Supplementary Material: Achieving Rotational Invariance with Bessel-Convolutional Neural Networks

Valentin Delchevalerie,[1] **Adrien Bibal,**[2] **Benoît Frénay,**[2*] **Alexandre Mayer**[3*]
[1] PReCISE, NADI & naXys institutes, University of Namur, Belgium
[2] PReCISE, NADI institute, University of Namur, Belgium
[3] Department of Physics, naXys institute, University of Namur, Belgium
{valentin.delchevalerie, adrien.bibal, benoit.frenay, alexandre.mayer}@unamur.be

A Orthonormality of the Bessel basis

In this section, the mathematical proof for the orthonormality of the Bessel basis is provided. To do so, one needs two preliminary results.

A.1 Lemma 1

$$\int_0^{2\pi} e^{i\theta(\nu'-\nu)}d\theta = \int_0^{2\pi} \cos\left(\theta\left(\nu'-\nu\right)\right) + i\int_0^{2\pi} \sin\left(\theta\left(\nu'-\nu\right)\right) \tag{1}$$

$$= 2\pi\delta_{\nu,\nu'}, \tag{2}$$

since $\nu' - \nu$ is always an integer.

A.2 Lemma 2

We will need to use the Lommel's integrals, which are in our particular case

$$\int_0^R \rho J_\nu\left(k_{\nu,j}\rho\right) J_\nu\left(k_{\nu,j'}\rho\right) d\rho =$$

$$\begin{cases} \frac{1}{k_{\nu,j'}^2 - k_{\nu,j}^2}\left[\rho\left(k_{\nu,j}J'_\nu\left(k_{\nu,j}\rho\right)J_\nu\left(k_{\nu,j'}\rho\right) - k_{\nu,j'}J'_\nu\left(k_{\nu,j'}\rho\right)J_\nu\left(k_{\nu,j}\rho\right)\right)\right]_0^R \text{ if } k_{\nu,j} \neq k_{\nu,j'}, \\ \left[\frac{\rho^2}{2}\left[J'^2_\nu\left(k_{\nu,j}\rho\right) + \left(1 - \frac{\nu^2}{k_{\nu,j}^2\rho^2}\right)J_\nu\left(k_{\nu,j}\rho\right)\right]\right]_0^R \text{ otherwise.} \end{cases} \tag{3}$$

By evaluating the $[.]_0^R$, and by taking into account that $J'_\nu\left(k_{\nu,j}R\right) = 0$ (the condition we imposed to set the $k_{\nu,j}$), these integrals lead to

$$\int_0^R \rho J_\nu\left(k_{\nu,j}\rho\right) J_\nu\left(k_{\nu,j'}\rho\right) d\rho = \begin{cases} 0 \text{ if } k_{\nu,j} \neq k_{\nu,j'} \\ \left(\frac{R^2}{2} - \frac{\nu^2}{2k_{\nu,j}^2}\right) J_\nu\left(k_{\nu,j}R\right) \text{ otherwise.} \end{cases} \tag{4}$$

$$= \left(\frac{R^2}{2} - \frac{\nu^2}{2k_{\nu,j}^2}\right) J_\nu\left(k_{\nu,j}R\right)\delta_{j,j'} \tag{5}$$

[*] B. Frénay and A. Mayer are co-last authors.

35th Conference on Neural Information Processing Systems (NeurIPS 2021).

A.3 Proof

$$\int_0^{2\pi}\int_0^R \rho\left[N_{\nu,j}J_\nu\left(k_{\nu,j}\rho\right)e^{i\nu\theta}\right]^*\left[N_{\nu',j'}J_{\nu'}\left(k_{\nu',j'}\rho\right)e^{i\nu'\theta}\right]d\theta d\rho \tag{6}$$

$$=\int_0^{2\pi} e^{i\theta\left(\nu'-\nu\right)}\int_0^R \rho N_{\nu,j}J_\nu\left(k_{\nu,j}\rho\right)N_{\nu',j'}J_{\nu'}\left(k_{\nu',j'}\rho\right)d\theta d\rho. \tag{7}$$

By using Lemma 1, it leads to

$$\int_0^{2\pi}\int_0^R \rho\left[N_{\nu,j}J_\nu\left(k_{\nu,j}\rho\right)e^{i\nu\theta}\right]^*\left[N_{\nu',j'}J_{\nu'}\left(k_{\nu',j'}\rho\right)e^{i\nu'\theta}\right]d\theta d\rho \tag{8}$$

$$=2\pi\delta_{\nu,\nu'}\int_0^R \rho N_{\nu,j}J_\nu\left(k_{\nu,j}\rho\right)N_{\nu,j'}J_\nu\left(k_{\nu,j'}\rho\right)d\theta d\rho, \tag{9}$$

and by using Lemma 2, to

$$\int_0^{2\pi}\int_0^R \rho\left[N_{\nu,j}J_\nu\left(k_{\nu,j}\rho\right)e^{i\nu\theta}\right]^*\left[N_{\nu',j'}J_{\nu'}\left(k_{\nu',j'}\rho\right)e^{i\nu'\theta}\right]d\theta d\rho \tag{10}$$

$$=2\pi N_{\nu,j}^2\left(\frac{R^2}{2}-\frac{\nu^2}{2k_{\nu,j}^2}\right)J_\nu\left(k_{\nu,j}R\right)\delta_{\nu,\nu'}\delta_{j,j'}. \tag{11}$$

To conclude this proof, one can show that

$$N_{\nu,j}^2=\frac{1}{2\pi\int_0^R \rho J_\nu^2\left(k_{\nu,j}\rho\right)d\rho} \tag{12}$$

$$=\frac{1}{2\pi\left(\frac{R^2}{2}-\frac{\nu^2}{2k_{\nu,j}^2}\right)J_\nu\left(k_{\nu,j}R\right)}, \tag{13}$$

and then finally,

$$\int_0^{2\pi}\int_0^R \rho\left[N_{\nu,j}J_\nu\left(k_{\nu,j}\rho\right)e^{i\nu\theta}\right]^*\left[N_{\nu',j'}J_{\nu'}\left(k_{\nu',j'}\rho\right)e^{i\nu'\theta}\right]d\theta d\rho=\delta_{\nu,\nu'}\delta_{j,j'}.\square \tag{14}$$

B Checklist

1. For all authors...
   (a) Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope? [Yes] The claim is that we achieve rigorous rotational invariance for $\alpha \in [0, 2\pi)$, which is theoretically proved in Section 4 and tested in Section 5.
   (b) Did you describe the limitations of your work? [Yes] In Section 4.4, we discuss the fact that using B-CNNs generally leads to an increase in the number of parameters in each layer. In Section 6, we point out that a deeper analysis needs to be performed to more easily chose the meta-parameters of B-CNNs.
   (c) Did you discuss any potential negative societal impacts of your work? [N/A]
   (d) Have you read the ethics review guidelines and ensured that your paper conforms to them? [Yes]
2. If you are including theoretical results...
   (a) Did you state the full set of assumptions of all theoretical results? [Yes]
   (b) Did you include complete proofs of all theoretical results? [Yes] Except for Equation (6), because it would be too long, but the proof is provided in the supplementary material.
3. If you ran experiments...

(a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? [Yes] Our implementation of B-CNNs is provided in the supplemental material. All the instructions and codes to reproduce the results are provided, and all the data used are available following the references.

(b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? [Yes]

(c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? [Yes] In our experiments, 99% confidence intervals are provided (obtained over 40 independent runs). Standard deviations are also given for experiments in Section 5

(d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? [No] The training for all the architectures is performed using GPUs. However, the total amount of compute and the type of GPUs is not provided as our implementation of B-CNNs is not, for the moment, as well optimized as the two other methods we tested (G-CNNs and CNNs). The comparison would therefore not make much sense.

4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets...

(a) If your work uses existing assets, did you cite the creators? [Yes] For data, references are given, and for the use of Group-CNNs, authors are cited.

(b) Did you mention the license of the assets? [N/A]

(c) Did you include any new assets either in the supplemental material or as a URL? [Yes] Our B-CNNs implementation is provided in the supplemental material

(d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? [N/A]

(e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? [N/A]

5. If you used crowdsourcing or conducted research with human subjects...

(a) Did you include the full text of instructions given to participants and screenshots, if applicable? [N/A]

(b) Did you describe any potential participant risks, with links to Institutional Review Board (IRB) approvals, if applicable? [N/A]

(c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? [N/A]