Published under CC BY 4.0 International
https://creativecommons.org/licenses/

Data-based Identifiability and Observability Assessment for Nonlinear Control Systems using the Profile Likelihood Method

Thomas Schmitt * **Bastian Ritter** **

* *Control Methods & Robotics Lab, Technische Universität Darmstadt, Landgraf Georg Str. 4, 64283 Darmstadt, Germany (e-mail: thomas.schmitt@rmr.tu-darmstadt.de).*
** *Control Systems & Mechatronics Lab, Technische Universität Darmstadt, Landgraf Georg Str. 4, 64283 Darmstadt, Germany (e-mail: britter@iat.tu-darmstadt.de)*

Abstract: This paper introduces the profile likelihood method in order to assess simultaneously the parameter identifiability and the state observability for nonlinear dynamic state-space models with constant parameters. While a formal definition of a parameter's identifiability has been used before, the novel idea is to investigate also the state's observability by the identifiability of its initial value. Using the profile likelihood, both qualitative as well as quantitative statements are drawn from the analysis based on the nonlinear model and (possibly noisy) sensor data. A simplified wind turbine model is presented and used as an application example for the profile likelihood approach in order to investigate the model's usability for state and parameter estimation. It is shown that the critical model parameters and initial states are identifiable in principle. The analysis with more complex models and realistic data reveals the limitations when assumptions are deliberately violated in order to meet reality.

Keywords: Nonlinear systems, identifiability analysis, observability analysis, parameter estimation, profile likelihood.

1. INTRODUCTION

Today, in many control applications the system is described by ordinary differential equations (ODEs) and its states are used by the control algorithm, e. g. state space or model predictive control. Since the complete state vector is often not directly measurable, state estimators are used to observe them presuming the full observability of the states. However, for nonlinear systems, a state's observability is not an invariant system property (as is the case for linear systems), but depends on the system's stimulation. Let us assume a nonlinear model described by ODEs of the form

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t), \boldsymbol{\theta}), \tag{1a}$$

$$\boldsymbol{y}(t) = \boldsymbol{g}(\boldsymbol{x}(t), \boldsymbol{\theta}), \tag{1b}$$

where $\boldsymbol{x}(t) \in \mathbb{R}^n$ is the state vector, $\boldsymbol{u}(t) \in \mathbb{R}^m$ the input vector, $\boldsymbol{y}(t) \in \mathbb{R}^r$ the output vector. Furthermore, we respect the possibility of unknown parameters $\boldsymbol{\theta}$. Then, whether a state x_i is observable depens on the input $\boldsymbol{u}$ and the selection of unknown parameters $\boldsymbol{\theta}$. Moreover, a state can be more or less *well* observable, depending on *how* the system is stimulated.

If such unknown parameters $\boldsymbol{\theta}$ exist, they need to be identified at the same time. However, their identifiability depends on the stimulation, too. Thus, the problem becomes even more complex and the question arises how both identifiability and observability in such a setting can be assessed for real-world applications.

In this paper, we present a method by assessing the states' $\boldsymbol{x}$ observability by identifying their initial values $\boldsymbol{x}_0$ using the profile likelihood (PL) approach. Namely, by treating $\boldsymbol{x}_0$ as part of the unknown parameter vector $\boldsymbol{\theta}$, we convert the identifiability and observability problem into a mere parameter estimation problem. Moreover, using the PL, measurement data (either from simulation or real-world measurements) are used and both qualitative as well as quantitative assessments are derived.

Various other methods to solve such parameter estimation problems exist. However, most have significant drawbacks. More specifically, most of them do not allow both qualitative and quantitative statements at the same time by use of measurement data. The power series expansion (Pohjanpalo (1978)), similarity transformation (Vajda et al. (1989)), direct test (Denis-Vidal and Joly-Blanchard (2000)), implicit function theorem (Xia and Moog (2003)) and using differential algebra (Miao et al. (2011)), despite having other significant drawbacks (Schmitt (2017)), all only lead to qualitative results whether a model's parameter is identifiable in principle or not. There are several sensitivity based methods which do lead to quantitative results, but do not respect possible limitations of measurement data. The same holds for Monte Carlo simulations or the correlation matrix (Miao et al. (2011)). Markov Chain Monte Carlo methods do not have one of these limitations. However, they are complex to apply and typically computationally intensive, which can be prohibitive for more complex models (Vanlier et al. (2012)). The use of the

PL for parameter estimation problems as described above has been introduced in Raue et al. (2009). The authors distinguish between *structural* and *practical* identifiability and provide a formal definition for the latter one. In this paper, this definition for the case of structural identifiability is extended to the cases of structural and practical *observability* of the dynamic states.

The remainder of this paper is structured as follows. The methodology of the PL approach is discussed in Section 2 and the formal definitions of identifiability and observability are presented. Section 3 provides the application example of a wind turbine control system. In Section 4, the PL approach is applied to the given models using different data sets in order to assess both identifiability of parameters as well as the states' observability. Section 5 summarizes the relevant outcomes and provides the conclusions.

Notation: Vectors are written in bold letters, e. g. $\boldsymbol{\theta}$. $\chi^2_{\alpha,df}$ refers to the common χ^2-distribution. $\chi^2(\boldsymbol{\theta})$ refers to a likelihood in dependence of the parameter vector $\boldsymbol{\theta}$. $\chi^2_{\mathrm{PL,i}}(\mu)$ refers to the PL of parameter θ_i to have a value μ.

2. METHODOLOGY

2.1 Profile Likelihood Approach

The profile likelihood approach provides an attractive alternative to Fisher Information based confidence intervals. It has been shown that especially for nonlinear models, Fisher Information based Confidence intervals are often not appropriate (Raue et al. (2009)).

Its derivation starts with the maximum likelihood estimator (MLE). Assume a system as in (1a) and (1b) and d_k measurements of its r outputs at time points t_l. Assuming Gaussian measurement noise σ^D_{kl} for the k-th output at t_l, the likelihood to measure data $\boldsymbol{y}^D$ given a parameter vector $\boldsymbol{\theta}$ is given by

$$L(\boldsymbol{\theta}) = \prod_{k=1}^{r}\prod_{l=1}^{d_k} \frac{1}{\sqrt{2\pi}\sigma^D_{kl}} \exp\left(-\frac{1}{2}\left(\frac{y^D_{kl} - y_k(\boldsymbol{\theta}, t_l)}{\sigma^D_{kl}}\right)^2\right). \quad (2)$$

To find the MLE, we could maximize (2). Alternatively, we minimize two times the negative logarithm of this equation, i. e.

$$\chi^2(\boldsymbol{\theta}) := -2\log(L(\boldsymbol{\theta})). \quad (3)$$

On the one hand, this has numerical advantages. On the other hand, using the negative log likelihood has another useful characteristic. It can be shown that if regularity conditions apply, then for $n \to \infty$ data points it follows a χ^2-distribution with $\dim(\boldsymbol{\theta})$ degrees of freedom (see Meeker and Escobar (1995)). Equation (3) can be rewritten as

$$\chi^2(\boldsymbol{\theta}) = C + \sum_{k=1}^{r}\sum_{l=1}^{d} 2\log\left(\sigma^D_{kl}\right) + \sum_{k=1}^{r}\sum_{l=1}^{d_k}\left(\frac{y^D_{kl} - y_k(\boldsymbol{\theta}, t_l)}{\sigma^D_{kl}}\right)^2 \quad (4)$$

where C subsumes parts of the likelihood function which do not depend on the parameters and as such do not influence the location of the minimum. If the measurement noise σ^D_{kl} is constant, then the second term in (4) is constant as well. Minimizing this expression results in the MLE,

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \left[\chi^2(\boldsymbol{\theta})\right]. \quad (5)$$

Once a model sufficiently predicts the experimental data, confidence intervals (CIs) can be computed. Such CIs can be obtained using the Profile Likelihood approach. To calculate the profile likelihood for one specific parameter, we initialize at the best parameter fit. Subsequently, the parameter to be profiled is incremented iteratively, while re-optimizing all other parameters,

$$\chi^2_{\mathrm{PL,i}}(\mu) = \min_{\boldsymbol{\theta}\in\{\boldsymbol{\theta}|\theta_i=\mu\}} \left[\chi^2(\boldsymbol{\theta})\right]. \quad (6)$$

This procedure yields an optimal path through parameter space for each θ_i, thereby providing information on the minimum likelihood change induced by forcing θ_i away from its best fit parameter value. It can be shown that, under weak assumptions, the region for which

$$D := \chi^2_{\mathrm{PL,i}}(\mu) - \chi^2(\hat{\boldsymbol{\theta}}) \le \chi^2_{\alpha,df} \quad (7)$$

holds yields the CI for the parameter θ_i to a given confidence level α (see Feder (1968)). The threshold $\chi^2_{\alpha,df}$ is the α-quantile of the χ^2-distribution with df degrees of freedom. [1] Usually, 1 D-profiles are calculated, i. e. pointwise CIs with $df = 1$ are derived (see also Meeker and Escobar (1995)). Note that the order of parameters being profiled is arbitrary.

2.2 Structural and Practical Identifiability

In nonlinear models, it is possible that the CI for a parameter is unbounded (non-identifiable). This means that no lower and/or upper limit can be calculated. Thereby, two types of non-identifiabilities can be detected. The first, often referred to as *structural non-identifiability*, refers to profiles which are completely flat, remaining constant for arbitrary values of the profiled parameter. Such non-identifiabilities occur when can analytically not be distinguished in the measured observables $\boldsymbol{y}$. The second type of non-identifiability often termed *practical non-identifiability* is related to the data (Raue et al. (2009)). In both cases, the profile parameter paths can be informative for determining how to appropriate reduce the model (see Maiwald et al. (2016)).

To the authors' knowledge, so far only practical identifiability has been defined formally. Thus, we define structural identifiability as follows:

Definition 1. A parameter θ_i is *structurally non-identifiable*, if its likelihood-based confidence interval is infinite, $[-\infty, +\infty]$, and its profile likelihood has no unique minimum, but is flat over all θ_i.

The definition for practical identifiability shall be cited from Raue et al. (2009):

Definition 2. A parameter θ_i is *practically non-identifiable*, if the likelihood-based confidence region is infinitely extended in increasing and/or decreasing direction of θ_i, although the likelihood has a unique minimum for this parameter.

Both cases are illustrated in Figure 1. Note that the absolute value of χ^2_{PL} is arbitrary; only whether and where it exceeds the CI threshold $\chi^2_{\alpha,df}$ matters.

[1] The χ^2-distribution does not coincide with (4). The nomenclature may be unfortunate, but has been adopted from Raue et al. (2009) for better comparability.

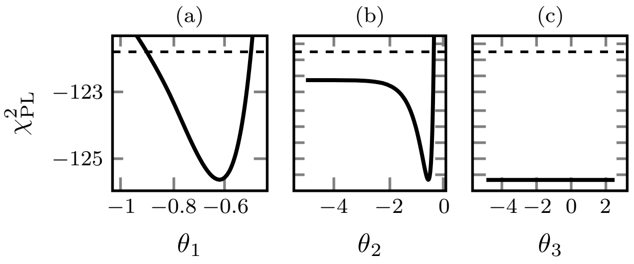

Fig. 1. Exemplary profile likelihood curves for different parameters θ_1, θ_2 and θ_3. The dashed line indicates the bounds for the confidence interval. θ_1 is structurally and practically identifiable, θ_2 is structurally, but not practically identifiable and θ_3 is neither practically nor structurally identifiable, i. e. non-identifiable.

2.3 Structural and Practical Observability

While various definitions for the observability of a nonlinear dynamical system exist; in any case it can be expressed as the initial state vector's distinguishability. By recalling the definition from Adamy (2018), it becomes clear that a system's observability depends on whether the initial state vector $\boldsymbol{x}_0$ can be determined or not:

Definition 3. A system as described in (1) shall be defined for $\boldsymbol{x} \in D_x \subseteq \mathbb{R}^n$ and $\boldsymbol{u} \in C_u \subseteq \mathbb{C}^{n-1}$ with $\boldsymbol{y} \in \mathbb{R}^r$ and $\boldsymbol{x}(t_0) = \boldsymbol{x}_0$. If all starting vectors $\boldsymbol{x}_0 \in D_x$ are uniquely determinable from the knowledge of $\boldsymbol{u}(t)$ and $\boldsymbol{y}(t)$ in a finite time interval $[t_0,\ t_1 < \infty]$ for all $\boldsymbol{u} \in C_u$, then the system is called *observable.*

Note that the unknown parameter vector $\boldsymbol{\theta}$ in (1a) is not respected in Adamy (2018). However, this does not affect the interpretation of observability. Moreover, the initial value of a single state $x_{\mathrm{i},0}$ is eventually a model parameter. Thus, the qualitative observability of a state can be determined by the identifiability of the initial value. Furthermore to formalize this approach, two definitions are introduced. First, the structural observability of a state is defined analogously to the structural identifiability of a parameter.

Definition 4. A state x_i is *structurally unobservable*, if the likelihood-based confidence interval of its initial value $x_{i,0}$ is infinite, $[-\infty, +\infty]$, and its profile likelihood has no unique minimum, but is flat over all x_i. Otherwise, it is *structurally observable.*

Note that the term practical observability has already been used in Kreutz et al. (2012), but with a different interpretation. There, CIs for a model's 'observation', which can be any combination of model information, are derived for specific points in time. Here, we refer to observability of a control system's state as defined in Definition 3 and formalize practical observability as follows.

Definition 5. A state x_i is *practically unobservable*, if the likelihood-based confidence region of its initial value is infinitely extended in increasing and/or decreasing direction of x_i, although the likelihood has a unique minimum for this parameter.

3. WIND TURBINE APPLICATION EXAMPLE

As an illustrative application example, this paper investigates the nonlinear system dynamics of a horizontal axis wind turbine. A simplified dynamic model reads

$$\ddot{\varphi}_{\mathrm{r}} = \frac{\varrho}{2}\frac{\pi R^3}{\Theta} C_{\mathrm{M}}(\lambda)\big(v_{\mathrm{w}}(t) - \dot{x}_{\mathrm{T}}\big)^2 - \frac{i_{\mathrm{gb}}}{\Theta} M_{\mathrm{g}} \tag{8a}$$

$$\ddot{x}_{\mathrm{T}} = \frac{\varrho}{2}\frac{\pi R^2}{m_{\mathrm{T}}} C_{\mathrm{T}}(\lambda)\big(v_{\mathrm{w}}(t) - \dot{x}_{\mathrm{T}}\big)^2 - 2D\omega_0\dot{x}_{\mathrm{T}} - \omega_0^2 x_{\mathrm{T}} \tag{8b}$$

which captures the drive-train dynamics as well as the nacelle motion for the partial load regime appropriately (cf. Ritter et al. (2016a) for more details).

The three dynamic states are the rotor angular velocity $\dot{\varphi}_{\mathrm{r}}$, the nacelle velocity $\dot{x}_{\mathrm{T}}$ and the position x_{T}. Two unknown model parameters are most relevant, the eigenfrequency ω_0 and the wind velocity v_{w}, where the first changes slowly and the second more quickly over time. The rotor aerodynamics are represented by the coefficients of torque $C_{\mathrm{M}}(\lambda)$ and thrust $C_{\mathrm{T}}(\lambda)$ which depend in this simplified case solely on the tip-speed ratio λ (TSR), defined as

$$\lambda = \frac{\dot{\varphi}_{\mathrm{r}} R}{v_{\mathrm{w}} - \dot{x}_{\mathrm{T}}}. \tag{9}$$

Equation (9) constitutes the dimensionless rotor speed of a wind turbine which shows a distinguished value λ^* that is optimal for electrical power production. A simplified approximation for the aerodynamic coefficients is

$$C_{\mathrm{M}}(\lambda) = c_{\mathrm{m},2}\lambda^2 + c_{\mathrm{m},1}\lambda + c_{\mathrm{m},0} \tag{10a}$$

$$C_{\mathrm{T}}(\lambda) = c_{\mathrm{t},2}\lambda^2 + c_{\mathrm{t},1}\lambda + c_{\mathrm{t},0} \tag{10b}$$

where both are assumed as second-order polynomials in λ. All model parameters are gathered and explained in Tab. 1. The numerical values therein have been evaluated for a well-known reference wind turbine, defined by Jonkman et al. (2009), with a rated electrical power 5 MW.

Table 1. Model parameters and its numerical values

Symbol	Description	Value	Unit
ϱ	air mass density	1.225	kg/m^3
ω_0	nacelle eigenfrequency	2.1	1/s
m_{T}	equiv. tower top mass	$450{\cdot}10^3$	kg
R	blade tip radius	63	m
D	modal damping ratio	0.01	-
i_{gb}	drive-train gear-box ratio	97	-
Θ	drive-train inertia	$4.05{\cdot}10^7$	kg m^2
λ^*	optimal tip-speed-ratio	7.7	-
k_{P}^*	optimal controller gain	$23.8{\cdot}10^{-3}$	Nm/rpm^2
$c_{\mathrm{m},0}$	constant torque coeff.	$10{\cdot}10^{-2}$	-
$c_{\mathrm{m},1}$	linear torque coeff.	$-17{\cdot}10^{-4}$	-
$c_{\mathrm{m},2}$	quadratic torque coeff.	$-40{\cdot}10^{-5}$	-
$c_{\mathrm{t},0}$	constant thrust coeff.	$-16{\cdot}10^{-2}$	-
$c_{\mathrm{t},1}$	linear thrust coeff.	$18{\cdot}10^{-2}$	-
$c_{\mathrm{t},2}$	quadratic thrust coeff.	$-77{\cdot}10^{-4}$	-

The measured outputs of the nonlinear system, cf. (8), are the generator speed $n_{\mathrm{g}} = (30/\pi)\, i_{\mathrm{gb}}\dot{\varphi}_{\mathrm{r}}$ and the nacelle acceleration $\ddot{x}_{\mathrm{T}}$. The only controlled input is the electrical generator torque M_{g}. The disturbance input is the rotor-effective wind velocity $v_{\mathrm{w}}(t)$ which is also considered as unknown time-varying input parameter in the following. If required, a more complex wind turbine model can be found in Ritter et al. (2016b).

In order to perform reasonable simulations of the wind turbine control system with changing wind velocity v_w, a simple and ideally an optimal control law is needed. For this purpose, the nonlinear controller

$$M_\mathrm{g} = k_\mathrm{P}^* n_\mathrm{g}^2 = k_\mathrm{P}^* \left(30/\pi\right)^2 i_\mathrm{gb}^2 \dot{\varphi}_\mathrm{r}^2 \tag{11}$$

with proportional feedback proves to be useful, see e. g. Burton et al. (2011) (p. 485 ff). Note that this control law is reference-free and thus the optimal operating point is found automatically in order to track λ^*. Optimality is considered in this context only for the power production and not for mechanical loads. Figure 2 shows illustrative simulation results obtained with the optimal torque controller. Due to the time delay in the generator speed with respect to the unknown and time-varying wind speed, λ^* is only

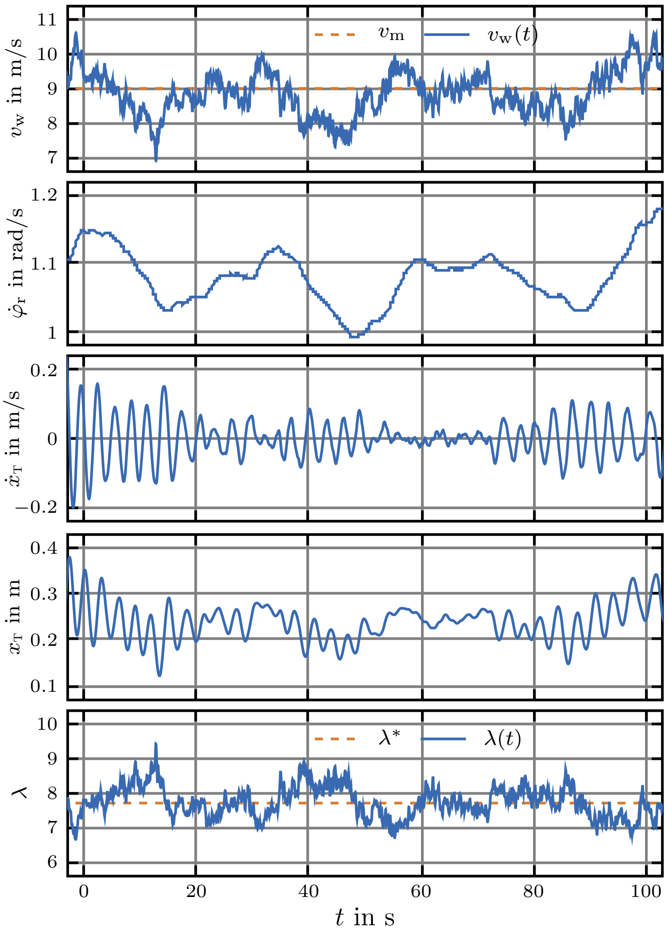

Fig. 2. Closed-loop simulation results for a turbulent wind field with an average wind speed $v_\mathrm{m} = 9\,\mathrm{m/s}$

reached on average. These simulation data are used in Sect. 4.2 to assess the identifiability when the assumption of a constant disturbance input is violated deliberately.

The reason for conducting this analysis is to make sure that the relevant states and parameters are identifiable based on the two measurable outputs. In this context, an analysis based on empirical Gramian matrices has already been conducted in Ritter et al. (2018) using a similar model. In the following, the focus now rests on the novel approach using the profile likelihood approach.

4. RESULTS AND DISCUSSION

The nonlinear wind turbine model (8) has been implemented in the toolbox *Data2Dynamics* (D2D, Raue et al. (2015)). D2D ships with a comprehensive functionality to apply the PL approach to nonlinear models. However, so far it has mainly been used for identification purposes of biological systems, i. e. many unknown parameters and few known data. Thus, some enhancements were necessary for this work, e. g. the efficient implementation of large input data sets by splines. Further details on the implementation can be found in Schmitt (2017).

4.1 Constant Wind Velocity

In all following simulation results, we identify two model parameters, i. e. the eigenfrequency ω_0 and the wind velocity v_w and all three initial states. For this section, we first use data generated with a constant v_w. The data consist of 200 measurement points with a sampling time of 0.1 s from three different sources. First, we use simulation data from the 3-states model (8) itself, which we call the *basic* model. Second, we use simulation data from a more complex wind turbine model with 8 states, which we refer to as the *standard* model. Third, we use simulation data from a sophisticated model with 21 states, called the *advanced* model. In any case, the simulation data is overlayed with normally distributed noise. Thereby, a signal-to-noise ratio of 30 dB is chosen and the same noise realization is used for every data set. The measurement error parameters σ_{kl}^D from (4) are treated as free parameters and estimated together with all other parameters, i. e. by solving (5). However, once the MLE is found, they are set constant and are not further adapted when calculating the PL for all other parameters. Note that D2D allows to formulate a priori knowledge about the distribution of each parameter θ_i, such as minimum and maximum possible values. Thus, together with the possibility to re-run the optimization of (5) with different initial values, no convergence problems occur with the model presented here.

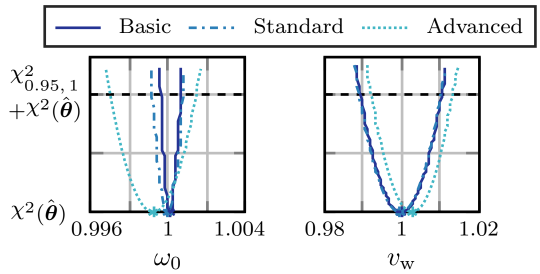

Fig. 3. PLs of the normalized unknown parameters ω_0 and v_w with pointwise CIs at $\alpha = 5\,\%$

For ease of presentation, all parameters and initial states have been normalized to 1. Namely, the true value which should be identified is 1. Figure 3 shows the PLs for both unknown parameters. Both parameters are structurally as well as practically identifiable with every data set, since their CIs are finite. Using noisy data from the basic model itself, both parameters are identified with an error $\leq 0.012\,\%$ and the CIs contain the true values. Using data from the standard model, the MLE errors are similar, but the CI for ω_0 increases. For data from the realistic advanced model, the estimation erorrs for both parameters and the CI for ω_0 increase significantly. However, the errors are still acceptably small with 0.07 % for ω_0 and 0.27 % for v_w.

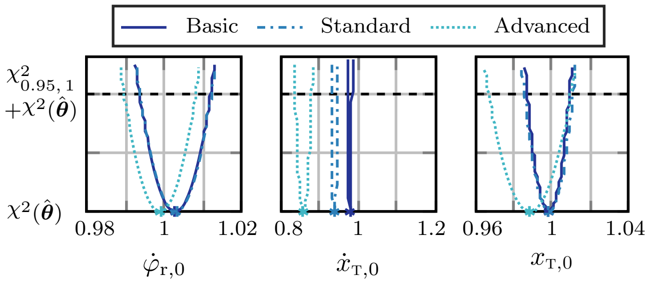

Fig. 4. PLs of the normalized initial states with pointwise CIs at $\alpha = 5\,\%$. Note that every data set (basic, standard and advanced) had a different initial value.

Figure 4 shows the results for initial states. All states are both structurally and practically observable, even with a constant wind velocity. However, the estimation errors are larger compared to the parameters ω_0 and v_w. For $\dot{\varphi}_{\mathrm{r},0}$, the estimation errors are all $\leq 0.3\,\%$ and all CIs contain the true value. For $x_{\mathrm{T},0}$, the PLs for data from the basic and standard model are comparable, i. e. the estimation errors are negligible with $\approx 0.16\,\%$ and the CI for the standard model data is only slightly larger. However, for data from the advanced model, the estimation error increases to 1.16 % and the CI is significant increased. For $\dot{x}_{\mathrm{T},0}$, the differences between the different data sets are most severe. In general, $\dot{x}_{\mathrm{T}}$ is less well observable. Namely, the estimation errors are 2.18 %, 6.18 % and 14.25 % for data from the basic, standard and advanced model, respectively. Moreover, this is the only parameter for which the CIs do not contain the true value. Note that the 95 %-confidence level is not valid for the measurements from the standard and advanced model in general, since the PL approach assumes only measurement noise but no model errors. For the data from the basic model itself, the necessary assumption of $n \to \infty$ data points is still violated, but further simulations showed that raising the confidence level to e. g. 99.99 % leads to a CI which does include the true value.

Figure 5 shows the model's outputs for measurement data from the basic model. Despite the significant noise on $\ddot{x}_{\mathrm{T}}$, the real trajectories for both $\ddot{x}_{\mathrm{T}}$ and n_g are followed very accurately.

Concluding, the results show the method's validity if all assumptions are met, i. e. there is no model error and v_w really is a constant parameter. Next, it is analyzed how well the method is applicable for turbulent wind.

4.2 Turbulent Wind Field

In reality, the wind velocity v_w is not constant but can be highly turbulent. However, the PL approach is designed for constant parameters only. Thus, in this section we test its applicability under real conditions and use the turbulent wind field shown in Figure 2. Measurement data is used only from the advanced model and the sampling time is reduced to 0.01 s to increase the total number of data points. Otherwise, the settings are the same as before.

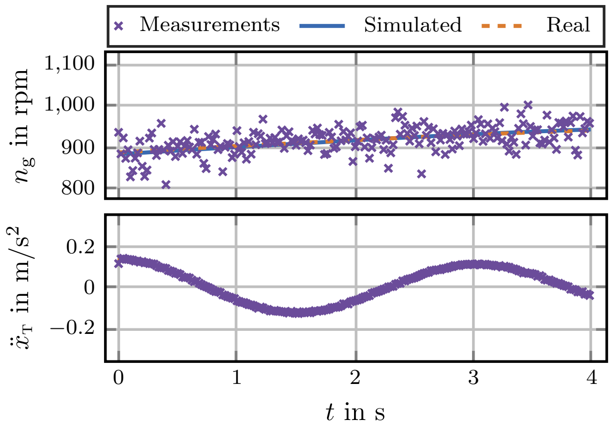

Fig. 5. Outputs from wind turbine model simulated with estimated parameters and initial states. Constant $v_w = 9$ m/s and noisy data from the basic model itself.

Figure 6 shows system outputs for a simulation time of 4 s. v_w is estimated to 9.08 m/s and the resulting n_g still follows the real one very accurately, since the wind turbulence's impact on n_g has a high time delay. However, in comparison to Figure 5, the simulated $\ddot{x}_{\mathrm{T}}$ is estimated significantly worse and cannot follow the abrupt changes. This is due to the direct impact of v_w on $\ddot{x}_{\mathrm{T}}$, see (8b).

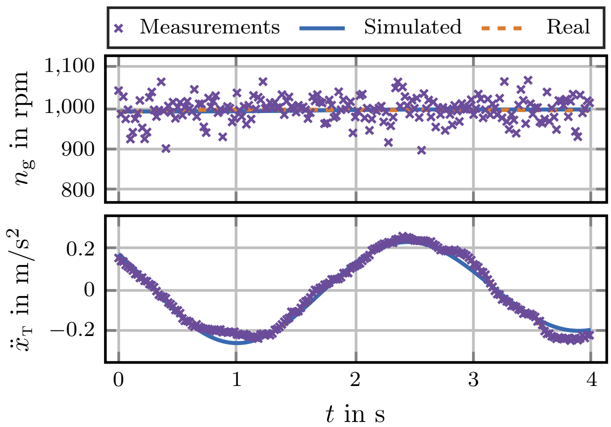

Fig. 6. Outputs from wind turbine model simulated with estimated parameters and initial states. Turbulent wind velocity and noisy data from the advanced model over 4 s.

This effect becomes more severe for longer simulation times. Figure 7 shows the results for a period of 10 s. Since the wind velocity is not only turbulent, but increases over time, the simulated $\ddot{x}_{\mathrm{T}}$ overestimates the first and underestimates the last oscillations. This is in addition to the abrupt changes which the model with a constant v_w cannot follow. Furthermore, the badly estimated peaks for $\ddot{x}_{\mathrm{T}}$ now also lead to a detectable estimation error for n_g.

5. CONCLUSION

This contribution has newly introduced the formal definition of structural and practical observability of nonlinear system states based on the PL method. The state's observability is evaluated by the identifiability of its initial value.

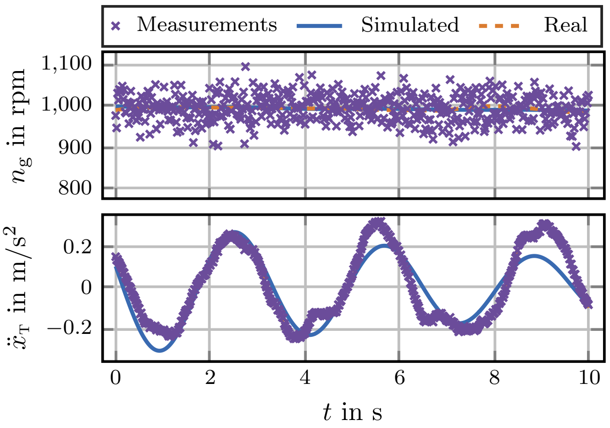

Fig. 7. Outputs from wind turbine model simulated with estimated parameters and initial states. Turbulent wind velocity and noisy data from the advanced model over 10 s.

Thereby, the observability is assessable qualitatively and quantitatively with realistic measurement data and sensor noise. Thus, it offers more potential for a priori analyses compared to sensitivity based methods. The PL approach is applied successfully to a nonlinear wind turbine control system with three states and two unknown critical model parameters. The results show that if both the generator speed and the nacelle tower acceleration are observed, all three states are structurally and practically observable and the eigenfrequency and wind velocity are structurally and practically identifiable. This is true for data from the model itself as well as from more complex models. However, while the PL is designed to handle Gaussian measurement noise and identify constant parameters, an increasing modeling error and time-varying parameters lead to larger estimation errors and CIs become invalid. For the presented application, this leads to increased but still acceptable estimation errors for control purposes.

REFERENCES

Adamy, J. (2018). *Nichtlineare Systeme und Regelungen*. Springer Vieweg, Berlin Heidelberg, 3. edition. doi: 10.1007/978-3-642-00794-1.

Burton, T., Jenkins, N., Sharpe, D., and Bossanyi, E. (2011). *Wind Energy Handbook*. John Wiley and Sons, 2. edition.

Denis-Vidal, L. and Joly-Blanchard, G. (2000). An easy to check criterion for (un)identifiability of uncontrolled systems and its applications. *IEEE Transactions on Automatic Control*, 45(4), 768–771.

Feder, P.I. (1968). On the distribution of the log likelihood ratio test statistic when the true parameter is ”near” the boundaries of the hypothesis regions. *The Annals of Mathematical Statistics*, 39(6), 2044–2055.

Jonkman, J., Butterfield, S., Musial, W., and Scott, G. (2009). Definition of a 5-MW reference wind turbine for offshore system development. Technical Report Technical Report NREL/TP-500-38060 February 2009, National Renewable Energy Laboratory (NREL).

Kreutz, C., Raue, A., and Timmer, J. (2012). Likelihood based observability analysis and confidence intervals for predictions of dynamic models. *BMC Systems Biology*, 6(1), 120. doi:10.1186/1752-0509-6-120.

Maiwald, T., Hass, H., Steiert, B., Vanlier, J., Engesser, R., Raue, A., Kipkeew, F., Bock, H.H., Kaschek, D., Kreutz, C., et al. (2016). Driving the model to its limit: Profile likelihood based model reduction. *PloS one*, 11(9), e0162366.

Meeker, W.Q. and Escobar, L.A. (1995). Teaching about approximate confidence regions based on maximum likelihood estimation. *The American Statistician*, 49(1), 48–53.

Miao, H., Xia, X., Perelson, A.S., and Wu, H. (2011). On identifiability of nonlinear ODE models and applications in viral dynamics. *SIAM Review*, 53(1), 3–39. doi: 10.1137/090757009.

Pohjanpalo, H. (1978). System identifiability based on the power series expansion of the solution. *Mathematical biosciences*, 41(1-2), 21–33.

Raue, A., Kreutz, C., Maiwald, T., Bachmann, J., Schilling, M., Klingmller, U., and Timmer, J. (2009). Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics*, 25(15), 1923. doi: 10.1093/bioinformatics/btp358.

Raue, A., Steiert, B., Schelker, M., Kreutz, C., Maiwald, T., Hass, H., Vanlier, J., Tnsing, C., Adlung, L., Engesser, R., Mader, W., Heinemann, T., Hasenauer, J., Schilling, M., Hfer, T., Klipp, E., Theis, F., Klingmller, U., Schberl, B., and Timmer, J. (2015). Data2Dynamics: a modeling environment tailored to parameter estimation in dynamical systems. *Bioinformatics*, 31(21), 3558. doi:10.1093/bioinformatics/btv405.

Ritter, B., Mora, E., and Schild, A. (2018). Observability analysis for horizontal axis wind turbines using empirical gramian matrices. *Journal of Physics: Conference Series*, 1102(012021), 1–11. doi:10.1088/1742-6596/1102/1/012021.

Ritter, B., Schild, A., Feldt, M., and Konigorski, U. (2016a). The design of nonlinear observers for wind turbine dynamic state and parameter estimation. *Journal of Physics: Conference Series*, 753(052029), 1–12. doi: 10.1088/1742-6596/753/5/052029.

Ritter, B., Schild, A., and Konigorski, U. (2016b). Making nonlinear state estimation techniques ready for use in industrial wind turbine control systems. In *Proc. of the WindEurope Summit 2016, Hamburg*.

Schmitt, T. (2017). Identifiability and observability assessment for nonlinear wind turbine control systems. Master's thesis. Technische Universität Darmstadt. URL `https://tuprints.ulb.tu-darmstadt.de/9264/`.

Vajda, S., Godfrey, K.R., and Rabitz, H. (1989). Similarity transformation approach to identifiability analysis of nonlinear compartmental models. *Mathematical biosciences*, 93(2), 217–248.

Vanlier, J., Tiemann, C.A., Hilbers, P.A.J., and van Riel, N.A.W. (2012). A Bayesian approach to targeted experiment design. *Bioinformatics*, 28(8), 1136. doi:10.1093/bioinformatics/bts092.

Xia, X. and Moog, C.H. (2003). Identifiability of nonlinear systems with application to HIV/AIDS models. *IEEE transactions on automatic control*, 48(2), 330–336.