Abstract—Two approaches, the extended Kalman filter (EKF) and moving horizon estimation (MHE), are discussed for state estimation for nonlinear dynamical systems under packet-dropping networks. For the EKF, we provide sufficient conditions that guarantee a bounded EKF error covariance. For MHE, a natural scheme on organizing the finite horizon window is proposed to handle the intermittent observations. A large-scale nonlinear programming software package, SNOPT, is employed in MHE and the formulation for constraints is discussed in detail. Examples and simulation results are presented.

Index Terms—State estimation, nonlinear systems, extended Kalman filter, moving horizon estimation, packet-dropping networks.

I. INTRODUCTION

Because of the technological advances in communication and computation, networks have become ubiquitous in today’s environment and the theory of networked control systems (NCSs) is an active area of research [1]. Fig. 1 shows the structure of a typical NCS. Unlike traditional control theory, measurements and control signals in an NCS travel through non-ideal communication networks in which information may be delayed, re-ordered, or even dropped. In this work we study an interesting problem within the NCS theory: online state estimation for nonlinear dynamical systems over packet-drop networks.

Some exciting progress has been reported in the area of linear estimation over packet-dropping networks. Studies on filtering with intermittent observations can be traced back to [2] and [3]. Other researchers tried to model Kalman filters with missing observations as jump linear systems, which are stochastic hybrid systems with linear dynamics and discrete Markov chains. Certain convergence conditions are given for expected estimation error covariance in [4] and [5]. More recently, Sinopoli et al. in [6] studied the behavior of the Kalman filter over an i.i.d. Bernoulli packet-dropping channel. They noticed a critical dropping probability, (i.e. a phase transition), above which the estimation error covariance diverges. Jin et al. in [7] showed that multiple-description codes can dramatically improve stability and performance of the Kalman filter over packet-dropping links while efficiently using the communication bandwidth. Although the work of [6] and [7] provide great insights into the state estimation problem with unreliable communication links, their models are restricted to linear dynamics with Gaussian noises. The estimation of a plant with nonlinear dynamics and non-Gaussian disturbances over a packet-dropping communication link has not been widely investigated.

Part of the difficulty with nonlinear dynamics and non-Gaussian noises is that theoretical guarantees are often hard to come by. Many strategies exist for online nonlinear estimation and we focus on two such strategies in this paper: the extended Kalman filter (EKF) and moving horizon estimation (MHE). Due to the ease of implementation and a widespread range of applications, many theoretical properties of the EKF have been explored: [8] shows that, when either the initial estimate is close enough to the true value or the nonlinearity of the system is weak enough, the EKF converges locally. Sufficient conditions that guarantee stochastic stability are derived in [9]. In [10], the authors linked the convergence behavior of the EKF to the derivative of the nonlinear dynamics. However, as noted by some authors, the existing convergence conditions are generally too conservative so that they are mainly of the theoretical interest.

MHE is an approach for online state estimation problem with nonlinear dynamics, constrained variables, and non-quadratic costs. The computation complexity is bounded by using a finite-size moving horizon window. As new measurements become available, old measurements are discarded, and the state estimation problem is resolved inside the horizon window. It has been shown in [11] that this approach can be used in some applications where the EKF is not appropriate. If the arrival cost is known exactly, then MHE provides the optimal Bayesian estimate. However, the arrival cost is difficult to compute in practice, and thus must be approximated. In such case, there are no known optimality guarantees. For a more in-depth discussion of MHE, we refer readers to [11].
Despite the lack of general performance guarantees, nonlinear systems arise often in practice. In today’s networked environment, it is important to understand the performance of nonlinear estimation schemes. We investigate the performance of the EKF and MHE for estimating the state of nonlinear dynamical systems with white Gaussian disturbance and observation noise over packet-dropping networks. We use the large-scale nonlinear programming software package, SNOPT, to solve the numerical optimization problem that arises in MHE.

The remainder of this paper is organized as follows. In Section II, the formulation of extended Kalman filter with the i.i.d. packet-dropping is presented. We give a sufficient condition on the boundedness of the expected EKF error covariance. In Section III, we discuss the method of moving horizon estimation with the details of SNOPT programming, where our estimation strategy for packet-dropping is proposed. Examples and simulation results are provided in Section IV and we conclude with remarks on future research directions in Section V.

II. EXTENDED KALMAN FILTER WITH OBSERVATION LOSS

For simplicity, we consider a nonlinear discrete-time dynamical system without control inputs

\[
\begin{align*}
    x_{k+1} &= f(x_k) + w_k \\
    y_k &= h(x_k) + v_k
\end{align*}
\]

(1)

where \( x_k \in \mathbb{R}^n \) is the state, \( y_k \in \mathbb{R}^m \) is the output, \( w_k \) and \( v_k \) are independent zero-mean, white Gaussian noise processes with covariances \( Q > 0 \) and \( R > 0 \), respectively. We assume that \( f(\cdot) \) and \( h(\cdot) \) are at least twice differentiable.

A. EKF without packet-dropping

The extended Kalman filter can be represented in two parts: the time update

\[
\begin{align*}
    \hat{x}_{k+1}^- &= f(\hat{x}_k) \\
    P_{k+1}^- &= A_k P_k A_k^T + Q
\end{align*}
\]

(2)

and the measurement update

\[
\begin{align*}
    \hat{x}_k &= \hat{x}_{k}^- + K_k (y_k - h(\hat{x}_k^-)) \\
    P_k &= (I - K_k C_k) P_{k}^-
\end{align*}
\]

(3)

where

\[
A_k = \frac{\partial f}{\partial x}(\hat{x}_k),
\]

\[
C_k = \frac{\partial h}{\partial x}(\hat{x}_k),
\]

\[
K_k = P_{k}^- C_k^T (C_k P_{k}^+ C_k^T + R)^{-1}.
\]

(4)

Let \( g_k(\cdot) \) denote the Riccati update for the error covariance

\[
g_k(X) = A_k X A_k^T + Q - A_k X C_k^T (C_k X C_k^T + R)^{-1} C_k X A_k^T.
\]

(5)

Since \( A_k \) and \( C_k \) are time-variant and they depend on the estimate at each step, it is difficult to give general conditions on uniform boundedness of the error covariance. Let us define the map \( H : \mathbb{R}^n \to \mathbb{R}^{m \times n} \) as

\[
H(x) = [h(x); h(f(x)); \cdots; h(f^{n-1}(x))]
\]

(6)

where

\[
f^{n-1}(x) = f(f(\cdots f(\cdot)))
\]

denotes function composition.

A nonlinear system is said to satisfy the observability rank condition if the rank of

\[
\frac{\partial H}{\partial x}(x_0) = \begin{bmatrix}
    \frac{\partial h}{\partial x}(x_0) \\
    \frac{\partial h}{\partial x}(x_1) \frac{\partial f}{\partial x}(x_0) \\
    \vdots \\
    \frac{\partial h}{\partial x}(x_{n-1}) \frac{\partial f}{\partial x}(x_{n-2}) \cdots \frac{\partial f}{\partial x}(x_0)
\end{bmatrix}
\]

(7)

equals \( n \) for any \( x_0 \in \mathbb{R}^n \). According to [8], [9], if the system (1) satisfies the observability rank condition, the uniformly bounded error covariance of the associated EKF is a sufficient condition so that the estimation error \( e_k = x_k - \hat{x}_k \) of the EKF is exponentially bounded, as long as either the initial guess is close enough to the true value or the function \( f(\cdot) \) is only weakly nonlinear. For a precise statement of the sufficient conditions, we refer the reader to Theorem 3.1 of [9] or Theorem 5.2 of [8]. We omit the exact statement to avoid excess notations.

B. EKF with packet-dropping

When the dynamical system and the estimator are spatially separated and connected by a communication network, additional conditions are required to bound the error covariance due to packet-dropping. We model the packet-dropping as an i.i.d. Bernoulli random process. A sequence of Bernoulli random variables \( \gamma_k \) is used to indicate whether a packet is successfully transmitted at each time step \( k \). More precisely, if \( \gamma_k = 1 \) then the packet goes through the communication network; otherwise, \( \gamma_k = 0 \) and the packet is dropped. This random process is characterized by a single parameter \( \lambda \):

\[
\gamma_k = \begin{cases} 
    1 & \text{with probability } \lambda \\
    0 & \text{with probability } 1 - \lambda
\end{cases}
\]

(8)

When a packet is lost, we proceed naturally with the time-update step. In the case of linear systems and Gaussian noise, this has been shown to be optimal in [6]. We have no such guarantee in nonlinear systems. Using this strategy, the Riccati update for the EKF becomes

\[
g_k^0(X) = A_k X A_k^T + Q
\]

(9)

when \( \gamma_k = 0 \), and

\[
g_k(X) = A_k X A_k^T + Q - A_k X C_k^T (C_k X C_k^T + R)^{-1} C_k X A_k^T.
\]

(10)

when \( \gamma_k = 1 \). Thus, the error covariance recurrence of EKF is stochastic and we have the following theorem.

**Theorem 2.1:** Consider the nonlinear system (1) with the following properties

1) The system satisfies the observability rank condition;

2) The first-order derivative \( \frac{\partial f}{\partial x} \) is invertible for any \( x \in \mathbb{R}^n \);
3) There exists a detectible pair \((A, C)\) such that
\[
A \geq \frac{\partial f}{\partial x} \bigg|_{x=x_0} \quad \text{and} \quad C^T R^{-1} C \leq \frac{\partial h^T}{\partial x} \bigg|_{x=x_0} \frac{\partial h}{\partial x}
\]
for all \(x_0 \in \mathbb{R}^n\).

Then, the expected error covariance \(E[P_k^-]\) is uniformly bounded if
\[
\lambda > 1 - 1/\rho(A)^2
\]
where \(\rho(A)\) is the spectral radius of \(A\).

**Proof:** The first two properties guarantee that \(P_k^-\) is uniformly bounded without packet drops. In order to show \(E[P_k^-] > 0\) is uniformly bounded with packet drops, we need to find an upper bound. Let
\[
\tilde{g}^0(X) = AXA^T + Q
\]
and
\[
\tilde{g}^1(X) = AXA^T + Q - AXC^T(CXC^T + R)^{-1}CXTA^T.
\]

It is true that
\[
\begin{align*}
\tilde{g}^0_k(X) & \leq g^0(X) \\
\tilde{g}^1_k(X) & \leq \tilde{g}^1(X)
\end{align*}
\]
for any \(k\). The first inequality is obvious from the definition of \(A\). For the second inequality, note that the update for error covariance in EKF can be re-written as
\[
\begin{align*}
P_{k+1}^- & = A_k P_k^- A_k^T + Q \\
P_k^- & = (P_k^-)^{-1} + C_k^T R^{-1} C_k.
\end{align*}
\]

Thus, with the same initial conditions, the error covariance \(P_k^-\) is bounded by the error covariance \(\tilde{P}_k^-\). Here, \(\tilde{P}_k^-\) corresponds to the Kalman filter error covariance for the linear system
\[
\begin{align*}
x_{k+1} & = Ax_k + w_k \\
y_k & = Cx_k + v_k.
\end{align*}
\]

So we have
\[
E[\tilde{P}_k^-] \leq E[P_k^-].
\]

According to [6], [7], the expected value \(E[\tilde{P}_k^-]\) evolves according to the modified ARE
\[
\begin{align*}
g_\lambda(X) & = AXA^T + Q - \lambda AXC^T(CXC^T + R)^{-1}CXTA^T.
\end{align*}
\]

And it has been shown in [6], [7] that \(E[\tilde{P}_k^-]\) converges to a unique positive definite matrix, i.e. uniformly bounded, as \(k \to \infty\) if the packet-dropping rate \(1 - \lambda\) satisfies
\[
1 - \lambda < 1/\rho(A)^2.
\]

This theorem states a sufficient condition on the uniform boundedness of the error covariance of EKF with packet-dropping. It indicates that the EKF exhibits a similar phase transition as the Kalman filter with respect to the packet drops. However, we have the following comments on this result:

- First, this condition is conservative since the behavior of the EKF is bounded by a Kalman filter of an approximate linear system. This conservativeness is verified in Section IV by simulation results.
- It is well known that the Riccati update in the EKF is only a first-order approximation to the true error covariance. In other words, the uniform boundedness of \(P_k^-\) does not necessarily indicate the boundedness of \(\tilde{P}_k^-\). Other conditions on the linearity of \(f(\cdot)\) and \(h(\cdot)\) as well as the precision of the initial guess must be considered. For the EKF with packet-dropping, Theorem 2.1 can only be used to judge the boundedness of \(E[P_k^-]\). The behavior of \(E[\tilde{P}_k^-]\) is still under investigation.

### III. Moving Horizon Estimation with Packet-Dropping

Other than EKF, moving horizon estimation (MHE) is another online method to estimate the state of the nonlinear system (1), which is formulated as an optimization problem to handle constraints explicitly. The optimization problem solves
\[
\begin{aligned}
\min_{x_0, \{w_k\}_{k=0}^{T-1}} \sum_{k=0}^{T-1} \|w_k\|_{Q_k}^2 + \|v_k\|_{R_k}^2 + \|x_0 - \hat{x}_0\|_{\Pi_{-1}}^2
\end{aligned}
\]
at time \(T\). As \(T\) increases, more observation data are taken into account and the optimization increases in size. To limit the amount of computation, MHE considers a finite-size horizon window. Fig. 2 illustrates this concept. When the time step increases by one, the horizon window moves one step to the right by including one new observation data and discarding the oldest one. More precisely, MHE solves the following optimization at each time \(T\)
\[
\begin{aligned}
\min_{x_0, \{w_k\}_{k=0}^{T-1}} \sum_{k=0}^{T-1} \|w_k\|_{Q_k}^2 + \|v_k\|_{R_k}^2 + \mathcal{Z}_{T-N}(x_{T-N})
\end{aligned}
\]
where \(N\) is the horizon window size and \(\mathcal{Z}_{T-N}(x_{T-N})\) is called the arrival cost, which summarizes the past information up to time \(T - N\). For general nonlinear systems with the form (1), it is difficult to determine the true arrival cost. As often done in practice, we use the weighted deviation from the EKF trajectory as the arrival cost. In symbols,
\[
\mathcal{Z}_{T-N}(x_{T-N}) = \|x_{T-N} - \hat{x}_{T-N}\|^2 \cdot \Pi_{-1}
\]
where \(\Pi_{-1}\) is the error covariance of EKF.

#### A. Formulation of MHE with SNOPT

A MHE scheme needs to solve an optimization problem at each step. We use SNOPT [12], a general-purpose software package for solving optimization problems as our numerical solver. In the cost function (16), there are \(2N\) variables total and they are \(\{x_{T-N}, w_{T-N}, \cdots, w_{T-1}, v_{T-N}, \cdots, v_T\}\). Our goal is to solve (16) subject to the following \(N - 1\)
nonlinear constraints:
\[
\begin{cases}
f(x_{T-N}) + v_{T-N+1} + w_{T-N} - y_{T-N+1} = 0 \\
f(y_{T-N+1} - v_{T-N+1}) + v_{T-N+2} + w_{T-N+1} - y_{T-N+2} = 0 \\
\vdots \\
f(y_{T-1} - v_{T-1}) + v_{T} + w_{T-1} - y_{T} = 0
\end{cases}
\] (18)
and one linear constraint
\[x_{T-N} + v_{T-N} - y_{T-N} = 0.\] (19)
These equality constraints arise from the system dynamics. Additional inequality constraints on variables can be introduced to model, for example, bounded noise.

To compute the Jacobian matrix of those constraints, we order those variables as follows
\[
\begin{pmatrix}
x_{T-N}, v_{T-N+1}, \cdots, v_{T}, w_{T-N}, \cdots, w_{T-1}, v_{T-N}
\end{pmatrix}
\]
and to yield the sparse Jacobian matrix as
\[
\frac{\partial f}{\partial x_{T-N}} \begin{pmatrix}
1 \\
\frac{\partial f}{\partial v_{T-N+1}} \\
\ddots \\
\frac{\partial f}{\partial v_{T}} \\
\frac{\partial f}{\partial w_{T-1}} \\
1
\end{pmatrix}
\begin{pmatrix}
1 \\
1 \\
\ddots \\
1 \\
1
\end{pmatrix}
\] (20)
where the left-upper \((N-1) \times (N-1)\) sub-matrix is the nonlinear Jacobian matrix and the last row corresponds to the linear constraint.

After solving this optimization problem, the estimation of \(x_T\) can be calculated by
\[
\begin{align*}
x_{T-N+1} &= f(X_{T-N}) + w_{T-N} \\
x_{T-N+2} &= f(X_{T-N+1}) + w_{T-N+1} \\
x_T &= f(X_{T-1}) + w_{T-1}.
\end{align*}
\] (21)

B. Estimation strategy for packet-dropping

When a packet is dropped in the communication network, the estimator has to predict the state value at that time step. For the EKF, it is natural to proceed with the time update step until a packet is successfully received. For MHE, we propose the following strategy to handle packet loss.

Suppose at time \(k\) the packet is dropped, i.e., the estimator does not receive \(y_k\). We estimate the state at time \(k\) as
\[\hat{x}_k = f(\hat{x}_{k-1}).\] (22)
This one-step propagation method is used whenever the packet is dropped. If consecutive packets are dropped, we perform this time update multiple times.

Suppose at time \(k\) the packet is received, the estimator uses the latest \(N\) received observation data as nonlinear constraints. Because multiple packets may be dropped in succession, the time indices of the last \(N\) received packets may not be consecutive, i.e., the time intervals between two received packets could be larger than 1. Fig. 3 shows this strategy for MHE with packet-drops. The width of the horizon may vary at each step, but the number of valid, successfully received observation packets inside the window is constant. When the time step increases by one, the horizon window may or may not moves to the right by discarding the oldest observation data. This depends on whether the estimator receives a new packet. Based on this strategy, the estimator only needs to store the latest \(N\) received packets at any time step \(k\). Table I shows the estimator memory and the list of the time intervals.

Due to the packet loss, we must use a different cost function in the optimization as
\[
\min_{x_0, \{w_k\}_{k=1}^{N}} \sum_{k=i_i}^{i_N} \|w_k\|_Q^2 + \|v_k\|^2_{R-1} + \|x_{i_i} - \hat{x}_{i_i}\|^2 \cdot \Pi_{i_i}^{-1}.
\] (23)
The arrival cost is again based on the output of the EKF at
As a comparison, we first consider the case with no packet-dropping since the trajectory according to MHE. The green curve shows the actual states, the cyan noise and stable dynamics, the EKF is almost as good as processes with covariances $Q$ and $R$.

The difference between constraints in (18) and (24) is that the nonlinear function $f(\cdot)$ is replaced by compositions $f^m(\cdot)$. While the Jacobian matrix has the same form as in Equ. (20), the derivatives of the function compositions need to be calculated recursively since

$$
\frac{\partial f^m}{\partial x} = \frac{\partial f(f^{m-1})}{\partial x} \cdot \frac{\partial f^{m-1}}{\partial x}.
$$

IV. EXAMPLES AND SIMULATION RESULTS

In this section, we apply the aforementioned estimation strategies to two scalar nonlinear systems as examples.

A. Example: stable nonlinear system

Let us first consider the system

$$
\begin{cases}
  x_{k+1} = x_k - 0.001 \cdot x_k(x_k + 2)(x_k - 5) + w_k \\
  y_k = x_k + v_k,
\end{cases}
$$

(26)

where $w_k$ and $v_k$ are zero-mean white Gaussian noise processes with covariances $Q = 0.01$ and $R = 1$, respectively. Without noise, this system satisfies the observability rank condition and has three equilibrium points $\{0, -2, 5\}$, where $-2$ and $5$ are stable equilibrium points and $0$ is unstable. As a comparison, we first consider the case with no packet-drops. Fig. 4 shows the performance of the EKF and MHE with window size 70. It is apparent that with Gaussian noise and stable dynamics, the EKF is almost as good as MHE. Fig. 5 shows the horizon window of MHE. The red interval represents the estimated state values inside the horizon window based on the results of the numerical solver, SNOPT. The green curve shows the actual states, the cyan dots are the noisy observations, and the blue is the estimated trajectory according to MHE.

Fig. 6 and 7 show the simulation results of the EKF and MHE, respectively, under various packet-dropping conditions. Large packet drop rates degrade estimator performance regardless which approach is used. Since the dynamics are stable, both the EKF and MHE eventually converges to the stable state. The output of estimator breaks into discontinuous pieces at high packet-dropping rate. Each piece corresponds to an interval of prediction due to continuous packet drops. When a packet successfully received, the estimator updates its output and the estimated trajectory jumps. Fig. 8 shows a typical horizon window of MHE.

The red cross represents the estimated state values inside the horizon window. The green curve shows the true states, the cyan dots denote the received noisy outputs which is quite sparse due to the high loss rate (40%), and the black is the estimated trajectory. Since the system (26) is rather tame, we can attribute the comparable performance to the fact that the arrival cost of MHE is dependant on the EKF.

B. Example: unstable nonlinear system

The second system that we consider is

$$
\begin{cases}
  x_{k+1} = 1.1 \cdot x_k + 0.2 \cdot \sin(x_k) + w_k \\
  y_k = x_k + v_k,
\end{cases}
$$

(27)

where $w_k$ and $v_k$ are zero-mean Gaussian white noise processes with covariances $Q = 1$ and $R = 1$, respectively. This system only has one equilibrium point at 0 and it is unstable.

Obviously, the nonlinear system (27) satisfies the observability rank condition since $y_k = x_k + v_k$. The derivative of
$f(\cdot)$ is bounded by

$$0.9 \leq \frac{\partial f}{\partial x} \leq 1.3.$$  

According to Theorem 2.1, the sufficient condition for uniform boundedness of the expected error covariance is

$$\lambda > 0.41,$$  

i.e., the packet-dropping rate should be below 59%. Fig. 9 shows the simulation result for the EKF where the average error covariance does not start to diverge until the packet-dropping rate is over 80%. This is a good example to show how conservative the condition is.

For the estimated trajectories of the EKF and MHE, they are similar to Fig. 6 and 7. We omit them due to space limitations.

V. CONCLUSIONS AND FUTURE WORK

In this paper, we investigated the state estimation problem for nonlinear dynamical systems with packet-drops. For the EKF, we state the sufficient condition on the uniform boundedness of the expected error covariance. Even though this condition is conservative, it indicates the existence of a phase transition in EKF, which is similar to the linear case. For MHE, we introduced an estimation strategy to compensate for the intermittent data. When packet-dropping occurs, the estimator conducts prediction. When a packet is received successfully, the packet history of estimator is updated and the nonlinear constraints in MHE are reformulated correspondingly. Simulation results are presented for two scalar nonlinear dynamics for both EKF and MHE. The simulation results verify the expected behaviors of these two estimators.

The future work includes a few of issues. First of all, we would like to study both estimators on general unstable nonlinear dynamics with non-Gaussian noise so that we can compare the statistical behaviors of error covariance of EKF and MHE; Second, it will be interesting to investigate the
behavior of expected error $e_k$ of the EKF with packet-dropping. Lastly, we would like to investigate whether multiple description codes can improve the performance of the EKF and MHE.

VI. ACKNOWLEDGEMENTS

The authors would like to thank Dr. Lars Cremean for the help on SNOPT programming. They also want to thank Ling Shi and Henrik Sandberg for useful feedbacks.

REFERENCES


