Proceedings of Machine Learning Research vol 242:890–901, 2024

Conditions for Parameter Unidentifiability of Linear ARX Systems for Enhancing Security

Xiangyu Mao MAOXY20@SJTU.EDU.CN
Shanghai Jiao Tong University, Shanghai, China

Jianping He JPHE@SJTU.EDU.CN
Shanghai Jiao Tong University, Shanghai, China

Chengpu Yu YUCHENGPU@BIT.EDU.CN
Beijing Institute of Technology, Beijing, China

Chongrong Fang CRFANG@SJTU.EDU.CN
Shanghai Jiao Tong University, Shanghai, China

Abstract

For an adversarial observer of parametric systems, the identifiability of parameters reflects the possibility of inferring the system dynamics and then affects the performance of attacks against the systems. Hence, achieving unidentifiability of the parameters, which makes the adversary unable to get identification with low variance, is an attractive way to enhance security. In this paper, we propose a quantitative definition to measure the unidentifiability based on the lower bound of identification variance. The lower bound is given via the analysis of the Fisher Information Matrix (FIM). Then, we propose the necessary and sufficient condition for unidentifiability and derive the explicit form of the unidentifiability condition for linear autoregressive systems with exogenous inputs (ARX systems). It is proved that the unidentifiability of linear ARX systems can be achieved through quadratic constraints on inputs and outputs. Finally, considering an optimal control problem with security concerns, we apply the unidentifiability constraint and obtain the optimal controller. Simulations demonstrate the effectiveness of our method.

Keywords: System Identification, Unidentifiability, Fisher Information, Linear ARX Systems

1. Introduction

System identification is an active and diverse branch of system theory (Nguyen and Wood, 1982). One of the attractive applications of system identification is to identify the parameters for agent models (Gautier et al., 2013). Then, the identification result provides an avenue for an adversarial observer to infer the control law or optimization task of the agent (Grover et al., 2021). Recently, a substantial amount of research has been proposed on attacking an agent based on the prior knowledge of the model parameters (Phillips et al., 2019; de Sá et al., 2020). Note that the convergence and accuracy of identification are important to the performance of attacks (Yuan and Mo, 2015). Hence, we can enhance the security of the agents by achieving unidentifiability of the parameters.

In the literature, a multitude of research efforts has been devoted to analyzing the identifiability (Bellman and Åström, 1970; Brun et al., 2001; Wieland et al., 2021). The necessary and sufficient conditions for the identifiability of parametric systems of various kinds of models are proposed (Walter and Lecourtier, 1982; Walter and Pronzato, 1996; Karlsson et al., 2012). However, there remain some notable issues for achieving unidentifiability despite the above contributions. First, system unidentifiability does not have a recognized definition temporarily. Existing definitions of

identifiability are mainly qualitative definitions, which reflect the possibility of identification under a best-case scenario, rather than give the quantitative measurement of identification performance (Eisenberg and Hayashi, 2014). For security concerns, it is necessary to define unidentifiability quantitatively to measure the identification performance of the adversary, which is closely connected to the performance of attacks (Teixeira et al., 2015; Teixeira, 2019). Second, few of the existing papers consider control problems with security concerns from the perspective of unidentifiability. Existing literature mainly considers optimal and robust control under conditions of limited identification performance (Stojanovic et al., 2016; Lale et al., 2021). Optimal control with constraints for limitation of identification performance remains an open issue.

The above concerns motivate us to enhance the security of parametric systems by proposing a quantitative unidentifiability definition and achieving unidentifiability. We take linear ARX systems as an example of analysis. Compared with the existing methods used to enhance the security of parametric systems against adversarial attacks, the proposed method has the advantage that the proposed definition of unidentifiability does not rely on the adversarial identification method. The main contributions in this paper are summarized as follows.

- We propose a quantitative definition for unidentifiability based on the evaluation of the lower bound of identification variance. For linear ARX systems, we derive the explicit form of conditions for unidentifiability. We prove that the unidentifiability of linear ARX systems can be achieved through quadratic constraints on inputs and outputs.
- For linear quadratic regulator (LQR) control problem for linear ARX systems, we provide a solution with security concern of unidentifiability conditionin this paper. We proved that this problem can be relaxed as a quadratic constrained quadratic programming (QCQP) problem. Simulations demonstrate the effectiveness of our method.

In Section 2, we provide the basic model, the scenario considered in this paper, and the definition of unidentifiability. Section 3 proposes the conditions and the realization for unidentifiability. In Section 4, we present numerical simulation examples, followed by conclusions in Section 5.

2. Preliminaries and Definition of Unidentifiability

2.1. Notations

We use capital letters with subscript t (e.g., Y_t) as vectors or matrices at time t and lower-case letters and subscript (i) or (i,j) (e.g., $y_{t,(i)}$)as the i-th element in the vector or the element at the i-th row and j-th column of the matrix, respectively. We define $A * B$ as the matrix obtained by multiplying the matrices with the same dimension A and B by elements. We denote the dimension of a matrix $(\cdot)$ by $\dim(\cdot)$. We use '$\succeq$' as an operator of matrix inequality where $A \succeq B$ means for any vector $x \in \mathbb{R}^{dim(A)}$, $x^\top (A-B)x \geq 0$. We denote the expectation of a random variable $(\cdot)$ by $\mathbb{E}[(\cdot)]$.

2.2. Model Description

The basic model investigated in this paper is a linear discrete-time, parametric, autoregressive system $\mathcal{S}$ with exogenous inputs (ARX(p,q)), which is given by

$$\mathcal{S}: Y_t = \sum_{i=1}^{p} A_i^* Y_{t-i} + \sum_{j=1}^{q} B_j^* U_{t-j} + V_t, \tag{1}$$

where $t \in \mathbb{N}$, $Y_t \in \mathbb{R}^m$ represents the output, $U_t \in \mathbb{R}^l$ is the input vector, $V_t \in \mathbb{R}^m$ is the noise and $A_i^* \in \mathbb{R}^{m\times m}$, $B_i^* \in \mathbb{R}^{m\times l}$ are matrices of parameters to be identified.

For brevity, we define $n = pm + ql$ and let

$$X_t = \left[Y_{t-1}^\top, \ldots, Y_{t-p}^\top, U_{t-1}^\top, \ldots, U_{t-q}^\top\right]^\top, \ \Theta^* = \left[A_1^*, \ldots, A_p^*, B_1^*, \ldots, B_q^*\right].$$

Then, the system model can be generalized as a system with mn parameters Θ with the true value $\Theta = \Theta^*$ as follows.

$$\mathcal{S} : Y_t = \Theta X_t + V_t. \tag{2}$$

The following assumption is made throughout the paper, which is commonly used for the analysis of system identification (de Sá et al., 2020; Nguyen and Wood, 1982).

Assumption 1 *The noise term V_t is an i.i.d. sequence of noise variables with each element obeying Gaussian distribution $\mathcal{N}(0, \sigma_v^2)$.*

2.3. Scenario and Definition

Considering a given control sequence $U_{0:T} = \left[U_0^\top, U_1^\top, \ldots, U_{T-1}^\top\right]^\top$, the system has a stochastic trajectory/output sequence, $Y_{0:T} = \left[Y_0^\top, Y_1^\top, \ldots, Y_{T-1}^\top\right]^\top$.

We suppose that there is an adversary, who knows that the orders of the ARX system model, i.e., m, l, and (p, q), and can passively observe $\{Y_{0:T}, U_{0:T}\}$. The objective of the adversary is to use the observed data $\{Y_{0:T}, U_{0:T}\}$ to identify the parameters Θ of $\mathcal{S}$, which is described as follows.

- **Parameter Identification**: Derive an unbiased estimation function of $\Theta \in \mathbb{R}^{m\times n}$ for any given observed data $\{Y_{0:T}, U_{0:T}\}$, i.e., derive

$$\hat{\Theta} : \{\mathbb{R}^{mT}, \mathbb{R}^{lT}\} \to \mathbb{R}^{m\times n}, \text{ s.t., } \forall\Theta, \ \mathbb{E}\left[\hat{\Theta}(Y_{0:T}, U_{0:T})\right] = \Theta. \tag{3}$$

Note that $Y_{0:T}$ is a vector of random variables since $V_{0:T}$ are unknown sequences. It follows that the estimator $\hat{\Theta}(Y_{0:T}, U_{0:T})$ is also a matrix of random variables. Hence, we can use the covariance matrix of $\hat{\Theta}(Y_{0:T}, U_{0:T})$ to define unidentifiability as follows.

Definition 1 (Parameter Unidentifiability) *Given a control sequence $\{U_{0:T}\}$, a semi-positive definite matrix $\Sigma \in \mathbb{R}^{mn\times mn}$, $\mathcal{S}$ is Σ-**unidentifiable** iff for any unbiased parameter estimator $\hat{\Theta}(Y_{0:T}, U_{0:T})$,*

$$\mathbb{E}\left[\left(\hat{\boldsymbol{\theta}}(Y_{0:T}, U_{0:T}) - \boldsymbol{\theta}^*\right)\left(\hat{\boldsymbol{\theta}}(Y_{0:T}, U_{0:T}) - \boldsymbol{\theta}^*\right)^\top\right] \succeq \Sigma, \tag{4}$$

where $\hat{\boldsymbol{\theta}}, \boldsymbol{\theta}^* \in \mathbb{R}^{mn\times 1}$ are the reshaped vectors which are obtained by expanding $\hat{\Theta}$ and Θ^* by rows, respectively.

By Definition 1, we know that if a system is Σ-unidentifiable, then the identification of parameters obtained by the adversary always has a variance matrix larger than Σ. Specifically, we have $\hat{\theta}_{(i)}$ satisfy $\mathbb{E}\|\hat{\theta}_{(i)} - \theta_{(i)}^*\| \geq \sigma_{(i)}$, $i = 1, 2, \ldots, mn$, where $(\sigma_{(1)}^2, \sigma_{(2)}^2, \ldots, \sigma_{(mn)}^2)$ are the main diagonal elements of Σ.

2.4. Problem Formulation

We aim to derive the condition for unidentifiability and apply the condition to an LQR control problem for security. We assume that the control sequence can be fully designed $U_{0:T}$ by us but the noises are unknown both to us and to the adversary. The main difficulty is that we do not know the distributions of $\hat{\boldsymbol{\theta}}$ since we do not know the estimator of the adversary. Hence, we are supposed to design the controller to ensure unidentifiability for all unbiased estimators just like Definition 1. The LQR control problem considered in this paper is a finite, discrete-time control problem with cost of trajectory tracking and energy of control signal, which is given by

$$\begin{aligned} P_1 : \min_{U_{0:T}} & \left(\sum_{t=0}^{T} (Y_t - Y_{\text{ref},t})^\top Q (Y_t - Y_{\text{ref},t}) + \sum_{t=0}^{T-1} U_t^\top R U_t \right), \\ \text{s.t. } & \mathbb{E}\left[\left(\hat{\boldsymbol{\theta}}(Y_{0:T}, U_{0:T}) - \boldsymbol{\theta}^* \right) \left(\hat{\boldsymbol{\theta}}(Y_{0:T}, U_{0:T}) - \boldsymbol{\theta}^* \right)^\top \right] \succeq \Sigma, \\ & Y_t = \Theta^* X_t + V_t, \ \ t = 1, 2, \ldots, T. \end{aligned} \tag{5}$$

3. System Unidentifiability: Conditions and Realization

In this section, we investigate the conditions for parameter unidentifiability. We start with the definition of FIM of linear ARX models. Then, we analyze the bound of the identification variance by FIM and give the conditions for parameter unidentifiability. Finally, we derive the solution to P_1 to achieve unidentifiability.

3.1. Fisher Information Matrix in System Identification

The FIM plays a significant role in parameter estimation. It is a symmetric matrix that represents the amount of information contained in the observation data. Hence, this paper starts from FIM and derives the specific form of FIM for parameter identification problems under linear ARX models.

First, we give the basic definition of FIM. We consider a vector of random processes Y with unknown parameters $\boldsymbol{\theta}$, assuming that the probability density function $\boldsymbol{f}$ of Y is smooth and $\log(\boldsymbol{f})$ is continuously differentiable w.r.t $\boldsymbol{\theta}$. Denote $\boldsymbol{f} = \left[f_{(1)}, f_{(2)}, \ldots, f_{(m)} \right]^\top$. Supposing that we make observations of Y and get $Y_{0:T}$. The likelihood function vector $L(Y; \boldsymbol{\theta})$ and the score function matrix $S(Y; \boldsymbol{\theta})$ are defined as follows.

$$L(Y; \boldsymbol{\theta}) = \left[\prod_{t=1}^{T} f_{(1)}(Y_t|\boldsymbol{\theta}), \ldots, \prod_{t=1}^{T} f_{(m)}(Y_t|\boldsymbol{\theta}) \right]^\top, \ \ S(Y; \boldsymbol{\theta}) = \sum_{t=1}^{T} \frac{\partial \log \boldsymbol{f}(Y_t|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^\top}. \tag{6}$$

Then, the FIM, denoted by $I(\boldsymbol{\theta})$, is defined by

$$I(\boldsymbol{\theta}) = \mathbb{E}\left[S(Y; \boldsymbol{\theta})^\top S(Y; \boldsymbol{\theta}) \right] = -\mathbb{E}\left[\frac{\partial^2}{\partial \boldsymbol{\theta}^\top \partial \boldsymbol{\theta}} \log \left(L(Y; \boldsymbol{\theta}) \right) \right]. \tag{7}$$

Then, we derive the form of the FIM in the parameter identification problem of linear ARX models. We have the following theorem (see the proof in the appendix).

Theorem 1 *For a linear ARX system $\mathcal{S}$ defined as (2) with parameters Θ and observation $\{Y_{0:T}, U_{0:T}\}$, the FIM, $I(\Theta)$ is a blocked diagonal matrix with m identical blocks as follows.*

$$I(\Theta) = diag(\frac{1}{\sigma_v^2}\mathbb{E}\left[\sum_{t=1}^{T} X_t X_t^\top\right], \ldots, \frac{1}{\sigma_v^2}\mathbb{E}\left[\sum_{t=1}^{T} X_t X_t^\top\right]). \tag{8}$$

Remark 1 *Existing literature on system identification (Eisenberg and Hayashi, 2014) has a different definition of FIM compared to this paper, where the FIM is defined by sensitivity matrix W, i.e., $I = W^\top W$, where $W = \left[\frac{\partial Y}{\partial \theta_{(1)}}, \ldots, \frac{\partial Y}{\partial \theta_{(mn)}}\right]$. This definition ignores the impact of observation errors on identification variance, and it is generally used to analyze identifiability. Considering that we need to determine the bound of the identification variance, we adopt the definition of (8).*

3.2. On the Unidentifiability of Parameters

In this subsection, we investigate the condition of unidentifiability based on the FIM. Considering that Definition 1 makes sense only when the unbiased estimator of Θ exists, first, we give the condition for the existence of unbiased estimators. We define observationally equivalence as follows.

Definition 2 (Observationally equivalence,(Rothenberg, 1971)) *Two parameter points of Θ_1 and Θ_2 are said to be observationally equivalent iff for all X_t in the observation data, $\Theta_1 X_t = \Theta_2 X_t$.*

Then, we propose the following lemma (see the proof in the appendix).

Lemma 1 *Supposing that the parameters taking value in Ω, an unbiased estimator of Θ exists iff all the parameter points in Ω are not observationally equivalent to each other.*

Next, the condition for the existence of an unbiased estimator of Θ is given by the following lemma.

Lemma 2 *Unbiased estimator of Θ exists iff $I(\Theta)$ has full rank.*

Lemma 2 can be directly obtained by combining Lemma 1 with Theorem 1 in (Rothenberg, 1971). It implies that Definition 1 can be discussed when $\text{rank}(\sum_{t=1}^{T} X_t X_t^\top) = n$.

Then, we use FIM to derive the variance bound of parameter identification.

Theorem 2 *Supposing that $I(\Theta)$ has full rank, for all unbiased estimator $\hat{\Theta}$, we have*

$$\mathbb{E}\left[\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^*\right)\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^*\right)^\top\right] \succeq I^{-1}(\boldsymbol{\theta}). \tag{9}$$

Furthermore, for a linear ARX system $\mathcal{S}$ defined as (2) with parameters Θ and observation $\{Y_{0:T}, U_{0:T}\}$, we have $I^{-1}(\boldsymbol{\theta}) = \inf \mathbb{E}\left[\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^*\right)\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^*\right)^\top\right]$.

The first part of Theorem 2 is known as the Cramér-Rao lower bound (Cramér, 1999). Note that the Cramér-Rao lower bound only provides a lower bound of the variance of unbiased estimators, which is not an infimum of the variance. To prove the second part of Theorem 2, we find that the Maximum Likelihood Estimation (MLE) for ARX system models that has the minimum variance $I^{-1}(\boldsymbol{\theta})$ (see the proof in appendix). The MLE is given by

$$\hat{\Theta}_{\text{MLE}} = \left(\sum_{t=1}^{T} Y_t X_t^\top\right)\left(\sum_{t=1}^{T} X_t X_t^\top\right)^{-1}. \tag{10}$$

It is proved that $\hat{\Theta}_{\text{MLE}}$ is an unbiased estimator with $\mathbb{E}\left[\left(\hat{\boldsymbol{\theta}}_{\text{MLE}}-\boldsymbol{\theta}^*\right)\left(\hat{\boldsymbol{\theta}}_{\text{MLE}}-\boldsymbol{\theta}^*\right)^\top\right]=I^{-1}(\boldsymbol{\theta})$.

Next, from Theorem 2, we can get the following theorem for parameter unidentifiability.

Theorem 3 *Supposing that $I(\Theta)$ has full rank, given a control sequence $\{U_{0:T}\}$ and $\Sigma \succeq 0$, $\mathcal{S}$ is Σ-unidentifiable iff*

$$I^{-1}(\boldsymbol{\theta}) \succeq \Sigma. \tag{11}$$

3.3. Controller Design for Parameter Unidentifiability

We can re-write the controller design problem P_1 as P_1' by Theorem 3.

$$\begin{aligned} P_1': \min_{U_{0:T}} & \left(\sum_{t=0}^{T}(Y_t-Y_{\text{ref},t})^\top Q(Y_t-Y_{\text{ref},t})+\sum_{t=0}^{T-1}U_t^\top R U_t\right), \\ \text{s.t. } & I^{-1}(\boldsymbol{\theta}) \succeq \Sigma,\ Y_t=\Theta^* X_t+V_t,\ t=1,2,\ldots,T. \end{aligned} \tag{12}$$

To avoid the computation of $I^{-1}(\boldsymbol{\theta})$ which is complicated in practice, we give the following corollary.

Corollary 1 *Given $\Sigma=diag(\sigma_{(1)}^2,\sigma_{(2)}^2,\ldots,\sigma_{(mn)}^2)$, supposing $\sigma_{\max}=\max(\sigma_{(1)},\ldots,\sigma_{(mn)})>0$, $\mathcal{S}$ is Σ-unidentifiable if $I(\Theta)$ has rank n and $\forall t=0,1,\ldots,T-1$,*

$$\|U_t\|_\infty^2 \le \frac{\sigma_v^2}{T\sigma_{\max}^2},\ \mathbb{E}\left[\|Y_t\|_\infty^2\right] \le \frac{\sigma_v^2}{T\sigma_{\max}^2}. \tag{13}$$

Corollary 1 gives a sufficient condition for parameter unidentifiability. Since sometimes it is hard to limit the bound of initial output, we can release the constraint (13) in Corollary 1 as follows, which also serves as a sufficient condition for Σ-unidentifiable (please see the proof in appendix).

$$\|\sum_{t=0}^{T-1}U_t*U_t\|_\infty \le \frac{\sigma_v^2}{\sigma_{\max}^2},\ \mathbb{E}\left[\|\sum_{t=0}^{T-1}Y_t*Y_t\|_\infty\right] \le \frac{\sigma_v^2}{\sigma_{\max}^2}. \tag{14}$$

Then, we can apply the condition for parameter unidentifiability in the controller design process. Equation (14) serves as a released form of sufficient condition for Σ-unidentifiable. Next, we can relax the constraint in the controller design problem P_1 as follows.

$$\begin{aligned} P_2: \min_{U_{0:T}} & \left(\sum_{t=0}^{T}(Y_t-Y_{\text{ref},t})^\top Q(Y_t-Y_{\text{ref},t})+\sum_{t=0}^{T-1}U_t^\top R U_t\right), \\ \text{s.t. } & \|\sum_{t=0}^{T-1}U_t*U_t\|_\infty \le \frac{\sigma_v^2}{\sigma_{\max}^2},\ \mathbb{E}\left[\|\sum_{t=0}^{T-1}Y_t*Y_t\|_\infty\right] \le \frac{\sigma_v^2}{\sigma_{\max}^2}, \\ & Y_t=\Theta^* X_t+V_t,\ t=1,2,\ldots,T. \end{aligned} \tag{15}$$

Since the constraint and the cost function of P_2 are convex, P_2 can be seen as a convex relaxation of P_1. Note that P_2 is a QCQP problem. Hence, it can be solved by solvers such as the CVX toolbox in Matlab, which means the optimal control problem with security concern P_2 is solvable.

3.4. Other Conditions for Parameter Unidentifiability

By Theorem 3, we can also give a necessary condition for Σ-unidentifiable as follows.

Corollary 2 *Given $\Sigma = diag(\sigma^2_{(1)}, \sigma^2_{(2)}, \dots, \sigma^2_{(mn)})$, supposing that $\mathcal{S}$ is Σ-unidentifiable, we have $\forall i = 1 + pm, 2 + pm, \dots, ql + pm$ and $\forall j = 1, 2, \dots, pm$,*

$$\begin{aligned} &\max(\sigma^2_{(i)}, \sigma^2_{(i+n)}, \dots, \sigma^2_{(i+(m-1)n)}) \sum_{t=0}^{T-1} u^2_{t,(i-pm)} \le \sigma_v^2, \\ &\max(\sigma^2_{(j)}, \sigma^2_{(j+n)}, \dots, \sigma^2_{(j+(m-1)n)}) \mathbb{E}\left[\sum_{t=0}^{T-1} y^2_{t,(j)}\right] \le \sigma_v^2. \end{aligned} \tag{16}$$

Corollary 3 *Given $\Sigma = diag(\sigma^2_{(1)}, \dots, \sigma^2_{(mn)})$, supposing that $\max(\sigma^2_{(j)}, \dots, \sigma^2_{(j+(m-1)n)}) > 0$, there always exist T_M, s.t. when $T > T_M$, system $\mathcal{S}$ is not Σ-unidentifiable.*

Corollary 2 gives the necessary condition for parameter unidentifiability. While Corollary 3 shows that Σ-unidentifiable cannot be achieved when $T \to \infty$, which means that we can always find a convergent unbiased estimator for linear ARX parametric systems.

Moreover, by Theorem 3, we can obtain the bound of output prediction as follows.

Corollary 4 *Given Σ, supposing that $\mathcal{S}$ is Σ-unidentifiable, we have*

$$\mathbb{E}\left[\left(\hat{Y}_{T+1} - Y_{T+1}\right)\left(\hat{Y}_{T+1} - Y_{T+1}\right)^\top\right] \succeq J_T^\top \Sigma J_T + \sigma_v^2 I_m, \tag{17}$$

where $\hat{Y}_{T+1} = \hat{\Theta} X_T$, $\hat{\Theta}$ is an unbiased estimator of Θ and J_T is a blocked diagonal matrix with m blocks, i.e., $J_T = diag(X_T, X_T, \dots, X_T)$.

Corollary 4 can be proved by combining Assumption 1 and the fact that $\frac{\partial \hat{Y}_{T+1}}{\partial \hat{\boldsymbol{\theta}}^\top} = J_T$. Corollary 4 shows that when the system is Σ-unidentifiable, the prediction of system output by parameters is bounded by the Jacobian matrix and FIM.

4. Numerical Simulation

This section uses a numerical example of solving P_2 with simulation to verify the conditions for parameter unidentifiability. Then, to evaluate security, we use a backpropagation (BP) neural network to predict the output of an unidentifiable system.

4.1. LQR Control Considering Parameter Unidentifiability

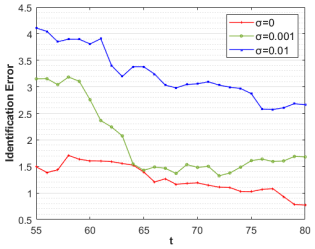
First, we consider an LQR control problem with security concerns defined as P_2. We consider a double-input double-output system $\mathcal{S}$ with an ARX$(9, 9)$ model, with $Y_t \in \mathbb{R}^2$, $U_t \in \mathbb{R}^2$, $V_t \in \mathbb{R}^2$ is the noise obeying $\mathcal{N}(0, 10^{-4})$. $A_i^* \in \mathbb{R}^{2\times 2}$, $B_i^* \in \mathbb{R}^{2\times 2}$ are constant matrices which are randomly generated. We let the initial state $Y_0 = \mathbf{0}$, the maximum time length of observation $T = 100$, and the trajectory Y_{ref} that needs to be tracked is defined as $Y_{\text{ref},t,1} = 0.97t/T$, $Y_{\text{ref},t,2} = 0.83\sin(\pi t/T)$.

For the identification method of Θ, we use the MLE of Θ defined as (10). As is proved in Theorem 2, the MLE of Θ is an unbiased estimator with the minimum identification variance.

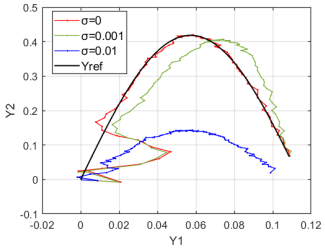
Then, we design the controller which is the solution to P_2. We conduct experiments under $\sigma_{\max} = 10^{-2}, 10^{-3}$ and $\sigma_{\max} = 0$ and get three controllers.

In Fig.1(*a*), the identification errors of MLE under the three controllers are compared. It intuitively illustrates the effectiveness of the unidentifiability condition. It can be seen that the identification error under all controllers tends to decrease. However, the controller designed under $\sigma_{\max} = 10^{-2}$ makes the performance of identification much worse than that of the controller under $\sigma_{\max} = 10^{-3}$, followed by the best performance of controller without unidentifiability condition.

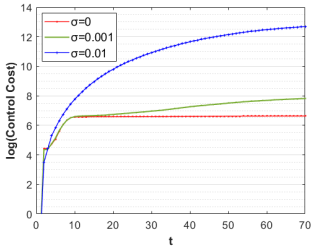
Fig.1(*b*) shows the control cost under the three controllers and Fig.1(*c*) shows the system trajectories. It follows from Fig.1(*b*) that using larger $\sigma_{\max}$ has more control cost. This is intuitively illustrated in Fig.1(*c*) where under the larger $\sigma_{\max}$, the trajectory is harder to track Y_{ref}. Hence, in practice, it is necessary to make a trade-off between unidentifiability and control performance.



(*a*) The identification errors of MLE under the three different controllers.



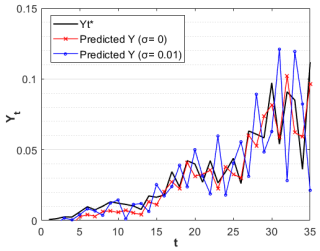
(*b*) The system trajectories under the three different controllers.



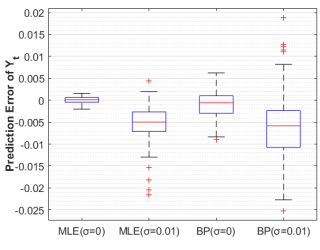
(*c*) The control cost under the three different controllers.

Figure 1: Experiment of LQR control under the three different controllers.

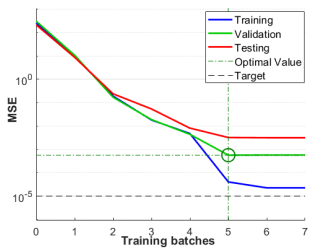
4.2. Parameter Unidentifiability Against BP Neural Network Prediction



(*a*) Output prediction of the BP network under different unidentifiability conditions.



(*b*) Box-plot of output prediction of Y_t in 100 Monte Carlo runs by the MLE predictor and the BP network.



(*c*) Training process of the BP network under $\sigma_{\max} = 0.01$ in one of the Monte Carlo runs in Fig.2(*b*).

Figure 2: Experiment of BP network output prediction under unidentifiability conditions.

To prove that the condition of unidentifiability works for other estimators, in Fig.2, we use a backpropagation (BP) neural network to predict the output of $\mathcal{S}_1$ under the two controllers with $\sigma_{\max} = 10^{-2}$ and $\sigma_{\max} = 0$. We use a fully connected neural network with four layers, with 45 and 5 nodes in the two hidden layers for identification. Fig.2(*a*) shows the output prediction of the BP network under different unidentifiability conditions. It is shown that the prediction under larger σ in the unidentifiability condition tends to have a large error. This conclusion is shown more obviously in Fig.2(*b*), which provides the box plot of the prediction of Y and shows that the

prediction variance is higher under larger σ. Fig.2(*c*) gives the training process of the BP network, which shows that the optimal value of the mean square error of the BP network cannot convergent to the target 10^{-5} under $\sigma_{\max} = 10^{-2}$. It can be seen from these figures that the unidentifiability conditions work for the BP neural network.

5. Conclusion

In this paper, we propose a definition of unidentifiability which gives the lower bound of identification variance of system parameters. Then, we obtain the condition for achieving unidentifiability and apply this condition to linear ARX systems for an example to get explicit expressions of the unidentifiability conditions. The condition is provided by the analysis of the Cramér-Rao lower bound, which is defined by the inverse of FIM. It is proved that the unidentifiability of linear ARX systems constraint is a quadratic constraint on inputs and outputs. Next, we derive the solution to an LQR controller design problem with conditions of unidentifiability, which serves as an example of the combination of optimal control and security concerns. Finally, we use numerical simulations to illustrate the effectiveness of unidentifiability. It shows that the unidentifiability condition can protect the parameters from being identified not only by the MLE or least-squares methods but also BP networks, which enhance the security of the system.

Appendix A. Proof of Theorem 1

From the system model (2) and Assumption 1, we have

$$f_{(i)}(Y = \boldsymbol{y}|\boldsymbol{\theta}) = \frac{1}{\sqrt{2\pi}\sigma_v} exp\left(-\frac{\left(\boldsymbol{e}_{(i)}^{\top}(\boldsymbol{y} - \Theta X_t)\right)^2}{2\sigma_v^2} \right),\ i = 1, 2, \ldots, m, \tag{18}$$

where $\boldsymbol{e}_{(i)} \in \mathbb{R}^m$ is a vector with the i-th element being 1 and the remaining elements being 0. We define $w_{t,(i)} = \boldsymbol{e}_{(i)}^{\top}(Y_t - \Theta X_t)$. It follows that

$$\log\left(L(Y;\boldsymbol{\theta})\right) = -\frac{1}{2\sigma_v^2}\sum_{t=1}^{T}\left[w_{t,(1)}^2, \ldots, w_{t,(m)}^2\right]^{\top} - T\log(\sqrt{2\pi}\sigma_v)\left[1, 1, \ldots, 1\right]^{\top}. \tag{19}$$

Then, combining with the system model, we have $I(\Theta)_{(i+kn,j+kn)} = \frac{1}{\sigma_v^2}\mathbb{E}\left[\sum_{t=1}^{T} x_{t,(i)}x_{t,(j)}\right]$, where $i, j = 1, 2, \ldots, n$ and $k = 0, 1, \ldots, m-1$, $I(\Theta)_{(i+kn,j+kn)}$ is the i-th row and j-th column element of $I(\Theta)$ and $x_{t,(i)}$ is the i-th element of vector X_t. For $k_1 \neq k_2$, we obtain that $I(\Theta)_{(i+k_1n,j+k_2n)} = 0$. Hence, Theorem 1 is proved.

Appendix B. Proof of Lemma 1

Suppose that there exist two observationally equivalent parameter points $\Theta_1 \neq \Theta_2$ and an unbiased estimator $\hat{\Theta}$. When $\Theta^* = \Theta_1$, given $U_{0:T}$, we have $\mathbb{E}\left[\hat{\Theta}(\boldsymbol{y}_1, U_{0:T})\right] = \Theta_1$.

Similarly, when $\Theta^* = \Theta_2$, we have $\mathbb{E}\left[\hat{\Theta}(\boldsymbol{y}_2, U_{0:T})\right] = \Theta_2$.

Since $\Theta_1 X_t = \Theta_2 X_t$, by Assumption 1, we have $\mathbb{E}\left[\hat{\Theta}(\boldsymbol{y}_1, U_{0:T})\right] = \mathbb{E}\left[\hat{\Theta}(\boldsymbol{y}_2, U_{0:T})\right]$, which means $\Theta_1 = \Theta_2$. Hence, Lemma 1 is proved.

Appendix C. Proof of Theorem 2

It follows from (19) that $\frac{\partial \log L(Y;\Theta)}{\partial \Theta^\top} = -\frac{1}{\sigma_v^2}\sum_{t=1}^{T}(Y_t - \Theta X_t)X_t^\top$. By letting $\frac{\partial \log L(Y;\boldsymbol{\theta})}{\partial \Theta^\top} = \mathbf{0}$, we have the Maximum Likelihood Estimation (MLE) of Θ as defined in (10). This estimator has the same expression as the Ordinary Least Square (OLS) estimator, which is known as an unbiased estimator, i.e., $\mathbb{E}[\hat{\boldsymbol{\theta}}_{MLE}] = \boldsymbol{\theta}^*$.

Since the variance of the MLE is $I^{-1}(\boldsymbol{\theta})$, we have that for ARX models and parameters Θ, $\hat{\Theta}_{MLE}$ is an ubiased estimator that has the minimum variance $I^{-1}(\boldsymbol{\theta})$.

Hence, Theorem 2 is proved.

Appendix D. Proof of Corollary 1

From Theorem 3 and (8), given $\Sigma = \text{diag}(\sigma^2_{(1)}, \sigma^2_{(2)}, \ldots, \sigma^2_{(mn)})$, $\mathcal{S}$ is Σ-unidentifiable iff

$$\text{diag}(\frac{1}{\sigma_v^2}\mathbb{E}\left[\sum_{t=1}^{T} X_t X_t^\top\right], \ldots, \frac{1}{\sigma_v^2}\mathbb{E}\left[\sum_{t=1}^{T} X_t X_t^\top\right])^{-1} \succeq \Sigma,$$

which is equivalent to $\forall k = 0, 1, \ldots, m-1$,

$$\mathbb{E}\left[\sum_{t=1}^{T} X_t X_t^\top\right]^{-1} \succeq \frac{1}{\sigma_v^2}\text{diag}(\sigma^2_{(1+kn)}, \ldots, \sigma^2_{(n+kn)}). \tag{20}$$

Supposing that $\sigma_{\max} = \max(\sigma_{(1)}, \ldots, \sigma_{(mn)}) > 0$, we can derive a sufficient condition for (20), i.e., $\frac{\sigma^2_{\max}}{\sigma_v^2} I_n \succeq \mathbb{E}\left[\sum_{t=1}^{T} X_t X_t^\top\right]$. It follows that for all $e \in \mathbb{R}^n$ where $\|e\|_2 = 1$, we have that $\mathbb{E}\left[\sum_{t=1}^{T} \|e^\top X_t\|_2^2\right] \leq \frac{\sigma^2_{\max}}{\sigma_v^2}$, which means

$$\|\sum_{t=0}^{T-1} U_t * U_t\|_\infty \leq \frac{\sigma_v^2}{\sigma^2_{\max}}, \ \mathbb{E}\left[\|\sum_{t=0}^{T-1} Y_t * Y_t\|_\infty\right] \leq \frac{\sigma_v^2}{\sigma^2_{\max}}. \tag{21}$$

Hence, we prove the release form of Corollary 1 in (14). Then, we can directly obtain Corollary 1, which is a sufficient condition for (14).

Appendix E. Proof of Corollary 2

Similar to the proof of Corollary 1, we have that$\mathcal{S}$ is Σ-unidentifiable iff (20) holds. Then, we define the matrix formed by the main diagonal elements of matrix of $\mathbb{E}\left[\sum_{t=1}^{T} X_t X_t^\top\right]^{-1}$ as $D(X)$. Then, by expanding the expression of X, we prove Corollary 2.

Appendix F. Proof of Corollary 3

Note that $y_{t,(j)} = \boldsymbol{e}_{(j)}^\top(\sum_{i=1}^{p} A_i^* Y_{t-i} + \sum_{j=1}^{q} B_j^* U_{t-j} + V_t)$, where $\boldsymbol{e}_{(j)} \in \mathbb{R}^m$ is a vector with the i-th element being 1 and the remaining elements being 0. By Assumption 1, we have that $\mathbb{E}[y^2_{t,(j)}] \geq \mathbb{E}[v^2_{t,(j)}] = \delta_v^2$. Hence, we have $\mathbb{E}\left[\sum_{t=0}^{T-1} y^2_{t,(j)}\right] \geq T\sigma_v^2$. Combining this inequality with Corollary 2, Corollary 3 is proved.

References

R. Bellman and K.J. Åström. On structural identifiability. *Mathematical Biosciences*, 7(3-4):329–339, April 1970. ISSN 00255564.

Roland Brun, Peter Reichert, and Hans R. Künsch. Practical identifiability analysis of large environmental simulation models. *Water Resources Research*, 37(4):1015–1030, 2001. ISSN 1944-7973.

Harald Cramér. *Mathematical methods of statistics*, volume 43. Princeton university press, 1999.

Alan Oliveira de Sá, Luiz F. R. da C. Carmo, and Raphael C. S. Machado. Bio-inspired Active System Identification: a Cyber-Physical Intelligence Attack in Networked Control Systems. *Mobile Networks and Applications*, 25(5):1944–1957, October 2020. ISSN 1572-8153.

Marisa C. Eisenberg and Michael A.L. Hayashi. Determining identifiable parameter combinations using subset profiling. *Mathematical Biosciences*, 256:116–126, October 2014. ISSN 00255564.

Maxime Gautier, Alexandre Janot, and Pierre-Olivier Vandanjon. A new closed-loop output error method for parameter identification of robot dynamics. *IEEE Transactions on Control Systems Technology*, 21(2):428–444, 2013.

Jaskaran Grover, Changliu Liu, and Katia Sycara. Parameter identification for multirobot systems using optimization-based controllers. In *2021 International Symposium on Multi-Robot and Multi-Agent Systems (MRS)*, pages 173–180, 2021.

Johan Karlsson, Milena Anguelova, and Mats Jirstrand. An Efficient Method for Structural Identifiability Analysis of Large Dynamic Systems*. *IFAC Proceedings Volumes*, 45(16):941–946, July 2012. ISSN 14746670.

Sahin Lale, Kamyar Azizzadenesheli, Babak Hassibi, and Anima Anandkumar. Finite-time system identification and adaptive control in autoregressive exogenous systems. In *Proceedings of the 3rd Conference on Learning for Dynamics and Control*, volume 144 of *Proceedings of Machine Learning Research*, pages 967–979. PMLR, 07 – 08 June 2021.

V. V. Nguyen and E. F. Wood. Review and Unification of Linear Identifiability Concepts. *SIAM Review*, 24(1):34–51, January 1982. ISSN 0036-1445, 1095-7200.

Tyler Phillips, Hoda Mehrpouyan, John Gardner, and Stephen Reese. A Covert System Identification Attack on Constant Setpoint Control Systems. In *2019 Seventh International Symposium on Computing and Networking Workshops (CANDARW)*, pages 367–373, November 2019.

Thomas J. Rothenberg. Identification in parametric models. *Econometrica*, 39(3):577–591, 1971. ISSN 00129682, 14680262.

Vladimir Stojanovic, Novak Nedic, Dragan Prsic, and Ljubisa Dubonjic. Optimal experiment design for identification of ARX models with constrained output in non-Gaussian noise. *Applied Mathematical Modelling*, 40(13):6676–6689, July 2016. ISSN 0307-904X.

André Teixeira, Iman Shames, Henrik Sandberg, and Karl Henrik Johansson. A secure control framework for resource-limited adversaries. *Automatica*, 51:135–148, 2015. ISSN 0005-1098.

André M. H. Teixeira. Data Injection Attacks against Feedforward Controllers. In *2019 18th European Control Conference (ECC)*, pages 2233–2239, June 2019.

Eric Walter and Yves Lecourtier. Global approaches to identifiability testing for linear and nonlinear state space models. *Mathematics and Computers in Simulation*, 24(6):472–482, December 1982. ISSN 03784754.

Eric Walter and Luc Pronzato. On the identifiability and distinguishability of nonlinear parametric models. *Mathematics and Computers in Simulation*, 42(2):125–134, October 1996. ISSN 0378-4754.

Franz-Georg Wieland, Adrian L. Hauber, Marcus Rosenblatt, Christian Tönsing, and Jens Timmer. On structural and practical identifiability. *Current Opinion in Systems Biology*, 25:60–69, March 2021. ISSN 24523100.

Ye Yuan and Yilin Mo. Security in cyber-physical systems: Controller design against Known-Plaintext Attack. In *2015 54th IEEE Conference on Decision and Control (CDC)*, pages 5814–5819, Osaka, December 2015. IEEE. ISBN 978-1-4799-7886-1.