

# Making Convolutional Networks Shift-Invariant Again

Richard Zhang  
Adobe Research



# Example classifications

86.7



P(correct class)

69.2



P(correct class)

# Deep Networks are not Shift-Invariant

46.3



P(correct class)

18.0



P(correct class)

# Deep Networks are not Shift-Invariant

46.3



P(correct class)

18.0



P(correct class)

Azulay and Weiss. *Why do deep convolutional networks generalize so poorly to small image transformations?* In ArXiv, 2018.  
Engstrom, Tsipras, Schmidt, Madry. *Exploring the Landscape of Spatial Robustness.* In ICML, 2019.

# Why is shift-invariance lost?

Why is shift-invariance lost?

“Convolutions are shift-equivariant”

Why is shift-invariance lost?

“Convolutions are shift-equivariant”

“Pooling builds up shift-invariance”

# Why is shift-invariance lost?

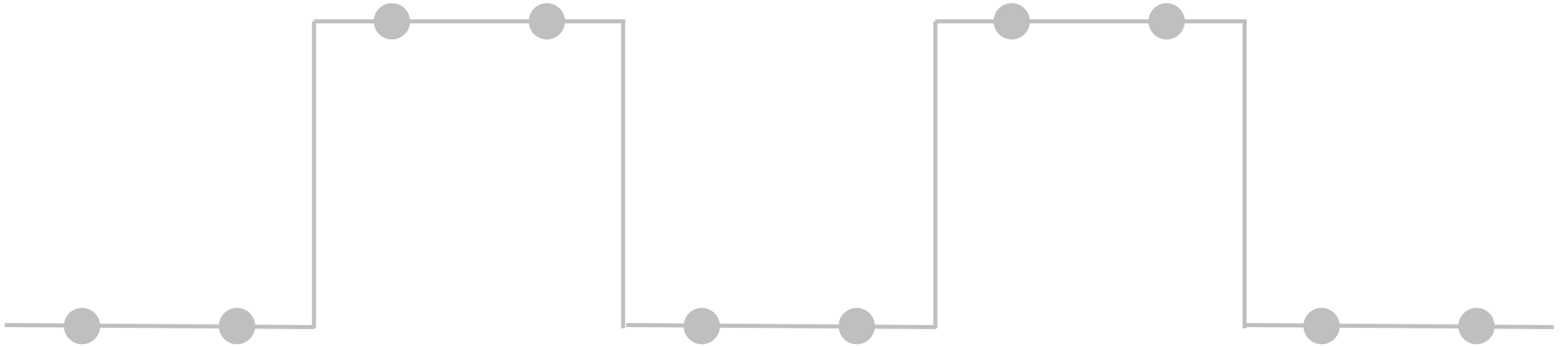
“Convolutions are **shift-equivariant**”

“Pooling builds up **shift-invariance**”

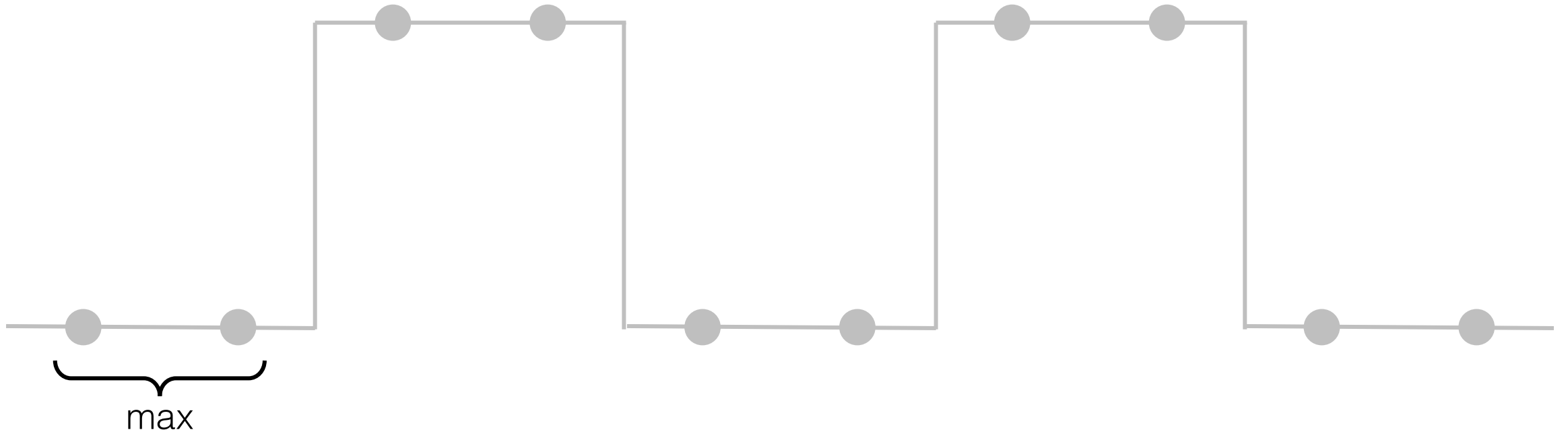
...but striding ignores Nyquist sampling theorem  
and **aliases**



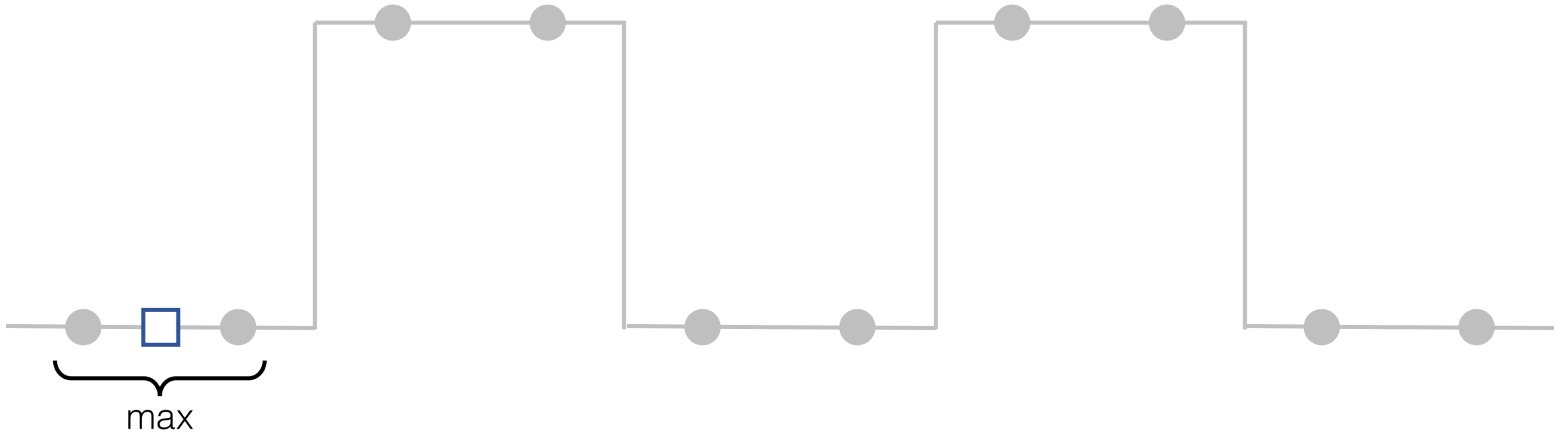
# Re-examining Max-Pooling



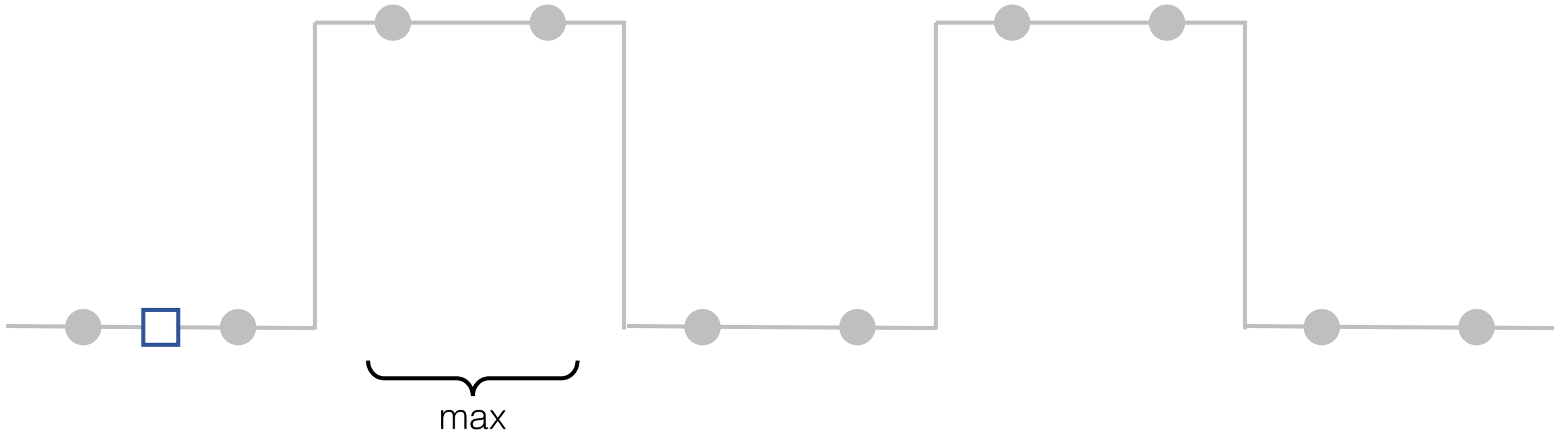
# Re-examining Max-Pooling



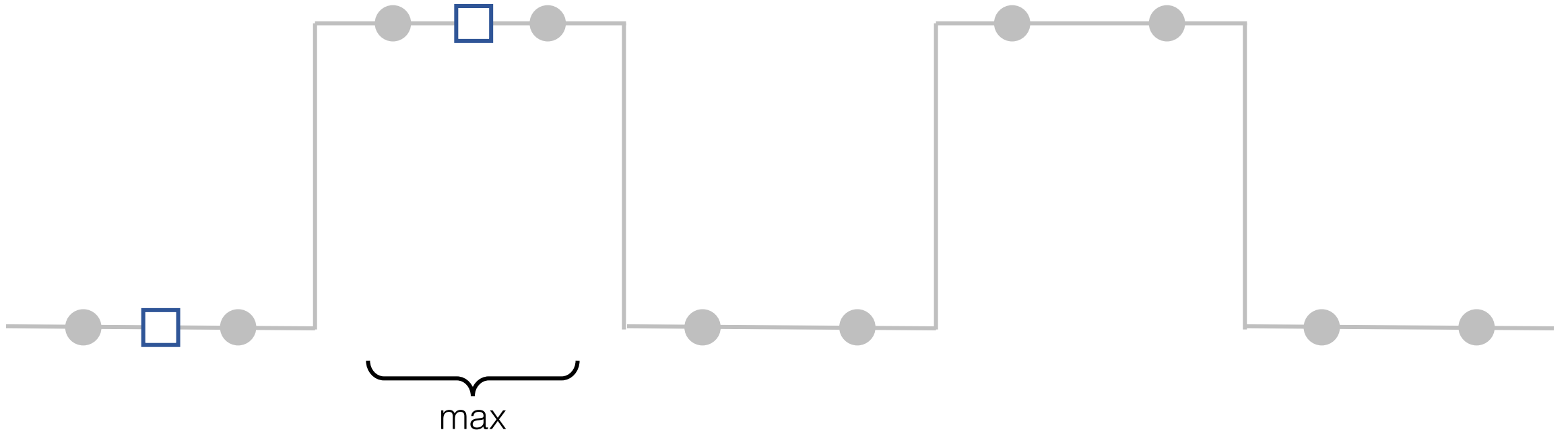
# Re-examining Max-Pooling



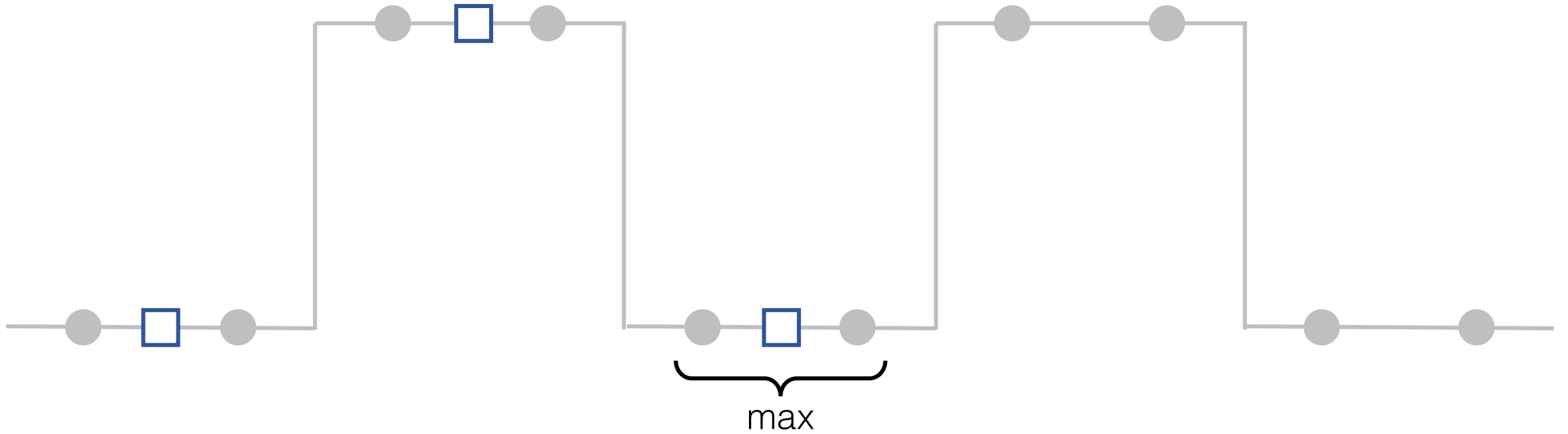
# Re-examining Max-Pooling



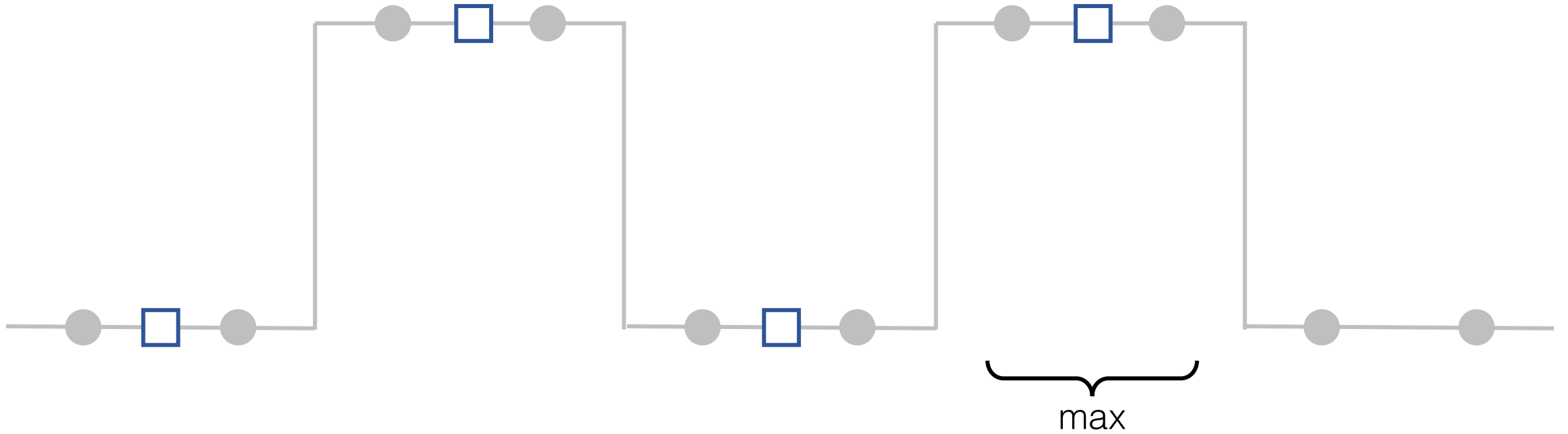
# Re-examining Max-Pooling



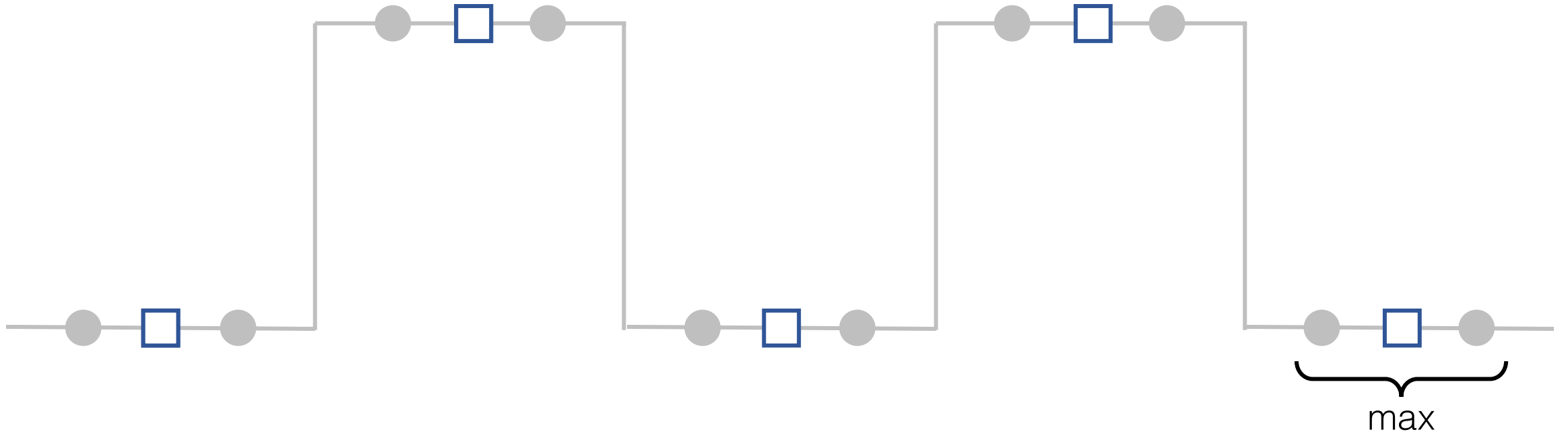
# Re-examining Max-Pooling



# Re-examining Max-Pooling

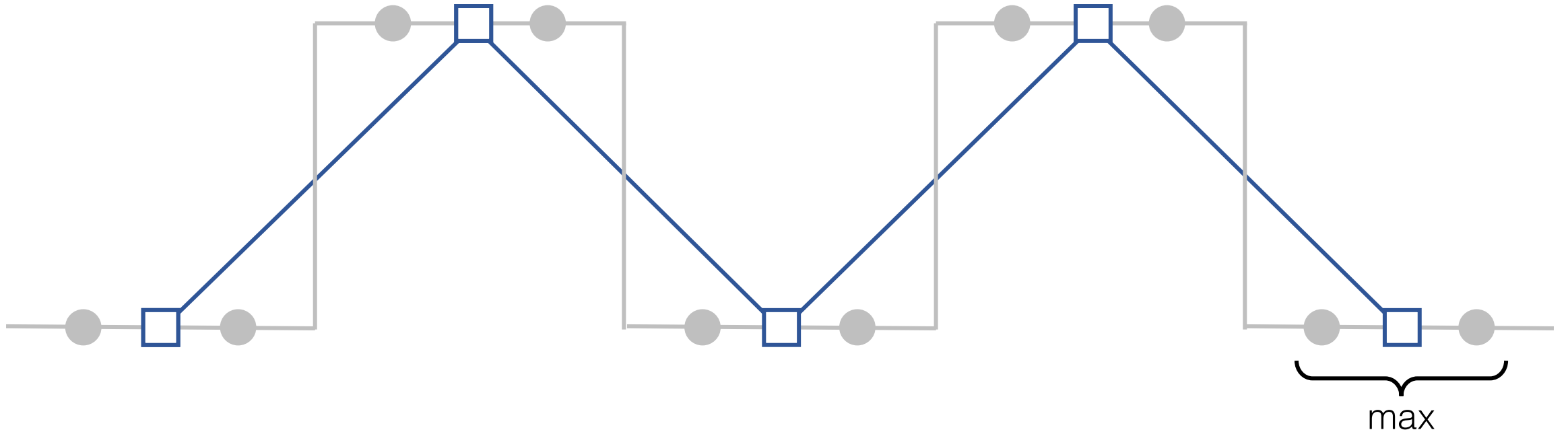


# Re-examining Max-Pooling

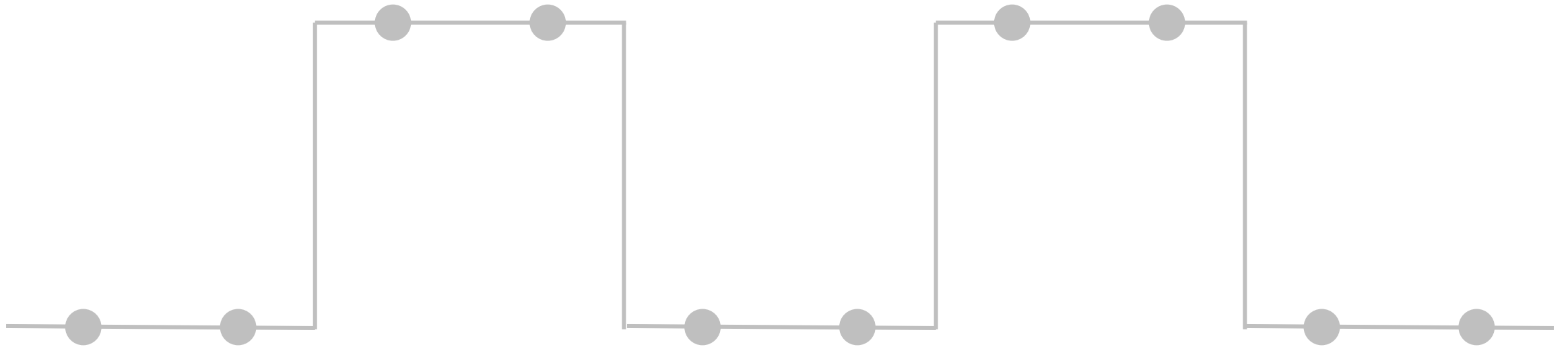




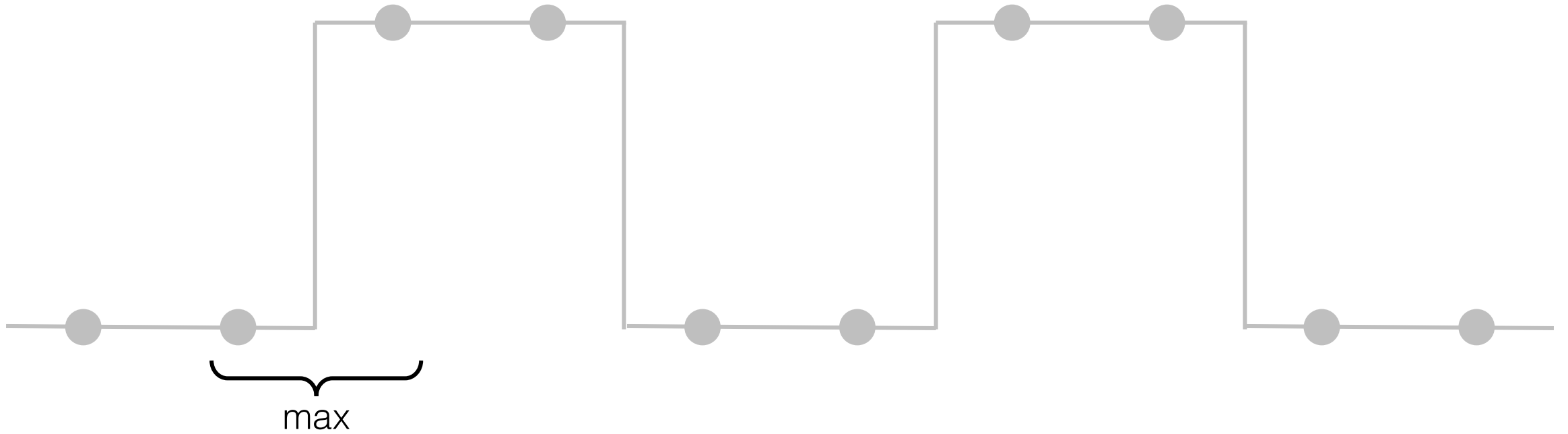
# Re-examining Max-Pooling



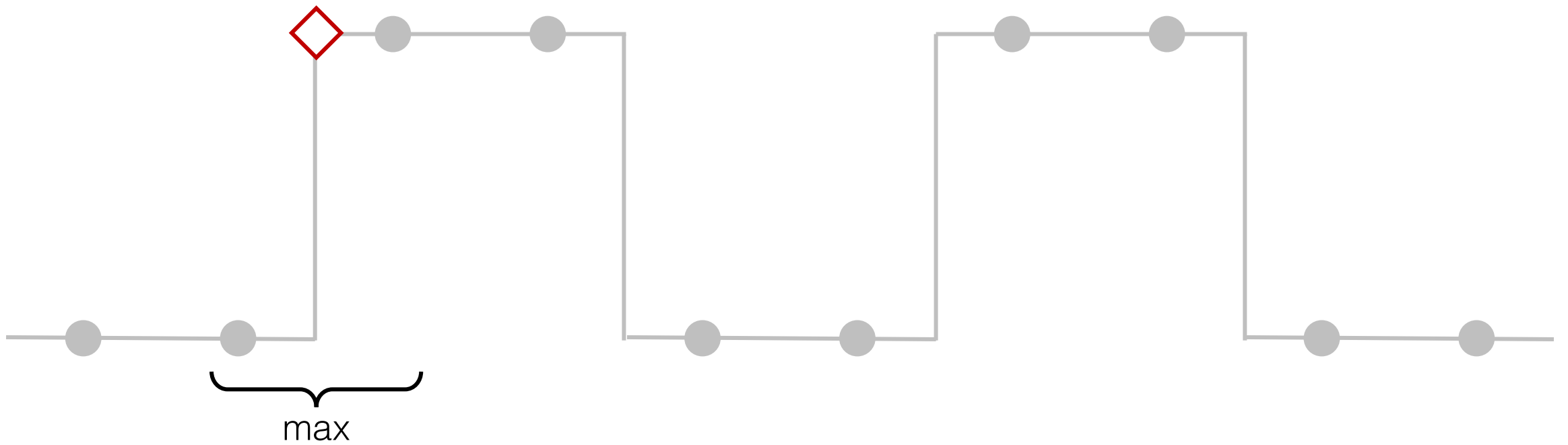
# Re-examining Max-Pooling



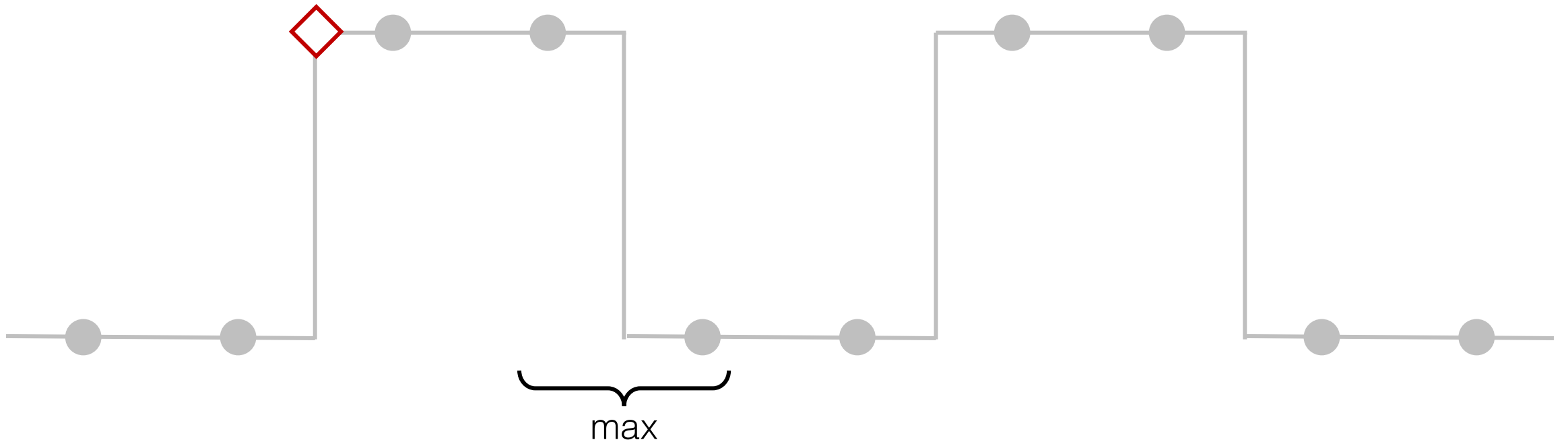
# Re-examining Max-Pooling



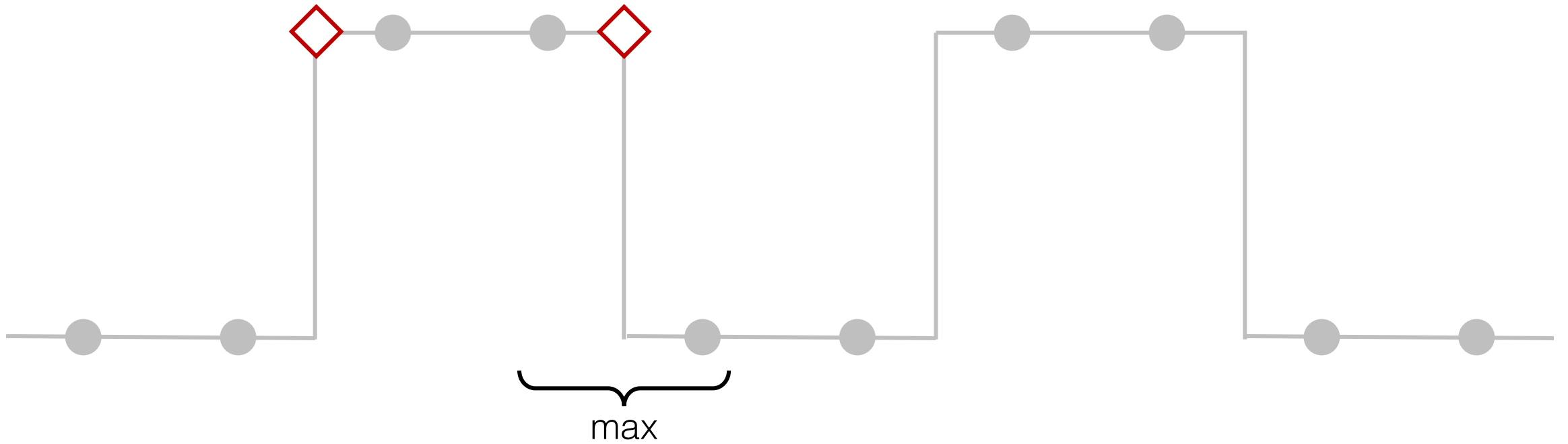
# Re-examining Max-Pooling



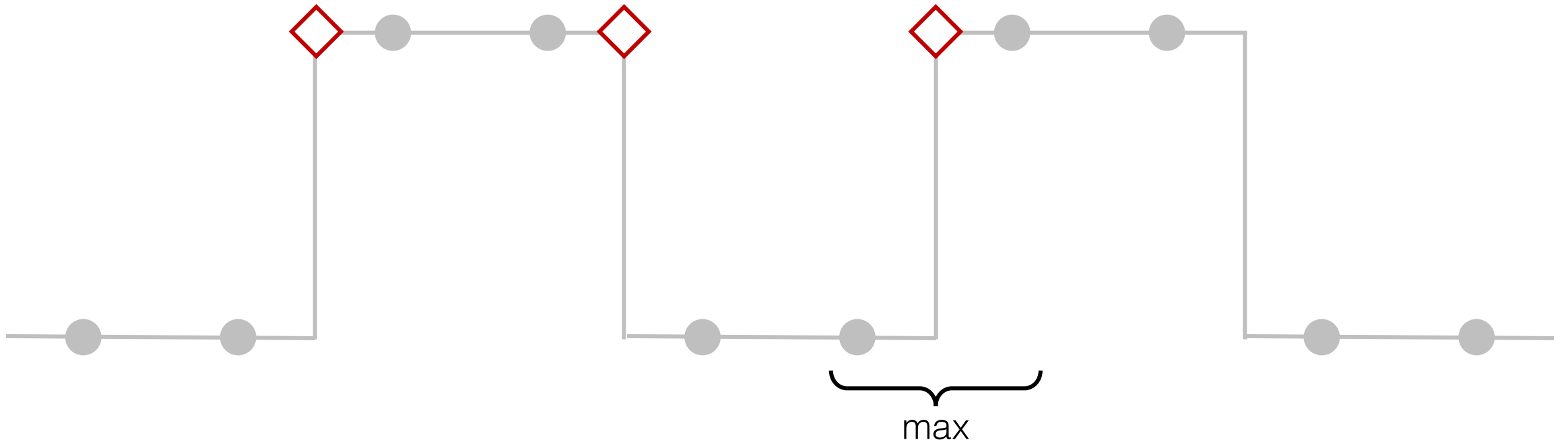
# Re-examining Max-Pooling



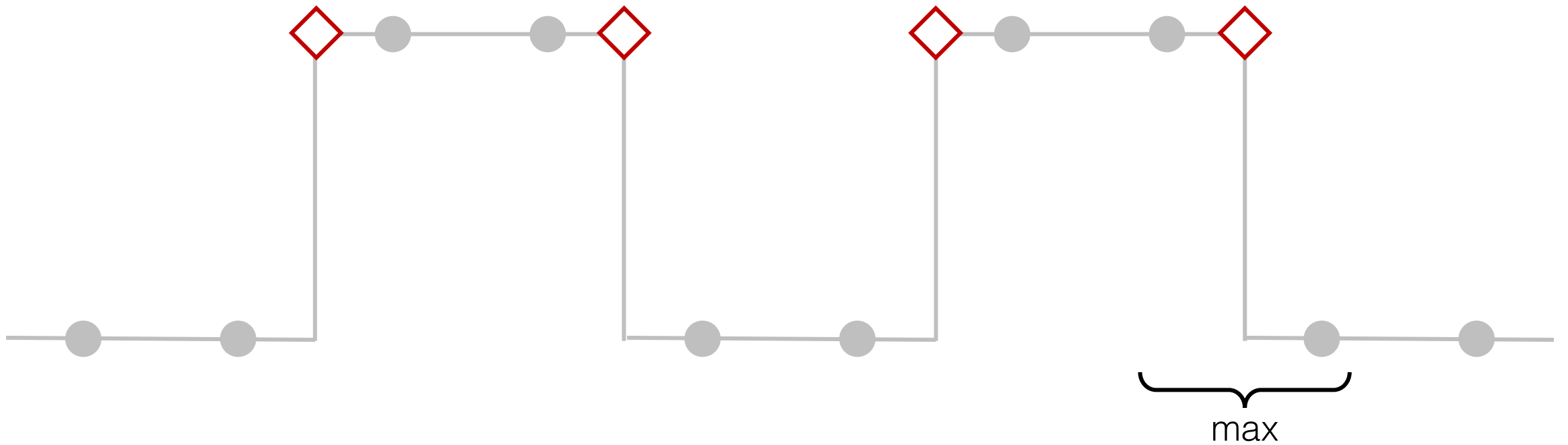
# Re-examining Max-Pooling



# Re-examining Max-Pooling

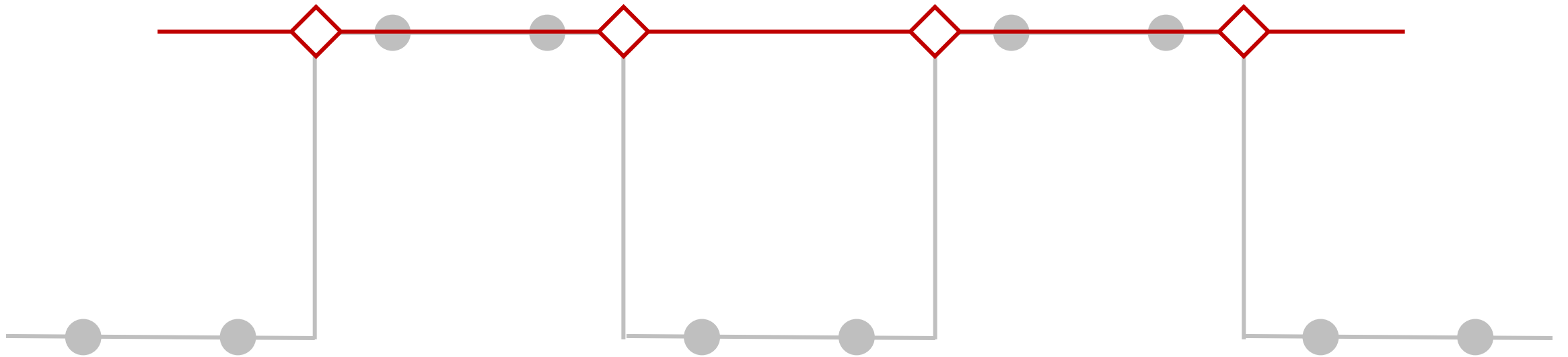


# Re-examining Max-Pooling





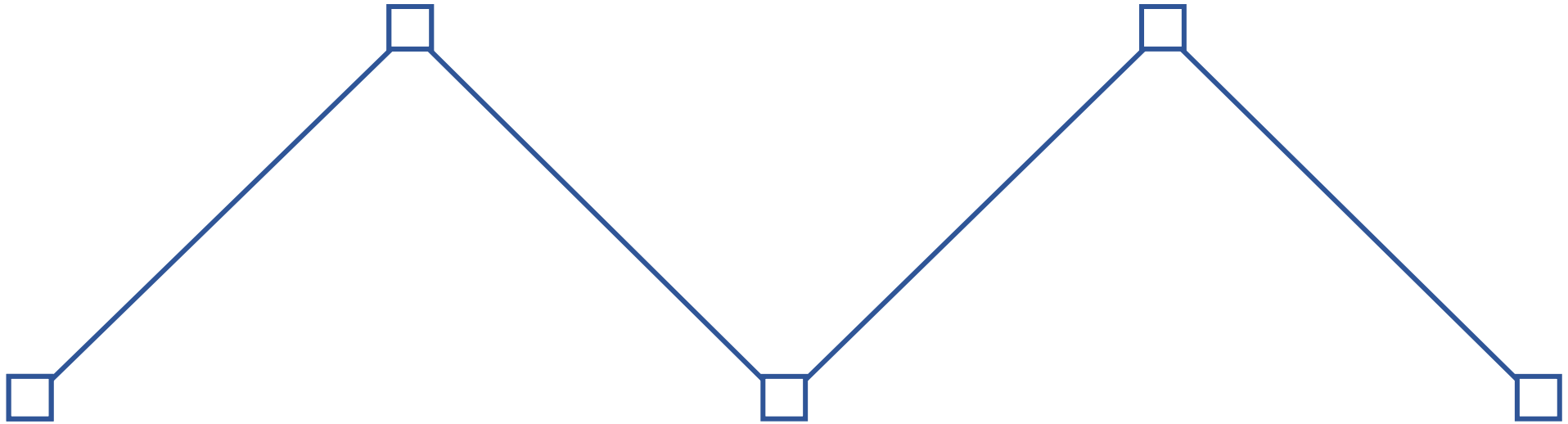
# Re-examining Max-Pooling



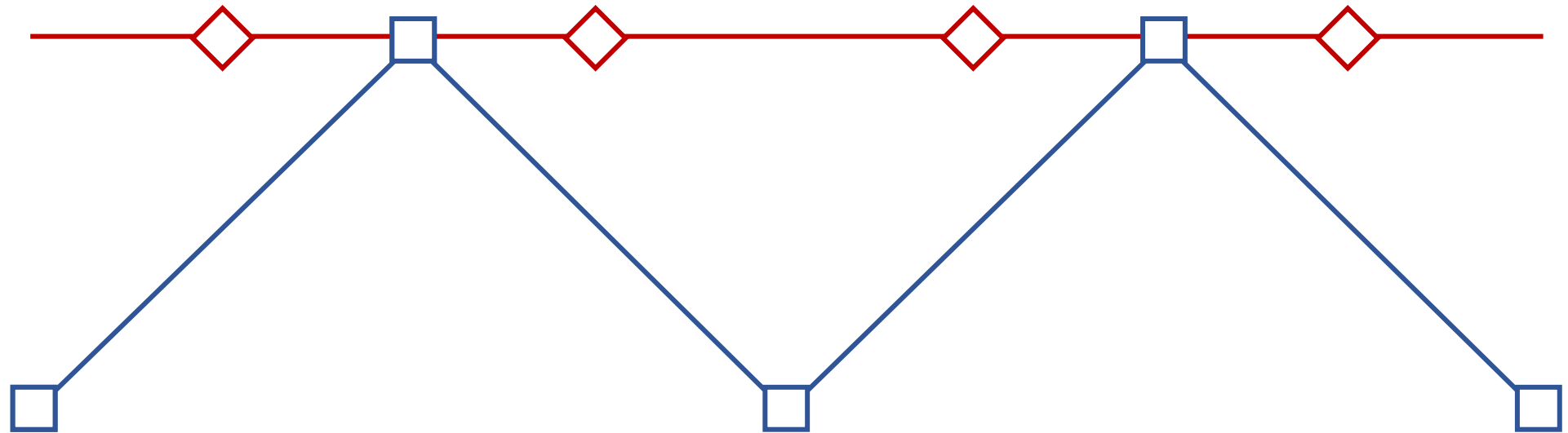
# Re-examining Max-Pooling



# Re-examining Max-Pooling



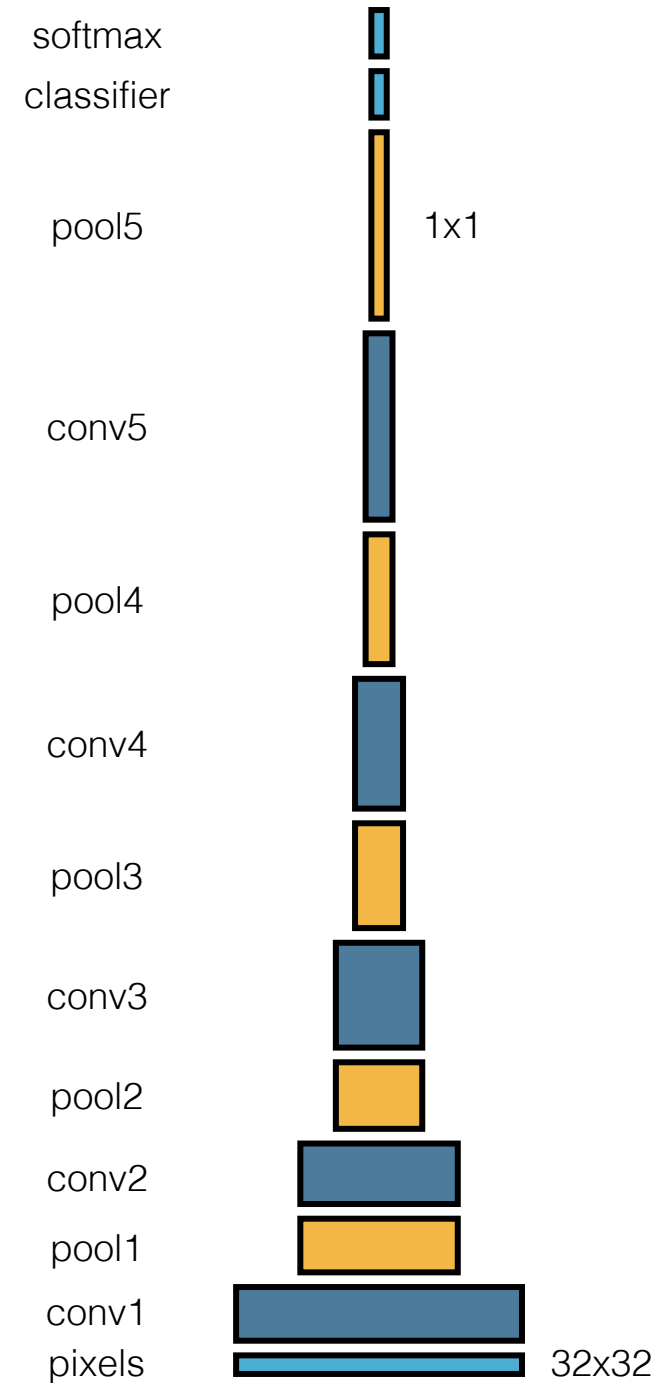
# Re-examining Max-Pooling



Max-pooling breaks shift-equivariance

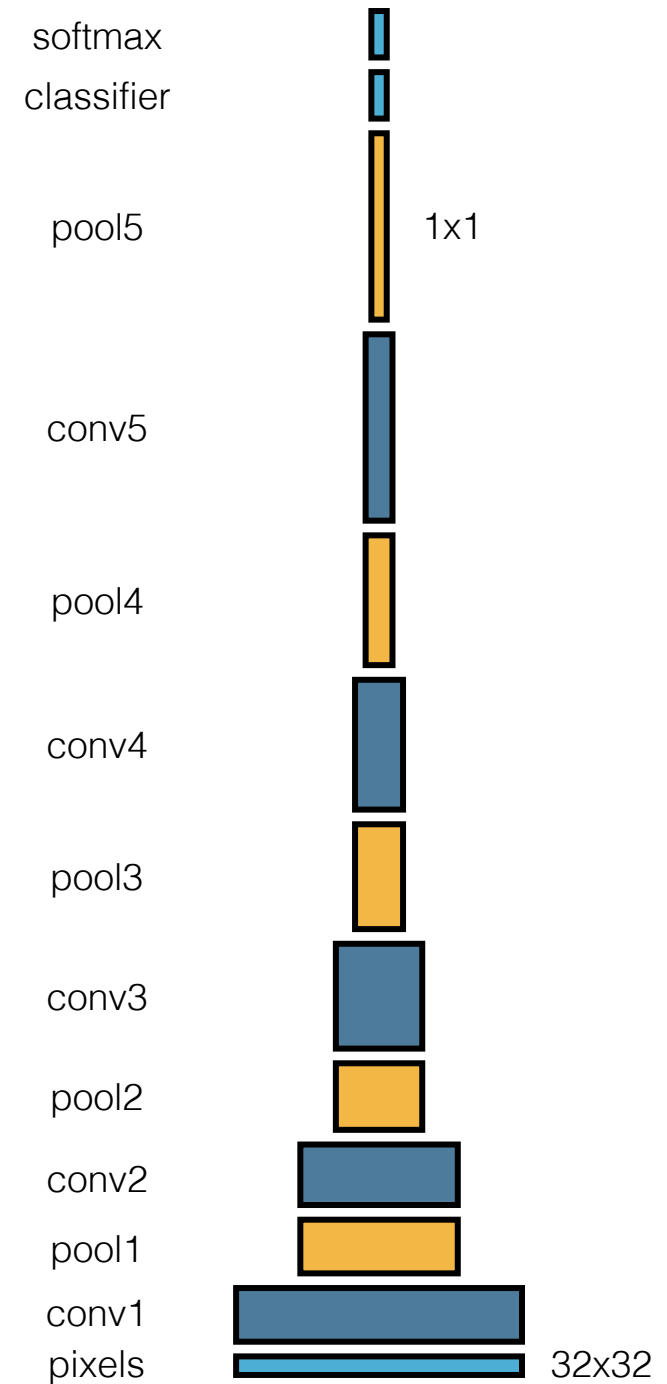
# Shift-equivariance in VGG

- CIFAR
- VGG network
  - 5 max-pools

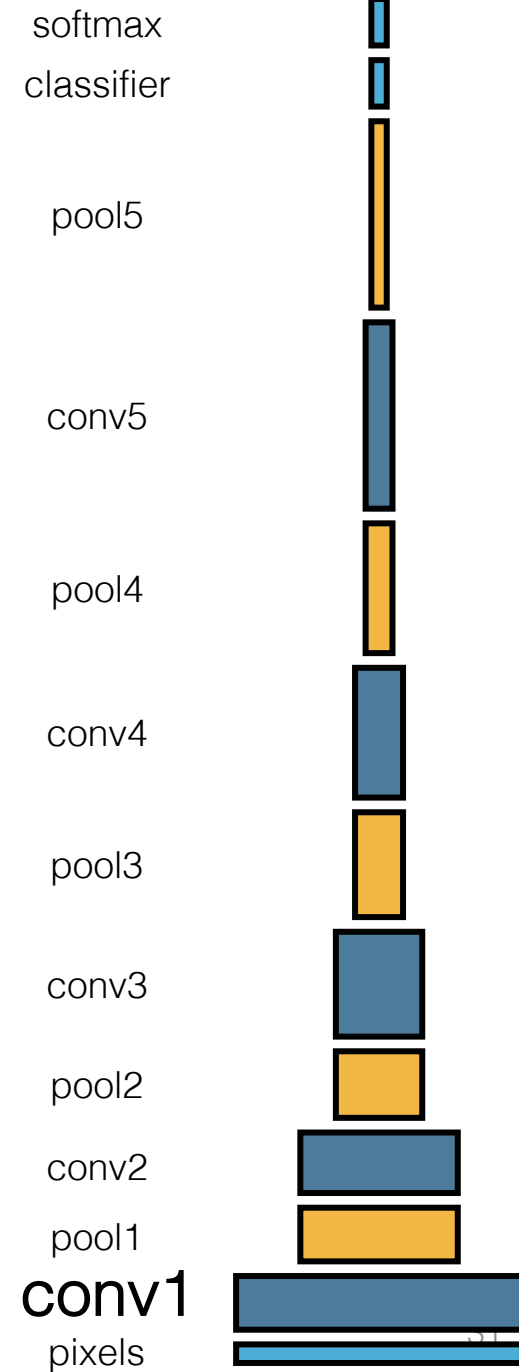
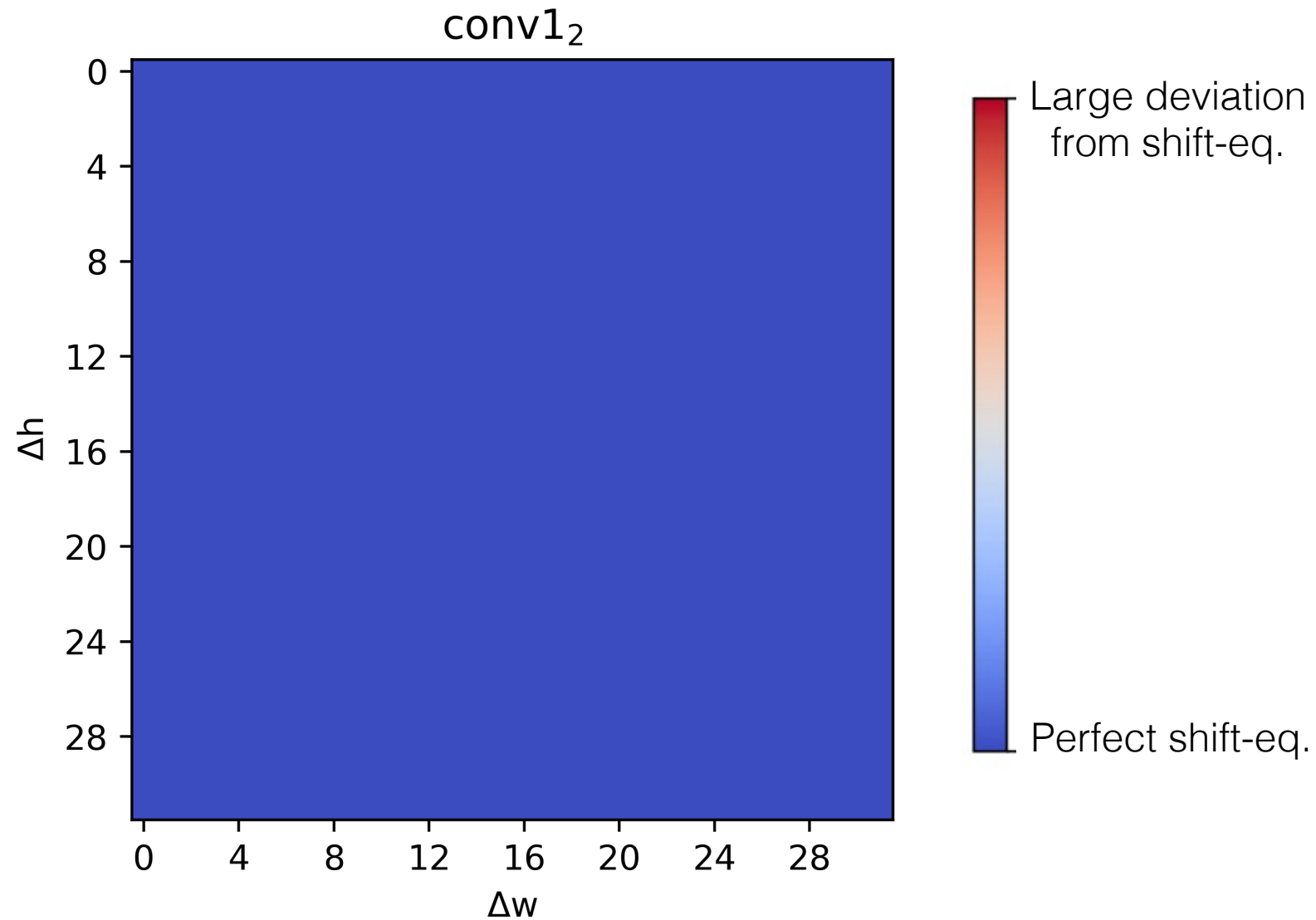


# Shift-equivariance in VGG

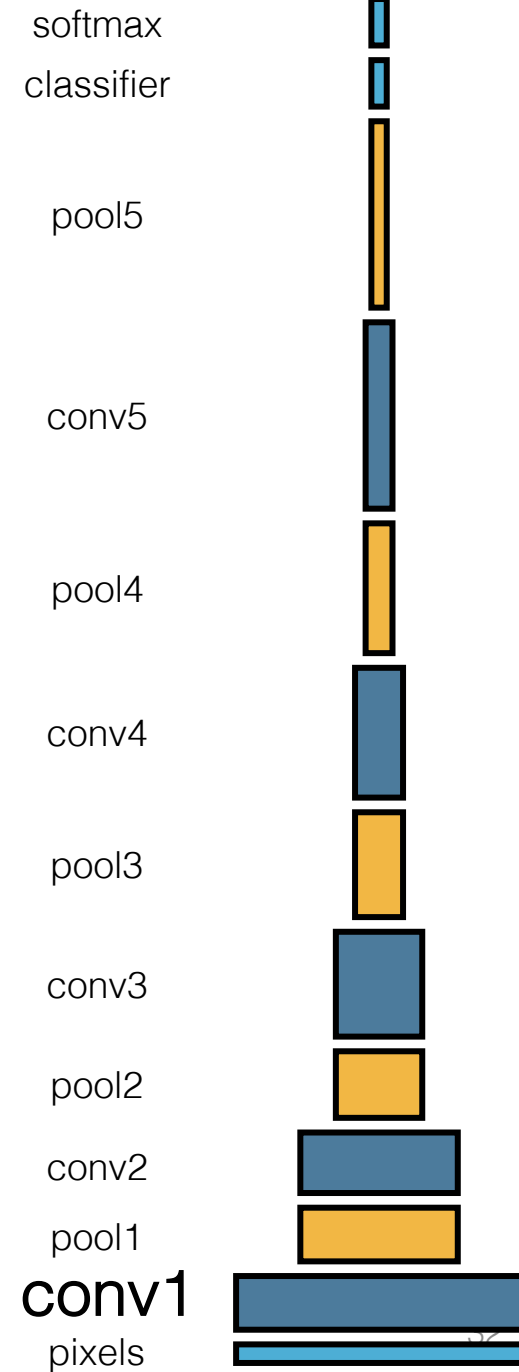
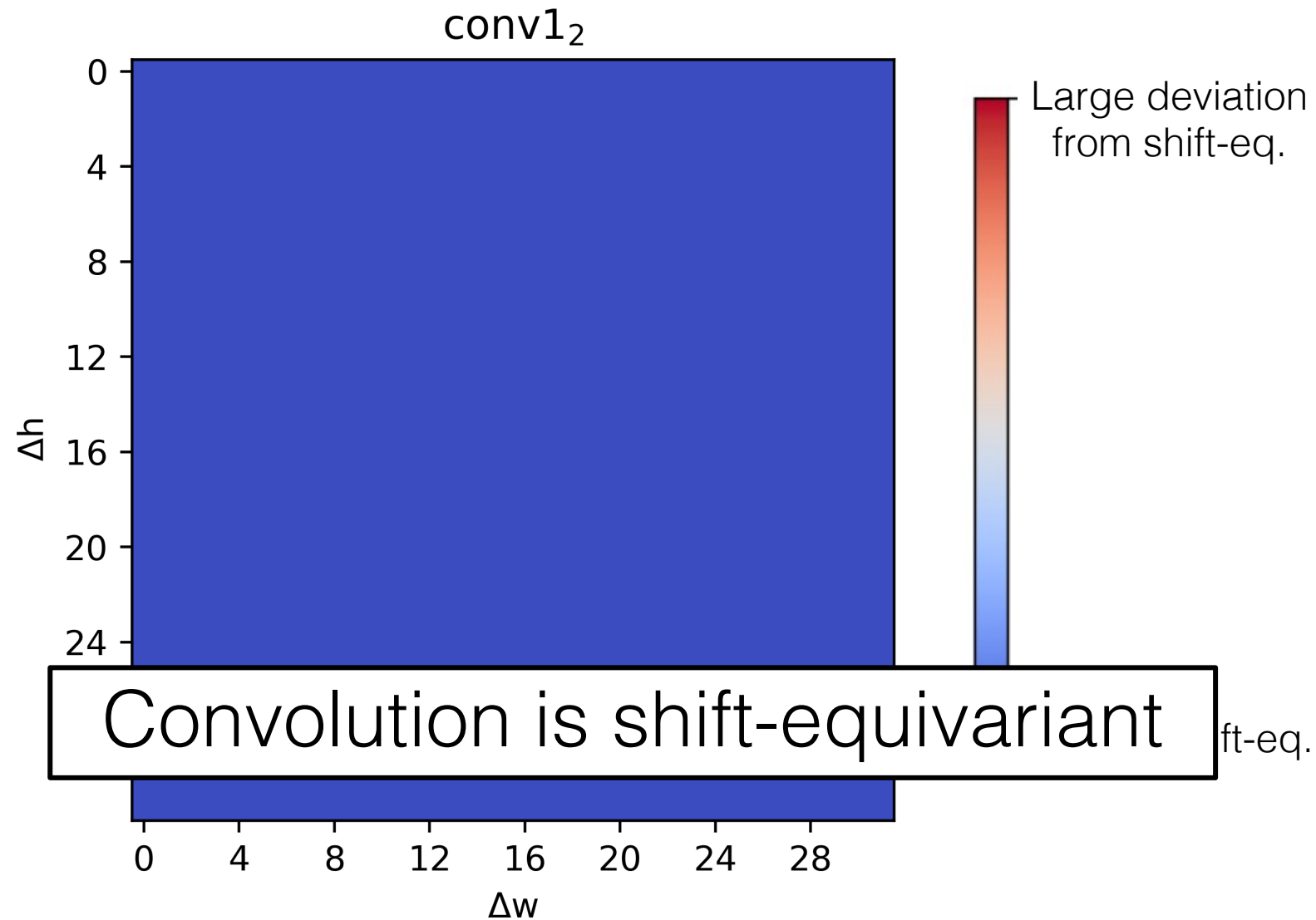
- CIFAR
- VGG network
  - 5 max-pools
- Test shift-equivariance condition
  - $\text{dist}(F(\text{Shift}_{\Delta h, \Delta w}(X)), \text{Shift}_{\Delta h, \Delta w}(F(X)))$



# Shift-equivariance, per layer

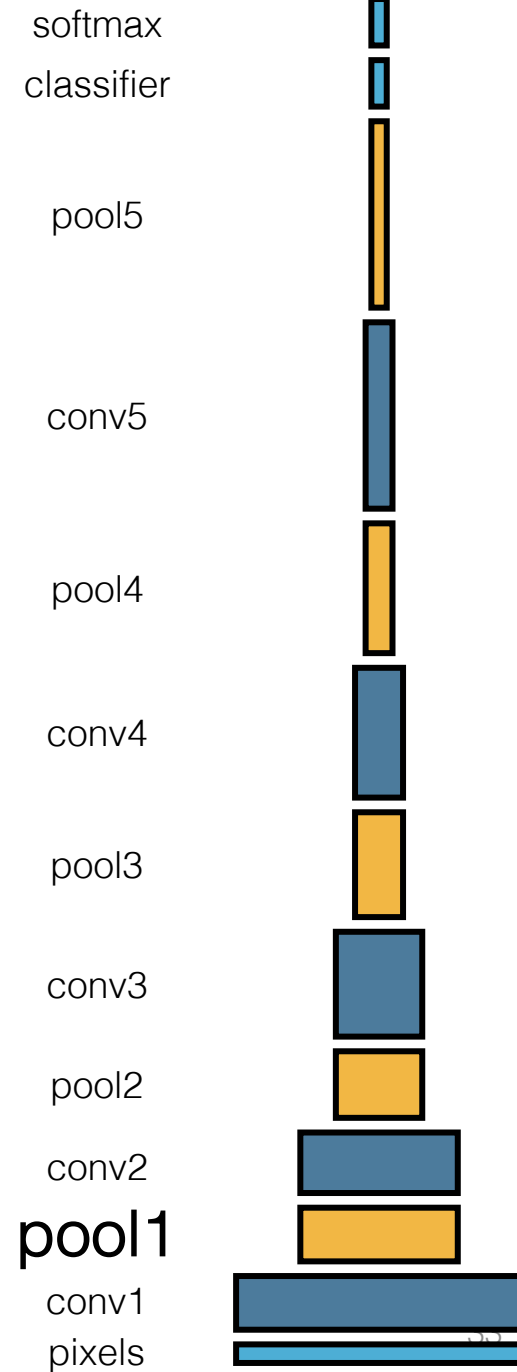
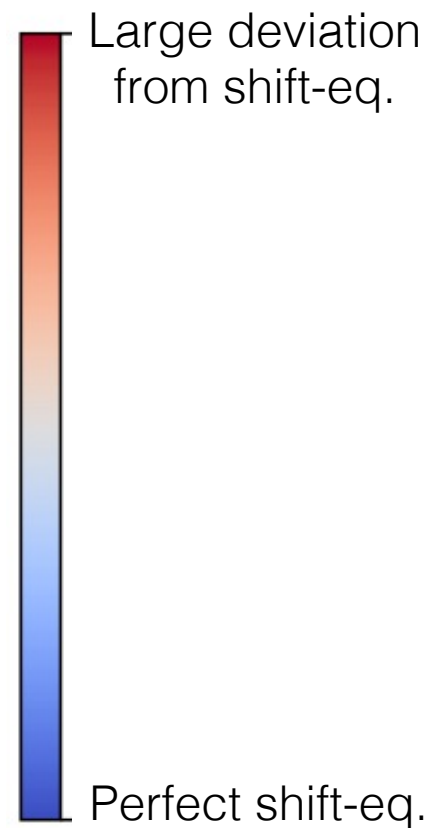
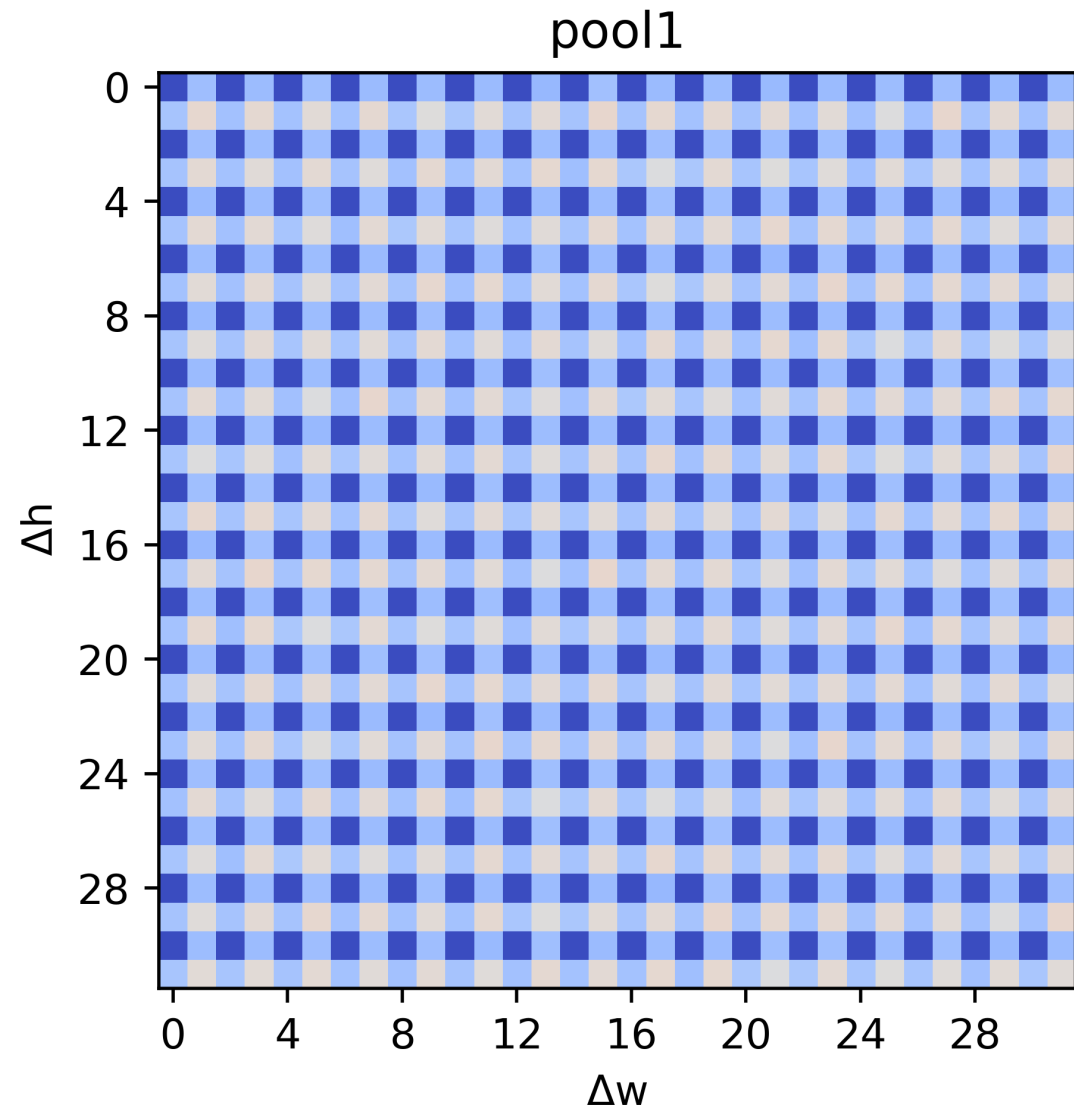


# Shift-equivariance, per layer

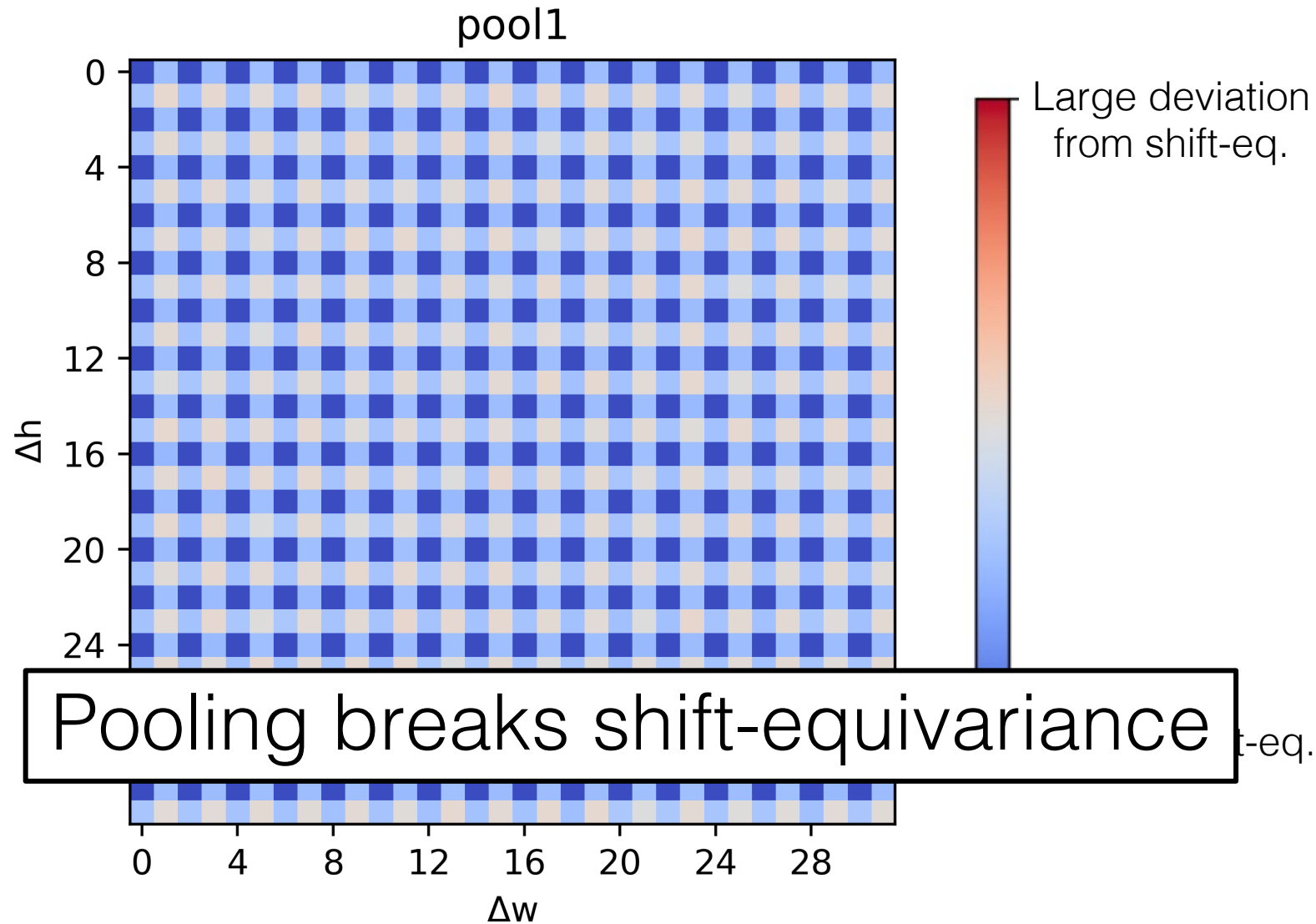




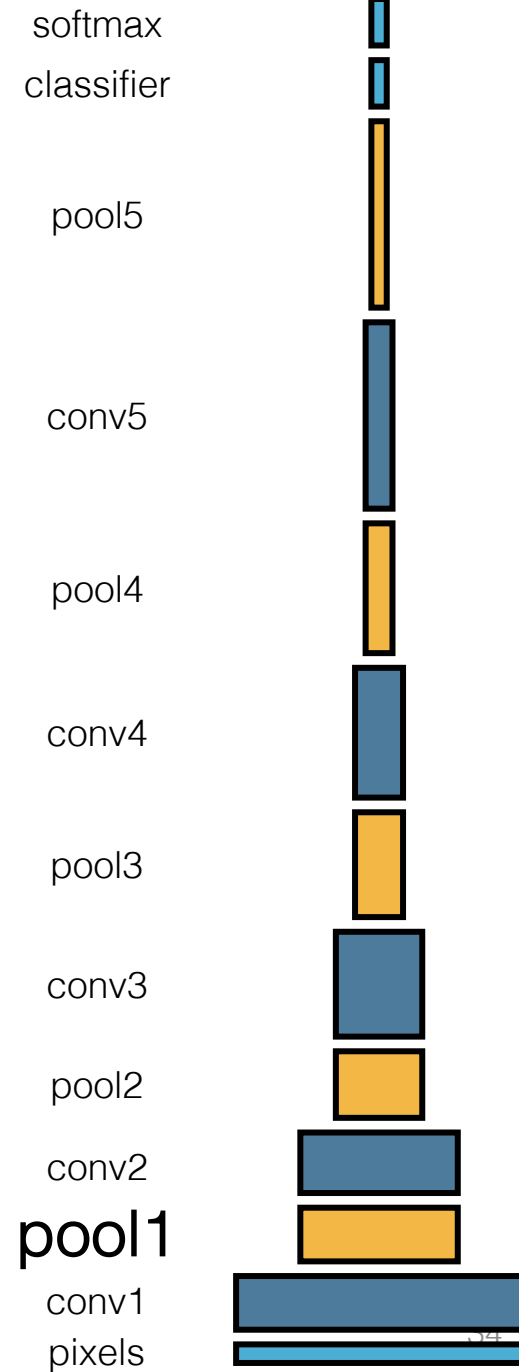
# Shift-equivariance, per layer



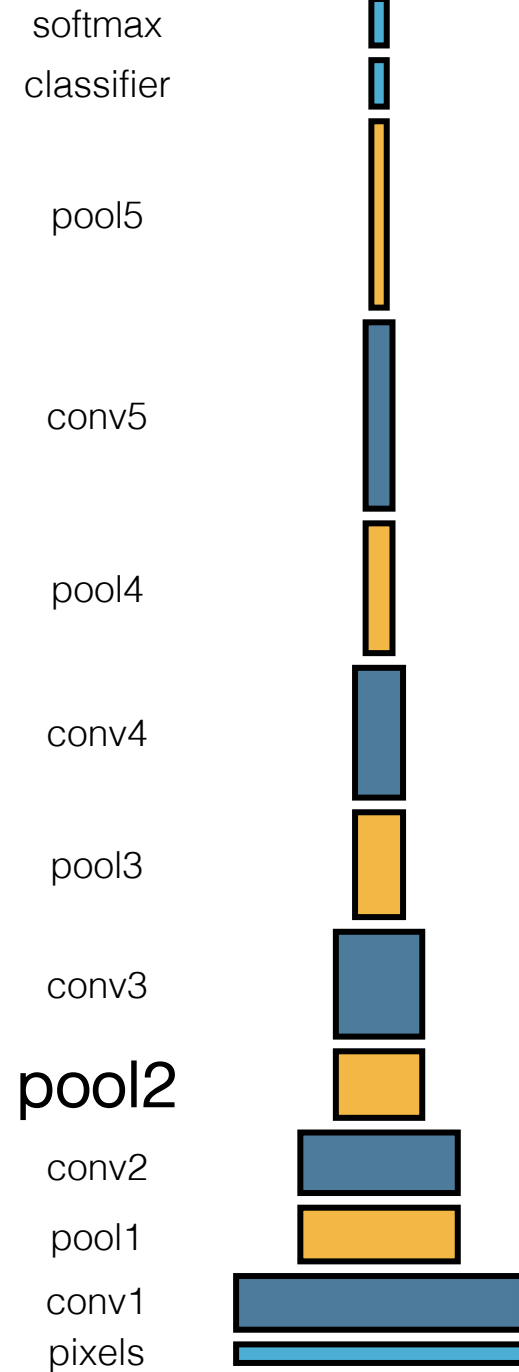
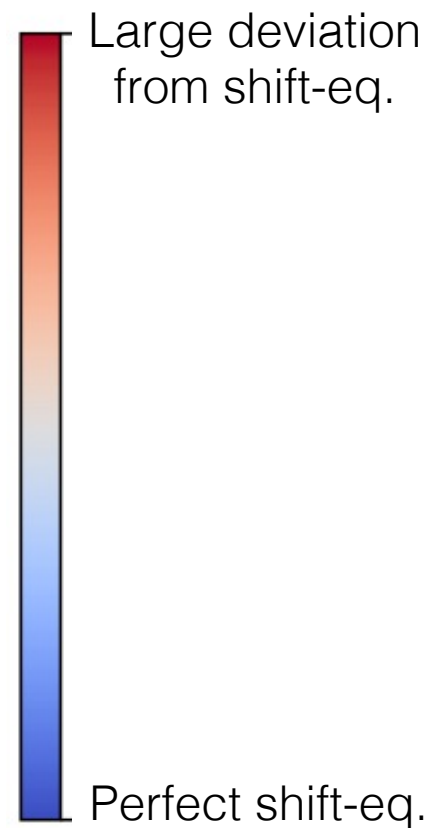
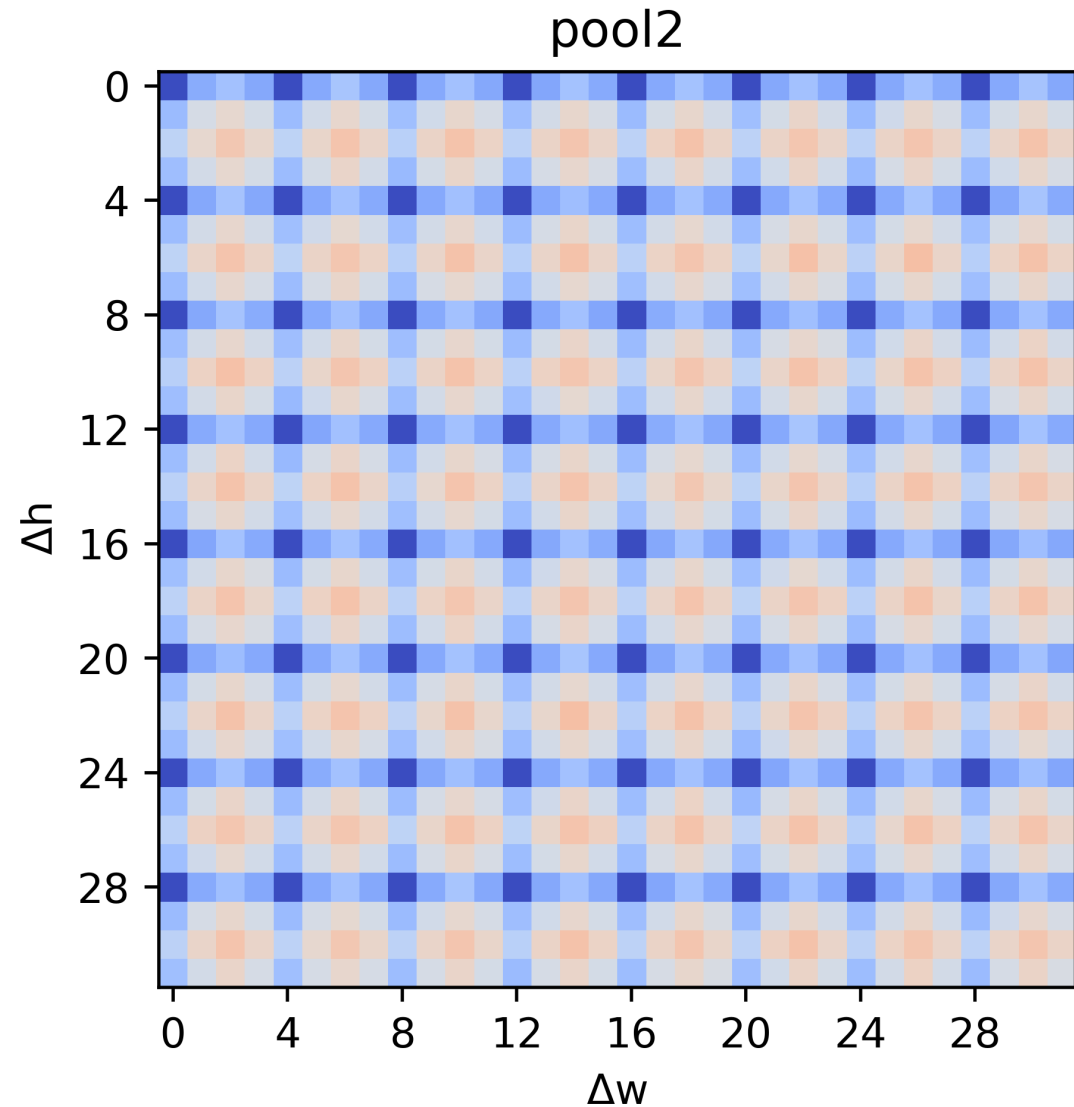
# Shift-equivariance, per layer



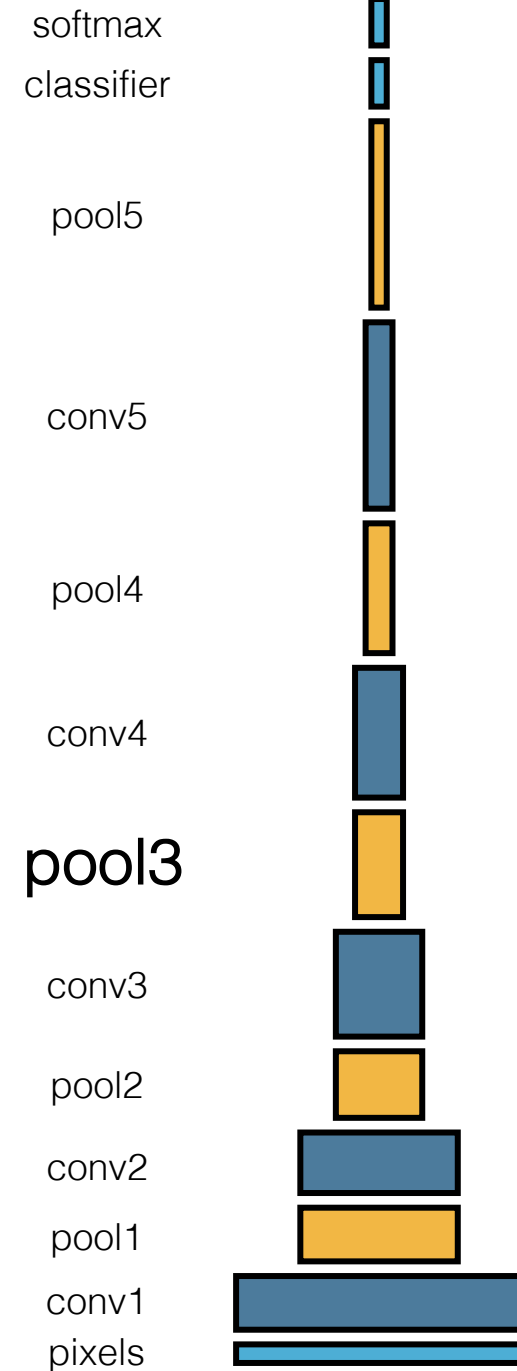
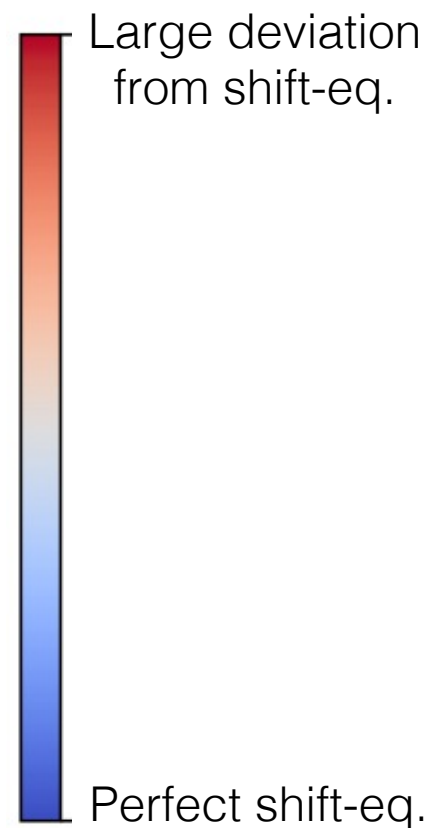
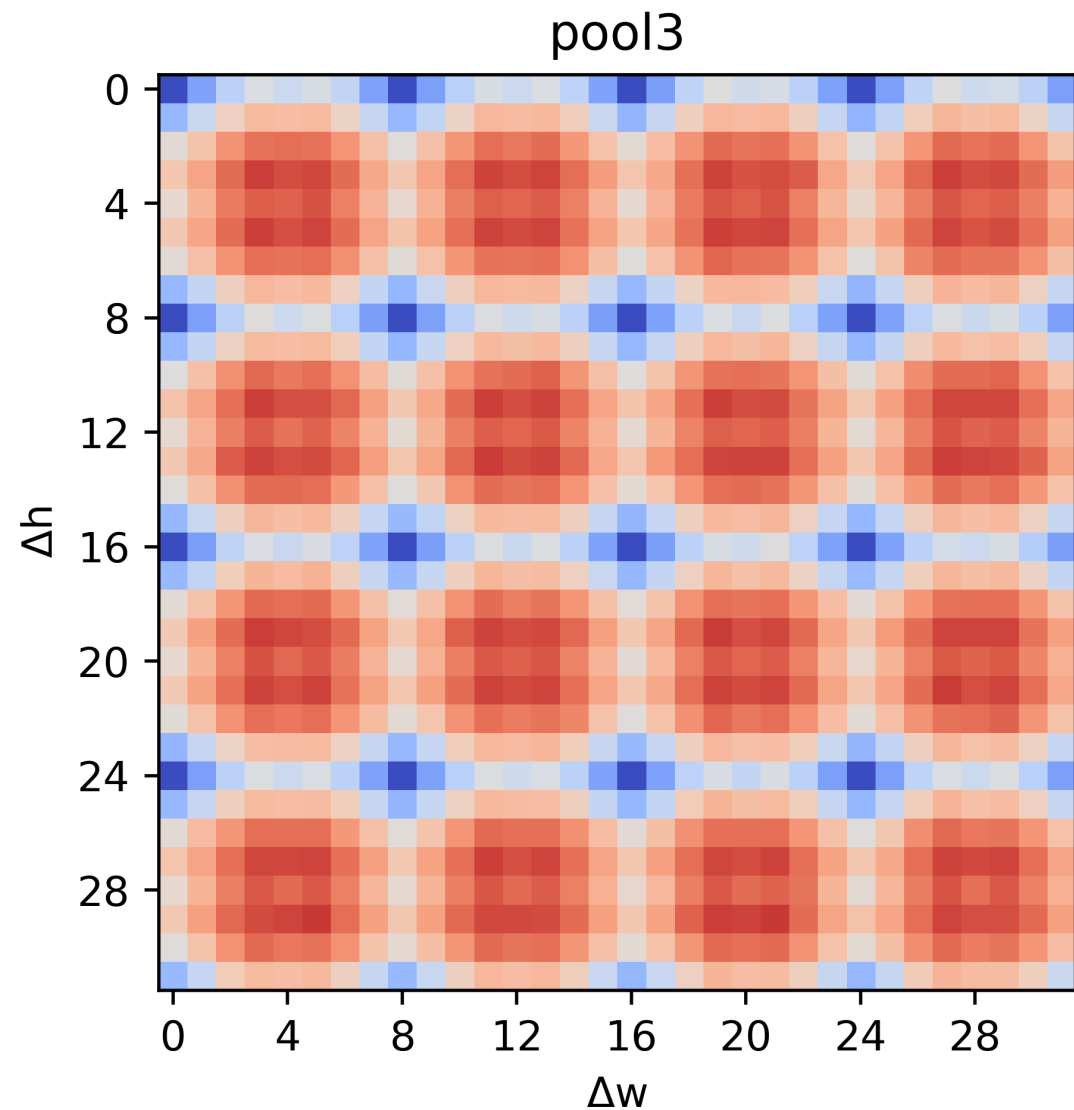
Pooling breaks shift-equivariance



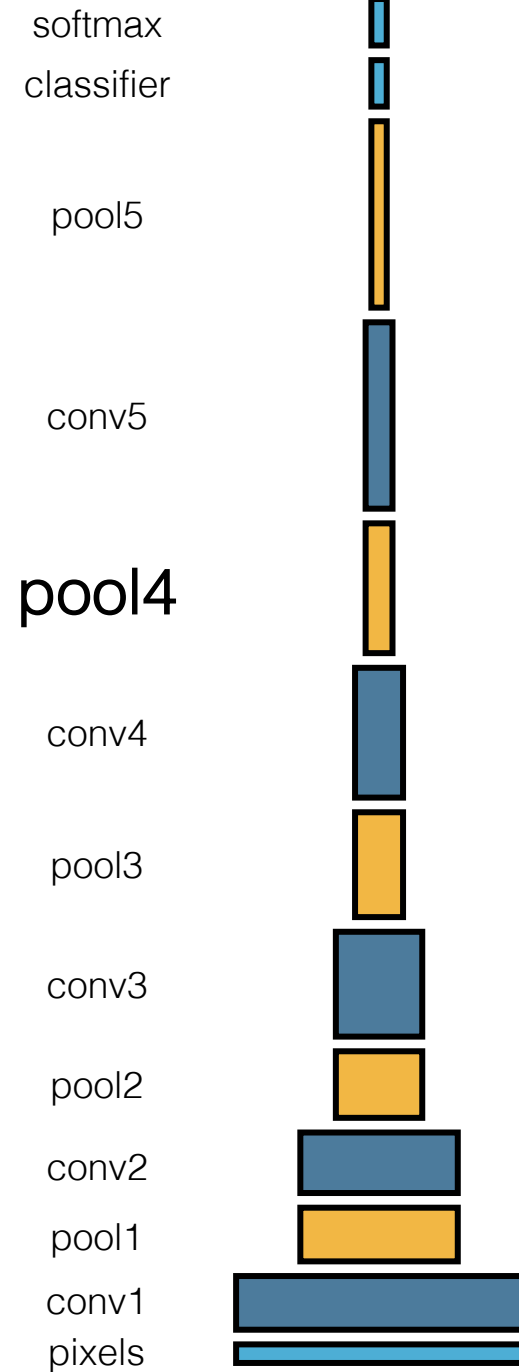
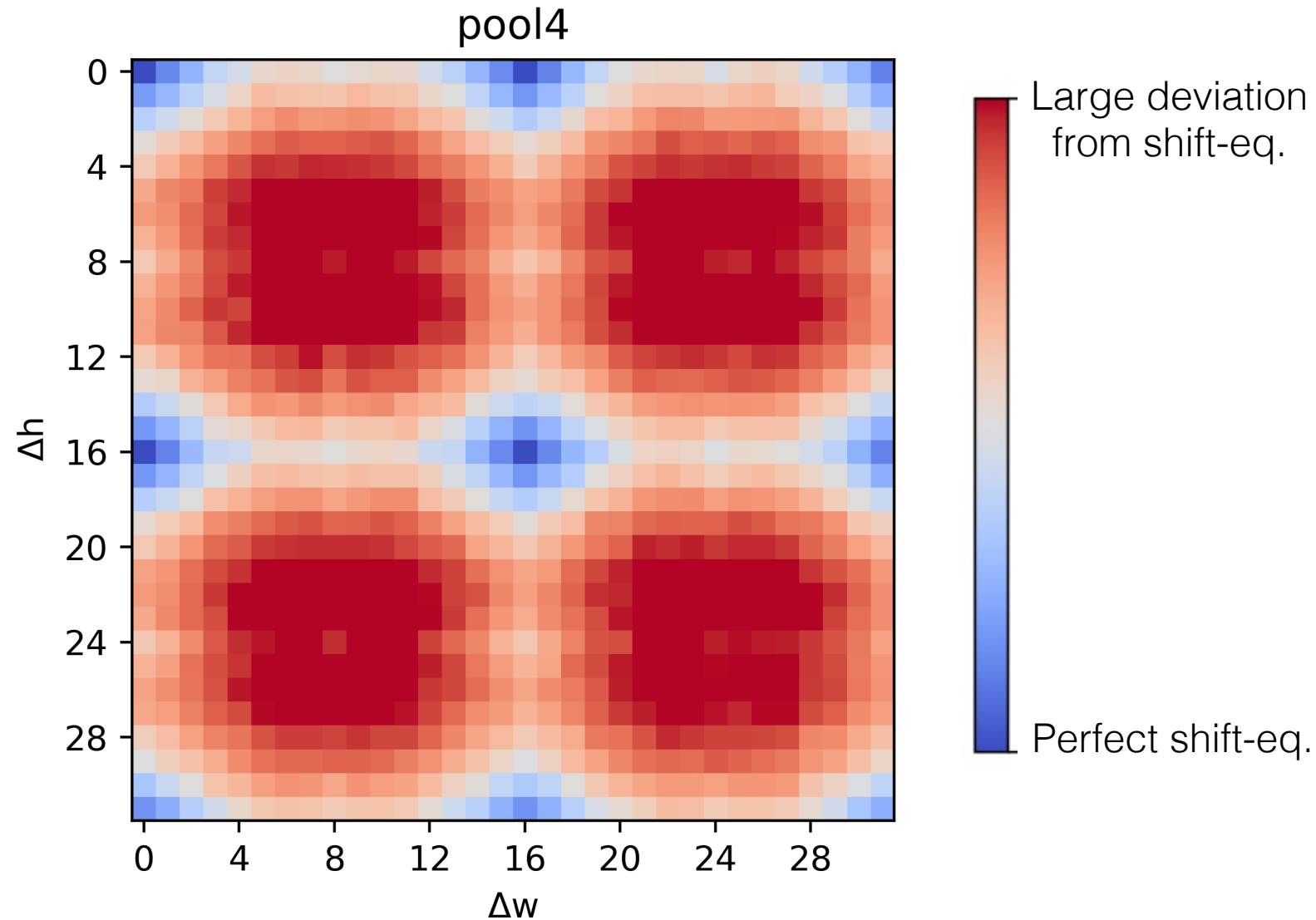
# Shift-equivariance, per layer



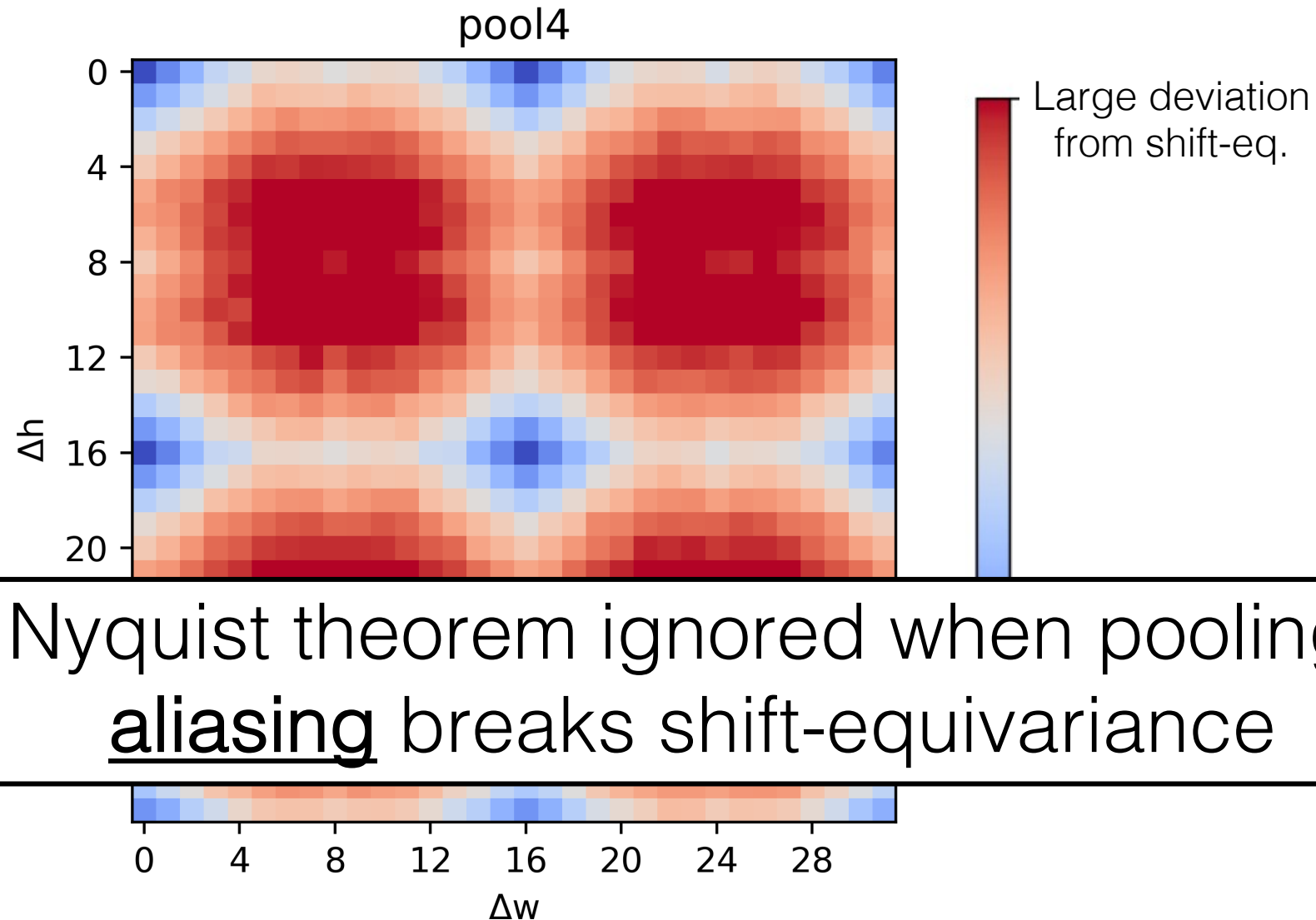
# Shift-equivariance, per layer



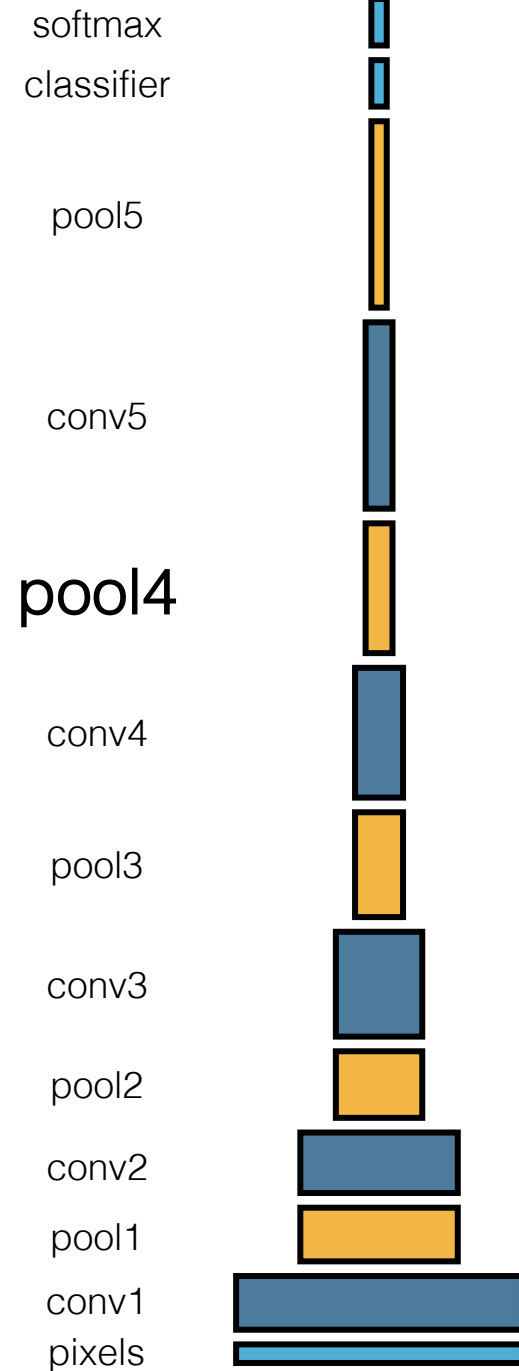
# Shift-equivariance, per layer



# Shift-equivariance, per layer



Nyquist theorem ignored when pooling;  
aliasing breaks shift-equivariance



# Alternative downsampling methods

- Blur+subsample
  - Antialiasing in signal processing; image processing; graphics

# Alternative downsampling methods

- Blur+subsample
  - Antialiasing in signal processing; image processing; graphics
- Max-pooling
  - Performs better in deep learning applications [Scherer 2010]

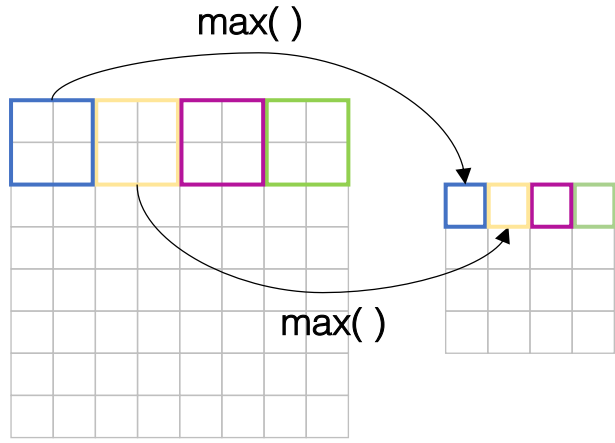


# Alternative downsampling methods

- Blur+subsample
  - Antialiasing in signal processing; image processing; graphics
- Max-pooling
  - Performs better in deep learning applications [Scherer 2010]

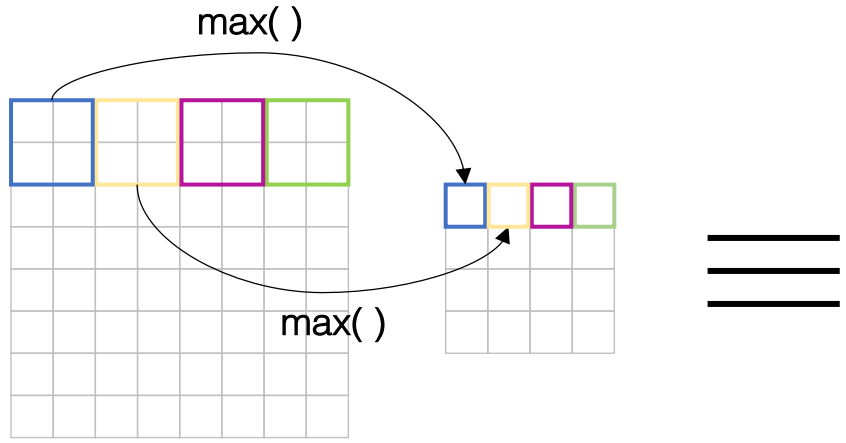
Reconcile antialiasing with max-pooling

Baseline  
(MaxPool)



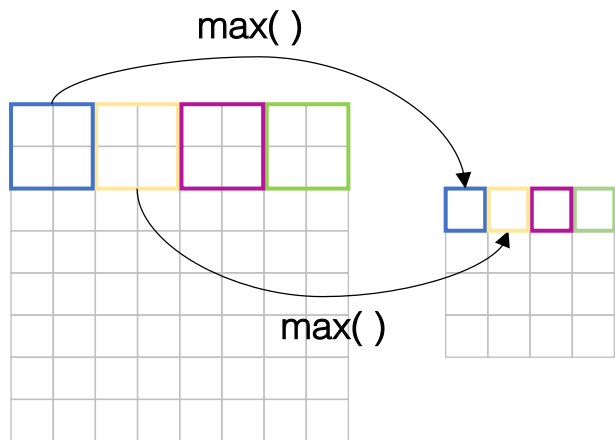
heavy aliasing

Baseline  
(MaxPool)



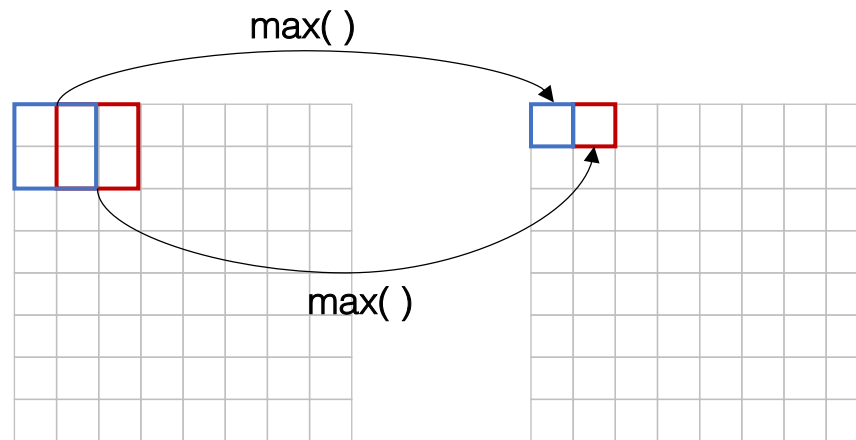
heavy aliasing

Baseline  
(MaxPool)



heavy aliasing

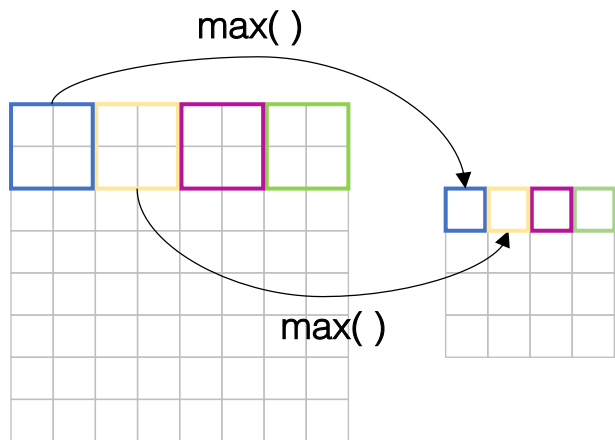
≡



(1) Max (dense evaluation)

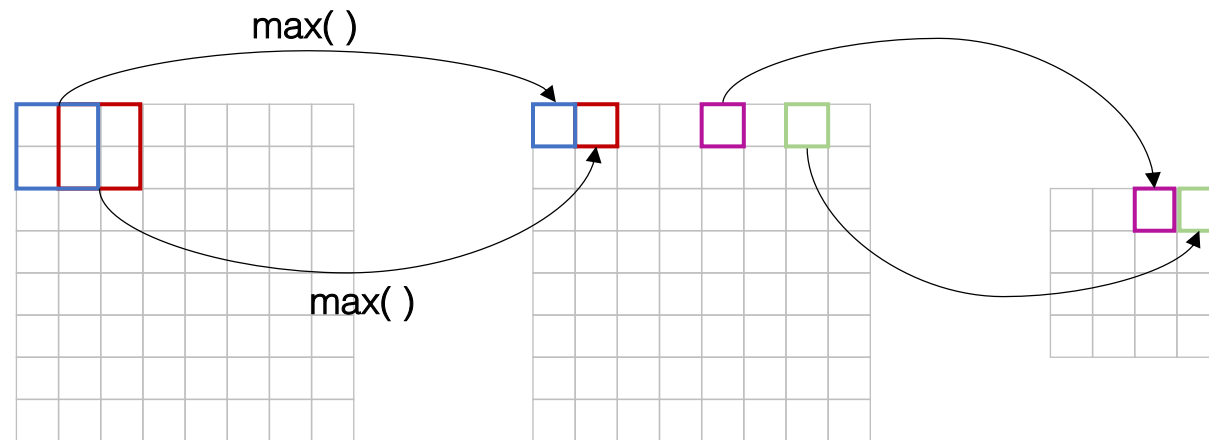
no aliasing

Baseline  
(MaxPool)



heavy aliasing

≡



(1) Max (dense evaluation)

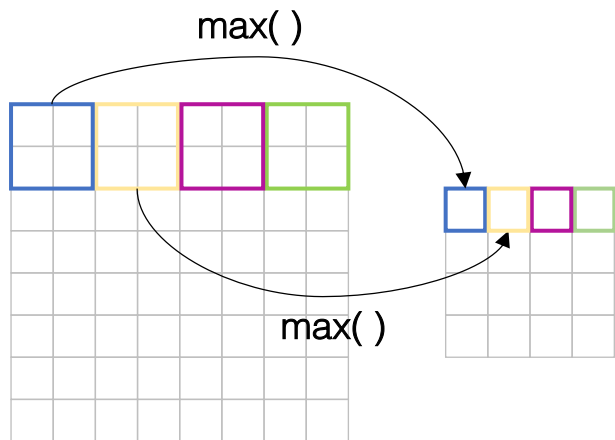
no aliasing

+

(2) Subsampling

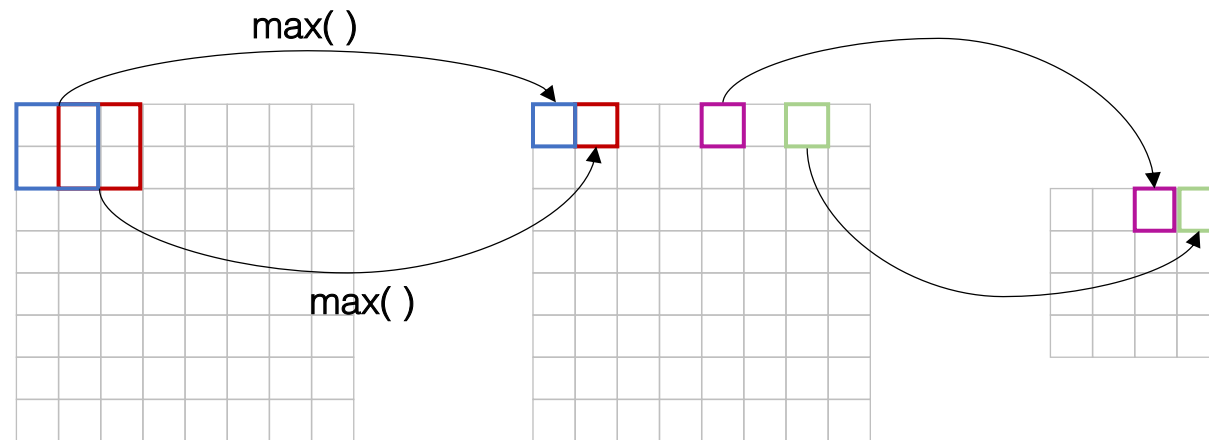
heavy aliasing

Baseline  
(MaxPool)



heavy aliasing

≡



(1) Max (dense evaluation)

no aliasing

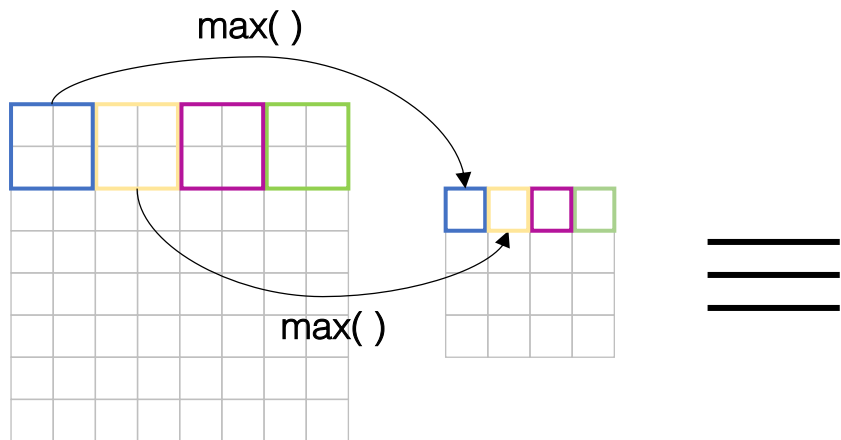
+

(2) Subsampling

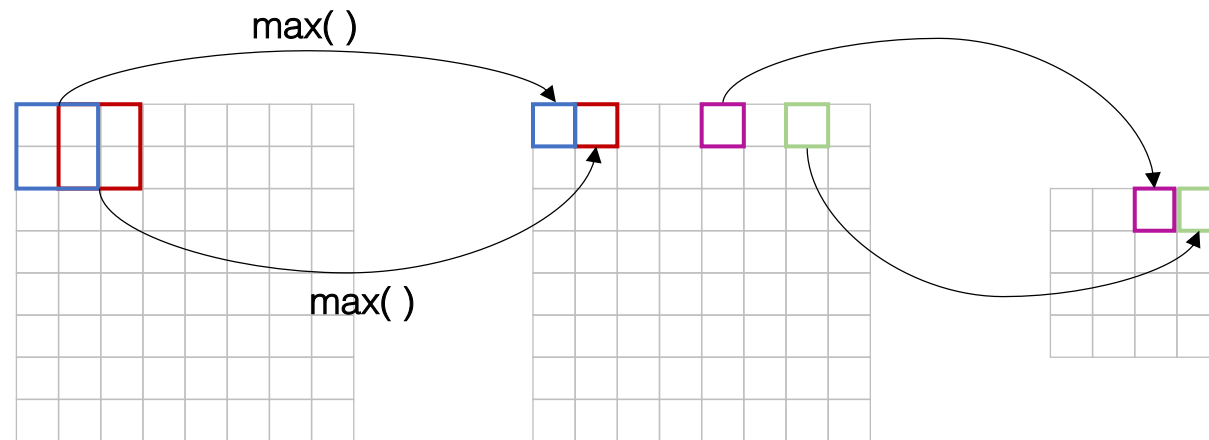
heavy aliasing



Baseline  
(MaxPool)

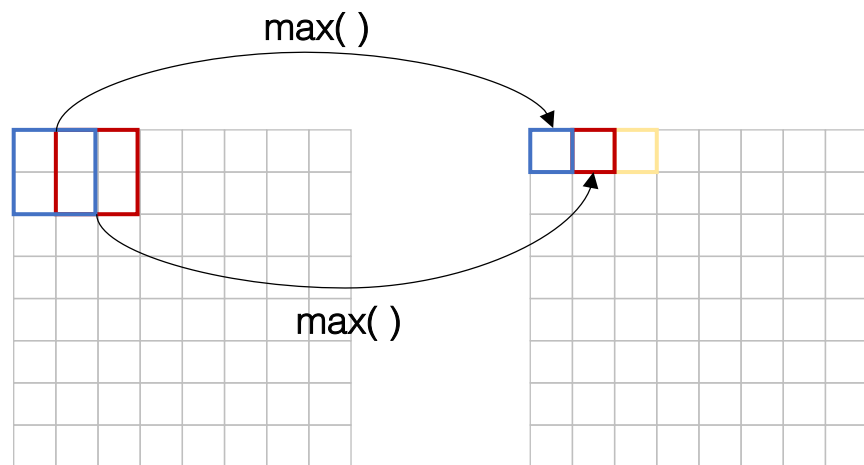


heavy aliasing



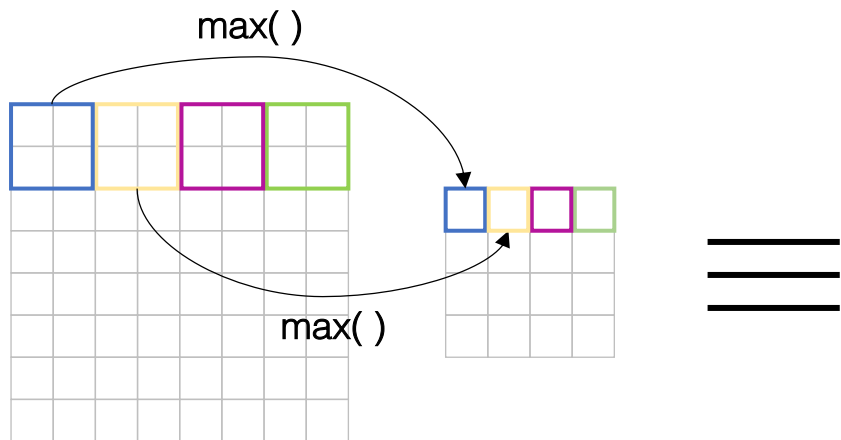
(1) Max (dense evaluation) + (2) Subsampling  
no aliasing heavy aliasing

Anti-aliased  
(MaxBlurPool)

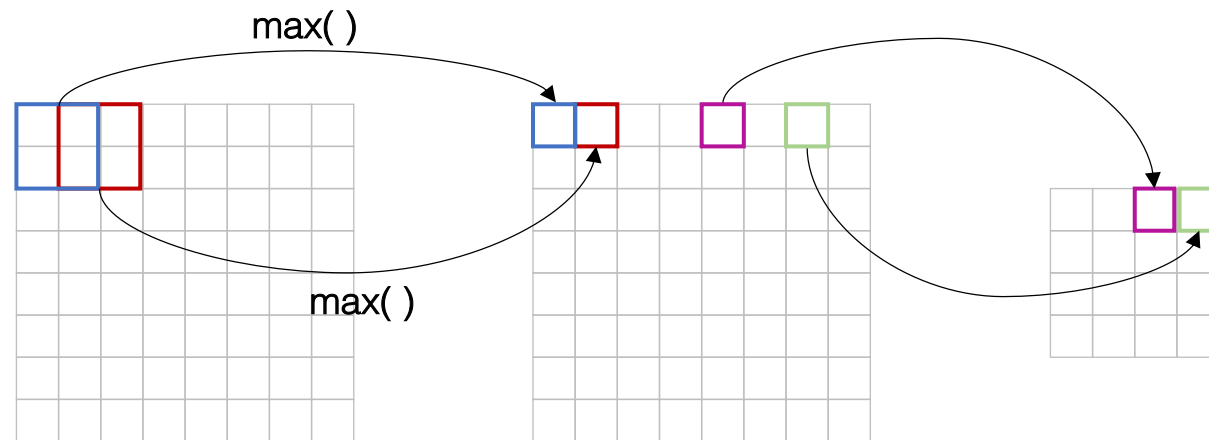


(1) Max (dense evaluation)  
no aliasing

Baseline  
(MaxPool)

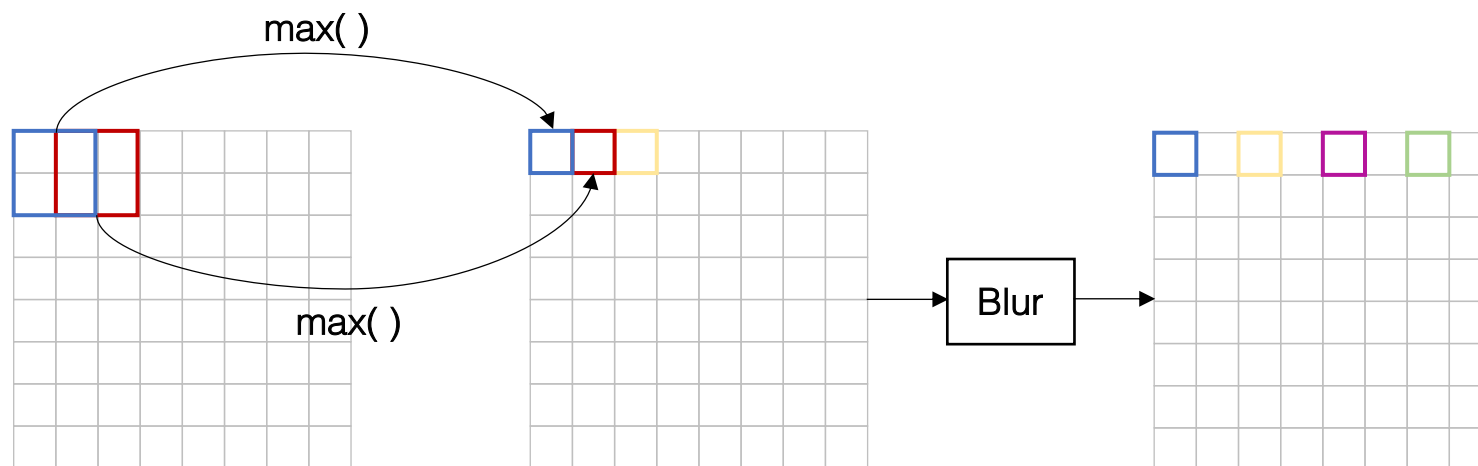


heavy aliasing



(1) Max (dense evaluation)  $+$  (2) Subsampling  
no aliasing heavy aliasing

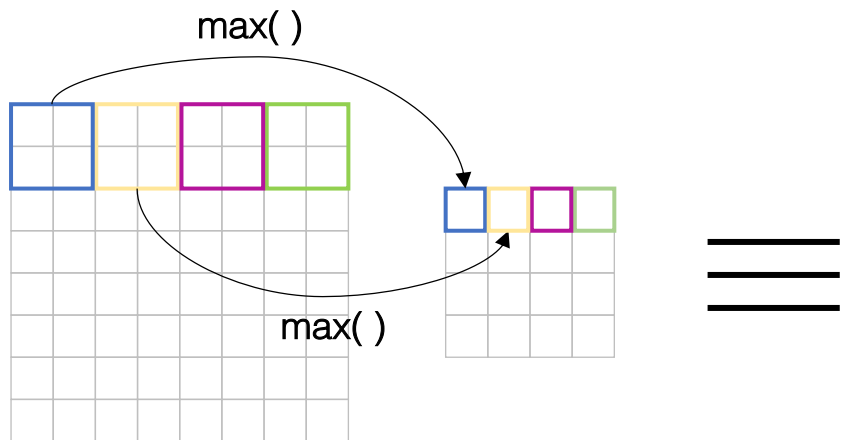
Anti-aliased  
(MaxBlurPool)



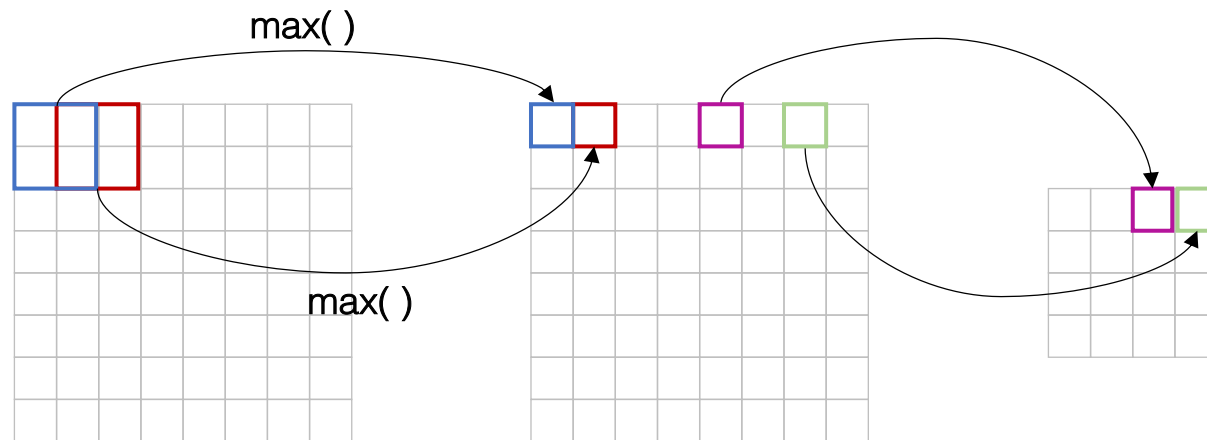
(1) Max (dense evaluation)  $+$  (2) Anti-aliasing filter  
no aliasing no aliasing



Baseline  
(MaxPool)

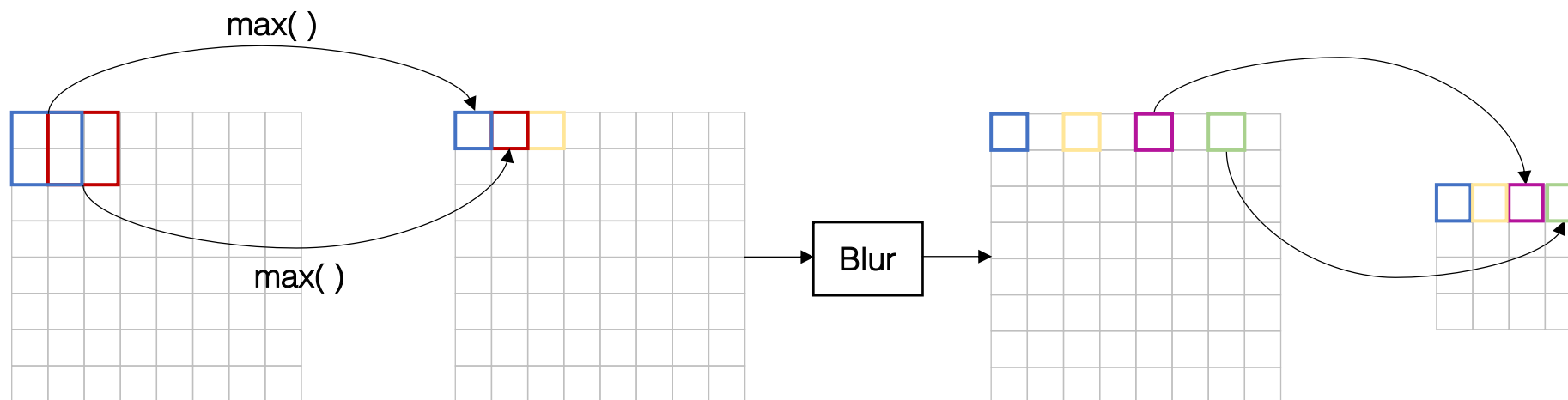


heavy aliasing



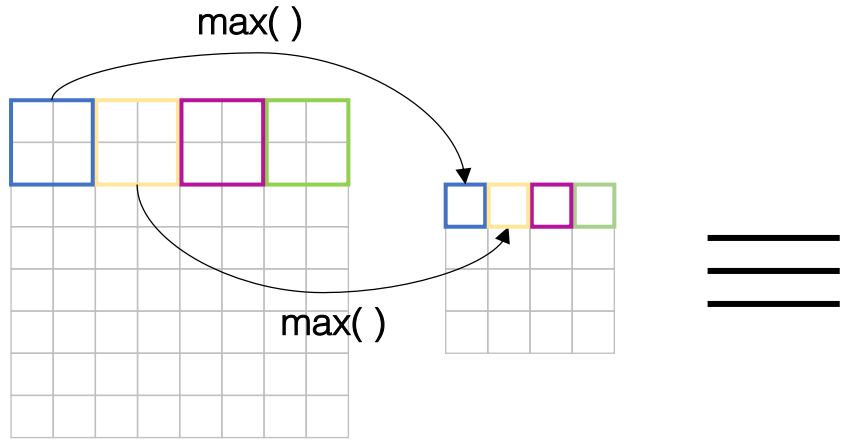
(1) Max (dense evaluation) + (2) Subsampling  
no aliasing heavy aliasing

Anti-aliased  
(MaxBlurPool)

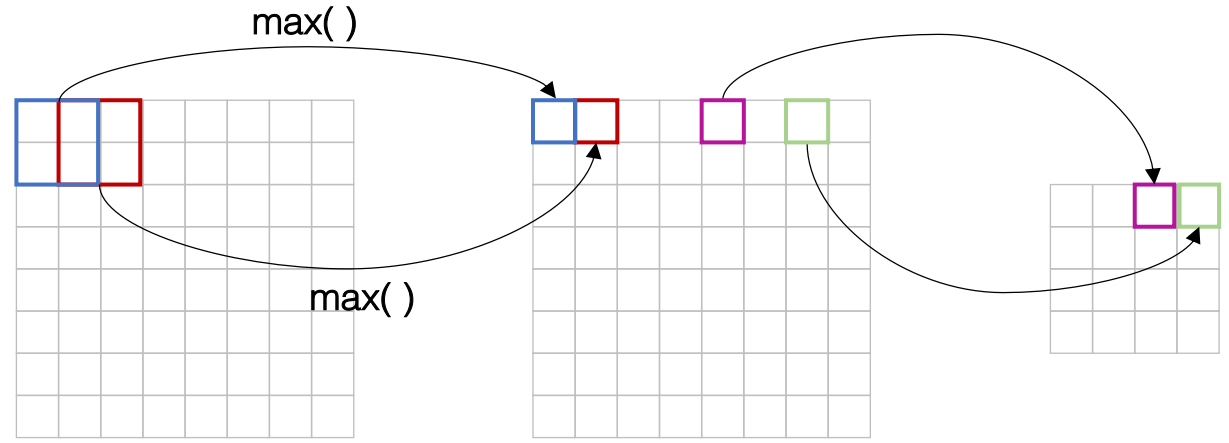


(1) Max (dense evaluation) + (2) Anti-aliasing filter + (3) Subsampling  
no aliasing no aliasing reduced aliasing

Baseline  
(MaxPool)

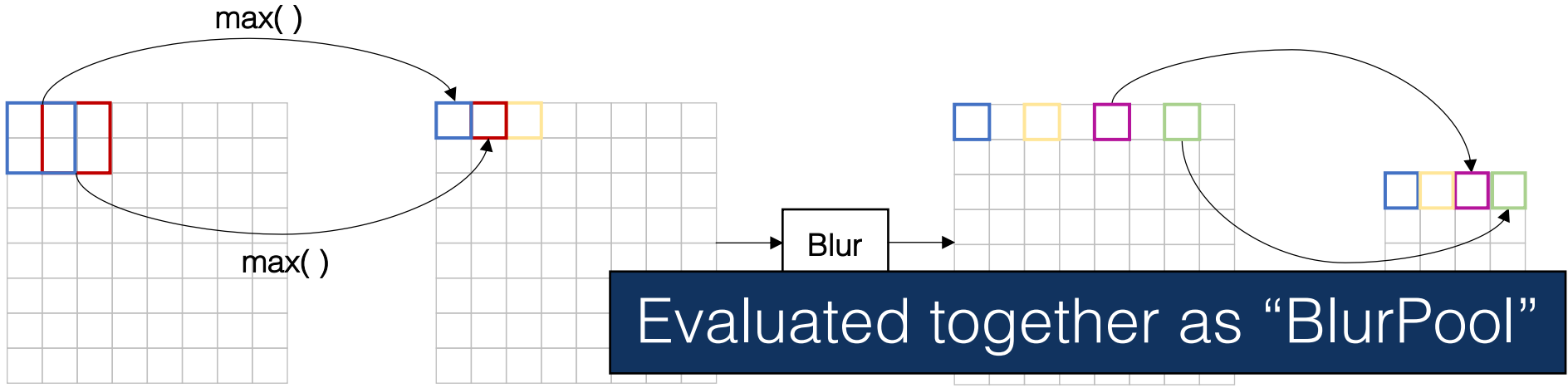


heavy aliasing



(1) Max (dense evaluation)  $+$  (2) Subsampling  
no aliasing heavy aliasing

Anti-aliased  
(MaxBlurPool)



(1) Max (dense evaluation)  $+$  (2) Anti-aliasing filter  $+$  (3) Subsampling  
no aliasing no aliasing reduced aliasing

# Antialiasing any downsampling layer

- Max Pool
  - VGG, Alexnet

# Antialiasing any downsampling layer

- Max Pool
  - VGG, Alexnet
- Strided Convolution
  - Resnet, MobileNetv2

# Antialiasing any downsampling layer

- Max Pool
  - VGG, Alexnet
- Strided Convolution
  - Resnet, MobileNetv2
- Average Pool
  - DenseNet

# ImageNet

Shift-invariance



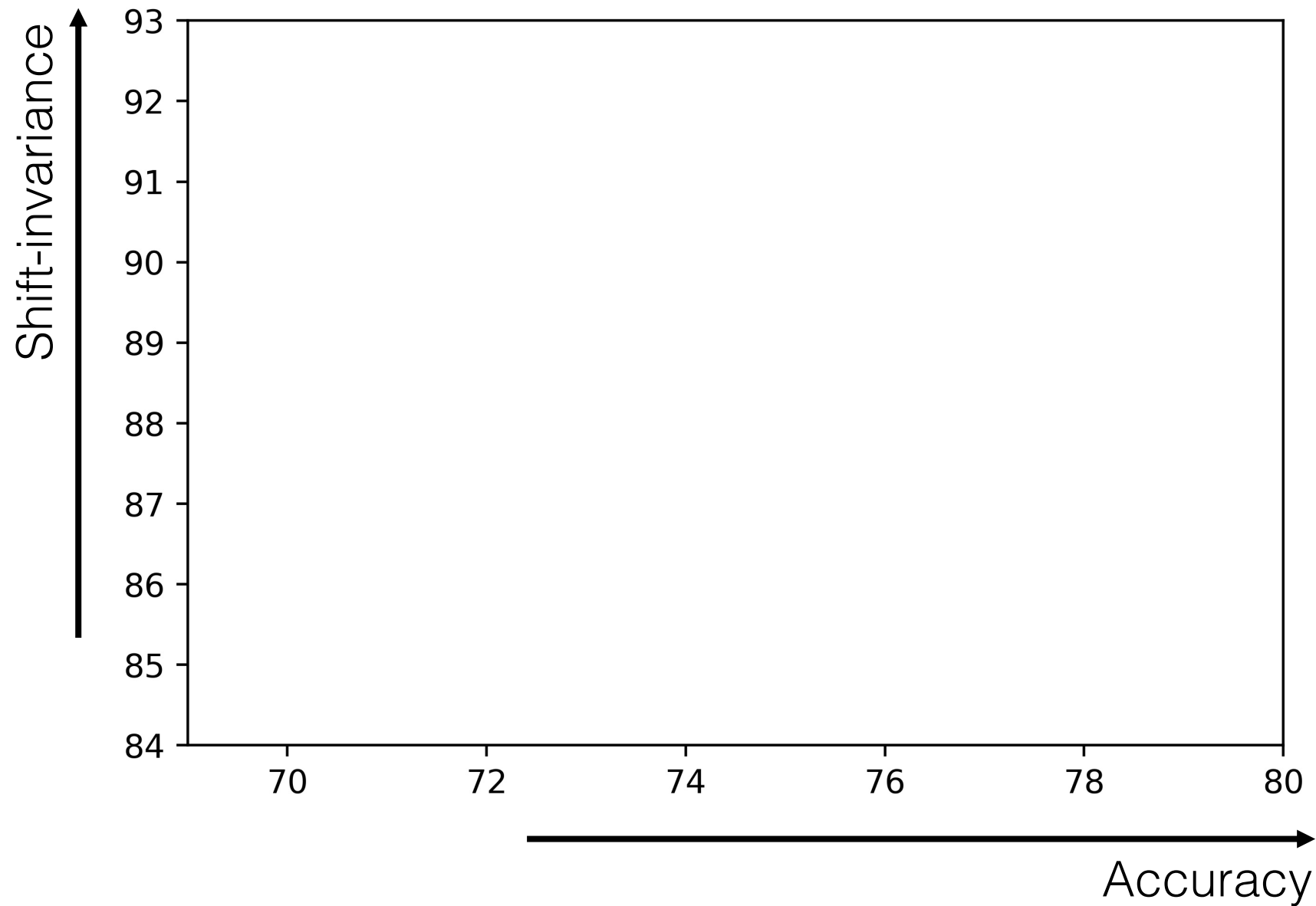
# ImageNet

Shift-invariance



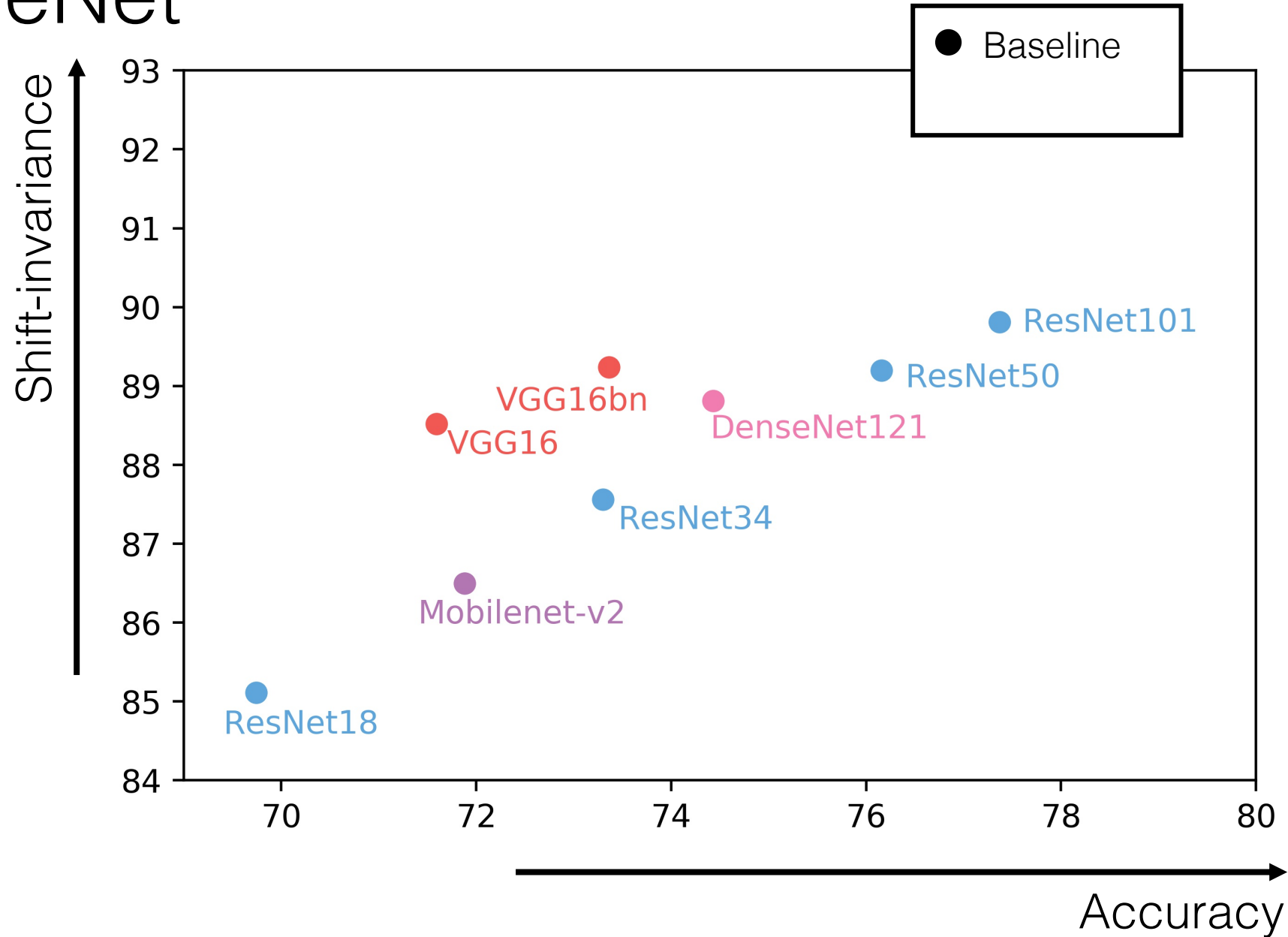
Accuracy

# ImageNet

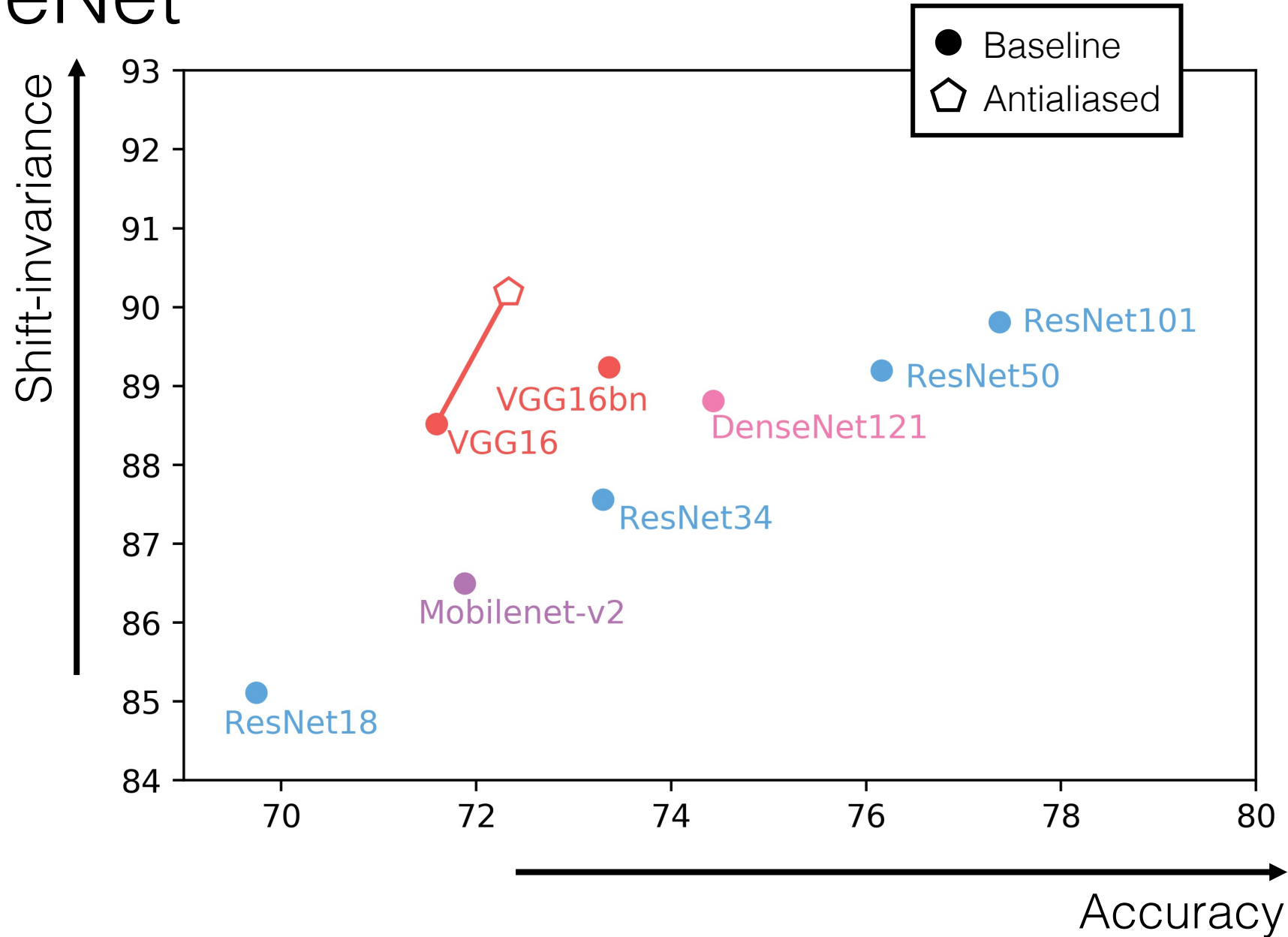




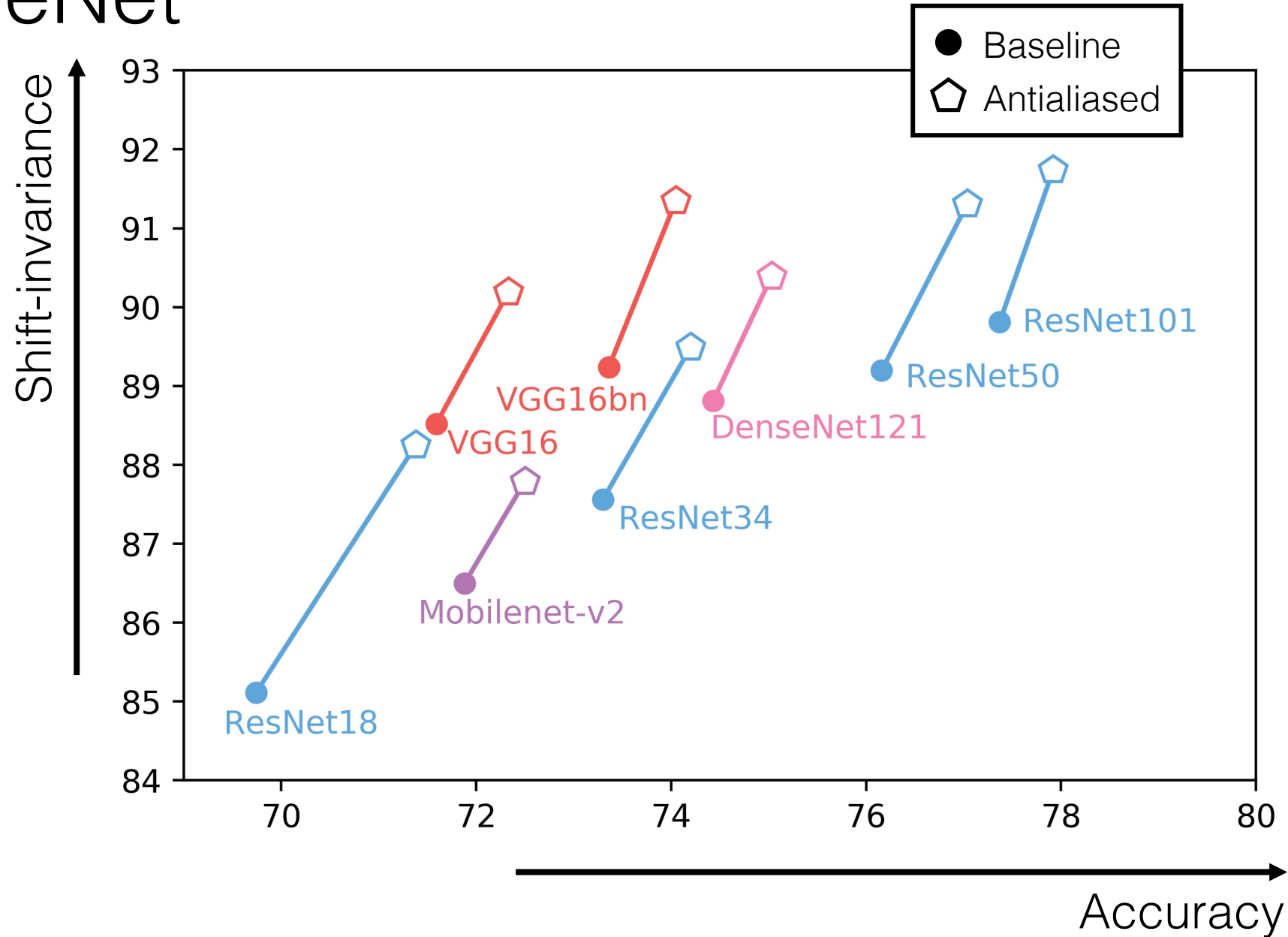
# ImageNet



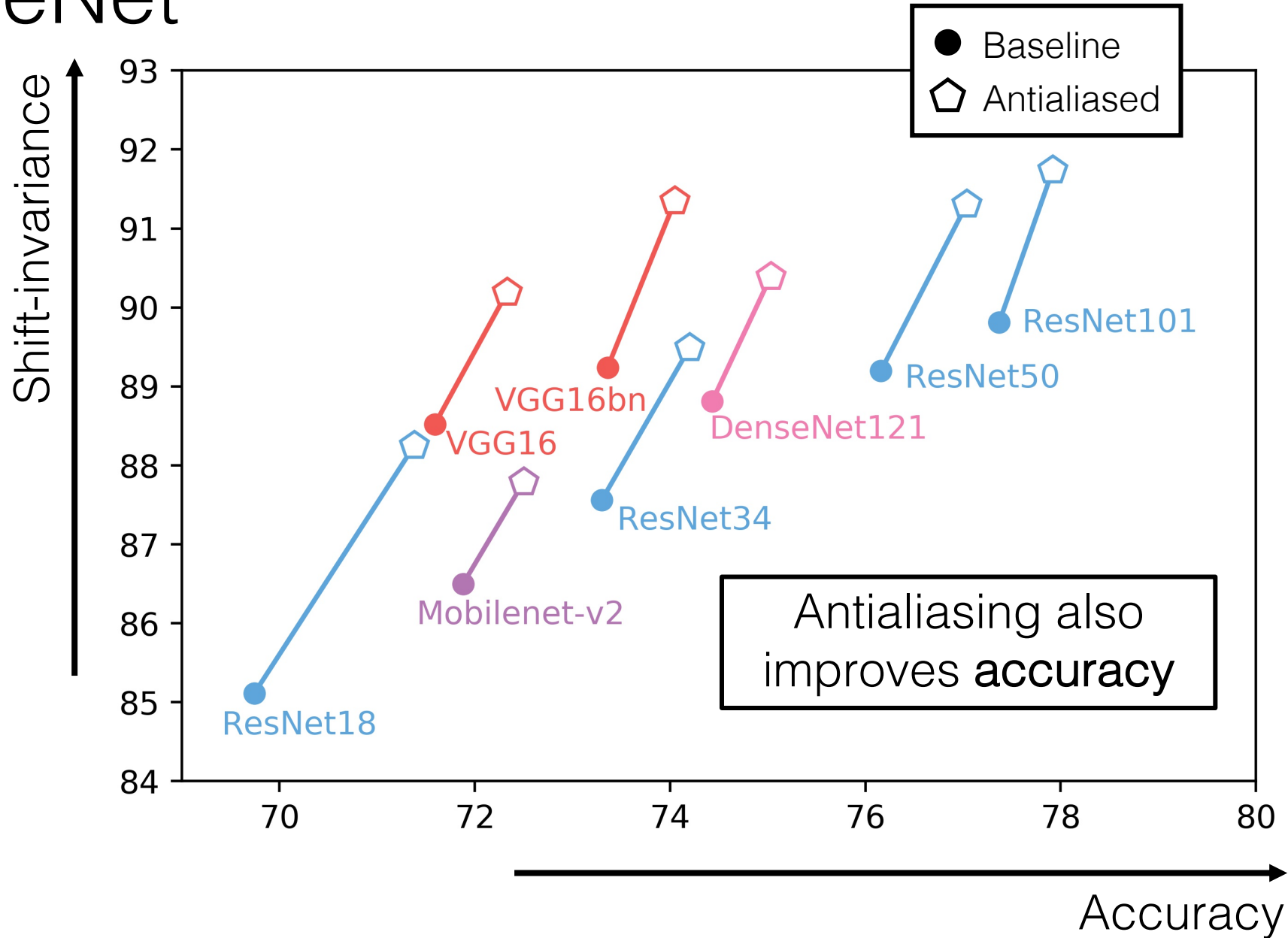
# ImageNet



# ImageNet



# ImageNet



# Discussion

Striding aliases (`stride=2`)

# Discussion

Striding aliases (`stride=2`)

Add antialiasing filter

# Discussion

Striding aliases (`stride=2`)

Add antialiasing filter

+ Improved shift-equivariance

# Discussion

Striding aliases (`stride=2`)

Add antialiasing filter

- + Improved shift-equivariance
- + Improved accuracy



# Discussion

Striding aliases (`stride=2`)

Add antialiasing filter

- + Improved shift-equivariance
- + Improved accuracy

Additionally

- + Improved stability to other perturbations

# Discussion

Striding aliases (`stride=2`)

Add antialiasing filter

- + Improved shift-equivariance
- + Improved accuracy

Additionally

- + Improved stability to other perturbations
- + Improved robustness

# Discussion

Striding aliases (`stride=2`)

Add antialiasing filter

- + Improved shift-equivariance
- + Improved accuracy

Additionally

- + Improved stability to other perturbations
- + Improved robustness

Antialiasing code, pretrained models  
<https://richzhang.github.io/antialiased-cnns/>

# Discussion

Striding aliases (`stride=2`)

Add antialiasing filter

- + Improved shift-equivariance
- + Improved accuracy

Additionally

- + Improved stability to other perturbations
- + Improved robustness

Antialiasing code, pretrained models  
<https://richzhang.github.io/antialiased-cnns/>

# Thank you!