Article

The International Journal of Robotics Research
2018, Vol. 37(10) 1152–1167

DOI: 10.1177/0278364918760698
journals.sagepub.com/home/ijr

Recursive decentralized localization for multi-robot systems with asynchronous pairwise communication

Lukas Luft[1], Tobias Schubert[1], Stergios I. Roumeliotis[2] and Wolfram Burgard[1]

Abstract

This paper provides a fully decentralized algorithm for collaborative localization based on the extended Kalman filter. The major challenge in decentralized collaborative localization is to track inter-robot dependencies, which is particularly difficult when sustained synchronous communication between the robots cannot be guaranteed. Current approaches suffer from the need for particular communication schemes, extensive bookkeeping of measurements, overly conservative assumptions, or the restriction to specific measurement models. This paper introduces a localization algorithm that is able to approximate the inter-robot correlations while fulfilling all of the following conditions: communication is limited to two robots that obtain a relative measurement, the algorithm is recursive in the sense that it does not require storage of measurements and each robot maintains only the latest estimate of its own pose, and it supports generic measurement models. The fact that the proposed approach can handle these particularly difficult conditions ensures that it is applicable to a wide range of multi-robot scenarios. We provide mathematical details on our approximation. Extensive experiments carried out using real-world datasets demonstrate the improved performance of our method compared with several existing approaches.

Keywords

Collaborative localization, cross-correlation, extended Kalman filter, over-confidence, multi-robot systems

1. Introduction

Localization is one of the most fundamental tasks for mobile robots. For a team of multiple robots, a paradigm called *collaborative localization* significantly improves the localization performance of the individual teammates, see Fox et al. (2000) and Roumeliotis and Bekey (2002).

In collaborative localization, robots detect each other and communicate their pose estimates, which correlates the estimates of their individual poses. It is of fundamental importance to keep track of these dependencies to avoid the problem of *double-counting* or *data incest*, which occurs when two robots treat shared information as uncorrelated. Double-counting puts the localization reliability at risk as it leads to overly optimistic estimates.

Collaborative localization can be formulated as the task of fusing information gathered from proprioceptive (e.g., wheel encoders) and exteroceptive sensors (e.g., cameras or laser scanners) to calculate the posterior of the joint pose of all involved robots. A popular solution to this problem is the *Bayes filter*. However, it is impossible to implement the Bayes filter for arbitrary distributions of continuous random variables as these distributions can possess infinitely many dimensions. Therefore, researchers have used approximations of the Bayes filter and related information fusing techniques. Fox et al. (2000), Prorok et al. (2012), and Howard et al. (2003) used particle filters, Nerurkar et al. (2009) employed a maximum *a posteriori* algorithm, and Howard et al. (2002) utilized a maximum-likelihood algorithm.

In addition to the already-mentioned algorithms, another popular approximation of the Bayes filter is the *extended Kalman filter* (EKF), which we leverage in the proposed approach. Under the assumption of Gaussian noise for motions and measurements, applying the EKF to the joint system state space containing the poses of all team members yields the optimal estimator (apart from linearization errors, see Martinelli and Siegwart (2005) and Huang et al. (2011)).

[1]University of Freiburg, Department of Computer Science, Freiburg, Germany
[2]University of Minnesota, Department of Computer Science and Engineering, Minneapolis, USA

Corresponding author:
Lukas Luft, Autonomous Intelligent Systems, Faculty of Engineering, Georges-Köhler-Allee 080, 79110 Freiburg, Germany.
Email: luft@cs.uni-freiburg.de

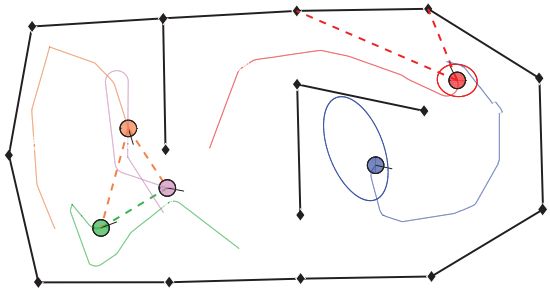

Fig. 1. A five robot localization problem in an indoor environment according to the publicly available data set recorded by Leung et al. (2011). The three robots on the left observe each other, while the rightmost robot observes two landmarks. Accordingly, its error ellipse is small. The fifth (blue) robot has a large error covariance as it did not detect any landmarks, nor was it able to improve its pose estimate with the help of pairwise observations in the most recent steps.

The EKF naturally tracks inter-robot correlations in terms of the off-diagonal blocks of the covariance matrix. This prevents double-counting. The disadvantage lies in the fact that it requires a central processing unit for updating the joint filter. In addition, all robots need to instantaneously send all their motion and measurement data to this unit introducing a substantial communication overhead.

However, in many relevant multi-robot applications, for example underwater, in mines, or in large-scale environments, communication might be energy consuming, error-prone, slow, or simply not available at all times. Therefore, decentralized architectures that reduce the need for communication to a minimum are desirable. Particularly appealing are algorithms restricting exchange of information between two robots to points in time where they actually meet. Establishing the correct joint estimate in such a context requires storage and bookkeeping of measurements, as in Leung et al. (2010). If the memory needed to store measurements is limited and the robots are only able to communicate at rare occasions, it is impossible to reproduce the centralized EKF. In particular, the covariance matrix of the joint system can only be approximated. Here, the double-counting problem can occur if the off-diagonal terms of the matrix are underestimated or neglected completely. Accordingly, in such a situation the estimates of the poses cannot be guaranteed to be consistent.

Our contribution is an EKF-based localization algorithm that is able to approximate the covariance under the above-mentioned constraints on communication and memory.

- Exchange of information takes place only between two robots obtaining a relative measurement. Our approach does not require communication with further teammates. The communication complexity[1] is thus $\mathcal{O}(1)$ per relative measurement.
- The algorithm is recursive in the sense that no measurements have to be stored. For a team of N robots, each robot only updates the belief about its own pose and $N-1$ matrices that carry information about the correlations with the estimates of the teammates.
- In addition to relative observations, the algorithm also supports absolute positioning of individual robots. It furthermore allows for generic measurement models.

This paper is a revised and substantially extended version of our previous conference publication Luft et al. (2016). Compared with this prior publication, the present paper introduces more comprehensive theoretical elaborations. First, we prove that all introduced approximations maintain well-behaved joint distributions of all two-robot subsystems. Second, we show that the corresponding approximation errors are explicitly bounded by the inter-robot correlations. In addition, we introduce an alternative update strategy (28) and present augmented experimental evaluations including a paired sample t-test, an evaluation of the ANEES (defined below) as additional metric, and comparisons with two additional approaches: the new update strategy (28) and the Schmidt–Kalman filter as introduced by Schmidt (1966). While the former is decentralized as the proposed approach, the latter relies on all-to-all communication to avoid the approximation introduced in Equation (26), which is essential for the proposed approach.

2. Related work

The problem of multi-robot localization has received considerable attention in the past. Roumeliotis and Bekey (2002) presented a distributed algorithm that is equivalent to the joint EKF. In their approach, each robot processes its own odometry and measurements are communicated to all teammates. For a team of N robots, this results in communication complexity of $\mathcal{O}(N)$ per robot–robot perception. In contrast, our approach is of communication complexity $\mathcal{O}(1)$ and is applicable even if the exchange of information is limited to robots performing a relative measurement. Similar communication networks without centralized entity are used in gossip protocols (see, e.g., Boyd et al., 2006). As opposed to our method, all centralized-equivalent approaches share the need for either a particular structure of the communication graph, similarly to Bailey et al. (2011), Kia et al. (2014), Martinelli (2007), and Wanasinghe et al. (2015), or the storage of measurement. Examples for the latter are Leung et al. (2010) for localization and Leung et al. (2012) for simultaneous localization and mapping (SLAM).

In the following, we concentrate on EKF-based collaborative localization algorithms that meet our assumptions: namely, algorithms that are not dependent on the storage of measurement information and rely only on pairwise communication in the context of a relative measurement. There are approaches in which the robots treat the poses of teammates as deterministic parameters and cross-correlations are simply neglected, as in Panzieri et al. (2006). This entails the risk of becoming over-confident, or forces the

robots to follow certain motion patterns, as in Kurazume et al. (1994).

A more elaborate approach is to treat incorporation of relative measurements as fusion of estimates with unknown correlations. To avoid over-confidence, *covariance intersection* (CI) or *split covariance intersection* (SCI) techniques constitute a natural choice. Carrillo-Arce et al. (2013) presented an approach based on CI, which treats estimates as if they were maximally correlated. Li and Nashashibi (2013) and Wanasinghe et al. (2014) presented approaches based on SCI, which splits the covariances into a dependent and an independent component. The fusion of consistent estimates by CI or SCI provably preserves consistency. However, as a price, they are in general overly pessimistic. Moreover, in contrast to our approach, the implementations as presented in the literature have to assume perception systems capable of identifying the relative poses of neighboring robots. More precisely, they require relative range and orientation measurements and cannot be applied to systems that, for example, only provide relative range or bearing measurements. For details on different measurement models, we refer the reader to the work of Martinelli et al. (2005).

Karam et al. (2006) presented an approach based on state exchange. It only allows independent states to be fused and is thus not optimal. Moreover, it requires relative pose detection. Unlike the already-mentioned approaches, the interleaved update algorithm by Bahr et al. (2009) can handle general measurement models. In this approach, each vehicle of a team of N robots has to maintain and update up to 2^N estimates and covariances and has to keep track of their dependencies with all the team members. Robot i can only use those estimates of robot j that are not (directly or indirectly via other robots) correlated to robot i. Compared with this algorithm, our method has the advantage that each robot has to maintain only one estimate and N matrices (as opposed to 2^N estimates and matrices). Moreover, our algorithm is recursive in the sense that only the most recent estimates are used, which separates it from the approaches of Bahr et al. (2009) and Karam et al. (2006).

In addition to the mentioned recursive localization methods related to the paper at hand, there is a lot of work on multi-robot localization and SLAM, which estimate the whole trajectories. As examples, see the works of Walls and Eustice (2013), Cunningham et al. (2012), Cunningham et al. (2013), Indelman et al. (2012), and Choudhary et al. (2016). A popular approach to optimize robot poses and landmark positions uses factor graphs. These graphs maintain measurement information over the whole robot trajectories, such that the latter can be optimized. To avoid double-counting, Cunningham et al. (2012) maintained for each robot a map derived from its local measurements and a map merged from neighboring robot's measurements. The authors proposed a method to solve the data association between landmarks observed from different robots. To avoid the overly conservative separation between local and neighborhood map, Cunningham et al. (2013) extended this work by introducing so-called anti-factors as a tool to avoid double-counting. Indelman et al. (2012) introduce another approach to multi-robot localization, where the cross-correlations can be calculated explicitly from the graph of all previous measurement updates. These graph-based methods have the advantage that they optimize over the full robot trajectories given the history of measurements. As a drawback they have to store these measurements, as opposed to our recursive method.

3. The proposed approach

In a team of N robots, let X_i for $i \in \{1, \dots, N\}$ be the pose of robot i in a common fixed reference frame. Throughout this paper, we consider the problem of localizing N vehicles navigating on a plane, although our approach is not limited to this case. Under this assumption, a robot pose is a triple $X_i = [x_i, y_i, \theta_i]^{\mathrm{T}}$ with Cartesian coordinates $[x_i, y_i]$ and orientation θ_i. The number N can change over time and does not need to be known to the robots beforehand. Each robot is provided with an estimate $\hat{X}_i^t$ of its own pose and the corresponding covariance matrix Σ_{ii}^t at each time step t. We call $\mathrm{bel}_i^t = \{\hat{X}_i^t, \Sigma_{ii}^t\}$ the *belief* of robot i at time t. The belief of the joint system is

$$Bel^t = \{\hat{X}^t, \Sigma^t\} \tag{1}$$

with $\hat{X}^t = [\hat{X}_1^t; \dots; \hat{X}_N^t]$ and $\Sigma^t = \left[\Sigma_{ij}^t\right]_{1 \le i,j \le N}$.

At the beginning of their mission the robots might be uncorrelated, which means $\Sigma_{ij}^t = 0$ for all $i \neq j$, and the individual beliefs are sufficient to represent Bel^t. When robot i detects robot j at time t, we have in general $\Sigma_{ij}^{t+1} \neq 0$. Inspired by Roumeliotis and Bekey (2002), we decompose

$$\Sigma_{ij}^{t+1} = \sigma_{ij}^{t+1} \left(\sigma_{ji}^{t+1}\right)^{\mathrm{T}}$$

where we can choose any possible decomposition. In Algorithm 3 given below, for example, we set $\sigma_{ij}^{t+1} = \Sigma_{ij}^{t+1}$ and $\sigma_{ji}^{t+1} = \mathbb{I}$. In the following, robots i and j independently update σ_{ij}^{t+1} and σ_{ji}^{t+1}, respectively. We present the details in Sections 3.1 to 3.3.

We assume the belief (1) to be distributed among the team such that each robot i carries bel_i^t and $\{\sigma_{ij}^t\}_{j \in \{1,\dots,N\} \setminus \{i\}}$. During their mission, the robots perform the following.

- **Motion** (Algorithm 1): Each robot has access to a control u, which is a velocity command or odometry information. We allow for generic motion models.
- **Private measurements** (Algorithm 2): Any of the robots might or might not perform measurements with output z that are not shared with the group, e.g., detection of known landmarks or GPS. The algorithm allows for a heterogeneous robot team and generic measurement models. In our experiments, the robots conduct range and bearing measurements to a set of previously known, distinguishable landmarks.

Algorithm 1 Motion Update for Robot i

Input: $\hat{X}_i^t,\ \Sigma_{ii}^t,\ \{\sigma_{ij}^t\}_{j\in\{1,\dots,N\}\setminus\{i\}},\ u$
Output: $\hat{X}_i^{t+1},\ \Sigma_{ii}^{t+1},\ \{\sigma_{ij}^{t+1}\}_{j\in\{1,\dots,N\}\setminus\{i\}}$

$G = \frac{\partial g(X,u)}{\partial X}(\hat{X}_i^t, u)$ ▷ motion model g is known
$\hat{X}_i^{t+1} = g\left(\hat{X}_i^t, u\right)$
$\Sigma_{ii}^{t+1} = G\,\Sigma_{ii}^t\,G^{\mathrm{T}} + R$ ▷ noise covariance R is known
for $1 \le j \le N, j \neq i$ **do**
    $\sigma_{ij}^{t+1} = G\,\sigma_{ij}^t$
end for

Algorithm 2 Private Measurement Update for Robot i

Input: $\hat{X}_i^t,\ \Sigma_{ii}^t,\ \{\sigma_{ij}^t\}_{j\in\{1,\dots,N\}\setminus\{i\}},\ z$
Output: $\hat{X}_i^{t+1},\ \Sigma_{ii}^{t+1},\ \{\sigma_{ij}^{t+1}\}_{j\in\{1,\dots,N\}\setminus\{i\}}$

$H = \frac{\partial h(X)}{\partial X}(\hat{X}_i^t)$ ▷ measurement model h is known
$K_i = \Sigma_{ii}^t H^{\mathrm{T}}\left(H\Sigma_{ii}^t H^{\mathrm{T}} + Q\right)^{-1}$ ▷ noise covariance Q is known
$\hat{X}_i^{t+1} = \hat{X}_i^t + K_i\left[z - h(\hat{X}_i^t)\right]$
$\Sigma_{ii}^{t+1} = (\mathbb{I} - K_i H)\,\Sigma_{ii}^t$
for $1 \le j \le N, j \neq i$ **do**
    $\sigma_{ij}^{t+1} = (\mathbb{I} - K_i H)\,\sigma_{ij}^t$
end for

Algorithm 3 Relative Measurement Update for Robot i

Input: $\hat{X}_i^t,\ \Sigma_{ii}^t,\ \{\sigma_{ij}^t\}_{j\in\{1,\dots,N\}\setminus\{i\}},\ r$
Output: $\hat{X}_i^{t+1},\ \Sigma_{ii}^{t+1},\ \{\sigma_{ij}^{t+1}\}_{j\in\{1,\dots,N\}\setminus\{i\}}$

// robot i detects robot j

if robot i conducts relative measurement r to robot j **then**
    receive from robot j: $\hat{X}_j^t,\ \Sigma_{jj}^t,\ \sigma_{ji}^t$
    $\Sigma_{ij}^t = \sigma_{ij}^t\left(\sigma_{ji}^t\right)^{\mathrm{T}}$
    $\Sigma_{aa}^t = \begin{bmatrix} \Sigma_{ii}^t & \Sigma_{ij}^t \\ \left(\Sigma_{ij}^t\right)^{\mathrm{T}} & \Sigma_{jj}^t \end{bmatrix}$
    $F_a = \left[\frac{\partial f(X_i,X_j)}{\partial X_i}(\hat{X}_i^t,\hat{X}_j^t),\ \frac{\partial f(X_i,X_j)}{\partial X_j}(\hat{X}_i^t,\hat{X}_j^t)\right]$
    $K_a = \Sigma_{aa}^t F_a^{\mathrm{T}}(F_a\Sigma_{aa}^t F_a^{\mathrm{T}} + Q)^{-1}$ ▷ Q is known
    $\begin{bmatrix} \hat{X}_i^{t+1} \\ \hat{X}_j^{t+1} \end{bmatrix} = \begin{bmatrix} \hat{X}_i^t \\ \hat{X}_j^t \end{bmatrix} + K_a\left[r - f(\hat{X}_i^t,\hat{X}_j^t)\right]$
    $\begin{bmatrix} \Sigma_{ii}^{t+1} & \Sigma_{ij}^{t+1} \\ \Sigma_{ji}^{t+1} & \Sigma_{jj}^{t+1} \end{bmatrix} = (\mathbb{I} - K_a F)\,\Sigma_{aa}^t$
    send to robot j: $\hat{X}_j^{t+1},\ \Sigma_{jj}^{t+1}$
    $\sigma_{ij}^{t+1} = \Sigma_{ij}^{t+1}$
    for $1 \le k \le N, k \notin \{i,j\}$ **do**
        $\sigma_{ik}^{t+1} = \Sigma_{ii}^{t+1}\left(\Sigma_{ii}^t\right)^{-1}\sigma_{ik}^t$
    end for
end if

// robot i is detected by robot j

if robot i is detected by robot j **then**
    send to Robot j: $\hat{X}_i^t,\ \Sigma_{ii}^t,\ \sigma_{ij}^t$
    receive from robot j: $\hat{X}_i^{t+1},\ \Sigma_{ii}^{t+1}$
    $\sigma_{ij}^{t+1} = \mathbb{I}$
    for $1 \le k \le N, k \notin \{i,j\}$ **do**
        $\sigma_{ik}^{t+1} \leftarrow \Sigma_{ii}^{t+1}\left(\Sigma_{ii}^t\right)^{-1}\sigma_{ik}^t$
    end for
end if

- **Relative measurements** (Algorithm 3): Any of the robots might or might not perform measurements with output r that includes itself and a teammate and depends on the two robot's current poses only. Again, we allow for varying and generic detection models. In particular, we do not demand access to relative poses.
- **Communication**: Whenever a relative measurement takes place, the two involved robots communicate. This is the only occasion where communication takes place. When robot i detects robot j, both share their individual beliefs bel_i, bel_j and the terms σ_{ij}, σ_{ji}. We do not require communication with further teammates. This assumes that robot i knows that he detects robot j.

Algorithms 1–3 provide an overview of how the robots perform the corresponding updates. We derive the underlying update schemes in Sections 3.1–3.3.

3.1. Motion update

We assume the motion of different robots to be independent and corrupted by Gaussian noise ν_u

$$X_i^{t+1} = g\left(X_i^t, u + \nu_u\right)$$

with differentiable motion model g, where u can either be a velocity command or an odometry measurement. If robot i performs a motion, the correct prediction step for the joint system can be derived from the standard EKF and is given by

$$\hat{X}_i^{t+1} = g\left(\hat{X}_i^t, u\right) \tag{2}$$
$$\Sigma_{ii}^{t+1} = G\Sigma_{ii}^t G^{\mathrm{T}} + R \tag{3}$$
$$\Sigma_{ij}^{t+1} = G\Sigma_{ij}^t \tag{4}$$
$$\hat{X}_j^{t+1} = \hat{X}_j^t \tag{5}$$
$$\Sigma_{jj}^{t+1} = \Sigma_{jj}^t \tag{6}$$

for all $j \neq i$ with linearization $G = \frac{\partial g(X,u)}{\partial X}(\hat{X}_i^t, u)$ and positive-definite noise covariance R. These equations correspond to (24) in Roumeliotis and Bekey (2002). In the case of odometry measurements with noise distributed as $\nu_u \sim \mathcal{N}(0; M)$, for example, we have $R = VMV^{\mathrm{T}}$, with

$V = \frac{\partial g(X,u)}{\partial u}(\hat{X}_i^t, u)$. The crucial insight here, as already pointed out in Roumeliotis and Bekey (2002), is that robot i can correctly update the cross-correlations to all teammates by updating

$$\sigma_{ij}^{t+1} = G\sigma_{ij}^t$$

The reason is that in case the robots i and j meet, they can reproduce their cross-correlation by the matrix multiplication

$$\sigma_{ij}^{t+1}\left(\sigma_{ji}^{t+1}\right)^{\mathrm{T}} = G\sigma_{ij}^t\left(\sigma_{ji}^t\right)^{\mathrm{T}} = \Sigma_{ij}^{t+1}$$

Please note that G and R are not constant. The important fact is that robot i can access these quantities. Algorithm 1 is the realization of the exact update equations (2)–(6).

3.2. Private measurement update

We define a private measurement to be a function of the observing robot's pose, corrupted by Gaussian noise

$$z = h\left(X_i^t\right) + \nu_p$$

with $\nu_p \sim \mathcal{N}(0, Q)$, positive-definite noise covariance Q, and differentiable measurement model h. Thus, the measurement probability is conditionally independent of all other variables given the observing robot's current pose. In our experiments, the robots perform range and bearing measurements to a set of known distinguishable landmarks. In the case of a range measurement to a landmark with position P, e.g. we have

$$h\left(X_i^t\right) = \|[x_i^t, y_i^t]^{\mathrm{T}} - P\|_2$$

with the Euclidean norm $\|\cdot\|_2$.

In the following, we derive the equations used in our approach for the private measurement update. The derivations prove that this component of our method is identical to the Schmidt–Kalman filter as introduced by Schmidt (1966). Assume robot i receives a private measurement and $j \in \{1, \ldots, N\} \setminus \{i\}$ indicate a remaining robot. Then, as we show in Appendix A, the exact correction step for the joint system is

$$\hat{X}_i^{t+1} = \hat{X}_i^t + K_i\left[z - h(\hat{X}_i^t)\right] \tag{7}$$

$$\Sigma_{ii}^{t+1} = (\mathbb{I} - K_i H)\,\Sigma_{ii}^t \tag{8}$$

$$\Sigma_{ij}^{t+1} = (\mathbb{I} - K_i H)\,\Sigma_{ij}^t \tag{9}$$

$$\hat{X}_j^{t+1} = \hat{X}_j^t + K_j\left[z - h(\hat{X}_i^t)\right] \tag{10}$$

$$\Sigma_{jj}^{t+1} = \Sigma_{jj}^t - K_j H \Sigma_{ij}^t \tag{11}$$

with the linearization $H = \frac{\partial h(X)}{\partial X}(\hat{X}_i^t)$ and the blocks

$$K_i = \Sigma_{ii}^t H^{\mathrm{T}} S^{-1}$$
$$K_j = \Sigma_{ji}^t H^{\mathrm{T}} S^{-1}$$

of the Kalman gain with the invertible residual covariance $S = H\Sigma_{ii}^t H^{\mathrm{T}} + Q$. We can execute the correction updates (7)–(9) without communication because all necessary terms are known to robot i.

The cross-correlation update (9) is realized by the matrix multiplication

$$\sigma_{ij}^{t+1} = (\mathbb{I} - K_i H)\,\sigma_{ij}^t$$

similar to the prediction step in Section 3.1.

To avoid communication, we approximate updates (10) and (11) to leave the corresponding beliefs unchanged, i.e.

$$\hat{X}_j^{t+1} \approx \tilde{X}_j^{t+1} := \hat{X}_j^t \tag{12}$$

$$\Sigma_{jj}^{t+1} \approx \tilde{\Sigma}_{jj}^{t+1} := \Sigma_{jj}^t \tag{13}$$

In summary, we arrive at a set of equations equivalent to the Schmidt-Kalman filter, see Schmidt (1966).[2] Algorithm 2 is the realization of the update equations (7)–(9), (12), and (13).

In expectation, this approximation is consistent with respect to the joint system. This is due to the following two observations: given the correct belief Bel^t, we have the expectation $E_z\left(\hat{X}_j^{t+1}\right) = \hat{X}_j^t$; the difference of the approximated $\tilde{\Sigma}^{t+1}$ and the exact covariance Σ^{t+1} is a positive-semidefinite matrix, i.e.,

$$\tilde{\Sigma}^{t+1} - \Sigma^{t+1} = \begin{pmatrix} 0 & 0 \\ 0 & \Sigma_{ji}^t H^{\mathrm{T}} S^{-1} H \Sigma_{ij}^t \end{pmatrix} \succeq 0$$

3.3. Relative measurements

We define a relative measurement r, most generally, as a noisy observation depending on the state of two robots, i.e.,

$$r = f\left(X_i^t, X_j^t\right) + \nu_r$$

with differentiable measurement model f. To meet the conditions for applying the EKF, we assume the noise term to be normally distributed $\nu_r \sim \mathcal{N}(0; Q)$. Thus, the measurement probability is conditionally independent of all other variables given the two observing robot's current poses. For ease of notation we use the same noise covariance as for the private measurement model. Obviously, we can also model the relative measurement with a different covariance matrix. In our experiments, the robots perform range and bearing measurements to each other. In the case of a range measurement between robots i and j, for example, we have

$$f\left(X_i^t, X_j^t\right) = \|[x_j^t, y_j^t]^{\mathrm{T}} - [x_i^t, y_i^t]^{\mathrm{T}}\|_2$$

We emphasize again that we neither require the relative measurement to be a function of pose differences of two robots $f\left(X_i^t, X_j^t\right) = f\left(X_i^t - X_j^t\right)$ nor do we assume that the relative measurement provides the observing robot

with an estimate of the observed robot's pose. Both are special cases of our general model. In particular, we can handle the practically relevant case of range and/or bearing measurements.

Assume that a robot obtains a relative measurement, without loss of generality, i observes j. Define now $a = \{i,j\}$ and let $b = \{1,\dots,N\} \setminus \{i,j\}$ be the indices of the $N-2$ robots that do not participate at the relative measurement. The exact correction step for the joint system can be written as

$$\hat{X}_a^{t+1} = \hat{X}_a^t + K_a \left[r - f(\hat{X}_a^t) \right] \tag{14}$$

$$\Sigma_{aa}^{t+1} = (\mathbb{I} - K_a F_a)\, \Sigma_{aa}^t \tag{15}$$

$$\begin{aligned} \Sigma_{ab}^{t+1} &= (\mathbb{I} - K_a F_a)\, \Sigma_{ab}^t \\ &\overset{(15)}{=} \underbrace{\Sigma_{aa}^{t+1} \left(\Sigma_{aa}^t\right)^{-1}}_{=:A} \Sigma_{ab}^t \end{aligned} \tag{16}$$

with

$$\begin{aligned} F_a &= \left[\frac{\partial f(X_i,X_j)}{\partial X_i}(\hat{X}_i^t,\hat{X}_j^t), \frac{\partial f(X_i,X_j)}{\partial X_j}(\hat{X}_i^t,\hat{X}_j^t) \right] \\ &=: \left[F_i, F_j\right] \\ K_a &= \begin{bmatrix} \Sigma_{ii}^t F_i^{\mathrm{T}} + \Sigma_{ij}^t F_j^{\mathrm{T}} \\ \Sigma_{jj}^t F_j^{\mathrm{T}} + \Sigma_{ji}^t F_i^{\mathrm{T}} \end{bmatrix} \left(F_a \Sigma_{aa}^t F_a^{\mathrm{T}} + Q\right)^{-1} \\ &=: \begin{bmatrix} K_i \\ K_j \end{bmatrix} \end{aligned}$$

and $f(\hat{X}_a^t) = f(\hat{X}_i^t,\hat{X}_j^t)$. We present the derivation in Appendix A. The update equations (14)–(16) generalize the update equations by Roumeliotis and Bekey (2002) to arbitrary measurement models. In more explicit form, these equations read

$$\hat{X}_i^{t+1} = \hat{X}_i^t + K_i \left[r - f(\hat{X}_i^t,\hat{X}_j^t) \right] \tag{17}$$

$$\Sigma_{ii}^{t+1} = (\mathbb{I} - K_i F_i)\, \Sigma_{ii}^t - K_i F_j \Sigma_{ji}^t \tag{18}$$

$$\Sigma_{ij}^{t+1} = (\mathbb{I} - K_i F_i)\, \Sigma_{ij}^t - K_i F_j \Sigma_{jj}^t \tag{19}$$

$$\Sigma_{ik}^{t+1} = (\mathbb{I} - K_i F_i)\, \Sigma_{ik}^t - K_i F_j \Sigma_{jk}^t \tag{20}$$

and, equivalently, if we exchange i and j (and keep the arguments in f).

As argued in Section 3.2, we approximate

$$\hat{X}_b^{t+1} \approx \hat{X}_b^t \tag{21}$$

$$\Sigma_{bb}^{t+1} \approx \Sigma_{bb}^t \tag{22}$$

and arrive at update equations equivalent to the Schmidt–Kalman filter. Communication between robot i and robot j is allowed at the time of the relative measurement and they share the measurement r and their beliefs bel_i^t and bel_j^t. In addition, they reproduce their correlation $\Sigma_{ij}^t = \sigma_{ij}^t \left(\sigma_{ji}^t\right)^{\mathrm{T}}$ by sharing the two terms on the right-hand side. Thus, they have access to all elements of Σ_{aa}^t and can exactly update their beliefs with (14) and (15). The new, exactly calculated cross-correlation is decomposed into $\Sigma_{ij}^{t+1} = \sigma_{ij}^{t+1} \left(\sigma_{ji}^{t+1}\right)^{\mathrm{T}}$ and distributed among the two participating robots. In Algorithm 3, we choose the decomposition $\sigma_{ij}^{t+1} = \Sigma_{ij}^{t+1}$ and $\sigma_{ji}^{t+1} = \mathbb{I}$.

The only terms we have not taken care of yet are the correlations between participating and non-participating robots Σ_{ab}^{t+1}. In more explicit form, the exact update (16) reads

$$\begin{bmatrix} \Sigma_{ik}^{t+1} \\ \Sigma_{jk}^{t+1} \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbb{I} - K_i F_i & -K_i F_j \\ -K_j F_i & \mathbb{I} - K_j F_j \end{bmatrix}}_{=A} \begin{bmatrix} \Sigma_{ik}^t \\ \Sigma_{jk}^t \end{bmatrix} \tag{23}$$

for all $k \in \{1,\dots,N\} \setminus \{i,j\}$. In contrast to Roumeliotis and Bekey (2002), we do not rely on communication with other teammates. The problem in this case is that, unlike (9) for private measurements, the update cannot be written as a simple matrix multiplication of the form $\Sigma_{ik}^{t+1} = M\, \Sigma_{ik}^t$. In particular, we cannot reproduce the terms $\Sigma_{ik}^t = \sigma_{ik}^t (\sigma_{ki}^t)^{\mathrm{T}}$ and $\Sigma_{jk}^t = \sigma_{jk}^t (\sigma_{kj}^t)^{\mathrm{T}}$ that are necessary for the correct update. The reason is that we do not rely on communication with robot k, which carries the terms σ_{ki}^t and σ_{kj}^t. To arrive at the general form of approximations that maintain the decentralized structure, we have to use a block-diagonal approximation of A.

A naive approximation of A and thus of (16), or equivalently (23), can be achieved by simply neglecting the off-diagonal blocks of A, which leads to

$$\begin{aligned} \Sigma_{ik}^{t+1} &= (\mathbb{I} - K_i F_i)\, \Sigma_{ik}^t - K_i F_j \Sigma_{jk}^t \\ &\approx (\mathbb{I} - K_i F_i)\, \Sigma_{ik}^t \end{aligned} \tag{24}$$

This approximation is exact if the robots j and k are uncorrelated, i.e., $\Sigma_{jk}^t = 0$.

However, in our experiments, approximation (24) yields unsatisfying results. Thus, we need a more reasoned approximation that also covers the case that the robots j and k are correlated before the relative measurement. The key problem here is to find a relation between the two unavailable terms Σ_{ik}^t and Σ_{jk}^t, that leads to a reasoned block-diagonal approximation A. In the following, we introduce the corresponding approximation and present an informal derivation. In Section 3.4, we present theoretical details.

The main idea is that we obtain the relation

$$\Sigma_{jk}^t \approx \Sigma_{ji}^t \left(\Sigma_{ii}^t\right)^{-1} \Sigma_{ik}^t \tag{25}$$

if robot i is strongly correlated to at least one of the other robots (j and/or k) before the relative measurement. We formalize and prove this in Section 3.4. We insert this into (23) and obtain

$$\begin{aligned} \Sigma_{ik}^{t+1} &\overset{(25)}{\approx} (\mathbb{I} - K_i F_i)\, \Sigma_{ik}^t - K_i F_j \Sigma_{ji}^t \left(\Sigma_{ii}^t\right)^{-1} \Sigma_{ik}^t \\ &\overset{(18)}{=} \Sigma_{ii}^{t+1} \left(\Sigma_{ii}^t\right)^{-1} \Sigma_{ik}^t \\ &=: \tilde{\Sigma}_{ik}^{t+1} \end{aligned} \tag{26}$$

This leads to one of the key contributions of this paper, namely the more reasoned approximation

$$A \approx \tilde{A} := \begin{bmatrix} \Sigma_{ii}^{t+1} \left(\Sigma_{ii}^{t}\right)^{-1} & 0 \\ 0 & \Sigma_{jj}^{t+1} \left(\Sigma_{jj}^{t}\right)^{-1} \end{bmatrix} \tag{27}$$

Before we present our theoretical results on the approximation (27), we sum up the update procedure for the relative measurement. The proposed approach uses the exact update equations (14) and (15) for the subsystem of the two robots involved in the relative measurement; it leaves the subsystem of the robots that are not involved unchanged according to (21) and (22); and it uses the approximation (26) for the cross-correlations between involved and non-involved robots. This procedure is summarized in Algorithm 3.

Please note that multiplying $\tilde{A}$ in (27) by a scalar factor λ, or equivalently modifying (26) to

$$\Sigma_{ik}^{t+1} \approx \lambda\, \Sigma_{ii}^{t+1} \left(\Sigma_{ii}^{t}\right)^{-1} \Sigma_{ik}^{t} \tag{28}$$

results in an alternative update strategy which can be implemented in the same decentralized structure as the proposed algorithm. In particular, setting $\lambda = 0$ means that each robot only maintains the cross-correlation to the teammate it encountered most recently, which is useful for memory restricted systems with a large number of robots.

The updates (26) or (28) require the existence of $\left(\Sigma_{ii}^{t}\right)^{-1}$ for all t. The standard EKF with finite measurement noise always maintains a positive-definite covariance matrix. In that case, the inverse exists. For the proposed approximation of the EKF, the following proposition ensures the existence of $\left(\Sigma_{ii}^{t}\right)^{-1}$.

Proposition 1. *For a matrix Σ^t where all two-robot submatrices are positive-definite, the proposed approximate update according to Algorithm 3 maintains the positive-definiteness off all two-robot submatrices. The same also holds true if we use* (28) *with $|\lambda| \leq 1$ instead of* (26)*.*

Proof. See Appendix B. □

Starting with a positive-definite covariance matrix, repeated application of Proposition 1 ensures that $\left(\Sigma_{ii}^{t}\right)^{-1}$ exists for all t. Note that beyond (26) and (28), all approximations used in the proposed approach also maintain the positive-definiteness as they can be derived from the Schmidt–Kalman filter.

Please note that there are instances where every approximation which relies on a multiplicative update, as (24), (26), and (28) do, underestimates the cross-correlations. Assume, for example, $\Sigma_{ik}^{t} = 0$ and $\Sigma_{jk}^{t} \neq 0$, before the relative measurement. Then, (23) implies in general $\Sigma_{ik}^{t+1} \neq 0$ after the relative measurement. However, with any block-diagonal approximation of A, it remains $\Sigma_{ik}^{t+1} = 0$. However, our experiments indicate, that (26) does not only outperform (24) and (28) for $\lambda \neq 1$, but also leads to good overall localization performance in practice. Moreover, at each time step in the experiments, the positive-definiteness of the joint covariance matrix is preserved.

3.4. Error bounds

The goal of this section is twofold. First, we introduce Lemma 1 which relates the correlations between three jointly distributed random variables. This lemma directly quantifies the error of the approximation in (25). Second, we introduce Corollary 1 to compute an upper bound on the error introduced by the proposed approximation (26) as a function of the correlations between the robots.

If not defined otherwise, $\|A\|$ denotes the spectral norm for a matrix A, i.e.,

$$\|A\| := \sqrt{\lambda_{\max}\left(A^{\mathrm{T}}A\right)} \tag{29}$$

where $\lambda_{\max}(A^{\mathrm{T}}A)$ denotes the maximal eigenvalue of $A^{\mathrm{T}}A$.

First, we formalize and derive the relation introduced in (25). Therefore, we need the following definition and lemma.

Definition 1. *For two d-dimensional, jointly distributed random variables X_i and X_j with the $(2d \times 2d)$-dimensional covariance matrix of the joint system*

$$\Sigma = \begin{pmatrix} \Sigma_{ii} & \Sigma_{ij} \\ \Sigma_{ji} & \Sigma_{jj} \end{pmatrix}$$

*we define X_i and X_j to be ϵ-*correlated*, if the following relation for the Schur complement of Σ_{jj} in Σ is valid:*

$$\|\Sigma_{ii} - \Sigma_{ij}\Sigma_{jj}^{-1}\Sigma_{ji}\| = \epsilon$$

In other words, for ϵ-correlated variables, ϵ is the spectral norm of the covariance of X_i conditioned on X_j. Please note that $0 \leq \epsilon$, with equality for maximally correlated variables, and $\epsilon \leq \|\Sigma_{ii}\|$, with equality for uncorrelated variables. The latter inequality is due to the fact that Σ_{ii}, $\Sigma_{ij}\Sigma_{jj}^{-1}\Sigma_{ji}$, and $\Sigma_{ii} - \Sigma_{ij}\Sigma_{jj}^{-1}\Sigma_{ji}$ are positive-semidefinite.

Lemma 1. *Assume three jointly distributed d-dimensional random variables $\{X_l\}_{l\in\{i,j,k\}}$ with the positive-definite $(3d \times 3d)$-dimensional covariance matrix*

$$\begin{pmatrix} \Sigma_{ii} & \Sigma_{ij} & \Sigma_{ik} \\ \Sigma_{ji} & \Sigma_{jj} & \Sigma_{jk} \\ \Sigma_{ki} & \Sigma_{kj} & \Sigma_{kk} \end{pmatrix}$$

Further, assume that X_j and X_i are ϵ_{ji}-correlated and X_k and X_i are ϵ_{ki}-correlated. Then, we have

$$\|\Sigma_{jk} - \Sigma_{ji}\left(\Sigma_{ii}\right)^{-1}\Sigma_{ik}\| \leq \sqrt{\epsilon_{ji}\epsilon_{ki}}$$

Proof. See Appendix C. □

With the assumption of a positive-definite covariance matrix, this lemma directly implies that approximation (25) and, thus, (26) are exact for sufficiently strongly correlated robots, i.e., sufficiently small ϵ_{ji} and ϵ_{ki}. More explicitly, for the error introduced by (26), Lemma 1 implies the following corollary.

Corollary 1. *With the assumption that* $\{X_l\}_{l\in\{i,j,k\}}$ *are jointly distributed with a positive-definite covariance matrix, and the random variables* X_j *and* X_i *are* ϵ_{ji}*-correlated and* X_k *and* X_i *are* ϵ_{ki}*-correlated, the following bound for the approximation error introduced by* (26) *is valid*

$$\|\Sigma_{ik}^{t+1} - \tilde{\Sigma}_{ik}^{t+1}\| \leq \|K_i F_j\| \sqrt{\epsilon_{ji}\epsilon_{ki}}$$

Proof. See Section 5.1 □

In the following, we demonstrate that with the assumption of a positive-definite covariance matrix, Corollary 1 provides error bounds for all possible combinations of inter-robot correlations which can be taken care of by a multiplicative cross-correlation update. Assuming that two robots can either be uncorrelated or correlated, there are 2^N possible combinations of correlations before the measurement. Without loss of generality, we look at $N = 3$. Corollary 1 provides error bounds for the cases that j and i or k and i are correlated.[3] This covers six of the eight cases. If all robots are uncorrelated, it is trivial to see that the proposed approximation is exact. The remaining combination is that only j and k are correlated. In this case, any multiplicative update yields the same error

$$\|\Sigma_{ik}^{t+1} - \tilde{\Sigma}_{ik}^{t+1}\| \leq \|K_i F_j\| \|\Sigma_{jk}^{t}\|$$

as can be seen by plugging $\Sigma_{ik}^{t} = \Sigma_{ij}^{t} = 0$ into the proof of Corollary 1 in Section 5.1.

4. Experiments with real-world data

We test our approach on the publicly available *utias multi-robot cooperative localization and mapping dataset*, recorded by Leung et al. (2011). A fleet of five (two-wheel differential drive) robots obtain range and bearing measurements with known correspondences to each other and up to 15 distinguishable landmarks. The dataset includes odometry and measurements together with pose ground truth. The duration is over 4.7 hours, spread over nine different runs with different trajectories ranging from 15 to 70 minutes. In the ninth run, the environment contains some barriers to reduce the number of measurements, as depicted in Figure 1. We let the robots perform localization based on eight different strategies (listed below), using dead reckoning and the relative measurements. In addition, we allow one of the robots to process its landmark measurements. To increase the amount of data fivefold, we consecutively allow each of the robots (one at a time) to use the landmark measurements. This provides us with over 23.6 hours of (not entirely independent) data. Then, we drop all relative bearing measurements, which leaves the robots with relative range only measurements. This serves to demonstrate that our approach can deal with generic measurement models. Then we replace all relative measurements by simulated relative pose measurements, consisting of the pose of the observed robot in the observing robot's reference frame. This experiment allows us to compare against CI (Carrillo-Arce et al., 2013), which relies on relative measurements providing the observing robot with an estimate of the observed robot's pose. Table 1 gives an overview on the number of velocity commands, private measurements, and relative measurements per robot in each scenario. The values in parentheses are the rates as a percentage.

Table 1. Overview of how often the different algorithms have to be run per robot in the different scenarios. The values in parentheses are the rates as a percentage.

	Algorithm 1	Algorithm 2	Algorithm 3
1	320,656 (90.43)	27,470 (7.75)	6,477 (1.83)
2	424,754 (90.13)	37,330 (7.92)	9,169 (1.95)
3	339,503 (88.98)	35,125 (9.21)	6,940 (1.82)
4	282,464 (87.75)	32,542 (10.11)	6,873 (2.14)
5	531,936 (91.78)	36,450 (6.29)	11,183 (1.93)
6	193,086 (90.88)	15,374 (7.24)	4,003 (1.88)
7	189,383 (90.33)	16,057 (7.66)	4,225 (2.02)
8	841,204 (89.95)	75,584 (8.08)	18,355 (1.96)
9	462,895 (91.47)	36,516 (7.22)	6,641 (1.31)
sum	3,585,881 (90.27)	312,448 (7.87)	73,866 (1.86)

We compare the following EKF-based approaches.

- Collaborative localization using the centralized joint EKF, e.g., the approach of Roumeliotis and Bekey (2002) (ground truth).
- Schmidt–Kalman filter (SK): identical to our proposed approach with the difference, that SK uses the exact update (16) instead of the decentralized approximation (26), see Schmidt (1966).
- Our approach (DCL): We use (26).
- Our approach with a scaling factor λ for the cross correlations (D[λ]): We use (28) with different values for λ instead of (26).
- Total naive collaborative localization (NCL): cross-correlations are neglected.
- Naive version of our approach (NDCL): we use approximation (24) instead of (26).
- Single Robot Localization (SL): robots can detect landmarks but no other robots.
- CI: information to be fused is assumed to be maximally correlated, see Carrillo-Arce et al. (2013).

Note that to realize the centralized equivalent approach (EKF), each measurement demands for $N - 1 = 4$ edges in the communication graph. In contrast, our approach (DCL) requires only one edge for each relative measurement and no communication at all for private measurements. The Schmidt–Kalman filter (SK) needs $N - 1 = 4$ edges for each relative measurement and no communication for private measurements. The datasets include a total number of $n_{\text{rel}} = 73{,}866$ relative measurements and $n_{\text{priv}} \approx 62{,}490$ private measurements on average for one of the robots performing landmark measurements. Thus, EKF needs $(N-1)(n_{\text{rel}} + n_{\text{priv}})/(n_{\text{rel}}) \approx 7.38$ times more edges, and SK needs four times more edges than our approach.

4.1. Accuracy analysis

As a measure for the accuracy of an estimator, we calculate the distance between the system's true and estimated position and average it over all time steps

$$\mathrm{D} := \frac{1}{T}\sum_{t=1}^{T}\left\|X^t - \hat{X}^t\right\|_2$$

where T is the absolute number of evaluated time steps.

As a measure for consistency, we choose the average normalized estimation error squared

$$\mathrm{ANEES} := \frac{1}{T}\sum_{t=1}^{T}(X^t - \hat{X}^t)^{\mathrm{T}}\left(\Sigma^t\right)^{-1}(X^t - \hat{X}^t)$$

For the quantitative evaluations in Tables 2 and 3, we normalize both metrics D and ANEES by the corresponding values of the exact EKF. When normalized accordingly, we refer to these metrics as $\overline{\mathrm{D}}$ and $\overline{\mathrm{ANEES}}$. This choice is justified by the following observation. Owing to linearization errors, approximate measurement and motion models, and randomness in the system, even EKF sometimes produces incorrect estimates. In these cases it is unlikely (though possible) that any approximation of EKF yields better results. This is illustrated in Figure 2, where we do not normalize the metrics to demonstrate the relation between the different approaches and the EKF. This figure shows D and ANEES in the second run, where the relative measurements consist of range information only. In the time between seconds 400 to 500, all algorithms (including the correct EKF) encounter problems. For the rest, our approach (DCL) yields localization with D below around 60 cm (mostly below 30 cm). After around 1,200 seconds, EKF yields bad localization results, while our algorithm remains stable. This is due to a robot processing landmark measurements with wrong associations. In EKF the robot instantaneously communicates these measurements which corrupts all other robots' estimates. In this particular example, DCL outperforms EKF due to its decentralized structure by not communicating the faulty measurement information.[4] The rest of the time, it performs almost identically well as EKF. Using single robot localization (SL), the error diverges due to the fact that only a subset of robots detect landmarks. At first, the naive approach (NCL) keeps up with DCL. However, due to overconfidence its localization performance starts to deteriorate after around 1,200 seconds.

Table 2 shows the $\overline{\mathrm{D}}$ for each of the approximate algorithms in the nine scenarios, normalized to the corresponding value of the exact EKF. For each scenario, we perform five runs where each time another robot processes private measurements. For each run, the result of the best decentralized algorithm is in bold print. Please note that the Schmidt–Kalman filter (SK) is not fully decentralized but depends on all-to-all communication for each relative measurement. The mean is taken over all 45 runs weighted

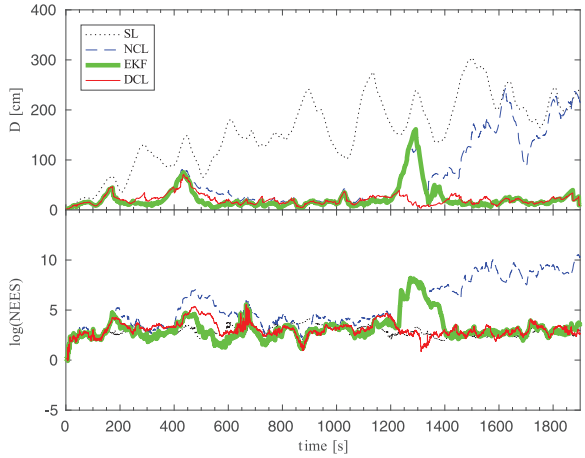

Fig. 2. Mean error D and ANEES in the second run. With our approach (DCL), the error remains below 50 cm throughout the whole experiment. Using single robot localization (SL), the error diverges due to the fact that only a subset of robots detect landmarks. At first, the naive approach (NCL) keeps up with DCL. However, due to overconfidence its localization performance starts to deteriorate after around 1,200 seconds.

with the respective duration. We perform a one-tailed, paired-sample t-test. The resulting p-value is the probability that the corresponding algorithm is outperformed by the proposed DCL. With a p-value above 90%, the proposed approach outperforms the alternative update strategy NDCL. The latter also performs well in most of the runs but completely fails in some runs. Note that for NDCL, the p-value cannot be interpreted as a probability because NDCL violates a condition to apply the t-test: results have to be normally distributed. We attribute the outliers in NDCL to the fact that there is no result like Proposition 1 for NDCL. With probability near to one, the proposed approach outperforms all other decentralized algorithms. On an average, its error is 12% higher than the error of the exact EKF and 3% higher than the error of SK, although the latter two algorithms need four (respectively, nearly five) times more communication links.

We repeat each run five times with the initial estimate chosen randomly from a Gaussian with mean at the true pose to obtain 225 runs. We erase all relative bearing measurements to obtain another 225 runs with range-only relative measurements. We generate another 225 runs by replacing all relative measurements by simulated relative pose measurements. Table 3 shows the mean $\overline{\mathrm{D}}$ and $\overline{\mathrm{ANEES}}$ over all 225 runs (weighted by the duration) for the different measurement types and algorithms. Lower values of $\overline{\mathrm{D}}$ indicate better localization accuracy. Values of $\overline{\mathrm{ANEES}}$ greater (smaller) than one indicate that the filter is less (more) conservative than the exact EKF.

To illustrate the influence of the communication frequency on the estimate, we repeat the whole experiment multiple times, such that each robot only processes

Table 2. Mean error $\overline{\mathrm{D}}$ on real-world data with range and bearing measurements, normalized by the corresponding values of the exact EKF. For each run, the best value is in bolt, where SK is not considered as it is not fully decentralized.

Scenario	SK	DCL	D[0.75]	D[0.5]	D[0.25]	D[0]	NCL	NDCL	SL
	1.03	**1.12**	1.39	1.49	1.54	1.57	1.60	1.19	17.34
	1.12	1.14	1.23	1.27	1.29	1.31	1.37	**1.13**	22.61
1	0.99	**1.04**	1.20	1.26	1.29	1.31	1.30	1.13	16.99
	1.07	**1.12**	1.34	1.39	1.43	1.46	1.66	1.20	20.54
	0.86	0.86	0.87	0.85	0.84	**0.83**	0.85	74.46	9.59
	1.11	**1.17**	1.34	1.42	1.45	1.47	1.47	1.19	17.25
	1.11	**1.11**	1.22	1.28	1.32	1.36	1.43	1.14	20.74
2	1.04	**1.05**	1.09	1.12	1.14	1.16	1.21	1.09	18.03
	1.24	**1.28**	1.30	1.34	1.36	1.38	1.38	1.38	13.79
	1.04	**1.05**	1.13	1.16	1.18	1.19	1.31	1.23	23.55
	0.96	**0.97**	1.27	1.44	1.55	1.62	1.99	71.41	21.98
	1.04	**1.05**	1.20	1.30	1.35	1.39	1.70	1.05	17.63
3	1.00	**0.99**	1.06	1.11	1.15	1.17	1.49	1.00	11.37
	0.65	1.02	0.91	0.92	0.92	0.92	**0.83**	6.37	7.71
	1.04	**1.02**	1.02	1.03	1.03	1.03	1.11	1.02	15.62
	1.03	1.12	1.46	1.47	1.47	1.47	1.84	**1.11**	11.81
	1.26	**1.22**	1.25	1.27	1.29	1.30	1.52	26.02	9.88
4	1.14	1.22	1.20	1.21	1.22	1.22	1.28	**1.19**	11.40
	0.97	**1.06**	1.37	1.39	1.36	1.36	1.50	595.81	8.99
	1.29	1.26	**1.09**	1.09	1.10	1.10	1.32	1.14	11.80
	1.06	1.12	1.20	1.24	1.27	1.30	1.46	**1.11**	14.68
	1.08	**1.16**	1.18	1.19	1.20	1.22	1.40	1.18	16.92
5	1.12	**1.13**	1.39	1.52	1.58	1.61	1.65	2.10	16.43
	1.09	**1.05**	1.14	1.16	1.16	1.15	1.35	1.88	13.12
	1.06	**1.05**	1.17	1.21	1.22	1.23	1.34	1.24	16.57
	1.12	**1.11**	1.16	1.27	1.35	1.41	1.47	1.56	13.17
	1.03	**1.02**	1.02	1.03	1.04	1.05	1.31	19.48	14.61
6	1.01	**1.05**	1.11	1.14	1.16	1.18	1.24	3.45	13.77
	1.16	1.24	**1.23**	1.29	1.33	1.37	1.51	1.88	12.70
	1.10	1.12	1.14	1.17	1.19	1.20	1.38	**1.11**	16.19
	1.09	**1.14**	1.24	1.22	1.22	1.22	1.40	1.18	9.57
	1.08	1.12	1.17	1.19	1.20	1.22	1.28	**1.11**	11.90
7	1.06	1.08	1.11	1.13	1.14	1.16	1.28	**1.06**	11.51
	1.03	1.22	1.31	1.33	1.35	1.37	1.51	**1.20**	8.60
	1.14	**1.16**	1.19	1.21	1.21	1.22	1.22	1.17	11.41
	1.13	**1.21**	1.53	1.61	1.65	1.68	1.97	1.37	19.94
	1.10	**1.11**	1.19	1.20	1.21	1.22	1.39	2.06	26.95
8	1.08	**1.10**	1.21	1.27	1.30	1.32	1.44	1.19	24.76
	1.02	**1.09**	1.30	1.38	1.43	1.46	1.75	1.18	14.96
	1.12	**1.13**	1.24	1.27	1.29	1.30	1.39	1.15	23.03
	1.22	1.03	1.05	1.06	1.07	**1.02**	1.15	1.46	5.16
	0.87	0.90	**0.85**	0.86	0.93	0.88	0.96	1.03	4.15
9	1.24	1.29	**1.15**	1.39	1.39	1.19	1.25	1.43	4.29
	1.18	1.35	1.26	**1.15**	1.40	1.46	1.44	3.16	3.94
	1.11	1.09	1.11	1.14	1.13	1.17	**1.02**	1.16	3.57
Mean	1.09	**1.12**	1.15	1.18	1.23	1.22	1.29	10.95	9.53
Variance	0.97	0.98	0.93	0.96	1.03	0.99	1.01	19.95	3.75
p-value	0.0012	-	1.0000	1.0000	1.0000	1.0000	1.0000	0.9049	1.0000

Table 3. $\overline{\mathrm{D}}$ and $\overline{\mathrm{ANEES}}$ over the entire dataset for three different measurement models, normalized by the corresponding values of the exact EKF.

Range and bearing (all relative measurements)	SK	DCL	D[0.75]	D[0.5]	D[0.25]	D[0]	NCL	NDCL	SL	CI
$\overline{\mathrm{D}}$	1.09	**1.12**	1.15	1.18	1.23	1.22	1.29	4.63	9.53	–
$\overline{\mathrm{ANEES}}$	0.28	0.30	0.68	0.79	0.87	0.91	1.33	1590.40	0.21	–
Range and bearing (50% of relative measurements)	SK	DCL	D[0.75]	D[0.5]	D[0.25]	D[0]	NCL	NDCL	SL	CI
$\overline{\mathrm{D}}$	1.01	**1.06**	1.10	1.11	1.12	1.12	1.21	2.47	8.36	–
$\overline{\mathrm{ANEES}}$	0.61	0.66	1.00	1.14	1.23	1.28	1.96	35.44	0.49	–
Range and bearing (10% of relative measurements)	SK	DCL	D[0.75]	D[0.5]	D[0.25]	D[0]	NCL	NDCL	SL	CI
$\overline{\mathrm{D}}$	1.07	**1.07**	1.11	1.11	1.11	1.13	1.14	1.11	6.60	–
$\overline{\mathrm{ANEES}}$	0.96	0.97	1.06	1.07	1.09	1.11	1.22	0.99	0.84	–
Range only	SK	DCL	D[0.75]	D[0.5]	D[0.25]	D[0]	NCL	NDCL	SL	CI
$\overline{\mathrm{D}}$	1.01	**1.00**	1.11	1.13	1.14	1.16	1.19	1.55	4.09	–
$\overline{\mathrm{ANEES}}$	0.69	0.71	1.14	1.31	1.40	1.53	1.89	6.45	0.42	–
Simulated relative pose	SK	DCL	D[0.75]	D[0.5]	D[0.25]	D[0]	NCL	NDCL	SL	CI
$\overline{\mathrm{D}}$	1.07	**1.13**	1.21	1.24	1.25	1.26	1.36	1.61	8.53	2.23
$\overline{\mathrm{ANEES}}$	0.72	0.96	2.92	3.44	3.74	3.96	4.87	12.94	0.33	0.40

and communicates only a subset of its relative measurements. Note that without relative measurements, all presented decentralized algorithms become identical to the Schmidt–Kalman filter. In particular, if the initial correlations between robots are zero, all algorithms become identical to the EKF. This tendency can be seen in Figure 3 and Table 3.

To illustrate the influence of faulty landmark detections, we repeat the whole experiment twice, such that 5% (10%) of the landmark detections are faulty. Therefore, we replace the correct association by randomly drawing the id of one of the other landmarks. The results in Figure 4 indicate that a growing number of faulty landmark detections harms the estimate of the EKF more than the estimates of the other algorithms.

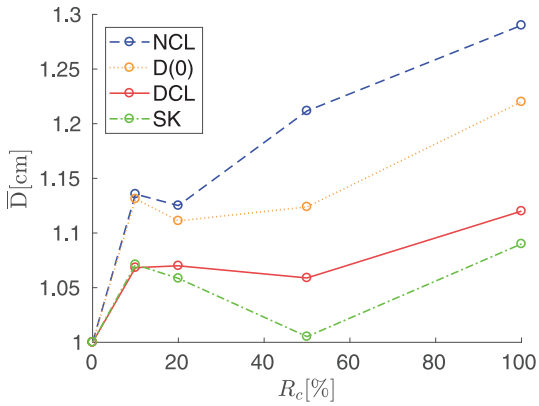

Fig. 3. Influence of the communication frequency on the localization accuracy. We show the error $\overline{\mathrm{D}}$ in relation to the error of the centralized EKF depending on the ratio R_c of processed relative measurements. With decreasing R_c, the relative error $\overline{\mathrm{D}}$ decreases, while the absolute error D, not shown in this figure, obviously increases.

4.2. Robustness analysis

As stated before, even the correct EKF might diverge. However, if the covariance is not underestimated, the filter might recover from that after subsequent measurements. To demonstrate robustness of our approximation, we compare mean time to failure and the relative number of recoveries after failures for the different approaches. We define a failure to be a point in time where the root-mean-square error of the joint system exceeds 50 cm. We define a recovery as a point in time where the root-mean-square error of the joint system drops below 10 cm after a failure. The recovery ratio R is the number of recoveries divided by the number of total failures. Mean time to failure T is the mean time difference between a recovery (or the start of a run) and the next failure. Figure 5 serves as an example. It shows the estimation

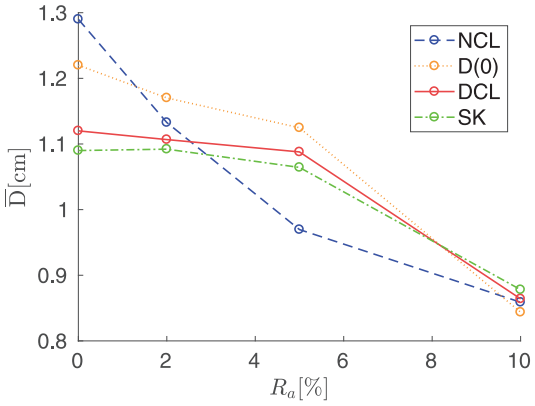

Fig. 4. Influence of faulty landmark associations on the localization accuracy. We show the error $\overline{\mathrm{D}}$ in relation to the error of the centralized EKF depending on the ratio R_a of faulty landmark associations. With increasing R_a, the absolute error D, not shown in this figure, obviously increases for all algorithms. However, the relative error $\overline{\mathrm{D}}$ decreases, as the wrong information affects the centralized EKF more than the decentralized algorithms.

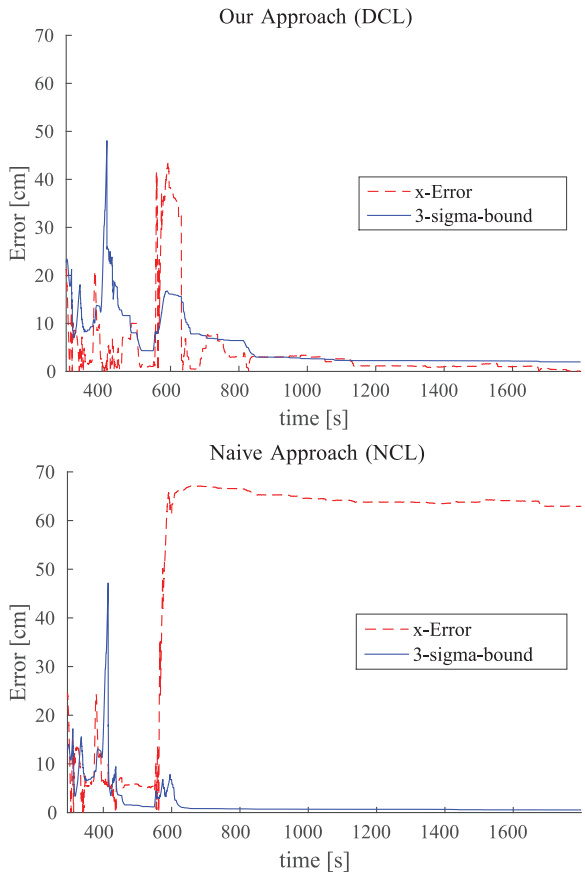

Fig. 5. Absolute difference $|x|$ between the correct x-component (of robot 4 in run 3) and the corresponding value estimated with our algorithm (DCL, plot on top) and the naive approach (NCL, plot at the bottom). This is an example, where all algorithms (including the correct centralized EKF) lose positioning around time 550–650 s. In contrast to the naive one, our approach is able to recover from that and yields accurate subsequent localization.

error of the x-component over time for one of the robots that only rely on dead reckoning and relative measurements. All compared algorithms, including EKF (not shown), lose positioning. In contrast to the naive approach (NCL), our approach (DCL) is able to recover from the bad localization around $t = 600$ s, and yield an accurate estimate later on.

Table 4 displays the mean time to failure in minutes and the recovery ratio as a percentage. For each measurement type, these values are taken over the complete 23.6 hours of data. Our algorithm (DCL) outperforms the EKF with regards to mean time to failure. With respect to recovery rate, our approach is approximately identical to EKF. Note that mean time to failure is very low for CI. We accredit this to the fact that in CI only the observed robot profits from a relative measurement. However, conservative approximation in CI achieves a very high recovery ratio. The experiments demonstrate our algorithm to be as robust as the correct EKF while it uses nearly five times less communication links and does not rely on the storage of measurements or global communication.

5. Conclusion

This paper is a revised and substantially extended version of Luft et al. (2016), which introduced a fully decentralized, EKF-based algorithm for collaborative localization. To the best of the authors' knowledge, this is the first approach to track inter-robot correlations while fulfilling all of the following relevant conditions: the algorithm does not require storage of measurements, it supports generic measurement models, and communication is limited to pairs of robots that obtain a relative measurement. Each robot maintains only the latest estimate of its own pose. The fact that the proposed approach can handle these particularly challenging requirements make it suitable for a wide range of applications, such as underwater or in mines.

For the approximations with respect to the standard EKF, we present explicit error bounds depending on the inter-robot correlations. As extensive tests with real-world datasets show, our approach outperforms existing decentralized EKF-based algorithms in terms of both accuracy and robustness. Albeit our estimator is not provably consistent in general, our experiments imply that it is as robust as the standard EKF, which – in contrast to our method – rests on the assumption of persistent availability of global communication or has substantial memory requirements.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work has been partially supported by the European Commission under ERC-AG-PE7-267686-LIFENAV, the Graduate School of Robotics in Freiburg, and the State Graduate Funding Program of Baden-Württemberg.

Table 4. Mean time to failure T in minutes and recovery ratio R as a percentage for real-world data.

Range and bearing	EKF	SK	DCL	D[0.75]	D[0.5]	D[0.25]	D[0]	NCL	NDCL	SL	CI
T	42.37	42.43	46.89	47.78	43.43	41.61	41.56	37.68	24.09	2.62	–
R	75.00	75.00	72.00	75.00	76.92	77.78	77.78	71.43	76.09	0	–
Range only	**EKF**	**SK**	**DCL**	**D[0.75]**	**D[0.5]**	**D[0.25]**	**D[0]**	**NCL**	**NDCL**	**SL**	**CI**
T	11.36	11.86	11.80	10.33	9.42	9.28	8.94	7.98	10.37	2.65	–
R	56.45	52.63	57.63	46.55	45.00	39.66	38.98	40.00	49.18	0	–
Simulated relative pose	**EKF**	**SK**	**DCL**	**D[0.75]**	**D[0.5]**	**D[0.25]**	**D[0]**	**NCL**	**NDCL**	**SL**	**CI**
T	38.79	47.75	45.14	45.20	42.38	42.32	40.56	42.44	37.34	2.40	5.10
R	80.00	75.00	76.00	79.17	76.00	76.00	76.92	80.00	72.41	0	93.60

Notes

1. We define *communication complexity* as the number of edges required in the communication graph, see, e.g., Kia et al. (2014).
2. Please note that Julier (2001) and Eustice et al. (2006) have used the same structure to reduce the computational cost of the related problem of *SLAM*.
3. Corollary 1 is also valid if both j and i, and k and i are uncorrelated. However, the stronger the correlations between j and i, and k and i are, the tighter is the error bound.
4. We emphasize that it is not our goal to outperform EKF. In general, the estimate profits from robots communicating their local information to the team. Only in the particular case of a faulty data association does it harm the estimate of the joint EKF. We perform experiments with faulty landmark associations, see Section 4.

References

Bahr A, Walter MR and Leonard JJ (2009) Consistent cooperative localization. In: *Proceedings of IEEE International Conference on Robotics and Automation (ICRA)*, pp. 3415–3422.

Bailey T, Bryson M, Mu H, Vial J, McCalman L and Durrant-Whyte HF (2011) Decentralised cooperative localisation for heterogeneous teams of mobile robots. In: *Proceedings of IEEE International Conference on Robotics and Automation (ICRA)*, pp. 2859–2865.

Boyd S, Ghosh A, Prabhakar B and Shah D (2006) Randomized gossip algorithms. *IEEE/ACM Transactions on Networking* 14: 2508–2530.

Carrillo-Arce LC, Nerurkar ED, Gordillo JL and Roumeliotis SI (2013) Decentralized multi-robot cooperative localization using covariance intersection. In: *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 1412–1417.

Choudhary S, Carlone L, Nieto C, Rogers J, Christensen HI and Dellaert F (2016) Distributed trajectory estimation with privacy and communication constraints: A two-stage distributed Gauss–Seidel approach. In: *2016 IEEE International Conference on Robotics and Automation (ICRA)*. pp. 5261–5268.

Cunningham A, Indelman V and Dellaert F (2013) DDF-SAM 2.0: Consistent distributed smoothing and mapping. *2013 IEEE International Conference on Robotics and Automation*, pp. 5220–5227.

Cunningham A, Wurm KM, Burgard W and Dellaert F (2012) Fully distributed scalable smoothing and mapping with robust multi-robot data association. In: *2012 IEEE International Conference on Robotics and Automation*, pp. 1093–1100.

Eustice R, Singh H, Leonard JJ and Walter M (2006) Visually mapping the RMS Titanic: Conservative covariance estimates for SLAM information filters. *The International Journal of Robotics Research* 25: 1223–1242.

Fox D, Burgard W, Kruppa H and Thrun S (2000) A probabilistic approach to collaborative multi-robot localization. *Autonomous Robots* 8(3): 325–344.

Howard A, Matarić MJ and Sukhatme G (2002) Localization for mobile robot teams using maximum likelihood estimation. In: *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 434–439.

Howard A, Matarić MJ and Sukhatme G (2003) Putting the 'i' in 'team': an ego-centric approach to cooperative localization. In: *Proceedings of IEEE International Conference on Robotics and Automation (ICRA)*, pp. 868–874.

Huang GP, Trawny N, Mourikis AI and Roumeliotis SI (2011) Observability-based consistent EKF estimators for multi-robot cooperative localization. *Autonomous Robots* 30(1): 99–122.

Indelman V, Gurfil P, Rivlin E and Rotstein H (2012) Graph-based distributed cooperative navigation for a general multi-robot measurement model. *The International Journal of Robotics Research* 31(9): 1057–1080.

Julier SJ (2001) A sparse weight Kalman filter approach to simultaneous localisation and map building. In: *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 1251–1256.

Karam N, Chausse F, Aufrere R and Chapuis R (2006) Cooperative multi-vehicle localization. In: *IEEE Intelligent Vehicles Symposium*, pp. 564–570.

Kia SS, Rounds SF and Martinez S (2014) A centralized-equivalent decentralized implementation of extended Kalman filters for cooperative localization. In: *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 3761–3766.

Kurazume R, Nagata S and Hirose S (1994) Cooperative positioning with multiple robots. In: *Proceedings of IEEE International Conference on Robotics and Automation (ICRA)*, pp. 1250–1257.

Lavergne P (2008) A Cauchy–Schwarz inequality for expectation of matrices. Discussion papers, Department of Economics, Simon Fraser University.

Leung KYK, Barfoot TD and Liu HHT (2010) Decentralized localization of sparsely-communicating robot networks: A centralized-equivalent approach. *IEEE Transactions on Robotics* 26(1): 62–77.

Leung KYK, Barfoot TD and Liu HHT (2012) Decentralized cooperative SLAM for sparsely-communicating robot networks: A centralized-equivalent approach. *Journal of Intelligent and Robotic Systems* 66(3): 321–342.

Leung KYK, Halpern Y, Barfoot TD and Liu HHT (2011) The utias multi-robot cooperative localization and mapping dataset. *The International Journal of Robotics Research* 30(8): 969–974.

Li H and Nashashibi F (2013) Cooperative multi-vehicle localization using split covariance intersection filter. *IEEE Intelligent Transportation Systems Magazine* 5(2): 33–44.

Luft L, Schubert T, Roumeliotis SI and Burgard W (2016) Recursive decentralized collaborative localization for sparsely communicating robots. In: *Proceedings of Robotics: Science and Systems*, Ann Arbor, Michigan.

Martinelli A (2007) Improving the precision on multi robot localization by using a series of filters hierarchically distributed. In: *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 1053–1058.

Martinelli A, Pont F and Siegwart R (2005) Multi-robot localization using relative observations. In: *Proceedings of IEEE International Conference on Robotics and Automation (ICRA)*, pp. 2797–2802.

Martinelli A and Siegwart R (2005) Observability analysis for mobile robot localization. In: *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1471–1476.

Nerurkar ED, Roumeliotis SI and Martinelli A (2009) Distributed maximum a posteriori estimation for multi-robot cooperative localization. In: *Proceedings of IEEE International Conference on Robotics and Automation (ICRA)*, pp. 1402–1409.

Panzieri S, Pascucci F and Setola R (2006) Multirobot localisation using interlaced extended Kalman filter. In: *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. pp. 2816–2821.

Prorok A, Bahr A and Martinoli A (2012) Low-cost collaborative localization for large-scale multi-robot systems. In: *Proceedings of IEEE International Conference on Robotics and Automation (ICRA)*, pp. 4236–4241.

Roumeliotis SI and Bekey GA (2002) Distributed multirobot localization. *IEEE Transactions on Robotics and Automation* 18(5): 781–795.

Schmidt SF (1966) Applications of state space methods to navigation problems. *Advances in Control Systems* 3: 293–340.

Thrun S, Burgard W and Fox D (2005) *Probabilistic Robotics.* Cambridge, MA: The MIT Press.

Walls JM and Eustice RM (2013) An exact decentralized cooperative navigation algorithm for acoustically networked underwater vehicles with robustness to faulty communication: Theory and experiment. In: *Proceedings of the Robotics: Science and Systems Conference*, Berlin, Germany.

Wanasinghe TR, Mann GKI and Gosine RG (2015) Distributed leader-assistive localization method for a heterogeneous multirobotic system. *IEEE Transactions on Automation Science and Engineering* 12(3): 795–809.

Wanasinghe TR, Mann GKI and Raymond GG (2014) Decentralized cooperative localization for heterogeneous multi-robot system using split covariance intersection filter. In: *Proceedings of IEEE International Conference on Computer and Robot Vision (CRV)*, pp. 167–174.

Appendix A: Derivation of the update equations

We derive the update equations from the standard Kalman filter, which is described for example in Thrun et al. (2005, Ch. 1). We partition the covariance of the joint system into blocks

$$\Sigma^t = \begin{bmatrix} \Sigma^t_{aa} & \Sigma^t_{ab} \\ \Sigma^t_{ba} & \Sigma^t{}_{bb} \end{bmatrix}$$

such that a is the set of indices of robots that are involved in a measurement and $b = \{1, \ldots, N\} \setminus a$. Thus, with the corresponding partition, we have the Jacobian of the measurement model $F = [F_a, 0]$. For a relative measurement between robots 1 and 2, for example, we have $a = \{1, 2\}$ and $F = [F_1, F_2, 0, \ldots, 0]$. For a private measurement of robot 1, we have $a = 1$ and $F = [H, 0, 0, \ldots, 0]$. The Kalman gain is calculated as

$$\begin{aligned} K &= \Sigma^t F^{\mathrm{T}} (F \Sigma^t F^{\mathrm{T}} + Q)^{-1} \\ &= \begin{bmatrix} \Sigma^t_{aa} & \Sigma^t_{ab} \\ \Sigma^t_{ba} & \Sigma^t{}_{bb} \end{bmatrix} \begin{bmatrix} F_a^{\mathrm{T}} \\ 0 \end{bmatrix} \\ &\quad \cdot \left([F_a, 0] \begin{bmatrix} \Sigma^t_{aa} & \Sigma^t_{ab} \\ \Sigma^t_{ba} & \Sigma^t{}_{bb} \end{bmatrix} \begin{bmatrix} F_a^{\mathrm{T}} \\ 0 \end{bmatrix} + Q \right)^{-1} \\ &= \begin{bmatrix} \Sigma^t_{aa} F_a^{\mathrm{T}} \\ \Sigma^t_{ba} F_a^{\mathrm{T}} \end{bmatrix} \cdot \left(F_a \Sigma^t_{aa} F_a^{\mathrm{T}} + Q \right)^{-1} \\ &=: \begin{bmatrix} K_a \\ K_b \end{bmatrix} \end{aligned}$$

For the updated covariance Σ^{t+1}, we have

$$\begin{aligned} \Sigma^{t+1} &= (\mathbb{I} - KF)\, \Sigma^t \\ &= \begin{bmatrix} \mathbb{I} - K_a F_a & 0 \\ -K_b F_a & \mathbb{I} \end{bmatrix} \begin{bmatrix} \Sigma^t_{aa} & \Sigma^t_{ab} \\ \Sigma^t_{ba} & \Sigma^t{}_{bb} \end{bmatrix} \end{aligned}$$

Element-wise evaluation yields

$$\begin{aligned} \Sigma^{t+1}_{aa} &= (\mathbb{I} - K_a F_a)\, \Sigma^t_{aa} \\ \Sigma^{t+1}_{ab} &= (\mathbb{I} - K_a F_a)\, \Sigma^t_{ab} \\ \Sigma^{t+1}_{bb} &= \Sigma^t_{bb} - K_b F_a \Sigma^t_{ab} \end{aligned}$$

Now we can easily transform these equations back into the representation where each block corresponds to a single robot's pose. The update for $\hat{X}$ can be derived in the same, straightforward manner.

Appendix B: Proof of Proposition 1

Without loss of generality, we prove Proposition 1 for a three-robot system with

$$\Sigma^t = \begin{bmatrix} \Sigma^t_{11} & \Sigma^t_{12} & \Sigma^t_{13} \\ \Sigma^t_{21} & \Sigma^t_{22} & \Sigma^t_{23} \\ \Sigma^t_{31} & \Sigma^t_{32} & \Sigma^t_{33} \end{bmatrix}$$

With the proposed approach, an approximate update for a relative measurement between robots 1 and 2 consists of Equations (14)-(15), (22), and (26). We prove the proposition for (28) with $|\lambda| \leq 1$, which is a generalization of (26). After the approximate update, we have

$$\tilde{\Sigma}^{t+1} = \begin{bmatrix} \Sigma^{t+1}_{11} & \Sigma^{t+1}_{12} & \tilde{\Sigma}^{t+1}_{13} \\ \left(\Sigma^{t+1}_{12}\right)^T & \Sigma^{t+1}_{22} & \tilde{\Sigma}^{t+1}_{23} \\ \left(\tilde{\Sigma}^{t+1}_{13}\right)^{\mathrm{T}} & \left(\tilde{\Sigma}^{t+1}_{23}\right)^{\mathrm{T}} & \Sigma^t_{33} \end{bmatrix}$$

The upper left two-by-two block corresponds to the subsystem of the two robots involved in the relative measurement. It is updated according to the exact EKF, thus it is positive-definite. To complete the proof, we have to show that the other two two-by-two blocks are also positive-definite. We show the proof for

$$\tilde{\Sigma}^{t+1}_{\{1,3\}\{1,3\}} := \begin{bmatrix} \Sigma^{t+1}_{11} & \tilde{\Sigma}^{t+1}_{13} \\ \left(\tilde{\Sigma}^{t+1}_{13}\right)^{\mathrm{T}} & \Sigma^t_{33} \end{bmatrix} \tag{30}$$

For $\tilde{\Sigma}^{t+1}_{\{2,3\}\{2,3\}}$, the lower right block of $\tilde{\Sigma}^{t+1}$, the proof is equivalent. The matrix $\tilde{\Sigma}^{t+1}_{\{1,3\}\{1,3\}}$ is positive definite if and only if both Σ^{t+1}_{11} and

$$\Sigma^t_{33} - \left(\tilde{\Sigma}^{t+1}_{13}\right)^{\mathrm{T}} \left(\Sigma^{t+1}_{11}\right)^{-1} \tilde{\Sigma}^{t+1}_{13}$$

(the Schur complement of Σ^{t+1}_{11} in $\tilde{\Sigma}^{t+1}_{\{1,3\}\{1,3\}}$) are positive-definite. The matrix Σ^{t+1}_{11} is positive-definite, as it results from the exact EKF update. Before we investigate the Schur complement, we recall the exact update equation for the whole system

$$\Sigma^{t+1} = \Sigma^t - \underbrace{\Sigma^t F^{\mathrm{T}} (F \Sigma^t F^{\mathrm{T}} + Q)^{-1} F \Sigma^t}_{:=P}$$

where P is positive-semidefinite. Thus, we can write

$$\Sigma^{t+1}_{11} = \Sigma^t_{11} - P_{11} \tag{31}$$

with some positive-semidefinite matrix P_{11}.

Now, we investigate the Schur complement

$$\begin{aligned}
&\Sigma^t_{33} - \left(\tilde{\Sigma}^{t+1}_{13}\right)^{\mathrm{T}} \left(\Sigma^{t+1}_{11}\right)^{-1} \tilde{\Sigma}^{t+1}_{13} \\
&\overset{(28)}{=} \Sigma^t_{33} - \lambda \left(\tilde{\Sigma}^{t+1}_{13}\right)^{\mathrm{T}} \underbrace{\left(\Sigma^{t+1}_{11}\right)^{-1} \Sigma^{t+1}_{11}}_{\mathbb{I}} \left(\Sigma^t_{11}\right)^{-1} \Sigma^t_{13} \\
&\overset{(28)}{=} \Sigma^t_{33} - \lambda^2 \left(\Sigma^{t+1}_{11} \left(\Sigma^t_{11}\right)^{-1} \Sigma^t_{13}\right)^{\mathrm{T}} \left(\Sigma^t_{11}\right)^{-1} \Sigma^t_{13} \\
&= \Sigma^t_{33} - \lambda^2 \left(\Sigma^t_{13}\right)^{\mathrm{T}} \left(\Sigma^t_{11}\right)^{-1} \Sigma^{t+1}_{11} \left(\Sigma^t_{11}\right)^{-1} \Sigma^t_{13} \\
&\overset{(31)}{=} \Sigma^t_{33} - \lambda^2 \left(\Sigma^t_{13}\right)^{\mathrm{T}} \left(\Sigma^t_{11}\right)^{-1} \Sigma^t_{13} \\
&\quad + \lambda^2 \left(\Sigma^t_{13}\right)^{\mathrm{T}} \left(\Sigma^t_{11}\right)^{-1} P_{11} \left(\Sigma^t_{11}\right)^{-1} \Sigma^t_{13} \\
&= (1 - \lambda^2) \Sigma^t_{33} \\
&\quad + \lambda^2 \underbrace{\left(\Sigma^t_{33} - \left(\Sigma^t_{13}\right)^{\mathrm{T}} \left(\Sigma^t_{11}\right)^{-1} \Sigma^t_{13}\right)}_{=:P'} \\
&\quad + \lambda^2 \underbrace{\left(\Sigma^t_{13}\right)^{\mathrm{T}} \left(\Sigma^t_{11}\right)^{-1} P_{11} \left(\Sigma^t_{11}\right)^{-1} \Sigma^t_{13}}_{=:P''}
\end{aligned}$$

The matrix P' is the Schur complement of Σ^t_{11} in $\Sigma^t_{\{1,3\}\{1,3\}}$. It is positive-definite according to the requirement of the proposition. The matrix P'' is positive-semidefinite and Σ^t_{33} is positive-definite. For $|\lambda| < 1$, we have the sum of two positive-definite matrices and a positive-semidefinite matrix, and for $|\lambda| = 1$, we have the sum of a positive-definite matrix and a positive-semidefinite matrix. In both cases, the sum is positive-definite.

Appendix C: Proof of Lemma 1

Before we can prove Lemma 1, we have to introduce the following general lemma related to the usual Cauchy–Schwarz inequality for scalar products, and a resulting corollary.

Lemma 2. *For d-dimensional random variables X and Y, with definition*

$$\begin{aligned}
\langle X, Y \rangle &:= cov(X, Y) \\
&= E\left((X - E(X))(Y - E(Y))^{\mathrm{T}}\right)
\end{aligned}$$

the induced matrix norm $\|\cdot\|$, and the assumption that $\langle X, X \rangle$ and $\langle Y, Y \rangle$ are positive-definite, we obtain

$$\|\langle Y, X \rangle\| \leq \|\langle Y, Y \rangle^{\frac{1}{2}}\| \|\langle X, X \rangle^{\frac{1}{2}}\| \tag{32}$$

Proof. With Lavergne (2008, Corollary 1), we obtain

$$v^{\mathrm{T}} \langle X, Y \rangle \langle Y, Y \rangle^{-1} \langle Y, X \rangle v \leq v^{\mathrm{T}} \langle X, X \rangle v$$

for all $v \in \mathbb{R}^d$. With the help of the symmetric and positive-definite square root, this is equivalent to

$$\left\| \langle Y, Y \rangle^{-\frac{1}{2}} \langle Y, X \rangle v \right\| \leq \left\| \langle X, X \rangle^{\frac{1}{2}} v \right\|$$

for all $v \in \mathbb{R}^d$. Using the induced matrix norm, we obtain

$$\begin{aligned}\|\langle Y,X\rangle\| &:= \max_{\|v\|=1} \|\langle Y,X\rangle v\| \\ &\leq \|\langle Y,Y\rangle^{\frac{1}{2}}\| \max_{\|v\|=1} \|\langle Y,Y\rangle^{-\frac{1}{2}}\langle Y,X\rangle v\| \\ &\leq \|\langle Y,Y\rangle^{\frac{1}{2}}\| \|\langle X,X\rangle^{\frac{1}{2}}\|\end{aligned}$$

□

Corollary 2. *With the notation of the present paper, we obtain*

$$\begin{aligned}&\|\Sigma_{jk} - \Sigma_{ji}\Sigma_{ii}^{-1}\Sigma_{ik}\| \\ \leq &\|\left(\Sigma_{jj} - \Sigma_{ji}\Sigma_{ii}^{-1}\Sigma_{ij}\right)^{\frac{1}{2}}\| \|\left(\Sigma_{kk} - \Sigma_{ki}\Sigma_{ii}^{-1}\Sigma_{ik}\right)^{\frac{1}{2}}\|\end{aligned}$$

Proof. For three jointly distributed d-dimensional random variables X, Y, and Z, we decompose Z and Y in an X-dependent part and a part orthogonal to X with respect to $\langle\cdot,\cdot\rangle$, namely

$$\begin{aligned}Z &= \langle Z,X\rangle\langle X,X\rangle^{-1}X + \widehat{Z} \\ Y &= \langle Y,X\rangle\langle X,X\rangle^{-1}X + \widehat{Y}\end{aligned}$$

so that $\langle\widehat{Y},X\rangle = \langle\widehat{Z},X\rangle = 0$. Here, we assume that

$$\begin{pmatrix}\langle X,X\rangle & \langle X,Y\rangle & \langle X,Z\rangle \\ \langle Y,X\rangle & \langle Y,Y\rangle & \langle Y,Z\rangle \\ \langle Z,X\rangle & \langle Z,Y\rangle & \langle Z,Z\rangle\end{pmatrix} = \begin{pmatrix}\Sigma_{ii} & \Sigma_{ij} & \Sigma_{ik} \\ \Sigma_{ji} & \Sigma_{jj} & \Sigma_{jk} \\ \Sigma_{ki} & \Sigma_{kj} & \Sigma_{kk}\end{pmatrix}$$

is positive-definite, which ensures the existence of $\langle X,X\rangle^{-1}$. The goal is to apply (32) to $\langle\widehat{Y},\widehat{Z}\rangle$. Therefore, we need to prove that $\langle\widehat{Z},\widehat{Z}\rangle$ and $\langle\widehat{Y},\widehat{Y}\rangle$ are positive-definite. We obtain that

$$\langle\widehat{Z},\widehat{Z}\rangle = \langle Z,Z\rangle - \langle Z,X\rangle\langle X,X\rangle^{-1}\langle X,Z\rangle$$

and

$$\langle\widehat{Y},\widehat{Y}\rangle = \langle Y,Y\rangle - \langle Y,X\rangle\langle X,X\rangle^{-1}\langle X,Y\rangle$$

are Schur complements of a positive-definite matrix and, thus, they are positive-definite. Thus, we can apply (32) and obtain

$$\begin{aligned}&\|\Sigma_{jk} - \Sigma_{ji}\Sigma_{ii}^{-1}\Sigma_{ik}\| \\ =&\|\langle\widehat{Y},Z\rangle\| \\ =&\|\langle\widehat{Y},\widehat{Z}\rangle\| \\ \leq&\|\langle\widehat{Y},\widehat{Y}\rangle^{\frac{1}{2}}\| \|\langle\widehat{Z},\widehat{Z}\rangle^{\frac{1}{2}}\| \\ =&\|\left(\Sigma_{jj} - \Sigma_{ji}\Sigma_{ii}^{-1}\Sigma_{ij}\right)^{\frac{1}{2}}\| \cdot \|\left(\Sigma_{kk} - \Sigma_{ki}\Sigma_{ii}^{-1}\Sigma_{ik}\right)^{\frac{1}{2}}\|\end{aligned}$$

□

Lemma 1 immediately follows from Corollary 2 together with the general equality $\|B^{1/2}\| = \|B\|^{1/2}$ for a positive-definite matrix B and the spectral norm.

5.1. Proof of Corollary 1

We use

$$\Sigma_{ik}^{t+1} = (\mathbb{I} - K_iF_i)\,\Sigma_{ik}^t - K_iF_j\Sigma_{jk}^t$$

from (20) and

$$\tilde{\Sigma}_{ik}^{t+1} = (\mathbb{I} - K_iF_i)\,\Sigma_{ik}^t - K_iF_j\Sigma_{ji}^t\left(\Sigma_{ii}^t\right)^{-1}\Sigma_{ik}^t$$

from (26) to obtain

$$\begin{aligned}&\|\Sigma_{ik}^{t+1} - \tilde{\Sigma}_{ik}^{t+1}\| \\ =&\left\|K_iF_j\left(\Sigma_{jk}^t - \Sigma_{ji}^t\left(\Sigma_{ii}^t\right)^{-1}\Sigma_{ik}^t\right)\right\| \\ \leq&\|K_iF_j\|\sqrt{\epsilon_{ji}\epsilon_{ki}}\end{aligned}$$

For the inequality, we use Lemma 1.