

On Bilinear Time-Domain Identification and Reduction in the Loewner Framework



D. S. Karachalios, I. V. Gosea, and A. C. Antoulas

Abstract The *Loewner framework* (LF) in combination with *Volterra series* (VS) offers a non-intrusive approximation method that is capable of identifying bilinear models from time-domain measurements. This method uses harmonic inputs which establish a natural way for data acquisition. For the general class of nonlinear problems with VS representation, the growing exponential approach allows the derivation of the generalized kernels, namely, *symmetric generalized frequency response functions* (GFRFs). In addition, the homogeneity of the Volterra operator determines the accuracy in terms of how many kernels are considered. For the weakly nonlinear setup, only a few kernels are needed to obtain a good approximation. In this direction, the proposed adaptive scheme is able to improve the estimations of the computationally nonzero kernels. The Fourier transform associates these measurements with the derived GFRFs and the LF makes the connection with system theory. In the linear case, the LF associates the so-called S-parameters with the linear transfer function by interpolating in the frequency domain. The goal of the proposed method is to extend identification to the case of bilinear systems from time-domain measurements and to approximate other general nonlinear systems (by means of the Carleman bilinearization scheme). By identifying the linear contribution with the LF, a considerable reduction is achieved by means of the SVD. The fitted linear system has the same

D. S. Karachalios (✉) · I. V. Gosea · A. C. Antoulas
Max Planck Institute for Dynamics of Complex Technical Systems, Data-Driven System
Reduction and Identification (DRI), Magdeburg, Germany
e-mail: karachalios@mpi-magdeburg.mpg.de

I. V. Gosea
e-mail: gosea@mpi-magdeburg.mpg.de

A. C. Antoulas
e-mail: aca@rice.edu

A. C. Antoulas
Electrical and Computer Engineering Department, Rice University, Houston, TX 77005, USA
Baylor College of Medicine, Houston, TX 77030, USA

McMillan degree as the original linear system. Then, the performance of the linear model is improved by augmenting a special nonlinear structure. In a nutshell, we learn reduced-dimension bilinear models directly from a potentially large-scale system that is simulated in the time domain. This is done by fitting first a linear model, and afterward, by fitting the corresponding bilinear operator.

1 Introduction

In natural sciences, evolutionary phenomena can be modeled as dynamical systems. An ever-increasing need for improving the approximation accuracy has motivated including more involved and detailed features in the modeling process, thus inevitably leading to large-scale dynamical systems [3]. To overcome this problem, efficient finite methods heavily rely on *model reduction*. Model reduction methods can be classified into two broad categories, namely, *SVD based* and *Krylov based* (moment matching).

The most prominent among the SVD-based methods is *balanced truncation* (BT). In general, balancing methods are based on the computation of controllability and observability *gramians* and lead to the elimination of state variables which are difficult to reach and to observe. Besides having high-computational cost of solving the associated matrix Lyapunov equations, the advantages of balancing methods include the preservation of stability and an a priori computable error bound. For more details on these topics as well as on other model reduction methods not treated here (e.g., proper orthogonal decomposition (POD)/reduced basis (RB)), we refer the reader to the book [3] and the surveys [9, 15].

One way to perform model reduction is by employing *tangential interpolation*. These methods are known as rational *Krylov methods* or *moment-matching* methods. Krylov-based methods are numerically efficient and have lower computational cost, but, in general, the preservation of other properties (e.g., stability or passivity) is not automatic. For an extensive study in *interpolatory model reduction*, we refer the reader to the recent book [4]. In what follows, we will consider exclusively *interpolatory model reduction methods* and, in particular, the LF. For recent surveys on the LF, see [2, 6, 29]. The sensitivity to noise in the LF was already discussed in [19, 30].

When *input-output* data are offered, *data-driven* methods, such as the *Loewner framework* (LF), *dynamic mode decomposition* (DMD) [37], *sparse identification of nonlinear systems (with control)* (SINDYc) in [27], *vector fitting* (VF) [23], *Hankel* [25] or *subspace* methods [8, 24, 26], *moment-matching* [36], and *operator inference* [13, 14, 32], remain the only feasible approaches for recovering the hidden information.

DMD-based methods represent viable alternatives that require state-derivative estimations.

While the underlying dynamical system acts as a black box, model identification tools are important for the reliability of the discovered models (i.e., stability,

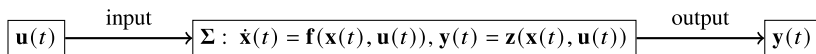


Fig. 1 Mathematical formalism for evolutionary phenomena

prediction). At the same time, these discovered models might have large dimension and hence are not suitable for fast numerical simulation and control. The LF is a direct data-driven interpolatory method able to identify and reduce models derived directly from measurements. For measured data in the frequency domain, the LF is well established for linear and nonlinear systems (e.g., bilinear or quadratic-bilinear systems) see [5, 22]. In the case of time-domain data, the LF was already applied for approximating linear models [21, 24, 31]. As the aim of this paper is to extend the identification and reduction procedure to the class of bilinear systems from time-domain data, we start our analysis by introducing the mathematical description of the *input*- $u(t)$ to *output*- $y(t)$ relation as depicted in Fig. 1. The differential and algebraic operators are denoted with \mathbf{f} and, respectively, with \mathbf{z} . To achieve this goal, all the important steps from nonlinear system theory and interpolatory model reduction are summarized.

1.1 Outline of the Paper

The rest of the paper is organized as follows:

- Section 2 contains a brief description of system theory starting from the linear case followed by extensions to the nonlinear case by means of the Volterra series representation. The single-input and single-output case is addressed for both frequency- and time-domain representations.
- Section 3 introduces the Loewner framework as an interpolatory tool for model approximation; the results that are presented here actually set the foundation for identification and reduction of linear time-invariant systems.
- Section 4 introduces a special class of nonlinear systems, e.g., bilinear systems. The theoretical discussion for analyzing such systems starts with the growing exponential approach and the derivation of the generalized frequency response functions (GFRFs) up to the case where a double-tone input is assumed. In addition, the kernel separation strategy for improving the measurements and the linear identification/reduction part is presented. A concise algorithm that summarizes the method is presented.
- Section 5 presents the numerical experiments performed in order to illustrate the practical applicability of the newly proposed method. This section includes both a simple (low-dimensional) example and a large-scale example, compared to another state-of-the-art method.
- Section 6 presents the concluding remarks and also some potential future developments of the current method.

2 System Theory Preliminaries

In this section, we will briefly present some important material from system theory starting from the linear case.

2.1 Linear Systems

Consider SISO linear, time-invariant systems with n internal variables (called “states” whenever the matrix \mathbf{E} is non-singular).

$$\Sigma_l : \begin{cases} \mathbf{E} \dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t), \\ y(t) = \mathbf{c}\mathbf{x}(t), \quad t \geq 0, \end{cases} \quad (1)$$

where $\mathbf{E}, \mathbf{A} \in \mathbb{R}^{n \times n}$, $\mathbf{b} \in \mathbb{R}^{n \times 1}$, $\mathbf{c} \in \mathbb{R}^{1 \times n}$. In the sequel, we will restrict our attention to invertible matrix \mathbf{E} and with zero d-term ($d = 0$) in the state-output equation¹. The explicit solution with the *convolution integral*² notation and the time-domain linear kernel $h(t)$ as the *impulse response* of the system can be written as

$$y(t) = \mathbf{c}e^{\mathbf{A}t}\mathbf{x}(0) + (h * u)(t), \quad t \geq 0, \quad (2)$$

where multiplication with \mathbf{E}^{-1} from the left has been performed in the differential part of Eq.(1). Also, we keep the same notation for the remaining matrix \mathbf{A} and vector \mathbf{b} . By assuming zero initial conditions and performing a *Laplace transform*, we obtain the transfer function description:

$$H(s) = \frac{Y(s)}{U(s)} = \mathbf{c}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{b}, \quad s \in \mathbb{C}, \quad (3)$$

where $Y(s), U(s)$ stand for the *input* and the *output* in the frequency domain.

2.2 Nonlinear Systems

A large class of nonlinear systems can be described by means of the Volterra-Wiener approach in [35]. Other relevant works on nonlinear systems and nonlinear modeling/identification include Schetzen (1980), Chen and Billings (1989), Boyd and Chua (1985) et. al.

The aim in this study is to identify and reduce special types of nonlinear systems (s.a., bilinear) from time-domain measurements. By knowing only the input and the

¹ The state-output equation often is represented as $y(t) = \mathbf{c}\mathbf{x}(t) + du(t)$.

² $(h * u)(t) = \int_{-\infty}^{\infty} h(\tau)u(t - \tau)d\tau$.

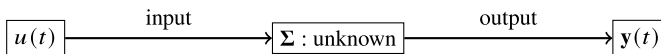


Fig. 2 The input-output mapping from the data-driven perspective with the unknown system Σ . Specific structures of the unknown system can be assumed/inspired by the physical problem. For instance, if the underlying physical phenomenon is fluid flow inside a control volume, quadratic models should be constructed, e.g., [22]

simulated or measured output in the time domain as in Fig. 2, we will identify the hidden model. In such situations where only snapshots are available, beyond the linear fit which is well established a nonlinear fit of a special type will be developed.

2.2.1 Approximation of Nonlinear Systems (Volterra Series)

The *input-output* relationship for a wide class of nonlinear systems [35] can be approximated by a Volterra series for sufficiently high N as

$$y(t) = \sum_{n=1}^N y_n(t), \quad y_n(t) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} h_n(\tau_1, \dots, \tau_n) \prod_{i=1}^n u(t - \tau_i) d\tau_i, \quad (4)$$

where $h_n(\tau_1, \dots, \tau_n)$ is a *real-valued* function of τ_1, \dots, τ_n known as the *n*th-order Volterra kernel.

Definition 1 The *n*th-order generalized frequency response function (GFRF) is defined as

$$H_n(j\omega_1, \dots, j\omega_n) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} h_n(\tau_1, \dots, \tau_n) e^{(-j\sum_{i=1}^n \omega_i \tau_i)} d\tau_1 \cdots d\tau_n, \quad (5)$$

which is the multidimensional Fourier³ transform of $h_n(\tau_1, \dots, \tau_n)$.

By applying the inverse Fourier transform of the *n*th-order GFRF, Eq. (5) can be written as

$$y_n(t) = \frac{1}{(2\pi)^n} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} H_n(j\omega_1, \dots, j\omega_n) \prod_{i=1}^n U(j\omega_i) e^{j(\omega_1 + \cdots + \omega_n)t} d\omega_i. \quad (6)$$

The *n*th Volterra operator is defined as

$$V_n(u_1, u_2, \dots, u_n) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} h_n(\tau_1, \dots, \tau_n) \prod_{i=1}^n u_i(t - \tau_i) d\tau_i, \quad (7)$$

³ With ($j^2 = -1$), as the frequency $s = j\omega$ lies on the imaginary axis, the Laplace transform simplifies in most cases to Fourier transform (e.g., for square-integrable functions).

so that $y_n = V_n(u, u, \dots, u)$ holds true.

➤ Homogeneity of the Volterra operator

The map $u(t) \rightarrow y_n(t)$ is homogeneous of degree n , that is, $\alpha u \rightarrow \alpha^n y_n$, $\alpha \in \mathbb{C}$. Each Volterra kernel $h_n(t)$ determines a symmetric multi-linear operator. Small amplitudes (e.g., $|\alpha| < \epsilon$) will allow ordering the nonlinear terms in such a way that terms with large powers of the amplitude (α^n) will be negligible. That is precisely the sense of approximating weakly nonlinear systems with Volterra series.

2.2.2 A Single-Tone Input

Consider the excitation of a system with an input consisting of two complex exponentials as in Eq. (8). Such inputs are typically used in chemical engineering applications as [33].

$$u(t) = A \cos(\omega t) = \left(\frac{A}{2}\right) e^{j\omega t} + \left(\frac{A}{2}\right) e^{-j\omega t}. \quad (8)$$

By using the above input in Eq. (4), we can derive the first Volterra term with $n = 1$ as

$$\begin{aligned} y_1(t) &= \int_{-\infty}^{\infty} h_1(\tau_1)[u(t - \tau_1)]d\tau_1 \\ &= \frac{A}{2} e^{j\omega t} \underbrace{\int_{-\infty}^{\infty} h_1(\tau_1)e^{-j\omega\tau_1} d\tau_1}_{H_1(j\omega)} + \frac{A}{2} e^{-j\omega t} \underbrace{\int_{-\infty}^{\infty} h_1(\tau_1)e^{j\omega\tau_1} d\tau_1}_{H_1(-j\omega)} \Rightarrow \\ y_1(t) &= \frac{A}{2} \left(e^{j\omega t} H_1(j\omega) + e^{-j\omega t} H_1(-j\omega) \right). \end{aligned} \quad (9)$$

Similarly, for the second term, we can derive

$$y_2(t) = \left(\frac{A}{2}\right)^2 \left[e^{2j\omega t} H_2(j\omega, j\omega) + 2e^{0j\omega t} H_2(j\omega, -j\omega) + e^{-2j\omega t} H_2(-j\omega, -j\omega) \right]. \quad (10)$$

Remark 1 (Conjugate symmetry): $H_2^*(j\omega, -j\omega) = H_2(-j\omega, j\omega)$, $\forall \omega \in \mathbb{R}$.

The input amplitude is A , the angular frequency is ω , the imaginary unit is j , the first-order response function is $H_1(j\omega)$, and $H_n(j\omega, \dots, j\omega)$, for $n \geq 2$, are the higher order FRFs or GFRFs. Then, the n th Volterra term can be written as

$$y_n(t) = \left(\frac{A}{2}\right)^n \sum_{p+q=n} {}^n C_q H_n^{p,q}(j\omega) e^{j\omega_{p,q}t}, \quad \omega_{p,q} = (p - q)\omega. \quad (11)$$

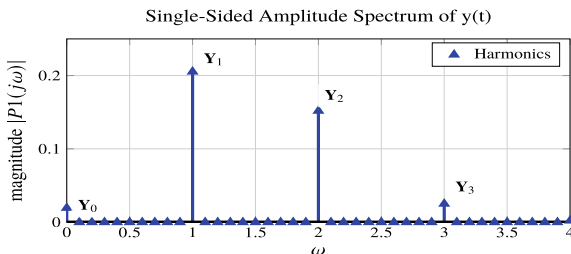


Fig. 3 An instance of the single-sided power spectrum with a singleton input with $\omega = 1$ is depicted. The underlying system is nonlinear and as a result higher harmonics appeared with a DC (direct current—non-periodic) term as well

where the following notations have been used:

$$H_n^{p,q}(j\omega) = H_n(\underbrace{j\omega, \dots, j\omega}_{p\text{-times}}; \underbrace{-j\omega, \dots, -j\omega}_{q\text{-times}}), \omega_{p,q} = (p - q)\omega, {}^n C_q = \frac{n!}{q!(n - q)!}. \quad (12)$$

2.2.3 Time-Domain Representation of Harmonics

The m th harmonic in the *time domain* can be computed by collecting the identical exponential power coefficients from Eq. (13) and by setting $p - q = m$, with $p = m + i - 1$ and $q = i - 1$ in Eq. (11). Hence, it follows that

$$y_m^{th}(t) = \sum_{i=1}^{\infty} \left(\frac{A}{2}\right)^{m+2i-2} {}^{m+2i-2} C_{i-1} H_{m+2i-2}^{m+i-1, i-1}(j\omega) e^{jm\omega t}. \quad (13)$$

2.2.4 Frequency-Domain Representation of Harmonics

The m th harmonic in the *frequency domain* by applying single-sided Fourier transform in Eq. (13) is the following:

$$Y_m^{th}(jm\omega) = \sum_{i=1}^{\infty} \left(\frac{A}{2}\right)^{m+2i-2} {}^{m+2i-2} C_{i-1} H_{m+2i-2}^{m+i-1, i-1}(j\omega) \delta(jm\omega), \quad (14)$$

where $\delta(\cdot)$ is the Dirac delta distribution. When a single-tone input excites a nonlinear dynamical system, the steady-state frequency response is characterized by a spectrum with higher harmonics (as can be seen, for example, in Fig. 3). This behavior is not observed in the linear case, where only one harmonic appears at the input frequency.

3 The Loewner Framework

We start with an account of the Loewner framework (LF) in the linear case [2, 6, 29]. The LF is an interpolatory method that seeks reduced models whose transfer function matches that of the original system at selected interpolation points. An important attribute is that it provides a trade-off between accuracy of fit and complexity of the model. It constructs models from given frequency data in a straightforward manner. In the case of SISO systems, we have the rational scalar interpolation problem to solve.

Consider a given set of complex data as

$$\{(s_k, f_k(s_k)) \in \mathbb{C} \times \mathbb{C} : k = 1, \dots, 2n\}.$$

We partition the data in two disjoint sets:

$$\mathbf{S} = [\underbrace{s_1, \dots, s_n}_{\mu}, \underbrace{s_{n+1}, \dots, s_{2n}}_{\lambda}], \quad \mathbf{F} = [\underbrace{f_1, \dots, f_n}_{\mathbb{V}}, \underbrace{f_{n+1}, \dots, f_{2n}}_{\mathbb{W}}],$$

where $\mu_i = s_i$, $\lambda_i = s_{n+i}$, $v_i = f_i$, $w_i = f_{n+i}$ for $i = 1, \dots, n$.

The objective is to find $H(s) \in \mathbb{C}$, such that

$$H(\mu_i) = v_i, \quad i = 1, \dots, n, \quad \text{and} \quad H(\lambda_j) = w_j, \quad j = 1, \dots, n. \quad (15)$$

The *left dataset* is denoted as

$$\mathbf{M} = [\mu_1, \dots, \mu_n] \in \mathbb{C}^{1 \times n}, \quad \mathbb{V} = [v_1, \dots, v_n]^T \in \mathbb{C}^{n \times 1}, \quad (16)$$

while the *right dataset* as

$$\mathbf{A} = [\lambda_1, \dots, \lambda_n]^T \in \mathbb{C}^{n \times 1}, \quad \mathbb{W} = [w_1, \dots, w_n] \in \mathbb{C}^{1 \times n}. \quad (17)$$

Interpolation points are determined by the problem or are selected to achieve given model reduction goals. For ways of choosing the interpolation grids and of partitioning the data into the left and right sets, we refer the reader to the recent survey [29].

3.1 The Loewner Matrix

Given a row array of complex numbers (μ_j, v_j) , $j = 1, \dots, n$, and a column array, (λ_i, w_i) , $i = 1, \dots, n$, (with λ_i and the μ_j mutually distinct) the associated *Loewner matrix* \mathbb{L} and the shifted *Loewner matrix* \mathbb{L}_s are defined as

$$\mathbb{L} = \begin{bmatrix} \frac{v_1 - w_1}{\mu_1 - \lambda_1} & \dots & \frac{v_1 - w_n}{\mu_1 - \lambda_n} \\ \vdots & \ddots & \vdots \\ \frac{v_n - w_1}{\mu_n - \lambda_1} & \dots & \frac{v_n - w_n}{\mu_n - \lambda_n} \end{bmatrix} \in \mathbb{C}^{n \times n}, \quad \mathbb{L}_s = \begin{bmatrix} \frac{\mu_1 v_1 - \lambda_1 w_1}{\mu_1 - \lambda_1} & \dots & \frac{\mu_1 v_1 - \lambda_n w_n}{\mu_1 - \lambda_n} \\ \vdots & \ddots & \vdots \\ \frac{\mu_n v_n - \lambda_1 w_1}{\mu_n - \lambda_1} & \dots & \frac{\mu_n v_n - \lambda_n w_n}{\mu_n - \lambda_n} \end{bmatrix} \in \mathbb{C}^{n \times n}.$$

Definition 2 If g is rational, i.e., $g(s) = \frac{p(s)}{q(s)}$, for appropriate polynomials p, q , the McMillan degree or the complexity of g is $\deg g = \max\{\deg(p), \deg(q)\}$.

Now, if $w_i = g(\lambda_i)$ and $v_j = g(\mu_j)$ are *samples* of a rational function g , the *main property* of Loewner matrices asserts the following.

Theorem 1 [2] Let \mathbb{L} be as above. If $k, q \geq \deg g$, then $\text{rank } \mathbb{L} = \deg g$.

In other words, the rank of \mathbb{L} encodes the complexity of the underlying rational function g . Furthermore, the same result holds for matrix-valued functions g .

3.2 Construction of Interpolants

If the pencil $(\mathbb{L}_s, \mathbb{L})$ is regular, then $\mathbf{E} = -\mathbb{L}$, $\mathbf{A} = -\mathbb{L}_s$, $\mathbf{b} = \mathbb{V}$, $\mathbf{c} = \mathbb{W}$, is a minimal realization of an interpolant for the data, i.e., $H(s) = \mathbb{W}(\mathbb{L}_s - s\mathbb{L})^{-1}\mathbb{V}$. Otherwise, as shown in [2], the problem in Eq. (15) has a solution provided that

$$\text{rank } [s\mathbb{L} - \mathbb{L}_s] = \text{rank } [\mathbb{L}, \mathbb{L}_s] = \text{rank } \begin{bmatrix} \mathbb{L} \\ \mathbb{L}_s \end{bmatrix} = r,$$

for all $s \in \{\mu_i\} \cup \{\lambda_j\}$. Consider then the thin SVDs:

$$[\mathbb{L}, \mathbb{L}_s] = \mathbf{Y} \widehat{\Sigma}_r \widetilde{\mathbf{X}}^*, \quad \begin{bmatrix} \mathbb{L} \\ \mathbb{L}_s \end{bmatrix} = \widetilde{\mathbf{Y}} \Sigma_r \mathbf{X}^*,$$

where $\widehat{\Sigma}_r, \Sigma_r \in \mathbb{R}^{r \times r}$, $\mathbf{Y} \in \mathbb{C}^{n \times r}$, $\mathbf{X} \in \mathbb{C}^{n \times r}$, $\widetilde{\mathbf{Y}} \in \mathbb{C}^{2n \times r}$, $\widetilde{\mathbf{X}} \in \mathbb{C}^{r \times 2n}$.

Remark 2 r can be chosen as the *numerical rank* (as opposed to the *exact rank*) of the Loewner pencil.

Theorem 2 The quadruple $(\tilde{\mathbf{A}}, \tilde{\mathbf{b}}, \tilde{\mathbf{c}}, \tilde{\mathbf{E}})$ of size, $r \times r$, $r \times 1$, $1 \times r$, $r \times r$, given by

$$\tilde{\mathbf{E}} = -\mathbf{Y}^T \mathbb{L} \mathbf{X}, \quad \tilde{\mathbf{A}} = -\mathbf{Y}^T \mathbb{L}_s \mathbf{X}, \quad \tilde{\mathbf{b}} = \mathbf{Y}^T \mathbf{V}, \quad \tilde{\mathbf{c}} = \mathbf{W} \mathbf{X},$$

is a descriptor realization of an (approximate) interpolant of the data with McMillan degree $r = \text{rank}(\mathbb{L})$, where $\tilde{H}(s) = \tilde{\mathbf{c}}(s\tilde{\mathbf{E}} - \tilde{\mathbf{A}})^{-1}\tilde{\mathbf{b}}$.

For more details on the construction/identification of linear systems with the LF, we refer the reader to [4, 6, 29] where both the SISO and MIMO cases are addressed together with other more technical aspects (e.g., how to impose the construction of real-valued models, etc.).

4 The Special Case of Bilinear Systems

In recent years, projection-based Krylov methods have extensively been applied for model reduction of bilinear systems. We mention the following contributions [1, 5, 7, 10–12, 17, 20, 34] and the references within.

Scalar bilinear systems are described by the set of matrices; $\Sigma_b = (\mathbf{A}, \mathbf{N}, \mathbf{b}, \mathbf{c}, \mathbf{E})$ and characterized by the following equations:

$$\Sigma_b : \begin{cases} \mathbf{E}\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{N}\mathbf{x}(t)u(t) + \mathbf{b}u(t), \\ y(t) = \mathbf{c}\mathbf{x}(t), \end{cases} \quad (18)$$

where $\mathbf{E}, \mathbf{A}, \mathbf{N} \in \mathbb{R}^{n \times n}$, $\mathbf{b} \in \mathbb{R}^{n \times 1}$, $\mathbf{c} \in \mathbb{R}^{1 \times n}$, and $\mathbf{x} \in \mathbb{R}^{n \times 1}$, $u, y \in \mathbb{R}$. In what follows, we restrict our analysis to systems with non-singular \mathbf{E} matrices (e.g., identity matrix).

4.1 The Growing Exponential Approach

The properties of the growing exponential approach can be adapted readily to the problem of *finding transfer functions* for constant-parameter (stationary) state equations. Let us consider the bilinear model in Eq. (18) with zero initial conditions. A single-tone input with amplitude $A < 1$ is considered as in Eq. (8).

$$u(t) = A \cos(\omega t) = \frac{A}{2} e^{j\omega t} + \frac{A}{2} e^{-j\omega t} = a e^{j\omega t} + a e^{-j\omega t}, \quad (19)$$

where $a = A/2$ and $a \in (0, \epsilon)$ with $0 < \epsilon < 1/2$ and for all $t \geq 0$. The steady-state solution for the differential equation in Eq. (18) can be written as follows:

$$\mathbf{x}(t) = \sum_{p,q \in \mathbb{N}} \mathbf{G}_n^{p,q}(\underbrace{j\omega, \dots, j\omega}_{p\text{-times}}, \underbrace{-j\omega, \dots, -j\omega}_{q\text{-times}}) a^{p+q} e^{j\omega(p-q)t}. \quad (20)$$

The symbol⁴ $\mathbf{G}_n^{p,q}$ denotes the n th input to state frequency response containing p -times the frequency ω and q -times the frequency $-\omega$. By substituting in Eq. (18) and collecting the terms of the same exponential (as the $e^{j\omega_m t}$), we can derive the input to state frequency responses \mathbf{G}_n for every n as follows:

$$\begin{aligned} \sum_{p,q \in \mathbb{N}} (j\omega(p-q)\mathbf{E} - \mathbf{A}) \mathbf{G}_n^{p,q} a^{p+q} e^{j\omega(p-q)t} &= \mathbf{b}(ae^{j\omega t} + ae^{-j\omega t}) + \\ + \mathbf{N} \left(\sum_{p,q \in \mathbb{N}} \mathbf{G}_n^{p,q} a^{p+q+1} e^{j\omega(p+1-q)t} + \sum_{p,q \in \mathbb{N}} \mathbf{G}_n^{p,q} a^{p+q+1} e^{j\omega(p-q-1)t} \right). \end{aligned}$$

For the first choices of p and q up to $p+q \leq 2$, $(1, 0)$, $(0, 1)$, $(2, 0)$, $(0, 2)$, $(1, 1)$ and by denoting the resolvent $\Phi(j\omega) = (j\omega\mathbf{E} - \mathbf{A})^{-1} \in \mathbb{C}^{n \times n}$, c.t. conjugate terms, we derive the first set of terms

$$\begin{aligned} \Phi(j\omega)^{-1} \mathbf{G}_1^{1,0} a e^{j\omega t} + \Phi(2j\omega)^{-1} \mathbf{G}_2^{2,0} a^2 e^{2j\omega t} + \Phi(0)^{-1} \mathbf{G}_2^{1,1} a^2 + c.t. + \dots = \\ \mathbf{N} \mathbf{G}_1^{1,0} a^2 e^{2j\omega t} + \mathbf{N} \mathbf{G}_2^{2,0} a^3 e^{3j\omega t} + \mathbf{N} \mathbf{G}_2^{1,1} a^3 e^{j\omega t} + c.t. + \dots + \mathbf{b} a e^{j\omega t} + c.t. \end{aligned}$$

Collecting the same powers in both exponential and polynomial magnitudes, we compute the first and the second time/input-invariant GFRFs:

$$\begin{aligned} \mathbf{G}_1^{1,0}(j\omega) &= \Phi(j\omega)\mathbf{b}, \\ \mathbf{G}_2^{2,0}(j\omega) &= \Phi(2j\omega)\mathbf{N}\mathbf{G}_1^{1,0} = \Phi(2j\omega)\mathbf{N}\Phi(j\omega)\mathbf{b}. \end{aligned} \quad (21)$$

Then, the following input to state transfer functions \mathbf{G}_n using induction are

$$\begin{aligned} \mathbf{G}_n^{n,0}(j\omega) &= \Phi(nj\omega)\mathbf{N}\Phi((n-1)j\omega)\mathbf{N}\dots\mathbf{N}\Phi(j\omega)\mathbf{b}, \\ \mathbf{G}_n^{0,n}(j\omega) &= \Phi(-nj\omega)\mathbf{N}\Phi(-(n-1)j\omega)\mathbf{N}\dots\mathbf{N}\Phi(-j\omega)\mathbf{b}, \\ \mathbf{G}_n^{p,q}(j\omega) &= \Phi((p-q)j\omega)\mathbf{N} \left[\mathbf{G}_{n-1}^{p,q-1}(j\omega) + \mathbf{G}_{n-1}^{p-1,q}(j\omega) \right], \quad p, q \geq 1, \end{aligned} \quad (22)$$

for $n \geq 1$ and $p+q = n$. By multiplying with the output vector \mathbf{c} , we can further derive the input-output generalized frequency responses GFRFs as

⁴ $\mathbf{G}_n^{p,q} = \mathbf{G}(j\omega, \dots, j\omega; \underbrace{-j\omega, \dots, -j\omega}_{q\text{-times}})$.

$$\begin{aligned}
H_n^{n,0}(j\omega) &= \mathbf{c}\Phi(nj\omega)\mathbf{N}\Phi((n-1)j\omega)\mathbf{N}\cdots\mathbf{N}\Phi(j\omega)\mathbf{b}, \\
H_n^{0,n}(j\omega) &= \mathbf{c}\Phi(-nj\omega)\mathbf{N}\Phi(-(n-1)j\omega)\mathbf{N}\cdots\mathbf{N}\Phi(-j\omega)\mathbf{b}, \\
H_n^{p,q}(j\omega) &= \mathbf{c}\Phi((p-q)j\omega)\mathbf{N}\left[\mathbf{G}_{n-1}^{p,q-1}(j\omega) + \mathbf{G}_{n-1}^{p-1,q}(j\omega)\right], \quad p, q \geq 1.
\end{aligned} \tag{23}$$

At this point, we can write the Volterra series by using the above specific structure of the GFRFs that were derived with the growing exponential approach for the bilinear case. An important property to notice is that the n th kernel is a multivariate function of order n . It is obvious that the identification of the n th-order FRF involves an n -dimensional frequency space. For that reason, next, we derive the general second symmetric kernel for the bilinear case with a double-tone input. Consider:

$$u(t) = A_1 \cos(\omega_1 t) + A_2 \cos(\omega_2 t) = \sum_{i=1}^2 \alpha_i (e^{j\omega_i t} + e^{-j\omega_i t}), \tag{24}$$

where $\alpha_1 = \frac{A_1}{2}$ and $\alpha_2 = \frac{A_2}{2}$. In that case, with the growing exponential approach the state solution in steady state is

$$\mathbf{x}(t) = \sum_{m_1, \dots, m_4 \in \mathbb{N}} \mathbf{G}_n^{m_1, m_2, m_3, m_4} \alpha_1^{m_1+m_2} \alpha_2^{m_3+m_4} e^{j((m_1-m_2)\omega_1 + (m_3-m_4)\omega_2)t}. \tag{25}$$

We are looking for the input to state frequency response $\mathbf{G}(j\omega_1, j\omega_2)$. By substituting to the bilinear model in Eq. (18) and collecting the appropriate terms while at the same time using the symmetry $\mathbf{G}(j\omega_1, j\omega_2) = \mathbf{G}(j\omega_2, j\omega_1)$, we conclude that

$$\mathbf{G}_2(j\omega_1, j\omega_2) = \frac{1}{2} [(j\omega_1 + j\omega_2)\mathbf{E} - \mathbf{A}]^{-1} \mathbf{N} [(j\omega_1\mathbf{E} - \mathbf{A})^{-1} \mathbf{b} + (j\omega_2\mathbf{E} - \mathbf{A})^{-1} \mathbf{b}], \tag{26}$$

where by using the resolvent notation and multiplying with \mathbf{c} , we derive the *second-order symmetric generalized frequency response function* as

$$H_2(j\omega_1, j\omega_2) = \frac{1}{2} \mathbf{c}\Phi(j\omega_1 + j\omega_2)\mathbf{N} [\Phi(j\omega_1)\mathbf{b} + \Phi(j\omega_2)\mathbf{b}]. \tag{27}$$

4.2 The Kernel Separation Method

One way to deduce Volterra kernels is by means of interpolation. This problem is equivalent to that of estimating a polynomial with noisy coefficients. This interpolation scheme builds a linear system with a Vandermonde matrix which is invertible since the amplitudes are distinct and nonzero. The inverse of a Vandermonde matrix can be explicitly computed and there are stable ways to solve these equations [16]. The recently proposed method presented in [18] solves the exponentially

ill-condition problem of the Vandermonde matrix with Arnoldi orthogonalization. The m th harmonic in the frequency domain is derived by applying a (single-sided) Fourier transform. More precisely, the explicit formulation is as follows:

$$\begin{aligned}
 Y_{m^{\text{th}}}(jm\omega) &= \sum_{i=1}^{\infty} \underbrace{\left(\frac{A}{2}\right)^{m+2i-2} C_{i-1} H_{m+2i-2}^{m+i-1, i-1}(j\omega)}_{\alpha^{m+2i-2}} \delta(jm\omega) \\
 &= \sum_{i=1}^{\infty} \alpha^{m+2i-2} H_{m+2(i-1)}^{m+i-1, i-1}(j\omega) \delta(jm\omega).
 \end{aligned} \tag{28}$$

We simplify the notation in order to reveal the adaptive method that will help us to estimate the GFRFs up to a specific order. Next, write the linear system of equations that connects the harmonic information with the higher Volterra kernels as follows:

$$\underbrace{\begin{bmatrix} Y_0(0j\omega) \\ Y_1(1j\omega) \\ Y_2(2j\omega) \\ Y_3(3j\omega) \\ \vdots \\ Y_m(mj\omega) \end{bmatrix}}_{\mathbf{Y}_{(\alpha, \omega)}} = \left\{ \underbrace{\begin{bmatrix} \alpha^0 & \alpha^2 & \alpha^4 & \dots \\ \alpha^1 & \alpha^3 & \alpha^5 & \dots \\ \alpha^2 & \alpha^4 & \alpha^6 & \dots \\ \alpha^3 & \alpha^5 & \alpha^7 & \dots \\ \vdots & \vdots & \vdots & \vdots \\ \alpha^m & \alpha^{m+2} & \alpha^{m+4} & \dots \end{bmatrix}}_{\mathbf{M}_{\alpha}} \odot \underbrace{\begin{bmatrix} H_0^{0,0} & H_2^{1,1} & H_4^{2,2} & \dots \\ H_1^{1,0} & H_3^{2,1} & H_5^{3,2} & \dots \\ H_2^{2,0} & H_4^{3,1} & H_6^{4,2} & \dots \\ H_3^{3,0} & H_5^{4,1} & H_7^{5,2} & \dots \\ \vdots & \vdots & \vdots & \vdots \\ H_n^{n,0} & H_{n+2}^{n+1,1} & H_{n+4}^{n+2,2} & \dots \end{bmatrix}}_{\mathbf{P}_{\omega}} \right\} \underbrace{\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}}_{\mathbf{e}_{n+1,1}}. \tag{29}$$

By introducing the Hadamard product notation⁵ and by substituting the δ 's with ones, we can compactly rewrite the above system in the following form:

$$\mathbf{Y}_{(\alpha, \omega)} = [\mathbf{M}_{\alpha} \odot \mathbf{P}_{\omega}] \cdot \mathbf{e}_{n+1,1}. \tag{30}$$

The above system offers the level of approximation we want to achieve. Note that the frequency response \mathbf{Y} depends on both the amplitude and the frequency, while the right-hand side of Eq. (30) reveals the separation of the aforementioned quantities. As we neglect higher order Volterra kernels, the measurement set tends to be corrupted by noise.

➤ Kernel separation and stage ℓ -approximation

For a given system, the procedure consists in exciting it with a single-tone input. By varying the driving frequency, as well as the amplitude, we can approximate the GFRFs by minimizing the (2-norm) of the remaining systems.

⁵ The Hadamard product is denoted with “ \odot ”; the matrix multiplication is performed element-wise.

$$\mathbf{Y}_{m+1,\ell}(jm\omega, \alpha_\ell) = [\mathbf{M}_{m+1,\ell}(\alpha_\ell) \odot \mathbf{P}_{m+1,\ell}(jm\omega)] \cdot \mathbf{e}_{n+1,1}. \quad (31)$$

The m -“direction” gives us the threshold up to the specific harmonic that we measure while the ℓ -“direction” gives us the level of the kernel separation that we want to achieve. For instance, for the second stage approximation, it holds $\ell = 2$ with $Y_m \approx 0$, $\forall m$ with $\ell = 2 < m = 3, 4, \dots$

4.3 Identification of the Matrix \mathbf{N}

The difference between linear and bilinear models is the presence of the product between the input and the state that is scaled by the matrix \mathbf{N} . As the LF is able to identify the linear part (\mathbf{A} , \mathbf{b} , \mathbf{c} , \mathbf{E}) of the bilinear model the only thing that remains is the identification of the matrix \mathbf{N} . The matrix \mathbf{N} enters linearly in the following kernels (as \mathbf{E} has been considered invertible, for simplicity, it is assumed $\mathbf{E} = \mathbf{I}$):

- With a single-tone input the kernel $H_2^{1,1}$ can be written as

$$H_2(j\omega_1, -j\omega_1) = \frac{1}{2} \mathbf{c} (-\mathbf{A})^{-1} \mathbf{N} ((j\omega_1 \mathbf{I} - \mathbf{A})^{-1} \mathbf{b} + (-j\omega_1 \mathbf{I} - \mathbf{A})^{-1} \mathbf{b}) \quad (32)$$

and the kernel $H_2^{2,0}$ as

$$H_2(j\omega_1, j\omega_1) = \mathbf{c} (2j\omega_1 \mathbf{I} - \mathbf{A})^{-1} \mathbf{N} (j\omega_1 \mathbf{I} - \mathbf{A})^{-1} \mathbf{b}. \quad (33)$$

- While with a double-tone input the general kernel H_2 can be written as

$$H_2(j\omega_1, j\omega_2) = \frac{1}{2} \mathbf{c} \left((j\omega_1 + j\omega_2) \mathbf{I} - \mathbf{A} \right)^{-1} \mathbf{N} \left((j\omega_1 \mathbf{I} - \mathbf{A})^{-1} \mathbf{b} + (j\omega_2 \mathbf{I} - \mathbf{A})^{-1} \mathbf{b} \right). \quad (34)$$

We introduce the following notation:

$$\begin{aligned} \mathcal{O}(j\omega_1, j\omega_2) &= \frac{1}{2} \mathbf{c} \left((j\omega_1 + j\omega_2) \mathbf{I} - \mathbf{A} \right)^{-1} \in \mathbb{C}^{1 \times n}, \\ \mathcal{R}(j\omega_1, j\omega_2) &= \left((j\omega_1 \mathbf{I} - \mathbf{A})^{-1} \mathbf{b} + (j\omega_2 \mathbf{I} - \mathbf{A})^{-1} \mathbf{b} \right) \in \mathbb{C}^{n \times 1}. \end{aligned} \quad (35)$$

Then, Eq. (34) can be compactly rewritten as

$$H_2(j\omega_1, j\omega_2) = \mathcal{O}(j\omega_1, j\omega_2) \mathbf{N} \mathcal{R}(j\omega_1, j\omega_2). \quad (36)$$

Assume that k measurements of the function H_2 are available (measured) for k different pairs (ω_1, ω_2) . By vectorizing in respect to the measurement set, we have

for the k th measurement:

$$\underbrace{H_2(j\omega_1^{(k)}, j\omega_2^{(k)})}_{\mathbf{Y}^{(k)}} = \underbrace{\mathcal{O}(j\omega_1^{(k)}, j\omega_2^{(k)})}_{\mathcal{O}_{1,n}^{(k)}} \underbrace{\mathbf{N}}_{n \times n} \underbrace{\mathcal{R}(j\omega_1^{(k)}, j\omega_2^{(k)})}_{\mathcal{R}_{n,1}^{(k)}},$$

$$\text{For all } k \text{ measurements} \rightarrow \mathbf{Y}_{(1:k,1)} = \underbrace{\left(\mathcal{O}_{(1,n)}^{(k)} \otimes \mathcal{R}_{(1,n)}^{T(k)} \right)}_{(1:k,n^2)} \underbrace{\text{vec}(\mathbf{N})}_{(1:n^2,1)}. \quad (37)$$

Note that Eqs. (32), (33), (34) can be equivalently rewritten as the one linear matrix equation given in Eq. (37). By filling out the above matrix $[\mathcal{O} \otimes \mathcal{R}^T]$ with the information from $H_2(j\omega_1, -j\omega_1)$ and from $H_2(j\omega_1, j\omega_1)$ as well, the solution can be improved. Hence, we are able to solve Eq. (37) with full rank and identify the matrix \mathbf{N} . All the symmetry properties of the kernels are appropriately used, e.g., conjugate-real symmetry. For n denoting the dimension of the bilinear model and k the number of measurements, we have the following two cases⁶:

1. $k < n^2$ underdetermined \rightarrow least-squares (LS) solution (minimizing the 2-norm) as in [28],
2. $k \geq n^2$ determined-rank completion \rightarrow identification of \mathbf{N} ,

Proposition 1 *Let $\Sigma_b = (\mathbf{A}, \mathbf{N}, \mathbf{b}, \mathbf{c}, \mathbf{E})$ be a bilinear system of dimension n for which the linear subsystem $\Sigma_l = (\mathbf{A}, \mathbf{b}, \mathbf{c}, \mathbf{E})$ is fully controllable and observable. Then, for $k \geq n^2$ measurements so that $(j\omega_1^{(k)}, j\omega_2^{(k)})$ are distinct complex pairs with $(\omega_1^{(k)}, \omega_2^{(k)}) \in \mathbb{R}_+^2$ and $\omega_1^{(k)} \neq \omega_2^{(k)}$, the following holds:*

$$\text{rank} \left(\underbrace{\begin{bmatrix} \mathcal{O}^{(1)} \otimes \mathcal{R}^{T(1)} \\ \mathcal{O}^{(2)} \otimes \mathcal{R}^{T(2)} \\ \vdots \\ \mathcal{O}^{(k)} \otimes \mathcal{R}^{T(k)} \end{bmatrix}}_{(1:k \geq n^2, n^2)} \right) = n^2. \quad (38)$$

As the above result indicates, one would need at least n^2 measurements to identify the matrix \mathbf{N} corresponding to bilinear system of dimension n .

4.4 A Separation Strategy for the second Kernel

To identify the n th Volterra kernel, we need an n -tone input signal. As we want to identify the second kernel, the input signal needs to be chosen as a double-tone Eq. (24). The propagating harmonics are $e^{j(m_1 - m_2)\omega_1 + j(m_3 - m_4)\omega_2} t$ or more compactly

⁶ The vectorization is row-wise, $\text{vec}(\mathbf{N}) = [\mathbf{N}(1, 1:n) \cdots \mathbf{N}(n, 1:n)]^T \in \mathbb{R}^{n^2 \times 1}$.

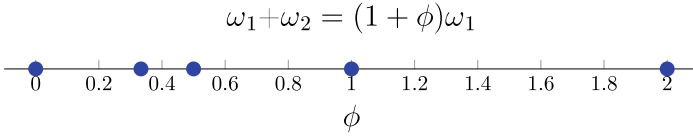


Fig. 4 This figure shows the constraints of ϕ (e.g., $\phi = 0, 1/3, 1/2, 1, 2, 3, \dots$, etc.). By choosing ϕ 's within the blue dots, we construct frequency bandwidths with a unique $(\omega_1 + \omega_2)$

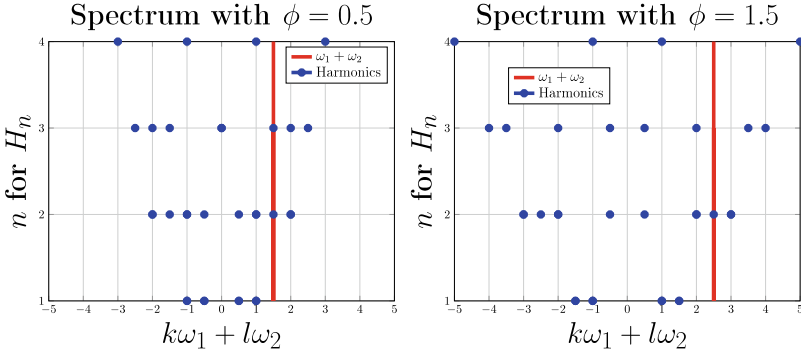


Fig. 5 Left pane: Overlapping kernels contributing to the same harmonic with invalid $\phi = 0.5$. Right pane: Uniquely defined harmonic at $(\omega_1 + \omega_2)$ with valid $\phi = 1.5$. Here, it holds $(n = k + l)$

$e^{\pm(kj\omega_1 \pm lj\omega_2)t}$, where $k, l \in \mathbb{N}$. The aim is to differentiate the $(\omega_1 + \omega_2)$ harmonic from the others harmonics. More precisely, we want the following result to hold:

$$\omega_1 + \omega_2 \neq k\omega_1 + l\omega_2, \quad \forall(k, l) \in \mathbb{Z} \times \mathbb{Z} \setminus \{1, 1\}. \quad (39)$$

Suppose $\omega_2 = \phi\omega_1$, $\phi \in \mathbb{R}$. The suitable ϕ 's where Eq. (39) holds are

$$\omega_1 + \phi\omega_1 = k\omega_1 + l\phi\omega_1 \Rightarrow 1 + \phi = k + l\phi \Rightarrow \phi = \frac{k-1}{1-l}, \quad k, l \in \mathbb{Z} \setminus \{1\}. \quad (40)$$

By choosing ϕ so that the equality in Eq. (40) doesn't hold, with harmonic mixing index $m = k + l$, it makes the harmonic $(\omega_1 + \omega_2)$ uniquely defined in the frequency spectrum up to the m th kernel.

To visualize this feature, we choose $\omega_1 = 1$, and $\omega_2 = \omega_1\phi = \phi$, for harmonic mixing index $m = 4$. Then, the constraints of ϕ are depicted in Fig. 4 with blue dots.

Next, in Fig. 5 and on the left pane, one ϕ constraint that occurs commensurate harmonics is depicted with the second and the third kernel to contribute at the same harmonic. On the right pane, the harmonic is uniquely defined at $(\omega_1 + \omega_2)$ from the second kernel up to the mixing order $m = 4$.

The next result allows us to construct sweeping frequency schemes to get enough measurements for the $H_2(j\omega_1, j\omega_2)$. So, for every $\omega_1 > 0$ the following should hold:

$$\omega_2 \in (\phi_{i-1}\omega_1, \phi_i\omega_1), \quad i = 1, \dots \quad (41)$$

where ϕ_i are the constraints (see Fig. 4 blue dots).

Remark 3 Note that in the proposed framework, the separation of the kernels that contribute at $(\omega_1 + \omega_2)$ harmonic is forced only under a specific mixing order m . We do not offer any general solution to this separation problem for multi-tone input, although techniques have been introduced such as in [16]. Therefore, it was also stated that the solution of the full separation of harmonics is, in general, not possible.

4.5 The Loewner-Volterra Algorithm for Time-Domain Bilinear Identification and Reduction

We start with a set of single-tone inputs $u(t) = \alpha_\ell \cos(\omega_1^{(i)} t)$, $i = 1, \dots, k$, with $\alpha_\ell < 1$. For those k measurements, we can estimate the linear kernel $H_1(j\omega_1^{(i)})$, the $H_2(j\omega_1^{(k)}, j\omega_1^{(k)})$ and the $H_2(j\omega_1^{(k)}, -j\omega_1^{(k)})$ by simply measuring the first harmonic as \mathbf{Y}_1 , the second harmonic as \mathbf{Y}_2 , and the DC term as \mathbf{Y}_0 , from the frequency spectrum as shown in Fig. 3. To improve the accuracy of the estimations for the aforementioned kernels, we could further upgrade to an ℓ -stage approximation by varying the amplitude α_ℓ as explained in Sect. 4.2. This approach is necessary whenever higher harmonics are considered to be numerically nonzero, hence meaningful. The reason for this is that the first harmonic is hence corrupted by noise introduced by the term $H_3^{2,1}$ and the rest of the terms which appears on the second row of matrix \mathbf{P}_ω in Eq. (29).

Since the LF reveals the underlying order of the linear system denoted with r , the value of k should be at least equal to $2r$. Then, we can take the decision on what will be the order r of the reduced system by analyzing the singular value decay. Up to the previous step, we have identified the linear part with the LF, and we have filled the LS problem Eq. (37) with measurements from the diagonal of the second kernel and from the the perpendicular to the diagonal axis $(\omega_1, -\omega_1)$. Those measurements contribute to the problem, but with an underdetermined (rank deficient) LS problem.

We need more measurements of H_2 to reach the full rank (r^2) solution that will lead to the identification of \mathbf{N} . So, we proceed by measuring the H_2 out of the diagonal ($\omega_1 \neq \omega_2$) with a double-tone input as $u(t) = \alpha_\ell \cos(\omega_1^{(k)} t) + \beta_\ell \cos(\omega_2^{(k)} t)$, for a set of frequency pairs (ω_1, ω_2) up to r^2 . The kernel separation problem for the frequency (ω_1, ω_2) appears now. To deal with this problem, we follow the solution proposed in Sect. 4.4 (up to a mixing degree). Last, we solve the real⁷ full-rank LS problem described in Eq. (37) by using all the symmetric properties of these kernels (i.e., real symmetry, conjugate symmetry, and the fact that $H_2(j\omega_1, j\omega_2) = H_2(j\omega_2, j\omega_1)$). An algorithm that summarizes the above procedure is presented below.

⁷ Enforcing real-valued models has been discussed in [6, 29]; here, we follow the same approach.

(Algorithm) The Loewner-Volterra algorithm for bilinear identification and reduction from time-domain data.

Input/Data acquisition: Use as control input the signals: $u(t) = \alpha_\ell \cos(\omega_1^{(k)} t) + \beta_\ell \cos(\omega_2^{(k)} t)$, $t \geq 0$, by sweeping the small amplitudes (< 1) and a particular range of frequencies.

Output: A bilinear system of dimension- r : $\Sigma_{b_r} : (\mathbf{A}_r, \mathbf{N}_r, \mathbf{b}_r, \mathbf{c}_r, \mathbf{E}_r)$

1. Apply one-tone input $u(t)$ with $\beta_\ell = 0$, $\omega_1^{(k)}$ for $k = 1, \dots, n$, and collect the snapshots $y(t)$ in steady state.
 2. Apply Fourier transform and collect the following measurements:
 - DC term: $Y_O(0 \cdot j\omega_1^{(k)})$,
 - 1st harmonic: $Y_I(1 \cdot j\omega_1^{(k)})$,
 - 2nd harmonic: $Y_{II}(2 \cdot j\omega_1^{(k)})$,
 - \vdots
 - m th harmonic: $Y_{m^{th}}(m \cdot j\omega_1^{(k)})$ (last numerically nonzero harmonic).
 3. If the second harmonic or higher harmonics are nonzero, the system is nonlinear. By sweeping the amplitude and using the adaptive scheme (stage ℓ -approximation) in Eq. (30), the estimations of the first and the second kernels can be improved. If the second and higher harmonics are equal to zero, the bilinear matrix \mathbf{N} remains zero and the underlying system is linear.
 4. Apply the linear LF, see Algorithm 1 in [29] by using the measurements (e.g., $H_1(j\omega_1^{(k)}) \approx 2Y_I(j\omega_1^{(k)})/\alpha_\ell$ for the second stage approximation $Y_m \approx 0$ for $m > 2$) and get the order r linear model.
 5. If the system is nonlinear, by fitting a bilinear matrix \mathbf{N} will improve the accuracy. Apply the two-tone input $u(t) = \alpha_\ell \cos(\omega_1^{(k)} t) + \beta_\ell \cos(\omega_2^{(k)} t)$ to get enough measurements ($\leq r^2$) to produce a full-rank LS problem. Measure the $(\omega_1 + \omega_2)$ harmonic as explained in Sect. 4.4 and get the estimations for the second kernel as: $H_2(j\omega_1^{(k)}, j\omega_2^{(k)}) \approx 2Y_{II}(j\omega_1^{(k)}, j\omega_2^{(k)})/(\alpha_\ell \beta_\ell)$.
 6. Solve the full-rank least-squares problem as described in Eq. (37) and compute the real-valued bilinear matrix \mathbf{N} . When the inversion is not exact due to numerical issues, the least-squares solution is obtained with a thresholding SVD.
-

4.6 Computational Effort of the Proposed Method

In this section, we discuss the computational effort of the proposed method by analyzing each step. We comment on the applicability of large-scale problems and the relation with real-world scenarios.

Simulation of processes with harmonic inputs constitutes a classical technique which is applied in many engineering applications; data acquisition in the time domain is a common procedure. Nevertheless, using advanced electronic devices such as vector network analyzers (VNAs), frequency-domain data can also be obtained (directly). The Loewner framework applied in the case where frequency-domain data that are obtained from VNAs offers an excellent identification and reduction tool in the linear case (with many applications in electrical, mechanical, or civil engineering). In the context of the current paper, we deal with time-domain data for a special class of nonlinear problems.

For the purpose of identifying and reducing bilinear systems from time-domain measurements, the most expensive procedure is that of data collection. This is done by simulating time-domain models with Euler's method (bilinear models such as the ones approximating Burgers' equation). Nevertheless, the heavy computational cost of simulating large dimensional systems in time domain could be alleviated using parallel processing (e.g., for multiple computational clusters). The process of estimating transfer functions values by computing the Fourier transform hence remains robust. In addition, the LF can adaptively detect the decay of the singular values and hence the procedure can be terminated for a specific reduced order $r \ll n$.

In the beginning, a linear system of reduced dimension r is fitted using the LF. For the rest of the proposed algorithm, note that we will use the lower dimension r to our advantage, and hence the method remains robust. The next step is to compute the matrix \mathbf{N} that characterizes the nonlinearity of bilinear systems. As the fitted linear system is of dimension r , we hence need to detect exactly r^2 unknowns (the entries of matrix \mathbf{N}). As presented in Sect. 4.3, this boils down to solving a full-rank LS problem that can be easily dealt with.

The aim of the newly proposed method is to accurately train bilinear models from time-domain data. We offer a first step approach toward complete identification of such systems within the Volterra series approximation approach. In many cases, large-scale systems are sparse (due to spatial domain semi-discretization) and hence reduction techniques can be applied. The new method deals with the inherent redundancies through the linear subsystem (compression by means of SVD). Afterward, it updates the nonlinear behavior by introducing an appropriate low-dimensional bilinear matrix that improves the overall approximation. Note also that the new method relies on the *controllability/observability* of the fitted linear system. Additionally, noise values up to a particular threshold can be handled as presented in Sect. 5; further analysis on noise-related issues is left for future research.

5 Numerical Examples

Example 1 (Identifying a low-order bilinear toy example) The aim of this experiment is to identify a simple bilinear model from time-domain measurements. Consider the following controllable/observable bilinear model Eq. (18) of dimension-2 with a *non-symmetric* matrix \mathbf{N} , zero initial condition and matrices as

$$\mathbf{E} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \mathbf{A} = \begin{bmatrix} -1 & -10 \\ 10 & -1 \end{bmatrix}, \mathbf{N} = \begin{bmatrix} 1 & -2 \\ 3 & -4 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{C} = [1 \ 1]. \quad (42)$$

We simulate the system in the time domain with an input as: $u(t) = A \cos(\omega t)$, magnitude $A = 0.01$, frequency $\omega \in [0.5 \ 1 \ 1.5 \ 2] 2\pi$, and time step $dt = 1e - 4$. Next, the second-stage approximation results for the linear kernel \tilde{H}_1 in comparison with the theoretical values of H_1 are presented in Table 1.

With the estimations of the linear transfer function and by using the LF as the data-driven identification and reduction tool for linear systems, we identify the linear system $(\tilde{\mathbf{A}}, \tilde{\mathbf{b}}, \tilde{\mathbf{c}}, \tilde{\mathbf{E}})$. We stopped at the fourth measurement due to the fact that the underlying system is of second order (McMillan degree 2). Otherwise, more measurements will be needed to have a sufficient decay of the singular values as shown in Fig. 6. The singular value decay offers a choice of reduction. As long as the simulation of the system is done, with time step $dt = 1e - 4$, the singular values with magnitude below that threshold are neglected.

Construction of the linear system with order $r = 2$, by using the theoretical noise-free measurements (subscript “t”) appears next:

$$\tilde{\mathbf{A}}_t = \begin{bmatrix} -1.4513 & -8.8181 \\ 11.363 & -0.54868 \end{bmatrix}, \tilde{\mathbf{B}}_t = \begin{bmatrix} -0.92979 \\ 1.3967 \end{bmatrix}, \tilde{\mathbf{C}}_t = [-0.76857 \ 0.9203], \quad (43)$$

while by using the measured data with second-stage approximation results to the following:

$$\tilde{\mathbf{A}} = \begin{bmatrix} -1.458 & -8.8137 \\ 11.367 & -0.55162 \end{bmatrix}, \tilde{\mathbf{B}} = \begin{bmatrix} -0.9342 \\ 1.4 \end{bmatrix}, \tilde{\mathbf{C}} = [-0.7675 \ 0.91611]. \quad (44)$$

Table 1 Measurements of the first (linear) kernel

Frequency ω	$\tilde{H}_1(j\omega)$ -second stage	$H_1(j\omega)$ -theoretical
$0.5 \cdot 2\pi$	$+0.026606 + 0.067106i$	$+0.026574 + 0.067115i$
$1.0 \cdot 2\pi$	$+0.071503 + 0.189600i$	$+0.071258 + 0.189700i$
$1.5 \cdot 2\pi$	$+0.752720 + 0.377300i$	$+0.754030 + 0.380870i$
$2.0 \cdot 2\pi$	$+0.134070 - 0.381970i$	$+0.133780 - 0.382520i$

^aWith 2nd-stage approximation $\tilde{H}_1(j\omega) \approx 2Y_1(j\omega)/A$

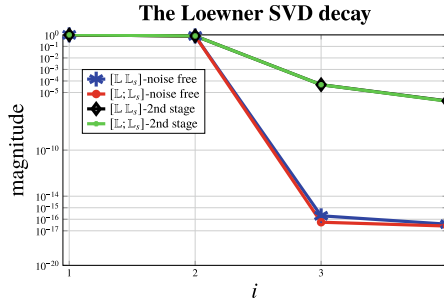


Fig. 6 The singular value decay of the LF as a fundamental characterization of the McMillan degree of the underlying linear system. Here, a truncation scheme of order $r = 2$ is recommended where the second stage approximation gave $\sigma_3/\sigma_1 = 4.721 \cdot 10^{-5}$, while for the noise-free case the third singular values have reached the machine precision

Table 2 Measurements of the H_2 on the diagonal and perpendicular to the diagonal

Freq. ω	$\tilde{H}_2(j\omega, j\omega)$	$H_2(j\omega, j\omega)$	$\tilde{H}_2(j\omega, -j\omega)$	$H_2(j\omega, -j\omega)$
$0.5 \cdot 2\pi$	$+0.026440 - 0.124490i$	$+0.026570 - 0.124440i$	$+0.032190$	$+0.032177$
$1.0 \cdot 2\pi$	$-0.184590 + 0.298430i$	$-0.184510 + 0.298910i$	$+0.045648$	$+0.045641$
$1.5 \cdot 2\pi$	$+0.178080 + 0.305840i$	$+0.178160 + 0.307170i$	$+0.063936$	$+0.064350$
$2.0 \cdot 2\pi$	$+0.062642 - 0.054219i$	$+0.062588 - 0.054423i$	-0.044927	-0.044998

^bThe estimations of the second kernel are given as: $\tilde{H}_2(j\omega, j\omega) \approx 4Y_2(j\omega, j\omega)/A^2$, on the diagonal, and $\tilde{H}_2(j\omega, -j\omega) \approx 2Y_2(j\omega, -j\omega)/A^2$, which is the DC term

> Identified linear dynamics

Even if the coordinate system is different, one crucial qualitative result is to compute the poles and zeros of the linear transfer function. For the identified system with the theoretical measurements (noise free), the poles and zeros are exactly as the original: $\tilde{p}_t = -1 \pm 10i$ and the zero is: $\tilde{z}_t = -1$ while for the second-stage approximation to the linear system, the corresponding results are: $\tilde{p} = -1.0048 \pm 9.9989i$, $\tilde{z} = -1.0042$.

At this point, we have recovered the linear part of the bilinear system up to an accuracy due to the truncation of Volterra series. The inexact simulations of the continuous system which are done with a finite time step $dt = 1e - 4$, and the Fourier accuracy led to quite accurate results with a perturbation of the order $\sim O(1e - 3)$ by comparing the theoretical poles and zeros. We proceed by collecting the measurements of the second kernel. Table 2, contains measurements of the second kernel with one-tone input.

We can get \mathbf{N} by solving the least-squares problem by just minimizing the 2-norm as in [28]. This result was not toward the identification of the matrix \mathbf{N} and here is the new approach working toward the identification of bilinear systems.

? Can we identify the matrix \mathbf{N} ?

The improvement relies on the rank deficiency problem that is produced by getting the least-squares solution without taking under consideration measurements out of the diagonal of the second kernel H_2 . By filling in the least-squares problem in Eq. (37) with these extra equations, as Proposition 1 indicates, the problem solution upgrades to a full-rank inversion and the answer is affirmative.

Back to our introductory example, the rank of the least-squares problem is less than $r^2 = 4$. So, we need to increase the rank. We take measurements (≤ 4) out of the diagonal from the second kernel by using the input $u(t) = A_1 \cos(\omega_1) + B_1 \cos(\omega_2)$. Table 3 includes the theoretical and measured results.

The full-rank least-squares solution gave for the theoretical noise-free case and for the second-stage approximation the following results, respectively:

$$\tilde{\mathbf{N}}_r = \begin{bmatrix} -4.1542 & -2.0998 \\ 3.236 & 1.1542 \end{bmatrix}, \quad \tilde{\mathbf{N}} = \begin{bmatrix} -4.1557 & -2.1084 \\ 3.2284 & 1.1513 \end{bmatrix} \quad (45)$$

> Coordinate transformation

By transforming all the matrices to the same coordinate system as in [26], we conclude to the

• Noise-free case—exact identification

$$\check{\mathbf{A}}_r = \begin{bmatrix} -1.0 & -10.0 \\ 10.0 & -1.0 \end{bmatrix}, \quad \check{\mathbf{N}}_r = \begin{bmatrix} 1.0 & -2.0 \\ 3.0 & -4.0 \end{bmatrix}, \quad \check{\mathbf{B}}_r = \begin{bmatrix} 1.0 \\ 1.0 \end{bmatrix}, \quad \check{\mathbf{C}}_r = \begin{bmatrix} 1.0 \\ 1.0 \end{bmatrix}^T. \quad (46)$$

• Simulated case—approximated identification

$$\check{\mathbf{A}} = \begin{bmatrix} -1.0037 & -9.9941 \\ 10.004 & -1.0059 \end{bmatrix}, \quad \check{\mathbf{N}} = \begin{bmatrix} 0.99525 & -1.997 \\ 3.006 & -3.9997 \end{bmatrix}, \quad \check{\mathbf{B}} = \begin{bmatrix} 0.99925 \\ 1.0003 \end{bmatrix}, \quad \check{\mathbf{C}} = \begin{bmatrix} 1.0 \\ 1.0 \end{bmatrix}^T. \quad (47)$$

Next, in Fig. 7, evaluation results for the linear and the second-order generalized transfer function are presented:

Finally, time-domain simulations for each system performed in Fig. 8 with a larger amplitude than the probing one.

Table 3 Measurements of the second kernel (out of the diagonal)

Frequencies (ω_1, ω_2)	$\tilde{H}_2(j\omega_1, j\omega_2)$	$H_2(j\omega_1, j\omega_2)$
$(0.2 \cdot 2\pi, 0.3 \cdot 2\pi)$	$+0.030440 - 0.039259i$	$+0.030429 - 0.039237i$
$(0.2 \cdot 2\pi, 0.6 \cdot 2\pi)$	$+0.031002 - 0.080364i$	$+0.031037 - 0.080315i$
$(0.4 \cdot 2\pi, 0.3 \cdot 2\pi)$	$+0.030948 - 0.062869i$	$+0.030961 - 0.062835i$
$(0.4 \cdot 2\pi, 0.6 \cdot 2\pi)$	$+0.026417 - 0.125320i$	$+0.026554 - 0.125260i$

^cThe estimation of the second kernel as $\tilde{H}_2(j\omega_1, j\omega_2) \approx 2Y_2(j\omega_1, j\omega_2)/(A_1 B_1)$. Here we use $\phi = 1.5$, to avoid the harmonic overlapping as explained in Sect. 4.4 and amplitudes as $A_1 = B_1 = 0.01$

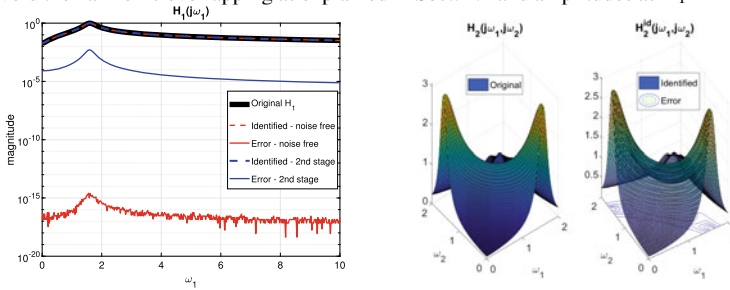


Fig. 7 The identified first and second kernel with second-stage approximation in comparison with the theoretical kernels

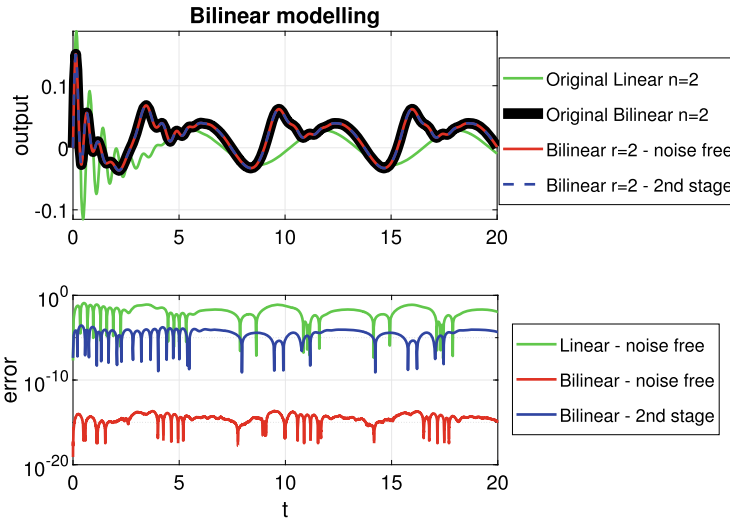


Fig. 8 The evaluation of the models with order $r = 2$ performed with input as $u(t) = \cos(t)$, $t \in [0, 20]$. The noise-free case has reached machine precision

Example 2 Time-domain reduction of the Burgers’ Equation. This example illustrates the bilinear modeling and reduction concepts proposed in [5] for the viscous Burgers’ equation from time-domain simulations. We simulate the system with 40 measurements as $\omega_k = j2\pi[0.1, 0.2, \dots, 4]$. We present the corre-

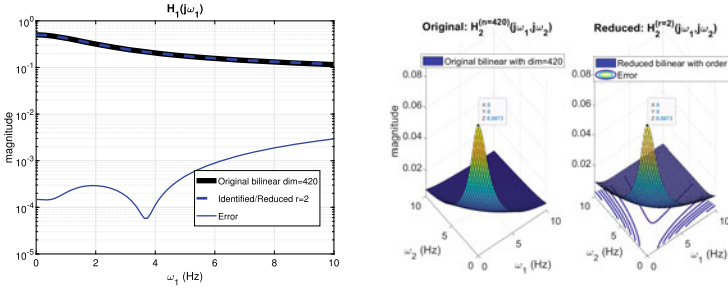


Fig. 9 The first and the second kernel evaluations in comparison with the originals

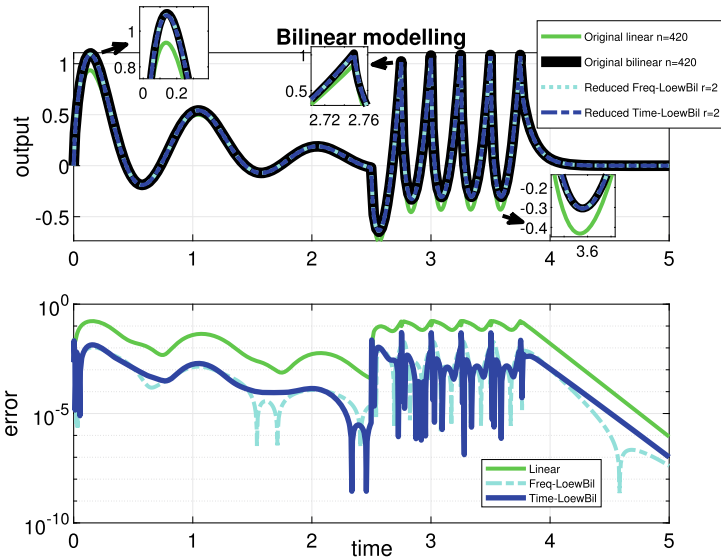


Fig. 10 Time-domain simulation for the Burgers’ equation example; viscosity parameter ν is set as 1 and the dimension of the semi-discretized model is chosen to be 420. A comparison among the identified/reduced bilinear of order $r = 2$ with the linear and with the frequency-domain Loewner bilinear is depicted. The input is chosen as: $u(t) = (1 + 2 \cos(2\pi t))e^{-t}$, $t \in [0, 2.5]$, $u(t) = 4\text{sawtooth}(8\pi t)$, $t \in [2.5, 3.75]$, $u(t) = 0$, $t \in [3.75, 5]$

sponding results with initial system dimension $n = 420$ reduced by the proposed method to order $r = 2$ with the first normalized neglected singular value to be $\sigma_3/\sigma_1 = 4.6255 \cdot 10^{-4}$. As the order was chosen $r = 2$, the reduced bilinear matrix \tilde{N} was introduced by using the following measurements as $\omega_1 = j2\pi [0.2, 0.4]$ and $\omega_2 = j2\pi [0.3, 0.6]$. In Fig. 9, evaluation results are presented.

Lastly, in Fig. 10, a time-domain simulation reveals that the proposed method can improve the accuracy by fitting a nonlinear model. Table 4 contains approximation results both in the frequency and, also in the time-domain. For the example presented (dimension reduction from $n = 420$ to $r = 2$), we offer a comparison of the

Table 4 Summary of the results from the two examples with Time-LoewBil and comparison with [5] for Burgers' Example 2 of dimension $n = 420$

Error quantification	Time-LoewBil Example 1	Time-LoewBil Example 2	Freq-LoewBil Example 2
$\max_{\omega} \ H_1(j\omega) - \tilde{H}_1(j\omega)\ $	$5.077 \cdot 10^{-3}$	$2.937 \cdot 10^{-3}$	$4.430 \cdot 10^{-3}$
$\max_t \ y(t) - y_l(t)\ $	$1.213 \cdot 10^{-1}$	$1.699 \cdot 10^{-1}$	$1.699 \cdot 10^{-1}$
$\max_{(\omega_1, \omega_2)} \ H_2(j\omega_1, j\omega_2) - \tilde{H}_2(j\omega_1, j\omega_2)\ $	$2.794 \cdot 10^{-2}$	$3.077 \cdot 10^{-3}$	$2.991 \cdot 10^{-3}$
$\max_t \ y(t) - \tilde{y}_b(t)\ $	$2.739 \cdot 10^{-4}$	$5.032 \cdot 10^{-2}$	$5.278 \cdot 10^{-2}$

^dThe evaluations of the kernels and the outputs (y_l : linear, \tilde{y}_b reduced bilinear ($r = 2$)) took place over the domains depicted in Figs. 7, 8, 9, 10

newly proposed method (Time-LoewBil) with another method, i.e., the frequency-domain bilinear Loewner framework introduced in [5] (Freq-LoewBil). The common frequency grid was selected as described above while the sampling values of the transfer functions (in the frequency-domain) were corrupted with white-noise. The noise magnitude of the latter was selected to match the noise values introduced by performing time-domain simulations with a time step of $dt = 1e - 4$.

Remark 4 (Computational cost for the discretized Burgers' model of dimension 420) The proposed time-domain Loewner bilinear method uses measurements corresponding to symmetric transfer functions. Such values can be directly inferred from time-domain data by processing the spectral domain, i.e., by computing the FFT of the observed output signals for oscillatory input signals. All experiments were performed on a computer with 12 GB RAM and an Intel(R) Core(TM) i7-10510U CPU running at 1.80 GHz, 2304 Mhz, 4 Cores, 8 Logical Processors. To simulate a system of dimension 420, each measurement took ~ 3 min. So, the data acquisition cost was reported in the range of 1 or 2h where the identification/reduction part was almost direct. The proposed method seems to be efficient for moderate dimensions; for large-scale problems, the computational issues that appear belong to the class of "embarrassingly parallel" tasks; as the simulations are independent to each other, one can easily speed up the whole process by using instead parallel clusters.

Remark 5 (Discussion and comparison between the two methods) In what follows, we will state the pluses and minuses of the two methods applied for the second numerical example.

The frequency Loewner bilinear framework (Freq-LoewBil)

- Pluses: recovers the original bilinear system with high accuracy, incorporates linear and nonlinear transfer function measurements in a coupled way ("all at once"), can be easily extended to cope with higher order regular kernels, can also be viewed as a Petrov-Galerkin projection-based moment-matching approach.

- Minuses: It is not completely clear how to measure/obtain the frequency-domain data needed for this method; it uses measurements of regular transfer functions which cannot be (directly) inferred from time-domain simulations.

The time-Loewner bilinear framework (Time-LoewBil)

- Pluses: It uses measurements corresponding to symmetric transfer functions. Such values can be directly inferred from time-domain data by processing the spectral domain, i.e., by computing the FFT of the observed output signals for oscillatory input signals.
- Minuses: The fitted bilinear model is as good as the fitted linear model (it relies on the linear fit). As opposed to the first method, it fits the linear and nonlinear parts separately (not “all at once”). It introduces additional errors due to conversion from the time domain to the frequency domain. The latter disadvantage could also occur for the method in [5], provided that “regular transfer function” measurements could be successfully inferred from time-domain data.

6 Conclusion

The proposed method offers approximate bilinear system identification from time-domain measurements, since it is not possible to measure the corresponding kernels exactly. An adaptive scheme that improves the estimation of the kernels was presented. Our proposed method uses only *input-output* measurements without requiring state-space access. What makes this algorithm feasible is the combination of the data-driven Loewner framework with the nonlinear Volterra series framework.

We have shown that for the noise-free case, the proposed method achieves system identification from time-domain measurements through the symmetric kernels. Further study is required to quantify the effects of the noise introduced by the truncation of the Volterra series (in the ℓ -stage approximation). All the time-domain numerical simulations have been implemented by means of the backward Euler approximation scheme which certifies that this method can handle some level of numerical noise. Multi-stepping methods, e.g., Runge-Kutta can offer a significance improvement to the results and reduce the influence of numerical noise.

The variational approach is a theoretical method to identify regular kernels which are appropriate for system identification purposes [35]. However, these kernels do not have a physical meaning, i.e., cannot be directly measured from time-domain simulations. This is not an issue for the growing exponential approach. The derived transfer functions by means of this method can be measured from time-domain data. The difficulty in combining both derivations, i.e., symmetric and regular is also explained from the n th-dimensional integral that connects those through the triangular kernels. Extensions to the MIMO case and to other nonlinearity structures, e.g., quadratic or bilinear quadratic etc., are promising endeavors that will be the matter of future research.

References

1. Ahmad, M.I., Baur, U., Benner, P.: Implicit Volterra series interpolation for model reduction of bilinear systems. *J. Comput. Appl. Math.* 316, 15–28 (2017). <https://doi.org/10.1016/j.cam.2016.09.048>
2. Anderson, B.D.O., Antoulas, A.C.: Rational interpolation and state-variable realizations. *Linear Algebra Appl.* 137/138, 479–509 (1990)
3. Antoulas, A.C.: Approximation of large-scale dynamical systems. *Advances in Design and Control*, vol. 6. SIAM Publications, Philadelphia, PA (2005). <https://doi.org/10.1137/1.9780898718713>
4. Antoulas, A.C., Beattie, C.A., Güğercin, S.: *Interpolatory Methods for Model Reduction*. Society for Industrial and Applied Mathematics, Philadelphia, PA (2020). <https://doi.org/10.1137/1.9781611976083>
5. Antoulas, A.C., Gosea, I.V., Ionita, A.C.: Model reduction of bilinear systems in the Loewner framework. *SIAM J. Sci. Comput.* **38**(5), B889–B916 (2016). <https://doi.org/10.1137/15M1041432>
6. Antoulas, A.C., Lefteriu, S., Ionita, A.C.: Chapter 8: A Tutorial Introduction to the Loewner Framework for Model Reduction, pp. 335–376. <https://doi.org/10.1137/1.9781611974829.ch8>
7. Bai, Z., Skoogh, D.: A projection method for model reduction of bilinear dynamical systems. *Linear Algebra Appl.* 415(2–3), 406–425 (2006)
8. Bartee, J.F., Georgakis, C.: Bilinear identification of nonlinear processes. *IFAC Proc. Vol.* **27**(2), 47–52 (1994). [https://doi.org/10.1016/S1474-6670\(17\)48128-6](https://doi.org/10.1016/S1474-6670(17)48128-6), <http://www.sciencedirect.com/science/article/pii/S1474667017481286>. IFAC Symposium on Advanced Control of Chemical Processes, Kyoto, Japan, 25–27 May 1994
9. Baur, U., Benner, P., Feng, L.: Model order reduction for linear and nonlinear systems: a system-theoretic perspective **21**(4), 331–358 (2014). <https://doi.org/10.1007/s11831-014-9111-2>
10. Benner, P., Breiten, T.: Interpolation-based \mathcal{H}_2 -model reduction of bilinear control systems. *SIAM J. Matrix Anal. Appl.* 33(3), 859–885 (2012)
11. Benner, P., Breiten, T., Damm, T.: Generalized tangential interpolation for model reduction of discrete-time MIMO bilinear systems. *Internat. J. Control* 84(8), 1398–1407 (2011). DOI: 10.1080/00207179.2011.601761
12. Benner, P., Damm, T.: Lyapunov equations, energy functionals, and model order reduction of bilinear and stochastic systems. *SIAM J. Control Optim.* 49(2), 686–711 (2011). DOI: 10.1137/09075041X
13. Benner, P., Goyal, P., Heiland, J., Pontes Duff, I.: Operator inference and physics-informed learning of low-dimensional models for incompressible flows. e-prints 2010.06701, arXiv (2020). <http://arxiv.org/abs/2010.06701>. Math.DS
14. Benner, P., Goyal, P., Kramer, B., Peherstorfer, B., Willcox, K.: Operator inference for non-intrusive model reduction of systems with non-polynomial nonlinear terms. *Comp. Meth. Appl. Mech. Eng.* 372, 113433 (2020). DOI: 10.1016/j.cma.2020.113433
15. Benner, P., Gugercin, S., Willcox, K.: A survey of model reduction methods for parametric systems. *SIAM Review* **57**(4), 483–531 (2015). <https://doi.org/10.1137/130932715>
16. Boyd, S., Shing Tang, Y., Chua, L.O.: *Measuring Volterra Kernels* (1983)
17. Breiten, T.: Interpolatory methods for model reduction of large-scale dynamical systems. Dissertation, Department of Mathematics, Otto-von-Guericke University, Magdeburg, Germany (2013)
18. Brubeck, P.D., Nakatsukasa, Y., Trefethen, L.N.: *Vandermonde with Arnoldi* (2019)
19. Drmač, Z., Peherstorfer, B.: Learning low-dimensional dynamical-system models from noisy frequency-response data with Loewner rational interpolation (2019)
20. Flagg, G.M., Gugercin, S.: Multipoint Volterra series interpolation and \mathcal{H}_2 optimal model reduction of bilinear systems. *SIAM J. Numer. Anal.* 36(2), 549–579 (2015). <https://doi.org/10.1137/130947830>
21. Fosong, E., Schulze, P., Unger, B.: From time-domain data to low-dimensional structured models (2019)

22. Gosea, I.V., Antoulas, A.C.: Data-driven model order reduction of quadratic-bilinear systems. *Numerical Linear Algebra Appl.* **25**(6), e2200 (2018). <https://doi.org/10.1002/nla.2200>. E2200 nla.2200
23. Gustavsen, B., Semlyen, A.: Rational approximation of frequency domain responses by vector fitting. *IEEE Transactions on Power Delivery* **14**(3), 1052–1061 (1999). <https://doi.org/10.1109/61.772353>
24. Ionita, A.C.: Matrix pencils in time and frequency domain system identification, pp. 79–88 (2012). https://doi.org/10.1049/pbce076e_ch9
25. Isidori, A.: Direct construction of minimal bilinear realizations from nonlinear input-output maps. *IEEE Transactions on Automatic Control* **18**(6), 626–631 (1973)
26. Juang, J.N.: Continuous-time bilinear system identification. *Nonlinear Dynamics* **39**(1), 79–94 (2005). <https://doi.org/10.1007/s11071-005-1915-z>
27. Kaiser, E., Kutz, J.N., Brunton, S.L.: Sparse identification of nonlinear dynamics for model predictive control in the low-data limit. *Proc. R. Soc. A: Math. Phys. Eng. Sci.* **474**(2219), 20180335 (2018). <https://doi.org/10.1098/rspa.2018.0335>
28. Karachalios, D.S., Gosea, I.V., Antoulas, A.C.: A bilinear identification-modeling framework from time domain data. *Proc. Appl. Math. Mech.* **19**(1), e201900246 (2019). DOI: 10.1002/pamm.201900246
29. Karachalios, D.S., Gosea, I.V., Antoulas, A.C.: The Loewner framework for system identification and reduction. In: Benner, P., Grivet-Talocia, S., Quarteroni, A., Rozza, G., Schilders, W.H.A., Silveira, L.M. (eds.), *Handbook on Model Reduction*, volume I of *Methods and Algorithms* (in press)
30. Lefteriu, S., Ionita, A.C., Antoulas, A.C.: *Modeling Systems Based on Noisy Frequency and Time Domain Measurements*, pp. 365–378. Springer, Berlin, Heidelberg (2010). https://doi.org/10.1007/978-3-540-93918-4_33
31. Peherstorfer, B., Gugercin, S., Willcox, K.: Data-driven reduced model construction with time-domain Loewner models. *SIAM J. Sci. Comput.* **39**(5), A2152–A2178 (2017). <https://doi.org/10.1137/16M1094750>
32. Peherstorfer, B., Willcox, K.: Data-driven operator inference for nonintrusive projection-based model reduction. *Comput. Methods Appl. Mech. Eng.* **306**, 196–215 (2016). <https://doi.org/10.1016/j.cma.2016.03.025>, <http://www.sciencedirect.com/science/article/pii/S0045782516301104>
33. Petkovska, M., Nikolić, D., Seidel-Morgenstern, A.: Nonlinear frequency response method for evaluating forced periodic operations of chemical reactors. *Israel J. Chem.* **58**(6-7), 663–681 (2018). <https://doi.org/10.1002/ijch.201700132>
34. Phillips, J.R.: Projection-based approaches for model reduction of weakly nonlinear, time-varying systems **22**(2), 171–187 (2003)
35. Rugh, W.J.: *Nonlinear System Theory: The Volterra/Wiener Approach*. The Johns Hopkins University Press, Baltimore (1981)
36. Scarciotti, G., Astolfi, A.: Data-driven model reduction by moment matching for linear and nonlinear systems. *Automatica* **79**, 340–351 (2017). <https://doi.org/10.1016/j.automatica.2017.01.014>, <http://www.sciencedirect.com/science/article/pii/S0005109817300249>
37. Schmid, P.J.: Dynamic mode decomposition of numerical and experimental data. *Journal of Fluid Mechanics* **656**, 5–28 (2010). <https://doi.org/10.1017/S0022112010001217>