

---

# Integration-free kernels for equivariant Gaussian fields with application in dipole moment prediction

---

**Tim Steinert**  
University of Bern  
tim.steinert@unibe.ch

**David Ginsbourger**  
University of Bern  
david.ginsbourger@unibe.ch

**August Lykke-Møller**  
University of Aarhus  
alm@chem.au.dk

**Ove Christiansen**  
University of Aarhus  
ove@chem.au.dk

**Henry Moss**  
University of Cambridge  
& Lancaster University  
hm493@cam.ac.uk

## Abstract

We develop a Gaussian Process model for accurate prediction of the dipole moments of water molecules by incorporating their equivariance under rotations. While kernels guaranteeing such equivariances have been investigated in previous work, their evaluation is often computationally prohibitive due to required integrations over the involved groups. In this work, we propose an alternative integration-free construction for equivariant kernels, relying on fundamental domain ideas previously explored in the scalar-valued invariant case, establishing a data-efficient and computationally lightweight GP model for dipole moments.

## 1 Motivation

The incorporation of structural knowledge such as physical laws into statistical machine learning models has gained significant attention for improving predictive accuracy and realism. In Gaussian process (GP) modelling, various invariances and equivariances can be encoded through tailored covariance kernels [Ginsbourger et al., 2012, Scheuerer and Schlather, 2012, Ginsbourger et al., 2016, van der Wilk et al., 2018, Holderrieth et al., 2021, Henderson, 2023]. GPs offer closed-form posterior distributions that facilitate uncertainty quantification and active learning for scientific tasks [Ranković et al., 2024]. However, constructing covariance functions that account for equivariance is computationally challenging, e.g. evaluating equivariant matrix-valued kernels [Reisert and Burkhardt, 2007] may require nested integrations.

Across molecular chemistry, invariances and equivariances are ubiquitous — applying simultaneous rigid motions on underlying atoms often results in unchanged scalar properties and in vectorial properties also invariant or affected by analogous motions (i.e. equivariance). While quantum chemical methods typically guarantee exact invariance and equivariance due to their theoretical foundations, ensuring these properties in computationally efficient models is less straightforward. Accordingly, we study how GP models trained on such properties can produce invariant and equivariant predictions for improved accuracy and efficiency.

The electric dipole moment, a vector indicating the imbalance in a molecule’s electron distribution, is key to understanding intermolecular interactions [Israelachvili, 2011, Stone, 2013] and predicting IR spectra intensities [Califano, 1976]. Estimating the dipole moment across a molecular surface is computationally intensive, often requiring recalculations for numerous configurations, which can take days even for small molecules. Therefore, accurate statistical models are crucial, as they reduce

the number of required calculations while still effectively describing the dipole surface, thereby making it feasible to predict IR spectra and manage computational costs for larger molecules.

Our main contributions are: (1) a novel class of computationally efficient integration-free equivariant kernels (2) a demonstration of this kernel class over SO(2)- and SO(3)- equivariance, and (3) an empirical validation on predicting dipole moments in quantum chemistry.

## 2 Background

**Preliminaries** For a detailed summary of the necessary background on the interplay between matrix-valued kernels and group theory, we refer the reader to Reisert and Burkhardt [2007]. In this work, we interchangeably denote multivariate Gaussian random fields as Gaussian Processes (GP).

**Dipole moments** The electric dipole moment of a molecule is here modelled as a vector function  $\boldsymbol{\mu} : \boldsymbol{x} \in D \subset \mathbb{R}^{3 \times s} \rightarrow \boldsymbol{\mu}(\boldsymbol{x}) \in \mathbb{R}^3$ , where  $\boldsymbol{x} = [\boldsymbol{x}_1, \dots, \boldsymbol{x}_s]$  encodes the position vectors in Euclidean space of the  $s$  atoms ( $\boldsymbol{x}_i, i = 1, \dots, s$ ) within the considered molecule. From physical principles,  $\boldsymbol{\mu}$  is known to be translation-invariant and rotation-equivariant, i.e. for all  $\boldsymbol{x} \in D$ ,

$$\begin{cases} \boldsymbol{\mu}(t \star_1 \boldsymbol{x}) = \boldsymbol{\mu}(\boldsymbol{x}) & \forall t \in \mathbb{R}^3, \\ \boldsymbol{\mu}(g \star_2 \boldsymbol{x}) = \rho_g \boldsymbol{\mu}(\boldsymbol{x}) & \forall g \in \text{SO}(3), \end{cases} \quad (1)$$

where  $\star_1$  denotes the action of translations (encoded by elements of  $\mathbb{R}^3$ ) on  $\mathbb{R}^3$ ,  $\star_2$  denotes the usual action of SO(3) on  $\mathbb{R}^3$ , and  $\rho_g \in \mathbb{R}^{3 \times 3}$  is the rotation matrix (representation) canonically associated with  $g \in \text{SO}(3)$ . The actions are extended to  $D$  with the conventions:

$$t \star_1 \boldsymbol{x} = [t \star_1 \boldsymbol{x}_1, \dots, t \star_1 \boldsymbol{x}_s], \quad g \star_2 \boldsymbol{x} = [g \star_2 \boldsymbol{x}_1, \dots, g \star_2 \boldsymbol{x}_s].$$

In the considered case of water molecules where  $s = 3$  and  $\boldsymbol{x}_2, \boldsymbol{x}_3$  both stand for hydrogen atoms, there is also permutation-invariance with respect to these two columns. Specifically, for all  $\boldsymbol{x} \in D$ ,

$$\boldsymbol{\mu}(\eta \star_3 \boldsymbol{x}) = \boldsymbol{\mu}(\boldsymbol{x}) \quad \eta \in \mathbb{Z}/2\mathbb{Z}, \quad (2)$$

where  $\star_3$  stands for the action of  $\mathbb{Z}/2\mathbb{Z}$  that swaps (for  $\eta = \bar{1}$ ) the second and third atoms.

## 3 Equivariant random field modeling

Due to the spatial dependence of dipole moments on the molecular configuration  $\boldsymbol{x}$ , we model their distribution as a  $\mathbb{R}^3$ -valued centred Gaussian random field  $Z = (Z_{\boldsymbol{x}})_{\boldsymbol{x} \in D}$ , where  $D \subset \mathbb{R}^{3 \times s}$  represents the configuration space, and  $Z_{\boldsymbol{x}}$  denotes the random dipole moment vector at configuration  $\boldsymbol{x}$ .  $Z$  is characterized by a covariance kernel matrix function  $K : D \times D \rightarrow \mathbb{R}^{3 \times 3}$ , which encodes the covariance structure of the dipole moments across different configurations.

**Equivariant Kernels** To encode our structural knowledge,  $Z$  must satisfy the invariance and equivariance conditions given in (1) and (2). Specifically, we can establish that, for a group  $G$  represented on  $\mathbb{R}^{p \times p}$  by  $\rho : G \rightarrow \mathbb{R}^{p \times p}$  and some action  $\star$  of  $G$  on  $D$ , second-order centred  $\mathbb{R}^p$ -valued random fields have equivariant paths (up to a modification), if their matrix-valued kernel  $K : D \times D \rightarrow \mathbb{R}^{p \times p}$ ,  $D \subset \mathbb{R}^d$  satisfies the following condition [Reisert and Burkhardt, 2007]:

$$\forall \boldsymbol{x}, \boldsymbol{x}' \in D, g, h \in G, \quad K(g \star \boldsymbol{x}, h \star \boldsymbol{x}') = \rho_g K(\boldsymbol{x}, \boldsymbol{x}') \rho_h^T. \quad (3)$$

For a linear, compact and unimodular group  $G$ , where the representation  $g \mapsto \rho_g$  is continuous if  $G$  is continuous, such an equivariant matrix-valued kernel function can be constructed by Haar integration of a base kernel matrix function  $K_o$  over  $G$ :

$$K_f(\boldsymbol{x}, \boldsymbol{x}') = \int_{G^2} \rho_g^T K_o(g \star \boldsymbol{x}, h \star \boldsymbol{x}') \rho_h dg dh. \quad (4)$$

However, the absence of an analytical form for the integrations in (4) necessitates expensive quadrature methods, which have so far made the incorporation of equivariances in GP models computationally infeasible and unstable. This challenge was previously addressed by Glielmo et al. [2017], who

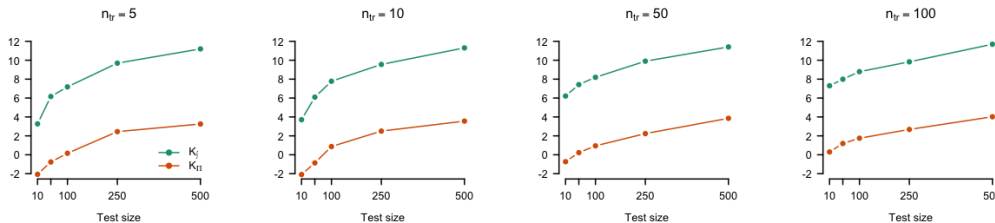


Figure 1: Comparison of the log computational time for the equivariant kernels.

aimed to model interatomic force fields using equivariant multivariate GP regression. Restricting to atoms of a single chemical species allowed them to represent molecule configurations as a scalar-valued linear combination of Gaussian functions (via kernel embedding), yielding analytical forms for the resulting kernels between configurations. We are not aware of extensions of this elegant approach to systems involving several chemical species.

**Integration Free Equivariant Kernels** We now introduce a new class of equivariant matrix-valued kernels that overcomes the problems of computational inefficiency and instability with existing equivariant kernels by combining a fundamental domain approach with a base matrix-valued kernel. The following proposition formalises the construction of this efficient kernel, and we then provide a first example for the case  $G = \text{SO}(2)$  demonstrating its significant computational advantages.

**Proposition 3.1.** *Let  $G$  be a linear group acting on  $D$  via  $\star$  and possessing a unitary group representation  $\rho : G \rightarrow \mathbb{R}^{p \times p}$ . Let  $A \subset D$  be a fundamental domain of  $\star$  and  $\Pi : D \rightarrow A$  be mapping any  $\mathbf{x} \in D$  to its representer in  $A$ . Let furthermore  $\tilde{D} = \{\mathbf{x} \in D : |\text{Stab}_\star(\mathbf{x})| = 1\}$  and  $s_\Pi : D \rightarrow G$  be a map such that for any  $\mathbf{x} \in D$ ,  $\Pi(\mathbf{x}) = s_\Pi(\mathbf{x}) \star \mathbf{x}$ . Then, for any matrix-valued kernel  $K_A$  on  $A$ ,  $K_\Pi$  below defines an equivariant (w.r.t  $\star$  and  $\rho$ ) matrix-valued kernel on  $\tilde{D}$ :*

$$K_\Pi(\mathbf{x}, \mathbf{x}') = \rho_{s_\Pi(\mathbf{x})}^T K_A(\Pi(\mathbf{x}), \Pi(\mathbf{x}')) \rho_{s_\Pi(\mathbf{x}')}. \quad (5)$$

**Remark 3.2.** *While there are several ways to define  $s_\Pi$  on points possessing a stabilizer with cardinality  $\geq 2$ ,  $s_\Pi(\mathbf{x})$  is uniquely defined on  $\tilde{D}$ . We refer to any such  $s_\Pi$  as a section.*

**Example 3.3.** *Assume  $D = \mathbb{R}^2$ ,  $p = 2$ , and  $G = \text{SO}(2)$  (with  $g \star \mathbf{x}$  being plainly defined as  $\mathbf{x}$  rotated by  $g$ ). A fundamental domain is then given by  $A = \{(x, 0), x \geq 0\}$  and its projector is defined by  $\Pi(\mathbf{x}) = (r(\mathbf{x}), 0)$ , where  $r(\mathbf{x}) = \|\mathbf{x}\|_2$ . Here, for  $\mathbf{x} \neq (0, 0)$ ,  $s_\Pi(\mathbf{x})$  is represented by the unique rotation matrix sending  $\mathbf{x}$  to  $\Pi(\mathbf{x})$ . For  $\mathbf{x} = (0, 0)$ , any rotation matrix would do; we conventionally set  $s_\Pi((0, 0))$  equal to the identity matrix. In Figure (1) we compare the total computational times of the covariance matrices of the posterior distribution of the GPs  $\text{GP}(0, K_\Pi)$  and  $\text{GP}(0, K_f)$  on  $[-1, 1]^2$ , for different training and test set sizes. Here,  $K_f$  is computed by adaptive integration with the `adaptIntegrate` function in  $\mathbf{x}$  on a maximum of 1000 function evaluations. The base kernel  $K_o$  is chosen to be the simple diagonal RBF kernel matrix function  $K_o = \text{diag}\left(\exp - \frac{\|\mathbf{x} - \mathbf{x}'\|_2^2}{2}, i \in \{1, 2\}\right)$ . The time difference for moderate training and test sizes of 100 and 500 are substantial; while the posterior distribution of  $\text{GP}(0, K_\Pi)$  requires a total of 55 seconds to compute,  $\text{GP}(0, K_f)$  requires 33 hours.*

## 4 Water molecule dipole moment prediction

We now provide an explicit example of how to apply our Eq. (5) to a quantum chemistry problem. In particular, we consider the case of water molecules, where  $\mathbf{x} \in D \subset \mathbb{R}^{3 \times 3}$  now represents the position vectors of an oxygen and two hydrogen atoms. As baseline kernel, we take diagonal squared exponential matrix-valued kernels:

$$K(\cdot; \boldsymbol{\theta}) : D \times D \rightarrow \mathbb{R}^{p \times p}, \quad (\mathbf{x}, \mathbf{x}') \mapsto \text{diag}\left(\sigma_i^2 \exp - \frac{\|\mathbf{x} - \mathbf{x}'\|_2^2}{2\ell_i^2}, i \in \{1, 2, 3\}\right),$$

where  $\boldsymbol{\theta} = (\ell_1, \sigma_1^2, \dots, \ell_3, \sigma_3^2)$  denotes the vector containing the tunable kernel lengthscale and variance hyperparameters. For ease of notation we write  $K = K(\cdot; \boldsymbol{\theta})$ .

**Constructing a tailored matrix-valued kernel** To construct matrix-valued kernels that lead to computationally efficient GP models honoring invariances and equivariances from Equations (1) and (2), we construct an integration-free rotation-equivariant (as well as translation and hydrogen swap invariant) kernel matrix function  $K_{\Pi}$ . Our construction requires a fundamental domain endowed with a matrix-valued kernel  $K_A$ , as well as a suitably defined projection mapping  $\Pi: D \rightarrow A$ , also expressible via an associated section  $s_{\Pi}: D \rightarrow \text{SO}(3)$  (such that for  $\mathbf{x} \in D$ ,  $\Pi(\mathbf{x}) = s_{\Pi}(\mathbf{x}) \star_2 \mathbf{x}$ ).

For the translation-invariance of (1), we arrange the configurations relative to the position of the oxygen atom and consider only  $\bar{\mathbf{x}} = (-\mathbf{x}_1) \star_1 \mathbf{x} = \mathbf{x} - \mathbf{x}_1$ . To honor (2), we ensure (by swapping columns when necessary) that  $\bar{r}_1 = \|\bar{\mathbf{x}}_1\|_2 \geq \|\bar{\mathbf{x}}_2\|_2 = \bar{r}_2$ . We then build a rotation matrix  $\mathcal{R}(\mathbf{x}) \in \mathbb{R}^{3 \times 3}$  based on three elementary rotations, such that for some  $a, b \in \mathbb{R}$ ,  $\mathcal{R}(\mathbf{x})\bar{\mathbf{x}}_1 = (0, \bar{r}_1, 0)$  and  $\mathcal{R}(\mathbf{x})\bar{\mathbf{x}}_2 = (a, b, 0)$ . We eventually obtain  $K_{\Pi}$  by taking

$$K_{\Pi}(\mathbf{x}, \mathbf{x}') = \mathcal{R}(\mathbf{x})^{\top} K_A(\Pi(\mathbf{x}), \Pi(\mathbf{x}')) \mathcal{R}(\mathbf{x}'),$$

where  $\Pi(\mathbf{x})$  is parametrised by  $(\bar{r}_1, a, b)$  only. Notice that  $\Pi$  is invariant under  $\star_1, \star_2, \star_3$ . Let us remark that choices of  $A$ ,  $\Pi$  and  $s_{\Pi}$  are not unique. Here,  $K_A$  is chosen of the same form as  $K(\cdot; \theta)$  (yet with inputs in  $\mathbb{R}^3$ ).

**Experimental Results** We utilise a dataset of dipole moments  $\boldsymbol{\mu}$  obtained from 850 water molecule configurations  $\mathbf{x}$ , which were computed using quantum chemical methods. For details on the dataset generation process, see Appendix B. To enhance the dataset’s diversity and representativeness, we applied random rotations to the data points. See Appendix B for a visualisation of this augmented dataset. We compare the predictive accuracy of GP models for the dipole moments of water molecules using the baseline kernel  $K$  and the proposed kernel  $K_{\Pi}$ .

For additional comparison, we include invariances separately with the following kernels:

1. **Translation-invariance:**

$$K_2(\mathbf{x}, \mathbf{x}') = K(\bar{\mathbf{x}}, \bar{\mathbf{x}}')$$

2. **Permutation-invariance:**

$$K_4(\mathbf{x}, \mathbf{x}') = K(\Pi_{\star_3}(\mathbf{x}), \Pi_{\star_3}(\mathbf{x}')),$$

$$\Pi_{\star_3}(\mathbf{x}) = \begin{cases} [\mathbf{x}_1, \mathbf{x}_3, \mathbf{x}_2], & \text{if } \bar{r}_1 < \bar{r}_2, \\ \mathbf{x}, & \text{otherwise} \end{cases}$$

3. **Permutation and translation-invariance:**

$$K_3(\mathbf{x}, \mathbf{x}') = K(\Pi_{\star_{1,3}}(\mathbf{x}), \Pi_{\star_{1,3}}(\mathbf{x}')),$$

$$\Pi_{\star_{1,3}}(\mathbf{x}) = (-\mathbf{x}_1) \star_1 \Pi_{\star_3}(\mathbf{x}).$$

We train the GPs using Adam optimisation over 1000 iterations, with all kernel hyperparameters initialised to 1, maximising the training likelihood. In Figure (2), we see that the proposed integration-free equivariant GP consistently outperforms the base GP across all training set sizes, the small order of magnitudes of their RMSE suggest that our proposed model is accurate enough for use in quantum chemistry.

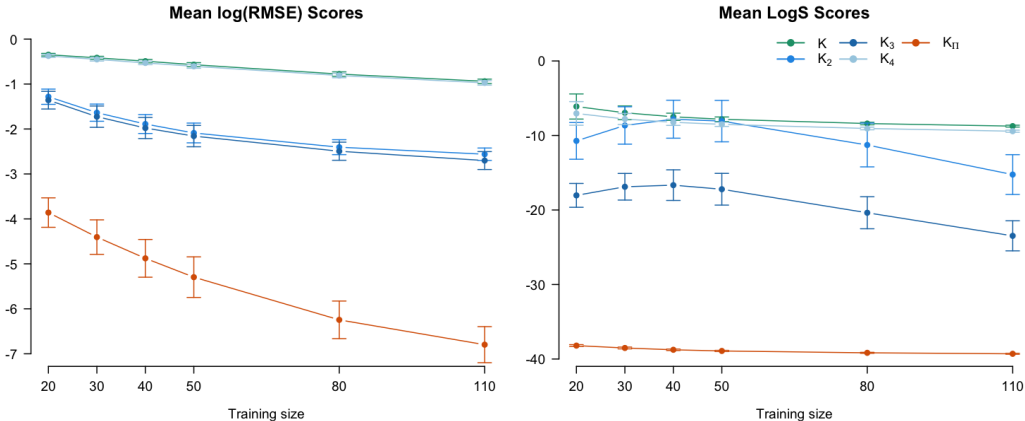


Figure 2: Predictive scores (mean  $\pm$  sd over  $10^3$  reps) of GP models versus training set size.

## 5 Conclusion and perspectives

We have introduced an integration-free method to construct a class of equivariant GP models. By leveraging fundamental domains, we eliminate the need for cumbersome integration over groups, significantly reducing the computational burden whilst leveraging prior knowledge of invariances and equivariances, enabling accurate predictions in small training data regimes. Our approach significantly enhances the modelling of dipole moments of water molecules by incorporating physical principles directly into the Gaussian process framework. The obtained numerical results are in fact very promising and open avenues of research on the short and longer term. Ongoing work include testing different scenarios on  $K_A$  and its hyperparameters, looking into equivariances with respect to reflections, comparing  $K_{\Pi}$  to  $K_f$  in feasible settings in terms of both accuracy and computational cost, and also analysing more precisely how the benefits are split among the account of the different invariances and equivariances. Investigating potential synergies with recent developments pertaining to kernels on graphs, Lie groups, and other structures is also of interest (See, e.g., Azangulov et al. [2024]). Future work include implementing resulting kernels into open source software toward GP prediction of molecular properties and active learning (see for instance [Moss and Griffiths, 2020, Griffiths et al., 2024]) and extend our approach to other molecular systems and tensor properties like polarisability volume, broadening the applicability of equivariant GP modeling in scientific research. Furthermore, future work may include comparisons between the fundamental domain approach and extensions of Glielmo et al. [2017] in dipole moment estimation and potentially in further contexts. On the theoretical side, further studying the proposed class of kernels and their properties calls for in-depth investigations on fundamental domains (potentially calling for variations in associated definitions and assumptions).

**Acknowledgements:** Tim Steinert and David Ginsbourger acknowledge the support of the Digitization Commission (DigiK) of the University of Bern Perception via the project “Statistics, Econometrics and Probability”. The authors would also like to thank the reviewers and all persons having helped improving the paper.

## References

- Turbomole v7.8. A development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007, 2023.
- Iskander Azangulov, Andrei Smolensky, Alexander Terenin, and Viacheslav Borovitskiy. Stationary kernels and gaussian processes on lie groups and their homogeneous spaces ii: non-compact symmetric spaces. *Journal of Machine Learning Research*, 25(281):1–51, 2024.
- Keld L Bak, Jürgen Gauss, Trygve Helgaker, Poul Jørgensen, and Jeppe Olsen. The accuracy of molecular dipole moments in standard electronic structure calculations. *Chemical Physics Letters*, 319(5):563–568, 2000.
- Sree Ganesh Balasubramani, Guo P. Chen, Sonia Coriani, Michael Diedenhofen, Marius S. Frank, Yannick J. Franzke, Filipp Furche, Robin Grotjahn, Michael E. Harding, Christof Hättig, Arnim Hellweg, Benjamin Helmich-Paris, Christof Holzer, Uwe Huniar, Martin Kaupp, Alireza Marefat Khah, Sarah Karbalaei Khani, Thomas Müller, Fabian Mack, Brian D. Nguyen, Shane M. Parker, Eva Perlt, Dmitriy Rappoport, Kevin Reiter, Saswata Roy, Matthias Rückert, Gunnar Schmitz, Marek Sierka, Enrico Tapavicza, David P. Tew, Christoph van Wüllen, Vamsee K. Voora, Florian Weigend, Artur Wodyński, and Jason M. Yu. Turbomole: Modular program suite for ab initio quantum-chemical and condensed-matter simulations. *Journal of Chemical Physics*, 152(18):184107, 2020.
- Renato Berlinghieri, Brian L. Trippe, David R. Burt, Ryan Giordano, Kaushik Srinivasan, Tamay Özgökmen, Junfei Xia, and Tamara Broderick. Gaussian processes at the helm(holtz): a more fluid model for ocean currents. In *Proceedings of the 40th International Conference on Machine Learning*. JMLR.org, 2023.
- S. Califano. *Vibrational States*. Wiley, 1976.

- Ove Christiansen, Denis G. Artiukhin, Frederik D. Bader, Ian H. Godtlielsen, Eduard M. Gras, Werner Györfy, Mikkel B. Hansen, Mads B. Hansen, Mads G. Højlund, Nicolai M. Høyer, Andreas B. Jensen, Emil L. Klinting, Jacob Kongsted, Carolin König, Diana Madsen, Niels K. Madsen, Kasper Monrad, Gunnar Schmitz, Peter Seidler, Kristian Sneskov, Manuel Sparta, Bo Thomsen, Daniele Toffoli, and Alberto Zoccante. Midascpp, version 2024.04.0, 2024.
- Thom H. Jr. Dunning. Gaussian basis sets for use in correlated molecular calculations. i. the atoms boron through neon and hydrogen. *The Journal of Chemical Physics*, 90(2):1007–1023, 1989.
- David Ginsbourger, Xavier Bay, Olivier Roustant, and Laurent Carraro. Argumentwise invariant kernels for the approximation of invariant functions. *Annales de la Faculté des sciences de Toulouse : Mathématiques*, Ser. 6, 21(3):501–527, 2012.
- David Ginsbourger, Olivier Roustant, and Nicolas Durrande. On degeneracy and invariances of random fields paths with applications in gaussian process modelling. *Journal of Statistical Planning and Inference*, 170:117–128, 2016.
- Aldo Glielmo, Peter Sollich, and Alessandro De Vita. Accurate interatomic force fields via machine learning with covariant kernels. *Phys. Rev. B*, 95:214302, Jun 2017.
- Ryan-Rhys Griffiths, Leo Klarner, Henry Moss, Aditya Ravuri, Sang Truong, Yuanqi Du, Samuel Stanton, Gary Tom, Bojana Rankovic, Arian Jamasb, et al. Gauche: a library for gaussian processes in chemistry. *Advances in Neural Information Processing Systems*, 36, 2024.
- Iain Henderson. *PDE constrained kernel regression methods*. PhD thesis, INSA Toulouse; Université de Toulouse-Toulouse III-UPS, 2023.
- Peter Holderrieth, Michael J Hutchinson, and Yee Whye Teh. Equivariant learning of stochastic fields: Gaussian processes and steerable conditional neural processes. In *International Conference on Machine Learning*, pages 4297–4307. PMLR, 2021.
- Jacob N. Israelachvili, editor. *Intermolecular and Surface Forces*. Academic Press, San Diego, third edition edition, 2011.
- Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization, 2017. URL <https://arxiv.org/abs/1412.6980>.
- Reiner Lenz. *Group theoretical methods in image processing*, volume 413. Springer, 1990.
- Jonathan M. Lilly and Paula Pérez-Brunius. GulfDrifters: A consolidated surface drifter dataset for the Gulf of Mexico (1.1.0), 2021. URL <https://doi.org/10.5281/zenodo.4421585>.
- Henry B Moss and Ryan-Rhys Griffiths. Gaussian process molecule property prediction with flowmo. *arXiv preprint arXiv:2010.01118*, 2020.
- Chr. Møller and M. S. Plesset. Note on an approximation treatment for many-electron systems. *Physical Review*, 46:618–622, 1934.
- Bojana Ranković, Ryan-Rhys Griffiths, Henry B Moss, and Philippe Schwaller. Bayesian optimisation for additive screening and yield improvements—beyond one-hot encoding. *Digital Discovery*, 3(4):654–666, 2024.
- Marco Reiser and Hans Burkhardt. Learning equivariant functions with matrix valued kernels. *Journal of Machine Learning Research*, 8(15):385–408, 2007.
- Michael Scheuerer and Martin Schlather. Covariance models for divergence-free and curl-free random vector fields. *Stochastic Models*, 28:433–451, 07 2012.
- Anthony Stone. *The Theory of Intermolecular Forces*. Oxford University Press, 01 2013. ISBN 9780199672394.
- D. Toffoli, J. Kongsted, and O. Christiansen. Automatic generation of potential energy and property surfaces of polyatomic molecules in normal coordinates. *The Journal of Chemical Physics*, 127(20):204106, 11 2007.

Mark van der Wilk, Matthias Bauer, ST John, and James Hensman. Learning invariances using the marginal likelihood. *Advances in Neural Information Processing Systems*, 31, 2018.

Florian Weigend and Marco Häser. RI-MP2: first derivatives and global consistency. *Theoretical Chemistry Accounts*, 97(1):331–340, 1997.

## A Group and representation theory

In this section we briefly describe the preliminaries on groups and group representations which are necessary to construct equivariant kernel matrix functions.

**Definition A.1.** A group  $(G, \circ)$  is a set  $G$  equipped with a binary operation

$$\circ : G \times G \rightarrow G, \quad (a, b) \mapsto a \circ b$$

which satisfies

1.  $\forall a, b, c \in G, a \circ (b \circ c) = (a \circ b) \circ c$
2.  $\exists e \in G$  s.t.  $\forall a \in G, a \circ e = e \circ a = a$
3.  $\forall a \in G, \exists a^{-1} \in G$  s.t.  $a \circ a^{-1} = a^{-1} \circ a = e$ .

**Definition A.2.** If the set  $G$  is furthermore a topological space and the group operation  $\circ$  as well as its inverse  $\circ^{-1}$  are continuous with respect to the topology of  $G$ , we call  $G$  a topological group. In addition, a topological group  $G$  is compact if it is compact with respect to its topology, i.e. if each open cover of  $G$  admits a finite open subcover.

**Definition A.3.** A group representation  $\rho$  maps elements from a group  $G$  to  $\mathcal{L}(V)$  where  $V$  is a finite dimensional Hilbert space and  $\mathcal{L}(V)$  the space of linear transformations on  $V$ . Furthermore  $\rho$  is a homomorphism, i.e. for any  $g_1, g_2 \in G, \rho(g_1 \circ g_2) = \rho(g_1)\rho(g_2)$ .

**Definition A.4.** A group  $G$  is called a linear group if there exists an injective homomorphism  $\phi : G \rightarrow GL(n, K)$  to the general linear group  $GL(n, K)$  for some integer  $n$  and some field  $K$ , such that the image  $\phi(G)$  is a closed set in the natural topology on  $GL(n, K)$ , which corresponds to the topology induced by the standard norm on  $K^n$ .

As consequence, a linear group admits invertible matrix-valued representations.

**Definition A.5.** A linear group representation is called unitary if for any  $g \in G$  it holds  $\rho_{g^{-1}} = \rho_g^\dagger$ .

Following Reisert and Burkhardt [2007], it can be shown that any representation of a finite group or of a compact continuous group with continuous representation is equivalent to a unitary representation.

**Definition A.6.** A measure  $\mu$  defined on the  $\sigma$ -algebra generated by the open sets of a compact group  $G$ , called the Borel algebra of  $G$ , is called left translation-invariant if for any open subset  $S \subset G$  and  $g \in G$  it holds  $\mu(gS) = \mu(S)$ , where  $gS = \{g \circ s \mid s \in S\}$ .

By Haar's theorem, on any compact group there exists a unique left translation-invariant measure, called the left Haar measure. Analogously, there exists a unique right translation-invariant measure, called the right Haar measure.

**Definition A.7.** A compact group is said to be unimodular, if its right and left Haar measure coincide.

It can be shown that representations of compact, linear, unimodular groups have determinant one, which simplifies reparametrisations in Haar integrals  $\int_G f(g) dg$ .

**Definition A.8.** A left group action of a group  $G$  on a set  $X$  is a function

$$\begin{aligned} \Phi : G \times X &\rightarrow X \\ (g, x) &\mapsto g.x \end{aligned}$$

satisfying for any  $x \in X, g, h \in G$  :

1.  $e.x = x$

$$2. g.(h.x) = (g \circ h).x$$

If  $X$  is a finite dimensional vector space and  $G$  is compact, linear and has a continuous representation  $\rho$  of the same dimension as  $X$ , we can express the action of  $G$  on  $X$  as left multiplication, i.e.  $g.x = \rho_g x$ .

### Fundamental domains

**Definition A.9.** The orbit of a point  $x \in X$  under the action  $\Phi$  is the set

$$\mathcal{O}_\Phi(x) = \{g.x | g \in G\}.$$

A point  $x \in X$  is said to be a fixed point of  $\Phi$  if  $g.x = x$  for any  $g \in G$ . For any  $x \in X$ , we further define the stabilizer (or stabilizer group) of  $x$  under  $\Phi$  to be  $\text{Stab}_\Phi(x) = \{g \in G : g.x = x\}$ .

**Definition A.10.** A subset  $A \subset X$  is said to be a fundamental domain of  $\Phi$ , if it is a system of representatives of the orbits of  $\Phi$ .

**Remark A.11.** With the latter fundamental domain definition,  $A$  is a subset of  $X$  that contains exactly one representative from each orbit under the action of  $G$ . That is, for any two distinct elements  $x, y \in A$ , there is no group element  $g \in G$  such that  $g.x = y$ . Furthermore it holds that  $\bigcup_{g \in G} g.A = X$ . For each  $x \in X$  there exists a unique representative in the fundamental domain, denoted here  $\Pi(x) \in \mathcal{O}_\Phi(x) \cap A$ . We refer to  $\Pi(x)$  as the representer of  $x$  in  $A$ .

### A.1 Proof of Proposition (3.1)

First, it directly follows that  $K_\Pi$  is positive definite as for any two sets of  $n$  vectors  $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{R}^p$  and  $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{R}^d$ ,

$$\begin{aligned} \sum_{1 \leq i, j \leq n} \mathbf{a}_i^T K_\Pi(\mathbf{x}_i, \mathbf{x}_j) \mathbf{a}_j &= \sum_{1 \leq i, j \leq n} \mathbf{a}_i^T \rho_{s_\Pi(\mathbf{x}_i)}^T K_A(\Pi(\mathbf{x}_i), \Pi(\mathbf{x}_j)) \rho_{s_\Pi(\mathbf{x}_j)} \mathbf{a}_j \\ &= \sum_{1 \leq i, j \leq n} \mathbf{b}_i^T K_A(\Pi(\mathbf{x}_i), \Pi(\mathbf{x}_j)) \mathbf{b}_j \\ &\geq 0, \end{aligned}$$

where  $\mathbf{b}_i := \rho_{s_\Pi(\mathbf{x}_i)} \mathbf{a}_i$  and the last inequality follows from positive definiteness of  $K_A$ . Since the projector  $\Pi$  is constant on the orbits it holds

$$\begin{aligned} K_\Pi(g \star \mathbf{x}, h \star \mathbf{x}') &= \rho_{s_\Pi(g \star \mathbf{x})}^T K_A(\Pi(g \star \mathbf{x}), \Pi(h \star \mathbf{x}')) \rho_{s_\Pi(h \star \mathbf{x}')} \\ &= \rho_{s_\Pi(g \star \mathbf{x})}^T K_A(\Pi(\mathbf{x}), \Pi(\mathbf{x}')) \rho_{s_\Pi(h \star \mathbf{x}')} \end{aligned}$$

For  $\mathbf{x} \in \tilde{D}$ ,  $s_\Pi(g \star \mathbf{x}) = s_\Pi(\mathbf{x}) \circ g^{-1}$  and thus  $\rho_{s_\Pi(g \star \mathbf{x})} = \rho_{s_\Pi(\mathbf{x})} \rho_{g^{-1}}$ . Hence,

$$\begin{aligned} K_\Pi(g \star \mathbf{x}, h \star \mathbf{x}') &= \rho_g \rho_{s_\Pi(\mathbf{x})}^T K_A(\Pi(\mathbf{x}), \Pi(\mathbf{x}')) \rho_{s_\Pi(\mathbf{x}')} \rho_h^T \\ &= \rho_g K_A(\mathbf{x}, \mathbf{x}') \rho_h^T. \end{aligned}$$

□

## B Data generation

The dipole moments were generated by optimising an initial guess for the geometry using the RI-MP2 module in the TURBOMOLE quantum chemistry program. The theory for this module is described in the original MP2 paper by Møller and Plesset [1934] and the RI-MP2 paper by Weigend and Häser [1997]. Furthermore, a cc-pVDZ basis set was used, as created by Dunning [1989]. Using this optimised structure, a numerical hessian was found. diagonalising the mass weighted Hessian then gave a set of three Normal coordinates, corresponding to bending, symmetric and asymmetric bond stretch. Using the static grid method [Toffoli et al., 2007] in the MIDASCPP program package a grid of twenty points was constructed along each normal coordinate. These points were linearly spaced between the classical turning points of the tenth excited state of the quantum mechanical oscillator approximation which can be constructed from the mass weighted hessian. Furthermore,



three twenty by twenty grids were constructed describing displacements along two out of the three coordinates simultaneously. Combined with the optimised structure this gives 1261 geometries. For each of these geometries, the electric dipole moment was calculated using the RI-MP2 method and a cc-pVDZ basis set in TURBOMOLE. To ensure dataset efficiency, 410 pairs of points that were rotations of each other were identified and removed. The remaining 851 points were then randomly rotated to cover a larger region of Cartesian space. This approach samples the geometries of chemical interest around the optimised structure and then rotates them in space resulting in a diverse and representative dataset.

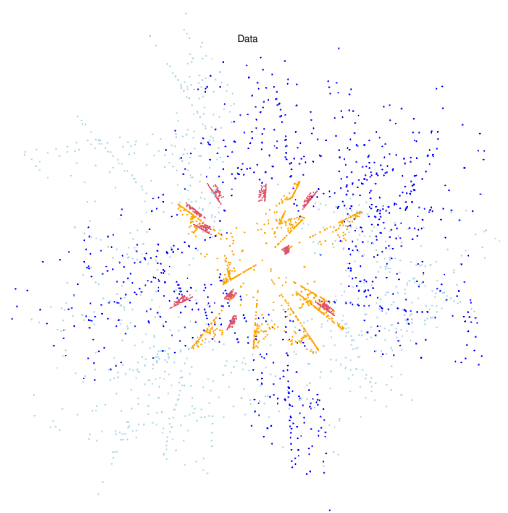


Figure 3: A realisation of the dataset: (light) blue indicate positions of the hydrogen atoms, red the position of the oxygen atoms and orange the dipole moments.