This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LCSYS.2021.3064931, IEEE Control Systems Letters

Robust Linearly Constrained Invariant Filtering for a Class of Mismatched Nonlinear Systems

Paul Chauchat, Jordi Vilà-Valls, Eric Chaumette

***Abstract*— Standard state estimation techniques require a perfect knowledge of the system's model, that is, process and measurement equations, inputs and the corresponding noise statistics. In practice this assumption does not hold and therefore robust filtering methods must be accounted for. A new approach to tackle a potential model mismatch via linear constraints within the Kalman filter, and its extended version (EKF), has been recently shown to be a very promising solution. This article further explores robust linearly constrained filtering in the context of the Invariant EKF (InEKF). With respect to the EKF, the InEKF is a recent filtering technique which has been shown to better handle the particular structure of a class of problems through the use of Lie groups. In this contribution, a new robust linearly constrained InEKF is introduced, together with particular parametric mismatched models and mitigation strategies through linear constraints. Numerical results for an illustrative navigation example are provided to show the performance improvement and support the discussion.**

***Index Terms*— Kalman filtering, Estimation, Robust control**

I. Introduction

STATE estimation, i.e, inference of dynamics states of a system from noisy observations, is a fundamental brick in a wide variety of applications such as robotics, tracking and navigation systems [1], [2]. For linear systems, the Kalman filter (KF) is the best linear minimum mean square error (MSE) estimator [2]. Many approaches were proposed to tackle nonlinear systems, such as the popular extended KF (EKF) [2], the family of sigma-point filters [3] or particle filters [4]. However, all these methods in their standard form assume a perfect knowledge of nonlinear process and measurement functions and their respective noise statistics (i.e., the two first moments for KF-type methods), which is not a realistic assumption in real-life applications. To overcome this limitation, a plethora of robust filtering techniques have been developed in the past decades, most of them focusing on counteracting unexpected or unknown noise behaviours and parameters: using noise covariance estimation techniques [5], the use of robust statistics to counteract outliers [6] or exploiting heavy-tailed distributions [7]. A number of strategies to design sliding-mode observers with fault estimation were also proposed [8], [9]. In contrast, few contributions explored how to counteract a mismatch on system process and measurement functions, thus being an important missing point, which is the main interest of this contribution.

Regarding the mitigation of system functions mismatch, a recent line of work explored the use of non-stationary linear equality constraints within the KF framework, for process and/or measurement model mismatch, generalizing the results in [10], and leading to a general linearly constrained KF (LCKF) formulation [11], [12]. This was then extended to mitigate a potential parametric mismatch in nonlinear systems through a linearly constrained EKF (LC-EKF) [13]. It has been shown in these contributions that using linear constraints is an effective way to robustify the KF.

In parallel, within the navigation community, new nonlinear filters appeared to better handle the nonlinear structure of the system under consideration, the so-called linear observed systems on groups [14]. In particular the Invariant EKF (InEKF), which came with strong theoretical guarantees [15], and has been successfully used in various applications [16], [17]. However, as the InEKF reduces the linearization errors, it should be all the more sensitive to a possible system model mismatch. Recently, it was proposed to apply the H_∞ methodology to the invariant filtering framework in order to compensate some disturbances on the process model [18], but this does not solve the system model mismatch problem.

Our main contributions explore the applicability of linear constraints as a model mismatch mitigation strategy for the class of systems handled by the InEKF:

- We derive a class of model mismatch based on the structure of linear observed systems on groups.
- We propose a mitigation strategy introducing a new robust linearly constrained InEKF (LC-InEKF).
- We illustrate the method and compare it to standard solutions on an illustrative navigation example.

II. Linear Observed Systems on Groups and Invariant Filtering

In this work, we consider a state $\boldsymbol{\chi}$ defined on a matrix Lie group G of dimension q. Its Lie algebra $\mathfrak{g}$ is identified with $\mathbb{R}^q$. The Lie algebra element associated with a vector $\boldsymbol{\xi} \in \mathbb{R}^q$ is noted $\boldsymbol{\xi}^\wedge$. Therefore we consider the Lie group exponential map to be defined as $\exp : \mathbb{R}^q \to G$, where $\exp(\boldsymbol{\xi}) = \exp_m(\boldsymbol{\xi}^\wedge)$, $\exp_m$ being the matrix exponential, and its (local) inverse being the $\log$. A last classical tool is the (inner) automorphism $\Psi_a \in Aut(G)$ defined for each $a \in G$ as $\Psi_a : g \mapsto aga^{-1}$. Its differential at the identity element $\mathbf{Id}$

This work has been partially supported by the DGA/AID funded projects 2018.60.0072.00.470.75.01 and 2019.65.0068.00.470.75.01.

Authors are with ISAE-SUPAERO/University of Toulouse, 10 avenue Edouard Belin, 31055 Toulouse, France; e-mail: paul.chauchat@isae-supaero.fr, jordi.vila-valls@isae-supaero.fr, eric.chaumette@isae-supaero.fr

of G is called *adjoint operator* and denoted by $\mathbf{Ad}_a : \mathfrak{g} \mapsto \mathfrak{g}$. This satisfies:

$$\forall a \in G, \mathbf{u} \in \mathfrak{g},\ a\exp(\mathbf{u})a^{-1} = \exp(\mathbf{Ad}_a\mathbf{u}) \tag{1a}$$

$$\Leftrightarrow a\exp(\mathbf{u}) = \exp(\mathbf{Ad}_a\mathbf{u})a, \tag{1b}$$

which gives a way to commute elements. We present hereafter the notion of linear observed systems on groups [14].

A. Uncertainties on Lie Groups and System Considered

To define random variables on Lie groups, we adopt the framework of [19]. The probability law $\boldsymbol{\chi} \sim \mathcal{N}_L(\bar{\boldsymbol{\chi}}, \mathbf{P})$ for the random variable $\boldsymbol{\chi} \in G$ is defined as $\boldsymbol{\chi} = \bar{\boldsymbol{\chi}}\exp(\boldsymbol{\xi})$, with $\boldsymbol{\xi} \sim \mathcal{N}(\mathbf{0}, \mathbf{P})$, where $\mathcal{N}(.,.)$ is the classical Gaussian distribution in Euclidean space and $\mathbf{P} \in \mathbb{R}^{q\times q}$ is a covariance matrix. In the following we consider a discrete trajectory of the state, $(\boldsymbol{\chi}_k)_{1\leq k\leq n}$, satisfying the system

$$\boldsymbol{\chi}_0 \sim \mathcal{N}_L(\bar{\boldsymbol{\chi}}, \mathbf{P}_0), \tag{2}$$

$$\boldsymbol{\chi}_k = f_{k-1}(\boldsymbol{\chi}_{k-1})\exp(\mathbf{w}_{k-1}), \tag{3}$$

$$\mathbf{y}_k = h_k(\boldsymbol{\chi}_k) + \mathbf{v}_k, \tag{4}$$

where f_{k-1} are group-affine dynamics and h_k equivariant outputs, two notions defined in the sequel; $\mathbf{w}_{k-1}$ and $\mathbf{v}_k$ are white noises of covariance $\mathbf{Q}_{k-1}$ and $\mathbf{N}_k$, respectively.

B. Group-affine Dynamics

Group affine dynamics were introduced in continuous time in [15], and in discrete time in [14] which ensure that, thanks to preintegration, we can focus on the latter only. The main idea is that they are to systems on Lie groups what affine dynamics are to systems on vector spaces.

Definition 1: A dynamics f is group-affine if there exist a group automorphism $\Theta : G \to G$, that is, a bijection satisfying $\Theta(ab) = \Theta(a)\Theta(b)$ and an element $\boldsymbol{\kappa} \in G$ s.t.

$$\forall a \in G,\ f(a) = \Theta(a)\ \boldsymbol{\kappa}. \tag{5}$$

The Lie group Lie algebra correspondence ensures that no linearization error is made during the propagation.

Theorem 1 ([14]): Let f be group-affine on G, and Φ be the matrix such that $\forall\boldsymbol{\xi} \in \mathfrak{g},\ \Theta(\exp(\boldsymbol{\xi})) = \exp(\Phi\boldsymbol{\xi})$ with the notations of (5). Then, letting $\mathbf{F} = \mathbf{Ad}_{\boldsymbol{\kappa}^{-1}}\Phi$,

$$\forall\boldsymbol{\chi} \in G,\ \boldsymbol{\xi} \in \mathfrak{g},\ f(\boldsymbol{\chi}\exp(\boldsymbol{\xi})) = f(\boldsymbol{\chi})\exp(\mathbf{F}\boldsymbol{\xi}), \tag{6}$$

As it corresponds to the shift of $f(\boldsymbol{\chi})$ under a perturbation $\boldsymbol{\xi}$, $\mathbf{F}$ is referred to as the Jacobian of f.

C. Outputs as Group Actions

Linear observed systems on groups are defined as having group-affine dynamics, and observations modeled as group actions [14]. The measurement belongs to another space, Y, on the elements of which the group "acts". Think for instance of the matrix-vector product. Here we only consider left group actions on a vector space. That is, the noise free observation can be modeled as the action on a given $\mathbf{b}_k$:

$$\mathbf{y}_k = h(\boldsymbol{\chi}_k, \mathbf{b}_k) = h_k(\boldsymbol{\chi}_k) = \boldsymbol{\chi}_k \star \mathbf{b}_k, \tag{7}$$

where $\star : G \times Y \to Y$ satisfies, for all $\mathbf{b} \in Y$, $\boldsymbol{\chi}_1, \boldsymbol{\chi}_2 \in G$:

$$\boldsymbol{\chi}_1 \star (\boldsymbol{\chi}_2 \star \mathbf{b}) = (\boldsymbol{\chi}_1\boldsymbol{\chi}_2) \star \mathbf{b}, \quad \mathbf{Id} \star \mathbf{b} = \mathbf{b}. \tag{8}$$

D. Invariant Filtering

The (left-)InEKF was developed precisely to estimate the state $\boldsymbol{\chi}$ of linear observed systems on a group G. The state is estimated through a mean $\hat{\boldsymbol{\chi}} \in G$, and a covariance matrix $\mathbf{P}$, such that $\boldsymbol{\chi} \sim \mathcal{N}_L(\hat{\boldsymbol{\chi}}, \mathbf{P})$, that is:

$$\boldsymbol{\chi} = \hat{\boldsymbol{\chi}}\exp(\boldsymbol{\xi}) \Leftrightarrow \boldsymbol{\xi} = \log(\hat{\boldsymbol{\chi}}^{-1}\boldsymbol{\chi}), \tag{9}$$

where $\boldsymbol{\xi} \sim \mathcal{N}(\mathbf{0}, \mathbf{P})$. In particular, $\boldsymbol{\xi} = \log(\exp(\boldsymbol{\xi}))$. We see, first, that (6) propagates (9) keeping the same form.

Second, the main advantage of group actions as outputs is that, similarly to group-affine dynamics, their linearization is particularly suited to the considered parametrization, albeit a change on the noise. Indeed, following [14], consider the nonlinear innovation defined as $\mathbf{i}_k = \hat{\boldsymbol{\chi}}_{k|k-1}^{-1} \star \mathbf{y}_k - \mathbf{b}_k$. Proposition 12 of [14] in turn brings

$$\mathbf{i}_k = \exp(\boldsymbol{\xi}_k) \star \mathbf{b}_k - \mathbf{b}_k + \hat{\mathbf{v}}_k \approx \mathbf{H}_k\boldsymbol{\xi}_k + \hat{\mathbf{v}}_k, \tag{10}$$

where $\mathbf{H}_k$ is the Jacobian at zero of $\boldsymbol{\xi} \mapsto \exp(\boldsymbol{\xi}) \star \mathbf{b}_k$ (remember that $\mathbf{Id} \star \mathbf{b} = \mathbf{b}$), and $\hat{\mathbf{v}}_k$ is the noise modified by the impact of $\hat{\boldsymbol{\chi}}_{k|k-1}^{-1}$, which must be specified in each case, for which we note $\hat{\mathbf{v}}_k \sim \mathcal{N}(\mathbf{0}, \widehat{\mathbf{N}}_k)$. See Section VI-C and [14] for examples. Thus, the linearization does not depend on the current estimate, although it is not exact as for group-affine dynamics.

As for the EKF, (10) motivates using standard Kalman theory to estimate the error $\boldsymbol{\xi}$. To preserve the form of the random variable, the exponential update is used. This is summarized in the following propagation-correction sequence

$$\hat{\boldsymbol{\chi}}_{k|k-1} = \Theta(\hat{\boldsymbol{\chi}}_{k-1|k-1})\boldsymbol{\kappa}, \tag{11a}$$

$$\mathbf{P}_{k|k-1} = \mathbf{F}_{k-1}\mathbf{P}_{k-1}\mathbf{F}_{k-1}^T + \mathbf{Q}_{k-1}, \tag{11b}$$

$$\mathbf{S}_k = \mathbf{H}_k\mathbf{P}_{k|k-1}\mathbf{H}_k^T + \widehat{\mathbf{N}}_k, \tag{11c}$$

$$\mathbf{K}_k = \mathbf{P}_{k|k-1}\mathbf{H}_k^T\mathbf{S}_k^{-1}, \tag{11d}$$

$$\hat{\boldsymbol{\chi}}_{k|k} = \hat{\boldsymbol{\chi}}_{k|k-1}\exp\left[\mathbf{K}_k(\hat{\boldsymbol{\chi}}_{k|k-1}^{-1}\mathbf{y}_k - \mathbf{b}_k)\right], \tag{11e}$$

$$\mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_{k|k-1}. \tag{11f}$$

The main advantage of the InEKF is that the involved jacobians $\mathbf{F}_{k-1}$ and $\mathbf{H}_k$ do not depend on the current estimate $\hat{\boldsymbol{\chi}}$. While this allowed proving strong theoretical guarantees [15], it also means that the filter might be all the more sensitive to model mismatch (with respect to the standard KF and EKF). This is the problem which is explored in the remainder of this work.

III. Mismatches for Systems on Groups

System mismatches can be of various origins and forms. Here we concentrate on misspecifications of group-affine dynamics (5) and group actions (7).

A. Mismatch on the Process Mean Noise Value

Consider the mismatched group-affine dynamics:

$$\boldsymbol{\chi}_k = \Theta(\boldsymbol{\chi}_{k-1})\boldsymbol{\kappa}\boldsymbol{\delta}_{\boldsymbol{\kappa}}\exp(\mathbf{w}_{k-1}), \tag{12}$$

where the mismatch arises from $\boldsymbol{\delta_\kappa}$. Then, the nonlinear error $\boldsymbol{\xi}$ evolves as, denoting $\circ$ the composition of applications

$$\begin{aligned}\boldsymbol{\xi}_{k|k-1} &= \log(\Psi_{\boldsymbol{\kappa}^{-1}} \circ \Theta(\hat{\boldsymbol{\chi}}_{k-1|k-1}^{-1}\boldsymbol{\chi}_{k-1})\boldsymbol{\delta_\kappa}\exp(\mathbf{w}_{k-1}))\\ &= \log(\exp(\mathbf{Ad}_{\boldsymbol{\kappa}^{-1}}\Phi\boldsymbol{\xi}_{k-1|k-1})\boldsymbol{\delta_\kappa}\exp(\mathbf{w}_{k-1}))\\ &\approx \mathbf{F}_{k-1}\boldsymbol{\xi}_{k-1|k-1} + \mathbf{d}_{\boldsymbol{\kappa}} + \mathbf{w}_{k-1}. \end{aligned} \tag{13}$$

with $\mathbf{F}_{k-1}$ as in (6), $\mathbf{d}_{\boldsymbol{\kappa}} = \log(\boldsymbol{\delta_\kappa})$, and using (9) and the first-order approximation $\log(\exp(\boldsymbol{\xi})\exp(\boldsymbol{\zeta})) \approx \boldsymbol{\xi} + \boldsymbol{\zeta}$ as in [15]. Notably, we assume the mismatch $\boldsymbol{\delta_\kappa}$ to be small.

B. Mismatch on Group Actions

1) Error on the Group Action Model

If the true measurement relies on a different group action than the assumed one, then (10) does not hold anymore, and extra terms will appear. Following [13], we consider a parametric family of group actions $(\star_s)$, with a vectorial parameter s for the sake of simplicity. The measurement function is extended as $\mathbf{y} = h(\boldsymbol{\chi}, \mathbf{b}, s) = \boldsymbol{\chi} \star_s \mathbf{b}$. Consider an assumed model $\star_{\hat{s}}$, while the true model is $\star_s$, with $s = \hat{s} + \delta_s$. We thus have $\mathbf{y} \approx \boldsymbol{\chi} \star_{\hat{s}} \mathbf{b} + \frac{\partial h}{\partial s}\big|_{\boldsymbol{\chi},\mathbf{b},\hat{s}} \delta_s$. Therefore, the innovation (10) becomes, with the assumed model and in the noise free case:

$$\begin{aligned}\mathbf{i}_k &\approx \hat{\boldsymbol{\chi}}_{k|k-1}^{-1} \star_{\hat{s}} \left(\boldsymbol{\chi}_k \star_{\hat{s}} \mathbf{b}_k + \frac{\partial h}{\partial s}\bigg|_{\boldsymbol{\chi}_k,\mathbf{b}_k,\hat{s}} \delta_s\right) - \mathbf{b}_k\\ &\approx \hat{\boldsymbol{\chi}}_{k|k-1}^{-1} \star_{\hat{s}} \boldsymbol{\chi}_k \star_{\hat{s}} \mathbf{b}_k - \mathbf{b}_k\\ &+ \frac{\partial h}{\partial \mathbf{b}}\bigg|_{\hat{\boldsymbol{\chi}}_{k|k-1}^{-1}, \boldsymbol{\chi}_k \star_{\hat{s}} \mathbf{b}_k, \hat{s}} \frac{\partial h}{\partial s}\bigg|_{\boldsymbol{\chi}_k,\mathbf{b}_k,\hat{s}} \delta_s\\ &\approx \mathbf{H}_k\boldsymbol{\xi}_{k|k-1} + \frac{\partial h}{\partial \mathbf{b}}\bigg|_{\hat{\boldsymbol{\chi}}_{k|k-1}^{-1}, \boldsymbol{\chi}_k \star_{\hat{s}} \mathbf{b}_k, \hat{s}} \frac{\partial h}{\partial s}\bigg|_{\boldsymbol{\chi}_k,\mathbf{b}_k,\hat{s}} \delta_s \end{aligned} \tag{14}$$

The noise is added as in (10), since it is not impacted.

2) Error from the Action of a Perturbation

Another form of mismatch can be expressed as the action of some group element on the output, that is, in the noise free case, $\mathbf{y} = \boldsymbol{\delta_\chi} \star \boldsymbol{\chi} \star \mathbf{b}$, $\boldsymbol{\delta_\chi} \in G$. Therefore, the innovation becomes $\mathbf{i}_k = \hat{\boldsymbol{\chi}}_{k|k-1}^{-1}\boldsymbol{\delta_\chi}\boldsymbol{\chi}_k \star \mathbf{b}_k - \mathbf{b}_k$. The induced error is obtained thanks to the adjoint operator. Noting $\boldsymbol{\delta_\chi} = \exp(\boldsymbol{\mu})$, we have

$$\begin{aligned}\mathbf{i}_k &= \left(\hat{\boldsymbol{\chi}}_{k|k-1}^{-1}\boldsymbol{\chi}_k\exp(\mathbf{Ad}_{\boldsymbol{\chi}_k^{-1}}\boldsymbol{\mu})\right) \star \mathbf{b}_k - \mathbf{b}_k\\ &\approx \exp(\boldsymbol{\xi}_{k|k-1}) \star \left(\mathbf{b}_k + \mathbf{H}_k\mathbf{Ad}_{\boldsymbol{\chi}_k^{-1}}\boldsymbol{\mu}\right) - \mathbf{b}_k\\ &\approx \mathbf{H}_k\boldsymbol{\xi}_{k|k-1} + \mathbf{H}_k\mathbf{Ad}_{\boldsymbol{\chi}_k^{-1}}\boldsymbol{\mu}, \end{aligned} \tag{15}$$

using (8) and that $\frac{\partial h}{\partial \boldsymbol{\chi}}\big|_{\exp(\boldsymbol{\xi}),\cdot} \approx \frac{\partial h}{\partial \boldsymbol{\chi}}\big|_{\mathbf{Id},\cdot} = \mathbf{Id}$.

If both mismatch forms are simultaneously present, the additional innovation error is taken as the addition of the individual errors in (14) and (15). That is, we neglect the effect of higher-order terms involving both δ_s and $\boldsymbol{\mu}$.

IV. On the Mismatch Mitigation

In the following we show the model mismatch impact, and propose a mitigation strategy by exploiting linear constraints.

A. Mismatch Impact on the Estimation Error

The different mismatches previously described can greatly harm the estimation process. Indeed, after a propagation-update sequence using an arbitrary gain $\mathbf{L}$, the error becomes, according to (13), (14) and (15):

$$\begin{aligned}\boldsymbol{\xi}_{k|k}(\mathbf{L}) &= \log(\exp(-\mathbf{L}\mathbf{i}_k)\hat{\boldsymbol{\chi}}_{k|k-1}^{-1}\boldsymbol{\chi}_k) \approx \boldsymbol{\xi}_{k|k-1} - \mathbf{L}\mathbf{i}_k\\ &\approx (\mathbf{I} - \mathbf{L}\mathbf{H}_k)(\mathbf{F}_{k-1}\boldsymbol{\xi}_{k-1|k-1} + \mathbf{w}_{k-1})\\ &- \mathbf{L}\mathbf{v}_k - \boldsymbol{\varepsilon}_k(\mathbf{L}), \end{aligned} \tag{16}$$

$$\begin{aligned}\boldsymbol{\varepsilon}_k(\mathbf{L}) = \mathbf{L}\Bigg(&\frac{\partial h}{\partial \mathbf{b}}\bigg|_{\hat{\boldsymbol{\chi}}_{k|k-1}^{-1}}, \boldsymbol{\chi}_k \star_{\hat{s}} \mathbf{b}_k, \hat{s}\ \frac{\partial h}{\partial s}\bigg|_{\boldsymbol{\chi}_k,\mathbf{b}_k,\hat{s}} \delta_s\\ &+\mathbf{H}_k\mathbf{Ad}_{\boldsymbol{\chi}_k^{-1}}\boldsymbol{\mu}\Bigg) - (\mathbf{I} - \mathbf{L}\mathbf{H}_k)\mathbf{d}_{\boldsymbol{\kappa}}. \end{aligned} \tag{17}$$

where $\boldsymbol{\varepsilon}_k(\mathbf{L})$ is the additional error induced by the mismatch.

B. Model Mismatch Mitigation

Without priors on the mismatches, [12] proposes to impose $\mathbf{L}\boldsymbol{\varepsilon}_k(\mathbf{L}) = \mathbf{0}$ in order to mitigate their impact. However, since $\boldsymbol{\varepsilon}_k(\mathbf{L})$ depends on $\boldsymbol{\chi}_k$, whose value is unknown, $\mathbf{L}\boldsymbol{\varepsilon}_k(\mathbf{L}) = \mathbf{0}$ could only be achieved by nullifying the term for all possible values, which would be too restrictive. Therefore, we follow here [13], and aim at mitigating these biases by constraining the gain $\mathbf{L}$ to satisfy $\mathbb{E}[\boldsymbol{\varepsilon}_k(\mathbf{L})] = \mathbf{0}$, where the mean is taken over the only probabilistic term, $\boldsymbol{\chi}_k$. However, this mean is also unknown. Thus, we further assume that the mean value of the true state $\boldsymbol{\chi}_k$ is well approximated by $\hat{\boldsymbol{\chi}}_{k|k-1}$, as done in [13] for the LC-EKF. That means that it is subject to the mismatch (12), which was supposed small and should not, in one time step, induce a large error.

From (17), we can see that we need two constraints to be able to mitigate the mismatch impact on the estimation error,

$$(\mathbf{I} - \mathbf{L}\mathbf{H}_k)\mathbf{d}_{\boldsymbol{\kappa}} = \mathbf{0} \Leftrightarrow \mathbf{L}\mathbf{H}_k\mathbf{d}_{\boldsymbol{\kappa}} = \mathbf{d}_{\boldsymbol{\kappa}}, \tag{18a}$$

$$\begin{aligned}&\mathbf{L}\left[\tfrac{\partial h}{\partial \mathbf{b}}\big|_{\hat{\boldsymbol{\chi}}_{k|k-1}^{-1}, \hat{\boldsymbol{\chi}}_{k|k-1}\star_{\hat{s}}\mathbf{b}_k, \hat{s}}\ \tfrac{\partial h}{\partial s}\big|_{\hat{\boldsymbol{\chi}}_{k|k-1},\mathbf{b}_k,\hat{s}} \quad \mathbf{H}_k\mathbf{Ad}_{\hat{\boldsymbol{\chi}}_{k|k-1}^{-1}}\boldsymbol{\mu}\right]\\ &= \mathbf{L}\mathbf{dh}_k = \mathbf{0}. \end{aligned} \tag{18b}$$

These constraints can be aggregated in a compact form $\mathbf{L}\boldsymbol{\Delta}_k = \mathbf{T}_k$, where

$$\boldsymbol{\Delta}_k = \begin{bmatrix}\mathbf{H}_k\mathbf{d}_{\boldsymbol{\kappa}} & \mathbf{dh}_k\end{bmatrix}, \tag{18c}$$

$$\mathbf{T}_k = \begin{bmatrix}\mathbf{d}_{\boldsymbol{\kappa}} & \mathbf{0}\end{bmatrix}. \tag{18d}$$

Then, the goal is to obtain the filter gain $\mathbf{L}_k$ which verifies these constraints,

$$\begin{aligned}&\mathbf{L}_k = \arg\min_{\mathbf{L}}\left\{\mathbf{P}_{k|k}(\mathbf{L})\right\} \text{ s.t. } \mathbf{L}\boldsymbol{\Delta}_k = \mathbf{T}_k,\\ &\mathbf{P}_{k|k}(\mathbf{L}) = \mathbb{E}\left[\boldsymbol{\xi}_{k|k}(\mathbf{L})\boldsymbol{\xi}_{k|k}(\mathbf{L})^T\right], \end{aligned} \tag{19}$$

computed from a "constrained" KF update, to be detailed in the next section. However, as for the standard (In)EKF, since the gain is based on first-order approximations, the minimum will not be reached, but the constraint will be satisfied.

V. Robust Linearly Constrained InEKF

From the previous analysis we have all the theoretical results needed to build the new LC-InEKF. As previously stated, it is unlikely that the practitioner has a full knowledge of the system, and therefore we want to cope with the situation where there is a true model, $\mathcal{M}_T$, and a mismatched/assumed one, $\mathcal{M}_A$, accounting for the mismatches described in Section III:

$$\mathcal{M}_A : \begin{cases} \boldsymbol{\chi}_k = \Theta(\boldsymbol{\chi}_{k-1})\boldsymbol{\kappa}\exp(\mathbf{w}_{k-1}), \\ \mathbf{y}_k = \boldsymbol{\chi}_k \star_{\hat{s}} \mathbf{b}_k + \mathbf{v}_k, \end{cases} \tag{20}$$

$$\mathcal{M}_T : \begin{cases} \boldsymbol{\chi}'_k = \Theta(\boldsymbol{\chi}'_{k-1})\boldsymbol{\kappa}\boldsymbol{\delta}_{\boldsymbol{\kappa}}\exp(\mathbf{w}_{k-1}), \\ \mathbf{y}_k = \boldsymbol{\delta}_{\boldsymbol{\chi}} \star_s \boldsymbol{\chi}'_k \star_s \mathbf{b}_k + \mathbf{v}_k, \end{cases} \tag{21}$$

where the mismatch arises because of $\boldsymbol{\delta}_{\boldsymbol{\kappa}}$, $s = \hat{s} + \delta_s$ and $\boldsymbol{\delta}_{\boldsymbol{\chi}}$. Since the InEKF of $\boldsymbol{\chi}_k$ is based on the measurements and our knowledge of the system model, any mismatch between the true model $\mathcal{M}_T$ and the assumed one $\mathcal{M}_A$ leads to a suboptimal filter, and possibly to a filter with bad performance, as the discrepancy between the two models increases. The estimation error for the mismatched InEKF $\boldsymbol{\xi}_{k|k}(\mathbf{L})$ is given in (16), and the additional error term $\varepsilon_k(\mathbf{L})$ induced by the mismatch is detailed in (17). In order to mitigate such impact we can look for the set of constraints $\mathbf{L}\boldsymbol{\Delta}_k = \mathbf{T}_k$ which leads to $\mathbb{E}[\varepsilon_k(\mathbf{L})] = \mathbf{0}$. These constraints are derived in (18). Finally, the new robust LC-InEKF is formulated as follows: at every time k,

Prediction using $\mathcal{M}_A$:

$$\textit{Compute } \hat{\boldsymbol{\chi}}_{k|k-1}, \mathbf{P}_{k|k-1}, \mathbf{S}_k, \mathbf{K}_k, \textit{ from } (11) \tag{22a}$$

Linear Constraints

$$\boldsymbol{\Delta}_k = \begin{bmatrix}\mathbf{H}_k\mathbf{d}_{\boldsymbol{\kappa}} & \mathbf{dh}_k\end{bmatrix}, \ \mathbf{T}_k = \begin{bmatrix}\mathbf{d}_{\boldsymbol{\kappa}} & \mathbf{0}\end{bmatrix},$$

$$\boldsymbol{\Gamma}_k = \mathbf{T}_k - \mathbf{K}_k\boldsymbol{\Delta}_k, \ \Psi_k = \boldsymbol{\Delta}_k^T(\mathbf{S}_k)^{-1}\boldsymbol{\Delta}_k,$$

Constrained Gain

$$\mathbf{L}_k = \mathbf{K}_k + \boldsymbol{\Gamma}_k\Psi_k^{-1}\boldsymbol{\Delta}_k^T(\mathbf{S}_k)^{-1}, \tag{22b}$$

Constrained Filter Update using $\mathcal{M}_A$

$$\hat{\boldsymbol{\chi}}_{k|k} = \hat{\boldsymbol{\chi}}_{k|k-1}\exp\left[\mathbf{L}_k(\hat{\boldsymbol{\chi}}_{k|k-1}^{-1} \star_{\hat{s}} \mathbf{y}_k - \mathbf{b}_k)\right], \tag{22c}$$

$$\mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_{k|k-1} + \boldsymbol{\Gamma}_k\Psi_k^{-1}\boldsymbol{\Gamma}_k^T,$$

Notice that $\mathbf{K}_k$ is the standard "unconstrained" Kalman gain in (11d) computed from the assumed mismatched model, as well as the filter prediction and update. By imposing the linear equality constraints (18), the estimate obtained with the mismatched model (21) is matched to the true observation (20). Indeed, the LC-InEKF minimises the MSE associated to the true state $\boldsymbol{\chi}_k$, matching the true observations to the assumed model.

VI. Systems on Groups in Practice: 2D and 3D Localisation

A. The 2D and 3D Rigid Transformation Groups

The groups of 2D and 3D rigid transformations are particularly well-suited for modeling localisation problems [20]. Denoted as $SE(n)$ with $n \in \{2, 3\}$, they are the groups of orientation-position pairs $(\mathbf{R}, \mathbf{x})$, where $\mathbf{R} \in SO(n)$, $\mathbf{x} \in \mathbb{R}^n$, and are classically used to represent a mobile system's frame. They form matrix Lie groups, since they can be represented by matrices, using the homogeneous coordinates: $(\mathbf{R}, \mathbf{x}) \equiv \begin{bmatrix}\mathbf{R} & \mathbf{x} \\ \mathbf{0} & 1\end{bmatrix}$. This readily provides the composition and inversion rules: $(\mathbf{R}_1, \mathbf{x}_1)(\mathbf{R}_2, \mathbf{x}_2) = (\mathbf{R}_1\mathbf{R}_2, \mathbf{R}_1\mathbf{x}_2 + \mathbf{x}_1)$ and $(\mathbf{R}, \mathbf{x})^{-1} = (\mathbf{R}^T, -\mathbf{R}^T\mathbf{x})$. The exponential, logarithm and adjoint operators for $SE(n)$ and its algebra $\mathfrak{se}(n)$ have closed forms, which we give for the 2D case, as it will serve as our running example. They are defined for $\boldsymbol{\chi} \in SE(2)$ and $\boldsymbol{\xi} \in \mathfrak{se}(2)$ as $\exp(\boldsymbol{\xi}) = (\mathbf{R}(\xi_0), \mathbf{V}_{\xi_0}\begin{bmatrix}\xi_1 & \xi_2\end{bmatrix}^T)$, $\log(\boldsymbol{\chi}) = \begin{bmatrix}\theta & (\mathbf{V}_\theta^{-1}x)^T\end{bmatrix}^T$, $\mathbf{Ad}_{\boldsymbol{\chi}} = \begin{bmatrix}1 & \mathbf{0} \\ -\mathbf{J}x & \mathbf{R}(\theta)\end{bmatrix}$, where $\mathbf{V}_\alpha = \frac{1}{\alpha}\begin{bmatrix}\sin(\alpha) & -1 + \cos(\alpha) \\ 1 - \cos(\alpha) & \sin(\alpha)\end{bmatrix}$, $\mathbf{J} = \begin{bmatrix}0 & -1 \\ 1 & 0\end{bmatrix}$, and $\mathbf{R}(\theta)$ is the 2D rotation matrix of angle θ. We will conveniently identify $\mathbf{R}(\theta_j) = \mathbf{R}_j$ and its associated scalar angle θ_j (since matrices are upper case bold it will always be clear which one is being considered).

B. Odometry is Group-affine

A propagation model for 2D or 3D odometry given angular and linear increments, $\boldsymbol{\Omega}_{k-1}$ and $\mathbf{u}_{k-1}$, is given by

$$\begin{aligned}(\mathbf{R}_k, \mathbf{x}_k) &= (\mathbf{R}_{k-1}\boldsymbol{\Omega}_{k-1}, \mathbf{x}_{k-1} + \mathbf{R}_{k-1}\mathbf{u}_k) \\ &= (\mathbf{R}_{k-1}, \mathbf{x}_{k-1})(\boldsymbol{\Omega}_{k-1}, \mathbf{u}_k).\end{aligned} \tag{23}$$

Identifying with (5) brings $\Theta = \mathbf{Id}$ and $\boldsymbol{\kappa} = (\boldsymbol{\Omega}_{k-1}, \mathbf{u}_k)$. Thus, the propagation Jacobian is $\mathbf{Ad}_{(\boldsymbol{\Omega}_{k-1}^T, -\boldsymbol{\Omega}_{k-1}^T\mathbf{u}_{k-1})}$.

C. Some Group Actions on SE(n)

A particular but useful class of group actions is given by

$$(\mathbf{R}, \mathbf{x}) \star \mathbf{b} = \mathbf{x} + \mathbf{R}\mathbf{b}. \tag{24}$$

Using the matrix form, this is the same as computing $\boldsymbol{\chi}\begin{pmatrix}\mathbf{b}^T & 1\end{pmatrix}^T$, and deleting its last element, formulation used in [15]. Here, we extend this class to a family of group actions based on a scaling of the position part, $\mathbf{x}$:

$$(\mathbf{R}, \mathbf{x}) \star_\epsilon \mathbf{b} = (1 + \epsilon)\mathbf{x} + \mathbf{R}\mathbf{b}. \tag{25}$$

It is indeed a group action, for a *fixed* ϵ, as we have for two elements $\boldsymbol{\chi}_i = (\mathbf{R}_i, \mathbf{x}_i)$, $i = 1, 2$,

$$\begin{aligned}\boldsymbol{\chi}_1 \star_\epsilon (\boldsymbol{\chi}_2 \star_\epsilon \mathbf{b}) &= (\mathbf{R}_1, \mathbf{x}_1) \star_\epsilon ((1+\epsilon)\mathbf{x}_2 + \mathbf{R}_2\mathbf{b}) \\ &= (1+\epsilon)\mathbf{x}_1 + \mathbf{R}_1((1+\epsilon)\mathbf{x}_2 + \mathbf{R}_2 b) \\ &= (1+\epsilon)(\mathbf{x}_1 + \mathbf{R}_1\mathbf{x}_2) + \mathbf{R}_1\mathbf{R}_2\mathbf{b} \\ &= (\boldsymbol{\chi}_1\boldsymbol{\chi}_2) \star_\epsilon \mathbf{b}.\end{aligned} \tag{26}$$

For this class, the innovation is written

$$\mathbf{i}_k = \hat{\boldsymbol{\chi}}_{k|k-1}^{-1} \star_\epsilon \mathbf{b}_k - \mathbf{b}_k = \hat{\mathbf{R}}_{k|k-1}^T((1+\epsilon)(\mathbf{y}_k - \hat{\mathbf{x}}_{k|k-1})) - \mathbf{b}_k. \tag{27}$$

In terms of the error $\boldsymbol{\xi}$, this becomes $\mathbf{i}_k = \exp(\boldsymbol{\xi}_k) \star_\epsilon \mathbf{b}_k - \mathbf{b}_k + \hat{\mathbf{R}}_{k|k-1}^T\mathbf{v}_k$. Therefore, the modified noise is $\hat{\mathbf{v}}_k = \hat{\mathbf{R}}_{k|k-1}^T\mathbf{v}_k \sim \mathcal{N}(\mathbf{0}, \hat{\mathbf{R}}_{k|k-1}^T\mathbf{N}_k\hat{\mathbf{R}}_{k|k-1})$. Note that, if $\mathbf{N}_k$ is isotropic, then $\hat{\mathbf{v}}_k$ and $\mathbf{v}_k$ share the same distribution.

The first-order expansion of $\exp(\boldsymbol{\xi})$ used to obtain $\mathbf{H}_k$ depends on n in general. In 2D and 3D, one has $\mathbf{H}_k = \begin{bmatrix}(\mathbf{b})_\times & \mathbf{I}_n\end{bmatrix}$, where $(\mathbf{b})_\times = \mathbf{J}\mathbf{b}$ if $n = 2$, and the $n \times n$ skew-symmetric matrix associated to $\mathbf{b}$ if $n = 3$.

D. Studied Mismatches

1) Mismatch for the odometry: Consider the 2D navigation problem in which, similarly to (25), the input linear velocity is scaled by ϵ_u, for instance due to deflated wheels, or slippery terrain, as was already considered in [18]:

$$\mathbf{x}_k = \mathbf{x}_{k-1} + (1+\epsilon_i)\mathbf{R}_{k-1}\mathbf{u}_{k-1}. \tag{28}$$

In the Lie group setting, this rewrites in the form of (12) $\boldsymbol{\chi}_k = \boldsymbol{\chi}_{k-1}(\boldsymbol{\Omega}_{k-1}, \mathbf{u}_{k-1})(\mathbf{Id}, \epsilon_u \boldsymbol{\Omega}_{k-1}^T \mathbf{u}_{k-1})$. Thus

$$\log(\boldsymbol{\delta_\kappa}) = \mathbf{d_\kappa} = (0, \epsilon_u(\boldsymbol{\Omega}_{k-1}^T\mathbf{u}_{k-1})^T)^T. \tag{29}$$

The associated constraint is written, following (18):

$$\mathbf{LH}_k(\mathbf{0}, (\boldsymbol{\Omega}_{k-1}^T\mathbf{u}_{k-1})^T)^T = (0, (\boldsymbol{\Omega}_{k-1}^T\mathbf{u}_{k-1})^T)^T \tag{30}$$

2) Mismatch for observations:

a) Error on the Group Action Model: Consider the model (25), parameterised by $\epsilon = \hat{\epsilon} + \delta_\epsilon$ for the true model and $\hat{\epsilon}$ for the assumed one. In this case we have $\frac{\partial h}{\partial \mathbf{b}}\big|_{\boldsymbol{\chi},\cdot,\cdot} = \mathbf{R}$ and $\frac{\partial h}{\partial \epsilon}\big|_{\boldsymbol{\chi},\cdot,\cdot} = \mathbf{x}$ for all $\boldsymbol{\chi}$. Thus, the error is exactly

$$\frac{\partial h}{\partial \mathbf{b}}\bigg|_{\hat{\boldsymbol{\chi}}_{k|k-1}^{-1}}, \boldsymbol{\chi}_k \star_{\hat{\epsilon}} \mathbf{b}_k, \hat{\epsilon}\ \frac{\partial h}{\partial \epsilon}\bigg|_{\boldsymbol{\chi}_k, \mathbf{b}_k, \hat{\epsilon}} \delta_\epsilon = \delta_\epsilon \hat{\mathbf{R}}_{k|k-1}^T \mathbf{x}_k. \tag{31}$$

The induced constraint is then obtained from (18) as

$$\mathbf{L}\hat{\mathbf{R}}_{k|k-1}^T \hat{\mathbf{x}}_{k|k-1} = \mathbf{0} \tag{32}$$

b) Error from the action of a perturbation: Consider the action of a purely rotational perturbation, that is $\boldsymbol{\delta_\chi} = (\boldsymbol{\delta_R}, \mathbf{0}) = \exp(\boldsymbol{\delta}_\theta, \mathbf{0})$. Then the constraint (18) is written

$$\mathbf{LJ}\hat{\mathbf{R}}_{k|k-1}^T \hat{\mathbf{x}}_{k|k-1} = \mathbf{0} \tag{33}$$

Note that this produces a constraint orthogonal to (32). This was expected, since the action of $\boldsymbol{\delta}_R$ rotates the measurement, while δ_ϵ expands it.

VII. Numerical Illustrative Example

To illustrate the impact of mismatch on linear observed systems on groups, and the effectiveness of the proposed mitigation strategy, consider a 2D localization problem as presented in Section VI, in which the orientation θ and position $\mathbf{x}$ of a mobile system are to be estimated. Suppose that its attached sensors give access to its odometry and position measurements. The latter follows the models of (24) (or (25) if scaled) by simply taking $\mathbf{b} = \mathbf{0}$.

The proposed mitigation strategy is thought for cases where the constraints do not create a degenerate minimization problem (19), that is, when there is some redundancy among the sensors. Indeed, the constraints induced by the proposed strategy reduce the effective dimension of the output, as it is nullified along some directions. Thus, sensor redundancy is necessary if one hopes to carry out efficient constrained navigation. Therefore, we consider that the system contains two sensors providing position measurements.

In this experiment we considered three different mismatches: scale and frame mismatch for the outputs, and scale mismatch for the inputs. Therefore, the actual system functions are

$$(\theta_k, \mathbf{x}_k) = (\theta_{k-1} + \omega_{k-1}, \mathbf{x}_{k-1} + (1+\epsilon_k^u)\mathbf{R}_{k-1}\mathbf{u}_{k-1}),$$
$$\mathbf{y}_k^0 = \mathbf{R}(\delta_k^\psi)^T\mathbf{x}_k, \quad \mathbf{y}_k^1 = (1+\epsilon_k^y)\mathbf{x}_k,$$

where δ_k^ψ and ϵ_k^y denote respectively the frame and scale mismatch of the outputs, and ϵ_k^u the scale mismatch of the input. The estimated system simply uses $\hat{\delta}_k^\psi = \hat{\epsilon}_k^u = \hat{\epsilon}_k^y = 0$. The main idea is that it is theoretically possible to reconstruct a faithful position measurement out of $\mathbf{y}^0$ and $\mathbf{y}^1$, but the corresponding parameters are unknown, with arbitrary distributions. With two measurements, the induced constraints are given by, noting $\boldsymbol{\alpha}_k = \boldsymbol{\Omega}_{k-1}^T\mathbf{u}_{k-1}$ and $\boldsymbol{\beta}_k = \hat{\mathbf{R}}_{k|k-1}^T\hat{\mathbf{x}}_{k|k-1}$:

$$\boldsymbol{\Delta}_k = \begin{bmatrix} \boldsymbol{\alpha}_k & \boldsymbol{\beta}_k & \mathbf{0} \\ \boldsymbol{\alpha}_k & \mathbf{0} & \mathbf{J}\boldsymbol{\beta}_k \end{bmatrix}, \ \mathbf{T}_k = \begin{bmatrix} 0 & \mathbf{0} \\ \boldsymbol{\alpha}_k & \end{bmatrix}. \tag{34}$$

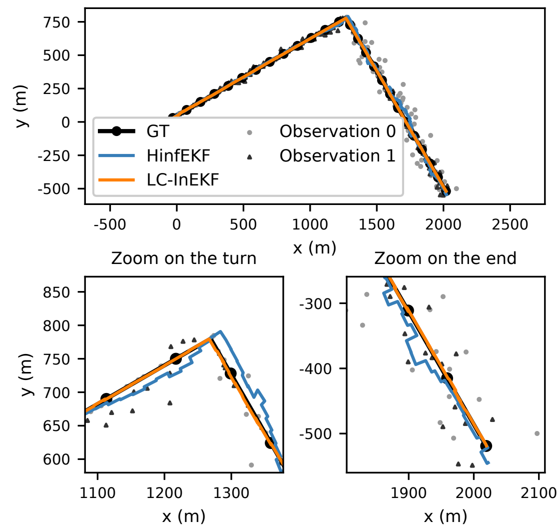

Fig. 1. Ground truth, estimated trajectories and outputs

A. Experimental Setup and Compared Filters

The simulated situation is that of a mobile robot going at 10 m/s in a straight line, taking a right angle turn and then going straight again. In order to evaluate the performance of the LC-InEKF, it is compared to three other filters: the InEKF, the LC-EKF along the lines of [13], and finally the IH$_\infty$F following [18], which had been introduced as a way to make the InEKF more robust. The root MSE (RMSE) is obtained from 50 Monte Carlo runs. The InEKF has been shown to be particularly interesting, and outperform the standard EKF, in low noise setting [15]. Thus, the considered noises are diagonal, with standard deviations given by $\sigma_u = 0.01$ m/s, $\sigma_\omega = 0.001$ rad/s, $\sigma_y = 3$ m. As the H$_\infty$ formulation requires knowing the maximum possible noise magnitude, its associated noise matrices were taken as 3 times the respective standard deviations, following [18], and the associated parameter γ was set to 0.02. The true initial state is located at $(-30\text{ m}, 30\text{ m})$, with a heading of $\pi/6$ rad. The initial heading and position estimates are taken as random following Gaussian distributions centered around the true state, with standard deviations $\sigma_{\theta_0} = 0.03$ rad and $\sigma_{x_0} = 3$ m respectively. The initial covariance is initialised accordingly: $\mathbf{P}_0 = diag(\sigma_{\theta_0}, \sigma_{x_0}, \sigma_{x_0})$. The mismatch are taken randomly, with $\epsilon_k^u \sim \mathcal{U}(-0.07, 0.15)$, $\epsilon_k^y \sim \mathcal{U}(-0.1, 0.1)$ and $\delta_k^\psi \sim \mathcal{U}(-0.05, 0.05)$ rad. In particular, note that the mismatch scale for the odometry is not centered.

B. Results

An example of the ground truth (GT) trajectory, the corrupted observations, and the estimates of the IH$_\infty$F and LC-InEKF are given in Figure 1. It clearly appears that the IH$_\infty$F deeply suffers from the mismatches, as it tries to accommodate at the middle of each pair of observations, while compensating the scaled inputs which tend to make it go too far. Figure 2 shows the position and heading RMSE of all filters. While it shows that the LC-InEKF has no problem mitigating all three mismatch sources, it also correlates with earlier work on the InEKF. First, the IH$_\infty$F performs slightly better than the InEKF, similarly to what was reported in [18]. Then, the LC-EKF manages to mitigate the mismatch, but its heading may get stuck in local minima on some runs, making the heading RMSE significant. The heading error distribution is actually bimodal, which had already been observed for invariant smoothing [21]. Indeed, $\mathbf{Q}$ is not large enough to cover the linearisation errors occuring due to the wrong heading initialisation. Overall, the new LC-InEKF is a powerful tool to robustify standard EKF and InEKF solutions, and a promising alternative to cope with misspecified system models.

VIII. Conclusion

Robust filtering has been largely explored when dealing with imprecise noise statistics or outliers, but much less so regarding model mismatch on the system functions. In this contribution, we proposed to extend the use of linear constraints for robust filter design in the particular field of invariant filtering. The possible mismatches for systems on groups were discussed, as well as mitigation strategies, leading to a new linearly constrained InEKF. This filter proved efficient in greatly reducing the impact of a class of dynamics and output model mismatch, as shown on a 2D localisation problem with mismatched inputs and outputs. The derived filter also preserved the advantages of invariant filtering over the standard EKF for precise navigation. Future work includes better approximating (19), and reducing the errors in the computation of (18).

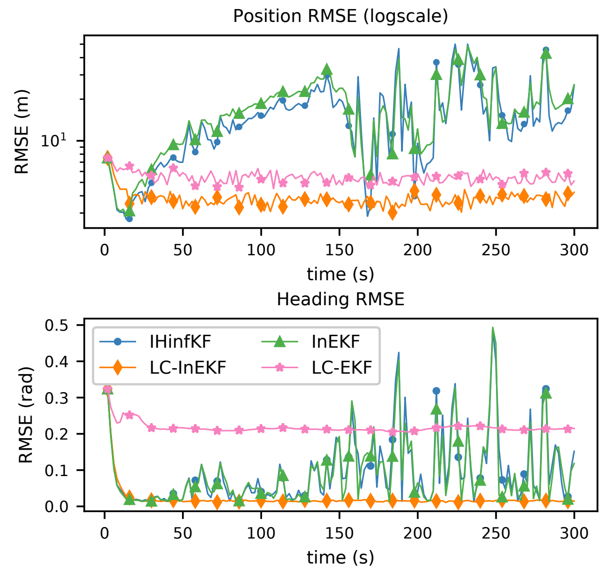

Fig. 2. Heading and position RMSE

References

[1] J. L. Crassidis and J. L. Junkins, *Optimal Estimation of Dynamic Systems*, CRC Press, Taylor & Francis Group, 2nd edition, 2012.

[2] D. Simon, *Optimal State Estimation: Kalman, H-infinity, and Nonlinear Approaches*, John Wiley & Sons, 2006.

[3] I. Arasaratnam and S. Haykin, “Cubature Kalman Filters,” *IEEE Trans. Automatic Control*, vol. 54, no. 6, pp. 1254–1269, 2009.

[4] P. Djurić, J. H. Kotecha, J. Zhang, Y. Huang, T. Ghirmai, M. F. Bugallo, and J. Míguez, “Particle Filtering,” *IEEE Signal Processing Magazine*, vol. 140, no. 2, pp. 19–38, 2003.

[5] J. Dunik, O. Straka, O. Kost, and J. Havlik, “Noise Covariance Matrices in State-Space Models: A Survey and Comparison of Estimation Methods - Part I,” *Intl. J. of Adaptive Control and Signal Process.*, vol. 31, pp. 1505–1543, 2017.

[6] A. M. Zoubir, V. Koivunen, E. Ollila, and M. Muma, *Robust Statistics for Signal Processing*, Cambridge Univ. Press, Cambridge, U.K., 2018.

[7] Y. Huang, Y. Zhang, P. Shi, Z. Wu, J. Qian, and J.A. Chambers, “Robust Kalman Filters Based on Gaussian Scale Mixture Distribution with Application to Target Tracking,” *IEEE Trans. Systems, Man. And Cybernetics*, vol. 49, no. 10, pp. 2082–2096, 2019.

[8] Z. Gao and S. X. Ding, “Actuator Fault Robust Estimation and Fault-tolerant Control for a Class of Nonlinear Descriptor Systems,” *Automatica*, vol. 43, no. 5, pp. 912 – 920, 2007.

[9] J. Lan, “Asymptotic Estimation of State and Faults for Linear Systems with Unknown Perturbations,” *Automatica*, vol. 118, pp. 108955, 2020.

[10] B. O. S. Teixeira et al., “Gain-Constrained Kalman Filtering for Linear and Nonlinear Systems,” *IEEE Trans. Signal Process.*, vol. 56, no. 9, pp. 4113–4123, 2008.

[11] E. Chaumette, F. Vincent, and J. Vilà-Valls, “Linearly Constrained Wiener Filter Estimates for Linear Discrete State-Space Models,” in *Asilomar Conference on Signals, Systems, and Computers*, 2018.

[12] J. Vilà-Valls, D. Vivet, E. Chaumette, F. Vincent, and P. Closas, “Recursive Linearly Constrained Wiener Filter for Robust Multi-Channel Signal Processing,” *Signal Processing*, vol. 167, 2020.

[13] E. Hrustic, R. Ben Abdallah, J. Vilà-Valls, D. Vivet, G. Pagès, and E. Chaumette, “Robust Linearly Constrained Extended Kalman Filter for Mismatched Nonlinear Systems,” *Int. J. Robust Nonlinear Control, early access*, 2020.

[14] A. Barrau and S. Bonnabel, “Linear Observed Systems on Groups,” *Systems & Control Letters*, vol. 129, pp. 36 – 42, 2019.

[15] A. Barrau and S. Bonnabel, “The Invariant Extended Kalman Filter as a Stable Observer,” *IEEE Trans. Autom. Control*, vol. 62, no. 4, pp. 1797–1812, 2017.

[16] R. Hartley, M. Ghaffari, R. M. Eustice, and J. W. Grizzle, “Contact-Aided Invariant EKF for Robot State Estimation,” *Int. J. Rob. Res.*, vol. 39, no. 4, pp. 402–430, 2020.

[17] M. Brossard, A. Barrau, and S. Bonnabel, “Exploiting Symmetries to Design EKFs With Consistency Properties for Navigation and SLAM,” *IEEE Sensors J.*, vol. 19, no. 4, pp. 1572–1579, 2019.

[18] M. Lavoie, J. Arsenault, and J. R. Forbes, “An Invariant Extended H∞ Filter,” in *IEEE Conference on Decision and Control*, 2019.

[19] T. D. Barfoot, *State Estimation for Robotics*, Cambridge University Press, 2017.

[20] A. W. Long, K. C. Wolfe, M. J. Mashner, and G. S. Chirikjian, “The Banana Distribution is Gaussian: A Localization Study with Exponential Coordinates,” *Robotics: Science and Systems VIII*, vol. 265, 2013.

[21] P. Chauchat, A. Barrau, and S. Bonnabel, “Invariant Smoothing on Lie Groups,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2018.