An efficient algorithm to test the observability of rational nonlinear systems with unmeasured inputs

X. Shi[a,*], M.N. Chatzis[b]

[a] Research Fellow, School of Civil and Environmental Engineering, Nanyang Technological University, Singapore
[b] Associate Professor, Department of Engineering Science, University of Oxford, Oxford, UK

Abstract

This work proposes an efficient algorithm to examine the observability and identifiability of rational nonlinear systems in the presence of unmeasured and unknown inputs. The proposed algorithm allows for determining whether the dynamic states, unknown parameters and unmeasured inputs of a dynamical system can be, in theory, successfully identified from a given set of input-output measurements. The underlying theory of the algorithm is based on a further extension of the recently suggested extended Observability Rank Condition while focusing on rational instead of analytic nonlinearities. For the robust development of the algorithm, a power series based framework is established for computing the observability matrix efficiently. The occurring framework substantially alleviates the computational burden of the standard implementations of the extended Observability Rank Condition approaches, which allows for applications to real-world engineering systems that are often large and complex. Several examples of large-scale and high-complexity engineering structures are used to demonstrate the performance and capability of the algorithm. Furthermore, the proposed algorithm is used to investigate the feasibility of monitoring a sub-system that is independently separated from a full system under the introduction of additional unmeasured inputs.

Keywords: Observability, Identifiability, Unmeasured inputs, Extended

[1] *Corresponding author. E-mail address: xiaodong.shi@ntu.edu.sg

Observability Rank Condition, System identification

1. Introduction

System Identification (SI) is of vital importance in reconstructing and inferring dynamic states, unknown parameters and even unmeasured inputs of dynamical systems from measured input and output signals [1], and in minimising the gaps between real engineering systems and their mathematical models. A fundamental requirement for a successful SI campaign is that these dynamic states, unknown parameters and unmeasured inputs must be observable based on the mathematical description of the system and the given setup of sensors [2]. Otherwise any effort on identifying the unobservable quantities of a dynamical system using any SI method would not be a productive exercise, and could lead to potentially erroneous results which have been proven to be associated with Lie symmetries of the system [3, 4, 5]. As one of the consequences of identifying an unobservable system, engineers may mistakenly adopt the incorrect identification results of those unobservable quantities as their true solutions without gaining the knowledge of their observability properties. To avoid such problematic situations, it is thus necessary to undertake a priori observability testing when proposing an identification campaign. Effective observability analyses would bring further benefits of suggesting how to design the pattern of sensor installations, adjust the mathematical model of the system and/or make suitable model assumptions so as to obtain an observable system.

Methods for examining the theoretical observability of nonlinear systems with fully measured inputs have been well-explored over the past decades. It should be noted that, theoretical observability considered throughout this paper does not take the effect of measurement noise into account, which stands at the opposite of so called practical observability/identifiability [6, 7]. Identifiability is a term used as a particular case of observability to account for whether unknown parameters can be identified. Hermann and Krener [8] introduced the well-known Observability Rank Condition (ORC) to analyse the local weak observability of analytic nonlinear systems. Implementation of the ORC is through rank evaluation of the observability matrix constructed from successive Lie derivative computations [9, 10]. Other approaches exist for various nonlinearities, which can be found in [11] based on exploring the existence of algebraic relations between states and measurements, in [12] based on the power series method, in [13] based on the local state variable isomorphism method, in [14] based on the characteristic set or standard bases computation, in [15] based on the elimination method in differential algebra, etc. Numerous works have also been devoted into specific engineering applications, such as studying the global identifiability of shear-type structures [16], the observability of non-smooth systems [2] and modally reduced order models [17] among others.

On the other hand, joint state-parameter-input estimation has received increasing attention in recent years. The motivation arises from the fact that it is often too difficult, expensive or even impossible to measure all inputs or excitations of a system in practice, for instance during a Structural Health Monitoring (SHM) campaign on a large-scale infrastructure. Research has been oriented toward development of robust identification algorithms with the objective of relaxing the requirement for input measurements [18, 19, 20]. Accordingly, it was not until very recently that the extensions of the existing observability approaches for systems where some or potentially all of their inputs are unmeasured drew attention.

Martinelli [21] introduced the extended ORC or EORC to gain an insight into the observability of input-affine nonlinear systems with unknown inputs. Maes et al. [22] provided a further extension, namely the ORC-DF, where the existence of direct feedthrough in output measurements was taken into consideration. The situation of output measurements affected by direct feedthrough is often encountered in practice, as for example the case of acceleration measurement in structural dynamics. The effectiveness and usefulness of both the EORC and ORC-DF approaches have been demonstrated through a number of classical mechanical and civil engineering models of modest size. However, the default implementations of the EORC and ORC-DF require symbolic computations of the Lie derivatives of an augmented form of a tested system. Such symbolic implementations on a computer of standard specifications are expensive needing tremendously high physical memory, especially when they are used for large and complex systems [23]. The applicability of the EORC and ORC-DF to real-life engineering systems is thus limited to a significant extent. The similar means of state augmentation, as in [21, 22], was adopted by Villaverde et al. [24, 25] to develop a Full Input-State-Parameter Observability algorithm, FISPO, where assumptions were made for unknown inputs being polynomial functions of time up to a specific order to suit the purpose of biological applications. The FISPO algorithm suffers from the same computation issue due to its symbolic nature.

While a considerable amount of works have been focused on tackling the computational complexity problem of observability for the case of all inputs being measured [3, 26, 27, 28], there is still a lack of attention on the case of unmeasured inputs. This paper proposes an efficient algorithm to assess the observability of rational nonlinear systems driven by measured and unmeasured inputs. The underlying theory of the algorithm is based on an extension of the ORC-DF removing the need for systems being affine in inputs. A power series based framework is obtained to construct the observability matrix by making use of an auxiliary variational system of differential equations and the method of Newton's iteration. The occurring framework opens up the route for using random numerical realizations of variables and modular operations, resulting in a semi-numerical implementation. A MATLAB implementation of the proposed algorithm is freely available at the following link in footnote [2]. Superior performance of the algorithm is demonstrated through applications to several suitably chosen examples of engineering models.

[2] https://eng.ox.ac.uk/non-lineardynamics/hidden-resources-mssp-xp2

2. The extended Observability Rank Condition for rational nonlinear systems with unmeasured inputs and direct feedthrough

The ORC-DF presented in [22] introduced an approach to examine the observability of input-affine nonlinear systems with unmeasured inputs and direct feedthrough. This section further relaxes the ORC-DF's requirement for systems being affine in inputs, but more specifically focuses on rational nonlinearities. Without loss of generality, the considered systems herein can be written in the following state-space representation:

$$\begin{aligned} \dot{\boldsymbol{x}} &= \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{\theta}, \boldsymbol{u}, \boldsymbol{w}), \;\; \dot{\boldsymbol{\theta}} = \boldsymbol{0} \\ \boldsymbol{y} &= \boldsymbol{h}(\boldsymbol{x}, \boldsymbol{\theta}, \boldsymbol{u}, \boldsymbol{w}) \end{aligned} \tag{1}$$

where $\boldsymbol{x} = [x_1, ..., x_n]^T$ is the vector of dynamic states, $\boldsymbol{\theta} = [\theta_1, ..., \theta_l]^T$ is the vector of unknown time-invariant parameters, $\boldsymbol{u} = [u_1, ..., u_m]^T$ is the vector of measured inputs, $\boldsymbol{w} = [w_1, ..., w_r]^T$ is the vector of unmeasured and unknown inputs and $\boldsymbol{y} = [y_1, ..., y_p]^T$ is the output measurement vector. The output measurements can be affected by direct feedthrough, that is, they are functions of both the measured and unmeasured inputs. $\boldsymbol{f}$ and $\boldsymbol{h}$ are vectors of rational nonlinear functions, where a rational function is any function defined by a fraction such that both its numerator and denominator are polynomial functions. For the system described in equation (1), it is of importance to determine whether it is theoretically feasible to estimate the unknown parameters $\boldsymbol{\theta}$, and track the dynamic states $\boldsymbol{x}$ and the unmeasured inputs $\boldsymbol{w}$ over time given the input-output $\boldsymbol{u} - \boldsymbol{y}$ measurements.

The basic idea for achieving the derivation of the observability criterion is to augment the considered system by treating the unmeasured inputs and their time derivatives as additional states [22], which then allows for discussion of the property of the Lie derivatives of the augmented system leading to the use of the rank condition from [8]. Following what was done in [22], the system in equation (1) can be augmented by including $\boldsymbol{x}$, $\boldsymbol{\theta}$, $\boldsymbol{w}$ and its time derivatives up to order k in a common state vector, $\boldsymbol{x^k}$:

$$\boldsymbol{x^k} = [\boldsymbol{x}^T, \boldsymbol{\theta}^T, \boldsymbol{w}^T, \dot{\boldsymbol{w}}^T, ..., \boldsymbol{w}^{(k)T}]^T \tag{2}$$

such that the state-space equations with respect to $\boldsymbol{x^k}$ become:

$$\dot{\boldsymbol{x}}^{\boldsymbol{k}} = \begin{bmatrix} \boldsymbol{f(x,\theta,u,w)} \\ \boldsymbol{0}_{l\times 1} \\ \dot{\boldsymbol{w}} \\ \vdots \\ \boldsymbol{w}^{(k)} \\ \boldsymbol{0}_{r\times 1} \end{bmatrix} + \begin{bmatrix} \boldsymbol{0}_{n\times 1} \\ \boldsymbol{0}_{l\times 1} \\ \boldsymbol{0}_{r\times 1} \\ \vdots \\ \boldsymbol{0}_{r\times 1} \\ \boldsymbol{w}^{(k+1)} \end{bmatrix} = \boldsymbol{F^k}(\boldsymbol{x^k}, \boldsymbol{u}, \boldsymbol{w}^{(k+1)}) \tag{3}$$

where $\boldsymbol{w}^{(k)} = \frac{d^k \boldsymbol{w}}{dt^k}$, while the output equations of the system remain the same. It should be noted that the augmented system in equation (3) still contains r unmeasured inputs which now coincide with the $(k+1)^{th}$ order time derivatives of the original unmeasured inputs, i.e. $\boldsymbol{w}^{(k+1)}$.

To proceed with the derivation, it is necessary to define the Lie derivatives, also referred to as the extended Lie derivatives as in [27], of the output functions of the augmented system using the following formulation:

$$\boldsymbol{L_f^i h} = \frac{\partial \boldsymbol{L_f^{i-1} h}}{\partial \boldsymbol{x}} \boldsymbol{f} + \sum_{j=1}^{i} \frac{\partial \boldsymbol{L_f^{i-1} h}}{\partial \boldsymbol{w}^{(j-1)}} \boldsymbol{w}^{(j)} + \sum_{j=1}^{i} \frac{\partial \boldsymbol{L_f^{i-1} h}}{\partial \boldsymbol{u}^{(j-1)}} \boldsymbol{u}^{(j)} \tag{4}$$

where $\boldsymbol{L_f^i h}$ is defined as the i^{th} order (extended) Lie derivative of $\boldsymbol{h}$ associated with the vector field $\boldsymbol{f}$, which can be calculated recursively given the zero-order $\boldsymbol{L_f^0 h} = \boldsymbol{h(x,\theta,u,w)}$. Throughout this paper, the operation $\frac{\partial \boldsymbol{V}}{\partial \boldsymbol{S}}$ for a column vector $\boldsymbol{V} = [v_1, v_2, ..., v_n]^T$ with respect to a column or row vector $\boldsymbol{S} = [s_1, s_2, ..., s_m]$ refers to the Jacobian:

$$\frac{\partial \boldsymbol{V}}{\partial \boldsymbol{S}} = \begin{bmatrix} \frac{\partial v_1}{\partial s_1} & \frac{\partial v_1}{\partial s_2} & \cdots & \frac{\partial v_1}{\partial s_m} \\ \frac{\partial v_2}{\partial s_1} & \frac{\partial v_2}{\partial s_2} & \cdots & \frac{\partial v_2}{\partial s_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial v_n}{\partial s_1} & \frac{\partial v_n}{\partial s_2} & \cdots & \frac{\partial v_n}{\partial s_m} \end{bmatrix} \tag{5}$$

It should be noted that the Lie derivatives herein are introduced for smooth $\boldsymbol{u}$, i.e. infinitely differentiable with respect to time, and thus involve the time derivatives of $\boldsymbol{u}$, while [22] expressed $\boldsymbol{u}$ as piecewise-constant over infinitesimally short intervals. The two representations of $\boldsymbol{u}$ are equivalent to yield identical observability results, since the measurements of input signals will be treated as infinitely fast.

A straightforward calculation based on the chain rule shows that the i^{th} order Lie derivative is equivalent to:

$$\boldsymbol{L_f^i h} = \frac{d\boldsymbol{L_f^{i-1} h}}{dt} = \frac{d^i \boldsymbol{h}}{dt^i} \tag{6}$$

Let $\boldsymbol{x_0^k} = [\boldsymbol{x_0}^T, \boldsymbol{\theta}^T, \boldsymbol{w_0}^T, \dot{\boldsymbol{w}}_{\boldsymbol{0}}{}^T, ..., \boldsymbol{w_0}^{(k)T}]^T$ denote the initial condition of the augmented state vector $\boldsymbol{x^k}$ at $t = 0$, where $\boldsymbol{x_0}$ and $\boldsymbol{w_0}^{(k)}$ are respectively the initial conditions of $\boldsymbol{x}$ and $\boldsymbol{w}^{(k)}$. The flow $\boldsymbol{\Phi}(t, \boldsymbol{x_0^k}, \boldsymbol{u}, \boldsymbol{w}^{(k+1)})$ of the vector field $\boldsymbol{F^k}$ is used to describe the trajectory of $\boldsymbol{x^k}$ over time, which satisfies:

$$\frac{d\boldsymbol{\Phi}}{dt} = \boldsymbol{F^k}(\boldsymbol{\Phi}, \boldsymbol{u}, \boldsymbol{w}^{(k+1)}), \;\; \boldsymbol{\Phi}(0, \boldsymbol{x_0^k}, \boldsymbol{u}, \boldsymbol{w}^{(k+1)}) = \boldsymbol{x_0^k} \tag{7}$$

Let $\boldsymbol{\Phi_x}(t, \boldsymbol{x_0^k}, \boldsymbol{u}, \boldsymbol{w}^{(k+1)})$ and $\boldsymbol{\Phi_w}(t, \boldsymbol{x_0^k}, \boldsymbol{u}, \boldsymbol{w}^{(k+1)})$ denote the components of $\boldsymbol{\Phi}$ used to describe respectively the trajectories of $\boldsymbol{x}$ and $\boldsymbol{w}$, with $\boldsymbol{\Phi_x}(0, \boldsymbol{x_0^k}, \boldsymbol{u}, \boldsymbol{w}^{(k+1)}) = \boldsymbol{x_0}$ and $\boldsymbol{\Phi_w}(0, \boldsymbol{x_0^k}, \boldsymbol{u}, \boldsymbol{w}^{(k+1)}) = \boldsymbol{w_0}$. Then for convenience, $\boldsymbol{x}$ and $\boldsymbol{w}$ reached at time t starting from the initial condition may be expressed as:

$$\begin{aligned} \boldsymbol{x}(t; \boldsymbol{x_0^k}) &= \boldsymbol{\Phi_x}(t, \boldsymbol{x_0^k}, \boldsymbol{u}, \boldsymbol{w}^{(k+1)}) \\ \boldsymbol{w}(t; \boldsymbol{x_0^k}) &= \boldsymbol{\Phi_w}(t, \boldsymbol{x_0^k}, \boldsymbol{u}, \boldsymbol{w}^{(k+1)}) \end{aligned} \tag{8}$$

Consider $t \in [0, t_1 + t_2 + ... + t_i]$ which consists of i small time intervals, i.e. $[0, t_1]$, $[t_1, t_1 + t_2]$, ..., $[t_1 + ... + t_{i-1}, t_1 + ... + t_i]$ with $0 \leq t_i < \epsilon$. The system outputs at $t = t_1 + t_2 + ... + t_i$ can thus be expressed as:

$$\boldsymbol{y}(t_1 + ... + t_i) = \boldsymbol{h}(\boldsymbol{x}(t_1 + ... + t_i; \boldsymbol{x_0^k}), \boldsymbol{\theta}, \boldsymbol{u}(t_1 + ... + t_i), \boldsymbol{w}(t_1 + ... + t_i; \boldsymbol{x_0^k})) \tag{9}$$

Indistinguishable states in the presence of unmeasured inputs are defined in [21] as a pair of initial conditions, $\boldsymbol{x_{0,a}^k}$ and $\boldsymbol{x_{0,b}^k}$, producing identical system outputs for any measured inputs $\boldsymbol{u}$, under in general two different vectors of unmeasured inputs. For any i, the indistinguishable states must satisfy:

$$\begin{aligned} &\boldsymbol{h}\big(\boldsymbol{x}(t_1 + ... + t_i; \boldsymbol{x_{0,a}^k}), \boldsymbol{\theta_a}, \boldsymbol{u}(t_1 + ... + t_i), \boldsymbol{w}(t_1 + ... + t_i; \boldsymbol{x_{0,a}^k})\big) \\ =&\boldsymbol{h}(\boldsymbol{x}\big(t_1 + ... + t_i; \boldsymbol{x_{0,b}^k}), \boldsymbol{\theta_b}, \boldsymbol{u}(t_1 + ... + t_i), \boldsymbol{w}(t_1 + ... + t_i; \boldsymbol{x_{0,b}^k})\big) \end{aligned} \tag{10}$$

for all possible $t_1, ..., t_i$. From this differentiating equation (10) with respect to $t_1, ..., t_i$ gives:

$$\left.\frac{d^i \boldsymbol{h}\big(\boldsymbol{x}(t_1 + ... + t_i; \boldsymbol{x_{0,a}^k}), \boldsymbol{\theta_a}, \boldsymbol{u}(t_1 + ... + t_i), \boldsymbol{w}(t_1 + ... + t_i; \boldsymbol{x_{0,a}^k})\big)}{dt_1 dt_2 ... dt_i}\right|_{t_1,...,t_i \to 0}$$
$$= \left.\frac{d^i \boldsymbol{h}\big(\boldsymbol{x}(t_1 + ... + t_i; \boldsymbol{x_{0,b}^k}), \boldsymbol{\theta_b}, \boldsymbol{u}(t_1 + ... + t_i), \boldsymbol{w}(t_1 + ... + t_i; \boldsymbol{x_{0,b}^k})\big)}{dt_1 dt_2 ... dt_i}\right|_{t_1,...,t_i \to 0} \tag{11}$$

Through successive applications of the chain rule, the following equality can be obtained for $i = 0, ..., k$:

$$\left.\frac{d^i \boldsymbol{h}\big(\boldsymbol{x}(t_1 + ... + t_i; \boldsymbol{x_0^k}), \boldsymbol{\theta}, \boldsymbol{u}(t_1 + ... + t_i), \boldsymbol{w}(t_1 + ... + t_i; \boldsymbol{x_0^k})\big)}{dt_1 dt_2 ... dt_i}\right|_{t_1,...,t_i \to 0}$$
$$= \left.\boldsymbol{L_f^i h}\right|_{\boldsymbol{x^k} = \boldsymbol{x_0^k}, \boldsymbol{u}, ..., \boldsymbol{u}^{(k)} = \boldsymbol{u_0}, ..., \boldsymbol{u_0}^{(k)}} \tag{12}$$

where $\boldsymbol{u_0}^{(k)}$ is the initial condition of $\boldsymbol{u}^{(k)}$ at $t = 0$. It should be noted that the Lie derivatives up to order k are independent of $\boldsymbol{w}^{(k+1)}$ and its higher order time derivatives, as can be easily deduced from equation (4). Equation (11) therefore becomes:

$$\left.\boldsymbol{L_f^i h}\right|_{\boldsymbol{x^k} = \boldsymbol{x_{0,a}^k}, \boldsymbol{u}, ..., \boldsymbol{u}^{(k)} = \boldsymbol{u_0}, ..., \boldsymbol{u_0}^{(k)}}$$
$$= \left.\boldsymbol{L_f^i h}\right|_{\boldsymbol{x^k} = \boldsymbol{x_{0,b}^k}, \boldsymbol{u}, ..., \boldsymbol{u}^{(k)} = \boldsymbol{u_0}, ..., \boldsymbol{u_0}^{(k)}} \tag{13}$$

for $i = 0, ..., k$. For the sake of brevity, equation (13) may be written as $\boldsymbol{L_f^i h}(\boldsymbol{x_{0,a}^k}) = \boldsymbol{L_f^i h}(\boldsymbol{x_{0,b}^k})$, which implies that the Lie derivatives of the augmented system up to order k are equal when evaluated for indistinguishable states.

Similar as in [22], the afore-shown property of the Lie derivatives allows for applying the rank condition from [8] (see Theorem 3.1 and Lemma 3.2) to establish the observability criterion for the augmented system in equation (3).

Arrange the Lie derivatives for $i = 0, ..., k$ in a column vector:

$$\boldsymbol{\Omega^k} = \begin{bmatrix} \boldsymbol{L_f^0 h}(\boldsymbol{x_0^k}) \\ \boldsymbol{L_f h}(\boldsymbol{x_0^k}) \\ \vdots \\ \boldsymbol{L_f^k h}(\boldsymbol{x_0^k}) \end{bmatrix} \tag{14}$$

The Jacobian of $\boldsymbol{\Omega^k}$ with respect to $\boldsymbol{x_0^k}$ results in the so called k-row observability matrix $d\boldsymbol{\Omega^k}$:

$$d\boldsymbol{\Omega^k} = \frac{\partial \boldsymbol{\Omega^k}}{\partial \boldsymbol{x_0^k}} = [d\boldsymbol{\Omega^k_x}, d\boldsymbol{\Omega^k_\theta}, d\boldsymbol{\Omega^k_{w_1}}, ..., d\boldsymbol{\Omega^k_{w_r}}] \tag{15}$$

where

$$d\boldsymbol{\Omega^k_x} = \frac{\partial \boldsymbol{\Omega^k}}{\partial \boldsymbol{x_0}} = \begin{bmatrix} \frac{\partial \boldsymbol{L_f^0 h}}{\partial x_{1,0}} & \cdots & \frac{\partial \boldsymbol{L_f^0 h}}{\partial x_{n,0}} \\ \vdots & \ddots & \vdots \\ \frac{\partial \boldsymbol{L_f^k h}}{\partial x_{1,0}} & \cdots & \frac{\partial \boldsymbol{L_f^k h}}{\partial x_{n,0}} \end{bmatrix}, \quad d\boldsymbol{\Omega^k_\theta} = \frac{\partial \boldsymbol{\Omega^k}}{\partial \boldsymbol{\theta}} = \begin{bmatrix} \frac{\partial \boldsymbol{L_f^0 h}}{\partial \theta_1} & \cdots & \frac{\partial \boldsymbol{L_f^0 h}}{\partial \theta_l} \\ \vdots & \ddots & \vdots \\ \frac{\partial \boldsymbol{L_f^k h}}{\partial \theta_1} & \cdots & \frac{\partial \boldsymbol{L_f^k h}}{\partial \theta_l} \end{bmatrix}$$

$$d\boldsymbol{\Omega^k_{w_j}} = \begin{bmatrix} \frac{\partial \boldsymbol{L_f^0 h}}{\partial w_{j,0}} & 0 & \cdots & 0 \\ \frac{\partial \boldsymbol{L_f h}}{\partial w_{j,0}} & \frac{\partial \boldsymbol{L_f h}}{\partial \dot{w}_{j,0}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \boldsymbol{L_f^k h}}{\partial w_{j,0}} & \frac{\partial \boldsymbol{L_f^k h}}{\partial \dot{w}_{j,0}} & \cdots & \frac{\partial \boldsymbol{L_f^k h}}{\partial w_{j,0}^{(k)}} \end{bmatrix} \quad (j = 1, ..., r) \tag{16}$$

If $d\boldsymbol{\Omega^k}$ is a full-rank matrix, i.e. $rank(d\boldsymbol{\Omega^k}) = n + l + (k+1)r$, the Inverse Function Theorem [29] ensures that within the neighbourhood of $\boldsymbol{x_0^k}$, there does not exist any indistinguishable state which results in identical outputs and Lie derivatives as $\boldsymbol{x_0^k}$. This consequently means that it is possible to distinguish and observe $\boldsymbol{x_0^k}$ from its neighbours using the given input-output measurements. Following the above fact, the observability criterion can be concluded: the augmented system in equation (3) is said to satisfy the extended Observability Rank Condition such that the augmented state vector $\boldsymbol{x^k}$ is (locally weakly) observable, if and only if there exists a k such that $d\boldsymbol{\Omega^k}$ is of full-rank for all possible $\boldsymbol{x_0^k} \in \mathbb{R}^{n+l+(k+1)r}$ (normally except for some singular realizations of $\boldsymbol{x_0^k}$ where ranks of the matrix would be reduced). If $\boldsymbol{x^k}$ is observable, then $\boldsymbol{x}$, $\boldsymbol{\theta}$ and $\boldsymbol{w}$ are equivalently observable as its components and therefore the considered system in equation (1) is observable. From a practical standpoint and for the purposes of this paper, further distinctions between the concepts of locally weakly observable, locally observable, weakly observable and globally observable, as in [8], are not considered herein.

If the k-row observability matrix for a chosen $k = k_0$ is not a full-rank matrix, i.e. $rank(d\boldsymbol{\Omega^k}) < n + l + (k+1)r$, it is then of interest to detect the observability property of each individual component of $\boldsymbol{x^k}$. This can be achieved by removing the i^{th} $(i = 1, ..., n + l + (k+1)r)$ column of $d\boldsymbol{\Omega^k}$: if the rank of the occurring matrix is smaller than the rank of $d\boldsymbol{\Omega^k}$, then the i^{th} state (dynamic state/parameter/unmeasured input) of $\boldsymbol{x^k}$ is observable; otherwise, the state is so-called k-row unobservable as defined in [22]. It should be noted that a k-row unobservable state might be detected observable or remain k-row unobservable for a larger $k > k_0$, while an (k-row) observable state can be concluded immediately. For this reason, the suggested implementation of the ORC-DF involves iterative increment of k starting from k_0 ($k_0 = 0$ in [22]) with repeated rank evaluation and observability detection of the corresponding observability matrix at each increment of k. In theory, such increment is continued until the detected observability properties of all the components of $\boldsymbol{x^k}$ remain invariant, i.e. converge, for any larger k and thus conclude the observability results.

The following algorithm summarizes the procedure of observability testing for the nonlinear system in equation (1).

Algorithm I

Initialization: Set $k = 0$, $\boldsymbol{x_0^k} = [\boldsymbol{x_0}^T, \boldsymbol{\theta}^T, \boldsymbol{w_0}^T]^T$, $\boldsymbol{L_f^k h}(\boldsymbol{x_0^k}) = \boldsymbol{h(x_0, \theta, u_0, w_0)}$, $d\boldsymbol{\Omega^k} = \frac{\partial \boldsymbol{L_f^k h}(\boldsymbol{x_0^k})}{\partial \boldsymbol{x_0^k}}$.

1. Set $k = k + 1$;

2. Set $\boldsymbol{x_0^k} = [\boldsymbol{x_0^{k-1}}^T, \boldsymbol{w_0}^{(k)T}]^T$;

3. Compute $\boldsymbol{L_f^k h} = \frac{\partial \boldsymbol{L_f^{k-1} h}}{\partial \boldsymbol{x}}\boldsymbol{f} + \sum_{j=1}^{k} \frac{\partial \boldsymbol{L_f^{k-1} h}}{\partial \boldsymbol{w}^{(j-1)}}\boldsymbol{w}^{(j)} + \sum_{j=1}^{k} \frac{\partial \boldsymbol{L_f^{k-1} h}}{\partial \boldsymbol{u}^{(j-1)}}\boldsymbol{u}^{(j)}$;

4. Compute and arrange $d\boldsymbol{\Omega^k} = \begin{bmatrix} d\boldsymbol{\Omega^{k-1}} & \boldsymbol{0} \\ \multicolumn{2}{c}{\frac{\partial \boldsymbol{L_f^k h}(\boldsymbol{x_0^k})}{\partial \boldsymbol{x_0^k}}} \end{bmatrix}$;

5. Compute the rank of $d\boldsymbol{\Omega^k}$, and if $rank(d\boldsymbol{\Omega^k}) < n + l + (k+1)r$, detect the k-row observability of $\boldsymbol{x}$, $\boldsymbol{\theta}$ and $\boldsymbol{w}$;

6. End if $rank(d\boldsymbol{\Omega^k}) = n + l + (k+1)r$, or $\boldsymbol{x}$, $\boldsymbol{\theta}$ and $\boldsymbol{w}$ are observable, or the detected observability of $\boldsymbol{x}$, $\boldsymbol{\theta}$ and $\boldsymbol{w}$ has been convergent;

7. Go to step **1**;

The above algorithm should be implemented symbolically, i.e. symbols are used as generic representations of the involved variables in the algorithm. Such a symbolic implementation is computationally cumbersome and suffers from the issues related to high physical memory requirements and low processing speed, especially when used for large systems [23]. Even if a tested system is of modest size, the size of the augmented form of the system for a large k might still result in an intractable computation problem on a standard computer. Therefore in order to maximise the applicability of the observability testing approach to real-world engineering systems that are often large and complex, a more computationally efficient framework for implementing the approach, more specifically for calculating the observability matrix, is developed and presented in the following section.

3. Efficient algorithm

The computation and rank evaluation of $d\boldsymbol{\Omega^k}$ could be implemented numerically by using random values as the generic representations of the involved variables, which would significantly save computational resources. This can be done by making:

$$\boldsymbol{x_0} = \tilde{\boldsymbol{x}}_{\boldsymbol{0}},\ \ \boldsymbol{\theta} = \tilde{\boldsymbol{\theta}},\ \ \boldsymbol{w_0}, ..., \boldsymbol{w_0}^{(k)} = \tilde{\boldsymbol{w}}_{\boldsymbol{0}}, ..., \tilde{\boldsymbol{w}}_{\boldsymbol{0}}{}^{(k)},\ \ \boldsymbol{u_0}, ..., \boldsymbol{u_0}^{(k)} = \tilde{\boldsymbol{u}}_{\boldsymbol{0}}, ..., \tilde{\boldsymbol{u}}_{\boldsymbol{0}}{}^{(k)} \tag{17}$$

where a variable with $\tilde{.}$ stands for a set of nonzero numbers. These numbers are randomly chosen in order to avoid the aforementioned singular realizations of $\boldsymbol{x_0^k}$. This section presents a power series-based framework which allows for using such random numerical realizations to calculate the elements of $d\boldsymbol{\Omega^k}$ without the need to perform the repeated Jacobian operations.

Making use of equation (6), the following power series expansion of $\boldsymbol{h}$ in t

truncated at order k is considered:

$$\boldsymbol{L_f^0 h(x_0^k)} + \boldsymbol{L_f h(x_0^k)} t + ... + \boldsymbol{L_f^k h(x_0^k)} \frac{t^k}{k!} + O(t^{k+1}) \tag{18}$$

Let $\boldsymbol{w_{j,0}^k}$ denote the initial condition of the vector $\boldsymbol{w_j^k} = [w_j, ..., w_j^{(k)}]$, and taking derivatives of equation (18) with respect to $\boldsymbol{x_0}$, $\boldsymbol{\theta}$ and $\boldsymbol{w_{j,0}^k}$ respectively yields the power series expansions of $\frac{d\boldsymbol{h}}{d\boldsymbol{x_0}}$, $\frac{d\boldsymbol{h}}{d\boldsymbol{\theta}}$ and $\frac{d\boldsymbol{h}}{d\boldsymbol{w_{j,0}^k}}$:

$$\begin{gathered}
\frac{\partial \boldsymbol{L_f^0 h(x_0^k)}}{\partial \boldsymbol{x_0}} + \frac{\partial \boldsymbol{L_f h(x_0^k)}}{\partial \boldsymbol{x_0}} t + ... + \frac{\partial \boldsymbol{L_f^k h(x_0^k)}}{\partial \boldsymbol{x_0}} \frac{t^k}{k!} + O(t^{k+1}) \\
\frac{\partial \boldsymbol{L_f^0 h(x_0^k)}}{\partial \boldsymbol{\theta}} + \frac{\partial \boldsymbol{L_f h(x_0^k)}}{\partial \boldsymbol{\theta}} t + ... + \frac{\partial \boldsymbol{L_f^k h(x_0^k)}}{\partial \boldsymbol{\theta}} \frac{t^k}{k!} + O(t^{k+1}) \\
\frac{\partial \boldsymbol{L_f^0 h(x_0^k)}}{\partial \boldsymbol{w_{j,0}^k}} + \frac{\partial \boldsymbol{L_f h(x_0^k)}}{\partial \boldsymbol{w_{j,0}^k}} t + ... + \frac{\partial \boldsymbol{L_f^k h(x_0^k)}}{\partial \boldsymbol{w_{j,0}^k}} \frac{t^k}{k!} + O(t^{k+1})
\end{gathered} \tag{19}$$

Comparing equation (19) to equations (15) and (16), it is immediately observed that the coefficients of the power series correspond to the elements of $d\boldsymbol{\Omega^k}$. In particular, the coefficients of the power series of $\frac{d\boldsymbol{h}}{d\boldsymbol{x_0}}$, $\frac{d\boldsymbol{h}}{d\boldsymbol{\theta}}$ and $\frac{d\boldsymbol{h}}{d\boldsymbol{w_{j,0}^k}}$ correspond to the elements of $d\boldsymbol{\Omega_x^k}$, $d\boldsymbol{\Omega_\theta^k}$ and $d\boldsymbol{\Omega_{w_j}^k}$ respectively. This suggests that finding a means of computing the power series in equation (19) can lead to the construction of $d\boldsymbol{\Omega^k}$. For the sake of brevity, in the rest of this paper the power series expansion of any variable z in t, truncated at order k, will be denoted by the notation $z[t,k]$, i.e.:

$$z[t,k] = z_0 + z_1 t + ... + z_k t^k \tag{20}$$

where $z_0, ..., z_k$ are constant coefficients.

Applying the chain rule to the power series in equation (19), now denoted as $\frac{d\boldsymbol{h}}{d\boldsymbol{x_0}}[t,k]$, $\frac{d\boldsymbol{h}}{d\boldsymbol{\theta}}[t,k]$ and $\frac{d\boldsymbol{h}}{d\boldsymbol{w_{j,0}^k}}[t,k]$, the following equations are obtained:

$$
\begin{aligned}
\frac{d\boldsymbol{h}}{d\boldsymbol{x_0}}[t,k] &= (\frac{\partial\boldsymbol{h}}{\partial\boldsymbol{x}}\frac{\partial\boldsymbol{x}}{\partial\boldsymbol{x_0}})[t,k] \\
&= \frac{\partial\boldsymbol{h}}{\partial\boldsymbol{x}}[t,k]\frac{\partial\boldsymbol{x}}{\partial\boldsymbol{x_0}}[t,k] \ \ (mod\ t^{k+1}) \\
\frac{d\boldsymbol{h}}{d\boldsymbol{\theta}}[t,k] &= (\frac{\partial\boldsymbol{h}}{\partial\boldsymbol{x}}\frac{\partial\boldsymbol{x}}{\partial\boldsymbol{\theta}}+\frac{\partial\boldsymbol{h}}{\partial\boldsymbol{\theta}})[t,k] \\
&= \frac{\partial\boldsymbol{h}}{\partial\boldsymbol{x}}[t,k]\frac{\partial\boldsymbol{x}}{\partial\boldsymbol{\theta}}[t,k]+\frac{\partial\boldsymbol{h}}{\partial\boldsymbol{\theta}}[t,k] \ \ (mod\ t^{k+1}) \\
\frac{d\boldsymbol{h}}{d\boldsymbol{w_{j,0}^k}}[t,k] &= (\frac{\partial\boldsymbol{h}}{\partial\boldsymbol{x}}\frac{\partial\boldsymbol{x}}{\partial\boldsymbol{w_{j,0}^k}}+\frac{\partial\boldsymbol{h}}{\partial w_j}\frac{\partial w_j}{\partial\boldsymbol{w_{j,0}^k}})[t,k] \\
&= \frac{\partial\boldsymbol{h}}{\partial\boldsymbol{x}}[t,k]\frac{\partial\boldsymbol{x}}{\partial\boldsymbol{w_{j,0}^k}}[t,k]+\frac{\partial\boldsymbol{h}}{\partial w_j}[t,k]\frac{\partial w_j}{\partial\boldsymbol{w_{j,0}^k}}[t,k] \ \ (mod\ t^{k+1})
\end{aligned}
\tag{21}
$$

where $mod\ t^{k+1}$ denotes the truncation of the occurring power series at order k. Equation (21) can then be used to calculate $\frac{d\boldsymbol{h}}{d\boldsymbol{x_0}}[t,k]$, $\frac{d\boldsymbol{h}}{d\boldsymbol{\theta}}[t,k]$ and $\frac{d\boldsymbol{h}}{d\boldsymbol{w_{j,0}^k}}[t,k]$ through obtaining $\frac{\partial\boldsymbol{h}}{\partial\boldsymbol{x}}[t,k]$, $\frac{\partial\boldsymbol{h}}{\partial\boldsymbol{\theta}}[t,k]$, $\frac{\partial\boldsymbol{h}}{\partial w_j}[t,k]$, $\frac{\partial\boldsymbol{x}}{\partial\boldsymbol{x_0}}[t,k]$, $\frac{\partial\boldsymbol{x}}{\partial\boldsymbol{\theta}}[t,k]$, $\frac{\partial\boldsymbol{x}}{\partial\boldsymbol{w_{j,0}^k}}[t,k]$ and $\frac{\partial w_j}{\partial\boldsymbol{w_{j,0}^k}}[t,k]$. $\frac{\partial\boldsymbol{h}}{\partial\boldsymbol{x}}$, $\frac{\partial\boldsymbol{h}}{\partial\boldsymbol{\theta}}$ and $\frac{\partial\boldsymbol{h}}{\partial w_j}$ can be calculated symbolically and are generally expressions of $\boldsymbol{x}$, $\boldsymbol{\theta}$, $\boldsymbol{u}$ and $\boldsymbol{w}$. Their power series expansions $\frac{\partial\boldsymbol{h}}{\partial\boldsymbol{x}}[t,k]$, $\frac{\partial\boldsymbol{h}}{\partial\boldsymbol{\theta}}[t,k]$ and $\frac{\partial\boldsymbol{h}}{\partial w_j}[t,k]$, under the random numerical realizations, can be obtained by substituting $\boldsymbol{x}=\boldsymbol{x}[t,k]$, $\boldsymbol{\theta}=\tilde{\boldsymbol{\theta}}$, $\boldsymbol{u}=\boldsymbol{u}[t,k]$ and $\boldsymbol{w}=\boldsymbol{w}[t,k]$ and truncating at order k. It should be noted that if $\frac{\partial\boldsymbol{h}}{\partial\boldsymbol{x}}$, $\frac{\partial\boldsymbol{h}}{\partial\boldsymbol{\theta}}$ and $\frac{\partial\boldsymbol{h}}{\partial w_j}$ are polynomial expressions, such substitutions result directly in $\frac{\partial\boldsymbol{h}}{\partial\boldsymbol{x}}[t,k]$, $\frac{\partial\boldsymbol{h}}{\partial\boldsymbol{\theta}}[t,k]$ and $\frac{\partial\boldsymbol{h}}{\partial w_j}[t,k]$; if they are rational expressions, obtaining $\frac{\partial\boldsymbol{h}}{\partial\boldsymbol{x}}[t,k]$, $\frac{\partial\boldsymbol{h}}{\partial\boldsymbol{\theta}}[t,k]$ and $\frac{\partial\boldsymbol{h}}{\partial w_j}[t,k]$ may require additional power/Taylor series expansion operations. Amongst all the power series needed to calculate $\frac{d\boldsymbol{h}}{d\boldsymbol{x_0}}[t,k]$, $\frac{d\boldsymbol{h}}{d\boldsymbol{\theta}}[t,k]$ and $\frac{d\boldsymbol{h}}{d\boldsymbol{w_{j,0}^k}}[t,k]$ in equation (21), $\boldsymbol{u}[t,k]$, $\boldsymbol{w}[t,k]$ and $\frac{\partial w_j}{\partial\boldsymbol{w_{j,0}^k}}[t,k]$ are readily obtained as:

$$
\begin{aligned}
\boldsymbol{u}[t,k] = \tilde{\boldsymbol{u}}_{\boldsymbol{0}} + ... + \tilde{\boldsymbol{u}}_{\boldsymbol{0}}{}^{(k)}\frac{t^k}{k!},\ \ \boldsymbol{w}[t,k] = \tilde{\boldsymbol{w}}_{\boldsymbol{0}} + ... + \tilde{\boldsymbol{w}}_{\boldsymbol{0}}{}^{(k)}\frac{t^k}{k!} \\
\frac{\partial w_j}{\partial\boldsymbol{w_{j,0}^k}}[t,k] = [1, t, ..., \frac{t^k}{k!}]
\end{aligned}
\tag{22}
$$

The remaining unknowns $\boldsymbol{x}[t,k]$, $\frac{\partial\boldsymbol{x}}{\partial\boldsymbol{x_0}}[t,k]$, $\frac{\partial\boldsymbol{x}}{\partial\boldsymbol{\theta}}[t,k]$ and $\frac{\partial\boldsymbol{x}}{\partial\boldsymbol{w_{j,0}^k}}[t,k]$ can be calculated indirectly by making use of an auxiliary system of ordinary differential equations.

3.1. Variational system

This subsection defines an auxiliary system of ordinary differential equations for deriving a means of calculating $\boldsymbol{x}[t,k]$, $\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{x_0}}[t,k]$, $\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{\theta}}[t,k]$ and $\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{w_{j,0}^k}}[t,k]$. The system of ODEs, so-called the variational system, is an extension of the equations derived in [3] for rational nonlinear systems with fully measured inputs, i.e. the number of unmeasured inputs $r = 0$. Similar to what was done in [3], the rational functions $\boldsymbol{f}$ are expressed as fractions with their numerators and denominators being polynomial functions:

$$\boldsymbol{f(x,\theta,u,w)} = \begin{bmatrix} \frac{f_{nu,1}}{f_{de,1}} \\ \vdots \\ \frac{f_{nu,n}}{f_{de,n}} \end{bmatrix} = \begin{bmatrix} \frac{1}{f_{de,1}} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \frac{1}{f_{de,n}} \end{bmatrix} \begin{bmatrix} f_{nu,1} \\ \vdots \\ f_{nu,n} \end{bmatrix} \tag{23}$$

where $f_{nu,1}$, ..., $f_{nu,n}$ and $f_{de,1}$, ..., $f_{de,n}$ are all polynomial functions of $\boldsymbol{x}$, $\boldsymbol{\theta}$, $\boldsymbol{u}$ and $\boldsymbol{w}$. The state-space equations of the system in (1) can then be written as a set of polynomial ordinary differential equations, denoted by $\boldsymbol{P} = \boldsymbol{0}$, in the following form:

$$\boldsymbol{P(x,\dot{x},\theta,u,w)} = \begin{bmatrix} f_{de,1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & f_{de,n} \end{bmatrix} \begin{bmatrix} \dot{x}_1 \\ \vdots \\ \dot{x}_n \end{bmatrix} - \begin{bmatrix} f_{nu,1} \\ \vdots \\ f_{nu,n} \end{bmatrix} = \boldsymbol{0}_{n\times 1} \tag{24}$$

Taking derivatives of $\boldsymbol{P} = \boldsymbol{0}$ with respect to $\boldsymbol{x_0}$, $\boldsymbol{\theta}$ and $\boldsymbol{w_{j,0}^k}$ respectively yields:

$$\begin{aligned} \frac{d\boldsymbol{P}}{d\boldsymbol{x_0}} &= \frac{\partial \boldsymbol{P}}{\partial \dot{\boldsymbol{x}}}\frac{\partial \dot{\boldsymbol{x}}}{\partial \boldsymbol{x_0}} + \frac{\partial \boldsymbol{P}}{\partial \boldsymbol{x}}\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{x_0}} = \boldsymbol{0}_{n\times n} \\ \frac{d\boldsymbol{P}}{d\boldsymbol{\theta}} &= \frac{\partial \boldsymbol{P}}{\partial \dot{\boldsymbol{x}}}\frac{\partial \dot{\boldsymbol{x}}}{\partial \boldsymbol{\theta}} + \frac{\partial \boldsymbol{P}}{\partial \boldsymbol{x}}\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{\theta}} + \frac{\partial \boldsymbol{P}}{\partial \boldsymbol{\theta}} = \boldsymbol{0}_{n\times l} \\ \frac{d\boldsymbol{P}}{d\boldsymbol{w_{j,0}^k}} &= \frac{\partial \boldsymbol{P}}{\partial \dot{\boldsymbol{x}}}\frac{\partial \dot{\boldsymbol{x}}}{\partial \boldsymbol{w_{j,0}^k}} + \frac{\partial \boldsymbol{P}}{\partial \boldsymbol{x}}\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{w_{j,0}^k}} + \frac{\partial \boldsymbol{P}}{\partial w_j}\frac{\partial w_j}{\partial \boldsymbol{w_{j,0}^k}} = \boldsymbol{0}_{n\times k} \end{aligned} \tag{25}$$

Using equations (24) and (25), the variational system, $\nabla \boldsymbol{P}$, is defined as the following set of differential equations:

$$\nabla \boldsymbol{P}(\boldsymbol{x}, \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{x_0}}, \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{\theta}}, \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{w}^{\boldsymbol{k}}_{\boldsymbol{j,0}}}, \boldsymbol{\theta}, \boldsymbol{u}, \boldsymbol{w}) : \begin{cases} \boldsymbol{P}(\boldsymbol{x}, \dot{\boldsymbol{x}}, \boldsymbol{\theta}, \boldsymbol{u}, \boldsymbol{w}) = \mathbf{0}_{n\times 1} \\ \frac{\partial \boldsymbol{P}}{\partial \dot{\boldsymbol{x}}} \frac{\partial \dot{\boldsymbol{x}}}{\partial \boldsymbol{x_0}} + \frac{\partial \boldsymbol{P}}{\partial \boldsymbol{x}} \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{x_0}} = \mathbf{0}_{n\times n} \\ \frac{\partial \boldsymbol{P}}{\partial \dot{\boldsymbol{x}}} \frac{\partial \dot{\boldsymbol{x}}}{\partial \boldsymbol{\theta}} + \frac{\partial \boldsymbol{P}}{\partial \boldsymbol{x}} \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{\theta}} + \frac{\partial \boldsymbol{P}}{\partial \boldsymbol{\theta}} = \mathbf{0}_{n\times l} \\ \frac{\partial \boldsymbol{P}}{\partial \dot{\boldsymbol{x}}} \frac{\partial \dot{\boldsymbol{x}}}{\partial \boldsymbol{w}^{\boldsymbol{k}}_{\boldsymbol{j,0}}} + \frac{\partial \boldsymbol{P}}{\partial \boldsymbol{x}} \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{w}^{\boldsymbol{k}}_{\boldsymbol{j,0}}} + \frac{\partial \boldsymbol{P}}{\partial w_j} \frac{\partial w_j}{\partial \boldsymbol{w}^{\boldsymbol{k}}_{\boldsymbol{j,0}}} = \mathbf{0}_{n\times k} \end{cases} \tag{26}$$

where $\frac{\partial \boldsymbol{P}}{\partial \dot{\boldsymbol{x}}}$, $\frac{\partial \boldsymbol{P}}{\partial \boldsymbol{x}}$, $\frac{\partial \boldsymbol{P}}{\partial \boldsymbol{\theta}}$ and $\frac{\partial \boldsymbol{P}}{\partial w_j}$ can be calculated symbolically and are generally expressions of $\boldsymbol{x}$, $\dot{\boldsymbol{x}}$, $\boldsymbol{\theta}$, $\boldsymbol{u}$ and $\boldsymbol{w}$. $\nabla \boldsymbol{P}$ is a system of first order polynomial ODEs with the dependent variables $\boldsymbol{x}$, $\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{x_0}}$, $\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{\theta}}$ and $\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{w}^{\boldsymbol{k}}_{\boldsymbol{j,0}}}$ and the independent variable t. The associated initial conditions are:

$$\boldsymbol{x}\Big|_{t=0} = \boldsymbol{x_0}, \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{x_0}}\Big|_{t=0} = \boldsymbol{I}_{n\times n}, \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{\theta}}\Big|_{t=0} = \mathbf{0}_{n\times l}, \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{w}^{\boldsymbol{k}}_{\boldsymbol{j,0}}}\Big|_{t=0} = \mathbf{0}_{n\times k} \tag{27}$$

Using the random numerical realizations and letting $\boldsymbol{\theta} = \tilde{\boldsymbol{\theta}}$, $\boldsymbol{u} = \boldsymbol{u}[t,k]$, $\boldsymbol{w} = \boldsymbol{w}[t,k]$ and $\frac{\partial w_j}{\partial \boldsymbol{w}^{\boldsymbol{k}}_{\boldsymbol{j,0}}} = \frac{\partial w_j}{\partial \boldsymbol{w}^{\boldsymbol{k}}_{\boldsymbol{j,0}}}[t,k]$, the occurring polynomial differential system $\nabla \boldsymbol{P}(\boldsymbol{x}, \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{x_0}}, \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{\theta}}, \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{w}^{\boldsymbol{k}}_{\boldsymbol{j,0}}}, \tilde{\boldsymbol{\theta}}, \boldsymbol{u}[t,k], \boldsymbol{w}[t,k])$ then allows for the use of Newton's iteration method to obtain the power series solutions of its dependent variables up to order k, i.e. $\boldsymbol{x}[t,k]$, $\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{x_0}}[t,k]$, $\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{\theta}}[t,k]$ and $\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{w}^{\boldsymbol{k}}_{\boldsymbol{j,0}}}[t,k]$.

3.2. Newton's iteration

The details about the use of Newton's iteration for finding power series solutions of polynomial ODEs up to an order of interest can be seen in Chapter 4 of [30], Section 5 of [31] and Section 3 of [3].

This subsection, first of all, places the attention on the first equation of $\nabla \boldsymbol{P}(\boldsymbol{x}, \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{x_0}}, \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{\theta}}, \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{w}^{\boldsymbol{k}}_{\boldsymbol{j,0}}}, \tilde{\boldsymbol{\theta}}, \tilde{\boldsymbol{u}}[t,k], \tilde{\boldsymbol{w}}[t,k])$, i.e.:

$$\boldsymbol{P}(\boldsymbol{x}, \dot{\boldsymbol{x}}, \tilde{\boldsymbol{\theta}}, \boldsymbol{u}[t,k], \boldsymbol{w}[t,k]) = \mathbf{0}_{n\times 1} \tag{28}$$

The objective of Newton's iteration is to solve equation (28) for the power series expansion of $\boldsymbol{x}$ up to order k, i.e. $\boldsymbol{x}[t,k]$, through a number of iterative computations. The iterations start from the initial condition $\boldsymbol{x_0}$ which is the first term of $\boldsymbol{x}[t,k]$. At each iteration, the power series expansion of $\boldsymbol{x}$ is computed up to a certain order and it is then used for the next iteration to compute the power series up to a higher order; the iterations are terminated when order k is reached. Let $\boldsymbol{x_q}$ denote the power series computed at the q^{th} iteration of the Newton's method. The formula of Newton's iteration for equation (28) is obtained by means of linearization of the equation around the point $\boldsymbol{x} = \boldsymbol{x_{q-1}}$, yielding:

$$\frac{\partial \boldsymbol{P}}{\partial \dot{\boldsymbol{x}}}\bigg|_{\boldsymbol{x}=\boldsymbol{x_{q-1}}}(\dot{\boldsymbol{x}}_{\boldsymbol{q}}-\dot{\boldsymbol{x}}_{\boldsymbol{q-1}})+\frac{\partial \boldsymbol{P}}{\partial \boldsymbol{x}}\bigg|_{\boldsymbol{x}=\boldsymbol{x_{q-1}},\dot{\boldsymbol{x}}=\dot{\boldsymbol{x}}_{\boldsymbol{q-1}}}(\boldsymbol{x_q}-\boldsymbol{x_{q-1}})+\boldsymbol{P}\bigg|_{\boldsymbol{x}=\boldsymbol{x_{q-1}},\dot{\boldsymbol{x}}=\dot{\boldsymbol{x}}_{\boldsymbol{q-1}}} = \boldsymbol{0}_{n\times 1} \tag{29}$$

Equation (29) is a first order linear ODE of the form:

$$\boldsymbol{C_1}(t)\dot{\boldsymbol{v}} + \boldsymbol{C_2}(t)\boldsymbol{v} + \boldsymbol{C_3}(t) = \boldsymbol{0} \tag{30}$$

Equation (30) can be solved analytically using the method of integrating factor, giving the solution:

$$\boldsymbol{v}(t) = -e^{-\int \boldsymbol{C_1}^{-1}\boldsymbol{C_2}dt}\int e^{\int \boldsymbol{C_1}^{-1}\boldsymbol{C_2}dt}(\boldsymbol{C_1}^{-1}\boldsymbol{C_3})dt + \boldsymbol{C_4}e^{-\int \boldsymbol{C_1}^{-1}\boldsymbol{C_2}dt} \tag{31}$$

where $\boldsymbol{C_1}$ is an invertible matrix, and $\boldsymbol{C_4}$ is a matrix of constants which is determined by knowing the initial condition $\boldsymbol{v}(0)$. Following equation (31), the correction term $\boldsymbol{E_q}(t)$ defined as:

$$\boldsymbol{E_q}(t) = \boldsymbol{x_q} - \boldsymbol{x_{q-1}}, \;\; \boldsymbol{E_q}(0) = \boldsymbol{0} \tag{32}$$

is determined as an analytical solution to equation (29), and $\boldsymbol{x_q}$ can then be expressed in terms of $\boldsymbol{x_{q-1}}$ as:

$$\begin{aligned}\boldsymbol{x_q} &= \boldsymbol{x_{q-1}} + \boldsymbol{E_q}(t) \\ &= \boldsymbol{x_{q-1}} - e^{-\int (\frac{\partial \boldsymbol{P}}{\partial \dot{\boldsymbol{x}}})^{-1}\frac{\partial \boldsymbol{P}}{\partial \boldsymbol{x}}dt}\int e^{\int (\frac{\partial \boldsymbol{P}}{\partial \dot{\boldsymbol{x}}})^{-1}\frac{\partial \boldsymbol{P}}{\partial \boldsymbol{x}}dt}(\frac{\partial \boldsymbol{P}}{\partial \dot{\boldsymbol{x}}}^{-1}\boldsymbol{P})dt\end{aligned} \tag{33}$$

Within equation (33), the power series of the matrix exponentials and inverses can be computed efficiently by using nested Newton's iterations (see details in [3] and [31]). For example, if $\boldsymbol{A_q}$ and $\boldsymbol{B_q}$ represent the power series expansions of $e^{-\int (\frac{\partial \boldsymbol{P}}{\partial \dot{\boldsymbol{x}}})^{-1}\frac{\partial \boldsymbol{P}}{\partial \boldsymbol{x}}dt}$ and its inverse $e^{\int (\frac{\partial \boldsymbol{P}}{\partial \dot{\boldsymbol{x}}})^{-1}\frac{\partial \boldsymbol{P}}{\partial \boldsymbol{x}}dt}$ respectively, they can be obtained through the following algorithm.

Algorithm II

Initialization: Set $g = 0$ and $\boldsymbol{A_0} = \boldsymbol{B_0} = \boldsymbol{I}_{n\times n}$.

1. Set $g = g + 1$;

2. Compute $\boldsymbol{A_g} = \boldsymbol{A_{g-1}} + \boldsymbol{A_{g-1}} \int \boldsymbol{B_{g-1}}(-(\frac{\partial \boldsymbol{P}}{\partial \dot{\boldsymbol{x}}})^{-1} \frac{\partial \boldsymbol{P}}{\partial \boldsymbol{x}})\boldsymbol{A_{g-1}} - \boldsymbol{B_{g-1}}\dot{\boldsymbol{A}}_{\boldsymbol{g-1}} dt$ truncated at order $2^g - 1$;

3. Compute $\boldsymbol{B_g} = 2\boldsymbol{B_{g-1}} - \boldsymbol{B_{g-1}}\boldsymbol{A_g}\boldsymbol{B_{g-1}}$ truncated at order $2^g - 1$;

4. End if $g = q$. Otherwise go to step **1**;

Using Algorithm II equation (33) becomes:

$$\boldsymbol{x_q} = \boldsymbol{x_{q-1}} - \boldsymbol{A_q} \int \boldsymbol{B_q}(\frac{\partial \boldsymbol{P}}{\partial \dot{\boldsymbol{x}}}^{-1} \boldsymbol{P})dt \tag{34}$$

Starting from $\boldsymbol{x_0} = \tilde{\boldsymbol{x}}_{\boldsymbol{0}}$ under the numerical realizations, equation (34) can readily be used to update $\boldsymbol{x_q}$ from $\boldsymbol{x_{q-1}}$ by calculating more terms of power series. It has been proven that the convergence of Newton's iteration is quadratic, which means the additional correct terms computed at the q^{th} iteration are twice as many as those at the $(q-1)^{th}$ iteration [31]. At the q^{th} iteration, the first 2^q terms of the obtained $\boldsymbol{x_q}$ are correct, and thus $\boldsymbol{x_q}$ should be truncated at order $2^q - 1$, i.e. $\boldsymbol{x_q} = \boldsymbol{x}[t, 2^q - 1]$, to remove the incorrect terms. The iterative process is continued until the q_0^{th} iteration when the maximum order of correct terms $2^{q_0} - 1 \geq k$. The computed power series $\boldsymbol{x_{q_0}} = \boldsymbol{x}[t, 2^{q_0} - 1]$ is then truncated at order k leading to the result of $\boldsymbol{x}[t, k]$.

Having obtained $\boldsymbol{x}[t, k]$ from the first equation of $\nabla \boldsymbol{P}$, $\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{x_0}}[t, k]$, $\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{\theta}}[t, k]$ and $\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{w}^{\boldsymbol{k}}_{\boldsymbol{j,0}}}[t, k]$ are computed based on the remaining three equations, i.e. equation (25) under the numerical realizations. Substituting $\boldsymbol{x}[t, k]$ into the equations gives the following first order linear ODEs:

$$\begin{aligned}
\left.\frac{\partial \boldsymbol{P}}{\partial \dot{\boldsymbol{x}}}\right|_{\boldsymbol{x}=\boldsymbol{x}[t,k]} \frac{\partial \dot{\boldsymbol{x}}}{\partial \boldsymbol{x_0}} + \left.\frac{\partial \boldsymbol{P}}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x}=\boldsymbol{x}[t,k],\dot{\boldsymbol{x}}=\dot{\boldsymbol{x}}[t,k]} \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{x_0}} &= \boldsymbol{0}_{n\times n} \\
\left.\frac{\partial \boldsymbol{P}}{\partial \dot{\boldsymbol{x}}}\right|_{\boldsymbol{x}=\boldsymbol{x}[t,k]} \frac{\partial \dot{\boldsymbol{x}}}{\partial \boldsymbol{\theta}} + \left.\frac{\partial \boldsymbol{P}}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x}=\boldsymbol{x}[t,k],\dot{\boldsymbol{x}}=\dot{\boldsymbol{x}}[t,k]} \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{\theta}} + \left.\frac{\partial \boldsymbol{P}}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{x}=\boldsymbol{x}[t,k],\dot{\boldsymbol{x}}=\dot{\boldsymbol{x}}[t,k]} &= \boldsymbol{0}_{n\times l} \\
\left.\frac{\partial \boldsymbol{P}}{\partial \dot{\boldsymbol{x}}}\right|_{\boldsymbol{x}=\boldsymbol{x}[t,k]} \frac{\partial \dot{\boldsymbol{x}}}{\partial \boldsymbol{w}^{\boldsymbol{k}}_{\boldsymbol{j,0}}} + \left.\frac{\partial \boldsymbol{P}}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x}=\boldsymbol{x}[t,k],\dot{\boldsymbol{x}}=\dot{\boldsymbol{x}}[t,k]} \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{w}^{\boldsymbol{k}}_{\boldsymbol{j,0}}} + \left.\frac{\partial \boldsymbol{P}}{\partial w_j}\right|_{\boldsymbol{x}=\boldsymbol{x}[t,k],\dot{\boldsymbol{x}}=\dot{\boldsymbol{x}}[t,k]} \frac{\partial w_j}{\partial \boldsymbol{w}^{\boldsymbol{k}}_{\boldsymbol{j,0}}}[t,k] &= \boldsymbol{0}_{n\times k}
\end{aligned} \tag{35}$$

with the given initial conditions of $\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{x_0}}$, $\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{\theta}}$ and $\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{w_{j,0}^k}}$ in equation (27). Using equation (31) the solutions of equation (35) are given by:

$$\begin{aligned}
\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{x_0}} &= e^{-\int (\frac{\partial \boldsymbol{P}}{\partial \dot{\boldsymbol{x}}})^{-1} \frac{\partial \boldsymbol{P}}{\partial \boldsymbol{x}} dt} \\
\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{\theta}} &= -e^{-\int (\frac{\partial \boldsymbol{P}}{\partial \dot{\boldsymbol{x}}})^{-1} \frac{\partial \boldsymbol{P}}{\partial \boldsymbol{x}} dt} \int e^{\int (\frac{\partial \boldsymbol{P}}{\partial \dot{\boldsymbol{x}}})^{-1} \frac{\partial \boldsymbol{P}}{\partial \boldsymbol{x}} dt} (\frac{\partial \boldsymbol{P}}{\partial \dot{\boldsymbol{x}}}^{-1} \frac{\partial \boldsymbol{P}}{\partial \boldsymbol{\theta}}) dt \\
\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{w_{j,0}^k}} &= -e^{-\int (\frac{\partial \boldsymbol{P}}{\partial \dot{\boldsymbol{x}}})^{-1} \frac{\partial \boldsymbol{P}}{\partial \boldsymbol{x}} dt} \int e^{\int (\frac{\partial \boldsymbol{P}}{\partial \dot{\boldsymbol{x}}})^{-1} \frac{\partial \boldsymbol{P}}{\partial \boldsymbol{x}} dt} (\frac{\partial \boldsymbol{P}}{\partial \dot{\boldsymbol{x}}}^{-1} \frac{\partial \boldsymbol{P}}{\partial w_j} \frac{\partial w_j}{\partial \boldsymbol{w_{j,0}^k}}[t,k]) dt
\end{aligned} \tag{36}$$

With the power series manipulations of integration, differentiation, matrix exponential and inverse, the power series forms of $\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{x_0}}$, $\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{\theta}}$ and $\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{w_{j,0}^k}}$ can be obtained from equation (36). These power series contain correct terms up to order k, leading to the results of $\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{x_0}}[t,k]$, $\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{\theta}}[t,k]$ and $\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{w_{j,0}^k}}[t,k]$ through truncation.

3.3. The algorithm and modular operations

It is now sufficient to summarize and present the procedure of computing $d\boldsymbol{\Omega^k}$ for a choice of k based on the power series methods in the following algorithm.

Algorithm III

Preprocessing: Compute $\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}}$, $\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{\theta}}$ and $\frac{\partial \boldsymbol{h}}{\partial w_j}$ $(j = 1, ..., r)$ symbolically. Convert the state-space equations to $\boldsymbol{P} = \boldsymbol{0}$ using equation (24), and compute $\frac{\partial \boldsymbol{P}}{\partial \dot{\boldsymbol{x}}}$, $\frac{\partial \boldsymbol{P}}{\partial \boldsymbol{x}}$, $\frac{\partial \boldsymbol{P}}{\partial \boldsymbol{\theta}}$ and $\frac{\partial \boldsymbol{P}}{\partial w_j}$ symbolically.

Initialization: Choose the sets of nonzero values randomly $\tilde{\boldsymbol{x}}_{\boldsymbol{0}}$, $\tilde{\boldsymbol{\theta}}$, $\tilde{\boldsymbol{w}}_{\boldsymbol{0}}, ..., \tilde{\boldsymbol{w}}_{\boldsymbol{0}}{}^{(k)}$ and $\tilde{\boldsymbol{u}}_{\boldsymbol{0}}, ..., \tilde{\boldsymbol{u}}_{\boldsymbol{0}}{}^{(k)}$. Let $\boldsymbol{u}[t,k] = \tilde{\boldsymbol{u}}_{\boldsymbol{0}} + ... + \tilde{\boldsymbol{u}}_{\boldsymbol{0}}{}^{(k)} \frac{t^k}{k!}$, $\boldsymbol{w}[t,k] = \tilde{\boldsymbol{w}}_{\boldsymbol{0}} + ... + \tilde{\boldsymbol{w}}_{\boldsymbol{0}}{}^{(k)} \frac{t^k}{k!}$, and evaluate $\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}}$, $\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{\theta}}$, $\frac{\partial \boldsymbol{h}}{\partial w_j}$, $\boldsymbol{P}$, $\frac{\partial \boldsymbol{P}}{\partial \dot{\boldsymbol{x}}}$, $\frac{\partial \boldsymbol{P}}{\partial \boldsymbol{x}}$, $\frac{\partial \boldsymbol{P}}{\partial \boldsymbol{\theta}}$ and $\frac{\partial \boldsymbol{P}}{\partial w_j}$ at $\boldsymbol{\theta} = \tilde{\boldsymbol{\theta}}, \boldsymbol{u} = \boldsymbol{u}[t,k], \boldsymbol{w} = \boldsymbol{w}[t,k]$. Set $\boldsymbol{x_0} = \tilde{\boldsymbol{x}}_{\boldsymbol{0}}$ and $q = 0$.

1. Compute $\boldsymbol{x}[t,k]$ through Newton's iteration:

- **a.** Set $q = q + 1$;
- **b.** Evaluate $\boldsymbol{P}$, $\frac{\partial \boldsymbol{P}}{\partial \dot{\boldsymbol{x}}}$ and $\frac{\partial \boldsymbol{P}}{\partial \boldsymbol{x}}$ at $\boldsymbol{x} = \boldsymbol{x_{q-1}}$, $\dot{\boldsymbol{x}} = \dot{\boldsymbol{x}}_{\boldsymbol{q-1}}$, truncated at order $2^q - 1$;
- **c.** Compute $\boldsymbol{A_q}$ and $\boldsymbol{B_q}$ using Algorithm II;
- **d.** Compute $\boldsymbol{x_q}$ using equation (34) truncated at order $2^q - 1$;

e. If $2^q - 1 \geq k$, $\boldsymbol{x}[t,k] = \boldsymbol{x_q}$ truncated at order k and go to step **2**.

Otherwise go to step **a**;

2. Evaluate $\frac{\partial \boldsymbol{P}}{\partial \dot{\boldsymbol{x}}}$, $\frac{\partial \boldsymbol{P}}{\partial \boldsymbol{x}}$, $\frac{\partial \boldsymbol{P}}{\partial \boldsymbol{\theta}}$ and $\frac{\partial \boldsymbol{P}}{\partial w_j}$ at $\boldsymbol{x} = \boldsymbol{x}[t,k], \dot{\boldsymbol{x}} = \dot{\boldsymbol{x}}[t,k]$, truncated at order k;

3. Compute $\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{x_0}}[t,k]$, $\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{\theta}}[t,k]$ and $\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{w}_{\boldsymbol{j},\boldsymbol{0}}^{\boldsymbol{k}}}[t,k]$ using equation (36) and Algorithm II, truncated at order k;

4. Compute $\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}}[t,k]$, $\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{\theta}}[t,k]$ and $\frac{\partial \boldsymbol{h}}{\partial w_j}[t,k]$ by evaluating $\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}}$, $\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{\theta}}$ and $\frac{\partial \boldsymbol{h}}{\partial w_j}$ at $\boldsymbol{x} = \boldsymbol{x}[t,k]$, truncated at order k. Take Taylor series expansions if necessary;

5. Compute $\frac{d\boldsymbol{h}}{d\boldsymbol{x_0}}[t,k]$, $\frac{d\boldsymbol{h}}{d\boldsymbol{\theta}}[t,k]$ and $\frac{d\boldsymbol{h}}{d\boldsymbol{w}_{\boldsymbol{j},\boldsymbol{0}}^{\boldsymbol{k}}}[t,k]$ using equation (21);

6. Extract the coefficients of $\frac{d\boldsymbol{h}}{d\boldsymbol{x_0}}[t,k]$, $\frac{d\boldsymbol{h}}{d\boldsymbol{\theta}}[t,k]$ and $\frac{d\boldsymbol{h}}{d\boldsymbol{w}_{\boldsymbol{j},\boldsymbol{0}}^{\boldsymbol{k}}}[t,k]$ to construct $d\boldsymbol{\Omega}_{\boldsymbol{x}}^{\boldsymbol{k}}$, $d\boldsymbol{\Omega}_{\boldsymbol{\theta}}^{\boldsymbol{k}}$ and $d\boldsymbol{\Omega}_{\boldsymbol{w_j}}^{\boldsymbol{k}}$ respectively based on equations (16) and (19);

7. Arrange the observability matrix $d\boldsymbol{\Omega}^{\boldsymbol{k}} = [d\boldsymbol{\Omega}_{\boldsymbol{x}}^{\boldsymbol{k}}, d\boldsymbol{\Omega}_{\boldsymbol{\theta}}^{\boldsymbol{k}}, d\boldsymbol{\Omega}_{\boldsymbol{w_1}}^{\boldsymbol{k}}, ..., d\boldsymbol{\Omega}_{\boldsymbol{w_r}}^{\boldsymbol{k}}]$.

Similar to what was discussed in [3] for the case of all inputs being measured, in Algorithm III all the computations between power series can be done with modular operations so as to further improve the efficiency. The means of achieving so is through selecting a large positive prime number p at the initialization step, and choosing random integers within the range 1 to $p-1$ for $\tilde{\boldsymbol{x}}_{\boldsymbol{0}}$, $\tilde{\boldsymbol{\theta}}$, $\tilde{\boldsymbol{w}}_{\boldsymbol{0}}, ..., \tilde{\boldsymbol{w}}_{\boldsymbol{0}}{}^{(k)}$ and $\tilde{\boldsymbol{u}}_{\boldsymbol{0}}, ..., \tilde{\boldsymbol{u}}_{\boldsymbol{0}}{}^{(k)}$. For all the occurring power series throughout the algorithm, modular operations are applied to their coefficients, i.e. the values of coefficients *mod* p. The arithmetic of modular operations can be simply described as follows [32]: an integer a *mod* p calculates the least positive remainder of a divided by p, resulting in an integer between 0 and $p-1$; a fraction of two integers b/c *mod* p, based on the rule of modular multiplicative inverse, also gives an integer between 0 and $p-1$. As a consequence of the modular operations, all the occurring coefficients of power series throughout the algorithm are restricted within the range 0 to $p-1$. The problem of number growth is thus effectively controlled to prevent expensive computations between very large numbers.

If the modular operations are applied, the output from Algorithm III is equivalent to the following matrix, due to the distributive property of modular arithmetic in polynomial calculations [28]:

$$d\boldsymbol{\Omega}_{\boldsymbol{p}}^{\boldsymbol{k}} = d\boldsymbol{\Omega}^{\boldsymbol{k}} \;\; mod \;\; p \tag{37}$$

where the *mod* operations are used element-wise. The observability is then determined by evaluating the rank of $d\boldsymbol{\Omega}_{\boldsymbol{p}}^{\boldsymbol{k}}$ over a finite field $\mathbb{F}_p$ through applying the modular operations to Gaussian elimination [3, 32, 33, 34]. Similarly as discussed in [3, 28], $d\boldsymbol{\Omega}_{\boldsymbol{p}}^{\boldsymbol{k}}$ and $d\boldsymbol{\Omega}^{\boldsymbol{k}}$ contain almost identical rank information over the corresponding fields and can hence deliver agreed observability results with a very high probability of success. The probability of disagreement, when the observability results obtained with and without the modular operations do not coincide, is vanishingly small if p is selected to be large as proven in [3]. This probability can even become practically negligible if the algorithm is repeated using different choices of p and numerical realizations.

Algorithm III demonstrates how to obtain the k-row observability matrix after pre-selecting a value of k of interest. However, if the user realizes that the higher order Lie derivatives, i.e. a larger k, are required to compute for understanding the observability of a tested system, the drawback of Algorithm III is that some computations would need to be repeated. In the following, a recursive version of Algorithm III, termed as Algorithm IV, is presented where k is incrementally changed to match the order of the leading term of the power series produced from Newton's iteration, i.e. $k = 2^q - 1$. At each increment of k, the k-row observability properties are detected from the corresponding matrix.

Algorithm IV

Preprocessing: Compute $\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}}$, $\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{\theta}}$ and $\frac{\partial \boldsymbol{h}}{\partial w_j}$ $(j = 1, ..., r)$ symbolically. Convert the state-space equations to $\boldsymbol{P} = \boldsymbol{0}$ using equation (24), and compute $\frac{\partial \boldsymbol{P}}{\partial \dot{\boldsymbol{x}}}$, $\frac{\partial \boldsymbol{P}}{\partial \boldsymbol{x}}$, $\frac{\partial \boldsymbol{P}}{\partial \boldsymbol{\theta}}$ and $\frac{\partial \boldsymbol{P}}{\partial w_j}$ symbolically.

Initialization: Select a large positive prime number p, and choose the sets of integers randomly $\tilde{\boldsymbol{x}}_{\boldsymbol{0}}$, $\tilde{\boldsymbol{\theta}}$, $\tilde{\boldsymbol{w}}_{\boldsymbol{0}}, ..., \tilde{\boldsymbol{w}}_{\boldsymbol{0}}^{(k)}$ and $\tilde{\boldsymbol{u}}_{\boldsymbol{0}}, ..., \tilde{\boldsymbol{u}}_{\boldsymbol{0}}^{(k)}$ within the range 1 to $p-1$. Let $\boldsymbol{u}[t,k] = \tilde{\boldsymbol{u}}_{\boldsymbol{0}} + ... + \tilde{\boldsymbol{u}}_{\boldsymbol{0}}^{(k)} \frac{t^k}{k!}$, $\boldsymbol{w}[t,k] = \tilde{\boldsymbol{w}}_{\boldsymbol{0}} + ... + \tilde{\boldsymbol{w}}_{\boldsymbol{0}}^{(k)} \frac{t^k}{k!}$, and evaluate $\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}}$, $\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{\theta}}$, $\frac{\partial \boldsymbol{h}}{\partial w_j}$, $\boldsymbol{P}$, $\frac{\partial \boldsymbol{P}}{\partial \dot{\boldsymbol{x}}}$, $\frac{\partial \boldsymbol{P}}{\partial \boldsymbol{x}}$, $\frac{\partial \boldsymbol{P}}{\partial \boldsymbol{\theta}}$ and $\frac{\partial \boldsymbol{P}}{\partial w_j}$ at $\boldsymbol{\theta} = \tilde{\boldsymbol{\theta}}, \boldsymbol{u} = \boldsymbol{u}[t,k], \boldsymbol{w} = \boldsymbol{w}[t,k]$. Set

$\boldsymbol{x_0} = \tilde{\boldsymbol{x}}_{\boldsymbol{0}}$ and $q = 0$.

1. Set $q = q + 1$;

2. Evaluate $\boldsymbol{P}$, $\frac{\partial \boldsymbol{P}}{\partial \dot{\boldsymbol{x}}}$ and $\frac{\partial \boldsymbol{P}}{\partial \boldsymbol{x}}$ at $\boldsymbol{x} = \boldsymbol{x_{q-1}}$, $\dot{\boldsymbol{x}} = \dot{\boldsymbol{x}}_{\boldsymbol{q-1}}$, truncated at order $2^q - 1$;

3. Compute $\boldsymbol{A_q}$ and $\boldsymbol{B_q}$ using Algorithm II;

4. Compute $\boldsymbol{x}[t,k] = \boldsymbol{x_q}$ using equation (34) truncated at order $k = 2^q - 1$;

5. Evaluate $\frac{\partial \boldsymbol{P}}{\partial \dot{\boldsymbol{x}}}$, $\frac{\partial \boldsymbol{P}}{\partial \boldsymbol{x}}$, $\frac{\partial \boldsymbol{P}}{\partial \boldsymbol{\theta}}$ and $\frac{\partial \boldsymbol{P}}{\partial w_j}$ at $\boldsymbol{x} = \boldsymbol{x}[t,k], \dot{\boldsymbol{x}} = \dot{\boldsymbol{x}}[t,k]$, truncated at order k;

6. Compute $\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{x_0}}[t,k]$, $\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{\theta}}[t,k]$ and $\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{w^k_{j,0}}}[t,k]$ using equation (36) and Algorithm II, truncated at order k;

7. Compute $\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}}[t,k]$, $\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{\theta}}[t,k]$ and $\frac{\partial \boldsymbol{h}}{\partial w_j}[t,k]$ by evaluating $\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}}$, $\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{\theta}}$ and $\frac{\partial \boldsymbol{h}}{\partial w_j}$ at $\boldsymbol{x} = \boldsymbol{x}[t,k]$, truncated at order k. Take Taylor series expansions if necessary;

8. Compute $\frac{d\boldsymbol{h}}{d\boldsymbol{x_0}}[t,k]$, $\frac{d\boldsymbol{h}}{d\boldsymbol{\theta}}[t,k]$ and $\frac{d\boldsymbol{h}}{d\boldsymbol{w^k_{j,0}}}[t,k]$ using equation (21);

9. Extract the coefficients of $\frac{d\boldsymbol{h}}{d\boldsymbol{x_0}}[t,k]$, $\frac{d\boldsymbol{h}}{d\boldsymbol{\theta}}[t,k]$ and $\frac{d\boldsymbol{h}}{d\boldsymbol{w^k_{j,0}}}[t,k]$ to construct $d\boldsymbol{\Omega^k_x}$, $d\boldsymbol{\Omega^k_\theta}$ and $d\boldsymbol{\Omega^k_{w_j}}$ respectively based on equations (16) and (19);

10. Arrange the observability matrix $d\boldsymbol{\Omega^k} = [d\boldsymbol{\Omega^k_x}, d\boldsymbol{\Omega^k_\theta}, d\boldsymbol{\Omega^k_{w_1}}, ..., d\boldsymbol{\Omega^k_{w_r}}]$;

11. End if $rank(d\boldsymbol{\Omega^k}) = n + l + (k+1)r$, and $\boldsymbol{x}$, $\boldsymbol{\theta}$ and $\boldsymbol{w}$ are detected observable;

12. If $rank(d\boldsymbol{\Omega^k}) < n + l + (k+1)r$, detect the k-row observability properties of the components of $\boldsymbol{x^k}$;

13. End if $\boldsymbol{x}$, $\boldsymbol{\theta}$ and $\boldsymbol{w}$ are detected observable. Otherwise go to step **1**;

Algorithms III and IV can be implemented efficiently in a variety of software with symbolic computing tools, such as MATLAB with symbolic toolbox [35], Maple [36] and Mathematica [37]. Algorithm IV succeeds in identifying an observable system or the observable quantities (dynamic states, parameters or inputs) of an unobservable system. However it has not been feasible to theoretically conclude the observability of those k-row unobservable quantities. In the examples of this paper, a practical guideline is provided on how to certify unobservable quantities with sufficient confidence by making use of the submatrices of the observability matrix of a large k. Theoretically addressing this convergence problem will be focus of future extensions of this work.

4. Illustrative examples

4.1. A 2 degrees of freedom (DOFs) nonlinear system

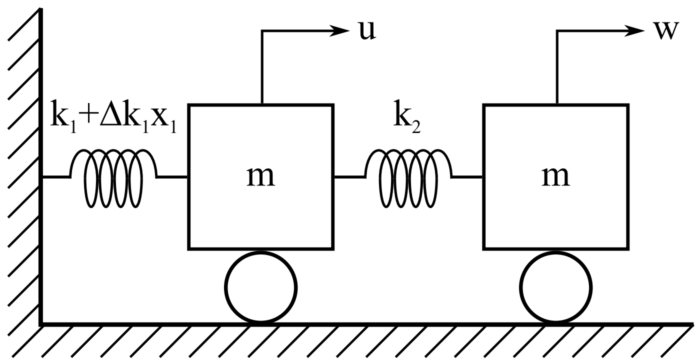

Figure 1: A 2 DOFs nonlinear system driven by measured and unmeasured inputs.

This example demonstrates the use of the proposed algorithm, with the versions of Algorithms III and IV, to examine the observability of a 2 DOFs nonlinear system driven by both measured and unmeasured inputs as shown in Figure 1. Let x_1, x_2, v_1, v_2 and $\dot{v}_1$, $\dot{v}_2$ denote the displacements, velocities and accelerations of the two masses m of the system respectively. A measured input u is applied at the first mass and an unmeasured input w is applied at the second mass. The stiffness of the nonlinear spring that connects the first mass to the fixed support is displacement-proportional and given by $k_1 + \Delta k_1 x_1$. The linear spring connecting the two masses has the stiffness of k_2. The dynamic states of the system are x_1, x_2, v_1 and v_2, and the unknown parameters to be identified are k_1, Δk_1, k_2 and m. Therefore:

$$\boldsymbol{x} = [x_1, x_2, v_1, v_2]^T, \;\; \boldsymbol{\theta} = [k_1, \Delta k_1, k_2, m]^T \tag{38}$$

The state-space equations of the system are given by:

$$\dot{\boldsymbol{x}} = \frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \\ v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ \frac{-(k_1+\Delta k_1 x_1)x_1 + k_2(x_2-x_1)+u}{m} \\ \frac{k_2(x_1-x_2)+w}{m} \end{bmatrix}, \;\; \dot{\boldsymbol{\theta}} = \frac{d}{dt}\begin{bmatrix} k_1 \\ \Delta k_1 \\ k_2 \\ m \end{bmatrix} = \mathbf{0}_{4\times 1} \tag{39}$$

Two different sensor setups are considered with the measurements taken in the setups denoted as $\boldsymbol{y_1}$ and $\boldsymbol{y_2}$ respectively:

$$\begin{aligned}
\boldsymbol{y_1} &= \begin{bmatrix} x_1 \\ \dot{v}_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ \frac{k_2(x_1-x_2)+w}{m} \end{bmatrix} \\
\boldsymbol{y_2} &= \begin{bmatrix} \dot{v}_1 \\ \dot{v}_2 \end{bmatrix} = \begin{bmatrix} \frac{-(k_1+\Delta k_1 x_1)x_1+k_2(x_2-x_1)+u}{m} \\ \frac{k_2(x_1-x_2)+w}{m} \end{bmatrix}
\end{aligned} \tag{40}$$

Using the measurement of the first sensor setup $\boldsymbol{y_1}$, Algorithm IV terminates at $k = 7$ reporting that the dynamical system is fully observable, that is, $\boldsymbol{x}$, $\boldsymbol{\theta}$ and w are all observable. The observability result is in agreement with those reported from the symbolic Algorithm I and the ORC-DF algorithm presented in [22]. With respect to the measurement of the second sensor setup $\boldsymbol{y_2}$, however, the system is not detected observable. To analyse the k-row observability properties of $\boldsymbol{x}$, $\boldsymbol{\theta}$ and w and their convergence, Algorithm III is used to calculate the observability matrix $d\boldsymbol{\Omega^k}$ with a choice of $k = 15$. The resulting $d\boldsymbol{\Omega^{15}}$ then allows for obtaining $d\boldsymbol{\Omega^0}$, ..., $d\boldsymbol{\Omega^{14}}$ by making use of its submatrices. Figure 2 plots the ranks of $d\boldsymbol{\Omega^k}$ in (a) and the k-row observability properties detected from $d\boldsymbol{\Omega^k}$ in (b) for $k = 0, ..., 15$. As can be seen for the measurement $\boldsymbol{y_1}$, $rank(d\boldsymbol{\Omega^k}) = n + l + (k+1)r$ occurs and all the $\boldsymbol{x}$, $\boldsymbol{\theta}$ and w correspondingly become observable at $k = 7$. In the case of measuring $\boldsymbol{y_2}$, v_1, v_2, Δk_1, k_2 and m become observable at $k = 6$, but x_1, x_2, k_1 and w remain k-row unobservable throughout all the k considered. From $k = 6$ to 15, $rank(d\boldsymbol{\Omega^k})$ (red-dashed line) and the rank condition $n + l + (k+1)r$ (black-dashed line) increase at the same rate leading to two parallel lines. The occurrence of the parallel lines is in indication that it is highly likely that x_1, x_2, k_1 and w would remain k-row unobservable for larger $k \to \infty$, which implies that these quantities are in fact unobservable. An advantage of using the proposed algorithm is that $d\boldsymbol{\Omega^k}$ of a larger k (e.g. $k = 100$) can be calculated to further investigate the ranks and the k-row observability properties (for e.g. $k = 16$ to 100), such that the observability of those k-row unobservable quantities can be concluded with more sufficient confidence.

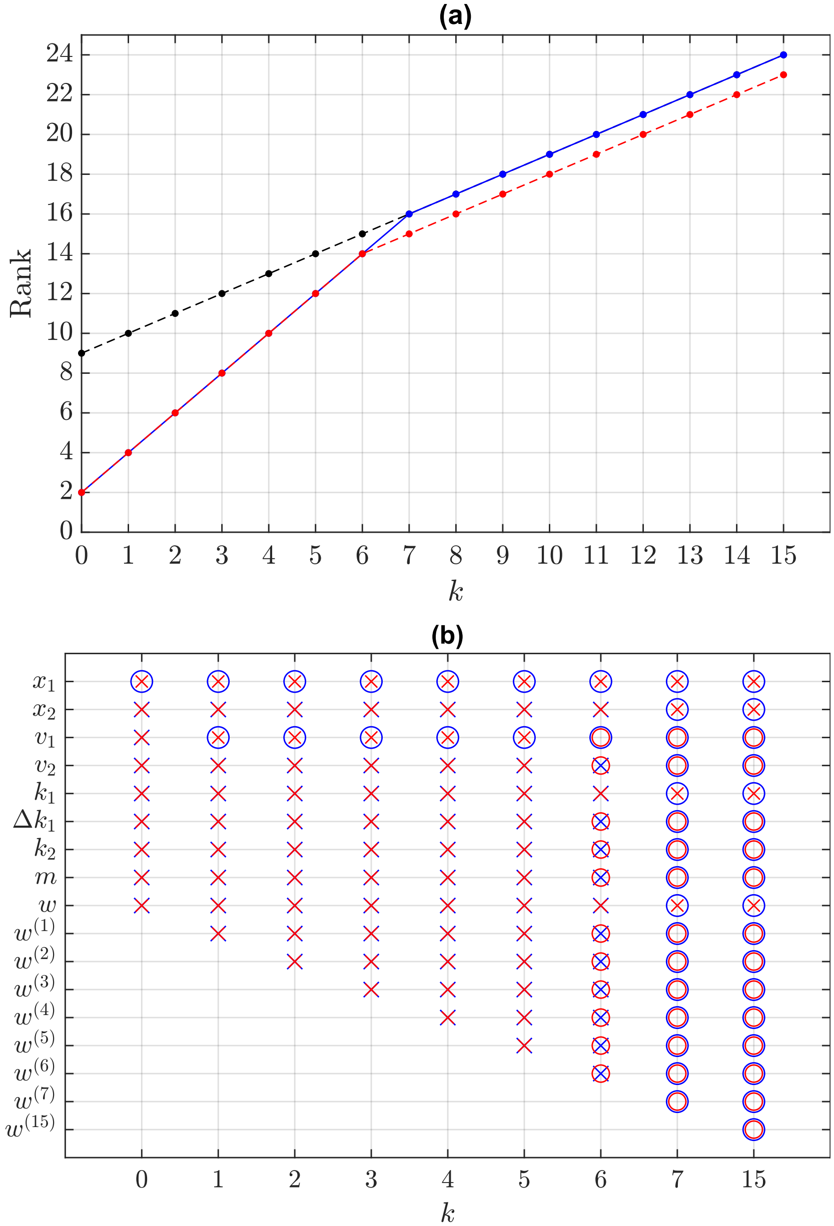

Figure 2: (a) The ranks of $d\boldsymbol{\Omega^k}$ for the case of $\boldsymbol{y_1}$ measurement (blue line) and the case of $\boldsymbol{y_2}$ measurement (red-dashed line), with comparison to the rank condition $n+l+(k+1)r$ (black-dashed line), (b) the observable quantities (circles) and k-row unobservable quantities (x-marks) for the case of $\boldsymbol{y_1}$ measurement (blue) and the case of $\boldsymbol{y_2}$ measurement (red).

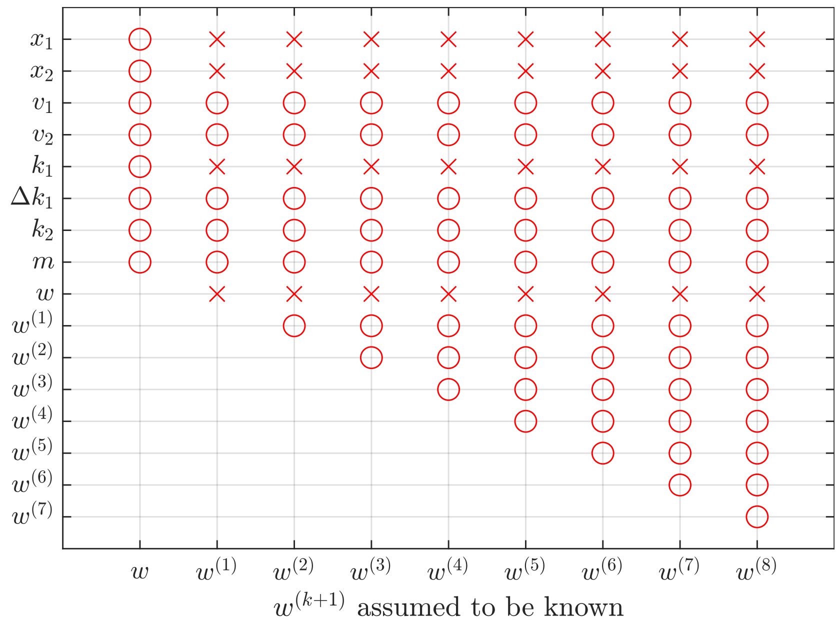

Figure 3: The observability of the augmented form of the system with $w^{(k+1)}$ assumed to be known for the case of $\boldsymbol{y_2}$ measurement.

The afore-described graphical method for determining unobservable quantities with a high level of confidence is suggested in [22]. This example illustrates an alternative way to conclude that x_1, x_2, k_1 and w are indeed unobservable with respect to the measurement $\boldsymbol{y_2}$. Assuming $w^{(k+1)}$ is known, the corresponding augmented form (3) of the considered system can be treated as a system with fully known or measured inputs. For such a system, Algorithm IV can be used as the regular ORC algorithm to examine its observability. When the algorithm operates up to a convergent point, the k-row unobservable quantities found can be immediately concluded to be unobservable (see [2, 3, 8] for the details of the ORC). Figure 3 summarizes the results of observability with assuming $w^{(k+1)}$ to be known for $k+1 = 0, ..., 8$. As shown in the figure, x_1, x_2, k_1 and w are detected unobservable at $k+1 = 1$ and remain unobservable further. If a quantity is unobservable for a known $w^{(1)}$, it is guaranteed to be unobservable for an unknown $w^{(1)}$ which is the case of this example. Therefore it can be concluded that x_1, x_2, k_1 and w are certainly unobservable, while the remaining quantities are observable as shown in Figure 2. It is not possible to properly estimate x_1, x_2, k_1 and w based on the measurements of u and $\boldsymbol{y_2}$ using any system identification method. In the MATLAB implementation available [2], the authors have included an 'Example' file that allows the reader to use the code to reproduce the observability results of this example under the setup $\boldsymbol{y_2}$.

4.2. A 101 DOFs nonlinear system with a pendulum tuned mass damper

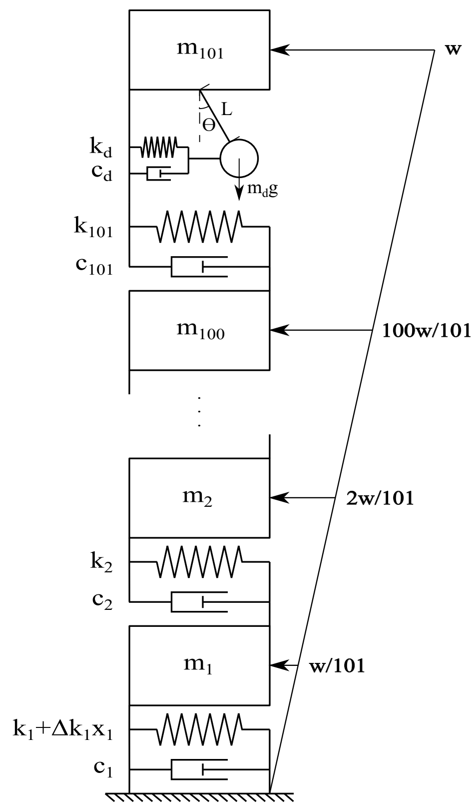

Figure 4: A 101-storey shear building with a pendulum TMD.

The proposed algorithm is now applied to test the observability of a large dynamical system. Increasing the number of masses of the system in Example 4.1 to 101, the occurring 101 DOFs nonlinear system can be viewed as the model of a 101-storey shear building. In this example, the effects of damping on the dynamics of the shear building are taken into consideration. As shown in Figure 4, linear viscous dampers with their coefficients c_i are connected successively between the floors of the model; a pendulum tuned mass damper (TMD), with its mass, stiffness and damping parameters being m_d, k_d and c_d respectively, is suspended from the top floor. The modelling of this example is inspired by the design of the Taipei 101 skyscraper in Taiwan. It is further assumed that unmeasured wind load with linearly increased magnitude is applied along the building. The maximum magnitude of the wind, w, appears at the level of top floor, and the load applied at the i^{th} floor is $\frac{iw}{101}$.

Let x_i and v_i denote the displacement and velocity of the mass m_i respectively. θ and v_θ denote respectively the sway angle and angular velocity of the pendulum TMD. The dynamic states of the considered system are therefore:

$$\boldsymbol{x} = [x_1, ..., x_{101}, \theta, v_1, ..., v_{101}, v_\theta]^T \tag{41}$$

k_i is the stiffness of the linear spring between m_{i-1} and m_i, and the stiffness of the nonlinear spring at the bottom is again given by $k_1 + \Delta k_1 x_1$. The unknown parameters are assumed to be:

$$\boldsymbol{\theta} = [k_1, ..., k_{101}, \Delta k_1, c_1, ..., c_{101}, m_d, k_d, c_d]^T \tag{42}$$

The equations used to describe the dynamics of the pendulum TMD are given by:

$$\begin{aligned}
\dot{\theta} =& v_\theta \\
\dot{v}_\theta =& \frac{k_{101}(x_{101} - x_{100})}{m_{101}L} + \frac{c_{101}(v_{101} - v_{100})}{m_{101}L} - \frac{m_{101} + m_d}{m_{101}m_d}(k_d + \frac{m_d g}{L})\theta \\
& - \frac{m_{101} + m_d}{m_{101}m_d} c_d v_\theta - \frac{w}{m_{101}L}
\end{aligned} \tag{43}$$

The following equations describe the dynamics of m_{101}:

$$\begin{aligned}
\dot{x}_{101} =& v_{101} \\
\dot{v}_{101} =& - \frac{k_{101}(x_{101} - x_{100})}{m_{101}} - \frac{c_{101}(v_{101} - v_{100})}{m_{101}} + \frac{(k_d L + m_d g)\theta}{m_{101}} \\
& + \frac{c_d L v_\theta}{m_{101}} + \frac{w}{m_{101}}
\end{aligned} \tag{44}$$

while the state-space equations of the remaining parts of the system can be written referring to the standard modelling of shear buildings [38]. Ten sensors are installed at every ten floors along the building to measure five displacements and five accelerations. Specifically, the measured outputs are:

$$\boldsymbol{y} = [\dot{v}_{101}, x_{91}, \dot{v}_{81}, x_{71}, \dot{v}_{61}, x_{51}, \dot{v}_{41}, x_{31}, \dot{v}_{21}, x_{11}]^T \tag{45}$$

Algorithm IV reports that the shear building with the pendulum TMD is fully observable, that is, $\boldsymbol{x}$, $\boldsymbol{\theta}$ and w are all observable for the measurement $\boldsymbol{y}$. It should be noted that the symbolic Algorithm I, the ORC-DF algorithm in [22] and the FISPO algorithm in [24] are far from being capable of handling this example on a standard computer, due to their limitation of high memory requirements. In fact, most systems studied using the ORC-DF in [22] have only 2-3 DOFs which is at the same level of dimensionality as Example 4.1. The limit of symbolic Algorithm I is expected to be similar to that of the ORC-DF. As shown in [25], the FISPO algorithm can reach the ability of assessing a biological model with 25 dynamic states, 26 unknown parameters, 5 unknown inputs and 15 output measurements, which is however a consequence of applying the model decomposition strategy and some other manual treatments, e.g. performing several rounds of observability testing where in each the Lie derivatives of a first few orders are calculated and parts of identifiable parameters detected are then assumed to be known. Moreover, additional assumptions on the properties of inputs may facilitate the study of certain systems of larger size, by both the ORC-DF and FISPO; such assumptions, e.g. assuming that the inputs are polynomial with respect to time and of a certain low order, however affect the observability results and are avoided in this work. In comparison, the proposed algorithm in this work is efficient in the usage of physical memory and is implemented relatively fast to test such a large system which contains a total of 410 dynamic states and unknown parameters. The efficiency of the algorithm could be further improved through implementing in a high-performance computer and making use of more computer cores.

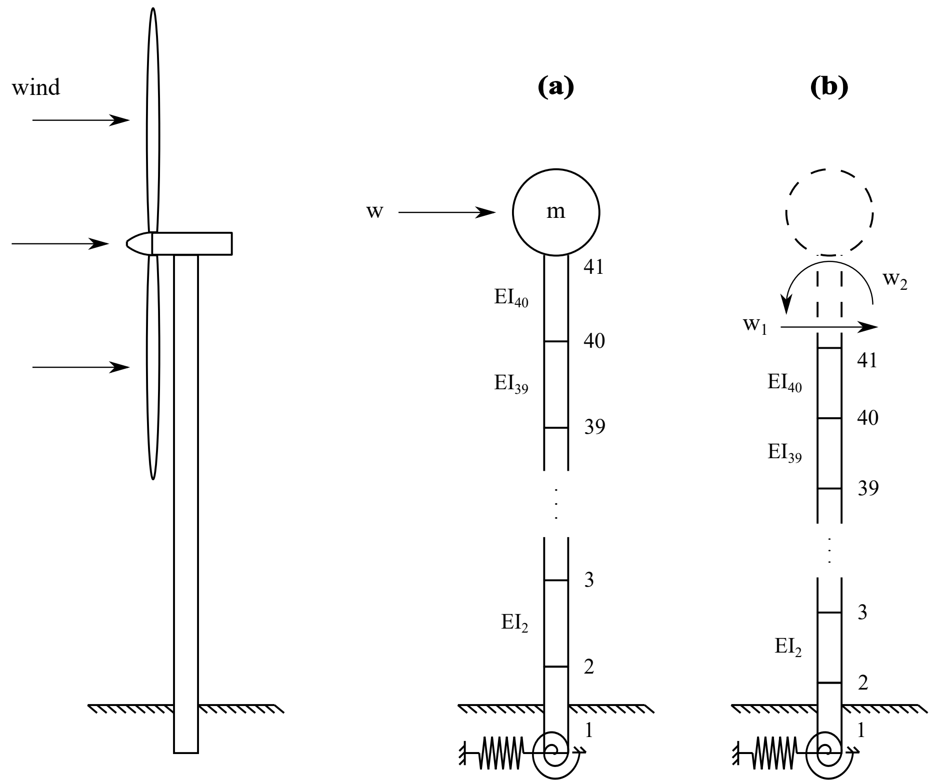

Figure 5: (a) A 2D FE model of wind turbine, and (b) a substructural model of the wind turbine.

4.3. A 2D finite element (FE) model of wind turbine

In this example, the observability of a wind turbine model as shown in Figure 5 is examined using the proposed algorithm. Figure 5(a) demonstrates a simplified 2D FE model of the wind turbine, where the rotor blades and nacelle are lumped at the top of the tower as a mass m with the associated moment of inertia I_m. It is assumed that m is known but I_m is an unknown parameter to be identified. The mass is subject to a lateral thrust force w which is unmeasured. It is assumed that the torsional stiffness of the wind turbine is infinitely large and the torsional vibration is not taken into consideration in the modelling. The tower is discretized into 40 uniform Euler-Bernoulli beam elements. Let x_i and θ_i denote the displacement and rotation of the beam node i respectively, and

v_i and $v_{\theta i}$ the translational and rotational velocities respectively. The dynamic state vector of the model is a collection of all the nodal displacements, rotations and the corresponding velocities:

$$\boldsymbol{x} = [x_1, \theta_1, ..., x_{41}, \theta_{41}, v_1, v_{\theta 1}, ..., v_{41}, v_{\theta 41}]^T \tag{46}$$

The stiffness and mass matrices of the i^{th} beam element between nodes i and $i+1$ are given by:

$$\begin{aligned} \boldsymbol{K_e} &= \frac{EI_i}{L^3} \begin{bmatrix} 12 & 6L & -12 & 6L \\ 6L & 4L^2 & -6L & 2L^2 \\ -12 & -6L & 12 & -6L \\ 6L & 2L^2 & -6L & 4L^2 \end{bmatrix} \\ \boldsymbol{M_e} &= \frac{\rho A_i L}{420} \begin{bmatrix} 156 & 22L & 54 & -13L \\ 22L & 4L^2 & 13L & -3L^2 \\ 54 & 13L & 156 & -22L \\ -13L & -3L^2 & -22L & 4L^2 \end{bmatrix} \end{aligned} \tag{47}$$

where EI_i is the unknown stiffness parameter of the element, and ρ, A_i and L are the known density, cross-section area and length respectively. The foundation of the wind turbine is modelled as the following macro-element:

$$\begin{bmatrix} F_{spr} \\ M_{spr} \end{bmatrix} = \begin{bmatrix} k_{11} & k_{12} \\ k_{12} & k_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ \theta_1 \end{bmatrix} \tag{48}$$

where F_{spr} and M_{spr} are the resistant force and moment respectively provided by the springs. The stiffness parameters k_{11}, k_{12} and k_{22} are to be identified. For the wind turbine model, the unknown parameter vector is:

$$\boldsymbol{\theta} = [I_m, EI_1, ..., EI_{40}, k_{11}, k_{12}, k_{22}]^T \tag{49}$$

and the state-space equations are written in the following form:

$$\dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}w, \;\; \dot{\boldsymbol{\theta}} = \boldsymbol{0} \tag{50}$$

where

$$\boldsymbol{A} = \begin{bmatrix} \mathbf{0}_{82\times 82} & \boldsymbol{I}_{82\times 82} \\ -\boldsymbol{M}^{-1}\boldsymbol{K} & \mathbf{0}_{82\times 82} \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} \mathbf{0}_{82\times 1} \\ \boldsymbol{M}^{-1} \begin{bmatrix} \mathbf{0}_{80\times 1} \\ 1 \\ 0 \end{bmatrix} \end{bmatrix} \tag{51}$$

where $\boldsymbol{K}$ and $\boldsymbol{M}$ are respectively the model stiffness and mass matrices constructed from the assembly [38] of the element stiffness and mass matrices of the lumped mass, the beam elements of the tower and the springs of the foundation.

Figure 5(b) shows a substructural model of the wind turbine which occurs by introducing a cut below the nacelle. A shear force w_1 and a bending moment w_2, which are unmeasured, are applied as a result of the cut to account for the dynamic effects from the upper-structure. This model is used to investigate whether it is theoretically possible to identify this substructure of interest using only measurements obtained from it. An advantage of the treatment in Figure 5(b) is that modelling of the rotor blades and nacelle is not required and the associated mass properties do not need to be known. In this case, the state-space equations of the substructural model are written as:

$$\begin{aligned} \dot{\boldsymbol{x}} &= \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B_1} w_1 + \boldsymbol{B_2} w_2 \\ \dot{\boldsymbol{\theta}} &= \frac{d}{dt}\left[EI_1, ..., EI_{40}, k_{11}, k_{12}, k_{22}\right]^T = \mathbf{0} \end{aligned} \tag{52}$$

where

$$\boldsymbol{B_1} = \begin{bmatrix} \mathbf{0}_{82\times 1} \\ \boldsymbol{M}^{-1} \begin{bmatrix} \mathbf{0}_{80\times 1} \\ 1 \\ 0 \end{bmatrix} \end{bmatrix}, \quad \boldsymbol{B_2} = \begin{bmatrix} \mathbf{0}_{82\times 1} \\ \boldsymbol{M}^{-1} \begin{bmatrix} \mathbf{0}_{80\times 1} \\ 0 \\ 1 \end{bmatrix} \end{bmatrix} \tag{53}$$

For both the models in Figure 5(a) and Figure 5(b), four sensors are installed at every ten nodes along the tower. Four different sensor setups are examined with the measurements taken denoted as $\boldsymbol{y_1}$, $\boldsymbol{y_2}$, $\boldsymbol{y_3}$ and $\boldsymbol{y_4}$ respectively:

$$\begin{aligned} \boldsymbol{y_1} &= [\dot{v}_{41}, \dot{v}_{31}, \dot{v}_{21}, \dot{v}_{11}], \quad \boldsymbol{y_2} = [\theta_{41}, \dot{v}_{31}, \dot{v}_{21}, \dot{v}_{11}] \\ \boldsymbol{y_3} &= [\theta_{41}, \theta_{31}, \dot{v}_{21}, \dot{v}_{11}], \quad \boldsymbol{y_4} = [\theta_{40}, \theta_{31}, \dot{v}_{21}, \dot{v}_{11}] \end{aligned} \tag{54}$$

Measurements		$\boldsymbol{y_1}$	$\boldsymbol{y_2}$	$\boldsymbol{y_3}$	$\boldsymbol{y_4}$
(a)	Observable	$\boldsymbol{\theta}$	$\boldsymbol{x}$, $\boldsymbol{\theta}$ and w	$\boldsymbol{x}$, $\boldsymbol{\theta}$ and w	$\boldsymbol{x}$, $\boldsymbol{\theta}$ and w
	Unobservable	$\boldsymbol{x}$ and w	-	-	-
(b)	Observable	$\boldsymbol{\theta}$	$\boldsymbol{\theta}$, θ_{41} and $v_{\theta 41}$	$\boldsymbol{x}$, $\boldsymbol{\theta}$, w_1 and w_2	the remaining
	Unobservable	$\boldsymbol{x}$, w_1 and w_2	the remaining	-	-
	k-row unobservable	-	-	-	EI_{40}, x_{41}, v_{41}, θ_{41} $v_{\theta 41}$, w_1 and w_2

Table 1: The observability of the full model in Figure 5(a) and the substructural model in Figure 5(b) with respect to the measurements $\boldsymbol{y_1}$, $\boldsymbol{y_2}$, $\boldsymbol{y_3}$ and $\boldsymbol{y_4}$.

In the measurement $\boldsymbol{y_1}$ all the sensors are measuring lateral accelerations, while in $\boldsymbol{y_2}$ an accelerometer at the top of the tower is replaced by an inclinometer. Two inclinometers are used in $\boldsymbol{y_3}$. In $\boldsymbol{y_4}$ an inclinometer is moved from the top to node 40.

The observability results obtained from the proposed algorithm are summarized in Table 1. When all the four sensors are accelerometers as in the measurement $\boldsymbol{y_1}$, both the full and substructural models are found unobservable: neither their dynamic states nor the unmeasured forces applied are observable, but all the unknown parameters are identifiable. The use of inclinometer(s) instead of accelerometer(s) improves the observability properties: the full model becomes observable for the measurements $\boldsymbol{y_2}$, $\boldsymbol{y_3}$ and $\boldsymbol{y_4}$, and the substructural model becomes fully observable for the measurement $\boldsymbol{y_3}$; the more unmeasured forces exist the more inclinometers are needed. Using the measurement $\boldsymbol{y_4}$ for the substructural model, 4 dynamic states, 1 parameter and the 2 unmeasured forces are found k-row unobservable for a large k using Algorithm III. A comparison between $\boldsymbol{y_3}$ and $\boldsymbol{y_4}$ implies that, a sensor should be considered to be placed at the location where the cut is introduced, i.e. node 41 in this example, in order to achieve feasible substructural identification.

4.4. A 2D FE model of a truss-beam bridge

The observability of a 2D FE model of a truss-beam bridge whose geometry is inspired by the design of Tokyo Gate Bridge [39] is examined in this example.

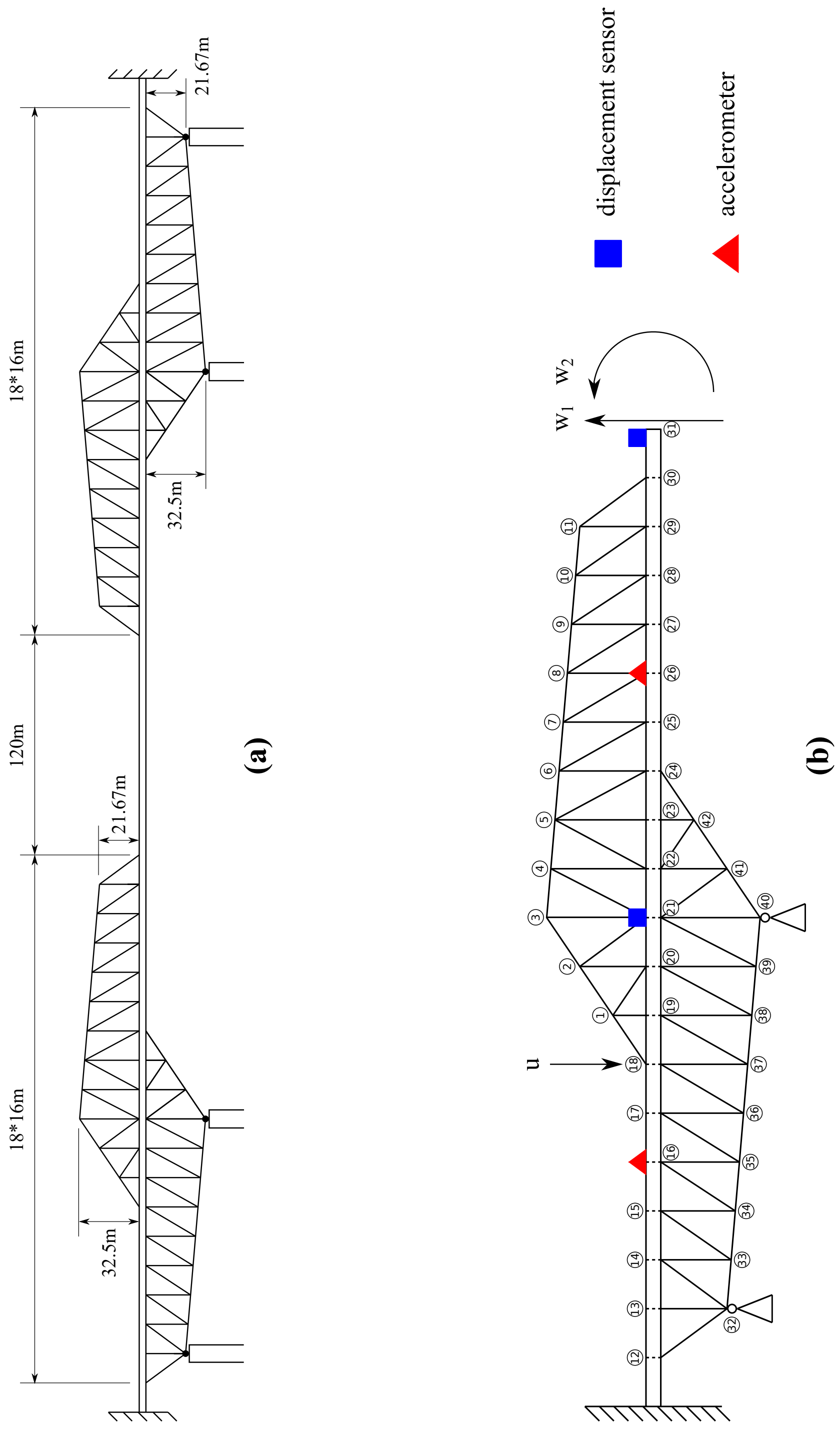

Figure 6: (a) The geometry of the truss-beam bridge, (b) the investigated substructural model of the bridge with a given sensor setup.

The bridge is a three-span structure of a steel box girder and steel bars in the arrangement of trusses, as shown in Figure 6(a). The investigated model is a substructure of the bridge, as shown in Figure 6(b), which occurs by introducing a cut at the central span. The model consists of the left 320m part of the girder and the associated trusses. An unmeasured shear force w_1 and an unmeasured bending moment w_2 are applied as a result of the cut to account for the dynamic effects from the remaining bridge structure. A measured excitation u is applied vertically at the marked position on the deck, which in practice could be provided by a hammer or an actuator.

The substructure of the bridge is modelled using linear frame-type macro-elements: the steel girder is discretized into 20 uniform Euler-Bernoulli beam elements, and the other pin-jointed members are treated as truss elements. Each of the truss nodes 1-11 and 32-42 has 2 DOFs, i.e. translational motions along horizontal and vertical axes. The axial stiffness of the truss element between nodes i and j is k_{ij}, which is a parameter to be identified. The mass matrix of each truss element is obtained by assuming the element mass is lumped as point masses along the translational DOFs at the ends. The point mass from all the jointed members lumped at each truss node is assumed to be known. Each of the beam nodes 12-31 has 2 DOFs, i.e. vertical motion and rotation with the axial motion being neglected. The stiffness and mass matrices of the beam element between nodes i and j follow equation (47), where the corresponding stiffness parameter EI_{ij} is to be identified. The left end of the girder is assumed to be fixed and the bottom truss is connected with pin-joints to the piers.

The state-space equations of the model are written in the following form:

$$\dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B_1}w_1 + \boldsymbol{B_2}w_2 + \boldsymbol{B_3}u, \quad \dot{\boldsymbol{\theta}} = \boldsymbol{0} \tag{55}$$

where the dynamic state vector $\boldsymbol{x}$ contains all the nodal displacements, rotations and the corresponding velocities. The parameter vector $\boldsymbol{\theta}$ includes all the unknown stiffness parameters of the truss and beam elements. The form of the system matrix $\boldsymbol{A}$ follows equation (51), where $\boldsymbol{K}$ and $\boldsymbol{M}$ are respectively the model stiffness and mass matrices constructed from the assembly of all the element stiffness and mass matrices of the truss and beam elements. $\boldsymbol{B_1}$, $\boldsymbol{B_2}$ and $\boldsymbol{B_3}$ are the input-affine vectors for the measured and unmeasured inputs to the model. Four sensors are installed on the deck at the positions shown in Figure 6(b), where two vertical displacements and two vertical accelerations are measured to record the response of the bridge to the applied excitation u.

The proposed algorithm reports that the substructural model of the bridge is fully observable. Based on the measurements obtained from the given sensor setup, it is theoretically feasible to estimate the unknown properties of the substructure and track its vibrations, as well as the unmeasured forces, using some suitably chosen system identification methods. Similarly to what was found in Example 4.3, the result of this example suggests the theoretical possibility of an alternative strategy that can be used to identify or perform health monitoring on a large structure (e.g. as for the bridge), especially when there is a lack of sensors. A substructure of the large structure, whose properties and dynamics are of interest, could be properly selected, modelled and identified by separating this substructure from the remaining structure through a carefully chosen cut, as long as the occurring structure is observable for the examined sensor setup. A counter-intuitive result is that a cut may lead to a subsystem whose observability properties are improved over the original full system despite the introduction of some unmeasured inputs. Such a strategy of substructural identification would help in tackling problems related to dimensionality and in using the available sensors more efficiently. In the cases of substructural identification and many other realistic scenarios, the observability algorithm proposed in this paper would allow researchers to freely investigate the viability of various system models and sensor patterns without concerning about any computational constraint.

5. Conclusions

In this paper, an efficient algorithm is developed to implement the extended Observability Rank Condition for testing the observability and identifiability of rational nonlinear systems subject to measured and unmeasured inputs. The framework of the algorithm aims to construct the observability matrix by extracting the coefficients of the power series associated with system outputs. To obtain the target power series, Newton's iteration is used to solve a variational system of ordinary differential equations for power series solutions. The occurring computational framework allows for using random numerical realizations of the involved variables and applying modular operations to control number growth, leading to a semi-numerical implementation that can handle robustly very large and complex systems. Compared to the standard symbolic implementation of the observability approach, repeated Jacobian operations are avoided throughout the proposed algorithm; the arithmetics of power series with numerical coefficients are substantially cheaper than those computations between symbolic expressions. Also due to the quadratic convergence property of Newton's iteration, a large number of the Lie derivatives can be computed through only a small number of iterations.

The proposed algorithm is applied to several illustrative examples of engineering systems including a high-rise shear building, finite element models of a wind turbine and a truss-beam bridge. Meanwhile, a practical guideline is provided for determining the unobservable quantities of a tested system. The examples demonstrate the superior capability of the algorithm to handle real-world engineering systems of large-size and high-complexity. To the best of authors' knowledge, none of the existing observability methods can have even close performance on the shown models. In addition, the investigation into substructural identification problems suggests that, it is feasible to identify a sub-system that is separated from a full system under properly designed sensor placement and introduction of unmeasured inputs. The result recommends a meaningful strategy for monitoring large systems from a theoretical observability/identifiability perspective.

Acknowledgement

The first author would like to gratefully acknowledge the financial support from the China Scholarship Council (CSC) for his Ph.D. studies.

References

[1] L. Ljung, System Identification: Theory for the User, 2nd Edition, Prentice Hall PTR, 1999.

[2] M. N. Chatzis, E. N. Chatzi, A. W. Smyth, On the observability and identifiability of nonlinear structural and mechanical systems, Structural Control and Health Monitoring 22 (2015) 574–593.

[3] A. Sedoglavic, A Probabilistic Algorithm to Test Local Algebraic Observability in Polynomial Time, Journal of Symbolic Computation 33 (2002) 735–755.

[4] B. Merkt, J. Timmer, D. Kaschek, Higher-order Lie symmetries in identifiability and predictability analysis of dynamic models, Physical Review E 92 (2015) 012920 1–9.

[5] X. Shi, M. N. Chatzis, Lie symmetries, observability and model transformation of nonlinear systems with unknown inputs, Proceedings of the XI International Conference on Structural Dynamics, EURODYN, Athens, Greece (2020) 901–911.

[6] H. Miao, X. Xia, A. S. Perelson, H. Wu, On Identifiability of Nonlinear ODE Models and Applications in Viral Dynamics, Society for Industrial and Applied Mathematics 53 (1) (2011) 3–39.

[7] A. Raue, C. Kreutz, T. Maiwald, J. Bachmann, M. Schilling, U. Klingmuller, J. Timmer, Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood, Bioinformatics 25 (15) (2009) 1923–1929.

[8] R. Hermann, A. J. Krener, Nonlinear Controllability and Observability, IEEE Transactions on Automatic Control AC-22 (5) (1977) 728–740.

[9] M. Anguelova, Observability and identifiability of nonlinear systems with applications in biology, Ph.D. thesis, Department of Mathematical Sciences,

Division of Mathematics, Chalmers University of Technology and Goteborg University (2007).

[10] A. Isidori, Nonlinear Control Systems, 3rd Edition, Springer-Verlag London, 1995.

[11] S. Diop, M. Fliess, Nonlinear observability, identifiability, and persistent trajectories, Proceedings of the 30th IEEE Conference on Decision and Control, Brighton, UK (1991) 714–719.

[12] H. Pohjanpalo, System Identifiability Based on the Power Series Expansion of the Solution, Mathematical Biosciences 41 (1978) 21–33.

[13] S. Vajda, K. R. Godfrey, H. Rabitz, Similarity transformation approach to identifiability analysis of nonlinear compartmental models, Mathematical Biosciences 93 (1989) 217–248.

[14] L. Ljung, T. Glad, On global identifiability for arbitrary model parametrizations, Automatica 30 (1994) 265–276.

[15] F. Boulier, D. Lazard, F. Ollivier, M. Petitot, Computing representations for radicals of finitely generated differential ideals, Applicable Algebra in Engineering, Communication and Computing 20 (2009) 73–121.

[16] G. Franco, R. Betti, R. W. Longman, On the uniqueness of solutions for the identification of linear structural systems, Journal of Applied Mechanics 73 (1) (2005) 153–162.

[17] K. Maes, M. N. Chatzis, R. Vandebril, G. Lombaert, Observability of modally reduced order models with unknown parameters, Mechanical Systems and Signal Processing 146 (2021) 1–16.

[18] S. Gillijns, B. De Moor, Unbiased minimum-variance input and state estimation for linear discrete-time systems with direct feedthrough, Automatica 43 (2007) 934–937.

[19] F. Naets, J. Croes, W. Desmet, An online coupled state/input/parameter estimation approach for structural dynamics, Computer Methods in Applied Mechanics and Engineering 283 (2015) 1167–1188.

[20] V. K. Dertimanis, E. N. Chatzi, S. Eftekhar Azam, C. Papadimitriou, Input-state-parameter estimation of structural systems from limited output information, Mechanical Systems and Signal Processing 126 (2019) 711–746.

[21] A. Martinelli, Extension of the Observability Rank Condition to Nonlinear Systems Driven by Unknown Inputs, Proceedings of the 23rd Mediterranean Conference on Control and Automation (MED), Torremolinos, Spain (2015) 589–595.

[22] K. Maes, M. N. Chatzis, G. Lombaert, Observability of nonlinear systems with unmeasured inputs, Mechanical Systems and Signal Processing 130 (2019) 378–394.

[23] B. De Jager, The use of symbolic computation in nonlinear control: is it viable?, IEEE Transactions on Automatic Control 40 (1) (1995) 84–89.

[24] A. F. Villaverde, N. Tsiantis, J. R. Banga, Full observability and estimation of unknown inputs, states and parameters of nonlinear biological models, The Royal Society Publishing 16 (2019) 20190043.

[25] N. Martinez, A. F. Villaverde, Nonlinear Observability Algorithms with Known and Unknown Inputs: Analysis and Implementation, Mathematics 8 (2020) 1876.

[26] A. Kugi, K. Schlacher, R. Novak, Symbolic computation for the analysis and synthesis of nonlinear control systems, Transactions on Engineering Sciences 22 (1999) 255–264.

[27] J. Karlsson, M. Anguelova, M. Jirstrand, An efficient method for structural identifiability analysis of large dynamic systems, Proceedings of the 16th

International Federation of Automatic Control, Brussels, Belgium 45 (2012) 941–946.

[28] X. Shi, M. S. Williams, M. N. Chatzis, A robust algorithm to test the observability of large linear systems with unknown parameters, Mechanical Systems and Signal Processing 157 (2021) 1–19.

[29] M. Spivak, Calculus on Manifolds: A Modern Approach to Classical Theorems of Advanced Calculus, Avalon Publishing, 1965.

[30] L. B. Rall, Computational Solution of Nonlinear Operator Equations, John Wiley & Sons, Inc., New York, 1969.

[31] R. P. Brent, H. T. Kung, Fast algorithms for manipulating formal power series, Journal of the Association for Computing Machinery 25 (4) (1978) 581–595.

[32] V. Shoup, A Computational Introduction to Number Theory and Algebra, 2nd Edition, Cambridge University Press, 2008.

[33] C. Bouillaguet, C. Delaplace, Sparse Gaussian Elimination Modulo p: an Update, International Workshop on Computer Algebra in Scientific Computing (2016) 101–116.

[34] I. Shparlinski, Finite fields: theory and computation: the meeting point of number theory, computer science, coding theory and cryptography, Springer Netherlands, 1999.

[35] MATLAB, version R2019b, The MathWorks Inc., Natick, Massachusetts, 2019.

[36] Maple, Maplesoft, a division of Waterloo Maple Inc., Maplesoft, Waterloo, Ontario, 2019.

[37] Wolfram, Mathematica, Wolfram Research Inc., Champaign, Illinois, 2019.

[38] A. K. Chopra, Dynamics of Structures: Theory and Applications to Earthquake Engineering, 4th Edition, Pearson Education, 2012.

[39] T. Yoneyama, Y. Fujii, Fabrication and erection of Tokyo Gate Bridge, IABSE-JSCE Joint Conference on Advances in Bridge Engineering-III (2015) 268–277.