KOOPMAN MEETS KALMAN: A DEEP LEARNING APPROACH TO MODEL RF CAVITY DETUNING

Andrei Maalberg, Andriy Ushakov, Pablo Echevarria, Axel Neumann, Jens Knobloch

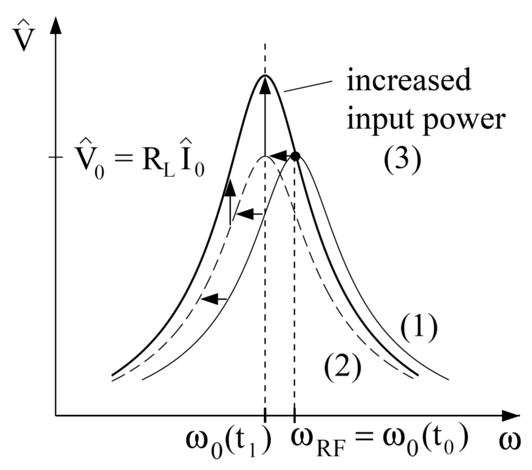
Institute for Science and Technology of Accelerating Systems

ABSTRACT: In the light of recent developments to employ high-Q superconducting cavities to reduce the energy consumption of particle accelerators, the problem of minimizing cavity detuning becomes highly relevant. Meanwhile, RF cavities are known for their non-stationary behavior, so finding a proper modeling approach is crucial for any model-based detuning control. In this contribution, we present a data-driven modeling framework that combines two complimentary perspectives on dynamical systems: the Koopman operator approach, which captures global patterns in cavity behavior, and Kalman-inspired ideas, which enable local, adaptive adjustments in response to changing cavity conditions. Following this, our architecture 1) separates long-term structure from local variability and then 2) blends them using a tunable weighting mechanism. We demonstrate the effectiveness of this approach on both synthetic cavity data and real-world cavity measurements. The results show a potential to make the detuning control more robust to the non-stationary cavity behavior.

MOTIVATION

Detuning of an RF cavity from a physics perspective

- The change of resonance frequency results in a decreasing amplitude at the operating frequency [1].
- The decreasing amplitude can be compensated by increasing input RF power, but this is highly inefficient in terms of energy consumption.



Principle of cavity detuning. Adapted from [1].

Detuning modeling and control

- Despite numerous works detuning compensation remains an open issue [2].
- There must be a new algorithm for detuning modeling that is
  - ✓ robust to time-varying cavity behavior,
  - ✓ able to learn cavity dynamics online,
  - ✓ interpretable to avoid black box modeling.

METHODS

Linear prediction in terms of a Koopman operator

Consider one-dimensional time series $\{x_t\}_{t=1}^{T}$ that represent an observable state of a discrete-time detuning process

$$x_{t+1} = f(x), x \in \mathcal{X} \subset \mathbb{R},$$

where x_t is a state at time t and $\mathcal{X}$ is a set of all physically realizable real-valued states. Even though the underlying equations of f are well-known, the direct modeling of f is often intractable due to ist nonlinear and time-varying nature.

Koopman operator theory offers an alternative: we lift the system into a higher-dimensional observal space, where the dynamics evolve linearly

$$\mathcal{K}g(x_t) = g(x_{t+1}) = g\big(f(x_t)\big),$$

where g is a set of observable functions $g : \mathcal{X} \to \mathbb{R}$, and where $\mathcal{K}$ is an approximation of an infinite-dimensional Koopman operator.

Kalman-inspired blending of dynamics prediction

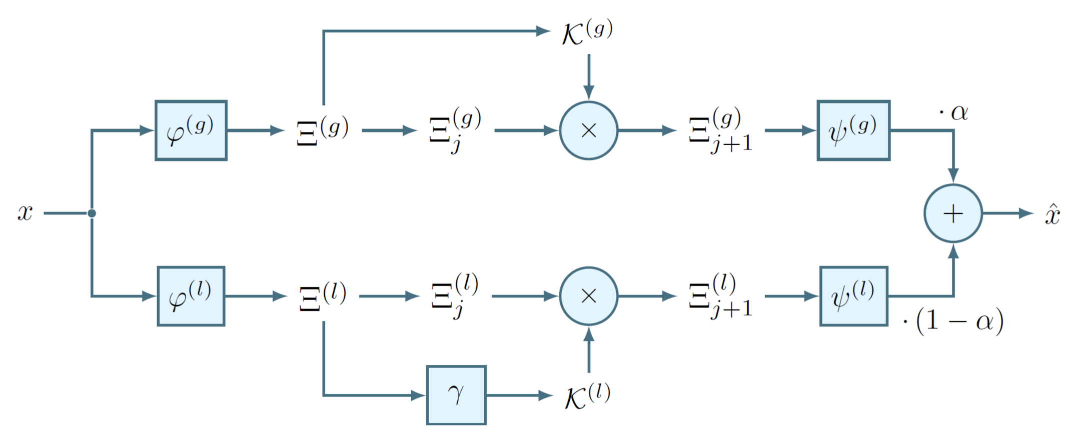
Consider a blend of stationary and transient dynamics

$$x_{t+1} = \alpha \cdot x_t^{DMD} + (1-\alpha) \cdot x_t^{Transformer}, \alpha \in [0,1],$$

where x_t^{DMD} and $x_t^{Transformer}$ are predictions for stationary and transient dynamics, and where α and x_{t+1} are a tunable blending weight and a blended prediction, respectively. Then let $\alpha = 1 - K_t$ and rearrange terms accordingly

$$x_{t+1} = x_t^{DMD} + K_t \cdot \left(x_t^{Transformer} - x_t^{DMD}\right),$$

which this is a classical Kalman update, so we can interpret α as a Kalman gain. In a neural network implementation, this blending can look as follows



Neural network architecture for our model.

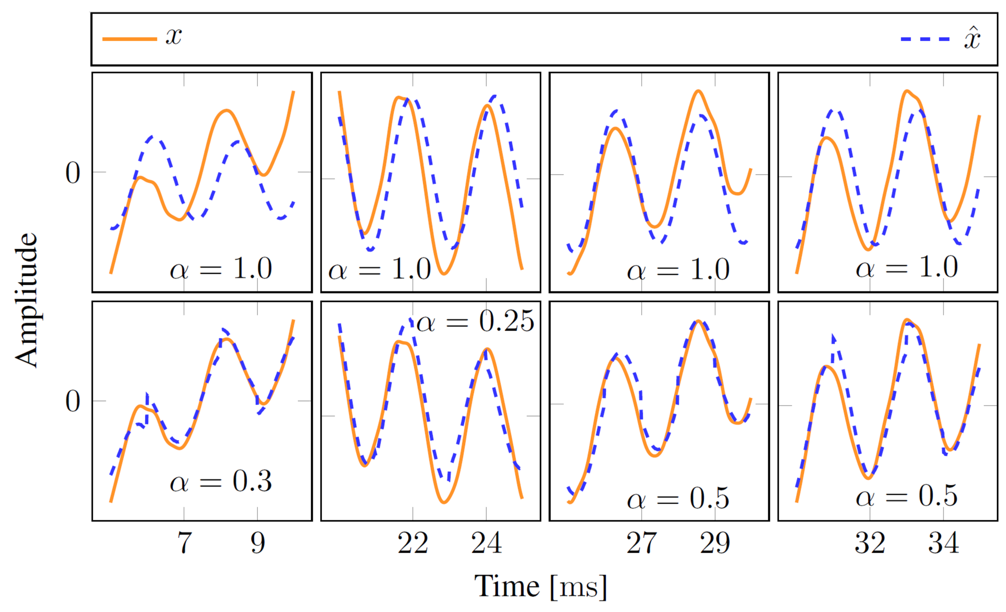
Real-world cavity measurement

	Frequency	Bandwidth	Quality factor
Pillbox cavity	1.3 GHz	708 kHz	1.8 kU

RESULTS

Validation of Kalman-inspired blending

As a proof-of-concept we set α manually when testing our principle on transient time series. As shown below, the DMD, which was trained on stationary data, cannot handle transients. In contrast, the Transformer demonstrates its capability to capture transient behavior. Hence, blending the two allows to improve the overall prediction.



Comparison of model predictions on transitional real-world time series. Upper row: DMD only. Bottom row: DMD blended with Transformer.

CONCLUSION

- We proposed a novel architecture that blends a global, DMD-like, Koopman operator with a local, attention-driven operator, inspired by the structure of a Kalman filter.
- This novel approach aims to increase the interpretability of a machine learning-based solution, which is crucial in control-oriented applications.
- We demonstrated the viability of this approach on real-world RF cavity data.
- This work opens promising directions for future research, including the development of a principled mechanism for dynamic operator weighting in place of manual tuning.

REFERENCES

[1] T. Schilcher. Vector Sum Control of Pulsed Accelerating Fields in Lorentz Force Detuned Superconducting Cavities. Ph. D. dissertation, Hamburg University, 1998.

[2] A. Bellandi. LLRF Control Techniques for the European XFEL Continuous Wave Upgrade. Ph. D. dissertation, Hamburg University, 2021.

ACKNOWLEDGEMENT AND PARTNERS

This work receives funding from European Commission's Horizon Europe Research and Innovation programme under Grant Agreement n°101131435.

MORE INFORMATION

https://isas.ijclab.in2p3.fr

Dr. Andrei Maalberg

andrei.maalberg@helmholtz-berlin.de