Moving-Horizon Estimation for Discrete-Time Systems with Measurements Subject to Outliers

Moath Awawdeh
University of Genoa - Italy

Presenting at IET Bahrain Local Network

The preliminary work of this presentation was introduced at the IEEE 53rd CDC, Los Angeles in December 16, 2014
awawdeh@dime.unige.it

February 1, 2015
Overview

1. Overview
2. Outlier Detection
3. Moving-Horizon Estimation MHE
4. Kalman Filtering
5. State Estimation with Measurement Corrupted by Outliers
6. Standard Solution based on the Kalman Filter (KF)
7. Proposed Solution: Moving-Horizon Estimation (with Guaranteed Robustness)
8. Stability of Proposed Moving-Horizon Filter MHF
9. Simulation and Comparison with Kalman Filtering
10. Conclusion
Across a wide variety of fields, data are being collected and accumulated more and more easily at an increasingly cheaper cost.

There is an urgent need for a new generation of computational theories and tools to assist human in extracting useful information from the rapidly growing volumes of digital data.

As a part, measurements technology is advancing in the industry. Innovation such as wireless transmitter, reduced cost of measurement technology, and active monitoring have the effect of increasing number of available measurements.

To success in extracting knowledge and building systems We need the data (measures) to be CLEAN and CORRECT!
Outlier Detection (cont’d)

- **Data-mining approach** - *data and statistics*: In many data analysis tasks a large number of variables are being recorded or sampled. One of the first steps towards obtaining a coherent analysis is the detection of outlaying observations.

- **Control system approach** - *measures and noise*: In numerous applications there exists the problem of dealing with large deviations in the measurements because of sensor malfunctions, wrong replacement of measures, or large non-Gaussian noises. Such abnormal data are called *Outliers*.

**Hawkins (Hawkins, 1980)**

defines an outlier as an observation that deviates so much from other observations as to arouse suspicion that it was generated by a different mechanism.
Outlier Detection Methods

- There is no rigid mathematical definition of what constitutes an outlier.
- There are many outlier detection methods and they can be divided generally between *univariate vs multivariate* and *parametric vs non parametric* methods.
- Since different outlier detection algorithms are based on disjoints sets of assumption, a direct comparison between them is not always possible.
- Different models are based on different assumptions to model outliers.
- Different models will produce different results.

Example (Outlier Detection Methods for GPS Networks)

Conventional Statistical Test Methods such as "The τ Test", Fuzzy Logic Method, M, L, and R estimators,...etc.
Outlier Detection Applications such as

Intrusion detection and cloud management

Positioning systems

Intrusion detection system

Clinical alerting and heart surgery

Process control and fault detection

Network anomaly detection
State Estimation
The state estimation problem is to determine the current state based on a sequence of past and current measurements at discrete-time instants, and the use of dynamic model.

MHE is an optimization-based strategy for process monitoring and state estimation
MHE computes an estimate at the current instant by solving a least squares problem, based on a limited amount of most recent information collected over a finite horizon.
Kalman Filter (KF)

- An estimator for what is called the linear-quadratic Gaussian problem, which is the problem of estimating the instantaneous "state" of a linear dynamic system perturbed by Gaussian white noise, by using measurements linearly related to the state, but corrupted by Gaussian white noise.

- It is recursive so that new measurements can be processed as they arrive.

- The resulting estimator is statistically optimal with respect to any quadratic function of estimator error.

- The applications of Kalman filtering encompass many fields, but its use as a tool is almost exclusively for two purposes:

  *Estimating the state of dynamic systems.*

  *Analysis of estimation systems.*
Kalman Filter (KF) (cont’d)

\[
\begin{align*}
x_k &= A x_{k-1} + B u_k + w_{k-1} \\
y_k &= H x_k + v_k
\end{align*}
\]

**Figure:** System, measurement model, and discrete-time Kalman filter, with \( \hat{x}_k(-) = A \hat{x}_{k-1}(+) \)
KF equations

\[ w \sim \mathcal{N}(0, Q) \quad \text{and} \quad v \sim \mathcal{N}(0, R) \]

**Prediction (Time Update)**

\[
\hat{x}_{k|k-1} = A x_{k-1|k-1} + B u_k
\]

\[
P_{k|k-1} = A P_{k-1|k-1} A^T + Q
\]

**Correction (Measurement Update)**

\[
\mathcal{E}_k = z_k - H \hat{x}_{k|k-1}
\]

\[
S_k = H P_{k-1|k-1} H^T + R
\]

\[
K_k = P_{k|k-1} H^T S_k^{-1}
\]

\[
\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k \mathcal{E}_k
\]

\[
P_{k|k} = (I - K_k H) P_{k|k-1}
\]
The performance of the Kalman filter degrades when the observed data contains outliers.

- Addressing the sensitivity of the square error criterion to outliers make the KF more robust to outlier.
- It is possible to model the observation and state noise as non-Gaussian, heavy-tailed distributions to account for non-Gaussian noise and outliers.
- Another class of methods uses a weighted least squares approach, where the measurement residual error is assigned some statistical.
- The required tuning of threshold parameters for optimal performance, maybe difficult to assure correct or accurate estimates because of bad choices of the weights, and thus, may lead to deteriorated performance property.
Problem Statement

- System model:

\[ x_{t+1} = A x_t + B u_t + w_t \]
\[ y_t = C x_t + v_t \]

where \( t = 0, 1, \ldots \), \( x_t \in \mathbb{R}^n \) is the state vector, \( u_t \in \mathbb{R}^m \) is the control vector, \( w_t \in \mathbb{R}^n \) is the system noise vector, \( y_t \in \mathbb{R} \) is the scalar output available at time \( t \), and \( v_t \in \mathbb{R} \) is the measurement noise.

- Additional information:
  - the initial state \( x_0 \) is Gaussian with mean \( \bar{x}_0 \) and covariance \( P_0 = P_0^\top > 0 \)
  - \( w_t \) is zero-mean Gaussian with covariance \( Q = Q^\top > 0 \)
  - \( x_0, \{w_t\}, \) and \( \{v_t\} \) are assumed to be uncorrelated
  - \( v_t \) is zero-mean Gaussian with covariance \( r > 0 \) except in case of outliers that occur in an unpredictable way and have covariance that is unknown and much larger than \( r \).
Maximum Likelihood Estimation With Measurements Affected by Outliers

- At each $t$ the state estimate is obtained by maximizing the conditional probability density w.r.t. $x_0, \ldots, x_t$ with given measures $y_0, \ldots, y_t$, while rejecting too large noises that might be outliers.

- The decision of rejection can be taken on the basis of a check on the residual error as compared with a suitable threshold $\sigma > 0$.

- Problem formulation:

\[
\begin{align*}
\min_{x_0^t, w_0^t, v_0^t} & \quad (x_0 - \bar{x}_0)^\top P_0^{-1} (x_0 - \bar{x}_0) + \sum_{i=0}^{t} v_i^2 / r \\
& \quad + \begin{cases} 
0, & \text{for } t = 0 \\
\sum_{i=0}^{t-1} w_i^\top Q^{-1} w_i, & \text{for } t = 1, 2, \ldots
\end{cases}
\end{align*}
\]
s.t.
\[ x_{i+1} = A x_i + B u_i + w_i, \ i = 0, 1, \ldots, t - 1 \]
\[ y_i = C x_i + v_i, \ i = 0, 1, \ldots, t \]
\[ -\sigma \leq v_i, \ i = 0, 1, \ldots, t \]
\[ v_i \leq \sigma, \ i = 0, 1, \ldots, t. \]

Unfortunately, the solution of this problem entails the exploration of the combinations of all the active sets due to the inequality constraints, and hence it is too computationally demanding.

To overcome the computational difficulty, a suboptimal approach based on the idea of performing only a one-step ahead optimization at each time \( t \) can be followed that results into a Kalman filter with covariance update driven by a threshold check.
MHE with Measurements affected by Outliers

- System model:

\[
x_{t+1} = Ax_t + Bu_t + w_t \\
y_t = Cx_t + v_t
\]

where \( t = 0, 1, \ldots \).

- The measurement noise is “small” except on rare occurrences: there exists \( r_v > 0 \) such that \( |v_t| \leq r_v, \ t = 0, 1, \ldots \) with \( t \neq \bar{t}_i \), where \( \bar{t}_i \in \mathbb{N}_{>0}, \ i = 0, 1, \ldots \), is a strictly increasing sequence such that \( \inf_{i \geq 0} (\bar{t}_{i+1} - \bar{t}_i) > 0 \) and \( |v_{\bar{t}_i}| \gg r_v \).

- The abnormal measurement disturbances at the instants \( \bar{t}_i \) are assumed to be bounded: there exists \( M \gg r_v \) such that, for all \( i = 0, 1, \ldots \), \( |v_{\bar{t}_i}| \leq M \).

- The system noise is “small” as compared with the dynamics, i.e., bounded and taking zero or around zero values: there exists \( r_w > 0 \) such that, for all \( t = 0, 1, \ldots \), \( \|w_t\| \leq r_w \).
Only information obtained in the recent past is used, i.e.,

\[ y_{t-N}, \ldots, y_t, u_{t-N}, \ldots, u_{t-1}. \]

We aim at estimating \( x_{t-N}, \ldots, x_t \) on the basis of such information and of a “prediction” \( \tilde{x}_{t-N} \) of the state \( x_{t-N} \) at the beginning of the moving window.

The estimates of \( x_{t-N}, \ldots, x_t \) at time \( t \) are denoted by \( \hat{x}_{t-N|t}, \ldots, \hat{x}_t|t \), respectively.

In principle we can deal with an arbitrary number of outliers, but here we restrict to the case of at most only one measurement affected by outlier in the batch of measures included in sliding window:

**Assumption**

The system outputs are such that

\[
\inf_{i \geq 0} (\tilde{t}_{i+1} - \tilde{t}_i) > N + 1
\]

for all \( i = 0, 1, \ldots \).
MHE with Measurements affected by Outliers (cont’d)

\[
\begin{align*}
Y_{t-N} & \quad Y_{t-N+1} & \cdots & \quad Y_t \\
\times & \quad \times & \cdots & \quad \times \\
Y_{t-N} & \quad Y_{t-N+1} & \cdots & \quad Y_t \\
\vdots \\
Y_{t-N} & \quad Y_{t-N+1} & \cdots & \quad Y_t
\end{align*}
\]

\[
\begin{align*}
J_t^0 & \quad \rightarrow J_t^1 \\
J_t^1 & \quad \rightarrow J_t^2 \\
& \vdots \\
J_t^N+1 & \quad \rightarrow \hat{x}_{t-N}
\end{align*}
\]

\[\min_{k=0,1,\ldots,N+1} J_t^k\]

**Figure:** Leaving-out Strategy
MHE with Measurements affected by Outliers (cont’d)

If an outlier corrupts the \( k \)-th measure of the batch 1, 2, \ldots, \( N+1 \), a least-squares cost function that leaves out such a measure is

\[
J^k_t (x_{t-N}) = \mu \| x_{t-N} - \bar{x}_{t-N} \|^2 + \frac{1}{N} \sum_{\substack{i=t-N \atop i \neq t-N+k-1}}^{t} (y_i - C x_i)^2
\]

where \( \mu \geq 0 \) and \( k = 1, 2, \ldots, N+1 \).

The cost is to be minimized together with the constraints

\[
x_{i+1} = A x_i + B u_i, \ i = t-N, \ldots, t-1.
\]

The case in which no outlier affects the measurements is accounted by using all the measures:

\[
J^0_t (x_{t-N}) = \mu \| x_{t-N} - \bar{x}_{t-N} \|^2 + \frac{1}{N+1} \sum_{i=t-N}^{t} (y_i - C x_i)^2
\]

to be minimized under the constraints given by the dynamics.
At each time $t = N, N + 1, \ldots$ we have to solve $N + 2$ problems as follows:

$$\min_{x \in \mathbb{R}^n \text{ s.t. } x_{i+1} = A x_i + B u_i} \quad k = 0, 1, \ldots, N + 1$$

and compare the optimal costs: the best of such costs is associated with the estimate. Thus, the estimate of $x_{t-N}$ at time $t$ is

$$\hat{x}_{t-N|t} = \hat{x}_{t-N}$$

where

$$k(t)^* \in \arg\min_{k=0,1,\ldots,N+1} J^k_t \left( \hat{x}^k_{t-N} \right)$$

and

$$\hat{x}^k_{t-N} \in \arg\min_{x \in \mathbb{R}^n \text{ s.t. } x_{i+1} = A x_i + B u_i} J^k_t (x).$$
Suppose that the matrix $F_k$ is of full rank for $k = 0, 1, \ldots, N + 1$ where

$$F_0 = \begin{pmatrix} C \\ CA \\ \vdots \\ CA^N \end{pmatrix}, \quad H_0 = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ C & 0 & \cdots & 0 \\ CA & C & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ CA^{N-1} & CA^{N-2} & \cdots & C \end{pmatrix}.$$ 

and $F_k$ and $H_k$ for $k \neq 0$ obtained from $F_0$ and $H_0$ by deleting the $k$-th block row, respectively. Moreover, let

$$\delta = \min_{k=0,1,\ldots,N+1} \lambda \left( F_k^T F_k \right) > 0.$$ 

Then the sequence $\{\zeta_t\}$ given by
Theorem (cont’d)

\[
\zeta_0 = \frac{2}{\mu + \frac{\delta}{N+1}} \left( \mu \|x_0 - \bar{x}_0\|^2 + c \right)
\]

\[
\zeta_{t+1} = a(\mu) \zeta_t + b(\mu) , \ t = 0, 1, \ldots
\]

is such that \( \|e_{t-N}\|^2 \leq \zeta_t \) for \( t = N, N + 1, \ldots \), where

\[
a(\mu) = \frac{8\mu \|A\|^2}{\mu + \frac{\delta}{N+1}} \quad b(\mu) = \frac{8\mu r_w^2 + 2c}{\mu + \frac{\delta}{N+1}}
\]

\[
c = \frac{2(2N + 1)}{N(N+1)} \left( N^2 r_w^2 \max_{k=0,1,\ldots,N+1} \|H_k\|^2 + (Nrv + M)^2 \right).
\]

If \( \mu \) is chosen such that \( a(\mu) < 1 \), the sequence \( \{\zeta_t\} \) converges to \( b(\mu)/(1 - a(\mu)) \), and is strictly decreasing if \( \zeta_0 > b(\mu)/(1 - a(\mu)) \).
Second-order oscillator with damping ratio $\xi$, (undamped) natural pulsation $\omega$ and measurements of only the first state variable:

$$A = \begin{pmatrix} 1 & T \\ -T\omega^2 & -2\omega\xi T + 1 \end{pmatrix}, \quad C = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

and $T > 0$ is the sample time.

Root mean square error:

$$RMSE(t) = \left( \frac{1}{M} \sum_{i=1}^{M} \frac{\|e_{t,i}\|^2}{M} \right)^{1/2}$$

where $e_{t,i}$ is the estimation error at time $t$ in the $i$-th simulation run, and $M$ is the number of simulation runs.

Outliers occur at fixed time instants.

MHE with $\mu = 0.6$ and a Kalman filter (KF) with recursive estimate update based on residual check with thresholds $\sigma_t \times \text{"integer number"}$.
Figure: $x_1$, its measure, and estimates with 100 simulation runs.
Simulation result (cont’d)

Table: Means of RMSEs over 100 runs

<table>
<thead>
<tr>
<th>$r$</th>
<th>MHE</th>
<th>KF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\mu = 0.1$</td>
<td>$\mu = 1$</td>
</tr>
<tr>
<td>0.01</td>
<td>0.0135</td>
<td>0.0160</td>
</tr>
<tr>
<td>0.1</td>
<td>0.1229</td>
<td>0.1461</td>
</tr>
<tr>
<td>1.0</td>
<td>1.2435</td>
<td>1.4774</td>
</tr>
</tbody>
</table>

Table: Means of computational time in seconds over 100 runs

<table>
<thead>
<tr>
<th>$r$</th>
<th>MHE</th>
<th>KF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\mu = 0.1$</td>
<td>$\mu = 1$</td>
</tr>
<tr>
<td>0.01</td>
<td>0.4260</td>
<td>0.4195</td>
</tr>
<tr>
<td>0.1</td>
<td>0.4153</td>
<td>0.4682</td>
</tr>
<tr>
<td>1.0</td>
<td>0.4203</td>
<td>0.4340</td>
</tr>
</tbody>
</table>
Figure: RMSEs of MHF and KF over 100 simulation runs for different choices of $\mu$ and $\sigma_t$. 
**$\ell_2$-Norm MHE advantages**

1. Least squares is intuitive and simple to implement
2. Model constrained can be added to model to improve the estimation accuracy.
3. Improving performances by an optimal tuning of the parameter $\mu$.

With the proposed method we can add

4. **Robustness to outliers** which it was hidden before.

**Challenge**

Efficient solution of the MHE approach is important for solving large-scale problems of industrial significant. Because of increase in data availability in the industry, advances in technology, improved networking, and the need of additional monitoring.
Motivations of nonlinear Moving-Horizon Estimation

- Investing the success of MHE in the state estimation and process monitoring for linear and nonlinear systems, efficient solution of MHE approach is important for solving the problem of outlier in industrial systems.

- MHE outperforms the Extended Kalman Filter (EKF) in the presence of constraints.

- Recent approaches for state estimation in model predictive control, employ MH estimator for estimation of the current state.

Robust estimator for nonlinear control systems in case of outlier occurrences is needed.
Problem Statement

- We address the estimation of the state vector adopting a MH strategy for nonlinear systems when the measurements are affected by outliers.

- Let us consider a dynamic system described by the discrete-time equations

\[
\begin{align*}
x_{t+1} &= f(x_t, u_t) + \xi_t \\
y_t &= h(x_t) + \eta_t
\end{align*}
\]

\(\xi_t \in \mathbb{R}^n\) is the system noise vector, and \(\eta_t \in \mathbb{R}^p\) is the measurement noise. Such a noise is assumed to be small except on rare occurrences, likewise in the linear case.
We use the same methodology of the linear part with some changes on the definitions and additive assumptions to meet the nonlinear approach.

Let us denote the set of admissible controls by $U$ and the sets from which the vectors $\xi_t$ and $\eta_t$ take their values by $\Xi$ and $H$, respectively.

The following assumptions are needed:

A1. $\Xi$, $H$, and $U$ are compact set.
A2. The initial state $x_0$ and the control sequence $\{u_t\}$ are such that, for any possible sequence of disturbances $\{\xi_t\}$, the system trajectory $\{x_t\}$ lies in a compact set $X$. Let $X'$ be the closed convex hull of the set $X$.
A3. The functions $f$ and $h$ are $C^2$ functions with respect to $x$ on $X'$ for every $u \in U$. 
Least-squares costs function

If an outlier corrupts the $k$-th measure of the batch $1, 2, \ldots, N + 1$

$$J^k_t (x_{t-N}) = \mu \| x_{t-N} - \bar{x}_{t-N} \|^2 + \frac{1}{N} \sum_{i=t-N+1}^{t} (y_i - h(x_i))^2$$

where $\mu \geq 0$ and $k = 1, 2, \ldots, N + 1$. to be minimized together with the constraints

$$x_{i+1} = f(x_i, u_i), \ i = t - N, \ldots, t - 1.$$ 

If no outlier effects the measurements of the batch

$$J^0_t (x_{t-N}) = \mu \| x_{t-N} - \bar{x}_{t-N} \|^2 + \frac{1}{N+1} \sum_{i=t-N}^{t} (y_i - h(x_i))^2$$
At each time \( t = N, N + 1, \ldots \) we have to solve \( N + 2 \) problems as follows:

\[
\min_{x \in \mathbb{R}^n} J^k_t (x), \quad k = 0, 1, \ldots, N + 1
\]

\[\text{s.t. } x_{i+1} = f(x_i, u_i), \quad i = t - N, \ldots, t - 1 \text{ hold}\]

and compare the optimal costs: the best of such costs is associated with the estimate. Thus, the estimate of \( x_{t-N} \) at time \( t \) is

\[
\hat{x}_{t-N|t} = \hat{x}^{k(t)*}_{t-N}
\]

where

\[
k(t)* \in \arg\min_{k=0,1,\ldots,N+1} J^k_t \left( \hat{x}^k_{t-N} \right)
\]

and

\[
\hat{x}^k_{t-N} \in \arg\min_{x \in \mathbb{R}^n} J^k_t (x) \quad \text{s.t. } x_{i+1} = f(x_i, u_i), \quad i = t - N, \ldots, t - 1 \text{ hold}\]
If we consider the collections of measurements at time steps $t - N, t - N + 1, \ldots, t$, it is straightforward to obtain

$$y_{t-N}^t|k = F_k(x_{t-N}) + D_{\xi_k}(x_{t-N}) \xi_{t-N}^{t-1}|k + \eta_{t-N}^t|k$$

for $k = 0, 1, \ldots, N + 1$, with $F_k$ and $D_{\xi_k}$ for $k \neq 0$ obtained from $F$ and $D$ by deleting the $k$-th block row, respectively, where

$$F \left( x_{t-N}, u_{t-N}^{t-1} \right) \triangleq \begin{bmatrix}
    h \left( x_{t-N} \right) \\
    h \circ f_{u_{t-N}}^{t-1} \left( x_{t-N} \right) \\
    \vdots \\
    h \circ f_{u_{t-1}} \circ \cdots \circ f_{u_{t-N}} \left( x_{t-N} \right)
\end{bmatrix}$$

for $t = N, N + 1, \ldots$, where $u_{t-N}^{t-1} \triangleq \text{col}(u_{t-N}, \ldots, u_{t-1})$, “$\circ$” denotes function composition, and $f_{u_i}(x_i) \triangleq f \left( x_i, u_i \right)$. 
Stability of MH estimator (cont’d)

\[ D_\xi(x_{t-N}, u_{t-1}^{t-N}) \triangleq \begin{bmatrix} \frac{\partial h_\circ f(x_{t-N})}{\partial \xi_{t-N}} & 0 & \cdots & 0 \\ \frac{\partial h_\circ f(x_{t-N})}{\partial \xi_{t-N+1}} & \frac{\partial h_\circ f(x_{t-N})}{\partial \xi_{t-N+2}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial h_\circ f(x_{t-N})}{\partial \xi_{t-N}} & \frac{\partial h_\circ f(x_{t-N})}{\partial \xi_{t-N+1}} & \cdots & \frac{\partial h_\circ f(x_{t-N})}{\partial \xi_{t-1}} \end{bmatrix} \]

\[ \Delta_\xi \triangleq \max_{x_{t-N} \in X, u_{t-1}^{t-N} \in U^N} \| D_\xi(x_{t-N}, u_{t-1}^{t-N}) \|, \]

\[ r_\xi \triangleq \max_{\xi \in \Xi} \| \xi \|, \quad r_\eta \triangleq \max_{\eta \in H} \| \eta \|, \quad d_x \triangleq \max_{x_0, \bar{x}_0 \in X} \| x_0 - \bar{x}_0 \|. \]
Stability of MH estimator (cont’d)

**Definition**

Given a set \( S \subseteq \mathbb{R}^n \), the system is said to be \( S \)-observable in \( N + 1 \) steps if there exists a \( K \)-function \( \varphi(\cdot) \), such that

\[
\varphi\left( \| x_1 - x_2 \|^2 \right) \leq \| F(x_1, \bar{u}) - F(x_2, \bar{u}) \|^2 \\
\forall x_1, x_2 \in S, \quad \forall \bar{u} \in U^N.
\]

The system is \( \mathcal{X} \)-observable in \( N + 1 \) steps with a \( K \)-function \( \varphi \). \( F_k \) is of full rank. Suppose that the \( K \)-function \( \varphi \) satisfies the following condition

\[
\delta \triangleq \inf_{x_1, x_2 \in \mathcal{X}; \; x_1 \neq x_2} \frac{\varphi\left( \| x_1 - x_2 \|^2 \right)}{\| x_1 - x_2 \|^2} > 0.
\]

Then the square norm of the estimation error is bounded as

\[
\| \varepsilon_{t-N} \|^2 \leq \zeta_{t-N}
\]
Stability of MH estimator (cont’d)

Definition (cont’d)

where \( \{\zeta_t\} \) is a sequence generated by

\[
\begin{align*}
\zeta_0 &= \beta_0 \\
\zeta_{t+1} &= \alpha \zeta_t + \beta, \quad t = 0, 1, \ldots
\end{align*}
\]

\[
\alpha \triangleq \frac{8 k_f^2 \mu}{\mu + \frac{\delta}{N+1}},
\]

\[
\beta \triangleq \frac{2}{\mu + \frac{\delta}{N+1}} \left( 4 \mu r_\xi^2 + c \right),
\]

\[
\beta_0 \triangleq \frac{2}{\mu + \frac{\delta}{N+1}} \left( \mu d_x^2 + c \right)
\]

\[
c \triangleq \frac{2(2N + 1)}{(N + 1)} \left( \Delta_{\xi_k}^2 N^2 r_\xi^2 + (N r_\eta + M)^2 + ||\Omega||^2 \right)
\]

\( k_f \) is an upper bound on the Lipschitz constant of \( f(x, u) \)
We consider the problem of estimating the state variables of a second-order oscillating system by using only the measures of the first variable, without knowledge of the damping coefficient $\zeta$.

The discrete nonlinear system and linear observation equations

\begin{align*}
    x_1^{t+1} &= x_1^t + x_2^t + \xi_1^t \\
    x_2^{t+1} &= -\omega^2 T x_1^t + (1 - 2\omega^2 T) x_2^t + \xi_2^t \\
    x_3^{t+1} &= x_3^t + \xi_3^t, \quad T > 0 \\
    y_t &= x_1^t + \eta_t
\end{align*}

Where $\xi_t \in \mathbb{R}^n$ is the system noise vector, and $\eta_t \in \mathbb{R}^p$ is the measurement noise.

Outliers are randomly positioned over 100 time steps.

EKF with recursive estimation update based on residual check.
Simulation result

Figure: $x_1$, its measures, and estimates

MHF with $\mu = 0.5$ and EKF with $\sigma_t = \sqrt{S_t}$

Outliers are randomly generated and positioned over 100 time steps
The distributions of the noises are taken according to the linear example except in case of outlier occurrences. All the state variables are affected by $\xi_t$. 

**Figure:** True value and estimates of the states
Figure: RMSEs of MHF and EKF over 100 simulation runs
Simulation result (cont’d)

Table: Means of RMSEs for MHF with $\mu = 0.5$ and EKF with $\sigma_t = \sqrt{S_t}$

<table>
<thead>
<tr>
<th>$r$</th>
<th>MHE</th>
<th>KF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$x_1$</td>
<td>$x_2$</td>
</tr>
<tr>
<td>0.01</td>
<td>0.1760</td>
<td>0.7584</td>
</tr>
<tr>
<td>0.10</td>
<td>0.2124</td>
<td>0.7262</td>
</tr>
<tr>
<td>1.00</td>
<td>0.5125</td>
<td>1.6230</td>
</tr>
</tbody>
</table>

Table: Means of Computational time in seconds over 100 runs with different $r$

<table>
<thead>
<tr>
<th>$r$</th>
<th>MHF $\mu = 0.5$</th>
<th>KF $\sigma_t = \sqrt{S_t}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>6.4228</td>
<td>0.0456</td>
</tr>
<tr>
<td>0.1</td>
<td>6.5228</td>
<td>0.0674</td>
</tr>
<tr>
<td>1.0</td>
<td>7.5520</td>
<td>0.0833</td>
</tr>
</tbody>
</table>
The results of linear MHE are presented in


Nonlinear MHE is treated in my Ph.D thesis.

Further reading:

Alessandri et al. (2008)
"Moving-Horizon state estimation for nonlinear discrete-time systems: New stability results and approximation schemes"

Christopher V. Rao, James B. Rawlings (2012).
"Constrained Process Monitoring: Moving-Horizon Approach".

Hans-Peter Kriegel et al. 2010
"Outlier Detection Techniques".
*SIAM International Conference on Data Mining, SDM, 2010.*
Conclusions

- State estimation for discrete-time systems with measurements affected by outliers:
  - moving horizon
  - generalized least-squares cost
  - "best optimal cost" estimation criterion.

- Stability of the estimation error provided by the moving-horizon estimator:
  - mild assumptions
  - rigorous proof.

- Comparison with the Kalman filter:
  - threshold selection
  - computational effort.

- On-going work:
  - reduction of the computational burden
  - rigorous proof of robustness
  - applications to industrial systems.
The End