Symmetry-aware models beyond energies and forces

Martin Uhrín

Computational Atomistic Methods & Machine Learning Group, UGA

Grenoble Alpes

Multidisciplinary Institute
In Artificial Intelligence

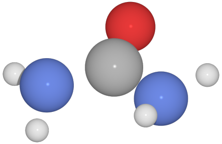
Physics inspired machine learning models

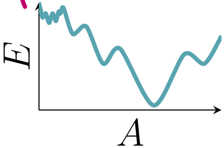
Accelerating property prediction with machine learning

$$A := \{\mathbf{r}_I, Z_I\} \xrightarrow{\hat{H}|\Psi\rangle = E|\Psi\rangle} E(\{\mathbf{r}_I, Z_I\})$$

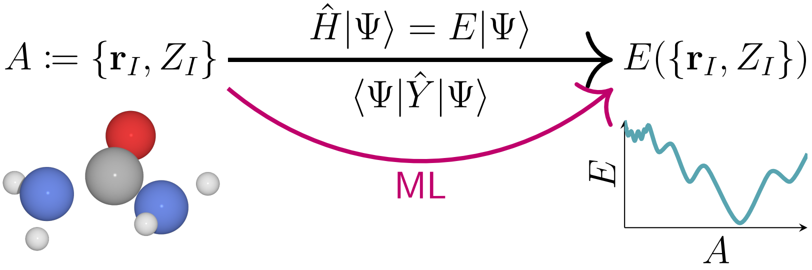
$$\langle\Psi|\hat{Y}|\Psi\rangle$$

ML

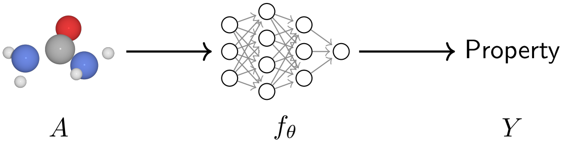




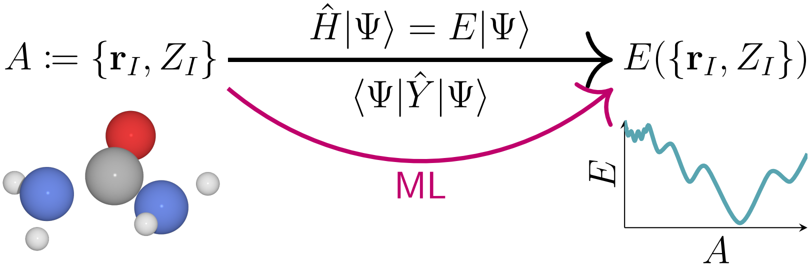
Accelerating property prediction with machine learning



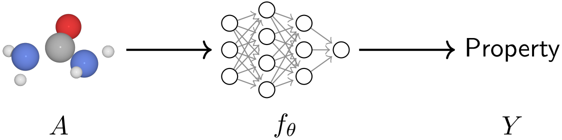
Our ML models learns to map
from atomic coordinates to properties: $f_\theta : A \rightarrow Y$,



Accelerating property prediction with machine learning



Our ML models learns to map from atomic coordinates to properties: $f_\theta : A \to Y$,

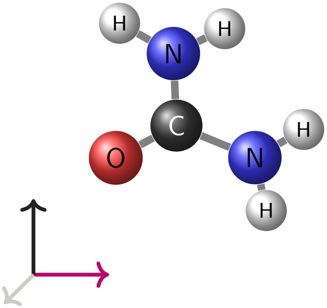


but A is not a suitable input. We want something that respects the symmetries of the underlying physical laws:

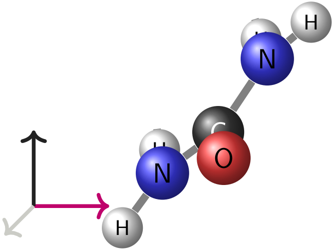
- permutation,
- translation,
- rotation,

and that is smooth, continuous and complete.

Symmetry awareness



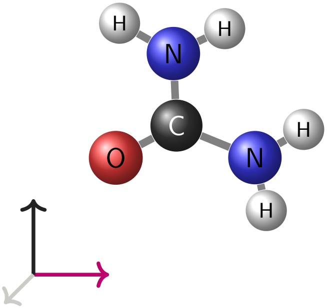
a)



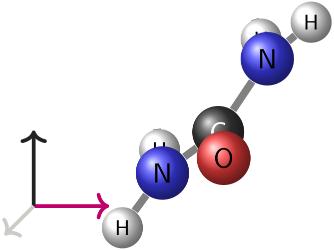
b)

Regular learning model will *see* a) and b) as completely different inputs

Symmetry awareness



a)

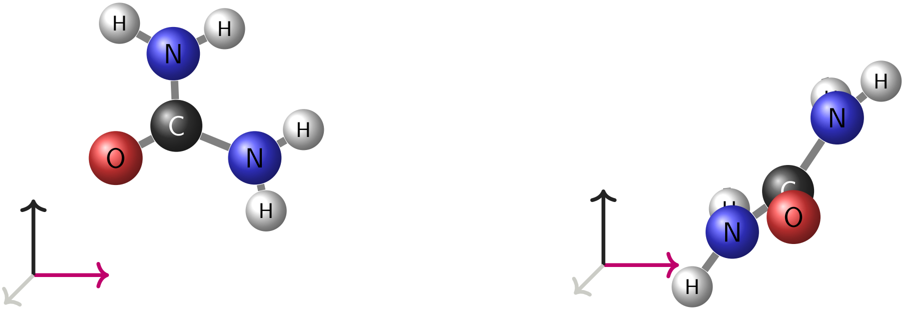


b)

Regular learning model will *see* a) and b) as completely different inputs

Invariant model will *see* the same inputs.

Symmetry awareness



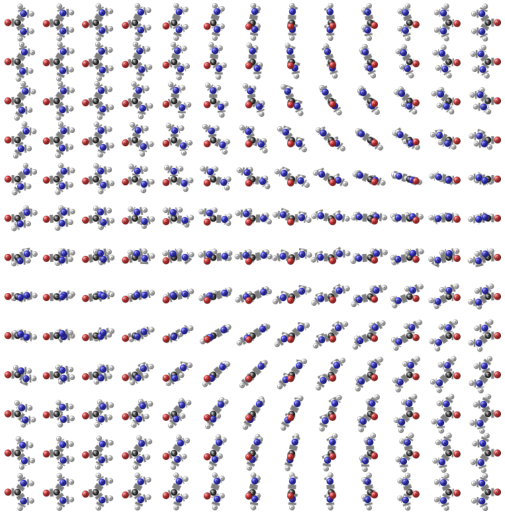
a) b)

Regular learning model will *see* a) and b) as completely different inputs

Invariant model will *see* the same inputs.

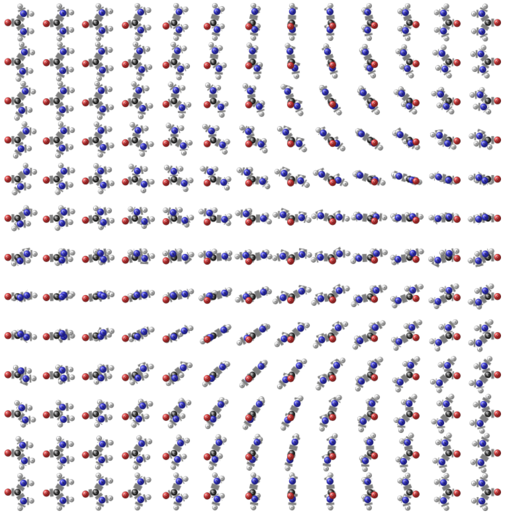
Equivariant model will *see* the same inputs, just described differently.

Learning with no symmetry awareness



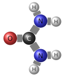
requires hundreds fold augmentation

Learning with no symmetry awareness



requires hundreds fold augmentation

Learning with symmetry-awareness



Lifting the hood on equivariant neural networks

Invariant model

$$f_\theta(D_X[g]x) = f_\theta(x)$$

for the group g.

Lifting the hood on equivariant neural networks

Invariant model

$$f_\theta(D_X[g]x) = f_\theta(x)$$

for the group g. Equivariant model:

$$D_Y[g]f_\theta(x) = f_\theta(D_X[g]x) \quad \forall g \in G, \forall x \in X$$

Natural way of extending ML models to handled non-scalar inputs and outputs (vectors, tensors),

Lifting the hood on equivariant neural networks

Invariant model

$$f_\theta(D_X[g]x) = f_\theta(x)$$

for the group g. Equivariant model:

$$D_Y[g]f_\theta(x) = f_\theta(D_X[g]x) \quad \forall g \in G, \forall x \in X$$

Natural way of extending ML models to handled non-scalar inputs and outputs (vectors, tensors),

Fully-connected layer

$$\phi\left(\begin{bmatrix} x_1 \\ y_1 \\ z_1 \\ x_2 \\ y_2 \\ z_2 \end{bmatrix}^\top \begin{bmatrix} w_1 & \\ w_2 & \\ w_3 & \cdots \\ w_4 & \\ w_5 & \\ w_6 & \end{bmatrix}\right) = \begin{bmatrix} x_3 & \cdots \end{bmatrix}$$

$\phi(x)$ - activation function

Lifting the hood on equivariant neural networks

Equivariant layer

$$\begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} \otimes \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = \begin{bmatrix} x_1x_2 & x_1y_2 & x_1z_2 \\ y_1x_2 & y_1y_2 & y_1z_2 \\ z_1x_2 & z_1y_2 & z_1z_2 \end{bmatrix}$$

Irreducible form:

$l = 0$ scalar - $w_1(x_1x_2 + y_1y_2 + z_1z_2)$

$l = 1$ vector - $w_2 \begin{bmatrix} y_1z_2 - z_1x_2 \\ z_1x_2 - x_1z_2 \\ x_1y_2 - x_2y_1 \end{bmatrix}$

$l = 2$ vector - $w_3 \begin{bmatrix} x_1z_2 + z_1x_2 \\ x_1y_2 + y_1x_2 \\ 2y_1y_2 - x_1x_2 - z_1z_2 \\ y_1z_2 - z_1y_2 \\ z_1z_2 - x_1x_2 \end{bmatrix}$

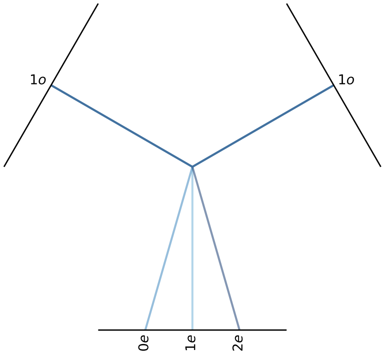
Fully-connected layer

$$\phi\left(\begin{bmatrix} x_1 \\ y_1 \\ z_1 \\ x_2 \\ y_2 \\ z_2 \end{bmatrix}^\top \begin{bmatrix} w_1 & \\ w_2 & \\ w_3 & \cdots \\ w_4 & \\ w_5 & \\ w_6 & \end{bmatrix}\right) = \begin{bmatrix} x_3 & \cdots \end{bmatrix}$$

$\phi(x)$ - activation function

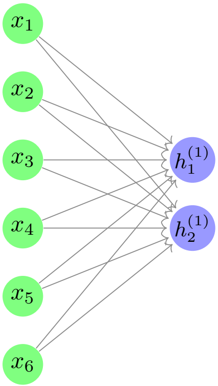
Lifting the hood on equivariant neural networks

Equivariant layer



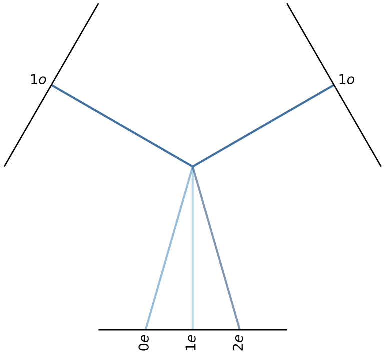
Outputs: $i \times 0e \oplus j \times 1e \oplus k \times 2e$

Fully-connected layer



Lifting the hood on equivariant neural networks

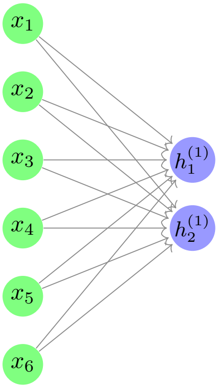
Equivariant layer



Outputs: $i \times 0e \oplus j \times 1e \oplus k \times 2e$

- Maintains symmetry (cannot *rotate* input)
- Minimal amount of parameters

Fully-connected layer



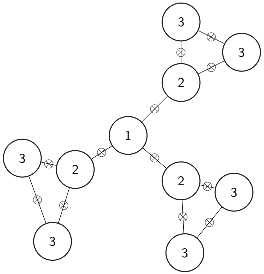
- No constraint on transformation (can scale/rotate/translate inputs)
- Many parameters

python ecosystem for development of equivariant learning

Complete ecosystem built in JAX and `e3nn-jax`:

tensorial

- Native support for equivariant graph convolutions
- Complete flexibility over input and output types:
  - per-node, per-edge or global
  - arbitrary node attributes (scalars, tensors)
- Implemented in JAX: 4-20x speedup over pytorch



reax

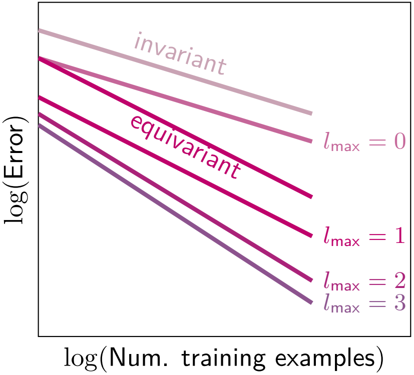
- pytorch lightning-like model training

e3md

- Built in implementations of
  - NequIP
  - Allegro
  - MACE

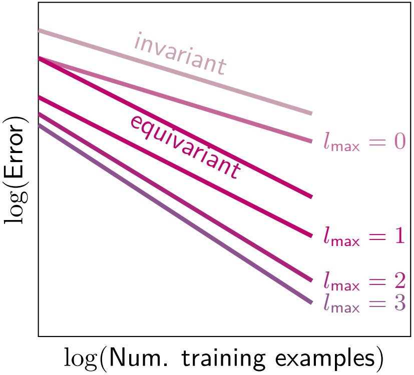
```
def f(x):
  return ** 2

# Evaluate gradient
jax.grad(f)(5.)

# Multiple inputs, scalar output
jax.grad(dft_energy(atom_positions))

# Multiple inputs, multiple outputs
jax.jacfwd(get_bands(atom_positions))
```

Data efficiency



Geiger M., Smidt T., ..., MU et al., (2022). e3nn/e3nn: 2022-04-13 (0.5.0). Zenodo. https://doi.org/10.5281/zenodo.6459381

Data efficiency



System		NequIP[a)]	NequIP[b)]	NequIP[c)]	DeepMD
Liquid Water	Energy	-	1.6	1.7	1.0
	Forces	11.9	49.4	11.6	40.4
Ice Ih (b)	Energy	-	2.5	4.3	0.7
	Forces	10.2	55.8	9.9	43.3
Ice Ih (c)	Energy	-	3.9	10.2	0.7
	Forces	12.0	27.7	11.7	26.8
Ice Ih (d)	Energy	-	2.6	12.7	0.8
	Forces	9.8	23.2	9.5	25.4

NequIP model trained on $< 0.1\%$ of DeepMD model

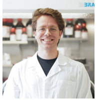
S. Batzner et al., Nature Communications **13**, 2453 (2022)

Geiger M., Smidt T., . . . , MU et al., (2022). e3nn/e3nn: 2022-04-13 (0.5.0). Zenodo. https://doi.org/10.5281/zenodo.6459381

Electric and magnetic response

Electric response

Lorenzo Bastonero
Uni Bremen



Mattia Ragni
UGA

Alessandro D'Urso
EPFL

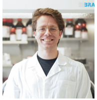
$$E = U - \vec{P} \cdot \vec{\epsilon}$$

$\vec{P}$ - polarization

$\vec{\epsilon}$ - external electric field

Electric response

Lorenzo Bastonero
Uni Bremen



Mattia Ragni
UGA

Alessandro D'Urso
EPFL

$$E = U - \vec{P} \cdot \vec{\epsilon}$$

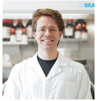
$\vec{P}$ - polarization

$\vec{\epsilon}$ - external electric field

Born effective charge tensor

$$\mathcal{Z}^{I}_{\alpha\beta} = \frac{d^2 E}{\partial \epsilon_\alpha \partial \tau^{I}_{\beta}} = \vec{\epsilon} \otimes \vec{\tau}^{I}$$

Electric response



Lorenzo Bastonero
Uni Bremen

Mattia Ragni
UGA

Alessandro D'Urso
EPFL

$$E = U - \vec{P} \cdot \vec{\epsilon}$$

$\vec{P}$ - polarization

$\vec{\epsilon}$ - external electric field

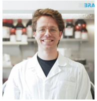
Born effective charge tensor

$$\mathcal{Z}^I_{\alpha\beta} = \frac{d^2 E}{\partial \epsilon_\alpha \partial \tau^I_\beta} = \vec{\epsilon} \otimes \vec{\tau}^I$$

Raman tensor - Can be calculated via derivatives of the enthaly:

$$\chi^I_{\alpha\beta\gamma} = -\frac{1}{\Omega} \frac{\partial^3 E}{\partial \epsilon_\alpha \partial \epsilon_\beta \partial \tau^I_\gamma} = -\frac{1}{\Omega} \vec{\epsilon} \otimes \vec{\epsilon} \otimes \vec{\tau}^I$$

Electric response


Lorenzo Bastonero
Uni Bremen

Mattia Ragni
UGA

Alessandro D'Urso
EPFL

$$E = U - \vec{P} \cdot \vec{\epsilon}$$

$\vec{P}$ - polarization

$\vec{\epsilon}$ - external electric field

Born effective charge tensor

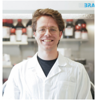
$$\mathcal{Z}^I_{\alpha\beta} = \frac{d^2 E}{\partial \epsilon_\alpha \partial \tau^I_\beta} = \vec{\epsilon} \otimes \vec{\tau}^I$$

rank 2 non-symmetric tensor, irreps $0e \oplus 1e \oplus 2e$

Raman tensor - Can be calculated via derivatives of the enthaly:

$$\chi^I_{\alpha\beta\gamma} = -\frac{1}{\Omega} \frac{\partial^3 E}{\partial \epsilon_\alpha \partial \epsilon_\beta \partial \tau^I_\gamma}$$
$$= -\frac{1}{\Omega} \vec{\epsilon} \otimes \vec{\epsilon} \otimes \vec{\tau}^I$$

Electric response



Lorenzo Bastonero
Uni Bremen

Mattia Ragni
UGA

Alessandro D'Urso
EPFL

$$E = U - \vec{P} \cdot \vec{\epsilon}$$

$\vec{P}$ - polarization

$\vec{\epsilon}$ - external electric field

Born effective charge tensor

$$\mathcal{Z}^I_{\alpha\beta} = \frac{d^2 E}{\partial \epsilon_\alpha \partial \tau^I_\beta} = \vec{\epsilon} \otimes \vec{\tau}^I$$

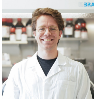
rank 2 non-symmetric tensor, irreps $0e \oplus 1e \oplus 2e$

Raman tensor - Can be calculated via derivatives of the enthaly:

$$\chi^I_{\alpha\beta\gamma} = -\frac{1}{\Omega} \frac{\partial^3 E}{\partial \epsilon_\alpha \partial \epsilon_\beta \partial \tau^I_\gamma} = -\frac{1}{\Omega} \vec{\epsilon} \otimes \vec{\epsilon} \otimes \vec{\tau}^I$$

rank 3 tensor symmetric in α, β irreps $2 \times 1o \oplus 2o \oplus 3o$

Electric response

Lorenzo Bastonero
Uni Bremen

Mattia Ragni
UGA

Alessandro D'Urso
EPFL

$$E = U - \vec{P} \cdot \vec{\epsilon}$$

$\vec{P}$ - polarization

$\vec{\epsilon}$ - external electric field

Born effective charge tensor

$$\mathcal{Z}^I_{\alpha\beta} = \frac{d^2 E}{\partial \epsilon_\alpha \partial \tau^I_\beta} = \vec{\epsilon} \otimes \vec{\tau}^I$$

rank 2 non-symmetric tensor, irreps $0e \oplus 1e \oplus 2e$

Raman tensor - Can be calculated via derivatives of the enthaly:

$$\chi^I_{\alpha\beta\gamma} = -\frac{1}{\Omega} \frac{\partial^3 E}{\partial \epsilon_\alpha \partial \epsilon_\beta \partial \tau^I_\gamma}$$
$$= -\frac{1}{\Omega} \vec{\epsilon} \otimes \vec{\epsilon} \otimes \vec{\tau}^I$$

rank 3 tensor symmetric in α, β irreps $2 \times 1o \oplus 2o \oplus 3o$

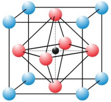
Model - double headed EGNN:

- R^I
- Z^I

→ tensorial EGNN →

- $\mathcal{Z}^I$
- χ^I

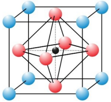
Electric response results

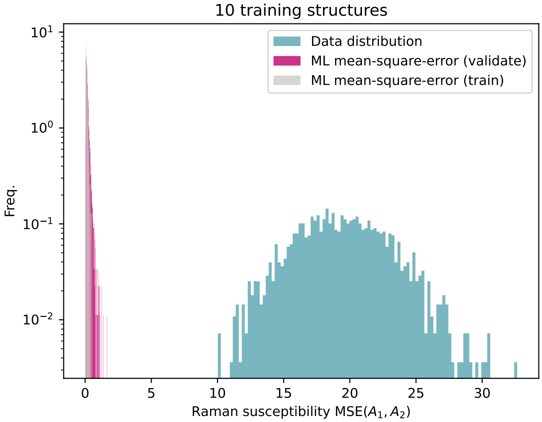


50 structures $BaTiO_3$, 135 atoms, at 400K with random displacements from equilibrium.

Electric response results

50 structures $BaTiO_3$, 135 atoms, at 400K with random displacements from equilibrium.

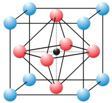


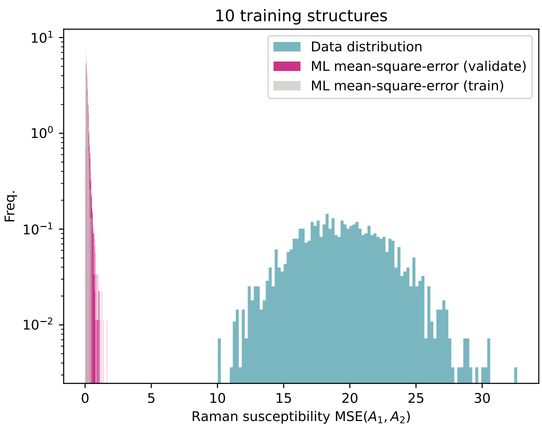


Raman susceptibilities

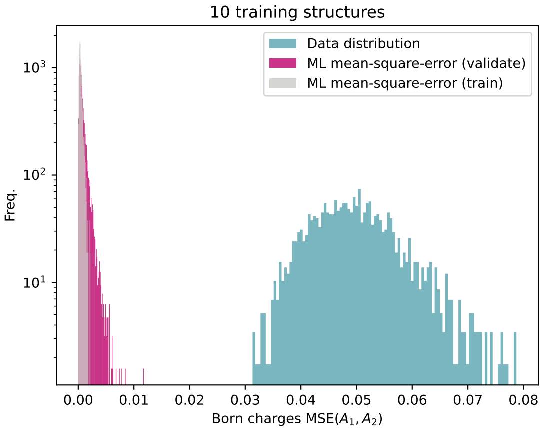
Electric response results

50 structures $BaTiO_3$, 135 atoms, at 400K with random displacements from equilibrium.





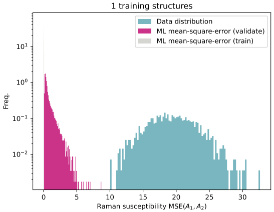
Raman susceptibilities

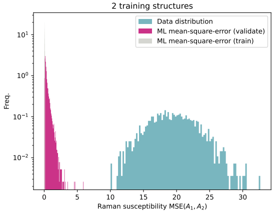


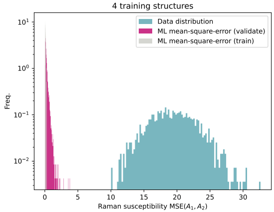
Born effective charges

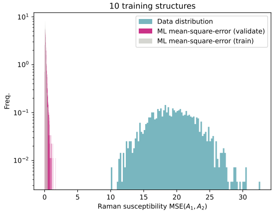
Accuracy as function of training set size

Raman susceptibilities

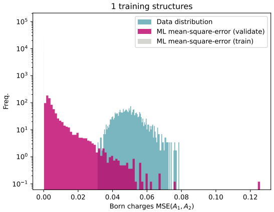


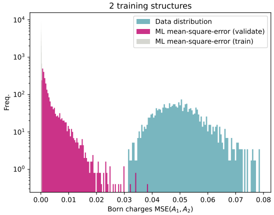


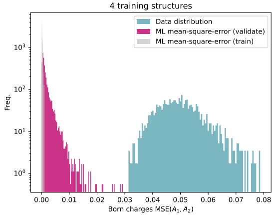


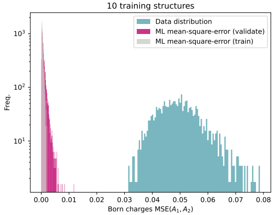


Born effective charges









Injecting more physics

Born effective charge tensor

$$\mathcal{Z}^{I}_{\alpha\beta} = \frac{d^2 E}{\partial\epsilon_\alpha \partial\tau^I_\beta}$$

Raman tensor

$$\chi^{I}_{\alpha\beta\gamma} = -\frac{1}{\Omega}\frac{\partial^3 E}{\partial\epsilon_\alpha \partial\epsilon_\beta \partial\tau^I_\gamma}$$

Injecting more physics

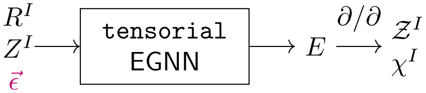
Born effective charge tensor

$$\mathcal{Z}^I_{\alpha\beta} = \frac{d^2 E}{\partial \epsilon_\alpha \partial \tau^I_\beta}$$

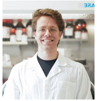
Raman tensor

$$\chi^I_{\alpha\beta\gamma} = -\frac{1}{\Omega} \frac{\partial^3 E}{\partial \epsilon_\alpha \partial \epsilon_\beta \partial \tau^I_\gamma}$$

Model



Magnetic response



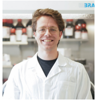
Mattia Ragni
UGA

$$E = U - \vec{\mu} \cdot \vec{B}$$

$\vec{\mu}$ - magnetic dipole moment

$\vec{B}$ - external magnetic field

Magnetic response



Mattia Ragni
UGA

NMR shielding tensor

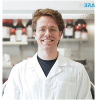
$$\sigma^I_{\alpha\beta} = \frac{d^2 E}{\partial B_\alpha \partial \mu^I_\beta} = \vec{B} \otimes \vec{\mu}^I$$

$$E = U - \vec{\mu} \cdot \vec{B}$$

$\vec{\mu}$ - magnetic dipole moment

$\vec{B}$ - external magnetic field

Magnetic response



Mattia Ragni
UGA

NMR shielding tensor

$$\sigma^I_{\alpha\beta} = \frac{d^2 E}{\partial B_\alpha \partial \mu^I_\beta} = \vec{B} \otimes \vec{\mu}^I$$

rank 2 non-symmetric tensor, irreps $0e \oplus 1e \oplus 2e$

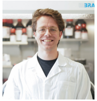
$$E = U - \vec{\mu} \cdot \vec{B}$$

$\vec{\mu}$ - magnetic dipole moment

$\vec{B}$ - external magnetic field

$$\sigma_{\alpha\beta} = \begin{bmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{bmatrix}$$

Magnetic response



Mattia Ragni
UGA

$$E = U - \vec{\mu} \cdot \vec{B}$$

$\vec{\mu}$ - magnetic dipole moment

$\vec{B}$ - external magnetic field

NMR shielding tensor

$$\sigma^I_{\alpha\beta} = \frac{d^2 E}{\partial B_\alpha \partial \mu^I_\beta} = \vec{B} \otimes \vec{\mu}^I$$

$$\sigma_{\alpha\beta} = \begin{bmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{bmatrix}$$

rank 2 non-symmetric tensor, irreps $0e \oplus 1e \oplus 2e$

Model - GNN:

- R^I
- Z^I

$\longrightarrow$ tensorial EGNN $\longrightarrow$

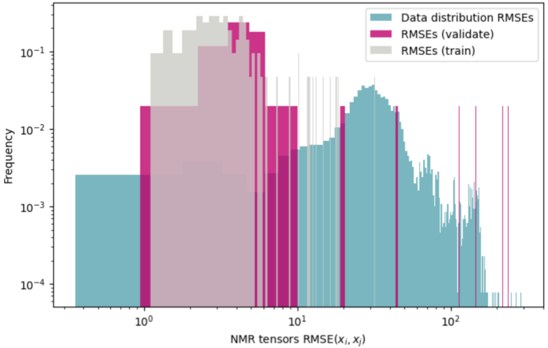
- σ^I

Magnetic response results

56 structures, 135 Si^{29} shieldings, containing N, O, Li, Na, Ni, Mg and S. 80:20 split.

Magnetic response results

56 structures, 135 Si^{29} shieldings, containing N, O, Li, Na, Ni, Mg and S. 80:20 split.



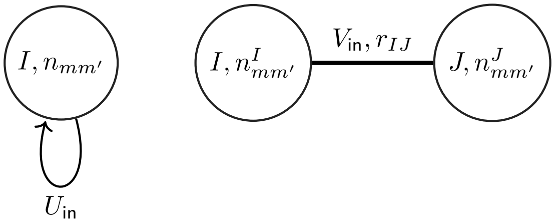
Shielding tensors

Hubbard parameters

Input representations

Relevant symmetries:

- Translation
- Rotation/reflection
- Permutation of labels

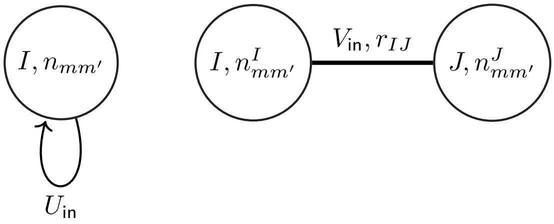


Input representations

Relevant symmetries:

- Translation
- Rotation/reflection
- Permutation of labels

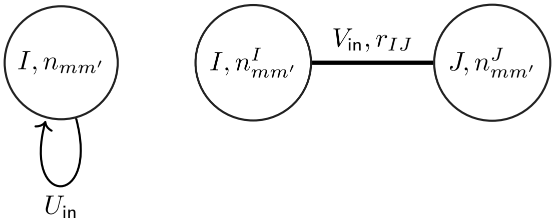
Translation $\phi^I_m(\mathbf{r}) \equiv \phi^{\gamma(I)}_m(\mathbf{r} - \mathbf{R}_I)$ ✓



Input representations

Relevant symmetries:

- Translation
- Rotation/reflection
- Permutation of labels



Translation $\phi^I_m(\mathbf{r}) \equiv \phi^{\gamma(I)}_m(\mathbf{r} - \mathbf{R}_I)$ ✓

Rotation Use equivariant neural network to respect $O(3)$ symmetry, express data terms of irreps.

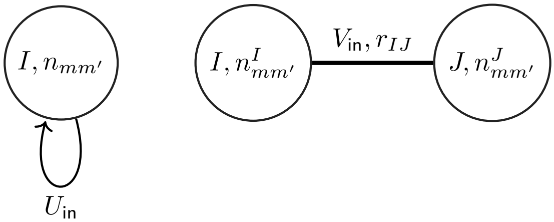
Occupation matrices

$$n^{IJ\sigma}_{mm'} = \sum_{v,\mathbf{k}} f^\sigma_{v,\mathbf{k}} \langle \psi^\sigma_{v,\mathbf{k}} | \phi^J_{m'} \rangle \langle \phi^I_m | \psi^\sigma_{v,\mathbf{k}} \rangle \,,$$

Input representations

Relevant symmetries:

- Translation
- Rotation/reflection
- Permutation of labels



Translation $\phi^I_m(\mathbf{r}) \equiv \phi^{\gamma(I)}_m(\mathbf{r} - \mathbf{R}_I)$ ✓

Rotation Use equivariant neural network to respect $O(3)$ symmetry, express data terms of irreps.
Occupation matrices

$$n^{IJ\sigma}_{mm'} = \sum_{v,\mathbf{k}} f^\sigma_{v,\mathbf{k}} \langle \psi^\sigma_{v,\mathbf{k}} | \phi^J_{m'} \rangle \langle \phi^I_m | \psi^\sigma_{v,\mathbf{k}} \rangle \, ,$$

representations can be written as $\sum D^{(l)} \otimes D^{(l)}$, which can be decomposed into irreps:

$$\begin{aligned} l &= 1: \quad D^{(0)} \oplus D^{(2)} \\ l &= 2: \quad D^{(0)} \oplus D^{(2)} \oplus D^{(4)} \end{aligned}$$

These tensors can be used directly as inputs to `e3nn` library.

Input representations

Permutation Want invariants wrt spin label, σ. Create permutationally invariant polynomials:

$$\begin{aligned} n^1_{mm'} &= \quad n^{\uparrow}_{mm'} + n^{\downarrow}_{mm'} \\ n^2_{mm'} &= \quad n^{\uparrow}_{mm'} \otimes n^{\downarrow}_{mm'} \end{aligned}$$

Use n^1_{p} as shorthand for $l = 1$ element and n^1_{d} for an $l = 2$.

Translation $\phi^I_m(\mathbf{r}) \equiv \phi^{\gamma(I)}_m(\mathbf{r} - \mathbf{R}_I)$ ✓

Rotation Use equivariant neural network to respect $O(3)$ symmetry, express data terms of irreps.

Occupation matrices

$$n^{IJ\sigma}_{mm'} = \sum_{v,\mathbf{k}} f^{\sigma}_{v,\mathbf{k}} \langle \psi^{\sigma}_{v,\mathbf{k}} | \phi^J_{m'} \rangle \langle \phi^I_m | \psi^{\sigma}_{v,\mathbf{k}} \rangle \,,$$

representations can be written as $\sum D^{(l)} \otimes D^{(l)}$, which can be decomposed into irreps:

$$\begin{aligned} l &= 1: \quad D^{(0)} \oplus D^{(2)} \\ l &= 2: \quad D^{(0)} \oplus D^{(2)} \oplus D^{(4)} \end{aligned}$$

These tensors can be used directly as inputs to `e3nn` library.

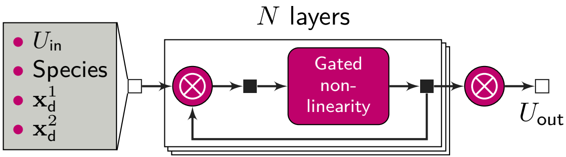
Hubbard U machine learning model

Austin Zadoks
EPFL

Luca Binci
UC Berkeley

Iurii Timrov
PSI

Body-order correlation induced by taking tensor products:



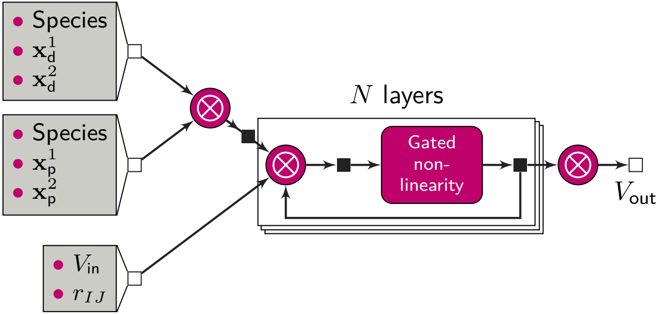
□ = Input/output tensor

■ = Hidden state

⊗ = Tensor product (with learnable weights)

Model for predicting Hubbard U values

Hubbard V machine learning model



Model for predicting Hubbard V values.

Model accuracy and transferability

Hubbard U training

Structure type	Chemical composition	#
Olivine	Li_xFePO_4	5
	Li_xMnPO_4	5
	$Li_xFe_{0.5}Mn_{0.5}PO_4$	5
Spinel	$Li_xMn_2O_4$	2
	$Li_xMn_{1.5}Ni_{0.5}O_4$	2
Layered	Li_xNiO_2	2
	Li_xMnO_2	2
Tunnel	α-MnO_2	1
Rutile	β-MnO_2	1
Perovskite	$YNiO_3$	1
	$PrNiO_3$	1
Total		27

Element	Number of data points	
	U	V
Ni	396 (124)	67,802 (63,332)
Mn	856 (284)	162,272 (153,232)
Fe	138 (120)	22,511 (21,483)
Total	1,390 (528)	252,585 (238,047)

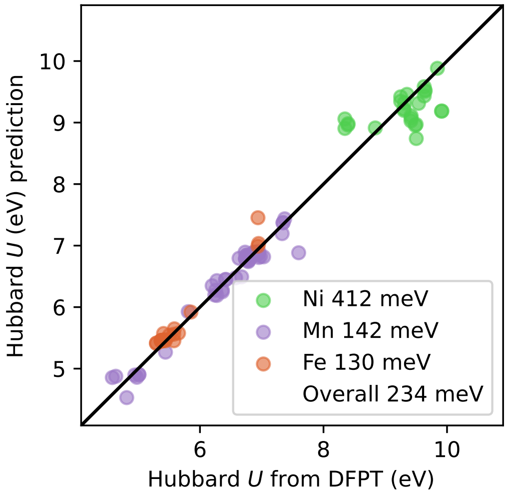
20:80 validate:train split

Hubbard U training

Structure type	Chemical composition	#
Olivine	Li_xFePO_4	5
	Li_xMnPO_4	5
	$Li_xFe_{0.5}Mn_{0.5}PO_4$	5
Spinel	$Li_xMn_2O_4$	2
	$Li_xMn_{1.5}Ni_{0.5}O_4$	2
Layered	Li_xNiO_2	2
	Li_xMnO_2	2
Tunnel	α-MnO_2	1
Rutile	β-MnO_2	1
Perovskite	$YNiO_3$	1
	$PrNiO_3$	1
Total		27

Element	Number of data points U	Number of data points V
Ni	396 (124)	67,802 (63,332)
Mn	856 (284)	162,272 (153,232)
Fe	138 (120)	22,511 (21,483)
Total	1,390 (528)	252,585 (238,047)

20:80 validate:train split

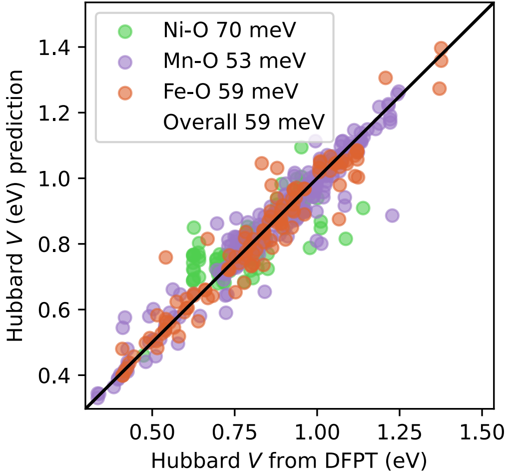


Hubbard U training

Structure type	Chemical composition	#
Olivine	Li_xFePO_4	5
	Li_xMnPO_4	5
	$Li_xFe_{0.5}Mn_{0.5}PO_4$	5
Spinel	$Li_xMn_2O_4$	2
	$Li_xMn_{1.5}Ni_{0.5}O_4$	2
Layered	Li_xNiO_2	2
	Li_xMnO_2	2
Tunnel	α-MnO_2	1
Rutile	β-MnO_2	1
Perovskite	$YNiO_3$	1
	$PrNiO_3$	1
Total		27

Element	Number of data points U	Number of data points V
Ni	396 (124)	67,802 (63,332)
Mn	856 (284)	162,272 (153,232)
Fe	138 (120)	22,511 (21,483)
Total	1,390 (528)	252,585 (238,047)

20:80 validate:train split



Model validation

Differences in properties between ML and DFT

x	Property	Li_xMnPO_4	Li_xFePO_4
$0-1$	$\Delta\Phi$	-0.26%	0.00%
0	Δm	-0.05%	0.01%
3	Δm	0.01%	0.01%

open-circuit voltages Φ (in V)
magnetic moments for TM elements (in μ_{B})

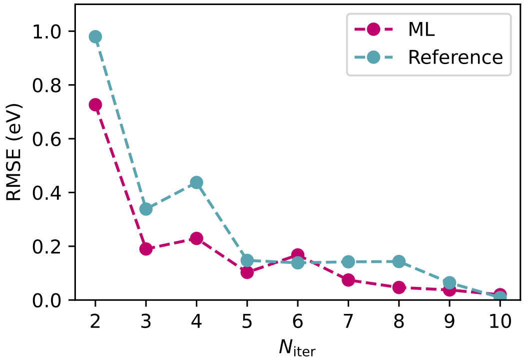
Model validation

Differences in properties between ML and DFT

x	Property	Li_xMnPO_4	Li_xFePO_4
$0-1$	$\Delta\Phi$	-0.26%	0.00%
0	Δm	-0.05%	0.01%
3	Δm	0.01%	0.01%

open-circuit voltages Φ (in V)
magnetic moments for TM elements (in μ_B)

Accuracy as function of num. training iterations



Different occupations - Hubbard U

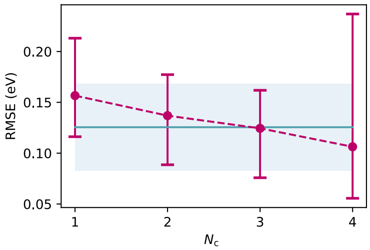
Mn	16	32	16	16	16
Fe	16	20	16	8	12
	0	0.25	0.5	0.75	1

x in $Li_x\{Fe, Mn, Fe_{0.5}Mn_{0.5}\}PO_4$

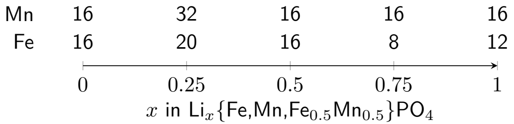
Different occupations - Hubbard U

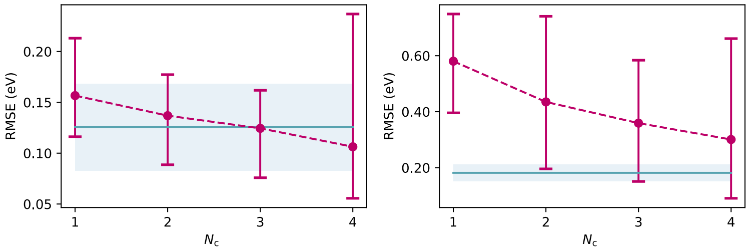
Mn	16	32	16	16	16
Fe	16	20	16	8	12
x	0	0.25	0.5	0.75	1

x in $Li_x\{Fe,Mn,Fe_{0.5}Mn_{0.5}\}PO_4$



Different occupations - Hubbard U

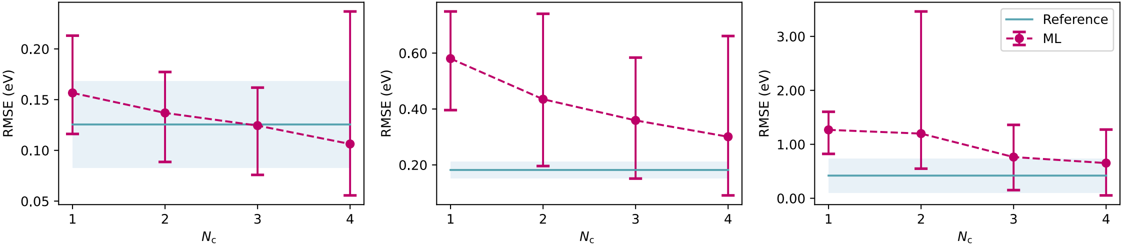




Different occupations - Hubbard U

Mn	16	32	16	16	16
Fe	16	20	16	8	12
	0	0.25	0.5	0.75	1

x in $Li_x\{Fe,Mn,Fe_{0.5}Mn_{0.5}\}PO_4$



Good transferability across different lithium concentrations.

Summary

Conclusion

Euclidean neural networks naturally encode the assumption that atomic systems exist in 3D Euclidean space

This inductive bias can make them more robust, data-efficient and accurate when predicting properties

GNNs give us a flexible way to express model inputs and outputs and perform multi-target learning

Conclusion

Euclidean neural networks naturally encode the assumption that atomic systems exist in 3D Euclidean space

This inductive bias can make them more robust, data-efficient and accurate when predicting properties

GNNs give us a flexible way to express model inputs and outputs and perform multi-target learning

Hubbard parameters

- Can accurately converged Hubbard parameters with error of 3-7%
- Good data efficiency: tens of HP calls sufficient
- Models show good transferability across:
  - Different stoichiometries (two example concentrations are enough)
  - Different self-consistent iterations (one is enough)
  - Different atomic environments

Uhrin et at., “Machine learning Hubbard parameters with equivariant neural networks” arXiv:2406.02457

Acknowledgements

Nicola Marzari
EPFL

Tess Smidt
MIT

Mario Geiger
Nvidia

Swiss National Science Foundation

MARVEL
NATIONAL CENTRE OF COMPETENCE IN RESEARCH

DOME 4.0

MIAI Grenoble Alpes
Multidisciplinary Institute In Artificial Intelligence