Technical report from Automatic Control at Linköpings universitet

On Indirect Input Measurements

Jonas Linder*, Martin Enqvist*

Division of Automatic Control

E-mail: jonas.linder@liu.se, maren@isy.liu.se

13th March 2015

Report no.: LiTH-ISY-R-3080

Submitted to the 17th IFAC Symposium on System Identification

Address:

*Department of Electrical Engineering
Linköpings universitet
SE-581 83 Linköping, Sweden

WWW: http://www.control.isy.liu.se

AUTOMATIC CONTROL
REGLERTEKNIK
LINKÖPINGS UNIVERSITET

Abstract

A common issue with many system identification problems is that the true input to the system is unknown. In this paper, a framework, based on indirect input measurements, is proposed to solve the problem when the input is partially or fully unknown, and cannot be measured directly. The approach relies on measurements that indirectly contain information about the unknown input. The resulting indirect model formulation, with both direct- and indirect input measurements as inputs, can be used to estimate the desired model of the original system. Due to the similarities with closed-loop system identification, an iterative instrumental variable method is proposed to estimate the indirect model. To show the applicability of the proposed method, it is applied to data from an inverted pendulum experiment with good results.

Keywords: System identification, Model structure, Physical models, Instrumental Variable, Closed-loop

1 Introduction

System identification is the science of building dynamical models from system data, i.e. signals. The objective is to estimate a model between an input and the resulting output. A common assumption is that the relevant input and output signals are available or known. One of the major challenges is to separate the effects of disturbances, present in the in- and output signals, from the relevant input-output behavior that would be preferable to use for estimating the model. A first example is the classical open-loop system identification problem, where the input is assumed to be known while the output is corrupted by an unknown disturbance (Ljung, 1999). Other examples are the classical closed-loop and errors-in-variables (EIV) problems, where the disturbances are of different characters and thus present different challenges. In the closed-loop problem, the input is typically known but is correlated with the same disturbance as the output (Forssell, 1999), and in the EIV problem, the input is assumed to be measured with noise and the focus is on the input's dependency on disturbances (Söderström, 2007).

In this paper, we are presenting a framework where the input vector u is assumed to be partially unknown but the effects of the inputs are visible in several measurements. The goal is to estimate a model $G_O(\mathsf{p})$ from the partially unknown input vector u to a set of the available measurements y_O. Two type of external signals are assumed to affect the system, user-controlled signals affecting the input, and unknown disturbances affecting both the in- and output. It is assumed that all elements of the input vector can be correlated with the external user-controlled signals.

The input is divided into known inputs u_K, inputs u_D that are directly measured with noise through a known transfer function and finally, inputs u_I that are indirectly measured with noise. An indirect input measurement is a signal that is affected by the input u_I and thus, indirectly contains information about u_I. An example is of course the output y_O, but here an indirect input measurements should be thought of as a measurement in addition to the output y_O. By formulating an indirect model, where the known input, the direct and the indirect input measurements are used as inputs, in terms of the original model $G_O(\mathsf{p})$, it is possible to estimate the dynamical model of interest.

The resulting estimation problem is similar to closed-loop system identification due to the correlation of the disturbance to both the input and output of the indirect model. In fact, due to the assumed generality of the framework, many classical and recently proposed models are recovered as special cases of the indirect model. For instance, if all inputs are assumed to be known, the open-loop or closed-loop case is recovered depending on the relations with the external disturbances. An approach based on the external user-controlled signal and an iterative instrumental variable method is proposed to estimate the wanted dynamical model.

The remainder of this paper is organized as follows: In Section 2 indirect input measurements are introduced, the indirect model is derived, a couple of examples of models with indirect input measurements are presented and finally, identifiability issues are discussed. In Section 3, an iterative instrumental variable method is proposed to estimate the indirect model. It is complemented with a discussion about discretization of the indirect model and how to overcome certain identifiability issues by using several datasets. The proposed method is applied to one of the examples from Section 2 together with experimental data in Section 4 and finally, in Section 5, the paper is summarized with conclusions.

2 Indirect input measurements

Figure 1 shows a system $G_O(\mathsf{p})$ with input $u \in \mathbb{R}^{n_u}$ and output $y_O \in \mathbb{R}^{n_O}$. The goal is to estimate a model of the transfer function matrix $G_O(\mathsf{p})$ from the relation

$$y_O = G_O(\mathsf{p})u + H_O(\mathsf{p})\tau + e_O \tag{1}$$

where p is the differential operator, $H_O(\mathsf{q})$ is a transfer function matrix, τ is a vector of unknown external disturbances and e_O is measurement noise. The input u is assumed to be given by

$$u = F_\delta(\mathsf{p})\delta + F_\tau(\mathsf{p})\tau \tag{2}$$

where $F_\delta(\mathsf{p})$ is a possibly unknown transfer function matrix, δ is a known vector of user-controllable external signals and $F_\tau(\mathsf{p})$ is a transfer function matrix. Note that (1) can be part of a bigger system since both u and y_O are dependent on τ, and that both $H_O(\mathsf{p})$ and $F_\tau(\mathsf{p})$ can have elements equal to zero. For instance, the input u and the output y_O can be signals in a system operating in closed-loop.

It is assumed that the input is partially unknown and that it can be divided into

$$u = \begin{bmatrix} u_K \\ \hline u_I \\ u_D \end{bmatrix} \begin{matrix} \} \text{ known} \\ \left.\vphantom{\begin{matrix} u_I \\ u_D \end{matrix}}\right\} \text{ unknown} \end{matrix} \tag{3}$$

with the transfer function matrix divided into

$$G_O(\mathsf{p}) = \begin{bmatrix} G_{OK}(\mathsf{p}) & G_{OI}(\mathsf{p}) & G_{OD}(\mathsf{p}) \end{bmatrix} \tag{4}$$

The unknown inputs are here divided into inputs $u_D \in \mathbb{R}^{n_D}$ directly measured with noise and inputs $u_I \in \mathbb{R}^{n_I}$ that are indirectly measured with noise. Here, indirect refers to that a measurement of the input u_I is not available, but that there is an extra measurement that is affected by the input u_I and thus, indirectly contains information about u_I.

Assuming that these direct and indirect measurements are available, (1) can be extended with the outputs $y_D \in \mathbb{R}^{n_D}$ corresponding to the directly

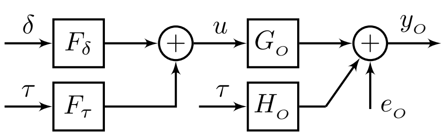

Figure 1: The system $G_O(\mathsf{p})$ represented as a block diagram showing the dependencies on the external user-controllable signal δ and external disturbance τ. Note that both the input u and the output y_O are assumed to be correlated with τ.

measured inputs and the outputs $y_I \in \mathbb{R}^{n_I}$ corresponding to the indirect input measurements, which gives

$$\underbrace{\begin{bmatrix} y_O \\ y_I \\ y_D \end{bmatrix}}_{y} = \underbrace{\left[\begin{array}{cc|c} G_{OK} & G_{OI} & G_{OD} \\ G_{IK} & G_{II} & G_{ID} \\ \hline 0 & 0 & G_{DD} \end{array}\right]}_{G(\mathsf{p})} \underbrace{\begin{bmatrix} u_K \\ u_I \\ u_D \end{bmatrix}}_{u} + \underbrace{\begin{bmatrix} H_O \\ H_I \\ 0 \end{bmatrix}}_{H(\mathsf{p})} \tau + e \tag{5}$$

where the transfer functions are of suitable dimensions, G_{II} is assumed to be invertible, G_{DD} is assumed to be known and invertible, and e is a vector of measurement noises. Here, the dependencies on p have been dropped for brevity.

2.1 The indirect model

The unknown inputs of (5), and thus of (1), can now be eliminated by algebraic manipulation. Firstly, the noise-free output is given by

$$\begin{bmatrix} \mathring{y}_O \\ \mathring{y}_I \\ \mathring{y}_D \end{bmatrix} = \begin{bmatrix} y_O \\ y_I \\ y_D \end{bmatrix} - \begin{bmatrix} H_O \\ H_I \\ 0 \end{bmatrix} \tau - e \tag{6}$$

Secondly, solving the third row for u_D gives

$$u_D = G_{DD}^{-1} \mathring{y}_D, \tag{7}$$

Thirdly, the second row is solved for u_I and combined with (7) which gives

$$u_I = G_{II}^{-1} \left[\mathring{y}_I - G_{IK} u_K - G_{ID} G_{DD}^{-1} \mathring{y}_D \right] \tag{8}$$

and finally, the first row is combined with (7) and (8) giving

$$\mathring{y}_O = \tilde{G}_{OK} u_K + \tilde{G}_{OI} \mathring{y}_I + \tilde{G}_{OD} \mathring{y}_D \tag{9}$$

where

$$\tilde{G}_{OI} = G_{OI} G_{II}^{-1}, \quad \tilde{G}_{OK} = G_{OK} - \tilde{G}_{OI} G_{IK} \tag{10}$$

and

$$\tilde{G}_{OD} = \left[G_{OD} - \tilde{G}_{OI} G_{ID} \right] G_{DD}^{-1} \tag{11}$$

In reality, the noise free outputs are not available and instead the noise-corrupted outputs need to be used. Equation (6) can be substituted into (9) giving the indirect model

$$\begin{aligned} y_O &= \tilde{G}_{OK} u_K + \tilde{G}_{OI} y_I + \tilde{G}_{OD} y_D \\ &\quad - \tilde{G}_{OI}(H_I \tau + e_I) - \tilde{G}_{OD} e_D + H_O \tau + e_O \\ &= \tilde{G}_{OK} u_K + \tilde{G}_{OI} y_I + \tilde{G}_{OD} y_D + \bar{\tau} \\ &= \tilde{G}_O \tilde{u} + \bar{\tau} \end{aligned} \tag{12}$$

where $\tilde{G}_O = \begin{bmatrix}\tilde{G}_{OK} & \tilde{G}_{OI} & \tilde{G}_{OD}\end{bmatrix}$, $\tilde{u} = \begin{bmatrix}u_K & y_I & y_D\end{bmatrix}^T$ and $\bar{\tau}$ is a vector of aggregated disturbances. Note that in this framework, there is some flexibility in the choice of the outputs since both y_O and y_I in (5) can act as the output in (12) (if G_{OI}, i.e. the new G_{II}, has the correct properties). Which measurements that are suitable to be selected as outputs depends on the application and many choices can lead to satisfactory results.

Note that even if the original problem (1) is open loop, i.e. that u is uncorrelated with τ and e_O, the indirect model might have correlation between $\bar{\tau}$ and $\tilde{u}$ since the direct and indirect input measurements are used as inputs. The indirect model formulation can be seen as having artificial loops and is in this sense similar to a closed-loop system identification problem. Hence, even if the original system is open-loop, it is sensible to treat the estimation using (12) as a closed-loop problem. Furthermore, in contrast to a system with an actual controller in the loop, the indirect model might have a direct term in this artificial loop and some closed-loop identification methods might fail if this is not considered (Linder et al., 2014; Dankers, 2014).

2.2 Connection to existing system identification problems

The indirect model (12) can in some sense be seen as a generalization of a number of classical problems considered in system identification literature and the connection to these are discussed below.

The most straight-forward cases are when the input is uncorrelated with the external disturbance, i.e. $F_\tau = 0$. The transfer function matrix F_δ can in these cases, for example, be seen as either some actuator dynamics or can be set to identity if the inputs are controlled directly. If all inputs are known, i.e. $n_K = n_u$, the classical open-loop case is recovered (Ljung, 1999) and the classic open-loop errors-in-variable case is recovered if $G_{DD} = 1$ and $n_D = n_u$ (Söderström, 2007).

The positive feedback closed-loop cases are recovered if $F_\delta = S$ and $F_\tau = FSH_O$ where F is the system in the feedback and $S = (1 - FG_O)^{-1}$ is the sensitivity function. In the same way as the open-loop cases, the classic closed-loop case is recovered if $n_K = n_u$ (Forssell, 1999) and the closed-loop errors-in-variable case is recovered if $G_{DD} = 1$ and $n_D = n_u$ (Söderström et al., 2013).

In some system identification problems, the input to the system is unknown. Blind, or sensor-only, system identification relies on knowledge of properties of the input signal. An alternative is the sensor-to-sensor system identification (S2SID) framework that is recovered if $F_\tau \tau = 0$ and $n_u = n_I$. In this case. the resulting $\tilde{G}_{OI}$ would correspond to the pseudo transfer function in the S2SID framework (D'Amato et al., 2009).

Another related and interesting model family concerns dynamic networks. In the framework of Dankers (2014), a dynamic network has a number of internal variables that are dynamically related to each other and a set of these variables is assumed to be measurable. There are also two types of external signals assumed to affect the network, firstly, variables that are known and can be manipulated by the user called external variables and, secondly, unmeasured disturbances, see Dankers (2014) for more details.

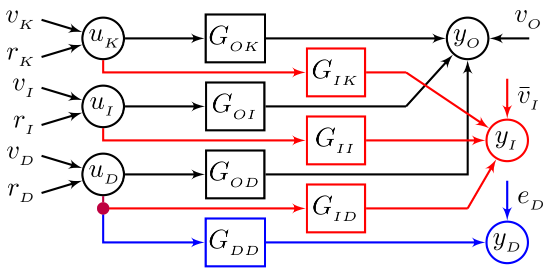

Figure 2: The model (5) seen as a dynamic network where circle corresponds to sums and blocks corresponds to the transfer functions. The black parts corresponds to the model of interest, red corresponds the indirect input measurements and blue corresponds to the direct input measurements.

By introducing the notation

$$\begin{bmatrix} r_K \\ r_I \\ r_D \end{bmatrix} = F_\delta(\mathsf{p})\delta, \quad \begin{bmatrix} v_K \\ v_I \\ v_D \end{bmatrix} = F_\tau(\mathsf{p})\tau \text{ and } \begin{bmatrix} v_O \\ \bar{v}_I \end{bmatrix} = \begin{bmatrix} H_O(\mathsf{p}) \\ H_I(\mathsf{p}) \end{bmatrix} \tau + \begin{bmatrix} e_O \\ e_I \end{bmatrix}, \tag{13}$$

then according to the framework of Dankers (2014), the model (5) can be written as the network model

$$\begin{bmatrix} y_O \\ y_I \\ y_D \\ \hline u_K \\ u_I \\ u_D \end{bmatrix} = \left[\begin{array}{ccc:c:cc} 0 & 0 & 0 & G_{OK} & G_{OI} & G_{OD} \\ 0 & 0 & 0 & G_{IK} & G_{II} & G_{ID} \\ 0 & 0 & 0 & 0 & 0 & G_{DD} \\ \hdashline 0 & 0 & 0 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] \begin{bmatrix} y_O \\ y_I \\ y_D \\ \hline u_K \\ u_I \\ u_D \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \hline r_K \\ r_I \\ r_D \end{bmatrix} + \begin{bmatrix} v_O \\ \bar{v}_I \\ 0 \\ \hline v_K \\ v_I \\ v_D \end{bmatrix} \tag{14}$$

which can also be seen in Figure 2. Since there are no direct paths from y_I and y_D to y_O, they can not be used directly to estimate G_O in this framework (Dankers, 2014). In some sense, the transformation from (5) to the indirect model (12) create artificial paths which makes it possible to utilize these measurements. The direct and indirect input measurements can with this point of view be seen as an generalization of the additive measurement error of the framework in Dankers (2014).

2.3 Examples

In this section two examples of parametric models with indirect input measurements will be presented. The first example will later be used in the experimental verification in Section 4. Here, the models are only summarized and for a complete derivation and deeper discussion, see Linder (2014).

2.3.1 A modified inverted pendulum

Figure 3 shows a modified inverted pendulum hinged on a cart. The mass of the pendulum is split into a nominal mass M and an additional load with mass m. The nominal mass has its center of gravity (CG) a distance z_g from the center of rotation (CR) and with the inertia I_x around its CG. The additional mass has the inertia $I_{x,m}$ and is positioned a distance z_m from the CR. There are two torsional torques acting on the pendulum, one linearly dependent on the angle ϕ and one linearly dependent on the angular velocity $\dot{\phi}$ corresponding to damping. There is an external unknown force acting on the cart which results in the acceleration a_y. Finally, a torque disturbance τ is assumed to act on the CR. A model of the system is given by

$$\ddot{\phi} = -\overbrace{\frac{k + Mgz_g + mgz_m}{I_1}\phi}^{\text{restoring torque}} \overbrace{-\frac{d}{I_1}\dot{\phi}}^{\text{damping}} + \overbrace{\frac{Mz_g + mz_m}{I_1}a_y + \frac{\tau}{I_1}}^{\text{disturbances}} \tag{15}$$

where $I_1 = I_x + Mz_g^2 + I_{x,m} + mz_m^2$. The transfer function from a_y to the angle ϕ is given by

$$G_1(\mathsf{p}) = \frac{b_0}{\mathsf{p}^2 + a_1\mathsf{p} + a_2} \tag{16}$$

where

$$b_0 = \frac{Mz_g + mz_m}{I_1}, a_1 = \frac{d}{I_1}, a_2 = \frac{k + Mgz_g + mgz_m}{I_1} \tag{17}$$

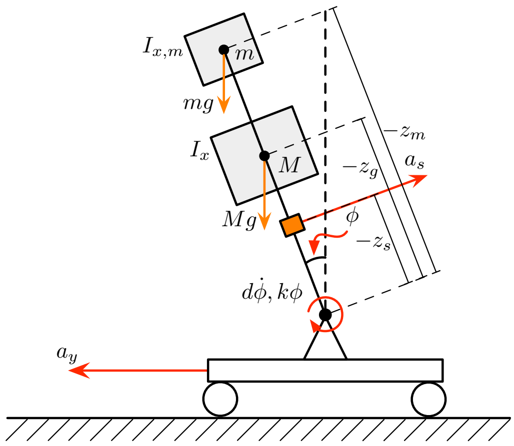

Figure 3: A sketch of the approximate inverted pendulum model showing the relation of the important quantities.

There is an inertial measurement unit (IMU) mounted on the cart a distance z_s from the CR and it is assumed to measure

$$\begin{aligned} y_1 &= \dot{\phi} + e_1 \\ y_2 &= a_s + e_2 = z_s\ddot{\phi} + g\phi - a_y + e_2 \end{aligned} \tag{18}$$

where a_s is the tangential acceleration (assuming small angles) and $e_i, i = 1, 2$ are measurement noise. In this example, there are no known or directly measured inputs and only the signals $u_I = a_y$, $y_I = y_2$ and $y_O = y_1$ are used. This gives

$$G_{OI}(\mathsf{p}) = G_1(\mathsf{p})\mathsf{p}, \quad G_{II}(\mathsf{p}) = G_1(\mathsf{p})(z_s\mathsf{p}^2 + g) - 1, \tag{19}$$

$$H_O(\mathsf{p}) = \frac{1}{b_0}G_{OI}(\mathsf{p}) \text{ and } H_I(\mathsf{p}) = H_O(\mathsf{p})(z_s\mathsf{p}^2 + g) \tag{20}$$

which results in

$$\tilde{G}_{OI}(\mathsf{p}) = G_{OI}(\mathsf{p})G_{II}^{-1}(\mathsf{p}) = \frac{\beta_0\mathsf{p}}{\mathsf{p}^2 + \alpha_1\mathsf{p} + \alpha_2} \tag{21}$$

where

$$\beta_0 = -\frac{Mz_g + mz_m}{I_2}, \quad \alpha_1 = \frac{d}{I_2}, \quad \alpha_2 = \frac{k}{I_2} \tag{22}$$

and $I_2 = I_x + Mz_g(z_g - z_s) + I_{x,m} + mz_m(z_m - z_s)$.

2.3.2 Ship's roll dynamics

Figure 4 shows a ship performing a turning maneuver. Due to its rudder angle δ, the ship's CR is affected by a force resulting in the acceleration a_y. There are a multitude of models for a ship's roll dynamics. Here, a nonlinear model developed and discussed in Blanke and Christensen (1993) and Perez (2005), and used in Linder (2014), is considered.

After linearization and a few manipulations, the model

$$\begin{aligned} A_1\ddot{\phi} = &-d\dot{\phi} - (k + Mgz_g + mgz_m)\phi \\ &+ (K_{\dot{v}} + Mz_g + mz_m)\dot{v} + (K_{ur} + Mz_g + mz_m)Ur \\ &+ K_\delta\delta + \tau \end{aligned} \tag{23}$$

is acquired, where $A_1 = A_x + Mz_g^2 + mz_m^2$, see Linder (2014) for a derivation. The input is considered to be $u = [\delta,\ \dot{v},\ r]^T$ where $\dot{v}$ is the sway acceleration and r is the angular velocity around the yaw axis.

The transfer function matrix from u to ϕ is given by

$$G_2(\mathsf{p}) = \begin{bmatrix} b_1 & b_2 & b_3 \end{bmatrix} \frac{1}{\mathsf{p}^2 + a_3\mathsf{p} + a_4} \tag{24}$$

where

$$\begin{aligned} b_1 &= \frac{K_\delta}{A_1}, b_2 = \frac{K_{\dot{v}} + Mz_g + mz_m}{A_1}, b_3 = \frac{K_{ur} + Mz_g + mz_m}{A_1}U, \\ a_3 &= \frac{d}{A_1} \text{ and } a_4 = \frac{k + Mgz_g + mgz_m}{A_1} \end{aligned} \tag{25}$$

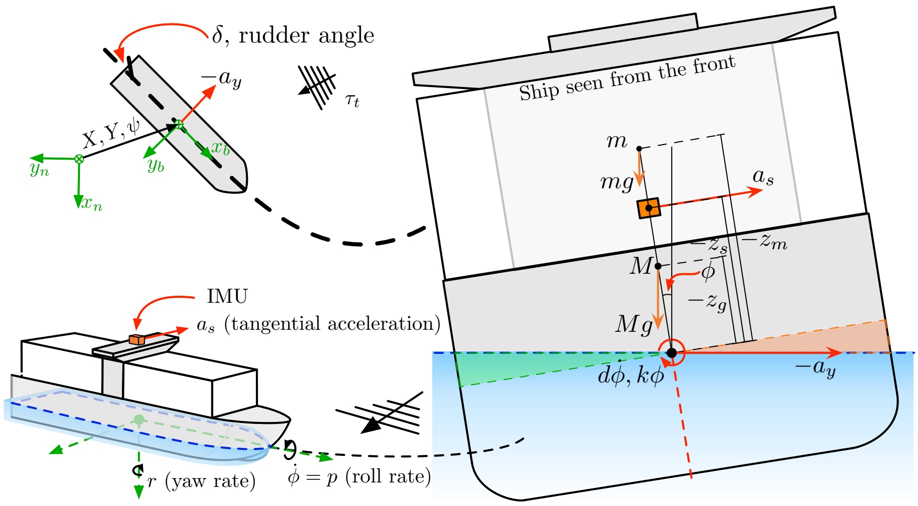

Figure 4: A sketch of a ship performing a turning maneuver which results in an unknown acceleration a_y acting on the CR. The shift in displaced water (green/red area) forces the ship back towards its equilibrium.

The motion of the ship is measured using an IMU. The measurements available are

$$
\begin{aligned}
y_1 &= \dot{\phi} + e_1 \\
y_2 &= a_s + e_2 = z_s\ddot{\phi} + g\phi \underbrace{-\dot{v} - Ur}_{-a_y} + e_2 \\
y_3 &= -r + e_3
\end{aligned}
\tag{26}
$$

where $\dot{\phi}$ is the angular velocity around the roll axis, a_s is the tangential acceleration and $e_i,\ i = 1, 2, 3$ are measurement noise. In this example, the signals are $u_{\scriptscriptstyle K} = \delta$, $u_{\scriptscriptstyle I} = \dot{v}$, $u_{\scriptscriptstyle D} = r$, $y_{\scriptscriptstyle O} = y_1$, $y_{\scriptscriptstyle I} = y_2$ and $y_{\scriptscriptstyle D} = y_3$, which gives

$$
\begin{aligned}
G_{\scriptscriptstyle O}(\mathsf{p}) &= G_2(\mathsf{p})\mathsf{p}, \\
G_{\scriptscriptstyle I}(\mathsf{p}) &= G_2(\mathsf{p})(z_s\mathsf{p}^2 + g) + \begin{bmatrix} 0 & -1 & -U \end{bmatrix} \text{ and} \\
G_{\scriptscriptstyle DD}(\mathsf{p}) &= -1,
\end{aligned}
\tag{27}
$$

that results in

$$
\tilde{G}(\mathsf{p}) = \begin{bmatrix} \beta_1 & \beta_2 & \beta_3 \end{bmatrix} \frac{1}{\mathsf{p}^2 + \alpha_3\mathsf{p} + \alpha_4} \tag{28}
$$

where

$$
\begin{aligned}
&\beta_1 = \frac{K_\delta}{A_2}, \beta_2 = -\frac{K_{\dot{v}} + Mz_g + mz_m}{A_2}, \beta_3 = \frac{K_r}{A_2}, \\
&\alpha_3 = \frac{d}{A_2}, \alpha_4 = \frac{k - K_{\dot{v}}g}{A_2},
\end{aligned}
\tag{29}
$$

$A_2 = A_x + Mz_g(z_g - z_s) + mz_m(z_m - z_s) - K_{\dot{v}}z_s$ and where $H_{\scriptscriptstyle O}$ and $H_{\scriptscriptstyle I}$ are nonzero but quite complex due to the coupling in the original model.

2.4 Identifiability of indirect model

The subject of unique estimation of the parameters in a chosen model structure is both connected to the model structure considered and the informativity of the data set used for identification (Bellman and Åström, 1970; Ljung and Glad, 1994; Bazanella et al., 2010). Here, (1) is considered to be a linear system, but note that the model structure might very well be nonlinearly parameterized. For an arbitrary nonlinear parameterization of (1), it is difficult to give an explicit statement of identifiability for the resulting model of the indirect model (12). There are however a few points worth stressing.

It is *not* necessary that the chosen model structure of (1) is structurally identifiable with respect to the parameters since the model of the indirect measurements might introduce more information. Consider, for instance, the model

$$\left.\begin{aligned} y_{O} &= b_1 u_{K} + b_2^2 u_{I} \\ y_{I} &= b_2 u_{I} \end{aligned}\right\} \Longrightarrow y_{O} = b_1 u_{K} + b_2 y_{I} \tag{30}$$

where the original model is not identifiable, due to the square, but the indirect model is. If the chosen model structure of (1) is structurally identifiable with respect to the parameters this does not necessarily imply identifiability of the resulting model structure of the indirect model (12) with respect to the parameters since cancellations might occur. Consider, for example, the model

$$\left.\begin{aligned} y_{O} &= b_1 u_{K} + b_2 u_{I} \\ y_{I} &= b_2 u_{I} \end{aligned}\right\} \Longrightarrow y_{O} = b_1 u_{K} + y_{I} \tag{31}$$

where the original model is identifiable but where b_2 is eliminated in the indirect model, which is thus not identifiable in the original parameters. Hence, the resulting structure of the indirect model (12) has to be identifiable with respect to the parameters and this has to be checked on a case to case basis.

Secondly, it is not sufficient that the input $\tilde{u}$ of the indirect model is informative since it can be correlated with τ and the setup is thus similar to identification in closed-loop. In the general case, i.e. when the input u is correlated with τ, a sufficient condition is that the external user-controlled signal $F_\delta \delta$ is informative enough. However, this is not a necessary condition since the required informativity of $F_\delta \delta$ depends on the complexity of $F_\tau(\mathsf{p})$, i.e. the feedback structure, see Bazanella et al. (2010).

3 Estimation of the indirect model

There are several methods to estimate the indirect model (12) but here we will focus on an instrumental variable (IV) method. The method is appealing due to low requirements on prior knowledge about the disturbances. The method is based on correlation between the signals in the estimation problem and the instruments created from the vector of external signals δ and thus handles both process and measurement disturbances.

3.1 Discretization using continuous-time parameters

All measurements are taken at discrete-time instances and a discrete-time model is needed to relate the measurements to the parameters. Several different discretization methods can be used, but in this paper the transfer functions are discretized with the bilinear transform $\mathsf{p} = \frac{2}{T}\frac{\mathsf{q}-1}{\mathsf{q}+1}$ where T is the sample period and q is the shift operator. The benefits of the bilinear transform are the simple algebraic relation between p and q while still being fairly accurate for short sampling periods and that the resulting discrete-time model is stable for all sampling periods. Given that a parameterization with the parameters ϑ has been chosen in continuous-time and that the transfer function is finite-dimensional, the discrete time matrix fraction description can be written

$$\tilde{G}_{O,d}(\mathsf{q},\vartheta) = \tilde{G}_{O}\left(\tfrac{2}{T}\tfrac{\mathsf{q}-1}{\mathsf{q}+1},\vartheta\right) = \tilde{A}^{-1}_{O,d}(\mathsf{q},\vartheta)\tilde{B}_{O,d}(\mathsf{q},\vartheta) \tag{32}$$

where the matrices are normalized such that

$$\begin{aligned}\tilde{A}_{O,d}(\mathsf{q},\vartheta) &= I + A_1(\vartheta)\mathsf{q}^{-1} + \ldots + A_{n_a}(\vartheta)\mathsf{q}^{-n_a}\\ \tilde{B}_{O,d}(\mathsf{q},\vartheta) &= B_0(\vartheta)\mathsf{q}^{-n_k} + \ldots + B_{n_b-1}(\vartheta)\mathsf{q}^{-n_b-n_k+1}\end{aligned} \tag{33}$$

which results in the difference equation

$$\begin{aligned}y_{O,t} = &- A_1(\vartheta)y_{O,t-1} - \ldots - A_{n_a}(\vartheta)y_{O,t-n_a}\\ &+ B_0(\vartheta)\tilde{u}_{t-n_k} + \ldots + B_{n_b-1}(\vartheta)\tilde{u}_{t-n_b-n_k+1} + \tilde{\tau}\end{aligned} \tag{34}$$

Assuming that the i^{th} component can be written

$$y^{(i)}_{O,t} = \varphi^T_{i,t}\theta_i(\vartheta) + \tilde{\tau}_{i,t} \tag{35}$$

then (34) can be cast as the linear regression

$$y_{O,t} = \boldsymbol{\varphi}^T_t\boldsymbol{\theta}(\vartheta) + \tilde{\tau} \tag{36}$$

where

$$\boldsymbol{\theta}(\vartheta) = \left[\theta^T_1(\vartheta)\ldots\theta^T_{n_y}(\vartheta)\right]^T \text{ and } \boldsymbol{\varphi}_t = \text{blkdiag}\left[\varphi_{1,t}\ldots\varphi_{n_y,t}\right] \tag{37}$$

Note that $\tilde{\tau}$ is a colored noise process even if $\bar{\tau}$ was white.

3.2 Identifiability issues and multiple datasets

In some situations it is possible to circumvent identifiability issues by introducing more information. A model can be made identifiable by introducing *a priori* knowledge or by introducing more measurements. For instance, consider the unidentifiable model

$$y_O^{(1)} = (a_1 + b_1)u_K \tag{38}$$

where both a_1 and b_1 are unknown. By introducing the second measurement

$$y_O^{(2)} = a_1 u_K \tag{39}$$

the extended model

$$y_O = \begin{bmatrix} y_O^{(1)} \\ y_O^{(2)} \end{bmatrix} = \begin{bmatrix} a_1 + b_1 \\ a_1 \end{bmatrix} u_K \tag{40}$$

becomes identifiable. If more measurements are not possible in the same experiment, measurements from several consecutive experiments can be used in a similar way if the parameters change in a known fashion between the experiments. If the measurements from (38) and (39) were collected from different experiments, the extended model becomes

$$y_O = \begin{bmatrix} y_O^{(1)} \\ y_O^{(2)} \end{bmatrix} = \begin{bmatrix} a_1 + b_1 & 0 \\ 0 & a_1 \end{bmatrix} \begin{bmatrix} u_K^{(1)} \\ u_K^{(2)} \end{bmatrix} \tag{41}$$

Hence, if we have several measurements which are related through the parameters, then (5) can be extended and cast into (36) and thus overcome certain identifiability issues.

3.3 Extended instrumental variable method

There are two terms contributing to the output of (36), one containing information about the interesting input-output relation and the second containing a contribution from disturbances. An instrumental variable method uses the instrument vector ζ_t to extract the interesting information from the data. In principle, the interesting information is estimated by requiring that

$$\frac{1}{N}\sum_{t=1}^{N} \zeta_t (y_{O,t} - \boldsymbol{\varphi}_t^T \boldsymbol{\theta}(\vartheta)) = 0 \tag{42}$$

i.e. that the sample covariance between ζ_t and the prediction error should be zero. A good instrument should be correlated with the regression vector $\boldsymbol{\varphi}_t$ but be uncorrelated with the disturbance $\tilde{\tau}_t$. This idea is generalized in the extended IV method, where the parameters are found by computing

$$\hat{\vartheta} = \operatorname*{argmin}_{\vartheta} \|Y_N - \Phi_N \boldsymbol{\theta}(\vartheta)\|_Q^2 \tag{43}$$

where $\|x\|_Q^2 = x^T Q x$, $Q \succeq 0$ is a weighting matrix,

$$\Phi_N = \frac{1}{N}[\zeta_1 \ \dots \ \zeta_N] \begin{bmatrix} \bar{\boldsymbol{\varphi}}_1^T \\ \vdots \\ \bar{\boldsymbol{\varphi}}_N^T \end{bmatrix}, \ Y_N = \frac{1}{N}[\zeta_1 \ \dots \ \zeta_N] \begin{bmatrix} \bar{y}_{O,1} \\ \vdots \\ \bar{y}_{O,N} \end{bmatrix}, \tag{44}$$

$\bar{y}_{O,t} = L(\mathsf{q})y_{O,t}$, $\bar{\boldsymbol{\varphi}}_t^T = L(\mathsf{q})\boldsymbol{\varphi}_t^T$ and $L(\mathsf{q})$ is a vector of stable prefilters. See, for instance, Söderström and Stoica (1989) or Ljung (1999) for more details.

The estimator (43) will be consistent under fairly general conditions and the real challenge is to choose the vector of instruments ζ_t and the vector of prefilters $L(\mathsf{q})$ such that the variance of the predictor is minimized. The covariance properties of the closed-loop IV methods have been investigated in, for instance, Forssell and Chou (1998), Gilson and Van den Hof (2005) and Gilson et al. (2011).

It is shown in (Gilson et al., 2011) that the optimal prefilter is given by the true inverse noise model and that the optimal instruments are given by the filtered noise-free regression vector. However, the optimal instruments are not implementable in reality since knowledge of the true system and perfect measurements would be required. Instead, a variant of the iterative approach presented in Gilson et al. (2011) is used. In this iterative approach, the current estimate of the parameters are used to create the instrument vector and prefilter which are then used to get a new estimate of the parameters, and the iterations are terminated when the parameters have converged. Here, the instrument vector is given by

$$\zeta_t = L(\mathsf{q}, \eta) \left[-\hat{y}_{O,t-1}^T \dots - \hat{y}_{O,t-\hat{n}_y}^T \; \hat{\mu}_t^T \dots \hat{\mu}_{t-\hat{n}_u+1}^T \right]^T \tag{45}$$

where $L(\mathsf{q}, \eta)$ is a prefilter, $\hat{F}_\delta$ is an estimate of F_δ,

$$\begin{bmatrix} \hat{y}_{O,t} \\ \hline \hat{\hat{u}}_t \end{bmatrix} = \begin{bmatrix} \hat{y}_{O,t} \\ \hline \hat{u}_{K,t} \\ \hat{y}_{I,t} \\ \hat{y}_{D,t} \end{bmatrix} = \begin{bmatrix} G_{OK}(\hat{\vartheta}) & G_{OI}(\hat{\vartheta}) & G_{OD}(\hat{\vartheta}) \\ \hline 1 & 0 & 0 \\ G_{IK}(\hat{\vartheta}) & G_{II}(\hat{\vartheta}) & G_{ID}(\hat{\vartheta}) \\ 0 & 0 & G_{DD}(\hat{\vartheta}) \end{bmatrix} \hat{F}_\delta \delta_t \tag{46}$$

are the simulated output and input, $\hat{\mu}_t$ is a linear combination of delayed inputs $\hat{\hat{u}}_t$, for instance, $\hat{\mu}_t = \hat{\hat{u}}_t - \hat{\hat{u}}_{t-2}$, and $\hat{\vartheta}$ is the current estimate of the parameters. The constants $\hat{n}_y$ and $\hat{n}_u$ are the number of time lags included in ζ_t for the output and the input, respectively, i.e. $\hat{n}_y = 0$ means that $\hat{y}_{O,t}$ is not included in ζ_t. The prefilter is given by

$$L(\mathsf{q}, \eta) = \hat{H}^{-1}(\mathsf{q}, \eta) \tag{47}$$

where $\hat{H}(\mathsf{q}, \hat{\eta})$ is an ARMA model estimated from the residual $\varepsilon_t = y_{O,t} - \boldsymbol{\varphi}_t^T \boldsymbol{\theta}(\hat{\vartheta})$. Note that the simulated signals in (46) are only used as instruments and thus, that an incorrect estimate of F_δ only affects the second-order properties of the estimator (Söderström and Stoica, 1989).

4 Experimental verification

To show the applicability of the suggested method, the modified inverted pendulum example was considered. Data, previously used in Linder et al. (2014), collected from the pendulum seen in Figure 5 was used in the experiment. The known physical quantities of the system, measured using a scale and caliper, and calculated from the basic measurements, are listed in on the first row of Table 1. Note that only the major components were considered when calculating the inertia and that screws and cables were not included.

Two IMUs were used, one mounted on the pendulum and one mounted on the cart. The angular velocity was measured using the x-axis gyro and the tangential acceleration was measured using the y-axis acceleration measurement of the IMU mounted on the pendulum. The y-axis acceleration measurement from the IMU on the cart was used as the reference signal δ which means that δ is actually measured with noise. However, this should not have an impact on the consistency since this measurement noise is uncorrelated with $\tilde{\tau}$.

The goals were to estimate the change in mass m and the change in center of mass z_m, but the inertia I_x, the damping coefficient d and the spring coefficient k were also considered to be unknown. Due to identifiability issues, two datasets were collected, one with the nominal mass and the other with the additional mass, and the model was extended according to Section 3.2. Both datasets were created by dragging the cart back and forth, are 10 minutes long and were down-sampled to 50Hz with a total of 30 000 samples in each dataset.

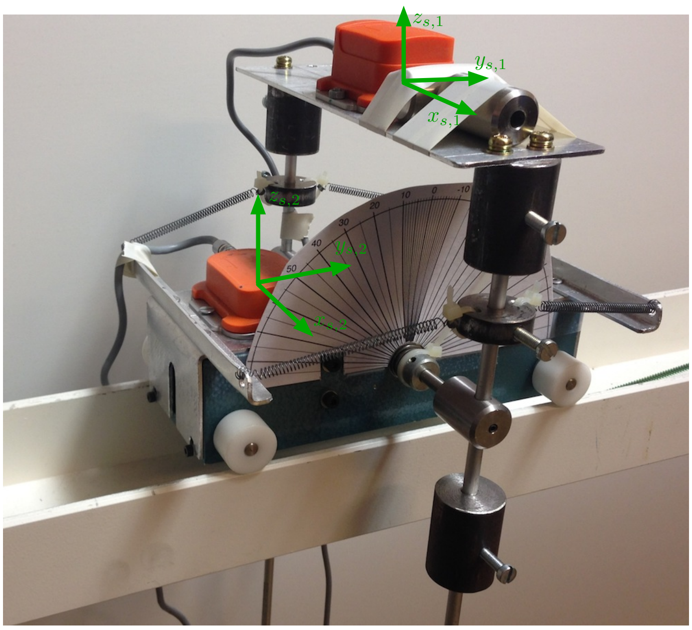

Figure 5: A picture of the modified inverted pendulum. Uppermost in the picture is the IMU (orange) measuring the rolling motion. In front of it is the additional load fastened with white tape. To the left on the cart is the reference IMU.

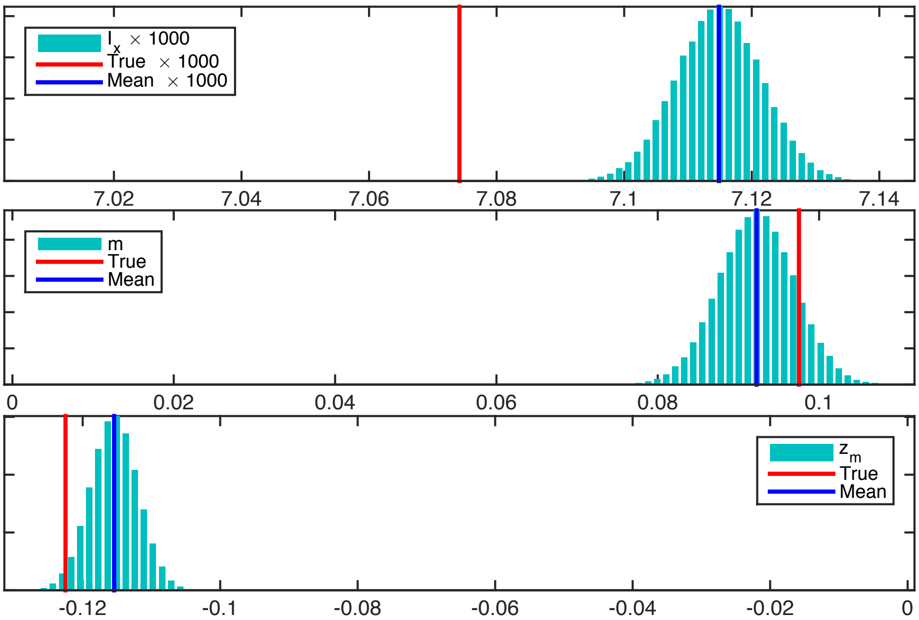

Figure 6: The result from the estimation of the parameters in Table 1. Here, only the parameters with known true values are shown. Red corresponds to the measured or calculated physical quantity in Table 1. Cyan is the histogram of the samples drawn from $\mathcal{N}(\hat{\vartheta}, \hat{P}_\vartheta)$ and blue corresponds to the the sample mean. Top: I_x, middle: m and bottom: z_m.

The noise model

$$\hat{H}_d(\mathsf{q}, \eta) = \frac{1 + \bar{c}_1\mathsf{q}^{-1} + \bar{c}_2\mathsf{q}^{-2} + \bar{c}_3\mathsf{q}^{-3} + \bar{c}_4\mathsf{q}^{-4}}{1 + \bar{\alpha}_1\mathsf{q}^{-1} + \bar{\alpha}_2\mathsf{q}^{-2}} \tag{48}$$

was used. Note that $1 + \bar{\alpha}_1\mathsf{q}^{-1} + \bar{\alpha}_2\mathsf{q}^{-2}$ is the denominator of (21) after discretization and hence, that the corresponding prefilter used in the estimator is

$$L(\mathsf{q}) = \frac{1}{1 + \bar{c}_1\mathsf{q}^{-1} + \bar{c}_2\mathsf{q}^{-2} + \bar{c}_3\mathsf{q}^{-3} + \bar{c}_4\mathsf{q}^{-4}} \tag{49}$$

Note that this noise model was flexible enough to capture the noise model of (12) and also gives some flexibility for coloring in $\bar{\tau}$, see Linder et al. (2014) for a discussion.

The result of the estimation can be seen in Figure 6 and on row two and three in Table 1. Figure 6 has been created by plotting the histogram of 100 000 samples from $\mathcal{N}(\hat{\vartheta}, \hat{P}_\vartheta)$, where $\hat{P}_\vartheta$ is the asymptotic covariance matrix. The cyan bars are for the IV estimator, the blue bars correspond to the mean and the red bars correspond to the true value. It can be seen that the true values of m and z_m are within the three standard deviations of the estimated value. The true value of the inertia I_x is outside the three standard deviations of the estimated value. This is partially due to sampling effects but also note that smaller parts, such as, screws or the cable, were not included when the true value was calculated. Also note that the relative error actually is quite small.

Table 1: The known physical quantities and the results from identification. The first row corresponds to the known values. The second row contains the estimated values and the third row contains the relative errors.

	M	z_g	I_x	d	k	m	z_m	mz_m	mz_m^2
True	1.324	0.0211	0.0071	–	–	0.098	−0.123	−0.0119	0.0015
Est.	–	–	0.0071	0.0032	1.84	0.0922	−0.115	−0.0126	0.0012
Err.	–	–	0.575%	–	–	5.41%	5.77%	10.9%	16.0%

5 Conclusions

In this paper, a framework using indirect input measurements has been proposed for certain system identification problems where the input is partially unknown. In this framework, the indirect model (12), where the unknown input was substituted with direct and indirect input measurements, have been proposed to obtain an estimate of the desired model (1). It has been shown that due to the generality of the framework, several already existing models are recovered as special cases of the indirect model (12) and two examples of physical models have been shown to fit into the framework.

An iterative instrumental variable method has been suggested to estimate the indirect model due to its similarity to closed-loop system identification, i.e. due to correlation between the disturbances and the input and output, and the proposed method has been applied to data from a real inverted pendulum experiment with good results.

In this paper, only the necessary number of signals are used to extend (1) and thus, create the indirect model (12). If additional measurements are available, the available information should probably be used for information fusion, for instance, using an observer, to possibly reduce the resulting variance of the estimator.

Acknowledgments

This work has been supported by the Vinnova Industry Excellence Center LINK-SIC.

References

Alexandre Sanfelice Bazanella, Michel Gevers, and Ljubiša Miškovic. Closed-loop identification of MIMO systems: A new look at identifiability and experiment design. *European Journal of Control*, 16(3):228–239, 2010. ISSN 0947-3580.

R. Bellman and K.J. Åström. On structural identifiability. *Mathematical Biosciences*, 7(3–4):329 – 339, 1970. ISSN 0025-5564. doi: 10.1016/0025-5564(70)90132-X.

Mogens Blanke and Anders C Christensen. Rudder-roll damping autopilot robustness to sway-yaw-roll couplings. In *Proceedings of the 10th Ship Control Systems Symposium*, Ottawa, Canada, 1993.

Anthony M D'Amato, Adam J Brzezinski, Matthew S Holzel, Jun Ni, and Dennis S Bernstein. Sensor-only noncausal blind identification of pseudo transfer functions. In *Proceedings of the 15th IFAC Symposium on System Identification*, pages 1698–1703, Saint-Malo, France, July 2009.

AG Dankers. *System Identification in Dynamic Networks.* PhD thesis, TU Delft, Delft University of Technology, 2014.

Urban Forssell. *Closed-loop identification: Methods, theory, and applications.* PhD thesis, Linköping University, Linköping, Sweden, 1999.

Urban Forssell and C.T. Chou. Efficiency of prediction error and instrumental variable methods for closed-loop identification. In *Proceedings of the 37th IEEE Conference on Decision and Control*, pages 1287–1288 vol.2, Tampa, Florida, USA, 1998.

Marion Gilson and Paul Van den Hof. Instrumental variable methods for closed-loop system identification. *Automatica*, 41(2):241–249, 2005.

Marion Gilson, Hugues Garnier, Peter C Young, and Paul MJ Van den Hof. Optimal instrumental variable method for closed-loop identification. *IET Control Theory & Applications*, 5(10):1147–1154, 2011.

Jonas Linder. *Graybox Modelling of Ships Using Indirect Input Measurements.* Linköping Studies in Science and Technology. Thesis 1681. 2014.

Jonas Linder, Martin Enqvist, and Fredrik Gustafsson. A closed-loop instrumental variable approach to mass and center of mass estimation using IMU data. In *Proceedings of the 53rd IEEE Conference on Decision & Control*, Los Angeles, CA, USA, December 2014.

L. Ljung. *System Identification: Theory for the User (2nd Edition).* Prentice Hall, 1999. ISBN 0136566952.

Lennart Ljung and Torkel Glad. On global identifiability for arbitrary model parametrizations. *Automatica*, 30(2):265 – 276, 1994. ISSN 0005-1098. doi: 10.1016/0005-1098(94)90029-9.

T. Perez. *Ship Motion Control: Course Keeping and Roll Stabilisation Using Rudder and Fins.* Advances in Industrial Control Series. Springer-Verlag London Limited, 2005. ISBN 9781846281570.

Torsten Söderström. Errors-in-variables methods in system identification. *Automatica*, 43(6):939 – 958, 2007. ISSN 0005-1098. doi: http://dx.doi.org/10.1016/j.automatica.2006.11.025.

Torsten Söderström, Liuping Wang, Rik Pintelon, and Johan Schoukens. Can errors-in-variables systems be identified from closed-loop experiments? *Automatica*, 49(2):681–684, 2013.

T.S. Söderström and P.G. Stoica. *System Identification.* Prentice Hall International Series In Systems And Control Engineering. Prentice Hall, 1989. ISBN 9780138812362.

Avdelning, Institution
Division, Department

Division of Automatic Control
Department of Electrical Engineering

Datum
Date

2015-03-13

Språk
Language

☐ Svenska/Swedish
☒ Engelska/English
☐ ______

Rapporttyp
Report category

☐ Licentiatavhandling
☐ Examensarbete
☐ C-uppsats
☐ D-uppsats
☒ Övrig rapport
☐ ______

ISBN
—

ISRN
—

Serietitel och serienummer
Title of series, numbering

ISSN
1400-3902

LiTH-ISY-R-3080

URL för elektronisk version

`http://www.control.isy.liu.se`

Titel
Title

On Indirect Input Measurements

Författare
Author

Jonas Linder, Martin Enqvist

Sammanfattning
Abstract

A common issue with many system identification problems is that the true input to the system is unknown. In this paper, a framework, based on indirect input measurements, is proposed to solve the problem when the input is partially or fully unknown, and cannot be measured directly. The approach relies on measurements that indirectly contain information about the unknown input. The resulting indirect model formulation, with both direct- and indirect input measurements as inputs, can be used to estimate the desired model of the original system. Due to the similarities with closed-loop system identification, an iterative instrumental variable method is proposed to estimate the indirect model. To show the applicability of the proposed method, it is applied to data from an inverted pendulum experiment with good results.

Nyckelord
Keywords

System identification, Model structure, Physical models, Instrumental Variable, Closed-loop