MAX PLANCK INSTITUTE FOR DYNAMICS OF COMPLEX TECHNICAL SYSTEMS MAGDEBURG

DRI DATA-DRIVEN SYSTEM REDUCTION AND IDENTIFICATION

Algorithms for Dimension and Complexity Reduction Workshop, Providence, March 23-27, 2020

Learning block-oriented nonlinear models of dynamical systems from data

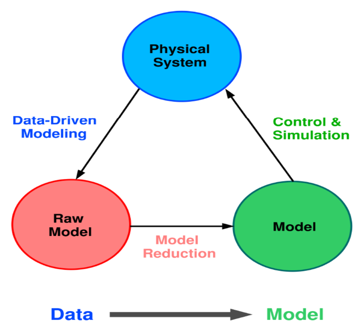
Ion Victor Gosea[1], Athanasios C. Antoulas[1,2,3]

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

Introduction, motivation and practical insight

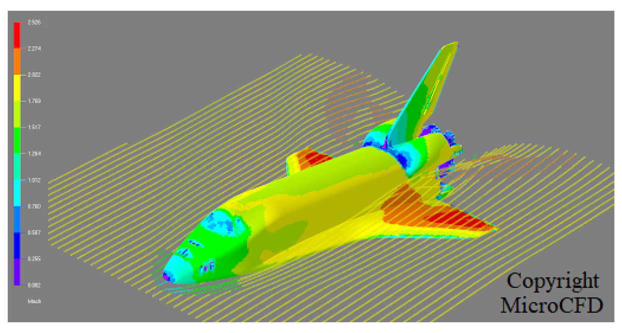
Approaches

- **Finite element/difference** schemes usually lead to large-scale high fidelity models and expensive computations in memory/time.
- Using **data-driven** methods, learn/reveal reduced-order models to be used as surrogates with cheap computation time/memory.



Type of data measurements

- **Time domain:** Easy to collect, commonly used in many fields, e.g., nonlinear PDEs, turbulent flow control, gas/energy networks.
- **Frequency domain:** Typically measured by DNS (direct numerical simulations) or with VNAs, e.g., the S(scattering) parameters.



Linear and nonlinear systems; block-oriented models

- Modeling nonlinear systems is challenging due to many different possible nonlinear structures, i.e., based on Volterra series, Wiener theory, NARMAX models, neural networks, etc.
- Different types of responses to an excitation, i.e., with subharmonics, bifurcation, chaos, etc.

Linear Systems

$$\begin{cases} \mathbf{E}\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}u(t), \\ y(t) = \mathbf{C}\mathbf{x}(t). \end{cases}$$

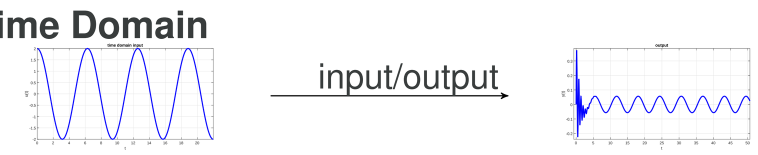
where $\mathbf{E}, \mathbf{A} \in \mathbb{R}^{n\times n}$, $\mathbf{B} \in \mathbb{R}^{n\times m}$, $\mathbf{C} \in \mathbb{R}^{p\times n}$.

Nonlinear Systems

$$\begin{cases} \mathbf{E}\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathcal{F}(\mathbf{x}(t), u(t)) + \mathbf{B}u(t), \\ y(t) = \mathbf{C}\mathbf{x}(t) + \mathcal{G}(\mathbf{x}(t), u(t)), \end{cases}$$

where $u(t) \rightsquigarrow$ **input** and $y(t) \rightsquigarrow$ **output**.

Time Domain (linear)



Time Domain (nonlinear)

Frequency Domain (linear)

Frequency Domain (nonlinear)

- Block-oriented models form a powerful and intuitive tool to handle nonlinear systems.
- Constructed from 2 blocks: a linear time-invariant (LTI) block and a static nonlinear block.

$u(t) \to$ [$f(x)$] $\to f[u(t)] \to$ [LTI] $\to y(t)$ — Hammerstein Model

$u(t) \to$ [LTI] $\to y(t) \to$ [$g(x)$] $\to g[y(t)]$ — Wiener Model

- Wiener models can approximate any nonlinear system with arbitrarily high accuracy [Boyd/Chua '85])
- Many different block combinations are possible: series, parallel or feedback connections.

$u(t) \to$ [$f(x)$] $\to f[u(t)] \to$ [LTI] $\to y(t) \to$ [$g(x)$] $\to g(y(t))$ — Hammerstein-Wiener Model

- Expand the static nonlinearities, i.e. the smooth nonlinear functions f, g, into Maclaurin series

$$f[u(t)] = \sum_{k=1}^{\infty} \frac{d^k f(u)}{du^k}\Big|_{u=0} \frac{u^k(t)}{k!}, \quad g[y(t)] = \sum_{k=1}^{\infty} \frac{d^k g(y)}{dy^k}\Big|_{y=0} \frac{y^k(t)}{k!}.$$

- Keep only the the first two terms, i.e. $f[u(t)] \approx \alpha_1 u(t) + \alpha_2 u^2(t)$ and $g[y(t)] \approx \beta_1 y(t) + \beta_2 y^2(t)$.

Wiener (Generalized) Model

$$\Sigma_W : \begin{cases} \mathbf{E}\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}u(t), \\ y(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{K}[\mathbf{x}(t) \otimes \mathbf{x}(t)]. \end{cases}$$

Hammerstein-Wiener (Generalized) Model

$$\Sigma_{HW} : \begin{cases} \mathbf{E}\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}u(t) + \mathbf{L}u^2(t), \\ y(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{K}[\mathbf{x}(t) \otimes \mathbf{x}(t)]. \end{cases}$$

In order that the Wiener/Hammerstein structure is conserved, $\exists a, b \in \mathbb{R}$, $\mathbf{K} = a(\mathbf{C} \otimes \mathbf{C})$ and $\mathbf{L} = b\mathbf{B}$.

Main contribution Given input-output data in the time domain (for purely oscillating control inputs), extract information from the spectrum and construct reduced-order models that explain the data.

The Loewner framework - linear systems

Aim: Construct reduced-order linear models directly from measurements - the Loewner framework.

- **Frequency domain:** linear [Mayo/Antoulas '07]) & nonlinear [Antoulas/G./Ionita '16]), [G. et al '18, '19]) .
- **Time domain:** linear [Peherstorfer/Gugercin/Willcox '17] & nonlinear [Peherstorfer/Gugercin '20] .

1. Given measurements $\{(\omega_k, f_k) : k = 1, \ldots, 2n\} \rightarrow$, divide the data into two disjoint sets:

$$\mathbf{S} = [\underbrace{\omega_1, \ldots, \omega_n}_{\mu}, \underbrace{\omega_{n+1}, \ldots, \omega_{2n}}_{\lambda}], \quad \mathbf{F} = [\underbrace{f_1(\omega_1), \ldots, f_n(\omega_n)}_{\mathbf{V}}, \underbrace{f_{n+1}(\omega_{n+1}), \ldots, f_{2n}(\omega_{2n})}_{\mathbf{W}}].$$

2. The associated **Loewner** & **shifted-Loewner** matrices $\mathbb{L}$ & $\mathbb{L}_s$ are introduced:

$$\mathbb{L}_{(i,j)} = \frac{\mathbf{v}_i - \mathbf{w}_j}{\mu_i - \lambda_j}, \; \mathbb{L}_{s(i,j)} = \frac{\mu_i \mathbf{v}_i - \lambda_j \mathbf{w}_j}{\mu_i - \lambda_j}, \; i, j = 1, \ldots, n.$$

- **Exact amount of data - regular Loewner pencil:**

$$H(s) = \mathbb{W}(\mathbb{L} - s\mathbb{L}_s)^{-1}\mathbb{V} = \begin{pmatrix} \mathbf{w}_{n+1} \\ \vdots \\ \mathbf{w}_{2n} \end{pmatrix}^* \left(\begin{bmatrix} \frac{\mathbf{v}_1 - \mathbf{w}_1}{\mu_1 - \lambda_1} & \cdots & \frac{\mathbf{v}_1 - \mathbf{w}_m}{\mu_1 - \lambda_m} \\ \vdots & \ddots & \vdots \\ \frac{\mathbf{v}_p - \mathbf{w}_1}{\mu_p - \lambda_1} & \cdots & \frac{\mathbf{v}_p - \mathbf{w}_m}{\mu_p - \lambda_m} \end{bmatrix} - s \begin{bmatrix} \frac{\mu_1\mathbf{v}_1 - \mathbf{w}_1\lambda_1}{\mu_1 - \lambda_1} & \cdots & \frac{\mu_1\mathbf{v}_1 - \mathbf{w}_m\lambda_m}{\mu_1 - \lambda_m} \\ \vdots & \ddots & \vdots \\ \frac{\mu_p\mathbf{v}_p - \mathbf{w}_1\lambda_1}{\mu_p - \lambda_1} & \cdots & \frac{\mu_p\mathbf{v}_p - \mathbf{w}_m\lambda_m}{\mu_p - \lambda_m} \end{bmatrix} \right)^{-1} \begin{pmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_n \end{pmatrix} = f(s).$$

- **Redundant data - singular Loewner pencil:**

Project using the singular vectors of the Loewner matrix, i.e., $[\mathbf{Y}, \mathbf{S}, \mathbf{X}] = \mathrm{svd}(\mathbb{L})$:

$$\underbrace{\{\mathbf{C}, \mathbf{E}, \mathbf{A}, \mathbf{B}\}}_{\text{model}} = \underbrace{\{\mathbb{W}, -\mathbb{L}, -\mathbb{L}_s, \mathbb{V}\}}_{\text{original}} \underbrace{\Rightarrow}_{\text{SVD}} \underbrace{\{\mathbb{W}\mathbf{X}, -\mathbf{Y}^*\mathbb{L}\mathbf{X}, -\mathbf{Y}^*\mathbb{L}_s\mathbf{X}, \mathbf{Y}^*\mathbb{V}\}}_{\text{reduced}} = \underbrace{\{\mathbf{C}_r, \mathbf{E}_r, \mathbf{A}_r, \mathbf{B}_r\}}_{\text{ROM}}.$$

Wiener models: transfer functions and lifting techniques

$u(t) \to$ [LTI] $\to$ [$g(x)$] $\to y(t) = g[x(t)]$

$u(t) = \sum_{\ell=1}^{L} e^{i\omega_\ell t}$ $\quad \mathbf{x}(t) = \sum_{\ell=1}^{L} \underbrace{(i\omega_\ell \mathbf{E} - \mathbf{A})^{-1}\mathbf{B}}_{\mathbf{G}_1(i\omega_\ell)} e^{i\omega_\ell t}$ $\quad g(\mathbf{x}) = \mathbf{C}\mathbf{x} + \mathbf{K}(\mathbf{x} \otimes \mathbf{x})$.

- The observed output of the generalized Wiener model can be split as follows

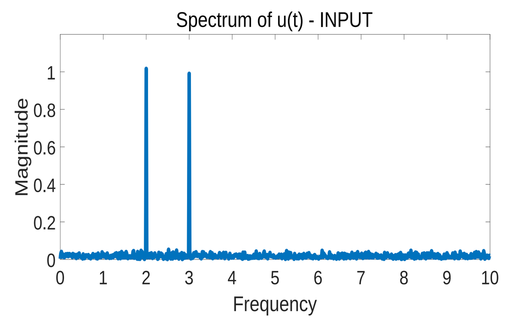
$$y(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{K}[\mathbf{x}(t) \otimes \mathbf{x}(t)] = \mathbf{C}\sum_{\ell=1}^{L} \mathbf{G}_1(i\omega_\ell)e^{i\omega_\ell t} + \mathbf{K}\left[\sum_{j=1}^{L} \mathbf{G}_1(i\omega_j)e^{i\omega_j t}\right] \otimes \left[\sum_{h=1}^{L} \mathbf{G}_1(i\omega_h)e^{i\omega_h t}\right]$$
$$= \sum_{\ell=1}^{L} H_1(i\omega_\ell)e^{i\omega_\ell t} + \sum_{j=1}^{L}\sum_{h=1}^{L} H_2(i\omega_j, i\omega_h)e^{i(\omega_j+\omega_h)t}.$$

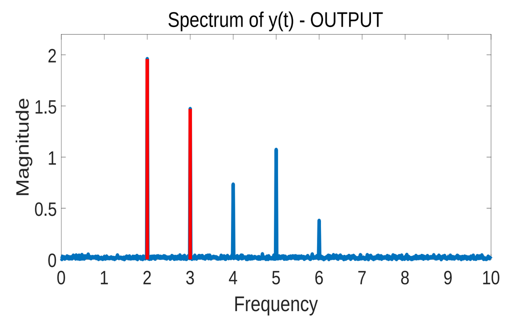
- The input-output behavior is characterized by two **transfer functions**

$$H_1(s_1) = \mathbf{C}(s_1\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}, \quad H_2(s_1, s_2) = \mathbf{K}\left[(s_1\mathbf{E} - \mathbf{A})^{-1}\mathbf{B} \otimes (s_2\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}\right].$$

Goal: Extend the classical Loewner framework by incorporating data as samples of both functions $H_1(s_1)$ and $H_2(s_1, s_2)$; measurements are acquired by simulating the system with oscillatory signals.

The **input** is $u(t) = e^{2it} + e^{3it}$;
The **output** is given by
$y(t) = H_1(2i)e^{2it} + H_1(3i)e^{3it}$
$+ H_2(2i, 2i)e^{4it} + H_2(3i, 3i)e^{6it}$
$+ 2H_2(2i, 3i)e^{5it}$.





Remark: By introducing new state variables (lifting), rewrite the Wiener model as bilinear or QB systems.

$$\bar{\mathbf{x}}(t) = \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{x}(t) \otimes \mathbf{x}(t) \end{bmatrix} \in \mathbb{R}^{n^2+n} \rightsquigarrow \quad \Sigma_W : \begin{cases} \dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}u(t), \\ y(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{K}[\mathbf{x}(t) \otimes \mathbf{x}(t)]. \end{cases} \underset{\text{Bilinear}}{\rightsquigarrow} \bar{\Sigma}_B : \begin{cases} \dot{\bar{\mathbf{x}}}(t) = \bar{\mathbf{A}}\bar{\mathbf{x}} + \bar{\mathbf{N}}\bar{\mathbf{x}}u(t) + \bar{\mathbf{B}}u(t), \\ y(t) = \bar{\mathbf{C}}\bar{\mathbf{x}}. \end{cases}$$

The system matrices of the **lifted bilinear** system $\bar{\Sigma}_B$ can be written as follows [Boyd/Chua '85]:

$$\bar{\mathbf{A}} = \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{A} \otimes \mathbf{I}_n + \mathbf{I}_n \otimes \mathbf{A} \end{bmatrix}, \quad \bar{\mathbf{N}} = \begin{bmatrix} \mathbf{0}_n & \mathbf{0}_{n,n^2} \\ \mathbf{B} \otimes \mathbf{I}_n + \mathbf{I}_n \otimes \mathbf{B} & \mathbf{0}_{n^2} \end{bmatrix}, \quad \bar{\mathbf{B}} = \begin{bmatrix} \mathbf{B} \\ \mathbf{0}_{n^2} \end{bmatrix}, \quad \bar{\mathbf{C}} = \begin{bmatrix} \mathbf{C} & \mathbf{K} \end{bmatrix}.$$

$$\tilde{\mathbf{x}}(t) = \begin{bmatrix} \mathbf{x}(t) \\ y(t) \end{bmatrix} \in \mathbb{R}^{n+1} \rightsquigarrow \quad \Sigma_W : \begin{cases} \dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}u(t), \\ y(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{K}[\mathbf{x}(t) \otimes \mathbf{x}(t)]. \end{cases} \underset{\text{QB}}{\rightsquigarrow} \tilde{\Sigma}_{QB} : \begin{cases} \dot{\tilde{\mathbf{x}}}(t) = \tilde{\mathbf{A}}\tilde{\mathbf{x}} + \tilde{\mathbf{Q}}\tilde{\mathbf{x}}(t) \otimes \tilde{\mathbf{x}}(t) + \tilde{\mathbf{N}}\tilde{\mathbf{x}}u(t) + \tilde{\mathbf{B}}u(t), \\ \tilde{y}(t) = \tilde{\mathbf{C}}\tilde{\mathbf{x}}. \end{cases}$$

The system matrices of the **lifted quadratic-bilinear** system $\tilde{\Sigma}_{QB}$ can be written as [Pulch/Narayan '19]

$$\tilde{\mathbf{A}} = \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{CA} & \mathbf{0} \end{bmatrix}, \quad \tilde{\mathbf{N}} = \begin{bmatrix} \mathbf{0}_n & \mathbf{0}_{n\times 1} \\ 2\mathbf{K}(\mathbf{B} \otimes \mathbf{I}_n) & \mathbf{0} \end{bmatrix}, \quad \tilde{\mathbf{Q}} = \begin{bmatrix} \mathbf{0}_{n^2+n,n^2} & \mathbf{0}_{n^2+n\times 1} \\ 2\mathbf{K}(\mathbf{A} \otimes \mathbf{I}_n) & \mathbf{0}_{1\times 2n+1} \end{bmatrix} \quad \tilde{\mathbf{B}} = \begin{bmatrix} \mathbf{B} \\ \mathbf{CB} \end{bmatrix}, \quad \tilde{\mathbf{C}} = \mathbf{e}_{n+1}.$$

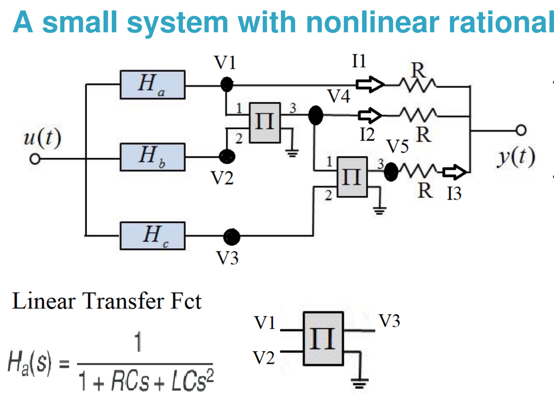
The second (symmetric) transfer function of a quadratic-bilinear system is given by:

$$H_2^{QB}(s_1, s_2) = \underbrace{\frac{1}{2}\mathbf{C}[(s_1 + s_2)\mathbf{I} - \mathbf{A}]^{-1}\mathbf{N}\left[(s_1\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} + (s_2\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}\right]}_{\text{bilinear part}} + \underbrace{\mathbf{C}[(s_1 + s_2)\mathbf{I} - \mathbf{A}]^{-1}\mathbf{Q}\left[(s_1\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} \otimes (s_2\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}\right]}_{\text{quadratic part}}$$

Goal: Approximate nonlinear dynamical systems (bilinear, QB etc.) with block-oriented models (H, W, HW).

Numerical Experiments

A small system with nonlinear rational output [Xiong/Jiang/Schutt-Ainé/Chew '17]



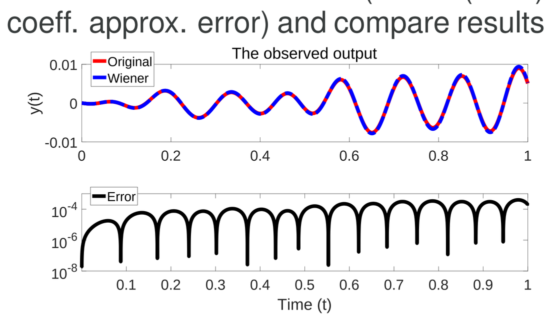
Time-domain multiplexer

$I_3(t) = [V_1(t)V_2(t) - V_3(t)]/Z_F$

The observed nonlinear output ($V_i = V$)

$$y(t) = \frac{4V(t) + 2V^2(t) + V^3(t)}{8 - V(t)} = \frac{1}{2}V(t) + \frac{5}{16}V^2(t) + \frac{21}{128}V^3(t) + \ldots$$

Fit a cubic Wiener model (with $\mathcal{O}(10^{-4})$ coeff. approx. error) and compare results



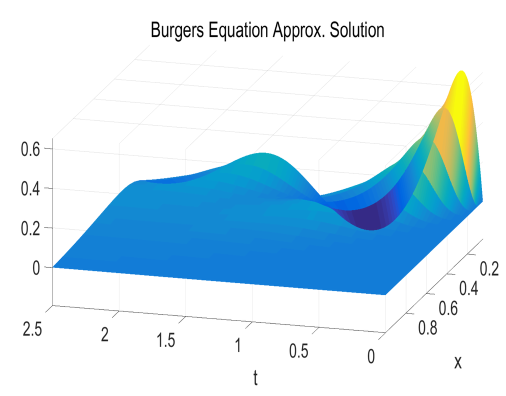
PDE from [Benner/Breiten '15]

Consider the viscous Burgers' equation

$$\frac{\partial v(x,t)}{\partial t} + v(x,t)\frac{\partial v(x,t)}{\partial x} = \frac{\partial}{\partial x}\left(\nu\frac{\partial v(x,t)}{\partial x}\right),$$

with boundary conditions $\forall t \geq 0$.

$$v(x, 0) = 0, \quad v(0, t) = u(t), \quad v(1, t) = 0.$$



Spatial discretization with a step size $h = 1/(n+1)$ produces a QB system

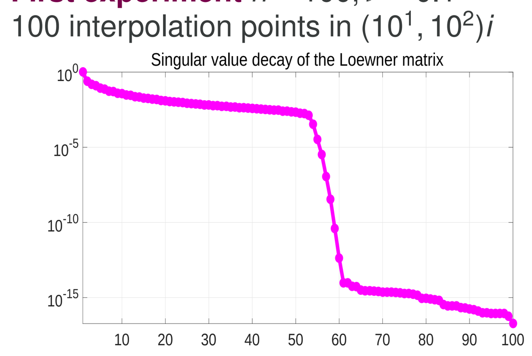
$$\dot{v}_k = \begin{cases} -\frac{v_1 v_2}{2h} + \frac{\nu}{h^2}(v_2 - 2v_1) + \left(\frac{v_1}{2h} + \frac{\nu}{h^2}\right)u, & k = 1, \\ -\frac{v_k}{2h}(v_{k+1} - v_{k-1}) + \frac{\nu}{h^2}(v_{k+1} - 2v_k + v_{k-1}), & 2 \leq k \leq n-1, \\ -\frac{v_n v_{n-1}}{2h} + \frac{\nu}{h^2}(-2v_n + 2v_{n-1}), & k = n. \end{cases}$$

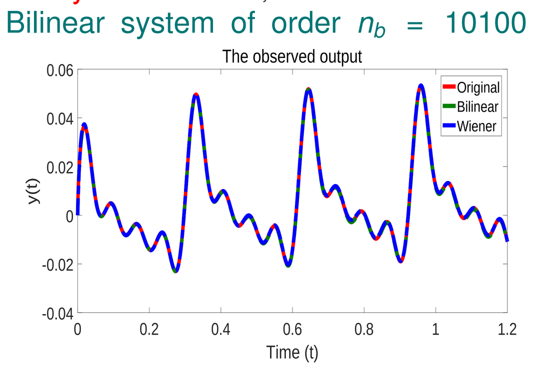
Rewrite as:
$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{Q}\mathbf{x}(t) \otimes \mathbf{x}(t) + \mathbf{N}\mathbf{x}(t)u(t) + \mathbf{B}u(t)$

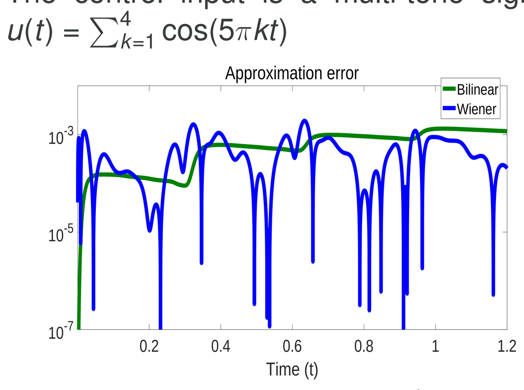
First experiment $n = 100$, $\nu = 0.1$
100 interpolation points in $(10^1, 10^2)i$

QB system $n = 100$, Wiener $r = 50$
Bilinear system of order $n_b = 10100$

The control input is a multi-tone signal $u(t) = \sum_{k=1}^{4} \cos(5\pi k t)$







Second experiment $n = 50$, $\nu = 0.02$
80 interpolation points in $(10^{-3}, 10^3)i$.

QB system $n = 50$, Wiener $r = 24$
Bilinear system of order $n_b = 2550$

The control input is $u(t) = e^{-t}\cos(4t)$.