Machine Learning Applied to the Blind Identification of Multiple Delays in Distributed Systems

Felipe Treviso, Riccardo Trinchero, Flavio G. Canavero

Dept. Electronics and Telecommunications, Politecnico di Torino, Italy

felipe.treviso@polito.it

https://emc.polito.it

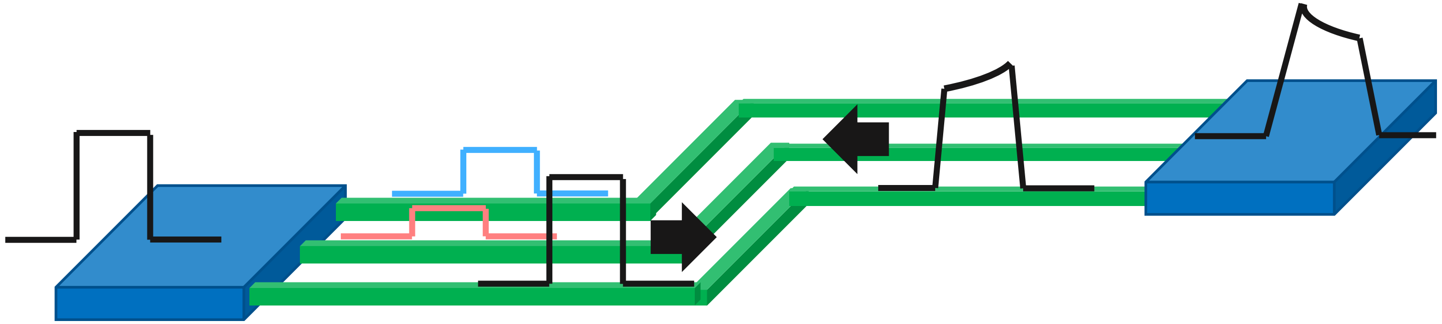
POLITECNICO
DI TORINO

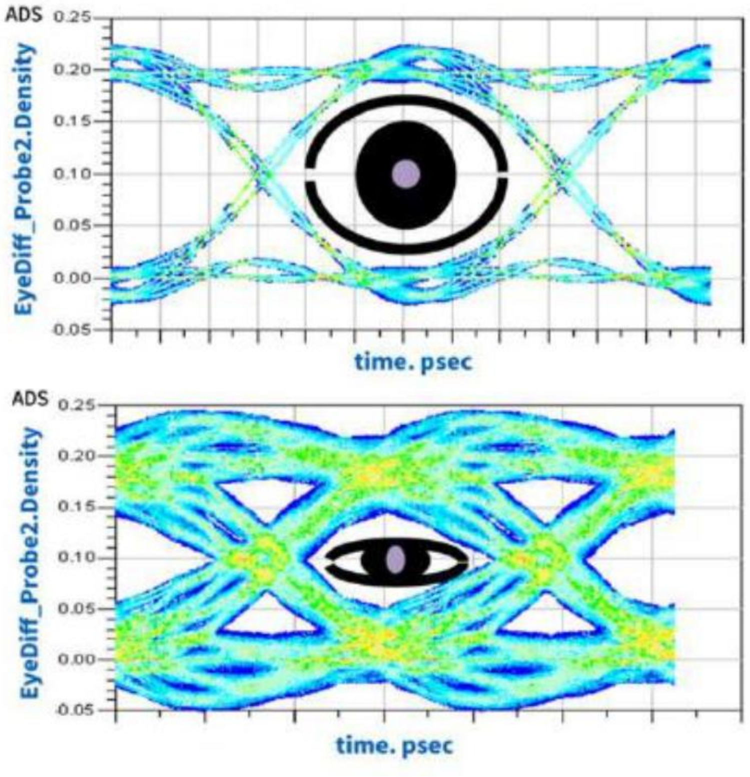
URSI GASS 2020

EMC Group

Signal Propagation Effects in Interconnects

- **Electrical interconnects** are responsible for a considerable part of **signal degradation** [1]



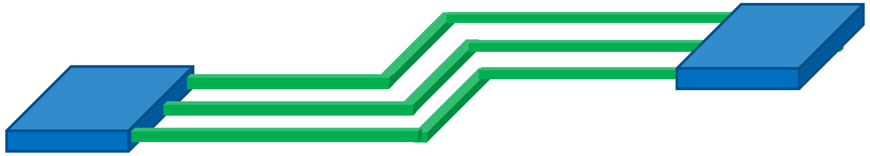


- **Distortion** (bits propagate and get distorted)
- **Reflection** (bits bounce back at all discontinuities)
- **Crosstalk** (the transmitted bits appear also on the other traces)

Interconnect models are essential to predict **signal integrity** of the channel during the **design phase** (via **simulations**) without requiring expensive prototyping→ **We need a model!!!**

Characterization

- Usually, **interconnects** are characterized by **tabulated frequency data** obtained from electromagnetic **simulations or measurements**



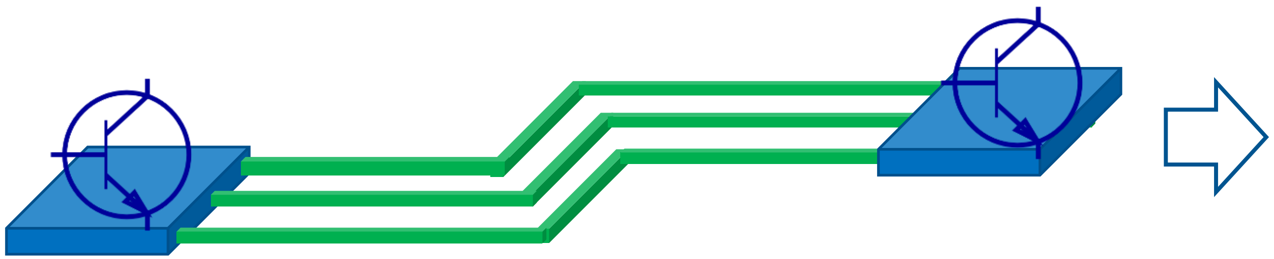
CST

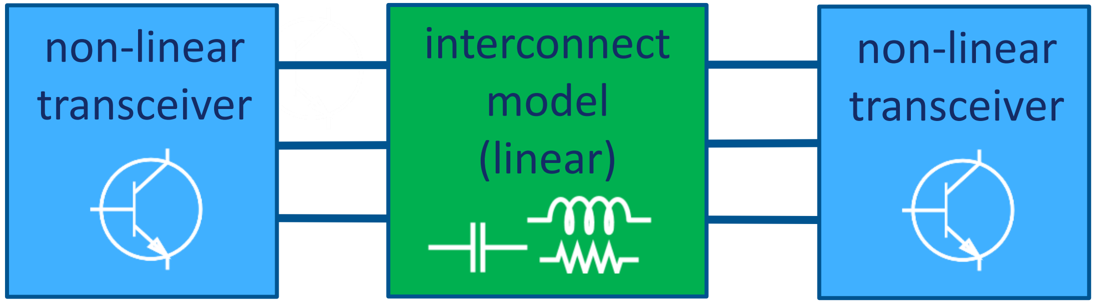
f	$H(j\omega)$
0	$H(0)$
f_1	$H(f_1)$
f_2	$H(f_2)$
f_3	$H(f_3)$
⋮	⋮

Why do we need a model?

We already have a **characterization** of the **linear interconnects**

- The link is made of **wires/PCB traces** (**linear**), while the **terminations** contain drivers, receivers, LDO and other **nonlinear components**





- Due to **nonlinear** elements, **signal integrity simulations** must be carried out in **<u>time-domain</u>**

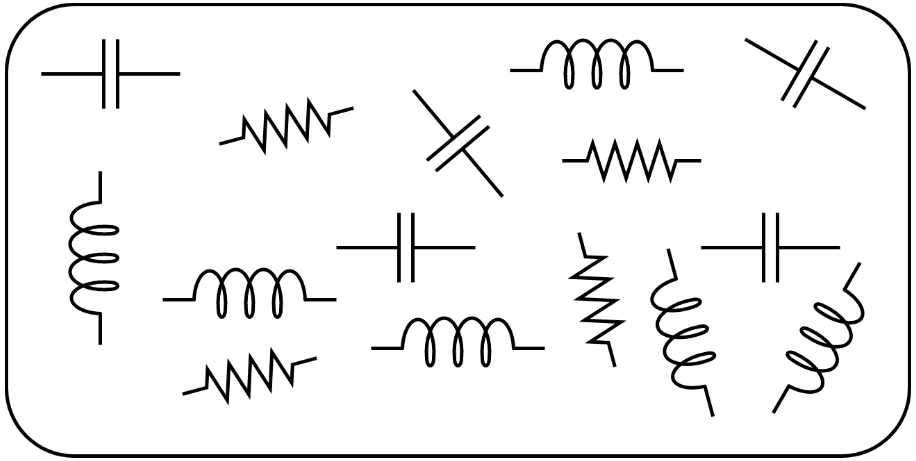
Interconnect models obtained from **frequency-domain data** should also be compatible with **time-domain circuit simulations**

- **Rational Models** are naturally adopted to model **linear structures**

$$H(j\omega) \approx \widetilde{H}(j\omega) = \sum_{j=1}^{n_p} \frac{r_j}{j\omega - p_j} + r_0$$

residues

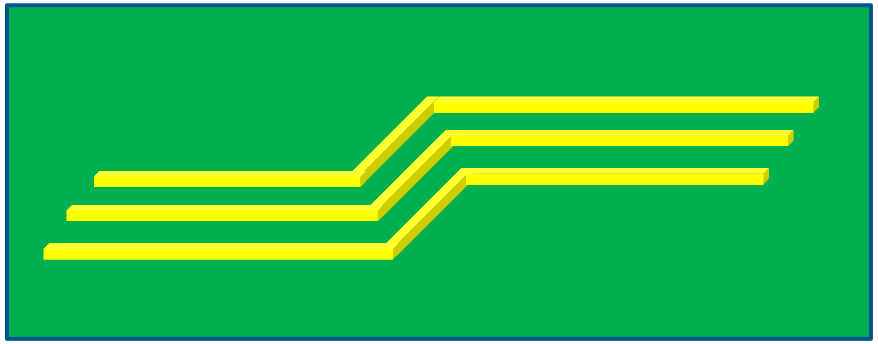
poles



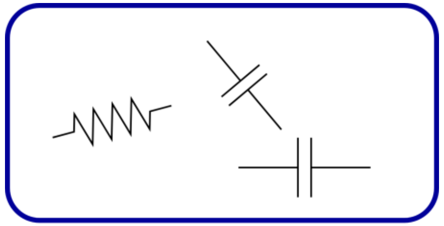
- It is a **linear expansion** of **rational basis functions**
  - Linear w.r.t. residues and nonlinear w.r.t. poles → iterative pole reallocation is used to select the optimal poles
  - An equivalent circuital representation is available

N.B. **1 pole = 1 dynamic element in the circuit** (capacitor/inductor)

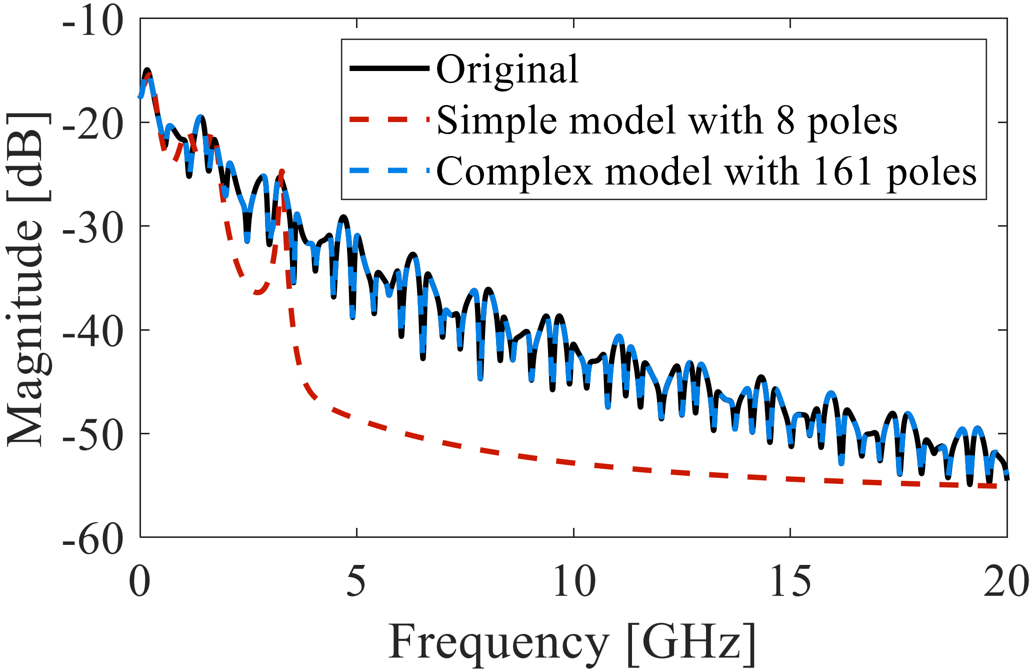
Accuracy versus Poles



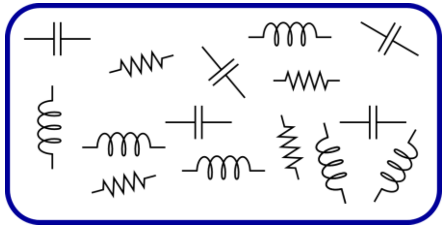
Simple model
(few poles)



Very inaccurate

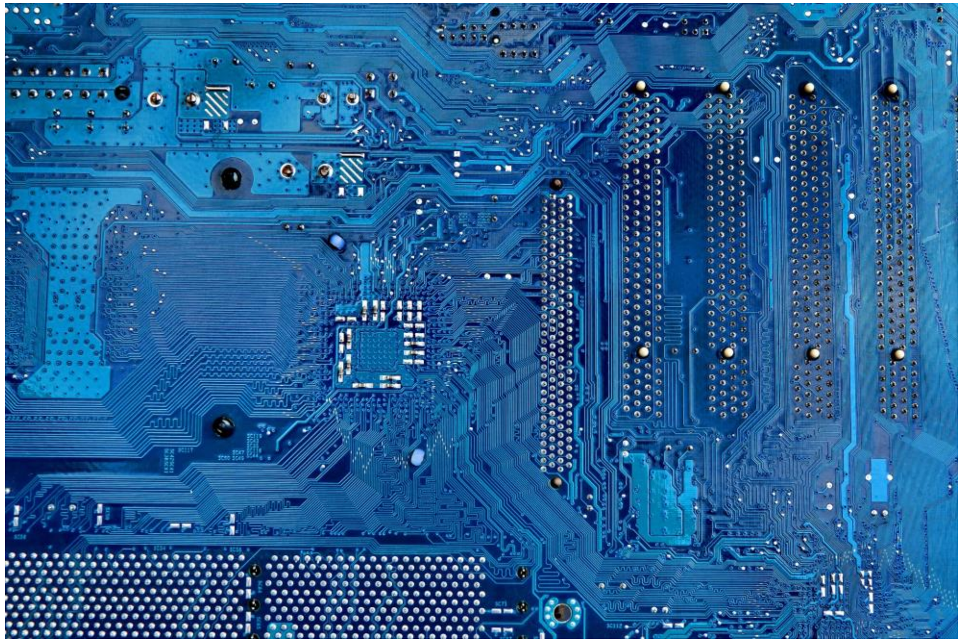


Complex model
(many poles)

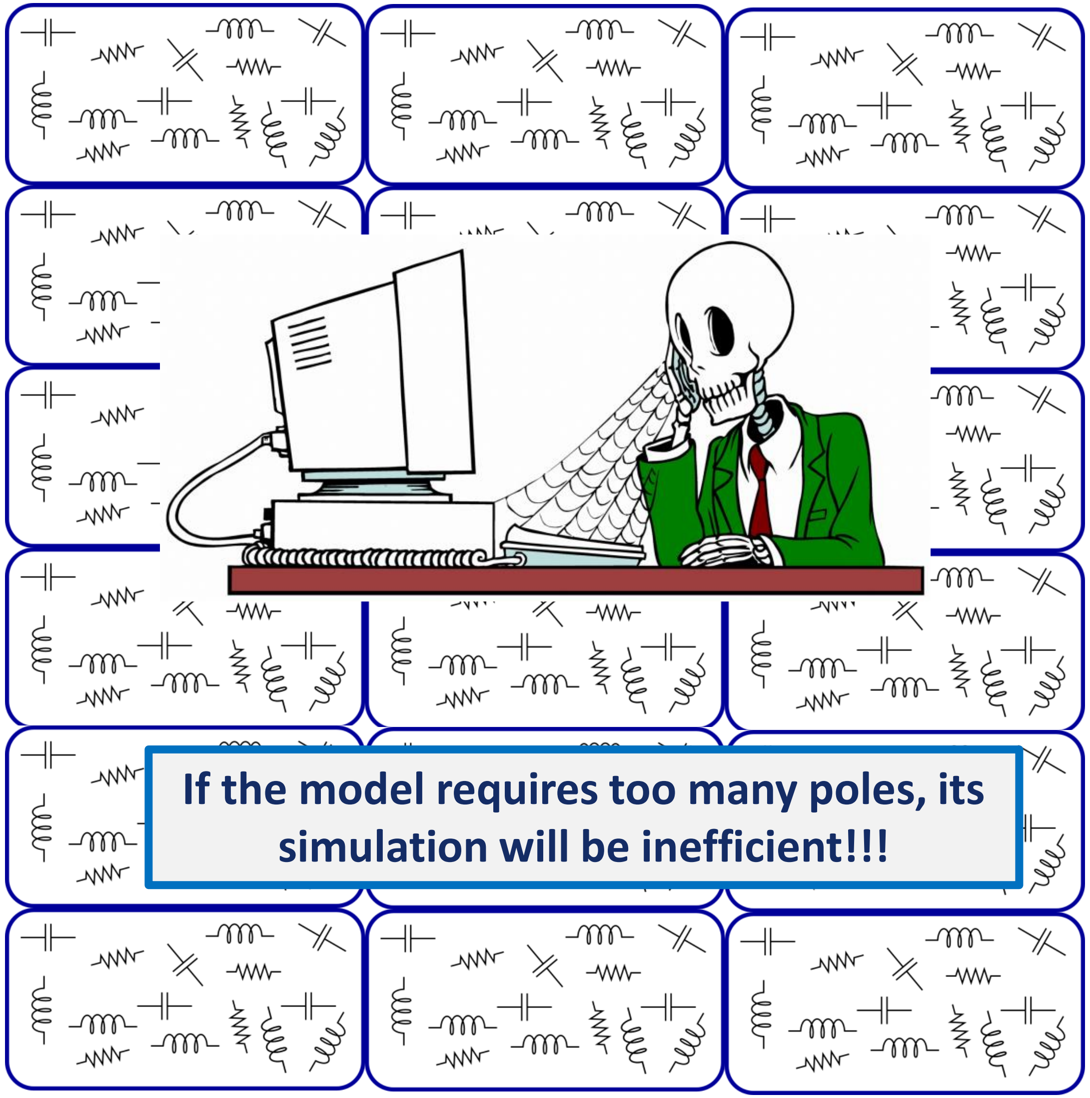


Circuital models sought for should be **accurate and fast to simulate**!!!

Larger System



- Supposing a simulation of 100 interconnects that need 161 poles each to achieve an accurate model
- 100 x 161 poles = 16,100 poles
- 16,100 dynamic elements in the simulation (capacitors/inductors)
- 16,100 additional states in the system



Delayed Rational Model (DRM)

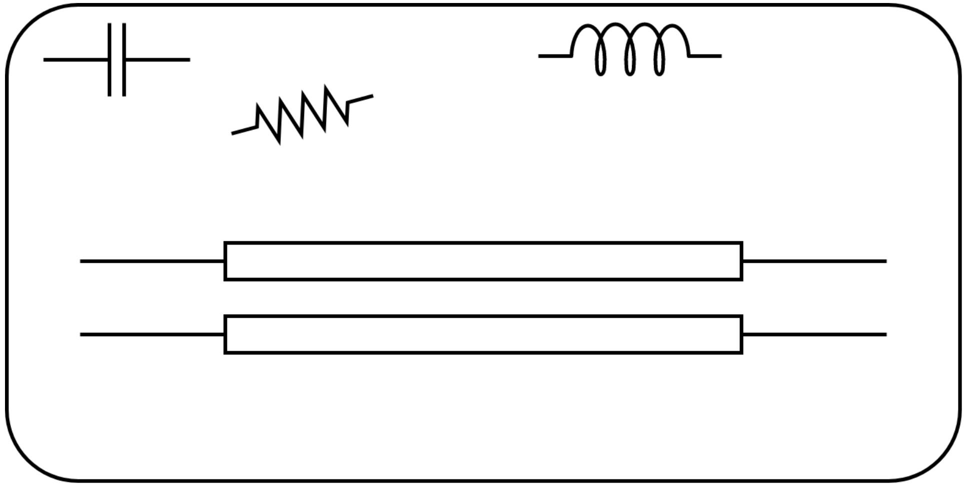
Delayed Rational Model:

$$\widetilde{H}(j\omega) = \sum_{i=1}^{n_\tau} \sum_{j=1}^{n_p} \frac{r_{ij}}{j\omega - p_{ij}} e^{-j\omega\tau_i} + r_0$$

residues

poles

delays



- **Linear expansion** of **delayed** **rational basis functions**

Advantages:

- ✓ Generally a **lower number** of poles w.r.t. the RM is required → **faster simulation time**
- ✓ **Causality** of the system is guaranteed by making $\tau_i > 0$
- ✓ **Linear** with respect to the **residues**
- ✓ **Explicit representation** of the **delayed behavior** of the transfer function

Disadvantages:

- x **Unpractical** to estimate both the **poles** and the **delays** together → usually delays are estimated first and poles afterwards
- x Generally requires **optimization** of the parameters to obtain a good model

Refs. [3]-[5]

How to find the appropriate poles and delays?

- Let us consider, as a **test function**, the transfer function

$$H(j\omega) = \frac{1}{j\omega + 3} e^{-j3\omega}$$

- We can **build a DRM** with **poles and delays** chosen on a **grid** in a $p - \tau$ **plane** (p is restricted to be real, for the sake of visualization)

- The **nodes** of the grid provide the candidate **poles and delays** to be considered with the **delayed rational model**:

To be estimated

$$\tilde{H}(j\omega) = \sum_{i=1}^{n_\tau} \sum_{j=1}^{n_{p,i}} \frac{r_{ij}}{j\omega - p_{ij}} e^{-j\omega\tau_i} + r_0$$

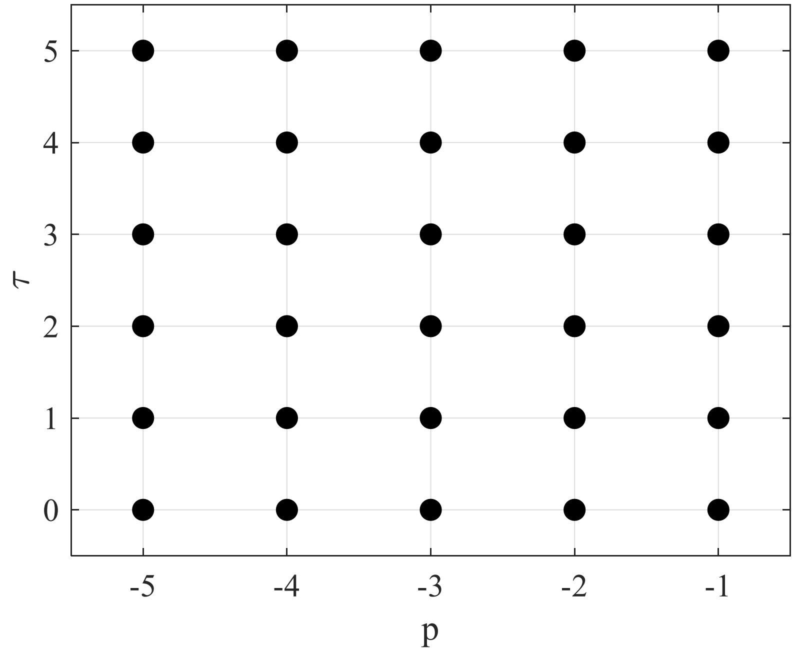
From the grid

Basis functions:

$$\tilde{H}(j\omega) = \sum w \, \varphi(\omega; p, \tau) + b \qquad \varphi(\omega; p, \tau) = \frac{c_{ij}}{j\omega - p_{ij}} e^{-j\omega\tau_i}$$

Linear parameter

Nonlinear parameters

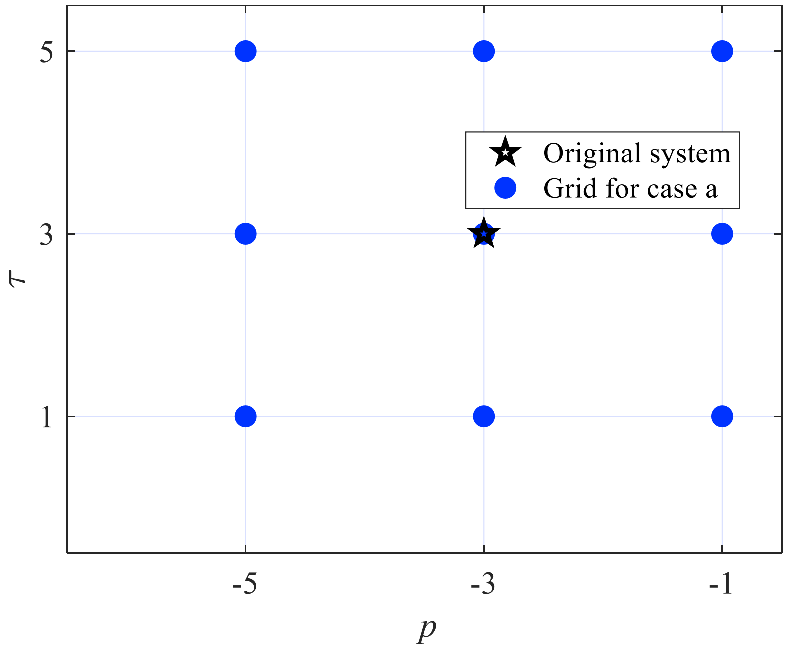


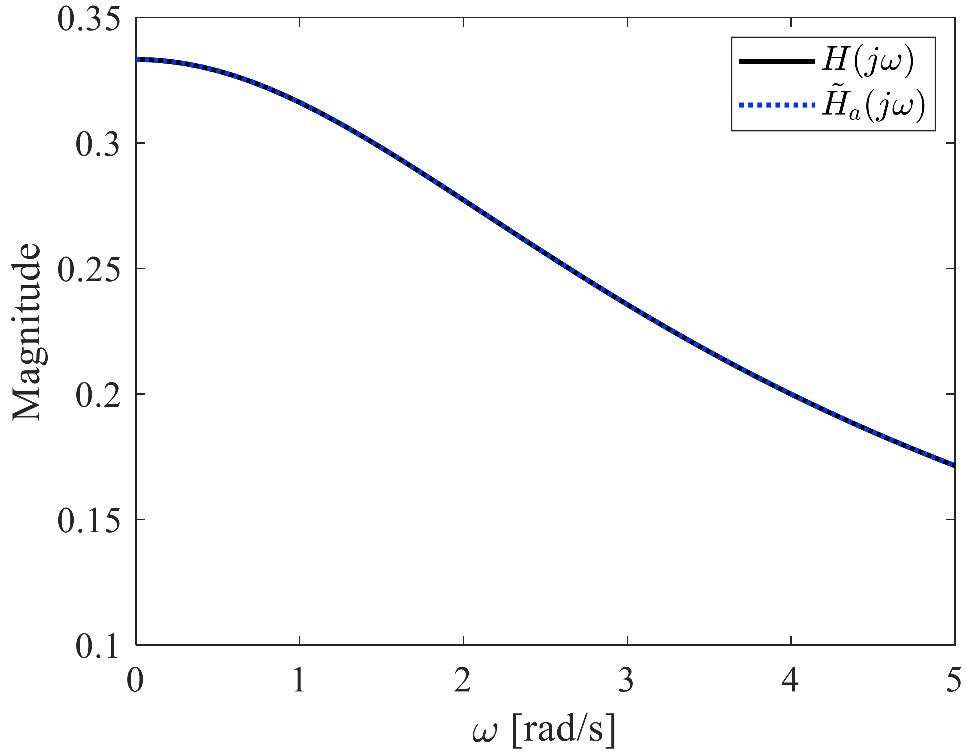
Grid Approximation – (a)

- Let us try to fit $H(j\omega) = \frac{1}{j\omega+3} e^{-j3\omega}$ by making a grid in a $p - \tau$ plane:

- If the grid captures **exact pole and delay**

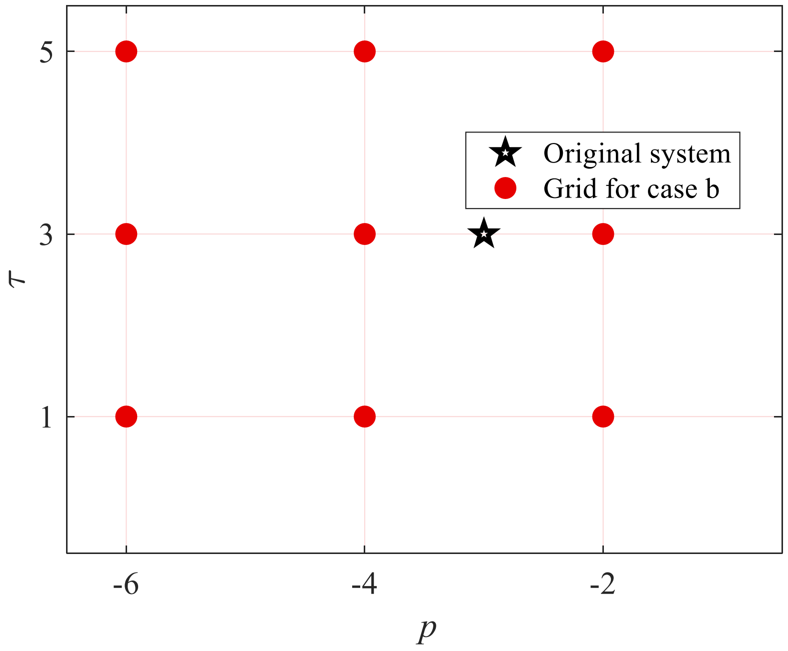
  - (a) - Approx. model is **essentially perfect**

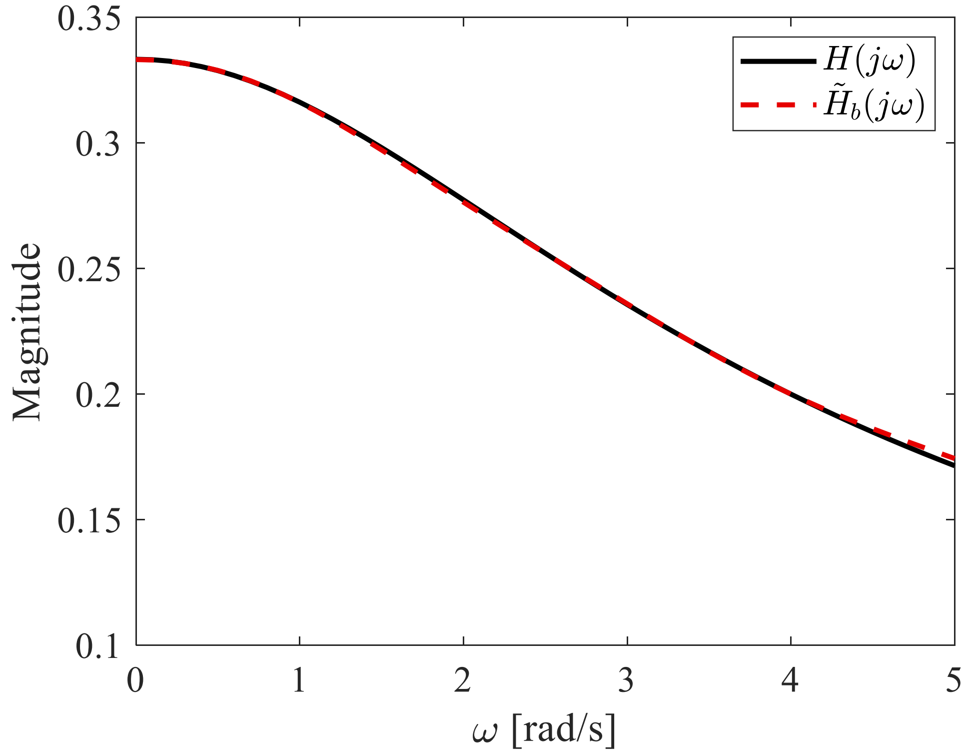




Grid Approximation – (b)

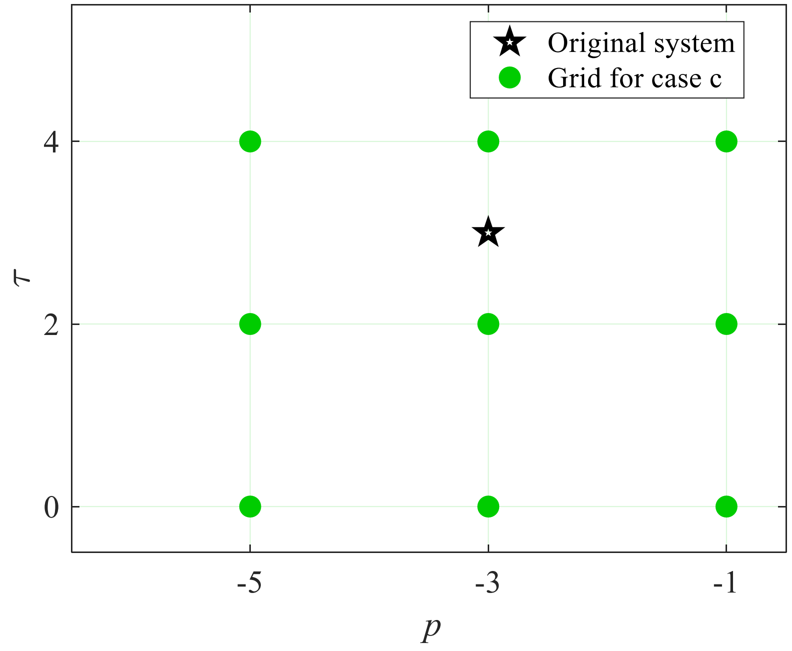
- Let us try to fit $H(j\omega) = \frac{1}{j\omega+3} e^{-j3\omega}$ by making a grid in a $p-\tau$ plane:
- If the grid captures **only the exact delay**
  - (b) – Approx. model is still **very good**

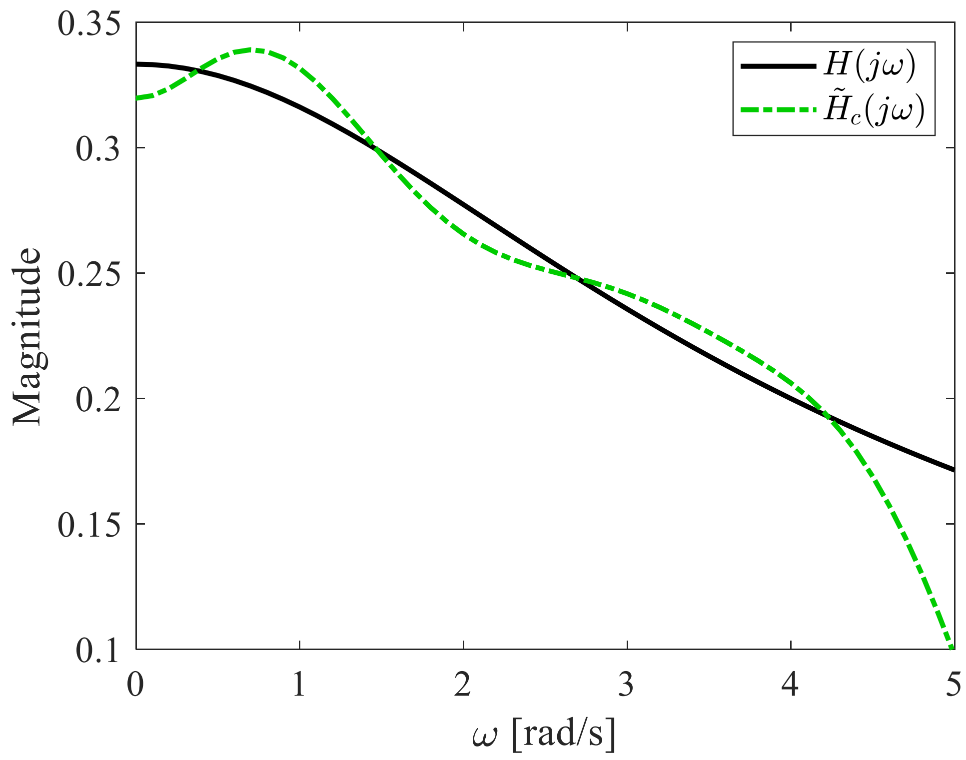




Grid Approximation – (c)

- Let us try to fit $H(j\omega) = \frac{1}{j\omega+3} e^{-j3\omega}$ by making a grid in a $p - \tau$ plane:
- If the grid captures **only the exact pole**
  - (c) – Approx. model is clearly **inaccurate**





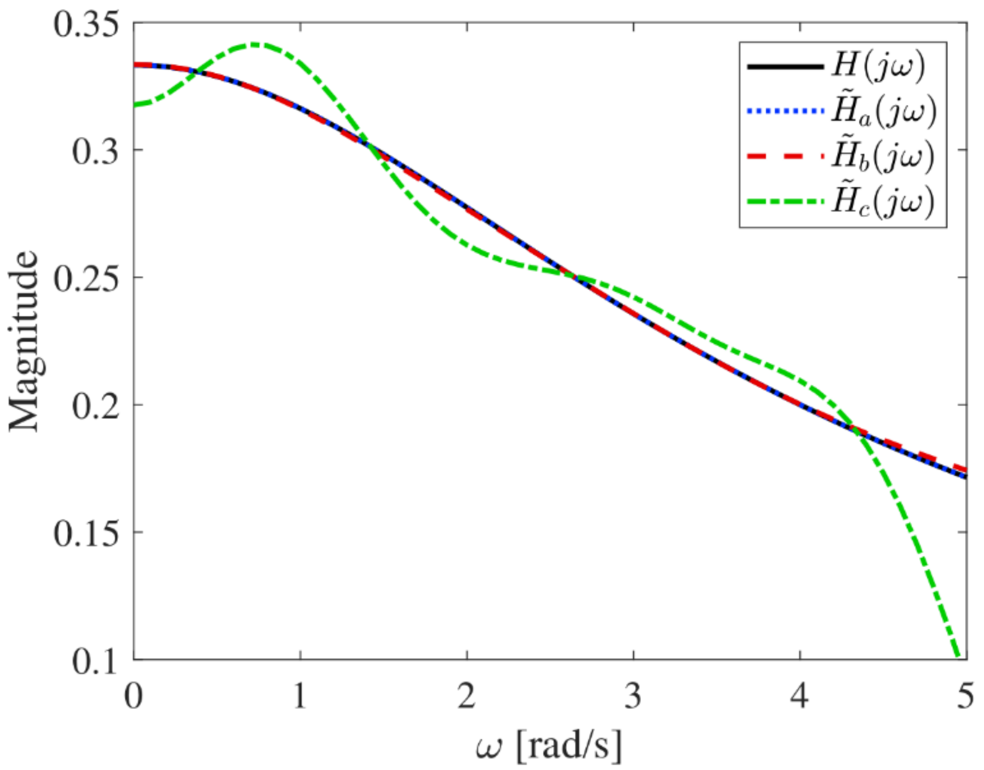
Grid Approximation - Summary

- Let us try to fit $H(j\omega) = \frac{1}{j\omega+3} e^{-j3\omega}$ by making a grid in a $p - \tau$ plane:

- Three cases considered:

  - (a) – Approx. model is **essentially perfect**
  - (b) – Approx. model is still **very good**
  - (c) – Approx. model is clearly **inaccurate**

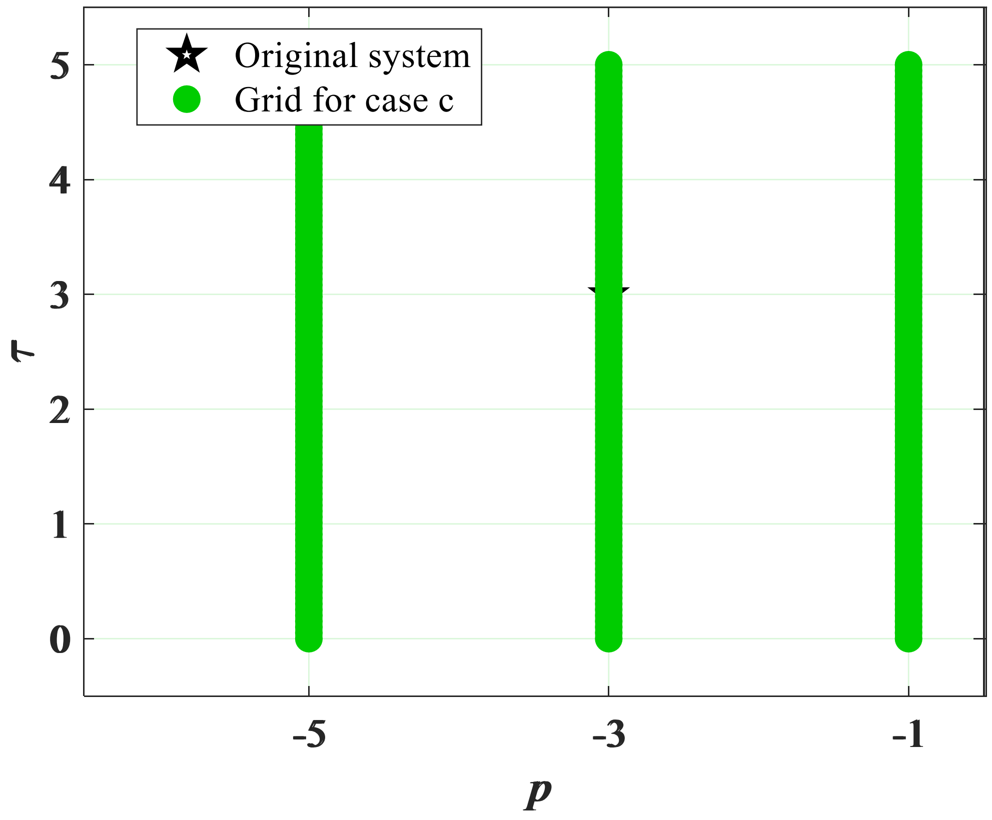
- A larger number of poles can compensate a non-exact estimation of the poles, but a wrong delay estimation generates an inaccurate model, even if it uses the right poles



An **accurate delay estimation** is essential to obtain an **accurate delayed rational model**

Grid Approximation – Extreme Case

- The DRM should contain the **exact delay** of the transfer function it approximates

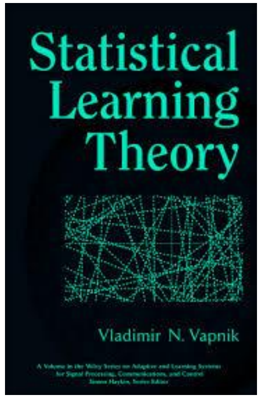


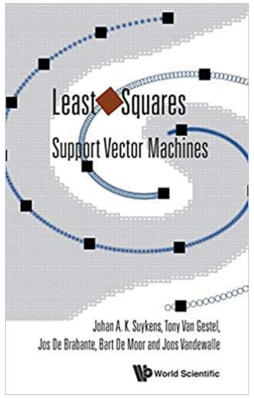
The only way to ensure that an unknown delay is included in the model is by considering an infinite number of delays

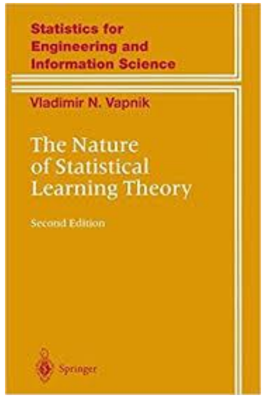
$$\widetilde{H}(j\omega) = \sum_{i=1}^{\infty} \sum_{j=1}^{n_{p,i}} \frac{r_{ij}}{j\omega - p_{ij}} e^{-j\omega\tau_i} + r_0$$

How can we estimate a model with an **infinite number of terms**?

❑ Support Vector Machines (SVMs) [6][7]

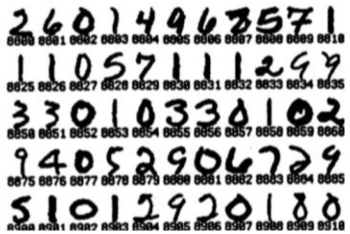




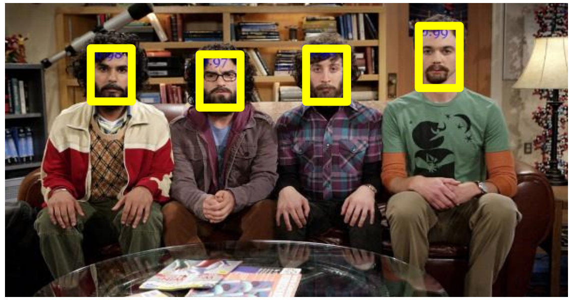


➢ Historical applications:

Handwriting recognition



Face detection



❑ **Model:**

$$\widetilde{H}(j\omega) = \sum_{k=1}^{K} \alpha_k \, k(\omega, \omega_k) + b$$

# of training samples (K)

$$k(\omega, \omega_k) = \langle \boldsymbol{\varphi}(\omega; \boldsymbol{p}, \boldsymbol{\tau}), \boldsymbol{\varphi}^*(\omega_k; \boldsymbol{p}, \boldsymbol{\tau}) \rangle$$

Inner product

Basis functions depending on the considered poles and delays

❑ Kernel is linked to a vector with the basis functions of a regression model → **vector can be infinite dimensional**!!! [6]

Duality

- The **Least Squares Support Vector Machine** (LS-SVM) regression has **two equivalent formulations** [7]:

Dual Space

$$\widetilde{H}(j\omega) = \sum_{k=1}^{K} \alpha_k \, k(\omega, \omega_k) + b \quad \Rightarrow \quad w = \sum_k \alpha_k \varphi^*(\omega_k) \quad \Rightarrow$$

- **Non-parametric model** → number of terms equal to the number of samples

Primal Space:

$$\widetilde{H}(j\omega) = \langle \boldsymbol{w}, \boldsymbol{\varphi}(\omega; \boldsymbol{p}, \boldsymbol{\tau}) \rangle + b$$

$$\downarrow$$

$$\widetilde{H}(j\omega) = \sum_{i,j} w_{ij} \, \varphi(\omega; p_{ij}, \tau_i) + b$$

- **Parametric model** → number of terms equal to the number of basis functions

Equivalence with DRM

- Weights w **are proportional to the residues** of a delayed-rational model

<u>ML model:</u> $$\tilde{H}(j\omega) = \sum w(\tau_i, p_{ij})\,\varphi(\omega; p_{ij}, \tau_i) + b$$

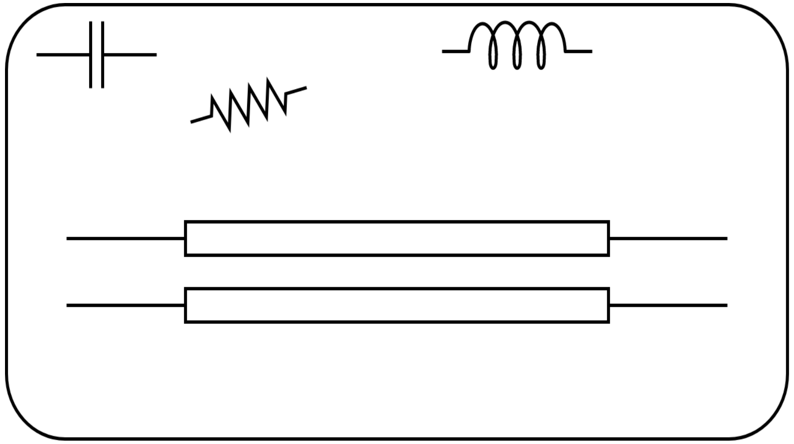
<u>DRM:</u> $$\tilde{H}(j\omega) = \sum_{i=1}^{n_\tau} \sum_{j=1}^{n_{p,i}} \frac{r_{ij}}{j\omega - p_{ij}} e^{-j\omega\tau_i} + r_0$$

- By looking at the values of **w as a function of τ**, we are able to see for which values of τ the w **is larger**, i.e., the **dominant propagation delays of the system**

Application of the Method

- The **identified propagation delays** can be employed to build **low-order delayed rational models**

$$\widetilde{H}(j\omega) = \sum_{i=1}^{n_\tau} \sum_{j=1}^{n_{p,i}} \frac{r_{ij}}{j\omega - p_{ij}} e^{-j\omega\tau_i} + r_0$$



Identify delays with the proposed method

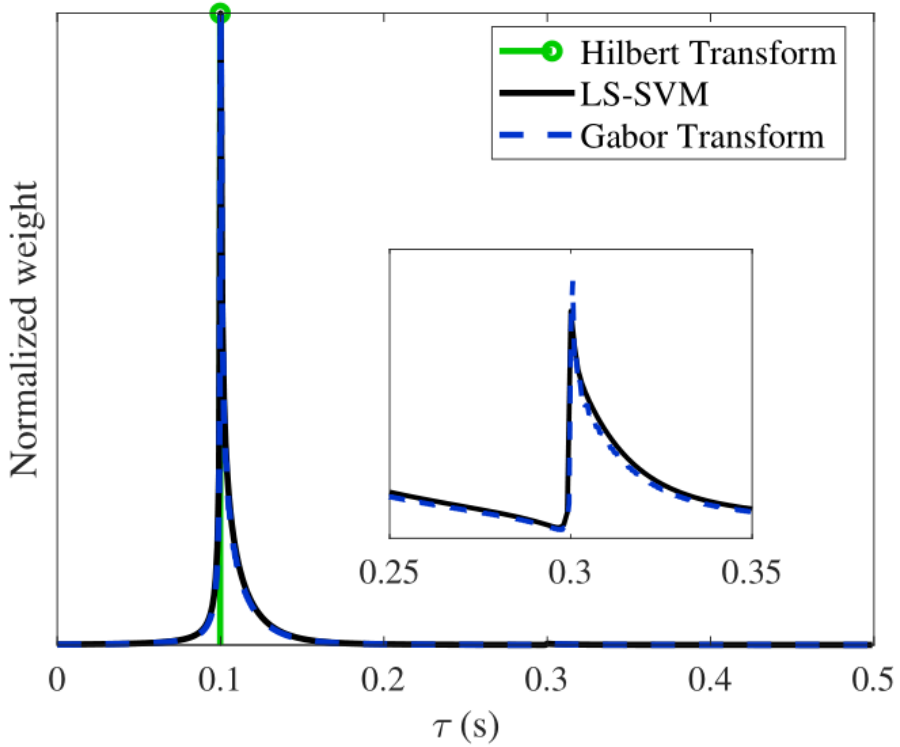
Employ such delays to obtain an accurate and efficient DRM

Application Examples – I

- Known transfer function:

$$H(j\omega) = \left(\frac{1}{j\omega + 60 + 20j} + \frac{1}{j\omega + 60 - 20j}\right) e^{-j\overset{\tau_1}{0.1}\omega} + \frac{0.075}{j\omega + 100} e^{-j\overset{\tau_2}{0.3}\omega}$$

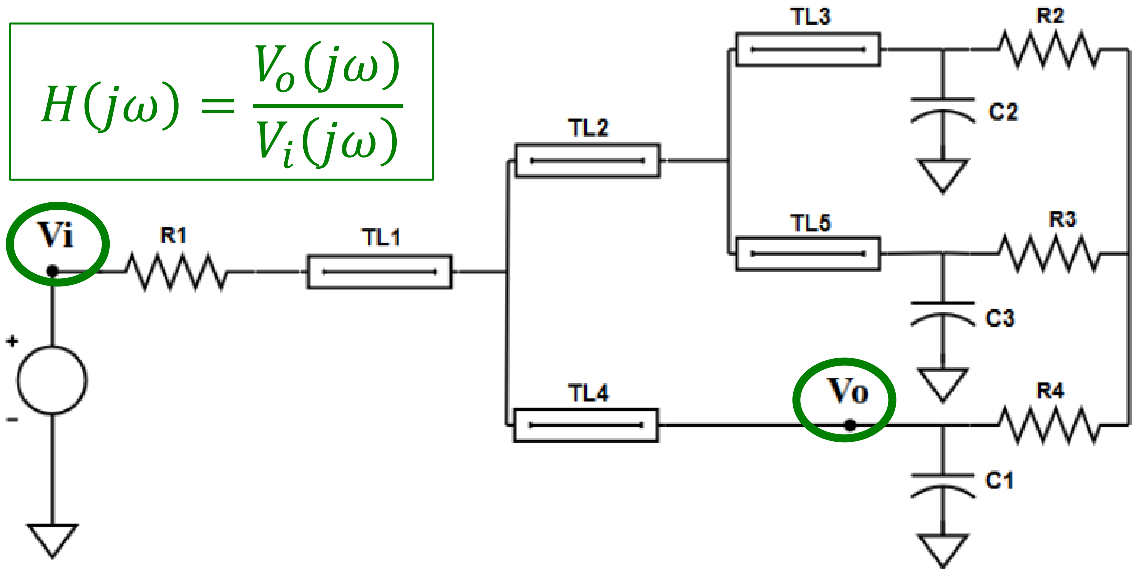
- **Real** and **complex-conjugate** poles
- **First delay** is the same as identified by **Hilbert transform** [8]
- Overall curve is like the one obtained by the **Gabor transform** [3]

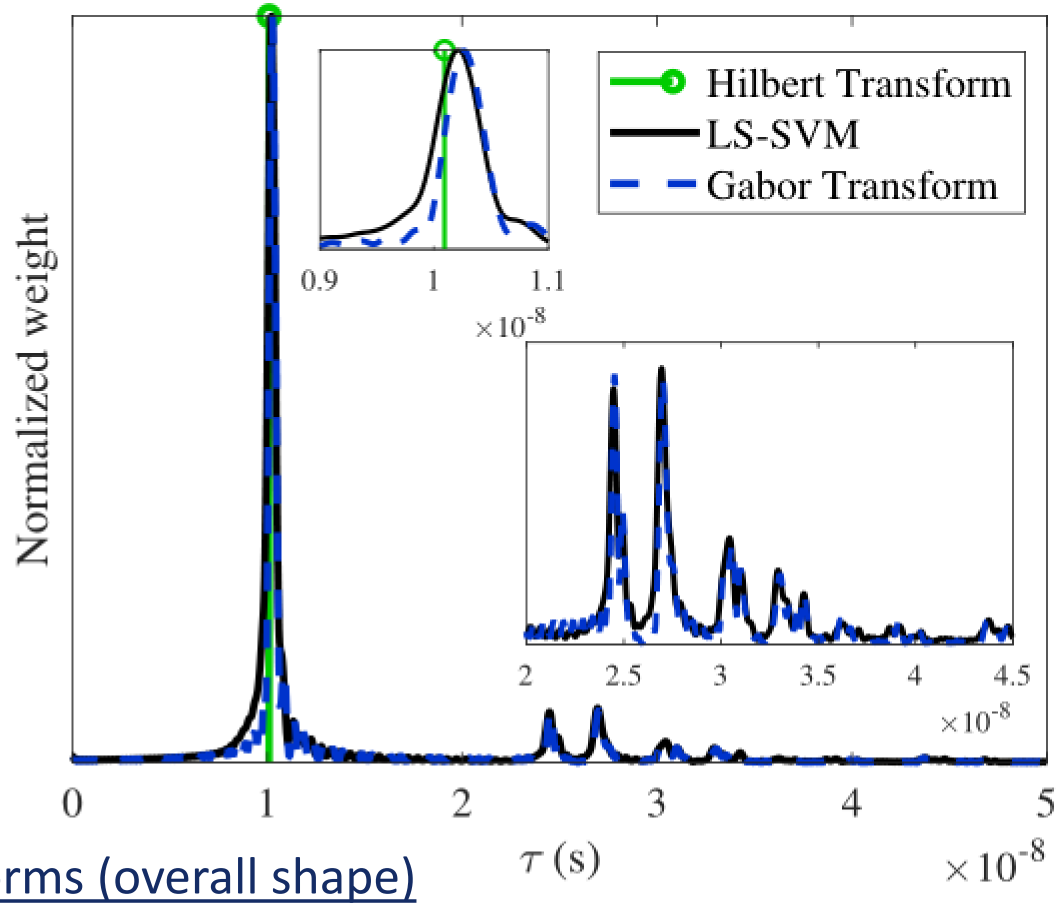


Application Examples – II

- Circuit with **multiple transmission lines**:

$$H(j\omega) = \frac{V_o(j\omega)}{V_i(j\omega)}$$



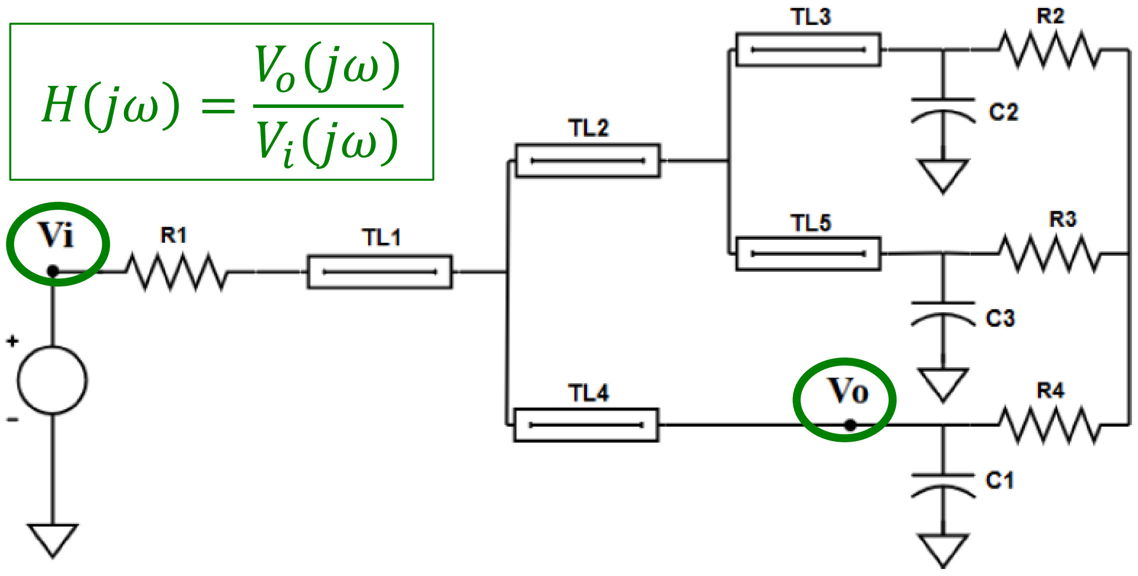


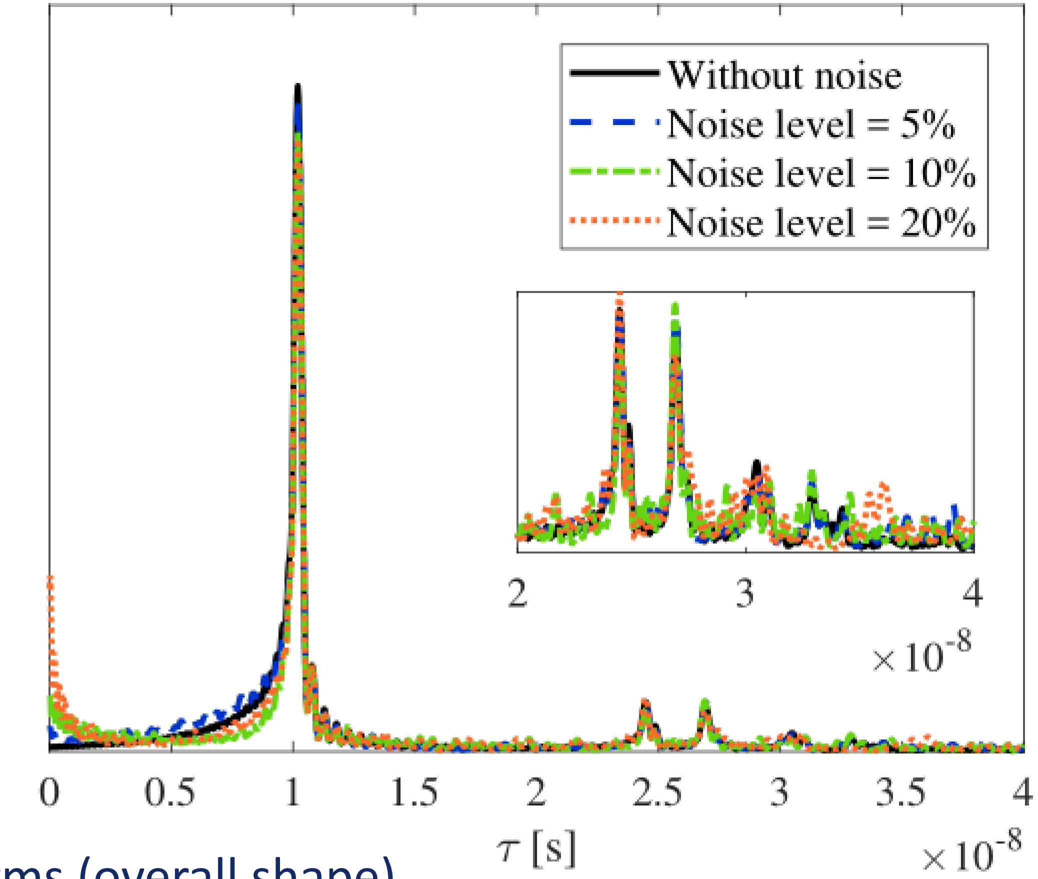
- Three paths from input to output, **wave reflection** at the **discontinuities**
- **Multiple delays** expected in $H(j\omega)$
- Accuracy similar to Hilbert (1^{st} delay) and Gabor transforms (overall shape)

Application Examples – II (with Noise)

- Circuit with **multiple transmission lines**:

$$H(j\omega) = \frac{V_o(j\omega)}{V_i(j\omega)}$$

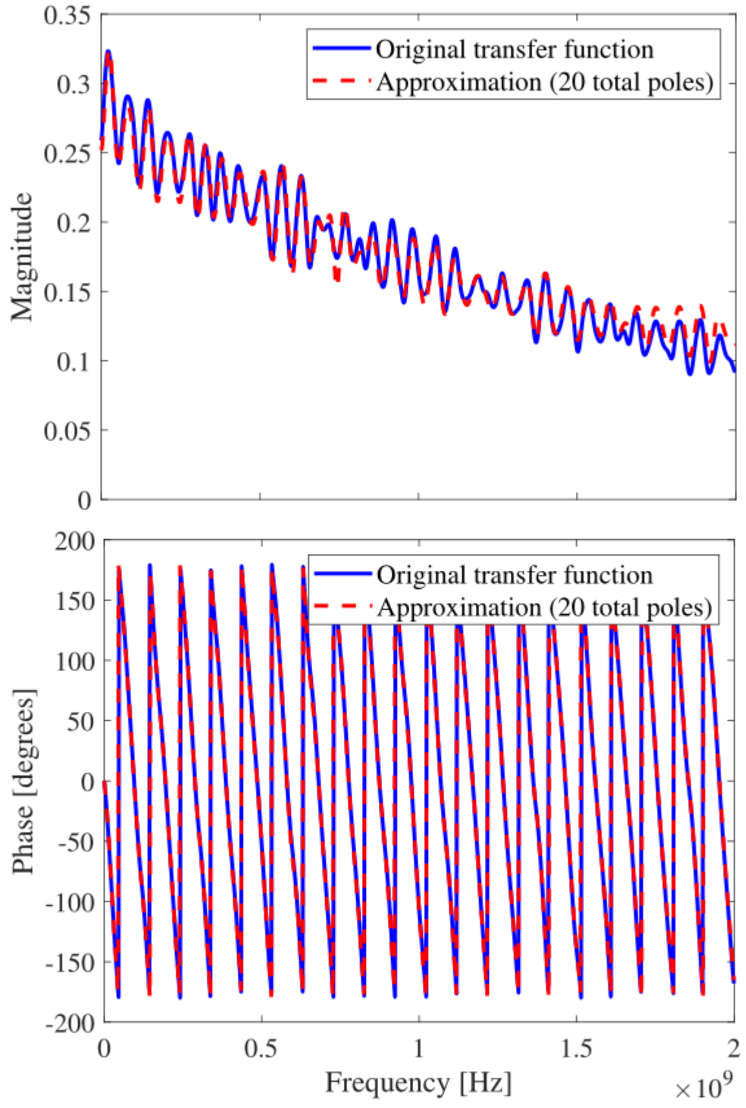


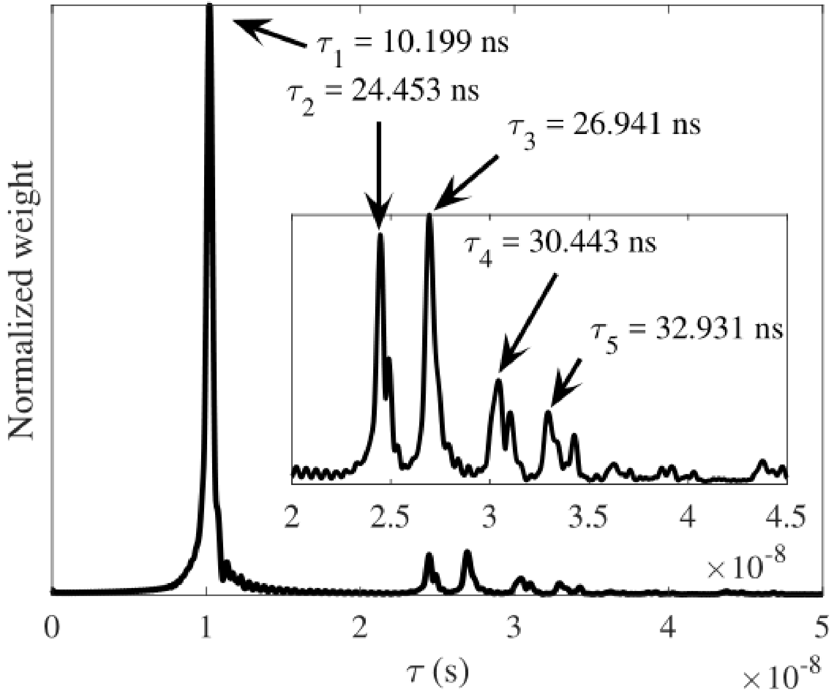


- Three paths from input to output, **wave reflection** at the **discontinuities**
- **Multiple delays** expected in $H(j\omega)$
- Accuracy similar to Hilbert (1st delay) and Gabor transforms (overall shape)
- Good performance with noisy data

Application Examples – II (Summary)

- Circuit with **multiple transmission lines**:





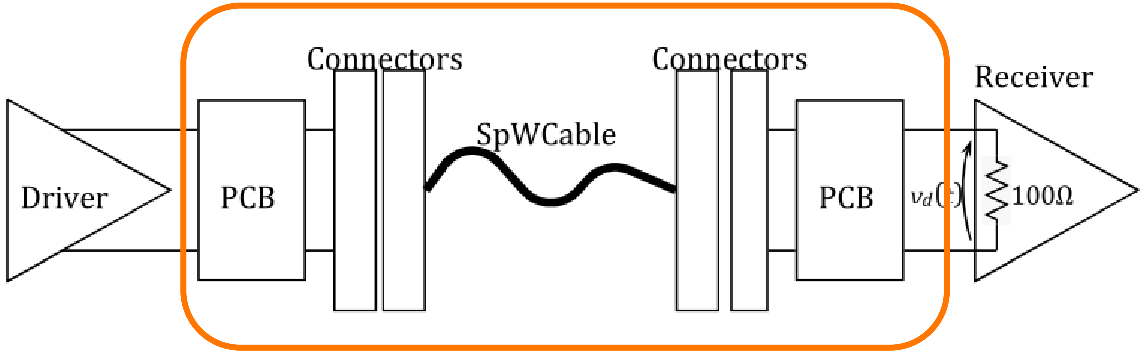
- Five delays and 20 poles distributed among them are sufficient to accurately model the original transfer function

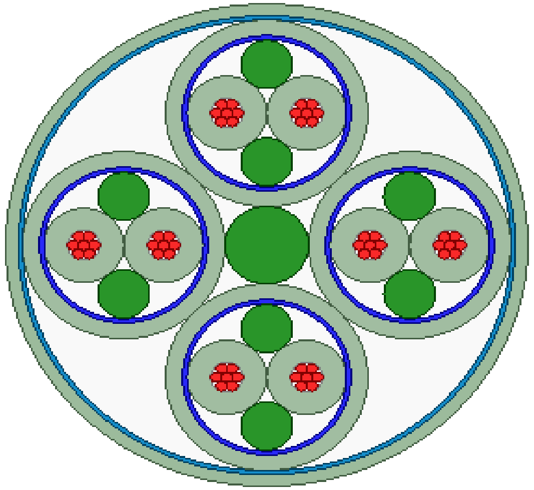
- A rational model with similar accuracy requires 54 poles

	Proposed model	VF model
Error - L_2−norm	0.3344	0.5200
Error - L_∞-norm	0.0261	0.0546
Order	20 total poles	54 total poles

Application Examples – III

- **SpaceWire (SpW)** cable link:

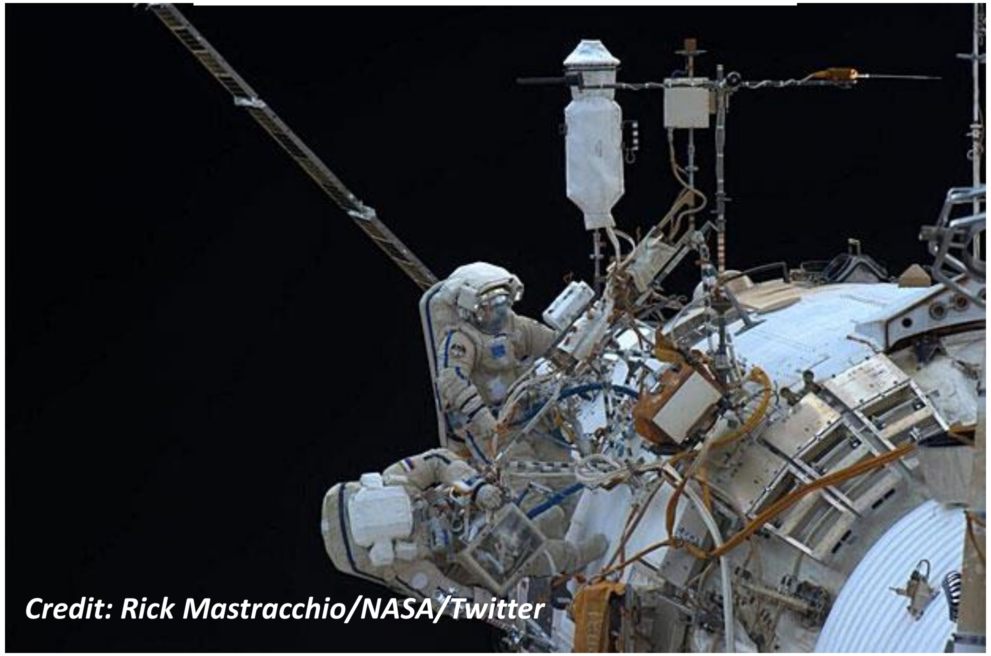




- **Scattering matrix** of two wires of the link is considered:

$$S(j\omega) = \begin{bmatrix} S_{1,1}(j\omega) & S_{1,2}(j\omega) \\ S_{2,1}(j\omega) & S_{2,2}(j\omega) \end{bmatrix}$$

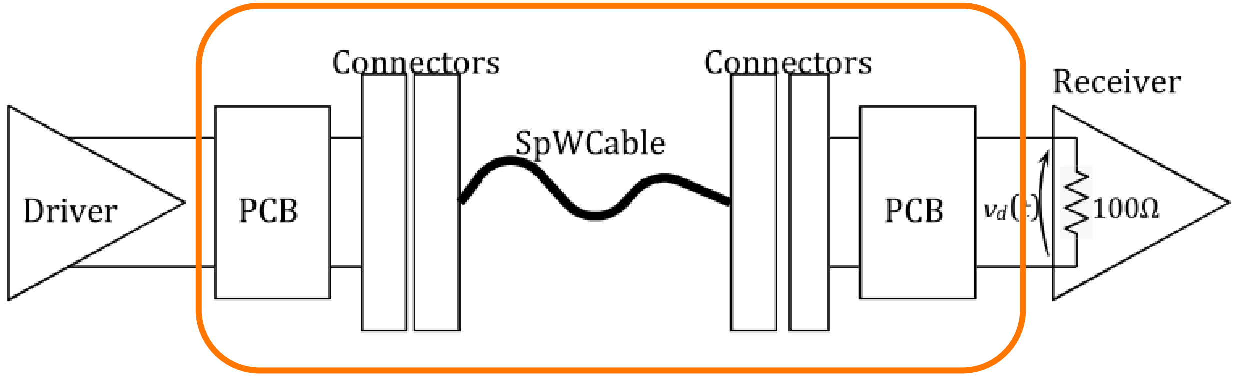
- Cable channel linking a driver and a receiver through:
  - **Striplines**
  - 9-pin Micro-D **connectors**
  - **SpW cable**
    - **4 twisted pairs** of wires
    - **1 inner shield** around each of the pairs
    - **1 outer shield**

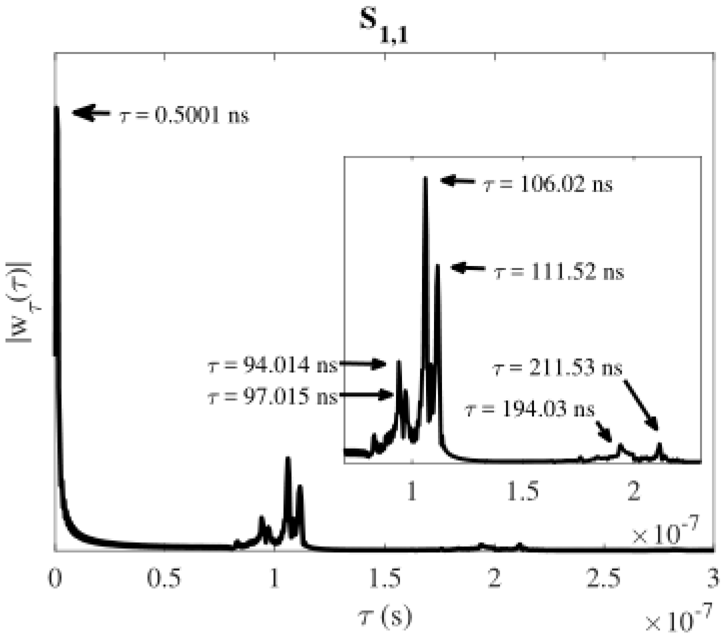


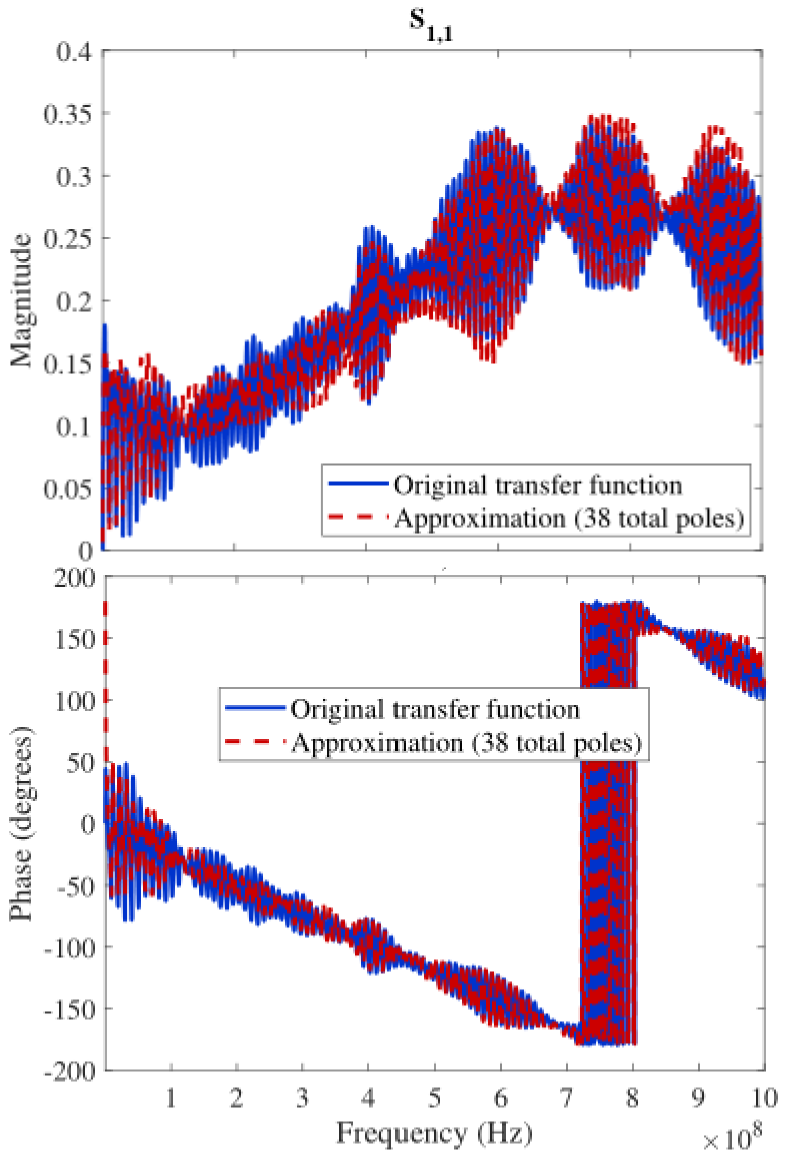
Credit: Rick Mastracchio/NASA/Twitter

Application Examples – III ($S_{1,1}$)

- **SpaceWire** cable link:

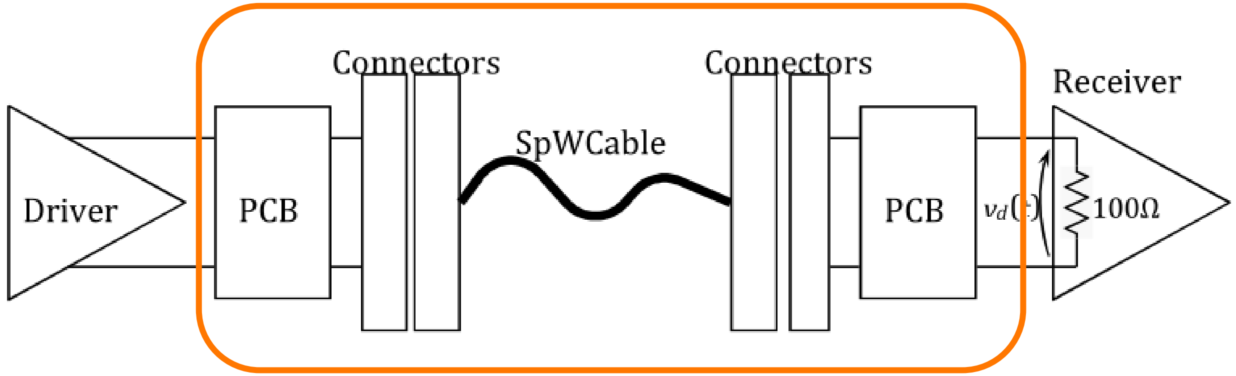


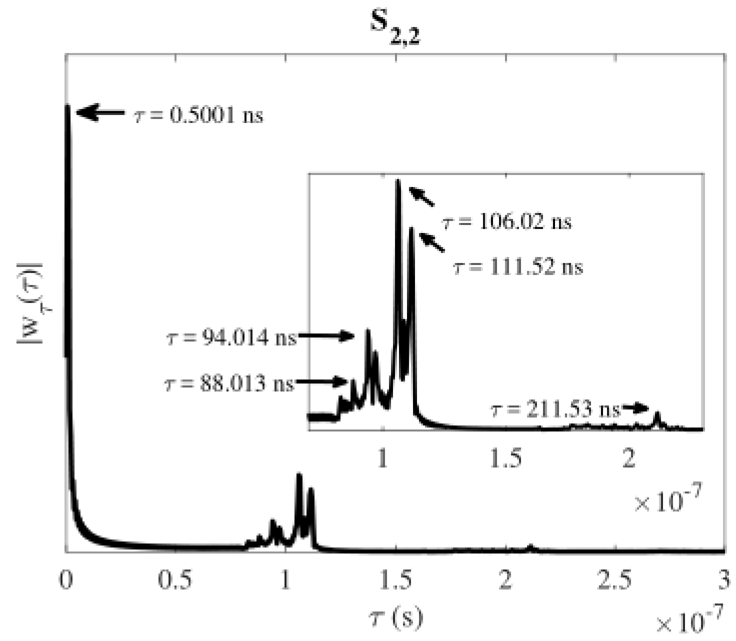


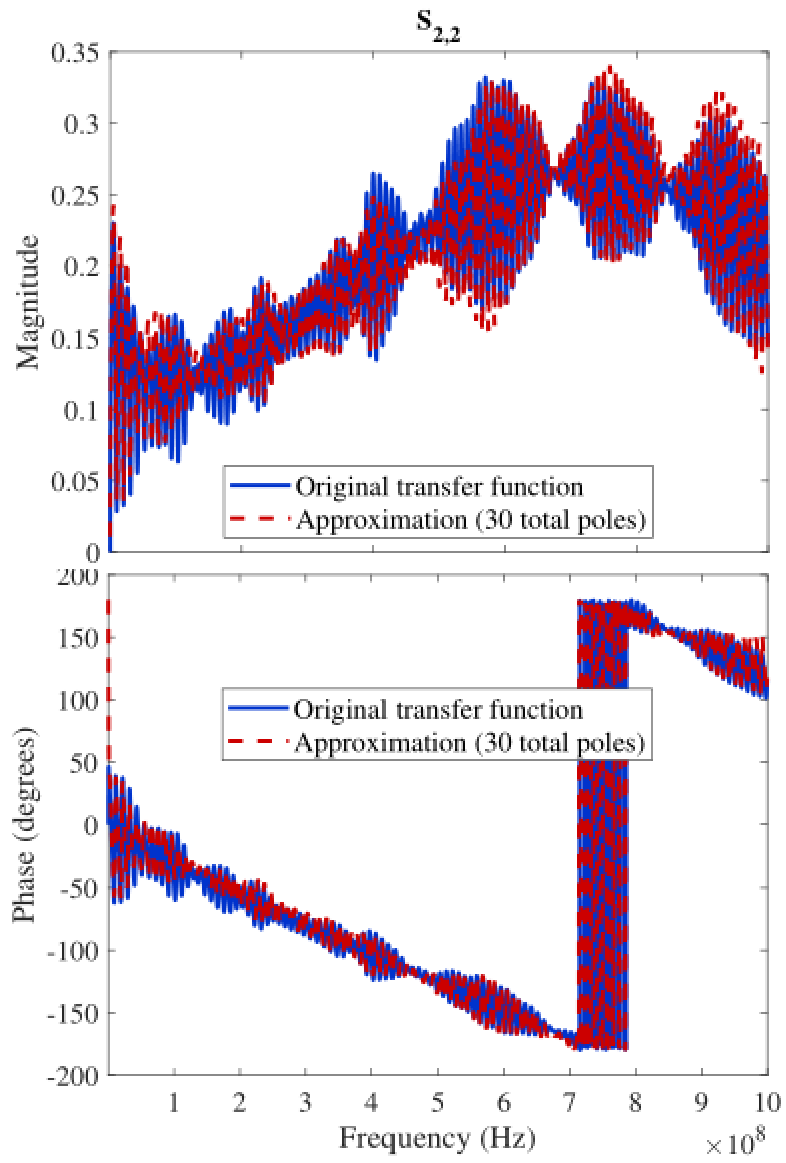


Application Examples – III ($S_{2,2}$)

- **SpaceWire** cable link:

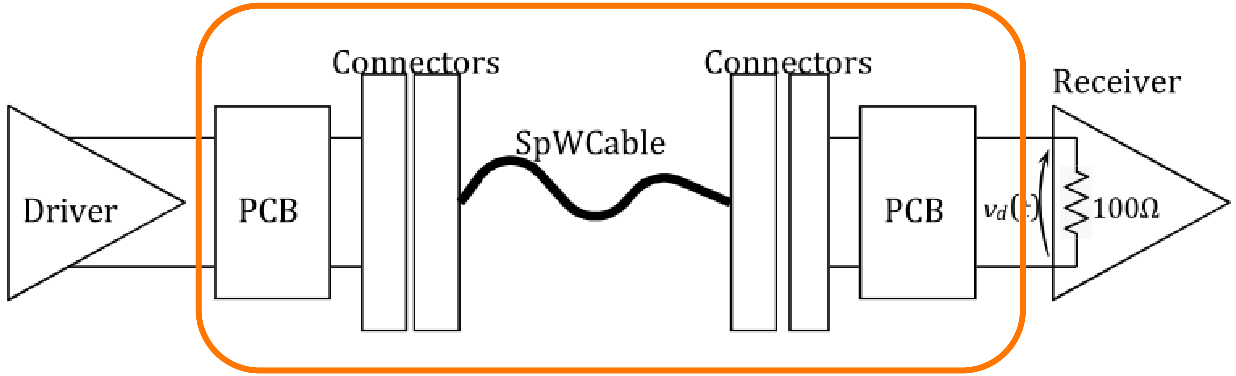


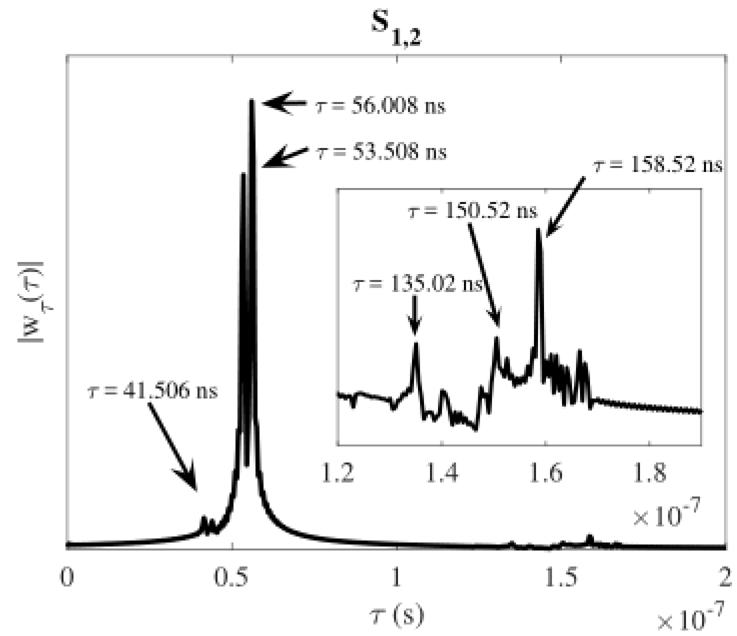


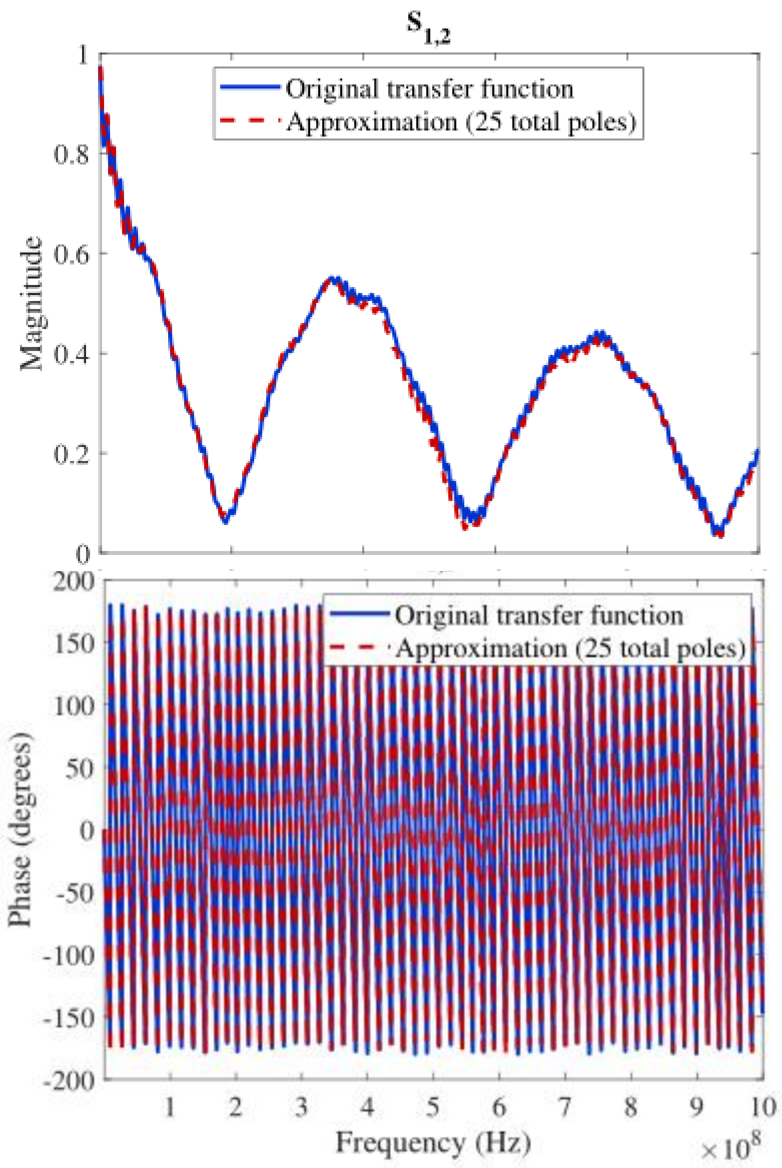


Application Examples – III ($S_{1,2}$)

- **SpaceWire** cable link:

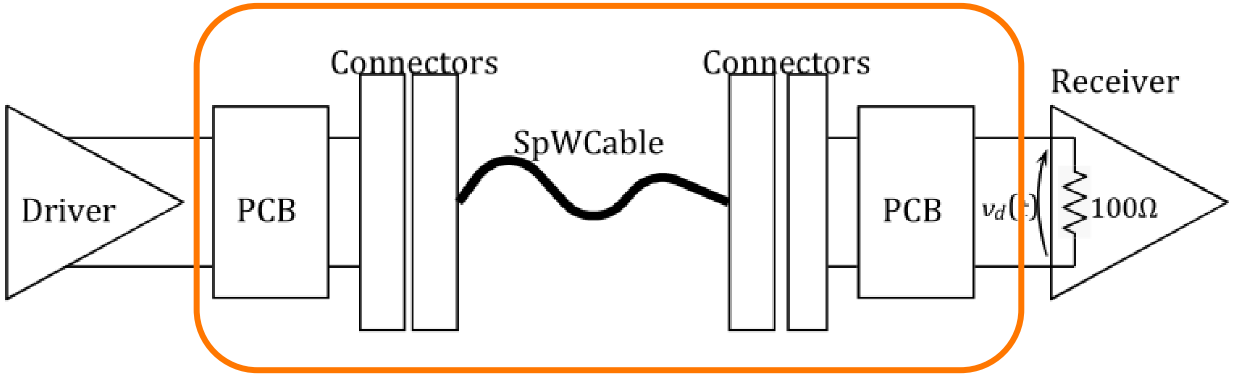






Application Examples – III (Summary)

- **SpaceWire** cable link:



	Proposed model	VF
$S_{1,1}$ error - L_2	0.780	0.776
$S_{1,1}$ error - L_∞	0.085	0.086
$S_{1,1}$ order	38	180
$S_{1,2}$ error - L_2	0.806	1.228
$S_{1,2}$ error - L_∞	0.041	0.177
$S_{1,2}$ - order	25	119
$S_{2,2}$ error - L_2	0.766	0.755
$S_{2,2}$ error - L_∞	0.053	0.078
$S_{2,2}$ - order	30	164

- All the **delayed-rational models** built with the identified delays require **less poles** than a pure **rational model** with similar accuracy

- Kernel depends only on **frequency points**, chosen poles, τ_m and τ_M:

$$k(\omega, \omega_k; \boldsymbol{p}, \tau_m, \tau_M)$$

  - Same kernel for all the 3 terms of the matrix!

Around **5x less poles**!

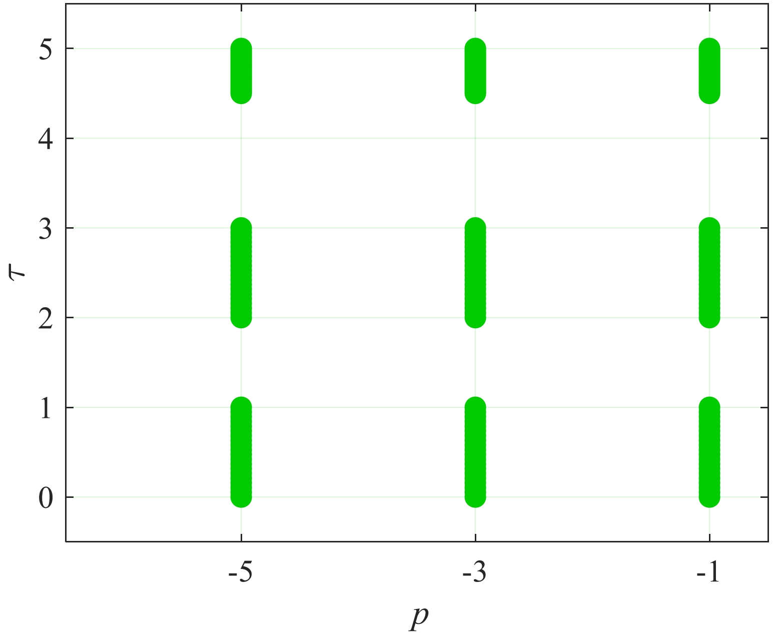
Conclusions

- **Delayed-rational models** allow reducing the **complexity** of **models** of **distributed systems.** Examples showed a **reduction of 2.5-5 times** in the total number of poles when comparing with rational models.

- **ML kernel-based regression** (e.g., **Least-Squares Support Vector Machine** (LS-SVM)) can be adopted for the estimation of the **dominant delays** in distributed systems

- The **LS-SVM approach** provides a very accurate identification of the network delays (comparable with **Hilbert transform** – when applicable – and with **Gabor transform** method for multiple delays), and generates a rational approximation with a number of poles significantly reduced w.r.t. conventional fitting methods

N.B:
The proposed methodology for the delay estimation is extremely flexible, i.e., poles and delay interval can be changed as the knowledge about the system increases. E.g., the model can consider multiple delay intervals:



Thank you!

Thank you very much for the attention!

Questions?

References

[1] R. Achar and M. S. Nakhla, "Simulation of high-speed interconnects," in *Proceedings of the IEEE*, vol. 89, no. 5, pp. 693-728, May 2001, doi: 10.1109/5.929650.

[2] S. N. Lalgudi, E. Engin, G. Casinovi and M. Swaminathan, "Accurate Transient Simulation of Interconnects Characterized by Band-Limited Data With Propagation Delay Enforcement in a Modified Nodal Analysis Framework," in *IEEE Transactions on Electromagnetic Compatibility*, vol. 50, no. 3, pp. 715-729, Aug. 2008, doi: 10.1109/TEMC.2008.924394.

[3] A. Chinea, P. Triverio and S. Grivet-Talocia, "Delay-Based Macromodeling of Long Interconnects From Frequency-Domain Terminal Responses," in *IEEE Transactions on Advanced Packaging*, vol. 33, no. 1, pp. 246-256, Feb. 2010, doi: 10.1109/TADVP.2008.2010525.

[4] A. Chinea *et al*., "Signal Integrity Verification of Multichip Links Using Passive Channel Macromodels," in *IEEE Transactions on Components, Packaging and Manufacturing Technology*, vol. 1, no. 6, pp. 920-933, June 2011, doi: 10.1109/TCPMT.2011.2138136.

[5] M. Sahouli and A. Dounavis, "Delay Extraction-Based Modeling Using Loewner Matrix Framework," in *IEEE Transactions on Components, Packaging and Manufacturing Technology*, vol. 7, no. 3, pp. 424-433, March 2017, doi: 10.1109/TCPMT.2017.2650138.

[6] Cristianini, N., & Shawe-Taylor, J. *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge: Cambridge University Press, 2000.

[7] Johan A K Suykens et al. *Least Squares Support Vector Machines*. Default Book Series. November 2002.

[8] R. Mandrekar and M. Swaminathan, "Causality enforcement in transient simulation of passive networks through delay extraction," *Proceedings. 9th IEEE Workshop on Signal Propagation on Interconnects, 2005.*, Garmisch-Partenkirchen, Germany, 2005, pp. 25-28, doi: 10.1109/SPI.2005.1500884.