TU/e EINDHOVEN UNIVERSITY OF TECHNOLOGY

Statistical Analysis of Block Coordinate Descent Algorithms for Linear Continuous-time System Identification

Citation for published version (APA):
González, R., Classens, K. H. J., Rojas, C. R., Welsh, J. S., & Oomen, T. A. E. (2024). Statistical Analysis of Block Coordinate Descent Algorithms for Linear Continuous-time System Identification. *IEEE Control Systems Letters*, *8*, 388-393. Article 10500885. https://doi.org/10.1109/LCSYS.2024.3389939

DOI:
10.1109/LCSYS.2024.3389939

Document status and date:
Published: 16/04/2024

Document Version:
Accepted manuscript including changes made at the peer-review stage

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Download date: 22. Mar. 2026

Rodrigo A. González *et al.*, Statistical Analysis of Block Coordinate Descent Algorithms for Linear Continuous-time System Identification. To appear in *IEEE Control Systems Letters (L-CSS)*, 2024, uploaded to Pure on April 13th, 2024

Statistical Analysis of Block Coordinate Descent Algorithms for Linear Continuous-time System Identification

Rodrigo A. González, Koen Classens, Cristian R. Rojas, James S. Welsh, Tom Oomen

***Abstract*— Block coordinate descent is an optimization technique that is used for estimating multi-input single-output (MISO) continuous-time models, as well as single-input single output (SISO) models in additive form. Despite its widespread use in various optimization contexts, the statistical properties of block coordinate descent in continuous-time system identification have not been covered in the literature. The aim of this paper is to formally analyze the bias properties of the block coordinate descent approach for the identification of MISO and additive SISO systems. We characterize the asymptotic bias at each iteration, and provide sufficient conditions for the consistency of the estimator for each identification setting. The theoretical results are supported by simulation examples.**

***Index Terms*— Continuous-time system identification; MISO models; Additive models; Block coordinate descent.**

I. Introduction

System identification deals with the problem of learning dynamical models using input and output data [1]. In this field, a fundamental choice is whether to work with discrete-time or continuous-time models based on the specific application. Continuous-time system identification [2] is advantageous if a requirement is that the model should have physical significance, and its methods can be adjusted to address irregular sampling and the estimation of stiff models [3].

For identifying continuous-time, single-input single-output (SISO) systems, a widely-used method is the simplified refined instrumental variable method (SRIVC, [4]), which iteratively computes instrumental variables estimates of the system parameters. A limitation is that the pseudolinear equation used in the SRIVC estimator cannot handle multiple-input single-output (MISO) models described with transfer functions of distinct denominator polynomials. For such case, [5] (see also [4, Sec. 5.2]) proposes a backfitting algorithm that estimates each submodel keeping all the other submodels fixed, and iterating until convergence. Such algorithm has been applied to model blood glucose [6] and wireless power systems [7], among other applications. Recently, a similar identification method for additive SISO models was introduced in [8]. This additive representation is particularly useful for systems interpreted in modal form, common in many mechanical systems. Unlike the approach presented in [5], in [8], instrumental variable iterations are performed for each submodel until convergence, aligning with the philosophy of the block coordinate descent approach for non-convex optimization [9]. This descent strategy has been also used for optimization over networks [10], and for training Neural ODEs [11].

The statistical analysis of continuous-time system identification methods is essential for their applicability, with substantial progress being made in recent years. The consistency and efficiency of the SRIVC estimator are studied in [12] and [13], and its relation to Maximum Likelihood is studied in [14]. Likewise, [15] extended these findings to closed-loop models. Despite this progress, the block coordinate descent method for MISO and additive SISO identification does not yet have this theoretical foundation. The method for MISO systems is believed to yield consistent estimates upon convergence [5, Section 3.3], though it lacks a formal proof. Similarly, conditions for consistent estimates have not been pursued in [8]. Although Monte Carlo simulations can provide evidence of the performance of these methods, a comprehensive theoretical analysis of their bias properties is essential to offer practical guidelines for their use under general conditions, and may help to design identification experiments that allow for consistent estimators.

In this paper, we analyze the statistical properties of the backfitting methods proposed in [5] and [8] in a unified manner. Our main contributions are outlined below.

- C1 We generalize the algorithms in [5] and [8] by deriving closed-form expressions for the Gauss-Newton (GN) and SRIVC iterations, which, upon convergence for finite sample size, are proven to deliver a stationary point of the cost function under mild conditions.
- C2 We characterize the asymptotic bias obtained at each descent iteration, which shows that consistent estimates can be achieved in one iteration in the MISO setup case.
- C3 We discuss the difference between the MISO and additive SISO setups in terms of the persistence of excitation conditions required for consistent estimates, and we provide sufficient conditions for the generic consistency of the block coordinate descent approach for each setup.

The remainder of this paper is structured as follows. Section II introduces the identification setup. Section III revisits the block coordinate descent method that is used for identifying MISO and additive SISO models. The statistical analysis is addressed in Section IV. Extensive simulations that verify the theoretical results can be found in Section V, while Section VI contains concluding remarks.

This work was supported by the research program VIDI with project number 15698, which is (partly) financed by the Netherlands Organization for Scientific Research (NWO), by the Digital Futures project EXTREMUM, and by the Swedish Research Council under contract number 2016-06079 (NewLEADS). R. A. González, K. Classens and T. Oomen are with the Department of Mechanical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands. C. R. Rojas is with the Division of Decision and Control Systems, KTH Royal Institute of Technology, Stockholm, Sweden. J. S. Welsh is with the School of Engineering, University of Newcastle, NSW, Australia. T. Oomen is also with the Delft Center for Systems and Control, Delft University of Technology, Delft, The Netherlands. E-mail of the corresponding author: r.a.gonzalez@tue.nl.

Notation: Vectors and matrices are written in bold. The Heaviside operator p is defined as $pu(t) = \frac{\mathrm{d}}{\mathrm{d}t}u(t)$, and $G(p)u(t_k)$ means that the continuous-time signal $u(t)$ is filtered through the transfer function $G(p)$ and evaluated at $t = t_k$. If $A(p)$ and $B(p)$ are polynomials of degrees n and m respectively, then $\mathcal{S}(A, B)$ is a $(n+m+1)\times(n+m+1)$ Sylvester matrix (see, e.g., Eq. (A3.9) of [16]).

II. Additive SISO and MISO system setups

Consider the following $K \in \mathbb{N}$ linear and time-invariant, causal, continuous-time systems

$$G_i^*(p) = \frac{B_i^*(p)}{A_i^*(p)}, \quad i = 1, \dots, K. \tag{1}$$

The numerator and denominator polynomials $B_i^*(p)$ and $A_i^*(p)$ are assumed to be coprime, i.e., they do not share roots, and

$$\begin{aligned} A_i^*(p) &= a_{i,n_i}^* p^{n_i} + a_{i,n_i-1}^* p^{n_i-1} + \cdots + a_{i,1}^* p + 1, \\ B_i^*(p) &= b_{i,m_i}^* p^{m_i} + b_{i,m_i-1}^* p^{m_i-1} + \cdots + b_{i,1}^* p + b_{i,0}^*, \end{aligned} \tag{2}$$

with $a_{i,n_i}^* \neq 0$ and $m_i \leq n_i$. The polynomials $A_i^*(p)$ and $B_i^*(p)$ are jointly described by the parameter vector

$$\boldsymbol{\theta}_i^* = \left[a_{i,1}^*,\ a_{i,2}^*,\ \dots,\ a_{i,n_i}^*,\ b_{i,0}^*,\ b_{i,1}^*,\ \dots,\ b_{i,m_i}^*\right]^\top. \tag{3}$$

These systems are excited by inputs $u_i(t)$, $i = 1, \dots, K$, and a noisy measurement of the sum of the outputs is retrieved at every time instant $t = t_k$, $k = 1, \dots, N$, where $\{t_k\}_{k=1}^N$ are assumed to be evenly spaced in time. That is,

$$x(t) = \textstyle\sum_{i=1}^K G_i^*(p)u_i(t) \tag{4}$$

$$y(t_k) = x(t_k) + v(t_k), \tag{5}$$

where $v(t_k)$ is a zero-mean stationary random process of finite variance. Two frameworks are studied in this paper:

1) Continuous-time MISO identification: We assume that the inputs u_i are uncorrelated with each other, and we allow the denominator polynomials A_i to be distinct.
2) Continuous-time additive SISO identification: We assume that all inputs u_i are equal. In addition, we assume that at most one subsystem is biproper.

This paper presents a statistical analysis of block coordinate algorithms to obtain estimates of the joint parameter vector of the systems $G_i^*(p)$, i.e.,

$$\boldsymbol{\beta}^* := \left[\boldsymbol{\theta}_1^{*\top},\ \boldsymbol{\theta}_2^{*\top},\ \dots,\ \boldsymbol{\theta}_K^{*\top}\right]^\top.$$

We consider $\{u_1(t_k), \dots, u_K(t_k), y(t_k)\}_{k=1}^N$ as given. In the next section, we revisit the block coordinate descent algorithm that solves this identification problem for both MISO and additive SISO frameworks.

III. Block coordinate descent approach for continuous-time system identification

Towards the goal of computing an estimator for $\boldsymbol{\beta}^*$, we are interested in solving the following minimization problem:

$$\min_{\substack{\boldsymbol{\theta}_i \in \Omega_i, \\ i=1,\dots,K}} V(\boldsymbol{\beta}), \tag{6}$$

where $\Omega_i \subset \mathbb{R}^{n_i+m_i+1}$ is a compact set where the true parameters of the ith subsystem are assumed to lie, and

$$V(\boldsymbol{\beta}) = \frac{1}{N}\sum_{k=1}^N \left[y(t_k) - \sum_{i=1}^K G_i(p, \boldsymbol{\theta}_i)u_i(t_k)\right]^2. \tag{7}$$

The approach for solving (6) that we investigate consists of iteratively minimizing the cost with respect to $\boldsymbol{\theta}_i$ while leaving the other variables fixed. Algorithm 1 describes the general procedure, known as *block coordinate descent* [9].

Algorithm 1 Block coordinate descent

1: Choose an initial parameter vector $\boldsymbol{\theta}_i^1$ for each i
2: **for** $l = 1, 2, \dots$ **do**
3: **for** $i = 1, \dots, K$ **do**
4: $\boldsymbol{\theta}_i^{l+1} \leftarrow \underset{\boldsymbol{\theta}_i \in \Omega_i}{\arg\min}\, V_N(\boldsymbol{\theta}_1^{l+1}, \dots, \boldsymbol{\theta}_{i-1}^{l+1}, \boldsymbol{\theta}_i, \boldsymbol{\theta}_{i+1}^{l}, \dots, \boldsymbol{\theta}_K^{l})$
5: **end for**
6: **end for**

For each iteration (in l) of Algorithm 1, denote $\boldsymbol{\beta}^l := [(\boldsymbol{\theta}_1^l)^\top, \dots, (\boldsymbol{\theta}_K^l)^\top]^\top$ and the update rule $\boldsymbol{\beta}^{l+1} = \mathcal{A}(\boldsymbol{\beta}^l)$. The following result concerns the convergence of Algorithm 1 to a stationary point of the cost (6).

Theorem 3.1 (Global convergence of Algorithm 1): If the search along any coordinate direction $\boldsymbol{\theta}_i$ yields a unique minimum point of V, then the limit of any convergent subsequence of $\{\boldsymbol{\beta}^l\}$ obtained from $\boldsymbol{\beta}^{l+1} = \mathcal{A}(\boldsymbol{\beta}^l)$ belongs to the set of fixed points $\{\boldsymbol{\beta} \in \prod_{i=1}^K \Omega_i \colon \nabla V(\boldsymbol{\beta}) = \mathbf{0}\}$.

Proof: See, e.g., [9, Section 8.9]. ∎

Remark 3.1: The result in Theorem 3.1 still holds if the unique minimum assumption is replaced with the requirement that V decreases at each coordinate search.

IV. Statistical Analysis

This section addresses the statistical properties of Algorithm 1 applied to the identification of MISO and additive SISO continuous-time systems. We explore Gauss-Newton and refined instrumental variable methods for computing the descent, the parameter bias at each descent iteration, persistence of excitation issues, and statistical consistency. These topics will be covered in the following four subsections.

A. Two alternatives for computing the descent step

The block coordinate descent algorithm described in Algorithm 1 requires a way to compute, at each iteration,

$$\begin{aligned} \boldsymbol{\theta}_i^{l+1} = \underset{\boldsymbol{\theta}_i \in \Omega_i}{\arg\min} \frac{1}{N}\sum_{k=1}^N \Bigg[& y(t_k) - \sum_{j=1}^{i-1} G_j(p, \boldsymbol{\theta}_j^{l+1})u_j(t_k) \\ & - \sum_{j=i+1}^{K} G_j(p, \boldsymbol{\theta}_j^{l})u_j(t_k) - G_i(p, \boldsymbol{\theta}_i)u_i(t_k)\Bigg]^2 \end{aligned} \tag{8}$$

for $i = 1, 2, \dots, K$. For fixed values of $\{\boldsymbol{\theta}_j^{l+1}\}_{j=1}^{i-1}$ and $\{\boldsymbol{\theta}_j^{l}\}_{j=i+1}^{K}$, and for a sufficiently large compact parameter space Ω_i, the optimization problem in (8) reduces to a nonlinear least-squares problem that can be solved via iterations. Indeed, if we define the residual output of each submodel

$$\tilde{y}_i(t_k) := y(t_k) - \sum_{j=1}^{i-1} G_j(p, \boldsymbol{\theta}_j^{l+1})u_j(t_k) - \sum_{j=i+1}^{K} G_j(p, \boldsymbol{\theta}_j^{l})u_j(t_k),$$

then $\boldsymbol{\theta}_i^{l+1}$ must satisfy the first-order optimality condition

$$\frac{1}{N}\sum_{k=1}^{N}\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t_k,\boldsymbol{\theta}_i^{l+1})e(t_k,\boldsymbol{\theta}_i^{l+1})=\mathbf{0}, \tag{9}$$

where the gradient and total residual are denoted as

$$\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t_k,\boldsymbol{\theta}_i^{l+1})=\Bigg[\frac{-pB_i^{l+1}(p)}{[A_i^{l+1}(p)]^2}u_i(t_k),\ldots,\frac{-p^{n_i}B_i^{l+1}(p)}{[A_i^{l+1}(p)]^2}u_i(t_k), \frac{1}{A_i^{l+1}(p)}u_i(t_k),\ldots,\frac{p^{m_i}}{A_i^{l+1}(p)}u_i(t_k)\Bigg]^{\top}, \tag{10}$$

$$e(t_k,\boldsymbol{\theta}_i^{l+1})=\tilde{y}_i(t_k)-\frac{B_i^{l+1}(p)}{A_i^{l+1}(p)}u_i(t_k), \tag{11}$$

with B_i^{l+1}, A_i^{l+1} denoting, respectively, the numerator and denominator polynomials of the ith model evaluated at $\boldsymbol{\theta}_i^{l+1}$. Lemma 4.1 provides two alternatives that are proven to deliver the *same* stationary points of the cost function in (8) at convergence in iterations under mild conditions.

Lemma 4.1 (Gauss-Newton and SRIVC iterations): For an initial model parameter estimate $\boldsymbol{\theta}_{i,0}^{l+1}$ and $s=0,1,2,\ldots$, consider the Gauss-Newton (GN) iterations

$$\boldsymbol{\theta}_{i,s+1}^{l+1}=\left[\frac{1}{N}\sum_{k=1}^{N}\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t_k,\boldsymbol{\theta}_{i,s}^{l+1})\hat{\boldsymbol{\varphi}}_{\mathrm{f}}^{\top}(t_k,\boldsymbol{\theta}_{i,s}^{l+1})\right]^{-1} \times\left[\frac{1}{N}\sum_{k=1}^{N}\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t_k,\boldsymbol{\theta}_{i,s}^{l+1})\left(e(t_k,\boldsymbol{\theta}_{i,s}^{l+1})+\frac{B_{i,s}^{l+1}(p)}{[A_{i,s}^{l+1}(p)]^2}u_i(t_k)\right)\right], \tag{12}$$

and the SRIVC [4] iterations

$$\boldsymbol{\theta}_{i,s+1}^{l+1}=\left[\frac{1}{N}\sum_{k=1}^{N}\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t_k,\boldsymbol{\theta}_{i,s}^{l+1})\boldsymbol{\varphi}_{\mathrm{f}}^{\top}(t_k,\boldsymbol{\theta}_{i,s}^{l+1})\right]^{-1} \times\left[\frac{1}{N}\sum_{k=1}^{N}\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t_k,\boldsymbol{\theta}_{i,s}^{l+1})\tilde{y}_{\mathrm{f}}(t_k,\boldsymbol{\theta}_{i,s}^{l+1})\right], \tag{13}$$

where $\boldsymbol{\varphi}_{\mathrm{f}}$ and filtered residual output $\tilde{y}_{\mathrm{f}}$ are given by

$$\boldsymbol{\varphi}_{\mathrm{f}}(t_k,\boldsymbol{\theta}_{i,s}^{l+1})=\Bigg[\frac{-p}{A_{i,s}^{l+1}(p)}\tilde{y}_i(t_k),\ldots,\frac{-p^{n_i}}{A_{i,s}^{l+1}(p)}\tilde{y}_i(t_k), \frac{1}{A_{i,s}^{l+1}(p)}u_i(t_k),\ldots,\frac{p^{m_i}}{A_{i,s}^{l+1}(p)}u_i(t_k)\Bigg]^{\top}, \tag{14}$$

$$\tilde{y}_{\mathrm{f}}(t_k,\boldsymbol{\theta}_{i,s}^{l+1})=\frac{1}{A_{i,s}^{l+1}(p)}\tilde{y}_i(t_k). \tag{15}$$

If the matrices being inverted in (12) and (13) are non-singular for all large enough integers s, then any converging point (when s tends to infinity) satisfies (9).

Proof: We omit the $(\cdot)_i$ notation on $\boldsymbol{\theta}, A, B$, and u for simplicity, assuming that the expressions refer to the ith submodel. Upon convergence as $s\to\infty$, any converging point $\boldsymbol{\theta}^{l+1}:=\lim_{s\to\infty}\boldsymbol{\theta}_s^{l+1}$ of the GN iterations satisfies

$$\boldsymbol{\theta}^{l+1}=\left[\frac{1}{N}\sum_{k=1}^{N}\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t_k,\boldsymbol{\theta}^{l+1})\hat{\boldsymbol{\varphi}}_{\mathrm{f}}^{\top}(t_k,\boldsymbol{\theta}^{l+1})\right]^{-1} \times\left[\frac{1}{N}\sum_{k=1}^{N}\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t_k,\boldsymbol{\theta}^{l+1})\left(e(t_k,\boldsymbol{\theta}^{l+1})+\frac{B^{l+1}(p)}{[A^{l+1}(p)]^2}u(t_k)\right)\right],$$

which, thanks to the non-singularity assumption of the normal matrix above, is equivalent to

$$\frac{1}{N}\sum_{k=1}^{N}\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t_k,\boldsymbol{\theta}^{l+1})\Bigg(\hat{\boldsymbol{\varphi}}_{\mathrm{f}}^{\top}(t_k,\boldsymbol{\theta}^{l+1})\boldsymbol{\theta}^{l+1} - e(t_k,\boldsymbol{\theta}^{l+1})-\frac{B^{l+1}(p)}{[A^{l+1}(p)]^2}u(t_k)\Bigg)=\mathbf{0}. \tag{16}$$

The gradient vector $\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t_k,\boldsymbol{\theta}^{l+1})$ can be shown to satisfy $\hat{\boldsymbol{\varphi}}_{\mathrm{f}}^{\top}(t_k,\boldsymbol{\theta}^{l+1})\boldsymbol{\theta}^{l+1}=\frac{B^{l+1}}{(}p)[A^{l+1}(p)]^2u(t_k)$. Replacing this expression in (16) leads to (9), which is what we aimed to prove. For the SRIVC iterations, the expression (13) reduces to the following condition when $\boldsymbol{\theta}_s^{l+1}=\boldsymbol{\theta}_{s+1}^{l+1}=\boldsymbol{\theta}^{l+1}$:

$$\frac{1}{N}\sum_{k=1}^{N}\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t_k,\boldsymbol{\theta}^{l+1})\big[\tilde{y}_{\mathrm{f}}(t_k,\boldsymbol{\theta}^{l+1})-\boldsymbol{\varphi}_{\mathrm{f}}^{\top}(t_k,\boldsymbol{\theta}^{l+1})\boldsymbol{\theta}^{l+1}\big]=\mathbf{0}.$$

However, from the definitions of $\boldsymbol{\varphi}_{\mathrm{f}}$ and $\tilde{y}_{\mathrm{f}}$ in (14) and (15),

$$\tilde{y}_{\mathrm{f}}(t_k,\boldsymbol{\theta}^{l+1})-\boldsymbol{\varphi}_{\mathrm{f}}^{\top}(t_k,\boldsymbol{\theta}^{l+1})\boldsymbol{\theta}^{l+1}=e(t_k,\boldsymbol{\theta}^{l+1}). \tag{17}$$

Substituting this result above also leads to (9). ■

In the literature of both the MISO [5] and additive SISO [8] methods, only SRIVC iterations have been considered. Lemma 4.1 derives the Gauss-Newton iterations that also lead to the same stationary points as the SRIVC iterations. While the SRIVC iterations have been known to be numerically robust against poor initialization [2], the iterations depend on the specification of the intersample behavior of the output, which is not known beforehand, since only output samples are measured. The output intersample behavior can influence the convergence rate of the iterations, but it does not affect the converging point of the algorithm [12].

B. Bias analysis at a fixed descent iteration

For the following analysis, we assume that we compute the GN (12) or SRIVC iterations (13) until convergence, and study the converging point of the ith submodel for a *fixed* descent iteration l and large sample size. The theoretical analysis requires the following assumptions:

Assumption 4.1 (Persistence of excitation): Each input $u_i(t_k)$ is persistently exciting of order no less than $2n_i$.

Assumption 4.2 (Stationarity): The inputs $u_i(t_k)$ and disturbance $v(t_s)$ are stationary, and each input is independent of the disturbance for all integers k and s.

Assumption 4.3 (Stability and coprimeness): For every integer s, the model estimate $B_{i,s}^{l+1}(p)/A_{i,s}^{l+1}(p)$ is stable, and the numerator and denominator polynomials are coprime.

Assumption 4.4 (Non-singularity): The normal matrix of the GN iterations, $\mathbb{E}\{\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t_k,\boldsymbol{\theta}_i^{l+1})\hat{\boldsymbol{\varphi}}_{\mathrm{f}}^{\top}(t_k,\boldsymbol{\theta}_i^{l+1})\}$, or the modified normal matrix of the SRIVC iterations, $\mathbb{E}\{\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t_k,\boldsymbol{\theta}_i^{l+1})\boldsymbol{\varphi}_{\mathrm{f}}^{\top}(t_k,\boldsymbol{\theta}_i^{l+1})\}$, are generically non-singular with respect to all system and model denominators.

Assumptions 4.1 to 4.3 are standard in the literature on statistical properties of the SRIVC method for open and closed loop [12], [15], and extensive practical experience indicates that instabilities rarely arise, but they may occur if the data is poor [17]. Assumption 4.4 for the SRIVC iterations can be posed as a 2-norm condition on the interpolation error of the

output intersample behavior [12]. Since this condition holds for most cases of interest, Assumption 4.4 is not restrictive.

Theorem 4.2: Consider the block coordinate descent method in Algorithm 1, where V is computed using (7), and the coordinate minimization is performed using the GN or SRIVC iterations in Lemma 4.1. Suppose Assumptions 4.1 to 4.4 hold, and define $\boldsymbol{\eta}_j \in \mathbb{R}^{n_j+m_j+1}$ as the vector that contains the coefficients of $A_j^*(p)\bar{B}_j(p) - \bar{A}_j(p)B_j^*(p)$ in descending order of degree, where

$$[\bar{A}_j(p), \bar{B}_j(p)] = \begin{cases} [A_j^{l+1}(p), B_j^{l+1}(p)], & j = 1, \ldots, i \\ [A_j^{l}(p), B_j^{l}(p)], & j = i+1, \ldots, K. \end{cases} \quad (18)$$

If the GN or SRIVC iterations of the ith model converge (for a fixed descent iteration l) for all N sufficiently large, then the converging point of the ith submodel satisfies

$$\sum_{j=1}^{K} \mathbb{E}\left\{ \frac{1}{\bar{A}_i^2(p)} \mathbf{u}_i(t_k) \frac{1}{A_j^*(p)\bar{A}_j(p)} \mathbf{u}_j^\top(t_k) \right\} \boldsymbol{\eta}_j = \mathbf{0}, \quad (19)$$

where $\mathbf{u}_j(t_k)$ $(j = 1, \ldots, K)$ is given by

$$\mathbf{u}_j(t_k) = \begin{bmatrix} p^{n_j+m_j}, & p^{n_j+m_j-1}, & \ldots, & 1 \end{bmatrix}^\top u_j(t_k). \quad (20)$$

Proof: Define $\boldsymbol{\theta}_i^{l+1} := \lim_{s\to\infty} \boldsymbol{\theta}_{i,s}^{l+1}$, where $\boldsymbol{\theta}_{i,s}^{l+1}$ is computed from the GN (12) or SRIVC (13) iterations. For a fixed sample size N, the converging point $\boldsymbol{\theta}_i^{l+1}$ must satisfy (16) in the case of the GN iterations, or

$$\left[\frac{1}{N} \sum_{k=1}^{N} \hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t_k, \boldsymbol{\theta}_i^{l+1}) \boldsymbol{\varphi}_{\mathrm{f}}^\top(t_k, \boldsymbol{\theta}_i^{l+1}) \right]^{-1} \quad (21)$$

$$\times \left[\frac{1}{N} \sum_{k=1}^{N} \hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t_k, \boldsymbol{\theta}_i^{l+1}) \big(\tilde{y}_{\mathrm{f}}(t_k, \boldsymbol{\theta}_i^{l+1}) - \boldsymbol{\varphi}_{\mathrm{f}}^\top(t_k, \boldsymbol{\theta}_{i,s}^{l+1}) \boldsymbol{\theta}_i^{l+1} \big) \right] = \mathbf{0}$$

for the SRIVC iterations. Recall that the expressions in parentheses in (16) and in the second sum in (21) are simply $e(t_k, \boldsymbol{\theta}_i^{l+1})$, which follows from the proof of Lemma 4.1. Under Assumption 4.2, the sums in (21) converge to their expected values as $N \to \infty$ [18]. Thus, as $N \to \infty$,

$$\mathbb{E}\{\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t_k, \boldsymbol{\theta}_i^{l+1}) \hat{\boldsymbol{\varphi}}_{\mathrm{f}}^\top(t_k, \boldsymbol{\theta}_i^{l+1})\}^{-1} \mathbb{E}\{\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t_k, \boldsymbol{\theta}_i^{l+1}) e(t_k, \boldsymbol{\theta}_i^{l+1})\} = \mathbf{0}$$

for the GN iterations, and

$$\mathbb{E}\{\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t_k, \boldsymbol{\theta}_i^{l+1}) \boldsymbol{\varphi}_{\mathrm{f}}^\top(t_k, \boldsymbol{\theta}_i^{l+1})\}^{-1} \mathbb{E}\{\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t_k, \boldsymbol{\theta}_i^{l+1}) e(t_k, \boldsymbol{\theta}_i^{l+1})\} = \mathbf{0}$$

for the SRIVC iterations. Either way, under Assumption 4.4, the converging point must (generically) satisfy

$$\mathbb{E}\left\{ \hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t_k, \boldsymbol{\theta}_i^{l+1}) e(t_k, \boldsymbol{\theta}_i^{l+1}) \right\} = \mathbf{0}. \quad (22)$$

The instrument vector $\hat{\boldsymbol{\varphi}}_{\mathrm{f}}$ can be written as

$$\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t_k, \boldsymbol{\theta}_i^{l+1}) = \mathcal{S}(-B_i^{l+1}, A_i^{l+1}) \frac{1}{[A_i^{l+1}(p)]^2} \mathbf{u}_i(t_k), \quad (23)$$

where B_i^{l+1}, A_i^{l+1} denote the ith model polynomials evaluated at $\boldsymbol{\theta}_i^{l+1}$, $\mathbf{u}_i(t_k)$ has the form in (20), and the Sylvester matrix $\mathcal{S}(-B_i^{l+1}, A_i^{l+1})$ is non-singular thanks to B_i^{l+1} and A_i^{l+1} being coprime [16, Lemma A3.1]. On the other hand, $e(t_k, \boldsymbol{\theta}_i^{l+1})$ can be written as

$$e(t_k, \boldsymbol{\theta}_i^{l+1}) = \sum_{j=1}^{K} \left[\frac{B_j^*(p)}{A_j^*(p)} - \frac{\bar{B}_j(p)}{\bar{A}_j(p)} \right] u_j(t_k) + v(t_k), \quad (24)$$

where we have used the notation in (18) for defining $\bar{B}_j(p)/\bar{A}_j(p)$. By exploiting these alternative descriptions for $\hat{\boldsymbol{\varphi}}_{\mathrm{f}}$ and e, we can write the condition (22) as

$$\sum_{j=1}^{K} \mathbb{E}\left\{ \frac{1}{\bar{A}_i^2(p)} \mathbf{u}_i(t_k) \left[\frac{B_j^*(p)}{A_j^*(p)} - \frac{\bar{B}_j(p)}{\bar{A}_j(p)} \right] u_j(t_k) \right\} = \mathbb{E}\left\{ \frac{1}{\bar{A}_i^2(p)} \mathbf{u}_i(t_k) v(t_k) \right\} = \mathbf{0},$$

where the last equality holds because the signals u and v are uncorrelated due to Assumption 4.2. Thus, (22) reduces to

$$\sum_{j=1}^{K} \mathbb{E}\left\{ \frac{1}{\bar{A}_i^2(p)} \mathbf{u}_i(t_k) \left[\frac{B_j^*(p)}{A_j^*(p)} - \frac{\bar{B}_j(p)}{\bar{A}_j(p)} \right] u_j(t_k) \right\} = \mathbf{0}.$$

Finally, by rearranging the residual terms as

$$\left[\frac{B_j^*(p)}{A_j^*(p)} - \frac{\bar{B}_j(p)}{\bar{A}_j(p)} \right] u_j(t_k) = \frac{1}{A_j^*(p)\bar{A}_j(p)} \mathbf{u}_j^\top(t_k) \boldsymbol{\eta}_j,$$

the converging point $\boldsymbol{\theta}_i^{l+1}$ must satisfy (19). ■

Corollary 4.1: Under the same assumptions of Theorem 4.2, for the MISO system case, the converging point of the ith model for a fixed descent iteration as N tends to infinity generically satisfies $\bar{B}_i(p)/\bar{A}_i(p) = B_i^*(p)/A_i^*(p)$.

Proof: All expectations in (19) are zero except for when $j = i$, due to the assumption that the inputs u_i are uncorrelated. Thus, the expectation gives a matrix that is generically non-singular due to the genericity argument of Theorem 1 of [12]. We reach $\boldsymbol{\eta}_i = \mathbf{0}$, which occurs if and only if $\bar{B}_i(p)/\bar{A}_i(p) = B_i^*(p)/A_i^*(p)$. ■

Theorem 4.2 reveals that, at a fixed descent iteration l, the block coordinate descent algorithm for additive SISO models will lead to biased estimates in general, and more bias is expected if the other submodels deviate further from their true subsystem. On the other hand, Corollary 4.1 reveals that consistent estimates can be obtained in the MISO framework by a single descent iteration with GN or SRIVC iterations performed until convergence. Such approach is similar but not equivalent to the one in [5], which performs one SRIVC iteration for each submodel, and descent iterations until convergence. Descent is not guaranteed with only one SRIVC iteration, which suggests that the method in [5] can benefit from multiple GN or SRIVC iterations for each submodel.

C. Persistence of excitation of the input

Theorem 4.2 requires a persistence of excitation order of $2 \max_i n_i$ for the additive SISO case. However, such condition is not enough to avoid identifiability issues when trying to estimate each submodel separately. An example of this identifiability problem is detailed below.

Example 4.1: For the additive SISO case, let $G_1^*(p)$ and $G_2^*(p)$ be two strictly causal, asymptotically stable, first-order subsystems. If the input is a sampled sinusoid signal (with ZOH intersample behavior and sampling period h) of the form $u(t_k) = \sin(\omega k h)$, with $|\omega| < \pi/h$, then, as $N \to \infty$,

the GN or SRIVC iterations for estimating $G_1^*(p)$ lead to the following interpolation conditions:

$$G_{\mathrm{d},1}(e^{\pm \mathrm{i}\omega h}, \boldsymbol{\theta}_1^{l+1}) = G_{\mathrm{d},1}^*(e^{\pm \mathrm{i}\omega h}) + G_{\mathrm{d},2}^*(e^{\pm \mathrm{i}\omega h}) - G_{\mathrm{d},2}(e^{\pm \mathrm{i}\omega h}, \boldsymbol{\theta}_2^{l}),$$

where $G_{\mathrm{d},i}$ denotes the ZOH discrete-time equivalent of the continuous-time transfer function G_i. On the other hand, the GN or SRIVC iterations for estimating $G_2^*(p)$ converge to

$$G_{\mathrm{d},2}(e^{\pm \mathrm{i}\omega h}, \boldsymbol{\theta}_2^{l+1}) = G_{\mathrm{d},1}^*(e^{\pm \mathrm{i}\omega h}) + G_{\mathrm{d},2}^*(e^{\pm \mathrm{i}\omega h}) - G_{\mathrm{d},1}(e^{\pm \mathrm{i}\omega h}, \boldsymbol{\theta}_1^{l+1}).$$

These equations indicate that $G_{\mathrm{d},2}(e^{\pm \mathrm{i}\omega h}, \boldsymbol{\theta}_2^{l+1}) = G_{\mathrm{d},2}(e^{\pm \mathrm{i}\omega h}, \boldsymbol{\theta}_2^{l})$, i.e., $\boldsymbol{\theta}_2$ converges in one descent iteration. Thus, the parameters $\boldsymbol{\theta}_1^{l+1}$ and $\boldsymbol{\theta}_2^{l+1}$ must satisfy the underdetermined system of equations

$$G_{\mathrm{d},1}(e^{\pm \mathrm{i}\omega h}, \boldsymbol{\theta}_1^{l+1}) + G_{\mathrm{d},2}(e^{\pm \mathrm{i}\omega h}, \boldsymbol{\theta}_2^{l+1}) = G_{\mathrm{d},1}^*(e^{\pm \mathrm{i}\omega h}) + G_{\mathrm{d},2}^*(e^{\pm \mathrm{i}\omega h}),$$

which is the expected result for identification with a single sinusoidal input. Since an infinite number of solutions exist for the system above, the parameters are not identifiable. $\triangle$

Remark 4.1: A similar observation to Example 4.1 can be derived for *continuous-time* multisine inputs. In such case, for $u(t) = \sin(\omega_1 t), t \geq 0$, a tailored version of the SRIVC iterations suited for continuous-time multisines [19] results in $G_1(\pm \mathrm{i}\omega_1, \boldsymbol{\theta}_1^{l+1}) + G_2(\pm \mathrm{i}\omega_1, \boldsymbol{\theta}_2^{l+1}) = G_1^*(\pm \mathrm{i}\omega_1) + G_2^*(\pm \mathrm{i}\omega_1)$, which again indicates a problem of identifiability.

Towards the goal of solving this identifiability problem, the next subsection studies sufficient conditions for consistency of the coordinate descent method for both MISO and additive SISO identification.

D. Consistency analysis: infinite descent iterations

We first must investigate the nonsingularity of a certain matrix formed by filtered versions of the input signal.

Lemma 4.3: Let the input $u(t_k)$ be persistently exciting of order no less than $2\sum_{i=1}^{K} n_i$, and assume that the sampling frequency is more than twice that of the largest imaginary part of the zeros of $\prod_{j=1}^{K} \bar{A}_j(p) A_j^*(p)$. Moreover, suppose Assumptions 4.1 to 4.3 hold. Then, the matrix

$$\boldsymbol{\Phi} = \mathbb{E}\left\{ \begin{bmatrix} \frac{1}{A_1^2(p)}\mathbf{u}_1(t_k) \\ \frac{1}{A_2^2(p)}\mathbf{u}_2(t_k) \\ \vdots \\ \frac{1}{A_K^2(p)}\mathbf{u}_K(t_k) \end{bmatrix} \begin{bmatrix} \frac{1}{A_1^*(p)\bar{A}_1(p)}\mathbf{u}_1(t_k) \\ \frac{1}{A_2^*(p)\bar{A}_2(p)}\mathbf{u}_2(t_k) \\ \vdots \\ \frac{1}{A_K^*(p)\bar{A}_K(p)}\mathbf{u}_K(t_k) \end{bmatrix}^{\top} \right\}, \quad (25)$$

with $\mathbf{u}_j(t_k)$ defined in (20), is generically non-singular with respect to the system and model denominator polynomials.

Proof: Direct from Lemma 3 of [20] when, following their notation, we set $S(q) = 1$ and $z(t_k) = u(t_k)$. ■

Theorem 4.4 shows that the block coordinate descent method possesses the consistency property in both MISO and additive SISO setups, provided the persistence of excitation order of the input in the additive SISO case is at least twice the number of poles of the aggregated model.

Theorem 4.4: Consider Algorithm 1, and suppose that Assumptions 4.1 to 4.3 hold. In addition, assume that the input $u(t_k)$ is persistently exciting of order no less than $2\sum_{i=1}^{K} n_i$ for the additive SISO setup, and that the sampling frequency is more than twice that of the largest imaginary part of the zeros of $\prod_{j=1}^{K} \bar{A}_j(p) A_j^*(p)$. If the GN or SRIVC iterations are performed until convergence, and the descent iterations $l \to \infty$, then the true parameters $\{\boldsymbol{\theta}_i^*\}_{i=1}^{K}$ are the unique converging point (generically) as $N \to \infty$.

Proof: As the number of GN or SRIVC iterations tends to infinity, the model parameters satisfy (19). At convergence in descent iterations, the condition $\boldsymbol{\Phi}\boldsymbol{\eta} = \mathbf{0}$ is obtained, where $\boldsymbol{\Phi}$ is given by (25) and $\boldsymbol{\eta} := [\boldsymbol{\eta}_1^{\top}, \ldots, \boldsymbol{\eta}_K^{\top}]^{\top}$ is the joint bias vector. Lemma 4.3 shows that $\boldsymbol{\Phi}$ is generically non-singular, which implies that $\boldsymbol{\eta} = \mathbf{0}$ generically. This expression means that $\bar{B}_i(p)/\bar{A}_i(p) = B_i^*(p)/A_i^*(p)$ for all $i = 1, \ldots, K$, which proves the statement. ■

Remark 4.2: There exists a parameter identifiability issue in the additive SISO setup when a pair of subsystems shares the same model structure. Swapping such pair results in an identical additive model, altering the arrangement of parameters within $\boldsymbol{\beta}$. This property, however, does not impact the consistency of the additive model estimator as a whole.

V. Simulations

In this section, we will verify the theoretical results via Monte Carlo simulations. We consider the following systems:

$$G_1^*(p) = \frac{2}{0.25p^2 + 0.25p + 1}, \quad G_2^*(p) = \frac{1}{0.025p^2 + 0.01p + 1}.$$

The output is contaminated by white noise of variance 0.25, and is sampled every 0.02[s]. Each setup considers unit-variance Gaussian white noise inputs, both resulting in an output signal-to-noise ratio of 6.6 [dB]. The additive SISO and MISO estimators consider SRIVC iterations; GN iterations have shown no significant performance differences at convergence in iterations, in agreement with Lemma 4.1.

A. Bias at a fixed descent iteration

To verify the results in Theorem 4.2, we perform 300 Monte Carlo runs with varying input and noise realizations for 30 different sample sizes, ranging logarithmically from $N = 2000$ to $N = 50000$. On each run, one descent iteration is performed on $G_2(p)$ for a fixed model estimate $\bar{G}_1(p) = \bar{B}_1(p)/\bar{A}_1(p)$. The additive SISO and MISO setups are tested under two different model estimates for $\bar{G}_1(p)$:

$$\bar{G}_1^{\mathrm{a}}(p) = \frac{2.2}{0.2p^2 + 0.2p + 1}, \quad \bar{G}_1^{\mathrm{b}}(p) = \frac{1.7}{0.1p^2 + 0.27p + 1}.$$

The 2-norm errors of each initial model for $G_1^*(p)$ are $\|\bar{G}_1^{\mathrm{a}} - G_1^*\|_2 = 1.25$ and $\|\bar{G}_1^{\mathrm{b}} - G_1^*\|_2 = 2.11$. The empirical means of each parameter estimate of $G_2^*(p)$ are shown in Figure 1. As expected from Corollary 4.1, the MISO estimates converge to the true values for increasing sample size, independently of the initial estimate of $G_1^*(p)$. In contrast, the coordinate descent step for the additive SISO setup suffers from bias in the parameter estimates. Since $\bar{G}_1^{\mathrm{b}}(p)$ is a more inaccurate estimate of $G_1^*(p)$ compared to $\bar{G}_1^{\mathrm{a}}(p)$, its associated bias vector $\boldsymbol{\eta}_1$ in Theorem 4.2 is larger in magnitude, implying that $\boldsymbol{\eta}_2$ must also be large for (19) to be satisfied. The sample means for $G_2(p)$ are further away from their true values if $\bar{G}_1^{\mathrm{b}}(p)$ is used as an estimate of $G_1^*(p)$, which is in accordance to this insight.

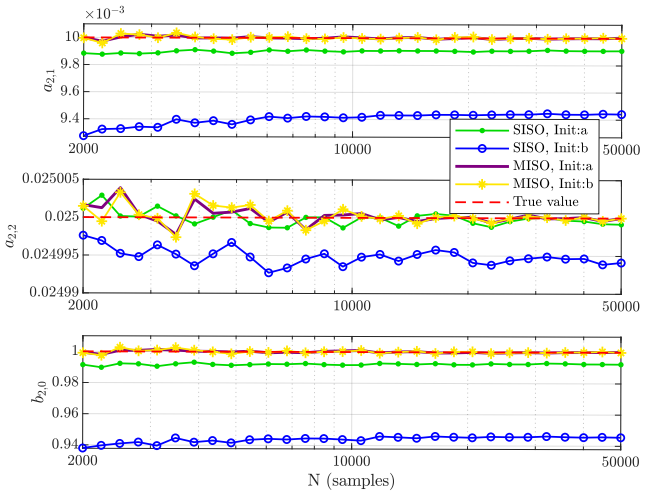


Fig. 1. Empirical mean of the estimate of each parameter of $G_2^*(p)$ for one descent iteration. The true values are given by $a_{2,1}^* = 0.01$, $a_{2,2}^* = 0.025$, and $b_{2,0}^* = 1$. The estimates of the MISO model setup converge to the true value, while bias is observed in for the additive SISO setup.

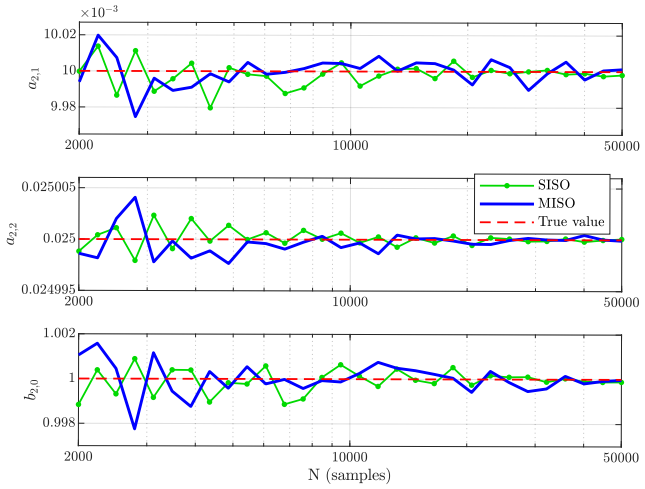


Fig. 2. Empirical mean of the estimate of each parameter of $G_2^*(p)$ using the full block coordinate descent method. The true values are given by $a_{2,1}^* = 0.01$, $a_{2,2}^* = 0.025$, and $b_{2,0}^* = 1$. Both the MISO and additive SISO estimators converge to the true values as the sample size grows.

B. Consistency

The consistency of the block coordinate descent estimator for additive SISO models is studied under a similar experimental condition to Section V-A. This time, the estimators are both initialized with $\bar{G}_1^{\mathrm{b}}(p)$, and ten SRIVC iterations are performed for each descent step. A maximum of 30 descent iterations are used for each estimate, and the iterations also may terminate if the relative error of the parameter vector with respect to its previous value is below 10^{-10}. The estimated parameters for $G_2^*(p)$ are analyzed. To avoid exchangeability issues for the additive SISO setup, we study the submodel parameters that are closest to $G_2^*(p)$ according to the 2-norm of the parameter error.

The results in Fig. 2 show that both MISO and additive SISO setups achieve consistent estimates of $G_2^*(p)$. This result is in accordance to Theorem 4.4. Note that the persistence of excitation is not an issue in this test, since the inputs are designed to be persistently exciting of any order.

VI. Conclusions

In this paper we have analyzed the statistical properties of the block coordinate descent method applied to the identification of MISO and additive SISO systems in open loop. The main difference between these setups is the nature of the input excitation, which reveals different bias properties and conditions for consistency. We have characterized the bias after a single descent for additive SISO systems, and we have used this insight to prove the generic consistency of the method for MISO and additive SISO setups in a unified manner. The results hold for both Gauss-Newton and SRIVC iterations. The numerical experiments here presented have verified the theoretical results.

References

[1] L. Ljung, *System Identification: Theory for the User,* 2nd Ed. Prentice-Hall, 1999.

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

[3] H. Garnier and P. C. Young, "The advantages of directly identifying continuous-time transfer function models in practical applications," *Int. J. of Control*, vol. 87, no. 7, pp. 1319–1338, 2014.

[4] P. C. Young and A. J. Jakeman, "Refined instrumental variable methods of recursive time-series analysis. Part III, Extensions," *Int. J. of Control*, vol. 31, no. 4, pp. 741–764, 1980.

[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 Eng. Pract.*, pp. 471–486, 2007.

[6] H. Kirchsteiger, S. Pölzer, R. Johansson, E. Renard, and L. del Re, "Direct continuous time system identification of MISO transfer function models applied to type 1 diabetes," in *50th IEEE Conference on Decision and Control*, pp. 5176–5181, 2011.

[7] F. Chen, P. C. Young, H. Garnier, Q. Deng, and M. K. Kazimierczuk, "Data-driven modeling of wireless power transfer systems with multiple transmitters," *IEEE Tran. on P. Electronics*, vol. 35, no. 11, pp. 11363–11379, 2020.

[8] R. A. González, C. R. Rojas, S. Pan, and J. Welsh, "Parsimonious identification of continuous-time systems: A block-coordinate descent approach," in *22nd IFAC World Congress (IFAC 2023)*, 2023.

[9] D. Luenberger and Y. Ye, *Linear and Nonlinear Programming,* 3rd Ed. Springer, 2008.

[10] I. Necoara, Y. Nesterov, and F. Glineur, "Random block coordinate descent methods for linearly constrained optimization over networks," *J. of Opt. Theory and Applications*, vol. 173, pp. 227–254, 2017.

[11] A. Quaglino, M. Gallieri, J. Masci, and J. Koutník, "SNODE: Spectral discretization of neural ODEs for system identification," *International Conference on Learning Representations*, 2020.

[12] 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*, vol. 113, Art. 108767, 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*, vol. 121, Art. 109196, 2020.

[14] P. C. Young, "Refined instrumental variable estimation: Maximum Likelihood optimization of a unified Box–Jenkins model," *Automatica*, vol. 52, pp. 35–46, 2015.

[15] R. A. González, C. R. Rojas, S. Pan, and J. S. Welsh, "Consistency analysis of refined instrumental variable methods for continuous-time system identification in closed-loop," *Automatica (in Press)*.

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

[17] P. C. Young, *Recursive Estimation and Time-Series Analysis: An Introduction for the Student and Practitioner,* 2nd Ed. Springer, 2012.

[18] T. Söderström, "Ergodicity results for sample covariances," *Problems of Control and Information Theory*, vol. 4, no. 2, pp. 131–138, 1975.

[19] R. A. González, C. R. Rojas, S. Pan, and J. S. Welsh, "Consistent identification of continuous-time systems under multisine input signal excitation," *Automatica*, vol. 133, Art. 109859, 2020.

[20] R. A. González, K. Classens, C. R. Rojas, J. Welsh, and T. Oomen, "Identification of additive continuous-time systems in open and closed-loop," *arXiv preprint arXiv:2401.01263*, 2024.