Theoretical and practical aspects of the convergence of the SRIVC estimator for over-parameterized models ⋆

Rodrigo A. González [a], Cristian R. Rojas [a], Siqi Pan [b], James S. Welsh [b]

[a] *Division of Decision and Control Systems, KTH Royal Institute of Technology, 10044 Stockholm, Sweden*

[b] *School of Electrical Engineering and Computing, University of Newcastle, Callaghan, 2308 NSW, Australia*

Abstract

The Simplified Refined Instrumental Variable method for Continuous-time systems (SRIVC) is one of the most popular direct methods for linear continuous-time system identification. Only recently, some optimal asymptotic properties of this method have been proven. Here, we provide a comprehensive analysis of the convergence of the SRIVC method for over-parameterized models. The results we derive are used to discuss practical aspects of the method and to analyze the behavior of the normalized error variance norm, which is a central part in Young's Information Criterion for model order selection. The theoretical findings are validated through simulation examples.

Key words: System identification; continuous-time systems; instrumental variable method; model order selection.

1 Introduction

Continuous-time system identification deals with the problem of estimating continuous-time systems based on sampled input and output data. To this end, the Refined Instrumental Variable method for Continuous-time systems (RIVC, [27]), and its simpler embodiment, the Simplified RIVC (SRIVC), are widely regarded as some of the most successful direct methods for continuous-time system identification. These methods have been successfully applied in diverse areas, such as distillation columns [5], semiconductors [3], environmental [26] and biological [12] modeling.

Despite more than forty years of use, formal theoretical treatments of the SRIVC algorithm have only been pursued in recent years. The generic consistency of the SRIVC estimator was proven in [11], and in [13] it was shown that this estimator is also asymptotically efficient for an output error model structure. An extension of the SRIVC estimator for input signals not necessarily reconstructed through zero or first order holds was proposed in [7]. These results have shown that the SRIVC estimator is generically consistent under the condition that the input is persistently exciting of a sufficient order and that the model structure contains the true system, with over-parameterization possibly occurring in the numerator or denominator of the transfer function description but not both simultaneously.

Analysis of identification methods when over-parameterization is present has been covered extensively in discrete-time system identification. For instance, global minimum points of the loss function in the Maximum Likelihood method describe the true transfer function with arbitrary pole-zero cancellations [15,17]. The possible convergence points of the Steiglitz-McBride algorithm have also been shown to enjoy the same property [19], as well as various instrumental variable variants [16]. Contrary to continuous-time iterative methods such as the SRIVC estimator, the discrete-time methods studied in the literature do not need to address the intersample behavior of the signals, nor take the sampling period into account in their analysis. These aspects are of particular interest in the SRIVC estimator, as the algorithm requires specifying the intersample behaviors of the input and output; an incorrect assumption on the input intersample behavior can lead to inconsistent estimates with asymptotic bias decreasing with smaller sampling periods [11]. In addition, the SRIVC algorithm computes an estimate of the covariance matrix of

⋆ This paper was not presented at any IFAC meeting. Corresponding author: R. A. González.

Email addresses: grodrigo@kth.se (Rodrigo A. González), crro@kth.se (Cristian R. Rojas), siqi.pan@uon.edu.au (Siqi Pan), james.welsh@newcastle.edu.au (James S. Welsh).

the parameter vector as a byproduct, which is used for computing the Young Information Criterion (YIC, [24]) for model order selection. Thus, understanding how the SRIVC estimator performs under over-parameterized scenarios can help interpret the normalized estimation error variance norm (NEVN), which is part of the YIC formula. It can also reveal the theoretical and practical problems associated with over-parameterization in continuous-time system identification.

In this work, we analyze the properties of the SRIVC estimator in situations where the model structure is over-parameterized in both the numerator and denominator polynomials of the transfer function. In summary, the main technical contributions of this paper are:

- We show that, with over-parameterization of both numerator and denominator, the numerical conditioning of the modified normal matrix in the SRIVC algorithm deteriorates for decreasing sampling periods;
- we prove that the SRIVC estimator, at a limit point of the iterative procedure with a fixed non-zero sampling period and as the sample size grows unbounded, returns the true transfer function description with additional arbitrary pole-zero cancellations;
- we perform a one-iteration analysis of the SRIVC method for large sample sizes and small sampling periods, which reveals that arbitrary pole-zero cancellations also arise after one step of the algorithm; and
- we derive the asymptotic behavior of the NEVN in the YIC expression for different cases of over-parameterization.

The paper is organized as follows. In the next section, we describe both the system and model. The SRIVC algorithm is briefly explained in Section 3. In Section 4, we present the main theoretical findings, and we corroborate them through simulations in Section 5. Section 6 concludes this work.

2 Problem formulation

Consider the linear time-invariant, asymptotically stable continuous-time system

$$x(t) = G^*(p)u(t),$$

with the system transfer function being parameterized as

$$G^*(p)=\frac{B^*(p)}{A^*(p)}=\frac{b^*_{m^*}p^{m^*}+b^*_{m^*-1}p^{m^*-1}+\cdots+b^*_0}{a^*_{n^*}p^{n^*}+a^*_{n^*-1}p^{n^*-1}+\cdots+a^*_1p+1},$$

where p is the derivative operator, i.e., $px(t) = \mathrm{d}x(t)/\mathrm{d}t$, and the polynomials $B^*(p)$ and $A^*(p)$ are assumed to be coprime. The true parameter vector is given by

$$\boldsymbol{\theta}^* = \left[a^*_1,\ \ldots,\ a^*_{n^*},\ b^*_0,\ \ldots,\ b^*_{m^*}\right]^\top.$$

The output is assumed to be contaminated with noise prior to sampling, i.e.,

$$y(t_k) = x(t_k) + v(t_k),$$

where the measurement noise $\{v(t_k)\}$ is a zero-mean stochastic process that is independent of the sampled input signal $\{u(t_k)\}$. The intersample behavior of the input applied to the system is assumed known and produced by a hold device such as a zero or first order hold (ZOH or FOH, respectively). Both input and output signals are sampled uniformly with sampling period h. Throughout this work, the output equation will also be written as

$$y(t_k) = \left\{\frac{B^*(p)}{A^*(p)}u(t)\right\}_{t_k} + v(t_k), \tag{1}$$

where the notation $\{G^*(p)u(t)\}_{t_k}$ means that the continuous-time input $\{u(t)\}$ is filtered through $G^*(p)$, and the resultant signal is evaluated at $t = t_k$. In contrast, the notation $G^*(p)u(t_k)$, often used in the description of continuous-time estimators, implies that the discrete-time signal $\{u(t_k)\}$ is interpolated using either a ZOH or FOH and the resulting output of the filter is sampled at t_k [6].

We retrieve the signals $\{u(t_k), y(t_k)\}_{k=1}^N$ from an identification experiment, where N is the number of samples, and a model of the form

$$G(p) = \frac{b_m p^m + b_{m-1}p^{m-1} + \cdots + b_0}{a_n p^n + a_{n-1}p^{n-1} + \cdots + a_1 p + 1}$$

is fit to the data. We assume that the model is over-parameterized in both the numerator and denominator polynomials, that is, $m > m^*$ and $n > n^*$. The goal is to determine the asymptotic properties of the SRIVC algorithm and to examine the practical implications of selecting over-parameterized models on the performance of this method.

3 The SRIVC estimator

The SRIVC estimator is an iterative instrumental variable algorithm first introduced in [27], in which parameter-dependent continuous-time filters are updated at each iteration. In each step, the instruments are computed using the parameter estimate obtained in the previous iteration until the model parameters have converged.

This method is described in Algorithm 1, where we denote the filtered regressor vector by $\boldsymbol{\varphi}_f(t_k, \boldsymbol{\theta}_j)$, the filtered instrument vector by $\hat{\boldsymbol{\varphi}}_f(t_k, \boldsymbol{\theta}_j)$, and the filtered output by $y_f(t_k, \boldsymbol{\theta}_j)$, all of which depend on the iterative estimate $\boldsymbol{\theta}_j$, with j being the iteration index. The first term of (5) inside the inverse is called the modified normal matrix of the SRIVC estimator, as it resembles the normal matrix present in Gauss-Newton iterations. For the filtering steps in (2), (3) and (4), the discrete-time signals $\{u(t_k)\}$ and $\{y(t_k)\}$ must be interpolated in some manner, e.g., using a ZOH or FOH, and the resultant output through the continuous-time filters is sampled at t_k. Contrary to the usual implementation of the SRIVC algorithm in the CONTSID Toolbox [4], here we consider an anti-monic denominator polynomial for the model, i.e., the constant term of the denominator polynomial is fixed to one. With this description, it is feasible that the model denominator converges to the system denominator even if it is over-parameterized, as the leading coefficient of the denominator is not fixed to one. The drawback of this implementation is that, for that same case, the leading coefficient will gravitate towards zero from the positive and negative real axes for large sample sizes, which might induce instability problems in the iterations.

Remark 1. *An extension of the SRIVC method has been proposed in [7] for continuous-time multisine inputs, which cannot be exactly reconstructed through ZOH or FOH devices. The analysis done in the present work also holds for that algorithm subject to some minor changes in the proofs. However, in this paper our attention is only focused on ZOH and FOH input signals.*

4 Analysis

In this section we present the main contributions of this paper. It is divided into four subsections: analysis of the modified normal matrix, convergence analysis of the stationary points, a one-iteration analysis of the SRIVC algorithm, and a review of the normalized error variance norm, which is a central part of the YIC criterion.

In [11], it was shown that the SRIVC estimator is generically consistent under mild conditions. One of these conditions is that the numerator or denominator may be over-parameterized, but not both. In other words, we must have

$$\min(n - n^*, m - m^*) = 0,$$

where n and m are the denominator and numerator polynomial degrees, respectively. One question we address is whether some notion of consistency is still achievable for model structures that are over-parameterized in both the numerator and denominator polynomials. Before stating our results, we introduce the assumptions used throughout this work.

Algorithm 1: SRIVC

Input: $\{u(t_k), y(t_k)\}_{k=1}^N$, model order (n, m), initial vector estimate $\boldsymbol{\theta}_1 \in \mathbb{R}^{n+m+1}$, tolerance ϵ and maximum number of iterations M

Using $\boldsymbol{\theta}_1$, form the model polynomials $A_1(p)$ and $B_1(p)$

$j \leftarrow 1$, flag $\leftarrow 1$

while flag $= 1$ and $j \leq M$ **do**

Prefilter $\{u(t_k)\}_{k=1}^N$ and $\{y(t_k)\}_{k=1}^N$ to form

$$\boldsymbol{\varphi}_f(t_k, \boldsymbol{\theta}_j) \leftarrow \left[\frac{-p}{A_j(p)} y(t_k), \ldots, \frac{-p^n}{A_j(p)} y(t_k), \frac{1}{A_j(p)} u(t_k), \ldots, \frac{p^m}{A_j(p)} u(t_k) \right]^\top \quad (2)$$

$$\hat{\boldsymbol{\varphi}}_f(t_k, \boldsymbol{\theta}_j) \leftarrow \left[\frac{-pB_j(p)}{A_j^2(p)} u(t_k), \ldots, \frac{-p^n B_j(p)}{A_j^2(p)} u(t_k), \frac{1}{A_j(p)} u(t_k), \ldots, \frac{p^m}{A_j(p)} u(t_k) \right]^\top \quad (3)$$

$$y_f(t_k, \boldsymbol{\theta}_j) \leftarrow \frac{1}{A_j(p)} y(t_k) \quad (4)$$

Compute the parameter estimate

$$\boldsymbol{\theta}_{j+1} \leftarrow \left[\sum_{k=1}^N \hat{\boldsymbol{\varphi}}_f(t_k, \boldsymbol{\theta}_j) \boldsymbol{\varphi}_f^\top(t_k, \boldsymbol{\theta}_j) \right]^{-1} \cdot \left[\sum_{k=1}^N \hat{\boldsymbol{\varphi}}_f(t_k, \boldsymbol{\theta}_j) y_f(t_k, \boldsymbol{\theta}_j) \right] \quad (5)$$

if $1/A_{j+1}(p)$ is unstable **then**

Reflect the unstable poles of $1/A_{j+1}(p)$ into the stable region of the complex plane

end if

if $\frac{\|\boldsymbol{\theta}_{j+1} - \boldsymbol{\theta}_j\|}{\|\boldsymbol{\theta}_j\|} < \epsilon$ **then**

flag $\leftarrow 0$

end if

$j \leftarrow j + 1$

end while

Output: $\boldsymbol{\theta}_j$ and its associated model $B_j(p)/A_j(p)$.

(A1) The true system $B^*(p)/A^*(p)$ is proper ($n^* \geq m^*$) and asymptotically stable with $A^*(p)$ and $B^*(p)$ being coprime.

(A2) The input sequence, $u(t_k)$, and disturbance, $v(t_i)$, are stationary and mutually independent for all integers k and i.

(A3) The input sequence is persistently exciting of order no less than $2n$.

(A4) The sampling frequency is more than twice the largest positive imaginary part of the zeros of $A_j(p)A^*(p)$.

(A5) The input is constant between samples (i.e., it is known exactly and it has a ZOH behavior).

Assumptions (A1) to (A5) are standard and have been used in other consistency works, such as in [11]. Analogous results for cases when the input is generated by a first-order hold device can also be obtained from our approach, under the condition that the input must be persistently exciting of order at least $2n+1$ instead of $2n$.

Aside from the previously stated assumptions on the stationarity of signals, persistence of excitation, sampling period and intersample behavior, we also require the following:

(A6) The degrees of the model polynomials, n and m, satisfy

$$\min(n-n^*, m-m^*) = n_{op} > 0, \tag{6}$$

and

$$2n^* + n_{op} \geq n+m. \tag{7}$$

(A7) For every j, the j-th iteration of the model polynomials $A_j(p)$ and $B_j(p)$ of the SRIVC algorithm has at most n_{op} pole-zero cancellations, with $A_j(p)$ being an asymptotically stable polynomial.

Condition (6) in Assumption (A6) holds when the model is over-parameterized in both numerator and denominator polynomials, which is the case of interest in this work. On the other hand, the technical condition (7) is set to limit the amount of over-parameterization allowed to ensure that the transfer functions in the matrix being inverted in the SRIVC algorithm are proper.

4.1 Ill-conditioning of the modified normal matrix of SRIVC

If the model order is known exactly, i.e., $n = n^*$ and $m = m^*$, then as N tends to infinity and under the stationarity conditions in Assumption (A2), the ergodic lemmas in [14, Lemma 3.1] and [18, Lemma A4.3] permit us to write (5) as

$$\boldsymbol{\theta}_{j+1} = \mathbb{E}\{\hat{\boldsymbol{\varphi}}_f(t_k,\boldsymbol{\theta}_j)\boldsymbol{\varphi}_f^\top(t_k,\boldsymbol{\theta}_j)\}^{-1}\mathbb{E}\{\hat{\boldsymbol{\varphi}}_f(t_k,\boldsymbol{\theta}_j)y_f(t_k,\boldsymbol{\theta}_j)\}. \tag{8}$$

For over-parameterized models, we show in Proposition 2 that the condition number of the matrix being inverted in (8) grows unbounded as the sampling period tends to zero.

Proposition 2. *Consider a continuous-time system with polynomial degrees n^* and m^* in the denominator and numerator, respectively. Suppose Assumptions (A1), (A2), (A6) and (A7) hold. Then, for a fixed input sequence, for all $K > 0$ there exists a sampling period $h > 0$ sufficiently small such that the condition number of the $(n+m+1)\times(n+m+1)$ modified normal matrix*

$$\mathbb{E}\{\hat{\boldsymbol{\varphi}}_f(t_k,\boldsymbol{\theta}_j)\boldsymbol{\varphi}_f^\top(t_k,\boldsymbol{\theta}_j)\} \tag{9}$$

is greater than K.

Proof. We introduce a theoretical regressor vector $\tilde{\boldsymbol{\varphi}}_f(t_k,\boldsymbol{\theta}_j)$ that has the form

$$\tilde{\boldsymbol{\varphi}}_f(t_k,\boldsymbol{\theta}_j) = \mathbf{v}_f(t_k) + \frac{1}{A_j(p)A^*(p)}\times \\ \left[-pB^*(p),\ldots,-p^nB^*(p),A^*(p),\ldots,p^mA^*(p)\right]^\top u(t_k),$$

where $\mathbf{v}_f(t_k)$ is a vector of filtered noise, which is uncorrelated with the filtered input signals present in $\hat{\boldsymbol{\varphi}}_f(t_k,\boldsymbol{\theta}_j)$. Since $\max(n+m^*, m+n^*) = n+m-n_{op}$, we can write

$$\tilde{\boldsymbol{\varphi}}_f(t_k,\boldsymbol{\theta}_j) = \mathbf{S}(-B^*,A^*)\frac{1}{A_j(p)A^*(p)}\mathbf{u}_{n+m-n_{op}}(t_k) + \mathbf{v}_f(t_k),$$

where

$$\mathbf{u}_{n+m-n_{op}}(t_k) := \left[p^{n+m-n_{op}}, p^{n+m-n_{op}-1},\ldots,1\right]^\top u(t_k), \tag{10}$$

and $\mathbf{S}(-B^*,A^*)$ is a $(n+m+1)\times(n+m-n_{op}+1)$ Sylvester-type rectangular matrix of the form

$$\mathbf{S}(-B^*,A^*) = \\ \begin{array}{l} \left[\begin{array}{cccccc} 0 & & -b^*_{m^*} & -b^*_{m^*-1} & \cdots & -b^*_0 \;\; 0 \\ & \iddots & \iddots & & \iddots & \iddots \\ -b^*_{m^*} & -b^*_{m^*-1} & \cdots & -b^*_0 & 0 & 0 \\ \hline 0 & & a^*_{n^*} & a^*_{n^*-1} & \cdots & a^*_1 \;\; 1 \\ & \iddots & \iddots & & \iddots & \iddots \\ a^*_{n^*} & a^*_{n^*-1} & \cdots & a^*_1 & 1 & 0 \end{array}\right] \end{array} \begin{array}{l} \left.\vphantom{\begin{array}{c}0\\ \iddots\\ 0\end{array}}\right\} n \text{ rows} \\ \left.\vphantom{\begin{array}{c}0\\ \iddots\\ 0\end{array}}\right\} m+1 \text{ rows}, \end{array} \tag{11}$$

where zeros are padded to the left in the upper or lower block matrix if $n_{op} = n-n^*$ or $n_{op} = m-m^*$ respectively. Next, we introduce a vector $\boldsymbol{\Delta}(t_k,\boldsymbol{\theta}_j)$, whose entries are given by

$$\boldsymbol{\Delta}_i(t_k,\boldsymbol{\theta}_j) = \\ \begin{cases} \dfrac{-p^i}{A_j(p)}\left\{\dfrac{B^*(p)}{A^*(p)}u(t)\right\}_{t_k} + \dfrac{p^iB^*(p)}{A_j(p)A^*(p)}u(t_k), & i=1,\ldots,n, \\ 0, & i=n+1,\ldots,n+m+1, \end{cases}$$

where we have used the notation introduced in (1). Now we can write

$$\boldsymbol{\varphi}_f(t_k,\boldsymbol{\theta}_j) = \tilde{\boldsymbol{\varphi}}_f(t_k,\boldsymbol{\theta}_j) + \boldsymbol{\Delta}(t_k,\boldsymbol{\theta}_j),$$

and therefore, the modified normal matrix in (9) is given by

$$\mathbb{E}\{\hat{\boldsymbol{\varphi}}_f(t_k,\boldsymbol{\theta}_j)\boldsymbol{\varphi}_f^\top(t_k,\boldsymbol{\theta}_j)\} = \mathbb{E}\{\hat{\boldsymbol{\varphi}}_f(t_k,\boldsymbol{\theta}_j)\boldsymbol{\Delta}^\top(t_k,\boldsymbol{\theta}_j)\} + \mathbb{E}\left\{\hat{\boldsymbol{\varphi}}_f(t_k,\boldsymbol{\theta}_j)\frac{1}{A_j(p)A^*(p)}\mathbf{u}_{n+m-n_{op}}^\top(t_k)\right\}\mathbf{S}^\top(-B^*,A^*).$$

Note that $\mathbf{S}(-B^*,A^*)$ has rank equal to $n+m+1-n_{op}$ (see Lemma 6 of the Appendix), which is strictly less than $n+m+1$ by Assumption (A6). Furthermore, each entry of $\boldsymbol{\Delta}(t_k,\boldsymbol{\theta}_j)$ goes to zero as the sampling period tends to zero, since they can be viewed as the difference between the output of a continuous-time cascaded system with a sampler between the cascaded sections and the output of the same system without the sampler. In other words, when $h\to 0$ we have $\boldsymbol{\Delta}(t_k,\boldsymbol{\theta}_j)\to\mathbf{0}$ for any bounded input signal. These observations imply that the matrix in (9) will converge to a singular matrix as h tends to zero, and thus the smallest singular value of (9) tends to zero for a decreasing sampling period.

Finally, in order to show that the condition number grows unbounded, we need to show that the largest singular value of the matrix in (9) does not tend to zero as $h\to 0$. Denote by $\mathbf{e}_i$ the i-th column of the identity matrix of an appropriate size, and by $\sigma_1(\cdot)$ the largest singular value of a matrix. By Theorem 5.6.2 of [10], we find that

$$\begin{aligned}\sigma_1&\left(\mathbb{E}\{\hat{\boldsymbol{\varphi}}_f(t_k,\boldsymbol{\theta}_j)\boldsymbol{\varphi}_f^\top(t_k,\boldsymbol{\theta}_j)\}\right)\\ &= \max_{\|\mathbf{x}\|_2=\|\mathbf{y}\|_2=1}|\mathbb{E}\{\mathbf{y}^\top\hat{\boldsymbol{\varphi}}_f(t_k,\boldsymbol{\theta}_j)\boldsymbol{\varphi}_f^\top(t_k,\boldsymbol{\theta}_j)\mathbf{x}\}|\\ &\geq |\mathbb{E}\{\mathbf{e}_{n+1}^\top\hat{\boldsymbol{\varphi}}_f(t_k,\boldsymbol{\theta}_j)\boldsymbol{\varphi}_f^\top(t_k,\boldsymbol{\theta}_j)\mathbf{e}_{n+1}\}|\\ &= \mathbb{E}\left\{\left(\frac{1}{A_j(p)}u(t_k)\right)^2\right\}.\end{aligned}$$

Since the spectrum of the input is assumed fixed for different sampling periods, this lower bound does not go to zero when h tends to zero. This fact, together with the smallest singular value tending to zero, leads to the desired result. □

4.2 Asymptotic analysis of the stationary points of the SRIVC estimator

Proposition 2 suggests that, for large sample sizes, the iterations of the SRIVC estimator become increasingly ill-conditioned for smaller sampling periods if an overparameterized model is considered. If the matrix being inverted in (5) becomes singular for a finite N, an SRIVC estimate can be computed as a solution of the system of equations

$$\sum_{k=1}^N \hat{\boldsymbol{\varphi}}_f(t_k,\boldsymbol{\theta}_j)\boldsymbol{\varphi}_f^\top(t_k,\boldsymbol{\theta}_j)\boldsymbol{\theta}_{j+1} = \sum_{k=1}^N \hat{\boldsymbol{\varphi}}_f(t_k,\boldsymbol{\theta}_j)y_f(t_k,\boldsymbol{\theta}_j). \tag{12}$$

Since the matrix on the left hand side of (12) may be singular or close to singular, we are interested in analyzing the set of stationary points of the algorithm, that is, the set of all parameter vectors $\bar{\boldsymbol{\theta}}$ that satisfy

$$\sum_{k=1}^N \hat{\boldsymbol{\varphi}}_f(t_k,\bar{\boldsymbol{\theta}})\boldsymbol{\varphi}_f^\top(t_k,\bar{\boldsymbol{\theta}})\bar{\boldsymbol{\theta}} = \sum_{k=1}^N \hat{\boldsymbol{\varphi}}_f(t_k,\bar{\boldsymbol{\theta}})y_f(t_k,\bar{\boldsymbol{\theta}}). \tag{13}$$

The asymptotic behavior of all stationary points is studied next.

Theorem 3. *As the sample size tends to infinity and under Assumptions (A1) to (A7), the stationary points of the SRIVC method generically*[1] *(with respect to the parameter space of denominator polynomials of degree n) describe a continuous-time model of the form*

$$\frac{L(p)B^*(p)}{L(p)A^*(p)},$$

where $L(p)$ is an arbitrary polynomial of degree n_{op}, and $A^*(p)$ and $B^*(p)$ are the true system polynomials.

Proof. Take $\bar{\boldsymbol{\theta}}$ as a stationary point of the SRIVC algorithm, which must belong to the solution set of (13). This vector forms the transfer function polynomials $\bar{A}(p)$ and $\bar{B}(p)$, which are written as

$$\bar{A}(p) := \hat{A}(p)L(p), \quad \bar{B}(p) := \hat{B}(p)L(p),$$

where $\hat{A}(p)$ and $\hat{B}(p)$ are coprime (i.e., they form the minimal realization of $\bar{B}(p)/\bar{A}(p)$), with $\hat{A}(p)$ being anti-monic. The polynomial $L(p)$ is described by $L(p) = 1+l_1p+\cdots+l_{n_l}p^{n_l}$, with $n_l\geq 0$ being the number of pole-zero cancellations of $\bar{B}(p)/\bar{A}(p)$. Note that this factorization can be done without loss of generality, since $n_l=0$ (i.e., $L(p)=1$) if $\bar{A}(p)$ and $\bar{B}(p)$ are coprime. Since

$$y_f(t_k)-\boldsymbol{\varphi}_f^\top(t_k,\bar{\boldsymbol{\theta}})\bar{\boldsymbol{\theta}} = y(t_k)-\frac{\bar{B}(p)}{\bar{A}(p)}u(t_k),$$

the SRIVC estimate must satisfy

$$\frac{1}{N}\sum_{k=1}^N \hat{\boldsymbol{\varphi}}_f(t_k,\bar{\boldsymbol{\theta}})\left(y(t_k)-\frac{\hat{B}(p)}{\hat{A}(p)}u(t_k)\right) = \mathbf{0}.$$

[1] A statement s is *generically true* [18] with respect to $\boldsymbol{\rho}\in\Omega$, if the set $\{\boldsymbol{\rho}|\boldsymbol{\rho}\in\Omega, s \text{ is not true}\}$ has Lebesgue measure zero in the open set Ω.

As the number of samples tends to infinity, under stationarity conditions and independence between noise and input signals in Assumption (A2), we obtain that

$$\mathbb{E}\left\{\hat{\boldsymbol{\varphi}}_f(t_k,\bar{\boldsymbol{\theta}})\left[\left(\frac{B^*(p)}{A^*(p)}-\frac{\hat{B}(p)}{\hat{A}(p)}\right)u(t_k)\right]\right\}=\mathbf{0}. \quad (14)$$

Furthermore, the filtered instrument vector is given by

$$\begin{aligned}\hat{\boldsymbol{\varphi}}_f(t_k,\bar{\boldsymbol{\theta}}) &= \left[\frac{-p\hat{B}(p)}{\bar{A}(p)\hat{A}(p)}u(t_k),\ \dots,\ \frac{-p^n\hat{B}(p)}{\bar{A}(p)\hat{A}(p)}u(t_k),\right.\\ &\qquad \left.\frac{\hat{A}(p)}{\bar{A}(p)\hat{A}(p)}u(t_k),\ \dots,\ \frac{p^m\hat{A}(p)}{\bar{A}(p)\hat{A}(p)}u(t_k)\right]^\top\\ &= \mathbf{S}(-\hat{B},\hat{A})\frac{1}{\bar{A}(p)\hat{A}(p)}\mathbf{u}_{n+m-n_l}(t_k), \end{aligned} \quad (15)$$

where $\mathbf{u}_{n+m-n_l}(t_k)$ and $\mathbf{S}(-\hat{B},\hat{A})$ are defined in a similar fashion to (10) and (11) respectively. By the same argument presented in Lemma 6 in the Appendix, we can show that $\text{rank}\{\mathbf{S}(-\hat{B},\hat{A})\}=n+m-n_l+1$. So, by (14), for every $i=0,1,\dots,n+m-n_l$ we must have

$$\mathbb{E}\left\{\frac{p^i}{L(p)\hat{A}^2(p)}u(t_k)\left[\left(\frac{B^*(p)}{A^*(p)}-\frac{\hat{B}(p)}{\hat{A}(p)}\right)u(t_k)\right]\right\}=0. \quad (16)$$

Due to the linearity of these conditions in p^i, we can study a subset of these conditions of the form

$$\mathbb{E}\left\{\frac{L(p)p^i}{L(p)\hat{A}^2(p)}u(t_k)\left[\left(\frac{B^*(p)}{A^*(p)}-\frac{\hat{B}(p)}{\hat{A}(p)}\right)u(t_k)\right]\right\}=0,$$

for $i=0,1,\dots,n+m-2n_l$. These equations are equivalent to

$$\mathbb{E}\left\{\frac{p^i}{\hat{A}^2(p)}u(t_k)\left[\left(\frac{B^*(p)}{A^*(p)}-\frac{\hat{B}(p)}{\hat{A}(p)}\right)u(t_k)\right]\right\}=0 \quad (17)$$

for $i=0,1,\dots,n+m-2n_l$. We have reached a set of conditions that depend only on $\hat{A}(p)$ and $\hat{B}(p)$, which are coprime polynomials of order $n-n_l$ and $m-n_l$, respectively. Next, we define $H(p):=\hat{A}(p)B^*(p)-\hat{B}(p)A^*(p)=h_rp^r+h_{r-1}p^{r-1}+\dots+h_0$, where

$$\begin{aligned} r &= \max(n+m^*,m+n^*)-n_l\\ &= n+m-n_{op}-n_l. \end{aligned}$$

With this, (17) can also be written as

$$\mathbb{E}\left\{\frac{1}{\hat{A}^2(p)}\mathbf{u}_{n+m-2n_l}(t_k)\frac{1}{A^*(p)\hat{A}(p)}\mathbf{u}^\top_{n+m-n_{op}-n_l}(t_k)\right\}\mathbf{h}=\mathbf{0}, \quad (18)$$

with $\mathbf{h}$ being the vector formed by the coefficients of $H(p)$. Denote the matrix of size $(n+m-2n_l)\times(n+m-n_{op}-n_l)$ in (18) as $\boldsymbol{\Phi}$. The next step is to prove that, generically,

$$\text{rank}\{\boldsymbol{\Phi}\}=n+m-n_{op}-n_l. \quad (19)$$

For this, we require $n_l\leq n_{op}$, which is true by Assumption (A7). The rank condition in (19) is proven in Lemma 7 in the Appendix. Thus, (18) implies $\mathbf{h}=\mathbf{0}$, which is in turn equivalent to

$$\begin{aligned} &\hat{A}(p)B^*(p)-\hat{B}(p)A^*(p)=0\\ \implies &\frac{\hat{B}(p)}{\hat{A}(p)}=\frac{B^*(p)}{A^*(p)}. \end{aligned} \quad (20)$$

Note that $\hat{A}(p)$ and $\hat{B}(p)$ have been treated as polynomials of degree (at most) $n-n_l$ and $m-n_l$, respectively. Since they are coprime and (20) is satisfied, we must have $n_l=n_{op}$, and the coefficients related to the possible excess of degree in the numerator or denominator of $\hat{B}(p)/\hat{A}(p)$ go to zero.

Regarding the polynomial $L(p)$, we see that the conditions in (16) are satisfied irrespective of the coefficients that form $L(p)$, due to (20). Thus, $L(p)$ is left unspecified, and with this we conclude the proof. □

Theorem 3 reveals that the SRIVC algorithm delivers a transfer function estimate whose minimal realization converges to the true one as N tends to infinity, even under over-parameterization. The excess of zeros and poles may play a big role in how this estimator reaches the stationary points as the stability of the excess poles is not guaranteed. This problem is usually solved in the SRIVC mechanism by reflecting the unstable poles as mentioned in line 7 of Algorithm 1. However, the user must exercise caution when choosing a high denominator order, as over-parameterization of the denominator may lead to difficulties in convergence in the practical implementations of the algorithm and inaccurate phase shifts in the estimated transfer function at each iteration. Note that if the model is unstable at some iteration of the SRIVC method, the pole reflection procedure in line 7 of Algorithm 1 changes the phase of the transfer function estimate. A more accurate stabilization procedure has been proposed in [8], which involves constraining the SRIVC estimate within ellipsoids that yield stable estimates.

4.3 *One-iteration analysis of the SRIVC method for small sampling periods*

We now study the iterations of the SRIVC method for large N. In Theorem 4, we show that one step of the SRIVC algorithm also leads to the true transfer function together with pole-zero cancellations for very small sampling periods.

Theorem 4. *Consider the SRIVC iterations given by* (12)*, and assume that Assumptions (A1) to (A7) hold. As the sample size tends to infinity, the asymptotic estimate (as $N \to \infty$) at the $(j+1)$-th iteration $\boldsymbol{\theta}_{j+1}$ satisfies*

$$A_{j+1}(p) = L(p)A^*(p) \quad \text{and} \quad B_{j+1}(p) = L(p)B^*(p)$$

as the sampling period goes to zero, where $L(p)$ is an arbitrary polynomial of degree n_{op}.

Proof. Denote the number of pole-zero cancellations of the transfer function at the j-th iteration, $B_j(p)/A_j(p)$, as $n_l \geq 0$. Its minimal realization transfer function description is denoted by $\hat{B}_j(p)/\hat{A}_j(p)$. Since

$$y_f(t_k,\boldsymbol{\theta}_j)-\boldsymbol{\varphi}_f^\top(t_k,\boldsymbol{\theta}_j)\boldsymbol{\theta}_{j+1}=\frac{A_{j+1}(p)}{A_j(p)}y(t_k)-\frac{B_{j+1}(p)}{A_j(p)}u(t_k),$$

the equation in (12) can be rewritten, as the number of samples tends to infinity, as

$$\mathbb{E}\left\{\hat{\boldsymbol{\varphi}}_f(t_k,\boldsymbol{\theta}_j)\left(\frac{A_{j+1}(p)}{A_j(p)}y(t_k)-\frac{B_{j+1}(p)}{A_j(p)}u(t_k)\right)\right\}=\mathbf{0}.$$

Since the input is assumed uncorrelated with the measurement noise, we have

$$\mathbb{E}\left\{\hat{\boldsymbol{\varphi}}_f(t_k,\boldsymbol{\theta}_j)\left(\frac{A_{j+1}(p)}{A_j(p)}\left\{\frac{B^*(p)}{A^*(p)}u(t)\right\}_{t_k}-\frac{B_{j+1}(p)}{A_j(p)}u(t_k)\right)\right\}=\mathbf{0}, \tag{21}$$

Note that, in terms of computing (21), the transfer functions $A_{j+1}(p)/A_j(p)$ and $B^*(p)/A^*(p)$ are discretized separately.

By defining

$$\varepsilon_u(t_k):=\frac{A_{j+1}(p)}{A_j(p)}\left\{\frac{B^*(p)}{A^*(p)}u(t)\right\}_{t_k}-\left(\frac{A_{j+1}(p)B^*(p)}{A_j(p)A^*(p)}\right)u(t_k), \tag{22}$$

the condition in (21) can be expressed as

$$\mathbb{E}\left\{\hat{\boldsymbol{\varphi}}_f(t_k,\boldsymbol{\theta}_j)\varepsilon_u(t_k)\right\}+ \\ \mathbb{E}\left\{\hat{\boldsymbol{\varphi}}_f(t_k,\boldsymbol{\theta}_j)\frac{A_{j+1}(p)B^*(p)-B_{j+1}(p)A^*(p)}{A_j(p)A^*(p)}u(t_k)\right\}=\mathbf{0}.$$

Note that in general we have $\varepsilon_u(t_k)\neq 0$. However, just as for $\boldsymbol{\Delta}_i(t_k,\boldsymbol{\theta}_j)$ in the proof of Proposition 2, as the sampling period goes to zero we have $\varepsilon_u(t_k)\to 0$ for any bounded input signal.

Similar to the proof of Theorem 3, we factor the model polynomials of the j-th iteration as $A_j(p)=\hat{A}_j(p)L(p)$ and $B_j(p)=\hat{B}_j(p)L(p)$, where $L(p)$ is a polynomial of degree n_l. With this notation in mind, we write the instrument vector $\hat{\boldsymbol{\varphi}}_f(t_k,\boldsymbol{\theta}_j)$ in the same form as in (15):

$$\hat{\boldsymbol{\varphi}}_f(t_k,\boldsymbol{\theta}_j)=\mathbf{S}(-\hat{B}_j,\hat{A}_j)\frac{1}{A_j(p)\hat{A}_j(p)}\mathbf{u}_{n+m-n_l}(t_k),$$

where $\mathbf{S}(-\hat{B}_j,\hat{A}_j)$ is a rectangular Sylvester-type matrix of size $(n+m+1)\times(n+m-n_l+1)$, which has rank equal to $n+m-n_l+1$. Next, we define a vector $\mathbf{h}\in\mathbb{R}^r$ that satisfies

$$A_{j+1}(p)B^*(p)-B_{j+1}(p)A^*(p)=\left[p^r,\,p^{r-1},\,\ldots,\,p,\,1\right]\mathbf{h},$$

where

$$r=\max(n+m^*,m+n^*)=n+m-n_{op}.$$

Thus, as the sample size tends to zero, condition (21) is equivalent to

$$\mathbf{0}=\mathbf{S}(-\hat{B}_j,\hat{A}_j)\times \\ \mathbb{E}\left\{\frac{1}{A_j(p)\hat{A}_j(p)}\mathbf{u}_{n+m-n_l}(t_k)\frac{1}{A_j(p)A^*(p)}\mathbf{u}_{n+m-n_{op}}^\top(t_k)\right\}\mathbf{h}.$$

Since $\mathbf{S}(-\hat{B}_j,\hat{A}_j)$ has full column rank, and the expectation above is a matrix with rank $n+m-n_{op}+1$ by the same logic used in Lemma 7, we must have $\mathbf{h}=\mathbf{0}$. Following the same steps as in the proof of Lemma 6, we see that $\mathbf{h}=\mathbf{0}$ is equivalent to

$$A_{j+1}(p)=L(p)A^*(p) \quad \text{and} \quad B_{j+1}(p)=L(p)B^*(p),$$

where $L(p)$ is an arbitrary polynomial of degree n_{op}. □

Remark 5. *Theorem 4 implies that, even if $B_j(p)/A_j(p)$ were to describe the true system (with the extra numerator and denominator terms being zero), there is no guarantee of convergence of the polynomials $B_j(p)$ and $A_j(p)$ separately. That is, the SRIVC algorithm does not guarantee that $B_{j+1}(p)=B_j(p)$ and $A_{j+1}(p)=A_j(p)$ when the sample size tends to infinity. Rather, the convergence can be in terms of the minimal realization of the transfer function $B_j(p)/A_j(p)$ but the polynomial coefficients will likely not converge. We refer to this as the parameter jumping effect. Also, there is no constraint on the stability of $L(p)$, indicating that the iterations may return unstable estimates in many occasions. In practice, the problem of unstable estimates at each iteration of the algorithm can be mitigated by forcing pole-zero cancellations through, e.g., the* `minreal` *command in MATLAB. However, additional poles must be added to this minimal realization for the next iteration, as the filter $p^n/A_j(p)$ must be causal for constructing the filtered regressor vector.*

The theoretical analysis done here suggests that an alternative stopping criterion could be proposed for mitigating the parameter jumping effect. The commonly used stopping criterion in line 10 of Algorithm 1 focuses on the parameter vector and not on the transfer function it describes. In the scenario studied in this work, the parameter vector may not converge to a fixed vector, but to an equivalence class formed by vectors with the same minimal transfer function representation. Thus, we propose a stopping criterion that might be more useful, i.e.,

$$\frac{\left\| F(p)\left(\frac{B_{j+1}(p)}{A_{j+1}(p)} - \frac{B_j(p)}{A_j(p)}\right)\right\|_2}{\left\| F(p)\frac{B_j(p)}{A_j(p)}\right\|_2} < \epsilon, \tag{23}$$

where $F(p)$ is a user-defined filter that has high gain in the bandwidth of interest of the continuous-time system, and $\|\cdot\|_2$ is the $\mathcal{L}_2$-norm for continuous-time systems. This stopping criterion is well justified, as it has been seen that upon convergence of the underlying transfer function, we may expect pole-zero quasi-cancellations on the $j+1$-th iteration and hence a very similar minimal realization to the previous iteration for large sample size. However, note that no matter what the stopping criterion is, jumps between transfer functions with quasi-cancellations will be frequent in the iteration steps.

4.4 Young Information Criterion revisited

In this subsection, we provide a theoretical justification of the normalized error variance norm, which is an important term in Young's Information Criterion (YIC) for model order selection [25].

Two widely used criteria for model selection are the Akaike and Bayesian information criteria (AIC and BIC respectively). For instrumental variable methods, a natural idea is to use the information of the covariance matrix of the parametric estimation errors, which is indirectly estimated in the instrumental variable procedure. This approach has led to the YIC, which is defined by

$$\text{YIC} = \log\left(\frac{\hat{\sigma}_e^2}{\hat{\sigma}_y^2}\right) + \log\left(\frac{\hat{\sigma}_e^2}{n+m+1}\sum_{i=1}^{n+m+1}\frac{\mathbf{P}_{ii}}{\bar{\boldsymbol{\theta}}_i^2}\right).$$

Here, the first term is related to the coefficient of determination $R^2 := 1-\hat{\sigma}_e^2/\hat{\sigma}_y^2$, while the second is the logarithm of the normalized estimation error variance norm (NEVN). The expressions $\hat{\sigma}_e^2$ and $\hat{\sigma}_y^2$ are the sample covariances of the residual and the model output respectively, and $\{\mathbf{P}_{ii}\}_{i=1}^{n+m+1}$ are the diagonal elements of the matrix

$$\mathbf{P} = \left[\sum_{k=1}^{N}\hat{\boldsymbol{\varphi}}_f(t_k,\bar{\boldsymbol{\theta}})\boldsymbol{\varphi}_f^\top(t_k,\bar{\boldsymbol{\theta}})\right]^{-1}. \tag{24}$$

The R^2 coefficient is useful for determining under-parameterization, and it usually reaches a plateau value once the model structure contains the true system [22]. On the other hand, the NEVN can detect over-parameterization, and thus our attention will be focused on this term.

It is argued in [21] and [22] that if the model is over-parameterized, the matrix being inverted in (24) will tend to be singular, which gives rise to at least one diagonal element of $\mathbf{P}$ being of high magnitude. These works consider biproper discrete-time transfer functions, in which only the degree of the model denominator is trying to be determined. Proposition 2 of this paper shows that ill-conditioning will occur in the continuous-time case for arbitrary proper transfer functions, when both the numerator and denominator are over-parameterized and the sampling period is small.

In the case where only one polynomial (numerator or denominator) is over-parameterized and the denominator is anti-monic, it has been shown in Theorem 1 of [11] that this matrix is non-singular as N tends to infinity. Thus, it is expected that $\{\mathbf{P}_{ii}\}_{i=1}^{n+m+1}$ will no longer reflect an over-parameterization problem in these cases. For the following analysis we will study two cases, depending on the existence of an element in $\bar{\boldsymbol{\theta}}$ that tends to zero as N tends to infinity. Such an element can describe the leading coefficient estimate of the numerator or denominator polynomials which go to zero almost surely due to the generic consistency of the SRIVC estimator [2], or any parameter estimate whose true value is zero.

First, we study how the NEVN behaves when consistency is achieved and the estimate does not have a parameter that tends to zero as N tends to infinity. Under stationarity assumptions, by the continuous mapping theorem [20, Theorem 2.3] we have

$$N\mathbf{P} \xrightarrow{a.s.} \mathbb{E}\{\hat{\boldsymbol{\varphi}}_f(t_k,\boldsymbol{\theta}^*)\boldsymbol{\varphi}_f^\top(t_k,\boldsymbol{\theta}^*)\}^{-1}.$$

Since $\bar{\boldsymbol{\theta}} \xrightarrow{a.s.} \boldsymbol{\theta}^*$ due to the consistency of the SRIVC estimator, the NEVN is expected to decrease as $1/N$ for large N.

Now, we assume that there exists at least one parameter estimate $\bar{\boldsymbol{\theta}}_i$ that tends to zero almost surely as N tends to infinity. Following similar lines to [13, Theorem 2], it can be shown that $\sqrt{N}\bar{\boldsymbol{\theta}}_i$ is asymptotically normally distributed with zero mean and variance denoted here as $\tilde{\mathbf{P}}_{ii}$. Thus, again by the continuous mapping theorem,

$$\frac{N\bar{\boldsymbol{\theta}}_i^2}{\tilde{\mathbf{P}}_{ii}} \xrightarrow{dist.} \chi^2(1),$$

[2] Recall that this effect does not necessarily occur when both numerator and denominator are over-parameterized, as proven in Theorem 3.

which implies that $N\bar{\boldsymbol{\theta}}_i^2$ is asymptotically gamma-distributed with shape $1/2$ and scale $2\tilde{\mathbf{P}}_{ii}$. Therefore, contrary to the previous scenario, the NEVN on average does not decrease as $1/N$.

In summary, we have found that the NEVN can correctly discriminate the correct number of poles and zeros among all models that contain the true system, as a decay of $1/N$ is expected for the correct model order, whereas for an excess of numerator and/or denominator order this behavior will not be observed. However, note that if there are coefficients that are zero in the numerator of the true transfer function (which arises when, e.g., there is a zero at the origin), then the NEVN term of the YIC will not decay as $1/N$, even if the correct number of poles and zeros is considered.

5 Simulations

We now confirm the theoretical results provided in the previous section through simulations. We start by exemplifying the problems associated with over-parameterization, and afterwards we show how the iterations of the SRIVC algorithm, implemented following Algorithm 1, behave for small sampling periods. This section ends with a simulation study of the NEVN. For the first three subsections, we consider a first-order system of the form

$$G^*(p) = \frac{10}{0.1p+1}.$$

The input is an i.i.d. Gaussian sequence with a unit variance, which is then interpolated with a ZOH device. The additive noise on the output is also an i.i.d. Gaussian sequence of unit variance which is uncorrelated with the input. The model structure we consider has one zero and two poles, that is, it is of the form

$$G(p) = \frac{b_1 p + b_0}{a_2 p^2 + a_1 p + 1}.$$

5.1 *Condition number for small sampling periods*

We test the condition number of the modified normal matrix in (9) given by one iteration of the SRIVC method, for five different sampling periods $h = 10^{-i}[s]$, with $i = 2, 3, \ldots, 6$. The length of each experiment is kept constant ($20[s]$), and 300 Monte Carlo simulations are performed for each sample size.

The average condition number of the modified normal matrix for each sampling period is reported in Table 1. The condition number is high for all sampling periods, and it grows even larger when the sampling period decreases to zero. This result is in line with Proposition 2, and it confirms that the SRIVC estimator is poorly conditioned when there exists over-parametrization and the sampling period is very low.

Table 1
Mean condition number of the modified normal matrix for different sampling periods.

Sampl. period	10^{-2} [s]	10^{-3} [s]	10^{-4} [s]	10^{-5} [s]	10^{-6} [s]
Cond. number	$1.72{\cdot}10^7$	$2.94{\cdot}10^7$	$5.60{\cdot}10^7$	$3.31{\cdot}10^8$	$1.08{\cdot}10^9$

5.2 *Difficulties associated with over-parameterization: a case study*

We consider one realization of the data with a sample size of $N = 2\cdot 10^5$, and a sampling period of $h = 0.01$[s]. The tolerance in the SRIVC algorithm is set to $\epsilon = 10^{-4}$, which is considered not very stringent, and the maximum number of iterations is set to 60. The SRIVC iterations are initialized with the LSSVF method [23] with the cut-off frequency of the prefilter being $\lambda = 10$[rad/s].

In Figure 1 we plot the poles, zero, and static gain of the models provided by each iteration of the SRIVC estimate. We show the poles before the step of stabilization of the SRIVC algorithm. Note that the number of unstable estimates is quite significant (28 out of 60 iterations gave unstable estimates), which requires the user to implement a stabilization procedure such as the one in line 7 of Algorithm 1.

Figure 1 shows that the static gain and pole are accurately estimated in each iteration. Pole-zero quasi-cancellations can be found in every step, which is in line with the analysis done in Section 4.3. However, these pole-zero pairs do not seem to converge to a fixed pair, at least in the first 60 iterations. After the first 50 iterations, the estimates oscillate indefinitely between two transfer functions with pole-zero quasi-cancellations.

In Figure 2, the convergence of the models estimated through SRIVC was tested under three different stopping criteria: the $\boldsymbol{\theta}$-stopping criterion mentioned in line 10 of Algorithm 1, the same criterion but with prior `minreal` pole-zero cancellation of the associated transfer functions with state-elimination tolerance [3] set to 10^{-3}, and the metric in (23). The last metric is computed with the `norm` command in MATLAB, and a stabilizing all-pass filter of unit norm is incorporated in case $(B_{j+1}(p)/A_{j+1}(p) - B_j(p)/A_j(p))$ is unstable. The filter $F(p)$ is chosen as a 6-th order Butterworth band-pass filter with cut-off frequencies 0.1 and 100[rad/s]:

[3] The `minreal` command in MATLAB cancels all pole-zero pairs that are within a specified tolerance. This tolerance should not be confused with the tolerance ϵ of the SRIVC algorithm (see line 10 of Algorithm 1).

$$F(p)=\frac{10^6p^3}{(p+10^{-2})(p^2+10^{-2}p+10^{-4})(p+10^2)(p^2+10^2p+10^4)}.$$

From Figure 2, we see that the estimates do not converge according to the $\boldsymbol{\theta}$-stopping criterion, but some benefit can be seen after around 50 iterations by considering minimal realizations of the models under a high tolerance for pole-zero cancellations. On the other hand, the metric in (23) quickly drops two orders of magnitude in the first iterations and later it fluctuates. No metric converges to zero, as convergence does not occur in the parameter space and the fluctuating pole-zero quasi-cancellations are never exact for a finite number of samples.

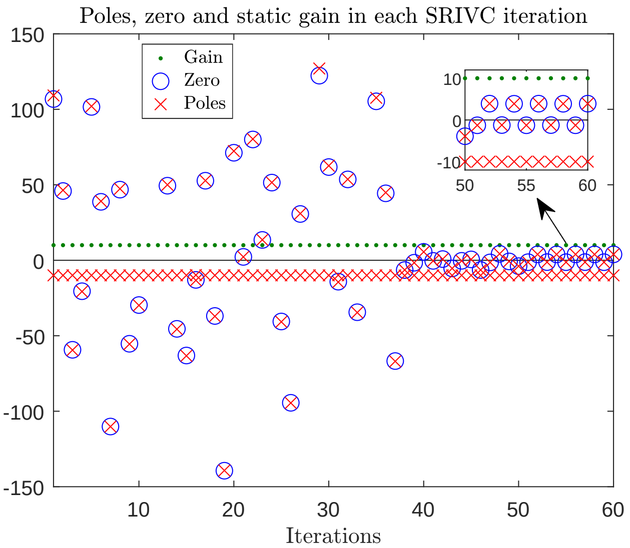

Fig. 1. Poles (red crosses), zero (blue circle) and static gain (green dashed line) of the estimates given in each iteration of the SRIVC algorithm.

5.3 *One iteration test and effect of the sampling period*

We now check the estimates provided by one SRIVC iteration for a large sample size. Using a sample size of $N = 2 \cdot 10^5$, we obtain an initial estimate of the parameter vector $\boldsymbol{\theta} = [a_2, a_1, b_1, b_0]^\top$ by the LSSVF method with the cut-off frequency, $\lambda = 10$[rad/s]. Next, one iteration of the SRIVC algorithm is performed without reflecting the unstable poles if the resulting transfer function is unstable. Three sampling periods are tested: $h = 0.005$[s], $h = 0.02$[s] and $h = 0.05$[s]. Note that according to the sampling criterion given in [2], the sampling period should be chosen between 0.01[s] and 0.025[s].

In Figure 3, we plot the zero and poles of eight estimates corresponding to eight different noise realizations. For the smaller sampling periods, pole-zero quasi-cancellations can be more clearly seen than for the largest sampling periods. This effect is correctly

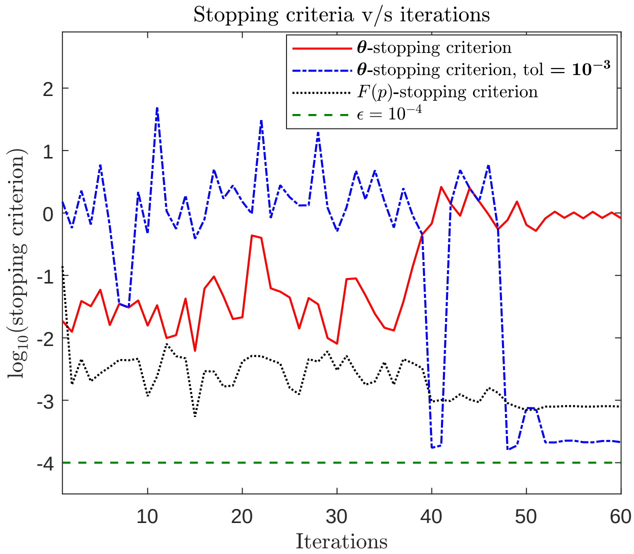

Fig. 2. $\boldsymbol{\theta}$-based and transfer function-based stopping criteria metrics evaluated at each iteration, compared to a baseline $\epsilon = 10^{-4}$.

predicted by Theorem 4 and the analysis of $\varepsilon_u(t_k)$ in (22). Note that the quasi-cancellations occur irregularly in the real line, which suggests that their locations are noise-dependent (and also dependent on the initial estimate) and therefore non-predictable, which agrees with our theoretical findings. We also plot the Bode diagrams of each estimated model in Figure 4. The estimated transfer functions are more accurate for smaller sampling periods; this occurs since $\varepsilon_u(t_k)$ is smaller in magnitude, which leads to estimates that are closer to the true transfer function with added pole-zero cancellations. Note that this behavior is not guaranteed for small sample sizes, in view of the fact that the data might miss the dominant dynamics of the true system if the duration of the experiment is too short [1].

5.4 *Behavior of the NEVN*

We now verify the properties of NEVN for different sample sizes and model structures through a series of simulation tests. Consider the second-order system

$$G^*(p) = \frac{1}{0.04p^2 + 0.2p + 1},$$

which is excited by a zero-mean Gaussian white noise of unit variance that is interpolated with a zero-order hold. The sampling period is set to $h = 0.1$[s], and an additive white noise of variance 0.1 contaminates the noiseless output. Three different model structures are tested: $n = 2, m = 0$ (exact parameterization), $n = 2, m = 1$ (over-parameterization of the numerator) and $n = 3, m = 1$ (over-parameterization of both numerator and denominator). Two hundred Monte Carlo runs are performed for six different sample sizes.

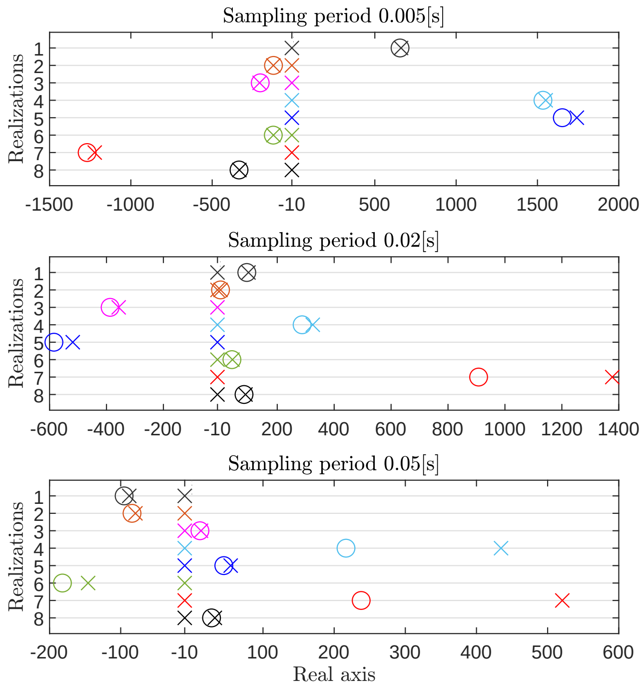

Fig. 3. Pole-zero plots of 8 estimates (separated through the vertical axis) of $G^*(p)$ for sampling periods $h = 0.005$[s], $h = 0.02$[s] and $h = 0.05$[s]. All the estimated poles are real numbers in this study.

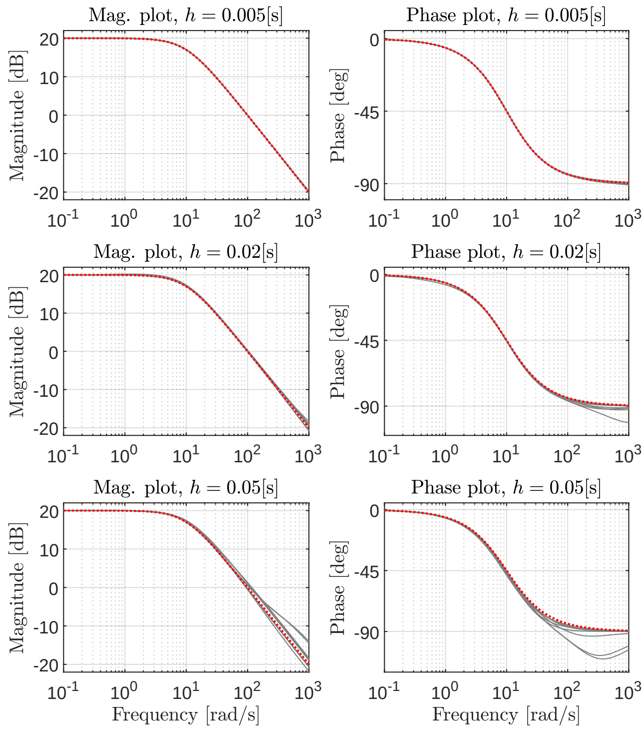

Fig. 4. Bode plots for 8 estimates of $G^*(p)$ in grey, for sampling periods $h = 0.005$[s] (upper), $h = 0.02$[s] (middle) and $h = 0.05$[s] (lower plots). In dashed red, the true system.

In Figure 5 we have plotted violin plots [9] for the values of log(NEVN) ranging in sample sizes for the three cases previously explained. Violin plots are modifications of the standard box plot that show estimated probability density distributions of the data, together with the median value as white circles and quartiles as grey lines around the median. Note that we have chosen to plot densities of log(NEVN) instead of NEVN in order to cover the wide range of values obtained in the simulation test. From Figure 5 we see that the distributions of log(NEVN) vary considerably among the three cases studied. For the exact parameterization, log(NEVN) decays as $\log(1/N)$ and the data points decrease in variance as N grows. These findings support the analysis made in Section 4.4. For the over-parameterized numerator case, log(NEVN) no longer decays for larger N. Moreover, the distribution of log(NEVN) seems to converge as N grows, which is also in line with Section 4.4. In contrast to the previous scenarios, the over-parameterization of both numerator and denominator produces a more erratic behavior, as values of NEVN range through several orders of magnitude. Note that the pole-zero quasi-cancellation can possibly be enforced by a vector $\bar{\boldsymbol{\theta}}$ that has very large parameters, which may reduce the NEVN by a considerable amount.

6 Conclusions

In this paper, we have studied the properties of the SRIVC estimator when the model is over-parameterized in both numerator and denominator polynomials. Statistical properties of the method have been proven, such as the ill-conditioning of its associated modified normal matrix and the exact transfer function described by its stationary points when the sample size tends to infinity. The pole-zero quasi-cancellations that the SRIVC estimator delivers (both in its stationary points and per iteration) have been discussed in theory and corroborated through simulations. It has also been shown that the NEVN is well-justified as a criterion for determining over-parameterization in the model numerator and/or denominator polynomials.

The theoretical developments have also led to practical implications for the SRIVC estimator. When over-parameterized models are used, this estimator may be numerically ill-conditioned and its iterates generally do not converge in the parameter space. Erratic behavior is to be expected in each iteration, and the standard stopping criterion used in the SRIVC algorithm fails to address these convergence problems.

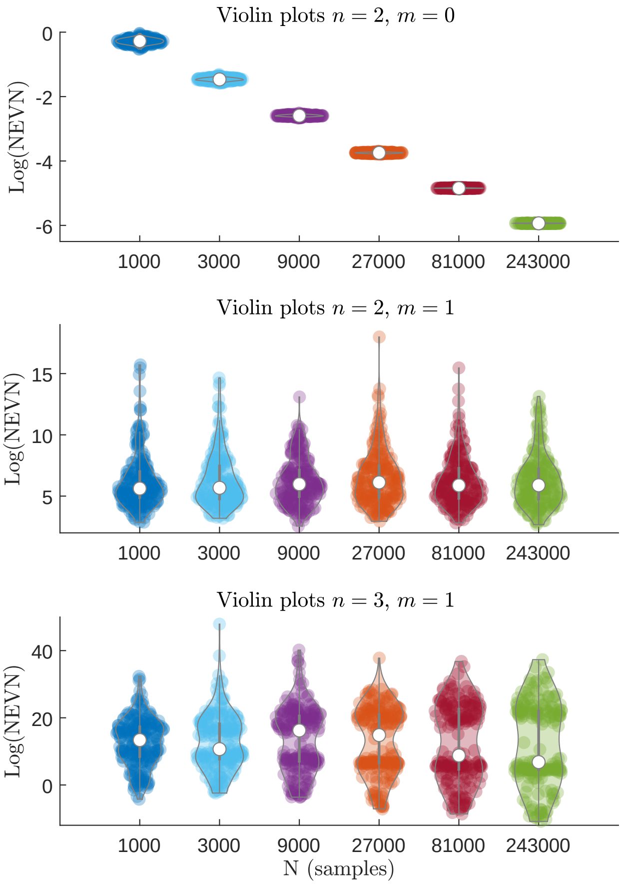

Fig. 5. Violin plots of log(NEVN) for different sample sizes. From top to bottom: exact model structure, numerator over-parameterization, and both numerator and denominator over-parameterization.

Acknowledgements

This work was partially supported by the Swedish Research Council under contract number 2016-06079 (NewLEADS) and partially by the Australian government Research Training Program (RTP) scholarship.

Appendix

Lemma 6 (Rectangular Sylvester matrices [17]). *Under Assumption (A6), consider the rectangular Sylvester-type matrix* $\mathbf{S}(-B^*, A^*)$ *in* (11) *where zeros are padded to the left in the upper or lower block matrix if* $n_{op} = n - n^*$ *or* $n_{op} = m - m^*$ *respectively. If* $A^*(p)$ *and* $B^*(p)$ *are coprime polynomials in* p*, then the rank of* $\mathbf{S}(-B^*, A^*)$ *is* $n + m - n_{op} + 1$*.*

Proof. Take $\mathbf{x} = [\alpha_1, \ldots, \alpha_n, \beta_0, \ldots, \beta_m]^\top$. The equation

$$\mathbf{x}^\top \mathbf{S}(-B^*, A^*) = \mathbf{0}$$

can be written as an equality between polynomials:

$$(\alpha_1 p + \cdots + \alpha_n p^n) B^*(p) = (\beta_0 + \cdots + \beta_m p^m) A^*(p).$$

Since the polynomials in the left and right hand side have the same zeros, it follows that the general solution must be

$$\alpha_1 p + \cdots + \alpha_n p^n = L(p) A^*(p),$$

and

$$\beta_0 + \cdots + \beta_m p^m = L(p) B^*(p),$$

where $L(p)$ is an arbitrary polynomial of degree $\min(n - n^*, m - m^*) = n_{op}$. Thus, $\mathbf{x}$ lies in an n_{op}-dimensional subspace, which is the left null-space of $\mathbf{S}(-B^*, A^*)$. By the Rank-Nullity Theorem [10], this dimension is equal to $n+m+1-\text{Rank}\{\mathbf{S}(-B^*, A^*)\}$, from which the lemma follows. $\square$

Lemma 7 (Rank of $\boldsymbol{\Phi}$). *Consider the matrix*

$$\boldsymbol{\Phi} := \mathbb{E}\left\{\frac{1}{\hat{A}^2(p)} \mathbf{u}_{n+m-2n_l}(t_k) \frac{1}{A^*(p)\hat{A}(p)} \mathbf{u}^\top_{n+m-n_{op}-n_l}(t_k)\right\},$$

and assume that Assumptions (A1), (A3), (A4) and (A6) hold. Then, the statement $\{\text{rank}\{\boldsymbol{\Phi}\} = n+m-n_{op}-n_l\}$ *is generically true with respect to the parameters forming* $\hat{A}(p)$*.*

Proof. To show that $\text{rank}\{\boldsymbol{\Phi}\} = n + m - n_{op} - n_l$, we can study the generic non-singularity of the following sub-matrix of $\boldsymbol{\Phi}$:

$$\boldsymbol{\Phi}_1 := \mathbb{E}\left\{\frac{1}{\hat{A}^2(p)} \mathbf{u}_{n+m-n_{op}-n_l}(t_k) \frac{1}{A^*(p)\hat{A}(p)} \mathbf{u}^\top_{n+m-n_{op}-n_l}(t_k)\right\}.$$

By Lemma 7 of [11], it can be shown that the matrix $\boldsymbol{\Phi}_1$ with $\hat{A}(p) = A^*(p)$ is positive definite when the input signal $\{u(t_k)\}$ is persistently exciting of order at least $2n$ for ZOH interpolation or $2n + 1$ for FOH interpolation. Note that condition (7) is used at this stage for the properness of the transfer functions in $\boldsymbol{\Phi}_1$ when $\hat{A}(p) = A^*(p)$, as the minimum relative degree of such transfer functions is given by

$$2n^* - (n - m - n_{op} - n_l) \geq 2n^* - n - m + n_{op} \geq 0.$$

Moreover, it has been shown in Lemma 9 of [11] that the entries of $\boldsymbol{\Phi}_1$ are analytic functions of the denominator parameters. Hence, by Lemma A2.3 of [18] and its corollary, we conclude that $\boldsymbol{\Phi}_1$ is generically nonsingular with respect to the parameters forming $\hat{A}(p)$. Since the columns of $\boldsymbol{\Phi}_1$ are generically linearly independent, it is generically true that the extended matrix $\boldsymbol{\Phi}$ has rank $n + m - n_{op} - n_l$. $\square$

References

[1] K. J. Åström. On the choice of sampling rates in parametric identification of time series. *Information Sciences*, 1(3):273–278, 1969.

[2] K. J. Åström and B. Wittenmark. *Computer Controlled Systems: Theory and Design.* Prentice-Hall, 1984.

[3] H. Garnier, R. R. Bitmead, and R. A. de Callafon. Direct continuous-time model identification of high-powered light-emitting diodes from rapidly sampled thermal step response data. *IFAC Proceedings Volumes*, 47(3):6430–6435, 2014.

[4] H. Garnier, M. Gilson, T. Bastogne, and M. Mensler. The CONTSID toolbox: A software support for data-based continuous-time modelling. In *Identification of Continuous-time Models from Sampled Data*, pages 249–290. Springer, 2008.

[5] H. Garnier, M. Gilson, P. C. Young, and E. Huselstein. An optimal IV technique for identifying continuous-time transfer function model of multiple input systems. *Control Engineering Practice*, 15(4):471–486, 2007.

[6] H. Garnier and L. Wang (Eds.). *Identification of Continuous-time Models from Sampled Data.* Springer, 2008.

[7] R. A. González, C. R. Rojas, S. Pan, and J. S. Welsh. Consistent identification of continuous-time systems under multisine input signal excitation. *Automatica,* Article 109859, 2021.

[8] H. Ha and J. S Welsh. Ensuring stability in continuous time system identification instrumental variable method for over-parameterized models. In *53rd IEEE Conference on Decision and Control*, pages 2597–2602, 2014.

[9] J. L. Hintze and R. D. Nelson. Violin plots: a box plot-density trace synergism. *The American Statistician*, 52(2):181–184, 1998.

[10] R. A. Horn and C. R. Johnson. *Matrix Analysis,* 2nd Ed. Cambridge University Press, 2012.

[11] S. Pan, R. A. González, J. S. Welsh, and C. R. Rojas. Consistency analysis of the simplified refined instrumental variable method for continuous-time systems. *Automatica,* Article 108767, 2020.

[12] S. Pan, J. S. Welsh, A. Brichta, H. Drury, and J. G. Stoddard. Estimating the membrane properties of vestibular type II hair cells using continuous-time system identification. In *Proceedings of the IFAC World Congress,* Berlin, 2020.

[13] S. Pan, J. S. Welsh, R. A. González, and C. R. Rojas. Efficiency analysis of the simplified refined instrumental variable method for continuous-time systems. *Automatica,* Article 109196, 2020.

[14] T Söderström. Ergodicity results for sample covariances. *Probl. Control Inf. Theory*, 4(2):131–138, 1975.

[15] T. Söderström. On the uniqueness of maximum likelihood identification. *Automatica*, 11(2):193–197, 1975.

[16] T. Söderström and P. Stoica. Comparison of some instrumental variable methods–Consistency and accuracy aspects. *Automatica*, 17(1):101–115, 1981.

[17] T. Söderström and P. Stoica. Some properties of the output error method. *Automatica*, 18(1):93–99, 1982.

[18] T. Söderström and P. Stoica. *Instrumental Variable Methods for System Identification.* Springer, 1983.

[19] P. Stoica and T. Söderström. The Steiglitz-McBride identification algorithm revisited–Convergence analysis and accuracy aspects. *IEEE Transactions on Automatic Control*, 26(3):712–717, 1981.

[20] A. W. Van der Vaart. *Asymptotic Statistics.* Cambridge University Press, 2000.

[21] P. E. Wellstead. An instrumental product moment test for model order estimation. *Automatica*, 14(1):89–91, 1978.

[22] P. Young, A. Jakeman, and R. McMurtrie. An instrumental variable method for model order identification. *Automatica*, 16(3):281–294, 1980.

[23] P. C. Young. Process parameter estimation and self adaptive control. *IFAC Proceedings Volumes*, 2(2):118–140, 1965.

[24] P. C. Young. Recursive estimation, forecasting and adaptive control. *Control and Dynamic Systems*, 30(3):119–165, 1989.

[25] P. C. Young. *Recursive Estimation and Time-Series Analysis: An Introduction for the Student and Practitioner.* Springer, 2012.

[26] P. C. Young and H. Garnier. Identification and estimation of continuous-time, data-based mechanistic (DBM) models for environmental systems. *Environmental Modelling & Software*, 21(8):1055–1072, 2006.

[27] P. C. Young and A. J. Jakeman. Refined instrumental variable methods of recursive time-series analysis. Part III, Extensions. *International Journal of Control*, 31(4):741–764, 1980.