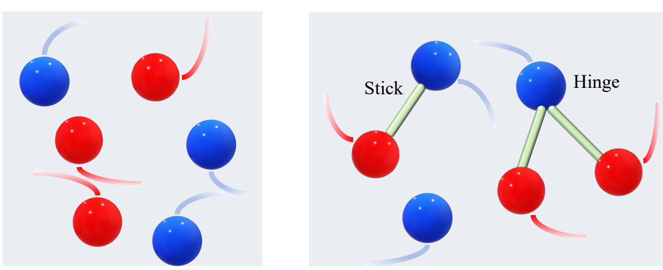
Equivariant Graph Mechanics Networks with Constraints

Wenbing Huang*[1], Jiaqi Han*[2], Yu Rong[3], Tingyang Xu[3], Fuchun Sun[2], Junzhou Huang[4]

AIR, Tsinghua University[1], Tsinghua University[2], Tencent AI Lab[3], University of Texas at Arlington[4], Equal Contributions*

Introduction

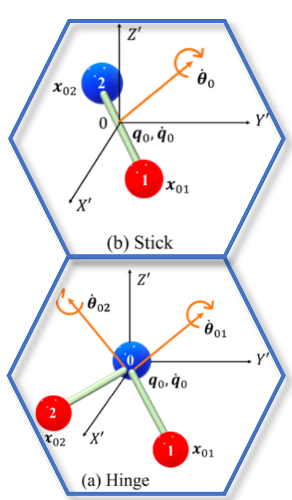
- TL; DR: We design a graph neural network that permits both geometrical constraints and Euclidean equivariance.
- Reasoning about the relations and dynamics of interacting objects is a core aspect of human intelligence.
- We consider the constrained N-body system by involving two crucial inductive biases:
  - Constraint satisfaction
  - Equivariance



(a) N-body System (b) Constrained N-body System

Graph Mechanics Networks

1. Constraint Satisfaction



Cartesian coordinate system

x_{01}, x_{02} $s.t. \|x_{01} - x_{01}\| = L$	2 variables 1 constraint
x_{00}, x_{01}, x_{02} $s.t. \|x_{00} - x_{01}\| = L$ $\|x_{00} - x_{02}\| = L$	3 variables 2 constraints

Generalized coordinate system

q_0, θ_0	2 variables No constraint
$q_0, \theta_{01}, \theta_{02}$	3 variables No constraint

- We propose to map the input Cartesian space into the generalized coordinate system for sticks and hinges.

- The flowchart of GMN:

Generalized coordinates $q_k^{l-1}, \dot{q}_k^{l-1}$

Forward Kinematics: $x_i', v_i' = \mathrm{FK}(q_k, \dot{q}_k)$

Cartesian coordinates $x_i^{l-1}, v_i^{l-1} (\dot{x}_i^{l-1})$

Interaction Force: $f_i, \; h_i = \sum_{j \in \mathcal{N}_i} \varphi_f(x_i, x_j, h_i, h_j)$

Interaction force f_i^l

Inverse Dynamics: $\ddot{q}_k = \varphi_q(f_i, x_{ki}, v_{ki})$

Generalized acceleration $\ddot{q}_k^l$

Coordinate Update: $q_k = q_k + \dot{q}_k, \dot{q}_k = \psi \dot{q}_k + \ddot{q}_k$

Generalized coordinates $q_k^l, \dot{q}_k^l$

- FK (Forward Kinematics):
  - Hand-crafted FK: Physically design the exact FK.
  - Learnable FK: Learn the FK with GMN layers.

2. Equivariant Message Passing

We derive a universal form of orthogonality-equivariant functions on matrix inputs. For a stack of dim-d vectors $Z \in \mathbb{R}^{d \times m}$,

$$\varphi(Z) = Z\sigma(Z^\top Z)$$

We present the following theorem, implying the universality.

If $m \geq d$ and $rank(Z) = d$, then for any continuous orthogonality-equivariant function $\hat{\varphi}(Z)$, there must exist an MLP σ_W satisfying $\|\varphi(Z) - \hat{\varphi}(Z)\|$ for arbitrarily small error ϵ.

We apply this formulation to φ and ρ, ensuring equivariance.

$$\ddot{q}_k = \sum_{i \in \mathcal{O}_k} \varphi_q(f_i, x_{ki}, v_{ki}) \quad \text{(Generalized acceleration)}$$

$$\ddot{x}_k = \rho(\ddot{q}_k, x_{ki}, f_i) \quad \text{(Learnable FK)}$$

Experiments

We evaluate the efficacy of GMN in three scenarios.

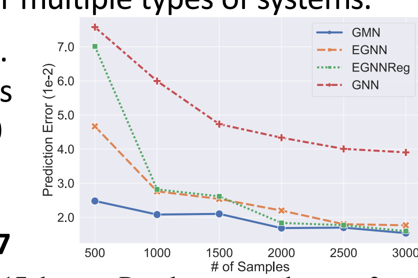
1. Simulation dataset: Constrained N-body

Table 1: Prediction error ($\times 10^{-2}$) on various types of systems. The header of each column "p, s, h" denotes the scenario with p isolated particles, s sticks and h hinges. Results averaged across 3 runs.

	|Train| = 500					|Train| = 1500				
	1,2,0	2,0,1	3,2,1	0,10,0	5,3,3	1,2,0	2,0,1	3,2,1	0,10,0	5,3,3
Linear	8.23±0.00	7.55±0.00	9.76±0.00	11.36±0.00	11.62±0.00	8.22±0.00	7.55±0.00	9.76±0.00	11.36±0.00	11.62±0.00
GNN	5.33±0.07	5.01±0.08	7.58±0.08	9.83±0.04	9.77±0.02	3.61±0.13	3.23±0.07	4.73±0.11	7.97±0.44	7.91±0.31
TFN	11.54±0.38	9.87±0.27	11.66±0.08	13.43±0.31	12.23±0.12	5.86±0.35	4.97±0.23	8.51±0.14	11.21±0.21	10.75±0.08
SE(3)-Tr.	5.54±0.06	5.14±0.03	8.95±0.04	11.42±0.01	11.59±0.01	5.02±0.03	4.68±0.05	8.39±0.02	10.82±0.03	10.85±0.02
RF	3.50±0.17	3.07±0.24	5.25±0.44	7.59±0.25	7.73±0.39	2.97±0.15	2.19±0.11	3.80±0.25	5.71±0.31	5.66±0.27
EGNN	<u>2.81</u>±0.12	<u>2.27</u>±0.04	<u>4.67</u>±0.07	<u>4.75</u>±0.05	<u>4.59</u>±0.07	<u>2.59</u>±0.10	1.86±0.02	<u>2.54</u>±0.01	<u>2.79</u>±0.04	3.25±0.07
EGNNReg	2.94±0.01	2.66±0.06	7.01±0.34	5.03±0.08	6.31±0.04	2.74±0.08	<u>1.58</u>±0.03	2.62±0.05	3.03±0.07	<u>3.07</u>±0.04
GMN	**1.84**±0.02	**2.02**±0.02	**2.48**±0.04	**2.92**±0.04	**4.08**±0.03	**1.68**±0.04	**1.47**±0.03	**2.10**±0.04	**2.32**±0.02	**2.86**±0.01

GMN performs the best over multiple types of systems.

GMN is highly data-efficient. The prediction error remains very low even with only 500 training samples.



2. Molecular Dynamics: MD17

Table 4: Prediction error ($\times 10^{-2}$) on MD17 dataset. Results averaged across 3 runs.

	Aspirin	Benzene	Ethanol	Malonaldehyde	Naphthalene	Salicylic	Toluene	Uracil
RF	10.94±0.01	103.72±1.29	<u>4.64</u>±0.01	13.93±0.03	0.50±0.01	1.23±0.01	10.93±0.04	<u>0.64</u>±0.01
TFN	12.37±0.18	58.48±1.98	4.81±0.04	13.62±0.08	0.49±0.01	1.03±0.02	10.89±0.01	0.84±0.02
SE(3)-Tr.	11.12±0.06	68.11±0.67	4.74±0.13	13.89±0.02	0.52±0.01	1.13±0.02	10.88±0.06	0.79±0.02
EGNN	14.41±0.15	62.40±0.53	<u>4.64</u>±0.01	13.64±0.01	0.47±0.02	1.02±0.02	11.78±0.07	<u>0.64</u>±0.01
EGNNReg	13.82±0.19	61.68±0.37	6.06±0.01	13.49±0.06	0.63±0.01	1.68±0.01	11.05±0.01	0.66±0.01
GMN	<u>10.14</u>±0.03	**48.12**±0.40	4.83±0.01	<u>13.11</u>±0.03	**0.40**±0.01	<u>0.91</u>±0.01	**10.22**±0.08	**0.59**±0.01
GMN-L	**9.76**±0.11	<u>54.17</u>±0.69	**4.63**±0.01	**12.82**±0.03	<u>0.41</u>±0.01	**0.88**±0.01	<u>10.45</u>±0.04	**0.59**±0.01

3. Human Motion: CMU Motion Capture

Table 5: Prediction error ($\times 10^{-2}$) on motion capture. Results averaged across 3 runs.

GNN	TFN	SE(3)-Tr.	RF	EGNN	EGNNReg	GMN	GMN-L
67.3±1.1	66.9±2.7	60.9±0.9	197.0±1.0	59.1±2.1	59.5±2.2	**43.9**±1.1	<u>50.9</u>±0.7

4. Visualization