On the Shift Invariance of Max Pooling Feature Maps in CNN

joint work with H. Leterme, K. Alahari et V. Perrier

Kévin Polisano
Convolutional neural networks

Convolutional neural networks

- Image classification: feature vectors are fed into a linear classifier

Convolutional neural networks

- Image classification: feature vectors are fed into a linear classifier
- Desired property of CNN: to remain invariant to small translations

Convolutional neural networks

- Image classification: feature vectors are fed into a linear classifier
- Desired property of CNN: to remain invariant to small translations
- Are extracted features maps stable to translations?

Are CNNs shift-invariant?

My cat Ada
Are CNNs shift-invariant?
Are CNNs shift-invariant?
Are CNNs shift-invariant?
Are CNNs shift-invariant?
Are CNNs shift-invariant?
Are CNNs shift-invariant?
Are CNNs shift-invariant?
Are CNNs shift-invariant?
Are CNNs shift-invariant?
Shift invariance

Input image $X$

Output $f(X) = 1$

Shifted input $\mathcal{T}_v X$

Output $f(\mathcal{T}_v X) = 1$
Shift invariance

Input image $X$

Output $f(X) = 1$

Shifted input $\mathcal{T}_v X$

Output $f(\mathcal{T}_v X) = 1$
How to make CNNs shift-invariant?

Training invariance by data augmentation

Input image X

Training set

'Cat'

Shifting

Rotating

Scaling
How to make CNNs shift-invariant?

- **Trained invariance** by data augmentation

Input image $X$

Training set

- Shifting
- Rotating
- Scaling

'Cat'
How to make CNNs shift-invariant?

- **Trained invariance** by data augmentation

![Diagram showing data augmentation with input image X, 'Cat', training set, shifting, rotating, scaling]
How to make CNNs shift-invariant?

- **Trained invariance** by data augmentation

![Diagram showing data augmentation on input images](image.png)
How to make CNNs shift-invariant?

- **Trained invariance** by data augmentation

![Diagram showing input image X being shifted, scaled, and rotated to train a 'Cat' classifier.](image-url)
How to make CNNs shift-invariant?

- Trained invariance by data augmentation

Input image $X$

Training set

- 'Cat'
- 'Cat'
- 'Cat'
- 'Cat'
How to make CNNs shift-invariant?

- **Trained invariance** by data augmentation

![Diagram showing data augmentation for training and testing a CNN]
How to make CNNs shift-invariant?

- Trained invariance by data augmentation

Input image $X$:
- Training set
- Test set

Test image:
- ‘Cat’

Figure: Shifting, rotating, and scaling transformations applied to the input image $X$ to train invariance.
How to make CNNs shift-invariant?

- **Trained invariance** by data augmentation
How to make CNNs shift-invariant?

- **Trained invariance** by data augmentation

![Diagram showing how data augmentation makes CNNs shift-invariant](Image)
How to make CNNs shift-invariant?

- **Online invariance at one-to-many locations**

**Training set**

Input image $X$

‘Cat’

Input image $X$

‘Cat’
How to make CNNs shift-invariant?

- **Online invariance** at one-to-many locations

Training set

Input image $X$ at one location

Input image $X$

'Cat'

At one location

(Images and text in the document are not provided, but the natural text is translated and formatted accordingly.)
How to make CNNs shift-invariant?

- **Online invariance** at one-to-many locations

**Training set**

Input image $X$

At one location

- 'Cat'
- 'Cat'
How to make CNNs shift-invariant?

- **Online invariance** at *one-to-many locations*

![Diagram showing input images and classification at one location for both training and test sets.](image-url)
How to make CNNs shift-invariant?

Online invariance at one-to-many locations

Input image $X$

'Cat'

At one location
How to make CNNs shift-invariant?

- **Online invariance** at one-to-many locations

![Diagram showing input images 'Cat' at different locations in training and test sets.](image-url)
How to make CNNs shift-invariant?

- Online invariance at one-to-many locations
How to make CNNs shift-invariant?

- Architectural online invariance

\[ \Phi(T_v x) \approx \Phi(x) \]

- What kind of linear and non-linear operators to consider?

- Are extracted features maps stable to translations?
How to make CNNs shift-invariant?

- **Architectural online invariance**

  J. Hinton, Y. LeCun

  Hierarchical invariance

- **Inductive biases**
  - Multichanneling
  - Weight sharing
  - Locality
  - Downsampling
Convolutional layers in CNN

Source: https://cs231n.github.io/convolutional-networks/
Convolutional layers in CNN

Source: https://cs231n.github.io/convolutional-networks/
Convolution & Max Pooling invariance?

Shift invariance

Input image $X$

Output $f(X) = 1$

Shifted input $\mathcal{T}_v X$

Output $f(\mathcal{T}_v X) = 1$
Shift invariance

Input image $X$

Output $f(X) = 1$

Shifted input $\mathcal{T}_v X$

Output $f(\mathcal{T}_v X) = 1$
Shift invariance

Input image $X$

Output $f(X) = 1$

Shifted input $\mathcal{T}_v X$

Output $f(\mathcal{T}_v X) = 1$

\textbf{Shift invariance $\neq$ Equivariance}
Shift equivariance
Shift equivariance

Input

Feature maps
Shift equivariance

\[ F_1 = \phi(X_1) \]

\[ F_2 = T(F_1) \]

\[ F_2 = \phi(X_2) \]

\[ X_1 \]

\[ X_2 = T(X_1) \]

Convolutions are shift-equivariant

![Diagram showing input, convolution, and output with shift-equivariance](image.png)
Convolutions are shift-equivariant
Convolutions are shift-equivariant
Convolutions are shift-equivariant
Convolutions are shift-equivariant

Input signal

Band-pass convolution kernel

Output signal
Convolutions are shift-equivariant

Input signal

Band-pass convolution kernel

Output signal

Shift equivariance, or covariance

Convolution operation
Convolutions are **shift-equivariant**

Source: https://www.doc.ic.ac.uk/~bkainz/teaching/DL/notes/equivariance.pdf
Convolutions are shift-equivariant
Max pooling layers

Source: https://cs231n.github.io/convolutional-networks/
Convolutions are followed by a max pooling
Convolutions are followed by a max pooling
Max pooling builds up shift-invariance
Invariance studies in CNN
The scattering transform builds shift-invariant feature vectors:

\[
\Phi(x) := S_J x = \left(\begin{array}{c}
x \ast \phi_{2J} \\
| x \ast \psi_{\lambda_1} | \ast \phi_{2J} \\
|| x \ast \psi_{\lambda_1} | \ast \psi_{\lambda_2} | \ast \phi_{2J} \\
| | | x \ast \psi_{\lambda_1} | \ast \psi_{\lambda_2} | \ast \psi_{\lambda_3} | \ast \phi_{2J} \\
\vdots \\
\lambda_1, \lambda_2, \lambda_3, \ldots
\end{array}\right)
\]
Invariance studies in CNN

The **scattering transform** builds shift-invariant feature vectors:

\[
\Phi(x) := S_J x = \begin{pmatrix}
x \ast \phi_{2J} \\
x \ast \psi_{\lambda_1} \ast \phi_{2J} \\
|x \ast \psi_{\lambda_1} \ast \psi_{\lambda_2} \ast \phi_{2J}| \\
|x \ast \psi_{\lambda_1} \ast \psi_{\lambda_2} \ast \psi_{\lambda_3} \ast \phi_{2J}| \\
\vdots \\
\lambda_1, \lambda_2, \lambda_3, \ldots
\end{pmatrix}
\]

Invariance studies in CNN

The scattering transform builds shift-invariant feature vectors:

\[ \Phi(x) := S_J x = \left( \begin{array}{c} x \ast \phi_{2J} \\ |x \ast \psi_{\lambda_1}| \ast \phi_{2J} \\ ||x \ast \psi_{\lambda_1}| \ast \psi_{\lambda_2}| \ast \phi_{2J} \\ |||x \ast \psi_{\lambda_1}| \ast \psi_{\lambda_2}| \ast \psi_{\lambda_3}| \ast \phi_{2J} | \\ \vdots \\ \lambda_1, \lambda_2, \lambda_3, \ldots \end{array} \right) \]

General deep convolutional neural networks also become more translation invariant with increasing network depth (proved in the continuous framework).

Invariance studies in CNN

- The **scattering transform** builds shift-invariant feature vectors:

\[
\Phi(x) := S_J x = \left(\begin{array}{c}
x \ast \phi_{2J} \\
|x \ast \psi_{\lambda_1} \ast \phi_{2J} \\
||x \ast \psi_{\lambda_1} \ast \psi_{\lambda_2} \ast \phi_{2J} \\
|||x \ast \psi_{\lambda_1} \ast \psi_{\lambda_2} \ast \psi_{\lambda_3} \ast \phi_{2J} \\
\vdots \\
\lambda_1, \lambda_2, \lambda_3, \ldots
\end{array}\right)
\]

- **General deep convolutional neural networks** also become more translation invariant with increasing network depth (proved in the *continuous framework*).


Invariance studies in CNN

- These results do not fully extend to the *discrete framework*
Invariance studies in CNN

- These results do not fully extend to the *discrete framework*

- **Strided** convolution and pooling operators may greatly diverge from shift invariance, due to **aliasing** when subsampling high-frequency signals

Source: https://community.sw.siemens.com/s/article/data-acquisition-anti-aliasing-filters
Invariance studies in CNN

- These results do not fully extend to the *discrete framework*

- **Strided** convolution and pooling operators may greatly diverge from shift invariance, due to **aliasing** when subsampling high-frequency signals


Source: https://community.sw.siemens.com/s/article/data-acquisition-anti-aliasing-filters
Invariance studies in CNN

- These results do not fully extend to the discrete framework.

- Strided convolution and pooling operators may greatly diverge from shift invariance, due to aliasing when subsampling high-frequency signals.

Subsampled convolution (retains one value out of two)

Input signal $\ast$ Band-pass convolution kernel $\downarrow 2 =$ Output signal
Invariance studies in CNN

- These results do not fully extend to the *discrete framework*

- **Strided convolution** and pooling operators may greatly diverge from shift invariance, due to **aliasing** when subsampling high-frequency signals.
Invariance studies in CNN

- These results do not fully extend to the *discrete framework*

- **Strided convolution** and pooling operators may greatly diverge from shift invariance, due to *aliasing* when subsampling high-frequency signals

![Diagram of convolution](image)

- *Input signal* 
- *Band-pass convolution kernel* 
- *Output signal*
Invariance studies in CNN

- These results do not fully extend to the *discrete framework*

- **Strided convolution** and pooling operators may greatly diverge from shift invariance, due to **aliasing** when subsampling high-frequency signals

Subsampling a high-frequency signal

- > **Aliasing effect**
- > **Instability to small input shifts**
Invariance studies in CNN

- These results do not fully extend to the *discrete framework*

- **Strided** convolution and **pooling operators** may greatly diverge from shift invariance, due to **aliasing** when subsampling high-frequency signals

---


Invariance studies in CNN

- These results do not fully extend to the *discrete framework*

- **Strided** convolution and pooling operators may greatly diverge from shift invariance, due to *aliasing* when subsampling high-frequency signals.


Invariance studies in CNN

- These results do not fully extend to the *discrete framework*

- **Strided** convolution and pooling operators may greatly diverge from shift invariance, due to *aliasing* when subsampling high-frequency signals


Invariance studies in CNN

- These results do not fully extend to the *discrete framework*

- **Strided** convolution and pooling operators may greatly diverge from shift invariance, due to *aliasing* when subsampling high-frequency signals

---


Invariance studies in CNN

- These results do not fully extend to the **discrete framework**

- **Strided** convolution and pooling operators may greatly diverge from shift invariance, due to **aliasing** when subsampling high-frequency signals

---


Invariance studies in CNN

- These results do not fully extend to the *discrete framework*

- **Strided** convolution and pooling operators may greatly diverge from shift invariance, due to **aliasing** when subsampling high-frequency signals

---


Invariance studies in CNN

- These results do not fully extend to the *discrete framework*

- **Strided** convolution and pooling operators may greatly diverge from shift invariance, due to *aliasing* when subsampling high-frequency signals

---


Invariance studies in CNN

These results do not fully extend to the **discrete framework**

**Strided** convolution and pooling operators may greatly diverge from shift invariance, due to **aliasing** when subsampling high-frequency signals

---


Invariance studies in CNN

- These results do not fully extend to the *discrete framework*

- **Strided** convolution and pooling operators may greatly diverge from shift invariance, due to *aliasing* when subsampling high-frequency signals


Invariance studies in CNN

- These results do not fully extend to the discrete framework.

- **Strided** convolution and pooling operators may greatly diverge from shift invariance, due to **aliasing** when subsampling high-frequency signals.


Invariance studies in CNN

- These results do not fully extend to the *discrete framework*

  - **Strided** convolution and pooling operators may greatly diverge from shift invariance, due to *aliasing* when subsampling high-frequency signals

---


Invariance studies in CNN

- These results do not fully extend to the *discrete framework*

- **Strided** convolution and pooling operators may greatly diverge from shift invariance, due to *aliasing* when subsampling high-frequency signals

---


Invariance studies in CNN

- These results do not fully extend to the \textit{discrete framework}.

- \textbf{Strided} convolution and pooling operators may greatly diverge from shift invariance, due to \textit{aliasing} when subsampling high-frequency signals.

\begin{itemize}
\end{itemize}
Invariance studies in CNN

- These results do not fully extend to the *discrete framework*

- **Strided** convolution and pooling operators may greatly diverge from shift invariance, due to **aliasing** when subsampling high-frequency signals


Invariance studies in CNN

- These results do not fully extend to the *discrete framework*

- **Strided** convolution and pooling operators may greatly diverge from shift invariance, due to *aliasing* when subsampling high-frequency signals


Invariance studies in CNN

- These results do not fully extend to the *discrete framework*

- **Strided** convolution and pooling operators may greatly diverge from shift invariance, due to **aliasing** when subsampling high-frequency signals

---


---

Invariance studies in CNN

- These results do not fully extend to the *discrete framework*

- **Strided** convolution and pooling operators may greatly diverge from shift invariance, due to *aliasing* when subsampling high-frequency signals


Invariance studies in CNN

- These results do not fully extend to the *discrete framework*

- **Strided** convolution and pooling operators may greatly diverge from shift invariance, due to **aliasing** when subsampling high-frequency signals

---


Invariance studies in CNN

- These results do not fully extend to the *discrete framework*

- Strided convolution and pooling operators may greatly diverge from shift invariance, due to *aliasing* when subsampling high-frequency signals


Invariance studies in CNN

- These results do not fully extend to the *discrete framework*

- **Strided** convolution and pooling operators may greatly diverge from shift invariance, due to **aliasing** when subsampling high-frequency signals

---


Aliasing breaks shift-invariance

Aliasing breaks shift-invariance

Blind spots in the shift-invariance studies
Blind spots in the shift-invariance studies

The **effect of the max pooling operator on network stability** under small input shifts has not been investigated, particularly when used in combination with Gabor-like convolutions.
Blind spots in the shift-invariance studies

- The effect of the max pooling operator on network stability under small input shifts has not been investigated, particularly when used in combination with Gabor-like convolutions.

- In the discrete case, the presence of subsampled convolutions with oriented band-pass filters can lead to aliasing artifacts. To our knowledge, the literature lacks theoretical studies that take these aliasing effects into account.
Blind spots in the shift-invariance studies

The **effect of the max pooling operator on network stability** under small input shifts has not been investigated, particularly when used in combination with Gabor-like convolutions.

In the discrete case, the presence of subsampled convolutions with oriented band-pass filters can lead to aliasing artifacts. To our knowledge, **the literature lacks theoretical studies that take these aliasing effects into account**.

Although extensive studies have been conducted on complex-valued convolutions followed by modulus, **a link is missing to extend these results to standard CNNs**, which implement real-valued convolutions and spatial pooling operators.
Focus on the first layer
Focus on the first layer

Input → Conv ↓ $m$ → Bias → ReLU → MaxPool ↓ 2 → ... → Classifier → Output

Affine operators
Nonlinear activation function
Nonlinear pooling operator
Focus on the first layer
Focus on the first layer
Focus on the first layer

Example: AlexNet (2012)
Focus on the first layer

Example: AlexNet (2012)

Band-pass “Gabor-like” filters
Focus on the first layer

Example: AlexNet (2012)


Subsampled convolutions, a real problem!

Input signal

\[ \begin{pmatrix}
1 \\
\vdots \\
-1
\end{pmatrix} \]

\[ \ast \]

Band-pass convolution kernel

\[ \begin{pmatrix}
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix} \]

Output signal

\[ \downarrow \ 2 \ = \]

Subsampled convolution
(retenets one value out of two)

\[ \mathbb{R}^{16} \rightarrow \text{Conv} \downarrow 2 \rightarrow \mathbb{R}^{8} \rightarrow \text{Output} \]
Subsampled convolutions, a real problem!

Input signal

Band-pass convolution kernel

Output signal

\[
\begin{pmatrix}
1 \\
0 \\
-1
\end{pmatrix}
\ast
\begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}
\downarrow 2
= 
\begin{pmatrix}
0.5 \\
0.5 \\
0.5 \\
0.5 \\
0.5 \\
0.5 \\
0.5 \\
0.5
\end{pmatrix}
\]
Subsampled convolutions, a real problem!

\[
\begin{array}{c}
\text{Input signal} \\
\end{array} \quad \star \quad \begin{array}{c}
\text{Band-pass convolution kernel} \\
\end{array} \quad \begin{array}{c}
\downarrow 2 \\
\end{array} = \begin{array}{c}
\text{Output signal} \\
\end{array}
\]
Subsampled convolutions, a real problem!

Subsampling a high-frequency signal

→ Aliasing effect

→ Instability to small input shifts
Subsampled convolutions, a real problem!

Input signal

Band-pass convolution kernel

Output signal

\[ \mathbb{R}^{16} \xrightarrow{\text{Conv} \downarrow 2} \mathbb{R}^{8} \xrightarrow{\text{MaxPool} \downarrow 2} \mathbb{R}^{4} \xrightarrow{\text{Output}} \]
Subsampled convolutions, a real problem!

Input signal

Band-pass convolution kernel

Output signal

Stabilizing power?
Complex-valued convolutions at rescue

Input signal

\[ \mathbb{C} \text{Conv} \downarrow 4 \rightarrow \mathbb{C}^4 \rightarrow \text{Modulus} \rightarrow \mathbb{R}^4 \rightarrow \text{Output} \]

\[ \mathbb{C} \text{Mod} \]

*
Complex-valued convolutions at rescue

\[ \text{Input signal} \xrightarrow{R^{16}} \mathbb{C}\text{Conv} \xrightarrow{4} \mathbb{C}^4 \xrightarrow{\text{Modulus}} R^4 \xrightarrow{} \text{Output} \]

\( \mathbb{C}\text{Mod} \)
Complex-valued convolutions at rescue

- Input signal
- Convolution
- Modulus
- Output

\[ \mathbb{C} \text{Mod} \]

Fourier transform modulus
Complex-valued convolutions at rescue

Input signal

Complex analytic convolution kernel

Input \xrightarrow{\mathbb{R}^{16}} \mathbb{C}\text{Conv} \downarrow 4 \xrightarrow{\mathbb{C}^{4}} \text{Modulus} \xrightarrow{\mathbb{R}^{4}} \text{Output}

\mathbb{C}\text{Mod}

Real part

Imaginary part

Hilbert transform

Fourier transform modulus

Kévin Polisano – DATA seminar – 23/11/2023
Complex-valued convolutions at rescue

Input signal

Complex analytic convolution kernel

Increased subsampling factor

Hilbert transform

Complex Mod

Output
Complex-valued convolutions at rescue

- Input signal
- Complex analytic convolution kernel
- Increased subsampling factor
- Hilbert transform
- Modulus operator

\( \mathbb{C}\text{Conv} \downarrow 4 \rightarrow \mathbb{C}^4 \rightarrow \text{Modulus} \rightarrow \mathbb{R}^4 \rightarrow \text{Output} \)
Complex-valued convolutions at rescue

Input signal

Complex analytic convolution kernel

Output signal

Increased subsampling factor

Hilbert transform

Modulus operator
Complex-valued convolutions at rescue

Input signal

Complex analytic convolution kernel

Increased subsampling factor

Modulus operator

Hilbert transform

Output signal
Complex-valued convolutions at rescue

**Objective:** establish conditions under which $\mathbb{R}^{\text{Max}} \approx \mathbb{C}\text{Mod}$. 

---

**Input signal**

**Complex analytic convolution kernel**

**Output signal**

**$\mathbb{C}\text{Mod}$**

Increased subsampling factor

Hilbert transform

Convolution operator

Real part

Imaginary part

$\mathbb{C}^4$ Modulus

$\mathbb{R}^4$ Output

$\mathbb{R}^{16}$ Input
Fundamental hypothesis on filter

- Band-pass, oriented and **analytic** Gabor-like filters $W$
Fundamental hypothesis on filter

- Band-pass, oriented and **analytic** *Gabor-like* filters $W$

\[ V = \]

![Diagram](image-url)

Kévin Polisano - DATA seminar - 23/11/2023
Fundamental hypothesis on filter

- Band-pass, oriented and **analytic** Gabor-like filters $W$

\[
V = \begin{array}{c}
\begin{array}{c}
\begin{array}{c}
0 \\
\pi \\
-\pi
\end{array}
\end{array}
\end{array}
\]

\[
W = \begin{array}{c}
\begin{array}{c}
\begin{array}{c}
V + i \mathcal{H}(V)
\end{array}
\end{array}
\end{array}
\]

\[
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
0 \\
\pi \\
-\pi
\end{array}
\end{array}
\end{array}
\]

\[
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\theta^+ \\
\theta^+
\end{array}
\end{array}
\end{array}
\]
Fundamental hypothesis on filter

- Band-pass, oriented and **analytic** Gabor-like filters $W$

\[ V = \quad W = \quad + i \quad V \quad H(V) \]
Fundamental hypothesis on filter

- Band-pass, oriented and **analytic** Gabor-like filters $W$

\[ V = \begin{bmatrix} \pi & \pi \\ 0 & 0 \end{bmatrix}, \quad W = \begin{bmatrix} V + i \end{bmatrix} \]

- Bandwidth
- Characteristic frequency
Fundamental hypothesis on filter

- Band-pass, oriented and **analytic** Gabor-like filters $W$

\[
V = \begin{bmatrix}
 & & \\
 & & \\
\end{bmatrix}
\quad \quad \quad W = \begin{bmatrix}
 & & \\
 & & \\
\end{bmatrix} + i \begin{bmatrix}
 & & \\
 & & \\
\end{bmatrix}
\]

- Bandwidth
- Characteristic frequency

$\mathcal{H}(V)$
Two operators to compare

\[
\begin{align*}
\text{Input} & \quad 224 \times 224 \\
\text{Conv} \downarrow m & \quad 56 \times 56 \\
V & \quad Y \quad \text{MaxPool} \downarrow 2 \\
\text{Output} & \quad Y_{\max}
\end{align*}
\]

\[
\begin{align*}
\text{Input} & \quad 224 \times 224 \\
\circ\text{Conv} \downarrow 2m & \quad 28 \times 28 \\
W & \quad Z \quad \text{Modulus} \\
\text{Output} & \quad Y_{\text{mod}}
\end{align*}
\]
Two operators to compare

\( \mathbb{R} \to \mathbb{M} \to \mathbb{A} \to \mathbb{X} \to \mathbb{V} \to \mathbb{Y} \to \mathbb{Y}^{\text{max}} \)

\( \mathbb{C} \to \mathbb{Ell} \to \mathbb{W} \to \mathbb{Z} \to \mathbb{Y}^{\text{mod}} \)
Two operators to compare

**RMax**

Input $\times 224 \times 224$ → Conv $\downarrow m$ → MaxPool $\downarrow 2$ → Output $\times 28 \times 28$

- $X \rightarrow V \rightarrow Y \rightarrow Y_{\text{max}}$

**CMod**

Input $\times 224 \times 224$ → Conv $\downarrow 2m$ → Modulus → Output $\times 28 \times 28$

- $X \rightarrow W \rightarrow Z \rightarrow Y_{\text{mod}}$
Two operators to compare

\( \mathbb{R} \) Max

As in standard CNNs

\( \mathbb{C} \) Mod
Two operators to compare

**RMax**

Input \(224 \times 224\) → Conv \(\downarrow m\) → MaxPool \(\downarrow 2\) → Output \(28 \times 28\)

- \(X\) → \(V\) → \(Y\) → \(Y_{\text{max}}\)

**CMod**

Input \(224 \times 224\) → \(\mathcal{C}\text{Conv} \downarrow 2m\) → Modulus → Output \(28 \times 28\)

- \(X\) → \(W\) → \(Z\) → \(Y_{\text{mod}}\) → No max pooling
Two operators to compare

\[ \mathbb{R} \]
\[ \text{Max} \]
\[ \text{Conv} \quad \downarrow m \quad \text{MaxPool} \quad \downarrow 2 \]
\[ X \quad V \quad Y \quad Y_{\text{max}} \]

\[ \mathbb{C} \]
\[ \text{Mod} \]
\[ \overline{\text{Conv}} \quad \downarrow 2m \quad \text{Modulus} \]
\[ X \quad W \quad Z \quad Y_{\text{mod}} \quad \text{No max pooling} \]
Two operators to compare

\[ R_{\text{Max}} \]

\[
\begin{array}{c}
\text{Input} \quad \begin{array}{c}
224 \times 224
\end{array} \\
\text{Conv} \downarrow m \\
X \rightarrow V
\end{array}
\begin{array}{c}
\text{MaxPool} \downarrow 2 \\
Y
\end{array}
\begin{array}{c}
\text{Output} \\
Y_{\text{max}}
\end{array}
\]

\[ C_{\text{Mod}} \]

\[
\begin{array}{c}
\text{Input} \quad \begin{array}{c}
224 \times 224
\end{array} \\
\text{CConv} \downarrow 2m \\
X \rightarrow W
\end{array}
\begin{array}{c}
\text{Modulus} \\
Z
\end{array}
\begin{array}{c}
\text{Output} \\
Y_{\text{mod}}
\end{array}
\]

\[ U_{m,q}^{\text{max}}[W] : X \leftrightarrow \text{MaxPool}_q \left( (X \ast \overline{\text{Re}W}) \downarrow m \right) \]
Two operators to compare

\textbf{RMax}

Input $224 \times 224$ $\rightarrow$ Conv $\downarrow m$ $\rightarrow$ MaxPool $\downarrow 2$ $\rightarrow$ Output $28 \times 28$

$X \rightarrow V \rightarrow Y \rightarrow Y_{\text{max}}$

\textbf{CMod}

Input $224 \times 224$ $\rightarrow$ CConv $\downarrow 2m$ $\rightarrow$ Modulus $\rightarrow$ Output $28 \times 28$

$X \rightarrow W \rightarrow Z \rightarrow Y_{\text{mod}}$

No max pooling

$U_{m,q}[W] : X \leftrightarrow \text{MaxPool}_q \left((X \ast \overline{\text{ReW}}) \downarrow m\right)$

$(X \ast \overline{V})[n] := \sum_{p \in \mathbb{Z}^2} X[p] \overline{V}[n - p] \quad \overline{V}[n] := V[-n]$

$\text{MaxPool}_q(Y)[n] := \max_{\|p\|_\infty \leq q} Y[2n + p] \quad (Y \downarrow m)[n] := Y[mn]$
Two operators to compare

$\mathbb{R}$\text{Max}

Input $\times 224$ → Conv $\downarrow m$ → MaxPool $\downarrow 2$ → Output $\times 28$

$X$ $\rightarrow$ $V$ $\rightarrow$ $Y$ $\rightarrow$ $Y_{\text{max}}$

$\mathbb{C}$\text{Mod}

Input $\times 224$ → Conv $\downarrow 2m$ → Modulus → Output $\times 28$

$X$ $\rightarrow$ $W$ $\rightarrow$ $Z$ $\rightarrow$ $Y_{\text{mod}}$

No max pooling

$U_{m,q}^{\text{max}}[W] : X \leftrightarrow \text{MaxPool}_q \left( (X \ast \overline{\text{Re}W}) \downarrow m \right)$
Two operators to compare

$\mathcal{R}_{\text{Max}}$

$\mathcal{C}_{\text{Mod}}$

$U_{m,q}^{\text{max}}[W]: X \leftrightarrow \text{MaxPool}_q \left( (X \ast \overline{\text{Re}W}) \downarrow m \right)$
Two operators to compare

**RMax**

Input $224 \times 224$ → Conv $\downarrow m$ → MaxPool $\downarrow 2$ → Output $28 \times 28$

X → V → Y → $Y^{\text{max}}$

**CMod**

Input $224 \times 224$ → CConv $\downarrow 2m$ → Modulus → Output $28 \times 28$

X → W → Z → $Y^{\text{mod}}$

No max pooling

$U_{m,q}^{\text{max}}[W] : X \mapsto \text{MaxPool}_q \left( (X \ast \text{Re}W) \downarrow m \right)$

$U_{m}^{\text{mod}}[W] : X \mapsto \left| (X \ast \text{Re}W) \downarrow (2m) \right|$
Roadmap
Show that, under the Gabor hypothesis, \( \mathcal{C} \text{Mod} \) is stable with respect to small input shifts.
Roadmap

- Show that, **under the Gabor hypothesis**, $\text{CMod}$ is **stable** with respect to small input shifts

- **Establish conditions** on the filter’s frequency and orientation under which $\text{RMax}$ and $\text{CMod}$ **produce comparable outputs**:

\[
U_{m,q}^{\max}[W](X) \approx U_{m}^{\text{mod}}[W](X)
\]
Show that, under the Gabor hypothesis, \( C\text{Mod} \) is stable with respect to small input shifts.

Establish conditions on the filter’s frequency and orientation under which \( R\text{Max} \) and \( C\text{Mod} \) produce comparable outputs:

\[
U_{m,q}^{\text{max}}[W](X) \approx U_{m}^{\text{mod}}[W](X)
\]

Deduce a measure of shift invariance for \( R\text{Max} \) operator, which benefits from the stability of \( C\text{Mod} \).
Roadmap

- Show that, under the Gabor hypothesis, \( \text{CMod} \) is stable with respect to small input shifts.

- Establish conditions on the filter’s frequency and orientation under which \( \text{RMax} \) and \( \text{CMod} \) produce comparable outputs:

\[
U_{m,q}^{\text{max}}[W](X) \approx U_{m}^{\text{mod}}[W](X)
\]

- Deduce a measure of shift invariance for \( \text{RMax} \) operator, which benefits from the stability of \( \text{CMod} \)

- Extend our results to multichannel operators (RGB images), such as implemented in conventional CNN architectures.
Roadmap

- Show that, **under the Gabor hypothesis**, \( C_{\text{Mod}} \) is **stable** with respect to small input shifts.

- **Establish conditions** on the filter’s frequency and orientation under which \( R_{\text{Max}} \) and \( C_{\text{Mod}} \) produce **comparable outputs**:

  \[
  U_{m,q}^{\text{max}}[W](X) \approx U_{m}^{\text{mod}}[W](X)
  \]

- Deduce a **measure of shift invariance** for \( R_{\text{Max}} \) operator, which benefits from the stability of \( C_{\text{Mod}} \).

- Extend our results to **multichannel operators** (RGB images), such as implemented in conventional CNN architectures.

- Experimental validation on a deterministic setting based on the **dual-tree** complex wavelet packet transform (DT-CWPT).
Detour via the continuous framework

- Using the **Shannon-Whittaker sampling theorem**

\[
\begin{align*}
(X * \text{Re}(W)) \downarrow m &= Y \\
\end{align*}
\]
Detour via the continuous framework

- Using the **Shannon-Whittaker sampling theorem**

\[ F \xrightarrow{\text{Shannon interpolation}} \mathbb{R} \mathbf{x} \]

\[ \mathbb{R}^\mathbb{M} \]

Uniform sampling

\[ \text{Re}(W) \downarrow m = Y \]

\[ Y^{\text{max}} \]
Detour via the continuous framework

Using the **Shannon-Whittaker sampling theorem**

Using the Shannon-Whittaker sampling theorem

\[
\text{Re}(W) \downarrow m = Y
\]

\( Y_{\text{max}} \)
Detour via the continuous framework

Using the **Shannon-Whittaker sampling theorem**

\[
F \ast \text{Sampling interval} \rightarrow \text{Re}(\Psi) \downarrow m = Y
\]

where \( \text{Re}(W) \) and \( Y_{\text{max}} \) are defined as:

\[
\text{Re}(W) \quad \downarrow m \quad Y
\]

Kévin Polisano - DATA seminar - 23/11/2023
Detour via the continuous framework

- Using the **Shannon-Whittaker sampling theorem**

\[ F \ast \text{Uniform sampling} \rightarrow \text{Re}(W) \]

\[ \downarrow m \]

\[ = Y \]

\[ \downarrow m_s \]

\[ = A \]
Detour via the continuous framework

- Using the **Shannon-Whittaker sampling theorem**
Detour via the continuous framework

Using the **Shannon-Whittaker sampling theorem**

\[ Y_{\text{max}}[n] = \max_{\|p\|_\infty \leq q} \left( F \star \text{Re}(\Psi) \right) (2msn + msp) \]

Using the Shannon-Whittaker sampling theorem, the maximum can be evaluated on a discrete grid.

\[ (X \star \text{Re}(W)) \downarrow m = Y \]

\[ F \star \text{Re}(\Psi) = A \]

Maximum evaluated on a discrete grid.
Detour via the continuous framework

- Using the **Shannon-Whittaker sampling theorem**

\[
Y_{\text{max}}[n] = \max_{\|p\|_{\infty} \leq q} (F \ast \text{Re}(\psi))(2msn + msp)
\]

Max pooling grid of size \((2q + 1) \times (2q + 1)\)

Using the **Shannon-Whittaker sampling theorem**

Uniform sampling

Maximum evaluated on a discrete grid
Detour via the continuous framework

Using the **Shannon-Whittaker sampling theorem**

\[ Y_{\text{max}}[n] = \max_{\|p\|_\infty \leq q} \left( F \star \text{Re}(\Psi) \right) \left( 2msn +msp \right) \]

Center of the max pooling window

Uniform sampling

\[ := x_n \]

\[ R_{\text{Max}} \]

\[ \sum \]

\[ 3ms \]
Detour via the continuous framework

Using the Shannon-Whittaker sampling theorem

\[ Y_{\text{max}}[n] = \max_{\|p\|_{\infty} \leq q} \left( F \ast \text{Re}(\psi) \right) \left( 2msn + msp \right) \]

\[ \overline{R_{\text{Max}}} \]

Browsing the max pooling grid

Uniform sampling

\[ Y_{\text{max}} \]
Detour via the continuous framework

What if we search for the maximum \textit{continuously} in the window?

\[
Y_{0}^{\max}[n] = \max_{\|h\|_{\infty} \leq \frac{3ms}{2}} (F \ast \text{Re}(\Psi))(x_n + h) \approx Y^{\max}[n]?
\]

\[
\begin{align*}
\text{(X)} & \ast \text{Re}(W) \downarrow m = Y \\
F & \ast \text{Re}(\Psi) = A \\
A^{\max} & \quad 2ms \\
\end{align*}
\]
Detour via the continuous framework

What if we search for the maximum \textit{continuously} in the window?

\[
Y_0^{\text{max}}[n] = \max_{\|h\|_\infty \leq \frac{3ms}{2}} (F \ast \text{Re}(\Psi))(x_n + h) \approx Y^{\text{max}}[n]?
\]
Detour via the continuous framework

What if we search for the maximum **continuously** in the window?

\[ Y_0^{\text{max}}[n] = \max_{\|h\|_\infty \leq \frac{3ms}{2}} (F \ast \text{Re}(\Psi))(x_n + h) \approx Y^{\text{max}}[n]? \]

Continuous window covering the max pooling grid

Browsing the whole window
Detour via the continuous framework

What if we search for the maximum *continuously* in the window?

\[ Y_{0\text{max}}[n] = \max_{\|h\|_{\infty} \leq \frac{3ms}{2}} (F \ast \text{Re}(\Psi))(x_n + h) \approx Y_{\text{max}}[n]? \]
The output $Y_{\text{mod}}$ can be obtained by a uniform sampling of $|F \ast \Psi|$.
Detour via the continuous framework

The output $Y^\text{mod}$ can be obtained by a uniform sampling of $|F \ast \Psi|$
Detour via the continuous framework

The output $Y^{\text{mod}}$ can be obtained by a uniform sampling of $|F \ast \Psi|$
The output $\mathbb{Y}^{\text{mod}}$ can be obtained by a uniform sampling of $|F \ast \Psi|$.
The output $Y^{\text{mod}}$ can be obtained by a uniform sampling of $|F \ast \Psi|$.
The output $Y_{\text{mod}}$ can be obtained by a uniform sampling of $|F \ast \Psi|$.
Detour via the continuous framework

The output $Y^{\text{mod}}$ can be obtained by a uniform sampling of $|F \ast \Psi|$.
From high to low-frequency
From high to low-frequency

\[ \mathcal{V}(\nu, \varepsilon) := \left\{ \Psi \in L^2_{\mathcal{C}}(\mathbb{R}^2) \mid \text{supp } \hat{\Psi} \subset B_{\infty}(\nu, \varepsilon/2) \right\}. \]
From high to low-frequency

\[ \mathcal{V}(\nu, \varepsilon) := \left\{ \Psi \in L^2_{\mathbb{C}}(\mathbb{R}^2) \mid \text{supp } \hat{\Psi} \subset B_\infty(\nu, \varepsilon/2) \right\}. \]

\[ F_0 : x \mapsto (F \ast \overline{\Psi})(x)e^{i \langle \nu, x \rangle} \quad \Psi \in \mathcal{V}(\nu, \varepsilon) \quad \text{supp } \hat{F}_0 \subset B_\infty(\varepsilon/2) \]
From high to low-frequency

- \( \mathcal{V}(\nu, \varepsilon) := \left\{ \Psi \in L^2_\mathbb{C}(\mathbb{R}^2) \middle| \text{supp } \widehat{\Psi} \subset B_\infty(\nu, \varepsilon/2) \right\} \).

- \( F_0 : \mathbf{x} \mapsto (F \ast \overline{\Psi})(\mathbf{x})e^{i\langle \nu, \mathbf{x} \rangle} \) \( \Psi \in \mathcal{V}(\nu, \varepsilon) \) \( \text{supp } \widehat{F_0} \subset B_\infty(\varepsilon/2) \)

- **Shift-invariance bound** for low-frequency functions:

\[
\| T_h F_0 - F_0 \|_{L^2} \leq \alpha(\varepsilon h) \| F_0 \|_{L^2} \quad \alpha : \tau \mapsto \frac{\| \tau \|_1}{2}
\]
Shift-invariance of CMod in the discrete framework
Shift-invariance of CMod in the discrete framework

**Theorem (Shift invariance of CMod)**

If $W \in J(\theta, \kappa)$ and $\kappa \leq \pi/m$

then for any input image with finite support $X \in l^2_{\mathbb{R}}(\mathbb{Z}^2)$

$$\left\| U_m^{\text{mod}}(TuX) - U_m^{\text{mod}}X \right\|_2 \leq \alpha(\kappa u) \left\| U_m^{\text{mod}}X \right\|_2$$
**Theorem** (Shift invariance of CMod)

If \( W \in \mathcal{J}(\theta, \kappa) \) and \( \kappa \leq \pi/m \)
then for any input image with finite support \( X \in l^2_\mathbb{R}(\mathbb{Z}^2) \)

\[
\|U^\text{mod}_m (T_u X) - U^\text{mod}_m X\|_2 \leq \alpha(\kappa u) \|U^\text{mod}_m X\|_2
\]

\( \alpha : \tau \mapsto \frac{\|\tau\|_1}{2} \)
From CMod to RMax in the continuous framework

I. Waldspurger intuition linking the two operators:
From CMod to RMax in the continuous framework

I. Waldspurger intuition linking the two operators:

\[ F_0 : \mathbf{x} \mapsto (F \ast \overline{\psi})(\mathbf{x}) e^{i\langle \nu, \mathbf{x} \rangle} \quad \psi \in \mathcal{V}(\nu, \varepsilon) \quad \supp \hat{F}_0 \subset B_\infty(\varepsilon/2) \]
From CMod to RMax in the continuous framework

I. Waldspurger intuition linking the two operators:

\[ F_0 : x \mapsto (F * \overline{\psi})(x)e^{i\langle \nu, x \rangle} \quad \text{supp} \widehat{F_0} \subset B_\infty(\varepsilon/2) \]

\[
(F * \text{Re} \overline{\psi})(x) = \text{Re}((F * \overline{\psi})(x)) = \text{Re}(F_0(x)e^{-i\langle \nu, x \rangle})
\]
From CMod to RMax in the continuous framework

I. Waldspurger intuition linking the two operators:

\[
F_0 : x \mapsto (F \ast \overline{\psi})(x)e^{i\langle \nu, x \rangle} \quad \Psi \in \mathcal{V}(\nu, \varepsilon)
\]

\[
\text{supp} \hat{F}_0 \subset B_{\infty}(\varepsilon/2)
\]

\[
(F \ast \text{Re} \overline{\psi})(x) = \text{Re}((F \ast \overline{\psi})(x)) = \text{Re}(F_0(x)e^{-i\langle \nu, x \rangle})
\]

\[
(F \ast \text{Re} \overline{\psi})(x + h) \approx \text{Re}(F_0(x)e^{-i\langle \nu, x+h \rangle}) = \text{Re}((F \ast \overline{\psi})(x)e^{-i\langle \nu, h \rangle})
\]
### I. Waldspurger intuition linking the two operators:

\[ F_0 : \mathbf{x} \mapsto (F \ast \overline{\Psi})(\mathbf{x})e^{i\langle \nu, \mathbf{x} \rangle} \quad \psi \in \mathcal{V}(\nu, \varepsilon) \]

\[ \text{supp} \widehat{F_0} \subset B_{\infty}(\varepsilon/2) \]

\[ \| \mathbf{h} \|_2 \ll 2\pi/\varepsilon \implies F_0(\mathbf{x} + \mathbf{h}) \approx F_0(\mathbf{x}) \]

\[ (F \ast \text{Re} \overline{\Psi})(\mathbf{x}) = \text{Re}((F \ast \overline{\Psi})(\mathbf{x})) = \text{Re}(F_0(\mathbf{x})e^{-i\langle \nu, \mathbf{x} \rangle}) \]

\[ (F \ast \text{Re} \overline{\Psi})(\mathbf{x} + \mathbf{h}) \approx \text{Re}(F_0(\mathbf{x})e^{-i\langle \nu, \mathbf{x} + \mathbf{h} \rangle}) = \text{Re}((F \ast \overline{\Psi})(\mathbf{x})e^{-i\langle \nu, \mathbf{h} \rangle}) \]
I. Waldspurger intuition linking the two operators:

\[ F_0 : x \mapsto (F \ast \overline{\Psi})(x)e^{i\langle \nu, x \rangle} \quad \Psi \in \mathcal{V}(\nu, \varepsilon) \]

\[ \text{supp} \widehat{F_0} \subset B_{\infty}(\varepsilon/2) \]

\[ (F \ast \overline{\Psi})(x) = |(F \ast \overline{\Psi})(x)|e^{-iH(x)} \quad \| h \|_2 \ll 2\pi/\varepsilon \implies F_0(x + h) \approx F_0(x) \]

\[ (F \ast \text{Re} \overline{\Psi})(x) = \text{Re}((F \ast \overline{\Psi})(x)) = \text{Re}(F_0(x)e^{-i\langle \nu, x \rangle}) \]

\[ (F \ast \text{Re} \overline{\Psi})(x + h) \approx \text{Re}(F_0(x)e^{-i\langle \nu, x + h \rangle}) = \text{Re}((F \ast \overline{\Psi})(x)e^{-i\langle \nu, h \rangle}) \]
From CMod to RMax in the continuous framework

I. Waldspurger intuition linking the two operators:

\[ F_0 : x \mapsto (F * \overline{\psi})(x)e^{i\langle \nu, x \rangle} \quad \psi \in V(\nu, \varepsilon) \]

\[ \text{supp} \hat{F_0} \subset B_\infty(\varepsilon/2) \]

\[ (F * \overline{\psi})(x) = |(F * \overline{\psi})(x)|e^{-iH(x)} \quad \|h\|_2 \ll 2\pi/\varepsilon \implies F_0(x + h) \approx F_0(x) \]

\[ (F * \text{Re} \overline{\psi})(x) = \text{Re}((F * \overline{\psi})(x)) = \text{Re}(F_0(x)e^{-i\langle \nu, x \rangle}) \]

\[ (F * \text{Re} \overline{\psi})(x + h) \approx \text{Re}(F_0(x)e^{-i\langle \nu, x + h \rangle}) = \text{Re}((F * \overline{\psi})(x)e^{-i\langle \nu, h \rangle}) \]

\[ (F * \text{Re} \overline{\psi})(x + h) \approx |(F * \overline{\psi})(x)| \cos(\langle \nu, h \rangle - H(x)) \]
I. Waldspurger intuition linking the two operators:

\[ F_0 : x \mapsto (F \ast \overline{\psi})(x)e^{i\langle \nu, x \rangle} \]

\[ \psi \in \mathcal{V}(\nu, \varepsilon) \]

\[ \text{supp} \widetilde{F}_0 \subset B_{\infty}(\varepsilon/2) \]

\[ (F \ast \overline{\psi})(x) = |(F \ast \overline{\psi})(x)|e^{-iH(x)} \quad \|h\|_2 \ll 2\pi/\varepsilon \implies F_0(x + h) \approx F_0(x) \]

\[ (F \ast \text{Re} \overline{\psi})(x) = \text{Re}((F \ast \overline{\psi})(x)) = \text{Re}(F_0(x)e^{-i\langle \nu, x \rangle}) \]

\[ (F \ast \text{Re} \overline{\psi})(x + h) \approx \text{Re}(F_0(x)e^{-i\langle \nu, x + h \rangle}) = \text{Re}((F \ast \overline{\psi})(x)e^{-i\langle \nu, h \rangle}) \]

\[ (F \ast \text{Re} \overline{\psi})(x + h) \approx |(F \ast \overline{\psi})(x)| \cos(\langle \nu, h \rangle - H(x)) \]

\[ G(x, h) \]
From CMod to RMax in the continuous framework

I. Waldspurger intuition linking the two operators:

\[(F \ast \text{Re} \overline{\Psi})(x + h) \approx |(F \ast \overline{\Psi})(x)| \cos(\langle \nu, h \rangle - H(x))\]
I. Waldspurger intuition linking the two operators:

\[(F \ast \text{Re} \Psi)(x + h) \approx |(F \ast \bar{\Psi})(x)| \cos(\langle \nu, h \rangle - H(x))\]

- \(U^\text{mod}[\Psi](F) : x \mapsto |(F \ast \bar{\Psi})(x)|\)
- \(U^\text{max}[\Psi](F) : x \mapsto \max_{||h||_\infty \leq r} (F \ast \text{Re}\bar{\Psi})(x + h)\)
I. Waldspurger intuition linking the two operators:

\[(F \ast \text{Re } \Psi)(x + h) \approx |(F \ast \Psi)(x)| \cos(\langle \nu, h \rangle - H(x))\]

- \(U^{\text{mod}}[\Psi](F) : x \mapsto |(F \ast \Psi)(x)|\)
- \(U^{\text{max}}[\Psi](F) : x \mapsto \max_{\|h\|_{\infty} \leq r} (F \ast \text{Re } \Psi)(x + h)\)

\[r \ll 2\pi/\varepsilon \implies U^{\text{max}}_r F(x) \approx U^{\text{mod}} F(x) \max_{\|h\|_{\infty} \leq r} G(x, h)\]
From CMod to RMax in the continuous framework

I. Waldspurger intuition linking the two operators:

\[(F \ast \text{Re } \overline{\psi})(x + h) \approx |(F \ast \overline{\psi})(x)| \cos(\langle \nu, h \rangle - H(x))\]

- \(U_{\text{mod}}[\psi](F) : x \mapsto |(F \ast \overline{\psi})(x)|\)
- \(U_{\text{max}}^r[\psi](F) : x \mapsto \max_{\|h\|_\infty \leq r} (F \ast \text{Re } \overline{\psi})(x + h)\)

\[r \ll 2\pi/\varepsilon \implies U_{r}^\text{max} F(x) \approx U_{\text{mod}} F(x) \max_{\|h\|_\infty \leq r} G(x, h)\]

- For \(r \geq \frac{\pi}{\|\nu\|_2}\) the cosine \(h \mapsto G(x, h)\) reaches 1 on \(B_\infty(r)\)
I. Waldspurger intuition linking the two operators:

\[(F \ast \text{Re} \overline{\Psi})(x + h) \approx |(F \ast \overline{\Psi})(x)| \cos(\langle \nu, h \rangle - H(x))\]

- \(U^{\text{mod}}[\Psi](F) : x \mapsto |(F \ast \overline{\Psi})(x)|\)
- \(U^{\text{max}}_r[\Psi](F) : x \mapsto \max_{\|h\|_\infty \leq r} (F \ast \text{Re} \overline{\Psi})(x + h)\)

\[r \ll 2\pi/\varepsilon \implies U^{\text{max}}_r F(x) \approx U^{\text{mod}} F(x) \max_{\|h\|_\infty \leq r} G(x, h)\]

- For \(r \geq \frac{\pi}{\|\nu\|_2}\) the cosine \(h \mapsto G(x, h)\) reaches 1 on \(B_\infty(r)\)

\[
\frac{\pi}{\|\nu\|_2} \leq r \ll \frac{2\pi}{\varepsilon} \implies U^{\text{max}}_r F(x) \approx U^{\text{mod}} F(x)
\]
From CMod to RMax in the continuous framework

I. Waldspurger intuition linking the two operators:
From CMod to RMax in the continuous framework

I. Waldspurger intuition linking the two operators:

\[ F \star \text{Re}(\Psi) \rightarrow A \]

\[ (F \star \text{Re}(\Psi))(x_n + h) \]
I. Waldspurger intuition linking the two operators:
From CMod to RMax in the continuous framework

I. Waldspurger intuition linking the two operators:

\[ (F \ast \text{Re}(\Psi))(x_n + h) \approx |(F \ast \Psi)(x_n)| \cos(\langle \nu, h \rangle - H(x_n)) \]
I. Waldspurger intuition linking the two operators:

\[ (F \ast \text{Re}(\Psi))(x_n + h) \approx |(F \ast \Psi)(x_n)| \cos(\langle \nu, h \rangle - H(x_n)) = A(x_n + h) \]
From CMod to RMax in the continuous framework

I. Waldspurger intuition linking the two operators:

\[
(F \ast \text{Re}(\Psi))(x_n + h) \approx (\text{Re}(\Psi)(x_n) \cos(\langle \nu, h \rangle - H(x_n)))\]
I. Waldspurger intuition linking the two operators:

\[(x_n + h, x_n \cdot x_n + h) \approx (F \ast \text{Re}(\Psi))(x_n + h) \leq \max_{\|h\|_{\infty} \leq \frac{3ms}{2}} (F \ast \Psi)(x_n) \cos(\nu, h) - H(x_n)\]
I. Waldspurger intuition linking the two operators:

\[
\begin{align*}
(F \ast \text{Re}(\Psi))(x_n + h) &\approx \max_{\|h\|_\infty \leq \frac{3ms}{2}} Y_{\text{mod}}^{\max}[n] \\
&\approx \cos(\langle \nu, h \rangle - H(x_n)) \max_{\|h\|_\infty \leq \frac{3ms}{2}}
\end{align*}
\]
From CMod to RMax in the continuous framework

I. Waldspurger intuition linking the two operators:
Adaptation to the discrete case

\[(F \ast \text{Re}(\Psi))(x_n + h_p) \approx |(F \ast \Psi)(x_n)| \cos(\langle \nu, h_p \rangle - H(x_n))\]
Adaptation to the discrete case

\[
(F \ast \text{Re}(\Psi))(x_n + h_p) \approx (F \ast \Psi)(x_n) \cos(\langle \nu, h_p \rangle - H(x_n)) \mod \lfloor n \rfloor
\]
Adaptation to the discrete case

\begin{align*}
(F \ast \text{Re}(\Psi))(x_n + h_p) &\approx \left(\left| (F \ast \Psi)(x_n) \right| \cos(\langle \nu, h_p \rangle - H(x_n)) \right) \cdot Y_{\text{mod}}[n]
\end{align*}
Adaptation to the discrete case

\[
\begin{align*}
(F \ast \text{Re}(\Psi))(x_n + h_p) & \approx (F \ast \Psi(x_n)) \cos(\langle \nu, h_p \rangle - H(x_n)) \\
\max_{\|p\| \leq q} & \quad \max_{\|p\| \leq q}
\end{align*}
\]
Adaptation to the discrete case

\[(F \ast \text{Re}(\Psi))(x_n + h_p) \approx \max_{\|p\| \leq q} (F \ast \Psi)(x_n) \cos(\langle \nu, h_p \rangle - H(x_n))\]

Maximum value over a discrete grid
\(q = 1\) in general
Adaptation to the discrete case

\[ y_{\text{max}}[n] \]

\[ (F \ast \text{Re}(\Psi))(x_n + h_p) \]

Maximum value over a discrete grid \((q = 1 \text{ in general})\)
Adaptation to the discrete case

\[ (F \ast \text{Re}(\Psi))(x_n + h_{p}) \]

Maximum value over a discrete grid (\( q = 1 \) in general)

\[ \max_{\|p\| \leq q} (F \ast \Psi)(x_n) \cos(\langle \nu, h_{p} \rangle - H(x_n)) \]

\[ G_{\max}(x_n) \leq 1 \]
Adaptation to the discrete case

\[ q \ll \frac{2\pi}{(m \kappa)} \implies U_{m, q}^{\text{max}} X[n] \approx U_{2m}^{\text{mod}} X[n] \max_{\|p\|_\infty \leq q} G_X(x_n, h_p), \]
Adaptation to the discrete case

\[ q \ll \frac{2\pi}{m \kappa} \quad \Rightarrow \quad U_{m, q}^{\text{max}} X[n] \approx U_{2m}^{\text{mod}} X[n] \]

\[ \max_{\|p\|_{\infty} \leq q} G_X(x_n, h_p), \]

Not necessarily reaches 1
Adaptation to the discrete case

\[ q \ll 2\pi/(m\kappa) \implies U_{m, q}^{\text{max}} X[n] \approx U_{2m}^{\text{mod}} X[n] \]

\[ \max_{\|p\|_\infty \leq q} G_X(x_n, h_p), \]

\[ q \ll 2\pi/(m\kappa) \implies \|U_{2m}^{\text{mod}} X - U_{m, q}^{\text{max}} X\|_2 \approx \|\delta_{m, q} X\|_2 \]

Not necessarily reaches 1
Adaptation to the discrete case

\[
q \ll \frac{2\pi}{(m\kappa)} \implies U_{m, q}^{\max} X[n] \approx U_{2m}^{\text{mod}} X[n] \quad \text{max}_{\|p\|_\infty \leq q} G_X(x_n, h_p),
\]

\[
q \ll \frac{2\pi}{(m\kappa)} \implies \|U_{2m}^{\text{mod}} X - U_{m, q}^{\max} X\|_2 \approx \|\delta_{m, q} X\|_2
\]

\[
\delta_{m, q} X[n] := U_{2m}^{\text{mod}} X[n] \left(1 - \max_{\|p\|_\infty \leq q} G_X(x_n, h_p)\right)
\]

Not necessarily reaches 1
Adaptation to the discrete case

$q \ll 2\pi/(m\kappa) \implies U_{m,q}^{\text{max}} X[n] \approx U_{2m}^{\text{mod}} X[n] \max_{\|p\|_\infty \leq q} G_X(x_n, h_p)$

$q \ll 2\pi/(m\kappa) \implies \|U_{2m}^{\text{mod}} X - U_{m,q}^{\text{max}} X\|_2 \approx \|\delta_{m,q} X\|_2$

$\delta_{m,q} X[n] := U_{2m}^{\text{mod}} X[n] \left(1 - \max_{\|p\|_\infty \leq q} G_X(x_n, h_p)\right)$

**Theorem** (Bound on the difference of $\text{CMod}$ and $\text{RMax}$)

If $\kappa \leq \pi/m$ and under another reasonable hypothesis

$$\|U_{2m}^{\text{mod}} X - U_{m,q}^{\text{max}} X\|_2 \leq \|\delta_{m,q} X\|_2 + \beta_q(m\kappa) \|U_{2m}^{\text{mod}} X\|_2$$

Not necessarily reaches 1
Adaptation to the discrete case

\[ q \ll 2\pi/(m\kappa) \implies U_{m, q}^{\text{max}} X[n] \approx U_{2m}^{\text{mod}} X[n] \]

\[ \max_{\|p\|_\infty \leq q} G_X(x_n, h_p), \]

\[ q \ll 2\pi/(m\kappa) \implies \|U_{2m}^{\text{mod}} X - U_{m, q}^{\text{max}} X\|_2 \approx \|\delta_{m, q} X\|_2 \]

\[ \delta_{m, q} X[n] := U_{2m}^{\text{mod}} X[n] \left(1 - \max_{\|p\|_\infty \leq q} G_X(x_n, h_p)\right) \beta_q : \kappa \mapsto q\kappa'. \]

**Theorem** (Bound on the difference of \( \text{CMod} \) and \( \text{RMax} \))

If \( \kappa \leq \pi/m \) and under another reasonable hypothesis

\[ \|U_{2m}^{\text{mod}} X - U_{m, q}^{\text{max}} X\|_2 \leq \|\delta_{m, q} X\|_2 + \beta_q(m\kappa) \|U_{2m}^{\text{mod}} X\|_2 \]
Pathological frequencies
Pathological frequencies

$x_n$

$x_n$

$x_{ni}$
Pathological frequencies

Need for a probabilistic framework where $X$ (resp. $F$) is seen as a discrete (resp. continuous) stochastic process on $\mathbb{Z}^2$ (resp. $\mathbb{R}^2$)
Pathological frequencies

Need for a probabilistic framework where $X$ (resp. $F$) is seen as a discrete (resp. continuous) stochastic process on $\mathbb{Z}^2$ (resp. $\mathbb{R}^2$)

Quantity of interest: $\mathbb{E}\left[\left(1 - G^{\text{max}}(x_n)\right)^2\right]$
Pathological frequencies

Need for a probabilistic framework where $X$ (resp. $F$) is seen as a discrete (resp. continuous) stochastic process on $\mathbb{Z}^2$ (resp. $\mathbb{R}^2$)

Quantity of interest: $\mathbb{E} \left[ (1 - G_{\text{max}}(x_n))^2 \right]$

with $G_{\text{max}}(x_n) = \max_{\|p\|_\infty \leq 1} \cos(\langle \nu, h_p \rangle - H(x_n))$
Pathological frequencies

Need for a probabilistic framework where $X$ (resp. $F$) is seen as a discrete (resp. continuous) stochastic process on $\mathbb{Z}^2$ (resp. $\mathbb{R}^2$)

Quantity of interest: $\mathbb{E} \left[ (1 - G_{\text{max}}(x_n))^2 \right]$

with $G_{\text{max}}(x_n) = \max_{\|p\|_\infty \leq 1} \cos(\langle \nu, h_p \rangle - H(x_n))$
Main result

MSE between $\text{CMod}$ and $\mathcal{RMax}$ output

\[
\mathbb{E} \left[ \left\| Y_{\text{max}} - Y_{\text{mod}} \right\|_2^2 \right] \leq (\beta_q(m\kappa) + \gamma_q(m\theta))^2
\]

$\theta \mapsto \gamma_q(m\theta)^2$

$q = 1$
Main result

\[ \text{MSE between } \mathcal{C} \text{Mod and } \mathcal{R} \text{Max output} \]

\[ \mathbb{E} \left[ \frac{\| Y^{\max} - Y^{\text{mod}} \|_2^2}{\| Y^{\text{mod}} \|_2^2} \right] \leq (\beta_q(m \kappa) + \gamma_q(m \theta))^2 \]

\[ \theta \mapsto \gamma_q(m \theta)^2 \]

\[ q = 1 \]
Main result

MSE between $\mathbb{C} \text{Mod}$ and $\mathbb{R} \text{Max}$ output

$$\mathbb{E} \left[ \left\| Y_{\text{max}} - Y_{\text{mod}} \right\|_2^2 \right] \leq \left( \beta_q(m \kappa) + \gamma_q(m \theta) \right)^2$$

Discrete nature of the max pooling grid

$$\mathbb{E} \left[ \left( 1 - G_{\text{max}}(x_n) \right)^2 \right] \quad \theta \leftarrow \frac{\gamma_q(m \theta)^2}{q = 1}$$
Main result

MSE between $\mathcal{C}Mod$ and $\mathcal{R}Max$ output

\[
\mathbb{E} \left[ \frac{\|Y_{\text{max}} - Y_{\text{mod}}\|_2^2}{\|Y_{\text{mod}}\|_2^2} \right] \leq (\beta_q(m\kappa) + \gamma_q(m\theta))^2
\]

Discrete nature of the max pooling grid

\[
\gamma_q(\omega) = \sqrt{\frac{3}{2} + \frac{1}{4\pi} \sum_{i=0}^{n-1} \left( \sin \delta H_i^{(q)}(\omega) - 8 \sin \frac{\delta H_i^{(q)}(\omega)}{2} \right)}
\]

$\theta \leftarrow \gamma_q(m\theta)^2$ for $q = 1$
Sketch of the proof

- **Reformulation** of the problem on the unit circle $S^1 \subset \mathbb{C}$
Sketch of the proof

- **Reformulation** of the problem on the unit circle $S^1 \subset \mathbb{C}$

  $[z, z']_{S^1} \subset S^1$ arc on the unit circle going from $z$ to $z'$
Sketch of the proof

**Reformulation** of the problem on the unit circle $S^1 \subset \mathbb{C}$

$[z, z']_{S^1} \subset S^1$ arc on the unit circle going from $z$ to $z'$

$Z_X : x \mapsto e^{i H_x(x)}$ and $Z_p : \omega \mapsto e^{i \langle \omega, p \rangle}$
Sketch of the proof

- **Reformulation** of the problem on the unit circle \( S^1 \subset \mathbb{C} \)

\[ [z, z']_{S^1} \subset S^1 \text{ arc on the unit circle going from } z \text{ to } z' \]

\[ Z_X : x \mapsto e^{iH_X(x)} \quad \text{and} \quad Z_p : \omega \mapsto e^{i\langle \omega, p \rangle} \]

\[ G_X(x_n, h_p) = \text{Re}(Z_X^*(x_n) Z_p(m\theta)) \quad \longrightarrow \quad \cos(\langle \nu, h_p \rangle - H(x_n)) \]
Sketch of the proof

- **Reformulation** of the problem on the unit circle $S^1 \subset \mathbb{C}$

$[z, z']_{S^1} \subset S^1$ arc on the unit circle going from $z$ to $z'$

$h_p = m \cdot p$

$Z_X : x \mapsto e^{iH_X(x)}$ and $Z_p : \omega \mapsto e^{i\langle \omega, p \rangle}$

- $G_X(x_n, h_p) = \text{Re}(Z_X^*(x_n) Z_p(m\theta)) \quad \rightarrow \quad \cos(\langle \nu, h_p \rangle - H(x_n))$
Sketch of the proof

- **Reformulation** of the problem on the unit circle $S^1 \subset \mathbb{C}$

$[z, z']_{S^1} \subset S^1$ arc on the unit circle going from $z$ to $z'$$$
h_p = m \cdot s \cdot p$$

$Z_X : x \mapsto e^{i H_X(x)}$ and $Z_p : \omega \mapsto e^{i \langle \omega, p \rangle}$$

$\nu = \theta / s$$$

- $G_X(x_n, h_p) = \text{Re}(Z_X^*(x_n) Z_p(m\theta)) \quad \xrightarrow{\text{blue}} \quad \cos(\langle \nu, h_p \rangle - H(x_n))$
Sketch of the proof

**Reformulation** of the problem on the unit circle $\mathbb{S}^1 \subset \mathbb{C}$

\[ [z, z']_{\mathbb{S}^1} \subset \mathbb{S}^1 \] arc on the unit circle going from $z$ to $z'$

- $h_p = m.s.p$
- $\nu = \theta / s$

\[ Z_X : x \mapsto e^{i H_X(x)} \quad \text{and} \quad Z_p : \omega \mapsto e^{i(\omega, p)} \]

- $G_X(x_n, h_p) = \Re(Z_X^*(x_n) Z_p(m \theta)) \quad \rightarrow \quad \cos(\langle \nu, h_p \rangle - H(x_n))$

- Sort $\{Z_p(\omega)\}_{p \in \{-q..q\}^2} \quad \rightarrow \quad (Z_i^{(q)}(\omega))_{i \in \{0..n_q-1\}} \quad n_q := (2q + 1)^2$
Sketch of the proof

- **Reformulation** of the problem on the unit circle $S^1 \subset \mathbb{C}$

$[z, z']_{S^1} \subset S^1$ arc on the unit circle going from $z$ to $z'$

$Z_X : x \mapsto e^{iH_X(x)}$ and $Z_p : \omega \mapsto e^{i\langle \omega, p \rangle}$

- $G_X(x_n, h_p) = \text{Re}(Z_X^*(x_n) Z_p(m\theta))$ \quad $\cos(\langle \nu, h_p \rangle - H(x_n))$

- Sort $\{Z_p(\omega)\}_{p \in \{\neg q..q\}^2}$ \quad $(Z_i^{(q)}(\omega))_{i \in \{0..n_q-1\}}$ \quad $n_q := (2q + 1)^2$

in ascending order of their argument:

$$0 = H_0^{(q)}(\omega) \leq \cdots \leq H_{n_q-1}^{(q)}(\omega) < 2\pi$$
Sketch of the proof

**Reformulation** of the problem on the unit circle $\mathbb{S}^1 \subset \mathbb{C}$

\[ [z, z']_{\mathbb{S}^1} \subset \mathbb{S}^1 \text{ arc on the unit circle going from } z \text{ to } z' \]

\[ Z_X : x \mapsto e^{iH_X(x)} \quad \text{and} \quad Z_p : \omega \mapsto e^{i\langle \omega, p \rangle} \]

\[ h_p = m.s.p \quad \nu = \theta / s \]

- \( G_X(x_n, h_p) = \Re(Z_X^*(x_n) Z_p(m\theta)) \)

- Sort \( \{Z_p(\omega)\}_{p \in \{-q..q\}^2} \rightarrow (Z_i^{(q)}(\omega))_{i \in \{0..n_q-1\}} \)

in ascending order of their argument:

\[ 0 = H_0^{(q)}(\omega) \leq \cdots \leq H_{n_q-1}^{(q)}(\omega) < 2\pi \]

\[ H_{n_q}^{(q)}(\omega) := 2\pi \quad Z_{n_q}^{(q)}(\omega) := 1 \]
Sketch of the proof

- **Reformulation** of the problem on the unit circle $S^1 \subset \mathbb{C}$

\[ [z, z']_{S^1} \subset S^1 \text{ arc on the unit circle going from } z \text{ to } z' \]

\[ Z_X : x \mapsto e^{iH_X(x)} \quad \text{and} \quad Z_p : \omega \mapsto e^{i\langle \omega, p \rangle} \]

- $G_X(x_n, h_p) = \text{Re}(Z_X^*(x_n) Z_p(m\theta)) \quad \xrightarrow{\rightarrow} \quad \cos(\langle \nu, h_p \rangle - H(x_n))$

- Sort $\{Z_p(\omega)\}_{p \in \{-q..q\}^2}$ \quad $\rightarrow$ \quad $\{Z_i^{(q)}(\omega)\}_{i \in \{0..n_q-1\}}$ \quad $n_q := (2q + 1)^2$

in ascending order of their argument:

\[ 0 = H_0^{(q)}(\omega) \leq \cdots \leq H_{n_q-1}^{(q)}(\omega) < 2\pi \]

- Split $S^1$ into $n_q$ arcs delimited by the $Z_i^{(q)}(\omega)$

\[ A_i^{(q)}(\omega) := \begin{cases} \left[ Z_i^{(q)}(\omega), Z_{i+1}^{(q)}(\omega) \right]_{S^1} & \text{if } H_{i+1}^{(q)}(\omega) - H_i^{(q)}(\omega) < 2\pi; \\ S^1 & \text{otherwise.} \end{cases} \]
Sketch of the proof

- **Reformulation** of the problem on the unit circle $S^1 \subset \mathbb{C}$
  
  $[z, z']_{S^1} \subset S^1$ arc on the unit circle going from $z$ to $z'$

  $$Z_X : x \mapsto e^{i H_X(x)} \quad \text{and} \quad Z_p : \omega \mapsto e^{i \langle \omega, p \rangle}$$

  - $G_X(x_n, h_p) = \text{Re}(Z_X^*(x_n) Z_p(m\theta)) \quad \rightarrow \quad \cos(\langle \nu, h_p \rangle - H(x_n))$

  - Sort $\{Z_p(\omega)\}_{p \in \{-q..q\}^2} \quad \rightarrow \quad (Z^{(q)}_i(\omega))_{i \in \{0..n_q-1\}}$
  
  $n_q := (2q + 1)^2$

  in ascending order of their argument:
  
  $$0 = H^{(q)}_0(\omega) \leq \cdots \leq H^{(q)}_{n_q-1}(\omega) < 2\pi$$

  - Split $S^1$ into $n_q$ arcs delimited by the $Z^{(q)}_i(\omega)$

  $$\mathcal{A}^{(q)}_i(\omega) := \left\{ \begin{array}{ll}
  \left[ Z^{(q)}_i(\omega), Z^{(q)}_{i+1}(\omega) \right]_{S^1} & \text{if } H^{(q)}_{i+1}(\omega) - H^{(q)}_i(\omega) < 2\pi; \\
  S^1 & \text{otherwise.}
  \end{array} \right.$$
Sketch of the proof

**Reformulation** of the problem on the unit circle $S^1 \subset \mathbb{C}$

$[z, z']_{S^1} \subset S^1$ arc on the unit circle going from $z$ to $z'\quad h_p = m.s.p$

$Z_X : x \mapsto e^{i H_X(x)} \quad$ and $\quad Z_p : \omega \mapsto e^{i \langle \omega, p \rangle} \quad \nu = \theta / s$

$G_X (x_n, h_p) = \text{Re} (Z_X^*(x_n) Z_p (m\theta)) \quad \Rightarrow \quad \cos (\langle \nu, h_p \rangle - H(x_n))$
### Sketch of the proof

**Reformulation** of the problem on the unit circle $S^1 \subset \mathbb{C}$

\[ [z, z'][S^1] \subset S^1 \] arc on the unit circle going from $z$ to $z'$

\[
Z_X : x \mapsto e^{iH_X(x)} \quad \text{and} \quad Z_p : \omega \mapsto e^{i\langle \omega, p \rangle}
\]

\[ h_p = ms_p \quad \nu = \theta / s \]

\[ G_X(x_n, h_p) = \text{Re}(Z_X^*(x_n) Z_p(m\theta)) \quad \text{cos}(\langle \nu, h_p \rangle - H(x_n)) \]

\[ G_X^{\text{max}}(x) = g_{\text{max}}(Z_X(x)) \quad g_{\text{max}} : z \mapsto \max_{\|p\|_\infty \leq q} \text{Re}(z^* Z_p) \]
Sketch of the proof

- **Reformulation** of the problem on the unit circle $S^1 \subset \mathbb{C}$

  $[z, z']_{S^1} \subset S^1$ arc on the unit circle going from $z$ to $z'$

  $h_p = mspd$

  $\nu = \theta/s$

- $G_X(x_n, h_p) = \text{Re}(Z_X^*(x_n) Z_p(m\theta)) \quad \rightarrow \quad \cos(\langle \nu, h_p \rangle - H(x_n))$

- $G_X^{\max}(x) = g_{\max}(Z_X(x)) \quad \rightarrow \quad g_{\max} : z \mapsto \max_{\|p\|_{\infty} \leq q} \text{Re}(z^* Z_p)$

- $Z_X(x)$ **uniformly distributed** on the unit circle (Hypothesis)
Sketch of the proof

- **Reformulation** of the problem on the unit circle $S^1 \subset \mathbb{C}$

  $[z, z']_{S^1} \subset S^1$ arc on the unit circle going from $z$ to $z' \quad h_p = msp$

  $Z_X : x \mapsto e^{iHx(x)}$ and $Z_p : \omega \mapsto e^{i\langle \omega, p \rangle} \quad \nu = \theta/s$

- $G_X(x_n, h_p) = \text{Re}(Z_X^*(x_n) Z_p(m\theta)) \quad \cos(\langle \nu, h_p \rangle - H(x_n))$

- $G_X^\max(x) = g_X^\max(Z_X(x)) \quad g_X^\max : z \mapsto \max_{\|p\|_\infty \leq q} \text{Re}(z^*Z_p)$

- $Z_X(x)$ **uniformly distributed** on the unit circle (Hypothesis)

- The $p$-th moment is given by

\[
\mathbb{E} \left[ G_X^\max(x)^p \right] = \frac{1}{2\pi} \int_{S^1} g_X^\max(z)^p \, d\theta(z) = \frac{1}{2\pi} \sum_{i=0}^{n_q-1} \int_{\mathbb{R}^i_{(q)}} g_X^\max(z)^p \, d\theta(z).
\]
Sketch of the proof

- **Reformulation** of the problem on the unit circle $S^1 \subset \mathbb{C}$

  $[z, z']_{S^1} \subset S^1$ arc on the unit circle going from $z$ to $z'$
  
  $h_p = msp$
  
  $\nu = \theta/s$

- $G_X(x_n, h_p) = \text{Re}(Z_X^*(x_n) Z_p(m\theta))$  \[\longrightarrow\]  $\cos(\langle \nu, h_p \rangle - H(x_n))$

- $G_X^{\max}(x) = g_{\max}(Z_X(x))$  \[\longrightarrow\]  $g_{\max} : z \mapsto \max_{\|p\|_\infty \leq q} \text{Re}(z^* Z_p)$

- $Z_X(x)$ **uniformly distributed** on the unit circle (Hypothesis)

- The $p$-th moment is given by

  \[
  \mathbb{E} [G_X^{\max}(x)^p] = \frac{1}{2\pi} \int_{S^1} g_{\max}(z)^p \, d\vartheta(z) = \frac{1}{2\pi} \sum_{i=0}^{n_q-1} \int_{\mathcal{U}_i^{(q)}} g_{\max}(z)^p \, d\vartheta(z). 
  \]

- $\forall z \in \mathcal{U}_i^{(q)}, g_{\max}(z) = \max \left( \text{Re}(z^* Z_i^{(q)}), \text{Re}(z^* Z_{i+1}^{(q)}) \right)$
Sketch of the proof

\[ g_{\text{max}} : z \mapsto \max_{\|p\|_\infty \leq q} \Re(z^* Z_p) \]
Sketch of the proof

- **Reformulation** of the problem on the unit circle $S^1 \subset \mathbb{C}$

- $\forall z \in \mathcal{A}_i^{(q)}, \ g_{\text{max}}(z) = \max \left( \Re(z^* Z_i^{(q)}), \ Re(z^* Z_{i+1}^{(q)}) \right)$
Sketch of the proof

- **Reformulation** of the problem on the unit circle $S^1 \subset \mathbb{C}$

  - $\forall z \in \mathcal{A}_i^{(q)}$, $g_{\text{max}}(z) = \max \left( \text{Re}(\bar{z}^* Z_i^{(q)}), \text{Re}(\bar{z}^* Z_{i+1}^{(q)}) \right)$

  - $\forall z \in \mathcal{A}_i^{(q)}$, $g_{\text{max}}(z) = \text{Re}(\bar{z}^* Z_i^{(q)})$
Sketch of the proof

**Reformulation** of the problem on the unit circle $\mathbb{S}^1 \subset \mathbb{C}$

- $\forall z \in \mathcal{A}_i^{(q)}$, $g_{\text{max}}(z) = \max \left( \text{Re}(z^* Z^{(q)}_i), \text{Re}(z^* Z^{(q)}_{i+1}) \right)$
- $\forall z \in \overline{\mathcal{A}}_i^{(q)}$, $g_{\text{max}}(z) = \text{Re}(z^* Z^{(q)}_i)$

\[
\int_{\mathcal{A}_i^{(q)}} g_{\text{max}}(z)^p \, d\vartheta(z) = 2 \int_{\overline{\mathcal{A}}_i^{(q)}} \text{Re}(z^* Z^{(q)}_i)^p \, d\vartheta(z)
\]

\[
= 2 \int_{H^{(q)}_i} \cos^p \left( \eta - H^{(q)}_i \right) \, d\eta
\]

\[
= 2 \int_0^{\delta H^{(q)}_i / 2} \cos^p \eta' \, d\eta'
\]
Sketch of the proof

- **Reformulation** of the problem on the unit circle $\mathbb{S}^1 \subset \mathbb{C}$

  - $\forall z \in \mathcal{A}_i^{(q)}$, $g_{\text{max}}(z) = \max \left( \Re(\bar{z}^* Z_i^{(q)}), \Re(\bar{z}^* Z_{i+1}^{(q)}) \right)$

  - $\forall z \in \overline{\mathcal{A}}_i^{(q)}$, $g_{\text{max}}(z) = \Re(z^* Z_i^{(q)})$

\[
\int_{\mathcal{A}_i^{(q)}} g_{\text{max}}(z)^p \, d\psi(z) = 2 \int_{\overline{\mathcal{A}}_i^{(q)}} \Re(z^* Z_i^{(q)})^p \, d\psi(z)
\]

\[
= 2 \int_{H_i^{(q)}} \cos^p \left( \eta - H_i^{(q)} \right) \, d\eta
\]

\[
= 2 \int_0^{\delta H_i^{(q)}/2} \cos^p \eta' \, d\eta'
\]

$z \leftarrow e^{i\eta}$
Sketch of the proof

- **Reformulation** of the problem on the unit circle $S^1 \subset \mathbb{C}$

- $\forall z \in \mathcal{A}_i^{(q)}$, $g_{\text{max}}(z) = \max \left( \text{Re}(z^* Z_i^{(q)}), \text{Re}(z^* Z_{i+1}^{(q)}) \right)$

- $\forall z \in \overline{\mathcal{A}}_i^{(q)}$, $g_{\text{max}}(z) = \text{Re}(z^* Z_i^{(q)})$

\[
\int_{\mathcal{A}_i^{(q)}} g_{\text{max}}(z)^p \, d\vartheta(z) = 2 \int_{\overline{\mathcal{A}}_i^{(q)}} \text{Re}(z^* Z_i^{(q)})^p \, d\vartheta(z)
\]

\[
z \leftarrow e^{i\eta}
\]

\[
= 2 \int_{H_i^{(q)}} \cos^p \left( \eta - H_i^{(q)} \right) \, d\eta \quad \overline{H}_i^{(q)} := \left( H_i^{(q)} + H_{i+1}^{(q)} \right) / 2
\]

\[
= 2 \int_0^{\delta H_i^{(q)}/2} \cos^p \eta' \, d\eta'
\]
Sketch of the proof

- **Reformulation** of the problem on the unit circle \( S^1 \subset \mathbb{C} \)

  - \( \forall z \in \mathcal{A}^{(q)}_i, \ g_{\text{max}}(z) = \max \left( \Re(z^* Z_i^{(q)}), \ Re(z^* Z_{i+1}^{(q)}) \right) \)
  
  - \( \forall z \in \overline{\mathcal{A}}^{(q)}_i, \ g_{\text{max}}(z) = \Re(z^* Z_i^{(q)}) \)

\[
\int_{\mathcal{A}^{(q)}_i} g_{\text{max}}(z)^p \, d\psi(z) = 2 \int_{\overline{\mathcal{A}}^{(q)}_i} \Re(z^* Z_i^{(q)})^p \, d\psi(z) \quad z \leftarrow e^{i\eta} \\
= 2 \int_{H_i^{(q)}} \overline{H}_i^{(q)} \cos^p \left( \eta - H_i^{(q)} \right) \, d\eta \\
= 2 \int_0^{\delta H_i^{(q)}/2} \cos^p \eta' \, d\eta' \quad \overline{H}_i^{(q)} := (H_i^{(q)} + H_{i+1}^{(q)})/2 \quad \eta' \leftarrow \eta - H_i^{(q)}
\]
Sketch of the proof
Sketch of the proof

\[ \mathbb{E} [G_X^{\text{max}}(x)] = \frac{1}{\pi} \sum_{i=0}^{n_q-1} \sin \frac{\delta H_i^{(q)}}{2}; \]

\[ \mathbb{E} [G_X^{\text{max}}(x)^2] = \frac{1}{2} + \frac{1}{4\pi} \sum_{i=0}^{n_q-1} \sin \delta H_i^{(q)}. \]
Sketch of the proof

\[ \mathbb{E}[G_X^{\text{max}}(x)] = \frac{1}{\pi} \sum_{i=0}^{n_q-1} \sin \frac{\delta H_i^{(q)}}{2}; \]

\[ \mathbb{E}[G_X^{\text{max}}(x)^2] = \frac{1}{2} + \frac{1}{4\pi} \sum_{i=0}^{n_q-1} \sin \delta H_i^{(q)}. \]

\[ Q_X := 1 - G_X^{\text{max}} \]
Sketch of the proof

\[\mathbb{E} [G_X^{\text{max}}(\mathbf{x})] = \frac{1}{\pi} \sum_{i=0}^{n_q-1} \sin \frac{\delta H_i^{(q)}}{2};\]

\[\mathbb{E} [G_X^{\text{max}}(\mathbf{x})^2] = \frac{1}{2} + \frac{1}{4\pi} \sum_{i=0}^{n_q-1} \sin \delta H_i^{(q)} \]

\[Q_X := 1 - G_X^{\text{max}}\]

By linearity of the expected value

\[\mathbb{E} [Q_X(\mathbf{x})^2] := \frac{3}{2} + \frac{1}{4\pi} \sum_{i=0}^{n_q-1} \left( \sin \delta H_i^{(q)} - 8 \sin \frac{\delta H_i^{(q)}}{2} \right)\]
Main result

- Shift invariance of $\mathbb{R}^\text{Max}$ outputs

$$\theta \mapsto \gamma_q (m \theta)^2$$

$$q = 1$$
Main result

- Shift invariance of $\mathbb{R}^\text{Max}$ outputs

$$\theta \mapsto \gamma_q(m\theta)^2$$

$q = 1$
Main result

- Shift invariance of $\mathbb{R}$Max outputs

\[ \mathbb{E} \left[ \frac{\|Y_1^{\text{max}} - Y_2^{\text{max}}\|_2}{\|Y_1^{\text{mod}}\|_2} \right] \leq 2(\beta_q(m\kappa) + \gamma_q(m\theta)) + \alpha(\kappa u) \]

\[ \theta \mapsto \gamma_q(m\theta)^2 \]

\[ q = 1 \]
Main result

- Shift invariance of $\mathbb{R}^\mathbb{M}_\mathbf{a} \times \mathbf{x}$ outputs

\[
\mathbb{E}\left[\frac{\|Y_1^{\text{max}} - Y_2^{\text{max}}\|_2}{\|Y_1^{\text{mod}}\|_2}\right] \leq 2\left(\beta_q(m\kappa) + \gamma_q(m\theta)\right) + \alpha(\kappa\mu)
\]

Divergence $\mathbb{R}^\text{Max-CMod}$

\[
\theta \mapsto \gamma_q(m\theta)^2
\]

$q = 1$
Main result

Shift invariance of $\mathbb{R}^{\text{Max}}$ outputs

\[ \mathbb{E} \left[ \frac{\|Y_1^{\text{max}} - Y_2^{\text{max}}\|_2}{\|Y_1^{\text{mod}}\|_2} \right] \leq 2(\beta_q(m\kappa) + \gamma_q(m\theta)) + o(\kappa\mu) \]

\[ \theta \mapsto \gamma_q(m\theta)^2 \]

\[ q = 1 \]
Experimental validation
Experimental validation

- What we need:
Experimental validation

- **What we need:**
  - A fully deterministic model with predefined convolution kernels
Experimental validation

What we need:

- A fully deterministic model with predefined convolution kernels
- A set of Gabor-like filters tiling the entire frequency plane and a unique bandwidth share across all filters
Experimental validation

**What we need:**

- A fully deterministic model with predefined convolution kernels
- A set of Gabor-like filters tiling the entire frequency plane and a unique bandwidth share across all filters
- No prediction, only the first layers are implemented
Experimental validation

What we need:

- A fully deterministic model with predefined convolution kernels
- A set of Gabor-like filters tiling the entire frequency plane and a unique bandwidth share across all filters
- No prediction, only the first layers are implemented
- Dataset: ImageNet-1K, validation set (50 000 images)
Experimental validation

What we need:
- A fully deterministic model with predefined convolution kernels
- A set of Gabor-like filters tiling the entire frequency plane and a unique bandwidth share across all filters
- No prediction, only the first layers are implemented
- Dataset: ImageNet-1K, validation set (50 000 images)

Proposed solution:
Experimental validation

What we need:

- A fully deterministic model with predefined convolution kernels
- A set of Gabor-like filters tiling the entire frequency plane and a unique bandwidth share across all filters
- No prediction, only the first layers are implemented
- Dataset: ImageNet-1K, validation set (50 000 images)

Proposed solution:

- The dual-tree complex wavelet packet transform (DT-CWPT)
Experimental validation

What we need:

- A fully deterministic model with predefined convolution kernels
- A set of Gabor-like filters tiling the entire frequency plane and a unique bandwidth share across all filters
- No prediction, only the first layers are implemented
- Dataset: ImageNet-1K, validation set (50,000 images)

Proposed solution:

- The dual-tree complex wavelet packet transform (DT-CWPT)

Experimental validation

What we need:

- A fully deterministic model with predefined convolution kernels
- A set of Gabor-like filters tiling the entire frequency plane and a unique bandwidth share across all filters
- No prediction, only the first layers are implemented
- Dataset: ImageNet-1K, validation set (50 000 images)

Proposed solution:

- The dual-tree complex wavelet packet transform (DT-CWPT)
- The bandwidth and subsampling are controlled by the depth J

Experiments

Filters generated by the DT-CWPT

Case $J = 2$ (two levels of dual-tree decomposition):
- $\kappa = \pi/2$
- $m = 2$
- 32 filters + complex conjugates.
Experiments

Filters generated by the DT-CWPT

Case $J = 2$ (two levels of dual-tree decomposition):

$\kappa = \pi/2$;
$m = 2$;
32 filters + complex conjugates.

![Diagram showing filters generated by the DT-CWPT]
Experiments

Filters generated by the DT-CWPT

Case $J = 2$ (two levels of dual-tree decomposition):

$\kappa = \pi/2$;

$m = 2$;

32 filters + complex conjugates.

 DT-CWPT (real part only)

Spatial domain

Fourier domain
Experiments

Filters generated by the DT-CWPT

Case $J = 2$ (two levels of dual-tree decomposition):

- $\kappa = \pi/2$;
- $m = 2$;
- 32 filters + complex conjugates.

![Diagram showing spatial and Fourier domain filters generated by the DT-CWPT](image)

ResNet-34

DT-CWPT (real part only)
Experiments

Filters generated by the DT-CWPT

**Case** $J = 3$ (three levels of dual-tree decomposition):

- $\kappa = \pi/4$;
- $m = 4$;
- 128 filters + complex conjugates.
Experiments

Filters generated by the DT-CWPT

Case $J = 3$ (three levels of dual-tree decomposition):

$\kappa = \pi/4$;
$\gamma = 4$;
$128$ filters + complex conjugates.

Filters generated by the DT-CWPT
Experiments

Filters generated by the DT-CWPT

Case $J = 3$ (three levels of dual-tree decomposition):

- $\kappa = \pi/4$;
- $m = 4$;
- 128 filters + complex conjugates.

DT-CWPT (subset, real part only)
Experiments

Filters generated by the DT-CWPT

Case $J = 3$ (three levels of dual-tree decomposition):

- $\kappa = \pi/4$;
- $m = 4$;
- 128 filters + complex conjugates.

DT-CWPT (subset, real part only)

AlexNet

Spatial domain

Fourier domain
Experiments

- Normalized MSE between $\text{CMod}$ and $\mathbb{R}_{\text{Max}}$

```
X
  Conv ↓ $m$
  Conv ↓ $2m$
  MaxPool ↓ 2
  Modulus

$Y_{\text{max}}$ $Y_{\text{mod}}$
```
Experiments

Normalized MSE between $\mathcal{C}Mod$ and $\mathcal{R}Max$

\[
\rho^2 = \frac{\|Y_{\text{max}} - Y_{\text{mod}}\|^2_2}{\|Y_{\text{mod}}\|^2_2}
\]
Experiments

- Normalized MSE between $\mathcal{CMod}$ and $\mathcal{RMax}$

$$\rho^2 = \frac{\|Y^{\max} - Y^{\mod}\|_2^2}{\|Y^{\mod}\|_2^2}$$

$$\theta \mapsto \gamma_q(m\theta)^2$$
Experiments

- Normalized MSE between $\mathcal{CMod}$ and $\mathbb{R}_{\text{Max}}$

Outside the scope of our study (low-pass filters)

\[
\rho^2 = \frac{\|Y_{\text{max}} - Y_{\text{mod}}\|_2^2}{\|Y_{\text{mod}}\|_2^2}
\]

\[
\theta \mapsto \gamma_q(m\theta)^2
\]
Experiments

- Normalized MSE between $\text{CMod}$ and $\text{RMax}$

[Diagram showing a ResNet-like scenario with $J = 2$, $m = 2$, $\kappa = \pi/2$.]

$$\rho^2 = \frac{\|Y_{\text{max}} - Y_{\text{mod}}\|_2^2}{\|Y_{\text{mod}}\|_2^2}$$

$$\theta \mapsto \gamma_q(m\theta)^2$$
Experiments

Normalized MSE between $\mathcal{C}\text{Mod}$ and $\mathbb{R}\text{Max}$

AlexNet-like scenario
$J = 3, m = 4, \kappa = \pi/4$

$\rho^2 = \frac{||Y_{\text{max}} - Y_{\text{mod}}||_2^2}{||Y_{\text{mod}}||_2^2}$

$\theta \mapsto \gamma_q(m\theta)^2$
Experiments

- **Shift-invariance of Max outputs**

$$\rho_{\text{max}}^m = \frac{\|Y_1^{\text{max}} - Y_2^{\text{max}}\|_2}{\|Y_1^{\text{mod}}\|_2}$$

$$\theta \mapsto \gamma_q (m\theta)^2$$
Experiments

- **Shift-invariance of $\mathbb{R}^\text{Max}$ outputs**

![Diagram showing shift-invariance](image)

\[ \rho_{\text{max}} = \frac{||Y_1^{\text{max}} - Y_2^{\text{max}}||_2}{||Y_1^{\text{mod}}||_2} \]

\[ \theta \mapsto \gamma_q(m\theta)^2 \]
Conclusion
Conclusion

- Exploration of the *shift invariance* properties captured by the *max pooling* operator, when applied on the *first layer*.
Conclusion

- Exploration of the shift invariance properties captured by the max pooling operator, when applied on the first layer.

- Bridge between complex and standard real CNNs.
Conclusion

- Exploration of the *shift invariance* properties captured by the *max pooling* operator, when applied on the *first layer*.

- Bridge between *complex* and standard *real* CNNs.

- Study of the *discrete* case in a *probabilist* framework.
Conclusion

- Exploration of the **shift invariance** properties captured by the **max pooling** operator, when applied on the **first layer**.

- Bridge between **complex** and standard **real** CNNs.

- Study of the **discrete** case in a **probabilist** framework.

- Establish a **validity domain** for near-shift invariance.
Conclusion

- Exploration of the **shift invariance** properties captured by the **max pooling** operator, when applied on the **first layer**.

- Bridge between **complex** and standard **real** CNNs.

- Study of the **discrete** case in a **probabilist** framework.

- Establish a **validity domain** for near-shift invariance.

- **Experimental setting** based on the **dual-tree** complex wavelet packet transform
Conclusion

- Exploration of the shift invariance properties captured by the max pooling operator, when applied on the first layer.

- Bridge between complex and standard real CNNs.

- Study of the discrete case in a probabilist framework.

- Establish a validity domain for near-shift invariance.

- Experimental setting based on the dual-tree complex wavelet packet transform

- \( \text{CMod} \) operator can serve as a stable proxy for \( \text{RMax} \) enabling to improve shift invariance in CNNs architecture while preserving high-frequency information.
Publications

- Hubert Leterme, Kévin Polisano, Valérie Perrier, Karteek Alahari. *Modélisation Parcimonieuse de CNNs avec des Paquets d'Ondelettes Dual-Tree*. ORASIS 2021 - Journées francophones des jeunes chercheurs en vision par ordinateur, Centre National de la Recherche Scientifique [CNRS], Sep 2021, Saint Ferréol, France. pp.1-9. ⟨hal-03339792v2⟩

- Hubert Leterme, Kévin Polisano, Valérie Perrier, Karteek Alahari. *On the Shift Invariance of Max Pooling Feature Maps in Convolutional Neural Networks*. 2023. ⟨hal-03779434v2⟩

- Hubert Leterme, Kévin Polisano, Valérie Perrier, Karteek Alahari. *From CNNs to Shift-Invariant Twin Models Based on Complex Wavelets*. 2023. ⟨hal-03880520v2⟩
Thank you!