Single range localization in 3D: observability and robustness issues

Giovanni Indiveri, Daniela De Palma and Gianfranco Parlangeli

***Abstract*—The problem of estimating the position of a 3D vehicle subject to a constant unknown velocity disturbance is addressed: the only available model output is assumed to be the distance (range) to a reference point. The vehicle's nominal velocity is also assumed to be known. An observability analysis is performed and an observer is designed. The proposed approach departs from alternative ones and leads to the definition of a linear time invariant state equation with a linear time varying output. The localization problem is solved using a novel outlier robust predictor - corrector state estimator. Numerical simulation examples are described to illustrate the performance of the method as compared to a standard Kalman filter.**

***Index Terms*—Observability; Localization; Navigation; Marine Vehicles; State estimation; Robustness; Outliers.**

I. Introduction

The problem of single range based localization consists in estimating an agent's position (or possibly pose, i.e. position and attitude) exploiting knowledge about its motion model together with a range measurement from a point and eventually other sensor readings. Most often the motion model is kinematic and the available sensor readings are relative to velocity and attitude. In some applications a position measurement along the vertical may be acquired through pressure gauges. The problem of single range based localization is particularly relevant in land [1], [2], aerial [3] and marine robotics [4] [5] [6] applications. The challenge of using single range information for localization is related to the fact that traditional trilateration algorithms used in systems as the Global Positioning System (GPS), long base line (LBL) or ultra short base line (USBL) underwater navigation systems are ill posed when only range from a single point is known. Yet fusing information from a motion model of the agent and a single range measurement can be sufficient to estimate the position of the agent. Finding the conditions on the agent's motion state that allow to estimate its position from successive measurements of the distance to a fixed beacon is an observability problem that needs to be tackled in order to eventually design an observer. Given that range is a nonlinear function of the position, even if the motion model of the vehicle should be linear, the observability issue is inherently nonlinear.

In reference [1] the authors consider the localization problem for two mobile robots equipped with different sensors. Using the Lie derivatives based observability rank criterion they determine the conditions under which the system is locally weakly observable and they achieve sufficient conditions for the observability of cooperative localization. Necessary and sufficient conditions for local weak observability are derived in [2]. In [3] the problem of observability is related to the one of opportunistic navigation. The problem of the observability of the environment is there introduced and different observability concepts are investigated and compared. In [4] local system observability for range-only measurement target tracking is established through the Fisher information matrix, and results are validated using an extended Kalman filter (EKF) to estimate the target states. In [6] the authors address observability issues in the context of the relative localization of autonomous underwater vehicles. After exploiting some nonlinear observability concepts, a metric based on the condition number of the observability matrix is proposed to quantitatively assess the degree of observability as a function of relative vehicle motion parameters.

The theoretical foundation on observability for nonlinear systems is the milestone paper [7] where the fundamental ideas and results about local and local weak observability are described: single range localization studies based on differential geometric tools need to tackle the difficulties related to local and local weak observability as opposed to the global observability concept known for linear systems. Such issues are clearly addressed, by example, in references [8], [9] and [10]. In [8] the authors propose an algebraic estimator for the estimation of an underwater vehicle using a single acoustic transponder. In [9] local weak observability for the kinematics of an underwater vehicle with range-only measurement is performed and an EKF is proposed. Simulation results evidence that the EKF shows exponential convergence when the filter's initial condition is sufficiently close to the actual one. In [10] the observability analysis is based on a linearization of the model.

An interesting approach to address the global observability analysis and observer design for single range localization is described in the work of Batista et al. [11] [12] [13]: they transform the original nonlinear system in a linear time varying (LTV) system through an augmented state technique. This leads to the remarkable result of allowing to study the global observability properties of the system with the tools of LTV systems theory [14] and of designing a Kalman filter for global state estimation. Moreover, the problem of range based localization is technically similar to the problem of source localization [15] [16] where a vehicle knowing its own position is asked to estimate the position of a source (or target)

This work was partially supported by the European Unions Horizon 2020 research and innovation programme under grant agreement No. 645141 (WiMUST project) and MIUR PRIN project MARIS prot. 2010FBLHRJ.

The authors are with Dipartimento Ingegneria Innovazione - ISME node - Università del Salento, via Monteroni, 73100 Lecce - Italy `name.lastname@unisalento.it`

from which it acquires range measurements. The single range localization problem is of particular interest for underwater navigation applications where range measurements are typically acquired through acoustic time of flight based sensors: the data so collected is often corrupted by outliers. Indeed the issue of designing outlier robust underwater navigation systems is of great importance and it has been addressed, for example, in [17] [18] [19] [20].

Motivated by the above arguments, the main goal of this paper is to propose an effective localization algorithm for a 3D vehicle based on single range only. The proposed solution is based on the construction of a fictitious LTI state equation defined on $\mathbb{R}^8$ with a time varying scalar output equation which allows to address the observability analysis and the state estimation filter design resorting to LTV systems theory. Furthermore, inspired by the problem of the presence of measurement outliers, a robust state estimator based on the Least Entropy-Like approach [21] is designed and illustrated.

The paper is organized as follows: the problem formulation, main ideas and methods are described in section II. A Kalman filter and a novel outlier robust state estimator are described in sections III and IV respectively. Implementation aspects of the proposed robust state estimator are addressed in section V. Finally, conclusions are summarized in section VI.

II. Single range localization

Consider an agent (or a vehicle, in the following) with position given by vector $\mathbf{p}$ and a source (or navigation reference) with fixed inertial position $\mathbf{s}$. Denoting with

$$\mathbf{r} := \mathbf{s} - \mathbf{p} \tag{1}$$

the relative position of the source with respect to the agent it is assumed that this can access the measurement y given by squared norm of $\mathbf{r}$, namely

$$y = \|\mathbf{r}\|^2. \tag{2}$$

Moreover, denoting with $\{I\}$ and $\{B\}$ an earth fixed and body fixed frames respectively, it is assumed that the agent has access to a measurement of its attitude, namely it can measure the rotation matrix ${}^I R_B \in \mathrm{SO}(3)$ thanks to an on board navigation system (Attitude and Heading Reference System - AHRS). The agent velocity is given by a superposition of a drift term $\mathbf{v}_f$ (unknown, but constant) and a controlled input term $\mathbf{v}_r$. For underwater vehicles the velocity term $\mathbf{v}_r$ can be generated through a guidance controller exploiting an on board navigation sensor as a Doppler Velocity Logger (DVL). The drift term $\mathbf{v}_f$ models a constant unknown ocean current. The resulting agent motion model expressed in frame $\{I\}$ is thus

$$\begin{aligned} \dot{\mathbf{p}} &= \mathbf{v}_r + \mathbf{v}_f && (3)\\ \dot{\mathbf{v}}_f &= \mathbf{0} && (4)\\ \dot{\mathbf{s}} &= \mathbf{0}. && (5) \end{aligned}$$

Consequently, the agents state equations in terms of the relative position $\mathbf{r}$ expressed in the fixed frame $\{I\}$ result in:

$$\begin{aligned} \dot{\mathbf{r}} &= -\mathbf{v}_r - \mathbf{v}_f && (6)\\ \dot{\mathbf{v}}_f &= \mathbf{0} && (7)\\ y &= \|\mathbf{r}\|^2. && (8) \end{aligned}$$

In the sequel we will refer to the model in (6, 7, 8) as the "*original*" model. The problem addressed in the paper can thus be formulated as in the following subsection II-A.

A. Problem formulation

Given the linear state equations (6 - 7) for the state vector $(\mathbf{r}^\top, \mathbf{v}_f^\top)^\top \in \mathbb{R}^6$ and the nonlinear scalar output y in (8) determine the conditions on the input $\mathbf{v}_r$ that guarantee observability of the state $(\mathbf{r}^\top, \mathbf{v}_f^\top)^\top$ and design a state estimator that is robust to possible outliers in the output y.

This problem corresponds to the one addressed in [13] with the only difference that here it is formulated in the inertial frame $\{I\}$ rather than in the body fixed frame $\{B\}$. Yet it should be noticed that the available information in the two setups is identical as both formulations require to have access to the rotation matrix ${}^I R_B \in \mathrm{SO}(3)$. In particular, in the present paper ${}^I R_B$ is needed to recover $\mathbf{v}_r$ in the $\{I\}$ frame from its measurement in the $\{B\}$ frame through a DVL (or other on board navigation sensors).

To tackle the formulated observability problem consider the integral of (6)

$$\begin{aligned} \mathbf{r}(t) - \mathbf{r}_0 &= -\mathbf{v}_f\,(t - t_0) - \int_{t_0}^{t} \mathbf{v}_r(\tau) d\tau = \\ &= -\mathbf{v}_f\,(t - t_0) - \boldsymbol{I}_{\mathbf{v}_r}(t_0, t) && (9) \end{aligned}$$

having defined the displacement $\boldsymbol{I}_{\mathbf{v}_r}(t_0, t) \in \mathbb{R}^{3\times 1}$ as

$$\boldsymbol{I}_{\mathbf{v}_r}(t_0, t) := \int_{t_0}^{t} \mathbf{v}_r(\tau)\, d\tau \tag{10}$$

and $\mathbf{r}_0 := \mathbf{r}(t)|_{t=t_0}$. Equation (9) allows to compute

$$\begin{aligned} &(\mathbf{r}(t) + \boldsymbol{I}_{\mathbf{v}_r}(t_0, t))^\top (\mathbf{r}(t) + \boldsymbol{I}_{\mathbf{v}_r}(t_0, t)) = \\ &= (\mathbf{r}_0 - \mathbf{v}_f\,(t - t_0))^\top (\mathbf{r}_0 - \mathbf{v}_f\,(t - t_0)) \end{aligned}$$

implying

$$\begin{aligned} &\|\mathbf{r}(t)\|^2 + \|\boldsymbol{I}_{\mathbf{v}_r}(t_0, t)\|^2 + 2\,\boldsymbol{I}_{\mathbf{v}_r}^\top(t_0, t)\,\mathbf{r}(t) = \\ &= \|\mathbf{r}_0\|^2 + \|\mathbf{v}_f\|^2\,(t - t_0)^2 - 2\,(\mathbf{r}_0^\top \mathbf{v}_f)\,(t - t_0) && (11) \end{aligned}$$

namely

$$\begin{aligned} &\|\mathbf{r}(t)\|^2 - \|\mathbf{r}_0\|^2 + \|\boldsymbol{I}_{\mathbf{v}_r}(t_0, t)\|^2 = && (12)\\ &= -2\,\boldsymbol{I}_{\mathbf{v}_r}^\top(t_0, t)\,\mathbf{r}(t) - 2\,(\mathbf{r}_0^\top \mathbf{v}_f)\,(t - t_0) + \|\mathbf{v}_f\|^2\,(t - t_0)^2. \end{aligned}$$

Notice that the left hand side of (12) is made of all known terms and it can be used as a new output map

$$\begin{aligned} \bar{y}(t) &= \|\mathbf{r}(t)\|^2 - \|\mathbf{r}_0\|^2 + \|\boldsymbol{I}_{\mathbf{v}_r}(t_0, t)\|^2 = \\ &= y(t) - y_0 + \|\boldsymbol{I}_{\mathbf{v}_r}(t_0, t)\|^2 && (13) \end{aligned}$$

and the right hand side of (12) can be expressed as a linear time varying (LTV) term in the new state variable $\mathbf{z} \in \mathbb{R}^{8\times 1}$

$$\mathbf{z} = (\mathbf{r}^\top, (\mathbf{r}_0^\top \mathbf{v}_f), \|\mathbf{v}_f\|^2, \mathbf{v}_f^\top)^\top, \tag{14}$$

i.e.

$$\bar{y}(t) = C(t)\,\mathbf{z} = \begin{bmatrix} -2\,\boldsymbol{I}_{\mathbf{v}_r}^\top(t_0,t) & -2\,\delta & \delta^2 & 0_{1\times 3} \end{bmatrix} \mathbf{z}. \tag{15}$$

being $\delta = (t - t_0)$. Given the definition of $\mathbf{z}$ in (14) and the model (6) - (7), its dynamic equation is linear time invariant (LTI):

$$\dot{\mathbf{z}} = A\,\mathbf{z} + B\,\mathbf{v}_r \tag{16}$$

namely

$$\dot{\mathbf{z}} = \frac{d}{dt}\begin{pmatrix} \mathbf{r} \\ (\mathbf{r}_0^\top \mathbf{v}_f) \\ \|\mathbf{v}_f\|^2 \\ \mathbf{v}_f \end{pmatrix} = \begin{bmatrix} 0_{3\times3} & 0_{3\times1} & 0_{3\times1} & -I_{3\times3} \\ 0_{1\times3} & 0 & 0 & 0_{1\times3} \\ 0_{1\times3} & 0 & 0 & 0_{1\times3} \\ 0_{3\times3} & 0_{3\times1} & 0_{3\times1} & 0_{3\times3} \end{bmatrix} \cdot$$
$$\cdot \begin{pmatrix} \mathbf{r} \\ (\mathbf{r}_0^\top \mathbf{v}_f) \\ \|\mathbf{v}_f\|^2 \\ \mathbf{v}_f \end{pmatrix} + \begin{bmatrix} -I_{3\times3} \\ 0_{1\times3} \\ 0_{1\times3} \\ 0_{3\times3} \end{bmatrix} \mathbf{v}_r. \tag{17}$$

The range-only localization problem of estimating $\mathbf{r}$ and the current velocity $\mathbf{v}_f$ from a measurement of $\|\mathbf{r}\|^2$ in (6, 7, 8) is hence reduced to a state estimation problem on a linear time invariant state equation (16) - (17) with an LTV output map (15), namely

$$\begin{cases} \dot{\mathbf{z}} = A\,\mathbf{z} + B\,\mathbf{v}_r \\ \bar{y}(t) = C(t)\,\mathbf{z}. \end{cases} \tag{18}$$

Estimating $\mathbf{z}$ will result in estimating both $\mathbf{r}$ and the current velocity $\mathbf{v}_f$. Moreover, in case that the absolute position of the source $\mathbf{s}$ is known a priori, by estimating $\mathbf{r}$ the absolute position of the vehicle could also be computed as $\mathbf{p} = \mathbf{s} - \mathbf{r}$.

B. Observability analysis

The observability properties of LTV system (18) can be studied through the observability Gramian

$$G(t_0,t) = \int_{t_0}^{t} e^{A^\top(\tau-t_0)}\, C^\top(\tau)\, C(\tau)\, e^{A(\tau-t_0)} d\tau. \tag{19}$$

Given the structure of the A matrix in (17), notice that $A^2 = 0_{8\times8}$ implying that the exponential matrix $e^{A(\tau-t_0)}$ is simply

$$e^{A(\tau-t_0)} = I_{8\times8} + A(\tau - t_0) \tag{20}$$

such that $C(\tau)e^{A(\tau-t_0)}$ results in

$$C(\tau)\, e^{A(\tau-t_0)} = \begin{bmatrix} -2\,\boldsymbol{I}_{\mathbf{v}_r}^\top(t_0,\tau) & -2\,\zeta & \zeta^2 & 2\zeta\,\boldsymbol{I}_{\mathbf{v}_r}^\top(t_0,\tau) \end{bmatrix}$$

and

$$e^{A^\top(\tau-t_0)}\, C^\top(\tau)\, C(\tau)\, e^{A(\tau-t_0)} =$$
$$\begin{bmatrix} 4\,\boldsymbol{I}_{\mathbf{v}_r}\boldsymbol{I}_{\mathbf{v}_r}^\top & 4\zeta\,\boldsymbol{I}_{\mathbf{v}_r} & -2\zeta^2\,\boldsymbol{I}_{\mathbf{v}_r} & -4\zeta\,\boldsymbol{I}_{\mathbf{v}_r}\boldsymbol{I}_{\mathbf{v}_r}^\top \\ 4\zeta\,\boldsymbol{I}_{\mathbf{v}_r}^\top & 4\,\zeta^2 & -2\zeta^3 & -4\zeta^2\,\boldsymbol{I}_{\mathbf{v}_r}^\top \\ -2\zeta^2\,\boldsymbol{I}_{\mathbf{v}_r}^\top & -2\,\zeta^3 & \zeta^4 & 2\zeta^3\,\boldsymbol{I}_{\mathbf{v}_r}^\top \\ -4\zeta\boldsymbol{I}_{\mathbf{v}_r}\boldsymbol{I}_{\mathbf{v}_r}^\top & -4\zeta^2\boldsymbol{I}_{\mathbf{v}_r} & 2\zeta^3\boldsymbol{I}_{\mathbf{v}_r} & 4\zeta^2\boldsymbol{I}_{\mathbf{v}_r}\boldsymbol{I}_{\mathbf{v}_r}^\top \end{bmatrix} \tag{21}$$

where $\zeta = \tau - t_0$ and the dependency of $\boldsymbol{I}_{\mathbf{v}_r}$ from t_0, τ has been omitted for the sake of notation compactness.

As for the observability conditions, following standard results for LTV systems [14], the LTV model in (18) will be observable in the time interval $[t_0, t]$ if and only if the Gramian given by equations (19) and (21) has full rank. Moreover, the structure of equation (21) implies that a necessary condition for the observability of the LTV model (18) in the time interval $[t_0, t]$ is that

$$G_{11}(t_0,t) := 4\int_{t_0}^{t} \boldsymbol{I}_{\mathbf{v}_r}(t_0,\tau)\, \boldsymbol{I}_{\mathbf{v}_r}^\top(t_0,\tau) d\tau \tag{22}$$

has full rank, i.e. three. Overall, the observability properties in the presence of constant currents can be summarized as follows.

Main Result - *Observability conditions.*
The LTV model in (18) is observable on $[t_0, t]$ if and only if the velocity signal $\mathbf{v}_r$ guarantees that the Gramian in equations (19) and (21) has full rank. Moreover a necessary condition for observability on $[t_0, t]$ is that the matrix $G_{11}(t_0,t) \in \mathbb{R}^{3\times3}$ in (22) has rank 3. Finally, the observability of the LTV model in (18) is a sufficient, but not necessary, condition to grant the observability of the original model in (6, 7, 8).

Proof of the Main Result
The necessary and sufficient conditions on the Gramian in equations (19) and (21) follow from standard LTV systems theory [14]. As for the necessary condition on the rank of the matrix $G_{11}(t_0,t) \in \mathbb{R}^{3\times3}$ in (22) it results that if $G_{11}(t_0,t)$ should *not* be full rank on $[t_0, t]$, there would exist a constant vector $\boldsymbol{\nu} \in \mathbb{R}^{3\times1}$, $\boldsymbol{\nu} \neq \mathbf{0}$ such that $\boldsymbol{I}_{\mathbf{v}_r}(t_0,\tau)^\top\boldsymbol{\nu} = 0 \quad \forall \quad \tau \in [t_0, t]$: this implies that any vector parallel to $\mathbf{z}^* = (\alpha\,\boldsymbol{\nu}^\top, 0, 0, \beta\,\boldsymbol{\nu}^\top)^\top \in \mathbb{R}^{8\times1}$ for any constant $\alpha, \beta \in \mathbb{R}$ would belong to the kernel of the Gramian (19) - (21) that, hence, would not be full rank. This proves that $\mathrm{rank}(G_{11}(t_0,t)) = 3 \; : \; G_{11}(t_0,t) \in \mathbb{R}^{3\times3}$ is defined in (22) is a necessary condition for the observability in $[t_0, t]$ of the LTV model in (18).
At last, given that:

1) all the possible state trajectories of the original model (6, 7, 8) are (by construction) a subset of components of specific state trajectories of the augmented LTV model in (18);
2) the two systems share the same identical input / output information,

the observability of the LTV system (18) implies (i.e. it is a sufficient condition) the observability of the original system (6, 7, 8). Notice that the opposite implication is false, namely the observability of the original system would not imply the observability of the LTV system. In fact, the state trajectories of the LTV system belong to a larger space ($\mathbb{R}^8$ in place of $\mathbb{R}^6$) and include, by example, cases where the fifth component z_5 of $\mathbf{z}$ could be negative: this would not correspond to any trajectory of the original system as the mapping $z_5 = \|\mathbf{v}_f\|^2$ would be obviously violated. It is worth highlighting that since the observability of the LTV system is a sufficient condition for the observability of the original nonlinear system, the full rank condition of the matrix $G_{11}(t_0,t)$ can not be a necessary condition for the observability of the original system as well.

Remark
As for the LTV system (18), the *Main Result* is simply

stating that it is observable if its observability Gramian is invertible. This is of course obvious and is not the point. The relevance of the *Main Result* is related to the fact that the (simple) derived observability conditions imply *global* observability for the original nonlinear system (6, 7, 8). This is remarkable as traditional nonlinear observability methods allow to determine only local observability results. In [12] [13] a state augmentation technique is proposed to address the same observability analysis and state estimation problem at hand. As a result the original nonlinear problem is mapped on a nine dimensional LTV system of the form

$$\dot{\mathbf{z}} = A(t, \mathbf{u}(t), y(t))\, \mathbf{z} + B\, \mathbf{u}(t) \tag{23}$$

$$y(t) = C\, \mathbf{z}. \tag{24}$$

The system matrix $A(t)$ in (23) explicitly depends, among the rest, on the output $y(t)$ that is the range measurement (i.e. the norm of the agent's position vector). More precisely, the $A(t)$ matrix is a function of terms proportional to $1/y(t)$: this poses both fundamental as well as implementation issues (singularity if $y(t) \approx 0$). These issues are absent in the approach described in this paper as the state equation matrix does not depend on $1/y(t)$. In this respect the proposed solution resembles the one in [11]. Notice that despite the difference between the augmented LTV models (18) and (23-24), the observability conditions are quite close. This is not surprising since they come from the same original problem. In particular, in [13] necessary conditions and sufficient conditions for the observability of (23-24) are derived in terms of linear independence of a set of function, whereas in II-B necessary and sufficient conditions for the LTV model in (18) are found in terms of rank of the observability Gramian (19).

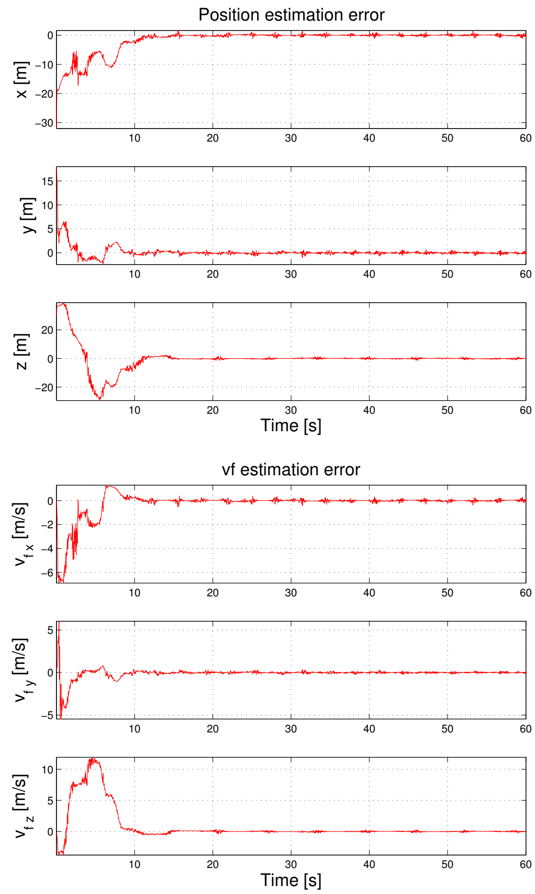

Fig. 1. Kalman filter estimation errors including currents: estimation error components of $\mathbf{p}$ (left) and of the current velocity $\mathbf{v}_f$ (right).

III. KALMAN FILTER DESIGN

With reference to the model in (18) assume that it is discretized and affected by state and output zero mean mutually independent disturbances respectively $\boldsymbol{\omega}_k$ and ε_k with covariances Q_k and R_k. Denoting with $\hat{\mathbf{z}}_{k|k}$ the Kalman estimate at step k and with $\hat{\mathbf{z}}_{k+1|k}$ the model prediction, the localization Kalman filter can be designed. In particular, a numerical experiment is performed using the same agent velocity profile $\mathbf{v}_r$ used in the examples presented in [15] and [13] namely $\mathbf{v}_r = (2\cos(t), -4\sin(2t), \cos(t/2))^\top$ [m/s].

The source beacon $\mathbf{s}$ is $\mathbf{s} = (2, 3, 1)^\top$ [m], the current is $\mathbf{v}_f = (0.2, 0.3, -0.1)^\top$ [m/s], and the initial position of the agent is $\mathbf{p}_0 = (2, 2, 0)^\top$ [m] such that the inertial position of the agent by $\mathbf{p}(t) = (2 + 0.2\,t + 2\sin(t), 2 + 0.3\,t + 2\cos(2t), -0.1\,t + 2\sin(0.5t))^\top$ [m]. Notice that, by direct calculation, the above $\mathbf{v}_r$ input satisfies the observability condition given in the *Main Result*. The covariances on the state $\mathbf{z}$ and output $\bar{y}(t)$ employed in the Kalman filter are $Q = (1e-2)\mathrm{diag}([1, 1, 1, 1e-4, 1e-6, (1e-2), (1e-2), (1e-2))$ and $R = 1$ respectively with proper units (i.e. [m^2] for position variables and [(m/s)^2] for velocity variables). Furthermore, assuming to acquire the velocity input through a DVL, a Gaussian noise with covariance $(1e-4)\mathrm{diag}([1, 1, 1])$ [m/s] has been added on $\mathbf{v}_r$. The filter is initialized with a position $\hat{\mathbf{p}}_0 = (-30, 20, 30)^\top$[m] as opposed to the real initial position $\mathbf{p}_0 = (2, 2, 0)^\top$[m] and a current estimate $\hat{\mathbf{v}}_f = (0.1, -0.1, 0.1)^\top$[m/s] as opposed to the real current $\mathbf{v}_f = (0.2, 0.3, -0.1)^\top$ [m/s]. A sampling time $T_s = (1/50)$[s] has been used in the implementation of the Kalman filter. The resulting time evolution of the position estimation error $\mathbf{p} - \hat{\mathbf{p}}$, being $\mathbf{p} = \mathbf{s} - \mathbf{r}$ and its estimate $\hat{\mathbf{p}} = \mathbf{s} - \hat{\mathbf{r}}$, are plotted in figure 1 together with the current estimation error $\mathbf{v}_f - \hat{\mathbf{v}}_f$. Uniform complete observability of the LTV system is required to ensure that the Kalman filter is asymptotically stable [22] [23] [24]. Moreover, the proposed solution allows to design a Kalman filter for state estimation on a system where all the system matrices (A, B and $C(t)$) are not affected by measurement noise as they do not depend from the output. Nevertheless, even if the velocity input $\mathbf{v}_r$ was perfectly known, i.e. noise free, the optimality of the Kalman filter is not preserved since the quantity actually measured is the range, not its square (2). In case of Gaussian noise for the range measurements, the noise affecting the squared range would have a chi-squared distribution rather than Gaussian. Therefore the Kalman filter can be considered a suboptimal solution rather than optimal for the LTV system. Consequently, of course, the proposed Kalman filter is not optimal for the original nonlinear model either. Moreover, the noise affecting the squared range is amplified for large distances due to the square operator.

Indeed, the output covariance is a function of the range itself. In particular, the covariance is increasing with range. Therefore special care needs to be taken in case of large distances where the performance of the estimation may be reduced. In this context the use of augmented descriptor systems [25] [26] to design observers that estimate simultaneously system states, measurement output noises and input uncertainties could be investigated. Also notice that the new output $\bar{y}(t)$ in (13) depends on the very first measurement $y(t_0)$. This dependency can impact on the robustness of the solution as a single bad measurement (as an outlier) at $t = t_0$ will affect the output for ever. A remedy to this issue can be found by periodically re-setting the initial measurement $y(t_0)$ with $y(t)$. This would also prevent possible uncertainties in the knowledge of $\mathbf{v}_r(t)$ from biasing unboundedly the displacement $\boldsymbol{I}_{\mathbf{v}_r}$ in (10) used to compute $\bar{y}$. A detailed analysis of this implementation detail goes beyond the scope of this paper and will not be addressed further. Besides this issue, notice that outliers are likely to be present in the range measurements in particular when acoustic sensors are employed.

IV. Outlier robust state estimation

Range measurements are often contaminated by outliers, namely data points that cannot be modelled by a single (eventually Gaussian) probability distribution function. To address this, one class of approaches exploits the equivalence between the Kalman filter and a weighted least square regression problem. It is known that the Kalman filter can be derived as a solution to the following minimization problem [24] [27]:

$$\begin{aligned}\hat{\mathbf{z}}_{k+1|k+1} = \\ \arg\min_{\mathbf{z}_{k+1}} \Big\{ \frac{1}{2}(\mathbf{z}_{k+1} - \hat{\mathbf{z}}_{k+1|k})^\top P^{-1}_{k+1|k}(\mathbf{z}_{k+1} - \hat{\mathbf{z}}_{k+1|k}) + \\ + \frac{1}{2}(\bar{y}_{k+1} - C_{k+1}\mathbf{z}_{k+1})^\top R^{-1}_{k+1}(\bar{y}_{k+1} - C_{k+1}\mathbf{z}_{k+1}) \Big\}\end{aligned} \tag{25}$$

where R_{k+1} is the covariance associated to the measurement $\bar{y}_{k+1}$ and $P_{k+1|k}$ is the covariance of the (model) predicted state $\hat{\mathbf{z}}_{k+1|k}$. In [28], [29] and [30] the authors solve this minimization problem in a robust manner replacing the second term of the objective function by robustifying functions used in the methodology of M-estimation (e.g., the Huber function [31]). They express the solution as a weighted least square approximation, where each weight indicates its contribution to the state estimate. Robustness is achieved trying to give a finite weight to single residuals that exceed a threshold. Each residual contributes to the objective function based on its bare value regardless the overall residual distribution. Other approaches as [32] and [33] model the observation noise through a heavy-tailed distribution assigning outliers a non-negligible probability: interestingly the resulting estimators still result in a weighted [32] or iterative re-weighted [33] least squares solution. A robust alternative technique is here proposed based on the robust parameter identification method known as LEL (Least Entropy-Like) [21]. The basic idea is to estimate the state minimizing the following non linear objective function

$$\begin{aligned}\hat{\mathbf{z}}_{k+1|k+1} &= \arg\min_{\mathbf{z}_{k+1}} \; J_{k+1} \\ J_{k+1} &= \underbrace{\frac{1}{2}\left((\mathbf{z}_{k+1} - \hat{\mathbf{z}}_{k+1|k})^\top P^{-1}_{k+1|k}(\mathbf{z}_{k+1} - \hat{\mathbf{z}}_{k+1|k})\right)}_{J_{dynamical\ model}} + \\ &+ \underbrace{\alpha\, H_{k+1}(r_1, \ldots, r_{k+1})}_{J_{LEL}}\end{aligned} \tag{26}$$

where $r_i = \bar{y}_i - C_i\hat{\mathbf{z}}_{i|i}$, $i = 1, \ldots, k$ and $r_{k+1} = \bar{y}_{k+1} - C_{k+1}\mathbf{z}_{k+1}$ denote the residuals. $H_{k+1}(\cdot)$ represents a residual loss function defined by exploiting the mathematical properties of Gibbs entropy. Notice that this is not to be confused with information theoretic or statistical based entropy methods. Refer to [21] for a detailed discussion about this point. Define D_{k+1} as the least squares cost

$$D_{k+1} = \sum_{j=1}^{k+1} {r_j}^2 \tag{27}$$

and the relative squared residual q_i as

$$\text{if } D_{k+1} \neq 0 \Rightarrow q_i := \frac{{r_i}^2}{\sum\limits_{j=1}^{k+1} {r_j}^2} : q_i \in [0, 1] \;\text{ and }\; \sum_{i=1}^{k+1} q_i = 1, \tag{28}$$

the residual loss function has the following definition:

$$H_{k+1} = \begin{cases} 0 & \text{if } D_{k+1} = 0 \\ -\frac{1}{\log(k+1)} \sum\limits_{i=1}^{k+1} q_i \log q_i & \text{otherwise.} \end{cases} \tag{29}$$

The main difference with respect to the Kalman filter and the methods based on M-estimators relies on the structure of the second term of the objective function (26). The aim of such loss function is to give a global measure of the scatter of the relative squared residuals. The idea behind the estimator is to make the relative squared residuals 'as little equally distributed as possible'. If this is the case, most residuals are small (with respect to the normalization constant D_{k+1}, i.e. the Least Squares cost) and 'a few' of the residuals are large. Data points corresponding to these large residuals are outlier candidates. It is worth highlighting that the structure of the LEL entropy-based loss function can not be resembled to the methodology of M-estimators. Indeed according to the definition of M-estimators, the contribution to the objective function of the i-th residual does not depend on the other residuals. This is not the case of the proposed estimator, since all residuals contribute to the objective function of the i-th residual through the relative squared residual q_i.

It should be noticed that due to the normalization factor $1/\log(k+1)$ in (29), $H_{k+1}(\cdot) \in [0, 1]$ by construction, so the parameter α in (26) is to be regarded as a tuning gain needed to make the two terms of J_{k+1} comparable.

As explicitly reported in (26), $H_{k+1}(\cdot)$ depends on the residuals $\{r_1, \ldots, r_{k+1}\}$: each residual r_i in this set is a function of the state estimate $\hat{\mathbf{z}}_{i|i}$ that is a fixed quantity for all $i < k+1$. Hence in terms of state variables, $H_{k+1}(\cdot)$ in (26) depends only on $\mathbf{z}_{k+1}$, namely

$$H_{k+1}(\mathbf{z}_{k+1}) = H_{k+1}(r_1(\hat{\mathbf{z}}_{1|1}), r_2(\hat{\mathbf{z}}_{2|2}), \ldots, r_{k+1}(\mathbf{z}_{k+1}))$$

that for the sake of notation compactness will be alternatively expressed as a function of $\mathbf{z}$, namely $\mathbf{z} \mapsto \mathbf{z}_{k+1}$ resulting in $H_{k+1}(\mathbf{z})$. In order to find the solution of the minimization problem (26), $H_{k+1}(\mathbf{z})$ can be approximated in a neighborhood of $\hat{\mathbf{z}}_{k|k}$ with a quadratic function by means of its second order Taylor series expansion.

$$H_{k+1}(\mathbf{z}) = H_{k+1}(\hat{\mathbf{z}}_{k|k}) + \nabla_{\mathbf{z}} H_{k+1}(\hat{\mathbf{z}}_{k|k})^\top (\mathbf{z} - \hat{\mathbf{z}}_{k|k}) + \quad (30)$$
$$+\frac{1}{2}(\mathbf{z} - \hat{\mathbf{z}}_{k|k})^\top \mathcal{H}\left[H_{k+1}(\hat{\mathbf{z}}_{k|k})\right](\mathbf{z} - \hat{\mathbf{z}}_{k|k}) + o(\|\mathbf{z} - \hat{\mathbf{z}}_{k|k}\|^2)$$

where the gradient and hessian $\mathcal{H}(\cdot)$ of the LEL cost function computed about the point $\hat{\mathbf{z}}_{k|k}$ are denoted respectively as

$$\nabla_{\mathbf{z}} H_{k+1}(\cdot)|_{\mathbf{z}=\hat{\mathbf{z}}_{k|k}} = \nabla_{\mathbf{z}} H_{k+1}(\hat{\mathbf{z}}_{k|k}) \quad (31)$$
$$\mathcal{H}[H_{k+1}(\cdot)]|_{\mathbf{z}=\hat{\mathbf{z}}_{k|k}} = \mathcal{H}[H_{k+1}(\hat{\mathbf{z}}_{k|k})] \quad (32)$$

fBy direct calculation assuming $D_{k+1} \neq 0$ the gradient and hessian of H_{k+1} result in:

$$\nabla_{\mathbf{z}} H_{k+1}(\mathbf{f}) = \frac{2}{D_{k+1}(\mathbf{f})\log(k+1)} \cdot$$
$$\cdot \left(\log r_{k+1}^2(\mathbf{f}) - \frac{S_{k+1}(\mathbf{f})}{D_{k+1}(\mathbf{f})}\right) C_{k+1}^\top r_{k+1}(\mathbf{f}) \quad (33)$$
$$\mathcal{H}[H_{k+1}(\mathbf{f})] = \frac{2}{D_{k+1}^2(\mathbf{f})\log(k+1)}\Bigg[2\,C_{k+1}^\top C_{k+1} r_{k+1}^2(\mathbf{f}) \cdot$$
$$\cdot\left(2\log r_{k+1}^2(\mathbf{f}) - 2\frac{S_{k+1}(\mathbf{f})}{D_{k+1}(\mathbf{f})} + 1\right) - C_{k+1}^\top C_{k+1} \cdot$$
$$\cdot \Big(D_{k+1}(\mathbf{f})\log r_{k+1}^2(\mathbf{f}) - S_{k+1}(\mathbf{f}) + 2\,D_{k+1}(\mathbf{f})\Big)\Bigg] \quad (34)$$

being $\mathbf{f} = \mathbf{z}_{k+1}$, D_{k+1} defined as in (27) and S_{k+1} as

$$S_{k+1}(\mathbf{f}) = \left(\sum_{j=1}^{k} r_j^2(\hat{\mathbf{z}}_{j|j})\log r_j^2(\hat{\mathbf{z}}_{j|j})\right) + r_{k+1}^2(\mathbf{f})\log r_{k+1}^2(\mathbf{f}). \quad (35)$$

Direct inspection of equations (33, 34) reveals that while the gradient $\nabla_{\mathbf{z}} H_{k+1}(\mathbf{z}_{k+1})$ is always well defined, the hessian matrix $\mathcal{H}[H_{k+1}(\mathbf{z}_{k+1})]$ is ill posed if $r_{k+1} = 0$ due to the presence of the term $\log r_{k+1}^2$ (not multiplied by r_{k+1}) in (34). A regularization technique for $\mathcal{H}[H_{k+1}(\mathbf{z}_{k+1})]$ will be discussed in section V. Assuming for the moment the hessian to be limited and well defined, the objective function J_{k+1} in (26) is approximated as

$$J_{k+1} \approx (\mathbf{z}_{k+1} - \hat{\mathbf{z}}_{k+1|k})^\top P_{k+1|k}^{-1}(\mathbf{z}_{k+1} - \hat{\mathbf{z}}_{k+1|k}) +$$
$$+\,\alpha\Bigg(H_{k+1}(\hat{\mathbf{z}}_{k|k}) + \nabla_{\mathbf{z}} H_{k+1}(\hat{\mathbf{z}}_{k|k})^\top(\mathbf{z}_{k+1} - \hat{\mathbf{z}}_{k|k}) +$$
$$+\frac{1}{2}(\mathbf{z}_{k+1} - \hat{\mathbf{z}}_{k|k})^\top \mathcal{H}[H_{k+1}(\hat{\mathbf{z}}_{k|k})](\mathbf{z}_{k+1} - \hat{\mathbf{z}}_{k|k})\Bigg) \quad (36)$$

Thus, setting the gradient of J_{k+1} with respect to $\mathbf{z}_{k+1}$ equal to zero, the filter equations result in:

$$\hat{\mathbf{z}}_{k+1|k} = A_d\hat{\mathbf{z}}_{k|k} + B_d\mathbf{v}_{rk} \quad (37)$$
$$P_{k+1|k} = A_d P_{k|k} A_d^\top + Q_k \quad (38)$$
$$K_{k+1} = \left(P_{k+1|k}^{-1} + \alpha\mathcal{H}[H_{k+1}(\hat{\mathbf{z}}_{k|k})]\right)^{-1}\alpha\mathcal{H}[H_{k+1}(\hat{\mathbf{z}}_{k|k})] \quad (39)$$
$$\hat{\mathbf{z}}_{k+1} = \hat{\mathbf{z}}_{k+1|k} + K_{k+1}(\hat{\mathbf{z}}_{k|k} - \hat{\mathbf{z}}_{k+1|k}) +$$
$$-\left(P_{k+1|k}^{-1} + \alpha\mathcal{H}[H_{k+1}(\hat{\mathbf{z}}_{k|k})]\right)^{-1}\alpha\,\nabla_{\mathbf{z}} H_{k+1}(\hat{\mathbf{z}}_{k|k}) \quad (40)$$
$$P_{k+1|k+1} = \left(P_{k+1|k}^{-1} + \alpha\mathcal{H}[H_{k+1}(\hat{\mathbf{z}}_{k|k})]\right)^{-1} \quad (41)$$

It is worth emphasizing that the term $P_{k+1|k}^{-1} + \alpha\mathcal{H}[H_{k+1}(\hat{\mathbf{z}}_{k|k})]$ is the hessian of the cost J_{k+1}. Assuming that $\mathcal{H}[H_{k+1}(\hat{\mathbf{z}}_{k|k})]$ is positive definite, the estimate $\hat{\mathbf{z}}_{k+1}$ in (40) is actually a local minimum for J_{k+1}. Notice also that besides the term

$$-\left(P_{k+1|k}^{-1} + \alpha\mathcal{H}[H_{k+1}(\hat{\mathbf{z}}_{k|k})]\right)^{-1}\alpha\nabla_{\mathbf{z}} H_{k+1}(\hat{\mathbf{z}}_{k|k}) \quad (42)$$

in (40), the resulting state estimator filter has a predictor - corrector structure as the standard Kalman filter. Indeed the term in (42) will be null if $\hat{\mathbf{z}}_{k|k}$ is a local minima of H_{k+1}. This method is characterized by a low computational effort making it suitable for real-time applications. In the next section further details about practical implementation are addressed.

In order to illustrate the robustness properties of the algorithm, it has been tested on the same numerical experiment performed in the previous section with the addition of some outliers in the range measurements. Indeed, if acoustic sensors are employed to acquire range measurements a significant source of non-Gaussian noise is the multi-path phenomena. Therefore building on (2), the outliers have been generated according to the following measurement model:

$$y(t) = (\mu(t)\,\|\mathbf{r}(t)\|)^2 + \epsilon, \qquad \epsilon \sim N(0,R) \quad (43)$$
$$\mu(t) = \begin{cases} 1 & \text{for inliers} \\ 2 & \text{for outliers.} \end{cases} \quad (44)$$

In the numerical simulations described in the following the outliers ($\mu = 2$) have been randomly generated with probability 1% (i.e. at each step the value of μ is 1 with probability 99% and 2 otherwise). As an exception to this strategy, a set of 50 consecutive outliers ($\mu = 2$) has been introduced starting at time 40 seconds in order to test the performance of the filter in extreme conditions. Outliers are depicted with a blue star in figure 2 for the sake of clarity, but of course the measurements were fed to the filters without making any use of the knowledge that a value was an outlier or not. The measured output and the resulting evolution of the position and current estimation errors obtained through both filters are plotted in figure 2 revealing the sensitivity of the standard Kalman filter to outliers as compared to the LEL filter. Notice that the Kalman filter generating the results in figure 2 is the ideal one, i.e. the measurement and state noise covariance matrices used to compute the Kalman gain are precisely the ones used in the model.

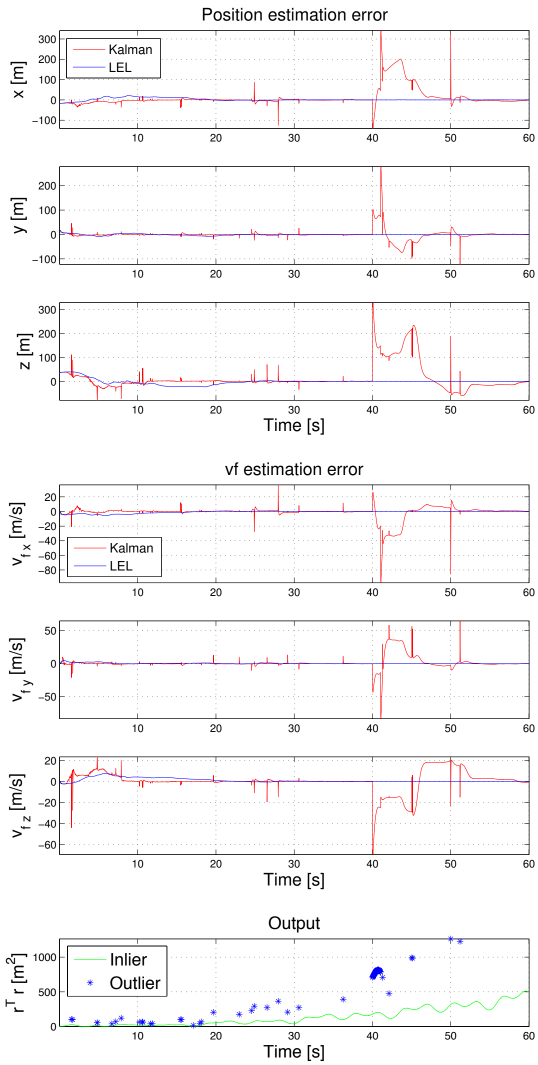

Fig. 2. Kalman and LEL filter estimations errors in presence of outliers: estimation error components of $\mathbf{p}$ and $\mathbf{v}_f$. Bottom plot: observed noisy output data with outliers (depicted as blue stars).

Besides the reported results, extensive simulations have been performed confirming that the LEL estimator outperforms the Kalman estimator in terms of outlier robustness.

V. Implementation issues for robust estimation

1) Recursive and finite memory formulations: The structures of the gradient (33) and hessian (34) of the LEL cost function reveal that they can be computed in a recursive form, exploiting the following update equations:

$$D_{k+1}(\mathbf{z}_{k+1}) = D_k + r^2_{k+1}(\mathbf{z}_{k+1}) \quad (45)$$

$$S_{k+1}(\mathbf{z}_{k+1}) = S_k + r^2_{k+1}(\mathbf{z}_{k+1}) \log r^2_{k+1}(\mathbf{z}_{k+1}). \quad (46)$$

This avoids storing all past data as only the described quantities at the current time step $k+1$ are needed making the computation suitable for real-time applications. Nevertheless, given the structure of (45 - 46) (notice that D_{k+1} is non decreasing) in line of principle there could be numerical overflow issues over time. To prevent these problems, an alternative strategy is a finite memory solution through a sliding window implementation. Indicating with N the size of such window, the update equations for D_{k+1} and S_{k+1} are

$$D_{k+1}(\mathbf{z}_{k+1}) = \left(\sum_{j=k+2-N}^{k} r^2_j(\hat{\mathbf{z}}_{j|j}) \right) + r^2_{k+1}(\mathbf{z}_{k+1}) \quad (47)$$

$$S_{k+1}(\mathbf{z}_{k+1}) = \left(\sum_{j=k+2-N}^{k} r^2_j(\hat{\mathbf{z}}_{j|j}) \log r^2_j(\hat{\mathbf{z}}_{j|j}) \right) + + r^2_{k+1}(\mathbf{z}_{k+1}) \log r^2_{k+1}(\mathbf{z}_{k+1}), \quad (48)$$

whereas (33) and (34) remain unchanged. As a result, the only values to be stored are the squared residuals $r^2_j(\hat{\mathbf{z}}_{j|j})$ relevant to the last N observations. This implementation has the additional advantage of forcing the filter to 'forget' measurements and eventually outliers in the distant past and placing more emphasis on recent measurements.

2) Regularization: As already noticed in analyzing the hessian matrix of the LEL cost function in (34), irrespective of the recursive or finite memory implementation of the filter, the LEL hessian matrix dependence from the term $\log r^2_{k+1}$ needs to be regularized to prevent unlikely, but theoretically possible divergences due to a perfectly null residual.

Notice that the perfect fitting case $D_{k+1} = 0$ can be thought as the limit case of complete absence of outliers (and otherwise ideal data). Indeed, given the properties of the LEL estimator [21], the complete absence of *any* outlier when $D_{k+1} > 0$ (non perfect fit) will result in a potentially poor estimation performance: this is because the LEL estimator is designed to find the fit that maximizes one relative squared residual (q_i in (28)) while minimizing the others so as to minimize the entropy-like loss function. Indeed, the absolute minimum of the LEL cost is achieved if all residuals but one are null. Hence, an additional regularization strategy that was implemented to derive the presented results consists in the following: in each time step with $D_{k+1} > 0$, and only for that time step, the largest residual $r^2_{i\,max}$ on the batch of N that are being processed is temporarily incremented by a certain scale factor, typically an order of magnitude. This implementation strategy guarantees the presence of a residual being sensibly larger than the others on the current batch of N residuals. Notice that N is either the size of the finite window or the total number of data depending on the chosen filter implementation (finite memory or not). Indeed, if the $r^2_{i\,max}$ is actually an outlier, the increase of its value is not expected to corrupt the estimation since it will still remain an outlier; otherwise, if $r^2_{i\,max}$ is an inlier (case of quasi-perfect measurements fit), it will become the only outlier in the observations set and the finite sample breakdown point [34] of the LEL function cost will continue to grant robustness.

As for the hessian divergence if $r^2_{k+1}(\mathbf{z}_{k+1}) = 0$ (but $D_{k+1} > 0$) this can be addressed by fixing an arbitrarily small threshold $\delta > 0$ and replacing $r^2_{k+1}(\mathbf{z}_{k+1})$ with δ in case $r^2_{k+1}(\mathbf{z}_{k+1})$ should be smaller than δ. Notice that due to the limited divergence speed of the $\log(\cdot)$ function for vanishing arguments, the δ value can be chosen to be very small. Indeed, in spite of the very extensive simulations ran to evaluate the proposed algorithm this regularization step was never necessary and never performed in practice.

3) Local nature of H: The entropy-like loss function $H_{k+1}(\cdot)$ is nonlinear and may have multiple local minima [21]. Moreover, its approximation by means of the second order Taylor series expansion has a local nature. As a consequence, special care needs to be taken in initializing the filter. Experience has shown that the Kalman filter estimate is a reliable candidate for initializing the proposed filter.

4) Tuning factor α: The loss function $H_{k+1}(\cdot) \in [0,1]$ by construction, so the parameter α in (26) is to be regarded as a tuning factor needed to make the two terms of J_{k+1} comparable. More precisely, the parameter α can be tuned trying to make the hessians of $J_{dynamical\ model}$ and J_{LEL} in (26) of the same order of magnitude. Following this guideline in the presented numerical experiments α has been kept constant and equal to 45. Future work will focus on algorithms to automatically and, perhaps, adaptively tune this gain online.

VI. Conclusions

The problem of single range based localization for a kinematics model of a $3D$ vehicle was addressed. The problem is relevant in several field robotics applications, particularly in underwater scenarios where ranges are measured acoustically and alternative localization devices as GPS are not available. The proposed solution allows to address the observability analysis and the state estimation filter design on an LTI state equation defined on $\mathbb{R}^8$ with a time varying scalar output equation. Moreover, to cope with possible outliers in the range measurements, a robust predictor - corrector state estimator has been proposed. Such filter builds on the novel Least Entropy-Like (LEL) parameter estimation paradigm illustrated in [21] that significantly departs from alternative robust state estimators based on M-estimators or heavy-tailed noise distributions.

References

[1] A. Martinelli and R. Siegwart, “Observability analysis for mobile robot localization,” in *Intelligent Robots and Systems, 2005. (IROS). 2005 IEEE/RSJ International Conference on*, aug. 2005, pp. 1471 – 1476.

[2] X. S. Zhou and S. I. Roumeliotis, “Robot-to-robot relative pose estimation from range measurements,” *IEEE Transactions on Robotics*, vol. 24, no. 6, pp. 1379–1393, 2008.

[3] Z. M. Kassas and T. E. Humphreys, “Observability analysis of opportunistic navigation with pseudorange measurements,” in *AIAA Guidance, Navigation, and Control Conference, AIAA GNC*, 2012.

[4] T. L. Song, “Observability of target tracking with range-only measurements,” *Oceanic Engineering, IEEE Journal of*, vol. 24, no. 3, pp. 383 – 387, jul 1999.

[5] T. Maki, T. Matsuda, T. Sakamaki, T. Ura, and J. Kojima, “Navigation method for underwater vehicles based on mutual acoustical positioning with a single seafloor station,” *IEEE Journal Of Oceanic Engineering*, vol. 38, no. 1, pp. 167 – 177, 2013.

[6] F. Arrichiello, G. Antonelli, A. P. Aguiar, and A. Pascoal, “Observability metric for the relative localization of AUVs based on range and depth measurements: Theory and experiments,” in *Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on*, sept. 2011, pp. 3166 –3171.

[7] R. Hermann and A. J. Krener, “Nonlinear controllability and observability,” *IEEE Transactions on Automatic Control*, vol. 22, no. 5, pp. 728 – 740, October 1977.

[8] J. Jouffroy and J. Reger, “An algebraic perspective to single-transponder underwater navigation,” in *Computer Aided Control System Design, 2006 IEEE International Conference on Control Applications, IEEE International Symposium on Intelligent Control*. IEEE, 2006, pp. 1789–1794.

[9] J. Jouffroy and A. Ross, “Remarks on the observability of single beacon underwater navigation,” in *Int. Symp. on Unmanned Untethered Submersible Technology (UUST 05)*, Durham, NH, August 2005, p. .

[10] A. S. Gadre and D. J. Stilwell, “Toward underwater navigation based on range measurements from a single location,” in *Proceedings of IEEE International Conference on Robotics and Automation (ICRA 2004)*, vol. 5, New Orleans, LA, USA, 26 April - 1 May 2004 2004, pp. 4472 – 4477.

[11] P. Batista, C. Silvestre, and P. Oliveira, “Single range navigation in the presence of constant unknown drifts,” in *Control Conference (ECC), 2009 European*, Aug 2009, pp. 3983–3988.

[12] ——, “Single beacon navigation: Observability analysis and filter design,” in *Proceedings of the 2010 American Control Conference, ACC2010*, Marriott Waterfront, Baltimore, MD, USA, June 30 – July 02 2010, pp. 6191 – 6196.

[13] ——, “Single range aided navigation and source localization: Observability and filter design,” *Systems & Control Letters*, vol. 60, pp. 665–673, 2011.

[14] W. J. Rugh, *Linear System Theory*. Prentice-Hall, 1996.

[15] S. H. Dandach, B. Fidan, S. Dasgupta, and B. D. Anderson, “A continuous time linear adaptive source localization algorithm, robust to persistent drift,” *Systems & Control Letters*, vol. 59, pp. 7–16, 2009.

[16] G. Indiveri, P. Pedone, and M. Cuccovillo, “Fixed target 3D localization based on range data only: A recursive least squares approach,” in *Proceedings of the 2012 IFAC Workshop on Navigation, Guidance and Control of Underwater Vehicles, IFAC-NGCUV2012*, Porto, Portugal, 10 - 12 April 2012, pp. 140 – 145.

[17] J. Vaganay, J. Leonard, and J. Bellingham, “Outlier rejection for autonomous acoustic navigation,” in *Proceedings of the 1996 IEEE International Conference on Robotics and Automation, ICRA 1996*, vol. 3. IEEE, 1996, pp. 2174–2181.

[18] S. Vike and J. Jouffroy, “Diffusion-based outlier rejection for underwater navigation,” in *Proceedings of MTS/IEEE OCEANS, 2005*, vol. 2. IEEE, 2005, pp. 1037–1042.

[19] A. Alcocer, P. Oliveira, and A. M. Pascoal, “Study and implementation of an EKF GIB-based underwater positioning system,” *Control engineering practice*, vol. 15, no. 6, pp. 689–701, 2007.

[20] J. Jouffroy and J. Opderbecke, “Underwater vehicle navigation using diffusion-based trajectory observers,” *Oceanic Engineering, IEEE Journal of*, vol. 32, no. 2, pp. 313 – 326, 2007.

[21] G. Indiveri, “An entropy-like estimator for robust parameter identification,” *Entropy*, vol. 11, no. 4, pp. 560–585, 2009.

[22] P. Batista, C. Silvestre, and P. Oliveira, “On the observability of linear motion quantities in navigation systems,” *Systems & Control Letters*, vol. 60, no. 2, pp. 101–110, 2011.

[23] L. M. Silverman and B. D. Anderson, “Controllability, observability and stability of linear systems,” *SIAM Journal on Control*, vol. 6, no. 1, pp. 121–130, 1968.

[24] A. H. Jazwinski, *Stochastic processes and filtering theory*. Courier Dover Publications, 2007.

[25] Z. Gao and D. Ho, “Proportional multiple-integral observer design for descriptor systems with measurement output disturbances,” *Control Theory and Applications, IEE Proceedings*, vol. 151, no. 3, pp. 279–288, May 2004.

[26] Z. Gao and H. Wang, “Descriptor observer approaches for multivariable systems with measurement noises and application in fault detection and diagnosis,” *Systems & Control Letters*, vol. 55, no. 4, pp. 304–313, 2006.

[27] D. Simon, *Optimal state estimation: Kalman, H infinity, and nonlinear approaches*. John Wiley & Sons, 2006.

[28] T. Cipra and R. Romera, “Kalman filter with outliers and missing observations,” *Test*, vol. 6, no. 2, pp. 379–395, 1997.

[29] Z. Ðurović and B. Kovačević, “Robust estimation with unknown noise statistics,” *Automatic Control, IEEE Transactions on*, vol. 44, no. 6, pp. 1292–1296, 1999.

[30] S. Chan, Z. Zhang, and K. Tse, “A new robust Kalman filter algorithm under outliers and system uncertainties,” in *Circuits and Systems. ISCAS 2005. IEEE International Symposium on*. IEEE, 2005, pp. 4317–4320.

[31] P. J. Huber, *Robust statistics*. Wiley, New York, 1981.

[32] J.-A. Ting, E. Theodorou, and S. Schaal, “A Kalman filter for robust outlier detection,” in *Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ International Conference on*. IEEE, 2007, pp. 1514–1519.

[33] G. Agamennoni, J. I. Nieto, and E. M. Nebot, “An outlier-robust Kalman filter,” in *Robotics and Automation (ICRA), 2011 IEEE International Conference on*. IEEE, May 2011, pp. 1551–1558.

[34] P. J. Rousseeuw and A. M. Leroy, *Robust regression and outlier detection*. John Wiley & Sons, 2005.