Inverse Extended Kalman Filter
Himali Singh, Arpan Chattopadhyay and Kumar Vijay Mishra

Abstract—Recent advances in counter-adversarial systems have garnered significant research interest in inverse filtering from a Bayesian perspective. For example, interest in estimating the adversary’s Kalman filter tracked estimate with the purpose of predicting the adversary’s future steps has led to recent formulations of inverse Kalman filter (I-KF). In this context of inverse filtering, we address the key challenges of non-linear process dynamics and unknown input to the forward filter by proposing inverse extended Kalman filter (I-EKF). We derive I-EKF with and without an unknown input by considering non-linearity in both forward and inverse state-space models. In the process, I-KF-with-unknown-input is also obtained. We then provide theoretical stability guarantees using both bounded non-linearity and unknown matrix approaches. We further generalize these formulations and results to the case of higher-order, Gaussian-sum, and dithered I-EKFs. Numerical experiments validate our methods for various proposed inverse filters using the recursive Cramér-Rao lower bound as a benchmark.

Index Terms—Bayesian filtering, counter-adversarial systems, extended Kalman filter, inverse filtering, nonlinear processes.

I. INTRODUCTION

In many engineering applications, it is desired to infer the parameters of a filtering system by observing its output. This inverse filtering is useful in applications such as system identification, fault detection, image deblurring, and signal deconvolution [1][2]. Conventional inverse filtering is limited to non-dynamic systems. However, applications such as cognitive and counter-adversarial systems [3] have recently been shown to require designing the inverse of classical stochastic filters such as hidden Markov model (HMM) filter [4] and Kalman filter (KF) [5]. The cognitive systems are intelligent units that sense the environment, learn relevant information about it, and then adapt themselves in real-time to optimally enhance their performance. For example, a cognitive radar [6] adapts both transmitter and receiver processing in order to achieve desired goals such as improved target detection [7] and tracking [8]. In this context, [9] recently introduced inverse cognition, in the form of inverse stochastic filters, to detect cognitive sensor and further estimate the information that the same sensor may have learnt. In this paper, we focus on inverse stochastic filtering for such inverse cognition applications.

At the heart of inverse cognition are two agents: ‘us’ (e.g., an intelligent target) and an ‘adversary’ (e.g., a sensor or radar) equipped with a Bayesian tracker. The adversary infers an estimate of our kinematic state and cognitively adapts its actions based on this estimate. ‘We’ observe adversary’s actions with the goal to predict its future actions in a Bayesian sense. In particular, [10] developed stochastic revealed preferences-based algorithms to ascertain if the adversary’s actions are consistent with optimizing a utility function; and if so, estimate that function.

If the target aims to guard against the adversary’s future actions, it requires an estimate of the adversary’s inference. This is precisely the objective of inverse Bayesian filtering. In (forward) Bayesian filtering, given noisy observations, a posterior distribution of the underlying state is obtained. An example is the KF, which provides optimal estimates of the underlying state in linear system dynamics with Gaussian measurement and process noises. The inverse filtering problem, on the other hand, is concerned with estimating this posterior distribution of a Bayesian filter given the noisy measurements of the posterior. An example of such a system is the recently introduced inverse Kalman filter (I-KF) [9]. Note that, historically, the Wiener filter – a special case of KF when the process is stationary – has been used for frequency-domain inverse filtering for deblurring in image processing [11]. Further, some early works [12] have investigated the inverse problem of finding cost criterion for a control policy.

Although KF and its continuous-time variant Kalman-Bucy filter [13] are highly effective in many practical applications, they are optimal for only linear and Gaussian models. In practice, many engineering problems involve nonlinear processes [14][15]. In these cases, a linearized KF, which employs a linear system whose states represent the deviations from a nominal trajectory of a nonlinear system, is used. The KF estimates the deviations from the nominal trajectory and obtains an estimate of the states of the nonlinear system. The linearized KF is extended to directly estimate the states of a nonlinear system in the extended KF (EKF) [15]. The linearization is locally at the state estimates through Taylor series expansion. This is very similar to the Volterra series filters [17] that are nonlinear counterparts of adaptive linear filters.

While inverse nonlinear filters have been studied for adaptive systems in some previous works [18][19], the inverse of nonlinear stochastic filters such as EKF remain unexamined so far. To address the aforementioned nonlinear inverse cognition scenarios, contrary to prior works which focus on only linear I-KF [9], our goal is to derive and analyze inverse EKF (I-EKF). The standard EKF has its limitations. Unlike KF, the filter and (filter) gain equations of EKF are not decoupled and hence offline computations of these EKF quantities are not possible. The EKF performs poorly when the dynamic system is significantly nonlinear. It is also very sensitive to initialization because of Taylor approximation, and may even completely fail [20].

These EKF drawbacks have led to the development of variants such as higher-order EKFs [21], which include terms beyond the first-order in the Taylor series. These filters reduce the linearization errors that are inherent to the EKF and may provide an improved estimation at the cost of higher complexity and computations. The second-order EKF (SOEKF) [22][23] performs better than EKF but its stability is not guaranteed [24]. Therefore, in this paper, we derive both the stability conditions of I-EKF under various scenarios and the inverses of a few prominent EKF variants.

Preliminary results of this work appeared in our conference publication [25], where only I-EKF was formulated. In this paper, we extend our methods to several other inverse filtering systems and applications. Our main contributions are:

1) I-KF and I-EKF with unknown inputs. In the inverse cognition scenario, the target may introduce additional motion or jamming that is known to the target but not to the adversarial cognitive sensor. While deriving I-EKF, we therefore consider this more general...
nonlinear system model with unknown input; this I-EKF simplifies in the absence of unknown excitation. In the process, we also obtain I-KF-with-unknown-input that was not examined in the I-KF of [9].

2) Augmented states for I-EKF. For systems with unknown inputs, the adversary’s state estimate depends on its estimate of the unknown input. As a result, the adversary’s forward filters vary with system models. We overcome this challenge by considering augmented states in the inverse filter so that the unknown estimation is performed jointly with state estimation, including for KF with direct feed-through. For different inverse filters, separate augmented states are considered depending on the state transitions for the inverse filter.

3) Stability of I-EKF. In general, stability and convergence results for nonlinear KFs, and more so for their inverses, are difficult to obtain. In this work, we show the stability of I-EKF using two techniques. The first approach is based on bounded nonlinearities, which has been earlier employed for proving stochastic stability of discrete-time [26] and continuous-time [27] EKFs. Here, the estimation error was shown to be both exponentially bounded in the mean square sense and with probability one. The second method relaxes the bound on the initial estimation error by introducing unknown matrices to model the linearization errors [28]. Besides providing the sufficient conditions for error boundedness, this approach also rigorously justifies the enlarging of the noise covariance matrices to stabilize the filter [29]. Because of the dependence of the I-EKF’s error dynamics on the forward filter’s recursive updates, the derivations of these theoretical guarantees are not straightforward.

4) Extensions to inverses of EKF variants. We also extend the I-EKF theory and stability proof to SOEKF. The presence of second order terms poses additional challenges in deriving the theoretical guarantees for both forward and inverse SOEKF. In addition, we consider the inverse of Gaussian-sum EKF (GS-EKF) which comprises of several individual EKFs [30] [31]. Here, the a posteriori density function is approximated by a sum of Gaussian density functions each of which has its own separate EKF. In situations where the estimation error is small, the a posteriori density is approximated adequately by one Gaussian density, and the Gaussian sum filter reduces to the EKF; the same is true for their inverses. Finally, we also examine the inverse of dithered EKF (DEKF) [32], which was shown to perform better when nonlinearities are cone-bounded. This filter introduces dither signals prior to the nonlinearities to tighten the cone-bounds and hence, improve stability properties. We validate the estimation errors of all inverse EKFs through extensive numerical experiments with recursive Cramér-Rao lower bound (RCRLB) [33] as the performance metric.

The rest of the paper is organized as follows. In the next section, we provide the background of inverse cognition model. The inverse EKF with unknown input is then derived in the Section III for the case of the forward EKF with and without direct feed-through. Here, we also obtain the standard I-EKF in the absence of unknown input. Then, similar cases are considered for inverse KF with unknown input in Section IV. We then discuss the inverse of some EKF variants in Section V before deriving the stability conditions in Section VI. In Section VII, we corroborate our results with numerical experiments before concluding in Section VIII.

Throughout the paper, we reserve boldface lowercase and uppercase letters for vectors (column vectors) and matrices, respectively. The notation [ai] is used to denote the i-th component of vector a and [Ai,j] denotes the (i,j)-th component of matrix A. The transpose operation is denoted by (·)T. The norm denoted by ||·||2 represents the L2 norm for a vector. The notation Tr(A), rank(A), ||A||, and ||A||∞, respectively, denote the trace, rank, spectral norm and maximum row sum norm of A. For matrices A and B, the inequality A ≤ B means that B − A is a positive semidefinite (p.s.d.) matrix. For a function f : Rn → Rm, ∇f denotes the Rm×n Jacobian matrix. Similarly, for a function f : Rn → R, ∇f and ∇2f denote the gradient vector (Rn×1) and Hessian matrix (Rn×n), respectively. A ‘n×n’ identity matrix is denoted by In and a an × m all zero matrix is denoted by 0n×m. The notation {ai}1≤i≤l denotes a set of elements indexed by integer i. The notation x ∼ N(μ, Q) represents a random variable drawn from a normal distribution with mean μ and covariance matrix Q while x ∼ U[ui, ui] means a random variable drawn from the uniform distribution over [ui, ui].

II. DESIDERATA FOR INVERSE COGNITION

Consider a discrete-time stochastic dynamical system as ‘our’ state evolution process {xk}k≥0, where xk ∈ Rn×1 is the state at the k-th time instant. Our current state is known to us perfectly. The control input uk ∈ Rm×1 is known to us but not to the adversary. In a linear state-space model, we denote the state-transition and control input matrices by F ∈ Rn×n and B ∈ Rn×m, respectively. Our state evolves as

\[ x_{k+1} = Fx_k + Bu_k + w_k, \]

where \( w_k \sim N(0_{n×1}, Q) \) is the process noise with covariance matrix Q ∈ Rn×n. At the adversary, the observation and control input matrices are given by H ∈ Rp×n and D ∈ Rp×m, respectively.

The adversary makes a noisy observation \( y_k \in Rp^+ \) at time k as

\[ y_k = Hx_k + Du_k + v_k, \]

where \( v_k \sim N(0_{p×1}, R) \) is the adversary’s measurement noise with covariance matrix R ∈ Rp×p.

The adversary uses \( \{y_j\}_{1≤j≤k} \) to compute the estimate \( \hat{x}_k \) of our current state \( x_k \) using a (forward) stochastic filter. The adversary then uses this estimate to administer an action matrix G ∈ Rn×p on \( \hat{x}_k \). We make noisy observations of this action as

\[ a_k = G\hat{x}_k + \epsilon_k \in R^{n×1}, \]

where \( \epsilon_k \sim N(0_{n×1}, \Sigma_k) \) is our measurement noise with covariance matrix \( \Sigma_k \in R^{n×n} \) of \( \hat{x}_k \) at time k. Finally, ‘we’ use \( \{a_j, x_j, u_j\}_{1≤j≤k} \) to compute the estimate \( \hat{x}_k \) ∈ Rn×1 of \( x_k \) in the (inverse) stochastic filter, where \( \hat{x}_k \) is the associated estimation covariance matrix of \( \hat{x}_k \).

Define \( \hat{u}_k \) to be the estimate of \( u_k \) as computed in the adversary’s forward filter, while \( \hat{u}_k \) is an estimate of \( u_k \) as computed by ‘our’ inverse filter. The noise processes \( \{w_k\}_{k≥0}, \{v_k\}_{k≥1} \) and \( \{\epsilon_k\}_{k≥1} \) are mutually independent and i.i.d. across time. These noise distributions are known to ‘us’ as well as the adversary. When the unknown input is absent, either B = 0n×m or D = 0p×m or both vanish. Throughout the paper, we assume that both parties (adversary and ‘us’) have perfect knowledge of the system model and parameters.

When the system dynamics are nonlinear, then the matrix pairs \{F, B\}, \{H, D\}, and the matrix G are replaced by nonlinear functions \( f(·, ·), h(·, ·) \), and \( g(·, ·) \), respectively, as

\[ x_{k+1} = f(x_k, u_k) + w_k, \]

\[ y_k = h(x_k, u_k) + v_k, \]

\[ a_k = g(\hat{x}_k) + \epsilon_k. \]

This is a direct feed-through (DF) model, wherein \( y_k \) depends on the unknown input. Without DF, the adversary’s observations are

\[ y_k = h(x_k) + v_k. \]

We show in the following Section III the presence or absence of the unknown input leads to different solution approaches towards forward and inverse filters. For simplicity, the presence of known exogenous inputs is also ignored in state evolution and observations.
However, it is trivial to extend the inverse filters developed in this paper for these modifications in the system model. Throughout the paper, we focus on discrete-time models.

III. I-EKF with Unknown Input

In order to extend KF to the nonlinear processes, EKF linearizes the model about the nominal values of the state vector and control input. The EKF is similar to the iterated least squares (ILS) method except that the former is for dynamical systems and the latter is not [34]. Note that the optimal forward EKFs with and without DF are conceptually different. In the latter case, while the observation $y_k$ is unaffected by the unknown input $u_k$, it is still dependent on $u_{k-1}$ through $x_k$; this induces a one-step delay in the adversary’s estimate of $u_k$. On the other hand, with DF, there is no such delay in estimating $u_k$. We now show that this difference results in different inverse filters for these two cases.

A. I-EKF-without-DF unknown input

Consider the non-linear system without DF given by [4] and [7]. Linearize the model functions as $F_k = \nabla_x f(x, u_{k-1})|_{x=x_k, u=u_{k-1}}$, $B_k = \nabla_u f(x_k, u)|_{u=u_{k-1}}$ and $H_k = \nabla_x h(x)|_{x=x_{k+1}}$.

1) Forward filter: The forward filter’s recursive state estimation procedure first obtains the prediction $\hat{x}_{k+1|k}$ of the current state using the previous state and input estimates, with $\Sigma_{k+1|k}$ as the associated state prediction error covariance matrix of $\hat{x}_{k+1|k}$. Then, the state and input gain matrices $K_{k+1}^F$ and $K_{k+1}^B$, respectively, are computed along with the input estimation (delay) covariance matrix $\Sigma_{k+1|k}$. Finally, the state $\hat{x}_{k+1}$, input $\hat{u}_k$, and covariance matrix $\Sigma_{k+1|k}$ are updated using current observation $y_{k+1}$, and gain matrices $K_{k+1}^F$ and $K_{k+1}^B$. Note that the current observation $y_{k+1}$ provides an estimate $\hat{u}_k$ of the input $u_k$ at the previous step time. The recursive equations of the adversary’s forward EKF are [35]:

**Prediction:** $\hat{x}_{k+1|k} = f(\hat{x}_k, \hat{u}_{k-1})$.

**Gain computation:** $\Sigma_{k+1|k}^F = F_k \Sigma_{k+1|k}^F F_k^T + Q_k$.

$K_{k+1}^F = \Sigma_{k+1|k}^F H_k^T (H_k \Sigma_{k+1|k}^F H_k^T + R_k)^{-1}$.

$\Sigma_{k+1}^u = (B_k^T H_k^T + R_k^{-1} (I_{p \times p} - H_k K_{k+1}^F H_k^T) H_k \Sigma_{k+1|k}^F B_k^T)^{-1}$.

**Update:** $\hat{x}_{k+1} = \hat{x}_{k+1|k} + K_{k+1}^F (y_{k+1} - h(\hat{x}_{k+1|k}))$.

$\hat{u}_k = K_{k+1}^B (y_{k+1} - h(\hat{x}_{k+1|k}) + H_k \hat{u}_{k-1})$.

**Covariance matrix update:** $\Sigma_{k+1}^u = (I_{n \times n} - K_{k+1}^F H_k \Sigma_{k+1}^F (I_{n \times n} - K_{k+1}^F H_k)^T)^{-1}$.

Forward filter exists if rank($\Sigma_{k+1}^u$) = $m$, for all $k \geq 0$, and $p \geq m$ [35].

2) Inverse filter: Consider a augmented state vector $z_k = [\hat{x}_k^T \hat{u}_{k-2}^T]^T$. ‘Our’ observation $z_k$ in (6) is the first observation that contains the information about unknown input estimate $\hat{u}_{k-2}$, because of the delay in forward filter input. Hence, the delayed estimate $\hat{u}_{k-2}$ is considered in the augmented state $z_k$.

Define $\bar{h}_k(\hat{u}_{k-2}, \hat{x}_{k-1}, x_k, v_k) = h_k(\hat{u}_{k-2}, \hat{x}_{k-1}, x_k, v_k)$, where the state transition equations of augmented state vector are $\hat{x}_{k+1} = \bar{f}_k(\hat{x}_k, \hat{u}_{k-2}, \hat{x}_{k-1}, x_k, x_{k+1}, v_k, v_{k+1})$ and $\hat{u}_{k-1} = \bar{h}_k(\hat{u}_{k-2}, \hat{x}_{k-1}, x_k, v_k)$. The state transition equations of augmented state vector are

\[
\begin{align*}
\bar{h}_k(\hat{u}_{k-2}, \hat{x}_{k-1}, x_k, v_k) &= K_{k-1}^u (H_k B_{k-1} u_{k-2} - h(f(\hat{x}_{k-1}, \hat{u}_{k-2})) + h(x_k) + v_k), \\
\bar{f}_k(\hat{x}_k, \hat{u}_{k-2}, \hat{x}_{k-1}, x_k, x_{k+1}, v_k, v_{k+1}) &= \phi_k(\hat{x}_k, h_k(\hat{u}_{k-2}, \hat{x}_{k-1}, x_k, v_k), x_{k+1}, v_{k+1}).
\end{align*}
\]

In these state transition equations, ‘our’ actual states $x_k$ and $x_{k+1}$ are perfectly known to us and henceforth treated as known exogenous inputs. Note that, unlike the forward filter, the process noise terms $v_k$ and $v_{k+1}$ are non-additive because the filter gains $K_{k+1}^F$ and $K_{k+1}^B$ depend on the previous estimates (through the Jacobians).

Denote $\tilde{z}_{k+1} = [\hat{z}_{k+1}^T \hat{u}_{k-1}^T]^T$. The state transition of the augmented state $\tilde{z}_{k+1}$ depends on the estimate $\hat{x}_{k-1}$ which we approximate by our previous estimate $\hat{x}_{k-1}$. With this approximation, $\hat{x}_{k-1}$ is treated as a known exogenous input for the inverse filter while the augmented process noise vector is $[v_k^T v_{k+1}^T]^T$.

Define the Jacobians $\bar{F}_k = \left[ \begin{array}{c} \nabla_{\hat{z}_{k+1}^T} f_k \\ \nabla_{\hat{v}_{k-2}^T} \bar{f}_k \\ \nabla_{\hat{v}_{k-1}^T} \bar{f}_k \\ \nabla_{\hat{z}_{k+1}^T} \bar{f}_k \end{array} \right]_{0_m \times n \times n \times n}$, and $G_{k+1} = \left[ \begin{array}{c} \nabla_{\hat{v}_{k-2}^T} \bar{f}_k \\ \nabla_{\hat{v}_{k-1}^T} \bar{f}_k \end{array} \right]_{0_m \times n \times n \times n}$ with respect to the augmented state; Jacobian

\[
\bar{F}_k = \left[ \begin{array}{c} \nabla_{\hat{z}_{k+1}^T} f_k \\ \nabla_{\hat{v}_{k-2}^T} \bar{f}_k \\ \nabla_{\hat{v}_{k-1}^T} \bar{f}_k \end{array} \right]_{0_m \times n \times n \times n}
\]

and $Q_{k+1} = \left[ \begin{array}{c} \nabla_{\hat{v}_{k-2}^T} \bar{f}_k \\ \nabla_{\hat{v}_{k-1}^T} \bar{f}_k \end{array} \right]_{0_m \times n \times n \times n}$.

Then, the recursive form of the I-EKF-without-DF yields the estimate $\hat{z}_k$ of the augmented state and the associated covariance matrix $\Sigma_k$ as:

Prediction: $\hat{x}_{k+1|k} = f(\hat{x}_k, \hat{u}_{k-2}, \hat{x}_{k-1}, x_k, x_{k+1}, 0_p \times 1_p, 0_p \times 1_p), \hat{x}_{k-1|k} = \hat{h}_k(\hat{u}_{k-2}, \hat{x}_{k-1}, x_k, 0_p \times 1_p), \hat{u}_{k-1} = K_{k+1}^F(\hat{y}_{k+1} - h(\hat{x}_{k+1|k}))$.

Update: $\hat{x}_{k+1} = G_{k+1} \hat{y}_{k+1} + \hat{x}_{k+1|k}$, $\hat{u}_{k} = K_{k+1}^B(\hat{y}_{k+1} - h(\hat{x}_{k+1|k}))$.

The recursions of I-EKF-without-DF take the same form as that of the standard EKF [36] but with modified system matrices. In particular, the former employs an augmented state such that the Jacobian of the state transition function with respect to the state is computed as $F_k$ while for the latter, it is simply $F_k = \nabla_x f(x_k, u_k)$.

B. I-EKF-with-DF unknown input

Consider the non-linear system with DF given by [4] and [5]. Linearize the functions as $F_k = \nabla_x f(x, u_k)|_{x=x_k}$, $H_{k+1} = \nabla_x h(x, u_k)|_{x=x_k, u=u_k}$.
1) Forward filter: denote the state and input estimation covariance and gain matrices to Section [H]. Here, the current observation \( y_{k+1} \) depends on the current unknown input \( u_{k+1} \) such that the forward filter infers \( u_{k+1} \) without any delay. For input estimation covariance without delay, we use \( \Sigma_{u}^u \). Then, the recursive form of forward EKF-with-DF is [37]

**Prediction:** \( x_{k+1|k} = f(x_k, u_k), \Sigma_{k+1|k} = F_k \Sigma_{u}^u F_k^T + Q 
\) (17)

**Gain computation:** \( K_{k+1} = \Sigma_{k+1|k} H_k^T \left( H_k \Sigma_{k+1|k} H_k^T + R \right)^{-1} \)

\( x_{k+1} = x_{k+1|k} + K_{k+1} \left( z_{k+1} - h(x_{k+1|k}) \right) \)

\( \Sigma_{k+1} = \left( I - K_{k+1} H_k \right) \Sigma_{k+1|k} \)

**Covariance matrix update:** \( \Sigma_{u}^u \) becomes \( \Sigma_{u}^u \) in place of the augmented noise vector of I-EKF-without-DF.

Here, unlike I-EKF-without-DF, the inverse filter’s prediction and update also result in a simplified process noise term \( \Sigma_{u}^u \) and \( \Sigma_{u}^u \) in forward EKF-without-DF yields forward EKF-without-unknown-input whose state prediction and updates are

\( \tilde{x}_{k+1} = f(\tilde{x}_k, x_{k+1|k}, u_{k+1}) \)

\( \tilde{x}_{k+1} = \tilde{x}_{k+1|k} + K_{k+1} (z_{k+1} - h(\tilde{x}_{k+1|k})) + \tilde{v}_{k+1} \) (22)

Denote \( \tilde{F}_k = H_k \Sigma_{k+1|k} H_k^T + R \) and gain \( \tilde{K}_{k+1} \) to replace with \( \Sigma_{k+1|k} \) and \( K_{k+1} \) respectively. (because only the state estimation covariances and gains are computed here). Hence, the I-EKF-without-DF’s state transition equations and recursions yield I-EKF-without-unknown-input. Dropping the input estimate term in the augmented state \( x_k \), the state transition equations become

\[ \tilde{x}_{k+1} = f(\tilde{x}_k, x_{k+1|k}, u_{k+1}) \]

\[ \tilde{x}_{k+1} = \tilde{x}_{k+1|k} + K_{k+1} (z_{k+1} - h(\tilde{x}_{k+1|k})) + \tilde{v}_{k+1} \] (23)

2) Inverse filter: Consider an augmented state vector \( x_k = \left[ x_k^T \ u_k^T \right]^T \) (note the absence of delay in the input estimate). Define

\[ \phi_k(\tilde{x}_k, \tilde{u}_k, x_{k+1|k}, u_{k+1}, v_{k+1}) = \]

\[ f(\tilde{x}_k, \tilde{u}_k) - K_{k+1} h(f(\tilde{x}_k, \tilde{u}_k), \tilde{u}_k) - K_{k+1} D_u u_{k+1} + K_{k+1} h(x_{k+1|k}, u_{k+1}) + K_{k+1} v_{k+1} \]

From [8] and [17], state transitions for inverse filter are

\[ \tilde{x}_{k+1} = \tilde{f}_k(\tilde{x}_k, \tilde{u}_k, x_{k+1|k}, u_{k+1}, v_{k+1}) \]

\[ \tilde{x}_{k+1} = \tilde{f}_k(\tilde{x}_k, \tilde{u}_k, x_{k+1|k}, u_{k+1}, v_{k+1}) \]

\[ \tilde{x}_{k+1} = x_{k+1|k} \]

\[ \tilde{x}_{k+1} = \tilde{f}_k(\tilde{x}_k, \tilde{u}_k, x_{k+1|k}, u_{k+1}, v_{k+1}) \]

\[ \tilde{x}_{k+1} = \tilde{x}_{k+1|k} \]

Examples of EKF with unknown inputs include fault detection and diagnosis, systems requiring stochastic state estimation or a simplified process noise term. The absence of delay in input estimation also results in a simplified process noise term \( v_{k+1} \), in place of the augmented noise vector of I-EKF-without-DF.

C. I-EKF without any unknown inputs

Consider a non-linear system model without unknown inputs in the system equations [4] and [7], i.e.,

\[ x_{k+1} = f(x_k) + w_k \] (20)

Linearize the functions as \( F_k = \nabla_x f(x)|_{x=x_k} \) and \( H_{k+1} = \nabla_z h(x)|_{x=x_{k+1|k}} \). Then, ceteris paribus, setting \( B_k = 0 \) and neglecting computation of \( \Sigma_{u}^u \) and \( \Sigma_{u}^u \) in forward EKF-without-DF yields forward EKF-without-unknown-input whose state prediction and updates are

\[ \tilde{x}_{k+1} = f(\tilde{x}_k) \]

\[ \tilde{x}_{k+1} = \tilde{x}_{k+1|k} + K_{k+1} (z_{k+1} - h(\tilde{x}_{k+1|k})) \]

IV. Inverse KF with unknown input

For linear Gaussian state-space models, our methods developed in the previous section are useful in extending the I-KF mentioned in [9].
to unknown input. Again, the forward KFs employed by the adversary with and without DF are conceptually different [39] because of the delay involved in input estimation. The forward KFs with unknown input provide unbiased minimum variance state and input estimates.

A. I-KF-without-DF

Consider the system in (1) and (2) with D = 0_{p×m}.

1) Forward filter: Unlike EKF-without-DF, the forward KF-without-DF considers an intermediate state update step using the estimated unknown input before the final state updates. In this step, the unknown input is first estimated (with one-step delay) using the current observation y_{k+1} and input estimation gain matrix M_{k+1}. In the update step, the current state estimate \( \hat{x}_{k+1} \) is computed again considering the current observation y_{k+1}. The forward KF-without-DF is [40]:

**Prediction:**
\[
\hat{x}_{k+1|k} = F \hat{x}_k + B u_k, \quad \Sigma_{k+1|k} = F \Sigma_k F^T + Q, \quad (27)
\]

**Unknown input estimation:**
\[
\hat{S}_{k+1} = H \Sigma_{k+1|k} H^T + R, \quad (28)
\]
\[
M_{k+1} = (B^T H^T S_{k+1}^{-1} H) B^{-1} B^T H^T S_{k+1}^{-1}, \quad (29)
\]
\[
\hat{u}_k = M_{k+1}(y_{k+1} - H \hat{x}_{k+1|k}), \quad (30)
\]
\[
\hat{x}_{k+1|k+1} = (I_n - B M_{k+1})(I_n - B M_{k+1})^T \Sigma_{k+1|k}^{-1} \quad (31)
\]
\[
\Sigma_{k+1|k+1} = (I_n - B M_{k+1}) H \Sigma_{k+1|k} (I_n - B M_{k+1})^T + B M_{k+1} R B^T, \quad (32)
\]

**Update:**
\[
\hat{x}_{k+1} = \hat{x}_{k+1|k+1} + K_{k+1}(y_{k+1} - H \hat{x}_{k+1|k+1} - \hat{D} u_k), \quad (33)
\]
\[
\Sigma_{k+1} = \Sigma_{k+1|k+1} - K_{k+1} H \Sigma_{k+1|k+1} - D \Sigma_{k+1|k+1} D^T K_{k+1}^T, \quad (34)
\]

The forward filter exists if rank(\(HH^T\)) = rank(\(B\)) = m which implies \(n \geq m\) and \(p \geq m\) such that \(\hat{u}_k\) is an unbiased minimum variance estimate of the unknown input [40]. Here, unlike I-EKFs, the gain matrices \(K_{k+1}\) and \(M_{k+1}\) are deterministic and completely determined by the model parameters and the initial covariance matrix similar to I-KF [9].

2) Inverse filter: Denote \(\bar{F}_k = (I_n - K_{k+1} H) (I_n - B M_{k+1} H) F\) and \(\bar{E}_k = B M_{k+1} R M_{k+1}^T B^T\). From (2) with \(D = 0_{n \times m}\), and (27)-(34), the state transition equation for I-KF-without-DF is

\[
\bar{F}_k \hat{x}_k + E_k H \hat{x}_k + E_k v_{k+1}. \quad (36)
\]

B. I-KF-with-DF

Consider the linear system model with DF given by (1) and (2).

1) Forward filter: Denote the state estimation covariance, input estimation (without delay) covariance, and cross-covariance of state and input estimates by \(\Sigma_{x}^x, \Sigma_{u}^u\), and \(\Sigma_{xu}^x\), respectively. The forward KF-with-DF is [39]:

**Prediction:**
\[
\hat{x}_{k+1|k} = F \hat{x}_k + B u_k, \quad \Sigma_{k+1|k} = \begin{bmatrix} \Sigma_{x}^x & \Sigma_{xu}^x \\ \Sigma_{u}^u & \Sigma_{uu} \end{bmatrix} = F \Sigma_k F^T + Q, \quad (41)
\]

**Gain computation:**
\[
M_{k+1} = (D^T S_{k+1}^{-1} D)^{-1} D^T S_{k+1}^{-1}, \quad (42)
\]
\[
K_{k+1} = \Sigma_{x+u}^x M_{k+1} H^T S_{k+1}^{-1}, \quad (43)
\]
\[
\hat{x}_{k+1} = \hat{x}_{k+1|k} + K_{k+1}(y_{k+1} - H \hat{x}_{k+1|k} - \hat{D} u_k), \quad (44)
\]
\[
\Sigma_{k+1} = \Sigma_{k+1|k} - K_{k+1} (S_{k+1} - D \Sigma_{k+1} D^T) K_{k+1}^T, \quad (45)
\]

The forward filter exists if rank(D) = m (which implies \(p \geq m\)). Again, the gain matrices \(K_{k+1}\) and \(M_{k+1}\) are deterministic and completely determined by the system model and initial covariance matrices estimates.

2) Inverse filter: Consider an augmented state vector \(z_k = [x_k^T \ u_k^T]^T\). Denote \(\bar{F}_k = (I_n - K_{k+1} H + K_{k+1} D M_{k+1} H) F\) and \(\bar{E}_k = K_{k+1} (I_{n-p} - D M_{k+1}) H\) and \(\bar{D}_k = -M_{k+1} H D\). From (2) and (41) [41], the state transition equations for IKF-with-DF are

\[
\bar{F}_k \hat{x}_k + \bar{B} \hat{u}_k + E_k H \hat{x}_k + E_k v_{k+1} \quad (37)
\]

Also, \([E_k v_{k+1}^T]^T (M_{k+1} v_{k+1} + x_{k+1}^T)\) \(| M_{k+1} v_{k+1} + x_{k+1}^T |\) is the augmented noise vector involved in this state transition with noise covariance matrix \(Q_k = \begin{bmatrix} E_k R E_k^T & E_k R M_{k+1}^T \\ M_{k+1} E_k R E_k^T & M_{k+1} E_k R M_{k+1}^T \end{bmatrix}\). Then, \(ceteris paribus\), following similar steps as in I-KF-without-DF, the I-KF-with-DF computes the estimate \(\hat{z}_k = [\hat{x}_k^T \ \hat{u}_k^T]^T\) of the augmented state vector using the observation \(u_k\) given by (3). The system matrices for the augmented state are \(\bar{F}_k = \bar{F}_k B_k H_k D_k\) and \(\bar{G} = [G \ 0_{n \times m}]\). The I-KF-with-DF predicts the augmented state as

\[
\hat{x}_{k+1|k} = \bar{F}_k \hat{x}_k + \bar{B} \hat{u}_k + E_k H \hat{x}_k + E_k v_{k+1}, \quad (38)
\]
\[
\hat{u}_{k+1} = \bar{D}_k \hat{x}_k + \bar{B} \hat{u}_k + M_{k+1} H \hat{x}_{k+1|k} + M_{k+1} D u_{k+1}, \quad (39)
\]
\[
\hat{z}_{k+1|k} = [\hat{x}_{k+1|k}^T \ \hat{u}_{k+1|k}^T]^T = \Sigma_{k+1|k}^{-1} = \bar{F}_k \Sigma_k \bar{F}_k^T + \bar{Q}_k, \quad (40)
\]

followed by the update procedure (38)-(40) with \(G\) and \(\hat{x}_{k+1}\) replaced by \(\bar{G}\) and \(\hat{z}_{k+1}\), respectively.

Since the observation \(y_k\) explicitly depends on the unknown input \(u_k\) for a system with DF, I-KF-with-DF and I-EKF-with-DF require perfect knowledge of the current input \(u_k\) as a known exogenous input to obtain their state and input estimates, which is not the case in I-KF-without-DF and I-EKF-without-DF.

V. EXTENSIONS TO INVERSES OF EKF VARIANTS

Several advanced versions of EKF exist, each aiming toward improving either the estimation accuracy (SOEKF or higher-order, GS-EKF), convergence (iterated EKF), stability (DEKF) or practical feasibility (hybrid EKF). In applications such as parameter estimation, EKF may not necessarily converge to true parameter values [41] unless some corrective terms are added. The iterated EKF, in which the correction equation is iterated a fixed number of times, was proposed to improve the convergence properties of the EKF [42]. The
We also derive the one-step prediction formulation of SOEKF and includes Hessian terms in the prediction equations.

$$\Sigma_{k+1|k} = F_k \Sigma_k F_k^T + Q$$

$$\begin{align*}
\Sigma_{k+1|k} = f(\hat{x}_{k+1|k}) + \frac{1}{2} \sum_{i=1}^{n} a_i \text{Tr} \left( \nabla^2 [f(\hat{x}_k)] \right) \Sigma_k \nabla^2 [f(\hat{x}_k)]_j \Sigma_k ) ,
\end{align*}$$

**Update:**

$$\hat{y}_{k+1|k} = h(\hat{x}_{k+1|k}) + \frac{1}{2} \sum_{i=1}^{p} b_i \text{Tr} \left( \nabla^2 [h(\hat{x}_{k+1|k})] \right) \Sigma_k \nabla^2 [h(\hat{x}_k)]_j \Sigma_k ) ,$$

\begin{align*}
\Sigma_{k+1|k} &= H_{k+1|k} \Sigma_{k+1|k} H_{k+1|k}^T + R,
\end{align*}

where $\hat{y}_{k+1|k}$ is the predicted observation. Unlike EKF, the SOEKF includes Hessian terms in the prediction equations.

**2) Inverse filter:** Define $f(\hat{x}, \Sigma) = f(\hat{x}) + \frac{1}{2} \sum_{i=1}^{n} a_i \text{Tr} \left( \nabla^2 [f(\hat{x})] \right) \Sigma_k \nabla^2 [f(\hat{x})]_j \Sigma_k )$ and $h(\hat{x}, \Sigma) = h(\hat{x}) + \frac{1}{2} \sum_{i=1}^{p} b_i \text{Tr} \left( \nabla^2 [h(\hat{x})] \right) \Sigma_k \nabla^2 [h(\hat{x})]_j \Sigma_k )$. Using (7) and (44)-(46), the state transition equation for the inverse SOEKF (I-SOEKF) is

$$\begin{align*}
\hat{x}_{k+1} &= \tilde{T}_h (\hat{x}_k, \Sigma_k, v_k) \\
&= f(\hat{x}_k, \Sigma_k) - K_{k+1} h(\hat{x}_k, \Sigma_k),
\end{align*}$$

where $K_{k+1} = \Sigma_{k+1|k} H_{k+1|k}^T \Sigma_k^{-1}$ is SOEKF’s gain matrix.

The I-SOEKF recursions now follow directly from forward SOEKF’s recursions treating $\Sigma_k$ as the observation. Denote the Jacobians and the Hessian with respect to the state estimate $\hat{x}_k$ as $F_k = \nabla_x f(\hat{x}_k, x, \Sigma)$ and $\Sigma_{k+1} \preceq \Sigma_k$. The process noise covariance matrix $Q_k$ is defined by $Q_k = \text{Tr}(\nabla^2 f(\hat{x}_k, x, \Sigma))$. The I-SOEKF’s prediction is

$$\begin{align*}
\hat{x}_{k+1|k} &= \tilde{T}_f (\hat{x}_k, \Sigma_k, v_{k+1}) \\
&= f(\hat{x}_k, \Sigma_k) + K_{k+1} h(\hat{x}_k, \Sigma_k),
\end{align*}$$

Here, besides the gain matrix $K_{k+1}$, the Hessian terms are also treated as time-varying parameters of the state transition and approximated by evaluating them using the previous I-SOEKF’s estimate $\hat{x}_k$. When the second-order terms are neglected, the forward SOEKF and I-SOEKF reduce to, respectively, forward EKF and I-EKF.

**3) One-step prediction formulation:** The one-step prediction formulation of SOEKF for the considered system model can be derived following the derivation of two-step recursion formulation of SOEKF as outlined in [46]. Denote $F_k = \nabla x f(\hat{x})_{|x=x_k}$ and $H_k = \nabla x h(\hat{x})_{|x=x_k}$, the recursive equations obtained are

$$\begin{align*}
\hat{x}_k &= f(\hat{x}_k) + \frac{1}{2} \sum_{i=1}^{n} a_i \text{Tr} \left( \nabla^2 [f(\hat{x}_k)] \right) \Sigma_k ) ,
\end{align*}$$

$$\begin{align*}
\hat{y}_k &= h(\hat{x}_k) + \frac{1}{2} \sum_{i=1}^{p} b_i \text{Tr} \left( \nabla^2 [h(\hat{x}_k)] \right) \Sigma_k \nabla^2 [h(\hat{x}_k)]_j \Sigma_k ) ,
\end{align*}$$

$$\begin{align*}
\Sigma_k &= F_k \Sigma_k F_k^T + Q,
\end{align*}$$

$$\begin{align*}
\Sigma_k &= H_k \Sigma_k H_k^T + R,
\end{align*}$$

$$\begin{align*}
\Sigma_{k+1} &= \Sigma_k - K_k \Sigma_k h_k^T,
\end{align*}$$

Again, the inverse filter follows directly from SOEKF’s one-step prediction formulation considering this state transition. The Jacobians with respect to the state estimate are $F_k = \nabla_x f(\hat{x}, x, \Sigma)_{|x=x_k}$ and $G_k = \nabla_x h(\hat{x})_{|x=x_k}$, for the process noise covariance matrix $Q_k = K_k R_k^T$ with the noise covariance matrix $Q_k$.

**B. Inverse GS-EKF**

The GS-EKF [35] assumes the posterior distribution of the state estimate to be a weighted sum of Gaussian densities, which are updated recursively based on the current observation.

**1) Forward filter:** In the forward GS-EKF, we consider $l$ Gaussians denoting the $l$-th Gaussian probability density function, with mean $x_{k,l}$ and covariance $\Sigma_{k,l}$ at $k$-th time-step, as $\gamma_{l} = \frac{1}{2\pi}^{-n/2} |\Sigma_{l}|^{-1/2} e^{-\frac{1}{2} (x_{k,l} - \Sigma_{k,l}^{-1} x_{k,l}^T)}$. Given observations $Y_k = \{y_j \}_{j=1}^l$ up to $k$-th time instant, the posterior distribution of state $x_k$ is approximated as $p(x_k|Y_k) = \sum_{l=1}^{l} c_{k,l} \gamma_{l}$, where $c_{k,l}$ is $i$-th Gaussian’s weight at $k$-th time step. Linearizing the functions as $F_{k,l} = \nabla x f(\hat{x})_{|x=x_{k,l}}$ and $H_{k,l} = \nabla x h(\hat{x})_{|x=x_{k,l}}$ for the $i$-th Gaussian, the means $\{x_{k,l,1}\}_{1 \leq i \leq l}$ and covariance matrices $\{\Sigma_{k,l,1}\}_{1 \leq i \leq l}$ are updated based on the current observation $y_{k+1}$ using independent EKF recursions.
for each Gaussian. Finally, the weights \( \{c_{i,k}\}_{1 \leq i \leq l} \) are updated as
\[
c_{i,k+1} = \frac{c_{i,k} \gamma [y_{k+1} - h(x_{i,k+1|k}), S_{i,k+1}]}{\sum_{i'=1}^{l} c_{i',k} \gamma [y_{k+1} - h(x_{i',k+1|k}), S_{i',k+1}]},
\]
where \( x_{i,k} \) is the \( i \)-th Gaussian’s predicted mean and \( S_{i,k} \) is the prediction covariance matrix of \( x_{i,k} \). With these updated Gaussians, the point-estimate \( \hat{x}_{k} \) and the associated covariance matrix \( \Sigma_{k} \) are
\[
\hat{x}_{k+1} = \sum_{i=1}^{l} c_{i,k} x_{i,k+1} \quad \text{and} \quad \Sigma_{k+1} = \sum_{i=1}^{l} c_{i,k} \left( \Sigma_{i,k+1} + \left( \Sigma_{i,k} - \Sigma_{i,k+1} \right)^{2} \right)^{2},
\]

2) Inverse filter: Consider an augmented state vector \( \mathbf{z}_{k} = \{ x_{i,k}, z_{i,k} \}_{1 \leq i \leq l} \) (means and weights of forward GS-EKF). Then, substituting for observation \( y_{k+1} \) from (7) in the forward filter’s updates similar to various inverse filters earlier and denoting the \( i \)-th EKF’s gain matrix as \( G_{i,k+1} = \Sigma_{i,k+1} H_{i}^{T} \Sigma_{i,k+1}^{-1} \), yields the state transition equations for inverse GS-EKF (I-GS-EKF)
\[
x_{i,k+1} = f(x_{i,k}) + K_{i,k+1} (y_{k+1} - h(x_{i,k})),
\]
\[
c_{i,k+1} = \gamma [h(x_{i,k+1}) + v_{k+1} - h(f(x_{i,k})), S_{i,k+1}].
\]

Treating \( \{ K_{i,k+1}, S_{i,k+1} \}_{1 \leq i \leq l} \) as time-varying parameters of the state transition equations, which are approximated in a similar way as we approximated the gain matrices for various inverse filters earlier, the overall state transition in terms of the augmented state is \( x_{k+1} = \mathbf{f}(x_{k}, x_{0:k}, v_{k+1}) \). Similarly, the observation \( z_{k} \) as a function of augmented state \( \mathbf{z}_{k} \) is \( z_{k} = g(x_{k}, e_{k}) = g \left( \sum_{i=1}^{l} c_{i,k} x_{i,k} \right) + e_{k} \).

The I-GS-EKF approximates the posterior distribution of the augmented state as a sum of \( l \) Gaussians with its recursions again following directly from forward GS-EKF’s recursions treating \( a_{k} \) as the observation. However, the inverse filter estimates an \( l \) \((n+1)\)-dimensional augmented state \( \hat{z}_{k} \) with the Jacobians with respect to the state denoted as \( \mathbf{F}_{j,k} \) and \( \mathbf{G}_{j,k+1} \), respectively. The process noise covariance matrix as \( \mathbf{Q}_{k} \) and \( \mathbf{V}_{k} \) for the \( j \)-th inverse filter’s Gaussian updates. The point estimate \( \hat{x}_{k} \) consists of the estimates \( \{ x_{i,k}, c_{i,k} \}_{1 \leq i \leq l} \) of the forward GS-EKF’s means and weights such that the point estimate \( \hat{x}_{k} \) of the forward filter’s estimate \( \hat{x}_{k} \) is \( \hat{x}_{k} = \sum_{i=1}^{l} c_{i,k} \hat{x}_{i,k} \).

When the filter considers only one Gaussian (\( l = 1 \)), the forward GS-EKF reduces to forward EKF with the only weight \( c_{1,k} = 1 \) for all \( k \). Hence, this weight need not be considered in the augmented state and \( z_{k} \) reduces to \( x_{1,k} \) which is the estimate \( \hat{x}_{k} \) itself. Similarly, I-GS-EKF also reduces to I-EKF if only one Gaussian is considered (\( l = 1 \)).

C. Inverse DEKF

Consider the adversary employing DEKF [22] as its forward filter. In DEKF, the output non-linearities are modified using dither signals so as to tighten the cone-bounds. Dithering tightens this cone such that the non-linearities are smoothed but it may also degrade the near-optimal performance of the EKF after the initial transient phase of estimation. Therefore, dithering is introduced only during the initial transient phase with the aim to improve the filter’s transient performance and avoid divergence. Denote the dither amplitude which controls the tightness of the cone-bounds by \( d \) and its amplitude probability density function by \( p(a) \). The observation function \( h(x) \) is dithered as \( h^{*}(x) = \int_{-d}^{d} h(a) p(a) da \). If \( d = d_{0} e^{-k/\tau} \), where \( d_{0} \) and \( \tau \) are constants and ‘\( k \)’ denotes the time index, then \( h^{*}(x) \rightarrow h(x) \) exponentially as \( k \rightarrow \infty \) during the transient phase.

The forward DEKF follows from conventional forward EKF of Section III-C by replacing \( h(\cdot) \) with \( h^{*}(\cdot) \) as the observation function of \( y_{k} \) during the initial transient phase and hence, the inverse DEKF (I-DEKF) also follows from I-EKF of Section III-C. The dither of the adversary’s filter is assumed to be known to us or may be estimated separately. Otherwise, the I-DEKF may also proceed with the unmodified observation function. We show in Section VII-D that these two formulations, labeled I-DEKF-1 and I-DEKF-2, respectively, generally vary in their estimation performances.

VI. Stability Analyses

For continuous-time nonlinear Kalman filtering, some convergence results were mentioned in [47]. In case of EKF, sufficient conditions for stability for non-linear systems with linear output map were described in [48]. The asymptotic convergence of EKF for a special class of systems, where EKF is applied for joint state and parameter estimation of linear stochastic systems, was studied in [41], [49]. If the nonlinearities have known bounds, then the Riccati equation is slightly modified to guarantee stability for the continuous-time EKF [50].

To derive the sufficient conditions for stochastic stability of non-linear filters, one of the common approaches is to introduce unknown instrumental matrices to account for the linearization errors [28]. It does not assume any bound on the estimation error, but its sufficient conditions for stability, the bounds assumed on the unknown matrices, are difficult to verify for practical systems.

Alternatively, [26] considers the one-step prediction formulation of the filter and provides sufficient conditions under which the state prediction error is exponentially bounded in mean-squared sense. We restate some results and a useful Lemma from [26].

Definition 1 (Exponential mean-squared boundedness [26]). A stochastic process \( \{ \zeta_{k} \}_{k \geq 0} \) is said to be exponentially bounded in mean-squared sense if there are real numbers \( \eta, \nu > 0 \) and \( 0 < \lambda < 1 \) such that \( \mathbb{E} \left( \| \zeta_{k} \|_{2}^{2} \right) \leq \eta \mathbb{E} \left( \| \zeta_{0} \|_{2}^{2} \right) \lambda^{k} + \nu \) holds for every \( k \geq 0 \).

Definition 2 (Boundedness with probability one [26]). A stochastic process \( \{ \zeta_{k} \}_{k \geq 0} \) is bounded with probability one if \( \sup_{k \geq 0} \mathbb{E} \left( \| \zeta_{k} \|_{2}^{2} \right) < \infty \) holds with probability one.

Lemma 1 (Boundedness of stochastic process [26, Lemma 2.1]). Consider a function \( V_{k}(\zeta_{k}) \) of the stochastic process \( \zeta_{k} \) and real numbers \( v_{\min}, v_{\max}, \mu, \) such that \( 0 < \lambda \leq 1 \) such that for all \( k \geq 0 \)
\[
v_{\min} \| \zeta_{k} \|_{2}^{2} \leq V_{k}(\zeta_{k}) \leq v_{\max} \| \zeta_{k} \|_{2}^{2},
\]
and
\[
\mathbb{E} \left( V_{k+1}(\zeta_{k+1}) \right) - V_{k}(\zeta_{k}) \leq \mu - \lambda V_{k}(\zeta_{k}).
\]

Then, the stochastic process \( \{ \zeta_{k} \}_{k \geq 0} \) is exponentially bounded in mean-squared sense, i.e.,
\[
\mathbb{E} \left( \| \zeta_{k} \|_{2}^{2} \right) \leq v_{\min} \mathbb{E} \left( \| \zeta_{0} \|_{2}^{2} \right) (1 - \lambda)^{k} + \frac{\mu}{v_{\min}} \sum_{i=1}^{k} (1 - \lambda)^{i},
\]
for every \( k \geq 0 \). Further, \( \{ \zeta_{k} \}_{k \geq 0} \) is also bounded with probability one.

In the bounded mean-squared sense, [26, Sec. III] showed that, while the two-step prediction and update recursion (described in previous sections) and one-step formulation of (forward) filters may differ in their performance and transient behaviour, they have same convergence properties. However, the conditions of Lemma 1 were proved to hold when the error remained within suitable bounds; the guarantees fail if the error exceeds this bound at any instant. However,
it was numerically shown [26] Sec. VI] that the bound on the error was only of theoretical interest and, in practice, the filter remained stable for much larger estimation errors.

In the following, we first derive stability conditions for I-KF-without-DF in which we rely on the stability of the forward KF-without-DF as proved in [51]. The procedure is similar for the stability of I-KF-with-DF and I-KF-without-unknown-input [9] and hence, we omit the details for these filters. For I-EKF stability, we employ both unknown matrix and bounded non-linearity approaches.

In the process, we also derive the forward EKF stability conditions using unknown matrix approach; note that the same was obtained using bounded non-linearity method in [26]. It is possible to extend these results to a general class of Gaussian filters [29], of which EKF using bounded non-linearity method in [26]. It is possible to extend these results to a general class of Gaussian filters [29], of which EKF is a special case, whose estimation error dynamics are represented by [22] eq. (2). Finally, we obtain stability conditions of forward and inverse SOEKF in the bounded mean-squared sense.

A. I-KF-with-unknown-input

Consider I-KF-without-DF of Section IV-A where the forward filter is asymptotically stable under the sufficient conditions provided by [51]. The following Theorem 1 states conditions for stability of the inverse filter.

**Theorem 1 (Stability of I-KF-without-DF).** Consider an asymptotically stable forward KF-without-DF [27] such that the gain matrices $M_k$ and $K_k$ asymptotically approach to limiting gain matrices $\overline{M}$ and $\overline{K}$, respectively. The measurement noise covariance matrix $\Sigma_\nu$ is positive definite (p.d.). Denote the limiting matrices $F = (I - K H) (I - MG H) F$ and $Q = \overline{E} \overline{F} \overline{E}^T$, where $\overline{E} = BM - KHBM + K$. Then, the I-KF-without-DF (7)-(10) is asymptotically stable under the assumption that pair $[\overline{F}, \overline{C}]$ is observable and the pair $[\overline{F}, \overline{C}]$ is controllable for the system given by (5) and (56), where $C$ is such that $\overline{Q} = C^T C$.

**Proof:** See Appendix A

Note that, for stability of I-KF-with-DF, the stability conditions of basic KF need to hold true for the augmented state considered in inverse filter formulation of Section IV-B. For the stability conditions of forward KF-with-DF, we refer the reader to [51].

B. I-EKF-without-unknown-input: Unknown matrix approach

Consider the two-step prediction and update formulation of I-EKF of Section III-C with EKF-without-unknown-input as the forward filter.

1) **Forward EKF stability:** Denote the forward EKF’s state prediction, state estimation and measurement prediction errors by $\tilde{x}_{k+1|k} - x_{k+1|k}$, $\tilde{x}_k - x_k$ and $\hat{y}_k - y_k$, with $\hat{y}_k = h(\tilde{x}_k(k-1)))$, respectively. Using (20), (21) and the Taylor series expansion of $f(\cdot)$ at $\tilde{x}_k$, we get

$$\tilde{x}_{k+1|k} = F_k (x_k - \tilde{x}_k) + w_k + O(\|x_k - \tilde{x}_k\|^2_2)$$

$$= F_k \tilde{x}_k + w_k.$$  

We consider the general case of time-varying process and measurement noise covariances and denote $Q, R$ and $\Sigma_\nu$ by $Q_k$, $R_k$ and $R_k$, respectively.

To account for the residuals and obtain an exact equality, we introduce an unknown instrumental diagonal matrix $U_k^y \in \mathbb{R}^{n \times n}$ [28] as

$$\tilde{x}_{k+1|k} = U_k^y F_k \tilde{x}_k + w_k.$$  

However, using (22), we have $\tilde{x}_k = \tilde{x}_{k|k-1} - K_k \hat{y}_k$, which when substituted in (56) yields

$$\tilde{x}_{k+1|k} = U_k^y F_k \tilde{x}_{k|k-1} - U_k^y F_k K_k \hat{y}_k + w_k.$$  

Similarly, using Taylor series expansion of $h(\cdot)$ at $\tilde{x}_{k+1|k}$ in (37) and introducing an unknown diagonal matrix $U_k^x \in \mathbb{R}^{n \times n}$ gives

$$\hat{y}_{k+1} = U_{k+1}^x H_{k+1} \tilde{x}_{k+1|k} + v_{k+1}.$$  

The prediction error dynamics of the forward EKF becomes

$$\tilde{x}_{k+1|k} = U_k^y F_k (I - K_k U_k^x H_k) \tilde{x}_{k|k-1} - U_k^y F_k K_k \hat{y}_k + w_k.$$  

Denote the true prediction covariance by $P_{k+1|k} = E[\tilde{x}_{k+1|k} \tilde{x}_{k+1|k}^T]$. Define $\delta P_{k+1|k}$ as the difference of estimated prediction covariance $\Sigma_{k+1|k}$ and the true prediction covariance $P_{k+1|k}$ while $\Delta P_{k+1|k}$ as the error in the approximation of the expectation $E[U_k^y F_k (I - K_k U_k^x H_k) \tilde{x}_{k|k-1} \tilde{x}_{k|k-1}^T (I - K_k U_k^x H_k) F_k^T U_k^y]$ by $U_k^y F_k (I - K_k U_k^x H_k) \tilde{x}_{k|k-1} \tilde{x}_{k|k-1}^T F_k^T U_k^y$. Denoting $Q_k = Q_k + U_k^y F_k K_k R_k K_k^T F_k^T U_k^y$ and following similar steps as in [28] [51], we have

$$\Sigma_{k+1|k} = U_k^y F_k (I - K_k U_k^x H_k) \Sigma_{k|k-1} (I - K_k U_k^x H_k) F_k^T U_k^y + Q_k.$$  

Similarly, denoting the true measurement prediction covariance and true cross-covariance by $F_k U_k^y$ and $F_k U_k^y$, respectively, we obtain

$$\Sigma_{k+1|k} = U_k^y F_k (I - K_k U_k^x H_k) \Sigma_{k|k-1} (I - K_k U_k^x H_k) F_k^T U_k^y + R_k.$$  

The following Theorem 2 provides stability conditions for the forward EKF using the unknown matrices $U_k^y$ and $U_k^y$.

**Theorem 2 (Stochastic stability of forward EKF).** Consider the non-linear stochastic system in (20) and (21). The two-step forward EKF formulation is as in Section III-C. Let the following assumptions hold true:

1) There exist positive real numbers $\overline{f}, \overline{r}, \overline{\sigma}, \overline{\tau}, \overline{q}, \overline{\nu}$ and $\overline{\sigma}$ such that the following bounds are fulfilled for all $k \geq 0$.

$$\|F_k\| \leq \overline{f}, \|H_k\| \leq \overline{r}, \|U_k^x\| \leq \overline{\sigma}, \|U_k^y\| \leq \overline{\tau},$$

$$\|U_k^x\| \leq \overline{q}, \overline{Q}_k \leq \overline{\nu}, \overline{R}_k \leq \overline{\sigma}, \overline{Q}_k \leq \overline{Q}_k,$$

$$\overline{f} \overline{r} + \overline{q} \overline{\nu} \leq \overline{\sigma}.$$

2) $U_k^y$ and $F_k$ are non-singular for every $k \geq 0$.

Then, the prediction error $\tilde{x}_{k|k-1}$ and the estimation error $\hat{y}_k$ of the forward EKF are exponentially bounded in mean-squared sense and bounded with probability one provided that the constants satisfy the inequality

$$\overline{f} \overline{r} \overline{\tau} \overline{\sigma}^2 < \overline{\nu}.$$  

**Proof:** See Appendix B
Theorem 3 (Stochastic stability of I-EKF). Consider the adversary’s forward EKF that is stable as per Theorem 2. Additionally, assume that the following hold true for all $k \geq 0$:

$$r_1 \preceq \mathbf{R}_k, \quad \| \mathbf{G}_k \| \leq \tilde{\gamma}, \quad \| \mathbf{U}_k \| \leq \tau, \quad \| \mathbf{U}_k^{pp} \| \leq \tilde{\tau},$$

$$\mathbf{R}_k \preceq r_1, \quad dI \preceq \hat{\mathbf{Q}}_k, \quad dI \preceq \hat{\mathbf{R}}_k, \quad \tilde{p} \mathbf{I} \preceq \mathbf{S}_{k-1} \preceq \tilde{p} \mathbf{I}.$$

for some real positive constants $r_1$, $\tilde{\gamma}$, $\tau$, $\tilde{\tau}, \tilde{p}, \tilde{\tau}$. Then, the state estimation error of I-EKF is exponentially bounded in mean-squared sense and bounded with probability one that the constants satisfy the inequality $\tilde{p} \mathbf{d} \mathbf{F}_\epsilon^2 \mathbf{d}^2 < d$.

Proof: See Appendix [C].

Note that Theorem 2 requires both $\hat{\mathbf{Q}}_k$ and $\hat{\mathbf{R}}_k$ to be p.d. In general, the difference matrices $\Delta \mathbf{P}_{k+1/k}$, $\Delta \mathbf{P}_{k+1}^{pp}$, and $\Delta \mathbf{P}_{k+1}^{pp}$ may not be p.d. One could enhance the stability of EKF by enlarging the noise covariance matrices by adding sufficiently large $\Delta \mathbf{Q}_k$ and $\Delta \mathbf{R}_k$ to $\mathbf{Q}_k$ and $\mathbf{R}_k$, respectively [28, 52]. The same argument also holds true for I-EKF noise covariance matrices.

C. I-EKF-without-unknown-input: Bounded non-linearity method

Consider the forward EKF’s one step prediction formulation (24)-(26). Using Taylor series expansion around the estimate $\hat{x}_k$, we have

$$f(x_k) = f(\hat{x}_k) + \phi(x_k, \hat{x}_k),$$

where $\phi(\cdot)$ and $\chi(\cdot)$ are suitable non-linear functions to account for the higher-order terms of the expansions. Denoting the estimation error by $e_k \triangleq x_k - \hat{x}_k$, the error dynamics of the forward filter is

$$e_{k+1} = (F_k - K_k H_k) e_k + r_k + s_k,$$

where $r_k = \phi(x_k, \hat{x}_k) - K_h \chi(\hat{x}_k, \hat{x}_k)$ and $s_k = w_k - K_h v_k$. The following Theorem (reproduced from [26]) provides sufficient conditions for the stochastic stability of forward EKF.

Theorem 4 (Exponential boundedness of forward EKF’s error). Consider a non-linear stochastic system defined by (20) and (27), and the one-step prediction formulation of forward EKF (24)-(26). Let the following assumptions hold true:

1) There exist positive real numbers $\tilde{\gamma}, \tilde{p}, \tilde{\tau}, \tilde{\epsilon}, \tilde{\delta}$ such that the following bounds are fulfilled for all $k \geq 0$:

$$\| F_k \| \leq \tilde{\gamma}, \quad \| F_k \| \leq \tilde{\gamma}, \quad \| H_k \| \leq \tilde{\tau},$$

2) $F_k$ is non-singular for every $k \geq 0$.

3) There exist positive real numbers $\kappa_\phi, \kappa_\chi, \kappa_\epsilon$ such that the non-linear functions $\phi(\cdot)$ and $\chi(\cdot)$ satisfy

$$\| \phi(x, \hat{x}) \| \leq \kappa_\phi \| x - \hat{x} \|^2$$

$$\| \chi(x, \hat{x}) \| \leq \kappa_\chi \| x - \hat{x} \|^2 \leq \kappa_\epsilon \| x - \hat{x} \|^2.$$

Then the estimation error given by (59) is exponentially bounded in mean-squared sense and bounded with probability one provided that the estimation error is bounded by suitable constant $\epsilon > 0$.

Theorem 3 guarantees that the estimation error remains exponentially bounded in mean-squared sense as long as the error is within suitable $\epsilon$ bounds. Further, the mean drift $\mathbb{E}[V_{k+1}(e_{k+1})\{e_{k+1} - V_k(e_k)\}$ for a suitably defined $V_k(\cdot)$ (for application of Lemma 1) is negative when $\epsilon \leq \| e_k \| \leq \epsilon$, which drives the system towards zero error in an expected sense. However, with some finite probability, the estimation error at some time-steps may be outside the $\epsilon$ bound. In this case, we cannot guarantee with probability one that the error will be within $\epsilon$ bound again at some future time-steps. As mentioned earlier, bounded non-linearity approach may not provide theoretical guarantees for the filter to be stable for all time-steps but, practically, the filter remains stable even if the estimation error is outside the $\epsilon$ bound provided that the assumed bounds on the system model are satisfied.

Finally, the error dynamics of the inverse filter, with the estimation error denoted by $\tilde{e}_k \triangleq x_k - \hat{x}_k$ and the inverse filter’s Kalman gain and estimation error covariance matrix by $\tilde{K}_k$ and $\tilde{S}_k$, respectively, is

$$\tilde{e}_{k+1} = (\tilde{F}_k - \tilde{K}_k \tilde{H}_k) \tilde{e}_k + r_k + s_k,$$

where $r_k = \tilde{\phi}_k(\hat{x}_k, \hat{x}_k) - \tilde{K}_h \tilde{\chi}_k(\hat{x}_k, \hat{x}_k)$ and $s_k = K_h v_k - \tilde{K}_h \epsilon_k$ with $\tilde{\phi}_k(\hat{x}_k, \hat{x}_k) = \phi(x_k, \hat{x}_k) - \chi(\hat{x}_k, \hat{x}_k)$. The following Theorem guarantees the stability of I-EKF. Note the additional assumption of $H_k$ to be full column rank for all $k \geq 0$, which implies $p \geq n$.

Theorem 5 (Exponential boundedness of I-EKF’s error). Consider the adversary’s forward one-step prediction EKF that is stable as per Theorem 2. Additionally, assume that the following hold true.

1) There exist positive real numbers $\tilde{\gamma}, \tilde{p}, \tilde{\tau}, \tilde{\epsilon}, \tilde{\delta}$ such that the following bounds are fulfilled for all $k \geq 0$.

$$\| G_k \| \leq \tilde{\gamma}, \quad \| I \| \leq \tilde{\tau}, \quad \| H_k \| \leq \tilde{\tau}.$$

2) $H_k$ is full column rank for every $k \geq 0$.

3) There exist positive real numbers $\kappa_\chi$ and $\kappa_\epsilon$ such that the non-linear function $\tilde{\chi}(\cdot)$ satisfies

$$\| \tilde{\chi}(\hat{x}, \hat{\hat{x}}) \| \leq \kappa_\chi \| \hat{x} - \hat{\hat{x}} \|^2 \quad \text{for} \quad \| \hat{x} - \hat{\hat{x}} \|^2 \leq \kappa_\epsilon.$$

Then, the estimation error for I-EKF given by (60) is exponentially bounded in mean-squared sense and bounded with probability one that the estimation error is bounded by suitable constant $\tau > 0$.

Proof: See Appendix [D].

D. I-SOEKF

The error dynamics of SOEKF cannot be expressed in a linear form [52, eq. (2)] for application of unknown matrix approach because of second-order terms. Therefore, we derive stability conditions using the bounded non-linearity approach here.

1) Forward SOEKF stability: Consider the one-step SOEKF’s formulation (47)-(53). Considering second-order terms as well, the Taylor series expansion of functions $f(\cdot)$ and $h(\cdot)$ at the estimate $\hat{x}_k$ are

$$f(x_k) = f(\hat{x}_k) + \frac{1}{2} \sum_{i=1}^{n} a_i (x_k - \hat{x}_k)^T \nabla^2 f(\hat{x}_k) (x_k - \hat{x}_k) + \phi(x_k, \hat{x}_k),$$

$$h(x_k) - h(\hat{x}_k) = H_k (x_k - \hat{x}_k) + \frac{1}{2} \sum_{i=1}^{p} b_i (x_k - \hat{x}_k)^T \nabla^2 h(\hat{x}_k) (x_k - \hat{x}_k) + \chi(x_k, \hat{x}_k),$$

where $\phi(\cdot)$ and $\chi(\cdot)$ are suitable non-linear functions to account for third and higher-order terms in the expansions. Using these expansions, the error dynamics of the forward filter with $e_k \triangleq x_k - \hat{x}_k$ is

$$e_{k+1} = (F_k - K_k H_k) e_k + r_k + q_k + s_k.$$  

(61)
where \( \mathbf{r}_k = \phi(x_k, \dot{x}_k) - K_k \chi(x_k, \dot{x}_k) \),
\[
\mathbf{q}_k = \frac{1}{2} \sum_{i=1}^{\mathbf{p}} \mathbf{a}_i \mathbf{e}_i^T \mathbf{G}_k^2 (\mathbf{f}(\hat{x}_k)), \mathbf{e}_k - \frac{1}{2} \sum_{i=1}^{\mathbf{p}} \mathbf{a}_i \mathbf{Tr} (\mathbf{G}_k^2 (\mathbf{f}(\hat{x}_k))) \mathbf{S}_k \\
- \frac{1}{2} \mathbf{K}_k \sum_{i=1}^{\mathbf{p}} \mathbf{b}_i \mathbf{e}_i^T \mathbf{G}_k^2 (\mathbf{h}(\hat{x}_k)), \mathbf{e}_k + \frac{1}{2} \mathbf{K}_k \sum_{i=1}^{\mathbf{p}} \mathbf{b}_i \mathbf{Tr} (\mathbf{G}_k^2 (\mathbf{h}(\hat{x}_k))) \mathbf{S}_k ,
\]
\( \mathbf{s}_k = \mathbf{w}_k - \mathbf{K}_k \mathbf{v}_k \).

The following Theorem 6 provides sufficient conditions for the stochastic stability of forward SOEKF.

**Theorem 6** (Exponential boundedness of forward SOEKF’s error). Consider the non-linear stochastic system defined by (20) and (7), and SOEKF’s one-step prediction formulation (47)-(55). Let the following assumptions hold true.

1) There exist positive real numbers \( \overline{\mathbf{f}}, \mathbf{f}, \overline{\mathbf{g}}, \mathbf{g}, \overline{\mathbf{b}}, \mathbf{b}, \delta \) and real numbers \( \overline{\mathbf{a}}, \overline{\mathbf{a}} \) (not necessarily positive) such that the following bounds are satisfied for all \( k \geq 0 \).

\[
\begin{align*}
\| \mathbf{f} \| & \leq \overline{\mathbf{f}} \\
\| \mathbf{g} \| & \leq \overline{\mathbf{g}} \\
\| \mathbf{b} \| & \leq \overline{\mathbf{b}}. 
\end{align*}
\]

2) \( \mathbf{F}_k \) is non-singular and \( \mathbf{F}_k^{-1} \) satisfies the following bound for all \( k \geq 0 \) for some positive real number \( \mathbf{f} \),

\[
\| \mathbf{F}_k^{-1} \| \leq \overline{\mathbf{f}}. 
\]

3) There exist positive real numbers \( \mathbf{c}_k, \mathbf{c}, \mathbf{c}_\chi, \mathbf{c}_x, \mathbf{c}_\chi \) such that the non-linear functions \( \phi(\cdot) \) and \( \chi(\cdot) \) satisfy

\[
\begin{align*}
\| \phi(\mathbf{x}, \dot{\mathbf{x}}) \| & \leq \mathbf{c}_k \| \mathbf{x} - \dot{\mathbf{x}} \|^2_2 \quad \text{for} \quad \| \mathbf{x} - \dot{\mathbf{x}} \| \leq \mathbf{c}_k \chi(\mathbf{x}, \dot{\mathbf{x}}) \| \mathbf{x} - \dot{\mathbf{x}} \|^2_2 \quad \text{for} \quad \| \mathbf{x} - \dot{\mathbf{x}} \| \leq \mathbf{c}_k. 
\end{align*}
\]

Then the estimation error given by (61) is exponentially bounded in mean-squared sense if the estimation error is within \( \epsilon \) bound for a suitable constant \( \epsilon > 0 \),

\[
\overline{\mathbf{f}} < \frac{2\epsilon^2}{\kappa_{\text{noise}}^2 n^{1/2} \rho},
\]

\[
\mathbf{q} \geq c, \quad \mathbf{q} \leq \overline{\mathbf{g}}.
\]

and

\[
\delta = \frac{1}{\kappa_{\text{noise}}} \left( \frac{\mathbf{c}_x^2}{2\gamma} - \epsilon c_\chi \right),
\]

for some \( \epsilon < \delta \), where \( \mathbf{c}, \alpha, \kappa_{\text{noise}} \) and \( c_\chi \) are constants that depend on the bounds assumed on the system.

**Proof:** See Appendix F.

2) Inverse SOEKF stability: Considering a suitable non-linear function \( \chi(\cdot) \), the Taylor series expansion of \( g(\cdot) \) at estimate \( \hat{x}_k \) of I-SOEKF’s one-step prediction formulation is

\[
g(\mathbf{x}_k) - g(\hat{x}_k) = \mathbf{G}_k (\hat{x}_k - \mathbf{x}_k) + \frac{1}{2} \sum_{i=1}^{\mathbf{p}} \mathbf{d}_i (\hat{x}_k - \mathbf{x}_k)^T \mathbf{G}_k^2 [g(\mathbf{x}_k)]_i (\hat{x}_k - \mathbf{x}_k) + \chi(\hat{x}_k, \dot{\hat{x}}_k),
\]

where \( \mathbf{d}_i \) is the \( i \)-th Euclidean basis vector in \( \mathbb{R}^{n_a \times 1} \). Finally, the error dynamics of the inverse filter with estimation error denoted by \( \mathbf{e}_k = \dot{x}_k - \mathbf{x}_k \) and the inverse filter’s Kalman gain and estimation error covariance matrix by \( \mathbf{K}_k \) and \( \mathbf{S}_k \), respectively, is

\[
\mathbf{e}_{k+1} = (\mathbf{F}_k - \mathbf{K}_k \mathbf{G}_k) \mathbf{e}_k + \mathbf{r}_k + \mathbf{q}_k + \mathbf{s}_k,
\]

where

\[
\mathbf{r}_k = \phi(\hat{x}_k, \dot{\hat{x}}_k) - \mathbf{K}_k \chi(\hat{x}_k, \dot{\hat{x}}_k),
\]

\[
\mathbf{q}_k = \frac{1}{2} \sum_{i=1}^{\mathbf{p}} \mathbf{a}_i \mathbf{e}_i^T \mathbf{G}_k^2 [j(\hat{x}_k)]_i, \mathbf{e}_k - \frac{1}{2} \sum_{i=1}^{\mathbf{p}} \mathbf{a}_i \mathbf{Tr} (\mathbf{G}_k^2 [j(\hat{x}_k)]_i) \mathbf{S}_k \\
- \frac{1}{2} \mathbf{K}_k \sum_{i=1}^{\mathbf{p}} \mathbf{b}_i \mathbf{e}_i^T \mathbf{G}_k^2 [h(\hat{x}_k)]_i, \mathbf{e}_k + \frac{1}{2} \mathbf{K}_k \sum_{i=1}^{\mathbf{p}} \mathbf{b}_i \mathbf{Tr} (\mathbf{G}_k^2 [h(\hat{x}_k)]_i) \mathbf{S}_k ,
\]

\[
\mathbf{s}_k = \mathbf{w}_k - \mathbf{K}_k \mathbf{v}_k,
\]

with \( \phi(\hat{x}_k, \dot{\hat{x}}_k) = \phi(\hat{x}_k, \dot{x}_k) - \mathbf{K}_k \chi(\hat{x}_k, \dot{x}_k) \) and \( \frac{1}{2} \sum_{i=1}^{\mathbf{p}} \mathbf{a}_i \mathbf{e}_i^T \mathbf{G}_k^2 [j(\hat{x}_k)]_i, \mathbf{e}_k = \frac{1}{2} \sum_{i=1}^{\mathbf{p}} \mathbf{a}_i \mathbf{e}_i^T \mathbf{G}_k^2 [j(\hat{x}_k)]_i \mathbf{e}_k - \frac{1}{2} \mathbf{K}_k \sum_{i=1}^{\mathbf{p}} \mathbf{b}_i \mathbf{e}_i^T \mathbf{G}_k^2 [h(\hat{x}_k)]_i, \mathbf{e}_k. 
\]

Then, the estimation error of I-SOEKF given by (62) is exponentially bounded in mean-squared sense if the estimation error is bounded by a suitable constant \( \overline{\mathbf{f}} > 0 \) and the bound constants also satisfy the equivalent conditions of (62), (63), and (64) for the inverse filter dynamics.

**Proof:** See Appendix F.

VII. NUMERICAL EXPERIMENTS

We illustrate the performance of proposed I-EKF considering two different examples systems. The efficacy of I-EKF is demonstrated by comparing the estimation error with CRLRB. The CRLRB provides a lower bound on mean-squared error (MSE) and is widely used to assess the performance of an estimator. For the discrete-time non-linear filtering, we employ the CRLRB [33]

\[
E \left[ \left( \mathbf{x}_k - \hat{\mathbf{x}}_k \right) \left( \mathbf{x}_k - \hat{\mathbf{x}}_k \right)^T \right] \geq J_k^{-1},
\]

where \( J_k \) is the Fisher information matrix

\[
J_k = \mathbb{E} \left[ \frac{\partial^2 \ln p(Y^k, X^k)}{\partial \mathbf{x}_k^2} \right].
\]

Here, \( X^k = \{ x_0, x_1, \ldots, x_k \} \) is the state vector series while \( Y^k = \{ y_0, y_1, \ldots, y_k \} \) are the noisy observations. Also, \( p(Y^k, X^k) \) is the joint probability density of pair \( (Y^k, X^k) \) and \( \mathbf{x}_k \) is an estimate of \( \mathbf{x}_k \) with \( \frac{\partial^2 \ln}{\partial \mathbf{x}_k^2} \) denoting the Hessian with second order partial derivatives. The sequence \( \{ J_k \} \) of information matrices can be computed recursively as [33].
\[ J_k = D_k^{22} - D_k^{21}(J_{k-1} + D_k^{11})^{-1}D_k^{12}, \tag{66} \]

where
\[ D_k^{11} = E \left[ \frac{\partial^2 \ln p(x_k|x_{k-1})}{\partial x_k^2} \right], \]
\[ D_k^{12} = E \left[ \frac{\partial^2 \ln p(x_k|x_{k-1})}{\partial x_k \partial x_{k-1}} \right], \]
\[ D_k^{21} = (D_k^{21})^T, \]
\[ D_k^{22} = E \left[ \frac{\partial^2 \ln p(x_k|x_{k-1})}{\partial x_k^2} \right] + E \left[ -\frac{\partial^2 \ln p(x_k|x_{k-1})}{\partial x_k \partial x_{k-1}} \right]. \]

For the non-linear system given by (20) and (7), the forward information matrices \( \{J_k\} \) recursions reduce to [28]
\[ J_{k+1} = Q_k^{-1} + H_k^T R_k^{-1} H_k + Q_k^{-1} F_k (J_k + F_k^T Q_k^{-1} F_k)^{-1} F_k^T Q_k^{-1}, \tag{67} \]
where \( F_k = \nabla_x f(x)|_{x=x_k} \) and \( H_k = \nabla_x h(x)|_{x=x_k} \). Note that, for the information matrices recursion, the Jacobians \( F_k \) and \( H_k \) are evaluated at the true state \( x_k \) while forward EKF recursions, these are evaluated at the estimates of the state. These recursions can be trivially extended to other system models considered in this paper as well as to compute the posterior information matrix \( \hat{J}_k \) for inverse filter’s estimate \( \hat{x}_k \).

Throughout all experiments, 100 time-steps (indexed by \( k \)) were considered. The initial information matrices \( J_0 \) and \( \hat{J}_0 \) were set to \( \Sigma_0^{-1} \) and \( \Sigma_0^{-1} \), respectively, unless mentioned otherwise.

A. Inverse KF with unknown inputs

Consider a discrete-time linear system without DF [54],
\[
x_{k+1} = \begin{bmatrix} 0.1 & 0.5 & 0.08 \\ 0.6 & 0.01 & 0.04 \\ 0.1 & 0.7 & 0.05 \end{bmatrix} x_k + \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} u_k + w_k,
\]
\[
y_k = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} \hat{x}_k + v_k,
\]
\[
a_k = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} \epsilon_k + u_k + \epsilon_k,
\]
with \( w_k \sim \mathcal{N}(0, 10I_3) \), \( v_k \sim \mathcal{N}(0, 20I_2) \) and \( \epsilon_k \sim \mathcal{N}(0, 25) \). The unknown input \( u_k \) was set to 50 for \( 1 \leq k \leq 50 \) and \( -50 \) thereafter. The initial state was \( x_0 = \begin{bmatrix} 1, 1, 1 \end{bmatrix}^T \). For the forward filter, the initial state estimate was set to \( \begin{bmatrix} 0, 0, 0 \end{bmatrix}^T \) with initial covariance \( \Sigma_0 = 10I_3 \). For the inverse filter, the initial state estimate was set to \( x_0 \) (known to ‘us’) itself with initial covariance \( \Sigma_0 = 15I_3 \).

For KF-with-DF, we considered the same linear system with a modified forward filter’s observations as [55]:
\[
y_k = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} x_k + \begin{bmatrix} 0 \\ 0 \end{bmatrix} u_k + v_k.
\]

Here, the initial input estimate was set to 10 with initial input estimate covariance \( \Sigma_0 = 10I_3 \) and initial cross-covariance \( \Sigma_{0w} = [0, 0, 0]^T \).

For the inverse filter, the initial estimate of the augmented state \( z_0 \) was set to \( [1, 1, 1, 50]^T \) with initial covariance \( \Sigma_0 = 15I_4 \).

Fig. 1 shows the time-averaged MSE (AMSE) \( \sqrt{\sum_{k=1}^N \|x_k - \hat{x}_k\|^2}/n \) for comparison here but omit it for later plots for clarity. The RCRLB value for state estimation is \( \sqrt{\text{tr}(J^{-1})} \) with \( J \) denoting the associated information matrix.

Fig. 1 shows that the effect of change in unknown input after 50 time-steps is negligible for KF-without-DF in both forward and inverse filters. However, for KF-with-DF, the sudden change in unknown input leads to an increase in state estimation error of the forward filter and, consequently, of the inverse filter. The estimation error of I-KF-without-DF is less than that of the corresponding forward filter for all time-steps. On the other hand, for KF-with-DF, the inverse filter’s estimation error converges to a higher value as compared to the forward filter, even though the initial estimation error assumed for the inverse filter is less than that assumed for the forward filter. Only I-KF-without-DF efficiently achieves the RCRLB bound on the estimation error.

B. Inverse filters for EKF, SOEKF and GS-EKF

Consider the discrete-time non-linear system model of FM demodulator without unknown inputs [36, Sec. 8.2]
\[
x_{k+1} = \begin{bmatrix} \lambda x_{k+1} + \theta_k \\ \theta_k \end{bmatrix} = \begin{bmatrix} \exp(-T/\beta) & 0 \\ -\beta \exp(-T/\beta) & 1 \end{bmatrix} \begin{bmatrix} \lambda_k \\ -1 \end{bmatrix} w_k,
\]
\[
y_k = \sqrt{2} \begin{bmatrix} \sin \theta_k \\ \cos \theta_k \end{bmatrix} + v_k, \quad \epsilon_k = \lambda_k^2 + \epsilon_k,
\]
with \( w_k \sim \mathcal{N}(0, 0.01) \), \( v_k \sim \mathcal{N}(0, 1) \), \( \epsilon_k \sim \mathcal{N}(0, 5) \), \( T = 2\pi/16 \) and \( \beta = 100 \). Here, the observation function \( g(\cdot) \) for the inverse filter is quadratic. Also, \( \lambda_k \) is the forward EKF’s estimate of \( \lambda_k \).

The initial state \( x_0 \) was set randomly with \( \lambda_0 \sim \mathcal{N}(0, 1) \) and \( \theta_0 \sim U[-\pi, \pi] \). All the initial state estimates of all forward and inverse filters including mean estimates for GS-EKF were also similarly drawn at random. The initial covariances were set to \( \Sigma_0 = 10I_2 \) and \( \Sigma_0 = 5I_4 \) for forward and inverse EKF as well as SOEKF. In the case of GS-EKF, we considered 5 Gaussians for the forward filter with the initial covariances and weights set to 10I_2 and

Fig. 2 shows that the effect of change in unknown input after 50 time-steps is negligible for KF-without-DF in both forward and inverse filters. However, for KF-with-DF, the sudden change in unknown input leads to an increase in state estimation error of the forward filter and, consequently, of the inverse filter. The estimation error of I-KF-without-DF is less than that of the corresponding forward filter for all time-steps. On the other hand, for KF-with-DF, the inverse filter’s estimation error converges to a higher value as compared to the forward filter, even though the initial estimation error assumed for the inverse filter is less than that assumed for the forward filter. Only I-KF-without-DF efficiently achieves the RCRLB bound on the estimation error.
that of the forward filter when the difference between AMSE and RCRLB for the inverse filters is being suboptimal filters, the forward as well as inverse EKF and which is also higher than that of the corresponding forward filters. Both I-EKF and I-effect of including second-order terms in the forward SOEKF is also reflected in the RCRLB of the inverse filters. The initial \( I_0 \) was taken close to the inverse of the steady state estimation covariance matrix of the forward filter. The initial \( I_0 \) only affects the RCRLB calculated for initial few time-steps. The RCRLB after these initial time-steps (around 20 for the considered system) shows same behaviour irrespective of the initial \( I_0 \). The Gaussian noise term \( \nu_{k+1} \) in I-GS-EKF’s state transition \( \{55\} \) is transformed through a non-linear function \( \gamma[\cdot, \cdot] \) such that \( \{67\} \) is not applicable. The RCRLB in this case is derived using the general \( J_k \) recursions given by \( \{66\} \).

Fig. 2 shows that the forward GS-EKF with \( l = 5 \) performs better than both forward EKF and forward SOEKF. This negligible effect of including second-order terms in the forward SOEKF is also reflected in the RCRLB of the inverse filters. Both I-EKF and I-SOEKF converge to the same steady-state estimation error values, which is also higher than that of the corresponding forward filters. Being suboptimal filters, the forward as well as inverse EKF and SOEKF do not achieve the RCRLB on the estimation error. However, the difference between AMSE and RCRLB for the inverse filters is less than that for the forward filters. We conclude that I-EKF and I-SOEKF are more efficient here.

For GS-EKF, the estimation error of the inverse filter is same as that of the forward filter when \( l = 2 \) but improves significantly when \( l = 5 \). Note that this improvement in performance comes at the expense of increased computational complexity because the inverse filter estimates an augmented state of dimension \( 'l(n+1)' \), which is larger than the forward filter’s state dimension \( 'n' \).

The I-EKF assumes initial covariance \( \Sigma_0 = 5I_2 \) (the true \( \Sigma_0 \) of forward EKF is \( 10I_2 \)) and a random initial state for these recursions. In spite of this difference in the initial estimates, I-EKF’s error performance is comparable to that of the forward EKF. Interestingly, despite similar differences in the initial estimates, I-GS-EKF with \( l = 5 \) outperforms the forward GS-EKF.

C. Inverse EKF with unknown inputs

For inverse EKF with unknown input, we modified the non-linear system model of Section \[ \text{VII-B} \] to include an unknown input \( u_k \) as

\[
x_{k+1} = \begin{bmatrix} \lambda_k+1 \\ \theta_{k+1} \end{bmatrix} = \begin{bmatrix} \exp(-T/\beta) & 0 \\ -\beta \exp(-T/\beta) - 1 & 1 \end{bmatrix} \begin{bmatrix} \lambda_k \\ \theta_k \end{bmatrix} + \begin{bmatrix} 0.001 \\ 1 \end{bmatrix} u_k + \begin{bmatrix} 1 \\ -\beta \end{bmatrix} \epsilon_k,
\]

where unknown input \( u_k \) was set to \( \pi/4 \) for \( 1 \leq k \leq 50 \) and \( -\pi/4 \) thereafter. The observation \( y_k \) of the forward EKF-without-DF was same as in Section \[ \text{VII-B} \]. Consider a linear measurement \( a_k \) for the inverse filter as \( a_k = \lambda_k + \epsilon_k \). For the forward filter, the initial input estimate was set to \( 0 \) while the inverse filter initial augmented state estimate consisted of the true state \( x_0 \) and true input \( u_0 \) (known to ‘us’) with initial covariance estimate \( \Sigma_0 = 15I_1 \).

Similarly, for system with DF, we again considered the same non-linear system (without any unknown input in \( x_k \) state transition) but with a modified forward filter’s observation \( y_k = \sqrt{2} \left[ \sin(\theta_k + u_k) \right] + \nu_k \). The input estimates \( \hat{u} \) and \( \hat{\theta} \) were also, as before, modulo \( 2\pi \). Again, the Gaussian noise terms in the inverse state transitions are transformed through non-linear functions such that \( \{67\} \) is not applicable. Fig. 3(a) shows that for both EKF with and without DF, the change in unknown input after 50 time-steps does not increase the estimation error (as for DF-with-DF in Fig. 1(b)). The estimation error of I-EKF-without-DF (I-EKF-with-DF) is higher (lower) than that of the corresponding forward filter. Any change in unknown input affects the inverse filter’s performance only when a significant change occurs in the forward filter’s performance.

D. Inverse DEKF

Consider the application of coordinate estimation of a stationary target from bearing observations taken by a moving sensor [32]. The actual coordinates of the stationary target are \( (X, Y) \) and that of the sensor at \( k \)-th time instant are \( (a_k, b_k) \). The constant velocity of sensor is \( s \). The forward and inverse EKF as well as DEKF were implemented in a modified coordinate basis with the state estimate

\[
x_k = [a_k/Y, s/Y, s, X/Y]^T \quad \text{and system model}
\]

\[
x_{k+1} = \begin{bmatrix} 1 & \Delta t & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} x_k + \begin{bmatrix} 0 \\ \Delta Y/\Delta t \\ \Delta t \\ 0 \end{bmatrix} w_k,
\]

where \( w_k \sim N(0, 1/2) \), \( \epsilon_k \sim N(0, 1) \), \( \epsilon_k \sim N(0, 1/2) \) and \( \Delta t = 20 \) s. The initial covariance estimates were set to \( \Sigma_0 = \text{diag}(4.44 \times 10^{-7}, 0.5 \times 10^{-6}, 1, 0.1) \) and \( \Sigma_0 = \text{diag}(10^{-6}, 6 \times 10^{-7}, 5, 0.5) \), respectively, for the forward and inverse filters. The initial state estimate for inverse filters were \( [0, 0.002, 200, 0]^T \). All other parameters of the system including the dither and the estimates were identical to those in [32].

With 200 time-steps, the modified observation function \( h^*(\cdot) \) in the forward DEKF replaced \( h(\cdot) \) up to 80 time-steps. Fig. 3(b) shows the absolute error and RCRLB, averaged over 400 runs, for estimation of \( X/Y \) whose estimate at the \( k \)-th time instant are given by \( \hat{x}_k \) and \( \hat{x}_k \) of the forward and inverse filters. The RCRLB at \( k \)-th time instant is \( \sqrt{J_k} \). The inverse filters’ estimation errors were significantly lower than that of the forward filters. While I-DEKF-1 and IDEKF-2 differ in their transient performance, they converge to the same steady-state error as I-EKF.
VIII. SUMMARY

We studied the inverse filtering problem for non-linear systems with and without unknown inputs in the context of counter-adversarial applications. These inverse filters allow ‘us’ to infer an estimate of the adversary’s estimate, given ‘our’ noisy observations of adversary’s actions. For systems with unknown inputs, the adversary’s observations may or may not be affected by the unknown input which is known to ‘us’ but not to the adversary. Addressing these two cases, we developed I-EKF and I-KF (each with and without DF) for non-linear and linear system dynamics, respectively. For systems without unknown inputs, we extended the theory of I-EKF to its variants, namely, I-SOEKF, I-GS-EKF and I-DEKF. These variants may provide improved estimation performance or stability depending on the system.

We investigated theoretical guarantees for the stability of I-KF-without-DF, I-EKF and I-SOEKF. In particular, stochastic stability of a forward filter with certain additional assumptions on the system is also sufficient for its inverse filter to be stable. For I-EKF, we considered two different approaches to study its stability, each with its own advantages. We also derived similar stability results for I-SOEKF. The asymptotic stability of I-KF-without-DF was obtained by extending standard KF stability results.

We demonstrated the efficacy of the different inverse filters through numerical examples using RCRLB as a performance measure. For the non-linear system without unknown inputs, we considered the FM demodulation and coordinate estimation applications. The FM demodulator model was also extended for systems with unknown inputs. Our experiments suggest that the impact of the unknown input on inverse filter’s performance highly depends on its impact on the forward filter’s performance. For certain systems, the inverse filter may perform more efficiently than the forward filter.

APPENDIX A

PROOF OF THEOREM[1]

Under the stability assumption of the forward filter, \( \hat{F}_k \) and \( E_k \) converge to \( F \) and \( E \), respectively, where \( F = (1 - KH)(I - BMH)F \) and \( E = BM - KBHM \) and \( R \), replaced by obtaining \( K_{k+1} \) and \( M_{k+1} \) by the limiting matrices \( K \) and \( M \), respectively, in \( F_k \) and \( E_k \). In this limiting case, the state transition equation (36) becomes

\[
\hat{x}_{k+1} = \hat{F} \hat{x}_k + E \hat{H} \hat{x}_{k+1} + E \nu_{k+1}.
\]

From (37), (38), and (40) and substituting the limiting matrices, the following Riccati equation is obtained

\[
\Sigma_{k+1|k} = \Sigma\Sigma|_{k+1|k}^{-1} \Sigma|_{k+1|k} - \Sigma|_{k+1|k}^{-1} G \Sigma|_{k+1|k}^{-1} F + Q,
\]

where \( Q = ERF^{-1} \). For the forward filter to be stable, covariance \( R \) needs to be p.d. \([51]\) and hence, \( Q \) is a p.s.d. matrix. With \( R \) being p.d. and the observability and controllability assumptions, \( \Sigma \) tends to a unique p.d. matrix \( \Sigma_0 \) satisfying

\[
\Sigma = F \Sigma \Sigma G T (GHG T + R)^{-1} (G \Sigma G T + R) \Sigma F T + Q,
\]

and \( F = \Sigma F \Sigma G T (G \Sigma G T + R)^{-1} G \Sigma F T + Q \) has eigenvalues strictly within the unit circle. These results follow directly from the application of \([56]\) Proposition 4.1, Sec. 4.1] similar to the stability and convergence results for the standard KF for linear systems \([56]\) Appendix E.4.

In this limiting case, the inverse filter prediction and update equations take the following asymptotic form

\[
\begin{align*}
\hat{x}_{k+1|k} &= \hat{F} \hat{x}_k + E \hat{H} \hat{x}_{k+1}, \\
\hat{x}_{k+1} &= \hat{x}_{k+1|k} + \hat{E} G \Sigma|_{k+1|k}^{-1}(\hat{n}_{k+1} - \hat{G} \hat{x}_{k+1|k}).
\end{align*}
\]

Denoting the inverse filter’s one-step prediction error as \( \tilde{r}_{k+1|k} = \hat{x}_{k+1} - \hat{x}_{k+1|k} \), the error dynamics for the inverse filter is obtained from this asymptotic form using (3) as

\[
\begin{align*}
\tilde{r}_{k+1|k} &= (F - F \Sigma \Sigma G T (G \Sigma G T + R)^{-1} G) \tilde{r}_{k+1} \\
&\quad - F \Sigma \Sigma G T (G \Sigma G T + R)^{-1} \hat{e}_{k} + \hat{E} \nu_{k+1}.
\end{align*}
\]

Since \( F - F \Sigma \Sigma G T (G \Sigma G T + R)^{-1} G \) has eigenvalues strictly within the unit circle, this error dynamics is asymptotically stable.

APPENDIX B

PROOF OF THEOREM[2]

For simplicity, we consider the case of \( n \geq p \) with \( U_{xy} \in \mathbb{R}^{n \times n} \). It is trivial to show that the proof remains valid for \( n < p \) as well. Using the expressions for \( \Sigma_{k+1|k} \) and \( S_{k+1} \), we have

\[
\begin{align*}
K_{k+1} &= \Sigma_{k+1|k} U_{xy} T H_{k+1} U_{xy} + R_{k+1}^{-1}, \\
\Sigma_{k+1|k} &= \Sigma_{k+1|k} - \Sigma_{k+1|k} U_{xy} T H_{k+1} U_{xy} \Sigma_{k+1|k}, \\
&\quad \ast \left( U_{xy} T H_{k+1} \Sigma_{k+1|k} H_{k+1} T U_{xy} + R_{k+1} \right)^{-1} \ast \\
&\quad \ast U_{xy} T H_{k+1} \left( U_{xy} T \Sigma_{k+1|k} U_{xy} \ast \right) \Sigma_{k+1|k}.
\end{align*}
\]

Define \( V_k (\tilde{X}_{k|k-1}) = \tilde{X}_{k|k-1} \Sigma_{k+1|k}^{-1} \tilde{X}_{k|k-1} \). Using the bounds assumed on \( \Sigma_{k+1|k} \), we have for all \( k \geq 0 \)

\[
\frac{1}{2} \| \tilde{X}_{k|k-1} \|_2^2 \leq V_k (\tilde{X}_{k|k-1}) \leq \frac{1}{2} \|\tilde{X}_{k|k-1}\|_2^2.
\]

Hence, the first condition of Lemma 1 is satisfied with \( \nu_{min} = 1/\sigma \) and \( \nu_{max} = 1/\sigma \).

Using (57) and the independence of noise terms, we have

\[
\begin{align*}
\mathbb{E} &\left[ V_k (\tilde{X}_{k+1|k}) | \tilde{X}_{k|k-1} \right] \\
&= \tilde{X}_{k|k-1} \Sigma_{k+1|k-1}^{-1} \tilde{X}_{k|k-1} + \mathbb{E} \left[ \nu_k T U_{xy} \tilde{F}_k K_k \right] U_{xy} \left( \tilde{U}_k \tilde{F}_k K_k \right) \nu_k | \tilde{X}_{k|k-1} \\
&+ \mathbb{E} \left[ \nu_k T U_{xy} \left( U_{xy} T \Sigma U_{xy} \right) \nu_k \right] | \tilde{X}_{k|k-1}.
\end{align*}
\]

(68)

The difference of two matrices \( A - B \) is invertible if maximum singular value of \( B \) is strictly less than the minimum singular value of \( A \). Using the assumed bounds, we have \( |K_k| \leq \hat{k} = (\sigma_3 / \sigma_4^2) / \hat{r} \). Hence, maximum singular value of \( K_k U_{xy} H_k \) is upper-bounded by \((\sigma_3 / \sigma_4^2) / \hat{r} \) and the inequality \([59]\) guarantees that \( I - K_k U_{xy} H_k \) is invertible (singular value of \( I \) is 1) such that

\[
\Sigma_{k+1|k} = U_{xy} F_k (I - K_k U_{xy} H_k).
\]

because \( U_{xy} \) and \( F_k \) are also assumed to be invertible. Again the assumed bounds, we have \( |U_{xy} F_k (I - K_k U_{xy} H_k)| \leq \sigma_4 / (1 + \sigma_3 / \hat{r}) \) which implies

\[
\begin{align*}
(U_{xy} F_k (I - K_k U_{xy} H_k))^{-1} Q_k (U_{xy} F_k (I - K_k U_{xy} H_k))^{-1} T^T &
\geq \frac{\hat{r}}{(\sigma_4 / (1 + \sigma_3 / \hat{r}))^2} I.
\end{align*}
\]

Using this bound in the expression of \( \Sigma_{k+1|k} \) as in \([53]\), we have

\[
\begin{align*}
(U_{xy} F_k (I - K_k U_{xy} H_k))^{-1} \Sigma_{k+1|k}^{-1} (U_{xy} F_k (I - K_k U_{xy} H_k))^{-1} T^T &
\geq (1 - \lambda) \Sigma_{k+1|k}^{-1},
\end{align*}
\]

where \( 1 - \lambda = \left(1 + \frac{\hat{r}}{(\sigma_4 / (1 + \sigma_3 / \hat{r}))^2} \right)^{-1} < 0 \) \( < 1 \). The last two expectation terms in (68) can be bounded by \( \mu = \left( \eta_3 / \sigma_4^2 \right)^2 \) \( + \eta_4 / \sigma_4^2 \) > 0 following similar steps as in \([53]\) such that...
\[ \mathbb{E} [V_{k+1}(\hat{x}_{k+1|k})|\hat{x}_{k|k-1}] - V_k(\hat{x}_{k|k-1}) \leq -\lambda V_k(\hat{x}_{k|k-1}) + \mu. \]

Hence, the second condition of Lemma 1 is also satisfied and the prediction error \( \hat{x}_{k|k-1} \) is exponentially bounded in mean-squared sense and bounded with probability one.

Furthermore, with the bounds assumed on various matrices, it is straightforward to show that
\[
\mathbb{E} \left[ \| \hat{x}_k \|^2 \right] \leq (1 + k\beta R) \mathbb{E} \left[ \| \hat{x}_{k|k-1} \|^2 \right] + T^2 \beta p.
\]

Finally, the exponential boundedness of \( \hat{x}_{k|k-1} \) leads to \( \hat{x}_k \) also being exponentially bounded in mean-squared sense as well as bounded with probability one.

**APPENDIX C**

**PROOF OF THEOREM 3**

We will show that the I-EKF’s dynamics also satisfies the assumptions of Theorem 2. For this, the following conditions C1-C13 need to hold true for all \( k \geq 0 \) for some real positive constants \( \alpha, \beta, \gamma, \delta, \tilde{a}, \tilde{b}, \tilde{c}, \tilde{d}, \tilde{e}, \tilde{f}, \tilde{g}, \tilde{h}, \tilde{p}, \tilde{q} \).

1. \( \| F_k^c \| \leq \tilde{\sigma}; \)
2. \( \| H_k \| \leq \tilde{\sigma}; \)
3. \( U_k^c \) is non-singular;
4. \( F_k^c \) is non-singular;
5. \( Q_k \preceq \tilde{\sigma}; \)
6. \( \| G_k \| \leq \tilde{\sigma}; \)
7. \( \| U_k^c \| \leq \tilde{\sigma}; \)
8. \( \| U_k^c \| \leq \tilde{\sigma}; \)
9. \( R_k \preceq \tilde{\sigma}; \)
10. \( d_k \preceq \tilde{\sigma}; \)
11. \( d_k \preceq \tilde{\lambda}; \)
12. \( \bar{p}_k \preceq \Sigma_{x|k-1} \preceq \bar{p} I; \) and
13. \( \tilde{\lambda} \) satisfies the inequality \( \beta \tilde{\lambda} \tilde{\lambda}^2 \prec \tilde{\sigma}^2. \)

Next, we prove that under the assumptions of Theorem 3, C1-C13 are satisfied. From the I-EKF’s state transition (23), the Jacobians \( F_k^c = F_k - K_k H_k F_k \) and \( F_k^c = K_k + 1 \) such that \( Q_k = K_k + 1 \) for all \( k \geq 0. \)

For C1, using \( \| K_{k+1} \| \leq \tilde{\sigma} \) (as proved in Theorem 2) and the bounds on \( F_k \) and \( H_k \) from the assumptions of Theorem 2, it is trivial to show that
\[
\| F_k^c \| = \| F_k - K_k H_k F_k \| \leq \tilde{J} + k\beta \tilde{J}.
\]

Therefore, C1 is satisfied with \( \tilde{\sigma} = \tilde{J} + k\beta \tilde{J} \).

For C2-C4, consider the unknown matrix \( U_k^c \) introduced to account for the residuals in linearization of \( f_k(\cdot) \). Let \( \tilde{x}_{k+1|k} \) and \( \tilde{x}_k \) denote the state prediction error and state estimation error of I-EKF. Similar to forward EKF with the introduction of the unknown matrix, we have
\[
\tilde{x}_{k+1|k} = U_k^c (F_k - K_k H_k F_k) \tilde{x}_k + K_k + 1 \nu_{k+1}. \tag{69}
\]

Also,
\[
\tilde{x}_{k+1|k} = f(\hat{x}_k) - f(\hat{x}_k) - K_k + 1 (h(f(\hat{x}_k)) - h(f(\hat{x}_k)))+ K_k + 1 \nu_{k+1}.
\]

Using the unknown matrices \( U_k^c \) and \( U_k^c \) introduced in the linearization of \( f(\cdot) \) and \( h(\cdot) \) respectively, we have
\[
\tilde{x}_{k+1|k} = (U_k^c F_k - K_k H_k F_k) \tilde{x}_k + K_k + 1 \nu_{k+1}.
\]

Comparing with (69), we have
\[
U_k^c (I - K_k H_k F_k) F_k = (I - K_k H_k F_k) U_k^c F_k. \tag{70}
\]

With the additional assumption of \( \tilde{p}_k \preceq R_k \) and using matrix inversion lemma as in proof of [25 Lemma 3.1], we have
\[
(I - K_k H_k F_k) F_k = (I - K_k H_k F_k) U_k^c F_k.
\]

Since \( \Sigma_{k+1|k} \) is invertible by the assumptions of Theorem 2, \( I - K_k H_k F_k \) is invertible for all \( k \geq 0 \) and
\[
(I - K_k H_k F_k)^{-1} = I + \Sigma_{k+1|k} H_k^T R_k^{-1} H_k + 1 H_k^T H_k.
\]

With the bounds assumed on various matrices, we have
\[
\| (I - K_k H_k F_k) F_k \| \leq 1 + \pi R^2.
\]

Furthermore, using this bound and the invertibility of \( I - K_k H_k F_k \), it is straightforward to show that \( U_k^c \) is invertible (both \( U_k^c \) and \( I - K_k H_k F_k \) are invertible under the assumptions of Theorem 2) and satisfies
\[
\| U_k^c \| \leq (1 + k\beta R)(1 + (\pi R^2)^2). \]

Also, since both \( I - K_k H_k F_k \) and \( F_k \) are invertible, \( \tilde{F}_k^c = F_k (I - K_k H_k F_k) \) is non-singular. Hence, C2-C4 are also satisfied with \( \tilde{\sigma} = (1 + k\beta R)(1 + (\pi R^2)^2). \)

For C5, using the upper bound on \( R_k \) from assumptions of Theorem 2, we have \( \tilde{Q}_k \preceq \tilde{\sigma} K_k + 1 K_k + 1 \). Since, \( \| K_{k+1} \| \leq \tilde{\sigma} \), the maximum eigenvalue of \( K_k + 1 K_k + 1 \) is bounded by \( \tilde{\sigma}^2 \) such that \( \tilde{Q}_k \preceq \tilde{\sigma} \tilde{Q}_k \). Hence, C5 is satisfied with \( \tilde{\sigma} \tilde{Q}_k \). The conditions C6-C13 are assumed to hold true in Theorem 3.

Hence, all the conditions hold true for the I-EKF dynamics and Theorem 2 is applicable for the I-EKF as well i.e. the estimation error is exponentially bounded in mean-squared sense and bounded with probability one.

**APPENDIX D**

**PROOF OF THEOREM 5**

We will show that the error dynamics of the I-EKF given by (69) satisfies the following conditions for all \( k \geq 0 \) for some real positive constants \( \kappa_\tilde{\sigma}, \epsilon_\tilde{\sigma} \).

1. \( \tilde{\sigma} I \preceq \tilde{Q}_k \);
2. \( \tilde{F}_k^c \) is non-singular matrix for all \( k \geq 0 \);
3. \( \| \bar{Q}_k(\hat{x}, \tilde{x}) \| \leq \kappa_\tilde{\sigma}_k \| \hat{x} - \tilde{x} \| \) for all \( \| \hat{x} - \tilde{x} \| \leq \epsilon_\tilde{\sigma} \) for some \( \kappa_\tilde{\sigma}, \epsilon_\tilde{\sigma} > 0 \).

All conditions of Theorem 4 can be proved to hold true for the I-EKF’s error dynamics under the assumptions of Theorem 5 following similar approach as in proof of Theorem 3 such that the estimation error given by (69) is exponentially bounded in mean-squared sense and bounded with probability one provided that the estimation error is bounded with \( \tilde{\sigma} \geq 0 \). \tilde{\sigma} depends on the various bounds in the same manner as \( \epsilon \) depends in the forward filter case.

For C1, using the bound on \( R_k \) from one of the assumptions of Theorem 3, we have
\[
\tilde{Q}_k = K_k R_k K_k^T \geq \tilde{\sigma} K_k K_k^T.
\]

Substituting for \( K_k \), we have
\[
K_k K_k^T = F_k \Sigma_k H_k^T (H_k \Sigma_k H_k^T + R_k)^{-1} H_k \Sigma_k F_k^T.
\]

With the assumption that \( H_k \) is full column rank, \( K_k K_k^T \) is p.d. as \( F_k \) is assumed to be non-singular in Theorem 3. Hence, there exists a constant \( \tilde{q} > 0 \) which is the minimum eigenvalue of \( K_k K_k^T \) such that \( K_k K_k^T \preceq \tilde{\sigma} I \) and \( \tilde{Q}_k \preceq \tilde{\sigma} I \). Hence, C1 is satisfied with \( \epsilon_\tilde{\sigma} = \tilde{\sigma} \). For C2, \( F_k^c = F_k - K_k H_k \) is proved to be invertible for all \( k \geq 0 \) as an intermediate result in the proof of Theorem 2 in [26 Lemma 3.1].
For $C3$, using $\|K_h\| \leq (\overline{7}\sigma T/\varepsilon)$ (proved in [26, Lemma 3.1]) and the bounds on functions $\phi(\cdot)$ and $\chi(\cdot)$ from the assumptions of Theorem 4 we have

$$\|\tilde{\phi}(\hat{x},\hat{x})\|_2 \leq \|\phi(\hat{x},\hat{x})\|_2 + \overline{7}\sigma T \|\chi(\hat{x},\hat{x})\|_2$$

$$\leq (\kappa_0 + \overline{7}\sigma T) \|\cdot - \hat{x}\|_2,$$

for $\|\hat{x} - \hat{x}\|_2 \leq \min(\epsilon_o, \epsilon_h)$. Hence, $C3$ is satisfied with $\kappa_\phi = \kappa_0 + (\overline{7}\sigma T/\varepsilon)\kappa_\chi$ and $\epsilon_\phi = \min(\epsilon_o, \epsilon_h)$.

**APPENDIX E**

**Proof of Theorem 6**

**A. Preliminaries to the Proof**

**Lemma 2.** Under the assumptions of Theorem 6 the following bounds hold true for all $k \geq 0$.

1. $\sum_{i=1}^n \sum_{j=1}^n a_i a_j^T \text{Tr} (\nabla^2 [f(\hat{x}_i)] \cdot \Sigma_k \nabla^2 [f(\hat{x}_j)]) \leq \pi^2 \sigma^2 n^2 I.$

2. $\sum_{i=1}^p \sum_{j=1}^p b_i b_j^T \text{Tr} (\nabla^2 [h(\hat{x}_i)] \cdot \Sigma_k \nabla^2 [h(\hat{x}_j)]) \leq \pi^2 \sigma^2 n^2 I.$

3. $\|M_k\| \leq \beta$ with $\beta > 0$.

**Proof:** Using the bounds from the assumptions of Theorem 6 we have for all $i, j \in \{1, 2, \ldots, n\}$

$$g^2 \sigma^2 I \leq \nabla^2 [f(\hat{x}_i)] \cdot \Sigma_k \nabla^2 [f(\hat{x}_j)] \leq \pi^2 \sigma^2 I,$$

which implies

$$g^2 \sigma^2 n \leq \text{Tr} (\nabla^2 [f(\hat{x}_i)] \cdot \Sigma_k \nabla^2 [f(\hat{x}_j)]) \leq \pi^2 \sigma^2 n.$$

Now, $\sum_{i=1}^n \sum_{j=1}^n a_i a_j^T$ is an $n \times n$ all-ones matrix with $n$ as one of its eigenvalue and all other $(n-1)$ eigenvalues are zero. Hence, $\sum_{i=1}^n \sum_{j=1}^n a_i a_j^T \text{Tr} (\nabla^2 [f(\hat{x}_i)] \cdot \Sigma_k \nabla^2 [f(\hat{x}_j)])$ is a p.s.d. matrix and satisfies the first bound in the lemma. Similarly, the bound on $\sum_{i=1}^p \sum_{j=1}^p b_i b_j^T \text{Tr} (\nabla^2 [h(\hat{x}_i)] \cdot \Sigma_k \nabla^2 [h(\hat{x}_j)])$ can be derived. Further, the maximum singular value of $\sum_{i=1}^n \sum_{j=1}^n a_i a_j^T \nabla^2$ and hence, using the bounds, we can show that $\|M_k\| \leq \beta$ with $\beta = \frac{1}{\pi} \sqrt{\overline{7}\sigma T} \sqrt{n} \sqrt{\overline{7}\sigma T}$.

**Lemma 3.** Under the assumptions of Theorem 6 there exists a real number $\alpha$ with $0 < \alpha < 1$ such that

$$(F_k - K_h H_k)^T \Sigma_k^{-1} (F_k - K_h H_k) \leq (1 - \alpha) \Sigma_k^{-1}.$$

**Proof:** Using (48) and (53), we have

$$\Sigma_{k+1} = F_k \Sigma_k F_k^T + Q_k - K_h S_k H_k^T + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n a_i a_j^T \text{Tr} (\nabla^2 [f(\hat{x}_i)] \cdot \Sigma_k \nabla^2 [f(\hat{x}_j)]) \cdot \Sigma_k).$$

Using the positive semi-definiteness of the last term (as proved in Lemma 2 and for substituting for $K_h S_k$ using (51)), we have

$$\Sigma_{k+1} \geq (F_k - K_h H_k) \Sigma_k (F_k - K_h H_k)^T + Q_k + K_h H_k \Sigma_k (F_k - K_h H_k)^T - M_k K_h^T.$$

Rearranging the terms,

$$\Sigma_{k+1} \geq (F_k - K_h H_k) \Sigma_k (F_k - K_h H_k)^T + Q_k + K_h H_k \Sigma_k (F_k - K_h H_k)^T - M_k K_h^T.$$
B. Proof of the Theorem

Consider $V_k(e_k) = e_k^T \Sigma_k^{-1} e_k$. Using the bounds on $\Sigma_k$ from assumptions of Theorem 6, we have

$$\frac{1}{\sigma} \| e_k \|^2 \leq V_k(e_k) \leq \frac{1}{\sigma} \| e_k \|^2.$$ 

Also, substituting for $e_{k+1}$ using (61), we have

$$V_{k+1}(e_{k+1}) = e_{k+1}^T \left((F_k - K_k H_k)^T \Sigma_k^{-1} (F_k - K_k H_k) e_k + s_k^T (\Sigma_k^{-1} e_k + 2s_k^T \Sigma_k^{-1} (F_k - K_k H_k) e_k + r_k)\right).$$

Using Lemma 3 and 5 we have for $\| e_k \| \leq \epsilon'$,

$$V_{k+1}(e_{k+1}) \leq (1 - \alpha) V_k(e_k) + \kappa_{sec} \| e_k \|^2 + \kappa_{sec} \| e_k \|^2 + c_{sec} + s_k^T \Sigma_k^{-1} (F_k - K_k H_k) e_k + r_k.$$ 

The last term $s_k^T \Sigma_k^{-1} (F_k - K_k H_k) e_k + r_k$ vanishes on taking expectation conditioned on $e_k$ and hence, for $\| e_k \| \leq \epsilon'$,

$$E[|V_{k+1}(e_{k+1}) - V_k(e_k)|] \leq \frac{1}{1 - \alpha} E[V_k(e_k)] + \kappa_{sec} \| e_k \|^2 + \kappa_{sec} \| e_k \|^2 + c_{sec} + \kappa_{noise} \delta,$$

where the bound of Lemma 5 is applied. But, for $\| e_k \| \leq \epsilon'$,

$$\kappa_{sec} \| e_k \|^2 + \kappa_{sec} \| e_k \|^2 \leq \kappa_{sec} \epsilon' \| e_k \|^2,$$

Choosing $\epsilon = \min \left(\epsilon', \frac{\alpha}{2 (\beta + \sqrt{\alpha (\sqrt{\beta} + \beta)})}\right)$, we have for $\| e_k \| \leq \epsilon$,

$$\kappa_{sec} \| e_k \|^2 + \kappa_{sec} \| e_k \|^2 \leq \frac{\alpha}{2} V_k(e_k),$$

which implies

$$E[V_{k+1}(e_{k+1}) - V_k(e_k)] \leq \frac{\alpha}{2} V_k(e_k) + c_{sec} + \kappa_{noise} \delta.$$ 

Hence, Lemma 1 is applicable. However, to have negative mean drift, we require $\delta$ to be small enough such that there exists some $\bar{\epsilon} < \epsilon$ to satisfy (64). This condition ensures that for $\| e_k \| \leq \epsilon$, $E[V_{k+1}(e_{k+1}) - V_k(e_k)] \leq 0$ is fulfilled and the estimation error $e_k$ remains exponentially bounded in mean-squared sense if the error is within $\epsilon$ bound. Note that the conditions of (63) and (64) suggest that $\delta$ needs to be chosen appropriately such that both the conditions could be satisfied simultaneously. However, the exact limits on $\delta$ depend on the other bounds of the system dynamics. Furthermore, as with the bounded non-linearity approach for EKF, these bounds may be very conservative and it is possible that the estimation error remains bounded outside this range as well (26 Sec. V).

### Appendix D

#### Proof of Theorem 7

From the state transition function of I-SOEFK (54), we have

$$\nabla^2 \left[f(\hat{x}_k)\right]_{i,j} \equiv \nabla^2 \left[f(\hat{x}_k)\right]_{i,j} - \sum_{j=1}^{p} \left[K_k\right]_{i,j} \nabla^2 \left[h(\hat{x}_k)\right]_{j,j}. $$

Using the upper bounds on Hessian matrices from the assumptions of Theorem 6, we have

$$\nabla^2 \left[f(\hat{x}_k)\right]_{i,j} \equiv \nabla^2 \left[f(\hat{x}_k)\right]_{i,j} - \sum_{j=1}^{p} \left[K_k\right]_{i,j} \nabla^2 \left[h(\hat{x}_k)\right]_{j,j}. $$

The proof of the remaining conditions of Theorem 6 for I-SOEFK dynamics follows from the proof of Theorem 5.
[56] D. P. Bertsekas, *Dynamic programming and optimal control*. Athena Scientific Belmont, 1995, vol. 1, no. 2.