Roto-Translation Equivariant Super-Resolution of Two-Dimensional Flows Using Convolutional Neural Networks

Yuki Yasuda

(*Electronic mail: yasuda.y.aa@m.titech.ac.jp)

(Dated: 24 February 2022)

Convolu-ational neural networks (CNNs) often process vectors as quantities having no direction like colors in images. This study investigates the effect of treating vectors as geometrical objects in terms of super-resolution of velocity on two-dimensional fluids. Vector is distinguished from scalar by the transformation law associated with a change in basis, which can be incorporated as the prior knowledge using the equivariant deep learning. We convert existing CNNs into equivariant ones by making each layer equivariant with respect to rotation and translation. The training data in the low- and high-resolution are generated with the downsampling or the spectral nudging. When the data inherit the rotational symmetry, the equivariant CNNs show comparable accuracy with the non-equivariant ones. Since the number of parameters is smaller in the equivariant CNNs, these models are trainable with a smaller size of the data. In this case, the transformation law of vector should be incorporated as the prior knowledge, where vector is explicitly treated as a quantity having direction. Two examples demonstrate that the symmetry of the data can be broken. In the first case, a downsampling method makes the correspondence between low- and high-resolution patterns dependent on the orientation. In the second case, the input data are insufficient to recognize the rotation of coordinates in the experiment with the spectral nudging. In both cases, the accuracy of the CNNs deteriorates if the equivariance is forced to be imposed, and the usage of conventional CNNs may be justified even though vector is processed as a quantity having no direction.

I. INTRODUCTION

In recent years, neural networks (NNs) have been actively applied to the field of fluid mechanics. The physical validity of inference of NNs is important not only in theory but also in application. One method to enhance the validity is to explicitly incorporate physics laws such as the Navier-Stokes equations into NNs, which is called the physics-informed neural network and has been extensively studied these days.

Vector is an essential quantity in physics. In applications of NNs to fluids, however, vectors are often processed as direct products of scalars such as colors in images, and convolutional neural networks (CNNs) are employed as in computer vision. Intuitively, vector has direction, whereas scalar does not. This distinction can be mathematically formulated by covariance in geometry. The covariance describes how the components of a tensor (e.g., scalar and vector) change under a change of basis associated with some coordinate transformation. These transformation laws distinguish the rank of tensors and can be regarded as a part of the definition of scalars and vectors. In other words, by changing the direction of observation through a coordinate transformation, one can determine whether a numerical array is a combination of scalars or a vector having a direction.

The covariance ensures that scalar and vector themselves are geometrically invariant objects, though their components may be changed under coordinate transformations. This fact leads to the invariance of physics laws described with scalars and vectors, that is, the forms of physics equations are independent of coordinate systems. If both covariance and physics laws are incorporated into NNs as the prior knowledge, the NNs can be applied to data on any coordinate system without losing accuracy and physical validity. Such NNs incorporating geometric symmetries have started to be applied to various systems including fluid systems.

Super-resolution (SR) refers to methods of estimating high-resolution images from low-resolution ones. Super-resolution is studied in computer vision as an application of NNs. The success of these NNs has led to an increasing number of studies on the fluid-related SR, idealized turbulent flows in two[23,24] and three dimensions[27,29,30,32,35] and Rayleigh–Bénard convection[33] flows in blood vessels[37] and atmospheric fields[38,39]. Physics laws such as the continuity equation can be taken into account as the prior knowledge, which makes super-resolved flows more accurate and physically valid[28,31]. However, SR models have not been investigated enough in terms of geometry, and the invariance of physics equations has not been fully exploited.

The covariance in super-resolution can be satisfied by imposing the equivariance on NNs. If an NN is equivariant, for instance, rotating the input leads to the output to be rotated in the same way (Fig. 1). Therefore, if the input is a vector field, the output satisfies the same transformation law, assuring that the output is a geometric vector field.

The success of conventional CNNs is attributable to the translation equivariance. CNNs with other types of equivariance (e.g., rotation) have been studied[49,55]. These CNNs can be understood within unified frameworks on the equivariant CNN[50,55] that are based on the group theory[56] and differential geometry[57]. The theories and techniques on the equivariant CNN are developed not only for the two-dimensional images but also for the three-dimensional volumetric data.
FIG. 1. Schematic illustration for the equivariance of super-resolution. If a super-resolution model is equivariant, the rotation is commutative with the super-resolution. The anti-clockwise rotation by 90° is considered here. The colors represent a scalar field and the arrows represent a vector field.

This study investigates the equivariance on the super-resolution of two-dimensional fluids by employing two existing CNNs by Fukami et al. and Bode et al. The equivariance to rotation and translation is taken into account by utilizing the theory and software library developed by Weiler and Cesa. We deal with two types of low- and high-resolution data. For the first type, only the high-resolution (HR) flow is generated from the fluid simulation, and the low-resolution (LR) one is obtained by downsampling. For the other type, the LR flow is first generated from the fluid simulation, and then the HR one is obtained from another simulation including a spectral nudging. The results presented here would be referred to when an SR model is built for climate data because large-scale flows in the ocean and atmosphere are regarded as two-dimensional (i.e., geostrophic flows).

This study demonstrates the following three things:

- The equivariance of super-resolution depends on the following: the intrinsic symmetry of fluid systems, the generation method of LR and HR data, and the choice of input data.

- When the super-resolution is equivariant, the accuracy of the equivariant CNNs is comparable to that of the non-equivariant ones. The equivariant CNNs are trainable with a smaller size of the data due to the reduction in the learnable parameters.

- When the super-resolution is not equivariant, the accuracy of the equivariant CNNs is lower than that of the non-equivariant ones. In this case, the use of conventional CNNs may be justified even though vector is treated as a quantity having no direction.

The present paper is organized as follows. Section II clarifies the covariance and equivariance and describes how to make the existing CNNs equivariant. Section II gives the methods of training and evaluating the CNNs. Section IV compares the equivariant and non-equivariant CNNs when the data have the rotational symmetry. Section V discusses two examples where the data do not have the rotational symmetry and the super-resolution is not equivariant. The conclusions are presented in Section VI. The source code is available on the GitHub repository.
II. ROTOTO-TRANSLATION EQUIVARIANT CNNS

We clarify the mathematical notations and terms used in this study and then make a discussion on equivariant CNNs. The covariance and equivariance are defined in Section II A. The non-equivariant and equivariant CNNs are presented in Section II B.

The present study treats the two-dimensional Euclidean space using the Cartesian coordinate system: \( x = (x, y)^T \in \mathbb{R}^2 \). A combination of translation and rotation is represented by the special Euclidean group \( SE(2) \) or a subgroup of \( SE(2) \). We consider scalar (i.e., tensor of rank 0) and vector (i.e., tensor of rank 1), but not higher-rank tensors. Contravariant and covariant vectors are identified because of the orthonormal basis on the Cartesian coordinates. There are two kinds of transformations in the tensor analysis: active and passive transformations. In the active one, scalar and vector themselves are transformed, whereas the coordinates remain fixed. The passive transform adopts the opposite treatment. We employ the active transformation following the previous studies on the equivariant CNNs. When the passive transform is used, we state it clearly.

A. Covariance and equivariance

Covariance and equivariance are clarified for this study. See the references on the mathematical details.

The translation and rotation are defined using a position vector \( x = (x, y)^T \):

\[
\tau_t x = x + t, \quad \rho_\theta x = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix},
\]

(1a)

(1b)

where \( t \) is a translation vector and \( \theta \) is a rotation angle. The inverse transforms are \( \tau_t^{-1} x = \tau_{-t} x \) and \( \rho_\theta^{-1} x = \rho_{-\theta} x \). A combination of translation and rotation is formulated as follows:

\[
\tau_t \rho_\theta x = (\rho_\theta x) + t.
\]

(2)

If there is no confusion, the subscripts of \( t \) and \( \theta \) are omitted.

The covariance gives the transformation laws for a scalar \( \omega(x) \) and a vector \( \nu(x) \) under rotation and translation:

\[
\begin{align*}
\text{scalar:} & \quad \omega(x) \mapsto \omega(\rho^{-1}(x - t)) \quad \text{[}3a\text{]} \quad \omega: \mathbb{R}^2 \to \mathbb{R}, \\
\text{vector:} & \quad \nu(x) \mapsto \rho\nu(\rho^{-1}(x - t)) \quad \text{[}3b\text{]} \quad \nu: \mathbb{R}^2 \to \mathbb{R}^2,
\end{align*}
\]

(3a)

(3b)

Only the referred position is changed for a scalar as in \(3a\), whereas not only the referred position but also the components are changed for a vector as in \(3b\). The change in the vector components describes the change in its direction. When scalars or vectors are piled into a numerical array, each one is independently transformed according to \(3a\) and \(3b\). Both scalar and vector satisfy the linearity. The transformation law, i.e., covariance, distinguishes between scalar and vector. The last equation \(3c\) can be used to any combination of any tensor fields, where the representation of \( \pi(\rho) \) depends on \( F \):

\[
\pi(\rho) = \begin{cases} 1 & \text{for a scalar}, \\
\rho & \text{for a vector}, \\
\text{a block diagonal matrix consisting of } \rho & \text{when } F \text{ is a stack of vectors}.
\end{cases}
\]

The equivariance is defined as the commutativity of transformation \( g \) and some function \( f \) such as an NN. Let us consider a function \( f: X_{in} \mapsto X_{out} \), where \( X_{in} \) is a scalar or vector field and for simplicity \( X_{out} \) is assumed to be a field of the same type as \( X_{in} \). The function \( f \) is equivariant if it satisfies

\[
(f \circ g)(X_{in}) = (g \circ f)(X_{in}) = g(X_{out}),
\]

(4)

where the representation of \( g \) is determined by \( X_{in} \) and \( X_{out} \). Eq. (4) means that \( f \) is equivariant if \( f \) is commutative with \( g \). When \( g \) represents only a rotation, the rotated input \( g(X_{in}) \) gives the output \( f(g(X_{in})) \), which is equal to the rotated output \( g(X_{out}) \) (see Fig. [1]). The whole neural network is equivariant if each layer satisfies the equivariance (4). Note that the invariance is a special case of the equivariance: \( (f \circ g)(X_{in}) = X_{out} \), where the output is a constant field.

B. Method of making equivariant CNNs

The equivariant CNNs used in this study were obtained from existing CNNs by making each layer equivariant. Many network architectures for super-resolution have been proposed. To make the best use of those results, it is necessary to examine the influence of converting the proposed networks to equivariant ones. We employ two CNNs: hybrid downsampled skip-connection/multi-scale (DSC/MS) model by Fukami et al.\cite{31} and residual in residual dense network (RRDN) by Bode et al.\cite{27,28} The details on the network architectures are presented in Appendix A.

The hybrid DSC/MS model (hereafter, DSC/MS) was proposed in pioneering work on the fluid super-resolution by Fukami et al.\cite{31} They combined their model of downsampling skip connection (DSC) with the multi-scale (MS) model.\cite{20} Multi-scale signals in a fluid field can be captured by the downsampling and skip connections in the DSC/MS. They reported that the DSC/MS super-resolved both velocity and vorticity fields in two dimensions. In the subsequent studies, the DSC/MS was applied to the three-dimensional velocity\cite{70}, the spatio-temporal super-resolution was performed by the successive use of DSC/MS in space and time\cite{23}. A deeper version of DSC/MS was proposed\cite{23}. Although the DSC/MS was proposed relatively early, it is still one of the most important CNNs in the fluid super-resolution.

The RRDN utilizes residual in residual dense block (RRDB)\cite{20,21} which extracts multi-scale structures in a fluid field with dense skip connections. Bode et al.\cite{31} proposed the RRDN as the generator of their physics-informed generative adversarial network (GAN) and applied their GAN to the three-dimensional velocity. A similar network to the RRDN was employed for the super-resolution of two-dimensional passive scalars\cite{20}. The RRDB, i.e., the core module of the RRDN, is utilized for the super-resolution not only in fluid.
mechanic\cite{26,31} but also in computer vision\cite{20,21}. It is important to investigate an equivariant CNN having RRDBs.

We discuss the discretization of rotation and the group representation in hidden layers before explaining how to make the DSC/MS and RRDN equivariant. In practice, scalar and vector fields need to be discretized in space, and the rotation acting on them is also discretized. The present study focuses on the discrete rotation by multiples of $90^\circ$ or $180^\circ$. These rotations do not require interpolation or extrapolation and are easily made commutative with downsampling such the local average (Section \textsection VA). The rotations by multiples of $90^\circ$ and $180^\circ$ are mathematically described by the cyclic groups of order 4 and 2, i.e., $C_4$ and $C_2$, respectively. CNNs equivariant to $C_4$ and $C_2$ are called $C_4$- and $C_2$-equivariant, respectively. Rotations with smaller intervals are briefly discussed in the context of local isotropy in Section \textsection VC.

One of the important hyperparameters in equivariant CNNs is the group representation used in hidden layers.\cite{33} We adopted the regular representation of $C_4$, considering the following three reasons. First, the regular representation consists of all irreducible representations, namely all minimal representations.\cite{26,31} Second, the CNNs using the regular representation showed high accuracy in image classification.\cite{26,31} Third, each channel can be transformed nonlinearly and independently when the regular representation is used.\cite{26,31} When the irreducible representation is employed, each channel cannot be transformed independently by a nonlinear function such as ReLU. A function such as Norm-ReLU\cite{33} is necessary, which acts to the vector norm and preserves the vector direction. Such a limitation may reduce the expressive power of CNNs and may not be suitable for super-resolution. An example of the regular representation is given in Appendix [B].

Equivariant versions of DSC/MS and RRDN are referred to as the Eq-DSC/MS and Eq-RRDN, respectively. The details of the Eq-DSC/MS and Eq-RRDN are described in Appendix [A]. These models were constructed by replacing usual convolution layers with the equivariant ones, while keeping the number of channels to make the calculation time nearly the same. For instance, the regular representation of the $C_4$ group is composed of the 4 elements. If both input and output of a convolution layer have 64 channels, the $C_4$-equivariant convolution employs 16 sets of the regular representation. The equivariant kernels are not completely independent across the channel dimension,\cite{33} reducing the number of trainable parameters. Generally, there is a trade-off between expressive power and equivariance. Two examples of the trade-offs are given in Appendix [B] together with a review on the equivariant convolution.

\section{Methodology}

This section gives the methods of training and testing the CNNs. The high-resolution (HR) data were generated from the numerical simulations that solved the governing equations (Section IIIA). The low-resolution (LR) data were obtained by downsampling the HR data (Section IIIB). The supervised training was performed using the pairs of HR and LR data (Section IIIC). The evaluation metrics are defined in Section IIID. In addition to the above datasets, we utilized the HR and LR data that were generated separately from HR and LR simulations, respectively, with the spectral nudging technique.\cite{27,28,69} The spectral nudging method is explained in Appendix [C]. The source code used in this study is available on the GitHub repository\cite{29}.

\subsection{Fluid dynamics simulations}

Two kinds of experiments were conducted: freely-decaying turbulence on a square flat torus and barotropic instability on a periodic channel. The governing equations for both experiments are as follows:

$$
\frac{\partial \omega}{\partial t} + u \frac{\partial \omega}{\partial x} + v \frac{\partial \omega}{\partial y} = -\nu(-\Delta)^n \omega, \tag{5a}
$$

$$
(u, v) = \left( \frac{\partial \psi}{\partial y}, \frac{\partial \psi}{\partial x} \right), \tag{5b}
$$

$$
\omega = \Delta \psi = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}, \tag{5c}
$$

where $t$ denotes time, $\nu$ is a positive real number, and $n$ is a positive integer. The velocity components $(u, v)$ are given by the stream function $\psi$ in (5b), which is obtained by solving the Poisson equation $\Delta \psi$ with the vorticity $\omega$. For the two-dimensional translation and rotation, the vorticity is regarded as scalar, whereas the velocity as vector. Note that large-scale flows in the ocean and atmosphere are described by similar equation\cite{23} (i.e., quasi-geostrophic equations); the only difference is that the Poisson equation is three-dimensional and includes the altitude dimension.

The governing equations were numerically solved with the spectral method. We employed a software library for spectral calculations, called ISPACK\cite{22} coded with Fortran. The discrete Fourier transform was used in a periodic direction, while the discrete sine/cosine transform was used in a direction bounded by two walls. The fourth-order Runge-Kutta method was utilized for the time integration.

The freely-decaying turbulence experiment was conducted following Fukami et al.\cite{23} and Taira et al.\cite{23}. We set $\nu = 10^{-2}$ and $n = 1$, leading the governing equations (5) to the two-dimensional Navier-Stokes equations. The doubly periodic boundary condition was imposed on the flow domain $[0, 2\pi] \times [0, 2\pi]$. This domain, namely a square flat torus, has the rotational symmetry for multiples of $90^\circ$, though it is not isotropic. The grid numbers were set to be $128 \times 128$, which determined the truncation wavenumber of 42 for alias-free simulations. The initial condition was given at random with keeping the energy spectrum of $E(k) = a \times k \exp(-k^2/k_0^2)$, where $k$ is the magnitude of wavenumber vector and $k_0 = 26.5$. For the training data $a = 9$ and for the test $a = 12$. Varying the random initial condition, we generated 1,000 simulation sets for the training and 100 simulations for the test. In each simulation, 10 snapshots were sampled between $1.1 \leq t \leq 2.0$ at the intervals of 0.1. Thus, the training and test
datasets consist of 10,000 and 1,000 snapshots, respectively. Similar test results to those shown below were obtained when the test data were halved, implying that the 1,000 snapshots are sufficient to estimate generalization errors.

Barotropic instability (or shear instability) is a fundamental mechanism to make the evolution of flows complicated.\(^\text{(24)}\) We set \(v = 10^{-30}\) and \(n_v = 8\), leading the governing equations \(^{\text{(5)}}\) to the two-dimensional Euler equations with the hyper-viscosity. The channel domain is \([0, 2\pi] \times [-\pi/2, \pi/2]\), where the \(x\) direction is periodic and the \(y\) direction is bounded by the two walls \([y = 0 \ (y = \pm \pi/2)]\). This domain has the rotational symmetry for multiples of 180° and is clearly not isotropic. The grid size was 128 × 65, which determined the truncation wavenumber of 42 for alias-free simulations. The initial condition was the superposition of an unstable laminar flow and random perturbations. The laminar flow has the shear region:

\[
  u = \begin{cases} 
    U & (\frac{\pi}{2} \leq y \leq \frac{\pi}{2}), \\
    U + 2U \left(\frac{y-\pi}{w}\right) & (-\frac{\pi}{2} < y < \frac{\pi}{2}), \\
    -U & (-\frac{3\pi}{4} \leq y \leq -\frac{\pi}{4}),
  \end{cases}
\]

where the shear is controlled by the width \(w\) and the speed \(U\). Note that the channel-wise velocity component \(v\) is zero. The speed was set either \(U = 0.25\) (positive shear) or \(-0.25\) (negative shear). The number of experiments of \(U = 0.25\) and \(-0.25\) was kept the same. For the training data \(w = 0.45\) and for the test 0.40, which controlled the most unstable mode and made the test data slightly finer.\(^{\text{(24)}}\) In the numerical experiments, the velocity profile \(^{\text{(5)}}\) was made smooth around \(y = \pm w/2\) to avoid the initial discontinuity of the vortex field. Each wavenumber perturbation had the constant amplitude of \(10^{-4}\) and a random phase. Varying the random phases, we generated 800 simulations for the training and 80 for the test. In each simulation, 14 snapshots were sampled between \(31 \leq t \leq 70\) at the intervals of 3. Thus, the training and test datasets consist of the 11,200 and 1,120 snapshots, respectively. Similar test results to those shown below were obtained when the test data were halved, implying that the 1,120 snapshots are sufficient to estimate generalization errors.

B. Data preparation

The HR snapshots of velocity or vorticity were generated from the fluid simulations in the previous subsection. The LR snapshots were obtained by downsampling the HR ones. Two downsampling algorithms were used: subsampling and local averaging. The subsampling here means the extraction of pixel values at certain intervals, which may be regarded as fluid observation such as particle image velocimetry (PIV)\(^{\text{(23,24)}}\). The local average can be used to compress HR fluid data and an SR model is then utilized to decompress the data.\(^{\text{(25)}}\)

The subsampling and local averaging are processed in four steps. First, a scale factor \(s\) is determined. When \(s = 2\), for instance, an LR image is twice as coarse as the HR image. Second, an HR snapshot of size \(H \times W\) is resized to \(H' \times W'\), where \(H' (W')\) is the multiple of \(s\) closest to \(H (W)\). Third, the downsampling algorithm is applied to the resized HR snapshot. In the subsampling, pixel values are extracted at the intervals of \(s \times s\). In the local average, each non-overlapping area of \(s \times s\) is replaced with its mean. Forth, the obtained snapshot of \(H' / s \times W' / s\) is resized to \(H \times W\). The bicubic interpolation\(^{\text{(22)}}\) was used in all the above resizing. For the velocity, the algorithm was applied to \(u\) or \(v\) separately, since each component of vector is added or subtracted independently in the Cartesian coordinates. The scale factor \(s\) was set to either 5, 9, or 13.

C. Training of CNNs

The CNNs were trained with supervised learning. The CNNs were implemented with PyTorch\(^{\text{(27)}}\) 1.8.0 and e2cnn\(^{\text{(58)}}\) 0.2.1. The latter is a software library for the equivariant deep learning of two-dimensional data.

An important hyperparameter of equivariant CNNs is the order \(N\) of a cyclic group \(C_N\). For instance, when \(N = 3\), a CNN is equivariant to the rotation by multiples of 360°/3 = 120°. We set \(N = 4\) in the decaying turbulence and \(N = 2\) in the barotropic instability experiment because the square torus has the rotational symmetry of 90° and the channel domain is symmetrical for the 180° rotation. Another important hyper-parameter is the group representation. The regular representation was employed, as discussed in Section II B. Appendix \(\text{[D]}\) investigates the effect of changing several hyperparameters.

The Adam optimizer was used together with the loss function of mean squared error (MSE). The MSE is invariant to rotation and translation for scalars and vectors. The mini-batch size for the DSC/MS and Eq-DSC/MS was 100, while that for the RRDN and Eq-RRDN was 32. The learning rate was between \(1.0 \times 10^{-3}\) and \(1.0 \times 10^{-4}\) and was basically the same for the DSC/MS and Eq-DSC/MS and for the RRDN and Eq-RRDN. Each training was finished by early stopping with the patience parameter of 30 epochs.

D. Evaluation metrics of CNNs

The following three metrics were utilized to evaluate the CNNs: norm error ratio, energy spectral error, and equivariance error ratio.

The norm error ratio (NER) is defined as follows:

\[
  \text{NER} = \frac{\sum_{\text{space}} \|Y - \hat{Y}\|}{\sum_{\text{space}} \|Y\|}.
\]

The summation is taken over the whole space. The symbol \(\| \cdot \|\) denotes the Euclidean norm when the inference \(\hat{Y}\) and ground truth \(Y\) are vectors, whereas it is the absolute value in the case of scalars. In the actual evaluation, the NER \(\text{(7)}\) was averaged over the entire test sample.

The energy spectral error (ESE) was introduced to evaluate the validity of spatial pattern by Wang et al.\(^{\text{(13)}}\) The NER is invariant against any pixel shuffling, indicating that the NER does not necessarily reflect the consistency of spatial patterns.
The ESE is equal to the root-mean-squared error of the logarithm of the energy spectra between the inference and ground truth:

\[ \text{ESE} = \sqrt{\frac{\sum_{\text{samples},k} \left[ \log E(k) - \log \hat{E}(k) \right]^2}{\sum_{\text{central region}}}}, \]

where \( E(\hat{E}) \) is the energy spectrum of the ground truth (inference) and the summation is taken over the wavenumber and the entire test sample. The ESE is interpreted as the error of spatial correlations through the convolution theorem.

The equivariance error was also introduced by Wang et al. [3] which measures the equivariance of any model and is calculated without referring to the ground truth. The equivariance error ratio (EER) is defined by

\[ \text{EER} := \frac{\sum_{\text{central region}} \| (f \circ g)(X_{in}) - (g \circ f)(X_{in}) \|^2}{\sum_{\text{central region}} \| (g \circ f)(X_{in}) \|^2}, \]

where \( f \) denotes a CNN and \( g \) is a representation of translation or rotation [3]. The summation was taken over the central 48 × 48 pixels to avoid the effect of extrapolation such as zero padding. In the actual evaluation, the EER [4] was averaged over the entire input sample. Only the rotational transformation is considered below because conventional CNNs possesses the translation equivariance and the usefulness of CNNs in super-resolution is confirmed by many studies (Section I).

IV. RESULTS

This section evaluates the equivariant and non-equivariant CNNs using the metrics in the previous subsection. We present only the results of super-resolution of velocity because the results on vorticity were similar. Appendix D 4 gives a brief comparison between the super-resolution of velocity and vorticity.

A. Freely-decaying turbulence experiment

Fig. 2 shows an example of the vorticity calculated from the super-resolved velocity, where the scale factor \( s = 9 \). The finite difference operation in deriving the vorticity enhances the small-scale structure in the super-resolved flow. For the baseline of bicubic interpolation, the shape of each vortex is ambiguous and the difference from the ground truth is large. All SR models reproduce well the small-scale structure, regardless of equivariant or non-equivariant.

Fig. 3 compares the NERs and ESEs for the locally averaged and subsampled velocity. As the scale factor \( s \) is increased, the errors tend to be increased. The NER and ESE values of the Eq-DSC/MS and Eq-RRDN are comparable to those of the DSC/MS and RRDN in all cases, respectively. The Eq-RRDN and RRDN tend to show slightly lower errors than the Eq-DSC/MS and DSC/MS, which is likely due to that the Eq-RRDN and RRDN are deeper and have more parameters (Table I).

Table I compares the numbers of the trainable parameters in the SR models. The number of parameters is generally smaller in equivariant models due to the constraint of the equivariance (Appendix B). Since the \( C_4 \) group contains \( C_2 \), the constraint on the \( C_4 \)-equivariant model is stricter and more parameters are reduced. The reduction in parameters implies that the expressive power of the equivariant SR models is lower in principle; however, the Eq-DSC/MS and Eq-RRDN show comparable performance with the DSC/MS and RRDN, respectively (Figs. 2 and 3). This result indicates that imposing the equivariance does not necessarily act to deteriorate the accuracy of super-resolution. The smaller number of parameters suggests that an equivariant model is trainable with a smaller size of data [5]. This suggestion was confirmed in Appendix D 2.

| Model name                        | Number of trainable parameters |
|-----------------------------------|-------------------------------|
| DSC/MS                            | 141,298                       |
| \( C_2 \)-equivariant DSC/MS       | 37,128                        |
| \( C_4 \)-equivariant DSC/MS       | 18,564                        |
| RRDN                              | 1,256,770                     |
| \( C_2 \)-equivariant RRDN         | 419,104                       |
| \( C_4 \)-equivariant RRDN         | 209,552                       |

Table I compares the numbers of the trainable parameters in the SR models. The number of parameters is generally smaller in equivariant models due to the constraint of the equivariance (Appendix B). Since the \( C_4 \) group contains \( C_2 \), the constraint on the \( C_4 \)-equivariant model is stricter and more parameters are reduced. The reduction in parameters implies that the expressive power of the equivariant SR models is lower in principle; however, the Eq-DSC/MS and Eq-RRDN show comparable performance with the DSC/MS and RRDN, respectively (Figs. 2 and 3). This result indicates that imposing the equivariance does not necessarily act to deteriorate the accuracy of super-resolution. The smaller number of parameters suggests that an equivariant model is trainable with a smaller size of data [5]. This suggestion was confirmed in Appendix D 2.

Fig. 4 shows the EER against the rotation angle. The subsampled velocity with the scaling factor \( s = 9 \) was super-resolved in this figure; similar results were obtained in the other cases. The square flat torus has the rotational symmetry for multiples of 90°, and a rotated flow is a solution of the governing equations (3). An SR model should be equivariant to the rotation by multiples of 90° (i.e., \( C_4 \)-equivariant). The original models (DSC/MS and RRDN) show the smaller EERs at the multiples of 90° than those at the other angles, implying that the conventional CNNs learn the approximate \( C_4 \) equivariance. However, the magnitude of EERs is not sufficiently low, compared to that of the equivariant models (Eq-DSC/MS and Eq-RRDN). Data augmentation may be useful for the original models. As for the rotation angles other than the multiples of 90°, the magnitude of EERs is nearly the same for the equivariant and original models because the \( C_4 \)-equivariance does not guarantee the equivariance to the other angles.

The equivariance should be explicitly incorporated into neural network designs as the prior knowledge if the symmetry of a physics system is clear. Although conventional CNNs may learn the equivariance from data, a larger size of data or data augmentation is necessary. The equivariance as the prior knowledge reduces the number of learnable parameters, and the model can be trained with a smaller size of data. This trainability is important in fluid mechanics because the computational cost of fluid simulations is often high and the sufficient size of data may not be available.
FIG. 2. An example of the vorticity calculated from the super-resolved velocity in the decaying turbulence experiment. The scale factor $s$ is 9, that is, the super-resolution increases the resolution by a factor of 9. The subsampled velocity was used as input.

FIG. 3. Test scores calculated from the locally averaged and subsampled velocity in the decaying turbulence experiment. The norm error ratio (NER) in (7) reflects pixel-wise errors. The energy spectrum error (ESE) in (8) reflects spatial pattern errors.

**B. Barotropic instability experiment**

The results of the barotropic instability experiment are briefly described below. The results are similar to those of the decaying turbulence. The $C_2$-equivariant CNNs were used here because the channel domain is symmetrical to the $180^\circ$ rotation.

Fig. 5 shows an example of the vorticity calculated from the super-resolved velocity. The vortices obtained with the bicubic interpolation have indistinct shapes. In contrast, the SR models reproduce well the cat’s eye structure associated with the collapse of the laminar flow having the strong shear region. The flows inferred by the equivariant models are nearly the same as those by the non-equivariant ones.

Fig. 6 compares the NERs and ESEs from the locally averaged and subsampled velocity. The figure suggests that the Eq-DSC/MS and Eq-RRDN are as accurate as the DSC/MS and RRDN, respectively. The Eq-RRDN and RRDN tend to show lower errors than the Eq-DSC/MS and DSC/MS as in the case of the decaying turbulence (Fig. 3).

Fig. 7 shows the EER against the rotation angle. The subsampled velocity with the scaling factor $s = 9$ was super-resolved in this figure; similar results were obtained in the other cases. The channel domain has the rotational symmetry of $180^\circ$. An SR model should be equivariant to the $180^\circ$ rotation (i.e., $C_2$-equivariant). The original models (DSC/MS and RRDN) learn the approximate $C_2$ equivariance; however, the EERs at $180^\circ$ are not sufficiently low, compared to those of the
FIG. 4. Equivariance error ratio (EER) against the rotation angle calculated with the super-resolution models of velocity in the decaying turbulence experiment. The EER is defined in (9). The scaling factor was 9 and the low-resolution data were generated by subsampling. The rotation angle interval is $15^\circ$, where $0^\circ$ and $360^\circ$ are omitted.

FIG. 5. An example of the vorticity calculated from the super-resolved velocity in the barotropic instability experiment. The scale factor $s$ was 9, that is, the super-resolution models made the resolution 9 times finer. The subsampled velocity was used as input.

equivariant ones (Eq-DSC/MS and Eq-RRDN). For the other rotation angles, the magnitude of EERs is nearly the same for the equivariant and original models.

V. DISCUSSIONS

This section demonstrates that the equivariance may be broken due to the method of data generation or the choice of input data regardless of the symmetry of fluid systems. If the equivariance between the input and output is broken, the super-resolved velocity cannot be regarded as a geometric vector field. In this case, the usage of conventional CNNs may be justified even though vector is treated as a quantity having no direction.

Section V A demonstrates that the equivariance can be broken by the data generation of subsampling. Section VB discusses an experiment using the spectral nudging technique. The data generated by the spectral nudging preserve the symmetry intrinsic to the governing equations (Appendix C). However, this symmetry can be broken depending on the choice of input data. The equivariance to local rotation is discussed in Section VC. This local equivariance is called the gauge equivariance.

A. A data generation breaking the equivariance

The $C_4$-equivariance was imposed on the DSC/MS and RRDN in the decaying turbulence experiment because both flow domain and governing equations have the rotational symmetry for multiples of $90^\circ$. The original DSC/MS and RRDN
FIG. 6. Test scores calculated from the locally averaged and subsampled velocity in the barotropic instability experiment. The norm error ratio (NER) in (7) reflects pixel-wise errors. The energy spectrum error (ESE) in (8) reflects spatial pattern errors.

FIG. 7. Equivariance error ratio (EER) against the rotation angle calculated with the super-resolution models of velocity in the barotropic instability experiment. The EER is defined in (9). The scaling factor was 9 and the low-resolution data were generated by subsampling. The rotation angle interval is $15^\circ$, where $0^\circ$ and $360^\circ$ are omitted.

Learned the approximate $C_4$-equivariance from the training data (Fig. 4), implying that the data themselves inherit the rotational symmetry from the flow domain and governing equations. However, this symmetry can be broken, depending on the method of data generation.

Fig. 8 is a schematic illustration that compares the subsampling of the scale factor $s = 2$ and 3. The LR image for $s = 2$, generated by subsampling, is varied by the orientation of the HR image, whereas the LR image for $s = 3$ is not. Local correspondence between LR and HR patterns depends on the orientation. Mathematically, the subsampling of an even number $s$ is not commutative with the rotation by multiples of $90^\circ$. This statement is formulated by

$$(d \circ g)(Y) \neq (g \circ d)(Y),$$

(10)

where $d$ is the subsampling operation, $g$ is the representation of rotation by multiples of $90^\circ$ [see Eq. (3)], and $Y$ is an HR velocity field. Note that the input LR field is given by $X = d(Y)$.

The $C_4$-equivariance is broken by the non-commutativity of subsampling (10). Let us consider an HR velocity $Y$ and...
FIG. 8. Schematic illustration showing the commutative property of subsampling with rotation. In the case of the scaling factor \( s = 2 \), the subsampling picks up the values from the upper left pixels (black tiles) in the windows of \( 2 \times 2 \). When \( s = 3 \), the subsampling extracts the central pixels (black tiles) regardless of the orientation.

The rotated version \( g(Y) \), both of which can exist because the flow domain and governing equations have the rotational symmetry. The LR fields are generated as \( X = d(Y) \) and \( X' = d(g(Y)) \). Importantly, \( X' \) is not equal to \( g(X) \) \([= g(d(Y))]\). The non-commutativity leads the inconsistency of LR and HR pairs. When an SR model is trained with the pairs of \((X, Y)\) and \((X', g(Y))\), it fails to learn the equivariance because \((X', g(Y))\) is not \((g(X), g(Y))\). In this case, the \( C_4 \)-equivariance should not be imposed on the SR model.

We demonstrate the effect of the non-commutativity \((10)\) by comparing the super-resolved velocity of the decaying turbulence for \( s = 8 \) and \( s = 9 \) (Fig. 9). The NERs of the equivariant models are much larger than those of the non-equivariant ones, only in the case of subsampling of \( s = 8 \). When \( s = 9 \), the subsampling is commutative with the rotation, and such large errors are not observed. As for the local average, the operation is always commutative with the rotation regardless of \( s \), and large errors are not found in both cases of \( s = 8 \) and 9.

The result implies that the SR model deteriorates if the equivariance is forced to be imposed when the data do not have the rotational symmetry. The ESE is comparable between the equivariant and non-equivariant models even for the subsampling of \( s = 8 \). The non-commutativity of subsampling largely affects pixel-wise values but not spatial patterns.

Fig. 9 shows the EER curves for the non-equivariant models (DSC/MS and RRDN) with \( s = 8 \) and 9. In the case of the local average, the models learn the approximate \( C_4 \)-equivariance regardless of \( s \). In the case of subsampling, the \( C_4 \)-equivariance is learned when \( s = 9 \) (solid curves) but not when \( s = 8 \) (dashed curves). The result indicates that the data do not have the equivariance in the latter case.

Even if a flow domain and governing equations have rotational symmetry, the symmetry may not carry over to LR and HR velocity fields. An SR model trained with these inconsistent data is not equivariant. If the coordinates are rotated (i.e., a passive transform), both LR and HR velocities should be rotated. However, since the SR model is not equivariant, the super-resolved velocity is not rotated according to the rotation of the input. The output velocity does not satisfy the covariance to rotation and cannot be regarded as a geometric vector.

In this case, the usage of conventional CNNs may be justified even if CNNs treat vector as a quantity having no direction.

B. An input choice breaking the equivariance

This subsection discusses that the equivariance can be broken depending on the choice of input data. The SR models are applied to the data generated with the spectral nudging method. Appendix C gives the details on the spectral nudging technique used in this study.

Both LR and HR data are generated from the fluid simulations. The LR data are generated by numerically solving
the governing equations (5). For the HR simulation, a linear relaxation toward the LR flow is added to the right-hand side of (5a). If the relaxation time is small enough, the large-scale flow pattern in the HR simulation becomes similar to that in the LR one. In addition to the large scales, the small-scale motions evolve in the HR simulation according to the governing equation (5a), where the relaxation term vanishes because these small scales are not resolved in the LR model. The spectral nudging technique is understood as a method of dynamical downscaling that reproduces fine-scale motions not resolved in LR simulations.

The spectral nudging may be useful to construct a super-resolution simulation system. The purpose of this system is to reduce the computational cost of HR simulations. After an SR model is trained, i.e., at the inference stage, only the LR fluid simulation is conducted and the HR results are obtained by feeding the LR snapshots to the SR model. A dataset generated with the spectral nudging can be utilized to train an SR model in the super-resolution simulation system.

The spectral nudging method was applied to the case of barotropic instability. The barotropic instability is more appropriate than the decaying turbulence because the initial laminar flow is clearly developed into smaller scales. The truncation wavenumber of the LR fluid model was 10 and that of HR was 42; the scale factor $s$ was regarded as 4. For comparison, another set of LR data was generated by locally averaging the HR flows obtained from the simulations with the spectral nudging. The value of $s$ was 4 for the local average.

Fig. 11 shows an example of the vorticity calculated from the super-resolved velocity. The non-equivariant models (DSC/MS and RRDN) reproduce the small-scale pattern in both cases of local averaging and spectral nudging. In contrast, the super-resolved flows inferred by the equivariant models (Eq-DSC/MS and Eq-RRDN) are clear in the local average but blurred in the spectral nudging.

Fig. 12 shows the NERs and ESEs in the cases of local averaging and spectral nudging. The errors tend to be larger in the spectral nudging than those in the local average, which is attributed to that the HR signals are completely not contained in the LR data of the spectral nudging. In the case of local average, the NERs and ESEs of the equivariant models are comparable to those of the non-equivariant ones, as in Fig. 6. In the case of spectral nudging, however, the NERs and ESEs of the equivariant models are larger than those of the non-equivariant ones.

The LR and HR velocities from the fluid simulations with the spectral nudging have the 180° rotational symmetry because the channel domain and governing equations have this symmetry even when the linear relaxation is added in the HR fluid model (Appendices C1 and C2). In other words, a set of LR and HR velocities rotated by 180° is a solution to the governing equations and can be obtained from the rotated initial condition. Moreover, the LR velocity generated with the local average inherits the rotational symmetry because the average operation is commutative with the 180° rotation (Section V A). Therefore, in both cases, an SR model should be $C_2$-equivariant.

Fig. 13 shows the EER curves of the non-equivariant models. The decrease in the EER at 180° is observed in the local average. On the contrary, the EER is the largest around 180° in the spectral nudging. The SR models learn the $C_2$-equivariance only in the case of local average. This result was confirmed even when the training data size was doubled, indicating that the result is not due to the lack of training data. The large NERs and ESEs of the equivariant models in the spectral nudging (Fig. 12) are attributed to that the equivariance is forced to be imposed even though the velocity data do not have the 180° rotational symmetry.

The result of the spectral nudging seems to contradict the fact that the velocity rotated by 180° is a solution to the gov-
FIG. 11. An example of the vorticity calculated from the super-resolved velocity in the barotropic instability experiment. The upper figures are from the case of local averaging and the bottom from the case of spectral nudging.

Let us consider the problem of inferring the HR velocity $v_n^H$. We first discuss the case where only the LR velocity $v_n^L$ is utilized as in the above experiment (Figs. 11-13), even though $v_n^H$ depends on the past states. The past LR velocities $v_t^L (t < n)$ are regarded as unobserved because they are not fed into an inference model. When $v_n^H$ is rotated by $180^\circ$, it is not clear whether all past states $v_t^L (t < n)$ are rotated or not. This unclarity suggests that the model is quite hard to learn the equivariance from the training data. Moreover, the graphical model indicates that $v_n^H$ is conditionally independent of the current and future LR states $[v_t^L (t \geq n)]$ when the previous LR state $v_{n-1}^L$ is given.

The usage of past states is consistent with the graphical model (Fig. 14) because the current HR velocity $v_n^H$ depends on all past LR velocities $v_t^L (t < n)$. A neural network taking into account the structure of the graphical model is necessary because it is not clear how many past time steps are necessary to infer the current state. When feeding time series, the Galilean equivariance should be incorporated into the model. Interestingly, the current HR state is conditionally independent of the future LR states, which is different from a situation that is applied by the Kalman smoother. Some SR models improved the accuracy of the fluid super-resolution by utilizing time series of flow fields. It would be a first step to consider the application of these models to data generated with the spectral nudging.

C. Equivariance to local rotation

The global equivariance has been discussed so far. This subsection focuses on the equivariance to local rotation, which is a kind of gauge equivariance. The local rotation means that the rotation angle depends on the position. Note that the gauge transformation is defined as a local change in basis vectors and is more general than local rotation.

Three things can be mentioned to motivate the discussion on the local equivariance. First, the local equivariance extends the application of super-resolution. If an SR model is equiv
FIG. 13. Equivariance error ratio (EER) against the rotation angle calculated with the non-equivariant super-resolution models of velocity in the barotropic instability experiment. The EER is defined in (9). The low-resolution velocity was generated with (left) the local average and (right) the spectral nudging. The rotation angle interval is $15^\circ$, where $0^\circ$ and $360^\circ$ are omitted.

FIG. 14. Graphical model of the velocity governed by the equations with the spectral nudging [Eqs. (C3) and (C4)]. The superscript $H$ ($L$) denotes the high- (low-) resolution. The subscript represents the time index. The circles of the low-resolution states are colored gray because they are observed at the inference step.

ariant to local rotation, the model can be applied to cropped data in any orientation without losing accuracy and physical validity. Second, if a convolution operation is globally equivariant to some combinations of rotation and translation, it is also locally equivariant and vice versa. This fact means that the $C_4$-equivariant models used above (Eq-DSC/MS and Eq-RRDN) are equivariant not only to the global but also to the local rotation by multiples of $90^\circ$. Third, SR models with the convolution often utilize the local correspondence between LR and HR patches. The locality can be interpreted with the sparse-coding-based method, which can be realized by CNNs. Even if a model can learn the non-locality, the convolution layer is often employed as a part of the non-local models.

The locality in two-dimensional fluids is first discussed from a physics point of view. Two-dimensional fluid is a long-range interacting system, where any local change can affect globally. This non-locality can be understood from the fact that the stream function and vorticity are linked by the Poisson equation. Even on the infinite plane, the Green function is proportional to the logarithm. Similarly, large-scale flows in the ocean and atmosphere at mid-and high-latitudes have non-locality because of the three-dimensional Poisson equation of the quasi-geostrophic system. Therefore, the super-resolution does not rigorously satisfy the local equivariance for these fluids. For instance, when a vortex shape is distorted locally in an LR field, the super-resolved flow is globally affected.

CNNs have succeeded in super-resolving two-dimensional fluids and large-scale flows in the ocean and atmosphere. These results suggest that the long-range interacting nature is not critical to the super-resolution in practice, and the gauge equivariant CNN may be useful for the fluid super-resolution. In applying a gauge equivariant CNN, it may be natural to assume the local isotropy even if the entire domain is not globally isotropic. Our results (Figs. 4 and 7) suggest that the local symmetry is strongly affected by the global one. The EER was evaluated over the local patch around the center by using the SR models (Eq-DSC/MS and Eq-RRDN) that had the local equivariance. The EERs were minimal at the rotation angles reflecting the global symmetries, and the local isotropy was not clear in Figs. 4 and 7. To investigate the effectiveness of gauge equivariant CNNs, it may be necessary to apply an SR model to complicated flows such as atmospheric flows, where the global symmetry is not obvious.

VI. CONCLUSIONS

Vector is an essential quantity in physics. In applications of convolutional neural networks (CNNs), vector is often processed as a quantity having no direction like color. We investigated the effect of treating vector as a geometrical quantity, in which a direction is recognized, in terms of super-resolution of two-dimensional flows. Vector is distinguished from scalar...
by the transformation law, i.e., the covariance in geometry. Equivariant CNNs incorporate the transformation laws as the prior knowledge and preserve the covariance of vector regardless of training. We converted two existing non-equivariant CNNs into equivariant ones using a theory of equivariant deep learning, where the rotation and translation equivariance were taken into account. To generate the training data, the two-dimensional fluid simulations were conducted and the pairs of low- and high-resolution (LR and HR) velocities were generated. The two types of data were used: in the first type, the LR data were generated by downsampling the HR ones; in the other type, both LR and HR data were generated from the numerical simulations with the spectral nudging method.

The contrasting results were obtained depending on the symmetry of the data. When the data inherit the rotational symmetry from the flow domain and governing equations, the equivariant CNNs show comparable accuracy with the non-equivariant ones. Since the number of learnable parameters is smaller in the equivariant CNNs, these models can be trained with a smaller size of the data. This trainability is important in fluid mechanics because the computational cost of fluid simulations is often high and the sufficient size of data may not be available. Therefore, when a dataset possesses the rotational symmetry, the transformation law of vector should be incorporated into a neural network design as the prior knowledge, where vector is treated as a quantity having direction.

The symmetry of the data can be broken. We demonstrated two examples. In the first case, the method of data generation, i.e., the subsampling, made the correspondence between LR and HR patterns dependent on the orientation, which broke the equivariance of super-resolution. In the second case, the choice of input data was not appropriate in the experiment with the spectral nudging. Although the HR velocity depended on all past LR states, only the current LR velocity was fed into the CNNs. This insufficient input might have broken the equivariance of super-resolution. When an existing dataset does not reflect the rotational symmetry of the governing equations, the accuracy of a CNN may deteriorate if the equivariance is forced to be imposed. In this case, the usage of conventional CNNs may be justified even though they treat vector as a quantity having no direction like color.

ACKNOWLEDGMENTS

This work was supported by the JSPS KAKENHI (grant numbers 20H05751 and 20H02074). The deep learning was conducted on the supercomputer of TSUBAME 3.0 at the Tokyo Institute of Technology.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request. The source code is available on the GitHub repository.

Appendix A: Implementation of the equivariant CNNs

This section gives the details on the implementation of the equivariant CNNs. We made two existing CNNs equivariant: the hybrid DSC/MS model by Fukami et al. and the RRDN by Bode et al. The architectures of the DSC/MS and RRDN are shown in Figs. 15 and 16 respectively.

FIG. 15. Network architecture of the hybrid DSC/MS model (Downsampled Skip-Connection and Multi-Scale). The architecture follows the CNN proposed by Fukami et al. All convolutions are two-dimensional. The symbol \( k \) denotes the spatial size of kernel; for instance, \( k = 3 \) means that the kernel size is 3 \( \times \) 3. The linear interpolation was used in upsampling.

FIG. 16. Network architecture of the RRDN (Residual in Residual Dense Network). The architecture follows the generator of the GAN proposed by Bode et al. All convolutions are two-dimensional. The symbol \( k \) denotes the spatial size of kernel; for instance, \( k = 3 \) means that the kernel size is 3 \( \times \) 3. The linear interpolation was used in upsampling.

If each layer is equivariant, the whole network is equivariant. A two-dimensional convolution layer is replaced with the equivariant convolution. A max-pooling layer is replaced with the norm max-pooling when the input is a vector field, while it remains unchanged when the input is scalar or described by the regular representation. In all cases, the pooling operation is equivariant. The linear or bicubic interpolation is used in resizing data, both of which are equivariant because the weights of interpolation depend only on the absolute difference in \( x \) or \( y \) coordinates. The skip connection is equivariant because it is an addition operation between tensors having the same type. Since all hidden layers utilize the regular representation (Section II B), a channel-wise nonlinear transformation such as ReLU is equivariant. The input or output representation is determined as follows. When the input or output is scalar, it can be
regarded as a feature field with the trivial representation, in which any conversion is not necessary. A vector field is naturally considered as a feature with the irreducible representation of order 1.

All the CNNs used in the present study were implemented with PyTorch\cite{pytorch} and e2cnn\cite{e2cnn}. The source code is available on the GitHub repository\cite{github}.

Appendix B: Review on equivariant convolution

This section reviews the equivariant convolution\cite{equivariant}. Our aim here is to explain a trade-off between the expressive power and equivariance, but not to give the details on the theory. We discuss a linear constraint that is satisfied by equivariant kernels and then give two examples to understand the meaning of the constraint.

1. Linear constraint on equivariant kernels

The convolution is defined by

\[ F_{\text{out}}(y) = \kappa * F_{\text{in}} = \int \kappa(y-x) F_{\text{in}}(x) \, dx , \]

where * stands for the convolution, the input \( F_{\text{in}} \) and output \( F_{\text{out}} \) are tensors generally having different ranks, \( \kappa \) is a kernel, and the integral is performed over the continuous Euclidean plane without boundary. The multiplication between \( \kappa \) and \( F_{\text{in}} \) is generally a matrix operation. Let us consider that the input \( F_{\text{in}} \) and output \( F_{\text{out}} \) are transformed by the action of an element of the special Euclidean group \( SE(2) \) as in Eq. (3):

\[ g_{\text{in}} : F_{\text{in}}(x) \mapsto \pi_{\text{in}}(\rho) F_{\text{in}}(\rho^{-1}(x-t)) , \]

\[ g_{\text{out}} : F_{\text{out}}(x) \mapsto \pi_{\text{out}}(\rho) F_{\text{out}}(\rho^{-1}(x-t)) , \]

where the representation of \( \pi \) depends on \( F_{\text{in}} \) and \( F_{\text{out}} \). If the kernel \( \kappa \) satisfies the following linear constraint, the convolution in (B1) is equivariant\cite{equivariant,linear}.

\[ \pi_{\text{out}}(\rho) \kappa(x) = \kappa(\rho x) \pi_{\text{in}}(\rho) . \]

Following Weiler and Cesa\cite{equivariant} the equivariance of the convolution is confirmed:

\[ \kappa * g_{\text{in}}(F_{\text{in}}) = \int \kappa(y-x) \pi_{\text{in}}(\rho) F_{\text{in}}((\rho^{-1}(x-t)) \, dx' , \]

\[ = \int \kappa(y-t-\rho x') \pi_{\text{in}}(\rho) F_{\text{in}}(x') \, dx' , \]

\[ = \int \kappa(\rho^{-1}(y-t-x')) \pi_{\text{in}}(\rho) F_{\text{in}}(x') \, dx' , \]

\[ = \pi_{\text{out}}(\rho) F_{\text{out}}(\rho^{-1}(y-t)) , \]

\[ = g_{\text{out}}(F_{\text{out}}) , \]

where we use the fact that the Jacobian is 1 under any coordinate transformation of \( SE(2) \). The constraint (B4) is not only sufficient but also necessary for the equivariance of convolution\cite{equivariant}. Moreover, a linear map is equivariant if and only if the map is an equivariant convolution\cite{equivariant}.

2. Example1: trivial representation with spatially varying kernel

We interpret the linear constraint (B4) using two simple examples. The effect of changes in position is first discussed. Let us consider that the input and output are scalars. Scalar has the trivial representation\cite{equivariant} and \( \pi(\rho) \) becomes the unity. The constraint (B4) becomes

\[ \kappa(x) = \kappa(\rho x) , \]

where \( \kappa(x) \) is a real number depending on \( x \) and \( \kappa(x) \) acts on a scalar field.

The rotation by multiples of 120° is considered as an example. A pattern having the wavenumber-3 structure such as \( \cos(3\theta) \) is a solution to (B6). Since such a kernel is invariant to the rotation, the output is rotated according to the rotation applied to the input. The radial structure is not determined by (B6), usually, some form such as Gaussian is assumed and its amplitude is optimized by training\cite{equivariant}. There is a trade-off between expressive power and equivariance. For the conventional convolution, a kernel can learn an azimuthal pattern from data, while it is not necessarily equivariant. This fact is in contrast to the equivariant kernels satisfying the constraint (B6).

3. Example2: \( C_3 \)-regular representation with spatially constant kernel

The effect of \( \pi(\rho) \) in (B4) is next discussed. Let us consider that the kernel is spatially constant and the input and output are described by the same regular representation. Eq. (B4) is simplified as follows:

\[ \pi(\rho) \kappa = \kappa \pi(\rho) , \]

where the subscripts of “out” and “in” are omitted and the matrix \( \kappa \) does not depend on position.

The rotation by multiples of 120° is considered as an example. In the regular representation, the rotation is formulated as a permutation\cite{equivariant}. Fig. 17 is a schematic illustration describing the action of the 120° rotation on a vector. Each component of the vector can be considered as a vertex of an equilateral triangle. The rotation is interpreted as the permutation of these vertices. The regular representation of \( C_3 \) consists of the following three matrices:

\[ \pi(\theta = 0) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} , \]

\[ \pi(\theta = 120^\circ) = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} , \]

\[ \pi(\theta = 240^\circ) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} , \]

where the argument of \( \pi \) is denoted by the rotation angle \( \theta \) instead of the rotation matrix \( \rho \).
Fig. 17. Schematic illustration describing the action of the 120° rotation on a vector \((a, b, c)^T\) with the regular representation. The matrix \(\pi(\theta = 120^\circ)\) in \((B8)\) acts on the vector.

The constraint \((B7)\) needs to hold for all three matrices of \(\pi\) in \((B8)\). The solution is

\[
\begin{pmatrix}
\kappa_1 & \kappa_2 & \kappa_3 \\
\kappa_3 & \kappa_1 & \kappa_2 \\
\kappa_2 & \kappa_3 & \kappa_1
\end{pmatrix},
\]

(B9)

where \(\kappa_1, \kappa_2,\) and \(\kappa_3\) are learnable parameters and determined by training. The kernel \((B9)\) has some symmetric structure. The kernel multiplication is commutative with the permutation representing the rotation (i.e., \(C_3\)-equivariant). The equivariant kernel \((B9)\) has only three learnable parameters, though it is a \(3 \times 3\) matrix. In the conventional convolution, when a kernel is \(3 \times 3\), that is, 3 channels are transformed into 3 channels, all 9 parameters are learnable. This is another example of trade-off between expressive power and equivariance.

Appendix C: Details on the spectral nudging method

This section gives the details on the spectral nudging technique used in this study. To introduce a nudging term, the governing equations \((3)\) are represented in the wavenumber space (Section \(C1\)). The covariance to rotation and translation are discussed in Sections \(C2\) and \(C3\), respectively. Section \(C4\) describes the configuration of the barotropic instability experiment with the spectral nudging. For simplicity, the infinite Euclidean plane is assumed in Sections \(C1\) - \(C3\) which makes wavenumbers continuous.

1. Governing equations in wavenumber space

We examine the governing equations \((3)\) in the Euclidean space before deriving the equations in the wavenumber space. Eqs. \((5a) - (5c)\) are transformed into the following:

\[
\frac{\partial \omega}{\partial t} + \frac{\partial \psi}{\partial x} \frac{\partial \omega}{\partial y} - \frac{\partial \omega}{\partial x} \frac{\partial \psi}{\partial y} = 0,
\]

(C1a)

\[
\nabla^2 \psi = \omega,
\]

(C1b)

where the viscosity term is omitted for simplicity and \(\nabla\) is the gradient operator \(\nabla = (\partial_x, \partial_y)^T\). Using the vertical unit vector \(e_z\), \((C1a)\) can be shown in the geometric formula:

\[
\frac{\partial \omega}{\partial t} + e_z \cdot (\nabla \psi \times \nabla \omega) = 0,
\]

(C2)

where the inner and cross products are used. Eq. \((C2)\) clarifies the covariance of the governing equations because this formula is independent of the coordinate system.

We derive the governing equations in the wavenumber space. The Fourier transform \(\hat{\omega}_k\) is distinguished with the subscript \(k\) from \(\omega(x)\). Eqs. \((C1a)\) and \((C1b)\) are combined into the following:

\[
\frac{\partial \hat{\omega}_k}{\partial t} + \int dl \int dm \hat{a}(l, m) \hat{\omega}_k \delta(l - m - k) = 0, \quad \text{(C3a)}
\]

\[
a(l, m) = \frac{(|l|^2 - |m|^2)}{2|l|^2|m|^2}, \quad \text{(C3b)}
\]

where \(k, l,\) and \(m\) are wavenumbers vectors (e.g., \(k = (k_x, k_y)^T\)) and \(\delta(x)\) is the Dirac delta function. The advection term, namely the second term in \((C3a)\), describes triad interactions \((22)\) where the coefficient \(a(l, m)\) is symmetric with respect to \(l\) and \(m\).

To add a nudging term, Eqs. \((C3)\) are first solved for a lower truncation wavenumber \(k_l\), and the obtained solution is denoted by \(\hat{\omega}_{k_l}\) where \(|k| \leq k_l\). In high-resolution simulations, a nudging term is added to the right-hand side of \((C3a)\):

\[
\frac{\partial \hat{\omega}_k}{\partial t} + \int dl \int dm \hat{a}(l, m) \hat{\omega}_k \delta(l + m - k) = -\lambda \left( \hat{\omega}_k - \hat{\omega}_{k_l} \right) \chi_{|k| \leq k_l},
\]

(C4)

where \(\chi_{|k| \leq k_l}\) is an indicator function, which is equal to 1 when \(|k| \leq k_l\) and zero otherwise. The nudging term vanishes for higher wavenumbers. The condition of \(\chi\) can be changed to be more complex. The coefficient \(\lambda (> 0)\) generally depends on time or wavenumber, while \(\lambda\) was constant in this study.

2. Covariance to rotation

We confirm the covariance of the governing equations \((C3a)\) and \((C4)\) with respect to rotation. A wavenumber vector is transformed in the same way as a position vector because the inner product is invariant: \(k \cdot x = k' \cdot x'\), where \(x' = \rho \theta x\), \(k' = \rho \theta k\), and \(\rho \theta\) is a rotation matrix. The Fourier transform of the vorticity is changed associated with the rotation \((\rho = \rho \theta)\) as follows:

\[
\hat{\omega}_k \rightarrow \hat{\omega}_{\rho^{-1}k},
\]

(C5)

which corresponds to \((3a)\) in the Euclidean space.

The covariance of \((C3a)\) without the nudging term is examined. The first term is clearly independent of rotation because it is the derivative of the scalar \(\hat{\omega}_k\) with respect to time. The second term also does not change its form under rotation. The delta function is invariant because the condition \(l + m - k = 0\) does not depend on the coordinates. The term \(l_m n_l - l_m n_l\) is the vertical component of the cross product \(l \times m\); hence, its form is preserved by two-dimensional rotation. The differential of \(dl \int dm\) is invariant because the Jacobian is the unity. The other parts in the second term such as \(|l|^2\) are invariant because they are scalars. Therefore, the whole equation of \((C3)\)
is covariant with respect to rotation, that is, the form of (C3) is independent of rotation.

The covariance of (C4) including the nudging term is examined. Since the left-hand side of (C4) is the same as that of (C3a), only the right-hand side is discussed. The indicator function is invariant because its condition $|k| \leq L^2$ is described by scalars. The prescribed solution $\omega_k^f$ is covariant because it is a solution of (C3). Therefore, the whole equation (C4) is covariant to rotation. The governing equations do not change the forms even when the spectral nudging term is added.

3. Covariance to translation

The covariance with respect to translation is confirmed. The vorticity $\omega(x)$ is converted to $\omega(x - r)$ with a given translation vector $r$. The Fourier transform is then converted as follows:

$$\omega_k \rightarrow \omega_k e^{-ik \cdot r}.$$  

(C6)

Note that the wavenumber vector $k$ itself is unchanged under translation.

The governing equations (C3) are covariant if the following quantity (i.e., the left-hand side) is zero when a translation acts:

$$\frac{\partial \omega_k}{\partial t} + \int dldm a(l, m) \omega_l \omega_m \delta(l + m - k).$$  

(C7)

Due to the delta function, the coefficient of exponentials of the second term is equal to that of the first term:

$$e^{-il \cdot r} e^{-im \cdot r} \delta(l + m - k) = e^{-ik \cdot r} \delta(l + m - k).$$  

(C8)

This leads to that (C7) is zero; hence, the governing equations (C3) do not change the form under translation.

The covariance still holds when the nudging term, namely the right-hand side of (C4), is added. The nudging term is linear with respect to the vorticity and is converted in the same way as the time derivative.

The governing equations are covariant even when a combination of rotation and translation acts because each transformation does not change the forms of the equations.

4. Configuration of the barotropic instability experiment

This subsection describes the configuration of the barotropic instability experiment in Section IVB. Basically, the configuration is the same as that of the barotropic instability experiment without the spectral nudging in Section IIIA.

The LR simulation was performed by numerically solving (C3) with the spectral model as in Section IIIA. The truncation wavenumber was 10, which determined the alias-free grids of $128 \times 65$, which was the same as in Section IIIA. In calculating the nudging term in (C4), the indicator function was replaced with $\chi_{|k| < L}$ or $|k| < L$, which is invariant only to the rotation by multiples of 90°. The linear interpolation in time was employed for the LR solution $\omega_k^f$. This interpolation makes $\omega_k^f$ dependent on $\omega_k^f$ in the graphical model of Fig. 14. The initial and boundary conditions were set to be the same as those in the LR simulation. The fluid system has the symmetries of rotation by 180° and translation along the x direction. In this case, a similar discussion to those in Sections C1 and C3 is possible if transformations are confined to a subgroup of $SE(2)$.

Appendix D: Additional analyses

This section presents additional analyses on the decaying turbulence experiment.

1. Dependency on the number of trainable parameters

We examine the test error dependency on the number of trainable parameters. The number of parameters was controlled here by changing the number of channels in the hidden layers. The number of channels was varied by the same factor in all hidden layers. Fig. 18 suggests that the equivariant models (Eq-DSC/MS and Eq-RRDN) are more accurate than the non-equivariant ones when the number of trainable parameters is in the same order. The crosses in the figure denote the original configurations used in Section IVB, where the total sizes of kernels in the convolution are almost the same for Eq-DSC/MS and DSC/MS and for Eq-RRDN and RRDN.

×9 SR of subsampled velocity

FIG. 18. Norm error ratios (NERs) against various numbers of trainable parameters in the super-resolution model of velocity in the decaying turbulence. The NER is defined by (7). The scale factor $s$ was 9 and the low-resolution data were created by subsampling. The crosses stand for the configurations used in Section IVB.
2. Dependency on the size of training data

A smaller number of trainable parameters implies that an equivariant model is trainable with a smaller size of training data. Fig. 19 shows the NERs against the various sizes of training data. As the data size is decreased, the NERs tend to increase for all SR models. The NERs of the equivariant models are smaller than those of the non-equivariant ones, and the difference in the NER tends to be larger as the data size is decreased. Therefore, if only a small size of data is available, an equivariant model should be used, i.e., the prior knowledge on the geometric symmetry should be incorporated into the neural network design.

3. Dependency on the unit of angle of discrete rotations

In practice, fluid fields need to be discretized for the calculation on computers. The order of discretization of rotations is controlled here by the order \( N \) of a cyclic group \( C_N \) used in the equivariant CNNs. For instance, if \( N = 3 \), the unit of angle of rotations is \( 120^\circ = (360^\circ / 3) \). This subsection demonstrates that the test errors do not necessarily decrease as \( N \) is increased, unlike in the image classification.

The spatial size of kernels usually needs to be larger when the order of \( C_N \) is increased because the interpolation is used to determine the kernel pattern. The RRDN is appropriate for an experiment varying both \( C_N \) and kernel size because all kernels have the same spatial size (Fig. 16). The RRDN consisting of one residual in residual dense block was employed here to reduce the training cost. In each hidden layer, eight sets of the regular representation were used for the input and output. For instance, the number of the input/output channels is 32 for \( C_4 \) and 96 for \( C_{12} \). Fig. 21 shows the NERs and ESEs against the various \( C_N \) and spatial sizes of kernels. The errors are about 10% smaller for the \( 5 \times 5 \) kernels than those of the \( 3 \times 3 \) kernels. This error reduction does not strongly depend on \( C_N \). We confirmed that the errors were also decreased for the non-equivariant RRDN when the \( 5 \times 5 \) kernels were used instead of \( 3 \times 3 \). The decrease in error was comparable between the RRDN and Eq-RRDN. If computational resources are enough, the \( 5 \times 5 \) kernels should be used for both models. A purpose of this study is to compare the equivariant CNNs with the non-equivariant ones. Since a large difference in accuracy was not observed with the \( 5 \times 5 \) kernels, we adopted the kernels of \( 3 \times 3 \) in Sections IV and V.
FIG. 21. Norm error ratios (NERs) and energy spectrum errors (ESEs) against the various \( C_N \) and spatial sizes of kernels used in the Eq-RRDN. The NER and ESE are defined in (7) and (8), respectively. The decaying turbulence velocity was used to create the figure, where the scale factor 9 was 9 and the low-resolution data were created by subsampling.

4. Comparison of super-resolution between velocity and vorticity

We have discussed the super-resolution of velocity so far. Similar results to those shown above were obtained in the super-resolution of vorticity. This subsection compares the vorticity calculated from the super-resolved velocity with the vorticity that is directly super-resolved. The former is called the velocity-SR, and the latter the vorticity-SR.

Fig. 22 compares the test scores between the velocity-SR and vorticity-SR. The vorticity spectrum error (VSE) was calculated instead of the ESE. The VSE is defined as in (3), in which the energy spectra are replaced with the enstrophy spectra. The scale factor 9 was 9 in the figure; similar results were obtained for the other 9. A clear tendency is not observed for the NER. In the case of subsampling, the NER is smaller in the velocity-SR. In the local average, the NER is smaller in the vorticity-SR for the DSC/MS and Eq-DSC/MS, while the NER is comparable for the RRDN and Eq-RRDN. Importantly, the difference in the NER between the velocity-SR and vorticity-SR is limited to about 10%. On the contrary, the VSE shows the clear tendency. The VSEs of the velocity-SR are about twice as small as those of the vorticity-SR. Therefore, we can conclude that the vorticity calculated from the super-resolved velocity is more physically valid than the vorticity that is super-resolved directly.

The lower accuracy of the vorticity-SR may not be attributed to the difference between scalar and vector. The equivariant models incorporate the transformation law of scalars or vectors, which assures that the vorticity calculated from the super-resolved velocity is a scalar field. In contrast, the non-equivariant models do not necessarily make such a guarantee. Despite this difference, the large VSEs are observed in both types of the models (Fig. 22). Vorticity fields generally have finer structures due to the differential operation than velocity fields. The difference in scale might have caused the difference in accuracy of the super-resolution.

References

1. L. Brunton, B. R. Noack, and P. Koumoutsakos, “Machine learning for fluid mechanics,” Annual Review of Fluid Mechanics 52, 477–508 (2020). https://doi.org/10.1146/annurev-fluid-010719-060214
2. D. Daruasisam, “Perspectives on machine learning-augmented Reynolds-averaged and large eddy simulation models of turbulence,” Phys. Rev. Fluids 6, 050504 (2021)
3. X. Vinyaes and S. L. Brunton, “The potential of machine learning to enhance computational fluid dynamics,” (2021), arXiv:2110.02085 [physics.flu-dyn]
4. S. L. Brunton, “Applying machine learning to study fluid mechanics,” Acta Mechanica Sinica (2022), 10.1007/s10409-021-01143-6
5. M. Raissi, P. Perdikaris, and G. Karniadakis, “Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations,” Journal of Computational Physics 378, 686–707 (2019)
6. K. Kashinath, M. Mustafa, A. Albert, J.-L. Wu, C. Jiang, S. Esmaeizadeh, K. Azizzadenesheli, R. Wang, A. Chattopadhyay, A. Singh, A. Manepalli, D. Chirila, R. Yu, R. Walters, B. White, H. Xiao, H. A. Tchelepi, P. Marcus, A. Anandkumar, P. Hassanzadeh, and n. Prabhat, “Physics-informed machine learning: case studies for weather and climate modelling,” Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 379, 20200093 (2021) https://royalsocietypublishing.org/doi/pdf/10.1098/rsta.2020.0093
7. S. Cai, Z. Mao, Z. Wang, M. Yin, George, and E. Karniadakis, “Physics-informed neural networks (pinn) for fluid mechanics: a review,” Acta Mechanica Sinica (2022), 10.1007/s10409-021-01148-1
8. B. F. Schutz, Geometrical Methods of Mathematical Physics, first edition ed. (Cambridge University Press, 1980).
9. M. Nakahara, Geometry, Topology and Physics, second edition ed. (CRC Press, 2003).
10. A. M. Dirac, General Theory of Relativity (Princeton University Press, 1996).
11. M. Bronstein, J. Bruna, T. Cohen, and P. Velickovic, “Geometric deep learning: Grids, groups, geodesics, and gauges,” CoRR abs/2010.13478 (2021) [physics.flu-dyn]
12. K. Atz, F. Grisoni, and G. Schneider, “Geometric deep learning on molecular representations,” Nature Machine Intelligence 3, 1023–1032 (2021)
13. R. Wang, R. Walters, and R. Yu, “Incorporating symmetry into deep dynamics models for improved generalization.” (2021).
14. J. S. Suk, P. de Haan, P. Lippe, C. Brune, and J. M. Wolterink, “Equivariant graph neural networks as surrogate for computational fluid dynamics in 3d artery models,” in Fourth Workshop on Machine Learning and the Physical Sciences (NeurIPS 2021).
15. B. Siddani, S. Balachandar, and R. Fang, “Rotational and reflection equivariant convolutional neural network for data-limited applications: Multiphase flow demonstration,” Physics of Fluids 33, 103323 (2021) https://doi.org/10.1063/5.0066049
16. S. Pawar, O. San, A. Rasheed, and P. Vedula, “Frame invariant neural network closures for Kraichnan turbulence,” (2022), arXiv:2201.02928 [physics.flu-dyn]
17. J. Han, X.-H. Zhou, and H. Xiao, “Vcnn-e: A vector-cloud neural network with equivariance for emulating Reynolds stress transport equations,” (2022), arXiv:2201.01287 [physics.flu-dyn]
18. C. Dong, C. C. Loy, K. He, and X. Tang, “Learning a deep convolutional network for image super-resolution,” in Computer Vision – ECCV 2014, edited by D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars (Springer International Publishing, Cham, 2014) pp. 184–199.
19. C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi, “Photo-realistic single image super-resolution using a generative adversarial network,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2017).
20. X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, C. Dong, Y. Qiao, and C. Change Loy, “Esrgan: Enhanced super-resolution generative adversarial networks,” in Proceedings of the European Conference on Computer Vision (ECCV) Workshops (2018).
FIG. 22. Norm error ratio (NER) and vorticity spectrum error (VSE) calculated from the super-resolved velocity and vorticity. For “Velocity-SR” the vorticity was calculated from the super-resolved velocity, while for “Vorticity-SR” the vorticity was directly super-resolved. The NER is defined in (7), where the vorticity is evaluated. The VSE is defined as in (8), in which the energy spectra are replaced with the enstrophy spectra. The scale factor $s$ was 9, that is, the SR models made the resolution 9 times finer.
downscaling time-evolving atmospheric fields with a generative adversarial network,” IEEE Transactions on Geoscience and Remote Sensing, 1–13 (2020).

48. K. Stengel, A. Glaws, D. Hettinger, and R. N. King, “Adversarial super-resolution of climatological wind and solar data,” Proceedings of the National Academy of Sciences 117, 16805–16815 (2020).

49. T. N. Wang, Z. Liu, I. Foster, W. Chang, R. Kettimuthu, and V. R. Kotamarti, “Fast and accurate learned multiresolution data downsampling for precipitation,” Geoscientific Model Development 14, 6355–6372 (2021).

50. Y. Yasuda, R. Onishi, Y. Hirokawa, D. Kolomenskiy, and D. Sugiyama, “Super-resolution of near-surface temperature utilizing physical quantities for real-time prediction of urban micrometeorology,” Building and Environment 209, 108597 (2022).

51. P. Cohen and M. Welling, “Group equivariant convolutional networks,” in Proceedings of the 33rd International Conference on Machine Learning, Proceedings of Machine Learning Research, Vol. 48, edited by M. F. Balcan and K. Q. Weinberger (PMLR, New York, New York, USA, 2016) pp. 2990–2999.

52. T. S. Cohen and M. Welling, “Steerable cnns,” 5th International Conference on Learning Representations, ICLR 2017 (2017).

53. D. Marcos, M. Volpi, N. Komodakis, and D. Tuia, “Rotation equivariant vector field networks,” in Proceedings of the IEEE International Conference on Computer Vision (ICCV) (2017).

54. D. E. Worrall, S. J. Garbin, D. Turmukhambetov, and G. J. Brostow, “Harmonic networks: Deep translation and rotation equivariance,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2017).

55. Y. Zhou, Q. Ye, Q. Qiu, and J. Jiao, “Oriented response networks,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2017).

56. K. T. Schütt, F. Arbabzadah, S. Chmiela, K. R. Müller, and A. Tkatchenko, “Article quantum-chemical insights from deep tensor neural networks,” Nature Communications (2017), 10.1038/ncomms13890.

57. E. J. Bekkers, M. W. Lafarge, M. Veta, K. A. J. Epenhof, J. P. W. Pluijm, and R. Duits, “Roto-translation covariant convolutional networks for medical image analysis,” in Medical Image Computing and Computer Assisted Intervention – MICCAI 2018, edited by A. F. Frangi, J. A. Schnabel, C. Davatzikos, C. Alberola-López, and G. Fichtinger (Springer International Publishing, Cham, 2018) pp. 440–448.

58. R. Kundor and S. Trivedi, “On the generalization of equivariance and convolution in neural networks to the action of compact groups,” in Proceedings of the 35th International Conference on Machine Learning, Proceedings of Machine Learning Research, Vol. 80, edited by J. Dy and A. Krause (PMLR, 2018) pp. 2747–2755.

59. T. S. Cohen, M. Geiger, and M. Weiler, “A general theory of equivariant cnns on homogeneous spaces,” in Advances in Neural Information Processing Systems, Vol. 32, edited by H. Wallach, H. Larochelle, A. Beygelzimer, F. d’Alché-Buc, E. Fox, and R. Garnett (Curran Associates, Inc., 2019).

60. M. Weiler and G. Cesa, “General e(2)-equivariant steerable cnns,” in Advances in Neural Information Processing Systems, Vol. 32, edited by H. Wallach, H. Larochelle, A. Beygelzimer, F. d’Alché-Buc, E. Fox, and R. Garnett (Curran Associates, Inc., 2019).

61. R. Gilmore, “Lie Groups, Physics, and Geometry: An Introduction for Physicists, Engineers and Chemists” (Cambridge University Press, 2008).

62. A. Zee, Group Theory in a Nutsheif for Physicists” (Princeton University Press, 2016).

63. M. Weiler, M. Geiger, M. Welling, W. Boomsma, and T. S. Cohen, “3d steerable cnns: Learning rotationally equivariant features in volumetric data,” in Advances in Neural Information Processing Systems, Vol. 31, edited by S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett (Curran Associates, Inc., 2018).

64. N. Thomas, T. Smidt, S. M. Kearnes, L. Yang, L. Li, K. Kohlhoff, and P. Riley, “Tensor field networks: Rotation- and translation-equivariant neural networks for 3d point clouds,” CoRR abs/1802.08219 (2018) arXiv:1802.08219.