

Introduction to group equivariant deep learning

Erik J Bekkers - <u>e.j.bekkers@uva.nl</u> - 17 August 2021 - DTU, DIKU & AAU Summer School on Geometric Deep Learning

Content

Part I: Introduction to group convolutions

- * Motivation
- * Introduction to group theory
- * Regular group convolutional neural networks
- * Applications

Part II: General theory for group equivariant deep learning

- * Group convolutions are all you need!
- * Deeper into group theory: representation theory, homogeneous spaces
- * Characterization of types of group equivariant layers

Part III: Steerable group convolutions

* Deep dive into group theory: irreducible representations, steerable operators and vector spaces * Examples of steerable group convolutions: Spherical data and Volumetric data/3D point clouds

Lecture notes, slides and exercises available at https://uvagedl.github.io

An Introduction to Equivariant Convolutional	II	Steerable group co
Neural Networks for Continuous Groups Lecture Notes	4	Mathematical tools: re Clebsch-Gordan tensor
dr.ir. Erik J. Bekkers		4.1 The groups E(3) and 4.1.1 The groups E 4.1.2 The sphere S ²
August 2021		4.2 Steerable vectors, Wig tions
Contents		4.3 Fourier transform on 4.3.1 Spherical Harn 4.3.2 Irreducible rep 4.4 Clebsch-Gordan prod
I Regular Group Convolutions 3	5	Steerable group convolu- 5.1 SO(3) Equivariant S
1 Introduction to convolutions 4		Data
1.1 Notations		5.1.1 Steerable isotr 5.1.2 Fourier-based
1.2 Convolution and cross-correlation 5 1.3 Convolutions/correlations seen as template matching 6		5.1.2 Fourier-based 5.1.3 Conclusion .
1.3 Convolutions/correlations seen as template matching 6 1.4 Group correlations 8		5.2 SE(3) Equivariant St
1.4 Group conclusions		5.2.1 Steerable func
Mathematical tools: basic group theory 9		5.2.2 Lifting convolu-
2.1 Groups		5.2.3 From regular
2.2 Group actions and representations (groups acting on vectors and		a spherical has
functions)		5.2.4 Steerable grou
2.3 Homogeneous space		5.2.5 Conclusion
2.4 Equivariance 18 2.5 Integration on G using the Haar measure 18	6	Steerable graph neural
3 Regular Group Convolutional Neural Networks 21		
3.1 Artifical Neural Networks for Vectors	II	I Lie group equivar
3.2 Artificial Neural Networks for Signals		
3.3 Equivariant Neural Networks for Signals		
3.4 Types of equivariant layers		
3.4.1 Isotropic & convolutions $(X = Y = K)$		
3.4.3 Group convolution layer		
3.4.4 Projection layer $(X = G, Y = \mathbb{R}^d)$		
3.4.5 Global pooling $(X = G, Y = \emptyset)$		
3.5 Designing equivariant neural networks		
1		
_		

convolutional neural networks	30
representation theory for $SO(3)$ and the	
or product	31
nd $O(3)$ and homogeneous space S^2	32
E(3) and $O(3)$	32
S^2 is a homogeneous space of $O(3)$ Wigner-D matrices and irreducible representa-	32
	33
on S^2 and $SO(3)$	34
Iarmonics	35
representations as a Fourier basis	36
roduct and steerable linear layers	38
volutions t Steerable Group Convolutions for Spherical	42
· · · · · · · · · · · · · · · · · · ·	42
sotropic S^2 convolutions $(X = Y = S^2)$	42
ed Group convolutions in vectorized form	43
Steerable Group Convolutions	45 46
inctions.	46
volution $(X = \mathbb{R}^d, Y = SE(3))$	46
ar to steerable group convolutions: kernels in	
harmonic basis.	48
roup convolutions	49
	50
al networks and Point Convolutions	51
variant neural networks	52

Exercise 2.4. Show transitivity (Definition 2.8) of the action of G given in Eq. (29).

Example 2.7 (Quotient space $\mathbb{R}^d = SE(d)/SO(d)$). Let $H = (\{0\} \times SO(d)$ the subgroup of rotations in SE(d), with 0 the identity element of the translationg roup $(\mathbb{R}^d, +)$. The the cosets gH are given by

> $gH = \{g \cdot (\mathbf{0}, \tilde{\mathbf{R}}) \mid \tilde{\mathbf{R}} \in SO(d)\}$ $= \{ (\mathbf{Re} + \mathbf{x}, \mathbf{R}\tilde{\mathbf{R}}) | h \in SO(d) \}$ $= \{(\mathbf{x}, \mathbf{R}\tilde{\mathbf{R}}) | \tilde{\mathbf{R}} \in SO(d)\}$ $= \{(\mathbf{x}, \tilde{\mathbf{R}}) | \tilde{\mathbf{R}} \in SO(d)\},\$

with $q = (\mathbf{x}, \mathbf{R})$. So, the cosets are given by all possible rotations for a fixed translation vector x, the vector x thus indexes these sets. We can therefore make the identification

 $\mathbb{R}^{d} \equiv SE(d)/SO(d)$.

We already saw in Exercise 2.1 that \mathbb{R}^d is a homogeneous space of SE(d), this is a consequence of Lemma 2.1.

Lemma 2.1 shows that a quotient space G/H of a group G with H is a homogeneous space. We can also approach this in the other direction and state that any homogeneous space of G is equivalent to a quotient space G/H for some H. This is stated in the following Lemma, for which we first need to introduce the notion of a stabilizer.

Definition 2.13 (Stabilizer). Let G act on X via the action \odot . For every $x \in X$, the stabilizer subgroup of G with respect to the point x is denoted with $Stab_G(x)$ is the set of all elements in G that fix x:

> $\operatorname{Stab}_G(x) = \{g \in G \mid g \odot x = x\}.$ (30)

Lemma 2.2. Let X be a homogeneous space of G. Then X can be identified with G/H with $H = \operatorname{Stab}_G(x_0)$ for any $x_0 \in X$.

Affine groups Finally when it comes to types of groups and homogeneous spaces we note that often we are interested in groups that act on \mathbb{R}^d , as most often one deal with data on \mathbb{R}^d . It is therefore useful to introduce the following class of groups.

Definition 2.14 (Affine groups). Affine groups $G = \mathbb{R}^d \rtimes H$ are a class of groups that are the semidirect product of a group $H \subseteq GL(\mathbb{R}^d)$ acting on \mathbb{R}^d , with $GL(\mathbb{R}^d$ the group of general linear transformations acting on \mathbb{R}^d .

The transformations in $H \subseteq GL(\mathbb{R}^d)$ are commonly represented as invertible matrices A which act on \mathbb{R}^d via matrix-vector multiplication, by which the group

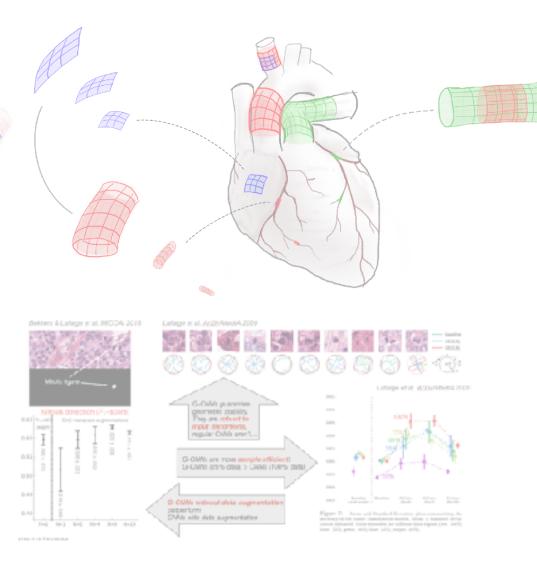


Part I

- 1. Why do we want equivariant learning models? - Geometric guarantees + weight sharing/sample efficiency
- 2. A group theoretical view on recognition by components (capsule nets) - Group theoretical prerequisites (group product and representations) - Group convolutions perform pattern recognition by components

- 3. Experimental examples







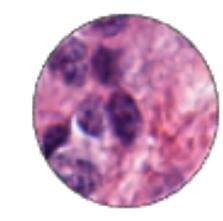
Content of this talk

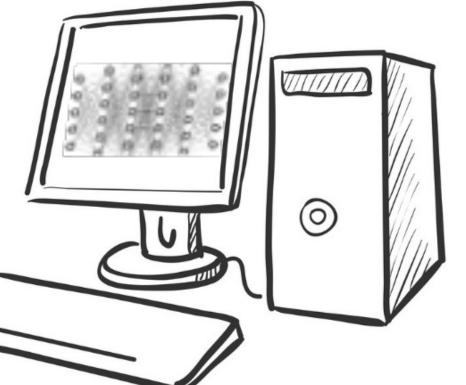
- 1. Why do we want equivariant learning models? - Geometric guarantees + weight sharing/sample efficiency
- 2. A group theoretical view on recognition by components (capsule nets) - Group theoretical prerequisites (group product and representations) - Group convolutions perform pattern recognition by components
- 3. Experimental examples

4. Theorem: Linear maps between feature maps are equivariant iff they are group convolutions

Motivation: Geometric Guarantees

Example: Detection of pathological cells









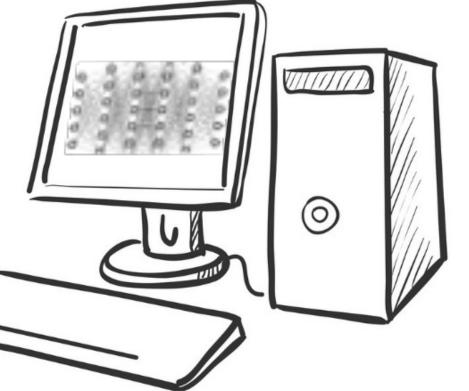
(invariance)

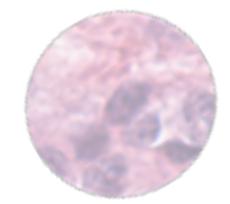


Motivation: Geometric Guarantees

Example: Detection of pathological cells











(invariance)

Healthy

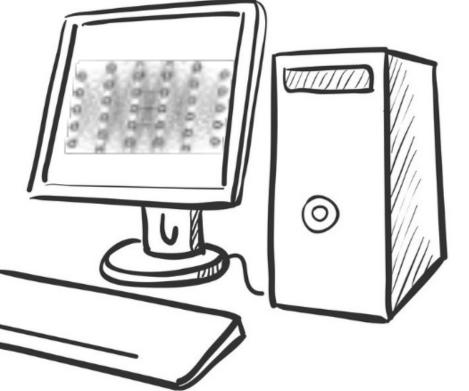
?

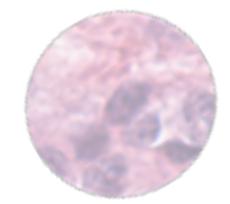


Motivation: Geometric Guarantees

Example: Detection of pathological cells











(invariance)

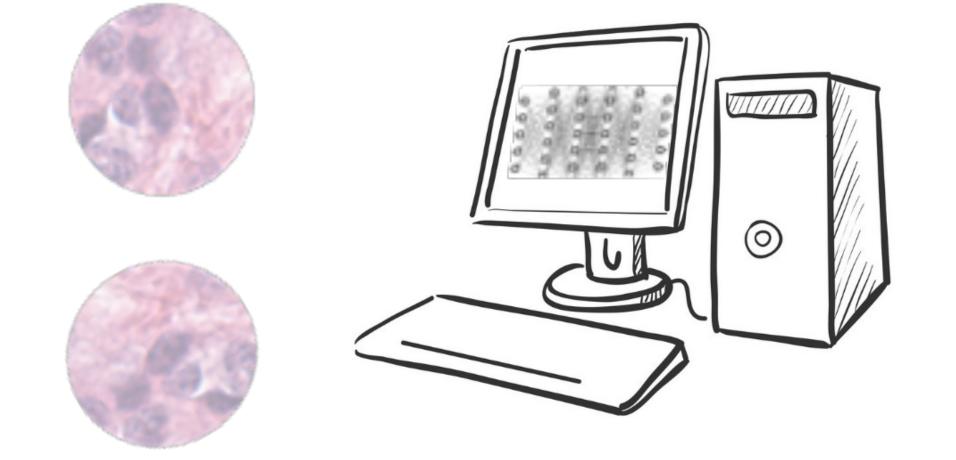
Healthy

? **Pathological**

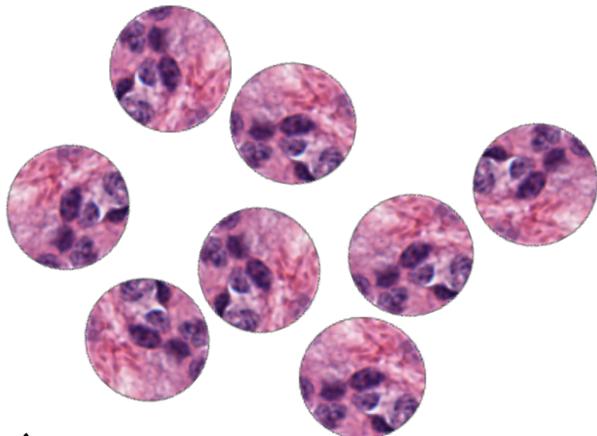


Motivation: Geometric Guarantees

Example: Detection of pathological cells



Common approach: data-augmentation





(invariance)

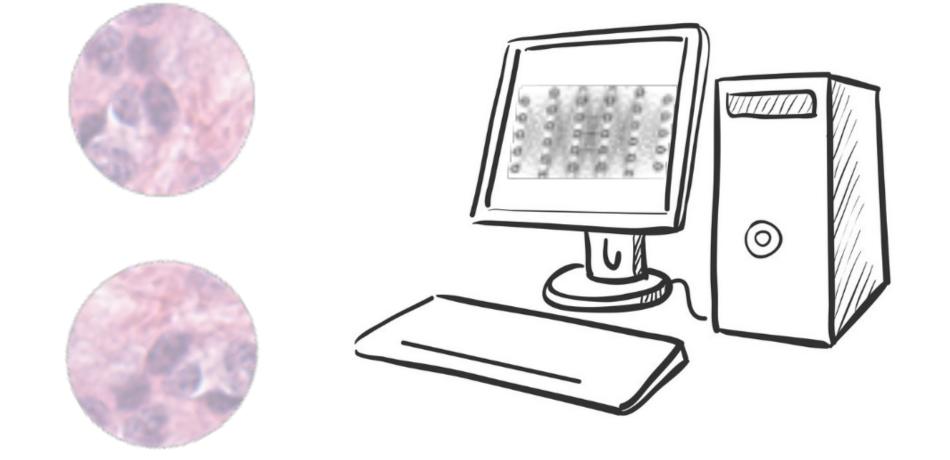
Healthy

? Pathological

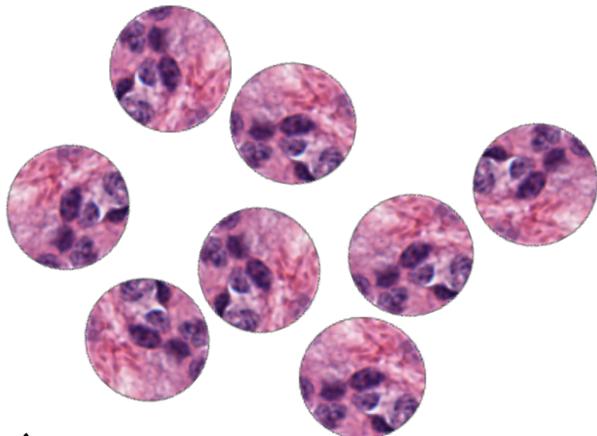


Motivation: Geometric Guarantees

Example: Detection of pathological cells



Common approach: data-augmentation





(invariance)

Healthy

?

Pathological

Issues:

- Still no guarantee of invariance
- Valuable net capacity is spend on learning invariance
- Redundancy in feature repr.



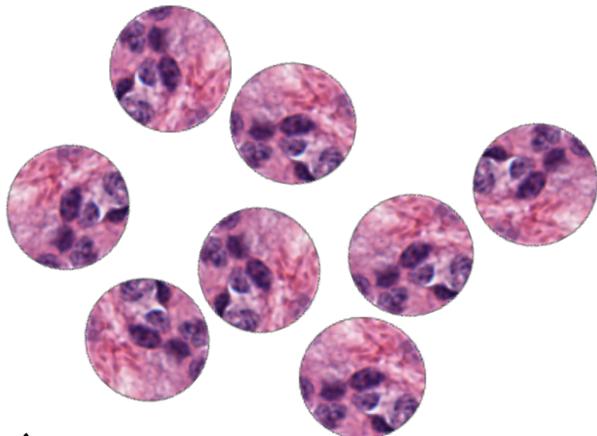


Motivation: Geometric Guarantees

Example: Detection of pathological cells



Common approach: data-augmentation



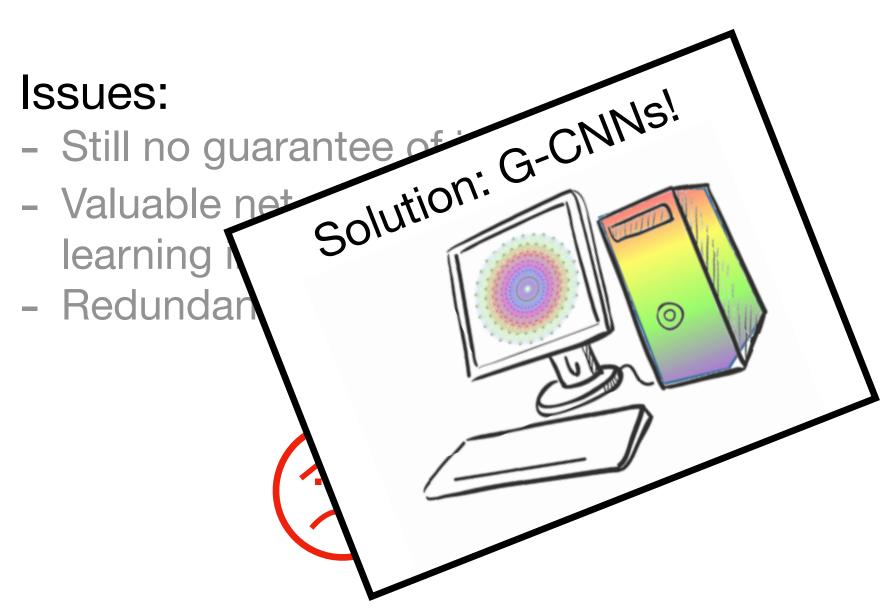


(invariance)

Healthy

?

Pathological





Motivation: Geometric Guarantees (equivariance)

CNNs are translation equivariant







Via convolutions



Motivation: Geometric Guarantees (equivariance)

CNNs are translation equivariant



Via convolutions







Motivation: Geometric Guarantees (equivariance)

CNNs are not rotation equivariant









Stabilized view





Motivation: Geometric Guarantees (equivariance)

CNNs are not rotation equivariant







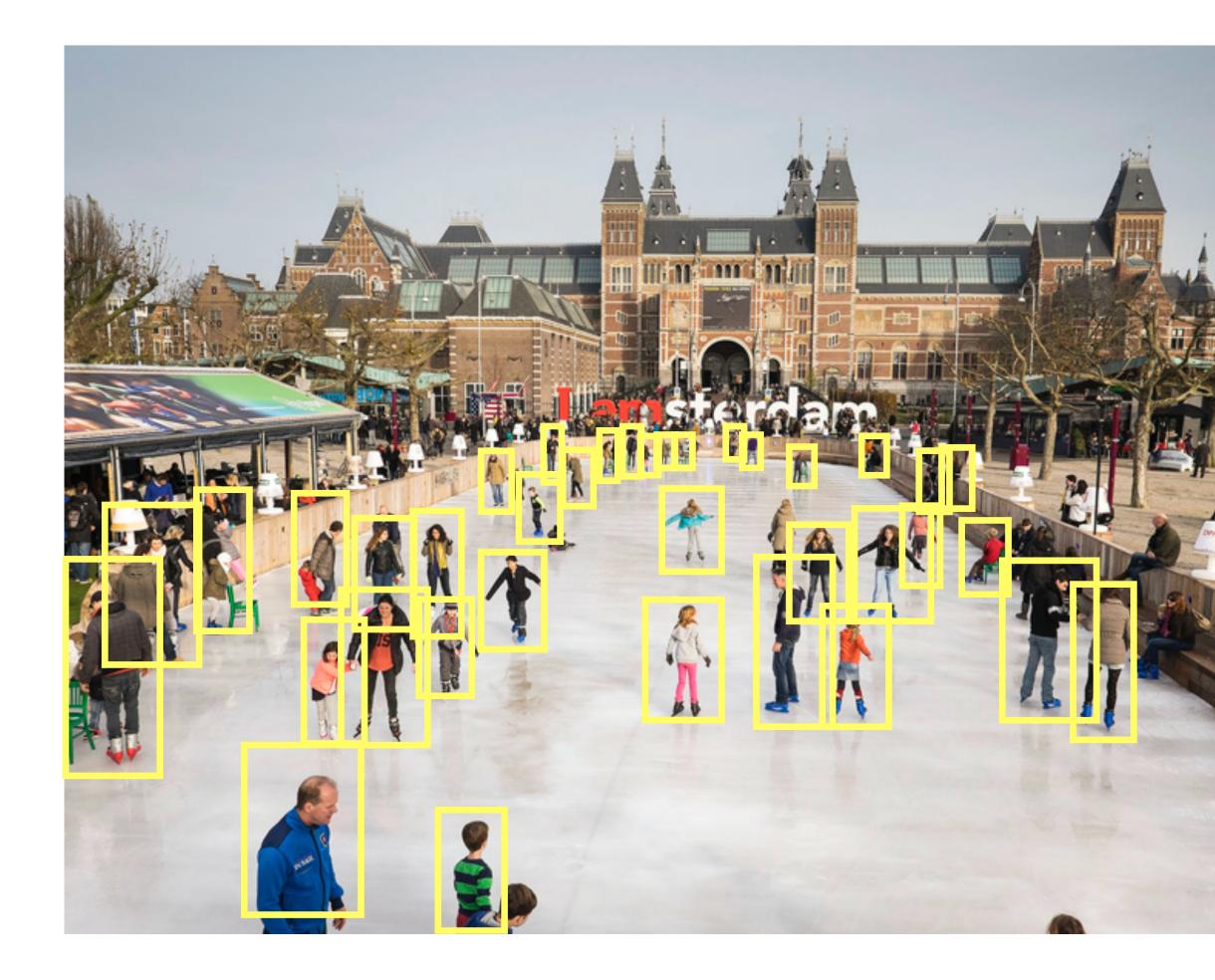


Stabilized view





Motivation: Geometric Guarantees (equivariance)





Importance of equivariance:

- No information is lost when the input is transformed
- Guaranteed stability to (local + global) transformations

Group convolutions:

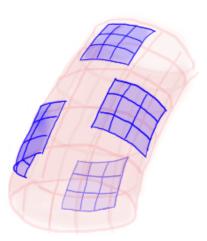
- Equivariance beyond translations
- Geometric guarantees
- Increased weight sharing

G-CNNs are not only relevant for invariant problems but for any type of structured data!

Motivation: Recognition by components In a group theoretical setting

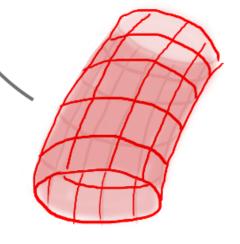
Ħ

Low-level features (e.g. local surfaces)



features can appear at arbitrary locations, angles, and scales

Low-level features arranged at relative angles and displacements form *mid-level features*



Mid-level features (e.g. vessel segments)

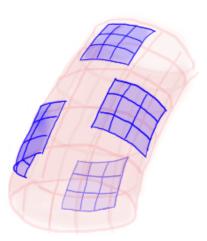




Motivation: Recognition by components In a group theoretical setting

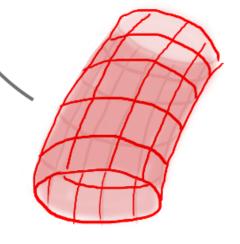
Ħ

Low-level features (e.g. local surfaces)



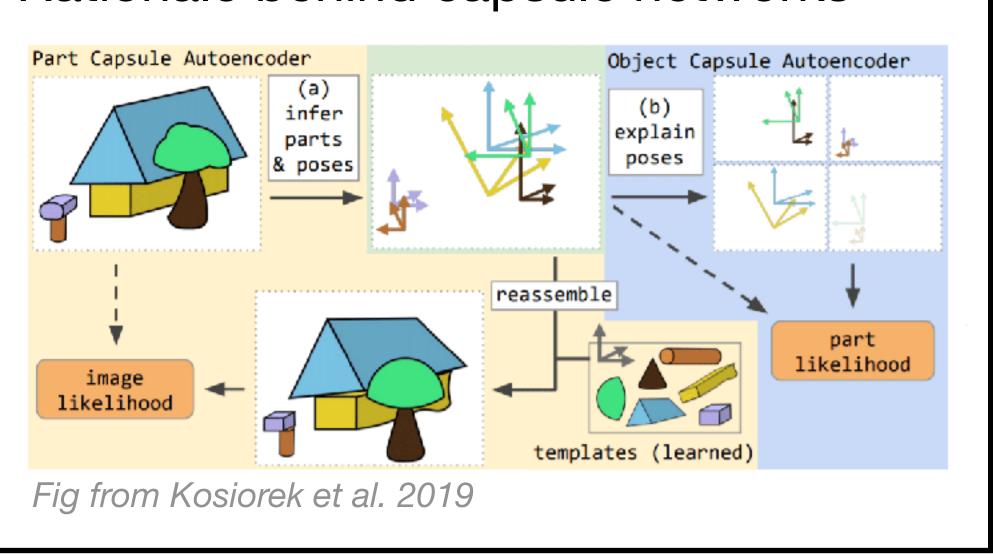
features can appear at arbitrary locations, angles, and scales

Low-level features arranged at relative angles and displacements form *mid-level features*



Mid-level features (e.g. vessel segments)

Rationale behind capsule networks





Content of this talk

1. Why do we want equivariant learning models? - Geometric guarantees + weight sharing/sample efficiency

2. A group theoretical view on **recognition by components** (capsule nets) - Group theoretical prerequisites (group product and representations) - Group convolutions perform pattern recognition by components

3. Experimental examples

4. Theorem: Linear maps between feature maps are equivariant iff they are group convolutions



What is a group?

operator, that satisfies the following four axioms:

- **Closure**: Given two elements g and h of G, the product $g \cdot h$ is also in G.
- Associativity: For $g, h, i \in G$ the product \cdot is associative, i.e., $g \cdot (h \cdot i) = (g \cdot h) \cdot i$.



A group (G, \cdot) is a set of elements G equipped with a group product \cdot , a binary

Identity element: There exists an identity element $e \in G$ such that $e \cdot g = g \cdot e = g$ for any $g \in G$.

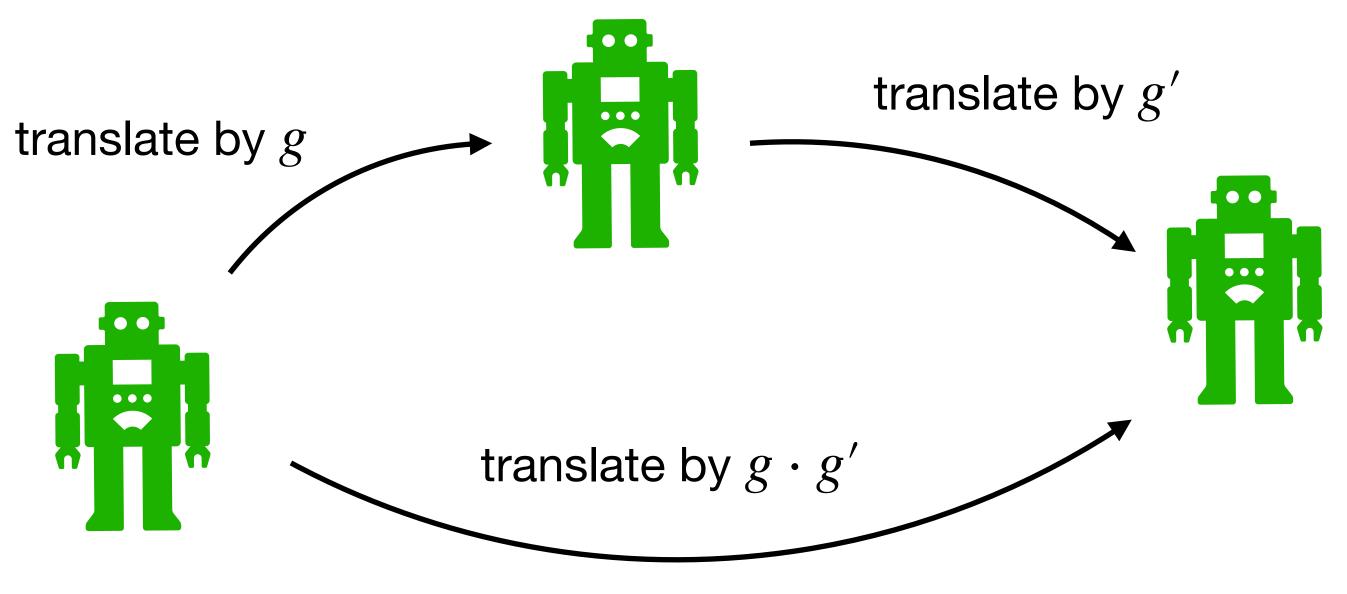
• Inverse element: For each $g \in G$ there exists an inverse element $g^{-1} \in G$ s.t. $g^{-1} \cdot g = g \cdot g^{-1} = e$.



The translation group (\mathbb{R}^2 , +)

The translation group consists of all possible translations in \mathbb{R}^2 and is equipped with the group product and group inverse:

with $g = (\mathbf{x}), g' = (\mathbf{x}')$ and $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^2$.



 $g \cdot g' = (\mathbf{x} + \mathbf{x}')$ $g^{-1} = (-\mathbf{x})$



The roto-translation group SE(2)

The group $SE(2) = \mathbb{R}^2 \rtimes SO(2)$ consists of the coupled space $\mathbb{R}^2 \times S^1$ of translations vectors in \mathbb{R}^2 , and rotations in SO(2) (or equivalently orientations in S^1), and is equipped with the group product and group inverse:

$$g \cdot g' = (\mathbf{x}, \mathbf{R}_{\theta}) \cdot (\mathbf{x}', \mathbf{R}_{\theta'}) = (\mathbf{R}_{\theta} \mathbf{x}' + \mathbf{x}, \mathbf{R}_{\theta + \theta'})$$

$$g^{-1} = (-\mathbf{R}_{\theta}^{-1} \mathbf{x}, \mathbf{R}_{\theta}^{-1})$$

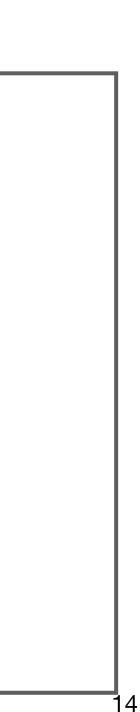
$$\mathbf{R}_{\theta'}).$$
roto-translate by g
roto-translate by $g \cdot g'$

with $g = (\mathbf{x}, \mathbf{R}_{\theta}), g' = (\mathbf{x}', \mathbf{R}_{\theta})$

$$e_{\theta} \cdot (\mathbf{x}', \mathbf{R}_{\theta'}) = (\mathbf{R}_{\theta}\mathbf{x}' + \mathbf{x}, \mathbf{R}_{\theta+\theta'})$$

o-translate by g
roto-translate by g'
roto-translate by $g \cdot g'$

2D **S**pecial **E**uclidean motion group



The scale-translation group $\mathbb{R}^2 \rtimes \mathbb{R}^+$

factors in \mathbb{R}^+ , and is equipped with the group product and group inverse:

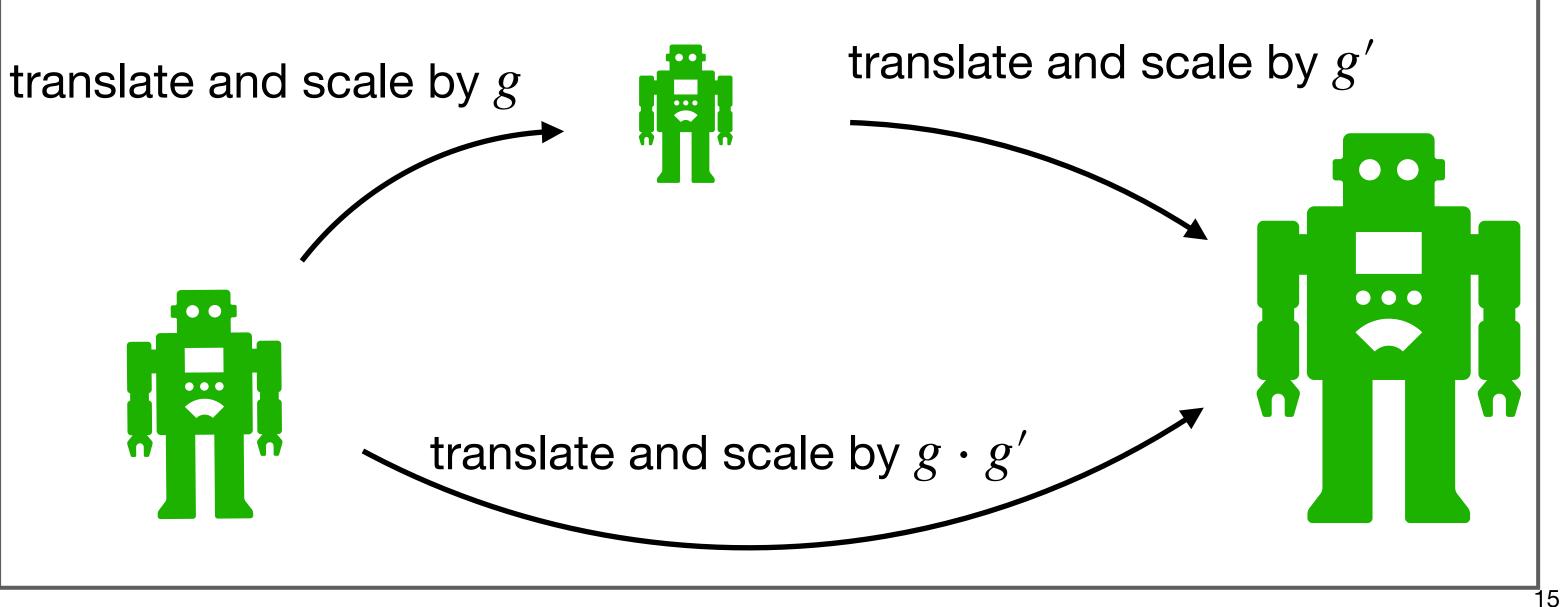
$$g \cdot g' = (\mathbf{x}, s) \cdot (\mathbf{x}', s') = (s\mathbf{x}' + \mathbf{x}, ss')$$
$$g^{-1} = \left(-\frac{1}{s}\mathbf{x}, \frac{1}{s}\right)$$

with
$$g = (x, s), g' = (x', s')$$
.

with $g \cdot g^{-1} = e = (0, 1)$ matrix repr: $\mathbf{G} = \begin{pmatrix} \mathbf{I}_s & \mathbf{X} \\ \mathbf{O}_T & \mathbf{I} \end{pmatrix}$

UNIVERSITY OF AMSTERDAM

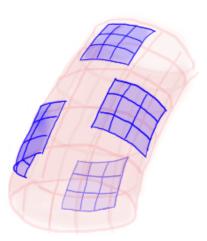
The scale-translation group of space $\mathbb{R}^2 \times \mathbb{R}^+$ of translations vectors in \mathbb{R}^2 and scale/dilation



Motivation: Recognition by components In a group theoretical setting

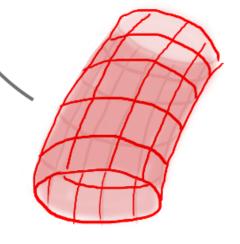
Ħ

Low-level features (e.g. local surfaces)



features can appear at arbitrary locations, angles, and scales

Low-level features arranged at relative angles and displacements form *mid-level features*



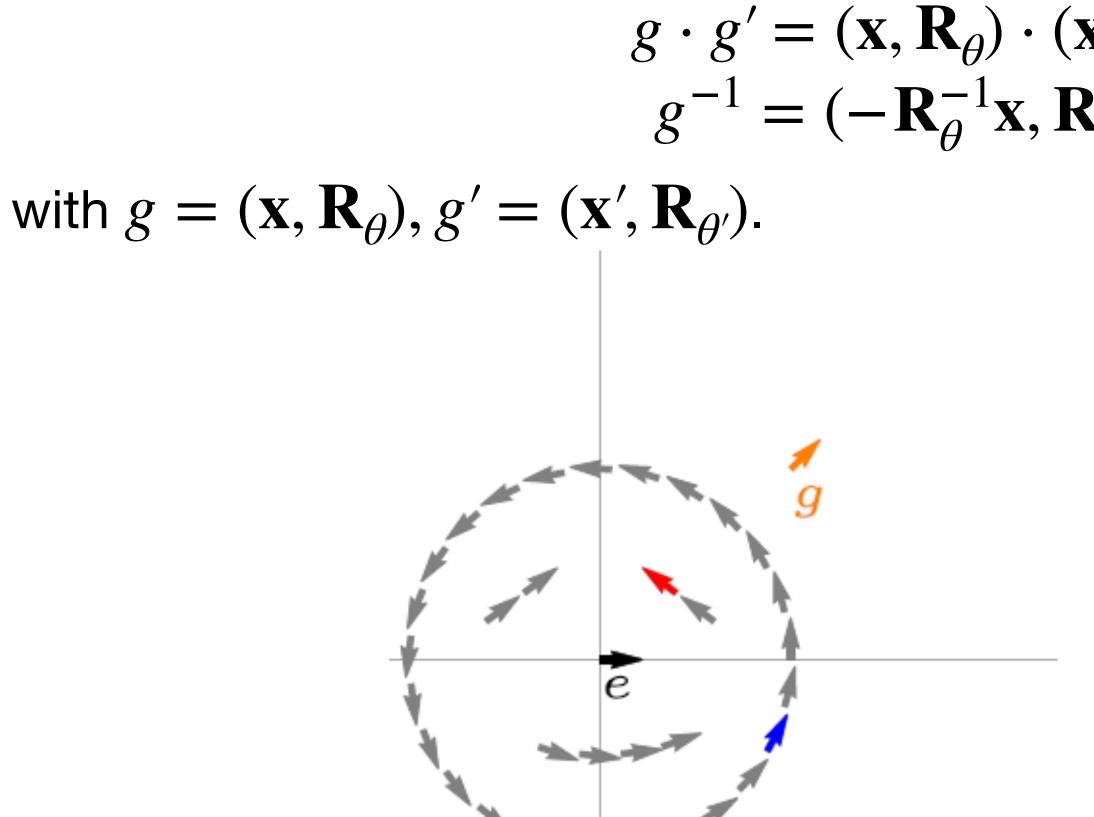
Mid-level features (e.g. vessel segments)





The roto-translation group SE(2)

The group $SE(2) = \mathbb{R}^2 \rtimes SO(2)$ consists of the coupled space $\mathbb{R}^2 \times S^1$ of translations vectors in \mathbb{R}^2 , and rotations in SO(2) (or equivalently orientations in S^1), and is equipped with the group product and group inverse:



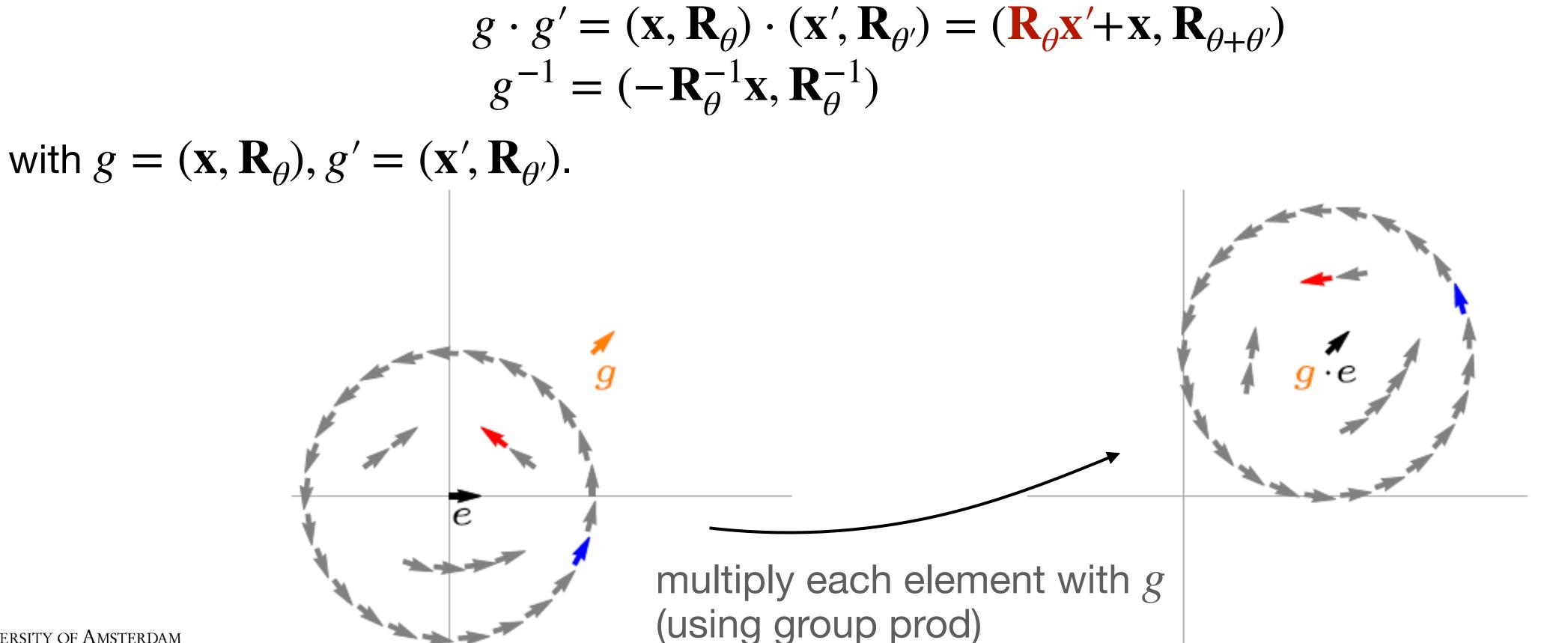
2D Special Euclidean motion group

$$\mathbf{x}', \mathbf{R}_{\theta'}) = (\mathbf{R}_{\theta}\mathbf{x}' + \mathbf{x}, \mathbf{R}_{\theta+\theta'})$$
$$\mathbf{R}_{\theta}^{-1})$$



The roto-translation group SE(2)

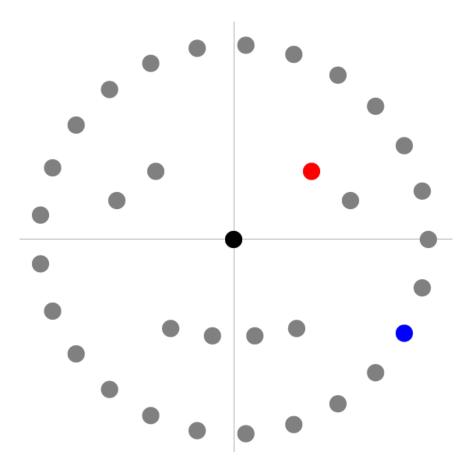
The group $SE(2) = \mathbb{R}^2 \rtimes SO(2)$ consists of the coupled space $\mathbb{R}^2 \times S^1$ of translations vectors in \mathbb{R}^2 , and rotations in SO(2) (or equivalently orientations in S^1), and is equipped with the group product and group inverse:



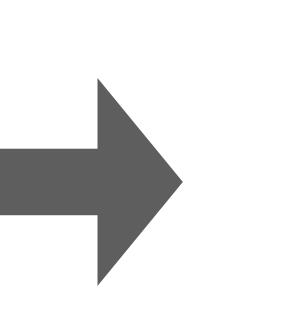
2D Special Euclidean motion group



Set of points (group elements)



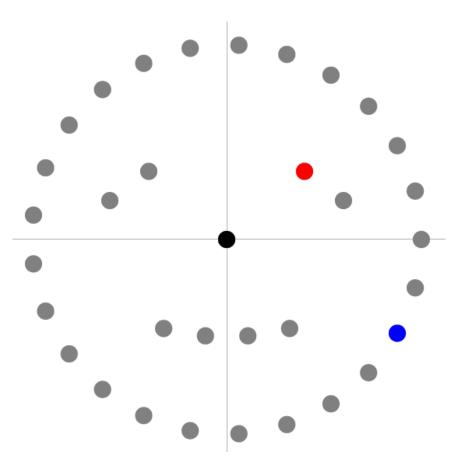
Convolution kernel







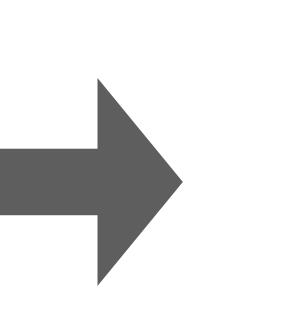
Set of points (group elements)



 $\{g_1, g_2, \dots\} \subset G = (\mathbb{R}^2, +)$

"A collection of parts in certain poses"

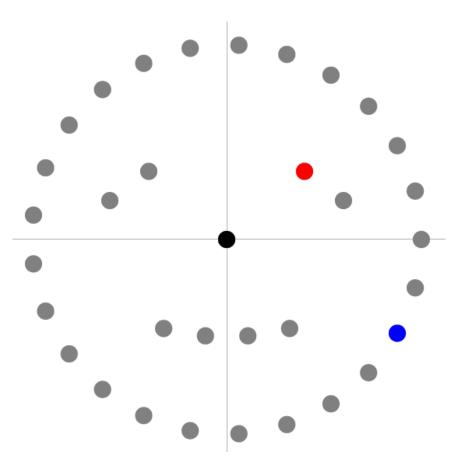
Convolution kernel







Set of points (group elements)



 $\{g_1, g_2, \dots\} \subset G = (\mathbb{R}^2, +)$

"A collection of parts in certain poses"

Convolution kernel



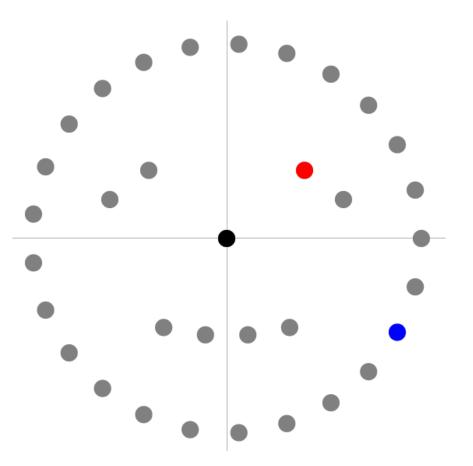
 $k \in \mathbb{L}_2(\mathbb{R}^2)$

"Assigning weights to relative poses"





Set of points (group elements)



 $\{g_1, g_2, \dots\} \subset G = (\mathbb{R}^2, +)$

"A collection of parts in certain poses"

Transforms via group product

Convolution kernel



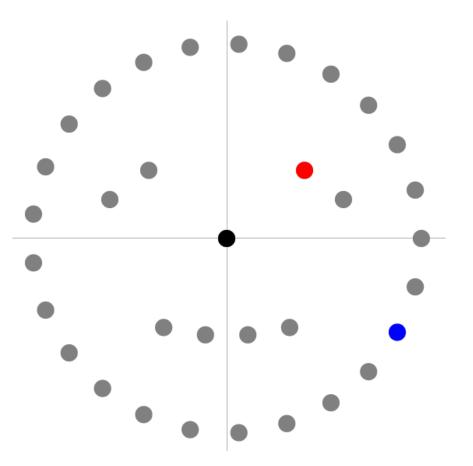
 $k \in \mathbb{L}_2(\mathbb{R}^2)$

"Assigning weights to relative poses"





Set of points (group elements)



 $\{g_1, g_2, \dots\} \subset G = (\mathbb{R}^2, +)$

"A collection of parts in certain poses"

Transforms via group product

Convolution kernel



 $k \in \mathbb{L}_2(\mathbb{R}^2)$

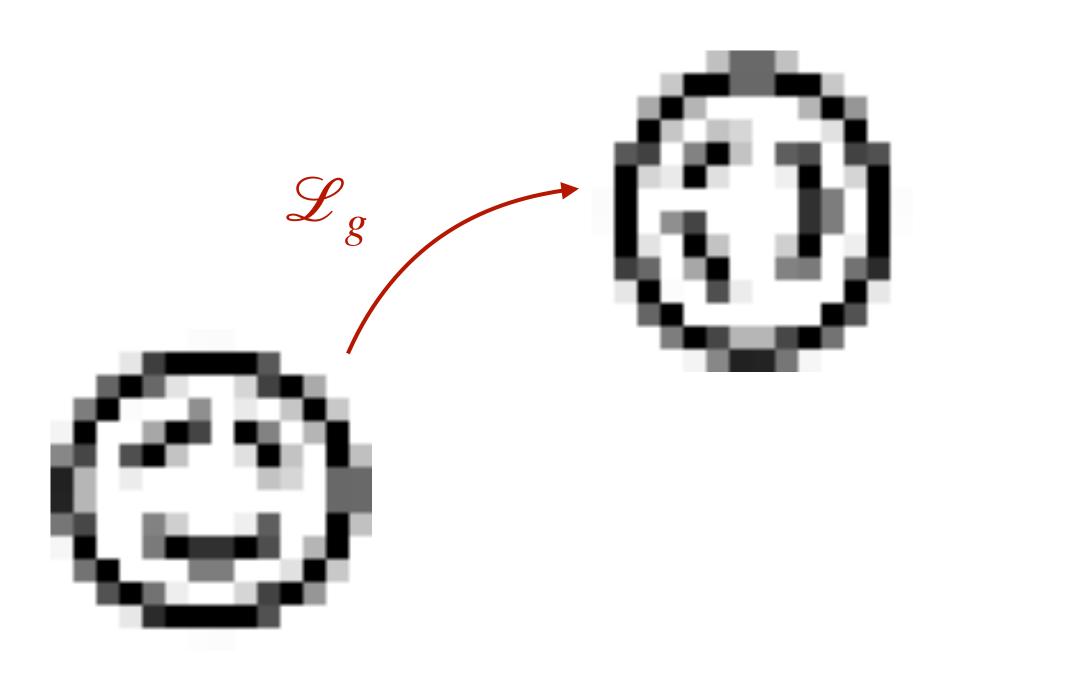
"Assigning weights to relative poses"

Transforms via group representations





Representations



A linear operator \mathscr{L}_g that is parameterized by group elements $g \in G$ that the group structure in the following way

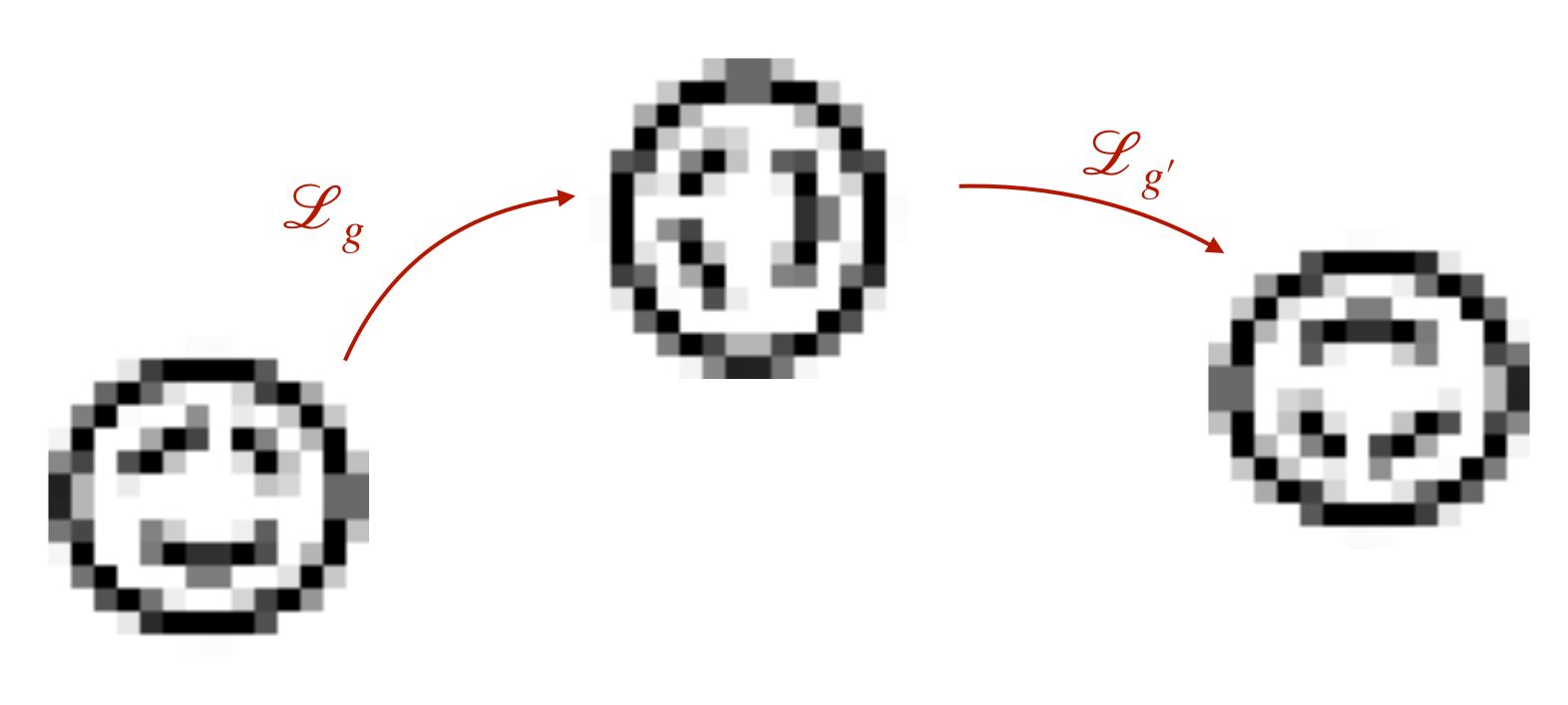


- transforms some object f (e.g. an image) is called a representation of G if it caries
 - $\mathscr{L}_{g'}(\mathscr{L}_g(f)) = \mathscr{L}_{g'\cdot g}(f)$





Representations



A linear operator \mathscr{L}_g that is parameterized by group elements $g \in G$ that the group structure in the following way

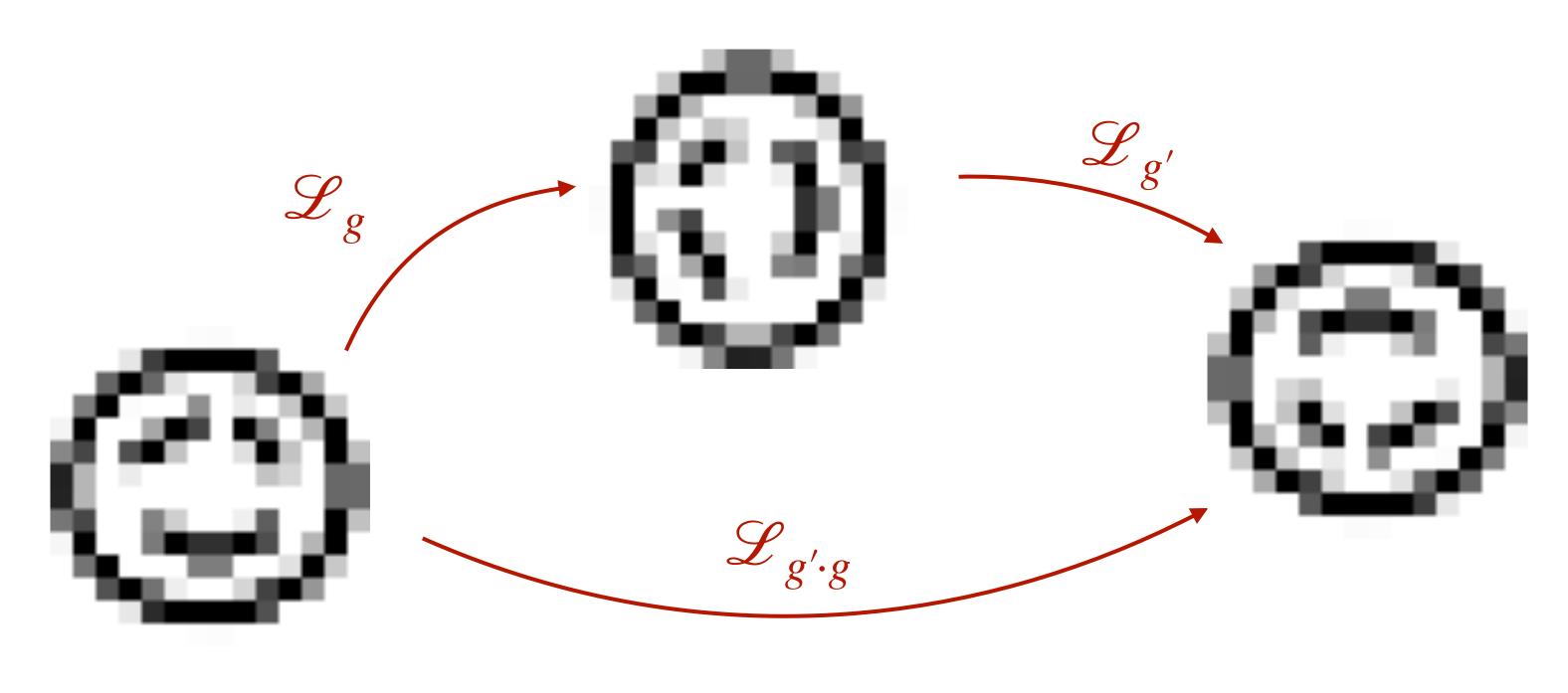


- transforms some object f (e.g. an image) is called a representation of G if it caries
 - $\mathscr{L}_{g'}(\mathscr{L}_g(f)) = \mathscr{L}_{g'\cdot g}(f)$





Representations



A linear operator \mathscr{L}_g that is parameterized by group elements $g \in G$ that the group structure in the following way



- transforms some object f (e.g. an image) is called a representation of G if it caries
 - $\mathscr{L}_{g'}(\mathscr{L}_g(f)) = \mathscr{L}_{g'\cdot g}(f)$





Left-regular representations

Example:

$$f \in \mathbb{L}_2(\mathbb{R}^2)$$

- a 2D image

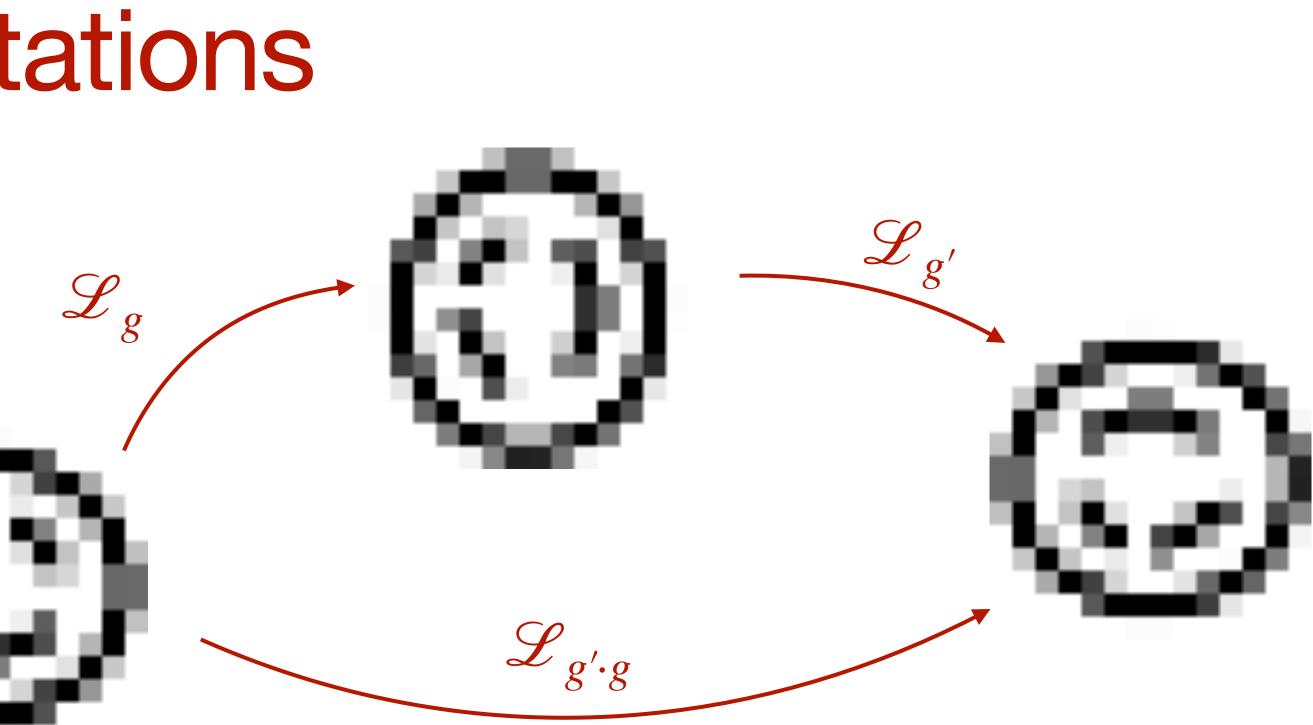
G = SE(2)- the roto-translation group



$\mathscr{L}_{g}(f)(\mathbf{y}) = f(\mathbf{R}_{\theta}^{-1}(\mathbf{y} - \mathbf{x}))$ - a roto-translation of the image

The left-regular representation of G transforms functions by acting on the domain on which they are defined via





"group action" equals group product when X = G





Group actions

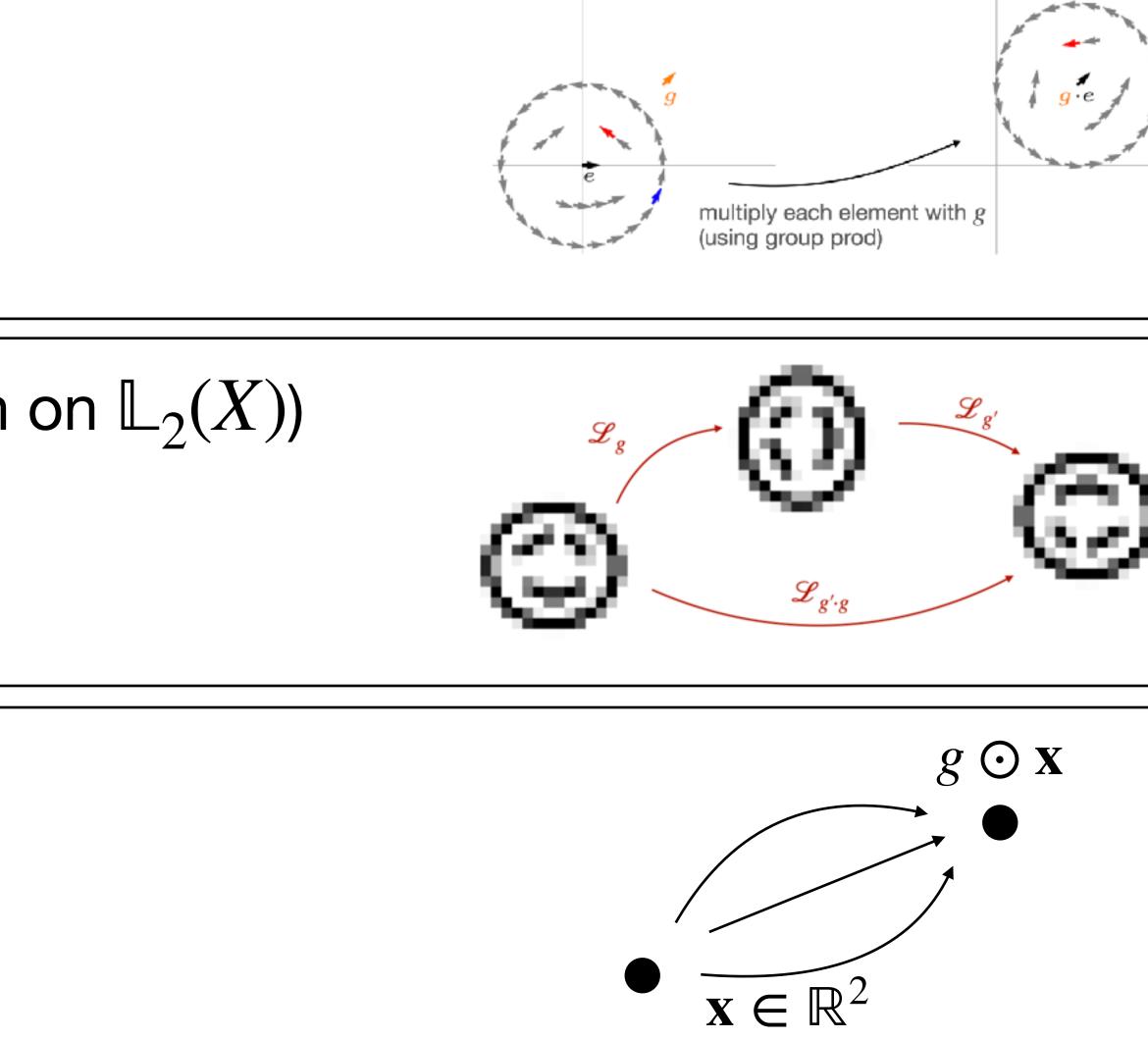
Group product (the action on G)

 $g \cdot g'$

Left regular representation (the action on $\mathbb{L}_2(X)$)

 $\mathscr{L}_{q}f$

Group action (the action on \mathbb{R}^d)





Group actions

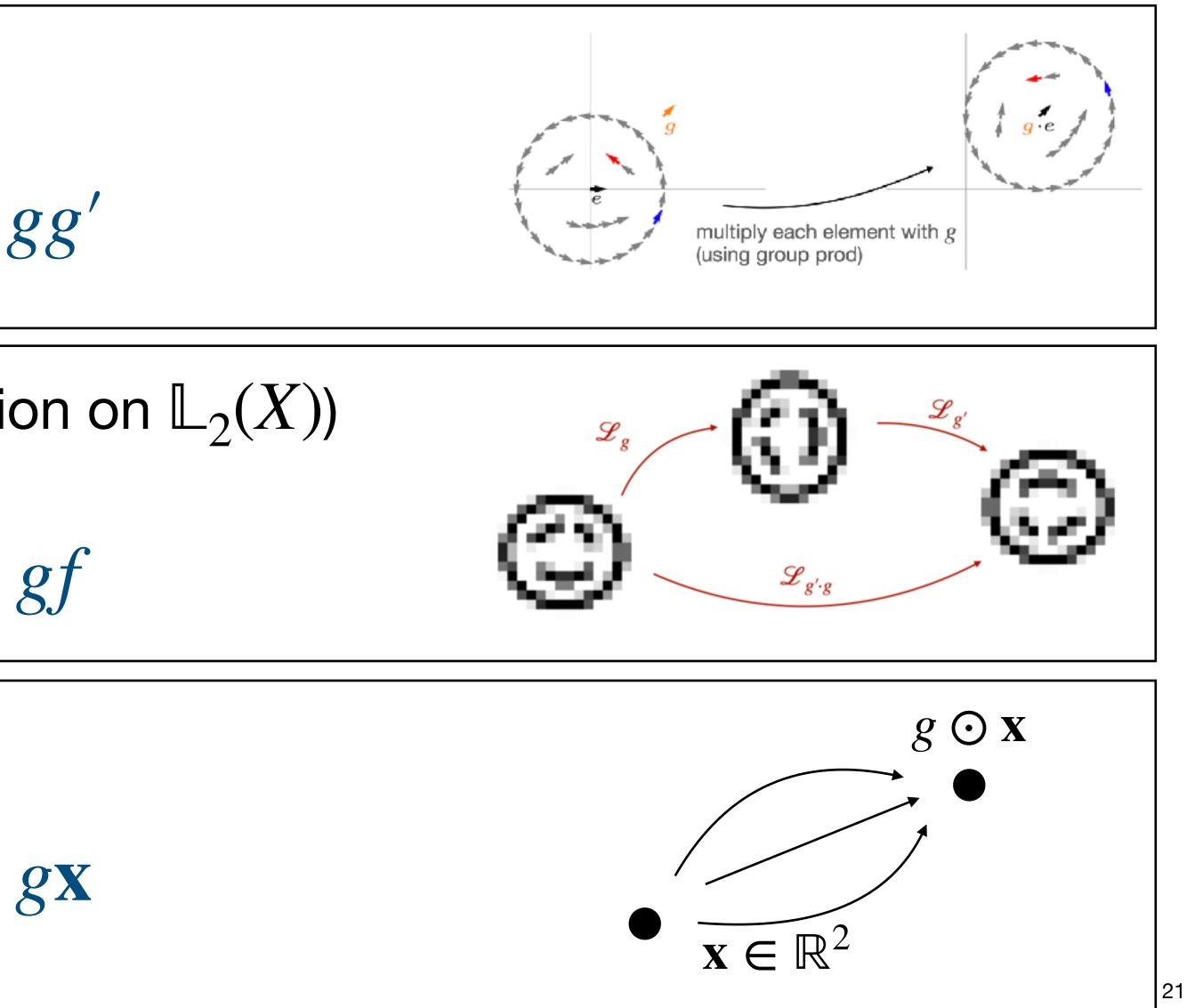
Group product (the action on G)

 $g \cdot g'$

Left regular representation (the action on $\mathbb{L}_2(X)$)

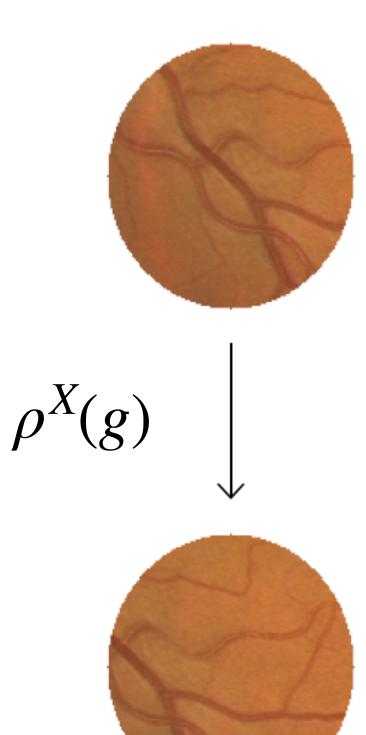
 $\mathscr{L}_{q}f$

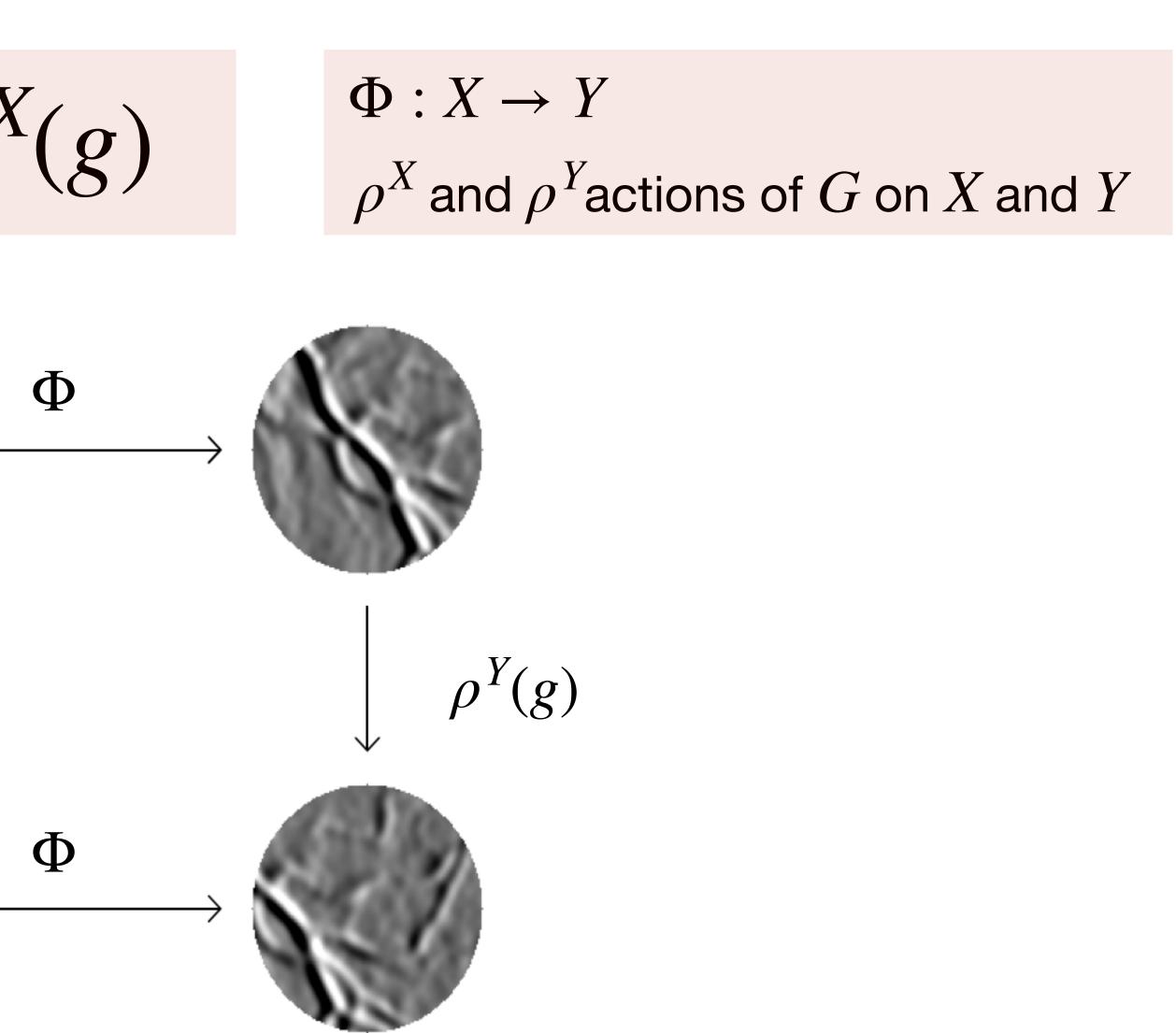
Group action (the action on \mathbb{R}^d)



Equivariance

$\rho^{Y}(g) \circ \Phi = \Phi \circ \rho^{X}(g)$





22

Content of this talk

1. Why do we want equivariant learning models? - Geometric guarantees + weight sharing/sample efficiency

- 2. A group theoretical view on **recognition by components** (capsule nets) - Group theoretical prerequisites (group product and representations)
 - Group convolutions perform pattern recognition by components

3. Experimental examples

4. Theorem: Linear maps between feature maps are equivariant iff they are group convolutions



Cross-correlations

 $(k \star_{\mathbb{R}^2} f)(\mathbf{x}) = \int_{\mathbb{R}^2} k(\mathbf{x}' - \mathbf{x}) f(\mathbf{x}') d\mathbf{x}'$

Are convolutions with reflected conv kernels (and vice versa)





Cross-correlations

 $(k \star_{\mathbb{R}^2} f)(\mathbf{x}) = \int_{\mathbb{R}^2} k(\mathbf{x}' - \mathbf{x}) f(\mathbf{x}') d\mathbf{x}' = (\mathscr{L}_g k, f)_{\mathbb{L}_2(\mathbb{R}^2)}$

Are convolutions with reflected conv kernels (and vice versa)

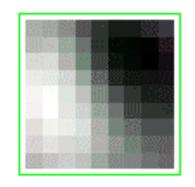
Representation of the translation group!



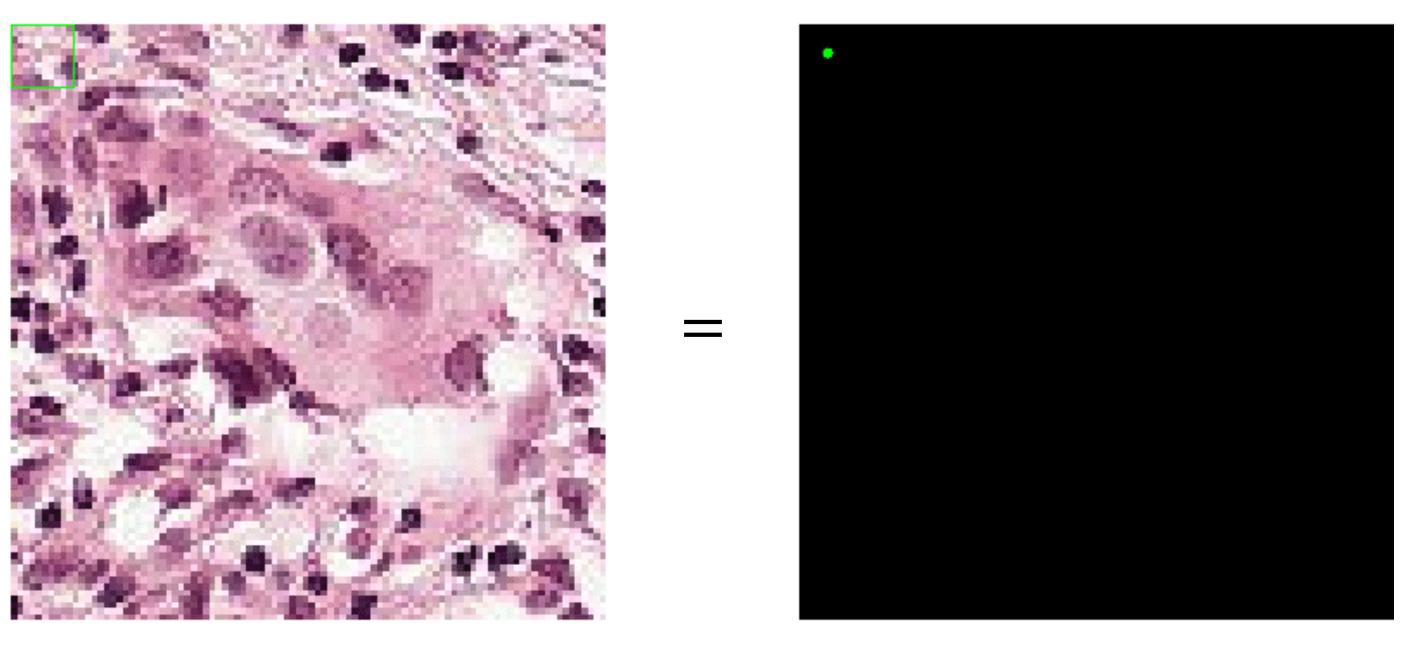


Cross-correlations

 $(k \star_{\mathbb{R}^2} f)(\mathbf{x}) = \int_{\mathbb{D}^2} k(\mathbf{x}' - \mathbf{x}) f(\mathbf{x}') d\mathbf{x}' = (\mathscr{L}_g k, f)_{\mathbb{L}_2(\mathbb{R}^2)}$



 $\star_{\mathbb{R}^2}$







University of Amsterdam

Are convolutions with reflected conv kernels (and vice versa)

Representation of the translation group!

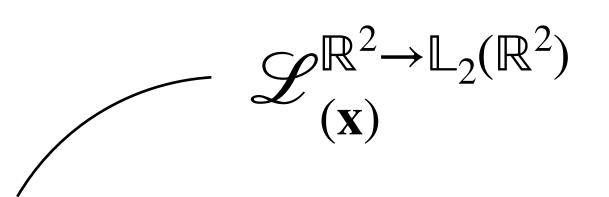
fout 2D feature map (after ReLU)





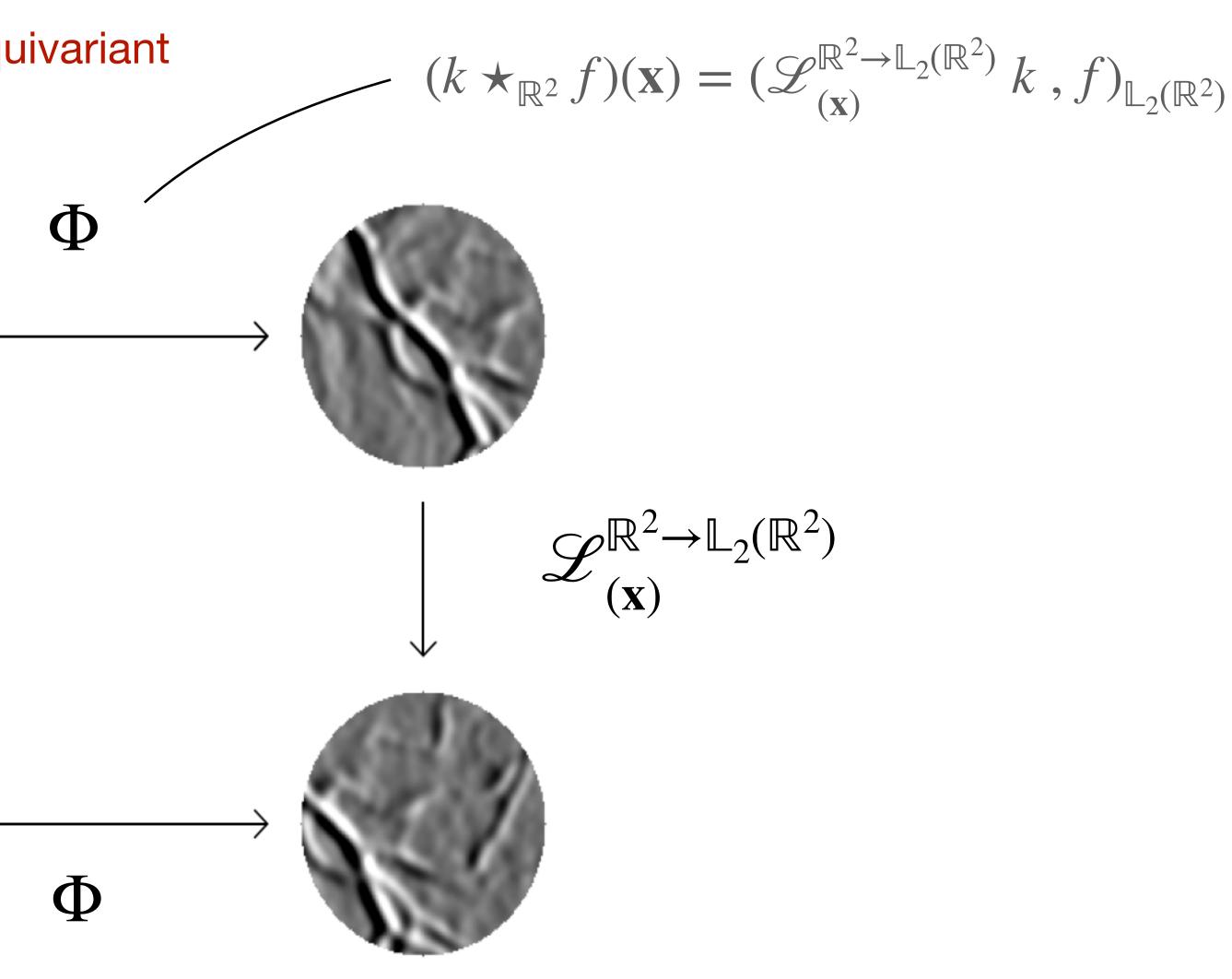
Convolutions/cross-correlations are translation equivariant





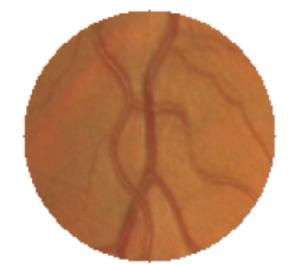
Representation of the translation group







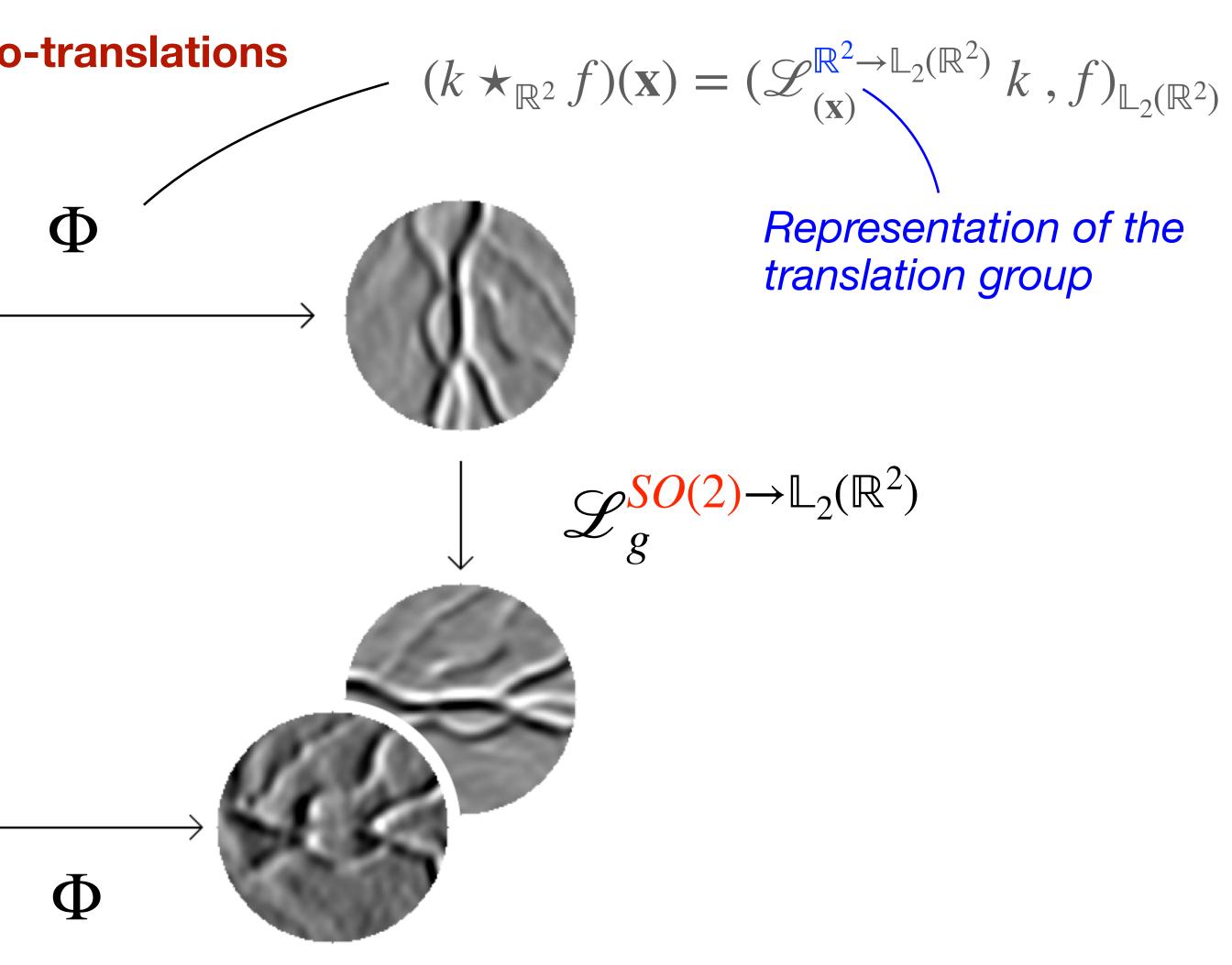
Convolutions are generally not equivariant to roto-translations



 $\mathscr{L}^{\mathcal{S}U(2)}_{\theta} \to \mathbb{L}_2(\mathbb{R}^2)$

Representation of the rotation group







Representation of the roto-translation group!

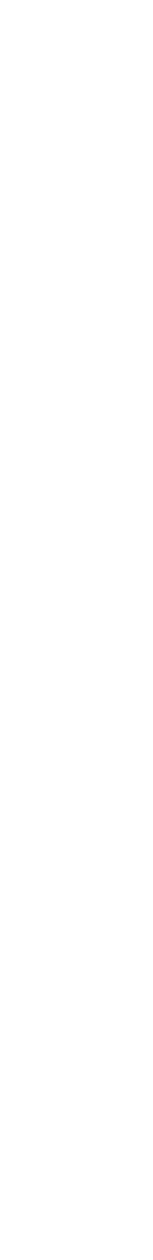
Lifting correlations: $(k \stackrel{\sim}{\star} f)(\mathbf{x}) = (\mathscr{L}_{g}^{SE(2) \to \mathbb{L}_{2}(\mathbb{R}^{2})} k, f)_{\mathbb{L}_{2}(\mathbb{R}^{2})}$



Representation of the roto-translation group!

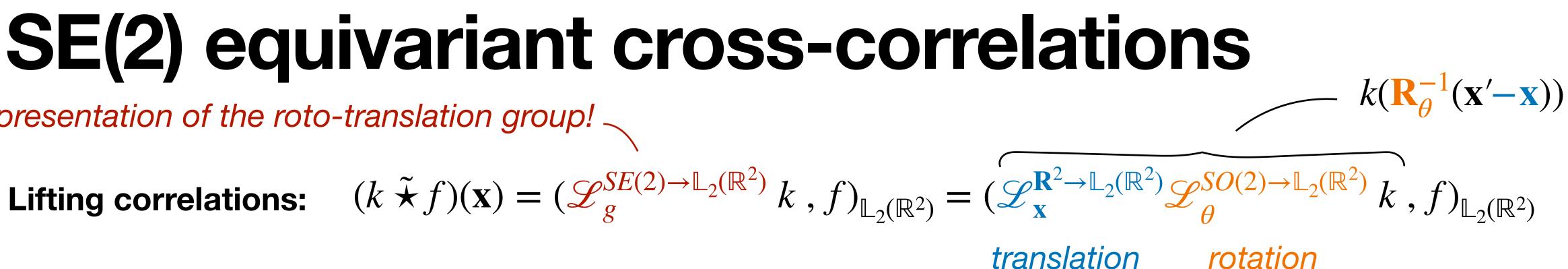
Lifting correlations: $(k \stackrel{\sim}{\star} f)(\mathbf{x}) = (\mathscr{L}_{g}^{SE(2) \to \mathbb{L}_{2}(\mathbb{R}^{2})} k, f)_{\mathbb{L}_{2}(\mathbb{R}^{2})} = (\mathscr{L}_{\mathbf{x}}^{\mathbb{R}^{2} \to \mathbb{L}_{2}(\mathbb{R}^{2})} \mathscr{L}_{\theta}^{SO(2) \to \mathbb{L}_{2}(\mathbb{R}^{2})} k, f)_{\mathbb{L}_{2}(\mathbb{R}^{2})}$

translation rotation



27

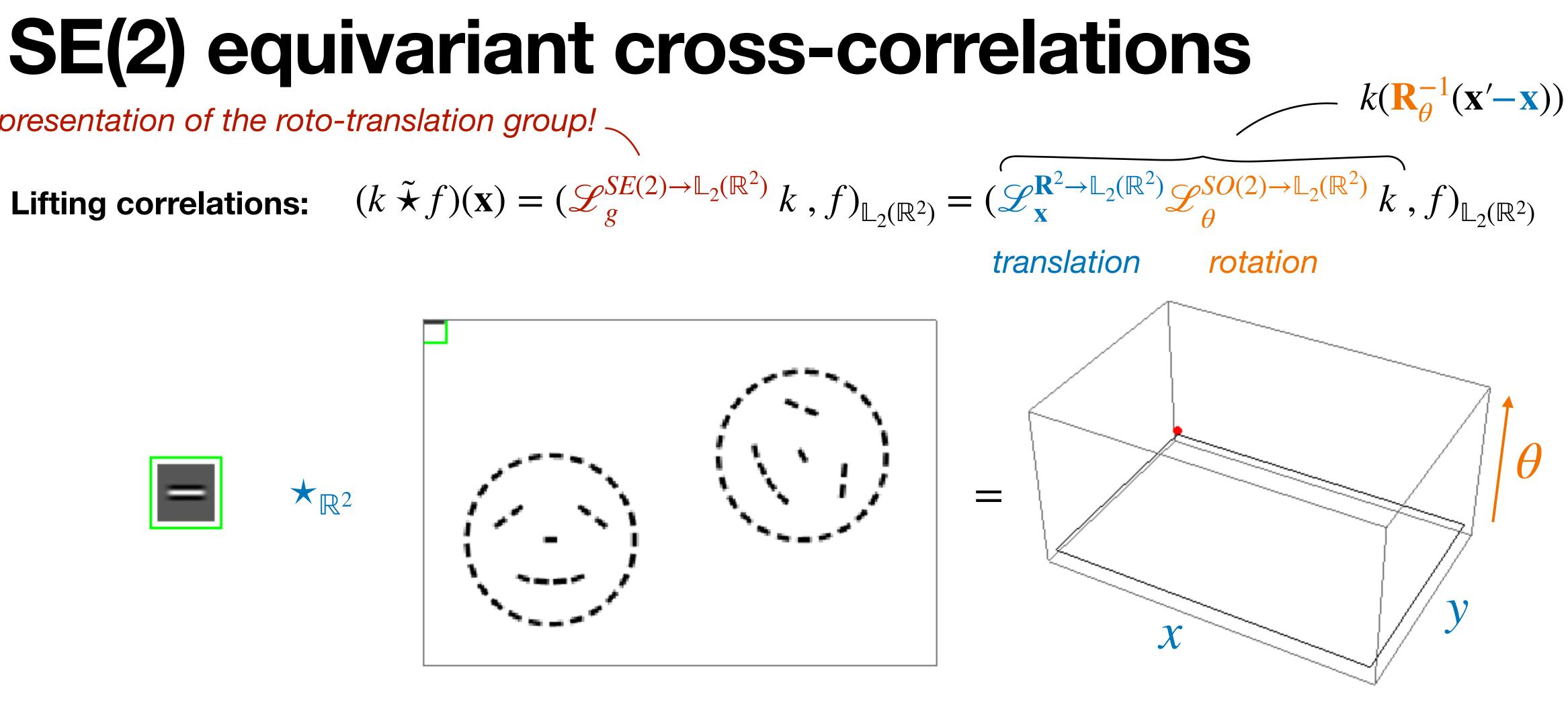
Representation of the roto-translation group!





Representation of the roto-translation group! ____

Lifting correlations:



 $\mathscr{L}^{SO(2) \to \mathbb{L}_2(\mathbb{R}^2)} k$ **Rotated** 2D convolution kernel

fln 2D feature map

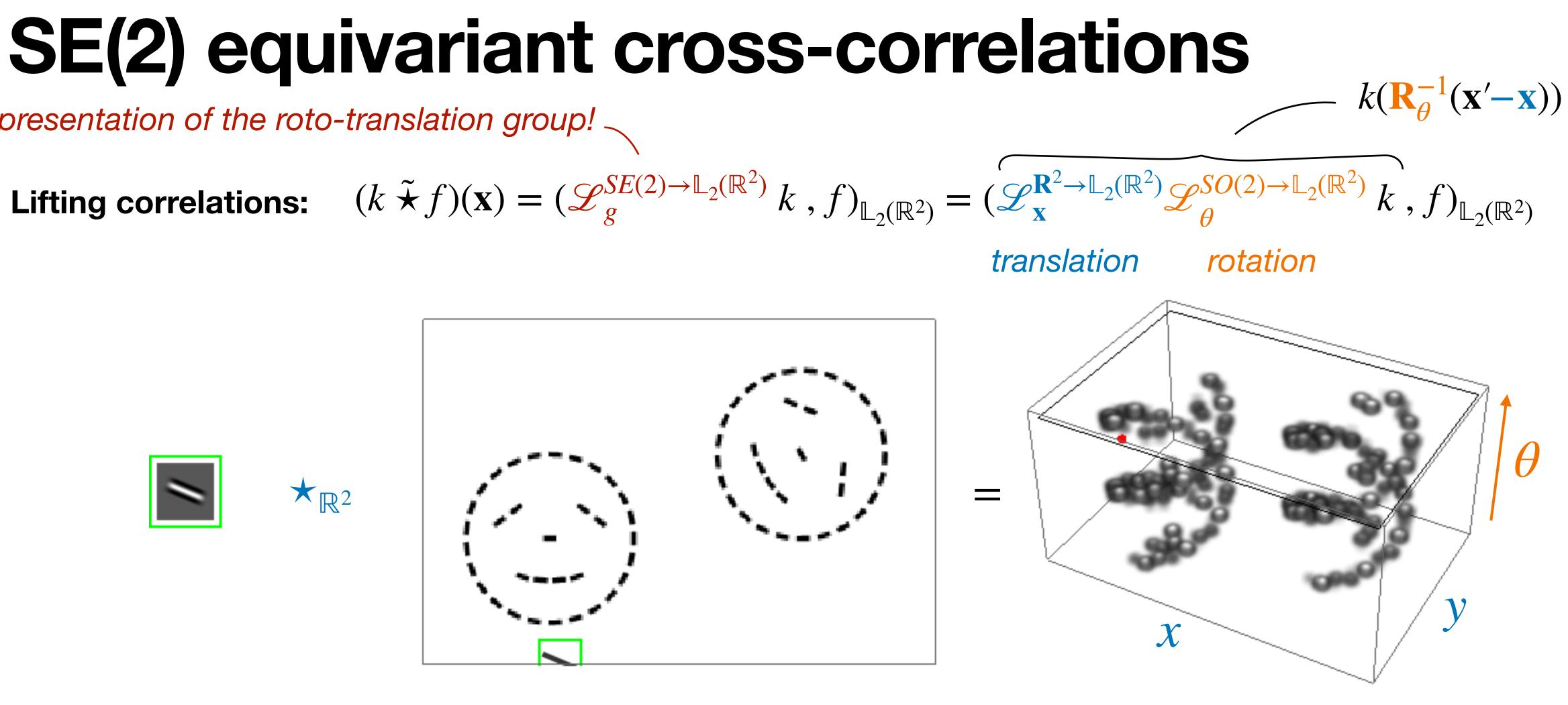
UNIVERSITY OF AMSTERDAM

fOUl 3D (SE(2)) feature map (after ReLU)



Representation of the roto-translation group! _

Lifting correlations:



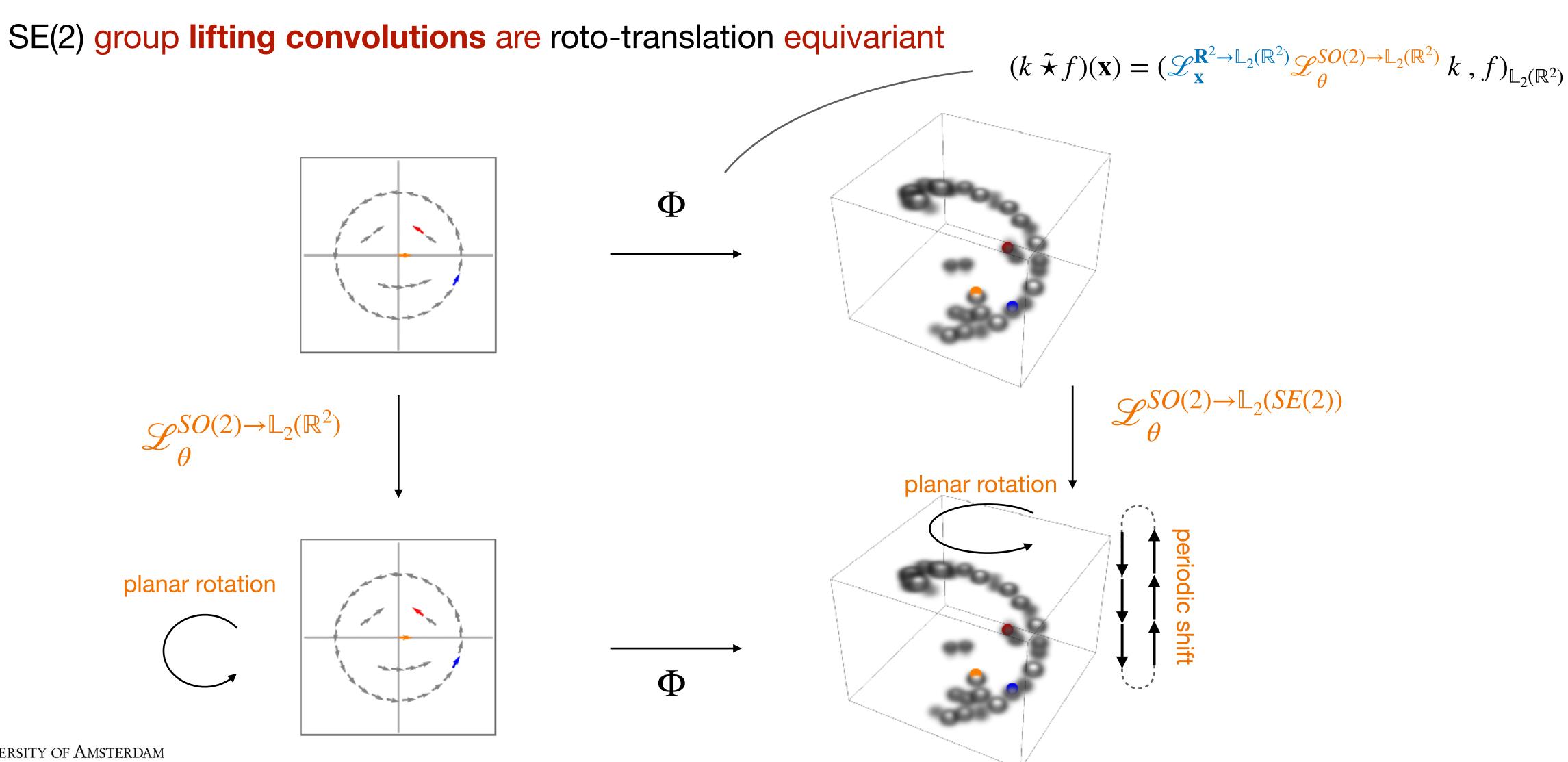
 $\mathscr{L}^{SO(2) \to \mathbb{L}_2(\mathbb{R}^2)} k$ Rotated 2D convolution kernel

fln 2D feature map

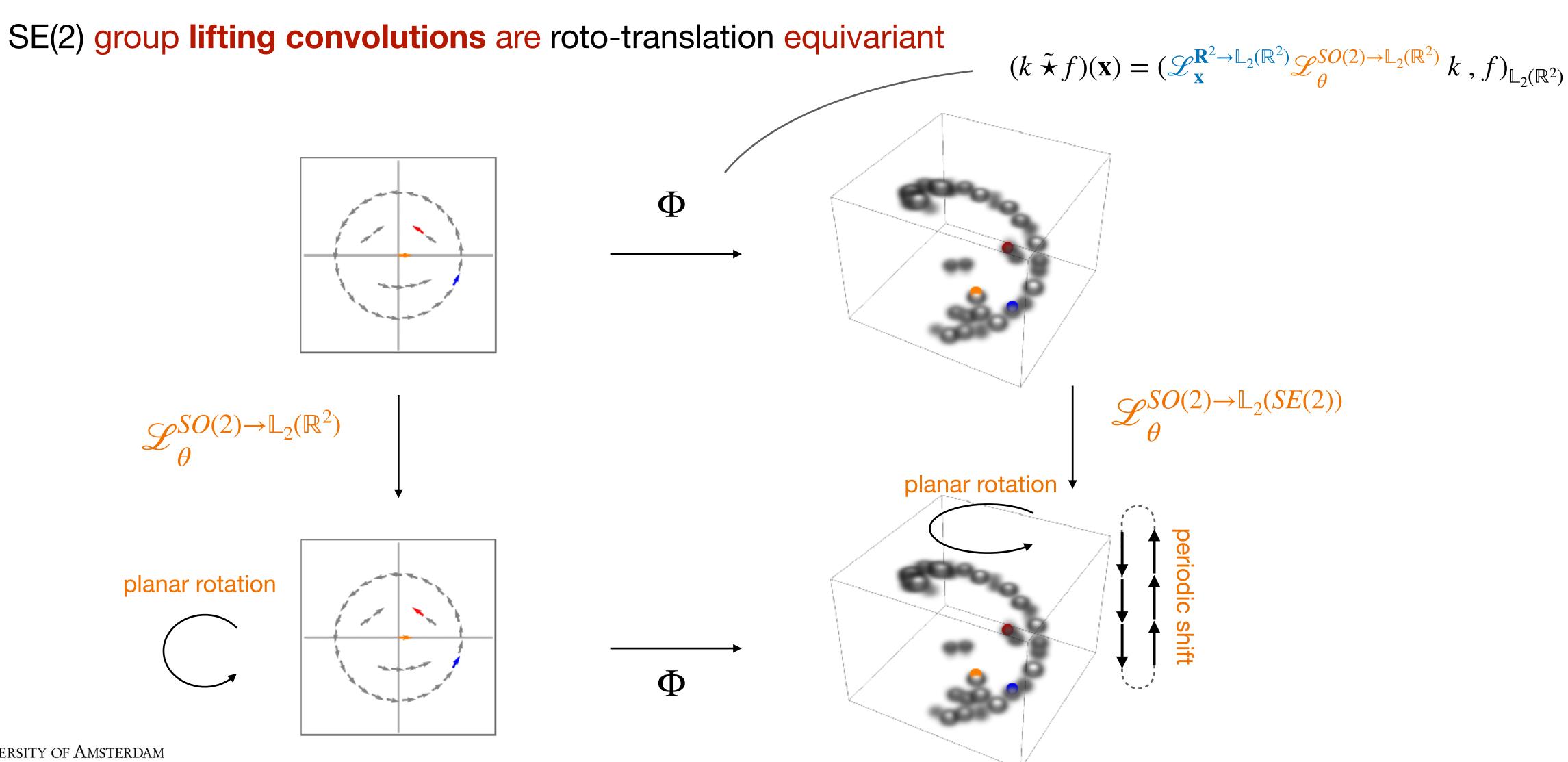
UNIVERSITY OF AMSTERDAM

fOUl 3D (SE(2)) feature map (after ReLU)

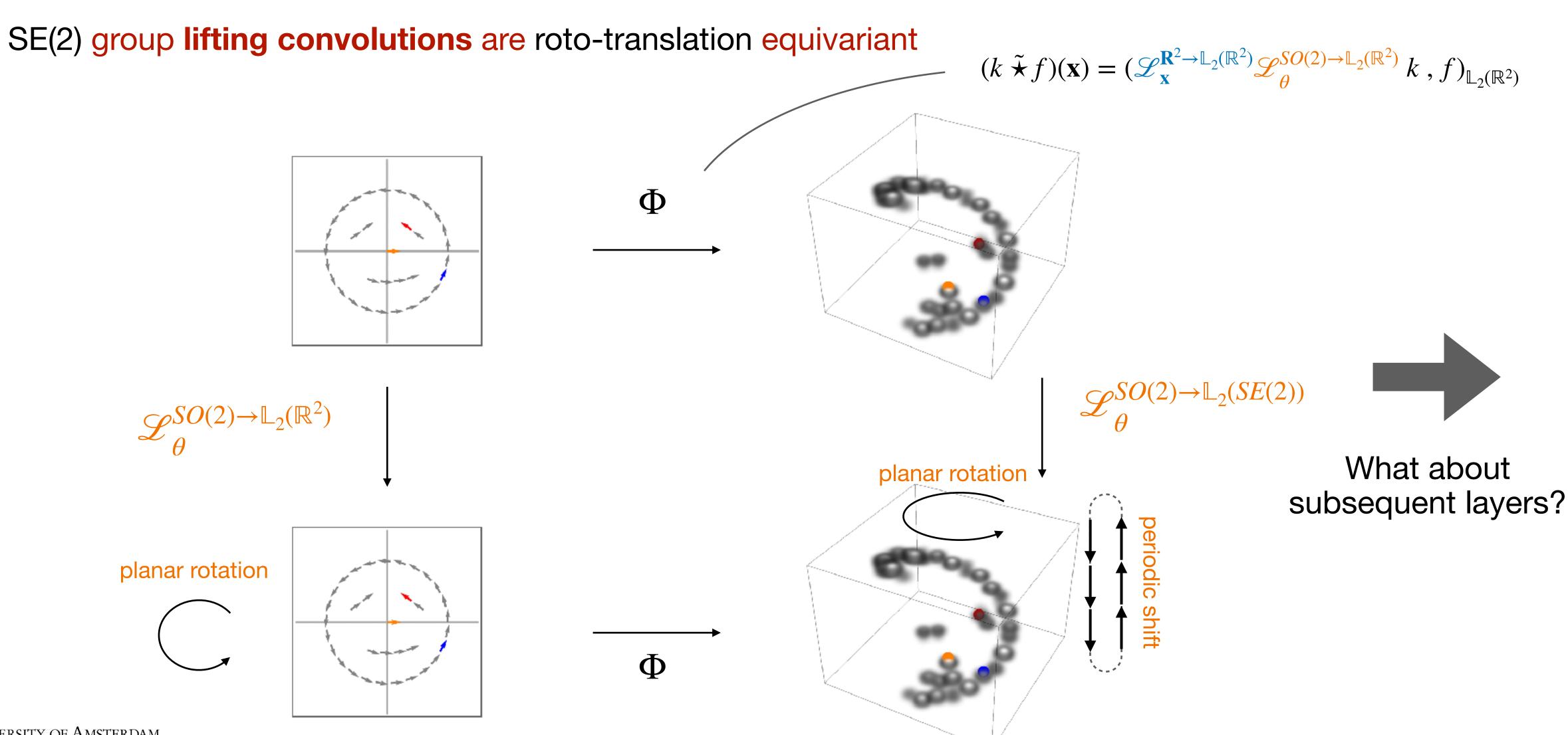








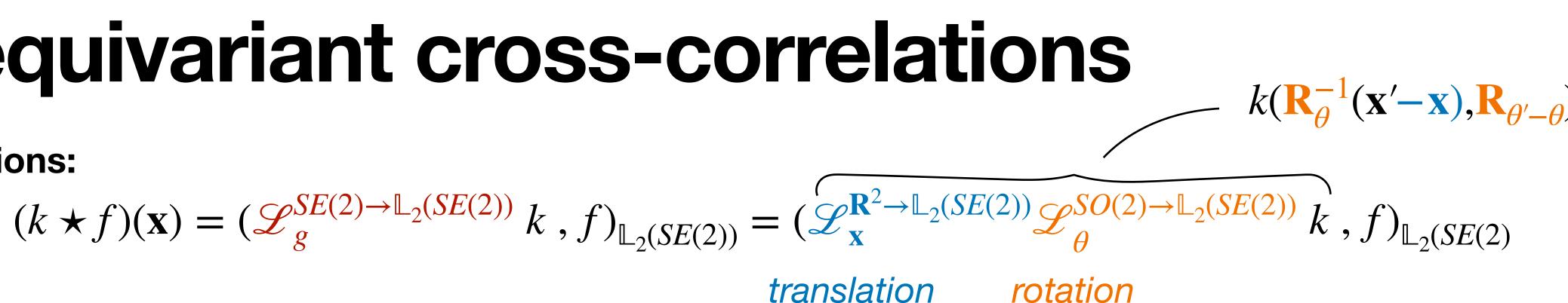






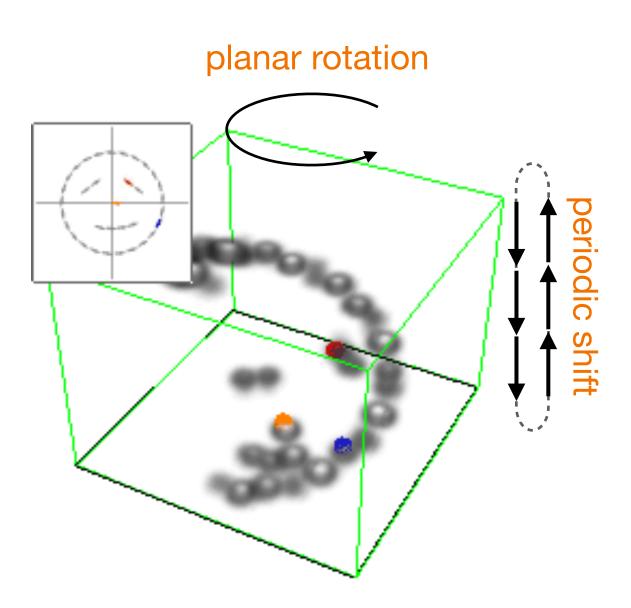


Group correlations:

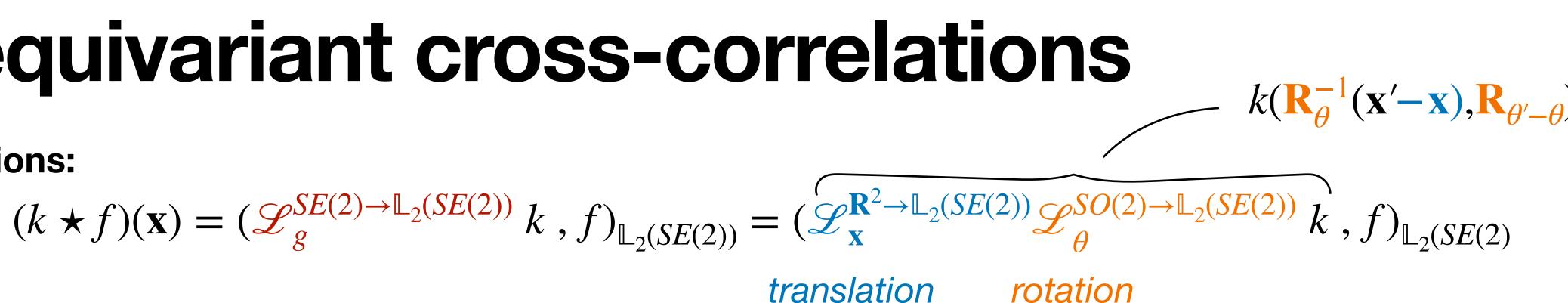




Group correlations:

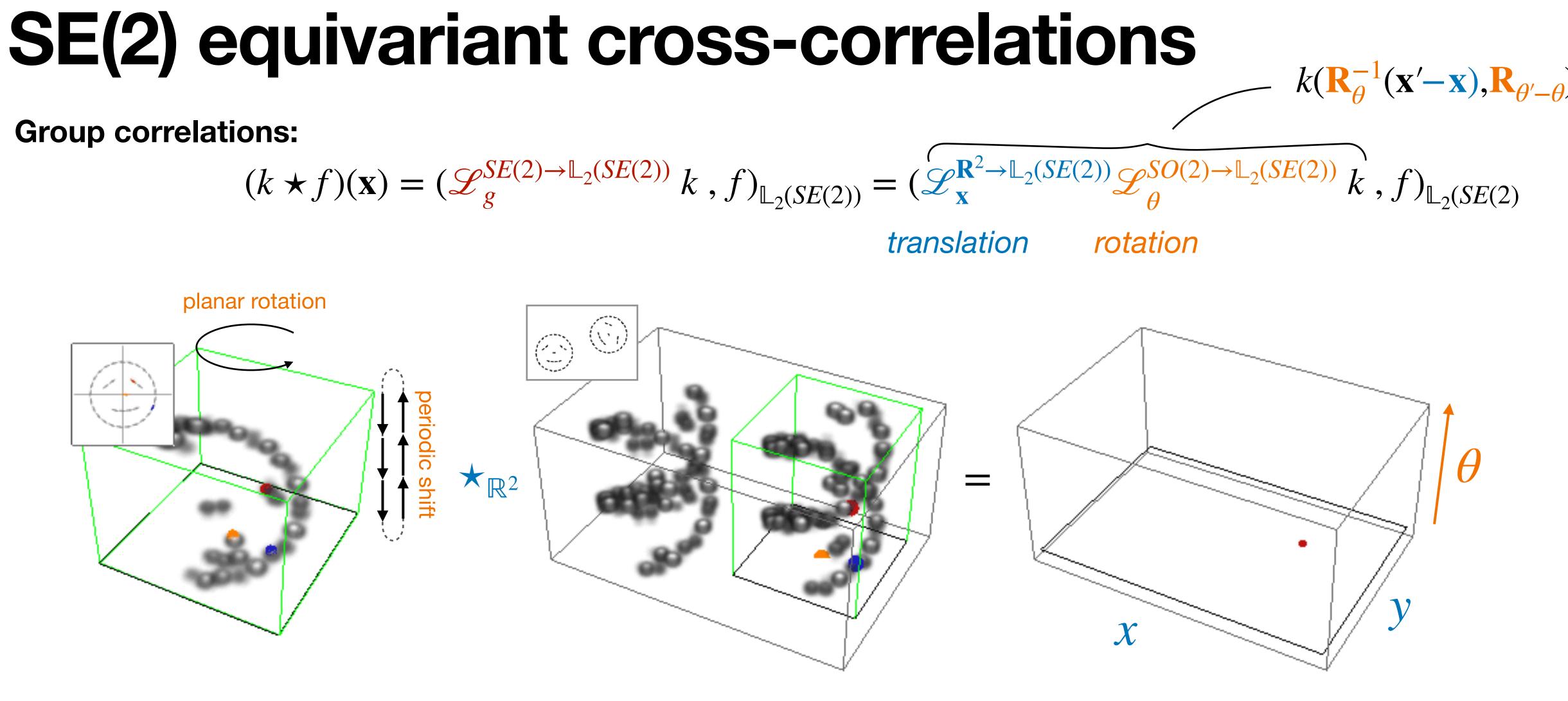


$\mathscr{L}_{\theta}^{SO(2) \to \mathbb{L}_{2}(SE(2))}k$ Rotated SE(2) convolution kernel





Group correlations:



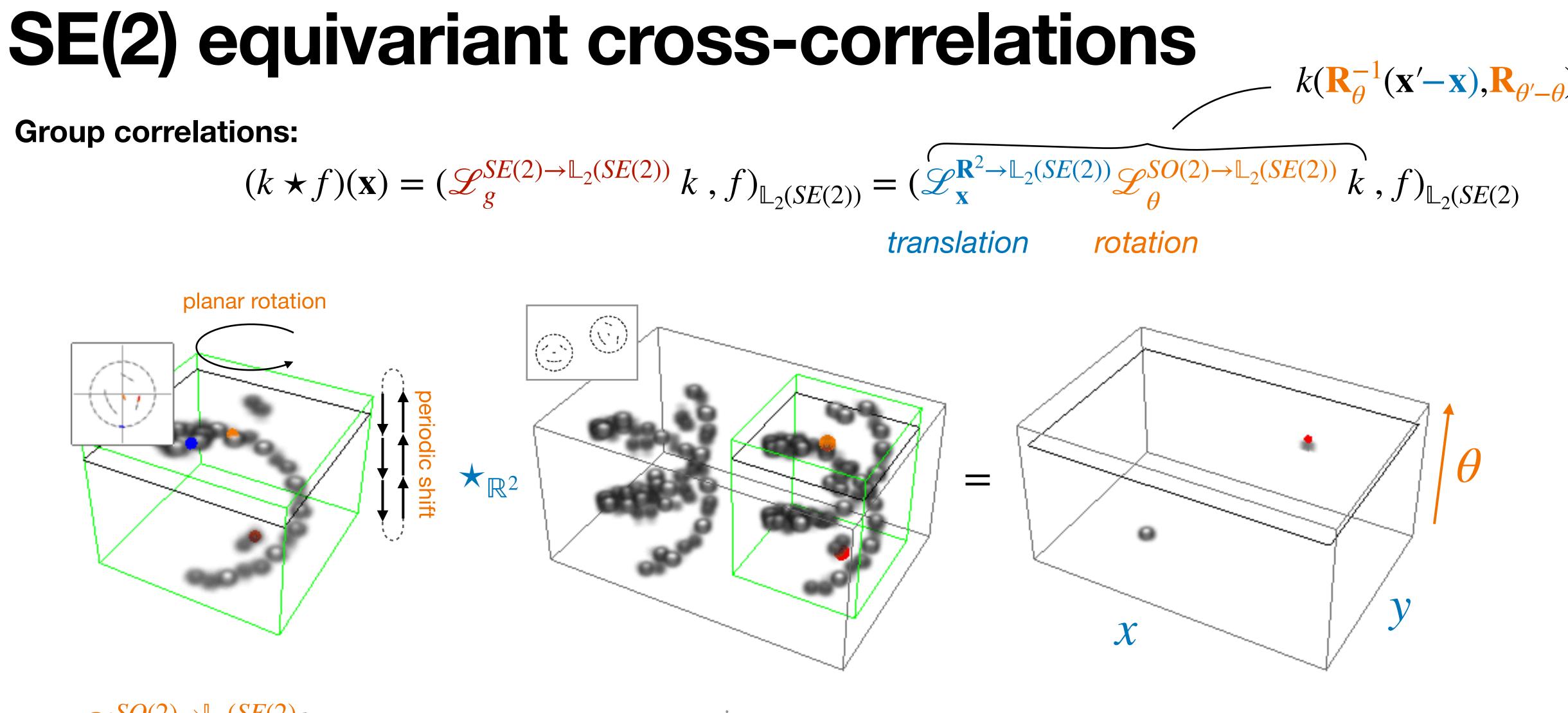
$\mathscr{L}^{SO(2) \to \mathbb{L}_2(SE(2))}_{A}k$ Rotated SE(2) convolution kernel

fin SE(2) feature map

fout SE(2) feature map (after ReLU)



Group correlations:

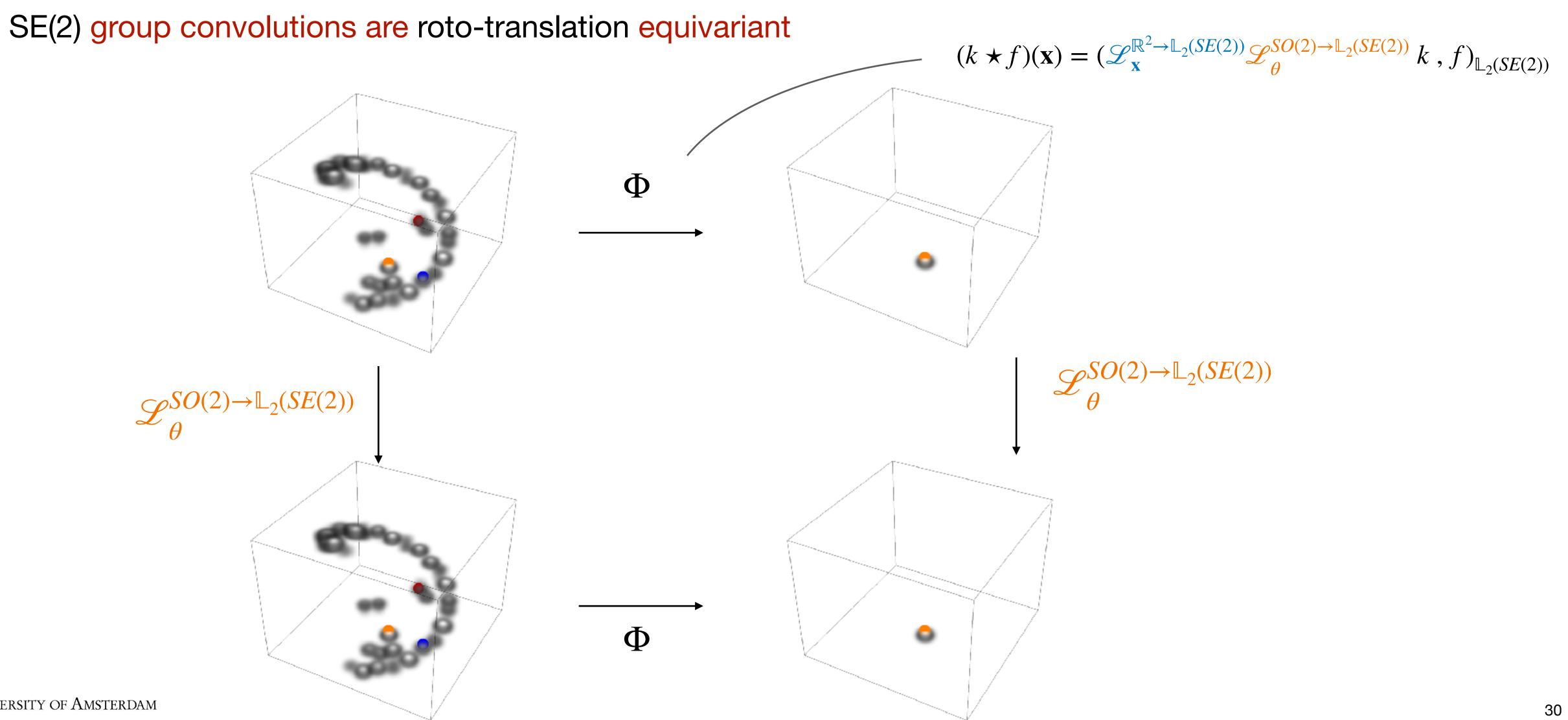


$\mathscr{L}^{SO(2) \to \mathbb{L}_2(SE(2))}_{A}k$ Rotated SE(2) convolution kernel

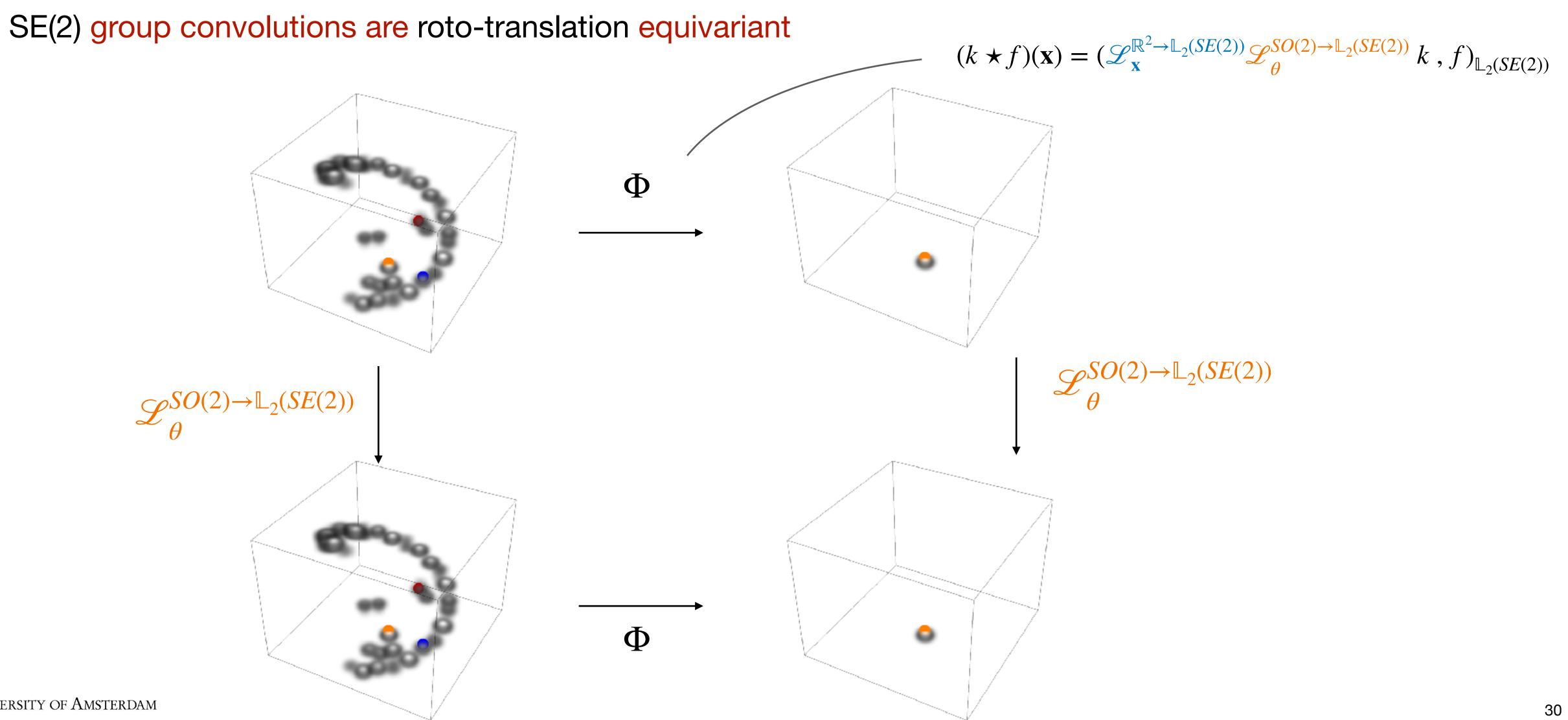
fin SE(2) feature map

fout SE(2) feature map (after ReLU)



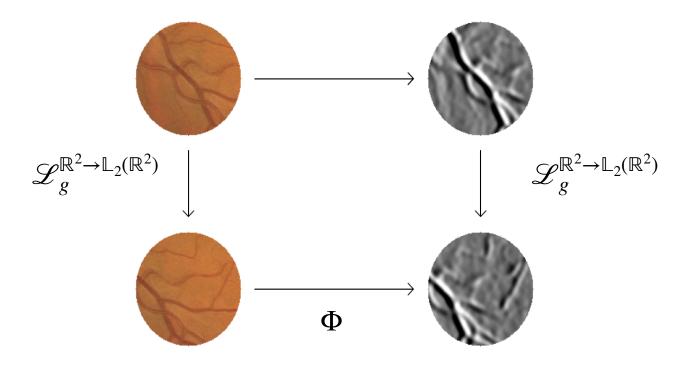


University of Amsterdam

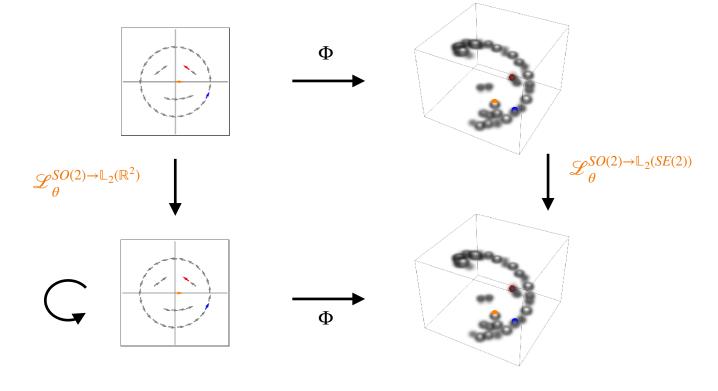


University of Amsterdam

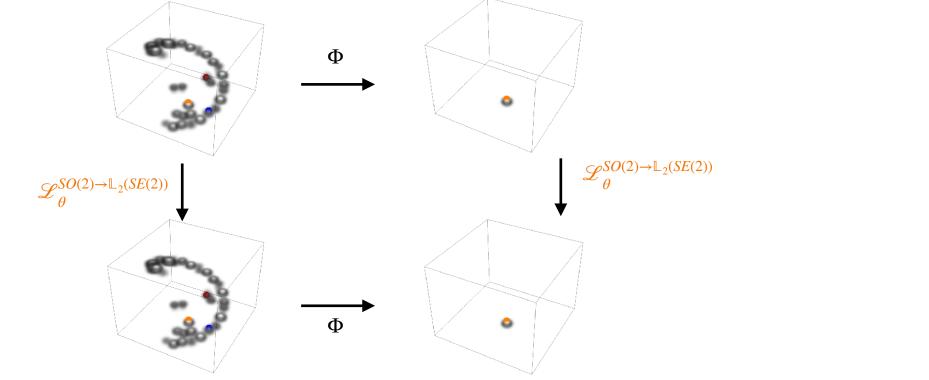
2D cross-correlation (translation equivariant)



SE(2) lifting correlations (roto-translation equivariant)



SE(2) G-correlations (roto-translation equivariant)



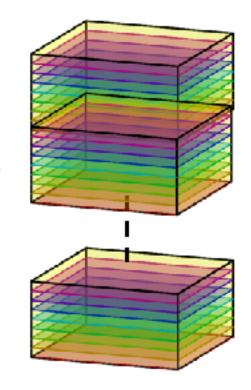
$$(k \star_{\mathbb{R}^2} f)(\mathbf{x}) = (\mathscr{L}_{\mathbf{x}}^{\mathbb{R}^2 \to \mathbb{L}_2(\mathbb{R}^2)} k, f)_{\mathbb{L}_2(\mathbb{R}^2)}$$
$$= \int_{\mathbb{R}^2} k(\mathbf{x}' - \mathbf{x}) f(\mathbf{x}') d\mathbf{x}' d\mathbf{x}'$$

$$k \stackrel{\sim}{\star} f)(\mathbf{x}) = (\mathscr{L}_{g}^{SE(2) \to \mathbb{L}_{2}(\mathbb{R}^{2})} k, f)_{\mathbb{L}_{2}(\mathbb{R}^{2})}$$
$$= \int_{\mathbb{R}^{2}} k(\mathbf{R}_{\theta}^{-1}(\mathbf{x}' - \mathbf{x})) f$$

$$k \stackrel{\sim}{\star} f)(\mathbf{x}) = (\mathscr{L}_{g}^{SE(2) \to \mathbb{L}_{2}(SE(2))} k, f)_{\mathbb{L}_{2}(SE(2))}$$
$$= \int_{\mathbb{R}^{2}} \int_{S^{1}} k(\mathbf{R}_{\theta}^{-1}(\mathbf{x}' - \mathbf{x}), \theta' - \theta \mod 2\pi) f(\mathbf{x}')$$

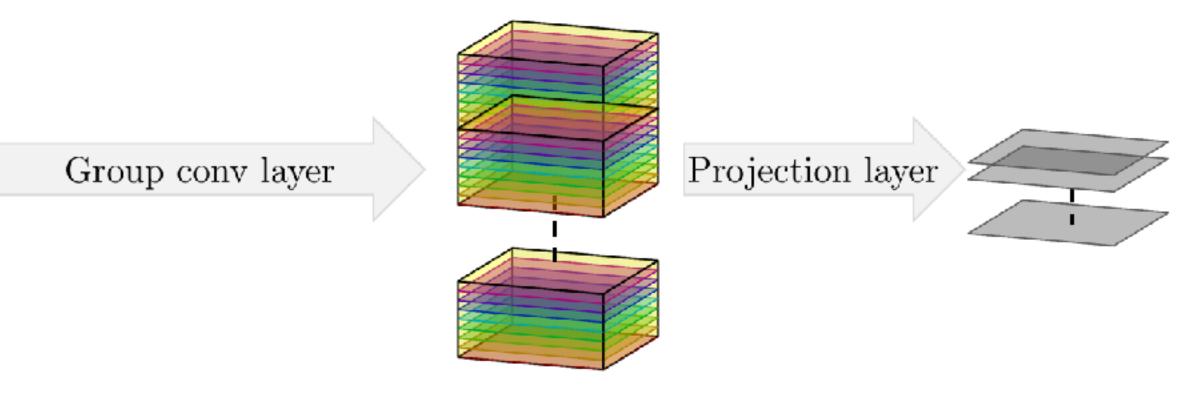


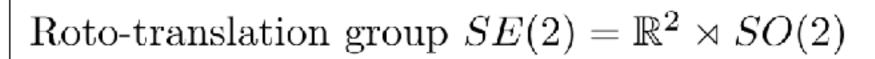
31



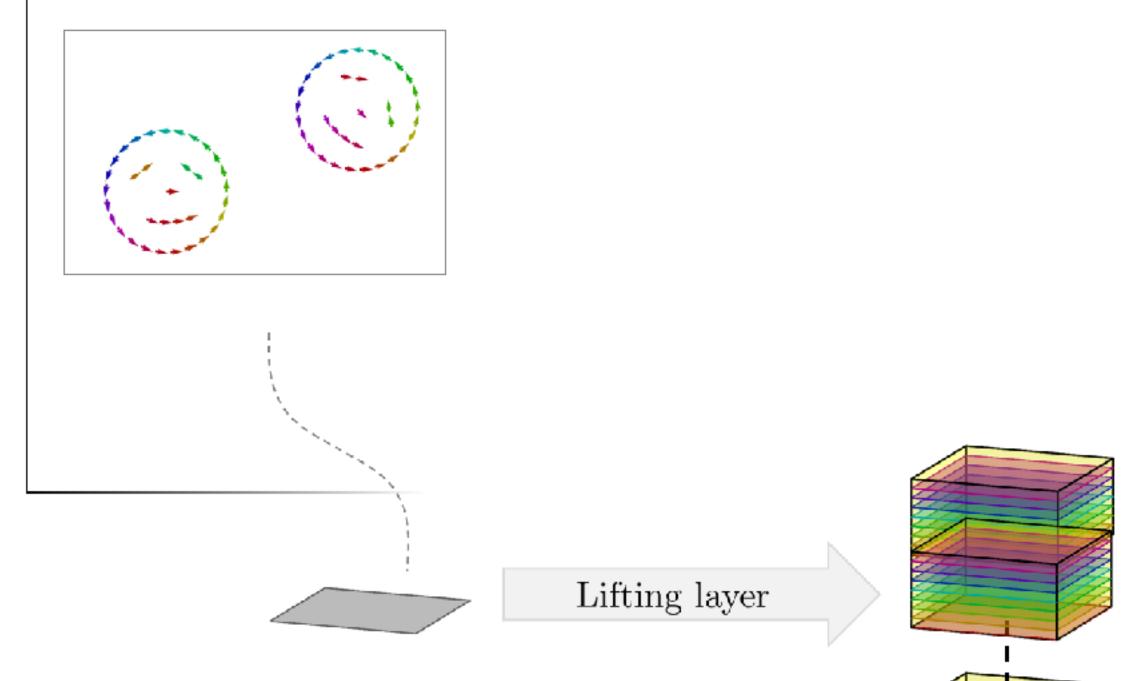


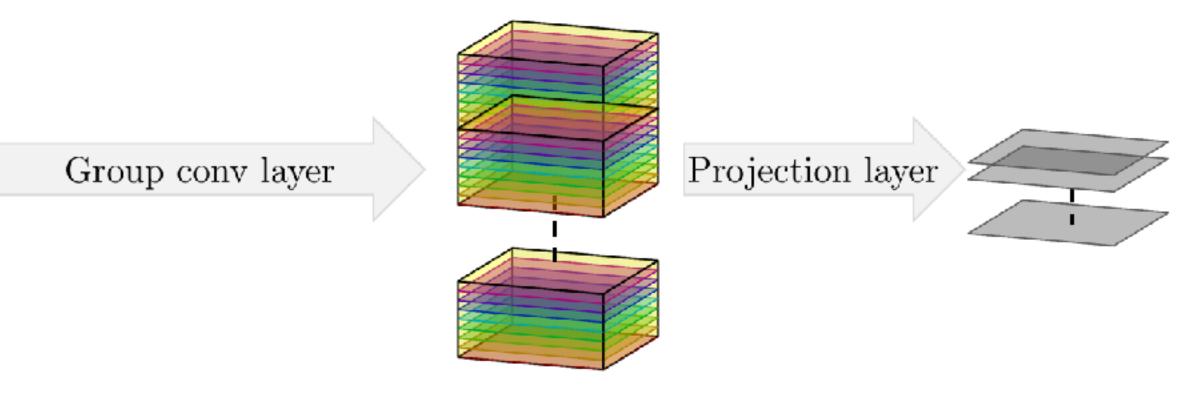
Lifting layer



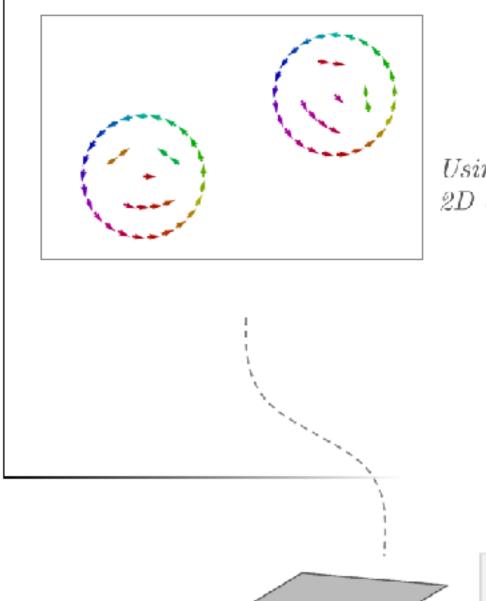


 $2\mathbf{D}$ feature map





2D feature map



Using a set of transformed by $2D\ conv\ kernels$

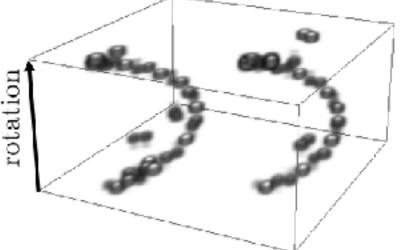
$$\theta = \frac{\pi}{2}$$

$$\theta = \frac{\pi}{4}$$

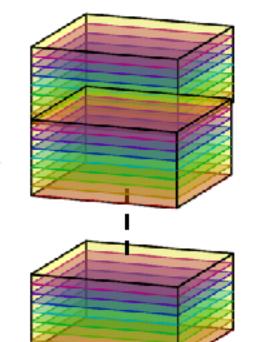
$$\theta = 0$$

Lifting layer

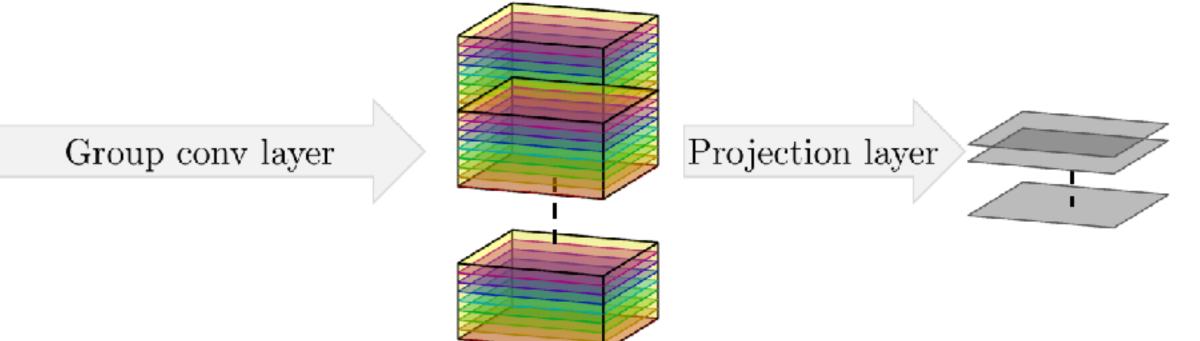
G feature map (activation for oriented structures at each position and rotation)



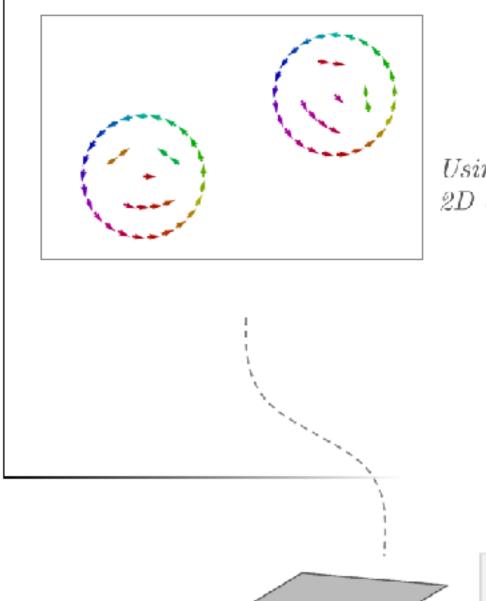
G-feature maps are equivariant w.r.t. translation and rotation of the input







$2\mathbf{D}$ feature map



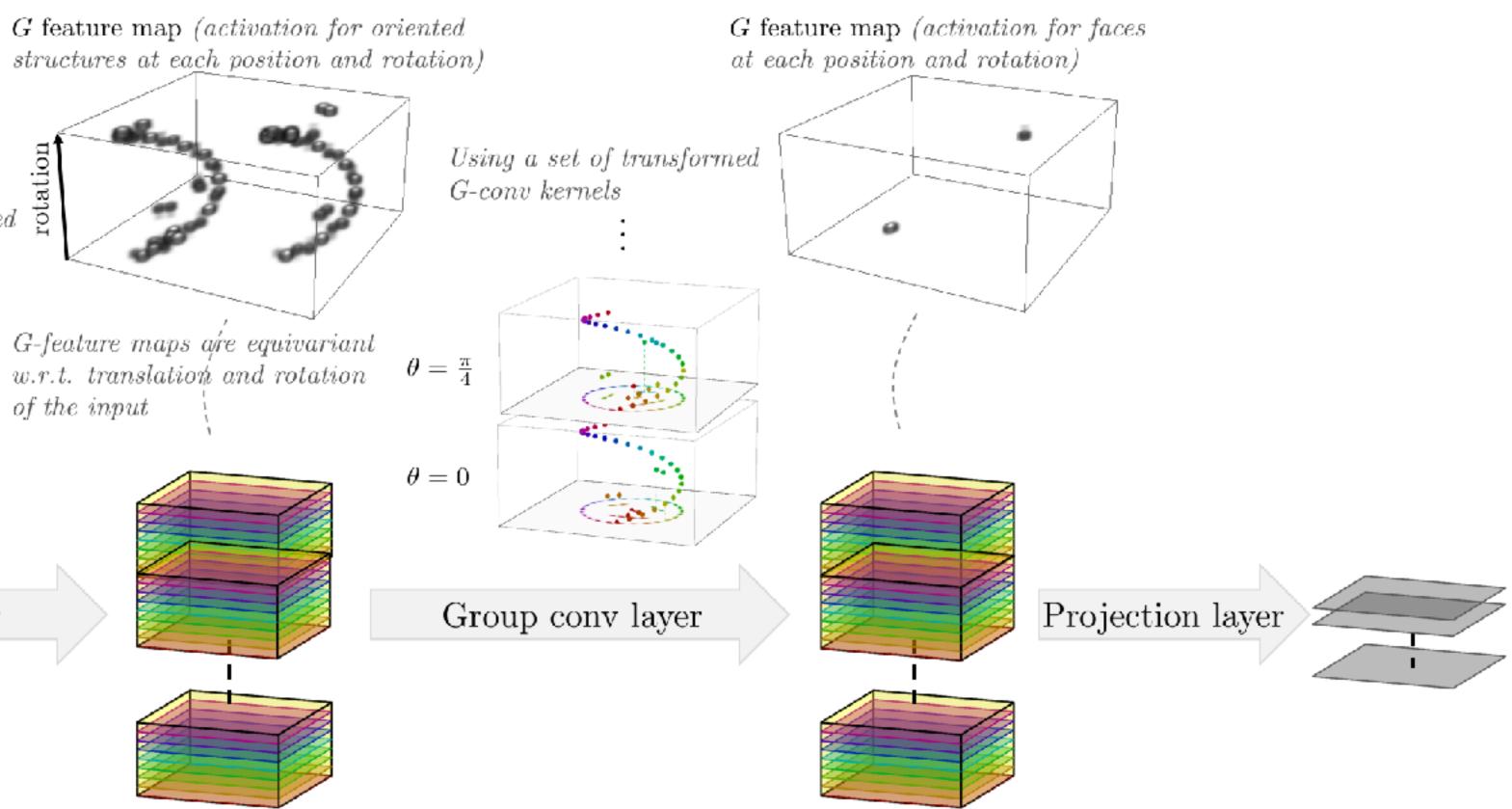
Using a set of transformed $2D\ conv\ kernels$

$$\theta = \frac{\pi}{2}$$

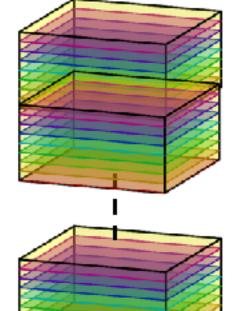
$$\theta = \frac{\pi}{4}$$

$$\theta = 0$$

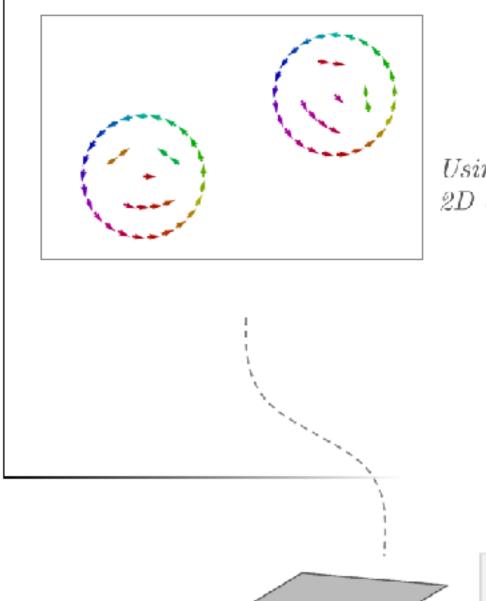
Lifting layer



G-feature maps are equivariant w.r.t. translation and rotation of the input



2D feature map



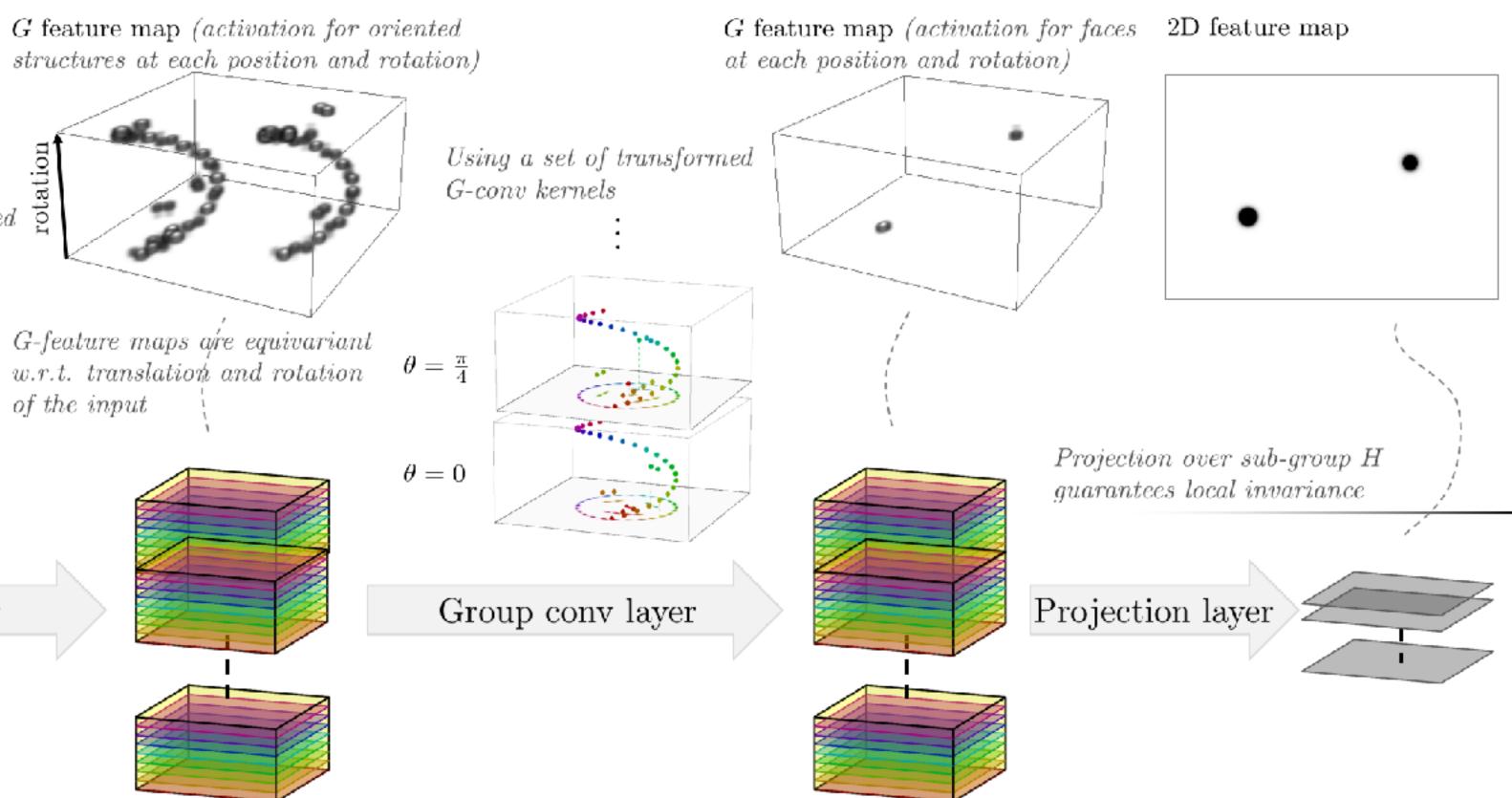
Using a set of transformed 2D conv kernels

$$\theta = \frac{\pi}{2}$$

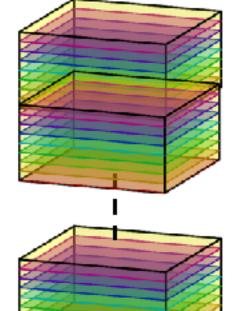
$$\theta = \frac{\pi}{4}$$

$$\theta = 0$$

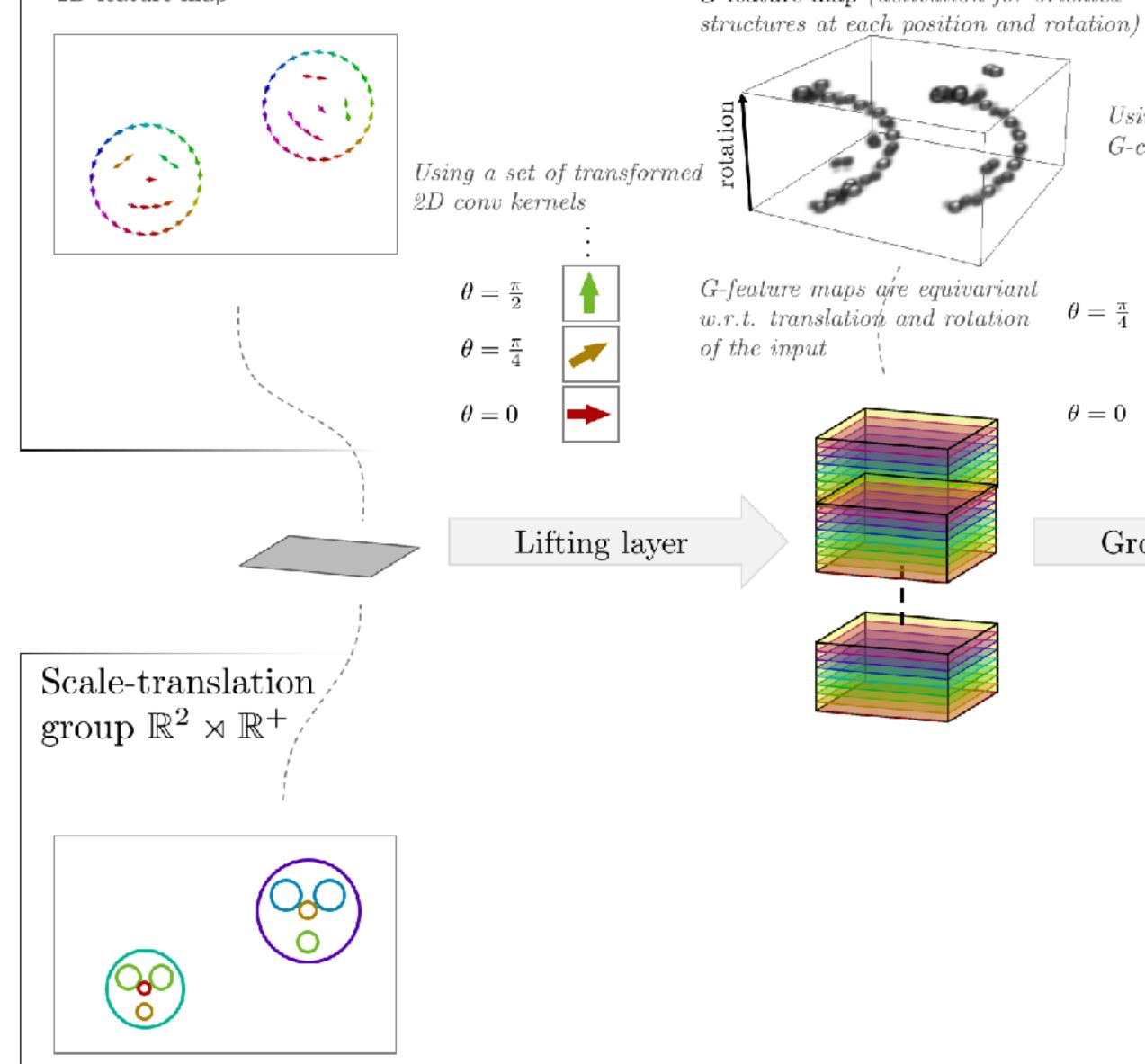
Lifting layer

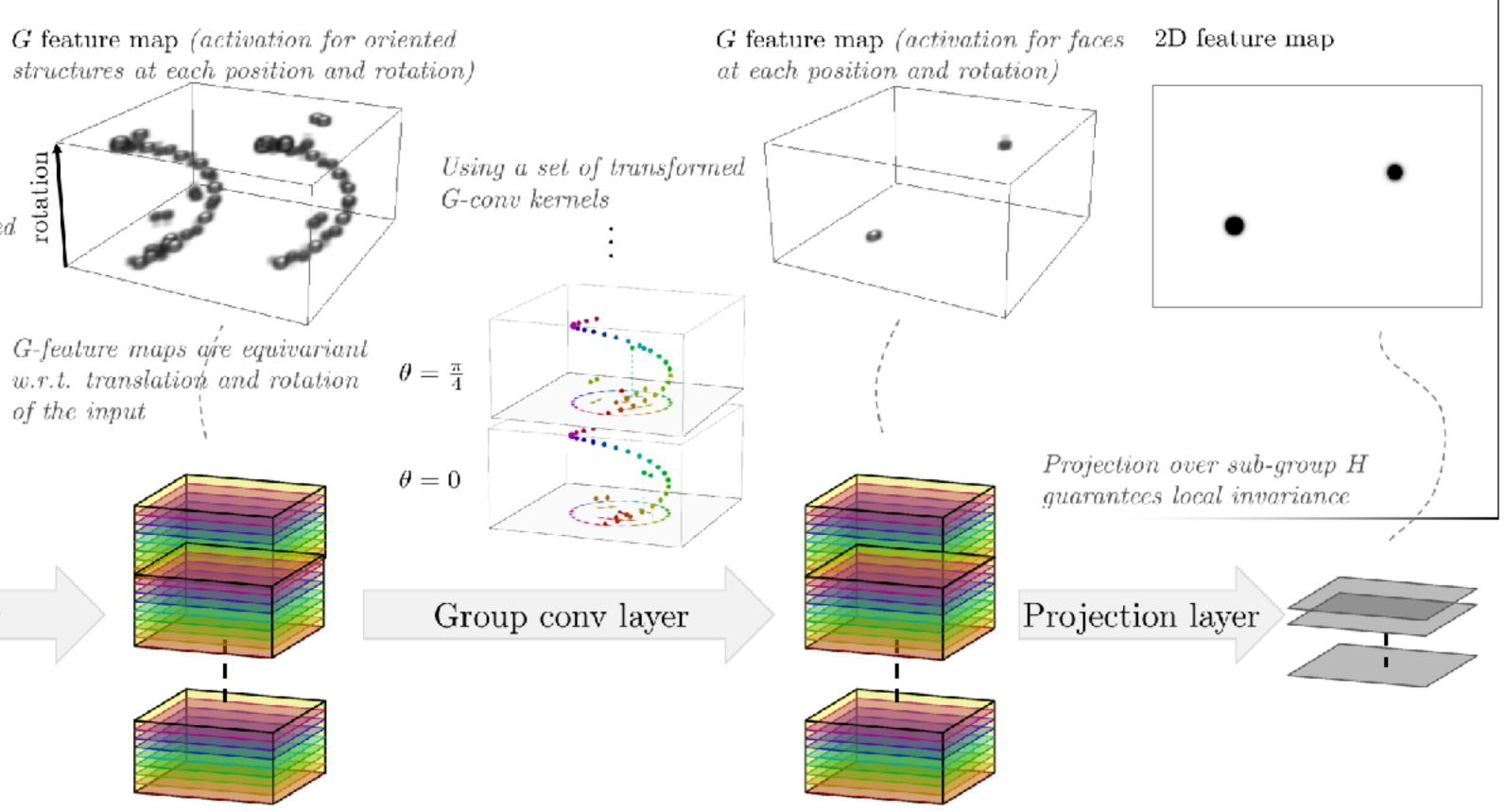


G-feature maps are equivariant w.r.t. translation and rotation of the input

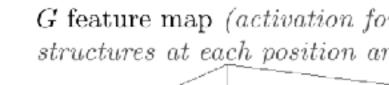


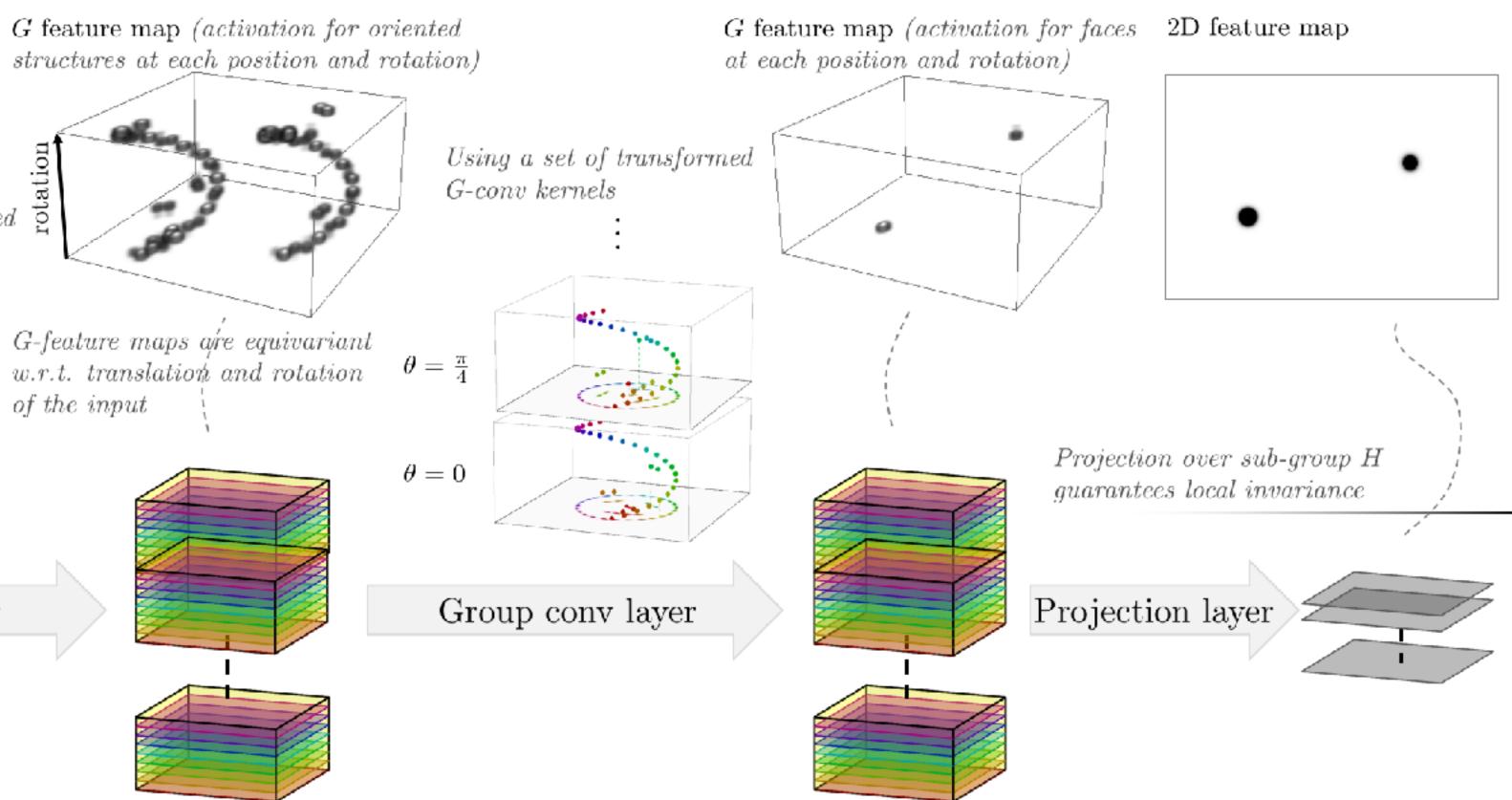
2D feature map





2D feature map

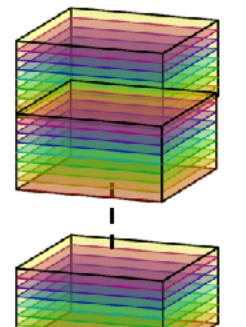




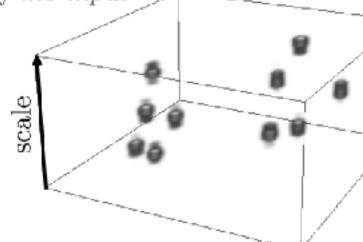
G-feature maps are equivariant w.r.t. translation and rotation of the input

Lifting layer

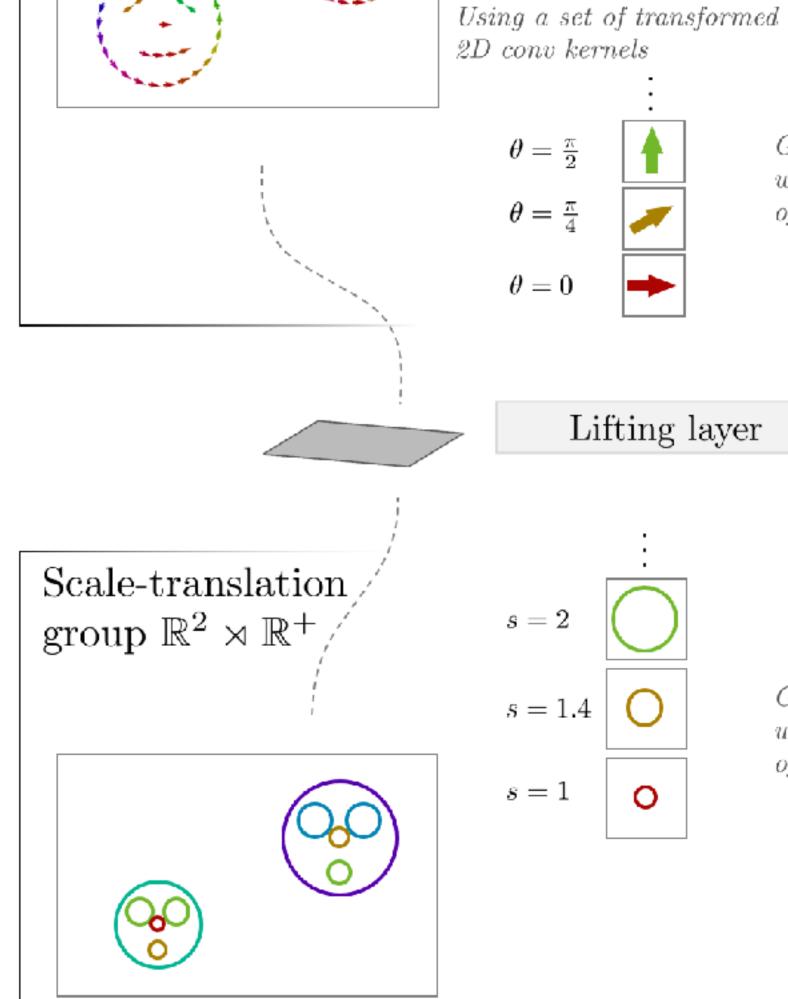
0



G-feature maps are equivariant w.r.t. translation and scaling of the input



Activation for circles at each position and scale

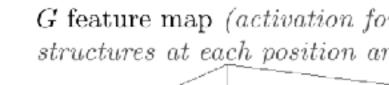


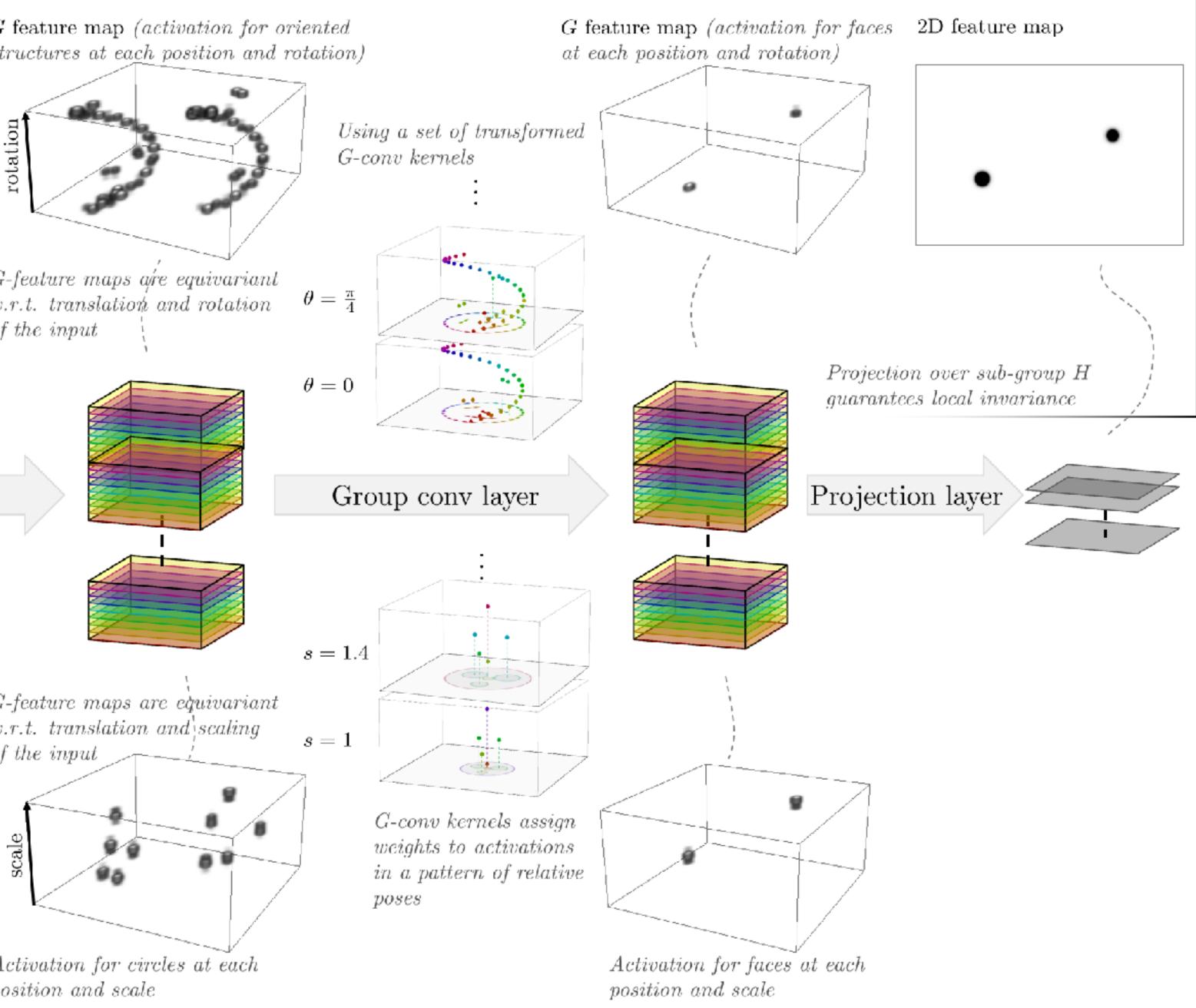
Using a set of transformed

Lifting layer

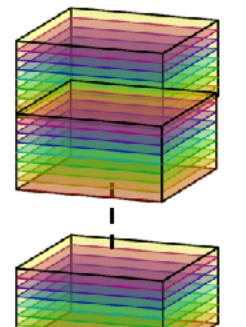
0

2D feature map

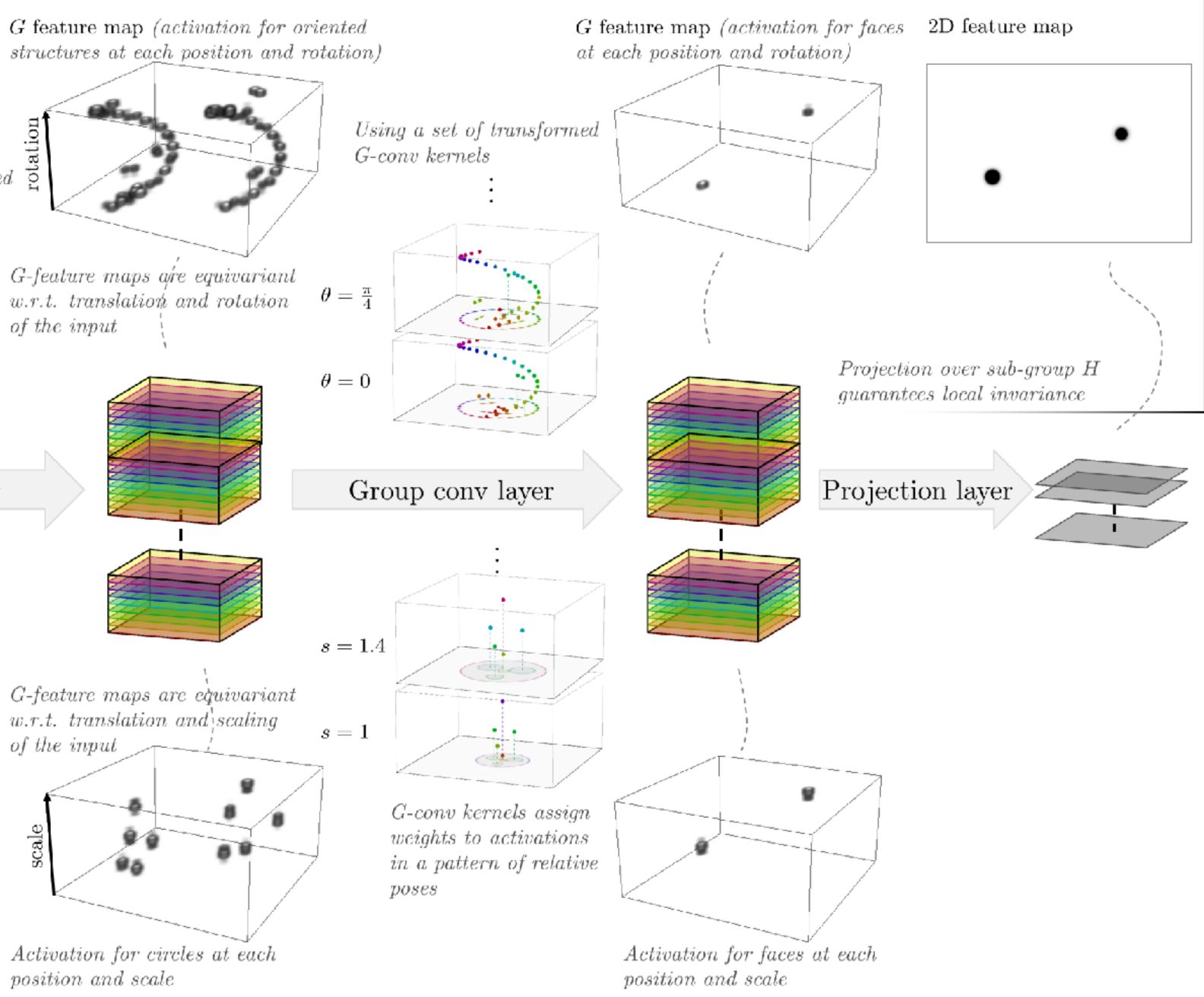




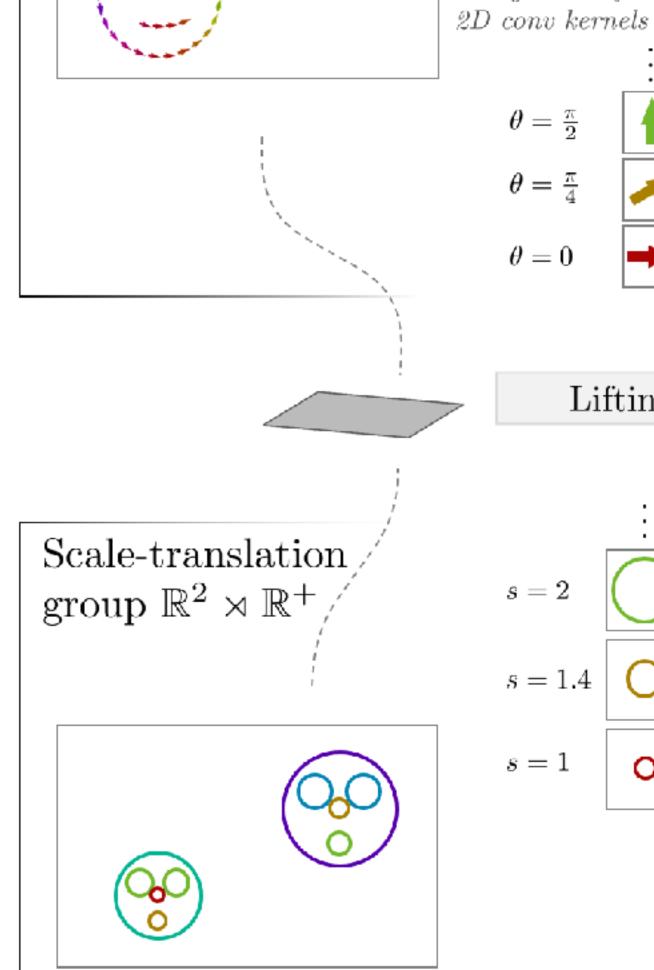
G-feature maps are equivariant w.r.t. translation and rotation of the input



G-feature maps are equivariant w.r.t. translation and scaling of the input



Activation for circles at each position and scale

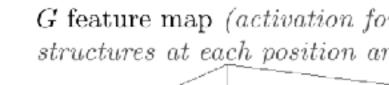


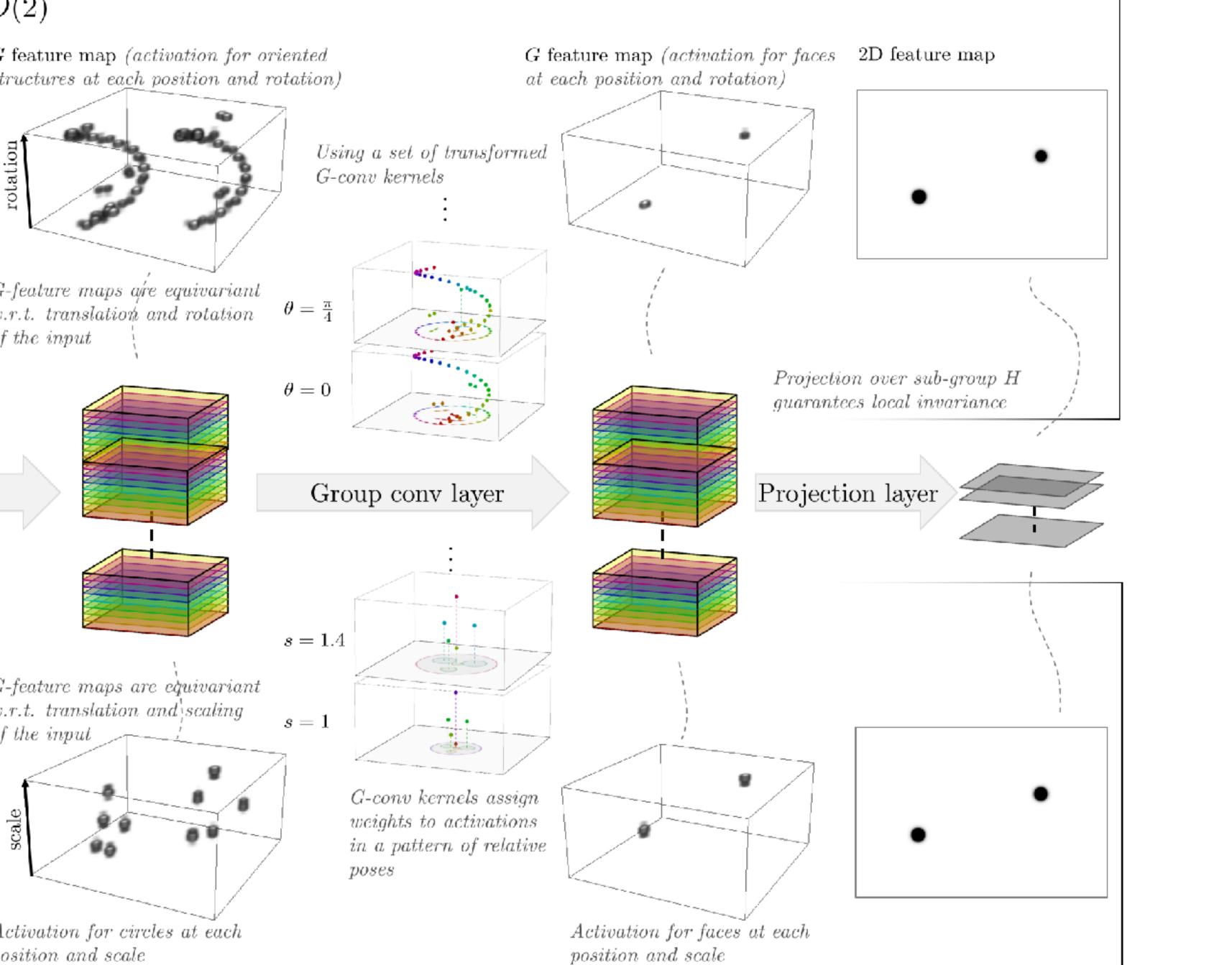
Using a set of transformed

Lifting layer

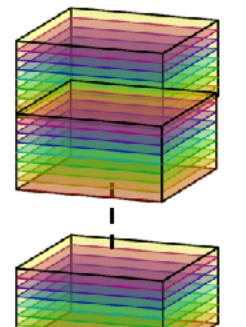
0

2D feature map

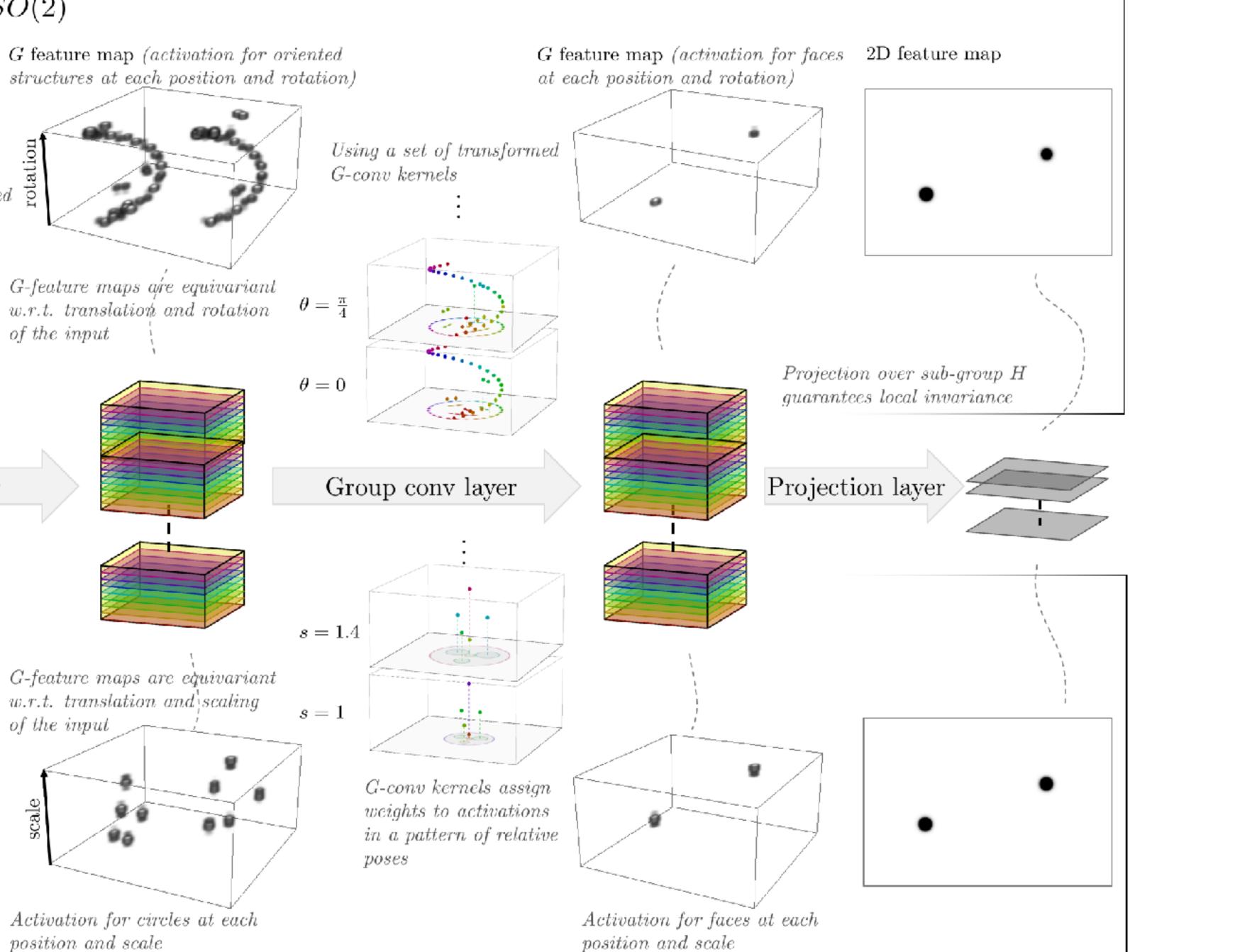




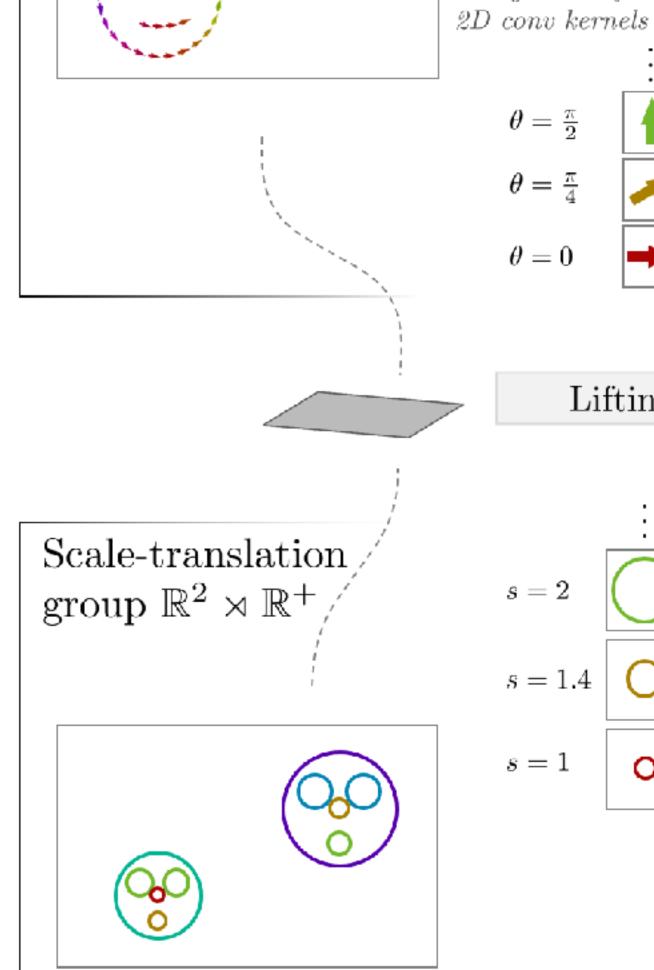
G-feature maps are equivariant w.r.t. translation and rotation of the input



G-feature maps are equivariant w.r.t. translation and scaling of the input



Activation for circles at each position and scale



Content of this talk

1. Why do we want equivariant learning models? - Geometric guarantees + weight sharing/sample efficiency

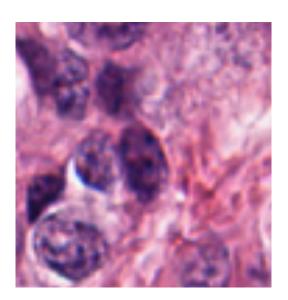
2. A group theoretical view on recognition by components (capsule nets) - Group convolutions perform pattern recognition by components

3. Experimental examples

4. Theorem: Linear maps between feature maps are equivariant iff they are group convolutions



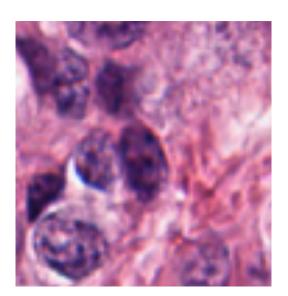
Architecture for rotation invariant mitotic cell detection



UNIVERSITY OF AMSTERDAM

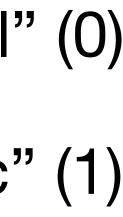


Architecture for rotation invariant mitotic cell detection



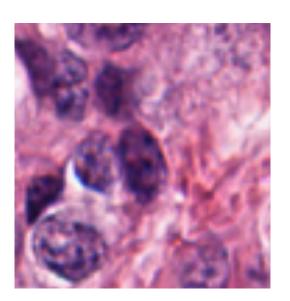
University of Amsterdam

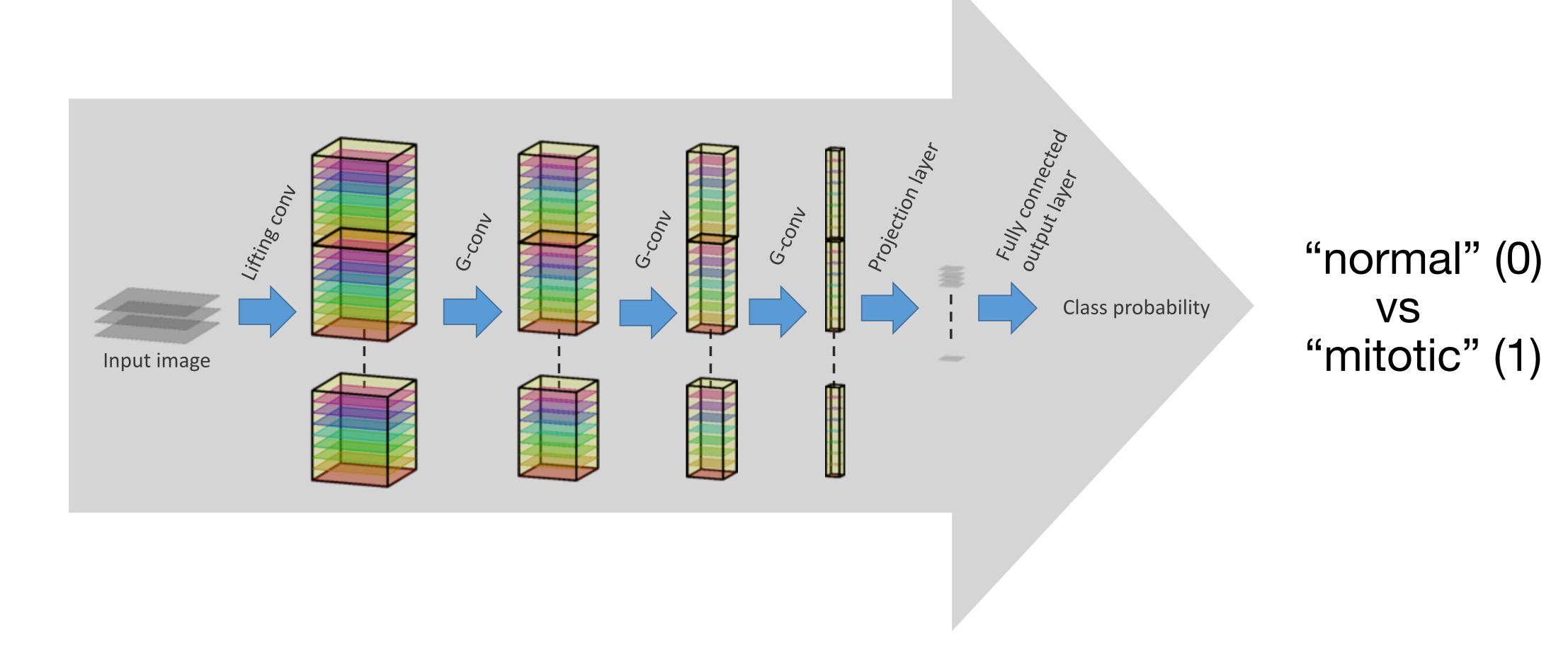
"normal" (0) VS "mitotic" (1)





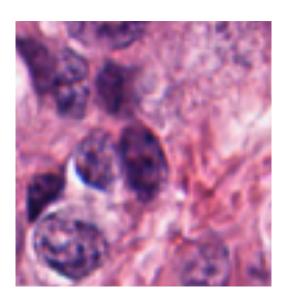
Architecture for rotation invariant mitotic cell detection

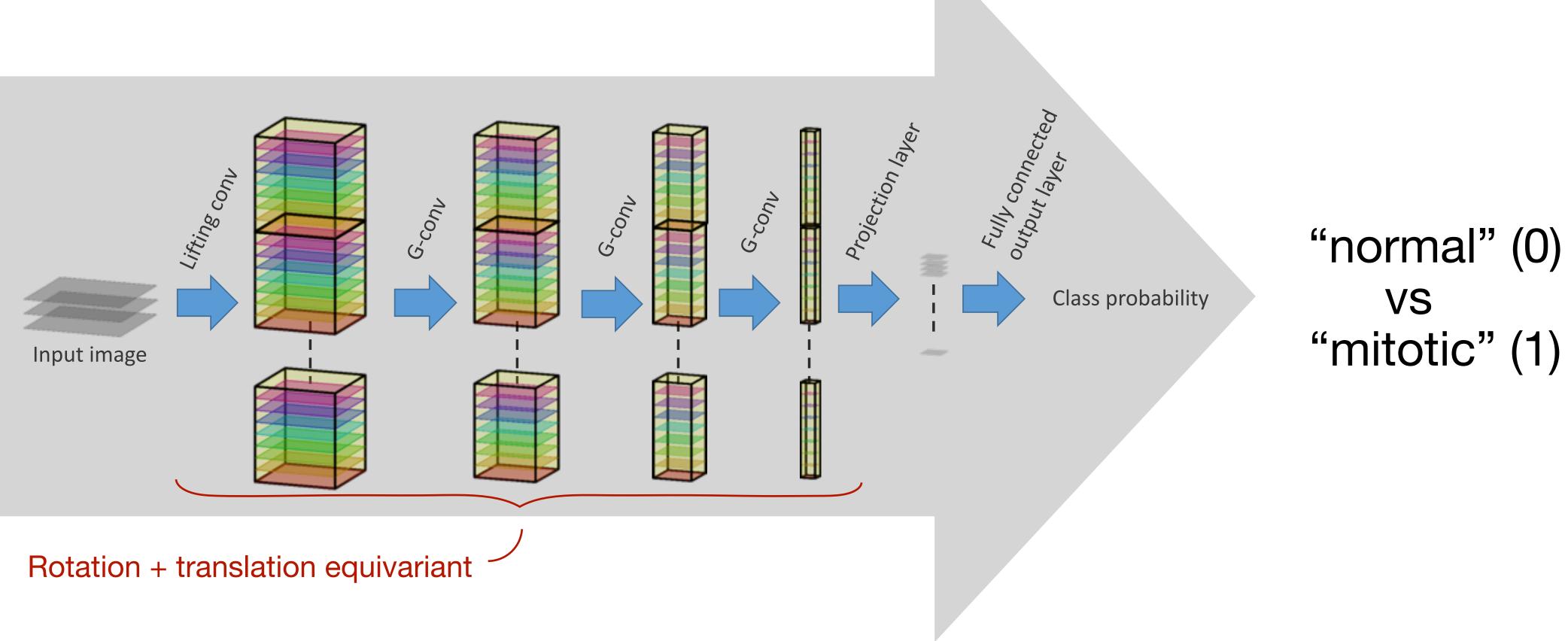




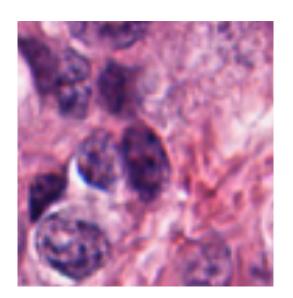


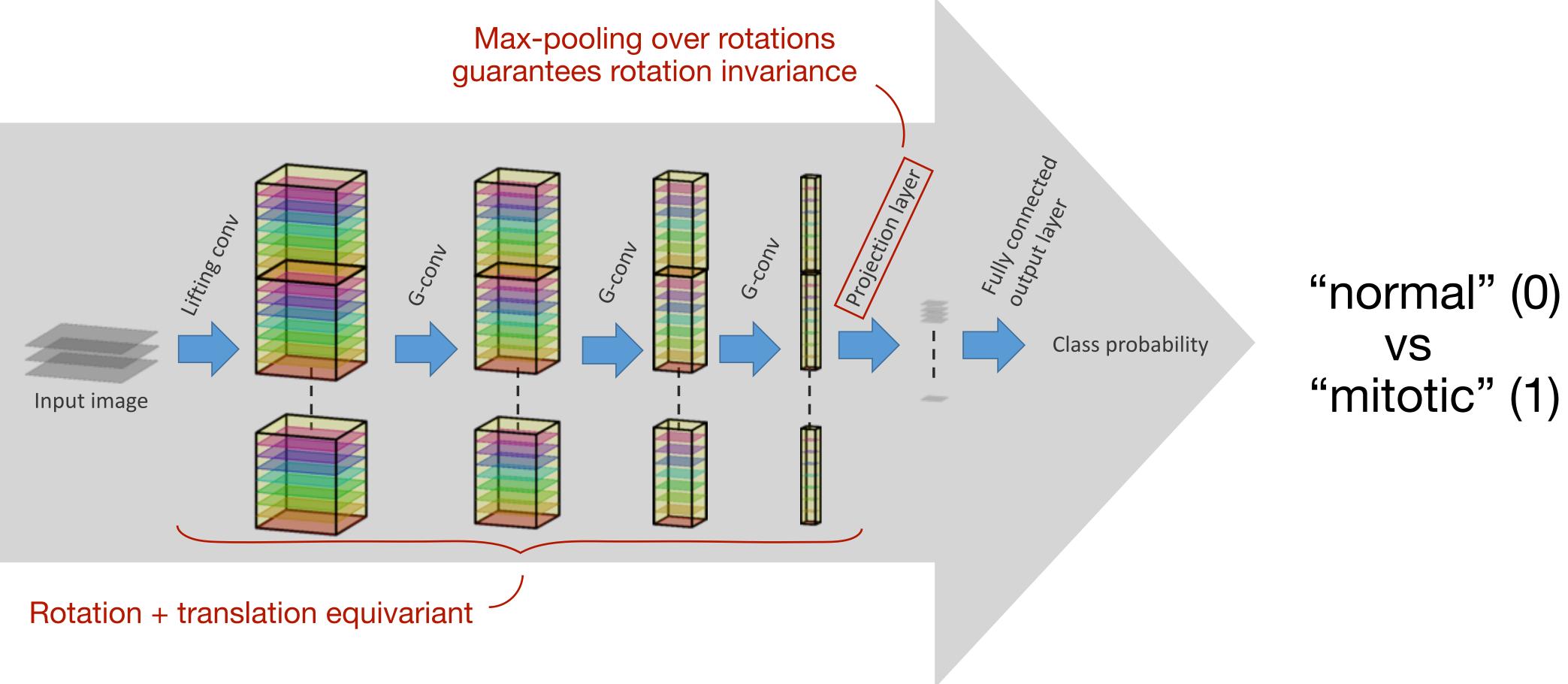






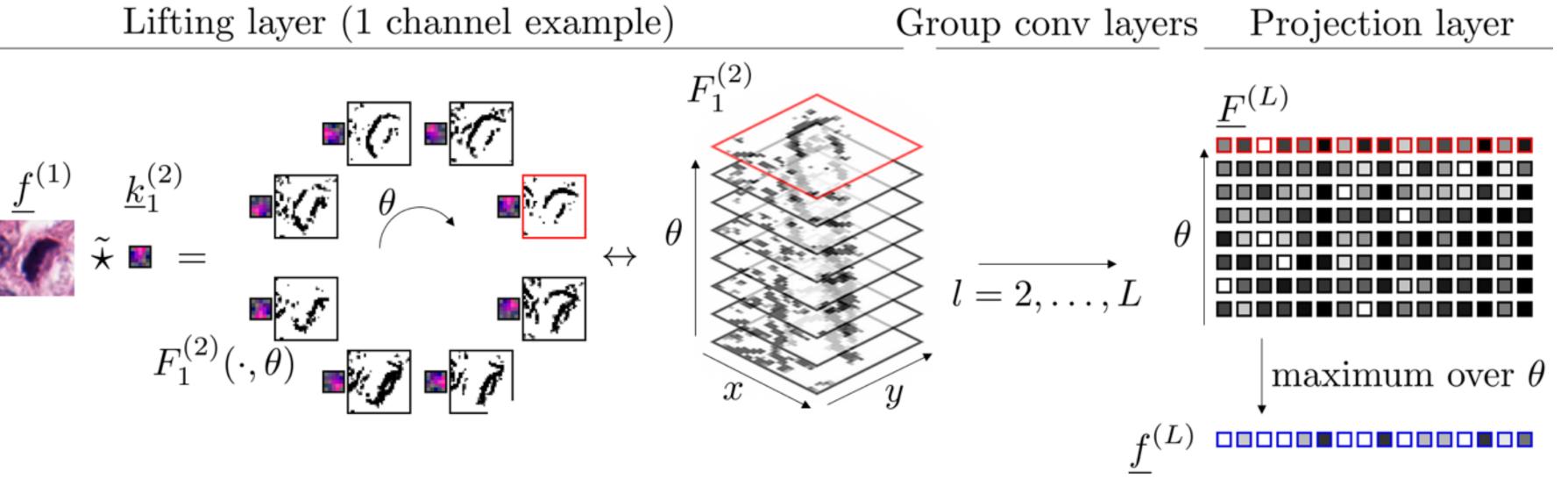






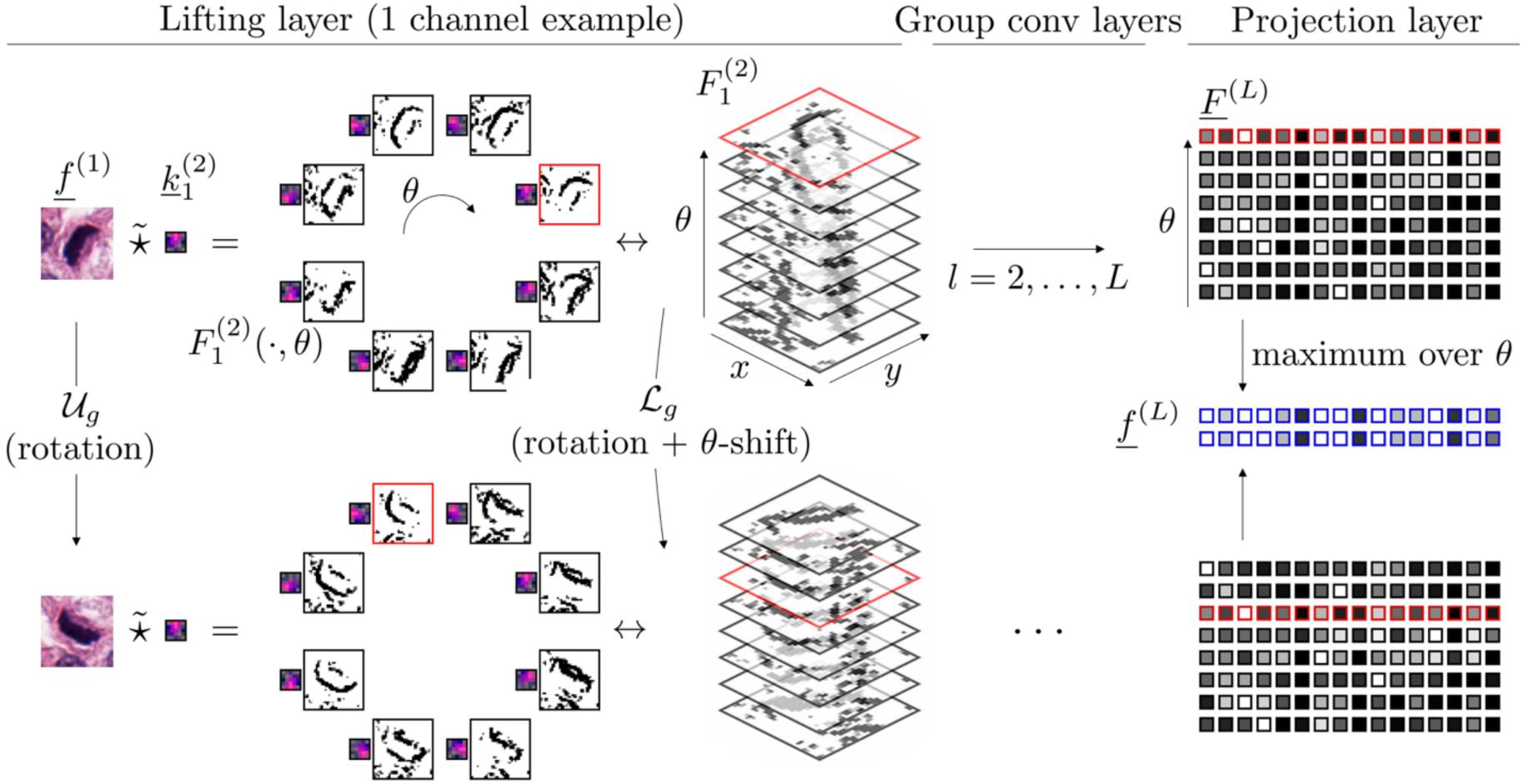


Bekkers & Lafarge et al. MICCAI 2018



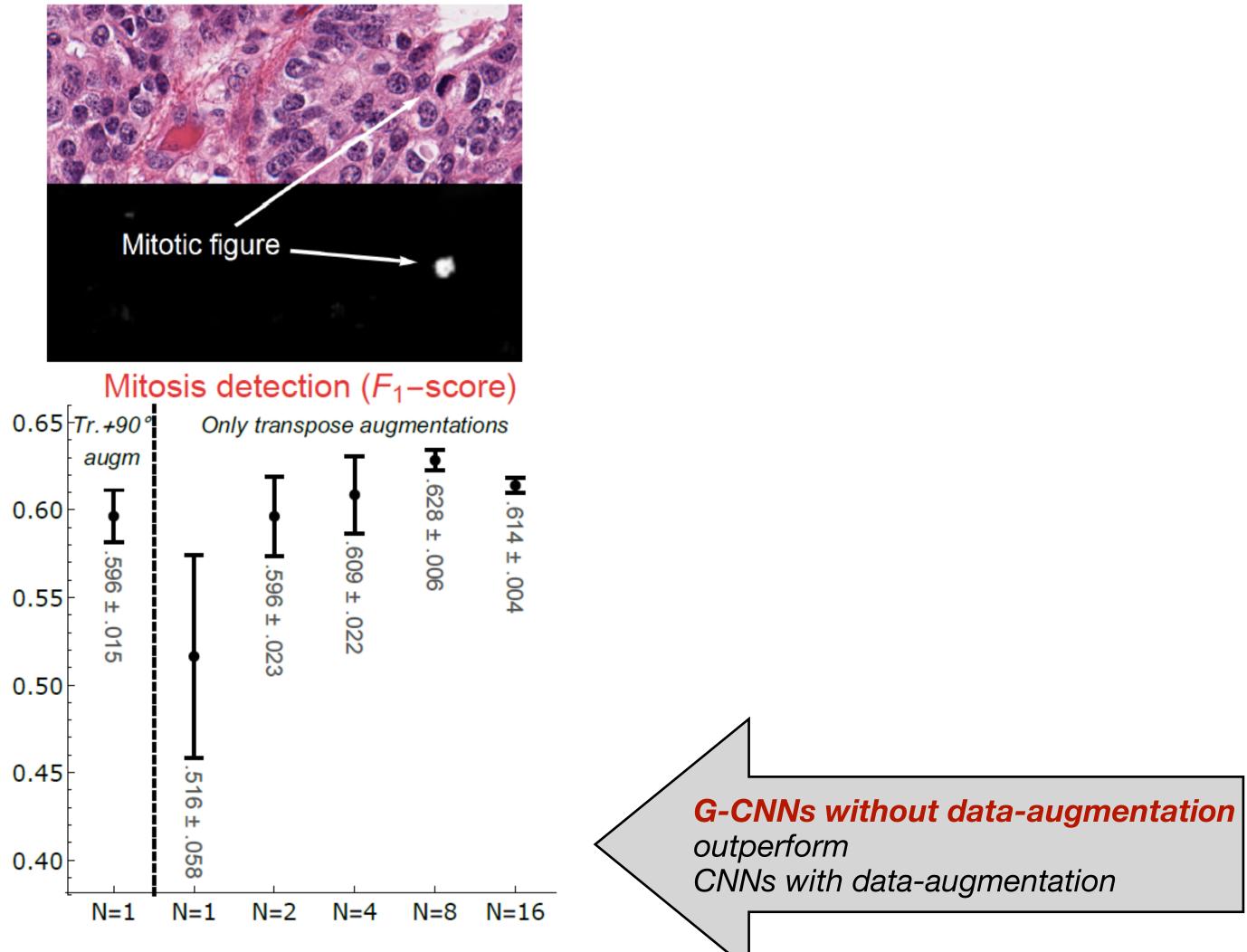


Bekkers & Lafarge et al. MICCAI 2018



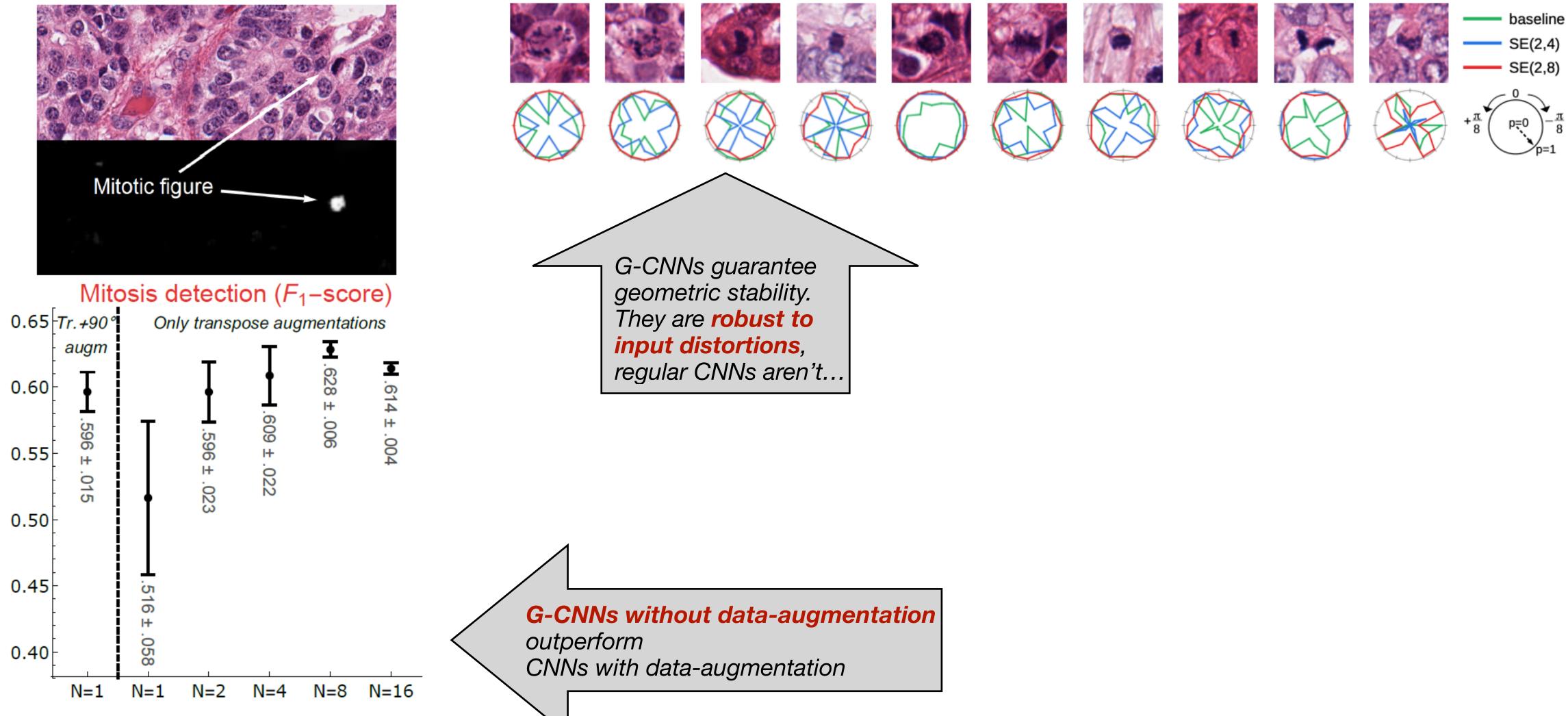


Bekkers & Lafarge et al. MICCAI 2018





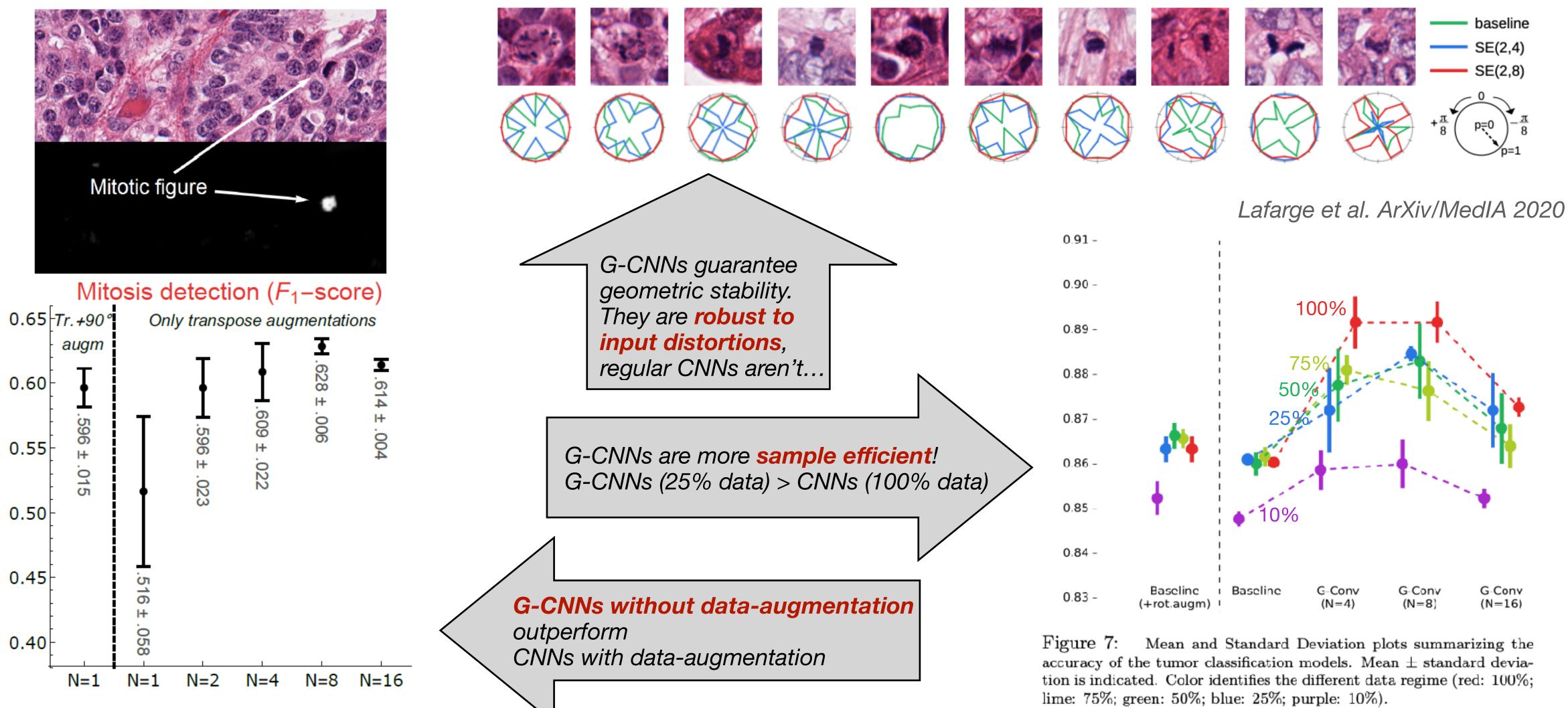
Bekkers & Lafarge et al. MICCAI 2018



Lafarge et al. MedIA 2020



Bekkers & Lafarge et al. MICCAI 2018

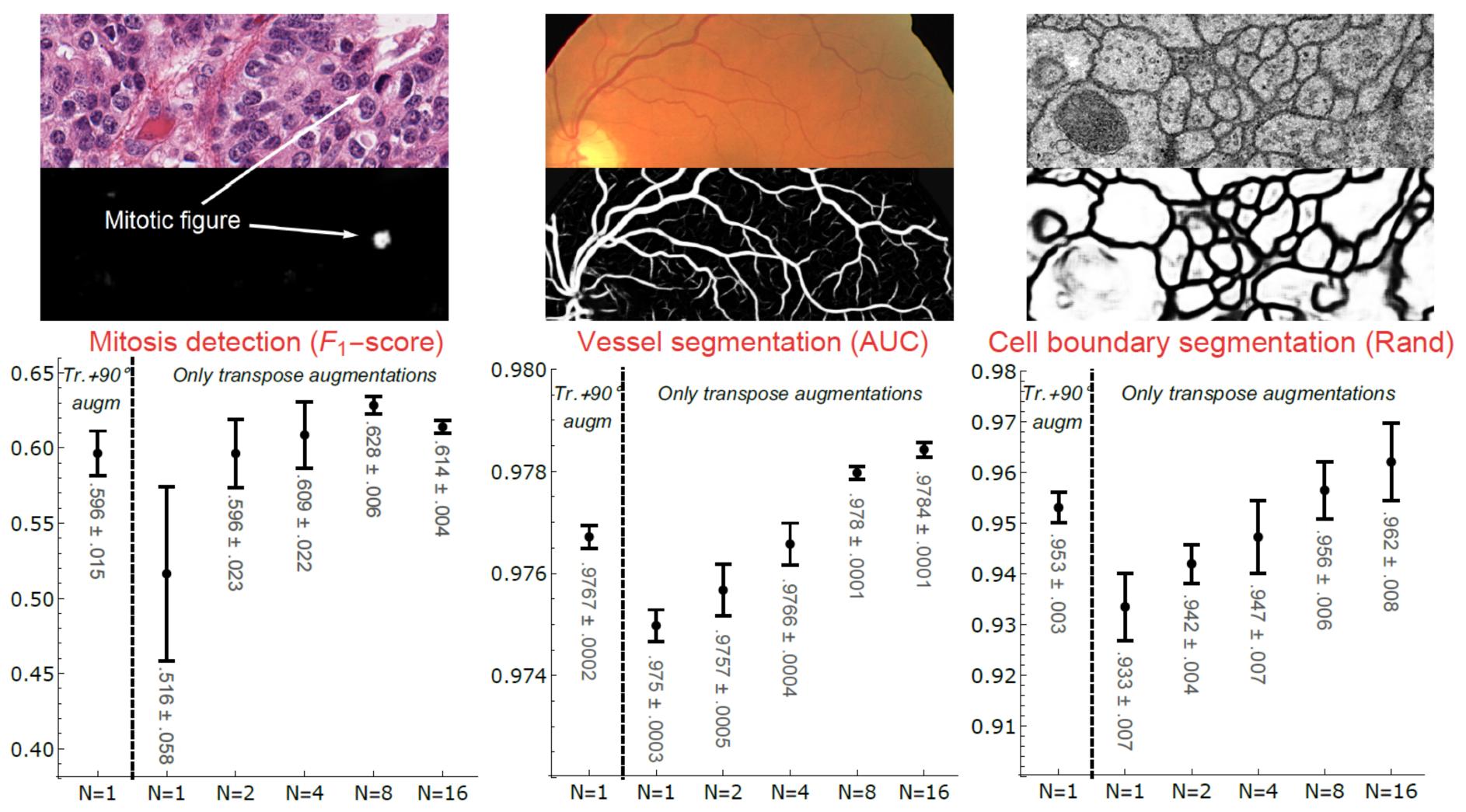


Lafarge et al. MedIA 2020



Experiments in medical image analysis

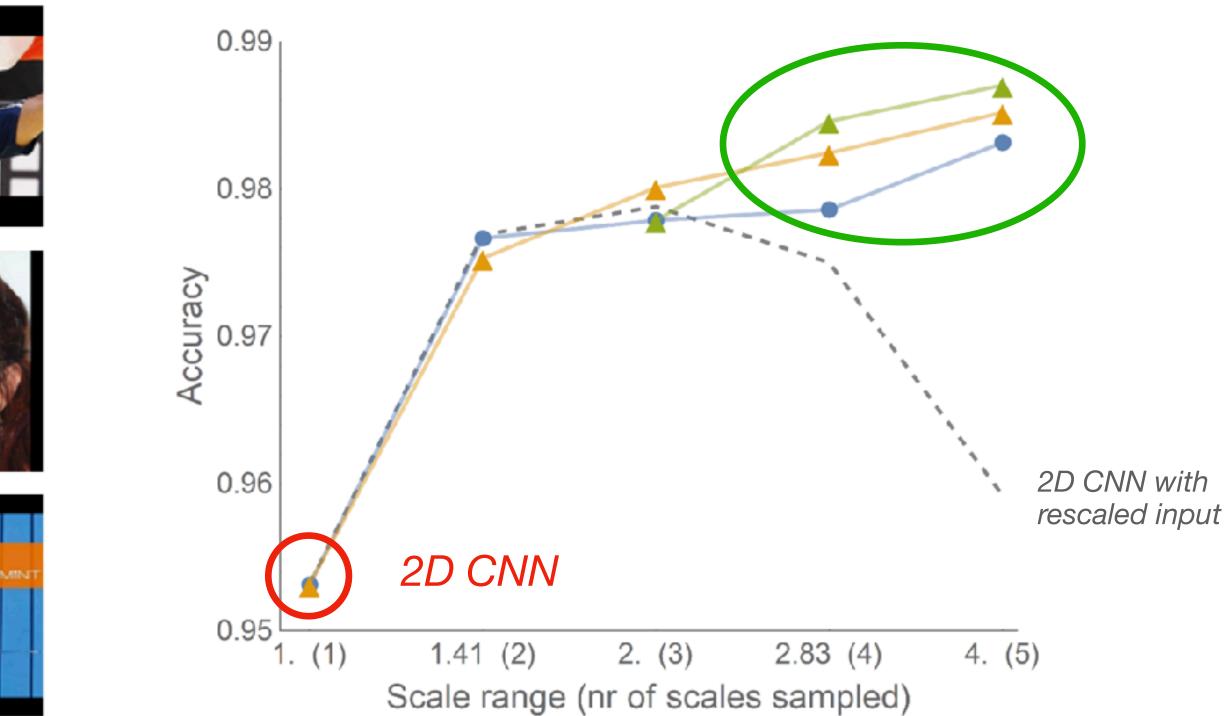
Bekkers & Lafarge et al. MICCAI 2018





From rotation to scale equivariant CNNs

Bekkers ICLR 2020



Translation + scale equivariant G-CNNs



G-CNNs rule!

- The right inductive bias: guaranteed equivariance (no loss of information)
- Performance gains that can't be obtained by data-augmentation alone (both local and global equivariance/invariance)
- Increased sample efficiency (increased weight sharing, no geometric augmentation necessary)



G-CNNs rule!

- The right inductive bias: guaranteed equivariance (no loss of information)
- Performance gains that can't be obtained by data-augmentation alone (both local and global equivariance/invariance)
- Increased sample efficiency (increased weight sharing, no geometric augmentation necessary)

Yanzhao Zhou ¹ , Qiziang Ye ¹ , Q ¹⁴ , Q ¹⁴ , Q ¹⁴ , Q ¹⁴ , ¹ University of Chinese A ¹ Duke Univ showynchestiff Wands see as the (qu'wy)			On the Generalization of Equivariance and Convolution in Neural N to the Action of Compart Groups					ice Esteves ¹ , C ¹ GRA (marke	
Bana Davrda Var	Abstract Accest Accessive USEANNE acc			Rist Ka	ndor ⁱ St	ubhandu Trivedi ¹		Translati	
Published as a	conference paper at ICLR	2020	Abstract			Reads dealers	No	ctworks fo	
B-SPLINE CNNS ON LIE GE Frik J. Bekters * Amteriar Machine Longing 1 (1)			Crossivitional neural is incready successful in the main because they or some town. How have been generalize this framewore shalling graphs are fram this paper we give a ri- ment of convolution and	Tensor field network Rotation- and translation-equiv networks for 3D point e			ation-equivar	ariant neura	
Austerdar ¹									
University	Roto-Tra	nelation	Equivariant Conv						
e. 5. heit Application Marme W. Lainge !, Enk 3.			to Histopathology Iokan ¹ , Jorian P.W. Ph	Notification Theorems" Stanford, Chilfornia, USA Stanford, Chilfornia, USA nothernas fortant/ord, add		General	General E(2) - Equ		
			to the former of the second						
Cr Sil 10 dia	² Devolute of Malace		Group Equivari	iant Co	nvolutie	onal Network	University of A	nice Weiler" Unsterdam, QUWA Liter@uss.ml	
an a	computational particulary approximation provides a second structure of the	University of Amateodam. Most Welling University of Amateodam. Evidentity of Colifornia Irvine					the spectring of focus has derei images. Here w the framework o	al success of group to great variety of a by been on rotation or give a general de of Americale CoNe. the convolution for	
of p is to	Introduction regimes to investigation sportly of marine harming in all images andly dependent of a presention to manue sharper	Kental isation ducon trist, G hyer B	Abstract relieve Group equivation: Consolution Networks (S. CNNA), a natural genera of conveloptional networks that a maple complexity by projecting symme- CNNA res G convolutions a new type at enjoys a nationarially higher degree shoring than update conveloption hype	ai re sef sef	Convol sector I Ecces in these ecigina the ecc other s code of s	Genera	alizing Convolu to Lie Group		
Group and b	richiations. For complex in sublemprocessing as near whose global orientation of preparation was down of instance in down arbitrarily as	0-cen af the i parare	starting the second conversion appendix entrands without in availing the number dets. Group convolution hyper are as and on the implemented with regligit	ng Af Ag	in the f In this process ing sole	Ma	re Finzi Samuel St	tanton – Pavel J New York U:	
signal data 🔐	clise on law any time-dis-	acrops	ational crederal for discrete groups ge		hey to i		Abstract		
habsviorum C	ancest, madelo witaso autean i le imput constitute a source		by translations, reflections and rotation is achieve state of the art results on 0		ean co After d	_			
the input da	fauch image contrained are a seeing that the output of a re-		and restated MNIST.		ber of a fering		uivariance of convolutie ational neural networks		
	s ingral is rolated.				la secti		ge problems. While tru		
not just train	Convolutional Nation Network Additional advector files inter-	1. Introd	ction		dan'i C	equivariance prov-	ides a powerful induct	tive bias	
performance to	the translation co-variance i	Deep corve	ilutional neural networks (UNNs, co	erana)	lations factoriat		en additionally desire e		
Part of the versional atres to	dusier operators. In theorem, Stern features in any oriente	have prover	to be very powerful models of armo-	ey data	for 4.		slormations, such as re image data. We propos		
the error H	or complet if a specific edge		ges, video, and audio. Although a site il network tionign in currently lacking.		standin For ge	eral method to co	onstruct a convolution	al layer	
structure this	is the bulk at hand. It is so its lifter to all possible directs	amouni of c	npirical evidence supports the notion th	hai both	ious ka		t to transformations fr		
world memory h	is after in the positive direction in relation in operation country yighterating additional random		al weight rhaning and depth (among of ordani for good predictive performance		ing and and res		p with a surjective exp- g equivariance to a new		
For most sike as	container that the model with		al weight sharing is effective because		with G		ing only the group cap		
	at the artificially shows of mains. Although data sugment		 assessed y in most perception to do and data distribution are both second 		Micros	and logarithm may	ps, enabling rapid prot	styping.	
	seques has see of sea	invariant to	and data distribution are both approxi- shifts. By using the same weights to a	eralyze	In social and UL		implicity and generalit; the same model archite		
Park Line	a. Eastlayer University of T		ch part of the image, a convolution hy conversition a fully connected one, wh		nder G		tick molecular data, and		
			apacity teleon many useful masfern		os sug ing pis	tonian dynamical	systems. For Hamilton	tian sys-	
		Precenting	g 6: 35 ⁻¹ International Conference on	Nation	impres-		nce of our models is as humanic compressions		
		Association, No.	(c) for the international state strugger and the at York, MP, USA, 2016. UMLIN: W38CP 17005 (syline and args).	solution.	peovikis votivatsi	and angular more	to exact conservation e miam.	ci datar	
		er of in the	consider a second de			v.		1	
								i	

Oriented Response Networks

1. Introduction

Symmetry prevales the natural world. The same law of gravitation governs a game of coach, the orbits of our planels, and the formation of galaxies. It is precisely because element. But in order to extend to continue of the universatival we can hope to understand finallike consolutional kernel to be a continu it. Once we started to understand the symmetries inherent — the group parameterized by a neural netwo in physic al laws, we could predict behavior in galaxies billions of light-years away by studying our own local region. of time and space. For statistical models to achieve their full potential, it is essential to incorporate our knowledge of naturally occurring symmetries into the design of algorithms and architectures. An example of this principle is the translation equivariance of convolutional layers in neural networks (LeCus et al., 1905): when an input (e.g. an image) is translated, the output of a convolutional layer is translated in the same way.

Group theory provides a machenism toreason about symmetry and equivariance. Corrolational layers are equivariant — all transformation groups and data types. I

Learning SO(3) Equivariant with Spherical C

Christine Allen-Dlanchette¹, Amee

LASP Laboratory, University of Pena hc,eller,kostas}@seas.openn.edu = r

tion Covariant Convolut for Medical Image Analy tiko Veta², Korn AJ moo Drotte' of Mathematics a ng, Eindhoven, Thi . n.w. laforgestr

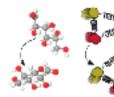
uivariant Steerable CNN:

Gabriely Cara*1 University of Amsterday ers.gabriele@gnuil.co

p equivariant networks has led in resent year ef equivation: network aschitectures. A particul-tion and reflection equivariant CNNs for plan description of E(2) -equivariant convolutions b. The theory of Stearable CNNs denoty yiel nels which depend on group representation

ral Networks for Equivarian rary Continuous Data

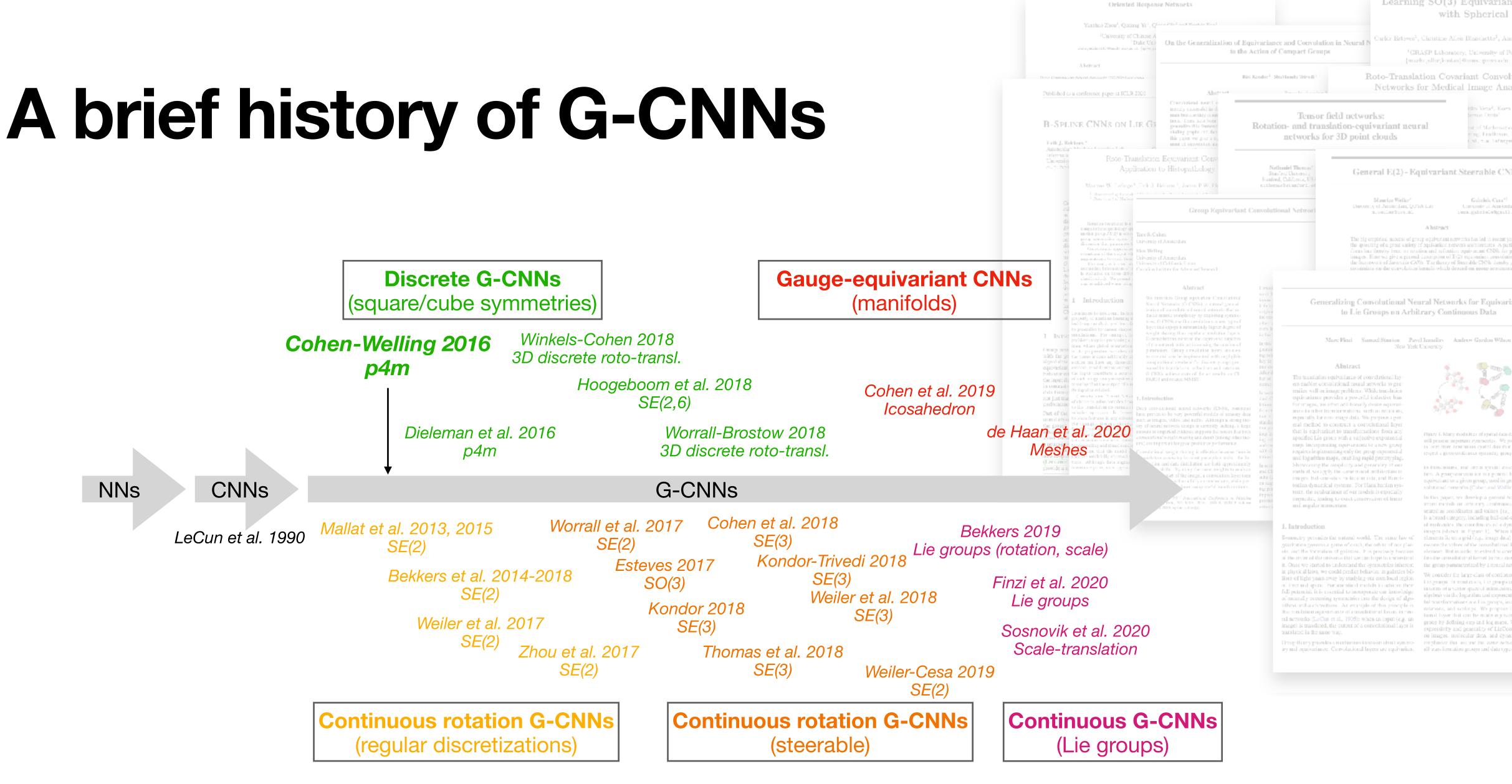
Andrew Gardon Wilson kmailey



Please J. Many modalities of spatial data do still possess important symmetries. We prope to learn from continuous control data that can respect a given confirmous symmetry group.

tion. A group convolution is a general line quivariant to a given group, used in group riant nodels or arbitrary continuous sented as coordinates and values $\{(x_i, f_i)\}$ is a broad category, including ball-and-stic of molecules, the coordinates of a dyna intega (shown in Figure D. When the elements lie en a grid (e.g., image data) en merate the values of the convolutional ker-We consider the large class of continuous ; Lie groups. In most cases, Lie groups can l interns of a vector space of infinitesimal as algebra) via the logarithm and exponential. ful transformations are Lie groups, inclurotations, and scalings. We propose Lief tional loyer that can be made equivarian group by defining exp and log maps. We expressivity and generality of LieConv w ou images, molecular data, and dynamic emphasize that we use the same network

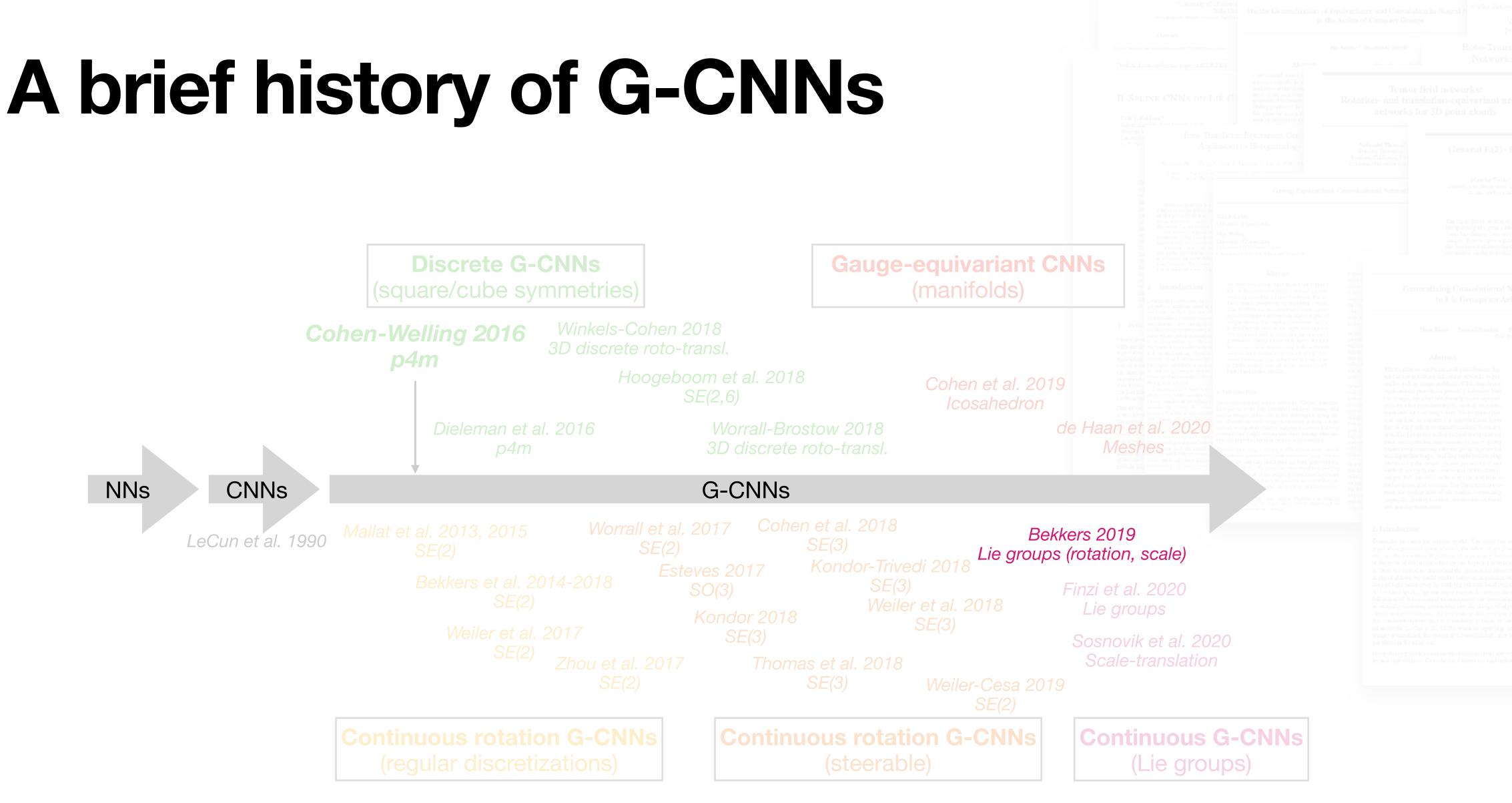




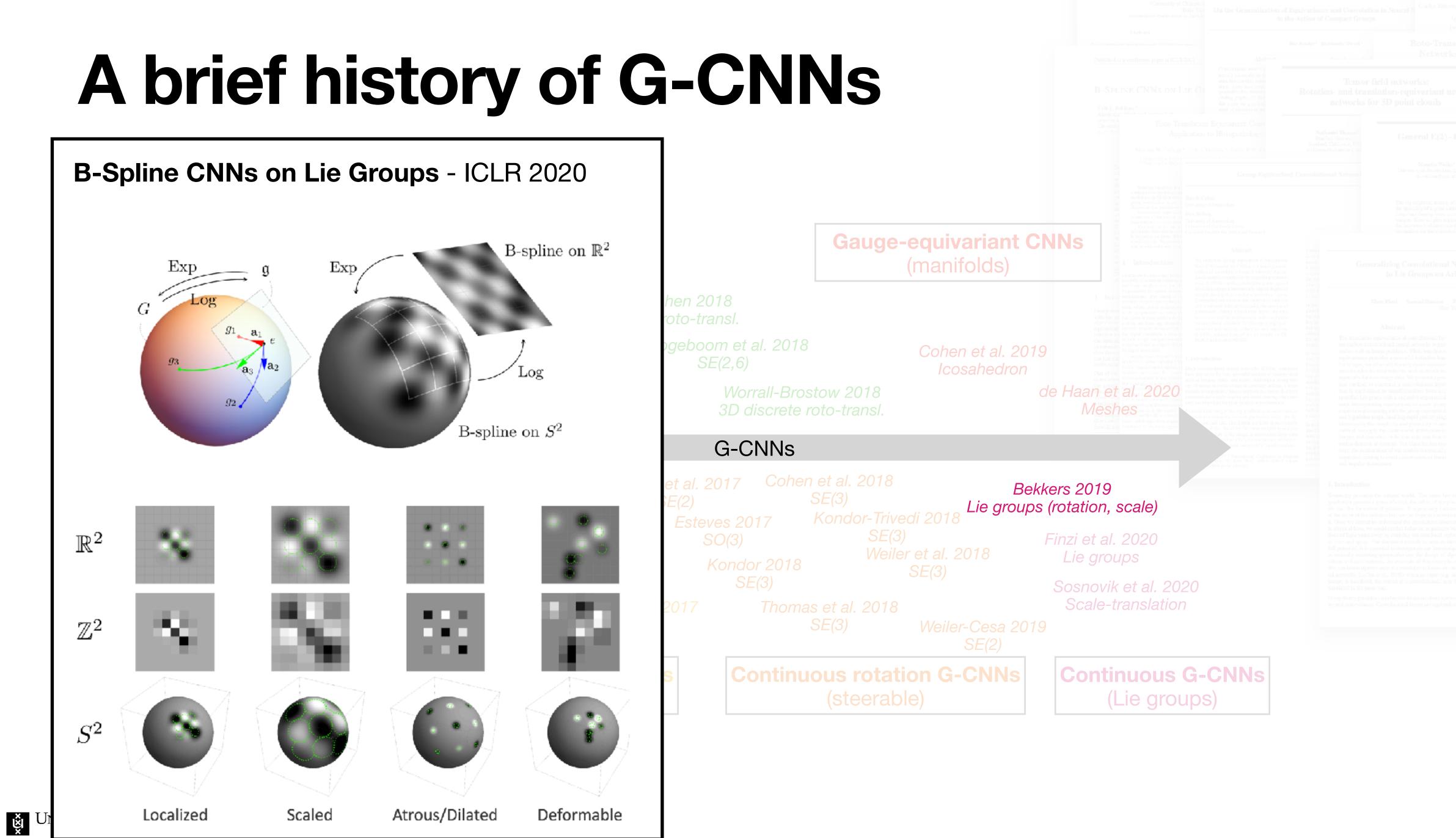
Learning SO(3) Equivariant with Spherical C

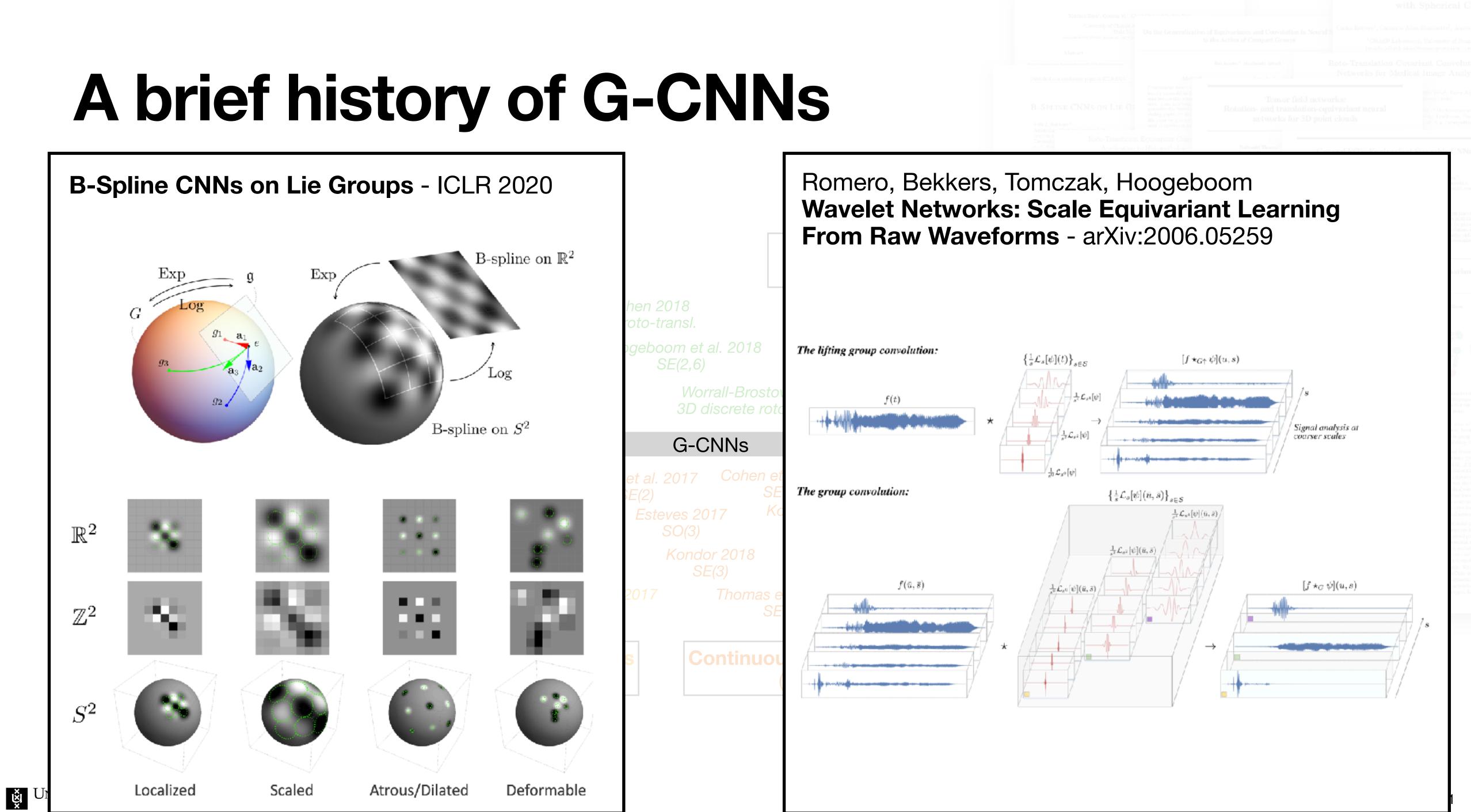


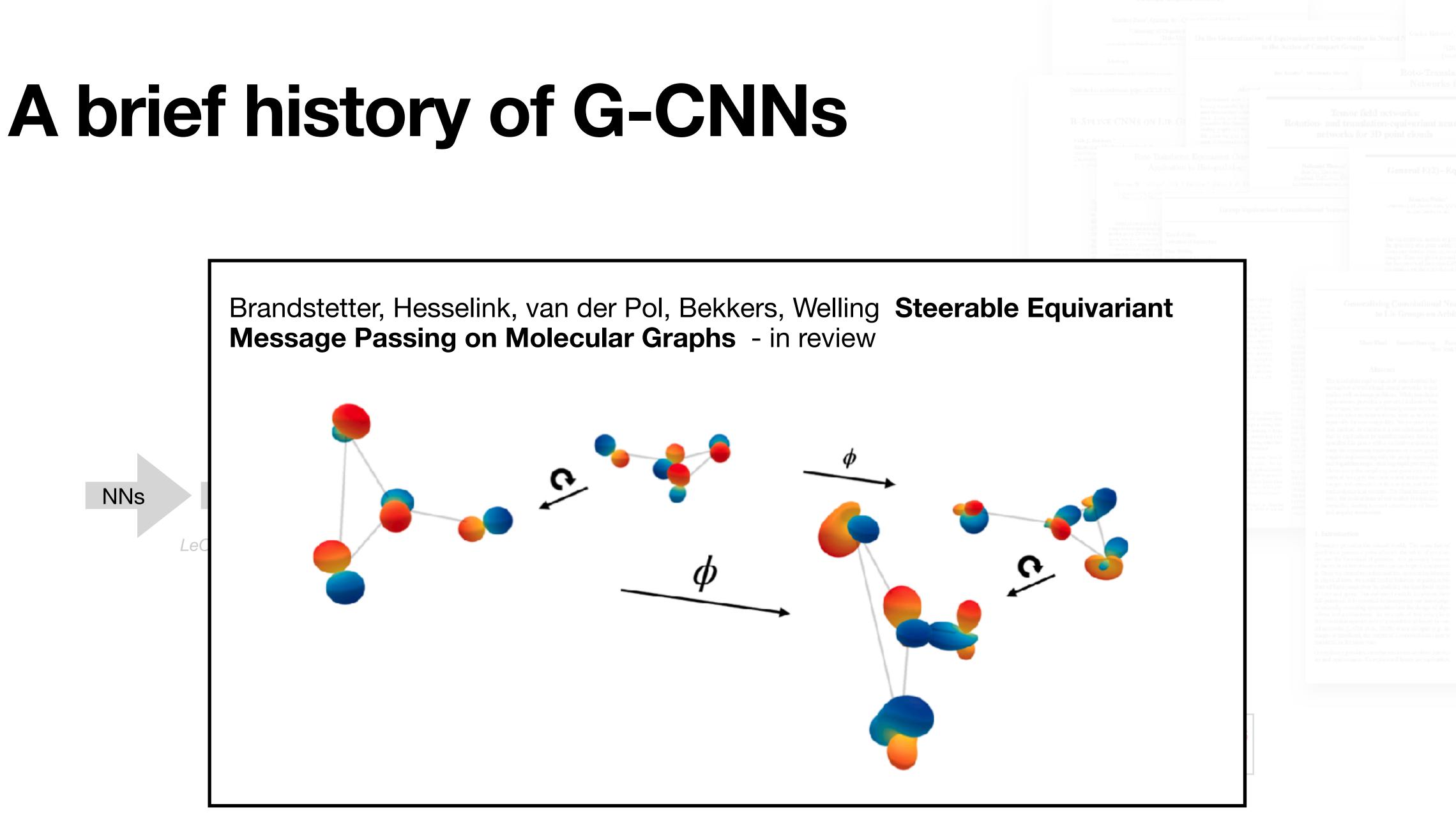




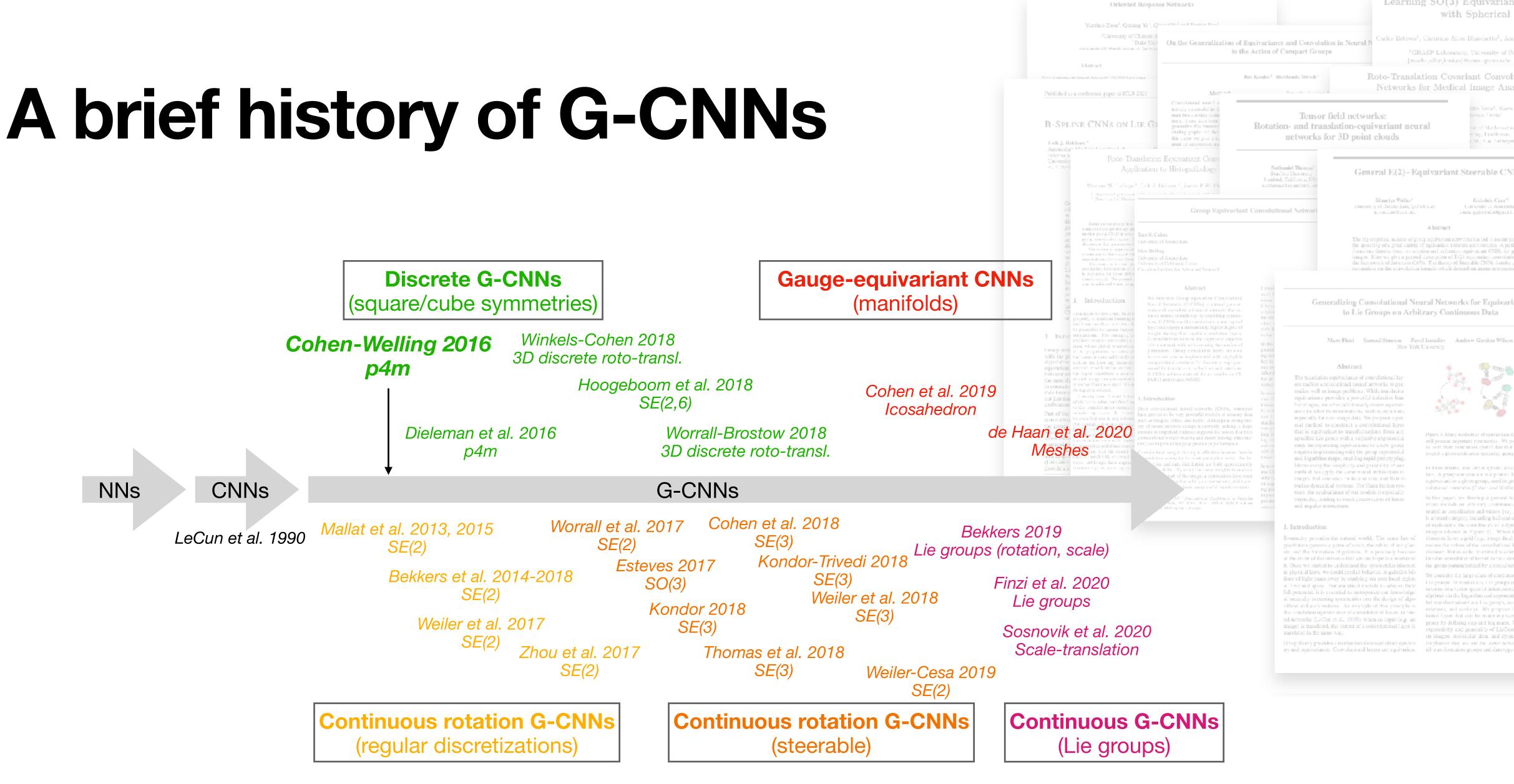
41







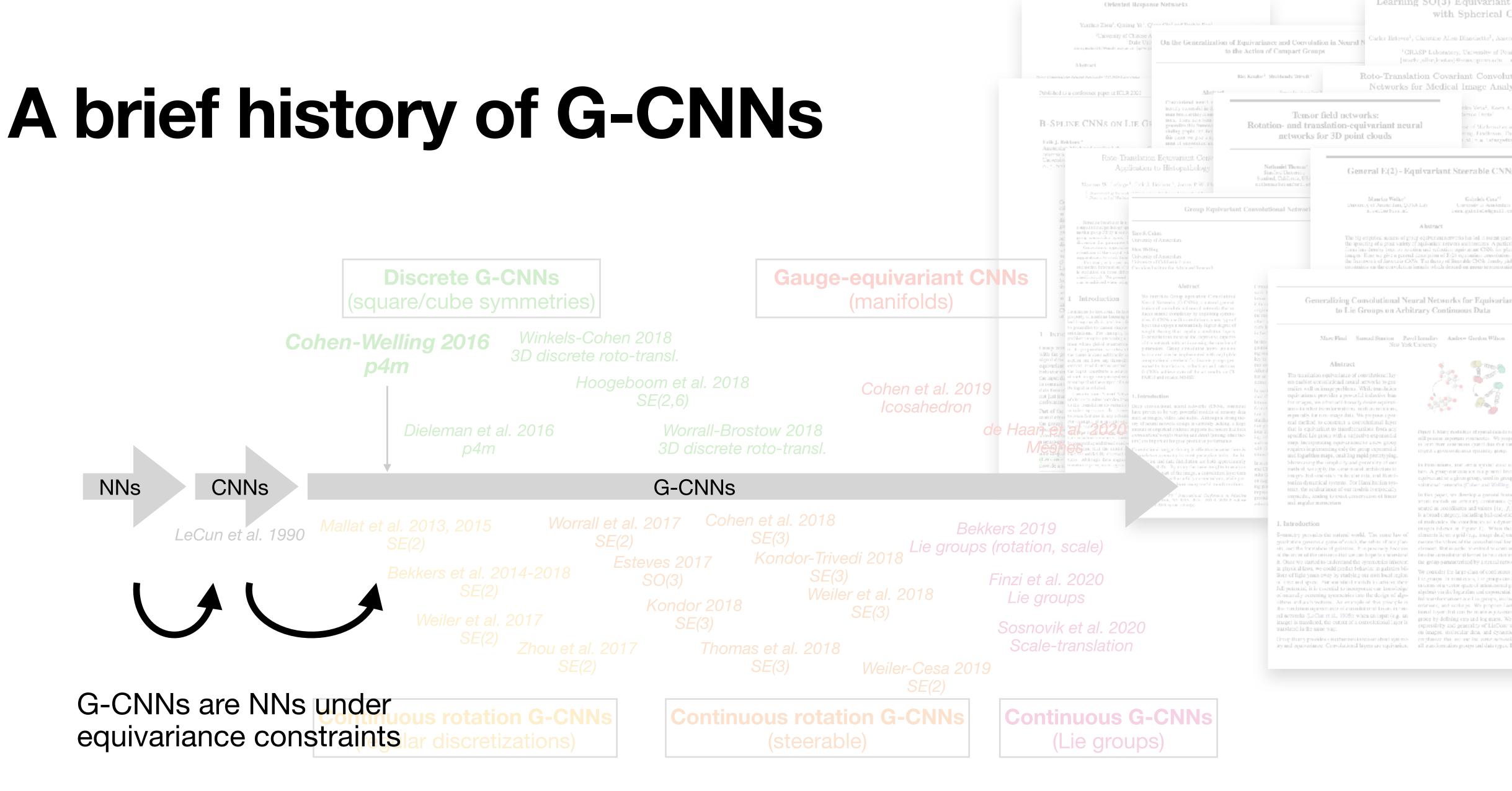




Learning SO(3) Equivariant with Spherical C







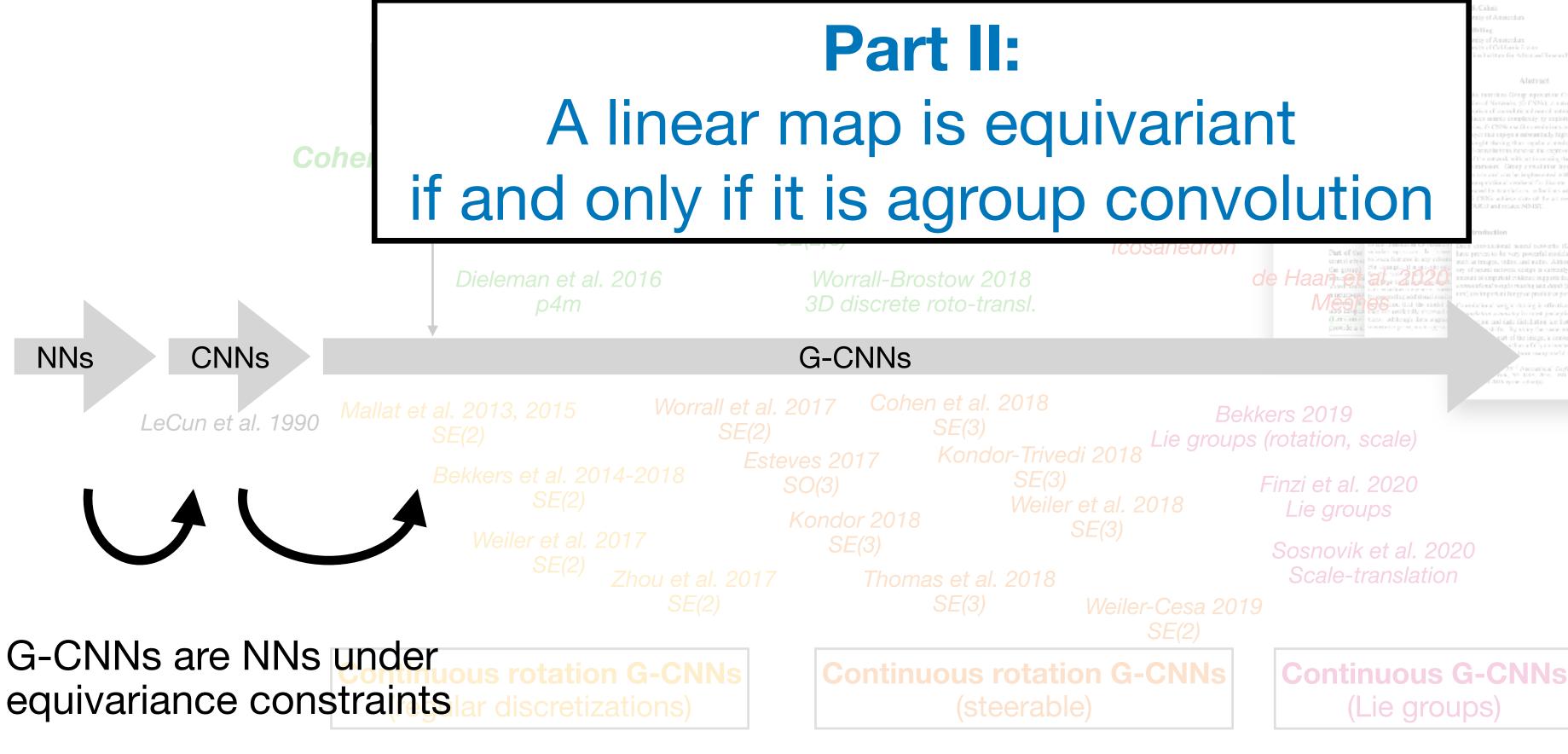
Learning SO(3) Equivariant with Spherical C

Roto-Translation Covariant Convolu





A brief history of G-CNNs



Oriented Response Network

B-Spline CNNs on Lie Gi

Tensor field networks:

networks for 3D point clouds

onian dynamical systems. For Hamiltonia

of Equivariance and Corvolution in to the Action of Compact Group

Learning SO(3) Equivariant with Spherical C

Rotation- and translation-equivariant neural

General E(2) - Equivariant Steerable CNN

Generalizing Convolutional Neural Networks for Equiva to Lie Groups on Arbitrary Continuous Data





Content

Part I: Introduction to group convolutions

- * Motivation
- * Introduction to group theory
- * Regular group convolutional neural networks
- * Applications

Part II: General theory for group equivariant deep learning

- * Group convolutions are all you need!
- * Deeper into group theory: representation theory, homogeneous spaces
- * Characterization of types of group equivariant layers

Part III: Steerable group convolutions

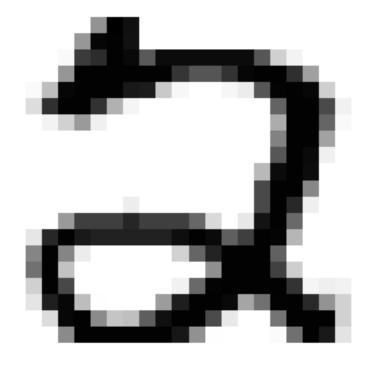
* Deep dive into group theory: irreducible representations, steerable operators and vector spaces * Examples of steerable group convolutions: Spherical data and Volumetric data/3D point clouds



44

Classical artificial neural networks

What's my input? $\underline{x}^0 \in \mathcal{X} = ?$



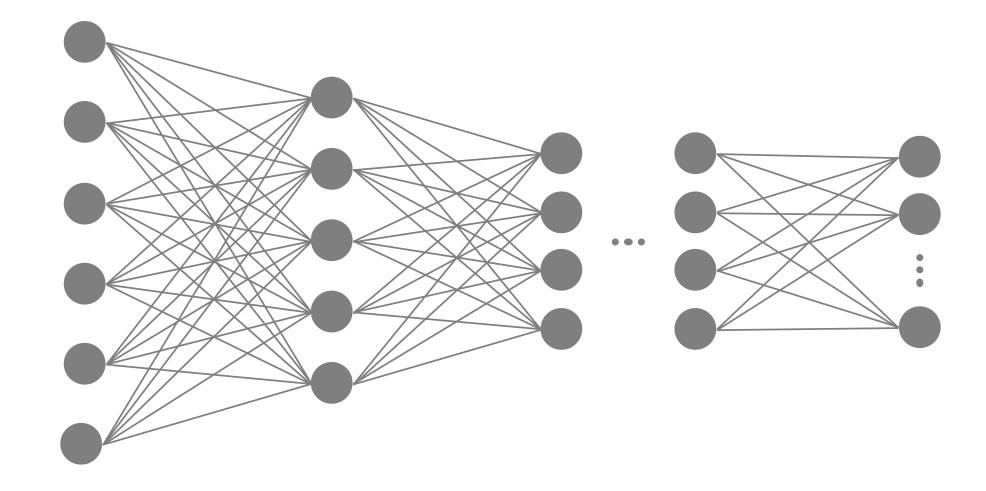


Image analyst: $\underline{x}^0 \in \mathcal{X} = \mathbb{L}_2(\mathbb{R}^2)$

Naive deep learner: $\underline{x}^0 \in \mathcal{X} = \mathbb{R}^{784}$



Classical artificial neural networks

input vector

 \underline{x}^{0}

What's my input? $\underline{x}^0 \in \mathcal{X} = ?$

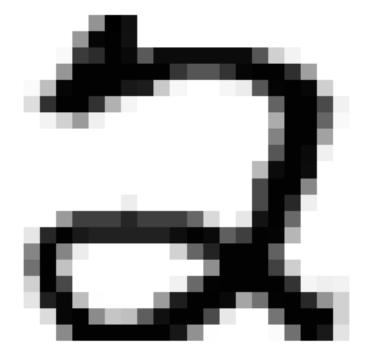
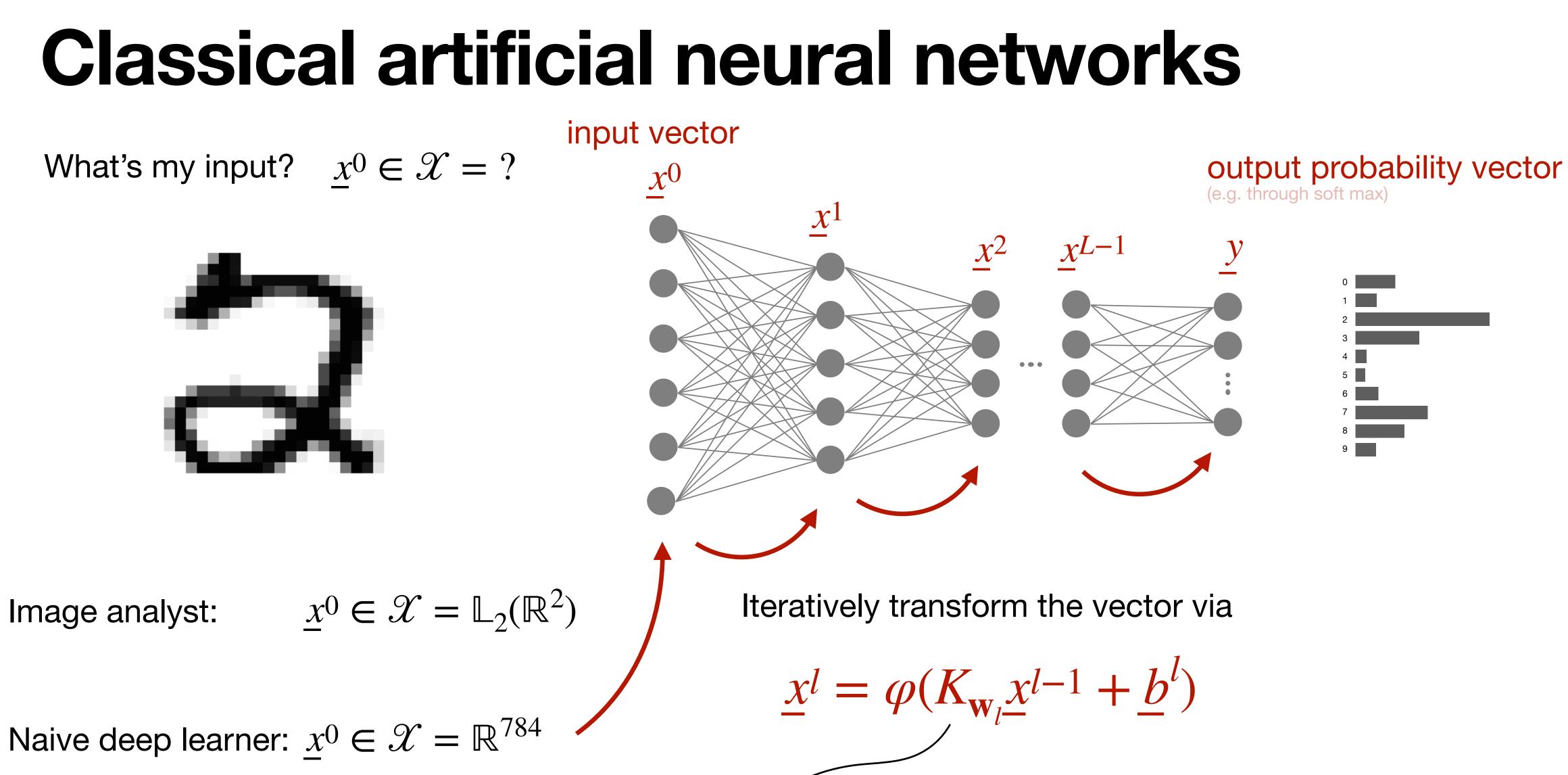


Image analyst: $\underline{x}^0 \in \mathcal{X} = \mathbb{L}_2(\mathbb{R}^2)$

Naive deep learner: $\underline{x}^0 \in \mathcal{X} = \mathbb{R}^{784}$



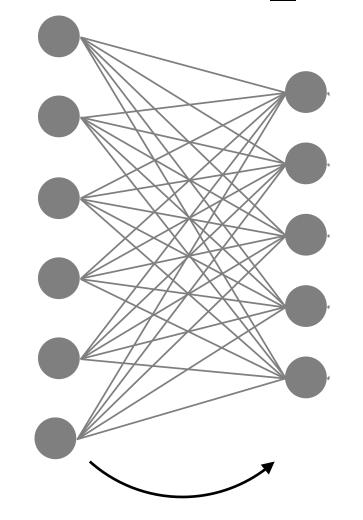


Linear map: matrix-vector multiplication with $K_{\mathbf{w}_{l}} \in \mathbb{R}^{N^{l} \times N^{l-1}}$



Classical artificial NNs in the continuous world

Working with vectors $x \in \mathcal{X} = \mathbb{R}^{N^x}$



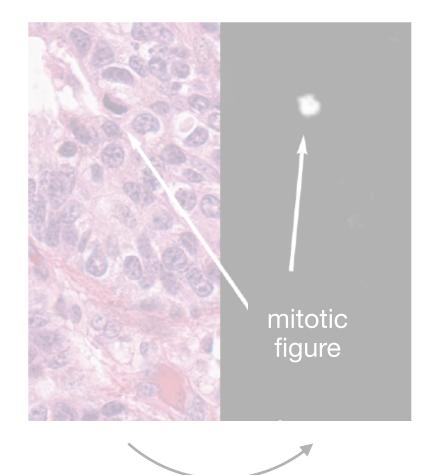
Iteratively transform the **vector** in \mathbb{R}^{N^x} via

$$\underline{y} = \varphi(K\underline{x} + \underline{b}^l)$$

Linear map: matrix-vector multiplication with $K \in \mathbb{R}^{N^{y} \times N^{x}}$

$$y_j = \sum_{i} K_{i,j} x_i$$

Working with feature maps $f \in \mathcal{X} = \mathbb{L}_2(X)$



Iteratively transform the **feature map** in $\mathbb{L}_2(X)$

$$f^{out} = \varphi(Kf^{in} + b^l)$$

Linear map: kernel operator with kernel in $\mathbb{L}_1(Y, X)$

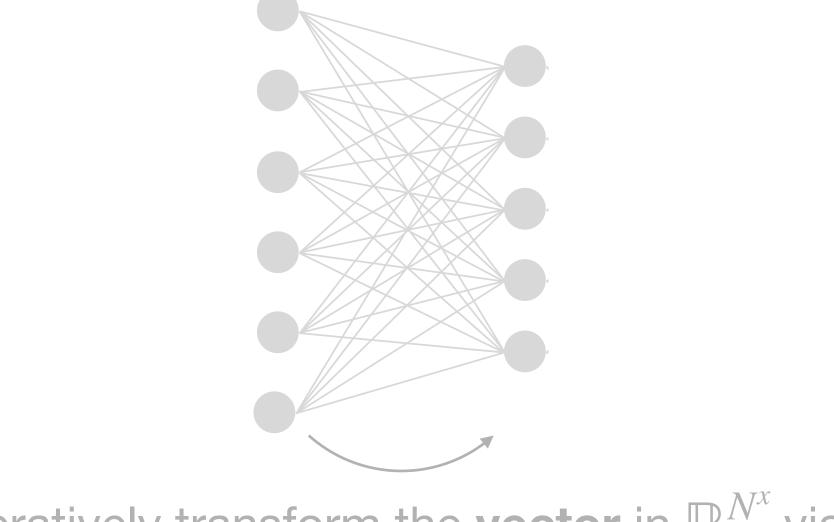
$$(Kf)(y) = \int_X k(y, x) f(x) dx$$





Classical artificial NNs in the continuous world

Working with vectors $x \in \mathcal{X} = \mathbb{R}^{N^x}$



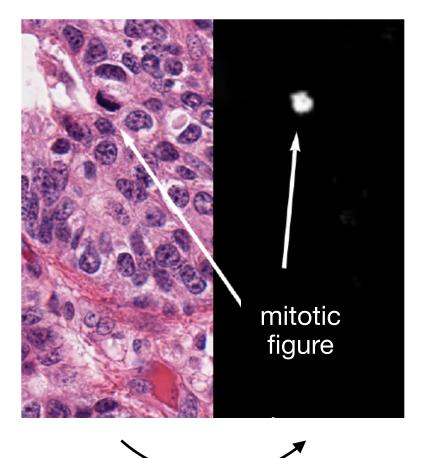
Iteratively transform the **vector** in \mathbb{R}^{N^x} via

$$\underline{y} = \varphi(K\underline{x} + \underline{b}^l)$$

Linear map: matrix-vector multiplication with $K \in \mathbb{R}^{N^{y} \times N^{x}}$

$$y_j = \sum_i K_{i,j} x_i$$

Working with feature maps $f \in \mathcal{X} = \mathbb{L}_2(X)$



Iteratively transform the **feature map** in $\mathbb{L}_2(X)$

$$f^{out} = \varphi(Kf^{in} + b^l)$$

Linear map: kernel operator with kernel in $\mathbb{L}_1(Y, X)$

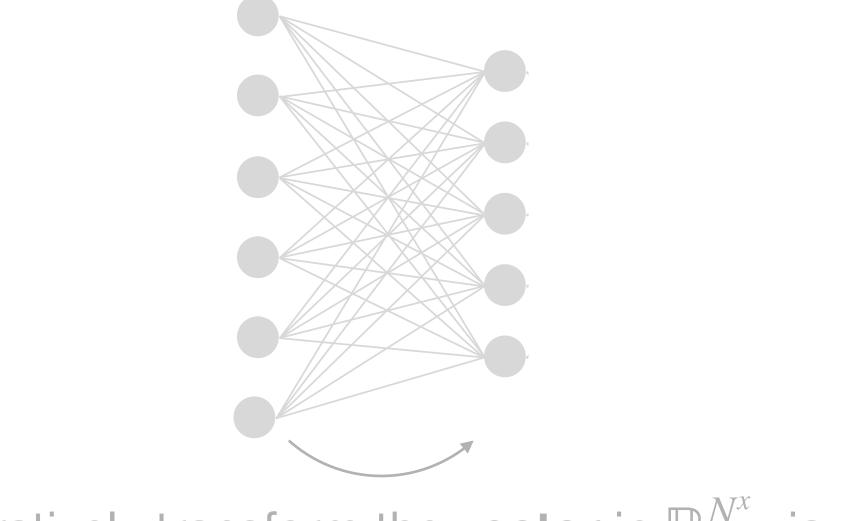
$$(Kf)(y) = \int_X k(y, x) f(x) dx$$





Classical artificial NNs in the continuous world

Working with vectors $x \in \mathcal{X} = \mathbb{R}^{N^x}$



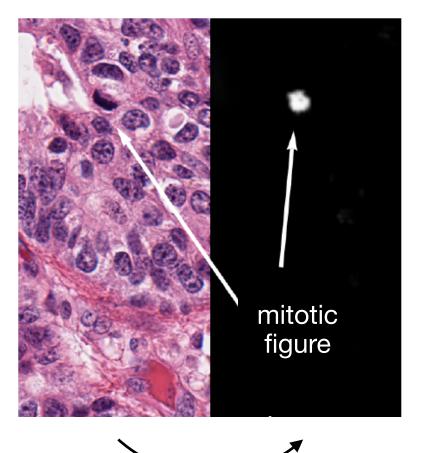
Iteratively transform the **vector** in \mathbb{R}^{N^x} via

We want K to be equivariant!

Linear map: matrix-vector multiplication with $K \in \mathbb{R}^{N^{y} \times N^{x}}$

$$y_j = \sum_i K_{i,j} x_i$$

Working with feature maps $f \in \mathcal{X} = \mathbb{L}_2(X)$



Iteratively transform the **feature map** in $\mathbb{L}_2(X)$

$$f^{out} = \varphi(Kf^{in} + b^l)$$

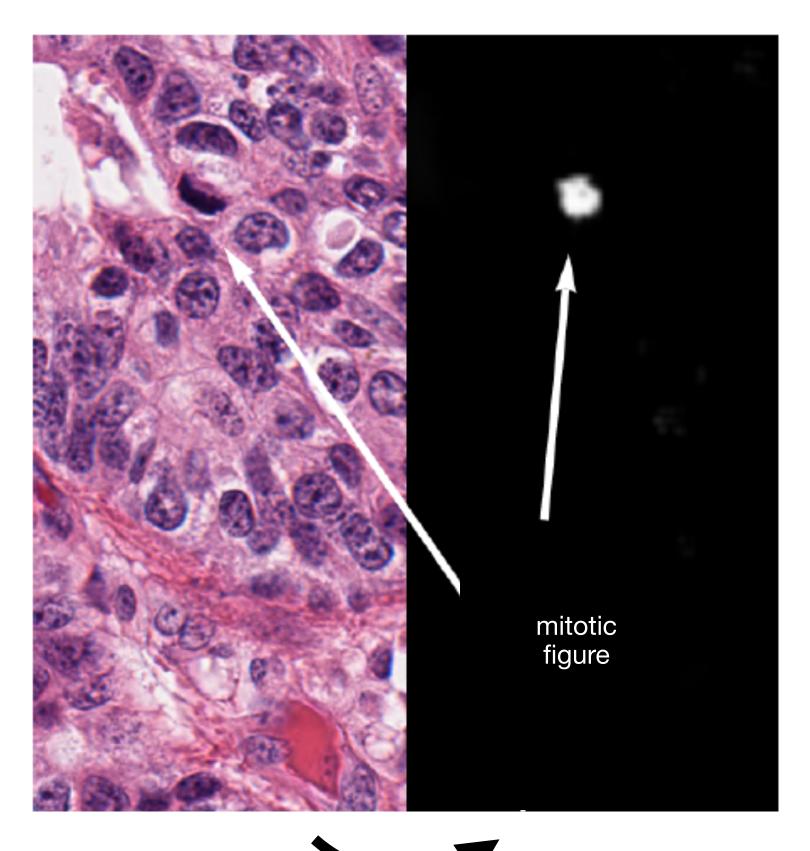
Linear map: kernel operator with kernel in $\mathbb{L}_1(Y, X)$

$$(Kf)(y) = \int_X k(y, x) f(x) dx$$





Neural Networks for Signal Data



The linear map has to be an integral transform with a two-argument kernel (Dunford-Pettis theorem)

 $\mathscr{K}: \mathbb{L}_2(X)^{N_l} \to \mathbb{L}_2(Y)^{N_{l+1}}$

Let's build neural networks for signal data via the layers of the form:

$$\underline{f}^{l+1} = \sigma(\mathscr{K}\underline{f}^l + \mathbf{b}^l)$$

$$(\mathcal{H}f)(y) = \int_X \mathbf{k}(y, x) f(x) dx$$





Lecture notes Theorem 3.2:

origin $y_0 \in Y$ and let $g_v \in G$ such that $\forall_{v \in Y} : y = g_v y_0$.

Then \mathcal{K} is equivariant to group G if and only if:

1. It is a group convolution:

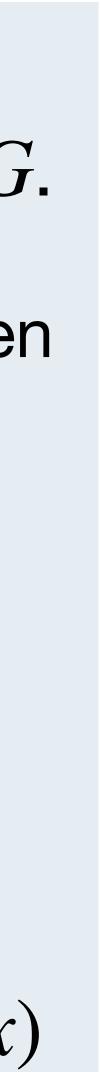
2. The kernel satisfies a symmetry constraint: $\forall_{h \in H}$: k(hx) = k(x)



Let $\mathscr{K}: \mathbb{L}_2(X) \to \mathbb{L}_2(Y)$ map between signals on homogeneous spaces of G.

Let homogeneous space $Y \equiv G/H$ such that $H = \operatorname{Stab}_G(y_0)$ for some chosen

$$\mathscr{K}f](y) = \int_X k(g_y^{-1}x)f(x)dx$$



Lecture notes Theorem 3.2:

origin $y_0 \in Y$ and let $g_v \in G$ such that $\forall_{v \in Y} : y = g_v y_0$.

Then \mathcal{K} is equivariant to group G if and only if:

1. It is a group convolution:

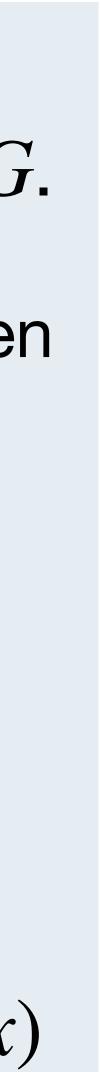
2. The kernel satisfies a symmetry constraint: $\forall_{h \in H}$: k(hx) = k(x)



Let $\mathscr{K}: \mathbb{L}_2(X) \to \mathbb{L}_2(Y)$ map between signals on homogeneous spaces of G.

Let homogeneous space $Y \equiv G/H$ such that $H = \operatorname{Stab}_G(y_0)$ for some chosen

$$\mathscr{K}f](y) = \int_X k(g_y^{-1}x)f(x)dx$$



Lecture notes **Theorem 3.2**: Let $\mathscr{K} : \mathbb{L}_2(X) \to \mathbb{L}_2(Y)$ map between

Let homogeneous space $Y \equiv G/H$ subscription $y_0 \in Y$ and let $g_y \in G$ such that

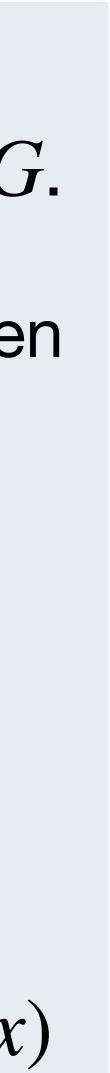
Then ${\mathscr K}$ is equivariant to group G if and only if:

1. It is a group convolution:

2. The kernel satisfies a symmetry constraint: $\forall_{h \in H}$: k(hx) = k(x)

where the
$$H = \operatorname{Stab}_{G}(y_0)$$
 for some choses
at $\forall_{y \in Y} : y = g_y y_0$.

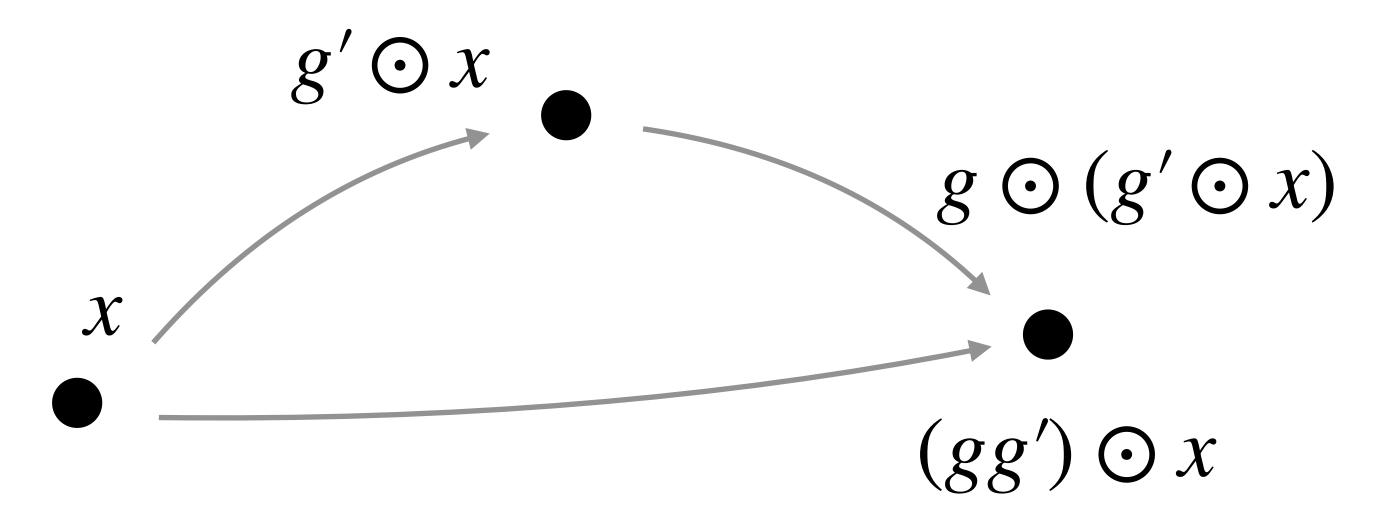
$$\mathscr{K}f](y) = \int_X k(g_y^{-1}x)f(x)dx$$



Group theory: Homogeneous spaces

Group action: An operator \odot : $G \times X \rightarrow X$ such that

$$\forall_{g,g'\in G,x\in X}: g \odot ($$



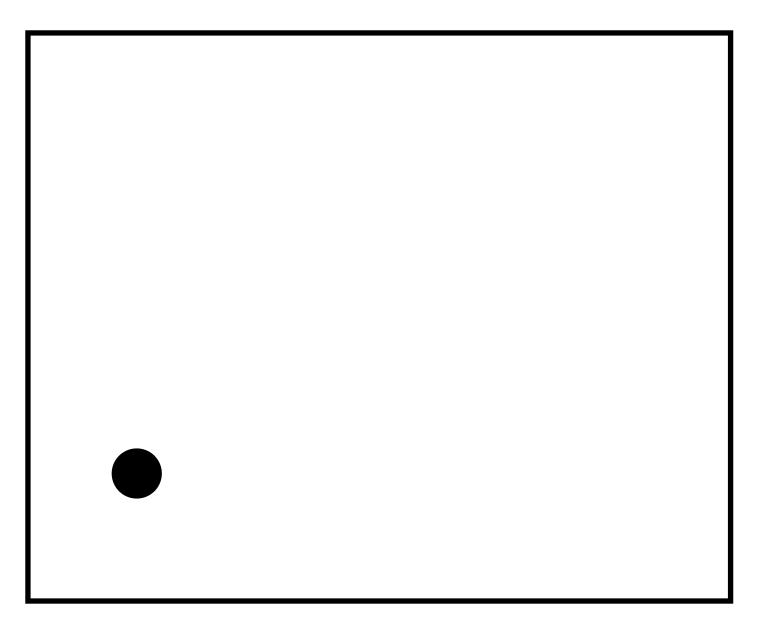
 $(g' \odot x) = (gg') \odot x$

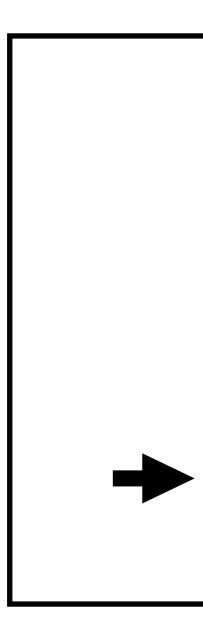


Group theory: Homogeneous spaces

$$\forall_{x_0, x \in X} \exists_{g \in G}$$

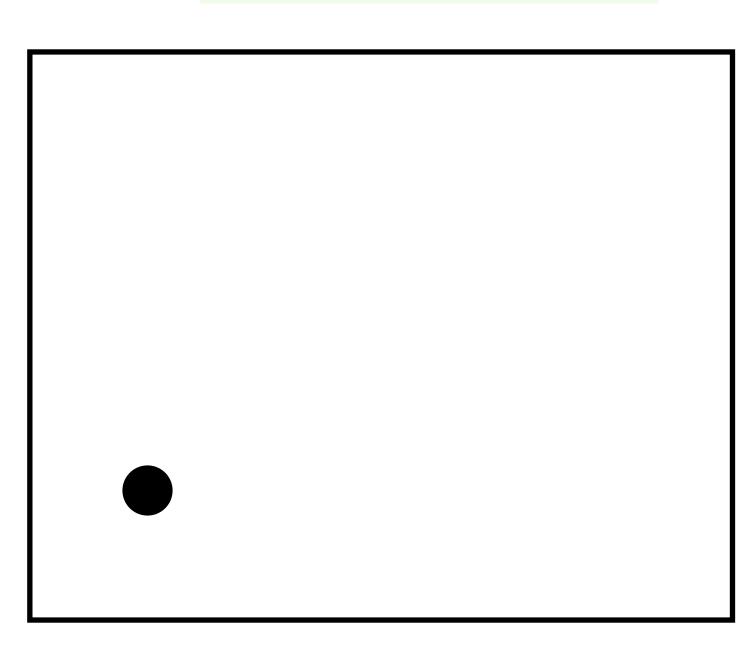






- **Transitive action**: An action \odot : $G \times X \rightarrow X$ such that
 - $x = g \odot x_0$
 - SE(2) acts transitively on \mathbb{R}^2

SO(2) does not ...

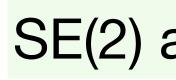


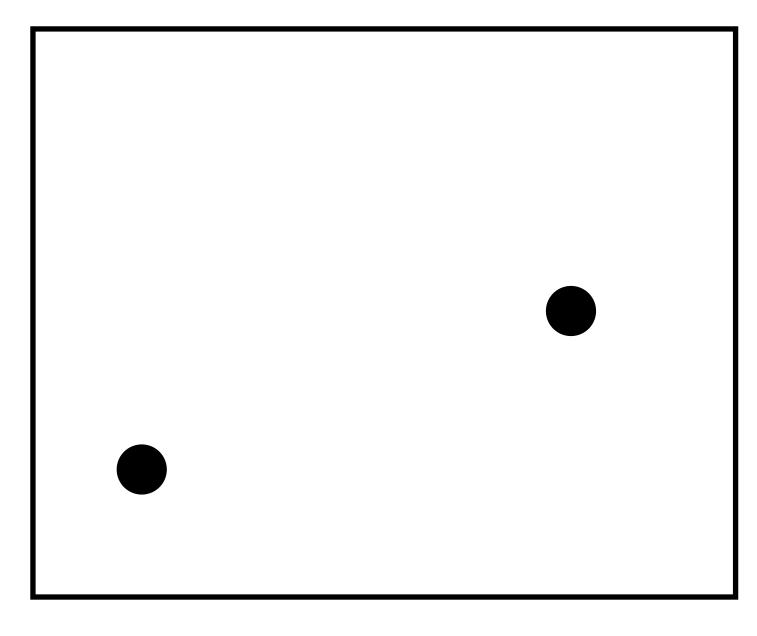


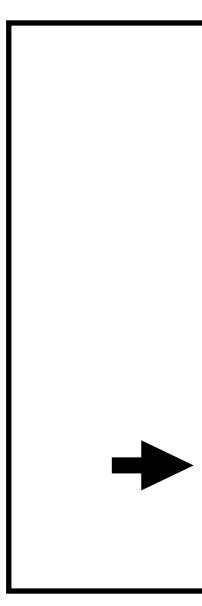


$$\forall_{x_0, x \in X} \exists_{g \in G}$$



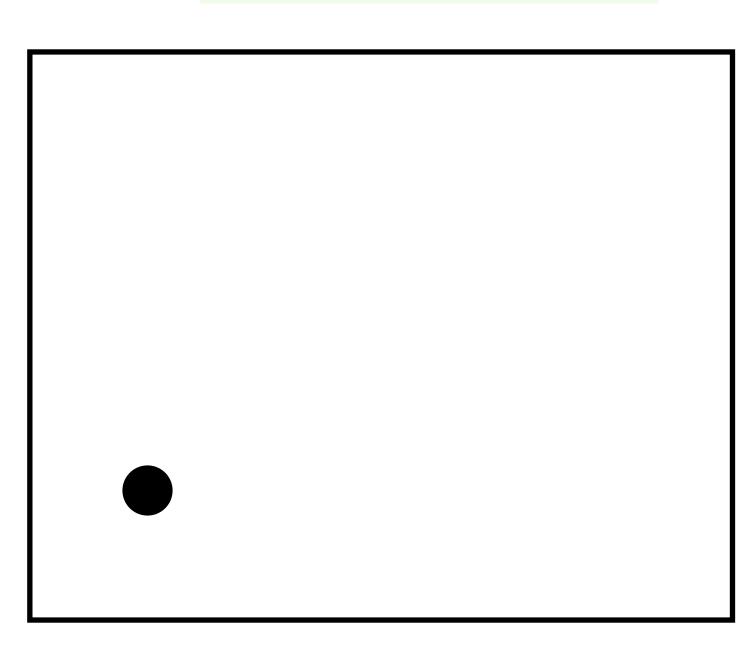




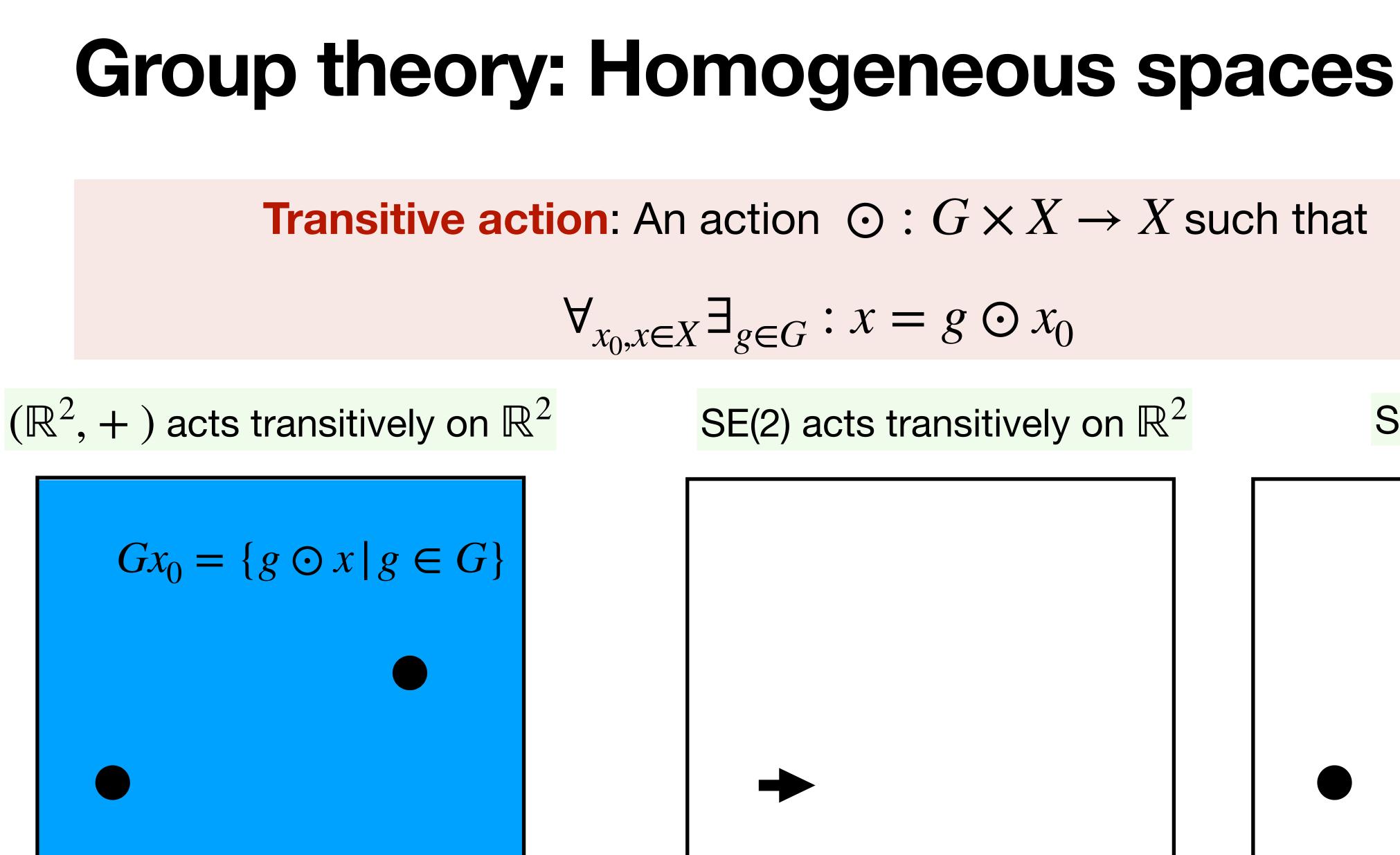


- **Transitive action**: An action \odot : $G \times X \rightarrow X$ such that
 - $x = g \odot x_0$
 - SE(2) acts transitively on \mathbb{R}^2

SO(2) does not ...

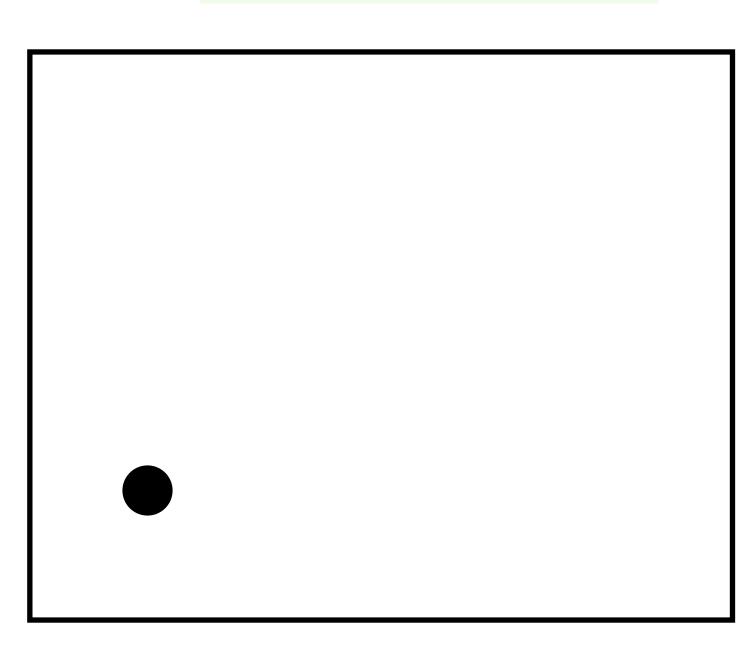




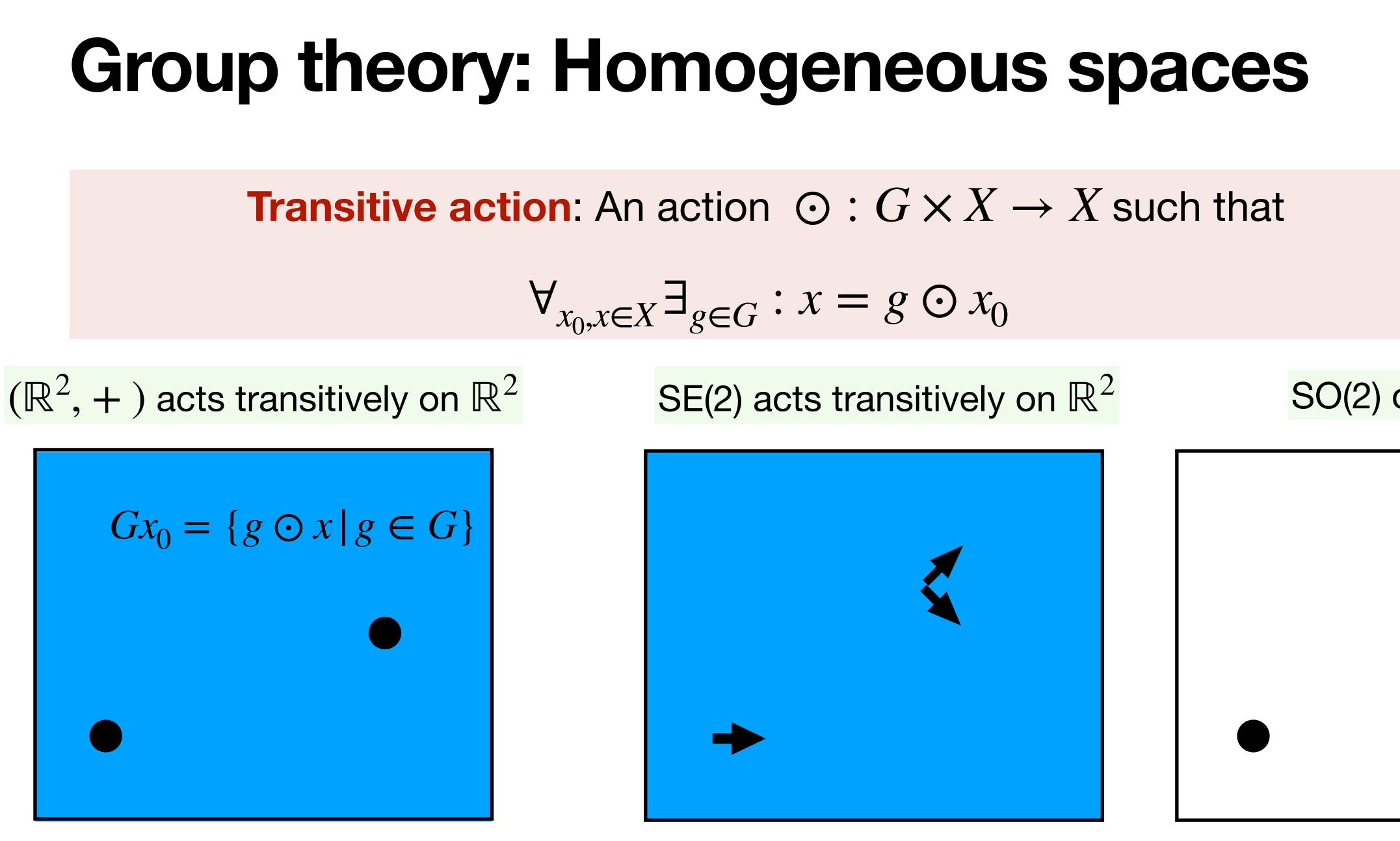


- **Transitive action**: An action \odot : $G \times X \rightarrow X$ such that
 - $\forall_{x_0, x \in X} \exists_{g \in G} : x = g \odot x_0$
 - SE(2) acts transitively on \mathbb{R}^2

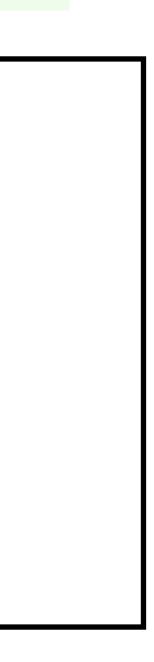
SO(2) does not ...



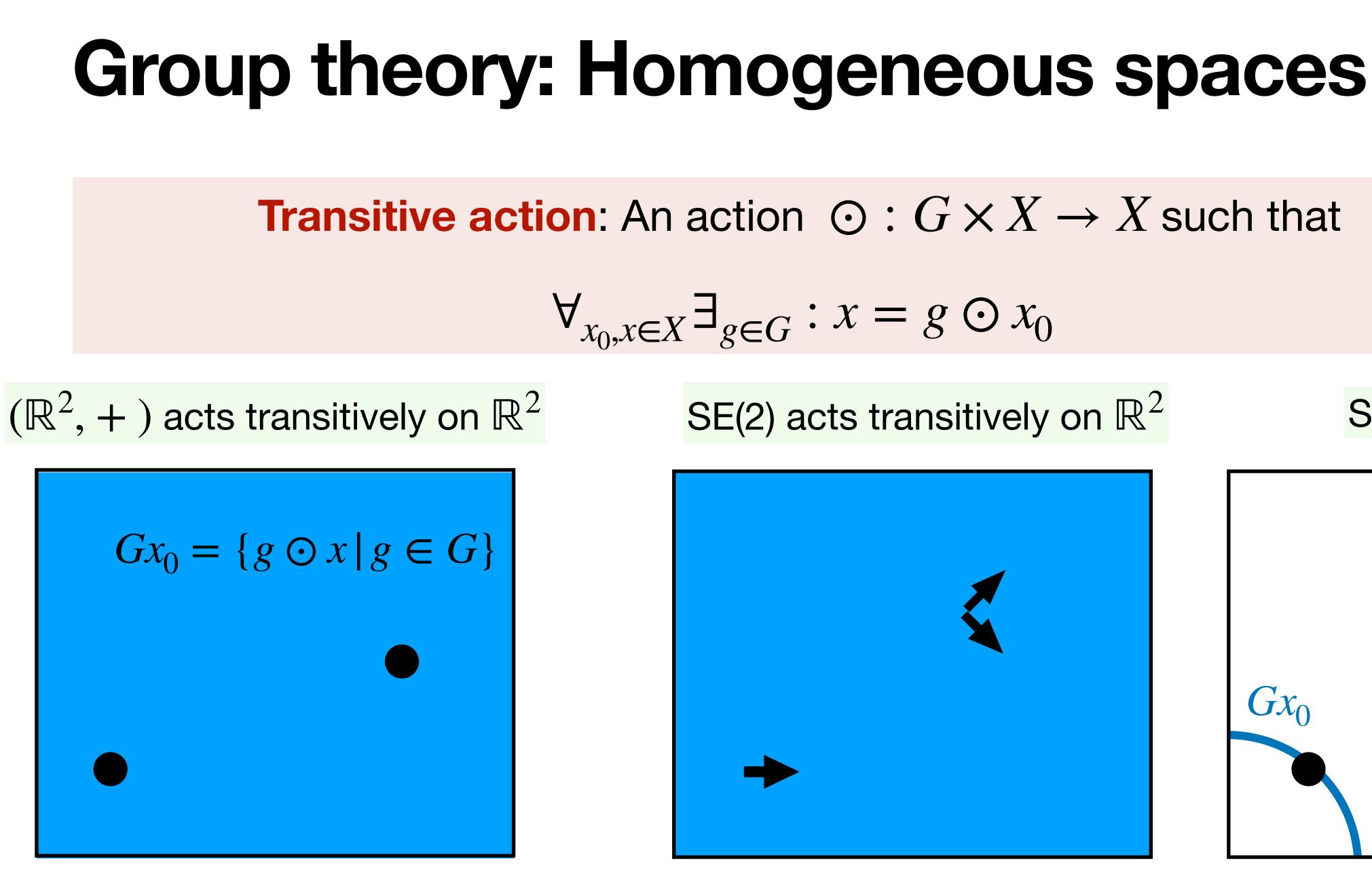




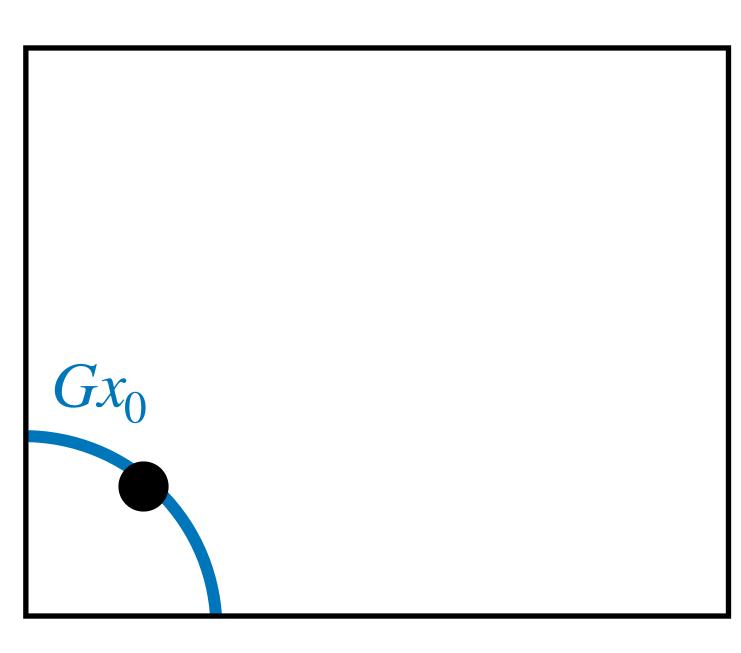




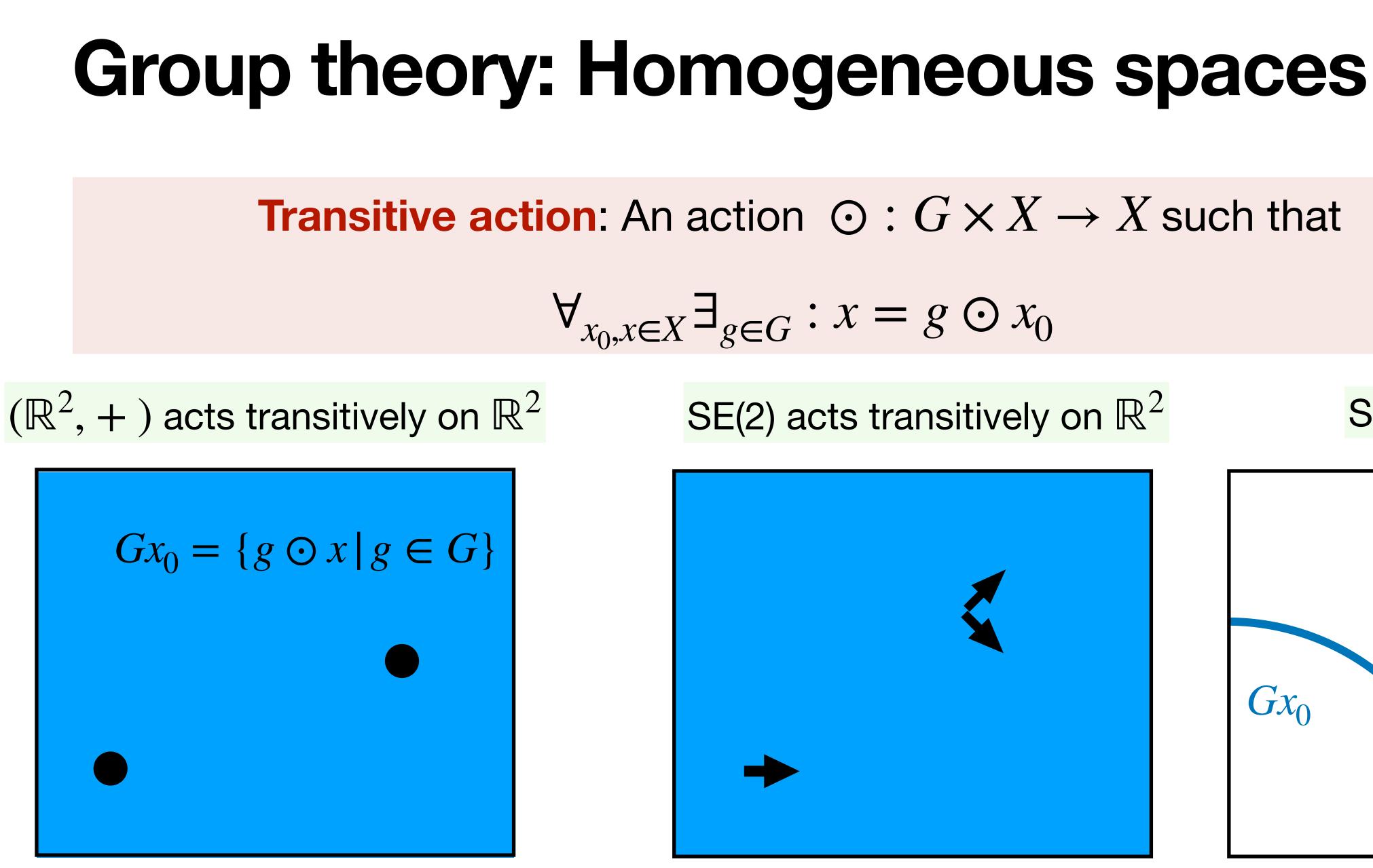




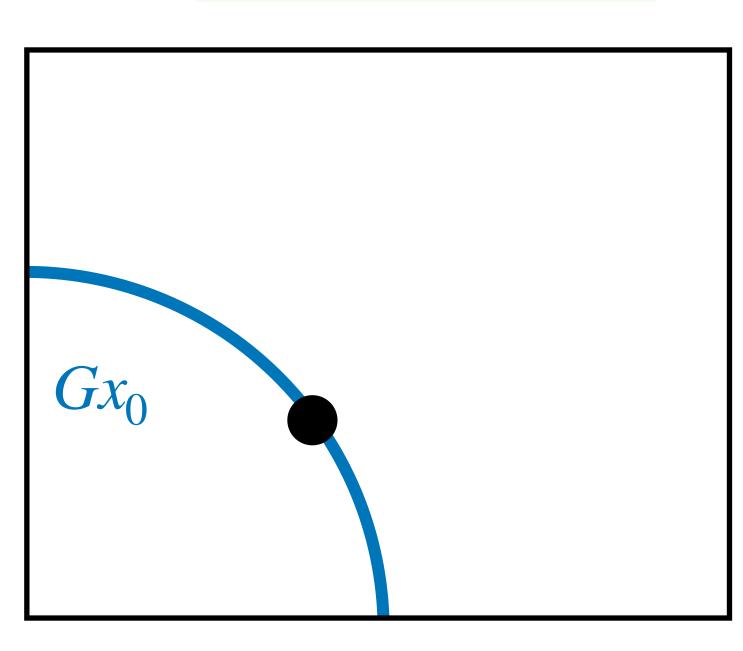




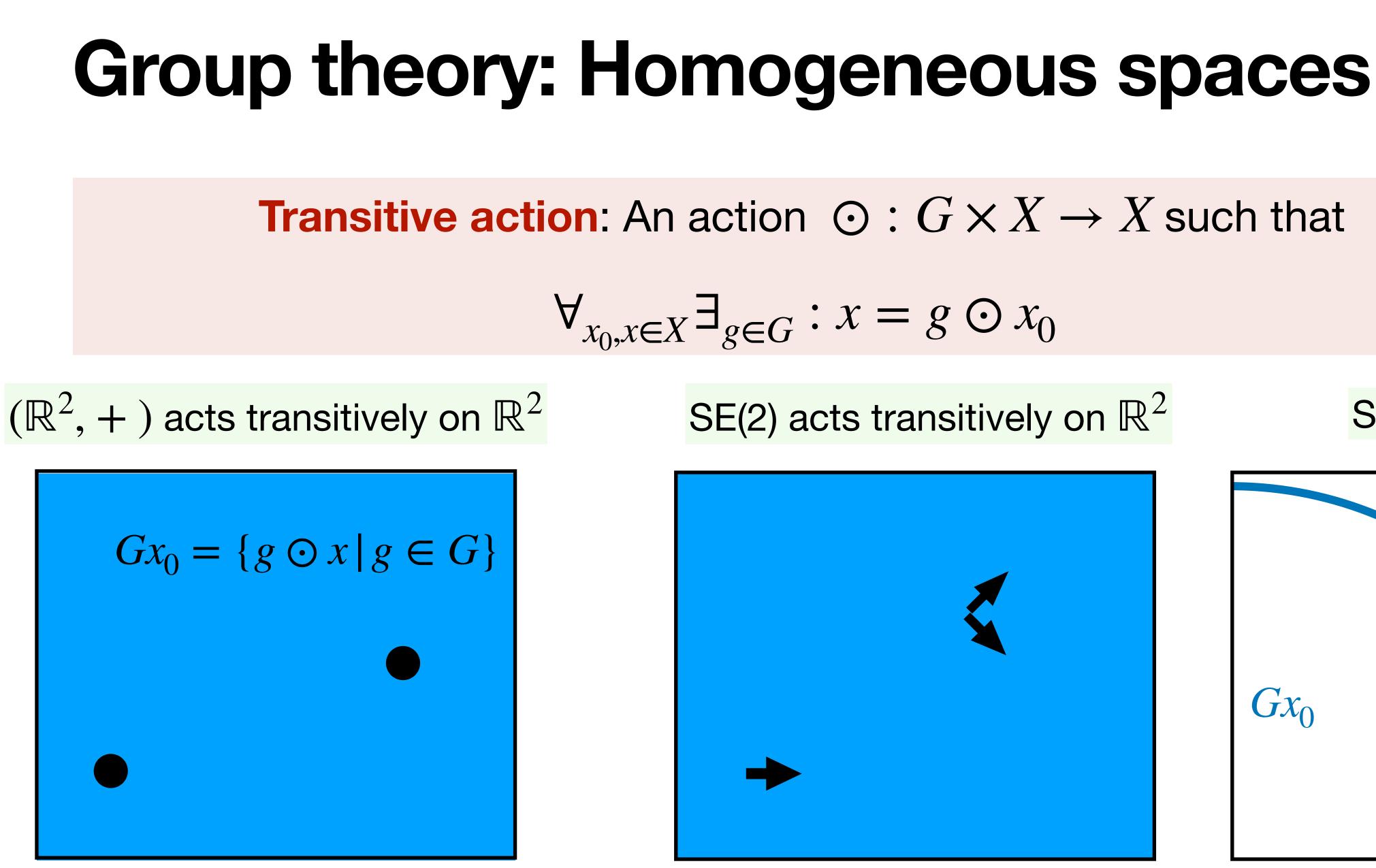




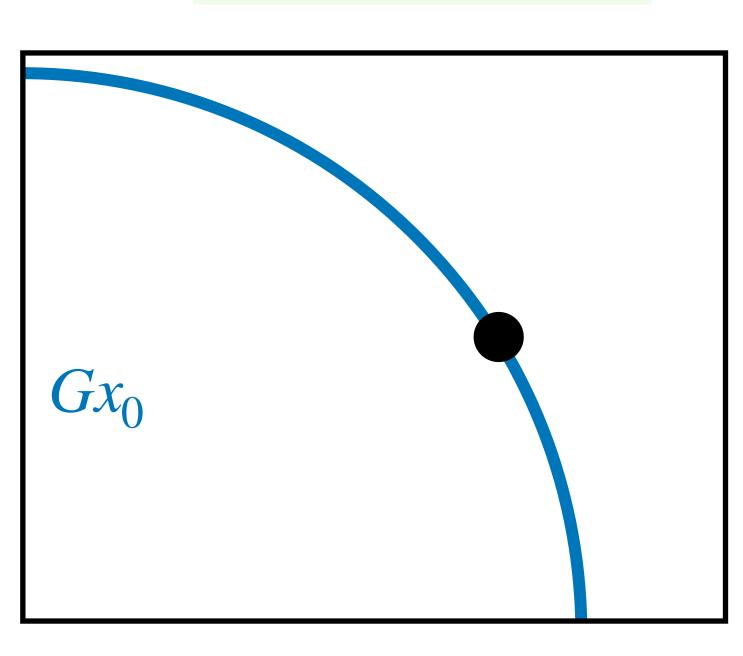








SO(2) does not ...





Group theory: Homogeneous spaces

Homogeneous space: A space on X on which G acts transitively.

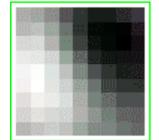
This is important as then we can guarantee that every part of the signal can be "seen" (probed by the convolution kernel)

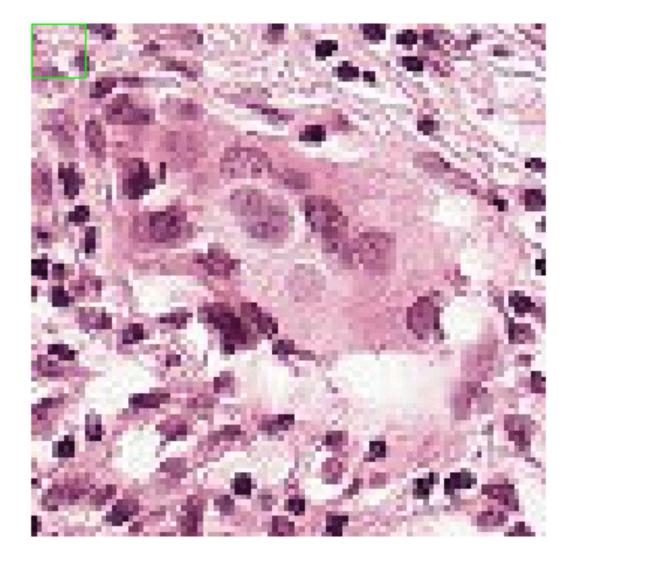


Group theory: Homogeneous spaces

Homogeneous space: A space on X on which G acts transitively.

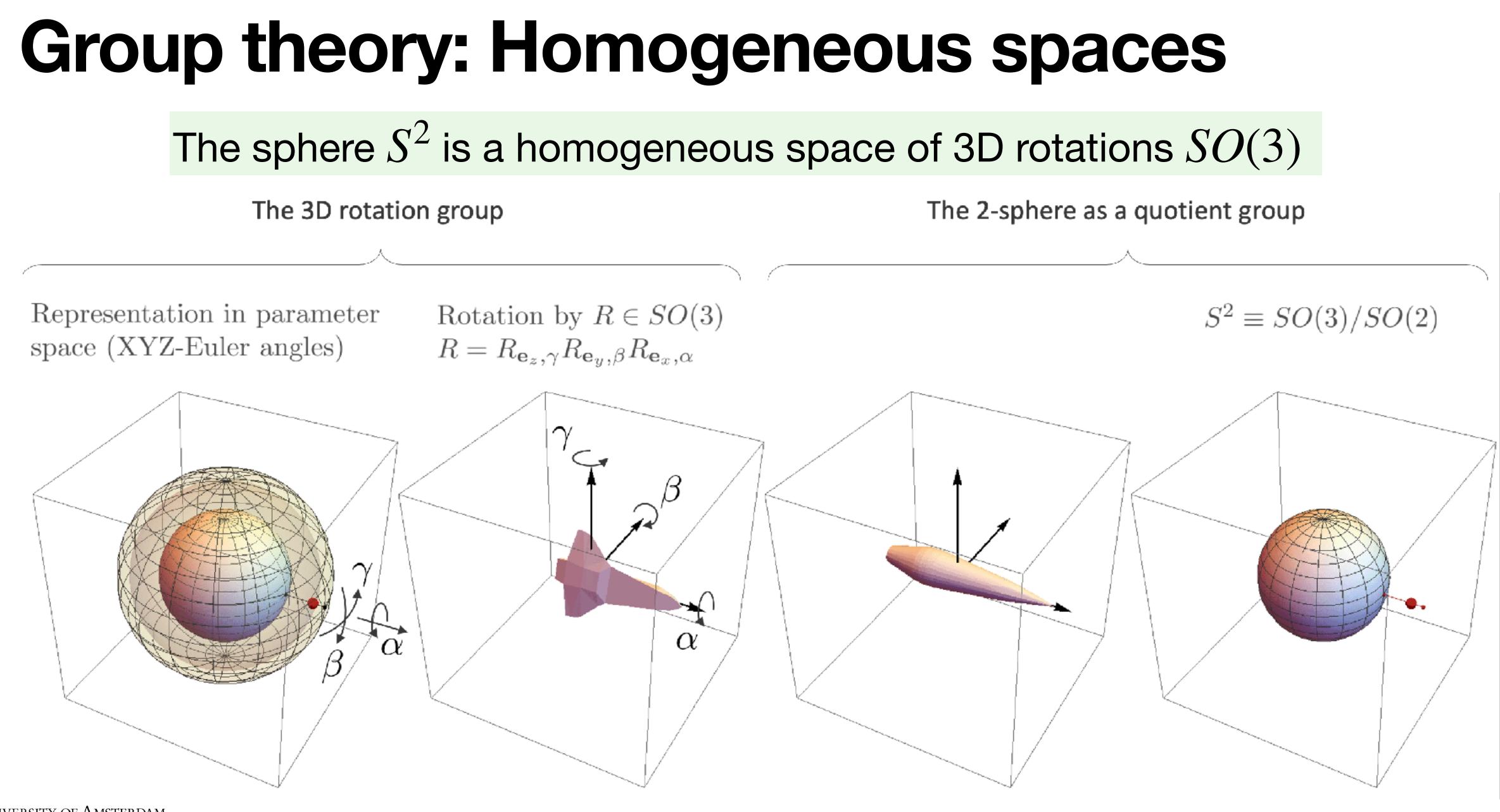
This is important as then we can guarantee that every part of the signal can be "seen" (probed by the convolution kernel)



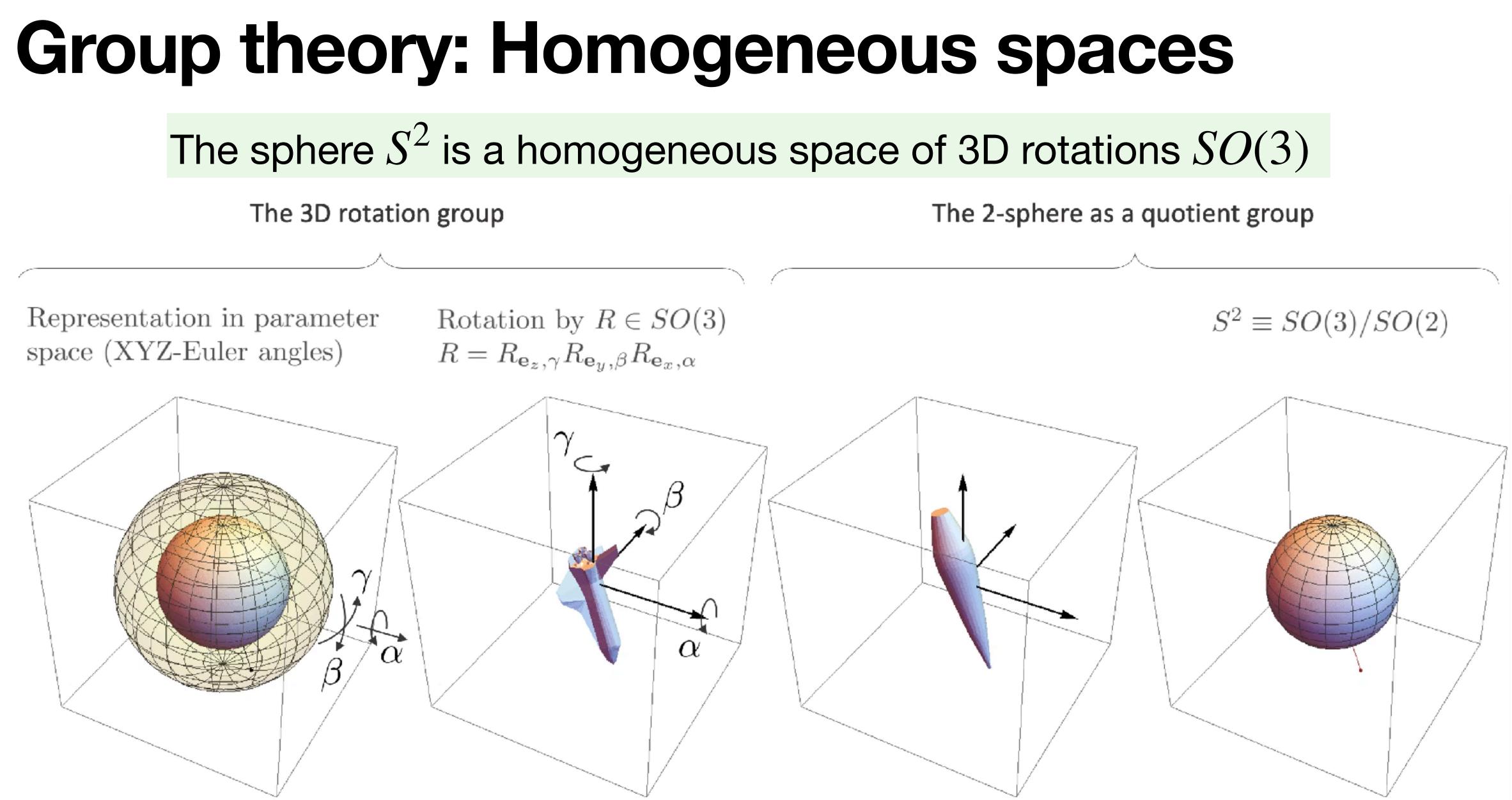








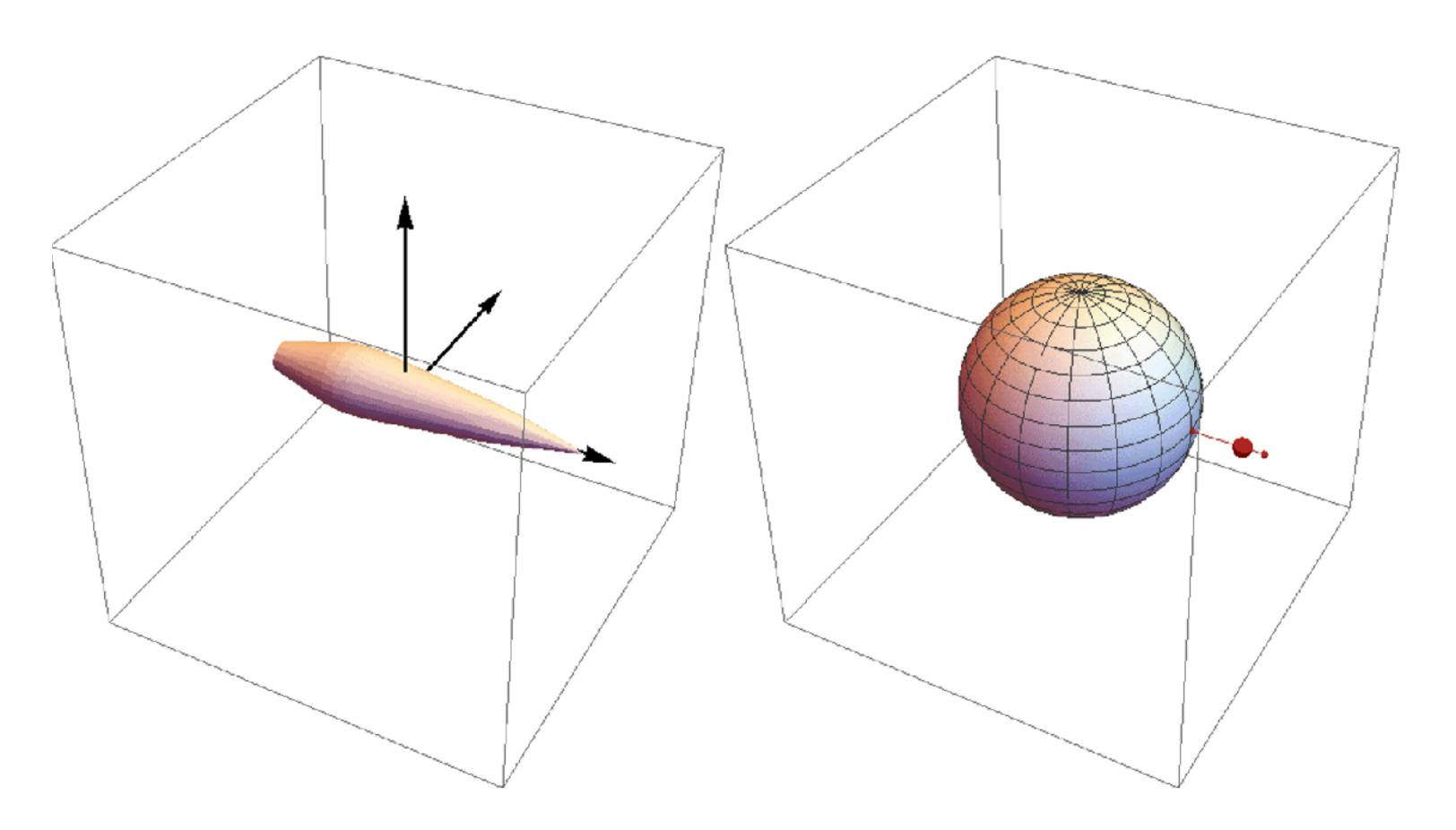




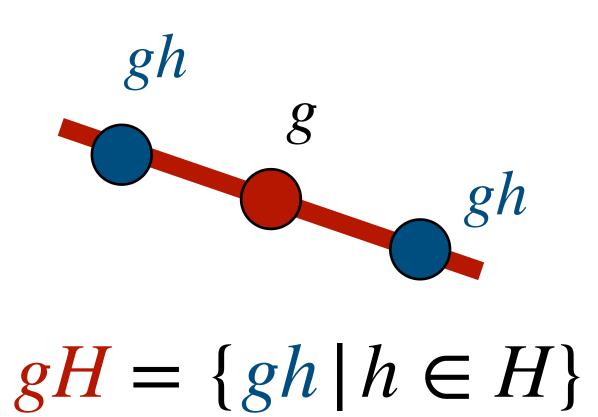


Group theory: Quotient spaces

of the space G/H are cosets.



Quotient space G/H: The space of unique cosets $gH = \{gh \mid h \in H\}$. Elements

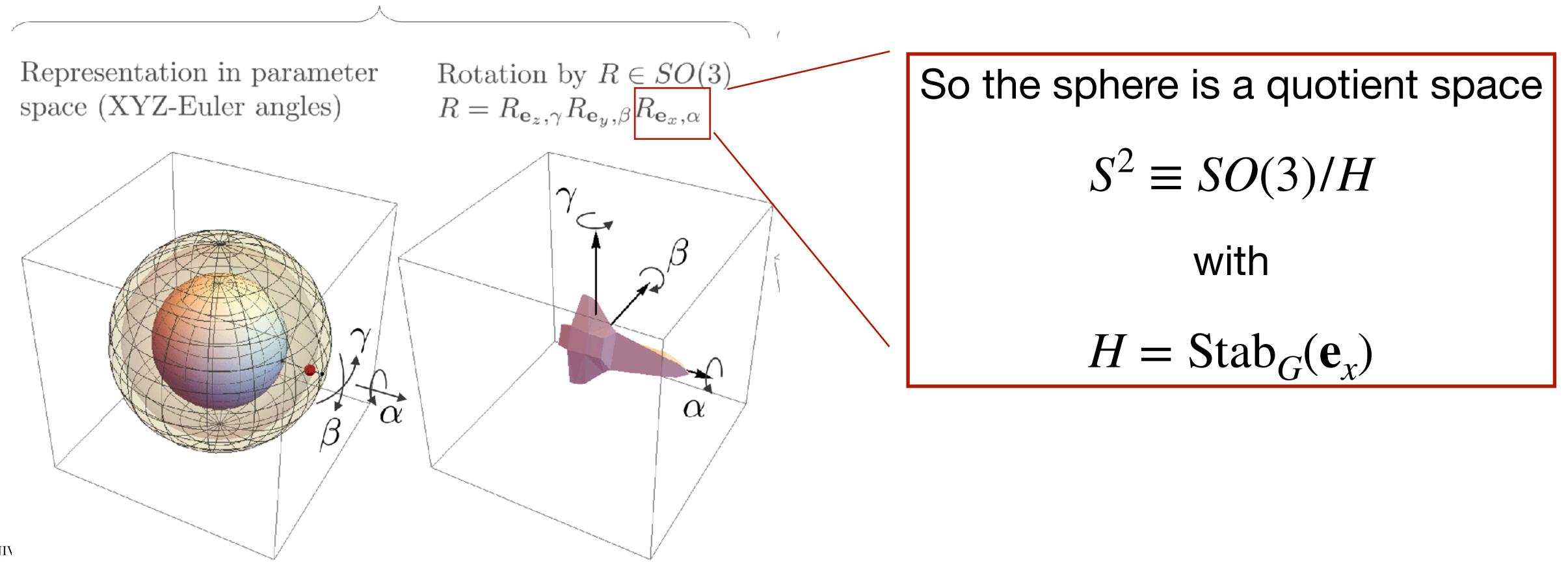




54

Group theory: Stabilizer **Stabilizer:** Stab_G(x_0) is a subset of G that leaves x_0 unchanged. I.e., $Stab_G(x_0) = \{ g \, | \, gx_0 = x_0 \}$

The 3D rotation group





Group theory: Homogeneous space \equiv Quotient space

Lemma 2.1: Any quotient space is a homogeneous space

Lemma 2.2: Any homogeneous space is a quotient space

UNIVERSITY OF AMSTERDAM



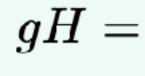


56

Group theory: Quotient spaces

Lecture notes Section 2.3

Example 2.7 (Quotient space $\mathbb{R}^d = SE(d)/SO(d)$). Let $H = (\{0\} \times SO(d)$ the subgroup of rotations in SE(d), with 0 the identity element of the translationg roup $(\mathbb{R}^d, +)$. The the cosets gH are given by



make the identification

We already saw in Exercise 2.1 that \mathbb{R}^d is a homogeneous space of SE(d), this is a consequence of Lemma 2.1.

 $gH = \{g \cdot (\mathbf{0}, \mathbf{\hat{R}}) \mid \mathbf{\hat{R}} \in SO(d)\}$ $= \{ (\mathbf{Re} + \mathbf{x}, \mathbf{R}\tilde{\mathbf{R}}) | h \in SO(d) \}$ $= \{ (\mathbf{x}, \mathbf{R}\tilde{\mathbf{R}}) | \ \tilde{\mathbf{R}} \in SO(d) \}$ $= \{ (\mathbf{x}, \tilde{\mathbf{R}}) | \ \tilde{\mathbf{R}} \in SO(d) \},\$

with $g = (\mathbf{x}, \mathbf{R})$. So, the cosets are given by all possible rotations for a fixed translation vector \mathbf{x} , the vector \mathbf{x} thus indexes these sets. We can therefore

 $\mathbb{R}^d \equiv SE(d)/SO(d)$.



Lecture notes Theorem 3.2:

origin $y_0 \in Y$ and let $g_v \in G$ such that $\forall_{v \in Y} : y = g_v y_0$.

Then \mathcal{K} is equivariant to group G if and only if:

1. It is a group convolution:

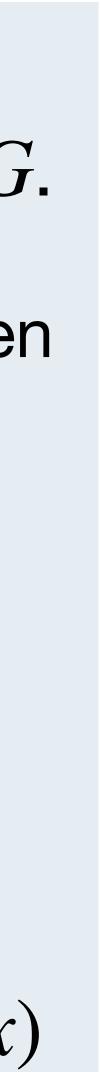
2. The kernel satisfies a symmetry constraint: $\forall_{h \in H}$: k(hx) = k(x)



Let $\mathscr{K}: \mathbb{L}_2(X) \to \mathbb{L}_2(Y)$ map between signals on homogeneous spaces of G.

Let homogeneous space $Y \equiv G/H$ such that $H = \operatorname{Stab}_G(y_0)$ for some chosen

$$\mathscr{K}f](y) = \int_X k(g_y^{-1}x)f(x)dx$$



Theorem: Let $\mathscr{K} : \mathbb{L}_2(X) \to \mathbb{L}_2(Y)$ map between signals on homogeneous spaces of G.

origin $y_0 \in Y$ and let $g_v \in G$ such that $\forall_{v \in Y} : y = g_v y_0$.

Then \mathcal{K} is equivariant to group G if and only if:

1. It is a group convolution:

2. The kernel satisfies a symmetry constraint: $\forall_{h \in H}$: k(hx) = k(x)

Let homogeneous space $Y \equiv G/H$ such that $H = \operatorname{Stab}_G(y_0)$ for some chosen

$$\mathscr{K}f](y) = \int_X k(g_y^{-1}x)f(x)dx$$





Theorem: Let $\mathscr{K} : \mathbb{L}_2(X) \to \mathbb{L}_2(Y)$ map betwee Let homogeneous space $Y \equiv G/H$ su origin $y_0 \in Y$ and let $g_y \in G$ such that

Then \mathcal{K} is equivariant to group G if and only if:

1. It is a group convolution:

2. The kernel satisfies a symmetry constraint: $\forall_{h \in H}$: k(hx) = k(x)



en signals on homogeneous spaces of
$$G$$

uch that $H = \operatorname{Stab}_G(y_0)$ for some chose
at $\forall_{y \in Y} : y = g_y y_0$.

$$\mathscr{K}f](y) = \int_X k(g_y^{-1}x)f(x)dx$$





Types of layers

(X = Y = G/H)

(X = G/H, Y = G)

(X = Y = G)

(X = G, Y = G/H)**Projection layer**. Mean pooling over *H*.

 $(X = G, Y = \emptyset)$ **Global pooling** over

Isotropic/Constraint convolutions on spaces of lower dimension than G, $\forall_{h \in H}$: k(hx) = k(x)

Lifting convolution. No constraints on k.

Group convolutions. No constraints on k.

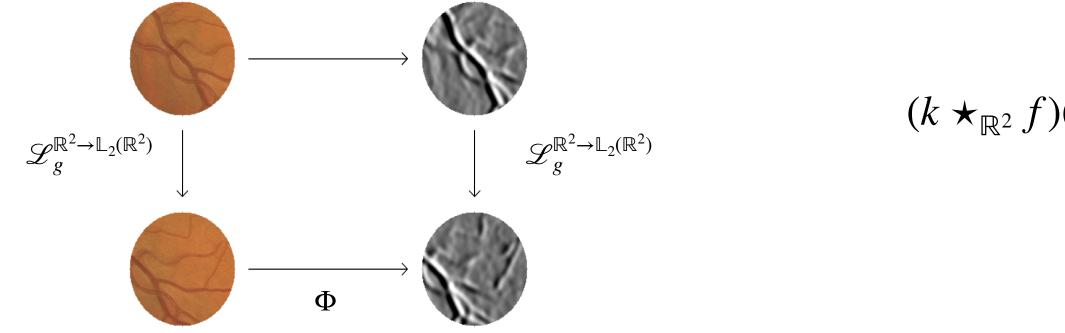


The case of SE(2) equivariant layers for signals on $\mathbb{R}^d \equiv SE(2)/SO(2)$

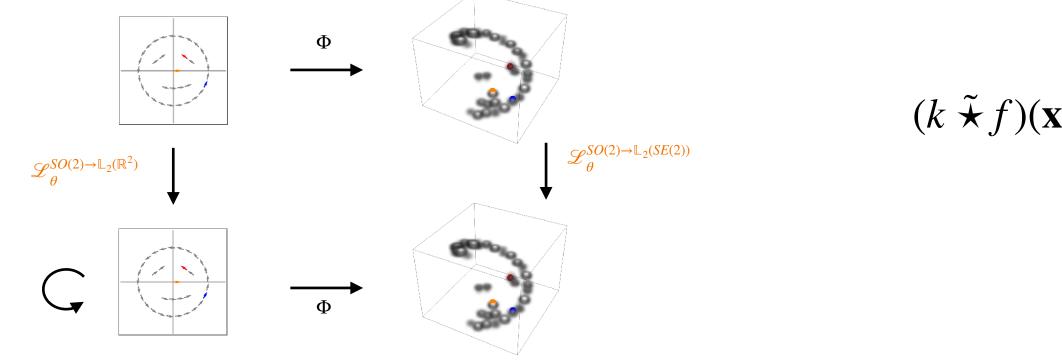




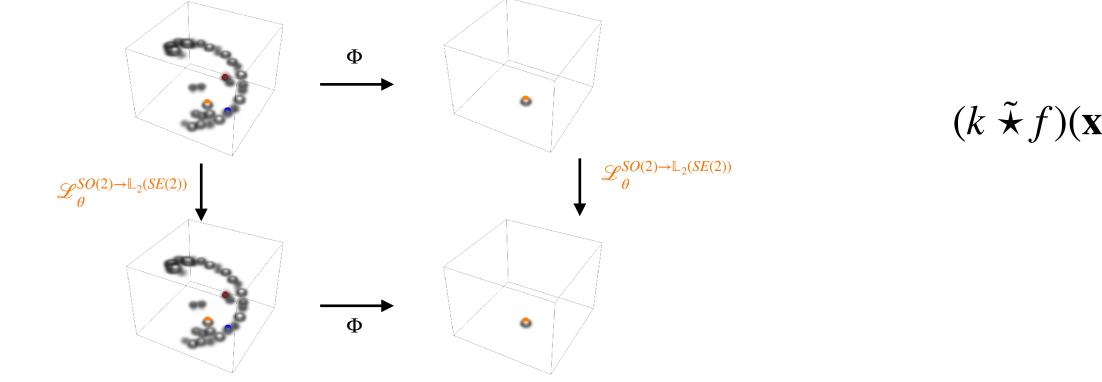
2D cross-correlation (translation equivariant) - $K : \mathbb{L}_2(\mathbb{R}^2) \rightarrow$



SE(2) lifting correlations - $K : \mathbb{L}_2(\mathbb{R}^2) \to \mathbb{L}_2(SE(2))$



SE(2) G-correlations – $K : \mathbb{L}_2(SE(2)) \rightarrow \mathbb{L}_2(SE(2))$



$$\rightarrow \mathbb{L}_2(\mathbb{R}^2)$$

$$f(\mathbf{x}) = (\mathscr{L}_{\mathbf{x}}^{\mathbb{R}^2 \to \mathbb{L}_2(\mathbb{R}^2)} k, f)_{\mathbb{L}_2(\mathbb{R}^2)}$$
$$= \int_{\mathbb{R}^2} k(\mathbf{x}' - \mathbf{x}) f(\mathbf{x}') d\mathbf{x}'$$

SE(2) equivariance iff

$$(\mathscr{L}_{\theta}^{SO(2) \to \mathbb{L}_{2}(\mathbb{R}^{2})} k)(\mathbf{x}) = I$$

$$\Leftrightarrow$$

$$k(\mathbf{R}_{\theta}^{-1}\mathbf{x}) = k(\mathbf{x})$$
since $Y = \mathbb{R}^{2} \equiv SE(2)/S$

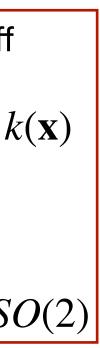
$$\mathbf{x} = (\mathscr{L}_g^{SE(2) \to \mathbb{L}_2(\mathbb{R}^2)} k, f)_{\mathbb{L}_2(\mathbb{R}^2)}$$

$$= \int_{\mathbb{R}^2} k(\mathbf{R}_{\theta}^{-1}(\mathbf{x}' - \mathbf{x})) f(\mathbf{x}') d\mathbf{x}$$

$$\mathbf{x} = (\mathscr{L}_g^{SE(2) \to \mathbb{L}_2(SE(2))} k, f)_{\mathbb{L}_2(SE(2))}$$

No constraints

$$= \int_{\mathbb{R}^2} \int_{S^1} k(\mathbf{R}_{\theta}^{-1}(\mathbf{x}' - \mathbf{x}), \theta' - \theta \mod 2\pi) f(\mathbf{x}', \theta') d\mathbf{x}'$$



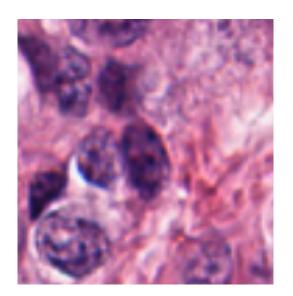


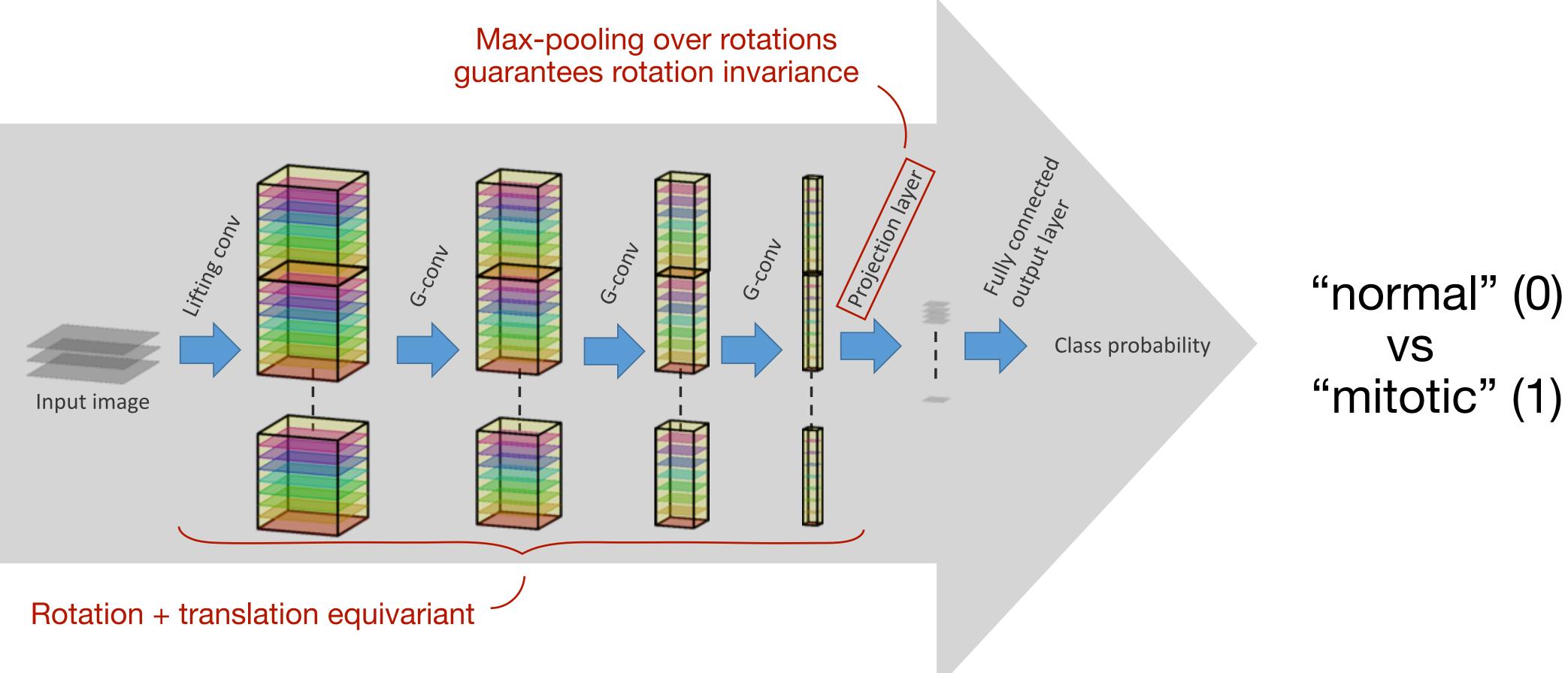


The most expressive group equivariant architectures are obtained by lifting the feature maps to the group



General group equivariant architecture







Content

Part I: Introduction to group convolutions

- * Motivation
- * Introduction to group theory
- * Regular group convolutional neural networks
- * Applications

Part II: General theory for group equivariant deep learning

- * Group convolutions are all you need!
- * Deeper into group theory: representation theory, homogeneous spaces
- * Characterization of types of group equivariant layers

Part III: Steerable group convolutions

* Deep dive into group theory: irreducible representations, steerable operators and vector spaces * Examples of steerable group convolutions: Spherical data and Volumetric data/3D point clouds

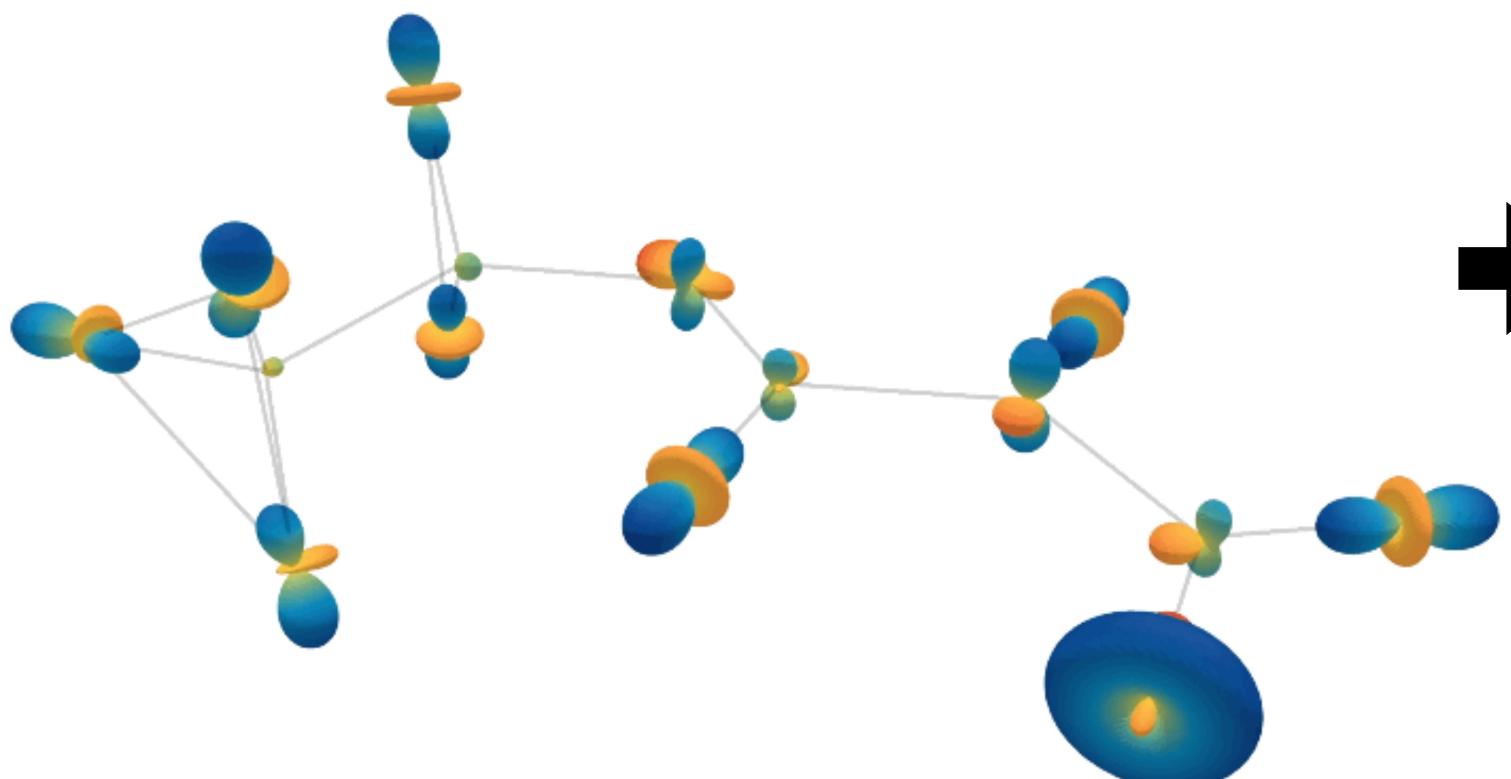


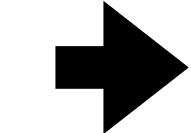
The need for steerable G-CNNs

- Steerable methods are designed for groups that involve the action of SO(d): • Are based on a Fourier convolution theorem on SO(d)
 - Avoids discretization of SO(d):
 - Numerically more precise than regular group convolutions
 - Exact equivariance lacksquare
 - Flexible to non-gridded data
 - Provide a roadmap to local equivariance on arbitrary manifolds through Gauge theory



Brandstetter, Hesselink, van der Pol, Bekkers, Welling Steerable **Equivariant Message Passing on Molecular Graphs**



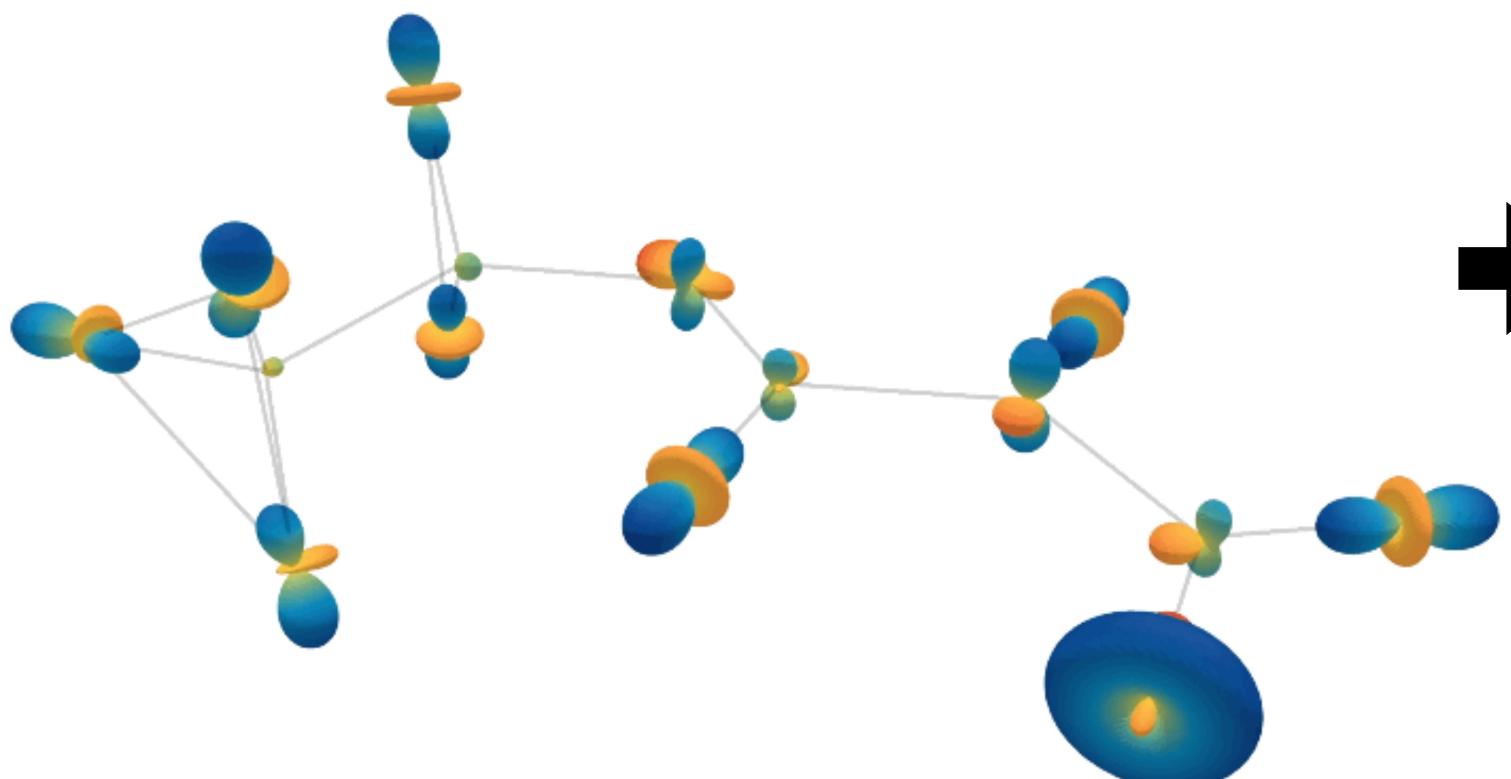


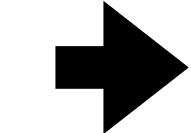
Molecular property prediction





Brandstetter, Hesselink, van der Pol, Bekkers, Welling Steerable **Equivariant Message Passing on Molecular Graphs**



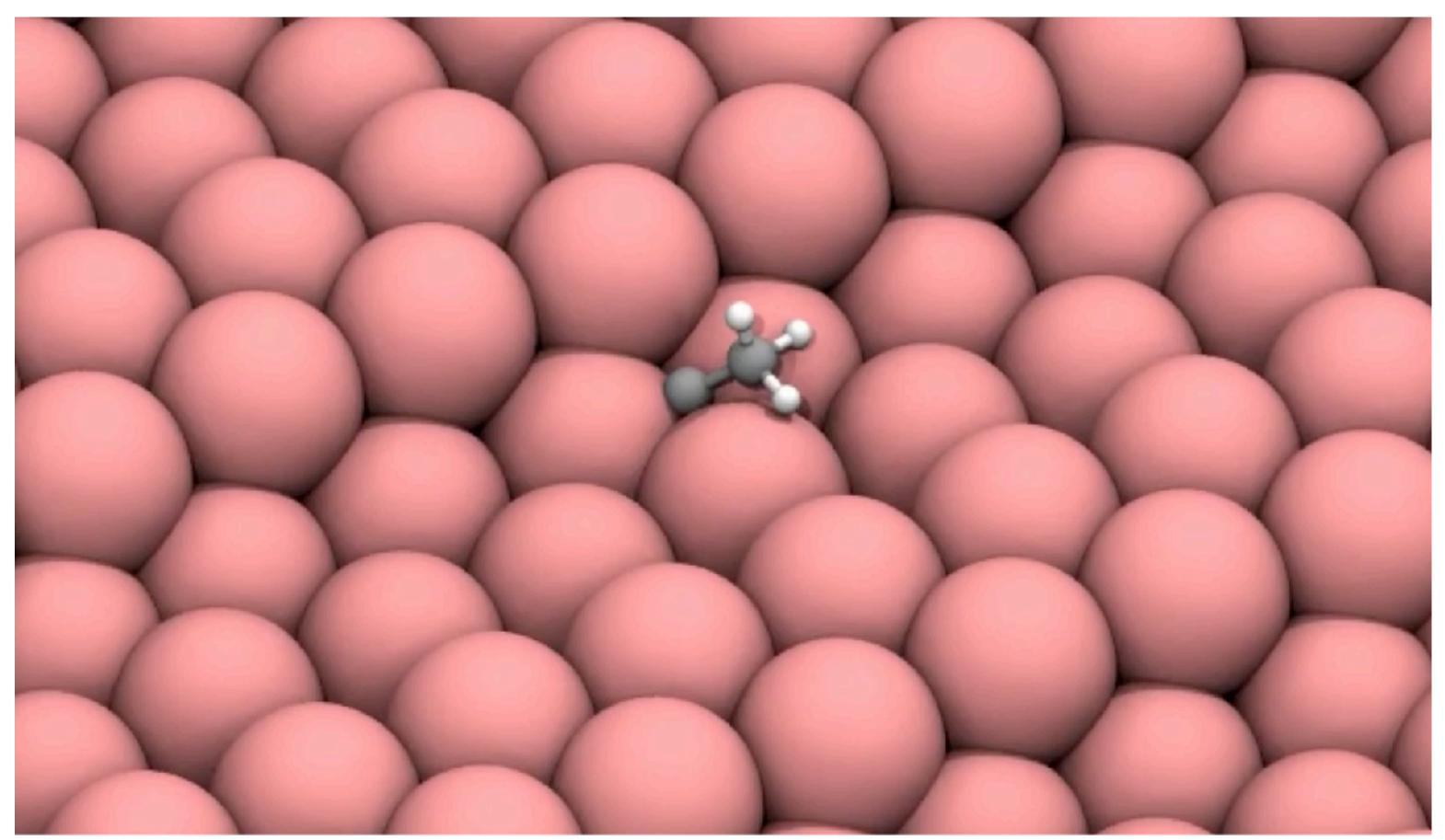


Molecular property prediction





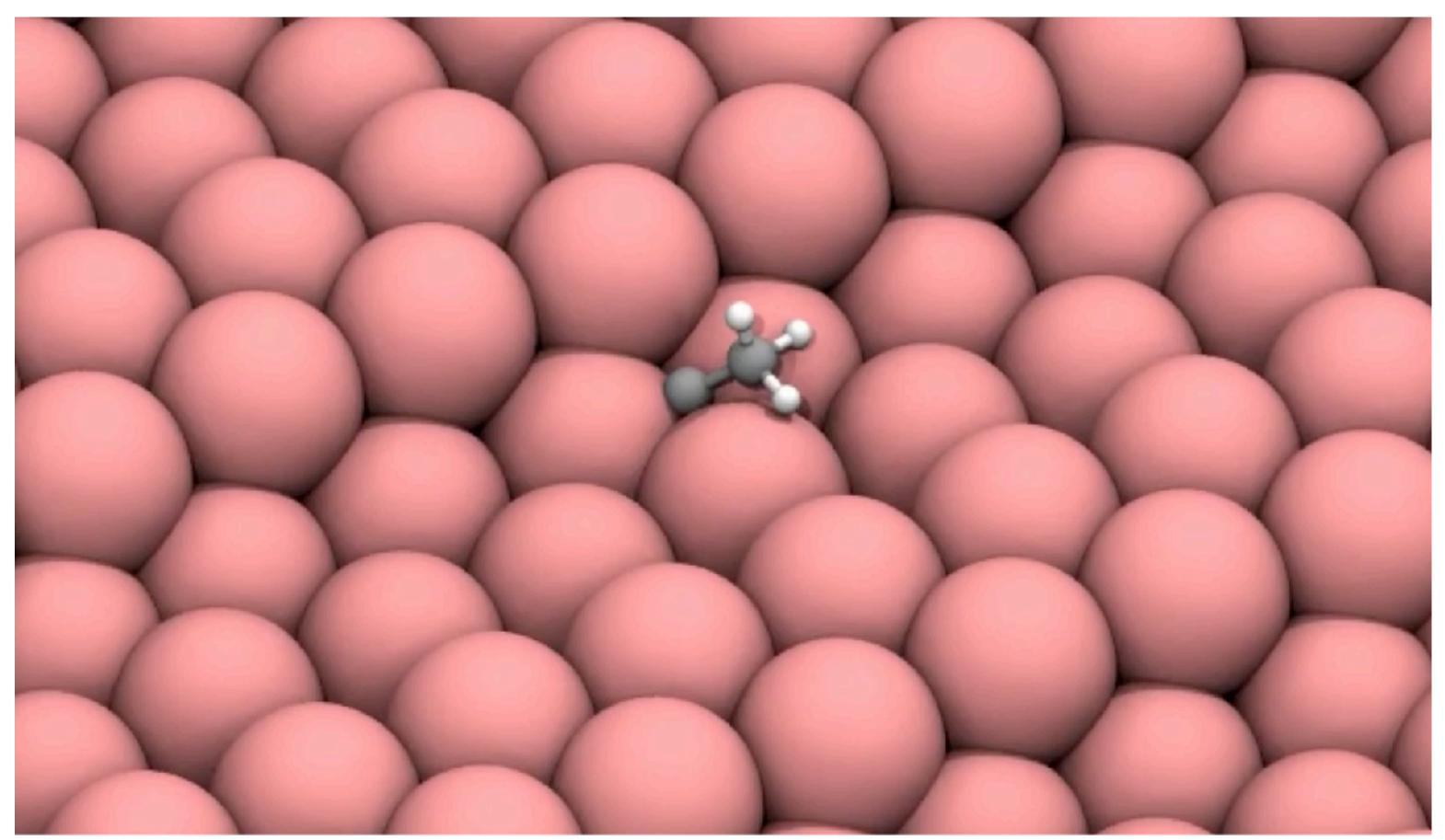
Brandstetter, Hesselink, van der Pol, Bekkers, Welling Steerable Equivariant Message Passing on Molecular Graphs



Video: Open Catalyst Project



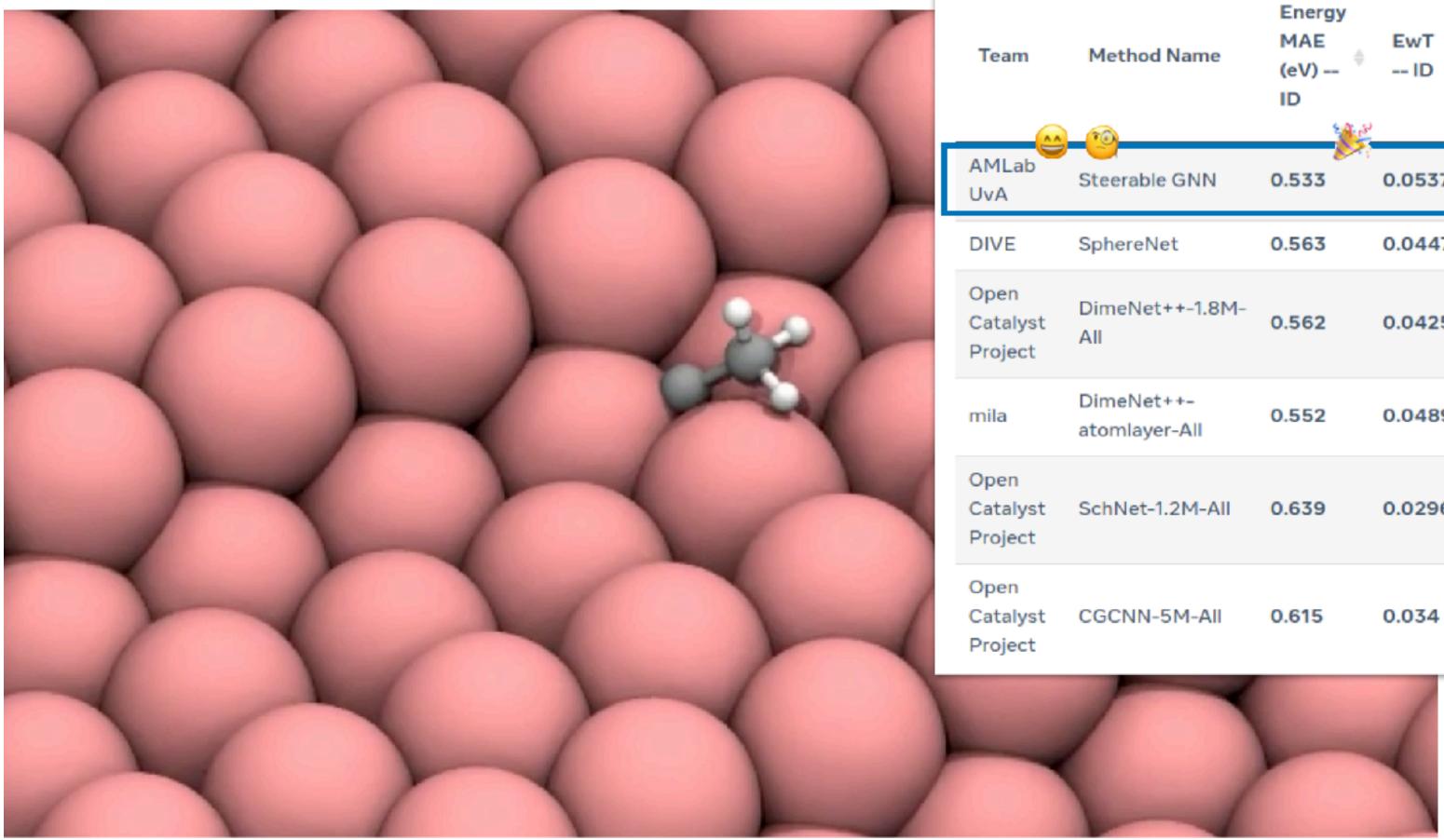
Brandstetter, Hesselink, van der Pol, Bekkers, Welling Steerable Equivariant Message Passing on Molecular Graphs



Video: Open Catalyst Project

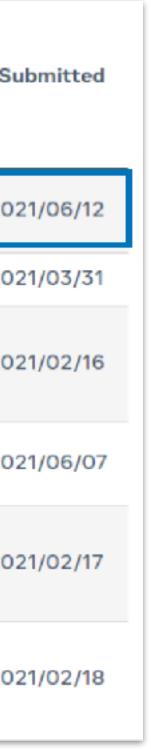


Brandstetter, Hesselink, van der Pol, Bekkers, Welling Steerable **Equivariant Message Passing on Molecular Graphs**



Video: Open Catalyst Project

eam	Method Name	Energy MAE (eV) * ID	EwT ID 🔶	Energy MAE (eV) 🔶 OOD Ads	EwT OOD ^{\$} Ads	Energy MAE (eV) 🔶 OOD Cat	EwT OOD ^{\$} Cat	Energy MAE (eV) 🔅 OOD Both	EwT OOD Both	Sı
1Lab A	Steerable GNN	0.533	0.0537	0.692	0.0246	0.537	0.0492	0.679	0.0263	20:
/E	SphereNet	0.563	0.0447	0.703	0.0229	0.571	0.0409	0.638	0.0241	20
en talyst oject	DimeNet++-1.8M- All	0.562	0.0425	0.725	0.0207	0.576	0.041	0.661	0.0241	202
а	DimeNet++- atomlayer-All	0.552	0.0489	0.747	0.0259	0.557	0.0459	0.688	0.0233	202
en talyst oject	SchNet-1.2M-All	0.639	0.0296	0.734	0.0233	0.662	0.0294	0.704	0.0221	20
en talyst oject	CGCNN-5M-All	0.615	0.034	0.915	0.0193	0.622	0.031	0.851	0.02	203





Group theoretical background

Irreducible Representations (spherical harmonics, Wigner-D matrices)

Fourier transform on SO(3)

UNIVERSITY OF AMSTERDAM



Convolution theorem + Clebsch-Gordan Tensor product

Steerable G-CNNs





Group theory: Irreducible Representations

similarity transform

in which **Q** carries out the change of basis.

Definition 4.1 (Equivalence of matrix representations). Any two matrix representations $\mathbf{D}(g)$ and $\mathbf{D}'(g)$ of a group G are equivalent if they relate via a

- $\mathbf{D}'(g) = \mathbf{Q}^{-1}\mathbf{D}(g)\mathbf{Q} ,$



Group theory: Irreducible Representations

similarity transform

in which **Q** carries out the change of basis.

sentation is called reducible if it can be written as

 $\mathbf{D}(g) = \mathbf{Q}^{-1}(\mathbf{D}_1(g) \oplus \mathbf{D}_2(g))\mathbf{Q}$

reducible they are called irreducible representations (irreps).



- **Definition 4.1** (Equivalence of matrix representations). Any two matrix representations $\mathbf{D}(g)$ and $\mathbf{D}'(g)$ of a group G are equivalent if they relate via a
 - $\mathbf{D}'(q) = \mathbf{Q}^{-1}\mathbf{D}(q)\mathbf{Q} ,$
- **Definition 4.2** (Reducible/irreducible matrix representation). A matrix repre-

$$\mathbf{Q} = \mathbf{Q}^{-1} \begin{pmatrix} \mathbf{D}_1(g) & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_2(g) \end{pmatrix} \mathbf{Q} ,$$

in which \mathbf{Q} carries out a change of basis. If the matrices \mathbf{D}_1 and \mathbf{D}_2 are not



Group theory: Wigner-D Matrices

Wigner-D matrices are the irreducible matrix representations of SO(3)

Every representation D(g) of SO(3) is block diagonalizable to a representation with Wigner-D matrices along the diagonal:

 $\mathbf{D}(g) = \mathbf{Q}^{-1}(\mathbf{D}^{(l_1)}(g) \oplus \mathbf{D}^{(l_2)}(g) \oplus$

$$(\exists \dots) \mathbf{Q} = \mathbf{Q}^{-1} \begin{pmatrix} \mathbf{D}^{(l_1)}(g) & \mathbf{D}^{(l_2)}(g) \\ & \mathbf{D}^{(l_2)}(g) \end{pmatrix}.$$





Group theory: Steerable vector space

Definition 4.3 (Wigner-D matrix). The Wigner-D matrices of type-l are the irreducible (2l+1) matrix representations of SO(3). A Wigner-D matrix of type l as a function of $q \in G$ will be denoted with $\mathbf{D}^{(l)}(q)$.

Wigner-D matrices generalize the notion of a rotation matrix for the rotation of (2l + 1)-dimensional vectors





Group theory: Steerable vector space

Definition 4.3 (Wigner-D matrix). The Wigner-D matrices of type-*l* are the irreducible (2l+1) matrix representations of SO(3). A Wigner-D matrix of type *l* as a function of $g \in G$ will be denoted with $\mathbf{D}^{(l)}(g)$.

Definition 4.4 (Wigner-D functions). The $(2l+1) \times (2l+1)$ components of the type-*l* Wigner-D matrices will be referred to as the type-*l* Wigner-D functions. The Wigner-D functions are denoted with $D_{mn}^{(l)}$ with *m* and *n* row and column index respectively.



Group theory: Steerable vector space

l as a function of $g \in G$ will be denoted with $\mathbf{D}^{(l)}(g)$.

index respectively.

dimensional vector $\mathbf{v} \in V_l$ will be called a type-l vector.



Definition 4.3 (Wigner-D matrix). The Wigner-D matrices of type-l are the irreducible (2l+1) matrix representations of SO(3). A Wigner-D matrix of type

Definition 4.4 (Wigner-D functions). The $(2l+1) \times (2l+1)$ components of the type-*l* Wigner-D matrices will be referred to as the type-*l* Wigner-D functions. The Wigner-D functions are denoted with $D_{mn}^{(l)}$ with m and n row and column

Definition 4.5 (Steerable vector spaces and steerable vectors). The (2l + 1)dimensional vector space on which a Wigner-D matrix of order l acts will be called a type l steerable vector space and is denoted with V_l . A (2l + 1)-



Group theory: Spherical Harmonics

- •Functions on the sphere
- •Solutions to Laplace's equation on S^2
- •The S^2 equivalent of the circular harmonics (1D Fourier basis)
- •Form orthonormal basis for $\mathbb{L}_2(S^2)$
- •Are Wigner-D functions:

Spherical harmonics $Y_m^{(l)}: S^2 \to \mathbb{R}$

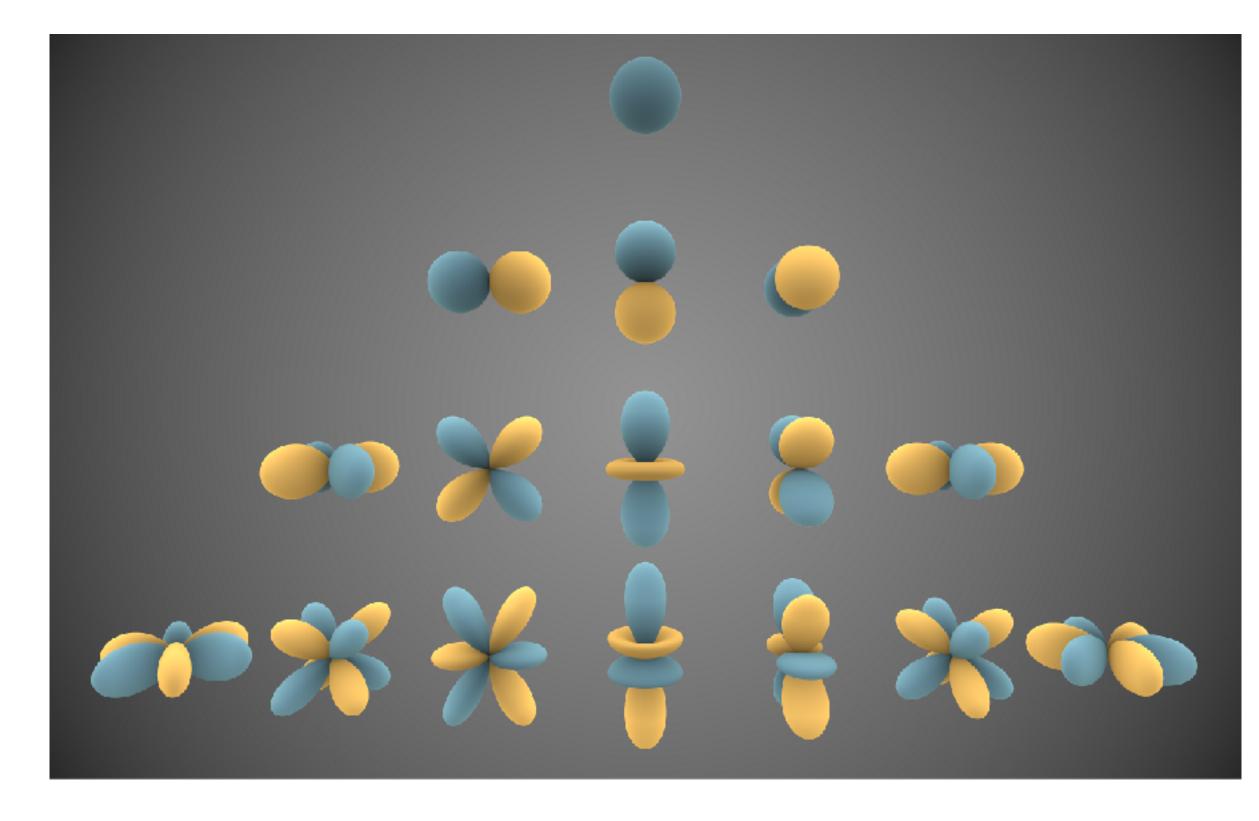


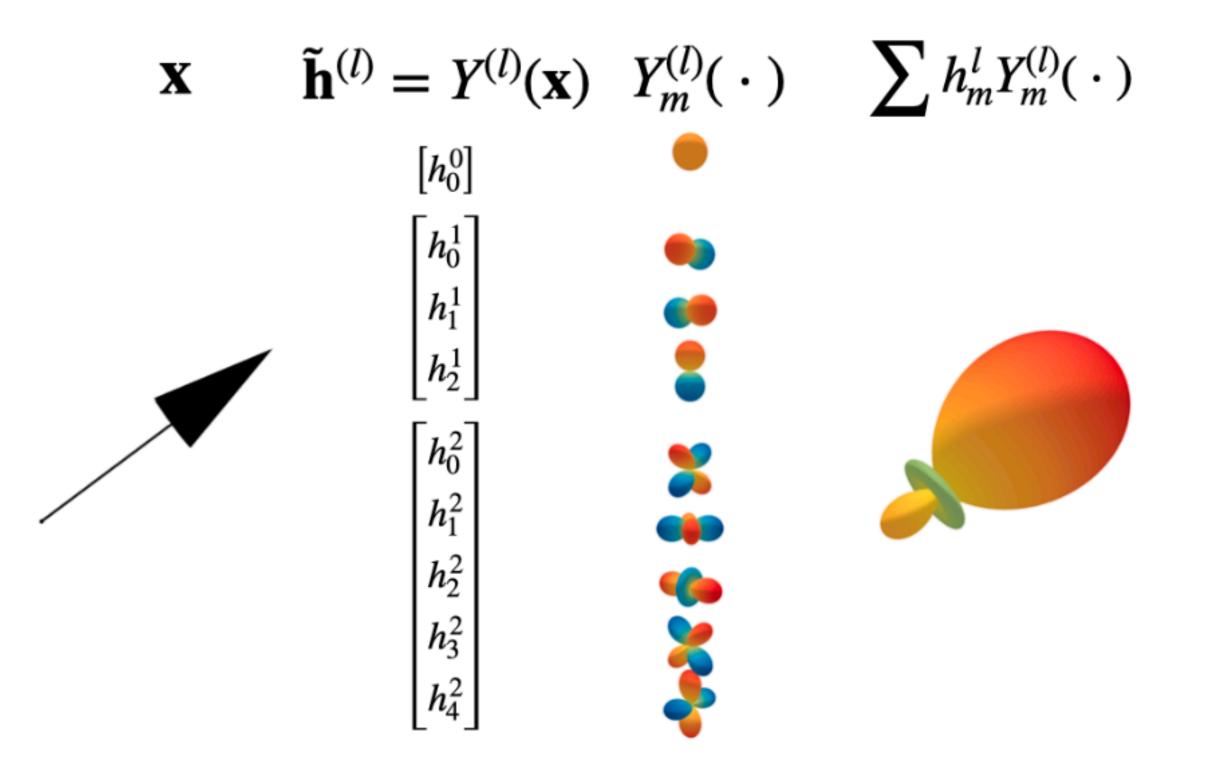
Image: wikipedia

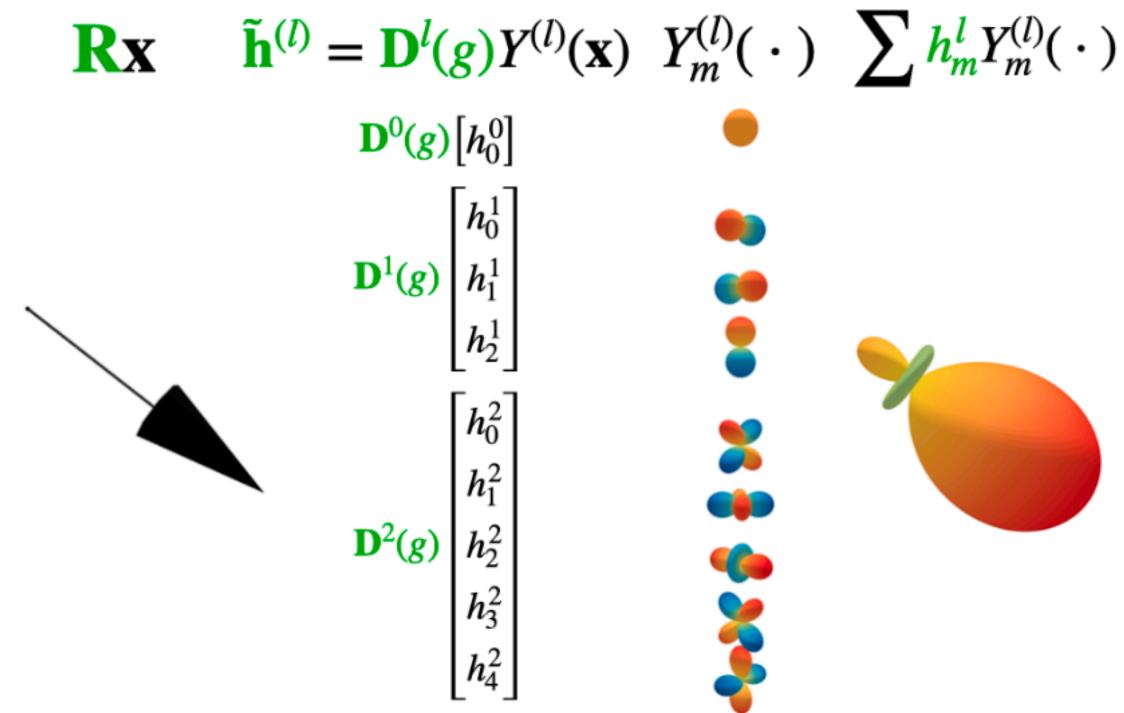


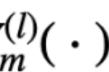




Group theory: Spherical Harmonics Form Steerable Vectors











Group theory: Fourier Transform on S^2

Definition 4.6 (Spherical Fourier transform). Let $f \in \mathbb{L}_2(S^2)$ be a spherical signal and let $\hat{f}(l) \in \mathbb{R}^{2l+1}$ denote the vector of Fourier coefficients of order l. We may refer to l as the frequency index. The forward and inverse Fourier transform are respectively given by

$$\hat{f}(l) = \int_{S^2} f(\mathbf{n}) \underline{Y}^{(l)}(\mathbf{n}) d\mathbf{n}$$

$$f(\mathbf{n}) = \sum_{l \ge 0} \hat{f}(l)^T \underline{Y}^{(l)}(\mathbf{n}),$$
(64)
(65)

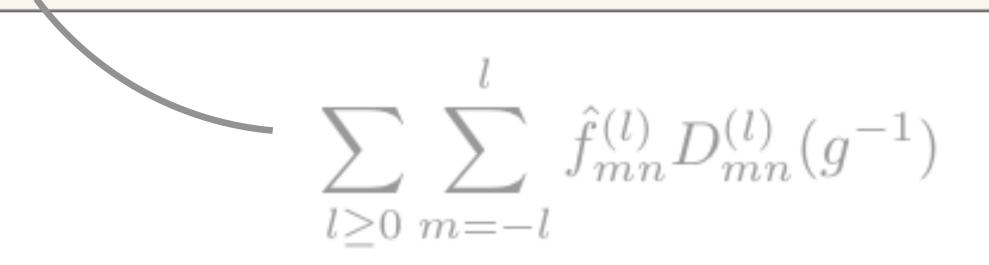
with $\underline{Y}^{(l)} = (Y_{-l}^{(l)}, \dots, Y_{l}^{(l)})^{T} \in \mathbb{L}_{2}(S^{2})^{2l+1}$ the vector of spherical harmonics.



Group theory: Fourier Transform on SO(3)

Definition 4.7 (SO(3) Fourier transform). Let $f \in \mathbb{L}_2(SO(3))$ be a spherical signal and let $\hat{f}(l) \in \mathbb{R}^{2l+1 \times 2l+1}$ denote the matrix of Fourier coefficients of order l. The number l may be referred to as frequency index. The forward and inverse Fourier transform are respectively given by

$$\hat{f}(l) = \int_{SO(3)} f(g) \mathbf{D}^{(l)}(g) \mathrm{d}g ,$$
$$f(g) = \sum_{l \ge 0} \operatorname{Tr}(\hat{f}(l) \mathbf{D}^{(l)}(g^{-1})) .$$





(68)

(69)

Group theory: SO(3) Fourier Theorems

Lemma 4.1 (Shift property). Let \mathcal{L}_g denote the left-regular representation of SO(3) on $\mathbb{L}_2(SO(3))$ and \hat{f} denote the SO(3) Fourier transform SO(3) (Definition 4.7). The SO(3) Fourier transform is equivariant via

 $\mathcal{L}_a f(l) =$



$$= \mathbf{D}^{(l)}(g)\hat{f}(l) .$$
 (70)



78

Group theory: SO(3) Fourier Theorems

tion 4.7). The SO(3) Fourier transform is equivariant via

 $\widehat{\mathcal{L}}_{a}f(l) =$

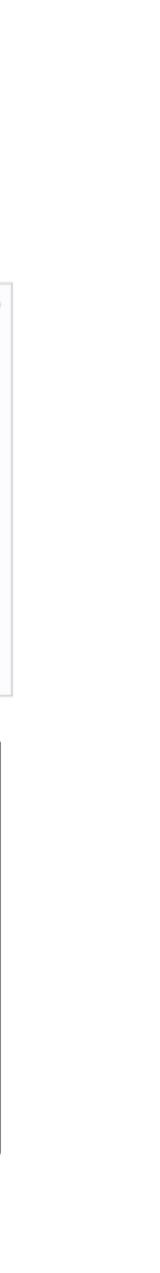
Theorem 4.1 (Convolution theorem on SO(3)). Let $k, f \in \mathbb{L}_2(SO(3))$ and let $\hat{k}(l), \hat{f}(l)$ denote their matrix valued Fourier coefficients. Then the Fourier transform of a (group) correlation of a k with f is given by

 $k \star f(l)$

Lemma 4.1 (Shift property). Let \mathcal{L}_q denote the left-regular representation of SO(3) on $\mathbb{L}_2(SO(3))$ and f denote the SO(3) Fourier transform SO(3) (Defini-

$$= \mathbf{D}^{(l)}(g)\hat{f}(l) .$$
 (70)

$$= \hat{f}(l)\hat{k}(l)^{T}$$
 (71)



78

General tensor product between two vectors:







General tensor product between two vectors:

 $\mathbf{h}_1 \otimes \mathbf{h}_2 = \mathbf{h}_1 \mathbf{h}_2^T = \begin{pmatrix} h_1 h_1 & h_1 h_2 & \dots \\ h_2 h_1 & h_2 h_2 & \dots \\ \vdots & \vdots & \ddots \end{pmatrix}$



We want the tensor product to be equivariant via

 $\mathbf{D}(g)(ilde{\mathbf{h}}_1\otimes ilde{\mathbf{h}}_2) = (\mathbf{D}^{(l_1)}(g) ilde{\mathbf{h}}_1)\otimes (\mathbf{D}^{(l_2)}(g) ilde{\mathbf{h}}_2)$

for some $\mathbf{D}(g)$



The tensor product between two steerable vectors results again in a steerable vector:

$$\operatorname{vec}\left(\left(\mathbf{D}^{(l_{1})}(g)\tilde{\mathbf{h}}_{1}\right)\left(\mathbf{D}^{(l_{2})}(g)\tilde{\mathbf{h}}_{2}\right)^{T}\right) = \operatorname{vec}\left(\left(\mathbf{D}^{(l_{1})}(g)\tilde{\mathbf{h}}_{1}\tilde{\mathbf{h}}_{2}^{T}\mathbf{D}^{(l_{2})T}(g)\right)\right)$$
$$= \left(\left(\mathbf{D}^{(l_{2})}(g)\otimes\mathbf{D}^{(l_{1})}(g)\right)\operatorname{vec}\left(\tilde{\mathbf{h}}_{1}\tilde{\mathbf{h}}_{2}^{T}\right)\right)$$

The resulting representation is reducible.

The CG-product \bigotimes_{cg} is defined in such a way that the output is directly obtained in direct sum of steerable vector spaces $\tilde{\mathbf{h}}_1 \bigotimes_{cg} \tilde{\mathbf{h}}_2 \in V_0 \oplus V_1 \oplus \dots$



Definition 4.8 (Clebsch-Gordan tensor product). Let $\tilde{\mathbf{h}}^{(l)} \in V_l = \mathbb{R}^{2l+1}$ denote a steerable vector of type l and $h_m^{(l)}$ its components with $m = -l, -l + 1, \ldots, l$. Then the Clebsch-Gordan tensor product is defined is a tensor product such tat the *m*-th component of the type *l* sub-vector of the output of the tensor product between two steerable vectors of type l_1 and l_2 is given by

$$(\tilde{\mathbf{h}}^{(l_1)} \otimes_{cg} \tilde{\mathbf{h}}^{(l_2)})_m^{(l)} = \sum_{m_1 = -l_1}^{l_1} \sum_{m_2 = -l_2}^{l_2} C^{(l,m)}_{(l_1,m_1)(l_2,m_2)} h^{(l_1)}_{m_1} h^{(l_2)}_{m_2} , \qquad (73)$$

vector $(\tilde{\mathbf{h}}^{(l_1)} \otimes_{c_q} \tilde{\mathbf{h}}^{(l_2)})^{(l)} \in \mathbb{R}^{2l+1}$ is a type-*l* steerable vector.

in which $C_{(l_1,m_1)(l_2,m_2)}^{(l,m_1)}$ are the Clebsch-Gordan coefficients. The *l*-th output



UNIVERSITY OF AMSTERDAM





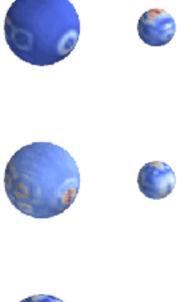
Steerable Neural Networks Method 1: Convolutions in the Fourier domain

Lecture notes Section 5.1

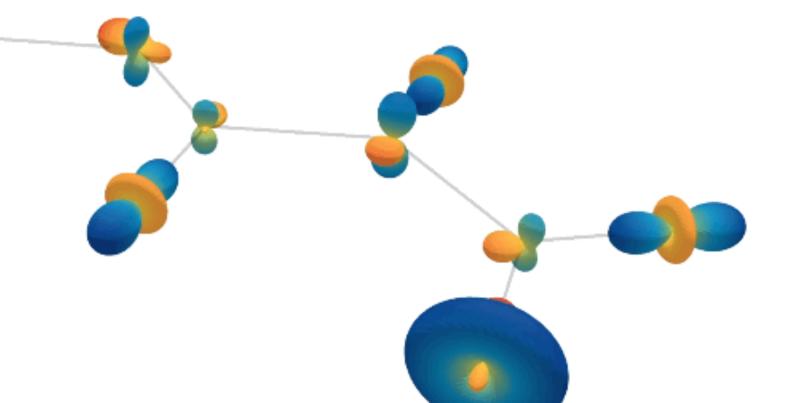


Lecture notes Section 5.2

UNIVERSITY OF AMSTERDAM









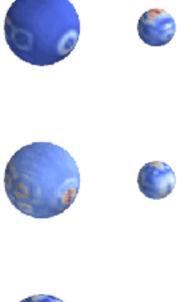
Steerable Neural Networks Method 1: Convolutions in the Fourier domain

Lecture notes Section 5.1

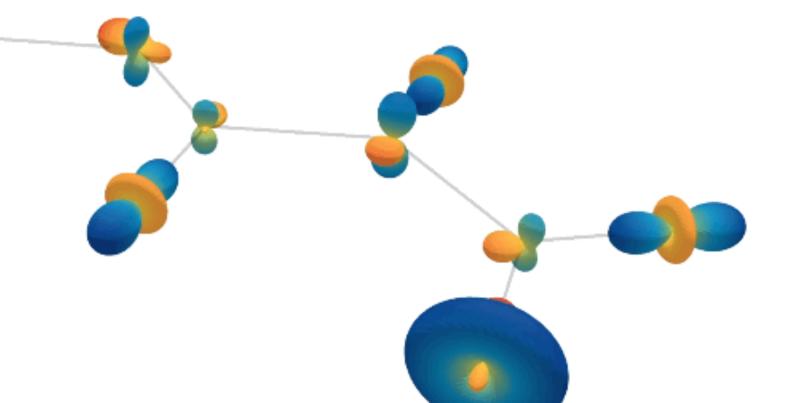


Lecture notes Section 5.2

UNIVERSITY OF AMSTERDAM



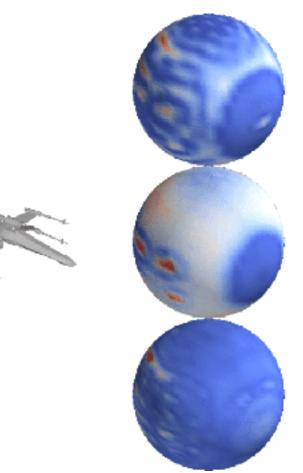






Steerable Neural Networks Method 1: Convolutions in the Fourier domain

Lecture notes Section 5.1



Method 2: Clebsch-Gordan tensor product

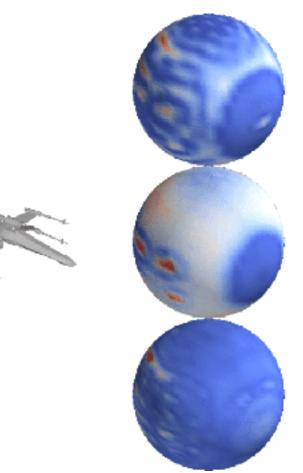
Figure from: <u>https://github.com/daniilidis-group/</u> spherical-cnn Esteves, C., Allen-Blanchette, C., Makadia, A., & Daniilidis, K. Learning SO(3) Equivariant

Representations with Spherical CNNs. European Conference on Computer Vision, ECCV 2018



Steerable Neural Networks Method 1: Convolutions in the Fourier domain

Lecture notes Section 5.1



Method 2: Clebsch-Gordan tensor product

Figure from: <u>https://github.com/daniilidis-group/</u> spherical-cnn Esteves, C., Allen-Blanchette, C., Makadia, A., & Daniilidis, K. Learning SO(3) Equivariant

Representations with Spherical CNNs. European Conference on Computer Vision, ECCV 2018



Spherical CNNs are build via convolutions in the Fourier domain

 $\widehat{k \star f}(l)$

In vectorized form this is simply a matrix vector multiplication

 $\operatorname{vec}(k \star f)(l)$

$$) = \hat{f}(l)\hat{k}^{T}(l)$$

$$= (\hat{w}(l) \otimes I) \operatorname{vec}(\hat{f}(l))$$



87

Spherical CNNs are build via convolutions in the Fourier domain

$$\widehat{k \star f}(l) = \widehat{f}(l)\widehat{k}^T(l)$$

Let us make it explicit and consider signals with only frequency l = 1. Such signals are represented with 3-dimensional vectors.

$$\operatorname{vec}(\hat{f})$$



$$= \begin{pmatrix} \hat{f}_{:,-1} \\ \hat{f}_{:,0} \\ \hat{f}_{:,+1} \end{pmatrix}$$





Spherical CNNs are build via convolutions in the Fourier domain

$$\widehat{k \star f}(l) = \widehat{f}(l)\widehat{k}^T(l)$$

Let us make it explicit and consider signals with only frequency l = 1. Such signals are represented with 3-dimensional vectors.

$$\operatorname{vec}(\hat{f})$$



$$= \begin{pmatrix} \hat{f}_{:,-1} \\ \hat{f}_{:,0} \\ \hat{f}_{:,+1} \end{pmatrix}$$

For
$$S^2$$
 signals: red



=zero



Spherical CNNs are build via convolutions in the Fourier domain

$$\widehat{k \star f}(l) = \widehat{f}(l)\widehat{k}^T(l)$$

Let us make it explicit and consider signals with only frequency l = 1. Then

$$\begin{pmatrix} \widehat{k \star f}_{:,-1} \\ \widehat{k \star f}_{:,0} \\ \widehat{k \star f}_{:,+1} \end{pmatrix} = \begin{pmatrix} w_{-1,-1}I & w_{-1,0}I & w_{-1,1}I \\ w_{0,-1}I & w_{0,0}I & w_{0,1}I \\ w_{1,-1}I & w_{1,0}I & w_{1,1}I \end{pmatrix} \begin{pmatrix} \widehat{f}_{:,-1} \\ \widehat{f}_{:,0} \\ \widehat{f}_{:,+1} \end{pmatrix}$$

Only three weights: the kernel is also a spherical harmonic!







Spherical CNNs are build via convolutions in the Fourier domain

$$\widehat{k \star f}(l) = \widehat{f}(l)\widehat{k}^T(l)$$

Let us make it explicit and consider signals with only frequency l = 1. Then

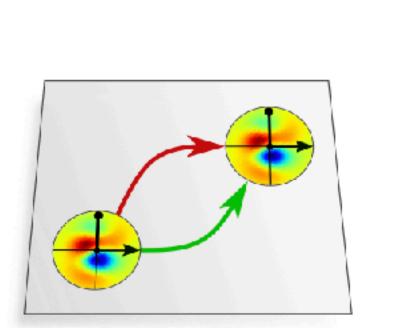
$$\begin{pmatrix} \widehat{k \star f}_{:,-1} \\ \widehat{k \star f}_{:,0} \\ \widehat{k \star f}_{:,+1} \end{pmatrix} = \begin{pmatrix} w_{-1,-1}I & w_{-1,0}I & w_{-1,1}I \\ w_{0,-1}I & w_{0,0}I & w_{0,1}I \\ w_{1,-1}I & w_{1,0}I & w_{1,1}I \end{pmatrix} \begin{pmatrix} \widehat{f}_{:,-1} \\ \widehat{f}_{:,0} \\ \widehat{f}_{:,+1} \end{pmatrix}$$

If we want the output to be a SH, only 1 weight can be non-zero!









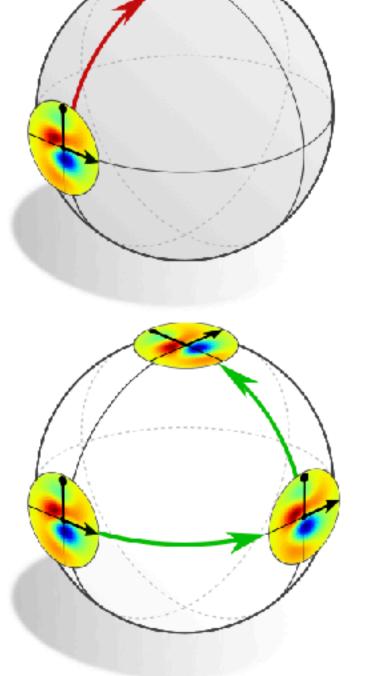


Figure from COORDINATE INDEPENDENT CONVOLUTIONAL NETWORKS, Weiler, Forré, Verlinde, Welling

Lemma 5.1. Consider the cases of SO(3) equivariant linear layers for signals on $X = S^2 \equiv SO(3)/SO(2)$ or X = SO(3). Such convolutions can be performed via the Fourier transform on SO(3)

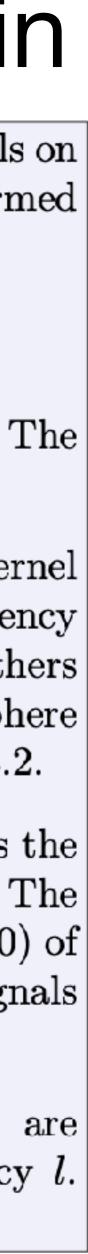
$$\widehat{k \star f}(l) = \widehat{f}(l)\widehat{w}^T(l),$$

with $\hat{w}(l) \in \mathbb{R}^{2l+1 \times 2l_1}$ the learnable parameters of the convolution kernel. The following can be said about the parametrization of the kernels through \hat{w} .

• Isotropic S^2 kernel convolutions $(X = Y = S^2)$ For isotropic kernel convolutions the kernel is parametrized by a single weight per frequency l. Namely the only possibly non-zero component is $\hat{w}_{00}(l)$ and all others have to be 0. This means that the kernels represent signals on the sphere S^2 that are symmetry around the \mathbf{n}_x axis, as required by Theorem 3.2.

• Lifting convolutions $(X = S^2, Y = SO(3))$ For lifting convolutions the kernel is parametrized by (2l+1) learnable weights per frequency l. The only possible non-zero weights are $\hat{w}_{:,0}$, i.e., the central column (n=0) of $\hat{w}(l)$. This means that the convolution kernels are unconstrained signals on the sphere S^2 .

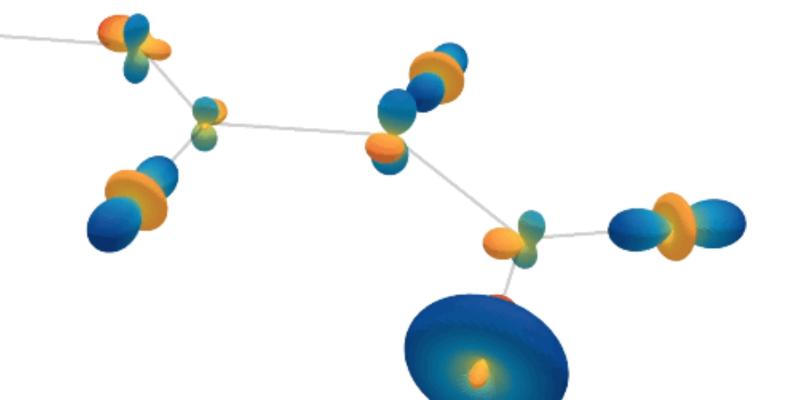
• Group convolution (X = Y = SO(3)) Group convolutions are parametrized by $(2l + 1) \times (2l + 1)$ learnable weights per frequency l. The kernels represent unconstrained functions on SO(3).





Method 2: Clebsch-Gordan tensor product(SE(3) gconvs)Lecture notes Section 5.2

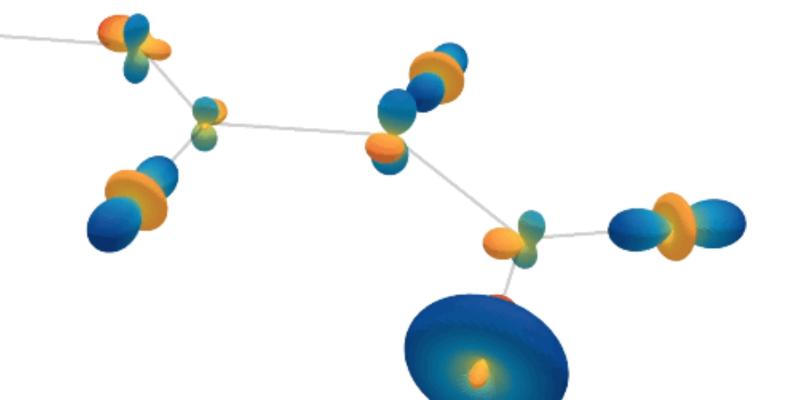
University of Amsterdam





Method 2: Clebsch-Gordan tensor product(SE(3) gconvs)Lecture notes Section 5.2

University of Amsterdam

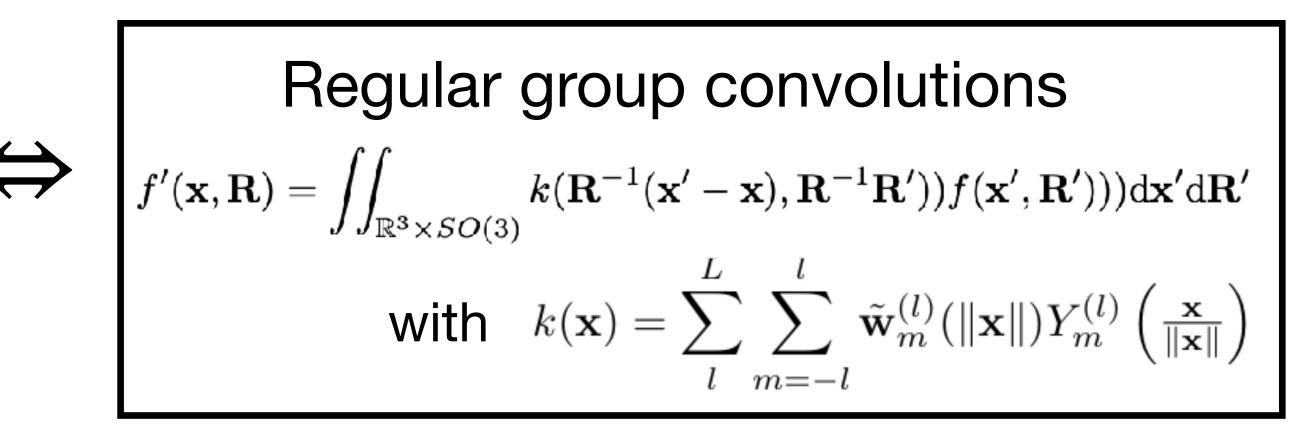


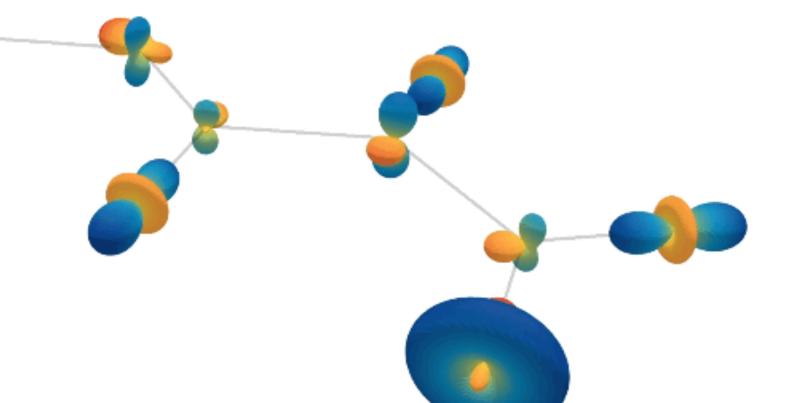


Steerable group convolutions

$$\tilde{\mathbf{f}}'(\mathbf{x}) = \int_{\mathbb{R}^3} \tilde{\mathbf{f}}(\mathbf{x}) \otimes_{cg}^{w(\|\mathbf{x}\|)} \underline{Y}^{(l)} \left(\frac{\mathbf{x}' - \mathbf{x}}{\|\mathbf{x}' - \mathbf{x}\|} \right) \mathrm{d}\mathbf{x}'$$

Method 2: Clebsch-Gordan tensor product Lecture notes Section 5.2



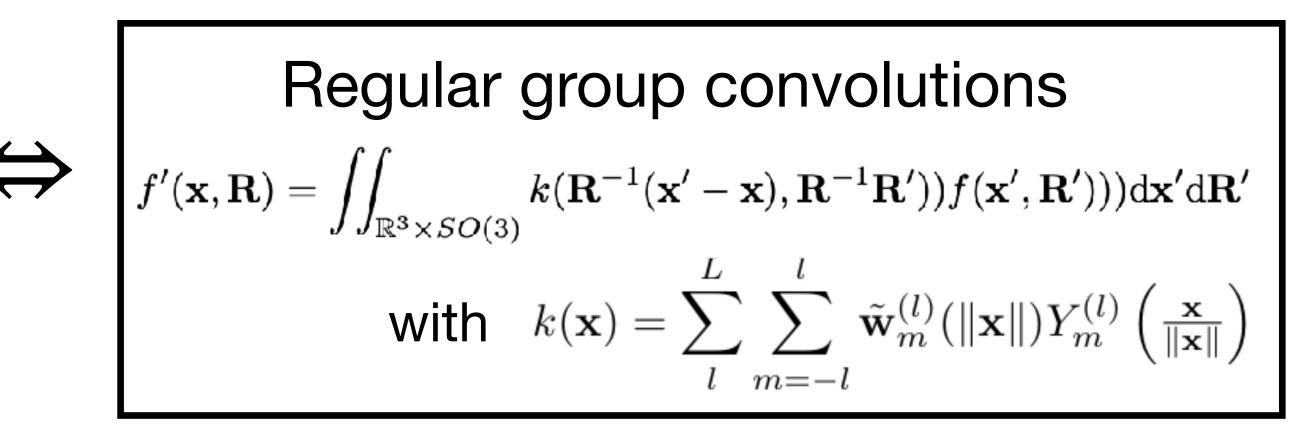


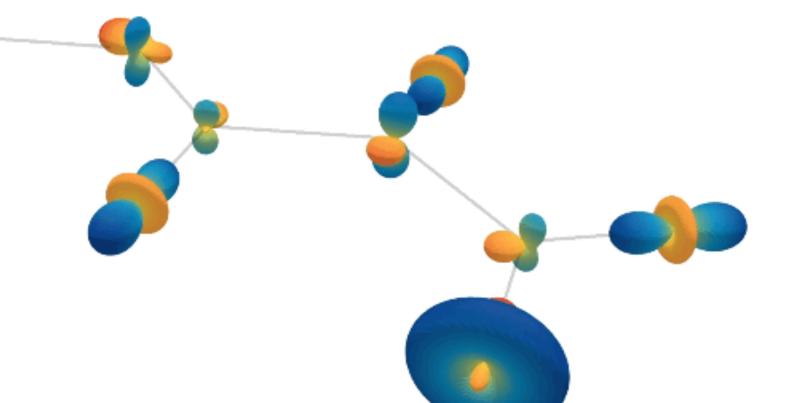


Steerable group convolutions

$$\tilde{\mathbf{f}}'(\mathbf{x}) = \int_{\mathbb{R}^3} \tilde{\mathbf{f}}(\mathbf{x}) \otimes_{cg}^{w(\|\mathbf{x}\|)} \underline{Y}^{(l)} \left(\frac{\mathbf{x}' - \mathbf{x}}{\|\mathbf{x}' - \mathbf{x}\|} \right) \mathrm{d}\mathbf{x}'$$

Method 2: Clebsch-Gordan tensor product Lecture notes Section 5.2



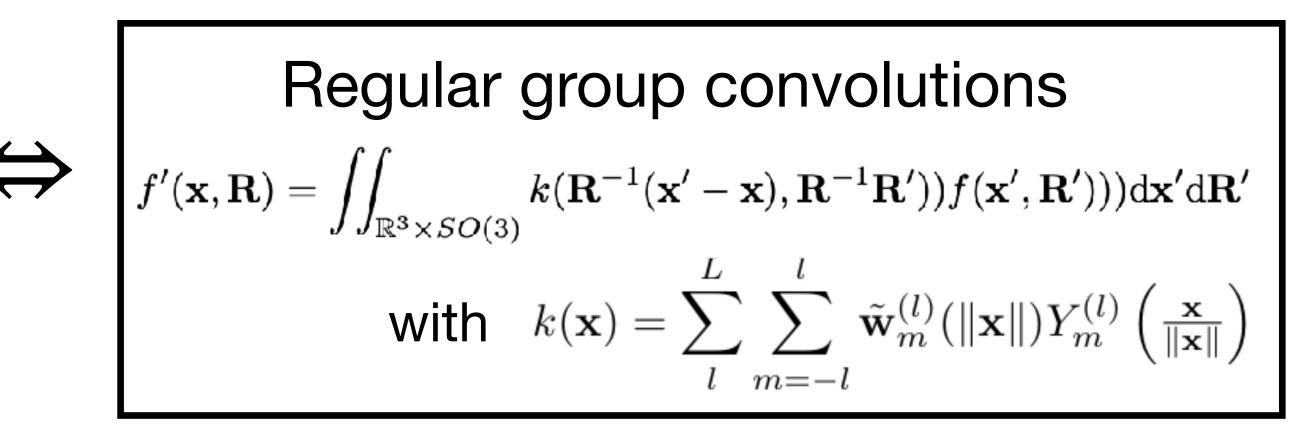


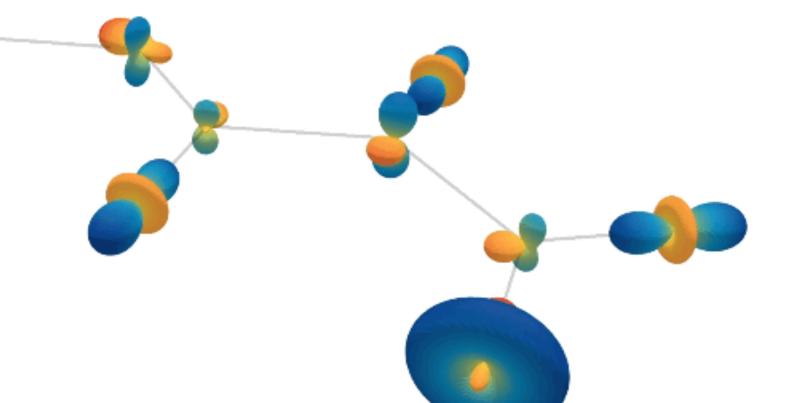


Steerable group convolutions

$$\tilde{\mathbf{f}}'(\mathbf{x}) = \int_{\mathbb{R}^3} \tilde{\mathbf{f}}(\mathbf{x}) \otimes_{cg}^{w(\|\mathbf{x}\|)} \underline{Y}^{(l)} \left(\frac{\mathbf{x}' - \mathbf{x}}{\|\mathbf{x}' - \mathbf{x}\|} \right) \mathrm{d}\mathbf{x}'$$

Method 2: Clebsch-Gordan tensor product Lecture notes Section 5.2



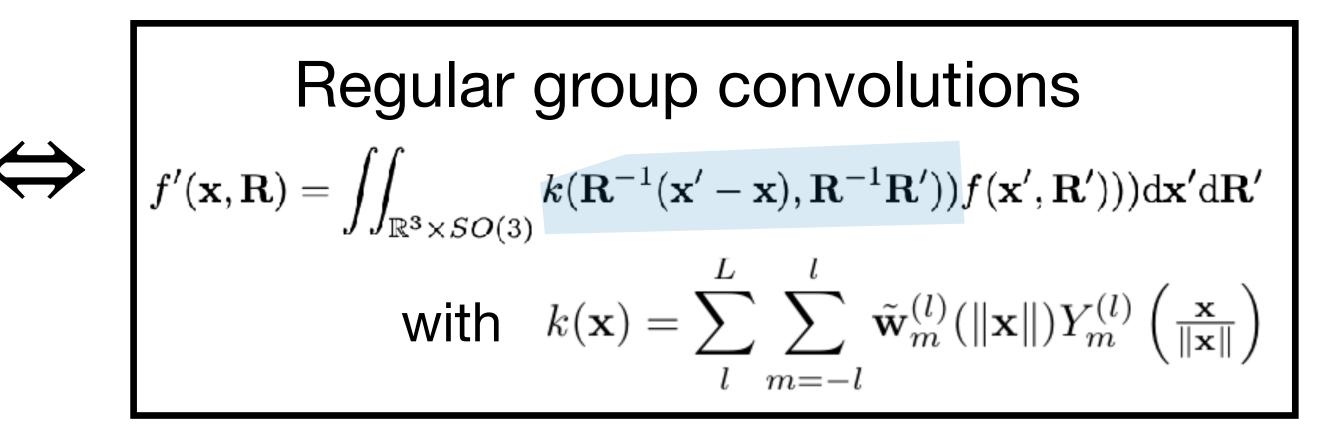


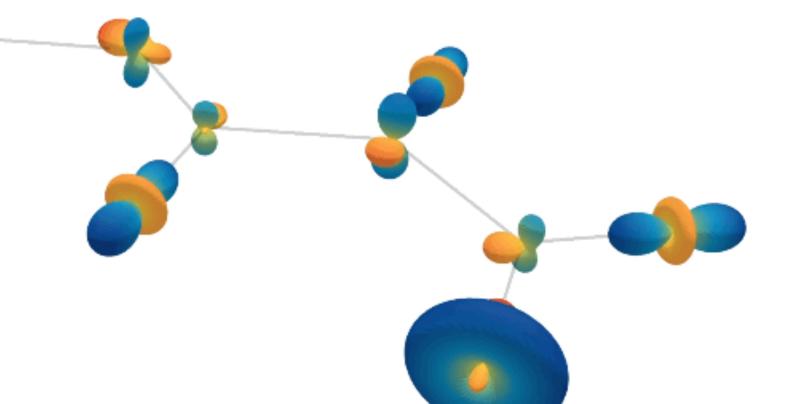


Steerable group convolutions

$$\tilde{\mathbf{f}}'(\mathbf{x}) = \int_{\mathbb{R}^3} \tilde{\mathbf{f}}(\mathbf{x}) \otimes_{cg}^{w(\|\mathbf{x}\|)} \underline{Y}^{(l)} \left(\frac{\mathbf{x}' - \mathbf{x}}{\|\mathbf{x}' - \mathbf{x}\|} \right) \mathrm{d}\mathbf{x}'$$

Method 2: Clebsch-Gordan tensor product Lecture notes Section 5.2







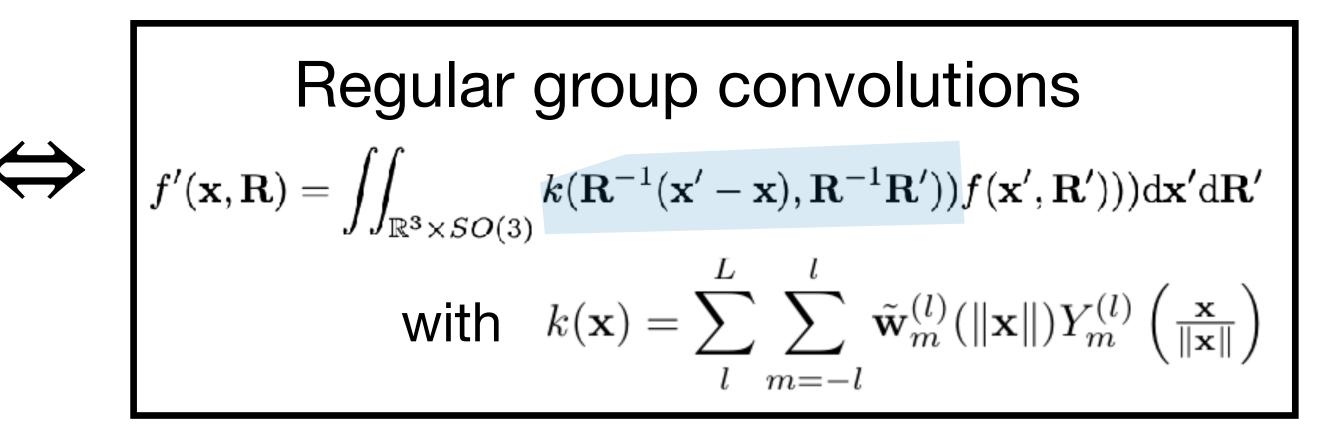
Steerable group convolutions

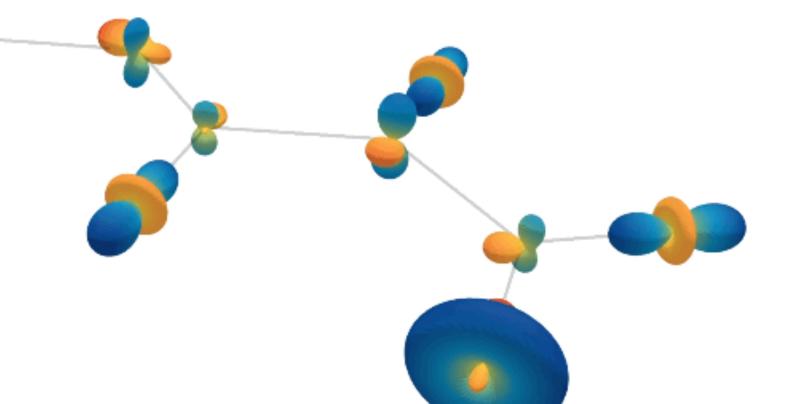
$$\tilde{\mathbf{f}}'(\mathbf{x}) = \int_{\mathbb{R}^3} \tilde{\mathbf{f}}(\mathbf{x}) \otimes_{cg}^{w(\|\mathbf{x}\|)} \underline{Y}^{(l)} \left(\frac{\mathbf{x}' - \mathbf{x}}{\|\mathbf{x}' - \mathbf{x}\|}\right) \mathrm{d}\mathbf{x}'$$

Angular part of k

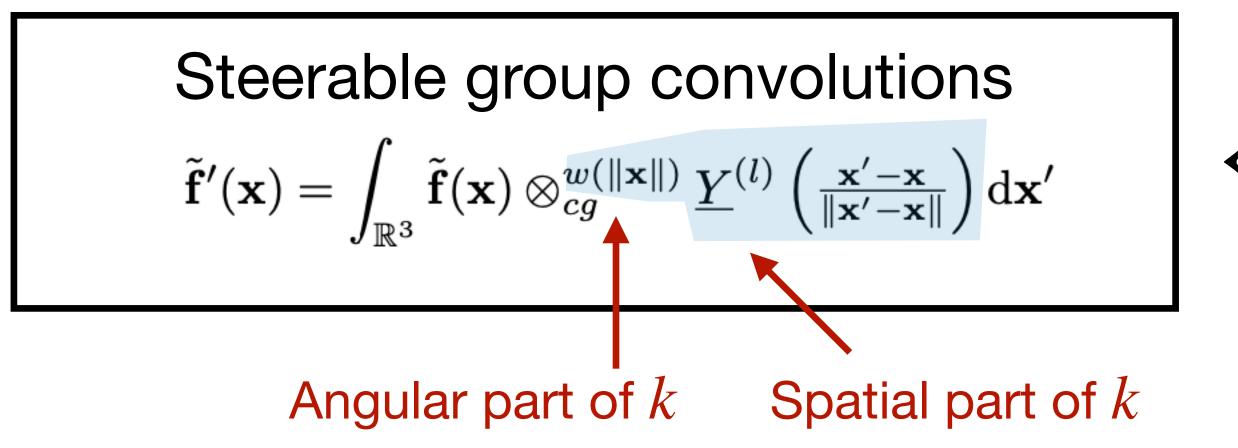
Method 2: Clebsch-Gordan tensor product Lecture notes Section 5.2

University of Amsterdam



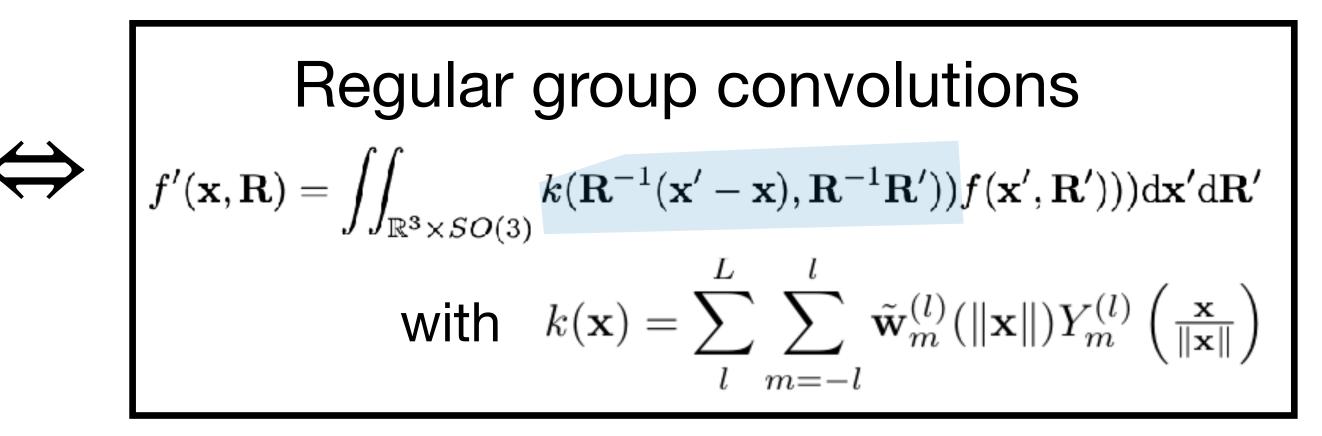


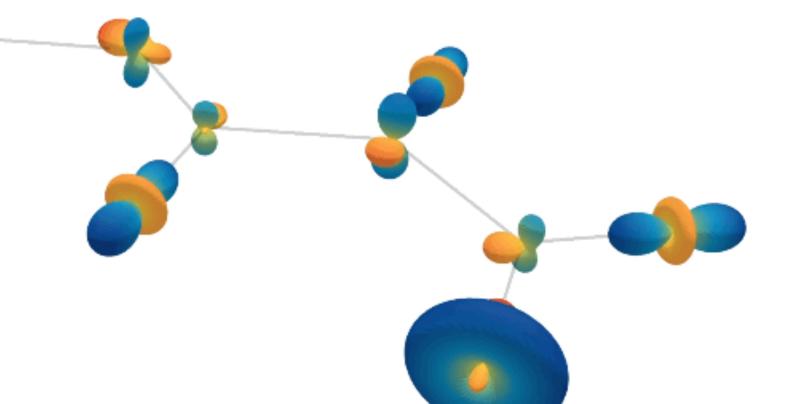




Method 2: Clebsch-Gordan tensor product (SE(3) gconvs)Lecture notes Section 5.2

University of Amsterdam



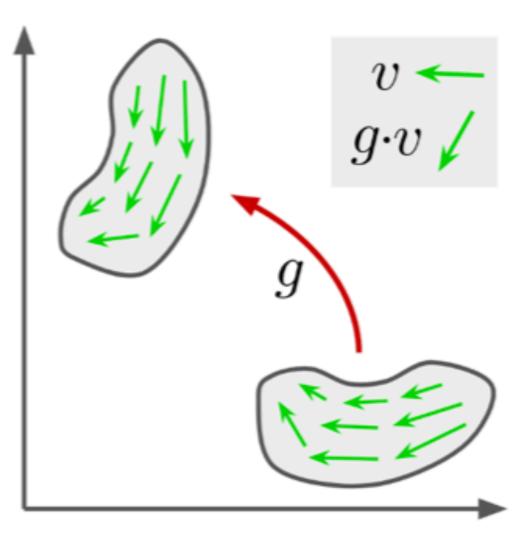


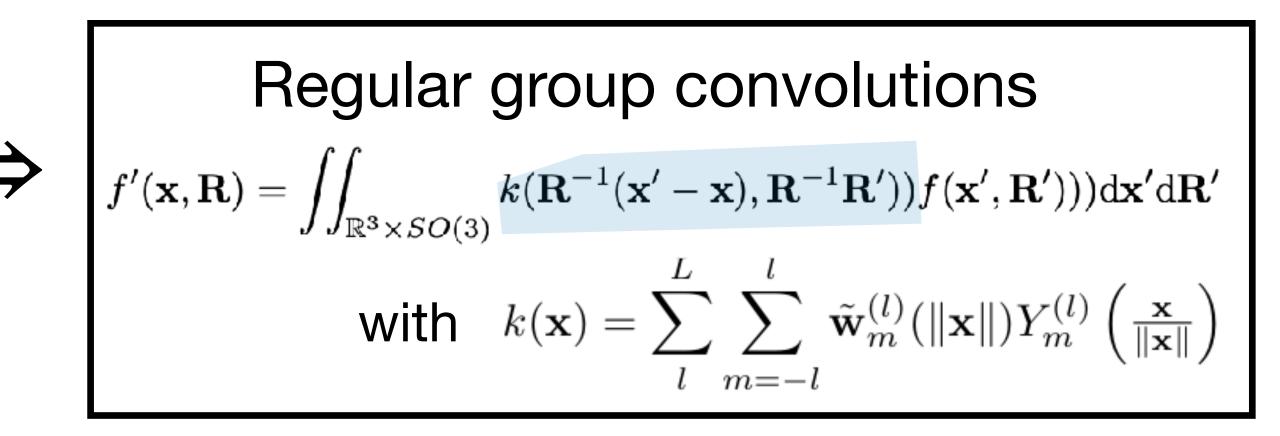


Steerable group convolutions

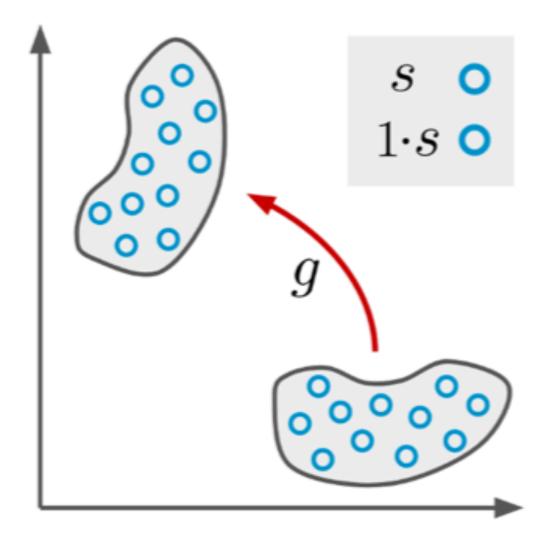
$$\tilde{\mathbf{f}}'(\mathbf{x}) = \int_{\mathbb{R}^3} \tilde{\mathbf{f}}(\mathbf{x}) \otimes_{cg}^{w(\|\mathbf{x}\|)} \underline{Y}^{(l)} \left(\frac{\mathbf{x}' - \mathbf{x}}{\|\mathbf{x}' - \mathbf{x}\|} \right) \mathrm{d}\mathbf{x}'$$

Feature maps: $f : \mathbb{R}^3 \to V_0 \oplus V_1 \oplus \dots$





Feature maps: $f : \mathbb{R}^3 \times S^2 \to \mathbb{R}$



Figures: <u>https://github.com/QUVA-Lab/e2cnn</u> 95





Libraries and repositories

e2cnn: <u>https://github.com/QUVA-Lab/e2cnn</u>

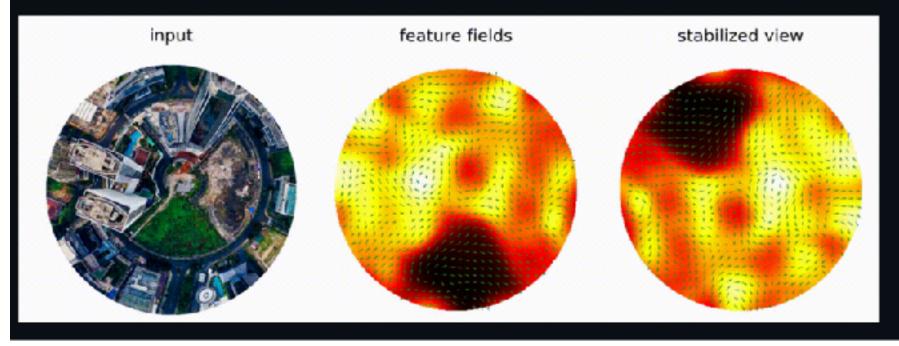
e3nn: https://github.com/e3nn/e3nn

e3cnn (released soon)

UNIVERSITY OF AMSTERDAM

Demo

Since E(2)-steerable CNNs are equivariant under rotations and reflections, their inference is independent from the choice of image orientation. The visualization below demonstrates this claim by feeding rotated images into a randomly initialized E(2)-steerable CNN (left). The middle plot shows the equivariant transformation of a feature space, consisting of one scalar field (color-coded) and one vector field (arrows), after a few layers. In the right plot we transform the feature space into a comoving reference frame by rotating the response fields back (stabilized view)

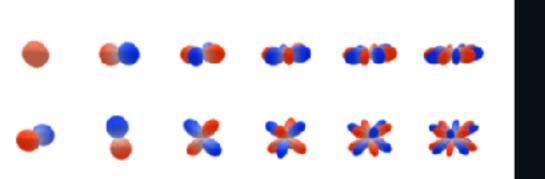


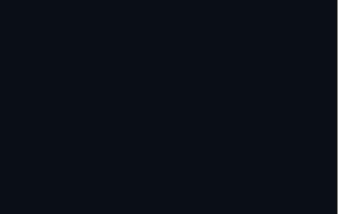
e3nn

coverage 97% DOI 10.5281/zenodo.5006322

Documentation | Code | ChangeLog | Colab

The aim of this library is to help the development of E3 equivariant neural networks. It contains fundamental nathematical operations such as tensor products and spherical harmonics.







Conclusion



Conclusion

- G-CNNs naturally arise from NNs under equivariance constraints
- G-CNNs improve upon classic CNNs by
 - Making data augmentation w.r.t. the group obsolete
 - No valuable network capacity needs to be spend on dealing w geometry
 - The added geometric structure allows to deal with context (recognition by components, relative poses)
 - The added geometric structure enables to reach performances that cannot be achieved with data augmentation alone
 - Have guaranteed geometric stability
 - Can be applied to many types of signal data (not covered today: equivariance to Lie groups and gauge equivariant methods

