On Bilinear Time Domain Identification

D.S. Karachalios[*], I.V. Gosea[*], and A.C. Antoulas[* † ‡]

[*]Max Planck Institute for Dynamics of Complex Technical Systems, Data-Driven System Reduction and Identification (DRI) group, Sandtorstrasse 1, 39106 Magdeburg, Germany.
[†]Rice University Houston, Electrical and Computer Engineering Department, 6100 Main St., Houston, TX 77005, USA.
[‡]Baylor College of Medicine, 1 Baylor Plaza, Houston, TX 77030, USA.

March 20, 2020

Abstract

The *Loewner framework*-(LF) in combination with *Volterra series*-(VS) offers a non-intrusive approximation method that is capable to identify bilinear models from time domain measurements.

1 Introduction

In natural sciences, evolutionary phenomena can be modelled as dynamical systems. An ever-increasing need for improving the approximation accuracy has motivated including more involved and detailed features in the modelling process, thus inevitably leading to largerâĂŘscale 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 states 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 [8, 12].

One way to perform model reduction is by employing *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, 24] and on the sensitivity of noise in the (LF) are given in [25, 15].

When *input-output* data are offered, *data-driven* methods such as the (LF), (Dynamic Mode Decomposition)-(DMD) [31], sparse identification of non-linear systems - SINDYc in [22], *Vector Fitting*-(VF) [19], Hankel or *subspace* methods [20, 21] and *operator inference* [27], remain the only feasible approaches for recovering the hidden information.

DMD-based methods represent viable alternatives that require state-derivative estimations. In such cases, the library construction allows a broad non-linear identification. Nevertheless, the dimension of this library could be very large when accurate discretization of the physical domain is performed.

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, prediction). At the same time, these discovered models might have large dimension and hence are not suitable for fast numerical simulation. 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 non-linear systems (e.g., bilinear or quadratic-bilinear systems) see [5, 18]. In the case of *time domain* data, the (LF) was already applied for approximating linear models [20, 26, 17]. The mathematical description of dynamical systems uses the concept of DAEs[1] which can simulate the *input*-$u(t)$ to *output*-$y(t)$ relation as in Fig. 1, where the differential operator denoted as $\mathbf{f}$ and the algebraic operator denoted as $\mathbf{z}$.

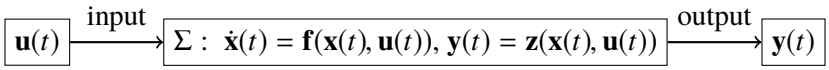

Figure 1: Mathematical formalism for evolutionary phenomena.

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).

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

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 equation[2]. The resulting system (1) is:

$$\Sigma_l : \begin{cases} \dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}u(t), \\ y(t) = \mathbf{C}\mathbf{x}(t),\ t \geq 0. \end{cases} \tag{2}$$

The explicit solution with the *convolution integral*[3] notation and the time domain linear kernel $h(t)$ as the *impulse response* of the system can be written as:

$$y(t) = \mathbf{C}e^{\mathbf{A}t}\mathbf{x}(0) + (\mathbf{h} * u)(t),\ t \geq 0. \tag{3}$$

By assuming zero initial conditions and performing a *Laplace transform*, we obtain the transfer function description:

$$H(s) = \frac{Y(s)}{U(s)} = \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B},\ s \in \mathbb{C}, \tag{4}$$

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 [30]. Other relevant works on nonlinear systems and nonlinear modelling/identification include Schetzen (1980), Chen and Billings (1989), Boyd and Chua (1985) et. al.

[1] DAEs: Differential Algebraic Equations.

[2] The state-output equation often is represented as $y(t) = \mathbf{C}\mathbf{x}(t) + Du(t)$.

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

The aim in this study is to identify and reduce special types of non-linear systems (s.a., bilinear) from time domain measurements. By knowing only the input and the simulated or measured output in the time domain as in Fig. 2, we will identify the hidden model which explains this phenomenon. In such situations where only snapshots are available, beyond the linear fit which is very well established to a series of methods, a non-linear fit of a special type will be able to capture the hidden non-linearity.

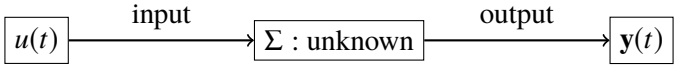

Figure 2: The input-output mapping from the data-driven perspective with the unknown system Σ. Although specific structures of the unknown system can be assumed/inspired by the physical problem. For instance, if the underlying physical phenomenon is fluid flow inside a control volume, quadratic models should be constructed e.g., [18].

2.2.1 Approximation of non-linear systems (Volterra series)

The *input-output* relationship for a wide class of non-linear systems [30] can be approximated by a Volterra series up to a sufficiently high N as:

$$y(t) = \sum_{n=1}^{N} y_n(t), \quad y_n(t) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} h_n(\tau_1, \cdots, \tau_n) \prod_{i=1}^{n} u(t-\tau_i) d\tau_i, \tag{5}$$

where $h_n(\tau_1, \cdots, \tau_n)$ is a *real-valued* function of $\tau_1, \cdots, \tau_n$ known as the n^{th} order Volterra kernel.

Definition 2.1. The n^{th} order generalized frequency response function (GFRF) is defined as:

$$H_n(j\omega_1, \cdots, j\omega_n) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} h_n(\tau_1, \cdots, \tau_n) e^{\left(-j \sum_{i=1}^{n} \omega_i \tau_i\right)} d\tau_1 \cdots d\tau_n, \tag{6}$$

which is the multidimensional Fourier[4] transform of $h_n(\tau_1, \cdots, \tau_n)$.

By applying the inverse Fourier transform of the n^{th} order GFRF, Eq. 6 can be written as:

$$y_n(t) = \frac{1}{(2\pi)^n} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} H_n(j\omega_1, \cdots, j\omega_n) \prod_{i=1}^{n} U(j\omega_i) e^{j(\omega_1 + \cdots + \omega_n)t} d\omega_1 \cdots \omega_n. \tag{7}$$

The n^{th} Volterra operator is defined:

$$V_n(u_1, u_2, ..., u_n) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} h_n(\tau_1, ..., \tau_n) \prod_{i=1}^{n} u_i(t-\tau_i) d\tau_i, \tag{8}$$

so that it holds $y_n = V_n(u, u, ..., u)$.

Remark 2.1. Homogeneity of the Volterra operator The map $u(t) \to y_n(t)$ is homogeneous of degree n, that is, $\alpha u \to \alpha^n y_n$. Each Volterra kernel $h_n(t)$ determines a symmetric multi-linear operator. This will allow to order the non-linear terms in such a way that larger powers of the amplitude (α^n) will eliminate the corresponding dynamics term. That is precisely the sense of approximating weakly non-linear systems with Volterra series (i.e., with $0 < \alpha < 1$, it holds $\alpha^n \to 0$ as $n \to \infty$).

[4] As the frequency $s = j\omega_1$ lies on the imaginary axis, the Laplace transform simplifies to Fourier transform.

2.2.2 A single sinusoidal input

Consider the excitation of a system with an input consisting of two complex exponentials as in Eq. 9. Such inputs are typically used in chemical engineering applications as [28].

$$u(t) = A\cos(\omega t) = \left(\frac{A}{2}\right)e^{j\omega t} + \left(\frac{A}{2}\right)e^{-j\omega t}, \;\; (j^2 = -1). \tag{9}$$

By using the above input in Eq. 5, we can derive the first Volterra term with $n = 1$ as:

$$\begin{aligned} y_1(t) &= \int_{-\infty}^{\infty} h_1(\tau_1)[u(t-\tau_1)]d\tau_1 \\ &= \frac{A}{2}e^{j\omega t}\underbrace{\int_{-\infty}^{\infty} h_1(\tau_1)e^{-j\omega\tau_1}d\tau_1}_{H_1(j\omega)} + \frac{A}{2}e^{-j\omega t}\underbrace{\int_{-\infty}^{\infty} h_1(\tau_1)e^{j\omega\tau_1}d\tau_1}_{H_1(-j\omega)} \Rightarrow \\ y_1(t) &= \frac{A}{2}\Big(e^{j\omega t}H_1(j\omega) + e^{-j\omega t}H_1(-j\omega)\Big). \end{aligned} \tag{10}$$

Similarly, for the second term we can derive:

$$y_2(t) = \left(\frac{A}{2}\right)^2\Big[e^{2j\omega}H_2(j\omega, j\omega) + 2e^0 H_2(j\omega, -j\omega) + e^{-2j\omega}H_2(-j\omega, -j\omega)\Big]. \tag{11}$$

Remark 2.2. (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 FRF[5] is $H_1(j\omega)$, and $H_n(j\omega, ..., j\omega)$, for $n \geq 2$, are the higher-order FRFs or GFRFs. Then, the n^{th} Volterra term can be written as:

$$y_n(t) = \left(\frac{A}{2}\right)^n \sum_{p+q=n} {}^nC_q H_n^{p,q}(j\omega)e^{j\omega_{p,q}t}, \; \omega_{p,q} = (p-q)\omega. \tag{12}$$

where the following brief notations have been used:

$$H_n^{p,q}(j\omega) = H_n(\underbrace{j\omega, ..., j\omega}_{p-times}; \underbrace{-j\omega, ..., -j\omega}_{q-times}), \; \omega_{p,q} = (p-q)\omega, \; {}^nC_q = \frac{n!}{q!(n-q)!}. \tag{13}$$

2.2.3 Time domain representation of harmonics

The n^{th} harmonic in the *time domain* is:

$$y_{n^{th}}(t) = \sum_{i=1}^{\infty} 2\left(\frac{A}{2}\right)^{n+2i-2} {}_{n+2i-2}C_{i-1}H_{n+2i-2}^{n+i-1,i-1}(j\omega)e^{jn\omega t}. \tag{14}$$

2.2.4 Frequency domain representation of harmonics

The n^{th} harmonic in the *frequency domain* by applying Fourier transform in Eq. 14 is the following:

$$Y_{n^{th}}(jn\omega) = \sum_{i=1}^{\infty} 2\left(\frac{A}{2}\right)^{n+2i-2} {}_{n+2i-2}C_{i-1}H_{n+2i-2}^{n+i-1,i-1}(j\omega)\delta(jn\omega). \tag{15}$$

where $\delta(\cdot)$ is the Dirac delta distribution.

[5]FRF: Frequency Response Function and GFRF: Generalized FRF.

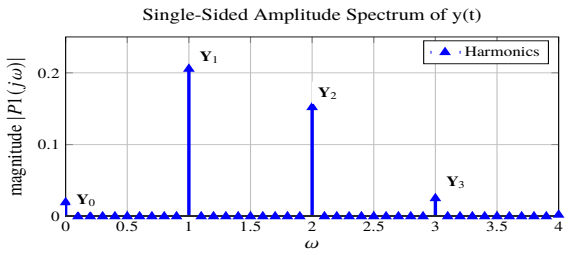

Figure 3: An instance of the single-sided power spectrum with a singleton input with $\omega = 1$ is depicted. The underlying system is non-linear and as a result higher harmonics appeared with a DC (Direct Current - non-periodic) term as well.

3 The Loewner framework

We start with an account of the Loewner framework (LF) in the linear case [2, 6, 24]. The (LF) as an interpolation method 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 data (abstract frequencies) as:

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

We partition the data in two disjoint sets:

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

where $\mu_i = s_i$, $\lambda_i = s_{n+i}$, $v_i = f_i$, $w_i = f_{n+i}$ for $i = 1, \ldots, n$.
The objective is to find $H(s) \in \mathbb{C}$ such that:

$$H(\mu_i) = v_i, \ i = 1, \ldots, n, \ \ H(\lambda_j) = w_j, \ j = 1, \ldots, n. \tag{16}$$

The *left data* is rearranged compactly as:

$$\mathbf{M} = [\mu_1, \cdots, \mu_n] \in \mathbb{C}^{1\times n}, \ \ \mathbb{V} = [v_1, \cdots, v_n] \in \mathbb{C}^{1\times n}, \tag{17}$$

while the *right data* as:

$$\mathbf{\Lambda} = [\lambda_1, \cdots, \lambda_n]^T \in \mathbb{C}^{n\times 1}, \ \ \mathbb{W} = [w_1, \cdots, w_n]^T \in \mathbb{C}^{n\times 1}. \tag{18}$$

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 [24].

3.1 The Loewner matrix

Given a row array of complex numbers (μ_j, v_j), $j = 1, \ldots, n$, and a column array, (λ_i, w_i), $i = 1, \ldots, n$, (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:

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

Definition 3.1. 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 3.1. [2] Let $\mathbb{L}$ be as above. If $k, q \geq \deg g$, then $\operatorname{rank}\mathbb{L} = \deg g$. 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., $H(s) = \mathbb{W}(\mathbb{L}_s - s\mathbb{L})^{-1}\mathbb{V}$, interpolates the data. Otherwise, as shown in [2], the problem in Eq. 16 has a solution provided that

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

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

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

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

Remark 3.1. r can be chosen as the *numerical rank* (as opposed to the *exact rank*) of the Loewner pencil.

Theorem 3.2. The quadruple $(\tilde{\mathbf{E}}, \tilde{\mathbf{A}}, \tilde{\mathbf{B}}, \tilde{\mathbf{C}})$ of size $r\times r$, $r\times r$, $r\times 1$, $1\times r$, given by:

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

is a descriptor realization of an (approximate) interpolant of the data with McMillan degree $r = rank(\mathbb{L})$, 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 [6, 24] 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 non-linear 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 [5, 1, 7, 9, 10, 11, 14, 16, 29] and the references within.

Bilinear systems are described by the following set of matrices; $\Sigma_b = (\mathbf{C}, \mathbf{E}, \mathbf{A}, \mathbf{N}, \mathbf{B})$ and characterized by the following equations:

$$\Sigma_b : \begin{cases}\mathbf{E}\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{N}\mathbf{x}(t)u(t) + \mathbf{B}u(t),\\ \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t),\end{cases} \tag{19}$$

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 $\mathbf{x} \in \mathbb{R}^n$, $u, y \in \mathbb{R}$. In what follows, we restrict our analysis to systems with non-singular $\mathbf{E}$ matrices. Moreover, this matrix can be considered to be the 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* descriptions from constant-parameter (stationary) state equations. Let us consider the bilinear model in Eq. 19 with zero initial conditions.
A single oscillatory input with amplitude $A < 1$ is considered as in Eq. 9.

$$u(t) = A\cos(\omega t) = \frac{A}{2}e^{j\omega t} + \frac{A}{2}e^{-j\omega t} = ae^{j\omega t} + ae^{-j\omega t}, \tag{20}$$

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. 19 can be written as:

$$\mathbf{x}(t) = \sum_{p,q\in\mathbb{N}}^{\infty} \mathbf{G}_n^{p,q}(\underbrace{j\omega, \ldots, j\omega}_{p-times}, \underbrace{-j\omega, ..., -j\omega}_{q-times})a^{p+q}e^{j\omega(p-q)t} \tag{21}$$

The symbol $\mathbf{G}_n^{p,q}$[6] is the n^{th} input to state frequency response containing p-times the frequency ω and q-times the frequency $-\omega$. By substituting in Eq. 19 and collecting the terms of the same exponential (as the $e^{j\omega_m t}$), we can derive the input to state frequency responses $\mathbf{G}$ for every n. By denoting the resolvent $\mathbf{\Phi}(j\omega) = (j\omega\mathbf{E} - \mathbf{A})^{-1} \in \mathbb{C}^{n\times n}$, we can explicitly derive the following input to state transfer functions $\mathbf{G}_n$:

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

for $n \geq 1$ and $p + q = n$. By multiplying with the output matrix $\mathbf{C}$, we can further derive the input-output generalized frequency responses GFRFs (generalized transfer functions) as:

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

At this point, we are capable to 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 involves an n-dimension frequency space. For that reason, next, we derive the general second symmetric kernel for the bilinear case with a 2-tone input.
A double oscillatory input as:

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

where $\alpha_1 = \frac{A_1}{2}$ and $\alpha_2 = \frac{A_2}{2}$. In that case, with the growing exponential approach, we can write the input to state solution in steady state as:

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

[6] $\mathbf{G}_n^{p,q} = \mathbf{G}(\underbrace{j\omega, ..., j\omega}_{p-times}; \underbrace{-j\omega, ..., -j\omega}_{q-times})$

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. 19 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 to:

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

where by using the resolvent notation and multiplying with $\mathbf{C}$, we derive the *2-order generalized frequency response (symmetric kernel)*

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

4.2 The kernel separation method

One way to measure Volterra kernels is by means of interpolation. This problem is equivalent to that of estimating the coefficients of a noisy polynomial. This interpolation scheme builds a linear system with a Vandermonde matrix. This matrix is invertible since the amplitudes are distinct and non-zero. The inverse of a Vandermonde matrix can be explicitly computed and there are stable ways to solve these equations. The nth harmonic in the frequency domain is derived by applying a (single-sided) Fourier transform. More precisely, the explicit formulation is as follows:

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

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

$$\begin{bmatrix} Y_\alpha(0j\omega) \\ Y_\alpha(1j\omega) \\ Y_\alpha(2j\omega) \\ \vdots \\ Y_\alpha(nj\omega) \end{bmatrix} = \begin{bmatrix} \alpha^0 & \alpha^2 & \alpha^4 & \dots \\ \alpha^1 & \alpha^3 & \alpha^5 & \dots \\ \alpha^2 & \alpha^4 & \alpha^6 & \dots \\ \vdots & \vdots & \vdots & \vdots \\ \alpha^n & \alpha^{n+2} & \alpha^{n+4} & \dots \end{bmatrix} \bigcirc \begin{bmatrix} H_0^{0,0} & H_2^{1,1} & H_4^{2,2} & \dots \\ H_1^{1,0} & H_3^{2,1} & H_5^{3,2} & \dots \\ H_2^{2,0} & H_4^{3,1} & H_6^{4,2} & \dots \\ \vdots & \vdots & \vdots & \vdots \\ H_n^{n,0} & H_{n+2}^{n+1,1} & H_{n+4}^{n+2,2} & \dots \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

By introducing the Hadamard product[7] notation and simplify the δ's contribution to ones, we can compactly rewrite the above system in the following form:

$$\mathbf{Y}_{(\alpha,\omega)} = [\mathbf{M}_\alpha \circ \mathbf{P}_\omega] \cdot \mathbf{e} \tag{29}$$

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. 29 reveals the separation of the aforementioned quantities. As we neglect higher order Volterra kernels, the measurement set tends to be more and more corrupted by noise.

Remark 4.1. Kernel separation and stage ℓ- approximation For a given system, the procedure consists in exciting it with an 1-tone input. By varying the driving frequency as well as, the amplitude, we can approximate GFRFs by minimizing the (2-norm) of the remaining over-determined systems.

$$\mathbf{Y}_{n+1,\ell}(jn\omega, \alpha_\ell) = \left[\mathbf{M}_{n+1,\ell}(\alpha_\ell) \circ \mathbf{P}_{n+1,\ell}(jn\omega)\right] \cdot \mathbf{e_{n+1,1}} \tag{30}$$

[7]Hadamard product denotes as "$\circ$" and means that the multiplication is element wise.

where the n-"direction" gives us the threshold up to the specific harmonic that we would like to measure. The ℓ-"direction" give us the level of the kernel separation we want to achieve. For instance, for the 2nd stage approximation $\ell = 2$ with $\alpha^m \approx 0,\ \forall m > \ell = 2$, and that results to $n = 2$ (up to 2nd harmonic).

4.3 Identification of the matrix N

The only difference between linear and bilinear models is the presence of the product between the input and the state that is scalled by the matrix $\mathbf{N}$. As the (LF) is able to identify the linear part ($\mathbf{A}, \mathbf{E}, \mathbf{B}, \mathbf{C}$) 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:

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

$$H_2(j\omega_1, -j\omega_1) = \frac{1}{2}\mathbf{C}\,(-\mathbf{A})^{-1}\,\mathbf{N}\left((j\omega_1\mathbf{I}-\mathbf{A})^{-1}\mathbf{B} + (-j\omega_1\mathbf{I}-\mathbf{A})^{-1}\mathbf{B}\right). \tag{31}$$

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

$$H_2(j\omega_1, j\omega_1) = \mathbf{C}\,(2j\omega_1\mathbf{I}-\mathbf{A})^{-1}\,\mathbf{N}(j\omega_1\mathbf{I}-\mathbf{A})^{-1}\mathbf{B}. \tag{32}$$

- While with a 2-tone input the general H_2 can be written:

$$H_2(j\omega_1, j\omega_2) = \frac{1}{2}\mathbf{C}\,(j(\omega_1+j\omega_2)\mathbf{I}-\mathbf{A})^{-1}\,\mathbf{N}\left((j\omega_1\mathbf{I}-\mathbf{A})^{-1}\mathbf{B} + (j\omega_2\mathbf{I}-\mathbf{A})^{-1}\mathbf{B}\right). \tag{33}$$

We introduce the following notations:

$$\begin{aligned} O(j\omega_1, j\omega_2) &= \frac{1}{2}\mathbf{C}\,(j(\omega_1+j\omega_2)\mathbf{I}-\mathbf{A})^{-1} \in \mathbb{C}^{1\times n}, \\ \mathcal{R}(j\omega_1, j\omega_2) &= \left((j\omega_1\mathbf{I}-\mathbf{A})^{-1}\mathbf{B} + (j\omega_2\mathbf{I}-\mathbf{A})^{-1}\mathbf{B}\right) \in \mathbb{C}^{n\times 1}. \end{aligned} \tag{34}$$

Note that, Eq. 33 be compactly rewritten as:

$$H_2(j\omega_1, j\omega_2) = O(j\omega_1, j\omega_2)\mathbf{N}\mathcal{R}(j\omega_1, j\omega_2). \tag{35}$$

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:

$$\begin{aligned} \underbrace{H_2(j\omega_1, j\omega_2)}_{\mathbf{Y}_{k,1}} &= \underbrace{O(j\omega_1, j\omega_2)}_{O_{k,n}}\underbrace{\mathbf{N}}_{n\times n}\underbrace{\mathcal{R}(j\omega_1, j\omega_2)}_{\mathcal{R}_{n,k}}, \\ \mathbf{Y}_{(1:k,1)} &= \underbrace{\left(O_{(1:k,1:n)} \otimes \mathcal{R}^T_{(1:k,1:n)}\right)}_{(1:k,1:n^2)}\underbrace{vec\,(\mathbf{N})}_{(1:n^2,1)} \end{aligned} \tag{36}$$

Note that Eqs. (31, 32, 33) can be equivalently rewritten as one linear matrix equation, i.e., the one given in Eq. 36. By filling out the matrix $[O \otimes \mathcal{R}]$ with the information from $H_2(j\omega_1, -j\omega_1)$ and from $H_2(j\omega_1, j\omega_1)$, has full rank. Hence, we are able to solve equation (36) and then 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 cases:

1. $k < n^2$ underdetermined $\rightarrow$ LS solution (minimizing the 2-norm) as in [23],
2. $k \geq n^2$ overdetermined $\rightarrow$ identification of $\mathbf{N}$,

Lemma 4.1. For $k \geq n^2$ measurements with $\omega_{1,k}, \omega_{2,k} \in \mathbb{C}$ distinct pairs, it holds:

$$rank\left(O_{(1:k,1:n)} \otimes \mathcal{R}_{(1:k,1:n)}\right) = n^2 \tag{37}$$

As the above lemma indicates, we need at least n^2 measurements to identify the matrix $\mathbf{N}$ from a bilinear system of dimension n.

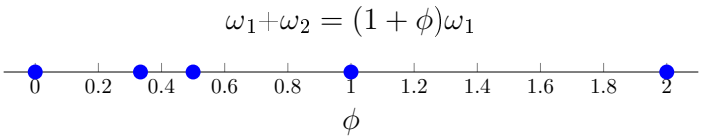

Figure 4: This figure shows the constrains of ϕ (e.g., $\phi = 0, 1/3, 1/2, 1, 2, 3, \ldots$, etc.). By choosing ϕ's within the blue dots, we construct frequency bandwidths with an unique $(\omega_1 + \omega_2)$.

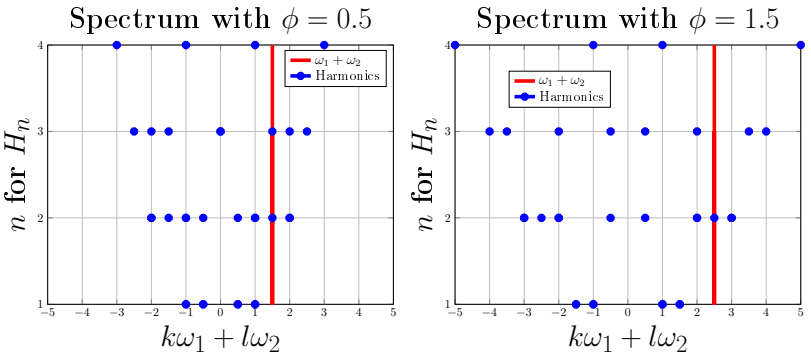

Figure 5: Left pane: Overlapping harmonics with invalid $\phi = 0.5$. Right pane: Uniquely defined harmonic at $(\omega_1 + \omega_2)$ with valid $\phi = 1.5$. Here, it holds $(n = k + l)$.

4.4 A harmonic separation strategy for the 2nd kernel

To identify the n^{th} Volterra kernel, we need an n-tone input signal. As we want to identify the 2nd kernel, the input signal needs to be chosen as a double sinusoidal 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 $e^{(\pm kj\omega_1 \pm lj\omega_2)t}$, where $k, l \in \mathbb{N}$. The aim is to differentiate the $(\omega_1+\omega_2)$ harmonic from the others harmonics. By succeeding that, there will be no overlapping (commensurate frequencies) and we know what exactly we measure in the spectrum. More precisely, we want the following result to hold:

$$\omega_1 + \omega_2 \neq k\omega_1 + l\omega_2, \ \forall (k, l) \in \mathbb{Z} \times \mathbb{Z} \setminus \{1, 1\}. \tag{38}$$

Suppose $\omega_2 = \phi\omega_1, \ \phi \in \mathbb{R}$. What are the suitable ϕ's where Eq. 38 holds?

$$\begin{aligned} \omega_1 + \phi\omega_1 &= k\omega_1 + l\phi\omega_1 \Rightarrow \\ 1 + \phi &= k + l\phi \Rightarrow \\ \phi &= \frac{k-1}{1-l}, \ k, l \in \mathbb{Z} \setminus \{1\}. \end{aligned} \tag{39}$$

By choosing ϕ that the equality in Eq. 39 doesn't hold, with harmonic mixing index as $m = k + l$ (i.e., for 2nd stage approximation $m = 2$), it makes the harmonic $(\omega_1 + \omega_2)$ uniquely defined in the frequency spectrum up to the m^{th} harmonic.

To visualize this feature, we choose $\omega_1 = 1$, and $\omega_2 = \omega_1\phi = \phi$, for harmonic mixing index $m = 2$. Then, the constrains of ϕ in Fig. 4 are depicted with blue dots.

Next, in Fig. 5 and on the left pane, one ϕ constraint that occurs commensurate harmonics is depicted, and on the right pane, one unique harmonic construction at $(\omega_1 + \omega_2)$ is presented.

The next theoretical 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 :

$$\omega_2 \in (\phi_{i-1}\omega_1, \phi_i\omega_1), \ i = 1, \ldots \tag{40}$$

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

Remark 4.2. With the framework that we have developed, we force the separation of $\omega_1 + \omega_2$ harmonic only under a specific mixing order m. We do not offer any general solution of the harmonic separation problem for multi-tone input, although techniques have been introduced such as in [13]. There, 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

We start with a set of 1-tone inputs as $u(t) = \alpha_1 \cos(\omega_i t)$, $i = 1, ..., n$, with $\alpha_1 < 1$. For those n measurements we can estimate the linear kernel $H_1(j\omega_i)$, the $H_2(j\omega_i, j\omega_i)$ and the $H_2(j\omega_i, -j\omega_i)$ 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 in Fig. 3.

We can increase the accuracy of the method to ℓ-stage by varying the amplitude α_ℓ as explained in section 4.2.

Since the (LF) reveals the underlying order of the system denoted with r, the value of n should be at least equal to $2r$. Then, we can take the decision by analysing 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. 36 with measurements from the diagonal of the 2nd kernel and from the the perpendicular to the diagonal axis $(\omega_1, -\omega_1)$. Those measurements contribute to the LS 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 off the diagonal with a 2-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 harmonic separation problem for the frequency (ω_1, ω_2) appears now. To deal with this problem, we follow the solution proposed in section 4.4 (up to a mixing degree). Last, we solve the LS problem described in Eq. 36 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.

Program 4.1. (Algorithm 2) A Loewner-Volterra algorithm for bilinear identification/reduction from time domain data.

Input: 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 reduced bilinear system of dimension-r: $\Sigma_r : (\mathbf{E}_r, \mathbf{A}_r, \mathbf{B}_r, \mathbf{C}_r, \mathbf{N}_r)$

1. Apply one-tone input $u(t)$ with $\beta_\ell = 0$, $\omega_{1,k}$ for $k = 1, \ldots, n$ and collect the snapshots $y(t)$ in steady state. Improve the measurements by solving system Eq. 29 by sweeping the amplitude for each frequency.

2. Apply Fourier transform for each frequency and collect the following k measurements:
   - DC term: $Y_O(0 \cdot j\omega_{1,k})$,
   - 1st harmonic: $Y_I(1 \cdot j\omega_{1,k})$,
   - 2nd harmonic: $Y_{II}(2 \cdot j\omega_{1,k})$.

3. Apply the linear the (LF), see Algorithm 1 in [24] by using the measurements (e.g., $H_1(j\omega_{1,k}) = 2Y_I(j\omega_{1,k})/\alpha_{1,\ell}$ for the 2nd stage approximation) and get the the order r reduced linear model. If the second harmonic and higher harmonics are equal with zero, the underlying system is linear so $\mathbf{N}$ is the zero matrix $\mathbf{0}_{\mathbf{r}\times\mathbf{r}}$ as well. Otherwise the system is nonlinear and computing the bilinear matrix $\mathbf{N}$ will improve the accuracy.

4. Apply the 2-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 section 4.4 and get the estimations for the 2nd kernel as: $H_2(j\omega_{1,k}, j\omega_{2,k}) = 2Y_{II}(j\omega_{1,k}, j\omega_{2,k})/(\alpha_\ell \beta_\ell)$.

5. Solve the full rank least square problem described in Eq. 36 and compute the real-valued bilinear matrix **N**.

5 Applications

Example 5.1 (A small bilinear identification)**.** The aim in this example is to identify the bilinear model from time domain measurements. Consider the following controllable/observable bilinear model Eq. 19 of dimension-2 with a *non-symmetric* matrix **N**, zero initial condition and matrices as:

$$\mathbf{E} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \mathbf{A} = \begin{bmatrix} -1 & -10 \\ 10 & -1 \end{bmatrix}, \mathbf{N} = \begin{bmatrix} 1 & -2 \\ 3 & -4 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{C} = \begin{bmatrix} 1 & 1 \end{bmatrix}. \tag{41}$$

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

Table 1: **Measurements of the first (linear) kernel**

Frequency ω(Hz)	$\hat{H}_1(j\omega)$**-2nd stage**	$H_1(j\omega)$**-theoretical**
0.5	0.019928 + 0.01223i	0.020033 + 0.012241i
1	0.020634 + 0.024733i	0.020742 + 0.024755i
1.5	0.021871 + 0.037798i	0.021988 + 0.037836i
2	0.023741 + 0.05176i	0.023874 + 0.051821i

[a] With 2nd-stage approximation $\tilde{H}_1(j\omega) = 2Y_1(j\omega)/\alpha$.

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 $(\tilde{\mathbf{E}}, \tilde{\mathbf{A}}, \tilde{\mathbf{B}}, \tilde{\mathbf{C}})$. We stopped at the 4th 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 reviling decay of the singular values as in the next Fig. 6. The singular values 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.

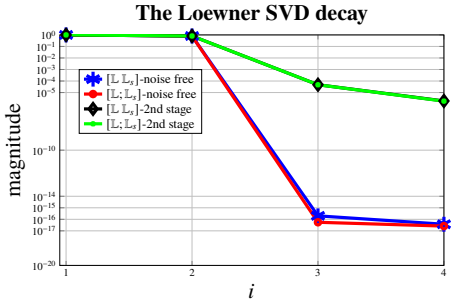

Figure 6: The singular value decay of the (LF) as a fundamental characterization of the McMillan degree of the underlying linear system. Here, a truncation scheme of order $r = 2$ is recommended where the 2nd stage approximation gave $\sigma_3/\sigma_1 = 4.721 \cdot 10^{-5}$, while for the noise free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

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

$$\tilde{\mathbf{A}}_t = \begin{bmatrix} -1.4513 & -8.8181 \\ 11.363 & -0.54868 \end{bmatrix}, \tilde{\mathbf{B}}_t = \begin{bmatrix} -0.92979 \\ 1.3967 \end{bmatrix}, \tilde{\mathbf{C}}_t = \begin{bmatrix} -0.76857 & 0.9203 \end{bmatrix}, \tag{42}$$

while by using the measured data with 2nd-stage approximation results to the following:

$$\tilde{\mathbf{A}} = \begin{bmatrix} -1.4581 & -8.8137 \\ 11.367 & -0.55163 \end{bmatrix}, \tilde{\mathbf{B}} = \begin{bmatrix} -0.93423 \\ 1.4 \end{bmatrix}, \tilde{\mathbf{C}} = \begin{bmatrix} -0.76751 & 0.9161 \end{bmatrix}. \quad (43)$$

Remark 5.1. 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: $\tilde{p}_t = -1 \pm 10\text{i}$ and the zero is: $\tilde{z}_t = -1$ while for the 2nd stage approximation to the linear system, the corresponding results are: $\tilde{p} = -1.0048 \pm 9.9989\text{i}$, $\tilde{z} = -1.0042$.

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. The following Table 2, contains measurements of the 2nd kernel with 1-tone input.

Table 2: **Measurements of the H_2 on the diagonal and perpendicular to the diagonal.**

Frequency 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	+0.15580 − 0.07747i	+0.15588 − 0.07729i	+0.10645	+0.10644
1	−0.26539 + 0.04057i	−0.26564 + 0.04093i	+0.15537	+0.15537
1.5	−0.03805 + 0.19171i	−0.03844 + 0.19233i	+0.30293	+0.30446
2	+0.03684 − 0.00528i	+0.036891 − 0.0054i	-0.12612	-0.12651

[b] The estimation of the second kernel with 2nd stage approximation on the diagonal as $\tilde{H}_2(j\omega, j\omega) = 2Y_2(j\omega, j\omega)/\alpha^2$ and perpendicular of the diagonal as $\tilde{H}_2(j\omega, -j\omega) = 2Y_2(j\omega, -j\omega)/\alpha^2$.

We can get a matrix $\mathbf{N}$ by solving the least square problem by just minimizing the 2-norm as in [23]. This result was not towards the identification of the matrix $\mathbf{N}$ and here is the new approach working towards the identification of bilinear systems.

Question 5.1. Can we identify the matrix $\mathbf{N}$? The improvement here relies on the rank deficiency problem is produced by solving the least square problem without taking under consideration measurements off the diagonal of the 2nd kernel H_2 along (ω_1, ω_2) with $(\omega_1 \neq \omega_2)$. By filling the LS problem Eq. 36 with a full rank, the answer is affirmative.

Back to our introductory example, the rank of the least square problem is less than $r^2 = 4$. So, we need to increase the rank. We take measurements (≤ 4) off the diagonal from the 2nd kernel by using the 2-tone input. The following Tab. 3 includes the theoretical and measured results.

Table 3: **Measurements of the second kernel (off 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, 0.3)	0.10735 − 0.021434i	0.10734 − 0.021388i
(0.2, 0.6)	0.13170 − 0.046818i	0.13171 − 0.046718i
(0.4, 0.3)	0.12011 − 0.035526i	0.12010 − 0.035449i
(0.4, 0.6)	0.15665 − 0.078110i	0.15671 − 0.077938i

[c] The estimation of the second kernel as $\tilde{H}_2(j\omega_1, j\omega_2) = Y_2(j\omega_1, j\omega_2)/(\alpha_1\alpha_2)$.
Here we use $\phi = 1.5$, to avoid the harmonic overlapping as explained in section 4.4.

The full rank LS solution gave for the theoretical noise free case and for the 2nd stage approximation the following results respectively:

$$\tilde{\mathbf{N}}_t = \begin{bmatrix} -4.6753 & -2.2605 \\ -1.5564 & -0.32474 \end{bmatrix}, \tilde{\mathbf{N}} = \begin{bmatrix} -4.6691 & -2.2667 \\ -1.5656 & -0.33364 \end{bmatrix} \quad (44)$$

Remark 5.2. Coordinate transformation By transforming all the matrices to the same coordinate system as in [21], we conclude:

- **Noise free case - exact identification**

$$\breve{\mathbf{A}}_t = \begin{bmatrix} -1.0 & -10.0 \\ 10.0 & -1.0 \end{bmatrix}, \breve{\mathbf{N}}_t = \begin{bmatrix} 1.0 & -2.0 \\ 3.0 & -4.0 \end{bmatrix}, \breve{\mathbf{B}}_t = \begin{bmatrix} 1.0 \\ 1.0 \end{bmatrix}, \breve{\mathbf{C}}_t = \begin{bmatrix} 1.0 \\ 1.0 \end{bmatrix}^T. \quad (45)$$

- **Simulated case - approximated identification**

$$\breve{\mathbf{A}} = \begin{bmatrix} -1.0038 & -9.9941 \\ 10.004 & -1.0059 \end{bmatrix}, \breve{\mathbf{N}} = \begin{bmatrix} -1.0055 & 2.0008 \\ 3.0038 & -3.9972 \end{bmatrix}, \breve{\mathbf{B}} = \begin{bmatrix} 0.99925 \\ 1.0004 \end{bmatrix}, \breve{\mathbf{C}} = \begin{bmatrix} 1.0 \\ 1.0 \end{bmatrix}^T. \quad (46)$$

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

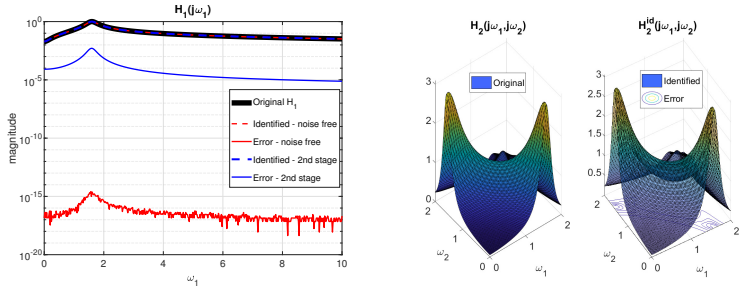

Figure 7: The identified first and second kernel with 2nd-stage approximation in comparison with the theoretical kernels.

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

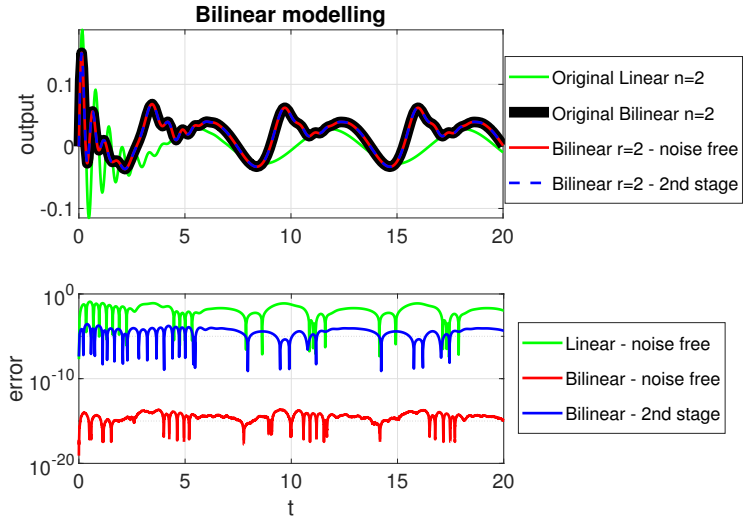

Figure 8: The evaluation of the models with order $r = 2$ performed with input as $u(t) = \cos(t)$, $t \in [0, 20]$. The noise free case has reached machine precision.

Example 5.2. Burger's Equation. This example illustrating the bilinear modeling and reduction concepts proposed in [5] for the viscous Burger's equation from time domain simulations. We kept the same set up, as in the introductory example and we present the corresponding results with initial system dimension $n = 110$ reduced by the proposed method to order $r = 2$ with the first neglected singular value to be $\sigma_3/\sigma_1 = 2.576 \cdot 10^{-5}$. In Fig. 9, evaluation results are presented. Last, in Fig. 10,

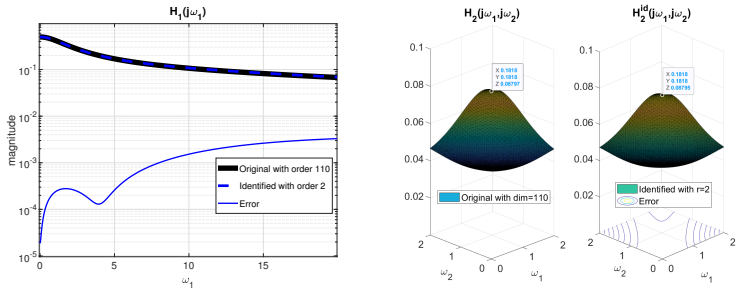

Figure 9: The first and the second kernel evaluations in comparison with the originals.

a simulation reveals that the proposed method can improve the accuracy by fitting a non-linear model.

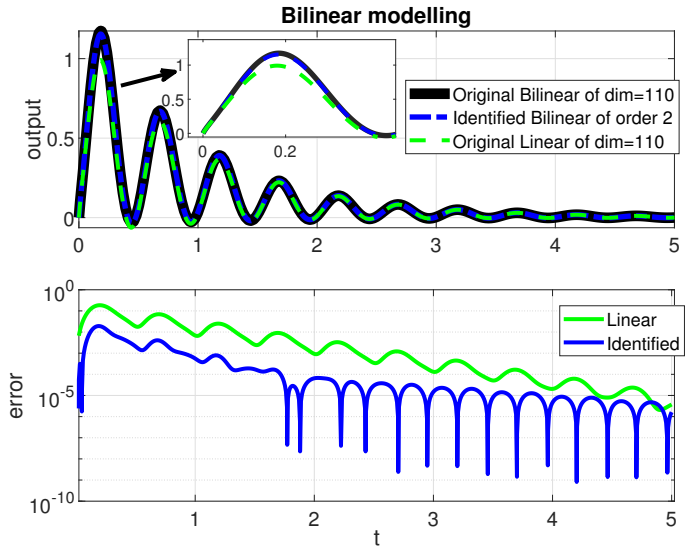

Figure 10: Time domain simulation for the original Burger's example with viscosity parameter as 1 and with dimension 110. A comparison among the identified/reduced bilinear of order $r = 2$ with the linear is depicted. The input is $u(t) = (1 + 2\sin(4\pi t))e^{-t}$ for $t \in [0, 5]$ and the maximum error for the reduced bilinear is of the order $O(10^{-2})$.

6 Conclusion

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

Table 4: **Summary of the results from the two examples.**

Error quantification	**Example 1**	**Example 2 (Burger's)**
$\|H_1 - \tilde{H}_1\|_{\max}$	$5.077 \cdot 10^{-3}$	$3.301 \cdot 10^{-3}$
$\|y(t) - y_l(t)\|_{\max}$	$1.213 \cdot 10^{-1}$	$1.856 \cdot 10^{-1}$
$\|H_2 - \tilde{H}_2\|_{\max}$	$2.794 \cdot 10^{-2}$	$8.089 \cdot 10^{-4}$
$\|y(t) - \tilde{y}_b(t)\|_{\max}$	$\mathbf{2.739 \cdot 10^{-4}}$	$\mathbf{1.917 \cdot 10^{-2}}$

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

We have shown that for the noise free case, the proposed method offers 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).

The variational approach is a theoretical method to identify regular kernels which are appropriate for system identification purposes [30]. However, these kernels do not have a physical meaning, i.e., can not 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.

Extension to the MIMO case and for other non-linearity structures (s.a., quadratic, bilinear-quadratic etc.) is under future research.

References

[1] M. I. Ahmad, U. Baur, and P. Benner. Implicit Volterra series interpolation for model reduction of bilinear systems. *J. Comput. Appl. Math.*, 316:15–28, 2017.

[2] B. D. O. Anderson and A. C. Antoulas. Rational interpolation and state-variable realizations. *Linear Algebra Appl.*, 137/138:479–509, 1990.

[3] A. C. Antoulas. *Approximation of Large-Scale Dynamical Systems*, volume 6 of *Adv. Des. Control*. SIAM Publications, Philadelphia, PA, 2005.

[4] A. C. Antoulas, C. A. Beattie, and S. GÃijÄ§ercin. *Interpolatory Methods for Model Reduction*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 2020.

[5] A. C. Antoulas, I. V. Gosea, and A. C. Ionita. Model reduction of bilinear systems in the loewner framework. *SIAM Journal on Scientific Computing*, 38(5):B889–B916, 2016.

[6] Athanasios C. Antoulas, Sanda Lefteriu, and A. Cosmin Ionita. *Chapter 8: A Tutorial Introduction to the Loewner Framework for Model Reduction*, pages 335–376.

[7] Z. Bai and D. Skoogh. A projection method for model reduction of bilinear dynamical systems. *Linear Algebra Appl.*, 415(2–3):406–425, 2006.

[8] U. Baur, P. Benner, and L. Feng. Model order reduction for linear and nonlinear systems: A system-theoretic perspective. 21(4):331–358, 2014.

[9] P. Benner and T. Breiten. Interpolation-based $\mathcal{H}_2$-model reduction of bilinear control systems. *SIAM J. Matrix Anal. Appl.*, 33(3):859–885, 2012.

[10] P. Benner, T. Breiten, and T. Damm. Generalized tangential interpolation for model reduction of discrete-time MIMO bilinear systems. *Internat. J. Control*, 84(8):1398–1407, 2011.

[11] P. Benner and T. Damm. Lyapunov equations, energy functionals, and model order reduction of bilinear and stochastic systems. *SIAM J. Control Optim.*, 49(2):686–711, 2011.

[12] P. Benner, S. Gugercin, and K. Willcox. A survey of model reduction methods for parametric systems. *SIAM Review*, 57(4):483–531, 2015.

[13] S. Boyd, Yaw shing Tang, and Leon O. Chua. Measuring volterra kernels. 1983.

[14] T. Breiten. *Interpolatory Methods for Model Reduction of Large-Scale Dynamical Systems*. Dissertation, Department of Mathematics, Otto-von-Guericke University, Magdeburg, Germany, 2013.

[15] Zlatko DrmaĂŊ and Benjamin Peherstorfer. Learning low-dimensional dynamical-system models from noisy frequency-response data with loewner rational interpolation, 2019.

[16] G. M. Flagg and S. Gugercin. Multipoint Volterra series interpolation and $\mathcal{H}_2$ optimal model reduction of bilinear systems. *SIAM J. Numer. Anal.*, 36(2):549–579, 2015.

[17] Elliot Fosong, Philipp Schulze, and Benjamin Unger. From time-domain data to low-dimensional structured models, 2019.

[18] Ion Victor Gosea and Athanasios C. Antoulas. Data-driven model order reduction of quadratic-bilinear systems. *Numerical Linear Algebra with Applications*, 25(6):e2200, 2018. e2200 nla.2200.

[19] B. Gustavsen and A. Semlyen. Rational approximation of frequency domain responses by vector fitting. *IEEE Transactions on Power Delivery*, 14(3):1052–1061, July 1999.

[20] A. Cosmin Ionita. Matrix pencils in time and frequency domain system identification. pages 79–88, 2012.

[21] Jer-Nan Juang. Continuous-time bilinear system identification. *Nonlinear Dynamics*, 39(1):79âĂŞ94, 2005.

[22] E. Kaiser, J. N. Kutz, and S. L. Brunton. Sparse identification of nonlinear dynamics for model predictive control in the low-data limit. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 474(2219):20180335, 2018.

[23] D. S. Karachalios, I. V. Gosea, and A. C. Antoulas. A bilinear identification-modeling framework from time domain data. *Proc. Appl. Math. Mech.*, 19(1):e201900246, 2019.

[24] D. S. Karachalios, I. V. Gosea, and A. C. Antoulas. The loewner framework for system identification and reduction. *In P. Benner, S. Grivet-Talocia, A. Quarteroni, G. Rozza, W. H. A. Schilders, and L. M. Silveira, editors, Handbook on Model Reduction, volume I of Methods and Algorithms*, in press.

[25] Sanda Lefteriu, Antonio C. Ionita, and Athanasios C. Antoulas. *Modeling Systems Based on Noisy Frequency and Time Domain Measurements*, pages 365–378. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.

[26] Benjamin. Peherstorfer, Serkan. Gugercin, and Karen. Willcox. Data-driven reduced model construction with time-domain loewner models. *SIAM Journal on Scientific Computing*, 39(5):A2152–A2178, 2017.

[27] Benjamin Peherstorfer and Karen Willcox. Data-driven operator inference for nonintrusive projection-based model reduction. *Computer Methods in Applied Mechanics and Engineering*, 306:196 – 215, 2016.

[28] Menka Petkovska, Daliborka NikoliĂĞ, and Andreas Seidel-Morgenstern. Nonlinear frequency response method for evaluating forced periodic operations of chemical reactors. *Israel Journal of Chemistry*, 58(6-7):663–681, 2018.

[29] J. R. Phillips. Projection-based approaches for model reduction of weakly nonlinear, time-varying systems. 22(2):171–187, 2003.

[30] W. J. Rugh. *Nonlinear System Theory: The Volterra/Wiener Approach*. The Johns Hopkins University Press, Baltimore, 1981.

[31] Peter J. Schmid. Dynamic mode decomposition of numerical and experimental data. *Journal of Fluid Mechanics*, 656:5âĂŞ28, 2010.