TU Informatics

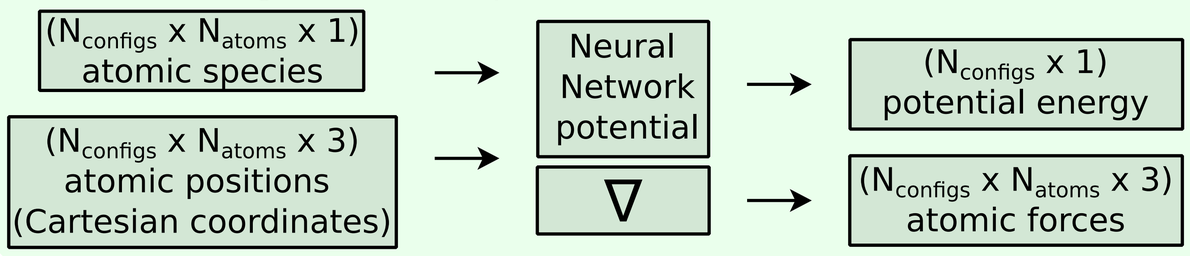
Sparse graph attention networks as efficient ionic liquid potentials

Stefan Hödl
Data Science

TU Wien Informatics
Institute of Materials Chemistry
Theoretical Materials Chemistry
Supervisor: Univ.Prof. Dr. Allan Hanbury
Co-Supervisor: Dr. Jesús Carrete Montaña
Co-Supervisor: Univ.Prof. Georg Madsen
Contact: stefano.hoedl@gmail.com

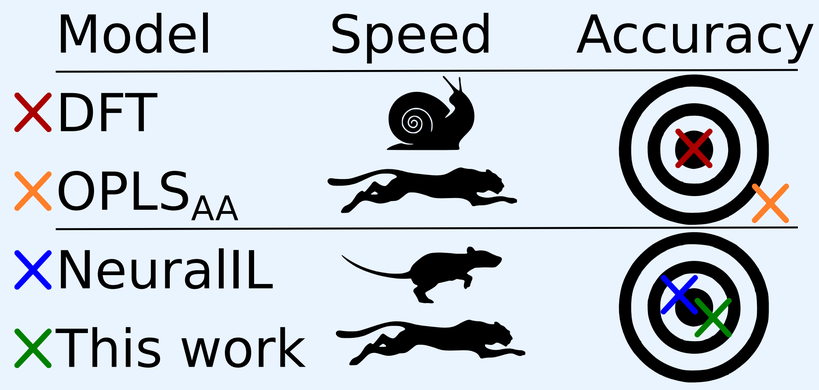
Problem context: What's a potential?

Domain: Computational (molecular) chemistry
Goal: Predict the energy of molecules
Task: Learn the potential energy surface (atomic positions → energy mapping)
Problem: "Exponential barrier" of electronic contribution, high-dimensional task (3Natoms-6 degrees of freedom)
Solution: Approximate the Schrödinger Equation using efficient Deep Neural Networks

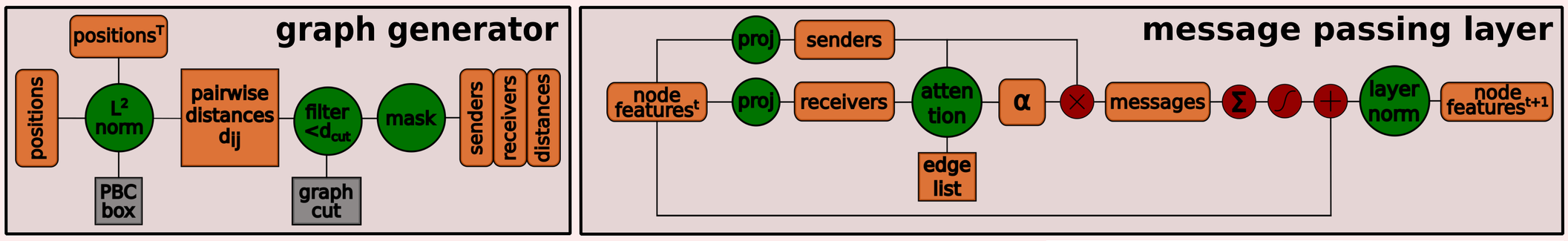


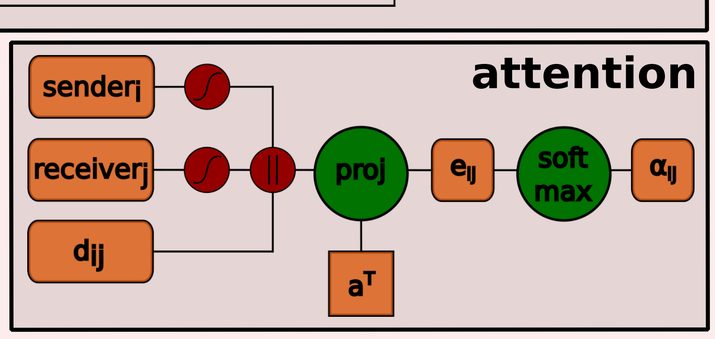
... but why? Molecular dynamics!

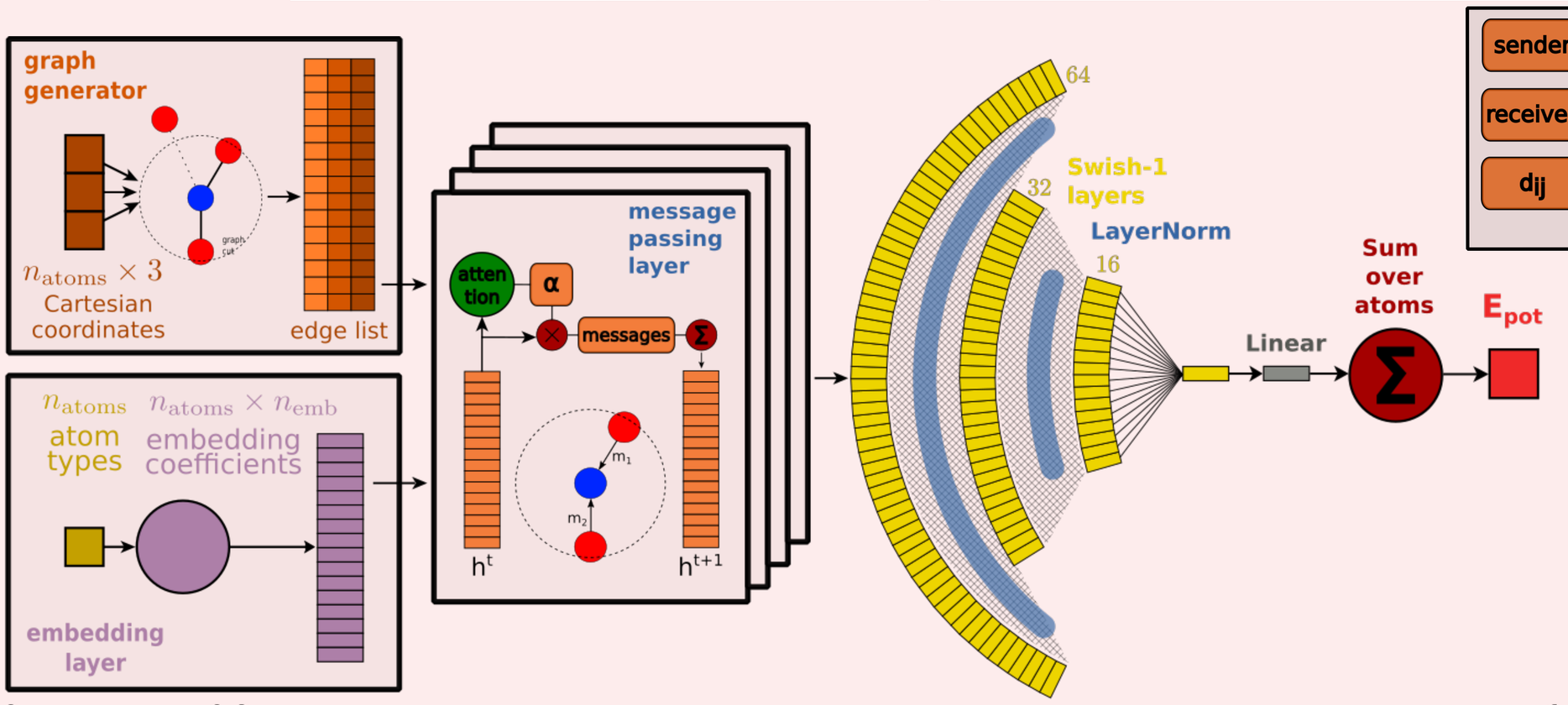
Obtain $3N_{atoms}$ forces as gradient of energy w.r.t. positions
- Use atomic forces to integrate Newton's equations of motion
- requires very small timestep Δt ~ fs, (10^{-15} s)
- 1M+ steps to watch something interesting happen...
- thus 1M+ model evaluations, making efficiency crucial
- model() & grad() for value (energy) and grad (forces)



Methods & Architecture







$$\left\{\frac{\partial E_{\text{pot}}}{\partial \mathbf{p}_\alpha}\right\}$$

$$\text{VJP}_i(\{\mathbf{r}\}, \square) = \sum_\alpha \square_\alpha \cdot \frac{\partial \mathbf{p}_\alpha}{\partial \mathbf{r}_i} \rightarrow \boxed{\mathbf{f}_i}$$

Approach & Architecture
- Adapt NeuralIL[1] implementation
- replace Bessel descriptors with a Message Passing Neural Network[2]
- with a sparse molecular graph
- using Graph Attention Networks[3]

Going fast
- replace expensive descriptors
- JAX, GPU & forces using VJP
- just-in-time (JIT) compilation
- sparse graph & self-attention

Features
- SE(3) + permutation invariant
- Periodic graph & PBC
- additive atomic contributions
- for energy and forces
- Locality for MD parallelization

[1] Montes-Campos, Hadrián, et al. "A Differentiable Neural-Network Force Field for Ionic Liquids." Journal of chemical information and modeling 62.1 (2021): 88-101.
[2] Gilmer, Justin, et al. "Neural message passing for quantum chemistry." International conference on machine learning. PMLR, 2017.
[3] Velickovic, Petar, et al. "Graph attention networks." stat 1050 (2017): 20.

Datasets

Ionic liquid EAN
- salt in liquid phase at room temperature
- 15 EAN pairs of anion (NO_3^-) & cation ($C_2H_5NH_3^+$) → 225 atoms
- 741 configurations sampled from OPLS-AA MD trajectory
- reference energy & forces calculated using DFT
- training on $3N_{atoms}$ atomic forces

ANI-1
- ~57K organic molecules in displaced states
- max. 29 {C, N, O, H} atoms with 1 - 8 heavy (non-H) atoms
- ~21M configurations sampled from vibration normal modes
- Sample 5% - 10% of subsets 1-7 to train & develop architecture
- final run: 5% of ANI-1, ~900K train, ~100K validation and test
- all $C_2H_6N_2$ and $C_4H_4N_4$ isomers excluded for separate evaluation
- training only on energy, DFT reference energies

Results

EAN
- best model uses 7 JAT layers, 1 head of dim 48
- 3K epochs in ~4:30h on GPU, ~100 configurations/sec
- forces MAE ~80-88 meV/Å, RMSE ~140-170 meV/Å
- 4x speed up against NeuralIL baseline
- reaches chemical accuracy but does not outperform baseline
- predicts dissociation

ANI-1
- best model uses 5 JAT layers, 4 heads of dim 48
- 80 epochs in ~9h on GPU, ~3300 configuration/sec
- energy MAE ~16-30 meV/atom (validation & test set)
- reaches chemical accuracy, especially for larger molecules
- MAE ~42 meV/atom on $C_2H_6N_2$, ~18 meV/atom on $C_4H_4N_4$
- excellent results scaling to larger training sets