



**HAL**  
open science

## A simple approach to rotationally invariant machine learning of a vector quantity

Jakub Martinka, Marek Pederzoli, Mario Barbatti, Pavlo O Dral, Jiří Pittner

► **To cite this version:**

Jakub Martinka, Marek Pederzoli, Mario Barbatti, Pavlo O Dral, Jiří Pittner. A simple approach to rotationally invariant machine learning of a vector quantity. *The Journal of Chemical Physics*, 2024, 10.1063/5.0230176 . hal-04763148

**HAL Id: hal-04763148**

**<https://hal.science/hal-04763148v1>**

Submitted on 1 Nov 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A simple approach to rotationally invariant machine learning of a vector quantity

Jakub Martinka,<sup>1,2</sup> Marek Pederzoli,<sup>1</sup> Mario Barbatti,<sup>3,4</sup> Pavlo O. Dral,<sup>5,6</sup> and Jiří Pittner<sup>1</sup>

<sup>1</sup>*J. Heyrovský Institute of Physical Chemistry,  
Academy of Sciences of the Czech Republic, v.v.i.,  
Dolejškova 3, 18223 Prague 8, Czech Republic*

<sup>2</sup>*Department of Physical and Macromolecular Chemistry,  
Faculty of Sciences, Charles University,  
Hlavova 8, 128 43 Prague 2, Czech Republic*

<sup>3</sup>*Aix Marseille University, CNRS, ICR, Marseille, France*

<sup>4</sup>*Institut Universitaire de France, 75231 Paris, France*

<sup>5</sup>*State Key Laboratory of Physical Chemistry of Solid Surfaces,  
College of Chemistry and Chemical Engineering,  
and Fujian Provincial Key Laboratory of Theoretical and Computational Chemistry,  
Xiamen University, Xiamen, Fujian 361005, China*

<sup>6</sup>*Institute of Physics, Faculty of Physics, Astronomy,  
and Informatics, Nicolaus Copernicus University in Toruń,  
ul. Grudziądzka 5, 87-100 Toruń, Poland*

(\*jiri.pittner@jh-inst.cas.cz)

## Abstract

Unlike with the energy, which is a scalar property, machine learning (ML) predictions of vector or tensor properties poses the additional challenge of achieving proper invariance (covariance) with respect to molecular rotation. For the energy gradients needed in molecular dynamics (MD), this symmetry is automatically fulfilled when taking analytic derivative of the energy, which is a scalar invariant (using properly invariant molecular descriptors). However, if the properties cannot be obtained by differentiation, other appropriate methods should be applied to retain the covariance. There have been several approaches suggested to properly treat this issue. For nonadiabatic couplings and polarizabilities, for example, it was possible to construct virtual quantities from which the above tensorial properties are obtained by differentiation and thus guarantee the covariance. Another possible solution is to build the rotational equivariance into the design of a neural network employed in the model. Here we propose a simpler alternative technique, which does not require construction of auxiliary properties or application of special equivariant ML techniques. We suggest a three-step approach, using the molecular tensor of inertia. In the first step, the molecule is rotated using the eigenvectors of this tensor to its principal axes. In the second step, the ML procedure predicts the vector property relative to this orientation, based on a training set where all vector properties were in this same coordinate system. As third step, it remains to transform the ML estimate of the vector property back to the original orientation. This rotate-predict-rotate (RPR) procedure should thus guarantee proper covariance of a vector property and is trivially extensible also to tensors such as polarizability. The PRP procedure has an advantage that the accurate models can be trained very fast for thousands of molecular configurations which might be beneficial where many trainings are required (e.g., in active learning). We have implemented the RPR technique, using the MLatom and Newton-X programs for ML and MD and performed its assessment on the dipole moment along MD trajectories of 1,2-dichloroethane.

## I. INTRODUCTION

Machine learning (ML) proved to be a powerful tool in computational chemistry<sup>1</sup>. While ML prediction of scalar quantities, such as potential energy to propagate molecular dynamics (MD) simulations, is a well-established procedure<sup>2,3</sup>, predicting vectors (rank-1 tensors) or higher rank tensors challenge the ML techniques to deliver quantities covariant with respect to symmetry operations<sup>4</sup>. This problem affects various properties such as dipole moment, (hyper)polarizability or forces, which are crucial for spectrum (infrared<sup>5</sup>; Raman<sup>6</sup>), machine learning potentials (MLPs)<sup>7</sup> and MD simulations. Other vectorial properties possess yet an additional demand of having a correct sign. This is the case of transition dipole moments<sup>8-12</sup> and nonadiabatic coupling vectors<sup>11,13,14</sup>, which are calculated for two different electronic states of an arbitrary global phase.

There are several approaches in the literature to handle orientation dependence. For dipole moments, neural networks were suggested that learn atomic charges to reconstruct the dipole moments as accurately as possible<sup>15,16</sup>. For polarizabilities and nonadiabatic couplings, such neural networks were built that can learn the virtual quantities, the analytical derivatives of which produce the required tensorial properties<sup>13,16,17</sup>. Similar in spirit is an approach based on kernel methods which was used to learn dipole moments as response properties.<sup>18</sup> Above approaches are inspired by MLPs where it is common to derive energy gradients (negative forces) as the derivatives of the machine learning function of potential energies. Many approaches rely on building equivariance into the design of the ML model, either into the local descriptors employed in neural network (NN) approaches, or in the design of the NN itself. This was done in case of Gaussian process regression (GPR), by starting with a kernel based on a smooth overlap of atomic positions (SOAP)<sup>19</sup> and deriving a hierarchy of  $\lambda$ -SOAP kernels preserving rotational symmetry<sup>20-22</sup>. This so-called symmetry-adapted GPR (SA-GPR) approach was used to paracetamol system leading to a covariant predictions<sup>23</sup>. The same is possible for the class of ML algorithms based on NNs as shown for neuroevolution potential (NEP) scheme<sup>24</sup>, which was modified into tensorial-NEP approach<sup>25</sup> or embedded atom NN (EANN)<sup>26</sup> transformed into tensorial EANN,<sup>17</sup> outperforming SA-GPR. Other examples introduced new rotationally covariant architectures<sup>27</sup>, equivariant message passing NN<sup>28</sup> or tensor field NN<sup>29</sup>. The NN-based approaches proved to be excellent tools for predicting infrared and Raman spectra<sup>30,31</sup>, even under the influence

of external fields<sup>32</sup>.

While above approaches all shown to work well for their test applications, the practical problem is their often high computational cost for training and predicting. The solution to this problem is often employing the kernel methods and, indeed, the kernel ridge regression (KRR) with a global molecular descriptor has already been suggested for treatment of such tensorial properties as multipole moments<sup>33</sup> and atomic forces<sup>34</sup>. In such approaches, it is common to rotate the molecular geometry to some reference frame<sup>35–37</sup>. Although such a rotation is straightforward to implement, care should be taken for flexible or dissociative systems<sup>21</sup>. Here we built upon this rotational idea to propose a three-step procedure to maintain proper covariance in the context of learning the vectorial properties for the set of conformations visited during molecular dynamics. Our procedure is similar in spirit to the one employed by Hu and Huo<sup>37</sup> in a related context, however, we define the reference frame by principal axes of the molecular tensor of inertia rather than by an arbitrarily selected triple of atoms, which should remove a potential source of ambiguity. We demonstrate this new method (named rotate-predict-rotate for reason that will be clear in the next section) on 1,2-dichloroethane, for which new descriptors were introduced improving the overall test set root mean square error (RMSE) values for both dipole moment and polarizability components. The proposed method is not just accurate but also computationally very efficient, which we demonstrate by timings for different sizes of the learning data set.

## II. METHODS

In this study, we used kernel ridge regression (KRR) to fit individual components of vectorial/tensorial properties, which are predicted using

$$y(\mathbf{x}') = \sum_{i=1}^{N_{\text{tr}}} \alpha_i K(\mathbf{x}', \mathbf{x}_i), \quad (1)$$

where  $N_{\text{tr}}$  is a number of training points,  $\alpha_i$  is a regression coefficient and,  $K$  represents kernel function which can has Gaussian (2a), Laplace (2b) or Matérn (2c) form

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(\frac{-\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}\right), \quad (2a)$$

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(\frac{-\|\mathbf{x}_i - \mathbf{x}_j\|}{\sigma}\right), \quad (2b)$$

$$K(\mathbf{x}_i, \mathbf{x}_j) = \sum_{k=0}^n \frac{(n+k)!}{(2n)!} \left( \frac{2\|\mathbf{x}_i - \mathbf{x}_j\|}{\sigma} \right)^{n-k} \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|}{\sigma}\right). \quad (2c)$$

The geometrical configuration of a molecule is represented by a molecular descriptor  $\mathbf{x}$ . There is a plethora of descriptors, usually suitable for various tasks. Common examples of global descriptors suitable for use with KRR are relative to equilibrium (RE, (3a)), Coulomb matrix (CM, (3b)) or inverse distance (ID, (3c))

$$\mathbf{x}_{\text{RE}}^\top = \left( \frac{r_{12}^{\text{eq}}}{r_{12}}, \dots, \frac{r_{23}^{\text{eq}}}{r_{23}}, \dots, \frac{r_{(N_{\text{at}}-1)N_{\text{at}}}^{\text{eq}}}{r_{(N_{\text{at}}-1)N_{\text{at}}}} \right), \quad (3a)$$

$$\mathbf{x}_{\text{CM}}^\top = \left( 0.5Z_1^{2,4}, \frac{Z_1Z_2}{r_{12}}, \dots, \frac{Z_2Z_1}{r_{21}}, 0.5Z_2^{2,4}, \dots \right), \quad (3b)$$

$$\mathbf{x}_{\text{ID}}^\top = \left( \frac{1}{r_{12}}, \frac{1}{r_{13}}, \dots \right). \quad (3c)$$

In the training, the regression coefficients are obtained by analytically solving the system of equations in matrix form  $(\mathbf{K} + \lambda I)\boldsymbol{\alpha} = \mathbf{y}$ . For each training, we also optimize the hyperparameters  $\lambda$  and  $\sigma$  (and in case of Matérn kernel also  $n$ ) by using the 20% of the training set as the validation set). The optimization of the hyperparameters was performed for each vector (or tensor) component independently. While there is a wide range of possible combinations of descriptors and kernels within KRR, RE descriptor with Gaussian kernel (the KREG model<sup>38,39</sup>) proved to be very accurate in learning potential energy surfaces (PESs)<sup>39,40</sup>.

If individual components of a vector or tensor quantity were fitted separately in this way, a proper covariance (transformation with respect to molecular rotations) would not in general be guaranteed. As a simple approach to achieve rotationally covariant results in the KRR method, we suggest the rotate-predict-rotate (RPR), a three-step approach based on the molecular tensor of inertia

$$I_{ij} = \sum_{\alpha}^{\text{atoms}} m_{\alpha} (\|\mathbf{r}^{\alpha}\|^2 \delta_{ij} - x_i^{\alpha} x_j^{\alpha}), \quad (4)$$

where  $\mathbf{r}^{\alpha} = [x_1^{\alpha}, x_2^{\alpha}, x_3^{\alpha}]$  is the radius-vector of atom  $\alpha$  with respect to the center of mass (alternatively, atomic charges and the center of charge can be employed too). This is a symmetric positive definite matrix and its diagonalization yields transformation matrices  $\mathbf{Q}$  from current orientation to the principal axes of inertia

$$\mathbf{I} = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^\top, \quad (5)$$

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0230176

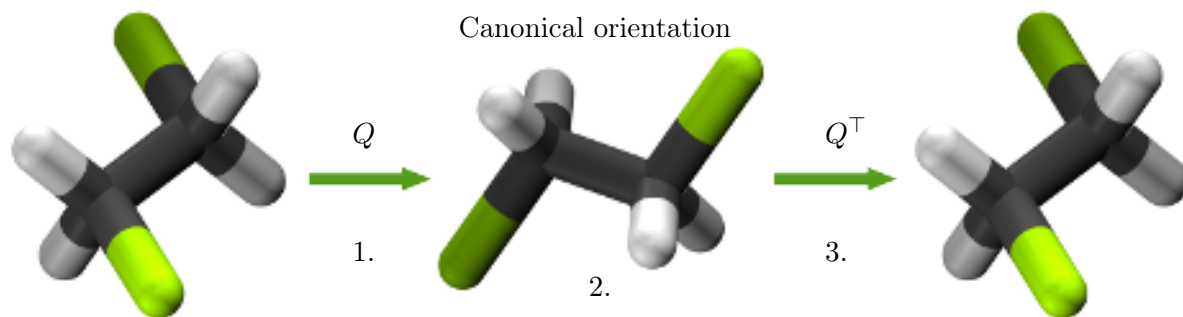


FIG. 1. Proposed RPR approach: 1. rotation of molecule and corresponding properties by  $Q$ , 2. prediction by a model trained on molecules rotated into their canonical orientation, 3. backward rotation into the initial orientation  $Q^T$ .

where  $Q$  is the orthogonal rotation matrix and  $\Lambda$  is a diagonal matrix of principal moments of inertia.

In the first step, the molecule is rotated using  $Q$  to its principal axes. I.e., if some vector property was already computed in the original orientation, it is rotated to the new canonical orientation before being used for ML training. In the second step, the ML procedure predicts the vector property relative to the canonical orientation, based on a training set where all vector properties were in this same coordinate system. In the third step, the ML estimate of the vector property is transformed back to the original orientation using  $Q^{-1} = Q^T$ . The RPR procedure, depicted in Fig. 1, should thus guarantee proper covariance of a vector property. Notice that thanks to the uniquely defined canonical orientation, the molecular descriptor itself does not need to be invariant to rotations. In fact, we tried even the set of cartesian coordinates with atomic charges as a descriptor and the covariance was properly maintained by the RPR procedure.

For higher rank tensors, the RPR procedure is easily generalized by application of the  $Q$ -transformation to each of its indices. When the general linear transformations of tensors are restricted to proper rotations, the cartesian tensors further split according to the irreducible representations of the  $SO(3)$  group. In particular, the second rank symmetric polarizability tensor decomposes to scalar polarizability being  $1/3$  of the trace and a traceless spherical tensor of rank 2 with 5 independent components. In the RPR procedure, we performed transformations of the original cartesian tensor, but verified by means of scatter plots also the results of the ML procedure on the aforementioned spherical components.



### III. NUMERICAL RESULTS

To numerically verify the RPR procedure, we have chosen 1,2-dichloroethane, which has a dipole moment and polarizability strongly dependent on the Cl-C-C-Cl dihedral angle. This makes it an ideal test system for ML of these properties. To assemble the training and testing data sets, we performed MD of this molecule in its ground state. For this purpose, the Newton-X<sup>41,42</sup> program, interfaced with Turbomole 7.3<sup>43</sup>, was used to compute the MD trajectories. All single point calculations were carried out using density functional theory (DFT) with the B3LYP functional and aug-cc-pVTZ basis set. The initial conditions were obtained from a harmonic-approximation Wigner distribution with temperature of 298 K, which was maintained during the MD using the Andersen thermostat. Geometries were then modified by varying the dihedral angle between Cl-C-C-Cl atoms by 5 degrees to yield geometries covering the whole dihedral range from 0 to 180 degrees. The change of the dihedral angle was performed by converting the Wigner's initial conditions to internal coordinates and changing the dihedral angle by an appropriate increment of a multiple of 5 degrees, while the other internal coordinates and the velocities were unchanged. While this procedure is somewhat artificial, its purpose was not to produce a physically motivated statistical ensemble, but to generate data for ML with widely varying Cl-C-C-Cl dihedral angles and consequently dipole moments.

In total we propagated seventy trajectories with a duration of 50 fs each, using a time step of 0.5 fs. This resulted in a dataset of 7,000 points (consisting of geometries, energies, dipole moments, and polarizability tensors). To analyze the learning behaviour, smaller training subsets were created by random selection from the full dataset. Each component is fitted by an independent model trained on a dataset of geometries, which were rotated, together with the corresponding properties, using rotational matrices  $\mathbf{Q}$  to the canonical orientation.

Firstly, we started with a benchmark of various combinations of descriptors and kernels to compare their performance. At this stage, all models were trained on a 1000-point dataset and evaluated on test set consisting of 1000 points selected from the trajectories. We employed the descriptors already implemented in MLatom<sup>24,45</sup>, i.e. RE, CM, ID. However, in addition to these well-established descriptors, new descriptors tailored for this molecule were introduced by incorporating cartesian or spherical coordinates of chlorine atoms (RExyzCl or REsphCl), dihedral angles (REdihedCl), or simply by utilizing cartesian coordinates multi-



This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0230176

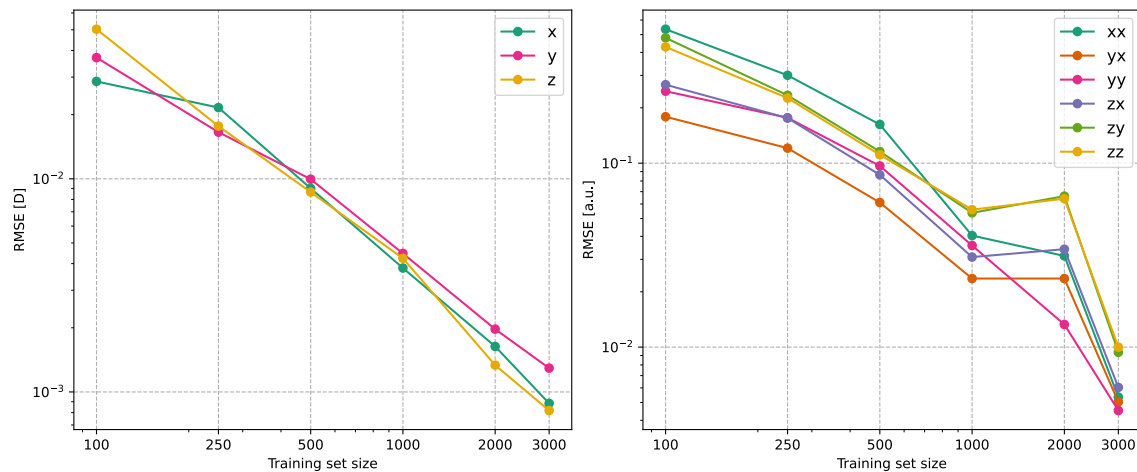


FIG. 2. Learning curves of dipole moment components in Debye (left) and polarizability components in atomic units (right). Accuracy of each model was determined on the test set consisting of 1000 points.

plied by atomic charges ( $xyzQ$ ); cartesian coordinates were taken in the canonical orientation. This was done to reflect the importance of the chlorine atoms since their orientation is crucial for the studied properties of 1,2-dichloroethane. The addition of chlorine coordinates to the descriptor turned out to have a large influence, where predictions of  $y$  and  $z$  components of the dipole moment with the **RExyzC1** descriptor produced RMSE values 20–40 times lower than the **RE** descriptor. This improvement was true for other descriptors as well, and the **RExyzC1** descriptor also outperforms **RE** in case of the polarizability components. Concerning the kernel function, Matérn kernel performed the best for dipole moment components, while the Gaussian kernel was superior for polarizability components.

The learning behaviour is shown in Fig. 2 for the dipole moment as well as for the polarizability. It can be seen that the training set consisting of 1000 points already achieved satisfactory accuracy and will thus be discussed in the following sections.

To validate the proposed approach, an independent test set covering all spatial orientations as well as dihedral angle values is required, since points of a test set constructed from the original dataset are strongly correlated with the training data, thus leading to overoptimistic RMSE. To achieve this, we fitted the ground state PES using the KREG model<sup>38,39</sup> and used Newton-X interfaced with MLatom to run multiple trajectories with the appropriate initial conditions. The KREG models were trained on the energies and energy gradients of the 1000-point dataset randomly selected from the 7000 training data points, for which

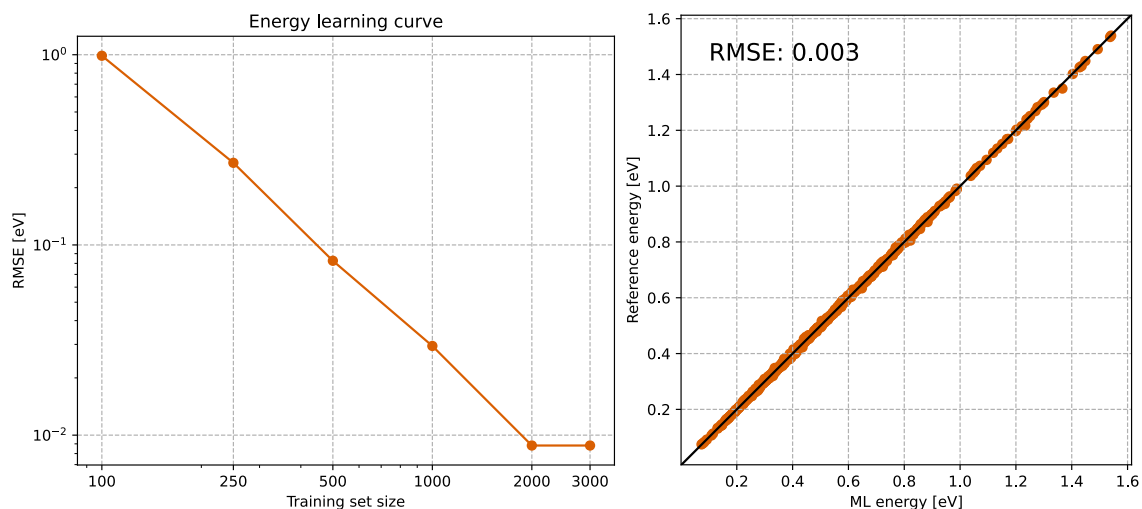


FIG. 3. Learning curve and scatter plot of a KREG model of  $S_0$  state trained on a dataset consisting of 1000 points. This model was used to propagate 24 trajectories in order to construct the test set.

learning curve and scatter plot in Fig. 3 were constructed using a test set of 1000 points.

Sampling of the dihedral angle by 15 degrees led to 24 initial geometries. A Boltzmann distribution with a temperature of 300 K was used to generate velocities, and an additional uniformly random rotation of the molecule as a whole was imposed to properly sample all molecular orientations. With a time step of 0.5 fs and a final time of 5 ps, a total of 240,000 points were obtained. This led to a dataset covering all molecular orientations in space and dihedral angles, which can be seen in Fig. 4, showing the time evolution of the Cl-C-C-Cl dihedral angle.

Geometries were taken after every 100 fs from this dataset, resulting in 1200-point test set for which reference dipole moments and polarizabilities were calculated using DFT. The RMSE for this dataset was found to be four times larger than that of the test set constructed from the original training data. This justifies the construction of the new test set, rather than taking points from initial set of trajectories. Nevertheless, the ML predictions of dipole moment as well as polarizability components demonstrated a good agreement with the reference values, as shown on Fig. 5, providing confidence in the model's predictive capabilities and the reliability of the proposed approach.

To illustrate the efficiency of the presented procedure, Fig. 6 shows the times for the prediction and training as a function of data set size. The training and prediction times were estimated on a single core of an Intel(R) Xeon(R) Gold 6226R CPU running at 2.90 GHz.

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0230176

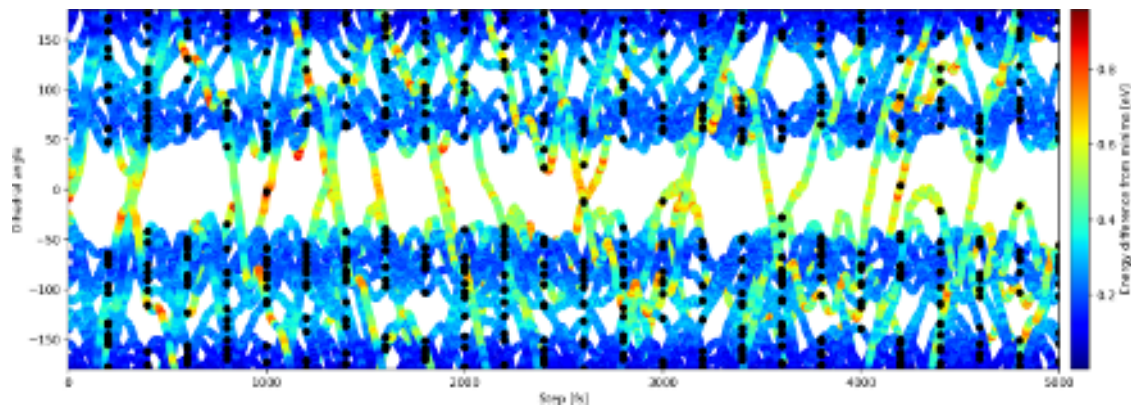


FIG. 4. Dihedral angle of Cl-C-C-Cl with respect to simulation time. Black points were taken as a test set consisting of 1200 points, for which dipole moments/polarizability tensors were calculated using DFT and accuracy of trained models was determined.

As can be seen, the training is extremely fast and takes less than 2 minutes for 3000 training points. The predictions are ca. a millisecond per point, which include the diagonalization of the inertia tensor and rotations of the molecular geometry, vectors and tensors that amount to only a fraction of the prediction time:  $2 \mu\text{s}$  for one diagonalization and  $0.08 \mu\text{s}$  for the rotation of one vector, thanks to the fact that only elementary linear algebra operations are involved and no calculations of goniometric or special functions is required, unlike in some of the equivariant approaches. This high speed combined with the excellent accuracy makes the approach a practical choice for learning vectorial properties. If the single-core computations become too time consuming for larger molecules or datasets, a parallelized implementation with a good scaling is available in MLATOM<sup>44</sup>. The RPR procedure is, e.g., particularly suitable for the transformation of a large number of vectors, for example when NAD couplings are involved, where for a  $N$ -atom molecule and  $M$  excited states there are  $O(NM^2)$  NAD coupling vectors to be processed.

While the rotation procedure scales linearly with the number of atoms, the training/prediction calculations with our descriptor scale as  $N(N-1)/2$  with the number of atoms  $N$ . This is, e.g., faster than alternative approaches based on another global descriptor Coloumb matrix which scales as  $N^2$ . We also know from the related benchmark of different machine learning potentials<sup>40</sup> that our RE descriptor with KRR provides the fastest (and, often, among the most accurate) solutions compared to methods with local descriptors such as SOAP or NN-based ANI, DeepPot, and particularly much faster than 'learned' descriptors

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0230176

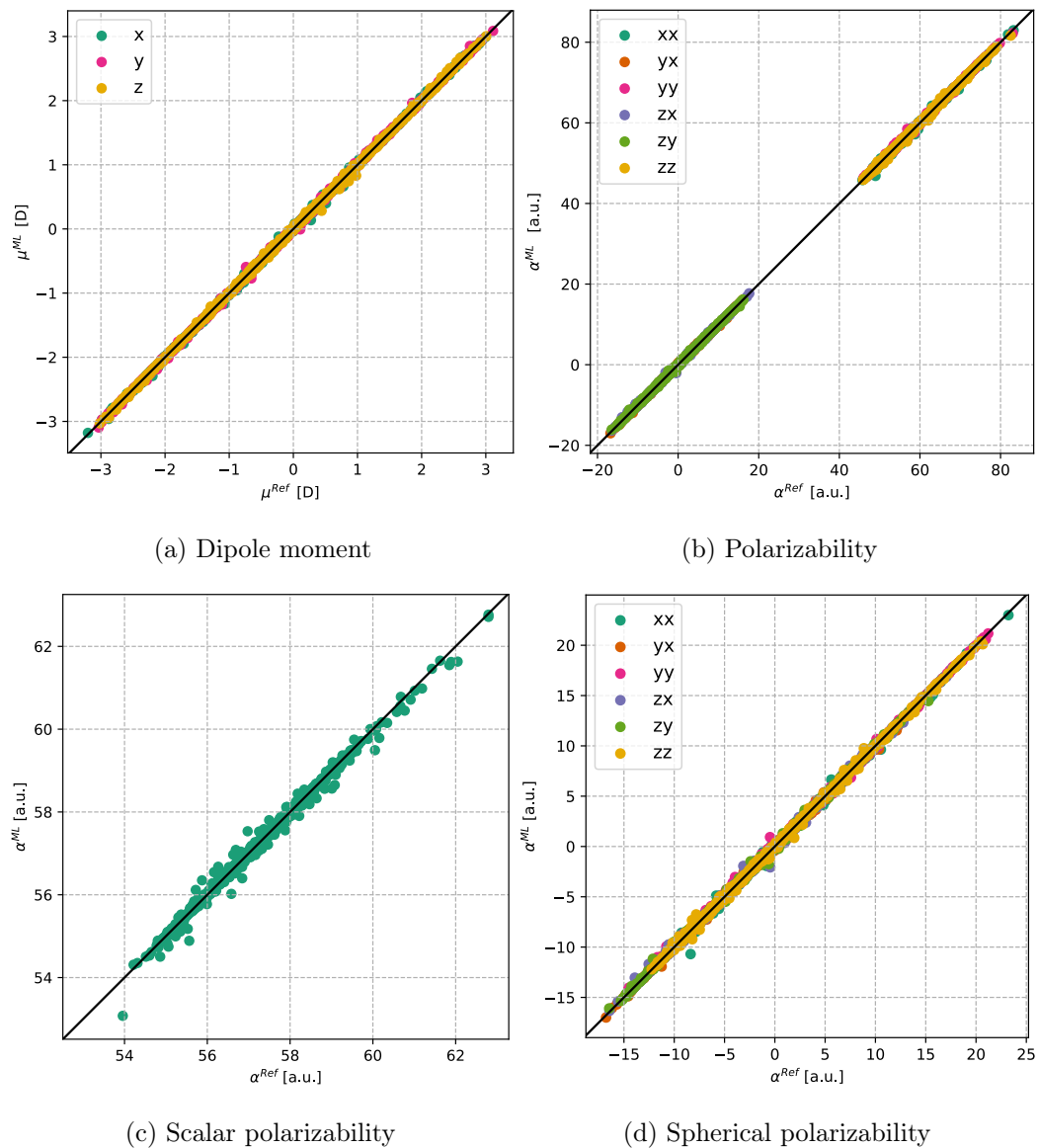


FIG. 5. Scatter plots showing the DFT dipole moment (a) and polarizability components (b) of the test set consisting of 1200 points with respect to ML predictions of the models trained on 1000 points. Dipole moment components were learned with Matérn kernel, while for polarizability Gaussian kernel performed better. In both cases RExyzC1 descriptor was used. Plots (c) and (d) are showing scalar and tensorial part of the polarizability tensor.

such as in PhysNet and should be much faster than even slower equivariant descriptors.

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0230176

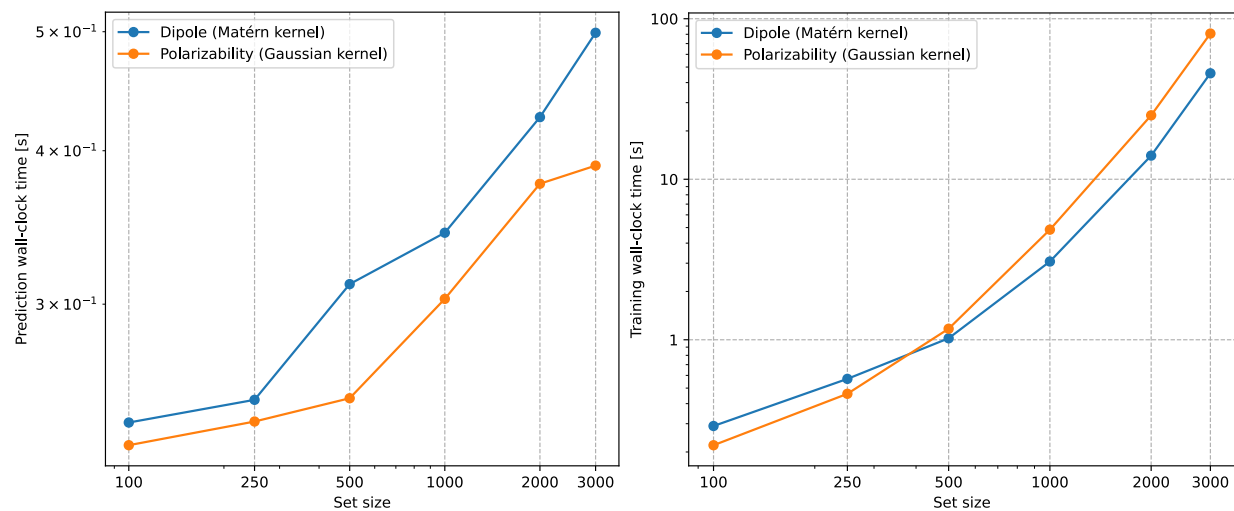


FIG. 6. Average prediction and training times for dipole moments and polarizability tensor components as a function of data set size.

#### IV. CONCLUSIONS

In this study, we addressed the challenge of machine learning vector or tensor properties. We proposed rotate-predict-rotate, a simple three-step technique to guarantee proper covariance without the need for complex equivariant ML architectures or the construction of auxiliary properties. The RPR method involves rotating the molecular geometry to a "canonical" orientation defined by the eigenvectors of the molecular tensor of inertia, training the ML model on properties in this standard orientation, and finally rotating the predicted properties back to the original orientation. We used kernel ridge regression to fit dipole moment and polarizability components of 1,2-dichloroethane and evaluated its performance. Our results demonstrate that global molecular descriptors enhanced with cartesian coordinates of chlorine atoms significantly improves ML models accuracy.

The learning curves for dipole moment and polarizability components confirmed that a training set of 1000 points was sufficient to achieve satisfactory accuracy. We used the KREG model to propagate MD simulations and constructed a test set covering all possible spatial orientations and dihedral angles. This method provides a straightforward and computationally efficient solution for achieving rotational covariance in ML models for vector and tensor properties and is also extensible to higher-rank tensors.

#### SUPPLEMENTARY MATERIAL

The supplementary material contains a Jupyter notebook and corresponding data related to descriptor benchmarks for energy, dipole moment, and polarizability components, along with other information required to generate the figures. This includes the evolution of dihedral angles, learning curves, scatter plots, predictions, and training times.

#### ACKNOWLEDGMENTS

The work of the Czech team has been supported by the *Czech Science Foundation* Grant 23-06364S and *Charles University* (project GAUK 6224). We also highly appreciate the generous admission to computing facilities owned by parties and projects contributing to the National Grid Infrastructure MetaCentrum provided under the program "Projects of Large Infrastructure for Research, Development, and Innovations" (no. LM2010005) and com-



puter time provided by the IT4I supercomputing center (Project ID:90254) supported by the Ministry of Education, Youth and Sports. MB thanks the funding provided by the European Research Council (ERC) Advanced grant SubNano (Grant agreement 832237). MB received support from the French government under the France 2030, as part of the initiative d'Excellence d'Aix-Marseille Université, A\*MIDEX (AMX-22-IN1-48). POD acknowledges funding by the National Natural Science Foundation of China (via the Outstanding Youth Scholars (Overseas, 2021) project).

- 
- <sup>1</sup> P. O. Dral, *The Journal of Physical Chemistry Letters* **11**, 2336 (2020).
  - <sup>2</sup> S. Lorenz, A. Groß, and M. Scheffler, *Chemical Physics Letters* **395**, 210 (2004).
  - <sup>3</sup> L. M. Raff, M. Malshe, M. Hagan, D. I. Doughan, M. G. Rockley, and R. Komanduri, *The Journal of Chemical Physics* **122** (2005), 10.1063/1.1850458.
  - <sup>4</sup> Y. Zhang, J. Jiang, and B. Jiang, in *Quantum Chemistry in the Age of Machine Learning*, edited by P. O. Dral (Elsevier, 2023) pp. 453–465.
  - <sup>5</sup> F. Perakis, L. De Marco, A. Shalit, F. Tang, Z. R. Kann, T. D. Kühne, R. Torre, M. Bonn, and Y. Nagata, *Chemical Reviews* **116**, 7590 (2016).
  - <sup>6</sup> X. Zhang, Q.-H. Tan, J.-B. Wu, W. Shi, and P.-H. Tan, *Nanoscale* **8**, 6435 (2016).
  - <sup>7</sup> Z. Cheng, H. Bi, S. Liu, J. Chen, A. J. Misquitta, and K. Yu, *Journal of Chemical Theory and Computation* **20**, 5598 (2024).
  - <sup>8</sup> S. Ye, W. Hu, X. Li, J. Zhang, K. Zhong, G. Zhang, Y. Luo, S. Mukamel, and J. Jiang, *Proceedings of the National Academy of Sciences* **116**, 11612 (2019).
  - <sup>9</sup> J. Zhang, S. Ye, K. Zhong, Y. Zhang, Y. Chong, L. Zhao, H. Zhou, S. Guo, G. Zhang, B. Jiang, S. Mukamel, and J. Jiang, *The Journal of Physical Chemistry B* **125**, 6171 (2021).
  - <sup>10</sup> S. Ye, K. Zhong, J. Zhang, W. Hu, J. D. Hirst, G. Zhang, S. Mukamel, and J. Jiang, *Journal of the American Chemical Society* **142**, 19071 (2020).
  - <sup>11</sup> J. Westermayr, M. Gastegger, M. F. S. J. Menger, S. Mai, L. González, and P. Marquetand, *Chemical Science* **10**, 8100 (2019).
  - <sup>12</sup> J. Westermayr, F. A. Faber, A. S. Christensen, O. A. von Lilienfeld, and P. Marquetand, *Machine Learning: Science and Technology* **1**, 025009 (2020).



- 13 J. Westermayr, M. Gastegger, and P. Marquetand, *The Journal of Physical Chemistry Letters* **11**, 3828 (2020).
- 14 J. O. Richardson, *The Journal of Chemical Physics* **158**, 011102 (2023).
- 15 O. T. Unke and M. Meuwly, *J. Chem. Theory Comput.* **15**, 3678– (2019).
- 16 Y. Chen, S. V. Pios, M. F. Gelin, and L. Chen, *Journal of Chemical Theory and Computation* **20**, 4703 (2024).
- 17 Y. Zhang, S. Ye, J. Zhang, C. Hu, J. Jiang, and B. Jiang, *The Journal of Physical Chemistry B* **124**, 7284 (2020).
- 18 A. S. Christensen, F. A. Faber, and O. A. von Lilienfeld, *J. Chem. Phys.* **150**, 064105 (2019).
- 19 A. P. Bartók, R. Kondor, and G. Csányi, *Physical Review B* **87**, 184115 (2013).
- 20 A. Glielmo, P. Sollich, and A. De Vita, *Physical Review B* **95**, 214302 (2017).
- 21 A. Grisafi, D. M. Wilkins, G. Csányi, and M. Ceriotti, *Physical Review Letters* **120**, 036002 (2018).
- 22 D. M. Wilkins, A. Grisafi, Y. Yang, K. U. Lao, R. A. DiStasio, and M. Ceriotti, *Proceedings of the National Academy of Sciences* **116**, 3401 (2019).
- 23 N. Raimbault, A. Grisafi, M. Ceriotti, and M. Rossi, *New Journal of Physics* **21**, 105001 (2019).
- 24 Z. Fan, Z. Zeng, C. Zhang, Y. Wang, K. Song, H. Dong, Y. Chen, and T. Ala-Nissila, *Physical Review B* **104**, 104309 (2021).
- 25 N. Xu, P. Rosander, C. Schäfer, E. Lindgren, N. Österbacka, M. Fang, W. Chen, Y. He, Z. Fan, and P. Erhart, *Journal of Chemical Theory and Computation* **20**, 3273 (2024).
- 26 Y. Zhang, C. Hu, and B. Jiang, *The Journal of Physical Chemistry Letters* **10**, 4962 (2019).
- 27 B. Anderson, T. S. Hy, and R. Kondor, 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).
- 28 K. T. Schütt, O. T. Unke, and M. Gastegger, in *International Conference on Machine Learning* (2021).
- 29 N. Thomas, T. Smidt, S. Kearnes, L. Yang, L. Li, K. Kohlhoff, and P. Riley, “Tensor field networks: Rotation- and translation-equivariant neural networks for 3d point clouds,” (2018), arXiv:1802.08219.
- 30 Y. Chen, S. V. Pios, M. F. Gelin, and L. Chen, *Journal of Chemical Theory and Computation* **20**, 4703 (2024).

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0230176

- <sup>31</sup> L. Zhang, M. Chen, X. Wu, H. Wang, W. E. , and R. Car, *Physical Review B* **102**, 041121 (2020).
- <sup>32</sup> Y. Zhang and B. Jiang, *Nature Communications* **14** (2023), 10.1038/s41467-023-42148-y.
- <sup>33</sup> T. Bereau, D. Andrienko, and O. A. von Lilienfeld, *J. Chem. Theory Comp.* **11**, 3225 (2015).
- <sup>34</sup> M. Rupp, R. Ramakrishnan, and O. A. von Lilienfeld, *J. Phys. Chem. Lett.* **6**, 3309 (2015).
- <sup>35</sup> T. Bereau, D. Andrienko, and O. A. von Lilienfeld, *Journal of Chemical Theory and Computation* **11**, 3225 (2015).
- <sup>36</sup> C. Liang, G. Tocci, D. M. Wilkins, A. Grisafi, S. Roke, and M. Ceriotti, *Physical Review B* **96**, 041407 (2017).
- <sup>37</sup> D. Hu and P. Huo, *J. Chem. Theory Comp.* **19**, 2353 (2023).
- <sup>38</sup> P. O. Dral, A. Owens, S. N. Yurchenko, and W. Thiel, *The Journal of Chemical Physics* **146**, 244108 (2017).
- <sup>39</sup> Y.-F. Hou, F. Ge, and P. O. Dral, *Journal of Chemical Theory and Computation* (2023), 10.1021/acs.jctc.2c01038.
- <sup>40</sup> M. Pinheiro Jr, F. Ge, N. Ferré, P. O. Dral, and M. Barbatti, *Chemical Science* **12**, 14396 (2021).
- <sup>41</sup> M. Barbatti, M. Bondanza, R. Crespo-Otero, B. Demoulin, P. O. Dral, G. Granucci, F. Kossoski, H. Lischka, B. Mennucci, S. Mukherjee, M. Pederzoli, M. Persico, M. P. Jr, J. Pittner, F. Plasser, E. S. Gil, and L. Stojanovic, *Journal of Chemical Theory and Computation* **18**, 6851 (2022).
- <sup>42</sup> M. Barbatti, G. Granucci, M. Ruckebauer, F. Plasser, R. Crespo-Otero, J. Pittner, M. Persico, and H. Lischka, “Newton-X - package for newtonian dynamics close to the crossing seam, <http://www.univie.ac.at/newtonx>,” (2008-2013).
- <sup>43</sup> “TURBOMOLE V7.3 2018, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from <http://www.turbomole.com>.” (2018).
- <sup>44</sup> P. O. Dral, *Journal of Computational Chemistry* **40**, 2339 (2019).
- <sup>45</sup> P. O. Dral, F. Ge, B.-X. Xue, Y.-F. Hou, M. Pinheiro Jr, J. Huang, and M. Barbatti, *Topics in Current Chemistry* **379**, 27 (2021).