CORMORANT:
COvaRiant MOleculaR Artificial Neural neTworks

[BMA, Hy & Kondor, NeurIPS 2019, arXiv:1906.04015]

N-Atom interactions

$$V(\mathbf{r}_1, \ldots, \mathbf{r}_N) = \underbrace{\sum_i u_1(\mathbf{r}_i)}_{\text{one-body}} + \underbrace{\sum_{i<j} u_2(\mathbf{r}_i, \mathbf{r}_j)}_{\text{two-body}} + \underbrace{\sum_{i<j<k} u_3(\mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_k)}_{\text{three-body}} + \ldots$$

- Goal: Build general architecture for learning physics
  - Exploits symmetries of problem
  - Exploits hierarchy of scales
  - Based upon intuition of N-Atom interactions

[Stillinger, Sakai, and Torquato, JCP (2002)]

Inspiration: Multipoles expansion

$$V = \frac{1}{4\pi\epsilon_0} \sum_{i,j} \frac{q_i q_j}{|\mathbf{r}_i - \mathbf{r}_j|}$$

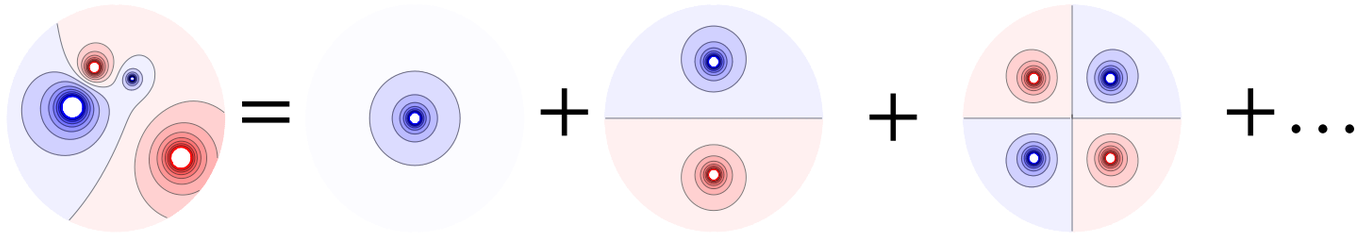
Multiple moments “one-body”:

$$Q_a^\ell = \sum_{i \in N(a)} q_i |\mathbf{r}_{ia}|^\ell Y^\ell(\hat{\mathbf{r}}_{ia})$$

$$V(\mathbf{R}) \sim Q^0 \cdot Y^0(\mathbf{R}) + Q^1 \cdot Y^1(\mathbf{R}) + Q^2 \cdot Y^2(\mathbf{R}) + \ldots$$

Under rotations:

$$Q^\ell \to D^\ell(\mathbf{R}) Q^\ell$$



Multiple energy “two-body” interaction

$$V = \frac{1}{4\pi\epsilon_0} \sum_{a,b} \sum_{\ell_1,\ell_2} \frac{Y^{\ell_1+\ell_2}(\mathbf{r}_{ab}) C_{\ell_1,\ell_2,\ell_1+\ell_2} Q_a^{\ell_1} \otimes Q_b^{\ell_2}}{|\mathbf{r}_{ab}|^{\ell_1+\ell_2+1}}$$

SO(3) Vectors: Generalized multipoles

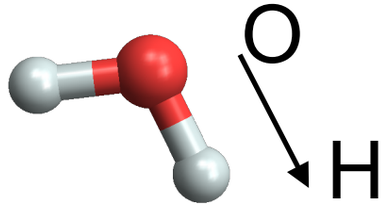
- Irreducible representation “irrep”:

$$Q_i^\ell \to F_i^\ell \in \mathbb{C}^{2\ell+1,\tau_\ell}$$

Multiplicity (number of copies of irrep): τ_ℓ

Transforms covariantly as: $F^\ell \to D^\ell(\mathbf{R})F^\ell$

Generalizes multipoles, e.g., O-H dipole/quadrupole moment



- Representation (list of irreps):

$$F_i = \bigoplus_\ell F_i^\ell$$
$$= \left(F_i^0, F_i^1, \ldots, F_i^{\ell_{\max}}\right)$$

- Multiplicity of each SO(3) Vector:

$$\tau = \left(\tau_0, \tau_1, \ldots, \tau_{\ell_{\max}}\right)$$

Covariant operations on SO(3) Vectors

- Clebsch-Gordan decomposition of two representations:

$$F_1 \otimes_{\text{cg}} F_2 = \bigoplus_{c} \bigoplus_{\ell=|\ell_1-\ell_2|}^{\ell_1+\ell_2} C_{\ell_1,\ell_2,\ell} \cdot \left(F_{1,c}^{\ell_1} \otimes F_{2,c}^{\ell_2} \right)$$

  - Multiplicity is treated as “channels”: $\tau = (c, c, \ldots, c)$

- SO(3) vectors can be concatenated:

$$F_1 \oplus F_2 = \bigoplus_{\ell} (F_1^{\ell} \oplus F_2^{\ell})$$

- SO(3) vectors can be mixed while preserving covariance

$$F \leftarrow \bigoplus_{\ell} F^{\ell} W^{\ell} \qquad W^{\ell} \in \mathbb{C}^{(\tau_{\text{out}},\tau_{\text{in}})}$$

Recipe for a Cormorant

- Goal: Build general architecture for learning covariant physics.

- Activations are SO(3)-Vectors F_i
  - Use CG, mixing, concatenation

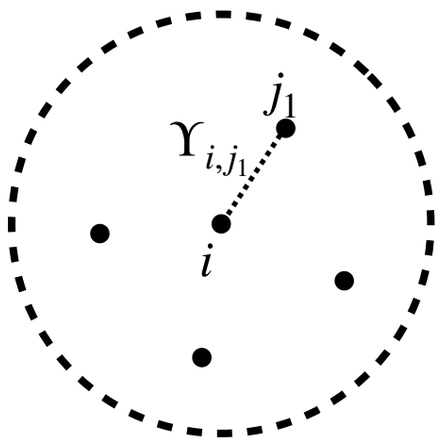
- Construct from general operations involving n=1, 2, 3, … Clebsch-Gordan operations.
  - We call these “interactions”

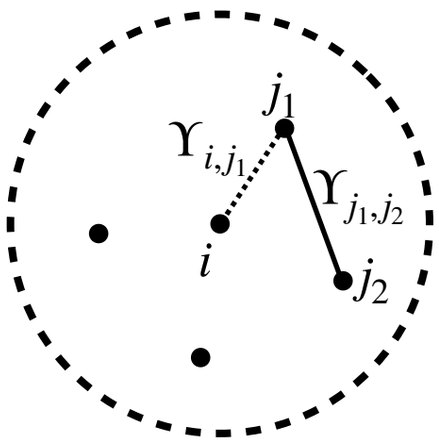
$$F_i^{s+1} = \left(\bigoplus_{n=1} \Phi_i^{(n)} \left(\left\{ F_j^s \right\} \right) \right) \cdot W$$

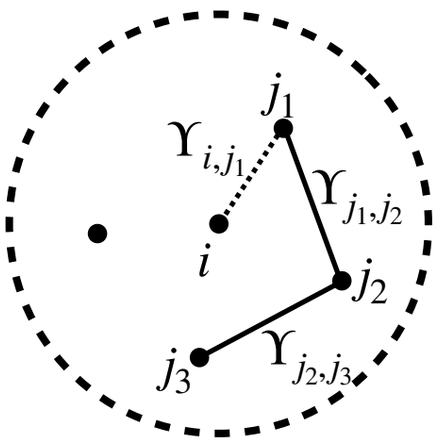
N-Atom interactions

- We base our Cormorant on *n*-atom interactions.
- Define interactions between an atom *i* and its neighbors $\{F_j\}$
- Must be permutation invariant for $j_k \neq i$
- Allow for a (learnable) transition amplitude $\Upsilon^{(n)}_{j_k j_{k+1}}$

$$\Phi^{(n)}_i\left(\left\{F_j\right\}\right) = \bigoplus_{\substack{j_1,\ldots,j_n = i \\ j_1,\ldots,j_n \in \mathcal{S}_i}} \sum_{\substack{j_1,\ldots,j_n \neq i \\ j_1,\ldots,j_n \in \mathcal{S}_i}} \bigotimes_{k=0}^{n-1} \left(\Upsilon^{(n)}_{j_k j_{k+1}} \otimes F_{j_{k+1}}\right)$$







Cormorant architecture (simplified)

Aggregation (one-body):

$$F_i^{(\mathrm{ag})} \leftarrow \sum_j \Upsilon_{ij} \otimes_{\mathrm{cg}} F_j$$

Identity (one-body):

$$F_i^{(\mathrm{id})} \leftarrow F_i$$

Non-linearity (two-body):

$$F_i^{(\mathrm{cg})} \leftarrow F_i \otimes_{\mathrm{cg}} F_i$$

Concatenation and mixing:

$$F_i^{(\mathrm{out})} \leftarrow (F_i^{(\mathrm{ag})} \oplus F_i^{(\mathrm{id})} \oplus F_i^{(\mathrm{cg})}) \cdot W$$

Network: Iterate this several times

Experimental results: MD-17 and QM-9

Table 1: Mean absolute error of various prediction targets on QM-9 (left) and conformational energies (in units of kcal/mol) on MD-17 (right). The best results within a standard deviation of three Cormorant training runs (in parenthesis) are indicated in bold.

	Cormorant		SchNet	NMP	WaveScatt
α (bohr^3)	**0.085**	(0.001)	0.235	0.092	0.160
$\Delta\epsilon$ (eV)	**0.061**	(0.005)	**0.063**	0.069	0.118
ϵ_{HOMO} (eV)	**0.034**	(0.002)	0.041	0.043	0.085
ϵ_{LUMO} (eV)	**0.038**	(0.008)	**0.034**	**0.038**	0.076
μ (D)	**0.038**	(0.009)	**0.033**	**0.030**	0.340
C_ν (cal/mol K)	**0.026**	(0.000)	0.033	0.040	0.049
G (eV)	0.020	(0.000)	**0.014**	0.019	0.022
H (eV)	0.021	(0.001)	**0.014**	0.017	0.022
R^2 (bohr^2)	0.961	(0.019)	**0.073**	0.180	0.410
U (eV)	0.021	(0.000)	**0.019**	0.020	0.022
U_0 (eV)	0.022	(0.003)	**0.014**	0.020	0.022
ZPVE (meV)	2.027	(0.042)	1.700	**1.500**	2.000

	Cormorant	DeepMD	DTNN	SchNet	GDML	sGDML
Aspirin	**0.098**	0.201	–	0.120	0.270	0.190
Benzene	**0.023**	0.065	0.040	0.070	0.070	0.100
Ethanol	**0.027**	0.055	–	0.050	0.150	0.070
Malonaldehyde	**0.041**	0.092	0.190	0.080	0.160	0.100
Naphthalene	**0.029**	0.095	–	0.110	0.120	0.120
Salicylic Acid	**0.066**	0.106	0.410	0.100	0.120	0.120
Toluene	**0.034**	0.085	0.180	0.090	0.120	0.100
Uracil	**0.023**	0.085	–	0.100	0.110	0.110

SchNet: [Schütt, et al, 2017]

NMP: [Gilmer, et al, 2017]

WaveScatt: [Eickenberg, et al, 2018]

DeepMD: [Zhang, et al, 2017]

DTNN: [Schütt, et al, 2017]

GDML: [Chmiela, et al, 2017]

sGDML: [Chmiela, et al, 2018]