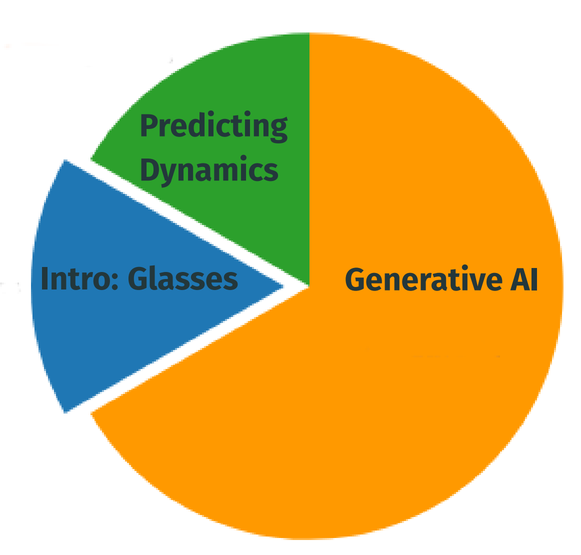
Can Generative AI Models Efficiently Sample Deeply Supercooled Liquids?

Gerhard Jung

Beijing, November 13th, 2024

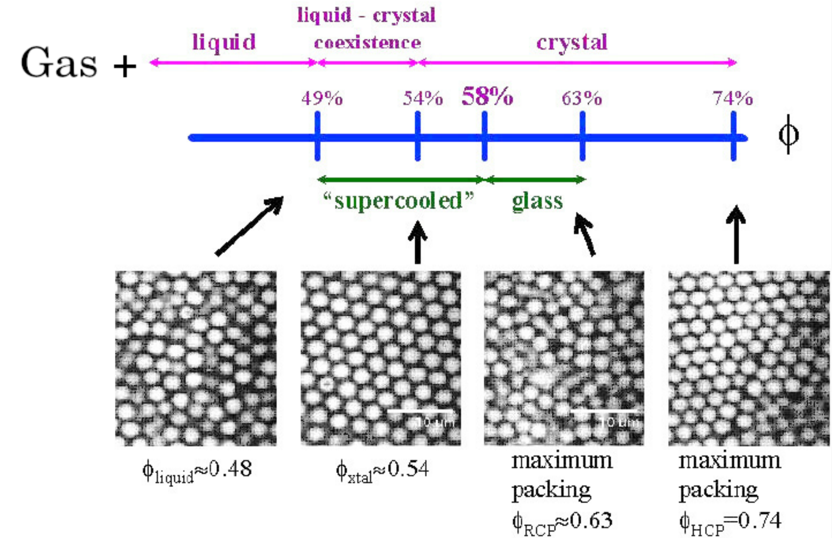
Artificial Intelligence for Theoretical Sciences

Laboratoire interdisciplinaire de Physique



Take-home message: Benchmarking is essential!

Introduction : Glass Physics



Gas

- Low density
  $\rightarrow$ Small parameters for statical and dynamical theories

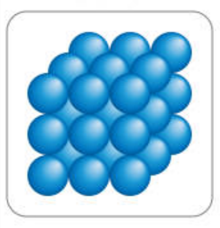
Very well understood

Gas

- Low density
  → Small parameters for statical and dynamical theories

Very well understood

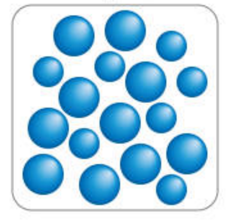
Crystal



- Small fluctuations around periodic crystal structure
- Relaxation around conserved dislocations (defects)

Well understood

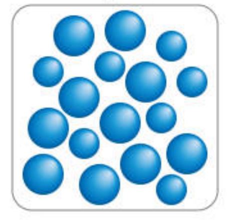
Liquid



There is still no general first-principle theory of liquids

- Amorphous structure
- No small parameter

Liquid



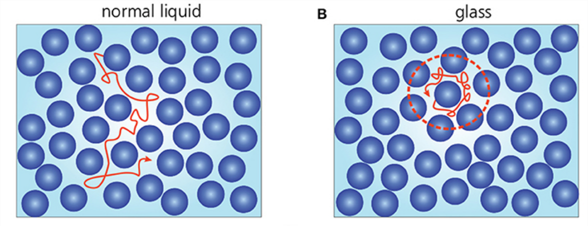
There is still no general first-principle theory of liquids

- Amorphous structure
- No small parameter

Nevertheless, well understood

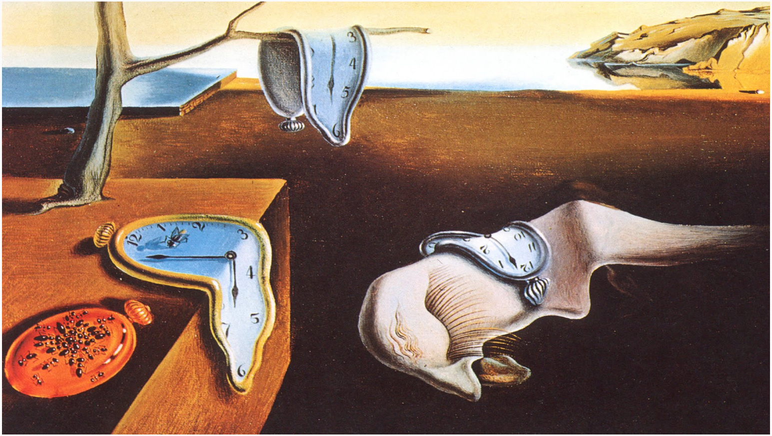
- Density-function theory (structure)
- Mode-coupling theory (dynamics)
- **Computer simulations**

Supercooled Liquids and Glasses



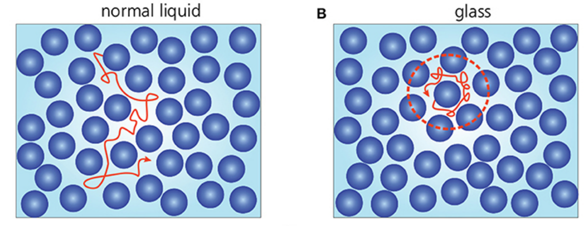
L. Janssen 6, 00097 (2018)

- **Supercooled liquid**: strongly cooled/compressed liquid
- Viscosity increases by orders of magnitude
- Glass (solid with amorphous structure)



Salvador Dali, The persistence of memory (1931)

Supercooled Liquids and Glasses

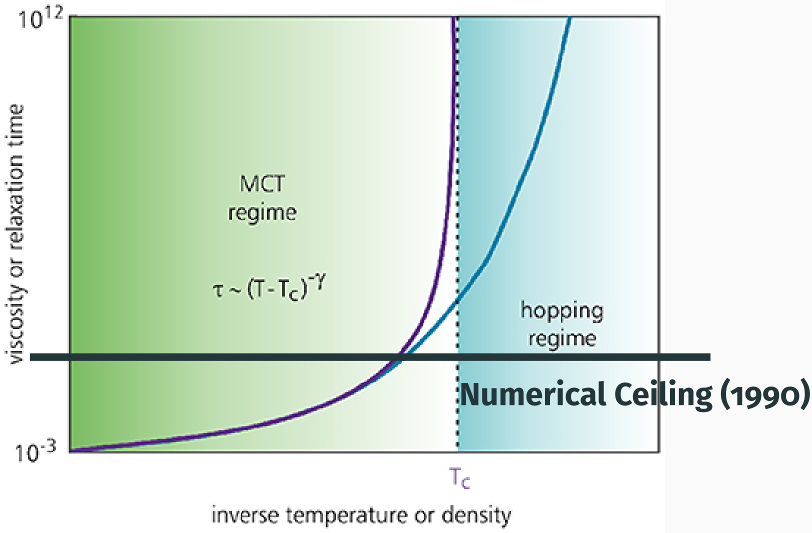


L. Janssen 6, 00097 (2018)

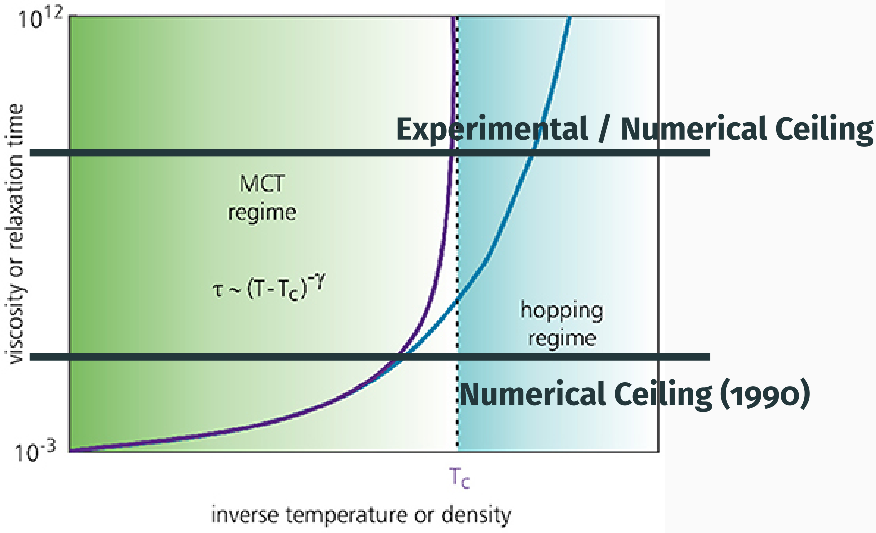
- **Supercooled liquid**: strongly cooled/compressed liquid
- Viscosity increases by orders of magnitude
- Glass (solid with amorphous structure)

Question 1: Is there an ideal, equilibrium, truly solid glass?

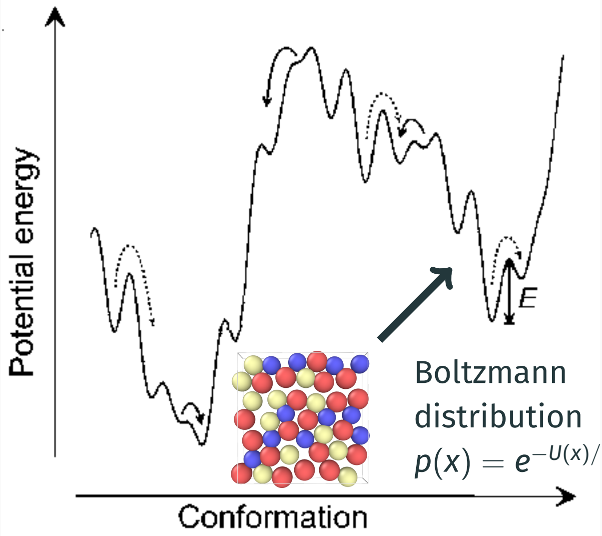
The Failure of Mode-Coupling Theory



The Failure of Mode-Coupling Theory



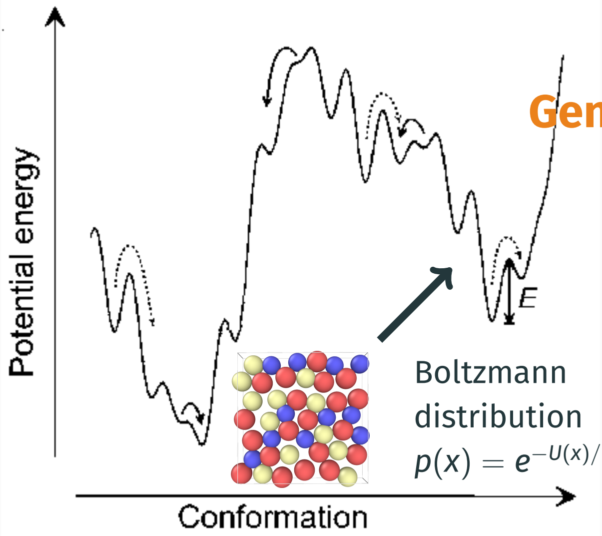
Our Sampling Problem



Boltzmann distribution

$p(x) = e^{-U(x)/T}$

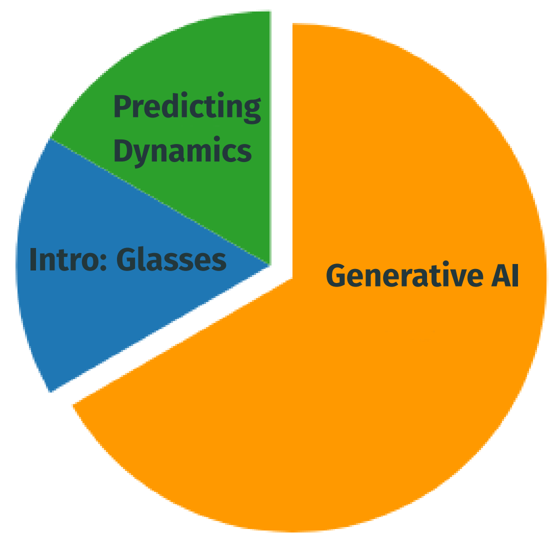
Our Sampling Problem



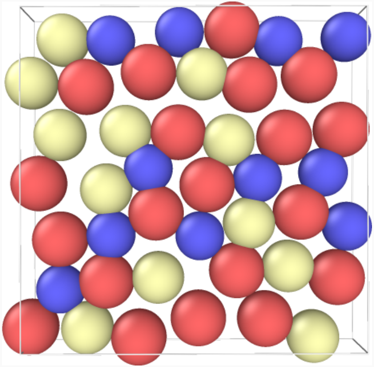
Generative AI!?

Boltzmann distribution

$p(x) = e^{-U(x)/T}$


Predicting Dynamics
Intro: Glasses
Generative AI

Sampling Supercooled Liquid Models



Ternary mixture

- 3 different types/radii (color) to prevent crystallization

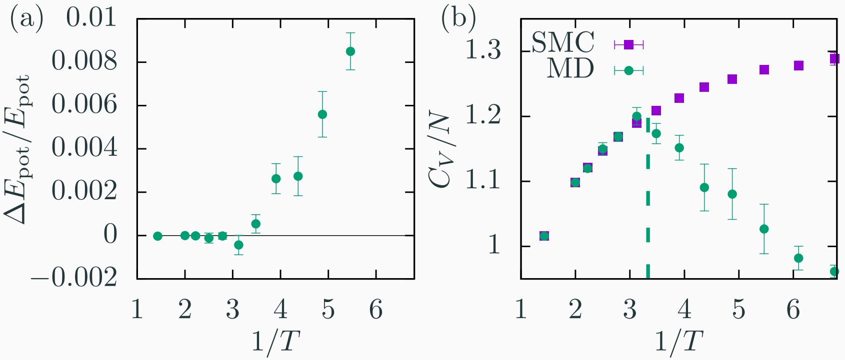
Molecular Dynamics (MD)

- Discretize and numerically solve Newtons e.o.m.

SWAP Monte Carlo (SMC)

- Accelerate equilibration by exchanging particles of different type

$\Delta E_{pot} = E_{MD} - E_{SMC}$ $\quad$ $C_V = T^{-2}(\langle E^2 \rangle - \langle E \rangle^2)$



Conclusion MD:

- Falls out of equilibrium at $T_{MD} = 0.3 \approx T_c$
- Can we do better with generative AI?

Generative AI

Learning to sample from high-dimensional probability distribution

You
Why would anyone be interested in studying glasses?

ChatGPT
Studying glasses, in the context of materials science or physics, can be of significant interest for several reasons:

1. **Understanding Material Properties:** Glasses exhibit unique material properties that differ from crystalline solids. Investigating these properties can help scientists understand fundamental aspects of solid-state physics and materials science.
2. **Technological Applications:** Glasses have a wide range of technological applications, including in optics (e.g., lenses, optical fibers), electronics (e.g., display screens, semiconductors), and materials engineering (e.g., glass ceramics, protective coatings). Studying glasses allows researchers to optimize these applications and develop new technologies.

ChatGPT, Dall-E, ... : Probability distribution **unknown** and **intractable**

↯

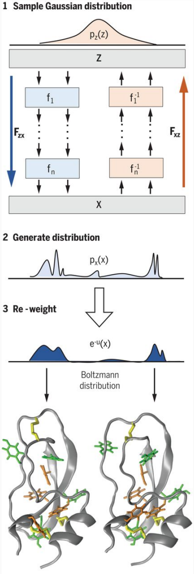
Statistical physics: Requires **control** over probability distribution $\left(P(x) \propto \exp(-U(x)/T)\right)$

ChatGPT, Dall-E, ... : Probability distribution **unknown** and **intractable**

↯

Statistical physics: Requires **control** over probability distribution ($P(x) \propto \exp(-U(x)/T)$)

Normalizing flows (NF): Invertible and analytically tractable



F. Noe *et al.* 365, eaaw1147 (2019)

Goal: Sampling from target distribution

$$\rho_*(x) = Z_*^{-1} \exp(-U(x)/T^*)$$

Prior (easy) distribution $\rho_P(x)$: High temperature samples

Goal: Sampling from target distribution

$$\rho_*(x) = Z_*^{-1} \exp(-U(x)/T^*)$$

Prior (easy) distribution $\rho_P(x)$: High temperature samples

Continuous Normalizing Flow: T

$$\frac{\mathrm{d}}{\mathrm{d}t} x_v(t) = v(x_v(t), t), \quad x_v(0) = x_0 \text{ drawn from } \rho_P(x)$$
$$T x_0 = x_v(t = 1).$$

$\rightarrow$ Natural choice for particle configurations

Continuous Normalizing Flow

$$\frac{\mathrm{d}}{\mathrm{d}t} x_v(t) = v(x_v(t), t), \quad x_v(0) = x_0 \text{ drawn from } \rho_P(x).$$

Approximation: Describe “flow” by pair-wise potential field

$$v(x(t), t) = \nabla_x \Phi(x(t), t),$$
$$\Phi(x(t), t) = \sum_{ij} \tilde{\Phi}(d_{ij}(t), t) \quad d_{ij}(t) = |x_i(t) - x_j(t)|.$$

J. Koehler *et al.*, arXiv:2006.02425 (2020)

Continuous Normalizing Flow

$$\frac{\mathrm{d}}{\mathrm{d}t} x_v(t) = v(x_v(t), t), \quad x_v(0) = x_0 \text{ drawn from } \rho_P(x).$$

Approximation: Describe “flow” by pair-wise potential field

$$v(x(t), t) = \nabla_x \Phi(x(t), t),$$
$$\Phi(x(t), t) = \sum_{ij} \tilde{\Phi}(d_{ij}(t), t) \quad d_{ij}(t) = |x_i(t) - x_j(t)|.$$

J. Koehler *et al.*, arXiv:2006.02425 (2020)

- **Equivariant**: Same symmetries as underlying distribution
- $\tilde{\Phi}(d, t)$: radial basis functions (free parameters)

Continuous Normalizing Flow

$$\frac{\mathrm{d}}{\mathrm{d}t}x_v(t) = v(x_v(t), t), \quad x_v(0) = x_0 \text{ drawn from } \rho_P(x).$$

Approximation: Describe “flow” by pair-wise potential field

$$v(x(t), t) = \nabla_x \Phi(x(t), t),$$
$$\Phi(x(t), t) = \sum_{ij} \tilde{\Phi}(d_{ij}(t), t) \quad d_{ij}(t) = |x_i(t) - x_j(t)|.$$

J. Koehler *et al.*, arXiv:2006.02425 (2020)

- **Equivariant**: Same symmetries as underlying distribution
- $\tilde{\Phi}(d, t)$: radial basis functions (free parameters)
- Differentiable $\rightarrow$ Backpropagation
- Loss: KL divergence(mixing of energy-based and maximum likelihood training)

Continuous Normalizing Flow

$$\frac{\mathrm{d}}{\mathrm{d}t} x_v(t) = v(x_v(t), t), \quad x_v(0) = x_0 \text{ drawn from } \rho_P(x).$$

Approximation: Describe "flow" by pair-wise potential field

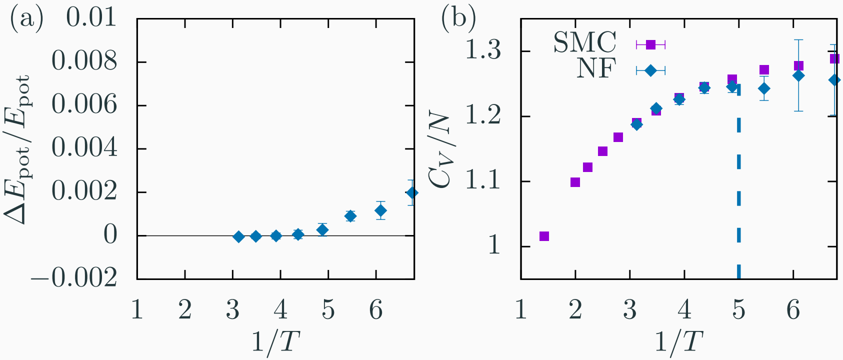
$$v(x(t), t) = \nabla_x \Phi(x(t), t),$$
$$\Phi(x(t), t) = \sum_{ij} \tilde{\Phi}(d_{ij}(t), t) \quad d_{ij}(t) = |x_i(t) - x_j(t)|.$$

J. Koehler *et al.*, arXiv:2006.02425 (2020)

- **Equivariant**: Same symmetries as underlying distribution
- $\tilde{\Phi}(d, t)$: radial basis functions (free parameters)
- Differentiable $\rightarrow$ Backpropagation
- Loss: KL divergence(mixing of energy-based and maximum likelihood training)
- Unbiasing with weights $w(x_0) \propto e^{-U(T x_0)/T_* - U(x_0)/T_P + \log\det|\nabla_x \bar{T}|}$

$$\Delta E_{pot} = E_{NF} - E_{SMC}$$

$$C_V = T^{-2}(\langle E^2 \rangle - \langle E \rangle^2)$$



Conclusion NF:

- Equilibration down to very low temperatures ($T_{NF} = 0.2$)
- **NF performs better than MD**
- Efficiency obtained by learning weights rather than transforming states

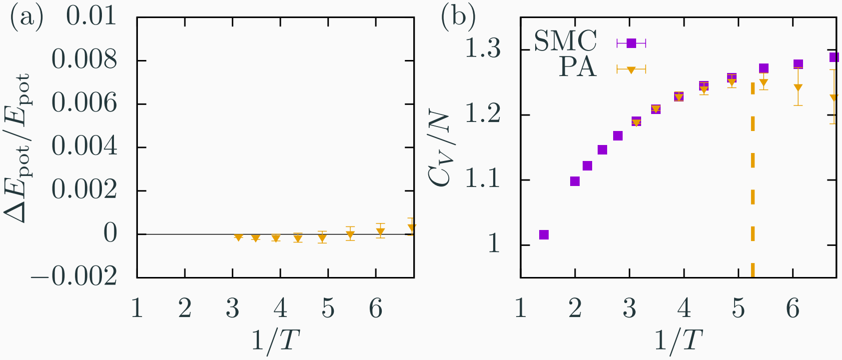
More benchmarking: Population annealing (**PA**)

More benchmarking: Population annealing (**PA**)

1. Create initial set of R configurations at high temperature T_0
2. Anneal configurations j to $T_{i+1} = T_i - \Delta T$ by weighting them as $e^{-\Delta T^{-1} E_j}$ (small ΔT!)
3. Relax by short MD simulations
4. Repeat iterativelly until desired low temperature is reached

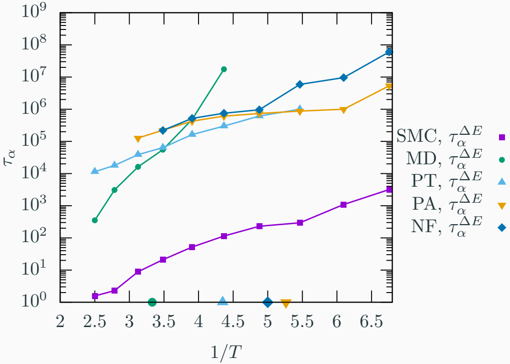
More benchmarking: Population annealing (**PA**)

1. Create initial set of R configurations at high temperature T_0
2. Anneal configurations j to $T_{i+1} = T_i - \Delta T$ by weighting them as $e^{-\Delta T^{-1} E_j}$ (small ΔT!)
3. Relax by short MD simulations
4. Repeat iterativelly until desired low temperature is reached



Can Generative AI Models Efficiently Sample Deeply Supercooled Liquids?

Can Generative AI Models Efficiently Sample Deeply Supercooled Liquids?



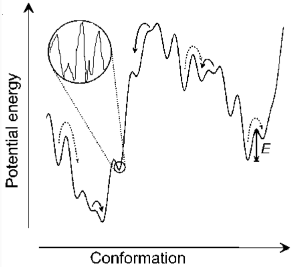
Yes, but not better than other enhanced sampling techniques!

Conclusion: Generative AI

- Enhanced sampling surprisingly efficient for small systems
- **Take-home message**: Benchmarking is essential!

Outlook

- Improve generative AI?

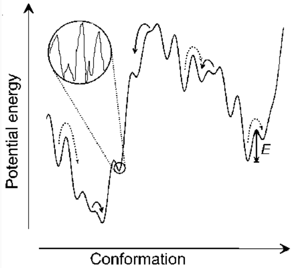


Conclusion: Generative AI

- Enhanced sampling surprisingly efficient for small systems
- **Take-home message**: Benchmarking is essential!

Outlook

- Improve generative AI?

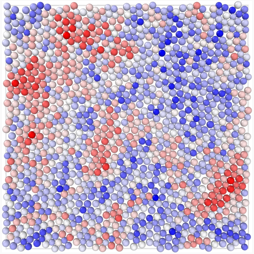


- Combine generative AI with other techniques
- Reinforcement learning improved SMC? (local moves)

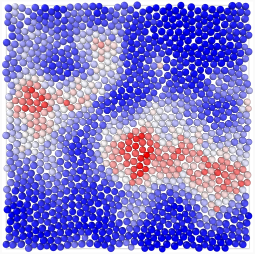
Dynamic Heterogeneity (DH)

- Strong contrast between active and passive regions
- Increases with decreasing T

High T

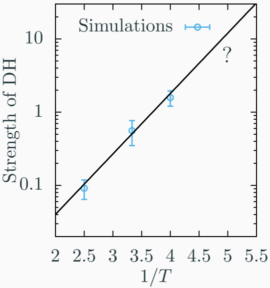


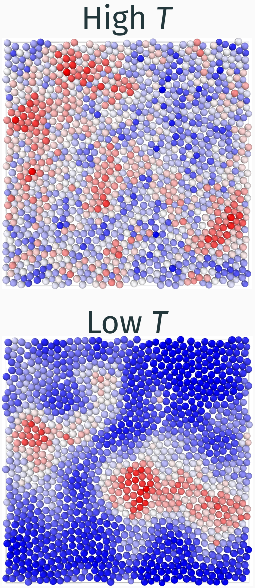
Low T



Dynamic Heterogeneity (DH)

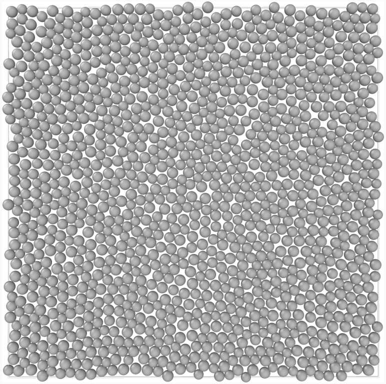
- Strong contrast between active and passive regions
- Increases with decreasing T
- Fate at low temperatures?





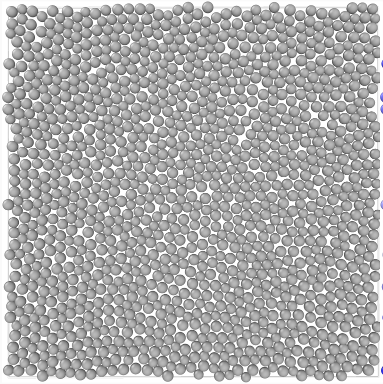
Our Inference Problem

Structure (Input)

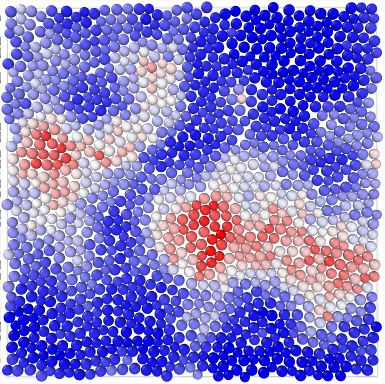


Our Inference Problem

Structure (Input)



Dynamics (Labels)

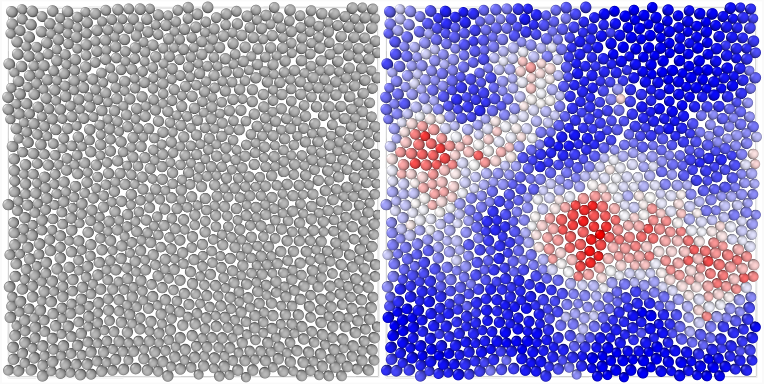


Question 2: Where will (long-time) structural relaxation set in?

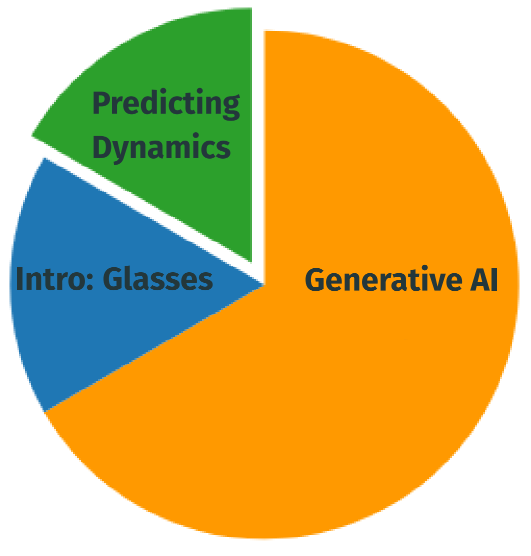
Our Inference Problem

Supervised Machine Learning!?

Structure (Input) Dynamics (Labels)



Question 2: Where will (long-time) structural relaxation set in?


Predicting Dynamics
Intro: Glasses
Generative AI

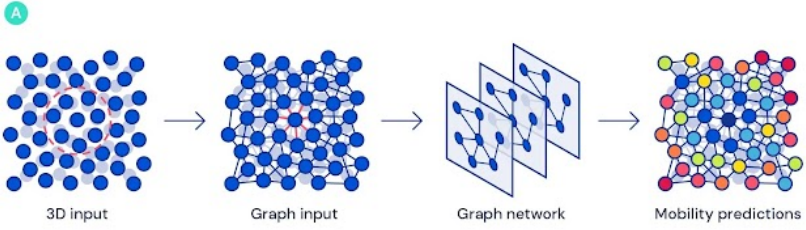
Human-made descriptors (density, potential energy, local packing)

- Weak correlation

Human-made descriptors (density, potential energy, local packing)

- Weak correlation

State-of-the-art: Graph neural networks (GNNs)



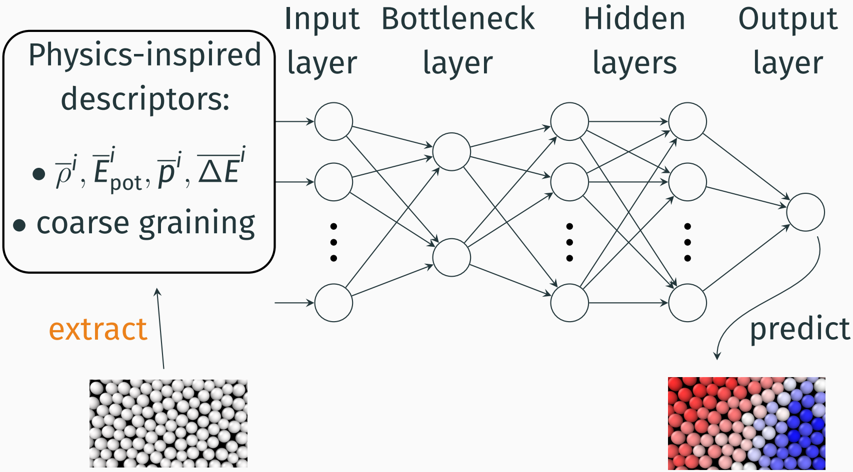
Bapst *et al.*, Nat. Phys. 16, 448 (2020)

- Computationally heavy

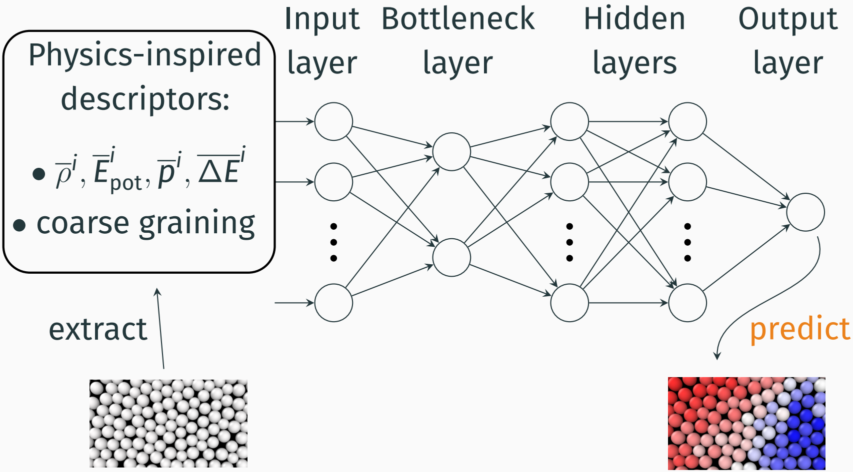
Physics-inspired machine learning: Combine different structural descriptors (inductive bias):

- Potential energy
- Voronoi volume/perimeter
- Local density
- Bond-order, locally favored structures, Tanaka's θ
- Soft modes
- ...

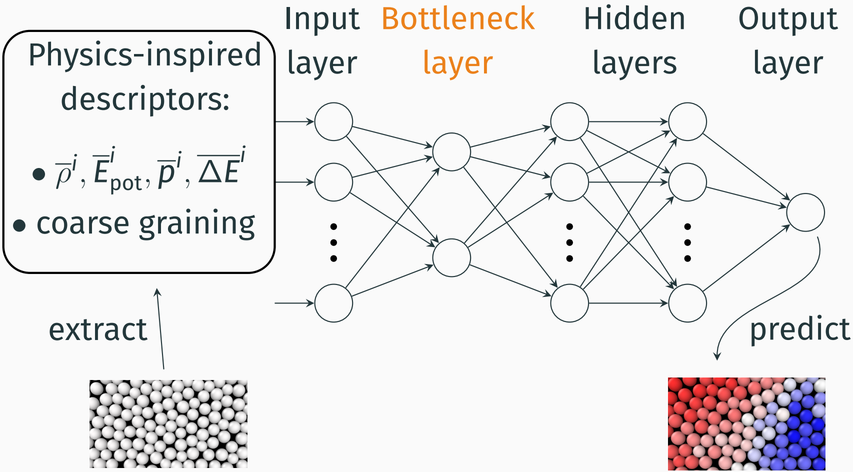
Network Geometry of GlassMLP



Network Geometry of GlassMLP

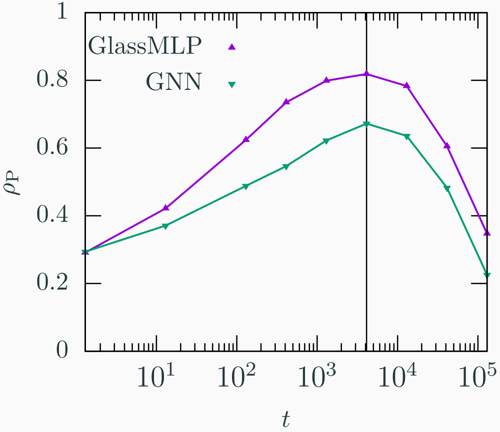


Network Geometry of GlassMLP

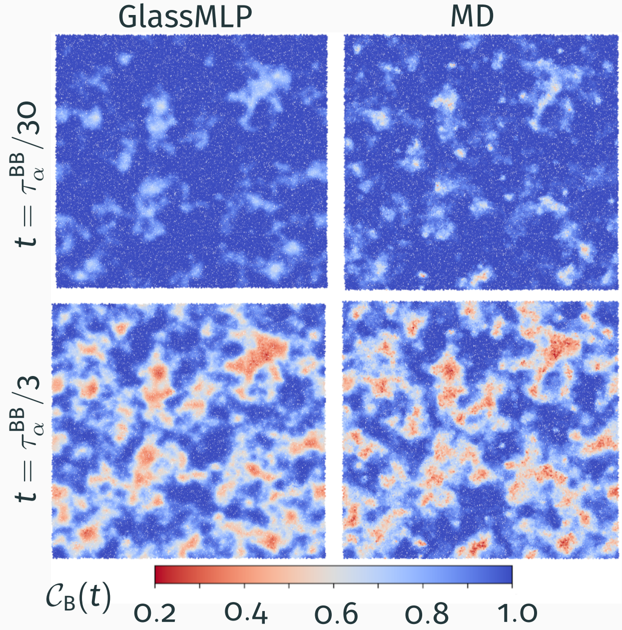


Benchmarking: Supervised learning on training set, analyze predictability on test set

Pearson: $\rho_P = \frac{\text{cov}(\mathcal{C}_B^i, Y_{\text{MLP}}^i)}{\sqrt{\text{var}(\mathcal{C}_B^i)\text{var}(Y_{\text{MLP}}^i)}}$



- Great performance
- Parsimonious (training: < 5 minutes)

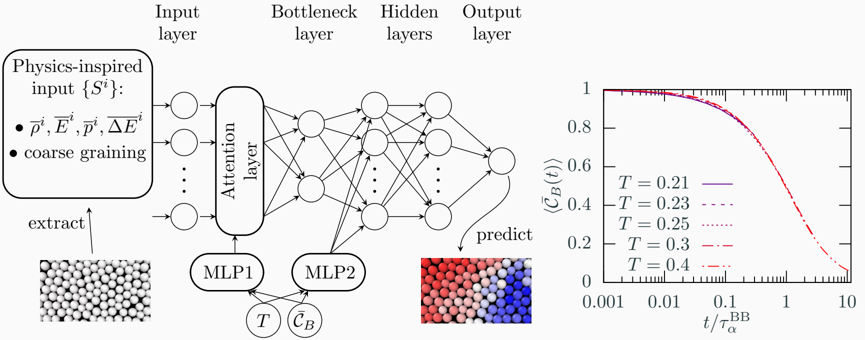

GlassMLP
MD
$t = \tau_\alpha^{BB}/30$
$t = \tau_\alpha^{BB}/3$
$\mathcal{C}_B(t)$
0.2
0.4
0.6
0.8
1.0

Until now: Learn one network for each state point (time, temperature)

New approach: Learn one network for all times and temperatures!

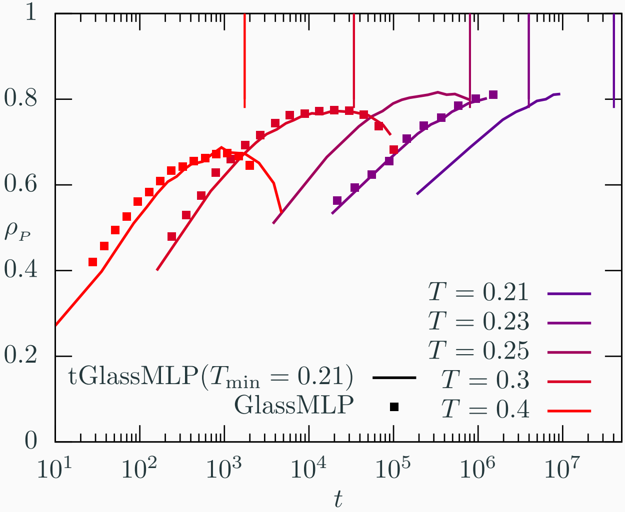
Until now: Learn one network for each state point (time, temperature)

New approach: Learn one network for all times and temperatures!



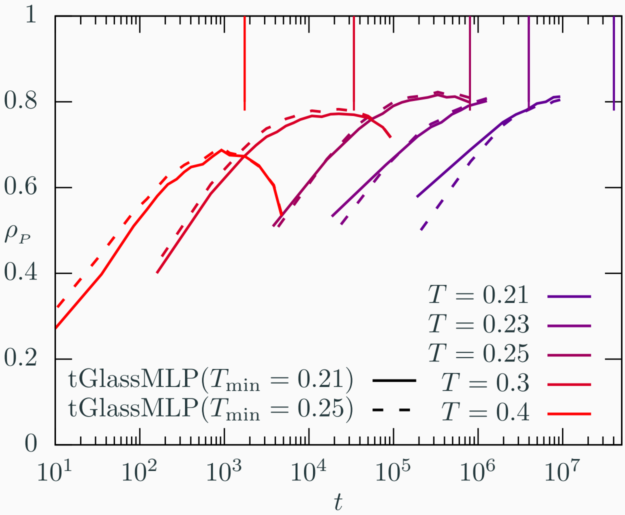
→ Only twice as many fitting parameters as before (≈ 1200)

Pearson correlation



tGlassMLP learns universal structural descriptors

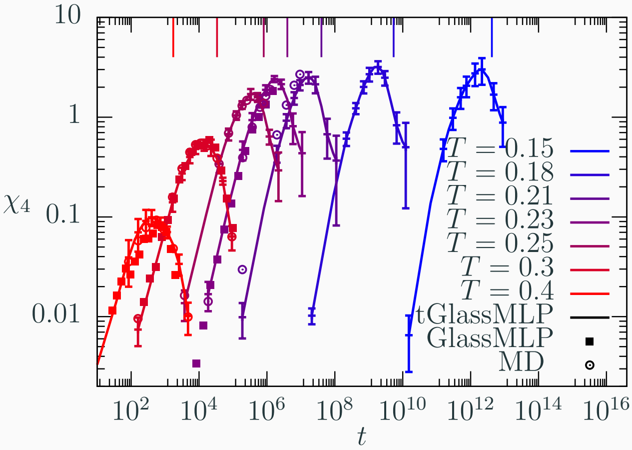
Pearson correlation



tGlassMLP transfers very accurately to lower temperatures

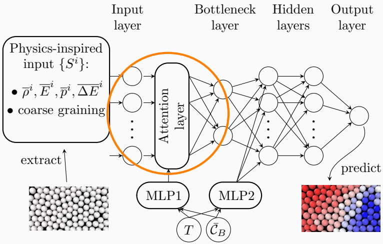
Dynamic susceptibility

$$\chi_4(t) = N\left(\langle \bar{\mathcal{C}}_B^2(t)\rangle - \langle \bar{\mathcal{C}}_B(t)\rangle^2\right) \text{ with } \bar{\mathcal{C}}_B(t) = \sum_i \mathcal{C}_B^i(t)$$



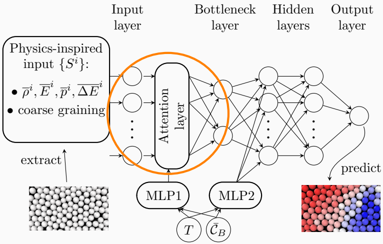
Transferability indicates crossover in susceptibility $\chi_4(T)$

Interpretation of trained networks

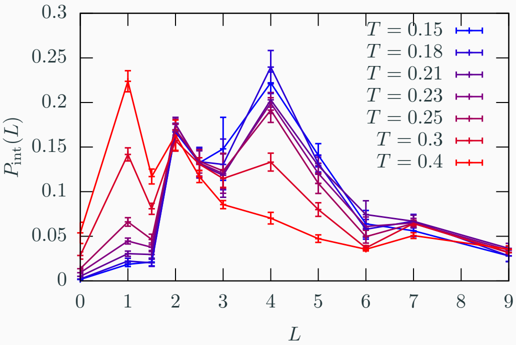


- Attention layer learns weights for each structural descriptor
- Extract relative weight of descriptor with length scale L

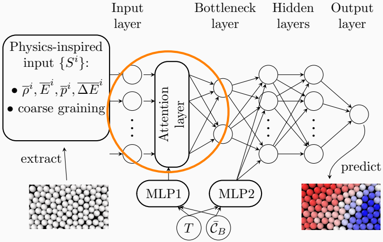
Interpretation of trained networks



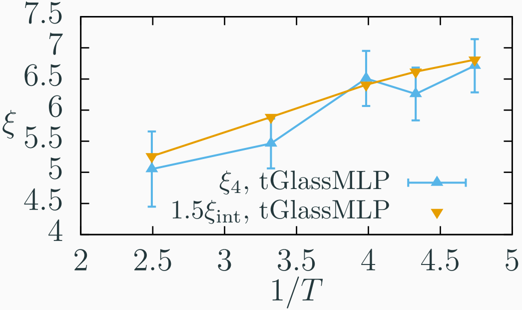
- Attention layer learns weights for each structural descriptor
- Extract relative weight of descriptor with length scale L



Interpretation of trained networks



- Attention layer learns weights for each structural descriptor
- Extract relative weight of descriptor with length scale *L*



Final Conclusions

Machine learning: useful tool to study glass physics

- *Generative AI* to create amorphous structures
- *Scalability* and *transferability*: Characterize dynamic heterogeneity at experimental glass transition temperature
- *Interpretability*: Extract growing length scale

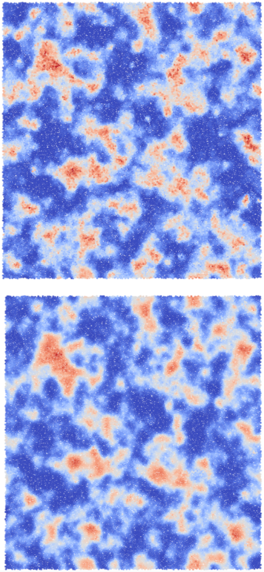
Final Conclusions

Machine learning: useful tool to study glass physics

- *Generative AI* to create amorphous structures
- *Scalability* and *transferability*: Characterize dynamic heterogeneity at experimental glass transition temperature
- *Interpretability*: Extract growing length scale

Outlook

- Create “Glass Simulator” (Generative AI + Dynamics)



References:

G. Jung, G. Biroli, L. Berthier;

PRL **130**, 238202 (2023)

PRB **109**, 064205 (2024)

MLST **5**, 035053 (2024)

Roadmap: arXiv:2311.14752

CNRS, Montpellier

- Ludovic Berthier

ENS, Paris

- Giulio Biroli

Thank you for your attention!