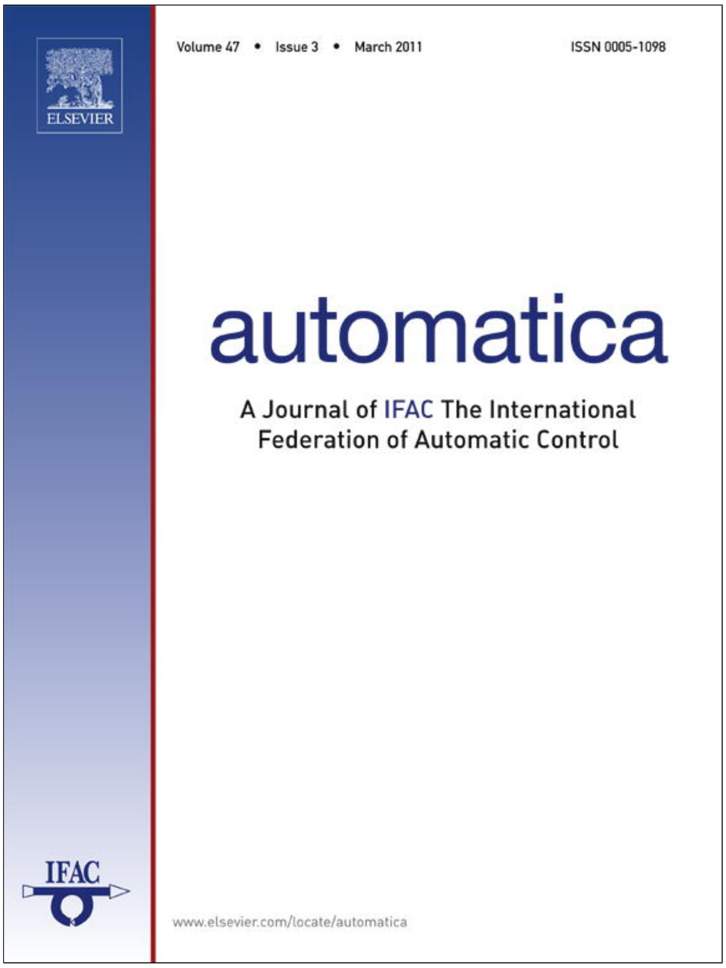

Volume 47 • Issue 3 • March 2011
ISSN 0005-1098
ELSEVIER
automatica
A Journal of IFAC The International Federation of Automatic Control
IFAC
www.elsevier.com/locate/automatica

Automatica 47 (2011) 545–551

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Brief paper

On the identifiability of errors-in-variables models with white measurement errors[☆]

Giulio Bottegal, Giorgio Picci *, Stefano Pinzoni

Department of Information Engineering, Università di Padova, Via Gradenigo 6/A, IT-35131 Padova, Italy

ARTICLE INFO

Article history:
Available online 26 January 2011

Keywords:
Identifiability
Errors-in-variables models
Frisch scheme

ABSTRACT

We discuss identifiability of dynamic SISO errors-in-variables (EIV) models with white measurement errors. Although this class of models turns out to be generically identifiable, it has been pointed out that in certain circumstances there may be *two* EIV models which are indistinguishable from external input–output experiments. This lack of (global) identifiability may be prejudicial to identification and needs better understanding. The identifiability conditions found in the literature guarantee uniqueness under certain coprimality assumptions on the (rational) transfer function of the ideal "true" system and the spectral density of the noiseless "true" input. Unfortunately these conditions are *not testable* since they concern precisely the unknowns of the problem which are not available to the experimenter. We provide new identifiability conditions which are instead expressible in terms of the external description of the observable signals, namely their joint power spectral densities.

1. Introduction

The identification of errors-in-variables (EIV) models is a classical subject which has been studied in the statistical literature since the beginning of the last century and has generated many papers, initially dealing with static EIV models, Frisch (1934), Gini (1921), Kalman (1982), Madansky (1959), and, more recently with dynamic models, see e.g. Anderson (1985), Söderström (1981). The interest is motivated by the fact that these models provide a more realistic description of systems where the input signal may also be affected by noise or by random errors of various kinds. This in contrast to the use of standard ARMAX or Box–Jenkins models, where the input signal is invariably supposed to be measured exactly by the data acquisition device. Yet, a main difficulty with EIV models is that they are generally non-identifiable. This is by now well-known, in particular for dynamic EIV models, and many papers have appeared dealing with identifiability of general dynamic EIV models such as Aguero and Goodwin (2008), Anderson and Deistler (1984), Deistler and Anderson (1989), Picci and Pinzoni (1986), Schachermayer and Deistler (1998), and Scherrer and Deistler (1998). In order to overcome this structural difficulty, dynamic EIV models with *white measurement errors*, often called *Frisch scheme* models, have recently been considered and identification of these models is now a rather active research subject. Although this model class is rather restricted because of the assumption of white measurement errors, it appears to be a natural generalization of ARMAX or Box–Jenkins models of the *output error (OE)* type, where the standard identification techniques may generalize naturally. Indeed the identification of these models has greatly advanced in recent years and seems to have reached the maturity to become a standard tool in applications, see e.g. Beghelli, Guidorzi, and Soverini (1990), Chen and Yang (2005), Fernando and Nicholson (1985), Söderström, Mahata, and Soverini (2003); Söderström, Soverini, and Mahata (2002), Zheng (1999), Zheng and Feng (1992), the recent survey paper Söderström (2007) and Diversi and Guidorzi (2009), Guidorzi and Diversi (2009), Söderström, Mossberg, and Hong (2009).

Although EIV models with white measurement errors turn out to be generically identifiable (where the attribute "generic" can here be given an intuitive meaning of "almost always"), it has been pointed out by Picci, Gei, and Pinzoni (1993) and Stoica and Nehorai (1987) that in certain circumstances there may be *two* EIV models which are indistinguishable from external input–output experiments. This lack of (global) identifiability, although should hopefully almost never be encountered in practice, is a fact which still seems rather obscure. Besides, for theoretical understanding of the phenomenon, we believe that a better comprehension of when two different models can describe the data equally well is needed for practical reasons, for example in order to avoid possible causes of ill-conditioning of identification algorithms. Now the identifiability conditions found in the literature, e.g. in Castaldi and Soverini (1996), Stoica and Nehorai (1987), guarantee uniqueness under certain coprimality assumptions on the (unknown) rational transfer function of the ideal "true" system and the (unknown)

[☆] The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Wolfgang Scherrer under the direction of Editor Torsten Söderström.

* Corresponding author. Tel.: +39 049 8277705; fax: +39 049 8277614.
E-mail addresses: bottegal@dei.unipd.it (G. Bottegal), picci@dei.unipd.it (G. Picci), pinzoni@dei.unipd.it (S. Pinzoni).

0005-1098/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.automatica.2010.12.004

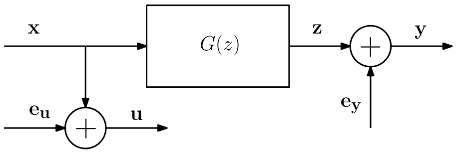


Fig. 1. Scheme of EIV model.

spectral density of the noiseless "true" input. Unfortunately these conditions are not testable since they concern precisely the unknowns of the problem which are not available to the experimenter. Ideally, identifiability conditions should instead be expressible in terms of the external description of the observable signals; namely their joint power spectral densities. Adhering to this point of view, in this paper we provide conditions on the spectral densities of the external (measurable) signals under which a SISO EIV structure with white measurement noises is *non-identifiable*. Our conditions state that a necessary condition for non-identifiability is the existence of a linear affine relation between the spectra of the two external signals. This condition turns out to be "almost" sufficient, modulo the nonlinear constraint of positivity of the variances of the additive noises. We provide some conditions on the parameters of the affine relation under which the condition is actually also sufficient. All of this will be explained in detail in Section 5.

2. Background on dynamic EIV models

In this paper, $t \in \mathbb{Z}$ denotes the discrete time variable and boldface symbols denote random variables or processes.

Consider a pair of real scalar second-order stationary zero-mean discrete-time stochastic processes $(\mathbf{y}, \mathbf{u})$, whose joint spectral density matrix function

$$S(z) = \begin{bmatrix} S_{\mathbf{y}}(z) & S_{\mathbf{yu}}(z) \\ S_{\mathbf{uy}}(z) & S_{\mathbf{u}}(z) \end{bmatrix}, \quad z \in \mathbb{C}, \tag{1}$$

will be assumed positive definite almost everywhere on the unit circle $\{|z| = 1\}$. Recall that by Hermitian symmetry of the spectrum we have $S_{\mathbf{yu}}(z) = S_{\mathbf{uy}}(z^{-1})$.

The background motivation for EIV models is to describe the pair $(\mathbf{y}, \mathbf{u})$ as measurements corrupted by additive noise of two *internal*[1] non observable, stochastic processes denoted $\mathbf{z}$ and $\mathbf{x}$, which are related by a time-invariant linear transfer function $G(z)$, $z \in \mathbb{C}$. We shall make no assumptions on $G(z)$ like causality, stability or other. A EIV model is thus described by the equations (see Fig. 1):

$$\begin{cases} \mathbf{y}(t) = G(z)\mathbf{x}(t) + \mathbf{e}_{\mathbf{y}}(t) \\ \mathbf{u}(t) = \mathbf{x}(t) + \mathbf{e}_{\mathbf{u}}(t), \end{cases} \tag{2}$$

where the processes $\mathbf{e}_{\mathbf{u}}(t)$ and $\mathbf{e}_{\mathbf{y}}(t)$ called *measurement noises* are mutually uncorrelated and uncorrelated also with the process $\mathbf{x}(t)$.

Note that even in the case when $G(z)$ is assumed causal, the causal appearance of (2) is actually misleading. According to the standard notions of causality in the literature Caines (1988), Granger (1963) it is in fact generally not true that $\mathbf{y}(t)$ is *caused* by $\mathbf{u}(t)$, as it is easy to check that for EIV models there is in general feedback from one variable to the other. Nevertheless it is common practice to call $\mathbf{u}$ the *input* and $\mathbf{y}$ the *output* processes.

For brevity, we shall say that an EIV model (2) is a *realization* of the joint spectrum of the $(\mathbf{y}, \mathbf{u})$ processes it represents. This joint spectrum is in a sense an *external* description which is uniquely attached to the $(\mathbf{y}, \mathbf{u})$ processes, while specifying an EIV description requires the introduction of additional non observable variables so that there are in general many EIV realizations of the same spectrum. A basic identifiability question of EIV models that has been studied in the literature and we shall also address in this paper is how many different EIV models can realize the same joint spectrum (1).

As it is well-known (Anderson & Deistler, 1984), any joint spectral density matrix $S(z)$ admits decompositions of the form, $S(z) = \hat{S}(z) + \tilde{S}(z)$ where the spectrum $\hat{S}(z)$ has rank one almost everywhere on the unit circle, and $\tilde{S}(z)$ is a diagonal spectral density. It is then easy to see that by defining the internal variables $\mathbf{x}$, $\mathbf{z}$, $\mathbf{e}_{\mathbf{u}}$, $\mathbf{e}_{\mathbf{y}}$ such that their spectral densities satisfy

$$\hat{S}(z) = \begin{bmatrix} S_{\mathbf{z}}(z) & S_{\mathbf{zx}}(z) \\ S_{\mathbf{xz}}(z) & S_{\mathbf{x}}(z) \end{bmatrix}, \qquad \tilde{S}(z) = \begin{bmatrix} S_{\mathbf{e}_{\mathbf{y}}}(z) & 0 \\ 0 & S_{\mathbf{e}_{\mathbf{u}}}(z) \end{bmatrix}, \tag{3}$$

and $S_{\mathbf{zx}}(z) := S_{\mathbf{yu}}(z)$, one has indeed a representation of the form (2) with

$$S_{\mathbf{z}}(z)S_{\mathbf{x}}(z) = S_{\mathbf{yu}}(z)S_{\mathbf{uy}}(z) \quad \forall z : |z| = 1 \tag{4}$$

and $G(z) := S_{\mathbf{zx}}(z)/S_{\mathbf{x}}(z)$. Hence all joint spectra (1) admit EIV representations.

From now on, we shall only consider EIV models where $\mathbf{e}_{\mathbf{u}}(t)$ and $\mathbf{e}_{\mathbf{y}}(t)$ are white processes; i.e. $S_{\mathbf{e}_{\mathbf{y}}}(z) = \sigma_{\mathbf{y}}^2, S_{\mathbf{e}_{\mathbf{u}}}(z) = \sigma_{\mathbf{u}}^2$ (not depending on z). This is a classical model class, often called the *Frisch scheme* which is discussed in many papers, see e.g the survey Söderström (2007). Naturally the Frisch scheme is a very simple model which will in general provide only an approximation of the real spectra. In particular it postulates that the equation

$$(S_{\mathbf{y}}(z) - S_{\mathbf{e}_{\mathbf{y}}}(z))(S_{\mathbf{u}}(z) - S_{\mathbf{e}_{\mathbf{u}}}(z)) = S_{\mathbf{yu}}(z)S_{\mathbf{uy}}(z), \tag{5}$$

in the unknowns $S_{\mathbf{e}_{\mathbf{y}}}, S_{\mathbf{e}_{\mathbf{u}}}$, should admit constant (positive) solutions, $S_{\mathbf{e}_{\mathbf{y}}} = \sigma_{\mathbf{y}}^2, S_{\mathbf{e}_{\mathbf{u}}} = \sigma_{\mathbf{u}}^2$, which of course will generically not happen.

3. The family of EIV models with white measurement errors (Frisch scheme)

It is well-known that, given the joint spectrum, the family of EIV models realizing it can be parametrized in terms of the two variances $(\sigma_{\mathbf{y}}^2, \sigma_{\mathbf{u}}^2)$, subjected to a *non-negativity* plus a *rank one condition* which we illustrate below. Letting

$$\begin{aligned} R_{\mathbf{y}}(z) &:= S_{\mathbf{y}}(z) - \frac{S_{\mathbf{yu}}(z)S_{\mathbf{uy}}(z)}{S_{\mathbf{u}}(z)} \\ R_{\mathbf{u}}(z) &:= S_{\mathbf{u}}(z) - \frac{S_{\mathbf{uy}}(z)S_{\mathbf{yu}}(z)}{S_{\mathbf{y}}(z)}, \end{aligned} \tag{6}$$

the *non-negativity constraint* (see e.g. Anderson, 1985) is

$$\begin{aligned} 0 &\le \sigma_{\mathbf{y}}^2 \le \bar{R}_{\mathbf{y}} := \min\{R_{\mathbf{y}}(z), z : |z| = 1\} \\ 0 &\le \sigma_{\mathbf{u}}^2 \le \bar{R}_{\mathbf{u}} := \min\{R_{\mathbf{u}}(z), z : |z| = 1\}. \end{aligned} \tag{7}$$

Given $(\sigma_{\mathbf{y}}^2, \sigma_{\mathbf{u}}^2)$, satisfying (7), let $S_{\mathbf{z}}(z) := S_{\mathbf{y}}(z) - \sigma_{\mathbf{y}}^2$ and $S_{\mathbf{x}}(z) := S_{\mathbf{u}}(z) - \sigma_{\mathbf{u}}^2$; then $S_{\mathbf{z}}(z)$ and $S_{\mathbf{x}}(z)$ are bona-fide spectral densities since they certainly satisfy the non-negativity constraints, $S_{\mathbf{y}}(z) - \sigma_{\mathbf{y}}^2 \ge 0$ and $S_{\mathbf{u}}(z) - \sigma_{\mathbf{u}}^2 \ge 0$ on the unit circle $\{z : |z| = 1\}$.

The *rank one constraint* comes from rewriting (4) as

$$(S_{\mathbf{y}}(z) - \sigma_{\mathbf{y}}^2)(S_{\mathbf{u}}(z) - \sigma_{\mathbf{u}}^2) = S_{\mathbf{yu}}(z)S_{\mathbf{uy}}(z) \quad \{z : |z| = 1\}. \tag{8}$$

It follows from a well-known result in the literature (see e.g. Anderson, 1985; Picci & Pinzoni, 1986) that if the noise variances $\sigma_{\mathbf{y}}^2, \sigma_{\mathbf{u}}^2$ satisfy these two constraints then they are valid noise variances of an EIV model realizing the given spectrum.

Since the very definition of an EIV model entails that the cross spectral density of $\mathbf{z}$ and $\mathbf{x}$ must coincide with that of $\mathbf{y}$ and $\mathbf{u}$, we can obtain $G(z)$ from

$$G(z) = \frac{S_{\mathbf{zx}}(z)}{S_{\mathbf{x}}(z)} = \frac{S_{\mathbf{yu}}(z)}{S_{\mathbf{u}}(z) - \sigma_{\mathbf{u}}^2},$$

the reciprocal formula providing the symmetric representation of $\mathbf{x}$ in terms of $\mathbf{z}$. Our problem then reduces to investigating how many pairs $(\sigma_{\mathbf{y}}^2, \sigma_{\mathbf{u}}^2)$ can lead to EIV (Frisch) realizations of a given

[1] These are often called "true" variables in the literature.

joint spectral density. To avoid trivial pathological cases of non uniqueness, from now on we shall assume that $S_{\mathbf{yu}}(z)S_{\mathbf{uy}}(z)$ is not identically zero and that neither $\mathbf{y}$ nor $\mathbf{u}$ are white noise processes. The following result (Picci et al., 1993; Stoica & Nehorai, 1987) lies at the background of our investigations.

Theorem 1. *There are at most two pairs of noise variances* $(\sigma_{\mathbf{y}}^2, \sigma_{\mathbf{u}}^2)$ *which satisfy condition (8). Equivalently, there are at most two EIV models (with white measurement errors) which are compatible with the joint spectrum (1).*

For ease of reference we also recall here the following obvious fact.

Lemma 2. *For every variance pair* $(\sigma_{\mathbf{y}}^2, \sigma_{\mathbf{u}}^2)$ *satisfying the rank one condition (8), one of the two variance values uniquely determines the value of the other.*

4. Conditions for non-identifiability

Conditions under which two EIV models exist both describing the same joint spectrum (non-identifiability), have been described in Stoica and Nehorai (1987). However these conditions are given in terms of unknown signal spectra and transfer functions and are not testable. It is therefore important to characterize this occurrence in terms of the available external spectral data.

An explicit necessary condition for non-identifiability is given in the following theorem.

Theorem 3. *If there are two EIV models realizing the same joint spectrum, then there are constants* $L > 0$ *and* K *such that the following* linear-affine *relation holds*

$$S_{\mathbf{y}}(z) = LS_{\mathbf{u}}(z) + K. \tag{9}$$

Assume the pair $(\sigma_{\mathbf{y}}'^2, \sigma_{\mathbf{u}}'^2)$ *parametrizes an EIV realization with* internal *signal spectra* $S_{\mathbf{z}}(z), S_{\mathbf{x}}(z)$*. Then if there is another model realizing the same joint spectrum, it must have the following structure:*

$$S_{\mathbf{y}}(z) = LS_{\mathbf{x}}(z) + \sigma_{\mathbf{y}}''^2, \qquad S_{\mathbf{u}}(z) = L^{-1}S_{\mathbf{z}}(z) + \sigma_{\mathbf{u}}''^2, \tag{10}$$

so that one model is obtained by switching and renormalizing the internal spectra of the other.

Proof.[2] Assume there are two distinct variance pairs $(\sigma_{\mathbf{y}}'^2, \sigma_{\mathbf{u}}'^2)$ and $(\sigma_{\mathbf{y}}''^2, \sigma_{\mathbf{u}}''^2)$ describing two EIV realizations of the same joint spectrum. From Eq. (5) it must hold that

$$(S_{\mathbf{y}}(z) - \sigma_{\mathbf{y}}'^2)(S_{\mathbf{u}}(z) - \sigma_{\mathbf{u}}'^2) = S_{\mathbf{yu}}(z)S_{\mathbf{uy}}(z)$$
$$(S_{\mathbf{y}}(z) - \sigma_{\mathbf{y}}''^2)(S_{\mathbf{u}}(z) - \sigma_{\mathbf{u}}''^2) = S_{\mathbf{yu}}(z)S_{\mathbf{uy}}(z)$$

are simultaneously true. Subtracting the second equation from the first we obtain

$$(\sigma_{\mathbf{u}}''^2 - \sigma_{\mathbf{u}}'^2)S_{\mathbf{y}}(z) = (\sigma_{\mathbf{y}}'^2 - \sigma_{\mathbf{y}}''^2)S_{\mathbf{u}}(z) + (\sigma_{\mathbf{y}}''^2\sigma_{\mathbf{u}}''^2 - \sigma_{\mathbf{y}}'^2\sigma_{\mathbf{u}}'^2). \tag{11}$$

Being $(\sigma_{\mathbf{u}}''^2 - \sigma_{\mathbf{u}}'^2) \neq 0$ we can rewrite (11) as

$$S_{\mathbf{y}}(z) = \frac{\sigma_{\mathbf{y}}'^2 - \sigma_{\mathbf{y}}''^2}{\sigma_{\mathbf{u}}''^2 - \sigma_{\mathbf{u}}'^2}S_{\mathbf{u}}(z) + \frac{\sigma_{\mathbf{y}}''^2\sigma_{\mathbf{u}}''^2 - \sigma_{\mathbf{y}}'^2\sigma_{\mathbf{u}}'^2}{\sigma_{\mathbf{u}}''^2 - \sigma_{\mathbf{u}}'^2},$$

which, denoting

$$L = \frac{\sigma_{\mathbf{y}}'^2 - \sigma_{\mathbf{y}}''^2}{\sigma_{\mathbf{u}}''^2 - \sigma_{\mathbf{u}}'^2}, \qquad K = \frac{\sigma_{\mathbf{y}}''^2\sigma_{\mathbf{u}}''^2 - \sigma_{\mathbf{y}}'^2\sigma_{\mathbf{u}}'^2}{\sigma_{\mathbf{u}}''^2 - \sigma_{\mathbf{u}}'^2},$$

leads to (9). From (11) we can also obtain

$$S_{\mathbf{y}}(z) = \frac{\sigma_{\mathbf{y}}'^2 - \sigma_{\mathbf{y}}''^2}{\sigma_{\mathbf{u}}''^2 - \sigma_{\mathbf{u}}'^2}(S_{\mathbf{u}}(z) - \sigma_{\mathbf{u}}'^2) + \sigma_{\mathbf{y}}''^2 = LS_{\mathbf{x}}(z) + \sigma_{\mathbf{y}}''^2$$

and $S_{\mathbf{u}}(z) = L^{-1}S_{\mathbf{z}}(z) + \sigma_{\mathbf{u}}''^2$. Finally, that L is always positive follows since, as pointed out in Anderson and Deistler (1984), the admissible variance pairs lay on a hyperbola. Hence whenever $\sigma_{\mathbf{y}}''^2 > \sigma_{\mathbf{y}}'^2$, necessarily $\sigma_{\mathbf{u}}''^2 < \sigma_{\mathbf{u}}'^2$, and conversely. One can then see that for any variance pair determining two (non-identifiable) EIV models one has $L > 0$. □

5. Sufficiency of the linear-affine relation

Theorem 3 provides a nice and simple necessary condition for non-identifiability of EIV models. In this section we shall take up the question of assessing when the linear-affine relation (9) is also sufficient for non-identifiability. Naturally, we shall have to *assume that the joint spectrum* (1) *admits EIV* (*Frisch-scheme*) realizations.

5.1. Checking for non-identifiability given a model

Let us consider a joint spectrum satisfying the linear-affine relation (9) admitting an EIV realization with variance pair $(\sigma_{\mathbf{y}}'^2, \sigma_{\mathbf{u}}'^2)$. Defining

$$\Sigma_{\mathbf{y}} := K + L\sigma_{\mathbf{u}}'^2 \tag{12}$$

and substituting $S_{\mathbf{u}}(z) = S_{\mathbf{x}}(z) + \sigma_{\mathbf{u}}'^2$ into (9), one gets a candidate alternative model

$$S_{\mathbf{y}}(z) = LS_{\mathbf{x}}(z) + \Sigma_{\mathbf{y}}, \tag{13}$$

which would prove non-identifiability just in case $\Sigma_{\mathbf{y}}$ turns out to be a valid variance, say $\sigma_{\mathbf{y}}''^2$. A similar argument leads to a candidate companion equation

$$S_{\mathbf{u}}(z) = L^{-1}S_{\mathbf{z}}(z) + \Sigma_{\mathbf{u}}, \quad \Sigma_{\mathbf{u}} := L^{-1}(\sigma_{\mathbf{y}}'^2 - K).$$

Hence the question of proving existence of a second valid EIV model reduces to discussing what range of parameters L and K guarantee that $\Sigma_{\mathbf{y}}$ in (13) is a valid output noise variance, that is, such that the positivity condition (7) is satisfied. It is actually easy to show that if one of the two variances, say $\sigma_{\mathbf{y}}^2$, satisfies the condition $\sigma_{\mathbf{y}}^2 \in [0, \bar{R}_{\mathbf{y}}]$, then the second inequality in (7) is automatically satisfied. For this reason we shall henceforth just concentrate on $\Sigma_{\mathbf{y}}$. We may distinguish three different situations:

- Either $\Sigma_{\mathbf{y}} < 0$ or $\Sigma_{\mathbf{y}} > \bar{R}_{\mathbf{y}}$: in this case $\Sigma_{\mathbf{y}}$ cannot be interpreted as a noise variance and (13) cannot give rise to a second EIV model. The given model is identifiable.
- $\Sigma_{\mathbf{y}} = \sigma_{\mathbf{y}}'^2$: by Lemma 2 it must also hold that $\Sigma_{\mathbf{u}} = \sigma_{\mathbf{u}}'^2$. This is the case in which the two EIV models coincide and $G(z)$ turns out to be all-pass with gain L; see Section 5.2. The model is identifiable.
- $0 \leq \Sigma_{\mathbf{y}} \leq \bar{R}_{\mathbf{y}}$, $\Sigma_{\mathbf{y}} \neq \sigma_{\mathbf{y}}'^2$: in this case $\Sigma_{\mathbf{y}}$ can be interpreted as output noise variance; i.e. $\Sigma_{\mathbf{y}} \equiv \sigma_{\mathbf{y}}''^2$. The decomposition

  $$S_{\mathbf{y}}(z) = LS_{\mathbf{x}}(z) + \Sigma_{\mathbf{y}} \tag{14a}$$

  $$S_{\mathbf{u}}(z) = L^{-1}S_{\mathbf{z}}(z) + \Sigma_{\mathbf{u}} \tag{14b}$$

  is a valid EIV realization of the given joint spectrum which is different from the given one as $\Sigma_{\mathbf{y}} \neq \sigma_{\mathbf{y}}'^2$, $\Sigma_{\mathbf{u}} \neq \sigma_{\mathbf{u}}'^2$. Therefore the model is non-identifiable.

Fig. 2 provides a graphical description of the three situations.

5.2. EIV models with an all-pass transfer function

Consider EIV models with an all-pass transfer function, namely a transfer function satisfying $G(e^{j\theta})G(e^{-j\theta}) = L$. Then the internal spectra must satisfy

$$S_{\mathbf{z}}(z) = G(z)G(z^{-1})S_{\mathbf{x}}(z) = LS_{\mathbf{x}}(z)$$

[2] We thank the associate editor for supplying this shorter and direct proof, which is much simpler than the original.

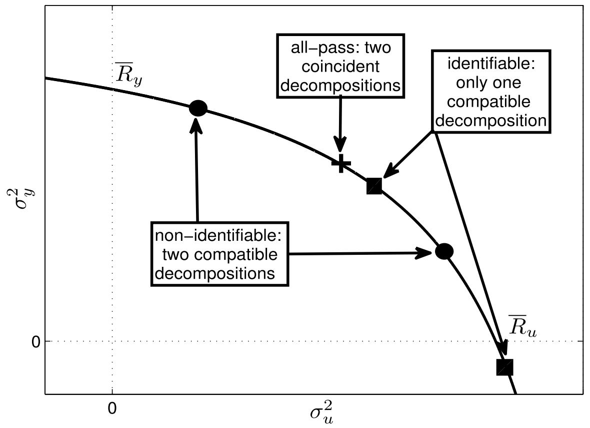


Fig. 2. Graphical interpretation of identifiability of EIV models.

and since $S_{\mathbf{y}}(z) = S_{\mathbf{z}}(z) + \sigma_{\mathbf{y}}^2$, summing $\sigma_{\mathbf{y}}^2$ to both members one gets $S_{\mathbf{y}}(z) = LS_{\mathbf{x}}(z) + \sigma_{\mathbf{y}}^2$. Further recalling that $S_{\mathbf{x}}(z) = S_{\mathbf{u}}(z) - \sigma_{\mathbf{u}}^2$, we arrive at

$$S_{\mathbf{y}}(z) = LS_{\mathbf{u}}(z) - L\sigma_{\mathbf{u}}^2 + \sigma_{\mathbf{y}}^2 = LS_{\mathbf{u}}(z) + K$$

where $K = \sigma_{\mathbf{y}}^2 - L\sigma_{\mathbf{u}}^2$. Hence EIV models with an all-pass transfer function satisfy the linear-affine relation. However it is easy to see that they are identifiable, since the expression of $\sigma_{\mathbf{y}}^2$ is equal to the one of the candidate alternative variances in (12).

These models are however quite special; in a sense they correspond to a limit situation, as explained in the following remark.

Remark 4. As it follows from Eq. (10), Theorem 3, for non-identifiable EIV models, the spectrum $S_{\mathbf{y}}(z)$ can be written $S_{\mathbf{y}}(z) = LS_{\mathbf{x}}(z) + \sigma_{\mathbf{y}}''^2$, but for all-pass transfer functions one also has $S_{\mathbf{z}}(z) = LS_{\mathbf{x}}(z)$ and therefore $\sigma_{\mathbf{y}}''^2 = \sigma_{\mathbf{y}}'^2$. For this reason even if formally there are two EIV realizations with the same all-pass transfer function, the *two realizations actually coincide*.

5.3. Checking non-identifiability from the joint spectrum

When (one of) the variance parameters $\sigma_{\mathbf{y}}^2, \sigma_{\mathbf{u}}^2$ of an EIV model is known, it is trivial to check identifiability by checking whether or not $\Sigma_{\mathbf{y}}$ lies in the feasible interval $[0, \bar{R}_{\mathbf{y}}]$. This test requires however knowledge of the variance parameters which in practice can only be estimated by some identification procedure and are therefore affected by noise. The result of the test may then be also uncertain and identification-method dependent. This is conceptually unsatisfactory, as identifiability should in principle be a property of a *model class* alone and should not depend on the outcome of a particular identification algorithm.

One would like to check identifiability on the basis of the external spectra only, assuming of course that these spectra are realizable by an EIV model. Assuming the linear affine relation (9) is satisfied, one would in particular like to check *a priori* if $\Sigma_{\mathbf{y}}$ belongs to the feasible interval just on the basis of the parameters L and K. Note that these are external parameters which can in practice be computed very accurately by running a linear regression in the frequency domain. See e.g. the simulation example in Section 7. Unfortunately there is not enough information about the joint spectrum in these parameters to provide a definite answer since $\Sigma_{\mathbf{y}}$ is a function also of the unknown value $\sigma_{\mathbf{u}}^2$. We can nevertheless obtain some loose sufficient conditions which may turn out useful in certain situations.

The following proposition, whose proof is skipped for reasons of space limitations, provides an instance of non-identifiability criteria based on L and K alone.

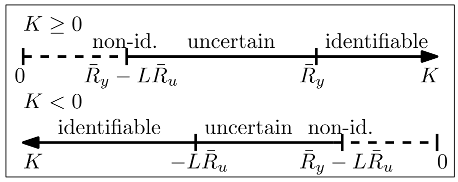


Fig. 3. Identifiability for various values of K.

Proposition 5. *Assume the joint spectrum (1) admits EIV realizations and that the linear-affine relation (9) is satisfied. Assume also that there is no all-pass relation between the internal processes. Then,*

- *if $K > \bar{R}_{\mathbf{y}}$ or $K < -L\bar{R}_{\mathbf{u}}$ we have identifiability;*
- *if $K = \bar{R}_{\mathbf{y}} - L\bar{R}_{\mathbf{u}}$ we have non-identifiability.*

The situation is described by the Fig. 3. We have excluded the presence of zeros on the unit circle of either $S_{\mathbf{y}}(z)$ or $S_{\mathbf{u}}(z)$. The presence of such zeros would in fact imply either that $\sigma_{\mathbf{y}}^2 = 0$ (i.e. $S_{\mathbf{z}}(z) = S_{\mathbf{y}}(z)$) or $\sigma_{\mathbf{u}}^2 = 0$ (and $S_{\mathbf{x}}(z) = S_{\mathbf{u}}(z)$). In both cases identifiability analysis would be superfluous.

Use of these conditions is illustrated in the examples of the next section.

6. Examples

To illustrate the results of this paper we shall first discuss two academic examples.

6.1. A non-identifiable model

Consider the following power spectra:

$$S_{\mathbf{y}}(z) = \frac{0.11z^2 - 4.864z - 14.57 - 4.864z^{-1} + 0.11z^{-2}}{z^2 - 0.138z - 2.83 - 0.138z^{-1} + z^{-2}}$$

$$S_{\mathbf{u}}(z) = \frac{0.01z^2 - 0.971z - 2.88 - 0.971z^{-1} + 0.01z^{-2}}{z^2 - 0.138z - 2.83 - 0.138z^{-1} + z^{-2}}$$

$$S_{\mathbf{yu}}(z) = \frac{0.033z^4 + 0.026z^3 + 0.005z^2 - 0.0002z}{z^4 + 0.2z^3 - 0.83z^2 - 0.084z + 0.176}$$

also represented in Fig. 4 (solid line). In this case the input and output spectra satisfy the linear-affine relation (9) with $L = 5$ and $K = 0.06$. In order to check identifiability we use Proposition 5. This requires a preliminary computation of $\bar{R}_{\mathbf{y}}$ and $\bar{R}_{\mathbf{u}}$. For this example we find $\bar{R}_{\mathbf{y}} = 2.6521$, $\bar{R}_{\mathbf{u}} = 0.5184$ and in this case we have exactly $K = \bar{R}_{\mathbf{y}} - L\bar{R}_{\mathbf{u}}$. In force of Proposition 5, the model is <u>non-identifiable</u>. As a check we may use the geometric method proposed in Beghelli, Castaldi, and Soverini (1997), interpreting (8) as the intersection of an infinite family of hyperbolas in the plane $\{\sigma_{\mathbf{u}}^2, \sigma_{\mathbf{y}}^2\}$. The intersection of all these branches in the plane is a point in the plane $\{\sigma_{\mathbf{u}}^2, \sigma_{\mathbf{y}}^2\}$ corresponding to the error variance pairs of candidate EIV models. Points of intersection, lying outside of the positive orthant do not correspond to valid EIV models. As we can see from Fig. 5 there are two nonnegative intersections $(\sigma_{\mathbf{u}}'^2, \sigma_{\mathbf{y}}'^2) = (0.3, 2.65)$, $(\sigma_{\mathbf{u}}''^2, \sigma_{\mathbf{y}}''^2) = (0.52, 1.56)$ and we may check that $\sigma_{\mathbf{y}}'^2 = \bar{R}_{\mathbf{y}}$ and $\sigma_{\mathbf{u}}''^2 = \bar{R}_{\mathbf{u}}$. The resulting spectra are drawn in Fig. 4.

6.2. An identifiable model satisfying the linear-affine relation

Assume the input, output and cross spectra are described by

$$S_{\mathbf{y}}(z) = \frac{-0.2z^3 + 30.9z^2 - 71.5z - 387.2 - 71.3z^{-1} + 30.9z^{-2} - 0.2z^{-3}}{z^3 - 6.35z^2 + 14.13z - 34.014 + 14.13z^{-1} - 6.35z^{-2} + z^{-3}}$$

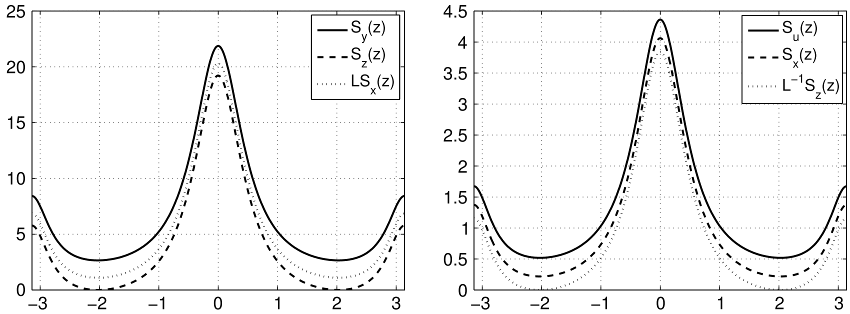


Fig. 4. Input–output and internal spectra for model 1 (dashed line) and model 2 (dotted line), Example 3.

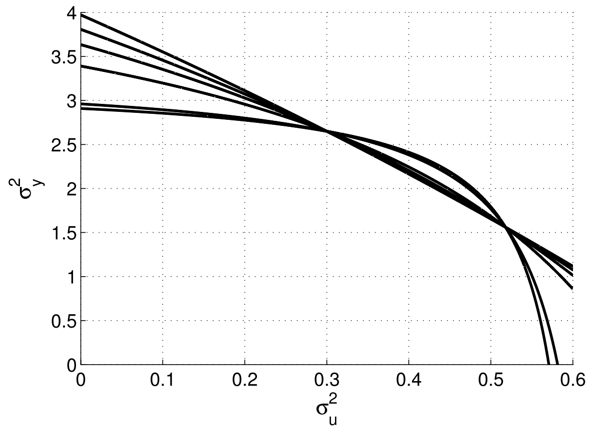


Fig. 5. Intersections of hyperbolas for example 6.1.

$$S_{\mathbf{u}}(z) = \frac{z^3 - 10.1z^2 - 24.04z - 252.9 - 24.04z^{-1} - 10.1z^{-2} + z^{-3}}{z^3 - 6.35z^2 + 14.13z - 34.014 + 14.13z^{-1} - 6.35z^{-2} + z^{-3}}$$

$$S_{\mathbf{yu}}(z) = \frac{9.29z^6 + 4.24z^5 - 0.95z^4 - 0.32z^3 - 0.037z^2 + 0.003z + 0.0001}{z^6 - 0.82z^5 + 0.59z^4 - 0.25z^3 + 0.074z^2 - 0.015z + 0.001}$$

and shown in Fig. 6. We see from the picture that there may be a linear affine relation between the two spectra. By imposing a relation of the type (9), we find $L = 1.8$, $K = -2$. In this case the necessary condition of Theorem 3 is satisfied. However non-identifiability is not guaranteed. According to Proposition 5, since here $K < 0$, we need to check if $K < -L\bar{R}_{\mathbf{u}}$, in which case the model would be identifiable.

Computing $\bar{R}_{\mathbf{u}}$ one gets $\bar{R}_{\mathbf{u}} = 0.902$, and so $-L\bar{R}_{\mathbf{u}} = -1.624$. Hence we have $K < -L\bar{R}_{\mathbf{u}}$, and the model is identifiable. We may in fact check that we have two possible variance pairs but only one of them, $(\sigma_{\mathbf{u}}^{\prime 2}, \sigma_{\mathbf{y}}^{\prime 2}) = (0.5, 0.7)$ is positive. The other, $(\sigma_{\mathbf{u}}^{\prime\prime 2}, \sigma_{\mathbf{y}}^{\prime\prime 2}) = (1.5, -1.1)$, is not feasible having a negative component. Note that this happens because $\sigma_{\mathbf{u}}^{\prime\prime 2} > \bar{R}_{\mathbf{u}}$ (the theoretical upper limit).

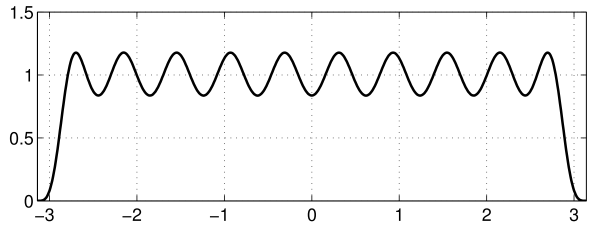


Fig. 7. Spectral profile of the noise filter.

7. A simulation experiment

We consider an experimental setup suggested by one of the referees. A vector time series realization of the bivariate process $[\mathbf{y}(t)\ \mathbf{u}(t)]$ is generated from an EIV model with correlated non-white additive noise errors obtained as a filtered linear combinations of two uncorrelated white noises $\mathbf{w}_1$ and $\mathbf{w}_2$ of unit variance, $\sigma_{\mathbf{w}_i}^2 = 1$, according to the scheme

$$\mathbf{e}_{\mathbf{y}}(t) = \sqrt{\sigma_{\mathbf{y}}^2} F(z)(c_{y,1}\mathbf{w}_1(t) + c_{y,2}\mathbf{w}_2(t))$$

$$\mathbf{e}_{\mathbf{u}}(t) = \sqrt{\sigma_{\mathbf{u}}^2} F(z)(c_{u,1}\mathbf{w}_1(t) + c_{u,2}\mathbf{w}_2(t)),$$

where $c_{i,1}^2 + c_{i,2}^2 = 1$ and $F(z)$ is a linear FIR filter, whose spectral profile is plotted in Fig. 7. The filter introduces a sort of realistic attenuation of the noise spectra at high frequencies.

From the sample time series of $\mathbf{y}(t)$, $\mathbf{u}(t)$ the power spectra of the simulated system are estimated by a standard non parametric method (Welch). These spectra will be called the "true" or "rough" spectra hereafter. The frequency plots of these spectra are the solid lines drawn in Figs. 8 and 9: naturally these true spectra do not comply with the Frisch scheme. One may produce a Frisch scheme approximation which is "best" according to some chosen identification/approximation procedure, see for example Söderström (2007). We come up with an estimated joint

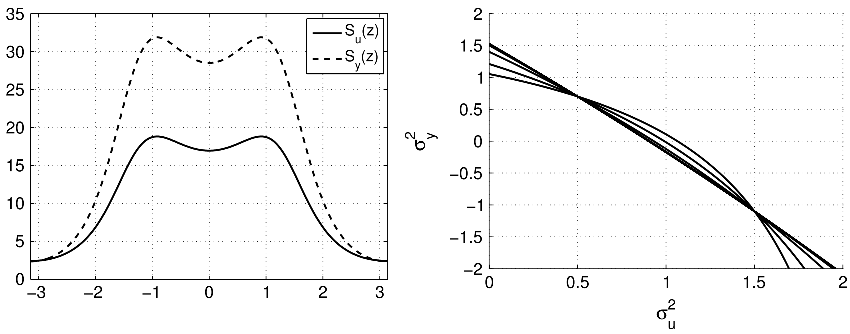


Fig. 6. Input–output spectra and intersection of hyperbolas, Example 2.

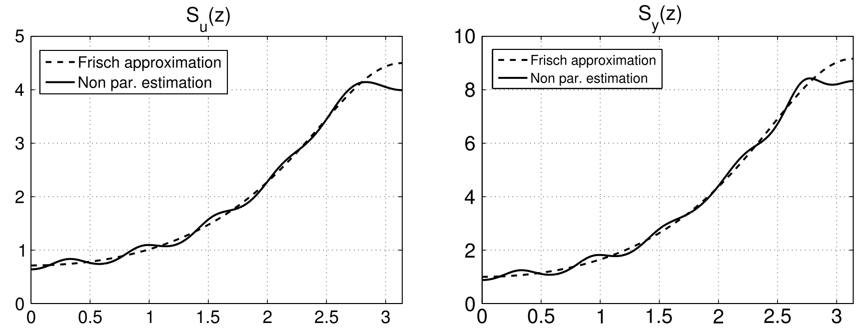


Fig. 8. Input–output estimated spectra and their approximations.

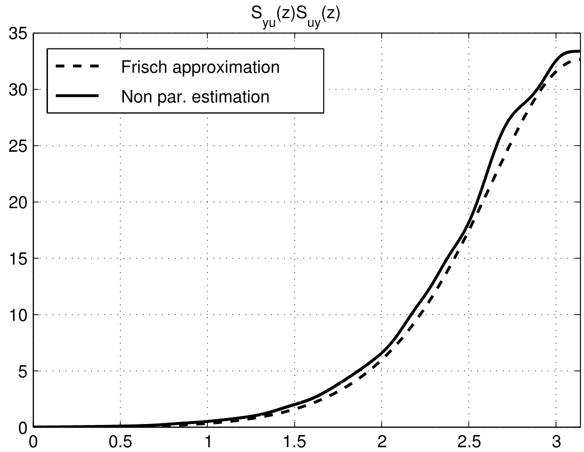


Fig. 9. Estimated cross-spectrum and its approximation.

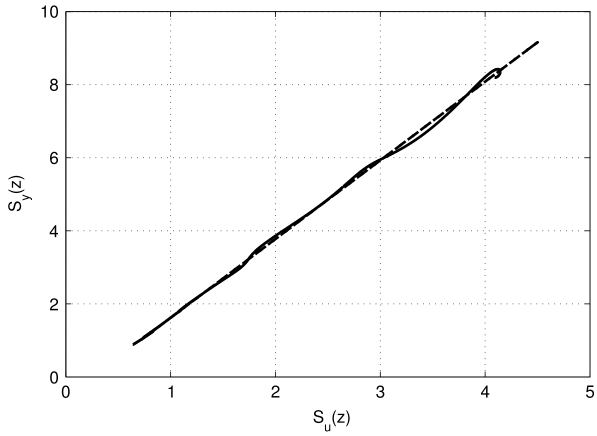


Fig. 10. Regression line on Frisch spectra vs rough spectra.

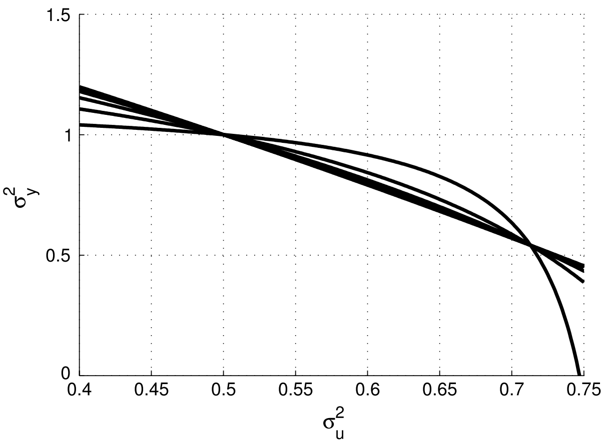


Fig. 11. Hyperbolas intersection for the spectra (15).

spectrum of the Frisch type described by[3]

$$\hat{S}_{\mathbf{y}}(z) = \frac{-2.333z + 10.3 - 2.333z^{-1}}{z + 3.633 + z^{-1}} \tag{15a}$$

$$\hat{S}_{\mathbf{u}}(z) = \frac{-0.25z + 1.705 - 0.25z^{-1}}{0.3z + 1.09 + 0.3z^{-1}} \tag{15b}$$

$$\hat{S}_{\mathbf{yu}}(z) = \frac{-1.333z^2 + 4.667z - 3.333}{z^2 + 3.633z + 1}. \tag{15c}$$

The plots of the approximate spectra are the dashed lines in Figs. 8 and 9.

At this point we may want to check for possible non-identifiability. The existence of a linear-affine relation between the two approximate output spectra (Theorem 3) is tested by fitting a linear regression of $\hat{S}_{\mathbf{y}}(e^{j\theta})$ versus $\hat{S}_{\mathbf{u}}(e^{j\theta})$ for various frequencies. The regression line is the dashed line shown in Fig. 10. In this case we can see that the linear regression is quite accurate. We find $L = 2.1556$ and $K = -0.537$. Hence we conclude that there may be another Frisch scheme model compatible with the joint spectra (15). One may argue that at this stage it may be simpler to use a specific algorithm to get estimates of $\sigma_{\mathbf{y}}^2, \sigma_{\mathbf{u}}^2$ from the given spectra (15), and thereby check directly for non-uniqueness. Alternatively one may check if the quantities $\Sigma_{\mathbf{y}}, \Sigma_{\mathbf{u}}$ are positive and are therefore interpretable as true noise variances. Experimental procedures of this kind may however be very imprecise and turn out estimates of the model variances which are affected by noise and ultimately provide wrong answers. For this reason we shall instead attempt to use the a priori criteria of Proposition 5 of Section 5.3, which only depend on the model spectra (15) and not on specific variance estimates. We find

$$\bar{R}_{\mathbf{y}} \simeq 1 \qquad \bar{R}_{\mathbf{u}} = 0.713;$$

and we see that $K = \bar{R}_{\mathbf{y}} - L\bar{R}_{\mathbf{u}}$, with good approximation, whence we can conclude, on the basis of Proposition 5, that the Frisch model describing the data (15) is *non-unique* and we have non identifiability. This can also be checked graphically by intersecting hyperbolas in the plane $\{\sigma_{\mathbf{y}}^2, \sigma_{\mathbf{u}}^2\}$ corresponding to different frequencies, see Fig. 11.

8. Discussion

Although testing for non-identifiability on the rough (true) spectra does not make sense since these spectra are in general not realizable by EIV (Frisch) models, still one may want to see how

[3] As we do not want to be tied up with any specific EIV identification procedure (each of which may give different estimates) we won't even mention which method was used in the experiment.

these tests perform on the rough data in order to get a feeling for the sensitivity of the procedure.

Running a linear regression of the rough spectrum $S_{\mathbf{y}}(z)$ on $S_{\mathbf{u}}(z)$, we obtain slightly different values of L and K, namely $L = 2.1572$, $K = -0.5706$, and the straight line gives an average error of fit of the linear approximation $\mathbf{Y} := LS_{\mathbf{u}}(z) + K$ versus the measured output spectrum,

$$e = \frac{1}{\sqrt{N}}\|\mathbf{Y} - S_{\mathbf{y}}(z)\|_2 = 0.0463,$$

($N = 4097$ is the sample size) which indicates that a linear-affine relation is a good approximation. Hence in this case we may conclude that there is a warning for the possible presence of two compatible models. Checking for actual non-identifiability cannot however be done on rough spectra and requires fitting a realizable spectrum to the data. The hyperbola intersection test is inconclusive due to sharp differences between rough and approximate spectra for certain frequencies, see Fig. 8 and also the computation of the bounds $\bar{R}_{\mathbf{y}}$ and $\bar{R}_{\mathbf{u}}$ on the rough spectra may easily become meaningless. This may happen either because of approximation errors, or also because of noise correlation. In our case we get the values $\bar{R}_{\mathbf{y}} = -0.1048$ and $\bar{R}_{\mathbf{u}} = -0.0516$, which are negative, and therefore meaningless.

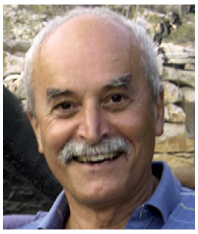
9. Conclusions

In this paper we have derived necessary conditions on the spectral densities of the external (measurable) signals under which a SISO EIV structure with white measurement noises (Frisch scheme) is *non-identifiable*. These conditions are almost sufficient, in the sense that they only disregard the nonlinear constraints of positivity of the variances of the additive noises. Our identifiability conditions for EIV models are expressible directly in terms of the external description of the observable signals, namely their joint power spectral densities. The linear-affine criterion for non-identifiability given in Theorem 3 is simple and direct and is believed to be new. Also identifiability criteria are derived in terms of the parameters K and L of the linear-affine relation.

References

Aguero, J. C., & Goodwin, G. C. (2008). Identifiability-of-errors in variables dynamic systems. *Automatica*, *55*, 371–382.
Anderson, B. D. O. (1985). Identification of scalar errors-in-variables models with dynamics. *Automatica*, *21*, 709–716.
Anderson, B. D. O., & Deistler, M. (1984). Identifiability of dynamic errors-in-variables models. *Journal of Time Series Analysis*, *5*, 1–13.
Beghelli, S., Castaldi, P., & Soverini, U. (1997). A frequential approach for errors-in-variables models. In: *4th ECC*. Brussels, Belgium.
Beghelli, S., Guidorzi, R. P., & Soverini, U. (1990). The Frisch scheme in dynamic system identification. *Automatica*, *26*(1), 171–176.
Caines, P. E. (1988). *Linear stochastic systems*. Wiley.
Castaldi, P., & Soverini, U. (1996). Identification of dynamic errors-in-variables models. *Automatica*, *32*(4), 631–636.
Chen, H., & Yang, J. (2005). Strongly consistent coefficient estimate for errors-in-variables models. *Automatica*, *41*, 1025–1033.
Deistler, M., & Anderson, B. D. O. (1989). Linear dynamic errors-in-variables models: some structure theory. *Journal of Econometrics*, *41*, 39–63.
Diversi, R., & Guidorzi, R. (2009). Frisch scheme-based identification of multi-variable errors-in-variables models. In: *15th IFAC symp. on sys. id.* Saint-Malo, France.
Fernando, K. V., & Nicholson, H. (1985). Identification of linear systems with input and output noise: the Koopmans–Levin method. In: *Proc. of IEE D*. Vol. 132 (pp. 30–36).
Frisch, R. (1934). Statistical confluence analysis of means of complete regression systems. *Technical report 5*. University of Oslo.
Gini, C. (1921). Sull'interpolazione di una retta quando i valori della variabile indipendente sono affetti da errori accidentali. *Metron*, 63–82.
Granger, C. W. J. (1963). Economic processes involving feedback. *Information and Control*, *6*, 28–48.
Guidorzi, R., & Diversi, R. (2009). A geometric approach to multivariable errors-in-variables identification. In: *15th IFAC symp. on sys. id.* Saint-Malo, France.
Kalman, R. E. (1982). System identification from noisy data. In A. Bednarek, & L. Cesari (Eds.), *Dynamical systems II* (pp. 135–164).
Madansky, A. (1959). The fitting of straight lines when both variables are subject to error. *Journal of the American Statistical Association*, *54*, 173.
Picci, G., Gei, F., & Pinzoni, S. (1993). Identification of errors-in-variables models with white measurement errors. In: *Proc. 2nd ECC* (pp. 2154–2158). Groningen, The Netherlands.
Picci, G., & Pinzoni, S. (1986). Dynamic factor-analysis models for stationary processes. *IMA Journal of Mathematical Control and Information*, *3*, 185–210.
Schachermayer, W., & Deistler, M. (1998). The set of observationally equivalent errors-in-variables models. *System & Control Letters*, *34*, 101–104.
Scherrer, W., & Deistler, M. (1998). A structure theory for linear dynamic errors-in-variables models. *SIAM Journal on Control and Optimization*, *36*(November), 2148–2175.
Söderström, T. (1981). Identification of stochastic linear systems in presence of input noise. *Automatica*, *17*, 713–725.
Söderström, T. (2007). Errors-in-variables methods in system identification. *Automatica*, *43*, 939–958.
Söderström, T., Mahata, K., & Soverini, U. (2003). Identification of dynamic errors-in-variables models: approaches based on two-dimensional ARMA modeling of the data. *Automatica*, *39*, 929–935.
Söderström, T., Mossberg, M., & Hong, M. (2009). A covariance matching approach for identifying errors-in-variables systems. *Automatica*, *45*, 2018–2031.
Söderström, T., Soverini, U., & Mahata, K. (2002). Perspectives on errors-in-variables estimation for dynamic systems. *Signal Processing*, *82*, 1139–1154.
Stoica, P., & Nehorai, A. (1987). On the uniqueness of prediction error models for systems with noisy input–output data. *Automatica*, *23*(4), 541–543.
Zheng, W. X. (1999). On least-square identification of stochastic linear systems with noisy input–output data. *International Journal of Adaptive Control and Signal Processing*, *13*, 131–143.
Zheng, W. X., & Feng, C. B. (1992). Identification of a class of dynamic errors-in-variables models. *International Journal of Adaptive Control and Signal Processing*, *6*, 431–440.

Giulio Bottegal is a Ph.D. student with the Department of Information Engineering of the University of Padova. He holds a Laurea degree (cum laude) from the University of Padova awarded in 2009. Since then has been working on identification of Errors in Variables systems.



Giorgio Picci is Full Professor with the University of Padova, Italy, Department of Information Engineering, since 1980. He graduated (cum laude) from the University of Padova in 1967 and since then has held several long term visiting appointments with various American, Japanese and European universities such as Brown University, M.I.T., the University of Kentucky, Arizona State University, the Center for Mathematics and Computer Sciences (C.W.I.) in Amsterdam, the Royal Institute of Technology, Stockholm Sweden, Kyoto University and Washington University in St. Louis, Mo. He has been contributing to Systems and Control theory mostly in the area of modeling, estimation and identification of stochastic systems and published over 100 papers and edited three books in this area. He has been involved in various joint research projects with industry and state agencies. He has been chairman of the IFAC Technical Committee on Stochastic Systems, a past member of the EUCA council, project manager of the Italian team for the Commission of the European Communities Network of Excellence System Identification (ERNSI) and general coordinator of the Commission of European Communities IST project RECSYS, in the fifth Framework Program. Giorgio Picci is a Life Fellow of the IEEE, a 2009 IFAC Fellow and a foreign member of the Swedish Royal Academy of Engineering Sciences.

Stefano Pinzoni holds a Laurea degree (1980) from the University of Padova and a Ph.D. (1989) from Arizona State University, both in Electrical Engineering. He worked as an Automation Engineer from 1982 to 1984 to the development of innovative technologies for sea and subsea applications in the oil and gas engineering. He had a research position from 1984 to 1992 at LADSEB-CNR in Padova. During the academic year 1992–93 he was an Associate Professor at the University of Calabria, Department of Electronics, Informatics and Systems. Since 1993 he has been an Associate Professor at the University of Padova, Faculty of Engineering, and in 1998 he joined the Department of Information Engineering. His research interests lie in the field of systems and control theory (linear time-varying models), stochastic modelling (EIV, FA models) and identification. He has collaborated with several international researchers and visited a number of research institutions abroad. He held a postdoctoral position in 1990–91 at the University of Groningen, Department of Mathematics.