Group Equivariant Deep Learning

Lecture 1 - Regular group convolutions

Lecture 1.3 - Regular group convolutions | Template matching viewpoint

General group convolutional NN design with example for roto-translation equivariance (SE(2))

Erik Bekkers, Amsterdam Machine Learning Lab, University of Amsterdam
This mini-course serves as a module with the UvA Master AI course Deep Learning 2 https://uvadl2c.github.io/

Are convolutions with reflected conv kernels (and vice versa)

Cross-correlations

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

Are convolutions with reflected conv kernels (and vice versa)

Cross-correlations

Representation of the translation group!

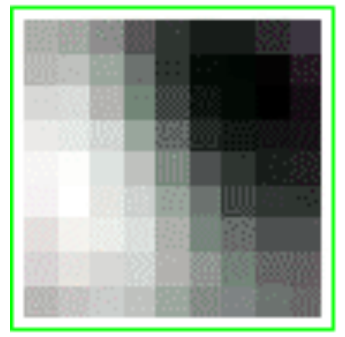
$$(k \star_{\mathbb{R}^2} f)(\mathbf{x}) = \int_{\mathbb{R}^2} k(\mathbf{x}' - \mathbf{x}) f(\mathbf{x}') d\mathbf{x}' = (\mathscr{L}_g\, k\, ,\, f)_{\mathbb{L}_2(\mathbb{R}^2)}$$

Are convolutions with reflected conv kernels (and vice versa)

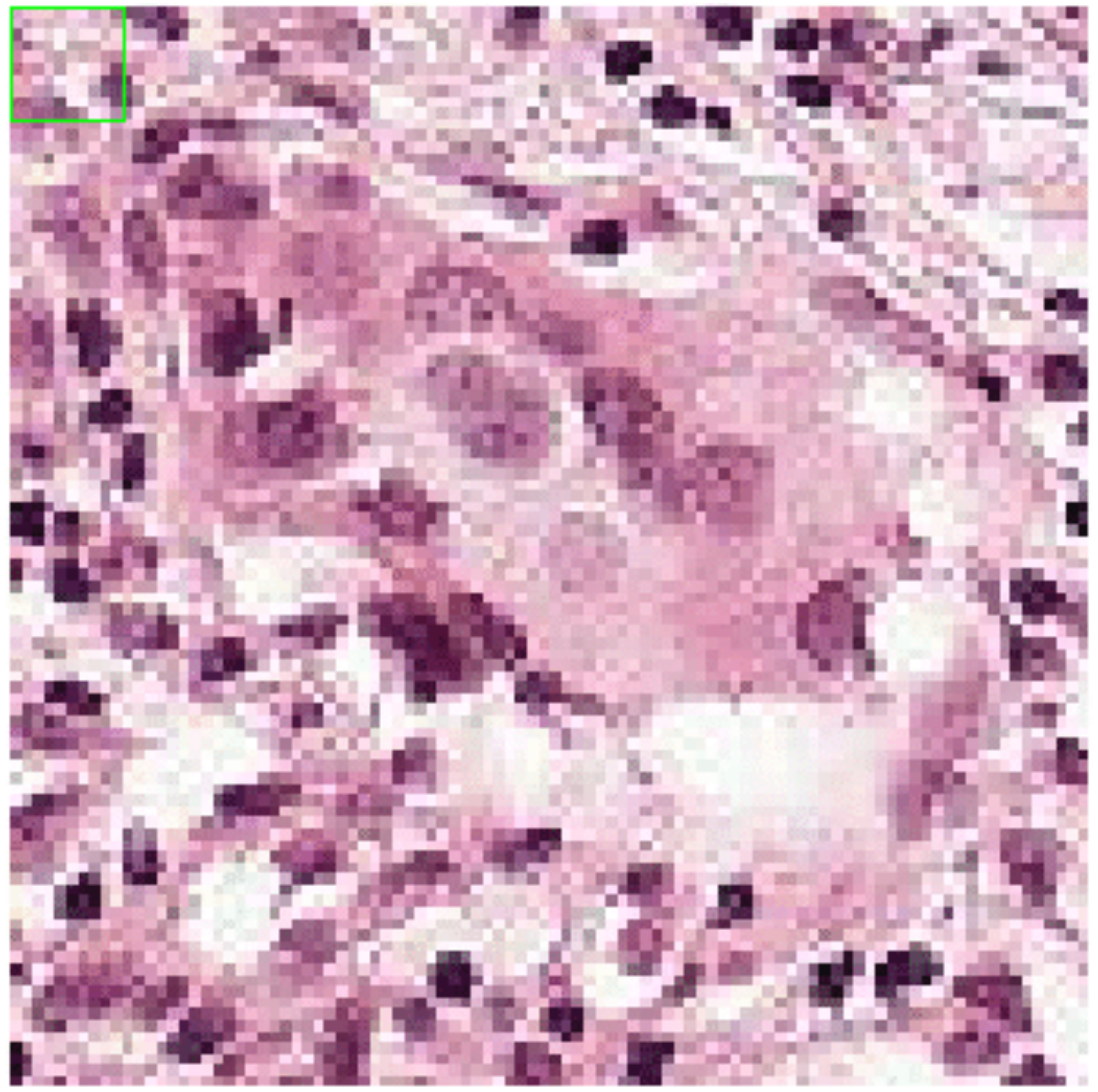
Cross-correlations

Representation of the translation group!

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



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



$=$

k
2D convolution kernel

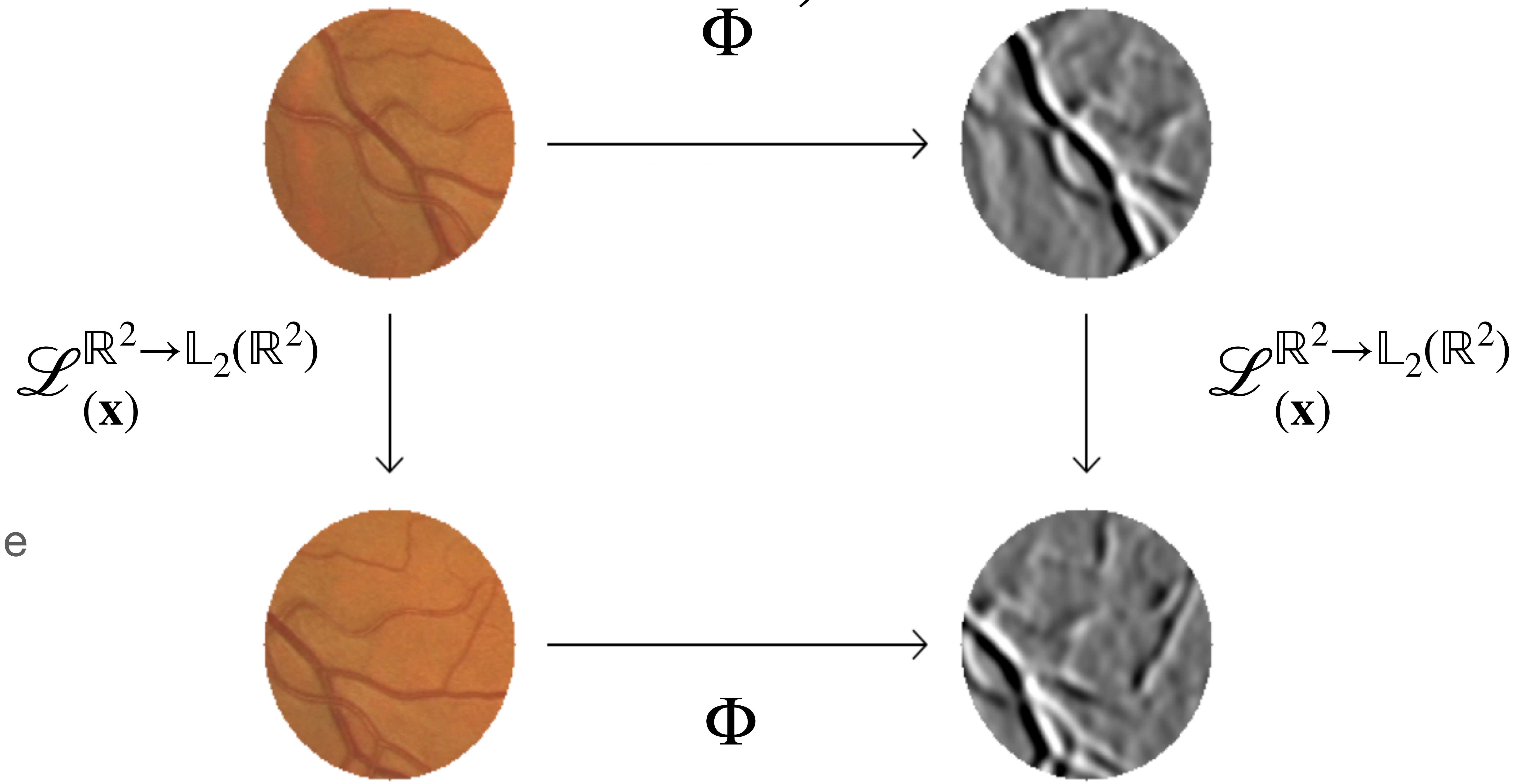
f^{in}
2D feature map

f^{out}
2D feature map (after ReLU)

Equivariance

Convolutions/cross-correlations are translation equivariant

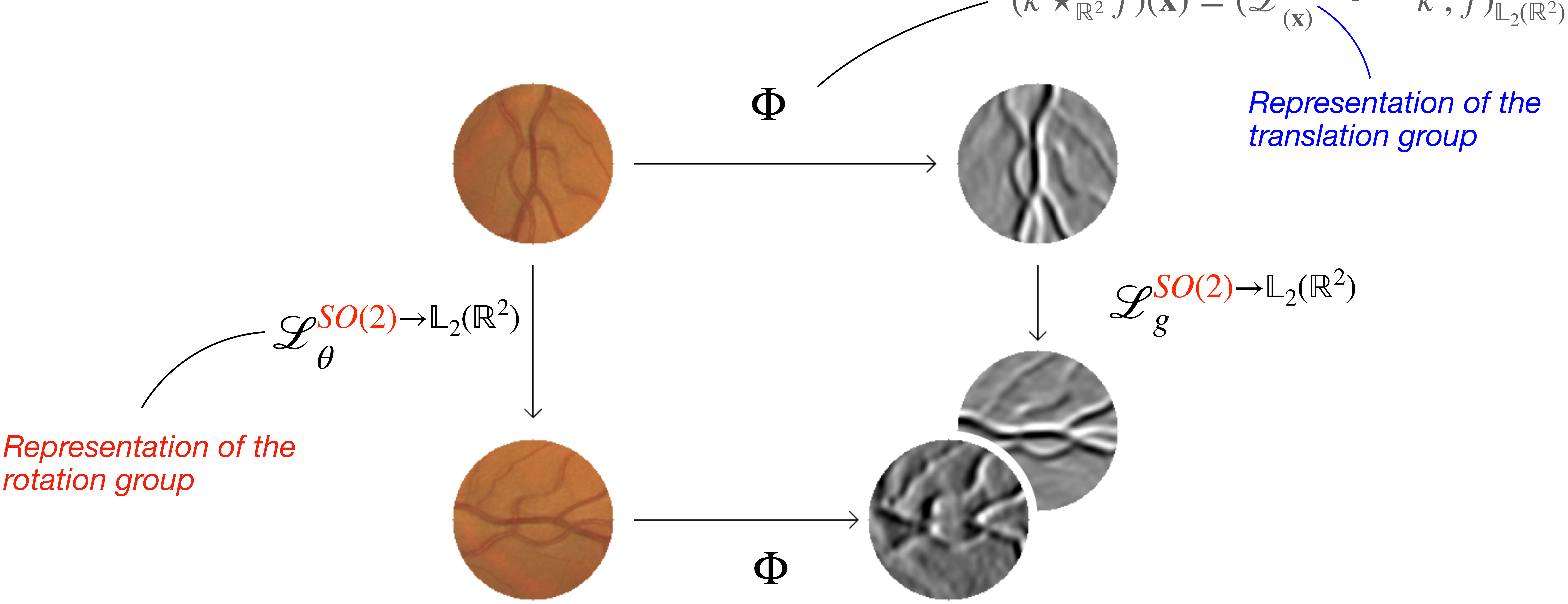
$$(k \star_{\mathbb{R}^2} f)(\mathbf{x}) = (\mathscr{L}^{\mathbb{R}^2 \to \mathbb{L}_2(\mathbb{R}^2)}_{(\mathbf{x})} k \,, f)_{\mathbb{L}_2(\mathbb{R}^2)}$$



Representation of the translation group

Equivariance

Convolutions are generally **not equivariant to roto-translations**



SE(2) equivariant cross-correlations

Representation of the roto-translation group!

Lifting correlations: $(k \tilde{\star} f)(\mathbf{x}, \theta) = (\mathscr{L}_g^{SE(2) \rightarrow \mathbb{L}_2(\mathbb{R}^2)} k \,, f)_{\mathbb{L}_2(\mathbb{R}^2)}$

SE(2) equivariant cross-correlations

Representation of the roto-translation group!

Lifting correlations: $(k \tilde{\star} f)(\mathbf{x}, \theta) = (\mathscr{L}_g^{SE(2) \to \mathbb{L}_2(\mathbb{R}^2)} k \, , f)_{\mathbb{L}_2(\mathbb{R}^2)} = (\mathscr{L}_{\mathbf{x}}^{\mathbf{R}^2 \to \mathbb{L}_2(\mathbb{R}^2)} \mathscr{L}_\theta^{SO(2) \to \mathbb{L}_2(\mathbb{R}^2)} k \, , f)_{\mathbb{L}_2(\mathbb{R}^2)}$

translation *rotation*

SE(2) equivariant cross-correlations

$k(\mathbf{R}_\theta^{-1}(\mathbf{x}' - \mathbf{x}))$

Representation of the roto-translation group!

Lifting correlations: $(k \tilde{\star} f)(\mathbf{x}, \theta) = (\mathscr{L}_g^{SE(2) \to \mathbb{L}_2(\mathbb{R}^2)} k \,, f)_{\mathbb{L}_2(\mathbb{R}^2)} = (\mathscr{L}_\mathbf{x}^{\mathbf{R}^2 \to \mathbb{L}_2(\mathbb{R}^2)} \mathscr{L}_\theta^{SO(2) \to \mathbb{L}_2(\mathbb{R}^2)} k \,, f)_{\mathbb{L}_2(\mathbb{R}^2)}$

translation *rotation*

SE(2) equivariant cross-correlations

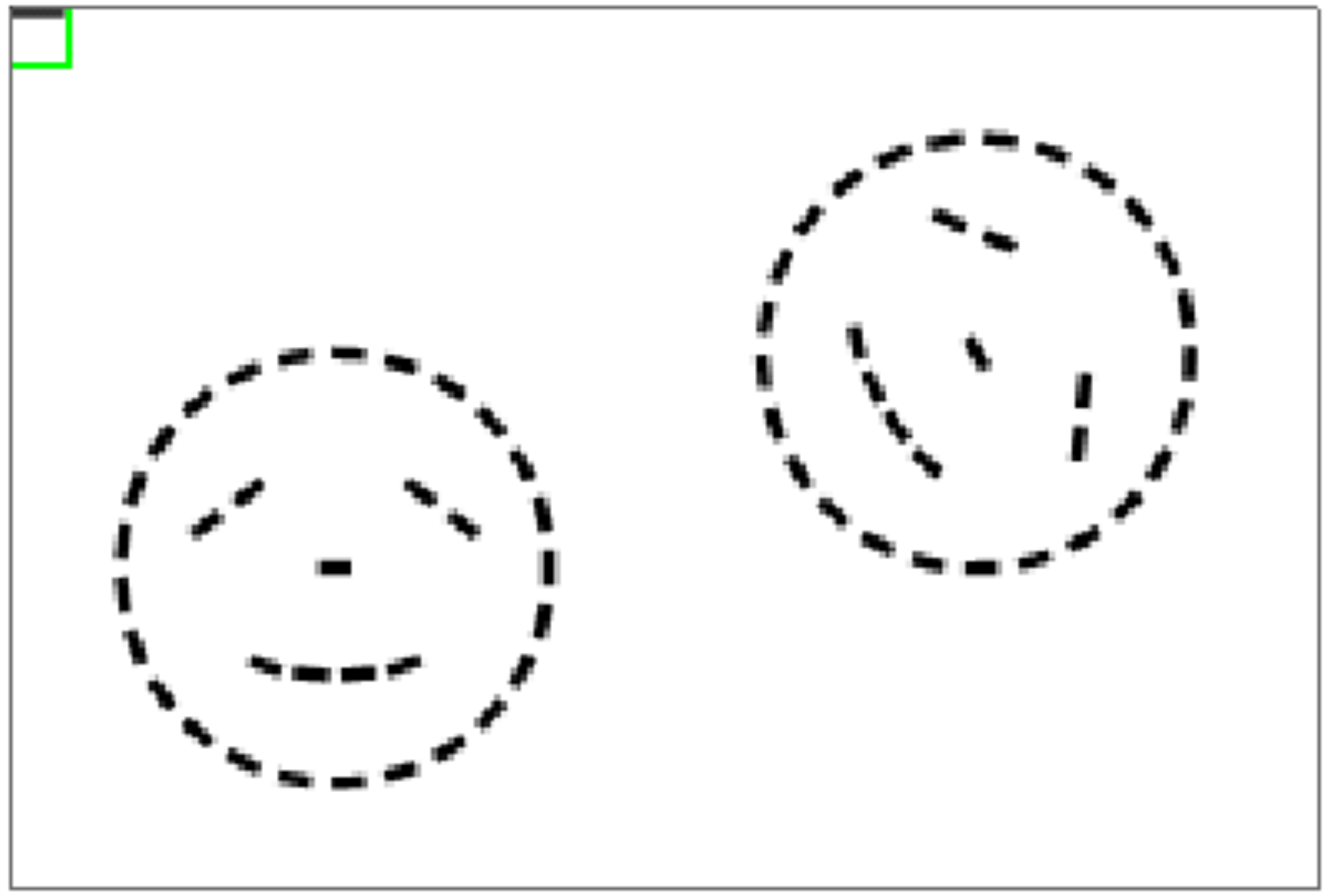
Representation of the roto-translation group!

$k(\mathbf{R}_\theta^{-1}(\mathbf{x}'-\mathbf{x}))$

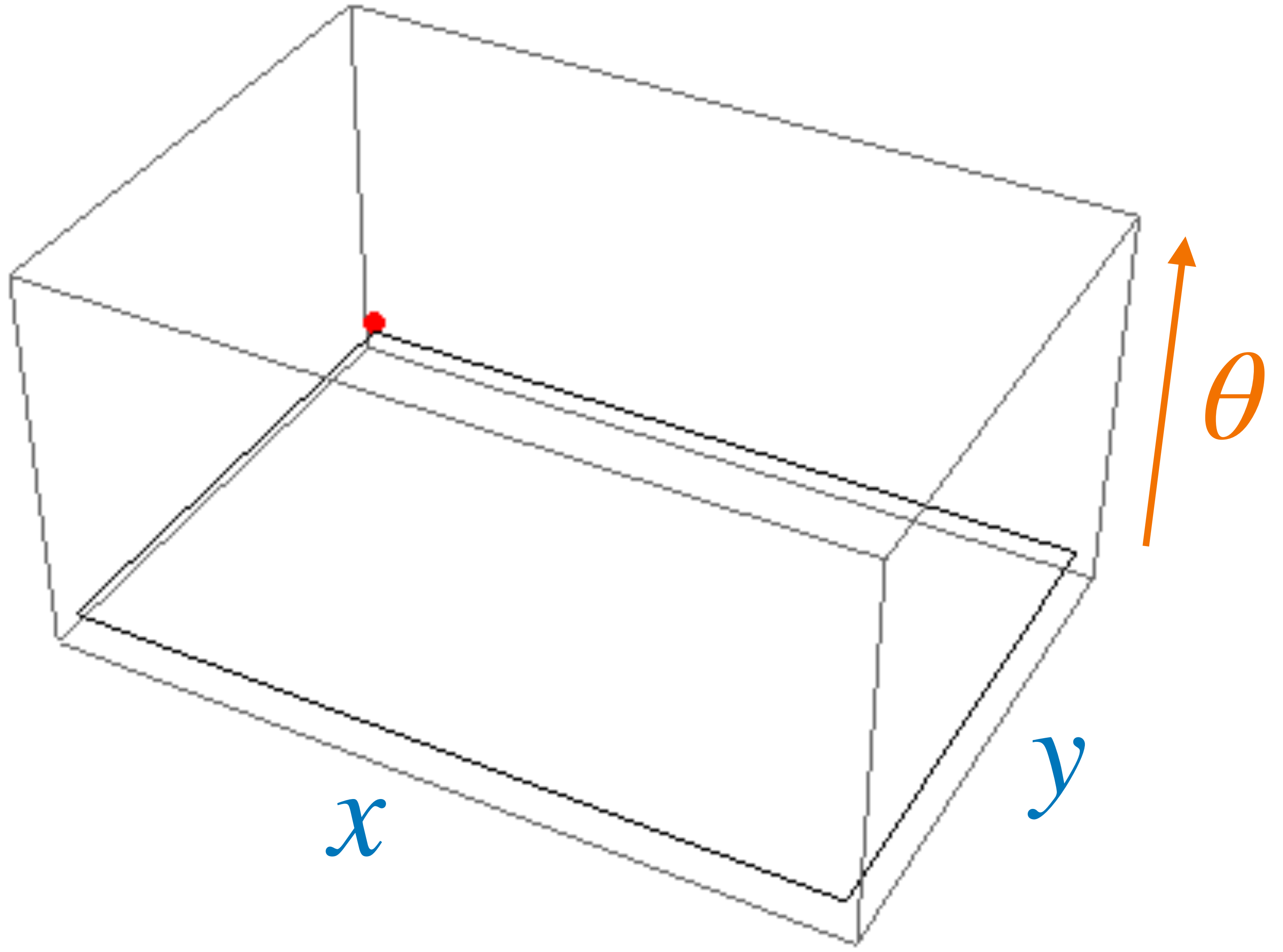
Lifting correlations: $(k \tilde{\star} f)(\mathbf{x},\theta) = (\mathscr{L}_g^{SE(2)\to\mathbb{L}_2(\mathbb{R}^2)} k \,, f)_{\mathbb{L}_2(\mathbb{R}^2)} = (\mathscr{L}_{\mathbf{x}}^{\mathbb{R}^2\to\mathbb{L}_2(\mathbb{R}^2)} \mathscr{L}_\theta^{SO(2)\to\mathbb{L}_2(\mathbb{R}^2)} k \,, f)_{\mathbb{L}_2(\mathbb{R}^2)}$

translation *rotation*

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



$=$



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

f^{in}
2D feature map

f^{out}
3D (*SE(2)*) feature map (after ReLU)

SE(2) equivariant cross-correlations

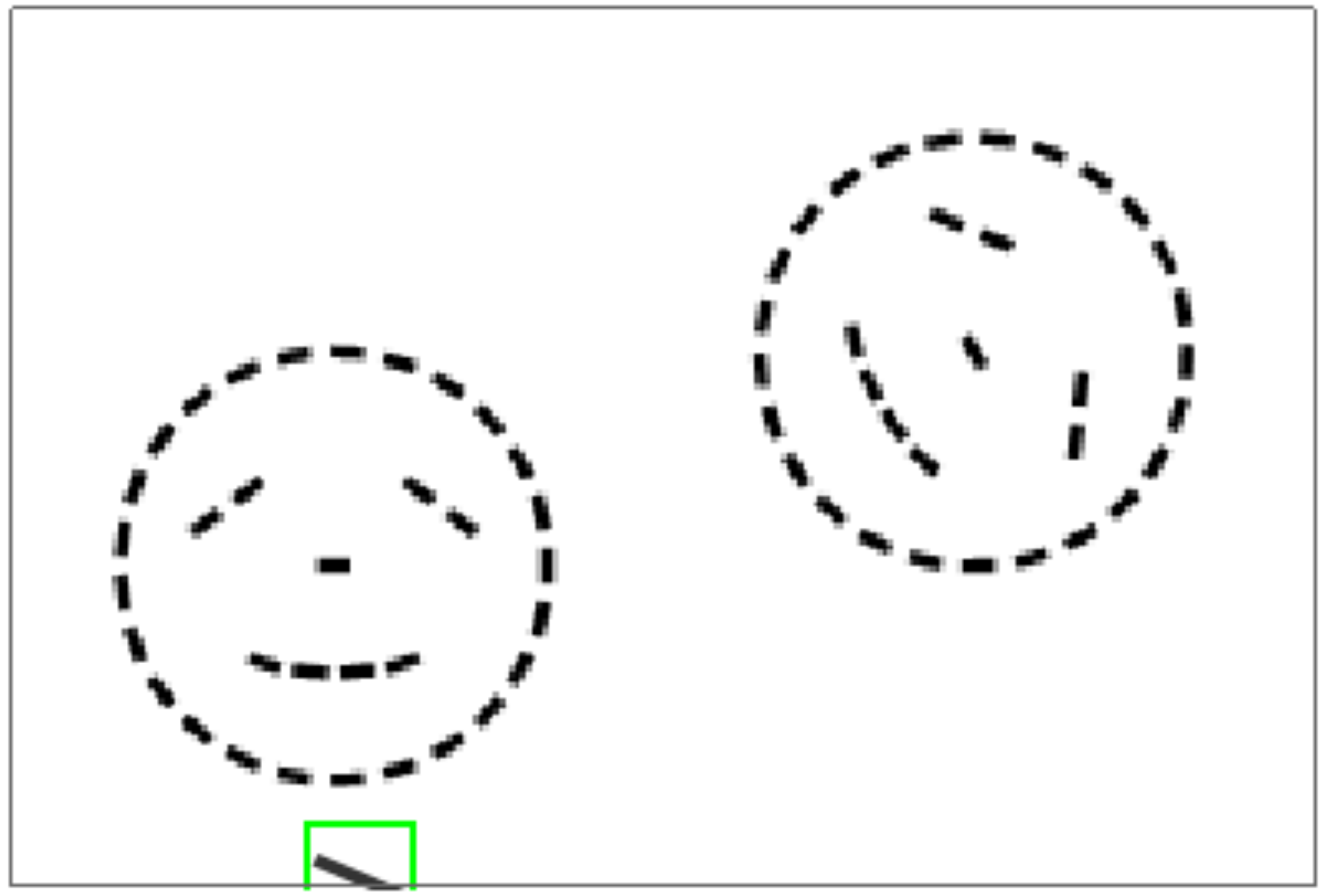
Representation of the roto-translation group!

$k(\mathbf{R}_\theta^{-1}(\mathbf{x}' - \mathbf{x}))$

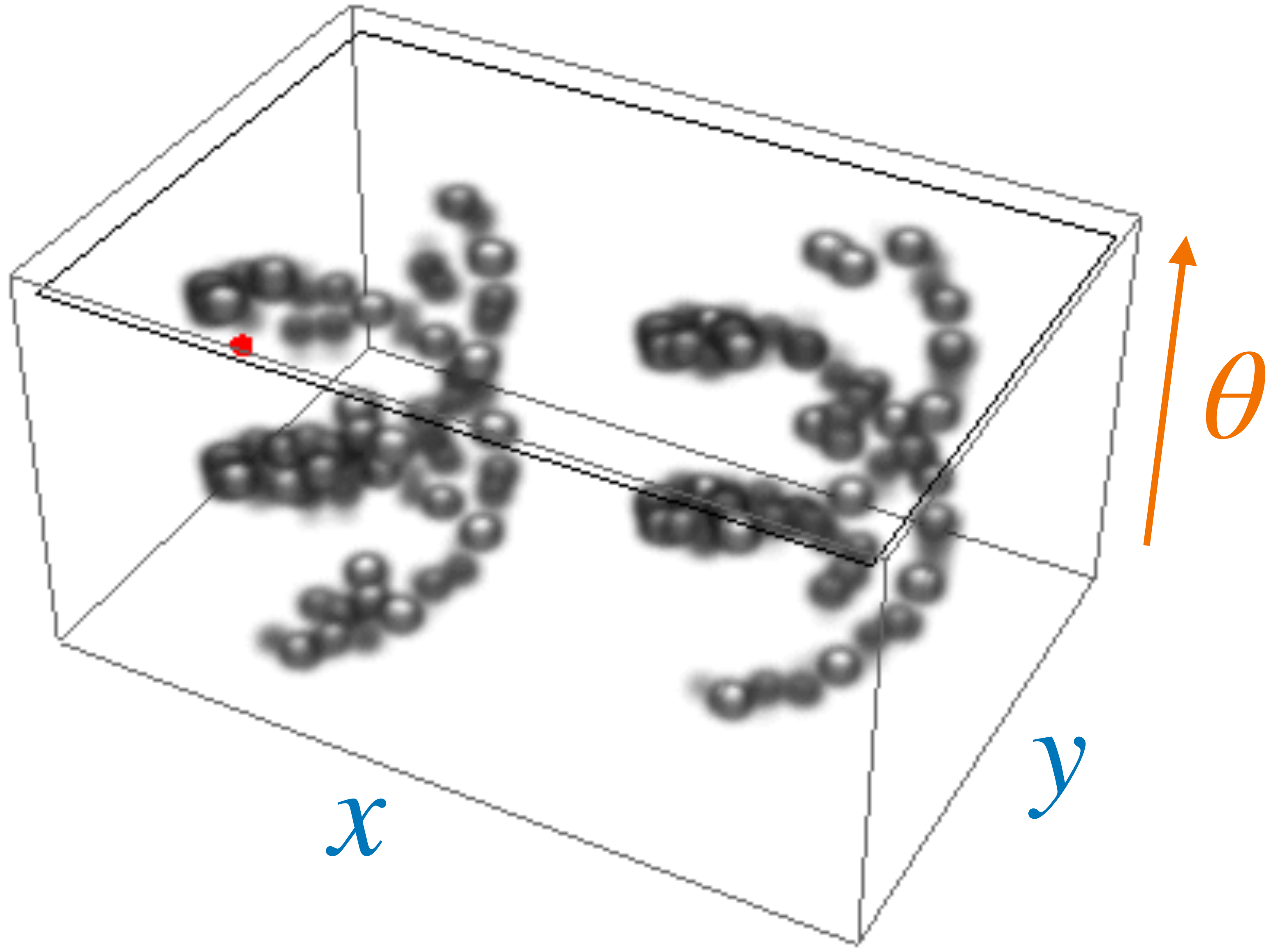
Lifting correlations: $(k \tilde{\star} f)(\mathbf{x}, \theta) = (\mathscr{L}_g^{SE(2) \to \mathbb{L}_2(\mathbb{R}^2)} k \,, f)_{\mathbb{L}_2(\mathbb{R}^2)} = (\mathscr{L}_{\mathbf{x}}^{\mathbb{R}^2 \to \mathbb{L}_2(\mathbb{R}^2)} \mathscr{L}_\theta^{SO(2) \to \mathbb{L}_2(\mathbb{R}^2)} k \,, f)_{\mathbb{L}_2(\mathbb{R}^2)}$

translation *rotation*

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



$=$



$\mathscr{L}_\theta^{SO(2) \to \mathbb{L}_2(\mathbb{R}^2)} k$

Rotated 2D convolution kernel

f^{in}

2D feature map

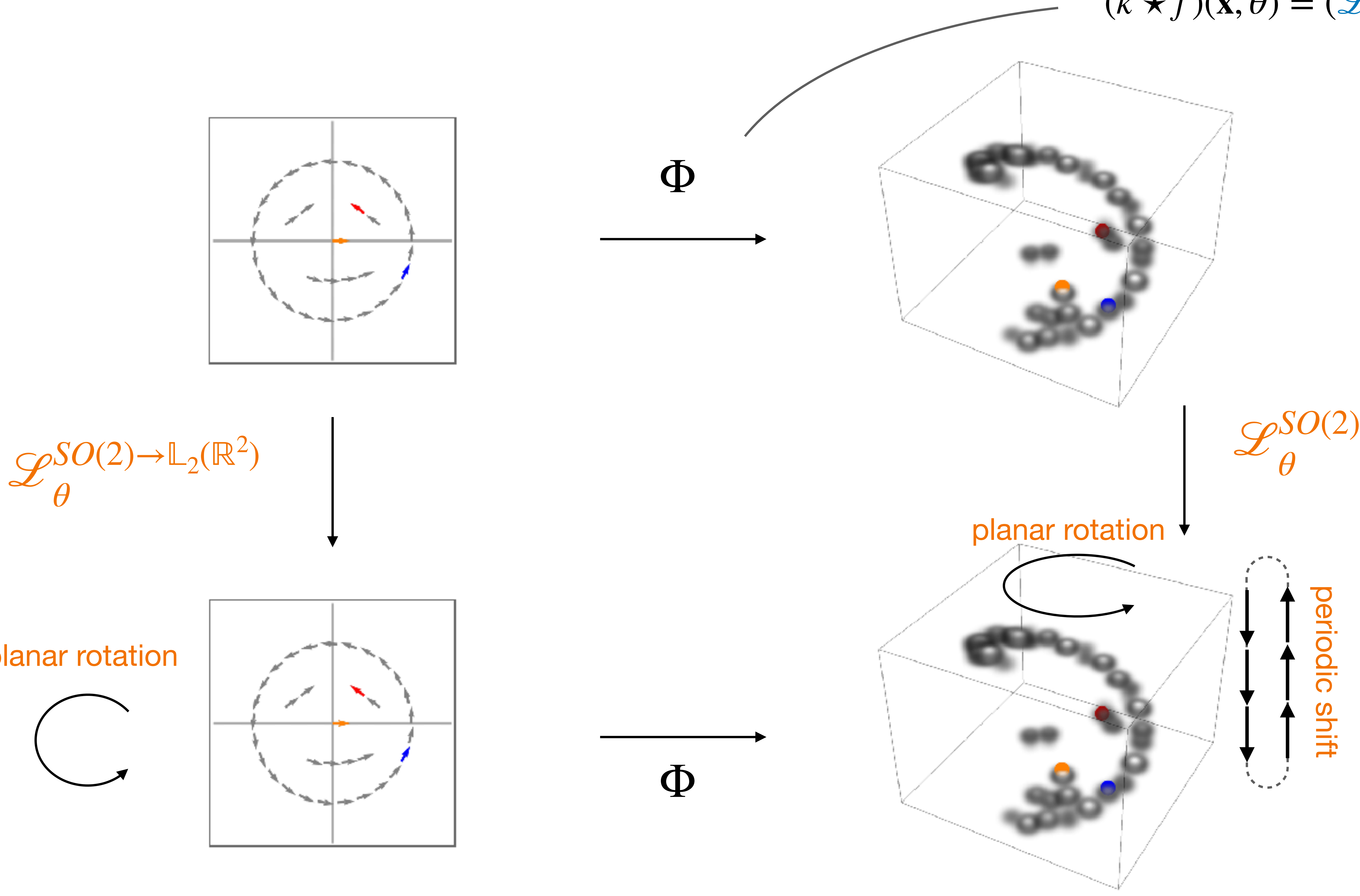
f^{out}

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

Equivariance

SE(2) group **lifting convolutions** are roto-translation equivariant

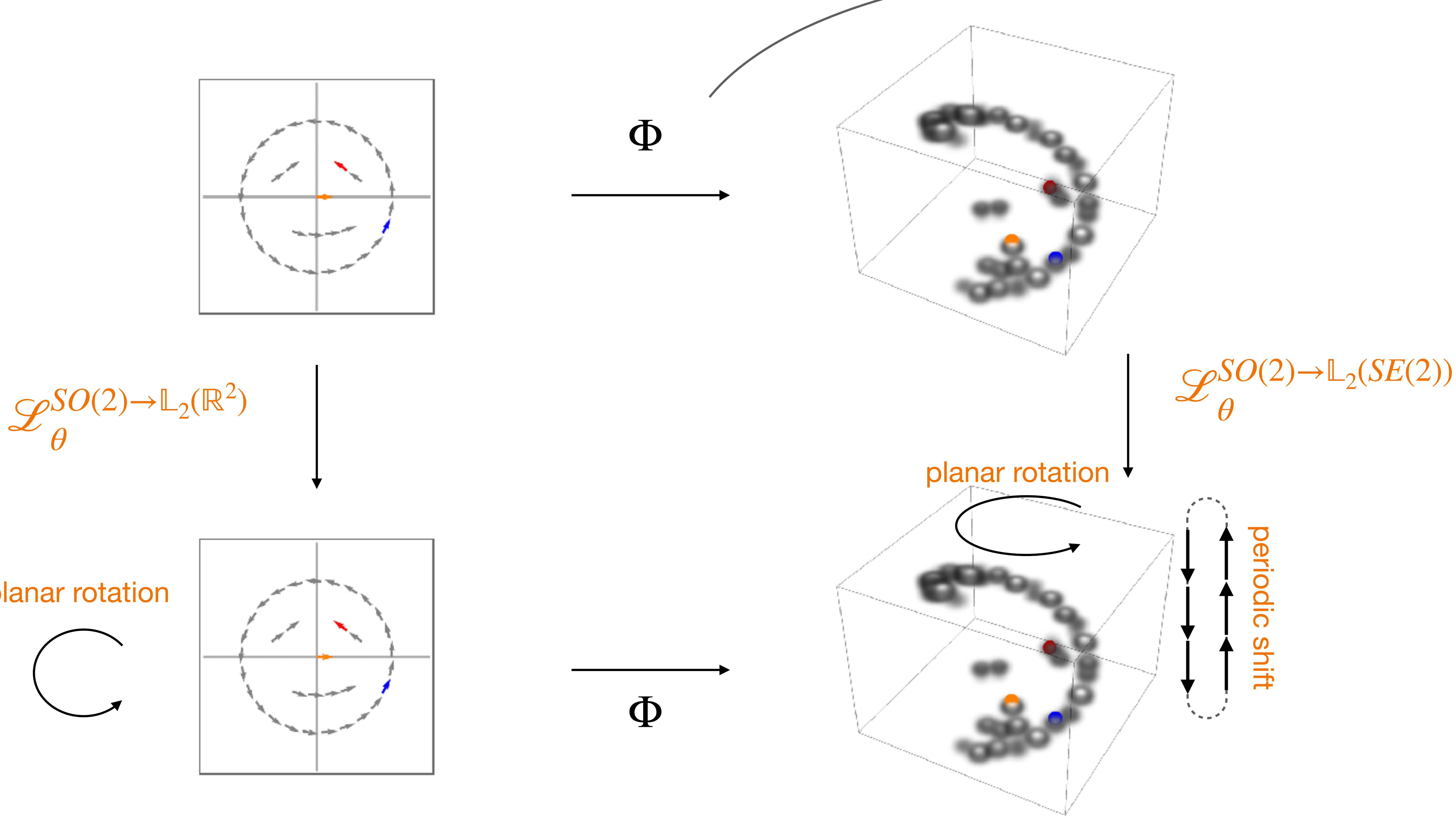
$$(k \,\tilde{\star}\, f)(\mathbf{x},\theta) = (\mathscr{L}_{\mathbf{x}}^{\mathbf{R}^2\to\mathbb{L}_2(\mathbb{R}^2)} \mathscr{L}_{\theta}^{SO(2)\to\mathbb{L}_2(\mathbb{R}^2)}\, k \,,\, f)_{\mathbb{L}_2(\mathbb{R}^2)}$$



Equivariance

SE(2) group **lifting convolutions** are roto-translation equivariant

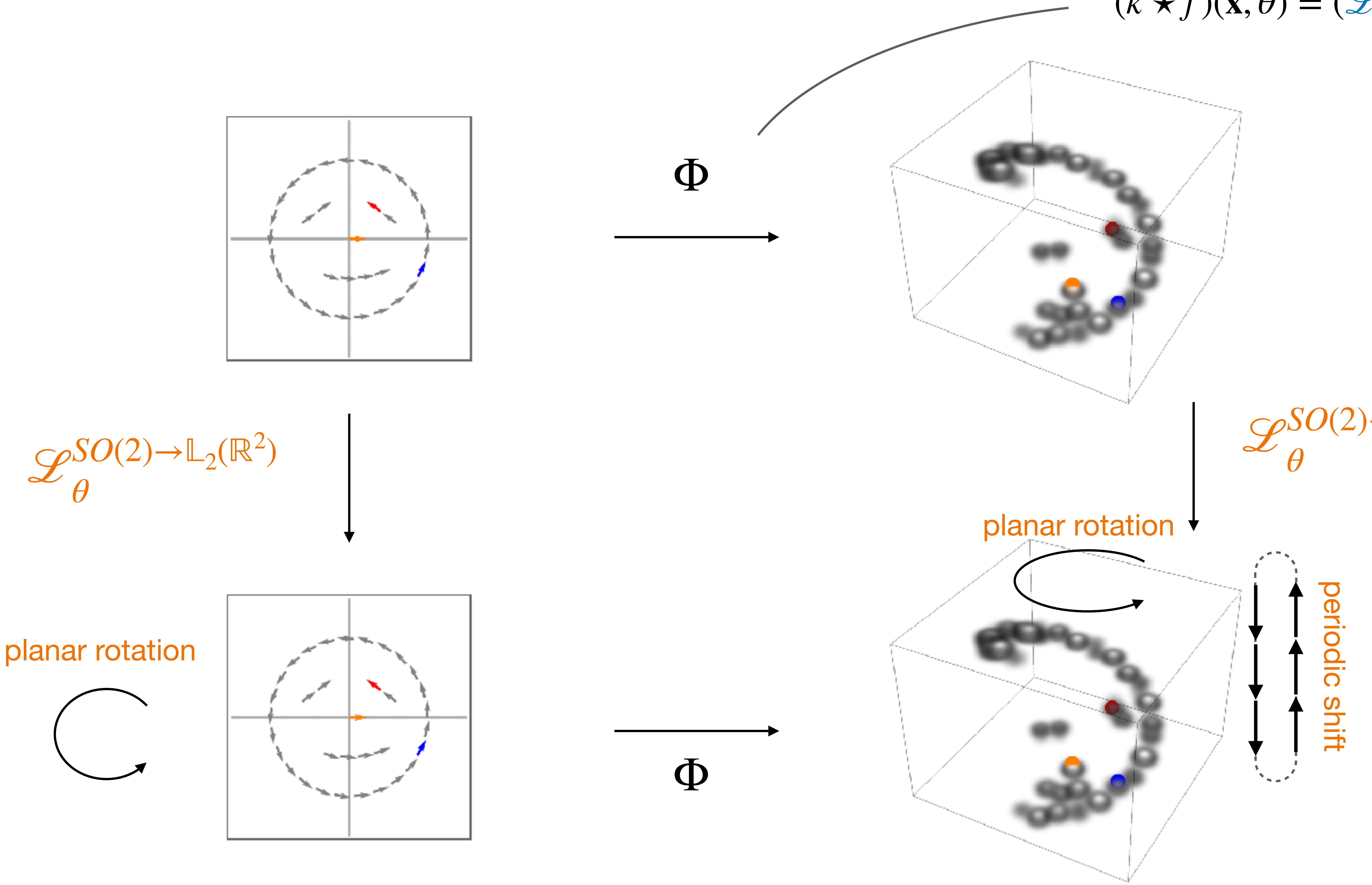
$$(k \,\tilde{\star}\, f)(\mathbf{x},\theta) = (\mathcal{L}_{\mathbf{x}}^{\mathbf{R}^2\to\mathbb{L}_2(\mathbb{R}^2)}\, \mathcal{L}_{\theta}^{SO(2)\to\mathbb{L}_2(\mathbb{R}^2)}\, k \,,\, f)_{\mathbb{L}_2(\mathbb{R}^2)}$$



Equivariance

SE(2) group **lifting convolutions** are roto-translation equivariant

$$(k \,\tilde{\star}\, f)(\mathbf{x},\theta) = (\mathscr{L}_{\mathbf{x}}^{\mathbf{R}^2\to\mathbb{L}_2(\mathbb{R}^2)} \mathscr{L}_\theta^{SO(2)\to\mathbb{L}_2(\mathbb{R}^2)} k \,,\, f)_{\mathbb{L}_2(\mathbb{R}^2)}$$



What about subsequent layers?

SE(2) equivariant cross-correlations

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

Group correlations:

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

translation *rotation*

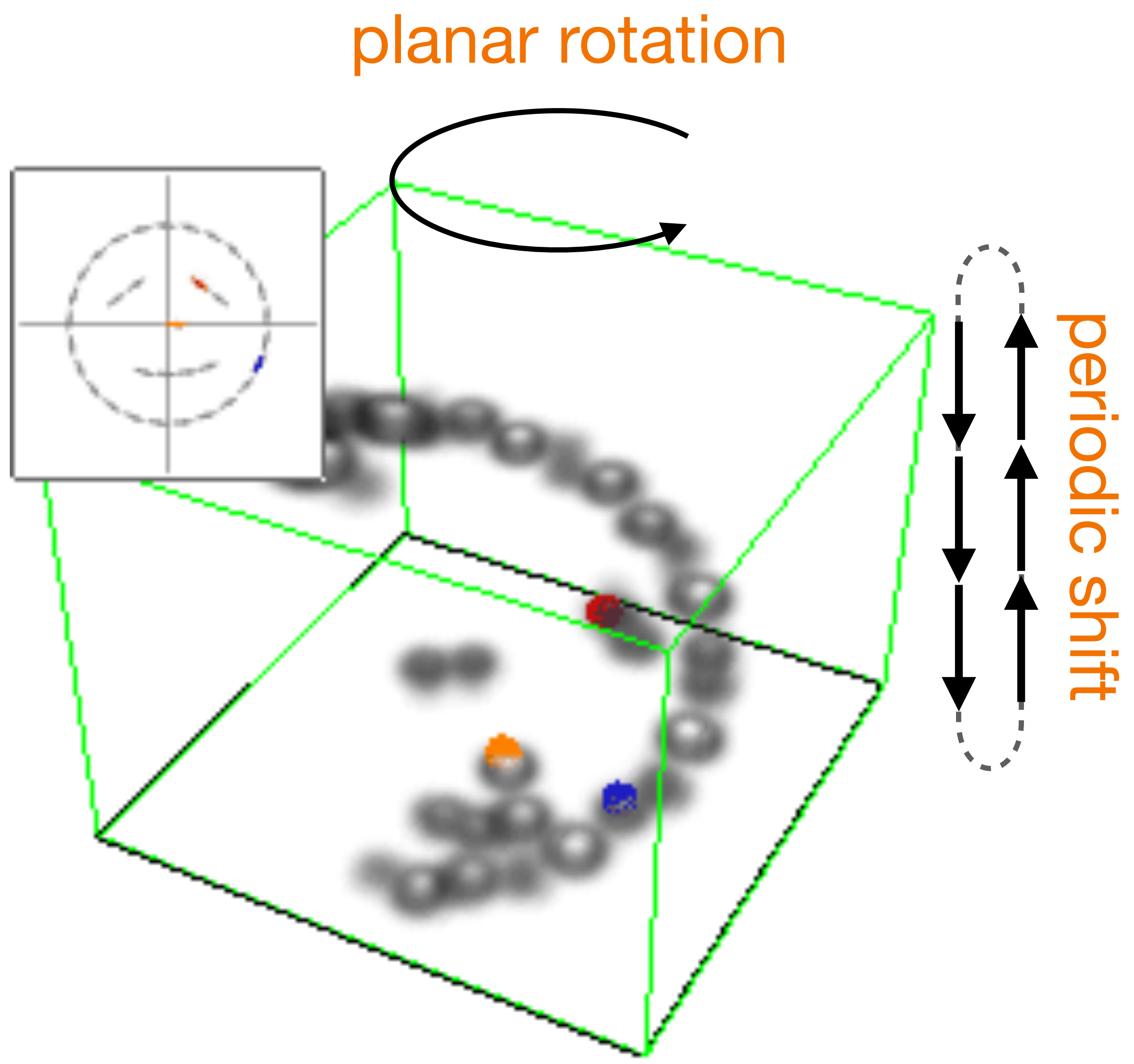
SE(2) equivariant cross-correlations

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

Group correlations:

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

translation *rotation*



$\mathscr{L}_\theta^{SO(2)\to\mathbb{L}_2(SE(2)}k$

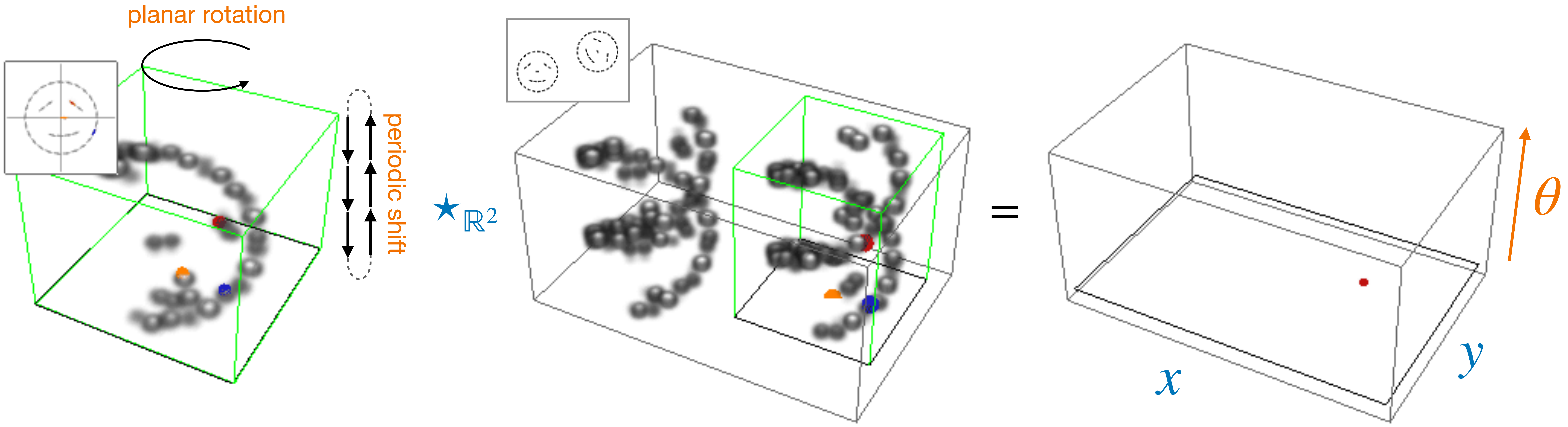
Rotated SE(2) convolution kernel

SE(2) equivariant cross-correlations

Group correlations:

$$(k \star f)(\mathbf{x},\theta) = (\mathcal{L}_g^{SE(2)\to\mathbb{L}_2(SE(2))}\, k \,,\, f)_{\mathbb{L}_2(SE(2))} = (\overbrace{\mathcal{L}_{\mathbf{x}}^{\mathbb{R}^2\to\mathbb{L}_2(SE(2))}\, \mathcal{L}_{\theta}^{SO(2)\to\mathbb{L}_2(SE(2))}\, k}^{k(\mathbf{R}_\theta^{-1}(\mathbf{x}'-\mathbf{x}),\mathbf{R}_{\theta'-\theta})} \,,\, f)_{\mathbb{L}_2(SE(2)}$$

translation *rotation*



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

f^{in}
SE(2) feature map

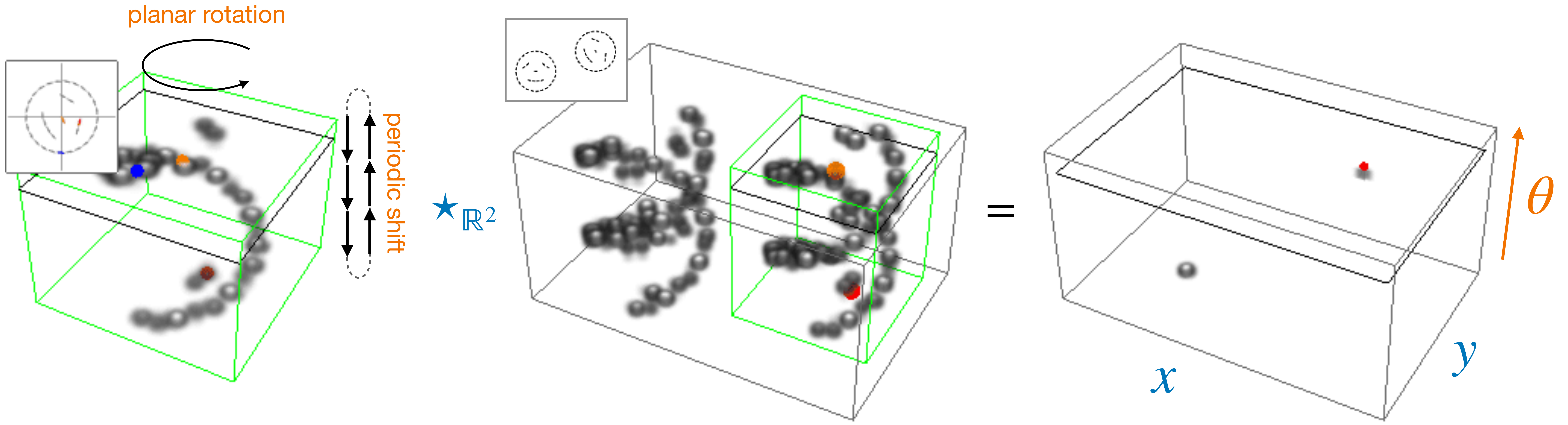
f^{out}
SE(2) feature map (after ReLU)

SE(2) equivariant cross-correlations

Group correlations:

$$(k \star f)(\mathbf{x},\theta) = (\mathscr{L}_g^{SE(2)\to\mathbb{L}_2(SE(2))}\, k\,,\, f)_{\mathbb{L}_2(SE(2))} = (\overbrace{\mathscr{L}_{\mathbf{x}}^{\mathbf{R}^2\to\mathbb{L}_2(SE(2))}\,\mathscr{L}_{\theta}^{SO(2)\to\mathbb{L}_2(SE(2))}\, k}^{k(\mathbf{R}_\theta^{-1}(\mathbf{x}'-\mathbf{x}),\mathbf{R}_{\theta'-\theta})}\,,\, f)_{\mathbb{L}_2(SE(2)}$$

translation *rotation*



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

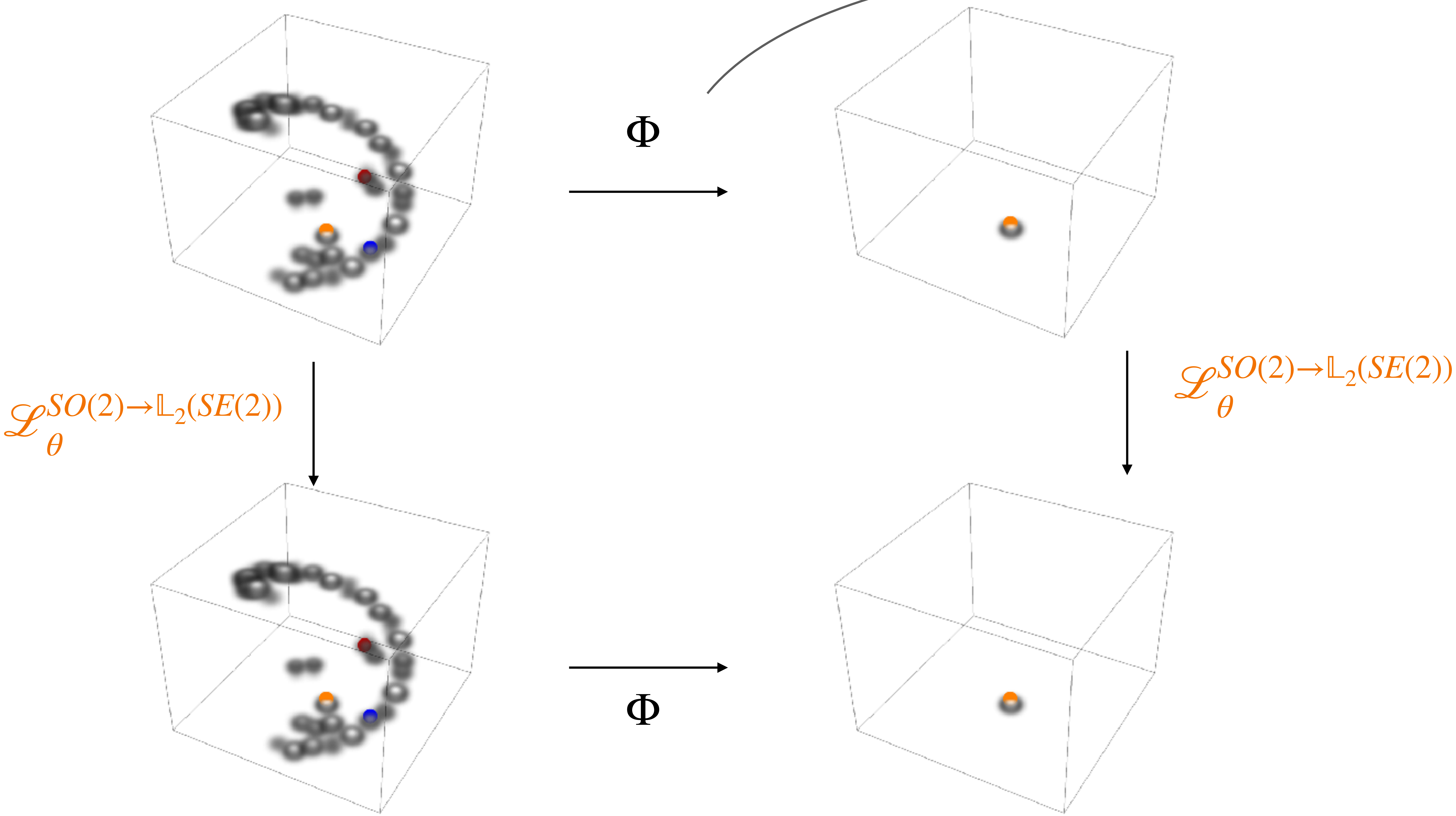
f^{in}
SE(2) feature map

f^{out}
SE(2) feature map (after ReLU)

Equivariance

SE(2) group convolutions are roto-translation equivariant

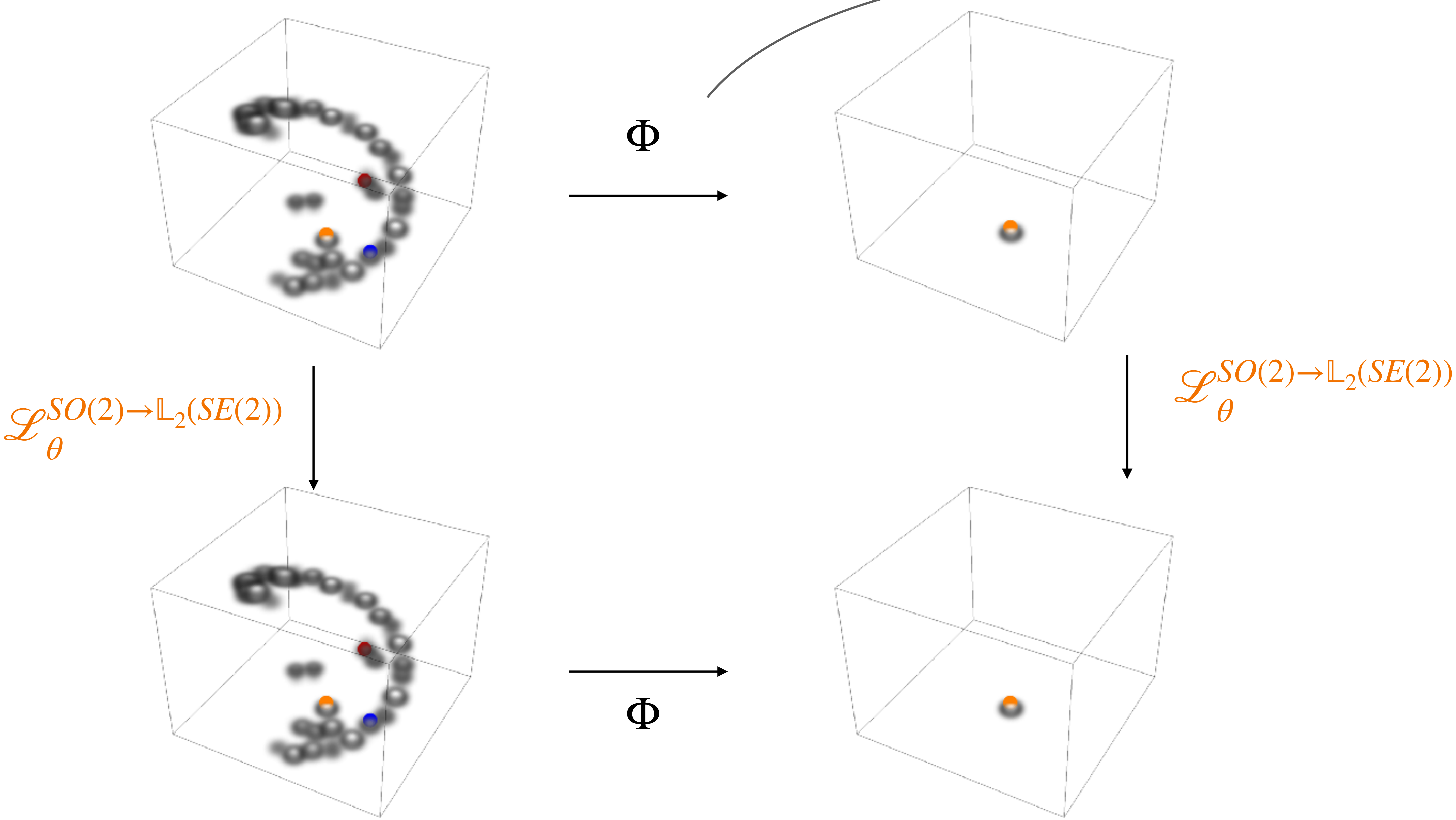
$$(k \star f)(\mathbf{x},\theta) = (\mathscr{L}_{\mathbf{x}}^{\mathbb{R}^2\to\mathbb{L}_2(SE(2))} \mathscr{L}_{\theta}^{SO(2)\to\mathbb{L}_2(SE(2))}\, k\, ,\, f)_{\mathbb{L}_2(SE(2))}$$



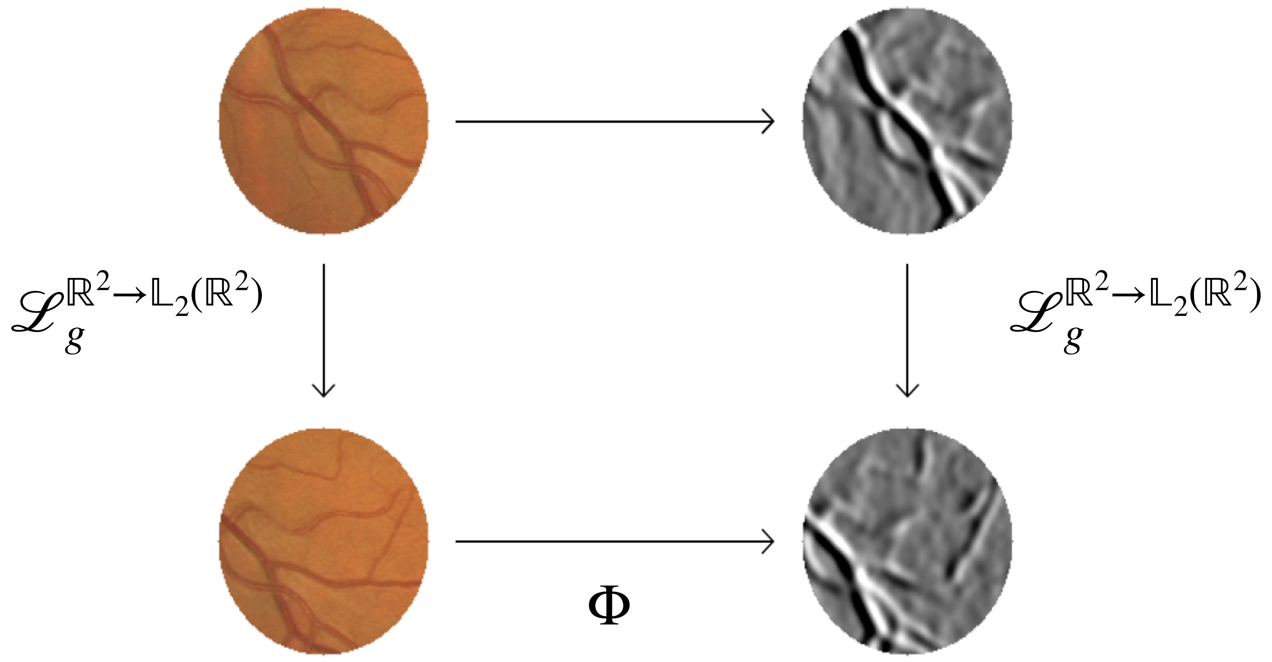
Equivariance

SE(2) group convolutions are roto-translation equivariant

$$(k \star f)(\mathbf{x},\theta) = (\mathcal{L}_{\mathbf{x}}^{\mathbb{R}^2\to\mathbb{L}_2(SE(2))} \mathcal{L}_{\theta}^{SO(2)\to\mathbb{L}_2(SE(2))} k \, , f)_{\mathbb{L}_2(SE(2))}$$



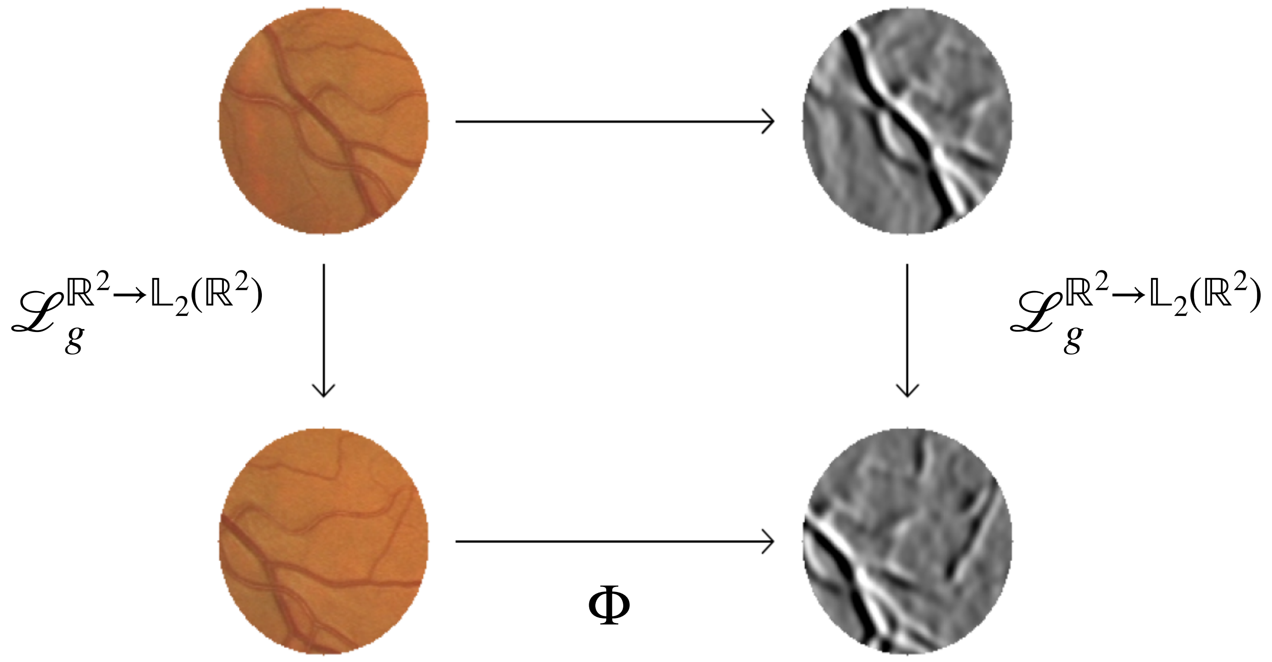
2D cross-correlation (translation equivariant)



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

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

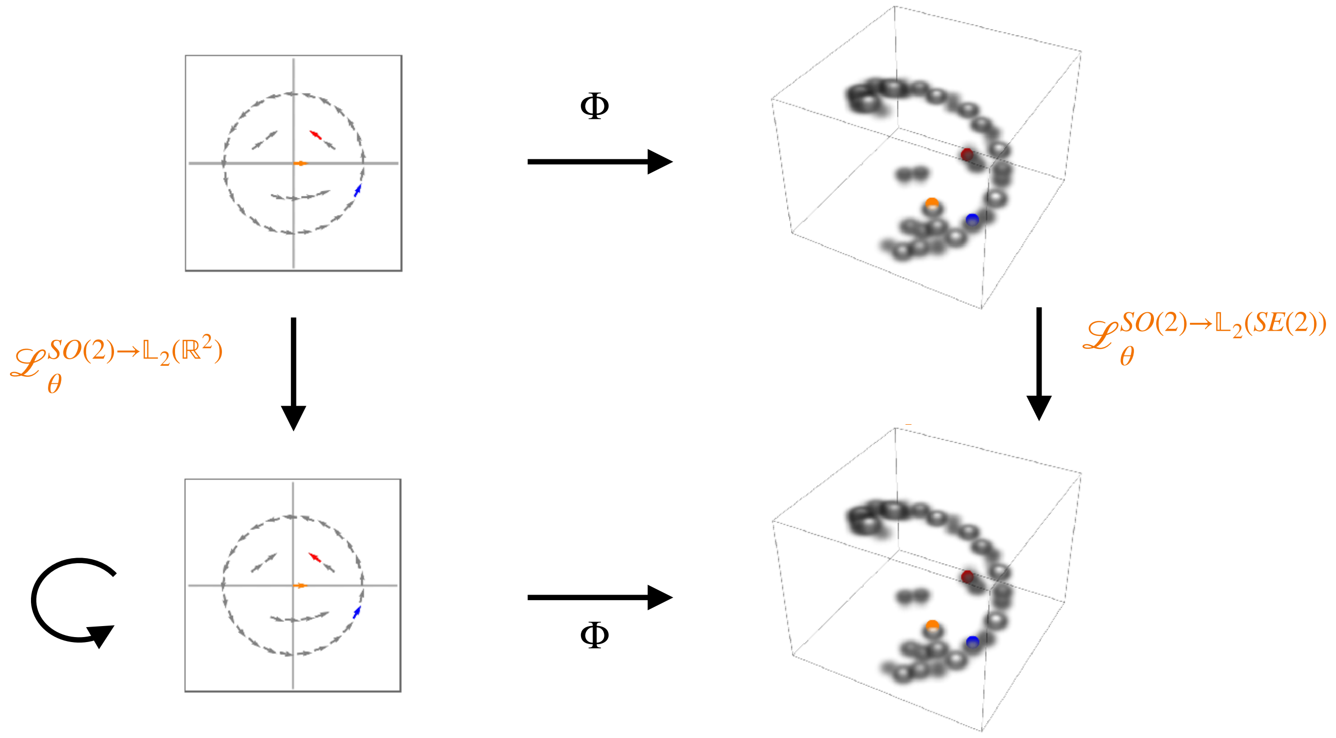
2D cross-correlation (translation equivariant)



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

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

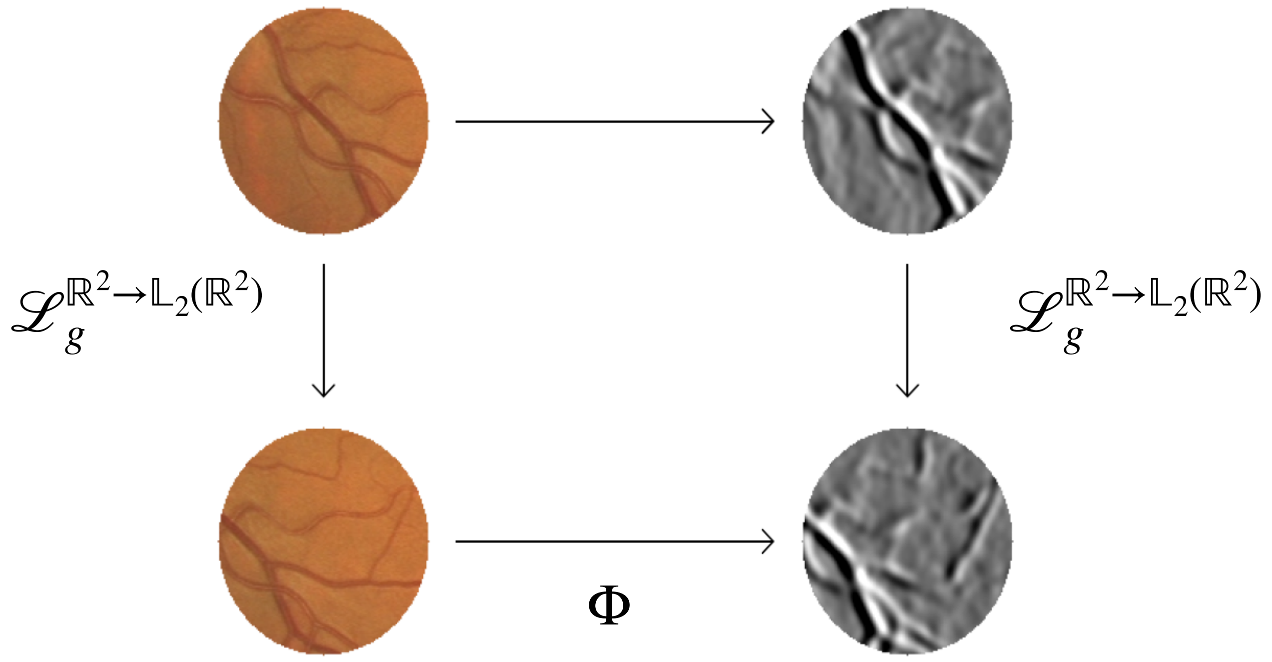
SE(2) lifting correlations (roto-translation equivariant)



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

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

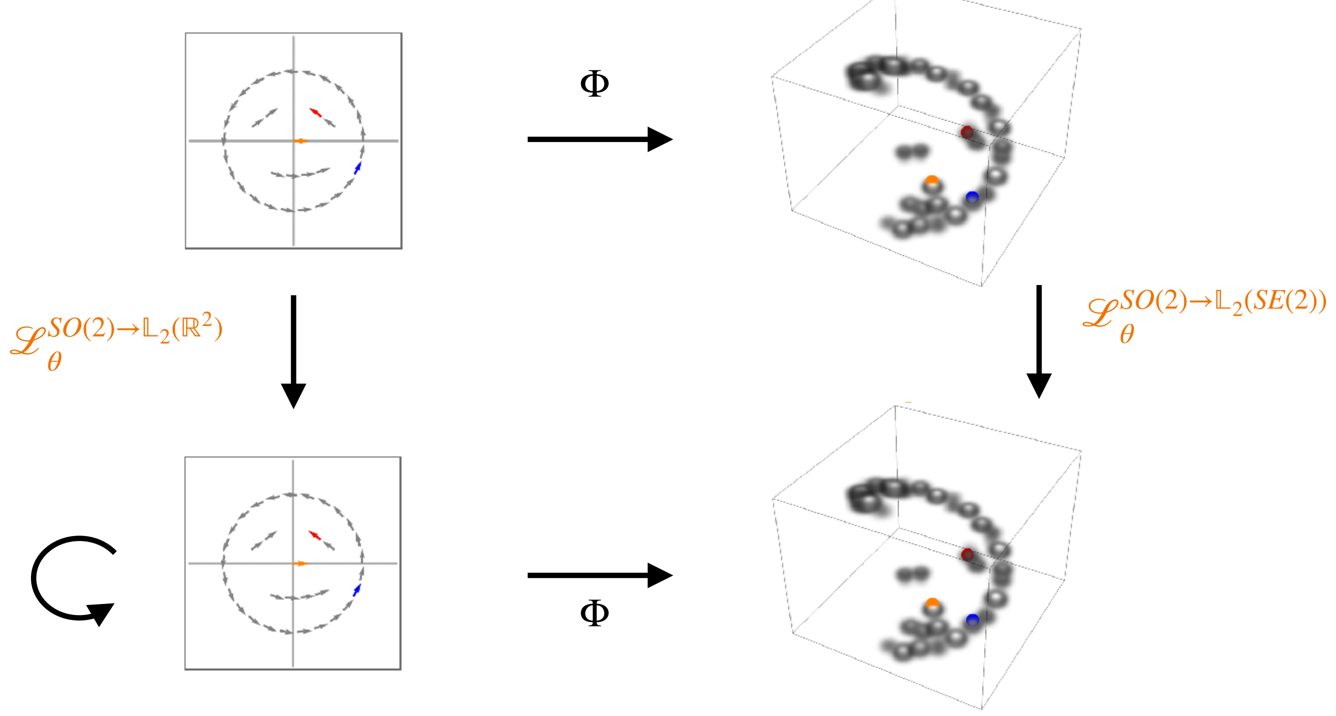
2D cross-correlation (translation equivariant)



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

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

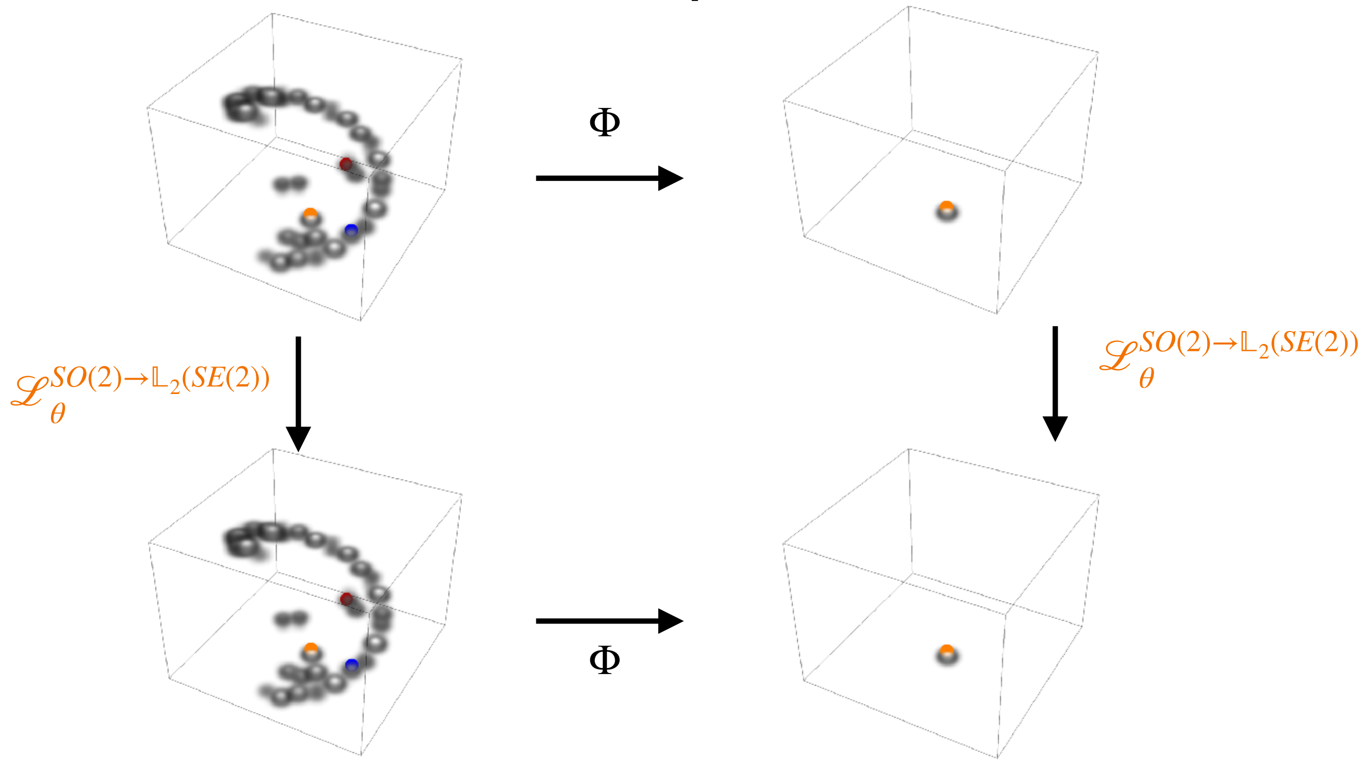
SE(2) lifting correlations (roto-translation equivariant)



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

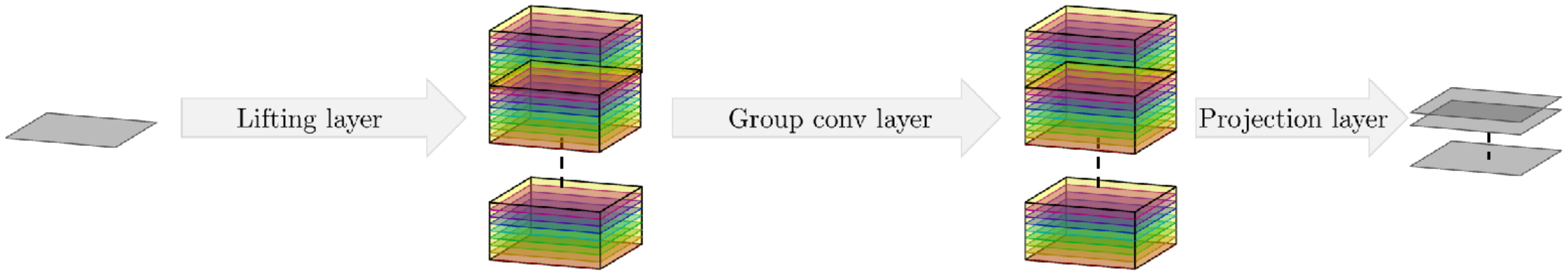
$$= \int_{\mathbb{R}^2} k(\mathbf{R}_\theta^{-1}(\mathbf{x}' - \mathbf{x})) f(\mathbf{x}') d\mathbf{x}'$$

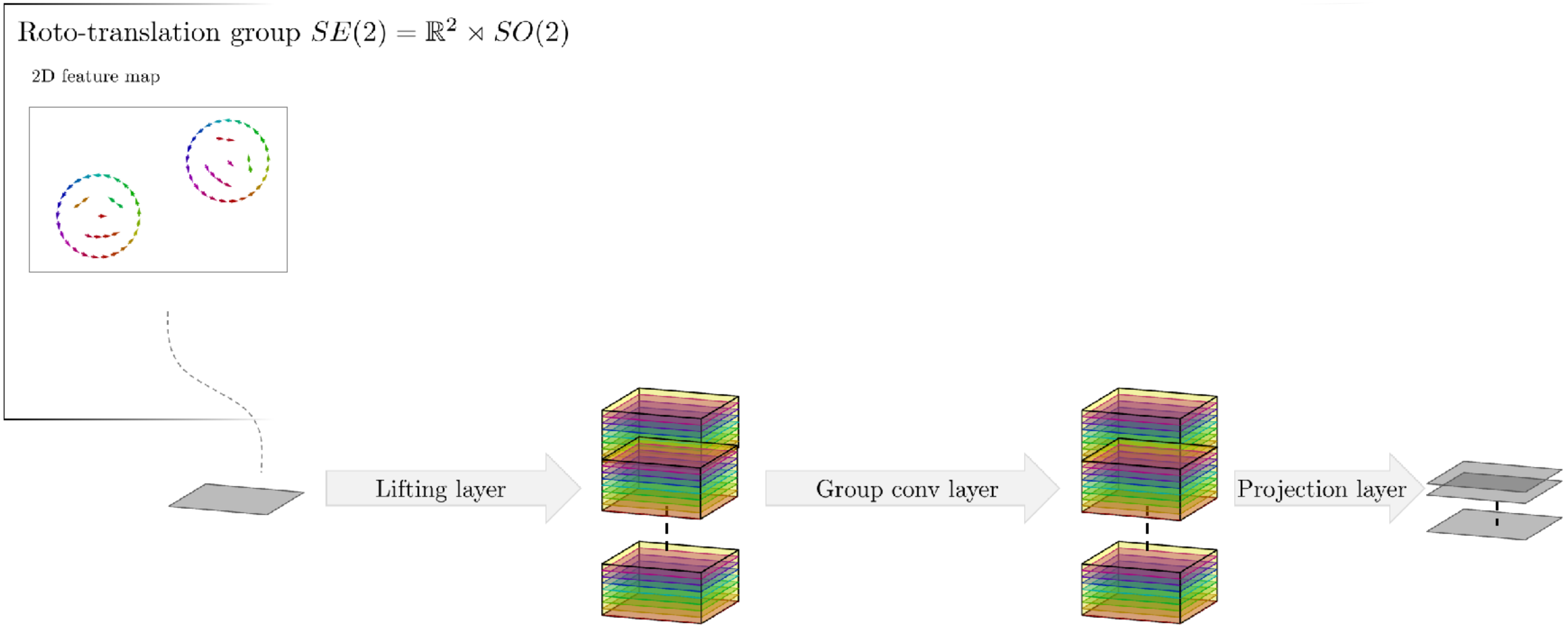
SE(2) G-correlations (roto-translation equivariant)

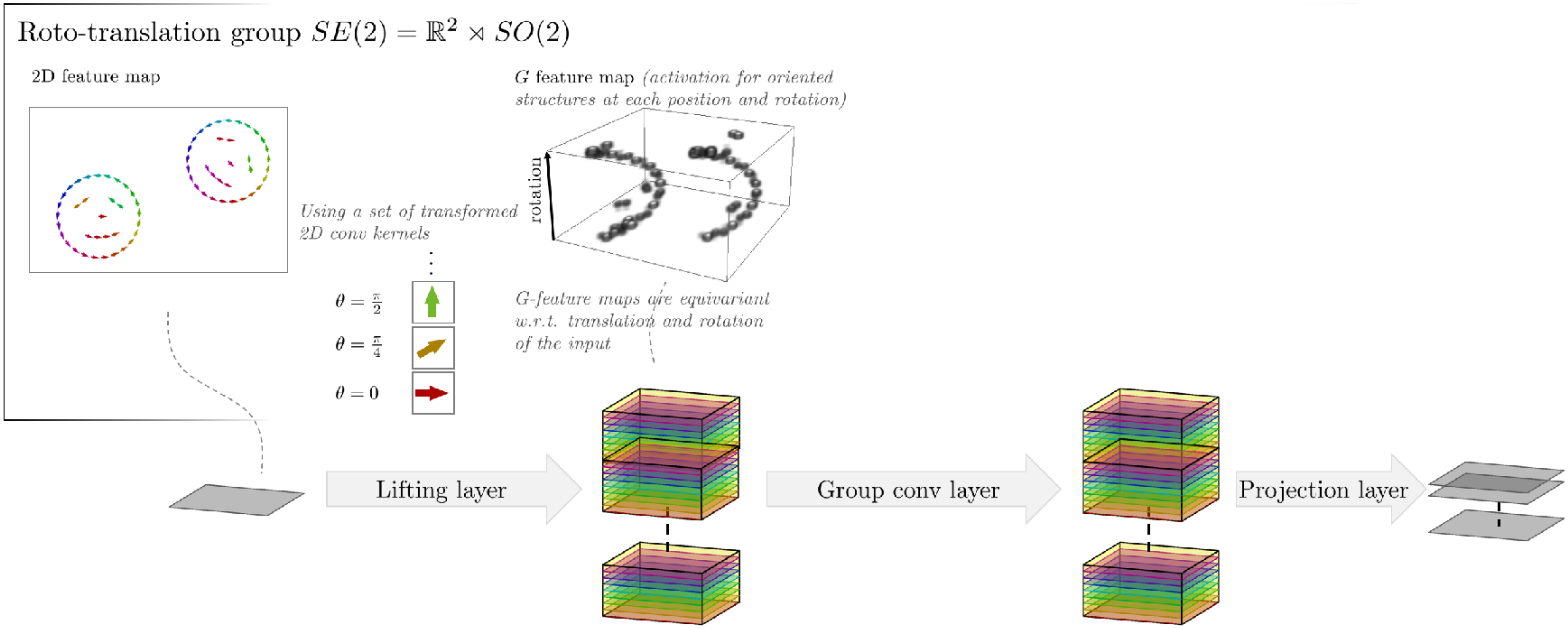


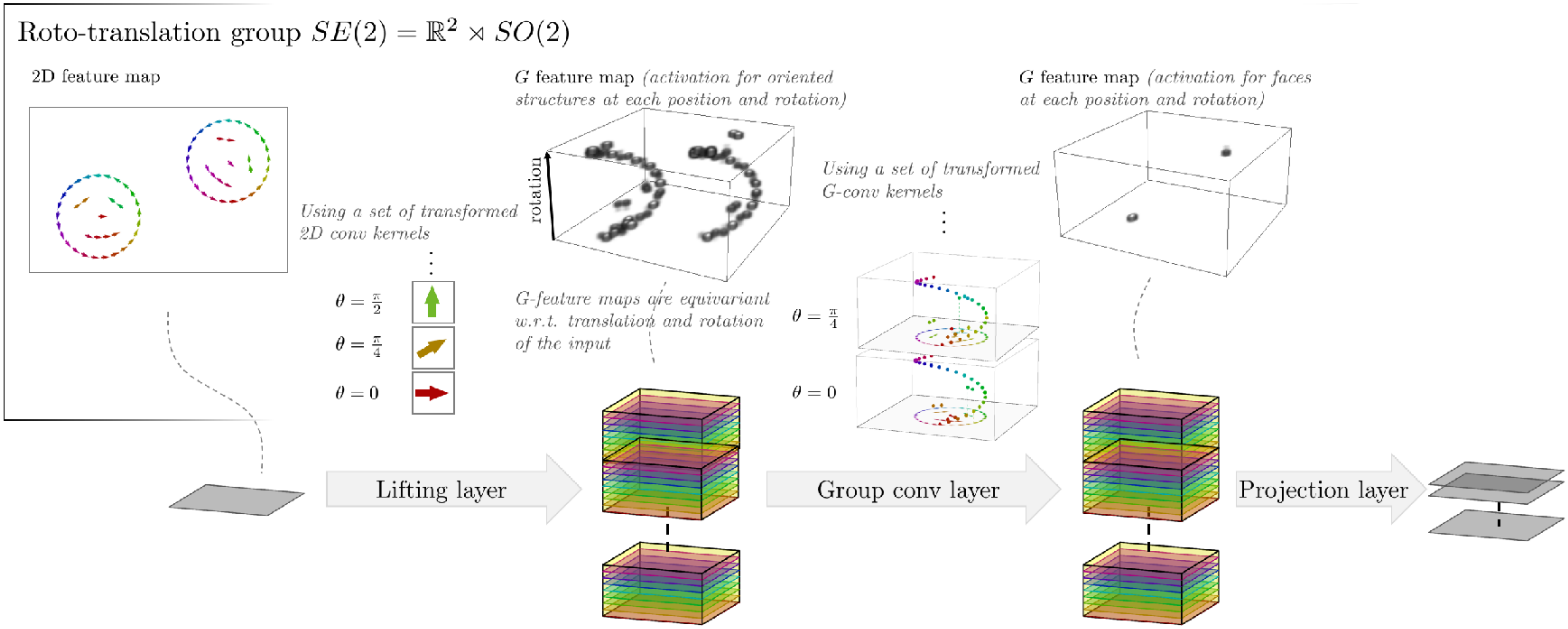
$$(k \tilde{\star} f)(\mathbf{x}, \theta) = (\mathcal{L}_g^{SE(2)\to\mathbb{L}_2(SE(2))} k, f)_{\mathbb{L}_2(SE(2))}$$

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

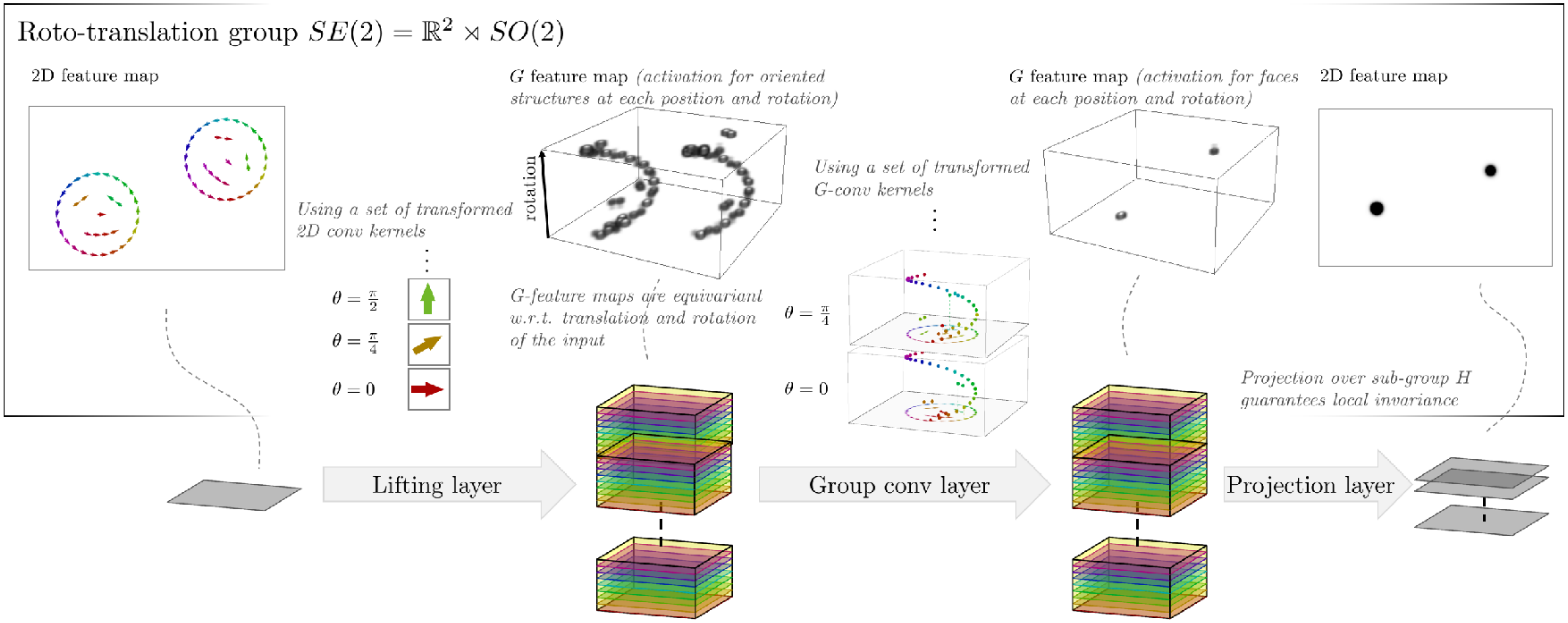

Lifting layer
Group conv layer
Projection layer

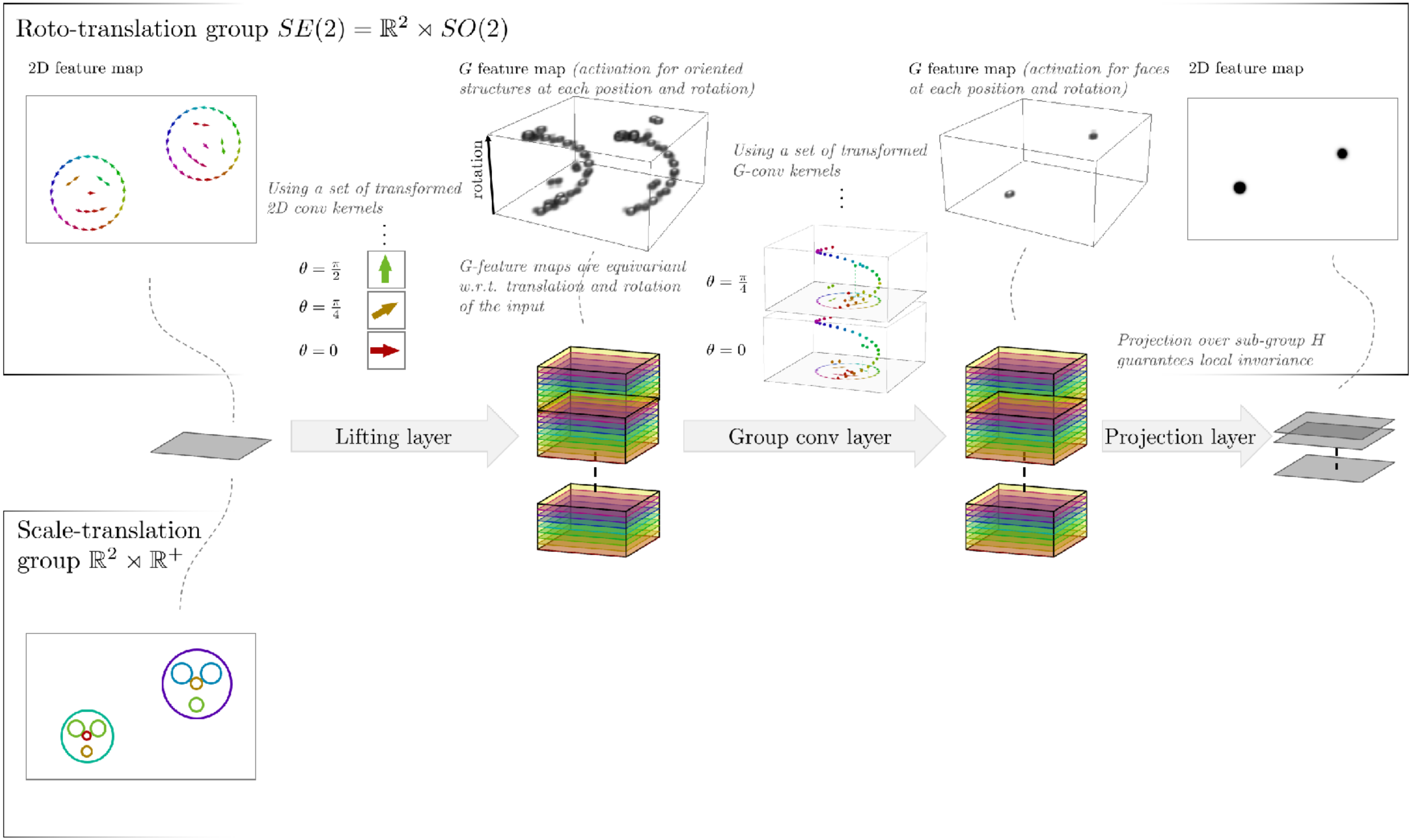

Roto-translation group $SE(2) = \mathbb{R}^2 \rtimes SO(2)$
2D feature map
Lifting layer
Group conv layer
Projection layer


Roto-translation group $SE(2) = \mathbb{R}^2 \rtimes SO(2)$
2D feature map
G feature map (activation for oriented structures at each position and rotation)
rotation
Using a set of transformed 2D conv kernels
$\theta = \frac{\pi}{2}$
$\theta = \frac{\pi}{4}$
$\theta = 0$
G-feature maps are equivariant w.r.t. translation and rotation of the input
Lifting layer
Group conv layer
Projection layer

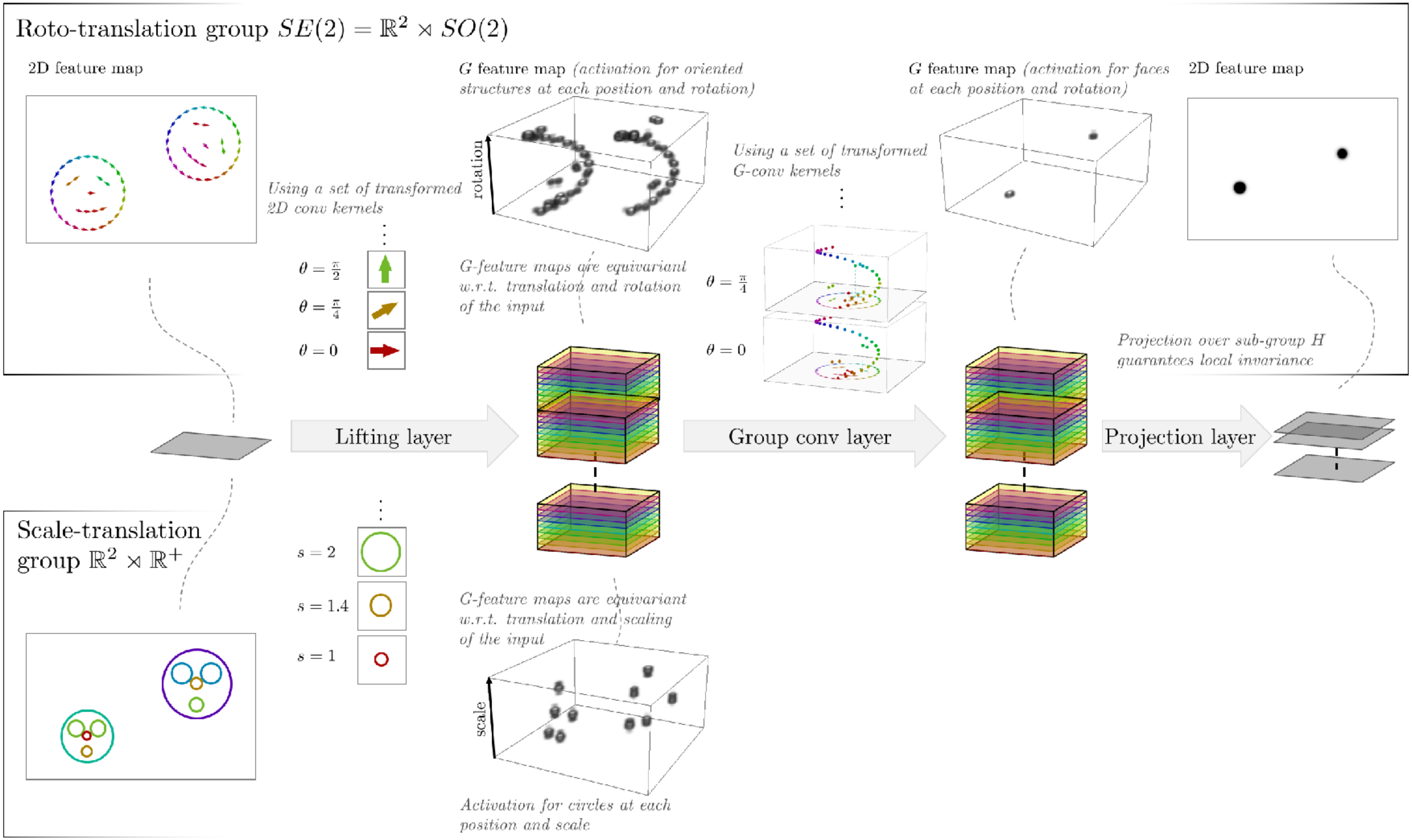


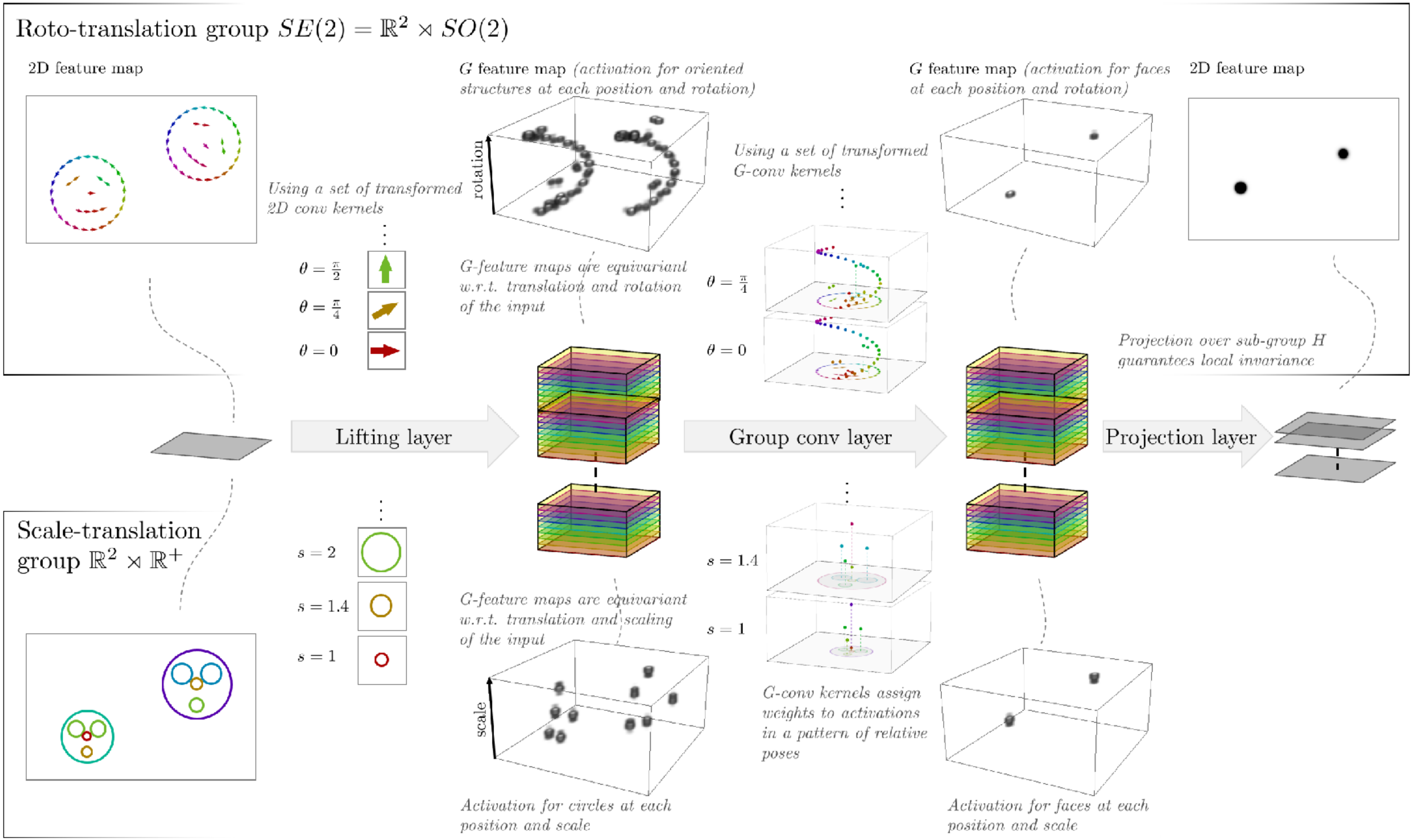
Roto-translation group $SE(2) = \mathbb{R}^2 \rtimes SO(2)$
2D feature map
Using a set of transformed 2D conv kernels
$\theta = \frac{\pi}{2}$
$\theta = \frac{\pi}{4}$
$\theta = 0$
G feature map (activation for oriented structures at each position and rotation)
rotation
G-feature maps are equivariant w.r.t. translation and rotation of the input
Using a set of transformed G-conv kernels
$\theta = \frac{\pi}{4}$
$\theta = 0$
G feature map (activation for faces at each position and rotation)
Lifting layer
Group conv layer
Projection layer

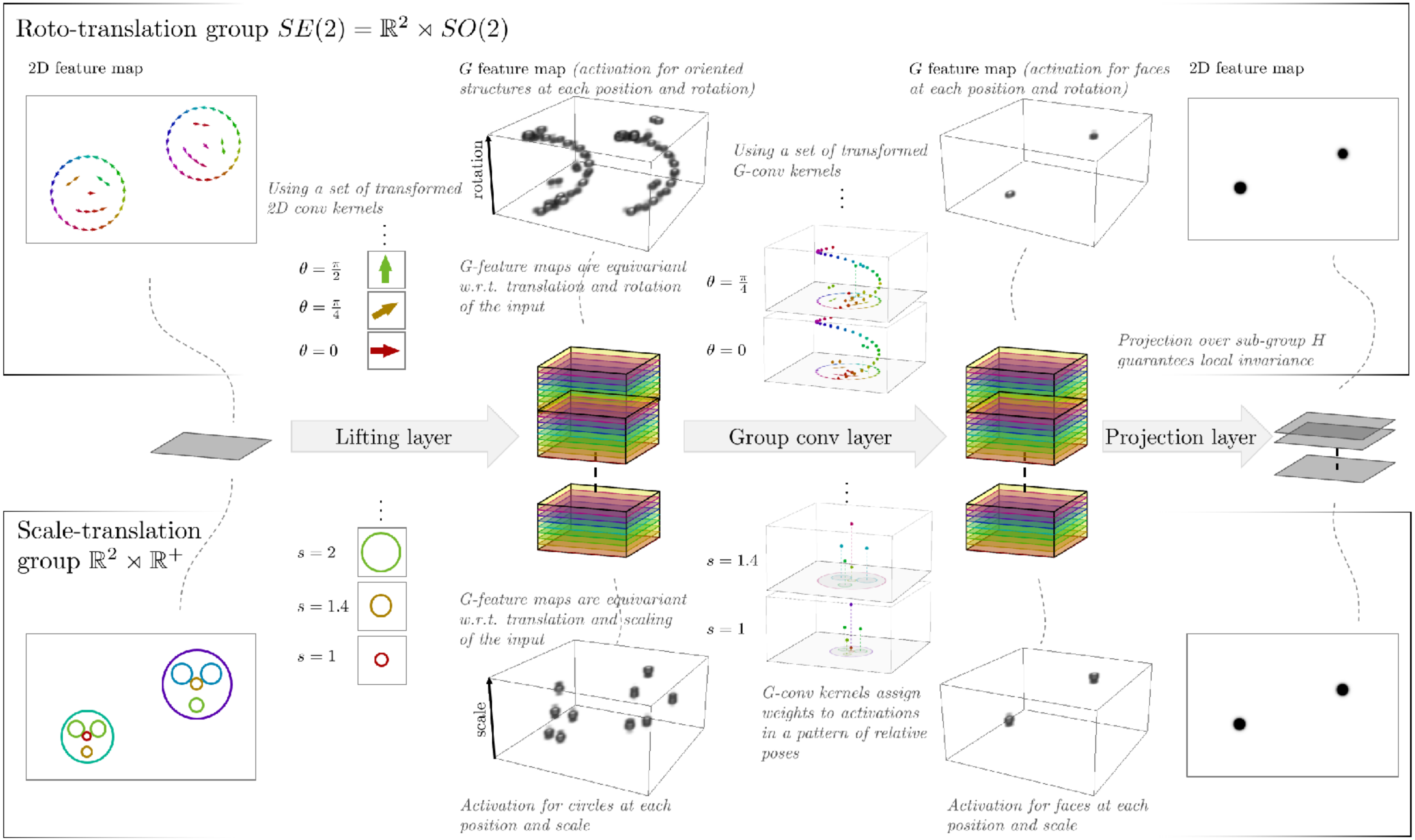

Roto-translation group $SE(2) = \mathbb{R}^2 \rtimes SO(2)$
2D feature map
Using a set of transformed 2D conv kernels
$\theta = \frac{\pi}{2}$
$\theta = \frac{\pi}{4}$
$\theta = 0$
G feature map (activation for oriented structures at each position and rotation)
rotation
G-feature maps are equivariant w.r.t. translation and rotation of the input
Using a set of transformed G-conv kernels
$\theta = \frac{\pi}{4}$
$\theta = 0$
G feature map (activation for faces at each position and rotation)
2D feature map
Projection over sub-group H guarantees local invariance
Lifting layer
Group conv layer
Projection layer



Roto-translation group $SE(2) = \mathbb{R}^2 \rtimes SO(2)$
2D feature map
Using a set of transformed 2D conv kernels
$\theta = \frac{\pi}{2}$
$\theta = \frac{\pi}{4}$
$\theta = 0$
G feature map (activation for oriented structures at each position and rotation)
rotation
G-feature maps are equivariant w.r.t. translation and rotation of the input
Using a set of transformed G-conv kernels
$\theta = \frac{\pi}{4}$
$\theta = 0$
G feature map (activation for faces at each position and rotation)
2D feature map
Projection over sub-group H guarantees local invariance
Lifting layer
Group conv layer
Projection layer
Scale-translation group $\mathbb{R}^2 \rtimes \mathbb{R}^+$


Roto-translation group $SE(2) = \mathbb{R}^2 \rtimes SO(2)$
2D feature map
Using a set of transformed 2D conv kernels
$\theta = \frac{\pi}{2}$
$\theta = \frac{\pi}{4}$
$\theta = 0$
G feature map (activation for oriented structures at each position and rotation)
rotation
G-feature maps are equivariant w.r.t. translation and rotation of the input
Using a set of transformed G-conv kernels
$\theta = \frac{\pi}{4}$
$\theta = 0$
G feature map (activation for faces at each position and rotation)
2D feature map
Projection over sub-group H guarantees local invariance
Lifting layer
Group conv layer
Projection layer
Scale-translation group $\mathbb{R}^2 \rtimes \mathbb{R}^+$
$s = 2$
$s = 1.4$
$s = 1$
G-feature maps are equivariant w.r.t. translation and scaling of the input
scale
Activation for circles at each position and scale

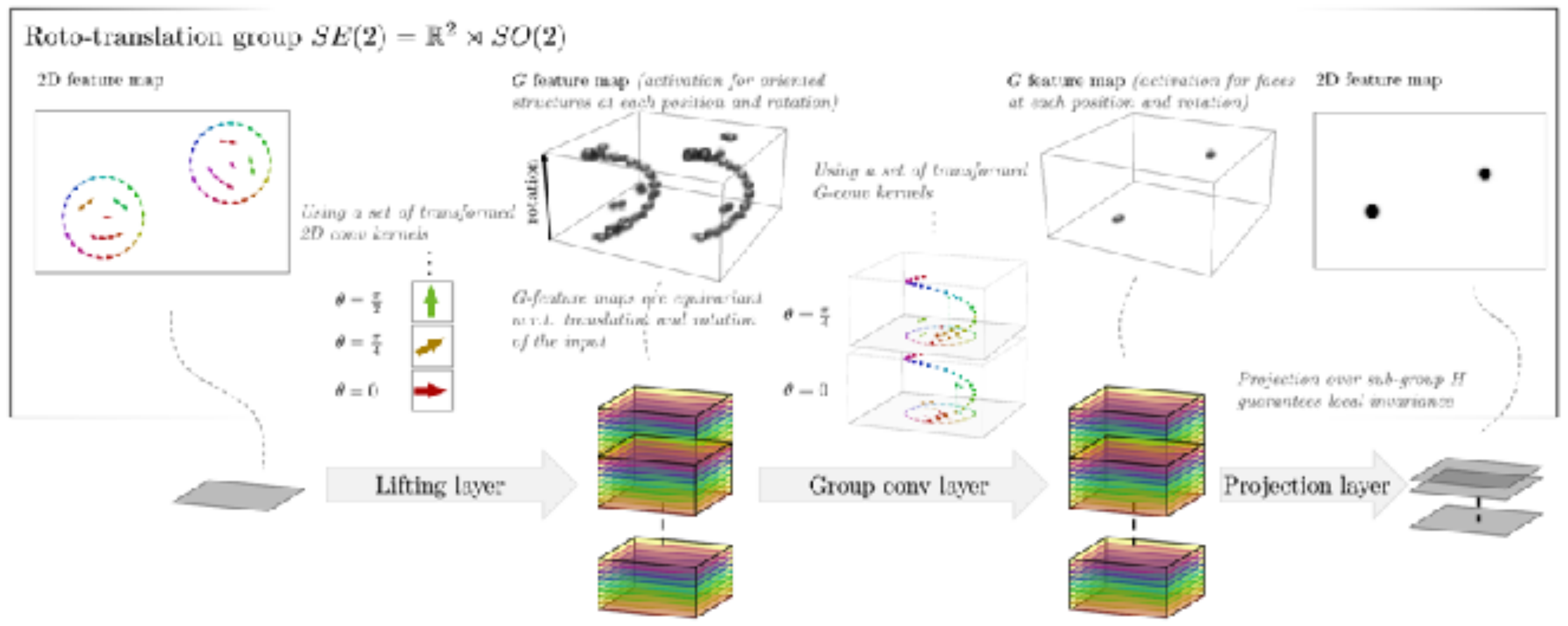

Roto-translation group $SE(2) = \mathbb{R}^2 \rtimes SO(2)$
2D feature map
G feature map (activation for oriented structures at each position and rotation)
rotation
G feature map (activation for faces at each position and rotation)
2D feature map
Using a set of transformed 2D conv kernels
Using a set of transformed G-conv kernels
$\theta = \frac{\pi}{2}$
$\theta = \frac{\pi}{4}$
$\theta = 0$
G-feature maps are equivariant w.r.t. translation and rotation of the input
$\theta = \frac{\pi}{4}$
$\theta = 0$
Projection over sub-group H guarantees local invariance
Lifting layer
Group conv layer
Projection layer
Scale-translation group $\mathbb{R}^2 \rtimes \mathbb{R}^+$
$s = 2$
$s = 1.4$
$s = 1$
G-feature maps are equivariant w.r.t. translation and scaling of the input
$s = 1.4$
$s = 1$
scale
G-conv kernels assign weights to activations in a pattern of relative poses
Activation for circles at each position and scale
Activation for faces at each position and scale



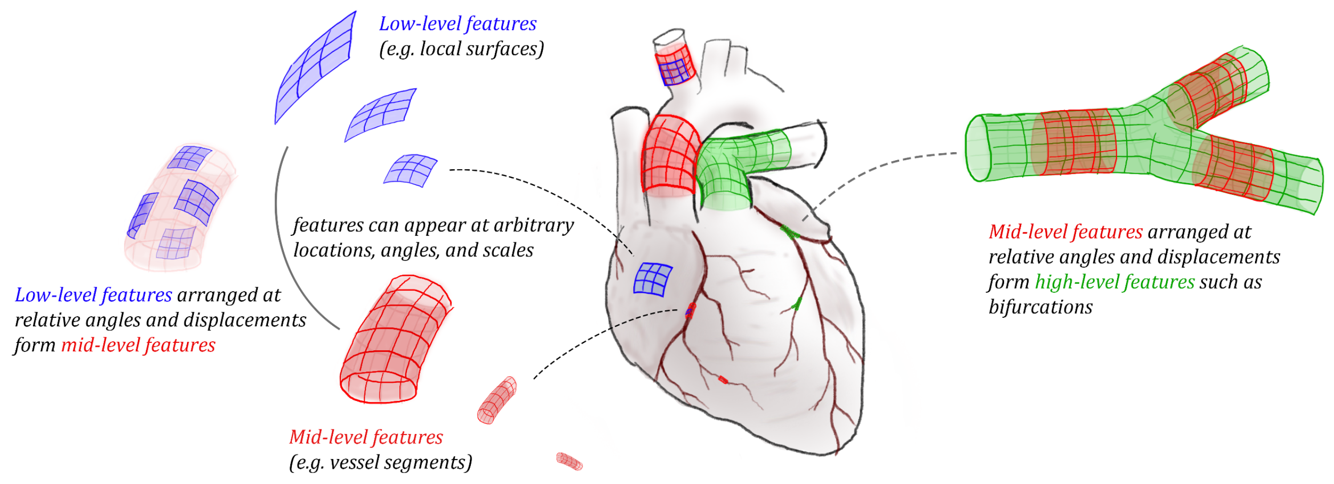
Roto-translation group $SE(2) = \mathbb{R}^2 \rtimes SO(2)$
2D feature map
G feature map (activation for oriented structures at each position and rotation)
rotation
G feature map (activation for faces at each position and rotation)
2D feature map
Using a set of transformed 2D conv kernels
Using a set of transformed G-conv kernels
$\theta = \frac{\pi}{2}$
$\theta = \frac{\pi}{4}$
$\theta = 0$
G-feature maps are equivariant w.r.t. translation and rotation of the input
$\theta = \frac{\pi}{4}$
$\theta = 0$
Projection over sub-group H guarantees local invariance
Lifting layer
Group conv layer
Projection layer
Scale-translation group $\mathbb{R}^2 \rtimes \mathbb{R}^+$
$s = 2$
$s = 1.4$
$s = 1$
$s = 1.4$
$s = 1$
G-feature maps are equivariant w.r.t. translation and scaling of the input
scale
G-conv kernels assign weights to activations in a pattern of relative poses
Activation for circles at each position and scale
Activation for faces at each position and scale

Summary

- Group convolutional neural networks intuitively perform template matching
- A template (kernel) is transformed and matched (inner-product) under all possible transformations in the group
- This creates higher-dimensional feature maps (functions on the group) on which we again define template matching via the group action
- In these higher dimensional feature maps we can detect advanced patterns in terms of features at relative poses!
- G-CNNs are based on equivariant layers (thus weight sharing) and guarantee invariance through pooling



recognition by components



leverage symmetries