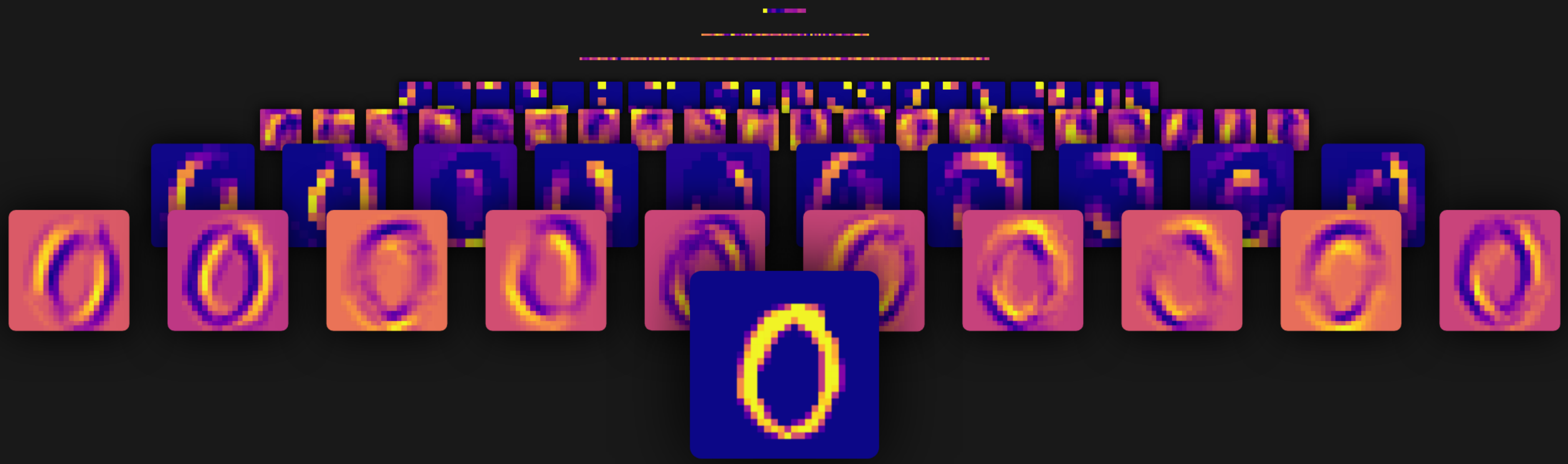
Fourier Analysis of Activations in Neural Networks

Bachelor-Thesis

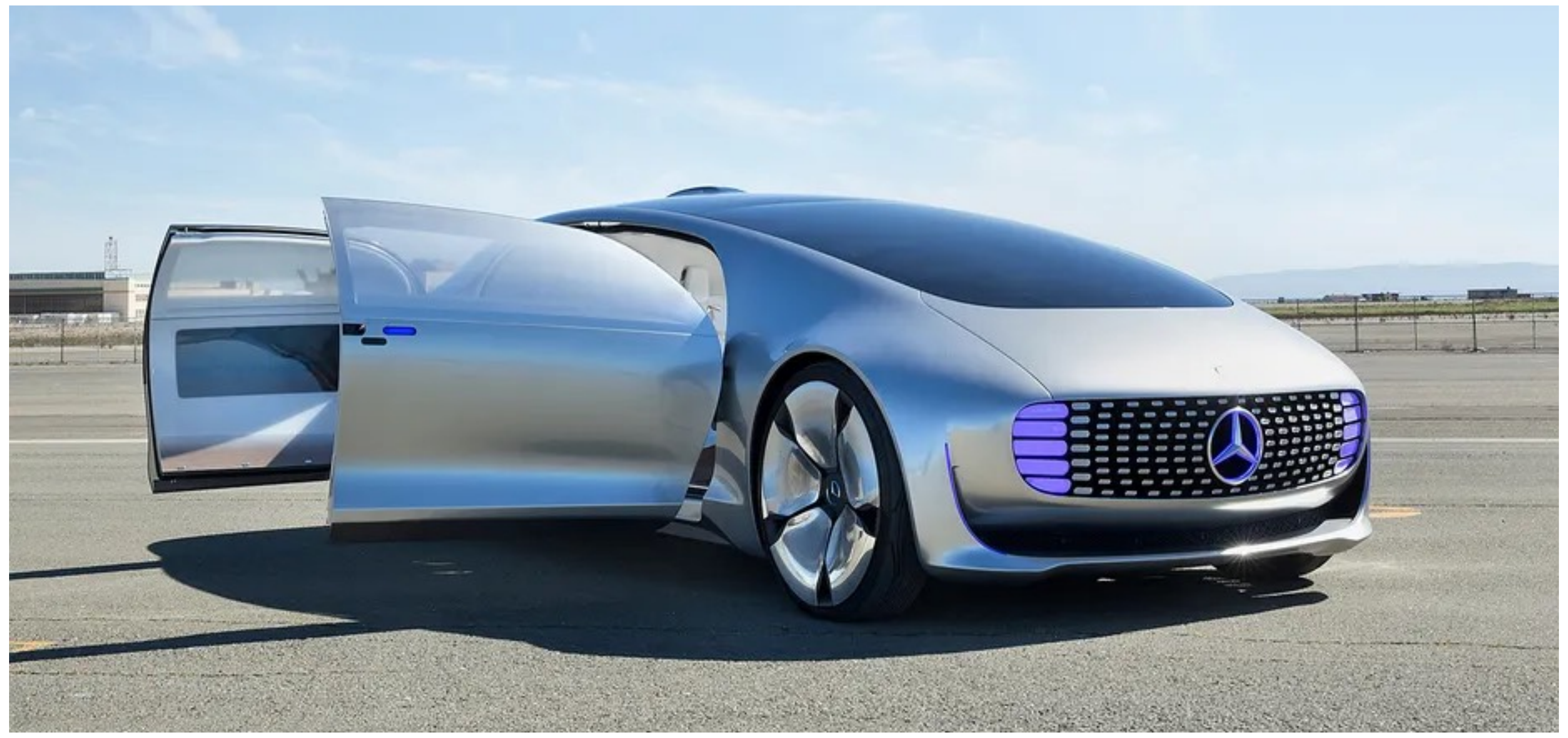


Felipa Schwarz, 2021

Analysis on Neural Networks

Motivation

Self-driving cars



https://www.wired.com/2017/04/mercedes-promises-self-driving-taxis-just-three-years/

Healthcare

https://www.medgadget.com/2019/08/healthcare-analytics-market-2019-industry-analysis-size-share-upcoming-trends-segmentation-forecast-to-2025-cagr-of-26-48.html

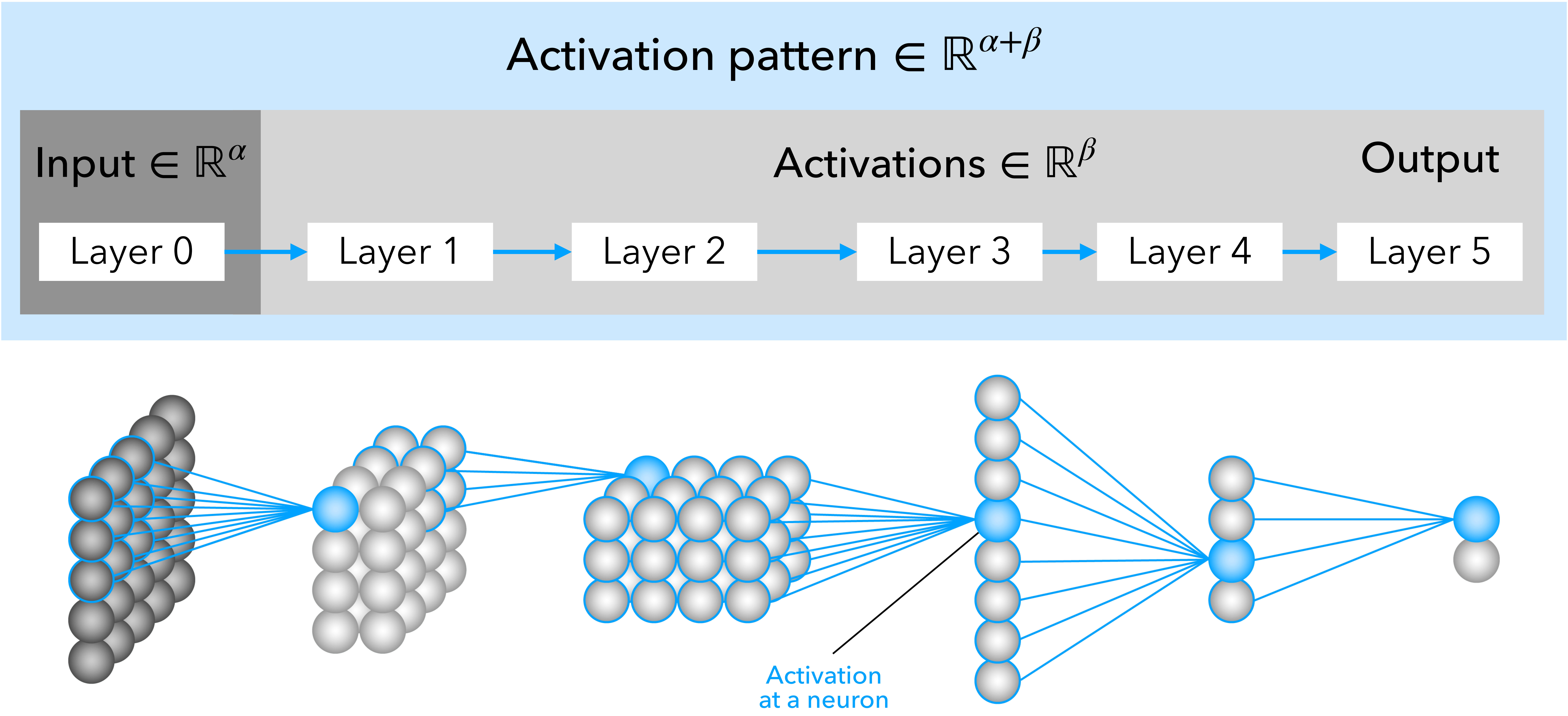
Cyber-security

https://www.news18.com/news/tech/hacker-group-darkside-suspected-of-carrying-cyber-attack-on-top-us-pipeline-operator-colonial-3723686.html

Ethics

https://todaysveterinarybusiness.com/practice-without-prejudice/

Neural Networks (NN)



A CAUSAL SHIFT AND FOURIER TRANSFORM ON DIRECTED ACYCLIC GRAPHS

Bastian Seifert, Chris Wendler, Markus Püschel

Department of Computer Science
ETH Zürich, Switzerland

ABSTRACT

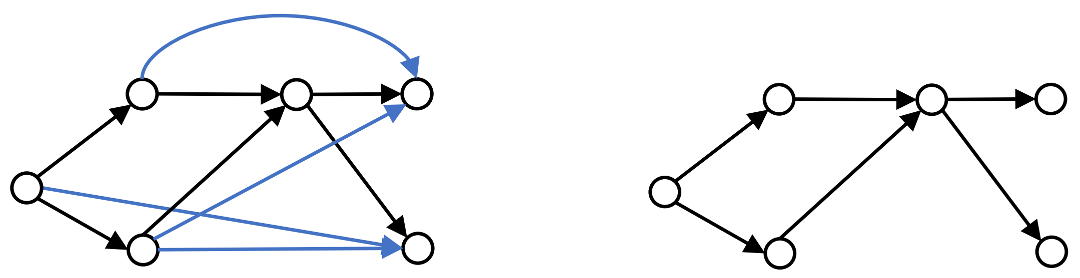
We introduce a notion of shift on signals indexed by the causal structure of a directed acyclic graph. We derive associated notions of convolution and causal Fourier transform. As application we learn Fourier-sparse signals on a dynamic contact tracing network, modeling the spread of an infectious disease. The proposed approach considerably outperforms similar approaches using graph signal processing.

Index Terms— directed acyclic graph, partially ordered set, algebraic signal processing, causality, infectious disease

1. INTRODUCTION

Causality is a recurring theme in science [1] and has recently gained considerable interest in machine learning [2–4]. The concept of causality is intimately connected to directed acyclic graphs (DAGs) that model the causal dependence between events. In this paper we are concerned with Fourier analysis of signals or data on DAGs.

Graph signal processing (GSP) [5, 6] is meanwhile well-



(a) General DAG D.

(b) Transitive reduction of D.

Fig. 1: The graph in (b) is the transitive reduction of the one in (a). The blue edges in (a) do not contain additional information about the partial order and thus can be removed.

GSP performs significantly worse.

2. DAGS AND POSETS

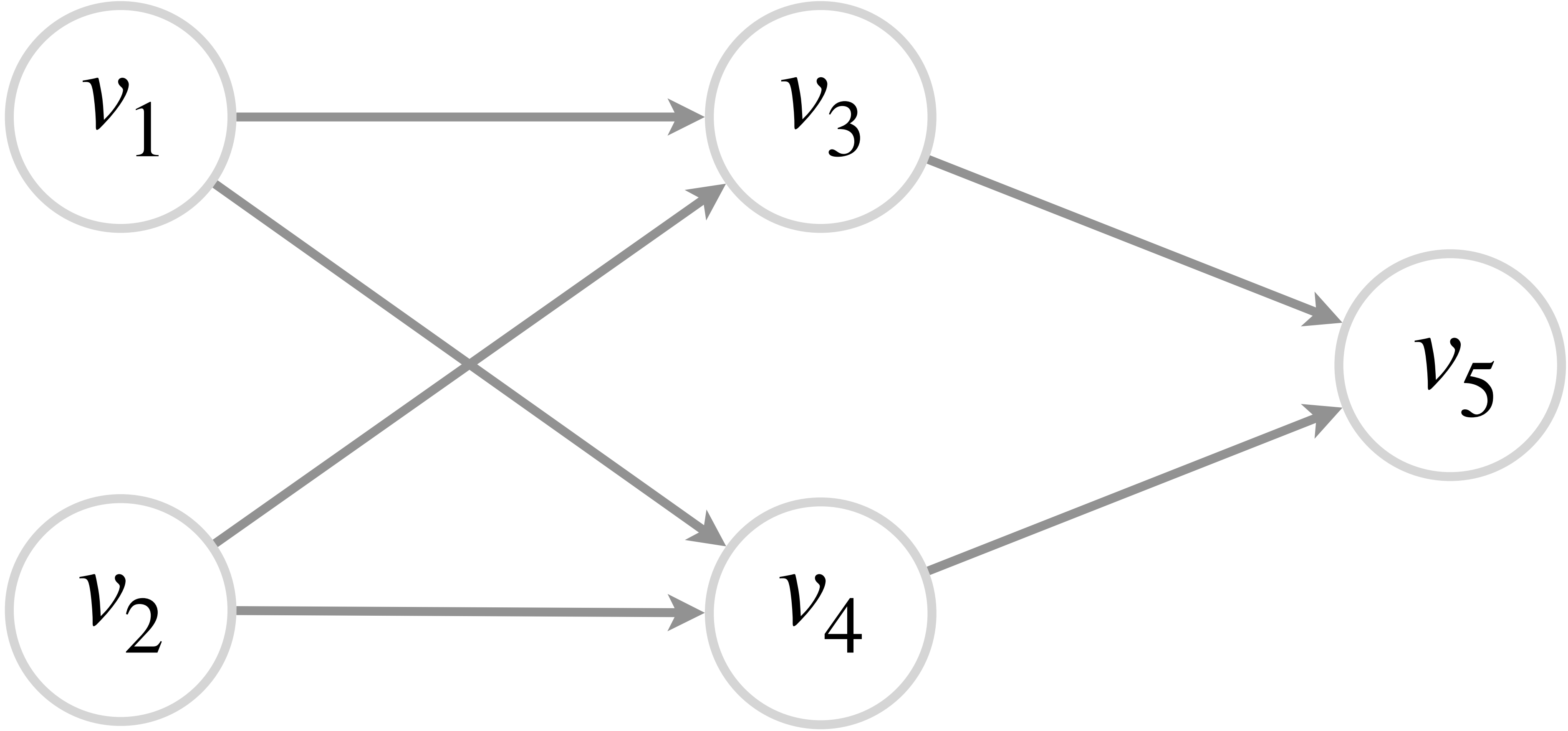
We provide the necessary background on DAGs and their relation to partially ordered sets.

DAGs. A directed graph $D = (V, E)$ consists of a set of nodes V and a set of directed edges $E = \{(v_i, v_j) \mid v_i, v_j \in$

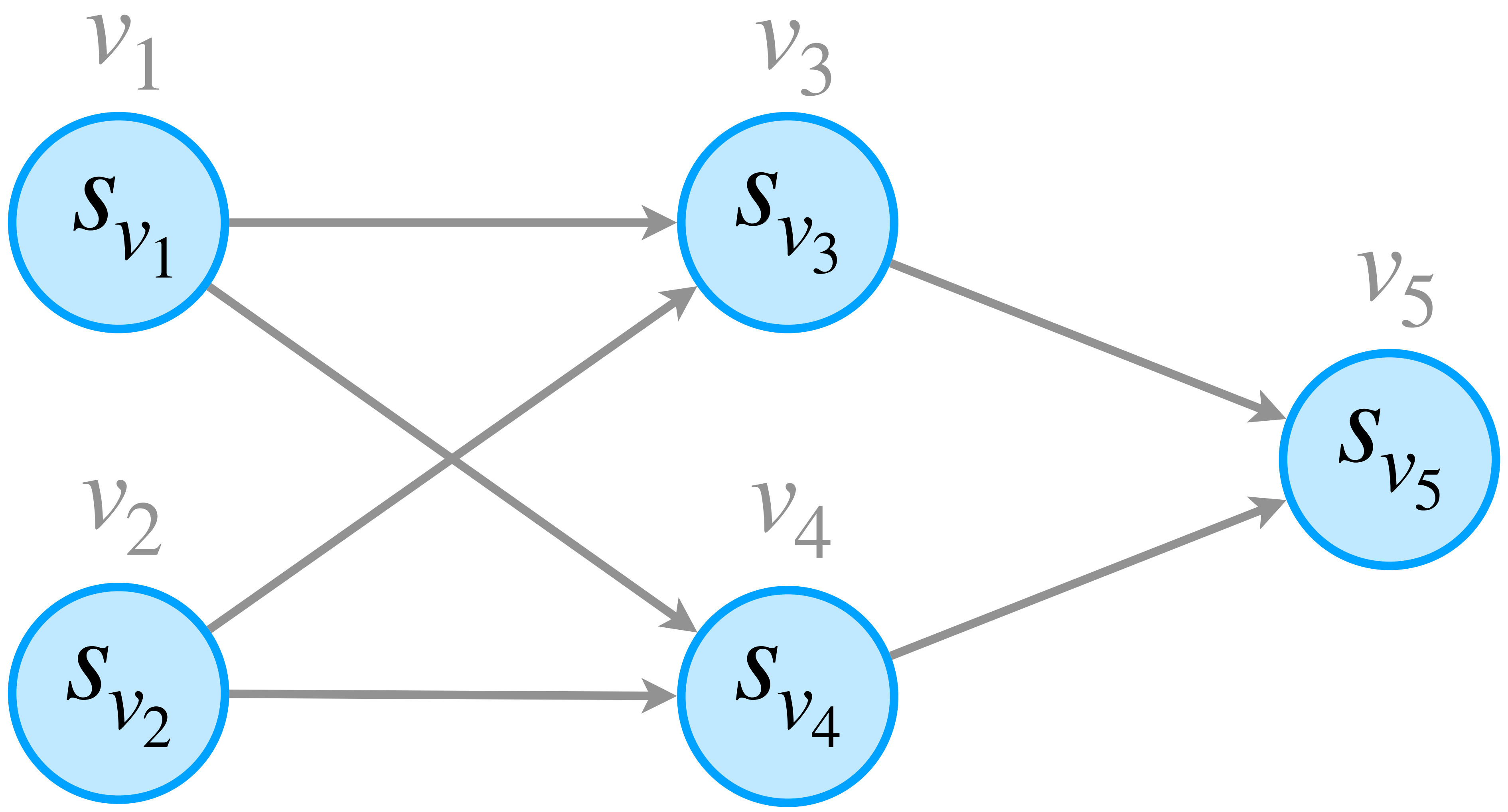
Causal Signal Processing (CSP)

Data-domain

Directed acyclic graph $G = (V, E)$



Signal $s \in \mathbb{R}^{|V|}$



Intuition: $$s_x = \boxed{\sum_{y \leq x,\ y \in V}} \hat{s}_y$$

Causes

Unobserved weight

$$s_{v_3} = \hat{s}_{v_1} + \hat{s}_{v_2} + \hat{s}_{v_3}$$

Causal Signal Processing (CSP)

Data-domain

Signal s

Zeta

$$s_x = \boxed{\sum_{y \le x,\ y \in V}} \hat{s}_y = \boxed{\sum_{y \in V} \zeta(x,y)} \hat{s}_y$$

$$s = \boxed{F^{-1}} \hat{s}$$

Inverse Fourier transform F^{-1}

$$\boxed{F^{-1}_{x,y} = \zeta(x,y) = \begin{cases} 1 & \text{if } y \le x \text{ ,} \\ 0 & \text{otherwise.} \end{cases}}$$

Fourier-domain

Fourier coefficients $\hat{s}$

Moebius

$$\hat{s}_x = \boxed{\sum_{y \in V} \mu(x,y)} s_y$$

$$\hat{s} = \boxed{F} s$$

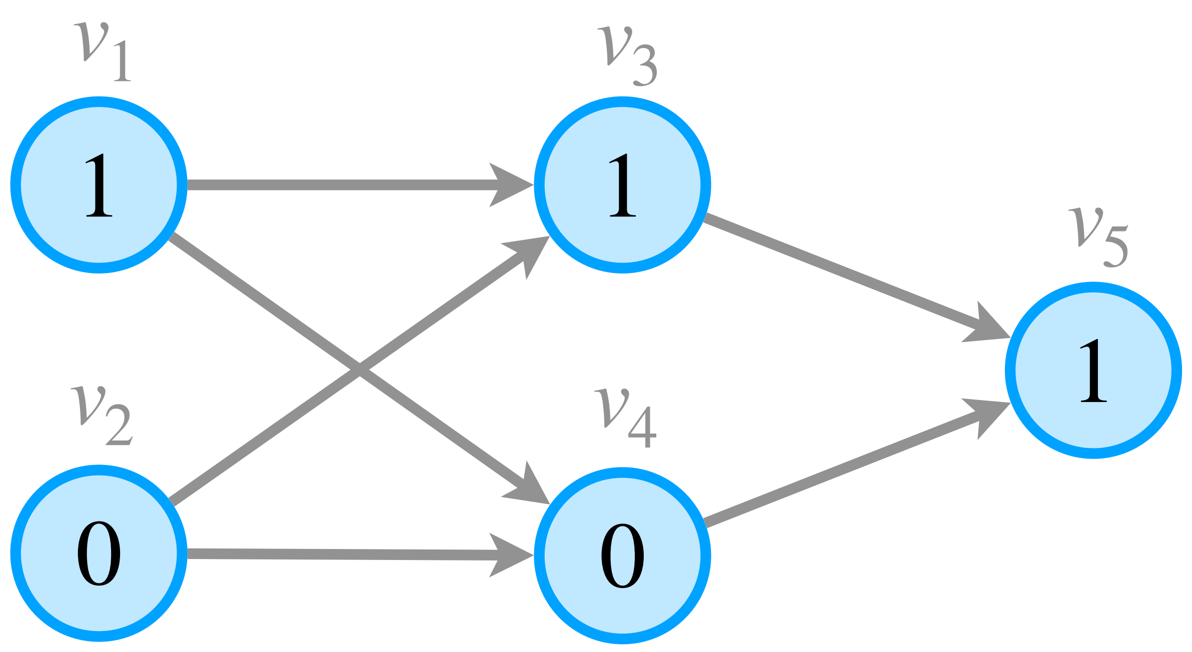
Fourier transform F

$$\boxed{F_{x,y} = \mu(y,x) = \begin{cases} 1 & \text{if } x = y \text{ ,} \\ -\sum_{y \le z < x} \mu(y,z) & \text{otherwise.} \end{cases}}$$

Causal Signal Processing (CSP)

Data-domain

Signal s

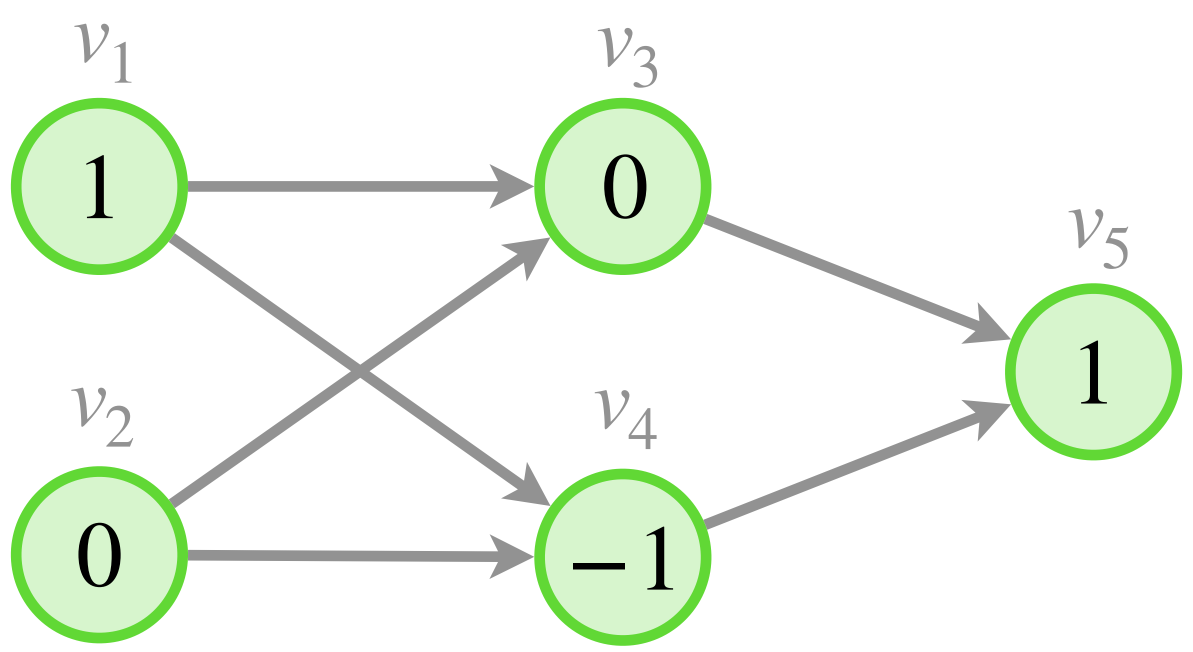


$$s = F^{-1}\,\hat{s}$$

$$\begin{bmatrix}1\\0\\1\\0\\1\end{bmatrix} = \begin{bmatrix}1&0&0&0&0\\0&1&0&0&0\\1&1&1&0&0\\1&1&0&1&0\\1&1&1&1&1\end{bmatrix}\begin{bmatrix}1\\0\\0\\-1\\1\end{bmatrix}\begin{array}{c}v_1\\v_2\\v_3\\v_4\\v_5\end{array}$$

Fourier-domain

Fourier coefficients $\hat{s}$



$$\hat{s} = F\,s$$

$$\begin{bmatrix}1\\0\\0\\-1\\1\end{bmatrix} = \begin{bmatrix}1&0&0&0&0\\0&1&0&0&0\\-1&-1&1&0&0\\-1&-1&0&1&0\\1&1&-1&-1&1\end{bmatrix}\begin{bmatrix}1\\0\\1\\0\\1\end{bmatrix}\begin{array}{c}v_1\\v_2\\v_3\\v_4\\v_5\end{array}$$

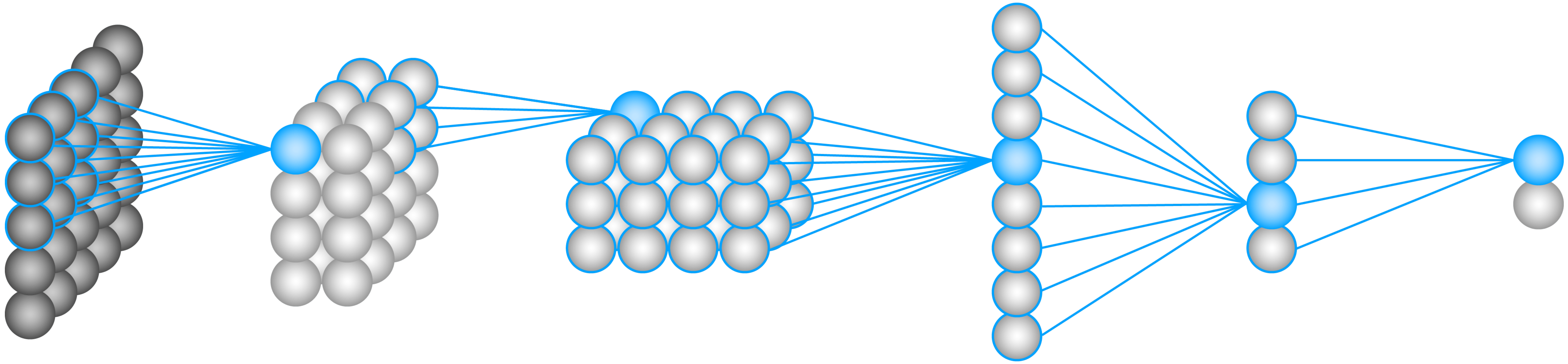
CSP on Neural Networks

Thesis Project

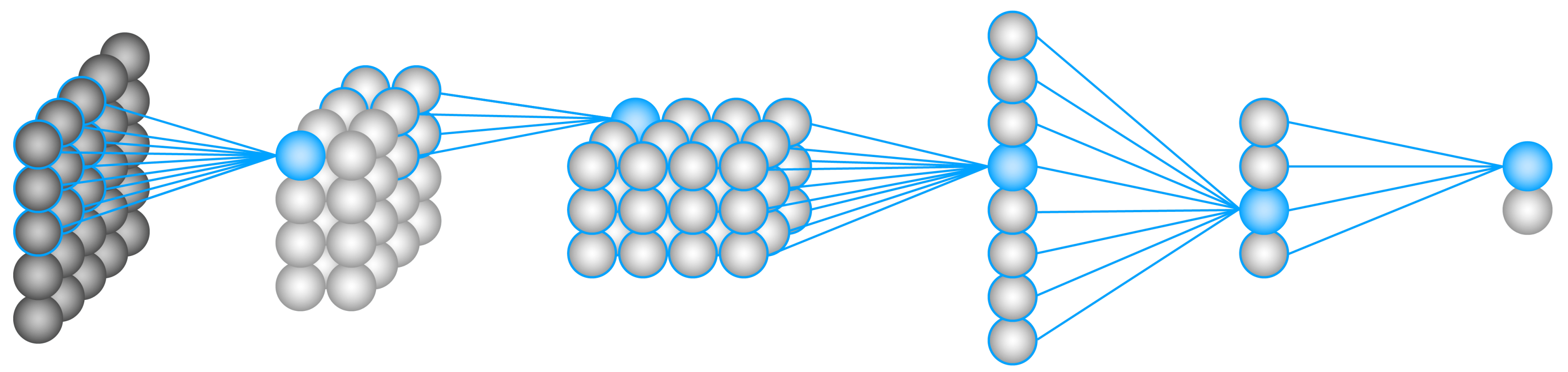
Causal Signal Processing

Neural Network $\Rightarrow$ DAG $G = (V, E)$ $\rightarrow$ Inverse Fourier Matrix F^{-1}

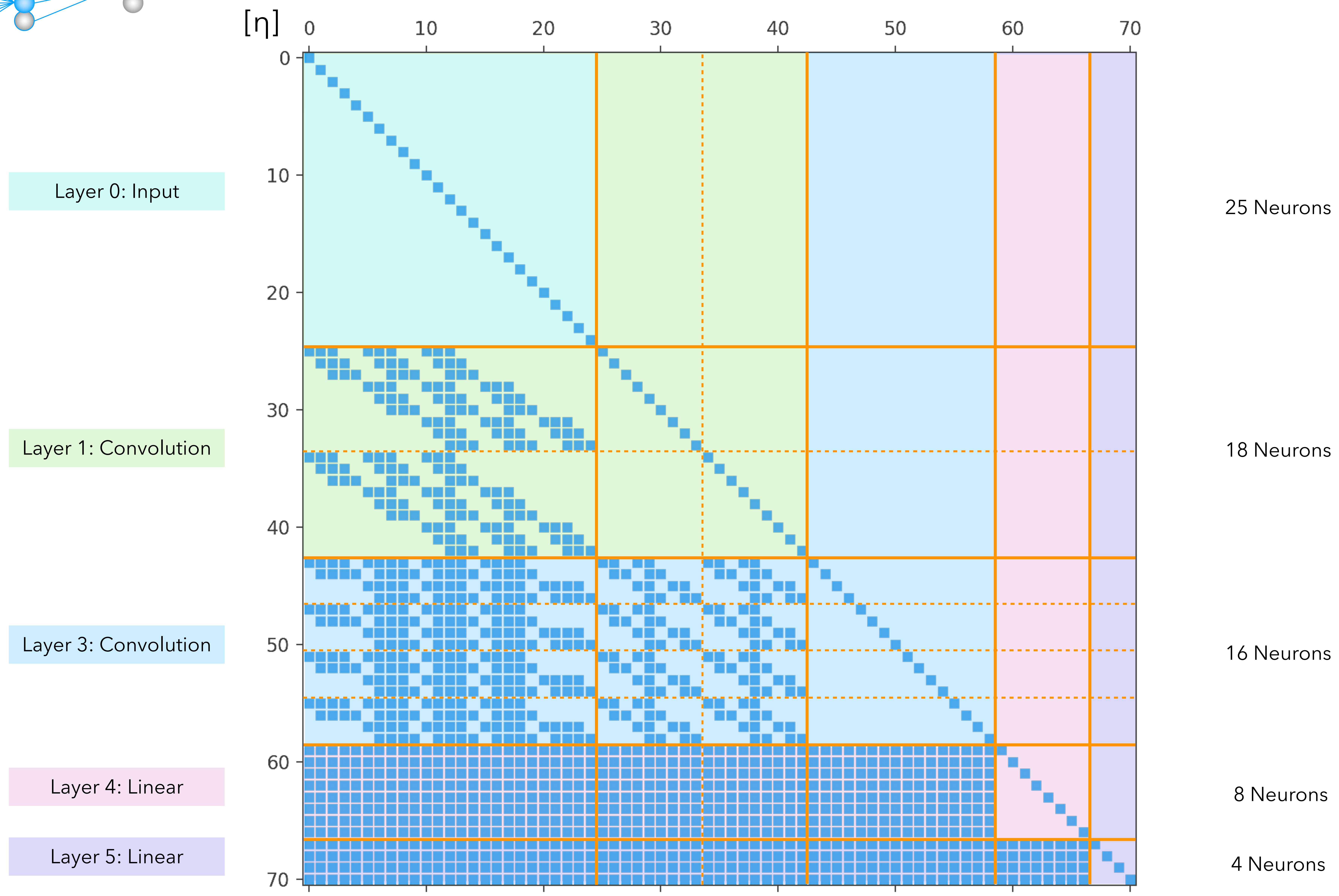
Activation pattern $\Rightarrow$ Signal s $\xrightarrow{\text{Solve: } F^{-1}\hat{s} = s}$ Fourier coefficients $\hat{s}$



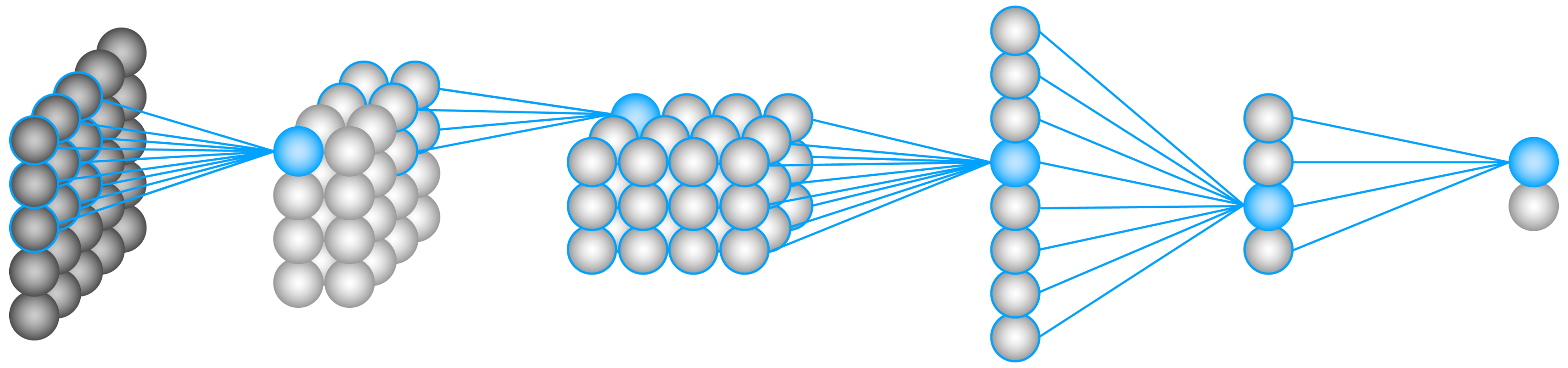
Inverse Fourier Matrix F^{-1}



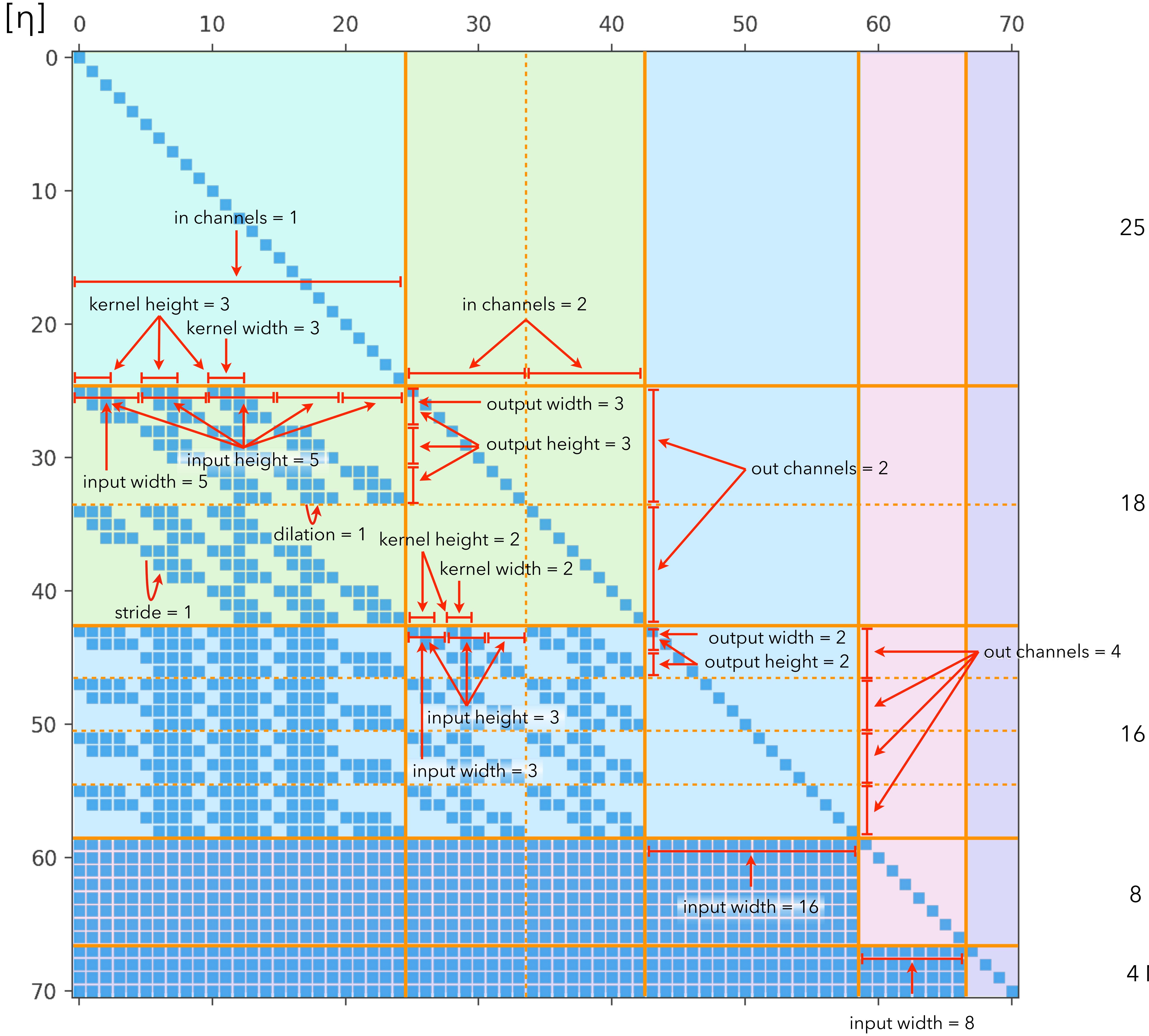
$$F^{-1}_{x,y} = \begin{cases} 1 & \text{if } y \leq x \\ 0 & \text{otherwise} \end{cases}$$



Inverse Fourier Matrix F^{-1}

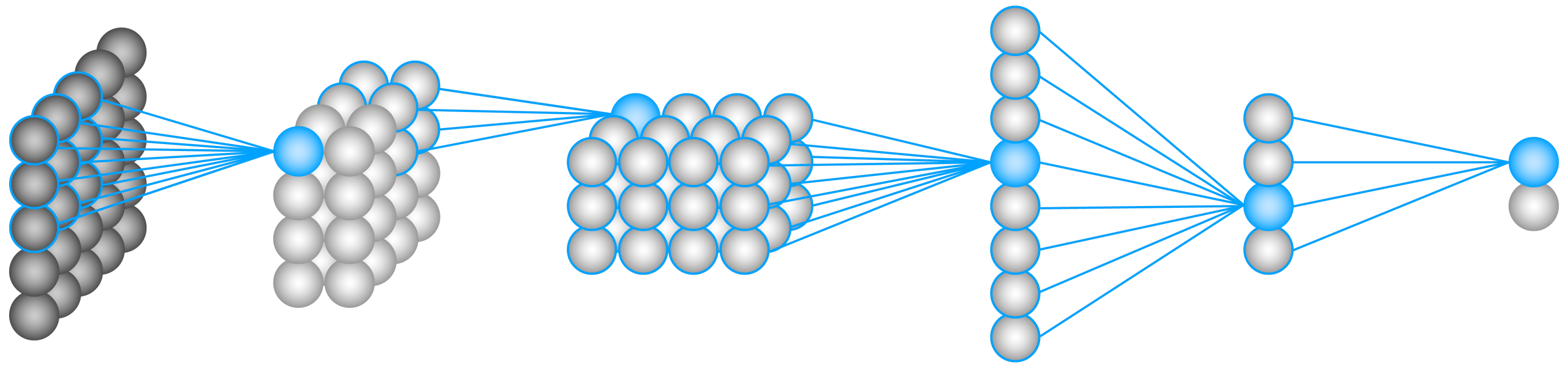


$$F^{-1}_{x,y} = \begin{cases} 1 & \text{if } y \leq x \\ 0 & \text{otherwise} \end{cases}$$

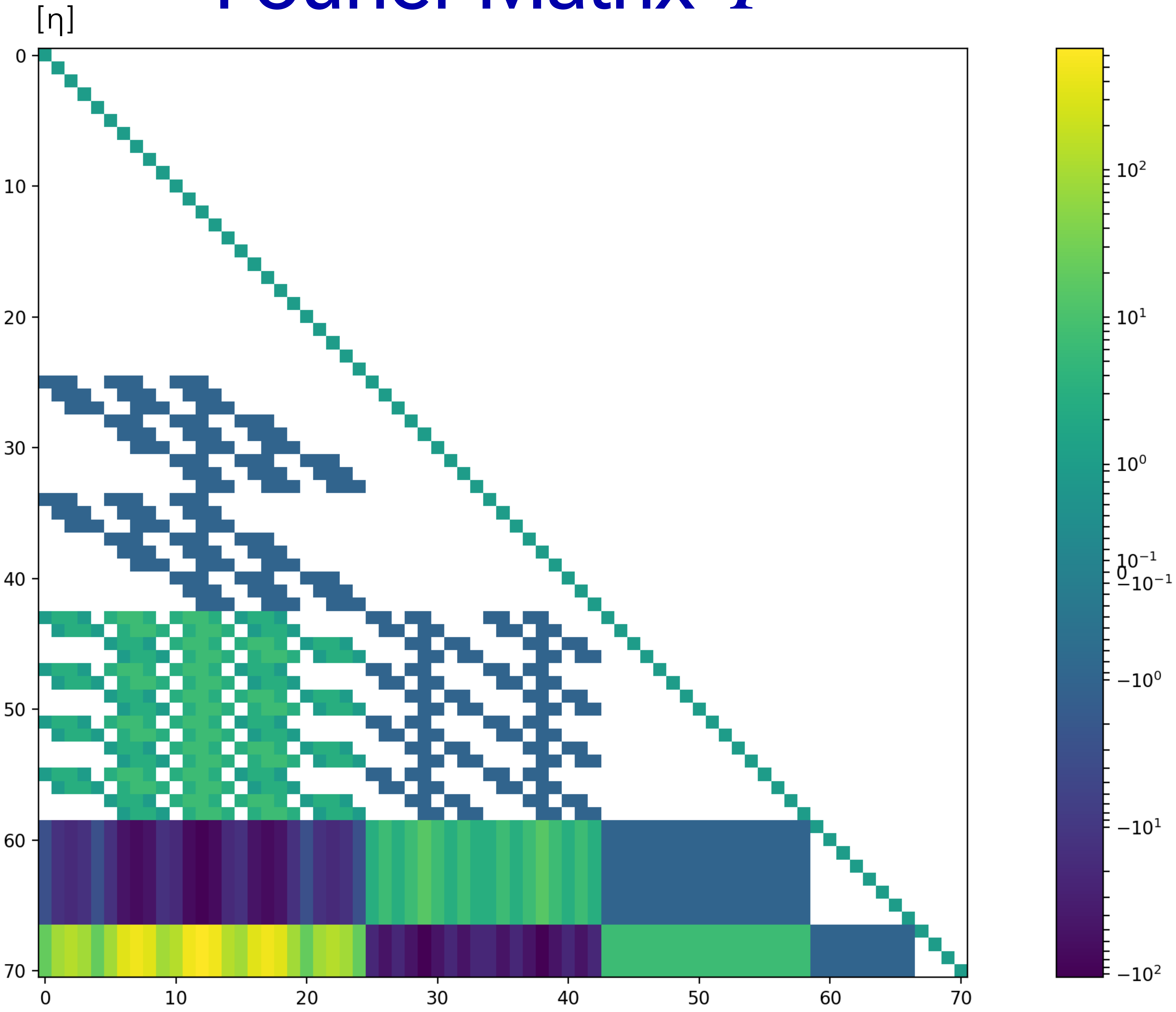


Layer 0: Input	25 Neurons
Layer 1: Convolution	18 Neurons
Layer 3: Convolution	16 Neurons
Layer 4: Linear	8 Neurons
Layer 5: Linear	4 Neurons

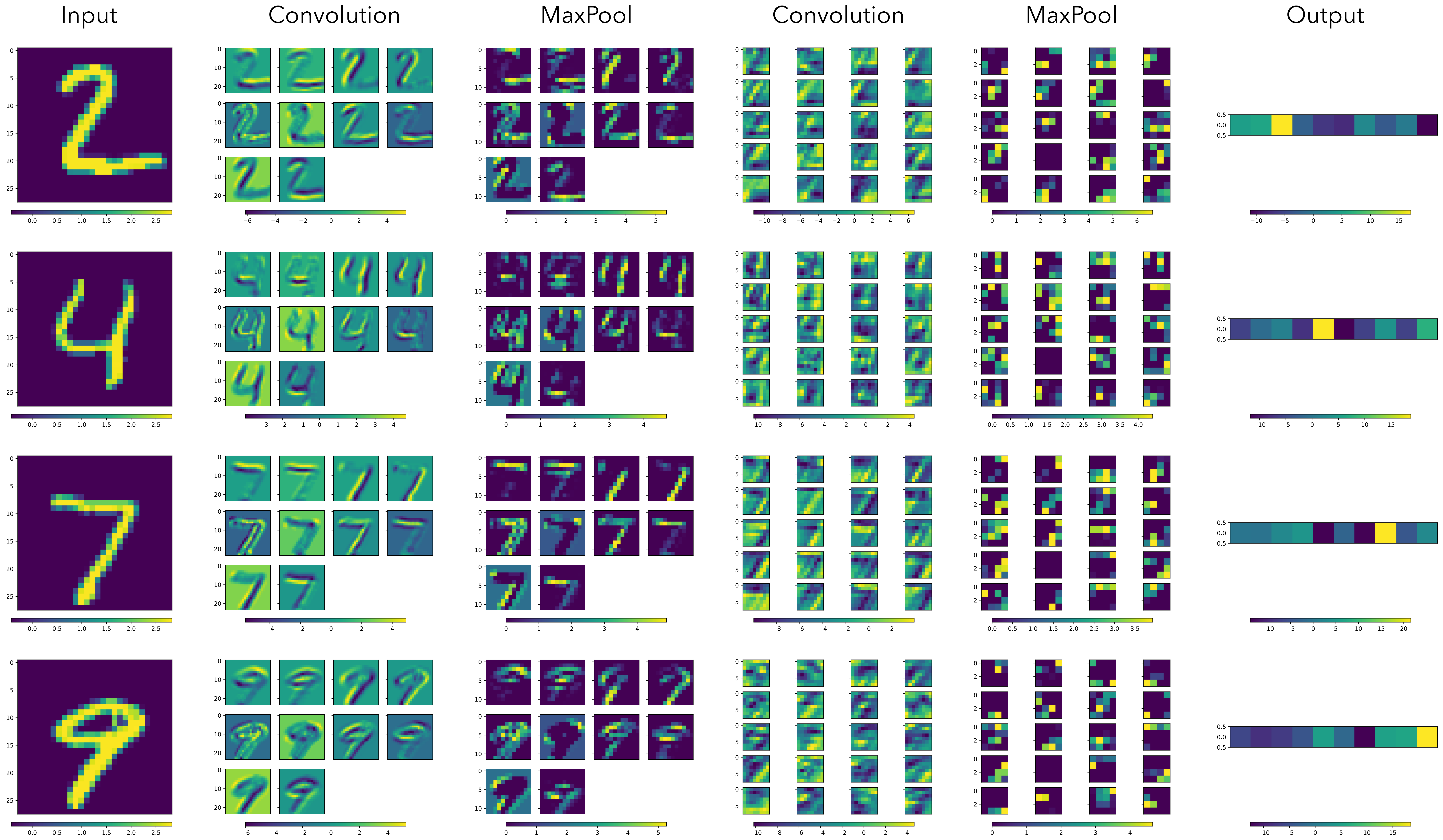
Fourier Matrix F



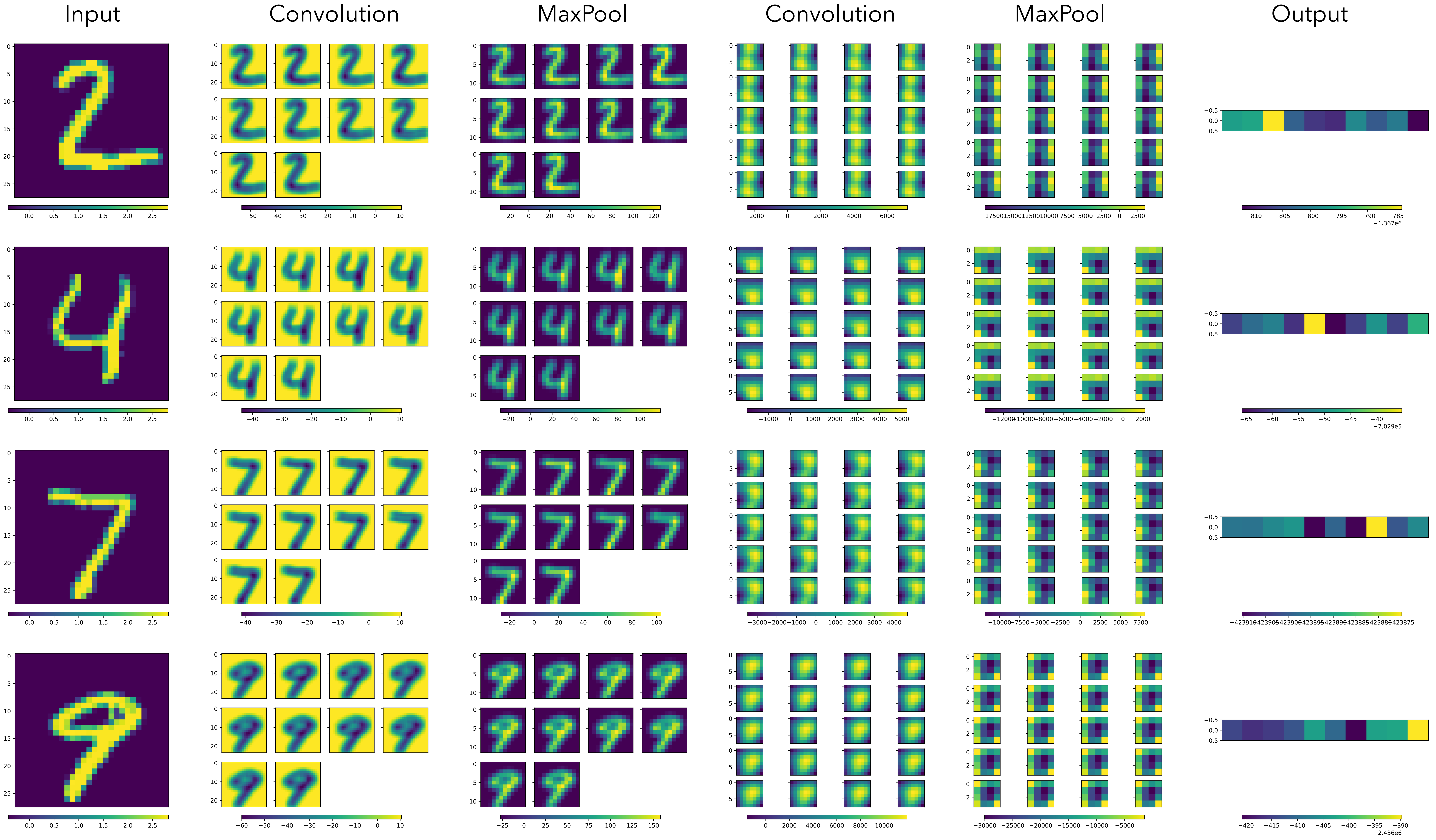
$$F_{x,y} = \begin{cases} 0 & \text{if } y \not\leq x \\ 1 & \text{if } x = y \\ -1 & \text{if } (y,x) \in E \\ -\sum_{\substack{y \leq w \leq z \\ (z,x) \in E}} \mu(y,w) & \text{otherwise} \end{cases}$$



Activations s



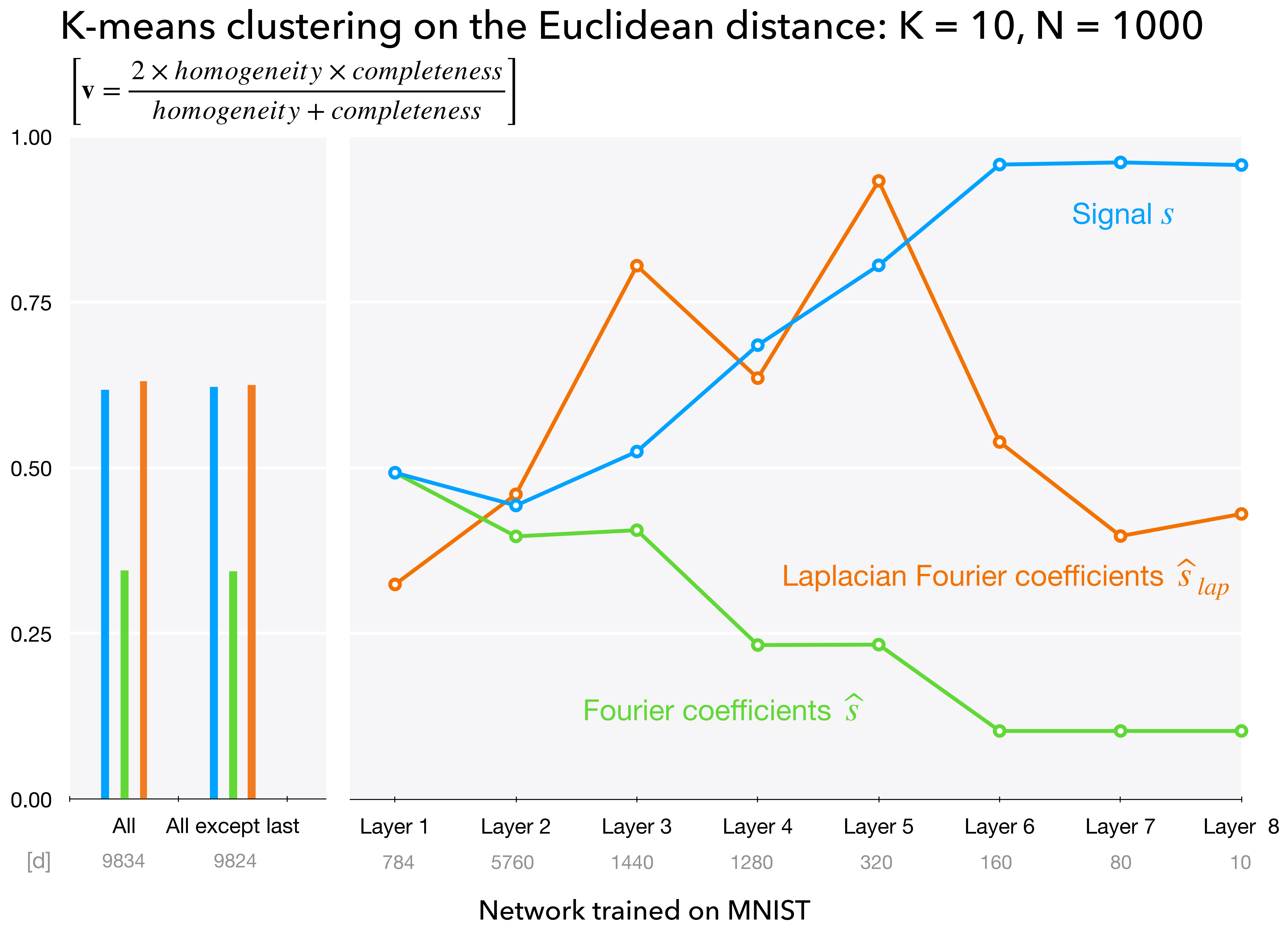
Fourier Coefficients $\hat{s}$



Clustering

$$\hat{s}_x = F_{x,:}\, s$$

$$\hat{s}_x = c_{s,i} + s_x$$



pip install nsp