Robust Extended Kalman Filtering for Systems with Measurement Outliers

Fang, Huazhen; Haile, Mulugeta; Wang, Yebin

TR2021-078 July 12, 2021

Abstract

Outliers can contaminate the measurement process of many nonlinear dynamic systems, which can be caused by sensor errors, model uncertainties, changes in ambient environment, data loss or malicious cyber attacks. When the extended Kalman filter (EKF) is applied to such systems for state estimation, the outliers can seriously reduce the estimation accuracy. This paper proposes an innovation saturation mechanism to make the EKF robust against outliers. This mechanism applies a saturation function to the innovation process that the EKF leverages to correct the state estimation. As such, when outliers occur, the distorted innovation is saturated so as not to undermine the state estimation. The mechanism features an adaptive adjustment of the saturation bounds. The design leads to the development robust EKF approaches for both continuous- and discrete-time systems. The stability of the proposed approaches when applied to linear systems is characterized, showing that they are capable of performing bounded-error estimation in the presence of bounded outlier disturbances in this case. A simulation study about mobile robot localization is presented to illustrate the efficacy of the proposed design. Compared to existing methods, the proposed approaches can effectively reject outliers of various magnitudes, types and durations, at significant computational efficiency and without requiring measurement redundancy.

IEEE Transactions on Control Systems Technology

Robust Extended Kalman Filtering for Systems with Measurement Outliers

Huazhen Fang, Mulugeta A. Haile and Yebin Wang

***Abstract*—Outliers can contaminate the measurement process of many nonlinear dynamic systems, which can be caused by sensor errors, model uncertainties, changes in ambient environment, data loss or malicious cyber attacks. When the extended Kalman filter (EKF) is applied to such systems for state estimation, the outliers can seriously reduce the estimation accuracy. This paper proposes an innovation saturation mechanism to make the EKF robust against outliers. This mechanism applies a saturation function to the innovation process that the EKF leverages to correct the state estimation. As such, when outliers occur, the distorted innovation is saturated so as not to undermine the state estimation. The mechanism features an adaptive adjustment of the saturation bounds. The design leads to the development robust EKF approaches for both continuous- and discrete-time systems. The stability of the proposed approaches when applied to linear systems is characterized, showing that they are capable of performing bounded-error estimation in the presence of bounded outlier disturbances in this case. A simulation study about mobile robot localization is presented to illustrate the efficacy of the proposed design. Compared to existing methods, the proposed approaches can effectively reject outliers of various magnitudes, types and durations, at significant computational efficiency and without requiring measurement redundancy.**

***Index Terms*—Kalman filter, extended Kalman filter, robust estimation, measurement outlier, localization.**

I. Introduction

The Kalman filter (KF) is arguably the most celebrated estimation technique in the literature, which can optimally estimate the state of a linear dynamic system on the basis of a model and a stream of noisy measurements. In practice, it is the extended KF (EKF), a nonlinear variant of the KF, that is the most widely used, because real-world systems usually involve nonlinearities [1]. The EKF's applications range from control systems to signal processing, system health monitoring, navigation and econometrics. However, a major challenge for high-quality estimation in practice is the measurement outliers, which can come from a diversity of sources, e.g., unreliable sensors, environmental variability, data dropouts in transmission, channel biases, incorrect assumptions about noises, model mismatch, and data falsification attacks from cyberspace [2]–[7]. When outliers corrupt the measurement data, the performance of the EKF can be seriously limited or degraded. As such, without adequate robustifcaition against outliers, the EKF will not be viable for a variety of real-world practical estimation problems.

H. Fang is with Department of Mechanical Engineering, University of Kansas, Lawrence, KS 66045, USA (e-mail: fang@ku.edu) .

M. A. Haile is with Vehicle Technology Directorate, US Army Research Laboratory, Aberdeen, MD 21005, USA (e-mail: mulugeta.a.haile.civ@mail.mil).

Y. Wang is with Mitsubishi Electric Research Laboratories, Cambridge, MA 02139, USA (e-mail: yebinwang@ieee.org).

Literature review. Robust state estimation against measurement outliers has attracted significant interest from researchers during the past years. A majority of the effort has been devoted to robustifying the standard linear KF. One can divide the existing methods mainly into three main categories. The first category models the measurement noises using heavy-tailed distributions rather than exponential distributions, e.g., the Gaussian distribution as often assumed in the classical KF, to capture the occurrence of an outlier. Heavy-tailed Gaussian-mixture [8], [9] and t-distributed noise models [10] have been used to modify the KF for better robustness. The study [11] view an outlier as a result of measurement noises with variable covariances. Assuming the noise covariances to follow an inverse Gamma distribution, it adds to the KF a procedure of adaptive identification of the key parameters involved in the inverse Gamma distribution. Methods of the second category seek to assign the measurement at each time with a weight, in an attempt to downweight outlying measurements. In [2], an expectation-maximization algorithm is used to enable adaptive determination of the weight for a measurement. The results are generalized in [12] and then extended to the smoothing problem. In [13], [14], a measurement-weighting-based pre-whitening procedure is designed to decorrelate outliers from normal measurements as a basis for building a robust KF. The third category of methods extends the KF to conduct simultaneous state and input estimation, regarding an outlier as an input added to the measurement and estimating it together with the state. To accomplish this, a few methods have been built on minimum variance unbiased estimation [15]–[20]. Further, Bayesian methods are developed in [21] to achieve joint state and outlier estimation through probabilistic filtering.

Although much attention has been given to outlier rejection for the linear KF, it is more important and pressing to robustify the EKF due to its practical significance. The EKF run relies on linearized approximations based on the most recent estimation. When measurement outliers arise, they would increase the state estimation error, which in turn will amplify the error involved in the linearized approximations. This may further drive away the estimation at the next time instant, potentially leading to divergence. Hence, the EKF is more vulnerable to outliers and needs effective ways to reject them. However, the current literature includes only few studies in this regard, due to the associated difficulty. In [22], the EKF is blended with a procedure that detects an outlier by evaluating the probability of its occurrence based on innovation statistics. The method in [13] is modified in [23] to deal with outliers affecting the

EKF run.

It is noteworthy that the above robust KF/EKF techniques, despite their importance, generally involve a dramatic increase in computational complexity, due to iterative optimization or other computationally expensive procedures to detect or suppress outliers at every time instant. Besides, some of them require measurement redundancy to differentiate outliers from normal measurements or estimate them directly. This, however, is not always possible, because a real system often allows only a limited number of sensors to be deployed.

In addition to the above outlier-robust KF/EKF methods, one can also find other types of estimation approaches in the literature to suppress outliers. Among them is the well-known $\mathcal{H}_\infty$ filtering [24], [25], which considers outliers as unknown yet bounded uncertainty. Yet, this approach introduces conservatism as it performs worst-case estimation by design. A stubborn observer is developed in [26], which employs a saturation function in the output injection signal to mitigate the influence of outliers. This method is not only computationally fast, but also can deal with very large outliers. Nonetheless, it is applicable to only linear systems suffering outliers that occur occasionally and individually.

Statement of contribution. In this work, we offer a new design to enhance the robustness of the EKF against measurement outliers, presenting a three-fold contribution. First, we propose a unique innovation saturation mechanism to reject outliers and ensure the performance of the EKF. The innovation plays a key role in correcting the state prediction in the EKF but can be distorted by outliers. To overcome this vulnerability, our mechanism saturates the innovation when it is unreasonably large in order to reduce the effects of outliers. At the core of the mechanism is a procedure for adaptively adjusting the saturation bounds to effectively grasp the change of the innovation. Along this line, we develop the innovation-saturated EKF (IS-EKF) for both continuous- and discrete-time systems. Second, we analyze the stability of the proposed IS-EKF for the linear case, proving that, if applied to linear systems, it produces bounded-error estimation when outlier disturbances are bounded. Finally, we apply the proposed IS-EKF to the problem of mobile robot localization, demonstrating its effectiveness in providing reliable estimation in the presence of measurement outliers.

We point that the idea of innovation saturation was first considered in [26] to suppress outliers affecting a linear state observer and then exploited in our previous work [27] to robustify the linear KF. However, both studies by design can only deal with outliers that appear singly or one at a time on a linear system, as is with various other methods in the literature. By contrast, our approaches can handle outliers of different magnitudes, types, and durations and imposed on nonlinear systems. They are structurally concise, computationally efficient, and free from requiring measurement redundancy. These advantages well lend them to practical application.

*Organization.*This paper is organized as follows. Section II develops the IS-EKF for nonlinear continuous-time systems and analyzes its stability for linear systems. Section III extends the results to discrete-time systems. A simulation example based on mobile robot localization is provided in Section IV to illustrate the usefulness of the proposed design. Finally, Section V summarizes the concluding remarks.

Notation: Notations used throughout this paper are standard. The n-dimensional Euclidean space is denoted as $\mathbb{R}^n$. For a vector, $\|\cdot\|$ denotes its 2-norm. The notation I is an identity matrix; $X > 0$ (≥ 0) means that X is a real, symmetric and positive definite (semidefinite) matrix; for a symmetric block matrix, we use a star ($\star$) to represent a symmetry-induced block in a matrix; the notation $\mathrm{diag}(\dots)$ stands for a block-diagonal matrix. The minimum and maximum eigenvalues of a real, symmetric matrix are denoted by $\underline{\lambda}(\cdot)$ and $\bar{\lambda}(\cdot)$, respectively. Matrices are assumed to be compatible for algebraic operations, if their dimensions are not explicitly stated.

II. IS-EKF for Continuous-Time Systems

This section develops the IS-EKF approach for a nonlinear continuous-time system and then offers analysis of its stability in the linear case.

A. IS-EKF Architecture

Consider the following model:

$$\begin{cases} \dot{x}_t = f(x_t) + w_t, \\ y_t = h(x_t) + Dd_t + v_t, \end{cases} \tag{1}$$

where $x \in \mathbb{R}^n$ is the state vector, $y \in \mathbb{R}^p$ the measurement vector, and $w_t \in \mathbb{R}^n$ and $v_t \in \mathbb{R}^p$ zero-mean, mutually independent noises with covariances given by $Q \geq 0$ and $R > 0$, respectively. The nonlinear mappings f and h represent the state evolution and measurement functions, respectively. Note that the measurement y_t is subjected to the outlier effects caused by an unknown disturbance $d_t \in \mathbb{R}^m$. The matrix D shows the relation between d_t and y_t and is assumed to be unknown. While an input-free model is considered as in (1), the state estimation design in sequel can be readily extended to an input-driven model.

Modifying the conventional EKF, we propose the following IS-EKF procedure:

$$\dot{\hat{x}}_t = f(\hat{x}_t) + K_t \cdot \mathrm{sat}_\sigma\left(y_t - h\left(\hat{x}_t\right)\right), \tag{2a}$$

$$K_t = P_t C_t^\top R^{-1}, \tag{2b}$$

$$\dot{P}_t = A_t P_t + P_t A_t^\top + Q - K_t R K_t^\top, \tag{2c}$$

where K_t is the estimation gain matrix, and P_t is a positive definite matrix that approximately represents the estimation error covariance in the standard EKF, and

$$A_t = \left.\frac{\partial f}{\partial x}\right|_{\hat{x}_t}, \quad C_t = \left.\frac{\partial h}{\partial x}\right|_{\hat{x}_t}.$$

Recall that, for the conventional EKF, the state estimation is corrected by the innovation $(y_t - h(\hat{x}_t))$. Its effectiveness, however, can be compromised if y_t is corrupted by an outlier. To address this issue, we use a saturated innovation instead, as shown in (2a). Specifically, it is defined as

$$\mathrm{sat}_\sigma\left(y_t - h(\hat{x}_t)\right) = \begin{bmatrix} \vdots \\ \mathrm{sat}_{\sqrt{\sigma_i}}\left(y_{i,t} - h_i(\hat{x}_t)\right) \\ \vdots \end{bmatrix}, \tag{3}$$

where $\sigma_i > 0$, y_i is the i-th element of y, and h_i the i-th element of h. For a variable r, the saturation function is defined as $\mathrm{sat}_\epsilon(r) = \max\{-\epsilon, \min\{\epsilon, r\}\}$. For (3), the saturation range $[-\sqrt{\sigma_i}, \sqrt{\sigma_i}]$ can be loosely viewed as an anticipated range of the innovation. If falling within this range, the innovation is considered as reasonable and applied without change to update the state estimation. Otherwise, it may be affected by an outlier and thus saturated to prevent the outlier from dragging the estimation away from a correct course.

For the EKF, it is observed that a saturation function with fixed upper and lower bounds is inadequate to reject outliers, as it may either confuse with an outlier a certain measurement generating a large innovation, or miss an outlier approximately falling within the saturation range. More often than not, one can also find it practically difficult to choose the fixed bounds, especially when knowledge of the outliers is scarce. Therefore, we propose the following procedure to adaptively adjust the saturation bounds:

$$\dot{\sigma}_{i,t} = \lambda_{1,i}\sigma_{i,t} + \gamma_{1,i}\varepsilon_{i,t}e^{-\varepsilon_{i,t}}, \quad \sigma_{i,0} > 0, \tag{4a}$$

$$\dot{\varepsilon}_{i,t} = \lambda_{2,i}\varepsilon_{i,t} + \gamma_{2,i}\left(y_{i,t} - h_i(\hat{x}_t)\right)^2, \quad \varepsilon_{i,0} > 0, \tag{4b}$$

for $i = 1, 2, \ldots, p$, where $\lambda_{1,i}, \lambda_{2,i} < 0$ and $\gamma_{1,i}, \gamma_{2,i} > 0$. For convenience of notation, we define

$$\Lambda_i = \mathrm{diag}([\cdots \ \lambda_{i,j} \ \cdots]), \ \Gamma_i = \mathrm{diag}([\cdots \ \gamma_{i,j} \ \cdots]),$$

for $i = 1, 2$ and $j = 1, 2, \cdots, p$.

Based on (4), σ_i will dynamically change driven by the innovation $(y_{i,t} - h_i(\hat{x}_t))$ to enable the adaptation of the saturation bounds. This mechanism specifically involves a double-layer structure. The lower layer, based on (4b), tracks the changes in the innovation signal — the variable ε_i will keep itself at an appropriate level when the innovation is normal but become large when the innovation is altered by outliers. The upper layer, based on (4a), is concerned with adjusting the saturation bounds. As is seen, σ_i will rapidly diminish when ε_i is large due to outliers; it will also be driven up by a relatively small ε_i. Adapting σ_i like this, the design proposed in (4) will achieve a discernment between an outlier and a normal measurement, filtering away innovation if it is corrupted by outliers and allowing it to pass through otherwise to correct the prediction.

B. Stability Analysis for Linear Systems

It has been widely acknowledged as a challenge to determine the exact conditions for the asymptotic stability of the EKF, even though there exist some studies. The analysis will be even more difficult for the IS-EKF, because of the added innovation saturation procedure and the nonlinear update of the saturation bounds. To formulate a tractable analysis, we restrict our attention to the asymptotic stability of the IS-EKF for a linear deterministic system:

$$\begin{cases} \dot{x}_t = Ax_t, \\ y_t = Cx_t + Dd_t. \end{cases}$$

For this system, the IS-EKF acts as a state observer, and the estimation is performed by the innovation-saturated KF. Here, we assume that $(A, Q^{\frac{1}{2}})$ is stabilizable and that (A, C) is detectable, as often needed for estimation.

Let us first define the state estimation error as $e_t = \hat{x}_t - x_t$. The dynamics of e_t is governed by

$$\dot{e}_t = Ae_t - K_t \cdot \mathrm{sat}_\sigma(Ce_t - Dd_t). \tag{5}$$

To proceed further, we define the following matrix

$$S_t = \begin{bmatrix} M_t - \alpha P_t^{-1} & -C^\top\left(R^{-1} + W\right) & C^\top(\Gamma_2 - R^{-1})D \\ \star & 2W & WD \\ \star & \star & U \end{bmatrix},$$

where $M_t = P_t^{-1}QP_t^{-1} + C^\top(R^{-1} - \Gamma_2)C$, W is a diagonal positive definite matrix, U a positive definite matrix, and $\alpha > 0$ a positive scalar. Furthermore, we recall a well-known fact [28]: if $(A, Q^{\frac{1}{2}})$ is stabilizable and (A, C) detectable, P_t for $P_0 \geq 0$ in (2c) will approach a unique positive-definite solution P_∞ satisfying

$$AP_\infty + P_\infty A^\top + Q - P_\infty C^\top R^{-1} CP_\infty = 0.$$

We obtain the following result regarding the stability of the proposed IS-EKF for the linear deterministic case.

Theorem 1. *Suppose $\|d_t\| \leq \mu < \infty$ and $\rho = e^{-1}\sum_i \gamma_{1,i} < \infty$, where $\mu, \rho > 0$. If there exist P_0, W, U, α and Γ such that $S_t \geq 0$ and $0 < \alpha \leq -\max\{\ldots, \lambda_{1,i}, \lambda_{2,i}, \ldots\}$ for $i = 1, 2, \ldots, p$, then the estimation error e_t is upper bounded with*

$$\|e_t\| \leq \sqrt{\frac{1}{c_2}\left[e^{-\alpha t}V_0 + \frac{1}{\alpha}(1 - e^{-\alpha t})(c_1\mu^2 + \rho)\right]}, \tag{6a}$$

$$\lim_{t\to\infty} \|e_t\| \leq \sqrt{\frac{c_1\mu^2 + \rho}{\alpha c_3}}, \tag{6b}$$

where $c_1 = \bar{\lambda}(U + D^\top\Gamma_2 D)$, $c_2 = \underline{\lambda}(P_t^{-1})$ and $c_3 = \underline{\lambda}(P_\infty^{-1})$.

Proof: We consider using the Lyapunov function

$$V_t = e_t^\top P_t^{-1} e_t + \textstyle\sum_i \sigma_{i,t} + \sum_i \varepsilon_{i,t}.$$

The first-order time derivative of V_t along (5) is

$$\begin{aligned} \dot{V}_t &= 2e_t^\top P_t^{-1}\dot{e}_t + e_t^\top \frac{\mathrm{d}\left(P_t^{-1}\right)}{\mathrm{d}t} e_t + \sum_i \dot{\sigma}_{i,t} + \sum_i \dot{\varepsilon}_{i,t} \\ &= 2e_t^\top P_t^{-1}\left[Ae_t - K_t \cdot \mathrm{sat}_\sigma\left(Ce_t - Dd_t\right)\right] \\ &\quad - e_t^\top P_t^{-1}\left(AP_t + P_tA^\top + Q - K_tRK_t^\top\right)P_t^{-1}e_t \\ &\quad + (Ce_t - Dd_t)^\top\Gamma_2(Ce_t - Dd_t) + \textstyle\sum_i \lambda_{1,i}\sigma_{i,t} \\ &\quad + \textstyle\sum_i \lambda_{2,i}\varepsilon_{i,t} + \sum_i \gamma_{1,i}\varepsilon_{i,t}e^{-\varepsilon_{i,t}} \\ &\leq -e_t^\top P_t^{-1}QP_t^{-1}e_t + e_t^\top C^\top(R^{-1} + \Gamma)Ce_t \\ &\quad - 2e_t^\top C^\top R^{-1}\mathrm{sat}_\sigma\left(Ce_t - Dd_t\right) - 2e_t^\top C^\top\Gamma_2 Dd_t \\ &\quad + d_t^\top D^\top\Gamma_2 Dd_t + \textstyle\sum_i \lambda_{1,i}\sigma_{i,t} + \sum_i \lambda_{2,i}\varepsilon_{i,t} + \rho, \end{aligned}$$

where the relation $\varepsilon_{i,t}e^{-\varepsilon_{i,t}} \leq e^{-1}$ is used. Let us define $s_t = Ce_t - Dd_t - \mathrm{sat}_\sigma(Ce_t - Dd_t)$. Then,

$$\begin{aligned} \dot{V}_t &\leq -e_t^\top M_t e_t + 2e_t^\top C^\top R^{-1}s_t + 2e_t^\top C^\top(R^{-1} - \Gamma_2)Dd_t \\ &\quad + d_t^\top D^\top\Gamma_2 Dd_t + \textstyle\sum_i \lambda_{1,i}\sigma_{i,t} + \sum_i \lambda_{2,i}\varepsilon_{i,t} + \rho. \end{aligned}$$

By [29, Lemma 1.6], we have

$$-s_t^\top W\left(s_t - Ce_t + Dd_t\right) \geq 0.$$

It then follows that

$$\begin{aligned}
\dot{V}_t &\leq \dot{V}_t - 2s_t^\top W\left(s_t - Ce_t + Dd_t\right) \\
&\leq -e_t^\top M_t e_t + 2e_t^\top C^\top\left(R^{-1}+W\right)s_t - 2s_t^\top W s_t \\
&\quad + 2e_t^\top C^\top (R^{-1}-\mathit{\Gamma}_2)Dd_t - 2s_t^\top W D d_t + \\
&\quad + d_t^\top D^\top \mathit{\Gamma}_2 D d_t + \textstyle\sum_i \lambda_{1,i}\sigma_{i,t} + \sum_i \lambda_{2,i}\varepsilon_{i,t} + \rho \\
&= -\begin{bmatrix} e_t \\ s_t \\ d_t \end{bmatrix}^\top \begin{bmatrix} M_t & -C^\top\left(R^{-1}+W\right) & C^\top(\mathit{\Gamma}_2 - R^{-1})D \\ \star & 2W & WD \\ \star & \star & U \end{bmatrix} \\
&\quad \cdot \begin{bmatrix} e_t \\ s_t \\ d_t \end{bmatrix} + d_t^\top\left(U + D^\top \mathit{\Gamma}_2 D\right)d_t + \textstyle\sum_i \lambda_{1,i}\sigma_{i,t} + \sum_i \lambda_{2,i}\varepsilon_{i,t} \\
&\quad + \rho.
\end{aligned}$$

If $S_t \geq 0$, we have

$$\begin{aligned}
\dot{V}(t) \leq& -\alpha e_t^\top P_t^{-1} e_t - \alpha \textstyle\sum_i \sigma_{i,t} - \alpha \sum_i \varepsilon_{i,t} \\
&+ d_t^\top\left(U + D^\top \mathit{\Gamma}_2 D\right)d_t + \textstyle\sum_i (\lambda_{1,i}+\alpha)\sigma_{i,t} \\
&+ \textstyle\sum_i (\lambda_{2,i}+\alpha)\varepsilon_{i,t} + \rho.
\end{aligned}$$

If $0 < \alpha \leq -\max(\lambda_i)$, one has $\lambda_i + \alpha \leq 0$. Then,

$$\begin{aligned}
\dot{V}_t &\leq -\alpha V_t + d_t^\top\left(U + D^\top \mathit{\Gamma}_2 D\right)d_t + \rho \\
&\leq -\alpha V_t + c_1\|d_t\|^2 + \rho \\
&\leq -\alpha V_t + c_1\mu^2 + \rho.
\end{aligned}$$

Hence,

$$V_t \leq e^{-\alpha t}V_0 + \frac{1}{\alpha}(1-e^{-\alpha t})(c_1\mu^2 + \rho).$$

Furthermore, $V_t \geq c_2\|e_t\|^2$. Then,

$$\|e_t\|^2 \leq \frac{1}{c_2}\left[e^{-\alpha t}V_0 + \frac{1}{\alpha}(1-e^{-\alpha t})(c_1\mu^2+\rho)\right],$$

which implies (7a). When $t \to \infty$, we can obtain (7b). •

Theorem 1 shows that, when applied to a noise-free linear system with upper-bounded outliers, the proposed IS-EKF scheme can lead to bounded-error estimation if some conditions are satisfied. Further, the following corollary can be developed.

Corollary 1. *Suppose $\|d_t\| \leq \mu < \infty$ for $\mu > 0$. If there exist P_0, W, U, α and $\mathit{\Gamma}$ such that $S_t \geq 0$ and $0 < \alpha \leq -\max\{\ldots, \lambda_{1,i}, \lambda_{2,i} + \gamma_{1,i}, \ldots\}$ for $i = 1, 2, \ldots, p$, then the estimation error e_t is upper bounded with*

$$\|e_t\| \leq \sqrt{\frac{1}{c_2}\left[e^{-\alpha t}V_0 + \frac{1}{\alpha}(1-e^{-\alpha t})c_1\mu^2\right]}, \tag{7a}$$

$$\lim_{t\to\infty}\|e_t\| \leq \sqrt{\frac{c_1}{\alpha c_3}}\mu. \tag{7b}$$

To prove this result, one can generally follow the proof of Theorem 1 while upper-bounding $\varepsilon_{i,t}e^{-\varepsilon_{i,t}}$ by $\varepsilon_{i,t}$ instead of $\leq e^{-1}$. Corollary 1 implies that e_t will exponentially approach zero as $t \to \infty$ if $\lim_{t\to\infty} d_t = 0$.

Remark 1. *According to Theorem 1 and Corollary 1, selection of P_0 can be important for making the condition $S_t \geq 0$ satisfied. Given the structure of S_t, it is cautioned that too large a P_0 may bring the risk of divergent estimation. However, P_t is monotonically non-decreasing for $P_0 = 0$ if $(A, Q^{\frac{1}{2}})$ is stabilizable and (A, C) detectable [28]. Leveraging this property, it is suggested that P_0 be set to be small enough (close to zero in particular if some accurate prior knowledge about the initial state is available) when the IS-EKF approach is to be implemented.* •

III. IS-EKF for Discrete-Time Systems

Extending Section II, this section investigates the development of IS-EKF for nonlinear discrete-time systems.

A. IS-EKF Architecture

Consider a nonlinear discrete-time model

$$\begin{cases} x_{k+1} = f(x_k) + w_k, \\ \quad y_k = h(x_k) + Dd_k + v_k. \end{cases} \tag{8}$$

The notations in above are the same as in Section II. Still, the noises w_k and v_k are zero-mean, mutually independent with covariances $Q \geq 0$ and $R > 0$, respectively.

For this system, we propose the IS-EKF as follows:

$$\hat{x}_{k|k-1} = f\left(\hat{x}_{k-1|k-1}\right), \tag{9a}$$

$$P_{k|k-1} = A_{k-1}P_{k-1|k-1}A_{k-1}^\top + Q, \tag{9b}$$

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k \cdot \mathrm{sat}_\sigma\left(y_k - h\left(\hat{x}_{k|k-1}\right)\right), \tag{9c}$$

$$K_k = P_{k|k-1}C_k^\top\left(C_k P_{k|k-1} C_k^\top + R\right)^{-1}, \tag{9d}$$

$$P_{k|k} = P_{k|k-1} - K_k\left(C_k P_{k|k-1} C_k^\top + R\right)K_k^\top, \tag{9e}$$

where $\hat{x}_{k|k-1}$ is the one-step-forward prediction of x_k, $\hat{x}_{k|k}$ the updated estimate when y_k arrives to correct the prediction, and

$$A_k = \left.\frac{\partial f}{\partial x}\right|_{\hat{x}_{k|k}}, \quad C_k = \left.\frac{\partial h}{\partial x}\right|_{\hat{x}_{k|k-1}}.$$

In addition, K_k is the estimation gain, and $P_{k|k-1}$ and $P_{k|k}$ the approximate estimation error covariances in the standard EKF. Akin to (2), we use an innovation saturation mechanism to deal with measurement outliers, as shown in (9c). In analogy to (4), σ is dynamically adjusted by

$$\sigma_{i,k+1} = \lambda_{1,i}\sigma_{i,k} + \gamma_{2,i}\varepsilon_{i,k}e^{-\varepsilon_{i,k}}, \ \sigma_{i,0} > 0, \tag{10a}$$

$$\varepsilon_{i,k+1} = \lambda_{2,i}\varepsilon_{i,k} + \gamma_{2,i}\left(y_{i,k} - C_i\hat{x}_{k|k-1}\right)^2, \ \varepsilon_{i,0} > 0, \tag{10b}$$

for $i = 1, 2, \ldots, p$, where $0 < \lambda_{1,i}, \lambda_{2,i} < 1$ and $\gamma_{1,i}, \gamma_{2,i} > 0$.

B. Stability Analysis for Linear Systems

We consider the stability analysis for the above IS-EKF when it is applied to a linear deterministic system:

$$\begin{cases} x_{k+1} = Ax_k, \\ \quad y_k = Cx_k + Dd_k. \end{cases}$$

Defining the state prediction error as $e_k = \hat{x}_{k|k-1} - x_k$, its dynamics can be expressed as

$$\begin{aligned} e_{k+1} &= \hat{x}_{k+1|k} - x_{k+1} \\ &= Ae_k - AK_k \cdot \mathrm{sat}_\sigma\left(Ce_k - Dd_k\right). \end{aligned} \tag{11}$$

Before proceeding further, we show some results that will be needed later. Suppose that $(A, Q^{\frac{1}{2}})$ is stabilizable and that (A, C) detectable. Then, $P_{k|k-1}$ will converge to a fixed positive definite matrix P_∞ that satisfies

$$\begin{aligned} P_\infty = AP_\infty A^\top + Q - AP_\infty C^\top \left(CP_\infty C^\top + R\right)^{-1} \\ \cdot CP_\infty A^\top. \end{aligned}$$

It is also known that $P_{k|k-1}$ is upper and lower bounded, and so is $P_{k|k}$. Hence, there should exist ϵ such that $P_{k|k}^{-1} \le \epsilon I$. Then, if A is invertible,

$$\begin{aligned} &P_{k+1|k}^{-1} = \left(AP_{k|k}A^\top + Q\right)^{-1} \\ &= A^{-\top}\left(P_{k|k} + A^{-1}QA^{-\top}\right)^{-1}A^{-1} \\ &= A^{-\top}\left[P_{k|k}^{-1} - P_{k|k}^{-1}\left(P_{k|k}^{-1} + A^\top Q^{-1}A\right)^{-1}P_{k|k}^{-1}\right]A^{-1} \\ &\le A^{-\top}\left[P_{k|k}^{-1} - P_{k|k}^{-1}\left(\epsilon I + A^\top Q^{-1}A\right)^{-1}P_{k|k}^{-1}\right]A^{-1} \\ &= A^{-\top}\left[P_{k|k}^{-1} - P_{k|k}^{-1}\bar{Q}P_{k|k}^{-1}\right]A^{-1}, \end{aligned}$$

where $\bar{Q} = \left(\epsilon I + A^\top Q^{-1}A\right)^{-1}$. We also define

$$Z_k = \begin{bmatrix} T_{1,k} - \alpha P_{k|k-1}^{-1} & T_{2,k} - C^\top W & T_{3,k} \\ \star & T_{4,k} + 2W & T_{5,k} + WD \\ \star & \star & U \end{bmatrix},$$

where W is a diagonal positive definite matrix, U a positive definite matrix, $\alpha > 0$ a positive scalar, and

$$\begin{aligned} T_{1,k} &= C^\top R^{-1}C + P_{k|k}^{-1}\bar{Q}P_{k|k}^{-1} - P_{k|k}^{-1}\bar{Q}C^\top R^{-1}C \\ &\quad - C^\top R^{-1}C\bar{Q}P_{k|k}^{-1} - C^\top R^{-1}C\left(P_{k|k} - \bar{Q}\right) \\ &\quad \cdot C^\top R^{-1}C - C^\top \Gamma_2 C, \\ T_{2,k} &= -C^\top R^{-1} + P_{k|k}^{-1}\bar{Q}C^\top R^{-1} \\ &\quad + C^\top R^{-1}C\left(P_{k|k} - \bar{Q}\right)C^\top R^{-1}, \\ T_{3,k} &= \Big[-C^\top R^{-1} + P_{k|k}^{-1}\bar{Q}C^\top R^{-1} \\ &\quad + C^\top R^{-1}C\left(P_{k|k} - \bar{Q}\right)C^\top R^{-1} + C^\top \Gamma_2\Big]D, \\ T_{4,k} &= -R^{-1}C\left(P_{k|k} - \bar{Q}\right)C^\top R^{-1}, \\ T_{5,k} &= -\left[R^{-1}C\left(P_{k|k} - \bar{Q}\right)C^\top R^{-1}\right]D. \end{aligned}$$

In addition, we consider another matrix

$$T_{6,k} = D^\top\left[R^{-1}C\left(P_{k|k} - \bar{Q}\right)C^\top R^{-1} + \Gamma_2\right]D.$$

Here, the upper boundedness of $P_{k|k}$ implies that $T_{6,k}$ is also upper bounded, with the bound denoted as $\bar{T}_6$.

The following theorem shows the result about the stability of the prediction error dynamics.

Theorem 2. *Suppose $\|d_k\| \le \mu < \infty$, $\rho = e^{-1}\sum_i \gamma_{1,i} < \infty$, and that A is invertible. If there exist $P_{0|-1}$, W, U, α and Γ_2 such that $Z_k \ge 0$ and $0 < \alpha \le 1 - \max\{\ldots, \lambda_{1,i}, \lambda_{2,i}, \ldots\}$ for $i = 1, 2, \ldots, p$, then the prediction error e_k is upper bounded with*

$$\|e_k\| \le \sqrt{\frac{1}{c_2}\left[(1-\alpha)^k V_0 + \frac{1-(1-\alpha)^{k-1}}{\alpha}(c_1\mu^2 + \rho)\right]}, \tag{12a}$$

$$\lim_{k\to\infty}\|e_k\| \le \sqrt{\frac{c_1\mu^2 + \rho}{\alpha c_3}}, \tag{12b}$$

where $c_1 = \bar{\lambda}(\bar{T}_6 + U)$, $c_2 = \underline{\lambda}(P_{k|k-1}^{-1})$, and $c_3 = \underline{\lambda}(P_\infty^{-1})$. Further, if $0 < \alpha \le 1 - \max\{\ldots, \lambda_{1,i}, \lambda_{2,i} + \gamma_{1,i}, \ldots\}$ for $i = 1, 2, \ldots, p$, then

$$\|e_k\| \le \sqrt{\frac{1}{c_2}\left[(1-\alpha)^k V_0 + \frac{1-(1-\alpha)^{k-1}}{\alpha}c_1\mu^2\right]}, \tag{13a}$$

$$\lim_{k\to\infty}\|e_k\| \le \sqrt{\frac{c_1}{\alpha c_3}}\mu. \tag{13b}$$

Proof: Consider a Lyapunov function

$$V_k = e_k^\top P_{k|k-1}^{-1}e_k + \textstyle\sum_i \sigma_{i,k} + \sum_i \varepsilon_{i,k}.$$

Using (11), we have

$$\begin{aligned} &V_{k+1} = e_{k+1}^\top P_{k+1|k}^{-1}e_{k+1} + \textstyle\sum_i \sigma_{i,k+1} + \sum_i \varepsilon_{i,k+1} \\ &\le \left[Ae_k - AK_k \cdot \mathrm{sat}_\sigma(CAe_k - Dd_k)\right]^\top \\ &\quad \cdot A^{-\top}\left[P_{k|k}^{-1} - P_{k|k}^{-1}\bar{Q}P_{k|k}^{-1}\right]A^{-1} \\ &\quad \cdot \left[Ae_k - AK_k \cdot \mathrm{sat}_\sigma(Ce_k - Dd_k)\right] \\ &\quad + \left(Ce_k - Dd_k\right)^\top \Gamma_2\left(Ce_k - Dd_k\right) + \textstyle\sum_i \lambda_{1,i}\sigma_{i,k} \\ &\quad + \textstyle\sum_i \lambda_{2,i}\varepsilon_{i,k} + \sum_i \gamma_{1,i}\varepsilon_{i,k}e^{-\varepsilon_{i,k}} \\ &\le e_k^\top\left[P_{k|k}^{-1} - P_{k|k}^{-1}\bar{Q}P_{k|k}^{-1}\right]e_k - 2e_k^\top\left[P_{k|k}^{-1} - P_{k|k}^{-1}\bar{Q}P_{k|k}^{-1}\right] \\ &\quad \cdot K_k \cdot \mathrm{sat}_\sigma(Ce_k - Dd_k) + \mathrm{sat}_\sigma^\top(Ce_k - Dd_k) \cdot K_k^\top \\ &\quad \cdot \left[P_{k|k}^{-1} - P_{k|k}^{-1}\bar{Q}P_{k|k}^{-1}\right]K_k \cdot \mathrm{sat}_\sigma(Ce_k - Dd_k) \\ &\quad + \left(Ce_k - Dd_k\right)^\top \Gamma_2\left(Ce_k - Dd_k\right) + \textstyle\sum_i \lambda_{1,i}\sigma_{i,k} \\ &\quad + \textstyle\sum_i \lambda_{2,i}\varepsilon_{i,k} + \rho. \end{aligned}$$

Let us define

$$s_k = Ce_k - Dd_k - \mathrm{sat}_\sigma(Ce_k - Dd_k).$$

In addition, we have $P_{k|k}^{-1} = P_{k|k-1}^{-1} + C^\top R^{-1}C$, $P_{k|k}^{-1}K_k = C^\top R^{-1}$ and $K_k^\top P_{k|k}^{-1}K_k = R^{-1}CP_{k|k}C^\top R^{-1}$. These relations can be readily proven. It then follows that

$$\begin{aligned} V_{k+1} \le\ & e_k^\top P_{k|k-1}^{-1}e_k - e_k^\top T_{1,k}e_k - 2e_k^\top T_2 s_k - 2e_k^\top T_{3,k}d_k \\ & - s_k^\top T_{4,k}s_k - 2s_k^\top T_{5,k}d_k + d_k^\top T_{6,k}d_k + \textstyle\sum_i \lambda_{1,i}\sigma_{i,k} \\ & + \textstyle\sum_i \lambda_{2,i}\varepsilon_{i,k} + \rho. \end{aligned}$$

According to [29, Lemma 1.6], we have

$$-s_k^\top W(s_k - Ce_k + Dd_k) \ge 0.$$

It can be obtained that

$$\begin{aligned} V_{k+1} &\le V_{k+1} - 2s_k^\top W(s_k - Ce_k + Dd_k) \\ &\le e_k^\top P_{k|k-1}^{-1}e_k - e_k^\top T_{1,k}e_k - 2e_k^\top(T_2 - C^\top W)s_k \\ &\quad - 2e_k^\top T_{3,k}d_k - s_k^\top\left(T_{4,k} + 2W\right)s_k \end{aligned}$$

$$
\begin{aligned}
&- 2s_k^\top \left(T_{5,k} + WD\right) d_k + d_k^\top T_{6,k} d_k \\
&+ \textstyle\sum_i \lambda_{1,i}\sigma_{i,k} + \sum_i \lambda_{2,i}\varepsilon_{i,k} + \rho \\
= {}& e_k^\top P_{k|k-1}^{-1} e_k - \begin{bmatrix} e_k \\ s_k \\ d_k \end{bmatrix}^\top \\
&\cdot \begin{bmatrix} T_{1,k} & T_{2,k} - C^\top W & T_{3,k} \\ \star & T_{4,k} + 2W & T_{5,k} + WD \\ \star & \star & U \end{bmatrix} \begin{bmatrix} e_k \\ s_k \\ d_k \end{bmatrix} \\
&+ d_k^\top \left(T_{6,k} + U\right) d_k + \textstyle\sum_i \lambda_{1,i}\sigma_{i,k} + \sum_i \lambda_{2,i}\varepsilon_{i,k} + \rho.
\end{aligned}
$$

If $Z_k \geq 0$, then

$$
\begin{aligned}
V_{k+1} \leq {}& (1-\alpha) e_k^\top P_{k|k-1}^{-1} e_k + (1-\alpha) \textstyle\sum_i \sigma_{i,k} \\
&+ (1-\alpha) \textstyle\sum_i \varepsilon_{i,k} + d_k^\top \left(T_{6,k} + U\right) d_k \\
&+ \textstyle\sum_i (\lambda_{1,i} + \alpha - 1)\sigma_{i,k} + \sum_i (\lambda_{2,i} + \alpha - 1)\varepsilon_{i,k} + \rho.
\end{aligned}
$$

Because $0 < \alpha \leq 1 - \max\{\lambda_{1,i}, \lambda_{2,i}\}$ for $i = 1, 2, \ldots, p$,

$$
\begin{aligned}
V_{k+1} &\leq (1-\alpha) V_k + d_k^\top \left(T_{6,k} + U\right) d_k + \rho \\
&\leq (1-\alpha) V_k + \bar{\lambda} \left(\bar{T}_6 + U\right) \mu^2 + \rho,
\end{aligned}
$$

from which one can easily obtain (13a)-(13b). •

Remark 2. *For the discrete-time case, it is also recommended that $P_{k|k-1}$ should be initialized to be small enough and near zero, in order to make the conditions in Theorem 2 satisfied more easily.* •

IV. Simulation Example

In this section, we offer an application example to show the effectiveness of the proposed IS-EKF. The example is concerned with mobile robot localization, the purpose of which is to enable a mobile robot to determine its position using GPS and onboard sensing data. However, a challenge for precise localization is the outliers due to weak or abnormal GPS signals and sensors noises or temporary failures [30].

Consider a mobile wheeled robot with the following dynamic model [31], [32]:

$$
\begin{aligned}
p_{x,k+1} &= p_{x,k} + \eta_k T \cos(\theta_k), \\
p_{y,k+1} &= p_{y,k} + \eta_k T \sin(\theta_k), \\
\theta_{k+1} &= \theta_k + T\delta_k,
\end{aligned}
$$

where $p_{x,k}$ and $p_{y,k}$ are the coordinates of the robot's center of mass, θ_k the heading angle, η_k the robot's speed at the center of mass, δ_k the steering angle, and T the sampling period. Thus, the state vector of the robot is $\left[p_{x,k}\ p_{y,k}\ \theta_k\right]^\top$, and the control input vector is $\left[\eta_k\ \delta_k\right]^\top$. Here, η_k and δ_k can be read from onboard meters. In addition, we let x_k and y_k be obtained using GPS and θ_k be measured by a compass. The measurement model hence is a linear equation, but the dynamic model is nonlinear. Process and measurement noises are included into the model to represent uncertainties.

When the robot is moving and sampled every $T = 0.1$ s, the measurements of the x-coordinate and the steering angle are corrupted by outliers from time to time. The outlier disturbance d_k is designed as

$$
d_k = \begin{cases} \begin{bmatrix} 5 & 1 \end{bmatrix}^\top & 150 < k \leq 200, \\ 2\zeta_k & 350 < k \leq 400, \\ \begin{bmatrix} 100 & 50 \end{bmatrix}^\top & 450 < k \leq 500, \\ \begin{bmatrix} 100 & 0 \\ 0 & 50 \end{bmatrix} \zeta_k & 550 < k \leq 600, \end{cases}
$$

where $\zeta_k \in \mathbb{R}^2$ is a uniform random vector. This design takes account of different types of outliers in four stages — d_k is small at Stages 1-2 and large Stages 3-4, and it is constant at Stages 1 and 3 and random at Stages 2 and 4. In this setting, synthetic measurements are generated and shown in Fig. 1, in which the outlier-corrupted measurements are displayed in shaded areas.

It can be easily verified that the conventional EKF will completely fail when the above outliers are imposed on the measurements. One way to improve the robustness of the EKF is to use the 3σ rule [33]. That is, the innovation $(y_k - h(\hat{x}_k))$ is considered as normal if it lies within the $\pm 3\sigma$ bounds based on the covariance matrix $(H_k P_{k|k-1} H_k^\top + R)$, and as outlying otherwise. It is reset to be zero in the latter case so as not to distort the update procedure. We call this method as EKF with 3σ outlier rejection and use it as a benchmark to compare with the IS-EKF. In addition, to apply the IS-EKF in (9)-(10), the following parameters are set for (10):

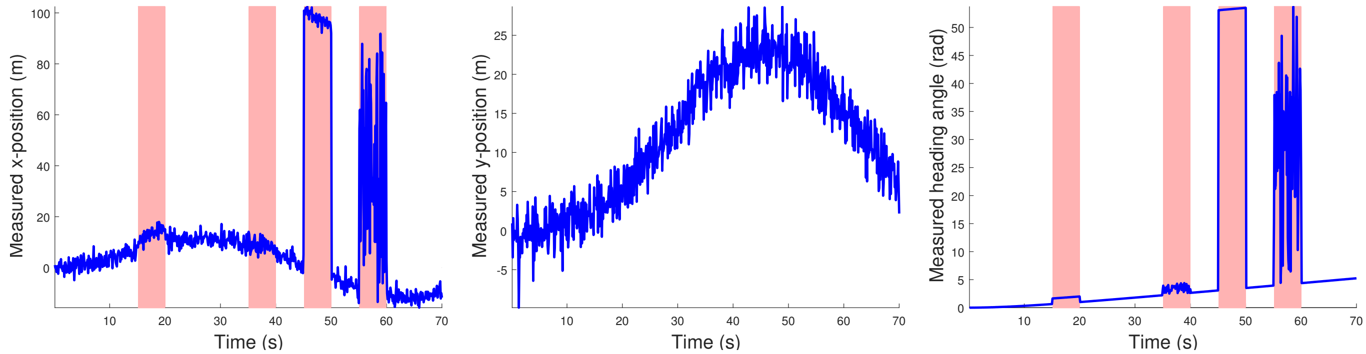
$$
\begin{aligned}
&\Lambda_1 = \mathrm{diag}(\begin{bmatrix} 0.5 & 0.5 & 0.1 \end{bmatrix}),\ \Lambda_2 = \mathrm{diag}(\begin{bmatrix} 0.1 & 0.1 & 0.1 \end{bmatrix}), \\
&\Gamma_1 = \mathrm{diag}(\begin{bmatrix} 10^2 & 10^2 & 5 \times 10^{-3} \end{bmatrix}),\ \Gamma_2 = \mathrm{diag}(\begin{bmatrix} 9 & 9 & 9 \end{bmatrix}).
\end{aligned}
$$

Figs. 2(a)-2(c) show the estimation of the x- and y-coordinates and the heading angle θ through time, respectively. It is seen that, with the 3σ outlier rejection, the EKF does not diverge seriously but still struggles with providing reliable state estimation. As a contrast, the IS-EKF demonstrates much better performance, maintaining a smooth and accurate estimation when the outliers appear. Further, Fig. 3(a) illustrates the estimated trajectories by the standard EKF, EKF with 3σ outlier rejection, and the IS-EKF, in comparison with the ground truth. The standard EKF completely fails to estimate the position due to the outliers in this case. The 3σ outlier rejection enhances its estimation considerably, but the IS-EKF reconstructs the trajectory at the best accuracy. These results reflect a considerable effectiveness of the IS-EKF in addressing the outliers..

Here, we provide some further remarks about the advantages and application of the IS-EKF.

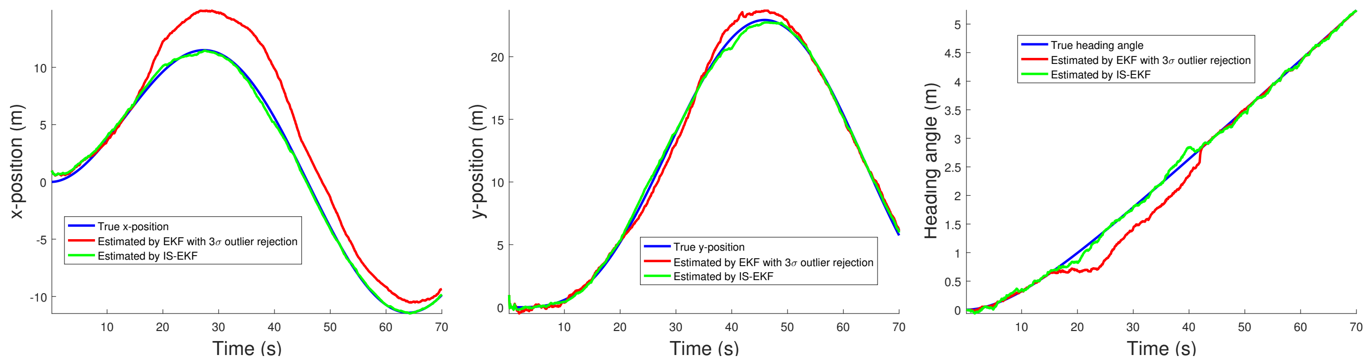
Remark 3. *In addition to the example presented above, we performed numerous simulations about robot localization in different settings (e.g., outliers, noises, initial estimation), and we found that the IS-EKF consistently offer satisfactory estimation. Two important observations include:*

- *The IS-EKF can well handle outliers that last for a relatively long period and vary in magnitude or type. This contrasts with many existing methods. For example, the stubborn observer in [26] can only reject occasional, singly outliers, and the EKF with 3σ outlier rejection*



(a) (b) (c)

Figure 1: Measurement profiles : (a) x-position; (b) y-position; (c) heading angle. The shaded areas represent the occurrence of outliers.



(a) (b) (c)

Figure 2: Estimation of the state variables: (a) estimation of the x-position; (b) estimation of the y-position; (c) estimation of the heading angle θ.

is less effective for dealing with small outliers that approximately align with the $\pm 3\sigma$ bounds. Note that the EKF may offer improved estimation by adjusting the outlier detection bounds from $\pm 3\sigma$ to $\pm \ell\sigma$ with $\ell \in \mathbb{R}^+$. However, for each selection of ℓ, it can still be futile for outliers that roughly lie within the bounds.

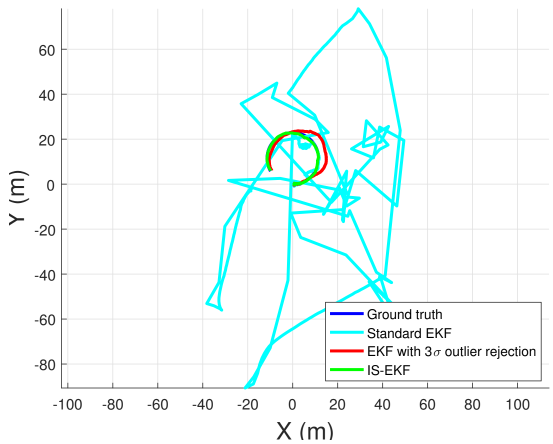
- *As an additional benefit, the IS-EKF demonstrates good robustness again initial state uncertainties. It is known that the EKF can easily fail if the initial guess is not accurate, although it is often difficult to obtain a precise guess in practice. However, the innovation saturation can often check the divergence caused by a poor initial guess, reducing the possibility of a complete failure.* •

Remark 4. *Application of the IS-EKF requires the selection of a set of parameters for the innovation saturation procedure. We have the following suggestions for practitioners:*

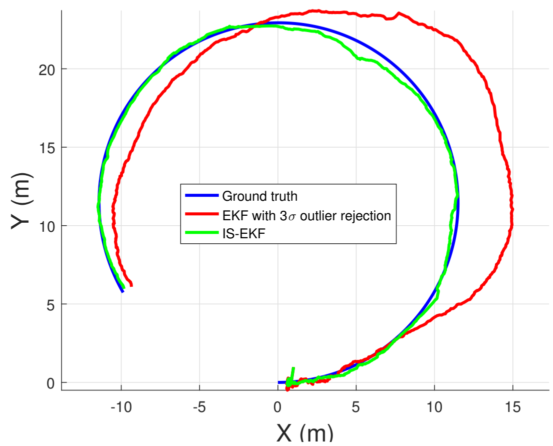
- *Let $\lambda_{1,i}, \lambda_{2,i} < 0$ for a continuous-time system and $0 < \lambda_{1,i}, \lambda_{2,i} < 1$ for a discrete-time system.*
- *It is useful to choose $\lambda_{2,i}$ such that the dynamics ε_i is fast in order to better track the change in innovation. This can be done by letting it take an appropriately small negative number in the continuous-time case or a number close to zero in the discrete-time case.*
- *Let $\gamma_{1,i}, \gamma_{2,i} > 0$ and $\gamma_{2,i} < 10$.*
- *The selection of $\gamma_{1,i}$ depends on the magnitude of the corresponding innovation process. Set $\gamma_{1,i}$ to be a large number if the innovation is usually large in the normal, i.e., outlier-free, case and a small number otherwise.* •

V. Conclusion

The EKF has gained wide application across different fields as a popular state estimation tool. However, its estimation accuracy can be severely hampered by measurement outliers due to sensor anomaly, model uncertainties, data transmission errors or cyber attacks. This paper presented a novel innovation saturation mechanism, IS-EKF, which is a robustified EKF, as an alternative to the conventional EKF. This mechanism saturates the innovation process, which is crucial for correcting the state estimation, thus ensuring a reasonable correction to be applied when outliers occur. We showed the IS-EKF architecture for both continuous- and discrete-time systems and provided theoretical analysis to derive useful stability properties of the proposed approaches for linear systems. We applied the discrete-time IS-EKF approach to mobile robot localization by simulation to illustrate its effectiveness. The numerical simulation results showed that the proposed approach brings about significant robustness for localization against GPS outliers. This indicated the viability of IS-EKF in effectively rejecting outliers of varying magnitude and



(a)



(b)

Figure 3: Estimated trajectories in comparison with the ground truth.

durations at a reasonable computational cost and without the need of measurement redundancy.

References

[1] H. Fang, N. Tian, Y. Wang, M. Zhou, and M. A. Haile, "Nonlinear bayesian estimation: from Kalman filtering to a broader horizon," *IEEE/CAA Journal of Automatica Sinica*, vol. 5, no. 2, pp. 401–417, 2018.

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

[3] Y. Zhang, N. Meratnia, and P. Havinga, "Outlier detection techniques for wireless sensor networks: A survey," *IEEE Communications Surveys Tutorials*, vol. 12, no. 2, pp. 159–170, 2010.

[4] M. A. Haile, J. C. Riddick, and A. H. Assefa, "Robust particle filters for fatigue crack growth estimation in rotorcraft structures," *IEEE Transactions on Reliability*, vol. 65, no. 3, pp. 1438–1448, 2016.

[5] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S. S. Sastry, "Kalman filtering with intermittent observations," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1453–1464, 2004.

[6] B. Jia and M. Xin, "Vision-based spacecraft relative navigation using sparse-grid quadrature filter," *IEEE Transactions on Control Systems Technology*, vol. 21, no. 5, pp. 1595–1606, 2013.

[7] Y. Li, L. Shi, and T. Chen, "Detection against linear deception attacks on multi-sensor remote state estimation," *IEEE Transactions on Control of Network Systems*, 2017, in press.

[8] C. Masreliez, "Approximate non-Gaussian filtering with linear state and observation relations," *IEEE Transactions on Automatic Control*, vol. 20, no. 1, pp. 107–110, 1975.

[9] H. Sorenson and D. Alspach, "Recursive bayesian estimation using Gaussian sums," *Automatica*, vol. 7, no. 4, pp. 465 – 479, 1971.

[10] R. J. Meinhold and N. D. Singpurwalla, "Robustification of Kalman filter models," *Journal of the American Statistical Association*, vol. 84, no. 406, pp. 479–486, 1989.

[11] S. Sarkka and A. Nummenmaa, "Recursive noise adaptive Kalman filtering by variational Bayesian approximations," *IEEE Transactions on Automatic Control*, vol. 54, no. 3, pp. 596–600, 2009.

[12] G. Agamennoni, J. I. Nieto, and E. M. Nebot, "An outlier-robust Kalman filter," in *IEEE International Conference on Robotics and Automation*, 2011, pp. 1551–1558.

[13] M. A. Gandhi and L. Mili, "Robust Kalman filter based on a generalized maximum-likelihood-type estimator," *IEEE Transactions on Signal Processing*, vol. 58, no. 5, pp. 2509–2520, 2010.

[14] D. De Palma and G. Indiveri, "Output outlier robust state estimation," *International Journal of Adaptive Control and Signal Processing*, vol. 31, no. 4, pp. 581–607, 2017.

[15] J. Keller and M. Darouach, "Optimal two-stage kalman filter in the presence of random bias," *Automatica*, vol. 33, no. 9, pp. 1745 – 1748, 1997.

[16] C.-S. Hsieh, "Robust two-stage Kalman filters for systems with unknown inputs," *IEEE Transactions on Automatic Control*, vol. 45, no. 12, pp. 2374–2378, 2000.

[17] S. Gillijns and B. D. Moor, "Unbiased minimum-variance input and state estimation for linear discrete-time systems with direct feedthrough," *Automatica*, vol. 43, no. 5, pp. 934 – 937, 2007.

[18] H. Fang, Y. Shi, and J. Yi, "On stable simultaneous input and state estimation for discrete-time linear systems," *International Journal of Adaptive Control and Signal Processing*, vol. 25, no. 8, pp. 671–686, 2011.

[19] S. Z. Yong, M. Zhu, and E. Frazzoli, "A unified filter for simultaneous input and state estimation of linear discrete-time stochastic systems," *Automatica*, vol. 63, pp. 321 – 329, 2016.

[20] D. Shi, T. Chen, and M. Darouach, "Event-based state estimation of linear dynamic systems with unknown exogenous inputs," *Automatica*, vol. 69, pp. 275 – 288, 2016.

[21] H. Fang, R. A. de Callafon, and J. Cortés, "Simultaneous input and state estimation for nonlinear systems with applications to flow field estimation," *Automatica*, vol. 49, no. 9, pp. 2805 – 2812, 2013.

[22] H.-Q. Mu and K.-V. Yuen, "Novel outlier-resistant extended Kalman filter for robust online structural identification," *Journal of Engineering Mechanics*, vol. 141, no. 1, p. 04014100, 2015.

[23] M. A. Gandhi, "Robust Kalman filters using generalized maximum likelihood-type estimators," Ph.D. dissertation, Virginia Polytechnic Institute and State University.

[24] D. Simon, *Optimal State Estimation: Kalman, $\mathcal{H}_\infty$, and Nonlinear Approaches*. Wiley-Interscience, 2006.

[25] X. Li, J. Lam, H. Gao, and J. Xiong, "$\mathcal{H}_\infty$ and $\mathcal{H}_2$ filtering for linear systems with uncertain markov transitions," *Automatica*, vol. 67, pp. 252 – 266, 2016.

[26] A. Alessandri and L. Zaccarian, "Stubborn state observers for linear time-invariant systems," *Automatica*, vol. 88, pp. 1 – 9, 2018.

[27] H. Fang, M. A. Haile, and Y. Wang, "Robustifying the Kalman filter against measurement outliers: An innovation saturation mechanism," in *Proceedings of IEEE Conference on Decision and Control*, 2018, pp. 6390–6395.

[28] V. Kucera, "A review of the matrix Riccati equation," *Kybernetika*, vol. 09, no. 1, pp. 42–61, 1973.

[29] S. Tarbouriech, G. Garcia, J. M. Gomes da Silva, and I. Queinnec, *Stability and Stabilization of Linear Systems with Saturating Actuators*. London: Springer London, 2011.

[30] P. Corke, *Robotics, Vision and Control*. Springer International Publishing.

[31] A. Sakai and Y. Kuroda, "Discriminative parameter training of unscented Kalman filter," *IFAC Proceedings Volumes*, vol. 43, no. 18, pp. 677 – 682, 2010, 5th IFAC Symposium on Mechatronic Systems.

[32] M. A. Ghadiri-Modarres, M. Mojiri, and H. R. Z. Zangeneh, "New schemes for GPS-denied source localization using a nonholonomic unicycle," *IEEE Transactions on Control Systems Technology*, vol. 25, no. 2, pp. 720–727, 2017.

[33] H. Liu, S. Shah, and W. Jiang, "On-line outlier detection and data cleaning," *Computers & Chemical Engineering*, vol. 28, no. 9, pp. 1635 – 1647, 2004.