

# Multirobot localization with unknown variance parameters using iterated Kalman filtering

Gianluigi Pillonetto and Stefano Carpin

**Abstract**—The multirobot localization problem is solved in this paper using an innovative approach related to Tikhonov regularization. We release the requirement that robots are equipped with sensors to estimate their own motion, as well as the requirement that covariance matrices describing system and measure noises are perfectly known. Robots are assumed to have a single sensor returning noisy measurements of mutual distances while they move along unknown paths. The proposed algorithm estimates online both the robots' poses as well as the unknown covariance parameters. In addition to the classical iterations of the well known iterated Kalman filter, we include iterations that propagate an approximation of the posterior marginal densities of the unknown variances. Simulation results provide evidence that the algorithm is capable of accurately estimating the variances online while at the same time keeping the localization error bounded.

## I. INTRODUCTION AND RELATED WORK

Networked robotics is envisioned to be one of the key research areas in the future of robotics research. Application scenarios are abundant: surveillance, environmental monitoring, wildlife observation, and homeland security are just few examples. In all these applications, the ability to localize sensors or robots is a key feature. While the NAVSTAR satellites implementing the GPS network offer in general a reliable solution, there exist significant fractions of the globe where this technology is not usable (e.g. indoor areas, forests, and urban canyons) In these situations, instead of relying on the existence of an external set of landmarks, the system can act as its own supporting structure, having some of its elements (sensors or robots) acting as reference points. This idea, that is one approach to solve the *cooperative localization* problem, was pioneered by Kurazume et al. [1]. In this paper we start from their problem setting, i.e. a scenario where certain robots move around, while certain others stay stationary to serve as reference points for *exteroceptive* measurements. We attack the multirobot localization problem removing some of the assumptions usually found in the related literature. First, we abandon the hypothesis that nodes are equipped with *proprioceptive* sensors returning estimates of their motion. We assume that the only form of external measurement available is the mutual distance between nodes, similarly to what can be obtained with time of flight sensors. Second, we drop the assumption that the covariance matrices characterizing state evolution and measurement noises are perfectly known. The core of the paper is indeed an algorithm that solves the

localization problem estimating these variances online. For this kind of problem, the simplest and most widely used technique consists of augmenting the dimension of the state space by interpreting variance parameters as additional states. Then, state/parameter estimation is performed by means of an extended Kalman filter [2, 3]. The on-line approach that we propose instead exploits and extends classical iterated Kalman filtering equations [4] in order to propagate an approximation of the marginal posterior of the unknown parameters over time.

Being one of the most fundamental problems in mobile robotics, *localization* has enjoyed a great amount of research. The recent book by Thrun et al. [5] provides a comprehensive coverage of the field. Most successful approaches to the localization problem can be regarded as special instances of the Bayes filter. Localization based on the extended Kalman filter (EKF) or any of its refinements has been popular since quite some time [6]. The posterior probability distribution is approximated by a Gaussian, i.e. a single modal distribution. A limitation affecting EKF based methods is the so called *data association*, i.e. the problem of matching perceived data with features in the environment. When the association cannot be univocally determined, multiple hypothesis should be tracked, and hence multi modal distributions should be used instead. Although this constraint can be somehow overcome using Gaussian mixtures, this is one of the main limitations of these methods. EKF based approaches have been used to solve both the single and multiple robot localization problem. Roumeliotis and Bekey [7] implemented a distributed solution for a system composed of robots able to sense each other as well as their own motion. Smaller communicating filters implement a global estimator. These methods are good in solving the *pose tracking* problem, i.e. to iteratively estimate the pose of the system once its initial value is known with a certain uncertainty, provided that the data association problem does not play a dominant role. *Monte Carlo* localization overcomes the single hypothesis limitation inherent in EKF based filters [8][9]. By using particle filters the authors propose an algorithm able to approximate an arbitrary posterior probability distribution. Monte Carlo based localization methods assume the availability of a *motion model* as well as of a *sensor model*. Monte Carlo methods are suitable to solve not only the pose tracking problem. They solve also the *global localization* problem, i.e. a localization task where initial pose estimation is not known at all, as well as the *kidnapped robot* problem, i.e. a localization problem where the initial pose estimation is completely wrong.

G. Pillonetto is with the Department of Information Engineering, University of Padova, Italy [giapi@dei.unipd.it](mailto:giapi@dei.unipd.it)

S. Carpin is with the School of Engineering, University of California, Merced, USA [carpin@ieee.org](mailto:carpin@ieee.org)

## II. STATEMENT OF THE PROBLEM

### A. General problem formulation

In the sequel all vectors are column vectors. Some of the notation used in the following of the paper is described in the following table:

Notation	Description
$p \in \mathbf{Z}_+$	dimension of parameter space
$n \in \mathbf{Z}_+$	dimension of state space
$m \in \mathbf{Z}_+$	dimension of output space
$\theta \in \mathbf{R}^p$	parameter vector
$x_0 \in \mathbf{R}^n$	initial value of state vector
$z_k \in \mathbf{R}^m$	measurement vector
$h_k : \mathbf{R}^n \rightarrow \mathbf{R}^m$	expected value of $z_k$ given $x_k$
$R_k : \mathbf{R}^p \rightarrow \mathbf{R}^{m \times m}$	variance of $z_k$ given $x_k$
$g_k : \mathbf{R}^n \rightarrow \mathbf{R}^n$	expected value of $x_k$ given $x_{k-1}$
$Q_k : \mathbf{R}^p \rightarrow \mathbf{R}^{n \times n}$	variance of $x_k$ given $x_{k-1}$

In particular, we assume that  $R_k$  and  $Q_k$  can be dependent on  $\theta$ . Now, consider the following state-space model

$$\begin{cases} x_k = g_k(x_{k-1}) + \omega_k \\ z_k = h_k(x_k) + \nu_k \end{cases} \quad (1)$$

*Assumption 1:* Given  $\theta$ , the random vectors  $\{\nu_k\}$  and  $\{\omega_k\}$  are all mutually independent, Gaussian, with mean zero and variance  $\{R_k(\theta)\}$  and  $\{Q_k(\theta)\}$  respectively.

Our problem consists of estimating on-line  $\{x_k\}$  and the unknown components of vector  $\theta$  starting from  $\{z_k\}$ .

### B. Multirobot localization problem

In the sequel we often use  $bd\{A_1, \dots, A_p\}$  to denote the square block matrix in which the diagonal elements are the square matrices  $A_1, \dots, A_p$  and the off-diagonal elements are 0.

We introduce the multirobot localization problem as a special instance of the problem previously stated. We have

$$\begin{cases} a_k^{(i)} = a_{k-1}^{(i)} + v_k^{(i)} \cos(c_k^{(i)}) + \omega_{k,a}^{(i)} \\ b_k^{(i)} = b_{k-1}^{(i)} + v_k^{(i)} \sin(c_k^{(i)}) + \omega_{k,b}^{(i)} \\ c_k^{(i)} = c_{k-1}^{(i)} + d_k^{(i)} + \omega_{k,c}^{(i)} \\ z_k^{(i,j)} = \Psi(a_k^{(i)}, a_k^{(j)}, b_k^{(i)}, b_k^{(j)}) + \nu_k^{(i,j)} \quad i \neq j \end{cases} \quad (2)$$

being

$$\Psi(a_k^{(i)}, a_k^{(j)}, b_k^{(i)}, b_k^{(j)}) = \sqrt{(a_k^{(i)} - a_k^{(j)})^2 + (b_k^{(i)} - b_k^{(j)})^2} \quad (3)$$

where integers  $i$  and  $j$  range from 1 to  $N$  ( $N$  is the number of nodes),  $a_k^{(i)}$  and  $b_k^{(i)}$  are the coordinates in the plane of the  $i$ -th robot at instant  $k$  while  $c_k^{(i)}$  represents its orientation. Furthermore,  $v_k^{(i)}$  and  $d_k^{(i)}$  denote robot's translational and rotational velocity, respectively. The reader should recognize in equation 2 the standard equations for a differential drive robot moving on a plane. In terms of model 1 the following correspondences hold (in this case  $m = N \times (N - 1)$ )

$$x_k = \left( a_k^{(1)} \ b_k^{(1)} \ c_k^{(1)} \ \dots \ a_k^{(N)} \ b_k^{(N)} \ c_k^{(N)} \right)^T$$

$$\omega_k = \left( \omega_{k,a}^{(1)} \ \omega_{k,b}^{(1)} \ \omega_{k,c}^{(1)} \ \dots \ \omega_{k,a}^{(N)} \ \omega_{k,b}^{(N)} \ \omega_{k,c}^{(N)} \right)^T$$

$$z_k = \left( z_k^{(1,2)} \ z_k^{(1,3)} \ \dots \ z_k^{(2,1)} \ z_k^{(2,3)} \ \dots \ z_k^{(N,N-1)} \right)^T$$

where the trailing  $T$  indicates the transpose. It is assumed that two kind of nodes may be present in the model. When node  $i$  acts as a landmark,  $v_k^i$  as well as the variances of the components of  $\omega_k$  are set to zero for every  $k$ . Otherwise, covariance of the random vectors  $[\omega_{k,a}^{(i)}, \omega_{k,b}^{(i)}, \omega_{k,c}^{(i)}]$  is given by

$$Q_k^{(i)}(\theta) = \begin{pmatrix} \sigma_a^2 & 0 & 0 \\ 0 & \sigma_b^2 & 0 \\ 0 & 0 & \sigma_c^2 \end{pmatrix}$$

where variances  $\sigma_a^2$ ,  $\sigma_b^2$  and  $\sigma_c^2$  could be components of  $\theta$ . We thus have  $Q_k = bd\{Q_k^{(1)}, Q_k^{(2)}, \dots, Q_k^{(N)}\}$ . It is assumed known which node is moving and which is a landmark. As concerns  $R_k$ , no particular structure is stated here. Notice that if both  $i$  and  $j$  represent landmarks, whose positions are assumed known, and the  $\{R_k\}$  do not depend on unknown components of  $\theta$ , measurements of their distance become irrelevant and can be removed from the model. In addition, depending on the particular problem under study, not all the measurements  $\{z_k\}$  listed above could be available, as also specified later on.

### C. Connection with Tikhonov regularization

We provide a connection between a simplified off-line version of our state-estimation problem and Tikhonov regularization. To simplify the notation, dependence of matrices on  $\theta$  is omitted in the rest of this sub-section. Let index  $k$  vary between 1 and  $K$  inclusive. We define the vectors  $X \in \mathbf{R}^{nK}$ ,  $Z \in \mathbf{R}^{mK}$  as follows

$$X = (x_1 \dots x_K)^T \quad Z = (z_1 \dots z_K)^T$$

The functions  $h : \mathbf{R}^{nK} \rightarrow \mathbf{R}^{mK}$  and  $g : \mathbf{R}^{nK} \rightarrow \mathbf{R}^{nK}$  are instead given by

$$h(X) = (h_1(x_1) \dots h_K(x_K))^T$$

$$g(X) = (g_1(x_0) \dots g_K(x_{K-1}))^T$$

Matrices  $Q \in \mathbf{R}^{(nK) \times (nK)}$  and  $R \in \mathbf{R}^{(mK) \times (mK)}$  are defined as follows

$$Q = bd\{Q_1, Q_2, \dots, Q_K\}$$

$$R = bd\{R_1, R_2, \dots, R_K\}$$

If  $W$  is a symmetric positive definite matrix,  $v \in \mathbf{R}^p$  and  $W \in \mathbf{R}^{p \times p}$ , we define  $\|v, W\|^2 = v^T W v$ . The following lemma provides an expression for the joint density of states and data. It can be easily proved using the probability chain

rule and considering that, by assumption,  $\{x_k\}$  is Markov and process  $\{z_k\}$  is white conditioned on the state process.

*Lemma 2:* If Assumption 1 holds, the negative log of the joint density of  $Z$  and  $X$  is given by the function  $L : \mathbf{R}^{nK} \times \mathbf{R}^{mK} \rightarrow \mathbf{R}$  defined by

$$\begin{aligned} L(X, Z) &= \frac{1}{2} \log \det(2\pi R) + \frac{1}{2} \|Z - h(X), R^{-1}\|^2 \\ &+ \frac{1}{2} \log \det(2\pi Q) + \frac{1}{2} \|X - g(X), Q^{-1}\|^2 \end{aligned} \quad (4)$$

The following theorem provides the desired connection between a version of the off-line multirobot localization problem and Tikhonov estimation [10].

*Theorem 3:* Consider model in eq.(2) with null initial condition. Assume also that no landmark nodes are present and that the  $N$  moving nodes are subject only to random inputs, i.e. the translational and rotational velocities  $\{v_k^{(i)}\}$  and  $\{a_k^{(i)}\}$  are all set to zero. In addition, let  $\sigma^2$  be the variance of the independent noises  $\nu_k^{(i,j)}$ . Then, conditioned on the values of the transition and measurement noise variances and on  $\{z_k\}$ , the maximum a posteriori estimates of  $\{a_k^{(i)}\}$  and  $\{b_k^{(i)}\}$  minimize the following objective function

$$\begin{aligned} &\sum_{k=1}^K \sum_{i,j,i \neq j} \left( z_k^{(i,j)} - \Psi \left( a_k^{(i)}, a_k^{(j)}, b_k^{(i)}, b_k^{(j)} \right) \right)^2 + \\ &+ \gamma_1 \sum_{i=1}^N (a^{(i)})^T F^T F a^{(i)} + \gamma_2 \sum_{i=1}^N (b^{(i)})^T F^T F b^{(i)} \end{aligned} \quad (5)$$

where  $\gamma_1 = \frac{\sigma^2}{\sigma_a^2}$ ,  $\gamma_2 = \frac{\sigma^2}{\sigma_b^2}$  and

$$a^{(i)} = \left( a_1^{(i)} a_2^{(i)} \dots a_K^{(i)} \right)^T \quad b^{(i)} = \left( b_1^{(i)} b_2^{(i)} \dots b_K^{(i)} \right)^T$$

while  $F$  is a  $K \times K$  lower triangular Toeplitz matrix whose first column is equal to  $[1 \quad -1 \quad 0 \quad \dots \quad 0]^T$ , i.e.  $F$  represents a discrete-time derivator.

The previous result is of interest since it shows how off-line localization problem enjoys theoretical interpretation in terms of regularization theory. In fact, the solution of an instance of the problem can be obtained by optimizing an objective function containing two kind of contrasting factors. The first term in equation 5 accounts for experimental evidence, while the second and third are penalty terms on the energy of the discrete derivative of the robot trajectory. Given this correspondence, we can think of  $\gamma_1$  and  $\gamma_2$  as two regularization parameters which have to properly balance the relative importance of the sum of squared residuals against the smoothness index. In regularization theory it is well known that the problem of finding a robust criterium for selecting the regularization parameter is key [11]. In fact, on one hand a too large value will correspond to a model that is not complex enough to describe experimental data and will return biased estimates. On the other hand, a too small value will produce over-fitting and will increase estimates variance [12]. When linear measurement models are used,

cross-validation or maximum likelihood represent robust and commonly used criteria to tune this crucial parameter. However, in our case the problem is more difficult since

- the measurements model as well as the state dynamics in the full model in eq.(2) are generally nonlinear
- more than one unknown regularization parameter may be present in the estimator, depending on the unknown components contained in  $\theta$
- estimation of robot trajectory and of regularization parameters must be performed in an on-line manner

The approach developed to overcome the above difficulties is described in the following Section.

### III. THE ALGORITHM

In this Section we describe the algorithm for on-line estimation of the state-vectors  $\{x_k\}$  and  $\theta$ . In the following, we think of  $\theta$  as a normal random variable whose realizations are constrained to be nonnegative and whose (possibly noninformative) density is specified before seeing the measurements and updated as time index  $k$  increases.

#### A. Time update

We use  $\hat{\theta}_{k-1}$  to denote the estimate of  $\theta$  at instant  $(k-1)$  while  $\hat{\Sigma}_{\theta,k-1}$  denotes the covariance matrix of the associated error. The estimate of  $x_{k-1}$  based on measurements collected up to instant  $(k-1)$  is instead  $\hat{x}_{k-1|k-1}$  with error covariance denoted  $\hat{\Sigma}_{k-1|k-1}$ . It is worth stressing that both  $\hat{x}_{k-1|k-1}$  and  $\hat{\Sigma}_{k-1|k-1}$  are quantities which are not interpreted as dependent on  $\theta$  at this level of the algorithm.

As usual, we use the linearized dynamics of the system around the current estimate to perform the time-update. We use the notation  $g'(y)$  to denote the Jacobian of a function  $g$  evaluated at  $y$ . Then

$$x_k \approx g_k(\hat{x}_{k-1|k-1}) + g'_k(\hat{x}_{k-1|k-1})(x_{k-1} - \hat{x}_{k-1|k-1}) + \omega_k \quad (6)$$

It comes from eq. (6) that  $x_k$  can be approximated as a Gaussian with mean given by  $\hat{x}_{k|k-1} = g_k(\hat{x}_{k-1|k-1})$  whose covariance matrix becomes now dependent on  $\theta$  and is given by

$$\begin{aligned} \hat{\Sigma}_{k|k-1}(\theta) &= g'_k(\hat{x}_{k-1|k-1}) \hat{\Sigma}_{k-1|k-1} (g'_k(\hat{x}_{k-1|k-1}))^T \\ &+ Q_k(\theta) \end{aligned}$$

#### B. Measurement update for state and parameter vector

The update problem amounts to achieve a better estimate for  $x_k$  and  $\theta$ , and the corresponding covariance matrices of the error affecting the estimates, starting from  $z_k$ .

*Proposition 4:* Assume that  $x_k$  is Gaussian with mean  $\hat{x}_{k|k-1}$  and covariance  $\hat{\Sigma}_{k|k-1}(\theta)$ . Then, conditioned on  $\theta$  and  $z_k$ , the maximum a posteriori estimate of  $x_k$  is the minimizer of

$$\begin{aligned}
l(x_k, z_k | \theta) &= \frac{1}{2} \log \det(2\pi R_k(\theta)) \\
&+ \frac{1}{2} \log \det(2\pi \hat{\Sigma}_{k|k-1}(\theta)) + \frac{1}{2} \|z_k - h(x_k), R_k(\theta)^{-1}\|^2 \\
&+ \frac{1}{2} \|x_k - \hat{x}_{k|k-1}, \hat{\Sigma}_{k|k-1}(\theta)^{-1}\|^2 \quad (7)
\end{aligned}$$

being  $l(x_k, z_k | \theta)$  the minus log of the joint density of  $x_k$  and  $z_k$  conditioned on  $\theta$ .

Optimization of the objective  $l(\cdot)$  can be performed by exploiting a Gauss-Newton method, see e.g. [13]. Recall that the iterated Kalman filter (IKF) update is exactly a Gauss-Newton method, i.e. given the same initial point the sequence of iterates generated by the Gauss-Newton method and the sequence of iterates generated by IKF are identical, see [4]. Recall also that IKF corresponds to the extended Kalman filter (EKF) when one performs just a single iteration. Thus, an approximation of the minimizer of  $l(\cdot)$  can be obtained by defining inductively the sequences  $x^i$  and  $\Sigma^i$  as follows

$$\begin{aligned}
x^0 &= \hat{x}_{k|k-1} & \Sigma^0 &= \hat{\Sigma}_{k|k-1}(\theta) \\
x^{i+1} &= \hat{x}_{k|k-1} + K^i (z_k - h(x^i) - H_i (\hat{x}_{k|k-1} - x^i)) \\
\Sigma^{i+1} &= (I - K_i H_i) \hat{\Sigma}_{k|k-1}(\theta) \\
H_i &= h'(x^i) \\
K_i &= \hat{\Sigma}_{k|k-1}(\theta) H_i^T (H_i \hat{\Sigma}_{k|k-1}(\theta) H_i^T + R_k(\theta))^{-1}
\end{aligned}$$

After computing a sufficient number of iterations to reach convergence, values of  $x^i$  and  $\Sigma^i$  provide the updated estimate  $\hat{x}_{k|k}(\theta)$  and covariance matrix of the error  $\hat{\Sigma}_{k|k}(\theta)$ , respectively. However, both of these quantities depend on  $\theta$  and the question now arises as how to estimate the unknown parameter vector. Let  $\pi(x_k, z_k | \theta)$  be the joint density for  $x_k$  and  $z_k$  conditioned on  $\theta$ , i.e.  $\pi(x_k, z_k | \theta) = \exp(-l(x_k, z_k | \theta))$ . Let instead  $\pi(z_k | \theta)$  denote the marginal likelihood of  $\theta$ , i.e.

$$\pi(z_k | \theta) = \int \pi(x_k, z_k | \theta) dx_k \quad (8)$$

This integral is useful since it allows to remove biases in parameter estimation [14]. If  $\pi_{k-1}(\theta)$  denotes the current "prior" for  $\theta$ , i.e. a Gaussian with mean  $\hat{\theta}_{k-1}$  and covariance  $\hat{\Sigma}_{\theta, k-1}$ , our target estimate for  $\theta$  is

$$\hat{\theta} = \operatorname{argmin}_{\theta} - \log [\pi(z_k | \theta) \pi_{k-1}(\theta)]$$

under nonnegative constraints for the components of  $\theta$ . However, due to the nonlinear nature of function  $h$ , evaluation of  $\pi(z_k | \theta)$  for a given  $\theta$  requires solution of an integral in eq.(8) which in general is analytically intractable. Now, we show how computations performed by IKF for a given  $\theta$  may provide an approximation for such an integral. To this aim let's consider the affine approximation of  $l(x_k, z_k | \theta)$  for  $x_k$  near  $y$  defined by

$$\begin{aligned}
\tilde{l}(x_k, y, z | \theta) &= \frac{1}{2} \log [\det(2\pi R_k(\theta) \det(2\pi \hat{\Sigma}_{k|k-1}(\theta)))] \\
&+ \frac{1}{2} \|z_k - h(y) - h'(y)(x_k - y), R_k(\theta)^{-1}\|^2 \\
&+ \frac{1}{2} \|x_k - \hat{x}_{k|k-1}, \hat{\Sigma}_{k|k-1}(\theta)^{-1}\|^2
\end{aligned}$$

The determinant of the Hessian of  $\tilde{l}$  with respect to  $x_k$  can be easily computed and turns out to be

$$\det \left[ \partial_{x_k}^2 \tilde{l}(x_k, y, z_k | \theta) \right] = \quad (9)$$

$$\det \left[ (h'(y))^T R_k^{-1}(\theta) h'(y) + \hat{\Sigma}_{k|k-1}^{-1}(\theta) \right]$$

The information matrix approximation for the marginal likelihood  $\pi(z_k | \theta)$  is denoted by  $\tilde{\pi}(z_k | \theta)$  and given by (see [15][16])

$$\det \left\{ \partial_{x_k}^2 \tilde{l}(x_k, \hat{x}_{k|k}(\theta), z_k | \theta) / (2\pi) \right\}^{-1/2} e^{-l(\hat{x}_{k|k}(\theta), z_k | \theta)} \quad (10)$$

In view of eq.(9) and eq.(10) it comes that such an approximation requires  $\hat{x}_{k|k}(\theta)$  and  $h'(\hat{x}_{k|k}(\theta))$  which represent quantities returned by IKF. In other words, for every  $\theta$  value the IKF can be used to evaluate an objective function whose optimization provides the estimate of  $\theta$ . Thus, we are in the position to define the estimate of  $\theta$ , computed after seeing data up to instant  $k$ , as the solution to the problem

$$\hat{\theta}_k = \operatorname{argmin}_{\theta} - \log [\tilde{\pi}(z_k | \theta) \pi_{k-1}(\theta)] \quad (11)$$

while  $\hat{\Sigma}_{\theta, k}$  is defined by

$$\hat{\Sigma}_{\theta, k} = \left( -\partial_{\theta}^2 \log \left[ (\tilde{\pi}(z_k | \hat{\theta}_k) \pi_{k-1}(\hat{\theta}_k)) \right] \right)^{-1} \quad (12)$$

and can be computed numerically. This completes the update for parameter  $\theta$ . Finally, the entire measurement update is completed by setting estimate of  $x_k$  to  $\hat{x}_{k|k}(\hat{\theta}_k)$  with covariance matrix of the error given by  $\hat{\Sigma}_{k|k}(\hat{\theta}_k)$ .

#### IV. EXPERIMENTAL RESULTS

To test the relative performance of our method, we consider two different cases where either matrices  $\{Q_k\}$  or  $\{R_k\}$  may depend on components of  $\theta$  which are unknown. We consider several Monte Carlo simulations involving a large set of localization problems. Every Monte Carlo study consists of 1000 runs and at every Monte Carlo run mobile nodes perform 100 steps. Parameter  $\theta$  is given an initial noninformative distribution which is always updated in the first 10 steps, and only one out of 10 steps in the remaining iterations. As concerns the moving nodes, their initial positions are randomly drawn from a uniform density with support on a square centered in the origin with side length equal to 2. In the model of eq. (2) all the  $v_k^{(i)}$  are constant and set to 0.4 while  $d_k^{(i)}$  equals 0.35 or  $-0.35$  with probability 0.5. These assumption are consistent with currently used mobile platforms. In fact, they amount to assuming that each node moves at most 0.4m and rotates at most 20 degrees

in each time step, which can be assumed to be 1s. These nominal values for the translational and rotational velocities, as well as the initial positions of the nodes, are assumed known during each simulated study. Finally, three nodes act as landmarks, being located so as to form an equilateral triangle with barycenter set in the origin and with side length equal to 40. Initial conditions of fixed and mobile nodes are assumed perfectly known. If  $a_{k,j}^{(i)}$  and  $b_{k,j}^{(i)}$  are the true coordinates of the  $i$ -th node at the  $j$ -th Monte Carlo run, we use  $\hat{a}_{k,j}^{(i)}$  and  $\hat{b}_{k,j}^{(i)}$  to denote the corresponding estimates. The errors affecting the estimates, as a function of  $j$ , are then denoted  $\{\varepsilon^j\}$  and given by

$$\varepsilon^j = \frac{1}{100} \sum_{k=1}^{100} \sqrt{\sum_{i=1}^N \left( a_{k,j}^{(i)} - \hat{a}_{k,j}^{(i)} \right)^2 + \left( b_{k,j}^{(i)} - \hat{b}_{k,j}^{(i)} \right)^2}$$

#### A. Possibly unknown transition noise covariance

We first consider case studies where  $\theta \in \mathbb{R}^3$  and its components are the three variances of the transition noise, i.e. the diagonal elements of matrices  $\{Q_k\}$ . Such variances may be unknown. Matrices  $\{R_k\}$  are instead assumed perfectly known. Several simulated studies (each consisting of 1000 Monte Carlo runs) are considered, differing in the number of moving nodes, which may be 3, 6 or 9, and in the values of the components of  $\theta$ , which may be  $[0.05^2 \ 0.05^2 \ 0.05^2]$ ,  $[0.1^2 \ 0.1^2 \ 0.1^2]$  or  $[0.2^2 \ 0.2^2 \ 0.2^2]$ . In each case, measurements regarding distance among nodes are corrupted by a Gaussian noise with a constant percentage coefficient of the error ( $CV\%$ ) equal to 10. Robot trajectories and orientation are then estimated by setting  $\theta_3$  to its true value<sup>1</sup> and

- assuming  $\theta_1$  and  $\theta_2$  known and using EKF (estimator named  $\Gamma_0$ )
- assuming  $\theta_1$  and  $\theta_2$  known and using IKF ( $\Gamma_1$ )
- assuming  $\theta_1$  and  $\theta_2$  unknown and using the new approach described in the previous Section based on IKF ( $\Gamma_2$ )
- setting  $\theta_1$  and  $\theta_2$  to their true values divided by 16 (SD of the transition noise is thus 4 times larger than the true value) and using IKF ( $\Gamma_3$ )
- setting  $\theta_1$  and  $\theta_2$  to their true values multiplied by 16 (SD of the transition noise is thus 4 times smaller than the true value) and using IKF ( $\Gamma_4$ )

In Fig. 1 we report values  $Av(\varepsilon)$ , as a function of the number of moving nodes employed in the Monte Carlo study, achieved by the five estimators when the true SD of the transition noise is 0.05 (top), 0.1 (middle) or 0.2 (bottom). First of all, it is worth stressing how the performance of  $\Gamma_0$  is significantly inferior than that regarding  $\Gamma_1$  and  $\Gamma_2$ . Obtained results thus show how in localization problems IKF (with variances of the transition noise either known or unknown and estimated by our approach) is much more robust than EKF (with variances assumed known). We now compare the performance of IKF-based estimators, i.e. of  $\{\Gamma_i\}$ . The

<sup>1</sup>experiments shown it is not worth trying to estimate variance  $\theta_3$ , due to the limited amount of output information

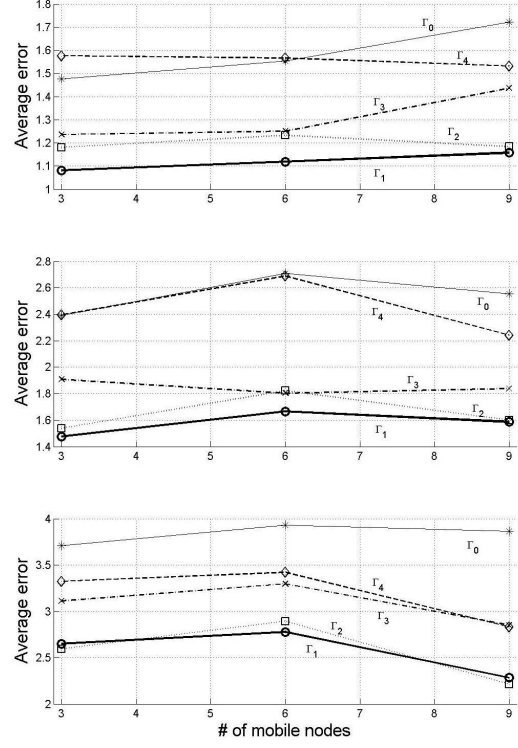


Fig. 1. Possibly unknown transition noise covariance parameters: average error, as a function of the moving nodes.

goodness of the proposed approach clearly surfaces since the performance of  $\Gamma_2$  always appears close to that of  $\Gamma_1$ . One can also see that the performance of  $\Gamma_2$  is closer and closer to that of  $\Gamma_1$  as the number of moving nodes increases. This points out how the proposed method is able to efficiently extract more and more useful information on the statistics of the state dynamics as the size of the data set augments. On the other hand, the use of  $\Gamma_3$  or  $\Gamma_4$  leads to a degradation of the quality of the estimates. However, it is worth noticing how the average error relative to  $\Gamma_3$  remains quite close to that obtained by employing the true variances. An interpretation of this result is that output measurements are able to compensate quite well overestimation of transition noise variances. In the next subsection, it will be shown that the situation becomes much more critical when undermodeling affects the statistics of the measurement noise.

#### B. Possibly unknown measurement noise covariance

Now, we consider case studies where matrices  $\{Q_k\}$  are assumed perfectly known while matrices  $\{R_k\}$  may be partly unknown. In particular, system state has to be reconstructed from measurements of two different types. Moving robots may measure their mutual distances and the variance of the noise affecting such measurements is known. To be specific, this kind of output data are corrupted by a Gaussian noise with a constant  $CV\%$  equal to 10. Other available measurements regard the distance between mobile robots and landmarks. We assume that such data are corrupted by a

Gaussian noise with a constant  $CV\%$  that is function of the fixed node and assume values equal to 5, 10 or 20. These  $CV$  values represent the components of  $\theta \in \mathbb{R}^3$  and may be unknown. As in the previous case, several simulated studies (each consisting of 1000 Monte Carlo runs) are considered, differing in the number of moving nodes, which may be 3, 6 or 9, and in the variance of the transition noise, which is the same for every moving node and may assume values  $0.05^2$ ,  $0.1^2$  or  $0.2^2$ . Robot trajectories are estimated

- assuming  $\theta$  known and using EKF (estimator  $\Gamma_0$ )
- assuming  $\theta$  known and using IKF ( $\Gamma_1$ )
- assuming all the components of  $\theta$  unknown and using the novel estimation approach described in the previous Section based on IKF ( $\Gamma_2$ )
- setting all the components of  $\theta$  so as to define a  $CV\%$  of the error affecting the measurements between all the mobile and fixed nodes equal to 5 and using IKF ( $\Gamma_3$ )
- setting all the components of  $\theta$  so as to define a  $CV\%$  of the error affecting the measurements between all the mobile and fixed nodes equal to 20 and using IKF ( $\Gamma_4$ )

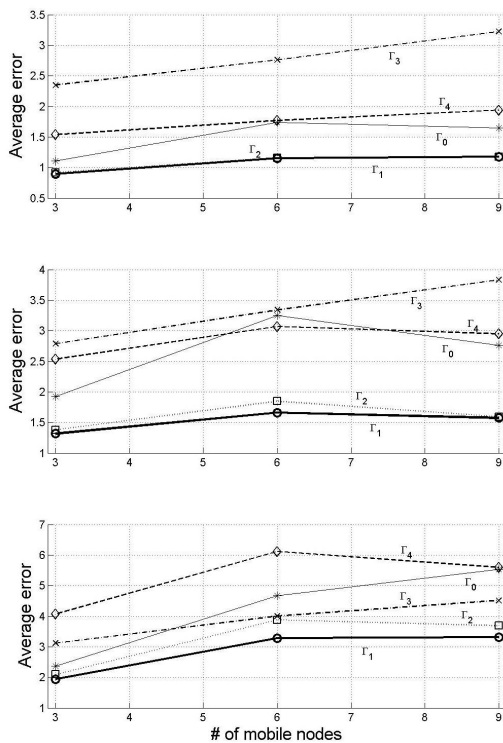


Fig. 2. Possibly unknown measurement noise covariance parameters: average error, as a function of the moving nodes, when the true SD of the transition noise is 0.05 (top), 0.1 (middle) or 0.2 (bottom).

Fig. 2 displays obtained results with the same rationale adopted in Fig. 1. Obtained results still point out how IKF-based schemes  $\Gamma_1$  and  $\Gamma_2$  are much more robust than  $\Gamma_0$  which relies upon EKF. As regards the performance of all the IKF-based estimators, it is remarkable the fact that  $\Gamma_2$  performs very well, returning an average error which is very similar to that achieved by  $\Gamma_1$ . On the other hand, one can

easily notice how the use of  $\Gamma_3$  or  $\Gamma_4$  leads to a large degradation of the quality of the estimates. As a matter of fact, the average error may turn out also 3 times larger than that obtained by using  $\Gamma_1$  or  $\Gamma_2$ .

## V. CONCLUSIONS

In this paper we have addressed the problem of cooperative localization of multi-robot systems. The presented approach is novel in many aspects. On the practical side, it drops the assumption of proprioceptive sensors. It further assumes that system and measurement noises are zero mean with covariances which may contain unknown parameters that are estimated online. On the theoretical side, the connection between the proposed estimation problem and Tikhonov regularization has been outlined. The importance of online parameters estimation is noteworthy. First, it relieves the task of estimating or measuring *a priori* system and measurement noises. We clearly show that the online algorithm outperforms estimation based on noise levels too high or too low, in particular when the statistics of the measurement sensors are uncertain. Second, in the light of recent results producing localization error bounds based on noise covariance matrices [17], the ability to precisely estimate these variances online allows the derivation of more precise bounds.

## REFERENCES

- [1] R. Kurazume and S. Hirose, "An experimental study of cooperative positioning system," *Autonomous Robots*, vol. 8, no. 1, pp. 43–52, 2000.
- [2] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Englewood Cliffs, N.J., USA: Prentice-Hall, 1979.
- [3] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*. Academic Press, New York, 1970.
- [4] B. Bell and F. Cathey, "The iterated kalman filter update as a gauss-newton method," *IEEE Trans. on Automatic Control*, vol. 38, no. 2, pp. 294–297, 1993.
- [5] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*. MIT Press, 2006.
- [6] J. Leonard and H. Durrant-Whyte, "Mobile robot localization by tracking geometric beacons," *IEEE Transaction on Robotics and Automation*, vol. 7, no. 3, pp. 376–382, 1991.
- [7] S. Roumeliotis and G. Bekey, "Distributed multirobot localization," *IEEE Transactions on Robotics and Automation*, vol. 18, no. 5, pp. 781–795, 2002.
- [8] C. Kwok, D. Fox, and M. Meila, "Real-time particle filters," *Proceedings of the IEEE*, vol. 92, no. 3, pp. 469–484, 2004.
- [9] S. Thrun, D. Fox, W. Burgard, and F. Dellart, "Robust monte carlo localization for mobile robots," *Artificial Intelligence*, vol. 128, no. 1–2, pp. 99–141, 2001.
- [10] A. Tikhonov and V. Arsenin, *Solutions of Ill-Posed Problems*. Washington, D.C.: Winston/Wiley, 1977.
- [11] T. Poggio and F. Girosi, "Networks for approximation and learning," *Proceedings of the IEEE*, vol. 78, pp. 1481–1497, 1990.
- [12] J. Rice, "Choice of smoothing parameter in deconvolution problems," *Contemp. Math.*, vol. 59, pp. 137–151, 1986.
- [13] D. Luenberger, *Linear and nonlinear programming*. Kluwer, 2003.
- [14] B. Bell, "The marginal likelihood for parameters in a discrete gauss-markov process," *IEEE Trans. on Signal Processing*, vol. 48, no. 3, pp. 870–873, 2000.
- [15] —, "Approximating the marginal likelihood estimate for models with random parameters," *Appl. Math. Comput.*, vol. 119, pp. 57–75, 2001.
- [16] B. Bell and G. Pilonetto, "Estimating parameters and stochastic functions of one variable using nonlinear measurements models," *Inverse Problems*, vol. 20, no. 3, pp. 627–646, 2004.
- [17] A. Mourikis and S. Roumeliotis, "Performance analysis of multirobot cooperative localization," *IEEE Transactions on robotics*, vol. 22, no. 4, pp. 666–681, 2006.