right) + {}^{R}\hat{\boldsymbol{t}}_{LT} - \left({}^{R}\boldsymbol{t}_{RF} + \boldsymbol{v}_{^{R}\boldsymbol{t}_{RF}} \right) \stackrel{!}{=} 0 \end{aligned} \quad (20)$$

The least squares adjustment claims that the weighted squared residuals $\boldsymbol{v}$ shall become a minimum ($\boldsymbol{v}^T \boldsymbol{Q}_{ll}^{-1} \boldsymbol{v} \rightarrow min$). It is solved by linearising equation (20) by Taylor series at approximate values $\boldsymbol{x}_0$ and $\boldsymbol{l}$ for a first solution. Subsequently, the points of expansion of the Taylor series are the estimated ones of the prior iteration step $i-1$: $\hat{\boldsymbol{x}}_i = \hat{\boldsymbol{x}}_{i-1} + \Delta\hat{\boldsymbol{x}}_i$ and $\hat{\boldsymbol{l}}_i = \boldsymbol{l} + \hat{\boldsymbol{v}}_i$ in [18, p. 5].

$$\begin{aligned} \boldsymbol{f}(\hat{\boldsymbol{x}}_i, \hat{\boldsymbol{l}}_i) = & \left.\frac{\partial \boldsymbol{f}}{\partial \hat{\boldsymbol{x}}}\right|_{\hat{\boldsymbol{x}}_{i-1}, \hat{\boldsymbol{l}}_{i-1}} (\hat{\boldsymbol{x}}_i - \hat{\boldsymbol{x}}_{i-1}) + \left.\frac{\partial \boldsymbol{f}}{\partial \hat{\boldsymbol{l}}}\right|_{\hat{\boldsymbol{x}}_{i-1}, \hat{\boldsymbol{l}}_{i-1}} (\hat{\boldsymbol{l}}_i - \boldsymbol{l}) \\ & + \left.\frac{\partial \boldsymbol{f}}{\partial \hat{\boldsymbol{l}}}\right|_{\hat{\boldsymbol{x}}_{i-1}, \hat{\boldsymbol{l}}_{i-1}} (\boldsymbol{l} - \hat{\boldsymbol{l}}_{i-1}) + \boldsymbol{f}(\hat{\boldsymbol{l}}_{i-1}, \hat{\boldsymbol{x}}_{i-1}) \end{aligned} \quad (21)$$

Equation (21) in matrix notation is presented in (22). It consists of the design matrix $\boldsymbol{A}_i$, the estimated parameter updates $\Delta\hat{\boldsymbol{x}}_i$, the condition matrix $\boldsymbol{B}_i$, the estimated residuals $\hat{\boldsymbol{v}}_i$ ($\hat{\boldsymbol{v}}_i = \hat{\boldsymbol{l}}_i - \boldsymbol{l}$) and the misclosure $\boldsymbol{w}_i$, which corresponds to the last two terms in equation (21).

$$\boldsymbol{A}_i \Delta\hat{\boldsymbol{x}}_i + \boldsymbol{B}_i \hat{\boldsymbol{v}}_i + \boldsymbol{w}_i = 0 \quad (22)$$

To derive the parameters update $\Delta\hat{\boldsymbol{x}}_i$ the least squares approach with the additional condition (22) needs to be solved. Minimising this Lagrange function Ω in (23) leads to the estimated parameter update $\Delta\hat{\boldsymbol{x}}_i$ in equation (25).

$$\Omega : \quad \boldsymbol{v}_i^T \boldsymbol{Q}_{ll,i}^{-1} \boldsymbol{v}_i - 2\boldsymbol{k}^T (\boldsymbol{A}_i \Delta\hat{\boldsymbol{x}}_i + \boldsymbol{B}_i \hat{\boldsymbol{v}}_i + \boldsymbol{w}_i) \rightarrow min \quad (23)$$

$$\boldsymbol{Q}_{\hat{x}\hat{x},i} = (\boldsymbol{A}_i^T (\boldsymbol{B}_i \boldsymbol{Q}_{ll,i} \boldsymbol{B}_i^T)^{-1} \boldsymbol{A}_i)^{-1} \quad (24)$$

$$\Delta\hat{\boldsymbol{x}}_i = (\boldsymbol{A}_i^T (\boldsymbol{B}_i \boldsymbol{Q}_{ll,i} \boldsymbol{B}_i^T)^{-1} \boldsymbol{A}_i)^{-1} \boldsymbol{A}_i^T (\boldsymbol{B}_i \boldsymbol{Q}_{ll,i} \boldsymbol{B}_i^T)^{-1} \boldsymbol{w}_i \quad (25)$$

The iterative estimation of $\Delta\hat{\boldsymbol{x}}_i$ proceeds until the absolute difference of consecutive parameters $\|\hat{\boldsymbol{x}}_i - \hat{\boldsymbol{x}}_{i-1}\|$ is smaller than a defined value ϵ of 1e-9 [12]. The number of iterations depend on the quality of the approximate values $\boldsymbol{x}_0$. The cofactor matrix of the estimated parameters $\boldsymbol{Q}_{\hat{x}\hat{x},i}$ (24) is obtained by variance propagation.

The functional and stochastic model is verified by the global test of adjustment, which is equivalent to the compatibility test in Section 2.3. It bases on the a priori variance factor σ_0^2 and the residual square sum, which is an unbiased estimator of the variance factor $\hat{o}_0^2$. If the variance factors do not originate from the same population, which will be tested on a significance level α, model errors are probable.

$$P\left(\frac{\boldsymbol{v}_i^T \boldsymbol{Q}_{ll,i}^{-1} \boldsymbol{v}_i}{(r-u)\sigma_0^2} = \frac{-\boldsymbol{k}_i^T (\boldsymbol{w}_i + \boldsymbol{A}_i \Delta\boldsymbol{x}_i)}{(r-u)\sigma_0^2} > F_{r-u,\infty} \right) = \alpha \quad (26)$$

3.3 Measured data

The correction tool based on the ANN has been developed for a KUKA robot arm. The high accuracy type is specified with a repeatability of 30 μm, a maximum load of 60 kg and a reach of 2033 mm. As already mentioned, the measurements are accomplished with a laser tracker Leica LTD800 in combination with a T-Cam and the T-Probe 3. Therefore, the probe has been mounted on the spindle head of the robot arm.

The measurements are carried out in a stop and go mode. Thereby, the robot moves in a regular grid to the next robot pose. In total, six raster have been observed, a data set with 779 positions has been derived. Their spatial distribution is presented in Figure 2. The realised variation in the encoder values is shown in Figure 3. The aim was to reach a good variation in the joint encoders and to move into the end positions of the encoders. The latter has been reached for the joints A2, A3 and A5, which can be verified by Figure 3 and Table 2.

The first preprocessing step includes the transformation of the laser tracker observations ${}^{LT}\boldsymbol{t}_P$ and ${}^{LT}\boldsymbol{R}_P$ into the robot frame. Therefore, the approach described in equa-

Brought to you by | TU Wien Bibliothek
Authenticated
Download Date | 3/26/19 2:46 PM

tion (16) is applied. Subsequently, robot position deviations are computable, which will be introduced in the ANN as output values for the training.

On basis of transformation points given in both frames (R and LT) the estimation of the unknown parameters is accomplished. The transformation points cover the area of interest and a variation in the robot pose is needed in order to separate the lever arm and the translational part of ${}^{R}\boldsymbol{T}_{LT}$. Thus, 20 points out of the raster data were selected, which had not already been considered for the computation of the robot deviations in Figure 2 and 3.

Sufficient initial values for the non-linear parameter estimation have been derived on basis of an iterative approach. First the lever arm has been determined by keeping a constant robot position while varying the robot orientation, as described in Section 3.1. Adding the lever arm to the laser tracker measurements gives the robot flange position in the laser tracker frame. This enables the estimation

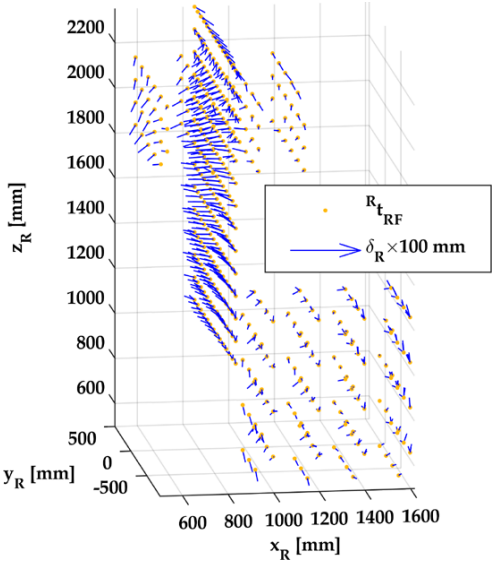

Figure 2: Raster positions ${}^{R}\boldsymbol{t}_{RF}$ and deviations $\boldsymbol{\delta}_R$ in the robot frame.

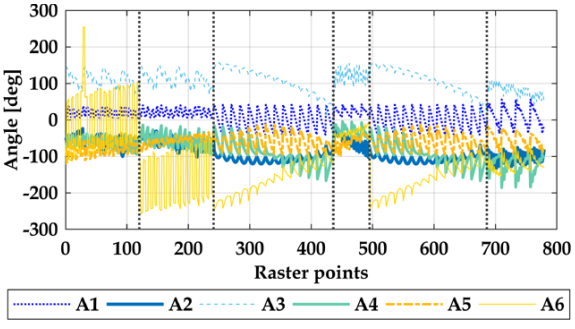

Figure 3: Encoder values of the six joints.

Table 2: Encoder range of values for the six joints.

Joint	Range
A1	±185°
A2	+35°/−135°
A3	+158°/−120°
A4	±350°
A5	±119°
A6	±350°

Table 3: Result of parameter estimation for the Kuka data set.

Parameter	$\hat{\boldsymbol{x}}$	$\hat{\boldsymbol{\sigma}}$
ω	−0.2120°	15″
ϕ	−0.3266°	18″
κ	−26.7056°	23″
${}^{R}t_{LT,x}$	2.6219 m	312 μm
${}^{R}t_{LT,y}$	−3.0266 m	213 μm
${}^{R}t_{LT,z}$	1.3262 m	202 μm
${}^{P}t_{RF,x}$	0.0452 m	50 μm
${}^{P}t_{RF,y}$	−0.0210 m	55 μm
${}^{P}t_{RF,z}$	−0.2377 m	180 μm

of approximate transformation parameters between robot and laser tracker frame.

Transformation points and sufficient approximate values enable the parameter estimation according to Section 3.1 and 3.2. Manufacturer specifications are assumed as a priori standard deviations – the laser tracker uncertainties are stated in Table 5, the robot arm is specified with a repeatability of 0.03 mm. However, after down weighting the outliers, the global test of the adjustment has been rejected on a significance level of 5 %. If the laser tracker measurements are accomplished in the uncertainty level as defined in Table 5, an a posteriori standard deviation of the robot of 0.2 mm is achieved, which obviously exceeds the manufacturer's specification by almost one order of magnitude. The corresponding result is presented in Table 3. Due to the lack of access to the robot arm it is not possible to repeat the measurements in order to check for the robot's standard deviation. Therefore, the stochastic model will be verified on basis of a simulation data set in the next section.

In order to obtain the robot position deviations $\boldsymbol{\delta}_R$ as well as their uncertainties $\boldsymbol{\sigma}_{\delta R}$, the model in (16) is used again. Now, the estimated transformation parameters of Table 3 are applied to compute the robot position deviations for the raster data according to equation (27). The robot position deviations are presented in Figure 2 and the time series in Figure 9. The largest robot position deviations $\boldsymbol{\delta}_R$ occur in the planar raster (${}^{R}\boldsymbol{t}_{RF} = 850\,mm$), which

Brought to you by | TU Wien Bibliothek
Authenticated
Download Date | 3/26/19 2:46 PM

has been observed twice. The volumetric raster shows some systematic deviations as well. The variance propagation of equation (27) results in a mean uncertainty of the robot deviations $\boldsymbol{\sigma}_{\delta R}$ of [0.48, 0.37, 0.38] mm.

$$\boldsymbol{\delta}_R = {}^R\boldsymbol{R}_{LT}\,{}^{LT}\boldsymbol{R}_P\,{}^P\boldsymbol{t}_{RF} + {}^R\boldsymbol{R}_{LT}\,{}^{LT}\boldsymbol{t}_P + {}^R\boldsymbol{t}_{LT} - {}^R\boldsymbol{t}_{RF} \qquad (27)$$

3.4 Simulation

The aim of the simulation is to proof the functional and stochastic model of Section 3.3. This includes the verification, whether the estimation is unbiased and the approach provides the expected precision of the results or the precision obtained in the empirical case respectively.

The computation of simulation data is summarised in Figure 4. Different to real data, the robot arm deficiences are analysed in more detail – there is a difference between the robot's sensed pose of its flange and the effective pose of the flange. Therefore, the robot position uncertainty is split up in a repeatability $\boldsymbol{\sigma}_{Rep}$ and a sensor noise $\boldsymbol{\sigma}_{f_{Enc}}$, which is mainly caused by the joint encoder noise. The robot arm arrives at a defined position repeatedly and the effective position varies according to $\boldsymbol{\sigma}_{Rep}$. The laser tracker always measures the effective position and the robot output varies about $\boldsymbol{\sigma}_{f_{Enc}}$. In [11] the position repeatability is defined for a constant pose. In this contribution the position repeatability refers to a constant robot flange position but a varying orientation. The orientation repeatability is not considered explicitly in ${}^R\boldsymbol{T}_{RF}$.

Thus, to compute laser tracker observations, the repeatability is explicitly considered in the robot position (Figure 4). Thereon, the laser tracker observations are calculated. The laser tracker noise ($\boldsymbol{\sigma}_\alpha$, $\boldsymbol{\sigma}_{{}^{LT}\boldsymbol{t}_P}$) is added to the simulated laser tracker measurements, whereby the translational laser tracker uncertainty $\boldsymbol{\sigma}_{{}^{LT}\boldsymbol{t}_P}$ splits up in the distance σ_{LT_d} and the bearing $\sigma_{LT_{ang}}$ measurement uncertainty. The robot observations correspond to the robot's output, only the sensor noise $\boldsymbol{\sigma}_{f_{Enc}}$ is added, as presented in the Scheme of Figure 4.

In order to verify the functional and the stochastic model, $D_1 \times D_2$ independent realisations of each observation pair n ($n = 1, \ldots, N$) are generated on basis of normal distributed random numbers scaled by the corresponding standard deviations. The observations are computed subsequently – a so called Monte-Carlo simulation is accomplished. Due to the successive computation of the laser tracker observations (Figure 4), D_1 robot positions ($d_1 = 1, \ldots, D_1$) with σ^2_{Rep} are generated and on each of these d_1 realisations D_2 ($d_2 = 1, \ldots, D_2$) independent laser tracker observations with σ^2_α and $\sigma^2_{{}^{LT}\boldsymbol{t}_P}$ are generated to cover all distributions of the random variables. The $D_1 \times D_2$ realisations for each robot position n are generated with $\sigma_{f_{Enc}}$ and serve as observation in the adjustment.

To estimate the nine unknown parameters from equation (17), at minimum three conditions of type (20) are needed as the condition consists of three coordinate components ($r = 3 \times N$). In accordance to the real data, 20 robot poses ($N = 20$) are used as an initial dataset (simulation set 1) and consequently 60 equations are introduced. These robot poses are distributed in space as presented in Figure 5. In order to compute the Monte-Carlo simulation the data has been generated according to the scheme in Figure 4. Therefore, the standard deviations given in Table 5 were applied. The $\sigma_{f_{Enc}}$ corresponds to a standard deviation of the robot's sensed encoder values, if equal robot positions are given. The same standard deviations are used

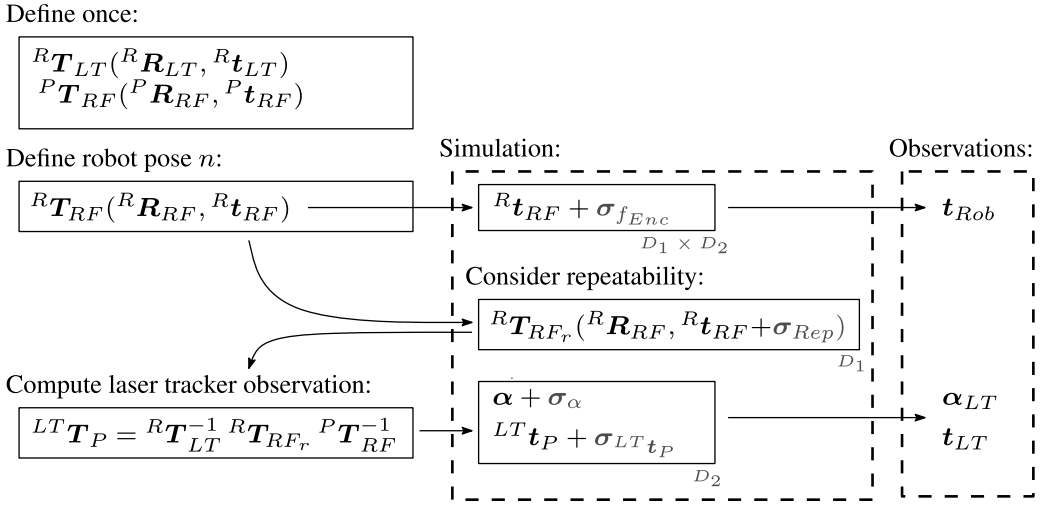

Figure 4: Scheme for data generation and the Monte-Carlo simulation.

Brought to you by | TU Wien Bibliothek
Authenticated
Download Date | 3/26/19 2:46 PM

for the stochastic model. The initial values of the parameters deviate from the true values for angles by about 5° and for translational measures by 500 mm.

The adjustment according to Section 3.2 delivers the results in the left part of Table 4 in case of a robot position repeatability of 30 μm. The global test of adjustment was not rejected. The deviations to the true values are nearly zero. Consequently, an unbiased estimation is guaranteed. The standard deviation of the worst determined parameter is estimated in the level of 1.5 to 2 times the repeatability.

If the data is generated on basis of a $\sigma_{Rep} = 0.2\,mm$, but the stochastic model still considers a repeatability of 0.03 mm, a similar case as for the real data is simulated. The global test of the adjustment was rejected. The $\hat{\sigma}_0^2$ results in 27. Compared to that, for the real data the $\hat{\sigma}_0^2$ amounts to 22, if a repeatability of 0.03 mm is assumed in the stochastic model.

If the adjustment is computed with a robot position repeatability of 0.2 mm – regarding both, the data generation as well as the stochastic model, the results in the right part of Table 4 are obtained. These estimated parameters show a similar uncertainty level as for the real data in Table 3. Consequently, the Kuka robot arm does not reach the specified repeatability of 0.03 mm. The reason for this can be found in the advanced age of the robot arm and its components. Moreover, the simulation shows, that the robot arm repeatability mainly influences the quality of the transformation parameters.

Examining the results in more detail, the correlation structure of the parameters in Figure 6 is conspicuous. The transformation parameters ${}^{R}\boldsymbol{T}_{LT}$ among themselves as well as ${}^{R}\boldsymbol{t}_{LT,x}$ and the lever arm component ${}^{P}\boldsymbol{t}_{RF,z}$ are highly correlated as shown in Figure 6. High correlations between parameters ($\rho_{\hat{x}\hat{x}} > 0.8$) indicate a possible under determined system. Two options to deal with this fact are:

Table 4: Results of simulation set 1 (mean values of 250.000 realisations).

Sim1: $\sigma_{Rep} = 30\,\mu m$			Sim1: $\sigma_{Rep} = 200\,\mu m$	
Parameter	$\tilde{x} - \hat{x}$	$\hat{\sigma}$	$\tilde{x} - \hat{x}$	$\hat{\sigma}$
ω	0″	4″	1″	22″
ϕ	0″	4″	−1″	22″
κ	0″	4″	0″	18″
${}^{R}t_{LT,x}$	0 μm	21 μm	−2 μm	110 μm
${}^{R}t_{LT,y}$	2 μm	49 μm	−5 μm	250 μm
${}^{R}t_{LT,z}$	−1 μm	57 μm	−6 μm	300 μm
${}^{P}t_{RF,x}$	0 μm	9 μm	0 μm	47 μm
${}^{P}t_{RF,y}$	0 μm	9 μm	0 μm	46 μm
${}^{P}t_{RF,z}$	−1 μm	18 μm	−1 μm	101 μm

either improving the geometric configuration of the observations or setting one of the two correlated parameters constant. In robotic literature, measures to describe inadequate data and unidentifiable parameters are summarised as observability and identifiability, see [10, p. 323 ff]. We will focus on the observability of the parameters on basis of the correlations of the parameter cofactor matrix $\boldsymbol{Q}_{\hat{x}\hat{x}}$.

In the chosen observation configuration, ${}^{P}\boldsymbol{t}_{RF,z}$ and ${}^{R}\boldsymbol{t}_{LT,x}$ are effective in the same direction – in the x direction of the robot coordinate system (see Figure 1). That originates from the alignment of the probe, mounted on the robot flange, to the laser tracker in order to be detectable for the laser tracker. In case of the simulation data the z-axis of the robot flange frame (blue coloured axis in Figure 5) needs to be aligned to the laser tracker ±45°. For a better separation, a higher variation in the robot flange orientation is necessary. The correlations between the el-

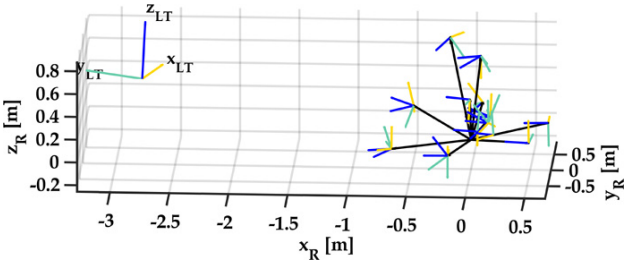

Figure 5: Geometry of simulation data set 1.

Table 5: Observation uncertainties.

Std.Dev.	Value
σ_{Rep}	30 μm
$\sigma_{f_{Enc}}$	20 μm
σ_{LT_D}	20 μm
$\sigma_{LT_{HZ,V}}$	2.6″
σ_{α}	0.01°

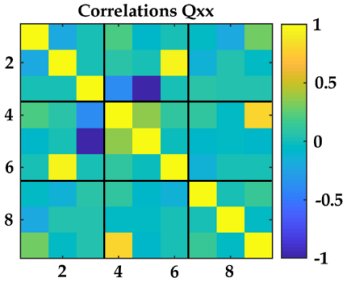

Figure 6: Correlations of the estimated parameters of simulation set 1 (order: $\boldsymbol{\omega}$, ${}^{R}\boldsymbol{t}_{LT}$, ${}^{T}\boldsymbol{t}_{RF}$).

Brought to you by | TU Wien Bibliothek
Authenticated
Download Date | 3/26/19 2:46 PM

ements of ${}^{R}\boldsymbol{T}_{LT}$ are not of that importance because they build one set of parameters.

Moreover, the aim is to reach minimal uncertainties for the computed robot deviations in order to describe the robot and not other components in further consequence. Therefore, the estimation shall be improved. One idea is to extend the parameter estimation by additional information like identical points. This shall enable a better separability of highly correlated parameters on the one hand and improve the parameter estimation in general on the other.

The information of identical robot flange positions is introduced on basis of an additional condition. Equating two identical robot flange realisations ${}^{R}\boldsymbol{t}_{RF,1} = {}^{R}\boldsymbol{t}_{RF,2}$ and substitute it by (16), leads to the reduced form in (28). The first laser tracker observation to an identical point equation is described by (20). For subsequent laser tracker observations to the same robot flange point but with a new orientation equation (28) is used as condition equation of the GHM.

$$\begin{aligned}\boldsymbol{f}_2 :&({}^{LT}\boldsymbol{R}_{P,1} + \boldsymbol{v}_{{}^{LT}\boldsymbol{R}_{P,1}})\,{}^{P}\hat{\boldsymbol{t}}_{RF} + ({}^{LT}\boldsymbol{t}_{P,1} + \boldsymbol{v}_{{}^{LT}\boldsymbol{t}_{P,1}}) \\ &- ({}^{LT}\boldsymbol{R}_{P,i1} + \boldsymbol{v}_{{}^{LT}\boldsymbol{R}_{P,i1}})\,{}^{P}\hat{\boldsymbol{t}}_{RF} - ({}^{LT}\boldsymbol{t}_{P,i1} + \boldsymbol{v}_{{}^{LT}\boldsymbol{t}_{P,i1}}) = 0 \end{aligned} \quad (28)$$

The additional condition (28) implies an identical robot flange position. As mentioned, this identity reflects the "believe" of the robot to achieve the same position. However, due to the repeatability variations σ_{Rep}, slightly different positions are reached effectively. To account for this effect the position repeatability is modelled stochastically. The laser tracker uncertainty for the additional laser tracker observations consists of the propagated polar measurement standard deviations $(\sigma_{LT_D}, \sigma_{LT_{HZ,V}})$ and the difference of two effective robot positions leading to $\sigma_{Rep}\sqrt{2}$.

Considering condition $\boldsymbol{f}_2$, 500×500 $(D_1 \times D_2)$ realisations with $\sigma_{Rep} = 0.1\,\text{mm}$ have been generated. The σ_{Rep} has been chosen as an average value. The structure of the condition matrix $\boldsymbol{B}$ is presented in Figure 7. The non-zero elements of $\boldsymbol{B}$ are coloured. The first 39 condition equations correspond to the standard approach in (20), the subsequent lines describe the identical constraint of equation (28). Thereby, the subsequent laser tracker measurements are introduced as new observations and the connection to the laser tracker observations, leading to an identical robot point is realised by the assignment.

The result of the Monte-Carlo simulation with $\sigma_{Rep} = 0.1\,\text{mm}$ is presented in Table 6. The global test is not acceptable. The level of significance α is again chosen to be 5 %. Despite the adaption of the stochastic model for $\boldsymbol{f}_2$ by $\sigma_{Rep}\sqrt{2}$, the global test is rejected in further 4 % of the cases. The reason for the not fulfilment can be found in the insufficient modelling of the repeatability. Condition (28) is valid for the "true" robot position. However, the laser tracker observes one robot realisation in the area of the repeatability. The estimation of a mean robot position could be a corrective.

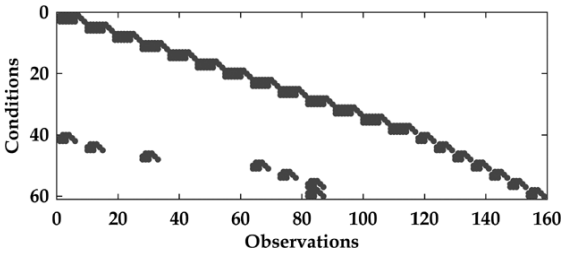

Figure 7: Structure of the condition matrix $\boldsymbol{B}$ of simulation set 1 considering $\boldsymbol{f}_2$ (order: $[\boldsymbol{\alpha}_1; {}^{LT}\boldsymbol{t}_{P,1}; {}^{R}\boldsymbol{t}_{RF,1}; \cdots \boldsymbol{\alpha}_N; {}^{LT}\boldsymbol{t}_{P,N}; {}^{R}\boldsymbol{t}_{RF,N}; \boldsymbol{\alpha}_{i1}; {}^{LT}\boldsymbol{t}_{P,i1}; \cdots \boldsymbol{\alpha}_{iN}; {}^{LT}\boldsymbol{t}_{P,iN}]$;).

If the repeatability σ_{Rep} of the robot results to 0.01 mm instead of 0.1 mm, the compatibility test can be accepted in 95 % of all cases. This confirms the insufficient description of the robot's repeatability in the model $\boldsymbol{f}_1$ and $\boldsymbol{f}_2$. However, the deviation between the true values of the parameter and the estimated one show similar good results as without condition $\boldsymbol{f}_2$ (Table 6). In contrast, the parameter uncertainties are determined more pessimistic.

For the separation of ${}^{P}\boldsymbol{t}_{RF,z}$ and ${}^{R}\boldsymbol{t}_{LT,x}$ a second data set has been simulated with a better observation configuration. Thereby, observations have been included, which are not realisable with the present measurement method – some of the robot flange poses do not fulfil the alignment criterion to the laser tracker. If we state, that the observations are performable, we receive the results in Table 7 and Figure 8. The correlation structure of the parameter covariance nearly shows no correlation between ${}^{P}\boldsymbol{t}_{RF,z}$ and ${}^{R}\boldsymbol{t}_{LT,x}$. This leads to the small uncertainties of these two parameters, which are much smaller than the repeatabil-

Table 6: Results of simulation set 1 without and with condition $\boldsymbol{f}_2$ for $\sigma_{Rep} = 0.1\,mm$ (mean values of 250.000 realisations).

Sim1 excl. $\boldsymbol{f}_2$			Sim1 incl. $\boldsymbol{f}_2$	
Parameter	$\tilde{\boldsymbol{x}} - \hat{\boldsymbol{x}}$	$\hat{\boldsymbol{\sigma}}$	$\tilde{\boldsymbol{x}} - \hat{\boldsymbol{x}}$	$\hat{\boldsymbol{\sigma}}$
ω	$0''$	$11''$	$0''$	$13''$
ϕ	$0''$	$11''$	$0''$	$14''$
κ	$0''$	$9''$	$0''$	$11''$
${}^{R}t_{LT,x}$	$-2\,\mu m$	$55\,\mu m$	$-1\,\mu m$	$61\,\mu m$
${}^{R}t_{LT,y}$	$-2\,\mu m$	$127\,\mu m$	$3\,\mu m$	$154\,\mu m$
${}^{R}t_{LT,z}$	$0\,\mu m$	$152\,\mu m$	$2\,\mu m$	$191\,\mu m$
${}^{P}t_{RF,x}$	$-1\,\mu m$	$24\,\mu m$	$-1\,\mu m$	$27\,\mu m$
${}^{P}t_{RF,y}$	$1\,\mu m$	$23\,\mu m$	$1\,\mu m$	$25\,\mu m$
${}^{P}t_{RF,z}$	$-2\,\mu m$	$51\,\mu m$	$2\,\mu m$	$52\,\mu m$

Brought to you by | TU Wien Bibliothek
Authenticated
Download Date | 3/26/19 2:46 PM

Table 7: Result of simulation set 2 for $\sigma_{Rep} = 0.1\,mm$ (mean values of 250.000 realisations).

Par.	$\tilde{x} - \hat{x}$	$\hat{\sigma}$
ω	$-1''$	$10''$
ϕ	$0''$	$11''$
κ	$0''$	$10''$
${}^{R}t_{LT,x}$	$1\,\mu m$	$32\,\mu m$
${}^{R}t_{LT,y}$	$-1\,\mu m$	$131\,\mu m$
${}^{R}t_{LT,z}$	$3\,\mu m$	$146\,\mu m$
${}^{P}t_{RF,x}$	$1\,\mu m$	$22\,\mu m$
${}^{P}t_{RF,y}$	$0\,\mu m$	$20\,\mu m$
${}^{P}t_{RF,z}$	$1\,\mu m$	$24\,\mu m$

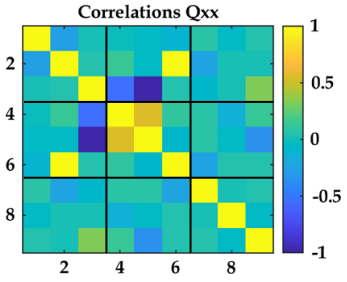

Figure 8: Correlations of the estimated parameters of simulation set 2.

ity. Thereby the uncertainty of the robot deviations (Section 3.3) is reduced by a factor of 1.5.

This observation configuration could be realised by using a mounting with three calibrated nests for corner cube reflectors (CCR) on the back of the T-probe adapter, which enables the determination of the robot pose in an opposite direction or by the method proposed in [23].

Another possibility to improve the geometric configuration is the introduction of a second laser tracker station, whereby a connection by identical robot points is not advisable due to the poor repeatability of the robot. Further investigations are needed, if the "true" robot position is estimated as an additional parameter and the identical robot point constraint $\boldsymbol{f}_2$ is adapted.

4 System identification

The assumed simplified robot arm model consists of six encoder values as inputs (Figure 3) and three determined position deviations $\boldsymbol{\delta}_R$ as outputs (Figure 9). To keep the complexity of the ANN low and due to the fact, that the zero offsets of the joints have the largest impact (>50 %) on the error budget [2], only the six joints are used as an error source.

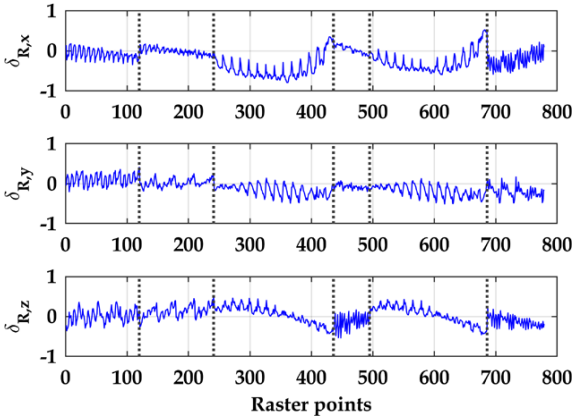

Figure 9: Robot position deviations $\boldsymbol{\delta}_R$ in the robot frame.

The generated data set bases on 779 raster points, distributed in the working field. The first four raster are used for training, which comprise 495 raster points. For the last two raster the spindle head compensation was turned off. Thus, a load influenced the positioning of the robot arm. These 284 raster points are used to test the generalization capability of the generated ANN. Thereby, it can even be better tested if the ANN already represents noise.

Before the training of the ANN, the structure of the network needs to be defined. Furthermore, the equivalence of LM and EKF is treated and the role of stochastic information in the EKF approach is evaluated in the next sections.

4.1 Model selection

The ANN structure is determined on basis of cross-validation. According to the universal approximation theorem [5], only one hidden layer is considered. The number of nodes in this hidden layer varies between three and 17 (Figure 10). For each structure 15 realisations with different weight initialisations are computed to avoid a decision on basis of a local minima. Therefore, the parameters are initialised to small random values from a zero-mean normal distribution. The EKF approach described in Section 2.3 has been applied. Thus, the compatibility test is applied after 10 iterations and if the test is rejected, then only q_{yy} is adapted. The cofactor matrix of the observations $\boldsymbol{Q}_{yy}$ is initialised as an identity matrix and is adjusted by q_{yy} ($\boldsymbol{Q}_{yy} = q_{yy}\tilde{\boldsymbol{Q}}_{yy}$). The process cofactor matrix $\boldsymbol{Q}_{pp}$ is a constant diagonal matrix ($\boldsymbol{Q}_{pp} = 0.001\boldsymbol{I}$).

Figure 10 presents the results for each ANN structure, leading to a minimal validation error. A good validation is achieved, if eight nodes are used in the hidden layer in the EKF case. In contrast, LM needs 11 nodes for the same

Brought to you by | TU Wien Bibliothek
Authenticated
Download Date | 3/26/19 2:46 PM

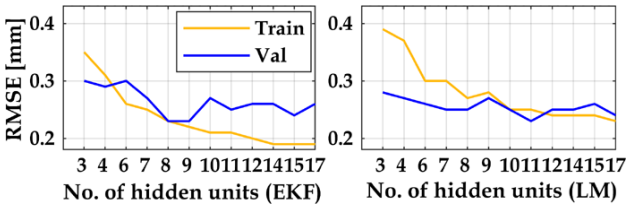

Figure 10: Model selection for EKF and LM – minimal validation error is reached with 8 and 11 hidden nodes in case of EKF and LM respectively.

level of validation error. The standard deviations of the 15 initialisations for these two model structures yield to 0.09 and 0.05 mm (EKF and LM).

A higher model complexity demands more data in order to achieve a general valid model. In case of 11 hidden nodes, 113 weights have to be estimated. These are 30 weights more than in the case of 8 hidden nodes. Consequently, much more data needs to be observed when the system identification is performed by LM estimation.

4.2 Equivalence of EKF and LM

The equivalence of LM and EKF under certain simplifications is discussed theoretically in Section 2.3. On basis of the robot data it shall be verified empirically. In Figure 11 the training process for both methods is plotted. Therefore, the EKF approach is simplified as follows: Cofactor matrices are set to identity matrices, no updating of the weight cofactor matrix is accomplished (6) and the variance factor of the observations q_{yy} is equal to the learning rate λ. The underlying ANN structure consists of 8 hidden nodes and thereon the same five weight initialisations have been computed for both methods. The curves in Figure 11 are identical in each iteration step of the training procedure. Thus, the equivalence of both approaches is confirmed empirically.

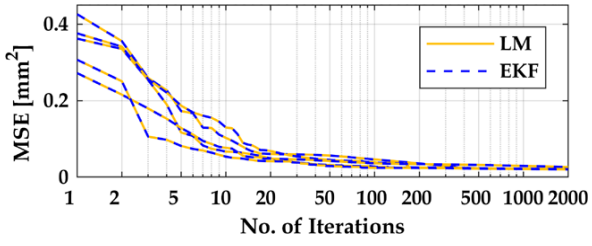

Figure 11: Output errors during training; Comparison of LM and simplified EKF approach.

4.3 Evaluating the influence of stochastic information

The estimation of the ANN in the EKF presents some advantages as stated in Section 2.3. The influence of the additional stochastic information on the weight estimation in the EKF is evaluated in this section.

Therefore, the cofactor matrices are initialised and the variance factors are defined. Thereon, the EKF versions and LM are compared and the influence of the stochastic information is discussed. In a third part, the weights are tested on their significance for the estimation.

4.3.1 Initialisation and quantification

A number of parameters need to be set for the different EKF versions. In the beginning, the cofactor matrices need to be initialised. The weight cofactor matrix $\boldsymbol{Q}_{\bar{w}\bar{w},1}$ is a diagonal matrix with components of 100 for the weights between input and hidden layer and 1000 for the weights between hidden and output layer. The process cofactor matrix $\boldsymbol{Q}_{pp,1}$ is initialised with a diagonal matrix with components of 0.001. The setting has been chosen on basis of recommendations in [6, p. 32–33].

In order to evaluate the influence of stochastic information, different initial observation cofactor matrices $\boldsymbol{Q}_{yy,1}$ are used. In Section 3.3 the cofactor matrix of the robot position deviations, corresponding to the observations in the EKF, have been computed. Three cases will be distinguished: The full cofactor matrix is introduced, labelled as $\boldsymbol{Q}_{yy,Kov}$. Only the variances are considered ($\boldsymbol{Q}_{yy}$). If nothing is specified, an identity matrix is used as observation cofactor matrix.

For all EKF versions the compatibility test is applied after 10 iterations, labelled as 'EKF-T'. If the test is rejected, the observation cofactor matrix $\boldsymbol{Q}_{yy}$ is varied by q_{yy}. If the test measure is larger than the upper quantile of the χ^2-distribution, $\widetilde{\boldsymbol{Q}}_{yy}$ is multiplied by $q_{yy} = 1.5$; if it is smaller, $\widetilde{\boldsymbol{Q}}_{yy}$ is divided by $q_{yy} = 1.5$. The same factor is used for the learning rate λ in case of LM. If the criterion of a decreasing total error in subsequent iterations is not fulfilled, λ is multiplied by 1.5 and reciprocal if it is fulfilled.

If q_{pp} is adapted additionally, it adjusts similar to q_{yy}. $\widetilde{\boldsymbol{Q}}_{pp}$ is multiplied by $q_{pp} = 1.01$, if the test measure is larger than the upper quantile of the χ^2-distribution and multiplied by $q_{pp} = 0.96$, if the test measure is smaller. The latter one is responsible for a smooth error curve.

The convergence is mainly controlled by q_{yy}, following the learning rate λ. q_{pp} is only used additionally to q_{yy}.

Brought to you by | TU Wien Bibliothek
Authenticated
Download Date | 3/26/19 2:46 PM

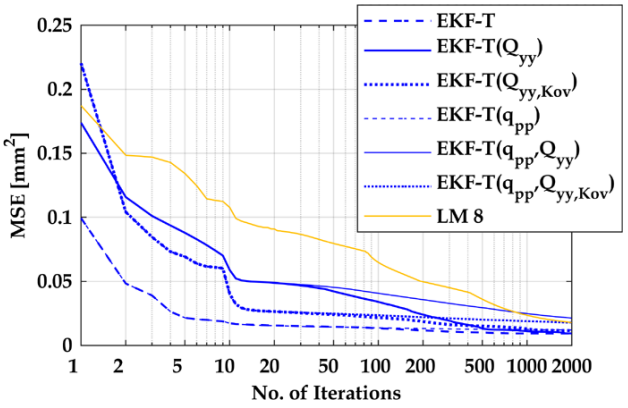

Figure 12: Mean output errors of 5 initialisations during training for the different versions.

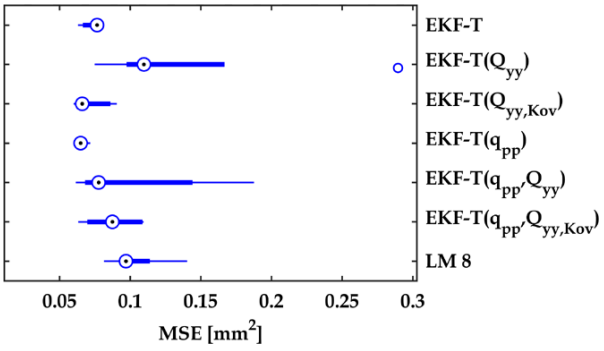

Figure 13: Results of test data for the different versions.

4.3.2 Discussion of the influence of stochastic information

One argument for the EKF algorithm is the faster achievable convergence, than with standard optimisation methods [6]. This can be confirmed on basis of the plotted mean training errors per iteration step in Figure 12. Nearly all EKF versions converge faster and to a smaller MSE than LM.

However, the criterion for choosing the best EKF version is the generalisation capability of an ANN. The evaluation is carried out on basis of a not already used data set – the test data. The MSE computed on basis of the test data corresponds to the test error. Figure 13 presents the test error of five initialisations in a box plot style. For all versions the same weight initialisations are used. The minimal MSE is reached in case of EKF including the full covariance matrix and no adaption of the process noise (EKF – T($\boldsymbol{Q}_{yy,Kov}$)). The smallest median value and interquartile range presents the version EKF – T(q_{pp}). The additional adaption of the process noise q_{pp} leads to superior results in case of EKF – T(q_{pp}) and (EKF – T(q_{pp}, $\boldsymbol{Q}_{yy}$)). It provides a smaller minimal and average test error than without adapting q_{pp}. The outlier of EKF – T($\boldsymbol{Q}_{yy}$) is now included in the whiskers, which corresponds to $\pm 2.7\sigma$ and a probability of 99.3 % and the interquartile range increases. On basis of equation (13) the effect of q_{pp} is discussed in the following. If the global test is rejected due to a smaller test measure than the lower χ^2-quantile, the observations are determined more precise than assumed. In case of the additional adaption of q_{pp}, not only $\widetilde{\boldsymbol{Q}}_{yy}$ is multiplied by 1/1.5, but also $\widetilde{\boldsymbol{Q}}_{pp}$ is multiplied by 0.96. In this case the Kalman gain $\boldsymbol{K}$ gets smaller ($K = 0.7462$) than without considering q_{pp} ($K = 0.7491$). If the test measure is larger than the upper χ^2-quantile, $\widetilde{\boldsymbol{Q}}_{yy}$ is multiplied by 1.5 and $\widetilde{\boldsymbol{Q}}_{pp}$ is multiplied by 1.01. The Kalman gain results in 0.5726 instead of 0.5714. The longer the estimation process lasts, the narrower the variation area of the Kalman Gain and the innovation covariance $\boldsymbol{Q}_{ee}$ becomes respectively. Of importance is the ratio between increasing and decreasing $\boldsymbol{Q}_{pp}$. In case of a too small test measure, q_{pp} shall lead to a larger decrease of $\boldsymbol{Q}_{pp}$ than in case of a too large test measure. It leads to a smoothed error curve and to test errors with a smaller variability.

Using the full covariance matrix $\boldsymbol{Q}_{yy,Kov}$ leads to less variance in the achieved test errors than reached by using only the variances $\boldsymbol{Q}_{yy}$. The EKF versions lead to smaller test errors and less variance of the achievable test errors than the standard optimisation by LM. If the ANN is trained on some parts of raster five or six, LM performs in the prediction mode similar to the best EKF versions. However, if the influence of the spindle head compensation is not included in the training data, LM is outperformed by some of the EKF versions.

The comparison of the test errors with the learning rate λ or the observation covariance factor q_{yy} in Figure 14, partly explains the level of the generalisation capability of the ANN. If the learning rate λ in the end of the training is proportionally small, the test error of LM is also small. The LM test error curve and the learning rate curve show a similar trend. A similar distinct tendency offers the EKF – T(q_{pp}, Q_{yy}). EKF-T and EKF – T(q_{pp}) show small observation variance factors q_{yy} at the end of training and small test errors. In case of EKF – T($\boldsymbol{Q}_{yy,Kov}$) and EKF – T(q_{pp}, $\boldsymbol{Q}_{yy,Kov}$) no distinct tendency of q_{yy} is visible. Improving $\sigma_{\delta R}$ will maybe lead to a clearer tendency in q_{yy}.

In Figure 15 the estimated output of the ANN, leading to the minimal test error is plotted. A mean square test error of 0.06 mm^2 is reached by EKF-T($\boldsymbol{Q}_{yy,Kov}$). The output of the ANN is compared to the original data, both have been shifted by +1 for a better visualisation. The errors in y- and z-direction (darkblue curves in Figure 15) of the first test

Brought to you by | TU Wien Bibliothek
Authenticated
Download Date | 3/26/19 2:46 PM

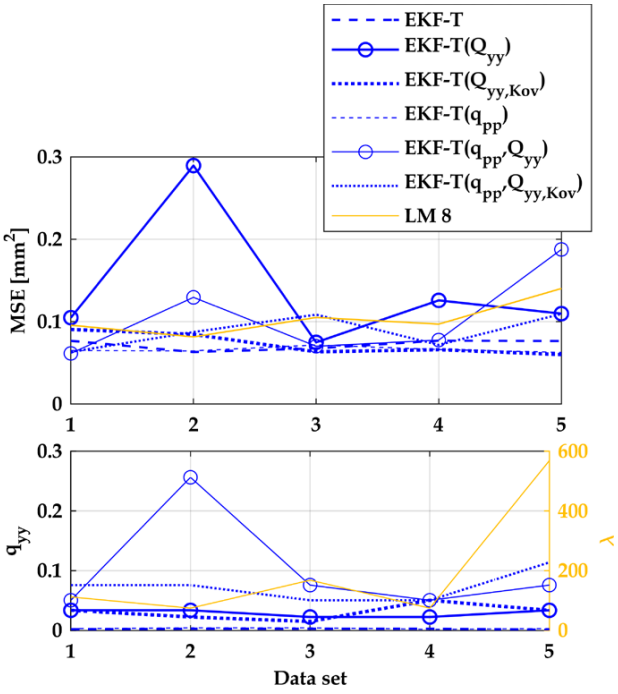

Figure 14: Test error for the different versions in comparison to the level of the learning rate λ and the observation variance factor q_{yy} at the end of training.

raster (raster five) are equivalent to the one of raster three. The difference between raster three and five is the use of the spindle head compensation, which mainly causes a deviation in x-direction. The second part of the test data represents complete new data for the ANN, which manifests also in increased errors mainly in x-direction. This is in accordance with the standard deviations $\sigma_{\delta R}$. While $\sigma_{\delta R,y}$ and $\sigma_{\delta R,z}$ amount to approximately 0.38 mm, $\sigma_{\delta R,x}$ varies between 0.45 and 0.55 mm. Table 3 as well shows the largest uncertainty in the x-direction of ${}^{R}\boldsymbol{t}_{LT}$, estimated with 0.3 mm.

4.3.3 Significance test of weights

The weight cofactor matrix $\boldsymbol{Q}_{\hat{w}\hat{w}}$ enables a testing of each weight on its significance, see Section 2.3. If they are significantly different from zero, the weights are denoted as significant for the estimation process. If weights are not significant in the end of the training, the corresponding weights are deleted and the network is trained again. This is repeated till only significant weights are detected.

The findings show, that only significant weights are obtained in case of using EKF – T(q_{pp}, $\boldsymbol{Q}_{yy}$) and nearly only significant weights in case of EKF – T(q_{pp}). The EKF version, considering q_{pp} rarely detects insignificant weights and then only few. It behaves different for EKF versions not considering q_{pp}. In three out of four cases of deleting insignificant weights an improvement of the test error can be reached. However, it is not always the last result leading to a smaller test error, but a selection in between. The largest improvement of 65 % has been achieved for the EKF – T($\boldsymbol{Q}_{yy,Kov}$). The MSE of the test data reduced to 0.06 mm^2, while the training error increased. It is estimated about 0.32 mm better, which corresponds to 1.5 times the repeatability of the robot arm. In Figure 16 the network is pictured, whereby the line width accords to the magnitude of the specific weight. Dashed lines correspond to eliminated weights. Six insignificant weights were removed in order to reach a better generalisation capability of the ANN.

It is not possible to identify systematics in the eliminated weights. The significance test of the weights is a tool to improve the estimation, but mainly if q_{pp} is not considered.

Summarising this subsection, the equivalence of LM and the simplified EKF as stated in Section 2.3 is confirmed empirically. The consideration of the full covariance matrix of the output data leads to superior results. The minimal test error has been achieved by EKF – T($\boldsymbol{Q}_{yy,Kov}$). The additional adaption of the process noise q_{pp} improves the test errors for this data set. The weight significance test leads to superior results. The main advantage is, that the number of weights can be reduced and in further consequence less data is needed for the training of the ANN.

5 Conclusion

The aim of this paper is to build a correction tool based on the learning method consisting of ANN and EKF. In order to model the robot position deviations as a function of the joint encoder values, the determination of transformation and lever arm parameters is required. An integrated estimation of both measures is introduced on basis of transformation points realised by the robot arm. It shows, that an unbiased estimation of the parameters is given and the robot's arm repeatability limits the achievable parameter uncertainty of this approach. In case of the test object, the estimated repeatability exceeds the manufacturers specifications by almost one order of magnitude, the reason may be the advanced age of the robot arm. Hence, the average uncertainty of the robot deviations only amounts to 0.41 mm. Therefore, possibilities to improve the parameter uncertainties are verified. On the one hand, the extension of the functional model by an identical point con-

Brought to you by | TU Wien Bibliothek
Authenticated
Download Date | 3/26/19 2:46 PM

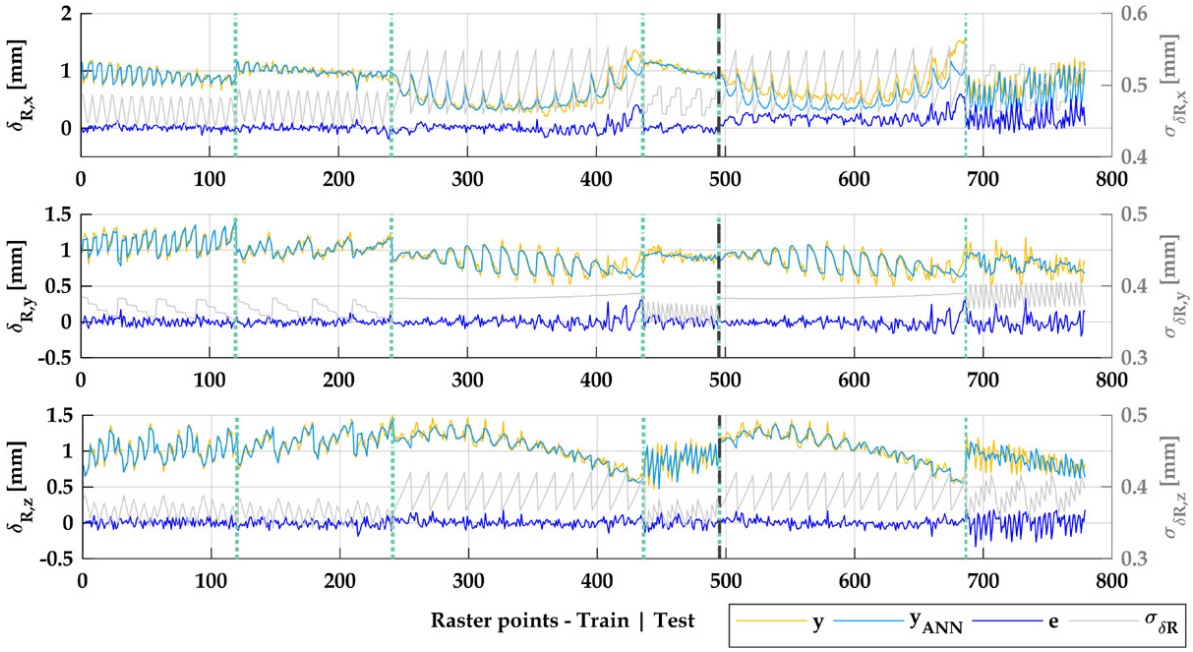

Figure 15: Minimal test error reached by EKF – T($\boldsymbol{Q}_{yy,Kov}$), whereby y and y_{ANN} are shifted by +1. The right axes show the standard deviations of the robot deviations, which have been considered in the training.

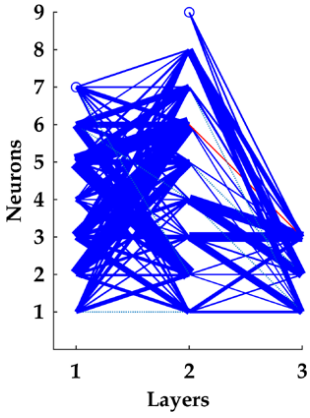

Figure 16: ANN leading to a minimal test error after deletion of the insignificant weights (dashed lines) of EKF – T($\boldsymbol{Q}_{yy,Kov}$); Line width corresponds to the weight value and the red highlighted weight is a further detected insignificant weight.

straint does not show the expected effect. The reason can be found in the insufficient description of the robot's repeatability. On the other hand, the simulation results indicate that a decorrelation of transformation parameters and lever arm reduces the robot deviations uncertainty $\boldsymbol{\sigma}_{\delta R}$ by 36 %.

On basis of the reached data set, the ANN was built up and trained in the integrated approach. The model selection already shows, that for the estimation in the EKF 30 weights less are needed than for LM. This manifests in reduced measurement costs. The equivalence of LM and a simplified EKF is confirmed theoretically as well as empirically on basis of the robot arm model. As stated by [6], a faster convergence is achieved by different EKF versions than by LM. However, the main advantage of estimating the ANN in the EKF algorithm is the additional consideration of the ANN's output uncertainty. Its effectiveness can be confirmed – considering uncertainty has a positive effect on the prediction capability of the ANN. The full covariance matrix of the observations in the integrated estimation leads to the minimal test error. The additional adaption of q_{pp} leads to smaller test errors and a lower variance of the test errors for this data set. EKF outperforms LM in case of testing on data, where a small load like the turned off spindle head compensation has been considered. Moreover, a first investigation on eliminating insignificant weights has been accomplished. The elimination improves the prediction capability of the ANN up to 65 %. However, it depends on the weight initialisation. In the moment no systematics can be identified. Further investigations shall figure out, if the weight elimination is somehow transferable to find superior model structures in the model selection task.

The number of labelled data pairs has been small dimensioned for this study. Future enhancements must include the generation of a larger data set and a method comparison on basis of a larger number of initialisations to reach more significant results.

Brought to you by | TU Wien Bibliothek
Authenticated
Download Date | 3/26/19 2:46 PM

References

[1] L. Alkaiem and H. Sternberg, Analysis of inclination measurement by means of artificial neural networks – A comparison of static and dynamic networks, in: *3rd Joint International Symposium on Deformation Monitoring*, Vienna, 2016.

[2] L. Beyer, *Genauigkeitssteigerung von Industrierobotern: insbesondere mit Parallelkinematik*, Ph. D. thesis, Shaker, 2005.

[3] V. Cherkassky and F. Mulier, *Learning from data: concepts, theory, and methods*, Wiley-Interscience, 1998.

[4] J. Dupuis, C. Holst and H. Kuhlmann, Improving the kinematic calibration of a coordinate measuring arm using configuration analysis, *Precision Engineering* 50 (2017), 171–182 (en).

[5] S. S. Haykin, *Neural Networks – A Comprehensive Foundation*, 2nd edition ed, Prentice Hall, 1999 (English).

[6] S. S. Haykin, *Kalman filtering and neural networks*, Wiley, New York, 2001 (English).

[7] M. Heinert and W. Niemeier, From fully automated observations to a neural network model inference: The Bridge Fallersleben Gate in Brunswick, Germany 1999–2006, *Journal of Applied Geodesy* 1 (2007).

[8] C. Herrmann, M. Hennes, M. Juretzko, C. Munzinger and M. Schneider, Positioning and synchronization of industrial robots, in: *International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pp. 15–17, 2010.

[9] O. Heunecke, H. Kuhlmann, W. Welsch, A. Eichhorn and H. Neuner, *Handbuch Ingenieurgeodäsie: Auswertung geodätischer Überwachungsmessungen*, 2., neu bearbeitete und erweiterte auflage ed, Wichmann Verlag, 2013.

[10] J. M. Hollerbach and C. W. Wampler, The calibration index and taxonomy for robot kinematic calibration methods, *The international journal of robotics research* 15 (1996), 573–591.

[11] International Organization for Standardization, *ISO 9283:1998 Manipulating industrial robots – Performance criteria and related test methods*, 1998.

[12] K.-R. Koch, Robust estimations for the nonlinear Gauss Helmert model by the expectation maximization algorithm, *Journal of Geodesy* 88 (2014), 263–271 (en).

[13] R. Laufer, *Prozedurale Qualitätsmodellierung und -management für Daten: ingenieurgeodätische verkehrstechnische Anwendungen*., Verlag der Bayerischen Akademie der Wissenschaften, München, 2011 (English).

[14] D. W. Marquardt, An Algorithm for Least-Squares Estimation of Nonlinear Parameters, *SIAM Journal on Applied Mathematics* 11 (1963), 431–441.

[15] H. Neuner, Model selection for system identification by means of artificial neural networks, *Journal of Applied Geodesy* 6 (2012), 117–124.

[16] H.-N. Nguyen, J. Zhou and H.-J. Kang, A calibration method for enhancing robot accuracy through integration of an extended Kalman filter algorithm and an artificial neural network, *Neurocomputing* 151 (2015), 996–1005 (en).

[17] A. Nubiola and I. A. Bonev, Absolute calibration of an ABB IRB 1600 robot using a laser tracker, *Robotics and Computer-Integrated Manufacturing* 29 (2013), 236–245 (en).

[18] A. Pope, Some Pitfalls to be avoided in the iterative adjustment of nonlinear problems, in: *Proceedings of the 38th Annual Meeting American Society of Photogrammetry*, Washingthon D.C., 1972.

[19] G. V. Puskorius and L. A. Feldkamp, Decoupled extended Kalman filter training of feedforward layered networks, in: *Neural Networks, 1991., IJCNN-91-Seattle International Joint Conference on*, 1, pp. 771–777, IEEE, 1991.

[20] A. Reiterer, U. Egly, T. Vicovac, E. Mai, S. Moafipoor, D. A. Grejner-Brzezinska and C. K. Toth, Application of artificial intelligence in Geodesy – A review of theoretical foundations and practical examples, *Journal of Applied Geodesy* 4 (2010).

[21] B. Siciliano and O. Khatib (eds.), *Springer handbook of robotics*, Springer, Berlin, 2008 (eng).

[22] S. Singhal and L. Wu, Training multilayer perceptrons with the extended Kalman algorithm, in: *Advances in neural information processing systems*, pp. 133–140, 1989.

[23] Y. Wu, A. Klimchik, S. Caro, B. Furet and A. Pashkevich, Geometric calibration of industrial robots using enhanced partial pose measurements and design of experiments, *Robotics and Computer-Integrated Manufacturing* 35 (2015), 151–168 (en).

[24] X. Zhong, J. Lewis and F. L. N-Nagy, Inverse robot calibration using artificial neural networks, *Engineering Applications of Artificial Intelligence* 9 (1996), 83–93 (en).

Brought to you by | TU Wien Bibliothek
Authenticated
Download Date | 3/26/19 2:46 PM