Digital Discovery

ROYAL SOCIETY OF CHEMISTRY

COMMUNICATION

View Article Online
View Journal | View Issue

Cite this: *Digital Discovery*, 2024, **3**, 649

Received 15th January 2024
Accepted 24th March 2024

DOI: 10.1039/d4dd00021h

rsc.li/digitaldiscovery

A message passing neural network for predicting dipole moment dependent core electron excitation spectra†

Kiyou Shibata * and Teruyasu Mizoguchi

Absorption near-edge structures in core electron excitation spectra reflect the anisotropy of orbitals in the final transition state and can be utilized for analyzing the local atomic environment, including its orientation. So far, the analysis of fine structures has primarily relied on fingerprint-matching with high-cost experimental or simulated spectra. If core electron excitation spectra, including their anisotropy, can be predicted at a low cost using machine learning, the application range of these spectra will be accelerated and extended to areas such as the orientation and electronic structure analysis of liquid crystals and organic solar cells at high spatial resolution. In this study, we introduce a message-passing neural network, named inversion symmetry-aware directional PaiNN (ISD-PaiNN) for predicting core electron excitation spectra using a unit direction vector in addition to molecular graphs as the input. Utilizing a database of calculated C K-edge spectra, we have confirmed that the network can predict core electron excitation spectra reflecting the anisotropy of molecules. Our model is expected to be expanded to other physical quantities in general that depend not only on molecular graphs but also on anisotropic vectors.

1 Introduction

Fine structures of core electron excitation spectra near the excitation edges, electron-energy loss near edge structure (ELNES) and X-ray absorption near edge structure (XANES or NEXAFS), have been widely used as some of the most effective fingerprints for determination of local atomic structures and electronic states. The fine structures reflect the excitation of an electron from an occupied core orbital to unoccupied orbitals. Since the fine structures differ depending on the symmetry of the orbital against the momentum transfer during the core electron excitation, orientation-dependence of the spectra can be used to analyze the local atomic environments including the anisotropic nature of orbitals. This kind of analysis has been applied for analyzing the orientation of molecules such as polymer films,[1] liquid crystals[2] and organic solar cells[3,4] through analysis of orientation dependence of characteristic peaks corresponding to anisotropic orbitals. Meanwhile, quantitative analysis of fine structures has conventionally been performed by fingerprint matching of reference spectra obtained from standard samples with known structures, but it is difficult to apply to unknown structures.

Recently, comparison with spectral shapes obtained by first-principles calculations has been used to analyze the experimental spectra.[5–7] As the reference spectra can be obtained faster and on a larger scale with calculations than with experiments, databases of simulated core electron excitation spectra have been constructed and published.[8–11] While experimental databases typically contain a few hundred spectra, simulated ones can include over 10 000 spectra. However, first-principles calculations are computationally expensive, and it is still difficult to obtain spectra exhaustively for a large number of candidate structures to identify unknown structures from experimental data without *a priori* structural information. It would be useful to develop a low-computational-cost alternative to first-principles calculations that can obtain spectra exhaustively for a large number of structures.

In this context, attempts have been made to predict spectra from molecular structures using machine learning models, and there have been several reports: prediction from molecular graphs using a message passing graph neural network (GNN),[12] a deep neural network,[13–15] and prediction using neural network ensemble through featurized local structural information.[16] However, as far as the authors know, the prediction of core electron excitation spectra considering anisotropy according to the momentum transfer has not been proposed.

From the perspective of predicting physical quantities in general, many machine learning models for material science, not limited to the spectral example above have been proposed.

Institute of Industrial Science, The University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8505, Japan. E-mail: kiyou@iis.u-tokyo.ac.jp; Fax: +81-3-5452-6319; Tel: +81-3-5452-6320

† Electronic supplementary information (ESI) available. See DOI: https://doi.org/10.1039/d4dd00021h

Previous research in GNN architecture includes notable models such as DimeNet[17] and GemNet[18] that leverage invariance, both emphasizing the utilization of invariance properties. Exploring equivariance, PaiNN[19] and NequIP[20] have been proposed, employing strict equivariance constraints. Moreover, attempts to enhance expressive capabilities by slightly relaxing equivariance constraints and incorporating nonlinear transformations have been performed in SCN[21] and eSCN.[22] However, most models predict physical quantities that depend solely on the structure, overlooking the explicit integration of material anisotropy and directional dependencies. It is crucial to consider such directional dependencies when dealing with anisotropic physical quantities, including strain tensors, resistivity tensors, dielectric polarization, and phenomena like dipole transitions. These physical quantities and phenomena reflect the anisotropic nature of materials, and neglecting them in model construction could limit accurate predictions of material properties.

In this study, we developed a message passing neural network[23] model named inversion symmetry-aware directional PaiNN (ISD-PaiNN) that can predict spectra due to the dipole transition. Our model can predict site-specific spectra which are dependent on not only the molecular graph but also the dipole moment direction relative to the molecular graph. Notably, our model consists of transformations that satisfy the physically required invariance of symmetric operations on the input pair of the molecular graph and dipole moment with respect to dipole transitions, ensuring an inductive bias with respect to the invariance.

2 Methods

2.1 Spatial symmetry constraints of dipole-transition spectra

In the limit of small angle electron scattering or long wavelength of the incident X-ray in ELNES and XANES, respectively, the core electron excitation process can be described in the form of a dipole transition. The transition matrix element for ELNES and XANES from an occupied initial state $|\psi_i\rangle$ to an unoccupied final state $|\psi_f\rangle$ is proportional to $\langle\psi_i|\mathbf{q}\cdot\mathbf{r}|\psi_f\rangle$ and $\langle\psi_i|\hat{\varepsilon}\cdot\mathbf{r}|\psi_f\rangle$, respectively, where $\mathbf{r}$ is the relative position of the excited electron from the core, $\mathbf{q}$ is the momentum transfer, and $\hat{\varepsilon}$ is the polarization vector. In either case, the generalized oscillator strength is proportional to the square of the transition matrix element, and the differential cross section of the energy loss spectrum can be expressed as a function of energy loss E as:

$$\frac{\partial\sigma(E)}{\partial E}\propto\sum_{i,f}\|\hat{\mathbf{n}}\cdot\langle\psi_i|\mathbf{r}|\psi_f\rangle\|^2\delta\big(E-\big(E_f-E_i\big)\big), \tag{1}$$

where $\hat{\mathbf{n}}$ is the unit directional vector parallel to the momentum transfer $\mathbf{q}$ or polarization vector $\hat{\varepsilon}$, and E_i and E_f are the energy eigenvalues of $|\psi_i\rangle$ and $|\psi_f\rangle$, respectively.

Based on eqn (1), we consider the constraints imposed on spectra associated with dipole transitions. Let $\mathcal{G}$ be the structural (molecular or crystal) graph defined by atomic numbers Z_i and positions $\vec{\mathbf{r}}_i$, $\hat{\mathbf{n}}$ the dipole vector, n site index as the input, and the target site-specific spectrum of n-th site in $\mathcal{G}$ for the transition regarding $\hat{\mathbf{n}}$, $S_n(\mathcal{G},\hat{\mathbf{n}})$. The invariance of $S_n(\mathcal{G},\hat{\mathbf{n}})$ against symmetry operations are as follows:

$$\begin{aligned}S_n(\mathcal{G},\hat{\mathbf{n}}) &= S_n(\mathcal{G},i(\hat{\mathbf{n}})) = S_n(i(\mathcal{G}),\hat{\mathbf{n}}) = S_n(R(\mathcal{G}),R(\hat{\mathbf{n}}))\\ &= S_n(T(\mathcal{G}),\hat{\mathbf{n}}),\end{aligned} \tag{2}$$

where i is the spatial inversion, $R\in$ SO(3) is a general 3D rotation, and $T\in T(3)$ is the displacement.

2.2 Model architecture

To satisfy the constraints in eqn (2), we employ the message passing neural network (MPNN)[23] and referred to its variant, the polarizable atom interaction neural network (PaiNN).[19]

2.2.1 PaiNN and its limitations on input. The PaiNN[19] model, classified among message passing neural networks,[23] employs rotationally equivariant representations. It is designed by modeling equivariant interactions in Cartesian space, avoiding equivariant convolutions relying on spherical harmonics and Clebsch–Gordan transforms, which contributes to its conceptual simplicity and computational efficiency.

The model takes atomic numbers $Z_i\in\mathbb{N}$ and atomic positions $\vec{\mathbf{r}}_i\in\mathbb{R}^3$ as inputs. Internally, it employs two node-wise features at each i-th site: an SO(3) equivariant 3D vector $\vec{\mathbf{v}}_i$ and an SO(3) invariant scalar $\mathbf{s}_i$. The features $\vec{\mathbf{v}}_i$ and $\mathbf{s}_i$ at the i-th site are initialized with zero vector $\vec{\mathbf{v}}_i^0=\vec{0}$ and embedded by atomic number Z_i, respectively. Subsequently, they are updated by a message constructed with relative positional vectors between site i and j, $\vec{\mathbf{r}}_{ij}=\vec{\mathbf{r}}_i-\vec{\mathbf{r}}_j$. The transformation of features $\vec{\mathbf{v}}_i$ and $\mathbf{s}_i$ while preserving their equivariance and invariance, respectively, is achieved by properly constructed message and update functions.

The design of the PaiNN implies a dependency of its model output solely on the input structural graph $\mathcal{G}$. Consequently, the PaiNN cannot incorporate additional information beyond $\mathcal{G}$, such as directional information respective to $\mathcal{G}$.

2.2.2 Element specific non-zero initialization of the node-wise vector feature. Expanding on the above, we introduce directional details by setting node-wise vectors with non-zero values. These initial vectors need only adhere to the symmetry requirements concerning their orientation, allowing flexibility in their magnitude. Consequently, we initialized vector $\vec{\mathbf{v}}_i$ using a scaled vector constructed using the embedded representation specific to each elemental species:

$$\vec{\mathbf{v}}_i^0=\mathbf{b}_{Z_i}\otimes\hat{\mathbf{n}}\in\mathbb{R}^{F\times 3}, \tag{3}$$

where $\mathbf{b}_{Z_i}\in\mathbb{R}^{F\times 1}$ is the embedded representation for the Z_i-th element, and F is the dimension of the embedded representation.

2.2.3 Message and update blocks. The invariance of the features under T on $\mathcal{G}$ is satisfied by using translational invariant $\vec{\mathbf{r}}_{ij}$ for the message block. The invariance and equivariance of features under R on both $\mathcal{G}$ and $\hat{\mathbf{n}}$ are also satisfied by the message function in the PaiNN. However, the message function for $\vec{\mathbf{v}}_i$ in the PaiNN is asymmetric under the inversion operation i on either $\mathcal{G}$ or $\hat{\mathbf{n}}$, which also results in $\mathbf{s}_i$ being asymmetric under i due to the mixing of $\vec{\mathbf{v}}_i$ and $\mathbf{s}_i$ in the message and update functions. To satisfy the constraints concerning invariance against space inversion symmetry operation i, *i.e.* $S_n(i(\mathcal{G}),\hat{\mathbf{n}})=S_n(\mathcal{G},i(\hat{\mathbf{n}}))=S_n(\mathcal{G},\hat{\mathbf{n}})$ intrinsically, the message

function for $\vec{\mathbf{v}}_i$, $\Delta\vec{\mathbf{v}}_i^m$ is modified slightly from one in the PaiNN as follows:

$$\Delta\vec{\mathbf{v}}_i^m = \sum_j \vec{\mathbf{v}}_j \circ \phi_{vv}(s_j) \circ \mathcal{W}_{vv}(\|\vec{r}_{ij}\|) + \sum_j (\vec{\mathbf{v}}_j \cdot \vec{r}_{ij}) \circ \phi_{vs}(s_j) \circ \mathcal{W}'_{vs}(\|\vec{r}_{ij}\|) \frac{\vec{r}_{ij}}{\|\vec{r}_{ij}\|}, \quad (4)$$

where ϕ_{vv} and ϕ_{vs} are the activation functions, and $\mathcal{W}'_{vv}$ and $\mathcal{W}'_{vs}$ are the weight functions following eqn (8) in ref. 19. The modification is only on the second term and is just multiplying the inner product $\vec{\mathbf{v}}_j \cdot \vec{r}_{ij}$ which is odd against $\vec{\mathbf{v}}_i$ and $\vec{r}_{ij}$ so that the overall message function is odd and even against the inversion operation i on $\vec{\mathbf{v}}_i$ and $\vec{r}_{ij}$, respectively.

The other message function for $\mathbf{s}_i$ and the update functions are the same as those of the PaiNN. This modification makes $\mathbf{s}_i$ to be updated with $\vec{\mathbf{v}}_i$, which reflects the input direction $\hat{\mathbf{n}}$. As a result, $\mathbf{s}_i$ possesses even symmetry against i on $\mathcal{G}$ and $\hat{\mathbf{n}}$, which is the required invariance for $S_n(\mathcal{G}, \hat{\mathbf{n}})$ as described in eqn (2).

2.2.4 Output block. The non-zero initialization of $\vec{\mathbf{v}}_i$ by using eqn (3) and the modified message function by using eqn (4) make $\mathbf{s}_i$ satisfy the required symmetry in eqn (2). The prediction of the site-specific spectrum is performed by converting $\mathbf{s}_i$ at the site of interest by using a multilayer perceptron consisting of two fully connected layers.

2.3 Dataset and training for evaluation

To validate the model, a spectral dataset that includes dipole vector dependent and site-specific oscillation intensity is needed. In this study, we validate our model with a simulated C K-edge spectral database of organic molecules,[10] which is constructed by first-principles calculations based on density functional theory (DFT) and contains 117 340 site-specific spectra for three typical directions of dipole vectors of symmetrically unique sites in 22 155 molecules with no more than eight non-hydrogen atoms (C, O, N, and F) in the QM9 database.[24,25] We preprocessed the spectra with a Gaussian smearing of 0.5 eV in the energy range of 288–310 eV sampled equally spaced as 256 dimensional vectors for the objective variable. Spectral intensities were normalized by scaling the entire data set so that the averaged intensity of each spectrum in the energy range considered was 1, while preserving the relative magnitude relationship between the spectra at each site.

The model was developed by modifying the PaiNN implemented in the Open Catalyst Project's library ocp[26] with PyTorch Geometric[27] and PyTorch,[28] based on the description in Section 2.2.2–2.2.4. Mini-batch learning with a batch size of 32 molecular graphs was performed. The loss function is the mean squared error (MSE) between the predicted and reference spectra on carbon sites with available spectra. We used the Adam[29] optimizer with a learning rate of 0.0001 for the learning process, implementing early stopping with a patience of 40 epochs.

3 Results and discussion

We evaluated the prediction performance using the processed C K-edge spectral dataset described in Section 2.3. The dataset was randomly divided: 80% for training and validation (in a ratio of 0.8 : 0.2) and 20% for testing based on molecules. The number of anisotropic and site-specific spectra for training, validation, and testing were 221 274, 55 653, and 69 477, respectively.

Fig. 1a shows the distribution of MSE, confirming that the over 95% data points have a MSE below 0.1. Fig. 1b–e show the predicted spectra (red line) and calculated results (gray line), sampled at typical percentiles at 0, 50, 95, and 100%, respectively as shown in Fig. 1a. For MSEs below 0.1 (Fig. 1b–d), the predicted spectra show good agreement with the calculated ones. Particularly, for the top 50% of MSEs, the predicted spectra closely resemble the calculated spectra. In the case of the worst-performing prediction (Fig. 1e), the general trend of peak distribution around 297 eV rising and falling at 305 eV is reproduced. However, accurate prediction of the major peaks is lacking. This specific spectrum corresponds to the trifluorocarbon site in molecule id #23880 (refer to ESI Fig. S7†). The poor prediction results might stem from the limited trifluorocarbon site data, consisting of only 87 sites within the 117 340 sites in the database, and the distinct spectral features originating from chemical interactions with highly electronegative fluorine. Concerning the value of the MSE and metrics for spectral similarity, we evaluated the MSE and its use by comparing it with another metric by simulating spectra with added noise (see the ESI†).

Predictions account for relative intensity enables the derivation of molecular spectra by aggregating site-specific ones. Fig. 2 displays both the site-specific spectra and the molecular spectra of a representative molecule (id #10578) from QM9, which represents the median MSE for the molecular spectra (refer to the ESI†). Not only the site-specific spectra but the molecular spectra are also predicted correctly including their dipole moment direction dependence, which confirms potential applications for molecular structure analysis taking advantage of the anisotropic dipole transition.

In order to investigate the effectiveness of non-zero initialization of $\vec{\mathbf{v}}_i$ as given in eqn (3) and the modified message block outlined in eqn (4), we conducted ablation experiments by systematically altering these components and assessing the resulting prediction performances of the models (see the ESI†).

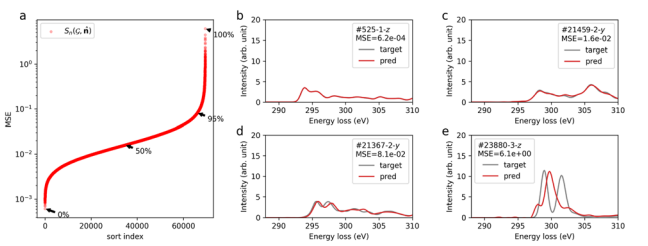

Fig. 1 Prediction results of site-specific anisotropic C K-edge spectra $S_n(\mathcal{G}, \hat{\mathbf{n}})$. (a) Sorted MSE of the prediction on the test dataset. (b–e) Predicted (red) and calculated (gray) spectra for typical percentiles in terms of MSE loss (0, 50, 95, and 100%) as denoted in (a). The insets of (b–e) shows the molecule id in QM9 corresponding $\mathcal{G}$, site index n, directional vector $\hat{\mathbf{n}}$, and the MSE value.

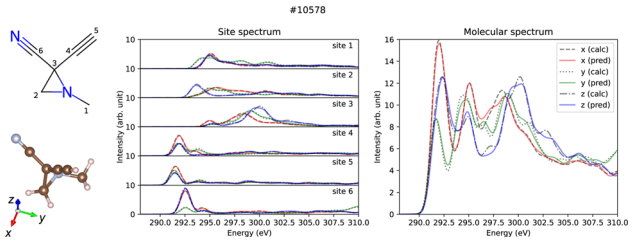

Fig. 2 Prediction results of site-specific and molecular anisotropic C K-edge spectra for molecule id #10578 in QM9. The left panel shows the molecular structure formula and the three-dimensional structure drawn by RDKit[30] and VESTA,[31] respectively. Middle and right panels show the predicted (solid lines) and calculated (dotted lines) spectra for site-specific spectra for each C site and molecular spectra, respectively.

The comparative analysis among the models reveals the crucial role of the non-zero initialization of $\vec{\mathbf{v}}_i$ in accurate predictions. In addition, the utilization of the symmetrized message block demonstrates its effectiveness in improving overall prediction performance. We also evaluated the generalization performance of the model by scaffold splitting, a test on larger molecules, and prediction of orientation dependence of a specific molecule (see the ESI†). The results confirm the model's generalization ability to predict the spectra of molecules with different sizes and shapes. For the efficiency and speed of spectrum prediction, we also compared the computational cost of our model with that of first-principles calculations, and comfirmed that our model is 10^6 times faster than first-principles calculations (see the ESI†).

4 Conclusion

We propose a message passing neural network based on the PaiNN for predicting core electron excitation spectra considering its constraints on the physically required symmetry. We tested the model on C K-edge spectra of organic molecules, and it can predict the general shape of most of the site-specific spectra and molecular spectra including their dependence on the dipole moment direction.

So far, analysis considering anisotropy of core electron excitation spectra requires a highly accurate reference, whether by calculations or an experiment. The cost of obtaining such a reference may become a bottleneck in the future, and our model, which is low-cost compared to first-principles calculations, will be useful for analyzing materials such as polymers, liquid cells, and organic solar cells.

Furthermore, the proposed model architecture is extended to be applied to targets other than ELNES/XANES by setting the required physical symmetry as an inductive bias to predict spectral data that physically depend on the orientation to molecules and crystals, which can advance GNN in materials science by explicitly incorporating directional dependencies in physical quantities alongside the representation of material graphs. This new approach complements the existing GNN paradigm by filling the gap in addressing directional aspects, fostering more accurate predictions in materials science. We anticipate that this novel strategy will contribute to an improved understanding of material anisotropy, offering innovative methodologies for detailed characterization.

Data availability

The code for the model IDS-PaiNN, scripts for training, and the data of the time comparison can be found at **https://github.com/nmdl-mizo/isdpainn** with **https://doi.org/10.5281/zenodo.10554765**. The simulated C-K edge spectral dataset for evaluation is originally published in ref. 10 and can be downloaded at **https://doi.org/10.6084/m9.figshare.c.5494395.v1**. The code of the dataset class for the C-K edge spectral dataset in PyTorch Geometric is available as a Python module C-K edge maker at **https://github.com/nmdl-mizo/ck_edge_maker** with **https://doi.org/10.5281/zenodo.6352029**. The weights of the trained model and MSEs for random split and scaffold split discussed in this paper are available at **https://doi.org/10.5281/zenodo.10547719**. The configuration, weights, and MSEs of the models for the ablation experiments are available at **https://doi.org/10.5281/zenodo.10566201**. The additional calculated data for evaluation of the amino acids and benzene rotation series are available on the NOMAD repository at **https://doi.org/10.17172/NOMAD/2024.01.23-1** and **https://doi.org/10.17172/NOMAD/2024.01.23-2**, respectively.

Author contributions

K. S.: conceptualization, data curation, formal analysis, investigation, methodology, software, writing. T. M.: project administration, supervision, funding acquisition.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This study was supported by the Grant-in-Aid for Scientific Research (Grant No. 19H05787) from the MEXT and CREST (Grant No. JPMJCR1993) from the JST. The authors thank Kento Nishio, Naoto Kawaguchi, and Izumi Takahara for their helpful discussions.

Notes and references

1 M. G. Samant, J. Stöhr, H. R. Brown, T. P. Russell, J. M. Sands and S. K. Kumar, *Macromolecules*, 1996, **29**, 8334–8342.
2 N. Kawatsuki, Y. Taniguchi, M. Kondo, Y. Haruyama and S. Matsui, *Macromolecules*, 2015, **48**, 2203–2210.
3 U. Aygül, D. Batchelor, U. Dettinger, S. Yilmaz, S. Allard, U. Scherf, H. Peisert and T. Chassé, *J. Phys. Chem. C*, 2012, **116**, 4870–4874.
4 H. Yao, L. Ye, H. Zhang, S. Li, S. Zhang and J. Hou, *Chem. Rev.*, 2016, **116**, 7397–7457.
5 J. J. Rehr and R. C. Albers, *Rev. Mod. Phys.*, 2000, **72**, 621–654.

6 J. Rehr and A. Ankudinov, *Coord. Chem. Rev.*, 2005, **249**, 131–140.
7 H. Ikeno and T. Mizoguchi, *Microscopy*, 2017, **66**, 305–327.
8 K. Mathew, C. Zheng, D. Winston, C. Chen, A. Dozier, J. J. Rehr, S. P. Ong and K. A. Persson, *Sci. Data*, 2018, **5**, 180151.
9 Y. Chen, C. Chen, C. Zheng, S. Dwaraknath, M. K. Horton, J. Cabana, J. Rehr, J. Vinson, A. Dozier, J. J. Kas, K. A. Persson and S. P. Ong, *Sci. Data*, 2021, **8**, 153.
10 K. Shibata, K. Kikumasa, S. Kiyohara and T. Mizoguchi, *Sci. Data*, 2022, **9**, 214.
11 H. Guo, M. R. Carbone, C. Cao, J. Qu, Y. Du, S.-M. Bak, C. Weiland, F. Wang, S. Yoo, N. Artrith, A. Urban and D. Lu, *Sci. Data*, 2023, **10**, 349.
12 M. R. Carbone, M. Topsakal, D. Lu and S. Yoo, *Phys. Rev. Lett.*, 2020, **124**, 156401.
13 M. M. M. Madkhali, C. D. Rankine and T. J. Penfold, *Phys. Chem. Chem. Phys.*, 2021, **23**, 9259–9269.
14 C. D. Rankine and T. J. Penfold, *J. Chem. Phys.*, 2022, **156**, 164102.
15 L. Watson, C. D. Rankine and T. J. Penfold, *Phys. Chem. Chem. Phys.*, 2022, **24**, 9156–9167.
16 A. Ghose, M. Segal, F. Meng, Z. Liang, M. S. Hybertsen, X. Qu, E. Stavitski, S. Yoo, D. Lu and M. R. Carbone, *Phys. Rev. Res.*, 2023, **5**, 013180.
17 J. Gasteiger, J. Groß and S. Günnemann, Directional Message Passing for Molecular Graphs, *arXiv*, 2022, preprint, arXiv:2003.03123, DOI: **10.48550/arXiv.2003.03123**.
18 J. Gasteiger, F. Becker and S. Günnemann, Universal Directional Graph Neural Networks for Molecules, *arXiv*, 2022, preprint, arXiv:2106.08903, DOI: **10.48550/arXiv.2106.08903**.
19 K. T. Schütt, O. T. Unke and M. Gastegger, Equivariant message passing for the prediction of tensorial properties and molecular spectra, *arXiv*, 2021, preprint, arXiv:2102.03150, DOI: **10.48550/arXiv.2102.03150**.
20 S. Batzner, A. Musaelian, L. Sun, M. Geiger, J. P. Mailoa, M. Kornbluth, N. Molinari, T. E. Smidt and B. Kozinsky, *Nat. Commun.*, 2022, **13**, 2453.
21 C. L. Zitnick, A. Das, A. Kolluru, J. Lan, M. Shuaibi, A. Sriram, Z. Ulissi and B. Wood, Spherical Channels for Modeling Atomic Interactions, *arXiv*, 2022, preprint, arXiv:2206.14331, DOI: **10.48550/arXiv.2206.14331**.
22 S. Passaro and C. L. Zitnick, Reducing SO(3) Convolutions to SO(2) for Efficient Equivariant GNNs, *arXiv*, 2023, preprint, arXiv:2302.03655, DOI: **10.48550/arXiv.2302.03655**.
23 J. Gilmer, S. S. Schoenholz, P. F. Riley, O. Vinyals and G. E. Dahl, *Proceedings of the 34th International Conference on Machine Learning*, 2017, pp. 1263–1272.
24 L. Ruddigkeit, R. van Deursen, L. C. Blum and J.-L. Reymond, *J. Chem. Inf. Model.*, 2012, **52**, 2864–2875.
25 R. Ramakrishnan, P. O. Dral, M. Rupp and O. A. von Lilienfeld, *Sci. Data*, 2014, **1**, 140022.
26 L. Chanussot, A. Das, S. Goyal, T. Lavril, M. Shuaibi, M. Riviere, K. Tran, J. Heras-Domingo, C. Ho, W. Hu, A. Palizhati, A. Sriram, B. Wood, J. Yoon, D. Parikh, C. L. Zitnick and Z. Ulissi, *ACS Catal.*, 2021, **11**(10), 6059–6072.
27 M. Fey and J. E. Lenssen, ICLR Workshop on Representation Learning on Graphs and Manifolds, *arXiv*, 2019, preprint, arXiv:1903.02428, DOI: **10.48550/arXiv.1903.02428**.
28 A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga and A. Lerer, *31st Conference on Neural Information Processing Systems (NIPS 2017)*, Long Beach, CA, USA, 2017.
29 D. P. Kingma and J. Ba, A Method for Stochastic Optimization, *arXiv*, 2017, preprint, arXiv:1412.6980, DOI: **10.48550/arXiv.1412.6980**.
30 RDKit: Open-source cheminformatics, **http://www.rdkit.org**.
31 K. Momma and F. Izumi, *J. Appl. Crystallogr.*, 2011, **44**, 1272–1276.