

# Sampling in Parametric and Nonparametric System Identification

**Citation for published version (APA):**

González, R., van Haren, M., Oomen, T. A. E., & Rojas, C. R. (2024). Sampling in Parametric and Nonparametric System Identification: Aliasing, Input Conditions, and Consistency. *IEEE Control Systems Letters*, 8, 2415-2420. Article 10737126. <https://doi.org/10.1109/LCSYS.2024.3487501>

**Document license:**

CC BY-NC-ND

**DOI:**

[10.1109/LCSYS.2024.3487501](https://doi.org/10.1109/LCSYS.2024.3487501)

**Document status and date:**

Published: 04/11/2024

**Document Version:**

Accepted manuscript including changes made at the peer-review stage

**Please check the document version of this publication:**

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

[Link to publication](#)

**General rights**

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

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

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

[www.tue.nl/taverne](http://www.tue.nl/taverne)

**Take down policy**

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

[openaccess@tue.nl](mailto:openaccess@tue.nl)

providing details and we will investigate your claim.

# Sampling in Parametric and Nonparametric System Identification: Aliasing, Input Conditions, and Consistency

Rodrigo A. González, Max van Haren, Tom Oomen and Cristian R. Rojas

**Abstract**—The sampling rate of input and output signals is known to play a critical role in the identification and control of dynamical systems. For slow-sampled continuous-time systems that do not satisfy the Nyquist-Shannon sampling condition for perfect signal reconstructability, careful consideration is required when identifying parametric and nonparametric models. In this letter, a comprehensive statistical analysis of estimators under slow sampling is performed. Necessary and sufficient conditions are obtained for unbiased estimates of the frequency response function beyond the Nyquist frequency, and it is shown that consistency of parametric estimators can be achieved even if input frequencies overlap after aliasing. Monte Carlo simulations confirm the theoretical properties.

**Index Terms**—Frequency-domain system identification; Undersampled systems; Frequency response function.

## I. INTRODUCTION

The exact reconstruction of continuous-time signals based on their samples is crucial when designing data-driven methods for estimating dynamical systems. According to the Nyquist-Shannon theorem, such reconstruction is possible if and only if the sampling rate is at least twice the bandwidth of the signal [1]. If the Nyquist-Shannon condition is not met, there exists an overlap of frequency components, known as aliasing. Since aliasing causes a loss of information for frequencies above the Nyquist frequency, system identification methods typically minimize the reconstruction error by using a small sampling period to ensure that the relevant dynamics remain below the Nyquist frequency [2].

When small sampling periods are not feasible, one way to mitigate aliasing effects is by multi-rate system identification. Methods have been developed to identify fast-dynamics models in multi-rate settings by exploiting the fast sampling of inputs [3], [4]. Polynomial transformation techniques [5] and state-space approaches [6] for dual-rate systems have also been proposed. These methods are inherently designed for discrete-time multi-rate models, and do not directly address continuous-time systems under slow sampling conditions.

For uniformly-sampled continuous-time systems, evidence suggests that slow sampling does not necessarily pose a

fundamental limitation for the identification of stochastic differential equations [7]. However, slow sampling can degrade performance of transfer function identification methods due to inaccuracies in estimating output derivatives [8]. Moreover, existing statistical analyses [9], [10] assume that the sampling period is fast enough to allow for a bijective transformation between continuous-time and discrete-time models [11], which may not hold in slow sampling scenarios.

Several methods for frequency-response function estimation beyond the Nyquist frequency have been developed. A practical nonparametric identification method beyond the Nyquist frequency is proposed in [12], which requires careful selection of input signals to avoid aliasing. Similarly, [13] develops a local polynomial approach to allow for disentanglement of the aliased contributions for slow-sampled systems assuming a smooth frequency response. While promising, these methods lack a rigorous statistical foundation, such as explicit conditions on the input signals for well-posedness of the methods, and optimality guarantees.

Although important developments have been made, a comprehensive statistical analysis of continuous-time system identification methods when the sampling violates the Nyquist-Shannon condition is still needed. This gap is addressed through our main contributions, which are as follows.

- C1 We prove that the least-squares estimator of the frequency response function is unbiased even when the Nyquist-Shannon sampling condition is violated, provided the input frequencies remain distinct after accounting for aliasing. We compute the covariance of this estimator, and provide an alternative interpretation of it as the quotient of output and input discrete-time Fourier transforms, evaluated at the input frequencies.
- C2 Using the least-squares estimator analyzed in Contribution C1, we provide an explicit relation, valid for finite sample size, between frequency domain and time domain identification in the slow-sampling regime.
- C3 We prove that the time-domain prediction error method is a consistent estimator of the parametric model under mild identifiability conditions, even if the input frequency lines are aliased, or if they overlap after aliasing.

The remainder of this letter is organized as follows. Section II introduces the identification setup. Section III presents the unbiased least-squares frequency response estimator, and Section IV links this estimator to parametric estimators and provides conditions for consistency. Simulation results are provided in Section V, and Section VI concludes this letter.

*Notation:* The Heaviside operator satisfies  $px(t) = \frac{dx(t)}{dt}$ .

This work was partly supported by the Swedish Research Council under contract number 2023-05170, and by the ECSEL Joint Undertaking under grant agreement 101007311 (IMOCO4.E). The Joint Undertaking receives support from the European Union Horizon 2020 research and innovation programme.

R. A. González, M. van Haren and T. Oomen are with the Control Systems Technology Section, Department of Mechanical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands. T. Oomen is also with the Delft Center for Systems and Control, Delft University of Technology, Delft, The Netherlands. C. R. Rojas is with the Division of Decision and Control Systems, KTH Royal Institute of Technology, Stockholm, Sweden. E-mail of the corresponding author: r.a.gonzalez@tue.nl.

All vectors, matrices, and multivariate signals are written in bold, and vectors are column vectors, unless transposed. The conjugate transpose of a vector  $\mathbf{x}$  is denoted as  $\mathbf{x}^H$ , and the diagonal matrix  $\text{diag}(\mathbf{x})$  or  $\text{diag}(x_1, \dots, x_n)$  is formed by the elements in  $\mathbf{x} = [x_1, \dots, x_n]^T$ . The identity matrix of size  $m$  is denoted as  $\mathbf{I}_m$ , and the null space of a matrix  $\mathbf{A}$  is denoted as  $\text{Ker}(\mathbf{A})$ . The Kronecker delta centered at  $n = 0$  is denoted as  $\delta(n)$ , and  $b \bmod m$  denotes the modulo operation. The expression  $\mathbb{E}\{x(kh)\} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N \mathbb{E}\{x(kh)\}$  denotes the expectation of a quasi-stationary signal  $x(kh)$ .

## II. PROBLEM FORMULATION

Consider a single-input single-output, linear time-invariant (LTI), asymptotically stable, continuous-time system

$$x(t) = G_0(p)u(t), \quad (1)$$

where  $x(t) \in \mathbb{R}$  is the output of the LTI system  $G_0(p)$  when subject to the input  $u(t) \in \mathbb{R}$ . We are interested in estimating  $G_0(p)$  from input and output data. Thus, we excite the system with a continuous-time multisine input of the form

$$u(t) = a_0 + \sum_{\ell=1}^M a_\ell \cos(\omega_\ell t + \phi_\ell), \quad (2)$$

where the angular frequencies  $\omega_\ell$  are positive and sorted in strictly increasing order, and  $\{a_\ell\}_{\ell=0}^M$  are all different from zero. We assume perfect knowledge of the continuous-time input signal and the absence of a hold device between the input source and the system, which implies that  $u(t)$  is a band-limited signal with bandwidth  $\omega_M$ .

We retrieve a noisy measurement of the output  $x(t)$  at equidistant time points  $t = h, 2h, \dots, Nh$ :

$$y(kh) = x(kh) + v(kh), \quad k = 1, \dots, N, \quad (3)$$

where  $v(kh) \in \mathbb{R}$  is a zero-mean i.i.d. white noise of variance  $\sigma^2$ . The key assumption in this work is that the sampling period  $h$  does not satisfy the Nyquist-Shannon criterion for perfect reconstructability of the input signal based only on its samples, i.e.,  $h \geq \pi/\omega_M$ . We refer to this experimental condition as the slow-sampling scenario.

Our goal is to obtain necessary and sufficient conditions for the identifiability of  $G_0(p)$ , and to derive consistent nonparametric and parametric estimators for this sampling regime.

*Remark 2.1:* A similar analysis to the one presented in this letter can be performed if a zero-order hold device is present, where a discrete-time formulation, not covered here, would be preferred.

## III. A LEAST-SQUARES ESTIMATE FOR THE FREQUENCY RESPONSE FUNCTION

A nonparametric model for the system  $G_0(p)$  can be obtained via the estimation of its frequency response function [14]. Since the input in (2) only excites the frequencies  $\omega = 0, \pm\omega_1, \dots, \pm\omega_M$ , we focus on estimating the frequency

points contained in the vector

$$\mathbf{G}_0^f := [G_0(0), G_0(-i\omega_1), G_0(i\omega_1), \dots, G_0(-i\omega_M), G_0(i\omega_M)]^T \in \mathbb{C}^{2M+1}. \quad (4)$$

We assume the output is measured in a stationary regime, i.e., after all distortions caused by transient effects have decayed to zero. Then, the noiseless output  $x(kh)$  is given by

$$\begin{aligned} x(kh) &= G_0(0)a_0 + \sum_{\ell=1}^M a_\ell |G_0(i\omega_\ell)| \cos(\omega_\ell kh + \phi_\ell + \angle G_0(i\omega_\ell)) \\ &= \boldsymbol{\zeta}^H(kh) \mathbf{G}_0^f, \end{aligned} \quad (5)$$

where

$$\boldsymbol{\zeta}(kh) = \begin{bmatrix} a_0 \\ \frac{a_1}{2} e^{i(\omega_1 kh + \phi_1)} \\ \frac{a_1}{2} e^{-i(\omega_1 kh + \phi_1)} \\ \vdots \\ \frac{a_M}{2} e^{i(\omega_M kh + \phi_M)} \\ \frac{a_M}{2} e^{-i(\omega_M kh + \phi_M)} \end{bmatrix}. \quad (6)$$

Thus, the following least-squares estimator to compute the frequency response vector  $\mathbf{G}_0^f$  is considered:

$$\hat{\mathbf{G}}^f = \left[ \sum_{k=1}^N \boldsymbol{\zeta}(kh) \boldsymbol{\zeta}^H(kh) \right]^{-1} \left[ \sum_{k=1}^N \boldsymbol{\zeta}(kh) y(kh) \right]. \quad (7)$$

### A. Statistical properties of the least-squares estimator

For the least-squares estimator in (7) to be well-defined in the slow-sampling scenario, we require the input frequencies  $\{\omega_\ell\}_{\ell=1}^M$  to not overlap after accounting for aliasing. This key condition is made explicit in Assumption 3.1.

*Assumption 3.1 (Non-overlapping condition):* The input frequencies  $\{\omega_\ell\}_{\ell=1}^M$  satisfy

$$\begin{cases} \omega_\ell \pm \omega_\tau \neq \frac{2n\pi}{h} & \text{for all } \ell, \tau = 1, \dots, M; \ell \neq \tau; n \in \mathbb{Z}, \\ \omega_\ell \neq \frac{n\pi}{h} & \text{for all } \ell = 1, \dots, M; n \in \mathbb{Z}. \end{cases} \quad (8)$$

Note that for each pair of frequencies below the Nyquist frequency, the inequalities in (8) are satisfied.

The well-posedness of the estimator defined in (7) and its unbiasedness when the sampling period does not satisfy the Nyquist-Shannon criterion are investigated in Theorem 3.1. These results constitute Contribution C1 of this letter.

*Theorem 3.1:* Consider the sampled input and output signals  $\{u(kh), y(kh)\}_{k=1}^N$ , where  $u(t)$  is given by (2),  $y(kh)$  is measured in a stationary regime,  $h \geq \pi/\omega_M$ , and  $N > 2M$ . Then, (7) is well-defined and is an unbiased estimator of the frequency response vector  $\mathbf{G}_0^f$  if and only if Assumption 3.1 holds. In such case, its covariance is given by

$$\text{Cov}\{\hat{\mathbf{G}}^f\} = \sigma^2 \left[ \sum_{k=1}^N \boldsymbol{\zeta}(kh) \boldsymbol{\zeta}^H(kh) \right]^{-1}. \quad (9)$$

*Proof:* The least-squares estimator (7) is well-defined if and only if its associated normal matrix

$$\mathbf{Z} := \sum_{k=1}^N \boldsymbol{\zeta}(kh) \boldsymbol{\zeta}^H(kh) \quad (10)$$

is nonsingular. Let  $\mathbf{w} \in \mathbb{C}^{2M+1}$  be arbitrary. We have  $\mathbf{w}^H \mathbf{Z} \mathbf{w} = 0$  if and only if  $\mathbf{K} \mathbf{w} = \mathbf{0}$ , where  $\mathbf{K} = [\zeta(h), \dots, \zeta(Nh)]^H$ . The matrix  $\mathbf{K}$  can be decomposed as  $\mathbf{K} = \mathbf{V} \text{diag}(\zeta^H(h))$ , where  $\mathbf{V} \in \mathbb{C}^{N \times (2M+1)}$  is a rectangular Vandermonde matrix given by

$$\mathbf{V} = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ 1 & e^{-i\omega_1 h} & e^{i\omega_1 h} & \cdots & e^{-i\omega_M h} & e^{i\omega_M h} \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 1 & e^{-i\omega_1(N-1)h} & e^{i\omega_1(N-1)h} & \cdots & e^{-i\omega_M(N-1)h} & e^{i\omega_M(N-1)h} \end{bmatrix}.$$

Vandermonde matrices such as the one above are known to have full column rank if  $N \geq 2M + 1$  and the complex numbers  $1, e^{\pm i\omega_1 h}, \dots, e^{\pm i\omega_M h}$  are distinct [15, Sec. 0.9.11]. This necessary and sufficient condition is equivalent to (8). Under the assumptions in (8),  $\mathbf{K}$  has full column rank and thus  $\mathbf{K} \mathbf{w} = \mathbf{0}$  implies  $\mathbf{w} = \mathbf{0}$ , which means that  $\mathbf{Z}$  is positive definite and hence nonsingular.

For the unbiasedness result, we compute the expected value of the least-squares estimator in (7) as

$$\mathbb{E}\{\hat{\mathbf{G}}^f\} = \mathbf{Z}^{-1} \left[ \sum_{k=1}^N \zeta(kh) x(kh) \right] = \mathbf{G}_0^f,$$

where the last equality is due to (5). Finally, by noting that

$$\hat{\mathbf{G}}^f - \mathbb{E}\{\hat{\mathbf{G}}^f\} = \mathbf{Z}^{-1} \left[ \sum_{k=1}^N \zeta(kh) v(kh) \right],$$

we have, using the fact that  $\{v(kh)\}$  is white noise,

$$\begin{aligned} \text{Cov}\{\hat{\mathbf{G}}^f\} &= \mathbf{Z}^{-1} \sum_{k=1}^N \sum_{\tau=1}^N \zeta(kh) \mathbb{E}\{v(kh)v(\tau h)\} \zeta^H(\tau h) \mathbf{Z}^{-1} \\ &= \sigma^2 \mathbf{Z}^{-1} \sum_{k=1}^N \zeta(kh) \zeta^H(kh) \mathbf{Z}^{-1} \\ &= \sigma^2 \mathbf{Z}^{-1}. \end{aligned} \quad \blacksquare$$

*Corollary 3.1:* Consider the same experimental conditions as in Theorem 3.1, and assume that Assumption 3.1 holds. Then, (7) is a consistent estimator of  $\mathbf{G}_0^f$  in (4). Furthermore, the frequency response estimates  $\hat{G}(0), \hat{G}(\pm i\omega_\ell), \ell = 1, \dots, M$  are asymptotically mutually uncorrelated, and the asymptotic covariance of  $\hat{\mathbf{G}}^f$  is given by

$$\lim_{N \rightarrow \infty} N \text{Cov}\{\hat{\mathbf{G}}^f\} = \sigma^2 \text{diag} \left( \frac{1}{a_0^2}, \frac{4}{a_1^2}, \frac{4}{a_1^2}, \dots, \frac{4}{a_M^2}, \frac{4}{a_M^2} \right). \quad (11)$$

*Proof:* An estimator is consistent if it is unbiased and its covariance decays to zero [16, p. 54]. Unbiasedness follows from Theorem 3.1. We now compute its asymptotic covariance. From (9), the asymptotic covariance is given by  $\sigma^2 \bar{\mathbf{Z}}^{-1}$ , where  $\bar{\mathbf{Z}} = \lim_{N \rightarrow \infty} \mathbf{Z}/N$ , with  $\mathbf{Z}$  being defined in (10). Calculating the entries of  $\bar{\mathbf{Z}}_{ij}$  of  $\bar{\mathbf{Z}}$ , we find  $\bar{\mathbf{Z}}_{11} = a_0^2$  and  $\bar{\mathbf{Z}}_{(2\ell)(2\ell)} = \bar{\mathbf{Z}}_{(2\ell+1)(2\ell+1)} = a_\ell^2/4$  for  $\ell \in \{1, 2, \dots, M\}$ . Lastly, every non-diagonal element of  $\bar{\mathbf{Z}}$  is of the form

$$\lim_{N \rightarrow \infty} \frac{C}{N} \sum_{k=1}^N e^{i(\tilde{\omega}kh + \tilde{\psi})} \quad (12)$$

for some real constants  $C, \tilde{\psi}$ . Due to the condition in (8), the constant  $\tilde{\omega}$ , which is the difference or sum of frequencies  $\omega_\ell$  (including the zero frequency), is not a multiple of  $2\pi/h$ . The following chain of inequalities is satisfied for  $N \geq 2M + 1$ :

$$0 \leq \left| \frac{C}{N} \sum_{k=1}^N e^{i(\tilde{\omega}kh + \tilde{\psi})} \right| = \left| \frac{C \sin(\frac{1}{2}[N+1]\tilde{\omega}h)}{N \sin(\frac{1}{2}\tilde{\omega}h)} \right| \leq \left| \frac{C}{N \sin(\frac{1}{2}\tilde{\omega}h)} \right|.$$

Since  $\tilde{\omega} \neq 2n\pi/h, n \in \mathbb{Z}$ , the denominator of the upper bound is nonzero. The squeeze theorem [17, Theorem 3.3.6] then implies that the limit in (12) is equal to zero. This means that all the off-diagonal elements of  $\bar{\mathbf{Z}}$  are zero, leading to (11). Finally, since the covariance of (7) decays to zero, it is consistent, concluding the proof.  $\blacksquare$

In summary, unbiased estimates of the frequency response function beyond the Nyquist frequency  $\pi/h$  can be obtained if the input frequencies do not overlap after aliasing, as indicated by the condition in (8). Frequencies equal to integer multiples of the Nyquist frequency, i.e.,  $\omega_\ell = n\pi/h$ , alias to the zero frequency if  $n$  is even, or a single frequency line if  $n$  is odd. In the latter case, after minor adjustments to the proof of Theorem 3.1, it can be shown that only the real part of  $G_0(i n\pi/h)$  can be estimated unbiasedly. Note that only one experiment is needed for the least-squares estimator in (7) to be computed, and no assumptions on the smoothness of the frequency response of the system are required.

## B. Frequency-domain interpretation

The least-squares estimator in (7) has a frequency-domain interpretation. Via Parseval's theorem [1], (7) is equivalent to

$$\hat{\mathbf{G}}^f = \left[ \sum_{n=1}^N \Psi[e^{i\frac{2\pi n}{N}}] \Psi^H[e^{i\frac{2\pi n}{N}}] \right]^{-1} \left[ \sum_{n=1}^N \Psi[e^{i\frac{2\pi n}{N}}] Y[e^{-i\frac{2\pi n}{N}}] \right], \quad (13)$$

where the discrete-time Fourier transforms (DTFTs) of  $\zeta(kh)$  and  $y(kh)$  are respectively given by

$$\Psi[e^{i\omega h}] = \sum_{k=1}^N \zeta(kh) e^{-ikh\omega}, \quad Y[e^{i\omega h}] = \sum_{k=1}^N y(kh) e^{-ikh\omega}. \quad (14)$$

To obtain an explicit formula for  $\hat{\mathbf{G}}^f$ , we introduce the following assumption.

*Assumption 3.2 (No spectral leakage):* There is no spectral leakage when computing  $\Psi[e^{i\omega h}]$ , i.e.,  $Nh$  is a multiple of the least common multiple of  $\{2\pi/\omega_\ell\}_{\ell=1}^M$ .

*Theorem 3.2:* Consider the sampled input and output signals  $\{u(kh), y(kh)\}_{k=1}^N$ , where  $u(t)$  is given by (2),  $y(kh)$  is measured in a stationary regime,  $h \geq \pi/\omega_M$ , and  $N > 2M$ . Assume that Assumptions 3.1 and 3.2 hold. Then, the least-squares estimator (7) can be computed as

$$\hat{\mathbf{G}}^f = \begin{bmatrix} \frac{Y[e^{i0}]}{U[e^{i0}]}, \frac{Y[e^{-i\omega_1 h}]}{U[e^{-i\omega_1 h}]}, \frac{Y[e^{i\omega_1 h}]}{U[e^{i\omega_1 h}]} \\ \vdots, \frac{Y[e^{-i\omega_M h}]}{U[e^{-i\omega_M h}]}, \frac{Y[e^{i\omega_M h}]}{U[e^{i\omega_M h}]} \end{bmatrix}^T, \quad (15)$$

where the DTFT of  $y(kh)$  is given by (14), and

$$U[e^{i0}] = Na_0, U[e^{\pm i\omega_\ell h}] = N \frac{a_\ell}{2} e^{\pm i\phi_\ell}, \ell = 1, \dots, M.$$

Furthermore, the frequency response estimates  $\hat{G}(0), \hat{G}(\pm i\omega_\ell), \ell = 1, \dots, M$  are mutually uncorrelated for any  $N \geq 2M + 1$ .

*Proof:* The discrete Fourier transform of  $\zeta(kh)$  is a  $N$ -periodic function described in one period by

$$\Psi[e^{i\frac{2\pi n}{N}}] = N \begin{bmatrix} a_0 \delta(n \bmod N) \\ \frac{a_1}{2} e^{i\phi_1} \delta((n - \frac{Nh\omega_1}{2\pi}) \bmod N) \\ \frac{a_1}{2} e^{-i\phi_1} \delta((n + \frac{Nh\omega_1}{2\pi}) \bmod N) \\ \vdots \\ \frac{a_M}{2} e^{i\phi_M} \delta((n - \frac{Nh\omega_M}{2\pi}) \bmod N) \\ \frac{a_M}{2} e^{-i\phi_M} \delta((n + \frac{Nh\omega_M}{2\pi}) \bmod N) \end{bmatrix}, \quad (16)$$

where we have used Assumption 3.2. Since (8) also holds, the matrix being inverted in (13) admits the expression

$$\sum_{n=1}^N \Psi[e^{i\frac{2\pi n}{N}}] \Psi^H[e^{i\frac{2\pi n}{N}}] = N^2 \text{diag} \left( a_0^2, \frac{a_1^2}{4}, \frac{a_1^2}{4}, \dots, \frac{a_M^2}{4}, \frac{a_M^2}{4} \right). \quad (17)$$

Via Parseval's theorem applied to (17), we find that  $\mathbf{Z}$  in (10) is also a diagonal matrix, and thus the correlation between frequency response estimates is zero.

Similarly, using (16) and the  $N$ -periodicity of  $Y[e^{i\frac{2\pi n}{N}}]$ ,

$$\sum_{n=1}^N \Psi[e^{i\frac{2\pi n}{N}}] Y[e^{-i\frac{2\pi n}{N}}] = N \begin{bmatrix} a_0 Y[e^{i0}] \\ \frac{a_1}{2} e^{i\phi_1} Y[e^{-i\omega_1 h}] \\ \frac{a_1}{2} e^{-i\phi_1} Y[e^{i\omega_1 h}] \\ \vdots \\ \frac{a_M}{2} e^{i\phi_M} Y[e^{-i\omega_M h}] \\ \frac{a_M}{2} e^{-i\phi_M} Y[e^{i\omega_M h}] \end{bmatrix}. \quad (18)$$

The computation of (13) via (17) and (18) leads to (15), which is what we aimed to prove.  $\blacksquare$

Theorem 3.2 states that an unbiased estimate of the frequency response at the frequencies above the Nyquist frequency can be computed directly from the empirical transfer function estimate (ETF) [18], evaluated at the aliased frequencies. This insight provides a theoretical basis for interpreting the ETF beyond the Nyquist frequency when the input suffers from aliasing.

*Remark 3.1:* The proof of Theorem 3.2 shows that  $\mathbf{Z}$  in (10) is a diagonal matrix when Assumption 3.2 holds. Thus, for a fixed sample size, the variance of the frequency response function estimates does not increase with increasing sampling periods, as long as no spectral leakage is introduced.

The frequency response function estimator in (7) or (15) can be used to compute parametric estimators that are accurate above the Nyquist frequency, as seen next.

#### IV. NONPARAMETRIC TO PARAMETRIC ESTIMATORS

Once the frequency response function estimate (7) has been computed, we propose a parametric model  $G(p, \boldsymbol{\theta})$  for the system  $G_0(p)$ , where  $\boldsymbol{\theta} \in \mathbb{R}^{n_\theta}$  contains the coefficients

associated with, e.g., numerator and denominator polynomials of an unfactored rational transfer function, or coefficients associated with a modal parametrization of the system [19]. Throughout this section, we assume that there exists a unique vector  $\boldsymbol{\theta}_0 \in \mathbb{R}^{n_\theta}$  such that  $G(p, \boldsymbol{\theta}_0) = G_0(p)$ .

For the sequel, we define the vector of frequency responses of the parametric model as

$$\mathbf{G}^f(\boldsymbol{\theta}) := [G(0, \boldsymbol{\theta}), G(-i\omega_1, \boldsymbol{\theta}), G(i\omega_1, \boldsymbol{\theta}), \dots, G(-i\omega_M, \boldsymbol{\theta}), G(i\omega_M, \boldsymbol{\theta})]^T. \quad (19)$$

A parametric model can be obtained from the least-squares estimate (7) by curve fitting in the frequency domain. This method, when weighted correctly, corresponds exactly to the time-domain prediction error method with a multisine predictor that suffers from aliasing. This result, contained in Theorem 4.1, constitutes Contribution C2 of this letter.

*Theorem 4.1:* Consider the sampled input and output signals  $\{u(kh), y(kh)\}_{k=1}^N$ , where  $u(t)$  is given by (2),  $y(kh)$  is measured in a stationary regime,  $h \geq \pi/\omega_M$ ,  $N > 2M$ , and assume that Assumption 3.1 holds. The following two cost functions share the same global minima:

$$V_f(\boldsymbol{\theta}) = (\mathbf{G}^f(\boldsymbol{\theta}) - \hat{\mathbf{G}}^f)^H [\text{Cov}\{\hat{\mathbf{G}}^f\}]^{-1} (\mathbf{G}^f(\boldsymbol{\theta}) - \hat{\mathbf{G}}^f), \quad (20)$$

$$V_t(\boldsymbol{\theta}) = \sum_{k=1}^N (y(kh) - \hat{y}(kh, \boldsymbol{\theta}))^2, \quad (21)$$

where  $\hat{\mathbf{G}}^f$  is from (7), and the predictor  $\hat{y}(kh, \boldsymbol{\theta})$  is given by

$$\hat{y}(kh, \boldsymbol{\theta}) = G(0, \boldsymbol{\theta})a_0 + \sum_{\ell=1}^M a_\ell |G(i\omega_\ell, \boldsymbol{\theta})| \cos(\omega_\ell kh + \phi_\ell + \angle G(i\omega_\ell, \boldsymbol{\theta})). \quad (22)$$

*Proof:* When analyzed as a function of  $\boldsymbol{\theta}$ , except for an additive constant,  $V_f(\boldsymbol{\theta})$  is proportional to

$$V_f(\boldsymbol{\theta}) \propto [\mathbf{G}^f(\boldsymbol{\theta})]^H \mathbf{Z} \mathbf{G}^f(\boldsymbol{\theta}) - 2\text{Re}\{(\hat{\mathbf{G}}^f)^H \mathbf{Z} \mathbf{G}^f(\boldsymbol{\theta})\}, \quad (23)$$

where we used the fact that  $\text{Cov}\{\hat{\mathbf{G}}^f\} = \sigma^2 \mathbf{Z}^{-1}$ , with  $\mathbf{Z}$  defined as in (10). First, by leveraging the result in (5),

$$[\mathbf{G}^f(\boldsymbol{\theta})]^H \mathbf{Z} \mathbf{G}^f(\boldsymbol{\theta}) = \sum_{k=1}^N |\zeta^H(kh) \mathbf{G}^f(\boldsymbol{\theta})|^2 = \sum_{k=1}^N \hat{y}^2(kh, \boldsymbol{\theta}).$$

On the other hand, the identity  $(\hat{\mathbf{G}}^f)^H \mathbf{Z} = \sum_{k=1}^N y(kh) \zeta^H(kh)$  leads to

$$\text{Re}\{(\hat{\mathbf{G}}^f)^H \mathbf{Z} \mathbf{G}^f(\boldsymbol{\theta})\} = \sum_{k=1}^N y(kh) \hat{y}(kh, \boldsymbol{\theta}).$$

Replacing  $[\mathbf{G}^f(\boldsymbol{\theta})]^H \mathbf{Z} \mathbf{G}^f(\boldsymbol{\theta})$  and  $\text{Re}\{(\hat{\mathbf{G}}^f)^H \mathbf{Z} \mathbf{G}^f(\boldsymbol{\theta})\}$  in (23) and completing the square leads to  $V_f(\boldsymbol{\theta}) \propto V_t(\boldsymbol{\theta})$ . Thus, these cost functions share the same global minima.  $\blacksquare$

Theorem 4.1 indicates that, provided the input frequency conditions in (8) are satisfied and  $N \geq 2M + 1$ , there is an equivalence between frequency domain and time domain parametric estimation for the slow-sampling scenario. The expressions also coincide with the standard prediction error method and maximum likelihood estimators for the Gaussian

noise case when the sampling rate is higher than the Nyquist frequency [2]. The parametric model can be obtained via numerical minimization of (20) or (21). For example, the minimization of (21) can be achieved using Gauss-Newton (GN) iterations [2, Sec. 10.2]:

$$\boldsymbol{\theta}_{j+1} = \boldsymbol{\theta}_j + \left[ \sum_{k=1}^N \hat{\varphi}_j(kh) \hat{\varphi}_j^\top(kh) \right]^{-1} \left[ \sum_{k=1}^N \hat{\varphi}_j(kh) \varepsilon_j(kh) \right],$$

where  $\hat{\varphi}_j(kh) = (\partial \hat{y}(kh, \boldsymbol{\theta}) / \partial \boldsymbol{\theta})|_{\boldsymbol{\theta}=\boldsymbol{\theta}_j}$  and  $\varepsilon_j(kh) = y(kh) - \hat{y}(kh, \boldsymbol{\theta}_j)$ , with  $\hat{y}(kh, \boldsymbol{\theta}_j)$  being defined in (22). As in the standard sampling case  $h < \pi/\omega_M$ , caution must be taken when initializing the iterative procedure, as convergence to a global minimum is not guaranteed [2].

*Remark 4.1:* Under Assumption 3.2 on the absence of spectral leakage, applying the result in (17) in  $V_f(\boldsymbol{\theta})$  gives

$$V_f(\boldsymbol{\theta}) \propto a_0^2 |\hat{G}(0) - G(0, \boldsymbol{\theta})|^2 + \sum_{\ell=1}^M \frac{a_\ell^2}{2} |\hat{G}(i\omega_\ell) - G(i\omega_\ell, \boldsymbol{\theta})|^2,$$

i.e., the least-squares cost function of the frequency response estimates, weighted by the magnitude of the input spectrum.

In the slow-sampling scenario, the prediction error method proposed in Theorem 4.1 provides consistent estimators of the parametric model  $G(p, \boldsymbol{\theta})$  under mild conditions. In other words, the parametric estimator that minimizes (20) and (21) converges to  $\boldsymbol{\theta}_0$  with probability 1. Interestingly, this property holds even if the aliasing constraint in Assumption 3.1 is not satisfied, i.e., if there are input frequencies that overlap after aliasing. Theorem 4.2 formalizes this result, which is Contribution C3 of this letter.

*Theorem 4.2:* Consider the sampled input and output signals  $\{u(kh), y(kh)\}_{k=1}^N$ , where  $u(t)$  is given by (2),  $y(kh)$  is measured in a stationary regime, and  $h \geq \pi/\omega_M$ . Assume that the model satisfies the condition

$$\mathbf{G}^f(\boldsymbol{\theta}) - \mathbf{G}_0^f \in \text{Ker}(\boldsymbol{\Phi}_\zeta) \implies \boldsymbol{\theta} = \boldsymbol{\theta}_0, \quad (24)$$

where  $\boldsymbol{\Phi}_\zeta := \overline{\mathbb{E}}\{\zeta(kh)\zeta^H(kh)\}$ . Then, as  $N$  tends to infinity, the parametric estimator that minimizes (21) among all asymptotically stable models converges to  $\boldsymbol{\theta}_0$  with probability 1.

*Proof:* By Lemma 8.2 of [2], the criterion  $V_f(\boldsymbol{\theta})/N$  converges uniformly and with probability 1 as  $N \rightarrow \infty$  to

$$\bar{V}(\boldsymbol{\theta}) = \overline{\mathbb{E}}\{(y(kh) - \hat{y}(kh, \boldsymbol{\theta}))^2\}.$$

Due to (5) and  $\hat{y}(kh, \boldsymbol{\theta}) = \zeta^H(kh)\hat{\mathbf{G}}^f(\boldsymbol{\theta})$ , we can compute

$$\begin{aligned} \bar{V}(\boldsymbol{\theta}) &= \overline{\mathbb{E}}\{(x(kh) - \hat{y}(kh, \boldsymbol{\theta}))^2\} + \sigma^2 \\ &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N |\zeta^H(kh)(\mathbf{G}^f(\boldsymbol{\theta}) - \mathbf{G}_0^f)|^2 + \sigma^2 \\ &= (\mathbf{G}^f(\boldsymbol{\theta}) - \mathbf{G}_0^f)^H \boldsymbol{\Phi}_\zeta (\mathbf{G}^f(\boldsymbol{\theta}) - \mathbf{G}_0^f) + \sigma^2. \end{aligned}$$

This cost function is bounded below by  $\sigma^2$ , with equality if and only if  $\boldsymbol{\Phi}_\zeta(\mathbf{G}^f(\boldsymbol{\theta}) - \mathbf{G}_0^f) = \mathbf{0}$ . By the identifiability condition in (24), the minimizer of  $\bar{V}(\boldsymbol{\theta})$  is unique and is equal to  $\boldsymbol{\theta}_0$ , concluding the proof. ■

Following similar steps from the proof of Corollary 3.1,

the matrix  $\boldsymbol{\Phi}_\zeta$  has a rank equal to the number of distinct input frequency lines (after accounting for aliasing) within the fundamental frequency band  $[-\pi/h, \pi/h)$ . If (8) is satisfied,  $\boldsymbol{\Phi}_\zeta$  is nonsingular, and (24) reduces to the standard identifiability condition  $\mathbf{G}^f(\boldsymbol{\theta}) = \mathbf{G}_0^f \implies \boldsymbol{\theta} = \boldsymbol{\theta}_0$ . Therefore, when no input frequencies overlap after aliasing,  $n_\theta \leq 2M + 1$  ensures identifiability for standard model parametrizations and consistency of the prediction error method estimator. If input frequency lines overlap, consistency is still guaranteed as long as the dimension of  $\boldsymbol{\theta}$  does not exceed the number of unique non-overlapping input frequency lines.

## V. SIMULATION STUDIES

In this section we verify the main theoretical findings via Monte Carlo simulations. The tests are conducted on a benchmark system known as the Rao-Garnier system [20]:

$$G_0(p) = \frac{-6400p + 1600}{p^4 + 5p^3 + 408p^2 + 416p + 1600}, \quad (25)$$

with  $\boldsymbol{\theta}_0 = [1600, 416, 408, 5, 1600, -6400]^\top$  as true parameter vector. The parametric model we consider is given by

$$G(p, \boldsymbol{\theta}) = \frac{b_1p + b_0}{p^4 + a_3p^3 + a_2p^2 + a_1p + a_0},$$

where  $\boldsymbol{\theta} = [a_0, a_1, a_2, a_3, b_0, b_1]^\top$ . The system in (25) is excited with an input as in (2), and the noise variance corresponds to an output signal-to-noise ratio of 10 [dB].

### A. Nonparametric system identification

To assess the statistical performance of the nonparametric estimator in (7), we conduct 2000 Monte Carlo runs with varying noise realizations. Each simulation generates  $N = 2000$  output data samples using a fixed input with unit amplitudes, random phases, and  $M = 13$  nonzero frequencies between 0.1 and 30 [rad/s]. The sampling period is set to  $h = 0.5$  [s], which is 100 times larger than the typical sampling period for this system [20]. Assumption 3.1 is satisfied under these experimental conditions.

Figure 1 shows the Bode plot of the true system and the mean value of the frequency response estimate at each input frequency. As a metric of dispersion of the estimates, we have also plotted the 95% confidence intervals associated with the magnitude and phase of the frequency response estimate. The estimator exhibits no noticeable bias at any frequency, including those above the Nyquist frequency.

### B. Parametric system identification

In this study we verify the consistency of the estimator derived from minimizing  $V_f(\boldsymbol{\theta})$  in (21). Specifically, we examine the sample mean and mean-square error (MSE) of the estimated parameters across 60 logarithmically spaced sample sizes ranging from  $N = 2 \cdot 10^3$  to  $2 \cdot 10^5$ . For each  $N$ , 1000 Monte Carlo runs are performed using a fixed random-phase multisine input with unit amplitudes and nonzero frequencies  $\pi/3, \pi, 7\pi/2$ , and  $5\pi$  [rad/s]. The frequency  $5\pi$  [rad/s] overlaps with  $\pi$  [rad/s], violating Assumption 3.1. The maximum number of GN iterations is set to 100, and

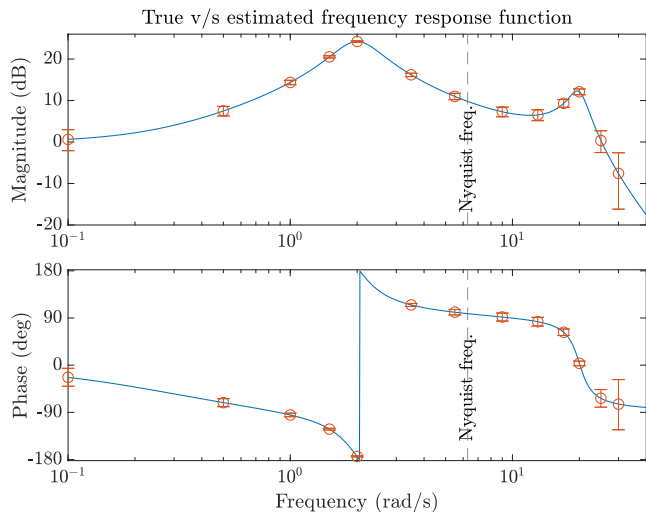


Fig. 1. Bode plot of the system (blue), and the mean value of the magnitude and phase of the estimated frequency response function via least-squares, with its 95% confidence interval (red). The least-squares approach provides an unbiased estimate of the system over the Nyquist frequency.

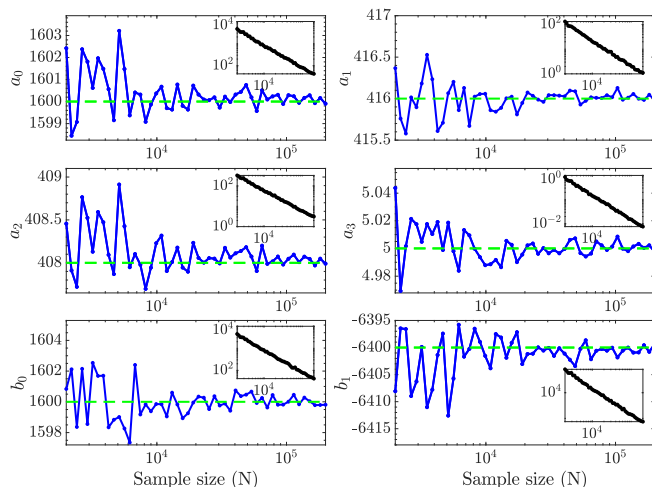


Fig. 2. Empirical means of the parameter estimates (blue), with corresponding MSEs (log-log subplots, black lines). Parameters  $a_i$  and  $b_i$  correspond to the denominator and numerator coefficients of  $G(p, \theta)$ , respectively. All estimates converge to their true values (green dashed lines), and MSEs decay as  $1/N$ , indicating the consistency of the estimator minimizing (21).

the initial parameter vector is randomly generated, deviating from the true parameter vector by at most 10% in each entry.

The sample means for increasing sample sizes are reported in Figure 2, where each plot contains a subplot with the MSE of each estimated parameter. For increasing values of  $N$ , all the parameter estimates converge to their true values, and the sample MSEs decay to zero as  $1/N$ . The estimator is consistent despite the input suffering from aliasing and frequency line overlapping, since the identifiability condition in (24) applied to this example indicates that a maximum number of 7 parameters can be estimated consistently. These results align with the findings in Theorem 4.2.

## VI. CONCLUSIONS

The results in this letter provide statistical guarantees for nonparametric and parametric identification methods for

slow-sampled continuous-time systems. Unbiased frequency response estimates can be obtained if the input frequencies avoid overlap after aliasing. For a fixed sample size and no spectral leakage, the covariance of these estimates does not deteriorate for increasing sampling periods. For parametric modelling, the prediction error method is consistent if the number of frequency lines below the Nyquist sampling rate is not less than the number of model parameters. Our methods and proofs extend to standard sampling rates ( $h < \pi/\omega_M$ ), where aliasing is absent, aligning with standard identifiability conditions for the prediction error method [2]. Future work includes extending these results to MIMO system identification, where multiple experiments must be considered to identify the frequency response function at each input frequency [14].

## REFERENCES

- [1] A. V. Oppenheim, A. S. Willsky, and S. H. Nawab, *Signals & Systems*, 2nd Edition. Pearson Education, 1997.
- [2] L. Ljung, *System Identification: Theory for the User*, 2nd Edition. Prentice-Hall, 1999.
- [3] Y. Zhu, H. Telkamp, J. Wang, and Q. Fu, "System identification using slow and irregular output samples," *Journal of Process Control*, vol. 19, no. 1, pp. 58–67, 2009.
- [4] Y. Liu, F. Ding, and Y. Shi, "An efficient hierarchical identification method for general dual-rate sampled-data systems," *Automatica*, vol. 50, no. 3, pp. 962–970, 2014.
- [5] J. Ding, Y. Shi, H. Wang, and F. Ding, "A modified stochastic gradient based parameter estimation algorithm for dual-rate sampled-data systems," *Digital Signal Processing*, vol. 20, no. 4, pp. 1238–1247, 2010.
- [6] M. Yin, A. Iannelli, and R. S. Smith, "Subspace identification of linear time-periodic systems with periodic inputs," *IEEE Control Systems Letters*, vol. 5, no. 1, pp. 145–150, 2020.
- [7] K. J. Åström, "On the choice of sampling rates in parametric identification of time series," *Information Sciences*, pp. 273–278, 1969.
- [8] L. Ljung, "Experiments with identification of continuous time models," in *15th IFAC Symposium on System Identification*, Saint Malo, France, pp. 1175–1180, 2009.
- [9] S. Pan, R. A. González, J. S. Welsh, and C. R. Rojas, "Consistency analysis of the simplified refined instrumental variable method for continuous-time systems," *Automatica*, vol. 113, Article 108767, 2020.
- [10] R. A. González, C. R. Rojas, S. Pan, and J. S. Welsh, "Consistent identification of continuous-time systems under multisine input signal excitation," *Automatica*, vol. 133, Article 109859, 2020.
- [11] 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*, pp. 14–19, 1996.
- [12] R. Ehrlich, C. Taussig, and D. Abramovitch, "Identification of sampled data systems at frequencies beyond the Nyquist rate," in *28th IEEE Conference on Decision and Control*, pp. 646–652, 1989.
- [13] M. van Haren, L. Mirkin, L. Blanken, and T. Oomen, "Beyond Nyquist in frequency response function identification: Applied to slow-sampled systems," *IEEE Control Systems Letters*, vol. 7, pp. 2131–2136, 2023.
- [14] R. Pintelon and J. Schoukens, *System Identification: A Frequency Domain Approach*, 2nd Edition. John Wiley & Sons, 2012.
- [15] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd Edition. Cambridge University Press, 2012.
- [16] E. L. Lehmann and G. Casella, *Theory of Point Estimation*, 2nd Edition. Springer, 1998.
- [17] H. H. Söhrab, *Basic Real Analysis*. Springer, 2003.
- [18] L. Ljung, "On the estimation of transfer functions," *Automatica*, vol. 21, no. 6, pp. 677–696, 1985.
- [19] R. A. González, K. Classens, C. R. Rojas, J. S. Welsh, and T. Oomen, "Identification of Additive Continuous-time Systems in Open and Closed-loop," *Automatica (in Press)*, 2024.
- [20] 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*, vol. 35, pp. 133–138, 2002.