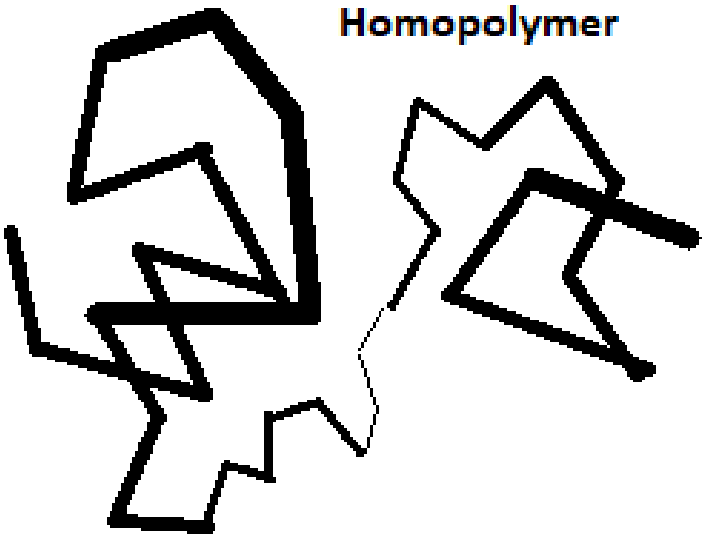
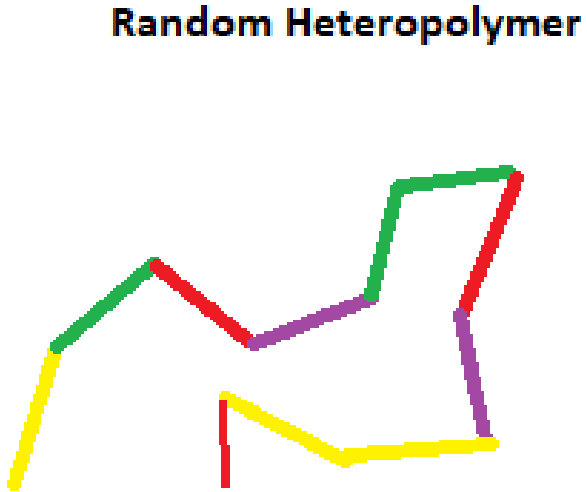
Polymer Physics: Representational Units & generative models. Outlook

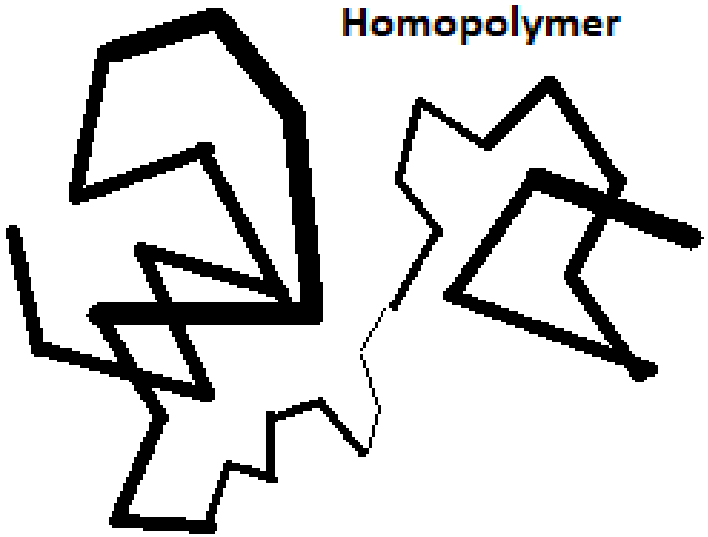
- Interest in polymers and need of computational models
- Generative models: Classical and attempts in integrating ML
- Involvement in Projects
  - Cross validation of Coarser models to Atomistic ones
  - Coupling of fluctuating local nematic ordering with density

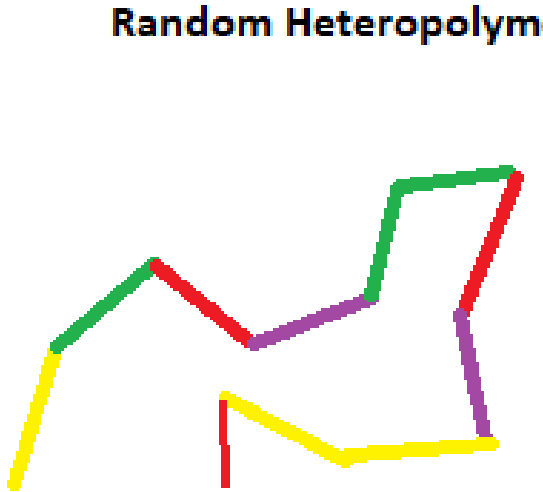
Polymer structure

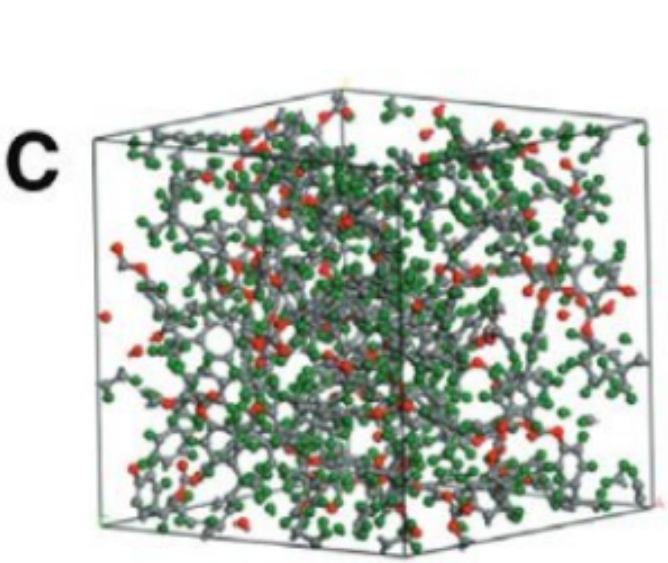


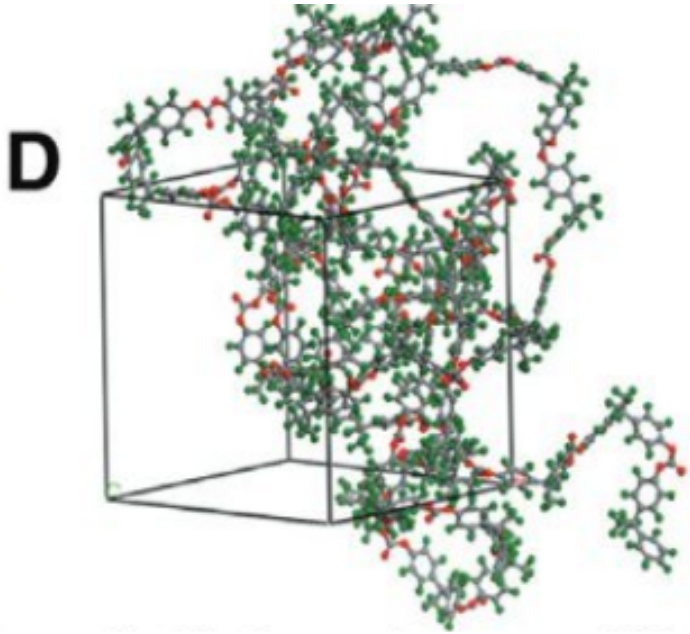


Polymer structure









Interest in polymers

Polymers are everywhere

Interest in polymers

Polymers are everywhere

Functions

- Packaging and commodity applications

Interest in polymers

Polymers are everywhere

Functions

- Packaging and commodity applications
- high end engineering applications

Interest in polymers

Polymers are everywhere

Functions

- Packaging and commodity applications
- high end engineering applications
- biomolecules

Interest in polymers

Polymers are everywhere

Functions

- Packaging and commodity applications
- high end engineering applications
- biomolecules

Visuals

Interest in polymers

Polymers are everywhere

Functions

- Packaging and commodity applications
- high end engineering applications
- biomolecules

Visuals

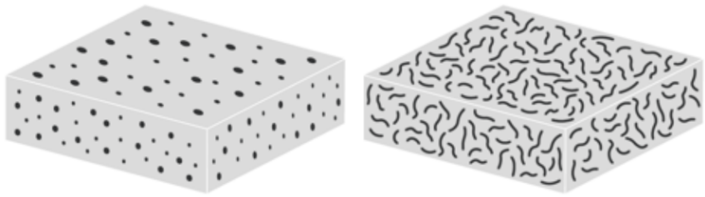
Interest in polymers

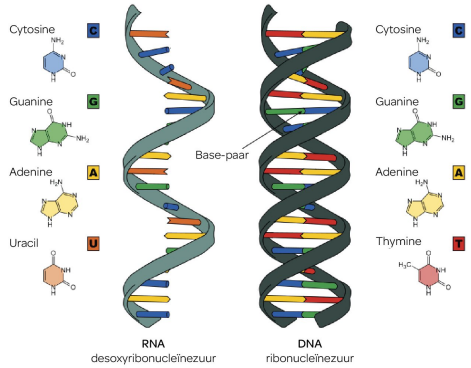
Polymers are everywhere

Functions

- Packaging and commodity applications
- high end engineering applications
- biomolecules

Visuals





Need for computational models

Limited experimental data on polymers due to their nature.
Assumptions on their configurations by their dynamics.

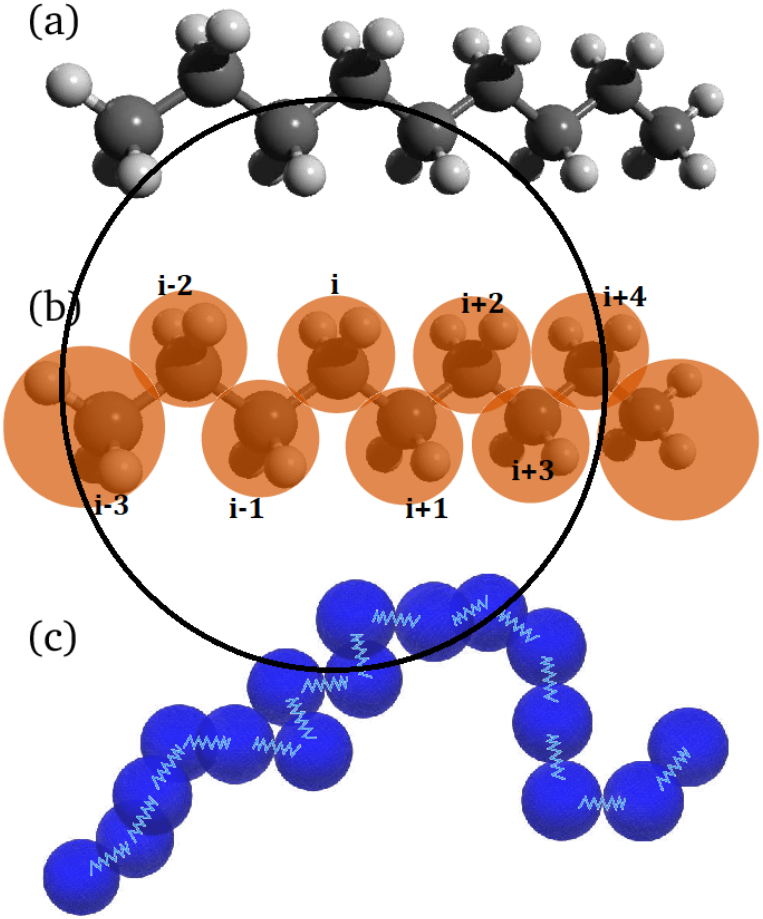
- Measurable quantities: Viscosities, bonded distributions and structure factor

Need for computational models

Limited experimental data on polymers due to their nature.
Assumptions on their configurations by their dynamics.

- Measurable quantities: Viscosities, bonded distributions and structure factor
- Indirect calculations for all other quantities

Representational Units



(d)

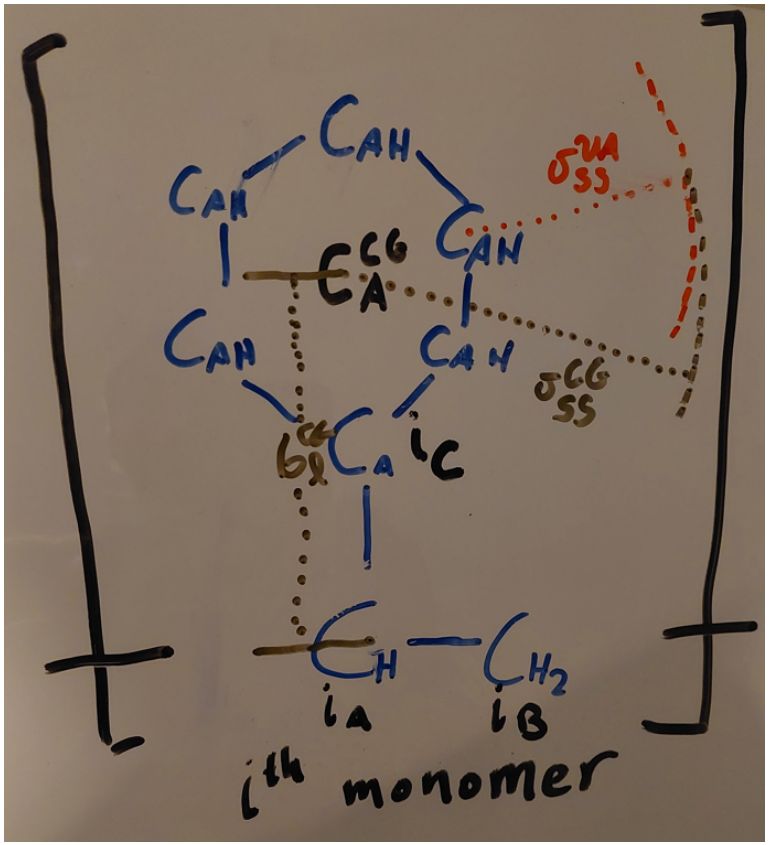


Figure: {a,b,c} are different representation scales of PE: (a) all-atoms (b) United-Atoms (c) CG models, (d) Representation of PS

Force Field at UA representation

- Bonded length Interactions
  $U_{b_l}(r_{ij}) = \frac{1}{2} k_{ij}(r_{ij} - r_{0_{ij}})^2$
- Bond angle interactions
  $U_{b_a}(\theta_{ijk}) = \frac{1}{2} k_{ijk}(\theta_{ijk} - \theta_{0_{ijk}})^2$
- Dihedral Angle interactions
  $U_{b_{dih}}(\phi_{ijkl}) =$
  $\sum_{n=0}^{3} C_{ijkl_n}(\cos(\phi_{ijkl}))^n$
- Intermolecular Interactions - LJ potential
  $U_{LJ}(r) = 4\epsilon_{ij}((\frac{\sigma_{ij}}{r})^{12} - (\frac{\sigma_{ij}}{r})^6)$

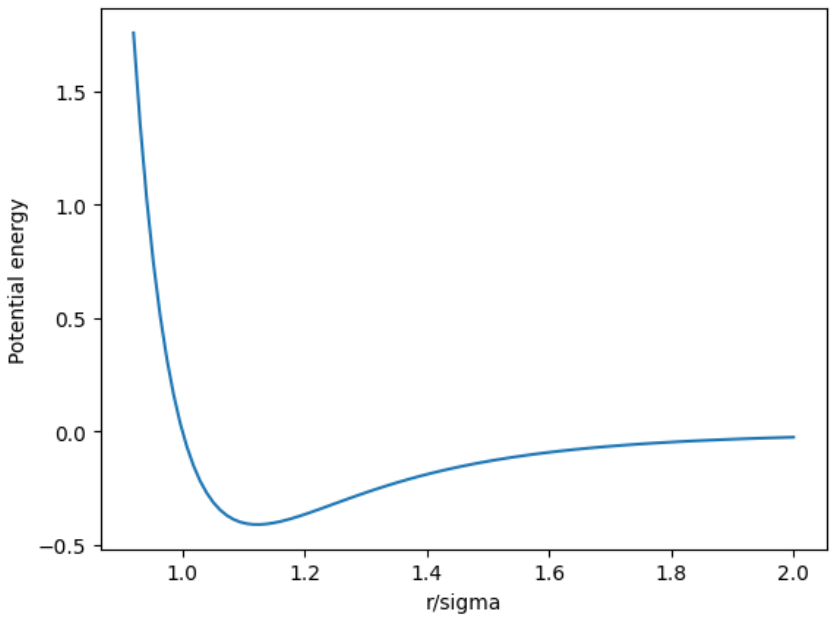


Figure: LJ potential

Phantom Chain Generation

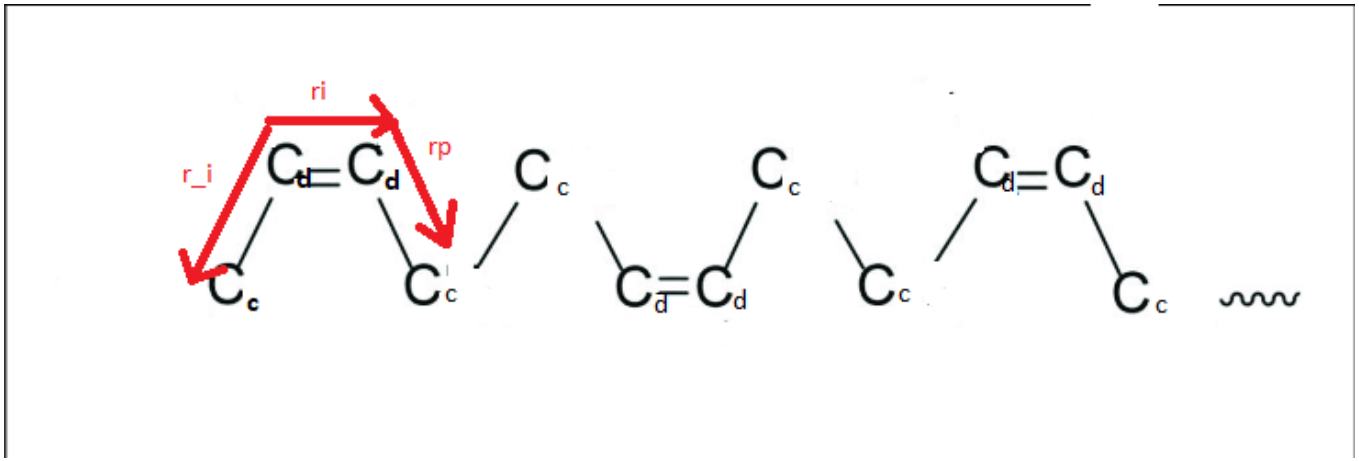
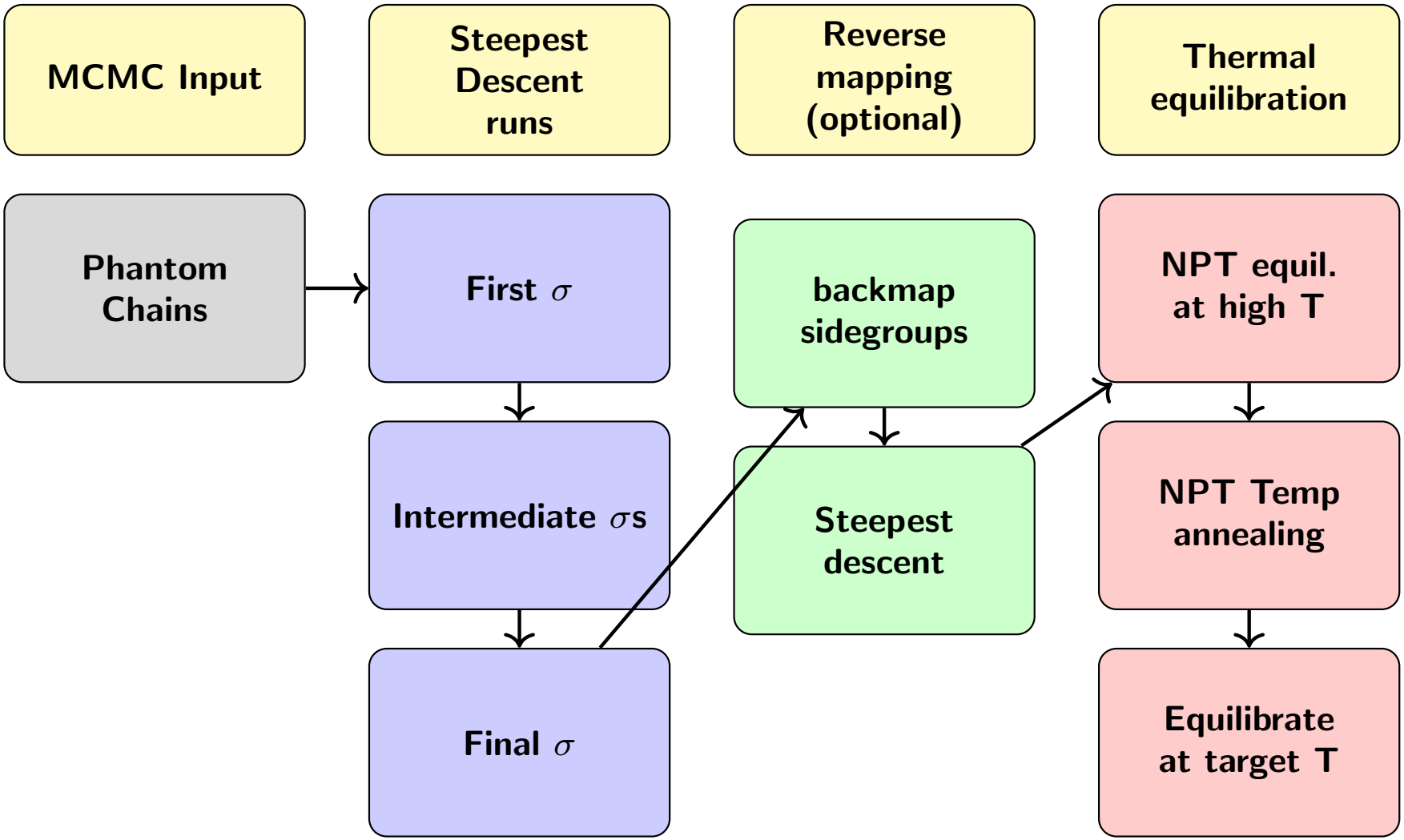


Figure: Representation of 3 monomers

- from bond length $r_{px}^2 + r_{py}^2 + r_{pz}^2 = r_p^2$
- from bond angle $-r_{ix} r_{px} - r_{iy} r_{py} - r_{iz} r_{pz} = r_p r_i \cos\theta_a$
- from dihedral angle $n_{1x} n_{2x} + n_{1y} n_{2y} + n_{1z} n_{2z} = |n_1||n_2| \cos\theta_d = |r_{-i}| \cdot |r_i|^2 \cdot |r_p| \sin\theta_{ai} \sin\theta_a \cos\theta_d$

Generation of phantom chains in a given box, to respect target density only and ignore inter molecular interactions, but respect intramolecular ones. Bonded and internal distances

Classical generative Pipeline



Conditionals: Acceptable and equilibrated

Acceptable

$$g^{(2)}_{N_{ij}}(r) = \frac{1}{N_i \rho_j} \frac{N_{in}(r)}{4\pi r^2 dr}$$

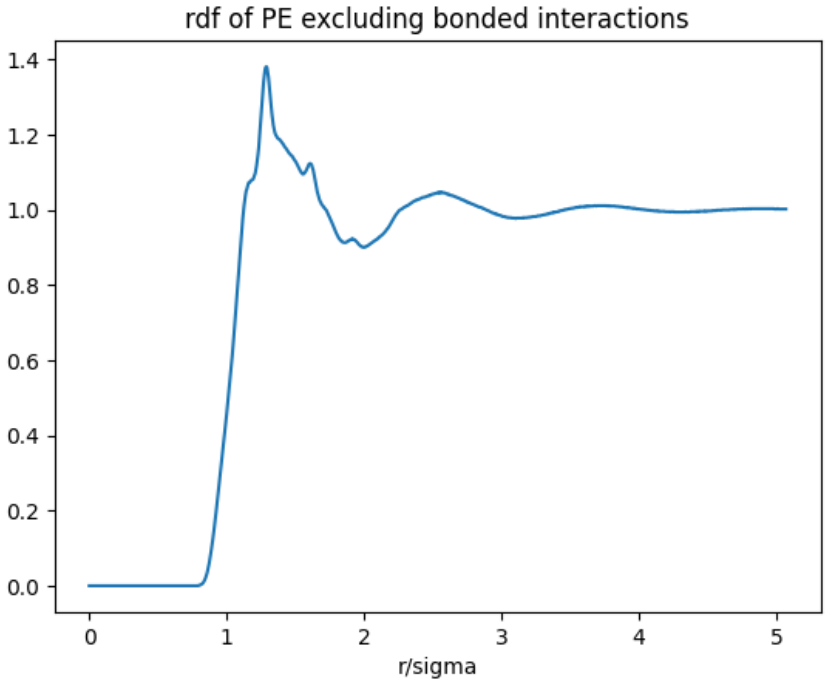


Figure: RDF of PE

Indicators of equilibration

$$R(n) = c \cdot \langle b \rangle \cdot \sqrt{n}$$

$$C(n) = \frac{\langle R^2 \rangle}{n \cdot \langle b \rangle^2}$$

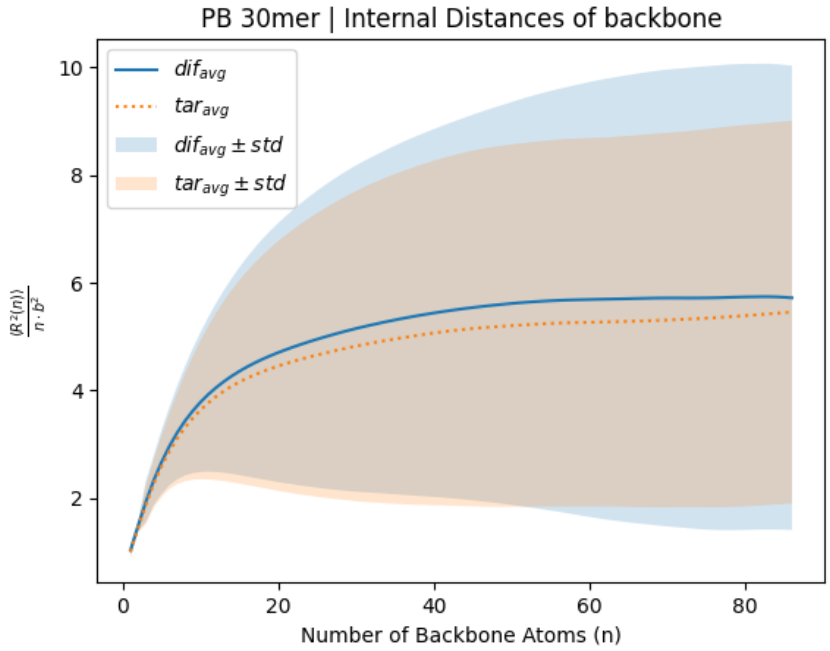


Figure: Internal distances of PB

Comments on pipeline

- This Pipeline works, regardless of system size.
- Modifications on the initial σ values may need to be considered depending on system size.
- Number of steps for thermal equilibration depend on system size and Internal distances distribution of Phantom Chains
- Despite considerable speedup, generation of large systems is still computationally expensive (even week per sample)

Generated systems usually contain from $20 \cdot 10^3$ to 10^6 particles

Integration of ML on the generative pipeline - Diffusion Models

credits to Dr Eleftherios Christofi
Generic diffusion equation with scheduled noice

$$q(\mathbf{z}_{1:T}|\mathbf{z}_0) = \prod_{t=1}^{T} q(\mathbf{z}_t|\mathbf{z}_{t-1}), \quad q(\mathbf{z}_t|\mathbf{z}_{t-1}) = \mathcal{N}(\mathbf{z}_t; \sqrt{1-\beta_t}\mathbf{z}_{t-1}, \beta_t\mathbf{I})$$

$q \in \mathcal{R}^{3N}$, $c \in \mathcal{N}^{M}$

- q - set of cartesian coordinates of N particles
- c - set of M monomers in the given sequence
- we train on the conditional probability of $P(\mathbf{q}|\mathbf{c})$

Generating the set of vectors, describing the polymer

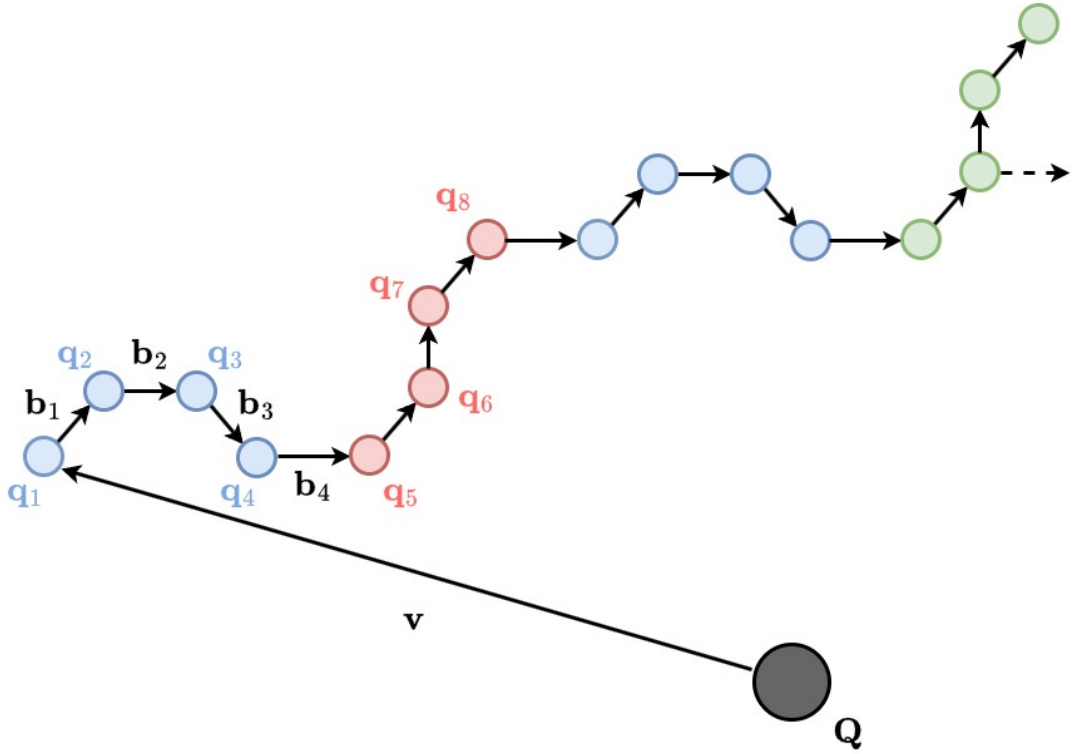


Figure: Schematic illustration of a tetramer section of a PB copolymer. The different colors represent the different monomer types. For *cis*-1,4 monomers the color is blue, for *trans*-1,4, it is red, and for *vinyl*-1,2, it is red.

Diffusion results

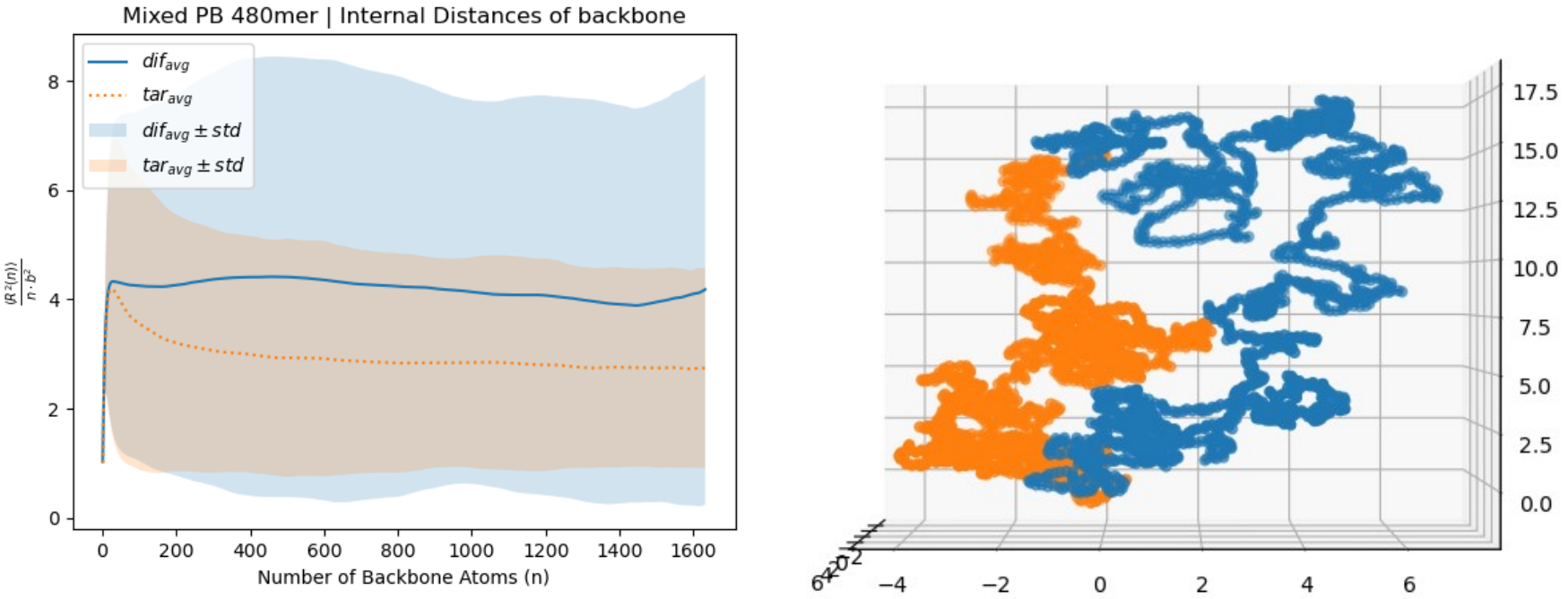


Figure: Internal distances of all mixed-PB 480mer and 3d representation of a random monomeric sequence. Blue is diffused data and orange is samples.

Comments on results

- By training on sets of 30mers, we managed to generate single molecules of arbitrary size, as long as their monomeric composition was part of the training set.
- The architecture model did not work on multiple molecules, despite adding more terms to the loss function (distogram of NxN dimensions, computational limiting factor)
- Need for different architectures on top of the diffusion model (innovation of AF3 still has limitations).

A reduction of dimensionality is also essential

Kremer Grest: a CG model of arbitrary scale

- KG is a model working on arbitrary dimensions per polymer type
- It connects to mean field theories of Polymer Physics as to replicate measurables of any given polymer type
- Adjustable temporal and spatial mapping to well studied Kuhn scale.
- Ability to replicate dynamics of entangled polymer melts

Scope of my take on this project is to relate atomistic simulations onto the KG model, accounting for the chemical specificity on the atomic scale that it lacks.

Spatial mapping

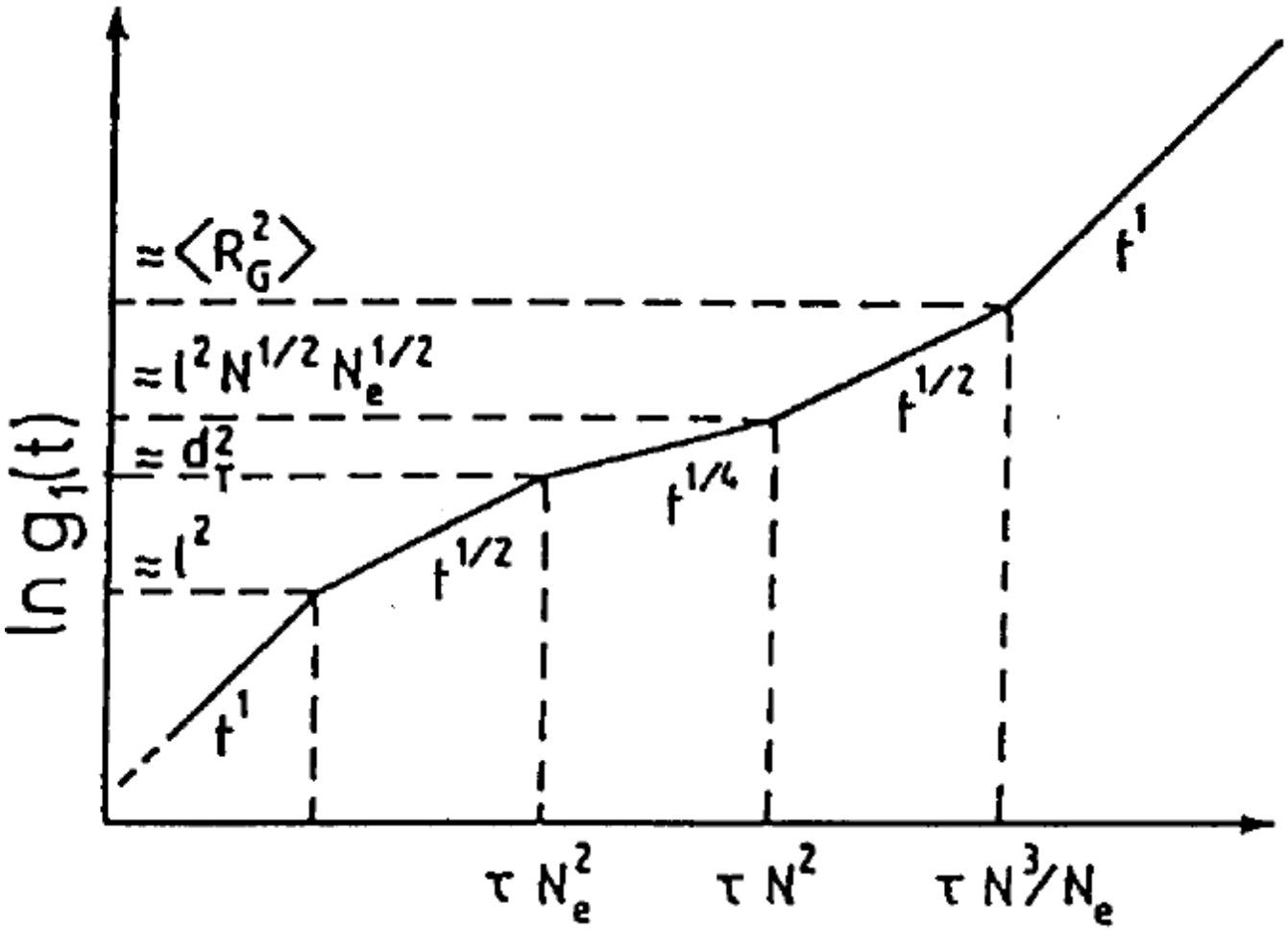
Definition of Kuhn Length l_K, and Kuhn Segments per Chain N_K and Kuhn number n_K.
Their model depends on n_K, which is the number of Kuhn segments per cube of dimension l_K.

$$l_K = \frac{\langle R^2(n) \rangle}{(n-1) \cdot b_l \cdot cos(\pi - \theta)}, \quad N_K = \frac{\langle R^2(n) \rangle}{{l_K}^2}, \quad n_K = \rho_K \cdot {l_K}^3$$

Through various experiments, they managed to

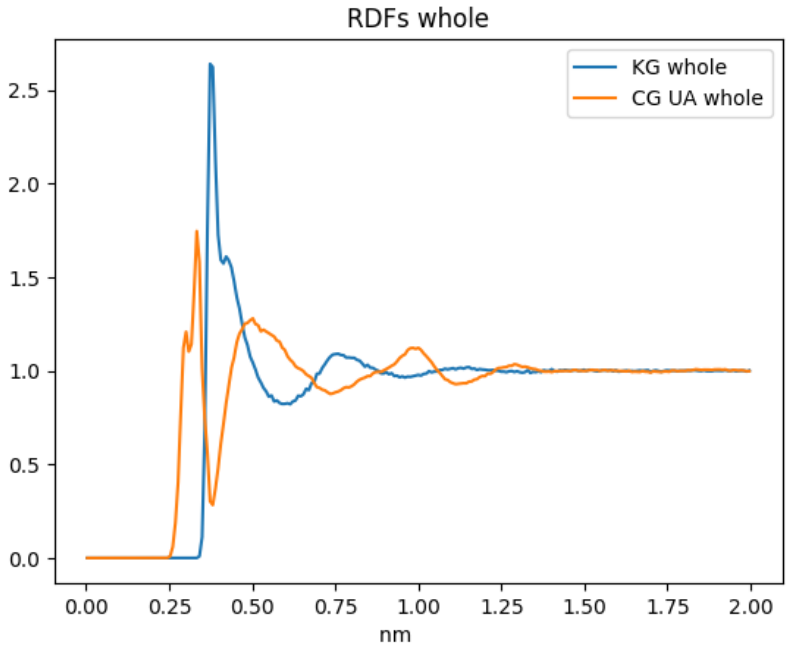
Temporal Mapping

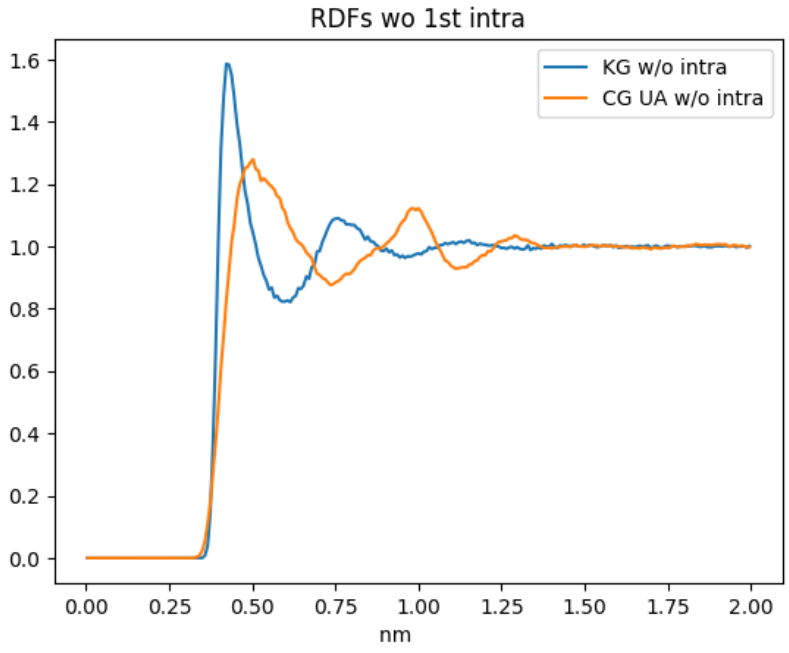
On a given entangled polymer systems, the diffusion dynamics obey certain power laws.



τ_ϵ can be calculated experimentally, and the other regions could be

Structural validation





Remaining progress on the KG model

- Validate structural aurocorrelations
- Run long simulations on systems way beyond the entanglement to validate dynamics
- possibly use such CG to reduce dimensionality of the ML models

Local Density-Nematic ordering flunctuations

Collaboration with Praprotnik M., Svensek D., Socan Y.
Provide generated data for them to run their analysis.
Data on PS (10^6) atoms per configuration
Data on PEO (10^5) atoms per configuration

In a nutshell

- Modernised Classical methods for generation of arbitrary polymer melts, of up to moderate topology complexity at the UA representation.
- Generate Phantom chains of arbitrary length via diffusion model
- Still need to
  - Validate KG model at structrural and dynamical level (and possibly use such representations to)
  - manage to create Bulk polymers via a new architecture (Diffusion within a Transformer model)

Many thanks to

You all for your attention

The ENGAGE program has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement number 101034267.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 810660