Becoming an “Atomic Architect”

From Cal to Google to LBL

Tess Smidt

Berkeley Lab Computing Sciences

Computational Materials, Chemistry, and Climate Group

2018.10.23

Becoming an “Atomic Architect”

From Cal to Google to LBL

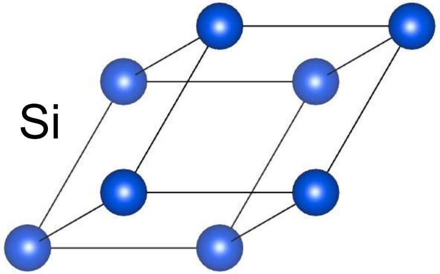
Overview of my research in deep learning
How did I get here?
- **My PhD in a nutshell**
- **Intern at Google for 1 year**
- **Choosing between Google and LBL**

Q&A -- Ask me anything!

Tess Smidt
Berkeley Lab Computing Sciences
Computational Materials, Chemistry, and Climate Group
2018.10.23

What a computational materials physicist does:

Given an atomic structure,

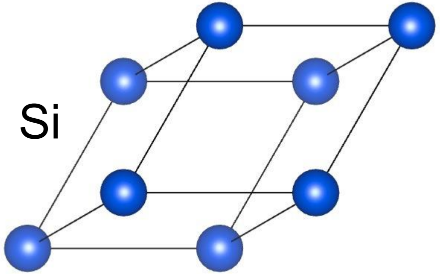


What a computational materials physicist does:

...use quantum theory and supercomputers to determine...

Given an atomic structure,

Si



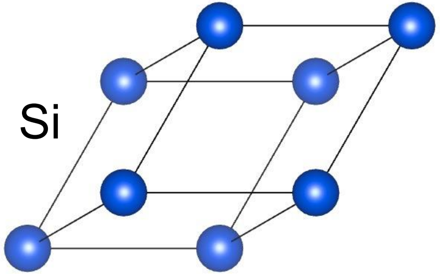
$$\hat{H}\,|\psi\rangle = E\,|\psi\rangle$$

NERSC
National Energy Research Scientific Computing Center

What a computational materials physicist does:

Given an atomic structure,

Si



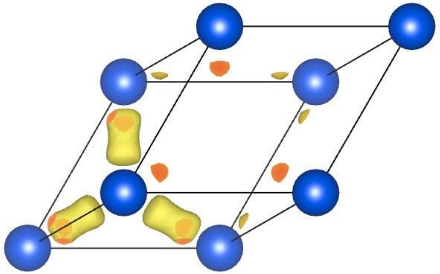
...use quantum theory and supercomputers to determine...

$$\hat{H}\,|\psi\rangle = E\,|\psi\rangle$$

NERSC

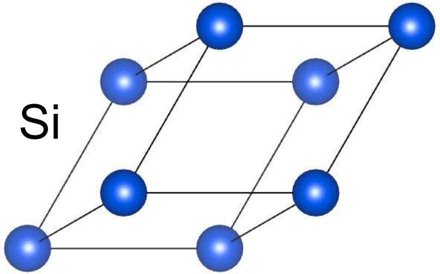
National Energy Research Scientific Computing Center

...where the electrons are...



What a computational materials physicist does:

Given an atomic structure,



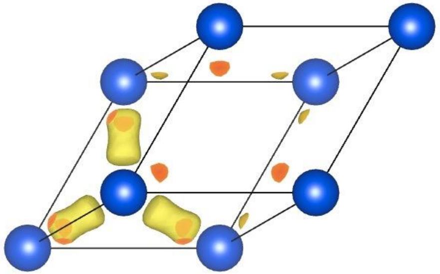
...use quantum theory and supercomputers to determine...

$$\hat{H}\,|\psi\rangle = E\,|\psi\rangle$$

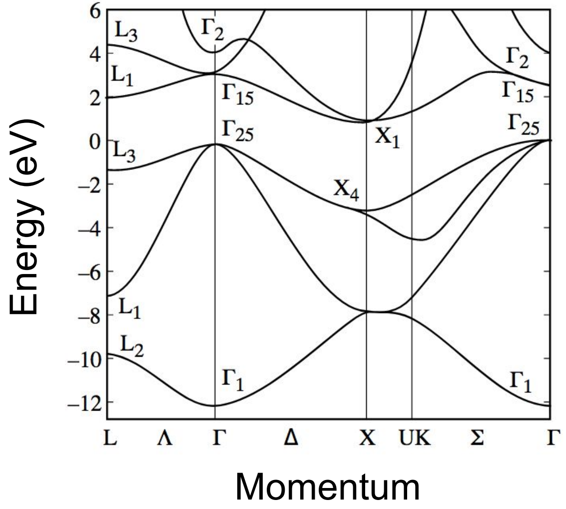
NERSC

National Energy Research Scientific Computing Center

...where the electrons are...



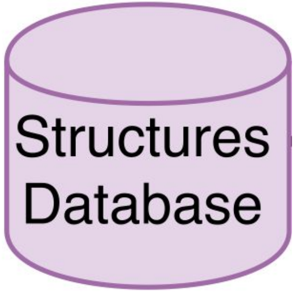
...and what the electrons are doing.



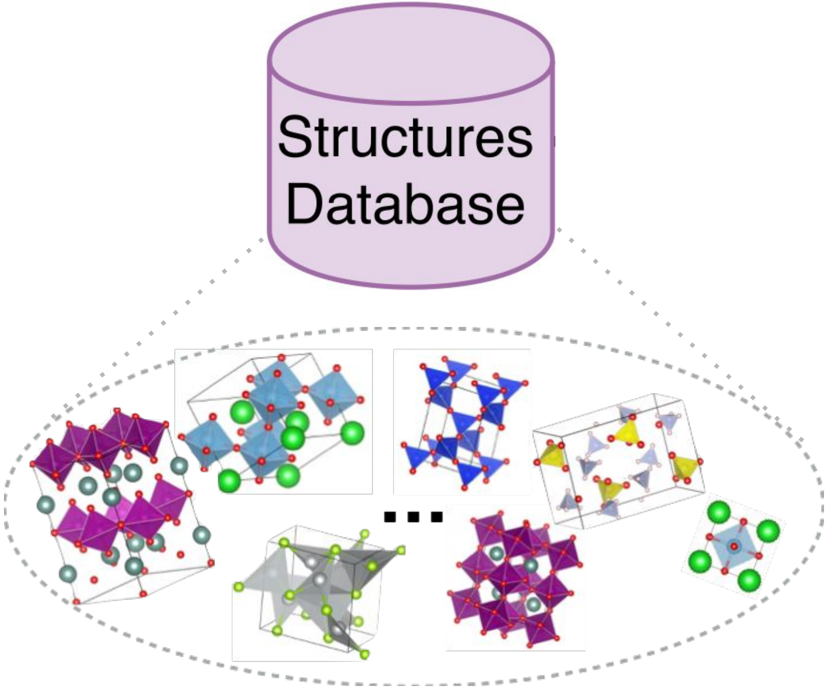
http://www.eecs.umich.edu/courses/eecs320/f00/bk7ch03.pdf

Workflows are automated recipes that encode best practices for calculating materials properties. We use them to screen materials for specific properties and applications.

Workflows are automated recipes that encode best practices for calculating materials properties. We use them to screen materials for specific properties and applications.

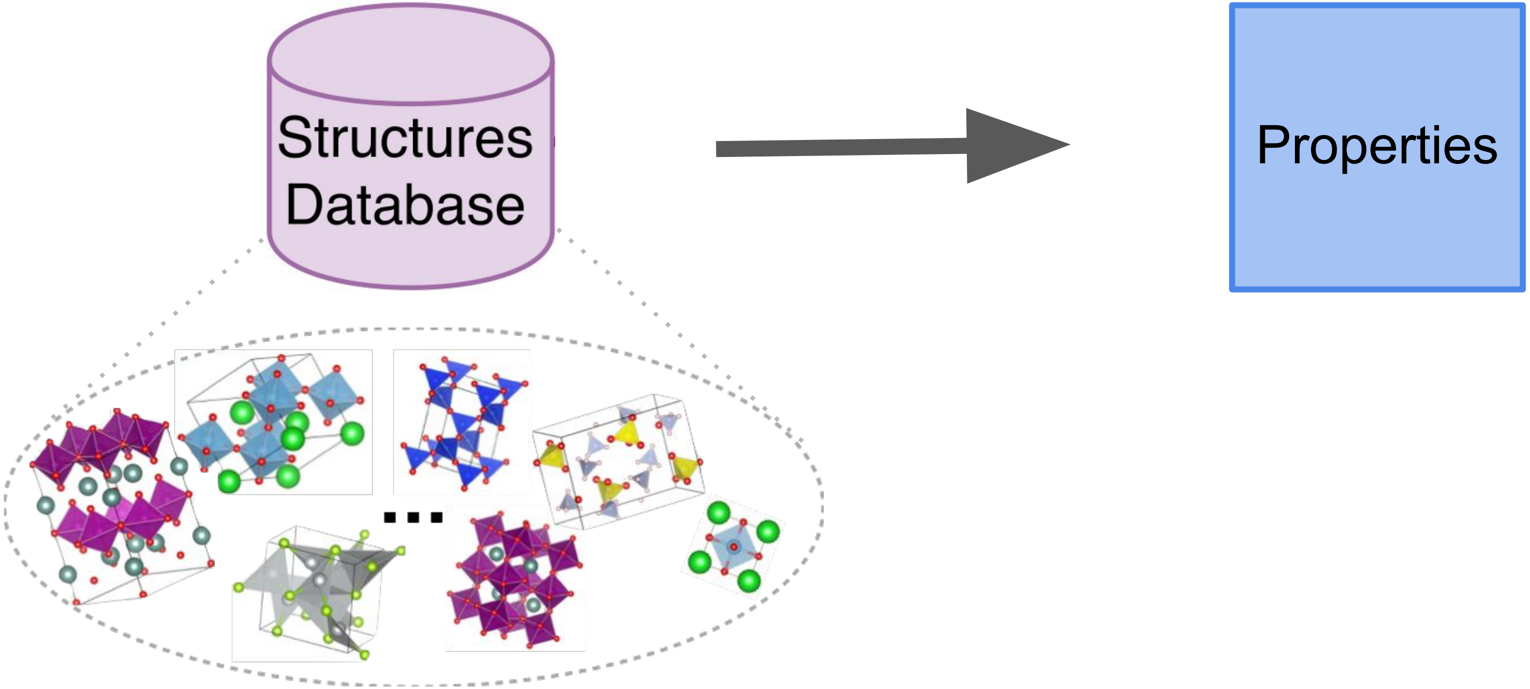


Workflows are automated recipes that encode best practices for calculating materials properties. We use them to screen materials for specific properties and applications.



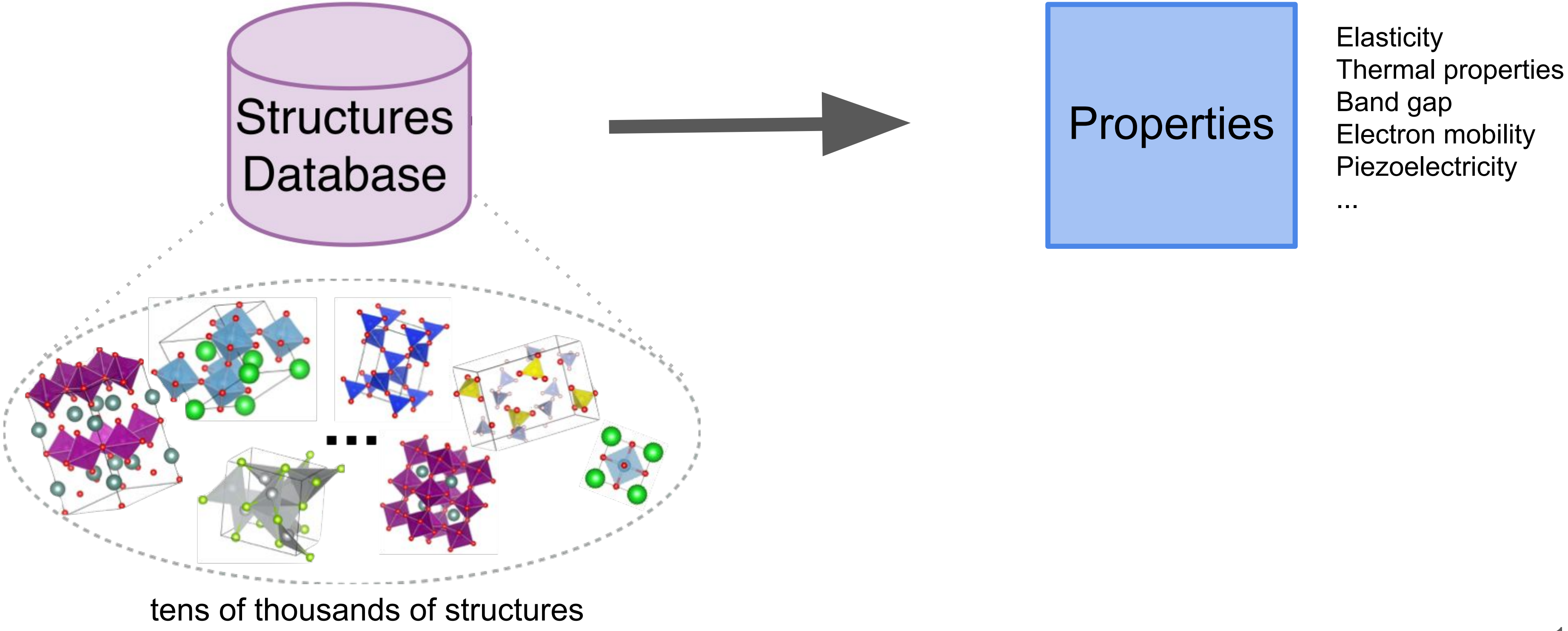
tens of thousands of structures

Workflows are automated recipes that encode best practices for calculating materials properties. We use them to screen materials for specific properties and applications.



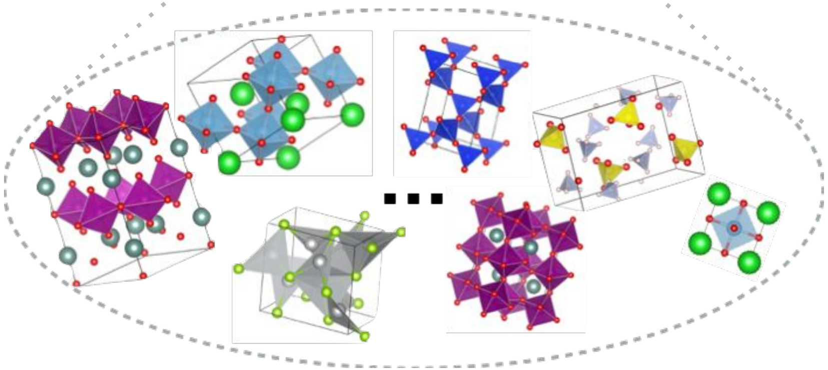
tens of thousands of structures

Workflows are automated recipes that encode best practices for calculating materials properties. We use them to screen materials for specific properties and applications.



Workflows are automated recipes that encode best practices for calculating materials properties. We use them to screen materials for specific properties and applications.

Structures Database



tens of thousands of structures

Properties

Elasticity
Thermal properties
Band gap
Electron mobility
Piezoelectricity
...

Photovoltaics

Magnetic materials

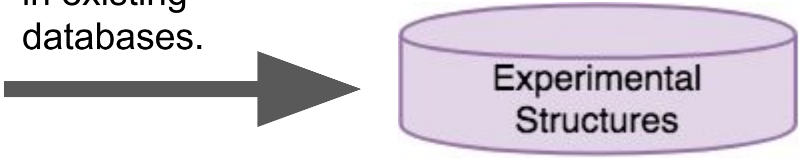
Batteries

Water-splitting catalysts

However, screening is bottlenecked limited by our ability to propose hypothetical atomic structures.

However, screening is bottlenecked limited by our ability to propose hypothetical atomic structures.

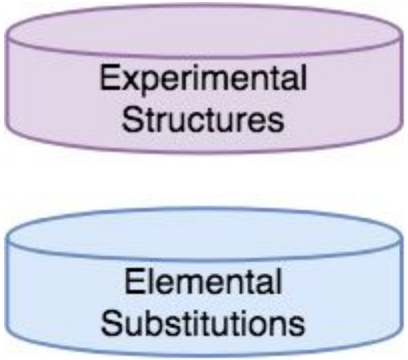
Materials
in existing
databases.



However, screening is bottlenecked limited by our ability to propose hypothetical atomic structures.

Materials
in existing
databases.

Small
modifications

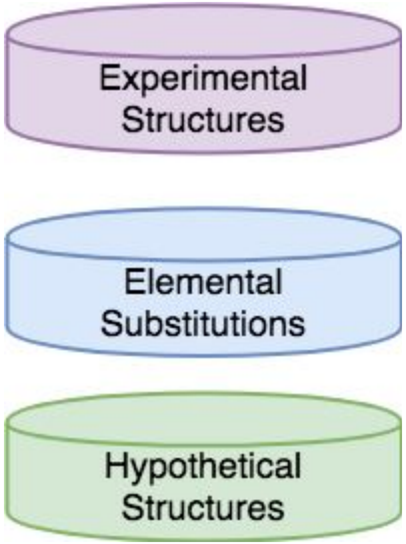


However, screening is bottlenecked limited by our ability to propose hypothetical atomic structures.

Materials
in existing
databases.

Small
modifications

***How do we
get these?***

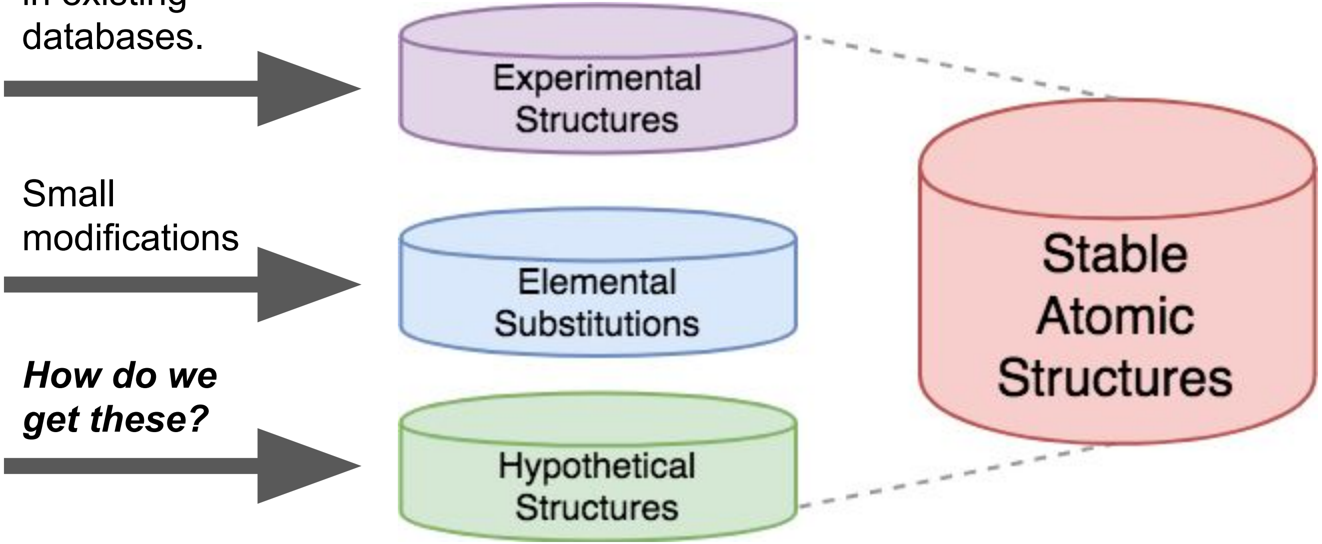


However, screening is bottlenecked limited by our ability to propose hypothetical atomic structures.

Materials in existing databases.

Small modifications

How do we get these?



Experimentalists are making new structures every day! These structures are not in existing databases.

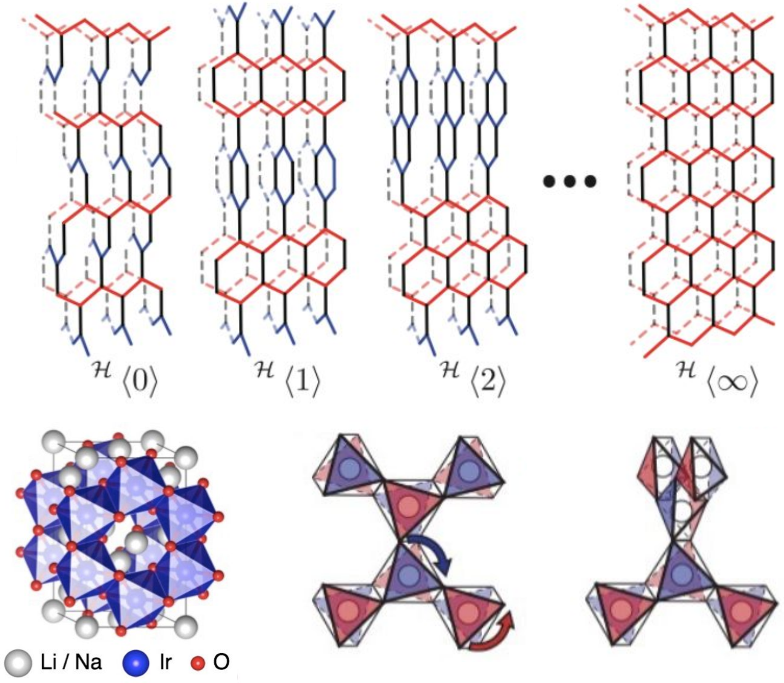
T. Smidt, S. Griffin, and J. B. Neaton, *Ab initio Studies of Structural and Energetic Trends in the Harmonic Honeycomb Iridates*, In preparation for submission to Physical Review: B (2018).

K. Modic, T. Smidt, I. Kimchi et al., *Realization of a three-dimensional spin-anisotropic harmonic honeycomb iridate*, Nature Communications 5 (2014). (arXiv:1402.3254)

J.N. Hohman, M. Collins, and T. Smidt, *Mithrene and methods of fabrication of mithrene*, (2017). International Patent App. PCT/US20l7/045609. Filed August 4, 2017.

Experimentalists are making new structures every day! These structures are not in existing databases.

Harmonic honeycomb iridates:
Frustrated quantum magnets



T. Smidt, S. Griffin, and J. B. Neaton, *Ab initio Studies of Structural and Energetic Trends in the Harmonic Honeycomb Iridates*, In preparation for submission to Physical Review: B (2018).

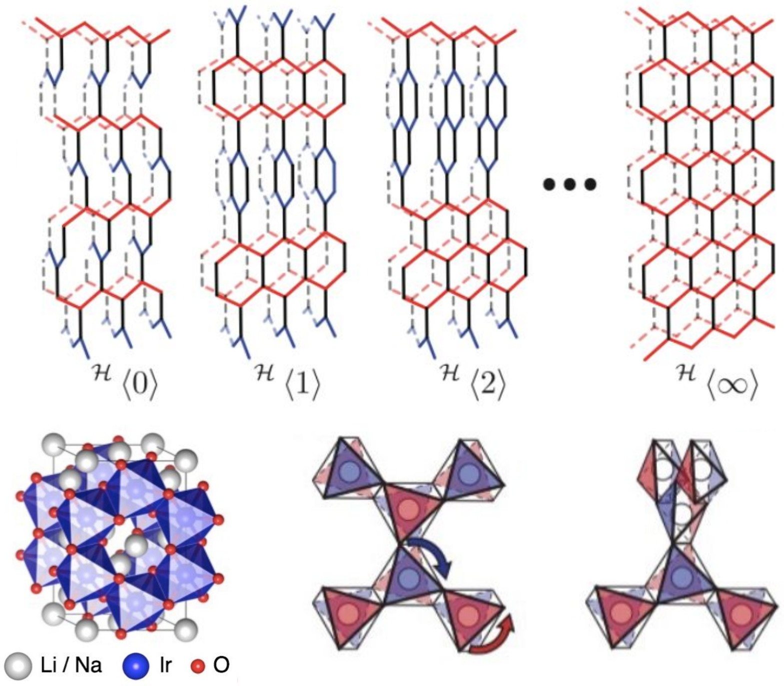
K. Modic, T. Smidt, I. Kimchi et al., *Realization of a three-dimensional spin-anisotropic harmonic honeycomb iridate*, Nature Communications 5 (2014). (arXiv:1402.3254)

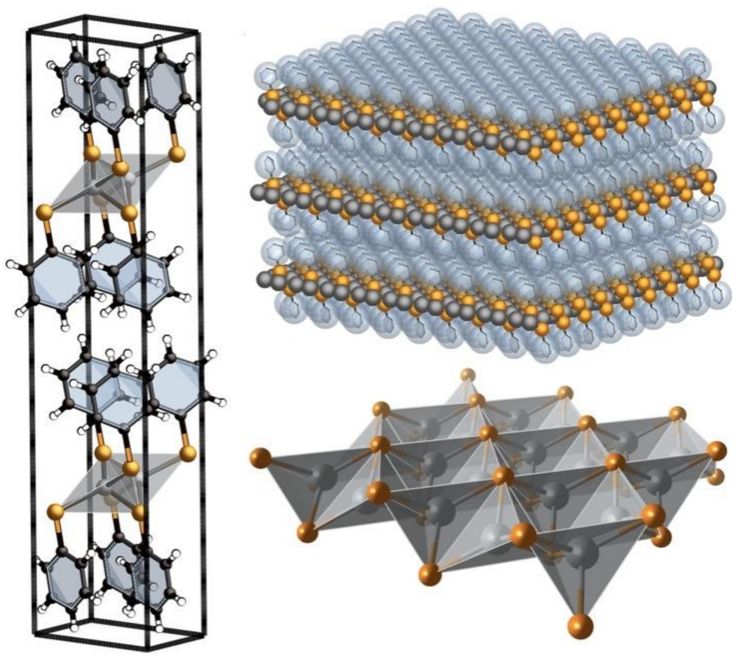
J.N. Hohman, M. Collins, and T. Smidt, *Mithrene and methods of fabrication of mithrene*, (2017). International Patent App. PCT/US20l7/045609. Filed August 4, 2017.

Experimentalists are making new structures every day! These structures are not in existing databases.

Harmonic honeycomb iridates:
Frustrated quantum magnets

Metal-organic chalcogenide assemblies (MOChAs):
2D electronic properties in a 3D crystal





T. Smidt, S. Griffin, and J. B. Neaton, *Ab initio Studies of Structural and Energetic Trends in the Harmonic Honeycomb Iridates*, In preparation for submission to Physical Review: B (2018).

K. Modic, T. Smidt, I. Kimchi et al., *Realization of a three-dimensional spin-anisotropic harmonic honeycomb iridate*, Nature Communications 5 (2014). (arXiv:1402.3254)

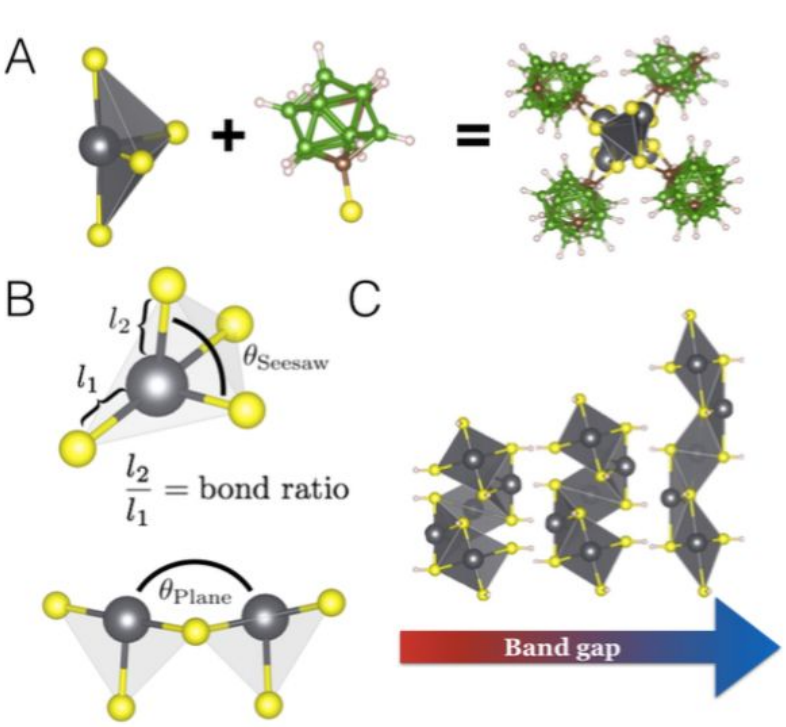
J.N. Hohman, M. Collins, and T. Smidt, *Mithrene and methods of fabrication of mithrene*, (2017). International Patent App. PCT/US20l7/045609. Filed August 4, 2017.

Materials are challenging to design because their 3D geometry and interactions are complex.

Ex: Hypothetical materials that I designed by hand (with parametric models).

Materials are challenging to design because their 3D geometry and interactions are complex.

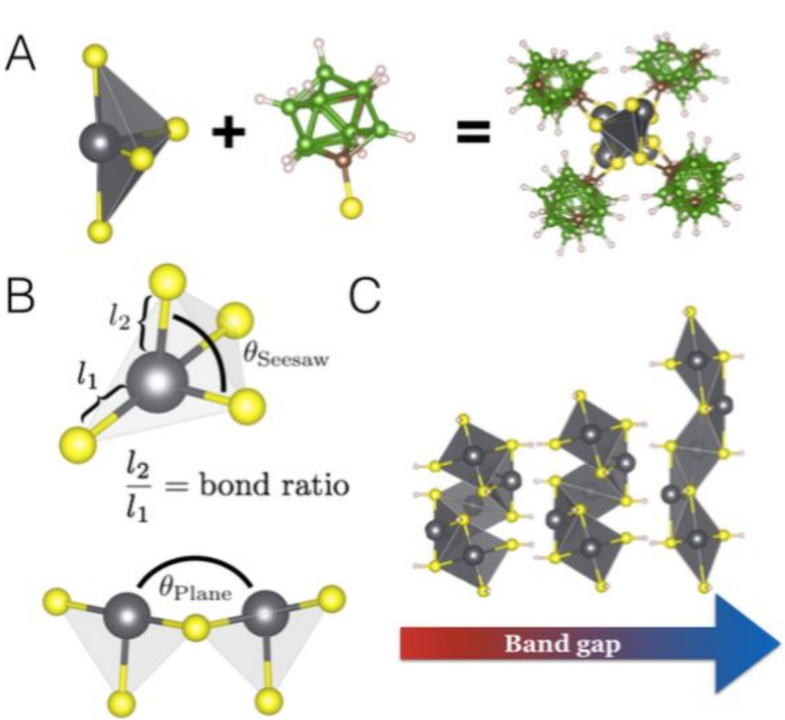
Ex: Hypothetical materials that I designed by hand (with parametric models).



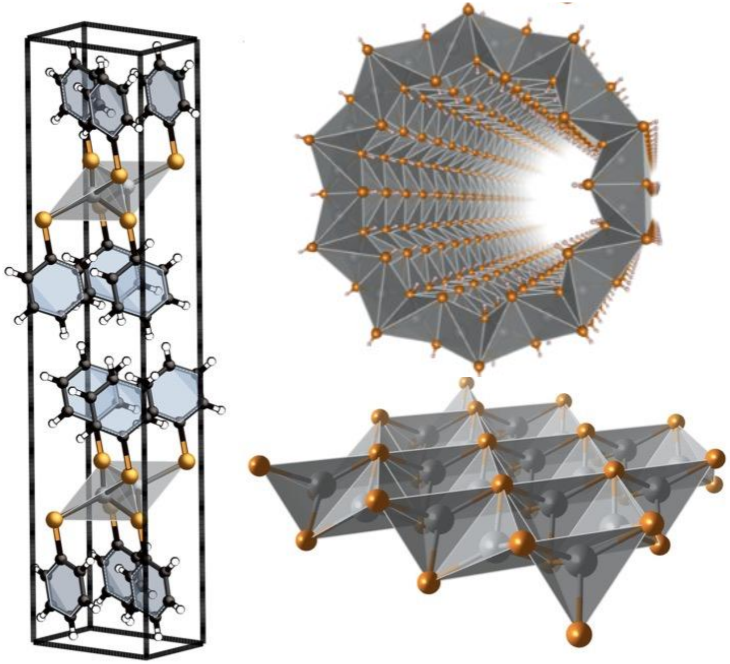
Distort subunits to tune properties.

Materials are challenging to design because their 3D geometry and interactions are complex.

Ex: Hypothetical materials that I designed by hand (with parametric models).



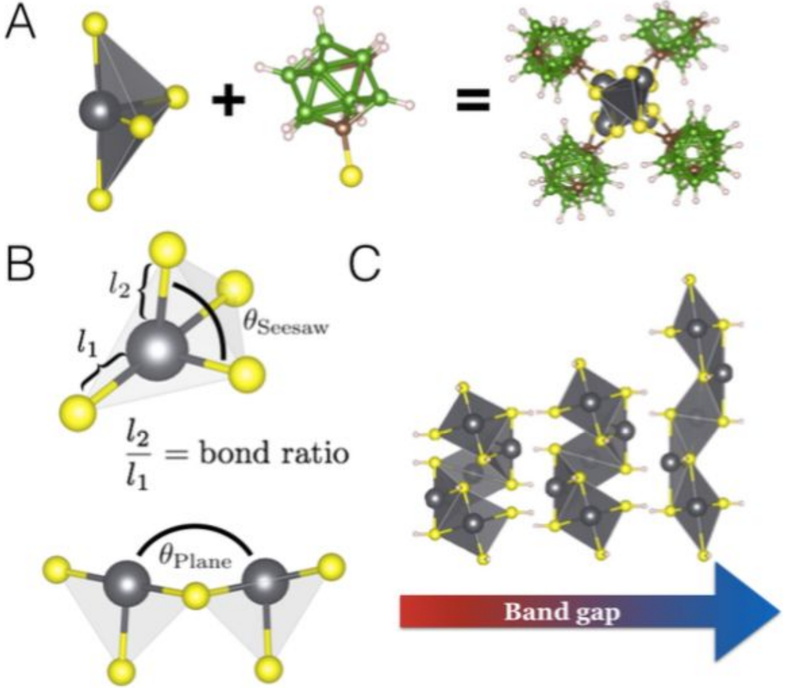
Distort subunits to tune properties.



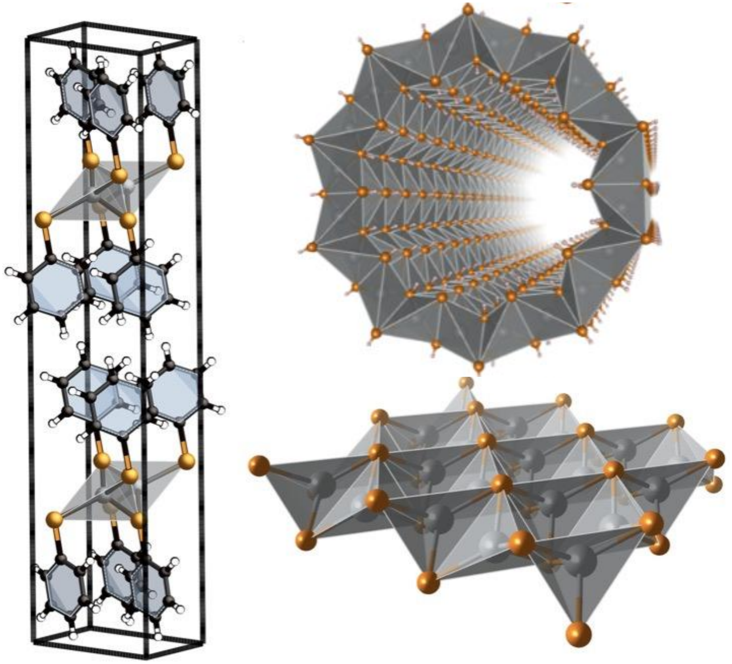
Produce new topologies that are chemically viable.

Materials are challenging to design because their 3D geometry and interactions are complex.

Ex: Hypothetical materials that I designed by hand (with parametric models).



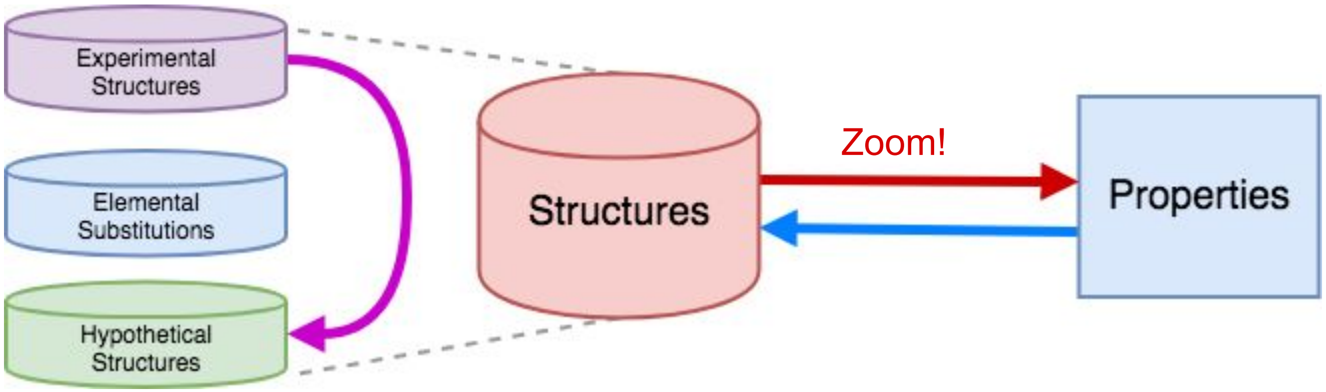
Distort subunits to tune properties.



Produce new topologies that are chemically viable.

We need better tools to systematically generate new hypothetical atomic structures.

Deep learning can help accelerate existing tools and create new capabilities for automating computational chemical and materials discovery.



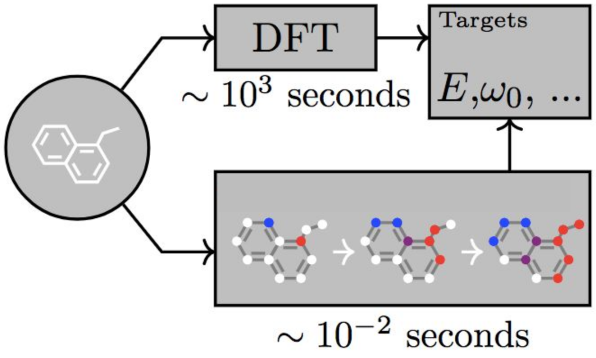
1) Help compute properties faster.

2) Generate hypothetical structures based on experimentally observed motifs.

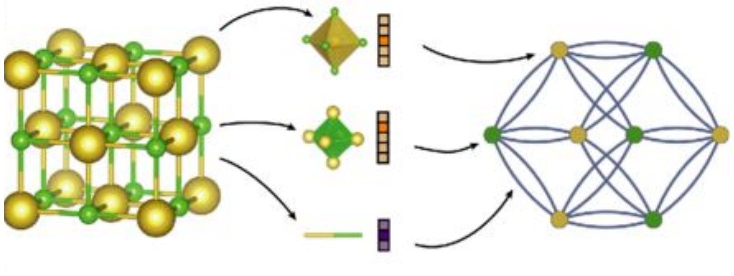
3) Generate structures with specific properties. [Need to do (2) first.]

Previous work

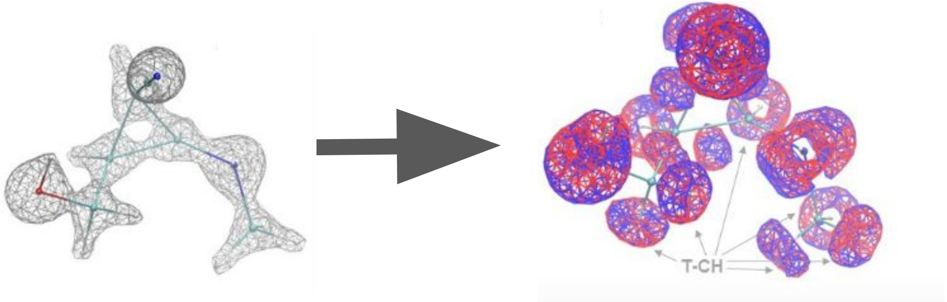
It is possible to train neural networks that can predict properties with the accuracy of quantum-mechanical calculations between 2 and 5 orders of magnitude faster.



J. Gilmer, S. S. Schoenholz, P. F. Riley, O. Vinyals, and G. E. Dahl, "Neural Message Passing for Quantum Chemistry." arXiv preprint arXiv:1704.01212 (2017).



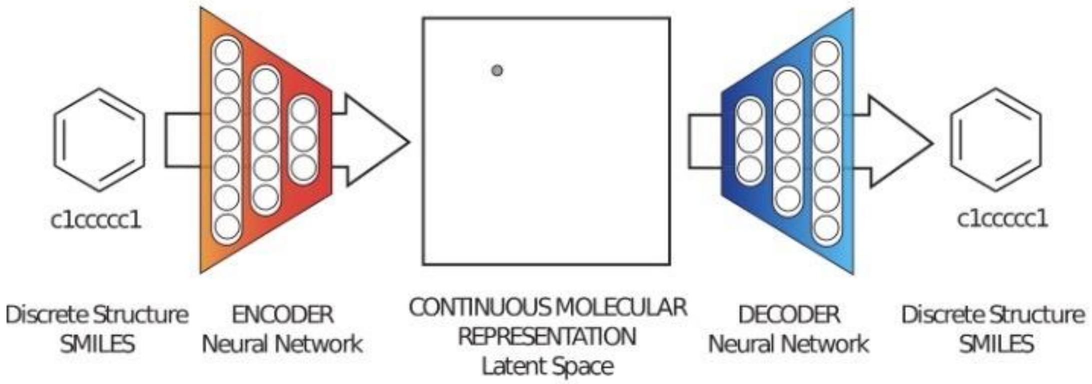
Tian Xie and Jeffrey C. Grossman. “Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties”, Phys. Rev. Lett. 120, 145301 (2018)



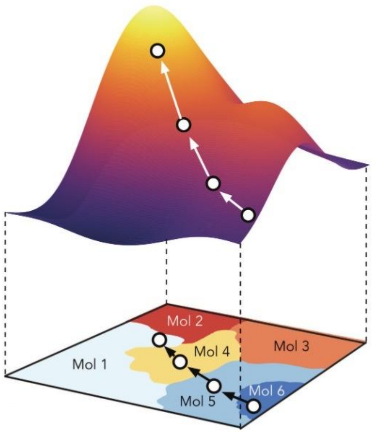
Anton V. Sinitskiy, Vijay S. Pande, “Deep Neural Network Computes Electron Densities and Energies of a Large Set of Organic Molecules Faster than Density Functional Theory (DFT)”, arXiv:1809.02723

Previous work

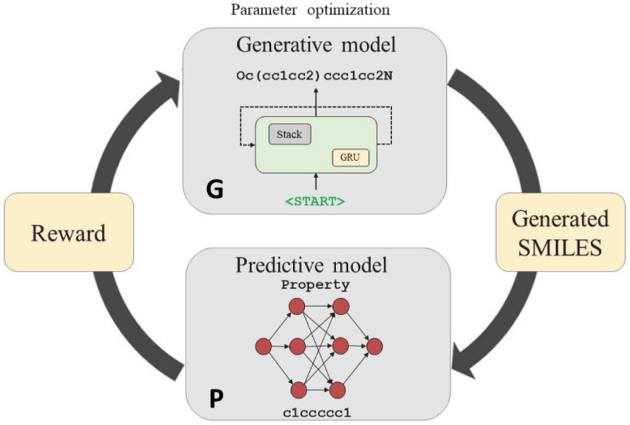
Deep learning techniques have also been used to generate new molecules for applications such as drugs and devices.



Properties can be optimized using learned continuous representation or reinforcement learning.



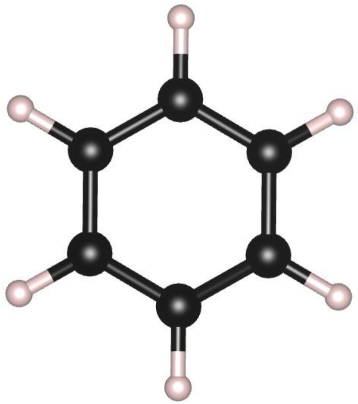
Gómez-Bombarelli, Rafael, et al. "Automatic chemical design using a data-driven continuous representation of molecules." ACS Cent. Sci., 4 (2), pp 268–276(2018)



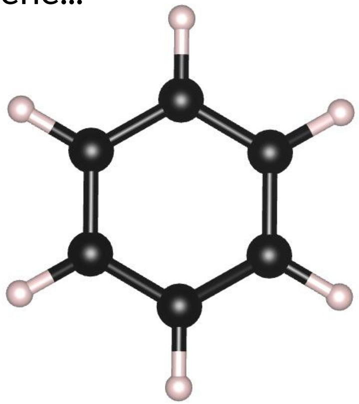
Mariya Popova, Olexandr Isayev, Alexander Tropsha, “Reinforcement learning for de novo drug design” Science Advances, Vol. 4, no. 7, eaap7885 (2018)

These examples used very different input representations. (strings, graphs, images)

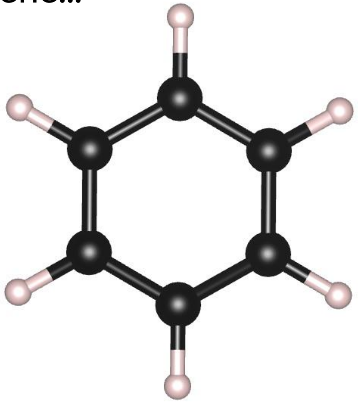
How to represent atomic systems to neural networks is an open question.



Many representations of benzene...

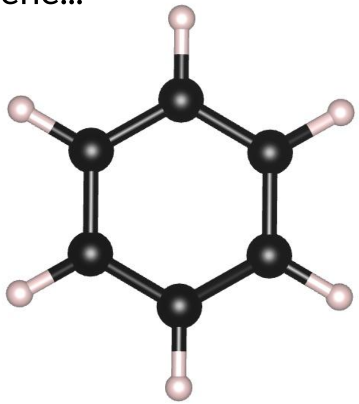


Many representations of benzene...



Vector (Fingerprint)

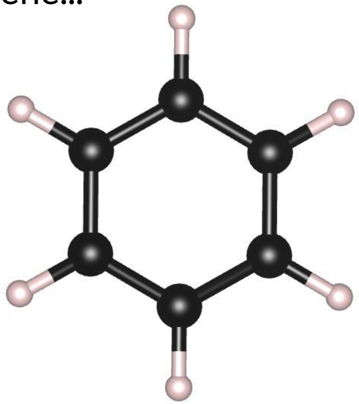
Many representations of benzene...



Vector (Fingerprint)

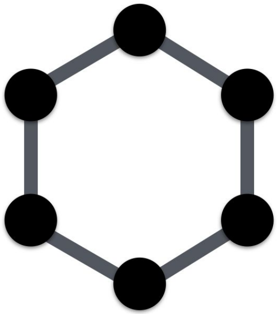
SMILES string:
C1=CC=CC=C1

Many representations of benzene...



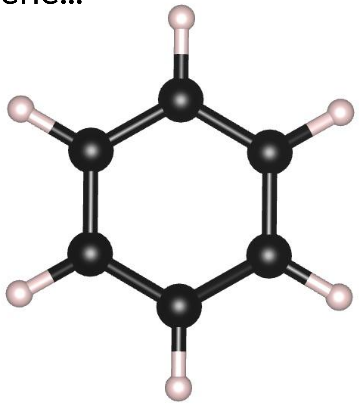
Vector (Fingerprint)

SMILES string:
C1=CC=CC=C1



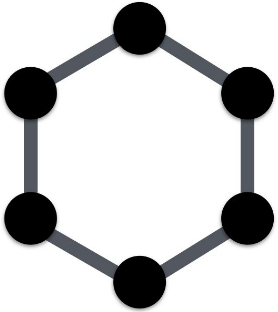
Graph of bonds

Many representations of benzene...



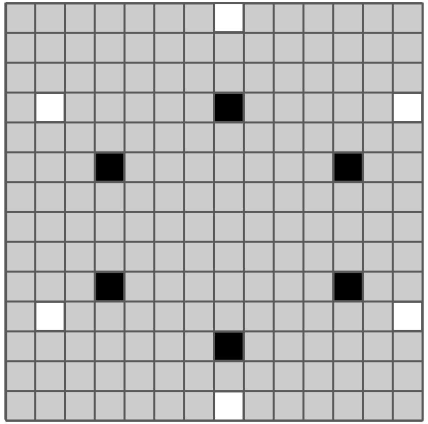
Vector (Fingerprint)

SMILES string:
C1=CC=CC=C1

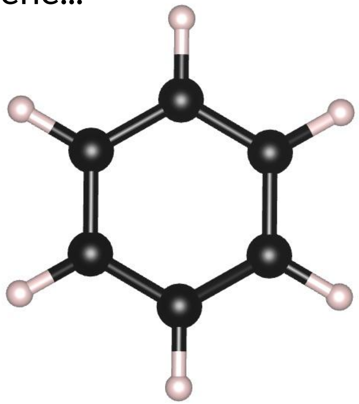


Graph of bonds

Image

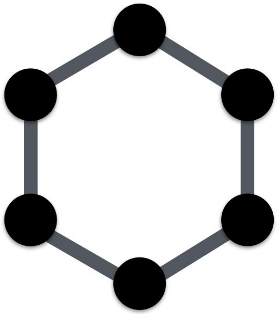


Many representations of benzene...



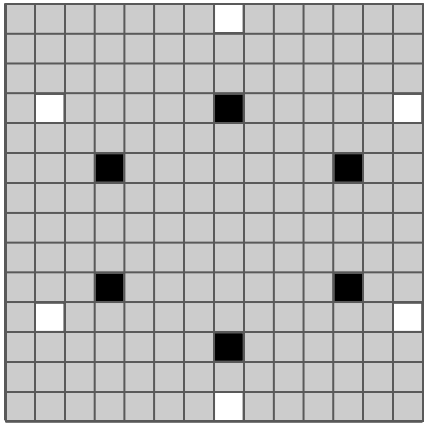
Vector (Fingerprint)

SMILES string:
C1=CC=CC=C1



Graph of bonds

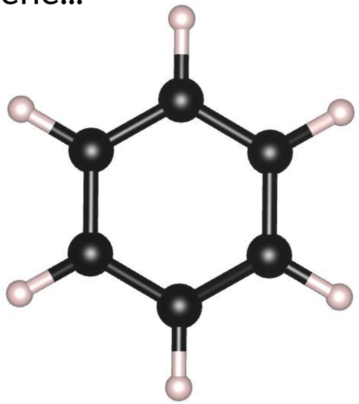
Image



3D Coordinates

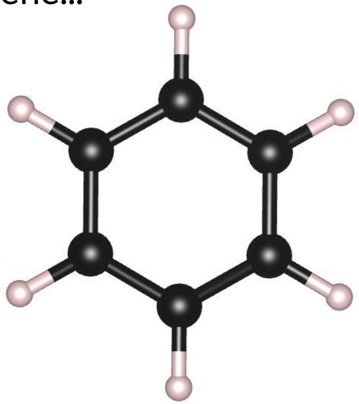
H	-0.21463	0.97837	0.33136
C	-0.38325	0.66317	-0.70334
C	-1.57552	0.03829	-1.05450
H	-2.34514	-0.13834	-0.29630
C	-1.78983	-0.36233	-2.36935
H	-2.72799	-0.85413	-2.64566
C	-0.81200	-0.13809	-3.33310
H	-0.98066	-0.45335	-4.36774
C	0.38026	0.48673	-2.98192
H	1.14976	0.66307	-3.74025
C	0.59460	0.88737	-1.66708
H	1.53276	1.37906	-1.39070

Many representations of benzene...



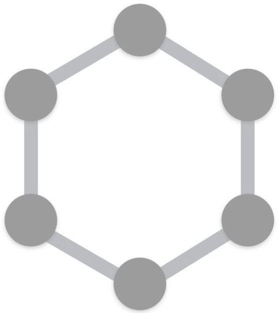
	Bonding	Geometry	Memory Efficient	Universality
Fingerprints	?	?	✓	?
SMILES	✓	X	✓	X
Graphs	✓	?	?	?
Images	X	✓	X	✓
Coordinates	X	✓	✓	✓

Many representations of benzene...

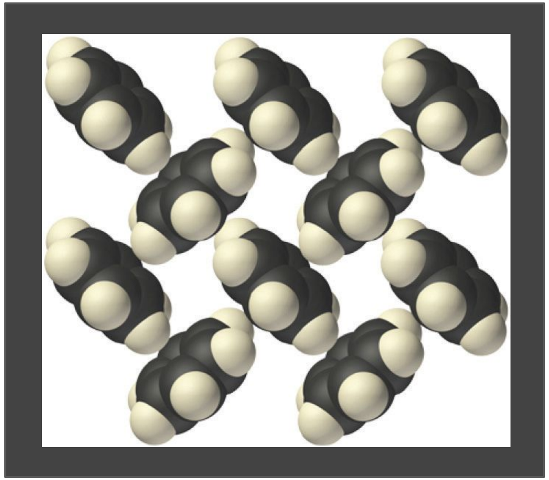


Vector (Fingerprint)

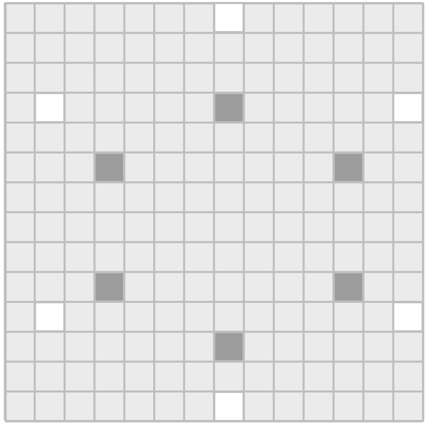
SMILES string:
C1=CC=CC=C1



Graph of bonds



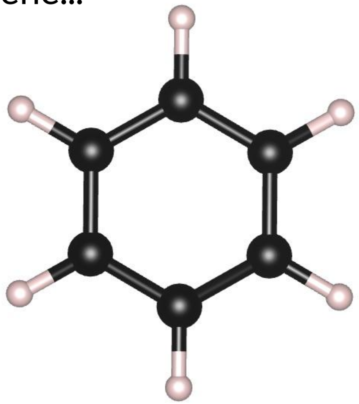
Image



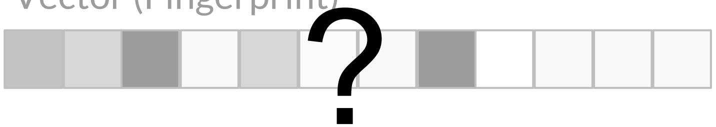
3D Coordinates

H	-0.21463	0.97837	0.33136
C	-0.38325	0.66317	-0.70334
C	-1.57552	0.03829	-1.05450
H	-2.34514	-0.13834	-0.29630
C	-1.78983	-0.36233	-2.36935
H	-2.72799	-0.85413	-2.64566
C	-0.81200	-0.13809	-3.33310
H	-0.98066	-0.45335	-4.36774
C	0.38026	0.48673	-2.98192
H	1.14976	0.66307	-3.74025
C	0.59460	0.88737	-1.66708
H	1.53276	1.37906	-1.39070

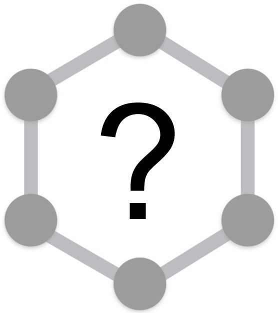
Many representations of benzene...



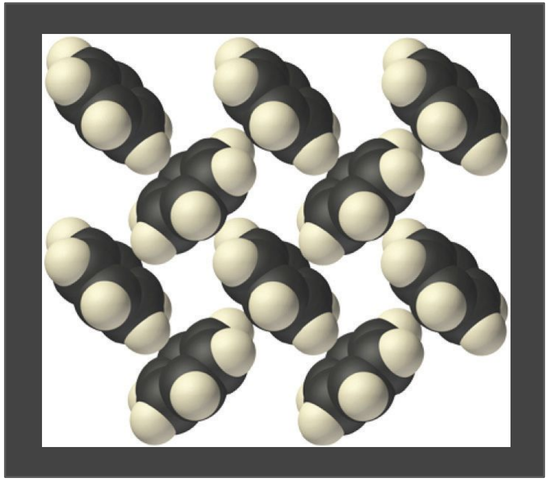
Vector (Fingerprint)



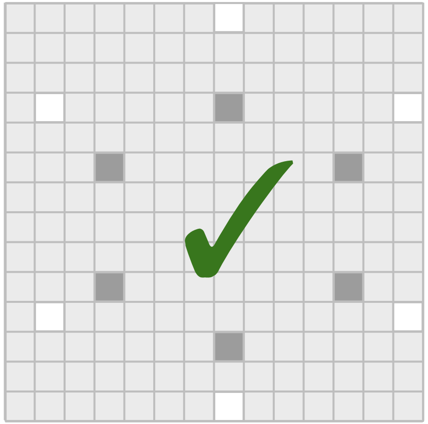
SMILES string:
C1=CC=CC=C1



Graph of bonds



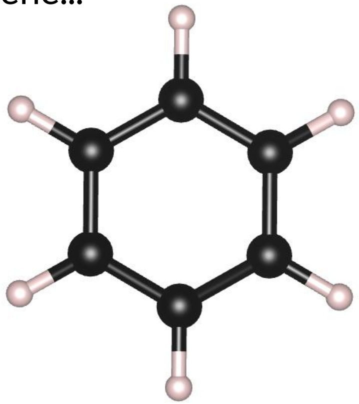
Image



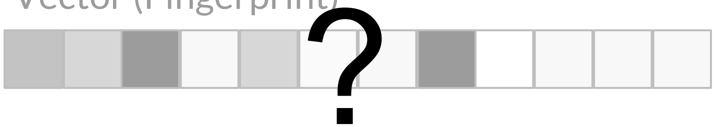
3D Coordinates

H	-0.21463	0.97837	0.33136
C	-0.38325	0.66317	-0.70334
C	-1.57552	0.03829	-1.05450
H	-2.34514	-0.13834	-0.29630
C	-1.78983	-0.36233	-2.36935
H	-2.72799	-0.85413	-2.64566
C	-0.81200	-0.13809	-3.33310
H	-0.98066	-0.45335	-4.36774
C	0.38026	0.48673	-2.98192
H	1.14976	0.66307	-3.74025
C	0.59460	0.88737	-1.66708
H	1.53276	1.37906	-1.39070

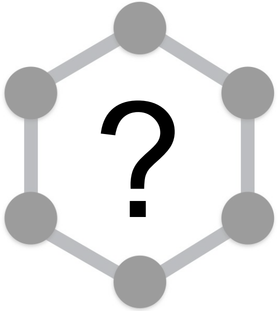
Many representations of benzene...



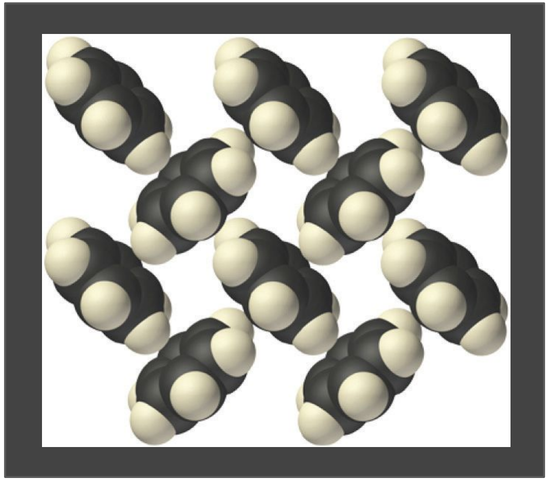
Vector (Fingerprint)



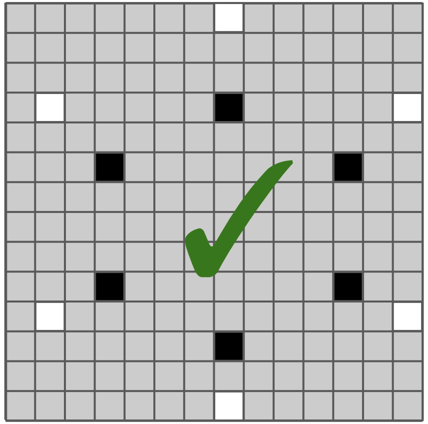
SMILES string:
C1=CC=CC=C1



Graph of bonds



Image

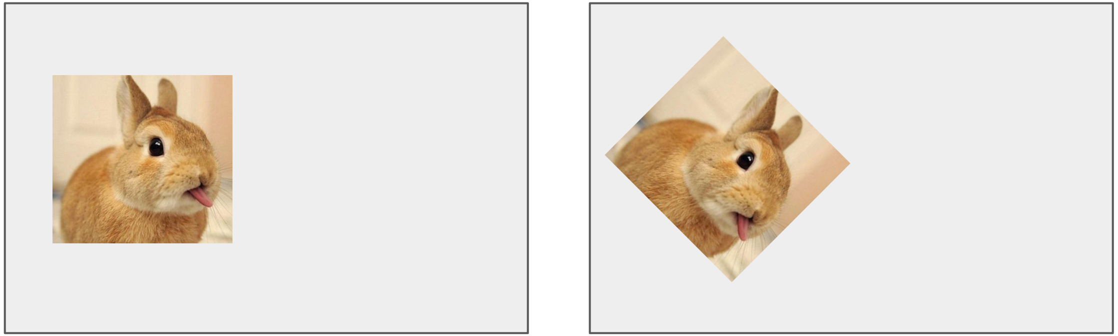


3D Coordinates

H	-0.21463	0.97837	0.33136
C	-0.38325	0.66317	-0.70334
C	-1.57552	0.03829	-1.05450
H	-2.34514	-0.13834	-0.29630
C	-1.78983	-0.36233	-2.36935
H	-2.72799	-0.85413	-2.64566
C	-0.81200	-0.13809	-3.33310
H	-0.98066	-0.45335	-4.36774
C	0.38026	0.48673	-2.98192
H	1.14976	0.66307	-3.74025
C	0.59460	0.88737	-1.66708
H	1.53276	1.37906	-1.39070

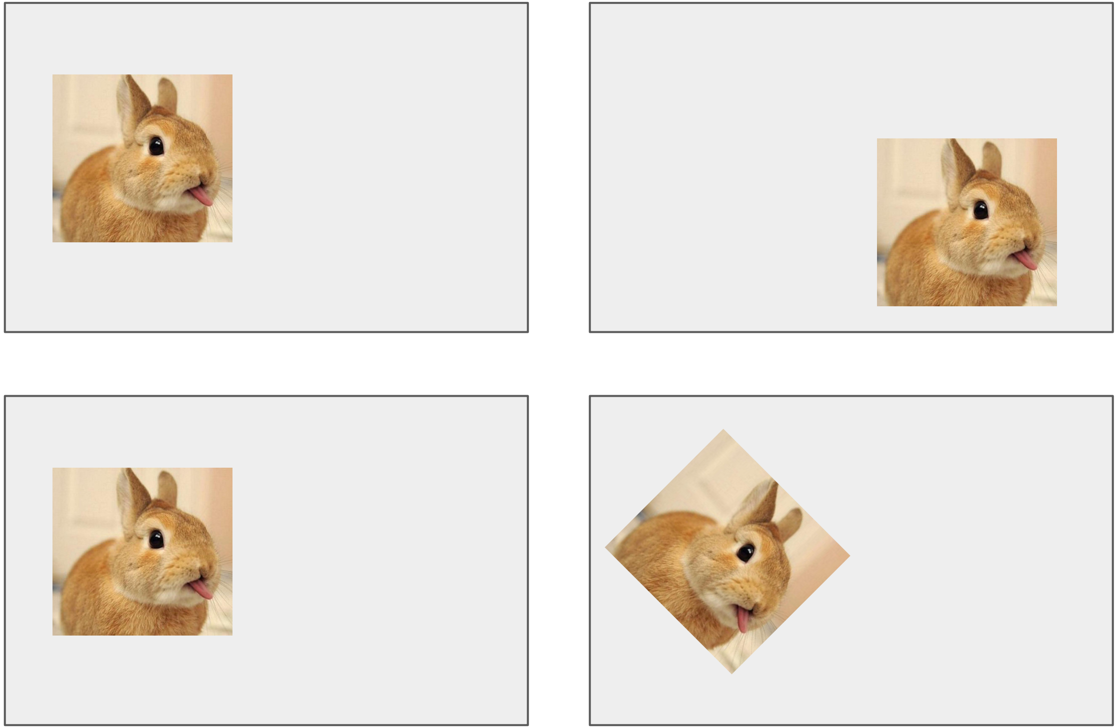
The following properties are useful for a network to have if one deals with geometry:

<u>Translation equivariance</u>



<u>Rotation equivariance</u>

The following properties are useful for a network to have if one deals with geometry:



Translation equivariance

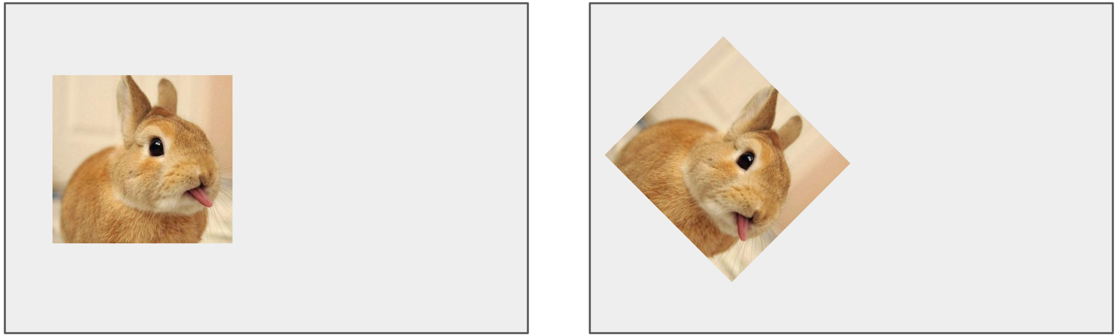
Convolutional neural network ✓

Rotation equivariance?

The following properties are useful for a network to have if one deals with geometry:

Translation equivariance

Convolutional neural network ✓



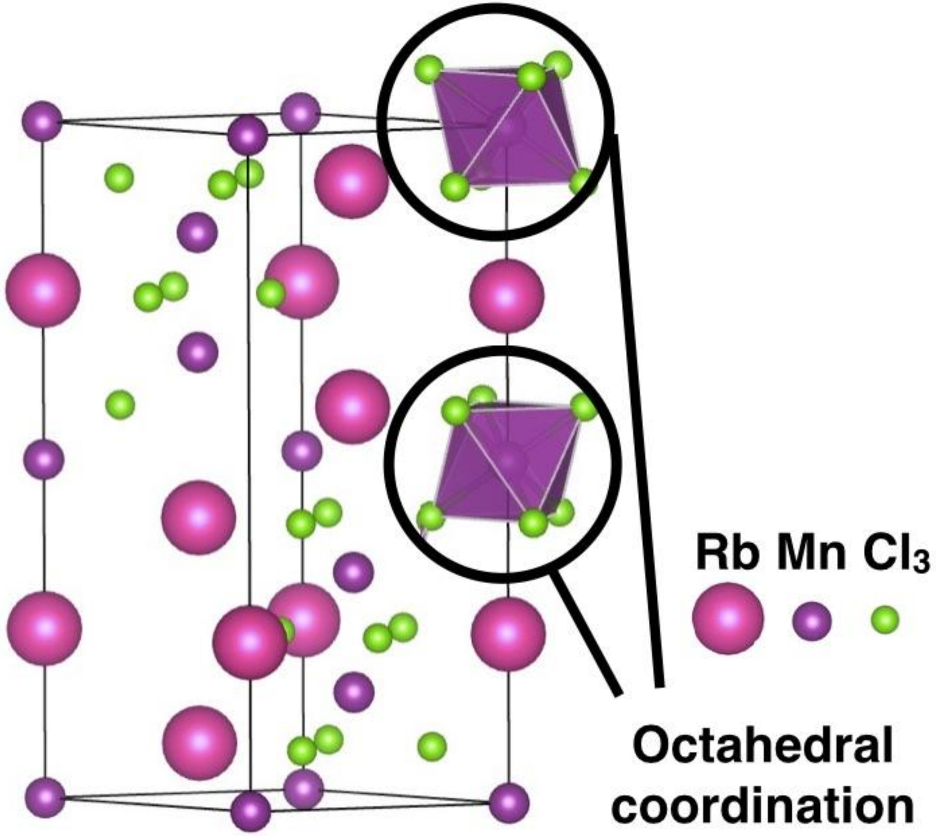
Rotation equivariance

~~Data augmentation~~

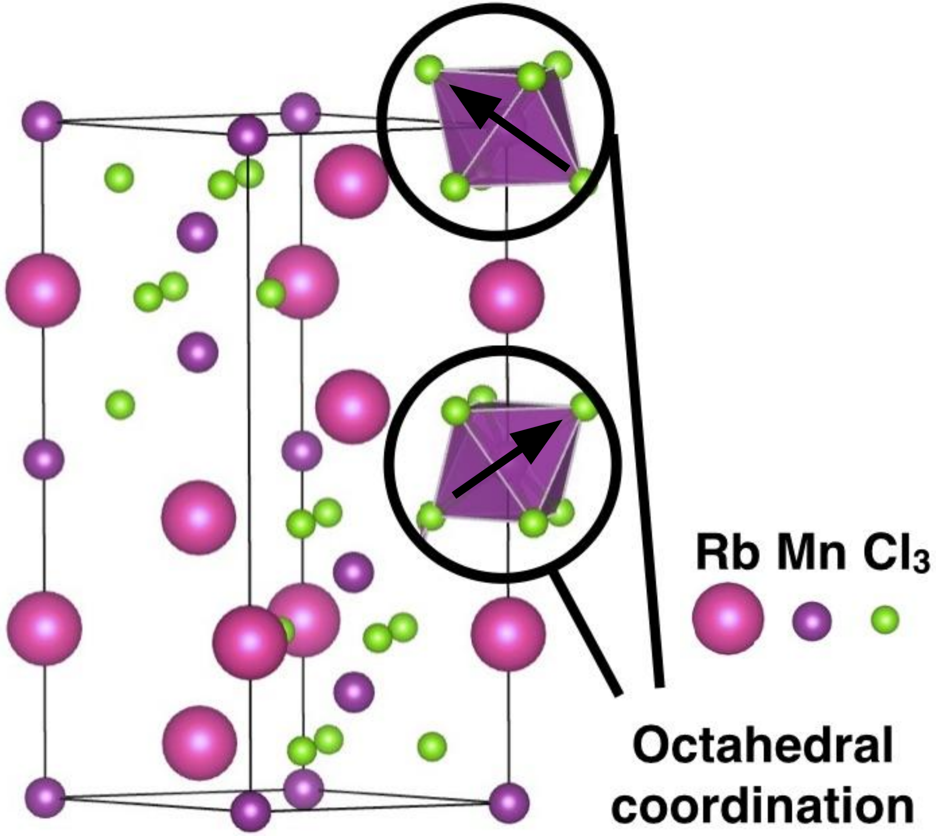
~~Radial functions~~

Want a network that both preserves geometry and exploits symmetry.

A network with 3D translation- and 3D rotation-equivariance allows us to identify chemical motifs in any location or orientation using the same filters.



A network with 3D translation- and 3D rotation-equivariance allows us to identify chemical motifs in any location or orientation using the same filters.



Previous networks applied to atomic systems **either**:

- cannot be extended to across all atomic systems *(molecules, materials, proteins, hybrid systems, nanoclusters, etc)* **or**
- throw out potentially useful geometric information.

Previous networks applied to atomic systems **either**:

- cannot be extended to across all atomic systems *(molecules, materials, proteins, hybrid systems, nanoclusters, etc)* **or**
- throw out potentially useful geometric information.

We created a framework for deep learning on atomic systems that can naturally handle 3D geometry and features of physical systems.

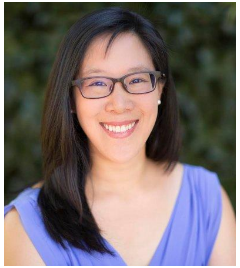
Stanford

Google Accelerated Science Team

Nate Thomas

Patrick Riley

Steve Kearnes



Lusann Yang

Li Li

Kai Kohlhoff

Previous networks applied to atomic systems **either**:

- cannot be extended to across all atomic systems *(molecules, materials, proteins, hybrid systems, nanoclusters, etc)* **or**
- throw out potentially useful geometric information.

We created a framework for deep learning on atomic systems that can naturally handle 3D geometry and features of physical systems.

arXiv:1802.08219

Tensor Field Networks: Rotation- and Translation-Equivariant Neural Networks for 3D Point Clouds

Nathaniel Thomas [*1] Tess Smidt [*234] Steven Kearnes [4] Lusann Yang [4] Li Li [4] Kai Kohlhoff [4] Patrick Riley [4]

Abstract

We introduce tensor field networks, which are locally equivariant to 3D rotations and translations (and invariant to permutations of points) at every layer. 3D rotation equivariance removes the need for data augmentation to identify features in arbitrary orientations. Our network uses filters built from spherical harmonics; due to the mathematical consequences of this filter choice, each

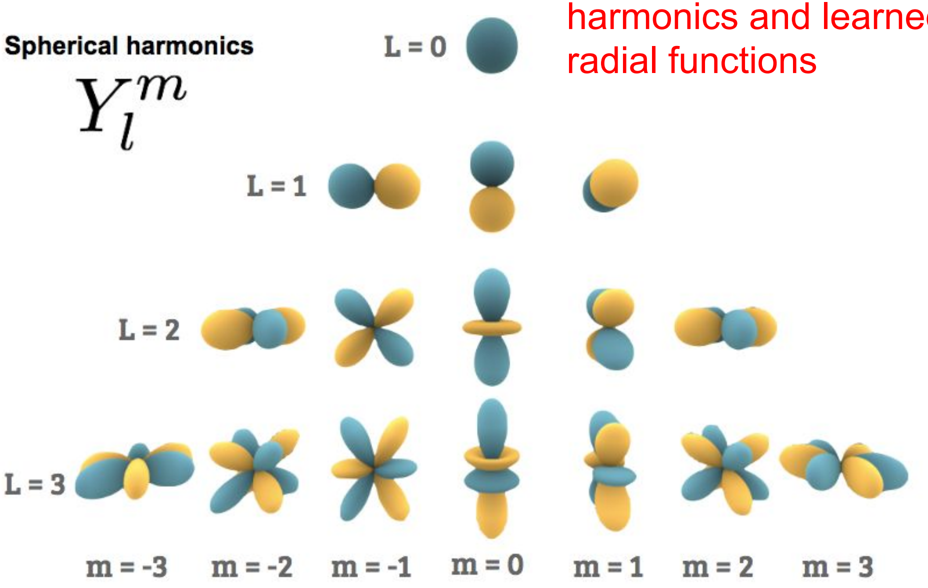
significantly more important in 3D than 2D. Without equivariant filters like those in our design, achieving an angular resolution of δ would require a factor of $\mathcal{O}(\delta^{-1})$ more filters in 2D but $\mathcal{O}(\delta^{-3})$ more filters in 3D.[1] Second, a 3D rotation- and translation-equivariant network can identify local features in different orientations and locations with the same filters, which is helpful for interpretability. Finally, the network naturally encodes geometric tensors (such as scalars, vectors, and higher-rank geometric objects), mathe-

Previous networks applied to atomic systems **either**:

- cannot be extended to across all atomic systems *(molecules, materials, proteins, hybrid systems, nanoclusters, etc)* **or**
- throw out potentially useful geometric information.

We created a framework for deep learning on atomic systems that can naturally handle 3D geometry and features of physical systems.

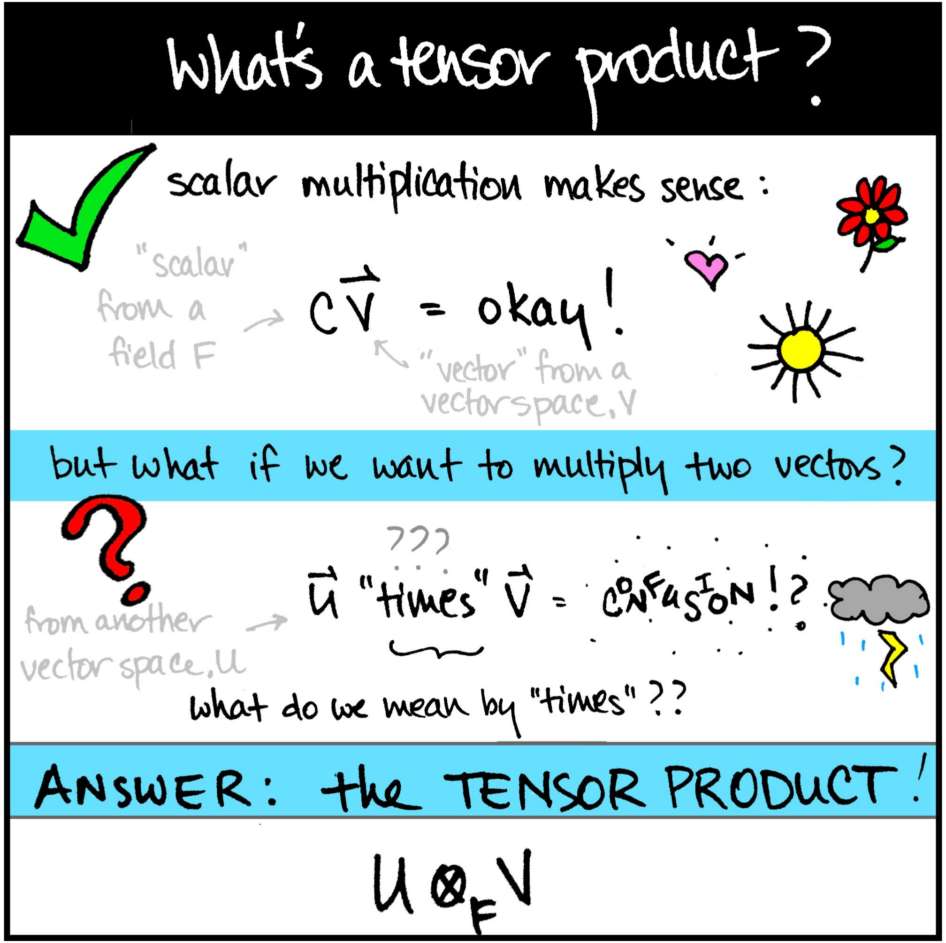
TLDR



Convolutional filters based on spherical harmonics and learned radial functions

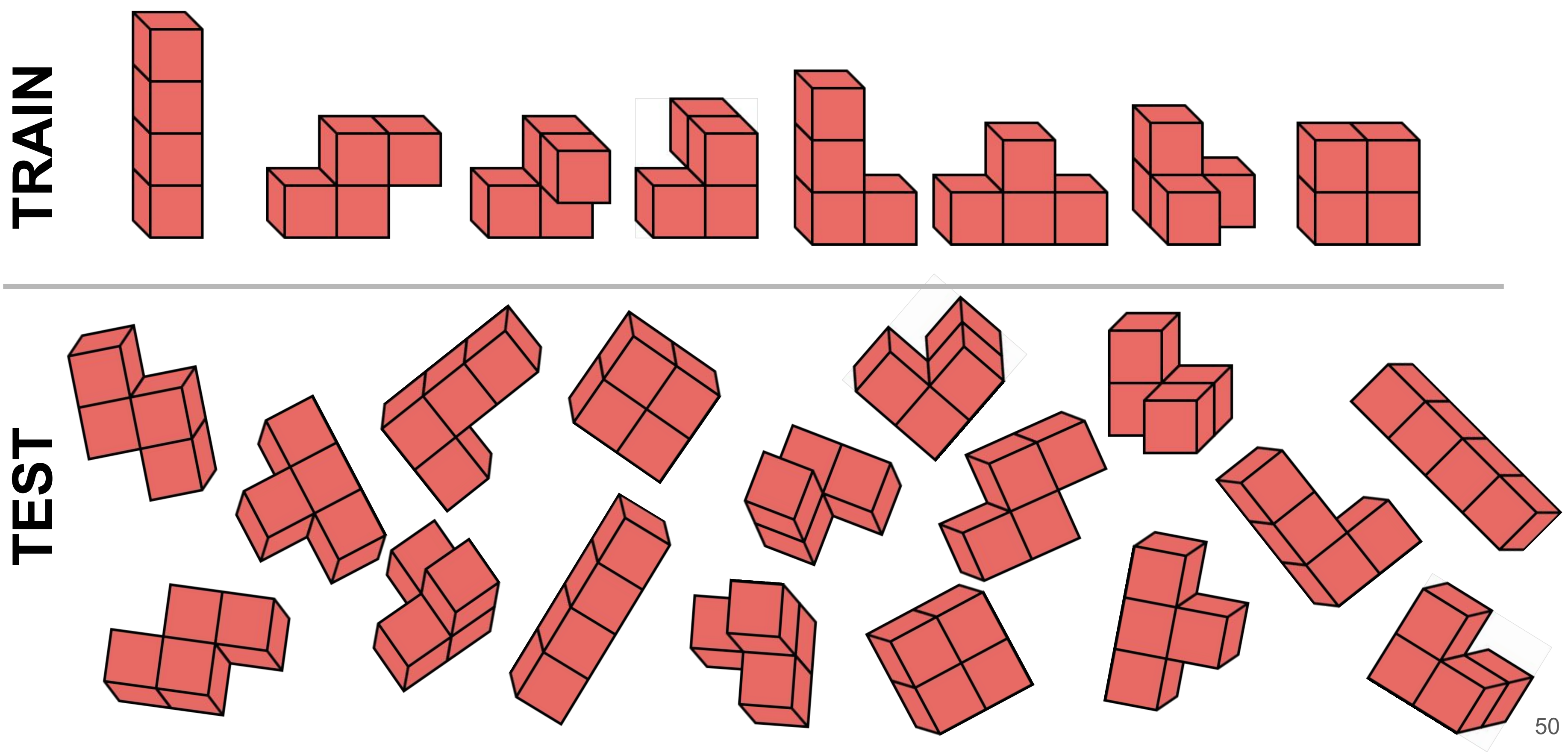
Everything in our network is a **geometric tensor**, so our network connectivity has to obey **tensor algebra**.

$$\otimes$$

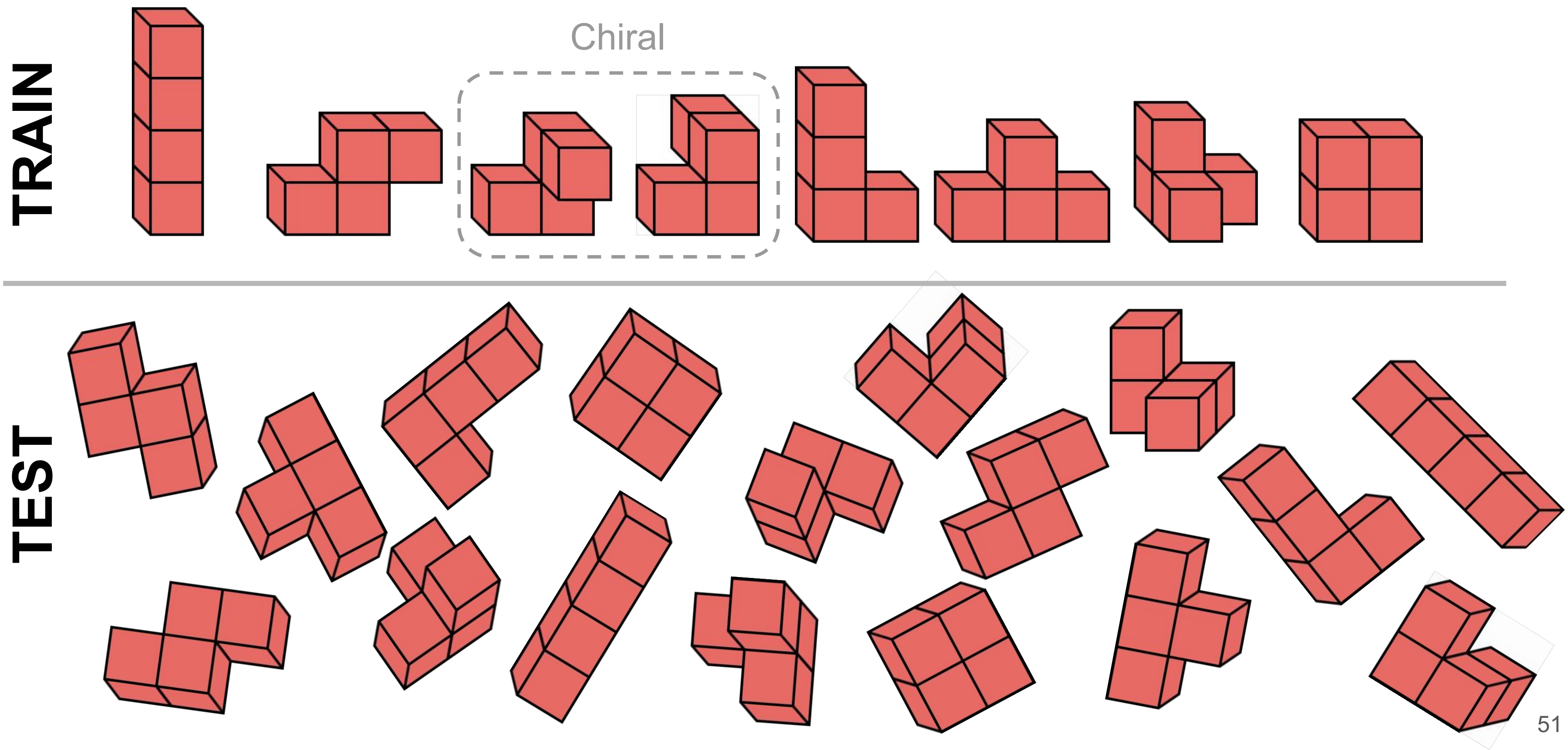

What's a tensor product?
scalar multiplication makes sense:
"scalar" from a field F
$c\vec{v}$ = okay!
"vector" from a vectorspace, V
but what if we want to multiply two vectors?
???
$\vec{u}$ "times" $\vec{v}$ = CONFUSION!?
from another vector space, U
what do we mean by "times"??
ANSWER: the TENSOR PRODUCT!
$U \otimes_F V$

Math3ma.com

Test of 3D rotation equivariance: Trained on 3D Tetris shapes in one orientation, our network can perfectly identify these shapes in any orientation.

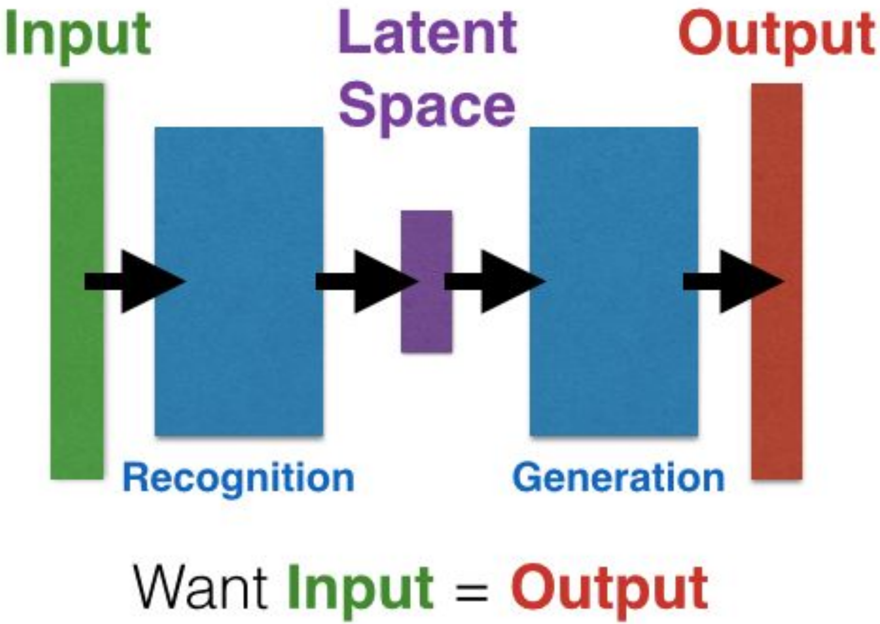


Test of 3D rotation equivariance: Trained on 3D Tetris shapes in one orientation, our network can perfectly identify these shapes in any orientation.



Autoencoders can learn how map data in its original representation to a new representation and back again.

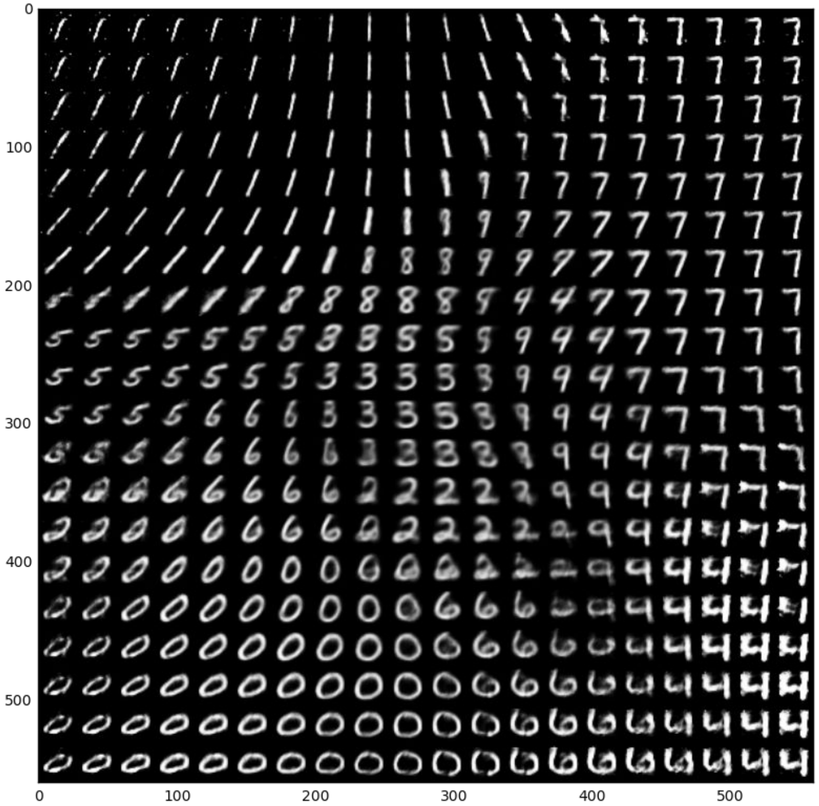
The learned representation is often very useful.



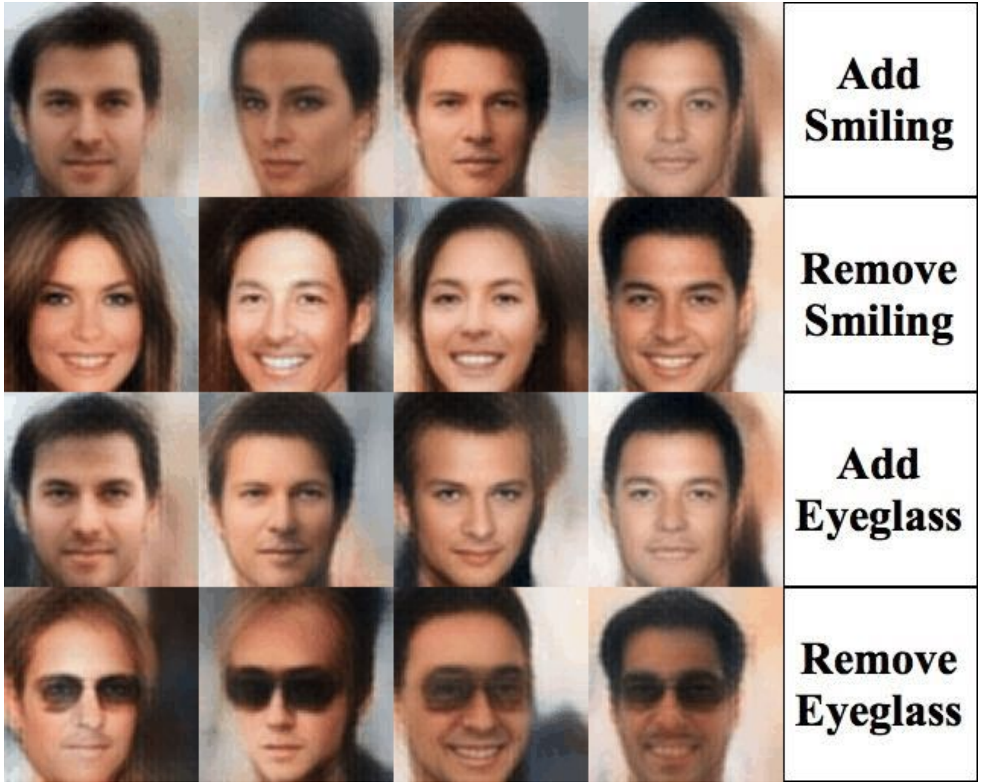
Latent space is either small or has a penalty to have a specified distribution.

Example MNIST digits:

2 dimensional latent space for autoencoder trained on MNIST handwritten digit images



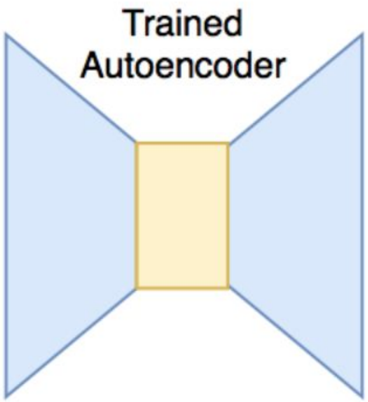
VAE Tutorial: https://jmetzen.github.io/2015-11-27/vae.html



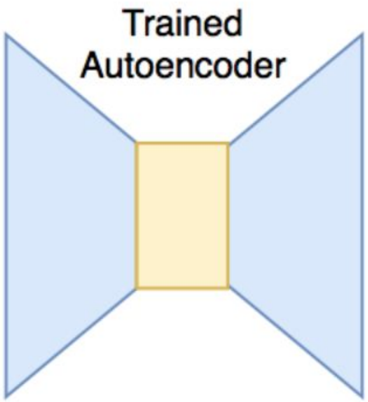
https://houxianxu.github.io/assets/project/dfcvae

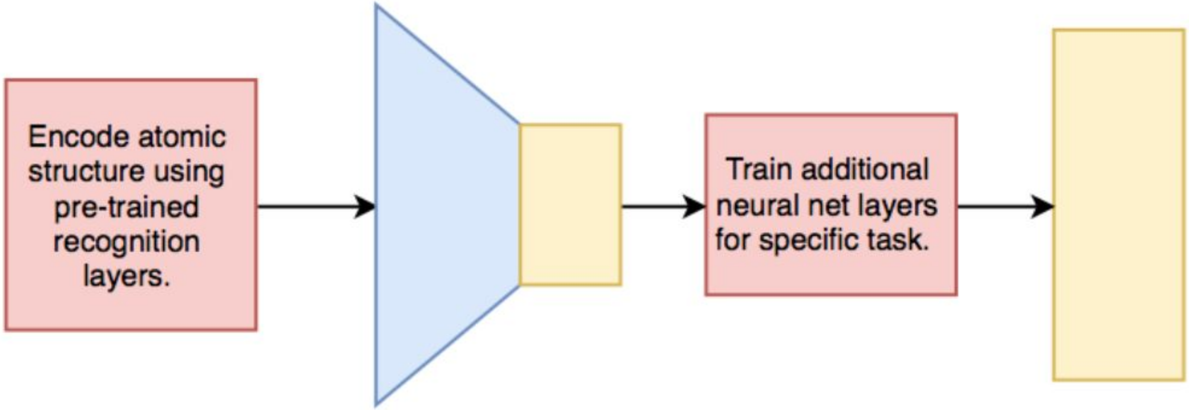
https://twitter.com/smilevector

An autoencoder trained on atomic systems would solve multiple problems at once.

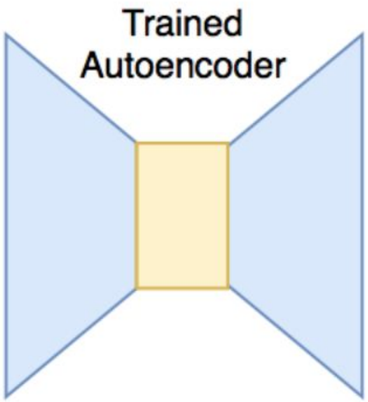


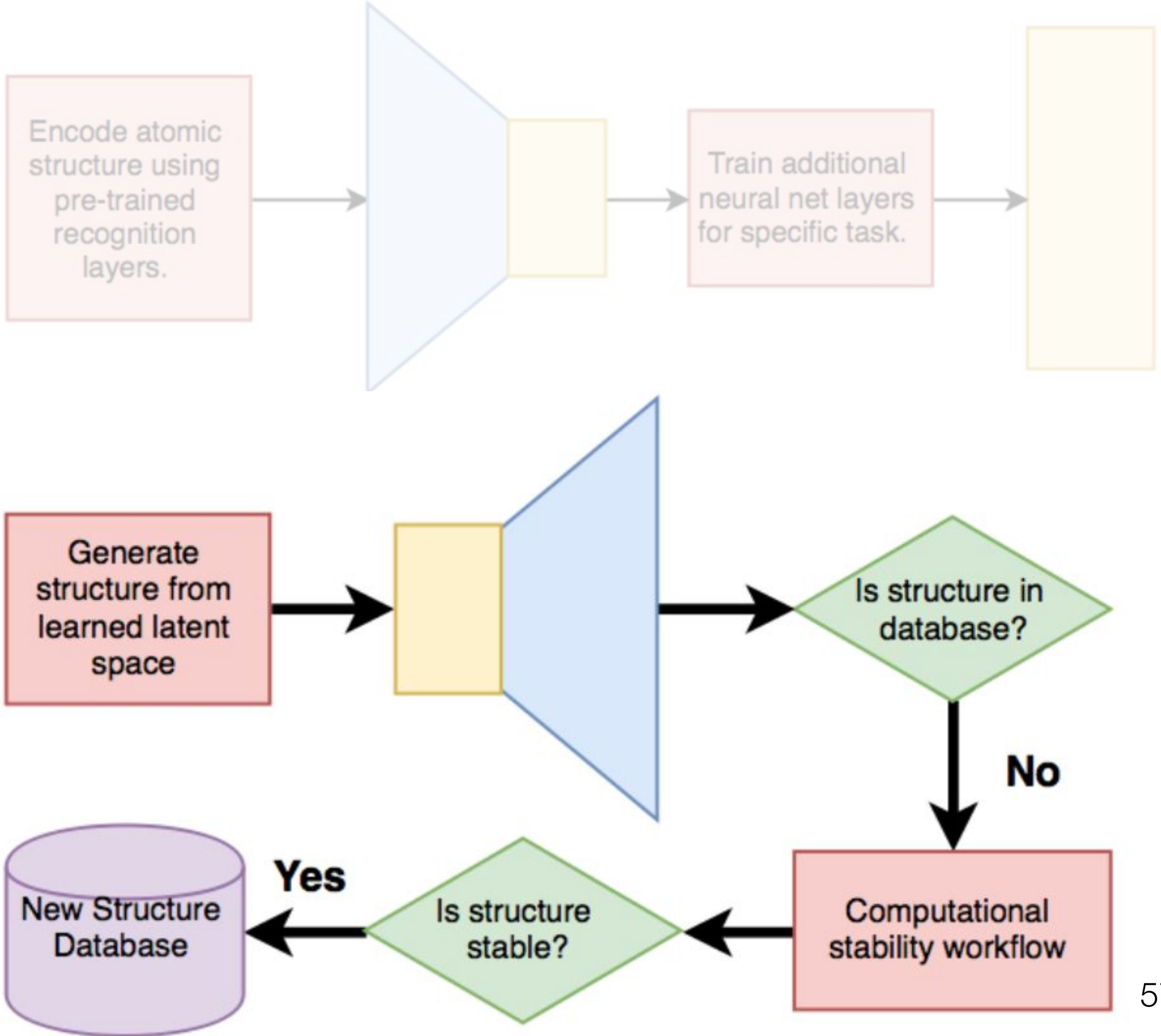
An autoencoder trained on atomic systems would solve multiple problems at once.



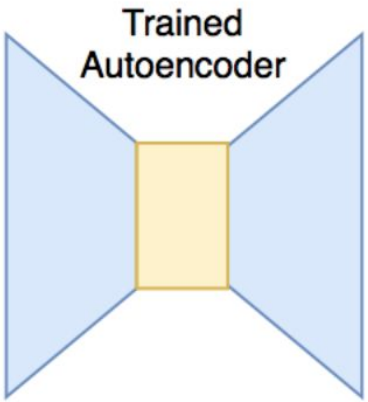


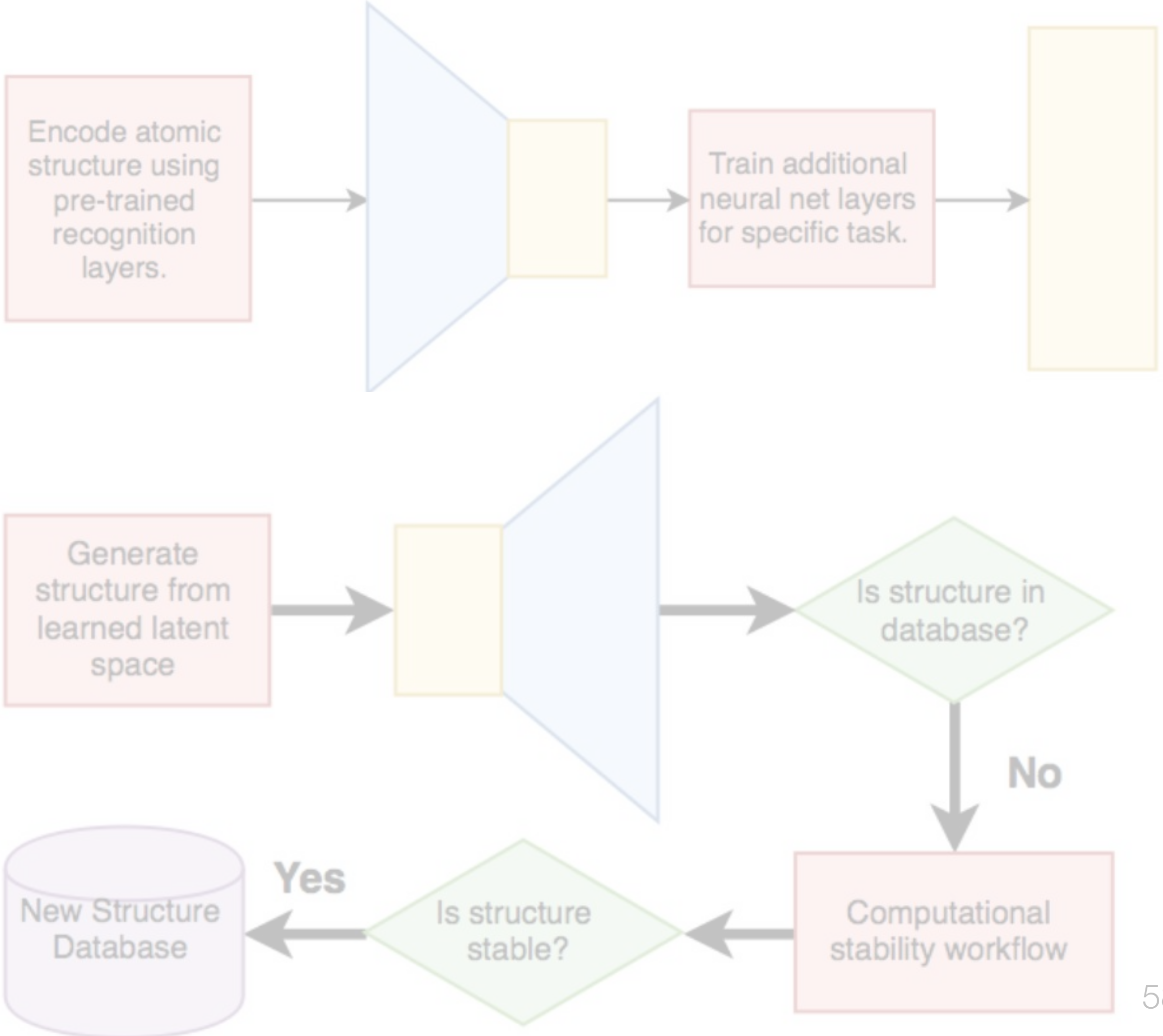
An autoencoder trained on atomic systems would solve multiple problems at once.





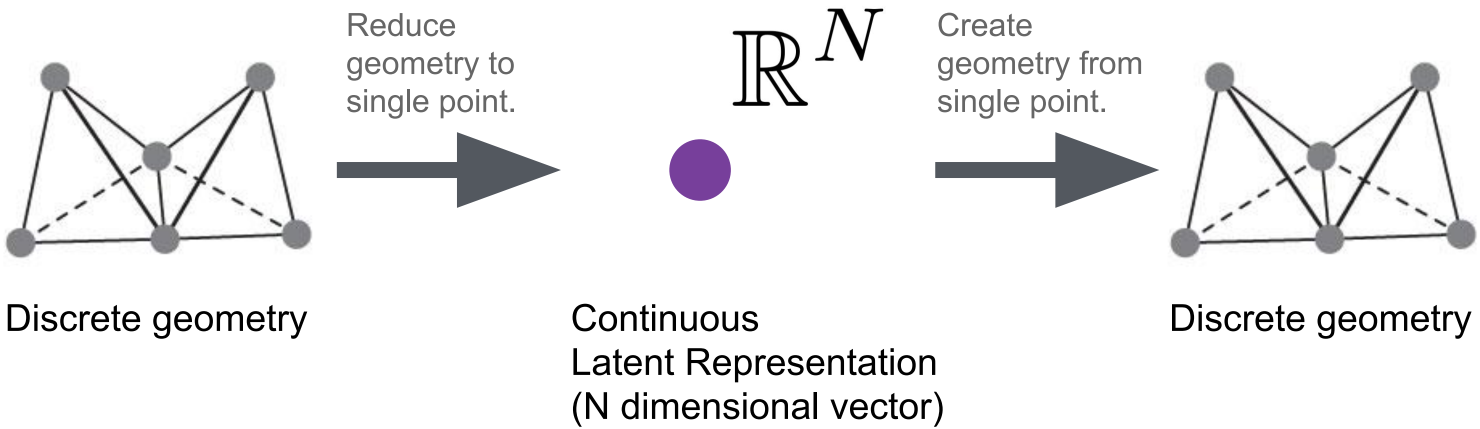
An autoencoder trained on atomic systems would solve multiple problems at once.



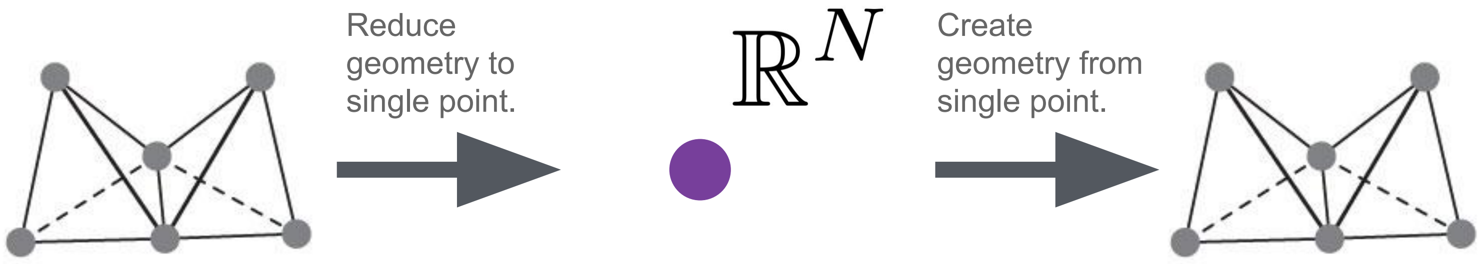


The latent space would provide a "materials map".

Creating an autoencoder for discrete geometry



Creating an autoencoder for discrete geometry



Discrete geometry

Continuous
Latent Representation
(N dimensional vector)

Discrete geometry

Atomic structures are hierarchical and can be constructed from geometric motifs.

- **Encode geometry ✓**
- Encode hierarchy
- Decode geometry
- Decode hierarchy

(Need to do this in a recursive manner)

Okay, so how did I get here?

My Thesis: Toward designing complex materials from structural motifs (The TLDR; version)

Ch 1: Introduction

Ch 2: Methods (DFT)

Ch 3: Realization of a three-dimensional spin-anisotropic harmonic honeycomb iridate

- K. Modic, T. Smidt et al, Nature Communications 5 (2014).

Ch 4: Ab initio Studies of Structural and Energetic Trends in the Harmonic Honeycomb Iridates

- T. Smidt et al, To be submitted to Physical Review B (2018)

Ch 5: Silver Benzeneselenolate is a Self-Assembling Direct-Gap Metal-Organic Chalcogenide Assembly

- M. Collins, T. Smidt et al, In preparation.

Ch 6: An Automatically Curated First-Principles Database of Ferroelectrics

- T. Smidt et al, Submitted to Nature Scientific Data (2018)

Ch 7: Tensor field networks: Rotation- and Translation-Equivariant Neural Networks for 3D Point Clouds

- N. Thomas*, T. Smidt* et al, arXiv:1802.08219

Ch 8: Outlook

My Thesis: Toward designing complex materials from structural motifs (The TLDR; version)

Ch 1: Introduction

Ch 2: Methods (DFT)

Ch 3: Realization of a three-dimensional spin-anisotropic harmonic honeycomb iridate
- K. Modic, T. Smidt et al, Nature Communications 5 (2014).

Ch 4: Ab initio Studies of Structural and Energetic Trends in the Harmonic Honeycomb Iridates
- T. Smidt et al, To be submitted to Physical Review B (2018)

(INSERT EXISTENTIAL RESEARCH CRISIS -- 3rd year)

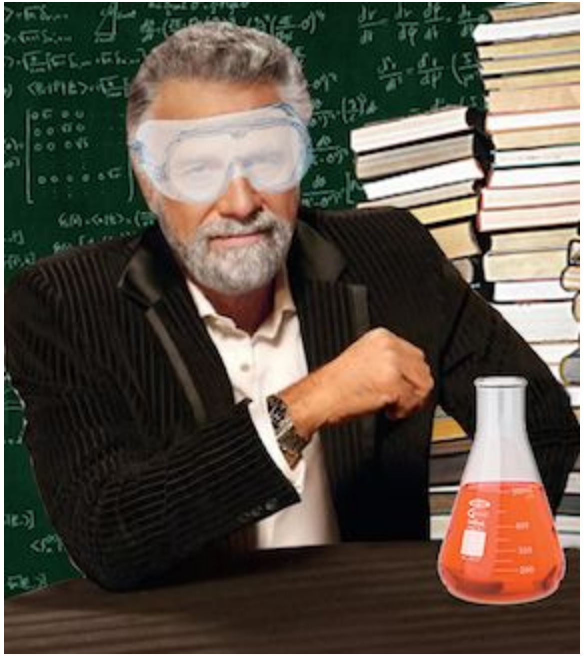
Ch 5: Silver Benzeneselenolate is a Self-Assembling Direct-Gap Metal-Organic Chalcogenide Assembly
- M. Collins, T. Smidt et al, In preparation.

Ch 6: An Automatically Curated First-Principles Database of Ferroelectrics
- T. Smidt et al, Submitted to Nature Scientific Data (2018)

Ch 7: Tensor field networks: Rotation- and Translation-Equivariant Neural Networks for 3D Point Clouds
- N. Thomas*, T. Smidt* et al, arXiv:1802.08219

Ch 8: Outlook

The Lunch Experiment: Randomized Controlled Lunches for Grad Students



**I don't always eat lunch, but when I do,
I prefer The Lunch Experiment.**

400+ participants
100+ lunches

Automated scheduling and invitation system maximizing for diversity of majors.

My Thesis: Toward designing complex materials from structural motifs (The TLDR; version)

Ch 1: Introduction

Ch 2: Methods (DFT)

Ch 3: Realization of a three-dimensional spin-anisotropic harmonic honeycomb iridate

- K. Modic, T. Smidt et al, Nature Communications 5 (2014).

Ch 4: Ab initio Studies of Structural and Energetic Trends in the Harmonic Honeycomb Iridates

- T. Smidt et al, To be submitted to Physical Review B (2018)

(INSERT EXISTENTIAL RESEARCH CRISIS -- 3rd year)

Ch 5: Silver Benzeneselenolate is a Self-Assembling Direct-Gap Metal-Organic Chalcogenide Assembly

- M. Collins, T. Smidt et al, In preparation.

Ch 6: An Automatically Curated First-Principles Database of Ferroelectrics

- T. Smidt et al, Submitted to Nature Scientific Data (2018)

Ch 7: Tensor field networks: Rotation- and Translation-Equivariant Neural Networks for 3D Point Clouds

- N. Thomas*, T. Smidt* et al, arXiv:1802.08219

Ch 8: Outlook

My Thesis: Toward designing complex materials from structural motifs (The TLDR; version)

Ch 1: Introduction

Ch 2: Methods (DFT)

Ch 3: Realization of a three-dimensional spin-anisotropic harmonic honeycomb iridate

- K. Modic, T. Smidt et al, Nature Communications 5 (2014).

Ch 4: Ab initio Studies of Structural and Energetic Trends in the Harmonic Honeycomb Iridates

- T. Smidt et al, To be submitted to Physical Review B (2018)

(INSERT EXISTENTIAL RESEARCH CRISIS -- 3rd year)

Ch 5: Silver Benzeneselenolate is a Self-Assembling Direct-Gap Metal-Organic Chalcogenide Assembly

- M. Collins, T. Smidt et al, In preparation.

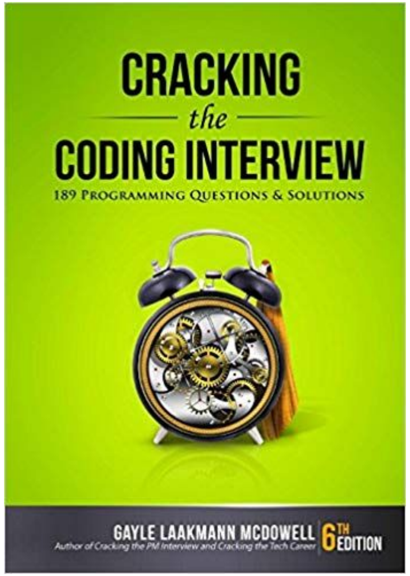
Ch 6: An Automatically Curated First-Principles Database of Ferroelectrics

- T. Smidt et al, Submitted to Nature Scientific Data (2018)

(DEEP LEARNING AND GOOGLE -- 5-6th years)

Ch 7: Tensor field networks: Rotation- and Translation-Equivariant Neural Networks for 3D Point Clouds

- N. Thomas*, T. Smidt* et al, arXiv:1802.08219



+ CS 182/282A =

Spring 2019 COMPSCI 282A 001 - LEC 001 offered through Electrical Engineering and Computer Sciences

Designing, Visualizing and Understanding Deep Neural Networks

John F Canny

M, W | 8:00 am - 9:29 am | Dwinelle 145

Class #: 31116 | Units: 4

Open Seats

30 Unreserved Seats

Deep Networks have revolutionized computer vision, language technology, robotics and control. They have growing impact in many other areas of science and engineering. They do not however, follow a closed or compact set of theoretical principles. In Yann Lecun's words they require "an interplay between intuitive insights, theoretical modeling, practical implementations, empirical studies, and...

My Thesis: Toward designing complex materials from structural motifs (The TLDR; version)

Ch 1: Introduction

Ch 2: Methods (DFT)

Ch 3: Realization of a three-dimensional spin-anisotropic harmonic honeycomb iridate
- K. Modic, T. Smidt et al, Nature Communications 5 (2014).

Ch 4: Ab initio Studies of Structural and Energetic Trends in the Harmonic Honeycomb Iridates
- T. Smidt et al, To be submitted to Physical Review B (2018)

(INSERT EXISTENTIAL RESEARCH CRISIS -- 3rd year)

Ch 5: Silver Benzeneselenolate is a Self-Assembling Direct-Gap Metal-Organic Chalcogenide Assembly
- M. Collins, T. Smidt et al, In preparation.

Ch 6: An Automatically Curated First-Principles Database of Ferroelectrics
- T. Smidt et al, Submitted to Nature Scientific Data (2018)

(DEEP LEARNING AND GOOGLE -- 5-6th years)

Ch 7: Tensor field networks: Rotation- and Translation-Equivariant Neural Networks for 3D Point Clouds
- N. Thomas*, T. Smidt* et al, arXiv:1802.08219

In summary...

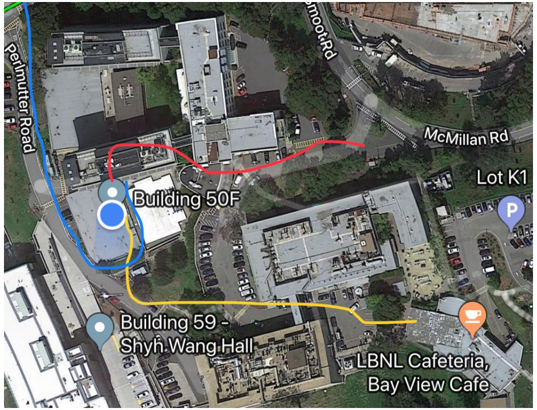
There's a lot of work to do in applying deep learning methods for tasks in atomic systems. Methods may not work out of the box. Many technical design choices to make and test.

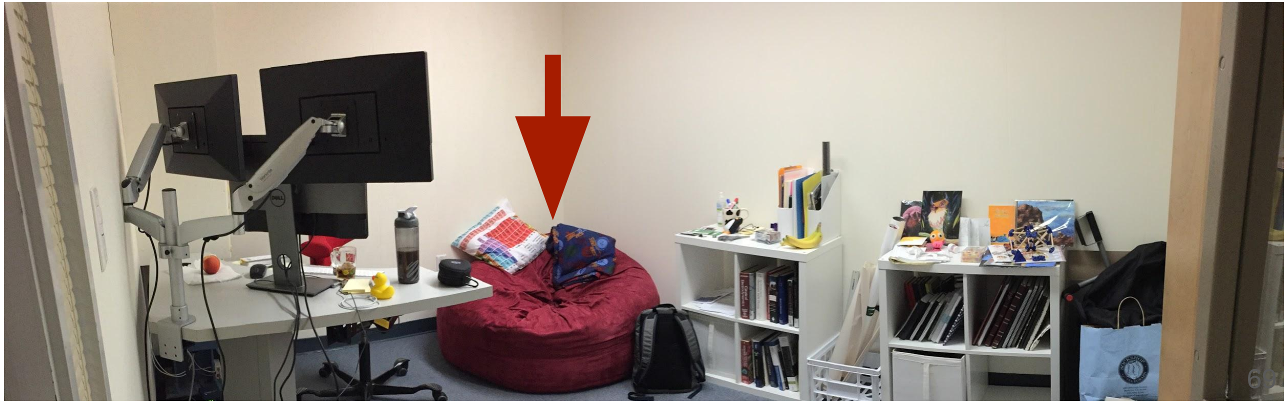
Google is an amazing place to work. I highly recommend interning during grad school if you can.

Berkeley Lab is in a great position to play a central role in how ML methods are adopted in the chemistry and materials communities.

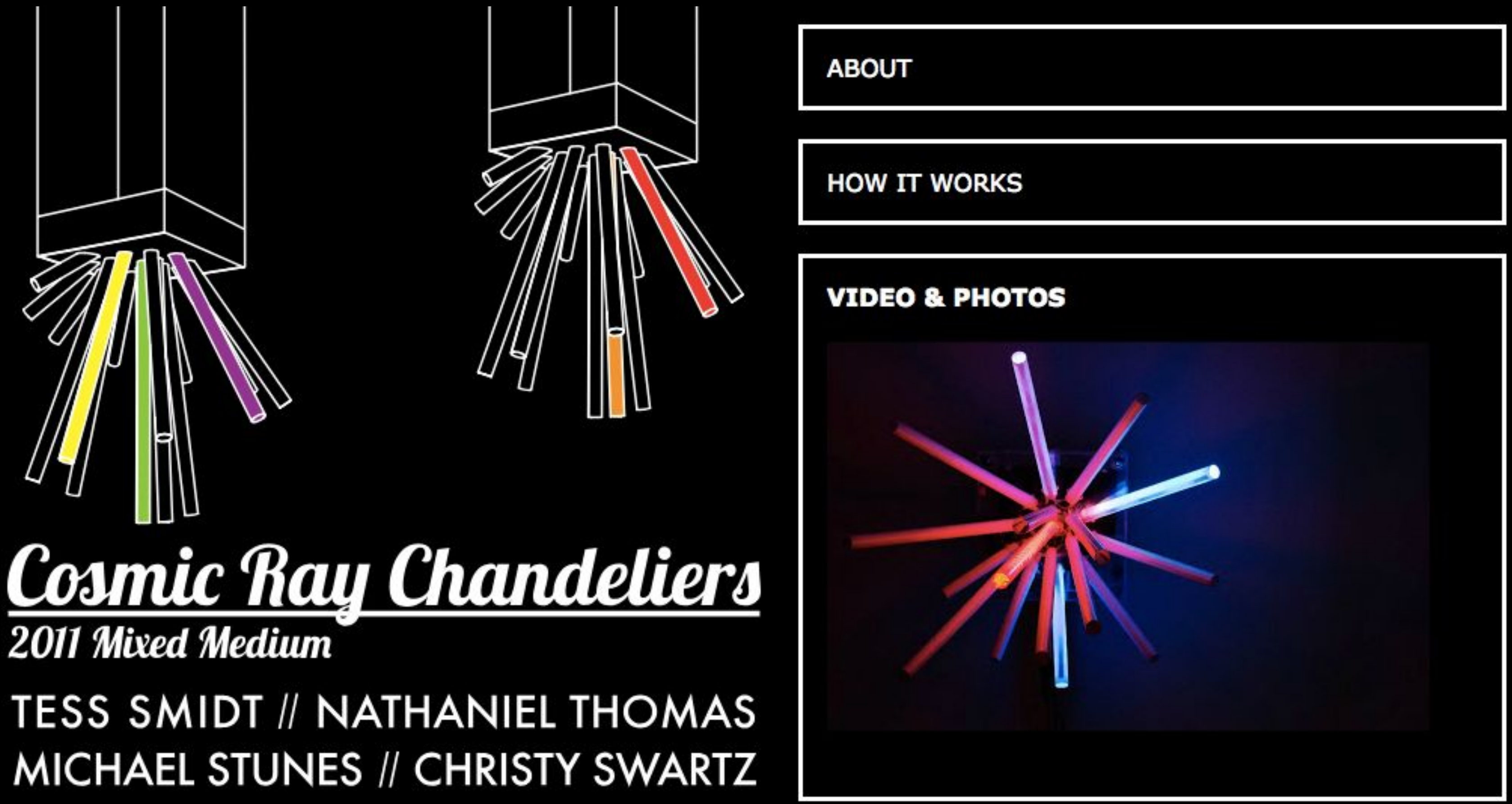
Review on ML for molecules and materials:
Machine learning for molecular and materials science
Keith T. Butler, Daniel W. Davies, Hugh Cartwright, Olexandr Isayev & Aron Walsh
Nature **559**, 547–555 (2018).
https://doi.org/10.1038/s41586-018-0337-2

Come visit and chat about DL for atomic systems! My office is 50F-1643.



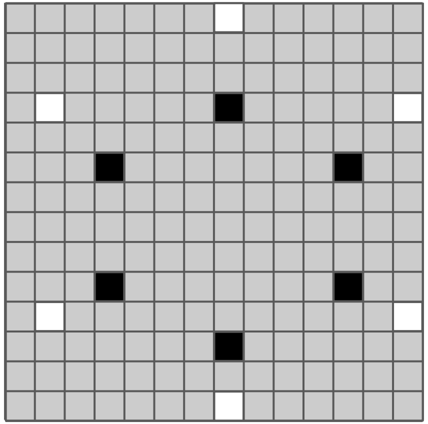


Calling in backup (slides)!

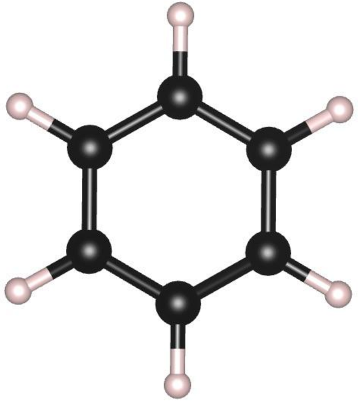


http://blondegeek.net/cosmicray

We use points. Images of atomic systems are sparse and imprecise.



vs.

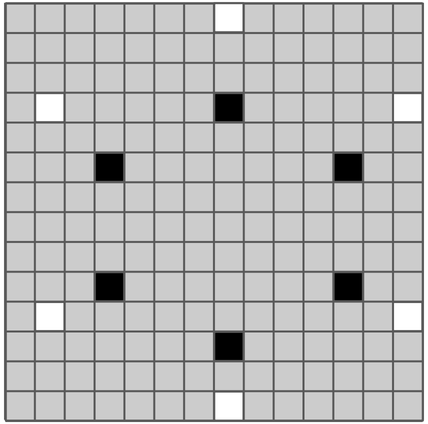


We use continuous convolutions with atoms as convolution centers.

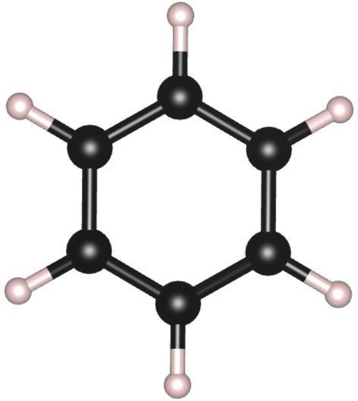
K. T. Schütt, P.-J. Kindermans, H. E. Sauceda, S. Chmiela, A. Tkatchenko, and K.-R. Müller, Adv. in Neural Information Processing Systems 30 (2017). (arXiv: 1706.08566)

We encode the symmetries of 3D Euclidean space (3D translation- and 3D rotation-equivariance).

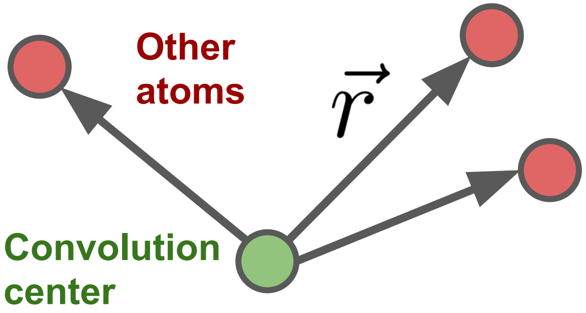
We use points. Images of atomic systems are sparse and imprecise.



vs.



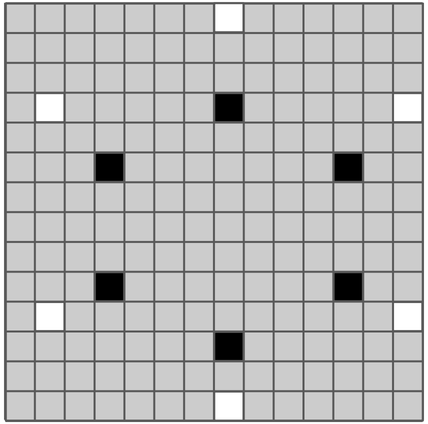
We use continuous convolutions with atoms as convolution centers.



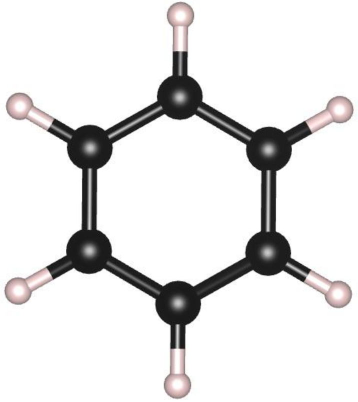
K. T. Schütt, P.-J. Kindermans, H. E. Sauceda, S. Chmiela, A. Tkatchenko, and K.-R. Müller, Adv. in Neural Information Processing Systems 30 (2017). (arXiv: 1706.08566)

We encode the symmetries of 3D Euclidean space (3D translation- and 3D rotation-equivariance).

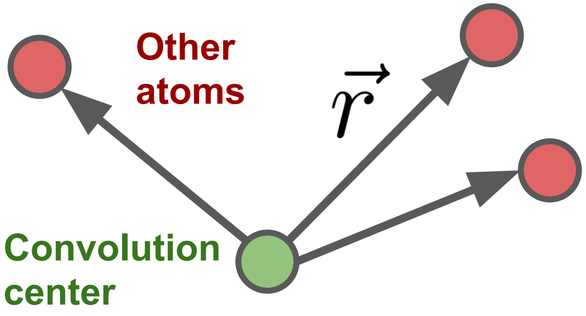
We use points. Images of atomic systems are sparse and imprecise.



vs.

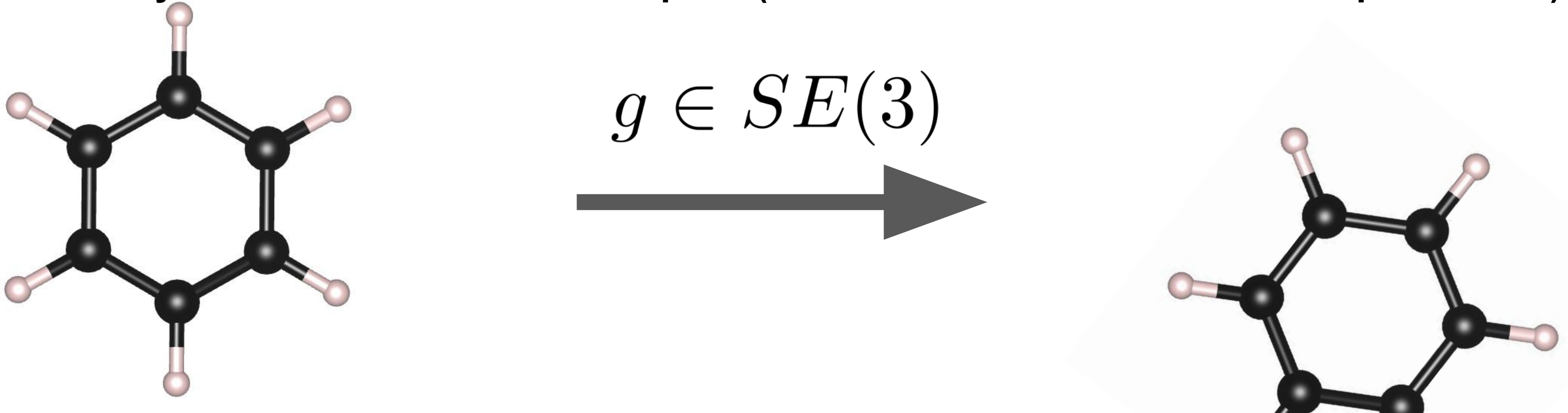


We use continuous convolutions with atoms as convolution centers.

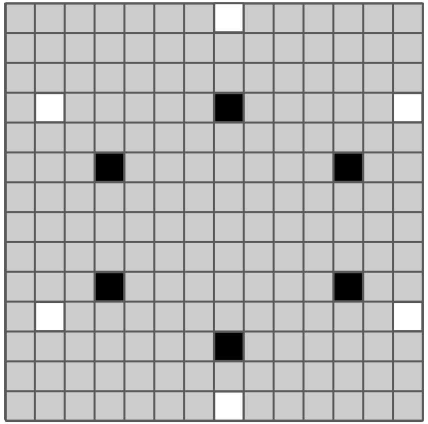


K. T. Schütt, P.-J. Kindermans, H. E. Sauceda, S. Chmiela, A. Tkatchenko, and K.-R. Müller, Adv. in Neural Information Processing Systems 30 (2017). (arXiv: 1706.08566)

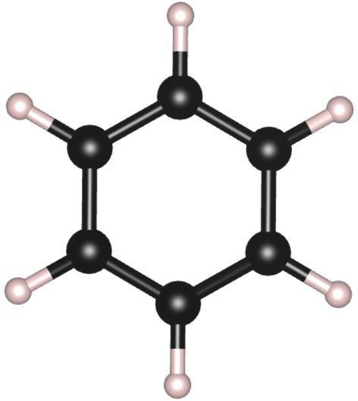
We encode the symmetries of 3D Euclidean space (3D translation- and 3D rotation-equivariance).



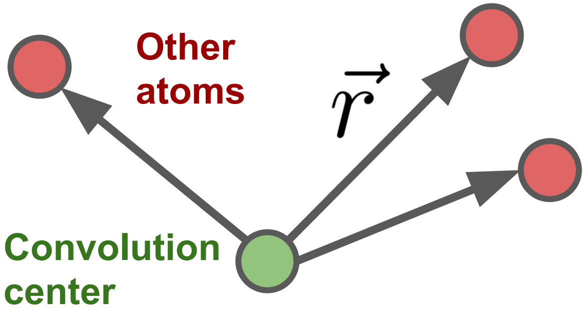
We use points. Images of atomic systems are sparse and imprecise.



vs.

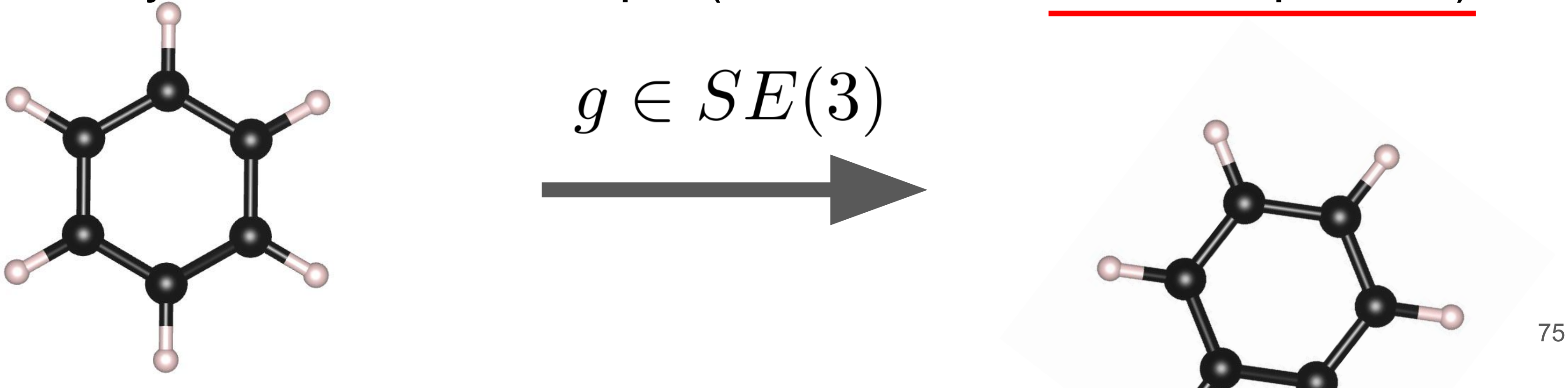


We use continuous convolutions with atoms as convolution centers.



K. T. Schütt, P.-J. Kindermans, H. E. Sauceda, S. Chmiela, A. Tkatchenko, and K.-R. Müller, Adv. in Neural Information Processing Systems 30 (2017). (arXiv: 1706.08566)

We encode the symmetries of 3D Euclidean space (3D translation- and 3D rotation-equivariance).



Convolutional kernels...

Learned Parameters

with no symmetry:

$$\boxed{W(\vec{r})}$$

with 3D rotation equivariance:

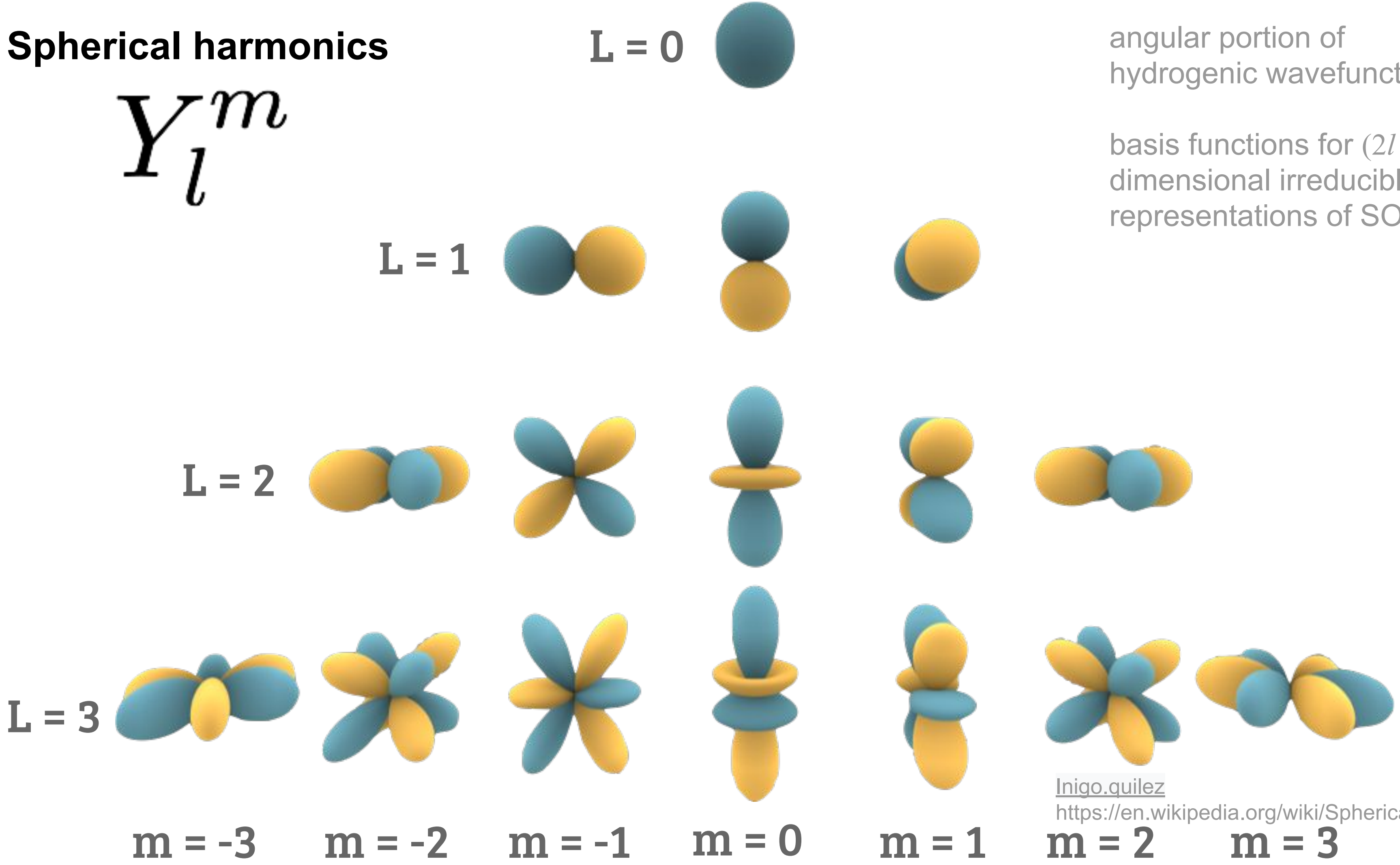
$$\boxed{R(r)}\,Y_l^m(\hat{r})$$

Spherical harmonics

$$Y_l^m$$

angular portion of hydrogenic wavefunctions

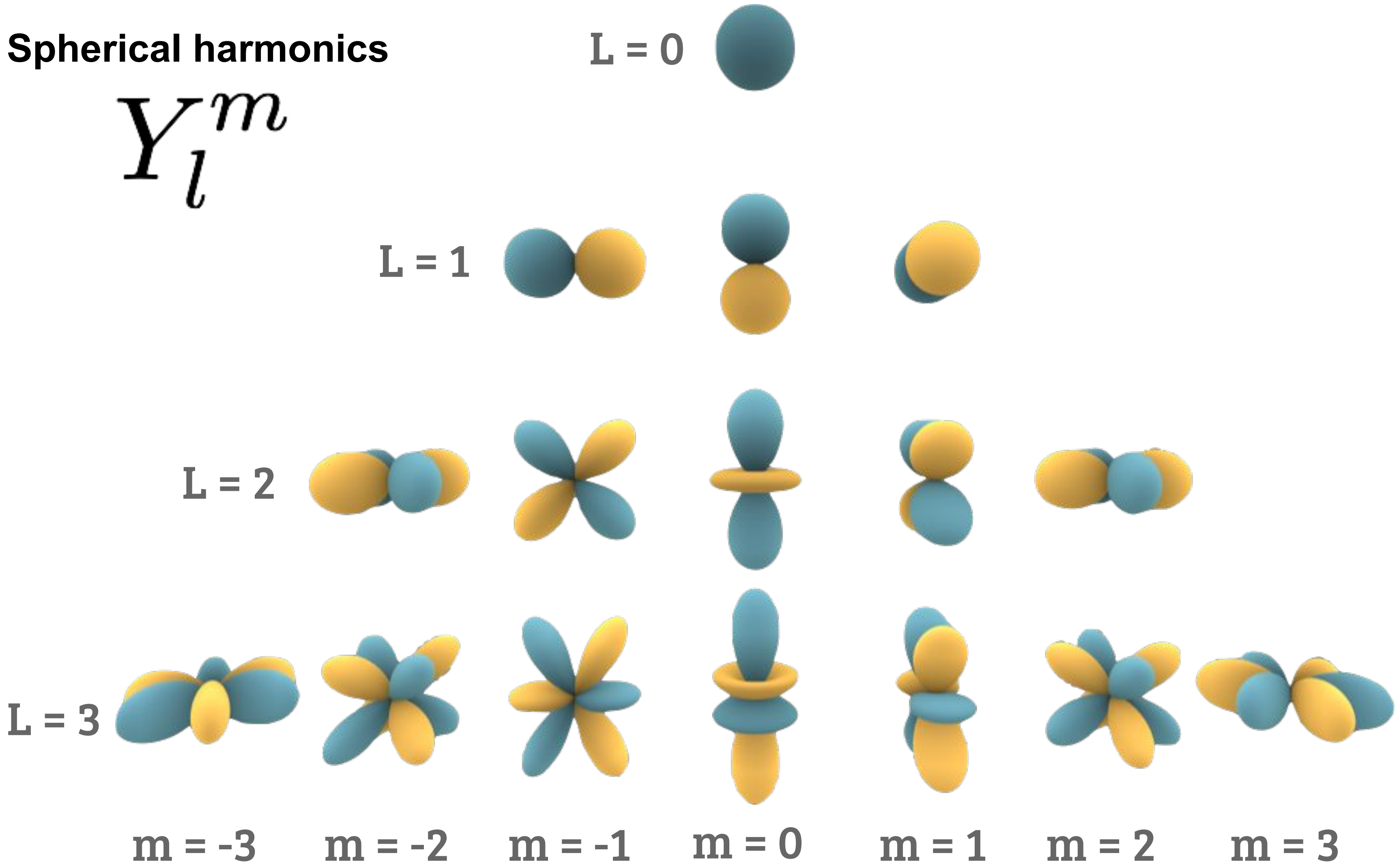
basis functions for $(2l+1)$ dimensional irreducible representations of SO(3)



Inigo.quilez
https://en.wikipedia.org/wiki/Spherical_harmonics

Spherical harmonics

$$Y_l^m$$



Convolutional kernels...

Learned Parameters

with no symmetry:

$$\boxed{W(\vec{r})}$$

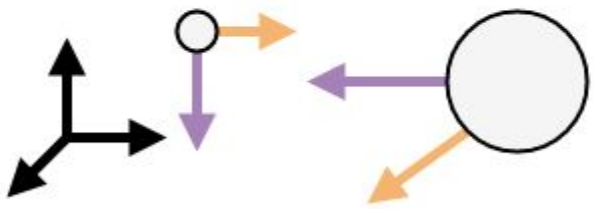
with 3D rotation equivariance:

$$\boxed{R(r)}Y_l^m(\hat{r})$$

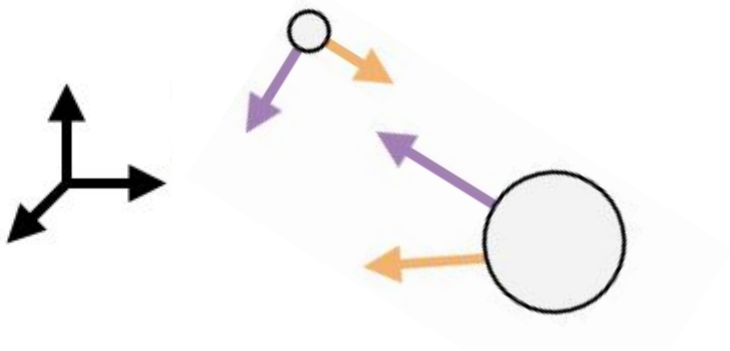
Our filter choice requires the input, filters, and output of our network to be **geometric tensors** and our network connectivity to be compatible with **tensor algebra**.
(Everything has L and M indices like the spherical harmonics.)

Geometric tensors transform predictably under 3D rotation.

Two point **masses** with **velocity** and **acceleration.**

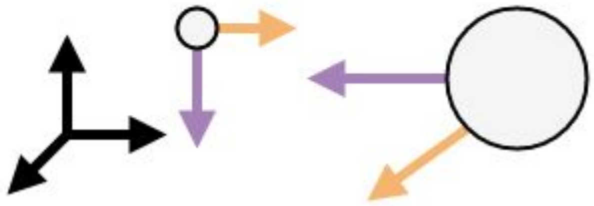


Same system, with rotated coordinates.

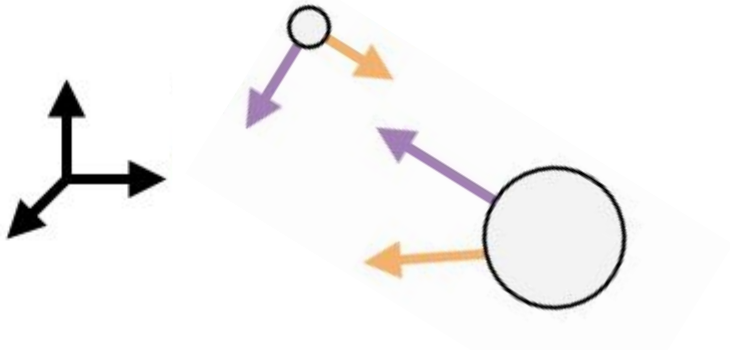


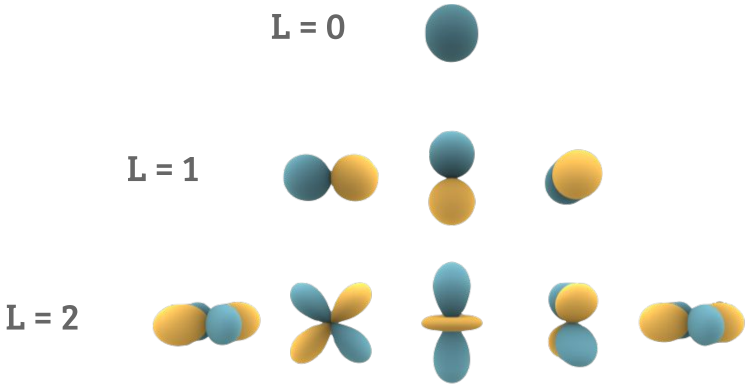
Geometric tensors transform predictably under 3D rotation.

Two point **masses** with **velocity** and **acceleration.**



Same system, with rotated coordinates.





Irreducible representations

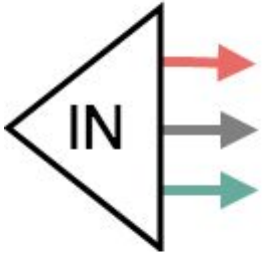
Scalars fields $l = 0$

Vectors fields $l = 1$

3x3 Matrix fields $l = 0 \oplus 1 \oplus 2$

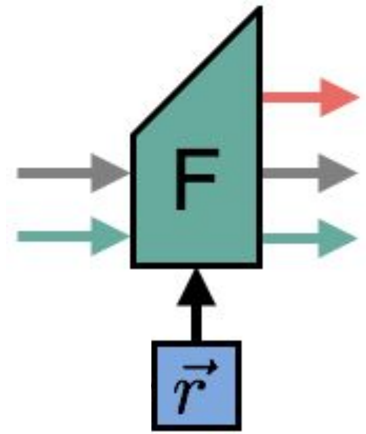
The input and output of our network is represented as tensors with **point** (or atom), **channel**, and **representation** indices organized by irreducible representation (L’s and M’s that index spherical harmonics).

$$V^{(l)}_{acm} =$$

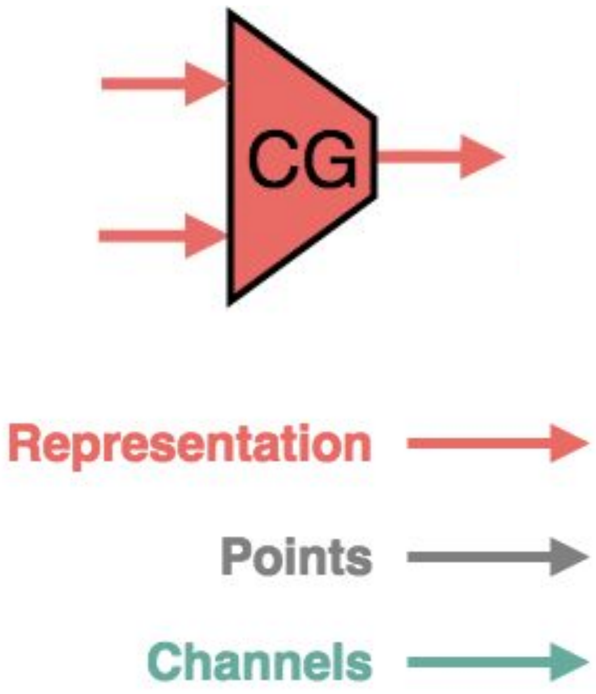


Filters contribute a **representation** index due to use of spherical harmonics.

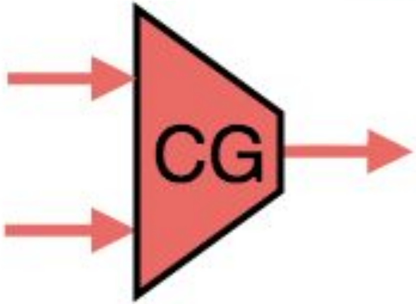
$$R(r)Y_l^m(\hat{r})$$



To combine two tensors to create one tensor, we uses Clebsch-Gordan coefficients.



To combine two tensors to create one tensor, we uses Clebsch-Gordan coefficients.



1/2 × 1/2

m_1	m_2	1 +1
+1/2	+1/2	1

m_1	m_2	1 0	0 0
+1/2	−1/2	1/2	1/2
−1/2	+1/2	1/2	−1/2

m_1	m_2	1 −1
−1/2	−1/2	1

1 × 1/2

m_1	m_2	3/2 +3/2
+1	+1/2	1

m_1	m_2	3/2 +1/2	1/2 +1/2
+1	−1/2	1/3	2/3
0	+1/2	2/3	−1/3

m_1	m_2	3/2 −1/2	1/2 −1/2
0	−1/2	2/3	1/3
−1	+1/2	1/3	−2/3

m_1	m_2	3/2 −3/2
−1	−1/2	1

2 × 1

m_1	m_2	3 +3
+2	+1	1

m_1	m_2	3 +2	2 +2
+2	0	1/3	2/3
+1	+1	2/3	−1/3

m_1	m_2	3 +1	2 +1	1 +1
+2	−1	1/15	1/3	3/5
+1	0	8/15	1/6	−3/10
0	+1	6/15	−1/2	1/10

m_1	m_2	3 0	2 0	1 0
+1	−1	1/5	1/2	3/10
0	0	3/5	0	−2/5
−1	+1	1/5	−1/2	3/10

m_1	m_2	3 −1	2 −1	1 −1
0	−1	6/15	1/2	1/10
−1	0	8/15	−1/6	−3/10
−2	+1	1/15	−1/3	3/5

m_1	m_2	3 −2	2 −2
−1	−1	2/3	1/3
−2	0	1/3	−2/3

m_1	m_2	3 −3
−2	−1	1

1 × 1

m_1	m_2	2 +2
+1	+1	1

m_1	m_2	2 +1	1 +1
+1	0	1/2	1/2
0	+1	1/2	−1/2

m_1	m_2	2 0	1 0	0 0
+1	−1	1/6	1/2	1/3
0	0	2/3	0	−1/3
−1	+1	1/6	−1/2	1/3

m_1	m_2	2 −1	1 −1
0	−1	1/2	1/2
−1	0	1/2	−1/2

m_1	m_2	2 −2
−1	−1	1

3/2 × 1

m_1	m_2	5/2 +5/2
+3/2	+1	1

m_1	m_2	5/2 +3/2	3/2 +3/2
+3/2	0	2/5	3/5
+1/2	+1	3/5	−2/5

m_1	m_2	5/2 +1/2	3/2 +1/2	1/2 +1/2
+3/2	−1	1/10	2/5	1/2
+1/2	0	3/5	1/15	−1/3
−1/2	+1	3/10	−8/15	1/6

m_1	m_2	5/2 −1/2	3/2 −1/2	1/2 −1/2
+1/2	−1	3/10	8/15	1/6
−1/2	0	3/5	−1/15	−1/3
−3/2	+1	1/10	−2/5	1/2

m_1	m_2	5/2 −3/2	3/2 −3/2
−1/2	−1	3/5	2/5
−3/2	0	2/5	−3/5

m_1	m_2	5/2 −5/2
−3/2	−1	1

2 × 1/2

m_1	m_2	5/2 +5/2
+2	1/2	1

m_1	m_2	5/2 3/2	3/2 +3/2
+2	−1/2	1/5	4/5
+1	+1/2	4/5	−1/5

m_1	m_2	5/2 +1/2	3/2 +1/2
+1	−1/2	2/5	3/5
0	+1/2	3/5	−2/5

m_1	m_2	5/2 −1/2	3/2 −1/2
0	−1/2	3/5	2/5
−1	+1/2	2/5	−3/5

m_1	m_2	5/2 −3/2	3/2 −3/2
−1	−1/2	4/5	1/5
−2	+1/2	1/5	−4/5

m_1	m_2	5/2 −5/2
−2	−1/2	1

3/2 × 1/2

m_1	m_2	2 +2
+3/2	+1/2	1

m_1	m_2	2 +1	1 +1
+3/2	−1/2	1/4	3/4
+1/2	+1/2	3/4	−1/4

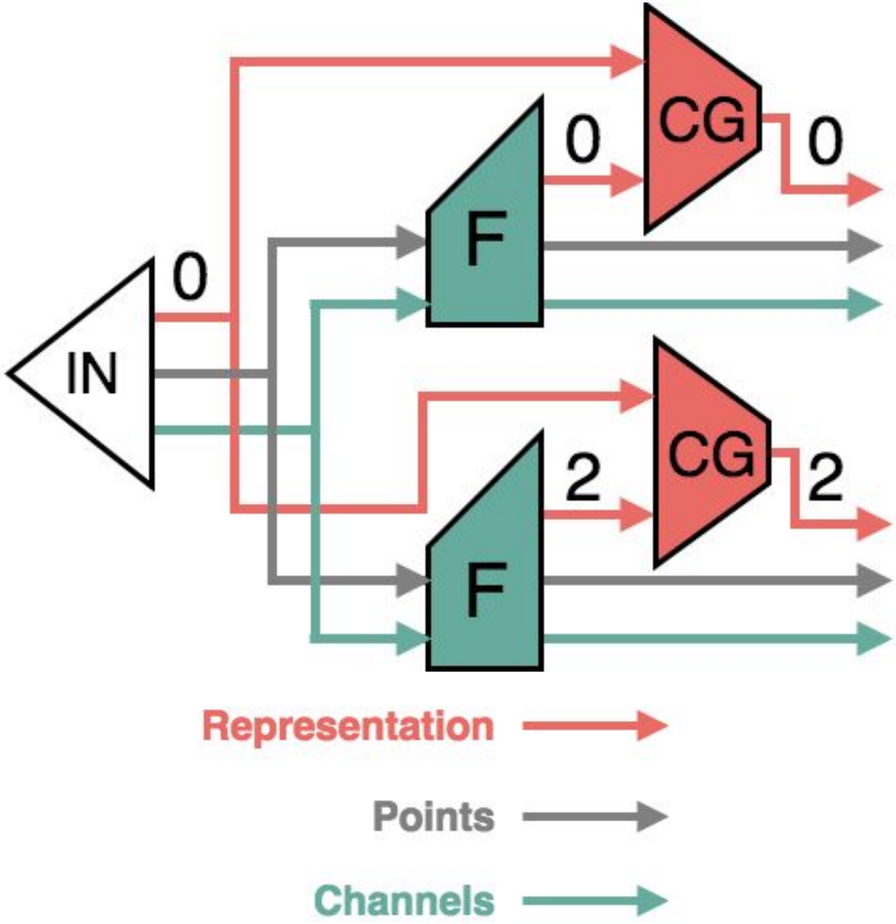
m_1	m_2	2 0	1 0
+1/2	−1/2	1/2	1/2
−1/2	+1/2	1/2	−1/2

m_1	m_2	2 −1	1 −1
−1/2	−1/2	3/4	1/4
−3/2	+1/2	1/4	−3/4

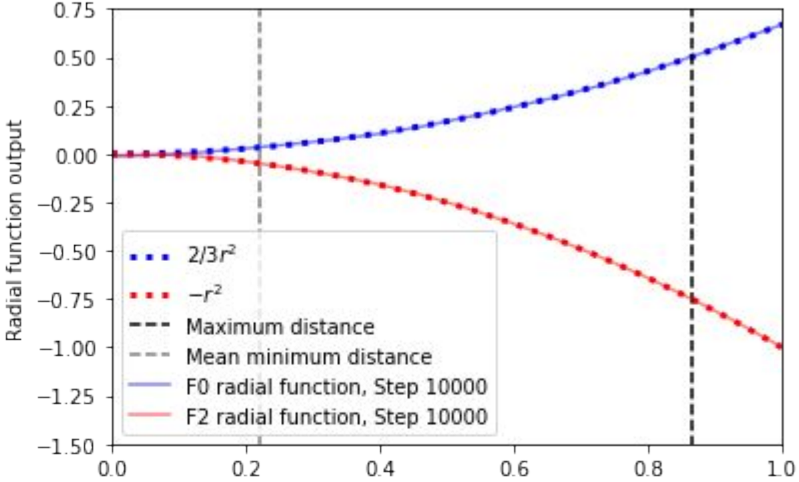
m_1	m_2	2 −2
−3/2	−1/2	1

D. Griffiths, Introduction to quantum mechanics

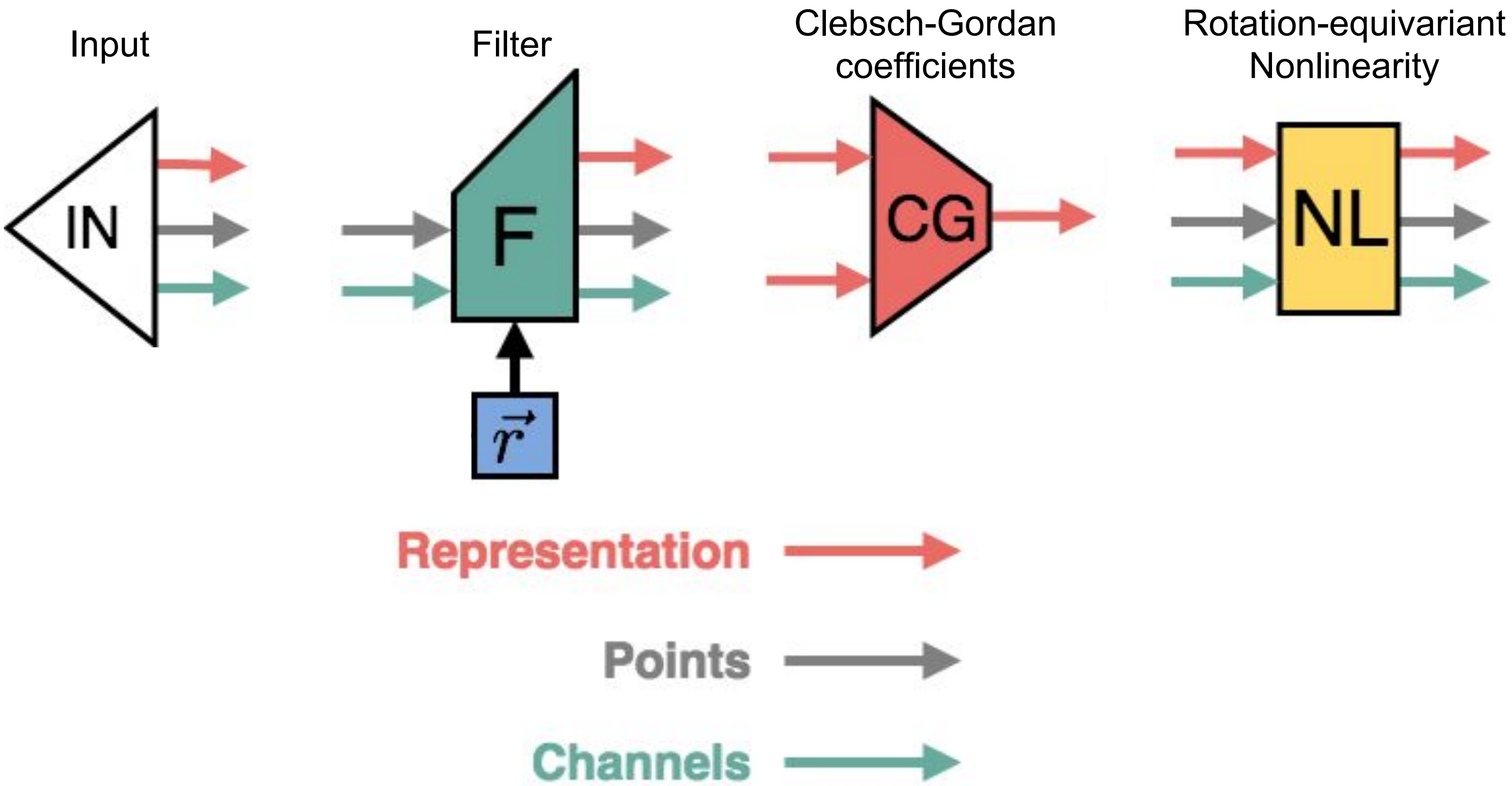
We can start with tensor input of any type and use filters to get tensor output of any type.
In this task, scalar masses are input and the moment of inertia tensor (a symmetric matrix) is output.



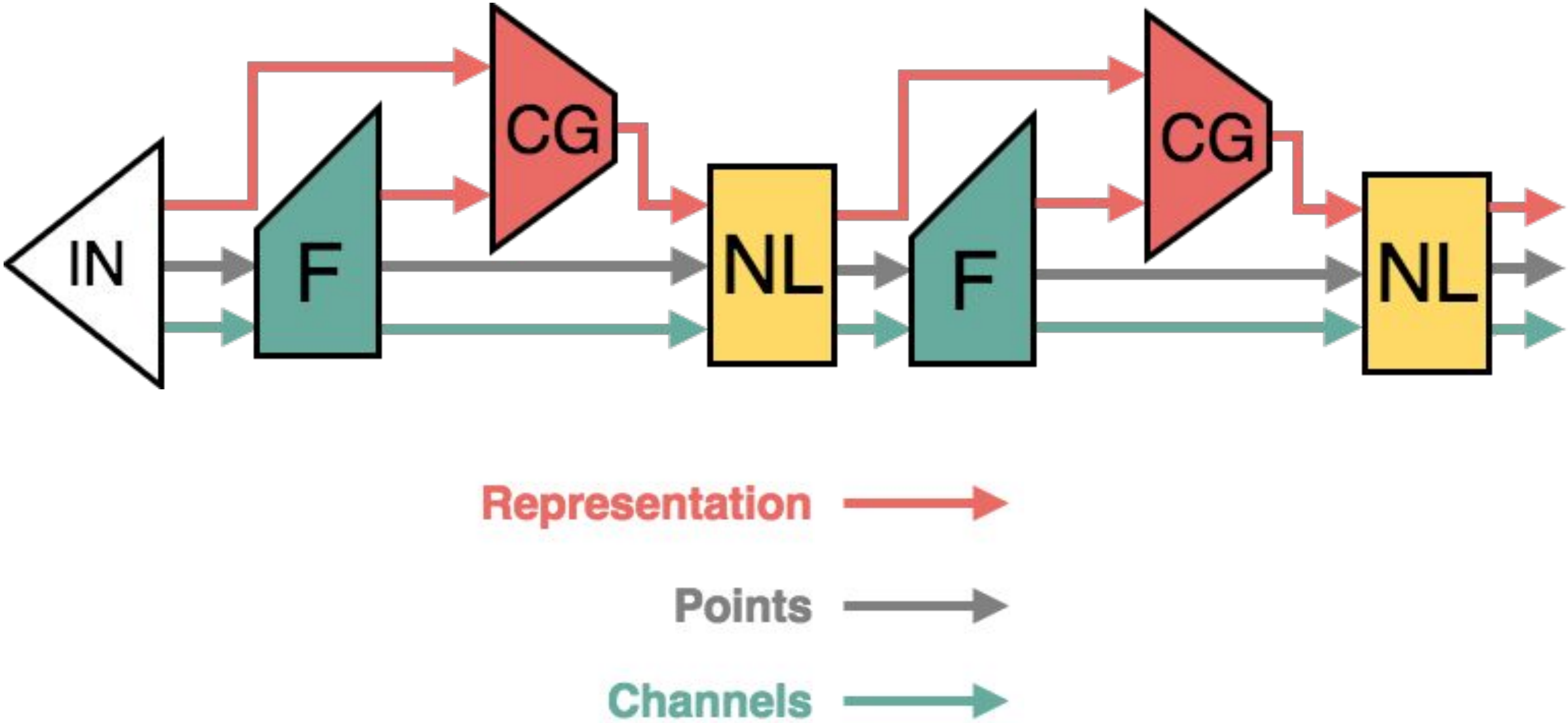
Moment of inertia:
0 (trace) + 2 (symmetric traceless)



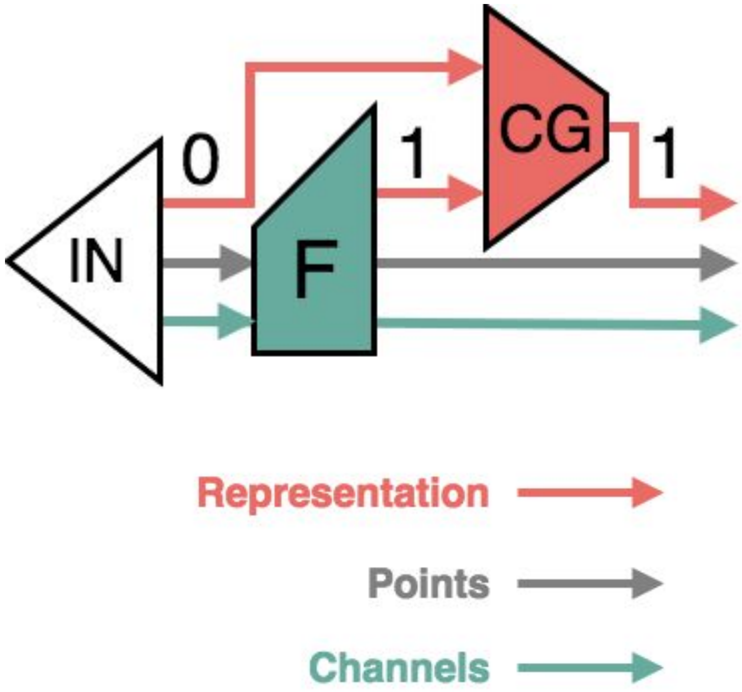
These are components of **tensor field networks**

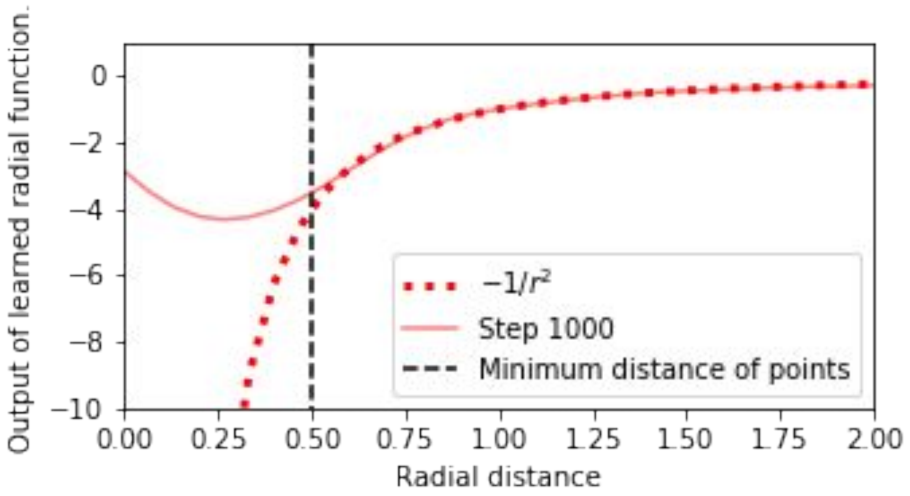


This is what a two-layer tensor field network looks like:

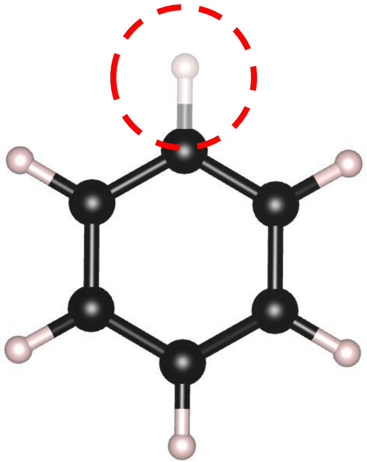


We can start with tensor input of any type and use filters to get tensor output of any type.
In this task, scalar masses are input and gravitational acceleration vectors are output.

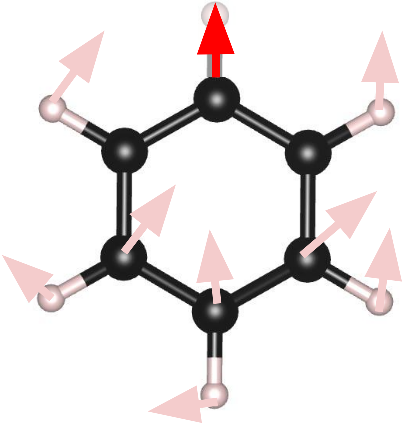




Given a small organic molecule with an atom removed, replace the correct element at the correct location in space.



Input coordinates with missing atom.



Network outputs
(N-1) atom type features (scalars),
(N-1) displacement vectors, and
(N-1) scalars indicating confidence probability used for "voting".

DATASET
QM9: *http://www.quantum-machine.org/datasets/*
134k molecules with 9 or less heavy atoms (non-hydrogen) and elements H, C, N, O, F.

TRAIN
1,000 molecules with 5-18 atoms

TEST
1,000 molecules with 19 atoms
1,000 molecules with 23 atoms
1,000 molecules with 25-29 atoms

Atoms	Number of predictions	Accuracy (%) ($\leq$ 0.5 Å and atom type)	Distance MAE in Å
5-18 (train)	15 947	92.6	0.16
19	19 000	94.7	0.15
23	23 000	96.9	0.14
25-29	25 404	97.8	0.17

Learns to replace atoms with over 90% accuracy across train and test by seeing the same 1,000 molecules 200 times.