Department of Mathematical Sciences
Faculty and Graduate Research Colloquium

Thursday, November 14, 2024
1:00 – 1:50 PM
RB 452

"Injectivity and Stability of ReLU-layers: Perspectives from Frame Theory"
Daniel Haider (University of Vienna)

Abstract:
Non-linearities frequently arise in applications, either due to technical constraints or as intentional elements of model design. A prominent instance of the latter is to use the composition of affine linear mappings and non-linear activation functions as layers of artificial neural networks. Among many layer designs, ReLU-layers - i.e., layers using ReLU}(t) = max(0,t) as activation - are the most widely used layer types due to their simplicity and effectiveness. By performing hard thresholding, the ReLU function naturally acts as a sparsifier, where a black-box machinery determines which and how much information of the input is suppressed. Assessing whether the original input can be reconstructed from the output is therefore crucial for improving the interpretability and functionality of the associated models. This makes the injectivity analysis of ReLU-layers both, a challenging and intriguing problem that has not yet been fully solved.

In this talk, we present a frame theoretic perspective to approach the problem. The main objective is to develop the most general characterization of the injectivity of ReLU-layers in terms of all three involved components: (i) the frame that determines the linear mapping, (ii) the bias that acts as affine offset, and (iii) the input domain where the data is drawn from. From (ii) we can derive a practical, numerically implementable methodology for studying information loss in ReLU-layers. Finally, for injective ReLU-layers, we derive explicit reconstruction formulas based on the duality concept from frame theory and establish novel stability bounds for the recovery map.

This is joint work with M. Ehler, D. Freeman, and P. Balazs.

Mini-bio:

Daniel Haider is a Ph.D. candidate at the University of Vienna in Austria and is conducting his project "Frame Theory meets Neural Networks" at the Acoustics Research Institute of the Austrian Academy of Sciences under the supervision of Peter Balazs. Daniel is very interested in working multidisciplinary across the fields of frame theory and harmonic analysis, audio signal processing, and machine learning. More specifically, he develops frame theoretic approaches to study the stability and invertibility of neural networks and proposes architecture designs and regularizers to ensure stable invertibility with a particular focus on audio applications. Currently, he is a research visitor at SLU, working with Daniel Freeman on non-linear frame expansions.