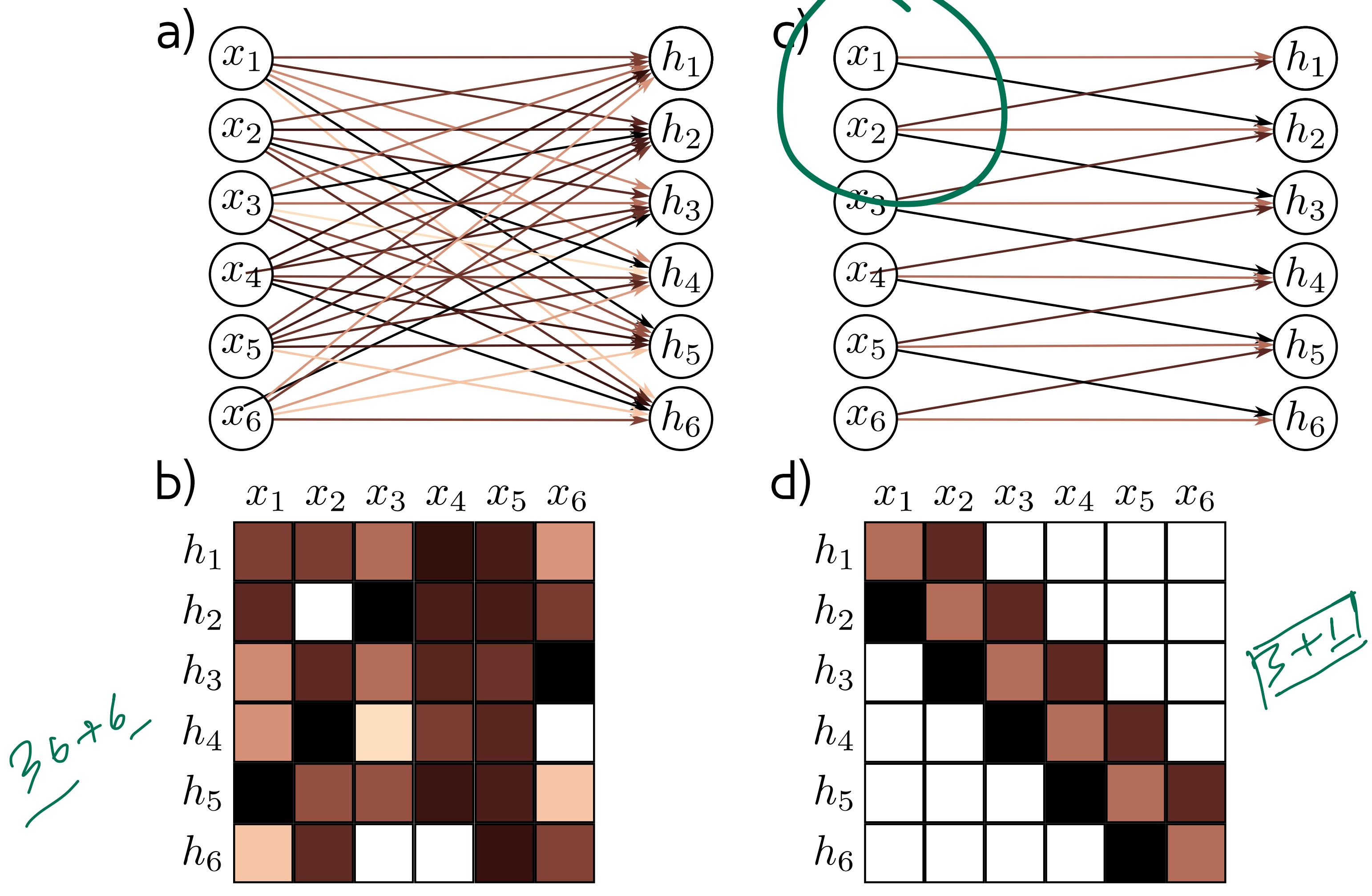
Special case of fully-connected network



Fully connected network

Convolution, kernel 3, stride 1, dilation 1

why CNNs?

→ sparser connection

→ weight sharing

1-D CNNs

0
0
0
0
0
0

0
0
0

i/p size
= I

Kernel size
= K
= M [incl. dilation]

o/p size
= $\boxed{I - M + 1}$

w_1 0
w_2 0
w_3 0

$D=3 \Rightarrow$

w_1
0
0
w_2
0
0
w_3

$M = K + (K-1)(D-1)$

PD

(dilation, padding, stride)
= 1 (def.)
= 0 (def.)
valid conv.
= 1 (def.)

$I - M + 1$

padding $P = \boxed{\left\lceil \frac{M-1}{2} \right\rceil}$ zeros at either side

o/p size $\quad I - M + 2P + 1 \quad (= I)$

Stride S : (A) $\dfrac{I - M + 2P + 1}{S}$ $\quad 5.5$ $\quad$ (B) $\left\lfloor \dfrac{I - M + 2P + 1}{S} \right\rfloor$ $\quad 5$

$I = 11 \quad M = 3 \quad P = 1, \quad S = 2$

(C) $\boxed{\left\lfloor \dfrac{I - M + 2P}{S} \right\rfloor + 1}$ ✓

0
0
0
11 0 $\quad$ 6
0
0
0

$S = 3 \to 4$
$I = 13$

0
0 →

0
0
0 →
0
0
0 →
0
0
0 →
0
0

4

$$\left\lfloor \frac{13-3}{3} \right\rfloor + 1$$

$$= 4$$

$I = 11 \quad M = 1 \quad P = 0 \quad S = 2$

$\rightarrow 6$

(c) 6 (d) 6

(n)

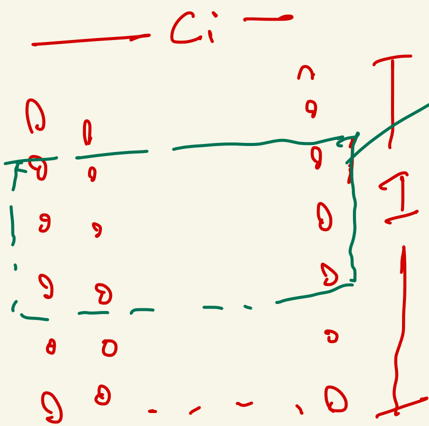
$$\left\lceil \frac{I - M + 2P + 1}{S} \right\rceil$$

$I = 11 \rightarrow 6$

$I = 12 \rightarrow 6$

1 bias

dim: $\boxed{C_i \times K}$ weights



Kernel size K

1 o/p channel

Ci i/p channel

multiple feature maps

Weights

$K \times C_o$

Biases

C_o

K

1 i/p channel

C_o o/p channels

1 feature map

C_i i/p

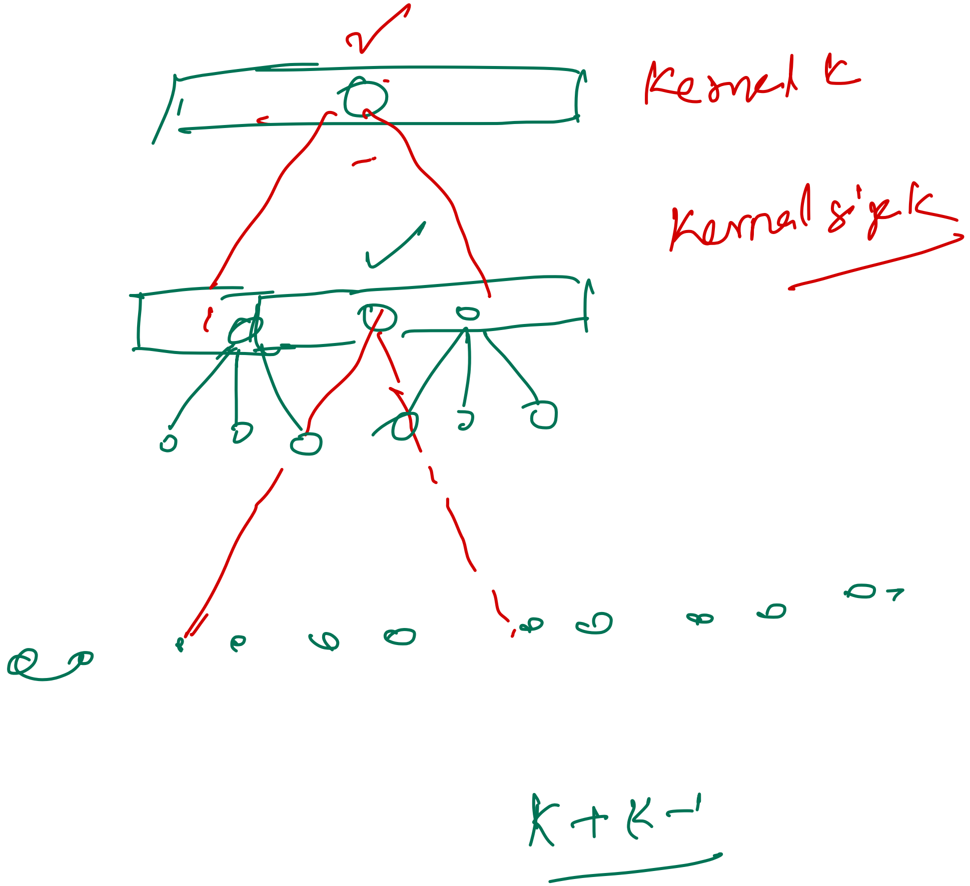
C_o o/p

$K \times C_i \times C_o$ weights

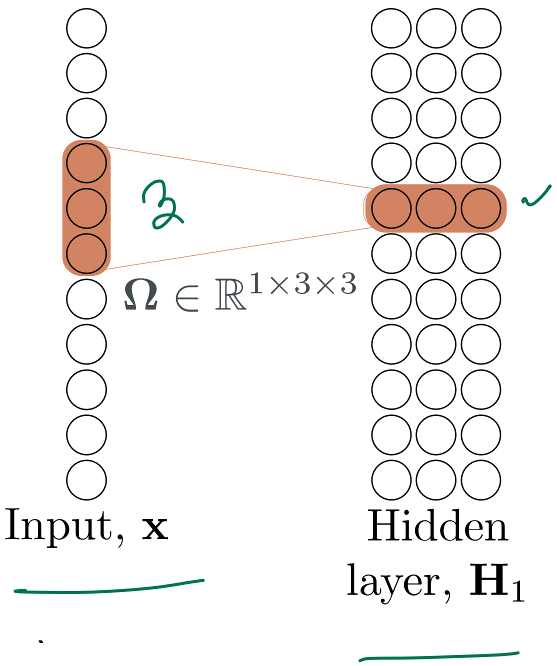
C_o biases

Convolutional networks

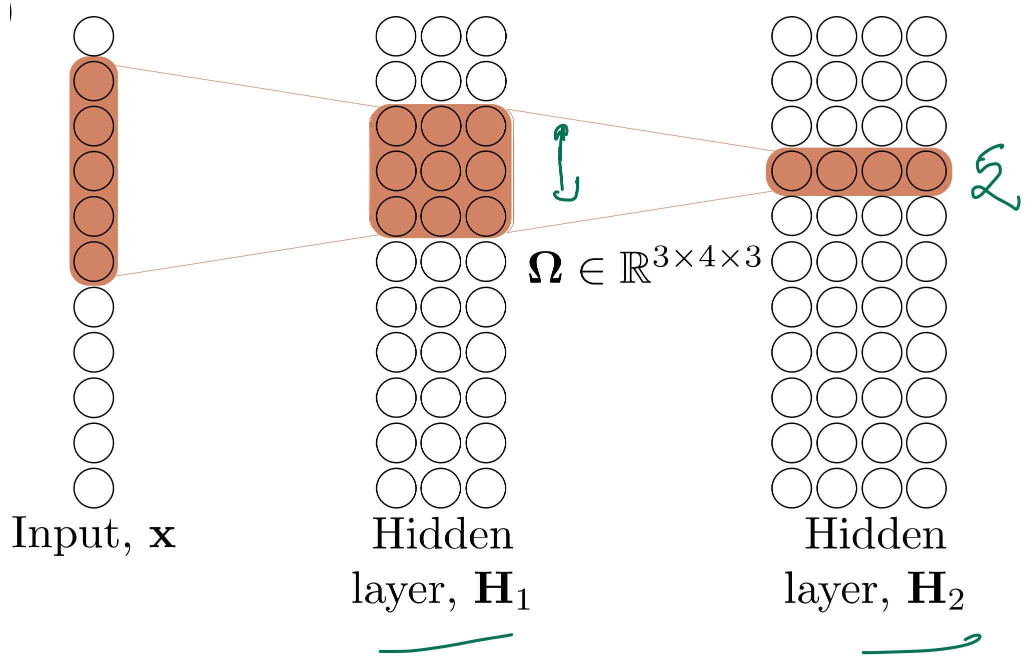
- Networks for images
- Invariance and equivariance
- 1D convolution
- Convolutional layers
- Channels
- Receptive fields
- Convolutional network for MNIST 1D



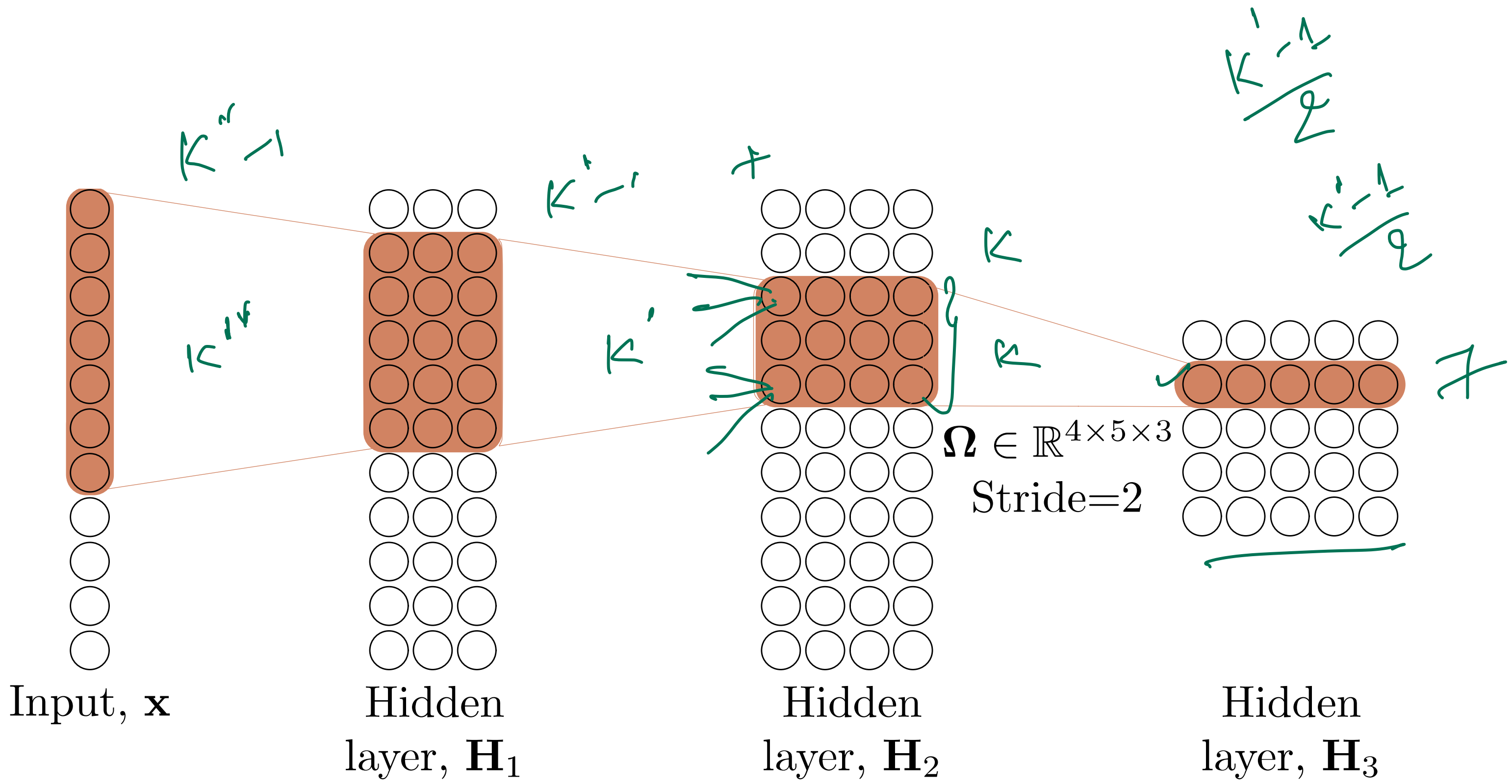
Receptive fields



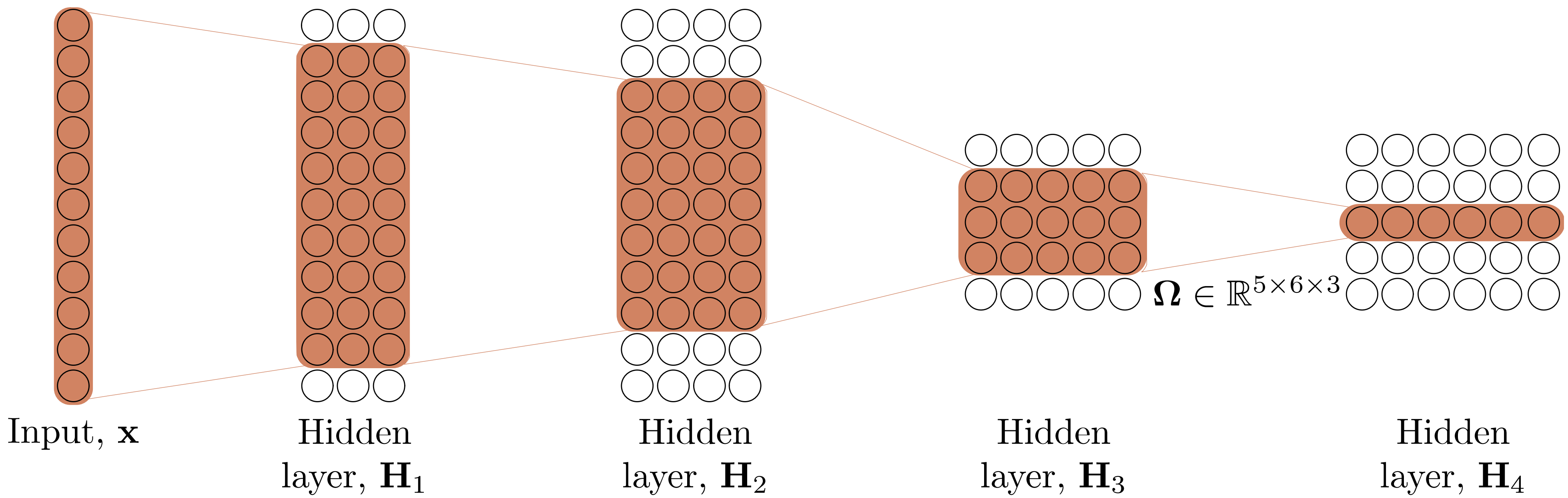
Receptive fields



Receptive fields



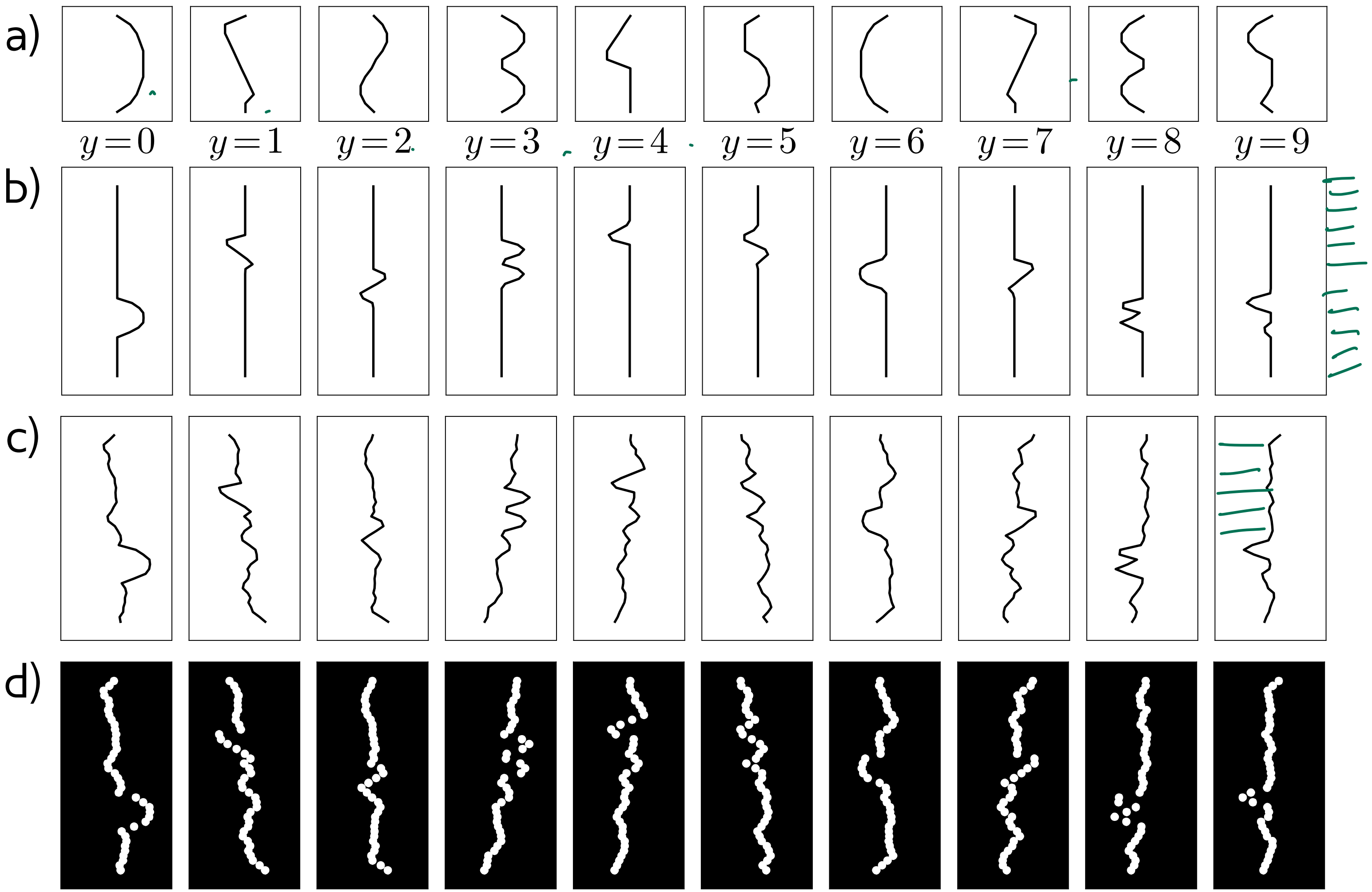
Receptive fields



Convolutional networks

- Networks for images
- Invariance and equivariance
- 1D convolution
- Convolutional layers
- Channels
- Receptive fields
- Convolutional network for MNIST 1D

MNIST 1D Dataset

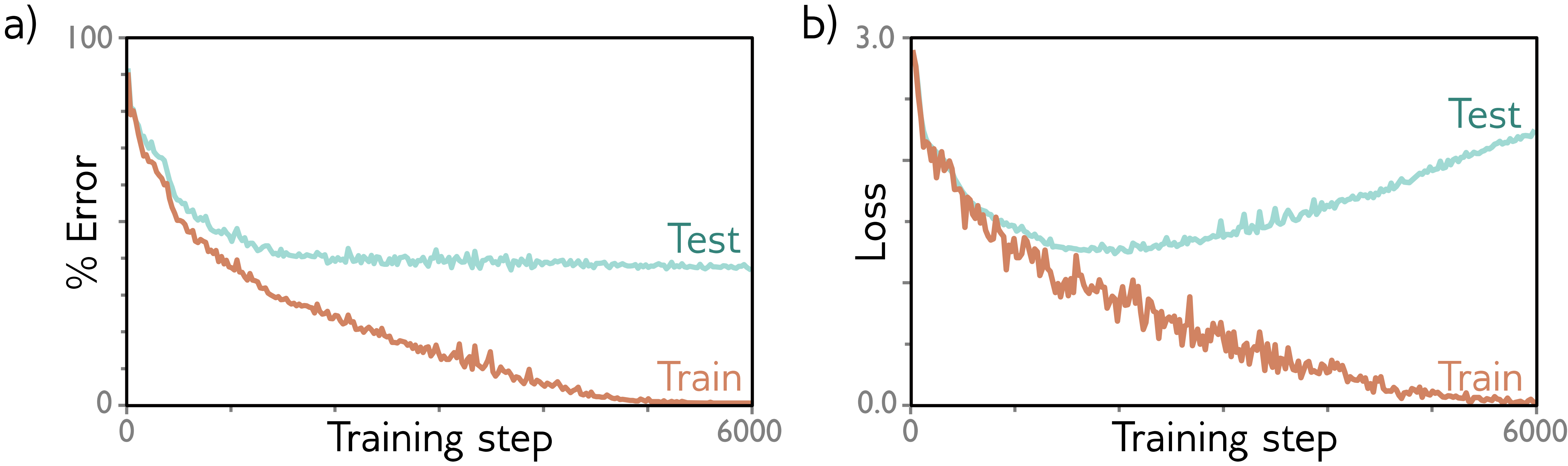


40 locations

$D_i = 40$

$D_o = 10$

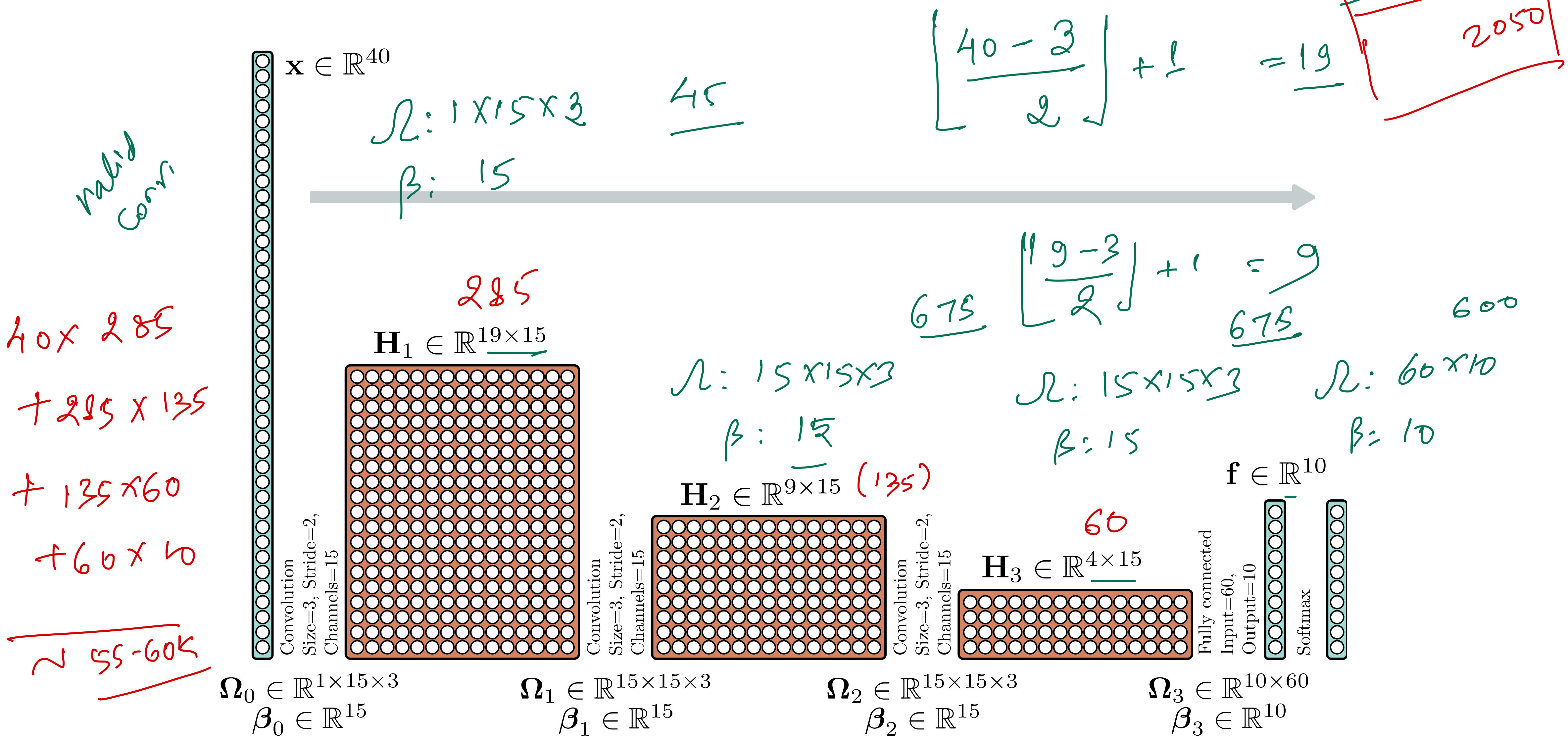
MNIST-1D results for fully-connected network



Convolutional network

- Four hidden layers
- Three convolutional layers
- One fully-connected layer
- Softmax at end
- Total parameters = 2050
- Trained for 100,000 steps with SGD, LR = 0.01, batch size 100

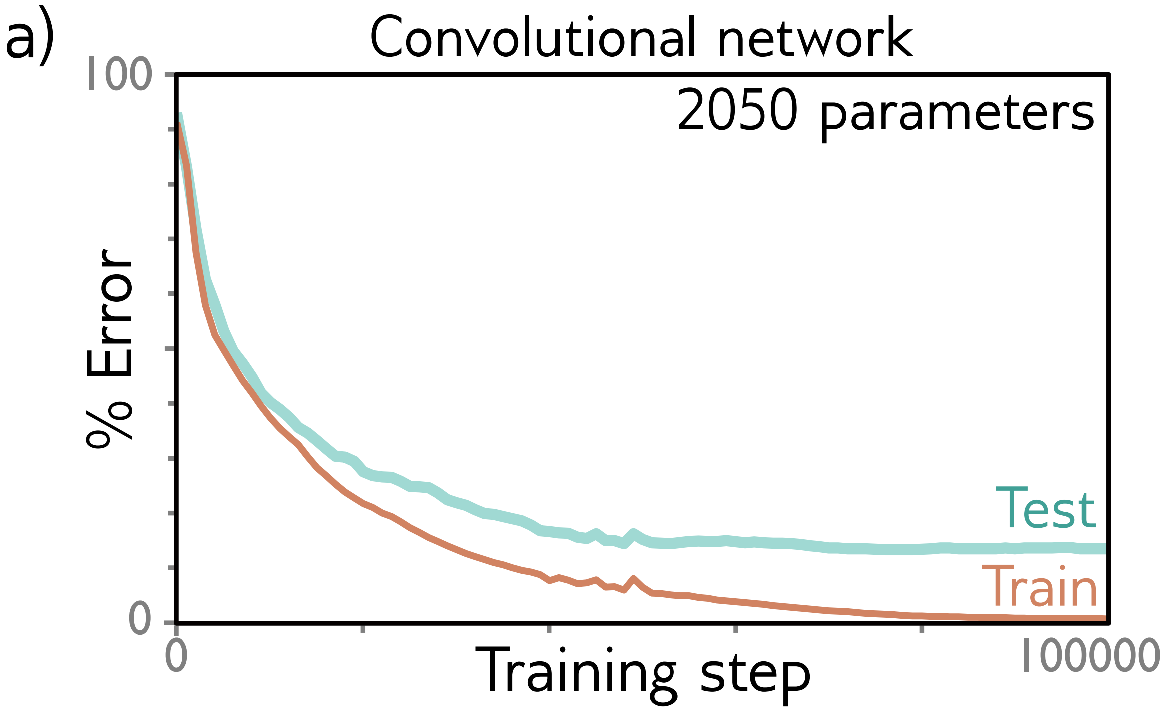
MNIST-1D convolutional network

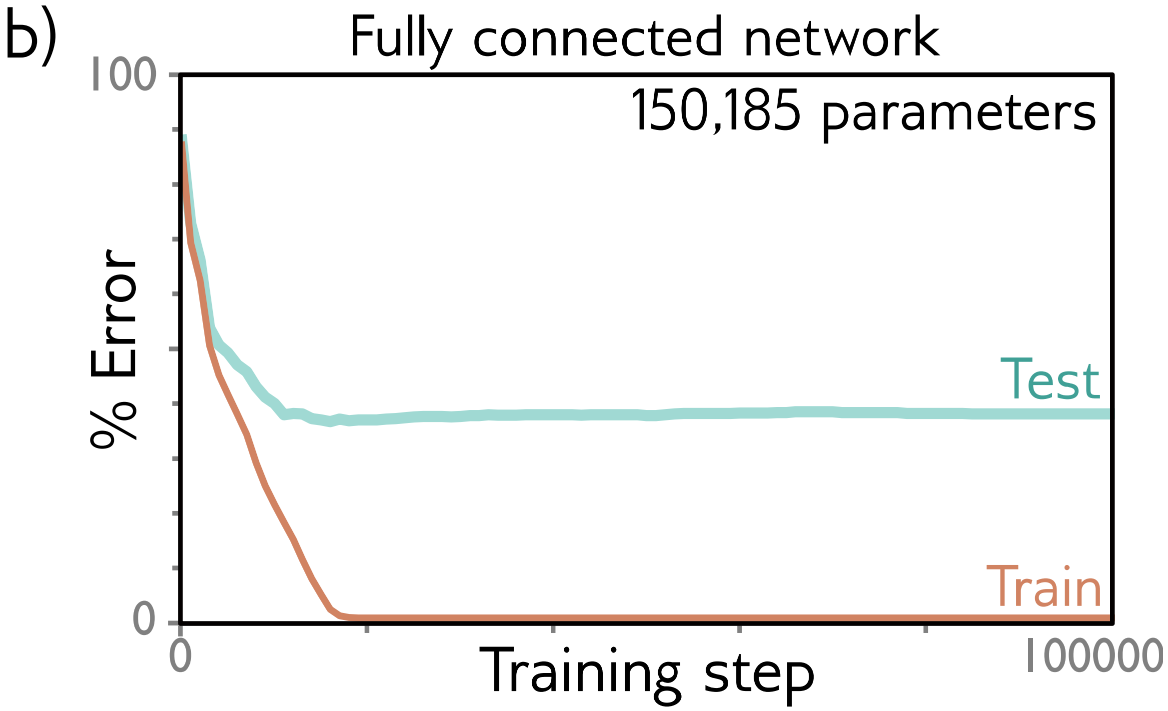


Fully connected network

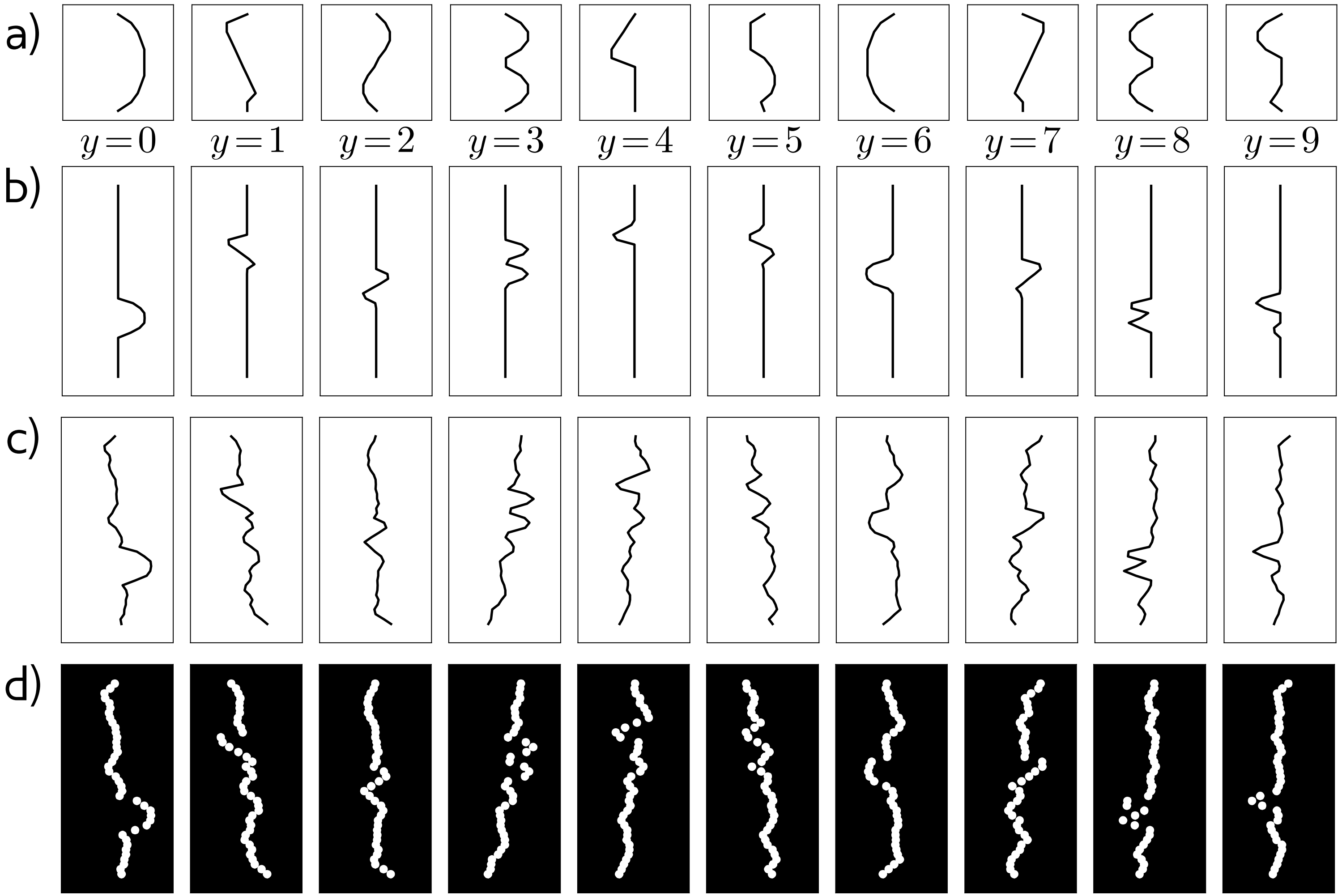
- Exactly same number of layers and hidden units
- All fully-connected layers
- Total parameters = 150,185

Performance





MNIST 1D Dataset



Why?

- Better inductive bias
- Forced the network to process each location similarly
- Shares information across locations
- Search through a smaller family of input/ouput mappings, all of which are plausible

Convolution #2

- 2D Convolution
- Downsampling and upsampling, 1x1 convolution
- Image classification
- Object detection
- Semantic segmentation
- Residual networks
- U-Nets and hourglass networks

2D Convolution

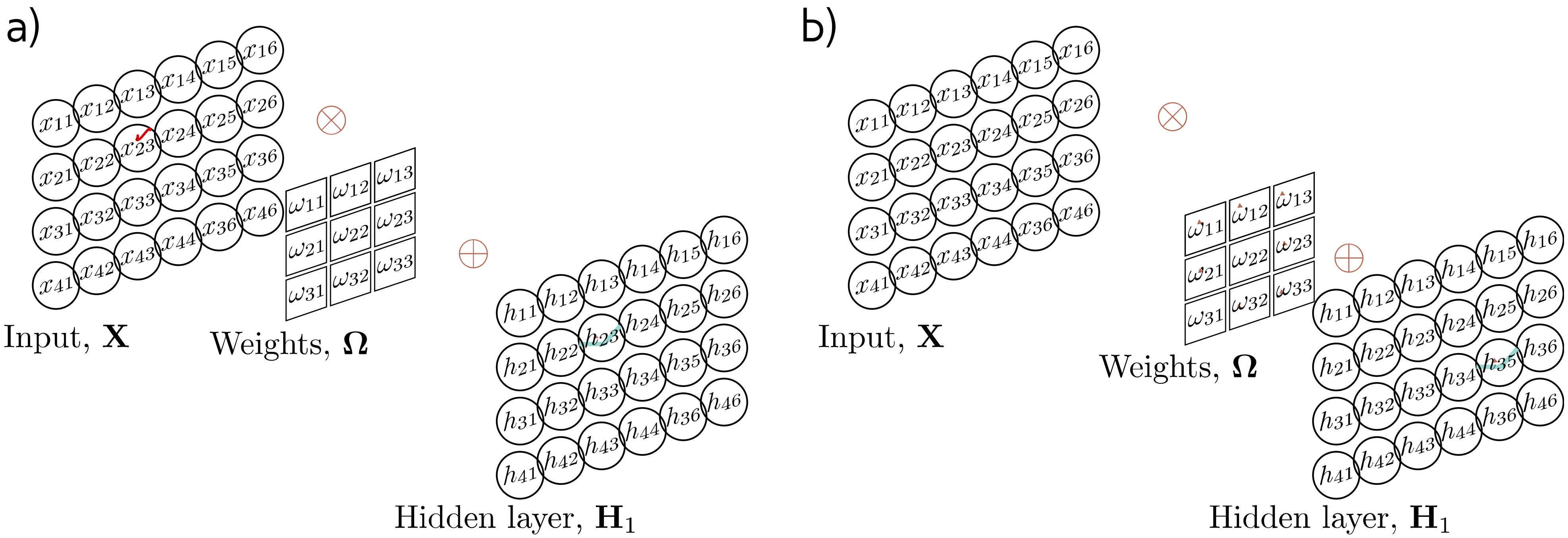
Square Kernel

K: odd

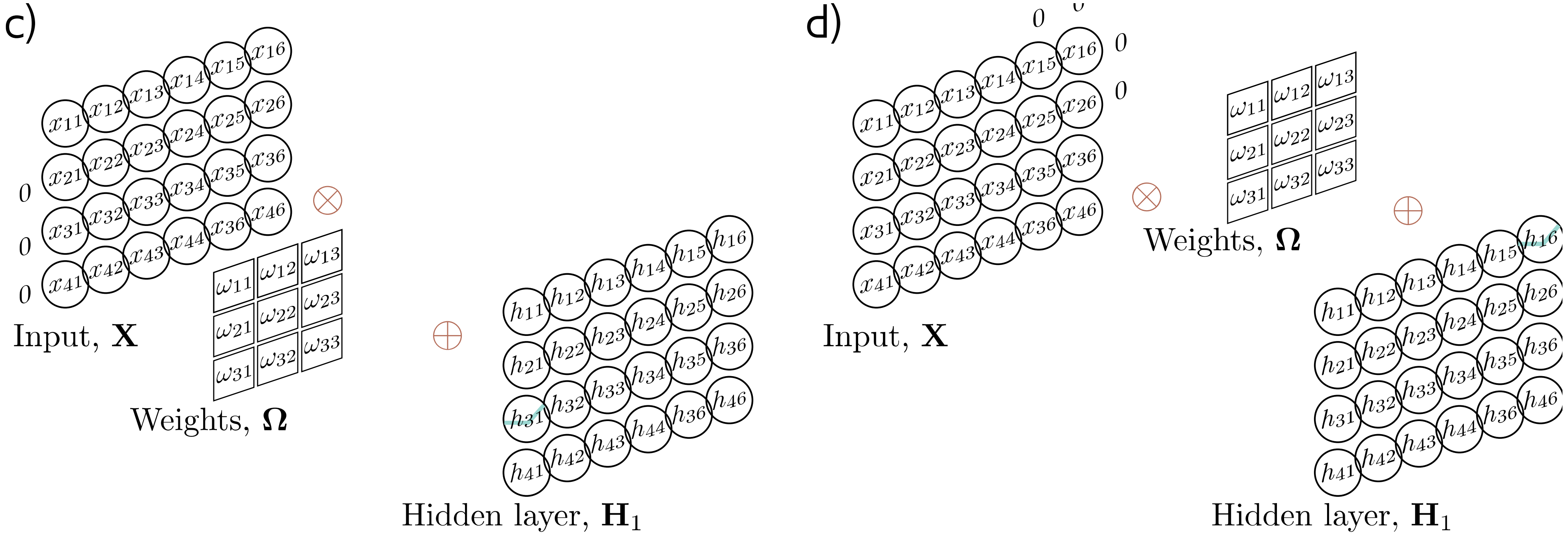
- Convolution in 2D
  - Weighted sum over a K x K region
  - K x K weights
- Build into a convolutional layer by adding bias and passing through activation function

$$h_{i,j} = \mathrm{a}\left[\beta + \sum_{m=1}^{3}\sum_{n=1}^{3} \underline{\omega_{m,n} x_{i+m-2,j+n-2}}\right]$$

2D Convolution



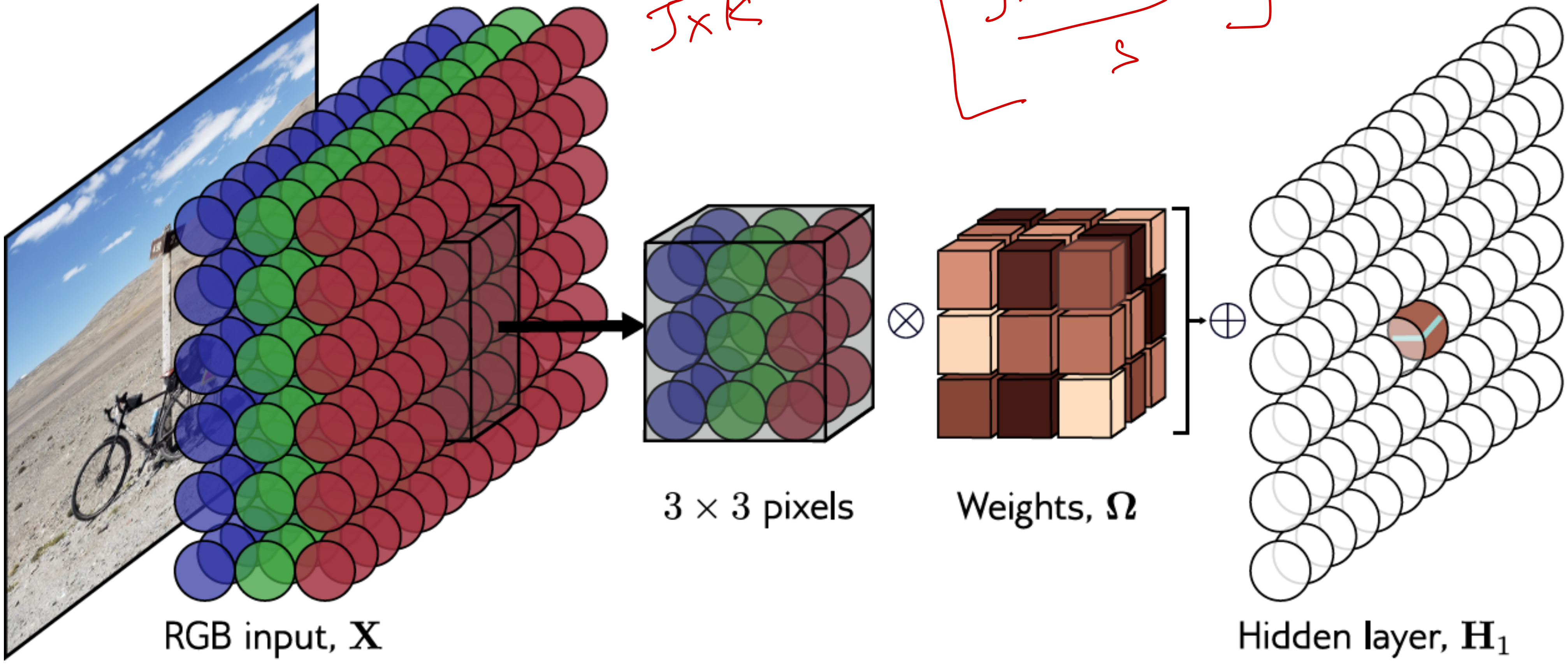
2D Convolution



Channels in 2D convolution

$J \times K$

$$\left\lfloor \frac{J-M+2P}{S} + 1 \right\rfloor \times \left\lfloor \frac{K-M+2P}{S} + 1 \right\rfloor$$



Kernel size, stride, dilation all
work as you would expect