XXXV Cycle

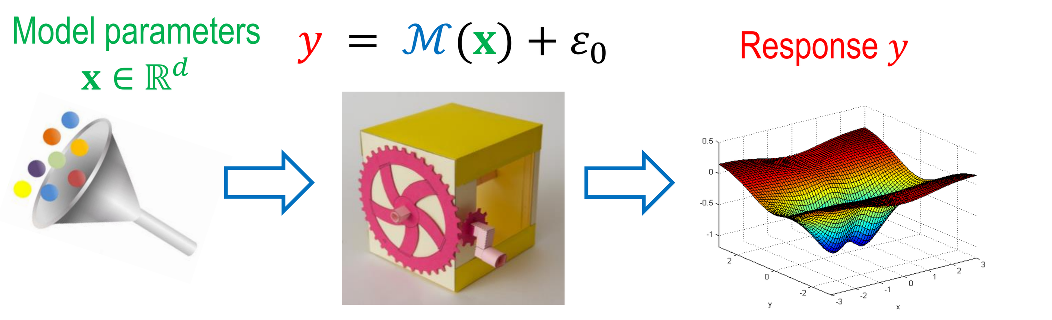
Bridging the Gap Between Artificial Neural Networks and Kernel Regressions in EM Applications

Nastaran Soleimani

Supervisors: Prof. Riccardo Trinchero & Prof. Flavio Canavero

Research context and motivation

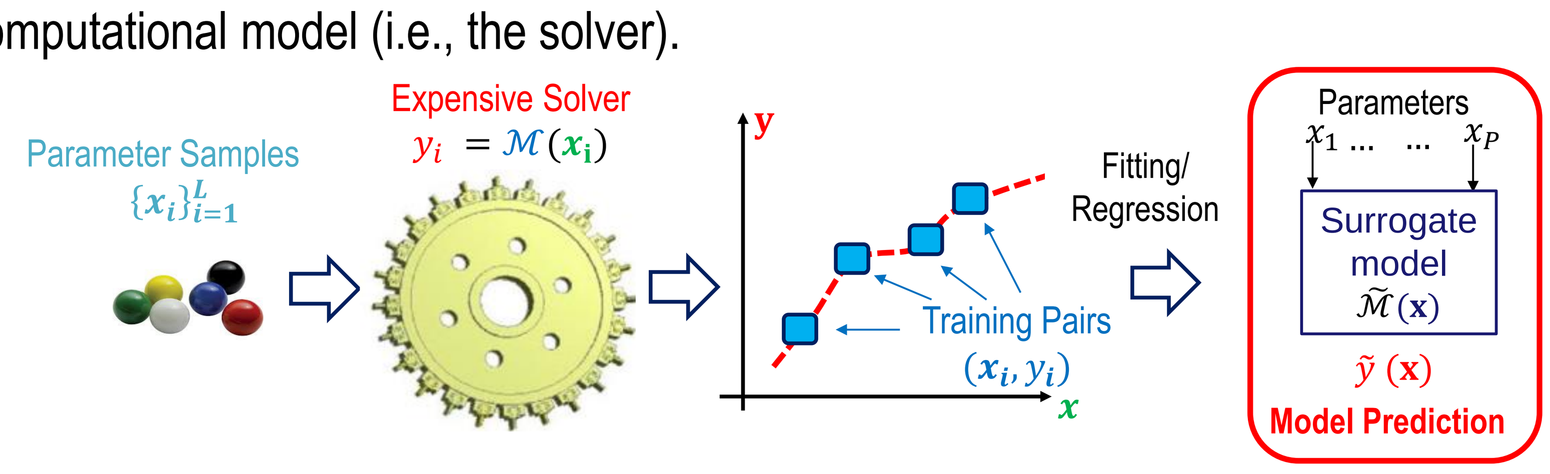
- **Optimization** and **uncertainty quantification** are key ingredients for the design of microwave structures and electronic devices.
- Such tasks are usually carried out synthetically via computer experiments (simulations), based on the **computational model.**
- **Computational model** is a procedure (e.g., a solver) able to compute quantities of interest from the input parameters (e.g., geometrical/electrical parameters).



WARNING
The **computational cost** of the computational model can be **huge!!!**

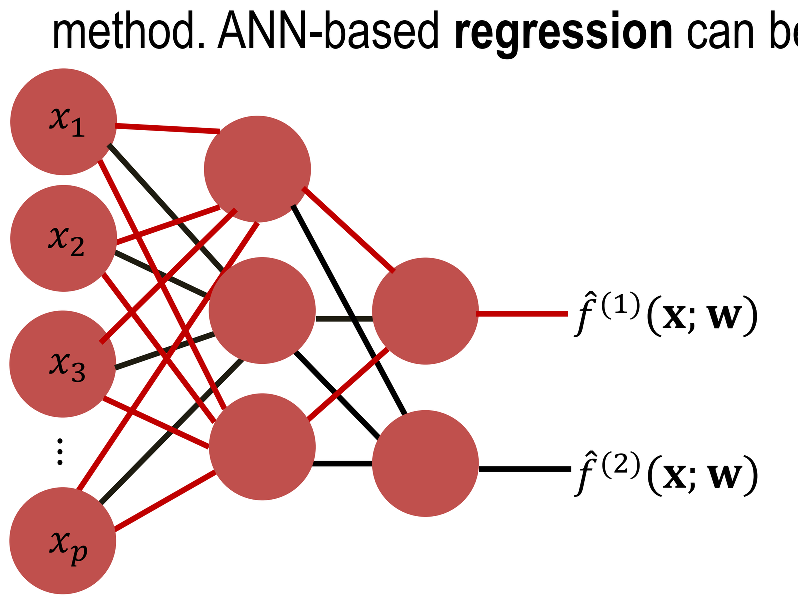
Addressed research questions/problems

- **Surrogate model** $\widetilde{\mathcal{M}}$ is "a model of a model", i.e., a **fast-to-evaluate model** of the computational model (i.e., the solver).



- In complex non-linear problem with dozen input parameters, the **accuracy** of the surrogate model depends on the fitting or regression techniques used to train it.

➢ Nowadays **Artificial Neural Network (ANN)** is the most popular Machine Learning method. ANN-based **regression** can be adopted to build accurate surrogate model.

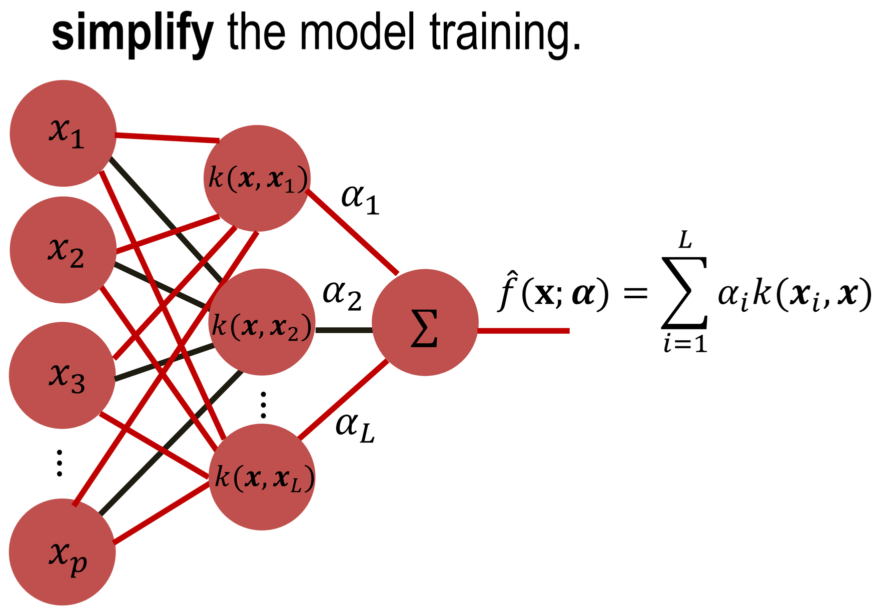


PROS:
- NON linear model structure
- Flexible topology
- Natural extension to multi-output

CONS:
- **NON-convex optimization problem**

→ Hard to train, data-hungry

➢ **Kernel regression** provides a clever alternative to ANN structure allowing to **heavily simplify** the model training.

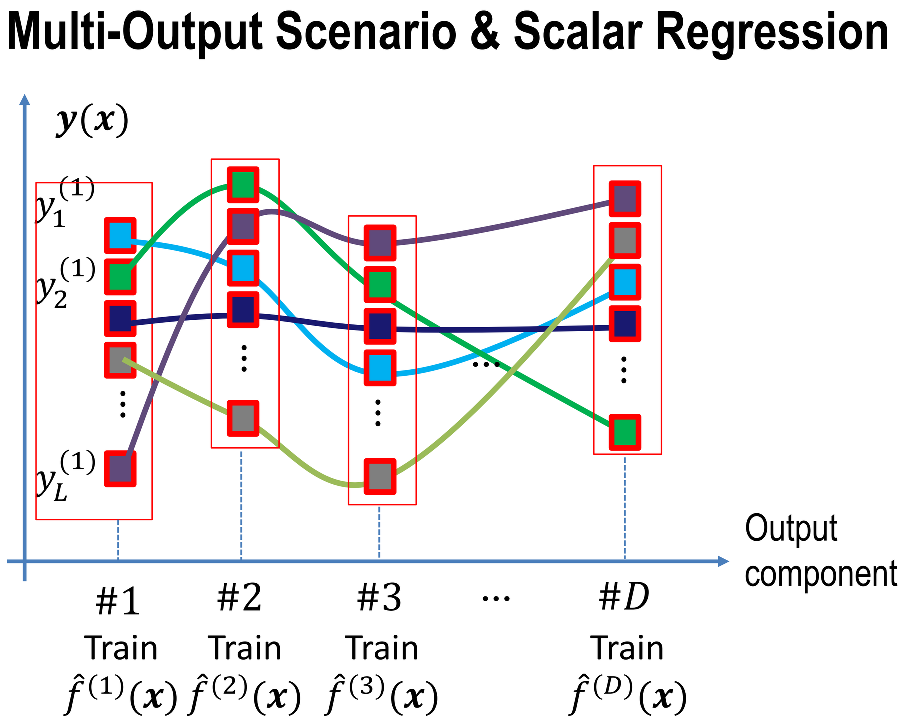


PROS:
- Linear model structure
- Convex optimization problem
  - Fast to train
  - Fast convergence w.r.t. training samples

CONS:
- Fixed topology
- **Scalar-valued methods**

WARNING: Most of the EM applications require a MULTI-OUTPUT formulation!!!

- **Multi-Output Scenario & Scalar Regression**

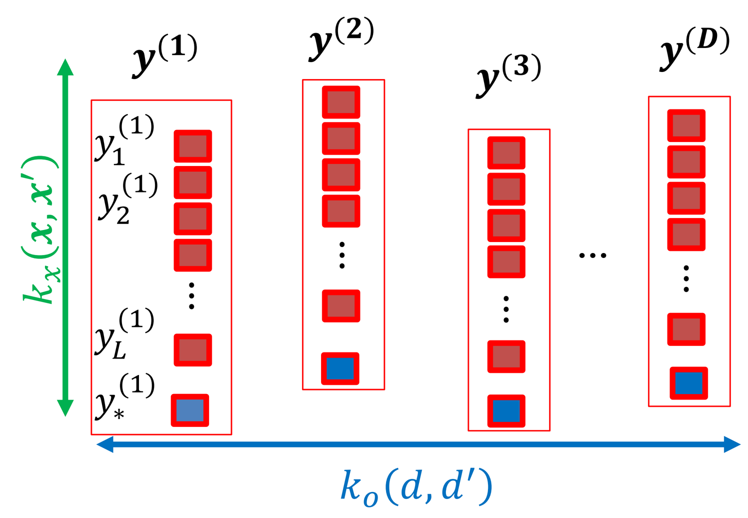


IDEA:
Use a scalar regression for each output components

Too many models and hyperparameters to tune!!! No protection against noise!!!

Methodology/Novel contributions

- **Multi-Output Kernel Ridge Regression:**



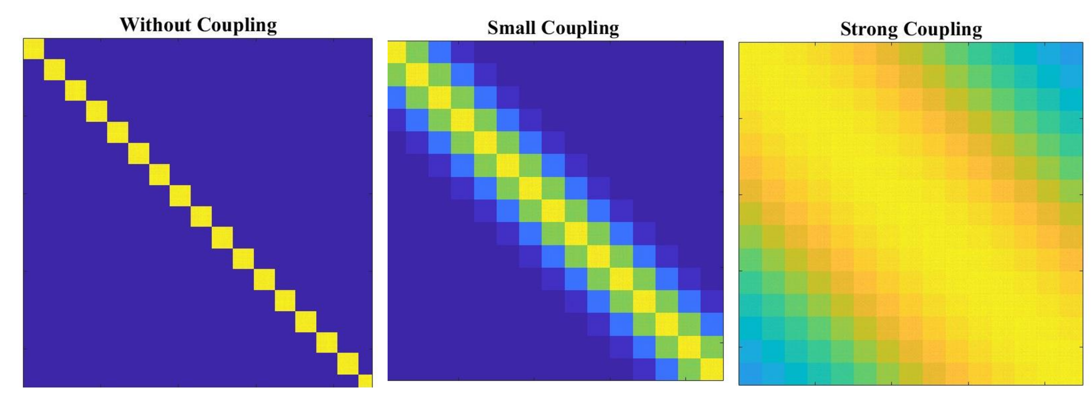
$$\hat{\mathbf{f}} = \underset{\bar{\mathbf{f}} \in \mathcal{H}}{\operatorname{argmin}} \sum_{d=1}^{D} \sum_{l=1}^{L} \left(y_l^{(d)} - \bar{f}^{(d)}(\mathbf{x}_l) \right)^2 + \lambda \|\bar{\mathbf{f}}\|^2$$

$$\hat{f}^{(d')}(\mathbf{x}_*) = \sum_{d=1}^{D} \sum_{i=1}^{L} k((\mathbf{x}_i, d), (\mathbf{x}_*, d')) c_{i,d}$$

where $k((\mathbf{x}_i, d), (\mathbf{x}_*, d'))$ is a **"new" kernel function** acting on both the input space and output components

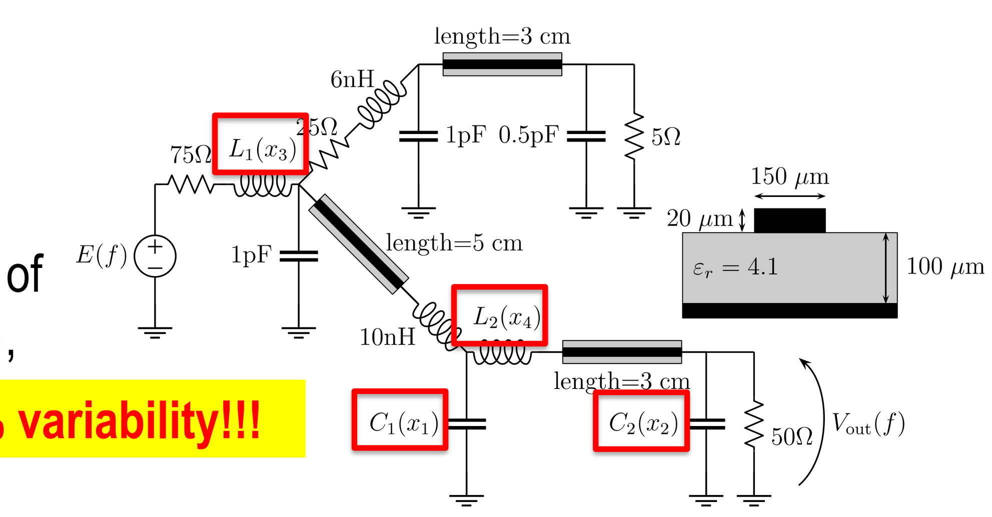
E.g., for a **separable kernel**:

$$k((\mathbf{x}, d), (\mathbf{x}', d')) = \\ = k_x(\mathbf{x}, \mathbf{x}') \cdot k_o(d, d')$$



Application Results

- **Example I**: High Speed Link

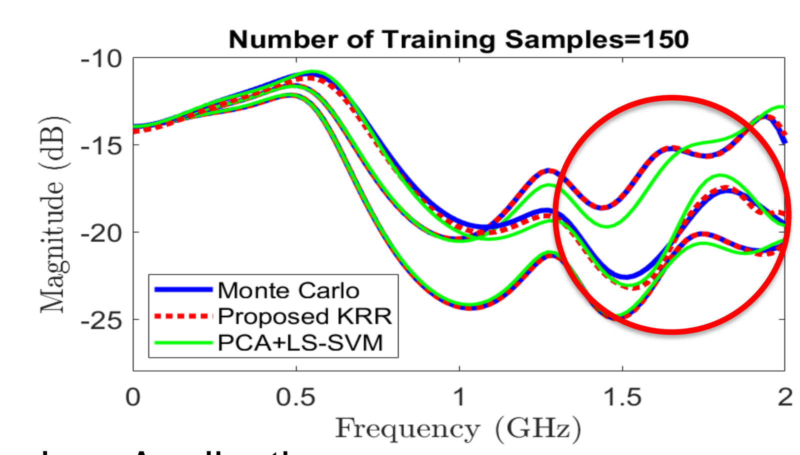


GOAL: Build a surrogate model for the **transfer function** $|y(f; \mathbf{x})|$ as function of the **circuit parameters** $C_1(x_1)$, $C_2(x_2)$, $L_1(x_3)$ and $L_2(x_4)$.

50% variability!!!

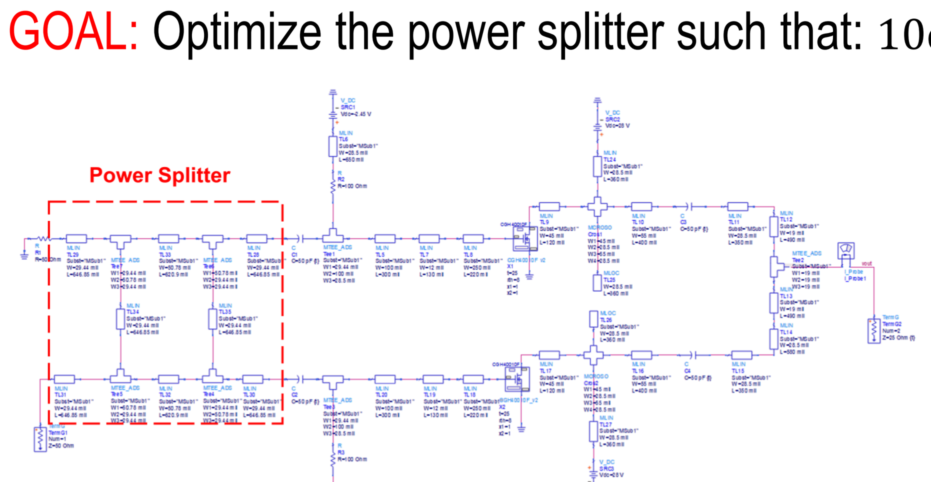
RESULT: Comparison among the **proposed multi-output kernel regression**, **ANN** and **PCA+LS-SVM** on 1000 test samples

Method	MSE $L = 50$	MSE $L = 100$	MSE $L = 150$
NN	8.01×10^{-4}	8.43×10^{-4}	7.12×10^{-4}
PCA+LS-SVM	2.73×10^{-5}	2.22×10^{-5}	2.13×10^{-5}
Proposed KRR	$\mathbf{1.45 \times 10^{-5}}$	$\mathbf{3.56 \times 10^{-6}}$	$\mathbf{2.02 \times 10^{-6}}$



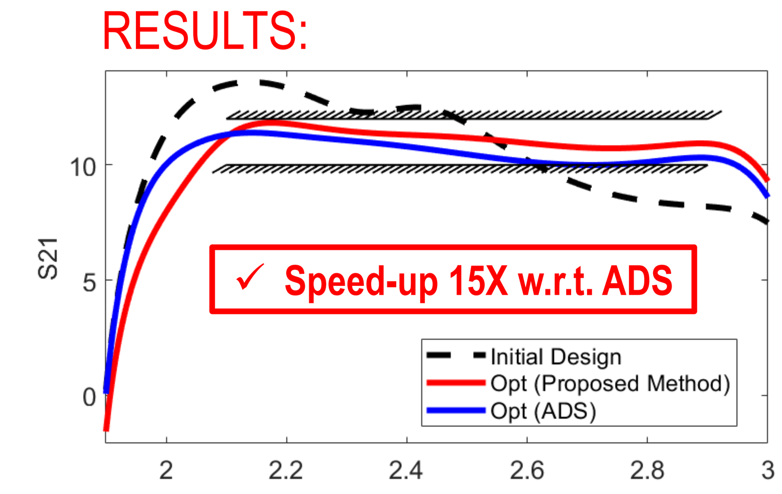
- **Example II** : Doherty Power Amplifier for Wireless Applications

GOAL: Optimize the power splitter such that: $10\text{dB} \leq S_{21}(f) \leq 12\text{dB}$ for $f \in [2.1, 2.9]$GHz



14 parameters

RESULTS:



✓ **Speed-up 15X w.r.t. ADS**

Future work

- How can we reduce the computational cost of the training phase?
- How can we automatically obtain the optimal structure of the multi-output kernel?

Submitted and published works

- N.Soleimani, R. Trinchero, F. Canavero, "*Bringing the Gap Between Neural Networks and Kernel Regressions for Vector-Valued Problems in Microwave Applications*", submitted to IEEE Transactions on Microwave Theory and Techniques, 2022.
- S. Kushwaha, N. Soleimani, et al, ''*Comparative Analysis of Prior Knowledge Based Machine Learning Metamodels for Modeling Hybrid Copper-Graphene On-Chip Interconnects*,'' IEEE Transactions on Electromagnetic Compatibility, 2022.
- N. Soleimani, R. Trinchero and F. Canavero, ''*Vector-Valued Kernel Ridge Regression for the Modeling of High-Speed Links*'', in Proc. IEEE MTT-S International Conference on Numerical Electromagnetic and Multiphysics Modeling and Optimization (NEMO2022}, Limoges (France), 2022. **BEST STUDENT PAPER AWARD**
- N. Soleimani and R. Trinchero, "*Compressed complex-valued least squares support vector machine regression for modeling of the frequency-domain responses of electromagnetic structures*," Electronics, vol. 11, no.4, 2022.
- M. Ahmadi, A. Sharifi, M. Jafarian Fard and N. Soleimani, ''*Detection of brain lesion location in MRI images using convolutional neural network and robust PCA*,'' International journal of neuroscience, 1-12, 2021.

POLITECNICO DI TORINO

PhD program in
Electrical, Electronics and Communications Engineering