

# Supplemental Material:

## Equivariant message passing for the prediction of tensorial properties and molecular spectra

### A Data

#### A.1 Reference data for substituted ferrocene

Reference computations for substituted ferrocene were carried out with the semi-empirical GFN2-xTB method[1] using the `xtb` package (<https://github.com/grimme-lab/xtb>). Training structures were generated by normal mode sampling[2] at 300 K starting from the global minimum structure and randomly rotating the cyclopentadienyl (Cp) moieties relative to each other around the Cp-Fe-Cp axis by angles uniformly sampled from  $[0, 2\pi]$ . The rotational potential energy profile was generated by sampling a full rotation around the Cp-Fe-Cp axis using 1k steps, starting from the global minimum structure and keeping all other degrees of freedom fixed at their equilibrium positions.

#### A.2 Reference data for infrared and Raman spectra

Electronic structure reference computations for aspirin were carried out at the PBE0/def2-TZVP[3, 4] level of theory using the ORCA quantum chemistry package [5]. SCF convergence was set to tight and integration grid levels of 4 and 5 were employed during SCF iterations and the final computation of properties, respectively. Computations were accelerated using the RIJK approximation [6]. The reference data for aspirin was generated by selecting 20 000 random configurations from the MD17 database [7] and recomputing them at the above level of theory. Reference data for the ethanol molecule was taken from reference [8], which employed the same level of theory bar the RIJK approximation.

### B Training details

All models were trained using the Adam optimizer [9]. The learning rate is decayed by a factor of 0.5 if the validation loss plateaus, starting with the largest learning rate that does not diverge in steps 1e-4, 5e-4, 1e-3, etc. We apply

exponential smoothing with factor 0.9 to the validation loss to reduce the impact of fluctuations which are particularly common when training with both energies and forces. We use smaller batches for datasets where we train on both energies and forces, since it is commonly observed that larger batch sizes converge to larger errors in this setting.

Table 1: Training parameters for all experiments.

DATA SET	BATCH SIZE	LEARNING RATE	DECAY PATIENCE	STOPPING PATIENCE	$r_{\text{cut}}$ [Å]
QM9	100	$5 \cdot 10^{-4}$	5	30	5.0
MD17	10	$1 \cdot 10^{-3}$	50	150	5.0
FERROCENE	10	$1 \cdot 10^{-3}$	10	30	2.5-4.0
SPECTRA	10	$5 \cdot 10^{-4}$	15	50	2.7

## C Computation of infrared and Raman spectra

All simulations were carried out with the molecular dynamics module implemented in SchNetPack [10]. Classical molecular dynamics simulations for ethanol and aspirin were carried out for 50 ps at a temperature of 300 K controlled via Nose-Hoover chain [11] thermostat with a chain length of 3 and time constant of 100 fs. The first 10 ps of these trajectories were then discarded. Ring polymer molecular dynamics using 64 replicas were performed using the same simulation and equilibration periods and temperature, but employed a specially adapted global Nose-Hoover chain as introduced in Ref. [12] for thermostating instead. The overall chain settings were kept the same as above. In all cases, the velocity Verlet algorithm and a time step of 0.2 fs were used to integrate the equations of motion. Infrared and Raman spectra were computed from the time-autocorrelation functions of the dipole moment and polarizability time derivatives (see Ref. [13]). All autocorrelation functions were computed with the Wiener-Khinchin theorem [14], using an autocorrelation depth of 2048 fs. A Hann window function [15] and zero-padding were applied to the autocorrelation functions in order to enhance the spectra. Raman spectra were calculated using a laser frequency of 514 nm and temperature of 300 K.

## References

- [1] Bannwarth, C., Ehlert, S. & Grimme, S. GFN2-xTB – An accurate and broadly parametrized self-consistent tight-binding quantum chemical method with multipole electrostatics and density-dependent dispersion contributions. *J. Chem. Theory Comput.* **15**, 1652–1671 (2019).
- [2] Smith, J. S., Isayev, O. & Roitberg, A. E. ANI-1: an extensible neural network potential with DFT accuracy at force field computational cost. *Chem. Sci.* **8**, 3192–3203 (2017).

- [3] Adamo, C. & Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **110**, 6158–6170 (1999).
- [4] Weigend, F. & Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **7**, 3297–3305 (2005).
- [5] Neese, F. The ORCA program system. *WIREs Comput. Mol. Sci.* **2**, 73–78 (2012).
- [6] Weigend, F. A fully direct RI-HF algorithm: Implementation, optimised auxiliary basis sets, demonstration of accuracy and efficiency. *Phys. Chem. Chem. Phys.* **4**, 4285–4291 (2002).
- [7] Chmiela, S. *et al.* Machine Learning of Accurate Energy-Conserving Molecular Force Fields. *Sci. Adv.* **3**, e1603015 (2017).
- [8] Gastegger, M., Schütt, K. T. & Müller, K.-R. Machine learning of solvent effects on molecular spectra and reactions. *arXiv preprint arXiv:2010.14942* (2020).
- [9] Kingma, D. P. & Ba, J. Adam: A method for Stochastic Optimization. *arXiv preprint arXiv:1412.6980* (2014).
- [10] Schütt, K. T., Sauceda, H. E., Kindermans, P.-J., Tkatchenko, A. & Müller, K.-R. SchNet—A deep learning architecture for molecules and materials. *J. Chem. Phys.* **148**, 241722 (2018).
- [11] Martyna, G. J., Klein, M. L. & Tuckerman, M. Nosé–Hoover chains: The canonical ensemble via continuous dynamics. *J. Chem. Phys.* **97**, 2635–2643 (1992).
- [12] Ceriotti, M., Parrinello, M., Markland, T. E. & Manolopoulos, D. E. Efficient stochastic thermostating of path integral molecular dynamics. *J. Chem. Phys.* **133**, 124104 (2010).
- [13] Thomas, M., Brehm, M., Fligg, R., Vöhringer, P. & Kirchner, B. Computing vibrational spectra from ab initio molecular dynamics. *Phys. Chem. Chem. Phys.* **15**, 6608–6622 (2013).
- [14] Wiener, N. Generalized harmonic analysis. *Acta Math.* **55**, 117–258 (1930).
- [15] Blackman, R. B. & Tukey, J. W. The measurement of power spectra from the point of view of communications engineering – Part I. *Bell Syst. Tech. J.* **37**, 185–282 (1958).