2017 11th Asian Control Conference (ASCC)
Gold Coast Convention Centre, Australia
December 17-20, 2017

Resilient Moving Horizon Estimation for Cyber-Physical Systems under Sensor Attacks

Yulei Wang[1], Jingxin Yuan[1], Shuyou Yu[1], Yufeng Hu[1], Hong Chen[1,*]

Abstract— **This paper addresses the problem of state estimation for linear dynamic cyber-physical systems (CPS) that is resilient against malicious integrity attacks on sensors. A resilient moving-horizon estimation (MHE) scheme is proposed to correctly estimate the states under sensor attacks by exploiting sensor redundancy, and it is optimal with a guarantee of prior knowledge in the form of both state and disturbance constraints. In this framework, the problem is formulated as a multistage optimal control problem from the perspective of probability theory. Then, it is solved by a special kind of optimization, the bi-level optimization, where the upper-level optimization task responds to the optimal state estimation, while the lower-level optimization task excludes the compromised sensors. Moreover, the strategy to reduce the computational burden is to develop a moving horizon approximation that has been used successfully to develop stabilizing estimation strategy. Numerical simulation is provided to illustrate the performance of the proposed state estimation scheme.**

I. INTRODUCTION

Recent years, automobile industry has witnessed a significant increase in the number of security-related reports on electronic control systems (ECS) [1], [2]. These high-profile attacks are in a wide range of ECS, from attacks on global positioning system (GPS) [3], inertial measurement units (accelerometers and gyroscopes) [4], light detection and ranging sensors [5], to attacks on engine control systems [6], electronic brake control systems [4] and keyless entry and start systems [7]. Various cyber security and communication cryptographic approaches have been proposed to prevent intrusions (see, e.g., [8] and the references therein), but an adversary can still affect the vehicle control systems via the computational nodes (e.g., ECU), communication networks and physical sensors (e.g., GPS), since resource constraints inherent in vehicle real-time control may prevent heavy-duty security approaches from being applied [9].

In this paper, a general problem is addressed that consists in estimating the state variables of a linear dynamic system by means of measures possibly attacked by an adversary. The estimation is performed by using a moving horizon estimation (MHE) approach, which will be set in such a way to make it resilient to sensor attacks. The first idea about what is currently denoted as MHE is presented in [10]. MHE determines state estimate online by solving a finite horizon optimization problem [11], [12]. As new measurements become available, the old measurements are discarded from the estimation window, and the finite horizon state problem is resolved to determine the update estimate of the state. The method is optimization based, so one may check abnormal measurements (caused by faults and/or attacks) via hardware or analytical redundancy [13], and handle explicitly state estimation with inequality constraints on the decision variables.

The problem of state estimation in the presence of sensor attacks has attracted more attention recent years [14]-[15]. Different from sensor faults or outliers, an intelligent adversary can invalidate the fault detection and isolation (FDI) systems and launch an attack with a number of compromised sensors to destroy the estimation and control performance [14], [16], [17]. For deterministic linear systems, the secure state estimation in the presence of sensor attacks can be obtained as the l_0 optimization problem [18]. They proved that if the attacker can manipulate less than half the measurements it is possible to accurately reconstruct the state variables of a system despite attacks. The similar conclusion was indicated in [19] as well. For stochastic systems, Pajic et al. [20], [21] proposed an l_0-norm moving horizon approach, in which the estimator will use the measurements from time $k-T+1$ to time k to estimate the current state $x(k)$ with a bound for the state estimation error. However, the measuring data before time $k-T$ are discarded in the l_0-norm-based state estimator, which may result in a degradation of the estimation performance. In addition, the idea of employing a bank of observers is developed for detecting compromised sensors and estimating states, which includes Luenberger observer [22], Kalman filter [23], high-gain observer [24], event-trigger observer [15], et al.

In this paper, we focus on the estimation problem for linear discrete-time systems with less than half measurements attacked by adversary. First we will propose a new framework of resilient moving-horizon estimation. Different from the preliminary moving horizon results [18], [20], [21], the full information estimate of state is utilized to obtain the optimal estimation performance, and the estimation is equivalent to a multistage optimal control problem from the perspective of probability theory. A bi-level optimization scheme is introduced to provide the solution as a two-step strategy, each of which can be formulated as a feasible programming problem. Furthermore, the strategy to reduce the computational burden is developed by a moving horizon approximation that has been used successfully to develop stabilizing estimation strategy. Numerical simulation is provided to illustrate the

*This work was supported in part by the National Natural Science Foundation of China under Grant 61603147, in part by China Postdoctoral Science Foundation under Grant 2014M561291, and in part by Department of Science and Technology of Jilin Province under Grant 20160520107JH.

[1]Y. Wang, et al. are with Faculty of the State Key Laboratory, Automotive Simulation and Control as well as Department of Control Science and Engineering, Jilin University, Changchun 130025, China (corresponding author: Hong Chen, e-mail: chenh@jlu.edu.cn)

performance of the proposed state estimation scheme.

The paper is organized as follows. Section II introduces the notations used throughout the paper and the necessary assumptions. In Section III, the proposed resilient MHE approach is described, where both the optimal property and the bi-level optimization solution are studied. Furthermore, we investigate a finite resilient MHE with moving horizon approximation, and sum up an algorithm. Finally, simulation results are given in Section IV and we provide concluding remarks in Section V

II. PRELIMINARIES

The denotations in this paper is as follows. $\mathbb{R}$ is used to denote the set of reals and the ith element of a vector x_k is denoted by $x_{k,i}$. For vector x and matrix A, $|x|$ and $|A|$ denote the vector and matrix whose elements are absolute values of the initial vector and matrix, respectively. For a vector v, $\|v\|_P^2 := v^T P v$ denotes its generalized Euclidean norm with symmetric positive definite matrix $P = P^T > 0$.For a vector $f \in \mathbb{R}^m$, the support of the vector is set $\text{supp}(f) = \{i | f_i \neq 0\} \subseteq \{1, 2, \ldots, m\}$, while the l_0 norm of vector f is the size of $\text{supp}(f)$, which implies $\|f\|_{l_0} = |\text{supp}(f)|$.

In this paper, we investigate online optimization strategies for estimating the state of compromised systems modeled by a linear time-invariant (LTI) system of the form

$$\begin{aligned} x_{k+1} &= Ax_k + w_k \\ y_k &= Cx_k + v_k + Ee_k \end{aligned} \tag{1}$$

where $x_k \in \mathbb{R}^n$ denotes the plant's state vector at time k, respectively, while $y_k \in \mathbb{R}^m$ describes the plant's output vector obtained from measurements of m sensors. $w_k \in \mathbb{R}^n$ and $v_k \in \mathbb{R}^m$ denote the process and measurement disturbance vector, respectively. Accordingly, $A \in \mathbb{R}^{n\times n}$, $C \in \mathbb{R}^{m\times n}$ are the state and output matrices, respectively. The matrix $E \in \mathbb{R}^{m\times p}$ denotes the projection from the attack $e_k \in \mathbb{R}^p$ to the sensor y_k. Given $m > p$, we have an SVD decomposition

$$E = \begin{bmatrix} U_1^T & U_2^T \end{bmatrix} \begin{bmatrix} I_p \\ 0_{m-p} \end{bmatrix} \tag{2}$$

such that $U_2 E = 0$ with $U_2 \in \mathbb{R}^{(m-p)\times m}$.

It is known that the state and disturbances satisfy the following constraints:

$$x_k \in \mathbb{X} \subseteq \mathbb{R}^n,\ w_k \in \mathbb{W} \subseteq \mathbb{R}^n,\ v_k \in \mathbb{V} \subseteq \mathbb{R}^m \tag{3}$$

where the constraints $\mathbb{W}$ and $\mathbb{V}$ are interpreted as a strategy for modeling bounded disturbances or random variables with truncated densities, and the state constraint $\mathbb{X}$ may be used to account for model inaccuracies or physical properties.

Let $x_k(z_l, \{w_j\})$ denote the state solution of the system (1) at time k when the initial state is z at time l and the state disturbance sequence is $\{w_j\}_{j=l}^k$, and let the vector $y_k(z_l, \{w_j\})$ denote the predicted output at time k when the initial condition at time l is z and the disturbance sequence is $\{w_j\}_{j=l}^k$.

We now introduce assumptions for the system (1).

Assumption 1: The dimension of e_k satisfies $2p < m$.

Assumption 2: The attack matrix E is unknown for the resilient estimator design, but the pair (U_2C, A) is observable.

Assumption 3: Disturbances w_k, v_k, initial state x_0 and attack vector e_k are independent.

Assumption 4: The priori estimate of initial state x_0 follows a normal distribution with mean $\bar{x}_0$ and variance P_0^{-1}.

Assumption 5: The disturbance sequences $\{w_k\}$ and $\{v_k\}$ are normal distributions with zero-mean and variances Q^{-1} and R^{-1}, respectively.

Assumption 1 states that strictly less than half of all the sensors in the system may be under integrity attack. This is a standard assumption for resilient state estimation [18], [19]. The rationale is that the adversaries who attack the sensors have limited resource only enough to compromise a subset of the sensors. Assumption 2 indicates that the form of attacks is unknown a prior, which is the main difference from faults, while the observability ensures that an estimator can be constructed. From system design point of view, one can select sensors that satisfy Assumption 2. In addition, Assumptions 3-5 are well-established for MHE to yield the exact conditional probability estimate in the special case that all errors are Gaussian-distributed.

III. SECURE MOVING HORIZON ESTIMATION

A. Full Information Estimate Problem

Notice that moving horizon estimation is an optimization approach that uses a series of measurements observed over time, containing noise (random variations) and other inaccuracies, and produces estimates of unknown variables or parameters. Unlike deterministic approaches like the Kalman filter, MHE requires an iterative approach that relies on linear programming or nonlinear programming solvers to find a solution. Unfortunately, the conventional cost function is not feasible in case of measurements compromised by attacks. To resilient estimation, the straightforward idea is that at each iteration we minimize a least-squares cost function, where the measurements that can be affected by attackers are left out.

Hence we formulate the resilient MHE problem, for $T \geq 0$, as the solution to the following optimal control problem:

$$\mathcal{P}_1(T):\ \Phi_T^* = \min_{x_0, \{w_k\}_{k=0}^{T-1}, U_2} \Phi_T\left(\{w_k\}_{k=0}^{T-1}, x_0, U_2\right) \tag{4}$$

subject to the the dynamic constraints

$$\hat{x}_{k+1} = A\hat{x}_k + Bw_k, \quad k \in [0, T-1] \tag{5}$$

and the time-domain constraints

$$\hat{x}_k \in \mathbb{X},\ k \in [0, T];\ w_k \in \mathbb{W},\ k \in [0, T-1] \tag{6}$$

$$y_k^c - U_2 C\hat{x}_k \in \mathbb{V}^c,\ k \in [0, T-1] \tag{7}$$

$$y_k^c = U_2 y_k,\ \mathbb{V}^c = \{v_k^c | v_k^c = U_2 v_k, v_k \in \mathbb{V}\} \tag{8}$$

where $R_c = U_2 R U_2^T$ and

$$\Phi_T\left(\{w_k\}_{k=0}^{T-1}, x_0, U_2\right) = \sum_{k=0}^{T-1} \|y_k^c - U_2 C\hat{x}_k\|_{R_c}^2 + \|w_k\|_Q^2 + \|x_0 - \bar{x}_0\|_P^2 \quad (9)$$

Different from the conventional MHE, the solution to the problem $\mathcal{P}_1(T)$ at time T is the triple

$$\left(\hat{x}_{0|T-1},\ \{\hat{w}_{k|T-1}\}_{k=0}^{T-1}, \hat{U}_{2|T-1}\right)$$

and the optimal triple yields an estimate $\{\hat{x}_{k|T-1}\}_{k=0}^{T-1}$ of the actual sequence $\{x_k\}$ from the iterative solution of (5) with the initial state $\hat{x}_{0|T-1}$ at time $k=0$ and disturbance sequence $\{\hat{w}_{k|T-1}\}_{k=0}^{T-1}$, i.e.,

$$\hat{x}_{k|T-1} := x_k\left(\hat{x}_{0|T-1}, \{\hat{w}_{k|T-1}\}_{k=0}^{T-1}\right) \quad (10)$$

Theorem 1: Given Assumptions 1-5 and the system measurements $\{y_k\}_{k=0}^{T-1}$ compromised by sensor attacks e_k, the solution of problem $\mathcal{P}_1(T)$ is the optimal moving horizon estimation.

Proof: To prove the optimal estimation, we formulate the state estimation problem $\mathcal{P}_1(T)$ from the perspective of probability theory. Due to sensor attack, the output signal y_k can be divided as $y_k^e = U_1 y_k = U_1 C x_k + U_1 v_k + e_k \in \mathbb{R}^p$ and $y_k^c = U_2 y_k = U_2 C x_k + U_2 v_k \in \mathbb{R}^{m-p}$. Due to the attack e_k, the output y_k^e is independent of x_k. Given Assumption 2, the conditional probability density function of the state evolution $\{x_0, x_1, \ldots, x_T\}$ given the measurements $\{y_0, y_1, \ldots, y_{T-1}\}$ is equivalent to that given the measurements $\{y_0^c, y_1^c, \ldots, y_{T-1}^c\}$, i.e.,

$$p(x_0, x_1, \ldots, x_T | y_0, y_1, \ldots, y_{T-1}) = p(x_0, x_1, \ldots, x_T | y_0^c, y_1^c, \ldots, y_{T-1}^c) \quad (11)$$

The optimal estimate of the state $\hat{x}_{k|T-1}$ at time k is then a functional L_T of conditional probability density function (11):

$$\{\hat{x}_{0|T-1}, \hat{x}_{1|T-1}, \ldots, \hat{x}_{T|T-1}\} = L_T\left(p(x_0, x_1, \ldots, x_T | y_0^c, y_1^c, \ldots, y_{T-1}^c)\right)$$

A typical choice for the functional L_T is the maximum a posteriori Bayesian (MAP) estimate:

$$\{\hat{x}_{0|T-1}, \hat{x}_{1|T-1}, \ldots, \hat{x}_{T|T-1}\} \in \underset{x_0, x_1, \ldots, x_T}{\operatorname{argmax}}\ p(x_0, x_1, \ldots, x_T | y_0^c, y_1^c, \ldots, y_{T-1}^c)$$

Using the Markov property, we can express the joint probability of the state as

$$p(x_0, \ldots, x_T) = p_{x_0}(x_0) \prod_{k=0}^{T-1} p(x_{k+1} | x_k)$$

where $p_{x_0}(x_0)$ denotes the prior information of the initial state. According to Assumption 3 that v_k is independent, using the sensor model y_k^c we have the relationship

$$p(y_0^c, \ldots, y_T^c | x_0, \ldots, x_{T-1}) = \prod_{k=0}^{T_1} p_{v_k}\left(y_k^c - U_2 C x_k\right)$$

Applying Bayes's rule, the conditional probability density function (11) can be rewritten as

$$p(x_0, x_1, \ldots, x_T | y_0^c, y_1^c, \ldots, y_{T-1}^c) \propto p_{x_0}(x_0) \prod_{k=0}^{T-1} p_{v_k}\left(y_k^c - U_2 C x_k\right) p(x_{k+1} | x_k)$$

and with the properties of logarithms and Assumption 4 and 5, we have

$$\begin{aligned} &\underset{x_0, x_1, \ldots, x_T}{\operatorname{argmax}}\ p(x_0, x_1, \ldots, x_T | y_0, y_1, \ldots, y_{T-1}) \\ &= \underset{x_0, x_1, \ldots, x_T}{\operatorname{argmax}}\ p(x_0, x_1, \ldots, x_T | y_0^c, y_1^c, \ldots, y_{T-1}^c) \\ &= \underset{x_0, x_1, \ldots, x_T}{\operatorname{argmax}}\ \ln p(x_0, x_1, \ldots, x_T | y_0^c, y_1^c, \ldots, y_{T-1}^c) \\ &= \underset{x_0, x_1, \ldots, x_T}{\operatorname{argmax}} \sum_{k=0}^{T_1} \left[\ln p_{v_k}(y_k^c - U_2 C x_k) + \ln p(x_{k+1} | x_k)\right] \\ &\quad + \ln p_{x_0}(x_0) = \underset{x_0, x_1, \ldots, x_T}{\operatorname{argmin}}\ \Phi_T\left(\{w_i\}, x_0, U_2\right) \end{aligned}$$

which indicates the performance index (9) in the problem $\mathcal{P}_1(T)$ is optimal for the probability estimation, and the proof is thus completed. ■

B. Bi-Level Optimization Problem

In the previous content, a state estimation problem $\mathcal{P}_1(T)$ is formulated, however, solving this optimization online is still computationally heavy (NP-hard) due to the lack of the knowledge of E or U_2. In this subsection, we will propose a bi-level optimization strategy, which transforms the problem $\mathcal{P}_1(T)$ into two easy programming problems.

Note that the bilinear products of E and e_k make the estimation problem as a non-convex optimization. To deal with this problem, the concept of over-parameterization is applied by introducing a new variable

$$f_k = E e_k \in \mathbb{R}^{m \times p} \quad (12)$$

with a property of $\|f_k\|_{l_0} = p$. From (1) it follows that for $k = 0, 1, \ldots, T-1$

$$y_k = C A^k x_0 + f_k + C \sum_{i=0}^{k-1} A^{k-1-i} w_i + v_k \quad (13)$$

Since both disturbances w_k and v_k are bounded, then there exists a matrix $\Delta_T = [\delta_0, \delta_1, \ldots, \delta_{T-1}] \in \mathbb{R}^{m \times T}$ containing positive thresholds $\delta_{k,j} > 0$ with $k = 0, 1, \ldots, T-1$ and $j = 1, \ldots, m$ such that

$$|y_k - C A^k x_0 - f_k| \le |C| \sum_{i=0}^{k-1} |A^{k-1-i}||w_i| + |v_k| \le \delta_k$$

Referring to [20], [21], under Assumption 1, the following mixed-integer linear programming problem is formulated to obtain the l_0 norm of attack sequence:

$$\mathcal{P}_2(T): \Psi_T^* = \min \Psi_T\left(x_0, \{f_k\}_{k=0}^{T-1}, \gamma\right) = \min \sum_{i=1}^{m} \gamma_i \quad (14)$$

subject to

$$-\delta_k \le y_k - CA^k x_0 - f_k \le \delta_k,\ k = 0, \dots, T-1 \quad (15)$$

$$-\gamma_i \alpha \le f_{k,i} \le \gamma_i \alpha\ , i = 1, \dots, m \quad (16)$$

$$\gamma = \left[\ \gamma_1, \cdots, \gamma_m\ \right] \in \{0,1\}^m \quad (17)$$

where α is a sufficiently large positive constant. Consequently, the solution to $P_2(T)$ at time T is the triple

$$\left(\hat{x}_{0|T-1}, \{\hat{f}_{k|T-1}\}_{k=0}^{T-1}, \hat{\gamma}_{|T-1}\right)$$

and the optimal triple yields an estimate $\hat{U}_{2|T-1}$ of the actual matrix U_2 in the form of

$$\hat{U}_{2|T-1} = I_m / \hat{\gamma}_{|T-1} \in \mathbb{R}^{(m-p)\times m} \quad (18)$$

where $I_m/\hat{\gamma}_{|T-1}$ specifies a matrix by deleting the ith rows of the identity matrix $I_m \in \mathbb{R}^{m\times m}$ with respect to $\hat{\gamma}_{i|T-1} = 1$.

Together with the problem $\mathcal{P}_1(T)$ and $\mathcal{P}_2(T)$, the following bi-level optimal state estimation problem can be written as follows:

$$\mathcal{P}_3(T):\ \Phi_T^* = \min_{x_0, \{w_k\}_{k=0}^{T-1}, U_2} \Phi_T\left(\{w_k\}_{k=0}^{T-1}, x_0, U_2\right) \quad (19)$$

subject to (5)-(8) and

$$U_2 \in \underset{\gamma}{\operatorname{argmin}} \left\{\Psi_T\left(x_0, \{f_k\}_{k=0}^{T-1}, \gamma\right)\ s.t.\ (15)\text{-}(17)\right\} \quad (20)$$

In the problem $\mathcal{P}_3(T)$, the optimal estimation is divided into two levels: the upper-level optimization task responds to the optimal state estimation, while the lower-level optimization task excludes the compromised sensors. Note that the lower-level optimization is a mix-integer linear programming with NP-hard, but the upper-level one is the standard quadratic programming with linear constraints. Hence, the problem $\mathcal{P}_3(T)$ is better than $\mathcal{P}_1(T)$ for a solution.

In the following theorem, the optimal property of the problem $\mathcal{P}_3(T)$ is investigated.

Theorem 2: Given Assumptions 1-5 and the system measurements $\{y_k\}_{k=0}^{T-1}$ compromised by sensor attacks e_k, the solution of the problem $\mathcal{P}_3(T)$ is equivalent to that of the problem $\mathcal{P}_1(T)$.

Proof: From the problem $\mathcal{P}_3(T)$, the event of solving the optimal U_2^* in the lower-level optimization is independent of the initial state x_0 and the disturbance sequence $\{w_k\}_{k=0}^{T-1}$ of the upper-level optimization. It also means that two events of Φ_T and Ψ_T are independent, i.e.,

$$p(x_0, \{w_k\}_{k=0}^{T-1}, U_2 | \{y_k\}_{k=0}^{T-1}) = p(x_0, \{w_k\}_{k=0}^{T-1} | \{y_k\}_{k=0}^{T-1}) p(U_2 | \{y_k\}_{k=0}^{T-1}) \quad (21)$$

Based on Bellman's Principle of Optimality [25], we have the conclusion in Theorem 2. ∎

C. Finite Secure Moving Horizon Estimation Problem

The formulation of problem $\mathcal{P}_3(T)$ is referred to the full information problem and $\hat{x}_k$ is the full information estimate of x_k. This problem has T states and disturbances, so the computational complexity scales at least linearly with T. Due to the time-domain constraints and the mixed-integer programming with NP-hard, the conline solution of $\mathcal{P}_3(T)$ is impractical because the computational burden increases with time. To make the problem tractable, one needs to finite secure MHE. One strategy to reduce $\mathcal{P}_3(T)$ is a fixed-dimension optimal control problem is to employ a moving horizon approximation [11]. Unlike the full information problem, finite MHE estimates the truncated sequence $\{x_k\}_{k=T-N}^{T}$ where the step error $T - N > 0$ determines the length of the estimate window. The key to preserving stability and performance is how one approximately summarizes the past data.

First, it is straightforward that the lower-level optimization problem $\mathcal{P}_2(T)$ can be rewritten as the window $[T-N, T-1]$ form of

$$\mathcal{P}_2'(N):$$

$$\Psi_N^* = \min \Psi_N\left(x_{T-N}, \{f_k\}_{k=T-N}^{T-1}, \gamma\right) = \min \sum_{i=1}^{m} \gamma_i \quad (22)$$

subject to

$$\left|y_{T-N+k} - CA^k x_{T-N} - f_{T-N+k}\right| \le \delta_{T-N+k} \quad (23)$$

$$|f_{k,i}| \le \gamma^T \alpha,\ \gamma = \left[\ \gamma_1, \cdots, \gamma_m\ \right] \in \{0,1\}^m \quad (24)$$

for $k = 0, \dots, T-N$ and $i = 1, \dots, m$, which implies that the estimate of U_2 in the problem $\mathcal{P}_2'(N)$ depends only on the set of compromised sensors during the estimate window $[T-N, T-1]$. Accordingly, one can use the estimate of U_2 from problem $\mathcal{P}_2'(N)$ instead of the one from problem $\mathcal{P}_2(N)$ to reduce the computational cost. It should be emphasized that the iterative form of $\mathcal{P}_2'(N)$ makes $\mathcal{P}_2'(N)$ robust against the change of U_2.

Different from the lower-level optimization, the upper-level objective function should be given by breaking the time interval into two pieces as follows:

$$\Phi_T\left(\{w_k\}, x_0, U_2\right) = \Phi_N\left(\{w_k\}_{k=T-N}^{T-1}, U_2\right) + \Phi_{T-N}\left(\{w_k\}_{k=0}^{T-N-1}, x_0, U_2\right)$$

Note that $\Phi_{T-N}\left(\{w_k\}, x_0, U_2\right)$ is the difference between $\Phi_T\left(\{w_k\}, x_0, U_2\right)$ and

$$\Phi_N\left(\{w_k\}, U_2\right) = \sum_{k=T-N}^{T-1} \|y_k^c - U_2 C \hat{x}_k\|_{R_c}^2 + \|w_k\|_Q^2$$

Exploiting the relation using forward dynamic programming, we are able to establish the equivalence between a full information problem and an estimation problem with a fixed-size estimation window.

Given the initial state $x_0 \in \mathbb{X}$, if there exists disturbances $\{w_k\}_{k=0}^{\tau-1}$ such that

$$x_\tau(x_0, \{w_k\}_{k=0}^{\tau-1}) = z \in \mathbb{X}$$

then we say the state z is reachable at time τ and the set of all reachable states is the corresponding reachable set, i.e.,

$$\mathbb{Z}_\tau = \left\{ x_\tau(x_0, \{w_k\}_{k=0}^{\tau-1}) : x_0 \in \mathbb{X}, \{w_k\} \in \mathbb{W} \right\}$$

Then we define the arrival cost at time τ for the state $z \in \mathbb{Z}_\tau$ as

$$\Theta_\tau = \min_{x_0, \{w_k\}} \left\{ \Phi_\tau \left(x_0, \{w_k\}, U_2 \right) : x_\tau(x_0, \{w_k\}_{k=0}^{\tau-1}) = z \right\}$$

Based on the above cost function, we can reformulate the bi-level optimization problem $P_3(T)$, for $T > N$, as the following equivalent optimal control problem:

$$\mathcal{P}_3'(T): \ \Phi_T^* = \min_{z, \{w_k\}, U_2} \left\{ \begin{array}{c} \Phi_T \left(\{w_k\}_{k=T-N}^{T-1}, U_2 \right) \\ +\Theta_{T-N}(x_{T-N}) : z \in \mathbb{Z}_{T-N} \end{array} \right\} \tag{25}$$

subject to (5)-(8) with $k = T-N, \ldots, T-1$ and

$$U_2 \in \underset{\gamma}{\operatorname{argmin}} \left\{ \Psi_N \left(x_{T-N}, \{f_k\}_{k=T-N}^{T-1}, \gamma \right) \ s.t. \text{ (23)-(24)} \right\} \tag{26}$$

Then an estimate $\{\hat{x}_k | T-1\}_{k=T-N}^{T-1}$ of the actual states is given by

$$\hat{x}_{k|T-1} := x_k \left(\hat{x}_{T-N|T-1}, \{\hat{w}_{k|T-1}\}_{k=T-N}^{T-1} \right) \tag{27}$$

for $k = T-N, \ldots, T-1$.

When the system is nonlinear or constrained, an algebraic expression for the arrival cost rarely exits, yet an approximation $\Theta'_{T-N}(x_{T-N})$ of the arrival cost $\Theta_{T-N}(x_{T-N})$ without constraints can be motivated by the standard Kalman estimate, i.e.,

$$\Theta'_{T-N}(x_{T-N}) = \| x_{T-N} - \bar{x}_{T-N} \|^2_{P_{T-N}} \tag{28}$$

where both P_{T-N} and $\bar{x}_{T-N}$ can be calculated from the previous iteration results as follows:

$$P_{T-N} = S_{T-N}^{-1} > 0, \ P_0 > 0 \tag{29}$$
$$S_{T-N} = (I - K_{T-N} U_2 C) \bar{S}_{T-N} \tag{30}$$
$$K_{T-N} = \bar{S}_{T-N} C^T U_2^T \tilde{S}_{T-N}^{-1} \tag{31}$$
$$\tilde{S}_{T-N} = R_c^{-1} + U_2 C \bar{S}_{T-N} C^T U_2^T \tag{32}$$
$$\bar{S}_{T-N} = A \bar{S}_{T-N-1} A^T + Q^{-1} \tag{33}$$
$$\bar{x}_{T-N} = A \hat{x}_{T-N-1|T-2} + \hat{w}_{T-N-1|T-2} \tag{34}$$

At the end, we provide a resilient moving horizon estimation algorithm by using the previous arguments.

Algorithm 1 Resilient MHE

1: Initialize Q, R and a priori estimate of x_0 and P_0 as well as the window length N. Set $T = 1$.
2: If $T \leq N$, solve the full information optimization problem $\mathcal{P}_3(T)$ (19) and obtain $\left(\hat{x}_{0|T-1}, \{\hat{w}_{k|T-1}\}_{k=0}^{T-1} \right)$. Otherwise, go to Step 4.
3: Estimate the states $\hat{x}_{k|T-1}$ with $k = 0, \ldots, T-1$ from (10).
4: If $T > N$, solve the fix-dimension optimization problem $\mathcal{P}_3'(T)$ (25) and obtain $\left(\hat{x}_{T-N|T-1}, \{\hat{w}_{k|T-1}\}_{k=T-N}^{T-1} \right)$.
5: Estimate the states $\hat{x}_{k|T-1}$ with $k = T-N, \ldots, T-1$ according to (27).
6: Compute the next-step matrix P_{T-N+1} and priori estimate $\bar{x}_{T-N+1}$ based on (29)-(34).
7: Collect the new sensor output y_T to update the measurement data. Set $T = T+1$, and go to Step 2.

IV. CASE STUDY: RESILIENT ESTIMATION ON ENGINE AIR PATH SYSTEMS

In this section, the use of the proposed resilient MHE scheme is illustrated on a vehicle engine air path system. The diagram of the system is depicted in Figure 1. The engine air path system describes the whole process of air flow from ambient to cylinders and reveals the physical dynamics. At first, the air flows through the throttle, where the air mass flow q_{air} can be controlled by the throttle angle actuator θ_{th}. Then the fresh air flows into the intake manifold and changes the manifold pressure P_m. Finally, the air ultimately flows into the cylinder and the air mass flow q_{cyl} into the manifold volume is governed by the engine speed N_e and manifold pressure P_m.

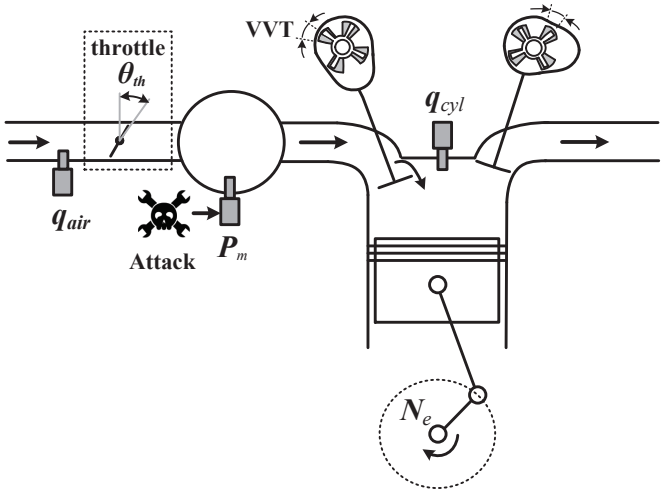

Fig. 1. Airpath scheme. The air mass flow q_{air}, the intake manifold pressure P_m, the aspirated mass air flow q_{cyl}, Opening throttle angle θ_{th}, and number of crankshaft revolutions N_e.

To obtain a dynamical model of the engine air path system, the standard mean value model [26] can be used. Consequently, the dynamical model of the engine air path system can be expressed by

$$\begin{aligned} \dot{P}_m &= \sigma_p \alpha_p(\theta_{th}) f(P_m) - \sigma_p \beta_p g(P_m, N_e) \\ q_{air} &= \alpha_p(\theta_{th}) f(P_m) \\ q_{cyl} &= \beta_p g(P_m, N_e) \end{aligned}$$

where $\sigma_p = \frac{RT_m}{V_m}$, $\alpha_p = \frac{q_{air,max}}{2} \sqrt{\frac{T_{ref}}{T_{atm}}} \frac{P_{atm}}{P_{ref}} (1 - \cos(2\theta_{th}))$, $\beta_p = \frac{V_{disp}}{120 R T_m}$, $f(P_m)$ and $g(P_m, N_e)$ are obtained by MAPs. The above denotations is omitted due to space limit. The readers can refer to [27] for details.

To estimate the state, three dependent sensors are employed to measure q_{air}, P_m and q_{cyl}, respectively. Assuming

that the throttle angle θ_{th} and engine speed N_e keep constant, a linear discrete model can be obtained at a set point with a sample time 10ms. To illustrate the use of the attack-resilient state estimator, the sensor for P_m is compromised from the 30th sample to end, while the sensors for q_{air} and q_{cyl} are always healthy.

In order to evaluate the effectiveness of the proposed resilient MHE scheme, we set $N = 1, \bar{x}_0 = 55240, Q = 500, R = \text{diag}([1, 0.1, 1]) \times 10^{-6}, \delta_k = 4 \times 10^{-3}, \alpha = 100$ and the constraint $x > 0$ and then apply Algorithm 1. The simulated results are given in Figure 2, where the behaviors of the lower- and upper-level optimizations for a randomly chosen simulation are shown. We can observe that the lower-level optimization is able to identify the compromised sensor, while the estimated state of MHE in the upper-level shows better performance than the noised state.

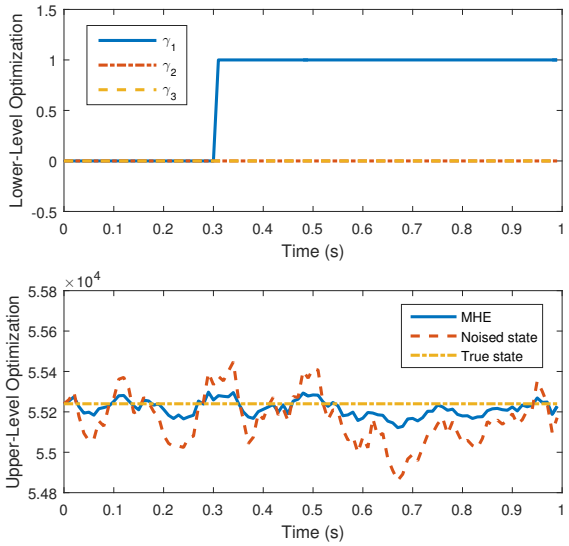

Fig. 2. Resilient MHE of the engine air path system with.

V. CONCLUSION

We have addressed the problem of state estimation for linear systems with measurements affected by adversaries by devising a novel approach based on a moving-horizon strategy, for which the optimality has been established. We have verified the effectiveness of the proposed approach via simulations, where the engine air path system is investigated. Further work will concern the method for nonlinear systems.

REFERENCES

[1] J. Petit and S. Shladover, "Potential cyberattacks on automated vehicles," *IEEE Trans. Intell. Transport. Syst.*, vol. 16, no. 23, pp. 546–556, 2015.

[2] E. Yagdereli, C. Gemci, and A. Z. Aktas, "A study on cyber-security of autonomous and unmanned vehicles," *JDMS*, vol. 12, no. 4, pp. 369–381, 2015.

[3] N. O. Tippenhauer, C. Poepper, K. B. Rasmussen, and S. Capkun, "On the requirements for successful GPS spoofing attacks," in *Proc. 18th ACM Conf. Comput. Commun. Secur.*, Chicago, Illinois, USA, Oct. 17–21, 2011, pp. 75–86.

[4] K. Koscher, A. Czeskis, F. Roesner, S. Patel, T. Kohno, S. Checkoway, D. McCoy, B. Kantor, D. Anderson, H. Shacham, and S. Savage, "Experimental security analysis of modern automobile," in *Proc. IEEE Symp. Secur. Privacy*, Berkeley/Oakland, CA, USA, May 16–19, 2010, pp. 447–462.

[5] P. Mundhenk, S. Steinhorst, M. Lukasiewycz, S. A. Fahmy, and S. Chakraborty, "Lightweight authentication for secure automotive networks," in *Proc. Design, Autom. Test Eur. Conf. Exhibit.*, Grenoble, France, 9–13, 2015, pp. 285–288.

[6] B. G. Stottelaar. (2015) Practical cyber-attacks on autonomous vehicles. [Online]. Available: http://essay.utwente.nl/66766/

[7] A. Francillon, B. Danev, and S. Capkun, "Relay attacks on passive keyless entry and start systems in modern cars," in *Proc. the 18th Annual Network and Distributed System Security Symposium*.

[8] L. F. Combita, J. Giraldo, A. A. Cardenas, and N. Quijano, "Response and reconfiguration of cyber-physical control systems: A survey," in *Proc. IEEE 2nd Colombian Conf. Automatic Control*, Manizales, Colombia, Oct. 14–16, 2015.

[9] B. Zheng, P. Deng, R. Anguluri, Q. Zhu, and F. Pasqualetti, "Cross-layer codesign for secure cyber-physical systems," *IEEE Trans. Computer-Aided Design*, vol. 35, no. 5, pp. 699–711, 2016.

[10] A. H. Jazwinski, "Limited memory optimal filtering," *IEEE Trans. Automat. Contr.*, vol. 13, no. 5, pp. 558–563, 1968.

[11] C. V. Rao, J. B. Rawlings, and D. Q. Mayne, "Constrained state estimation for nonlinear discrete-time systems: stability and moving horizon approximations," *IEEE Trans. Automat. Contr.*, vol. 48, no. 2, pp. 246–257, 2003.

[12] A. Alessandri, M. Baglietto, and G. Battistelli, "Receding-horizon estimation for discrete-time linear systems," *IEEE Trans. Automat. Contr.*, vol. 48, no. 3, pp. 473–478, 2003.

[13] S. X. Ding, *Model-Based Fault Diagnosis Techniques: Design Schemes, Algorithms and Tools*. Longdon: Springer-Verlag, 2013.

[14] L. Xie, Y. Mo, and B. Sinopoli, "Integrity data attacks in power market operations," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 659–666, 2011.

[15] Y. Shoukry and P. Tabuada, "Event-triggered state observers for sparse sensor noise/attack," *IEEE Control Syst. Mag.*, vol. 61, no. 8, pp. 2079–2091, 2016.

[16] F. Pasqualetti, F. Dorfler, and F. Bullo, "Attack detection and identification in cyber-physical systems," *IEEE Trans. Automat. Contr.*, vol. 58, no. 11, pp. 2715–2729, 2013.

[17] A. Teixeira, I. Shames, H. Sandberg, and K. H. Johansson, "A secure control framework for resource-limited adversaries," *Automatica*, vol. 51, pp. 135–148, 2015.

[18] H. Fawzi, P. Tabuada, and S. Diggavi, "Secure estimation and control for cyber-physical systems under adversarial attacks," *IEEE Trans. Automat. Contr.*, vol. 59, no. 6, pp. 1454–1467, 2014.

[19] Y. Mo and B. Sinopoli, "Secure estimation in the presence of integrity attacks," *IEEE Trans. Automat. Contr.*, vol. 60, no. 4, pp. 1145–1151, 2015.

[20] M. Pajic, I. Lee, and G. J. Pappas, "Attack-resilient state estimation for noisy dynamical systems," *IEEE Network*, vol. 4, no. 1, pp. 82–92, 2017.

[21] M. Pajic, J. Weimer, N. Bezzo, O. Sokolsky, G. J. Pappas, and I. Lee, "Design and implementation of attack-resilient cyberphysical systems with a focus on attack-resilient state estimators," *IEEE Control Syst. Mag.*, vol. 37, no. 2, pp. 66–81, 2017.

[22] C. Lee, H. Shim, and Y. Eun, "Secure and robust state estimation under sensor attacks, measurement noises, and process disturbances: observer-based combinatorial approach," in *Proc. 2015 European Control Conference*, Linz, Austria, 15–17, 2015.

[23] Y. Mo and E. Garone, "Secure dynamic state estimation via local estimators," in *Proc. 2016 IEEE 55th Conference on Decision and Control*, Las Vegas, USA, Dec. 12–14, 2016.

[24] J. Kim, C. Lee, H. Shim, Y. Eun, and J. H. Seo, "Detection of sensor attack and resilient state estimation for uniformly observable nonlinear systems," in *Proc. 2016 IEEE 55th Conference on Decision and Control*, Las Vegas, USA, Dec. 12–14, 2016.

[25] R. Bellman, "On the theory of dynamic programming," *Proceedings of the National Academy of Sciences*, vol. 38, no. 8, pp. 716–719, 1952.

[26] L. Guzzella and C. Onder, *Introduction to Modeling and Control of Internal Combustion Engine Systems*. Berlin: Springer-Verlag, 2004.

[27] Y. Li, Y. Hu, S. Wang, and H. Chen, "Nonlinear model predictive controller design for air system control of a gasoline engine," in *Proc. 2016 12th WCICA*, Guilin, China, 12–15, 2016.