



KTH ROYAL INSTITUTE
OF TECHNOLOGY

Doctoral Thesis in Electrical Engineering

Continuous-time System Identification: Refined Instrumental Variables and Sampling Assumptions

RODRIGO A. GONZÁLEZ

Continuous-time System Identification: Refined Instrumental Variables and Sampling Assumptions

RODRIGO A. GONZÁLEZ

Academic Dissertation which, with due permission of the KTH Royal Institute of Technology,
is submitted for public defence for the Degree of Doctor of Philosophy on Thursday the 5th May
2022, at 2:00 p.m. in F3, Lindstedtsvägen 26, Stockholm

Doctoral Thesis in Electrical Engineering
KTH Royal Institute of Technology
Stockholm, Sweden 2022

© Rodrigo A. González

ISBN: 978-91-8040-186-9
TRITA-EECS-AVL-2022:23

Printed by: Universitetsservice US-AB, Sweden 2022

Abstract

Continuous-time system identification deals with the problem of building continuous-time models of dynamical systems from sampled input and output data. There are two main approaches in this field: indirect and direct. In the indirect approach, a suitable discrete-time model is first determined, and then it is transformed into continuous-time. On the other hand, the direct approach obtains a continuous-time model from the sampled data without an intermediate discrete-time model. In both approaches there exists a dichotomy between discrete-time data and continuous-time models, which can induce robustness issues and complications in the theoretical analysis of identification methods. These difficulties are addressed in this thesis.

First, we consider the indirect approach to continuous-time system identification. For a zero-order hold sampling mechanism, this approach usually leads to an excess of model zeros when the true system has a relative degree greater than one. Inspired by the indirect prediction error method, we propose an indirect-approach estimator that guarantees stability in the model and enforces the desired number of poles and zeros in the continuous-time transfer function estimate.

The second part of this thesis concerns the asymptotic properties and extensions of direct continuous-time identification methods. We provide a comprehensive statistical analysis of the simplified refined instrumental variable method for continuous-time systems (SRIVC), which is a widely-used direct identification algorithm that applies an adaptive prefiltering to the sampled input and output data. We prove that the SRIVC estimator is generically consistent and asymptotically efficient under some mild conditions when taking into account the intersample behavior of the signals in the analysis, and we give conditions under which these statistical properties are not achieved. An extended analysis is provided for when the model is over-parameterized. Later, we propose and analyze the statistical properties of an extension of the SRIVC estimator that can deal with input signals that cannot be interpolated exactly via hold reconstructions. The standard SRIVC estimator and its extension for arbitrary inputs, together with other refined instrumental variable methods, are also investigated in closed-loop settings and are further enhanced to deal with the identification of unstable systems.

The last part of this thesis focuses on the analysis and identification of continuous-time systems subject to band-limited input excitations. The non-causal behavior of the band-limited discrete-time equivalent system is studied in detail, and the findings are later used for designing novel non-parametric and parametric identification methods for when the input is band-limited. Special treatment is given to identification with continuous-time multisine inputs. For that case, we investigate fundamental relations between prediction error methods, optimal refined instrumental variables, and interpolation and approximation of frequency response function estimates.

All of the methods and theoretical results are accompanied by extensive simulation tests that verify our findings.

Sammanfattning

Tidskontinuerlig systemidentifiering går ut på att bygga tidskontinuerliga modeller av dynamiska system genom att använda samplad data. Det finns två huvudsakliga tillvägagångssätt inom detta forskningsområde: indirekta och direkta. Med det indirekta tillvägagångssättet så hittar man först en lämplig tidsdiskret modell, för att sedan transformera den till kontinuerlig tid. Det direktta tillvägagångssättet, å andra sidan, identifierar en tidskontinuerlig modell från samplad data utan att först hitta en tidsdiskret modell som ett mellanliggande steg. Med båda tillvägagångssätten så finns en dikotomi mellan tidsdiskret data och tidskontinuerliga modeller, vilket kan skapa problem med robusthet samt komplikationer med den teoretiska analysen av identifieringsmetoder. Dessa svårigheter behandlas i denna avhandling.

Vi börjar med att betrakta det indirekta tillvägagångssättet för identifiering av tidskontinuerliga system. För en samplingsmekanisk med en nollte ordningens hållkrets så leder detta tillvägagångssätt vanligtvis till ett överskott av nollställen i modellen då det sanna systemet har relativt gradtal större än ett. Vi föreslår en skattare, inspirerad av den indirekta prediktionsfelmetoden, som med det indirekta tillvägagångssättet garanterar modellens stabilitet och framtvingar önskat antal poler och nollställen i skattningen av den tidskontinuerliga överföringsfunktionen.

En viktig del av denna avhandling berör direkta metoder för tidskontinuerlig systemidentifiering. Vi inkluderar en omfattande statistisk analys den förenklade raffinerade instrumentvariabelmetoden för tidskontinuerliga system (SRIVC), som är en direkt identifieringsalgoritm som används i stor utsträckning och som appliceras ett adaptivt förfilter till den samplade in- och utsignalen. Vi bevisar att SRIVC-skattaren är generiskt konsistent och asymptotiskt effektiv, under vissa milda antaganden. Vi presenterar även villkor under vilka detta antagande inte är uppfyllt. En utökad analys är inkluderad för fallet då modellen är överparametrerad. Senare så föreslår och analyserar vi även de statistiska egenskaperna av en utökning av SRIVC-skattaren som kan hantera insignaler som inte kan interpoleras exakt via rekonstruktioner som använder sig av hållkretsar. Den typiska SRIVC-skattaren och dess utökning för godtyckliga insignaler, tillsammans med andra raffinerade instrumentvariabelmetoder, undersöks också i fall då systemet är återkopplat, samt förbättras ytterligare för att hantera identifiering av instabila system.

Avhandlingens sista del fokuserar på analys och identifiering av tidskontinuerliga system med begränsad bandbredd av insignalen. Det icke-kausala beteendet av det ekvivalenta tidsdiskreta systemet med begränsad bandbredd undersöks i detalj, och resultaten används senare för att designa både nya icke-parametriska och parametriska identifieringsmetoder för fall då insignalen har begränsad bandbredd. Särskilt behandlas identifiering av tidskontinuerliga system då insignalen är summan av flera sinusvågor. För det fallet så undersöker vi fundamentala förhållanden mellan prediktionsfelmetoder, optimala raffinerade instrumentvariabelmetoder, samt interpolering och approximation av skattningar av frekvensfunktionen.

Alla metoder och teoretiska resultat presenteras tillsammans med omfattande simulerade test som verifierar våra fynd.

Acknowledgements

After more than four years of Ph.D. work, there are many people and moments for which one should be grateful. First of all, this thesis would have not been possible without the guidance and help from my supervisor Cristian Rojas. Cristian has supported me ever since I started my MSc. thesis in Chile back in 2016, and has given me the freedom to explore most of the topics that can be found in this monograph. I would also like to thank my co-supervisor Bo Wahlberg for the interesting discussions and lunches, and Christer Magnusson and Felicia Gustafsson for their support in daily Ph.D. endeavors.

The years I have spent as a member of the Division of Decision and Control at KTH have been full of pleasant memories thanks to many people. I want to deeply thank the SYSID group for all the numerous fikas, lunches, afterworks, workshops, and conferences we have enjoyed together. Special thanks to Matías Müller for supporting me since day one of my Ph.D. (and even much before that!), and for being such a great long-time friend. Thanks to Mina Ferizbegovic for always offering a helping hand whenever I needed it the most. Thanks to Robert Mattila and Mohamed Abdalmoaty for all the lunches we have shared, and for the discussions we have had about life, future, and personal growth. Thanks to Inês Lourenço for sharing your joy and charisma with everyone. Thanks to Robert Marczuk for sharing your contagious energy and for helping me with the Sammanfattning, and Javad Parsa for your kindness. I also want to thank Peter Varnai for his enthusiasm for being involved in fun initiatives with me. My gratitude also goes to my colleagues in the office: Fei Chen for the nice discussions about the Netherlands and NBA, Yuchao Li for sharing his enthusiasm about math and probability with me, Joana Fonseca for inspiring me to eat healthier, and Braghadeesh Lakshminarayanan for the weekly fikas we have had together. I would like to express my gratitude to all my colleagues at KTH, including Kaito Ariu, Erik Berglund, Manne Held, Yassir Jedra, Sarit Khirirat, Jacob Lindbäck, Hanxiao Liu, Othmane Mazhar, Alexandros Nikou, Rui Oliveira, Pedro Roque, Elis Stefansson, David Umsonst, and Yu Wang, for all the time shared as colleagues and friends.

An important part of the results developed here are thanks to my great collaborators around the world. Thanks to James Welsh, who kindly invited me to the University of Newcastle in New South Wales, Australia, for a 5-week research visit in 2019. James, I have enjoyed every conversation we have had, and I hope we can meet again to run together. Thanks to Siqi Pan for welcoming me in Newcastle, and for introducing me to the technical details of the SRIVC estimator which gave way to an essential part of this thesis. Thanks to Håkan Hjalmarsson for introducing to me the idea of non-causality in band-limited setups, and to Johan Schoukens for his interest in my work. Thanks to Angel Cedeño, María Coronel and Juan Carlos Agüero from the Santa María Technical University (UTFSM) in Valparaíso, Chile, for playing a fundamental part in the implementation of the EM algorithm for event-based setups. Special thanks go to Ricardo Rojas, Francisco Vargas, and Juan Yuz, for

guiding me during my formative years in Chile as an electronic engineering student.

I also want to express my gratitude to Patricio Valenzuela, Daniela Medina, and their daughters Fernanda and Amanda. Thank you for all the great moments we have shared together, and for being as close as a family I have had here in Sweden. Thanks to José Marroquín and Nina Salalaiko for their hospitality and for the fun times during our sails through Stockholm.

I would also like to thank all my friends in Chile. I am specially grateful to my friends from the UTFSM for the great moments we have shared when I am back home: José Allende, Danilo and Felipe Ávila, Guillermo Becerra, Guillermo Castro, René Fredes, Mario López, Julian Rojas and José Rojas. Thanks to Gonzalo Farías and José Ignacio Freire, for being emotionally close to me despite the distance barriers. Thanks also go to my high school friends Felipe Barrera, Marko Gasic, Ives and Ignacio Ibacache, Jorge Pinto, Alfonso Prado, Marcelo Rivera, Pablo Salgado and Camilo Silva.

Finally, I want to thank all the support and love I have received from my family. My life partner Antonia Murillo has encouraged and supported me from Chile ever since I started my Ph.D. journey; your love will always be with me, no matter where we are. I am also really grateful to my brother Eduardo and my sister-in-law Nicole Arraño for all the joyful moments we have shared across the years, and to my sister Paula for always being a source of good vibes. Lastly, I thank my parents Agustín and Cecilia for the unconditional love and support I have always received from them.

Rodrigo A. González

Stockholm, Sweden. April, 2022.

To my family

Contents

Acknowledgements	vii
Abbreviations	xiv
1 Introduction	1
1.1 What does continuous-time system identification have to offer?	3
1.2 Related work	6
1.3 Thesis outline and contributions	8
2 Background: Continuous-time system identification	15
2.1 Notation	15
2.2 System and model considerations	16
2.3 Continuous-time identification methods	24
2.4 Summary	39
3 Asymptotically optimal indirect approach to continuous-time system identification	41
3.1 Introduction	41
3.2 Problem formulation and preliminaries	42
3.3 Asymptotically optimal enforcement of relative degree	44
3.4 Enforcing stability in IPEM	49
3.5 Monte Carlo simulation studies	54
3.6 Conclusions	61
3.A Theorem proofs	63
4 Consistency and asymptotic efficiency of the SRIVC estimator	67
4.1 Introduction	67
4.2 System and model setup	70
4.3 Consistency analysis	70
4.4 Efficiency analysis	81
4.5 Monte Carlo simulation studies	89
4.6 Conclusions	96
4.A Supplementary material for consistency results	97

4.B	Supplementary material for efficiency results	101
4.C	A result from real analytic function theory	102
5	Theoretical and practical aspects of SRIVC for over-parameterized models	105
5.1	Introduction	105
5.2	System and model setup	106
5.3	Analysis of SRIVC for over-parameterized models	107
5.4	Simulation studies	118
5.5	Conclusions	123
5.A	Supplementary material	125
6	The SRIVC estimator for arbitrary continuous-time inputs	127
6.1	Introduction	127
6.2	System and model setup	129
6.3	The SRIVC method for arbitrary continuous-time inputs	130
6.4	Efficiency and iterations analysis of the extended SRIVC method	138
6.5	Computational aspects	144
6.6	Monte Carlo simulation studies	147
6.7	Conclusions	153
6.A	Supplementary material for consistency results	155
6.B	Supplementary material for efficiency and iterations analysis	159
7	Consistency of refined IV methods for continuous-time systems in closed-loop	163
7.1	Introduction	163
7.2	Closed-loop settings	165
7.3	Analysis of the SRIVC and CLSRIVC estimators for Setting 1	167
7.4	Analysis of the SRIVC and CLSRIVC estimators for Setting 2	172
7.5	Simulations	180
7.6	Conclusions	183
7.A	Supplementary material	185
8	Refined IV methods for unstable continuous-time systems	187
8.1	Introduction	187
8.2	Problem formulation	189
8.3	RIVC and SRIVC for unstable systems	190
8.4	The CLSRIVC-uns method	196
8.5	Estimating continuous-time systems with integral action	199
8.6	Monte Carlo simulation studies	201
8.7	Conclusions	206
8.A	Proofs of technical results	209
8.B	Filtered regressor and output when the input is an arbitrary signal	211
8.C	On computing the derivatives in (8.19)	212

9 Discrete-time representations of systems with band-limited inputs	215
9.1 Introduction	215
9.2 Preliminaries and main non-causality result	216
9.3 Properties of g_{BL} and oversampling variants	221
9.4 Conclusions	227
9.A Proofs of technical results	229
10 Identification of continuous-time systems with band-limited inputs	235
10.1 Introduction	235
10.2 Problem formulation	237
10.3 Non-parametric frequency response estimation: Non-causal regularized least-squares	238
10.4 From non-parametric to parametric estimators	248
10.5 Multisine inputs: analysis of non-parametric and parametric estimators	257
10.6 Conclusions	271
10.A Proofs of technical results	273
10.B Computation of the asymptotic variance of \hat{a}_1	281
11 Summary and future research directions	283
11.1 Thesis conclusions	283
11.2 Some research directions	286
Bibliography	289

Abbreviations

ARMA	<i>autoregressive moving-average</i>
ARMAX	<i>autoregressive moving-average with exogenous input</i>
ARX	<i>autoregressive with exogenous input</i>
a.s.	<i>almost surely</i>
BIBO	<i>bounded-input bounded-output</i>
CLSRIVC	<i>closed-loop simplified refined instrumental variables for continuous-time systems</i>
CRLB	<i>Cramér-Rao lower bound</i>
dist.	<i>(convergence in) distribution</i>
DTFT	<i>discrete-time Fourier transform</i>
EM	<i>expectation-maximization</i>
FIR	<i>finite impulse response</i>
FOH	<i>first-order hold</i>
GEE	<i>generalized equation error</i>
GN	<i>Gauss-Newton</i>
i.i.d.	<i>independent and identically distributed</i>
LMI	<i>linear matrix inequality</i>
LS	<i>least-squares</i>
LSSVF	<i>least-squares state-variable filter</i>
LTI	<i>linear time-invariant</i>
MIMO	<i>multi-input multi-output</i>
ML	<i>maximum likelihood</i>
MSE	<i>mean square error</i>
PDF	<i>probability density function</i>
PEM	<i>prediction error method</i>
PRBS	<i>pseudorandom binary sequence</i>
RIVC	<i>refined instrumental variable method for continuous-time systems</i>
RLS	<i>regularized least-squares</i>

SISO	<i>single-input single-output</i>
SNR	<i>signal-to-noise ratio</i>
SRIV	<i>simplified refined instrumental variable method (for discrete-time systems)</i>
SRIVC	<i>simplified refined instrumental variable method for continuous-time systems</i>
w.p. 1	<i>with probability 1</i>
WNSF	<i>weighted null-space fitting</i>
ZOH	<i>zero-order hold</i>

Chapter 1

Introduction

Since the second half of the previous century, modern technological systems have revolutionized most aspects of scientific, economic, and social endeavors. A central feature of such technology is the generation and use of data, which quantifies information about the status of processes. In this day and age, data is more ubiquitous than ever: it is well known that the amount of digital data that is being stored is growing at astonishing rates [38], and its manipulation is essential in many fields. Despite the wide access to data, what is truly useful is the knowledge we can extract from it to understand previous behaviors, predict future events, and make better decisions. For this, a common approach is to use the available data together with reasonable prior assumptions to construct *models*, which are useful representations of the underlying phenomena that drive the data. Through models, it is possible to understand cause-effect relationships and focus only on the desired properties of physical or abstract phenomena, avoiding excessive complexity but also oversimplifications.

In many areas of science and engineering, one is interested in modeling *systems*, which can be understood as entities that manipulate one or several variables to accomplish a function, thereby yielding observable variables or *outputs* [108]. The external stimuli that can be manipulated by the observer are called *inputs*, while the ones that cannot be controlled are called *disturbances*. A system can be dynamic (or dynamical), that is, it can depend on internal information (or *state*) present in it that represents a summary of the system's history. In some circumstances, and subject to suitable simplifications, it is possible to write down the physical laws that govern the behavior of a dynamical system, thus arriving at a mathematical model of it. Many times a more experimental approach is needed, in which the model is learned from the data collected from the system. The field that studies how to obtain mathematical models for dynamical systems based on data is called *system identification* [61, 138, 211]. Note that mathematical models, either obtained through physical laws or system identification, are only abstractions of the underlying phenomena and never fully describe the true systems.

The system identification paradigm consists of at least four steps [138], which can be reassessed during the process of obtaining a model. First, a data set is needed: the

external stimuli that act on the system of study must be designed, and the system's response must be recorded. The conditions under which this input is applied, and the input itself, must be such that the data becomes maximally informative. Next, a *model structure* must be chosen; that is, a set or family of models must be picked such that the true system has at least one suitable representative within it. Afterwards, the "best" model in the set must be determined given the data at hand. This is done by selecting an *identification method*, which is a mapping between the data and a particular model in the model set. Finally, the model that is obtained through an identification method must be validated to check if it is suitable for the intended purpose. Since a model should never be accepted as the true description of the system, the validation step must only confirm whether the model provides a good enough description of certain aspects of interest of the system.

Every step of the system identification procedure described above is challenging in its own way. The model selection step is arguably the most difficult since it requires the user to analyze the complexity of the data and combine that information with prior knowledge and engineering intuition. A trade-off arises between model flexibility and tractability, as more complex model structures have more degrees of freedom but may also lead to unnecessarily complicated computations for finding the appropriate model and for using it. In addition, more data is needed to fit them with a prescribed accuracy. Some techniques, called *non-parametric methods*, have been designed to avoid the selection of the model order by delivering generic curves, plots, frequency responses, or impulse/step responses [137, 172]. If the family of models is instead parameterized by a finite-dimensional vector that encompasses all models up to a certain order, the techniques that estimate the parameter vector are called *parametric methods*.

Before selecting a parametric method for system identification, the relationship of the model with time must be decided. A family of models can take various forms, such as differential equations, transfer functions, difference equations, or state-space formulations. A distinction is made between *discrete-time* and *continuous-time* models. In discrete-time modeling, it is assumed that a complete description of the underlying system can be made by only observing its behavior at specific time instants, which are usually equally spaced in time. The boom in digital technology has impelled the development of discrete-time modeling, despite the fact that most physical processes are continuous-time in nature. On the other hand, continuous-time modeling consists in deriving mathematical relations that reflect the properties of the system under study for any moment in time instead of only some fixed instants of time.

Regarding the choice of the identification method, there have been many algorithms developed both for discrete-time models [15, 138, 210, 211] as well as for continuous-time models [61, 186, 250]. Other algorithms are specifically designed for identifying systems that are in closed-loop feedback schemes [51, 227], or in network environments [229, 234]. The identification method (or *estimator*) that is chosen should ideally have statistical guarantees. These are usually formulated by assuming the availability of large data sets and the existence of a *true system*, that

is, the existence of a particular mathematical model that we assume has generated the collected data. One important property that good estimators enjoy is *consistency*, which is related to accuracy and asymptotic unbiasedness; more precisely, an estimator is said to be consistent when the resulting model converges to the true mathematical description of the system as the number of data points tends to infinity. *Asymptotic efficiency* is another measure of the quality of an estimator that is related to precision; that is, a consistent estimator is said to be asymptotically efficient when its asymptotic covariance matrix is the smallest in a positive semi-definite sense among all consistent estimators.

The dichotomy between sampled data and continuous-time modeling is central in the development of continuous-time system identification methods. Since usually only sampled data is retrieved from an identification experiment, it is more natural to exploit time shifts in the design of estimators (as in discrete-time system identification) instead of time derivatives. This argument was adopted to explore an indirect approach for continuous-time system identification [186], which consists in first estimating a discrete-time model using input and output data and later computing its continuous-time equivalent. Unfortunately, such methods are prone to initialization problems, lack of robustness, and over-parametrization. A direct approach [255] can be pursued instead, in which a continuous-time model is directly obtained from the sampled data. The main shortcoming of direct methods is that they require the time derivatives of the input and output signals to be approximated. This problem is usually solved by introducing continuous-time low-pass filters that are applied to sampled data by assuming that the input and output signals have a known intersample behavior. The use of continuous-time filters, in conjunction with discrete-time data, leads to difficulties in how to formally derive statistical properties of these methods when the effect of the intersample behavior of the signals must be taken into account.

This thesis covers identification methods for continuous-time systems. Our interest is in providing comprehensive statistical analyses of several indirect and direct identification methods for linear continuous-time systems, and to extend their robustness and optimal asymptotic properties to more intricate identification settings.

1.1 What does continuous-time system identification have to offer?

Despite the wide popularity and success of discrete-time system identification methods, there are many reasons why continuous-time system identification techniques may be more appropriate than the discrete-time ones in practical applications. Here we will review some arguments that support the use of continuous-time system identification methods. Extensive lists of advantages and benefits can be found in, e.g., [59, 61, 63].

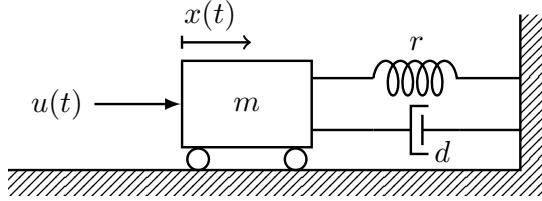


Figure 1.1: Mass-spring-damper system.

Physical insights and less over-parametrization

The parameters of physical processes often have a direct interpretation in continuous-time, whereas in discrete-time, these may not have any physical interpretation. The lack of physical interpretation of discrete-time models can complicate the validation of the model using expert knowledge since it is not immediately clear that the estimated parameters are in line with what is known *a priori* about the physical system. Furthermore, discrete-time models can lack parsimony in some cases, depending on the relative degree of the true continuous-time system.

To show these traits, take as an example the mass-spring-damper system in Figure 1.1. The spring and damper coefficients are $r[\text{N/m}]$ and $d[\text{Ns/m}]$, respectively, while the object has mass $m[\text{Kg}]$. If we wish to model the displacement $x(t)$ that is produced by an external force $u(t)$, by leveraging Newton's laws we may describe the system as

$$x(t) = \frac{1}{mp^2 + dp + r} u(t), \quad (1.1)$$

where p is the Heaviside operator, i.e., $px(t) = dx(t)/dt$. Now assume that the input force is constant between measurements, which are retrieved every $h[\text{s}]$. If the system exhibits a damped oscillatory response (i.e., $d < 2\sqrt{rm}$), we can then write the discrete-time equivalent of the mass-spring-damper system as

$$x(kh) = \frac{qb_1 + b_0}{q^2 - 2e^{-\frac{dh}{2m}} \cos\left(\frac{h}{m}\sqrt{mr - \frac{d^2}{4}}\right)q + e^{-\frac{dh}{m}}} u(kh), \quad (1.2)$$

where q is the forward shift operator, i.e., $qx(kh) = x(kh + h)$, and

$$\begin{aligned} b_0 &= \frac{e^{-\frac{dh}{m}}}{r} + \frac{e^{-\frac{dh}{2m}}}{r} \left(\frac{d}{\sqrt{4rm - d^2}} \sin\left[\frac{h}{m}\sqrt{rm - \frac{d^2}{4}}\right] - \cos\left[\frac{h}{m}\sqrt{rm - \frac{d^2}{4}}\right] \right), \\ b_1 &= \frac{1}{r} - \frac{e^{-\frac{dh}{2m}}}{r} \left(\cos\left[\frac{h}{m}\sqrt{rm - \frac{d^2}{4}}\right] + \frac{d}{\sqrt{4rm - d^2}} \sin\left[\frac{h}{m}\sqrt{rm - \frac{d^2}{4}}\right] \right). \end{aligned}$$

There are at least two conclusions that can be obtained by comparing the system descriptions (1.1) and (1.2). First, as we convert the model from continuous to discrete-time, the physical meaning of the coefficients is lost. It is more difficult to

check if the model parameters correspond to reasonable values in (1.2), and it is not clear how to relate the discrete-time system parameter estimates to values of m , d and r separately. Secondly, we see that four parameters must be fitted in (1.2), whereas only three parameters completely define the continuous-time model in (1.1). If this mismatch is not accounted for in a discrete-time system identification task, then the discrete-time estimates will not be as accurate as the ones obtained through continuous-time system identification methods since the variance of a model tends to increase with the number of parameters that need to be estimated.

Use of non-uniformly sampled data and time-delay systems

Instead of using the derivative operator for continuous-time systems, dynamical discrete-time systems are described by using (fixed) temporal displacements of signals. The displacement is dependent on the sampling period, thus, one continuous-time system leads to a different discrete-time representation for each sampling period that is chosen. If non-uniformly sampled data is collected and a linear and time-invariant model is sought, there is no adequate discrete-time model that can be naturally adjusted to the data without introducing time-variant elements or imposing unknown intersample behaviors. Identification using such types of data sets is needed in systems operating within a network environment [234], as well as for systems whose sampling is event-based such as the Lebesgue-sampled ones [12] where data is collected only when the measurements cross certain thresholds. Note that continuous-time models do not have problems with irregularly sampled data, as they are not tied to a sampling period.

Moreover, fractional time-delay systems—that is, systems that have an inherent time delay that is not a multiple of the sampling period—can be appropriately estimated in continuous-time, and several recent works have addressed this [29, 33, 35, 104]. The estimation of these delays may not be as natural in discrete-time modeling, since they will manifest themselves as zeros in the discrete-time system description that can be easily confused with sampling zeros or non-minimum phase zeros.

Robustness

Continuous-time identification methods have a natural advantage over discrete-time ones when the sampling frequency is high. This aspect is particularly relevant since most modern data acquisition devices can sample inputs and outputs at very high frequencies, leading to an almost continuous-time description of the signals of interest. For fast sampling rates, discrete-time methods using the shift operator can be ill-conditioned due to the clustering of discrete-time model poles around the point $(1, 0i)$ of the complex z -plane. As a remedy, the δ operator has been introduced for discrete-time system modeling and control [154]. On the other hand, continuous-time system identification can naturally handle high-frequency sampling and may lead to

better results as the sampling frequency increases, since the intersample behavior of the signals can be better approximated.

By a similar reasoning, systems that have poles that cover a wide dynamic range (i.e., stiff systems) are particularly difficult to model using discrete-time approaches with a fixed sampling period. The main reason is that it is often difficult to find a sampling period that can capture the system's dynamics without any compromise: numerical conditioning is affected when the sampling period is too small, and the high-frequency spectral content is poorly estimated when the sampling period is too large. In contrast, reliable continuous-time models can be obtained through fast sampling and their parameters do not depend on the sampling period.

1.2 Related work

There have been many surveys on identification of continuous-time systems [59, 186, 225, 226, 244], as well as books written on the subject [61, 203, 224, 250]. Here we will provide a brief history of the field and mention some contributions that have marked the direction of research, with a focus on the most relevant works related to this thesis.

One of the first studies on parameter estimation of continuous-time models was done in 1965 in a purely analog framework [241], in which an adaptive identification procedure was introduced for continuous-time data that included prefiltering of the input and output signals with a low-pass analog filter. This algorithm resembles the work in [198], where integrating steps were proposed. The availability of digital computers during the mid-nineteen sixties produced a rapid development of discrete-time control and system identification, as going “completely digital” was now feasible. Further developments followed due to the benefits of continuous-time modeling, now using sampled data. The least-squares state-variable-filter (LSSVF) method for discrete-time data was introduced in [242], which considers a differential equation to model the system and filters the discrete-time input and output data (and their derivatives) with a low pass filter. After the prefiltering step, the parameters are estimated using standard least-squares.

The simplified refined instrumental variable method for continuous-time systems (SRIVC) was first presented in [255], and it has impulsed the development and practical use of the direct approach for continuous-time modeling. The SRIVC algorithm can be viewed as an iterative procedure in which an instrumental variable estimate is obtained by using prefiltered input and output signals, whose prefilters are updated in every iteration. An extension of this method, the refined instrumental variable method for continuous-time systems (RIVC), admits colored noise modeling in a hybrid Box-Jenkins model structure [248]. These iterative algorithms have led to many contributions, such as extensions to multi-input single-output continuous-time systems [66] and closed-loop systems [68], as well as procedures that can handle non-uniformly sampled data for SRIVC [118] and RIVC in a continuous-time noise modeling framework [30]. Other works have proposed a unification with the discrete-

time versions of SRIVC and RIVC [252], tutorials and toolboxes with recursive estimation implementations [60, 62, 221], and comparisons with other identification methods [246]. The consistency of the RIVC estimator was analyzed in [135], where an extra filter was introduced for the purpose of discretizing the derivatives of the input signal. Recently, research has moved towards extending the iterative instrumental variable procedures to identify time-delay continuous-time systems, where special caution must be taken to avoid local minima [33–35, 104]. Also, interesting surveys have been written remarking the advantages of continuous-time system identification over discrete-time, with a focus on the merits of the SRIVC estimator and the direct approaches in general [59, 63].

Apart from the SRIVC-based methods, several other direct identification procedures for continuous-time systems have been developed. Frequency-domain identification methods have been studied in, e.g., [173, 179, 181]. In [149], a correlation method was introduced and analyzed, with close connections to the frequency-domain algorithms. The work in [120] pursued an algebraic approach, that consisted in introducing an operator λ which acts like a prefilter to the input and output signals. An extension of this idea was presented in [39], where continuous-time Laguerre functions were chosen as prefilters. As a way to avoid prefiltering, the idea of indirect inference [99] was introduced in continuous-time system identification in [239], with excellent results in simulation examples.

In parallel to the direct methods previously described, indirect methods for continuous-time system identification were developed thanks to the success and popularity of discrete-time system identification. One of the first indirect methods can be found in [204], in which a two-step procedure was presented: first, a discrete-time model was obtained through an eigenvector method for transfer function estimation, and later, a continuous-time model was derived by solving a set of equations with the coefficients of the estimated continuous-time model as unknowns. The problem of the excess of relative degree in the continuous-time models was acknowledged in that work and was solved via a generalized least squares approach over the overdetermined set of equations relating the numerator parameters. The bilinear z transformation for obtaining the continuous-time model from a discrete-time transfer function estimate was proposed in [199], while [202] extended the analysis to multivariable systems and different transformations from continuous to discrete-time. The selection of the sampling period for indirect and direct methods for continuous-time modeling was studied in [201], and state-space methods were reviewed in [200]. In [7] it was proposed that the continuous-time parameters can be uniquely determined by prefiltering the input sequence, and thus better conditioning of the model conversion step can be achieved. The contribution in [185] helped to clarify the shortcomings of indirect methods in favor of the SRIVC estimator, which is classified as a direct method. Further tests followed in [141], where it was confirmed that the indirect approach methods suffer from initialization problems (especially in the output error model structure), and that the discrete to continuous-time transformation can be ill-conditioned and not suitable for systems with a relative degree greater than one.

1.3 Thesis outline and contributions

In this section, we provide the outline of the thesis and indicate the contributions in each chapter.

Chapter 2

In this chapter, we detail the notation used throughout this thesis and review the main ideas and methods behind the indirect and direct approaches to continuous-time system identification in open and closed-loop. The material covered in this chapter is mostly based on [62, 138, 186] and [61].

Chapter 3

In this chapter, we propose an indirect method to continuous-time system identification that enforces a fixed relative degree in the transfer function estimate. By relating it to the indirect PEM method, we show that this estimator is consistent and asymptotically efficient. Furthermore, to cope with highly noisy data sets, we develop a refinement of this method that enforces stability on the model by optimizing over ellipsoidal inner approximations of the stability region in the parameter space. Extensive numerical simulations are put forward to show the performance of this estimator when contrasted with other indirect and direct methods for continuous-time system identification.

This chapter is based on the following publications:

- [C1] R. A. González, C. R. Rojas, and J. S. Welsh. An asymptotically optimal indirect approach to continuous-time system identification. In *Proceedings of the 57th IEEE Conference on Decision and Control (CDC'18)*, pages 638-643, 2018.
- [C2] R. A. González, J. S. Welsh and C. R. Rojas. Enforcing stability through ellipsoidal inner approximations in the indirect approach for continuous-time system identification. In *Proceedings of the 21st IFAC World Congress (IFAC 2020)*, pages 566-571, 2020.

Chapter 4

This chapter studies the asymptotic properties of the SRIVC estimator for inputs that are exactly reconstructed with zero or first-order hold devices. It is divided into two main sections. First, we present a comprehensive consistency analysis of the SRIVC estimator while taking into account the intersample behavior of the input signal. We show that the SRIVC estimator is generically consistent under some mild conditions, and we describe the conditions for when consistency is not achieved. Later, we derive the asymptotic Cramér-Rao lower bound for the continuous-time output error model structure and provide an analysis of the asymptotic efficiency

of the SRIVC estimator. We prove that the SRIVC estimator is, under some mild conditions, asymptotically efficient for the output error model structure and i.i.d. Gaussian measurement noise. Monte Carlo simulations are performed to verify the asymptotic properties we have derived.

The covered material is based on the following publications:

- [J1] 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*, 113, Article 108767, 2020.
- [J1b] S. Pan, R. A. González, J. S. Welsh and C. R. Rojas. Corrigendum to “Consistency analysis of the Simplified Refined Instrumental Variable method for Continuous-time systems” [Automatica 113 (2020) 108767]. *Automatica*, 136, Article 109946, 2021.
- [J2] 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*, 121, Article 109196, 2020.

Chapter 5

In this chapter, we provide a comprehensive convergence analysis of the SRIVC estimator for when the model is over-parameterized in both numerator and denominator polynomials. In particular, we show that the numerical conditioning of the modified normal matrix of the SRIVC algorithm deteriorates for decreasing sampling periods, and we prove that the SRIVC estimator, at a limit point of the iterative procedure and as the sample size tends to infinity, returns the true transfer function with arbitrary pole-zero cancellations. A one-iteration analysis of the method is also performed, and we derive the asymptotic behavior of the normalized estimation variance norm (NEVN), which is part of Young’s information criterion (YIC) for model order selection. The theoretical results are verified with numerical simulations.

This chapter is based on the following publication:

- [J3] R. A. González, C. R. Rojas, S. Pan and J. S. Welsh. Theoretical and practical aspects of the convergence of the SRIVC estimator for over-parameterized models. Accepted for publication in *Automatica*, 2022.

Chapter 6

In this chapter, we introduce an extension of the SRIVC estimator that can handle inputs that are not necessarily reconstructed by standard hold devices but whose intersample behavior is known in advance. We prove that this estimator yields generic consistency of the estimated model parameters for continuous-time multisine input signal excitations. We also derive the asymptotic Cramér-Rao lower bound for the continuous-time output error model structure when the full continuous-time

input signal is known, and we formally specify conditions under which the extended SRIVC method relates to the maximum likelihood estimator. Later, we study the effect of the hold device for reconstructing the output measurements on the iterations of the extended SRIVC estimator and establish connections between the extended SRIVC iterations, Gauss-Newton iterations, and stability enforcement procedures based on projected gradient iterations. Finally, we discuss the implementation aspects of the extended SRIVC estimator and test the method under various simulation settings.

An important part of this chapter is based on the following publications:

- [J4] R. A. González, C. R. Rojas, S. Pan and J. S. Welsh. Consistent identification of continuous-time systems under multisine input signal excitation. *Automatica*, 133, Article 109859, 2021.
- [C3] R. A. González, C. R. Rojas, S. Pan and J. S. Welsh. The SRIVC algorithm for continuous-time system identification with arbitrary input excitation in open and closed loop. In *Proceedings of the 60th IEEE Conference on Decision and Control (CDC'21)*, pages 3004-3009, 2021.

Chapter 7

In this chapter, we provide a consistency analysis of two refined instrumental variable methods when they are applied in closed-loop settings. More precisely, we study the generic consistency of the SRIVC and CLSRIVC estimators when discrete-time and continuous-time controllers are implemented in the feedback loop, and we propose an extension of the CLSRIVC estimator to handle the loss of consistency when a continuous-time controller is used. This chapter proves that the SRIVC and CLSRIVC estimators are not generically consistent when there is a continuous-time controller in the loop, and that generic consistency can be achieved when the control is in discrete-time. The bias of the SRIVC estimator for biproper models and colored noise is also characterized when it is applied in a discrete-time control loop. The theoretical results are corroborated via extensive Monte Carlo simulation studies.

This chapter is based on the following contributions:

- [C3] R. A. González, C. R. Rojas, S. Pan and J. S. Welsh. The SRIVC algorithm for continuous-time system identification with arbitrary input excitation in open and closed loop. In *Proceedings of the 60th IEEE Conference on Decision and Control (CDC'21)*, pages 3004-3009, 2021.
- [J5] R. A. González, S. Pan , C. R. Rojas and J. S. Welsh. Consistency analysis of refined instrumental variables for continuous-time system identification in closed-loop. Submitted for publication to the *European Journal of Control*, 2022.

Chapter 8

In this chapter we propose various refined instrumental variable methods for estimating unstable continuous-time models. The ill-conditioning problem of the filtered regressor and instrument vectors, as well as of the filtered output, is solved by including a tailor-made all-pass filter in the prefiltering step. This approach is used for obtaining an extension of the LSSVF method, in addition to extensions for the SRIVC and RIVC methods, that admit the identification of unstable systems and are shown to minimize the prediction error upon convergence in iterations and as the sample size tends to infinity. Several implementations of these estimators are proposed depending on the intersample behavior of the input (zero and first-order hold, multisine and arbitrary). The particular case when the plant has integral action is explicitly considered. An extension of the CLSRIVC method is also proposed and discussed in detail in a setting where the reference signal is known. Monte Carlo simulations are used to assess the performance of the estimators.

This chapter is based on the following contribution:

- [J6] R. A. González, C. R. Rojas, S. Pan and J. S. Welsh. Refined instrumental variable methods for unstable continuous-time systems. Accepted for publication in the *International Journal of Control*, 2022.

Chapter 9

This chapter investigates discrete-time representations of continuous-time systems under band-limited inputs. The interpolation property of band-limited input signals is shown to yield equivalent discrete-time system representations that are non-causal in general. This fact is explored in detail to reveal the effects of fast sampling in the band-limited equivalent impulse response coefficients, and to analyze the situations where non-causality plays a role in representation and estimation of linear systems under band-limited inputs.

An important part of this chapter is based on the following contributions:

- [C4] R. A. González, C. R. Rojas and H. Hjalmarsson. Non-causal regularized least-squares for continuous-time system identification with band-limited input excitations. In *Proceedings of the 60th IEEE Conference on Decision and Control (CDC'21)*, pages 114-119, 2021.
- [J7] R. A. González, C. R. Rojas, H. Hjalmarsson and J. Schoukens. Identification of continuous-time systems excited by band-limited inputs. Under preparation, 2022.

Chapter 10

The observations of Chapter 9 are exploited in this chapter to study non-parametric and parametric frequency response estimators of linear continuous-time

systems. As a first contribution, we study the properties of non-causal least-squares estimators for continuous-time system identification and propose a kernel-based non-causal regularized least-squares approach for estimating the band-limited equivalent impulse response. This analysis is later used for constructing a parametric estimate of the continuous-time system based on a frequency-domain refined instrumental variable method. Secondly, as a particular case, we analyze the properties of causal and non-causal least-squares estimators when the input is a continuous-time multisine signal. This analysis links the non-parametric frequency response resulting from least-squares to maximum likelihood estimation and refined instrumental variable methods. The estimators and theoretical findings are tested via extensive numerical simulations.

An important part of this chapter is based on the following contributions:

- [C4] R. A. González, C. R. Rojas and H. Hjalmarsson. Non-causal regularized least-squares for continuous-time system identification with band-limited input excitations. In *Proceedings of the 60th IEEE Conference on Decision and Control (CDC'21)*, pages 114-119, 2021.
- [J7] R. A. González, C. R. Rojas, H. Hjalmarsson and J. Schoukens. Identification of continuous-time systems excited by band-limited inputs. Under preparation, 2022.

Chapter 11

In Chapter 11, we summarize the main contributions of the thesis and provide suggestions for future research directions.

Contributions not included in this thesis

The following contributions have not been included in the thesis:

- [C5] R. A. González, F. J. Vargas and J. Chen. Stabilization of MIMO systems over additive correlated noise channels subject to multiple SNR-constraints. In *Proceedings of the 16th European Control Conference (ECC'18)*, pages 1493-1498, 2018.
- [C6] R. A. González and C. R. Rojas. A fully Bayesian approach to kernel-based regularization for impulse response estimation. In *Proceedings of the 18th IFAC Symposium on System Identification*, pages 186-191, 2018.
- [J8] R. A. González, F. J. Vargas and J. Chen. Necessary and sufficient conditions for mean square stabilization over MIMO SNR-constrained channels with colored and spatially correlated additive noises. In *IEEE Transactions on Automatic Control*, volume 64, pages 4825-4832, 2019.

- [J9] F. J. Vargas and R. A. González. On the existence of a stabilizing solution of modified algebraic Riccati equations in terms of standard algebraic Riccati equations and linear matrix inequalities. In *IEEE Control Systems Letters*, volume 4, pages 91-96, 2020.
- [C7] R. A. González and C. R. Rojas. Finite sample deviation and variance bounds for first order autoregressive processes. In *Proceedings of the 45th International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pages 5380-5384, 2020.
- [C8] R. A. González and C. R. Rojas. A finite-sample deviation bound for stable autoregressive processes. In *Proceedings of the 2nd Annual Conference on Learning for Dynamics and Control (L4DC 2020)*, pages 191-200, 2020.
- [J10] S. Pan, J. S. Welsh, R. A. González and C. R. Rojas. Consistency analysis and bias elimination of the Instrumental Variable Based State Variable Filter method. Provisionally accepted for publication in *Automatica*, 2022.
- [C9] R. A. González, A. L. Cedeño, M. Coronel, C. R. Rojas and J. C. Agüero. Identification of continuous-time state-space systems utilizing Lebesgue-sampled data. Submitted to the *61st IEEE Conference on Decision and Control (CDC'22)*, 2022.

Author's contributions

The order of the author names in the publications reflects the workload, where the first one has the most important contribution. In all listed publications, all authors were actively involved in formulating the problems, developing the solutions, evaluating the results, and writing the papers.

Chapter 2

Background: Continuous-time system identification

The purpose of this chapter is to provide the necessary background material to understand the contents of this thesis, and to introduce the essentials of continuous-time system identification that will give context to the contributions herein.

This chapter is structured as follows. In Section 2.1 we define the notation common to all the subsequent chapters. Section 2.2 describes the system and model structures that are of interest for this thesis, while Section 2.3 defines the statistical properties related to the performance of estimators and discusses several continuous-time system identification methods. We summarize this chapter in Section 2.4.

2.1 Notation

The imaginary number $\sqrt{-1}$ is written as i . The real and imaginary part of a complex number z are denoted by $\text{Re}\{z\}$ and $\text{Im}\{z\}$ respectively, and its complex conjugate is written as \bar{z} . All matrices and vectors are written in bold, and column vectors are utilized, unless transposed. We write the n -th dimensional (column) vector \mathbf{x} with entries x_1, x_2, \dots, x_n as $\mathbf{x} = [x_1, x_2, \dots, x_n]^\top$. If \mathbf{A} is a matrix, then \mathbf{A}^\top and \mathbf{A}^H denote its transpose and Hermitian (complex conjugated transpose), respectively. The Kronecker product between the matrices \mathbf{A} and \mathbf{B} is denoted by $\mathbf{A} \otimes \mathbf{B}$. The identity matrix of size n is written as \mathbf{I}_n , where the subscript can be absent if no ambiguity exists. The vector \mathbf{e}_k denotes the k -th column of the identity matrix of appropriate size. If \mathbf{A} and \mathbf{B} are symmetric matrices, the expression $\mathbf{A} \succ \mathbf{B}$ (resp., $\mathbf{A} \succeq \mathbf{B}$) means that the matrix $\mathbf{A} - \mathbf{B}$ is positive definite (resp., semi-definite). If $g(\mathbf{x})$ is a scalar function of the vector \mathbf{x} , then $\partial g(\mathbf{x})/\partial \mathbf{x}$ is a column vector denoting the gradient of the function g with respect to \mathbf{x} . If $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{Q} \in \mathbb{R}^{n \times n}$ is a positive definite matrix, then the 2-norm of \mathbf{x} is $\|\mathbf{x}\|_2 = \sqrt{\mathbf{x}^\top \mathbf{x}}$, and the weighted 2-norm of \mathbf{x} is $\|\mathbf{x}\|_{\mathbf{Q}} = \sqrt{\mathbf{x}^\top \mathbf{Q} \mathbf{x}}$. The quantities $\sigma_{\min}(\cdot)$ and $\sigma_{\max}(\cdot)$ represent the smallest and largest singular value of a matrix respectively, and $\text{cond}(\cdot)$ denotes its condition number.

If $\{f(kh)\}_{k=0}^{\infty}$ is a sequence $f(0), f(h), \dots$, then the operation $qf(kh)$ returns $f(kh + h)$, with q being called the forward shift operator. The Z-transform and delta transform of this sequence are denoted as $\mathcal{Z}\{f(kh)\}$ and $\mathcal{T}\{f(kh)\}$, respectively. In order to remark the time instant evaluation, we usually employ the notation $\{f(t_k)\}_{k=1}^N$ to denote the set of evaluations $\{f(t_1), f(t_2), \dots, f(t_N)\}$, where the time instants $\{t_k\}_{k=1}^N$ may or may not be evenly spaced, depending on the context. When the initial and terminal instants are not relevant, we may also use the notation $\{f(t_k)\}$ or $\{f(kh)\}$ to denote a sequence ranging in k . If the context does not leave ambiguity, we may simply refer to these sequences as $f(t_k)$ or $f(kh)$. Similarly to the discrete-time case, if $\{x(t)\}_{t \in (t_1, t_N)}$ is a continuous-time signal defined on the open interval (t_1, t_N) , then the Heaviside operator p satisfies $px(t) = dx(t)/dt$. The Laplace transform of $\{x(t)\}_{t \geq 0}$ is denoted as $\mathcal{L}\{x(t)\}$ and is a function of the complex variable s .

If x is a random variable, $\mathbb{E}\{x\}$ denotes its mean. For sequences of random variables $\{y_n\}$ and $\{r_n\}$, the notation $y_n = o_p(r_n)$ means that $y_n = x_n r_n$ for $\{x_n\}$ being some sequence of random variables that converges to zero in probability as n tends to infinity. The expressions $y_n \xrightarrow{p} y$ and $y_n \xrightarrow{a.s.} y$ denote convergence in probability and almost sure convergence, respectively.

2.2 System and model considerations

In this thesis we will assume that the system to be identified enjoys certain properties. First, the systems we study are assumed to be linear and time-invariant (LTI). A system is *linear* when its output to a linear combination of inputs is equal to the same linear combination of outputs obtained from the individual inputs. A system is *time-invariant* when its output does not depend explicitly on the instant of time when the input was applied. Furthermore, we typically assume that the systems we encounter are *causal*, which means that the output of a system at a given time instant t only depends on the values of the input up to time t . A system is *non-causal* if that property is not satisfied. Although several methods presented here can be extended to multi-input multi-output (MIMO) systems, we only study single-input single-output (SISO) systems.

One key aspect when describing dynamical systems is their nature in time. A causal, LTI *continuous-time* system is a mapping between a continuous-time input $u(t)$ and a continuous-time output $y(t)$ that can be expressed through a linear differential equation of the form

$$\frac{d^n y(t)}{dt^n} + a_{n-1} \frac{d^{n-1} y(t)}{dt^{n-1}} + \cdots + a_1 \frac{dy(t)}{dt} + a_0 y(t) = b_m \frac{d^m u(t)}{dt^m} + \cdots + b_0 u(t), \quad (2.1)$$

where m and n are non-negative integers such that $n \geq m$, and $a_0, a_1, \dots, a_{n-1}, b_0, b_1, \dots, b_m$ are real numbers. The roots of the polynomial $A(s) := s^n + a_{n-1}s^{n-1} + \cdots + a_1s + a_0$ define the asymptotic stability of the continuous-time system: it is *asymptotically stable* when $A(s)$ has all its (complex) roots in the open left half

plane. The difference $n - m$ is the *relative degree* of the continuous-time system. When $m \leq n$, which is the case studied in this thesis, we say that the system is *proper*, and it is *strictly proper* if $m < n$. In (2.1) we assume that the input does not have an inherent time delay.

In Chapters 4 and 5 we argue that in some situations it is convenient to study an alternative description of the continuous-time system which is equivalent to (2.1) when the system does not have poles at the origin:

$$a_n \frac{d^n y(t)}{dt^n} + a_{n-1} \frac{d^{n-1} y(t)}{dt^{n-1}} + \cdots + a_1 \frac{dy(t)}{dt} + y(t) = b_m \frac{d^m u(t)}{dt^m} + \cdots + b_0 u(t). \quad (2.2)$$

The system in (2.1) is written in its *monic* form, as the leading coefficient of the polynomial $A(s)$ is equal to 1. Conversely, (2.2) describes a system in *anti-monic* form. Both forms will be studied in this thesis, and the choice of form will depend on the context and application.

Sometimes it is convenient to use operator notation to denote the operation of differentiation. Consequently, we can use the Heaviside operator p to formally write the system in (2.2) as

$$G(p) = \frac{B(p)}{A(p)}, \quad (2.3)$$

where

$$\begin{aligned} B(p) &= b_m p^m + b_{m-1} p^{m-1} + \cdots + b_0, \\ A(p) &= a_n p^n + a_{n-1} p^{n-1} + \cdots + a_1 p + 1. \end{aligned}$$

We assume that $A(p)$ and $B(p)$ are coprime, i.e., when viewed as polynomials, they do not share roots. If $G(p)$ is not viewed as an operator but as a function of a complex variable s , it is the *transfer function* of the continuous-time system [97], and it is the Laplace transform of the impulse response of the system (2.2) with initial conditions equal to zero. Bear in mind that in some situations the roots of the A and B polynomials are studied; in such cases we sometimes use p and s (i.e., $A(p)$ or $A(s)$) interchangeably, if there is no possible ambiguity.

Alternatively to the continuous-time description introduced above, in many cases one can analyze dynamical systems by solely considering their behavior at certain time instants. If one considers input and output signals that are sampled with a period h greater than zero, a causal LTI *discrete-time* system is a mapping between a discrete-time input $u(kh)$ and a discrete-time output $y(kh)$ that can be written as a linear recursive equation of the form

$$y([k+n]h) + \alpha_{n-1}y([k+n-1]h) + \cdots + \alpha_0y(kh) = \beta_m u([k+m]h) + \cdots + \beta_0 u(kh), \quad (2.4)$$

where, similarly to the above, we can use operator formalism to rewrite the system as

$$y(kh) = \frac{B_d(q)}{A_d(q)} u(kh), \quad (2.5)$$

where

$$\begin{aligned} A_d(q) &= q^n + \alpha_{n-1}q^{n-1} + \cdots + \alpha_0, \\ B_d(q) &= \beta_m q^m + \beta_{m-1}q^{m-1} + \cdots + \beta_0. \end{aligned}$$

The expression $H(q) := B_d(q)/A_d(q)$, when not viewed as an operator but as a function of a complex variable z , is the transfer function of the discrete-time system and is the Z-transform of the impulse response of the system (2.4) with zero initial conditions. The transfer function $H(z)$ is said to be asymptotically stable if and only if the roots of $A_d(z)$ are all located in the open unit disk of the complex plane. The difference $n - m$ is the *relative degree* of the discrete-time system.

Another convenient way to write the discrete-time system (2.5) is by using the delta operator δ [154]. This operator is particularly useful when rapid sampling is used, since the parameters of systems in the δ -domain are not as numerically ill-conditioned as the ones in the q -domain. In the simplest embodiment¹ of the delta operator, the q and δ operators are linked by the relation

$$\delta = \frac{q - 1}{h}.$$

The transform variable in this case is denoted by γ , and the delta transform of a signal $f(kh)$ is defined by

$$\mathcal{T}\{f(kh)\} := h \sum_{k=0}^{\infty} f(kh)(1 + h\gamma)^{-k},$$

which leads to an alternative description of the discrete-time system (2.5) as $S(\gamma) = hH(1 + h\gamma)$, where $H(\cdot)$ is the discrete-time transfer function in the q -domain.

The notion of a true system

A key concept used throughout this work is the notion of a *true system*, that is, the system we intend to describe is in fact of the form (2.3) for certain (real) values of $a_1, a_2, \dots, a_n, b_0, b_1, \dots, b_m$. This is always an idealization, as no system representation of the form (2.3) can exactly describe the full complexity of nature [24]. However, this concept is helpful for building identification methods and it can be used to derive qualitative properties that “good” estimators should have, such as consistency (see Section 2.3.1). In that regard, true systems are also used for sanity checks: if estimators perform poorly under such an idealized condition, then we cannot expect them to perform well in practice either.

We will usually describe the true system in terms of the parameter vector

$$\boldsymbol{\theta}^* := \left[a_1^*, \ a_2^*, \ \dots, \ a_{n^*}^*, \ b_0^*, \ b_1^*, \ \dots, \ b_{m^*}^* \right]^\top, \quad (2.6)$$

¹We mention other variants of the delta operator in Subsection 2.3.2, all of which serve as tools for improving numerical robustness.

where n^* and m^* are the orders of the denominator and numerator polynomials of the true system. Following this notation, $G^*(p)$ denotes the (true) system under study, with $B^*(p)$ and $A^*(p)$ being the m^* -th order numerator and n^* -th order denominator polynomials of $G^*(p)$, respectively.

The goal in continuous-time system identification

The true parameter vector $\boldsymbol{\theta}^*$ of the continuous-time system, together with its dimension, is generally unknown to the practitioner. Thus, it must be estimated from data collected in an identification experiment. We seek to find a *model* of the continuous-time system $G^*(p)$ of the form (2.3) parameterized by the vector

$$\boldsymbol{\theta} = [a_1, \ a_2, \ \dots, \ a_n, \ b_0, \ b_1, \ \dots, \ b_m]^\top \quad (2.7)$$

that can represent the system faithfully, or that is valid with regard to the purpose in mind. As the model depends on the parameter vector we choose, we can make explicit this dependence as $G(p, \boldsymbol{\theta})$. We explore different methods for obtaining accurate models for $G^*(p)$ with particular attention to the theoretical properties that these methods enjoy. In particular, we study if there exist statistical guarantees of convergence to the true parameters as we acquire an increasing number of samples, or a notion of “asymptotically optimal precision” of our model with respect to the true system, which we will detail more precisely in Section 2.3.

Intersample behavior assumptions

Depending on the intersample behavior of the discrete-time input $u(t_k)$, it may or may not be possible to obtain an equivalence between the descriptions in (2.2) and (2.4). By equivalence we mean that, for a continuous-time input that has a known intersample behavior, we can find a discrete-time representation that delivers an exact representation of the continuous-time output at the sampling instants. Since the intersample behavior of the input may not be known, it is common to base our derivations on an intersample behavior guess or *assumption* on how a continuous-time signal is reconstructed from sampled data. From this reconstruction we can establish exact relationships between discrete-time and continuous-time systems. If the reconstruction matches the true intersample behavior of the signal, we say that the intersample behavior has been *correctly specified*. Otherwise, the intersample behavior is *misspecified*.

One typical intersample behavior assumption is provided by a *zero-order hold* (ZOH) device. This assumption considers the signal to be constant between consecutive samples. That is, for every t , a continuous-time version of a sampled signal $u(t_k)$ is constructed as

$$u(t) = u(t_k), \quad t_k \leq t < t_{k+1}, \quad k \in \mathbb{N}.$$

For sampled signals with a constant sampling period h , an exact relationship between continuous-time and discrete-time systems can be found if the input is constant

between samples [16]:

$$H_{\text{ZOH}}(z) = \frac{z-1}{z} \mathcal{Z} \left\{ \mathcal{L}^{-1} \left\{ \frac{G(s)}{s} \right\}_{t=kh} \right\}. \quad (2.8)$$

The ZOH assumption can be regarded as an extrapolation using a polynomial of degree zero [16]. Bear in mind that it is possible to use higher-order extrapolations that may achieve smaller reconstruction errors for smooth input signals. For example, a *first-order hold* (FOH) reconstruction is obtained by computing the line between the two most consecutive samples. A similar expression to (2.8) can then be obtained to relate the discrete-time and continuous-time transfer functions [54]:

$$H_{\text{FOH}}(z) = \frac{(z-1)^2}{zh} \mathcal{Z} \left\{ \mathcal{L}^{-1} \left\{ \frac{G(s)}{s^2} \right\}_{t=kh} \right\}.$$

The FOH intersample behavior assumption studied here leads to a causal discrete-time system, despite the reconstruction process being non-causal in time (see Figure 2.1, where the segment between samples depends on the next sample). Other FOH reconstruction mechanisms that are causal can be proposed [259], but they will not be reviewed here.

Apart from zero or first-order hold setups, we can also assume that the input is *band-limited* [181]. In this case the continuous-time input $u(t)$ is assumed to be a signal that has energy only on a finite band of frequencies, which permits full reconstruction of the intersample behavior based on its samples. Mathematically speaking we assume that $U(i\omega) = 0$ for $|\omega| > \omega_B$, where $U(i\omega)$ is the continuous-time Fourier transform

$$U(i\omega) = \int_{-\infty}^{\infty} u(t)e^{-i\omega t} dt,$$

and ω_B defines the bandwidth of the input signal. This setup is sometimes considered in an errors-in-variables framework where noisy samples of the band-limited input and output are measured [175].

In Figure 2.1 we show an example of how a sampled signal is reconstructed assuming each of the intersample behaviors we have mentioned above. Keep in mind that it is known that misspecifying the intersample behavior or model assumptions in system identification can greatly impact the performance of the estimation method in use [196].

Remark 2.1. The intersample behavior of the output may or may not play a role in continuous-time identification methods, depending on whether the estimator requires specifying an intersample behavior for the output or not. A rigorous analysis of this aspect for the SRIVC estimator is covered in Chapters 4 and 6.

The noise model

So far, we have described the true continuous-time system and transfer function models that are of interest in this work. We shall assume that in general we do not

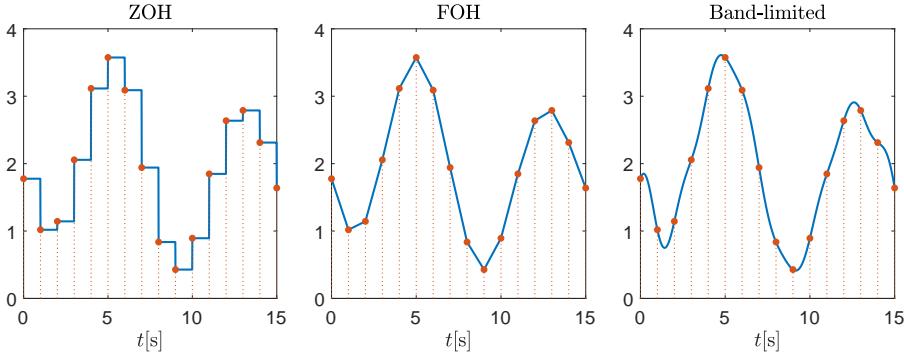


Figure 2.1: Reconstruction of a red-dotted sampled signal assuming a ZOH (left), FOH (center), or band-limited (right) intersample behavior.

have direct access to the output of a true continuous-time system, but to a noisy version of it. Following (2.2) and (2.3), one could consider

$$y(t) = G^*(p)u(t) + v(t), \quad (2.9)$$

where $\{v(t)\}$ is a pink noise process, or a band-limited continuous-time white noise process. Usually standard continuous-time white noise, understood as a stochastic process whose spectral density is constant in all the real frequency line, is not suitable as it carries theoretical and practical problems due to its infinite variance [9]. In this thesis we depart from (2.9) since we assume that only a sampled version of the continuous-time output is measured. We therefore instead follow a hybrid modeling approach, where a continuous-time plant model that is contaminated by discrete-time noise needs to be estimated from data. This description is quite popular and practical [67, 120, 255] since we typically only have access to the output at particular instants in time. In other words, the system we intend to model is of the form

$$x(t) = \frac{B^*(p)}{A^*(p)}u(t), \quad (2.10a)$$

$$y(t_k) = x(t_k) + v(t_k), \quad (2.10b)$$

where $k = 1, \dots, N$, with N being the number of samples extracted from an identification experiment, and $v(t_k)$ is assumed to be a zero-mean discrete-time random process. In the case when $v(t_k)$ is a white noise process, the structure in (2.10) is known as an *output error* structure, as the noise source is assumed to be present in the output measurement only.

More broadly, an estimate of the spectrum of the additive noise might also be of interest. In that case, following the notion of a true system described before, we may also assume that the discrete-time measurement noise is a filtered white noise and the filter can be exactly described as a discrete-time transfer function. Then,

the system has the form

$$x(t) = \frac{B^*(p)}{A^*(p)} u(t), \quad (2.11a)$$

$$y(t_k) = x(t_k) + \frac{C^*(q)}{D^*(q)} e(t_k), \quad (2.11b)$$

where $e(t_k)$ describes a zero-mean white noise stochastic process of finite variance, and

$$C^*(q) = 1 + c_1^* q^{-1} + c_2^* q^{-2} + \cdots + c_{m_c^*}^* q^{-m_c^*},$$

$$D^*(q) = 1 + d_1^* q^{-1} + d_2^* q^{-2} + \cdots + d_{n_d^*}^* q^{-n_d^*},$$

with the polynomial degrees satisfying $n_d^* \geq m_c^*$. The true parameter vector related to the noise filter is defined as

$$\boldsymbol{\eta}^* := \begin{bmatrix} c_1^*, & c_2^*, & \dots, & c_{m_c^*}^*, & d_1^*, & d_2^*, & \dots, & d_{n_d^*}^* \end{bmatrix}. \quad (2.12)$$

The model structure associated to the system (2.11) is known as a *hybrid Box-Jenkins* (hybrid BJ) structure, and it will be considered in Chapters 7, 8 and 10.

State-space descriptions

An alternative to (2.10) is to consider a system described in a state-space form. Given the difficulties associated with directly sampling continuous-time white noise [9], one option is to include a disturbance component only in the state equation, leading to

$$\begin{aligned} d\mathbf{x}(t) &= \mathbf{A}_c \mathbf{x}(t) dt + \mathbf{B}_c u(t) dt + d\mathbf{w}(t), \\ y(t) &= \mathbf{C}_c \mathbf{x}(t) + D_c u(t), \end{aligned} \quad (2.13)$$

where $\mathbf{A}_c, \mathbf{B}_c, \mathbf{C}_c$ and D_c are matrices of a suitable dimension, $\mathbf{x}(t)$ is the state vector, and $\mathbf{w}(t)$ is a Wiener process of finite incremental covariance. If noise is also considered in the output equation, then it is usually represented in discrete-time when only samples are accessible [10]:

$$y(t_k) = \mathbf{C}_c \mathbf{x}(t_k) + D_c u(t_k) + v(t_k), \quad (2.14)$$

where $v(t_k)$ is a zero-mean discrete-time stochastic process. Continuous-time systems given by (2.13) and (2.14) can be described exactly (in terms of the second-order moments of the output) by an equivalent discrete-time model by exploiting the same principle as in (2.8). We refer to [144] for more details.

A fully continuous-time output equation can also be considered in a linear stochastic differential equations description:

$$d\mathbf{x}(t) = \mathbf{A}_c \mathbf{x}(t) dt + \mathbf{B}_c u(t) dt + d\mathbf{w}(t), \quad (2.15a)$$

$$dy(t) = \mathbf{C}_c \mathbf{x}(t) dt + D_c u(t) dt + dv(t), \quad (2.15b)$$

with $v(t)$ being another Wiener process of finite incremental covariance. Sampling the output equation now requires an anti-aliasing step, resulting in what is called integrated sampling [144]. Identification in this framework is covered in, e.g., [260], where the data is assumed to be fast sampled so that the identified matrices in the δ -domain approximate well the true continuous-time state-space matrices. A more general setup, consisting of partially observed nonlinear stochastic differential equations, is considered in [124]. Other interesting results about identification of linear stochastic differential equations can be found in [126].

Identifiability and persistence of excitation

Experimental conditions are critical for the performance of identification methods for continuous and discrete-time systems. In this regard, a reasonable wish is to design the model structure and input signal such that the true parameter vector is recovered when sufficient samples of data are retrieved from the system in the absence of noise. In terms of the model structure, the question we need to address is whether different values of $\boldsymbol{\theta}$ can give equal models. If this generally does not occur, we say that the model structure is *identifiable*. A more precise definition of identifiability for a hybrid BJ model structure is given below, which is in line with the definition in [138] for a discrete-time BJ model structure.

Definition 2.1 (Identifiability for a hybrid BJ model structure). Denote $G(p, \boldsymbol{\theta})$ and $H(q, \boldsymbol{\eta})$ as the system and noise models parameterized by $\boldsymbol{\theta}$ and $\boldsymbol{\eta}$, respectively. A model structure is *globally identifiable* at $[\boldsymbol{\theta}^{*\top}, \boldsymbol{\eta}^{*\top}]$ if and only if

$$G(s, \boldsymbol{\theta}) = G(s, \boldsymbol{\theta}^*) \text{ and } H(z, \boldsymbol{\eta}) = H(z, \boldsymbol{\eta}^*) \text{ for almost all } s, z \implies [\boldsymbol{\theta}^\top, \boldsymbol{\eta}^\top] = [\boldsymbol{\theta}^{*\top}, \boldsymbol{\eta}^{*\top}].$$

In terms of the data retrieved from an identification experiment, we require it to be “informative enough” with respect to the model structure we use. For this to happen, the input must be *persistently exciting* of a certain degree. This concept is usually studied in inputs that are *quasi-stationary* [138]:

Definition 2.2 (Quasi-stationary signals). A signal $\{u(t_k)\}$ is said to be quasi-stationary if it is subject to the following conditions:

1. $\mathbb{E}\{u(t_k)\}$ is well defined and bounded for all k .
2. $\mathbb{E}\{u(t_k)u(t_r)\}$ is well defined and bounded for all k, r .
3. $\overline{\mathbb{E}}\{u(t_k)u(t_{k-\tau})\} = r_u(\tau)$ is well defined and only depends on τ , where

$$\overline{\mathbb{E}}\{\cdot\} := \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N \mathbb{E}\{\cdot\}.$$

Note that the spectrum $\Phi_u(\omega)$ of a quasi-stationary signal $u(t_k)$ is given by the discrete-time Fourier transform (DTFT) of $r_u(\tau)$.

Definition 2.3 (Persistence of excitation). A quasi-stationary signal $u(t_k)$, with spectrum $\Phi_u(\omega)$, is said to be *persistently exciting of order n* if, for all filters of the form $M_n(q) = m_1q^{-1} + \dots + m_nq^{-n}$, the relation $|M_n(e^{i\omega})|^2\Phi_u(\omega) = 0$ implies that $M_n(e^{i\omega}) \equiv 0$.

Note that the concept of persistence of excitation can also be interpreted in terms of the number of spectral lines of the spectrum of the input: $u(t_k)$ is persistently exciting of order n if $\Phi_u(\omega)$ is different from zero on at least n points in $[-\pi, \pi]$.²

2.3 Continuous-time identification methods

Continuous-time system identification can be undertaken in both time and frequency domains. We focus on methods based on time-domain data, even though there is an important body of work dedicated to continuous-time system identification in the frequency domain. An overview of popular methods for frequency-domain system identification is presented in [181], in which the frequency-domain Gaussian maximum likelihood estimator is suggested. The estimation of continuous-time power spectral densities has also been studied from the frequency-domain perspective in [72].

There are two main approaches in continuous-time system identification in the time domain, namely the indirect and direct approaches [186]. In the indirect approach, a discrete-time model is first identified using measured input and output data and then it is transformed into a continuous-time form. On the other hand, the direct approach identifies a continuous-time model directly from sampled data, thus avoiding the intermediate discrete-time model. Point estimators are used in both approaches for generating an estimate of $\boldsymbol{\theta}^*$ in (2.6). Before analyzing the indirect and direct approaches in more detail, we will first present some important asymptotic properties that good point estimators should enjoy. Afterwards, we will discuss the most popular methods in each approach. These methods and statistical concepts are important for the development of this thesis, as we cover both indirect (Chapter 3 and part of Chapters 9 and 10) and direct approaches (Chapters 4 to 8), and study their asymptotic properties. Since the indirect methods we study in Chapter 3 are constructed by considering the output error model structure, we focus our attention only to this form in the review of the indirect approach. Extensions of the indirect approach for more general model structures can be readily derived from, e.g., [138, 211].

2.3.1 Properties of estimators

In any parametric continuous-time identification method, a set of noisy input and output data that was obtained from the system is used to estimate the true parameter

²For simplicity we assume that $h = 1$. In the general case, the spectral lines can be found in $[-\pi/h, \pi/h]$.

vector $\boldsymbol{\theta}^*$ in (2.6). A *point estimator* is a function that maps the observed data into an estimate $\hat{\boldsymbol{\theta}}_N \in \Theta$ of $\boldsymbol{\theta}^*$, where N is the number of input and/or output data points used to form the estimate, and Θ is the parameter space. We will simply refer to them as *estimators* since only point estimators will be covered here. Their dependence on N will be made explicit in this section only for clarity purposes.

When studying the properties of an estimator, it is important to understand the relationship between $\hat{\boldsymbol{\theta}}_N$ and $\boldsymbol{\theta}^*$ for an increasing number of data points N . Since the parameter $\boldsymbol{\theta}^*$ is unknown, it is desired that the estimator enjoys properties that hold for a large set of possible true parameters. One of these properties is *consistency*, which is defined below.

Definition 2.4 (Consistency). An estimator sequence $\{\hat{\boldsymbol{\theta}}_N\}_{N=1}^\infty$ of a parameter $\boldsymbol{\theta}^*$ is *consistent* if

$$\hat{\boldsymbol{\theta}}_N \xrightarrow{a.s.} \boldsymbol{\theta}^* \quad \text{for all } \boldsymbol{\theta}^* \in \Theta$$

as $N \rightarrow \infty$, where $\xrightarrow{a.s.}$ denotes almost sure convergence [28].

Remark 2.2. Some authors make the distinction between weak and strong consistency depending on whether the convergence is in probability or almost surely respectively, and define consistency with convergence in probability [130]. We follow the definition in [211], where consistency is defined with almost sure convergence. Also, when analyzing asymptotic properties we shall frequently use the term *estimator* instead of the more accurate but cumbersome term *estimator sequence*.

It may be possible that, given an identification method, its consistency depends on certain factors of the identification experiment such as the system or noise model parameters. In some cases, estimators can be proven to be “almost always consistent”, which means that consistency will be achieved for all except (possibly) certain pathological cases. Such estimator is referred to as being *generically consistent*. We provide the precise definition of this property in Definitions 2.5 and 2.6.

Definition 2.5 (Generically true statement [210]). Let Ω be an open set in some Euclidean space \mathbb{R}^n . A statement s , which depends on the elements $\boldsymbol{\rho}$ of Ω , is *generically true with respect to Ω* if the set $M = \{\boldsymbol{\rho} \in \Omega \mid s(\boldsymbol{\rho}) \text{ is not true}\}$ has Lebesgue measure zero in Ω .

Definition 2.6 (Generic consistency). An estimator $\hat{\boldsymbol{\theta}}_N$ of a parameter $\boldsymbol{\theta}^*$ is *generically consistent* with respect to a set of possible values of $\boldsymbol{\theta}^* \in \Omega$ if the statement $s = \{\hat{\boldsymbol{\theta}}_N \text{ is consistent}\}$ is generically true with respect to Ω .

When assessing the precision of an estimator, the requirement of achieving a uniformly minimum mean square error is typically too stringent and unrealistic [95]. Therefore, it is common to concentrate on the class of consistent estimators of the unknown parameter $\boldsymbol{\theta}^*$, and on the *asymptotic covariance matrix* of those estimators, which is given by

$$\text{AsCov}(\hat{\boldsymbol{\theta}}_N) := \lim_{N \rightarrow \infty} N \mathbb{E}\{(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}^*)(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}^*)^\top\}. \quad (2.16)$$

Our interest is to construct consistent estimators that have the least asymptotic covariance possible. To formalize this goal, we first define a property that characterizes all consistent estimators that have optimal dispersion properties.

Definition 2.7 (Asymptotic efficiency). A consistent estimator $\hat{\theta}_N$ is said to be *asymptotically efficient* if, for all other estimators $\tilde{\theta}_N$ of θ^* , it holds that

$$\lim_{N \rightarrow \infty} N \mathbb{E}\{(\tilde{\theta}_N - \theta^*)(\tilde{\theta}_N - \theta^*)^\top\} \succeq \lim_{N \rightarrow \infty} N \mathbb{E}\{(\hat{\theta}_N - \theta^*)(\hat{\theta}_N - \theta^*)^\top\},$$

assuming the limits exist.

One way to show that an estimator is asymptotically efficient is to derive a lower bound on the asymptotic covariance of any consistent estimator, and later prove that this bound is achieved for that estimator. It can be shown (see, e.g., [130, Thm 2.6., p. 440]) that, under mild conditions, the asymptotic covariance (2.16) of a consistent estimator is lower bounded in a positive semi-definite sense by

$$\mathbf{P}_{\text{CR}} = \left[\lim_{N \rightarrow \infty} \frac{1}{N} \mathbb{E} \left\{ \left(\frac{\partial \log p(\mathbf{y}^N; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right) \left(\frac{\partial \log p(\mathbf{y}^N; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right)^\top \right\} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*} \right]^{-1}, \quad (2.17)$$

for all values of $\boldsymbol{\theta}^*$, possibly except for a set of Lebesgue measure zero. Here \mathbf{y}^N is the vector of the full data $y(t_1), \dots, y(t_N)$, and $p(\mathbf{y}^N; \boldsymbol{\theta})$ is the probability density function (PDF) of \mathbf{y}^N , which is parameterized by the vector $\boldsymbol{\theta}$. We call (2.17) the *asymptotic Cramér-Rao lower bound* (CRLB) on $\text{AsCov}(\hat{\theta}_N)$. The expression inside the inverse in (2.17) is known as the Fisher information matrix per sample.

Thus, if the estimator $\hat{\theta}_N$ is consistent and its asymptotic covariance coincides with the CRLB, then the estimator is asymptotically efficient.

2.3.2 Indirect approach

One approach to identify a continuous-time system is to first estimate the discrete-time model given the input and output data samples, and then translate this model into continuous-time. This is referred to as the indirect approach, and it relies on discrete-time rather than continuous-time system identification techniques to obtain a continuous-time model.

First step: discrete-time system identification

The first step of the indirect approach consists in obtaining a discrete-time model from the sampled data. Much has been written regarding this step [138, 211]. In the following, we will only focus on the most common methods for performing discrete-time system identification, which are the prediction error method (PEM) and the maximum likelihood (ML) estimator. These two estimators are closely linked, as we will see.

Before choosing a method, a discrete-time model structure should be selected. Prior knowledge, in conjunction with non-parametric methods and model selection tools such as the AIC [4], BIC [197] or MDL [188] criteria, can be used for choosing a model structure. Note that only the number of poles of the discrete-time model must be chosen at this stage. That is, the following fully-parameterized, causal, discrete-time model structure is usually proposed:

$$H(q) = \frac{\beta_{n-1}q^{n-1} + \beta_{n-2}q^{n-2} + \cdots + \beta_0}{q^n + \alpha_{n-1}q^{n-1} + \cdots + \alpha_1q + \alpha_0}, \quad (2.18)$$

where we denote the discrete-time parameter vector as

$$\boldsymbol{\rho} := [\alpha_0, \alpha_1, \dots, \alpha_{n-1}, \beta_0, \beta_1, \dots, \beta_{n-1}]^\top.$$

In contrast to the direct approach, in the indirect approach it is convenient to always estimate a discrete-time transfer function with relative degree less than or equal to one. This insight is derived from the fact that, when sampling a strictly proper continuous-time system through a ZOH, its equivalent discrete-time model will usually have relative degree equal to one, independent of the relative degree of the underlying continuous-time system. This claim is more precisely stated in the following proposition.

Proposition 2.1. Consider a strictly proper, LTI, continuous-time system $G(p)$. The relative degree of the ZOH equivalent of $G(p)$ is $r \geq 1$ if and only if $g(h) = g(2h) = \cdots = g([r-1]h) = 0$ and $g(rh) \neq 0$, where $g(t)$ is the step response of $G(p)$ and h is the sampling period.

Proof. According to (2.8), the ZOH equivalent of $G(p)$ can be computed by

$$\begin{aligned} H(z) &= \frac{z-1}{z} \mathcal{Z}\{g(kh)\} \\ &= \frac{z-1}{z} \left(\frac{g(h)}{z} + \frac{g(2h)}{z^2} + \cdots \right) \\ &= \frac{g(h)}{z} + \frac{g(2h)-g(h)}{z^2} + \frac{g(3h)-g(2h)}{z^3} + \cdots. \end{aligned}$$

Therefore, the transfer function $H(z)$ has relative degree r if and only if

$g(h) = g(2h) - g(h) = \cdots = g([r-1]h) - g([r-2]h) = 0$ and $g(rh) - g([r-1]h) \neq 0$,

from which the statement follows. \square

After deciding the model structure, the estimation method must be chosen; one possibility is the widely popular PEM method. The idea of PEM is to pick a parameter vector estimate that minimizes a cost function of the prediction errors—that is, the difference between the observed output data at a certain time and its predicted value with a certain model, based on past data.

In other words, we shall select the parameter vector that minimizes the cost

$$V_N(\boldsymbol{\rho}) = \frac{1}{N} \sum_{k=1}^N \ell(\varepsilon(t_k, \boldsymbol{\rho})),$$

where $\ell(\cdot)$ is a scalar non-negative function and $\varepsilon(t_k, \boldsymbol{\rho})$ is the prediction error at time t_k , which depends on the discrete-time model parameter vector $\boldsymbol{\rho}$. For a quadratic cost function and a causal system operating in an output error framework, we must minimize

$$V_N(\boldsymbol{\rho}) = \frac{1}{N} \sum_{k=1}^N \left(y(t_k) - \frac{B_d(q, \boldsymbol{\rho})}{A_d(q, \boldsymbol{\rho})} u(t_k) \right)^2. \quad (2.19)$$

Thus, the estimate of the underlying discrete-time system is computed as

$$\hat{\boldsymbol{\rho}}_N = \arg \min_{\boldsymbol{\rho}} V_N(\boldsymbol{\rho}).$$

The PEM estimator is closely linked to the ML estimator, which can also be used to obtain a discrete-time model for the indirect approach. The idea of ML is to describe the observed data set $\mathbf{y}^N := [y(t_1), \dots, y(t_N)]^\top$ as a realization of a stochastic vector, which has a certain PDF. Then, the ML estimate is the parameter vector that maximizes the likelihood that the data takes the observed values.

More formally, the ML estimator is given by

$$\begin{aligned} \hat{\boldsymbol{\rho}}_N &= \arg \max_{\boldsymbol{\rho}} p(\mathbf{y}^N; \boldsymbol{\rho}) \\ &= \arg \max_{\boldsymbol{\rho}} \log p(\mathbf{y}^N; \boldsymbol{\rho}), \end{aligned}$$

where the second equality holds because $\log(\cdot)$ is a monotonically increasing function. If the prediction errors $\varepsilon(t_k, \boldsymbol{\rho})$ are assumed to be zero-mean and Gaussian-distributed with variance σ^2 , then

$$\begin{aligned} \log p(\mathbf{y}^N; \boldsymbol{\rho}) &= \sum_{k=1}^N \log p(\varepsilon(t_k, \boldsymbol{\rho}); \boldsymbol{\rho}, \sigma^2) \\ &= -\frac{N}{2} \log(2\pi) - \frac{1}{2} \log(\sigma^2) - \frac{1}{2\sigma^2} \sum_{k=1}^N \varepsilon^2(t_k, \boldsymbol{\rho}). \end{aligned} \quad (2.20)$$

Regardless of having a known or unknown noise variance σ^2 , the ML estimator is in fact minimizing the quadratic cost in (2.19) for an output error model structure, which means that this estimator coincides with the PEM estimator in this case.

The ML and PEM estimators are known to give consistent and asymptotically efficient estimates under general conditions. In [14], ML was introduced for ARMAX models, and its consistency, asymptotic efficiency and normality were studied. Asymptotic properties of PEM were obtained in [27], and results on the asymptotic accuracy of approximate prediction error methods were developed in [142]. The main ideas behind the proofs of these properties can be found in [138, Ch. 9].

Second step: Transformation into a continuous-time model

The next step is to transform this discrete-time transfer function into an adequate continuous-time model. The return from the conventional discrete-time model to a continuous-time model is not possible without assumptions on the behavior of the signals between the sampling instants. For example, if it is known that the input has a ZOH intersample behavior, then the exact transformation (at the sampling instants) is the inverse ZOH mapping, i.e., the inverse operation of (2.8). This transformation is unique and straightforward when there is no negative real pole in the discrete-time transfer function and the angular sampling frequency is larger than twice the largest imaginary part of the continuous-time transfer function poles [122]. Usually it is done by taking the matrix logarithm of a partitioned state-space matrix [16]. More precisely, if \mathbf{A}_c and \mathbf{B}_c are the state transition matrices of the continuous-time system and \mathbf{A}_d and \mathbf{B}_d are their discrete-time counterparts, then

$$\begin{bmatrix} \mathbf{A}_c & \mathbf{B}_c \\ \mathbf{0} & \mathbf{0} \end{bmatrix} = \log \left(\begin{bmatrix} \mathbf{A}_d & \mathbf{B}_d \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \right). \quad (2.21)$$

If there is a negative real pole in the discrete-time model, then the partitioned matrix in the right hand side of (2.21) has at least one negative eigenvalue, which leads to the transformation not being unique. In fact, it can be shown that a discrete-time model with poles on the negative real axis may be interpreted as a sampled (at the Nyquist frequency) continuous-time system of a degree equal to the number of discrete poles, where poles on the negative real axis are counted twice [189]. This means that if a pole on the negative real axis is obtained, then the continuous-time system should be sampled faster to overcome missing information due to sampling at the Nyquist frequency.

If the intersample behavior of the input is not known, then it might be useful to extract data with a small sampling period to reduce the reconstruction error. However, this may lead to numerical issues in the z -domain as the estimated poles will all be mapped close to the point $(1, 0i)$ in the z -plane. A possible solution to this problem is provided by the delta operator [154], which provides an approximation of the derivative operator. By converting the discrete-time model into a model in the δ -domain, the models obtained through the delta operator converge to their original continuous-time models as the sampling period approaches zero [186]. Thus, the model where $\delta \approx p$ can be a reasonable continuous-time model of the underlying continuous-time system for fast sampling rates. Note that it is also possible to directly estimate the equivalent system in the δ -domain, as done in [46].

Remark 2.3. Changing the approximation we use for the derivative leads to different δ operators, which differ in which frequency bands are more accurately represented. Different approximations can be found in Table 2.1, as reported in [186].

Table 2.1: Different equivalences between the δ and q operators. Each relationship maps the stability region differently, and causes different frequency distortions. A comparison of these approximations is provided in, e.g., [16, Ch. 8].

Approximation of the derivative operator	δ
Forward Euler	$(q - 1)/h$
Backward Euler	$(q - 1)/(hq)$
Tustin	$(2/h)(q - 1)/(q + 1)$

Benefits and drawbacks of the indirect approach

As with any method, the estimators relying on the indirect approach have both benefits and drawbacks. On a positive note, the indirect approach is based on solid theoretical foundations, as most of the available system identification methods have been investigated in discrete-time scenarios [138, 211]. Also, even though continuous-time models are ultimately estimated, no handling of derivatives of inputs or outputs is needed.

There are also reasons why this approach is not always recommended. First, it may suffer from numerical inaccuracies at fast sampling rates, and it requires a precise initialization. Non-uniformly sampled signals also carry a problem, since discrete-time models would necessarily be time-varying [59]. In addition, it is not possible to easily select the desired numerator order of the continuous-time model, as the estimated discrete-time model will generally lead to a continuous-time model with relative degree equal to one in the case of sampling by a zero-order hold mechanism (see Proposition 2.1).

2.3.3 Direct approach

In direct continuous-time system identification, a continuous-time model is obtained from sampled data without an intermediate discrete-time model. Just as for the indirect approach, there are many procedures that have been developed (see, e.g., [151, 186, 192]). These methods have achieved remarkable success in many practical applications [61, 165, 250, 254]. This subsection provides a description of some of the most popular methods in the direct approach, which are based on least-squares and instrumental variables. It is by no means an exhaustive review of all the methods available; the main goal is to introduce the reader only to the methods we study carefully in this thesis.

We describe the following two methods as presented in [62].

Least-squares state-variable filter method

One of the main problems in the direct approach is how to reconstruct time-derivatives. If the derivatives of the input and output were known from the identification experiment, then estimating the parameters θ^* could be done by following

a least-squares approach to (2.2). This idea gave birth to the least-squares state-variable filter (LSSVF) method, which was one of the first methods ever developed for continuous-time system identification [241].

Let us consider the differential equation in (2.2) written in operator form:

$$a_n p^n y(t) + \cdots + a_1 p y(t) + y(t) = b_m p^m u(t) + \cdots + b_0 u(t). \quad (2.22)$$

Since only (possibly noisy) samples of $u(t)$ and $y(t)$ are retrieved, their derivatives $p^i y(t)$ and $p^l u(t)$ for $i = 1, \dots, n$ and $l = 1, \dots, m$ are not directly available. Even if a suitable intersample behavior were specified, the high-order derivatives of $y(t)$ and $u(t)$ might be highly inaccurate. Thus, instead of this, we apply a continuous-time low-pass filter $F(p)$ (i.e., a “state-variable filter”) to both sides of (2.22), resulting in

$$\begin{aligned} a_n p^n F(p) y(t) + \cdots + a_1 p F(p) y(t) + F(p) y(t) \\ = b_m p^m F(p) u(t) + \cdots + b_0 F(p) u(t), \end{aligned} \quad (2.23)$$

where we have used the fact that $F(p)$ commutes with $A(p)$ and $B(p)$ since the system is assumed to be time-invariant. The filter $F(p)$ is typically chosen to be a low-pass filter with a DC gain equal to one of the form $F(p) = \lambda^n / (p + \lambda)^n$, where $\lambda > 0$ is a user-defined breakpoint frequency. From now on, we denote the filtered reconstructed input $F(p)u(t)$ and filtered reconstructed output $F(p)y(t)$ as $u_f(t)$ and $y_f(t)$, and their i -th derivatives as $u_f^{(i)}(t)$ and $y_f^{(i)}(t)$, respectively.

Note that the filter $F(p)$ must have relative degree greater than or equal to n , so that the transfer functions $p^i F(p)$, $i = 1, \dots, n$, are all proper. This permits the filtering of the sampled input and output signals through causal continuous-time filters. Since these signals are only known at specific time instants t_k , it is necessary to specify or assume an intersample behavior of the input and output in order to perform the continuous-time prefiltering. Once the intersample behavior is assumed, the (causal) filtering can be done in continuous-time and the resultant signal can then be evaluated at t_k . If the sampling is done evenly in time and a ZOH or FOH assumption is used, then this process can be done fully in discrete-time by applying the ZOH or FOH equivalents of the continuous-time filters to the discrete-time data. On the other hand, if non-uniformly sampled data is considered, then an adaptive ordinary differential equation solver, such as the Dormand-Prince method [45], can be used to obtain accurate values of the filtered signals [118].

If additive noise is present in the measurements, we can rewrite (2.23) as a standard linear regression

$$y_f(t_k) = \varphi_f^\top(t_k) \boldsymbol{\theta} + \varepsilon(t_k), \quad (2.24)$$

with $\varphi_f(t_k)$ being the *filtered regressor vector*, which is given by

$$\varphi_f(t_k) = \left[-y_f^{(1)}(t_k), \dots, -y_f^{(n)}(t_k), u_f(t_k), u_f^{(1)}(t_k), \dots, u_f^{(m)}(t_k) \right]^\top,$$

while $\boldsymbol{\theta}$ is described by (2.7). Now, from N available samples of the input and output signals observed at times t_1, \dots, t_N , the LSSVF method delivers estimates that are computed by the least-squares formula

$$\hat{\boldsymbol{\theta}}^{\text{LS}} = \left[\frac{1}{N} \sum_{k=1}^N \boldsymbol{\varphi}_f(t_k) \boldsymbol{\varphi}_f^\top(t_k) \right]^{-1} \left[\frac{1}{N} \sum_{k=1}^N \boldsymbol{\varphi}_f(t_k) y_f(t_k) \right]. \quad (2.25)$$

Remark 2.4. The LSSVF method is closely related to the generalized Poisson moment functionals approach, which was developed later [194]. In fact, the minimal-order generalized Poisson moment functional method may be considered as a particular case of the LSSVF method [65]. Also, note that a $1/N$ factor is included in both sums in (2.25). These factors are useful when analyzing the asymptotic properties of the method, although they are not strictly needed when describing the estimator. From now on we will avoid the inclusion of $1/N$ in all the methods unless necessary.

The LSSVF method provides a simple way to obtain a rough estimate of the continuous-time system, and thus it can be used as an initialization estimate for other algorithms [250]. The intersample behavior of the input and output must be specified in this method, and λ is a tuning parameter that is typically chosen such that the bandwidth of the system of interest is contained in the pass-band of $F(p)$. One major drawback of this method is that, unless the additive noise $\varepsilon(t_k)$ in (2.24) is white, it is known not to be consistent. To see this, note that if the true system has the form (2.10) where $v(t_k)$ is white noise stochastic process, we can then write

$$y_f(t_k) = \boldsymbol{\varphi}_f^\top(t_k) \boldsymbol{\theta}^* + v_f(t_k), \quad (2.26)$$

where $v_f(t_k) = F(p)v(t_k)$. Thus, the bias of the LSSVF estimator is given by

$$\hat{\boldsymbol{\theta}}^{\text{LS}} - \boldsymbol{\theta}^* = \left[\sum_{k=1}^N \boldsymbol{\varphi}_f(t_k) \boldsymbol{\varphi}_f^\top(t_k) \right]^{-1} \left[\sum_{k=1}^N \boldsymbol{\varphi}_f(t_k) v_f(t_k) \right].$$

As N tends to infinity, under stationarity conditions we can see from the expression above that the bias will asymptotically depend on $\mathbb{E}\{\boldsymbol{\varphi}_f(t_k)v_f(t_k)\}$. Since $\boldsymbol{\varphi}_f(t_k)$ is a function of the output measurements $y(t_k)$, which in turn depend on the noise signal $v(t_k)$ and its previous values, this expectation is in general different from zero.

One possible solution to the asymptotic bias problem is to use instrumental variable methods, as we will discuss next.

Simplified refined instrumental variables

We will first introduce the idea of instrumental variables in the context of continuous-time system identification, as a prequel to presenting the simplified refined instrumental variable method for continuous-time systems (SRIVC). In the basic instrumental

variable method [243], the parameter vector is estimated by

$$\hat{\boldsymbol{\theta}}^{\text{IV}} = \left[\sum_{k=1}^N \hat{\boldsymbol{\varphi}}_f(t_k) \boldsymbol{\varphi}_f^\top(t_k) \right]^{-1} \left[\sum_{k=1}^N \hat{\boldsymbol{\varphi}}_f(t_k) y_f(t_k) \right],$$

where $\hat{\boldsymbol{\varphi}}_f(t_k)$ is called the *filtered instrument vector* or vector of filtered instruments [210]. By leveraging (2.26), the bias of this estimator satisfies

$$\hat{\boldsymbol{\theta}}^{\text{IV}} - \boldsymbol{\theta}^* = \left[\sum_{k=1}^N \hat{\boldsymbol{\varphi}}_f(t_k) \boldsymbol{\varphi}_f^\top(t_k) \right]^{-1} \left[\sum_{k=1}^N \hat{\boldsymbol{\varphi}}_f(t_k) v_f(t_k) \right].$$

We denote the matrix being inverted above as the modified normal matrix, as it resembles the normal matrix used in Gauss-Newton methods [161]. The key idea behind the instrumental variable approach is to choose $\hat{\boldsymbol{\varphi}}_f(t_k)$ such that the expected value $\mathbb{E}\{\hat{\boldsymbol{\varphi}}_f(t_k) v_f(t_k)\}$ is equal to zero, thus yielding a consistent estimator if the modified normal matrix $\mathbb{E}\{\hat{\boldsymbol{\varphi}}_f(t_k) \boldsymbol{\varphi}_f^\top(t_k)\}$ is non-singular. That is, we should consider instruments that are correlated with the regressors $\boldsymbol{\varphi}_f(t_k)$, while being uncorrelated with the (prefiltered) noise sequence $v_f(t_k)$. One possible instrument vector can be

$$\hat{\boldsymbol{\varphi}}_f(t_k) = \left[-\hat{x}_f^{(1)}(t_k), -\hat{x}_f^{(2)}(t_k), \dots, -\hat{x}_f^{(n)}(t_k), u_f(t_k), \dots, u_f^{(m)}(t_k) \right]^\top,$$

where $\hat{x}_f(t_k)$ is the estimated filtered noise-free sampled output, which is given by

$$\hat{x}_f(t_k) = F(p) \frac{B(p, \hat{\boldsymbol{\theta}}^{\text{LS}})}{A(p, \hat{\boldsymbol{\theta}}^{\text{LS}})} u(t_k), \quad (2.27)$$

with $\hat{\boldsymbol{\theta}}^{\text{LS}}$ denoting the LSSVF estimate.

Remark 2.5. In (2.27) we have introduced a mixed notation of continuous-time filters with discrete-time signals that will be used throughout the thesis. If $G(p)$ is a continuous-time filter and $\{x(t_k)\}$ is a sampled signal, then $G(p)x(t_k)$ implies that the signal $x(t_k)$ is interpolated using either a ZOH or FOH, and the resulting output of the filter is sampled at t_k . On the other hand, the expression $\{G(p)x(t)\}_{t=t_k}$ (or $[G(p)x(t)]_{t=t_k}$ in the vector-valued case) means that the continuous-time signal $x(t)$ is filtered through $G(p)$ and later sampled at $t = t_k$. Note that the implementation of these filtering operations can have an important impact on the performance of the estimator. In particular, if the intersample behavior of the input is not known, a high sampling frequency is often required for minimizing the effect of misspecifications of the behavior in between samples [62].

The final step is to choose $F(p)$ properly. For this we can shift our attention to the prediction error method. For the output error case, PEM tries to minimize a least-squares criterion of the error function (cf. (2.19))

$$\varepsilon(t_k) = y(t_k) - \frac{B(p)}{A(p)} u(t_k). \quad (2.28)$$

In the continuous-time system identification literature, the expression in (2.28) is commonly called the *generalized equation error* (GEE, [244]). The GEE can also be written as

$$\varepsilon(t_k) = A(p) \frac{1}{A(p)} y(t_k) - B(p) \frac{1}{A(p)} u(t_k), \quad (2.29)$$

or equivalently,

$$\varepsilon(t_k) = y_f(t_k) + a_1 y_f^{(1)}(t_k) + \cdots + a_n y_f^{(n)}(t_k) - b_0 u_f(t_k) - \cdots - b_m u_f^{(m)}(t_k), \quad (2.30)$$

where the continuous-time filter $F(p)$ was chosen as $1/A(p)$. Unfortunately, this filter depends on the parameters we are trying to estimate. Thus, instead of a standard linear regression, we have derived a *pseudo-linear regression* [214] of the form

$$y_f(t_k) = \boldsymbol{\varphi}_f^\top(t_k) \boldsymbol{\theta} + \varepsilon(t_k), \quad (2.31)$$

where *pseudo* refers to the fact that the parameter vector $\boldsymbol{\theta}$ is also present in the regressor vector $\boldsymbol{\varphi}_f(t_k)$ and in the filtered output $y_f(t_k)$.

One problem with this formulation is that the filter $F(p)$ is not known a priori. Also, we cannot always assume that the measurement noise in (2.10b) is white; it is likely that it is in fact colored. These issues are simultaneously solved by proposing an iterative instrumental variable method which in its basic embodiment is called the SRIVC method. To solve the first problem, the SRIVC estimator uses the previous estimate of the polynomial $A(p, \boldsymbol{\theta}_j)$ to form $F(p)$, which transforms the pseudo-linear regression in (2.31) into a simple linear regression problem. Instead of least squares, an instrumental variable approach is used to solve the possible bias problems with colored noise. Once the estimate $\boldsymbol{\theta}_{j+1}$ is obtained, it is used to improve the estimate for $F(p)$ and the process is repeated until a maximum number of iterations is reached or until the relative error between the previous and current estimates is smaller than a preset tolerance factor.

Initialization of the SRIVC estimator involves the selection of a suitable initial prefilter polynomial $1/A(p, \boldsymbol{\theta}_1)$. This can be done by obtaining a model estimate through the LSSVF method previously described, which requires the specification of the bandwidth λ . Alternatively, the SRIVC estimator can be initialized with the continuous-time equivalent of a model estimate provided by an algorithm for discrete-time model estimation. The discrete-time version of SRIVC, SRIV, is the preferred option in [59] for regularly sampled data. Further discussion on initialization aspects can be found in [253].

A detailed description of the SRIVC estimator can be found in Algorithm 2.1. This estimator will be thoroughly analyzed in Chapters 4, 5 and 7, and extensions of this method are proposed in Chapters 6 and 8. Note that it is important to reflect the unstable poles of the estimates at each iteration so that the filtering procedure at the next iteration is well conditioned. This reflection step, described in lines 7 to 14 of Algorithm 2.1, limits the method to provide accurate estimates for stable continuous-time systems only.

Algorithm 2.1: Simplified refined instrumental variable method for continuous-time systems (SRIVC)

- 1: Input: $\{u(t_k), y(t_k)\}_{k=1}^N$, model order (n, m) , initial vector estimate $\boldsymbol{\theta}_1$, tolerance factor ϵ and maximum number of iterations M
- 2: Using $\boldsymbol{\theta}_1$, form the estimated system denominator polynomial $A_1(p)$
- 3: $j \leftarrow 1$, flag $\leftarrow 1$
- 4: **while** flag = 1 and $j \leq M$ **do**
- 5: Prefilter the sampled input $\{u(t_k)\}_{k=1}^N$ and output $\{y(t_k)\}_{k=1}^N$ to form

$$\boldsymbol{\varphi}_f(t_k) \leftarrow \left[\frac{-p}{A_j(p)} y(t_k), \dots, \frac{-p^n}{A_j(p)} y(t_k), \frac{1}{A_j(p)} u(t_k), \dots, \frac{p^m}{A_j(p)} u(t_k) \right]^\top, \quad (2.32)$$

$$\hat{\boldsymbol{\varphi}}_f(t_k) \leftarrow \left[\frac{-pB_j(p)}{A_j^2(p)} u(t_k), \dots, \frac{-p^n B_j(p)}{A_j^2(p)} u(t_k), \frac{1}{A_j(p)} u(t_k), \dots, \frac{p^m}{A_j(p)} u(t_k) \right]^\top, \quad (2.33)$$

$$y_f(t_k) \leftarrow \frac{1}{A_j(p)} y(t_k) \quad (2.34)$$

- 6: Compute the parameter estimate

$$\boldsymbol{\theta}_{j+1} \leftarrow \left[\sum_{k=1}^N \hat{\boldsymbol{\varphi}}_f(t_k) \boldsymbol{\varphi}_f^\top(t_k) \right]^{-1} \left[\sum_{k=1}^N \hat{\boldsymbol{\varphi}}_f(t_k) y_f(t_k) \right] \quad (2.35)$$

- 7: **if** $1/A_{j+1}(p)$ is unstable **then**
 - 8: **for** $i = 1, \dots, n$ **do**
 - 9: **if** pole p_i of $1/A_{j+1}(p)$ is unstable **then**
 - 10: $\text{Re}\{p_i\} \leftarrow -\text{Re}\{p_i\}$
 - 11: **end if**
 - 12: **end for**
 - 13: $A_{j+1}(p) \leftarrow \prod_{i=1}^n (p - p_i) / \prod_{i=1}^n (-p_i)$
 - 14: **end if**
 - 15: **if** $\frac{\|\boldsymbol{\theta}_{j+1} - \boldsymbol{\theta}_j\|_2}{\|\boldsymbol{\theta}_j\|_2} < \epsilon$ **then**
 - 16: flag $\leftarrow 0$
 - 17: **end if**
 - 18: $j \leftarrow j + 1$
 - 19: **end while**
 - 20: Output: $\boldsymbol{\theta}_j$ and its associated model $B_j(p)/A_j(p)$.
-

Refined instrumental variables for hybrid BJ models

An extension of the SRIVC estimator for hybrid BJ models, called RIVC, has been suggested in [255]. Each iteration of this procedure consists in a noise model estimation step followed by a system parameter estimation step using instruments and regressors that are computed from a bank of prefilters that depend on the current system and noise models.

If the system is of the form (2.11), the RIVC estimator seeks to minimize a least-squares criterion of the following error function:

$$\begin{aligned}\varepsilon(t_k) &= \frac{D(q)}{C(q)} \left(A(p) \frac{1}{A(p)} y(t_k) - B(p) \frac{1}{A(p)} u(t_k) \right) \\ &= y_f(t_k) + a_1 y_f^{(1)}(t_k) + \cdots + a_n y_f^{(n)}(t_k) - b_0 u_f(t_k) - \cdots - b_m u_f^{(m)}(t_k),\end{aligned}$$

where the filtered output and input signals are now given by

$$y_f^{(i)}(t_k) = \frac{D(q)}{C(q)} \frac{p^i}{A(p)} y(t_k), \quad i = 0, \dots, n, \quad (2.36a)$$

$$u_f^{(l)}(t_k) = \frac{D(q)}{C(q)} \frac{p^l}{A(p)} u(t_k), \quad l = 0, \dots, m. \quad (2.36b)$$

Thus, the pseudo-linear regression we obtain now is of the form (2.31), but where

$$\varphi_f(t_k) = \frac{D(q)}{C(q)} \left[\frac{-p}{A(p)} y(t_k), \dots, \frac{-p^n}{A(p)} y(t_k), \frac{1}{A(p)} u(t_k), \dots, \frac{p^m}{A(p)} u(t_k) \right]^\top. \quad (2.37)$$

Since $1/A(p)$ and $D(q)/C(q)$ are unknown, as they depend on the unknown parameters, the RIVC method proposes to estimate them via iterations. To this end, assume that the estimate at the j -th iteration $\boldsymbol{\theta}_j$ is known, with its associated model $G_j(p) := B_j(p)/A_j(p)$. First, an estimate of the noise model $C_{j+1}(q)/D_{j+1}(q)$ with parameter vector $\boldsymbol{\eta}_{j+1}$ is obtained by fitting an autoregressive moving-average (ARMA) model to the estimated noise sequence $y(t_k) - G_j(p)u(t_k)$. This estimation step can be performed using standard discrete-time methods such as PEM [138] or IVARMA [247]. Afterwards, the filtering in (2.36) is done considering the inverse of the noise model previously obtained and continuous-time transfer functions of the form $p^l/A_j(p)$. Lastly, the instrumental variable approach in (2.35) is used to produce the next iterate of the system parameters, but where the filtered instrument vector is given by

$$\hat{\varphi}_f(t_k) = \frac{D_{j+1}(q)}{C_{j+1}(q)} \left[\frac{-p B_j(p)}{A_j^2(p)} u(t_k), \dots, \frac{-p^n B_j(p)}{A_j^2(p)} u(t_k), \frac{1}{A_j(p)} u(t_k), \dots, \frac{p^m}{A_j(p)} u(t_k) \right]^\top, \quad (2.38)$$

and the filtered regressor vector $\varphi_f(t_k)$ is obtained as in (2.32) but with $A_j(p)$, $C_{j+1}(q)$ and $D_{j+1}(q)$ in place of $A(p)$, $C(q)$ and $D(q)$, respectively.

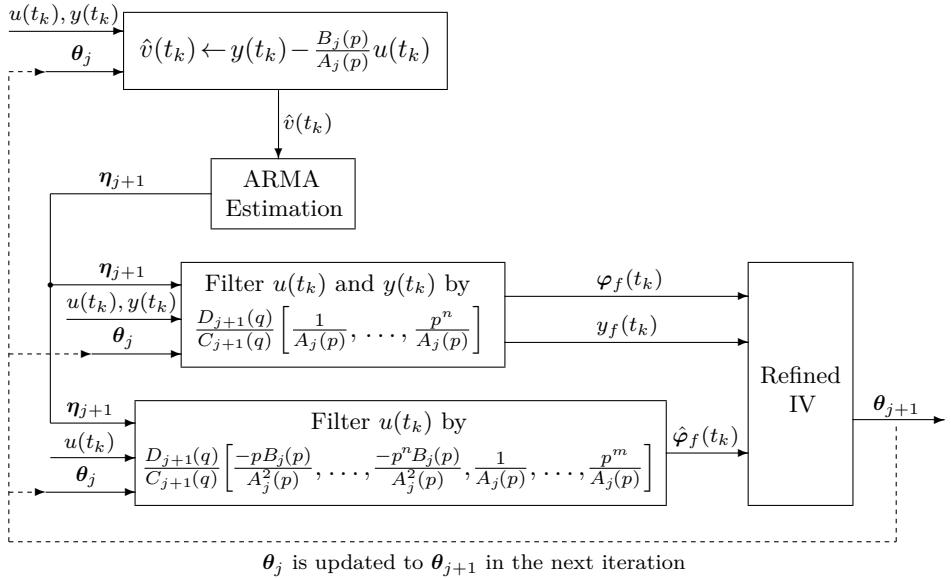


Figure 2.2: Implementation of the RIVC estimator. The algorithm is initialized with $\theta_j = \theta_1$ (i.e., $j = 1$). Note that the SRIVC algorithm follows from omitting the ARMA estimation step and using $D_{j+1}(q) = C_{j+1}(q) \equiv 1$.

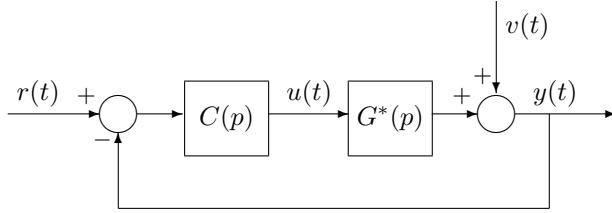
An outline of the RIVC estimator is presented in Figure 2.2, in which the stabilizing step in case the estimate is unstable is included in the Refined IV block. This procedure is the same one used for the SRIVC estimator, namely, lines 7 to 14 of Algorithm 2.1.

Direct approach for closed-loop system identification

The SRIVC and RIVC methods, developed initially for open-loop system identification, allow extensions for systems in closed-loop. In this work we study the extension of SRIVC for closed-loop settings, which is called CLSRIVC (i.e., closed-loop SRIVC [77]).

Consider the closed-loop system in Figure 2.3. In this formulation the user has knowledge of the continuous-time controller $C(p)$, the reference signal measurements $\{r(t_k)\}_{k=1}^N$, and the input and output samples $\{u(t_k), y(t_k)\}_{k=1}^N$. The reference signal is assumed to be provided by a ZOH device, i.e., $r(t)$ is constant between samples. As usual, the goal is to estimate the continuous-time transfer function $G^*(p)$. Note that an important aspect of this setup is that both the system and the controller are continuous in time; another setup that includes a discrete-time controller is discussed in detail in Chapter 7.

If the noise is assumed constant between samples, the input and output signals

**Figure 2.3:** Closed-loop setup.

satisfy

$$u(t_k) = \frac{C(p)}{1 + C(p)G^*(p)}r(t_k) - \frac{C(p)}{1 + C(p)G^*(p)}v(t_k), \quad (2.39a)$$

$$y(t_k) = \frac{C(p)G^*(p)}{1 + C(p)G^*(p)}r(t_k) + \frac{1}{1 + C(p)G^*(p)}v(t_k). \quad (2.39b)$$

Remark 2.6. The extensions and theoretical results we pursue can also be obtained for when there exists an external input between the controller and the plant. However, for simplicity we do not cover this case explicitly.

One naive way to estimate $G^*(p)$ in this scenario is to compute the SRIVC estimator using $\{u(t_k), y(t_k)\}_{k=1}^N$. However, the measurement noise is commonly correlated with the input in closed-loop setups [211], which can add bias to the SRIVC estimator. The key idea behind the CLSRIVC estimator is to use the reference input information to construct a filtered instrument that is not correlated with the measurement noise. In other words, the iterations for CLSRIVC are again of the form (2.35), with $\varphi_f(t_k)$ and $y_f(t_k)$ as in (2.32) and (2.34), respectively, but estimates of the noiseless input and output signals (i.e., (2.39a) and (2.39b) with $v(t_k) \equiv 0$) are used to form the filtered instrument $\hat{\varphi}_f(t_k)$:

$$\hat{\varphi}_f(t_k) = \left[\frac{-pT_{o,j}(p)}{A_j(p)}r(t_k), \dots, \frac{-p^n T_{o,j}(p)}{A_j(p)}r(t_k), \frac{S_{uo,j}(p)}{A_j(p)}r(t_k), \dots, \frac{p^m S_{uo,j}(p)}{A_j(p)}r(t_k) \right]^\top, \quad (2.40)$$

where the sensitivity functions $T_{o,j}(p)$ and $S_{uo,j}(p)$ are defined by

$$T_{o,j}(p) := \frac{G_j(p)C(p)}{1 + G_j(p)C(p)}, \quad S_{uo,j}(p) := \frac{C(p)}{1 + G_j(p)C(p)}.$$

Benefits and drawbacks of the direct approach

There have been interesting contributions stating the benefits of direct methods for continuous-time system identification, and in particular the benefits of instrumental variable methods [59, 63]. Here we will mention some advantages, and also several difficulties related to the direct approach.

First, in contrast to most indirect methods, the direct approach can naturally handle non-uniformly sampled data when operating with the output error model structure. This is because a discrete-time model is never computed, and hence no time-varying models are obtained. This property can be useful for event-based sampling since in these situations measurements are obtained sporadically [13]. Non-uniformly sampled data has also been used for estimating continuous-time ARMA and BJ models [30, 31]. Furthermore, the *a priori* knowledge of the relative degree of the continuous-time system can be easily incorporated, which leads to more parsimonious models than the indirect approach methods. Also, fast sampling usually does not lead to ill-conditioning: since these methods do not require a discrete-time equivalent system, they are not prone to the poor conditioning of model poles converging to the $(1, 0i)$ point in the z -plane, as explained in Subsection 2.3.2. This also explains why direct methods usually deliver better results for stiff systems compared to the indirect approach [141]. Finally, the direct methods are known to be reliable, having more than 40 years of use in practical applications [250].

The main difficulty present in the direct methods is the handling of derivatives, as they are not immediately available from discrete data points without amplifying noise [186]. The filtering procedures that these methods use require specifying the intersample behavior of the input and output, which are not always known and may have an impact on the overall performance of the algorithm. With regards to the theoretical foundations of these methods, convergence aspects and the influence of the intersample behaviors are still misinterpreted or not well understood by the system identification community.

2.4 Summary

In this chapter we first presented some notations and aspects of linear systems. Later, essential statistical tools were introduced, and the indirect and direct approaches to continuous-time system identification were described. Advantages and shortcomings of the methods related to each approach were underlined.

The contents exposed here mark the starting point of the developments presented in Chapters 3 to 10.

Asymptotically optimal indirect approach to continuous-time system identification

In the previous chapter we described the two main approaches for continuous-time system identification, indirect and direct. We now present a method that improves on the standard approach to indirect continuous-time system identification and analyze its asymptotic properties and robustness.

3.1 Introduction

In today's predominantly digital era, most system identification algorithms are studied in discrete-time setups [138, 211]. When the system of interest is continuous-time in nature instead of discrete-time, one natural approach to identify it is to first estimate its discrete-time equivalent and later convert the estimated model to continuous-time. This procedure is called the indirect approach for continuous-time system identification [186].

Even though the indirect approach seems straightforward to implement, as it is based on the solid theoretical and algorithmic foundations of discrete-time system identification, it has important limitations which have favored the use of the direct approach [59, 250]. Initialization of the indirect approach methods, together with their robustness, have been put into question in [185]. In [139] and [141] it has been suggested that a careful initialization of PEM must be set in order for the indirect approach to give promising results. Another difficulty present in the indirect approach is the handling of the relative degree of the estimate. As shown in Proposition 2.1 and observed in early works on indirect approach methods [204], even if a fully parameterized discrete-time model is chosen, the relative degree of the equivalent continuous-time model is usually one. Such property is independent of the relative degree of the true continuous-time system. Hence, an unnecessarily complex model structure is being estimated with the indirect approach. This leads to a loss in accuracy according to the parsimony principle [211, Ch. 11] which states that, among models that explain the data similarly well, the model that has the

smallest number of independent parameters usually is the most accurate.

This chapter focuses on how to solve the issues of relative degree and stability enforcement in the indirect approach method, and it is based on the works presented in [83] and [87]. In particular,

- we introduce a method that imposes a desired relative degree in the indirect approach to continuous-time system identification. For this, we rely on the indirect prediction error method (indirect PEM, [212]);
- we show that this procedure provides a consistent and asymptotically efficient estimator of the true parameter vector for inputs constructed via a zero-order-hold device;
- we develop a robustification step that enforces stability in the estimate. This is done by constructing ellipsoidal inner approximations of the continuous-time stability region [110]; and
- we compare the performance of this estimator against other indirect and direct methods for continuous-time system identification.

This chapter is organized as follows. In Section 3.2 the problem statement is described, and the indirect PEM method is introduced as a two-step optimization tool. We derive an estimator in Section 3.3 that optimally enforces the desired relative degree for the indirect approach, and determine its properties. We robustify the novel estimator in Section 3.4, and Section 3.5 illustrates the method with extensive numerical examples. Conclusions are drawn in Section 3.6. The proofs of Theorems 3.3 and 3.4 can be found in the appendix of this chapter.

3.2 Problem formulation and preliminaries

Consider an LTI, causal, asymptotically stable, SISO, continuous-time system

$$\begin{aligned} x(t) &= G^*(p)u(t) \\ &= \frac{b_{n^*-r}^* p^{n^*-r} + b_{n^*-r-1}^* p^{n^*-r-1} + \cdots + b_1^* p + b_0}{p^{n^*} + a_{n^*-1}^* p^{n^*-1} + \cdots + a_1^* p + a_0^*} u(t), \end{aligned} \tag{3.1}$$

where r ($\leq n^*$) is the relative degree of the system $G^*(p)$. The numerator and denominator polynomials of $G^*(p)$ are assumed to be coprime and hence no zero-pole cancellations occur. We denote $\boldsymbol{\theta}^* := [b_{n^*-r}^*, \dots, b_0^*, a_{n^*-1}^*, \dots, a_0^*]^\top$ as the true continuous-time system parameter vector¹.

¹This description is a permutation of the entries of (2.6), and it describes a monic denominator polynomial instead of an anti-monic one. The difference in notation is only present in this chapter, and it is used to simplify the computations.

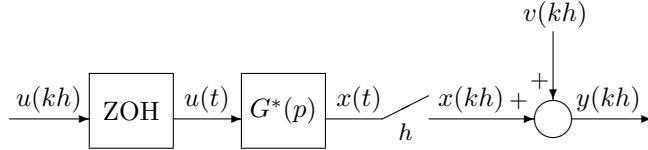


Figure 3.1: System description.

Suppose that the input and output signals are sampled with period h and the resulting output is contaminated by an additive zero-mean white noise sequence $\{v(kh)\}$ of variance σ^2 as depicted in Figure 3.1. That is,

$$y(kh) = x(kh) + v(kh), \quad k \in \mathbb{N}. \quad (3.2)$$

The goal of continuous-time system identification using sampled data is to obtain a continuous-time transfer function estimate for $G^*(p)$, given N input and output data measurements $\{u(kh), y(kh)\}_{k=1}^N$, knowledge about the physical characteristics of the system, and the intersample behavior of the signals. In this chapter we assume that the continuous-time input is a piecewise constant signal between samples, i.e., it has a ZOH behavior [16].

One way to obtain a model of $G^*(p)$ is to first identify the ZOH equivalent transfer function given the input and output data measurements by applying the standard PEM method in the discrete-time domain, and then return to the continuous-time domain via the ZOH equivalence relation (2.8). Although this procedure exhibits good statistical properties, the ZOH can lead to continuous-time models having numerator parameters that exceed the desired numerator polynomial degrees, which contributes to poor accuracy and high variance at high frequencies. This generally is the case when the relative degree of $G^*(p)$ is greater than one (see Proposition 2.1). One elementary way of treating this issue consists in setting the numerator coefficients that should be zero to zero, but it is known that this is not statistically optimal [141]. Our goal is to optimally impose this constraint, which should lead to a statistically improved estimate of the continuous-time system. We adopt ideas from the indirect PEM [212] technique to achieve this goal, which will be briefly explained next.

3.2.1 Indirect PEM

A variant of the prediction error method was developed in [212], called indirect PEM. This method provides an optimal way of reducing the size of the model structure one considers in PEM and is presented as follows. Consider two nested model structures \mathbf{M}_1 and \mathbf{M}_2 , where $\mathbf{M}_1 \subset \mathbf{M}_2$, and suppose that we are interested in estimating a model parameterized by $\boldsymbol{\theta}$ in the model structure \mathbf{M}_1 . Given measured data, suppose that the PEM estimate is relatively easy to obtain using the model parameterized by $\boldsymbol{\alpha}$ in the structure \mathbf{M}_2 . The idea of indirect PEM is then to produce an estimate in \mathbf{M}_1 using the PEM model parameters in \mathbf{M}_2 as the “data” for this estimation.

Before describing the procedure, we shall state the assumptions needed for the model structures \mathbf{M}_1 and \mathbf{M}_2 and the data. First the structures should be nested, and there should exist a smooth and injective map $\mathbf{f}: S_1 \rightarrow S_2$ that links $\boldsymbol{\theta}$ with $\boldsymbol{\alpha}$, where $S_1 \subset \mathbb{R}^{\dim \boldsymbol{\theta}}$ and $S_2 \subset \mathbb{R}^{\dim \boldsymbol{\alpha}}$, in which the predictors in the PEM formulation are equal. That is, if $\varepsilon_1(kh, \boldsymbol{\theta})$ and $\varepsilon_2(kh, \boldsymbol{\alpha})$ are the residuals for computing the PEM estimates in the model structures \mathbf{M}_1 and \mathbf{M}_2 , then $\varepsilon_2(kh, \mathbf{f}(\boldsymbol{\theta})) = \varepsilon_1(kh, \boldsymbol{\theta})$. Note that S_2 may be of lower dimension than the complete space $\mathbb{R}^{\dim \boldsymbol{\alpha}}$. We also must assume that the Jacobian matrix $\partial \mathbf{f}(\boldsymbol{\theta}) / \partial \boldsymbol{\theta}$ has full rank over S_1 , and that both model structures \mathbf{M}_1 and \mathbf{M}_2 give parameter identifiability (recall Section 2.2).

The indirect PEM procedure consists of two main steps. First, the standard PEM is applied to the model structure \mathbf{M}_2 , yielding an estimate $\hat{\boldsymbol{\alpha}}$. Next, we should check if $\hat{\boldsymbol{\alpha}} \in S_2$. If this is the case, then we simply obtain $\hat{\boldsymbol{\theta}}$, the indirect PEM estimate, as

$$\hat{\boldsymbol{\theta}} = \mathbf{f}^{-1}(\hat{\boldsymbol{\alpha}}), \quad (3.3)$$

where $\mathbf{f}^{-1}: S_2 \rightarrow S_1$ is the inverse function of \mathbf{f} . If $\hat{\boldsymbol{\alpha}} \notin S_2$, which occurs in most cases of interest², then (3.3) cannot be used. In such cases, the following optimization problem is proposed for finding the estimate $\hat{\boldsymbol{\theta}}$:

$$\hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta}} \frac{1}{2} (\hat{\boldsymbol{\alpha}} - \mathbf{f}(\boldsymbol{\theta}))^\top \hat{\Sigma}_{\boldsymbol{\alpha}}^{-1} (\hat{\boldsymbol{\alpha}} - \mathbf{f}(\boldsymbol{\theta})), \quad (3.4)$$

with $\hat{\Sigma}_{\boldsymbol{\alpha}}$ being a consistent estimate of the covariance matrix of $\hat{\boldsymbol{\alpha}}$. Note that solving (3.4) explicitly is not always needed, as a single Newton-Raphson iteration of (3.4) starting from a consistent estimate of $\boldsymbol{\theta}^*$ can yield an asymptotically efficient estimator [98, Ch. 23].

In terms of statistical performance, it is shown in [212] that indirect PEM delivers estimates that have the same asymptotic distribution as the PEM estimates obtained from the smallest model structure \mathbf{M}_1 . It is computationally fast, and it also has been observed in [212] that numerical issues and convergence properties for finite sample size can be better than the results obtained using a direct PEM procedure.

3.3 Asymptotically optimal enforcement of relative degree

In this section we develop an indirect estimator for the continuous-time parameter vector $\boldsymbol{\theta}^*$ that renders a transfer function estimate of a desired relative degree r . For this matter, we first focus on the PEM estimator of the ZOH equivalent model of $G^*(p)$.

For simplicity we assume that the correct model order is known, i.e., $n = n^*$ and r is given. A model with structure (2.18) is obtained by PEM, and the covariance matrix of the estimated discrete-time parameter vector

$$\hat{\boldsymbol{\rho}} = \left[\hat{\beta}_{n-1}, \hat{\beta}_{n-2}, \dots, \hat{\beta}_0, \hat{\alpha}_{n-1}, \hat{\alpha}_{n-2}, \dots, \hat{\alpha}_0 \right]^\top$$

²In fact, we typically have $\dim S_2 < \dim \boldsymbol{\alpha}$, which means that S_2 has Lebesgue measure zero over $\mathbb{R}^{\dim \boldsymbol{\alpha}}$ [191].

is also estimated. We know that the continuous-time ZOH equivalent of this model is in general given by a transfer function with relative degree equal to one, of the form

$$\hat{G}(p) = \frac{\hat{b}_{n-1}p^{n-1} + \hat{b}_{n-2}p^{n-2} + \cdots + \hat{b}_1p + \hat{b}_0}{p^n + \hat{a}_{n-1}p^{n-1} + \cdots + \hat{a}_1p + \hat{a}_0}.$$

In other words, $\hat{G}(p)$ is the model obtained by the standard indirect approach for continuous-time system identification. We define its parameter vector as

$$\hat{\boldsymbol{\theta}} = \left[\hat{b}_{n-1}, \dots, \hat{b}_0, \hat{a}_{n-1}, \dots, \hat{a}_0 \right]^\top,$$

and denote the true discrete-time parameter vector by $\boldsymbol{\rho}^* \in \mathbb{R}^{2n}$. The parameters in $\hat{\boldsymbol{\theta}}$ are related to $\hat{\boldsymbol{\rho}}$ by the zero-order hold equivalence equations that can be derived by using (2.8) and comparing coefficients. This relation is a nonlinear mapping $\boldsymbol{\rho}: \hat{\boldsymbol{\theta}} \mapsto \boldsymbol{\rho}(\hat{\boldsymbol{\theta}}) = \hat{\boldsymbol{\rho}}$ that is differentiable almost everywhere, and is bijective if the sampling frequency (in radians) is greater than twice the largest imaginary part of the s -domain poles [122]. The main idea of the proposed method is to apply the indirect PEM technique to the continuous-time parameter estimates of the indirect approach. For this, we shall define the smooth mapping \mathbf{f} , the subspaces S_1 and S_2 , compute the covariance matrix estimate $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\alpha}}$, and solve the optimization problem in (3.4).

Let $S_1 = \mathbb{R}^{2n-r+1}$ consist of vectors of the form $[b_{n-r}, \dots, b_0, a_{n-1}, \dots, a_0]^\top$, and let S_2 be defined as

$$S_2 = \{ [b_{n-1}, \dots, b_0, a_{n-1}, \dots, a_0]^\top \in \mathbb{R}^{2n} : b_{n-1} = b_{n-2} = \cdots = b_{n-r+1} = 0 \}.$$

Note that, from Proposition 2.1, the indirect approach estimate will in general not belong to S_2 . Provided that the model order is known, the function $\mathbf{f}: S_1 \rightarrow S_2$ is defined as:

$$\mathbf{f}([b_{n-r}, \dots, b_0, a_{n-1}, \dots, a_0]^\top) = [0, \dots, 0, b_{n-r}, \dots, b_0, a_{n-1}, \dots, a_0]^\top. \quad (3.5)$$

It can be readily checked that \mathbf{f} is smooth and injective, and delivers equal predictors in its domain.

Next, we must find an estimate of the covariance matrix of $\hat{\boldsymbol{\theta}}$, the continuous-time estimate from the indirect approach. For this we know that a consistent estimate of the covariance matrix of the discrete-time model parameters, $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\rho}}$, can be computed (see e.g. Section 9.3 of [138]) and is available through the PEM implementation in MATLAB [140]. Furthermore, as a consequence of the Delta method [230, Thm. 3.1], we find that $\sqrt{N}(\hat{\boldsymbol{\rho}} - \boldsymbol{\rho}^*)$ converges in law to a multivariate normal distribution with asymptotic covariance given by $\mathbf{J}\boldsymbol{\Sigma}_{\boldsymbol{\theta}}\mathbf{J}^\top$, where \mathbf{J} is the Jacobian matrix of $\boldsymbol{\rho}(\boldsymbol{\theta})$ evaluated at the naive estimation of $\boldsymbol{\theta}^*$, that is, throwing away the high order coefficients of the numerator of $\hat{G}(p)$ that should be zero. Note that the standard PEM estimate can also be used for this evaluation, yielding the same asymptotic relation.

Given that we derived all the ingredients for the proposed method based on indirect PEM, we now need to find an appropriate projection of $\hat{\boldsymbol{\theta}}$ onto S_1 . This is done by solving the following optimization problem:

$$\begin{aligned}\tilde{\boldsymbol{\theta}} := \arg \min_{\boldsymbol{\theta}} & \frac{1}{2} (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^\top \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}}^{-1} (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \\ \text{s.t. } & \begin{bmatrix} \mathbf{I}_{r-1} & \mathbf{0} \end{bmatrix} \boldsymbol{\theta} = \mathbf{0},\end{aligned}\quad (3.6)$$

where \mathbf{I}_{r-1} is the identity matrix of dimension $r - 1$, and $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}}^{-1} = \mathbf{J}^\top \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\rho}}^{-1} \mathbf{J}$. Bear in mind that the covariance matrix estimate $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}}$ does not depend on $\boldsymbol{\theta}$ for the purposes of minimizing the cost in (3.6). From Lagrange multiplier theory [25], we know that the optimization problem in (3.6) is equivalent to calculating, for a suitable $\boldsymbol{\lambda} \in \mathbb{R}^{2n-r+1}$, the estimator given by

$$\tilde{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta}} \frac{1}{2} (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^\top \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}}^{-1} (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) + \boldsymbol{\lambda}^\top \begin{bmatrix} \mathbf{I}_{r-1} & \mathbf{0} \end{bmatrix} \boldsymbol{\theta}.$$

Partitioning $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}}$ appropriately (easing the notation for simplicity), and differentiating with respect to $\boldsymbol{\theta}$, we obtain

$$\tilde{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}} - \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{12}^\top & \boldsymbol{\Sigma}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{I}_{r-1} \\ \mathbf{0} \end{bmatrix} \boldsymbol{\lambda}, \quad (3.7)$$

which leads to $\boldsymbol{\lambda} = \boldsymbol{\Sigma}_{11}^{-1} [\mathbf{I}_{r-1} \ \mathbf{0}] \hat{\boldsymbol{\theta}}$ after the constraints in (3.6) are imposed. If we denote by \mathbf{C} the Cholesky factorization matrix [115] of $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}}$ (i.e., a lower triangular matrix with positive diagonal entries such that $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}} = \mathbf{C} \mathbf{C}^\top$), and \mathbf{C}_{11} , \mathbf{C}_{21} and \mathbf{C}_{22} as its partition matrices of appropriate dimensions, we can write (3.7) as

$$\begin{aligned}\tilde{\boldsymbol{\theta}} &= \hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}} \begin{bmatrix} \mathbf{I}_{r-1} \\ \mathbf{0} \end{bmatrix} \boldsymbol{\Sigma}_{11}^{-1} \begin{bmatrix} \mathbf{I}_{r-1} & \mathbf{0} \end{bmatrix} \hat{\boldsymbol{\theta}} \\ &= \hat{\boldsymbol{\theta}} - \begin{bmatrix} \mathbf{C}_{11} & \mathbf{0} \\ \mathbf{C}_{21} & \mathbf{C}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{I}_{r-1} \\ \mathbf{0} \end{bmatrix} \mathbf{C}_{11}^{-1} \begin{bmatrix} \mathbf{I}_{r-1} & \mathbf{0} \end{bmatrix} \hat{\boldsymbol{\theta}} \\ &= \mathbf{C} \begin{bmatrix} \mathbf{0}_{r-1} & \mathbf{0} \\ \mathbf{0}^\top & \mathbf{I}_{2n-r+1} \end{bmatrix} \mathbf{C}^{-1} \hat{\boldsymbol{\theta}}.\end{aligned}\quad (3.8)$$

The proposed estimator (3.8), which we will call the IPEM estimator, can be seen as an asymptotically optimal \mathcal{L}^2 approximation to the indirect approach estimate $\hat{\boldsymbol{\theta}}$ that imposes the desired relative degree. Note that this estimator relies on PEM giving a good initial estimate of the discrete-time model parameters.

Remark 3.1. The relative degree of the continuous-time system is not always known. In some practical applications, physical knowledge about the system can give intuition. In addition, statistical measures, such as the coefficient of determination [59] or Young's information criterion [245], can be used to select the relative degree of the model.

3.3.1 Properties of the IPEM estimator

We present the most important properties of the IPEM estimator (3.8) in the following theorems.

Theorem 3.1 (Consistency and asymptotic efficiency of the IPEM estimator). *Consider the system described by (3.1) and (3.2), where $\{v(kh)\}_{k=1}^N$ is a Gaussian white noise sequence of variance σ^2 . Assume that the sampling frequency $2\pi/h$ is greater than twice the largest imaginary part of the s -domain poles and that there is no delay in the true continuous-time system.³ Then, (3.8) is a consistent and asymptotically efficient estimator of the real vector parameter $\boldsymbol{\theta}^*$, provided that the discrete-time model set contains the true discrete-time system and the relative degree chosen coincides with that of the true continuous-time system.*

Proof. The discrete-time PEM estimate can be interpreted as the maximum likelihood (ML) estimate under the Gaussian noise assumption. In addition, the $z \rightarrow s$ transformation is unique for sufficiently large sample size under the proposed sampling frequency [122]. Hence, by the invariance principle of ML estimators [261], the continuous-time system parameter vector obtained through the inverse of the ZOH transformation of the PEM estimate is also an ML estimate.

To prove the theorem, we only need to show that the assumptions in [212] are satisfied in this scenario so that we can directly apply the results obtained in Section 3 of [212] related to the indirect PEM procedure. First, note that the model structure given by this method contains the models with the desired relative degree (and the inclusion is proper if $r > 1$), with a smooth linear mapping \mathbf{f} in (3.5) that can be equivalently described by the matrix

$$\mathbf{T} = \begin{bmatrix} \mathbf{0}_{(r-1) \times (2n-r+1)} \\ \mathbf{I}_{2n-r+1} \end{bmatrix}.$$

Furthermore, provided that the discrete-time model set contains the true system, both structures give parameter identifiability. Also note that $\hat{\Sigma}_{\boldsymbol{\theta}}$, computed by $\mathbf{J}^{-1}\hat{\Sigma}_{\boldsymbol{\theta}}\mathbf{J}^{-\top}$, is a consistent estimate of the covariance matrix of $\hat{\boldsymbol{\theta}}$. Hence, the results in [212] follow. Namely, the normalized estimation errors⁴ $\sqrt{N}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^*)/\sigma$ and $\sqrt{N}(\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}^*)/\sigma$ are asymptotically normally distributed with zero means and their asymptotic covariance matrices satisfy the relation

$$[\text{AsCov}(\tilde{\boldsymbol{\theta}})]^{-1} = \mathbf{T}^\top [\text{AsCov}(\hat{\boldsymbol{\theta}})]^{-1} \mathbf{T}.$$

Moreover, following the steps in Section 3 of [212], the improved PEM estimate is consistent and has the same asymptotic distribution as the PEM estimate considering

³These conditions can be relaxed, as long as the sampling frequency is such that the $z \rightarrow s$ transformation is well defined.

⁴For simplicity, we discard the enforced zero elements of $\tilde{\boldsymbol{\theta}}$ in this analysis. We assume that the vector $\boldsymbol{\theta}^*$ has the appropriate dimension, where its zero elements have been discarded in the expression $\sqrt{N}(\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}^*)/\sigma$.

the model structure consisting of continuous-time models with relative degree at most r , thereby proving its asymptotic efficiency. \square

Theorem 3.2. *Consider the standard indirect approach estimator $\hat{\boldsymbol{\theta}}$ and the IPEM estimator $\tilde{\boldsymbol{\theta}}$ as described in (3.8). The asymptotic covariances of these estimators satisfy the following properties:*

1. $\lim_{N \rightarrow \infty} N \mathbb{E}\{(\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}^*)(\hat{\boldsymbol{\theta}} - \tilde{\boldsymbol{\theta}})^\top\} = \mathbf{0}$.
2. $\text{AsCov}(\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}^*) = \text{AsCov}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^*) - \text{AsCov}(\hat{\boldsymbol{\theta}} - \tilde{\boldsymbol{\theta}})$.

Proof. Both of these claims follow by applying Properties 10.5 and 10.6 from [98, Ch. 10] to this context. \square

These properties imply that, for a sufficiently large sample size, the proposed estimator can only decrease the covariance of the estimated parameters compared to the standard indirect approach method. The asymptotic covariances satisfy a Pythagorean relation, as for large sample size the PEM estimate is projected orthogonally onto the proper subspace where $\boldsymbol{\theta}^*$ lies.

Next, we establish that imposing a larger relative degree improves the precision of the estimates, provided that the highest relative degree model structure contains the true system.

Theorem 3.3. *Given $G^*(p)$ of order n and relative degree $r > 1$, consider two continuous-time candidate models of order n and relative degrees r_1 and r_2 , and their IPEM parameter vector estimates $\tilde{\boldsymbol{\theta}}^{r_1}$ and $\tilde{\boldsymbol{\theta}}^{r_2}$, respectively. If $r_1 < r_2 \leq r$, then their asymptotic covariance matrices satisfy $\text{AsCov}(\tilde{\boldsymbol{\theta}}^{r_2}) \preceq \text{AsCov}(\tilde{\boldsymbol{\theta}}^{r_1})$.*

Proof. See Appendix 3.A.1. \square

It is possible to obtain explicit expressions for the reduction in covariance by this method by making use of some asymptotic expressions. For analysis purposes only, here we assume knowledge of the true discrete-time system parameter vector $\boldsymbol{\rho}^*$.

Theorem 3.4. *Consider the estimators $\hat{\boldsymbol{\theta}}$ and $\tilde{\boldsymbol{\theta}}$ and their asymptotic covariances $\text{AsCov}(\hat{\boldsymbol{\theta}})$ and $\text{AsCov}(\tilde{\boldsymbol{\theta}})$, respectively.*

1. *If $\text{AsCov}(\hat{\boldsymbol{\theta}})$ is partitioned as*

$$\text{AsCov}(\hat{\boldsymbol{\theta}}) = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{12}^\top & \boldsymbol{\Sigma}_{22} \end{bmatrix}, \quad (3.9)$$

where $\boldsymbol{\Sigma}_{11} \in \mathbb{R}^{(r-1) \times (r-1)}$, and $\boldsymbol{\Sigma}_{22} \in \mathbb{R}^{(2n-r+1) \times (2n-r+1)}$, then the difference between the asymptotic covariances of the truly non-zero parameters of $\hat{\boldsymbol{\theta}}$ and $\tilde{\boldsymbol{\theta}}$ is given by $\boldsymbol{\Sigma}_{12}^\top \boldsymbol{\Sigma}_{11}^{-1} \boldsymbol{\Sigma}_{12} \succeq \mathbf{0}$.

2. If $\{u(kh)\}$ and $\{v(kh)\}$ are white noise sequences of variance λ^2 and σ^2 respectively, then $\text{AsCov}(\hat{\theta})$ can be computed explicitly as $\sigma^2\lambda^{-2}(\mathbf{J}^\top \mathbf{S}\mathbf{P}\mathbf{S}^\top \mathbf{J})^{-1}$, where \mathbf{J} is the Jacobian matrix of $\rho(\theta)$ evaluated at θ^* , $\mathbf{S} \in \mathbb{R}^{2n \times 2n}$ is the Sylvester matrix associated with the true discrete-time polynomials $A_d^*(q)$ and $B_d^*(q)$, i.e.,

$$\mathbf{S} = \begin{bmatrix} 1 & \alpha_{n-1}^* & \cdots & \alpha_1^* & \alpha_0^* & 0 \\ \ddots & \ddots & & \ddots & \ddots & \\ 0 & 1 & \alpha_{n-1}^* & \cdots & \alpha_1^* & \alpha_0^* \\ 0 & -\beta_{n-1}^* & \cdots & -\beta_1^* & -\beta_0^* & 0 \\ \ddots & \ddots & & \ddots & \ddots & \\ 0 & 0 & -\beta_{n-1}^* & \cdots & -\beta_1^* & -\beta_0^* \end{bmatrix}, \quad (3.10)$$

and

$$\mathbf{P} := \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{|A_d^*(e^{i\omega})|^4} \mathbf{\Gamma}(e^{i\omega}) \mathbf{\Gamma}^H(e^{i\omega}) d\omega, \quad (3.11)$$

where $\mathbf{\Gamma}(e^{i\omega}) := [e^{i\omega(2n-1)}, e^{i\omega(2n-2)}, \dots, 1]^\top$.

Proof. See Appendix 3.A.2. □

Theorem 3.4 quantifies the asymptotic gain in accuracy achieved with the proposed method compared to the standard indirect approach. Intuitively, the improvement is greater when the variance of the excess high order numerator coefficients is low, and when these coefficients are highly correlated with the estimates of the model parameters of interest.

3.4 Enforcing stability in IPEM

The IPEM estimator for continuous-time systems enjoys strong asymptotic statistical properties as seen in Theorems 3.1–3.4. However, when a limited number of data points is collected, or when the signal-to-noise ratio (SNR) is low, the high order numerator coefficients in the standard indirect estimate can be far from zero, which produces strong perturbations in the denominator coefficients of the IPEM estimate. This effect can lead to instability even if the standard indirect approach method provides a stable estimate, as seen in the following motivating example.

Example 3.1. Consider the following continuous-time system:

$$\begin{aligned} x(t) &= \frac{0.67p + 0.23}{p^3 + 1.83p^2 + 0.99p + 0.15} u(t), \\ y(kh) &= x(kh) + v(kh), \end{aligned}$$

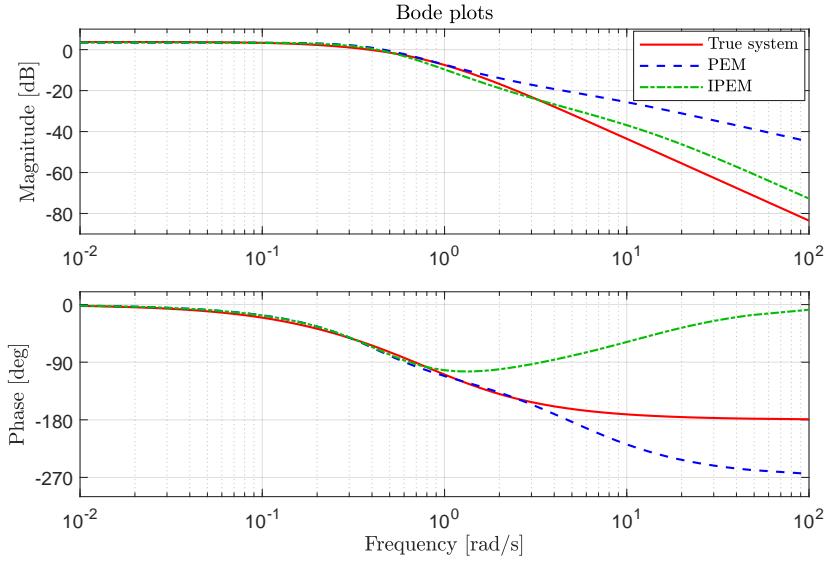


Figure 3.2: Bode plots of the system and models of Example 3.1. True system (solid, red), PEM estimate (dashed, blue), and IPEM estimate (dash-dotted, green).

where $\{u(kh)\}$ and $\{v(kh)\}$ are white noise processes, and $u(t)$ is piecewise constant. This system is asymptotically stable, with poles at $s \approx -0.97, s \approx -0.61$ and $s \approx -0.25$. The SNR between the noiseless output sequence and the additive noise is 3[dB], and 500 data points are collected. The standard indirect approach delivers an asymptotically stable model

$$\hat{G}(p) = \frac{-0.57p^2 + 1.70p + 1.72}{p^3 + 6.81p^2 + 4.51p + 1.16},$$

whereas the proposed IPEM estimator gives

$$\tilde{G}(p) = \frac{-2.33p - 4.10}{p^3 - 12.41p^2 - 9.67p - 2.75}.$$

Bode diagrams of these models can be found in Figure 3.2. The IPEM estimator delivers an unstable estimate even though both the system and the standard indirect approach estimates are asymptotically stable.

The goal in this section is to enforce stability in the IPEM estimate. In continuous-time systems, enforcing stability as a constraint on the parameter space can be done through the Routh-Hurwitz criterion [97]. However, the stability domain derived for polynomial orders greater than two is non-convex (see, e.g., [2, 110]), which introduces difficulties in optimization. This problem has been dealt with by obtaining convex bounds in an EM formulation for state-space models [223], or by introducing convex approximations of the stability region, like polyhedra in a robust control

framework [3], or ellipsoids [110]. In particular, ellipsoidal approximations have been used for imposing stability in the SRIVC estimator in [103], and for control design in [42].

The key idea behind the robustification step is to generate a closed convex stability domain in the space of the parameter coefficients, and to project the standard indirect approach estimate into the intersection of this domain with the subspace that yields the desired relative degree. Before presenting the robust IPEM algorithm, we introduce the techniques used to generate the domain of interest. Let

$$\mathcal{D} = \{s \in \mathbb{C}: a + b(s + \bar{s}) + cs\bar{s} < 0\} \quad (3.12)$$

be a given region in the complex plane, where $a, b, c \in \mathbb{R}$. We define the vectors $\mathbf{x} := [x_0, x_1, \dots, x_{n-1}]^\top$ and $\bar{\mathbf{x}} := [x_0, x_1, \dots, x_{n-1}, x_n]^\top$ to be the coefficients of the polynomial $x(s) = x_n s^n + \dots + x_1 s + x_0$, where we assume without loss of generality that $x_n = 1$. The following well-known result relates the location of the roots of $x(s)$ with a positive-definiteness condition, and it is an extension of Hermite's stability criterion [171].

Lemma 3.1. ([110, 131]) *The roots of the polynomial $x(s)$ lie in \mathcal{D} if and only if*

$$\mathbf{H}(\bar{\mathbf{x}}) = \sum_{i,j=0}^n x_i x_j \mathbf{H}_{ij} \succ 0,$$

where $\mathbf{H}_{ij} = \mathbf{H}_{ji}^\top \in \mathbb{R}^{n \times n}$ are constant matrices depending only on \mathcal{D} , which are computed by equating the coefficients of each side of the expression

$$\bar{\mathbf{x}}\bar{\mathbf{x}}^\top - \tilde{\mathbf{x}}\tilde{\mathbf{x}}^\top = a\mathbf{R}_l^\top \mathbf{H}(\bar{\mathbf{x}})\mathbf{R}_l + b(\mathbf{R}_l^\top \mathbf{H}(\bar{\mathbf{x}})\mathbf{R}_r + \mathbf{R}_r^\top \mathbf{H}(\bar{\mathbf{x}})\mathbf{R}_l) + c\mathbf{R}_r^\top \mathbf{H}(\bar{\mathbf{x}})\mathbf{R}_r, \quad (3.13)$$

where $\mathbf{R}_l = [\mathbf{I}_n \ \mathbf{0}_{n \times 1}]$, $\mathbf{R}_r = [\mathbf{0}_{n \times 1} \ \mathbf{I}_n]$, and $\tilde{\mathbf{x}} \in \mathbb{R}^{n+1}$ is the vector of coefficients of the polynomial

$$\tilde{x}(s) = \left(\frac{b + cs}{\sqrt{b^2 - ac}} \right)^n x \left(-\frac{a + bs}{b + cs} \right). \quad (3.14)$$

For the following result, we need to write the matrix $\mathbf{H}(\bar{\mathbf{x}})$ as

$$\mathbf{H}(\bar{\mathbf{x}}) = (\mathbf{I}_n \otimes \bar{\mathbf{x}})^\top \bar{\mathbf{H}} (\mathbf{I}_n \otimes \bar{\mathbf{x}}), \quad (3.15)$$

where $\bar{\mathbf{H}} \in \mathbb{R}^{n(n+1) \times n(n+1)}$ is a fixed matrix that is formed by replacing (3.15) in (3.13) and matching polynomial coefficients. In Theorem 3.5 we present a procedure for generating stable ellipsoids, which was first introduced in [110].

Theorem 3.5. *Let \mathcal{D} be a stability region with associated matrix $\bar{\mathbf{H}}$, and let $\mathbf{x}_C \in \mathbb{R}^n$ describe an n -th order monic polynomial with all its roots in \mathcal{D} . Solve the convex*

optimization problem

$$\begin{aligned} & \min_{\mathbf{P}, \mathbf{G}, \mathbf{D}} \text{trace}(\mathbf{P}) \\ \text{s.t. } & (\mathbf{D} \otimes \mathbf{I}_{n+1}) \bar{\mathbf{H}} = \bar{\mathbf{H}} (\mathbf{D} \otimes \mathbf{I}_{n+1}), \\ & (\mathbf{D} \otimes \mathbf{I}_{n+1}) \bar{\mathbf{H}} \succ \mathbf{I}_n \otimes \bar{\mathbf{P}} + \mathbf{G}, \\ & \mathbf{D} = \mathbf{D}^\top \succ \mathbf{0} \in \mathbb{R}^{n \times n}, \end{aligned} \quad (3.16)$$

where $\bar{\mathbf{P}}$ is a symmetric block matrix which is partitioned as

$$\bar{\mathbf{P}} = \begin{bmatrix} -\mathbf{P} & \mathbf{P}\mathbf{x}_C \\ \mathbf{x}_C^\top \mathbf{P} & 1 - \mathbf{x}_C^\top \mathbf{P} \mathbf{x}_C \end{bmatrix},$$

with $\mathbf{P} \succ \mathbf{0}, \mathbf{P} \in \mathbb{R}^{n \times n}$, and \mathbf{G} is a symmetric block matrix

$$\mathbf{G} = \begin{bmatrix} \mathbf{0} & \mathbf{G}_{21}^\top & \cdots & \mathbf{G}_{n1}^\top \\ \mathbf{G}_{21} & \mathbf{0} & \cdots & \mathbf{G}_{n2}^\top \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{G}_{n1} & \mathbf{G}_{n2} & \cdots & \mathbf{0} \end{bmatrix}$$

formed by skew-symmetric matrices $\mathbf{G}_{ij} = -\mathbf{G}_{ij}^\top \in \mathbb{R}^{(n+1) \times (n+1)}$.

Take $\mathbf{P} = \mathbf{P}_{\text{opt}}$ as the solution of the optimization problem stated in (3.16). Then, any vector \mathbf{x} such that $(\mathbf{x} - \mathbf{x}_C)^\top \mathbf{P}_{\text{opt}} (\mathbf{x} - \mathbf{x}_C) \leq 1$ parameterizes a polynomial $x(s)$ with all its roots in \mathcal{D} .

Proof. See [110]. □

By setting $a = 0$, $b = 1$ and $c = 0$, Theorem 3.5 provides a procedure for computing an ellipsoid with center at \mathbf{x}_C whose interior points represent stable polynomials. This is a convex constraint, and therefore it can be easily included in a convex optimization problem. Note that \mathbf{x}_C must describe a stable polynomial which is user-defined. Since the stability region (in the parameter space) for the polynomial $\tilde{A}(p) = p^n + \tilde{a}_{n-1}p^{n-1} + \cdots + \tilde{a}_1p + \tilde{a}_0$ is non-convex for $n > 2$, we shall approximate this non-convex stability region by ellipsoids and solve the minimization problem of indirect PEM in (3.6) for each convex region.

The general procedure we propose is described in Algorithm 3.1. Since the system $G^*(p)$ is assumed stable, the proposed method enjoys the same asymptotic properties as the IPEM method described in Theorem 3.1. This is due to the fact that IPEM will provide unstable models with decreasing probability as the number of data points tends to infinity, and thus, for large data sets, the robustified method will deliver the same estimate as the IPEM estimator in (3.8). However, it is expected that the robustified method shows greater improvement against the estimator (3.8) for small data sets or when the SNR is poor. Note that in line 6 of Algorithm 3.1

Algorithm 3.1: Robustified IPEM for continuous-time system identification

1: Input: $\{u(kh), y(kh)\}_{k=1}^N$, number of poles n , relative degree r , number of ellipsoids M
 2: Compute the estimate $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}^*$ by the standard indirect approach and its covariance matrix $\hat{\Sigma}_{\boldsymbol{\theta}}$
 3: Compute $\tilde{\boldsymbol{\theta}}$ by (3.8)
 4: **if** $\tilde{\boldsymbol{\theta}}$ describes an unstable model **then**
 5: Compute $\bar{\mathbf{H}}$ by equations (3.13) and (3.15)
 6: Pick $\boldsymbol{\alpha}^k \in \mathbb{R}^n$, $k = 1, \dots, M$, all of which describe stable polynomials
 7: **for** $i = 1, \dots, M$ **do**
 8: For $\mathbf{x}_C = \boldsymbol{\alpha}^k$, obtain $\mathbf{P}_{\text{opt}}^k$ by solving (3.16)
 9: Solve

$$\tilde{\boldsymbol{\theta}}^k = \arg \min_{\boldsymbol{\theta}} \frac{1}{2} (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^\top \hat{\Sigma}_{\boldsymbol{\theta}}^{-1} (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})$$

$$\text{s.t. } \begin{bmatrix} \mathbf{I}_{r-1} & \mathbf{0} \end{bmatrix} \boldsymbol{\theta} = \mathbf{0},$$

$$(\boldsymbol{\theta} - \boldsymbol{\alpha}^k)^\top \mathbf{P}_{\text{opt}}^k (\boldsymbol{\theta} - \boldsymbol{\alpha}^k) \leq 1$$

10: **end for**
 11: Compute

$$\tilde{\boldsymbol{\theta}} = \arg \min_{\{\tilde{\boldsymbol{\theta}}^k\}_{k=1}^M} \frac{1}{2} (\hat{\boldsymbol{\theta}} - \tilde{\boldsymbol{\theta}}^k)^\top \hat{\Sigma}_{\boldsymbol{\theta}}^{-1} (\hat{\boldsymbol{\theta}} - \tilde{\boldsymbol{\theta}}^k)$$

12: **end if**
 13: Output: $\tilde{\boldsymbol{\theta}}$ and its associated model $\tilde{G}(p) = \tilde{B}(p)/\tilde{A}(p)$.

we need to choose M stable polynomials, which correspond to the centers of the stable ellipsoids. These polynomials should be chosen such that the global optimum (among the whole stability region) belongs to at least one ellipsoid. In practice, the user may choose to reflect the unstable poles obtained from $\hat{\boldsymbol{\theta}}$ (as in, e.g., lines 7 to 14 of Algorithm 2.1), or may choose the coefficients from any other continuous-time identification method that provides a stable estimate given the data. In that regard, any consistent estimator with a stable estimate will be a good candidate for producing a stable ellipsoid around the true parameter vector.

Remark 3.2. Since the matrix $\bar{\mathbf{H}}$ only depends on the number of poles of the model, whose structure is assumed fixed, it must only be computed once in the algorithm.

Remark 3.3. The proposed procedure can also be extended to incorporate other constraints on the poles of the model. For example, if it is known that the true system does not have poles with real parts greater than $a^* < 0$, it is only needed to adjust the parameter a in (3.12), recompute $\bar{\mathbf{H}}$ by using (3.13), (3.14) and (3.15), and then compute the ellipsoids accordingly.

3.5 Monte Carlo simulation studies

We now study the performance of the proposed estimator under a series of experiments. This section is divided into two parts. First, the IPEM estimator in (3.8) is tested against the standard indirect approach estimator and the SRIVC estimator. Later, the improvement that enforces stability is tested.

3.5.1 Tests with IPEM

We first test the IPEM estimator in (3.8) on the Rao-Garnier system [185], which is a benchmark system for linear continuous-time identification. It is a linear, fourth-order, non-minimum phase system with complex poles that has been tested in many previous contributions (see, e.g., [59, 139, 239]). The system is described by

$$G^*(p) = \frac{-6400p + 1600}{p^4 + 5p^3 + 408p^2 + 416p + 1600}. \quad (3.17)$$

This system is interesting since it has two damped oscillatory modes at 2 and 20[rad/s] with damping ratios of 0.1 and 0.25 respectively, and a non-minimum phase zero at $s = 0.25$. It has been reported that this is a particularly difficult system to estimate by PEM and ML methods, since these methods may converge to a local minimum if they are not well initialized [139].

Three methods are compared: PEM (i.e., the standard indirect approach estimator), IPEM, and SRIVC. Each method is tested in K different Monte Carlo runs, and they are evaluated according to the average normalized square error of the system estimate

$$\text{MSE } \hat{G} = \frac{1}{K} \sum_{i=1}^K \frac{\|\hat{G}_i - G^*\|_2^2}{\|G^*\|_2^2},$$

and the average fit measure

$$\text{Fit} = \frac{100}{K} \sum_{i=1}^K \left[1 - \frac{\|\mathbf{y} - \hat{\mathbf{y}}^i\|_2}{\|\mathbf{y} - \bar{\mathbf{y}}\|_2} \right],$$

where \mathbf{y} is the noiseless output sequence, $\hat{\mathbf{y}}^i$ is the simulated output sequence of the i -th estimated model, and $\bar{\mathbf{y}}$ is a constant vector formed by the average value of $\{y(kh)\}_{k=1}^N$.

We run PEM using the MATLAB System Identification Toolbox [140] with the `oe` command, and assume that the correct order of the system is known. The search algorithm for PEM is initialized with the estimate given by the weighted null space fitting (WNSF) method⁵ [58]. We base IPEM on the PEM estimate previously obtained. The required Jacobian matrix is numerically calculated via

⁵PEM initialized with the ZOH equivalent of the SRIVC estimate (as in [141]) was also tested with similar results.

Table 3.1: Monte Carlo results for PRBS input of lengths $N = 7161$ and $N = 1533$.

$h[\text{s}]$	Method	MSE \hat{G} (7161)	Fit (7161)	MSE \hat{G} (1533)	Fit (1533)
0.01	PEM	$1.11 \cdot 10^{-4}$	98.97	$5.88 \cdot 10^{-4}$	97.78
	IPEM	$0.73 \cdot 10^{-4}$	99.12	$4.03 \cdot 10^{-4}$	98.07
	SRIVC	$0.73 \cdot 10^{-4}$	99.12	$4.02 \cdot 10^{-4}$	98.07
0.05	PEM	$2.00 \cdot 10^{-3}$	98.83	$8.62 \cdot 10^{-4}$	97.67
	IPEM	$1.91 \cdot 10^{-3}$	98.96	$5.30 \cdot 10^{-4}$	97.99
	SRIVC	$1.54 \cdot 10^{-3}$	98.97	$5.27 \cdot 10^{-4}$	97.99
0.1	PEM	$3.28 \cdot 10^{-3}$	98.94	$1.96 \cdot 10^{-2}$	97.74
	IPEM	$1.91 \cdot 10^{-3}$	99.09	$1.08 \cdot 10^{-2}$	98.03
	SRIVC	$1.92 \cdot 10^{-3}$	99.09	$9.99 \cdot 10^{-3}$	97.98

finite differences, and the correct relative degree is imposed. We use the command `d2c` of MATLAB in both cases. The SRIVC algorithm is implemented with the CONTSID toolbox version 7.2 for MATLAB [61, Ch. 9] with default initialization, and it is set to estimate a continuous-time model with four poles and two zeros.

Effect of the number of data points and sampling rate

In this study, we design the input as a pseudorandom binary sequence (PRBS) of amplitude switches between 0 and 2. Two scenarios are analyzed with different number of data points. For the first input sequence, the number of stages of the shift register is ten and the data length of the shortest interval is seven, leading to a sequence of $N = 7161$ data points. The second input sequence has nine stages and the data length of the shortest interval is three, resulting in an input of $N = 1533$ data points. The output disturbance is a zero-mean Gaussian white noise signal in both cases, where the variance is set such that the SNR between the noiseless output sequence and the noise equals approximately 10[dB].

Two Monte Carlo studies are performed with $K = 500$ runs of different noise realizations each one, for $h = 0.01, 0.05$ and $0.1[\text{s}]$. The results are shown in Table 3.1. The results in that table show that the IPEM estimator statistically improves the estimates given by PEM, and is a competitive method against SRIVC, even for high frequency sampling. The IPEM estimator also outperforms PEM under less data points for every sampling period, which indicates that the asymptotic properties studied in Section 3.3 can be observed in practical finite data cases as well. Note that for $h = 0.05$ the input for the smaller sample size has more energy between the frequencies 14 and 22[rad/s] than the input for $N = 7161$. This explains why the methods perform better in terms of mean MSE for $N = 1533$ when $h = 0.05[\text{s}]$.

Remark 3.4. We have discarded cases where PEM gave estimates with one pole in the negative real axis for the results in Table 3.1. Fortunately this scenario is

Table 3.2: Estimated parameter value means and standard deviations for each method, considering $h = 0.05[\text{s}]$ and $N = 7161$.

Method	Parameter True value	b_3 0	b_2 0	b_1 −6400	b_0 1600	a_3 5	a_2 408	a_1 416	a_0 1600
PEM	Mean	−0.04	0.30	−6403.01	1601.33	5.01	408.16	416.25	1600.49
	Std. Dev.	0.96	11.41	147.13	47.85	0.40	7.98	9.27	33.29
IPEM	Mean	0	0	−6397.25	1599.79	4.99	407.88	415.84	1599.27
	Std. Dev.	0	0	122.39	42.39	0.32	7.11	8.33	28.59
SRIVC	Mean	0	0	−6399.19	1600.49	5.01	407.99	416.61	1599.71
	Std. Dev.	0	0	132.05	44.21	0.34	7.75	8.98	31.1

very uncommon (9 cases seen in 3000 runs). A similar phenomenon was observed for $h = 0.1[\text{s}]$ with the SRIVC estimator, where two out of 3000 estimates gave negative fits. These runs were not considered either.

Mean and standard deviation of the estimated parameters

As established in Theorem 3.2, the IPEM estimator should have (at least asymptotically) a smaller parameter covariance than that of the standard indirect approach method. To test this, we obtain the mean and standard deviation of each parameter given by the Monte Carlo study under the same setup as before, with $K = 500$, $h = 0.05[\text{s}]$, $N = 7161$. The results in Table 3.2 show that the mean values are similar for all methods (except standard PEM, which does not estimate the correct model structure), and IPEM provides the lowest standard deviation for every parameter.

Direct comparison with standard PEM

We now analyze the improvements of the proposed method over the PEM estimates. In particular, we focus on the Bode plots of the continuous-time models obtained by PEM and IPEM to show the impact of enforcing the correct relative degree. In Figure 3.3, we plot the frequency response estimates of 100 Monte Carlo runs for $h = 0.05[\text{s}]$ and $N = 7161$. It is clear that the improvement over PEM is mainly at high frequencies; this is intuitive, since the relative degree determines the asymptotic slope of the Bode diagram of magnitude. The proposed method enforces the correct asymptotic slope, leading to an important gain in accuracy in both Bode plots.

Also, direct comparison plots between PEM and IPEM are shown in Figure 3.4. These plots compare the normalized model error and the fit of PEM and IPEM for each Monte Carlo run (500 in total). IPEM outperforms the standard indirect approach in most Monte Carlo runs, especially in the fit comparison, where 496 out of 500 trials have led to an increase in fit under the proposed refinement.

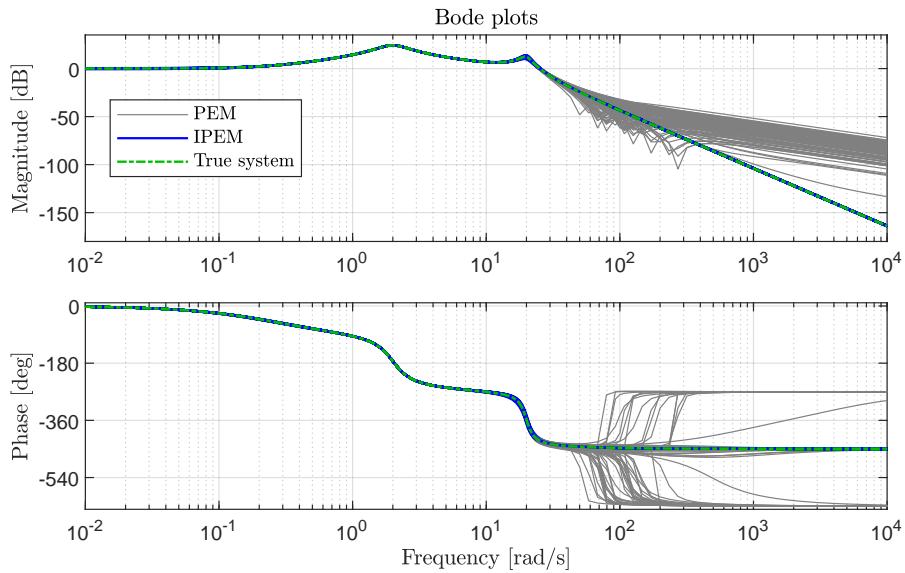


Figure 3.3: Bode diagrams of 100 estimates of $G^*(p)$ by PEM (grey), 100 estimates of $G^*(p)$ by IPEM (blue), and the true system (dash-dotted green).

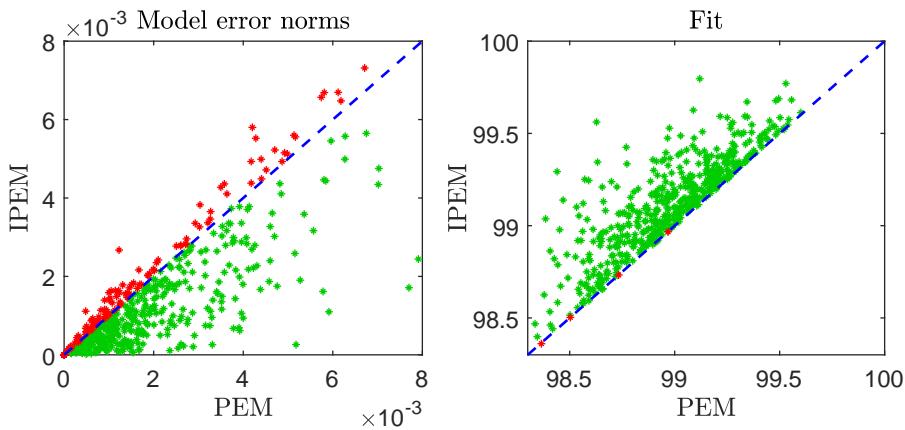


Figure 3.4: Direct comparison plots between PEM and IPEM for 500 Monte Carlo simulations. The green dots correspond to Monte Carlo simulations where IPEM outperforms PEM. Red dots represent the opposite, and the dashed blue line is the separatrix.

3.5.2 Tests with the robustified IPEM

Next, we test the robustified indirect PEM estimator, which we call IPEM-s. For the following simulations, the ellipsoids which yield stability are obtained through Theorem 3.5 by solving the LMI optimization problem with the CVX package [100] in MATLAB, using the SeDuMi solver. Two ellipsoids ($M = 2$) are computed for each stable estimate, with centers at the (reflected for stability, if necessary) standard indirect approach estimate and the reflected IPEM estimate. The SRIVC estimator is implemented with the default tolerance and initialization method provided in CONTSID 7.3 [60].

Several considerations are taken regarding the robustness of the indirect methods. PEM is initialized with the WNSF method [58]. If PEM returns an unstable estimate, we simply reflect the unstable poles. If the discrete-time estimate from PEM has a pole in the real negative axis, an n -th order approximation of the $(n + 1)$ -th order continuous-time estimate is obtained using the `balred` command in MATLAB.

Remark 3.5. The problems encountered here can be solved by enforcing stability directly in the discrete-time PEM estimate, proposing a different sampling period, or a different data length. These adjustments are not performed here only to simplify the comparisons.

Rao-Garnier system tests

We consider the Rao-Garnier system [185] described in (3.17). A PRBS signal of length $N = 1533$ is used as input and the sampling period is set to $h = 0.1[\text{s}]$, which is considered a large sampling period for this system. The additive white noise is Gaussian-distributed, with variance set to be half the variance of the noiseless output.

The box plots of the fit for each method is shown in Figure 3.5. All fits under 0 are grouped, and the number of outliers of this kind is recorded in the lower part of the box plots. IPEM returned 23 unstable estimates in this set of runs, which are all denoted as outliers in the box plot. Figure 3.5 shows that by forcing stability in an optimal manner, IPEM-s is the most robust method against bad outliers while being comparable to the other estimators in terms of median value.

We also compare the behavior of each estimator in the most challenging runs for each one. The worst 60 fits and normalized model errors are ordered according to increasing performance (ascending for fit, and descending for model errors) and registered in Figure 3.6. Here, unstable estimates by IPEM were bounded by -100 fit, and the normalized model error was bounded by 60 .⁶ These plots show that IPEM-s has a similar model error performance to SRIVC for the most challenging runs, and it is the preferred method in terms of fit.

⁶Here it is considered that unstable systems have 2-norm equal to infinity. Thus, for plotting the measure results, we set the values for unstable models at a fixed upper bound.

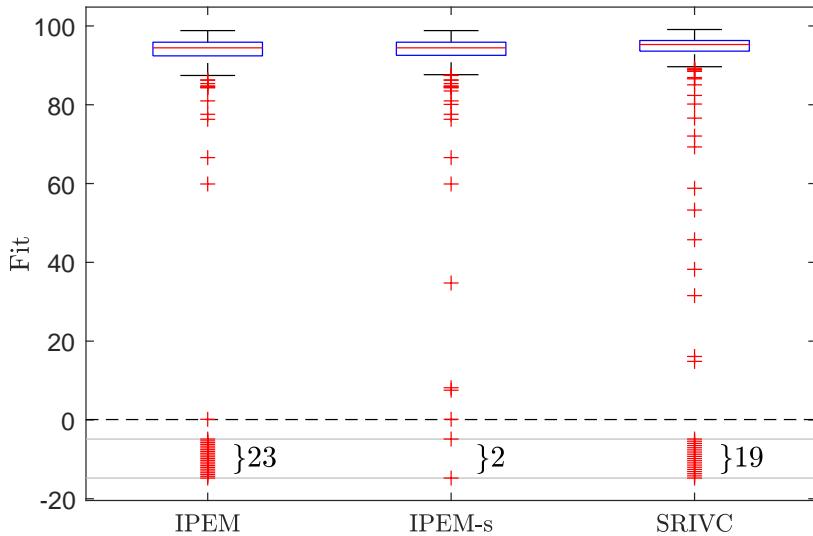


Figure 3.5: Fit box plots of IPEM, IPEM-s and the SRIVC method for the Rao-Garnier system experiment. Crosses in between the horizontal lines are compressed outliers (fits less than 0).

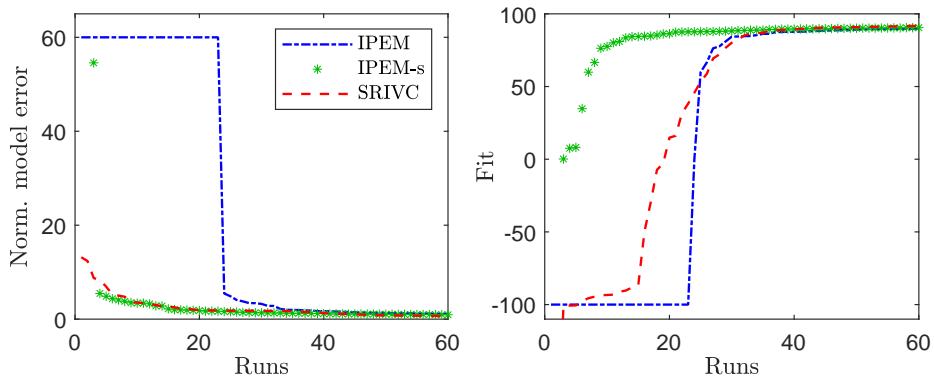


Figure 3.6: Performance in each metric of the worst 100 runs per method under the Rao-Garnier system experiments. IPEM (dash-dotted, blue), IPEM-s (asterisk, green), and SRIVC (dashed, red).

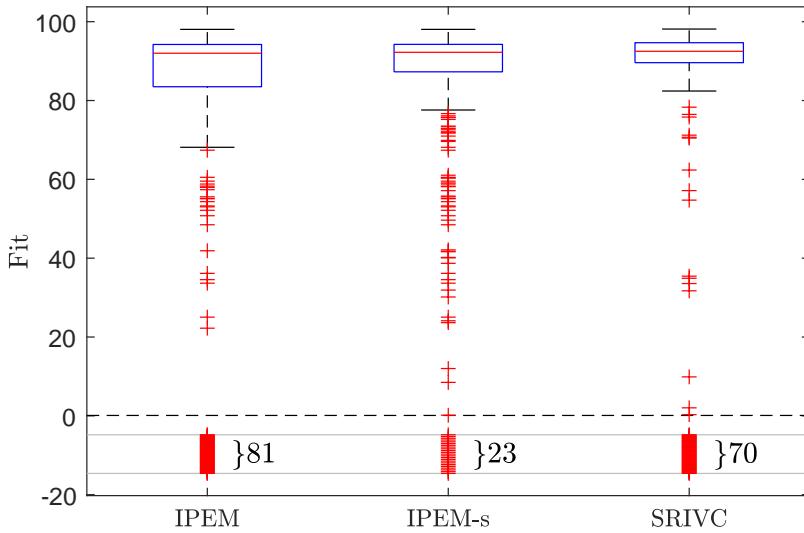


Figure 3.7: Fit box plots of IPEM, IPEM-s and the SRIVC method for the set of random systems. Crosses in between the horizontal lines are compressed outliers (fits less than 0).

Random systems tests

The IPEM-s estimator is tested on a set of 500 random continuous-time systems of order 3 and relative degree 2, which are generated with the `rss` command in MATLAB. The slowest pole of each continuous-time random system is set to have real part not larger than -0.2 . The input is a unit-variance Gaussian white noise of length $N = 500$, and the additive noise is also Gaussian and white, with variance such that the SNR is equal to 3[db]. The noisy output is sampled ten times faster than the fastest pole or zero of the true system. This set of systems is particularly challenging to model in this setup due to the poor signal-to-noise ratio and low number of data points.

We present box plots of the fit measure for all methods under study in Figure 3.7. IPEM-s reduces the number of outliers of IPEM from 81 to 23, and performs considerably better than SRIVC in terms of robustness. In addition, we have also determined the performance in the 100 most challenging runs in Figure 3.8. This time, unstable estimates from IPEM are chosen to have fit -100 and normalized model error equal to 12. From these plots we find that IPEM-s is the most robust method in terms of model errors and fit.

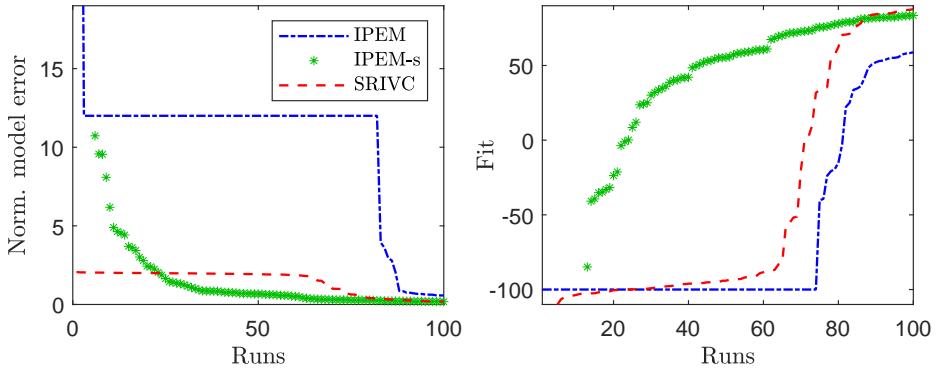


Figure 3.8: Performance in each metric of the worst 100 runs per method under the set of random systems. IPEM (dash-dotted, blue), IPEM-s (asterisk, green), and SRIVC (dashed, red).

3.6 Conclusions

We have proposed a refinement to the standard PEM estimator for indirect continuous-time system identification that has asymptotically optimal performance, and enforces the desired relative degree in the estimate. An explicit expression for this estimator has been found, and its consistency and asymptotic efficiency have been proven. Afterwards, a robustification procedure has been developed and discussed. By introducing convex inner approximations of the stability region in the indirect PEM framework, the robustified method guarantees the desired number of poles and zeros in the continuous-time model while enforcing stability in the estimate. Due to its construction, it enjoys optimal asymptotic properties and it is also robust for short and noisy data set scenarios where the IPEM estimates can be highly inaccurate. Extensive simulations using both standard benchmarks and random systems have been put forward with promising results for both proposed methods. These have shown that the refinement to the standard indirect approach leads to an important improvement in all the statistical metrics studied, while also being more robust compared to the other indirect and direct methods.

Appendix

3.A Theorem proofs

Here we present the proofs of Theorems 3.3 and 3.4.

3.A.1 Proof of Theorem 3.3

Proof. We partition the asymptotic covariance matrix of $\hat{\theta}$ as

$$\text{AsCov}(\hat{\theta}) = \left[\begin{array}{c|c|c|c} \Sigma_{00} & \Sigma_{01} & \Sigma_{02} & \Sigma_{03} \\ \hline \Sigma_{01}^\top & \Sigma_{11} & \Sigma_{12} & \Sigma_{13} \\ \hline \Sigma_{02}^\top & \Sigma_{12}^\top & \Sigma_{22} & \Sigma_{23} \\ \hline \Sigma_{03}^\top & \Sigma_{13}^\top & \Sigma_{23}^\top & \Sigma_{33} \end{array} \right],$$

where $\Sigma_{00} \in \mathbb{R}^{(r_1-1) \times (r_1-1)}$, $\Sigma_{11} \in \mathbb{R}^{(r_2-r_1) \times (r_2-r_1)}$, $\Sigma_{22} \in \mathbb{R}^{(r-r_2) \times (r-r_2)}$ and $\Sigma_{33} \in \mathbb{R}^{(2n-r+1) \times (2n-r+1)}$. We write $\text{AsCov}(\hat{\theta}) = \mathbf{C}\mathbf{C}^\top$, where \mathbf{C} is the Cholesky factor of $\text{AsCov}(\hat{\theta})$. After some algebra, the asymptotic covariance of $\tilde{\theta}^{r_1}$ can be written as

$$\begin{aligned} \text{AsCov}(\tilde{\theta}^{r_1}) &= \mathbf{C} \begin{bmatrix} \mathbf{0}_{r_1-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{2n-r_1+1} \end{bmatrix} \mathbf{C}^\top \\ &= \left[\begin{array}{c|c|c} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} \\ \hline \Sigma_{12}^\top & \Sigma_{22} & \Sigma_{23} \\ \hline \Sigma_{13}^\top & \Sigma_{23}^\top & \Sigma_{33} \end{array} \right] - \left[\begin{array}{c} \Sigma_{01}^\top \\ \hline \Sigma_{02}^\top \\ \hline \Sigma_{03}^\top \end{array} \right] \Sigma_{00}^{-1} \left[\begin{array}{c|c|c} \Sigma_{01} & \Sigma_{02} & \Sigma_{03} \end{array} \right]. \end{aligned}$$

Similarly, the asymptotic covariance of $\tilde{\theta}^{r_2}$ is

$$\text{AsCov}(\tilde{\theta}^{r_2}) = \left[\begin{array}{c|c} \Sigma_{22} & \Sigma_{23} \\ \hline \Sigma_{23}^\top & \Sigma_{33} \end{array} \right] - \left[\begin{array}{c|c} \Sigma_{02}^\top & \Sigma_{12}^\top \\ \hline \Sigma_{03}^\top & \Sigma_{13}^\top \end{array} \right] \left[\begin{array}{c|c} \Sigma_{00} & \Sigma_{01} \\ \hline \Sigma_{01}^\top & \Sigma_{11} \end{array} \right]^{-1} \left[\begin{array}{c|c} \Sigma_{02} & \Sigma_{03} \\ \hline \Sigma_{12} & \Sigma_{13} \end{array} \right].$$

Since we are only interested in the covariance of the truly non-zero parameters (which is a square matrix of $2n - r + 1$ columns), we consider the bottom right

submatrix of each asymptotic covariance matrix⁷:

$$\text{AsCov}(\tilde{\theta}^{r_1}) = \Sigma_{33} - \Sigma_{03}^\top \Sigma_{00}^{-1} \Sigma_{03},$$

$$\text{AsCov}(\tilde{\theta}^{r_2}) = \Sigma_{33} - \left[\begin{array}{c|c} \Sigma_{03}^\top & \Sigma_{13}^\top \end{array} \right] \left[\begin{array}{c|c} \Sigma_{00} & \Sigma_{01} \\ \hline \Sigma_{01}^\top & \Sigma_{11} \end{array} \right]^{-1} \left[\begin{array}{c} \Sigma_{03} \\ \hline \Sigma_{13} \end{array} \right].$$

Finally, by a matrix inequality related to Schur complements [262, Theorem 2.1] applied to $\text{AsCov}(\tilde{\theta}^{r_2})$, we conclude that

$$\text{AsCov}(\tilde{\theta}^{r_1}) - \text{AsCov}(\tilde{\theta}^{r_2}) =$$

$$\left[\begin{array}{c|c} \Sigma_{03}^\top & \Sigma_{13}^\top \end{array} \right] \left[\begin{array}{c|c} \Sigma_{00} & \Sigma_{01} \\ \hline \Sigma_{01}^\top & \Sigma_{11} \end{array} \right]^{-1} \left[\begin{array}{c} \Sigma_{03} \\ \hline \Sigma_{13} \end{array} \right] - \Sigma_{03}^\top \Sigma_{00}^{-1} \Sigma_{03} \succeq \mathbf{0},$$

which shows that, by considering a relative degree closer to its true value, the covariance matrix of the non-zero parameters can only decrease. \square

3.A.2 Proof of Theorem 3.4

Proof of Statement 1: We note that if the Cholesky factor of $\text{AsCov}(\hat{\theta})$ is denoted by \mathbf{C} , then the asymptotic covariance for $\tilde{\theta}$ is given by

$$\text{AsCov}(\tilde{\theta}) = \mathbf{C} \begin{bmatrix} \mathbf{0}_{r-1} & \mathbf{0} \\ \mathbf{0}^\top & \mathbf{I}_{2n-r+1} \end{bmatrix} \mathbf{C}^{-1} \text{AsCov}(\hat{\theta}) \mathbf{C}^{-\top} \begin{bmatrix} \mathbf{0}_{r-1} & \mathbf{0} \\ \mathbf{0}^\top & \mathbf{I}_{2n-r+1} \end{bmatrix} \mathbf{C}^\top$$

$$= \mathbf{C} \begin{bmatrix} \mathbf{0}_{r-1} & \mathbf{0} \\ \mathbf{0}^\top & \mathbf{I}_{2n-r+1} \end{bmatrix} \mathbf{C}^\top.$$

Due to the partition (3.9), we can write \mathbf{C} explicitly as

$$\mathbf{C} = \begin{bmatrix} \Sigma_{11}^{1/2} & \mathbf{0} \\ \Sigma_{12}^\top \Sigma_{11}^{-1/2} & (\Sigma_{22} - \Sigma_{12}^\top \Sigma_{11}^{-1} \Sigma_{12})^{1/2} \end{bmatrix},$$

leading to

$$\text{AsCov}(\hat{\theta}) - \text{AsCov}(\tilde{\theta}) = \mathbf{C} \begin{bmatrix} \mathbf{I}_{r-1} & \mathbf{0} \\ \mathbf{0}^\top & \mathbf{0}_{2n-r+1} \end{bmatrix} \mathbf{C}^\top = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12}^\top & \Sigma_{12}^\top \Sigma_{11}^{-1} \Sigma_{12} \end{bmatrix} \succeq \mathbf{0},$$

which proves the first statement. \square

Proof of Statement 2: For the particular case where the input is a white noise

⁷Here we permit an abuse of notation by now naming $\tilde{\theta}^{r_1}$ and $\tilde{\theta}^{r_2}$ as the parameter vectors containing only the truly non-zero parameters.

random sequence, the asymptotic covariance of the estimated discrete-time parameter vector $\hat{\boldsymbol{\eta}}$ can be computed by using the formula given in [138, pp. 291], which gives

$$\text{AsCov}(\hat{\boldsymbol{\eta}}) = \left[\frac{\lambda^2}{2\pi\sigma^2} \int_{-\pi}^{\pi} \frac{\partial H(e^{i\omega}, \boldsymbol{\eta})}{\partial \boldsymbol{\eta}} \Big|_{\boldsymbol{\eta}=\boldsymbol{\eta}^*} \left(\frac{\partial H(e^{i\omega}, \boldsymbol{\eta})}{\partial \boldsymbol{\eta}} \Big|_{\boldsymbol{\eta}=\boldsymbol{\eta}^*} \right)^H d\omega \right]^{-1}, \quad (3.18)$$

where $H(e^{i\omega}, \boldsymbol{\eta})$ is the frequency response of the discrete-time equivalent model. After some calculation, it can be found that

$$\frac{\partial H(e^{i\omega}, \boldsymbol{\eta})}{\partial \boldsymbol{\eta}} \Big|_{\boldsymbol{\eta}=\boldsymbol{\eta}^*} = \frac{1}{A^{*2}(e^{i\omega})} \mathbf{S}\boldsymbol{\Gamma}(e^{i\omega}),$$

where $\mathbf{S} \in \mathbb{R}^{2n \times 2n}$ is the Sylvester matrix associated to the polynomials $A^*(q)$ and $B^*(q)$ given by (3.10), which is non-singular since $A^*(q)$ and $B^*(q)$ are assumed coprime [210, Lemma A3.1], and $\boldsymbol{\Gamma}(e^{i\omega}) = [e^{i\omega(2n-1)}, e^{i\omega(2n-2)}, \dots, 1]^\top$. Hence, the integral in (3.18) can be related to the stationary covariance of the output of a single-input multi-output LTI model described by $\boldsymbol{\Gamma}(q)/A^{*2}(q)$ driven by a white noise of unit variance, which can be easily computed in a state-space formulation by solving a Lyapunov equation. This stationary covariance is precisely \mathbf{P} in (3.11). Thus, (3.18) reduces to $\text{AsCov}(\hat{\boldsymbol{\eta}}) = (\sigma^2/\lambda^2)(\mathbf{S}\mathbf{P}\mathbf{S}^\top)^{-1}$, from which we immediately conclude that $\text{AsCov}(\hat{\boldsymbol{\theta}}) = (\sigma^2/\lambda^2)(\mathbf{J}^\top \mathbf{S}\mathbf{P}\mathbf{S}^\top \mathbf{J})^{-1}$. \square

Chapter 4

Consistency and asymptotic efficiency of the SRIVC estimator

In the previous chapter we introduced a novel indirect approach algorithm for continuous-time system identification and studied its properties. In Chapters 4 to 8 we will focus our attention on the SRIVC method, which is one of the most popular direct methods for continuous-time system identification [59]. This chapter presents a comprehensive theoretical analysis of the consistency and asymptotic efficiency of the SRIVC estimator, and the results are corroborated through extensive simulations.

4.1 Introduction

The refined instrumental variable method for continuous-time systems (RIVC) and its simplest embodiment, the simplified RIVC (SRIVC) [255] are considered to be the most reliable algorithms in linear continuous-time system identification [61, 63] and have been used in practice for almost 40 years. These methods have been successfully applied in many application areas, such as in semiconductors [69], distillation columns [66], biological [165] and environmental modeling [254].

Despite the success that the RIVC and SRIVC algorithms have achieved in practical applications [250, 254], there has been a lack of theoretical support available for them. Most of the discussions with respect to the properties of these estimators, such as consistency and asymptotic efficiency, are based on empirical observations [61, 246, 248, 250, 257], or on well-developed theoretical results in discrete-time system identification [34, 135, 252] which do not provide a mechanism to include the intersample behavior of the continuous-time signals of interest. Including knowledge of the signals between measurements is crucial in continuous-time system identification, as it is known that an incorrect interpolation of the sampled signals can lead to unsatisfactory estimates of the system parameters [181]. Also, the intersample behavior of the input and output is commonly used for estimating (prefiltered versions of) the time derivatives of the signals in direct continuous-time methods.

Since these derivatives are in general unknown to the user, the continuous-time identification problem can become more difficult than its discrete-time counterpart, as the discrete-time equivalent of derivatives, time-shifts, are directly available in discrete-time frameworks.

The objective of this chapter is to provide some theoretical support in terms of consistency and asymptotic efficiency for the SRIVC estimator. The results are based on the works [164], [167], and [166]. In summary, in this chapter,

- we analyze the consistency of the SRIVC estimator by incorporating the intersample behavior of the input in the analysis. We prove that the SRIVC estimator is generically consistent under mild conditions in the presence of additive white or colored noise on the output;
- we show that the SRIVC estimator remains generically consistent if an incorrect assumption of the intersample behavior is used for generating the filtered signals in the instrument vector, but is generically not consistent if an incorrect assumption on the intersample behavior is used for filtering the input signal in the regressor vector;
- we derive an expression for the asymptotic Cramér-Rao lower bound (CRLB) and show that the asymptotic covariance of the SRIVC estimates coincides with the asymptotic CRLB when the intersample behavior of the input signal in the filtered regressor and instrument vectors matches that of the system input and the additive noise is white and Gaussian; and
- we confirm our findings through extensive Monte Carlo simulations.

The outline of this chapter is as follows. We first provide an overview of the related work in Section 4.1.1. The system and model definitions are stated in Section 4.2. We present the main theoretical results in Sections 4.3 and 4.4, namely, the consistency and asymptotic efficiency proofs respectively, together with their related corollaries and remarks. In Section 4.5, we provide simulation results that support the theoretical analysis, and the chapter is summarized in Section 4.6. Supplementary results can be found in the appendix related to this chapter.

4.1.1 Related work

In the literature there have been several discussions regarding the statistical properties of the SRIVC method. Here we will refer to the most relevant works, and detail the main differences between them and the results in this thesis.

The consistency analysis in the current chapter has some similarities to the work in [135] as both analyses are based on the consistency result found in Theorem 4.5 of [210] developed for the discrete-time bootstrapped instrumental variable method. The work in [135] analyzes the convergence of the RIVC estimator with an autoregressive noise model. There are, however, a few shortcomings associated with this analysis. An extra filter is introduced for the purpose of discretizing the

derivatives of the input signal, which is not part of the RIVC implementation. This extra filter introduces unnecessary complexity into the analysis and its role can be replaced by filters that are already part of the algorithm. In addition, non-causal filters may arise in the formulation of the proof in [135] since the system and model are allowed to be parameterized as biproper transfer functions. Furthermore, the first step to establish the convergence and consistency of the RIVC algorithm is to show that the estimator is well-posed. This relies on 1) the Sylvester matrices constructed from the system and model polynomials being non-singular, and 2) the expectation of the two filtered input vectors, denoted by Φ in [135], being non-singular. Firstly, in [135], the non-singularity of the Sylvester matrix does not comply with their given assumptions. The analysis assumes that the system and model are parameterized with monic denominator polynomials. It can be shown that this implicit assumption, together with Assumption A5 in [135] on the model order, results in the Sylvester matrix constructed from the system polynomials being singular when the degree of the model denominator is greater than that of the system. The proof of Theorem 1 in [135] with respect to Assumption A5 therefore cannot proceed once the Sylvester matrix is singular. Secondly, it is stated in [135] that proving the non-singularity of Φ relies on the matrix $\mathbb{E}\{UU^\top\}$ being non-singular, where U is the vector containing the input samples with sample size N (see (19) in [135]). We note that $\mathbb{E}\{UU^\top\}$ has dimension $N \times N$, while its rank is at most equal to the order of persistence of excitation of the input. Thus, in the asymptotic case, it is not possible to conclude that Φ is always non-singular under the persistent excitation assumption given in [135].

Other work related to consistency such as [34] assumes that the model structure is exactly known and does not take into account the intersample behavior of the input as part of the analysis. The work in [252] describes a unified refined instrumental variable approach for estimating discrete or continuous-time transfer functions characterized by a unified operator that can be interpreted in terms of a backward shift, derivative or delta operator. This unified approach is suggested to be optimal in maximum likelihood, prediction error minimization and instrumental variable terms under the Box-Jenkins model structure for both discrete and continuous-time [252]. However, only limited theoretical analysis is provided with respect to the consistency of the estimates by using an incremental implementation of the algorithm with no explicit mention of the intersample behavior of the signals. By neglecting the intersample behavior as part of the analysis, the results in [34], [135] and [252] have overlooked the possibility that the converging point of the estimator no longer corresponds to the true system parameters when the system input cannot be interpolated exactly.

With respect to the efficiency analysis of the SRIVC or RIVC estimators, the existing literature [30, 59, 61, 66] have claimed that these estimators are asymptotically efficient with additive colored noise and white noise on the output respectively. Even though CRLB expressions for the SRIVC and RIVC estimators exist in the literature, these expressions have been obtained directly through the results for discrete-time instrumental variable algorithms developed in [208, 210, 219], and

thus have overlooked the effects of the intersample behavior of the signals in the analysis. In fact, the CRLB expression stated in page 105 of [61] is dependent on the filtered version of the sampled noise-free system output, which means that the entries in the existing CRLB depend on the interpolation of the output signal prior to the continuous-time filtering operations. This implies that assuming a ZOH reconstruction on the output will yield a different CRLB expression than assuming an FOH. Hence, the existing results in the literature are somewhat ambiguous and do not properly account for the behavior between samples.

In summary, to the best of the author's knowledge, formal derivations of the consistency and asymptotic efficiency of the SRIVC estimator that carefully address the intersample behavior of the measured signals are not available in the literature.

4.2 System and model setup

Consider an LTI, causal, asymptotically stable, continuous-time system of the form

$$x(t) = \frac{B^*(p)}{A^*(p)} u(t), \quad (4.1)$$

where the system numerator and denominator polynomials are assumed coprime with orders given by m^* and n^* respectively, i.e.,

$$\begin{aligned} B^*(p) &= b_{m^*}^* p^{m^*} + b_{m^*-1}^* p^{m^*-1} + \cdots + b_0^*, \\ A^*(p) &= a_{n^*}^* p^{n^*} + a_{n^*-1}^* p^{n^*-1} + \cdots + a_1^* p + 1, \end{aligned}$$

and the system parameter vector is described by (2.6). The output observation equation of the continuous-time system (4.1) at a sample instant t_k is given by

$$y(t_k) = x(t_k) + v(t_k), \quad (4.2)$$

where $x(t_k)$ is the unobserved, noise-free output, and $v(t_k)$ can be described by a zero-mean stochastic process. The model is parameterized as a proper transfer function $B(p)/A(p)$ of the form (2.3), where the unknown parameter vector is given by (2.7). We assume that sampled input values $u(t_k)$ and noisy output measurements $y(t_k)$ are obtained from an identification experiment. In this setup, the goal is to provide a comprehensive proof of the consistency and asymptotic efficiency of the SRIVC estimator described in Algorithm 2.1, and to state the conditions under which these properties are not achieved.

4.3 Consistency analysis

In this section we develop a theorem that establishes the generic consistency of the SRIVC estimator, as well as some corollaries and remarks that examine the consistency under different intersample conditions. Some additional lemmas required

by the proof of the consistency theorem are presented in the appendix of this chapter. The analysis is presented for the case of a SISO, LTI, asymptotically stable system with regularly sampled data.

We next state the main assumptions required in Theorem 4.1 for the SRIVC estimator to be generically consistent. First, the true system must satisfy the following assumptions:

Assumption 4.1. The true system $B^*(p)/A^*(p)$ is proper ($n^* \geq m^*$) and asymptotically stable with $A^*(p)$ and $B^*(p)$ being coprime.

The input sequence is assumed to be independent of the noise and sufficiently rich in spectral content, as stated in Assumptions 4.2 and 4.3, respectively:

Assumption 4.2. The input sequence, $\{u(t_k)\}$, and disturbance, $\{v(t_l)\}$, are stationary and mutually independent for all integers k and l .

Assumption 4.3. The input sequence is persistently exciting of order no less than $2n + 1$.

Next, we state a stability assumption on the models obtained by each iteration of the SRIVC algorithm, which plays a role in the well-posedness of the SRIVC estimator.

Assumption 4.4. All the zeros of the j -th iteration of the model denominator estimate $A_j(p)$ have strictly negative real parts, with $A_j(p)$ and $B_j(p)$ being coprime.

Note that unstable zeros in the denominator polynomial may arise when estimating the parameters of a transfer function. A simple solution is to reflect the unstable zeros over the imaginary axis by the procedure detailed in lines 7 to 14 of Algorithm 2.1. Hence, Assumption 4.4 is commonly satisfied in practice.

To ensure a unique solution for the model parameters, we must assume that over-parameterization can only occur in the numerator or denominator but not in both simultaneously. Also, there is a limit on the amount of over-parameterization allowed in the model denominator to ensure that the transfer functions in the modified normal matrix of the SRIVC estimator are proper for theoretical analysis. These conditions are formalized in Assumption 4.5.

Assumption 4.5. The degrees n and m of the polynomials in the model satisfy $\min(n - n^*, m - m^*) = 0$ and $n - n^* \leq n^* - m^*$.

Finally, Assumption 4.6 avoids the problem of aliasing and ensures a meaningful model to be obtained according to the Shannon-Nyquist theorem:

Assumption 4.6. The sampling frequency is more than twice the largest positive imaginary part of the zeros of $A_j(p)A^*(p)$.

Before presenting the main theorem on generic consistency of the SRIVC estimator, we shall introduce some signals of interest. By substituting

$$y(t_k) = \frac{B^*(p)}{A^*(p)} u(t_k) + v(t_k)$$

into (2.32), we can express the filtered regressor vector of the SRIVC estimator as

$$\varphi_f(t_k) = \tilde{\varphi}_f(t_k) + \Delta(t_k) - \mathbf{v}_f(t_k), \quad (4.3)$$

where the noise-free, interpolation-error-free regressor vector is given by

$$\tilde{\varphi}_f(t_k) = \left[\frac{-pB^*(p)}{A_j(p)A^*(p)} u(t_k), \dots, \frac{-pnB^*(p)}{A_j(p)A^*(p)} u(t_k), \frac{1}{A_j(p)} u(t_k), \dots, \frac{p^m}{A_j(p)} u(t_k) \right]^\top, \quad (4.4)$$

the filtered noise vector $\mathbf{v}_f(t_k)$ is given by

$$\mathbf{v}_f(t_k) = \left[\frac{p}{A_j(p)} v(t_k), \dots, \frac{p^n}{A_j(p)} v(t_k), 0, \dots, 0 \right]^\top, \quad (4.5)$$

and $\Delta(t_k)$ is a vector that contains the interpolation errors that arise from constructing the filtered derivatives of the output with entries given by the difference between the noise-free version of the regressor vector and $\tilde{\varphi}_f(t_k)$ in (4.4), i.e.,

$$\Delta_i(t_k) = \begin{cases} \frac{p^i B^*(p)}{A_j(p)A^*(p)} u(t_k) - \frac{p^i}{A_j(p)} \left\{ \frac{B^*(p)}{A^*(p)} u(t) \right\}_{t=t_k} & \text{if } i = 1, \dots, n, \\ 0 & \text{otherwise.} \end{cases} \quad (4.6)$$

We are ready to present the main theorem of this section, namely, the generic consistency of the SRIVC estimator. For Theorem 4.1 we shall assume that the intersample behavior of the input is provided by an FOH device, but it can equally be chosen with a ZOH intersample behavior as shown later in Corollary 4.1. Furthermore, the prefilters for the output (see (2.32) and (2.34)) are assumed to be discretized with the same hold as the sampled input, i.e., an FOH.

Theorem 4.1 (Generic consistency of the SRIVC estimator). *Consider the SRIVC estimator described in Algorithm 2.1, and suppose Assumptions 4.1 to 4.6 hold. Then, for an input with an FOH intersample behavior, the following statements are true:*

1. *The matrix $\mathbb{E}\{\hat{\varphi}_f(t_k)\varphi_f^\top(t_k)\}$ is generically non-singular with respect to the system and model denominator provided that the condition*

$$\|\mathbb{E}\{\hat{\varphi}_f(t_k)\Delta^\top(t_k)\}\|_2 < \sigma_{\min}(\mathbb{E}\{\hat{\varphi}_f(t_k)\tilde{\varphi}_f^\top(t_k)\}) \quad (4.7)$$

holds, where $\hat{\varphi}_f(t_k)$, $\tilde{\varphi}_f(t_k)$ and $\Delta(t_k)$ are defined as in (2.33), (4.4) and (4.6), respectively.

2. If (4.7) is satisfied and the iterations of the SRIVC estimator converge for all N sufficiently large to, say, $\bar{\theta}^N$, then the true parameter θ^* is the unique converging point of θ^N as the sample size tends to infinity.
3. As the sample size N approaches infinity, the algorithm locally converges to θ^* provided that the matrix $\mathbf{Q}^{-1}\mathbf{R}$ has all eigenvalues with magnitude less than 1, where

$$\mathbf{Q} = \mathbb{E}\{\hat{\varphi}_f(t_k)\varphi_f^\top(t_k)\}, \quad (4.8)$$

and

$$\mathbf{R} = \mathbb{E}\{\hat{\varphi}_f(t_k)\bar{\Delta}^\top(t_k)\}, \quad (4.9)$$

where the vector $\bar{\Delta}(t_k)$ has entries

$$\bar{\Delta}_i(t_k) = \begin{cases} \frac{p^i \bar{B}(p)}{\bar{A}^2(p)} u(t_k) - \frac{p^i}{\bar{A}(p)} \left\{ \frac{B^*(p)}{A^*(p)} u(t) \right\}_{t=t_k} & \text{if } i = 1, \dots, n, \\ 0 & \text{otherwise.} \end{cases} \quad (4.10)$$

Proof of Statement 1: By exploiting the decomposition in (4.3), the modified normal matrix $\mathbb{E}\{\hat{\varphi}_f(t_k)\varphi_f^\top(t_k)\}$ can be written as

$$\mathbb{E}\{\hat{\varphi}_f(t_k)\varphi_f^\top(t_k)\} = \mathbb{E}\{\hat{\varphi}_f(t_k)\tilde{\varphi}_f^\top(t_k)\} - \mathbb{E}\{\hat{\varphi}_f(t_k)\mathbf{v}_f^\top(t_k)\} + \mathbb{E}\{\hat{\varphi}_f(t_k)\Delta^\top(t_k)\}.$$

The term $\mathbb{E}\{\hat{\varphi}_f(t_k)\mathbf{v}_f^\top(t_k)\}$, which is the term related to the filtered noise, is shown to be equal to zero in Lemma 4.1 of Appendix 4.A. Thus,

$$\mathbb{E}\{\hat{\varphi}_f(t_k)\varphi_f^\top(t_k)\} = \mathbb{E}\{\hat{\varphi}_f(t_k)\tilde{\varphi}_f^\top(t_k)\} + \mathbb{E}\{\hat{\varphi}_f(t_k)\Delta^\top(t_k)\}. \quad (4.11)$$

The approach we follow consists in proving that $\mathbb{E}\{\hat{\varphi}_f(t_k)\tilde{\varphi}_f^\top(t_k)\}$ is generically non-singular, and that the perturbation matrix $\mathbb{E}\{\hat{\varphi}_f(t_k)\Delta^\top(t_k)\}$ is small enough (in 2-norm) to not affect the non-singularity of the matrix $\mathbb{E}\{\hat{\varphi}_f(t_k)\tilde{\varphi}_f^\top(t_k)\}$. With regards to the role of $\mathbb{E}\{\hat{\varphi}_f(t_k)\Delta^\top(t_k)\}$, the additive perturbation result in Theorem 5.1 of [41] defines the smallest size of this perturbation matrix that would make (4.11) singular. The condition imposed by (4.7) therefore ensures the (generic) non-singularity of (4.11) provided that $\mathbb{E}\{\hat{\varphi}_f(t_k)\tilde{\varphi}_f^\top(t_k)\}$ is generically non-singular.

In the sequel we prove the generic non-singularity of $\mathbb{E}\{\hat{\varphi}_f(t_k)\tilde{\varphi}_f^\top(t_k)\}$. The noise-free, interpolation-error-free regressor vector (4.4) can be expressed as

$$\tilde{\varphi}_f(t_k) = \frac{1}{A_j(p)A^*(p)} \mathbf{r}(t_k), \quad (4.12)$$

where

$$\mathbf{r}(t_k) = \left[-pB^*(p)u(t_k), \dots, -p^n B^*(p)u(t_k), A^*(p)u(t_k), \dots, p^m A^*(p)u(t_k) \right]^\top.$$

Note that we write (4.12) with a slight abuse of notation, since in this case the \mathbf{r} vector is never evaluated at t_k without incorporating $[A_j(p)A^*(p)]^{-1}$ to the filters of each entry.

The highest order derivative in $\mathbf{r}(t_k)$ according to Assumption 4.5 is given by

$$\max(n + m^*, n^* + m) = n + m - \min(n - n^*, m - m^*) = n + m.$$

The vector $\mathbf{r}(t_k)$ can then be expressed as a product of an $(n + m + 1) \times (n + m + 1)$ Sylvester matrix and a vector containing the derivatives of $u(t_k)$, i.e.,

$$\mathbf{r}(t_k) = \mathbf{S}(-B^*(p), A^*(p))\mathbf{u}_{n+m}(t_k), \quad (4.13)$$

where

$$\mathbf{S}(-B^*(p), A^*(p)) = \begin{bmatrix} 0 & -b_{m^*}^* & -b_{m^*-1}^* & \cdots & -b_0^* & 0 \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ -b_{m^*}^* & -b_{m^*-1}^* & \cdots & -b_0^* & 0 & 0 \\ \hline 0 & a_{n^*}^* & a_{n^*-1}^* & \cdots & a_1^* & 1 \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ a_{n^*}^* & a_{n^*-1}^* & \cdots & a_1^* & 1 & 0 \end{bmatrix}, \quad (4.14)$$

where n rows consist of numerator coefficients and $m + 1$ rows of denominator coefficients (columns of zeros are padded to the left if necessary), and $\mathbf{u}_{n+m}(t_k)$ is formed by derivatives of the input signal, as

$$\mathbf{u}_{n+m}(t_k) = [p^{n+m}, p^{n+m-1}, \dots, 1]^\top u(t_k). \quad (4.15)$$

Note that the matrix in (4.14) is written as a permutation of a standard Sylvester matrix [210, Eq. A3.3]. From now on, the argument p in all Sylvester matrices is omitted for simplicity of notation. We require (4.14) to remain non-singular under the three possible cases imposed by Assumption 4.5, i.e., 1) the order of the true system is known exactly, 2) the numerator of the model is over-parameterized, and 3) the denominator of the model is over-parameterized. Under condition 1), the Sylvester matrix in (4.14) is non-singular when $B^*(p)$ and $A^*(p)$ are coprime [210, Lemma A3.1]. Under condition 2), when the numerator polynomial is over-parameterized, the first $m - m^*$ columns of the top half of (4.14) are filled with zeros. Similarly, under condition 3), when the denominator polynomial is over-parameterized, the first $n - n^*$ columns of the bottom half of (4.14) are filled with zeros. In all three cases, (4.14) does not lose rank since it is guaranteed that there is at least one non-zero entry in each column due to the anti-monic model denominator assumption. Therefore, $\mathbf{S}(-B^*, A^*)$ is non-singular under Assumption 4.5.

Now, following similar steps as above, the filtered instrument vector in (2.33) can be written as

$$\hat{\varphi}_f(t_k) = \mathbf{S}(-B_j, A_j) \frac{1}{A_j^2(p)} \mathbf{u}_{n+m}(t_k), \quad (4.16)$$

where $\mathbf{S}(-B_j, A_j)$ is an $(n+m+1) \times (n+m+1)$ Sylvester matrix defined in the same way as (4.14) with $a_1^*, \dots, a_{n^*}^*, b_0^*, \dots, b_{m^*}^*$ replaced by $a_1, \dots, a_n, b_0, \dots, b_m$, which is also non-singular, this time due to Assumption 4.4. Substituting (4.12), (4.13) and (4.16) into $\mathbb{E}\{\hat{\varphi}_f(t_k)\hat{\varphi}_f^\top(t_k)\}$, we obtain

$$\mathbb{E}\{\hat{\varphi}_f(t_k)\hat{\varphi}_f^\top(t_k)\} = \mathbf{S}(-B_j, A_j)\Phi\mathbf{S}^\top(-B^*, A^*), \quad (4.17)$$

where

$$\Phi := \mathbb{E}\left\{\frac{1}{A_j^2(p)}\mathbf{u}_{n+m}(t_k)\frac{1}{A_j(p)A^*(p)}\mathbf{u}_{n+m}^\top(t_k)\right\}. \quad (4.18)$$

Since $\mathbf{S}(-B_j, A_j)$ and $\mathbf{S}(-B^*, A^*)$ are non-singular, in order for (4.17) to be generically non-singular, it is sufficient to show that Φ is generically non-singular. In Appendix 4.A we show that Φ is positive definite when evaluated at the true system parameters (see Lemma 4.2). Moreover, for a fixed input signal, by Lemma 4.4 in the same appendix, every entry of Φ is a real analytic function in the joint variables $[a_1, \dots, a_n]$. Hence, for degrees of the model denominator and numerator polynomials that satisfy $n = n^*$ and $m \geq m^*$, it can be concluded by Lemma A2.3 of [210] and its corollary¹ that Φ is generically non-singular and thus $\mathbb{E}\{\hat{\varphi}_f(t_k)\hat{\varphi}_f^\top(t_k)\}$ is generically non-singular. In the case of the model denominator being over-parameterized under Assumption 4.5, Lemma 4.2 shows that Φ is non-singular at a point defined on the boundary of the domain given by the open set, since the over-parameterized component of $A_j(p)$ is set to zero when Φ is evaluated at $A_j(p) = A^*(p)$. By the continuity of the entries of Φ , there exists a perturbation sufficiently small such that Φ is non-singular when evaluated at $A_j(p) = A^*(p) + \epsilon(p)$, where the parameters of the polynomial $\epsilon(p)$ form the perturbation vector. The matrix Φ can therefore be shown to be generically non-singular with respect to the joint variables $[a_1, \dots, a_n]$, which concludes the proof. \square

Proof of Statement 2: Here we show that, upon convergence, the limiting point of the SRIVC estimator corresponds to the true parameters. Suppose $\bar{\theta}^N$ is a limiting point of the iteration in (2.35). The ergodic lemmas in [205, Lemma 3.1] and [210, Lemma A4.3] permit us to write (2.35), at the converging point and as N tends to infinity, as

$$\mathbb{E}\{\hat{\varphi}_f(t_k, \bar{\theta})\varphi_f^\top(t_k, \bar{\theta})\}^{-1}\mathbb{E}\{\hat{\varphi}_f(t_k, \bar{\theta})\varepsilon(t_k, \bar{\theta})\} = \mathbf{0}, \quad (4.19)$$

where $\bar{\theta} = \lim_{N \rightarrow \infty} \bar{\theta}^N$ and $\varepsilon(t_k, \bar{\theta})$ is the GEE (2.28) evaluated at the converging point. Since the matrix inverse in (4.19) is non-singular by Statement 1, the second expectation in (4.19) must be zero, i.e.,

$$\mathbb{E}\{\hat{\varphi}_f(t_k, \bar{\theta})\varepsilon(t_k, \bar{\theta})\} = \mathbf{0}. \quad (4.20)$$

¹This lemma and corollary are stated for completeness as Lemma 4.6 and Corollary 4.4 in Appendix 4.C.

Now, the GEE in (2.28), when evaluated at the converging point, can be rearranged to be

$$\varepsilon(t_k, \bar{\boldsymbol{\theta}}) = \frac{\bar{A}(p)B^*(p) - \bar{B}(p)A^*(p)}{\bar{A}(p)A^*(p)} u(t_k) + v(t_k), \quad (4.21)$$

where $\bar{B}(p)$ and $\bar{A}(p)$ are the corresponding polynomials of the model, i.e.,

$$\begin{aligned}\bar{B}(p) &= \bar{b}_m p^m + \bar{b}_{m-1} p^{m-1} + \cdots + \bar{b}_0, \\ \bar{A}(p) &= \bar{a}_n p^n + \bar{a}_{n-1} p^{n-1} + \cdots + \bar{a}_1 p + 1.\end{aligned}$$

Note that the polynomials $\bar{B}(p)$ and $\bar{A}(p)$ are coprime since $\bar{\boldsymbol{\theta}}$ satisfies the conditions in Statement 1. Let $\bar{A}(p)B^*(p) - \bar{B}(p)A^*(p) = h_r p^r + h_{r-1} p^{r-1} + \cdots + h_0$, where $r = \max(n+m^*, n^*+m) = n+m$. Then, the GEE can be expressed as

$$\varepsilon(t_k, \bar{\boldsymbol{\theta}}) = \frac{1}{\bar{A}(p)A^*(p)} \left[p^{n+m} u(t_k), \ p^{n+m-1} u(t_k), \ \dots, \ u(t_k) \right] \mathbf{h} + v(t_k), \quad (4.22)$$

where

$$\mathbf{h} := \left[h_{n+m}, \ h_{n+m-1}, \ \dots, \ h_0 \right]^\top. \quad (4.23)$$

Now, substituting (4.16) for $\hat{\boldsymbol{\varphi}}(t_k, \bar{\boldsymbol{\theta}})$ and (4.22) for $\varepsilon(t_k, \bar{\boldsymbol{\theta}})$ into the expectation in (4.20), we obtain

$$\mathbb{E}\{\hat{\boldsymbol{\varphi}}_f(t_k, \bar{\boldsymbol{\theta}})\varepsilon(t_k, \bar{\boldsymbol{\theta}})\} = \mathbf{S}(-\bar{B}, \bar{A})\bar{\boldsymbol{\Phi}}\mathbf{h} + \mathbb{E}\{\hat{\boldsymbol{\varphi}}_f(t_k, \bar{\boldsymbol{\theta}})v(t_k)\}, \quad (4.24)$$

where $\bar{\boldsymbol{\Phi}}$ is (4.18) evaluated at the converging point. By following the same procedure as the proof of Statement 1, we can show that $\bar{\boldsymbol{\Phi}}$ is generically non-singular and that $\mathbb{E}\{\hat{\boldsymbol{\varphi}}_f(t_k, \bar{\boldsymbol{\theta}})v(t_k)\} = \mathbf{0}$. Thus, for (4.24) to be zero, we must have $\mathbf{h} = \mathbf{0}$, implying $\bar{A}(p)B^*(p) - \bar{B}(p)A^*(p) = 0$ or equivalently $\bar{B}(p)/\bar{A}(p) = B^*(p)/A^*(p)$, i.e., $\bar{\boldsymbol{\theta}}^*$ is the unique limiting point. \square

Proof of Statement 3. Let $\bar{\boldsymbol{\theta}}$ be the limiting point. Then $\boldsymbol{\theta}_{j+1} - \bar{\boldsymbol{\theta}} = f_1(\boldsymbol{\theta}_j)f_2(\boldsymbol{\theta}_j)$ as N tends to infinity, where

$$f_1(\boldsymbol{\theta}_j) = \mathbb{E}\{\hat{\boldsymbol{\varphi}}_f(t_k, \boldsymbol{\theta}_j)\boldsymbol{\varphi}_f^\top(t_k, \boldsymbol{\theta}_j)\}^{-1},$$

and

$$f_2(\boldsymbol{\theta}_j) = \mathbb{E}\{\hat{\boldsymbol{\varphi}}_f(t_k, \boldsymbol{\theta}_j)(y_f(t_k, \boldsymbol{\theta}_j) - \boldsymbol{\varphi}_f^\top(t_k, \boldsymbol{\theta}_j)\bar{\boldsymbol{\theta}})\}.$$

To examine how the SRIVC estimate behaves around the limiting point, we can linearize $\boldsymbol{\theta}_{j+1}$ around $\bar{\boldsymbol{\theta}}$ using a first order Taylor series, that is,

$$\begin{aligned}\boldsymbol{\theta}_{j+1} - \bar{\boldsymbol{\theta}} &= f_1(\bar{\boldsymbol{\theta}})f_2(\bar{\boldsymbol{\theta}}) + \left(\frac{\partial f_1(\boldsymbol{\theta}_j)}{\partial \boldsymbol{\theta}_j} \Big|_{\boldsymbol{\theta}_j=\bar{\boldsymbol{\theta}}} f_2(\bar{\boldsymbol{\theta}}) + f_1(\bar{\boldsymbol{\theta}}) \frac{\partial f_2(\boldsymbol{\theta}_j)}{\partial \boldsymbol{\theta}_j} \Big|_{\boldsymbol{\theta}_j=\bar{\boldsymbol{\theta}}} \right) (\boldsymbol{\theta}_j - \bar{\boldsymbol{\theta}}) \\ &\quad + o_p(\|\boldsymbol{\theta}_j - \bar{\boldsymbol{\theta}}\|).\end{aligned}$$

At the limiting point, $f_2(\bar{\boldsymbol{\theta}}) = \mathbf{0}$ as given by (4.20). Hence,

$$\boldsymbol{\theta}_{j+1} - \bar{\boldsymbol{\theta}} = f_1(\bar{\boldsymbol{\theta}}) \frac{\partial f_2(\boldsymbol{\theta}_j)}{\partial \boldsymbol{\theta}_j} \Big|_{\boldsymbol{\theta}_j=\bar{\boldsymbol{\theta}}} (\boldsymbol{\theta}_j - \bar{\boldsymbol{\theta}}) + o_p(\|\boldsymbol{\theta}_j - \bar{\boldsymbol{\theta}}\|),$$

where

$$\frac{\partial f_2(\boldsymbol{\theta}_j)}{\partial \boldsymbol{\theta}_j} \Big|_{\boldsymbol{\theta}_j=\bar{\boldsymbol{\theta}}} = \mathbb{E} \left\{ \frac{\partial \hat{\varphi}_f(t_k, \boldsymbol{\theta}_j)}{\partial \boldsymbol{\theta}_j} \Big|_{\boldsymbol{\theta}_j=\bar{\boldsymbol{\theta}}} (y_f(t_k, \bar{\boldsymbol{\theta}}) - \varphi_f^\top(t_k, \bar{\boldsymbol{\theta}})\bar{\boldsymbol{\theta}}) \right\} \quad (4.25a)$$

$$+ \mathbb{E} \left\{ \hat{\varphi}_f(t_k, \bar{\boldsymbol{\theta}}) \left(\frac{\partial y_f(t_k, \boldsymbol{\theta}_j)}{\partial \boldsymbol{\theta}_j} \Big|_{\boldsymbol{\theta}_j=\bar{\boldsymbol{\theta}}} - \bar{\boldsymbol{\theta}}^\top \frac{\partial \varphi_f(t_k, \boldsymbol{\theta}_j)}{\partial \boldsymbol{\theta}_j} \Big|_{\boldsymbol{\theta}_j=\bar{\boldsymbol{\theta}}} \right) \right\}. \quad (4.25b)$$

Let Ψ_1 and Ψ_2 denote the expectations in (4.25a) and (4.25b), respectively. The expression $y_f(t_k, \bar{\boldsymbol{\theta}}) - \varphi_f^\top(t_k, \bar{\boldsymbol{\theta}})\bar{\boldsymbol{\theta}} = y_f(t_k, \boldsymbol{\theta}^*) - \varphi_f^\top(t_k, \boldsymbol{\theta}^*)\boldsymbol{\theta}^* = v(t_k)$ can be substituted into (4.25a), and after computing the vector differentiation of the filtered instrument vector we can express Ψ_1 as $\mathbb{E}\{\mathbf{M}(p)u(t_k)v(t_k)\}$, where

$$\mathbf{M}(p) = \frac{1}{\bar{A}^2(p)} \begin{bmatrix} 2\frac{p^2\bar{B}(p)}{\bar{A}(p)} & \dots & 2\frac{p^{n+1}\bar{B}(p)}{\bar{A}(p)} & -p & \dots & -p^{m+1} \\ \vdots & & \vdots & \vdots & & \vdots \\ 2\frac{p^{n+1}\bar{B}(p)}{\bar{A}(p)} & \dots & 2\frac{p^{2n}\bar{B}(p)}{\bar{A}(p)} & -p^n & \dots & -p^{n+m} \\ -p & \dots & -p^n & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ -p^{m+1} & \dots & -p^{n+m} & 0 & \dots & 0 \end{bmatrix}. \quad (4.26)$$

Similar to the procedure undertaken in Lemma 4.1, every element of Ψ_1 can be expressed as

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \tilde{G}_d(e^{i\omega}) \phi_{uv}(\omega) d\omega,$$

where \tilde{G}_d represents the FOH equivalent of the transfer function of an entry of the matrix in (4.26), and $\phi_{uv}(\omega)$ is the cross spectrum between $u(t_k)$ and $v(t_k)$. Since the input and noise are uncorrelated, $\phi_{uv}(\omega) = 0$ for all $\omega \in [-\pi, \pi]$. Thus, $\Psi_1 = \mathbf{0}$.

Now, after some further vector differentiations, we obtain

$$\frac{\partial y_f(t_k, \boldsymbol{\theta}_j)}{\partial \boldsymbol{\theta}_j} \Big|_{\boldsymbol{\theta}_j=\bar{\boldsymbol{\theta}}} = \left[\frac{-p}{\bar{A}^2(p)} y(t_k), \dots, \frac{-p^n}{\bar{A}^2(p)} y(t_k), 0, \dots, 0 \right],$$

and

$$\bar{\theta}^\top \frac{\partial \varphi_f(t_k, \theta_j)}{\partial \theta_j} \Big|_{\theta_j=\bar{\theta}} = \bar{\theta}^\top \frac{1}{A^2(p)} \begin{bmatrix} p^2 y(t_k) & \dots & p^{n+1} y(t_k) & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ p^{n+1} y(t_k) & \dots & p^{2n} y(t_k) & 0 & \dots & 0 \\ -pu(t_k) & \dots & -p^n u(t_k) & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ -p^{m+1} u(t_k) & \dots & -p^{n+m} u(t_k) & 0 & \dots & 0 \end{bmatrix},$$

which leads to

$$\frac{\partial y_f(t_k, \theta_j)}{\partial \theta_j} \Big|_{\theta_j=\bar{\theta}} - \bar{\theta}^\top \frac{\partial \varphi_f(t_k, \theta_j)}{\partial \theta_j} \Big|_{\theta_j=\bar{\theta}} = \bar{\Delta}^\top(t_k) + \bar{v}_f^\top(t_k),$$

where $\bar{\Delta}(t_k)$ is given in (4.10) and $\bar{v}_f(t_k)$ corresponds to (4.5) being evaluated at the limiting point. Following the lines of Lemma 4.1, we find that $\mathbb{E}\{\hat{\varphi}_f(t_k, \bar{\theta})\bar{v}_f^\top(t_k)\} = \mathbf{0}$. Therefore, we can write the Taylor series expansion as

$$\theta_{j+1} - \bar{\theta} = \mathbf{Q}^{-1} \mathbf{R}(\theta_j - \bar{\theta}) + o_p(\|\theta_j - \bar{\theta}\|),$$

where \mathbf{Q} and \mathbf{R} are defined in (4.8) and (4.9), respectively. Provided that all the eigenvalues of $\mathbf{Q}^{-1} \mathbf{R}$ have magnitude less than 1, θ_{j+1} is locally convergent to $\bar{\theta} = \theta^*$, and this completes the final part of the proof. \square

Corollary 4.1. *Assume that the input sequence is persistently exciting of order no less than $2n$ (instead of $2n+1$) and the model is strictly proper. When the FOH used in Theorem 4.1 is replaced with a ZOH device, Statements 1, 2 and 3 in Theorem 4.1 still hold.*

Proof. The proof follows the same procedure as that of Theorem 4.1. In this case, when the system and model transfer functions are strictly proper, the numerator degree of the discrete-time transfer function is at most $n-1$. Thus, the persistent excitation order in Assumption 4.3 can be relaxed to $2n$ according to the reasoning provided in Remark 4.3 (see Appendix 4.A). \square

Theorem 4.1 and Corollary 4.1 have established the consistency of the SRIVC estimator when the intersample behavior of the input signals in both the filtered regressor and instrument vectors as well as the output signal are assumed to be the same as that of the true system input. Next, Corollary 4.2 examines the effect on the consistency of the SRIVC estimates when there is a misspecification of the intersample behavior of 1) the input in the filtered instrument vector, and 2) the input in the filtered regressor vector, i.e., the model input. In the sequel, the true system input is assumed to have an FOH intersample behavior.

Corollary 4.2. *The SRIVC estimator*

1. remains generically consistent if a misspecification of the intersample behavior (i.e., ZOH instead of FOH) is used for generating the filtered signals in the instrument vector $\hat{\varphi}_f(t_k)$, and
2. is generically not consistent if a misspecification of the intersample behavior is used for filtering the input signal in the regressor vector $\varphi_f(t_k)$.

Proof of Statement 1: Part 1 of Theorem 4.1 still holds since the only change is that the FOH discretizations of the first transfer functions in (4.48) and (4.51) are replaced by their ZOH equivalents, which does not affect the way the analyticity of Φ is shown. Part 2 in Theorem 4.1 also remains unchanged since the misspecification of the intersample behavior of the input in the filtered instrument does not affect the formulation of the equation error in (4.21). For the same reason, Part 3 of Theorem 4.1 also remains unchanged. Thus, when the intersample behavior for the input signal in the filtered instrument vector $\hat{\varphi}_f(t_k)$ is misspecified, the SRIVC estimator remains generically consistent. \square

Proof of Statement 2: Let the input in the filtered regressor vector, denoted by $\tilde{u}(t_k)$, have a different intersample behavior than that of the true system input. Part 1 of Theorem 4.1 remains unchanged; however, we will show that Part 2 of Theorem 4.1 has been affected. Consider the GEE at the limiting solution $\bar{\theta}$:

$$\begin{aligned}\varepsilon(t_k, \bar{\theta}) &= y(t_k) - \frac{\bar{B}(p)}{\bar{A}(p)}\tilde{u}(t_k) \\ &= \frac{B^*(p)}{A^*(p)}u(t_k) + v(t_k) - \left(\frac{\bar{B}(p)}{\bar{A}(p)}u(t_k) + \varepsilon_u(t_k) \right) \\ &= \frac{1}{\bar{A}(p)A^*(p)}(\bar{A}(p)B^*(p) - \bar{B}(p)A^*(p))u(t_k) + v(t_k) - \varepsilon_u(t_k).\end{aligned}\quad (4.27)$$

We have introduced an input-dependent term $\varepsilon_u(t_k)$ into the modeled output in (4.27) to account for the interpolation error. At the limiting point, (4.20) holds. Substituting (4.16) for $\hat{\varphi}_f(t_k, \bar{\theta})$ and (4.27) for $\varepsilon(t_k, \bar{\theta})$ into (4.20), by using the same definition of \mathbf{h} from (4.23) we obtain

$$\mathbb{E}\{\hat{\varphi}_f(t_k, \bar{\theta})\varepsilon(t_k, \bar{\theta})\} = \mathbf{S}(-\bar{B}, \bar{A})\bar{\Phi}\mathbf{h} - \mathbf{S}(-\bar{B}, \bar{A})\mathbb{E}\left\{\frac{1}{\bar{A}^2(p)}\mathbf{u}_{n+m}(t_k)\varepsilon_u(t_k)\right\} = \mathbf{0}. \quad (4.28)$$

Denote the expectation on the right hand side of (4.28) as $\tilde{\Psi}_u$. Since the error $\varepsilon_u(t_k)$ is input dependent, $\tilde{\Psi}_u$ does not go to zero in general. Note that the matrix $\bar{\Phi}$ is generically non-singular by Part 1 of Theorem 4.1. Therefore, we can obtain \mathbf{H} by solving $\mathbf{H} = \bar{\Phi}^{-1}\tilde{\Psi}_u$. Now, define $H(p) := h_{n+m}p^{n+m} + h_{n+m-1}p^{n+m-1} + \dots + h_0$. We find that $\bar{A}(p)B^*(p) - \bar{B}(p)A^*(p) = H(p)$, which implies

$$\frac{\bar{B}(p)}{\bar{A}(p)} = \frac{B^*(p)}{A^*(p)} - \frac{H(p)}{\bar{A}(p)A^*(p)}. \quad (4.29)$$

The expression in (4.29) shows that the true parameters are no longer the limiting solution of the estimator, that is, $\bar{\theta} \neq \theta^*$.

It is implied by (4.20) that the input is uncorrelated with the GEE evaluated at the converging point. Thus, θ_{j+1} converges to the new limiting point $\bar{\theta}$ in (4.29) as $N \rightarrow \infty$ for every fixed $j \geq 1$. Since $\bar{\theta} \neq \theta^*$, we conclude that θ_{j+1} does not converge to the true parameters θ^* if the intersample behavior of the input signal is misspecified in the filtered regressor vector $\varphi_f(t_k)$. Hence, the SRIVC estimator is generically not consistent when the true system input cannot be interpolated exactly. \square

For discretization purposes, the true system input is assumed to have an FOH intersample behavior for the analysis above. We note that Statement 2 of Corollary 4.2 holds for any input that cannot be interpolated exactly, as stated in the following remark.

Remark 4.1. When the input to the true system $u(t)$ is a continuous function of time that cannot be interpolated exactly by a ZOH or FOH device, the SRIVC estimator is not consistent. This follows from the analysis in Corollary 4.2.

In situations where the input cannot be interpolated exactly by a ZOH or FOH device, $\varepsilon_u(t_k)$ will be non-zero and the asymptotic bias on the estimates is captured by $[\bar{A}(p)A^*(p)]^{-1}H(p)$ in (4.29). Since the polynomial $H(p)$ is proportional to the interpolation error $\varepsilon_u(t_k)$, which will decrease if the signals are sampled faster, it is implied that the bias of the estimates will generally decrease with the sampling period. This result falls in line with practical experience [63].

Next, we examine the effect of the intersample behavior of the sampled output on the consistency of the SRIVC estimator in the following remark.

Remark 4.2. At each SRIVC iteration j , the GEE $\varepsilon(t_k, \theta_j)$ can be expressed as

$$\varepsilon(t_k, \theta_j) = A_j(p) \frac{1}{A_{j-1}(p)} y(t_k) - B_j(p) \frac{1}{A_{j-1}(p)} u(t_k).$$

Upon convergence, the expression above becomes

$$\varepsilon(t_k, \bar{\theta}) = \bar{A}(p) \frac{1}{\bar{A}(p)} y(t_k) - \bar{B}(p) \frac{1}{\bar{A}(p)} u(t_k).$$

The SRIVC algorithm computes the prefiltered signals $(p^l/\bar{A}(p)) y(t_k) = y_f^{(l)}(t_k)$ for $l = 0, \dots, n$ (see, e.g., (2.30)), leading to

$$\bar{A}(p) \frac{1}{\bar{A}(p)} y(t_k) = a_n y_f^{(n)}(t_k) + a_{n-1} y_f^{(n-1)}(t_k) + \dots + y_f(t_k) = y(t_k).$$

The intersample behavior of the measured output does not affect the GEE at the converging point as the effect of filtering the measured output cancels to the polynomial $\bar{A}(p)$ appearing in both the numerator and denominator of the filter.

Thus, provided that (4.7) is satisfied, the intersample behavior assumption of the measured output does not affect the consistency of the SRIVC estimator. Note that the condition in (4.7) implicitly depends on the intersample behavior specification of the output through the term $\Delta(t_k)$, which implies that the SRIVC estimator may fail to converge if the intersample behavior assumption of the measured output differs greatly from that of the noiseless output.

In summary, we have shown that

- the SRIVC estimator is generically consistent under mild assumptions on the model structure, persistence of excitation of the input, sampling period, and intersample behavior assumption of the measured output;
- the SRIVC estimator is generically consistent even if an incorrect intersample assumption is used for constructing the filtered instrument vector, but is generically inconsistent if the filtered regressor vector is constructed with an incorrect intersample assumption on the input; and
- the intersample behavior of the measured output does not play a role in the consistency of the SRIVC estimator provided it does not depart much from that of the noiseless output.

4.4 Efficiency analysis

In this section, we derive the asymptotic Cramér-Rao lower bound (CRLB) for the continuous-time output error model structure and provide an analysis of the statistical efficiency of the SRIVC estimator based on sampled data. For this, we consider the model structure in (4.2) and Assumptions 4.1 to 4.6, but we replace Assumption 4.5 by a stronger one:

Assumption 4.7. The model order matches the system order, i.e., $n = n^*$ and $m = m^*$.

4.4.1 Asymptotic Cramér-Rao lower bound

The goal of this subsection is to develop an explicit expression for the asymptotic CRLB for the continuous-time output error model structure. It is shown that the derived expression is independent of the intersample behavior of the noise-free system output and hence it only depends on the intersample behavior of the system input.

Theorem 4.2 (Asymptotic Cramér-Rao Lower Bound). *Consider the prediction error*

$$\varepsilon(t_k, \boldsymbol{\theta}) = y(t_k) - \left\{ \frac{B(p)}{A(p)} u(t) \right\}_{t=t_k} \quad (4.30)$$

for an unknown parameter vector $\boldsymbol{\theta}$ formed using the model coefficients of $A(p)$ and $B(p)$. Assume the output observations come from an output error model structure

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

where $v(t_k)$ is an i.i.d. Gaussian noise with variance σ^2 , and $B^*(p)$ and $A^*(p)$ are the system polynomials. Then, under Assumptions 4.1, 4.2, and 4.7, the asymptotic Cramér-Rao lower bound is given by

$$\mathbf{P}_{\text{CR}} = \sigma^2 \mathbb{E} \{ \boldsymbol{\psi}(t_k, \boldsymbol{\theta}^*) \boldsymbol{\psi}^\top(t_k, \boldsymbol{\theta}^*) \}^{-1}, \quad (4.31)$$

where

$$\boldsymbol{\psi}(t_k, \boldsymbol{\theta}^*) = \left[-\frac{pB^*(p)}{A^{*2}(p)} u(t), \dots, -\frac{p^{n^*} B^*(p)}{A^{*2}(p)} u(t), \frac{1}{A^*(p)} u(t), \dots, \frac{p^{m^*}}{A^*(p)} u(t) \right]_{t=t_k}^\top. \quad (4.32)$$

Proof. Denote the vector of N samples of the measured output as \mathbf{y}^N . According to (4.30), the joint PDF of \mathbf{y}^N , and hence the likelihood of \mathbf{y}^N when viewed as a function of unknown parameters $\boldsymbol{\theta}$, is [138, Lemma 5.1]

$$p(\mathbf{y}^N; \boldsymbol{\theta}) = \prod_{k=1}^N p(\varepsilon(t_k, \boldsymbol{\theta})),$$

where $p(\varepsilon(t_k, \boldsymbol{\theta}))$ is the PDF of the prediction error $\varepsilon(t_k, \boldsymbol{\theta})$. The log-likelihood of the measurement vector \mathbf{y}^N is thus (recall (2.20)),

$$\begin{aligned} \ell(\boldsymbol{\theta}) &:= \log p(\mathbf{y}^N; \boldsymbol{\theta}) \\ &= \sum_{k=1}^N \log p(\varepsilon(t_k, \boldsymbol{\theta})) \\ &= -\frac{N}{2} \log(2\pi) - \frac{N}{2} \log(\sigma^2) - \frac{1}{2\sigma^2} \sum_{k=1}^N \left(y(t_k) - \left\{ \frac{B(p)}{A(p)} u(t) \right\}_{t=t_k} \right)^2. \end{aligned}$$

According to Lemma 3.2 of [156], the procedures of linearization and discretization commute. We can therefore differentiate $\ell(\boldsymbol{\theta})$ with respect to the continuous-time parameters and then discretize the transfer functions according to the intersample behavior of the input signal and the sampling period. Now, differentiating the log-likelihood function with respect to the denominator coefficients of the model and then evaluating at the system parameters gives

$$\begin{aligned} \frac{\partial \ell(\boldsymbol{\theta})}{\partial a_l} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*} &= -\frac{1}{\sigma^2} \sum_{k=1}^N \left(y(t_k) - \left\{ \frac{B^*(p)}{A^*(p)} u(t) \right\}_{t=t_k} \right) \left\{ \frac{p^l B^*(p)}{A^{*2}(p)} u(t) \right\}_{t=t_k} \\ &= -\frac{1}{\sigma^2} \sum_{k=1}^N v(t_k) \left\{ \frac{p^l B^*(p)}{A^{*2}(p)} u(t) \right\}_{t=t_k}, \end{aligned}$$

where $l = 1, \dots, n^*$. Similarly, differentiating the log-likelihood function with respect to the numerator coefficients yields

$$\begin{aligned} \frac{\partial \ell(\boldsymbol{\theta})}{\partial b_l} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*} &= -\frac{1}{\sigma^2} \sum_{k=1}^N \left(y(t_k) - \left\{ \frac{B^*(p)}{A^*(p)} u(t) \right\}_{t=t_k} \right) \left(-\left\{ \frac{p^l}{A^*(p)} u(t) \right\}_{t=t_k} \right) \\ &= \frac{1}{\sigma^2} \sum_{k=1}^N v(t_k) \left\{ \frac{p^l}{A^*(p)} u(t) \right\}_{t=t_k}, \end{aligned}$$

where $l = 0, \dots, m^*$. Hence,

$$\frac{\partial \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*} = \frac{1}{\sigma^2} \sum_{k=1}^N \psi(t_k, \boldsymbol{\theta}^*) v(t_k),$$

where $\psi(t_k, \boldsymbol{\theta}^*)$ is given by (4.32). Note that the input in (4.32) is treated as a continuous-time signal and is filtered by continuous-time transfer functions prior to sampling. Thus, it does not depend on the intersample behavior of the noise-free system output. The Fisher information matrix is then given by

$$\begin{aligned} \mathbf{I}_F &:= \mathbb{E} \left\{ \left(\frac{\partial \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*} \right) \left(\frac{\partial \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*} \right)^\top \right\} \\ &= \frac{1}{\sigma^4} \sum_{k=1}^N \sum_{s=1}^N \mathbb{E} \{ v(t_k) \psi(t_k, \boldsymbol{\theta}^*) v(t_s) \psi^\top(t_s, \boldsymbol{\theta}^*) \} \\ &= \frac{1}{\sigma^2} \sum_{k=1}^N \mathbb{E} \{ \psi(t_k, \boldsymbol{\theta}^*) \psi^\top(t_k, \boldsymbol{\theta}^*) \}, \end{aligned} \quad (4.33)$$

where we have used the fact that the sequences $\{u(t_k)\}$ and $\{v(t_k)\}$ are independent by Assumption 4.2, and that $\{v(t_k)\}$ is a Gaussian i.i.d. random process of variance σ^2 . Now, since $\psi(t_k, \boldsymbol{\theta}^*)$ is composed of stationary random processes, its ensemble average is equal to its time average as sample size approaches infinity [211, Appendix B.1], i.e.,

$$\mathbb{E} \{ \psi(t_k, \boldsymbol{\theta}^*) \psi^\top(t_k, \boldsymbol{\theta}^*) \} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{l=1}^N \psi(t_l, \boldsymbol{\theta}^*) \psi^\top(t_l, \boldsymbol{\theta}^*) \quad \text{a.s.} \quad (4.34)$$

for every $k = 1, \dots, N$. Then, thanks to (4.33) and (4.34), the asymptotic Cramér-Rao lower bound, given by the inverse of the Fisher information matrix per sample, can be expressed as

$$\mathbf{P}_{\text{CR}} = \left[\lim_{N \rightarrow \infty} \frac{1}{N} \mathbf{I}_F \right]^{-1} = \sigma^2 \mathbb{E} \{ \psi(t_k, \boldsymbol{\theta}^*) \psi^\top(t_k, \boldsymbol{\theta}^*) \}^{-1},$$

where $\psi(t_k, \boldsymbol{\theta}^*)$ is given by (4.32). \square

The CRLB obtained in Theorem 4.2 suggests that for an estimator to be asymptotically optimal in terms of its covariance, it should consider the intersample behavior of the input and its covariance should not depend on the intersample behavior of the output as the sample size tends to infinity. The question that remains is whether the SRIVC estimator satisfies these requisites for optimality. The answer is given in the next subsection.

4.4.2 Asymptotic distribution of the SRIVC estimates

Here we obtain the asymptotic distribution of the SRIVC estimates for the output error model structure. For derivation purposes, we introduce a theoretical SRIVC estimator, which is described in Definition 4.1.

Definition 4.1 (Theoretical SRIVC estimator). Define a theoretical SRIVC estimator given by the same procedure as in Algorithm 2.1 but with the filtered regressor vector and the filtered output given by

$$\begin{aligned} \dot{\varphi}_f(t_k) := & \left[- \left\{ \frac{p}{A_j(p)} \frac{B^*(p)}{A^*(p)} u(t) \right\}_{t=t_k} + \frac{p}{A_j(p)} v(t_k), \dots, \right. \\ & \left. - \left\{ \frac{p^n}{A_j(p)} \frac{B^*(p)}{A^*(p)} u(t) \right\}_{t=t_k} + \frac{p^n}{A_j(p)} v(t_k), \left\{ \frac{1}{A_j(p)} u(t) \right\}_{t=t_k}, \dots, \left\{ \frac{p^m}{A_j(p)} u(t) \right\}_{t=t_k} \right]^\top \end{aligned} \quad (4.35)$$

and

$$\dot{y}_f(t_k) := \left\{ \frac{1}{A_j(p)} \frac{B^*(p)}{A^*(p)} u(t) \right\}_{t=t_k} + \frac{1}{A_j(p)} v(t_k), \quad (4.36)$$

respectively. Note that (4.35) and (4.36) implicitly assume that the measured output is a continuous-time signal. Thus, this estimator cannot be implemented in practice.

We are now ready to present the main contribution of this section.

Theorem 4.3 (Asymptotic distribution of the SRIVC estimates). *Consider the SRIVC estimator given in Algorithm 2.1 under the output error model structure, where the output measurement noise sequence $\{v(t_k)\}$ in (4.2) is an i.i.d. Gaussian white noise of variance σ^2 , and the continuous-time input has a ZOH or FOH intersample behavior that is correctly specified in the SRIVC implementation. Suppose Assumptions 4.1–4.4, 4.6–4.7 hold, and assume that the estimator is consistent. Let $\bar{\theta}^N$ be the converging point (in iterations) of the SRIVC estimator for a fixed sample size N . Then, the SRIVC estimate is asymptotically Gaussian-distributed, i.e.,*

$$\sqrt{N}(\bar{\theta}^N - \theta^*) \xrightarrow{\text{dist.}} \mathcal{N}(\mathbf{0}, \mathbf{P}_{\text{SRIVC}}),$$

where the asymptotic covariance matrix is

$$\mathbf{P}_{\text{SRIVC}} = \sigma^2 \mathbb{E}\{\tilde{\varphi}_f(t_k, \theta^*) \tilde{\varphi}_f^\top(t_k, \theta^*)\}^{-1}, \quad (4.37)$$

with $\tilde{\varphi}_f(t_k, \boldsymbol{\theta}^*)$ given by

$$\tilde{\varphi}_f(t_k, \boldsymbol{\theta}^*) = \left[-\frac{pB^*(p)}{A^{*2}(p)}u(t), \dots, -\frac{p^{n^*}B^*(p)}{A^{*2}(p)}u(t), \frac{1}{A^*(p)}u(t), \dots, \frac{p^{m^*}}{A^*(p)}u(t) \right]_{t=t_k}^\top. \quad (4.38)$$

Proof. From (2.35), we know that

$$\begin{aligned} \bar{\boldsymbol{\theta}}^N - \boldsymbol{\theta}^* &= \left[\sum_{k=1}^N \hat{\varphi}_f(t_k, \bar{\boldsymbol{\theta}}^N) \varphi_f^\top(t_k, \bar{\boldsymbol{\theta}}^N) \right]^{-1} \left[\sum_{k=1}^N \hat{\varphi}_f(t_k, \bar{\boldsymbol{\theta}}^N) y_f(t_k, \bar{\boldsymbol{\theta}}^N) \right] - \boldsymbol{\theta}^* \\ &= \left[\sum_{k=1}^N \hat{\varphi}_f(t_k, \bar{\boldsymbol{\theta}}^N) \varphi_f^\top(t_k, \bar{\boldsymbol{\theta}}^N) \right]^{-1} \left[\sum_{k=1}^N \hat{\varphi}_f(t_k, \bar{\boldsymbol{\theta}}^N) (y_f(t_k, \bar{\boldsymbol{\theta}}^N) - \varphi_f^\top(t_k, \bar{\boldsymbol{\theta}}^N) \boldsymbol{\theta}^*) \right]. \end{aligned} \quad (4.39)$$

It has been stated in Remark 4.2 that the intersample behavior of $y(t_k)$ assumed in order to perform the filtering operations in the regressor vector does not affect the SRIVC estimates at the converging point of the algorithm. Furthermore, it is proven in Lemma 4.5 (see Appendix 4.B) that the standard SRIVC estimator is equivalent to the theoretical estimator given in Definition 4.1 at the converging point for a large sample size. Hence, according to Lemma 4.5, at $\bar{\boldsymbol{\theta}}^N$, the filtered regressor and the filtered output in (4.39) can be replaced by (4.35) and (4.36) evaluated at $\bar{\boldsymbol{\theta}}^N$ respectively. Note that the input $u(t_k)$ must have the same intersample behavior as the system input in order to obtain a consistent estimate according to Theorem 4.1. The second half of (4.39) can then be expressed as

$$\begin{aligned} y_f(t_k, \bar{\boldsymbol{\theta}}^N) - \varphi_f^\top(t_k, \bar{\boldsymbol{\theta}}^N) \boldsymbol{\theta}^* &= \left\{ \frac{A^*(p)}{\bar{A}_N(p)} \frac{B^*(p)}{A^*(p)} u(t) \right\}_{t=t_k} + \frac{A^*(p)}{\bar{A}_N(p)} v(t_k) - \left\{ \frac{B^*(p)}{\bar{A}_N(p)} u(t) \right\}_{t=t_k} \\ &= \frac{A^*(p)}{\bar{A}_N(p)} v(t_k), \end{aligned}$$

where $\bar{A}_N(p)$ is the denominator polynomial corresponding to $\bar{\boldsymbol{\theta}}^N$. Therefore, (4.39) simplifies to

$$\bar{\boldsymbol{\theta}}^N - \boldsymbol{\theta}^* = \left[\frac{1}{N} \sum_{k=1}^N \hat{\varphi}_f(t_k, \bar{\boldsymbol{\theta}}^N) \varphi_f^\top(t_k, \bar{\boldsymbol{\theta}}^N) \right]^{-1} \left[\frac{1}{N} \sum_{k=1}^N \hat{\varphi}_f(t_k, \bar{\boldsymbol{\theta}}^N) \frac{A^*(p)}{\bar{A}_N(p)} v(t_k) \right].$$

The next step consists in analyzing the behavior of the expression above for large N , similarly to what has been done for obtaining the asymptotic distribution of discrete-time bootstrapped instrumental variable estimators in [210, Ch. 5]. The

first-order Taylor series expansion of $\sqrt{N}(\bar{\boldsymbol{\theta}}^N - \boldsymbol{\theta}^*)$ can be written as

$$\begin{aligned} & \sqrt{N}(\bar{\boldsymbol{\theta}}^N - \boldsymbol{\theta}^*) \\ &= \left[\frac{1}{N} \sum_{k=1}^N \hat{\varphi}_f(t_k, \boldsymbol{\theta}^*) \dot{\hat{\varphi}}_f^\top(t_k, \boldsymbol{\theta}^*) \right]^{-1} \left[\frac{1}{\sqrt{N}} \sum_{k=1}^N \hat{\varphi}_f(t_k, \boldsymbol{\theta}^*) v(t_k) \right] \end{aligned} \quad (4.40a)$$

$$+ \frac{\partial}{\partial \bar{\boldsymbol{\theta}}^N} \left\{ \left[\frac{1}{N} \sum_{k=1}^N \hat{\varphi}_f(t_k, \bar{\boldsymbol{\theta}}^N) \dot{\hat{\varphi}}_f^\top(t_k, \bar{\boldsymbol{\theta}}^N) \right]^{-1} \right\} \Big|_{\bar{\boldsymbol{\theta}}^N = \boldsymbol{\theta}^*} \left[\frac{1}{\sqrt{N}} \sum_{k=1}^N \hat{\varphi}_f(t_k, \boldsymbol{\theta}^*) v(t_k) \right] (\bar{\boldsymbol{\theta}}^N - \boldsymbol{\theta}^*) \quad (4.40b)$$

$$+ \left[\frac{1}{N} \sum_{k=1}^N \hat{\varphi}_f(t_k, \boldsymbol{\theta}^*) \dot{\hat{\varphi}}_f^\top(t_k, \boldsymbol{\theta}^*) \right]^{-1} \left(\frac{1}{\sqrt{N}} \sum_{k=1}^N \frac{\partial \hat{\varphi}_f(t_k, \bar{\boldsymbol{\theta}}^N)}{\partial \bar{\boldsymbol{\theta}}^N} \Big|_{\bar{\boldsymbol{\theta}}^N = \boldsymbol{\theta}^*} v(t_k) \right) \quad (4.40c)$$

$$+ \frac{1}{\sqrt{N}} \sum_{k=1}^N \hat{\varphi}_f(t_k, \boldsymbol{\theta}^*) \frac{\partial}{\partial \bar{\boldsymbol{\theta}}^N} \left(\frac{1}{\bar{A}_N(p)} \right) \Big|_{\bar{\boldsymbol{\theta}}^N = \boldsymbol{\theta}^*} A^*(p) v(t_k) \Big) (\bar{\boldsymbol{\theta}}^N - \boldsymbol{\theta}^*) + o_p(\sqrt{N} \|\bar{\boldsymbol{\theta}}^N - \boldsymbol{\theta}^*\|). \quad (4.40d)$$

This expression can be rewritten by moving (4.40b)-(4.40d) to the left-hand side and factoring by $(\bar{\boldsymbol{\theta}}^N - \boldsymbol{\theta}^*)$, resulting in

$$\sqrt{N} \left(\mathbf{I} - \mathbf{T} \left[\frac{1}{N} \sum_{k=1}^N \hat{\varphi}_f(t_k, \boldsymbol{\theta}^*) v(t_k) \right] - \mathbf{R} \left[\frac{1}{N} \sum_{k=1}^N \frac{\partial \hat{\varphi}_f(t_k, \bar{\boldsymbol{\theta}}^N)}{\partial \bar{\boldsymbol{\theta}}^N} \Big|_{\bar{\boldsymbol{\theta}}^N = \boldsymbol{\theta}^*} v(t_k) \right] \right) \quad (4.41a)$$

$$- \mathbf{R} \left[\frac{1}{N} \sum_{k=1}^N \hat{\varphi}_f(t_k, \boldsymbol{\theta}^*) \frac{\partial}{\partial \bar{\boldsymbol{\theta}}^N} \left(\frac{1}{\bar{A}_N(p)} \right) \Big|_{\bar{\boldsymbol{\theta}}^N = \boldsymbol{\theta}^*} A^*(p) v(t_k) \right] (\bar{\boldsymbol{\theta}}^N - \boldsymbol{\theta}^*) \quad (4.41b)$$

$$= \mathbf{R} \left[\frac{1}{\sqrt{N}} \sum_{k=1}^N \hat{\varphi}_f(t_k, \boldsymbol{\theta}^*) v(t_k) \right] + o_p(\sqrt{N} \|\bar{\boldsymbol{\theta}}^N - \boldsymbol{\theta}^*\|), \quad (4.41c)$$

where \mathbf{I} is the identity matrix, \mathbf{T} is the gradient of the matrix inverse in (4.40b) which is a tensor of order 3, and \mathbf{R} is the matrix inverse in (4.40c). Note that the tensor \mathbf{T} premultiplies a vector, resulting in a matrix of size $(n+m+1) \times (n+m+1)$.

Next, we will examine each term in the parenthesis in (4.41a)-(4.41b) for a large sample size. Since the estimator is assumed to be consistent, the matrix \mathbf{R} in (4.41a), (4.41b) and (4.41c) is non-singular for N sufficiently large. Now, thanks to the fact that the input and disturbance are stationary, the ergodic lemma in [205, Lemma 3.1] holds for the second term of (4.41a). Thus, this term can be written as

$$\mathbf{T} \left[\frac{1}{N} \sum_{k=1}^N \hat{\varphi}_f(t_k, \boldsymbol{\theta}^*) v(t_k) \right] = \mathbf{T} \mathbb{E} \{ \hat{\varphi}_f(t_k, \boldsymbol{\theta}^*) v(t_k) \} + o_p(1)$$

for a large sample size. The filtered instrument vector $\hat{\varphi}_f(t_k, \boldsymbol{\theta}^*)$ consists of filtered inputs that are independent of the disturbance $v(t_k)$ according to Assumption 4.2.

Thus, by using the same method as in the proof of Theorem 4.1, we can show that $\mathbb{E}\{\hat{\varphi}_f(t_k, \boldsymbol{\theta}^*)v(t_k)\} = \mathbf{0}$. So, the second term in the parenthesis in (4.41a) is $o_p(1)$.

In addition, when performing vector differentiation, we obtain

$$\left. \frac{\partial \hat{\varphi}_f(t_k, \bar{\boldsymbol{\theta}}^N)}{\partial \bar{\boldsymbol{\theta}}^N} \right|_{\bar{\boldsymbol{\theta}}^N=\boldsymbol{\theta}^*} = \mathbf{M}^*(p)u(t_k),$$

where $\mathbf{M}^*(p)$ is the matrix (4.26) evaluated at the true parameters. Therefore, for a large sample size and by the same reasoning, the third term in the parenthesis in (4.41a) can be written as

$$\mathbf{R} \left[\frac{1}{N} \sum_{k=1}^N \left. \frac{\partial \hat{\varphi}_f(t_k, \bar{\boldsymbol{\theta}}^N)}{\partial \bar{\boldsymbol{\theta}}^N} \right|_{\bar{\boldsymbol{\theta}}^N=\boldsymbol{\theta}^*} v(t_k) \right] = \mathbf{R}\mathbb{E}\{\mathbf{M}^*(p)u(t_k)v(t_k)\} + o_p(1) = o_p(1).$$

Furthermore,

$$\left. \frac{\partial}{\partial \bar{\boldsymbol{\theta}}^N} \left(\frac{1}{A_N(p)} \right) \right|_{\bar{\boldsymbol{\theta}}^N=\boldsymbol{\theta}^*} = \frac{1}{A^{*2}(p)} [-p, \dots, -p^{n^*}, 0, \dots, 0].$$

Then, for a large sample size, the term in the parenthesis in (4.41b) is given by

$$\begin{aligned} & \mathbf{R} \left[\frac{1}{N} \sum_{k=1}^N \hat{\varphi}(t_k, \boldsymbol{\theta}^*) \frac{\partial}{\partial \bar{\boldsymbol{\theta}}^N} \left(\frac{1}{A_N(p)} \right) \Big|_{\bar{\boldsymbol{\theta}}^N=\boldsymbol{\theta}^*} A^*(p)v(t_k) \right] \\ &= \mathbf{R}\mathbb{E} \left\{ \hat{\varphi}(t_k, \boldsymbol{\theta}^*) \frac{1}{A^*(p)} [-p, \dots, -p^{n^*}, 0, \dots, 0] v(t_k) \right\} + o_p(1) \\ &= o_p(1). \end{aligned}$$

Therefore, (4.41) can be equivalently written as

$$\sqrt{N}(\bar{\boldsymbol{\theta}}^N - \boldsymbol{\theta}^*) = \left[\frac{1}{N} \sum_{k=1}^N \hat{\varphi}_f(t_k, \boldsymbol{\theta}^*) \hat{\varphi}_f^\top(t_k, \boldsymbol{\theta}^*) \right]^{-1} \left[\frac{1}{\sqrt{N}} \sum_{k=1}^N \hat{\varphi}_f(t_k, \boldsymbol{\theta}^*) v(t_k) \right] (1 + o_p(1)).$$

According to [205], as the sample size approaches infinity, we have

$$\frac{1}{N} \sum_{k=1}^N \hat{\varphi}_f(t_k) \hat{\varphi}_f^\top(t_k, \boldsymbol{\theta}^*) \xrightarrow{w.p.1} \mathbb{E}\{\hat{\varphi}_f(t_k, \boldsymbol{\theta}^*) \hat{\varphi}_f^\top(t_k, \boldsymbol{\theta}^*)\}.$$

Hence, by the continuous-mapping theorem [230, Thm. 2.3],

$$\left[\frac{1}{N} \sum_{k=1}^N \hat{\varphi}_f(t_k, \boldsymbol{\theta}^*) \hat{\varphi}_f^\top(t_k, \boldsymbol{\theta}^*) \right]^{-1} \xrightarrow{w.p.1} \mathbb{E}\{\hat{\varphi}_f(t_k, \boldsymbol{\theta}^*) \hat{\varphi}_f^\top(t_k, \boldsymbol{\theta}^*)\}^{-1}.$$

Let $\tilde{\varphi}_f(t_k, \boldsymbol{\theta}^*)$ in (4.38) be the noise-free version of $\hat{\varphi}_f(t_k, \boldsymbol{\theta}_j)$ given in (4.35) evaluated at $\boldsymbol{\theta}^*$. Then, for large N ,

$$\begin{aligned}\sqrt{N}(\bar{\boldsymbol{\theta}}^N - \boldsymbol{\theta}^*) &= \mathbb{E}\{\hat{\varphi}_f(t_k, \boldsymbol{\theta}^*)\hat{\varphi}_f^\top(t_k, \boldsymbol{\theta}^*)\}^{-1} \left[\frac{1}{\sqrt{N}} \sum_{k=1}^N \hat{\varphi}_f(t_k, \boldsymbol{\theta}^*)v(t_k) \right] + o_p(1) \\ &= \mathbb{E}\{\tilde{\varphi}_f(t_k, \boldsymbol{\theta}^*)\tilde{\varphi}_f^\top(t_k, \boldsymbol{\theta}^*)\}^{-1} \left[\frac{1}{\sqrt{N}} \sum_{k=1}^N \tilde{\varphi}_f(t_k, \boldsymbol{\theta}^*)v(t_k) \right] + o_p(1).\end{aligned}$$

Since both $\hat{\varphi}_f(t_k)$ and $v(t_k)$ are stationary and independent, by Lemma A4.1 of [210],

$$\frac{1}{\sqrt{N}} \sum_{k=1}^N \hat{\varphi}_f(t_k, \boldsymbol{\theta}^*)v(t_k) \xrightarrow{\text{dist.}} \mathcal{N}(\mathbf{0}, \mathbf{P}),$$

where

$$\begin{aligned}\mathbf{P} &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N \sum_{s=1}^N \mathbb{E}\{[\hat{\varphi}_f(t_k, \boldsymbol{\theta}^*)v(t_k)][\hat{\varphi}_f(t_s, \boldsymbol{\theta}^*)v(t_s)]^\top\} \\ &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N \sum_{s=1}^N \mathbb{E}\{\hat{\varphi}_f(t_k, \boldsymbol{\theta}^*)\mathbb{E}\{v(t_k)v(t_s)\}\hat{\varphi}_f^\top(t_s, \boldsymbol{\theta}^*)\} \\ &= \lim_{N \rightarrow \infty} \frac{\sigma^2}{N} \sum_{k=1}^N \mathbb{E}\{\hat{\varphi}_f(t_k, \boldsymbol{\theta}^*)\hat{\varphi}_f^\top(t_k, \boldsymbol{\theta}^*)\} \\ &= \sigma^2 \mathbb{E}\{\hat{\varphi}_f(t_k, \boldsymbol{\theta}^*)\hat{\varphi}_f^\top(t_k, \boldsymbol{\theta}^*)\}. \tag{4.42}\end{aligned}$$

Now, by Lemma A4.2 of [210] and its corollary,

$$\sqrt{N}(\bar{\boldsymbol{\theta}}^N - \boldsymbol{\theta}^*) \xrightarrow{\text{dist.}} \mathcal{N}(\mathbf{0}, \mathbf{P}_{\text{SRIVC}}),$$

where

$$\mathbf{P}_{\text{SRIVC}} = \mathbb{E}\{\hat{\varphi}_f(t_k, \boldsymbol{\theta}^*)\hat{\varphi}_f^\top(t_k, \boldsymbol{\theta}^*)\}^{-1} \mathbf{P} \mathbb{E}\{\tilde{\varphi}_f(t_k, \boldsymbol{\theta}^*)\tilde{\varphi}_f^\top(t_k, \boldsymbol{\theta}^*)\}^{-1},$$

with \mathbf{P} given in (4.42). Substituting (4.42) into (4.43), we can express the asymptotic covariance matrix as

$$\begin{aligned}\mathbf{P}_{\text{SRIVC}} &= \sigma^2 \mathbb{E}\{\hat{\varphi}_f(t_k, \boldsymbol{\theta}^*)\hat{\varphi}_f^\top(t_k, \boldsymbol{\theta}^*)\}^{-1} \mathbb{E}\{\hat{\varphi}_f(t_k, \boldsymbol{\theta}^*)\hat{\varphi}_f^\top(t_k, \boldsymbol{\theta}^*)\} \mathbb{E}\{\tilde{\varphi}_f(t_k, \boldsymbol{\theta}^*)\tilde{\varphi}_f^\top(t_k, \boldsymbol{\theta}^*)\}^{-1}, \tag{4.43}\end{aligned}$$

where $\tilde{\varphi}_f(t_k, \boldsymbol{\theta}^*)$ is given in (4.38).

Note that since the intersample behavior of the input signal in the filtered instrument vector, in the form of (2.33), is chosen to be exactly the same as that of the true continuous-time input signal, we have $\hat{\varphi}_f(t_k, \boldsymbol{\theta}^*) = \tilde{\varphi}_f(t_k, \boldsymbol{\theta}^*)$. Then, (4.43) simplifies to (4.37), concluding the proof. \square

From Theorems 4.2 and 4.3 we note that (4.38) is equal to (4.32) when the intersample behavior of the model input matches that of the system input. Hence, the asymptotic covariance of the SRIVC estimates in (4.37) coincides with the asymptotic CRLB in (4.31). We therefore conclude that the SRIVC estimator is asymptotically efficient under the output error model structure.

Corollary 4.3. *The SRIVC estimator is not asymptotically efficient if the intersample behavior of the input in the filtered instrument vector in (2.33) does not match that of the system input under the output error model structure.*

Proof. If the intersample behavior of the input in (2.33) does not match that of the system input in the SRIVC algorithm, then we claim that there does not exist a constant and nonsingular matrix \mathbf{M} such that $\tilde{\varphi}_f(t_k, \boldsymbol{\theta}^*) = \mathbf{M}\hat{\varphi}_f(t_k, \boldsymbol{\theta}^*)$ with probability 1. Indeed, since multiplying the filtered instrument vector by a constant and nonsingular matrix does not affect the SRIVC computations (see, e.g. (2.35) when $\hat{\varphi}_f(t_k, \boldsymbol{\theta}^*)$ is replaced by $\mathbf{M}\hat{\varphi}_f(t_k, \boldsymbol{\theta}^*)$), the existence of such matrix would imply that the SRIVC estimator is in fact using $\tilde{\varphi}_f(t_k, \boldsymbol{\theta}^*)$ instead of $\hat{\varphi}_f(t_k, \boldsymbol{\theta}^*)$ in its implementation. This contradicts the assumption of this corollary.

Thus, the proof follows from Lemma A3.9 in [210], which in this case states that

$$\mathbb{E}\{\tilde{\varphi}_f(t_k, \boldsymbol{\theta}^*)\tilde{\varphi}_f^\top(t_k, \boldsymbol{\theta}^*)\}^{-1} \prec \\ \mathbb{E}\{\hat{\varphi}_f(t_k, \boldsymbol{\theta}^*)\tilde{\varphi}_f^\top(t_k, \boldsymbol{\theta}^*)\}^{-1}\mathbb{E}\{\hat{\varphi}_f(t_k, \boldsymbol{\theta}^*)\tilde{\varphi}_f^\top(t_k, \boldsymbol{\theta}^*)\}\mathbb{E}\{\tilde{\varphi}_f(t_k, \boldsymbol{\theta}^*)\hat{\varphi}_f^\top(t_k, \boldsymbol{\theta}^*)\}^{-1}.$$

□

4.5 Monte Carlo simulation studies

In this section, the asymptotic properties of the SRIVC estimator are tested through extensive simulations. We divide our experiments in consistency and asymptotic efficiency tests.

4.5.1 Consistency tests

Monte Carlo simulations are performed on a second order system to support the theoretical results developed in Section 4.3. The system is chosen to be

$$G^*(p) = \frac{1}{0.04p^2 + 0.2p + 1}, \quad (4.44)$$

and the true parameters are given by $\boldsymbol{\theta}^* = [a_1^*, a_2^*, b_0^*]^\top = [0.2, 0.04, 1]^\top$. The sampling period is chosen as $h = 0.1[\text{s}]$, and the system input is a random binary signal with a ZOH intersample behavior. The additive noise on the output is an i.i.d. Gaussian sequence with a variance of 0.1. The consistency of the SRIVC estimator is investigated by examining the mean and variance of the estimates in a Monte Carlo study as the sample size N increases. Here, N is adjusted from 50 to 200000

in a logarithmic scale, where a total of 100 different sample sizes are used. Three hundred Monte Carlo simulations are performed for each value of N with the mean and variance of the estimates being empirically calculated. The maximum number of iterations of the SRIVC algorithm is set to 200, and the relative error bound ϵ in line 15 of Algorithm 2.1 is set to 10^{-7} . The mean and variance of the estimated parameters with respect to an increasing sample size are examined under four different cases by changing the intersample behavior of the signals when discretizing different filters in the SRIVC algorithm. These cases include

- setting the intersample behavior of all the signals in the algorithm to that of $u(t)$, which uses a ZOH device,
- setting only the intersample behavior of $u(t_k)$ in the filtered regressor vector to FOH,
- setting only the intersample behavior of $u(t_k)$ in the filtered instrument vector to FOH, and
- setting only the intersample behavior of $y(t_k)$ to FOH.

These cases correspond to the first four instances in Figures 4.1 and 4.2.

In another simulation, a continuous-time multisine input with angular frequencies $[\omega_1, \omega_2, \omega_3, \omega_4] = [0.5, 2, 5, 7]$ [rad/s] is used to excite the true system $G^*(p)$. The noise-free system output is computed using the sinusoidal steady state response, i.e.,

$$x(t) = \sum_{l=1}^4 |G^*(i\omega_l)| \sin(\omega_l t + \angle G^*(i\omega_l)).$$

The experimental conditions and simulation settings are exactly the same as the random binary signal case. The model input is interpolated using an FOH to approximate the continuous-time signal $u(t)$. The mean and variance of the estimates for each sample size are calculated to examine the consistency of the SRIVC estimator in situations when the input cannot be reconstructed exactly. This corresponds to the “Sine FOH” instance in Figures 4.1 and 4.2.

The mean and variance of the estimated parameters for the five instances are shown in Figures 4.1 and 4.2 respectively, and the true parameters are plotted with a dotted line in Figure 4.1. It can be seen in Figure 4.2 that the variance of the estimates decreases with an increasing sample size in all cases. More oscillations in the mean values are observed for small sample sizes, but the estimates eventually converge to the true parameters after approximately 10000 samples when the intersample behavior of the model input in the filtered regressor matches that of the true system input, i.e., instances “All ZOH”, “Instrument U = FOH” and “Y = FOH”. Together with the decreasing variance, this provides empirical evidence to the consistency result in Theorem 4.1. In addition, changing the intersample behavior of the input in the filtered instrument vector or the output does not affect the consistency of SRIVC, which aligns with Statement 1 of Corollary 4.2 and Remark 4.2, respectively.

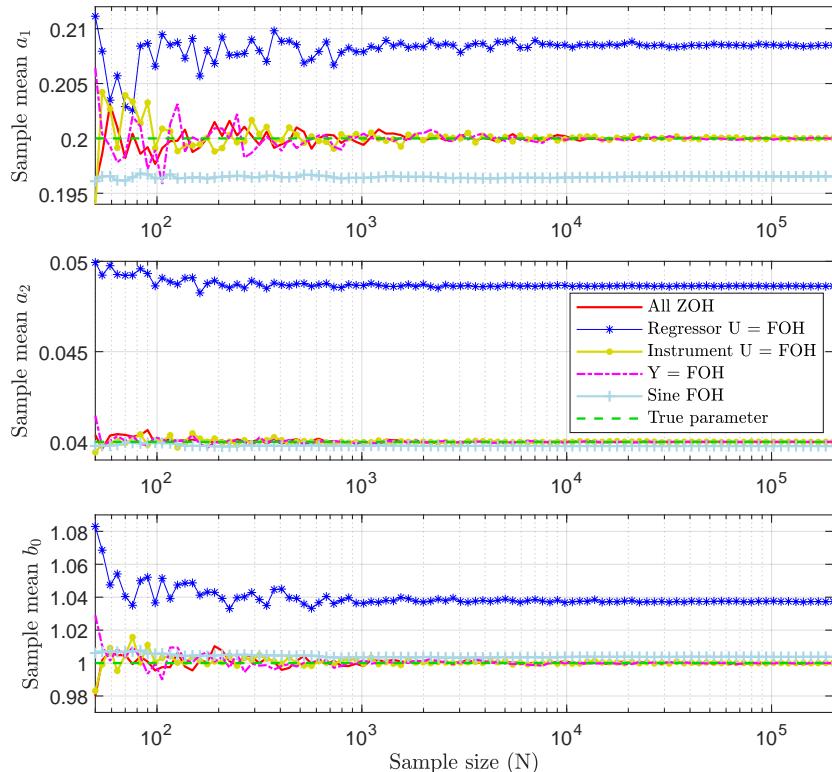


Figure 4.1: Mean of the estimated parameters for an increasing sample size.

It can also be seen that when the model input does not match the true system input, the estimates do not converge to the true parameters with an increasing sample size. The SRIVC estimator is not consistent in this case, which has been proven theoretically in Statement 2 of Corollary 4.2. Furthermore, when a continuous-time signal that is not interpolated exactly is used as the true system input, SRIVC can also be not consistent. We note that the bias on the SRIVC estimates can be reduced if more sophisticated interpolation methods other than a ZOH or FOH are used to reconstruct the input signal, as will be seen in Chapters 6 and 7. Furthermore, the bias will also decrease with a decreasing sampling period as seen from the proof of Statement 2 of Corollary 4.2.

4.5.2 Asymptotic efficiency tests

We now confirm the asymptotic efficiency results shown in Section 4.4. To this end, Monte Carlo simulations are performed on both first and second order systems.

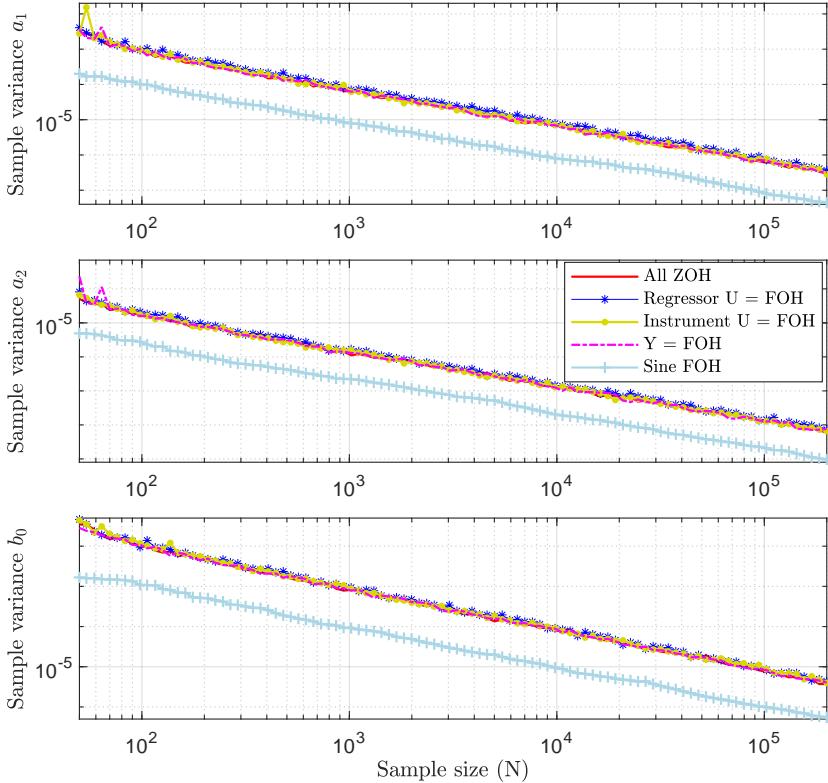


Figure 4.2: Variance of the estimated parameters for an increasing sample size.

Simulation 1: First order system

The first order system is chosen to be

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

and the parameter vector of true parameters is therefore $\boldsymbol{\theta}^* = [a_1^*, b_0^*]^\top = [0.1, 10]^\top$. The sampling period is chosen to be $h = 0.01[\text{s}]$ and the input is an i.i.d. Gaussian sequence with a unity variance, which is then interpolated with a ZOH device. The additive noise on the output is also an i.i.d. Gaussian sequence with a unity variance and is uncorrelated with the input. The sample size N is set to $2 \cdot 10^5$, and $5 \cdot 10^4$ Monte Carlo runs are performed with the SRIVC estimator initialized at $\boldsymbol{\theta}^*$. The maximum number of iterations of the SRIVC algorithm is set to 200, and the tolerance factor ϵ in line 15 of Algorithm 2.1 is set to 10^{-12} . The covariance of the asymptotic distribution of the SRIVC estimate given in (4.37) is then approximated

using $5 \cdot 10^4$ sets of estimates to be

$$\mathbf{P}_{\text{SRIVC}} = \begin{bmatrix} (8.0327 \pm 0.0508) \cdot 10^{-3} & 0.3996 \pm 0.0021 \\ 0.3996 \pm 0.0021 & 39.8223 \pm 0.2519 \end{bmatrix}, \quad (4.45)$$

where the mean and standard deviation of each covariance entry in (4.45) are determined using the method outlined in [211, Appendix B.9].

The expectation of two signals filtered by continuous-time transfer functions can be computed by first converting the continuous-time transfer functions to their discrete-time ZOH equivalents and then using the method outlined in Section 5 of [207]. The CRLB in (4.31) can then be computed analytically at $\boldsymbol{\theta}^*$ to be

$$\mathbf{P}_{\text{CR}} = \begin{bmatrix} 8.0334 \cdot 10^{-3} & 0.4010 \\ 0.4010 & 40.0333 \end{bmatrix}, \quad (4.46)$$

where (4.46) is accurate up to four decimal places. The analytical expression of the CRLB in (4.46) matches well with the approximated covariance matrix in (4.45) within the standard deviations.

Now, it has been stated in the existing literature [34, 61, 66, 69] that the SRIVC estimator is asymptotically efficient with the covariance matrix, and therefore also the CRLB, defined in the same form as (4.37) but with the filtered regressor given by (see e.g. [61, p. 105])

$$\begin{aligned} \tilde{\boldsymbol{\varphi}}_f(t_k) &= \frac{1}{A^*(p)} \left[-x^{(1)}(t_k), \dots, -x^{(n^*)}(t_k), u(t_k), \dots, u^{(m^*)}(t_k) \right]^\top \\ &= \frac{1}{A^*(p)} \left[-px(t_k), \dots, -p^{n^*}x(t_k), u(t_k), \dots, p^{m^*}u(t_k) \right]^\top, \end{aligned} \quad (4.47)$$

where $x(t_k)$ is the sampled version of the noise-free system output. It can be seen that (4.47) is different to the filtered regressor defined in (4.38) by Theorem 4.3. Computing the covariance matrix of the SRIVC estimator of the CRLB expression using (4.47) will lead to the ambiguous assumption of the intersample behavior of the noise-free system output $x(t_k)$, which is unknown in practice since only sampled signals are available as measurements. The key difference between the covariance expressions derived in Section 4.4 and the expressions given in the existing literature is that the evaluation of the filtered regressor in (4.38) implicitly assumes that the noise-free system output is a continuous-time signal, whereas (4.47) assumes the noise-free output $x(t_k)$ is interpolated in some manner where a mixed notation of continuous-time transfer function and sampled data is used [61, p. 96]. For instance, assuming a ZOH reconstruction for $x(t_k)$ during the filtering operations when computing the covariance matrix from the existing literature [61] will result in

$$\mathbf{P}_{\text{SRIVC}}^{\text{lit}} = \begin{bmatrix} 7.2629 \cdot 10^{-3} & 0.3813 \\ 0.3813 & 40.0333 \end{bmatrix},$$

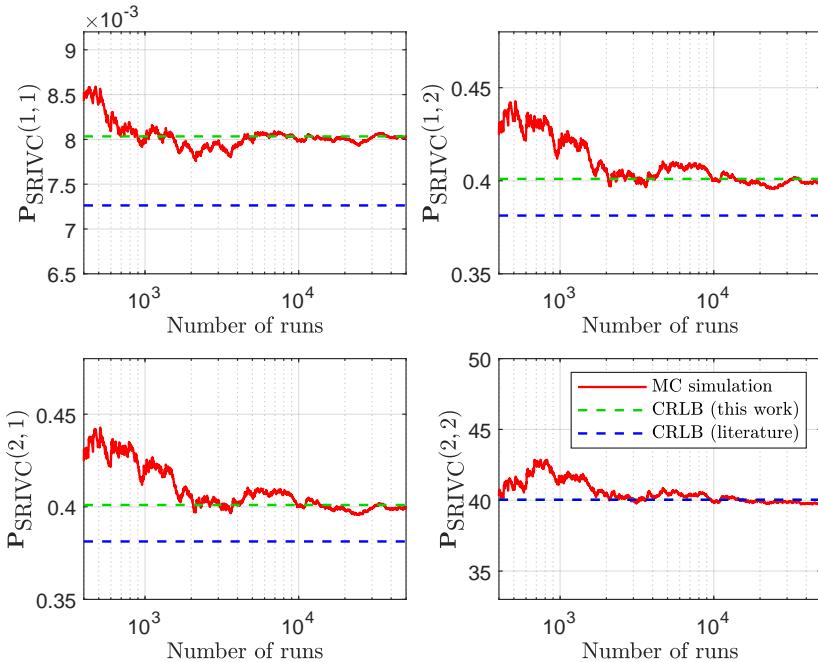


Figure 4.3: Covariance of the first-order transfer function estimates. Note that the dashed green and dashed blue lines match in the bottom right plot.

which does not match the asymptotic covariance matrix approximated using the Monte Carlo simulations in (4.45).

The covariance matrix approximated through the Monte Carlo runs is plotted against the number of runs as shown in Figure 4.3. It can be seen that the covariance matrix of the SRIVC estimator obtained in simulation converges to the CRLB derived in this chapter with an increasing number of Monte Carlo runs. This provides empirical evidence that the SRIVC estimator is asymptotically efficient under the output error model structure. On the other hand, the CRLB or the covariance expression in the existing literature does not match the simulated covariance results.

Simulation 2: Second order system

The second order system is chosen to be the one in (4.44). The input and additive noise settings as well as the relative error bound and maximum iterations of the SRIVC algorithm are exactly the same as the settings for the first-order system simulation. The sampling period is chosen to be $h = 0.1[\text{s}]$, the sample size N is varied from 10^3 to $2 \cdot 10^5$ in eight steps, and 10^4 Monte Carlo runs are performed for each sample size with the SRIVC estimator initialized at $\boldsymbol{\theta}^*$. The covariance matrix of the estimates are then approximated from the distributions generated by the Monte Carlo runs, and the three diagonal entries are plotted against N in Figure 4.4.

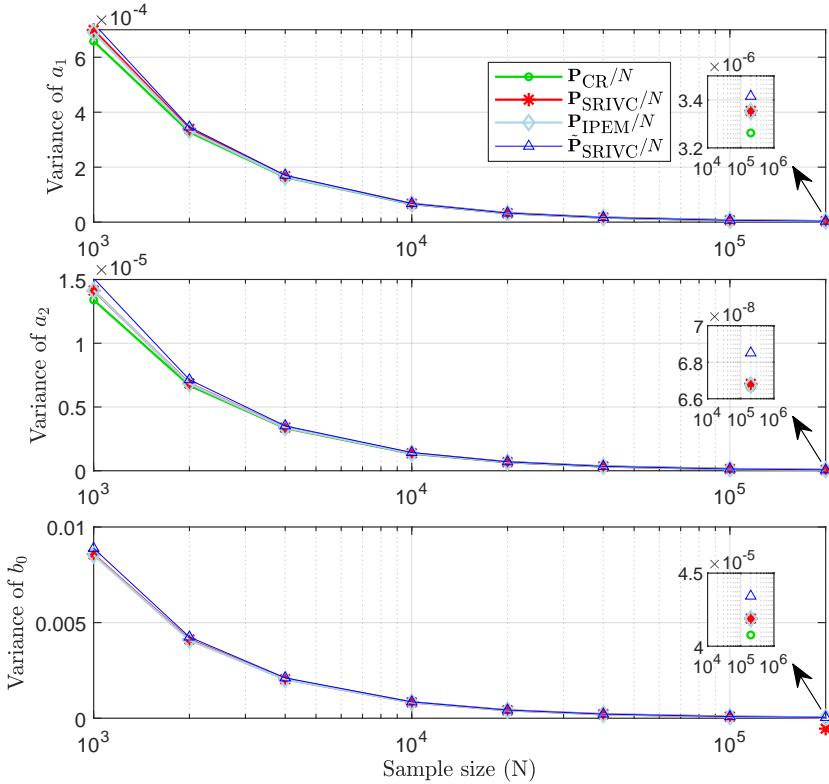


Figure 4.4: Variance of the second-order transfer function estimates.

In addition to the CRLB calculated using (4.31), the system parameters are also estimated with the asymptotically efficient IPEM method derived in Chapter 3 for comparison, as shown in Figure 4.4. Furthermore, the fourth instance in Figure 4.4 corresponds to an SRIVC estimator with an FOH input signal in the filtered instrument vector with the covariance matrix labeled as \tilde{P}_{SRIVC}/N .

It can be observed in Figure 4.4 that the variance of the SRIVC estimates converges quickly to the CRLB with an increasing sample size, and the small discrepancies are due to the finite sample approximation to the covariance matrix. The variance of the SRIVC estimates are indistinguishable from that of the asymptotically efficient IPEM estimates as seen in the upper right windows in Figure 4.4. These results provide empirical evidence to the asymptotic efficiency of the SRIVC estimator. On the other hand, when the intersample behavior of the input signal in the filtered instrument vector does not match the system input in the SRIVC estimator, a higher variance of the estimates can be observed, which indicates that the estimator is not efficient in this case. This agrees with Corollary 4.3.

4.6 Conclusions

In this chapter we have analyzed the consistency and asymptotic efficiency of the SRIVC estimator by taking into account the intersample behavior of the input signal, and conducted simulations to provide empirical observations to the theoretical results. The first result of the chapter is that the SRIVC estimator is generically consistent when the intersample behavior of the continuous-time input, ZOH or FOH, is known exactly and subsequently used in the implementation of the algorithm. It has been shown that when the intersample behavior of the input signal in the filtered regressor, i.e., the model input, does not match that of the true system input, the unique converging point of the estimator no longer corresponds to the true parameters, and thus the SRIVC estimator is generically not consistent. On the other hand, the intersample behavior of the input signal in the filtered instrument vector does not affect the consistency of the estimator, and neither does the intersample behavior of output signal in the filtered regressor vector, unless it is excessively different from that of the noiseless output.

We have also derived the asymptotic Cramér-Rao lower bound for the continuous-time output error model structure and provided an expression for the asymptotic covariance of the SRIVC estimator by explicitly incorporating the intersample behavior of the signals as part of the analysis. The asymptotic CRLB and the covariance expression derived in this chapter are both different to the results reported in the existing literature. It has been shown that the asymptotic CRLB is independent of the intersample behavior of the noise-free system output and hence it only depends on the intersample behavior of the system input. We conclude that the SRIVC estimator is asymptotically efficient under the output error model structure, i.e., the asymptotic covariance expression coincides with the asymptotic CRLB, when the intersample behavior of the input signal in both the filtered regressor and instrument vectors matches that of the system input.

Appendix

4.A Supplementary material for consistency results

Lemma 4.1. *Under Assumption 4.2, with the input $\{u(t_k)\}$ assumed to be interpolated by an FOH device, we state that $\mathbb{E}\{\hat{\varphi}_f(t_k)\mathbf{v}_f^\top(t_k)\} = \mathbf{0}$.*

Proof. By writing the instrument vector as in (4.16), it is sufficient to prove that

$$\Psi := \mathbb{E} \left\{ \frac{1}{A_j^2(p)} \mathbf{u}_{n+m}(t_k) \mathbf{v}_f^\top(t_k) \right\} = \mathbf{0}.$$

Denote the (r, l) -entry of Ψ as Ψ_{rl} . Since the last $m + 1$ elements of $\mathbf{v}_f(t_k)$ are zero, we have that $\Psi_{rl} = 0$ for $l = n + 1, n + 2, \dots, n + m + 1$. An arbitrary entry in the first n columns of Ψ can be written as

$$\begin{aligned} \Psi_{rl} &= \mathbb{E} \left\{ \frac{p^{n+m+1-r}}{A_j^2(p)} u(t_k) \frac{p^l}{A_j(p)} v(t_k) \right\} \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{C_1(e^{i\omega})}{A_{d,j}^2(e^{i\omega})} \frac{C_2(e^{-i\omega})}{A_{d,j}(e^{-i\omega})} \phi_{uv}(\omega) d\omega, \end{aligned} \quad (4.48)$$

where $C_1/A_{d,j}^2$ and $C_2/A_{d,j}$ are the FOH equivalents of their respective continuous-time transfer functions, $r = 1, \dots, n + m + 1, l = 1, \dots, n$, and $\phi_{uv}(\omega)$ is the cross-spectrum between $u(t_k)$ and $v(t_k)$. Since the input and noise are uncorrelated according to Assumption 4.2, it holds that $\phi_{uv}(\omega) = 0$. Thus, $\Psi_{rl} = 0$ for every $r, l = 1, \dots, n + m + 1$, hence $\Psi = \mathbf{0}$ and the statement of the lemma is true. \square

Lemma 4.2. *Consider an FOH input reconstruction, and that Assumptions 4.1, 4.3 and 4.6 hold. Then, the matrix Φ in (4.18) when evaluated at the true system parameters is positive definite, i.e.,*

$$\Phi^* := \mathbb{E} \left\{ \frac{1}{A^{*2}(p)} \mathbf{u}_{n+m}(t_k) \frac{1}{A^{*2}(p)} \mathbf{u}_{n+m}^\top(t_k) \right\} \succ \mathbf{0}. \quad (4.49)$$

Proof. For an arbitrary vector $\mathbf{z} \in \mathbb{R}^{n+m+1}$, we can write

$$\mathbf{z}^\top \Phi^* \mathbf{z} = \mathbb{E} \left\{ \left(\frac{B_\mathbf{z}(p)}{A^{*2}(p)} u(t_k) \right)^2 \right\} \geq 0, \quad (4.50)$$

where $B_{\mathbf{z}}(p)$ is an arbitrary polynomial of degree $n + m$. In the frequency domain, (4.50) can be written as

$$\mathbf{z}^\top \Phi^* \mathbf{z} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \frac{\tilde{B}_{\mathbf{z}}(e^{i\omega})}{A_d^{*2}(e^{i\omega})} \right|^2 dF_u(\omega),$$

where $\tilde{B}_{\mathbf{z}}(e^{i\omega})$ and $A_d^{*2}(e^{i\omega})$ are the FOH equivalent polynomials of $B_{\mathbf{z}}(p)/A^{*2}(p)$ evaluated at $z = e^{i\omega}$, and $F_u(\omega)$ is the spectral distribution of $\{u(t_k)\}$. Note that $\tilde{B}_{\mathbf{z}}(e^{i\omega})$ is a $2n$ -degree polynomial in the variable $e^{i\omega}$. We can also write (4.50) as

$$\mathbf{z}^\top \Phi^* \mathbf{z} = \frac{1}{2\pi} \int_{-\pi}^{\pi} |\tilde{B}_{\mathbf{z}}(e^{i\omega})|^2 dF_{\bar{u}}(\omega),$$

where the support of the spectral distribution function $F_{\bar{u}}$ consists of at least $2n + 1$ points, since filtering $u(t_k)$ by $1/A_d^{*2}(q)$ gives a signal which is also persistently exciting of order at least $2n + 1$. By the definition of persistence of excitation, $\mathbf{z}^\top \Phi^* \mathbf{z} = 0$ implies that $\tilde{B}_{\mathbf{z}}(e^{i\omega}) = 0$ for all $\omega \in [-\pi, \pi]$ (cf. sufficiency proof of [136, Thm. 1]). This means that $B_{\mathbf{z}}(p)/A^{*2}(p)$ gives a sampled model equal to zero at all sampling instants.

Now, assume that there exists a vector $\mathbf{z}^* \in \mathbb{R}^{n+m+1}$ such that, for all integers k , we have $[B_{\mathbf{z}^*}(p)/A^{*2}(p)]u(t_k) = 0$. Thus, by linearity, for all \mathbf{z} we reach

$$\frac{B_{\mathbf{z}}(p)}{A^{*2}(p)} u(t_k) = \frac{B_{\mathbf{z}}(p) + B_{\mathbf{z}^*}(p)}{A^{*2}(p)} u(t_k).$$

If $B_{\mathbf{z}^*}(p)$ were non-zero, the continuous-time model is not uniquely determined by the discrete-time model. However, this is not true under the sampling condition of the statement (see [122]). Thus, it is not possible for a polynomial $B_{\mathbf{z}}(p)$ different from zero to give a sampled model equal to zero. This means that (4.50) is strictly positive for any non-zero vector \mathbf{z} , which implies that Φ^* is positive definite. \square

Remark 4.3. Even though there are $n + m + 1$ parameters in the continuous-time transfer function to be identified, the input is assumed to be persistently exciting of order $2n + 1$ instead of $n + m + 1$ for an FOH reconstruction of the input. We want to clarify that Assumption 4.3 has not been shown to be a sufficient and necessary condition for the result in Lemma 4.2 to be true. In fact, an input that is persistently exciting of order $n + m + 1$ is enough in many practical scenarios, as long as the frequency lines of the input do not coincide with the zeros of the discrete-time equivalent of $1/A^{*2}(p)$. A similar conclusion has been obtained for general linear prefilters in continuous-time system algorithms in Section 5 of [193]. Proving the sufficiency of an input persistent excitation order of $n + m + 1$ is not an easy task, since the result is directly related to the zeros of the sampled model of a continuous-time system. These zeros are known to not have straightforward expressions with respect to the continuous-time system parameters [17], and the transfer functions of the zero and first-order holds do not provide much information on them.

Remark 4.4 (A counterexample for the non-singularity of Φ^*). Following the idea in [122], we can see that the sampling period plays an important role in the non-singularity of Φ^* in (4.49). If the sampling period is poorly chosen, the matrix Φ^* can be singular. As an example, consider an input reconstructed through a ZOH, $n = 2$, $m = 1$, and

$$A^*(p) = \frac{1}{\alpha^2 + (\pi/h)^2} p^2 + \frac{2\alpha}{\alpha^2 + (\pi/h)^2} p + 1,$$

where $\alpha > 0$ and h is the sampling period. We write

$$\mathbf{z}^\top \Phi^* \mathbf{z} = (\alpha^2 + (\pi/h)^2)^2 \mathbb{E} \left\{ \left(\frac{B_{\mathbf{z}}(p)}{(p^2 + 2\alpha p + \alpha^2 + (\pi/h)^2)^2} u(t_k) \right)^2 \right\}.$$

For the choice $\mathbf{z} = [0, 1, \alpha, 0]^\top$, we have

$$\mathbf{z}^\top \Phi^* \mathbf{z} = (\alpha^2 + (\pi/h)^2)^2 \mathbb{E} \left\{ \left(\frac{p(p + \alpha)}{(p^2 + 2\alpha p + \alpha^2 + (\pi/h)^2)^2} u(t_k) \right)^2 \right\}.$$

Denote the transfer function in the expectation above as $\bar{G}(p)$. The step response of this system is

$$\mathcal{L}^{-1} \left\{ \frac{\bar{G}(s)}{s} \right\} = \frac{h}{2\pi} t e^{-\alpha t} \sin \left(\frac{\pi}{h} t \right).$$

Evaluating this step function at $t = kh$, we obtain

$$\mathcal{L}^{-1} \left\{ \frac{\bar{G}(s)}{s} \right\}_{t=kh} = 0 \quad \text{for all integers } k,$$

from which we conclude that the zero-order hold equivalent of $\bar{G}(p)$ is exactly zero. Hence, $\mathbf{z}^\top \Phi^* \mathbf{z} = 0$ but clearly $\mathbf{z} \neq 0$. This implies that Φ^* is singular. This counterexample shows that the choice of the sampling period is important for the non-singularity of Φ^* , and therefore of Φ .

Lemma 4.3. Consider the transfer function $1/(a_n p^n + \dots + a_1 p + 1)$ and the denominator of its discrete-time equivalent, written as $q^n + \alpha_{n-1} q^{n-1} + \dots + \alpha_0$. The mapping between the continuous-time parameter vector $[a_1, \dots, a_n]$ and the discrete-time parameters $[\alpha_0, \dots, \alpha_{n-1}]$ is real analytic in $\{[a_1, \dots, a_n] \in \mathbb{R}^n : a_n \neq 0\}$.

Proof. For $a_n \neq 0$, we can write

$$\begin{aligned} A'(p) &:= p^n + \frac{a_{n-1}}{a_n} p^{n-1} + \dots + \frac{a_1}{a_n} p + \frac{1}{a_n} \\ &= p^n + a'_{n-1} p^{n-1} + \dots + a'_1 p + a'_0, \end{aligned}$$

where we note that the mapping $[a_1, a_2, \dots, a_n] \rightarrow [a'_0, a'_1, \dots, a'_{n-1}]$ is real analytic for $a_n \neq 0$. Let the state matrix of the continuous-time system $1/A'(p)$ be \mathbf{A}_c . The state matrix of the discrete-time equivalent is then $\mathbf{A}_d = \exp(\mathbf{A}_c h)$, where h is the sampling period. The (matrix) exponential function is also real analytic in the variables $[a'_0, a'_1, \dots, a'_{n-1}]$. It is known that the coefficients of the characteristic polynomial $A_d(z) = \det(z\mathbf{I} - \mathbf{A}_d)$ are polynomial expressions in the entries of the matrix \mathbf{A}_d (see e.g., [26]). This implies that $\{\alpha_k\}_{k=0}^{n-1}$ are real analytic functions of the entries of \mathbf{A}_d . The lemma then follows from the composition of real analytic functions. \square

Lemma 4.4 (Analyticity of Φ). *Each entry of Φ in (4.18) is an analytic function of a_1, \dots, a_n for $[a_1, \dots, a_n] \in \Omega$, where Ω denotes the subset of \mathbb{R}^n consisting of parameter vectors $[a_1, \dots, a_n]$ such that $A_j(p)$ has all zeros strictly in the left half-plane.*

Proof. An arbitrary element of Φ can be written as

$$\begin{aligned}\Phi_{rl} &= \mathbb{E} \left\{ \frac{p^{n+m+1-r}}{A_j^2(p)} u(t_k) \frac{p^{n+m+1-l}}{A_j(p) A^*(p)} u(t_k) \right\} \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\tilde{B}_r(e^{i\omega})}{A_{d,j}^2(e^{i\omega})} \frac{\tilde{B}_l(e^{-i\omega})}{A_{d,j}(e^{-i\omega}) A_d^*(e^{-i\omega})} dF_u(\omega).\end{aligned}\quad (4.51)$$

Denote the denominator of the FOH equivalent of $1/A_j(p)$ as $A_{d,j}(q) := q^n + \alpha_{n-1}q^{n-1} + \dots + \alpha_0$, and Ω_d as the subset of \mathbb{R}^n consisting of parameter vectors $[\alpha_0, \dots, \alpha_{n-1}]$ such that $A_{d,j}(q)$ has all zeros strictly inside the unit circle. By Lemma 4.3, there is an analytic mapping between $[a_1, \dots, a_n]$ and the discrete-time parameter vector $[\alpha_0, \dots, \alpha_{n-1}]$. Now, fixing $\alpha_1, \dots, \alpha_{n-1}$ allows us to define a region $\Omega_{d0} \subset \mathbb{R}$ where $\bar{\alpha}_0 \in \Omega_{d0}$ implies that $[\bar{\alpha}_0, \alpha_1, \dots, \alpha_{n-1}] \in \Omega_d$. Note that the integrand in (4.51) is a real analytic function of α_0 in Ω_{d0} , and from now on, we denote this integrand as $f(\bar{\alpha}_0, \omega)$.

Let \mathcal{C} be a closed contour in Ω_{d0} such that $\bar{\alpha}_0$ is interior to \mathcal{C} . Then,

$$f(\bar{\alpha}_0, \omega) = \frac{1}{2\pi i} \int_{\mathcal{C}} \frac{f(\alpha_0, \omega)}{\alpha_0 - \bar{\alpha}_0} d\alpha_0.$$

As a result,

$$\Phi_{rl}(\bar{\alpha}_0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{2\pi i} \int_{\mathcal{C}} \frac{f(\alpha_0, \omega)}{\alpha_0 - \bar{\alpha}_0} d\alpha_0 dF_u(\omega).$$

Since the function being integrated is bounded on $[-\pi, \pi] \times \mathcal{C}$, the order of integration can be changed by Fubini's Theorem [44], which yields

$$\begin{aligned}\Phi_{rl}(\bar{\alpha}_0) &= \frac{1}{2\pi i} \int_{\mathcal{C}} \frac{1}{\alpha_0 - \bar{\alpha}_0} \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\alpha_0, \omega) dF_u(\omega) d\alpha_0 \\ &= \frac{1}{2\pi i} \int_{\mathcal{C}} \frac{\Phi_{rl}(\alpha_0)}{\alpha_0 - \bar{\alpha}_0} d\alpha_0,\end{aligned}$$

from which we conclude that $\Phi_{rl}(\alpha_0)$ is real analytic in a neighborhood around $\alpha_0 = \bar{\alpha}_0$. Repeating this process for every α_k , $k = 1, \dots, n - 1$, we obtain that Φ_{rl} is an analytic function of the variables $\alpha_0, \alpha_1, \dots, \alpha_{n-1}$ separately. Since Φ_{rl} is a continuous function of $[\alpha_0, \dots, \alpha_{n-1}]$ in Ω_d , Φ_{rl} is a real analytic function of the joint variables $[\alpha_0, \dots, \alpha_{n-1}]$ by Osgood's Lemma [23, p. 139]. Hence, each entry of Φ is a real analytic function of a_1, \dots, a_n for $[a_1, \dots, a_n] \in \Omega$. \square

4.B Supplementary material for efficiency results

Lemma 4.5 (Equivalence between practical and theoretical SRIVC estimators). *Consider the SRIVC iterations in (2.35) for finite N , where we denote by $\varphi_f(t_k, \theta_j^N)$, $\hat{\varphi}_f(t_k, \theta_j^N)$ and $y_f(t_k, \theta_j^N)$ the expressions in (2.32), (2.33) and (2.34), respectively. Let the converging point of the iterations be $\bar{\theta}^N := \lim_{j \rightarrow \infty} \theta_j^N$, and assume that the modified normal matrix*

$$\frac{1}{N} \sum_{k=1}^N \hat{\varphi}_f(t_k, \bar{\theta}^N) \varphi_f^\top(t_k, \bar{\theta}^N) \quad (4.52)$$

is non-singular. Also, consider the theoretical SRIVC estimator given in Definition 4.1 at the converging point $\bar{\theta}^N$, and further assume that (4.52) remains non-singular with the filtered regressor replaced by (4.35). Then, there exists an integer N_0 such that the SRIVC estimator given in (2.35) is equivalent to the theoretical SRIVC estimator given in Definition 4.1 evaluated at the converging point for $N > N_0$, that is, (2.35) at the converging point can be expressed as

$$\bar{\theta}^N = \left[\sum_{k=1}^N \hat{\varphi}_f(t_k, \bar{\theta}^N) \dot{\varphi}_f(t_k, \bar{\theta}^N) \right]^{-1} \left[\sum_{k=1}^N \hat{\varphi}_f(t_k, \bar{\theta}^N) \dot{y}_f(t_k, \bar{\theta}^N) \right].$$

Proof. The converging point $\bar{\theta}^N$ of the SRIVC estimator must satisfy (2.35), i.e.,

$$\bar{\theta}^N = \left[\sum_{k=1}^N \hat{\varphi}_f(t_k, \bar{\theta}^N) \varphi_f^\top(t_k, \bar{\theta}^N) \right]^{-1} \left[\sum_{k=1}^N \hat{\varphi}_f(t_k, \bar{\theta}^N) y_f(t_k, \bar{\theta}^N) \right].$$

Equivalently,

$$\sum_{k=1}^N \hat{\varphi}_f(t_k, \bar{\theta}^N) y_f(t_k, \bar{\theta}^N) - \sum_{k=1}^N \hat{\varphi}_f(t_k, \bar{\theta}^N) \varphi_f^\top(t_k, \bar{\theta}^N) \bar{\theta}^N = \mathbf{0}. \quad (4.53)$$

Note that

$$\varphi_f^\top(t_k, \bar{\theta}^N) \bar{\theta}^N = \frac{-\bar{a}_n p^n - \dots - \bar{a}_1 p}{\bar{a}_n p^n + \dots + \bar{a}_1 p + 1} y(t_k) + \frac{\bar{B}_N(p)}{\bar{A}_N(p)} u(t_k),$$

where the filtering on $y(t_k)$ depends on the hold reconstruction that is chosen. Therefore, (4.53) reduces to

$$\sum_{k=1}^N \hat{\varphi}_f(t_k, \bar{\theta}^N) \left(y(t_k) - \frac{\bar{B}_N(p)}{\bar{A}_N(p)} u(t_k) \right) = \mathbf{0}. \quad (4.54)$$

Now, consider the theoretical SRIVC estimator evaluated at the converging point with the filtered regressor and the filtered output given by (4.35) and (4.36). We have

$$\begin{aligned} \dot{y}_f(t_k, \bar{\theta}^N) - \dot{\varphi}_f^\top(t_k, \bar{\theta}^N) \bar{\theta}^N &= \frac{B^*(p)}{\bar{A}_N(p) A^*(p)} u(t_k) + \frac{1}{\bar{A}_N(p)} v(t_k) + \frac{(\bar{a}_n p^n + \dots + \bar{a}_1 p)}{\bar{A}_N(p) A^*(p)} u(t_k) \\ &\quad + \frac{(\bar{a}_n p^n + \dots + \bar{a}_1 p)}{\bar{A}_N(p)} v(t_k) - \frac{\bar{B}_N(p)}{\bar{A}_N(p)} u(t_k) \\ &= \frac{\bar{A}_N(p) B^*(p)}{\bar{A}_N(p) A^*(p)} u(t_k) + \frac{\bar{A}_N(p)}{\bar{A}_N(p)} v(t_k) - \frac{\bar{B}_N(p)}{\bar{A}_N(p)} u(t_k) \\ &= y(t_k) - \frac{\bar{B}_N(p)}{\bar{A}_N(p)} u(t_k), \end{aligned}$$

which gives the same expression as the standard SRIVC estimator in (4.54). This means that both estimators solve the same equation for the parameters at the converging point for every sample size N . Following the proof in Statement 2 of Theorem 4.1, it can be shown that there is a unique converging point for the SRIVC estimator as the sample size approaches infinity. Therefore, replacing the filtered regressor vector and the filtered output in the SRIVC estimator by (4.35) and (4.36) will result in the same estimate $\bar{\theta}^N$ for all sufficiently large sample size. \square

4.C A result from real analytic function theory

The next lemma from real analytic function theory is of interest for proving generic consistency.

Lemma 4.6 (Lemma A2.3 of [210]). *Let $\Omega \subset \mathbb{R}^n$ be a open connected set, and $\mathbf{x} = [x_1, \dots, x_n]$. Consider a function $f: \Omega \rightarrow \mathbb{R}$ such that*

1. *$f(\mathbf{x})$ is real analytic in x_l ($l = 1, \dots, n$) for $\mathbf{x} \in \Omega$ and*
2. *there is a point $\mathbf{x}^* \in \Omega$ such that $f(\mathbf{x}^*) \neq 0$.*

Then, the set

$$M = \{\mathbf{x} \in \Omega: f(\mathbf{x}) = 0\}$$

has zero Lebesgue measure.

Proof. An elementary proof of this result can be found in [155]. \square

Corollary 4.4. Consider the square matrix $\mathbf{R}(\mathbf{x}) \in \mathbb{R}^{n \times n}$ as a function of the finite dimensional vector \mathbf{x} which belongs to a domain Ω . Assume that each entry $\mathbf{R}_{rl}(\mathbf{x})$ of $\mathbf{R}(\mathbf{x})$ is a real analytic function of \mathbf{x} in Ω . If there exists a vector \mathbf{x}^* such that $\mathbf{R}(\mathbf{x}^*)$ is nonsingular, then $\mathbf{R}(\mathbf{x})$ is generically non-singular with respect to Ω .

Proof. Take $f(\mathbf{x}) = \det(\mathbf{R}(\mathbf{x}))$, which is real analytic in Ω since it is the composition of real analytic functions. The corollary follows from Lemma 4.6. \square

Theoretical and practical aspects of SRIVC for over-parameterized models

In Chapter 4 we have proven the generic consistency and asymptotic efficiency of the SRIVC method under mild assumptions. One assumption for consistency involves limiting the amount of over-parametrization allowed in the model. Here we extend this analysis by studying the convergence of the SRIVC estimator for models that suffer over-parameterization in both numerator and denominator polynomials. The results we derive are later 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.

5.1 Introduction

Despite more than forty years of use, formal theoretical treatments of the SRIVC method have only been pursued in recent years. In this thesis, we 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-parametrization possibly occurring in the numerator or denominator of the transfer function description but not both simultaneously. When the model structure is not known in advance (as in most applications), most model order selection algorithms try estimating models of increasing complexity that are eventually over-parameterized in both numerator and denominator polynomials [138]. Thus, a natural follow-up to Chapter 4 is to derive statistical properties of the SRIVC estimator that cover these scenarios.

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 [206, 209]. The possible convergence points of the Steiglitz-McBride algorithm have also been

shown to enjoy the same property [218], as well as various instrumental variable variants [208]. 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 method since the algorithm requires specifying the intersample behaviors of the input and output; in fact, a misspecification of the input intersample behavior can lead to inconsistent estimates with asymptotic bias decreasing with smaller sampling periods as studied in the previous chapter. In addition, the SRIVC algorithm computes an estimate of the covariance matrix of the parameter vector as a byproduct, which is used for computing Young's information criterion (YIC, [245]) 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, and it can also reveal the theoretical and practical problems associated with over-parameterization in continuous-time system identification.

The main results of this chapter are based on [93]. In summary,

- we show that, under over-parametrization of both numerator and denominator, the numerical conditioning of the modified normal matrix being inverted in the SRIVC algorithm deteriorates for decreasing sampling periods;
- we prove that at a limit point of the iterative procedure and as the sample size tends to infinity, the SRIVC estimator 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-parametrization.

The rest of the chapter is organized as follows. In the next section we describe the system and model considerations. In Section 5.3 we present the main theoretical findings, and we corroborate them through simulations in Section 5.4. Conclusions are drawn in Section 5.5. Supplementary material concerning the proofs can be found in the appendix of this chapter.

5.2 System and model setup

Consider the following LTI, asymptotically stable, continuous-time system

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

with the system transfer function being parameterized as

$$G^*(p) = \frac{B^*(p)}{A^*(p)} = \frac{b_{m^*}^* p^{m^*} + b_{m^*-1}^* p^{m^*-1} + \cdots + b_0^*}{a_{n^*}^* p^{n^*} + a_{n^*-1}^* p^{n^*-1} + \cdots + a_1^* p + 1},$$

where the polynomials $B^*(p)$ and $A^*(p)$ are assumed to be coprime. The true parameter vector is given by $\theta^* = [a_1^*, \dots, a_{n^*}^*, b_0^*, \dots, b_{m^*}^*]^\top$. The output is assumed to be contaminated with noise prior to sampling, i.e.,

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

where we have used the notation introduced in Remark 2.5. The measurement noise $v(t_k)$ is described by a zero-mean stochastic process that is independent of the sampled input signal $u(t_k)$. The intersample behavior of the input is assumed to be known and produced by a hold device such as a ZOH or FOH, and both input and output signals are sampled regularly with sampling period h .

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 fitted to the data. The key assumption throughout this chapter is that the model is over-parameterized in both numerator and denominator polynomials, that is, $m > m^*$ and $n > n^*$. The goal is to determine the asymptotic properties of the SRIVC estimator in this case and to examine the practical implications of selecting over-parameterized models in the performance of this method.

5.3 Analysis of SRIVC for over-parameterized models

As mentioned in Chapter 2, the usual implementation of the SRIVC algorithm (as in the CONTSID toolbox [67]) considers a monic denominator for the model, i.e., the non-zero coefficient of highest degree is set to one. Throughout this thesis, and specially this chapter, we analyze the SRIVC estimator that uses an anti-monic denominator polynomial for the model, that is, the constant term of the denominator polynomial is fixed to one. The implementation of this method is given in Algorithm 2.1. With this description, it is feasible that the procedure converges to the true denominator if it is over-parameterized as the leading coefficient of the denominator is not fixed. 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 size, which might induce more instability problems in the iterations.

This section is divided into four subsections: an asymptotic analysis of the modified normal matrix of the SRIVC estimator, a convergence analysis of the

stationary points of the method, a one-iteration analysis of the SRIVC algorithm, and a review of the normalized error variance norm.

In Chapter 4 we have 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 $\min(n - n^*, m - m^*) = 0$, where n and m are the denominator and numerator polynomial degrees, respectively. One question we address here is whether some notion of consistency is still achievable for over-parameterized model structures. Before stating our results, we introduce the assumptions used throughout this chapter.

Assumption 5.1. The true system $B^*(p)/A^*(p)$ is proper ($n^* \geq m^*$) and asymptotically stable with $A^*(p)$ and $B^*(p)$ being coprime.

Assumption 5.2. The input sequence, $u(t_k)$, and disturbance, $v(t_l)$, are stationary and mutually independent for all integers k and l .

Assumption 5.3. The input sequence is persistently exciting of order no less than $2n + 1$.

Assumption 5.4. The sampling frequency is more than twice the largest positive imaginary part of the zeros of $A_j(p)A^*(p)$.

Assumption 5.5. The input is piece-wise linear between samples (i.e., it is exactly reconstructed from an FOH device), and it is correctly specified in the SRIVC algorithm.

Assumptions 5.1 to 5.5 are standard and have been used in the previous chapter. All of the results in this chapter also hold for cases when the input is generated by a zero-order hold device, under the condition that the input must be persistently exciting of order at least $2n$ instead of $2n + 1$ in case the model is strictly proper.

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

Assumption 5.6. The degrees of the model polynomials, n and m , satisfy

$$\min(n - n^*, m - m^*) =: n_{op} > 0, \quad (5.1)$$

and

$$2n^* + n_{op} \geq n + m. \quad (5.2)$$

Assumption 5.7. For every integer 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.

The condition (5.1) in Assumption 5.6 implies that the model is over-parameterized in both numerator and denominator polynomials, which is the case of interest in this work. Analogously to the results in Chapter 4, the technical condition (5.2) is

set in order 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. Note that Assumption 5.7 is equivalent to Assumption 4.4 of Chapter 4 for when $n_{op} = 0$, and it is commonly satisfied in practice.

5.3.1 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 5.2, the ergodic lemmas in [205, Lemma 3.1] and [210, Lemma A4.3] permit us to write the SRIVC iterations in (2.35) as

$$\boldsymbol{\theta}_{j+1} = \mathbb{E}\{\hat{\varphi}_f(t_k, \boldsymbol{\theta}_j)\varphi_f^\top(t_k, \boldsymbol{\theta}_j)\}^{-1}\mathbb{E}\{\hat{\varphi}_f(t_k, \boldsymbol{\theta}_j)y_f(t_k, \boldsymbol{\theta}_j)\}. \quad (5.3)$$

Recall that the matrix being inverted above is denoted as the modified normal matrix of the SRIVC method. For over-parameterized models, we show in Proposition 5.1 that the condition number of the matrix being inverted in (5.3) tends to infinity as the sampling period tends to zero.

Proposition 5.1. Consider a continuous-time system with polynomial degrees n^* and m^* in the denominator and numerator, respectively. Suppose Assumptions 5.1, 5.2, 5.6 and 5.7 hold. Then, for a fixed input sequence,

$$\lim_{h \rightarrow 0} \text{cond}(\mathbb{E}\{\hat{\varphi}_f(t_k, \boldsymbol{\theta}_j)\varphi_f^\top(t_k, \boldsymbol{\theta}_j)\}) \rightarrow \infty, \quad (5.4)$$

where $\text{cond}(\cdot)$ denotes the condition number of a matrix.

Proof. From Section 4.3, we recall the following description for the regressor vector of the SRIVC method:

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

where

$$\begin{aligned} \tilde{\varphi}_f(t_k, \boldsymbol{\theta}_j) &= \left[\frac{-pB^*(p)}{A_j(p)A^*(p)}u(t_k), \dots, \frac{-p^nB^*(p)}{A_j(p)A^*(p)}u(t_k), \frac{1}{A_j(p)}u(t_k), \dots, \frac{p^m}{A_j(p)}u(t_k) \right]^\top, \\ \mathbf{v}(t_k) &= \left[\frac{p}{A_j(p)}v(t_k), \dots, \frac{p^n}{A_j(p)}v(t_k), 0, \dots, 0 \right]^\top, \end{aligned}$$

and the entries of $\Delta(t_k, \boldsymbol{\theta}_j)$ are given by

$$\Delta_i(t_k, \boldsymbol{\theta}_j) = \begin{cases} \frac{p^iB^*(p)}{A_j(p)A^*(p)}u(t_k) - \frac{p^i}{A_j(p)}\left\{\frac{B^*(p)}{A^*(p)}u(t)\right\}_{t=t_k} & \text{if } i = 1, \dots, n, \\ 0 & \text{if } i = n+1, \dots, n+m+1. \end{cases}$$

Since $\max(n + m^*, m + n^*) = n + m - n_{op}$, we can write

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

where (recall $\mathbf{u}_{n+m}(t_k)$ in Chapter 4)

$$\mathbf{u}_{n+m-n_{op}}(t_k) := [p^{n+m-n_{op}}, p^{n+m-n_{op}-1}, \dots, 1]^\top u(t_k), \quad (5.5)$$

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

$$\mathbf{S}(-B^*, A^*) = \left[\begin{array}{cccccc} 0 & -b_{m^*}^* & -b_{m^*-1}^* & \dots & -b_0^* & 0 \\ \ddots & \ddots & & \ddots & \ddots & \\ -b_{m^*}^* & -b_{m^*-1}^* & \dots & -b_0^* & 0 & 0 \\ \hline 0 & a_{n^*}^* & a_{n^*-1}^* & \dots & a_1^* & 1 \\ \ddots & \ddots & & \ddots & \ddots & \\ a_{n^*}^* & a_{n^*-1}^* & \dots & a_1^* & 1 & 0 \end{array} \right] \left. \begin{array}{l} n \\ \text{rows} \\ m+1 \\ \text{rows,} \end{array} \right\} \quad (5.6)$$

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.

By noting that the vector of filtered noise $\mathbf{v}(t_k)$ is uncorrelated with the filtered input signal present in $\hat{\varphi}_f(t_k, \boldsymbol{\theta}_j)$, we can write the modified normal matrix in (5.4) as

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

Note that $\mathbf{S}(-B^*, A^*)$ has rank equal to $n + m + 1 - n_{op}$ (see Lemma 5.1 of the appendix), which is strictly less than $n + m + 1$ by Assumption 5.6. Furthermore, each entry of $\Delta(t_k, \boldsymbol{\theta}_j)$ 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 in between and the output of the same system without the sampler. Thus, the equivalent system tends to zero when $h \rightarrow 0$ and therefore $\Delta(t_k, \boldsymbol{\theta}_j) \rightarrow \mathbf{0}$ for any bounded input signal. These observations imply that the matrix in (5.4) will converge to a singular matrix as h tends to zero, and thus the smallest singular value of (5.4) tends to zero for a decreasing sampling period.

Finally, to show that the condition number tends to infinity, we need to show that the largest singular value of the matrix in (5.4) does not tend to zero as $h \rightarrow 0$.

By Theorem 5.6.2 of [115], we find that

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

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. \square

5.3.2 Asymptotic analysis of the stationary points of SRIVC

Proposition 5.1 suggests that, for large sample sizes, the iterations of the SRIVC estimator become more ill-conditioned for smaller sampling periods if an over-parameterized model is considered. If the matrix being inverted in (2.35) becomes singular for a finite N , one SRIVC estimate can be computed as a solution of the system of equations

$$\sum_{k=1}^N \hat{\varphi}_f(t_k, \boldsymbol{\theta}_j) \varphi_f^\top(t_k, \boldsymbol{\theta}_j) \boldsymbol{\theta}_{j+1} = \sum_{k=1}^N \hat{\varphi}_f(t_k, \boldsymbol{\theta}_j) y_f(t_k, \boldsymbol{\theta}_j). \quad (5.7)$$

Since the matrix on the left hand side of (5.7) 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 $\bar{\boldsymbol{\theta}}^N$ that satisfy

$$\sum_{k=1}^N \hat{\varphi}_f(t_k, \bar{\boldsymbol{\theta}}^N) \varphi_f^\top(t_k, \bar{\boldsymbol{\theta}}^N) \bar{\boldsymbol{\theta}}^N = \sum_{k=1}^N \hat{\varphi}_f(t_k, \bar{\boldsymbol{\theta}}^N) y_f(t_k, \bar{\boldsymbol{\theta}}^N). \quad (5.8)$$

The asymptotic behavior of all stationary points is studied next.

Theorem 5.1. *Consider any stationary point of the SRIVC algorithm with iterates that are obtained by solving (5.7) until convergence. As the sample size tends to infinity and under Assumptions 5.1 to 5.7, the stationary points of the SRIVC method generically (with respect to the parameter space of denominator polynomials of degree n) describe a continuous-time model with polynomials*

$$\bar{A}(p) = L(p) A^*(p) \text{ and } \bar{B}(p) = L(p) B^*(p)$$

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

Proof. Take $\bar{\boldsymbol{\theta}}^N$ as a stationary point of the SRIVC algorithm, which must belong to the solution set of (5.8). This vector forms the transfer function polynomials $\bar{A}_N(p)$

and $\bar{B}_N(p)$, which are written as $\bar{A}_N(p) := \hat{A}_N(p)L_N(p)$ and $\bar{B}_N(p) := \hat{B}_N(p)L_N(p)$, where $\hat{A}_N(p)$ and $\hat{B}_N(p)$ are coprime (i.e., they form the minimal realization of $\bar{B}_N(p)/\bar{A}_N(p)$), with $\hat{A}_N(p)$ being anti-monic. The polynomial $L_N(p)$ is monic of degree equal to the number of pole-zero cancellations of $\bar{B}_N(p)/\bar{A}_N(p)$. Note that this factorization can be done without loss of generality, since $L_N(p) \equiv 1$ if $\bar{A}_N(p)$ and $\bar{B}_N(p)$ are coprime. Since

$$y_f(t_k) - \varphi_f^\top(t_k, \bar{\theta}^N)\bar{\theta}^N = y(t_k) - \frac{\bar{B}_N(p)}{\bar{A}_N(p)}u(t_k),$$

the SRIVC estimate must satisfy

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

As the number of samples tends to infinity, under stationarity conditions and independence between noise and input signals in Assumption 5.2, we obtain that

$$\mathbb{E} \left\{ \hat{\varphi}_f(t_k, \bar{\theta}) \left[\left(\frac{B^*(p)}{A^*(p)} - \frac{\hat{B}(p)}{\hat{A}(p)} \right) u(t_k) \right] \right\} = \mathbf{0}, \quad (5.9)$$

where $\hat{B}(p)/\hat{A}(p)$ is the minimal realization of the transfer function described by $\bar{\theta}$, i.e., the polynomials satisfy $\bar{A}(p) = \hat{A}(p)L(p)$ and $\bar{B}(p) = \hat{B}(p)L(p)$ for some monic polynomial $L(p)$ of degree $n_l \geq 0$, and $\hat{A}(p), \hat{B}(p)$ are coprime. Now, the filtered instrument vector is given by

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

where $\mathbf{u}_{n+m-n_l}(t_k)$ and $\mathbf{S}(-\hat{B}, \hat{A})$ are defined in a similar fashion to (5.5) and (5.6), respectively. By the same argument presented in Lemma 5.1 in the appendix of this chapter, we can show that $\text{rank}\{\mathbf{S}(-\hat{B}, \hat{A})\} = n + m - n_l + 1$. So, by (5.9), for every $i = 0, 1, \dots, n + m - n_l$ we must have

$$\mathbb{E} \left\{ \frac{p^i}{L(p)\hat{A}^2(p)}u(t_k) \left[\left(\frac{B^*(p)}{A^*(p)} - \frac{\hat{B}(p)}{\hat{A}(p)} \right) u(t_k) \right] \right\} = 0. \quad (5.10)$$

Thanks 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

$$\mathbb{E} \left\{ \frac{p^i}{\hat{A}^2(p)} u(t_k) \left[\left(\frac{B^*(p)}{A^*(p)} - \frac{\hat{B}(p)}{\hat{A}(p)} \right) u(t_k) \right] \right\} = 0 \quad (5.11)$$

for $i = 0, 1, \dots, n + m - 2n_l$. We have reached a set of conditions that does not depend on $L(p)$ but 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 $H(p) := \hat{A}(p)B^*(p) - \hat{B}(p)A^*(p) = h_r p^r + h_{r-1} p^{r-1} + \dots + h_0$, where

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

With this, (5.11) can also be written as

$$\mathbb{E} \left\{ \frac{1}{\hat{A}^2(p)} \mathbf{u}_{n+m-2n_l}(t_k) \frac{1}{A^*(p)\hat{A}(p)} \mathbf{u}_{n+m-n_{op}-n_l}^\top(t_k) \right\} \mathbf{h} = \mathbf{0}, \quad (5.12)$$

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 (5.12) as Φ . The next step is to prove that, generically,

$$\text{rank } \{\Phi\} = n + m - n_{op} - n_l. \quad (5.13)$$

For this we require $n_l \leq n_{op}$, which is true by Assumption 5.7. The rank condition in (5.13) is proven in Lemma 5.2 in the appendix of this chapter. Thus, (5.12) implies $\mathbf{h} = \mathbf{0}$, which is in turn equivalent to $\hat{A}(p)B^*(p) - \hat{B}(p)A^*(p) = 0$, that is,

$$\frac{\hat{B}(p)}{\hat{A}(p)} = \frac{B^*(p)}{A^*(p)}. \quad (5.14)$$

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 (5.14) is satisfied, we must have $n_l = n_{op}$, and the coefficients related to the possible excess of degree in numerator or denominator of $\hat{B}(p)/\hat{A}(p)$ go to zero.

Regarding the polynomial $L(p)$, we see that the conditions in (5.10) are satisfied irrespective of the coefficients that form $L(p)$, due to (5.14). Thus, $L(p)$ is left unspecified, and with this we conclude the proof. \square

Theorem 5.1 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-parametrization of both numerator and denominator polynomials. 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 stipulated in line 10 of Algorithm 2.1. 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 10 of Algorithm 2.1 changes the phase of the transfer function estimate. One solution to this problem has been proposed in [103], which involves constraining the SRIVC estimate within ellipsoids that yield stable estimates.

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

We now study the iterations of the SRIVC method for large sample size. In Theorem 5.2 we show that one step of the SRIVC algorithm also leads to the true transfer function together with pole-zero cancellations for small sampling periods.

Theorem 5.2. *Consider the SRIVC iterations given by (2.35), and assume that Assumptions 5.1 to 5.7 hold. The limit estimate (as $N \rightarrow \infty$) of the iterate $\boldsymbol{\theta}_{j+1}$ generically (with respect to the parameter space of denominator polynomials of degree n) 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. 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 (5.7) can be rewritten, as the number of samples tends to infinity, as

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

Given that the input is assumed uncorrelated with the measurement noise, we have

$$\mathbb{E} \left\{ \hat{\boldsymbol{\varphi}}_f(t_k, \boldsymbol{\theta}_j) \left(\frac{A_{j+1}(p)}{A_j(p)} \left\{ \frac{B^*(p)}{A^*(p)}u(t) \right\}_{t=t_k} - \frac{B_{j+1}(p)}{A_j(p)}u(t_k) \right) \right\} = \mathbf{0}. \quad (5.15)$$

Note that, in terms of computing (5.15), the transfer functions $A_{j+1}(p)/A_j(p)$ and $B^*(p)/A^*(p)$ are discretized separately. By defining

$$\varepsilon_u(t_k) = \frac{A_{j+1}(p)}{A_j(p)} \left\{ \frac{B^*(p)}{A^*(p)}u(t) \right\}_{t=t_k} - \left(\frac{A_{j+1}(p)B^*(p)}{A_j(p)A^*(p)} \right) u(t_k), \quad (5.16)$$

the condition in (5.15) can be expressed as

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

Note that in general we have $\varepsilon_u(t_k) \neq 0$. However, similar to the case of $\Delta_i(t_k, \theta_j)$ in the proof of Proposition 5.1, as the sampling period goes to zero we have $\varepsilon_u(t_k) \rightarrow 0$ for any bounded input signal.

We now need an alternative expression for the filtered instrument vector $\hat{\varphi}_f(t_k, \theta_j)$. Similar to the proof of Theorem 5.1, 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 $n_l \geq 0$. In other words, the minimal realization of $B_j(p)/A_j(p)$ is denoted by $\hat{B}_j(p)/\hat{A}_j(p)$. Thus,

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

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

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

where $r = \max(n+m^*, m+n^*) = n+m-n_{op}$. Thus, as the sample size tends to zero, the condition in (5.15) is equivalent to

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

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 5.2, we must have $\mathbf{h} = \mathbf{0}$. Following the same steps as in the proof of Lemma 5.1, we see that $\mathbf{h} = \mathbf{0}$ is equivalent to $A_{j+1}(p) = L(p)A^*(p)$ and $B_{j+1}(p) = L(p)B^*(p)$, where $L(p)$ is an arbitrary polynomial of degree n_{op} . \square

Remark 5.1. Theorem 5.2 implies that, even if the model were to describe the true system (with the extra numerator and denominator terms being zero), there is no guarantee of convergence (in iterations) of the polynomials $B_j(p)$ and $A_j(p)$ separately. That is, for over-parameterized models satisfying Assumption 5.6, 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 over the stability of $L(p)$, indicating that the iterations may return unstable estimates in many occasions. In practice, the problem of unstable iterates 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 regressor.

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 15 of Algorithm 2.1 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, another stopping criterion might be more useful, such as

$$\frac{\left\| F(p) \left(\frac{B_{j+1}(p)}{A_{j+1}(p)} - \frac{B_j(p)}{A_j(p)} \right) \right\|_2}{\left\| F(p) \frac{B_j(p)}{A_j(p)} \right\|_2} < \epsilon, \quad (5.17)$$

where $F(p)$ is a user-defined filter that has a high gain on 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.

5.3.4 Young's 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 for model order selection [250].

Two widely used criteria for model selection are the Akaike and Bayesian information criteria (AIC [4] and BIC [197], respectively). A natural idea for instrumental variable methods is to use the information of the approximate covariance matrix of the parametric estimation errors, which is indirectly estimated in the instrumental variable procedure. This approach has led to the YIC, that is defined in [250, p. 178] as

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

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 simulated output respectively, $\bar{\theta}_i$ is the i -th element of $\bar{\theta}^N$, and $\{\mathbf{P}_{ii}\}_{i=1}^{n+m+1}$ are the diagonal elements of the inverse modified normal matrix

$$\mathbf{P} = \left[\sum_{k=1}^N \hat{\varphi}_f(t_k, \bar{\theta}^N) \hat{\varphi}_f^\top(t_k, \bar{\theta}^N) \right]^{-1}. \quad (5.18)$$

The R^2 coefficient is useful for determining under-parametrization, and it usually reaches a plateau value once the model structure contains the true system [256]. On the other hand, the NEVN can distinguish over-parametrization, and thus our attention will be focused on this term.

It is argued in [256] and [236] that, if the model is over-parameterized, the matrix being inverted in (5.18) will tend to become singular and at least one diagonal element of \mathbf{P} will be of high magnitude. These works consider biproper discrete-time transfer functions, in which only the number of poles is being determined in the model order selection process. Proposition 5.1 shows that ill-conditioning will occur in the continuous-time case for arbitrary proper transfer functions, as long as both 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 4.1 that this matrix is (generically) non-singular under mild conditions as N tends to infinity. Thus, it is expected that $\{\mathbf{P}_{ii}\}_{i=1}^{n+m+1}$ will no longer reflect an over-parametrization problem in these cases. For the following analysis we will study three cases, depending on the consistency of the SRIVC estimator and on the existence of an element in $\bar{\boldsymbol{\theta}}^N$ that tends to zero as N tends to infinity. Such element can describe the coefficient estimate of the leading numerator or denominator orders which go to zero almost surely due to the generic consistency of the SRIVC estimator¹, or any parameter estimate whose true value is zero.

First, we study how the NEVN behaves when consistency is achieved and the model estimates does not have a parameter that tends to zero as N tends to infinity. Under stationarity assumptions, by the ergodic theorem [205] and the continuous-mapping theorem [230, Theorem 2.3] we have

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

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

Now, assume that consistency is achieved, and 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 Theorem 4.3, 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,

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

which implies that $N\bar{\boldsymbol{\theta}}_i^2$ is asymptotically gamma-distributed [170] 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 $1/N$ for large sample size.

¹Recall that this effect does not necessarily occur when both numerator and denominator are over-parameterized, as proven in Theorem 5.1.

Finally, consider the case when $\bar{\theta}^N$ is no longer consistent, which in this study occurs when the model is over-parameterized in both numerator and denominator polynomials. The parameters are not expected to converge, and neither does its normalized approximate covariance matrix \mathbf{P} . In light of the above, we expect the NEVN to highly depend on the iteration in which the SRIVC algorithm terminated, and consequently, we do not expect the NEVN to decay 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 this behavior will not be observed for an excess of numerator and/or denominator order. 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 model is exactly parameterized.

5.4 Simulation studies

The theoretical results provided in the previous section are here verified through simulations. We start by exemplifying the problems associated with over-parametrization, and afterwards we show how the iterations of SRIVC 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 and is uncorrelated with the input. The model structure we consider has one zero and two poles, that is, it is of the form

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

5.4.1 Condition number for small sampling periods

We test the condition number of the modified normal matrix in (5.4) given by one iteration of the SRIVC method for five different sampling periods $h = 10^{-i}[\text{s}]$, with $i = 2, 3, \dots, 6$. The time length of each experiment is kept constant ($20[\text{s}]$), 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 5.1. 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 5.1, and it confirms that the SRIVC estimator is particularly poorly conditioned when the sampling period is small.

Table 5.1: Mean condition number of the modified normal matrix for different sampling periods.

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

5.4.2 Difficulties associated with over-parametrization: a case study

We consider one run with $N = 200000$ input and output data points and sampling period $h = 0.01$ [s]. The tolerance factor ϵ in the SRIVC algorithm is set to 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 [242] with cut-off frequency λ of the prefilter equal to 10[rad/s].

We plot the poles, zero, and static gain of the models provided by each iteration of the SRIVC estimate in Figure 5.1, with the poles being plotted before the stabilization step 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 10 of Algorithm 2.1. Figure 5.1 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 5.3.3. 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.

The convergence of the models estimated through SRIVC is tested in Figure 5.2 under three different stopping criteria: the θ -stopping criterion in line 15 of Algorithm 2.1, the same criterion but with prior `minreal` pole-zero cancellation of the associated transfer functions with state-elimination tolerance set to 10^{-3} , and the metric in (5.17). The last metric is computed with the `norm` command in MATLAB, and a stabilizing all-pass filter of unit norm is incorporated in case the transfer function $(B_{j+1}(p)/A_{j+1}(p) - B_j(p)/A_j(p))$ 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]:

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

From Figure 5.2, we see that the estimates do not converge according to the θ -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 (5.17) quickly drops two orders of magnitude in the first iterations and later 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.

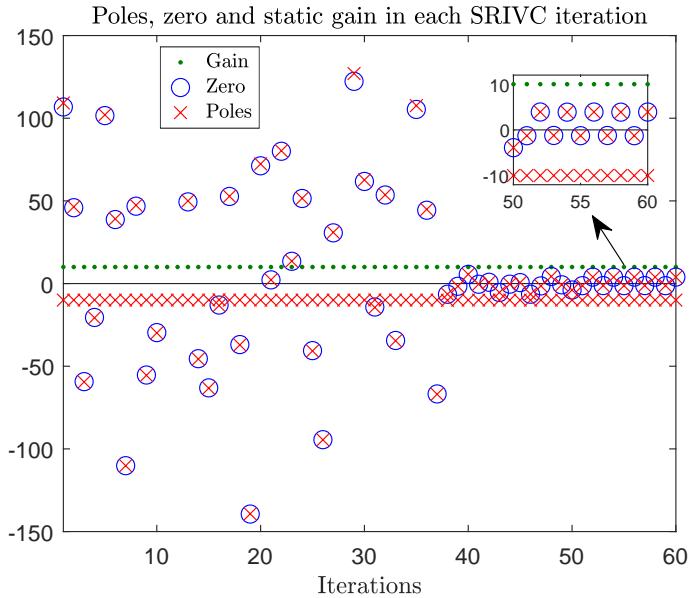


Figure 5.1: Poles (red crosses), zero (blue circle) and static gain (green dashed line) of the estimates given in each iteration of the SRIVC algorithm.

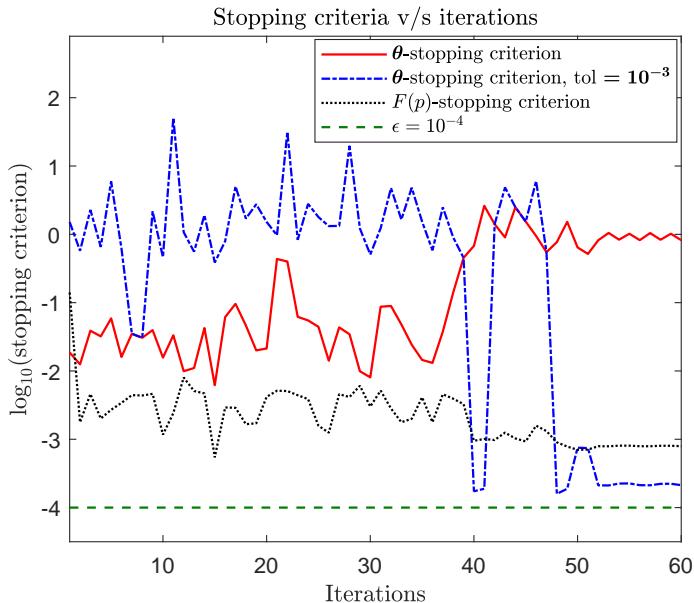


Figure 5.2: θ -based and transfer function-based stopping criteria metrics evaluated at each iteration, compared to a baseline $\epsilon = 10^{-4}$.

5.4.3 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 $N = 200000$ input and output data points, we obtain an initial estimate of the parameter vector $\boldsymbol{\theta} = [a_1, a_2, b_0, b_1]^\top$ by the LSSVF method with the cut-off frequency being $\lambda = 10[\text{rad/s}]$. Afterwards, one iteration of the SRIVC algorithm is performed without forcing stability on the estimate. Three sampling periods are tested: $h = 0.005[\text{s}]$, $h = 0.02[\text{s}]$ and $h = 0.05[\text{s}]$. Note that according to the sampling criterion given in [16], the sampling period should be chosen between $0.01[\text{s}]$ and $0.025[\text{s}]$.

We plot the zero and poles of eight estimates corresponding to eight different noise realizations in Figure 5.3. Pole-zero quasi-cancellations can be more clearly seen for the smaller sampling periods than for the larger ones. This effect is correctly predicted by Theorem 5.2 and the analysis of $\varepsilon_u(t_k)$ in (5.16). Note that the quasi-cancellations occur irregularly in the real line, which suggests that the location of them is noise-dependent (and also dependent on the initial condition) and therefore non-predictable, which aligns with our theoretical findings. We also plot the Bode diagrams of each estimated model in Figure 5.4. 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. 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 sampling period is too small [8].

5.4.4 Behavior of the NEVN

We now verify the properties of the NEVN for different sample sizes and model structures through a series of simulation tests. These 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 $h = 0.1[\text{s}]$, and an additive white noise of variance 0.1 contaminates the noiseless output. Three different model order structures are tested: $n = 2, m = 0$ (exact parametrization), $n = 2, m = 1$ (over-parametrization of the numerator) and $n = 3, m = 1$ (over-parametrization of both numerator and denominator). Two hundred Monte Carlo runs are performed for six different sample sizes.

In Figure 5.5, we have plotted violin plots [113] for the values of $\log(\text{NEVN})$ 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. It should be noted that we have chosen to plot densities of $\log(\text{NEVN})$ instead of NEVN in order to cover the wide

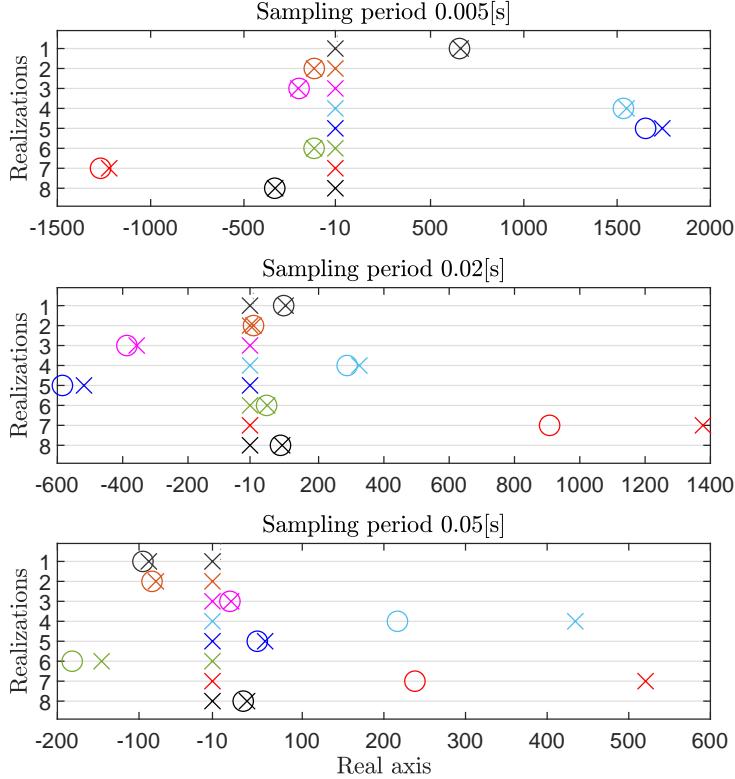


Figure 5.3: Pole-zero plots of 8 estimates (separated through the vertical axis) of $G^*(p)$ for sampling periods $h = 0.005[\text{s}]$, $h = 0.02[\text{s}]$ and $h = 0.05[\text{s}]$. All the estimated poles are real numbers in this study.

range of values obtained in the simulation test. We see in Figure 5.5 that the distributions of $\log(\text{NEVN})$ vary considerably among the three cases studied. For the exact parametrization, $\log(\text{NEVN})$ decays as $\log(1/N)$ and the data points decrease in variance as N grows. These findings support the analysis made in Section 5.3.4. For the over-parameterized numerator case, $\log(\text{NEVN})$ no longer decays for larger N . Moreover, the distribution of $\log(\text{NEVN})$ seems to converge as N grows, which is also in line with Section 5.3.4. In contrast to the previous scenarios, the over-parametrization 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 $\bar{\theta}^N$ that has large or small parameters, which may, respectively, reduce or increase the NEVN by a considerable amount.

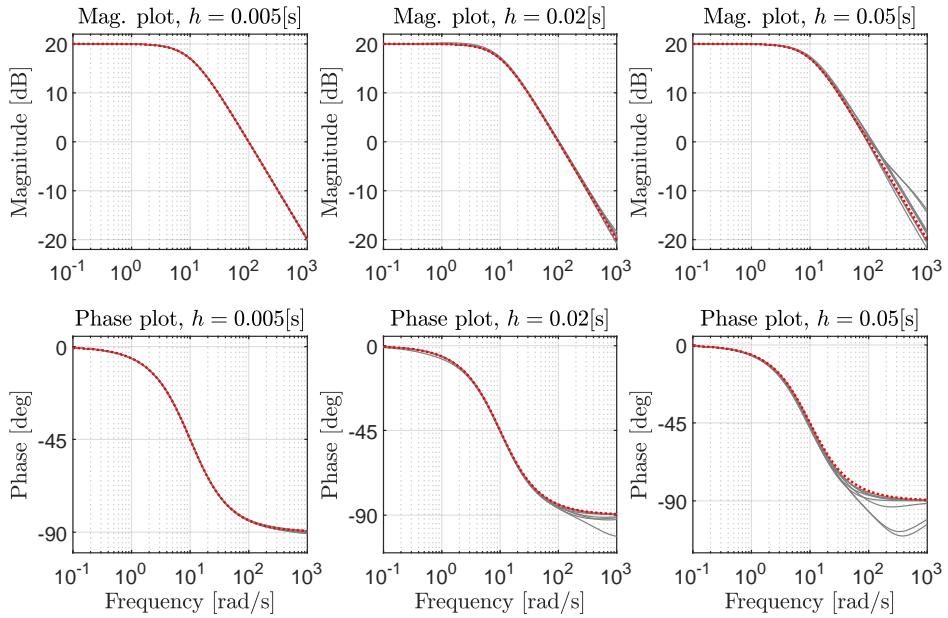


Figure 5.4: Bode plots for 8 estimates of $G^*(p)$ in grey, for sampling periods $h = 0.005[\text{s}]$ (left), $h = 0.02[\text{s}]$ (center) and $h = 0.05[\text{s}]$ (right plots). In dashed red, the true system.

5.5 Conclusions

In this chapter we have studied the properties of the SRIVC estimator when the model is over-parameterized. 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 point 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-parametrization 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.

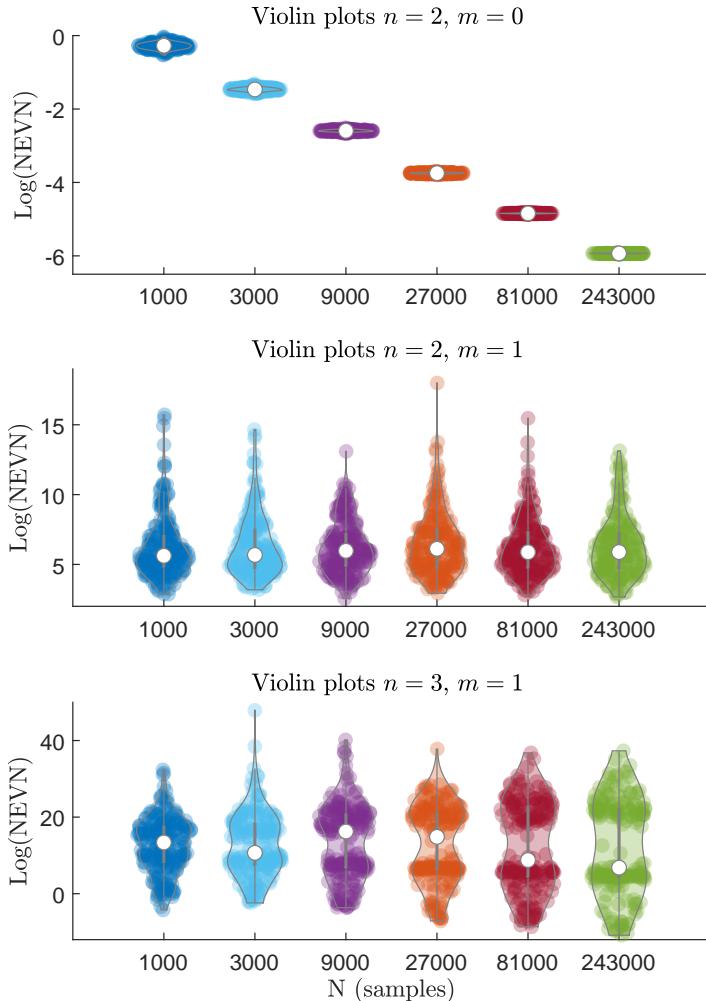


Figure 5.5: Violin plots of $\log(\text{NEVN})$ for different sample sizes. From top to bottom: exact model structure, numerator over-parametrization, and both numerator and denominator over-parametrization.

Appendix

5.A Supplementary material

Lemma 5.1 (Rectangular Sylvester matrices [209]). *Under Assumption 5.6, consider the rectangular Sylvester-type matrix $\mathbf{S}(-B^*, A^*)$ in (5.6) 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 $\mathbf{x} = [\alpha_1, \dots, \alpha_n, \beta_0, \dots, \beta_m]^\top$. The equation $\mathbf{x}^\top \mathbf{S}(-B^*, A^*) = \mathbf{0}$ can be written as an equality between polynomials:

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

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

$$\alpha_1 p + \dots + \alpha_n p^n = L(p) A^*(p), \quad \text{and} \quad \beta_0 + \dots + \beta_m p^m = L(p) B^*(p),$$

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 [115], this dimension is equal to $n+m+1-\text{Rank}\{\mathbf{S}(-B^*, A^*)\}$, from which the lemma follows. \square

Lemma 5.2 (Rank of Φ). *Consider the matrix*

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

and assume that Assumptions 5.1, 5.3 and 5.4 hold. Then, the statement $\{\text{rank}\{\Phi\} = n + m - n_{op} - n_l\}$ is generically true with respect to the parameters forming $\hat{A}(p)$.

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

$$\Phi_1 := \mathbb{E} \left\{ \frac{1}{\hat{A}^2(p)} \mathbf{u}_{n+m-n_{op}-n_l}(t_k) \frac{1}{A^*(p)\hat{A}(p)} \mathbf{u}_{n+m-n_{op}-n_l}^\top(t_k) \right\}.$$

By Lemma 4.2, it can be shown that the matrix Φ_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 (5.2) is used at this stage for the properness of the transfer functions in Φ_1 when $\hat{A}(p) = A^*(p)$, as the minimum relative degree of such transfer functions is given by

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

Furthermore, it has been shown in Lemma 4.4 that the entries of Φ_1 are analytic functions of the denominator parameters. Hence, by Lemma A2.3 of [210] and its corollary, we conclude that Φ_1 is generically non-singular with respect to the parameters forming $\hat{A}(p)$. Since the columns of Φ_1 are generically linearly independent, it is generically true that the extended matrix Φ has rank $n + m - n_{op} - n_l$. \square

Chapter 6

The SRIVC estimator for arbitrary continuous-time inputs

The asymptotic properties of the SRIVC estimator were discussed in detail in Chapters 4 and 5. In particular, the insights gained on its consistency can lead to formalize possible extensions of the SRIVC estimator when full knowledge of the intersample behavior is known in advance. In this chapter we prove the generic consistency of an extended SRIVC estimator based on such knowledge, study its connections to the maximum likelihood estimator, analyze its iterations, and discuss some implementation aspects.

6.1 Introduction

Despite decades of practical experience, the SRIVC estimator is still not fully understood. In Chapter 4, we have proven the generic consistency of this estimator for inputs that can be exactly reconstructed with zero or first order holds and stated conditions on the intersample behavior of the input and output in the instrument and regressor vectors that ensure this property. Moreover, in that same chapter, the asymptotic efficiency of this estimator has been proven for the case when the intersample behavior of the input signal is correctly specified in the SRIVC iterations. These two results provide theoretical support to the findings in [252], where it is stated that the refined instrumental variable estimator from which the SRIVC method is derived uses the optimal instruments in its recursive algorithm, and therefore it minimizes the prediction error and maximizes the likelihood function for a Gaussian noise distribution.

One of the shortcomings of the work in [252] is that, due to its generality, the role of the intersample behavior in the continuous-time system identification algorithm is not analyzed in detail. The behavior of the signals between samples is of particular interest in continuous-time identification due to the inherent continuous-time filtering of discrete-time signals; in particular, the measured input needs to be interpolated in some manner in order to simulate a continuous-time model output.

This interpolation must be carefully taken into account, as it has been discussed in [6] and [196] that assuming the incorrect input intersample behavior of the underlying data generating process may lead to severe estimation errors. A simple motivational example was provided in [196] to illustrate the modeling error induced in the estimation process when a band-limited input is assumed while the true system input is reconstructed via a zero-order hold. The takeaway from this example is that it is important to take into account the intersample behavior of the signals when dealing with continuous-time estimators, which has been somewhat overlooked in the existing literature of continuous-time system identification algorithms. We note, however, that there are some continuous-time identification methods that do not require the sampled signals to be interpolated. For example, higher order Padé approximation is used in [39] to approximate a discrete-time filter that produces the same output as the sampled continuous-time filter output, which avoids the need to reconstruct the continuous-time input signal, and continuous-time models are identified in [148] based on second-order statistics.

The user has several choices regarding the intersample behavior assumptions in the SRIVC method. More precisely, the intersample behavior of the input must be specified in both filtered regressors and instruments, and the output signal must be reconstructed in the filtered regressor and output. Usually the output is selected to have an FOH behavior, since it is argued that it typically gives rise to a satisfactory approximation if the sampling period is small [34]. It has been noted in [62] that the SRIVC estimator requires optimal interpolation of the input signal in order for the estimator to be truly optimal, although simple interpolation (i.e., ZOH or FOH) is commonly used for the SRIVC implementation [60]. The implications of this statement have only been studied formally in [164] and [166], which constitute Chapter 4 of this thesis. To the best of the author's knowledge, extensions of the statistical properties of the SRIVC estimator for inputs that are not perfectly interpolated with ZOH and FOH devices have not been considered in the literature.

In this chapter we incorporate arbitrary intersample behaviors in the SRIVC framework and analyze the properties of this extended SRIVC estimator. Focus will be put on the consistency of this estimator for continuous-time multisine inputs, and the effects of the intersample behavior of the output on the proposed estimator will be analyzed. In summary,

- we present an extension of the SRIVC estimator that admits an arbitrary intersample behavior for the input signal, and we show that it is generically consistent for continuous-time multisine inputs;
- we prove that, given knowledge of the continuous-time multisine input signal and measured output samples, the exact computation of the input regressors is necessary and sufficient for a generically consistent estimate of the continuous-time system;
- we establish the relationship between this estimator and the maximum likeli-

hood estimator that considers *full intersample knowledge* of the input of the true system;

- we interpret the proposed estimator as a quasi Gauss-Newton step for optimizing the likelihood function, study the effect of the hold device for reconstructing the output measurements, and derive ways to robustify the extended SRIVC method;
- we study computational aspects regarding the continuous-time filtering procedure in the extended SRIVC method; and
- we exemplify the properties we have derived through extensive Monte Carlo simulations.

The remainder of this chapter is organized as follows. The identification problem is formulated in Section 6.2. In Section 6.3 the extended SRIVC method is presented and its generic consistency is proven for multisine inputs. The relationship between this method and the maximum likelihood estimator, and the analysis of its iterations, are studied in Section 6.4. In Section 6.5 we describe computational aspects of the proposed method, while Section 6.6 illustrates the extended SRIVC method and its properties with extensive numerical examples. Conclusions are drawn in Section 6.7, and supplementary material can be found in the appendix related to this chapter.

6.2 System and model setup

Consider an LTI, causal, asymptotically stable, SISO, continuous-time system

$$x(t) = \frac{B^*(p)}{A^*(p)} u(t), \quad (6.1)$$

where the numerator and denominator polynomials are coprime and given by

$$\begin{aligned} B^*(p) &= b_{m^*}^* p^{m^*} + b_{m^*-1}^* p^{m^*-1} + \cdots + b_0^*, \\ A^*(p) &= a_{n^*}^* p^{n^*} + a_{n^*-1}^* p^{n^*-1} + \cdots + a_1^* p + 1. \end{aligned}$$

Suppose that the continuous-time input $u(t)$ is known from $t = t_1$ to $t = t_N$, and that N noisy measurements of the output $x(t)$ are obtained at the time instants $\{t_k\}_{k=1}^N$. In other words, the output observations are given by

$$y(t_k) = x(t_k) + v(t_k), \quad k = 1, \dots, N, \quad (6.2)$$

where it is assumed that the sampled noise sequence $\{v(t_k)\}$ can be described as a zero-mean and finite-variance random process. To identify the system, we propose the model structure

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

The goal is to obtain an adequate model of the continuous-time system given the knowledge of the output measurements with N samples and the (non-sampled) continuous-time input signal. Alternatively, we may also assume that samples of the input signal are obtained and the intersample behavior of such signal is arbitrary but exactly known. Note that the input signal is not limited to hold reconstructions in this framework; hence, the description includes the standard framework where $u(t)$ is assumed to be obtained via a ZOH or FOH, and allows for more general inputs such as multisines and band-limited signals [195].

The identification of the system $G^*(p)$ can be done by obtaining the data points $\{u(t_k), y(t_k)\}$ and applying a method for continuous-time system identification such as in [30], or as in [83, 151, 255] for regularly sampled data. In most of these algorithms, however, hold reconstructions of the input and output signals must be specified. In particular, the SRIVC algorithm uses interpolation of the input and output data to compute the filtered regressor and instrument vectors in its iterative estimation procedure. This reconstruction is usually done by assuming a ZOH or an FOH intersample behavior of the input and output signals. For inputs that can be described exactly with these reconstruction schemes, we have shown that the SRIVC estimator is generically consistent and asymptotically efficient in Chapter 4. However, when the intersample behavior assumption on the model input does not match that of the system input, the SRIVC estimator is generically inconsistent. This intersample behavior mismatch always occurs when the input cannot be described via hold reconstructions, which is the case for band-limited signals such as multisines. For instance, the complete continuous-time input signal is known to the practitioner when it is a continuous-time multisine signal, but the SRIVC procedure only performs simple interpolations of the input that impact its consistency regardless of the sampling period.

6.3 The SRIVC method for arbitrary continuous-time inputs

As shown in Chapter 4, a correct specification of the intersample behavior of the input (ZOH or FOH) in the filtered regressor vector $\varphi_f(t_k)$ guarantees generic consistency under mild conditions. Furthermore, if the intersample behavior of the input is correctly specified in the filtered instrument vector, then the resulting estimator is asymptotically efficient. The extension of this principle, together with the proof of the underlying statistical properties, constitutes the main contribution of this chapter. Before introducing the extended SRIVC method for arbitrary inputs, we provide intuition on the bias of the SRIVC method when the intersample behavior of the input is misspecified. Although the following arguments can be made formal following the ideas of Chapter 4, we decide to focus on the interpretation of the result more than the strict derivation.

For any fixed N , as the number of iterations j tends to infinity, the converging

point $\bar{\theta}^N$ of the SRIVC estimator satisfies

$$\left[\frac{1}{N} \sum_{k=1}^N \hat{\varphi}_f(t_k, \bar{\theta}^N) \varphi_f^\top(t_k, \bar{\theta}^N) \right]^{-1} \left[\frac{1}{N} \sum_{k=1}^N \hat{\varphi}_f(t_k, \bar{\theta}^N) (y_f(t_k, \bar{\theta}^N) - \varphi_f^\top(t_k, \bar{\theta}^N) \bar{\theta}^N) \right] = \mathbf{0}. \quad (6.3)$$

Note that

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

where $\bar{A}(p)$ and $\bar{B}(p)$ are the A and B polynomials described with the coefficients in $\bar{\theta}^N$. On the other hand, if the input is uncorrelated with the noise present in $y(t_k)$, we have

$$\mathbb{E} \{ \hat{\varphi}_f(t_k, \bar{\theta}) y(t_k) \} = \mathbb{E} \left\{ \hat{\varphi}_f(t_k, \bar{\theta}) \left\{ \frac{B^*(p)}{A^*(p)} u(t) \right\}_{t=t_k} \right\},$$

where we have used the notation introduced in Remark 2.5, and $\bar{\theta} = \lim_{N \rightarrow \infty} \bar{\theta}^N$. As the sample size tends to infinity and under mild assumptions, the sums in (6.3) converge to their expected values [205], thus leading to

$$\mathbb{E} \{ \hat{\varphi}_f(t_k, \bar{\theta}) \varphi_f^\top(t_k, \bar{\theta}) \}^{-1} \mathbb{E} \left\{ \hat{\varphi}_f(t_k, \bar{\theta}) \left(\left\{ \frac{B^*(p)}{A^*(p)} u(t) \right\}_{t=t_k} - \frac{\bar{B}(p)}{\bar{A}(p)} u(t_k) \right) \right\} = \mathbf{0}.$$

Provided some identifiability conditions concerning the model structure and persistence of excitation are met, we can show that $\mathbb{E}\{\hat{\varphi}_f(t_k, \bar{\theta}) \varphi_f^\top(t_k, \bar{\theta})\}$ is generically non-singular, which leads to

$$\mathbb{E} \left\{ \hat{\varphi}_f(t_k, \bar{\theta}) \left(\left\{ \frac{B^*(p)}{A^*(p)} u(t) \right\}_{t=t_k} - \frac{\bar{B}(p)}{\bar{A}(p)} u(t_k) \right) \right\} = \mathbf{0}. \quad (6.4)$$

This equation characterizes the asymptotic bias of the SRIVC estimator when the intersample behavior of the system input does not match the one used in the algorithms. From (4.16), we see that $\hat{\varphi}_f(t_k, \bar{\theta})$ can be computed as

$$\hat{\varphi}_f(t_k, \bar{\theta}) = \mathbf{S}(-\bar{B}, \bar{A}) \frac{1}{\bar{A}^2(p)} \mathbf{u}_{n+m}(t_k),$$

where $\mathbf{S}(-\bar{B}, \bar{A})$ is the Sylvester matrix formed by the polynomials $-\bar{B}(p)$ and $\bar{A}(p)$, which is non-singular when $\bar{B}(p)$ and $\bar{A}(p)$ are coprime [210, Lemma A3.1], and $\mathbf{u}_{n+m}(t_k)$ is defined in (4.15). We can also write

$$\frac{\bar{B}(p)}{\bar{A}(p)} u(t_k) - \frac{B^*(p)}{A^*(p)} u(t_k) = \frac{1}{\bar{A}(p) A^*(p)} \mathbf{u}_{n+m}^\top(t_k) \mathbf{h},$$

where \mathbf{h} is a vector that contains the coefficients of $A^*(p)\bar{B}(p) - \bar{A}(p)B^*(p)$ in descending order of degree. Thus, the condition in (6.4) is equivalent to

$$\mathbb{E} \left\{ \frac{1}{\bar{A}^2(p)} \mathbf{u}_{n+m}(t_k) \frac{1}{\bar{A}(p)A^*(p)} \mathbf{u}_{n+m}^\top(t_k) \right\} \mathbf{h} = \mathbb{E} \left\{ \frac{1}{\bar{A}^2(p)} \mathbf{u}_{n+m}(t_k) \Delta(t_k) \right\}, \quad (6.5)$$

where we have defined the error signal $\Delta(t_k)$ as

$$\Delta(t_k) := \left\{ \frac{B^*(p)}{A^*(p)} u(t) \right\}_{t=t_k} - \frac{B^*(p)}{A^*(p)} u(t_k).$$

In the case where the input is exactly reconstructed by a known hold device, we have $\Delta(t_k) = 0$ and thus $\mathbf{h} = \mathbf{0}$, since the matrix on the left hand side of (6.5) is known to be generically non-singular by the same procedure as in Theorem 4.1. So, we reach $\bar{B}(p)/\bar{A}(p) = B^*(p)/A^*(p)$, i.e., the SRIVC estimator is generically consistent. Otherwise, if the intersample behavior of the input is misspecified in the computation of the regressor, then $\Delta(t_k) \neq 0$ and the expectation on the right hand side of (6.5) is in general different from zero. This leads to $\mathbf{h} \neq \mathbf{0}$, which means that the SRIVC estimator will be asymptotically biased.

In the sequel we introduce a solution to this bias problem. We propose a method, which we call the extended SRIVC method, that computes the filtered regressors and instruments exactly for a wide class of input signals whose intersample behavior is known and we prove its consistency for multisine input excitations.

The generalized equation error for the proposed approach is (cf. (2.29))

$$\varepsilon(t_k) = A(p)y_f(t_k) - \{B(p)u_f(t)\}_{t=t_k}.$$

That is, the predicted output measurement is explicitly calculated by first computing the underlying continuous-time signal $B(p)u_f(t)$ (where $u_f(t)$ is the filtered input $A^{-1}(p)u(t)$), and later evaluating it at $t = t_k$. The extended SRIVC estimator is outlined in Algorithm 6.1. The proposed estimator follows the SRIVC procedure described in Algorithm 2.1, but the intersample behavior of the input in the regressor and instrument vectors is known and therefore used in the filtering process, instead of interpolating the input samples with a ZOH or FOH device. However, keep in mind that we do not require knowing the exact intersample behavior of the system output for our analysis and results.

Remark 6.1. The proposed estimator is equivalent to the standard SRIVC estimator for input signals that are reconstructed exactly through a ZOH or FOH device. Thus, for this selection of input signal, the extended SRIVC estimator described in Algorithm 6.1 is generically consistent under Assumptions 4.1 to 4.6 by Theorem 4.1, and it is asymptotically efficient under Assumptions 4.1–4.4, 4.6–4.7 by Theorem 4.3.

To further analyze the asymptotic properties of the proposed estimator, we first study its consistency for multisine inputs. Later, we discuss the extension for arbitrary input excitations.

Algorithm 6.1: The extended SRIVC method

- 1: Input: $\{u(t)\}_{t \in [t_1, t_N]}$, $\{y(t_k)\}_{k=1}^N$, model order (n, m) , initial vector estimate $\boldsymbol{\theta}_1$, tolerance factor ϵ and maximum number of iterations M
- 2: Using $\boldsymbol{\theta}_1$, form the estimated system polynomials $A_1(p)$ and $B_1(p)$
- 3: $j \leftarrow 1$, flag $\leftarrow 1$
- 4: **while** $\text{flag} = 1$ and $j \leq M$ **do**
- 5: Prefilter the continuous-time input $\{u(t)\}_{t \in [t_1, t_N]}$ and discrete-time output $\{y(t_k)\}_{k=1}^N$ by continuous-time filters to form

$$\boldsymbol{\varphi}_f(t_k) \leftarrow \left[\frac{-p}{A_j(p)} y(t_k), \dots, \frac{-p^n}{A_j(p)} y(t_k), \left\{ \frac{1}{A_j(p)} u(t) \right\}_{t=t_k}, \dots, \left\{ \frac{p^m}{A_j(p)} u(t) \right\}_{t=t_k} \right]^\top, \quad (6.6)$$

$$\hat{\boldsymbol{\varphi}}_f(t_k) \leftarrow \left[-\frac{pB_j(p)}{A_j^2(p)} u(t), \dots, -\frac{p^n B_j(p)}{A_j^2(p)} u(t), \frac{1}{A_j(p)} u(t), \dots, \frac{p^m}{A_j(p)} u(t) \right]_{t=t_k}^\top, \quad (6.7)$$

$$y_f(t_k) \leftarrow \frac{1}{A_j(p)} y(t_k)$$

- 6: Compute the parameter estimate

$$\boldsymbol{\theta}_{j+1} \leftarrow \left[\sum_{k=1}^N \hat{\boldsymbol{\varphi}}_f(t_k) \boldsymbol{\varphi}_f^\top(t_k) \right]^{-1} \left[\sum_{k=1}^N \hat{\boldsymbol{\varphi}}_f(t_k) y_f(t_k) \right] \quad (6.8)$$

- 7: **if** $1/A_{j+1}(p)$ is unstable **then**
- 8: **for** $i = 1, \dots, n$ **do**
- 9: **if** pole p_i of $1/A_{j+1}(p)$ is unstable **then**
- 10: $\text{Re}\{p_i\} \leftarrow -\text{Re}\{p_i\}$
- 11: **end if**
- 12: **end for**
- 13: $A_{j+1}(p) \leftarrow \prod_{i=1}^n (p - p_i) / \prod_{i=1}^n (-p_i)$
- 14: **end if**
- 15: **if** $\frac{\|\boldsymbol{\theta}_{j+1} - \boldsymbol{\theta}_j\|}{\|\boldsymbol{\theta}_j\|} < \epsilon$ **then**
- 16: flag $\leftarrow 0$
- 17: **end if**
- 18: $j \leftarrow j + 1$
- 19: **end while**
- 20: Output: $\boldsymbol{\theta}_j$ and its associated model $B_j(p)/A_j(p)$.

6.3.1 Analysis of the extended SRIVC method for multisine inputs

We consider multisine input signals of the form

$$u(t) = \alpha_0 + \sum_{l=1}^{m_u} \alpha_l \cos(\omega_l t + \psi_l), \quad (6.9)$$

where m_u , $\{\alpha_l\}_{l=0}^{m_u}$, $\{\omega_l\}_{l=1}^{m_u}$ and $\{\psi_l\}_{l=1}^{m_u}$ are input parameters. The frequencies ω_l are assumed to be positive and distinct, and without loss of generality we assume that every scalar α_l is positive as well. If this input passes through an asymptotically stable LTI filter $H(p)$, it is well known that the output at steady state is also a continuous-time multisine signal, given by

$$y(t) = H(0)\alpha_0 + \sum_{l=1}^{m_u} \alpha_l |H(i\omega_l)| \cos(\omega_l t + \psi_l + \angle H(i\omega_l)). \quad (6.10)$$

This property of LTI systems provides a natural way to obtain exact values for the signal evaluations in (6.6) and (6.7) while reducing the computational cost, since the prefiltering is directly obtained by evaluating (6.10) with the corresponding filter instead of performing a convolution step. Another advantage of this filtering formula is that it extends naturally to non-uniformly sampled data. For such type of data, the proposed method is not as computationally intensive as the standard SRIVC method since the algorithm only requires approximations of the filtered output $p^l A_j^{-1}(p)y(t_k)$, $l = 1, \dots, n$ instead of computing approximations of the filtered values of both $u(t_k)$ and $y(t_k)$. The filtered output computations can be carried out by, e.g., an adaptive Runge-Kutta method (as in [30]), or by any oversampling technique with intersample behavior assumptions (more about this can be found in Section 6.5).

We now prove the generic consistency of the extended SRIVC estimator for a multisine input. The assumptions considered for this analysis are detailed below.

Assumption 6.1. The true system $B^*(p)/A^*(p)$ is proper ($n^* \geq m^*$) and asymptotically stable with $A^*(p)$ and $B^*(p)$ being coprime.

Assumption 6.2. The disturbance sequence $\{v(t_k)\}$ is a zero-mean stationary random process.

Assumption 6.3. The number of sinusoids of the input, m_u , satisfies $m_u \geq (n+m)/2$, and the input offset, α_0 , is different from zero.¹

Assumption 6.4. For all $j \geq 1$, all the zeros of $A_j(p)$ have strictly negative real parts, and $n \geq m$, with $A_j(p)$ and $B_j(p)$ being coprime.

¹If no offset is considered (i.e., $\alpha_0 = 0$), then at least $(n+m+1)/2$ sinusoids are required for our results.

Assumption 6.5. The degrees n and m of the polynomials in the model satisfy $\min(n - n^*, m - m^*) = 0$.

By comparing these assumptions to Assumptions 4.1 to 4.6, we find that the persistence of excitation requirement in Assumption 4.3 is now explicit in Assumption 6.3. Instead of persistence of excitation of order $2n + 1$, we only require a persistently exciting input of order $m + n + 1$. This is due to the fact that no discrete-time equivalent is computed for the input filters in this consistency proof, thus avoiding the issue pointed out in Remark 4.3. Also, since the input is deterministic, statistical independence between the input and disturbance signals is no longer a necessary condition.

Similar to the derivation before Theorem 4.1, Theorem 6.1 requires expressing the regressor vector of the extended SRIVC estimator as

$$\varphi_f(t_k) = \tilde{\varphi}_f(t_k) - \mathbf{v}(t_k) + \Delta(t_k), \quad (6.11)$$

where the noise-free, interpolation-error-free regressor vector is given by

$$\tilde{\varphi}_f(t_k) = \left[-\frac{pB^*(p)}{A_j(p)A^*(p)} u(t), \dots, -\frac{p^n B^*(p)}{A_j(p)A^*(p)} u(t), \frac{1}{A_j(p)} u(t), \dots, \frac{p^m}{A_j(p)} u(t) \right]_{t=t_k}^\top, \quad (6.12)$$

the filtered noise vector $\mathbf{v}(t_k)$ is given by (4.5), and $\Delta(t_k)$ is a vector that contains the interpolation errors that arise from constructing the filtered derivatives of the output with entries given by the difference between the noise-free version of the regressor vector and $\tilde{\varphi}_f(t_k)$ in (6.12), i.e.,

$$\Delta_i(t_k) = \begin{cases} \left\{ \frac{p^i B^*(p)}{A_j(p)A^*(p)} u(t) \right\}_{t=t_k} - \frac{p^i}{A_j(p)} \left\{ \frac{B^*(p)}{A^*(p)} u(t) \right\}_{t=t_k} & \text{if } i = 1, \dots, n, \\ 0 & \text{otherwise.} \end{cases} \quad (6.13)$$

As deterministic inputs will be considered in conjunction with stochastic noise processes, our analysis uses the standard definition of expectation for quasi-stationary signals [138, pp. 34], which is

$$\overline{\mathbb{E}}\{g(t)\} := \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N \mathbb{E}\{g(t)\}.$$

Theorem 6.1 (Generic consistency of the extended SRIVC estimator for multisine inputs). *Consider the extended SRIVC estimator described in Algorithm 6.1 with a fixed sampling period h , and suppose that Assumptions 6.1 to 6.5 hold. Then, the following statements are true for a multisine input of the form (6.9):*

1. *The matrix $\overline{\mathbb{E}}\{\hat{\varphi}_f(t_k)\varphi_f^\top(t_k)\}$ is generically non-singular with respect to the system and model denominator provided that the condition*

$$\|\overline{\mathbb{E}}\{\hat{\varphi}_f(t_k)\Delta^\top(t_k)\}\|_2 < \sigma_{\min}(\overline{\mathbb{E}}\{\hat{\varphi}_f(t_k)\tilde{\varphi}_f^\top(t_k)\}) \quad (6.14)$$

holds, where $\hat{\varphi}_f(t_k)$, $\tilde{\varphi}_f(t_k)$ and $\Delta(t_k)$ are defined as in (6.7), (6.12) and (6.13), respectively.

2. If the iterations of the extended SRIVC estimator converge for all N sufficiently large to, say, $\bar{\theta}^N$, and (6.14) is satisfied, then the true parameter θ^* is the unique converging point of $\bar{\theta}^N$ as the sample size tends to infinity.
3. Denote $\bar{A}(p)$ as the model denominator at convergence. As the sample size tends to infinity, the estimator locally converges to θ^* provided that the matrix $(\Phi_1 + \Psi_2)^{-1}\Psi_2$ has spectral radius less than 1, where

$$\Phi_1 = \bar{\mathbb{E}}\{\hat{\varphi}_f(t_k)\tilde{\varphi}_f^\top(t_k)\}$$

and

$$\Psi_2 = \bar{\mathbb{E}}\{\hat{\varphi}_f(t_k)\bar{\Delta}^\top(t_k)\},$$

with $\tilde{\varphi}_f^\top(t_k)$ and $\bar{\Delta}(t_k)$ being defined as (6.12) and (6.13) respectively, but evaluated at the converging point.

Proof. Statements 2 and 3 follow from the same arguments as in the proof of Theorem 4.1 and are therefore omitted. The proof of Statement 1 goes as follows. By exploiting the decomposition in (6.11), the modified normal matrix $\bar{\mathbb{E}}\{\hat{\varphi}_f(t_k)\varphi_f^\top(t_k)\}$ can be written as

$$\bar{\mathbb{E}}\{\hat{\varphi}_f(t_k)\varphi_f^\top(t_k)\} = \bar{\mathbb{E}}\{\hat{\varphi}_f(t_k)\tilde{\varphi}_f^\top(t_k)\} - \bar{\mathbb{E}}\{\hat{\varphi}_f(t_k)\mathbf{v}^\top(t_k)\} + \bar{\mathbb{E}}\{\hat{\varphi}_f(t_k)\Delta^\top(t_k)\}.$$

The term $\bar{\mathbb{E}}\{\hat{\varphi}_f(t_k)\mathbf{v}^\top(t_k)\}$, which is the term related to the filtered noise, is shown to be equal to zero in Lemma 6.3 of Appendix 6.A.

By the same approach as in the proof of Statement 1 of Theorem 4.1, which relies on the matrix perturbation result in Theorem 5.1 of [41], the condition in (6.14) ensures the generic non-singularity of the modified normal matrix provided that $\bar{\mathbb{E}}\{\hat{\varphi}_f(t_k)\tilde{\varphi}_f^\top(t_k)\}$ is generically non-singular. This non-singularity result is proven next.

We can write

$$\tilde{\varphi}_f(t_k) = \mathbf{S}(-B^*, A^*) \left[\frac{1}{A_j(p)A^*(p)} \mathbf{u}_{n+m}(t) \right]_{t=t_k},$$

with $\mathbf{S}(-B^*, A^*)$ being the Sylvester matrix associated with the polynomials $-B^*(p)$ and $A^*(p)$, which is non-singular since $A^*(p)$ and $B^*(p)$ are coprime by Assumption 6.1. Similarly,

$$\hat{\varphi}_f(t_k) = \mathbf{S}(-B_j, A_j) \left[\frac{1}{A_j^2(p)} \mathbf{u}_{n+m}(t) \right]_{t=t_k},$$

where $\mathbf{S}(-B^*, A^*)$ is non-singular due to the same analysis done in the proof of Statement 1 of Theorem 4.1, where Assumption 6.4 is used. Thus,

$$\bar{\mathbb{E}}\{\hat{\varphi}_f(t_k)\tilde{\varphi}_f^\top(t_k)\} = \mathbf{S}(-B_j, A_j)\Phi\mathbf{S}^\top(-B^*, A^*),$$

where

$$\Phi := \bar{\mathbb{E}} \left\{ \left[\frac{1}{A_j^2(p)} \mathbf{u}_{n+m}(t) \right]_{t=t_k} \left[\frac{1}{A_j(p) A^*(p)} \mathbf{u}_{n+m}(t) \right]_{t=t_k}^\top \right\}.$$

It is shown in Lemma 6.4 of Appendix 6.A that Φ is generically non-singular, which means that $\bar{\mathbb{E}}\{\hat{\varphi}_f(t_k)\hat{\varphi}_f^\top(t_k)\}$ is generically non-singular. This concludes the proof of Statement 1. \square

Note that if the commonly used ZOH (or FOH) were chosen as the intersample behavior of the signals when discretizing the prefilters, the reconstruction of $u(t)$ would suffer from high-frequency distortion. Such distortion would lead to inaccuracies in the computation of $\varphi_f(t_k)$ and $\hat{\varphi}_f(t_k)$, thus compromising the performance of the method. As stated next, an inaccurate computation of the regressor vector $\varphi_f(t_k)$ causes generic inconsistency of the proposed method under continuous-time multisine input excitation, whereas the generic consistency is unaffected by inaccurate computations of the instrument vector $\hat{\varphi}_f(t_k)$.

Corollary 6.1. *Assume that the intersample behavior in the input signal is misspecified in the extended SRIVC method, but the reconstructed input nevertheless satisfies $G(p)u(t_k) = \{G(p)u(t)\}_{t=t_k}$ as $h \rightarrow 0$. The extended SRIVC estimator*

1. *remains generically consistent if a misspecification of the intersample behavior is used for generating the filtered signals in the instrument vector $\hat{\varphi}_f(t_k)$, and*
2. *is generically not consistent if a misspecification of the intersample behavior is used for filtering the input signal in the regressor vector $\varphi_f(t_k)$.*

Proof. Part 1 of this corollary follows from the same logic as in the proof of Corollary 4.2 and therefore it is omitted. Regarding Part 2, Statement 1 of Theorem 6.1 still holds by following the same steps as before, but this time the vector $\Delta(t_k)$ will have non-zero elements in its bottom $m+1$ entries. Namely, the i -th component of $\Delta(t_k)$, with $i = n+1, \dots, n+m+1$, is now

$$\Delta_i(t_k) = \frac{p^{i-n-1}}{A_j(p)} u(t_k) - \left\{ \frac{p^{i-n-1}}{A_j(p)} u(t) \right\}_{t=t_k},$$

which still satisfies $\Delta_i(t_k) \rightarrow 0$ as $h \rightarrow 0$ for $i = n+1, \dots, n+m+1$. Thus, Statement 1 of Theorem 6.1 is valid for this case as well. However, Statement 2 of Theorem 6.1 does not hold. This fact follows from the same analysis done in the proof of Statement 3 of Corollary 4.2. \square

In summary, we have shown that the extended SRIVC method provides generically consistent estimates when the input is a continuous-time multisine signal. For these input signals, this estimator overcomes the deficiency of the standard SRIVC method exposed in Corollary 4.1. Note that a similar procedure to (6.6) and (6.7) could be proposed for the computation of $y_f(t_k)$ by exploiting the fact

that the noiseless output also corresponds to a multisine (thus, a more adequate reconstruction scheme could be designed). However, provided the condition (6.14) is satisfied, the generic consistency of the SRIVC estimator does not depend on the output reconstruction mechanism. Thus, a more precise filtering of the output is usually not needed.

6.4 Efficiency and iterations analysis of the extended SRIVC method

The extended SRIVC method presented and analyzed above serves as a natural refinement of the SRIVC estimator for situations where the full intersample behavior of the input is known in advance. In this section we cover several aspects of this estimator. We first discuss how precise the extended SRIVC estimator is, by relating it to the well-known maximum likelihood estimator. Later, we characterize the influence of the intersample behavior assumption of the output measurements. Finally, we relate the extended SRIVC iterations to Gauss-Newton steps and extend its reliability for highly noisy identification scenarios.

6.4.1 Dependence on the interpolation of the output measurements and relation to maximum likelihood

The asymptotic CRLB for the output error model structure that was presented in Theorem 4.2 depends on the evaluation of filtered continuous-time signals. More precisely, it depends on the continuous-time signal $u(t)$, and it is not influenced by how the output signal $y(t_k)$ could possibly be reconstructed between samples. The CRLB suggests that asymptotically optimal estimators should correctly specify (at least asymptotically) the intersample behavior for the input, and should not depend on the intersample behavior of the output.

In the extended SRIVC method, only the output signal has an intersample behavior that can be specified by the user in the algorithm. This choice is important for constructing the regressor vector and filtered output (Equations (6.6) and (2.34), respectively). The output is selected to have an FOH behavior in the standard SRIVC method since it resembles the underlying smooth continuous-time output signal if the sampling period is small [34]. Although the iterative procedures of the SRIVC and extended SRIVC estimators depend on how the output measurements are interpolated, this effect disappears as the number of iterations tends to infinity, as seen in the following lemma.

Lemma 6.1. *Consider the iterations of the extended SRIVC method for finite N . If the iterative procedure of the extended SRIVC estimator converges to $\bar{\theta}^N$ and the matrix*

$$\frac{1}{N} \sum_{k=1}^N \dot{\varphi}_f(t_k, \bar{\theta}^N) \varphi_f^\top(t_k, \bar{\theta}^N) \quad (6.15)$$

is non-singular, then $\bar{\boldsymbol{\theta}}^N$ does not depend on the hold mechanism used on $y(t_k)$ for constructing the regressor vector and filtered output.

Proof. The proof follows the same lines as the proof of Lemma 4.5 and therefore it is omitted. \square

Note that the nonsingularity of (6.15) is not guaranteed for finite N ; however, the expression in (6.15) converges to the expectation $\mathbb{E}\{\hat{\varphi}_f(t_k, \bar{\boldsymbol{\theta}})\varphi_f^\top(t_k, \bar{\boldsymbol{\theta}})\}$ under mild conditions as the number of samples tends to infinity. Thus, the generic nonsingularity of this matrix as N tends to infinity follows from similar arguments to the ones used for proving Statement 1 of Theorems 4.1 and 6.1. Also, note that an important byproduct of this result is that, provided the iterations of the extended SRIVC estimator converge, the estimate must satisfy (cf. (4.54))

$$\sum_{k=1}^N \hat{\varphi}_f(t_k, \bar{\boldsymbol{\theta}}^N) \left(y(t_k) - \left\{ \frac{\bar{B}_N(p)}{\bar{A}_N(p)} u(t) \right\}_{t=t_k} \right) = \mathbf{0}, \quad (6.16)$$

where $\bar{B}_N(p)/\bar{A}_N(p)$ is the transfer function estimate described by $\bar{\boldsymbol{\theta}}^N$. Next, we will study the implications of Lemma 6.1.

Theorem 6.2. *Consider the system described by (6.1) and (6.2), and assume that the correct model structure is chosen. Under Assumption 6.1, if the iterations of the extended SRIVC method converge as $j \rightarrow \infty$, then the converging point is a stationary point of the likelihood function for an i.i.d. Gaussian noise.*

Proof. By repeating the derivation of the log-likelihood in the proof of Theorem 4.2, we have

$$\frac{\partial \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \frac{1}{\sigma^2} \sum_{k=1}^N \psi(t_k, \boldsymbol{\theta}) \left(y(t_k) - \left\{ \frac{B(p)}{A(p)} u(t) \right\}_{t=t_k} \right), \quad (6.17)$$

where

$$\psi(t_k, \boldsymbol{\theta}) = \left[-\frac{pB(p)}{A^2(p)} u(t), \dots, -\frac{p^n B(p)}{A^2(p)} u(t), \frac{1}{A(p)} u(t), \dots, \frac{p^m}{A(p)} u(t) \right]_{t=t_k}^\top. \quad (6.18)$$

On the other hand, by repeating the arguments that lead to (4.54) in Lemma 4.5, we know that the converging point of the extended SRIVC algorithm must satisfy (6.16). If the input is exactly reconstructed during the filtering process in (6.7), then $\psi(t_k, \boldsymbol{\theta}) = \hat{\varphi}_f(t_k, \boldsymbol{\theta})$, and thus by comparing (6.17) with (6.16) we find that the converging point of the extended SRIVC estimator belongs to the set of critical points of the likelihood function. \square

We have thus proven that the extended SRIVC estimator is closely linked to the stationary points of the likelihood function for the i.i.d. Gaussian noise case.

Remark 6.2. Assume that there exists an N_0 such that for all $N > N_0$, the iterations of the extended SRIVC method converge (i.e., $j \rightarrow \infty$ in Algorithm 6.1), and that assumptions for Theorem 4.1 (for the ZOH/FOH case) or Theorem 6.1 (for the multisine case) are satisfied. Then, as N tends to infinity, the extended SRIVC estimator maximizes the likelihood function with probability tending to 1. This follows from repeating the derivations in Section 3 of [206] for continuous-time estimates of an output error model structure, which show that, provided identifiability and persistence of excitation conditions are met, the global optimum of the likelihood function is given by the true parameters.

The convergence of the iterations of the extended SRIVC method will depend on how close the initial estimate is to the true system parameter vector. Intuitively, if the initial estimate is sufficiently close to the global maximum of the likelihood function and N is finite, then the extended SRIVC estimator is the maximum likelihood estimator (with full intersample knowledge of the system input) as the iterations of the method tend to infinity. For a sample size not large enough, the extended SRIVC estimator might not converge in iterations or may converge to another stationary point of the likelihood function, if at all. The number of samples needed for convergence of the extended SRIVC iterations to the maximum likelihood estimate will depend on the true system, the initial conditions, and the realization of the noise and input signals.

6.4.2 Quasi-Gauss-Newton interpretation

The results in Theorem 6.2 and Remark 6.2 show the importance of the initial estimate of the extended SRIVC method, and of the iterations themselves. We now study the iterations more closely by extending the analysis done in [228] and [252] for the standard SRIVC algorithm. We rewrite the extended SRIVC iterations in Lemma 6.2, and derive the Gauss-Newton (GN) [161] iteration for finding the maximum likelihood estimate.

Lemma 6.2 (Iteration formulas). *The iterations of the extended SRIVC method can be equivalently written as*

$$\boldsymbol{\theta}_{j+1} = \boldsymbol{\theta}_j + (\hat{\Phi}_j^\top \hat{\Phi}_j)^{-1} \hat{\Phi}_j^\top \mathbf{e}_j, \quad (6.19)$$

where

$$\hat{\Phi}_j := \begin{bmatrix} \left\{ \frac{-pB_j(p)}{A_j^2(p)} u(t) \right\}_{t=t_1}, \dots, \left\{ \frac{-p^n B_j(p)}{A_j^{n+1}(p)} u(t) \right\}_{t=t_1}, \left\{ \frac{1}{A_j(p)} u(t) \right\}_{t=t_1}, \dots, \left\{ \frac{p^m}{A_j(p)} u(t) \right\}_{t=t_1} \\ \vdots \quad \vdots \quad \vdots \quad \vdots \\ \left\{ \frac{-pB_j(p)}{A_j^2(p)} u(t) \right\}_{t=t_N}, \dots, \left\{ \frac{-p^n B_j(p)}{A_j^{n+1}(p)} u(t) \right\}_{t=t_N}, \left\{ \frac{1}{A_j(p)} u(t) \right\}_{t=t_N}, \dots, \left\{ \frac{p^m}{A_j(p)} u(t) \right\}_{t=t_N} \end{bmatrix}, \quad (6.20)$$

$$\Phi_j := \begin{bmatrix} \frac{-p}{A_j(p)} y(t_1), \dots, \frac{-p^n}{A_j(p)} y(t_1), \left\{ \frac{1}{A_j(p)} u(t) \right\}_{t=t_1}, \dots, \left\{ \frac{p^m}{A_j(p)} u(t) \right\}_{t=t_1} \\ \vdots \quad \vdots \quad \vdots \quad \vdots \\ \frac{-p}{A_j(p)} y(t_N), \dots, \frac{-p^n}{A_j(p)} y(t_N), \left\{ \frac{1}{A_j(p)} u(t) \right\}_{t=t_N}, \dots, \left\{ \frac{p^m}{A_j(p)} u(t) \right\}_{t=t_N} \end{bmatrix}, \quad (6.21)$$

and

$$\mathbf{e}_j := \left[y(t_1) - \left\{ \frac{B_j(p)}{A_j(p)} u(t) \right\}_{t=t_1}, \dots, y(t_N) - \left\{ \frac{B_j(p)}{A_j(p)} u(t) \right\}_{t=t_N} \right]^\top. \quad (6.22)$$

On the other hand, the GN-ML iteration (with step-size 1) for finding the maximizer of the likelihood function for the output error model structure with i.i.d. Gaussian noise is given by

$$\boldsymbol{\theta}_{j+1} = \boldsymbol{\theta}_j + (\hat{\Phi}_j^\top \hat{\Phi}_j)^{-1} \hat{\Phi}_j^\top \mathbf{e}_j, \quad (6.23)$$

with $\hat{\Phi}_j$ and \mathbf{e}_j given by (6.20) and (6.22), respectively.

Proof. See Appendix 6.B.1. □

Both recursive formulas of the form (6.19) and (6.23) have been presented before. First, [228] related these recursions in a discrete-time setting, and later [252] extended the analysis to a unified time domain, but without carefully analyzing the intersample behavior for the continuous-time case and therefore focusing only on input signals that are perfectly interpolated from the discrete-time data by ZOH or FOH devices. Analogously to what is stated in [252] for the standard SRIVC algorithm, the iterations in (6.19) are equivalent to the iterations of the extended SRIVC method in (6.8).

The only difference between the extended SRIVC method and the unitary step-size GN-ML iterations in (6.23) is the choice of the weighting matrix. The iterations of the extended SRIVC estimator can be interpreted as a quasi-GN method for optimizing the likelihood in the output error model structure with Gaussian noise. However, the matrix associated with the Hessian, $\hat{\Phi}_j^\top \hat{\Phi}_j$, is not positive definite (in fact, it is not symmetric). This might lead to convergence problems since the angle between the update and the gradient is not guaranteed to be acute. Furthermore, contrary to what is stated in [252, Eq. (50)], we do not find an asymptotic relationship between $(\hat{\Phi}_j^\top \hat{\Phi}_j)/N$ and $(\hat{\Phi}_j^\top \hat{\Phi}_j)/N$ as $j, N \rightarrow \infty$, because even if the estimates

converge to the true value $\boldsymbol{\theta}^*$, the difference between an element of the first n columns of the noiseless version of Φ_j and of $\hat{\Phi}_j$ as $j \rightarrow \infty$ would be of the form

$$\left\{ \frac{p^i B^*(p)}{A^{*2}(p)} u(t) \right\}_{t=t_k} - \frac{p^i}{A^*(p)} \left\{ \frac{B^*(p)}{A^*(p)} u(t) \right\}_{t=t_k}$$

which, in general, is different from zero for $i = 1, \dots, n$. However, the error between the elements in Φ_j and $\hat{\Phi}_j$ will be negligible as the sampling period tends to zero, as shown in Proposition 6.1 in Appendix 6.B. This implies that the noiseless version of Φ_j will coincide with $\hat{\Phi}_j$ as the sampling period is decreased to zero. Such insight is interesting, because it means that the iterations of the extended SRIVC method will approach the GN-ML iterations (on average) for small sampling periods and same initial point.

In terms of computational complexity, the iterations of the extended SRIVC method are more costly than GN-ML because they imply calculating an extra matrix per iteration (Φ_j). This can be important when dealing with irregular sampling, as the output filtering must be performed by solving the underlying differential equation or oversampling, with both cases producing computational burden. Note that either the extended SRIVC iterations or the GN-ML iterations can consider irregular sampling with no inherent problems.

On another note, an alternative description of (6.24) and (6.25) can reveal further differences between the extended SRIVC and GN-ML iterations.

Corollary 6.2. Let $\tilde{\boldsymbol{\theta}} = [\tilde{a}_1, \dots, \tilde{a}_n, \tilde{b}_0, \dots, \tilde{b}_m]^\top$, $A_j(p) = a_n^j p^n + a_{n-1}^j p^{n-1} + \dots + a_1^j p + 1$ and $B_j(p) = b_m^j p^m + b_{m-1}^j p^{m-1} + \dots + b_0^j$. The iterations of the extended SRIVC method can be equivalently written as

$$\boldsymbol{\theta}_{j+1} - \boldsymbol{\theta}_j = \arg \min_{\tilde{\boldsymbol{\theta}} \in \mathbb{R}^{n+m+1}} \left\| \begin{bmatrix} \left[\frac{(a_n^j + \tilde{a}_n)p^n + \dots + (a_1^j + \tilde{a}_1)p + 1}{A_j(p)} y(t_k) \right. \\ \left. - \left\{ \frac{(b_m^j + \tilde{b}_m)p^m + \dots + (b_0^j + \tilde{b}_0)}{A_j(p)} u(t) \right\}_{t=t_k} \right] \end{bmatrix}_{k=1,\dots,N} \right\|_{(\hat{\Phi}_j \hat{\Phi}_j^\top)}^2, \quad (6.24)$$

where $\|\mathbf{x}\|_\mathbf{Q}^2$ is defined as $\mathbf{x}^\top \mathbf{Q} \mathbf{x}$, and the notation $[f(t_k)]_{k=1,\dots,N}$ refers to an N -th dimensional vector of elements $f(t_k)$, $k = 1, \dots, N$. Analogously, the GN-ML iterations can also be expressed as

$$\boldsymbol{\theta}_{j+1} - \boldsymbol{\theta}_j = \arg \min_{\tilde{\boldsymbol{\theta}} \in \mathbb{R}^{n+m+1}} \left\| \begin{bmatrix} \left[y(t_k) + \left\{ \frac{(\tilde{a}_n p^n + \dots + \tilde{a}_1 p) B_j(p)}{A_j^2(p)} u(t) \right\}_{t=t_k} \right. \\ \left. - \left\{ \frac{(b_m^j + \tilde{b}_m)p^m + \dots + (b_0^j + \tilde{b}_0)}{A_j(p)} u(t) \right\}_{t=t_k} \right] \end{bmatrix}_{k=1,\dots,N} \right\|_{(\hat{\Phi}_j \hat{\Phi}_j^\top)}^2. \quad (6.25)$$

Proof. The iterations of the extended SRIVC method can be expressed as

$$\begin{aligned}\boldsymbol{\theta}_{j+1} - \boldsymbol{\theta}_j &= \arg \min_{\tilde{\boldsymbol{\theta}} \in \mathbb{R}^{n+m+1}} \left\| \hat{\Phi}_j^\top \mathbf{e}_j - \hat{\Phi}_j^\top \Phi_j \tilde{\boldsymbol{\theta}} \right\|^2 \\ &= \arg \min_{\tilde{\boldsymbol{\theta}} \in \mathbb{R}^{n+m+1}} \left\| \mathbf{e}_j - \Phi_j \tilde{\boldsymbol{\theta}} \right\|_{(\hat{\Phi}_j \hat{\Phi}_j^\top)}^2.\end{aligned}$$

Rewriting the expression above in terms of $A_j(p)$, $B_j(p)$ and $\tilde{\boldsymbol{\theta}}$, we obtain (6.24). For the alternative GN-ML iterations, we consider

$$\boldsymbol{\theta}_{j+1} - \boldsymbol{\theta}_j = \arg \min_{\tilde{\boldsymbol{\theta}} \in \mathbb{R}^{n+m+1}} \left\| \mathbf{e}_j - \hat{\Phi}_j \tilde{\boldsymbol{\theta}} \right\|_{(\hat{\Phi}_j \hat{\Phi}_j^\top)}^2$$

and expand similarly. \square

Corollary 6.2 provides a way to interpret the effect of the intersample behavior of the output. At first glance, the updates of the extended SRIVC method seem more natural because they can be seen as finding the optimal polynomials $A_{j+1}(p)$ and $B_{j+1}(p)$ that minimize

$$\left\| \left[\frac{A_{j+1}(p)}{A_j(p)} y(t_k) - \left\{ \frac{B_{j+1}(p)}{A_j(p)} u(t) \right\}_{t=t_k} \right]_{k=1,\dots,N} \right\|_{(\hat{\Phi}_j \hat{\Phi}_j^\top)}, \quad (6.26)$$

which in fact has been noted previously in Remark 4.2 of Chapter 4. However, the noise present in $y(t_k)$ in (6.24) and (6.26) is also being prefiltered, whereas the noise remains untouched in the GN-ML iterations. On the other hand, the GN-ML updates in (6.25) do not manipulate the noise in $y(t_k)$ but can lead to large updates if the estimate $A_j(p)$ is poor, because this inaccurate estimate is squared in the term that updates the denominator polynomial.

6.4.3 Refinements to the extended SRIVC iterations

Natural refinements of the iterations of the extended SRIVC method arise from the quasi-GN interpretation presented in the previous subsection. The goal of these refinements is to guide the initial estimate more smoothly towards the converging point so that a larger domain of attraction can be achieved. One possibility is to introduce a sequence of non-negative step factors, by which the iterations can be computed as

$$\begin{aligned}\boldsymbol{\theta}_{j+1} &= \boldsymbol{\theta}_j + \gamma_j (\hat{\Phi}_j^\top \Phi_j)^{-1} \hat{\Phi}_j^\top \mathbf{e}_j \\ &= (1 - \gamma_j) \boldsymbol{\theta}_j + \gamma_j (\hat{\Phi}_j^\top \Phi_j)^{-1} \hat{\Phi}_j^\top \mathbf{y}_j.\end{aligned} \quad (6.27)$$

This refinement can be useful if the initial point is thought to be a rough estimate of the true parameters, since the initial large updates can be controlled at the expense

of possibly more iterations to achieve the desired tolerance threshold. It also serves as a way to reduce the correction term if stability is lost in the next iteration. This interpretation has been noted in [209] as part of the analysis of a discrete-time estimation algorithm for the output error model structure.

Remark 6.3. A similar iterative method to (6.27) was proposed in [151], where the estimates are updated according to the Levenberg-Marquardt algorithm [157]. First, the instrument matrix $\hat{\Phi}_j$ is replaced by Φ_j in (6.23), and a constant matrix is added to the matrix that is being inverted. After a fixed number of steps, a standard GN iteration procedure with regularization is used. Although this work has received some criticism [251], the algorithm can also provide a robust way to obtain an maximum likelihood estimate of the system parameters, at least for ZOH input signals.

To conclude this analysis, we also link the descent interpretation to the stability enforcement algorithm proposed in [103] for the standard SRIVC estimator. In that work, it is proposed that the user should solve the following optimization problem at each iteration:

$$\min_{\theta \in \Omega} \|\hat{\Phi}_j^\top \mathbf{y}_j - \hat{\Phi}_j^\top \Phi_j \theta\|_2^2, \quad (6.28)$$

where

$$\Omega := \left\{ [a_1, \dots, a_n, b_0, \dots, b_m]^\top \in \mathbb{R}^{n+m+1} : \frac{1}{a_n p^n + a_{n-1} p^{n-1} + \dots + a_1 p + 1} \text{ is stable} \right\}.$$

That is, the algorithm in [103] enforces stability in each iteration by finding the parameter vector that is closest to the SRIVC estimate that ensures a stable model. On the other hand, if we compute a projected GN recursion from (6.19), we obtain

$$\theta_{j+1} = \underset{\Omega}{\text{proj}} \left(\theta_j + (\hat{\Phi}_j^\top \Phi_j)^{-1} \hat{\Phi}_j^\top \mathbf{e}_j \right) = \arg \min_{\theta \in \Omega} \|(\hat{\Phi}_j^\top \Phi_j)^{-1} \hat{\Phi}_j^\top \mathbf{y}_j - \theta\|_{\mathbf{Q}}^2, \quad (6.29)$$

where $\mathbf{Q} \succ \mathbf{0}$. Note that if \mathbf{Q} is chosen as $(\hat{\Phi}_j^\top \hat{\Phi}_j \hat{\Phi}_j^\top \Phi_j)$ in (6.29), we reach (6.28). In conclusion, the method in [103] can be interpreted as a projected quasi-GN algorithm, and can be also naturally incorporated into the extended SRIVC framework.

6.5 Computational aspects

The extended SRIVC method, as detailed in Algorithm 6.1, requires filtering the continuous-time input through continuous-time filters. This is not a straightforward task if the input does not have a simple intersample behavior (constant or piecewise linear, for example). For arbitrary input signals the filtering can be done by solving the associated ordinary differential equation related to each filter via, e.g., Runge-Kutta methods [45]. However, this may add extra computational burden and implementation difficulties when dealing with stiff systems. Here we show that the computations in (6.6) and (6.7) can be generalized for an arbitrary input signal

with arbitrary accuracy by introducing a delta transform description. Only regular sampling of the output is considered for this procedure, although extensions to irregular sampling are straightforward.

The proposed algorithm for computing $\varphi_f(t_k)$ and $\hat{\varphi}_f(t_k)$ at the j -th iteration of the extended SRIVC method goes as follows:

1. Given the sampling period h of $y(t)$, oversample $u(t)$ by a factor $S \gg 1$.
2. From θ_j , form a state-space realization of the prefilters of interest, namely $p^i B_j(p)/A_j^2(p)$ and $p^l/A_j(p)$ for $i = 1, \dots, n$; $l = 0, \dots, m$.
3. Compute the delta equivalent [154] of the prefilters in state-space form, i.e.,

$$\begin{aligned}\mathbf{x}_+(lh/S) &= \frac{h}{S} \mathbf{A}_\delta \mathbf{x}(lh/S) + \frac{h}{S} \mathbf{B}_\delta u(lh/S), \\ \mathbf{z}(lh/S) &= \mathbf{C}_\delta \mathbf{x}(lh/S) + \mathbf{D}_\delta u(lh/S),\end{aligned}$$

where $\mathbf{z}(lh/S)$ denotes the output of the filtering process, and $\mathbf{x}_+(lh/S) := \mathbf{x}([l+1]h/S) - \mathbf{x}(lh/S)$.

4. Calculate the response at the time instants $t_k = kh$ of each filter to the fast-sampled version of $u(t)$ by solving the recursion in point 3.

The extended SRIVC method with prefilters computed as in steps 1 to 4 above calculates the filtered regressor and instrument vectors accurately if the oversampling rate S is large, and exact simulation is achieved for $S \rightarrow \infty$. In practice, we have found that sampling 100 times faster is usually enough to provide reliable estimates. Potentially high sampling rates might make the delta operator description preferable for ameliorating rounding errors and ill-conditioning problems regarding the sensitivity of the coefficients of the prefilters.

6.5.1 Particular case: Band-limited signals

The method described above can be used in any situation when an arbitrary continuous-time input signal is recorded. Here we propose an alternative procedure to compute the filtering steps for when the input is a band-limited signal. For the following approach we only require even samples of the input or reference signal, not necessarily sampled at a faster rate than the output. We briefly review concepts of band-limited signals before we present the method.²

A *band-limited* signal is a continuous-time signal that does not have energy above a certain frequency ω_B . That is, $U(i\omega) = 0$ for $|\omega| > \omega_B$, where $U(i\omega)$ is the continuous-time Fourier transform

$$U(i\omega) = \int_{-\infty}^{\infty} u(t) e^{-i\omega t} dt.$$

²These concepts will be revisited and extensively studied in Chapters 9 and 10; however, for the sake of completeness, we will introduce the essential concepts here also.

If $u(t)$ is sampled every h seconds, where $h \leq \pi/\omega_B$, then its discrete-time Fourier transform pair is given by

$$U_h(e^{i\omega h}) = h \sum_{k=-\infty}^{\infty} u(kh)e^{-i\omega kh} \iff u(kh) = \frac{1}{2\pi} \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} U_h(e^{i\omega h})e^{i\omega kh} d\omega.$$

These expressions can be exploited so that the discrete-time Fourier transform is written in terms of the continuous-time one, which is known as Poisson's summation formula [19]

$$U_h(e^{i\omega h}) = \sum_{n=-\infty}^{\infty} U\left(i\omega + i\frac{2\pi n}{h}\right).$$

Due to $u(t)$ being band-limited, this formula indicates that $U_h(e^{i\omega h}) = U(i\omega)$ for $|\omega| < \pi/h$.

Our interest lies in computing the output of a single-input multi-output linear system $\mathbf{H}(p)$ such as the one in (6.7), or the last $m+1$ elements of (6.6), when the input $u(t)$ is band-limited. By noting that the continuous-time Fourier transform of the output of $H(p)$ is $\mathbf{Z}(i\omega) = \mathbf{H}(i\omega)U(i\omega)$, we also conclude that the output $\mathbf{z}(t)$ is band-limited and $\mathbf{Z}_h(e^{i\omega h}) = \mathbf{Z}(i\omega)$, also for $|\omega| < \pi/h$. By using these identities we can reconstruct the output of an LTI system with a band-limited excitation. Leveraging the inverse Fourier transform of $\mathbf{Z}_h(e^{i\omega h})$ yields

$$\begin{aligned} \mathbf{Z}(kh) &= \frac{1}{2\pi} \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} \mathbf{H}(i\omega)U_h(e^{i\omega h})e^{i\omega kh} d\omega \\ &= \frac{1}{\pi} \operatorname{Re} \left\{ \int_0^{\frac{\pi}{h}} \mathbf{H}(i\omega)U_h(e^{i\omega h})e^{i\omega kh} d\omega \right\}, \end{aligned}$$

where we have used the fact that the Fourier transform of a real signal is conjugate symmetric. Thus, the output can be computed by solving N integrals. This is computationally expensive, considering that such integration must be performed at each iteration of the extended SRIVC algorithm and these integrals usually do not have a closed form. Instead of this, we evaluate its associated $(M+1)$ -point Riemann sum

$$\mathbf{z}(kh) \approx \frac{1}{\pi h(M+1)} \operatorname{Re} \left\{ \sum_{l=0}^M \mathbf{H}\left(\frac{i\pi l}{hM}\right) U_h(e^{i\frac{\pi l}{M}}) e^{i\frac{\pi lk}{M}} \right\}.$$

Since the input samples $u(0), \dots, u([N-1]h)$ are known, we can compute $U_h(e^{i\frac{\pi l}{M}})$ as the zero-padded discrete Fourier transform (DFT) of the input vector following an FFT algorithm. More precisely,

$$U_h(e^{i\frac{\pi l}{M}}) = h \sum_{k=0}^{N-1} u(kh) e^{-i\frac{\pi lk}{M}}.$$

Once $\{U_h(e^{i\frac{\pi l}{M}})\}_{l=0}^M$ is computed, the output vector $\{\mathbf{z}(kh)\}_{k=0}^{N-1}$ is obtained via the (zero-padded) inverse DFT of $\mathbf{H}\left(\frac{i\pi l}{hM}\right) U_h(e^{i\frac{\pi l}{M}})$.

Remark 6.4. The choice of M dictates the accuracy of the filtering procedure described above: a larger value of M provides more accurate results at the expense of more arithmetic operations.

6.6 Monte Carlo simulation studies

We now compare the performance of the standard SRIVC method with the extended SRIVC method (labeled SRIVC-x in the following tests) via numerical simulations, and show the benefits of including a non-unitary step size in its implementation. The consistency of both methods is examined for a multisine input for different regular sampling periods, and also for irregular sampling. We also study the consistency of each method under an arbitrary input excitation for different regular sampling periods. For the tests in Subsections 6.6.1 to 6.6.5, we consider the system

$$G^*(p) = \frac{1.25}{0.25p^2 + 0.7p + 1}, \quad (6.30)$$

where the parameters of interest are $a_1^* = 0.7$, $a_2^* = 0.25$, and $b_0^* = 1.25$. Regarding the implementation of the standard SRIVC method, we use the `srivc` command from the CONTSID toolbox version 7.3 for MATLAB [60], under default initialization and tolerance settings. The correct model structure is set in the implementation, and an FOH is used as the intersample behavior of the input signal.

6.6.1 Multisine input: Regular sampling

We first test if the algorithms provide consistent estimates of the parameters $[a_1^*, a_2^*, b_0^*]$. The system in (6.30) is excited with the input

$$u(t) = \sin(0.714t) + \sin(1.428t) + \sin(2.142t).$$

The noiseless output is computed analytically by assuming that it corresponds to the output of the system in a stationary regime, i.e.,

$$x(t) = \sum_{l=1}^3 |G^*(i\omega_l)| \sin(\omega_l t + \angle G^*(i\omega_l)),$$

where $[\omega_1, \omega_2, \omega_3] = [0.714, 1.428, 2.142]\text{[rad/s]}$. This output is sampled at $h = 0.3\text{[s]}$ and is contaminated with additive noise, which is set as an i.i.d. Gaussian sequence of variance 0.1. Sixty different sample sizes are considered, ranging logarithmically from $N = 100$ to $N = 25500$, and 300 Monte Carlo runs are performed for each sample size.

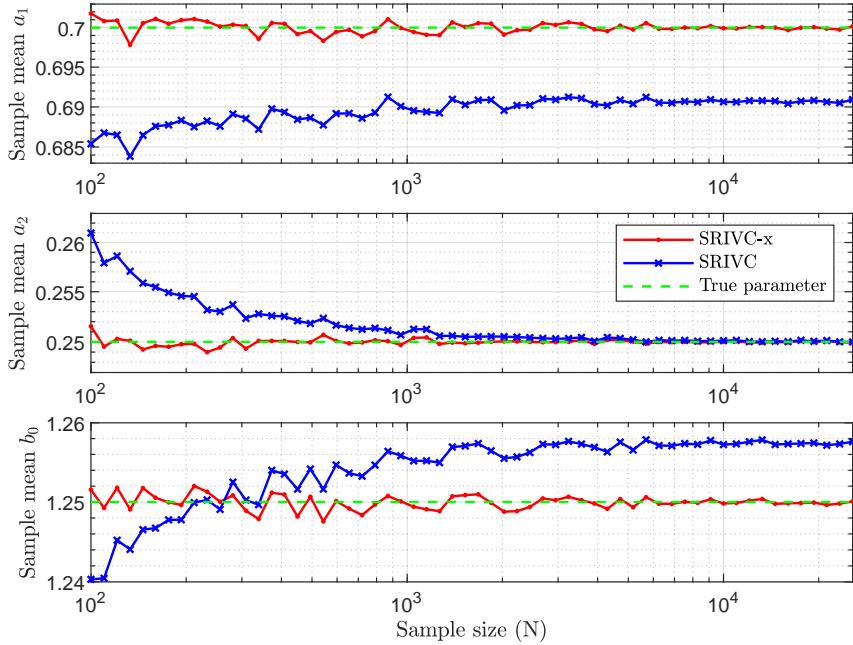


Figure 6.1: Consistency test using a multisine input under regular output sampling. Sample means of each estimated parameter for SRIVC-x (red), and standard SRIVC (blue). The true parameters are in dashed-green.

Figures 6.1 and 6.2 show the sample means and sample mean square errors (MSEs) of each estimated parameter, respectively. The extended SRIVC estimator accurately identifies all parameters while the standard SRIVC method fails to recover the true parameters as N increases. The MSEs for the parameters estimated with the extended SRIVC estimator decrease to zero as expected for consistency, while at least two out of the three estimated parameters given by the SRIVC method are biased. This observation empirically indicates that the SRIVC estimator is not consistent in this example.

6.6.2 Multisine input: Different sampling periods

We now study the effect of the input intersample behavior on the SRIVC and extended SRIVC estimates. Under the same input and noise variance as the previous simulation, we test the performance of each algorithm for a fixed number of output measurements ($N = 2000$) with different regular sampling periods. Since the rise time of the system is approximately 2 seconds, a good choice for the sampling period should be between 0.2 and 0.5 seconds according to the criterion suggested in [16]. We test with sampling periods $h = 0.06, 0.2, 0.6[\text{s}]$ to cover fast, normal and slow sampling, respectively.

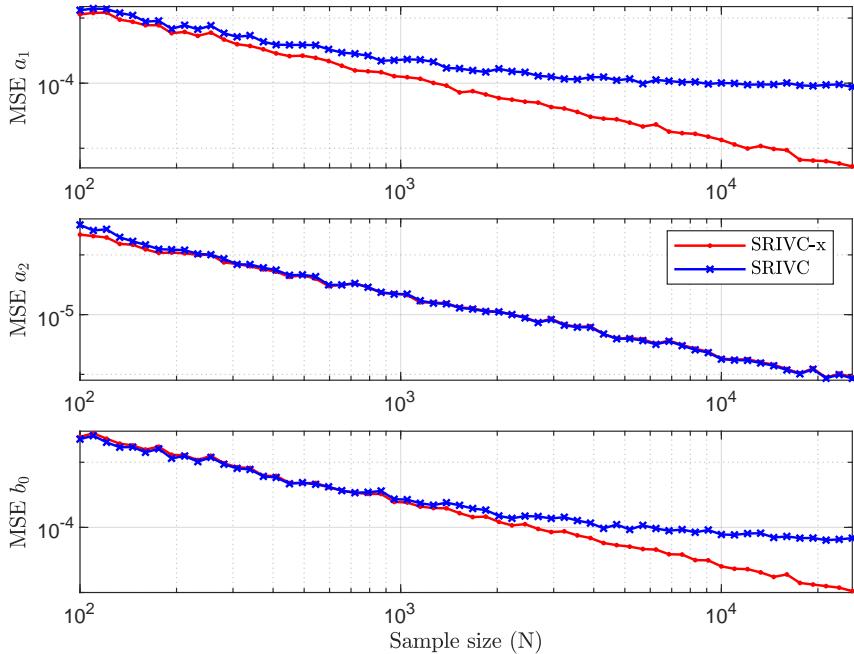


Figure 6.2: Consistency test using a multisine input under regular output sampling. Sample MSE of each estimated parameter for SRIVC-x (red), and standard SRIVC (blue).

The sample mean and mean square error of each parameter over 300 Monte Carlo runs for each sampling period are shown in Table 6.1. On average, the SRIVC-x estimator delivers the true values of every parameter for all sampling periods in this study, whereas the SRIVC estimator only performs well when the sampling period is small. The decline in performance of the SRIVC estimates for larger sampling periods is due to the severe differences in the FOH implementation of the multisine compared to the true continuous-time version. For $h = 0.6[\text{s}]$, the large sampling period exaggerates the interpolation error of the input signal in the standard SRIVC estimator, which severely degrades its performance. This is seen by the order of magnitude of difference in MSE of the parameters given by the two estimators.

6.6.3 Multisine input: Irregular sampling

We consider the same system described before, with the same input and noise variance. This time 2000 irregularly sampled output measurements are obtained. The sampling interval is distributed uniformly between h_{lb} and h_{ub} , where the lower bound is fixed at $h_{lb} = 0.05[\text{s}]$ and the upper bound is varied from 0.1 to 0.6[s]. A total of six Monte Carlo simulations are performed with each simulation containing 300 runs.

Table 6.1: Sample mean and MSE of each parameter, for SRIVC and SRIVC-x, when $h = 0.06, 0.2$ and $0.6[\text{s}]$.

Method	Param. (Value)	Stats.	$h = 0.06[\text{s}]$	$h = 0.2[\text{s}]$	$h = 0.6[\text{s}]$
SRIVC	$a_1(0.7)$	Mean MSE	0.697 $7.0 \cdot 10^{-5}$	0.694 $9.3 \cdot 10^{-5}$	0.668 $1.1 \cdot 10^{-3}$
	$a_2(0.25)$	Mean MSE	0.253 $1.8 \cdot 10^{-5}$	0.251 $1.1 \cdot 10^{-5}$	0.248 $1.6 \cdot 10^{-5}$
	$b_0(1.25)$	Mean MSE	1.244 $1.6 \cdot 10^{-4}$	1.251 $1.2 \cdot 10^{-4}$	1.286 $1.5 \cdot 10^{-3}$
SRIVC-x	$a_1(0.7)$	Mean MSE	0.700 $6.0 \cdot 10^{-5}$	0.699 $6.3 \cdot 10^{-5}$	0.700 $6.7 \cdot 10^{-5}$
	$a_2(0.25)$	Mean MSE	0.250 $1.1 \cdot 10^{-5}$	0.250 $1.1 \cdot 10^{-5}$	0.250 $1.2 \cdot 10^{-5}$
	$b_0(1.25)$	Mean MSE	1.249 $1.3 \cdot 10^{-4}$	1.250 $1.3 \cdot 10^{-4}$	1.250 $1.3 \cdot 10^{-4}$

Figure 6.3 shows the mean value of each parameter, with their standard deviation around this value. As expected, the extended SRIVC estimator provides accurate estimates for all sampling period ranges in this study. On the other hand, the SRIVC estimator has a degrading performance as the sampling range increases, which could be attributed to the approximation errors in the prefilter calculations due to the misspecification of the intersample behavior of the input. This observation is in line with the conclusion after Remark 4.1, that mentions that the bias of the SRIVC estimates decreases with the sampling period.

6.6.4 Arbitrary input: Chirp signal

The next goal is to check whether the extended SRIVC algorithm can provide accurate estimates for arbitrary input signals. Now the system in (6.30) is excited with an up-chirp signal, which is a continuous-time signal that increases in frequency with time. These signals are widely used in signal processing applications such as radar systems and seismology and have been used for system identification [158, 240]. The chirp signal used in this example is given by

$$u(t) = \cos \left(f_0 \left[\frac{f_1}{f_0} \right]^{t/T_f} 2\pi t \right),$$

where $[f_0, f_1] = [0.1, 0.6]\text{[Hz]}$ and $T_f = 500[\text{s}]$ is the length of one period of the chirp signal. We determine the true system output using the explicit Runge-Kutta formulas RK5(4) [45]. The output is sampled every $h = 0.5[\text{s}]$ with $S = 500$, and the measurement noise has variance 0.05, which corresponds to approximately 10% of

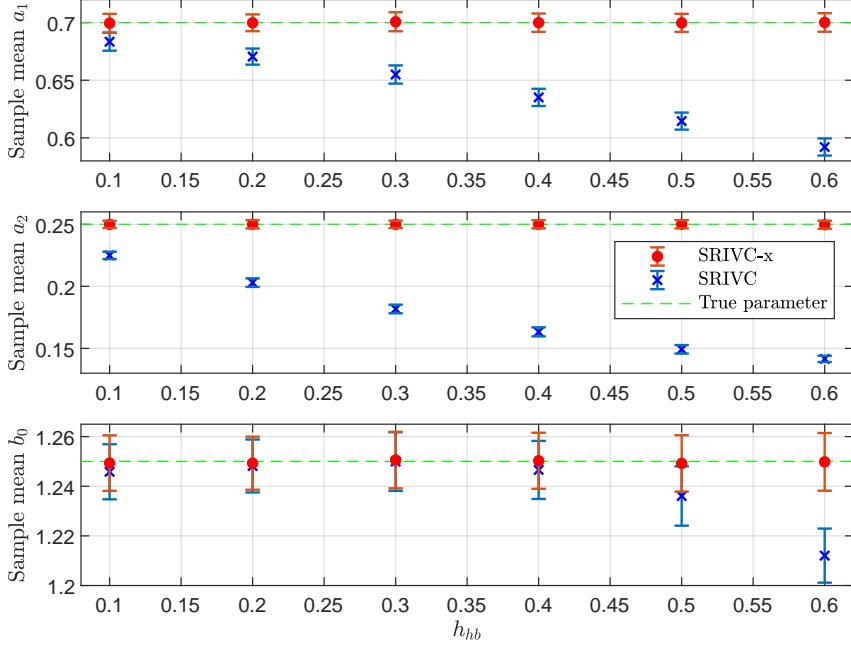


Figure 6.3: Consistency test using a multisine input under irregular output sampling. Sample means of each estimated parameter using SRIVC-x (red) and standard SRIVC (blue), with 1 standard deviation, for different sampling intervals. The true parameter values are in dashed green.

the energy of the noiseless output. We follow the algorithm described in Section 6.5 for the computation of the SRIVC-x estimate. The number of periods of the input varies from 1 to 10, which is equivalent to sample sizes ranging from 1000 to 10000, and 300 Monte Carlo runs are performed for each sample size.

The empirical evidence in Figure 6.4 suggests that the proposed procedure for computing the regressor and instrument vectors leads to accurate estimates of the parameters of interest, while the standard SRIVC method fails to deliver statistically consistent estimates. This improvement in accuracy is not without a drawback in computational efficiency: the total computation time of the extended SRIVC estimator tests is 30.2 hours, while the SRIVC estimates are computed in 1.3 hours. The simulations are performed on a laptop with an Intel Core i7-7600u 2.8Ghz processor.

6.6.5 Robustified extended SRIVC

We now show the advantage of using a non-unitary step size in the extended SRIVC iterations. We consider the Rao-Garnier system described in (3.17), which is excited by a multisine signal given by the sum of sine waves of angular frequencies

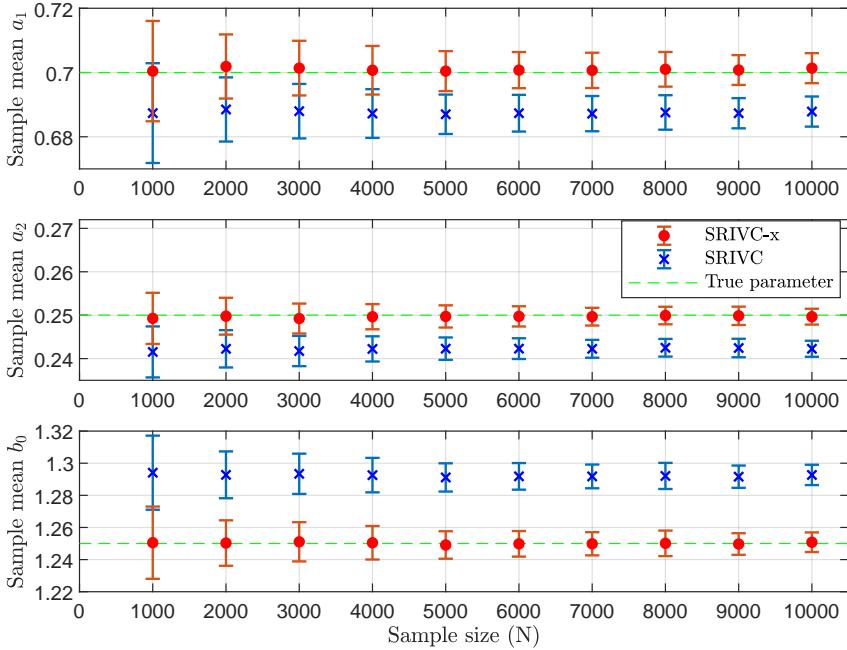


Figure 6.4: Test with chirp input signal. Sample means of each parameter with 1 standard deviation for SRIVC-x (red circles) and standard SRIVC (blue crosses). The true parameter values are in dashed green.

$\omega = 1, 1.9, 2.1, 18, 22[\text{rad/s}]$. The signal to noise ratio is set to 3[dB], and $N = 500$ output measurements were collected with a sampling period $h = 0.03[\text{s}]$. We test the SRIVC method with FOH intersample behavior assumptions in the input, the extended SRIVC method, and the extended SRIVC implementation in (6.27) with a decreasing step size $\gamma_j = 1/j$ (labeled as SRIVC-xr). All methods are initialized with the SVF estimate with cutoff frequency $\lambda = 25[\text{rad/s}]$. The tolerance factor and the maximum number of iterations are set to 10^{-5} and 200, respectively. The estimated model is chosen to have five poles and one zero, and 1000 Monte Carlo runs are performed.

The box plots of the fit for each method are shown in Figure 6.5. While SRIVC fails to provide good estimates, both the extended SRIVC method and its non-unitary step size version perform well even in this highly noisy and over-parameterized scenario. By choosing a decreasing step size for the extended SRIVC iterations, the number of outliers decreases compared to the unitary step size implementation. These simulation results confirm that it is possible to improve the robustness of the extended SRIVC procedure by reducing the step size, at the dispense of possibly more iterations needed for convergence.

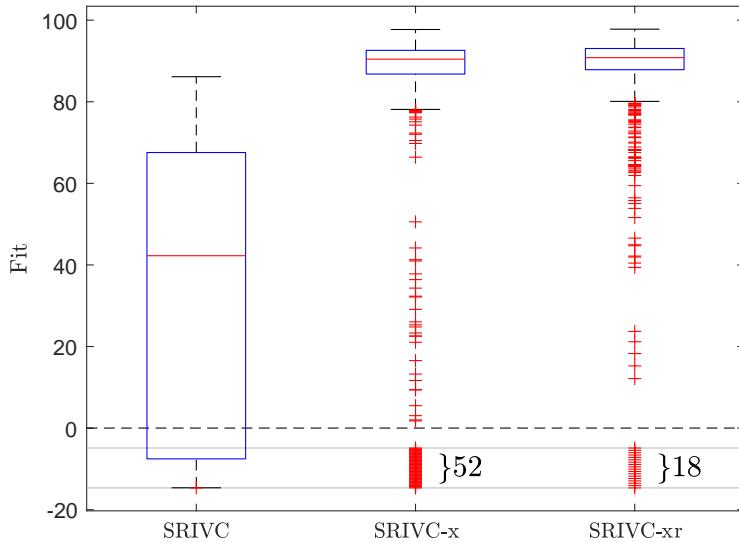


Figure 6.5: Fit box plots for the Rao-Garnier system. Crosses in between the horizontal lines are compressed outliers (fits less than zero). Numbers in the SRIVC-x and SRIVC-xr plots indicate the number of runs that obtained less than zero fit.

6.7 Conclusions

In this chapter we have derived an algorithm for continuous-time system identification that is proven to be generically consistent for a wide class of input signals that have a known intersample behavior. This estimator extends the applicability of the standard SRIVC method to inputs that are not exactly described by hold devices. We put forward a comprehensive analysis of the generic consistency of this estimator for multisine inputs, and exposed the relationship between this algorithm and the maximum likelihood method by showing that it provides an estimate that is a root of the likelihood equation in the output error model structure under Gaussian noise. Furthermore, we interpreted the extended SRIVC iterations as a quasi-Gauss-Newton procedure, which hinted on how to improve its robustness. Computational aspects regarding the filtering steps have also been covered. Extensive simulations have confirmed the theoretical findings, and revealed the advantages of the extended SRIVC estimator over the standard SRIVC method for when the user knows the exact intersample behavior of the input.

Appendix

6.A Supplementary material for consistency results

Lemma 6.3. Consider $u(t)$ as in (6.9), and $\mathbf{v}(t_k)$ and $\mathbf{u}_{n+m}(t_k)$ as defined in (4.5) and (4.15), respectively. Under Assumption 6.2, the matrix

$$\Psi = \bar{\mathbb{E}} \left\{ \left[\frac{1}{A_j^2(p)} \mathbf{u}_{n+m}(t) \right]_{t=t_k} \mathbf{v}^\top(t_k) \right\}$$

is equal to the null matrix.

Proof. We directly obtain that the elements Ψ_{rl} , with $l \geq n + 1$, are equal to zero from the definition of $\mathbf{v}(t_k)$. For the other entries of Ψ , we see that an arbitrary entry of this matrix is of the form

$$\Psi_{rl} = \bar{\mathbb{E}} \left\{ \left\{ \frac{p^{n+m+1-r}}{A_j^2(p)} u(t) \right\}_{t=t_k} \frac{p^l}{A_j(p)} v(t_k) \right\}. \quad (6.31)$$

If we define $g_r(t)$ as the inverse Laplace transform of $s^{n+m+1-r} A_j^{-2}(s)$, the first factor in the expectation in (6.31) can be written as

$$\left\{ \frac{p^{n+m+1-r}}{A_j^2(p)} u(t) \right\}_{t=t_k} = \int_0^{t_k} g_r(t_k - \tau) u(\tau) d\tau.$$

Note that this is a discrete-time signal that is a function of the time measurements $\{t_k\}_{k \in \mathbb{N}}$. On the other hand, the second factor in (6.31) can be described by

$$\frac{p^l}{A_j(p)} v(t_k) = \sum_{s=1}^k \beta_{k-s,l} v(t_s).$$

where $\{\beta_{s,l}\}_{s=0}^{k-1}$ are the first k values of the impulse response of the FOH discrete-

time equivalent of $p^l/A_j(p)$. So, we compute Ψ_{rl} as

$$\begin{aligned}\Psi_{rl} &= \bar{\mathbb{E}} \left\{ \int_0^{t_k} g_r(t_k - \tau) u(\tau) d\tau \sum_{s=1}^k \beta_{k-s,l} v(t_s) \right\} \\ &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N \sum_{s=1}^k \int_0^{t_k} g_r(t_k - \tau) u(\tau) \beta_{k-s,l} \mathbb{E}\{v(t_s)\} d\tau \\ &= 0,\end{aligned}$$

where we have used the fact that the disturbance signal has zero mean. Thus, every element Ψ_{rl} of Ψ is equal to zero, thereby proving the statement. \square

Lemma 6.4 (Generic non-singularity of $\bar{\Phi}$). *Under Assumptions 6.1 to 6.5, with $u(t)$ described as in (6.9), the following matrix is generically non-singular:*

$$\bar{\Phi} := \bar{\mathbb{E}} \left\{ \left[\frac{1}{\bar{A}^2(p)} \mathbf{u}_{n+m}(t) \right]_{t=t_k} \left[\frac{1}{\bar{A}(p) A^*(p)} \mathbf{u}_{n+m}^\top(t) \right]_{t=t_k} \right\}.$$

Proof. Similar to the proof of Statement 1 of Theorem 4.1, we follow an analyticity argument. We must first prove that the matrix

$$\bar{\Phi}^* := \bar{\mathbb{E}} \left\{ \left[\frac{1}{A^{*2}(p)} \mathbf{u}_{n+m}(t) \right]_{t=t_k} \left[\frac{1}{A^{*2}(p)} \mathbf{u}_{n+m}^\top(t) \right]_{t=t_k} \right\}$$

is positive definite. To this end, let $\mathbf{z} \in \mathbb{R}^{n+m+1}$. We write

$$\mathbf{z}^\top \bar{\Phi}^* \mathbf{z} = \bar{\mathbb{E}} \left\{ \left(\left\{ \frac{B_{\mathbf{z}}(p)}{A^{*2}(p)} u(t) \right\}_{t=t_k} \right)^2 \right\} \geq 0.$$

Since $u(t)$ is a multisine of the form (6.9), we have in stationary regime that

$$\frac{B_{\mathbf{z}}(p)}{A^{*2}(p)} u(t) = \tilde{\alpha}_0 + \sum_{l=1}^{m_u} \tilde{\alpha}_l \cos(\omega_l t + \tilde{\phi}_l),$$

where $\tilde{\alpha}_0 = \alpha_0 B_{\mathbf{z}}(0)/A^{*2}(0)$, $\tilde{\alpha}_l = \alpha_l |B_{\mathbf{z}}(i\omega_l)/A^{*2}(i\omega_l)|$, and the phases are $\tilde{\phi}_l = \phi_l + \angle B_{\mathbf{z}}(i\omega_l)/A^{*2}(i\omega_l)$. Therefore,

$$\begin{aligned}\mathbf{z}^\top \bar{\Phi}^* \mathbf{z} &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N \left(\tilde{\alpha}_0 + \sum_{l=1}^{m_u} \tilde{\alpha}_l \cos(\omega_l k h + \tilde{\phi}_l) \right)^2 \\ &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N \left(\tilde{\alpha}_0^2 + 2\tilde{\alpha}_0 \sum_{l=1}^{m_u} \tilde{\alpha}_l \cos(\omega_l k h + \tilde{\phi}_l) \right.\end{aligned}\tag{6.32a}$$

$$\left. + \sum_{r,l=1}^{m_u} \tilde{\alpha}_r \tilde{\alpha}_l \cos(\omega_r k h + \tilde{\phi}_r) \cos(\omega_l k h + \tilde{\phi}_l) \right). \tag{6.32b}$$

Recall the formula for a geometric series

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N \cos(\omega k + \phi) &= \operatorname{Re} \left\{ \lim_{N \rightarrow \infty} \frac{e^{i\phi}}{N} \sum_{k=1}^N e^{i\omega k} \right\} \\ &= \operatorname{Re} \left\{ \lim_{N \rightarrow \infty} \frac{e^{i\phi}}{N} \frac{e^{i\omega N} - e^{i\omega(N+1)}}{1 - e^{i\omega}} \right\} \\ &= 0. \end{aligned}$$

Using this result and the identity $\cos(\alpha) \cos(\beta) = [\cos(\alpha + \beta) + \cos(\alpha - \beta)]/2$, the second term in the sum in (6.32a) is zero. In addition, the term in (6.32b) for $r \neq l$ is a sum of sinusoids whose sum tends to zero as N tends to infinity, while for $r = l$ constants appear. Thus,

$$\mathbf{z}^\top \bar{\Phi}^* \mathbf{z} = \tilde{\alpha}_0^2 + \frac{1}{2} \sum_{l=1}^{m_u} \tilde{\alpha}_l^2. \quad (6.33)$$

This shows that $\mathbf{z}^\top \bar{\Phi}^* \mathbf{z} = 0$ occurs if and only if $\tilde{\alpha}_0 = \tilde{\alpha}_1 = \dots = \tilde{\alpha}_{m_u} = 0$, which in turn is equivalent to imposing

$$\frac{B_{\mathbf{z}}(0)}{A^{*2}(0)} = 0, \quad \frac{B_{\mathbf{z}}(i\omega_l)}{A^{*2}(i\omega_l)} = \frac{B_{\mathbf{z}}(-i\omega_l)}{A^{*2}(-i\omega_l)} = 0, \quad l = 1, \dots, m_u.$$

Since $m_u \geq (n+m)/2$, the only rational function that satisfies all of these restrictions is the null transfer function. Thus, $B_{\mathbf{z}}(p) = 0$ and $\mathbf{z} = \mathbf{0}$. With this we have shown that $\bar{\Phi}^*$ is positive definite.

Now we show that the entries of the matrix $\bar{\Phi}$ are real analytic functions of the real parameter vector $[\bar{a}_1, \dots, \bar{a}_n]$ in the domain where $\bar{A}(p)$ is a stable polynomial. We denote this domain as $\Omega \subset \mathbb{R}^n$. The entries of the matrix $\bar{\Phi}$ are given by

$$\bar{\Phi}_{il} := \bar{\mathbb{E}} \left\{ \left\{ \frac{p^{n+m+1-i}}{\bar{A}^2(p)} u(t) \right\}_{t=t_k} \left\{ \frac{p^{n+m+1-l}}{\bar{A}(p) A^*(p)} u(t) \right\}_{t=t_k} \right\},$$

where $i, l = 1, 2, \dots, n+m+1$. By computing the expectation in a similar way to the derivation of (6.33), we find that

$$\bar{\Phi}_{rl} = \tilde{\alpha}_0^r \tilde{\alpha}_0^l + \frac{1}{2} \sum_{s=1}^{m_u} \tilde{\alpha}_s^r \tilde{\alpha}_s^l \cos(\tilde{\phi}_s^r - \tilde{\phi}_s^l),$$

where

$$\begin{aligned} \tilde{\alpha}_0^r &= \begin{cases} 0 & \text{if } r < n+m+1, \\ \alpha_0 & \text{if } r = n+m+1, \end{cases} \quad \tilde{\alpha}_0^l = \begin{cases} 0 & \text{if } l < n+m+1, \\ \alpha_0 & \text{if } l = n+m+1, \end{cases} \\ \tilde{\alpha}_s^r &= \alpha_s \left| \frac{\omega_s^{n+m+1-r}}{\bar{A}^2(i\omega_s)} \right|, \quad \tilde{\phi}_s^r = \phi_s + \angle \left[\frac{(i\omega_s)^{n+m+1-r}}{\bar{A}^2(i\omega_s)} \right], \\ \tilde{\alpha}_s^l &= \alpha_s \left| \frac{\omega_s^{n+m+1-l}}{\bar{A}(i\omega_s) A^*(i\omega_s)} \right|, \quad \tilde{\phi}_s^l = \phi_s + \angle \left[\frac{(i\omega_s)^{n+m+1-l}}{\bar{A}(i\omega_s) A^*(i\omega_s)} \right]. \end{aligned}$$

The coefficient $\tilde{\alpha}_s^r$ can be equivalently expressed as

$$\begin{aligned}\tilde{\alpha}_s^r &= \frac{\alpha_s |\omega_s|^{n+m+1-r}}{\operatorname{Re}\{\bar{A}(i\omega_s)\}^2 + \operatorname{Im}\{\bar{A}(i\omega_s)\}^2} \\ &= -\frac{\alpha_s |\omega_s|^{n+m+1-r}}{\left(1 + \sum_{1 \leq l \leq n, l \text{ even}} \bar{a}_l \omega_s^l (-1)^{\frac{l}{2}}\right)^2 + \left(\sum_{1 \leq l \leq n, l \text{ odd}} \bar{a}_l \omega_s^l (-1)^{\frac{l-1}{2}}\right)^2}.\end{aligned}$$

From the expression above we see that the denominator of $\tilde{\alpha}_s^r$ is a multivariate polynomial in the variables $[\bar{a}_1, \dots, \bar{a}_n]$, which is strictly positive in Ω since we know that $\bar{A}(p)$ is a stable polynomial for any $[\bar{a}_1, \dots, \bar{a}_n] \in \Omega$. This shows that the denominator of $\tilde{\alpha}_s^r$ is real analytic in Ω , and since the quotient of real analytic functions is real analytic as long as the denominator does not vanish [123, Proposition 2.2.2], we have that $\tilde{\alpha}_s^r$ is real analytic in Ω .

Similarly, the coefficient $\tilde{\alpha}_s^l$ can be written as

$$\begin{aligned}\tilde{\alpha}_s^l &= \frac{\alpha_s |\omega_s|^{n+m+1-l}}{|A^*(i\omega_s)| \sqrt{\operatorname{Re}\{\bar{A}(i\omega_s)\}^2 + \operatorname{Im}\{\bar{A}(i\omega_s)\}^2}}, \\ &= \frac{\alpha_s |\omega_s|^{n+m+1-l}}{|A^*(i\omega_s)|} \exp\left(-\frac{1}{2} \log [\operatorname{Re}\{\bar{A}(i\omega_s)\}^2 + \operatorname{Im}\{\bar{A}(i\omega_s)\}^2]\right).\end{aligned}$$

Here, the argument of the logarithm is a multivariate polynomial in the variables $[\bar{a}_1, \dots, \bar{a}_n]$ and is strictly positive in Ω . Furthermore, the natural logarithm is a real analytic function on $(0, \infty)$ (see, e.g., [220, p. 93]). Thus, by Proposition 2.2.8 of [123], which states that the composition of real analytic functions is real analytic, we conclude that $\tilde{\alpha}_s^l$ is real analytic in Ω .³

Finally, note that $\phi_s^r - \phi_s^l = \pi(l - r)/2 + \angle A^*(i\omega_s) - \angle \bar{A}(i\omega_s)$, which leads to

$$\begin{aligned}\cos(\phi_s^r - \phi_s^l) &= \cos\left(\frac{\pi}{2}(l - r) + \angle A^*(i\omega_s)\right) \frac{\operatorname{Re}\{\bar{A}(i\omega_s)\}}{|\bar{A}(i\omega_s)|} \\ &\quad + \sin\left(\frac{\pi}{2}(l - r) + \angle A^*(i\omega_s)\right) \frac{\operatorname{Im}\{\bar{A}(i\omega_s)\}}{|\bar{A}(i\omega_s)|}.\end{aligned}$$

The functions $\operatorname{Re}\{\bar{A}(i\omega_s)\}/|\bar{A}(i\omega_s)|$ and $\operatorname{Im}\{\bar{A}(i\omega_s)\}/|\bar{A}(i\omega_s)|$ are real analytic for any $[\bar{a}_1, \dots, \bar{a}_n] \in \Omega$ by the same justification above. Therefore, $\cos(\phi_s^r - \phi_s^l)$ is real analytic for any $[\bar{a}_1, \dots, \bar{a}_n] \in \Omega$. Since any function defined by multiplication and addition of real analytical functions is real analytic, we conclude that $\bar{\Phi}_{rl}$ is real analytic in the variables $[\bar{a}_1, \dots, \bar{a}_n]$, in the domain Ω .

Thus, for $n = n^*$ and $m \geq m^*$, the matrix $\bar{\Phi}$ is generically non-singular with respect to $[\bar{a}_1, \dots, \bar{a}_n] \in \Omega$ by Lemma 4.6 and Corollary 4.4. If $n \geq n^*$ and $m = m^*$,

³Note that the coefficients in $A^*(i\omega_s)$ do not play a role in the analyticity of $\tilde{\alpha}_s^l$, since $\tilde{\alpha}_s^l$ is viewed as a function of $[\bar{a}_1, \dots, \bar{a}_n]$ only.

we note that $\mathbf{a}^* := [a_1^*, \dots, a_{n^*}^*, 0, \dots, 0] \in \mathbb{R}^n$ (where $n - n^*$ zeros are appended) belongs to the boundary of Ω . However, since $\det(\bar{\Phi})$ is a real analytic function in Ω , its continuity ensures the existence of a small perturbation vector $\boldsymbol{\eta} \in \mathbb{R}^n$ such that $\mathbf{a}^* + \boldsymbol{\eta} \in \Omega$ and $\bar{\Phi}$ is non-singular when $\mathbf{a}^* + \boldsymbol{\eta}$ are the parameters of the model denominator. Hence, generic non-singularity of $\bar{\Phi}$ also holds for this case by the same lemma cited above, and this concludes the proof. \square

6.B Supplementary material for efficiency and iterations analysis

6.B.1 Proof of Lemma 6.2

We can rewrite the extended SRIVC iteration equation in (6.8) as

$$\boldsymbol{\theta}_{j+1} = (\hat{\Phi}_j^\top \Phi_j)^{-1} \hat{\Phi}_j^\top \mathbf{y}_j, \quad (6.34)$$

where Φ_j and $\hat{\Phi}_j$ are defined in the statement of Lemma 6.2, and

$$\mathbf{y}_j := \left[\frac{1}{A_j(p)} y(t_1), \frac{1}{A_j(p)} y(t_2), \dots, \frac{1}{A_j(p)} y(t_N) \right]^\top.$$

On the other hand, we note that

$$\mathbf{y}_j = \Phi_j \boldsymbol{\theta}_j + \mathbf{e}_j, \quad (6.35)$$

where the residual \mathbf{e}_j is precisely (6.22). Replacing (6.35) into (6.34) leads to (6.19). Now, we shall derive a Gauss-Newton recursion for the maximum likelihood estimator of $\boldsymbol{\theta}^*$. This recursion has the form

$$\boldsymbol{\theta}_{j+1} = \boldsymbol{\theta}_j - \mathbf{H}_j^{-1} \mathbf{g}_j, \quad (6.36)$$

where the gradient \mathbf{g}_j is given by

$$\mathbf{g}_j = \frac{\partial}{\partial \boldsymbol{\theta}} \log p(\mathbf{y}^N; \boldsymbol{\theta}) \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_j} = \frac{1}{\sigma^2} \sum_{k=1}^N \left[y(t_k) - \left\{ \frac{B_j(p)}{A_j(p)} u(t) \right\}_{t=t_k} \right] \psi(t_k, \boldsymbol{\theta}_j),$$

with $\psi(t_k, \boldsymbol{\theta}_j)$ being (6.18) with $\boldsymbol{\theta}_j$ instead of $\boldsymbol{\theta}$. The Hessian \mathbf{H}_j is obtained as

$$\begin{aligned} \mathbf{H}_j &= \mathbb{E} \left\{ \frac{\partial^2}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} \log p(\mathbf{y}^N; \boldsymbol{\theta}) \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_j} \right\} \\ &= -\mathbb{E} \left\{ \left(\frac{\partial}{\partial \boldsymbol{\theta}} \log p(\mathbf{y}^N; \boldsymbol{\theta}) \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_j} \right) \left(\frac{\partial}{\partial \boldsymbol{\theta}} \log p(\mathbf{y}^N; \boldsymbol{\theta}) \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_j} \right)^\top \right\} \\ &= -\frac{1}{\sigma^4} \sum_{k=1}^N \sum_{l=1}^N \psi(t_k, \boldsymbol{\theta}_j) \psi^\top(t_l, \boldsymbol{\theta}_j) \mathbb{E} \left\{ \left(\left\{ \left(\frac{B^*(p)}{A^*(p)} - \frac{B_j(p)}{A_j(p)} \right) u(t) \right\}_{t=t_k} + v(t_k) \right) \right. \\ &\quad \times \left. \left(\left\{ \left(\frac{B^*(p)}{A^*(p)} - \frac{B_j(p)}{A_j(p)} \right) u(t) \right\}_{t=t_l} + v(t_l) \right) \right\}. \end{aligned}$$

The terms related to the input in the expectation above can be disregarded for $\bar{G}(p) \approx G^*(p)$. Thus, we arrive at

$$\mathbf{H}_j \approx -\frac{1}{\sigma^2} \sum_{k=1}^N \boldsymbol{\psi}(t_k, \boldsymbol{\theta}_j) \boldsymbol{\psi}^\top(t_k, \boldsymbol{\theta}_j).$$

Thus, by replacing \mathbf{g}_j and \mathbf{H}_j into (6.36), we obtain (6.25), concluding the proof. \square

Proposition 6.1. Consider the matrices (6.20) and (6.21) when evaluated at the true parameters, denoted by $\hat{\Phi}^*$ and Φ^* respectively. Furthermore, assume that there is no noise present in Φ^* , and that $y(t_k)$ is reconstructed using a ZOH device. We have that $\|\hat{\Phi}^* - \Phi^*\|_\infty \rightarrow 0$ as $\max_{k \in \{1, \dots, N\}} (t_k - t_{k-1}) \rightarrow 0$.

Proof. The difference $\hat{\Phi}^* - \Phi^*$ is a matrix whose non-zero (k, r) entries are of the form

$$\left\{ \frac{p^r B^*(p)}{A^{*2}(p)} u(t) \right\}_{t=t_k} - \frac{p^r}{A^*(p)} \left\{ \frac{B^*(p)}{A^*(p)} u(t) \right\}_{t=t_k}$$

for $r = 1, \dots, n$ and $k = 1, \dots, N$. By denoting $x(t)$ as in (6.1), we find that

$$\left| \left\{ \frac{p^r}{A^*(p)} x(t) \right\}_{t=t_k} - \frac{p^r}{A^*(p)} x(t_k) \right| = \left| \int_0^\infty h_r(\tau) [x(t_k - \tau) - \tilde{x}(t_k - \tau)] d\tau \right|,$$

where $h_r(\tau)$ is the impulse response of $p^r/A^*(p)$, and

$$\tilde{x}(t) := \begin{cases} \vdots & \\ x(t_{k-1}) & \text{if } t_{k-1} \leq t < t_k, \\ x(t_k) & \text{if } t_k \leq t < t_{k+1}, \\ \vdots & \end{cases}$$

Thus, we can bound the term $\|\hat{\Phi}^* - \Phi^*\|_\infty$ by

$$\begin{aligned} \|\hat{\Phi}^* - \Phi^*\|_\infty &= \max_{k \in \{1, \dots, N\}} \sum_{r=1}^n \left| \left\{ \frac{p^r}{A^*(p)} x(t) \right\}_{t=t_k} - \frac{p^r}{A^*(p)} x(t_k) \right| \\ &\leq \max_{0 \leq t \leq t_N} |x(t) - \tilde{x}(t)| \sum_{r=1}^n \int_0^\infty |h_r(\tau)| d\tau \\ &\leq \max_{k \in \{1, \dots, N\}} (t_k - t_{k-1}) \max_{t_1 \leq t \leq t_N} \left| \frac{dx(t)}{dt} \right| \sum_{r=1}^n \int_0^\infty |h_r(\tau)| d\tau. \end{aligned} \quad (6.37)$$

Since $G^*(p)$ is assumed to be strictly proper, the derivative in (6.37) can be bounded by a constant times $\|x\|_\infty$. Moreover, the integrals are also finite (and independent of t), since the transfer functions $p^r/A^*(p)$ are asymptotically stable. Therefore, if $\max_{k \in \{1, \dots, N\}} (t_k - t_{k-1}) \rightarrow 0$ then we have $\|\hat{\Phi}^* - \Phi^*\|_\infty \rightarrow 0$. \square

Remark 6.5. It is possible to make $\hat{\Phi}^* - \Phi^*$ exactly equal to zero by choosing the instrument vector to match exactly with the noiseless version of Φ_j with θ_j instead of θ^* . Although leading to a consistent estimator, this choice of instrument vector is not optimal in terms of efficiency, and therefore it is not considered in the extended SRIVC algorithm.

Consistency of refined IV methods for continuous-time systems in closed-loop

Refined instrumental variable methods have been broadly used for identification of continuous-time systems in both open and closed-loop. However, the theoretical properties of these methods are still yet to be fully understood for closed-loop settings. In this chapter we address the consistency of the SRIVC estimator and its closed-loop variant CLSRIVC when they are applied on data that are generated from two different standard feedback loop configurations. The theoretical findings are verified numerically through extensive simulation tests.

7.1 Introduction

Closed-loop system identification has mostly been studied in the discrete-time domain [51, 227]. As in the open-loop case, estimators based on instrumental variables for closed-loop identification have also been subject of extensive research. Methods based on this approach have been developed and discussed for discrete-time models in closed-loop [74, 76, 78], and an extension for discrete-time Hammerstein models can be found in [128]. Nevertheless, the user may be interested in a continuous-time description of the system for several reasons, such as a greater physical insight on the system parameters and controller tuning, ease of incorporating non-uniform and fast sampling, and more flexibility in the proposed model structure [63].

For continuous-time models, one option is to disregard the feedback and perform direct system identification using the refined instrumental variables method for continuous-time systems (RIVC), or its simplified embodiment, the SRIVC method [255]. Although the SRIVC estimator has been analyzed extensively in this thesis in open-loop, several questions still remain unsolved regarding its statistical properties in closed-loop settings. On the other hand, various algorithms have been proposed for incorporating the controller into the estimation process. A bias-eliminated least-squares method was derived in [64], while [75] explored instrumental variable techniques. These were later revisited in [258] and [248], where

the SRIVC and RIVC methods are implemented using simulation data for obtaining a model estimate in closed-loop. A three-stage estimator was detailed in [249], and another two-stage algorithm was introduced in [134]. An extensive review of instrumental variable methods for closed-loop continuous-time system identification can be found in [77], in which the authors introduce a closed-loop variant of SRIVC called CLSRIVC that requires the knowledge of the continuous-time controller. This estimator is argued to be asymptotically unbiased for white and colored output noise [61, pp. 154-156], although no proof of this statement is given.

There are two settings that are commonly addressed in the contributions cited above, which differ in the nature of the controller (continuous-time or discrete-time). This chapter analyzes the consistency of the SRIVC and CLSRIVC methods when they are applied with continuous-time and discrete-time controllers in the feedback loop, and clarifies the conditions that are sufficient for obtaining consistent estimates of the continuous-time system under study. The main results of this chapter can be found in [91]. In summary,

- we prove that the SRIVC and CLSRIVC estimators are not consistent when the data is generated from a fully continuous-time feedback loop and only sampled data are available as measurements. The lack of consistency of the CLSRIVC estimator in this scenario has been overlooked in the literature (see, e.g., [77, p. 154]) where it is claimed that CLSRIVC is asymptotically unbiased;
- we illustrate that the lack of asymptotic unbiasedness of the methods can be mitigated via oversampling techniques;
- we show that, under some technical assumptions, the SRIVC estimator is generically consistent when the controller in the feedback loop is in discrete-time. We also prove that consistency is lost if there is a direct path in the loop or if the noise is not white. In such cases, analogously to some fully discrete-time settings [235, 237], the discrete-time equivalent of the model estimate is shown to be biased towards the negative inverse of the controller;
- we derive a variant of the CLSRIVC estimator that is proven to be generically consistent when the controller works in discrete-time and is known in advance; and
- we provide empirical support of the theoretical results via extensive Monte Carlo simulations.

This chapter is organized as follows. Section 7.2 introduces the two feedback settings we investigate. Sections 7.3 and 7.4 present the main contributions of this chapter, namely, the asymptotic properties of the SRIVC and CLSRIVC methods for the feedback loops in Settings 1 and 2 respectively, as shown in Figure 7.1. Simulations that support our analysis can be found in Section 7.5, while conclusions are drawn in Section 7.6. The appendix of this chapter contains a technical lemma used in the theoretical analysis.

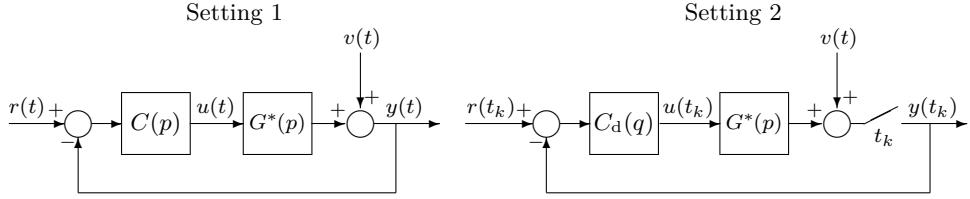


Figure 7.1: Block diagrams for the closed-loop settings 1 and 2.

7.2 Closed-loop settings

Consider a linear, time-invariant, causal, single-input single-output, continuous-time system $G^*(p) = B^*(p)/A^*(p)$, where

$$\begin{aligned} B^*(p) &= b_{m^*}^* p^{m^*} + \cdots + b_1^* p + b_0^*, \\ A^*(p) &= a_{n^*}^* p^{n^*} + \cdots + a_1^* p + 1, \end{aligned}$$

with the polynomial degrees satisfying $n^* \geq m^*$. The zero-order hold (ZOH) equivalent discrete-time system of $G^*(p)$ is denoted by $G_d^*(q)$. Given data, the main goal is to estimate the parameters of the transfer function $G^*(p)$ that are described by the vector $\boldsymbol{\theta}^*$ in (2.6). In this chapter we study the generic consistency of the SRIVC and CLSRIVC estimators for different closed-loop settings. The two settings we study are presented in Figure 7.1; both scenarios assume that the reference signal is reconstructed from a ZOH device (i.e., it is constant between samples). Our results naturally extend to signals with an FOH intersample behavior, provided we adapt the persistence of excitation requirements similar to what was done for Corollary 4.1. A detailed description of each setting is given below.

- *Setting 1:* The first setting assumes that the controller is a continuous-time transfer function $C(p) = F(p)/L(p)$ and that the system output is corrupted by a measurement noise $v(t)$. Since we typically do not have access to the noise behavior between the sampling time instants of the output, it is easiest to model the noise $v(t)$ as being constant between samples and that the noise samples $v(t_k)$ that perturb the output are a realization of some stochastic process. This assumption is used in [77], as well as in [94, 134, 233], where hybrid Box-Jenkins models are estimated. Other contributions have not used this assumption [64, 75], although limited analysis has been done in such context. In this work, the variability of the output noise is assumed to be discrete in time (or equivalently, the disturbance signal is assumed to be constant between the samples of the output). We note that our theoretical results could be extended for the continuous-time white noise case, but extra care must be taken to avoid the theoretical problems associated to continuous-time white noise [9].

As mentioned in Chapter 2, we can express $u(t_k)$ and $y(t_k)$ as a function of the signals $r(t_k)$ and $v(t_k)$ as follows:

$$\begin{aligned} u(t_k) &= S_{uo}^*(p)r(t_k) - S_{uo}^*(p)v(t_k), \\ y(t_k) &= T_o^*(p)r(t_k) + S_o^*(p)v(t_k), \end{aligned}$$

where we have used the standard notation of sensitivity functions [97]

$$T_o^*(p) := \frac{G^*(p)C(p)}{1 + G^*(p)C(p)}, \quad S_o^*(p) := \frac{1}{1 + G^*(p)C(p)}, \quad S_{uo}^*(p) := \frac{C(p)}{1 + G^*(p)C(p)}. \quad (7.1)$$

Keep in mind that we employ the notation of discrete-time signals filtered by continuous-time filters that is explained in Remark 2.5.

- *Setting 2:* The second setting is hybrid in nature, as it considers a discrete-time controller $C_d(q) = F_d(q)/L_d(q)$ acting upon a continuous-time system. The controller output $u(t_k)$ is a discrete-time signal that is reconstructed through a ZOH device before acting on $G^*(p)$, and the noisy system output is sampled before being fed back to the controller. This framework is used in [248, 249, 258].

In this setting the input and output samples satisfy

$$u(t_k) = S_{uo}^*(q)r(t_k) - S_{uo}^*(q)v(t_k), \quad (7.2a)$$

$$y(t_k) = T_o^*(q)r(t_k) + S_o^*(q)v(t_k), \quad (7.2b)$$

where the sensitivity functions are in the same form as those in (7.1), but with $G^*(p)$ and $C(p)$ being replaced by $G_d^*(q)$ and $C_d(q)$, respectively. Note that the CLSRIVC estimator, as introduced in [77] and defined in Section 2.3, cannot be directly implemented in this setting as it requires knowing a *continuous-time* controller in advance.

Both settings assume that the samples $\{u(t_k), y(t_k)\}_{k=1}^N$ (and possibly $\{r(t_k)\}_{k=1}^N$) are available as data for identification, and we define n_l^* and n_f^* as the number of poles and zeros, respectively, of $C(p)$ or $C_d(q)$. The general assumptions that we require for our analysis are:

Assumption 7.1. The true system $B^*(p)/A^*(p)$ is proper ($n^* \geq m^*$) and asymptotically stable with $A^*(p)$ and $B^*(p)$ being coprime. The closed-loop system is also asymptotically stable, i.e., the zeros of $1 + G^*(p)C(p)$ are in the open left half plane for Setting 1, and the zeros of $1 + G_d^*(q)C_d(q)$ are in the open unit disk for Setting 2.

Assumption 7.2. The external reference $r(t_k)$ and disturbance $v(t_l)$ are stationary and mutually independent for all integers k and l .

Assumption 7.3. The degrees n and m of the polynomials in the model satisfy $\min(n - n^*, m - m^*) = 0$. For Setting 1 we also require $n - n^* \leq n^* - m^* + n_l^* - n_f^*$, whereas for Setting 2 we also assume that $n - n^* \leq n^* - m^*$.

The stability of the system in open loop is needed for the filtering steps in both algorithms. Although it is shown later in Chapter 8 that this condition is no longer required if a more involved filtering technique is implemented, we do not pursue an analysis of such refinements here. We also note that the assumption on the model order is similar to Assumption 4.5 for the consistency analysis of the SRIVC estimator in Chapter 4, and it is used in the same manner in the forthcoming proofs.

This chapter provides a comprehensive analysis on the generic consistency of the SRIVC and CLSRIVC estimators in closed-loop for Settings 1 and 2.

7.3 Analysis of the SRIVC and CLSRIVC estimators for Setting 1

In this section we study the consistency of both refined instrumental variable methods for Setting 1. Most of our attention will be set on the properties of the CLSRIVC estimator, as they have not been carefully addressed in the literature prior to this thesis. With regards to the SRIVC estimator, the following result follows directly from Corollary 4.2.

Theorem 7.1 (Generic inconsistency of the SRIVC estimator for Setting 1). *Consider the SRIVC estimator with iterations (2.35), filtered regressor (2.32), filtered instrument (2.33), and filtered output (2.34), and suppose Assumptions 7.1 to 7.3 hold. Moreover, assume that $\mathbb{E}\{\hat{\varphi}_f(t_k)\varphi_f^\top(t_k)\}$ is generically non-singular with respect to the system and model parameters. The SRIVC estimator, provided it converges in iterations for all N sufficiently large, is not generically consistent under Setting 1 when only sampled data are available as measurements.*

Proof. The result follows as a direct consequence of Corollary 4.2, as it is well known that a continuous-time controller $C(p)$, as in Setting 1, produces a smooth system input that cannot be exactly reconstructed from a ZOH or FOH device for the computation of the filtered regressor vector. \square

Note that the generic non-singularity condition of the modified normal matrix $\mathbb{E}\{\hat{\varphi}_f(t_k)\varphi_f^\top(t_k)\}$ can be proven under some technical requirements. We do not pursue this proof here as it is similar to the open-loop case, which has been presented in Chapter 4.

The analysis for the CLSRIVC estimator is more intricate, and therefore we will state some additional technical assumptions that are related to the stability of the estimates, the reference excitation, and the sampling frequency. More precisely, the assumptions that we consider are the following:

Assumption 7.4. For all $j \geq 1$, the zeros of $A_j(p)$ have strictly negative real parts, $n \geq m$, with $A_j(p)$ and $B_j(p)$ being coprime. The model of the closed-loop system is also asymptotically stable, i.e., the zeros of $A_j(p)L(p) + B_j(p)F(p)$ have strictly negative real parts.

Assumption 7.5. The external reference is persistently exciting of order no less than $2n + n_l^*$ if $n > m$ and no less than $2n + n_l^* + 1$ if $n = m$.

Assumption 7.6. The sampling frequency is larger than twice the largest imaginary part of the zeros of $A_j(p)(A^*(p)L(p) + B^*(p)F(p))$.

We shall define some signals related to the filtered regressor vector before we present the main results of this section. Similar to the decomposition in (4.3), the filtered regressor vector can be written as

$$\varphi_f(t_k) = \tilde{\varphi}_f^r(t_k) + \Delta(t_k) - \mathbf{v}_f(t_k),$$

where the disturbance-free, interpolation error-free regressor vector is given by

$$\tilde{\varphi}_f^r(t_k) = \left[\frac{-pT_o^*(p)}{A_j(p)} r(t_k), \dots, \frac{-p^n T_o^*(p)}{A_j(p)} r(t_k), \frac{S_{uo}^*(p)}{A_j(p)} r(t_k), \dots, \frac{p^m S_{uo}^*(p)}{A_j(p)} r(t_k) \right]^\top, \quad (7.3)$$

and $\Delta(t_k)$ is a vector that contains the interpolation errors that arise from constructing the filtered disturbance-free derivatives of the output. The entries of $\Delta(t_k)$ are given by the difference between the noise-free version of the regressor vector and $\tilde{\varphi}_f^r(t_k)$ in (7.3), i.e.,

$$\Delta_i(t_k) = \begin{cases} \frac{p^i T_o^*(p)}{A_j(p)} r(t_k) - \frac{p^i}{A_j(p)} \{T_o^*(p)r(t)\}_{t=t_k} & \text{if } i = 1, \dots, n, \\ 0 & \text{otherwise.} \end{cases} \quad (7.4)$$

The contribution of the disturbance to the filtered regressor is given by

$$\mathbf{v}_f(t_k) = \left[\frac{p}{A_j(p)} \{S_o^*(p)v(t)\}_{t=t_k}, \dots, \frac{p^n}{A_j(p)} \{S_o^*(p)v(t)\}_{t=t_k}, \right. \\ \left. \frac{1}{A_j(p)} \{S_{uo}^*(p)v(t)\}_{t=t_k}, \dots, \frac{p^m}{A_j(p)} \{S_{uo}^*(p)v(t)\}_{t=t_k} \right]^\top.$$

In Lemma 7.1 we show that the modified normal matrix being inverted in each iteration of the CLSRIVC method is generically non-singular when N tends to infinity. This property is essential for establishing asymptotic properties of the converging point of the method, since it ensures that the converging point is well-defined in all but possibly some rare and isolated cases.

Lemma 7.1. Consider the CLSRIVC estimator with iterations (2.35), filtered regressor vector (2.32), filtered instrument vector (2.40) and filtered output (2.34). Suppose Assumptions 7.1 to 7.6 hold and the condition

$$\|\mathbb{E}\{\hat{\varphi}_f(t_k)\Delta^\top(t_k)\}\|_2 < \sigma_{\min}(\mathbb{E}\{\hat{\varphi}_f(t_k)\tilde{\varphi}_f^r(t_k)\}) \quad (7.5)$$

is satisfied, where $\tilde{\varphi}_f^r(t_k)$ and $\Delta(t_k)$ are given by (7.3) and (7.4), respectively. Then, the modified normal matrix $\mathbb{E}\{\hat{\varphi}_f(t_k)\varphi_f^r(t_k)\}$ of the CLSRIVC method is generically non-singular with respect to the system denominator and model parameters.

Proof. The filtered instrument vector and the disturbance-free, interpolation error-free regressor vector can be written as

$$\hat{\varphi}_f(t_k) = \mathbf{S}(-B_j, A_j) \frac{S_{uo,j}(p)}{A_j^2(p)} \mathbf{r}_{n+m}(t_k) \quad \text{and} \quad \tilde{\varphi}_f^r(t_k) = \mathbf{S}(-B^*, A^*) \frac{S_{uo}^*(p)}{A_j(p)A^*(p)} \mathbf{r}_{n+m}(t_k) \quad (7.6)$$

respectively, where $\mathbf{S}(-B_j, A_j)$ and $\mathbf{S}(-B^*, A^*)$ are $(n+m+1) \times (n+m+1)$ Sylvester matrices constructed using the coefficients of the respective B and A polynomials (see, e.g., (4.14) for their construction). These matrices are non-singular under Assumptions 7.4 and 7.1 due to Lemma A3.1 of [210]. The vector $\mathbf{r}_{n+m}(t_k)$ contains the derivatives of the reference signal, that is,

$$\mathbf{r}_{n+m}(t_k) = [p^{n+m}, p^{n+m-1}, \dots, 1]^\top r(t_k). \quad (7.7)$$

As $N \rightarrow \infty$, the ergodic lemma in [205, Lemma 3.1] permits expressing the modified normal matrix of the CLSRIVC estimator as

$$\mathbb{E}\{\hat{\varphi}_f(t_k)\varphi_f^\top(t_k)\} = \underbrace{\mathbf{S}(-B_j, A_j)\Phi\mathbf{S}^\top(-B^*, A^*)}_{\mathbb{E}\{\hat{\varphi}_f(t_k)\tilde{\varphi}_f^{r\top}(t_k)\}} + \mathbb{E}\{\hat{\varphi}_f(t_k)\Delta^\top(t_k)\} + \mathbb{E}\{\hat{\varphi}_f(t_k)\mathbf{v}_f^\top(t_k)\}, \quad (7.8)$$

where

$$\Phi := \mathbb{E} \left\{ \frac{S_{uo,j}(p)}{A_j^2(p)} \mathbf{r}_{n+m}(t_k) \frac{S_{uo}^*(p)}{A_j(p)A^*(p)} \mathbf{r}_{n+m}^\top(t_k) \right\}. \quad (7.9)$$

By the same arguments found in the proof of Lemma 4.1, the last expectation in (7.8) is equal to zero under Assumption 7.2. Thus, (7.8) simplifies to

$$\mathbb{E}\{\hat{\varphi}_f(t_k)\varphi_f^\top(t_k)\} = \mathbf{S}(-B_j, A_j)\Phi\mathbf{S}^\top(-B^*, A^*) + \mathbb{E}\{\hat{\varphi}_f(t_k)\Delta^\top(t_k)\}. \quad (7.10)$$

Theorem 5.1 of [41] ensures the (generic) non-singularity of $\mathbb{E}\{\hat{\varphi}_f(t_k)\varphi_f^\top(t_k)\}$ provided that condition (7.5) holds and that Φ is generically non-singular. The generic non-singularity of Φ follows from the arguments below. Let the set of parameters that describe $A_j(p)$ and $B_j(p)$ be defined as

$$\Omega = \{[a_1, \dots, a_n, b_0, \dots, b_m]^\top \in \mathbb{R}^{n+m+1} : A_j(p) \text{ is a stable polynomial}\}.$$

Define Φ^* as the matrix Φ in (7.9) with $A_j(p) = A^*(p)$ and $B_j(p) = B^*(p)$, i.e.,

$$\Phi^* := \mathbb{E} \left\{ \frac{S_{uo}^*(p)}{A^{*2}(p)} \mathbf{r}_{n+m}(t_k) \frac{S_{uo}^*(p)}{A^{*2}(p)} \mathbf{r}_{n+m}^\top(t_k) \right\}.$$

Note that all the transfer functions that form the elements of Φ^* are proper due to the second condition in Assumption 7.3. The denominator polynomial of $S_{uo}^*(p)/A^{*2}(p)$ has a degree of $n + \max(n + n_l^*, m + n_f^*) = 2n + n_l^*$. Then, by the same procedure as in the proof of Lemma 4.2, the matrix Φ^* can be shown to be positive definite

if the reference signal is persistently exciting of order $2n + n_l^*$ in case $n > m$, or $2n + n_l^* + 1$ in case $n = m$. Next, by leveraging the frequency domain description of covariance expressions, an arbitrary entry of Φ in (7.9) can be written as

$$\begin{aligned}\Phi_{il} &= \mathbb{E} \left\{ \frac{p^{n+m+1-i} S_{uo,j}(p)}{A_j^2(p)} r(t_k) \frac{p^{n+m+1-l} S_{uo}^*(p)}{A_j(p) A^*(p)} r(t_k) \right\} \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{C_1(e^{i\omega})}{A_{d,j}^2(e^{i\omega}) Q_{d,j}(e^{i\omega})} \frac{C_2(e^{-i\omega})}{A_{d,j}(e^{-i\omega}) A_d^*(e^{-i\omega}) Q_d^*(e^{-i\omega})} dF_r(\omega),\end{aligned}$$

where $F_r(\omega)$ is the spectral distribution of the reference signal, and $A_{d,j}, Q_{d,j}, A_d^*$ and Q_d^* are the denominator polynomials of the respective ZOH equivalents. The same steps as in the proof of Lemma 4.4 show that the entries of Φ are analytic functions of every parameter in Ω . Thus, invoking Lemma 4.6 and Corollary 4.4 leads to the conclusion that the modified normal matrix $\mathbb{E}\{\hat{\varphi}_f(t_k)\varphi_f^\top(t_k)\}$ is generically non-singular. \square

Next, we examine the consistency of the CLSRIVC estimator in Theorem 7.2.

Theorem 7.2 (Generic inconsistency of the CLSRIVC estimator for Setting 1). *Consider the CLSRIVC estimator with iterations (2.35), filtered regressor vector (2.32), filtered instrument vector (2.40) and filtered output (2.34). Suppose Assumptions 7.1 to 7.6 hold and the condition in (7.5) is satisfied. The CLSRIVC estimator, provided it converges in iterations for all N sufficiently large, is not generically consistent under Setting 1 when only sampled data are available as measurements.*

Proof. If we denote $\bar{\theta}^N$ as the converging point of the CLSRIVC estimator (with associated transfer function model $\bar{G}_N(p) = \bar{B}(p)/\bar{A}_N(p)$), it must satisfy

$$\bar{\theta}^N = \left[\frac{1}{N} \sum_{k=1}^N \hat{\varphi}_f(t_k, \bar{\theta}^N) \varphi_f^\top(t_k, \bar{\theta}^N) \right]^{-1} \left[\frac{1}{N} \sum_{k=1}^N \hat{\varphi}_f(t_k, \bar{\theta}^N) y_f(t_k, \bar{\theta}^N) \right]. \quad (7.11)$$

The ergodic lemmas in [205, Lemma 3.1] and [210, Lemma A4.3] permit writing (7.11) when N tends to infinity as

$$\mathbb{E}\{\hat{\varphi}_f(t_k, \bar{\theta}) \varphi_f^\top(t_k, \bar{\theta})\}^{-1} \mathbb{E}\{\hat{\varphi}_f(t_k, \bar{\theta})(y_f(t_k, \bar{\theta}) - \varphi_f^\top(t_k, \bar{\theta})\bar{\theta})\} = \mathbf{0},$$

where $\bar{\theta} = \lim_{N \rightarrow \infty} \bar{\theta}^N$. This parameter vector has an associated transfer function model $\bar{G}(p) = \bar{B}(p)/\bar{A}(p)$. The non-singularity of the modified normal matrix implies that

$$\begin{aligned}&\mathbb{E}\{\hat{\varphi}_f(t_k, \bar{\theta})(y_f(t_k, \bar{\theta}) - \varphi_f^\top(t_k, \bar{\theta})\bar{\theta})\} = \mathbf{0} \\ \iff &\mathbb{E}\left\{\hat{\varphi}_f(t_k, \bar{\theta})\left(\frac{1}{\bar{A}(p)}y(t_k) + \frac{\bar{a}_n p^n + \dots + \bar{a}_1 p}{\bar{A}(p)}y(t_k) - \frac{\bar{B}(p)}{\bar{A}(p)}u(t_k)\right)\right\} = \mathbf{0} \\ \iff &\mathbb{E}\{\hat{\varphi}_f(t_k, \bar{\theta})(y(t_k) - \bar{G}(p)u(t_k))\} = \mathbf{0} \quad (7.12) \\ \iff &\mathbb{E}\{\hat{\varphi}_f(t_k, \bar{\theta})(\{G^*(p)u(t)\}_{t=t_k} - \bar{G}(p)u(t_k))\} + \mathbb{E}\{\hat{\varphi}_f(t_k, \bar{\theta})v(t_k)\} = \mathbf{0}.\end{aligned}$$

Since the additive noise on the output is independent of the reference signal under Assumption 7.2, it follows that $\mathbb{E}\{\hat{\varphi}_f(t_k, \bar{\theta})v(t_k)\} = \mathbf{0}$. Hence, the converging point of the CLSRIVC estimator satisfies

$$\mathbb{E}\{\hat{\varphi}_f(t_k, \bar{\theta}) (\{G^*(p)u(t)\}_{t=t_k} - \bar{G}(p)u(t_k))\} = \mathbf{0} \quad (7.13)$$

$$\iff \mathbb{E}\left\{\hat{\varphi}_f(t_k, \bar{\theta}) \underbrace{\left(\{G^*(p)S_{uo}^*(p)r(t)\}_{t=t_k} - \bar{G}(p)\{S_{uo}^*(p)r(t)\}_{t=t_k}\right)}_{=:\varepsilon(t_k, \bar{\theta})}\right\} = \mathbf{0}. \quad (7.14)$$

Now, we introduce a reference-dependent term

$$\varepsilon_r(t_k, \bar{\theta}) := \{\bar{G}(p)S_{uo}^*(p)r(t)\}_{t=t_k} - \bar{G}(p)\{S_{uo}^*(p)r(t)\}_{t=t_k}. \quad (7.15)$$

Then, $\varepsilon(t_k, \bar{\theta})$ in (7.14) becomes

$$\begin{aligned} \varepsilon(t_k, \bar{\theta}) &= G^*(p)S_{uo}^*(p)r(t_k) - \bar{G}(p)S_{uo}^*(p)r(t_k) + \varepsilon_r(t_k, \bar{\theta}) \\ &= \frac{B^*(p)\bar{A}(p) - \bar{B}(p)A^*(p)}{A^*(p)\bar{A}(p)}S_{uo}^*(p)r(t_k) + \varepsilon_r(t_k, \bar{\theta}) \\ &= \frac{S_{uo}^*(p)}{\bar{A}(p)A^*(p)}\mathbf{r}_{n+m}^\top(t_k)\mathbf{h} + \varepsilon_r(t_k, \bar{\theta}), \end{aligned}$$

where $\mathbf{r}_{n+m}(t_k)$ is defined in (7.7) and \mathbf{h} is a vector of dimension $(n + m + 1)$ containing the coefficients of $B^*(p)\bar{A}(p) - \bar{B}(p)A^*(p)$ in descending order of degree. Using the fact that the filtered instrument vector can be written in terms of $\mathbf{r}_{n+m}(t_k)$ as in (7.6), with $\mathbf{S}(-\bar{B}, \bar{A})$ being non-singular given Assumption 7.4, the condition in (7.14) can be expressed as

$$\underbrace{\mathbb{E}\left\{\frac{\bar{S}_{uo}(p)}{\bar{A}^2(p)}\mathbf{r}_{n+m}(t_k)\frac{S_{uo}^*(p)}{\bar{A}(p)A^*(p)}\mathbf{r}_{n+m}^\top(t_k)\right\}\mathbf{h}}_{=:\bar{\Phi}} + \underbrace{\mathbb{E}\left\{\frac{\bar{S}_{uo}(p)}{\bar{A}^2(p)}\mathbf{r}_{n+m}(t_k)\varepsilon_r(t_k, \bar{\theta})\right\}}_{=:\bar{\Psi}} = \mathbf{0}.$$

We can show that $\bar{\Phi}$ is generically non-singular by using the same reasoning as in the non-singularity proof of Φ in Lemma 7.1. Note that $\bar{\Psi}$ is reference-dependent and is in general not equal to zero due to $\varepsilon_r(t_k, \bar{\theta}) \neq 0$. Since the matrix $\mathbf{S}(-\bar{B}, \bar{A})$ is non-singular under Assumption 7.4, we have

$$\mathbf{h} = -\bar{\Phi}^{-1}\bar{\Psi} \neq \mathbf{0}. \quad (7.16)$$

This in turn implies that

$$\frac{\bar{B}(p)}{\bar{A}(p)} = \frac{B^*(p)}{A^*(p)} - \frac{H(p)}{\bar{A}(p)A^*(p)},$$

i.e., the unique converging point does not correspond to the true parameter vector. We then conclude that the CLSRIVC estimator is generically not consistent. \square

7.3.1 Reducing the bias of refined IV methods in closed-loop via oversampling

Just as for the SRIVC estimator in the open-loop case, the lack of consistency of the CLSRIVC estimator is due to the unknown intersample behavior of the plant input $u(t)$. However, if the intersample behavior of the input were to be fully known and the filtered regressor vector of the CLSRIVC algorithm were to be computed as

$$\varphi_f(t_k, \theta_j) = \left[\frac{-p}{A_j(p)} y(t_k), \dots, \frac{-p^n}{A_j(p)} y(t_k), \left\{ \frac{1}{A_j(p)} u(t) \right\}_{t=t_k}, \dots, \left\{ \frac{p^m}{A_j(p)} u(t) \right\}_{t=t_k} \right]^\top, \quad (7.17)$$

then the converging point would satisfy (cf. (7.14))

$$\mathbb{E} \left\{ \hat{\varphi}_f(t_k, \bar{\theta}) \left(\{G^*(p)S_{uo}^*(p)r(t)\}_{t=t_k} - \{\bar{G}(p)S_{uo}^*(p)r(t)\}_{t=t_k} \right) \right\} = \mathbf{0}.$$

Thus, $\mathbf{h} = \mathbf{0}$ and the CLSRIVC estimator would be generically consistent for this case. The uncertainty related to the intersample behavior of the input can be mitigated by choosing a faster sampling rate; however, from a theoretical standpoint, this can only alleviate the bias to a certain degree. If the input can be sampled at a faster rate than the output, then computing a more precise approximation of (7.17) using the computational aspects covered in Section 6.5 constitutes a useful refinement of the CLSRIVC estimator. In other words, oversampling procedures reduce the norm of the reference-dependent term $\varepsilon_r(t_k, \bar{\theta})$ that contributes to the bias of the CLSRIVC estimator.

Remark 7.1. Note that (7.17) has also been proposed in (6.6) for the extended SRIVC estimator in Chapter 6. This estimator can be shown to be generically consistent in the closed-loop scenario under more stringent conditions, similar to the ones established in the next section. We do not pursue the complete analysis of the extended SRIVC estimator here, as it follows similar lines to Theorems 4.1, 6.1 and 7.1.

In summary, we have proven in this section that the SRIVC and CLSRIVC estimators are not consistent when there is a continuous-time controller in the feedback loop and only sampled data are available as measurements. However, these estimators may benefit from a more sophisticated filtering that makes a more adequate assumption on the intersample behavior of the input signal $u(t)$.

7.4 Analysis of the SRIVC and CLSRIVC estimators for Setting 2

We now analyze the statistical properties of the refined instrumental variable methods when a discrete-time controller is present in the loop instead of a continuous-time controller. This means that the intersample behavior of the system input is known

to be a ZOH¹, which is beneficial for the implementation and performance of the algorithms.

The analysis below relies on leveraging the filtered external reference and filtered disturbance, which are defined as

$$\tilde{r}(t_k) := \frac{C_d(q)}{1 + G_d^*(q)C_d(q)} r(t_k), \quad \tilde{v}(t_k) := \frac{C_d(q)}{1 + G_d^*(q)C_d(q)} v(t_k). \quad (7.18)$$

The assumptions we consider are the following:

Assumption 7.7. The transfer function $G_d^*(q)C_d(q)$ is strictly proper.

Assumption 7.8. For all $j \geq 1$, all the zeros of the j -th iteration of the model denominator estimate $A_j(p)$ have strictly negative real parts, with $A_j(p)$ and $B_j(p)$ being coprime.

Assumption 7.9. The external reference is persistently exciting of order no less than $2n + n_f^*$ if $n > m$ and no less than $2n + n_f^* + 1$ if $n = m$.

Assumption 7.10. The sampling frequency is more than twice the largest positive imaginary part of the zeros of $A_j(p)A^*(p)$.

We must decompose the filtered regressor and instrument vectors similarly to the derivation in the previous section. For the sake of simplifying the notation, we will repeat the notation used in Section 7.3 despite some vectors having different definitions. The filtered regressor vector can now be written as

$$\varphi_f(t_k) = \tilde{\varphi}_f^r(t_k) + \Delta(t_k) - \mathbf{v}_f(t_k), \quad (7.19)$$

where

$$\tilde{\varphi}_f^r(t_k) = \left[\frac{-pB^*(p)}{A_j(p)A^*(p)} \tilde{r}(t_k), \dots, \frac{-p^n B^*(p)}{A_j(p)A^*(p)} \tilde{r}(t_k), \frac{1}{A_j(p)} \tilde{r}(t_k), \dots, \frac{p^m}{A_j(p)} \tilde{r}(t_k) \right]^\top. \quad (7.20)$$

The entries of $\Delta(t_k)$ are given by

$$\Delta_i(t_k) = \begin{cases} \frac{p^i G^*(p)}{A_j(p)} \tilde{r}(t_k) - \frac{p^i}{A_j(p)} [G_d^*(q) \tilde{r}(t_k)] & \text{if } i = 1, \dots, n, \\ 0 & \text{otherwise,} \end{cases} \quad (7.21)$$

and the contribution of the disturbance $v(t_k)$ in the filtered regressor is decomposed into two parts $\mathbf{v}_f(t_k) = \mathbf{v}_{f1}(t_k) + \mathbf{v}_{f2}(t_k)$, which are given by

$$\mathbf{v}_{f1}(t_k) = \left[\frac{p}{A_j(p)}, \dots, \frac{p^n}{A_j(p)}, 0, \dots, 0 \right]^\top \frac{1}{1 + G_d^*(q)C_d(q)} v(t_k) \quad (7.22)$$

¹It is worth noting that results when the intersample behavior is an FOH can be readily obtained from our approach, with only some minor modifications on the persistence of excitation condition of the external reference (see, e.g., Corollary 4.1).

and

$$\mathbf{v}_{f2}(t_k) = \left[0, \dots, 0, \frac{1}{A_j(p)} \tilde{v}(t_k), \dots, \frac{p^m}{A_j(p)} \tilde{v}(t_k) \right]^\top. \quad (7.23)$$

On the other hand, the filtered instrument vector can be decomposed as the sum of a filtered version of the external reference and disturbance:

$$\hat{\varphi}_f(t_k) = \hat{\varphi}_f^r(t_k) - \hat{\mathbf{v}}_f(t_k), \quad (7.24)$$

where

$$\hat{\varphi}_f^r(t_k) = \left[\frac{-pB_j(p)}{A_j^2(p)} \tilde{r}(t_k), \dots, \frac{-p^n B_j(p)}{A_j^2(p)} \tilde{r}(t_k), \frac{1}{A_j(p)} \tilde{r}(t_k), \dots, \frac{p^m}{A_j(p)} \tilde{r}(t_k) \right]^\top, \quad (7.25)$$

and the contribution of the disturbance is again decomposed as $\hat{\mathbf{v}}_f(t_k) = \hat{\mathbf{v}}_{f1}(t_k) + \mathbf{v}_{f2}(t_k)$, with $\mathbf{v}_{f2}(t_k)$ being given by (7.23) and

$$\hat{\mathbf{v}}_{f1}(t_k) = \left[\frac{-pB_j(p)}{A_j^2(p)} \tilde{v}(t_k), \dots, \frac{-p^n B_j(p)}{A_j^2(p)} \tilde{v}(t_k), 0, \dots, 0 \right]^\top. \quad (7.26)$$

7.4.1 Consistency analysis of the SRIVC estimator for Setting 2

Sufficient conditions for generic consistency of the SRIVC estimator under Setting 2 are derived in Theorem 7.3.

Theorem 7.3 (Generic consistency of the SRIVC estimator for Setting 2). *Consider the SRIVC estimator described by the iterations in (2.35), filtered regressor (2.32), filtered instrument (2.33), and filtered output (2.34), and suppose Assumptions 7.1 to 7.3 and 7.7 to 7.10 hold. Furthermore, assume that $v(t_k)$ is a white noise stochastic process. Then, the following statements are true:*

1. *The matrix $\mathbb{E}\{\hat{\varphi}_f(t_k)\varphi_f^\top(t_k)\}$ of the SRIVC method is generically non-singular with respect to the system and model denominator provided that the condition*

$$\begin{aligned} & \left\| \mathbb{E}\left\{ \hat{\varphi}_f^r(t_k) \Delta^\top(t_k) + \hat{\mathbf{v}}_{f1}(t_k) \mathbf{v}_f^\top(t_k) + \mathbf{v}_{f2}(t_k) \mathbf{v}_{f1}^\top(t_k) \right\} \right\|_2 \\ & < \sigma_{\min} \left(\mathbb{E} \left\{ \hat{\varphi}_f^r(t_k) \hat{\varphi}_f^{r\top}(t_k) + \mathbf{v}_{f2}(t_k) \mathbf{v}_{f2}^\top(t_k) \right\} \right) \end{aligned} \quad (7.27)$$

holds, where $\tilde{\varphi}_f^r(t_k)$, $\Delta(t_k)$, $\mathbf{v}_{f1}(t_k)$, $\mathbf{v}_{f2}(t_k)$, $\hat{\varphi}_f^r(t_k)$, and $\hat{\mathbf{v}}_{f1}(t_k)$ are defined as in (7.20), (7.21), (7.22), (7.23), (7.25), and (7.26), respectively.

2. *If (7.27) is satisfied and the iterations of the SRIVC estimator converge for all N sufficiently large to, say, $\bar{\theta}^N$, then the true parameter θ^* is the unique converging point of $\bar{\theta}^N$ as the sample size tends to infinity.*

Proof of Statement 1: By exploiting the decompositions in (7.19) and (7.24), the modified normal matrix $\mathbb{E}\{\hat{\varphi}_f(t_k)\varphi_f^\top(t_k)\}$ can be written as

$$\begin{aligned}\mathbb{E}\{\hat{\varphi}_f(t_k)\varphi_f^\top(t_k)\} &= \mathbb{E}\left\{\hat{\varphi}_f^r(t_k)(\tilde{\varphi}_f^r(t_k) + \Delta(t_k))^\top + \hat{\mathbf{v}}_f(t_k)\mathbf{v}_f^\top(t_k)\right\} \\ &\quad - \mathbb{E}\left\{\hat{\mathbf{v}}_f(t_k)(\tilde{\varphi}_f^r(t_k) + \Delta(t_k))^\top + \hat{\varphi}_f^r(t_k)\mathbf{v}_f^\top(t_k)\right\}.\end{aligned}$$

Due to the fact that the external reference and disturbance are statistically independent by Assumption 7.2, the second expectation in the equation above is equal to zero. Thus,

$$\mathbb{E}\{\hat{\varphi}_f(t_k)\varphi_f^\top(t_k)\} = \mathbb{E}\left\{\hat{\varphi}_f^r(t_k)\tilde{\varphi}_f^{r\top}(t_k) + \mathbf{v}_{f2}(t_k)\mathbf{v}_{f2}^\top(t_k)\right\} \quad (7.28a)$$

$$+ \mathbb{E}\left\{\hat{\varphi}_f^r(t_k)\Delta^\top(t_k) + \hat{\mathbf{v}}_{f1}(t_k)\mathbf{v}_f^\top(t_k) + \mathbf{v}_{f2}(t_k)\mathbf{v}_{f1}^\top(t_k)\right\}. \quad (7.28b)$$

The approach for proving Statement 1 is similar to the proof of Lemma 7.1. Under condition (7.27), the matrix in (7.28b) is small enough (in 2-norm) to not affect the non-singularity of the matrix $\mathbb{E}\{\hat{\varphi}_f(t_k)\varphi_f^\top(t_k)\}$. Theorem 5.1 of [41] ensures the generic non-singularity of the modified normal matrix provided that the condition in (7.27) holds and that $\mathbb{E}\{\hat{\varphi}_f^r(t_k)\tilde{\varphi}_f^{r\top}(t_k) + \mathbf{v}_{f2}(t_k)\mathbf{v}_{f2}^\top(t_k)\}$ is generically non-singular.

We now prove the generic non-singularity of $\mathbb{E}\{\hat{\varphi}_f^r(t_k)\tilde{\varphi}_f^{r\top}(t_k) + \mathbf{v}_{f2}(t_k)\mathbf{v}_{f2}^\top(t_k)\}$. It is known that (cf. (7.6))

$$\hat{\varphi}_f^r(t_k) = \mathbf{S}(-B_j, A_j) \frac{1}{A_j^2(p)} \tilde{\mathbf{r}}_{n+m}(t_k) \quad \text{and} \quad \tilde{\varphi}_f^r(t_k) = \mathbf{S}(-B^*, A^*) \frac{1}{A_j(p)A^*(p)} \tilde{\mathbf{r}}_{n+m}(t_k), \quad (7.29)$$

where

$$\tilde{\mathbf{r}}_{n+m}(t_k) := \left[p^{n+m}, p^{n+m-1}, \dots, 1 \right]^\top \tilde{r}(t_k), \quad (7.30)$$

and the Sylvester matrices $\mathbf{S}(-B^*, A^*)$ and $\mathbf{S}(-B_j, A_j)$ are non-singular due to Lemma A3.1 of [210], since these polynomials are coprime by Assumptions 7.1 and 7.8 respectively. Substituting (7.29) into the first summand in (7.28a) yields

$$\mathbb{E}\left\{\hat{\varphi}_f^r(t_k)\tilde{\varphi}_f^{r\top}(t_k) + \mathbf{v}_{f2}(t_k)\mathbf{v}_{f2}^\top(t_k)\right\} = \mathbf{S}(-B_j, A_j) \Phi \mathbf{S}(-B^*, A^*) + \mathbb{E}\{\mathbf{v}_{f2}(t_k)\mathbf{v}_{f2}^\top(t_k)\},$$

where

$$\Phi = \mathbb{E}\left\{\frac{1}{A_j^2(p)} \tilde{\mathbf{r}}_{n+m}(t_k) \frac{1}{A_j(p)A^*(p)} \tilde{\mathbf{r}}_{n+m}^\top(t_k)\right\}. \quad (7.31)$$

We show the generic non-singularity of $\mathbb{E}\{\hat{\varphi}_f^r(t_k)\tilde{\varphi}_f^{r\top}(t_k) + \mathbf{v}_{f2}(t_k)\mathbf{v}_{f2}^\top(t_k)\}$ by proving that every entry of such matrix is a real analytic function in the joint variables of the model and system denominator, and that it is positive definite when evaluated at the true system parameters. The analyticity of each entry of the

matrix in question again follows from applying the ideas in Lemma 4.4 to each entry in (7.31) and $\mathbb{E}\{\mathbf{v}_{f2}(t_k)\mathbf{v}_{f2}^\top(t_k)\}$ and using the fact that the sum and multiplication of real analytic functions is real analytic. For the positive definiteness requirement, we require that $\tilde{r}(t_k)$ is persistently exciting of order no less than $2n$ for Lemma 4.2 to be applied. Since

$$\frac{C_d(q)}{1 + G_d^*(q)C_d(q)} = \frac{A_d^*(q)F_d(q)}{A_d^*(q)L_d(q) + B_d^*(q)F_d(q)},$$

we find that at least n zeros of this transfer function are minimum-phase (namely, the zeros of $A_d^*(q)$). Thus, Assumption 7.9 ensures that the spectrum of $\tilde{r}(t_k)$ has at least $2n$ distinct frequency lines, and therefore Lemma 4.2 ensures that Φ is positive definite when evaluated at the true parameters. Since $\mathbb{E}\{\mathbf{v}_{f2}(t_k)\mathbf{v}_{f2}^\top(t_k)\}$ is positive semi-definite by construction, we have that $\mathbb{E}\{\hat{\varphi}_f^r(t_k)\hat{\varphi}_f^{r\top}(t_k) + \mathbf{v}_{f2}(t_k)\mathbf{v}_{f2}^\top(t_k)\}$ is positive definite when evaluated at the true parameters. With this, by Lemma A2.3 of [210] and its corollary we obtain that $\mathbb{E}\{\hat{\varphi}_f^r(t_k)\tilde{\varphi}_f^{r\top}(t_k) + \mathbf{v}_{f2}(t_k)\mathbf{v}_{f2}^\top(t_k)\}$ is generically non-singular. This concludes the proof of Statement 1. \square

Proof of Statement 2: The ergodic results in [205, Lemma 3.1] and [210, Lemma A4.3] permit us to write (2.35), at the converging point and as N tends to infinity, as

$$\bar{\boldsymbol{\theta}} = \mathbb{E}\{\hat{\varphi}_f(t_k, \bar{\boldsymbol{\theta}})\varphi_f^\top(t_k, \bar{\boldsymbol{\theta}})\}^{-1}\mathbb{E}\{\hat{\varphi}_f(t_k, \bar{\boldsymbol{\theta}})y_f(t_k, \bar{\boldsymbol{\theta}})\}, \quad (7.32)$$

where $\bar{A}(p)$ and $\bar{B}(p)$ are the corresponding denominator and numerator polynomials of the converging point as N tends to infinity. These polynomials are coprime by Assumption 7.8. Since the matrix inverse in (7.32) is non-singular by Statement 1, the following condition must hold:

$$\mathbb{E}\{\hat{\varphi}_f(t_k, \bar{\boldsymbol{\theta}})(y_f(t_k, \bar{\boldsymbol{\theta}}) - \varphi_f^\top(t_k, \bar{\boldsymbol{\theta}})\bar{\boldsymbol{\theta}})\} = \mathbf{0}. \quad (7.33)$$

From (2.32) and (2.34), after some computations, we find that

$$y_f(t_k, \bar{\boldsymbol{\theta}}) - \varphi_f^\top(t_k, \bar{\boldsymbol{\theta}})\bar{\boldsymbol{\theta}} = y(t_k) - \frac{\bar{B}(p)}{\bar{A}(p)}u(t_k) = (G^*(p) - \bar{G}(p))u(t_k) + v(t_k).$$

By Lemma 7.2 in the appendix of this chapter, we have $\mathbb{E}\{\hat{\varphi}_f(t_k, \bar{\boldsymbol{\theta}})v(t_k)\} = \mathbf{0}$ and therefore (7.33) is equivalent to

$$\mathbb{E}\{\hat{\varphi}_f(t_k, \bar{\boldsymbol{\theta}})[G^*(p) - \bar{G}(p)](\tilde{r}(t_k) - \tilde{v}(t_k))\} = \mathbf{0},$$

where the notation in (7.18) has been used. Note that

$$\begin{aligned} [G^*(p) - \bar{G}(p)](\tilde{r}(t_k) - \tilde{v}(t_k)) &= \frac{\bar{A}(p)B^*(p) - A^*(p)\bar{B}(p)}{A^*(p)\bar{A}(p)}(\tilde{r}(t_k) - \tilde{v}(t_k)) \\ &= \frac{1}{A^*(p)\bar{A}(p)}(\tilde{\mathbf{r}}_{n+m}^\top(t_k) - \tilde{\mathbf{v}}_{n+m}^\top(t_k))\mathbf{h}, \end{aligned} \quad (7.34)$$

where $\tilde{\mathbf{v}}_{n+m}(t_k)$ has the same form as (7.30) but $\tilde{r}(t_k)$ is replaced by $\tilde{v}(t_k)$. The vector \mathbf{h} is formed by the coefficients of the polynomial $\bar{A}(p)B^*(p) - \bar{B}(p)A^*(p)$. On the other hand, we can express the filtered instrument vector as

$$\hat{\varphi}_f(t_k, \bar{\boldsymbol{\theta}}) = \mathbf{S}(-\bar{B}, \bar{A}) \frac{1}{\bar{A}^2(p)} (\tilde{\mathbf{r}}_{n+m}(t_k) - \tilde{\mathbf{v}}_{n+m}(t_k)). \quad (7.35)$$

Since $\mathbf{S}(-\bar{B}, \bar{A})$ is non-singular and the external reference signal $r(t_k)$ and disturbance $v(t_k)$ are independent by Assumption 7.2, we can use (7.34) and (7.35) to write condition (7.33) as $\bar{\Phi}\mathbf{h} = \mathbf{0}$, with

$$\bar{\Phi} = \mathbb{E} \left\{ \frac{1}{\bar{A}^2(p)} \tilde{\mathbf{r}}_{n+m}(t_k) \frac{1}{\bar{A}(p)A^*(p)} \tilde{\mathbf{r}}_{n+m}^\top(t_k) \right\} + \mathbb{E} \left\{ \frac{1}{\bar{A}^2(p)} \tilde{\mathbf{v}}_{n+m}(t_k) \frac{1}{\bar{A}(p)A^*(p)} \tilde{\mathbf{v}}_{n+m}^\top(t_k) \right\}. \quad (7.36)$$

The generic non-singularity of $\bar{\Phi}$ follows from the same procedure as in the proof of Statement 1. Thus, we must have $\mathbf{h} = \mathbf{0}$, implying $\bar{A}(p)B^*(p) - \bar{B}(p)A^*(p) = 0$ or equivalently $\bar{G}(p) = G^*(p)$, i.e., $\boldsymbol{\theta}^*$ is the unique limiting point. \square

Remark 7.2. The condition in (7.27) is sufficient but not necessary for the generic non-singularity of the modified normal matrix $\mathbb{E}\{\hat{\varphi}_f(t_k)\varphi_f^\top(t_k)\}$. This requirement can always be satisfied as long as the signal-to-noise ratio (SNR) between the external reference and the disturbance is high enough, and the interpolation error $\Delta(t_k)$ is not significant. Intuitively, the SRIVC estimator is generically consistent for loops with no direct feedthrough if the SNR is high enough and the sampling frequency is not too low.

Corollary 7.1 shows that generic consistency is lost if there is a direct path in the closed-loop, or if the noise is not white. We assume that the modified normal matrix is non-singular; this fact can be proven generically provided that a condition similar to (4.7) is satisfied. We do not pursue that proof here.

Corollary 7.1. *Assume that the SRIVC iterations converge for all N sufficiently large and that $\mathbb{E}\{\hat{\varphi}_f(t_k)\varphi_f^\top(t_k)\}$ is non-singular. Under Assumptions 7.1 to 7.3 and 7.8 to 7.10, if $v(t_k)$ is not white noise, or if Assumption 7.7 is not satisfied, then generically $\bar{G}(p) \neq G^*(p)$ as N goes to infinity.*

Proof. Following the same steps as in the proof of Statement 2 of Theorem 7.3, we find that the limiting point of the SRIVC estimator must satisfy, as N tends to infinity,

$$\mathbb{E} \{ \hat{\varphi}_f(t_k, \bar{\boldsymbol{\theta}}) [G^*(p) - \bar{G}(p)] u(t_k) \} = -\mathbb{E} \{ \hat{\varphi}_f(t_k, \bar{\boldsymbol{\theta}}) v(t_k) \}. \quad (7.37)$$

By exploiting (7.34), (7.35) and (7.43), we can rewrite (7.37) as

$$\bar{\Phi}\mathbf{h} = \mathbb{E} \left\{ \left[\frac{\bar{\mathbf{B}}_{n+m}^\text{d}(q) C_\text{d}(q)}{\bar{A}_\text{d}^2(q)(1 + G_\text{d}^*(q)C_\text{d}(q))} v(t_k) \right] v(t_k) \right\},$$

where $\bar{\mathbf{B}}_{n+m}^d(q)$ is defined in (7.42). Since $\bar{\Phi}$ is generically non-singular and the right-hand side is different from zero in general, we have $\mathbf{h} \neq \mathbf{0}$, i.e., $\bar{G}(p) \neq G^*(p)$ as N tends to infinity. \square

Remark 7.3. A well-known fact in both non-parametric and parametric discrete-time system identification is that the model is biased towards the negative inverse of the controller in some closed-loop settings [96, 109]. Since this aspect is intrinsic to the feedback loop configuration, it is not surprising to find similar relationships in hybrid settings with discrete-time controllers and continuous-time systems. In fact, again assuming the non-singularity of $\mathbb{E}\{\hat{\varphi}_f(t_k)\varphi_f^\top(t_k)\}$, the converging point must satisfy, as N tends to infinity,

$$\begin{aligned} & \mathbb{E}\left\{\frac{1}{A^2(p)}\tilde{\mathbf{r}}_{n+m}(t_k)[G^*(p)-\bar{G}(p)]\tilde{r}(t_k)\right\} + \mathbb{E}\left\{\frac{1}{\bar{A}^2(p)}\tilde{\mathbf{v}}_{n+m}(t_k)[G^*(p)-\bar{G}(p)]\tilde{v}(t_k)\right\} \\ &= \mathbb{E}\left\{\frac{1}{\bar{A}^2(p)}\tilde{\mathbf{v}}_{n+m}(t_k)\tilde{v}(t_k)\right\}. \end{aligned}$$

After some computations, this equation can be rewritten as

$$\mathbb{E}\left\{\frac{1}{\bar{A}^2(p)}\tilde{\mathbf{r}}_{n+m}(t_k)[G^*(p)-\bar{G}(p)]\tilde{r}(t_k)\right\} + \mathbb{E}\left\{\frac{1}{\bar{A}^2(p)}\tilde{\mathbf{v}}_{n+m}(t_k)[-C_d^{-1}(q)-\bar{G}_d(q)]\tilde{v}(t_k)\right\} = \mathbf{0}. \quad (7.38)$$

Thus, we find that the converging transfer function of the SRIVC estimator will tend towards $G^*(p)$ for high SNR. Indeed, for that case, the first expectation in (7.38) dictates the value of $\bar{G}(p)$ and by (7.34) we reach $\bar{\Phi}_1 \mathbf{h} \approx \mathbf{0}$, where $\bar{\Phi}_1$ is the first expectation in (7.36). Hence, we obtain $\bar{G}(p) \approx G^*(p)$. On the contrary, for low SNR the second expectation in (7.38) impacts the value of $\bar{G}(p)$ the most, and the ZOH equivalent of $\bar{G}(p)$ will tend to $-1/C_d(q)$ by a similar reasoning as before. This property is exemplified later in Section 7.5.

7.4.2 The CLSRIVC estimator for Setting 2 and its consistency analysis

The CLSRIVC estimator, as introduced in [61], assumes that a *continuous-time* controller is known in advance for the implementation of the algorithm. Therefore, it is only suitable for Setting 1. In this work we introduce and analyze an extension of this estimator which is valid for Setting 2 that has a filtered instrument given by

$$\hat{\varphi}_f(t_k, \theta_j) = \left[\frac{-pB_j(p)}{A_j^2(p)}, \dots, \frac{-p^n B_j(p)}{A_j^2(p)}, \frac{1}{A_j(p)}, \dots, \frac{p^m}{A_j(p)} \right]^\top S_{uo,j}(q)r(t_k), \quad (7.39)$$

where $S_{uo,j}(q) = C_d(q)/[1 + G_{d,j}(q)C_d(q)]$. The other signals of interest at each iteration ($\varphi_f(t_k, \theta_j)$ and $y_f(t_k, \theta_j)$) have the same form as the standard SRIVC estimator, that is, they are given by (2.32) and (2.34), and the iterations are computed by (2.35). We prove the generic consistency of this variant of the CLSRIVC estimator for Setting 2 in Theorem 7.4.

Theorem 7.4 (Generic consistency of the CLSRIVC estimator for Setting 2). *Consider the CLSRIVC estimator described by the iterations in (2.35) with filtered instrument vector (7.39), and suppose Assumptions 7.1 to 7.3 and 7.7 to 7.10 hold. Then, the following statements are true:*

1. *The matrix $\mathbb{E}\{\hat{\varphi}_f(t_k)\varphi_f^\top(t_k)\}$ is generically non-singular with respect to the system and model denominator provided that the condition*

$$\|\mathbb{E}\{\hat{\varphi}_f(t_k)\Delta^\top(t_k)\}\|_2 < \sigma_{\min}(\mathbb{E}\{\hat{\varphi}_f(t_k)\hat{\varphi}_f^{r\top}(t_k)\}) \quad (7.40)$$

holds, where $\tilde{\varphi}_f^r(t_k)$, $\Delta(t_k)$ and $\hat{\varphi}_f(t_k)$ are defined as in (7.20), (7.21), and (7.39), respectively.

2. *If (7.40) is satisfied and the iterations of the CLSRIVC estimator converge for all N sufficiently large to, say, $\bar{\theta}^N$, then the true parameter θ^* is the unique converging point of $\bar{\theta}^N$ as the sample size tends to infinity.*

Proof of Statement 1: Statement 1 follows from a similar reasoning to the proof of Lemma 7.1. This time, the modified normal matrix can be also written as (7.10) but with

$$\Phi = \mathbb{E} \left\{ \left[S_{uo,j}(q) \frac{1}{A_j^2(p)} \mathbf{r}_{n+m}(t_k) \right] \left[S_{uo}^*(q) \frac{1}{A_j(p)A^*(p)} \mathbf{r}_{n+m}^\top(t_k) \right] \right\}. \quad (7.41)$$

Note that the continuous-time transfer functions involved in the construction of the elements of Φ are well defined when evaluating at the true parameters due to Assumption 7.3. Furthermore, the matrix Φ in (7.41) is equivalent to (7.31) when they are evaluated at the true parameters. Thus, its positive-definiteness follows from the persistence of excitation condition in Assumption 7.9, and the analyticity of the entries of Φ again follow from Lemma 4.4. Thus, by Lemma A2.3 of [210] and its corollary we conclude that the modified normal matrix $\mathbb{E}\{\hat{\varphi}_f(t_k)\varphi_f^\top(t_k)\}$ is generically non-singular. \square

Proof of Statement 2: The first lines of the proof of Statement 2 are the same as in the proof of Theorem 7.2 until Equation (7.13). Since the input $u(t)$ is perfectly reconstructed from a ZOH device, we have $\{G^*(p)u(t)\}_{t=t_k} = G^*(p)u(t_k)$ and thus

$$\mathbb{E}\{\hat{\varphi}_f(t_k, \bar{\theta}) ([G^*(p) - \bar{G}(p)]\tilde{r}(t_k))\} = \mathbf{0}.$$

Exploiting the expressions in (7.34) and (7.35), together with the non-singularity of $\mathbf{S}(-\bar{B}, \bar{A})$, yields

$$\mathbb{E}\left\{\frac{1}{\bar{A}^2(p)}\tilde{\mathbf{r}}_{n+m}(t_k)\frac{1}{\bar{A}(p)A^*(p)}\tilde{\mathbf{r}}_{n+m}^\top(t_k)\right\}\mathbf{h} = \mathbf{0}.$$

The matrix above is known to be generically non-singular by leveraging the same ideas as in Statement 1. Therefore we must have $\mathbf{h} = \mathbf{0}$, implying $\bar{A}(p)B^*(p) - \bar{B}(p)A^*(p) = 0$ or equivalently $\bar{G}(p) = G^*(p)$, i.e., θ^* is the unique limiting point. \square

In summary, we have found that the SRIVC and CLSRIVC estimators are generically consistent for Setting 2. Contrary to Setting 1, the estimators do not require oversampling to alleviate bias issues since the intersample behavior of the system input is exactly known in advance. The SRIVC estimator requires stricter conditions in the model and noise statistics for consistency. When these are not satisfied, the discrete-time equivalent of the estimate is biased towards the negative inverse of the controller.

7.5 Simulations

We now corroborate the theoretical results via extensive Monte Carlo simulations. This section is divided in two parts: the consistency tests (that is, the verification of Theorems 7.1, 7.2, 7.3 and 7.4), and an example of the bias behavior of the SRIVC estimator for Setting 2.

7.5.1 Consistency tests

We consider the continuous-time system

$$G^*(p) = \frac{-0.25p + 0.5}{0.5p^2 + 0.707p + 1},$$

where the true parameters are given by $\boldsymbol{\theta}^* = [a_1^*, a_2^*, b_0^*, b_1^*]^\top = [0.707, 0.5, 0.5, -0.25]^\top$. The sampling period is chosen as $h = 0.1[\text{s}]$ for all the tests, and the reference signal is white noise of variance 1 interpolated with a ZOH device. The output noise is also white, of variance 0.01 for both settings. The consistency of the SRIVC and CLSRIVC estimators is investigated for an increasing sample size, and they are compared with their extended versions that incorporate the exact intersample behavior of the input (i.e., the SRIVC-x estimator of Chapter 6 and the CLSRIVC estimator with filtered regressor vector given by (7.17), which is labeled as the CLSRIVC-x estimator). The sample size N is adjusted from 200 to 200000 in a logarithmic scale, where a total of 40 different sample sizes are considered. Three hundred Monte Carlo runs are performed for each sample size, and the sample mean for each parameter is recorded. The estimators are initialized with the LSSVF estimator [242]. The maximum number of iterations of each method is set to 200, and the tolerance factor is set to $\epsilon = 10^{-7}$. The extended estimators consider input data that is sampled $S = 100$ times faster than the output, and the computation of the filtered regressor vector (7.17) follows Section 6.5. The following controllers are used for Settings 1 and 2, respectively:

$$C(p) = 1.896 \cdot 10^{-4} + \frac{0.7278}{p}, \quad C_d(q) = \frac{0.416(q - 0.7452)}{q - 1}.$$

Figure 7.2 empirically shows that the SRIVC and CLSRIVC estimators are not consistent under Setting 1, as the sample means of the estimated parameters do not

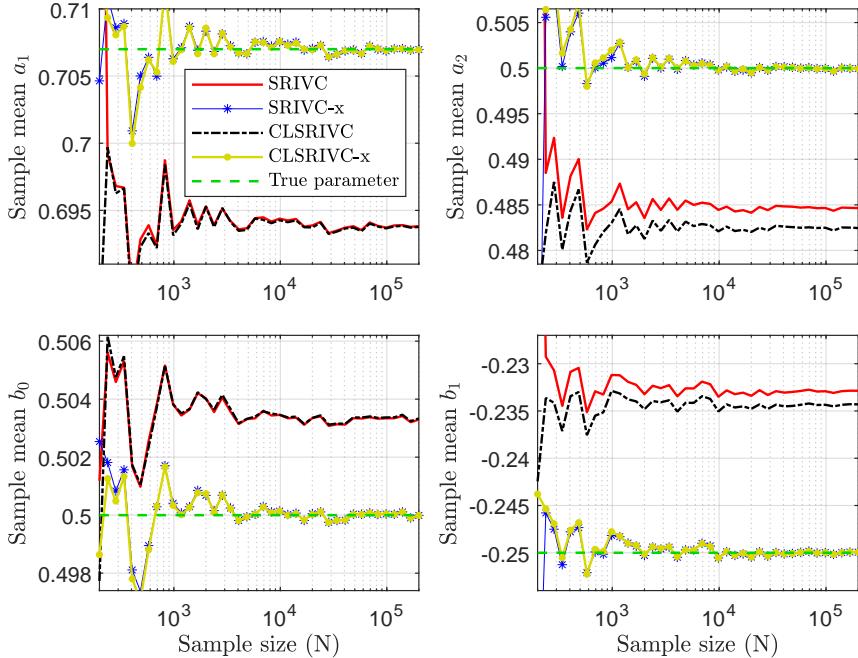


Figure 7.2: Mean of the estimated parameters for an increasing sample size, Setting 1.

approach the true parameter values when the sample size increases. This observation is aligned with Theorems 7.1 and 7.2. In contrast, the SRIVC-x and CLSRIVC-x estimators do not exhibit noticeable bias for large sample size.

Note that the bias of SRIVC and CLSRIVC can also be reduced by jointly decreasing the sampling period of the input and output signals. This fact can be verified theoretically for the CLSRIVC estimator from observing that the term $\varepsilon_r(t_k, \theta)$ in (7.15) (and therefore $\bar{\Psi}$ in (7.16)) is less significant for smaller sampling periods. In fact, if $N = 200000$ and $h = 0.02[\text{s}]$ instead of $0.1[\text{s}]$, we obtain the sample means presented in Table 7.1. The bias has been reduced as anticipated, but we should not expect it to decay to zero. Identification with small sampling periods might explain why some authors have considered the CLSRIVC estimator to be asymptotically unbiased when analyzing its performance through simulations [77]. We clarify that this property does not hold from a theoretical standpoint for any sampling period greater than zero.

In another simulation study, we test the consistency of the methods under Setting 2. The intersample behavior of $u(t_k)$ is now correctly specified in the regressor vector of the algorithms, thus yielding generically consistent estimators according to Theorems 7.3 and 7.4. Indeed, Figures 7.3 and 7.4 show that the sample mean of each parameter estimate converges to the true value for an increasing sample size, and that the variance of those estimates is decreasing. This provides

Table 7.1: Sample mean of the model parameters obtained with 300 Monte Carlo runs with $N = 200000$ and $h = 0.02[\text{s}]$, Setting 1.

Method	a_1 (0.707)	a_2 (0.5)	b_0 (0.5)	b_1 (-0.25)
SRIVC	0.7044	0.4966	0.5008	-0.2469
CLSRIVC	0.7044	0.4965	0.5008	-0.2469

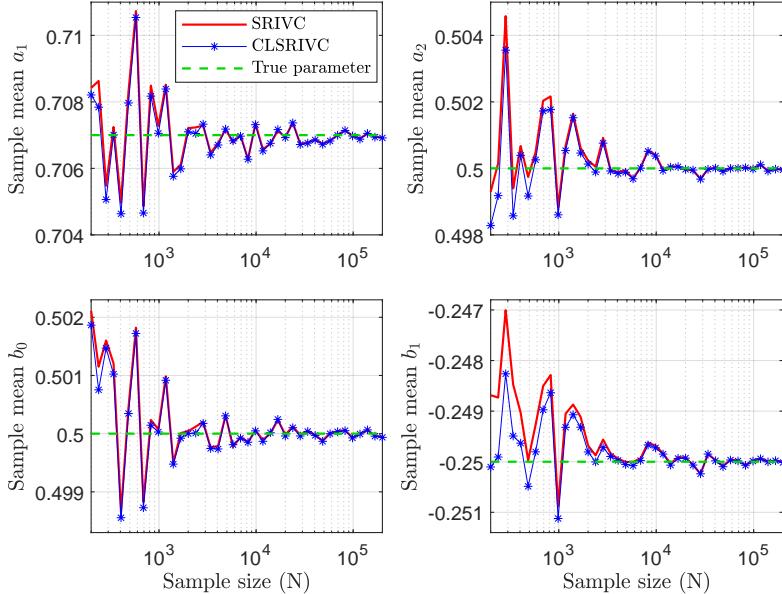


Figure 7.3: Mean of the estimated parameters for an increasing sample size, Setting 2.

empirical evidence to the consistency results of Section 7.4. Note that although the SRIVC estimator is consistent in this simulation test, its consistency in closed-loop is limited to cases where the output noise is white and there is no direct feedthrough in the loop. In contrast, the CLSRIVC estimator can handle colored noise and biproper transfer functions without a loss in generic consistency.

7.5.2 Bias behavior of SRIVC in Setting 2

To conclude this section, we study the bias of the SRIVC estimator when $G_d^*(q)C_d(q)$ is biproper. Based on (7.38) we expect a bias of the ZOH equivalent of the model estimates toward $-1/C_d(q)$. We verify this property for the following system and controller:

$$G^*(p) = \frac{-0.3p + 1}{p + 1}, \quad C_d(q) = \frac{2.15(q - 0.9949)}{q - 1},$$

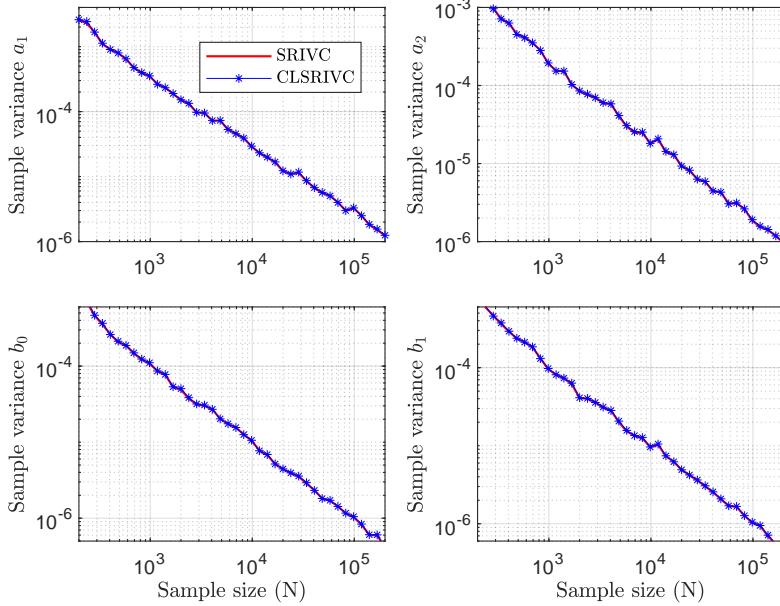


Figure 7.4: Variance of the estimated parameters for an increasing sample size, Setting 2.

with $h = 0.1[\text{s}]$ and $N = 100000$. Forty values that are logarithmically-spaced and ranging from 10^{-3} to 10^3 are considered for the SNR, which is defined here as the quotient of the reference variance σ_r^2 and the disturbance variance σ_v^2 . An average model is computed among 300 Monte Carlo runs for each SNR value, and Figure 7.5 reveals the normalized bias behavior of the discrete-time equivalent with respect to the SNR in the feedback loop. As expected, the ZOH equivalent of the mean model from the SRIVC method approaches $-1/C_d(q)$ for low SNR and there is a transition towards unbiasedness for high SNR. This curve, now obtained in a hybrid closed-loop scenario, is aligned with what is known for indirect non-parametric estimators in closed-loop (see, e.g., Figure 4 of [237]).

7.6 Conclusions

This chapter has focused on the generic consistency properties of the SRIVC and CLSRIVC estimators for two well-established closed-loop settings. The theoretical results show that the SRIVC and CLSRIVC estimators are generically consistent when there is a discrete-time controller in the loop, and this consistency is lost if a continuous-time controller is implemented. The bias for the fully continuous-time case may be partially overcome by oversampling methods that compute a more adequate regressor vector, which is directly related to the findings in Chapter 6. The

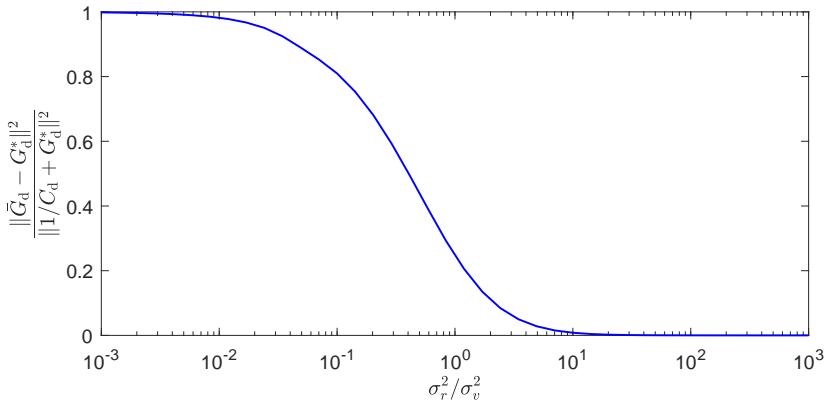


Figure 7.5: Normalized bias of the SRIVC estimator versus SNR, Setting 2.

consistency of the CLSRIVC estimator has been shown to extend to more general scenarios than that of the SRIVC estimator, such as when there is a direct path in the loop or when the output disturbance is not white. The bias of the SRIVC estimator has also been characterized in the presence of a discrete-time control feedback for such scenarios. Monte Carlo simulations have verified the theoretical findings.

Appendix

7.A Supplementary material

Lemma 7.2. *Assume that $\{v(t_k)\}$ is a white noise stochastic process and that Assumptions 7.2 and 7.7 hold. Then, $\mathbb{E}\{\hat{\varphi}_f(t_k, \bar{\theta})v(t_k)\} = \mathbf{0}$.*

Proof. We can write the filtered instrument in discrete-time form as

$$\hat{\varphi}_f(t_k, \bar{\theta}) = \mathbf{S}(-\bar{B}, \bar{A}) \frac{1}{\bar{A}_d^2(q)} \mathbf{u}_{n+m}^d(t_k),$$

where the polynomial $\bar{A}_d(q)$ is the denominator of the ZOH equivalent of $\bar{B}(p)/\bar{A}(p)$, and the vector $\mathbf{u}_{n+m}^d(t_k)$ is given by

$$\mathbf{u}_{n+m}^d(t_k) = \underbrace{\left[\bar{B}_d^{n+m}(q), \bar{B}_d^{n+m-1}(q), \dots, \bar{B}_d^0(q) \right]^\top}_{=: \bar{\mathbf{B}}_{n+m}^d(q)} u(t_k), \quad (7.42)$$

with $\bar{B}_d^i(q)/\bar{A}_d^2(q)$ being the ZOH equivalent of $p^i/\bar{A}^2(p)$, $i = 0, 1, \dots, n+m$. Thus,

$$\hat{\varphi}_f(t_k, \bar{\theta}) = \mathbf{S}(-\bar{B}, \bar{A}) \frac{\bar{\mathbf{B}}_{n+m}^d(q)}{\bar{A}_d^2(q)} (\tilde{r}(t_k) - \tilde{v}(t_k)).$$

By exploiting Assumption 7.2, we can write $\mathbb{E}\{\hat{\varphi}_f(t_k, \bar{\theta})v(t_k)\}$ as

$$\mathbb{E}\{\hat{\varphi}_f(t_k, \bar{\theta})v(t_k)\} = -\mathbf{S}(-\bar{B}, \bar{A}) \mathbb{E} \left\{ \left[\frac{\bar{\mathbf{B}}_{n+m}^d(q)C_d(q)}{\bar{A}_d^2(q)(1 + G_d^*(q)C_d(q))} v(t_k) \right] v(t_k) \right\}. \quad (7.43)$$

Since the transfer function $G_d^*(q)C_d(q)$ is strictly proper by Assumption 7.7, we have $n > m$ (which means that $\bar{\mathbf{B}}_{n+m}^d(q)/\bar{A}_d^2(q)$ is formed solely by strictly proper transfer functions), or $C_d(q)$ is strictly proper. Either way, the transfer function on the left-hand side of (7.43) is stable and strictly proper and therefore it can be decomposed as

$$\frac{\bar{\mathbf{B}}_{n+m}^d(q)C_d(q)}{\bar{A}_d^2(q)(1 + G_d^*(q)C_d(q))} = \mathbf{g}_1 q^{-1} + \mathbf{g}_2 q^{-2} + \dots, \quad (7.44)$$

for some constant vectors $\mathbf{g}_1, \mathbf{g}_2, \dots$. Since $v(t_k)$ is assumed to be white noise, the expectation in (7.43) is given by

$$\mathbb{E}\{\hat{\varphi}_f(t_k, \bar{\theta})v(t_k)\} = -\mathbf{S}(-\bar{B}, \bar{A})\mathbb{E}\left\{[(\mathbf{g}_1q^{-1} + \mathbf{g}_2q^{-2} + \dots)v(t_k)]v(t_k)\right\} = \mathbf{0}, \quad (7.45)$$

which proves the result. \square

Remark 7.4. Note that, in general, $\mathbb{E}\{\hat{\varphi}_f(t_k, \bar{\theta})v(t_k)\} \neq \mathbf{0}$ when $G_d^*(q)C_d(q)$ is biproper or when $\{v(t_k)\}$ is not white. This can be seen from the proof of Lemma 7.2, as the constant term of the transfer function in (7.44) is different from zero if $G_d^*(q)C_d(q)$ is biproper, and (7.45) does not hold in general if $v(t_k)$ is not white noise.

Chapter 8

Refined IV methods for unstable continuous-time systems

One of the limitations of the refined instrumental variable methods is that they require filtering the input and output by a stable model at each iteration. As seen in Chapter 5, their convergence is also compromised when the stabilizing step is performed. In this chapter we present extensions of the refined instrumental variable methods for continuous-time systems in open and closed loop that admit the estimation of unstable models. Some asymptotic properties of these variants are covered, and several implementation aspects with regards to the intersample behavior of the input are studied. The proposed methods are tested via extensive numerical simulations.

8.1 Introduction

The greatest difficulty surrounding the estimation of unstable linear and time-invariant models is that most system identification algorithms, directly or indirectly, use predictors or regressors that become unstable during the identification process. An illustrating example of this problem is encountered in the SRIVC method [255] when input and output data are used for identification. This estimator has been proven to be generically consistent and asymptotically efficient in Chapter 4 of this work under mild conditions, including the assumption that the true system and the model at every iteration of the algorithm are asymptotically stable. The stability of the model iterates is needed for the filtering process that must be performed: since the input and output signals are prefiltered by auxiliary estimates of the true system, such a procedure becomes severely ill-conditioned if these estimates are unstable.

Naturally, identification of unstable systems must be done in a closed-loop setting. If the controller is linear and time-invariant and known *a priori*, then an attractive approach is to apply the CLRIVC or CLSRIVC methods as proposed in [77]. In such case, the idea is to formulate the instrument vector by prefiltering the reference

signal with estimates obtained in the previous iteration of the closed-loop transfer function multiplied by the open-loop denominator transfer function. Again, these methods are not suitable for identifying unstable systems. Several extensions of closed-loop methods suffer from the same issue: the authors in [160] developed an instrumental variable-based method for estimating Hammerstein-Wiener models, while the CLRIVC method was adapted in [32] to estimate continuous-time systems with non-uniformly sampled data. The work in [233] proposes using instrumental variables for closed-loop identification in an errors-in-variables framework, and [163] explores recursive identification. Among other findings, in Chapters 6 and 7 we extend the applicability of both SRIVC and CLSRIVC estimators for arbitrary inputs. All the methods mentioned above are based on refined instrumental variables and thus cannot be directly applied if the open-loop system is unstable.

In contrast to the aforementioned work, the main contributions of this chapter are:

- we present an extension of the LSSVF estimator that can serve as an initialization step for any subsequent identification method for unstable continuous-time systems;
- we propose an extension to the SRIVC and RIVC methods for estimating continuous-time unstable systems in output error and hybrid BJ model structures, respectively. These refinements admit ZOH and FOH input signals, as well as continuous-time multisines and arbitrary inputs whose intersample behaviors are known;
- we discuss how to handle unstable systems for indirect closed-loop continuous-time system identification using the CLSRIVC method and its variant for discrete-time controllers introduced in Chapter 7;
- we provide variants of the SRIVC, RIVC and CLSRIVC estimators that can deal with plants with integral action. This type of systems is a limit case that is not adequately covered in the extensions previously proposed; and
- we test the proposed methods via extensive numerical simulations.

The rest of this chapter is organized as follows. In Section 8.2 we present the framework of our work and state the identification problem. Section 8.3 contains the main contribution of this chapter, namely, the derivation of the RIVC and SRIVC estimators that admit unstable continuous-time models. Section 8.4 studies the applicability of our approach to indirect closed-loop continuous-time system identification methods, while Section 8.5 covers identification with integral action in the plant. In Section 8.6 extensive simulations are performed to test the proposed estimators under different experimental conditions, and we provide concluding comments in Section 8.7. Supplementary material can be found in the appendix of this chapter.

8.2 Problem formulation

Consider the proper, single-input single-output, linear and time-invariant, continuous-time system

$$x(t) = \frac{B^*(p)}{A^*(p)} u(t).$$

The system numerator and denominator polynomials are assumed to be coprime with orders m^* and n^* , respectively, i.e.,

$$\begin{aligned} B^*(p) &= b_{m^*}^* p^{m^*} + b_{m^*-1}^* p^{m^*-1} + \cdots + b_1^* p + b_0^*, \\ A^*(p) &= a_{n^*}^* p^{n^*} + a_{n^*-1}^* p^{n^*-1} + \cdots + a_1^* p + 1, \end{aligned} \quad (8.1)$$

where the degrees of the polynomials satisfy $m^* \leq n^*$, and the system parameter vector is described by (2.6). A noisy measurement of the sampled output is retrieved for identification:

$$y(t_k) = x(t_k) + \frac{C^*(q)}{D^*(q)} e(t_k),$$

where $e(t_k)$ describes a zero-mean white noise stochastic process of finite variance, and

$$\begin{aligned} C^*(q) &= 1 + c_1^* q^{-1} + c_2^* q^{-2} + \cdots + c_{m_c^*}^* q^{-m_c^*}, \\ D^*(q) &= 1 + d_1^* q^{-1} + d_2^* q^{-2} + \cdots + d_{n_d^*}^* q^{-n_d^*}, \end{aligned}$$

with $n_d^* \geq m_c^*$. The parameters in the noise model can be described by the vector in (2.12). Throughout this chapter, the transfer function $G^*(p) := B^*(p)/A^*(p)$ is assumed to be *unstable*. More precisely, following the definition of instability in [97, p. 77], we assume at least one zero of the polynomial $A^*(p)$ is in the closed right-half of the complex plane. We also assume that the discrete-time filter $H^*(q) := C^*(q)/D^*(q)$ is stable and minimum-phase. Recall that by the spectral factorization theorem [170], any colored noise with rational, strictly positive and integrable spectrum can be described by a noise white stochastic process filtered by a stable and minimum-phase transfer function.

Our goal is to estimate the transfer function $G^*(p)$ and the noise model $H^*(q)$ using sampled data. We require N samples of input and output data $\{u(t_k), y(t_k)\}_{k=1}^N$, where the sampling instants $\{t_k\}_{k=1}^N$ are assumed to be evenly spaced in time with a sampling period h . Unstable systems must be stabilized prior to identifying them due to safety concerns regarding their operation when performing an identification experiment. Therefore, we assume that the data is obtained with a stabilizing feedback. Bear in mind that this is not a limitation of our approach; it is a standard assumption when identifying unstable systems [52, 57]. However, at this stage, no information about the closed-loop is assumed. We further assume that the intersample behavior of the input is known but can be arbitrary (i.e., it is not necessarily perfectly described via ZOH or FOH devices). This aspect, along with the nature

of the transfer functions, separates this analysis from the standard discrete-time approaches [51], which typically assume that the intersample behavior of the signals has already been addressed prior to the description of the systems.

The model we consider is the following:

$$y(t_k) = \left\{ \frac{B(p)}{A(p)} u(t) \right\}_{t=t_k} + \frac{C(q)}{D(q)} \varepsilon(t_k), \quad (8.2)$$

where the signal $\varepsilon(t_k)$ is denoted as the residual, and the transfer functions $G(p) = B(p)/A(p)$ and $H(q) = C(q)/D(q)$ are described by polynomials that are parameterized¹ by vectors $\boldsymbol{\theta}$ and $\boldsymbol{\eta}$ of the same form as (2.6) and (2.12), respectively. The system model is parameterized by polynomials of degree n and m in the denominator and numerator respectively, and the noise model has m_c zeros and n_d poles. An exact parametrization of the models (i.e., $n = n^*$, $m = m^*$, $m_c = m_c^*$ and $n_d = n_d^*$) is not assumed.

8.3 RIVC and SRIVC for unstable systems

It has been widely reported [52, 55, 138] that unstable systems can be estimated in a prediction error framework only if the predictors are stable. Refined instrumental variable estimators are not an exception to this rule: in fact, there is a close resemblance between the Gauss-Newton iterations of the prediction error method and refined instrumental variable methods as they both require computing the gradient of the predictor and some form of the residual $\varepsilon(t_k)$ at each iteration. Following the derivations in [252] and Lemma 6.2, the refined instrumental variable iterations in (2.35) can be written in an incremental form:

$$\boldsymbol{\theta}_{j+1} = \boldsymbol{\theta}_j + (\hat{\Phi}_j^\top \Phi_j)^{-1} \hat{\Phi}_j^\top \mathbf{e}_j,$$

where Φ_j is a $N \times (n+m+1)$ matrix formed by stacking the filtered regressor vectors $\{\varphi_f^\top(t_k)\}_{k=1}^N$, and $\hat{\Phi}_j$ is formed with the instrument vector by the same manner. The vector \mathbf{e}_j is given by the residual at each time instant using the estimate at the j -th iteration, and it is computed by filtering through an unstable model if the system is unstable. The same problem is encountered with the computation of Φ_j and $\hat{\Phi}_j$. Keep in mind that the refined instrumental variable methods typically project the parameter vector estimate at each iteration to the region of stability (see, e.g., lines 7 to 14 of Algorithm 2.1); this procedure improves the robustness of the methods but it will lead to useless results if the underlying system is stable.

In this section we derive extensions of the RIVC, SRIVC, and LSSVF methods that can identify stable or unstable continuous-time plants. For this matter, we first study the optimal one-step-ahead predictor for a hybrid BJ model structure

¹In some sections of this chapter we will write $A(p, \boldsymbol{\theta})$ instead of $A(p)$, or $C(q, \boldsymbol{\eta})$ instead of $C(q)$, to show explicit dependence of the model polynomials on the unknown parameters $\boldsymbol{\theta}$ or $\boldsymbol{\eta}$, respectively.

with unstable continuous-time systems. We then propose the filtered regressor and output vectors to be used for the extension of the RIVC and SRIVC methods, and present an adequate initialization procedure based on the LSSVF method. Finally, we study the instrument vector constructed such that a PEM estimate is obtained upon convergence.

8.3.1 One-step-ahead predictor and refined least-squares

The RIVC algorithm implicitly uses the predictor

$$\hat{y}(t_k | \boldsymbol{\theta}) = \frac{D(q)}{C(q)} \left\{ \frac{B(p)}{A(p)} u(t) \right\}_{t=t_k} + \left(1 - \frac{D(q)}{C(q)} \right) y(t_k). \quad (8.3)$$

Our first goal is to extend the applicability of this estimator to the unstable system case. This extension must admit ZOH and FOH input signals as well as arbitrary input signals whose intersample behavior is known, such as in the case of continuous-time multisines or band-limited signals (recall Chapter 6). Before we derive the stable predictor, we must define some polynomials of interest. The denominator of the discrete-time equivalent of $1/A(p)$ is denoted as $A_d(q)$, and it is assumed to be of the form $A_d(q) = 1 + a_1^d q^{-1} + \dots + a_n^d q^{-n}$. This polynomial can be written in terms of the coefficients of $A(p)$ via the following relation:

$$1 + a_1^d q^{-1} + \dots + a_n^d q^{-n} = q^{-n} \det (q \mathbf{I}_n - \exp(\mathbf{A}h)), \quad (8.4)$$

where

$$\mathbf{A} = \begin{bmatrix} -\frac{a_{n-1}}{a_n} & \dots & -\frac{a_1}{a_n} & -\frac{1}{a_n} \\ 1 & & 0 & 0 \\ & \ddots & & \vdots \\ 0 & & 1 & 0 \end{bmatrix}. \quad (8.5)$$

We also consider the decomposition $A_d(q) = A_{d,s}(q)A_{d,a}(q)$, where $A_{d,s}(q)$ denotes the stable part of $A_d(q)$ and $A_{d,a}(q)$ its antistable part. In particular, we write

$$\begin{aligned} A_{d,s}(q) &= 1 + a_1^{d,s} q^{-1} + \dots + a_{n_s}^{d,s} q^{-n_s}, \\ A_{d,a}(q) &= 1 + a_1^{d,a} q^{-1} + \dots + a_{n_a}^{d,a} q^{-n_a}, \end{aligned}$$

where $n_s + n_a = n$ and $a_{n_a}^{d,a}$ is assumed to be different from zero without loss of generality. Note that with the notation

$$\tilde{a}_k^{d,s} := \begin{cases} 1 & \text{if } k = 0, \\ a_k^{d,s} & \text{if } 1 \leq k \leq n_s, \\ 0 & \text{otherwise;} \end{cases} \quad \tilde{a}_k^{d,a} := \begin{cases} 1 & \text{if } k = 0, \\ a_k^{d,a} & \text{if } 1 \leq k \leq n_a, \\ 0 & \text{otherwise,} \end{cases}$$

the following relation holds:

$$a_k^d = \sum_{j=0}^n \tilde{a}_j^{d,s} \tilde{a}_{k-j}^{d,a}. \quad (8.6)$$

Furthermore, we denote by $\bar{A}_{d,a}(q)$ the monic polynomial whose zeros are those of $A_{d,a}(q)$ reflected into the unit disc, i.e.,

$$\bar{A}_{d,a}(q) = 1 + \frac{a_{n_a-1}^{d,a}}{a_{n_a}^{d,a}} q^{-1} + \cdots + \frac{a_1^{d,a}}{a_{n_a}^{d,a}} q^{-n_a+1} + \frac{1}{a_{n_a}^{d,a}} q^{-n_a}.$$

Our approach relies on studying the following model structure:

$$y(t_k) = \left\{ \frac{B(p)}{A(p)} u(t) \right\}_{t=t_k} + \frac{\tilde{C}(q)}{\tilde{D}(q)} \varepsilon(t_k), \quad (8.7)$$

where

$$\tilde{C}(q) = C(q) \bar{A}_{d,a}(q), \text{ and } \tilde{D}(q) = D(q) A_{d,a}(q). \quad (8.8)$$

A model structure with a tailor-made noise model as in (8.7) was introduced for discrete-time system identification in [52], and it has been used for identifying unstable systems with the WNSF method [56], as well as for ARX modeling [57]. To the best of our knowledge, this approach has not been applied in continuous-time system identification until now.

With the previous definitions in mind, we are ready to derive a stable predictor and its properties for this case. In the sequel, we will explicitly write the dependence of the polynomials and vectors on the parameter vectors $\boldsymbol{\theta}$ and $\boldsymbol{\eta}$.

Proposition 8.1. Consider the model structure (8.7), where $\varepsilon(t_k)$ is assumed to be white noise. The optimal one-step-ahead predictor is given by

$$\hat{y}(t_k | \boldsymbol{\theta}, \boldsymbol{\eta}) = \frac{\tilde{D}(q, \boldsymbol{\eta}, \boldsymbol{\theta})}{\tilde{C}(q, \boldsymbol{\eta}, \boldsymbol{\theta})} \left\{ \frac{B(p, \boldsymbol{\theta})}{A(p, \boldsymbol{\theta})} u(t) \right\}_{t=t_k} + \left(1 - \frac{\tilde{D}(q, \boldsymbol{\eta}, \boldsymbol{\theta})}{\tilde{C}(q, \boldsymbol{\eta}, \boldsymbol{\theta})} \right) y(t_k). \quad (8.9)$$

Proof. See Appendix 8.A.1. □

The difference between Proposition 8.1 and the standard result for discrete-time models (see, e.g., Chapter 3 of [138]) is that the plant is a continuous-time transfer function. Thus, the cascading of discrete-time and continuous-time transfer functions must be done with caution: the model output must be treated as a discrete-time signal instead of a continuous-time one. Nevertheless, the proof follows the same lines as its discrete-time version.

Similar to the discrete-time case, minimizing the 2-norm of the prediction error of the model structure in (8.7) can be shown to be asymptotically equivalent to minimizing the 2-norm of the prediction error for the model structure (8.2).

Proposition 8.2. When applying the prediction error method to model structures (8.2) and (8.7), the resulting estimates of the two model structures will asymptotically (as $N \rightarrow \infty$) be the same.

Proof. The proof follows the same lines as the proof of Proposition 1 of [52]. \square

Proposition 8.1 provides a path to writing a modification of the GEE that has stable filters, which is essential for computing the RIVC estimate. We write the residual $\varepsilon(t_k)$ as

$$\varepsilon(t_k) = y(t_k) - \hat{y}(t_k | \boldsymbol{\theta}, \boldsymbol{\eta}) \quad (8.10)$$

$$= \frac{\tilde{D}(q, \boldsymbol{\eta}, \boldsymbol{\theta})}{\tilde{C}(q, \boldsymbol{\eta}, \boldsymbol{\theta})} A(p, \boldsymbol{\theta}) \frac{1}{A(p, \boldsymbol{\theta})} y(t_k) - \frac{\tilde{D}(q, \boldsymbol{\eta}, \boldsymbol{\theta})}{\tilde{C}(q, \boldsymbol{\eta}, \boldsymbol{\theta})} \left\{ \frac{B(p, \boldsymbol{\theta})}{A(p, \boldsymbol{\theta})} u(t) \right\}_{t=t_k}. \quad (8.11)$$

Thus, a pseudo-linear regression equation can now be written as

$$y_{f,uns}(t_k) = \boldsymbol{\varphi}_{f,uns}^\top(t_k) \boldsymbol{\theta} + \varepsilon(t_k), \quad (8.12)$$

where

$$y_{f,uns}(t_k) = \frac{\tilde{D}(q, \boldsymbol{\eta}, \boldsymbol{\theta})}{\tilde{C}(q, \boldsymbol{\eta}, \boldsymbol{\theta})} \frac{1}{A(p, \boldsymbol{\theta})} y(t_k), \quad (8.13)$$

$$\boldsymbol{\varphi}_{f,uns}^\top(t_k) = \frac{\tilde{D}(q, \boldsymbol{\eta}, \boldsymbol{\theta})}{\tilde{C}(q, \boldsymbol{\eta}, \boldsymbol{\theta})} \left[\frac{-p}{A(p, \boldsymbol{\theta})} y(t_k), \dots, \frac{-p^n}{A(p, \boldsymbol{\theta})} y(t_k), \left\{ \frac{1}{A(p, \boldsymbol{\theta})} u(t) \right\}_{t=t_k}, \dots, \left\{ \frac{p^m}{A(p, \boldsymbol{\theta})} u(t) \right\}_{t=t_k} \right]. \quad (8.14)$$

Note that the goal of this description is to render filtered signals that are bounded, despite possible instabilities due to filtering by $1/A(p, \boldsymbol{\theta})$. Before deriving the proposed extension for the RIVC estimator (which requires deriving a tailor-made filtered instrument vector), a refined least-squares method will be introduced. Assuming knowledge of an initial model estimate, the refined least-squares method we study consists of four steps:

1. Based on the j -th iteration of the system model (i.e., $\boldsymbol{\theta}_j$), fit an ARMA model $v(t_k) = [C(q, \boldsymbol{\eta}_{j+1})/D(q, \boldsymbol{\eta}_{j+1})]e(t_k)$ to the estimated colored noise sequence

$$\hat{v}(t_k) = \frac{A_{d,a}(q, \boldsymbol{\theta}_j)}{\bar{A}_{d,a}(q, \boldsymbol{\theta}_j)} \left(y(t_k) - \left\{ \frac{B(p, \boldsymbol{\theta}_j)}{A(p, \boldsymbol{\theta}_j)} u(t) \right\}_{t=t_k} \right). \quad (8.15)$$

2. Form the filtered output $y_{f,uns}(t_k)$ and regressor $\boldsymbol{\varphi}_{f,uns}(t_k)$ from (8.13) and (8.14), where $\boldsymbol{\eta}$ and $\boldsymbol{\theta}$ are set to $\boldsymbol{\eta}_{j+1}$ and $\boldsymbol{\theta}_j$, respectively.
3. Compute the estimate

$$\boldsymbol{\theta}_{j+1} = \left[\sum_{k=1}^N \boldsymbol{\varphi}_{f,uns}(t_k) \boldsymbol{\varphi}_{f,uns}^\top(t_k) \right]^{-1} \left[\sum_{k=1}^N \boldsymbol{\varphi}_{f,uns}(t_k) y_{f,uns}(t_k) \right]. \quad (8.16)$$

4. Go back to Step 1 and repeat until a maximum number of iterations is reached or until the relative error between the previous and current estimates is smaller than a preset tolerance factor ϵ .

Several implementation issues regarding the filtering process are discussed next.

- a) The discrete-time filtering in (8.15) is needed for stability since the model estimate will usually be unstable. The unstable poles of the model estimate are canceled with the zeros of $A_{d,a}(q, \theta_j)$, rendering a signal $\hat{v}(t_k)$ that is quasi-stationary if the input and noise are quasi-stationary [138]. Note that the pole-canceling filter only affects the gain of the spectrum of $\hat{v}(t_k)$ since it is designed to be all-pass. Moreover, a byproduct of Proposition 8.2 is that other pole-canceling filters can be proposed with the same prediction error being minimized asymptotically, as long as the discrete-time filters are all-pass.
- b) Poles of model estimates that are on the imaginary axis are not canceled with our approach. This is not an important setback in a practical setting if the true system does not contain integrators or undamped oscillatory modes, since cases where a pole of the model is found exactly on the imaginary axis occur with probability zero. In Section 8.5 we study two different approaches for the identification of systems with integral effect that are complementary to the methods presented in this section.
- c) The input signal filtering can be done entirely in discrete-time for ZOH or FOH inputs. In practice, this is done by transforming the continuous-time transfer function into its ZOH or FOH equivalent, multiplying the transfer functions in the q -domain, and later computing the filtered signal via discrete-time filtering. This approach is needed for the cancellation of unstable poles present in $1/A(p, \theta_j)$ with the zeros in $A_{d,a}(q, \theta_j)$. Note that for such types of inputs, discrete-time filtering (i.e., with ZOH or FOH equivalent transfer functions) is commonly used for computing the regressor and filtered output [61].
- d) For continuous-time multisine input signals, it is sufficient to compute the continuous-time filtering in (8.14) assuming steady state such as in (6.10), and later performing the discrete-time filtering by the inverted noise model assuming a ZOH or FOH intersample behavior for the signals $\{[p^l/A(p, \theta_j)]u(t)\}_{t=t_k}$, $l = 0, 1, \dots, m$. Note that such inputs are rarely encountered if there is noise present in the loop, which is the case for most practical applications.
- e) The unstable poles of the bank of continuous-time filters are not canceled by $A_{d,a}(q, \theta_j)$ when the input is an arbitrary continuous-time signal that is not a multisine, ZOH or FOH. To solve this issue, we propose modifying the oversampling technique suggested in Chapters 6 and 7 for this scenario. The previously proposed method for solving the instability problem using the noise model in (8.8) does not work for this case, since the poles of the equivalent

fast-sampled delta description do not cancel with the zeros of $A_{d,a}(q, \boldsymbol{\theta}_j)$. The solution to this issue is described in detail in Appendix 8.B.

Remark 8.1. It can be shown that the estimator presented in (8.16) is asymptotically biased, and does not minimize the sum of squares of the prediction errors. However, it is useful for understanding the refinement nature of the instrumental variable methods that we introduce in the next subsection, and its use can be justified if there are computational constraints that limit the use of the other methods we study in this chapter. Furthermore, one iteration of the estimator (8.16) with $j = 1$, $C_2(q) = D_2(q) \equiv 1$ and a fixed $1/A_1(p)$ polynomial provides a refinement of the LSSVF method that can admit an unstable state variable filter $1/A_1(p)$. We propose this method, which we call LSSVF-uns, as an alternative for initializing refined instrumental variable methods for unstable systems.

In the sequel, we eliminate the asymptotic bias of the refined LS method by constructing the instrument vector that delivers a PEM estimate for this context, which leads to novel refined instrumental variable methods that admit unstable models.

8.3.2 Instrument vector

We are interested in estimators that are computed iteratively using instrumental variables:

$$\boldsymbol{\theta}_{j+1} = \left[\sum_{k=1}^N \hat{\boldsymbol{\varphi}}_{f,uns}(t_k) \boldsymbol{\varphi}_{f,uns}^\top(t_k) \right]^{-1} \left[\sum_{k=1}^N \hat{\boldsymbol{\varphi}}_{f,uns}(t_k) y_{f,uns}(t_k) \right]. \quad (8.17)$$

Here the filtered instrument vector is $\hat{\boldsymbol{\varphi}}_{f,uns}(t_k)$, and uses filters that depend on the estimate of the previous iteration $\boldsymbol{\theta}_j$ and the current noise model estimate $\boldsymbol{\eta}_{j+1}$.

Lemma 8.1. Consider the iterative method in (8.17) but where the noise model is fixed, and assume that at the converging point (i.e., $\lim_{j \rightarrow \infty} \boldsymbol{\theta}_j$) the matrix

$$\frac{1}{N} \sum_{k=1}^N \hat{\boldsymbol{\varphi}}_{f,uns}(t_k) \boldsymbol{\varphi}_{f,uns}^\top(t_k)$$

is non-singular, where the instrument vector $\hat{\boldsymbol{\varphi}}_{f,uns}(t_k)$ at each iteration is given by

$$\hat{\boldsymbol{\varphi}}_{f,uns}(t_k) = \frac{\partial}{\partial \boldsymbol{\theta}} \hat{y}(t_k | \boldsymbol{\theta}, \boldsymbol{\eta}) \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_j}, \quad (8.18)$$

with $\hat{y}(t_k | \boldsymbol{\theta}, \boldsymbol{\eta})$ as in (8.9). If the algorithm in (8.17) converges as $j \rightarrow \infty$ (with N fixed and finite), then the converging point minimizes the sum of squares of the residuals.

Proof. See Appendix 8.A.2. □

Remark 8.2. If the model estimates at each iteration are stable, then Lemma 8.1 applied to the predictor (8.3) leads to the RIVC estimator with instrument vector given by (2.38), or in the arbitrary input case,

$$\hat{\varphi}_f(t_k) = \frac{D_{j+1}(q)}{C_{j+1}(q)} \left[\frac{-pB_j(p)}{A_j^2(p)} u(t), \dots, \frac{-p^n B_j(p)}{A_j^2(p)} u(t), \frac{1}{A_j(p)} u(t), \dots, \frac{p^m}{A_j(p)} u(t) \right]_{t=t_k}^\top.$$

This supports the findings of [252], in which it is suggested that the RIVC estimator provides a maximum likelihood estimate for hybrid BJ models with filtered Gaussian noise.

We introduce the estimator computed by (8.17), where $\hat{\varphi}_{f,uns}(t_k)$ is computed from (8.18), as the RIVC estimator for stable and unstable continuous-time systems. This estimator is called RIVC-uns from now on. Following this notation, the SRIVC-uns estimator is computed when the noise model is fixed to $C(q) = D(q) \equiv 1$. An implementation diagram for the RIVC-uns estimator can be found in Figure 8.1. Note that the gradient step can also be performed with the residual $\varepsilon(t_k)$ instead of the one-step predictor $\hat{y}(t_k|\boldsymbol{\theta}, \boldsymbol{\eta})$, due to the relation between these signals in (8.10).

The computation of (8.18) is not direct due to the mixture of continuous and discrete-time transfer functions. For simplicity, we will only analyze the case when the input has a ZOH or FOH intersample behavior. For the numerator components present in $\boldsymbol{\theta}$, the derivative can be directly evaluated as

$$\frac{\partial}{\partial b_l} \hat{y}(t_k|\boldsymbol{\theta}, \boldsymbol{\eta}) = \frac{\tilde{D}(q, \boldsymbol{\eta}, \boldsymbol{\theta})}{\tilde{C}(q, \boldsymbol{\eta}, \boldsymbol{\theta})} \frac{p^l}{A(p, \boldsymbol{\theta})} u(t_k), \quad l = 0, 1, \dots, m.$$

In other words, the last $m + 1$ components of $\hat{\varphi}_f(t_k, \boldsymbol{\theta})$ coincide with those of the filtered regressor $\varphi_f(t_k, \boldsymbol{\theta})$. The first n components of $\hat{\varphi}_f(t_k, \boldsymbol{\theta})$, associated with the denominator polynomial of the model, are more difficult to compute. They are given by

$$\frac{\partial}{\partial a_l} \hat{y}(t_k|\boldsymbol{\theta}, \boldsymbol{\eta}) = \frac{D(q, \boldsymbol{\eta})}{C(q, \boldsymbol{\eta})} \frac{\partial}{\partial a_l} \left\{ \frac{A_{d,a}(q, \boldsymbol{\theta})}{A_{d,a}(q, \boldsymbol{\theta})} \left(\left\{ \frac{B(p, \boldsymbol{\theta})}{A(p, \boldsymbol{\theta})} u(t) \right\}_{t=t_k} - y(t_k) \right) \right\}, \quad l = 1, 2, \dots, n. \quad (8.19)$$

Computing the derivative via numerical approximation is sufficient for most practical applications. If a gradient is needed in explicit form, the required computations are provided in Appendix 8.C.

8.4 The CLSRIVC-uns method

In this section we extend the ideas of Section 8.3 to study indirect closed-loop continuous-time system identification when the open-loop plant is unstable. We will not estimate the noise model in this case, although our approach can also be extended to handle that scenario by modifying the CLRIVC method in [77]. Let us consider

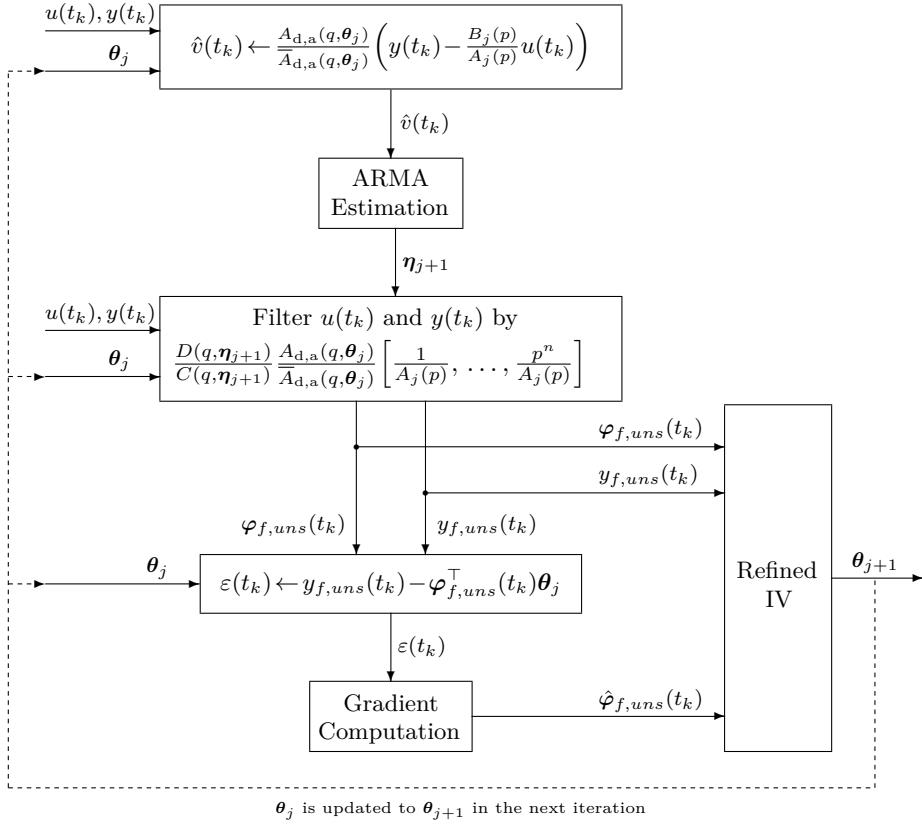


Figure 8.1: Implementation of the RIVC-uns algorithm for when the input has a ZOH or FOH behavior. Note that in this case, the filtering process is done in discrete-time by discretizing the bank of continuous-time filters $p^i/A_j(p)$ and cascading them with the all-pass filter and inverse noise model. The algorithm is initialized with $\theta_j = \theta_1$ (i.e., $j = 1$). Note that the SRIVC-uns algorithm follows from omitting the ARMA estimation step and using $D(q) = C(q) \equiv 1$.

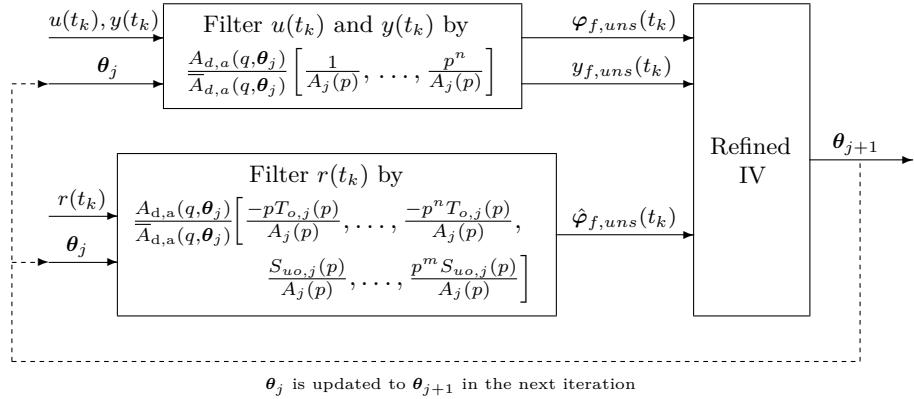


Figure 8.2: Implementation of the CLSRIVC-uns algorithm for when the input has a ZOH or FOH intersample behavior and the control is in continuous-time. The method is initialized with $\theta_j = \theta_1$ (i.e., $j = 1$).

again the two closed-loop settings in Figure 7.1, which have been carefully studied in Chapter 7. An exact implementation of the CLSRIVC is considered, that is, the model estimates of each iteration of the CLSRIVC estimator are given by (2.35), the filtered regressor vector is computed as (7.17), and the filtered instrument vector is given by

$$\hat{\varphi}_f(t_k) = \left[\frac{-pT_{o,j}(p)}{A_j(p)} r(t_k), \dots, \frac{-p^nT_{o,j}(p)}{A_j(p)} r(t_k), \frac{S_{uo,j}(p)}{A_j(p)} r(t_k), \dots, \frac{p^mS_{uo,j}(p)}{A_j(p)} r(t_k) \right]^\top \quad (8.20)$$

for Setting 1 in Figure 7.1 (i.e., a continuous-time control law), and

$$\hat{\varphi}_f(t_k, \theta_j) = \left[\frac{-pB_j(p)}{A_j^2(p)}, \dots, \frac{-p^nB_j(p)}{A_j^2(p)}, \frac{1}{A_j(p)}, \dots, \frac{p^m}{A_j(p)} \right]^\top S_{uo,j}(q) r(t_k) \quad (8.21)$$

for Setting 2 in Figure 7.1 (i.e., a discrete-time control law).

The proposed solution to the instability problem in the CLSRIVC method is illustrated in Figure 8.2. The idea is to use $y_{f,uns}(t_k, \theta_j)$ and $\varphi_{f,uns}(t_k, \theta_j)$ as in (8.13) and (8.14), respectively (but with $C_j(q) = D_j(q) \equiv 1$), and to construct the modified instrument vector

$$\hat{\varphi}_{f,uns}(t_k) = \frac{A_{d,a}(q, \theta_j)}{A_{d,a}(q, \theta_j)} \hat{\varphi}_f(t_k), \quad (8.22)$$

where $\hat{\varphi}_f(t_k)$ is defined in (8.20) or (8.21), depending on the feedback setup. The iterations are given by (8.17) for both cases. If a continuous-time controller is used and the intersample behavior of the input is known, the filtering procedure follows the same ideas discussed in Appendix 8.B. If a discrete-time control law is known in advance, Lemma 8.2 characterizes the asymptotic behavior of the estimator.

Lemma 8.2. Consider the CLSRIVC-uns method with filtered output (8.13), filtered regressor vector (8.14) and filtered instrument vector (8.22) with $\hat{\varphi}_f(t_k)$ computed as (8.21), and assume that the reference, input, and noise are ZOH signals and are stationary. Furthermore, assume that the CLSRIVC-uns method converges in iterations for all N sufficiently large, and that the matrix

$$\mathbb{E} \left\{ \hat{\varphi}_{f,uns}(t_k) \varphi_{f,uns}^\top(t_k) \right\} \quad (8.23)$$

is non-singular. Then, at the converging point and as $N \rightarrow \infty$, the CLSRIVC-uns method and its standard version (i.e., (2.35) with (2.40) as filtered instrument vector) solve the same pseudo-linear regression equation, namely,

$$\mathbb{E} \left\{ \hat{\varphi}_f(t_k, \bar{\theta}) \left(y(t_k) - \frac{\bar{B}(p)}{\bar{A}(p)} u(t_k) \right) \right\} = \mathbf{0}.$$

Proof. See Appendix 8.A.3. \square

Remark 8.3. From the consistency analysis in Chapter 7, we know that the CLSRIVC method proposed in Section 7.4.2 is generically consistent (with respect to the model and system parameters) even when the noise is not white, as long as it is uncorrelated with the reference signal. Thus, under that condition and the assumptions in Lemma 8.2, we find that the CLSRIVC-uns method is generically consistent for both stable and unstable systems. It should be noted that, provided some identifiability conditions are met, the matrix in (8.23) can be shown to be generically non-singular following the same approach as in Chapter 7.

Remark 8.4. The three-stage method for closed-loop continuous-time system identification presented in [258] is claimed to induce statistical efficiency for stable open-loop systems. However, under its current implementation it cannot be used when the open-loop system is unstable. The all-pass filtering technique exploited in the present thesis can also be applied to this method, thus extending it to open-loop unstable systems as well. After the generation of an estimate of the noise-free input $\hat{u}(t_k)$, the extended RIVC or SRIVC methods presented in Section 8.3 permit the identification of an unstable estimate $\hat{G}(p) = \hat{B}(p)/\hat{A}(p)$ using $\hat{u}(t_k)$ and $y(t_k)$. Finally, a refinement of this model estimate can be obtained based on the filtered estimates of the noise-free input and output:

$$\tilde{u}(t_k) = \frac{A_{d,a}(q, \hat{\theta})}{\bar{A}_{d,a}(q, \hat{\theta})} \hat{u}(t_k), \quad \text{and} \quad \tilde{y}(t_k) = \frac{A_{d,a}(q, \hat{\theta})}{\bar{A}_{d,a}(q, \hat{\theta})} \left(y(t_k) - \hat{G}(p)[u(t_k) - \hat{u}(t_k)] \right).$$

8.5 Estimating continuous-time systems with integral action

In many applications such as platoon vehicle modeling [182], modeling of DC motors [146], or position tracking [97], prior knowledge of the system's dynamics

suggests that there is an integrator in the plant. Unfortunately, the all-pass filtering procedure introduced in Section 8.3 only stabilizes strictly unstable poles of the model. Here we propose two solutions for identifying systems with integration.

- *Solution 1* (Use the extended versions of LSSVF, SRIVC, RIVC, and CLSRIVC for a monic denominator polynomial of the model): Instead of deriving the estimators using an anti-monic denominator polynomial for the system (as in (8.1)), we can consider a monic polynomial for $A^*(p)$ of the form

$$A^*(p) = p^{n^*} + a_{n^*-1}^* p^{n^*-1} + \cdots + a_1^* p + a_0^*.$$

With this description the regression equation is still of the form (8.12), but with

$$\begin{aligned} y_{f,uns}(t_k) &= \frac{\tilde{D}(q, \boldsymbol{\eta}, \boldsymbol{\theta})}{\tilde{C}(q, \boldsymbol{\eta}, \boldsymbol{\theta})} \frac{p^n}{A(p, \boldsymbol{\theta})} y(t_k), \\ \boldsymbol{\varphi}_{f,uns}^\top(t_k) &= \frac{\tilde{D}(q, \boldsymbol{\eta}, \boldsymbol{\theta})}{\tilde{C}(q, \boldsymbol{\eta}, \boldsymbol{\theta})} \left[\frac{-1}{A(p, \boldsymbol{\theta})} y(t_k), \dots, \frac{-p^{n-1}}{A(p, \boldsymbol{\theta})} y(t_k), \right. \\ &\quad \left. \left\{ \frac{1}{A(p, \boldsymbol{\theta})} u(t) \right\}_{t=t_k}, \dots, \left\{ \frac{p^m}{A(p, \boldsymbol{\theta})} u(t) \right\}_{t=t_k} \right]. \end{aligned}$$

The SRIVC-uns and RIVC-uns estimators with monic denominator polynomial for the model are obtained by computing (8.17) iteratively, and the CLSRIVC-uns estimator for a continuous-time control law uses the filtered instrument vector

$$\boldsymbol{\varphi}_f(t_k) = \left[\frac{-T_{o,j}(p)}{A_j(p)} r(t), \dots, \frac{-p^{n-1} T_{o,j}(p)}{A_j(p)} r(t), \frac{S_{uo,j}(p)}{A_j(p)} r(t), \dots, \frac{p^m S_{uo,j}(p)}{A_j(p)} r(t) \right]_{t=t_k}^\top.$$

The advantage of this formulation is that the estimate for a_0^* can go to zero, contrary to the anti-monic formulation which fixes the constant term of the denominator to one. The drawback of this approach is that integration is not forced explicitly, which can lead to an increase in covariance of the parameter estimates compared to estimators that can fix an integrator in the model, as the ones proposed in Solution 2. Bear in mind that the anti-monic formulation considered in Sections 8.3 and 8.4 is favored when the model structure is not known, since it can lead to consistent estimators despite possible overparametrization in the denominator as studied in Chapters 4 and 5.

- *Solution 2* (Force a pole at zero in the SRIVC-uns, RIVC-uns and CLSRIVC-uns estimators): The key idea behind this approach is that the mean value of the system input is zero if the closed-loop system is stable and the system has an integrator,. In other words, filtering the input $u(t_k)$ through a filter with

integration yields a bounded signal, which can be used in the construction of the filtered regressor and instrument vectors. The procedure proposed here resembles the one in [165], where a known stable pole was forced into the SRIVC algorithm to identify a biological system.

For simplicity, assume that the system has one integrator. If the true system is of the form $B^*(p)/[p\tilde{A}^*(p)]$, where $\tilde{A}^*(p)$ is a monic polynomial of order n , then the RIVC-uns algorithm with a fixed pole at $p = 0$ can be obtained from the regression equation (8.12), where

$$y_{f,uns}(t_k) = \frac{\tilde{D}(q, \boldsymbol{\eta}, \boldsymbol{\theta})}{\tilde{C}(q, \boldsymbol{\eta}, \boldsymbol{\theta})} \frac{p^{n-1}}{\tilde{A}(p, \boldsymbol{\theta})} y(t_k),$$

$$\boldsymbol{\varphi}_{f,uns}^\top(t_k) = \frac{\tilde{D}(q, \boldsymbol{\eta}, \boldsymbol{\theta})}{\tilde{C}(q, \boldsymbol{\eta}, \boldsymbol{\theta})} \left[\frac{-1}{\tilde{A}(p, \boldsymbol{\theta})} y(t_k), \dots, \frac{-p^{n-2}}{\tilde{A}(p, \boldsymbol{\theta})} y(t_k), \right.$$

$$\left. \left\{ \frac{1}{p\tilde{A}(p, \boldsymbol{\theta})} u(t) \right\}_{t=t_k}, \dots, \left\{ \frac{p^{m-1}}{\tilde{A}(p, \boldsymbol{\theta})} u(t) \right\}_{t=t_k} \right],$$

and the CLSRIVC-uns estimator for a continuous-time control setup has an instrument vector given by

$$\boldsymbol{\varphi}_f(t_k) = \left[\frac{-T_{o,j}(p)}{\tilde{A}_j(p)} r(t), \dots, \frac{-p^{n-2}T_{o,j}(p)}{\tilde{A}_j(p)} r(t), \frac{S_{uo,j}(p)}{p\tilde{A}_j(p)} r(t), \dots, \frac{p^{m-1}S_{uo,j}(p)}{\tilde{A}_j(p)} r(t) \right]_{t=t_k}^\top.$$

These formulations lead to model estimates with a fixed pole at $p = 0$, and also admit other unstable poles in the estimation process, if present.

8.6 Monte Carlo simulation studies

In this section we show the applicability of the proposed methods to identify unstable continuous-time systems. For the first three tests we consider the following unstable plant:

$$G^*(p) = \frac{-0.5p + 1}{p^2 + 2\xi\omega_n p + \omega_n^2}, \quad (8.24)$$

where $\omega_n^2 = 2$ and $\xi = -0.05$. The output of this system is sampled at $h = 0.1$ [s], unless indicated otherwise.

8.6.1 Continuous-time control, output error model structure

We test the performance of the proposed SRIVC-uns estimator for arbitrary input signal excitation, as in Setting 1 of Chapter 7. This time, a continuous-time PID controller is used to control the system in (8.24) of the form

$$C(p) = 0.0249 + \frac{1.42 \cdot 10^{-4}}{p} + \frac{0.974p}{0.00897p + 1}.$$

Table 8.1: Mean and standard deviation of the model parameters obtained with 300 Monte Carlo runs with a continuous-time controller and an output error model structure.

h	Method	Stats.	$a_1(-0.0707)$	$a_2(0.5)$	$b_0(0.5)$	$b_1(-0.25)$
0.1[s]	LSSVF-x-uns	Mean	-0.0642	0.4944	0.4912	-0.2512
		S. Dev.	$4.01 \cdot 10^{-3}$	$4.62 \cdot 10^{-3}$	$2.06 \cdot 10^{-2}$	$2.72 \cdot 10^{-3}$
	SRIVC-x-uns	Mean	-0.0661	0.5018	0.5018	-0.2553
		S. Dev.	$1.20 \cdot 10^{-3}$	$7.23 \cdot 10^{-4}$	$2.45 \cdot 10^{-3}$	$5.58 \cdot 10^{-4}$
0.02[s]	LSSVF-x-uns	Mean	-0.0676	0.4929	0.4894	-0.2485
		S. Dev.	$3.75 \cdot 10^{-3}$	$3.88 \cdot 10^{-3}$	$1.28 \cdot 10^{-2}$	$2.10 \cdot 10^{-3}$
	SRIVC-x-uns	Mean	-0.0701	0.5003	0.4991	-0.2527
		S. Dev.	$1.24 \cdot 10^{-3}$	$8.40 \cdot 10^{-4}$	$2.55 \cdot 10^{-3}$	$5.56 \cdot 10^{-4}$

Two sampling periods are tested, $h = 0.1[\text{s}]$ and $h = 0.02[\text{s}]$. The total time of both experiments is the same ($500[\text{s}]$), and the statistics of the reference signal and disturbance noise are the same for both experiments. The goal is to test the effect of the sampling period on the biasedness of the LSSVF-uns and SRIVC-uns estimators. To this end, a test of 300 Monte Carlo runs are performed for each scenario.

Table 8.1 presents the mean and standard deviation of the parameters for the LSSVF-uns and SRIVC-uns estimators adapted for arbitrary input signals (labeled LSSVF-x-uns and SRIVC-x-uns respectively, following the notation of Chapter 6). This adaptation is done following the δ -domain mechanism proposed in Appendix 8.B. The results show that the SRIVC-x-uns estimator is preferred over its initialization method LSSVF-x-uns, and that the bias of SRIVC-x-uns is reduced when the sampling period decreases. This is in line with the analysis in Appendix 8.B, which indicates that the misspecification of the model structure (i.e., (8.28) instead of (8.7)) is less severe with small sampling periods. Note that fast-rate sampling for accurate estimates is only required for the continuous-time control case, as the model structure we use in this scenario cannot enforce stable predictors and also be equivalent to (8.2) at the same time.

8.6.2 Discrete-time control, output error model structure

The system in (8.24) is now controlled by the discrete-time PID feedback controller

$$C_d(q) = 0.0272 + \frac{2.54 \cdot 10^{-5}}{1 - q^{-1}} + 7.29(1 - q^{-1}),$$

which includes a digital-to-analog device at its output, rendering a continuous-time input $u(t)$ that is constant between samples. The noise is also assumed to be constant between samples, and thus the input and output samples can be generated entirely

Table 8.2: Mean and standard deviation of the model parameters obtained with 300 Monte Carlo runs with a discrete-time controller and an output error model structure.

Method	Stats.	a_1 (-0.0707)	a_2 (0.5)	b_0 (0.5)	b_1 (-0.25)
SRIVC	Mean	0.1132	0.5071	0.5059	-0.2555
	S. Dev.	$2.36 \cdot 10^{-1}$	$4.50 \cdot 10^{-2}$	$3.46 \cdot 10^{-1}$	$4.45 \cdot 10^{-2}$
LSSVF-uns	Mean	-0.0614	0.4672	0.4523	-0.2252
	S. Dev.	$8.95 \cdot 10^{-3}$	$1.24 \cdot 10^{-2}$	$5.27 \cdot 10^{-2}$	$9.65 \cdot 10^{-3}$
Refined LS-uns	Mean	-0.0669	0.4726	0.4569	-0.2280
	S. Dev.	$1.74 \cdot 10^{-3}$	$2.70 \cdot 10^{-3}$	$6.29 \cdot 10^{-3}$	$2.03 \cdot 10^{-3}$
SRIVC-uns	Mean	-0.0708	0.5001	0.5002	-0.2501
	S. Dev.	$1.49 \cdot 10^{-3}$	$1.00 \cdot 10^{-3}$	$3.82 \cdot 10^{-3}$	$1.17 \cdot 10^{-3}$
CLSRIVC-uns	Mean	-0.0708	0.5001	0.5002	-0.2501
	S. Dev.	$1.53 \cdot 10^{-3}$	$1.12 \cdot 10^{-3}$	$3.96 \cdot 10^{-3}$	$2.10 \cdot 10^{-3}$

by the formulas in (7.2). The reference and measurement noise signals are white noise of variance 1 and 0.05, respectively.

Data from 300 Monte Carlo runs, each with $N = 15000$ input and output samples, are used for identification. Four algorithms that have been introduced in this work are evaluated: LSSVF for unstable systems (labeled LSSVF-uns), a refined least-squares method (labeled Refined LS-uns), SRIVC and closed-loop SRIVC for unstable systems (SRIVC-uns and CLSRIVC-uns, respectively). These estimators are compared against the standard SRIVC estimator as implemented in Algorithm 2.1 of Chapter 2. The LSSVF-uns estimate is given by one iteration of (8.16), and is initialized with a filter of the form $1/(\tilde{a}_2 p^2 + \tilde{a}_1 p + 1)$, where \tilde{a}_1 and \tilde{a}_2 are the true parameters perturbed by noise of standard deviation 0.02 and 0.1, respectively. The refined LS-uns, SRIVC-uns and CLSRIVC-uns estimators are computed based on the initialization given by LSSVF-uns, and the instrument vector of the SRIVC-uns algorithm is computed by a first-order numerical differentiation with step size 10^{-8} [216]. The standard SRIVC estimator is initialized with the LSSVF-uns estimate but with its unstable poles reflected into the stable region of the complex plane. The maximum number of iterations of all algorithms is set to 50, with a tolerance factor ϵ of 10^{-8} .

The mean and standard deviation of the parameters are given in Table 8.2, and 10 realizations of both the SRIVC and SRIVC-uns estimators are provided in Figure 8.3. Out of the 300 runs, the SRIVC-uns failed to produce an estimate in one realization (which was not considered in the statistics), while the other methods did not have errors. The LSSVF-uns method provides a reasonable estimate for the initialization of the SRIVC-uns and CLSRIVC-uns methods, which achieve great accuracy in the estimated parameters. These results mimic the behavior of their stable counterparts for stable system identification, as it has been proven in

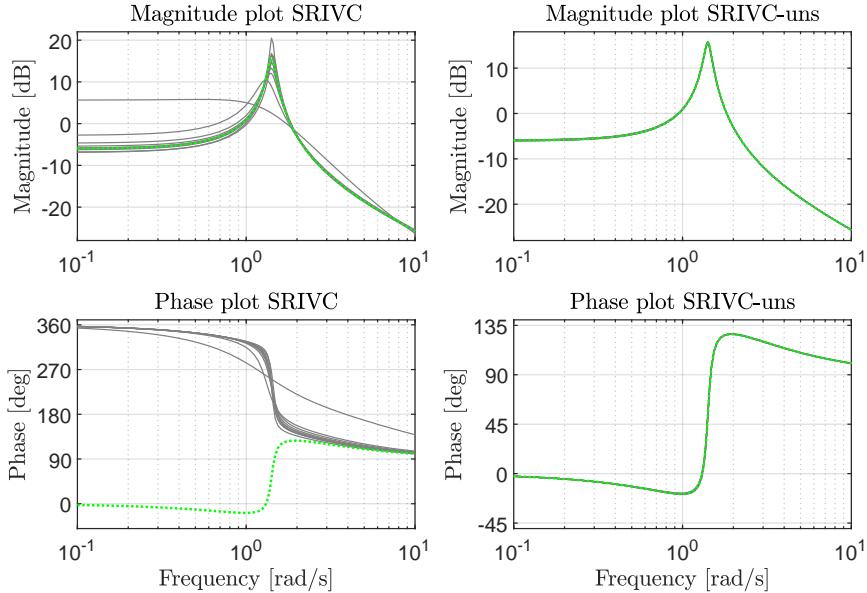


Figure 8.3: Bode plots for 10 realizations of the SRIVC estimator (grey, left plots), 10 realizations of the SRIVC-uns estimator (grey, right plots), and the true system (dashed, green).

Chapter 7 that CLSRIVC is generically consistent for white or colored noise while SRIVC is generically consistent in closed-loop if the disturbance is white and there is no direct feedthrough in the loop. On the other hand, the refined least-squares estimate is more precise than its one-iteration counterpart but does not minimize the 2-norm of the prediction error, which leads to biasedness in this example. As seen in Figure 8.3, the standard SRIVC shows poor performance in the phase plots compared to the SRIVC-uns estimates since it only computes stable estimates. Furthermore, the magnitude plot for the SRIVC estimates show that the realizations are much more volatile than the SRIVC-uns counterparts, which is in agreement with the large standard deviations of the SRIVC parameters displayed in Table 8.2.

8.6.3 Discrete-time control, hybrid Box-Jenkins model structure

Instead of white noise, we now consider a colored noise in the loop and include it in the identification procedure. The discrete-time ARMA noise process is of the form

$$v(t_k) = \frac{1 + 0.5q^{-1}}{1 - 0.85q^{-1}} e(t_k),$$

where $e(t_k)$ is white noise of variance 0.01. Just as in the previous simulation examples, 15000 samples were obtained for identification and 300 Monte Carlo runs

Table 8.3: Mean and standard deviation of the system and noise model parameters obtained with 300 Monte Carlo runs with a discrete-time controller and a hybrid Box-Jenkins model structure.

Method	Stats.	$a_1(-0.0707)$	$a_2(0.5)$	$b_0(0.5)$	$b_1(-0.25)$	$c_1(0.5)$	$d_1(-0.85)$
RIVC	Mean S.Dev.	0.1593 $1.96 \cdot 10^{-2}$	0.4910 $6.20 \cdot 10^{-3}$	0.6045 $2.54 \cdot 10^{-2}$	-0.2532 $3.67 \cdot 10^{-3}$	0.5435 $3.24 \cdot 10^{-3}$	-0.9252 $8.98 \cdot 10^{-3}$
LSSVF-uns	Mean S.Dev.	-0.0645 $1.20 \cdot 10^{-2}$	0.4806 $1.15 \cdot 10^{-2}$	0.3976 $3.96 \cdot 10^{-2}$	-0.2436 $7.83 \cdot 10^{-3}$		
SRIVC-uns	Mean S.Dev.	-0.1022 $7.66 \cdot 10^{-3}$	0.5026 $2.96 \cdot 10^{-3}$	0.4807 $1.29 \cdot 10^{-2}$	-0.2616 $2.50 \cdot 10^{-3}$		
CLSRIVC-uns	Mean S.Dev.	-0.0707 $5.52 \cdot 10^{-3}$	0.5003 $3.72 \cdot 10^{-3}$	0.5004 $1.21 \cdot 10^{-2}$	-0.2499 $7.04 \cdot 10^{-3}$		
RIVC-uns	Mean S.Dev.	-0.0701 $4.35 \cdot 10^{-3}$	0.5003 $2.74 \cdot 10^{-3}$	0.4998 $8.43 \cdot 10^{-3}$	-0.2501 $1.43 \cdot 10^{-3}$	0.4998 $4.77 \cdot 10^{-3}$	-0.8494 $7.38 \cdot 10^{-3}$

were performed. The LSSVF-uns estimator is used as an initialization point for the RIVC, SRIVC-uns, CLSRIVC-uns and RIVC-uns estimators, and the ARMA estimation step of the RIVC-uns estimator is computed using the `armax` command of the System Identification Toolbox in MATLAB [143].

In Table 8.3 we present the mean and standard deviation of each parameter for each estimator under study. The SRIVC-uns estimator fails to provide accurate estimates due to the colored noise, which introduces bias in the method. This is not a problem for the CLSRIVC-uns estimator, as it is generically consistent even if the disturbance is colored. Again, these results mimic the behavior of the standard version of these estimators for stable, continuous-time system identification. The RIVC-uns method achieves excellent performance for estimating unstable systems in a hybrid BJ model structure, while the standard RIVC method is biased and does not deliver unstable models.

8.6.4 System with integral effect

In this final test we identify a continuous-time system with integral action that is studied in both open and closed-loop. The system and controller we consider in this simulation study are given by

$$G^*(p) = \frac{p-2}{p(p+1)}, \quad C_d(q) = -0.249 - 2.07(1 - q^{-1}),$$

with sampling period $h = 0.1[s]$. The additive noise $v(t_k)$ is a zero-mean Gaussian noise noise of variance 0.2. Two experimental conditions are tested: open-loop (i.e., without controller, and with input $u(t_k)$ being a zero-mean Gaussian noise of variance 1), and closed-loop (with the reference $r(t_k)$ also being zero-mean Gaussian noise noise with variance 1). The same number of samples and Monte Carlo runs as in the previous test are used to obtain estimates of the extended versions of

Table 8.4: Bias and MSE of the model parameters obtained from a system with integral action with 300 Monte Carlo runs.

Method	Stats.	a_0 (0)	a_1 (1)	b_0 (-2)	b_1 (1)
LSSVF-uns-1	Bias	$4.24 \cdot 10^{-7}$	$4.86 \cdot 10^{-3}$	$-9.24 \cdot 10^{-3}$	$6.68 \cdot 10^{-3}$
	MSE	$7.63 \cdot 10^{-11}$	$4.36 \cdot 10^{-4}$	$1.18 \cdot 10^{-3}$	$7.97 \cdot 10^{-4}$
SRIVC-uns-1	Bias	$-2.88 \cdot 10^{-8}$	$4.43 \cdot 10^{-4}$	$-9.09 \cdot 10^{-4}$	$1.22 \cdot 10^{-3}$
	MSE	$2.31 \cdot 10^{-12}$	$1.08 \cdot 10^{-4}$	$4.30 \cdot 10^{-4}$	$5.08 \cdot 10^{-4}$
SRIVC-uns-2	Bias	0	$1.65 \cdot 10^{-4}$	$-3.82 \cdot 10^{-4}$	$-8.25 \cdot 10^{-4}$
	MSE	0	$1.10 \cdot 10^{-4}$	$4.37 \cdot 10^{-4}$	$5.04 \cdot 10^{-4}$
CLSRIVC-uns-1	Bias	$6.60 \cdot 10^{-5}$	$-3.73 \cdot 10^{-3}$	$3.53 \cdot 10^{-3}$	$-3.85 \cdot 10^{-3}$
	MSE	$7.28 \cdot 10^{-7}$	$1.50 \cdot 10^{-3}$	$3.27 \cdot 10^{-3}$	$1.35 \cdot 10^{-3}$
CLSRIVC-uns-2	Bias	0	$-3.42 \cdot 10^{-3}$	$3.16 \cdot 10^{-3}$	$-3.62 \cdot 10^{-3}$
	MSE	0	$1.47 \cdot 10^{-3}$	$3.24 \cdot 10^{-3}$	$1.33 \cdot 10^{-3}$

LSSVF and SRIVC for open-loop, and CLSRIVC for closed-loop. The estimators that are computed following Solution 1 of Section 8.5 are labeled as LSSVF-uns-1, SRIVC-uns-1 and CLSRIVC-uns-1, and the ones computed from Solution 2 are labeled SRIVC-uns-2 and CLSRIVC-uns-2. The LSSVF-uns-1 estimator is initialized with a random polynomial $A_1(p) = p^2 + \tilde{a}_{1p} + \tilde{a}_0$, with \tilde{a}_0 and \tilde{a}_1 being the true parameters perturbed by white noise of standard deviation 0.1. The model from this estimator is used as initialization for the extended SRIVC and CLSRIVC algorithms.

The bias and mean square error (MSE) of each parameter is recorded and shown in Table 8.4. Forcing one pole to zero leads to a decrease in bias for the open and closed loop tests, while having similar MSE compared to the SRIVC and CLSRIVC variants obtained from Solution 1. At any rate, we verify that both solutions introduced in Section 8.5 are adequate for estimating systems with integration in open and closed-loop. The SRIVC and CLSRIVC estimators using Solution 1 are also seen to be robust despite the perturbation noise being of high variance, thanks to the ad-hoc stable filtering that is performed whenever the previous iteration returns an unstable estimate.

8.7 Conclusions

In this chapter we have presented extensions of refined instrumental variable methods that allow the direct and indirect identification of unstable continuous-time systems in closed-loop. The main technique consists of considering an ad-hoc noise model which iteratively cancels the unstable poles present in the model iterates. This tool has been applied successfully to the SRIVC and RIVC estimators, as well as to its initialization method, the LSSVF estimator. The same technique has been shown to

be promising for the CLSRIVC estimator when data from the reference signal is also collected. These estimators have also been extended to identify systems with integral action. In the case of continuous-time controllers without a hold device at their output, the approach we propose allows unstable models at the expense of bias in the parameters, and one way to mitigate this bias is to reduce the sampling period.

Appendix

8.A Proofs of technical results

8.A.1 Proof of Proposition 8.1

The model (8.7) can also be written as

$$y(t_k) = \left\{ \frac{B(p, \boldsymbol{\theta})}{A(p, \boldsymbol{\theta})} u(t) \right\}_{t=t_k} + \left(\frac{\tilde{C}(q, \boldsymbol{\eta}, \boldsymbol{\theta})}{\tilde{D}(q, \boldsymbol{\eta}, \boldsymbol{\theta})} - 1 \right) \varepsilon(t_k) + \varepsilon(t_k).$$

On the other hand,

$$\varepsilon(t_k) = \frac{\tilde{D}(q, \boldsymbol{\eta}, \boldsymbol{\theta})}{\tilde{C}(q, \boldsymbol{\eta}, \boldsymbol{\theta})} \left(y(t_k) - \left\{ \frac{B(p, \boldsymbol{\theta})}{A(p, \boldsymbol{\theta})} u(t) \right\}_{t=t_k} \right),$$

which leads to

$$\begin{aligned} y(t_k) &= \left\{ \frac{B(p, \boldsymbol{\theta})}{A(p, \boldsymbol{\theta})} u(t) \right\}_{t=t_k} + \left(\frac{\tilde{C}(q, \boldsymbol{\eta}, \boldsymbol{\theta})}{\tilde{D}(q, \boldsymbol{\eta}, \boldsymbol{\theta})} - 1 \right) \frac{\tilde{D}(q, \boldsymbol{\eta}, \boldsymbol{\theta})}{\tilde{C}(q, \boldsymbol{\eta}, \boldsymbol{\theta})} \left(y(t_k) - \left\{ \frac{B(p, \boldsymbol{\theta})}{A(p, \boldsymbol{\theta})} u(t) \right\}_{t=t_k} \right) + \varepsilon(t_k) \\ &= \frac{\tilde{D}(q, \boldsymbol{\eta}, \boldsymbol{\theta})}{\tilde{C}(q, \boldsymbol{\eta}, \boldsymbol{\theta})} \left\{ \frac{B(p, \boldsymbol{\theta})}{A(p, \boldsymbol{\theta})} u(t) \right\}_{t=t_k} + \left(1 - \frac{\tilde{D}(q, \boldsymbol{\eta}, \boldsymbol{\theta})}{\tilde{C}(q, \boldsymbol{\eta}, \boldsymbol{\theta})} \right) y(t_k) + \varepsilon(t_k). \end{aligned} \quad (8.25)$$

The one-step-ahead optimal predictor is given by the conditional expectation of $y(t_k)$ given the previous values of the input and output. Since $\varepsilon(t_k)$ is white noise, and the right-hand side of (8.25) depends solely on previous values of $y(t_k)$ and $u(t)$, the result follows. \square

8.A.2 Proof of Lemma 8.1

For a fixed noise model parameter vector $\boldsymbol{\eta}$, the prediction error cost can be written as

$$V(\boldsymbol{\theta}) = \frac{1}{N} \sum_{k=1}^N (y(t_k) - \hat{y}(t_k | \boldsymbol{\theta}))^2.$$

Its stationary points must satisfy the first-order condition $\partial V(\boldsymbol{\theta}) / \partial \boldsymbol{\theta} = \mathbf{0}$, that is,

$$\frac{2}{N} \sum_{k=1}^N (y(t_k) - \hat{y}(t_k | \boldsymbol{\theta}, \boldsymbol{\eta})) \frac{\partial}{\partial \boldsymbol{\theta}} \hat{y}(t_k | \boldsymbol{\theta}, \boldsymbol{\eta}) = \mathbf{0}.$$

On the other hand, a converging point $\bar{\boldsymbol{\theta}}^N$ of the iterative estimator in (8.17) satisfies

$$\bar{\boldsymbol{\theta}}^N = \left[\sum_{k=1}^N \hat{\boldsymbol{\varphi}}_{f,uns}(t_k, \bar{\boldsymbol{\theta}}^N) \boldsymbol{\varphi}_{f,uns}^\top(t_k, \bar{\boldsymbol{\theta}}^N) \right]^{-1} \left[\sum_{k=1}^N \hat{\boldsymbol{\varphi}}_{f,uns}(t_k, \bar{\boldsymbol{\theta}}^N) y_{f,uns}(t_k, \bar{\boldsymbol{\theta}}^N) \right]. \quad (8.26)$$

The dependence on $\bar{\boldsymbol{\theta}}^N$ in the filtered instrument, filtered regressor and filtered output has been made explicit in the expression above. Provided that the matrix inverse exists, (8.26) can be written as

$$\sum_{k=1}^N \hat{\boldsymbol{\varphi}}_{f,uns}(t_k, \bar{\boldsymbol{\theta}}^N) (y_{f,uns}(t_k, \bar{\boldsymbol{\theta}}^N) - \boldsymbol{\varphi}_{f,uns}^\top(t_k, \bar{\boldsymbol{\theta}}^N) \bar{\boldsymbol{\theta}}^N) = \mathbf{0}.$$

Now, note that

$$\begin{aligned} & y_{f,uns}(t_k, \bar{\boldsymbol{\theta}}^N) - \boldsymbol{\varphi}_{f,uns}^\top(t_k, \bar{\boldsymbol{\theta}}^N) \bar{\boldsymbol{\theta}}^N \\ &= \frac{\tilde{D}(q, \boldsymbol{\eta}, \bar{\boldsymbol{\theta}}^N)}{\tilde{C}(q, \boldsymbol{\eta}, \bar{\boldsymbol{\theta}}^N)} \frac{1 + \bar{a}_1 p + \cdots + \bar{a}_n p^n}{\bar{A}_N(p)} y(t_k) - \frac{\tilde{D}(q, \boldsymbol{\eta}, \bar{\boldsymbol{\theta}}^N)}{\tilde{C}(q, \boldsymbol{\eta}, \bar{\boldsymbol{\theta}}^N)} \left\{ \frac{\bar{b}_0 + \bar{b}_1 p + \cdots + \bar{b}_m p^m}{\bar{A}_N(p)} u(t) \right\}_{t=t_k} \\ &= \frac{\tilde{D}(q, \boldsymbol{\eta}, \bar{\boldsymbol{\theta}}^N)}{\tilde{C}(q, \boldsymbol{\eta}, \bar{\boldsymbol{\theta}}^N)} \left(y(t_k) - \left\{ \frac{\bar{B}_N(p)}{\bar{A}_N(p)} u(t) \right\}_{t=t_k} \right), \end{aligned}$$

where \bar{B}_N and \bar{A}_N are the numerator and denominator polynomials described by $\bar{\boldsymbol{\theta}}^N$. The last expression above coincides with the residual in (8.11), which means that the converging point of the estimator in (8.17) must solve

$$\frac{1}{N} \sum_{k=1}^N (y(t_k) - \hat{y}(t_k | \bar{\boldsymbol{\theta}}^N, \boldsymbol{\eta})) \hat{\boldsymbol{\varphi}}_{f,uns}(t_k, \bar{\boldsymbol{\theta}}^N) = \mathbf{0}.$$

Thus, setting $\hat{\boldsymbol{\varphi}}_{f,uns}(t_k, \boldsymbol{\theta})$ to be $\frac{\partial}{\partial \boldsymbol{\theta}} \hat{y}(t_k | \boldsymbol{\theta}, \boldsymbol{\eta})$ leads to the desired conclusion. \square

8.A.3 Proof of Lemma 8.2

At the converging point of the CLSIVC-uns estimator and as the sample size tends to infinity, the stationarity assumptions imply that the sums in (8.17) converge to their expected values [205]. This leads to

$$\bar{\boldsymbol{\theta}} = \mathbb{E} \{ \hat{\boldsymbol{\varphi}}_{f,uns}(t_k, \bar{\boldsymbol{\theta}}) \boldsymbol{\varphi}_{f,uns}^\top(t_k, \bar{\boldsymbol{\theta}}) \}^{-1} \mathbb{E} \{ \hat{\boldsymbol{\varphi}}_{f,uns}(t_k, \bar{\boldsymbol{\theta}}) y_{f,uns}(t_k, \bar{\boldsymbol{\theta}}) \}.$$

Under the assumption that (8.23) is non-singular, we find that the CLSIVC-uns estimator satisfies

$$\begin{aligned} & \mathbb{E} \{ \hat{\boldsymbol{\varphi}}_{f,uns}(t_k, \bar{\boldsymbol{\theta}}) (y_{f,uns}(t_k, \bar{\boldsymbol{\theta}}) - \boldsymbol{\varphi}_{f,uns}^\top(t_k, \bar{\boldsymbol{\theta}}) \bar{\boldsymbol{\theta}}) \} = \mathbf{0} \\ & \iff \mathbb{E} \left\{ \frac{A_{d,a}(q, \bar{\boldsymbol{\theta}})}{A_{d,a}(q, \bar{\boldsymbol{\theta}})} \hat{\boldsymbol{\varphi}}_f(t_k, \bar{\boldsymbol{\theta}}) \frac{A_{d,a}(q, \bar{\boldsymbol{\theta}})}{A_{d,a}(q, \bar{\boldsymbol{\theta}})} \left(y(t_k) - \frac{\bar{B}(p)}{\bar{A}(p)} u(t_k) \right) \right\} = \mathbf{0} \\ & \iff \mathbb{E} \left\{ \hat{\boldsymbol{\varphi}}_f(t_k, \bar{\boldsymbol{\theta}}) \left(y(t_k) - \frac{\bar{B}(p)}{\bar{A}(p)} u(t_k) \right) \right\} = \mathbf{0}, \end{aligned} \quad (8.27)$$

where the last equality holds since the filter $A_{d,a}(q, \boldsymbol{\theta})/\bar{A}_{d,a}(q, \boldsymbol{\theta})$ is all-pass. Comparing this with the derivation leading to (7.12), we see that (8.27) is the same pseudo-linear regression equation that the standard CLSRIVC method solves in the case when the estimates at each iteration are stable models. \square

8.B Filtered regressor and output when the input is an arbitrary signal

The following filtered regressor and output expressions are proposed to solve the instability problem when the input $u(t)$ is an arbitrary signal that can be oversampled with an oversampling rate S . Instead of considering the all-pass filter $A_{d,a}(q, \boldsymbol{\theta})/\bar{A}_{d,a}(q, \boldsymbol{\theta})$ for prefiltering the input, we introduce its fast-sampled delta version $A_{\Delta,a}(\delta, \boldsymbol{\theta})/\bar{A}_{\Delta,a}(\delta, \boldsymbol{\theta})$. If $1/A(p)$ has unstable poles $\{p_i\}_{i=1}^{n_a}$, the polynomial $A_{\Delta,a}(\delta, \boldsymbol{\theta})$ has roots of the form

$$\delta_i = \frac{S}{h} \left(e^{\frac{p_i h}{S}} - 1 \right), \quad i = 1, \dots, n_a.$$

Just like its q -domain version, the polynomial $\bar{A}_{\Delta,a}(\delta, \boldsymbol{\theta})$ is designed such that the filter $A_{\Delta,a}(\delta, \boldsymbol{\theta})/\bar{A}_{\Delta,a}(\delta, \boldsymbol{\theta})$ is all-pass. This means that $\bar{A}_{\Delta,a}(\delta, \boldsymbol{\theta})$ must be given by

$$\bar{A}_{\Delta,a}(\delta, \boldsymbol{\theta}) = \prod_{i=1}^{n_a} \left(\delta_i + \delta \left[1 + \frac{h}{S} \delta_i \right] \right).$$

We denote the oversampled input by \tilde{u} . With this in mind, we can fit an ARMA model to the estimated noise sequence

$$\hat{v}(t_k) = \frac{A_{d,a}(q, \boldsymbol{\theta}_j)}{A_{d,a}(q, \boldsymbol{\theta}_j)} y(t_k) - \frac{A_{\Delta,a}(\delta, \boldsymbol{\theta}_j)}{A_{\Delta,a}(\delta, \boldsymbol{\theta}_j)} \frac{B_{\Delta}(\delta, \boldsymbol{\theta}_j)}{A_{\Delta}(\delta, \boldsymbol{\theta}_j)} \tilde{u}(t_k),$$

where $B_{\Delta}(\delta, \boldsymbol{\theta}_j)/A_{\Delta}(\delta, \boldsymbol{\theta}_j)$ is the δ -domain equivalent [154] of the continuous-time transfer function $B(p, \boldsymbol{\theta}_j)/A(p, \boldsymbol{\theta}_j)$. After this estimation step, we compute the filtered signals in the δ -domain as

$$\begin{aligned} y_{f,uns}(t_k) &= \frac{D(q, \boldsymbol{\eta}_{j+1})}{C(q, \boldsymbol{\eta}_{j+1})} \frac{A_{d,a}(q, \boldsymbol{\theta}_j)}{\bar{A}_{d,a}(q, \boldsymbol{\theta}_j)} \frac{1}{A(p, \boldsymbol{\theta}_j)} y(t_k) \\ \boldsymbol{\varphi}_{f,uns}(t_k) &= \frac{D(q, \boldsymbol{\eta}_{j+1})}{C(q, \boldsymbol{\eta}_{j+1})} \left[\frac{A_{d,a}(q, \boldsymbol{\theta}_j)}{\bar{A}_{d,a}(q, \boldsymbol{\theta}_j)} \frac{-p}{A(p, \boldsymbol{\theta}_j)} y(t_k), \dots, \frac{A_{d,a}(q, \boldsymbol{\theta}_j)}{\bar{A}_{d,a}(q, \boldsymbol{\theta}_j)} \frac{-p^n}{A(p, \boldsymbol{\theta}_j)} y(t_k), \right. \\ &\quad \left. \frac{A_{\Delta,a}(\delta, \boldsymbol{\theta}_j)}{\bar{A}_{\Delta,a}(\delta, \boldsymbol{\theta}_j)} \frac{B_{\Delta}^0(\delta, \boldsymbol{\theta}_j)}{A_{\Delta}(\delta, \boldsymbol{\theta}_j)} \tilde{u}(t_k), \dots, \frac{A_{\Delta,a}(\delta, \boldsymbol{\theta}_j)}{\bar{A}_{\Delta,a}(\delta, \boldsymbol{\theta}_j)} \frac{B_{\Delta}^m(\delta, \boldsymbol{\theta}_j)}{A_{\Delta}(\delta, \boldsymbol{\theta}_j)} \tilde{u}(t_k) \right]^T, \end{aligned}$$

where $B_{\Delta}^i(\delta, \boldsymbol{\theta}_j)/A_{\Delta}(\delta, \boldsymbol{\theta}_j)$ is the δ -domain equivalent of $p^i/A(p, \boldsymbol{\theta}_j)$ for $i = 0, \dots, m$.

All the unstable poles of the estimated model at each iteration are canceled under this formulation. The main drawback of this approach is that different all-pass

filters are used for constructing the filtered output and filtered input signals. In other words, the model structure that is implicitly used is not (8.7), but rather

$$y(t_k) = \frac{\bar{A}_{d,a}(q)}{A_{d,a}(q)} \frac{A_{\Delta,a}(\delta)}{\bar{A}_{\Delta,a}(\delta)} \frac{B_{\Delta}(\delta)}{A_{\Delta}(\delta)} \tilde{u}(t_k) + \frac{\tilde{C}(q)}{\tilde{D}(q)} \varepsilon(t_k). \quad (8.28)$$

However, as the sampling period tends to zero, we recover (8.7) since the all-pass filters in the first summand cancel out and the delta description converges to the continuous-time one [154].

8.C On computing the derivatives in (8.19)

The following expressions can be used for the implementation of the refined instrumental variable method in (8.17). Denote the numerator of the ZOH equivalent of $B(p, \boldsymbol{\theta})/A(p, \boldsymbol{\theta})$ as $B_d(q, \boldsymbol{\theta})$. Since the noise model estimate $C(q, \boldsymbol{\eta})/D(q, \boldsymbol{\eta})$ does not depend on $\boldsymbol{\theta}$, we only need to compute derivatives of the form

$$\frac{\partial}{\partial a_l} \left\{ \frac{B_d(q, \boldsymbol{\theta})}{A_{d,s}(q, \boldsymbol{\theta}) \bar{A}_{d,a}(q, \boldsymbol{\theta})} \right\} = \frac{\frac{\partial B_d(q, \boldsymbol{\theta})}{\partial a_l}}{A_{d,s}(q, \boldsymbol{\theta}) \bar{A}_{d,a}(q, \boldsymbol{\theta})} - \frac{B_d(q, \boldsymbol{\theta}) \frac{\partial A_{d,s}(q, \boldsymbol{\theta}) \bar{A}_{d,a}(q, \boldsymbol{\theta})}{\partial a_l}}{[A_{d,s}(q, \boldsymbol{\theta}) \bar{A}_{d,a}(q, \boldsymbol{\theta})]^2}, \quad (8.29)$$

and

$$\frac{\partial}{\partial a_l} \left\{ \frac{A_{d,a}(q, \boldsymbol{\theta})}{\bar{A}_{d,a}(q, \boldsymbol{\theta})} \right\} = \frac{1}{\bar{A}_{d,a}(q, \boldsymbol{\theta})} \frac{\partial A_{d,a}(q, \boldsymbol{\theta})}{\partial a_l} - \frac{A_{d,a}(q, \boldsymbol{\theta})}{[\bar{A}_{d,a}(q, \boldsymbol{\theta})]^2} \frac{\partial \bar{A}_{d,a}(q, \boldsymbol{\theta})}{\partial a_l}.$$

The partial derivative of $A_{d,s}(q, \boldsymbol{\theta}) \bar{A}_{d,a}(q, \boldsymbol{\theta})$ with respect to a_l can be computed by

$$\frac{\partial A_{d,s}(q, \boldsymbol{\theta}) \bar{A}_{d,a}(q, \boldsymbol{\theta})}{\partial a_l} = \bar{A}_{d,a}(q, \boldsymbol{\theta}) \frac{\partial A_{d,s}(q, \boldsymbol{\theta})}{\partial a_l} + A_{d,s}(q, \boldsymbol{\theta}) \frac{\partial \bar{A}_{d,a}(q, \boldsymbol{\theta})}{\partial a_l},$$

which reduces the problem to differentiating $A_{d,s}(q, \boldsymbol{\theta})$, $\bar{A}_{d,a}(q, \boldsymbol{\theta})$ and $B_d(q, \boldsymbol{\theta})$. For $A_{d,s}(q, \boldsymbol{\theta})$, we exploit (8.6) and the chain rule for differentiation to obtain

$$\frac{\partial A_{d,s}(q, \boldsymbol{\theta})}{\partial a_l} = \sum_{j=1}^n \frac{\partial A_{d,s}(q, \boldsymbol{\theta})}{\partial a_j^d} \frac{\partial a_j^d}{\partial a_l} = \sum_{j=1}^n \sum_{k=0}^{n_s} \frac{\partial \tilde{a}_k^{d,s}}{\partial a_j^d} \frac{\partial a_j^d}{\partial a_l} q^{-k},$$

where $\partial a_j^d / \partial a_l$ can be obtained by differentiating (8.4), and

$$\frac{\partial \tilde{a}_k^{d,s}}{\partial a_j^d} = \begin{cases} \frac{1}{\tilde{a}_{j-k}^{d,a}} & \text{if } j - n_s \leq k \leq j, \\ 0 & \text{otherwise.} \end{cases}$$

On the other hand, by using the fact that $\bar{A}_{d,a}(q, \boldsymbol{\theta}) = q^{-n_a} A_{d,a}(q^{-1}, \boldsymbol{\theta}) / a_{n_a}^{d,a}$, the partial derivative of this polynomial with respect to a_l is given by

$$\begin{aligned}\frac{\partial \bar{A}_{d,a}(q, \boldsymbol{\theta})}{\partial a_l} &= \frac{q^{-n_a}}{a_{n_a}^{d,a}} \frac{\partial A_{d,a}(q^{-1}, \boldsymbol{\theta})}{\partial a_l} - \frac{q^{-n_a} A_{d,a}(q^{-1}, \boldsymbol{\theta})}{(a_{n_a}^{d,a})^2} \frac{\partial a_{n_a}^{d,a}}{\partial a_l} \\ &= \frac{q^{-n_a}}{a_{n_a}^{d,a}} \sum_{j=1}^n \sum_{k=0}^{n_a} \frac{\partial \tilde{a}_k^{d,a}}{\partial a_j^d} \frac{\partial a_j^d}{\partial a_l} q^k - \frac{q^{-n_a} A_{d,a}(q^{-1}, \boldsymbol{\theta})}{(a_{n_a}^{d,a})^2} \frac{\partial a_{n_a}^{d,a}}{\partial a_l},\end{aligned}$$

where

$$\frac{\partial \tilde{a}_k^{d,a}}{\partial a_j^d} = \begin{cases} \frac{1}{\tilde{a}_{j-k}^{d,s}} & \text{if } j - n_a \leq k \leq j, \\ 0 & \text{otherwise.} \end{cases}$$

Finally, for the derivative of $B_d(q, \boldsymbol{\theta})$ in (8.29), we compute this polynomial in terms of the continuous-time parameter vector $\boldsymbol{\theta}$. To illustrate the idea, we assume that $A(p, \boldsymbol{\theta})$ has no roots at $p = 0$ and that $m < n$. We have

$$B_d(q, \boldsymbol{\theta}) = \frac{q^{-n}}{a_n} \mathbf{C} \text{adj}(q\mathbf{I} - \exp(\mathbf{A}h)) \mathbf{A}^{-1} (\exp(\mathbf{A}h) - \mathbf{I}) \mathbf{B},$$

where \mathbf{A} is given by (8.5), $\mathbf{B} = [1, 0, \dots, 0]^\top$, and $\mathbf{C} = [0, \dots, 0, b_m, \dots, b_0]$. The equations above are explicit in the parameter a_l and therefore allow direct differentiation of the polynomial $B_d(q, \boldsymbol{\theta})$ for $l = 1, \dots, n$.

Chapter 9

Discrete-time representations of systems with band-limited inputs

This chapter departs from the analyses and extensions of the instrumental variable methods presented in Chapters 4 to 8, and focuses on discrete-time descriptions of continuous-time systems when band-limited inputs excite them. Here we show that the band-limited equivalent impulse response of a continuous-time system is non-causal, and we study the properties of this impulse response.

9.1 Introduction

Even though usually only sampled data can be extracted from a measured signal, it is necessary to also take into consideration the intersample behavior that can be assumed from the signal under study. Three main assumptions can be commonly found: that the signal is piecewise constant (zero-order hold), piecewise linear (first-order hold), or band-limited, which means that the power spectrum of the signal is zero above a certain frequency [175].

Due to the advantages provided by the Nyquist-Shannon reconstruction theorem, which permits the exact intersample behavior of the signal to be known based on discrete-time samples, band-limited signals have been studied extensively in signal processing, filter theory, and spectral theory [111, 169, 222]. In this chapter we show that the equivalent discrete-time LTI system is non-causal when the input is assumed to be band-limited. A similar observation has been made in [49], where it is stated that the reconstruction of band-limited signals is a non-causal filtering procedure. In signal processing, the convolution operation with band-limited signals is usually carried out in the frequency domain [22], where the description of these signals is more natural. Most of the work in this area has focused on the approximation of band-limited signals [20, 21], and the convergence of their series representations [105]. In system identification, however, the implications of this sampling result to the equivalent discrete-time system description seem to have been overlooked in the literature. For example, a strictly causal discrete-time model was obtained under

band-limited measurement conditions in [187], and the effect of setting the direct term to zero was analyzed. Equation (2) of [187] adds a direct term to the discrete-time impulse response instead of a fully non-causal discrete-time representation. On a similar note, band-limited input signals were used in [184] for estimating a discrete-time finite-impulse response (FIR) filter; this work was done in the discrete-time domain and the non-causal components that arise in the system description were also neglected.

The present chapter explores the properties of band-limited equivalent discrete-time systems, and lays the theoretical foundations for the identification of continuous-time systems with band-limited inputs that is covered in Chapter 10. In summary,

- we show that the equivalent discrete-time system arising from a band-limited input signal is non-causal, and we discuss what conditions lead to more significant non-causal terms;
- we provide several formulas for computing the band-limited discrete-time impulse response equivalent and study some of its properties with respect to rational transfer functions; and
- we prove that oversampling in conjunction with a trapezoidal anti-aliasing filter leads to uniform convergence of the band-limited equivalent to the continuous-time impulse response as the sampling period tends to zero.

The rest of this chapter is organized as follows. Essentials on band-limited signals and the main discrete-time representation result can be found in Section 9.2. In Section 9.3 we explore more properties of this discrete-time equivalent and analyze the effect of some anti-aliasing filters in the representations, while Section 9.4 concludes this chapter.

9.2 Preliminaries and main non-causality result

In this section we discuss the topic of sampling band-limited signals from a continuous-time system standpoint. In particular, we recall the concept of a band-limited signal, and introduce the non-causal impulse response that is obtained from the equivalent discrete-time system under band-limited assumptions in the input signal.

Consider the system description

$$x(t) = \int_0^\infty g(\tau)u(t - \tau)d\tau, \quad (9.1)$$

where $u(t)$ is a scalar continuous-time input of the asymptotically stable, single-input single-output, linear and time-invariant system that has a causal impulse response $g(t)$, and where $x(t)$ is the output. The frequency response $G(i\omega)$ is the Fourier transform of $g(t)$, i.e.,

$$G(i\omega) := \mathcal{F}\{g(t)\} = \int_{-\infty}^{\infty} g(t)e^{-i\omega t}dt.$$

The key assumption used throughout the following two chapters is that the input signal is *band-limited*, that is, the signal $u(t)$ does not have energy above a certain frequency ω_B . In other words, $U(i\omega) = 0$ for $|\omega| > \omega_B$, where $U(i\omega)$ is the continuous-time Fourier transform, and $\omega_B > 0$ is the (finite) bandwidth of the input signal.

Assume that $u(t)$ is sampled every h seconds, where $h \leq \pi/\omega_B$. Then, we can use its discrete-time Fourier transform pair

$$U_h(e^{i\omega h}) = h \sum_{k=-\infty}^{\infty} u(kh)e^{-i\omega kh} \iff u(kh) = \frac{1}{2\pi} \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} U_h(e^{i\omega h}) e^{i\omega kh} d\omega \quad (9.2)$$

to write

$$\begin{aligned} u(kh) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} U(i\tilde{\omega}) e^{i\tilde{\omega} kh} d\tilde{\omega} \\ &= \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \int_{\frac{2\pi n}{h} - \frac{\pi}{h}}^{\frac{2\pi n}{h} + \frac{\pi}{h}} U(i\tilde{\omega}) e^{i\tilde{\omega} kh} d\tilde{\omega} \\ &\stackrel{(\tilde{\omega}=\omega+\frac{2\pi n}{h})}{=} \frac{1}{2\pi} \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} \sum_{n=-\infty}^{\infty} U\left(i\omega + i\frac{2\pi n}{h}\right) e^{i\omega kh} d\omega. \end{aligned}$$

Comparing with (9.2), we find that

$$U_h(e^{i\omega h}) = \sum_{n=-\infty}^{\infty} U\left(i\omega + i\frac{2\pi n}{h}\right),$$

which is known as Poisson's summation formula [19]. Due to $u(t)$ being band-limited, this formula indicates that $U_h(e^{i\omega h}) = U(i\omega)$ for $|\omega| \leq \pi/h$. Since $X(i\omega) = G(i\omega)U(i\omega)$, we find that $x(t)$ is also band-limited and therefore $X_h(e^{i\omega h}) = X(i\omega)$ in the same domain. Using these identities we can exactly reconstruct a continuous-time band-limited signal based on its samples. By leveraging the inverse Fourier transform of $U_h(e^{i\omega h})$, we have

$$\begin{aligned} u(t) &= \frac{1}{2\pi} \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} U_h(e^{i\omega h}) e^{i\omega t} d\omega \\ &= \frac{1}{2\pi} \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} h \sum_{n=-\infty}^{\infty} u(nh) e^{-i\omega nh} e^{i\omega t} d\omega \\ &= \sum_{n=-\infty}^{\infty} u(nh) \frac{h}{2\pi} \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} e^{i\omega(t-nh)} d\omega \\ &= \sum_{n=-\infty}^{\infty} u(nh) \text{sinc}\left(\frac{t-nh}{h}\right), \end{aligned} \quad (9.3)$$

where the sinc function is defined as $\text{sinc}(t) := (\pi t)^{-1} \sin(\pi t)$. Replacing this description of $u(t)$ in (9.1) and interchanging summation and integration, the system equation can then be rewritten as

$$x(t) = \sum_{n=-\infty}^{\infty} u(nh) \int_0^{\infty} g(\tau) \text{sinc}\left(\frac{t-\tau-nh}{h}\right) d\tau. \quad (9.4)$$

Thus, we have derived the central result of this chapter.

Theorem 9.1. *The equivalent discrete-time model of a system whose input is a band-limited signal with bandwidth ω_B and sampling period $h \leq \pi/\omega_B$ is described by the impulse response*

$$g_{\text{BL}}(kh) := \frac{1}{h} \int_0^{\infty} g(\tau) \text{sinc}\left(\frac{kh-\tau}{h}\right) d\tau, \quad k \in \mathbb{Z}. \quad (9.5)$$

If we introduce the non-causal discrete-time transfer function

$$G_{\text{BL}}(z) := \sum_{k=-\infty}^{\infty} g_{\text{BL}}(kh) z^{-k} \quad (9.6)$$

with $g_{\text{BL}}(kh)$ being defined as in (9.5), then the continuous-time frequency response of the system satisfies

$$G(i\omega) = h G_{\text{BL}}(e^{i\omega h}) \quad \text{for all } \omega \in [-\pi/h, \pi/h]. \quad (9.7)$$

The band-limited equivalent is unique. That is, given a band-limited equivalent there is only one finite-order transfer function $G(i\omega)$ that satisfies (9.7).

Proof. Equation (9.7) follows directly from computing the discrete-time Fourier transform of (9.4) when evaluated at $t = kh$ and recalling that $x(t)$ is a band-limited signal. Finally, the uniqueness result follows from supposing that two finite-order transfer functions $G_i(p) = B_i(p)/A_i(p)$, $i = 1, 2$ satisfy (9.7). Then $B_1(i\omega)/A_1(i\omega) - B_2(i\omega)/A_2(i\omega) = 0$ for $|\omega| \leq \pi/h$, which implies that $B_1(i\omega)A_2(i\omega) - B_2(i\omega)A_1(i\omega) = 0$ holds over some open set in $(-\pi/h, \pi/h)$. Since G_1 and G_2 are of finite order, we must have $G_1(i\omega) = G_2(i\omega)$. This concludes the proof of Theorem 9.1. \square

On another note, we see from (9.7) that any pole s of the continuous-time system that satisfies $|s| \leq \pi/h$ is mapped to a pole $z = e^{sh}$ of $G_{\text{BL}}(z)$, whereas poles that are not captured by the bandwidth given by the Nyquist frequency cannot be accounted for in $G_{\text{BL}}(z)$.

Remark 9.1. The band-limited equivalent impulse response $g_{\text{BL}}(kh)$ can also provide an exact description of the continuous-time output signal based on the

input samples. It can be readily seen from (9.4) that the following convolution result holds:

$$\int_0^\infty g(\tau)u(t-\tau)d\tau = h \sum_{k=-\infty}^{\infty} u(kh)g_{BL}(t-kh).$$

This result is well known for when g is a band-limited function with bandwidth π/h . In that case, we have $g(t) = g_{BL}(t)$ and we recover Equation (1.3) of [159, Ch. 3]. However, this work deals with systems that are not band-limited, that is, continuous-time systems whose impulse response is not a band-limited signal.

The expression in (9.5) indicates that the impulse response $\{g_{BL}(kh)\}_{k \in \mathbb{Z}}$ is non-causal in general. That is, *a causal continuous-time system behaves like a non-causal system when sampled assuming a band-limited intersample behavior in the input*. Intuitively, this non-causal behavior can be deduced by how the intersample behavior of the input is formed: by (9.3), we see that the sinc interpolation of the input must take into consideration the contributions of all the future values of the input at the sampling instants. Thus, the output of the system will be a function of these future input values as well. It should be emphasized that this phenomenon has been observed in signal processing [222] but with particular focus on band-limited systems, i.e., linear systems whose impulse responses are band-limited signals. Here we take a different approach as the linear system G is assumed to not be band-limited in general, and we do not study approximation of band-limited signals. Rather, we investigate what properties the equivalent sampled-data system enjoys and how to exploit them for identification purposes.

A more precise statement regarding the non-causal representation is given in Proposition 9.1, where we show that a causal band-limited equivalent impulse response may only be obtained under a stringent condition.

Proposition 9.1. Consider the asymptotically stable, causal impulse response $\{g(t)\}_{t \geq 0}$ and its band-limited impulse response equivalent $\{g_{BL}(kh)\}_{k \in \mathbb{Z}}$. The equivalent band-limited impulse response is causal (i.e., $g_{BL}(kh) = 0$ for $k < 0$) if and only if

$$\int_0^\infty g(\tau)f(\tau)d\tau = 0$$

for every $f \in \mathcal{P}$, where

$$\mathcal{P} = \left\{ f: \mathbb{R} \rightarrow \mathbb{R}; f(t) = \sum_{l=-\infty}^{-1} f(lh) \text{sinc}\left(\frac{lh-t}{h}\right), \{f(lh)\}_{l \in \mathbb{Z}} \in \ell_2 \right\}.$$

Proof. See Appendix 9.A.1. □

Remark 9.2. A time-delay system with impulse response $g(t) = \delta(t-nh)$, where n is a fixed non-negative integer, has an equivalent band-limited impulse response $g_{BL}(kh) = h^{-1}\delta_K(k-n)$. By extrapolating this observation, it is easy to see that any system whose output can be represented as a weighted sum of continuous-time

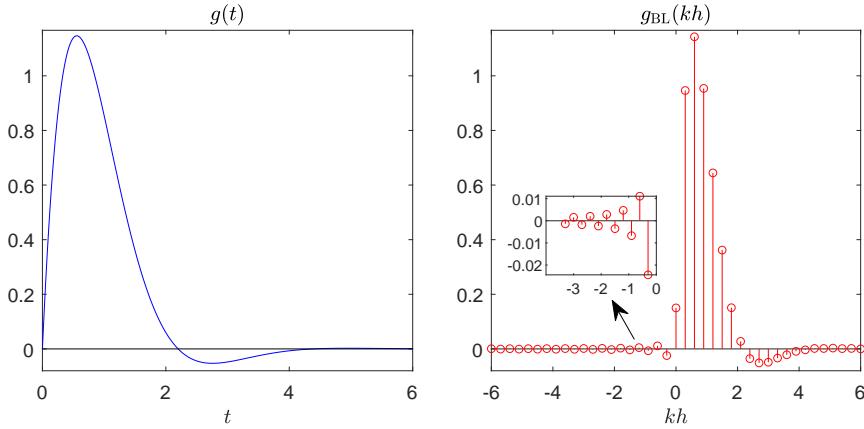


Figure 9.1: Left: Impulse response of a second order continuous-time system $g(t) = 3.5e^{-1.4t} \sin(1.43t)$. Right: Impulse response of its discrete-time band-limited equivalent.

delayed inputs admits a causal representation. On another note, a similar result to Proposition 9.1 can be obtained for describing the set of causal impulse responses that have a finite-time discrete-time equivalent impulse response (i.e., $g_{BL}(kh) = 0$ for $k < -M_{nc}$ and $k > M_c$, with $-M_{nc} \leq M_c$). We do not state this result explicitly in our work.

An example of this non-causal behavior can be seen in Figure 9.1, where we have plotted the impulse response $g(t)$ of a second order continuous-time system and the impulse response of its discrete-time band-limited equivalent $g_{BL}(kh)$, which is non-causal. With regards to the behavior of the non-causal part of $\{g_{BL}(kh)\}_{k \in \mathbb{Z}}$, we find that

- A significant non-causal part is present if the impulse response $g(t)$ correlates with $\text{sinc}(t/h - k)$. For example, consider $g(t) = -e^{-0.2t} \sin(\frac{\pi}{1.1}t)$ and sampling period $h = 1[\text{s}]$. The cutoff frequency is $\pi/1.1[\text{rad/s}]$, which is less than the Nyquist frequency $\pi/h = \pi[\text{rad/s}]$. As seen in Figure 9.2, the impulse response has an important overlap with the sinc function, which induces considerable non-causal values in $g_{BL}(kh)$. Note that such correlation is more likely to occur when the sampling frequency is close to the cutoff frequency, as in this example.
- As the sampling period tends to zero, the non-causal part vanishes. In fact, since $h^{-1}\text{sinc}([t-\tau]h)$ converges weakly to $\delta(t-\tau)$ (see, e.g., [121]), we see that for any fixed t where g is continuous we have $g_{BL}(t) \rightarrow g(t)$ as h tends to zero.
- The convergence of $g_{BL}(t)$ to $g(t)$ is not uniform. This can be observed by noting that $g_{BL}(t)$ can be interpreted as a convolution between $g(t)$ and the sinc kernel, which is not a “good kernel” (as defined in [215, p. 139]) since the sinc

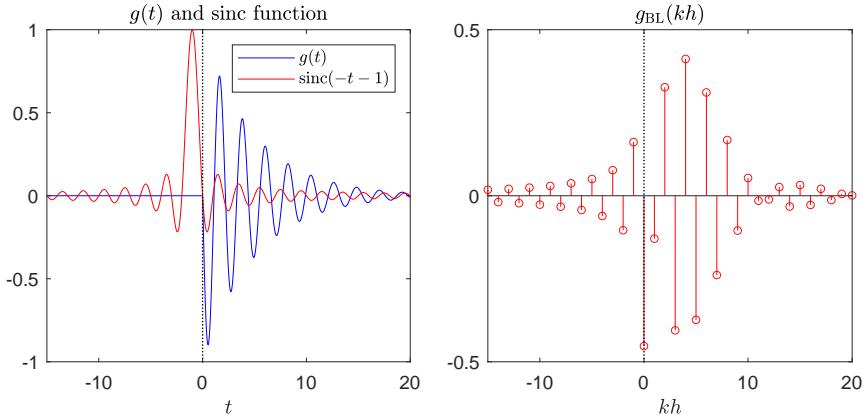


Figure 9.2: Left: Impulse response $g(t)$ with the sinc function that is used for computing $g_{BL}(-1)$ with (9.5). Since the most pronounced lobes are synchronized, the resulting impulse response coefficient is significant. Right: Band-limited equivalent impulse response of $g(t)$.

function is not absolutely integrable. Therefore, uniform convergence results such as Corollary 1.7 of [215] do not apply in this context. This discussion resembles the one for Dirichlet kernels in Fourier analysis [48, Ch. 5].

9.3 Properties of g_{BL} and oversampling variants

In this section we will expand on the insights obtained from Theorem 9.1. We first write alternative forms for $g_{BL}(kh)$, and we find what type of systems admit closed-form expressions for this impulse response equivalent. After this, we study the effect of oversampling and the extra degrees of freedom that oversampling provides for designing better reconstruction filters.

9.3.1 Alternative representations and properties

Corollary 9.1. *We can write the band-limited equivalent impulse response as*

$$g_{BL}(kh) = \frac{1}{2\pi} \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} G(i\omega) e^{i\omega kh} d\omega = \frac{1}{\pi} \int_0^{\frac{\pi}{h}} \operatorname{Re}\{G(i\omega) e^{i\omega kh}\} d\omega. \quad (9.8)$$

The non-causal part of the band-limited equivalent impulse response is also given by

$$g_{BL}(-kh) = -\frac{1}{2\pi} \int_{|\omega|>\frac{\pi}{h}} G(i\omega) e^{-i\omega kh} d\omega, \quad k > 0, \quad (9.9)$$

and the frequency response pertaining the non-causal part, denoted as $\{G(i\omega)\}_-$, can be computed as

$$\begin{aligned}\{G(i\omega)\}_- &= h \sum_{k=-\infty}^{-1} g_{BL}(kh) e^{-i\omega kh} \\ &= \frac{G(i\omega)}{2} - \frac{h}{4\pi} \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} G(iv) dv - \frac{1}{4\pi i} \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} G(iv) \cot\left(\frac{[\omega-v]h}{2}\right) dv.\end{aligned}$$

Proof. See Appendix 9.A.2. \square

An expression similar to the first equality in (9.8) has been derived in, e.g., [22], where convergence of the convolution sums with band-limited signals is studied. It is worth noting that the expression in (9.8) is closely related to the zero-order hold equivalent impulse response $\{g_{ZOH}(kh)\}_{k=0}^{\infty}$ with equivalent transfer function $G_{ZOH}(e^{i\omega h})$, which can be written as

$$g_{ZOH}(kh) = \frac{h}{2\pi} \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} G_{ZOH}(e^{i\omega h}) e^{i\omega kh} d\omega. \quad (9.10)$$

Also, (9.9) offers another interpretation of the non-causal part of g_{BL} : $-g_{BL}(kh)$ for $k < 0$ is the impulse response corresponding to the tail of the transfer function $G(i\omega)$ that is not captured by the Nyquist bandwidth $[-\pi/h, \pi/h]$. The interval of integration in (9.9) shrinks as h decreases, in which case small values for the non-causal part of g_{BL} are expected. Inappreciable values of $g_{BL}(kh)$ for $k < 0$ will also be obtained if the frequency response $G(i\omega)$ does not have much energy outside the Nyquist frequency band.

The next two corollaries provide more properties of $G_{BL}(z)$ in (9.6) and the band-limited continuous-time reconstruction of $g_{BL}(kh)$, and they also follow from Theorem 9.1.

Corollary 9.2. *Assume that the system G is strictly proper and asymptotically stable. Then, the \mathcal{H}_2 -norm of the system G satisfies $\|G\|_2^2 \geq h\|G_{BL}\|_2^2$.*

Proof. See Appendix 9.A.3. \square

Corollary 9.3. *Consider the band-limited continuous-time reconstruction of $g_{BL}(kh)$, i.e.,*

$$\tilde{g}_{BL}(t) := \sum_{k=-\infty}^{\infty} g_{BL}(kh) \operatorname{sinc}\left(\frac{t-kh}{h}\right).$$

Its Fourier transform is given by $\mathcal{F}\{\tilde{g}_{BL}(t)\} = G(i\omega) \operatorname{rect}(\omega h / 2\pi)$, where the rectangular function $\operatorname{rect}(\cdot)$ is defined by

$$\operatorname{rect}(x) = \begin{cases} 1 & \text{if } |x| < 1/2, \\ 1/2 & \text{if } |x| = 1/2, \\ 0 & \text{otherwise.} \end{cases}$$

The supremum of $|\tilde{g}_{\text{BL}}(t)|$ over $t \in \mathbb{R}$ is upper-bounded by $\|G_{\text{BL}}\|_2$. Furthermore, if $u(t)$ is band-limited with bandwidth $\omega_B \leq \pi/h$, then for any integer k we have

$$\int_{-\infty}^{\infty} \tilde{g}_{\text{BL}}(\tau) u(kh - \tau) d\tau = h \sum_{n=-\infty}^{\infty} g_{\text{BL}}(nh) u([k-n]h). \quad (9.11)$$

Proof. See Appendix 9.A.4. \square

Another topic of interest is whether it is possible to explicitly solve the integrals that define $\{g_{\text{BL}}(kh)\}_{k \in \mathbb{Z}}$ in (9.5) given a continuous-time transfer function of finite order. We claim that any transfer function $G(i\omega)$ that is a rational function in $e^{i\omega h}$ has an explicit expression for $g_{\text{BL}}(kh)$ that can easily be obtained via the residue theorem¹[191]; the same conclusion can be obtained by observing (9.8) from Corollary 9.1 and comparing it with the zero-order hold formula in (9.10). In contrast, even simple rational continuous-time transfer functions in $i\omega$ do not admit straightforward expressions for the band-limited equivalent $g_{\text{BL}}(kh)$. An example of this fact is provided below.

Example 9.1. Consider the following frequency response functions

$$G_1(i\omega) = \frac{h}{e^{i\omega h} - a}; \quad G_2(i\omega) = \frac{1}{i\omega + b},$$

where $|a| < 1$, $b > 0$, and h is the sampling period. Thanks to the residue theorem, the band-limited equivalent impulse response of $G_1(i\omega)$ is given by

$$g_{\text{BL},1}(kh) = \frac{1}{2\pi i} \oint_{|z|=1} \frac{z^{k-1}}{z-a} dz = a^{k-1} \mu(k-1),$$

where $\mu(\cdot)$ is the Heaviside function. Thus, the band-limited equivalent transfer function $G_{\text{BL},1}(z)$ is equal to

$$G_{\text{BL},1}(z) = \frac{1}{a} \sum_{k=1}^{\infty} \left(\frac{a}{z}\right)^k = \frac{1}{z-a}, \quad |z| > a.$$

Note that $G_{\text{BL},1}(z)$ describes a causal discrete-time transfer function. On the other hand, the band-limited equivalent impulse response of $G_2(i\omega)$ can be written as

$$g_{\text{BL},2}(kh) = \frac{1}{h} \int_0^{\infty} e^{-b\tau} \text{sinc}\left(\frac{kh-\tau}{h}\right) d\tau = e^{-bhk} \int_{-k}^{\infty} e^{-bx} \text{sinc}(x) dx.$$

This last integral can be written in terms of the exponential integral function [1] with a complex-valued argument. Unfortunately, these integrals are known to not have descriptions in terms of elementary functions. Consequently, the transfer function $G_{\text{BL},2}(z)$ is not a rational transfer function of the complex variable z .

¹In fact, the coefficients of the band-limited equivalent impulse response can be related to a Laurent series of $G_{\text{BL}}(z)$ around $z = 0$, but we will not exploit this connection explicitly here.

9.3.2 Oversampling and sampling variants

So far we have not made any consideration on how close the bandwidth ω_B of the input signal is to the Nyquist frequency. Here we take a look at the implications of having a smaller input bandwidth than the one allowed by the Nyquist-Shannon theorem, which is equivalent to analyzing the effect of oversampling the input signal. With a smaller input bandwidth there is an extra degree of freedom in the selection of the anti-aliasing filter that is implicitly being considered, and therefore other discrete-time equivalents can be computed.

Let us introduce a filter $F(i\omega)$ that is equal to one when $|\omega| \leq \omega_B$, zero when $|\omega| > \pi/h$, and left unspecified otherwise. It is known that

$$u(t) = \frac{1}{2\pi} \int_{-\omega_B}^{\omega_B} U(i\omega) e^{i\omega t} d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} U(i\omega) F(i\omega) e^{i\omega t} d\omega.$$

Since $u(t)$ is band-limited, its continuous-time Fourier transform can be described as $U_h(e^{i\omega h})$ in (9.2), which after some computations results in

$$u(t) = h \sum_{n=-\infty}^{\infty} u(nh) f(t - nh),$$

where $f(t)$ is the inverse Fourier transform of $F(i\omega)$, i.e.,

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(i\omega) e^{i\omega t} d\omega.$$

Proposition 9.2. Consider a continuous-time LTI system g and the anti-aliasing filter $F(i\omega)$ defined above. The discrete-time equivalent impulse response of g for band-limited inputs of bandwidth $\omega_B < \pi/h$ is given by

$$g_{BL}^f(kh) = \int_0^\infty g(\tau) f(kh - \tau) d\tau = \frac{1}{2\pi} \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} G(i\omega) F(i\omega) e^{i\omega kh} d\omega = h \sum_{l=-\infty}^{\infty} g_{BL}(lh) f([k-l]h), \quad (9.12)$$

where $g_{BL}(lh)$ is defined in (9.5). The impulse response equivalent $g_{BL}^f(kh)$ satisfies

$$h \sum_{k=-\infty}^{\infty} g_{BL}^f(kh) e^{-i\omega kh} = G(i\omega) F(i\omega), \quad \text{for all } \omega \in [-\pi/h, \pi/h], \quad (9.13)$$

and

$$h \sum_{k=-\infty}^{\infty} |g_{BL}^f(kh)|^2 = \frac{1}{2\pi} \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} |G(i\omega) F(i\omega)|^2 d\omega.$$

Proof. The first equality in (9.12) can be proven by following the same ideas as in Theorem 9.1 and Corollary 9.1, and the second equality follows from writing

the convolution in the first equality in the frequency domain. By noting that $f(t)$ can be expressed in the form (9.3), inserting this formula into the first integral in (9.12) leads to the convolution in the third equality. In addition, (9.13) is direct from computing the DTFT in any of the expressions in (9.12), and the last identity follows by the same reasoning as in the proof of Corollary 9.2. \square

The user can design the filter $F(i\omega)$ in an arbitrary way as long as the aforementioned conditions are satisfied and $\overline{F(i\omega)} = F(-i\omega)$ for all real ω , which ensures a real impulse response of the filter in the time domain. For instance, one might be interested in finding a filter $F(i\omega)$ that minimizes the non-causal terms. Ideally, we would desire a filter $F(i\omega)$ such that we obtain $g(kh) \approx 0$ for $k < 0$. This is equivalent to

$$\operatorname{Re} \left\{ \int_{\omega_B}^{\infty} G(i\omega) e^{-i\omega kh} d\omega \right\} \approx \operatorname{Re} \left\{ \int_{\omega_B}^{\frac{\pi}{h}} G(i\omega) F(i\omega) e^{-i\omega kh} d\omega \right\},$$

where we have exploited (9.12), the definition of $F(i\omega)$, and the causality of the impulse response $g(t)$. However, this problem is system-dependent and there seems to be no straightforward way to choose $F(i\omega)$ for this goal when the sampling period h is not very small.

Example 9.2. To illustrate the derivations above, we consider three different filters:

$$F_1(i\omega) = \operatorname{rect}\left(\frac{\omega}{2\omega_B}\right), \quad F_2(i\omega) = \operatorname{rect}\left(\frac{\omega h}{2\pi}\right), \quad F_3(i\omega) = \begin{cases} 1 & \text{if } |\omega| < \omega_B, \\ -0.3|\omega| + 1.75 & \text{if } \omega_B \leq |\omega| \leq \pi/h, \\ 0 & \text{otherwise.} \end{cases}$$

The filter $F_1(i\omega)$ is the minimum-width filter in the frequency domain that can be chosen without perturbing the input spectrum, and $F_2(i\omega)$ is the filter whose bandwidth is given by the Nyquist frequency π/h . In between these two possibilities, we have a hexagonally-shaped filter $F_3(i\omega)$. In Figure 9.3 we plot the band-limited equivalent impulse response of the continuous-time system $g(t) = -e^{-0.2t} \sin(\frac{\pi}{1.1}t)$ for the different filter reconstructions using $h = 1[\text{s}]$ and $\omega_B = 2.5[\text{rad/s}]$. The impulse response associated with $F_2(i\omega)$ has less energy in its non-causal part, at the expense of also smaller magnitudes in its causal part. It is visibly different from the other impulse responses, as the cut-off frequency of the system is not captured in this frequency band. The ad-hoc anti-aliasing filter $F_3(i\omega)$ provides a better ratio between total energy and the energy of the non-causal part of the impulse response, as detailed in Table 9.1.

Example 9.2 shows that there can be a benefit in choosing specific shapes for the filter $F(i\omega)$. In fact, we can prove that contrarily to the rectangular filter used for generating g_{BL} in (9.5), a trapezoidally-shaped filter $F(i\omega)$ allows a uniform convergence to the band-limited equivalent impulse response $g_{BL}^f(t)$ to $g(t)$, as we

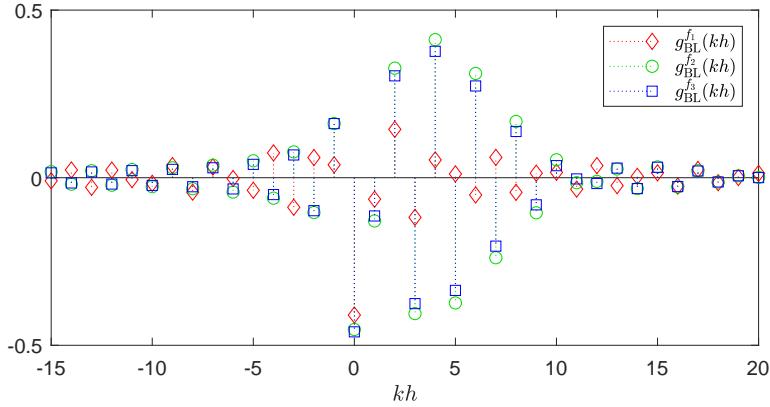


Figure 9.3: Band-limited equivalent impulse responses with anti-aliasing filters $F_1(i\omega)$ ($g_{BL}^{f_1}(kh)$), $F_2(i\omega)$ ($g_{BL}^{f_2}(kh)$) and $F_3(i\omega)$ ($g_{BL}^{f_3}(kh)$).

Table 9.1: 2-norms of the non-causal and total band-limited impulse response equivalents for different reconstruction filters, and the ratio of non-causal energy.

Filter	$\sum_{k=-\infty}^{-1} g_{BL}^f(kh) ^2$	$\sum_{k=-\infty}^{\infty} g_{BL}^f(kh) ^2$	$\frac{\sum_{k=-\infty}^{-1} g_{BL}^f(kh) ^2}{\sum_{k=-\infty}^{\infty} g_{BL}^f(kh) ^2}$
$F_1(i\omega)$	0.03102	0.2592	0.1197
$F_2(i\omega)$	0.06159	1.0677	0.0577
$F_3(i\omega)$	0.05282	0.9165	0.0576

see next. Consider the trapezoidal filter

$$F_h(i\omega) = \begin{cases} 1 & \text{if } |\omega| \leq \omega_B, \\ \frac{\frac{\pi}{h} - \omega}{\frac{\pi}{h} - \omega_B} & \text{if } \omega_B < |\omega| \leq \pi/h, \\ 0 & \text{otherwise.} \end{cases} \quad (9.14)$$

This filter has an impulse response given by

$$f_h(t) = \frac{1}{2\pi} \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} F(i\omega) e^{i\omega t} d\omega = \frac{\cos(\omega_B t) - \cos(\frac{\pi}{h}t)}{(\frac{\pi}{h} - \omega_B)\pi t^2}.$$

Note that we have made explicit the dependence on h on both the filter and its impulse response. The key aspect of the impulse response $f_h(t)$ is that it decays as $1/t^2$ instead of $1/t$, which is the decay of the standard sinc filter. Thus, the series that reconstructs the band-limited signals can converge faster if there exists oversampling. Before proving the uniform convergence of $g_{BL}^{f_h}$ to g , we need to show some properties of $f_h(t)$.

Lemma 9.1 (f_h is a “good kernel” as $h \rightarrow 0$). Consider the reconstruction filter $F_h(i\omega)$ in (9.14) and its associated impulse response $f_h(t)$. The following statements are true:

- i) $\int_{-\infty}^{\infty} f_h(t)dt = 1$ for all $h \in (0, \pi/\omega_B)$.
- ii) For all h sufficiently small, there exists $M < \infty$ such that $\int_{-\infty}^{\infty} |f_h(t)|dt \leq M$.
- iii) For every $\eta > 0$, we have $\int_{|t|>\eta} |f_h(t)|dt \rightarrow 0$ as $h \rightarrow 0$.

Proof. See Appendix 9.A.5. □

Theorem 9.2 (Uniform convergence of $g_{BL}^{f_h}$ to g). Suppose that g is a bounded, piecewise continuous, causal impulse response, and let I be any finite interval in which g is continuous. Then, $g_{BL}^{f_h}$ (i.e., (9.12) when kh is replaced by $t \in \mathbb{R}$), with reconstruction filter $F_h(i\omega)$ as in (9.14), converges uniformly to g as $h \rightarrow 0$ on the interval I .

Proof. See Appendix 9.A.6. See also Corollary 1.7 of [215] for more details. □

Remark 9.3. In fact, the proof of Theorem 9.2 reveals a stronger result: Take any $\epsilon > 0$ and let δ be any positive number (not dependent on t) such that $|g(t - \tau) - g(t)| < \epsilon\pi^2/(4\pi^2 + 12)$ when $|\tau| < \delta$. Then, $|g_{BL}^{f_h}(t) - g(t)| \leq \epsilon$ whenever

$$h < \min \left\{ \frac{\pi^2 \delta \epsilon}{\omega_B \pi \delta \epsilon + 16 \|g\|_\infty}, \frac{\pi}{2\omega_B} \right\},$$

where $\|g\|_\infty := \sup_{t \geq 0} |g(t)|$ is the infinity norm of g . Bear in mind that Theorem 9.2 can also be related to Jackson-type results [119], although we do not explore these connections here.

Theorem 9.2 suggests that the trapezoidal filter in (9.14) might be a adequate anti-aliasing filter for estimating the *continuous-time* impulse response $g(t)$ from FIR models that estimate the band-limited equivalent impulse response in discrete-time.

9.4 Conclusions

This chapter has addressed discrete-time representations of continuous-time systems with band-limited inputs. We began this study by showing that the equivalent discrete-time system for band-limited inputs is non-causal. This phenomenon, often overlooked in the system identification literature, has been exposed with theoretical support and case studies. We have pursued alternative expressions for the band-limited equivalent, and some of its properties have been proven. We also showed that it may be beneficial to exploit the degree of freedom introduced by the anti-aliasing filter when a band-limited input is oversampled, as this can lead to less significant non-causal terms and uniform convergence properties.

The non-causality of the discrete-time representations give way to novel continuous-time system identification methods with band-limited input excitation, which are covered in the next chapter.

Appendix

9.A Proofs of technical results

9.A.1 Proof of Proposition 9.1

First assume that $g_{\text{BL}}(kh) = 0$ for all integers $k < 0$. Then, for any sequence $\{f(-kh)\}_{k=1}^{\infty} \in \ell_2$, the following equalities are equivalent:

$$\begin{aligned} h \sum_{k=-\infty}^{-1} f(kh) g_{\text{BL}}(kh) = 0 &\iff \int_0^\infty g(\tau) \sum_{k=-\infty}^{-1} f(kh) \operatorname{sinc}\left(\frac{kh - \tau}{h}\right) d\tau = 0 \\ &\iff g \perp \mathcal{P}, \end{aligned}$$

where we have used Fubini's theorem to interchange summation and integral in the second step, which is valid due to the absolute convergence of the series forming $f(\tau)$ [111, Thm. 1.3]. The converse statement follows from the same steps above and later choosing $f(kh) = \delta_K([k-n]h)$ with $n < 0$, which leads to $g_{\text{BL}}(nh) = 0$. Repeating this logic with every negative integer n , we conclude that $g_{\text{BL}}(nh) = 0$ for all $n < 0$. \square

9.A.2 Proof of Corollary 9.1

The first equality in (9.8) follows from the fact that (9.5) is a convolution in the time domain, and the second equality is due to conjugate symmetry of $G(i\omega)$. For the expression defining the non-causal part of the impulse response, we see that for any integer k the continuous-time impulse response satisfies

$$g(-kh) = \frac{1}{2\pi} \int_{-\infty}^{\infty} G(i\omega) e^{-i\omega kh} d\omega.$$

Since the impulse response g is causal, we know that $g(-kh) = 0$ for any $k > 0$. Thus,

$$g_{\text{BL}}(-kh) + \frac{1}{2\pi} \int_{|\omega| > \frac{\pi}{h}} G(i\omega) e^{-i\omega kh} d\omega = g(-kh) = 0,$$

and therefore an equivalent expression for $g_{\text{BL}}(-kh)$ is

$$g_{\text{BL}}(-kh) = -\frac{1}{2\pi} \int_{|\omega| > \frac{\pi}{h}} G(i\omega) e^{-i\omega kh} d\omega.$$

Finally, for the alternative description of $\{G(i\omega)\}_-$, Lemma 2.1 of [133] leads to the following Hilbert transform relation:

$$h \sum_{k=-\infty}^{\infty} \operatorname{sgn}(kh) g_{\text{BL}}(kh) e^{-i\omega kh} = \frac{1}{2\pi i} \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} G(iv) \cot\left(\frac{[\omega - v]h}{2}\right) dv.$$

This relation yields

$$\begin{aligned} \{G(i\omega)\}_- &= \frac{h}{2} \sum_{k=-\infty}^{\infty} g_{\text{BL}}(kh) e^{-i\omega kh} - \frac{h}{2} g_{\text{BL}}(0) - \frac{h}{2} \sum_{k=-\infty}^{\infty} \operatorname{sgn}(kh) g_{\text{BL}}(kh) e^{-i\omega kh} \\ &= \frac{G(i\omega)}{2} - \frac{h}{4\pi} \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} G(iv) dv - \frac{1}{4\pi i} \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} G(iv) \cot\left(\frac{[\omega - v]h}{2}\right) dv, \end{aligned}$$

where we have replaced $g_{\text{BL}}(0)$ by its alternative representation following (9.8). \square

Remark 9.4. The result in (9.9) can also be proven directly in the frequency domain by noting that, for any $h > 0$,

$$\int_{-\infty}^{\infty} \operatorname{Re}\{G(i\omega)\} \cos(\omega kh) d\omega = - \int_{-\infty}^{\infty} \operatorname{Im}\{G(i\omega)\} \sin(\omega kh) d\omega.$$

This identity can be proven using the Kramers-Kronig relations [116]. We leave the details for the reader.

9.A.3 Proof of Corollary 9.2

Expressing the 2-norm of G in the frequency domain yields

$$\|G\|_2^2 \geq \frac{1}{2\pi} \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} |G(i\omega)|^2 d\omega = \frac{h^2}{2\pi} \sum_{k=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} g_{\text{BL}}(kh) g_{\text{BL}}(nh) \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} e^{i\omega h(n-k)} d\omega.$$

Since

$$\int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} e^{i\omega h(n-k)} d\omega = \begin{cases} \frac{2\pi}{h} & \text{if } k = n, \\ 0 & \text{if } k \neq n, \end{cases}$$

we obtain

$$\frac{1}{2\pi} \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} |G(i\omega)|^2 d\omega = h \sum_{k=-\infty}^{\infty} |g_{\text{BL}}(kh)|^2 = h \|G_{\text{BL}}\|_2^2, \quad (9.15)$$

concluding the proof. \square

9.A.4 Proof of Corollary 9.3

We compute

$$\begin{aligned}\mathcal{F}\{\tilde{g}(t)\} &= \int_{-\infty}^{\infty} \sum_{k=-\infty}^{\infty} g_{\text{BL}}(kh) \text{sinc}\left(\frac{t-kh}{h}\right) e^{-i\omega t} dt \\ &\stackrel{(t=h(k+\tau))}{=} h \sum_{k=-\infty}^{\infty} g_{\text{BL}}(kh) e^{-i\omega kh} \int_{-\infty}^{\infty} \text{sinc}(\tau) e^{-i\omega h\tau} d\tau.\end{aligned}$$

The result concerning $\mathcal{F}\{\tilde{g}(t)\}$ follows from replacing (9.7) in the expression above and the fact that $\mathcal{F}\{\text{sinc}(t)\} = \text{rect}(\omega/2\pi)$. The bound on $|\tilde{g}(t)|$ is derived by invoking the Cauchy-Schwartz inequality and exploiting the identity in (9.15):

$$|\tilde{g}(t)| = \frac{1}{2\pi} \left| \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} G(i\omega) e^{i\omega t} d\omega \right| \leq \sqrt{\frac{1}{2\pi h} \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} |G(i\omega)|^2 d\omega} = \|G_{\text{BL}}\|_2.$$

Finally, for proving (9.11), we see that

$$\begin{aligned}\int_{-\infty}^{\infty} \tilde{g}(\tau) u(kh - \tau) d\tau &= \sum_{n=-\infty}^{\infty} g_{\text{BL}}(nh) \int_{-\infty}^{\infty} u(kh - \tau) \text{sinc}\left(\frac{\tau-nh}{h}\right) d\tau \\ &= \sum_{n=-\infty}^{\infty} g_{\text{BL}}(nh) \sum_{l=-\infty}^{\infty} u(lh) \int_{-\infty}^{\infty} \text{sinc}\left(\frac{kh-\tau-lh}{h}\right) \text{sinc}\left(\frac{\tau-nh}{h}\right) d\tau \\ &= h \sum_{n=-\infty}^{\infty} g_{\text{BL}}(nh) u([k-n]h),\end{aligned}$$

where we have exploited the orthogonality of the sinc basis [107] in the last step. \square

9.A.5 Proof of Lemma 9.1

Proof of i): We directly find that

$$\int_{-\infty}^{\infty} f_h(t) dt = F_h(0) = 1.$$

\square

Proof of ii): We first separate the integral in two parts

$$\int_{-\infty}^{\infty} |f_h(t)| dt = 2 \int_0^{\frac{2\pi h}{\pi + \omega_B h}} |f_h(t)| dt + 2 \int_{\frac{2\pi h}{\pi + \omega_B h}}^{\infty} |f_h(t)| dt. \quad (9.16)$$

After some embellishments, a bound on the first integral in (9.16) can easily be computed:

$$\begin{aligned} 2 \int_0^{\frac{2\pi h}{\pi + \omega_B h}} |f_h(t)| dt &= 4 \int_0^{\frac{2\pi h}{\pi + \omega_B h}} \frac{|\sin(\frac{\pi + \omega_B h}{2h}t) \sin(\frac{\pi - \omega_B h}{2h}t)|}{(\frac{\pi}{h} - \omega_B)\pi t^2} dt \\ &= \frac{\pi + \omega_B h}{\pi h} \int_0^{\frac{2\pi h}{\pi + \omega_B h}} \text{sinc}\left(\frac{\pi + \omega_B h}{2\pi h}t\right) \text{sinc}\left(\frac{\pi - \omega_B h}{2\pi h}t\right) dt \\ &\leq 2, \end{aligned} \quad (9.17)$$

where we have used the fact that the sinc function is upper bounded by 1. On the other hand,

$$2 \int_{\frac{2\pi h}{\pi + \omega_B h}}^{\infty} |f_h(t)| dt \leq \frac{4}{(\frac{\pi}{h} - \omega_B)\pi} \int_{\frac{2\pi h}{\pi + \omega_B h}}^{\infty} \frac{1}{t^2} dt = \frac{2(\pi + \omega_B h)}{\pi^2(\pi - \omega_B h)}. \quad (9.18)$$

This last expression is upper bounded by $6/\pi^2$ when h is picked such that $h < \pi/(2\omega_B)$. Combining (9.17) and (9.18), we reach the required result with $M = 2 + 6/\pi^2$. \square

Proof of iii): Fix $\eta > 0$. Similar to (9.18), we compute the bound

$$\int_{|t|>\eta} |f_h(t)| dt \leq \frac{4}{(\frac{\pi}{h} - \omega_B)\pi} \int_{\eta}^{\infty} \frac{1}{t^2} dt = \frac{4}{(\frac{\pi}{h} - \omega_B)\pi\eta} \xrightarrow{h \rightarrow 0} 0,$$

which proves the final statement. \square

9.A.6 Proof of Theorem 9.2

Note that

$$\begin{aligned} |g_{BL}^f(t) - g(t)| &= \left| \int_{-\infty}^{\infty} g(\tau) f_h(t - \tau) d\tau - g(t) \right| \\ &= \left| \int_{-\infty}^{\infty} f_h(\tau) g(t - \tau) d\tau - g(t) \int_{-\infty}^{\infty} f_h(\tau) d\tau \right| \\ &\leq \int_{-\infty}^{\infty} |f_h(\tau)| |g(t - \tau) - g(t)| d\tau, \end{aligned}$$

where we have used Fact i) of Lemma 9.1. Since g is continuous in I , it is uniformly continuous [190, Thm. 4.19]. Therefore, given any $\epsilon > 0$, there exists a positive number δ (not dependent on t) such that $|g(t - \tau) - g(t)| < \epsilon/(2M)$ when $|\tau| < \delta$, where M is the bound given in Fact ii) of Lemma 9.1. Also, since g is bounded, we have $|g(t)| < C$ for any $t \in \mathbb{R}$. Thus, by exploiting Facts ii) and iii) of Lemma 9.1,

we find that

$$\begin{aligned} |g_{\text{BL}}^{f_h}(t) - g(t)| &\leq \int_{|\tau|>\delta} |f_h(\tau)| |g(t-\tau) - g(t)| d\tau + \int_{-\delta}^{\delta} |f_h(\tau)| |g(t-\tau) - g(t)| d\tau \\ &\leq 2C \int_{|\tau|>\delta} |f_h(\tau)| d\tau + \frac{\epsilon}{2M} \int_{-\delta}^{\delta} |f_h(\tau)| d\tau \\ &\leq \frac{8C}{(\frac{\pi}{h} - \omega_B)\pi\delta} + \frac{\epsilon}{2}. \end{aligned}$$

By picking

$$h < \min \left\{ \frac{\pi^2 \delta \epsilon}{\omega_B \pi \delta \epsilon + 16C}, \frac{\pi}{2\omega_B} \right\},$$

we have $|g_{\text{BL}}^{f_h}(t) - g(t)| \leq \epsilon/2 + \epsilon/2 = \epsilon$. Since ϵ is arbitrary and the bound on h does not depend on t , we conclude that $\max_{t \in I} |g_{\text{BL}}^{f_h}(t) - g(t)| \xrightarrow{h \rightarrow 0} 0$, i.e., $g_{\text{BL}}^{f_h}$ converges uniformly to g in I as h tends to zero. \square

Identification of continuous-time systems with band-limited inputs

As proven in Chapter 9, the interpolation property of band-limited input signals yields equivalent discrete-time system representations that are non-causal in general. This observation is exploited in this chapter to study non-parametric and parametric frequency response estimators of linear continuous-time systems. The methods and theoretical findings are tested via extensive numerical simulations.

10.1 Introduction

In both continuous and discrete-time system identification, methods should be picked according to the assumptions the user makes on the input signal. Piecewise constant assumptions are ubiquitous in the system identification literature [138, 211], although input and output signals are filtered through anti-aliasing filters in many practical situations, which produce (approximately) band-limited data [175, Ch. 13]. This cautionary remark has been noted in [177], where a general framework has been presented with the differences between the ZOH and band-limited assumptions, and the modeling errors that can be encountered if the intersample behavior of an input signal is misspecified.

Parametric continuous-time system identification under band-limited input assumptions has been carried out mostly in the frequency domain [173, 178], since this framework leads to natural descriptions of the data at hand. In [231] a general framework for reconstructing time-derivatives based on filters was presented, with applications to inputs passed through anti-aliasing filters. For the specific case of multisine inputs, a frequency-domain method based on least-squares that has been extensively used is Levy's method [132], while the extended SRIVC method for continuous-time multisine inputs derived in Chapter 6 is a promising time-domain algorithm. In terms of analysis, the bias and variance of the frequency response function under correlated input and output measurements was obtained in [174]. In [180] the probability density function of the frequency response function was

derived, together with its Gaussian approximation. On the other hand, [176] covers many methods for parametric estimation in the frequency domain with the assumption that frequency response data is measured. A review and analysis of maximum-likelihood and least-squares estimators with frequency domain data was put forward in [152], and [73] presents several approaches for direct frequency domain identification, and some approximation formulas for the frequency response function for fast and possibly irregularly sampled data.

In this chapter we take an alternative approach to the works above, and exploit the non-causal discrete-time equivalent system description in Chapter 9 to estimate the non-causal impulse response, and therefore the continuous-time frequency response, based on sampled band-limited data. Our results comprise non-parametric and parametric estimation using general band-limited inputs and a comprehensive analysis of least-squares estimators when the input is a continuous-time multisine. It should be noted that non-causal models have been previously studied for the identification of unstable systems [50], closed-loop systems [5, 53], and for applications in cross direction modeling of paper machines [147]. These contributions differ from our work as the non-causality issues they tackle do not arise due to properties of band-limited signals. On another note, our analysis for multisine inputs differs from the non-parametric tools developed in e.g., [106, 153], as we do not seek to propose methods that reduce leakage. Rather, our intention when analyzing the estimation methods with multisines is to clarify the relationships between time and frequency domain estimation, signal interpolation and intersample behavior, and the maximum likelihood estimate using sampled time-domain data for continuous-time system identification.

Part of the contributions found in this chapter can be found in [88]. In summary,

- we show that a non-causal least-squares method delivers the true band-limited equivalent impulse response. Later, non-causal regularization methods are applied for estimating the non-causal impulse response that represents the continuous-time system. Advantages with respect to causal least-squares estimation are confirmed via extensive simulations;
- we use the non-causal, non-parametric estimators developed in Section 10.3 as a starting point for designing parametric estimators of continuous-time frequency responses. Based on an indirect PEM philosophy [212], we derive an estimator that can be computed via the refined instrumental variables approach in the frequency domain. Numerical simulations confirm the applicability of this method and its advantages in terms of computational efficiency compared to other approaches; and
- we analyze the statistical optimality of the least-squares estimator with a continuous-time multisine input excitation, which is a singular case of a band-limited signal due to its discrete spectrum. The statistical properties of the least-squares estimator are posed in terms of the frequency response function, and an optimization step is proposed to convert the non-parametric

estimate into a continuous-time transfer function estimate that is shown to be asymptotically optimal. Here, a link between maximum likelihood, the indirect prediction error method, and optimal refined instrumental variables is derived. These relationships are later used to analyze the finite-time behavior of the maximum likelihood estimator. Similar to the previous contributions, extensive numerical simulations are performed to verify the theoretical findings.

The rest of this chapter is organized as follows. Section 10.2 states the identification problem we study. A non-parametric estimator using non-causal least-squares and its regularized variant are developed in Section 10.3. Section 10.4 discusses several parametric estimators that use a non-parametric estimate as a first step, and Section 10.5 focuses on identification using continuous-time multisines. Conclusions are drawn in Section 10.6. The appendix of this chapter contains the proofs of the theoretical results and some supplementary material.

10.2 Problem formulation

Consider the asymptotically stable, single-input single-output, causal, linear and time-invariant system

$$x(t) = \int_0^\infty g^*(\tau)u(t-\tau)d\tau, \quad (10.1)$$

where $u(t)$ is a band-limited input signal with bandwidth ω_B . The frequency response of the system, given by the Fourier transform of the impulse response $g^*(t)$, is denoted by $G^*(i\omega)$. Assume that we retrieve noisy measurements of the output of the form

$$y(kh) = x(kh) + v(kh), \quad k = 1, 2, \dots, N, \quad (10.2)$$

where h is the sampling period and is less than or equal to π/ω_B , and $\{v(kh)\}_{k=1}^N$ is a zero-mean stochastic process of variance σ^2 that is independent of the sampled input. The question we address is how to estimate the continuous-time frequency response $G^*(i\omega)$ (or equivalently, the band-limited equivalent impulse response $g_{BL}^*(kh)$) from sampled values of the input and noisy output.

Remark 10.1. It is well known that band-limited signals must extend infinitely in time [127]. Equivalently, a time-limited signal is not band-limited. Thus, in practice we will encounter *approximately* band-limited signals, which are commonly obtained via anti-aliasing filters [175].

The following three sections cover different topics of this identification problem, namely, non-parametric estimation with non-causal regularized least-squares, parametric estimation, and identification with continuous-time multisines as inputs.

10.3 Non-parametric frequency response estimation: Non-causal regularized least-squares

Our first contribution concerns the non-parametric estimation of the frequency response of the continuous-time system based on non-causal regularized least-squares. Before deriving this estimator, we first take a look at a least-squares estimator for the estimation of the non-causal equivalent band-limited impulse response.

10.3.1 Non-causal least-squares

One of the consequences of Theorem 9.1 is that sampled outputs of systems driven by band-limited inputs can be expressed exactly by a doubly-infinite convolution sum. Therefore, the measured output $y(kh)$ in (10.2) can be written as

$$y(kh) = h \sum_{l=-M_{nc}}^{M_c} u([k-l]h) g_{BL}^*(lh) + w(kh), \quad (10.3)$$

where the integers M_{nc} and M_c satisfy $M_{nc} + M_c \geq 0$. These integers respectively indicate the number of non-causal and strictly causal terms of the impulse response that will be estimated. Additionally, the signal $\{w(kh)\}_{k=1}^N$ is a residual term accounting for the noise sequence $\{v(kh)\}_{k=1}^N$, the approximation error of the series in (9.4) evaluated at $t = kh$, and possible transient effects. The equations that the output data satisfy can be put in matrix form as $\mathbf{y} = \Phi \boldsymbol{\rho}^* + \mathbf{w}$, where

$$\mathbf{y} = \begin{bmatrix} y(h), y(2h), \dots, y(Nh) \end{bmatrix}^\top, \\ \Phi = h \begin{bmatrix} u([1+M_{nc}]h) & u(M_{nc}h) & \dots & u([1-M_c]h) \\ u([2+M_{nc}]h) & u([1+M_{nc}]h) & \dots & u([2-M_c]h) \\ \vdots & \vdots & & \vdots \\ u([N+M_{nc}]h) & u([N-1+M_{nc}]h) & \dots & u([N-M_c]h) \end{bmatrix}, \quad (10.4)$$

$$\boldsymbol{\rho}^* = \begin{bmatrix} g_{BL}^*(-M_{nc}h), g_{BL}^*([1-M_{nc}]h), \dots, g_{BL}^*(M_ch) \end{bmatrix}^\top, \quad (10.5)$$

$$\mathbf{w} = \begin{bmatrix} w(h), w(2h), \dots, w(Nh) \end{bmatrix}^\top.$$

Since the goal is to provide an estimate for $G^*(i\omega)$, we shall focus on estimating the vector of coefficients of its truncated band-limited equivalent impulse response, $\boldsymbol{\rho}^*$. To this end we consider the least-squares estimate of $\boldsymbol{\rho}^*$, that is given by

$$\hat{\boldsymbol{\rho}}_N = (\Phi^\top \Phi)^{-1} \Phi^\top \mathbf{y} \\ = \left[\sum_{k=1}^N \boldsymbol{\varphi}(kh) \boldsymbol{\varphi}^\top(kh) \right]^{-1} \left[\sum_{k=1}^N \boldsymbol{\varphi}(kh) y(kh) \right], \quad (10.6)$$

where the regressor vector is given by the columns of Φ^\top , i.e.,

$$\varphi(kh) = h \left[u([k + M_{nc}]h), u([k - 1 + M_{nc}]h), \dots, u([k - M_c]h) \right]^\top.$$

The impulse response estimate in (10.6) generates a non-parametric frequency response estimate according to Theorem 9.1 of the form

$$\hat{G}_N(i\omega) = \hat{\rho}_N^\top \Gamma(e^{i\omega h}), \quad (10.7)$$

where $\Gamma(e^{i\omega h})$ is given by

$$\Gamma(e^{i\omega h}) = h \left[e^{i\omega M_{nc}h}, e^{i\omega(M_{nc}-1)h}, \dots, e^{-i\omega M_ch} \right]^\top. \quad (10.8)$$

Note that, by construction, the proposed estimate satisfies the conjugacy property

$$\overline{\hat{G}_N(i\omega)} = \hat{\rho}_N^\top \overline{\Gamma(e^{i\omega h})} = \hat{\rho}_N^\top \Gamma(e^{-i\omega h}) = \hat{G}_N(-i\omega).$$

In the sequel we study the properties of this estimator when the input is constructed by a discrete-time white noise sequence that is interpolated with sinc functions, and propose a non-causal regularized least-squares method that can deliver more accurate models in terms of goodness of fit. These novel estimators are later tested via extensive numerical simulations.

10.3.2 Asymptotic properties of least-squares for band-limited white noise inputs

For practical purposes, we shall consider that the non-causal samples of the input are all equal to zero (i.e., $u(kh) = 0$ for $k < 0$). Bear in mind that this does not mean that the continuous-time input to the system is causal, as every band-limited signal must extend infinitely in both directions in time [217, Ch. 2.8].

The next result concerns the consistency of the least-squares estimator in (10.6) when the input is discrete-time white noise interpolated through sinc functions of the form

$$u(t) = \sum_{n=1}^N e(nh) \operatorname{sinc}\left(\frac{t-nh}{h}\right), \quad (10.9)$$

where $\{e(nh)\}_{n=1}^N$ is white noise of finite variance λ^2 . Such signal has an autocorrelation function given by

$$R_u(t_1, t_2) = \mathbb{E}\{u(t_1)u(t_2)\} = \lambda^2 \sum_{n=1}^N \operatorname{sinc}\left(\frac{t_1-nh}{h}\right) \operatorname{sinc}\left(\frac{t_2-nh}{h}\right). \quad (10.10)$$

Theorem 10.1. Consider the system (10.1) with measured output (10.2) and input (10.9), where $\{e(nh)\}_{n=1}^N$ is white noise of finite variance. Then, for any integers M_{nc}, M_c such that $M_{nc} + M_c \geq 0$, we have $\hat{\rho}_N \xrightarrow{a.s.} \rho^*$, where $\hat{\rho}_N$ and ρ^* are defined in (10.6) and (10.5), respectively.

Proof. See Appendix 10.A.1. \square

Remark 10.2. Theorem 10.1 shows that when causal FIR models are being fit to data with a band-limited white noise input, the coefficients that will be estimated converge to the ones provided by the band-limited equivalent, and not the zero-order hold equivalent, which is commonly assumed when discrete-time data is obtained. This fact has implications on the accuracy of the model, as the band-limited equivalent has non-causal coefficients that are in general different from zero but are usually left unmodeled. These non-causal terms may only be neglected if the sampling period is sufficiently small.

The next result concerns the asymptotic distribution of the least-squares estimate.

Theorem 10.2. Consider the system (10.1) with measured output (10.2), where the input is given by (10.9), and $\{e(nh)\}_{n=1}^N$ is white noise of variance λ^2 that is independent of the output noise sequence $\{v(kh)\}_{k=1}^N$. The least-squares estimate $\hat{\rho}_N$ in (10.6) is asymptotically Gaussian-distributed, i.e.,

$$\sqrt{N}(\hat{\rho}_N - \rho^*) \xrightarrow{\text{dist.}} \mathcal{N}(\mathbf{0}, \Sigma),$$

where the asymptotic covariance is given by

$$\Sigma := \frac{1}{h^4 \lambda^4} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N \sum_{n=1}^N \mathbb{E}\{\varphi(kh)w(kh)w(nh)\varphi^\top(nh)\}. \quad (10.11)$$

Proof. See Appendix 10.A.2. \square

Remark 10.3. Note that if the tails of the band-limited impulse response are neglected, the asymptotic covariance in (10.11) can be well approximated by

$$\Sigma \approx \frac{\sigma^2}{h^4 \lambda^4} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N \mathbb{E}\{\varphi(kh)\varphi^\top(kh)\} = \frac{\sigma^2}{h^2 \lambda^2} \mathbf{I}_{M_{nc}+M_c+1},$$

where \mathbf{I}_n denotes the identity matrix of size n . This implies that the estimated impulse coefficients are approximately uncorrelated if the sampled input sequence is white.

10.3.3 Non-causal regularized least-squares for continuous-time system identification with band-limited inputs

The least-squares estimator proposed in (10.6) provides consistent estimates of the band-limited equivalent impulse response as the sample size tends to infinity. However, usually the practitioner is interested in the finite-time behavior of estimators, where the number of parameters to be estimated can be of the order of the number of samples. Another situation that may occur is that the number of parameters is

larger than the persistence of excitation order of the input signal. In both of these cases, it is convenient to use regularized least-squares estimators [172].

The regularized least-squares estimate of ρ^* , denoted here by $\hat{\rho}_N^r$, is given by

$$\begin{aligned}\hat{\rho}_N^r &= \arg \min_{\rho} \|\mathbf{y} - \Phi \rho\|_2^2 + \gamma \rho^\top \mathbf{P}_r^{-1} \rho \\ &= (\mathbf{P}_r \Phi^\top \Phi + \gamma \mathbf{I}_{M_{nc}+M_c+1})^{-1} \mathbf{P}_r \Phi^\top \mathbf{y},\end{aligned}\quad (10.12)$$

where $\mathbf{P}_r \succeq 0$ is a regularization matrix and γ is a positive scalar. The problem of choosing the best regularization matrix for causal FIR models has been thoroughly studied during the past years [37, 145, 172]; this problem is challenging since it is known that the optimal regularization matrix depends on the true system. More precisely, for causal FIR models the regularization matrix that minimizes the MSE in a positive definite sense is given by $\gamma^{\text{opt}} = \sigma^2$ and $\mathbf{P}_r^{\text{opt}} = \rho^* \rho^{*\top}$. An analogous result holds for non-causal FIR models, as stated in Proposition 10.1.

Proposition 10.1. Consider the system with sampled output as in (10.3), and assume that the signal $w(kh)$ is white noise of variance σ^2 . The regularization term that minimizes the MSE matrix in a positive definite sense is given by $\gamma^{\text{opt}} = \sigma^2$ and $\mathbf{P}_r^{\text{opt}} = \rho^* \rho^{*\top}$, and the corresponding optimal regularized estimate is

$$\hat{\rho}_N^{r,\text{opt}} = (\rho^* \rho^{*\top} \Phi^\top \Phi + \sigma^2 \mathbf{I}_{M_{nc}+M_c+1})^{-1} \rho^* \rho^{*\top} \Phi^\top \mathbf{y}. \quad (10.13)$$

Proof. The result follows by the same reasoning as in the proof of Theorem 1 of [37]. \square

Since the optimal regularization matrix is not known a priori, the matrix \mathbf{P}_r is typically parameterized by a low-dimensional hyperparameter vector β (defined in a region \mathcal{B} of the Euclidean space of appropriate dimension) according to what can be assumed about the impulse response. The difference between the estimation problem in this work and causal FIR estimation is that the impulse response has a causal and non-causal decay, which induces changes in the way kernels are designed. Recently [18], the kernel-design problem for general non-causal systems was studied with the goal of identifying inverse systems for feedforward control. Here we recall the findings in [18] and apply them to our context.

Let

$$b_k := \begin{cases} \lambda_{nc}^{-2k} & \text{if } k < 0 \\ \lambda_c^{2k} & \text{if } k \geq 0, \end{cases}$$

with $\lambda_{nc}, \lambda_c \in [0, 1]$. The non-causal tuned/correlated (TC) kernel and the second-order stable spline (SS) kernel yield the following regularization matrices:

TC kernel: $\mathbf{P}_{r,(j,l)}(\beta) = \alpha \min(b_{j-M_{nc}-1}, b_{l-M_{nc}-1})$;

$$\beta = [\lambda_{nc}, \lambda_c, \alpha]^\top; \quad \mathcal{B} = \{\beta \in \mathbb{R}^3 : \alpha > 0, 0 \leq \lambda_{nc}, \lambda_c < 1\}.$$

$$\begin{aligned} \text{SS kernel: } \mathbf{P}_{r,(j,l)}(\boldsymbol{\beta}) &= \alpha \frac{\min(b_{j-M_{nc}-1}, b_{l-M_{nc}-1})^2}{6} \\ &\quad \times [3 \max(b_{j-M_{nc}-1}, b_{l-M_{nc}-1}) - \min(b_{j-M_{nc}-1}, b_{l-M_{nc}-1})]; \\ \boldsymbol{\beta} &= [\lambda_{nc}, \lambda_c, \alpha]^\top; \quad \mathcal{B} = \{\boldsymbol{\beta} \in \mathbb{R}^3 : \alpha > 0, 0 \leq \lambda_{nc}, \lambda_c < 1\}. \end{aligned}$$

Remark 10.4. Note that for any real numbers a and b , the identities $2 \min(a, b) = a + b - |a - b|$ and $2 \max(a, b) = a + b + |a - b|$ are true. Using these equalities it can be shown that the kernels above are equivalent to the standard TC and SS causal kernels if λ_{nc} is set to zero [18].

To compute the regularized estimator in (10.12), all that is left to know is how to tune the hyperparameters in $\boldsymbol{\beta}$. This can be done using marginal likelihood optimization with respect to the data, as in the causal impulse response estimation case¹ [172]. The procedure is detailed as follows. If we fix $\gamma = \sigma^2$ and the parameter vector (denoted here as $\boldsymbol{\rho}$) is assumed to be distributed as a zero-mean Gaussian $\mathcal{N}(\mathbf{0}, \mathbf{P}_r(\boldsymbol{\beta}))$ independent of the additive noise $\{v(kh)\}_{k=1}^N$, then the random variables \mathbf{y} and $\boldsymbol{\rho}$ will be jointly Gaussian:

$$\begin{bmatrix} \boldsymbol{\rho} \\ \mathbf{y} \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{P}_r(\boldsymbol{\beta}) & \mathbf{P}_r(\boldsymbol{\beta})\Phi^\top \\ \Phi\mathbf{P}_r(\boldsymbol{\beta}) & \Phi\mathbf{P}_r(\boldsymbol{\beta})\Phi^\top + \sigma^2\mathbf{I}_N \end{bmatrix} \right).$$

This means that the conditional distribution of the data given $\boldsymbol{\beta}$ can be expressed as

$$p(\mathbf{y}|\boldsymbol{\beta}) = \frac{1}{\sqrt{\det(2\pi\mathbf{Z}(\boldsymbol{\beta}))}} \exp \left(-\frac{1}{2} \mathbf{y}^\top [\mathbf{Z}(\boldsymbol{\beta})]^{-1} \mathbf{y} \right),$$

where $\mathbf{Z}(\boldsymbol{\beta}) := \Phi\mathbf{P}_r(\boldsymbol{\beta})\Phi^\top + \sigma^2\mathbf{I}_N$. Thus, the hyperparameter vector $\boldsymbol{\beta}$ can be obtained by maximizing the marginalized log-likelihood

$$\begin{aligned} \hat{\boldsymbol{\beta}}_{\text{ML}} &= \arg \max_{\boldsymbol{\beta} \in \mathcal{B}} \log(p(\mathbf{y}|\boldsymbol{\beta})) \\ &= \arg \min_{\boldsymbol{\beta} \in \mathcal{B}} \mathbf{y}^\top [\mathbf{Z}(\boldsymbol{\beta})]^{-1} \mathbf{y} + \log \det(\mathbf{Z}(\boldsymbol{\beta})). \end{aligned}$$

As reported in Remark 5 of [172], the variance of the additive noise σ^2 can be included as a hyperparameter in $\boldsymbol{\beta}$ or it can be estimated separately by computing the sample variance that results from ARX or FIR modeling with least-squares. For completeness, we provide the computational considerations that must be taken into account for including σ^2 as a hyperparameter, most of which are included in [36].

First, we must factor the regularization matrix as $\mathbf{P}_r(\boldsymbol{\beta})/\sigma^2 = \mathbf{L}(\boldsymbol{\beta})\mathbf{L}^\top(\boldsymbol{\beta})$. Afterwards, we consider the thin QR factorization [115, Thm. 2.1.14]

$$\begin{bmatrix} \Phi\mathbf{L}(\boldsymbol{\beta}) & \mathbf{y} \\ \mathbf{I}_{M_{nc}+M_c+1} & \mathbf{0} \end{bmatrix} = \mathbf{Q}(\boldsymbol{\beta}) \begin{bmatrix} \mathbf{R}_1(\boldsymbol{\beta}) & \mathbf{R}_2(\boldsymbol{\beta}) \\ \mathbf{0} & r(\boldsymbol{\beta}) \end{bmatrix}, \quad (10.14)$$

¹An alternative, albeit more computationally demanding strategy is to proceed via a fully Bayesian philosophy as in [80, 183].

where $\mathbf{Q}(\beta)$ is a rectangular orthogonal matrix, $\mathbf{R}_1(\beta) \in \mathbb{R}^{(M_{nc}+M_c+1) \times (M_{nc}+M_c+1)}$ is an upper triangular matrix, and $r(\beta)$ is a positive scalar. The diagonal of $\mathbf{R}_1(\beta)$ is assumed to be positive without loss of generality. Note that the following identities are satisfied:

$$\mathbf{R}_1^\top(\beta)\mathbf{R}_1(\beta) = \mathbf{L}^\top(\beta)\Phi^\top\Phi\mathbf{L}(\beta) + \mathbf{I}_{M_{nc}+M_c+1}, \quad (10.15a)$$

$$\mathbf{R}_1^\top(\beta)\mathbf{R}_2(\beta) = \mathbf{L}^\top(\beta)\Phi^\top\mathbf{y}, \quad (10.15b)$$

$$\mathbf{R}_2^\top(\beta)\mathbf{R}_2(\beta) + r(\beta)^2 = \mathbf{y}^\top\mathbf{y}.$$

After some technical manipulations, the equalities above lead to expressing the log-likelihood cost as

$$\mathbf{y}^\top[\mathbf{Z}(\beta)]^{-1}\mathbf{y} + \log \det(\mathbf{Z}(\beta)) = \frac{r^2(\beta)}{\sigma^2} + N \log(\sigma^2) + 2 \log \det(\mathbf{R}_1(\beta)). \quad (10.16)$$

Since the TC and SS regularization matrices are already factored by a scalar constant α , the dependence on σ in \mathbf{L} (and therefore in \mathbf{R}_1) is redundant for the optimization of the marginal likelihood with respect to β and σ^2 . Therefore, we can concentrate the cost function by minimizing (10.16) over σ^2 , which leads to

$$\hat{\beta}_{ML} = \arg \min_{\beta \in \mathcal{B}} N \log(r(\beta)) + \log \det(\mathbf{R}_1(\beta)). \quad (10.17)$$

Finally, thanks to (10.15a) and (10.15b), the regularized least-squares estimate in (10.12) can be computed by

$$\begin{aligned} \hat{\rho}_N^r &= \mathbf{L}(\hat{\beta}_{ML}) [\mathbf{L}^\top(\hat{\beta}_{ML})\Phi^\top\Phi\mathbf{L}(\hat{\beta}_{ML}) + \mathbf{I}_{M_{nc}+M_c+1}]^{-1}\Phi^\top\mathbf{y} \\ &= \mathbf{L}(\hat{\beta}_{ML}) [\mathbf{R}_1(\hat{\beta}_{ML})]^{-1}\mathbf{R}_2(\hat{\beta}_{ML}). \end{aligned} \quad (10.18)$$

To conclude this subsection, we provide a pseudo-code for computing the regularized least-squares estimator for continuous-time systems under band-limited inputs in Algorithm 10.1.

10.3.4 Simulations

We now present simulation tests that confirm the applicability of the proposed methods. Two different tests are performed: first on two second-order systems studied previously in this work, and later on random systems.

Two examples

We consider the following two second-order systems:

$$G_1^*(p) = \frac{1.25}{0.25p^2 + 0.7p + 1}, \quad G_2^*(p) = \frac{-\pi/1.1}{(p + 0.2)^2 + \pi^2/1.1^2}. \quad (10.19)$$

Algorithm 10.1: Non-causal regularized least-squares for continuous-time system identification with band-limited inputs

- 1: Input: $\{u(kh), y(kh)\}_{k=1}^N$, model order (N_{nc}, N_c) , kernel (TC or SS)
 - 2: Form Φ as in (10.4)
 - 3: Factor $\mathbf{P}(\beta) = \mathbf{L}(\beta)\mathbf{L}^\top(\beta)$
 - 4: Perform the thin QR factorization in (10.14)
 - 5: Compute
$$\hat{\beta}_{\text{ML}} = \arg \min_{\beta \in \mathcal{B}} N \log(r(\beta)) + \log \det(\mathbf{R}_1(\beta))$$
 - 6: Output: $\hat{\rho}_N^r = \mathbf{L}(\hat{\beta}_{\text{ML}})[\mathbf{R}_1(\hat{\beta}_{\text{ML}})]^{-1}\mathbf{R}_2(\hat{\beta}_{\text{ML}})$ and $\hat{G}_N^r(i\omega) = \mathbf{\Gamma}^\top(e^{i\omega h})\hat{\rho}_N^r$, where $\mathbf{\Gamma}(e^{i\omega h})$ is defined in (10.8).
-

These systems have been used for generating the band-limited equivalent impulse responses in Figures 9.1 and 9.2 of Chapter 9, respectively. The sampling periods are $h_1 = 0.3[\text{s}]$ and $h_2 = 1[\text{s}]$ respectively, and the inputs to both systems are given by (10.9), where $e(kh)$ is white noise of unit variance. The noiseless continuous-time output is simulated by sampling the input 100 times faster and assuming an FOH intersample behavior. White noise is added to the samples of the simulated output, with variance corresponding to an amplitude signal-to-noise ratio of approximately five.

One hundred causal samples are obtained from each system. We compute the noiseless output $x(t)$ starting from $t = -Nh$ to capture the effect of the non-causal part of the input, and only the causal sequence $\{x(kh)\}_{k=1}^N$ is contaminated with noise and used for identification. Three causal and four non-causal estimators are tested, all with 40 parameters each; the non-causal ones use $(M_{nc}, M_c) = (15, 24)$. The following estimators are considered:

- a) Causal least-squares (C-LS);
- b) Non-causal least-squares (NC-LS);
- c) Causal TC-regularized least-squares (C-TC);
- d) Non-causal TC-regularized least-squares (NC-TC);
- e) Causal SS-regularized least-squares (C-SS);
- f) Non-causal SS-regularized least-squares (NC-SS);
- g) Optimal non-causal regularized least-squares (Oracle).

The causal TC and SS-regularized least-squares estimators are obtained via the `impulseest` command in MATLAB, while the non-causal TC and SS kernels are tuned by solving (10.17) and evaluating (10.18). The unrealizable oracle estimator is computed by (10.13). The performance of these estimators is compared via Monte

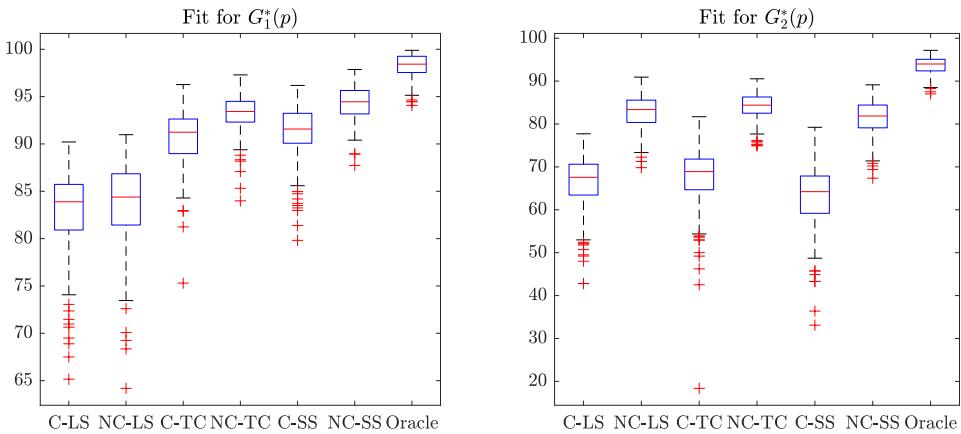


Figure 10.1: Fit box plots of seven different impulse response estimators. Left: $G_1^*(p)$; right: $G_2^*(p)$. These systems are described in (10.19).

Carlo simulations, in which 300 different noise realizations are considered. Validation data with a different input realization and noiseless output are generated to compute the fit metric

$$\text{Fit} = 100 \left(1 - \frac{\|\mathbf{y} - \Phi\hat{\rho}_N\|_2}{\|\mathbf{y} - \bar{\mathbf{y}}\|_2} \right),$$

where $\bar{\mathbf{y}}$ is the sample mean of the validation output data \mathbf{y} (expressed as a constant vector).

The box plots of the fit of each method for $G_1^*(p)$ and $G_2^*(p)$ are shown in Figure 10.1. For both systems there is an advantage in considering non-causal parameters versus fixing them to zero. Only a modest improvement can be observed for $G_1^*(p)$ by estimating the non-causal terms in the least-squares estimate, as the sampling frequency is large compared to the bandwidth of the system and thus the non-causal component of the impulse response is not significant. On the other hand, the gain in performance in the test with $G_2^*(p)$ is substantial: this is explained by the fact that an important part of the band-limited equivalent impulse response is non-causal. As expected for small sample sizes, an increase in performance is observed on both systems if regularization is included. It must be noted that including regularization in the causal least-squares estimate improves the fit of the causal model but will anyway disregard the significant non-causal components, leading to a worse performance compared to the non-causal regularized least-squares estimators.

Concerning the unbiasedness of the least-squares method as presented in Theorem 10.1, the mean value of each impulse response estimate is shown in Figure 10.2. From this figure we observe that the proposed estimator captures the non-causal part accurately, leading to better goodness of fit compared to the causal model according to Figure 10.1.

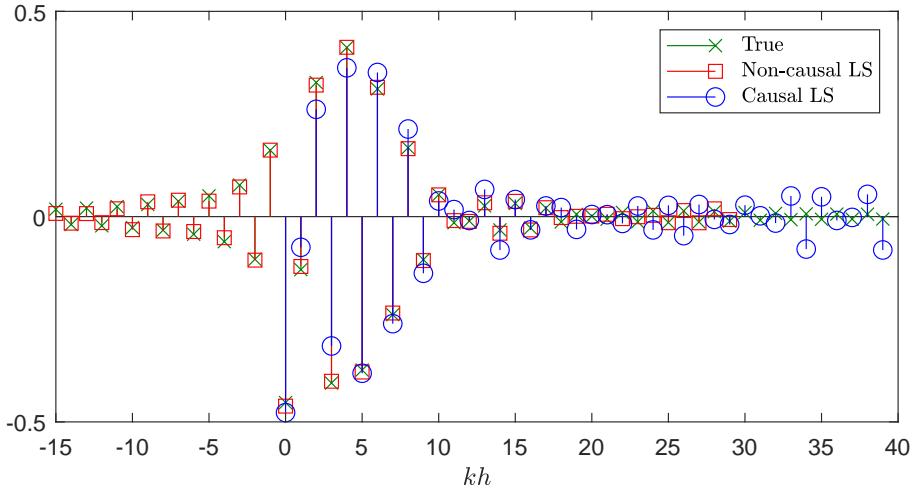


Figure 10.2: In green, the true band-limited equivalent impulse response of $G_2^*(p)$. Red squares indicate the sample mean of the non-causal least-squares estimate, and the sample mean for the causal least-squares estimate is shown in blue circles.

Random systems

We also test the proposed methods on a set of random systems. Similar to the bank of test systems presented in [37], a number of continuous-time systems of order 30 are generated using the `rss` command in MATLAB, and the systems are sampled at a frequency $f = 1/h$ equal to three times their bandwidth. The systems are split into 300 “fast” systems whose poles have real parts not greater than $\log(0.95)/h$, and 300 “slow” systems which have at least one pole with real part greater than $\log(0.95)/h$. The systems are excited with the same input as in Section 10.3.4, and $N = 500$ samples are obtained for identification. Gaussian white noise contaminates the output measurements, first with an amplitude SNR of 10, and later with an SNR of 20, resulting in 1200 different simulation runs.

All estimators tested previously (except the oracle) are assessed in these new scenarios. The fit metric for the fast and slow systems with an SNR equal to 10 is presented in the upper plots of Figure 10.3, while the lower plots show the same but for an SNR equal to 20. In all cases the non-causal estimators are superior in terms of median fit, which provides strong evidence for the adequacy of including non-causal terms in a general continuous-time system identification setup with band-limited input excitation. Note that the “fast” systems exhibit on average a greater improvement if non-causal terms are included. This can be explained by the fact that the sampling period is relatively large compared to the dominant time constant for this case, which can induce greater non-causal values for the band-limited equivalent impulse response as studied in Chapter 9.

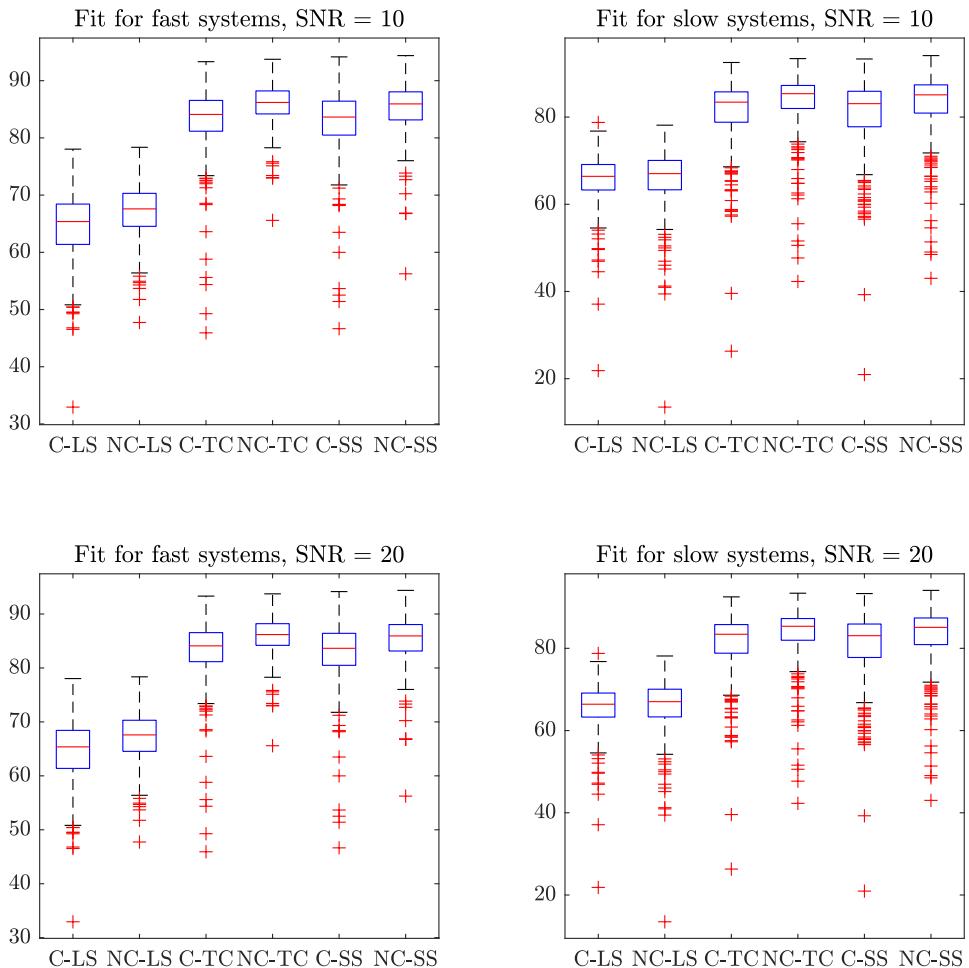


Figure 10.3: Fit box plots of different impulse response estimators for SNR equal to 10 (upper plots), and for SNR equal to 20 (lower plots): causal least-squares (C-LS), non-causal least-squares (NC-LS), causal TC-regularized LS (C-TC), non-causal TC-regularized LS (NC-TC), causal SS-regularized LS (C-SS), and non-causal SS-regularized LS (NC-SS). Left: “fast” random systems; right: “slow” random systems.

10.4 From non-parametric to parametric estimators

Theorem 9.1 also provides a direct relationship between the continuous-time frequency response and the band-limited impulse response $\{g_{BL}(kh)\}_{k \in \mathbb{Z}}$. In this section, we exploit this theorem to formulate parametric estimators for the continuous-time system based on band-limited inputs and sampled output data. We assume that the parametric model $G(i\omega, \boldsymbol{\theta})$ we seek is a frequency response function of the form

$$G(i\omega, \boldsymbol{\theta}) = \frac{B(i\omega, \boldsymbol{\theta})}{A(i\omega, \boldsymbol{\theta})} = \frac{b_m(i\omega)^m + \dots + b_1(i\omega) + b_0}{a_n(i\omega)^n + \dots + a_1(i\omega) + 1},$$

where $n \geq m$, and where the parameter vector is $\boldsymbol{\theta} = [a_1, \dots, a_n, b_0, \dots, b_m]^\top$.

10.4.1 Parametric estimator using an indirect PEM philosophy

Following the indirect approach ideas of Chapter 3, one might be interested in first obtaining a discrete-time rational model and then transforming it to continuous-time. One difficulty arising from this approach for this context is that such transfer functions do not relate to continuous-time rational models from the transformation in Theorem 9.1 (recall Example 9.1, where it was observed that even an elementary continuous-time transfer function cannot be characterized by a band-limited equivalent discrete-time rational model in z). In other words, if a straightforward indirect approach method is applied for band-limited inputs, the estimates will be biased and no finite order rational model structure will contain the true discretized system.

The key idea behind our approach consists in letting the non-parametric estimate of the band-limited equivalent impulse response $\hat{\rho}_N$ (or its regularized variant, $\hat{\rho}_N^r$) act as a proxy for a sufficient statistic for $\boldsymbol{\theta}$. Equivalently, this approach assumes that $\hat{\rho}_N$ estimates a model with greater flexibility than $G(i\omega, \boldsymbol{\theta})$, and thus we can use ideas from the indirect PEM approach [212] to compute $\hat{\boldsymbol{\theta}}$, an estimate for $\boldsymbol{\theta}$.

For now we will first consider the non-causal least-squares estimator $\hat{\rho}_N$, and assume that the sampled input is quasi-stationary and the measurement noise is white with variance σ^2 . Based on what we mention above, we are interested in obtaining the minimizer of the cost function

$$V(\boldsymbol{\theta}) := \frac{1}{N} (\boldsymbol{\rho}(\boldsymbol{\theta}) - \hat{\boldsymbol{\rho}}_N)^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{\rho}(\boldsymbol{\theta}) - \hat{\boldsymbol{\rho}}_N), \quad (10.20)$$

where $\boldsymbol{\Sigma}$ is the covariance matrix of $\hat{\boldsymbol{\rho}}_N$, and $\boldsymbol{\rho}(\boldsymbol{\theta})$ is the band-limited equivalent impulse response of $G(p, \boldsymbol{\theta})$ of the form (10.5). For the sake of simplicity, we denote $\tilde{g}_{BL}(kh, \boldsymbol{\theta}), k = -M_{nc}, \dots, M_c$ as the elements of the vector $\boldsymbol{\rho}(\boldsymbol{\theta}) - \hat{\boldsymbol{\rho}}_N$. Minimizing the cost function in (10.20) is not direct, and therefore we will try to obtain simpler expressions for it. First, for large M_{nc}, M_c and N , the derivations that led to Theorem 10.2 now lead to

$$\boldsymbol{\Sigma} \approx \frac{\sigma^2}{N} \bar{\mathbb{E}} \{ \boldsymbol{\varphi}(kh) \boldsymbol{\varphi}^\top(kh) \}^{-1}. \quad (10.21)$$

Next, define the covariance function of the sampled input $r_u(kh) := \overline{\mathbb{E}}\{u([l+k]h)u(lh)\}$ and its associated spectrum

$$\phi_u(e^{i\omega h}) = h \sum_{k=-\infty}^{\infty} r_u(kh) e^{-i\omega kh} \iff r_u(kh) = \frac{1}{2\pi} \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} \phi_u(e^{i\omega h}) e^{i\omega kh} d\omega. \quad (10.22)$$

Note that the covariance function is different in nature to (10.10) as we are solely considering the input samples instead of viewing it as a continuous-time signal. By leveraging the fact that Σ^{-1} has elements $\Sigma_{k,l}^{-1} = Nh^2 r_u([k-l]h)/\sigma^2$, we find that for large M_{nc}, M_c and N , the cost in (10.20) can be well approximated by

$$\begin{aligned} V(\boldsymbol{\theta}) &\approx \frac{h^2}{\sigma^2} \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \tilde{g}_{BL}(kh, \boldsymbol{\theta}) \tilde{g}_{BL}(lh, \boldsymbol{\theta}) r_u([k-l]h) \\ &= \frac{h^2}{2\pi\sigma^2} \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \tilde{g}_{BL}(kh, \boldsymbol{\theta}) \tilde{g}_{BL}(lh, \boldsymbol{\theta}) \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} \phi_u(e^{i\omega h}) e^{i\omega(k-l)h} d\omega \\ &= \frac{1}{2\pi\sigma^2} \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} \left(h \sum_{k=-\infty}^{\infty} \tilde{g}_{BL}(kh, \boldsymbol{\theta}) e^{i\omega kh} \right) \left(h \sum_{l=-\infty}^{\infty} \tilde{g}_{BL}(lh, \boldsymbol{\theta}) e^{-i\omega lh} \right) \phi_u(e^{i\omega h}) d\omega \\ &= \frac{1}{2\pi\sigma^2} \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} |\tilde{G}(i\omega, \boldsymbol{\theta})|^2 \phi_u(e^{i\omega h}) d\omega, \end{aligned} \quad (10.23)$$

where $\tilde{G}(i\omega, \boldsymbol{\theta})$ is given by $\hat{G}_N(i\omega) - G(i\omega, \boldsymbol{\theta})$, with $\hat{G}_N(i\omega)$ being the frequency response estimate corresponding to the band-limited impulse response $\hat{\rho}_N$. Note that this derivation is equivalent to the one in the proof of Corollary 9.2 when Σ is a scalar matrix. With (10.23) in mind, we come to the conclusion that the following estimator provides a good approximation of the minimizer of (10.20):

$$\hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta}} \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} |\hat{G}_N(i\omega) - G(i\omega, \boldsymbol{\theta})|^2 \phi_u(e^{i\omega h}) d\omega, \quad (10.24)$$

where $\hat{G}_N(i\omega)$ is given by (10.7).

Remark 10.5. Another alternative is to pick $\hat{G}_N^r(i\omega)$ (i.e., the frequency response estimate obtained through the regularized least-squares estimator) instead of $\hat{G}_N(i\omega)$ for computing $\hat{\boldsymbol{\theta}}$ in (10.24). In such case, we also suggest using (10.21) as the covariance matrix since it is well known that the regularized least-squares estimator has the same asymptotic distribution as its non-regularized version. This fact follows the intuition that regularization is a transient phenomenon [145].

All we need to know now is how to compute (10.24). In this work, we propose to compute $\hat{\boldsymbol{\theta}}$ by following a refined instrumental approach in the frequency domain. This method differs from other instrumental variable methods in the frequency domain such as [79] in several aspects:

- The measured data provides a continuous and smooth frequency response, not a collection of data points. This curve is obtained through the non-causal least-squares procedure from Section 10.3 based on time-domain input and output data, and therefore we do not rely on additional hardware equipment to obtain the non-parametric estimate.
- The estimate is obtained by optimizing over a continuum of frequencies instead of over a discrete frequency grid.
- The algorithm explicitly incorporates the input spectrum in the analysis. Conversely, in [79] the uncertainty of the frequency response estimate at each frequency is not explicitly taken into account in the algorithm.
- We do not employ frequency localizing basis functions [238] to improve the conditioning of the modified normal matrix. Such technique might be applicable in the proposed method, although we do not explore this tool here.

The main idea behind our method is derived below. First, we factor the input spectrum as $\phi_u(e^{i\omega h}) = \lambda^2 |L(e^{i\omega h})|^2$, where $L(e^{i\omega h})$ is a causal, asymptotically stable transfer function. Then, a stationary point of the cost in (10.24) must satisfy the first order condition

$$\begin{aligned} & \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} L(e^{i\omega h}) \frac{\partial G(i\omega, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \overline{L(e^{i\omega h})} \overline{(G(i\omega, \boldsymbol{\theta}) - \hat{G}_N(i\omega))} d\omega = \mathbf{0} \\ \iff & \left[\int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} \hat{\Psi}_f(i\omega, \boldsymbol{\theta}) \Psi_f^H(i\omega, \boldsymbol{\theta}) d\omega \right] \boldsymbol{\theta} = \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} \hat{\Psi}_f(i\omega, \boldsymbol{\theta}) \overline{\nu_f(i\omega, \boldsymbol{\theta})} d\omega, \end{aligned} \quad (10.25)$$

where

$$\Psi_f(i\omega, \boldsymbol{\theta}) = \frac{L(e^{i\omega h})}{A(i\omega, \boldsymbol{\theta})} \left[-(i\omega) \hat{G}_N(i\omega), \dots, -(i\omega)^n \hat{G}_N(i\omega), 1, \dots, (i\omega)^m \right]^T, \quad (10.26a)$$

$$\hat{\Psi}_f(i\omega, \boldsymbol{\theta}) = \frac{L(e^{i\omega h})}{A(i\omega, \boldsymbol{\theta})} \left[-(i\omega) \frac{B(i\omega, \boldsymbol{\theta})}{A(i\omega, \boldsymbol{\theta})}, \dots, -(i\omega)^n \frac{B(i\omega, \boldsymbol{\theta})}{A(i\omega, \boldsymbol{\theta})}, 1, \dots, (i\omega)^m \right]^T, \quad (10.26b)$$

$$\nu_f(i\omega, \boldsymbol{\theta}) = L(e^{i\omega h}) \frac{\hat{G}_N(i\omega)}{A(i\omega, \boldsymbol{\theta})}. \quad (10.26c)$$

Following the notation from the refined instrumental variable methods investigated in the previous chapters, the vectors $\Psi_f(i\omega, \boldsymbol{\theta})$, $\hat{\Psi}_f(i\omega, \boldsymbol{\theta})$ and $\nu_f(i\omega, \boldsymbol{\theta})$ in (10.26) are referred to as the filtered regressor vector, filtered instrument vector, and filtered output, respectively. Since $\boldsymbol{\theta}$ in (10.25) cannot be obtained in closed form, we propose to solve (10.25) via iterations. The proposed parametric method can be found in Algorithm 10.2.

Remark 10.6. In practice, the integrals in (10.27) can be replaced by sums using, e.g., Riemann or trapezoidal approximations [216], and the number of points can be set as another user-defined parameter.

Algorithm 10.2: Frequency-domain simplified refined instrumental variable method for continuous-time systems with band-limited inputs (FSRIVC-BL)

- 1: Input: Data $\{u(kh), y(kh)\}_{k=1}^N$, model order (n, m) , initial parameter estimate $\boldsymbol{\theta}_1$, tolerance factor ϵ and maximum number of iterations M
- 2: Using $\{u(kh), y(kh)\}_{k=1}^N$, compute $\hat{G}_N(i\omega)$ from (10.7) or $\hat{G}_N^r(i\omega)$ from Algorithm 10.1
- 3: Compute the factor $L(e^{i\omega h})$ from (10.22)
- 4: Using $\boldsymbol{\theta}_1$, form the estimated system polynomials $A(p, \boldsymbol{\theta}_1)$ and $B(p, \boldsymbol{\theta}_1)$
- 5: $j \leftarrow 1$, flag $\leftarrow 1$
- 6: **while** $\text{flag} = 1$ and $j \leq M$ **do**
- 7: Form $\Psi_f(i\omega, \boldsymbol{\theta}_j)$, $\hat{\Psi}_f(i\omega, \boldsymbol{\theta}_j)$ and $\nu_f(i\omega, \boldsymbol{\theta}_j)$ from (10.26a), (10.26b) and (10.26c), respectively. For this, use $\hat{G}_N(i\omega)$ or $\hat{G}_N^r(i\omega)$ if desired
- 8: Compute the parameter estimate

$$\boldsymbol{\theta}_{j+1} \leftarrow \left[\int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} \hat{\Psi}_f(i\omega, \boldsymbol{\theta}_j) \Psi_f^H(i\omega, \boldsymbol{\theta}_j) d\omega \right]^{-1} \left[\int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} \hat{\Psi}_f(i\omega, \boldsymbol{\theta}_j) \overline{\nu_f(i\omega, \boldsymbol{\theta}_j)} d\omega \right] \quad (10.27)$$

- 9: **if** $1/A_{j+1}(p)$ is unstable **then**
 - 10: **for** $i = 1, \dots, n$ **do**
 - 11: **if** pole p_i of $1/A_{j+1}(p)$ is unstable **then**
 - 12: $\text{Re}\{p_i\} \leftarrow -\text{Re}\{p_i\}$
 - 13: **end if**
 - 14: **end for**
 - 15: $A_{j+1}(p) \leftarrow \prod_{i=1}^n (p - p_i) / \prod_{i=1}^n (-p_i)$
 - 16: **end if**
 - 17: **if** $\frac{\|\boldsymbol{\theta}_{j+1} - \boldsymbol{\theta}_j\|}{\|\boldsymbol{\theta}_j\|} < \epsilon$ **then**
 - 18: $\text{flag} \leftarrow 0$
 - 19: **end if**
 - 20: $j \leftarrow j + 1$
 - 21: **end while**
 - 22: Output: $\boldsymbol{\theta}_j$ and its associated parametric model $B_j(p)/A_j(p)$.
-

What has not been discussed is how to obtain $\boldsymbol{\theta}_1$, the initialization of the method in Algorithm 10.2. One way to proceed is to estimate the bandwidth λ of the continuous-time system from input and output data, and later compute

$$\boldsymbol{\theta}_1 = \left[\int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} \Psi_f(i\omega, \boldsymbol{\theta}_{\text{init}}) \Psi_f^H(i\omega, \boldsymbol{\theta}_{\text{init}}) d\omega \right]^{-1} \left[\int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} \Psi_f(i\omega, \boldsymbol{\theta}_{\text{init}}) \overline{\nu_f(i\omega, \boldsymbol{\theta}_{\text{init}})} d\omega \right], \quad (10.28)$$

where $\Psi_f(i\omega, \boldsymbol{\theta}_{\text{init}})$ and $\nu_f(i\omega, \boldsymbol{\theta}_{\text{init}})$ are computed by (10.26a) and (10.26c), and $\boldsymbol{\theta}_{\text{init}}$ is such that $A(i\omega, \boldsymbol{\theta}_{\text{init}}) = (i\omega/\lambda + 1)^n$. This estimator can be interpreted as a frequency-domain variant of the LSSVF estimator [242] for band-limited signals.

10.4.2 Extension to Hybrid Box-Jenkins models

The FSRIVC-BL estimator described in Algorithm 10.2 has been constructed considering an output error model structure. In case the spectrum of the noise is also of interest, we can extend the ideas of the previous subsection to hybrid Box-Jenkins models. Assume that we now intend to estimate a model of the form

$$\begin{aligned} x(t) &= G(p, \boldsymbol{\theta})u(t), \\ y(kh) &= x(kh) + H(q, \boldsymbol{\eta})e(kh), \end{aligned}$$

where $e(kh)$ describes a zero-mean white noise stochastic process of finite variance σ^2 , independent of the input sequence, and $H(q, \boldsymbol{\eta}) = C(q, \boldsymbol{\eta})/D(q, \boldsymbol{\eta})$ is an asymptotically stable and minimum-phase transfer function where

$$\begin{aligned} C(q, \boldsymbol{\eta}) &= 1 + c_1 q^{-1} + c_2 q^{-2} + \cdots + c_{m_c} q^{-m_c}, \\ D(q, \boldsymbol{\eta}) &= 1 + d_1 q^{-1} + d_2 q^{-2} + \cdots + d_{n_d} q^{-n_d}, \end{aligned}$$

with the degrees of the polynomials satisfying $n_d \geq m_c$. The parameters in the noise model can be described by the vector $\boldsymbol{\eta}$ of the form (2.12). The goal is to develop a method that can also find an estimate of the parameter vector $\boldsymbol{\eta}$. This can be done by again computing a minimizer for (10.20), where this time the covariance matrix of $\hat{\rho}_N$ for large N can be approximated by

$$\boldsymbol{\Sigma} \approx (\boldsymbol{\Phi}^\top \boldsymbol{\Phi})^{-1} (\boldsymbol{\Phi}^\top \mathbf{R} \boldsymbol{\Phi}) (\boldsymbol{\Phi}^\top \boldsymbol{\Phi})^{-1}, \quad (10.29)$$

where \mathbf{R} is a Toeplitz matrix that has elements $\mathbf{R}_{k-l} = \mathbb{E}\{H(q, \boldsymbol{\eta})e(kh)H(q, \boldsymbol{\eta})e(lh)\}$. As M_{nc} , M_c , and N tend to infinity, all three matrices in parenthesis in (10.29) are asymptotically Toeplitz. Moreover, the symbol of the Toeplitz operator characterized by $\boldsymbol{\Phi}^\top \boldsymbol{\Phi}/N$ is $\phi_u(e^{i\omega h})$, while the symbol for $\boldsymbol{\Phi}^\top \mathbf{R} \boldsymbol{\Phi}/N$ is $\sigma^2 |H(e^{i\omega h}, \boldsymbol{\eta})|^2 \phi_u(e^{i\omega h})$. Thus, by the asymptotic results for Toeplitz matrices in [102, Ch. 5], we see that $\boldsymbol{\Sigma}$ is asymptotically Toeplitz with symbol $\phi_u(e^{i\omega h})/(\sigma^2 |H(e^{i\omega h}, \boldsymbol{\eta})|^2)$. Following a

similar derivation as in (10.23), we conclude that

$$\begin{aligned} V(\boldsymbol{\theta}, \boldsymbol{\eta}) &= \frac{1}{N} [\boldsymbol{\rho}(\boldsymbol{\theta}) - \hat{\boldsymbol{\rho}}_N]^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{\eta}) [\boldsymbol{\rho}(\boldsymbol{\theta}) - \hat{\boldsymbol{\rho}}_N] \\ &\approx \frac{1}{2\pi\sigma^2} \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} \frac{|\hat{G}_N(i\omega) - G(i\omega, \boldsymbol{\theta})|^2}{|H(e^{i\omega h}, \boldsymbol{\eta})|^2} \phi_u(e^{i\omega h}) d\omega, \end{aligned} \quad (10.30)$$

which is a natural result, as it is similar to the prediction error method cost for discrete-time systems [138] but with a mixture of continuous and discrete-time models.

The cost function in (10.30) can be minimized following a similar principle as in Section 10.4.1; however, $\boldsymbol{\eta}$ is another variable that must be taken into consideration. The approach we follow here is similar to the one used in the RIVC method for identification in the time domain. That is, for a fixed $\boldsymbol{\theta}_j$, we first obtain an estimate of $\boldsymbol{\eta}$ by fitting an ARMA model to the residual

$$\hat{v}(kh, \boldsymbol{\theta}_j) = y(kh) - \left\{ \frac{B_j(p)}{A_j(p)} u(t) \right\} \Big|_{t=kh}. \quad (10.31)$$

This estimate is denoted as $\boldsymbol{\eta}_{j+1}$. Later, the next iteration of $\boldsymbol{\theta}_j$ (denoted as $\boldsymbol{\theta}_{j+1}$) is computed by an iteration of the instrumental variable method in Algorithm 10.2, but with $L(e^{i\omega h})$ in (10.26) replaced by $L(e^{i\omega h})/H(e^{i\omega h}, \boldsymbol{\eta}_{j+1})$. This entire procedure can be seen as an RIVC algorithm in the frequency domain for band-limited signals, and it is depicted in Figure 10.4. We call this method FRIVC-BL.

10.4.3 Simulations

Here we devise two simulation experiments to assess the parametric estimators we have derived.

Tests on two systems in an output error model structure

The first example considers the same two systems in (10.19), under the same experimental conditions as in Subsection 10.3.4, that is, $N = 100$, $M_{nc} = 15$, $M_c = 24$, and the signal-to-noise ratio is approximately five. The input sequence is white noise of unit variance, which is interpolated by sinc functions as in (10.9). The sampling periods for the systems $G_1^*(p)$ and $G_2^*(p)$ are $h_1 = 0.3[\text{s}]$ and $h_2 = 1[\text{s}]$, respectively.

We perform 300 Monte Carlo runs and obtain the box plots for the fit metric corresponding to the following estimators:

- a) Non-causal least-squares with $(M_{nc}, M_c) = (15, 24)$ (NC-LS);
- b) Non-causal TC-regularized least-squares with $(M_{nc}, M_c) = (15, 24)$ (NC-TC);
- c) Initial parametric estimate given by (10.28) (FLSSVF);

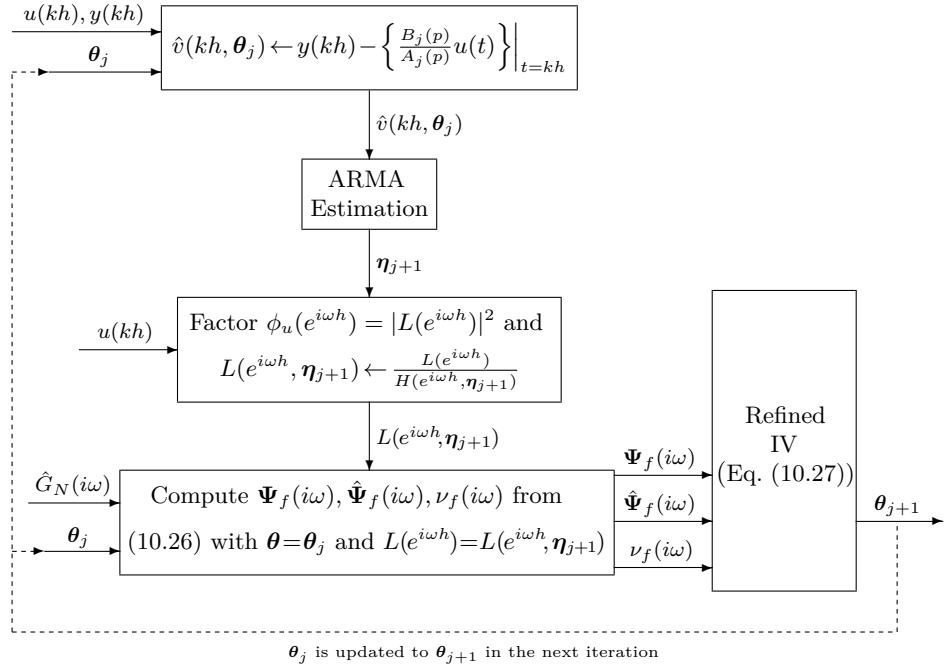


Figure 10.4: Implementation of the frequency-domain RIVC method for band-limited signals. The algorithm is initialized with $\theta = \theta_1$ (i.e., $j = 1$). The frequency response estimate $\hat{G}_N(i\omega)$ can be obtained from (10.7) or from $\hat{G}_N^r(i\omega)$ in Algorithm 10.1.

- d) FSRIVC-BL with NC-LS as its base non-parametric estimate (FSRIVC-LS);
- e) FSRIVC-BL with NC-TC as its base non-parametric estimate (FSRIVC-TC);
- f) Extended SRIVC for band-limited inputs (SRIVC-x, see Chapter 6).

The extended SRIVC estimate we consider is computed following Algorithm 6.1, and the filtering is done with the steps described in Section 6.5.1 with an oversampling factor M of 100 (that is, the input is sampled 100 times faster). We initialize the FSRIVC-BL and SRIVC-x estimators with the FLSSVF estimate, and the maximum number of iterations for the FSRIVC-BL and SRIVC-x methods is set to 50, with tolerance factor $\epsilon = 10^{-8}$. The integrals in (10.27) are computed from a Riemann approximation of 20000 points.

The results in Figure 10.5 show that the parametric estimators based on non-causal impulse responses achieve excellent performance in both systems, and they are comparable with SRIVC-x for $G_1^*(p)$. The system $G_2^*(p)$ is more difficult to estimate for these estimators, possibly due to the fact that a more significant non-causal part leads to greater bias in the frequency response when not enough non-causal coefficients are computed.

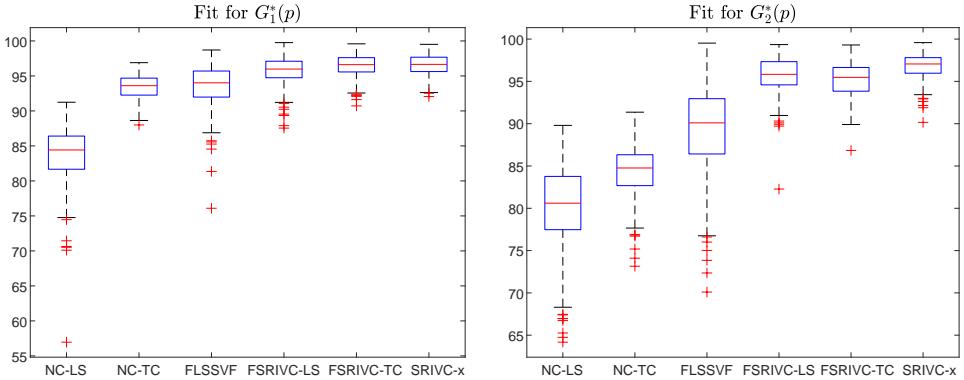


Figure 10.5: Fit box plots of six different estimators: NC-LS, NC-TC, FLSSVF, FSRIVC-LS, FSRIVC-TC, and SRIVC-x. Left: $G_1^*(p)$; right: $G_2^*(p)$.

Table 10.1: Sample sizes considered for each Monte Carlo simulation study, and the causal and non-causal terms used to compute $\hat{G}_N(i\omega)$ for each sample size.

N	100	200	500	1000	2000	5000	10000	20000
M_c	25	50	100	200	300	400	500	600
M_{nc}	15	30	70	140	200	250	300	400

Table 10.2: Average time (in seconds) per Monte Carlo run for FSRIVC-LS and SRIVC-x.

N	100	200	500	1000	2000	5000	10000	20000
FSRIVC-LS	0.22	0.24	0.18	0.23	0.39	1.04	2.62	8.15
SRIVC-x	0.42	0.80	1.65	3.48	7.29	15.44	27.74	55.29

One of the computational advantages of using FSRIVC-BL instead of SRIVC-x is that the FSRIVC-BL estimator only depends on N when computing $\hat{G}_N(i\omega)$, which can be easily obtained by least-squares. After this, the iterations are in the frequency domain; that is, only evaluations of frequency responses are required, and the integrals in (10.27) have a fixed integration interval and thus do not depend on the sample size. On the other hand, the SRIVC-x estimator requires continuous-time filtering that can be computationally intensive when the sample size is large, as seen in Section 6.6.4. To illustrate this we test the FSRIVC-BL (with an LS non-parametric estimate) and SRIVC-x estimators by tracking the mean fit and average time per iteration for an increasing sample size. We have recorded the sample size and causal/non-causal terms used in each simulation in Table 10.1. The mean fit is plotted in Figure 10.6, while Table 10.2 shows the average time per iteration for each method. We see that the mean fit of both FSRIVC-LS and SRIVC-x nearly match as N grows, while the FSRIVC-LS estimates are at least 7 times faster to generate than the SRIVC-x estimates.

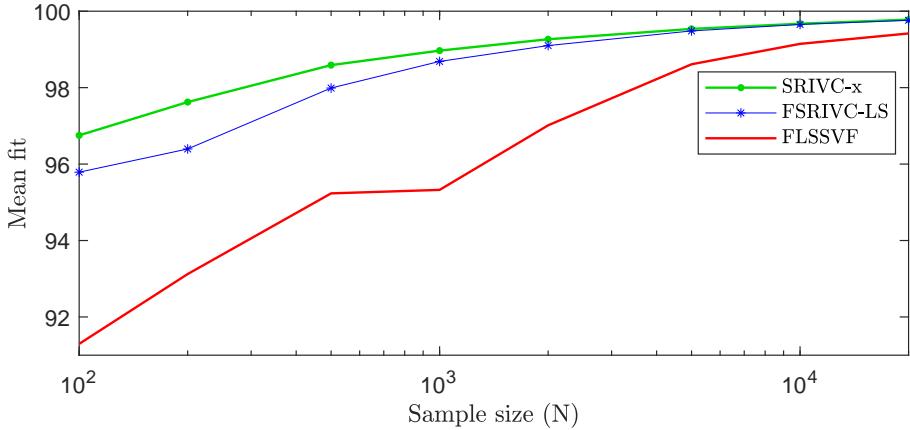


Figure 10.6: Mean fits of each method per sample size.

Tests with a hybrid Box-Jenkins model structure

We now focus on identifying $G_2^*(p)$ in (10.19) with an additive noise on the output that is given by

$$v(kh) = \frac{1 + 0.5q^{-1}}{1 - 0.85q^{-1}} e(kh),$$

where $e(kh)$ is white noise of variance 0.05. The number of samples obtained for identification is $N = 5000$, and 300 Monte Carlo runs are performed to assess the performance of the FRIVC-BL estimator and its initialization given by the FLSSVF method, compared to the standard RIVC estimator as implemented in the CONTSID toolbox version 7.4 [70]. The ARMA estimation step of the FRIVC-BL estimator is computed using the `armax` command of the System Identification Toolbox in MATLAB [143], and the continuous-time filtering needed for computing the residual in (10.31) is implemented according to Section 6.5. Both instrumental variable methods use a tolerance factor of $\epsilon = 10^{-8}$, and the maximum number of iterations is set to 50.

In Table 10.3 we present the mean and standard deviation of each estimated parameter using the standard RIVC method, the FLSSVF method, and the FRIVC-BL method with the least-squares estimator providing the base non-parametric estimate. The proposed FRIVC-BL estimator achieves great precision in all estimated parameters, and the FLSSVF method provides a good initial point for the iterations of the FRIVC-BL method in this study. The standard RIVC method fails to estimate adequate models; this poor performance can be explained by two factors: a) a low sampling frequency rate induces problems in computing the simulated output of continuous-time systems using the `lsim` command in MATLAB, and b) the standard RIVC estimator does not consider the correct intersample behavior of the input.

Table 10.3: Mean and standard deviation of the estimated system and noise model parameters obtained with 300 Monte Carlo runs with a hybrid Box-Jenkins model structure.

Method	Parameter True value	a_1 0.0488	a_2 0.1220	b_0 −0.3484	c_1 0.5	d_1 −0.85
RIVC	Mean	56.65	154.55	$5.65 \cdot 10^{-2}$	0.4715	0.7649
	Std. Dev.	$2.70 \cdot 10^2$	$1.34 \cdot 10^3$	$8.01 \cdot 10^{-1}$	$4.06 \cdot 10^{-2}$	$8.26 \cdot 10^{-2}$
FLSSVF	Mean	0.0482	0.1226	−0.3467		
	Std. Dev.	$2.49 \cdot 10^{-3}$	$6.71 \cdot 10^{-4}$	$1.75 \cdot 10^{-2}$		
FRIVC-BL	Mean	0.0488	0.1220	−0.3482	0.4976	−0.8499
	Std. Dev.	$3.93 \cdot 10^{-4}$	$2.84 \cdot 10^{-5}$	$2.77 \cdot 10^{-3}$	$8.08 \cdot 10^{-3}$	$1.34 \cdot 10^{-2}$

10.5 Multisine inputs: analysis of non-parametric and parametric estimators

This section departs from the design of parametric estimators of Section 10.4 and focuses on multisine inputs, which are particular cases of band-limited signals. These signals are advantageous for identification purposes due to their flexibility regarding power spectrum design, time domain averaging possibilities, simplification of the model validation step, and finite-sample estimation performance [195]. The input signals we study are assumed without loss of generality to be of the form

$$u(t) = \alpha_0 + \sum_{l=1}^M \alpha_l \cos(\omega_l t + \psi_l), \quad (10.32)$$

where $\{\alpha_l\}_{l=0}^M$, $\{\omega_l\}_{l=1}^M$ and $\{\psi_l\}_{l=1}^M$ are user-defined parameters. The frequencies ω_l are assumed to be positive and sorted in strictly ascending order, and we also assume that α_l is different from zero for $l = 0, 1, \dots, M$. Clearly, this is a band-limited signal with bandwidth $\omega_B = \omega_M$ that has $2M + 1$ frequency lines [138].

10.5.1 Frequency response estimator from least-squares

To begin our analysis, we propose studying the frequency response estimator that arises from the least-squares estimation of the band-limited equivalent impulse response. Since we have shown that this impulse response is non-causal in general, we consider estimating $g_{BL}(-M_{nc}h)$ to $g_{BL}([2M - M_{nc}]h)$, where M_{nc} is the number of non-causal terms. The least-squares estimate of g_{BL} is given by (10.6) and its associated frequency response estimate is $\hat{G}_N(i\omega) = \hat{\rho}_N^\top \Gamma(e^{i\omega h})$, where $\Gamma(e^{i\omega h})$ is given by (10.8). We first show that this frequency-response estimate does not depend on the integer M_{nc} at the frequencies excited by the input. For the sequel, we denote

$$\tilde{\Gamma} := \left[\Gamma(1), \Gamma(e^{i\omega_1 h}), \Gamma(e^{-i\omega_1 h}), \dots, \Gamma(e^{i\omega_M h}), \Gamma(e^{-i\omega_M h}) \right].$$

Note that this matrix is non-singular. Indeed, it can be written as

$$\tilde{\Gamma} = h \begin{bmatrix} 1 & 1 & 1 & \dots & 1 & 1 \\ 1 & e^{-i\omega_1 h} & e^{i\omega_1 h} & \dots & e^{-i\omega_M h} & e^{i\omega_M h} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 1 & e^{-2i\omega_1 Mh} & e^{2i\omega_1 Mh} & \dots & e^{-2i\omega_M Mh} & e^{2i\omega_M Mh} \end{bmatrix} \times \begin{bmatrix} 1 & & & & 0 \\ e^{i\omega_1 M_{nc} h} & & & & \\ & e^{-i\omega_1 M_{nc} h} & & & \\ & & \ddots & & \\ & & & e^{i\omega_M M_{nc} h} & \\ 0 & & & & e^{-i\omega_M M_{nc} h} \end{bmatrix}, \quad (10.33)$$

from which the invertibility is revealed, as it is the product of a non-singular Vandermonde matrix [115] with a diagonal matrix with non-zero diagonal entries. With this description we can decompose $\varphi(kh)$ as

$$\varphi(kh) = \tilde{\Gamma} \zeta(kh), \quad (10.34)$$

where $\zeta(kh)$ is defined as

$$\zeta(kh) = \left[\alpha_0, \frac{\alpha_1}{2} e^{i(\omega_1 kh + \psi_1)}, \frac{\alpha_1}{2} e^{-i(\omega_1 kh + \psi_1)}, \dots, \frac{\alpha_M}{2} e^{i(\omega_M kh + \psi_M)}, \frac{\alpha_M}{2} e^{-i(\omega_M kh + \psi_M)} \right]^T. \quad (10.35)$$

Note that the vector $\zeta(kh)$ does not depend on M_{nc} .

Lemma 10.1. *Consider the frequency response estimate $\hat{G}_N(i\omega)$ in (10.7), and assume that the input to the continuous-time system is a multisine of the form (10.32). Then, the frequency response does not depend on the value of M_{nc} at $\omega = 0, \pm\omega_1, \pm\omega_2, \dots, \pm\omega_M$.*

Proof. See Appendix 10.A.3. □

Lemma 10.1 shows that the user can choose any FIR model (of consecutive coefficients) of length $2M + 1$, and the frequency response estimates from least-squares will be invariant to this choice for any sample size, and independently of whether the output is in a stationary regime or not. Intuitively, the causality of the model does not play a role due to the periodicity of the input signal; however, there is an observable difference in the frequency response estimate evaluated at frequencies not present in the input since the trigonometric polynomial in (10.8) depends on M_{nc} .

The next result concerns the accuracy of this estimator. In the sequel we say that the output is in a *stationary regime* when the noiseless output is also a multisine, of

the form

$$x(t) = \alpha_0 G(0) + \sum_{l=1}^M \alpha_l |G(i\omega_l)| \cos(\omega_l t + \psi_l + \angle G(i\omega_l)).$$

Theorem 10.3. *Assume that the input to the continuous-time system described in (10.1) is a multisine of the form (10.32), and that the output is in a stationary regime. Consider the least-squares estimate in (10.6) for any integer M_{nc} and $N \geq 2M + 1$. Then, its associated frequency response estimate $\hat{G}_N(i\omega)$ is unbiased at $\omega = 0, \pm\omega_1, \pm\omega_2, \dots, \pm\omega_M$.*

Proof. See Appendix 10.A.4. \square

As stated in Theorem 10.3, the least-squares estimator of the non-causal impulse response leads to unbiased estimates of the frequency response of the continuous-time system at the frequencies of interest. For the frequencies that are not excited by the multisine input, the least-squares method delivers an interpolation of the frequency response that is usually reliable if the frequencies that are excited are sufficiently densely distributed among the bandwidth of interest.

To conclude this first batch of results, we are also interested in determining the covariance of these estimates. This covariance is computed explicitly in Theorem 10.4.

Theorem 10.4. *Consider the least-squares estimate in (10.6), and its associated frequency response estimate $\hat{G}_N(i\omega) = \boldsymbol{\Gamma}^\top(e^{i\omega h})\hat{\boldsymbol{\rho}}_N$. Assume that the noisy output $\{y(kh)\}_{k=1}^N$ is in a stationary regime, and that the sampling period is less than π/ω_M . Then, the following result holds for any $N \geq 2M + 1$:*

$$\text{Cov} \left\{ \begin{bmatrix} \hat{G}_N(0) \\ \hat{G}_N(i\omega_1) \\ \hat{G}_N(-i\omega_1) \\ \vdots \\ \hat{G}_N(i\omega_M) \\ \hat{G}_N(-i\omega_M) \end{bmatrix} \right\} = \sigma^2 \tilde{\boldsymbol{\Gamma}}^\top \left[\sum_{k=1}^N \boldsymbol{\varphi}(kh) \boldsymbol{\varphi}^\top(kh) \right]^{-1} \tilde{\boldsymbol{\Gamma}} = \sigma^2 \left[\sum_{k=1}^N \overline{\boldsymbol{\zeta}(kh)} \boldsymbol{\zeta}^\top(kh) \right]^{-1}, \quad (10.36)$$

where $\boldsymbol{\zeta}(kh)$ is given by (10.35). Furthermore, if Nh is chosen to be a multiple of the least common multiple of $\{2\pi/\omega_l\}_{l=1}^M$ (if such quantity is finite), then the frequency response estimates are uncorrelated.

Proof. See Appendix 10.A.5. \square

If the noise is Gaussian-distributed, it follows from the proof of Theorem 10.4 that the frequency response estimates are also Gaussian-distributed for any $N \geq 2M + 1$. The least common multiple constraint in Theorem 10.4 for uncorrelatedness is related to spectral leakage: for such values of Nh , the multisine input signal is applied for a total time equal to a multiple of the period of the signal. Even if

this condition is not satisfied, the frequency estimates will still be asymptotically uncorrelated, as stated in Corollary 10.1.

Corollary 10.1. *Consider the frequency response estimate in (10.7), obtained from using (10.32) as an input signal. If the sampling period is less than π/ω_M , the frequency response estimates $\hat{G}_N(0), \hat{G}_N(\pm i\omega_1), \dots, \hat{G}_N(\pm i\omega_M)$ are asymptotically uncorrelated. Furthermore, the asymptotic covariance is given by*

$$\text{AsCov} \left\{ \begin{bmatrix} \hat{G}_N(0) \\ \hat{G}_N(i\omega_1) \\ \hat{G}_N(-i\omega_1) \\ \vdots \\ \hat{G}_N(i\omega_M) \\ \hat{G}_N(-i\omega_M) \end{bmatrix} \right\} = \sigma^2 \begin{bmatrix} \alpha_0^{-2} & & & & & 0 \\ & 4\alpha_1^{-2} & & & & \\ & & 4\alpha_1^{-2} & & & \\ & & & \ddots & & \\ & & & & 4\alpha_M^{-2} & \\ 0 & & & & & 4\alpha_M^{-2} \end{bmatrix}.$$

Proof. See Appendix 10.A.6. \square

10.5.2 From non-parametric to parametric estimators: multisine case

The finite-time properties stated in Theorems 10.3 and 10.4 provide a first step to obtaining a parametric estimator based on the non-parametric one, similar to the ideas explored in Section 10.4. For the following, consider the model structure

$$G(p, \boldsymbol{\theta}) = \frac{B(p, \boldsymbol{\theta})}{A(p, \boldsymbol{\theta})} = \frac{b_m p^m + b_{m-1} p^{m-1} + \dots + b_0}{a_n p^n + a_{n-1} p^{n-1} + \dots + a_1 p + 1}, \quad (10.37)$$

where $n \geq m$, and $\boldsymbol{\theta} = [a_1, \dots, a_n, b_0, \dots, b_m]^\top$. For the sequel, we define the vectors

$$\mathbf{G}(\boldsymbol{\theta}) := \left[G(0, \boldsymbol{\theta}), G(i\omega_1, \boldsymbol{\theta}), G(-i\omega_1, \boldsymbol{\theta}), \dots, G(i\omega_M, \boldsymbol{\theta}), G(-i\omega_M, \boldsymbol{\theta}) \right]^\top \quad (10.38)$$

and

$$\mathbf{G}^r(\boldsymbol{\theta}) := \left[G(0, \boldsymbol{\theta}), \text{Re}\{G(i\omega_1, \boldsymbol{\theta})\}, \text{Im}\{G(i\omega_1, \boldsymbol{\theta})\}, \dots, \text{Re}\{G(i\omega_M, \boldsymbol{\theta})\}, \text{Im}\{G(i\omega_M, \boldsymbol{\theta})\} \right]^\top. \quad (10.39)$$

We are interested in obtaining an estimate $\hat{\boldsymbol{\theta}}_N$ whose associated model $G(p, \hat{\boldsymbol{\theta}}_N)$ accurately represents the true continuous-time system $G^*(p)$. A key result is given next, that relates the least-square estimates of the frequency response with the sampled input and output data.

Theorem 10.5. *Consider the system description (10.2), where $\{v(kh)\}_{k=1}^N$ is a Gaussian white noise signal of variance σ^2 , and the output is in a stationary regime.*

Then,

$$\hat{\mathbf{G}}_N^r := \left[\hat{G}_N(0), \operatorname{Re}\{\hat{G}_N(i\omega_1)\}, \operatorname{Im}\{\hat{G}_N(i\omega_1)\}, \dots, \operatorname{Re}\{\hat{G}_N(i\omega_M)\}, \operatorname{Im}\{\hat{G}_N(i\omega_M)\} \right]^\top$$

is a sufficient statistic for $\boldsymbol{\theta}$.

Proof. See Appendix 10.A.7. See also Remark 10.8 in that appendix for an alternative proof. \square

In other words, if the output is in a stationary regime, then the knowledge that the data $\{u(kh), y(kh)\}_{k=1}^N$ contains about $\boldsymbol{\theta}$ is completely captured by the least-squares estimate of the vector of frequency responses

$$\hat{\mathbf{G}}_N := \left[\hat{G}_N(0), \hat{G}_N(i\omega_1), \hat{G}_N(-i\omega_1), \dots, \hat{G}_N(i\omega_M), \hat{G}_N(i\omega_M) \right]^\top.$$

This result has several implications on the methods proposed for parametric system identification, some of which are explicitly stated in the following corollary.

Corollary 10.2. *Assume that the additive measurement noise $\{v(kh)\}_{k=1}^N$ is Gaussian and white, and that the output is in a stationary regime. Then, the maximum likelihood estimator of $\boldsymbol{\theta}$ based on the data $\{u(kh), y(kh)\}_{k=1}^N$, where $\{u(kh)\}_{k=1}^N$ are samples of the multisine (10.32), is given by*

$$\hat{\boldsymbol{\theta}}_N = \arg \min_{\boldsymbol{\theta}} \frac{1}{2} \left(\hat{\mathbf{G}}_N - \mathbf{G}(\boldsymbol{\theta}) \right)^\text{H} \left[\operatorname{Cov}(\hat{\mathbf{G}}_N) \right]^{-1} \left(\hat{\mathbf{G}}_N - \mathbf{G}(\boldsymbol{\theta}) \right), \quad (10.40)$$

where $\operatorname{Cov}(\hat{\mathbf{G}}_N)$ and $\mathbf{G}(\boldsymbol{\theta})$ are given by (10.36) and (10.38), respectively. This estimator also minimizes the prediction error

$$V_N(\boldsymbol{\theta}) := \frac{1}{2} \sum_{k=1}^N (y(kh) - y(kh, \boldsymbol{\theta}))^2, \quad (10.41)$$

where

$$y(kh, \boldsymbol{\theta}) = \alpha_0 G(0, \boldsymbol{\theta}) + \sum_{l=1}^M \alpha_l |G(i\omega_l, \boldsymbol{\theta})| \cos(\omega_l kh + \psi_l + \angle G(i\omega_l, \boldsymbol{\theta})).$$

Proof. The result follows directly from Theorem 10.5. \square

Corollary 10.2 implies that the direct and indirect prediction error methods [212] provide the exact maximum likelihood estimator when the input is a continuous-time multisine signal. This relationship is not asymptotic: it is valid for any $N \geq 2M+1$ if the output is in a stationary regime. This finding is directly related to the extended SRIVC estimator analyzed in Chapter 6. Via Theorem 6.2, we know that such method obtains a stationary point of the likelihood equation at convergence for the

multisine input case and thus, according to Corollary 10.2, it is a way to solve the minimization problem in (10.40) via iterations.

The relationship between the prediction error method and its indirect version in (10.40) is not new; it has been previously stated for some specific cases in [212]. However, in this framework it permits us to obtain extra information on the maximum likelihood estimator that to the best of our knowledge has not been thoroughly investigated in continuous-time settings. More precisely, we can observe that for $n + m \geq 2M$ the minimum of the cost function in (10.40) is zero, and it is achieved for a parameter vector $\tilde{\theta}_N$ satisfying $\hat{\mathbf{G}}_N = \mathbf{G}(\tilde{\theta}_N)$.

Corollary 10.3. *Consider the system description in (10.2) with (10.32) as input signal, and assume that the output is in a stationary regime. Moreover, assume that $n + m \geq 2M$, and that there exists² a transfer function of the form (10.37) with a real-valued parameter vector $\tilde{\theta}_N$ such that*

$$G(0, \tilde{\theta}_N) = \hat{G}_N(0); \quad G(\pm i\omega_l, \tilde{\theta}_N) = \hat{G}_N(\pm i\omega_l), \quad l = 1, 2, \dots, M.$$

Then, the following statements are true:

1. *The vector $\tilde{\theta}_N$ minimizes the prediction error costs in (10.40) and (10.41), i.e., $\tilde{\theta}_N = \hat{\theta}_N$.*
2. *The parametric model $G(i\omega, \tilde{\theta}_N)$ is unbiased at the frequencies $\omega = 0, \pm\omega_1, \pm\omega_2, \dots, \pm\omega_M$, and the joint covariance of these estimates is given by (10.36).*

Proof. The first statement follows from the fact that $\mathbf{G}(\tilde{\theta}_N) = \hat{\mathbf{G}}_N$, which means that the cost function in (10.40) reaches zero, its minimum value. Note that $\tilde{\theta}_N$ is not necessarily unique; it is unique only if $n + m = 2M$. On the other hand, Statement 2 follows directly from Statement 1 and Theorems 10.3 and 10.4. \square

In other words, for $n + m \geq 2M$, an asymptotically optimal parametric estimator can be constructed by first forming the least-squares estimate of the frequency response at the frequencies of the input via (10.7) and later computing the rational function interpolation that has n poles and m zeros. As noted in [40], if $n + m > 2M$, it is possible to construct a transfer function with n' poles and m' zeros ($n' \leq n$ and $m' \leq m$) such that the interpolation conditions are satisfied. This implies that the optimization problem in (10.40) is likely to be ill-conditioned in this case, and algorithms that compute $\hat{\theta}_N$ (such as the extended SRIVC estimator in Chapter 6) can deliver transfer functions with zero-pole quasi-cancellations. This phenomenon is also related to the singularity of the normal matrix that is computed in the extended SRIVC algorithm, which can occur when the persistence of excitation of the input is not of sufficient order.

²It can be shown that such transfer function exists for all vectors $\hat{\mathbf{G}}_N \in \mathbb{C}^{2M-1}$, except at most for a set of Lebesgue measure zero.

With regards to the computation of the rational interpolation, there exists a wide range of literature on this mathematical problem (see, e.g., [40, 101]). One of the simplest approaches is to describe the interpolation problem as a set of linear equations. For $n + m \geq 2M$, the vector $\hat{\theta}_N = [\hat{a}_1, \dots, \hat{a}_n, \hat{b}_0, \dots, \hat{b}_m]^\top$ must satisfy

$$\begin{aligned}\hat{b}_0 &= \hat{G}_N(0), \\ \frac{\hat{b}_m(i\omega_l)^m + \hat{b}_{m-1}(i\omega_l)^{m-1} + \dots + \hat{b}_0}{\hat{a}_n(i\omega_l)^n + \hat{a}_{n-1}(i\omega_l)^{n-1} + \dots + \hat{a}_1(i\omega_l) + 1} &= \hat{G}_N(i\omega_l), \quad l = 1, 2, \dots, M, \\ \frac{\hat{b}_m(-i\omega_l)^m + \hat{b}_{m-1}(-i\omega_l)^{m-1} + \dots + \hat{b}_0}{\hat{a}_n(-i\omega_l)^n + \hat{a}_{n-1}(-i\omega_l)^{n-1} + \dots + \hat{a}_1(-i\omega_l) + 1} &= \hat{G}_N(-i\omega_l), \quad l = 1, 2, \dots, M.\end{aligned}$$

Thus, $\hat{\theta}_N$ is the solution of the set of linear equations described by

$$\left[\begin{array}{ccccccccc} 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ -(i\omega_1)\hat{G}_N(i\omega_1) & \dots & -(i\omega_1)^n\hat{G}_N(i\omega_1) & 1 & (i\omega_1) & \dots & (i\omega_1)^m \\ -(-i\omega_1)\hat{G}_N(-i\omega_1) & \dots & -(-i\omega_1)^n\hat{G}_N(-i\omega_1) & 1 & (-i\omega_1) & \dots & (-i\omega_1)^m \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ -(i\omega_M)\hat{G}_N(i\omega_M) & \dots & -(i\omega_M)^n\hat{G}_N(i\omega_M) & 1 & (i\omega_M) & \dots & (i\omega_M)^m \\ -(-i\omega_M)\hat{G}_N(-i\omega_M) & \dots & -(-i\omega_M)^n\hat{G}_N(-i\omega_M) & 1 & (-i\omega_M) & \dots & (-i\omega_M)^m \end{array} \right] \hat{\theta}_N = \hat{\mathbf{G}}_N. \quad (10.42)$$

Bear in mind that the solution of the linear equations above is real-valued, since the conjugate of $\hat{\theta}_N$ can be shown to satisfy the same equations above. Also, the parameter \hat{b}_0 will be unbiased (for finite sample size) if the input has a non-zero constant term, due to the fact that $\hat{G}_N(0)$ is unbiased.

Remark 10.7. This method has great similarities with Levy's method for rational function estimation [47]. In fact, the equation that is solved in (10.42) is the same as the one in, e.g., Problem 2.3. of [95]. For $n+m \geq 2M$, we see that these two methods are identical. However, for $n+m < 2M$, Levy's method solves the minimization problem

$$\min_{\theta} |A(0, \theta)\hat{G}_N(0) - B(0, \theta)|^2 + \sum_{l=1}^M \left| A(i\omega_l, \theta)\hat{G}_N(i\omega_l) - B(i\omega_l, \theta) \right|^2,$$

which does not have the same solution as (10.40), since the gradient of the cost function in (10.40) is not linear in the parameter vector θ .

A case study in finite-sample behavior

So far in this section, we have discussed how to obtain non-parametric estimators of frequency responses from multisine inputs and sampled output data. The least-squares estimator proposed here serves as a base for computing the maximum

likelihood estimate of the model structure (10.37), which is known to be a consistent and asymptotically efficient estimator in many scenarios [138]. The question we address here is: Can anything be affirmed about the non-asymptotic behavior of $\hat{\theta}_N$ in case the true system can be exactly characterized by a transfer function with n poles and m zeros? Via extensive simulations, we have found evidence that the mean value of $\hat{\theta}_N$ for finite sample size is close to the true parameter vector when $2M \geq n + m$ and the measurement noise is white. This estimator also provides a transfer function whose mean is close to $G^*(p)$ even when there exists over-parametrization in the one of the polynomials of the model structure (denominator or numerator).

An example is given below with some explicit results on the finite-sample distribution of the maximum likelihood estimate. We show that, even in a simple case, the distribution of the parameter estimate may not have any finite moments and that the unbiasedness of the frequency response values does not in general translate to (finite-sample) unbiasedness in the parameter estimates.

Example 10.1. Consider the true transfer function

$$G^*(p) = \frac{b_0^*}{a_1^*p + 1}$$

with $a_1^* > 0$ and $b_0^* \neq 0$. We are interested in obtaining an estimate for the parameter vector $\theta^* = [a_1^*, b_0^*]^\top$ when the output of $G^*(p)$ is corrupted by a Gaussian white noise of variance σ^2 . To this end, we consider the input $u(t) = \alpha_1 \sin(\omega_1 t)$ with $\alpha_1, \omega_1 > 0$. For simplicity, we take Nh to be a multiple of $2\pi/\omega_1$ so that the resulting frequency response estimates are uncorrelated according to Theorem 10.3.

By Corollary 10.3, the maximum likelihood estimator must satisfy

$$\hat{G}_N(i\omega_1) = \frac{\hat{b}_0}{i\hat{a}_1\omega_1 + 1} \quad \text{and} \quad \hat{G}_N(-i\omega_1) = \frac{\hat{b}_0}{-i\hat{a}_1\omega_1 + 1}.$$

These equations lead to

$$\begin{bmatrix} -i\omega_1 \hat{G}_N(i\omega_1) & 1 \\ i\omega_1 \hat{G}_N(-i\omega_1) & 1 \end{bmatrix} \begin{bmatrix} \hat{a}_1 \\ \hat{b}_0 \end{bmatrix} = \begin{bmatrix} \hat{G}_N(i\omega_1) \\ \hat{G}_N(-i\omega_1) \end{bmatrix} \implies \begin{bmatrix} \hat{a}_1 \\ \hat{b}_0 \end{bmatrix} = \begin{bmatrix} -\frac{\text{Im}\{\hat{G}_N(i\omega_1)\}}{\omega_1 \text{Re}\{\hat{G}_N(i\omega_1)\}} \\ \frac{|\hat{G}_N(i\omega_1)|^2}{\text{Re}\{\hat{G}_N(i\omega_1)\}} \end{bmatrix}.$$

Let us focus on \hat{a}_1 . This parameter estimate is given by the quotient of two random variables that are jointly Gaussian-distributed:

$$\begin{bmatrix} \text{Re}\{\hat{G}_N(i\omega_1)\} \\ \text{Im}\{\hat{G}_N(i\omega_1)\} \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} \text{Re}\{G^*(i\omega_1)\} \\ \text{Im}\{G^*(i\omega_1)\} \end{bmatrix}, \mathbf{P} \right),$$

where the covariance matrix \mathbf{P} is given by

$$\begin{aligned}\mathbf{P} &= \sigma^2 \begin{bmatrix} \text{Re}\{\boldsymbol{\Gamma}^\top(e^{i\omega_1 h})\} \\ \text{Im}\{\boldsymbol{\Gamma}^\top(e^{i\omega_1 h})\} \end{bmatrix} \left[\sum_{k=1}^N \boldsymbol{\varphi}(kh) \boldsymbol{\varphi}^\top(kh) \right]^{-1} \begin{bmatrix} \text{Re}\{\boldsymbol{\Gamma}(e^{i\omega_1 h})\} & \text{Im}\{\boldsymbol{\Gamma}(e^{i\omega_1 h})\} \end{bmatrix} \\ &= \frac{2\sigma^2}{N\alpha_1^2 \sin^2(\omega_1 h)} \begin{bmatrix} 1 & \cos(\omega_1 h) \\ 0 & -\sin(\omega_1 h) \end{bmatrix} \begin{bmatrix} 1 & -\cos(\omega_1 h) \\ -\cos(\omega_1 h) & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \cos(\omega_1 h) & -\sin(\omega_1 h) \end{bmatrix} \\ &= \frac{2\sigma^2}{N\alpha_1^2} \mathbf{I}_2.\end{aligned}$$

Thus, the distribution of \hat{a}_1 for finite N is given by the ratio of independent normal variables. Distributions of this type have been previously studied in, e.g., [71, 112]. Hinkley [112] provides the exact distribution for this case. It is in general a heavy-tailed distribution that does not have finite moments [150]; however, it can be approximated by a normal distribution under certain conditions involving the probability of observing negative (resp., positive) values of $\text{Re}\{\hat{G}_N(i\omega_1)\}$ if $\text{Re}\{G^*(i\omega_1)\}$ is positive (resp., negative). This approximation is centered around the quotient of the mean values of the numerator and denominator random variables. A precise statement concerning the normal approximation is given in Theorem 1 of [43], but for the sake of this example we consider a practical result. Using the criterion³ stated in [125], the normal approximation is reasonable whenever the condition

$$\frac{\sigma}{\alpha_1 |\text{Re}\{G^*(i\omega_1)\}|} \sqrt{\frac{2}{N}} \leq 0.1 \quad (10.43)$$

holds, and the normal distribution that approximates the distribution of \hat{a}_1 (see Equation (9) of [43]) is

$$\hat{a}_1 \sim \mathcal{N} \left(-\frac{\text{Im}\{G^*(i\omega_1)\}}{\omega_1 \text{Re}\{G^*(i\omega_1)\}}, \frac{2\sigma^2}{N\alpha_1^2 \omega_1^2 \text{Re}\{G^*(i\omega_1)\}^2} \left[1 + \frac{\text{Im}\{G^*(i\omega_1)\}^2}{\text{Re}\{G^*(i\omega_1)\}^2} \right] \right).$$

Rewriting this distribution in terms of the true parameters a_1^* and b_0^* , we conclude that \hat{a}_1 as a random variable can be approximated by

$$\hat{a}_1 \sim \mathcal{N} \left(a_1^*, \frac{2\sigma^2(1+a_1^{*2}\omega_1^2)^3}{N\alpha_1^2 \omega_1^2 b_0^{*2}} \right). \quad (10.44)$$

Three interesting aspects can be noted from the normal distribution approximation derived above:

1. For the normal approximation to be reasonable, we require a balance between $\text{Re}\{G^*(i\omega_1)\}$ and \sqrt{N} in (10.43). If the frequency of the input signal is high

³Other criteria are possible, but they are usually related to how small the relative standard deviation of the random variable of the denominator is. In this case, the random variable in the denominator is $\omega_1 \text{Re}\{\hat{G}_N(i\omega_1)\}$.

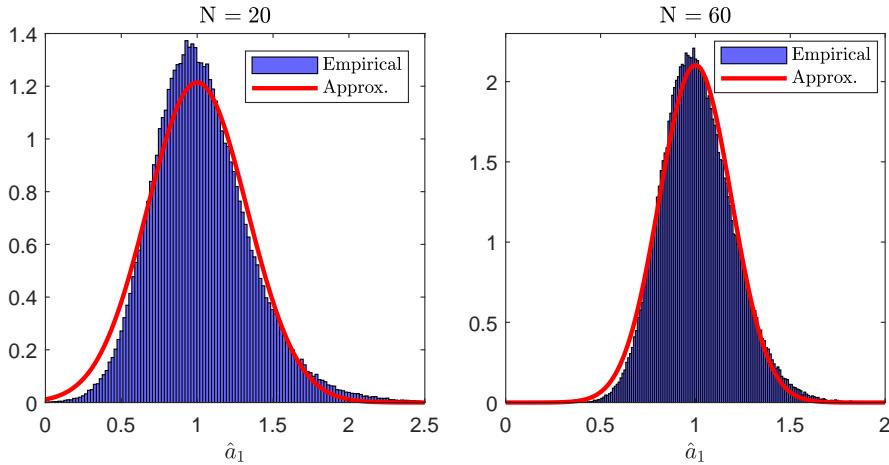


Figure 10.7: Empirical distribution of \hat{a}_1 (histogram) compared to its approximate normal distribution (red curve). Left: $N = 20$; right: $N = 60$.

compared to the bandwidth of the system, then $\text{Re}\{G^*(i\omega_1)\}$ will be small in magnitude and more data points will be required for achieving a reasonable normal approximation. Also, the signal-to-noise ratio α_1^2/σ^2 plays a natural role in (10.43); if such ratio is large, then less data points are needed for the approximation to be sound.

2. The mean of the approximate normal distribution of \hat{a}_1 is the true parameter a_1^* . This property can be observed in finite time, even with less than 100 samples. In Figure 10.7 we compare the empirical distribution (using $2 \cdot 10^5$ Monte Carlo runs) with the normal approximation for two different sample sizes $N = 20$ and $N = 60$. The following parameters are considered: $a_1^* = 1$; $b_0^* = 2$; $\sigma = 0.8$; $\alpha_1 = 1$; $\omega_1 = 1/\sqrt{2}$. Note that the SNR is poor, and very few samples are used for generating the estimate.
3. The angular frequency that minimizes the (approximate) variance of the parameter \hat{a}_1 is $\omega_1 = 1/(\sqrt{2}a_1^*)$, and the resulting approximate variance can be computed as

$$\text{Cov}(\hat{a}_1) \approx \frac{27\sigma^2 a_1^{*2}}{2N\alpha_1^2 b_0^{*2}}.$$

The approximate variance in (10.44) coincides with the asymptotic variance obtained through standard methods [138] (see Appendix 10.B for the details). However, our expression was obtained through an analysis with finite sample size, leading to greater insight about when the estimates can be considered to have reached their asymptotic distribution.

Finally, an unanswered question that arises from the analysis presented in this subsection concerns the behavior of the maximum likelihood estimator when the order of excitation of the multisine input is strictly greater than the number of parameters to be estimated in the parametric model. Bear in mind that it has been shown in Corollary 10.3 that, for not more frequency lines than parameters, the parametric estimator $G(i\omega, \hat{\theta}_N)$ is unbiased at the frequencies present in the input. We conjecture that this result does not hold when there are more frequency lines in the input than parameters. An illustrative example is presented in the simulations subsection below.

10.5.3 Simulations

To conclude this chapter, we will provide empirical evidence for the theoretical findings presented throughout Section 10.5. The system $G_1^*(p)$ in (10.19) is excited by the continuous-time signal $u(t) = 1 + \sin(2t)$, and the noiseless output is computed analytically by assuming that it is in a stationary regime, i.e.,

$$x(t) = G_1^*(0) + |G_1^*(2i)| \sin(2t + \angle G_1^*(2i)).$$

This output signal is sampled every $h = 0.3[\text{s}]$ and is corrupted by additive white Gaussian noise of variance 0.5. Two noise realizations are generated to form two sets of one hundred input and output samples. These realizations are used for estimating several non-parametric responses as shown in Figure 10.8, where we have plotted the true frequency response, two realizations of the non-causal frequency response estimate with $M_{nc} = 1$, two realizations of the causal frequency response estimate with $M_{nc} = 0$, and two realizations of the strictly causal frequency response estimate with $M_{nc} = -1$. For both noise realizations, the estimates at $\omega = 0$ and $\omega = 2[\text{rad/s}]$ are independent of the causality of the proposed non-parametric estimator. This effect, observed by the intersection of the curves in these frequencies, is correctly predicted by Lemma 10.1. The behavior of the estimates between the excitation frequencies depends on the geometric interpolation that is implicitly fixed by the number of non-causal terms M_{nc} .

Next, we study the parametric estimator that can be obtained via rational function interpolation of the frequency estimates $\hat{G}_N(0)$ and $\hat{G}_N(2i)$ (that is, we obtain the maximum likelihood estimate $\hat{\theta}_N$ by solving (10.42)). We use 100000 Monte Carlo runs for the sample size $N = 10$ (which represents approximately one cycle of the input signal) and $\sigma^2 = 0.2$. We plot the sample mean of the frequency response and its variance using an increasing number of realizations in Figure 10.9, and in Figure 10.10 we show the sample mean of the parameter estimates. Even for the very low number of samples used in this simulation study, Figure 10.9 suggests that the frequency response estimates are unbiased and their variances match the theoretical values computed via Theorem 10.4. On the other hand, the sample mean of the parameter estimates in Figure 10.10 indicate that there is a small bias in \hat{a}_1 and \hat{a}_2 but not in \hat{b}_0 , which concurs with the insights developed in Section 10.5.2.

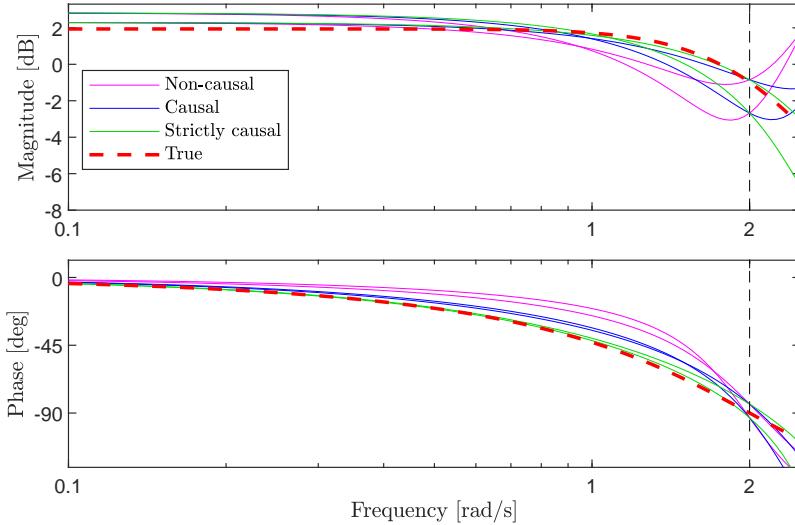


Figure 10.8: In magenta, two non-causal frequency response estimates. In blue, two causal frequency response estimates. In green, two strictly causal frequency response estimates. In dashed red, the true system. The dashed black vertical line indicates the frequency of the sinusoid in the input.

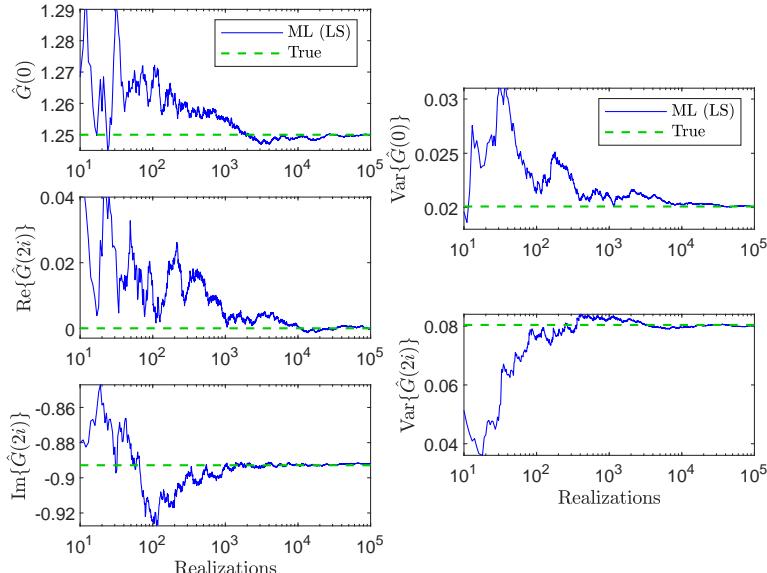


Figure 10.9: In blue, the sample mean of the estimates $\hat{G}(0)$, $\text{Re}\{\hat{G}(2i)\}$ and $\text{Im}\{\hat{G}(2i)\}$ obtained via least-squares (left), and their sample variance (right) for an increasing number of realizations with a fixed sample size $N = 10$. In dashed green, the true values.

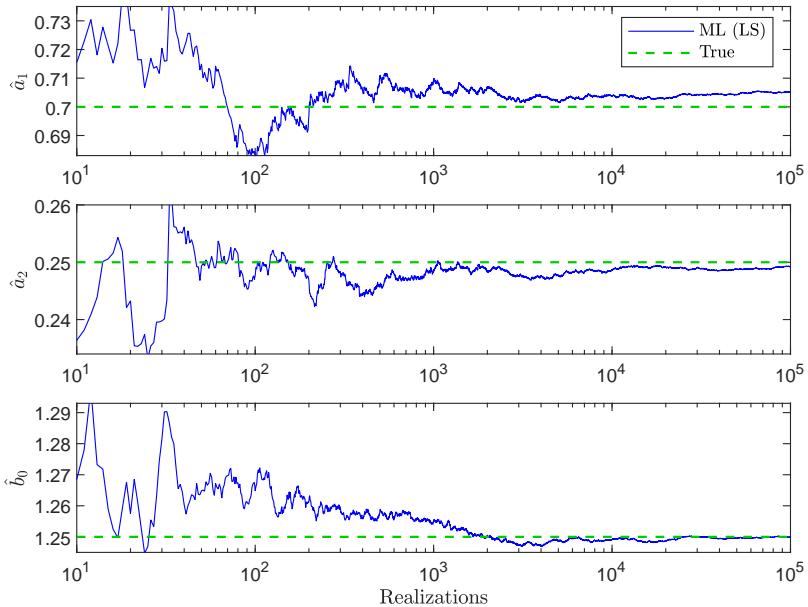


Figure 10.10: In blue, the sample mean of the parameter estimates \hat{a}_1 , \hat{a}_2 and \hat{b}_0 obtained via rational function interpolation, for an increasing number of realizations with a fixed sample size $N = 10$. In dashed green, the true values.

To conclude this section, we perform Monte Carlo runs for when the input is given by $u(t) = 1 + \sin(t) + \sin(2t)$. Contrary to the first case, $\hat{\theta}_N$ does not admit a straightforward expression as the maximum likelihood estimator obtained via (10.40) does not have a closed form. Thus, we use the extended SRIVC method of Chapter 6 to compute it instead. The maximum number of iterations of the extended SRIVC method is set to 500, and the tolerance factor ϵ is fixed at 10^{-15} . This time $N = 100$ samples are used for identification with noise variance $\sigma^2 = 0.2$, and 100000 Monte Carlo runs are performed to compute the empirical means of the frequency and parameter estimates.

Figures 10.11 and 10.12 contain the sample means of the frequency response values $\hat{G}_N(0, \hat{\theta}_N)$, $\text{Re}\{\hat{G}_N(2i, \hat{\theta}_N)\}$ and $\text{Im}\{\hat{G}_N(2i, \hat{\theta}_N)\}$ and the parameter estimates for an increasing number of realizations. Figure 10.11 suggests that the frequency response estimates that arise from evaluating the maximum likelihood estimator are biased in this case, thereby providing evidence that non-asymptotic unbiasedness in the frequency domain is lost when converting the non-parametric estimator into a parametric one via (10.40). The parameter estimates shown in Figure 10.12 are close to their true values on average even for this relatively small sample size. This shows that, similar to Example 10.1, asymptotic properties of the maximum likelihood estimator for continuous-time systems and inputs (and therefore of the extended SRIVC estimator) can be observed with very few samples.

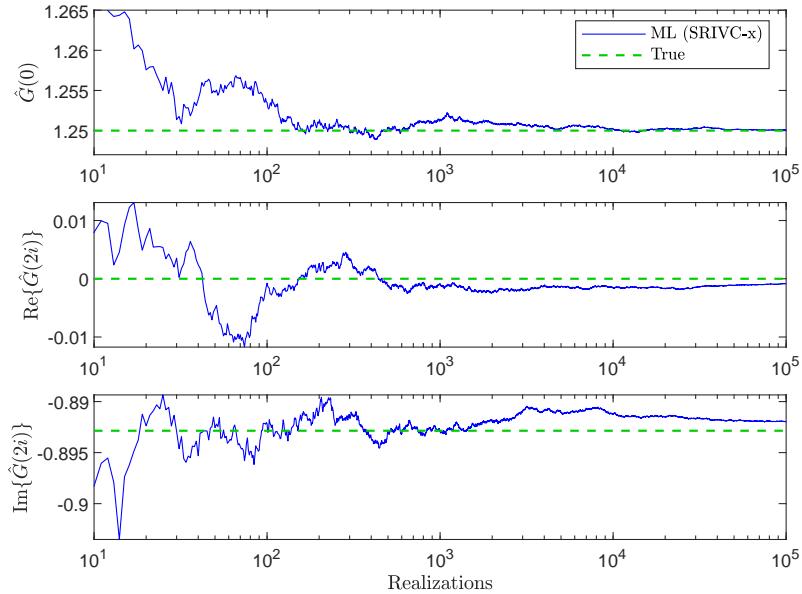


Figure 10.11: In blue, the sample mean of $\hat{G}(0)$, $\text{Re}\{\hat{G}(2i)\}$ and $\text{Im}\{\hat{G}(2i)\}$ obtained by evaluating the maximum likelihood estimate $\hat{G}(p, \hat{\theta}_N)$ for an increasing number of realizations with fixed sample size $N = 100$. In dashed green, the true values.

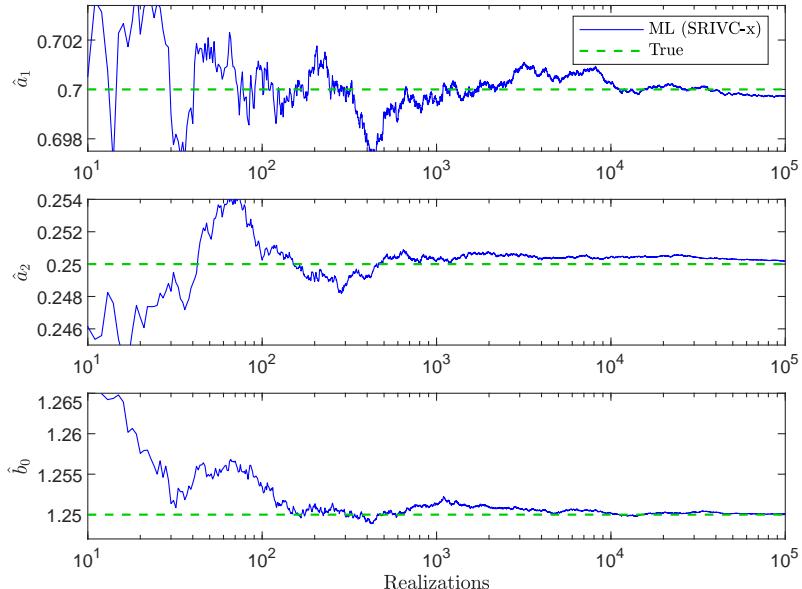


Figure 10.12: In blue, the sample mean of the parameter estimates \hat{a}_1 , \hat{a}_2 and \hat{b}_0 obtained by SRIVC-x, for an increasing number of realizations with fixed sample size $N = 100$. In dashed green, the true values.

10.6 Conclusions

In this chapter we have studied methods for estimating the frequency response of a continuous-time system based on band-limited input excitations. The observation that the equivalent discrete-time system for band-limited inputs is non-causal led to studies on how to estimate the equivalent impulse response and frequency response based on least-squares. For general band-limited inputs, we analyzed the properties of the non-causal least-squares estimator and proposed a kernel-based method for estimating the band-limited equivalent impulse response, and with this the frequency response of the continuous-time system. This non-parametric estimator was later leveraged to design parametric estimators for the continuous-time system using an instrumental variable approach in the frequency domain, for both output error and hybrid Box-Jenkins model structures.

A specialized treatment was given to multisines as they are the only band-limited signals that have a discrete spectrum. By estimating an impulse response via least-squares, we first described the finite-time statistics of the frequency response estimate at the input frequencies and studied a non-parametric estimator for this frequency response that results from trigonometric polynomial interpolation. Later, with the theoretical findings previously obtained, we showed that the prediction error method, optimal refined instrumental variables for multisine inputs, and the maximum likelihood estimator of any parametric model, are directly linked with the interpolation and approximation of frequency response functions. This insight has implications on how parametric estimators can be computed and sheds light on some finite-sample properties.

Extensive simulations have shown the applicability of non-causal regularized least-squares and the parametric estimators developed for estimation using band-limited signals. Simulation tests have also been useful to verify the theoretical results concerning continuous-time system identification using continuous-time multisines as inputs.

Appendix

10.A Proofs of technical results

10.A.1 Proof of Theorem 10.1

By using (10.3), we can write the least-squares estimate as

$$\hat{\rho}_N = \rho^* + \left[\sum_{k=1}^N \varphi(kh) \varphi^\top(kh) \right]^{-1} \left[\sum_{k=1}^N \varphi(kh) w(kh) \right],$$

where $w(kh)$ is given by

$$w(kh) = h \sum_{\{n : n < -M_{nc} \text{ or } n > M_c\}} u([k-n]h) g_{BL}^*(nh) + v(kh).$$

The ergodic lemma in [205, Lemma 3.1] and the continuous-mapping theorem [230, Thm. 2.3] permit us to write, as the sample size tends to infinity,

$$\hat{\rho}_N \xrightarrow{a.s.} \rho^* + \mathbb{E}\{\varphi(kh) \varphi^\top(kh)\}^{-1} \mathbb{E}\{\varphi(kh) w(kh)\}, \quad (10.45)$$

where the matrix being inverted is a positive-definite scalar matrix. Since the system in (10.1) is asymptotically stable, the band-limited impulse response has finite 2-norm⁴ by Corollary 9.2. Thus, thanks to [205, Lemma 3.1] and the fact that $u(t)$ is white noise at the sampling instants, the following expectation can be computed:

$$\mathbb{E}\{\varphi(kh) w(kh)\} = h^2 \sum_{\{n : n < -M_{nc} \text{ or } n > M_c\}} \begin{bmatrix} \mathbb{E}\{u([k-n]h)u([k+M_{nc}]h)\}g_{BL}^*(nh) \\ \mathbb{E}\{u([k-n]h)u([k-1+M_{nc}]h)\}g_{BL}^*(nh) \\ \vdots \\ \mathbb{E}\{u([k-n]h)u([k-M_c]h)\}g_{BL}^*(nh) \end{bmatrix} = \mathbf{0}.$$

This result, together with (10.45), leads to the desired conclusion. \square

⁴It can be shown that the 2-norm of the band-limited impulse response is also finite when the continuous-time system is biproper.

10.A.2 Proof of Theorem 10.2

We first write the parameter error as

$$\sqrt{N}(\hat{\rho}_N - \rho^*) = \left[\frac{1}{N} \sum_{k=1}^N \varphi(kh) \varphi^\top(kh) \right]^{-1} \left[\frac{1}{\sqrt{N}} \sum_{k=1}^N \varphi(kh) w(kh) \right].$$

Since the first sum converges to its expected value for large N , we can write

$$\begin{aligned} \sqrt{N}(\hat{\rho}_N - \rho^*) &= \mathbb{E}\{\varphi(kh) \varphi^\top(kh)\}^{-1} \left[\frac{1}{\sqrt{N}} \sum_{k=1}^N \varphi(kh) w(kh) \right] + o_p(1) \\ &= \frac{1}{h^2 \lambda^2} \left[\frac{1}{\sqrt{N}} \sum_{k=1}^N \varphi(kh) w(kh) \right] + o_p(1). \end{aligned}$$

By Lemma A4.1 of [210],

$$\frac{1}{\sqrt{N}} \sum_{k=1}^N \varphi(kh) w(kh) \xrightarrow{\text{dist.}} \mathcal{N}(\mathbf{0}, \mathbf{P}),$$

where \mathbf{P} is given by

$$\mathbf{P} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N \sum_{n=1}^N \mathbb{E}\{\varphi(kh) w(kh) w(nh) \varphi^\top(nh)\}.$$

The asymptotic covariance Σ in (10.11) is obtained by applying Lemma A4.2 of [210] and its corollary, concluding the proof. \square

10.A.3 Proof of Lemma 10.1

Thanks to the decomposition of $\varphi(kh)$ in (10.34), we can write the least-squares estimator as

$$\begin{aligned} \hat{\rho}_N &= \left[\sum_{k=1}^N \varphi(kh) \varphi^\top(kh) \right]^{-1} \left[\sum_{k=1}^N \varphi(kh) y(kh) \right] \\ &= \left[\tilde{\Gamma} \sum_{k=1}^N \zeta(kh) \zeta^\top(kh) \tilde{\Gamma}^\top \right]^{-1} \left[\tilde{\Gamma} \sum_{k=1}^N \zeta(kh) y(kh) \right] \\ &= \tilde{\Gamma}^{-\top} \left[\sum_{k=1}^N \zeta(kh) \zeta^\top(kh) \right]^{-1} \left[\sum_{k=1}^N \zeta(kh) y(kh) \right]. \end{aligned}$$

Multiplying by $\tilde{\Gamma}^\top$ on both sides from the left leads to the desired result, namely, that $\tilde{\Gamma}^\top \hat{\rho}_N$ does not depend on M_{nc} . \square

10.A.4 Proof of Theorem 10.3

We start by showing that the least-squares estimate is well-defined for $N \geq 2M + 1$. This will be the case if the matrix being inverted in (10.6) is positive definite. In other words, we need to show that

$$\mathbf{x}^\top \sum_{k=1}^N \boldsymbol{\varphi}(kh) \boldsymbol{\varphi}^\top(kh) \mathbf{x} = 0 \implies \mathbf{x} = \mathbf{0}.$$

The first equality above holds if and only if $|\boldsymbol{\varphi}^\top(kh)\mathbf{x}|^2 = 0$ for all $k = 1, 2, \dots, N$, which is equivalent to $\boldsymbol{\varphi}^\top(kh)\mathbf{x} = 0$ for $k = 1, 2, \dots, N$. Using the decomposition $\boldsymbol{\varphi}(kh) = \tilde{\boldsymbol{\Gamma}}\boldsymbol{\zeta}(kh)$ with $\boldsymbol{\zeta}(kh)$ being defined as in (10.35), we see that $\boldsymbol{\zeta}^\top(kh)\tilde{\mathbf{x}} = 0$ for $k = 1, 2, \dots, N$, where $\tilde{\mathbf{x}} = \tilde{\boldsymbol{\Gamma}}^\top \mathbf{x}$. In other words,

$$\begin{bmatrix} \boldsymbol{\zeta}^\top(h) \\ \boldsymbol{\zeta}^\top(2h) \\ \vdots \\ \boldsymbol{\zeta}^\top([2M+1]h) \end{bmatrix} \tilde{\mathbf{x}} = \mathbf{0}. \quad (10.46)$$

Similar to what was done in (10.33), the matrix on the left-hand side of (10.46) can be decomposed as a non-singular Vandermonde matrix multiplied by a diagonal matrix, and therefore it is non-singular. This fact implies $\tilde{\mathbf{x}} = \mathbf{0}$, and therefore $\mathbf{x} = \mathbf{0}$. Hence, the least-squares estimate in (10.6) is well-defined in this case.

The expected value of the least-squares estimator is given by

$$\mathbb{E}\{\hat{\mathbf{p}}_N\} = \left[\frac{1}{N} \sum_{k=1}^N \boldsymbol{\varphi}(kh) \boldsymbol{\varphi}^\top(kh) \right]^{-1} \left[\frac{1}{N} \sum_{k=1}^N \boldsymbol{\varphi}(kh) x(kh) \right],$$

where we have used the assumption that the measurement noise $v(kh)$ has zero mean. Since the normal matrix is non-singular, the solution of the set of equations in $\boldsymbol{\rho}$

$$\sum_{k=1}^N \boldsymbol{\varphi}(kh) [x(kh) - \boldsymbol{\varphi}^\top(kh)\boldsymbol{\rho}] = \mathbf{0} \quad (10.47)$$

is unique and equal to $\mathbb{E}\{\hat{\mathbf{p}}_N\}$. After standard computations, we can write $\boldsymbol{\varphi}^\top(kh)$ as

$$\boldsymbol{\varphi}^\top(kh) = \frac{1}{2} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_M \end{bmatrix}^\top \left(\begin{bmatrix} \boldsymbol{\Gamma}^\top(1) \\ e^{i(\omega_1 kh + \psi_1)} \boldsymbol{\Gamma}^\top(e^{i\omega_1 h}) \\ \vdots \\ e^{i(\omega_M kh + \psi_M)} \boldsymbol{\Gamma}^\top(e^{i\omega_M h}) \end{bmatrix} + \begin{bmatrix} \boldsymbol{\Gamma}^\top(1) \\ e^{-i(\omega_1 kh + \psi_1)} \boldsymbol{\Gamma}^\top(e^{-i\omega_1 h}) \\ \vdots \\ e^{-i(\omega_M kh + \psi_M)} \boldsymbol{\Gamma}^\top(e^{-i\omega_M h}) \end{bmatrix} \right).$$

Also,

$$x(kh) = \frac{1}{2} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_M \end{bmatrix}^\top \left(\begin{bmatrix} G^*(0) \\ G^*(i\omega_1)e^{i(\omega_1 kh + \psi_1)} \\ \vdots \\ G^*(i\omega_M)e^{i(\omega_M kh + \psi_M)} \end{bmatrix} + \begin{bmatrix} G^*(0) \\ G^*(-i\omega_1)e^{-i(\omega_1 kh + \psi_1)} \\ \vdots \\ G^*(-i\omega_M)e^{-i(\omega_M kh + \psi_M)} \end{bmatrix} \right),$$

which leads to

$$x(kh) - \boldsymbol{\varphi}^\top(kh)\boldsymbol{\rho} = \frac{1}{2} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_M \end{bmatrix}^\top \left(\begin{bmatrix} G^*(0) - \boldsymbol{\Gamma}^\top(1)\boldsymbol{\rho} \\ e^{i(\omega_1 kh + \psi_1)}[G^*(i\omega_1) - \boldsymbol{\Gamma}^\top(e^{i\omega_1 h})\boldsymbol{\rho}] \\ \vdots \\ e^{i(\omega_M kh + \psi_M)}[G^*(i\omega_M) - \boldsymbol{\Gamma}^\top(e^{i\omega_M h})\boldsymbol{\rho}] \end{bmatrix} + \begin{bmatrix} G^*(0) - \boldsymbol{\Gamma}^\top(1)\boldsymbol{\rho} \\ e^{-i(\omega_1 kh + \psi_1)}[G^*(-i\omega_1) - \boldsymbol{\Gamma}^\top(e^{-i\omega_1 h})\boldsymbol{\rho}] \\ \vdots \\ e^{-i(\omega_M kh + \psi_M)}[G^*(-i\omega_M) - \boldsymbol{\Gamma}^\top(e^{-i\omega_M h})\boldsymbol{\rho}] \end{bmatrix} \right).$$

A sufficient condition for (10.47) to hold is that $x(kh) - \boldsymbol{\varphi}^\top(kh)\boldsymbol{\rho} = 0$ for all integers k . From the equation above, we see that this is achieved for a vector $\boldsymbol{\rho}$ that satisfies

$$\tilde{\boldsymbol{\Gamma}}^\top \boldsymbol{\rho} = \left[G^*(0), G^*(i\omega_1), G^*(-i\omega_1), \dots, G^*(i\omega_M), G^*(-i\omega_M) \right]^\top. \quad (10.48)$$

Since $\tilde{\boldsymbol{\Gamma}}$ is non-singular, (10.48) has a unique solution. This means that $\mathbb{E}\{\hat{\boldsymbol{\rho}}_N\}$ must satisfy (10.48), which implies that the trigonometric polynomial $\mathbb{E}\{\hat{G}_N(i\omega)\} = \boldsymbol{\Gamma}^\top(e^{i\omega h})\mathbb{E}\{\hat{\boldsymbol{\rho}}_N\}$ satisfies $\mathbb{E}\{\hat{G}_N(0)\} = G^*(0)$, and $\mathbb{E}\{\hat{G}_N(\pm i\omega_l)\} = G^*(\pm i\omega_l)$ for $l = 1, 2, \dots, M$. \square

10.A.5 Proof of Theorem 10.4

We can write

$$\begin{aligned} & \begin{bmatrix} \hat{G}_N(0) - G^*(0) \\ \hat{G}_N(i\omega_1) - G^*(i\omega_1) \\ \hat{G}_N(-i\omega_1) - G^*(-i\omega_1) \\ \vdots \\ \hat{G}_N(i\omega_M) - G^*(i\omega_M) \\ \hat{G}_N(-i\omega_M) - G^*(-i\omega_M) \end{bmatrix} = \tilde{\boldsymbol{\Gamma}}^\top (\hat{\boldsymbol{\rho}}_N - \mathbb{E}\{\hat{\boldsymbol{\rho}}_N\}) \\ &= \tilde{\boldsymbol{\Gamma}}^\top \left[\sum_{k=1}^N \boldsymbol{\varphi}(kh) \boldsymbol{\varphi}^\top(kh) \right]^{-1} \left[\sum_{k=1}^N \boldsymbol{\varphi}(kh) (y(kh) - \boldsymbol{\varphi}^\top(kh) \mathbb{E}\{\hat{\boldsymbol{\rho}}_N\}) \right]. \end{aligned}$$

Following the same lines as in the proof of Theorem 10.3, we can show that $x(kh) - \boldsymbol{\varphi}^\top(kh)\mathbb{E}\{\hat{\boldsymbol{\rho}}_N\}$ is equal to zero for any integer k . Thus,

$$\tilde{\boldsymbol{\Gamma}}^\top (\hat{\boldsymbol{\rho}}_N - \mathbb{E}\{\hat{\boldsymbol{\rho}}_N\}) = \tilde{\boldsymbol{\Gamma}}^\top \left[\sum_{k=1}^N \boldsymbol{\varphi}(kh) \boldsymbol{\varphi}^\top(kh) \right]^{-1} \left[\sum_{k=1}^N \boldsymbol{\varphi}(kh) v(kh) \right].$$

Therefore, the covariance matrix of interest is given by

$$\begin{aligned}
& \mathbb{E} \left\{ \tilde{\Gamma}^\top (\hat{\rho}_N - \mathbb{E}\{\hat{\rho}_N\}) (\hat{\rho}_N - \mathbb{E}\{\hat{\rho}_N\})^\top \tilde{\Gamma} \right\} \\
&= \tilde{\Gamma}^\top \left[\sum_{k=1}^N \varphi(kh) \varphi^\top(kh) \right]^{-1} \left[\sum_{k=1}^N \sum_{n=1}^N \varphi(kh) \mathbb{E}\{v(kh)v(nh)\} \varphi^\top(nh) \right] \left[\sum_{k=1}^N \varphi(kh) \varphi^\top(kh) \right]^{-1} \tilde{\Gamma} \\
&= \sigma^2 \tilde{\Gamma}^\top \left[\sum_{k=1}^N \varphi(kh) \varphi^\top(kh) \right]^{-1} \left[\sum_{k=1}^N \varphi(kh) \varphi^\top(kh) \right] \left[\sum_{k=1}^N \varphi(kh) \varphi^\top(kh) \right]^{-1} \tilde{\Gamma} \\
&= \sigma^2 \tilde{\Gamma}^\top \left[\sum_{k=1}^N \varphi(kh) \varphi^\top(kh) \right]^{-1} \tilde{\Gamma}.
\end{aligned}$$

The second equality in (10.36) follows from noting that $\varphi(kh) = \overline{\tilde{\Gamma} \zeta(kh)}$ and $\varphi^\top(kh) = \zeta^\top(kh) \tilde{\Gamma}^\top$.

Now, assume that Nh is a multiple of the least common multiple of $\{2\pi/\omega_l\}_{l=1}^M$. We are interested in computing the non-diagonal elements of the inverse of the covariance matrix of the frequency estimates. An arbitrary non-diagonal element of the first row or column of this matrix will be of the form

$$\frac{\alpha_0 \alpha_l e^{i\psi_l}}{2\sigma^2} \sum_{k=1}^N e^{i\frac{2\pi km_l}{N}} \quad \text{or} \quad \frac{\alpha_0 \alpha_l e^{-i\psi_l}}{2\sigma^2} \sum_{k=1}^N e^{-i\frac{2\pi km_l}{N}}, \quad (10.49)$$

where $m_l = Nh\omega_l/(2\pi)$ is an positive integer less than $N/2$. Both cases can be written as the sum of the m_l -th power of the N -th roots of unity, which is known to be equal to zero for when m_l is not a multiple of N (see, e.g., Example 3 of [129, Ch. 3]). Similarly, an arbitrary non-diagonal element of the remaining rows and columns will have the form

$$\frac{\alpha_{l_1} \alpha_{l_2} e^{i(\psi_{l_1} + \psi_{l_2})}}{4\sigma^2} \sum_{k=1}^N e^{i(\omega_{l_1} + \omega_{l_2})kh} \quad \text{or} \quad \frac{\alpha_{l_1} \alpha_{l_2} e^{-i(\psi_{l_1} + \psi_{l_2})}}{4\sigma^2} \sum_{k=1}^N e^{-i(\omega_{l_1} + \omega_{l_2})kh}$$

for some $l_1, l_2 \in \{1, 2, \dots, M\}$, or

$$\frac{\alpha_{l_1} \alpha_{l_2} e^{i(\psi_{l_1} - \psi_{l_2})}}{4\sigma^2} \sum_{k=1}^N e^{i(\omega_{l_1} - \omega_{l_2})kh}$$

for some $l_1, l_2 \in \{1, 2, \dots, M\}$, $l_1 \neq l_2$. In all of these cases, the sums will also be zero by the same reasoning detailed above for (10.49). With this, we conclude that if Nh is a multiple of the least common multiple of $\{2\pi/\omega_l\}_{l=1}^M$ the matrix being inverted in the right-hand side of (10.36) is diagonal, which leads to the desired result. \square

10.A.6 Proof of Corollary 10.1

Following similar lines to the proof of Theorem 10.4, we find that

$$\text{AsCov} \left\{ \begin{bmatrix} \hat{G}_N(0) \\ \hat{G}_N(i\omega_1) \\ \hat{G}_N(-i\omega_1) \\ \vdots \\ \hat{G}_N(i\omega_M) \\ \hat{G}_N(-i\omega_M) \end{bmatrix} \right\} = \sigma^2 \mathbf{P}^{-1},$$

where

$$\mathbf{P} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N \mathbb{E} \left\{ \overline{\zeta(kh)} \zeta^\top(kh) \right\}.$$

We now compute the entries of \mathbf{P} . For \mathbf{P}_{11} , we directly have $\mathbf{P}_{11} = \alpha_0^2$. Similarly, for $l \in \{1, 2, \dots, M\}$, we have $\mathbf{P}_{(2l)(2l)} = \alpha_l^2/4$ and $\mathbf{P}_{(2l+1)(2l+1)} = \alpha_l^2/4$. Lastly, every element that is not in the diagonal will be of the form

$$C \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N e^{i(\tilde{\omega}kh + \tilde{\psi})}$$

for some real constants $C, \tilde{\psi}$. The constant $\tilde{\omega}$, which is the difference or addition between distinct frequencies ω_l , will satisfy $0 < |\tilde{\omega}| < 2\pi/h$ by the band-limited sampling condition.

For any $N > 0$, the following chain of inequalities is satisfied:

$$0 \leq \left| C \frac{1}{N} \sum_{k=1}^N e^{i(\tilde{\omega}kh + \tilde{\psi})} \right| = \frac{1}{N} \left| \frac{C \sin([N+1]\tilde{\omega}h/2)}{\sin(\tilde{\omega}h/2)} \right| \leq \frac{1}{N} \left| \frac{C}{\sin(\tilde{\omega}h/2)} \right|.$$

Since $0 < |\tilde{\omega}| < 2\pi/h$, the denominator of the upper bound must go to infinity. Thus, by the squeeze theorem [213, Thm. 3.3.6], we conclude that

$$C \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N e^{i(\tilde{\omega}kh + \tilde{\psi})} = 0,$$

which implies that all the elements outside the diagonal of \mathbf{P} are zero, concluding the proof. \square

10.A.7 Proof of Theorem 10.5

To resemble the notation of [28], in this proof we will adopt the notation $\hat{\mathbf{G}}_N(\mathbf{y})$ to refer to \hat{G}_N . The idea is to obtain the probability density functions $p(\mathbf{y}|\boldsymbol{\theta})$ and $q(\hat{\mathbf{G}}_N(\mathbf{y})|\boldsymbol{\theta})$ and later conclude via the Factorization Theorem [28, Thm. 6.2.6].

From (10.2), we see that the output samples $\{y(kh)\}_{k=1}^N$ are jointly Gaussian-distributed:

$$p(\mathbf{y}|\boldsymbol{\theta}) = (2\pi\sigma^2)^{-N/2} \exp \left\{ \frac{-1}{2\sigma^2} \sum_{k=1}^N [y(kh) - x(kh, \boldsymbol{\theta})]^2 \right\},$$

where the dependence on $\boldsymbol{\theta}$ of the mean value of the output $x(kh, \boldsymbol{\theta})$ has been made explicit. On the other hand, following the same reasoning as in Theorems 10.3 and 10.4 and under the Gaussianity condition on the output noise, we find that $\hat{\mathbf{G}}_N^r(\mathbf{y})$ is also jointly Gaussian-distributed with mean equal to $\mathbf{G}^r(\boldsymbol{\theta})$ (previously defined in (10.39)) and covariance

$$\text{Cov}(\hat{\mathbf{G}}_N^r) = \sigma^2 \tilde{\boldsymbol{\Gamma}}_r^\top \left[\sum_{k=1}^N \boldsymbol{\varphi}(kh) \boldsymbol{\varphi}^\top(kh) \right]^{-1} \tilde{\boldsymbol{\Gamma}}_r, \quad (10.50)$$

where

$$\tilde{\boldsymbol{\Gamma}}_r = h \begin{bmatrix} 1 & 1 & 0 & \dots & 1 & 0 \\ 1 & \cos(\omega_1 h) & -\sin(\omega_1 h) & \dots & \cos(\omega_M h) & -\sin(\omega_M h) \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 1 & \cos(2M\omega_1 h) & -\sin(2M\omega_1 h) & \dots & \cos(2M\omega_M h) & -\sin(2M\omega_M h) \end{bmatrix}.$$

In other words, the probability density function of $\hat{\mathbf{G}}_N^r(\mathbf{y})$ is given by

$$\begin{aligned} q(\hat{\mathbf{G}}_N^r(\mathbf{y})|\boldsymbol{\theta}) &= \frac{1}{\sqrt{\det[2\pi\text{Cov}(\hat{\mathbf{G}}_N^r)]}} \\ &\times \exp \left\{ -\frac{1}{2} (\hat{\mathbf{G}}_N^r(\mathbf{y}) - \mathbf{G}^r(\boldsymbol{\theta}))^\top [\text{Cov}(\hat{\mathbf{G}}_N^r)]^{-1} (\hat{\mathbf{G}}_N^r(\mathbf{y}) - \mathbf{G}^r(\boldsymbol{\theta})) \right\}. \end{aligned} \quad (10.51)$$

The next step is to rewrite the exponent in (10.51). By leveraging (10.50) and the fact that $\hat{\boldsymbol{\theta}}_N = \tilde{\boldsymbol{\Gamma}}_r^{-\top} \hat{\mathbf{G}}_N^r(\mathbf{y})$ where $\hat{\boldsymbol{\theta}}_N$ is given by (10.6), we obtain

$$[\text{Cov}(\hat{\mathbf{G}}_N^r)]^{-1} \hat{\mathbf{G}}_N^r(\mathbf{y}) = \frac{1}{\sigma^2} \tilde{\boldsymbol{\Gamma}}_r^{-1} \sum_{k=1}^N \boldsymbol{\varphi}(kh) y(kh).$$

On the other hand, we also have

$$[\text{Cov}(\hat{\mathbf{G}}_N^r)]^{-1} \mathbf{G}^r(\boldsymbol{\theta}) = \frac{1}{\sigma^2} \tilde{\boldsymbol{\Gamma}}_r^{-1} \sum_{k=1}^N \boldsymbol{\varphi}(kh) \boldsymbol{\varphi}^\top(kh) \tilde{\boldsymbol{\Gamma}}_r^{-\top} \mathbf{G}^r(\boldsymbol{\theta}).$$

By noting that we can decompose $\varphi(kh)$ as

$$\varphi(kh) = \tilde{\Gamma}_r \begin{bmatrix} \alpha_0 \\ \alpha_1 \cos(\omega_1 kh + \psi_1) \\ -\alpha_1 \sin(\omega_1 kh + \psi_1) \\ \vdots \\ \alpha_M \cos(\omega_M kh + \psi_M) \\ -\alpha_M \sin(\omega_M kh + \psi_M) \end{bmatrix},$$

we find that

$$\begin{aligned} & \varphi^\top(kh) \tilde{\Gamma}_r^{-\top} \mathbf{G}^r(\boldsymbol{\theta}) \\ &= \alpha_0 G(0, \boldsymbol{\theta}) + \sum_{l=1}^M \alpha_l (\cos(\omega_l kh + \psi_l) \operatorname{Re}\{G(i\omega_l, \boldsymbol{\theta})\} - \sin(\omega_l kh + \psi_l) \operatorname{Im}\{G(i\omega_l, \boldsymbol{\theta})\}) \\ &= \alpha_0 G(0, \boldsymbol{\theta}) + \sum_{l=1}^M \alpha_l |G(i\omega_l, \boldsymbol{\theta})| \cos(\omega_l kh + \psi_l + \angle G(i\omega_l, \boldsymbol{\theta})). \end{aligned} \quad (10.52)$$

In other words, we see that $\varphi^\top(kh) \tilde{\Gamma}_r^{-\top} \mathbf{G}^r(\boldsymbol{\theta})$ is simply the mean value $x(kh, \boldsymbol{\theta})$. With these partial results, we compute

$$\begin{aligned} & \left(\hat{\mathbf{G}}_N^r(\mathbf{y}) - \mathbf{G}^r(\boldsymbol{\theta}) \right)^\top \left[\operatorname{Cov}(\hat{\mathbf{G}}_N^r) \right]^{-1} \left(\hat{\mathbf{G}}_N^r(\mathbf{y}) - \mathbf{G}^r(\boldsymbol{\theta}) \right) \\ &= -\frac{1}{\sigma^2} \left(\hat{\mathbf{G}}_N^r(\mathbf{y}) \right)^\top \tilde{\Gamma}_r^{-1} \sum_{k=1}^N \varphi(kh) \varphi^\top(kh) \tilde{\Gamma}_r^{-\top} \mathbf{G}^r(\boldsymbol{\theta}) + \frac{1}{\sigma^2} \left(\hat{\mathbf{G}}_N^r(\mathbf{y}) \right)^\top \tilde{\Gamma}_r^{-1} \sum_{k=1}^N \varphi(kh) y(kh) \\ &+ \frac{1}{\sigma^2} \sum_{k=1}^N (x^2(kh, \boldsymbol{\theta}) - y(kh)x(kh, \boldsymbol{\theta})). \end{aligned}$$

The first term in the last equality above can be expressed as

$$\begin{aligned} -\frac{1}{\sigma^2} \left(\hat{\mathbf{G}}_N^r(\mathbf{y}) \right)^\top \tilde{\Gamma}_r^{-1} \sum_{k=1}^N \varphi(kh) \varphi^\top(kh) \tilde{\Gamma}_r^{-\top} \mathbf{G}^r(\boldsymbol{\theta}) &= -\frac{1}{\sigma^2} \hat{\theta}_N^\top \sum_{k=1}^N \varphi(kh) \varphi^\top(kh) \tilde{\Gamma}_r^{-\top} \mathbf{G}^r(\boldsymbol{\theta}) \\ &= -\frac{1}{\sigma^2} \sum_{k=1}^N y(kh) \varphi^\top(kh) \tilde{\Gamma}_r^{-\top} \mathbf{G}^r(\boldsymbol{\theta}) \\ &= -\frac{1}{\sigma^2} \sum_{k=1}^N y(kh) x(kh, \boldsymbol{\theta}). \end{aligned}$$

Hence,

$$\frac{p(\mathbf{y}|\boldsymbol{\theta})}{q(\hat{\mathbf{G}}_N^r(\mathbf{y})|\boldsymbol{\theta})} = \sqrt{\frac{\operatorname{Cov}(\hat{\mathbf{G}}_N^r)}{\sigma^{2N}(2\pi)^{N-2M-1}}} \exp \left\{ \frac{1}{2\sigma^2} \sum_{k=1}^N \left[(\hat{\mathbf{G}}_N^r(\mathbf{y}))^\top \tilde{\Gamma}_r^{-1} \varphi(kh) y(kh) - y^2(kh) \right] \right\}. \quad (10.53)$$

Since the right-hand side of (10.53) does not depend on $\boldsymbol{\theta}$, we conclude by the Factorization Theorem [28] that $\hat{\mathbf{G}}_N^r(\mathbf{y})$ is a sufficient statistic for $\boldsymbol{\theta}$. \square

Remark 10.8. The theorem also follows from a useful result in statistics [130]: if the data can be described as $y(kh) = \boldsymbol{\psi}^\top(kh)\mathbf{f}(\boldsymbol{\theta}) + v(kh)$, where $v(kh)$ is i.i.d. and Gaussian, and \mathbf{f} depends only on $\boldsymbol{\theta}$, then $\boldsymbol{\theta}$ has sufficient statistic

$$\hat{\mathbf{f}}(\mathbf{y}) = \left[\sum_{k=1}^N \boldsymbol{\psi}(kh) \boldsymbol{\psi}^\top(kh) \right]^{-1} \left[\sum_{k=1}^N \boldsymbol{\psi}(kh) y(kh) \right].$$

In this context, (10.52) shows that the output can be written as

$$y(kh) = \boldsymbol{\varphi}^\top(kh) \tilde{\boldsymbol{\Gamma}}_r^{-\top} \mathbf{G}^r(\boldsymbol{\theta}) + v(kh).$$

Exploiting this fact together with the identity $\hat{\boldsymbol{\theta}}_N = \tilde{\boldsymbol{\Gamma}}_r^{-\top} \hat{\mathbf{G}}_N^r(\mathbf{y})$ leads to writing a sufficient statistic for $\boldsymbol{\theta}$ as

$$\begin{aligned} & \left[\sum_{k=1}^N \tilde{\boldsymbol{\Gamma}}_r^{-1} \boldsymbol{\varphi}(kh) \boldsymbol{\varphi}^\top(kh) \tilde{\boldsymbol{\Gamma}}_r^{-\top} \right]^{-1} \left[\sum_{k=1}^N \boldsymbol{\varphi}^\top(kh) \tilde{\boldsymbol{\Gamma}}_r^{-\top} y(kh) \right] \\ &= \tilde{\boldsymbol{\Gamma}}_r^\top \left[\sum_{k=1}^N \boldsymbol{\varphi}(kh) \boldsymbol{\varphi}^\top(kh) \right]^{-1} \left[\sum_{k=1}^N \boldsymbol{\varphi}^\top(kh) y(kh) \right] \\ &= \tilde{\boldsymbol{\Gamma}}_r^\top \hat{\boldsymbol{\theta}}_N \\ &= \hat{\mathbf{G}}_N^r(\mathbf{y}). \end{aligned}$$

This again proves that $\hat{\mathbf{G}}_N^r(\mathbf{y})$ is a sufficient statistic for $\boldsymbol{\theta}$.

10.B Computation of the asymptotic variance of \hat{a}_1

In this section we provide the computations that lead to the variance in (10.44) through standard asymptotic arguments. It is known (see, e.g., [138]) that the asymptotic covariance of $\hat{\boldsymbol{\theta}}$ is given by

$$\text{AsCov}(\hat{\boldsymbol{\theta}}) \approx \sigma^2 \left[\bar{\mathbb{E}} \left\{ \left(\frac{\partial \hat{y}(kh|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*} \right) \left(\frac{\partial \hat{y}(kh|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*} \right)^\top \right\} \right]^{-1},$$

where the predictor $\hat{y}(kh|\boldsymbol{\theta})$ in this case can be written as

$$\hat{y}(kh|\boldsymbol{\theta}) = \alpha_1 |G(i\omega_1, \boldsymbol{\theta})| \sin(\omega_1 t + \angle G(i\omega_1, \boldsymbol{\theta})).$$

We compute the gradient

$$\frac{\partial \hat{y}(kh|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*} = \begin{bmatrix} \frac{-\alpha_1 b_0^* |\omega_1|}{a_1^{*2} \omega_1^2 + 1} \sin(\omega_1 t + \phi_1) \\ \frac{\alpha_1}{\sqrt{a_1^{*2} \omega_1^2 + 1}} \sin(\omega_1 t + \phi_2), \end{bmatrix},$$

where the angles ϕ_1 and ϕ_2 are given by $\pi/2 + 2\angle(1 - ia_1^*\omega_1)$ and $\angle(1 - ia_1^*\omega_1)$, respectively. The identity $\bar{\mathbb{E}} \{\sin(\omega_1 t + \phi_1) \sin(\omega_1 t + \phi_2)\} = \cos(\phi_1 - \phi_2)/2$ yields

$$\begin{aligned} & \bar{\mathbb{E}} \left\{ \left(\frac{\partial \hat{y}(kh|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*} \right) \left(\frac{\partial \hat{y}(kh|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*} \right)^\top \right\}^{-1} \\ &= \frac{4(a_1^{*2}\omega_1^2 + 1)^3}{\alpha_1^4 b_0^{*2} \omega_1^2 \sin^2(\phi_1 - \phi_2)} \begin{bmatrix} \frac{\alpha_1^2}{2(a_1^{*2}\omega_1^2 + 1)} & \frac{\alpha_1^2 b_0^* |\omega_1|}{2(a_1^{*2}\omega_1^2 + 1)^{3/2}} \cos(\phi_1 - \phi_2) \\ \frac{\alpha_1^2 b_0^* |\omega_1|}{2(a_1^{*2}\omega_1^2 + 1)^{3/2}} \cos(\phi_1 - \phi_2) & \frac{\alpha_1^2 b_0^{*2} \omega_1^2}{2(a_1^{*2}\omega_1^2 + 1)^2} \end{bmatrix}, \end{aligned}$$

where $\sin^2(\phi_1 - \phi_2) = \cos^2(\angle(1 - ia_1^*\omega_1)) = (a_1^{*2}\omega_1^2 + 1)^{-1}$. Thus, we conclude that

$$\text{AsCov}(\hat{a}_1) = \frac{2\sigma^2(a_1^{*2}\omega_1^2 + 1)^3}{\alpha_1^2 b_0^{*2} \omega_1^2}, \quad \text{AsCov}(\hat{b}_0) = \frac{2\sigma^2(a_1^{*2}\omega_1^2 + 1)^2}{\alpha_1^2}.$$

Chapter 11

Summary and future research directions

In this thesis we have studied algorithms for continuous-time system identification. In particular, we have developed and analyzed methods for both indirect and direct approaches, in open and closed-loop, and under different intersample behaviors of the signals. Several robustness issues have been addressed in these methods, including stability enforcement and over-parametrization. Extensions have also been proposed to tackle stability limitations of the refined instrumental variables and possible misspecifications of the intersample behavior of the input signal. A delicate treatment of systems excited by band-limited signals has also been pursued, and new promising ways to identify continuous-time systems in this context have been proposed.

We have focused our attention on the asymptotic properties of continuous-time identification methods. In this regard, the present thesis aimed to fill in the theoretical gaps in the literature, to clarify some misconceptions, and to explain the strengths and limitations of the methods with rigorous mathematical and statistical machinery. The comprehensive proofs of consistency and asymptotic efficiency of the proposed indirect approach method, as well as of the refined instrumental variable methods, provide the necessary theoretical support for the use of these estimators in practical applications and have also led to discovering what experimental conditions may limit their performance.

In the following section we give a short summary of the main results of the chapters herein.

11.1 Thesis conclusions

Indirect approach for continuous-time system identification

In Chapter 3 we investigated the model flexibility and robustness of the indirect approach for continuous-time system identification. The first problem rises due to the fact that the inverse ZOH transformation leads to model mismatches if the true system has relative degree greater than one. The second difficulty was

encountered after solving the first problem: if the relative degree information is taken into account, sometimes the resulting estimate yields an unstable model. To solve the model flexibility issue, we proposed a refinement to the standard indirect approach by applying ideas from the indirect PEM estimator. This refinement has an closed-form expression in this case, and it was shown that the estimator is consistent and asymptotically efficient. Afterwards, an extension was proposed that preserves the asymptotic properties of the basic refinement previously derived while also enforcing stability in the model estimate.

Asymptotic properties of the SRIVC estimator in open-loop

We studied the asymptotic properties of the SRIVC estimator in open-loop in Chapters 4 and 5. The SRIVC estimator was proven to be generically consistent when the intersample behavior of the continuous-time input (ZOH or FOH) is known exactly and subsequently used in the implementation of the algorithm, and consistency was shown to be lost if the intersample behavior of the input signal in the filtered regressor vector is misspecified. In terms of efficiency, we derived the asymptotic Cramér-Rao lower bound for the continuous-time output error model structure and provided the asymptotic covariance expression of the SRIVC estimator. We proved that the SRIVC estimator is asymptotically efficient under the output error model structure if the intersample behavior of the input signal is correctly specified in the algorithm. The properties of the SRIVC estimator when the model is over-parameterized were also studied. We proved that the modified normal matrix is ill-conditioned for small sampling periods in this case, and that non-minimal realizations of the true transfer function are obtained from the stationary points of the SRIVC method when the sample size tends to infinity. It was also shown that the NEVN is well justified as a criterion for determining over-parametrization in the model numerator and/or denominator polynomials.

The SRIVC estimator for arbitrary input excitations

In Chapter 6 we derived an extension of the SRIVC method that can deal with input signals with arbitrary (but known) intersample behavior. This procedure proposes an exact computation of the continuous-time filtering of the instrument and regressor vectors of the SRIVC algorithm, and it was proven to yield generically consistent estimates for multisine inputs. Afterwards, we established several connections with this estimator and other well-known procedures. First, we characterized the extended SRIVC estimator in terms of the equation it satisfies at its converging point, and noted that as more data is collected, it corresponds to the maximum likelihood estimator when full knowledge of the input signal is considered in the probabilistic description of the model. Next, we related the iterations of this method to the Gauss-Newton iterations for maximizing the likelihood function, and analyzed the differences between these two methods. The role of the intersample behavior of the output was studied, and various computational aspects were discussed.

Asymptotic properties of the SRIVC and CLSRIVC estimators in closed-loop

We studied the generic consistency of the SRIVC and CLSRIVC estimators for two well-established closed-loop setups in Chapter 7. More precisely, we showed that the SRIVC and CLSRIVC estimators are generically consistent when there is a discrete-time controller in the loop, and that consistency is lost if a continuous-time controller is implemented. The bias for the fully continuous-time case may be overcome by oversampling methods that compute a more adequate regressor vector. The consistency of the CLSRIVC estimator was shown to extend to more general scenarios than that of the SRIVC estimator, and the bias of the SRIVC estimator towards the negative inverse of the controller in the discrete-time control setup was characterized.

Refined instrumental variables for unstable systems

In Chapter 8 we presented extensions of refined instrumental variable methods that allow the direct and indirect identification of unstable continuous-time systems in closed-loop. An ad-hoc noise model has been proposed such that the unstable poles of the system model are canceled at each iteration of the refined instrumental variable algorithms. This tool enforces stability in the bank of prefilters of the LSSVF, SRIVC, RIVC and CLSRIVC estimators, leading to novel extensions of these estimators that can be applied for the identification of stable or unstable systems. Model estimation with integral action was also covered.

Analysis and identification of continuous-time systems with band-limited inputs

Chapters 9 and 10 addressed sampling properties of systems with band-limited inputs, and how to identify them. We first proved that the equivalent discrete-time system for band-limited inputs is non-causal. This result led to deriving several alternative expressions for the band-limited equivalent and its non-causal part, and some properties related to oversampling the input signal were studied. The non-causality of the discrete-time representation gave way to exploring novel continuous-time system identification methods with band-limited input excitation, such as non-causal impulse response estimators with regularized least-squares, and frequency-domain refined instrumental variables with the frequency response estimate as a proxy for the data. A comprehensive statistical analysis for when the input is a continuous-time multisine signal was also provided. This analysis revealed that the prediction error method, optimal refined instrumental variables for multisine inputs, and the maximum likelihood estimator of any parametric model are directly linked with the interpolation and approximation of frequency functions. These relationships have implications on how these methods can be computed and shed light on some attractive finite-sample properties.

11.2 Some research directions

To conclude, we hint at future research directions that relate to the contents of this thesis.

Indirect approach for time-delay systems and hybrid Box-Jenkins models

The problem of estimating continuous-time systems with time delays is difficult due to the existence of several local minima in most standard choices of cost functions. In direct continuous-time system identification, there has been recent interest in proposing solutions to this problem [29, 33–35, 104]. An interesting idea for future research is to develop a method that can estimate continuous-time models with an inherent time delay using a refined version of the indirect approach. Such method could provide an alternative to the direct methods and may have computational advantages, as the description of time-delay systems in discrete-time might be better conditioned for estimation. Possibly the IPEM estimator studied in Chapter 3 could be extended to this case. On another note, the proposed IPEM estimator could also be extended to estimate hybrid Box-Jenkins models.

Bias analysis of SRIVC

Chapters 4 to 7 focused on obtaining asymptotic properties of the SRIVC estimator and its extension. For finite sample size, initial conditions of the filters and the signal reconstruction techniques may introduce biases. A bias analysis for the LSSVF estimator was recently presented in [117], where an upper bound on the estimation error was obtained. A similar analysis for the SRIVC estimator has not been done yet due to difficulties in analyzing the iterations. A bias analysis, even if only feasible for first order systems, could reveal the effects of the sampling time, initialization, and initial conditions of the filters, on the performance of the SRIVC estimator for a finite amount of data. Moreover, a deeper analysis of the bias of the SRIVC method when a wrong hold mechanism is used for generating the input as a function of the sampling period would be of practical interest, and it may proceed from similar techniques to the ones presented through Chapters 4 to 7.

Robustifying the SRIVC method

Although the SRIVC method has been widely regarded as a reliable algorithm for continuous-time system identification [59], some robustness issues are worth exploring even further. Examples of this include studying the effect of the initial conditions of the prefilters for estimation with small data sets, analyzing more carefully the implications of choosing a monic or anti-monic denominator polynomial, and proposing numerically-robust means for computing the iterations when the normalized normal matrix is poorly conditioned. Furthermore, the estimation of high-order and highly resonant systems are a challenge, and carefully addressing

the issues above might have direct applications in, e.g., identification of motion systems [162].

Persistence of excitation conditions for continuous-time methods

Persistence of excitation is known to be a fundamental concept in the literature on identification of dynamical systems. While the necessary and sufficient requirements in terms of persistence of excitation have been thoroughly explored for discrete-time methods, the work in Chapter 4 (in particular, Remark 4.3) has revealed that it is not clear whether a persistence of excitation order larger than the number of parameters is needed for continuous-time system identification. A discussion on this topic can also be found in [193]. Although theoretical proofs and simulation tests show that a persistence of excitation order of $n + m + 1$ is necessary and sufficient for some low-order systems, a proof of this conjecture for any LTI system has been elusive and is of practical and theoretical interest.

Identification of systems with band-limited inputs

Although many different aspects of the identification of systems with band-limited inputs have been covered in Chapters 9 and 10, some questions remained unaddressed. First, the implications of having “approximately band-limited” signals for identification can be analyzed in detail, as most of the theoretical results consider “exact” band-limited signals that extend infinitely in both directions of time. In addition, asymptotic properties of the FSRIVC estimator, as well as transparent covariance expressions for $\hat{G}_N(i\omega)$ similar to the ones in [114], can also be pursued. Future work also concerns exploiting oversampling techniques for identification with band-limited inputs, the identification of discrete-time non-causal rational transfer functions for simulation purposes, closed-loop extensions, and bounds on approximation errors for causal versus non-causal representations.

Continuous-time multi-step least-square methods

Recently, multi-step least-squares methods for discrete-time system identification such as the WNSF method have been developed and analyzed [55]. One advantage of the WNSF method is that, contrary to the SRIVC estimator [250], it can provide asymptotically efficient estimates with only a finite number of least-square steps, and it does not depend on iterations. The WNSF procedure has also proven useful for settings where the PEM cost function is highly non-convex, avoiding local minima issues. One interesting line of research consists in developing a similar procedure to directly estimate continuous-time systems. Frequency-domain variants of WNSF for band-limited inputs is yet another extension of the results in Chapter 10 that is worth exploring.

Identification in event-based sampling schemes

As an alternative to equidistant sampling in time (“Riemann” sampling), Lebesgue sampling [11] is a type of event-based sampling which consists in sampling the continuous-time signal whenever it crosses fixed and regularly partitioned thresholds. Continuous-time system identification when the output is Lebesgue-sampled has not yet been fully explored. Even though SRIVC for irregularly sampled data [118] can be applied, it does not use the fact that the sampled output is known to be in a specific bounded interval between sampling instants, which leads to a loss of optimality. Several questions can be posed in this topic, such as how to identify continuous-time systems in a state-space description (or stochastic differential equation models, see (2.15)) under Lebesgue sampling, input design for Lebesgue-sampled systems, and threshold selection. Preliminary results on this topic can be found in our recent preprint [90], where an EM algorithm is derived for identifying continuous-time state-space models with Lebesgue-sampled output measurements.

Bibliography

- [1] M. Abramowitz and I. A. Stegun. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. US Government printing office, 1964.
- [2] J. Ackermann. *Robust control: Systems with uncertain physical parameters*. Springer, 1993.
- [3] J. Ackermann and D. Kaesbauer. Stable polyhedra in parameter space. *Automatica*, 39(5):937–943, 2003.
- [4] H. Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723, 1974.
- [5] K. F. Aljanaideh and D. S. Bernstein. Closed-loop identification of unstable systems using noncausal FIR models. *International Journal of Control*, 90(2):168–185, 2017.
- [6] T. Andersson, P. Pucar, and L. Ljung. Identification aspects of inter-sampling behavior. Technical report, Linköping University, 1994.
- [7] R. E. Araújo, A. V. Leite, and D. S. Freitas. Indirect parameter estimation of continuous-time systems using discrete time. In *29th Annual Conference of the IEEE Industrial Electronics Society*, pages 600–605, 2003.
- [8] K. J. Åström. On the choice of sampling rates in parametric identification of time series. *Information Sciences*, 1(3):273–278, 1969.
- [9] K. J. Åström. *Introduction to Stochastic Control Theory*. Academic Press, 1970.
- [10] K. J. Åström. Maximum likelihood and prediction error methods. *Automatica*, 16(8):551–574, 1979.
- [11] K. J. Åström and B. Bernhardsson. Comparison of periodic and event based sampling for first-order stochastic systems. In *Proceedings of the 14th IFAC World Congress*, volume 11, pages 301–306, 1999.

- [12] K. J. Åström and B. Bernhardsson. Systems with Lebesgue sampling. In A. Rantzer and C. I. Byrnes (Eds.). *Directions in Mathematical Systems Theory and Optimization*, pages 1–13. Springer, 2003.
- [13] K. J. Åström and B. M. Bernhardsson. Comparison of Riemann and Lebesgue sampling for first order stochastic systems. In *41st IEEE Conference on Decision and Control (CDC)*, volume 2, pages 2011–2016, 2002.
- [14] K. J. Åström and T. Bohlin. Numerical identification of linear dynamic systems from normal operating records. *IFAC Proceedings Volumes*, 2(2):96–111, 1965.
- [15] K. J. Åström and P. Eykhoff. System identification—A survey. *Automatica*, 7(2):123–162, 1971.
- [16] K. J. Åström and B. Wittenmark. *Computer Controlled Systems: Theory and Design*. Prentice-Hall, 1984.
- [17] K. J. Åström, P. Hagander, and J. Sternby. Zeros of sampled systems. *Automatica*, 20(1):31–38, 1984.
- [18] L. Blanken and T. Oomen. Kernel-based identification of non-causal systems with application to inverse model control. *Automatica*, 114, Article 108830, 2020.
- [19] R. P. Boas Jr. Summation formulas and band-limited signals. *Tohoku Mathematical Journal, Second Series*, 24(2):121–125, 1972.
- [20] H. Boche and U. J. Mönich. General behavior of sampling-based signal and system representation. In *IEEE International Symposium on Information Theory*, pages 2439–2443, 2008.
- [21] H. Boche and U. J. Mönich. Sampling of deterministic signals and systems. *IEEE Transactions on Signal Processing*, 59(5):2101–2111, 2011.
- [22] H. Boche and U. J. Mönich. System representations for the Paley–Wiener Space PW_π^2 . *Journal of Fourier Analysis and Applications*, 24(1):285–308, 2018.
- [23] S. Bochner and W. T. Martin. *Several Complex Variables*. Princeton University Press, 1948.
- [24] G. E. P. Box. Science and statistics. *Journal of the American Statistical Association*, 71(356):791–799, 1976.
- [25] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

- [26] B. P. Brooks. The coefficients of the characteristic polynomial in terms of the eigenvalues and the elements of an $n \times n$ matrix. *Applied Mathematics Letters*, 19(6):511–515, 2006.
- [27] P. E. Caines and L. Ljung. Prediction error estimators: Asymptotic normality and accuracy. In *IEEE Conference on Decision and Control including the 15th Symposium on Adaptive Processes*, pages 652–658, 1976.
- [28] G. Casella and R. L. Berger. *Statistical Inference*, 2nd Edition. Duxbury, 2002.
- [29] F. Chen and P. C. Young. A simple robust method of fractional time-delay estimation for linear dynamic systems. *Automatica*, 137, Article 110117, 2022.
- [30] F. Chen, H. Garnier, and M. Gilson. Refined instrumental variable identification of continuous-time OE and BJ models from irregularly sampled data. In *11th IFAC International Workshop on Adaptation and Learning in Control and Signal Processing (ALCOSP)*, Caen, France, 2013.
- [31] F. Chen, H. Garnier, M. Gilson, J. C. Agüero, and B. I. Godoy. Identification of continuous-time transfer function models from non-uniformly sampled data in presence of colored noise. In *IFAC Proceedings Volumes*, volume 47, pages 10379–10384, 2014.
- [32] F. Chen, M. Gilson, J. C. Agüero, H. Garnier, and J. Schorsch. Closed-loop identification of continuous-time systems from non-uniformly sampled data. In *2014 European Control Conference (ECC)*, pages 19–24, 2014.
- [33] F. Chen, H. Garnier, and M. Gilson. Robust identification of continuous-time models with arbitrary time-delay from irregularly sampled data. *Journal of Process Control*, 25:19–27, 2015.
- [34] F. Chen, M. Gilson, H. Garnier, and T. Liu. Robust time-domain output error method for identifying continuous-time systems with time delay. *Systems & Control Letters*, 102:81–92, 2017.
- [35] F. Chen, H. Garnier, A. Padilla, and M. Gilson. Recursive IV identification of continuous-time models with time delay from sampled data. *IEEE Transactions on Control Systems Technology*, 28(3):1074–1082, 2019.
- [36] T. Chen and L. Ljung. Implementation of algorithms for tuning parameters in regularized least squares problems in system identification. *Automatica*, 49(7):2213–2220, 2013.
- [37] T. Chen, H. Ohlsson, and L. Ljung. On the estimation of transfer functions, regularizations and Gaussian processes—Revisited. *Automatica*, 48(8):1525–1535, 2012.

- [38] X.-W. Chen and X. Lin. Big data deep learning: Challenges and perspectives. *IEEE Access*, 2:514–525, 2014.
- [39] C. T. Chou, M. Verhaegen, and R. Johansson. Continuous-time identification of SISO systems using Laguerre functions. *IEEE Transactions on Signal Processing*, 47(2):349–362, 1999.
- [40] A. Cuyt and L. Wuytack. *Nonlinear Methods in Numerical Analysis*. Elsevier, 1987.
- [41] M. Dahleh, M. A. Dahleh, and G. Verghese. *Lectures on Dynamic Systems and Control*. Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, 2002.
- [42] S. Datta, D. Pal, and D. Chakraborty. Partial pole placement and controller norm optimization over polynomial stability region. *IFAC Proceedings Volumes*, 44(1):10129–10134, 2011.
- [43] E. Díaz-Francés and F. J. Rubio. On the existence of a normal approximation to the distribution of the ratio of two independent normal random variables. *Statistical Papers*, 54(2):309–323, 2013.
- [44] E. DiBenedetto. *Real Analysis*. Springer, 2002.
- [45] J. R. Dormand and P. J. Prince. A family of embedded Runge-Kutta formulae. *Journal of Computational and Applied Mathematics*, 6(1):19–26, 1980.
- [46] G. Durgaprasad, G. P. Rao, A. Patra, and S. Mukhopadhyay. Indirect methods of parameter estimation of discrete-time models. *IFAC Proceedings Volumes*, 25(15):343–348, 1992.
- [47] S. A. Dyer and J. S. Dyer. Least-squares fitting of data by rational functions: Levy’s method (part 1). *IEEE Instrumentation & Measurement Magazine*, 12(6):40–43, 2009.
- [48] R. E. Edwards. *Fourier Series: A Modern Introduction, Volume 1*, 2nd Edition. Springer, 1979.
- [49] A. Feuer and G. C. Goodwin. *Sampling in Digital Signal Processing and Control*. Birkhäuser, 1996.
- [50] U. Forssell and H. Hjalmarsson. Maximum likelihood estimation of models with unstable dynamics and non-minimum phase noise zeros. In *14th IFAC World Congress*, pages 3784–3789, 1999.
- [51] U. Forssell and L. Ljung. Closed-loop identification revisited. *Automatica*, 35(7):1215–1241, 1999.

- [52] U. Forssell and L. Ljung. Identification of unstable systems using output error and Box-Jenkins model structures. *IEEE Transactions on Automatic Control*, 45(1):137–141, 2000.
- [53] U. Forssell and L. Ljung. A projection method for closed-loop identification. *IEEE Transactions on Automatic Control*, 45(11):2101–2106, 2000.
- [54] G. F. Franklin, J. D. Powell, and M. L. Workman. *Digital Control of Dynamic Systems*, 3rd Edition. Addison-Wesley, 1998.
- [55] M. Galrinho. *System Identification with Multi-Step Least-Squares Methods*. PhD thesis, KTH Royal Institute of Technology, 2018.
- [56] M. Galrinho, C. R. Rojas, and H. Hjalmarsson. A weighted least squares method for estimation of unstable systems. In *55th IEEE Conference on Decision and Control (CDC)*, pages 341–346, 2016.
- [57] M. Galrinho, N. Everitt, and H. Hjalmarsson. ARX modeling of unstable linear systems. *Automatica*, 75:167–171, 2017.
- [58] M. Galrinho, C. R. Rojas, and H. Hjalmarsson. Parametric identification using weighted null-space fitting. *IEEE Transactions on Automatic Control*, 64(7):2798–2813, 2018.
- [59] H. Garnier. Direct continuous-time approaches to system identification. Overview and benefits for practical applications. *European Journal of control*, 24:50–62, 2015.
- [60] H. Garnier and M. Gilson. CONTSID: a Matlab toolbox for standard and advanced identification of black-box continuous-time models. *IFAC-PapersOnLine*, 51(15):688–693, 2018.
- [61] H. Garnier and L. Wang (Eds.). *Identification of Continuous-time Models from Sampled Data*. Springer, 2008.
- [62] H. Garnier and P. C. Young. Time-domain approaches to continuous-time model identification of dynamical systems from sampled data. In *American Control Conference (ACC)*, pages 667–672, 2004.
- [63] H. Garnier and P. C. Young. The advantages of directly identifying continuous-time transfer function models in practical applications. *International Journal of Control*, 87(7):1319–1338, 2014.
- [64] H. Garnier, M. Gilson, and W. X. Zheng. A bias-eliminated least-squares method for continuous-time model identification of closed-loop systems. *International Journal of Control*, 73(1):38–48, 2000.

- [65] H. Garnier, M. Mensler, and A. Richard. Continuous-time model identification from sampled data: implementation issues and performance evaluation. *International Journal of Control*, 76(13):1337–1357, 2003.
- [66] 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.
- [67] H. Garnier, M. Gilson, T. Bastogne, and M. Mensler. The CONTSID toolbox: A software support for data-based continuous-time modelling. In H. Garnier and L. Wang (Eds.) *Identification of Continuous-time Models from Sampled Data*, pages 249–290. Springer, 2008.
- [68] H. Garnier, P. C. Young, and M. Gilson. Simple refined IV methods of closed-loop system identification. *IFAC Proceedings Volumes*, 42(10):1151–1156, 2009.
- [69] 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.
- [70] H. Garnier, M. Gilson, H. Muller, and F. Chen. A new graphical user interface for the CONTSID toolbox for Matlab. *IFAC-PapersOnLine*, 54(7):397–402, 2021.
- [71] R. C. Geary. The frequency distribution of the quotient of two normal variates. *Journal of the Royal Statistical Society*, 93(3):442–446, 1930.
- [72] J. Gillberg and L. Ljung. Frequency-domain identification of continuous-time ARMA models from sampled data. *Automatica*, 45(6):1371–1378, 2009.
- [73] J. Gillberg and L. Ljung. Frequency domain identification of continuous-time output error models, Part I: Uniformly sampled data and frequency function approximation. *Automatica*, 46(1):1–10, 2010.
- [74] M. Gilson. What has instrumental variable method to offer for system identification? *IFAC-PapersOnLine*, 48(1):354–359, 2015.
- [75] M. Gilson and H. Garnier. Continuous-time model identification of systems operating in closed-loop. *IFAC Proceedings Volumes*, 36(16):405–410, 2003.
- [76] M. Gilson and P. M. J. Van den Hof. Instrumental variable methods for closed-loop system identification. *Automatica*, 41(2):241–249, 2005.
- [77] M. Gilson, H. Garnier, P. C. Young, and P. M. J. Van den Hof. Instrumental variable methods for closed-loop continuous-time model identification. In H. Garnier and L. Wang (Eds.). *Identification of continuous-time models from sampled data*, pages 133–160. Springer, 2008.

- [78] M. Gilson, H. Garnier, P. C. Young, and P. M. J. Van den Hof. Optimal instrumental variable method for closed-loop identification. *IET Control Theory & Applications*, 5(10):1147–1154, 2011.
- [79] M. Gilson, J. S. Welsh, and H. Garnier. A frequency localizing basis function-based IV method for wideband system identification. *IEEE Transactions on Control Systems Technology*, 26(1):329–335, 2017.
- [80] R. A. González and C. R. Rojas. A fully Bayesian approach to kernel-based regularization for impulse response estimation. *IFAC-PapersOnLine*, 51(15):186–191, 2018.
- [81] R. A. González and C. R. Rojas. A finite-sample deviation bound for stable autoregressive processes. In *Proceedings of the 2nd Annual Conference on Learning for Dynamics and Control (L4DC)*, pages 191–200, 2020.
- [82] R. A. González and C. R. Rojas. Finite sample deviation and variance bounds for first order autoregressive processes. In *Proceedings of the 45th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 5380–5384, 2020.
- [83] R. A. González, C. R. Rojas, and J. S. Welsh. An asymptotically optimal indirect approach to continuous-time system identification. In *57th IEEE Conference on Decision and Control (CDC)*, pages 638–643, 2018.
- [84] R. A. González, F. J. Vargas, and J. Chen. Stabilization of MIMO systems over additive correlated noise channels subject to multiple SNR-constraints. In *2018 European Control Conference (ECC)*, pages 1493–1498, 2018.
- [85] R. A. González, F. J. Vargas, and J. Chen. Mean square stabilization over SNR-constrained channels with colored and spatially correlated additive noises. *IEEE Transactions on Automatic Control*, 64(11):4825–4832, 2019.
- [86] R. A. González, C. R. Rojas, S. Pan, and J. S. Welsh. Consistent identification of continuous-time systems under multisine input signal excitation. *Automatica*, 133, Article 109859, 2020.
- [87] R. A. González, J. S. Welsh, and C. R. Rojas. Enforcing stability through ellipsoidal inner approximations in the indirect approach for continuous-time system identification. In *21st IFAC World Congress (IFAC 2020)*, pages 566–571, 2020.
- [88] R. A. González, C. R. Rojas, and H. Hjalmarsson. Non-causal regularized least-squares for continuous-time system identification with band-limited input excitations. In *60th IEEE Conference on Decision and Control (CDC)*, pages 114–119, 2021.

- [89] R. A. González, C. R. Rojas, S. Pan, and J. S. Welsh. The SRIVC algorithm for continuous-time system identification with arbitrary input excitation in open and closed loop. In *60th IEEE Conference on Decision and Control (CDC)*, pages 3004–3009, 2021.
- [90] R. A. González, A. L. Cedeño, M. Coronel, C. R. Rojas, and J. C Agüero. Identification of continuous-time state-space systems utilizing Lebesgue-sampled data. In Submitted to the *61th IEEE Conference on Decision and Control (CDC)*, 2022.
- [91] R. A. González, S. Pan, C. R. Rojas, and J. S. Welsh. Consistency analysis of refined instrumental variables for continuous-time system identification in closed-loop. Submitted for publication to the European Journal of Control, 2022.
- [92] R. A. González, C. R. Rojas, H. Hjalmarsson, and J. Schoukens. Identification of continuous-time systems excited by band-limited inputs. In preparation, 2022.
- [93] R. A. González, C. R. Rojas, S. Pan, and J. S. Welsh. Theoretical and practical aspects of the convergence of the SRIVC estimator for over-parameterized models. Accepted for publication in *Automatica*, 2022.
- [94] R. A. González, C. R. Rojas, S. Pan, and J. S. Welsh. Refined instrumental variable methods for unstable continuous-time systems. Accepted for publication in the International Journal of Control, 2022.
- [95] G. C. Goodwin and R. L. Payne. *Dynamic System Identification: Experiment Design and Data Analysis*. Academic Press, 1977.
- [96] G. C. Goodwin and J. S. Welsh. Bias issues in closed loop identification with application to adaptive control. *Communications in Information and Systems*, 2(4):349–370, 2002.
- [97] G. C. Goodwin, S. F. Graebe, and M. E. Salgado. *Control System Design*. Prentice Hall, 2001.
- [98] C. Gourieroux and A. Monfort. *Statistics and Econometric Models*, volumes 1 and 2. Cambridge University Press, 1995.
- [99] C. Gourieroux, A. Monfort, and E. Renault. Indirect inference. *Journal of Applied Econometrics*, 8(S1):S85–S118, 1993.
- [100] M. Grant and S. Boyd. CVX: MATLAB software for disciplined convex programming, version 2.1, 2014.
- [101] P. R. Graves-Morris and T. R. Hopkins. Reliable rational interpolation. *Numerische Mathematik*, 36(2):111–128, 1980.

- [102] R. M. Gray. Toeplitz and Circulant matrices: A review. *Foundations and Trends in Communications and Information Theory*, 2(3):155–239, 2006.
- [103] 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 (CDC)*, pages 2597–2602, 2014.
- [104] H. Ha, J. S. Welsh, and M. Alamir. Useful redundancy in parameter and time delay estimation for continuous-time models. *Automatica*, 95:455–462, 2018.
- [105] M. K. Habib. Digital representations of operators on band-limited random signals. *IEEE Transactions on Information Theory*, 47(1):173–177, 2001.
- [106] P. Hägg, J. Schoukens, M. Gevers, and H. Hjalmarsson. The transient impulse response modeling method for non-parametric system identification. *Automatica*, 68:314–328, 2016.
- [107] G. H. Hardy. Notes on special systems of orthogonal functions (IV): the orthogonal functions of Whittaker's cardinal series. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 37, pages 331–348, 1941.
- [108] S. Haykin and B. Van Veen. *Signals and Systems*. John Wiley & Sons, 2007.
- [109] W. P. Heath. Bias of indirect non-parametric transfer function estimates for plants in closed loop. *Automatica*, 37(10):1529–1540, 2001.
- [110] D. Henrion, D. Peaucelle, D. Arzelier, and M. Sebek. Ellipsoidal approximation of the stability domain of a polynomial. *IEEE Transactions on Automatic Control*, 48(12):2255–2259, 2003.
- [111] J. R. Higgins. *Sampling Theory in Fourier and Signal Analysis: Foundations*. Oxford University Press, 1996.
- [112] D. V. Hinkley. On the ratio of two correlated normal random variables. *Biometrika*, 56(3):635–639, 1969.
- [113] J. L. Hintze and R. D. Nelson. Violin plots: a box plot-density trace synergism. *The American Statistician*, 52(2):181–184, 1998.
- [114] H. Hjalmarsson and B. Ninness. Least-squares estimation of a class of frequency functions: A finite sample variance expression. *Automatica*, 42(4):589–600, 2006.
- [115] R. A. Horn and C. R. Johnson. *Matrix Analysis*, 2nd Edition. Cambridge University Press, 2012.
- [116] B. Y.-K. Hu. Kramers–Kronig in two lines. *American Journal of Physics*, 57(9):821, 1989.

- [117] X.-L. Hu and J. S. Welsh. Continuous-time model identification from filtered sampled data: Error analysis. *IEEE Transactions on Automatic Control*, 65(10):4005–4015, 2020.
- [118] E. Huselstein and H. Garnier. An approach to continuous-time model identification from non-uniformly sampled data. In *41st IEEE Conference on Decision and Control (CDC)*, volume 1, pages 622–623, 2002.
- [119] D. Jackson. *The Theory of Approximation*. American Mathematical Society, 1930.
- [120] R. Johansson. Identification of continuous-time models. *IEEE Transactions on Signal Processing*, 42(4):887–897, 1994.
- [121] R. P. Kanwal. *Generalized Functions: Theory and Applications*, 3rd Edition. Springer, 2011.
- [122] I. Kollár, G. Franklin, and R. Pintelon. On the equivalence of z -domain and s -domain models in system identification. In *Instrumentation and Measurement Technology Conference*, Brussels, Belgium, volume 1, pages 14–19, 1996.
- [123] S. G. Krantz and H. R. Parks. *A Primer of Real Analytic Functions*. Springer, 2002.
- [124] N. R. Kristensen, H. Madsen, and S. B. Jørgensen. Parameter estimation in stochastic grey-box models. *Automatica*, 40(2):225–237, 2004.
- [125] D. O. Kuethe, A. Caprihan, H. M. Gach, I. J. Lowe, and E. Fukushima. Imaging obstructed ventilation with NMR using inert fluorinated gases. *Journal of Applied Physiology*, 88(6):2279–2286, 2000.
- [126] E. K. Larsson. *Identification of Stochastic Continuous-time Systems*. PhD thesis. Uppsala University, Division of Systems and Control, 2004.
- [127] B. P. Lathi and R. A. Green. *Essentials of Digital Signal Processing*. Cambridge University Press, 2014.
- [128] V. Laurain, M. Gilson, and H. Garnier. Refined instrumental variable methods for identifying Hammerstein models operating in closed loop. In *Proceedings of the 48h IEEE Conference on Decision and Control (CDC) held jointly with the 28th Chinese Control Conference*, pages 3614–3619, 2009.
- [129] W. Ledermann. *Complex Numbers*. Springer, 1962.
- [130] E. L. Lehmann and G. Casella. *Theory of Point Estimation*, 2nd Edition. Springer, 1998.

- [131] H. Lev-Ari, Y. Bistritz, and T. Kailath. Generalized bezoutians and families of efficient zero-location procedures. *IEEE Transactions on Circuits and Systems*, 38(2):170–186, 1991.
- [132] E. C. Levy. Complex-curve fitting. *IRE Transactions on Automatic Control*, AC-4(1):37–43, 1959.
- [133] P. Li. Singular integral equations of convolution type with Hilbert kernel and a discrete jump problem. *Advances in Difference Equations*, 1:1–13, 2017.
- [134] Q. Li, D. Li, and L. Cao. Closed-loop identification of systems using hybrid Box–Jenkins structure and its application to PID tuning. *Chinese Journal of Chemical Engineering*, 23(12):1997–2004, 2015.
- [135] X. Liu, J. Wang, and W. X. Zheng. Convergence analysis of refined instrumental variable method for continuous-time system identification. *IET Control Theory & Applications*, 5(7):868–877, 2011.
- [136] L. Ljung. Characterization of the concept of “persistently exciting” in the frequency domain. *Report TFRT*, 3038, 1971.
- [137] L. Ljung. On the estimation of transfer functions. *Automatica*, 21(6):677–696, 1985.
- [138] L. Ljung. *System Identification: Theory for the User*, 2nd Edition. Prentice-Hall, 1999.
- [139] L. Ljung. Initialisation aspects for subspace and Output-Error identification methods. In *European Control Conference (ECC)*, Cambridge, UK, pages 773–778, 2003.
- [140] L. Ljung. *System Identification Toolbox 7: Getting Started Guide*. The MathWorks, 2008.
- [141] L. Ljung. Experiments with identification of continuous time models. In *15th IFAC Symposium on System Identification*, Saint Malo, France, pages 1175–1180, 2009.
- [142] L. Ljung and P. E. Caines. Asymptotic normality of prediction error estimators for approximate system models. *Stochastics*, 3(1-4):29–46, 1980.
- [143] L. Ljung and R. Singh. Version 8 of the MATLAB System Identification Toolbox. In *16th IFAC Symposium on System Identification*, Brussels, Belgium, pages 1826–1831, 2012.
- [144] L. Ljung and A. Wills. Issues in sampling and estimating continuous-time models with stochastic disturbances. *Automatica*, 46(5):925–931, 2010.

- [145] L. Ljung, T. Chen, and B. Mu. A shift in paradigm for system identification. *International Journal of Control*, 93(2):173–180, 2020.
- [146] L. Ljung, T. Glad, and A. Hansson. *Modeling and Identification of Dynamic Systems*. Studentlitteratur, 2021.
- [147] Q. Lu, P. D. Loewen, R. B. Gopaluni, M. G. Forbes, J. U. Backström, G. A. Dumont, and M. S. Davies. Identification of symmetric noncausal processes. *Automatica*, 103:515–530, 2019.
- [148] D. Marelli and M. Fu. Exact identification of continuous-time systems from sampled data. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 757–760, 2007.
- [149] D. Marelli and M. Fu. A continuous-time linear system identification method for slowly sampled data. *IEEE Transactions on Signal Processing*, 58(5):2521–2533, 2010.
- [150] G. Marsaglia. Ratios of normal variables. *Journal of Statistical Software*, 16(4):1–10, 2006.
- [151] I. Maruta and T. Sugie. Projection-based identification algorithm for grey-box continuous-time models. *Systems & Control Letters*, 62(11):1090–1097, 2013.
- [152] T. McKelvey. Frequency domain identification methods. *Circuits, Systems and Signal Processing*, 21(1):39–55, 2002.
- [153] T. McKelvey and G. Guérin. Non-parametric frequency response estimation using a local rational model. *IFAC Proceedings Volumes*, 45(16):49–54, 2012.
- [154] R. H. Middleton and G. C. Goodwin. *Digital Control and Estimation: A Unified Approach*. Prentice-Hall, 1990.
- [155] B. S. Mityagin. The zero set of a real analytic function. *Matematicheskie Zametki*, 107(3):473–475, 2020.
- [156] S. Monaco and D. Normand-Cyrot. Zero dynamics of sampled nonlinear systems. *Systems & Control Letters*, 11(3):229–234, 1988.
- [157] J. J. Moré. The Levenberg-Marquardt algorithm: implementation and theory. In *Conference on Numerical Analysis*, pages 105–116, 1978.
- [158] S. Müller and P. Massarani. Transfer-function measurement with sweeps. *Journal of the Audio Engineering Society*, 49(6):443–471, 2001.
- [159] F. Natterer. *The Mathematics of Computerized Tomography*. SIAM, 2001.

- [160] B. Ni, M. Gilson, and H. Garnier. Two-stage refined instrumental variable method for identifying Hammerstein-Wiener continuous-time models in closed loop. In *16th IFAC Symposium on System Identification, Brussels, Belgium*, pages 25–30, 2012.
- [161] J. Nocedal and S. Wright. *Numerical Optimization*. Springer, 2006.
- [162] T. Oomen. Control for precision mechatronics. In *Encyclopedia of Systems and Control*. Springer, 2019.
- [163] A. Padilla, H. Garnier, P. C. Young, and J. Yuz. Recursive online IV method for identification of continuous-time slowly time-varying models in closed loop. *IFAC-PapersOnLine*, 50(1):4008–4013, 2017.
- [164] 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*, 113, Article 108767, 2020.
- [165] S. Pan, J. S. Welsh, A. M. Brichta, H. R. Drury, and J. G. Stoddard. Estimating the membrane properties of vestibular type II hair cells using continuous-time system identification. *IFAC-PapersOnLine*, 53(2):548–553, 2020.
- [166] 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*, 121, Article 109196, 2020.
- [167] S. Pan, R. A. González, J. S. Welsh, and C. R. Rojas. Corrigendum to “Consistency analysis of the Simplified Refined Instrumental Variable method for Continuous-time systems” [Automatica 113(2020) 108767]. *Automatica*, 136, Article 109946, 2021.
- [168] S. Pan, J. S. Welsh, R. A. González, and C. R. Rojas. Consistency analysis and bias elimination of the Instrumental Variable Based State Variable Filter method. Provisionally accepted for publication in *Automatica*, 2022.
- [169] A. Papoulis. *Signal Analysis*. McGraw-Hill, 1977.
- [170] A. Papoulis and S. U. Pillai. *Probability, Random Variables, and Stochastic Processes*, 4th Edition. McGraw-Hill, 2002.
- [171] P. C. Parks and V. Hahn. *Stability theory*. Prentice-Hall, 1993.
- [172] G. Pillonetto, F. Dinuzzo, T. Chen, G. De Nicolao, and L. Ljung. Kernel methods in system identification, machine learning and function estimation: A survey. *Automatica*, 50(3):657–682, 2014.
- [173] R Pintelon and J. Schoukens. Identification of continuous-time systems using arbitrary signals. *Automatica*, 33(5):991–994, 1997.

- [174] R. Pintelon and J. Schoukens. Measurement of frequency response functions using periodic excitations, corrupted by correlated input/output errors. *IEEE Transactions on Instrumentation and Measurement*, 50(6):1753–1760, 2001.
- [175] R. Pintelon and J. Schoukens. *System Identification: A Frequency Domain Approach*, 2nd Edition. John Wiley & Sons, 2012.
- [176] R. Pintelon, P. Guillaume, Y. Rolain, J. Schoukens, and H. Van Hamme. Parametric Identification of Transfer Functions in the Frequency Domain—A Survey. *IEEE Transactions on Automatic Control*, 39(11):2245–2260, 1994.
- [177] R. Pintelon, J. Schoukens, and H. Chen. On the basic assumptions in the identification of continuous time systems. *IFAC Proceedings Volumes*, 27(8):1105–1114, 1994.
- [178] R. Pintelon, J. Schoukens, and G. Vandersteen. Frequency domain system identification using arbitrary signals. *IEEE Transactions on Automatic Control*, 42(12):1717–1720, 1997.
- [179] R. Pintelon, J. Schoukens, and Y. Rolain. Box–Jenkins continuous-time modeling. *Automatica*, 36(7):983–991, 2000.
- [180] R. Pintelon, Y. Rolain, and W. Van Moer. Probability density function for frequency response function measurements using periodic signals. In *19th IEEE Instrumentation and Measurement Technology Conference*, volume 2, pages 869–874, 2002.
- [181] R. Pintelon, J. Schoukens, and Y. Rolain. Frequency-domain approach to continuous-time system identification: Some practical aspects. In H. Garnier and L. Wang (Eds.). *Identification of continuous-time models from sampled data*, pages 215–248. Springer, 2008.
- [182] J. Ploeg, B. T. M. Scheepers, E. Van Nunen, N. Van de Wouw, and H. Nijmeijer. Design and experimental evaluation of cooperative adaptive cruise control. In *14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pages 260–265, 2011.
- [183] G. Prando, D. Romeres, G. Pillonetto, and A. Chiuso. Classical vs. Bayesian methods for linear system identification: Point estimators and confidence sets. In *2016 European Control Conference (ECC)*, pages 1365–1370, 2016.
- [184] L. Rabiner, R. Crochiere, and J. Allen. FIR system modeling and identification in the presence of noise and with band-limited inputs. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 26(4):319–333, 1978.
- [185] G. P. Rao and H. Garnier. Numerical illustrations of the relevance of direct continuous-time model identification. In *15th Triennial IFAC World Congress on Automatic Control*, Barcelona, Spain, volume 35, pages 133–138, 2002.

- [186] G. P. Rao and H. Unbehauen. Identification of continuous-time systems. *IEE Proceedings-Control Theory and Applications*, 153(2):185–220, 2006.
- [187] R. Relan and J. Schoukens. Recursive discrete-time models for continuous-time systems under band-limited assumptions. *IEEE Transactions on Instrumentation and Measurement*, 65(3):713–723, 2016.
- [188] J. Rissanen. Modeling by shortest data description. *Automatica*, 14(5):465–471, 1978.
- [189] R. A. Rojas and C. R. Rojas. The inverse of sampling revisited. In *IASTED Conference on Intelligent Systems and Control (ISC-2001)*, 2001.
- [190] W. Rudin. *Principles of Mathematical Analysis*, 3rd Edition. McGraw-Hill, 1976.
- [191] W. Rudin. *Real and Complex Analysis*, 3rd Edition. McGraw-Hill, 1986.
- [192] S. Sagara and Z. Y. Zhao. Numerical integration approach to on-line identification of continuous-time systems. *Automatica*, 26(1):63–74, 1990.
- [193] S. Sagara and Z. Y. Zhao. Application of digital filtering techniques. In *Identification of continuous-time systems*, pages 291–325. Springer, 1991.
- [194] D. C. Saha and G. P. Rao. *Identification of Continuous Dynamical Systems—The Poisson Moment Functional (PMF) approach*. Springer, 1983.
- [195] J. Schoukens, R. Pintelon, and P. Guillaume. On the advantages of periodic excitation in system identification. *IFAC Proceedings Volumes*, 27(8):1115–1120, 1994.
- [196] J. Schoukens, R. Pintelon, and H. Van Hamme. Identification of linear dynamic systems using piecewise constant excitations: use, misuse and alternatives. *Automatica*, 30(7):1153–1169, 1994.
- [197] G. Schwarz. Estimating the dimension of a model. *The Annals of Statistics*, 6(2):461–464, 1978.
- [198] M. Shinbrot. On the analysis of linear and nonlinear systems. *Transactions of the American Society of Mechanical Engineers*, 79(3):547–552, 1957.
- [199] N. K. Sinha. Estimation of transfer function of continuous system from sampled data. In *Proceedings of the Institution of Electrical Engineers (IET)*, volume 119, pages 612–614, 1972.
- [200] N. K. Sinha. Identification of continuous-time systems from samples of input-output data: An introduction. *Sadhanā*, 25(2):75–83, 2000.

- [201] N. K. Sinha and S. Puthenpura. Choice of the sampling interval for the identification of continuous-time systems from samples of input/output data. In *IEE Proceedings D-Control Theory and Applications*, volume 132, pages 263–267, 1985.
- [202] N. K. Sinha and Z. Qi-Jie. Discrete-time approximation of multivariable continuous-time systems. In *IEE Proceedings D-Control Theory and Applications*, volume 130, pages 103–110, 1983.
- [203] N. K. Sinha and G. P. Rao (Eds.). *Identification of Continuous-Time Systems: Methodology and Computer Implementation*. Springer, 1991.
- [204] F. Smith. System Laplace-transform estimation from sampled data. *IEEE Transactions on Automatic Control*, 13(1):37–44, 1968.
- [205] T. Söderström. Ergodicity results for sample covariances. *Problems of Control and Information Theory*, 4(2):131–138, 1975.
- [206] T. Söderström. On the uniqueness of maximum likelihood identification. *Automatica*, 11(2):193–197, 1975.
- [207] T. Söderström. Computational methods for evaluating covariance functions. Technical report, Systems and Control Group, Uppsala University, 2003.
- [208] T. Söderström and P. Stoica. Comparison of some instrumental variable methods—consistency and accuracy aspects. *Automatica*, 17(1):101–115, 1981.
- [209] T. Söderström and P. Stoica. Some properties of the output error method. *Automatica*, 18(1):93–99, 1982.
- [210] T. Söderström and P. Stoica. *Instrumental Variable Methods for System Identification*. Springer, 1983.
- [211] T. Söderström and P. Stoica. *System Identification*. Prentice-Hall, 1989.
- [212] T. Söderström, P. Stoica, and B. Friedlander. An indirect prediction error method for system identification. *Automatica*, 27(1):183–188, 1991.
- [213] H. H. Sohrab. *Basic Real Analysis*. Springer, 2003.
- [214] V. Solo. *Time series recursions and stochastic approximation*. PhD. thesis. The Australian National University, 1978.
- [215] E. M. Stein and R. Shakarchi. *Fourier Analysis: An Introduction*, volume 1. Princeton University Press, 2011.
- [216] J. Stoer and R. Bulirsch. *Introduction to Numerical Analysis*, 3rd Edition. Springer, 2002.

- [217] P. Stoica and R. L. Moses. *Spectral Analysis of Signals*. Prentice-Hall, 2005.
- [218] 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.
- [219] P. Stoica and T. Söderström. Optimal instrumental variable estimation and approximate implementations. *IEEE Transactions on Automatic Control*, 28(7):757–772, 1983.
- [220] T. Tao. *Analysis II*, 3rd Edition. Springer, 2015.
- [221] C. J. Taylor, P. C. Young, W. Tych, and E. D. Wilson. New developments in the CAPTAIN Toolbox for Matlab with case study examples. *IFAC-PapersOnLine*, 51(15):694–699, 2018.
- [222] G. C. Temes, V. Barcilon, and F. C. Marshall. The optimization of bandlimited systems. *Proceedings of the IEEE*, 61(2):196–234, 1973.
- [223] J. Umenberger, J. Wågberg, I. R. Manchester, and T. B. Schön. Maximum likelihood identification of stable linear dynamical systems. *Automatica*, 96:280–292, 2018.
- [224] H. Unbehauen and G. P. Rao. *Identification of continuous systems*. Elsevier, 1987.
- [225] H. Unbehauen and G. P. Rao. Continuous-time approaches to system identification—a survey. *Automatica*, 26(1):23–35, 1990.
- [226] H. Unbehauen and G. P. Rao. A review of identification in continuous-time systems. *Annual Reviews in Control*, 22:145–171, 1998.
- [227] P. M. J. Van den Hof. Closed-loop issues in system identification. *Annual Reviews in Control*, 22:173–186, 1998.
- [228] P. M. J. Van den Hof and S. G. Douma. An IV-based iterative linear regression algorithm with optimal output error properties. Technical report, Delft Center for Systems and Control, 2008.
- [229] P. M. J. Van den Hof, A. Dankers, P. S. C. Heuberger, and X. Bombois. Identification of dynamic models in complex networks with prediction error methods—Basic methods for consistent module estimates. *Automatica*, 49(10):2994–3006, 2013.
- [230] A. W. Van der Vaart. *Asymptotic Statistics*. Cambridge University Press, 2000.

- [231] H. Van Hamme, R. Pintelon, and J. Schoukens. Discrete-time modeling and identification of continuous-time systems: a general framework. In H. Garnier and L. Wang (Eds.). *Identification of Continuous-Time Systems*, pages 17–77. Springer, 1991.
- [232] F. J. Vargas and R. A. González. On the existence of a stabilizing solution of modified algebraic Riccati equations in terms of standard algebraic Riccati equations and linear matrix inequalities. *IEEE Control Systems Letters*, 4(1):91–96, 2019.
- [233] S. Victor, A. Diudichi, and P. Melchior. Closed-loop continuous-time model identification with noisy input-output. *IFAC-PapersOnLine*, 50(1):12853–12858, 2017.
- [234] J. Wang, W. X. Zheng, and T. Chen. Identification of linear dynamic systems operating in a networked environment. *Automatica*, 45(12):2763–2772, 2009.
- [235] P. E. Wellstead. Reference signals for closed-loop identification. *International Journal of Control*, 26(6):945–962, 1977.
- [236] P. E. Wellstead. An instrumental product moment test for model order estimation. *Automatica*, 14(1):89–91, 1978.
- [237] J. S. Welsh and G. C. Goodwin. Finite sample properties of indirect nonparametric closed-loop identification. *IEEE Transactions on Automatic Control*, 47(8):1277–1292, 2002.
- [238] J. S. Welsh and G. C. Goodwin. Frequency localising basis functions for wide-band identification. In *2003 European Control Conference (ECC)*, pages 376–381, 2003.
- [239] J. S. Welsh, J. C. Agüero, and M. Alimir. Continuous-time system identification using indirect inference. In *15th IFAC Symposium on System Identification*, Saint Malo, France, pages 1169–1174, 2009.
- [240] X.-G. Xia. System identification using chirp signals and time-variant filters in the joint time-frequency domain. *IEEE Transactions on Signal Processing*, 45(8):2072–2084, 1997.
- [241] P. C. Young. The determination of the parameters of a dynamic process. *IERE Journal of Radio and Electronic Engineering*, 29:345–361, 1965.
- [242] P. C. Young. Process parameter estimation and self adaptive control. In P. H. Hammond (Ed.). *Theory of Self-Adaptive Control Systems*, pages 118–140. Springer, 1966.
- [243] P. C. Young. An instrumental variable method for real-time identification of a noisy process. *Automatica*, 6(2):271–287, 1970.

- [244] P. C. Young. Parameter estimation for continuous-time models—A survey. *Automatica*, 17(1):23–39, 1981.
- [245] P. C. Young. Recursive Estimation, Forecasting and Adaptive Control. *Control and Dynamic Systems*, 30(3):119–165, 1989.
- [246] P. C. Young. Optimal IV identification and estimation of continuous-time TF models. *IFAC Proceedings Volumes*, 35(1):109–114, 2002.
- [247] P. C. Young. An instrumental variable approach to ARMA model identification and estimation. *IFAC Proceedings Volumes*, 39(1):410–415, 2006.
- [248] P. C. Young. The refined instrumental variable method. *Journal Européen des Systèmes Automatisés*, 42(2-3):149–179, 2008.
- [249] P. C. Young. A three stage refined IV algorithm for closed loop identification and estimation. Technical Report TR/210a, Faculty of Science and Technology, Lancaster University, 2009.
- [250] P. C. Young. *Recursive Estimation and Time-Series Analysis: An Introduction for the Student and Practitioner*, 2nd Edition. Springer, 2012.
- [251] P. C. Young. Comment on ‘Projection-based identification algorithm for grey-box continuous-time models’ by Ichiro Maruta and Toshiharu Sugie. *Systems & Control Letters*, 69:62–64, 2014.
- [252] P. C. Young. Refined instrumental variable estimation: Maximum Likelihood optimization of a unified Box–Jenkins model. *Automatica*, 52:35–46, 2015.
- [253] P. C. Young. On the initialization of refined instrumental variable estimation algorithms. Technical report, Lancaster Environment Centre, Lancaster University, UK, 2016.
- [254] 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.
- [255] 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.
- [256] P. C. Young, A. Jakeman, and R. McMurtrie. An instrumental variable method for model order identification. *Automatica*, 16(3):281–294, 1980.
- [257] P. C. Young, H. Garnier, and M. Gilson. An optimal instrumental variable approach for identifying hybrid continuous-time Box-Jenkins models. *IFAC Proceedings Volumes*, 39(1):225–230, 2006.

- [258] P. C. Young, H. Garnier, and M. Gilson. Simple refined IV methods of closed-loop system identification. In *15th IFAC Symposium on System Identification*, Saint Malo, France, pages 1151–1156, 2009.
- [259] J. I. Yuz and G. C. Goodwin. *Sampled-Data Models for Linear and Nonlinear Systems*. Springer, 2014.
- [260] J. I. Yuz, J. Alfaro, J. C. Agüero, and G. C. Goodwin. Identification of continuous-time state-space models from non-uniform fast-sampled data. *IET Control Theory & Applications*, 5(7):842–855, 2011.
- [261] P. W. Zehna. Invariance of maximum likelihood estimators. *Annals of Mathematical Statistics*, 37(3):744, 1966.
- [262] Y. Zi-Zong. Schur complements and determinant inequalities. *Journal of Mathematical Inequalities*, 3(2):161–167, 2009.

