

Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

Tomilola M. Obadiya

Date

Correlating Structure with Dynamics in Supercooled Liquids Using Machine
Learning Tools

By

Tomilola M. Obadiya
Doctor of Philosophy

Physics

Daniel Sussman, Ph.D.
Advisor

Gordon Berman, Ph.D.
Committee Member

Justin Burton, Ph.D.
Committee Member

Eric Weeks, Ph.D.
Committee Member

Carl Yang, Ph.D.
Committee Member

Accepted:

Kimberly J. Arriola, Ph.D.
Dean of the James T. Laney School of Graduate Studies

Date

Correlating Structure with Dynamics in Supercooled Liquids Using Machine
Learning Tools

By

Tomilola M. Obadiya
B.Sc., Obafemi Awolowo University, Osun Nigeria, 2015
M.Sc., Creighton University, NE, 2019

Advisor: Daniel Sussman, Ph.D.

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Physics
2025

Abstract

Correlating Structure with Dynamics in Supercooled Liquids Using Machine Learning Tools

By Tomilola M. Obadiya

Understanding the relationship between structure and dynamics in supercooled liquids remains a central challenge in glass physics. Machine learning techniques, particularly Support Vector Classification (SVC), have provided insights into structural predictors of dynamics, such as the “softness” order parameter, which correlates with energy barriers for particle rearrangements. This dissertation critically examines the methodology and interpretability of machine learning models in this context, focusing on their ability to predict rearrangement probabilities and energy barriers. We first investigate whether classification hyperplanes trained on structural data from high-temperature, diffusive regimes can predict energy barriers in the supercooled regime. By introducing a Z-score-based binning approach, we demonstrate that structural features associated with purely diffusive motion retain predictive power for activated events, challenging conventional assumptions about structure-dynamics correlations at high temperatures. Building on this, we explore the physical interpretability of various regression-based machine learning models, including Ridge Regression, Support Vector Regression, and Multilayer Perceptron, in predicting energy barriers. Our analysis, leveraging the iso-configurational ensembles, shows that these models capture similar structural signatures as SVC, reinforcing the idea that predictive success is rooted in the ability to learn high-dimensional structure-dynamics relationships. We then extend our investigation to the role of memory effects in supercooled liquids, using softness as a structural order parameter to probe system responses under thermal cycling. Preliminary findings suggest that supercooled liquids exhibit memory effects typically associated with glasses, raising new questions about the glass transition. This dissertation highlights key challenges in using machine learning to understand glassy dynamics, particularly regarding feature selection, model interpretability, and the physical significance of learned representations. Our findings contribute to ongoing efforts to develop more robust and interpretable machine learning frameworks for studying supercooled liquids and other complex systems.

Correlating Structure with Dynamics in Supercooled Liquids Using Machine
Learning Tools

By

Tomilola M. Obadiya
B.Sc., Obafemi Awolowo University, Osun Nigeria, 2015
M.Sc., Creighton University, NE, 2019

Advisor: Daniel Sussman, Ph.D.

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Physics
2025

Acknowledgments

I am deeply grateful to my advisor, Dr. Daniel M. Sussman, for his insightful mentorship and unwavering support throughout this doctoral journey. His expertise in Statistical Mechanics and Thermodynamics, in general, is very respectable and his ability to foster critical thinking have been instrumental in shaping my professional development and this dissertation. Daniel has taught me to identify mathematical basis to concepts, as I am more intuitive, which has improved how I do science. I am equally grateful for the dynamics of our interactions which is laden with mutual respect. Thank you, Daniel, for the opportunity to join the lab. As always, “Thank you for your time”.

I also express my sincere appreciation to the members of my dissertation committee, Drs. Eric Weeks, Justin Burton, Carl Yang, and Gordon Berman, for their constructive criticism, valuable perspectives, and insightful advice, which significantly enhanced the quality of this work. I am thankful to have benefited from your wealth of experience and wisdom which you freely gave.

I am grateful to my collaborators, Sean A. Ridout and Claudia Schmidt, for their contributions to our work that lead to Chapter 3 of this dissertation. I am particularly grateful for the insightful conversations I had with Sean A. Ridout and Arabind Swain; these conversations influence the work in this dissertation in various ways.

I thank Helen Ansell, Charles Packard, Toler Webb, and Chengling Li, current members of the SussmanLab research group, and Haicen Yue, a past member, for their camaraderie and intellectual support. The positive atmosphere within the group made the research process both enjoyable and uplifting. Because of you all, I am getting better at presenting science.

I am thankful for my cohort here at the Department of Physics at Emory University. We made great memories together that will last a lifetime. I know the future is bright for us all. I am equally grateful for everyone in the department with whom I have

interacted, and perhaps shared a smile with. Additionally, I am thankful for my colleagues at the Soft Matter Journal Club and the Theoretical BioPhysics Journal Club; I have learned a lot from our meetings.

I am thankful to the National Science Foundation (through Daniel!), supported by the American taxpayer for their support. It is a privilege to be paid to explore research questions that ignite my curiosity.

I am eternally grateful to my darling wife, my loving parents and family and my dependable friends for their unwavering support and encouragement throughout this journey. Their belief in me provided the strength and motivation to persevere.

Contents

1	Introduction	1
1.1	Background and Motivation	1
1.2	Kob-Andersen glass model	4
1.3	Dynamical Heterogeneity	5
1.4	Relaxation Behavior in Supercooled Liquids	7
1.5	The Potential Energy Landscape	13
1.6	Challenge in defining an order parameter correlating structure with dynamics	15
1.6.1	Particle Mobility from Iso-configurational ensemble	16
1.6.2	Correlating structure and dynamics through machine learning	17
1.7	Dissertation: Challenges in correlating structure with dynamics in su- percooled liquids using machine learning tools	23
1.8	Dissertation Outline and Contributions	24
2	Using fluid structures to encode predictions of glassy dynamics	27
2.1	Introduction	27
2.2	Methods	30
2.2.1	Model and simulations	30
2.2.2	Local structure and dynamics	32
2.2.3	Machine learning protocol	33

2.3	Classification in the fluid phase	35
2.4	Interpretability of fluidity and softness	38
2.5	Discussion	40
3	Machine Learning of Energy Barriers to Rearrangements from Local Structure in Supercooled Liquids	46
3.1	Introduction	47
3.2	Methods	50
3.2.1	Model and Simulations	50
3.2.2	Identifying Rearrangements	52
3.2.3	Structural descriptors	52
3.2.4	Standard training datasets	54
3.3	Inferring energy barriers with different data-driven approaches	54
3.3.1	Correlating structure with dynamics using linear models	54
3.3.2	Inferring energy barriers	56
3.3.3	Correlating structure with dynamics using non-linear models . . .	61
3.4	Impact of structural descriptors on model performance	62
3.5	Impact of dynamical labels on model performance	65
3.6	Conclusion	66
4	Ongoing Work on Memory Effect in Supercooled Liquid	70
4.1	Introduction	70
4.2	Methods	74
4.2.1	Model and Simulations	74
4.2.2	Temperature Cycle	75
4.3	Results	76
4.4	Discussion	76
4.5	Future Directions	78

5	Conclusion	80
5.1	Open Questions and Future Directions	81
5.2	Final Remarks	83
Appendix A	Supplemental Information to Chapter 3	84
A.0.1	Mean trend for softness and log of the probability of rearrange- ment	84
A.0.2	Definition of Cumulative Squared Displacement (CSD)	85
A.0.3	Influence of the ratio of training examples to model parameters	86
	Bibliography	87

List of Figures

- 1.1 **Dynamics varying heterogeneously in space in a supercooled liquid.** Particles are colored based on how much they have displaced from their initial position. The deep red particles have moved several particle diameters while the deep blue particles have not moved at all. Particles with intermediate displacements are colored correspondingly. The dynamics of particles varies heterogeneously across the volume of the supercooled liquid. This figure is from [1]. 6
- 1.2 **Radial distribution function for a KA model as a function of radial distance.**The radial distribution function shows nearly identical peak positions and shape between liquid and supercooled states, with only a slight increase in nearest, next-nearest etc. neighbors upon cooling, indicating minimal changes in average spatial arrangement. Legend represents different temperatures. The cartoon illustrates how the local density at a distance r from a reference particle (in red) is calculated using a thin spherical shell of width dr 8

1.3	Self-Intermediate scattering function for Kob-Andersen glass model at density $\rho = 1.2$ for different temperatures. A two-step relaxation process become more obvious as temperature (legend) decreases. On supercooling, structural relaxation develops a stretched-exponential form, deviating from an exponential process and reflecting the complexity that develops. The dashed horizontal line show when $F_s(q, t) = 1/e$	9
1.4	Alpha-relaxation time against inverse temperature for the Kob-Andersen model. The relaxation time for the KA model increases in a super-Arrhenius form as temperature decreases. For the temperature range for which structure changes linearly, there is a super-Arrhenius change in dynamics. The data points for this figure are from [2] for a KA model at density $\rho = 1.2$	11
1.5	The viscosity of several glass-formers as a function of temperature. The viscosity, so also the relaxation time, for strong glass-formers grow in an Arrhenius fashion. Fragile glass-formers exhibit super-Arrhenius growth as temperature decreases. This figure is from [3].	12
1.6	The Arrhenius dependence of $P(R S)$ on temperature. The probability of rearrangements for different soft particles becomes orders of magnitude apart with temperature decrease. At the onset temperature, T_o (around 0.8 to 1.2), dynamics is uncorrelated with structure as shown by the merge of the Arrhenius fit (dotted line). The colored lines represent values of $S = -3$ (blue) to $S = +3$ (red). The inset show a collapse of $P(R S)/P_o$ when plotted against $\Delta E/T$. Here, $P_o = e^\Sigma$. This Figure is from [4].	20

2.1	Fluidity classifies rare events at high and low temperatures.	
	The points show the 5-fold cross-validation accuracy of linear SVMs trained on extreme diffusive samples at $T = 1.0, 1.2, 1.4, 1.8, 2.0$ (dark blue to light red) as a function of the classifier's bias. Each classifier achieves near-peak accuracy for small values of the bias. In contrast, the inset shows the test classification accuracy of the "softness" classifier trained on activated dynamics at $T=0.45$ and applied to the extremes of diffusive events at different temperatures, for which very different values of the bias optimize performance.	36
2.2	A transfer learning approach connects extreme diffusive events above T_0 with activated dynamics below T_0. The main figure shows the test accuracy of linear SVMs, trained on extreme diffusive samples at $T = 1.0, 1.2, 1.4, 1.8, 2.0$ (dark blue to light red), as applied to a test set of activated dynamics at $T = 0.47$ as a function of the classifier's bias relative to the optimal bias for that choice of temperature. The inset show the self part of the Van-Hove correlation function for large particles at a time scale of 10 LJ time units. (Black dots are for $T=0.45$.)	37
2.3	The distribution of fluidity at different test temperatures for two considered training temperatures. Training at both $T = 1.2$ (a) and $T = 2.0$ (b) the distribution of fluidity is approximately Gaussian. The mean is a monotonic function of the test temperature (legend).	43

2.4	The probability of rearrangement conditioned on fluidity reveals energy barriers below the onset temperature. (a) The log probability of rearrangement conditioned on $\mathcal{F}^{2.0}$ vs inverse temperature. Point colors correspond to different bins of fluidity, as indicated in the legend. Part (b) shows the same features for rearrangements conditioned on S . In all cases, dotted lines are Kramers-form fits. . . .	44
2.5	Collapse of inferred landscape features from different training temperatures. The energy barrier as a function of distance to optimized classifier, ΔE vs. x_α , as inferred from Kramers fits to $P_R(S)$ shows little variation across a wide range of classifier training temperatures. The inset showing the entropic contribution similarly collapses in this representation.	45
3.1	Predictions from the SVR model strongly correlates with Softness. The distribution of the log of the rearrangement probability predicted by the SVR model follows an approximately Gaussian shape. This distribution shifts progressively to the right as the temperature (legend) increases. The inset scatter plot further demonstrates a correlation between the SVR model predictions and Softness, suggesting that both quantities capture related structural information.	57
3.2	SVR predictions have an Arrhenius form. Binning particles by $\ln p_r^{(model)}$ and finding the fraction of the binned particles that rearranges at different temperature has an Arrhenius decomposition. The energy barrier, ΔE , and the entropic piece, Σ , are extracted by fitting to these Arrhenius form (dashed lines show a Kramer's fit). The gradient from light green to dark purple indicates $\ln p_r^{SVR}$ values ranging from -6.5 to -10.	58

3.3	Different models predict similar energy barriers to local rearrangements. Energy scale and (inset) entropic terms characterizing the local energy barrier to particle rearrangement. A strong correlation between regression models and Softness is observed. Different symbols correspond to different regression and classification models, as noted in the plot legend — the chosen symbol shapes consistently correspond to the indicated model in all future figures.	59
3.4	Fraction of the true variance explained by the ML models when applied to iso-configurational snapshots across a range of temperatures. The explained variance remains lesser than 0.5 across temperatures, indicating that both classification-based and regression-based models poorly capture relevant structural features of the energy landscape. The error bars are standard error on the mean from 40 bins each containing 177 uncorrelated particles.	61
3.5	SVR model interpretation as a function of the number of features considered. Through RFE, the (a) energy barrier height and (b) entropic contribution to the probability of rearranging as a function of the dimension of feature space barely changes. The fraction of the true variance learned by the model (c) increases modestly over this range.	63

3.6	Model predictions as a function of varying the feature space.	
	(a) The inferred energy barriers and entropic contributions of an SVR model using higher-order features are strongly correlated with those using the standard (Behler-Parrinello) features. (b) Using higher order features does not lead to better estimate of the variance of logit p_r^{true} . The chosen symbol shapes in (b) consistently correspond to the indicated features in (a). Error bars are standard error on the mean from 40 different bins each containing 177 uncorrelated particles.	68
3.7	Inferred local landscape features for different choices of dynamical label. The inferred energy barriers (main plot) and entropic contributions (inset) obtained from an SVR classifier are nearly indistinguishable when training is done with different dynamical labels. . .	69
3.8	Pearson correlation between true and predicted inherent state propensity. The Pearson correlation between true and predicted inherent state propensities for an SVR model trained on our thermal configurations (light cyan triangles) and an analogous model trained directly on inherent state data from Ref. [5] (black triangles).	69
4.1	The evolution of the average probability of rearrangement for different snapshots approach a common value at long timescale. The probability of rearrangement averaged for the big particles for three different thermal snapshots (Legend) evolve in such a way that the path they take depends on the snapshot.	71
4.2	The evolution of the average softness of particles grouped by their softness at time $t = 0$. The average softness of particles grouped based on their initial softness (Legend) evolve towards the system average. The paths each grouped particles follow towards the long-time depends on their initial value.	72

4.3	The normalized average softness for different particle groups (binned by initial softness) is compared with the self-intermediate scattering function ($F_s(k, t)$). The decay of the normalized average for various initial softness (thick lines) at short timescales, on the order of the β -relaxation time (τ_β), suggests a connection between local structural rearrangements and the fast-relaxation process. The dash line represents the $F_s(k, t)$ and the numbers on the legend represents initial softness.	73
4.4	The evolution of average softness for particles grouped by their initial softness. The red vertical line marks the time at which the temperature began to increase. The blue vertical line marks the time at which temperature is returned back to the initial value. There seem to be a consistent bump in the average structure for each bins during the temperature cycle. At the moment, it is too soon to tell. The bumps seem to show that both soft and hard particles become softer during the temperature cycle. The path of the particles return back to their initial state when the temperature is reverted back to the initial value.	77
A.1	Trend of the mean of the distribution from the models as a function of Temperature. The mean of the outputs from the models increases with temperature. However, the mean from the linearSVR and linearSVC models increases at almost the same rate as a function of temperature.	85
A.2	Details of the energy landscape inferred by the multilayer perceptron models. The energy barrier ΔE , and the entropic piece, Σ , are correlated for the different MLP.	86

List of Tables

- 2.1 **Table of the optimal bias, b , of the hyperplane during training; the standard deviation of fluidity (or softness for $T=0.45$) at the training temperatures; and the optimal bias, b^{best} , that maximizes test accuracy at $T=0.47$. The final column displays the projection of our classifiers onto the softness classifier.** . 40
- 3.1 **Features of linear and non-linear models** For linear and non-linear models we show (where appropriate) the projection of hyperplanes onto the softness hyperplane, the average of the predicted $\ln p_r$ at the training temperature, the bias, and the standard deviation of the distribution of the predicted $\ln p_r$ at the training temperature. . 60

Chapter 1

Introduction

1.1 Background and Motivation

The most familiar states of matter—gas, liquid, and solid—are not merely different forms but represent distinct physical phases characterized by their underlying symmetries and mechanical properties. Consider a simple system: an ensemble of particles confined within a container. In the gaseous state, these particles move freely, bouncing off each other and the container walls in a seemingly chaotic dance. This freedom of movement reflects the continuous translational and rotational symmetry of the gas: we can shift the entire system by any arbitrary distance or rotate it by any angle, and it would appear unchanged. Cooling this gas reduces the kinetic energy of the particles, decreasing the frequency and intensity of their collisions. Eventually, the particles become closely packed, forming a liquid.

While denser than a gas, a liquid retains the continuous translational and rotational symmetry. Like the gas, it can flow and adapt to the shape of its container. However, if we continue to cool the liquid, it will eventually reach its crystallization point. At this temperature, a profound transformation occurs: the continuous symmetries are broken, and the system adopts a crystalline solid structure. This solid is

characterized by discrete symmetries, meaning that only specific rotations and translations will leave the system looking the same. Crucially, this breaking of symmetry is accompanied by a dramatic change in the mechanical properties: the solid becomes rigid, able to resist deformation and support mechanical loads [6].

However, if we cool the liquid rapidly enough, we can bypass the crystallization point and enter a metastable state known as a supercooled liquid. This state is fascinating because it exhibits the structural characteristics of a liquid—disordered and lacking long-range order—yet displays solid-like rigidity at short timescales – typically around the picosecond to nanosecond range [3]. Imagine a dense crowd of people jostling about: at any given moment, an individual might feel “trapped” by their neighbors, unable to move freely. Yet, given enough time, the crowd as a whole can rearrange and flow. This is the essence of a supercooled liquid: it behaves like a solid on short timescales but flows like a liquid at longer timescales, relaxing any imposed stresses, whether thermal or mechanical. As we continue to cool, the time it takes for this relaxation to occur increases dramatically over a narrow range of temperature. Eventually, we reach a temperature, known as the glass transition temperature (T_g), where the relaxation time becomes so long that it exceeds any practical observation time. At this point, the system falls out of equilibrium with its surroundings, becoming trapped in a disordered, non-crystalline state: a glass. The system retains the continuous rotational and translational symmetry of the liquid, yet it is puzzlingly rigid.

Like all solids this rigidified systems will fail upon the application of sufficiently high mechanical load. In crystalline solids, failure typically originates from sites exhibiting local deviations from the crystalline order. These initiation sites are readily identifiable because we can define parameters that quantify local order. However, in glasses, the absence of a well-defined local order parameter makes the identification or prediction of failure points a significant challenge.

To address this challenge, researchers have sought to identify order parameters that can capture the underlying structural features governing the dynamics of supercooled liquids and glasses. Throughout this dissertation, the term “dynamics” specifically refers to particle rearrangements. Particle rearrangements are the fundamental microscopic events that initiate material failure. Unlike thermodynamic quantities like density which remain largely homogeneous, these rearrangements are highly heterogeneous and localized, typically involving only a few neighboring particles. Given that the dominant structural length scales in these disordered materials are indeed microscopic, extending at most a few particle diameters, these localized rearrangements become critical precursors to larger-scale material failure. These efforts have ranged from intuitive ideas like local free volume [7, 8, 9, 10] to data-driven methods like “softness”. As I will describe in more details in Section 1.6, “softness” was identified using machine learning techniques, specifically Support Vector Machines (SVM). Softness quantifies the local structural environment of a particle and has been shown to correlate with its propensity for rearrangement and stress relaxation. However, the precise reasons for the effectiveness of this order parameter are still not fully understood.

This dissertation delves into the intricate relationship between structure and dynamics in supercooled liquids, with a particular emphasis on understanding why softness is predictive of particle rearrangement from structure through a systematic test of the assumptions implemented in the softness machinery. We examine the underlying assumptions in the construction of the datasets employed to train the Support Vector Machine (SVM). Furthermore, we investigate the dependence of the machine-learned order parameter on the specific machine learning algorithm utilized. By exploring the assumptions employed in developing this order parameter, we aim to shed light on the fundamental mechanisms governing the behavior of these fascinating and technologically important materials.

In the remainder of this introductory chapter, I provide a comprehensive overview of the key concepts and challenges addressed in this dissertation. The rest of this chapter is structured as follows: Section 1.2 introduces the KA model, a canonical glass-forming system and the primary system used in our computational studies. Section 1.3 introduces the phenomenon of heterogeneous dynamics in supercooled liquids and glasses, highlighting its significance in understanding the complex relaxation behavior of glassy materials. Section 1.4 delves into the mechanisms of structural relaxation in supercooled liquids, examining the interplay between different relaxation processes and their temperature dependence, referencing the KA model. In Section 1.5, we explore the abstract potential energy landscape framework, illustrating how the exploration of this landscape governs the dynamics of supercooled liquids and connects the concepts introduced in the previous sections. Section 1.6 outlines the challenges in the field concerning the correlation of structure with dynamics in supercooled liquids, and introduces a promising approach based on machine learning. Section 1.7 focuses on specific challenges that we deem particularly important and aim to resolve within this dissertation. Finally, Section 1.8 provides a detailed outline of the dissertation’s structure, objectives, and key contributions.

1.2 Kob-Andersen glass model

Throughout this dissertation, our investigations utilize the Kob-Andersen (KA) model [11], a widely studied glass-forming system. This model comprises a binary mixture of two particle species, A and B, differing only in size. The mixture consists of 80% A particles and 20% B particles, interacting via the Lennard-Jones potential:

$$V_{ij}(r) = 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r} \right)^{12} - \left(\frac{\sigma_{ij}}{r} \right)^6 \right], \quad (1.1)$$

where $i, j \in \{A, B\}$ denote the particle species, r is the interparticle distance, ϵ_{ij} represents the interaction energy scale, and σ_{ij} represents the interaction length scale. The specific values of these parameters, namely $\epsilon_{ij} \in \{1.0, 1.5, 0.5\}$ and $\sigma_{ij} \in \{1.0, 0.8, 0.88\}$, for the AA , AB and BB pairs respectively, were chosen to imitate the nickel-phosphorus mixture [12]. These parameters were carefully selected to promote mixing and prevent phase separation between the two species, facilitating the formation of a structurally homogeneous supercooled liquid. Distance is measured in the unit of the large particle, σ_{AA} ; energy is measured in the unit of ϵ_{AA} . The mass for both particle specie is set to unity. With the Boltzmann's constant set to unity, temperature is measured in the unit of ϵ_{AA} (which has a value of unity). The computational simplicity of the KA model and its ability to capture key features of glassy dynamics has established it as a standard model for computational studies of the glass transition and the behavior of viscous liquids. For a reduced density of $\rho = 1.2$, the reduced melting temperature is observed to be $T_m \approx 1.028$ [12] – the exact value depends on pressure.

1.3 Dynamical Heterogeneity

Contrary to the uniform behavior of particles in crystalline solids, supercooled liquids and glasses exhibit heterogeneous dynamics, characterized by spatial and temporal fluctuations in particle mobility [13]. Fig. 1.1 shows an heterogeneous display of particle mobilities; mobile and immobile regions vary in space in a non-homogeneous fashion. This phenomenon, observed in a wide range of amorphous solids with glassy dynamics [14], emerges as the system is cooled below the onset temperature (T_o) of dynamical heterogeneity. The onset temperature, which is within the supercooled regime, is between the crystallization/freezing temperature and T_g . In three dimensions, the onset temperature of dynamical heterogeneity, T_o , for the KA model at

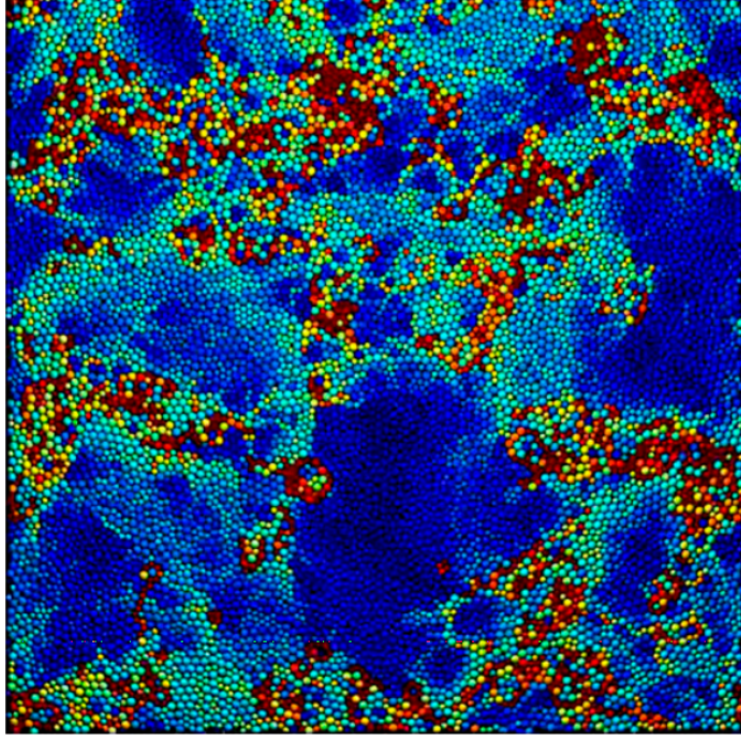


Figure 1.1: **Dynamics varying heterogeneously in space in a supercooled liquid.** Particles are colored based on how much they have displaced from their initial position. The deep red particles have moved several particle diameters while the deep blue particles have not moved at all. Particles with intermediate displacements are colored correspondingly. The dynamics of particles varies heterogeneously across the volume of the supercooled liquid. This figure is from [1].

$\rho = 1.2$ falls within the approximate range of $0.87 \leq T_o \leq 1.2$. While this onset temperature is not sharply defined in finite-dimensional systems, it marks the point where distinct differences in particle mobilities arise within the supercooled liquid [15, 13] leading to a broad distribution of relaxation times. Some regions of the liquid readily relax by orders of magnitude, compared to the system average, in response to perturbations, while others remain relatively immobile. The regions themselves fluctuate and evolve in time; a region can be fast-relaxing at one instant, and in the next instant, it is a slow-relaxing region.

This spatial heterogeneity in dynamics becomes increasingly pronounced upon cooling, with mobile particles forming clusters that grow significantly in size, often

exhibiting a string-like morphology [16, 17, 18]. This clustering behavior stands in stark contrast to the Gaussian distribution of particle mobilities observed in the high-temperature liquid phase. The spatial extent of these dynamic correlations can be substantial, with studies suggesting that a single-particle rearrangement can influence the dynamics of neighboring particles up to 5 to 10 particle diameters away [13, 16]. This lengthscale is long compared to how far local structural order extends (usually 1-2 particle diameters [19]), indicating the need for cooperative motion of many particles for structural rearrangements.

Since dynamical heterogeneity is not reflected in thermodynamic quantities like density, which remain largely homogeneous, one has to directly observe particle trajectories and their fluctuations over time [14]. The fact that the lengthscale for cooperative rearrangement is long compared to the lengthscale of structural order implies that simply observing the disordered arrangement of particles in a supercooled liquid or glass does not readily reveal which particles will exhibit greater or lesser mobility – what particle will rearrange or not. Given this challenge, a fundamental question arises: is there a hidden order parameter, encoded within the disordered structure, that can predict the observed dynamic heterogeneity [14]? In other words, is there a correlation between the local structural environment of a particle and its propensity for motion and relaxation? This question lies at the heart of efforts to understand the microscopic origins of heterogeneous dynamics and the complex relaxation behavior of supercooled liquids and glasses.

1.4 Relaxation Behavior in Supercooled Liquids

The radial distribution function, which characterizes the local density surrounding a particle at a given distance using a shell as shown in the cartoon in Figure 1.2, exhibits nearly identical peak positions and overall shape in both the high-temperature

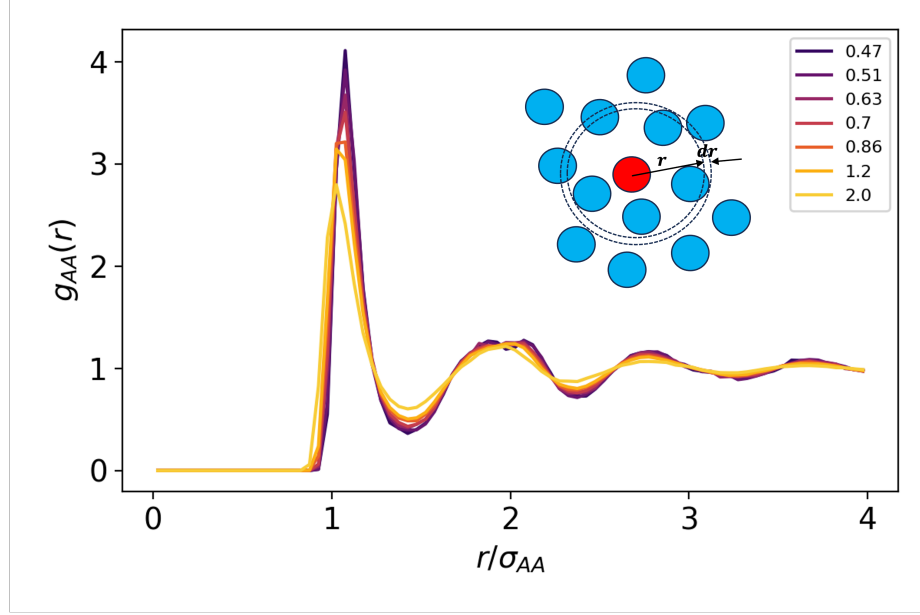


Figure 1.2: **Radial distribution function for a KA model as a function of radial distance.** The radial distribution function shows nearly identical peak positions and shape between liquid and supercooled states, with only a slight increase in nearest, next-nearest etc. neighbors upon cooling, indicating minimal changes in average spatial arrangement. Legend represents different temperatures. The cartoon illustrates how the local density at a distance r from a reference particle (in red) is calculated using a thin spherical shell of width dr .

liquid and supercooled states. As depicted in Fig. 1.2, while a slight increase in the number of nearest and next-nearest neighbors is observed upon cooling, their spatial positions remain largely invariant. This suggests that the average spatial arrangement of particles remains largely unchanged upon supercooling. The small linear change in the radial distribution over the temperature range does not commensurate with the (super-) exponential change in dynamics. The dynamic behavior of the liquid and supercooled states diverges dramatically. In a typical liquid, the response to a perturbation, such as an applied stress, decays exponentially, indicating a single relaxation process with a well-defined short timescale. Contrarily, supercooled liquids and glasses exhibit a profound slowing down of dynamics.

Additionally, on supercooling, dynamics qualitatively change in character, displaying a complex, multi-step relaxation process. This complex relaxation behav-

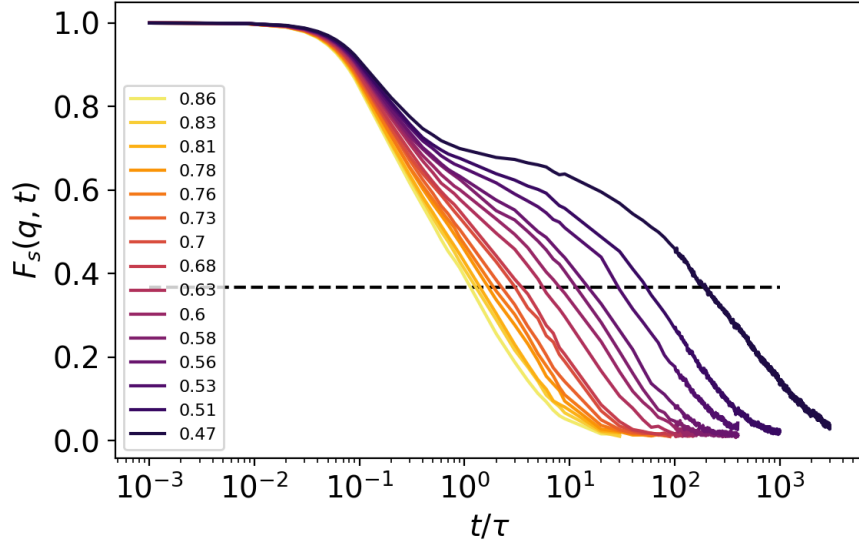


Figure 1.3: **Self-Intermediate scattering function for Kob-Andersen glass model at density $\rho = 1.2$ for different temperatures.** A two-step relaxation process become more obvious as temperature (legend) decreases. On supercooling, structural relaxation develops a stretched-exponential form, deviating from an exponential process and reflecting the complexity that develops. The dashed horizontal line show when $F_s(q, t) = 1/e$.

ior is evident in the self-intermediate scattering function, which probes the time-dependent correlations of particle displacements. Specifically, when we examine the self-intermediate scattering function, which is the Fourier transform of particle separations, at the wavevector corresponding to the first peak of the radial distribution function, a two-step relaxation pattern and a stretched-exponential profile emerge as the temperature is decreased, as illustrated in Figure 1.3.

The two-step relaxation process observed in supercooled liquids is characterized by distinct timescales associated with different modes of particle motion. The faster relaxation, termed the β -relaxation, is primarily attributed to the localized vibrational motion of particles within the “cages” formed by their neighbors. This process, with a characteristic timescale of τ_β , reflects the restricted mobility of particles in the dense supercooled state. The slower relaxation, known as the α -relaxation, attains a

stretched-exponential profile and arises from the cooperative rearrangement of particles. These cooperative relaxation lead to structural relaxation and the eventual flow of the supercooled liquid as indicated by the fact that the τ_α is related to the viscosity.

In “strong” glass formers, τ_α increases with decreasing temperature following an Arrhenius law:

$$\tau_\alpha \sim e^{\Delta E/T}, \quad (1.2)$$

where ΔE represents an effective activation energy barrier for structural rearrangements. This behavior reflects the thermally activated nature of particle rearrangements in these systems. That is, as temperature decreases, the probability of particles possessing sufficient thermal energy to overcome the fixed energy barrier decreases exponentially, which to the relaxation time increasing rapidly.

However, many glass formers exhibit a more complex temperature dependence, deviating from the Arrhenius trend as the temperature decreases. These “fragile” glass formers display a super-Arrhenius behavior, as illustrated in Figure 1.4 for the KA model and Figure 1.5 for several fragile glass-formers, where the increase in τ_α with decreasing temperature is more dramatic than predicted by the Arrhenius law. To capture this super-Arrhenius behavior, one can use equation 1.2, but with $\Delta E = \Delta E(T)$ to capture the super-Arrhenius behavior. Additionally, empirical functions such as the Vogel-Fulcher-Tammann (VFT) equation can be employed. The VFT equation, $\tau_\alpha \sim e^{A \frac{1}{T-T_{VFT}}}$, suggests a divergence of the relaxation time at T_{VFT} .

While the VFT equation is widely used, it is essential to recognize that it is an empirical fit, and other non-diverging functions [20, 21] can also be used to describe the observed temperature dependence of τ_α in fragile glass formers. The underlying mechanisms responsible for the super-Arrhenius behavior and the potential existence of a critical temperature remain active areas of research in the field of glass physics.

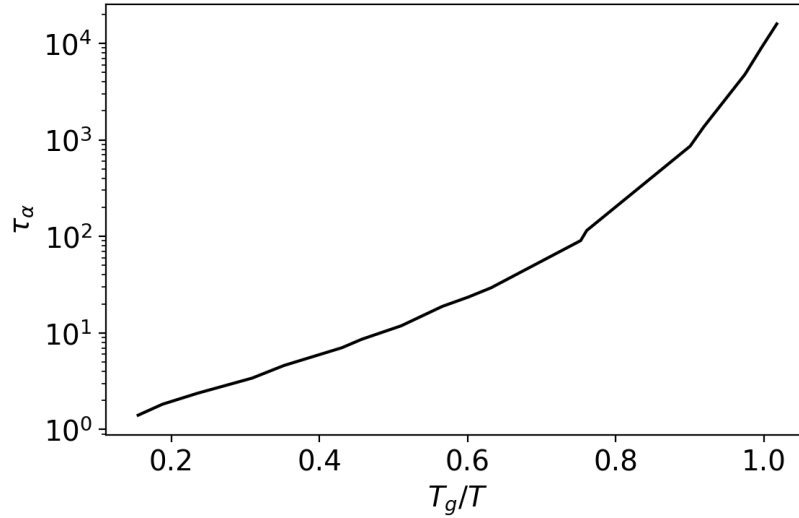


Figure 1.4: **Alpha-relaxation time against inverse temperature for the Kob-Andersen model.** The relaxation time for the KA model increases in a super-Arrhenius form as temperature decreases. For the temperature range for which structure changes linearly, there is a super-Arrhenius change in dynamics. The data points for this figure are from [2] for a KA model at density $\rho = 1.2$.

In contrast to the super-Arrhenius behavior exhibited by fragile glass formers, certain systems display a sub-Arrhenius temperature dependence of the relaxation time. In this regime, the relaxation time increases more gradually with decreasing temperature compared to the Arrhenius trend, indicating a weaker temperature sensitivity of the dynamics. While sub-Arrhenius behavior is very rare compared to the super-Arrhenius behavior, it has been observed in specific model systems, such as vitrimeric polymers [22]; Voronoi cell model; and the vertex model [23]. These models, which represent liquids as a collection of Voronoi cells or vertices, capture certain geometrical aspects of liquid structure and provide insights into the origins of sub-Arrhenius behavior.

The stark difference between the super-Arrhenius behavior of fragile glass formers and the sub-Arrhenius behavior of these model systems highlights the diversity of relaxation mechanisms in supercooled liquids. The dramatic growth of the relaxation time in fragile glass formers, without a corresponding significant change in the av-

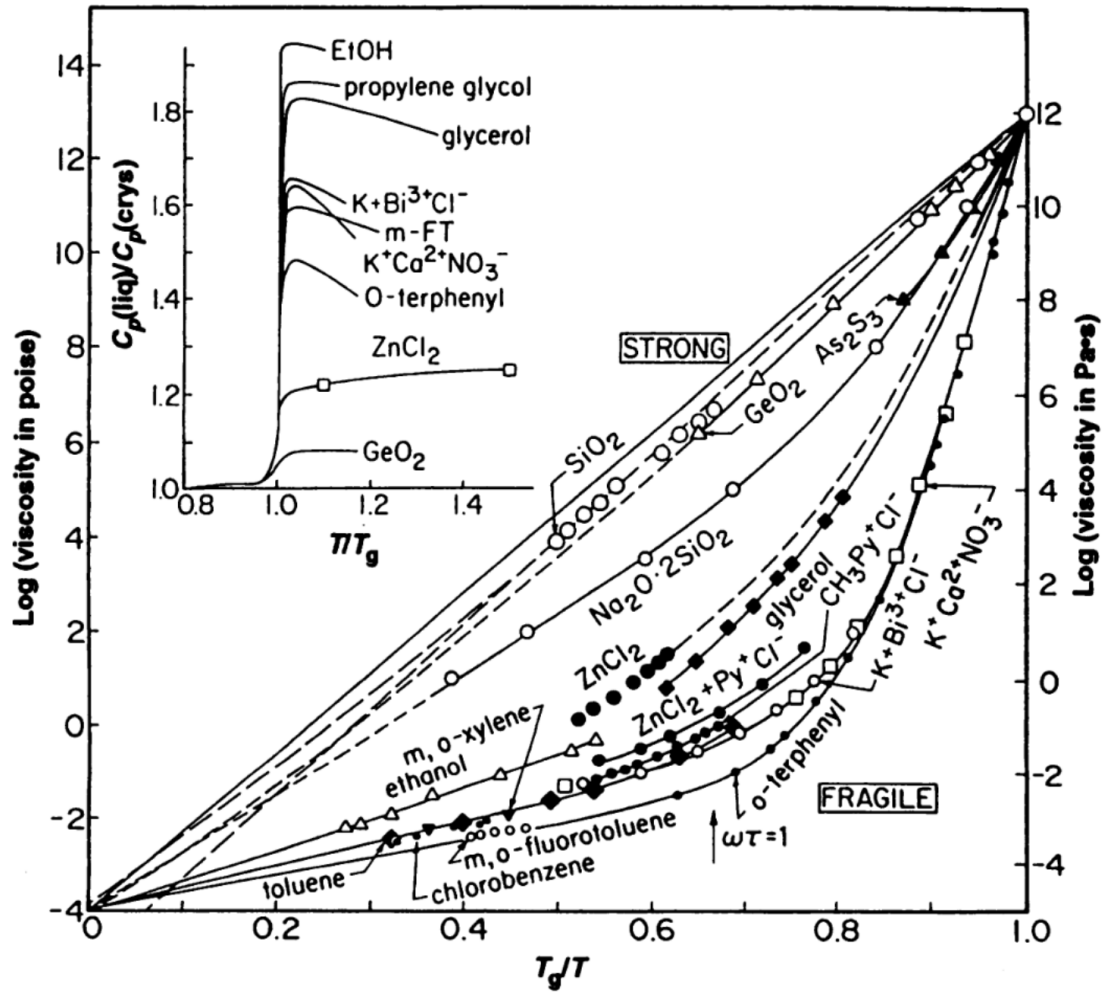


Figure 1.5: The viscosity of several glass-formers as a function of temperature. The viscosity, so also the relaxation time, for strong glass-formers grow in an Arrhenius fashion. Fragile glass-formers exhibit super-Arrhenius growth as temperature decreases. This figure is from [3].

erage structure, remains a puzzle and a subject of intense research. Understanding the microscopic origins of this super-Arrhenius behavior is crucial for developing a comprehensive theory of the glass transition and for predicting the behavior of these materials.

1.5 The Potential Energy Landscape

The dynamics of supercooled liquids and glasses can be conceptualized within the framework of the potential energy landscape, a high-dimensional abstract space introduced by Goldstein in 1968 [24]. This landscape, with its intricate topography of hills and valleys, represents the potential energy of the system as a function of the particle coordinates. The dimensionality of this landscape is determined by the degrees of freedom of the system; for a system of N particles in three-dimensional space with no internal degrees of freedom, the landscape is a function of $3N$ degrees of freedom. The $3N$ dimensions are the coordinates of the particles which describes the hypersurface and an extra one dimension is the height of the hypersurface.

The potential energy landscape provides a powerful tool for understanding the complex behavior of supercooled liquids. For a fixed number of particles N and fixed volume V , the landscape is also fixed. However, the way the landscape is sampled by the particle system is dictated by temperature [25, 15, 26, 27]. At high temperatures, the system explores the landscape through diffusive motion, readily crossing energy barriers and sampling a wide range of configurations. In this regime, the system is ergodic, meaning that the time-averaged behavior of a single particle is equivalent to the ensemble average at a given time. Mathematically, this is expressed as: $\langle A \rangle_{\text{time}} = \langle A(t) \rangle_{\text{ensemble}}$, where A represents a physical quantity, $\langle \dots \rangle_{\text{time}}$ denotes the time average, and $\langle \dots \rangle_{\text{ensemble}}$ denotes the ensemble average at time t .

As the temperature approaches the onset temperature of dynamical heterogeneity,

T_o , and decreases further, the system’s exploration of the landscape becomes increasingly constrained. The system begins to sample deeper minima, and the landscape’s topography starts to influence the dynamics. It is important to acknowledge the inherent complexity of the potential energy landscape, which may exhibit a hierarchical structure near the proposed Gardner transition [28, 29, 30], although the existence of this transition remains a topic of active debate. Due to the landscape’s rugged nature, characterized by a distribution of minima with varying energy barriers, the growing influence of the landscape on the system’s exploration leads to the emergence of complex relaxation behavior. With further cooling, the system eventually lacks the kinetic energy to overcome energy barriers, becoming trapped in local minima. At this point, ergodicity is broken, and the system’s dynamics are governed by the structure of the landscape [25].

This confinement to deeper minima leads to the stretched-exponential relaxation behavior observed in supercooled liquids, as exemplified by the self-intermediate scattering function shown in Figure 1.3 [25]. This stretched-exponential relaxation observed in supercooled liquids is commonly attributed to the spatial averaging of heterogeneous, locally exponential relaxation processes [31, 13]. This interpretation suggests that the non-exponential behavior arises from the superposition of different regions relaxing with distinct exponential timescales. However, recent studies have demonstrated that a combination of spatial averaging and locally non-exponential relaxation can also give rise to stretched-exponential behavior [32]. This finding highlights the complexity of relaxation dynamics in supercooled liquids and suggests that multiple mechanisms may contribute to the observed non-exponential behavior. However, at high temperatures where diffusion dominates, the system explores shallow minima, resulting in simple exponential relaxation.

In Chapter 2, we leverage on the concept that the exploration of the potential energy landscape is tied to local rearrangements of particles to demonstrate that even a

heated fluid, in the supercritical state where there is no distinction between liquid and gas phases, can retain information about the fine details of the landscape. That is, we investigate how local structure above T_o is predictive of the barriers between minima. This finding highlights the enduring influence of the landscape on the system's dynamics, even at high temperatures where the system exhibits seemingly simple diffusive behavior.

1.6 Challenge in defining an order parameter correlating structure with dynamics

While liquids and gases both exhibit continuous translational and rotational symmetry, they are distinguished by their density, a physical quantity that serves as an order parameter. Order parameters describe the degree of order in different phases of matter and play a crucial role in modeling phase transitions and ordering phenomena.

In the context of linking structure with dynamics in supercooled liquids and structural glasses, various order parameters have been explored. For example, the free-volume order parameter, which quantifies the amount of unoccupied space available for particle movement, has been shown to decrease with decreasing temperature, leading to reduced particle mobility [7, 8]. However, the correlation between free volume and dynamics has been shown to be weak [33, 34, 35, 36]. Another class of intuitive order parameters such as bond-orientational order parameters [37], has been used to identify locally favored structures in supercooled liquids and glasses. These order parameters quantify the degree to which the local arrangement of particles around a central particle resembles specific geometric motifs, such as icosahedra. While bond-orientational order parameters have shown some correlation with dynamics in certain glass-forming models [38, 39, 40, 41], the strength and nature of this correlation vary depending on the specific system under study [42].

These other parameters have been insightful, however, a key challenge remains: identifying an order parameter that explicitly and consistently correlates local structure with dynamics across a wide range of supercooled liquids and glasses. Subsection 1.6.1 provides a detailed account of a computational methodology employed to establish a correlation between structure and dynamics. Subsection 1.6.2 further elaborates on the “softness” parameter, a machine-learned order parameter central to our investigations.

1.6.1 Particle Mobility from Iso-configurational ensemble

One method to link structure with dynamics via particle mobility is through the iso-configurational ensemble [43, 44, 45]. The iso-configurational ensemble is implemented by keeping an initial configurational setup fixed and implementing various thermal trajectories of the configurational setup. The idea is as follows:

1. An initial configuration $R_0(t_o) = \{r_1^0(t_o), r_2^0(t_o), r_3^0(t_o), \dots, r_N^0(t_o)\}$ of N number of particles where r_i^0 is the position of particle i at the initial time $t = t_o$ is considered.
2. Various thermal trajectories of this initial configuration is made by drawing the initial velocities $V_k(t_o) = \{v_1^k(t_o), v_2^k(t_o), v_3^k(t_o), \dots, v_N^k(t_o)\}$ for trajectory k from the Maxwell-Boltzmann’s distribution.
3. The resulting set of configurations with position $R_k(t)$ and velocity $V_k(t)$ at time t is used to determine the impact of the local structure on dynamics by averaging out the noise associated with velocity. The isoconfigurational average of a dynamical quantity $A[R_k(t), V_k(t)]$ for M different thermal trajectories is then given as $\langle A(t) \rangle_{\text{iso}} = \frac{1}{M} \sum_{k=1}^M A[R_k(t), V_k(t)]$.

In our case, $A[R_k(t), V_k(t)]$ is the displacement of individual particles which we use to signify rearrangements. While the iso-configurational ensemble can be used to

test hypothesized order parameters connecting structure and dynamics [43, 45, 44], and can be used to bound the maximum effectiveness of any such order parameter [46], it does not on its own suggest what structural order parameter should be used. Nonetheless, the iso-configurational ensemble technique has proven to be quite useful in correlating structure with dynamics using machine learning tools as shown by works like [4, 47, 48, 49]. These works have used various machine learning models to show correlations between structure and the iso-configurational average of particle displacements. In chapter 3 we show how we utilize this methodology to identify a particle’s probability of rearrangements for which we analyzed the physical interpretability of various machine learning techniques.

1.6.2 Correlating structure and dynamics through machine learning

Data-driven approaches, particularly machine learning, have become increasingly valuable tools for unraveling the complex relationship between structure and dynamics in supercooled liquids. A seminal work by Schoenholz et al. [4] demonstrated the power of Support Vector Classification (SVC) to identify a structural order parameter that strongly correlates with particle dynamics. This order parameter, termed “softness,” is a quantitative measure of the local structural environment surrounding a particle.

The “softness” parameter is determined by encoding the local structural environment of particle i as a M -dimensional vector, F_i . The components of this vector can be quite generic. For instance, the environment is often quantified using techniques inspired by the pair correlation function, such as Gaussian weighting functions [50], which count the number of neighboring particles at various distances. For a particle

i , this two-point radial structure functions $G_X(i; r, \delta)$ [50] is defined as

$$G_X(i; r, \delta) = \sum_{j \in X} \exp \left(\frac{-(r - R_{ij})^2}{2\delta^2} \right), \quad (1.3)$$

where X denotes which of the components of the KA mixture is being considered, r is varied to describe local environment at different distances, and the parameter δ controls the width of the Gaussian shells. R_{ij} is the distance between particles i and j . Angular or other many-body correlations can also be incorporated through three-point functions, which capture the arrangement of triplets of particles. Subsequently, a dataset is constructed, comprising two distinct classes of local structures. Structures deemed unlikely to undergo rearrangement within a defined time window are labeled -1 , while those that rearrange in the near future are labeled $+1$. Since the correlation between structure and dynamics weakens at temperatures above the onset temperature, T_o , the training data set is typically drawn from deeper within the supercooled regime. This binary classification dataset is then employed to train a SVC algorithm. The SVC algorithm then seeks to find the optimal hyperplane, with normal vector w and bias b , that maximizes the separation between these two classes of structures in the feature space while minimizing misclassification. Mathematically, the softness S^i of particle i is defined as:

$$S^i = \sum_{\alpha=1}^M w_{\alpha} F_{\alpha}^i - b, \quad (1.4)$$

where F_{α}^i represents a component of the local structure F^i in the M -dimensional feature space.

Geometrically, the softness S represents the signed distance of a local structure from the hyperplane in the feature space. A positive softness value indicates a structure that is more likely to rearrange, while a negative value indicates a structure that is less likely to rearrange. This softness value, surprisingly, is then associated with an

energy barrier, ΔE , reflecting the energetic cost for a particle to undergo an activated rearrangement or “hopping” event. For example, based on the labeling of the training dataset stated above, a particle with a negative softness would be considered “hard” and require a higher energy barrier to rearrange, while a particle with a positive softness would be considered “soft” and require a lower energy barrier. This is a key result that has fueled extensive research in this area. There is no inherent reason, a priori, to expect that a distance in a classification hyperplane would have a direct physical interpretation as an energy scale to rearrangement.

To determine the relationship between softness and the energy barrier, the trained SVC model is applied to configurations at higher temperatures, up to T_o . The fraction of similar local structures that rearrange within a short time window, which represents the probability of rearrangement given a certain softness value, $P(R | S)$, is evaluated at each temperature. As shown in Figure 1.6, this probability exhibits an Arrhenius dependence on temperature:

$$P(R|S) = e^{\Sigma(S) - \Delta E(S)/T}, \quad (1.5)$$

where $e^{\Sigma(S)}$ is a prefactor related to the entropic contribution and the local curvature of the energy landscape. Studies have shown that the energy barrier ΔE is linearly related to the softness S , both in equilibrium [4] and out-of-equilibrium [51] conditions: $\Delta E = \epsilon_0 - \epsilon_1 S$.

This linear relationship allows for the construction of a free energy barrier, $\Delta F(S) = \Delta E(S) - T\Sigma(S)$, associated with the softness field. This free energy barrier field provides a valuable tool for understanding the dynamics of supercooled liquids, as it captures the interplay between energetic and entropic contributions to particle rearrangements. For example, [52] used this to show that the free energy for fragile glass formers increases on cooling while it decreases for a strong glass former.

The methodology of employing machine learning to identify structural predictors

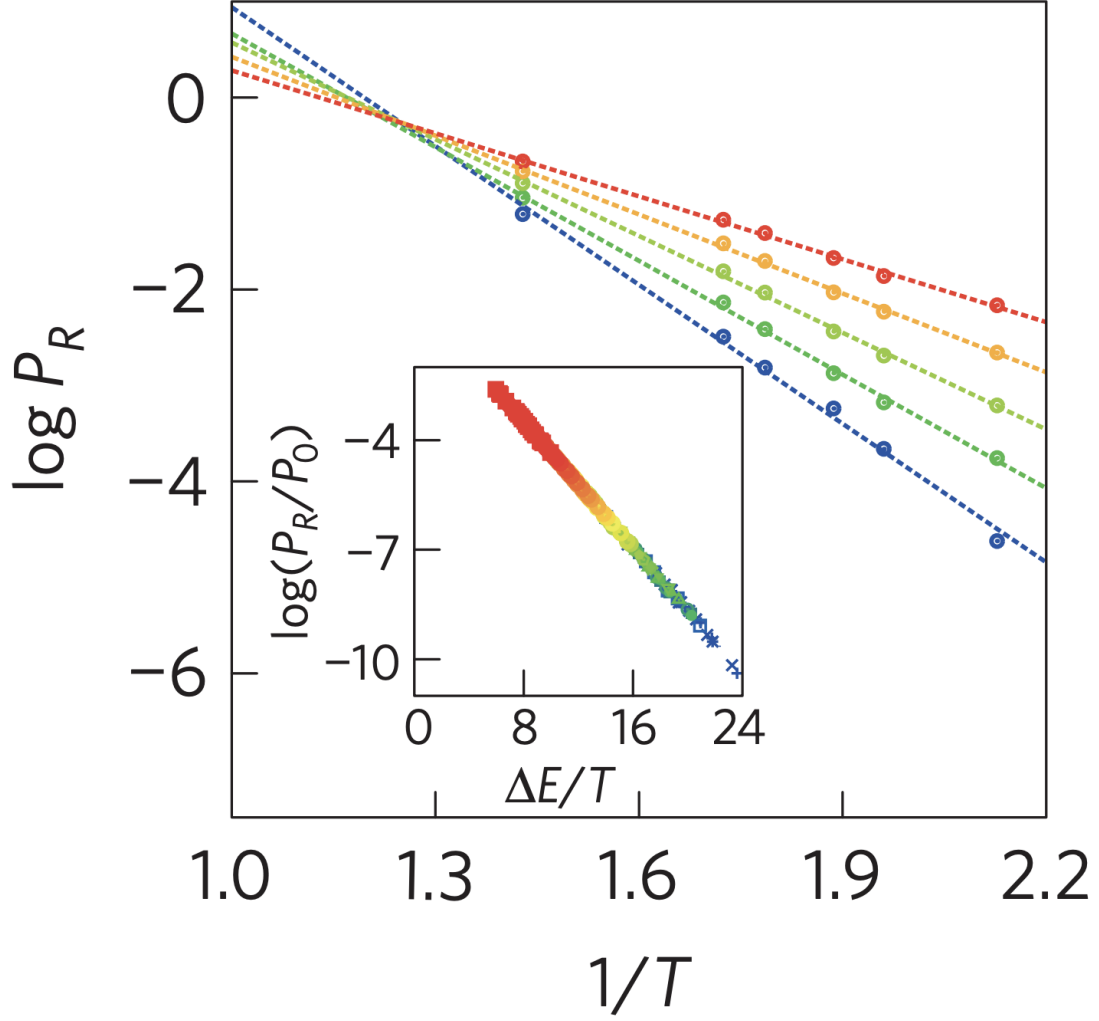


Figure 1.6: **The Arrhenius dependence of $P(R|S)$ on temperature.** The probability of rearrangements for different soft particles becomes orders of magnitude apart with temperature decrease. At the onset temperature, T_o (around 0.8 to 1.2), dynamics is uncorrelated with structure as shown by the merge of the Arrhenius fit (dotted line). The colored lines represent values of $S = -3$ (blue) to $S = +3$ (red). The inset show a collapse of $P(R|S)/P_o$ when plotted against $\Delta E/T$. Here, $P_o = e^\Sigma$. This Figure is from [4].

of dynamics, exemplified by the “softness” order parameter, has proven fruitful across a diverse range of supercooled liquid systems. This approach has been successfully applied to strong glass formers [53], systems with density-dependent fragility [54], and even sub-Arrhenius glass formers like the Voronoi cell model [54]. Furthermore, in the context of thin films, this methodology has been instrumental in disentangling the origins of faster dynamics near the film edges from the influence of local structure [55].

The ability to physically interpret the “softness” order parameter as encoding an energy barrier has enabled researchers to gain deeper insights into the underlying mechanisms governing glassy dynamics. For instance, “softness” has been used to explain the emergence of dynamic heterogeneity in supercooled liquids [56]. In another study, it provided crucial information about the role of near-field effects, specifically the changes in the local structure of neighboring particles, in controlling the ductility of glasses [57]. By tuning these near-field effects, for example, by introducing more randomness in the local structure, it is possible to manipulate the mechanical properties of the material.

Furthermore, “softness” has been employed to investigate the concept of facilitation, which is the idea that local rearrangements in a supercooled liquid can trigger or facilitate rearrangements in other regions of the system. Simple trap models which assume that particles hop between energy ‘traps’ with different heights, and that each hop is independent of the previous ones have been used to explain dynamics in glasses and supercooled liquids [58, 59, 60]. However, a trap-like model using ‘softness’ have shown that these simple models do not fully capture what happens in real supercooled liquids [61]. The models fail to account for the fact that one rearrangement can influence the likelihood of another. This study also suggests that supercooled liquids, even above the temperature where they become glasses, have a kind of ‘memory’ of past rearrangements. This ‘memory’ challenges the traditional view that such effects

are only important in the glassy state itself.

Beyond “softness,” other machine learning techniques have also contributed to our understanding of the structure-dynamics relationship in supercooled liquids. Bapst et al. [5] employed a graph neural network with a message-passing algorithm to predict the inherent state propensity, a measure of a particle’s long-time mobility. Their work revealed high correlations between the predicted and true propensities at timescales on the order of τ_α and showed that the length scale over which the model learns structural information grows with decreasing temperature, reflecting the increasing influence of structure on dynamics. Further work by Shiba et al. [47] showed that incorporating information about the edges of the graph into the loss function can further improve the correlation between the actual and predicted propensity.

The graph neural network implemented by Bapst et al. had at its core a message-passing algorithm which is the means by which a node learned by collecting information from other neighboring nodes. This message-passing algorithm initially posed challenges for physical interpretation, but it was demonstrated in [48] that it is equivalence to a particle receiving information about the average structure of its neighbors through a simpler Ridge regression model.

However, it is important to acknowledge that more sophisticated machine learning techniques do not always guarantee better answers to the question of correlating structure with dynamics. Alkemade et al. [62] demonstrated that a linear regression model, with fewer parameters than a graph neural network, can perform equally well in predicting dynamics when higher-order structural information, such as the average structure of a particle’s neighbors, is included. This finding suggests that while machine learning offers powerful tools for analyzing complex systems, there is a need for careful consideration of the relevant structural features and the choice of appropriate models.

1.7 Dissertation: Challenges in correlating structure with dynamics in supercooled liquids using machine learning tools

The need for careful feature and model selection, coupled with the broader challenges in understanding the physical meaning of its outputs, underscore the central motivation of this dissertation: to advance our understanding of the limitations and potential of machine learning in correlating structure with dynamics in supercooled liquids. While machine learning models like graph neural networks etc. have shown promise in predicting dynamics from structure in supercooled liquids, they often lack the clear physical interpretability afforded by methods like SVC. In the case of SVC, the distance of a structure from the classifying hyperplane was found to have a direct physical interpretation as an energy scale. This lack of interpretability in other models is often overlooked because the focus is primarily on the correlation between predicted and true dynamical quantities. However, a crucial open question remains: can these models, beyond simply predicting particle rearrangements, also provide insights into the underlying physics, such as the energy barriers associated with local structural rearrangements?

Even for SVC, where physical interpretability is more readily achieved, there are still gaps in our understanding of the process used to construct the classifying hyperplane. Previous work [5] utilized labeled structural data obtained from inherent states (energy-minimized configurations) to train the SVC model. However, obtaining a sufficient number of inherent state configurations for training can be computationally expensive. This raises several questions: Could training data derived from thermal configurations, which are more readily accessible, achieve comparable performance? Or do the thermal fluctuations inherent in these configurations obscure the relationship between structure and dynamics? Furthermore, is the assumption that structure

and dynamics are uncorrelated above the onset temperature, T_o , strictly valid? Or could structural features from the high-temperature, diffusive regime still encode information about the glassy dynamics at lower temperatures?

Beyond the training data, open questions regarding the nature of the feature space itself. While SVC partitions the feature space into regions that favor or disfavor mobility, the underlying organization of this space remains largely unexplored. Given that the basis vectors of the feature space are not orthogonal, to what extent are the different directions in this space correlated? How does the choice of kernel function influence the model’s performance and the structure of the feature space? And how does the dimensionality of the feature space impact the inferred energy barriers and the overall interpretability of the model? To summarize, significant open questions surround what aspect of the softness protocol are actually necessary and physical interpretability actually means.

This dissertation addresses these unanswered questions, delving deeper into the methodology of using SVC to predict dynamics in supercooled liquids. Through a combination of molecular dynamics simulations and advanced data analysis techniques, we investigate the influence of training data, the structure of the feature space, and the physical interpretation of the resulting models. By exploring these facets of the SVC approach, we aim to gain a more comprehensive understanding of the relationship between structure and dynamics in supercooled liquids and contribute to the development of more robust and interpretable machine learning models for complex systems.

1.8 Dissertation Outline and Contributions

This dissertation is structured as follows:

- Chapter 2: The objective of this chapter is to **investigate whether a classifi-**

cation hyperplanes trained on structural data from the liquid regime, where dynamics are purely diffusive, can discriminate between diffusive events and can effectively predict energy barriers for activated events in the supercooled regime. We introduce a novel approach based on binning local structures by their Z-score values to demonstrate the correlation between energy barriers predicted by different classification hyperplanes. These hyperplanes are trained on datasets from distinct temperature regimes, including the supercritical regime where dynamics are purely diffusive. Our analysis reveals that structural features associated with diffusive events can indeed predict energy barriers for activated events in the supercooled regime, challenging the conventional assumption that structure and dynamics are uncorrelated at high temperatures.

- Chapter 3: Building upon the Z-score approach developed in Chapter 2, this chapter centers on **investigating the physical interpretability of various machine regression-based learning models, beyond SVC, in terms of their ability to predict energy barriers for rearrangements in supercooled liquids**. We employ the iso-configurational ensembles to compare the performance of regression models like Ridge Regression, Support Vector Regression, and Multilayer Perceptron with Support Vector Classification (SVC) in predicting energy barriers. Our findings demonstrate that these regression models capture similar details of the energy landscape as SVC, further supporting the notion that the predictive power of these models stems from their ability to capture complex, high-dimensional relationships between structure and dynamics.
- Chapter 4: I present ongoing work focused on **investigating the presence of memory effects in supercooled liquids using the machine-learned**

order softness. This chapter explores the response of the system to thermal cycling protocols and discusses the potential implications of memory effects for understanding the dynamics of supercooled liquids and the nature of the glass transition. Our initial analysis of the response to thermal cycling protocols suggests the presence of memory effects in supercooled liquids, although further investigation is needed to confirm this observation. This would challenge the conventional view that memory effects are exclusive to the glassy state and may have implications for understanding the nature of the glass transition.

- Chapter 5: I summarize the key findings of this dissertation and discuss their implications for the field of supercooled liquids and glass physics. I also identify open questions and challenges that remain to be addressed, as well as promising avenues for future research. This chapter concludes by highlighting the broader implications of our work for understanding complex systems and developing more robust and interpretable machine learning models for materials science and other fields.

Chapter 2

Using fluid structures to encode predictions of glassy dynamics

This chapter draws from our research published in [63]. We investigate how the local arrangements of particles in a liquid, even in its freely flowing state, can be utilized to predict rearrangements upon cooling to the supercooled state, where flow is significantly restricted. Here, I will explain what motivated this work, how we created a training dataset from the liquid state, how our analysis compares to the established “softness” approach, and what our findings mean for our understanding of glassy systems.

2.1 Introduction

Consider a common liquid: its constituent molecules are in perpetual motion, continually shifting and rearranging. While the liquid appears macroscopically homogeneous, significant local structural variations exist at the molecular level. Certain molecules reside in environments that facilitate diffusion exceeding the average rate, while others become momentarily confined within regions of diminished mobility. Upon cooling, this liquid may undergo crystallization, forming a highly ordered lattice. However,

if crystallization is bypassed, the liquid may instead transition into a glassy state, akin to the glass windows. Glasses are intriguing because they retain the disordered, liquid-like structure [64], yet they exhibit the mechanical rigidity of a solid. They're not crystalline, but they do not flow like a liquid either, at least not on experimentally relevant timescales. This dichotomy arises because intermolecular interactions constrain molecular motion as temperature decreases, leading to molecules becoming kinetically arrested in specific configurations. Consequently, the dynamics of a glass are radically different from those of a simple liquid [3], although their structures may appear superficially similar.

As the liquid is cooled, a critical temperature known as the onset temperature (T_o) emerges, signaling a fundamental shift in molecular motion [15, 25, 65, 66]. Above T_o , the system behaves much like a normal liquid, with molecules moving in a largely uncorrelated, random manner. However, below T_o , the dynamics undergo a dramatic transformation. Instead of uniform motion, the system exhibits dynamical heterogeneity [20, 13, 14, 67, 12]: certain molecules become significantly more mobile, relaxing orders of magnitude faster than the system average, while others remain nearly frozen. This heterogeneity manifests both spatially, with some regions more mobile than others, and temporally, as individual molecules alternate between periods of mobility and confinement. Furthermore, below T_o , the local arrangement of molecules plays a critical role. The specific configuration of neighboring molecules strongly influences the way in which a molecule can move. Some local structures facilitate motion, while others effectively cage the molecule, hindering its movement. And for certain types of glass-forming liquids, known as "fragile" glass formers, the slowing down of molecular motion below T_o becomes super-exponential - much faster than a simple Arrhenius law.

A central theme in the study of supercooled liquids and glasses is the connection between structure and dynamics [46, 68, 69, 70, 71, 45]. How does the arrangement

of molecules influence their movement? Researchers are tackling this question using a variety of techniques, including sophisticated data-driven approaches. These methods often involve large-scale computer simulations to generate the necessary data for training machine learning algorithms such as Support Vector Machines (SVMs) [72, 4, 55] or Graph Neural Networks [5], among others [49, 48]. These techniques have shown considerable promise in identifying structural features that correlate strongly with dynamics across different timescales, and have been applied to various systems, including “strong” glass formers [53], “fragile” glass formers [4, 5, 73], and even more exotic, biologically-inspired models [54]. The work by [4] using linearSVMs is of particular interest. They showed that the SVMs learn to recognize patterns in the structure and assigns a “softness” value to each molecule. “Softness” acts like a local order parameter, quantifying how likely a molecule is to rearrange. Geometrically, it is the signed distance of a local structure to linearSVM classifier. Importantly, “softness” has been argued to be physically interpretable. This physical interpretation arises from training the algorithm at low temperatures and then applying it to data at other temperatures. By finding the strongest correlations between structure and dynamics, the algorithm effectively learns combinations of structural features that can be understood as representing a local energy barrier that must be overcome for a molecule to rearrange.

However, some question remains: Why do these methods work? What are they really telling us about the physics of glasses and supercooled liquids? And can we use these techniques to learn something about the liquid state before it becomes a glass? One key unresolved questions include why these particular approaches lead to what is apparently a local order parameter for the supercooled liquids, and how the learned energy barriers actually depend on the construction of the classifiers.

In this chapter, we explore this question by applying machine learning technique (SVM), typically used for supercooled liquids, to the liquid state above the onset

temperature. We ask: Can we find patterns in the liquid structure that predict how the molecules will behave later, when the liquid is cooled down? This is equivalent to asking if one can find patterns in the vapor/liquid phase that can predict dynamics in the supercooled or glass phase. Surprisingly, the answer is yes! We first show that the same machine learning techniques that have successfully correlated structure and dynamics in the supercooled phase can be used to classify “extreme diffusive” events even far above the onset temperature. In the spirit of a transfer learning approach, we show that these liquid-state classifiers can statistically identify activated events in the supercooled phase, even though the character of the activated dynamics below T_o changes dramatically. We further show that not only can accuracy on a classification task be maintained, but that the physical interpretability is maintained: apparently fluid-phase classifiers also learn energy barriers in the super-cooled phase.

2.2 Methods

In Section 2.2.1, we provide a detailed description of the glass model employed in our simulations and the simulation methodology itself. Section 2.2.2 elaborates on the methods used to quantify both dynamics and local structure. Finally, Section 2.2.3 outlines the procedure for constructing the training datasets used in our machine learning analysis, specifically detailing how we determine the datasets above and below the onset temperature.

2.2.1 Model and simulations

To study the behavior of our model liquid, we used computer simulations. Specifically, we performed a large number of molecular dynamics simulations, tracking the positions and displacements of $N = 4096$ particles. We used a well-established model for glass-forming liquids called the 80 : 20 Kob-Andersen model [11] (with a cutoff

distance of 2.5). This model represents a mixture of two types of particles (80% of type A and 20% of type B) that interact via the Lennard-Jones (LJ) potential. We set the density of the liquid to $\rho = 1.2$ which is typical of the density used in these studies. Our simulations were performed in a virtual box with periodic boundary conditions; periodic boundary conditions ensured that our results were not affected by the edges of our simulation box.

In our simulations, we used a standard set of dimensionless (reduced) units, often called Lennard-Jones (LJ) units, to measure distances, energies, and masses. These units can be thought of as convenient units tailored to the model. The base units for distance is measured in units of the large particle diameter, σ_{AA} ; energy is in units of the interaction parameter, ϵ_{AA} ; and mass is in units of the particle mass, m . For this model the dimensionless onset temperature is often reported as $T_0 \approx 0.87$ [20, 4]; given the broad crossover in the dynamics, values between 0.8 and 1.2 are also reasonable estimates for this temperature scale [25, 12].

Our simulations were performed at constant number of particles (N) and volume (V) which implies that we explore the same potential energy landscape. We also kept the temperature (T). Keeping N, V and T constant implies our simulations were performed in the canonical ensemble NVT . In this ensemble, the system is connected to a fictitious thermal reservoir, making temperature fluctuations possible while maintaining the average energy of the system. The temperature is kept constant using a deterministic Nose-Hoover thermostat [74] which is expected to not affect our main results. We expect that if anything our results would improve quantitatively if we used a constant N, V and energy (E) ensemble, that is, the microcanonical (NVE) ensemble. We examined how our system explores the energy landscape by varying temperature in the range $T \in [0.45, 2.0]$.

To prepare our simulations, we started with a random arrangement of particles and let the system equilibrate (i.e. settle into a stable state) for 5000τ at a temperature

of $T = 0.45$. We then used this equilibrated configuration as the starting point for simulations at other temperatures. At each temperature, we again allowed the system to equilibrate for 1000τ before collecting data. We saved the particle positions and other relevant information at regular intervals of 1τ during the simulations.

2.2.2 Local structure and dynamics

To understand the local environment around each particle, we used a function which, based on the radial distribution function, essentially count the number of neighboring particles of each type at different distances from a central particle. This two-point radial structure functions $G_X(i; r, \delta)$ [50] is defined for a target particle i as

$$G_X(i; r, \delta) = \sum_{j \in X} \exp\left(\frac{-(r - R_{ij})^2}{2\delta^2}\right), \quad (2.1)$$

where X denotes which of the components of the binary mixture is being considered, r is a parameter controlling the distance from which dominant contributions to the feature come, δ is a parameter controlling the width of the Gaussian shells, and R_{ij} is the distance between particles i and j . We characterize the local environment of particle i as a vector in a 100-dimensional feature space, $\vec{F}_i$, with $\delta = 0.2$, $0 < r < 5$ in increments of 0.1, and $X = A, B$. Each feature is standardized [75] so has zero mean and unit variance at the training temperature.

To measure how much particles moved, we used a quantity called p_{hop} ; to be consistent with work on activated dynamics as introduced in Ref. [76]. This quantity measures the displacement of a particle over a certain time window. We use an observational time window of 10 LJ time units, for which

$$p_{hop}(i, t) = \sqrt{\langle (\vec{r}_i(t) - \langle \vec{r}_i \rangle_{w_2})^2 \rangle_{w_1} \langle (\vec{r}_i(t) - \langle \vec{r}_i \rangle_{w_1})^2 \rangle_{w_2}},$$

where $w_1 = [t - 5, t]$, $w_2 = [t, t + 5]$, and thus $\langle \cdots \rangle_{w_i}$ averages over one half of the observation window. We chose this time window because it is long enough that it excludes intra-basin vibrations and short enough to identify rare rearrangements compared to numerous rearrangements that occur at τ_α for the supercooled regime. A large value of p_{hop} indicates that the particle has moved a lot, while a small value indicates that it has stayed close to its initial position. We do not believe that using p_{hop} as a dynamical label is crucial; we show in Chapter 3 that choosing instead to measure particle dynamics using their cumulative displacement over the same time window leads to qualitatively identical results.

2.2.3 Machine learning protocol

We train SVMs connecting structural features with dynamic observables largely following the “softness” methodology [4]. We build a training set by combing through MD trajectories for examples of dynamically active (“rearranging”) and inactive (“non-rearranging”) particles, and train a linear soft margin SVM (using the Scikit-learn package [77]) to classify these examples. We can then use the learned classifier (here: a hyperplane in feature space) to try to predict dynamics based on a particle’s instantaneous environment, and we define the softness of particle i at time t , $S_i(t)$, as the shortest distance between its vector of structural features and this classifying hyperplane.

The training set construction for our “softness” classifier closely followed the protocol outlined in Ref. [4]. We constructed a balanced 7600-sample training set using the coldest temperature considered ($T = 0.45$): 3800 rearranging samples and 3800 non-rearranging samples. We adopted the previously-used convention of associating the structural data of a particle i at time $t - 2\tau$ with the dynamical state at time t . We defined a rearranging particle if, at time t , $p_{hop}(i, t) > p_c$ where $p_c = 0.2$. We defined a non-rearranging particle by requiring its $p_{hop}(i, t)$ value to remain less than

a lower threshold of $p_l = 0.0085$ for at least 120τ duration of time. We then used the local structure of the non-rearranging particle in the middle of its time of low activity. Unlike in previous work, we take structure and p_{hop} values directly from the thermal configurations rather than quenching to the inherent states (in part because the fluid-phase simulations would be far from any minima). Unless otherwise stated, we used a soft-margin misclassification hyperparameter of $C = 10^{-2}$.

A major finding of Ref. [4] was that this signed distance – softness – encodes the probability the target particle would rearrange at a given temperature. The corresponding curves for the probability of rearranging at different values of S as a function of T all intersected at a common temperature, which in turn suggested the existence of an onset temperature above which structure was no longer predictive of dynamical events. The predicted value of T_0 was consistent with alternative definitions [25, 20, 4] and with the numerical values cited above. Before we return to this finding, we first ask: Can we learn to classify dynamical events based on structure not in the supercooled regime but at and even *above* the onset temperature?

For $T > T_0$ individual particle motion is diffusive rather than activated, and it is not clear that using p_{hop} as a dynamical label is the most natural choice. We continue to use it as an indicator function – it is still large for dynamical trajectories that move a particle far from its initial position and small for diffusive motions that stay near a particle’s initial position – and will show in a later work that this choice is not crucial to our results. To identify “extreme events” to classify, we select particles in high and low tails of the probability density function of p_{hop} at different temperatures. To have similarly sized training sets as in the case of softness, we choose lower and upper cutoffs (p_l and p_u) that captured the most extreme 0.033% of low- and high-activity events, respectively. We identify particle i at a given time t as “extremely diffusive” if $p_{hop}(i, t) > p_u$, and associate it with the particle’s local structure at time $t - 2\tau$. Similarly, if $p_{hop}(i, t) < p_l$, the particle is identified as “extremely non-diffusive,” and

its structure at time $t - 2\tau$ is included in the training set. The training set for each temperature we considered above T_0 contained 10400 balanced samples.

Aside from this difference in choosing “rearranging” and “non-rearranging” labels, we follow the methodology above: we find a linear soft margin SVM that best classifies a labeled training set, and then apply this classifier to new data. To distinguish it from softness we call the distance of a point in feature space to such a classifying hyperplane the “fluidity,” and we use $\mathcal{F}_i^T$ to denote the fluidity of particle i with respect to a classifier trained from data at temperature T . Given any of our classifiers, one can compute a particle’s softness or fluidity by computing its feature vector (which depends only on the instantaneous structure around the particle) and evaluating $\alpha_i(t) = \vec{w}_\alpha \cdot \vec{F}_i(t) - b_\alpha$, where $\vec{w}_\alpha$ is the normal vector and b_α the bias defining a classifying hyperplane, and where α refers to either softness S or a fluidity $\mathcal{F}^T$. We note that the term “fluidity” has previously been used to describe an average rate of plastic events in models of soft glassy rheology [78]; while our definition is different, we will see that highly “fluid” particles have more active dynamics at high temperatures and, indeed, are more likely to undergo plastic rearrangement events at low temperatures.

2.3 Classification in the fluid phase

We find that we *can* learn to classify extreme diffusive events even far above the onset temperature using local structure. Hyperplanes are characterized by a normal vector and a bias; the direction of the normal corresponds to the linear combination of features that has been learned, and the bias is an offset that best separates the training set given that direction. We expect the *direction* to encode the key physical features governing rearrangements, whereas we expect the *bias* may be strongly dependent on details such as the choice of time window or the temperature of the training set. For

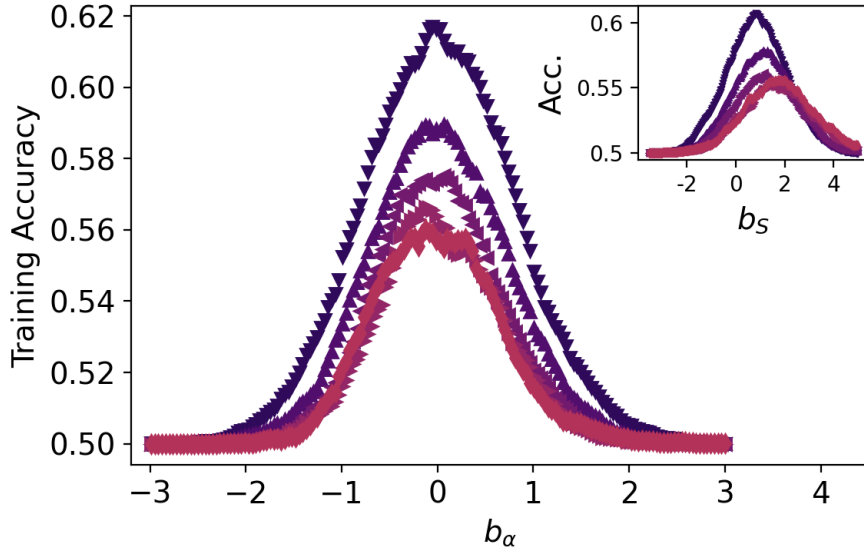


Figure 2.1: **Fluidity classifies rare events at high and low temperatures.** The points show the 5-fold cross-validation accuracy of linear SVMs trained on extreme diffusive samples at $T = 1.0, 1.2, 1.4, 1.8, 2.0$ (dark blue to light red) as a function of the classifier’s bias. Each classifier achieves near-peak accuracy for small values of the bias. In contrast, the inset shows the test classification accuracy of the “softness” classifier trained on activated dynamics at $T=0.45$ and applied to the extremes of diffusive events at different temperatures, for which very different values of the bias optimize performance.

instance: with our fixed-threshold definition of a rearrangement the total number of rearranging particles increases as T increases, so even if the same underlying structural variable controls rearrangements the optimal bias of the hyperplane will shift to maximize the soft margin in the training set data. Because of this, we want to remove the influence of the bias on our later results. In Fig. 2.1 we show the training accuracy of fluidity as a function of the bias, and during our transfer learning approach later we will select values that maximize our classification accuracy not on the training but on a low-temperature test set.

It is noteworthy that at such high temperatures, *any* structural features predictive of dynamics can be found. We find that even a softness classifier – i.e., a classifier trained on activated dynamics – has some ability to classify diffusive events in the fluid

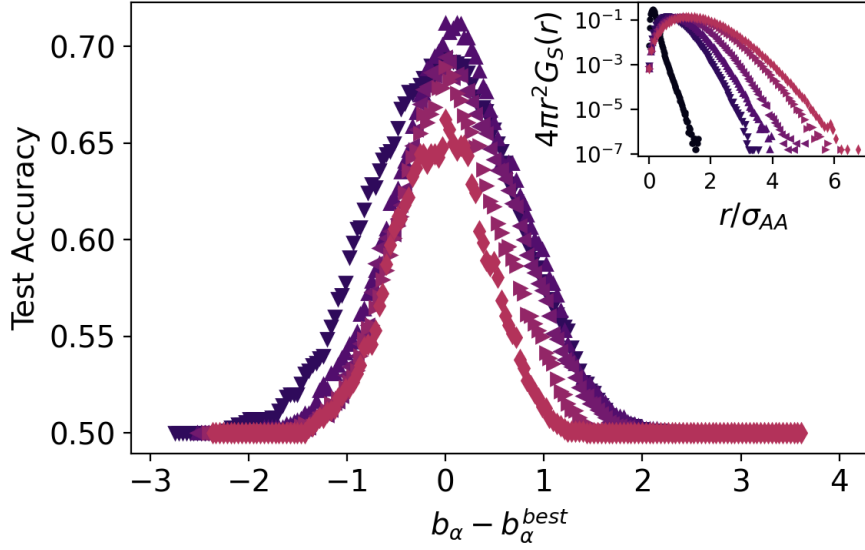


Figure 2.2: **A transfer learning approach connects extreme diffusive events above T_0 with activated dynamics below T_0 .** The main figure shows the test accuracy of linear SVMs, trained on extreme diffusive samples at $T = 1.0, 1.2, 1.4, 1.8, 2.0$ (dark blue to light red), as applied to a test set of activated dynamics at $T = 0.47$ as a function of the classifier’s bias relative to the optimal bias for that choice of temperature. The inset show the self part of the Van-Hove correlation function for large particles at a time scale of 10 LJ time units. (Black dots are for $T=0.45$.)

phase: as shown in the inset, the accuracy on the high- T training sets is almost as good as the classifiers trained at those temperatures. The optimal bias that needs to be chosen is quite different, but the *direction* in feature space learned is quite similar. This finding encourages us to more explicitly frame a transfer learning task from the high-temperature to the low-temperature regime. Concretely, we apply the fluid-phase classifiers – trained at temperatures ranging from $T = 1$ to $T = 2$ – to labeled data from $T = 0.47$. As shown in Fig. 2.2, even though we have trained on data well above T_0 we find that our classifiers maintain substantial classification accuracy. Again, the optimal bias varies strongly with training and testing temperature, but the *direction* in feature space is extremely highly correlated.

To highlight how surprising this is, in the inset we show the self part of the van

Hove function characterizing single-particle displacements, choosing as a time scale the same window we used for p_{hop} . At high temperatures this distribution is essentially Gaussian and involves a substantial numbers of particles moving *many* times their own size; at low temperatures this distribution is non-trivial and has an exponential tail corresponding to hopping motions whose size is less than a single particle diameter. This is reflective of the fact that diffusive events and activated events arise from mechanisms that are fundamentally different. Diffusive events are often driven by thermal fluctuations without significant energy barriers. In contrast, activated events involve rare, collective rearrangements that require surmounting an energy barrier, making them non-trivially different from diffusive motion. The successful prediction of activated events by the model suggests that the model has learned some underlying structural features that are shared between the two processes.

2.4 Interpretability of fluidity and softness

Using $T > T_0$ classifiers we are able to obtain reasonable accuracy on training sets (which are by definition constructed from atypical particles at the various training temperatures), but similar to the distribution of softness shown in Ref. [4], the distributions of fluidity, when measured at different test temperatures, remains approximately Gaussian. This is shown in Fig. 2.3 for two training temperatures ($T = 1.2$ and $T = 2.0$). The mean of the distribution behaves monotonically as the test temperature changes as was equally observed in [4].

Remarkably, we find that fluidity has the same kind of physical interpretability as softness. We define a rearrangement as a particle having an instantaneous value of $p_{hop} > p_c$, and fit the probability of rearranging, P_R , to a Kramers form [79]: $P_R = \frac{1}{T} \exp(\Sigma(\mathcal{F}^T)) \exp(-\Delta E(\mathcal{F}^T)/T)$. Just as for softness, we show in Fig. 2.4 that fluidity partitions the overall system dynamics into a collection of barrier-hopping

processes characterized by an energy barrier scale (ΔE) and an entropic contribution (Σ). We also find that our prediction of the onset temperature itself – whether from the intersection of the Kramers form fits or more qualitatively from where the data collapses – is the same across our softness and fluidity classifiers, suggesting that a consistent physical interpretation is being learned.

The identification of a scalar value – fluidity – that encodes the energy barrier characterizing an activated process by training a classifier on diffusive events is striking. Given the cross-over nature of the onset temperature, perhaps this qualitative result could have been expected for training temperatures close to T_0 , but it holds even when training far above T_0 , as shown in Fig. 2.4a. How do the energy barriers learned by these classifiers compare to the energy barriers learned by classifiers trained on supercooled data, i.e., to those from softness? A direct answer to this question is complicated by two aspects of the training and testing procedure.

The first is that there is no reason to think that the hyperplane *bias* should be held constant when moving from one task to another. This is implicit in the relatively large shifts in bias needed in the inset of Fig. 2.1 and in the test accuracy for sub-optimal choices of bias in Fig. 2.2. The second issue relates to the fact that we study systems across such a wide temperature range that the distribution of the structural features changes substantially (a similar issue arose in the context of applying classifiers to systems at different *densities* [52]). To account for these, we compare the physical interpretations of the different classifiers by defining $x_\alpha = \left(\vec{w}_\alpha \cdot \vec{F} - b_\alpha^{best} \right) / \sigma_\alpha$. That is, we adjust the bias to the optimal value when the classifier is applied to a common ($T = 0.47$) training set, and rescale the feature vector by the standard deviation of the distribution of fluidity (or softness) at the training temperature. With this choice, in Fig. 2.5 we show that the learned aspects of the landscape associated with particle structure – including both the energy barrier and entropic contribution – are almost *identical*. We note that fitting the data only in the regime unambiguously below the

T	b	σ_T	b^{best}	$\hat{w}_S \cdot \hat{w}_\alpha$
0.45	0.0096	1.1188	0.55710	1.0000
1.0	-0.0056	0.7653	-0.2385	0.8246
1.2	-0.0171	0.7028	-0.4644	0.8071
1.4	-0.0099	0.6760	-0.4644	0.6643
1.8	-0.0262	0.6286	-0.5774	0.7104
2.0	-0.0341	0.6286	-0.6151	0.7198

Table 2.1: **Table of the optimal bias, b , of the hyperplane during training; the standard deviation of fluidity (or softness for $T=0.45$) at the training temperatures; and the optimal bias, b^{best} , that maximizes test accuracy at $T=0.47$. The final column displays the projection of our classifiers onto the softness classifier.**

onset temperature – i.e., the points for which $T < 0.8$ – does not qualitatively change these results. We speculate that there may be some correlation between training at higher temperatures and a hint of a slight curvature in the data, but do not yet have sufficient data to confirm this.

Given these results, Table 2.1 reports several values that contribute to the formation of the results: the bias (i.e., the bias that achieves the highest accuracy during training), standard deviation of fluidity and softness, and the optimal choice of bias when the classifier is applied to a test set at $T = 0.47$. We also report, as a simple measure of the similarity of the classifiers, the dot product between the normal vector describing each classifier and that of the softness classifier. From the dot product, it is clear that these classifiers point nearly in the same direction in the high dimensional feature space.

2.5 Discussion

Taken together, our work establishes a surprising connection between the structural features that control activated events at low temperatures and those apparently responsible for the tails of the distribution of diffusive events above the onset temperature. Although many approaches have considered the link between local structural

arrangements and dynamical arrest in the supercooled regime [41], much of this knowledge is set aside when studying the liquid phase. Our finding that structure is relevant even above the temperature of liquid-gas critical point (roughly $T = 1.2$ in this model [80]) suggests that further pursuing this avenue of research may prove fruitful. The connections between structure and dynamics across temperatures that we find may be a consequence of the only modestly growing structural length scales over the temperature range studied, but we again emphasize that the qualitative character of the dynamics changes significantly over these same temperatures.

A natural hypothesis might be that our classification accuracy stems from an ability to identify fluid phase particles that do not diffuse very much: perhaps we are identifying rare particles that consistently sample a similar, high-barrier part of the energy landscape, and are not truly distinguishing both immobile and highly-mobile particles? We show in the Appendix that this hypothesis fails, and that using both tails of the diffusive-motion distribution is crucial to our results. We comment that our main finding – that one can take a classifier built on fluid-phase data without barrier-hopping dynamics, apply it to data in a dynamically heterogeneous phase, and infer the existence of energy barriers there – is reminiscent of the results reported in Ref. [54]. That work considered a biologically-inspired model with highly unusual glassy dynamics [23, 81, 82] meant to mimic the behavior of dense cellular materials. There it was speculated that it was the anomalous, sub-Arrhenius behavior of the model that was responsible for the success of the transfer learning task; the results presented here suggest an alternative explanation may be needed.

Our work highlights what we believe continue to be crucial unanswered questions: *why* do these machine learning methodologies learn simple structural order parameters that correspond to local energy barriers in disordered phases of matter? What aspects of the training lead to this result? And to what extent can we use this result to uncover new, relevant descriptions for the physics of amorphous solids? We note

that there is some indication that the specific methodology used here and earlier – linear SVMs – may not be crucial to recover this physical interpretation; Ref. [48] hinted at a similar result using a GNN-inspired linear-regression-based model. We believe it will be crucial to compare different machine learning techniques as applied to predicting glassy dynamics [62, 5] not only along dimensions of predictive capacity, generalizability, efficiency, and training cost, but *also* in terms of their physical interpretability.

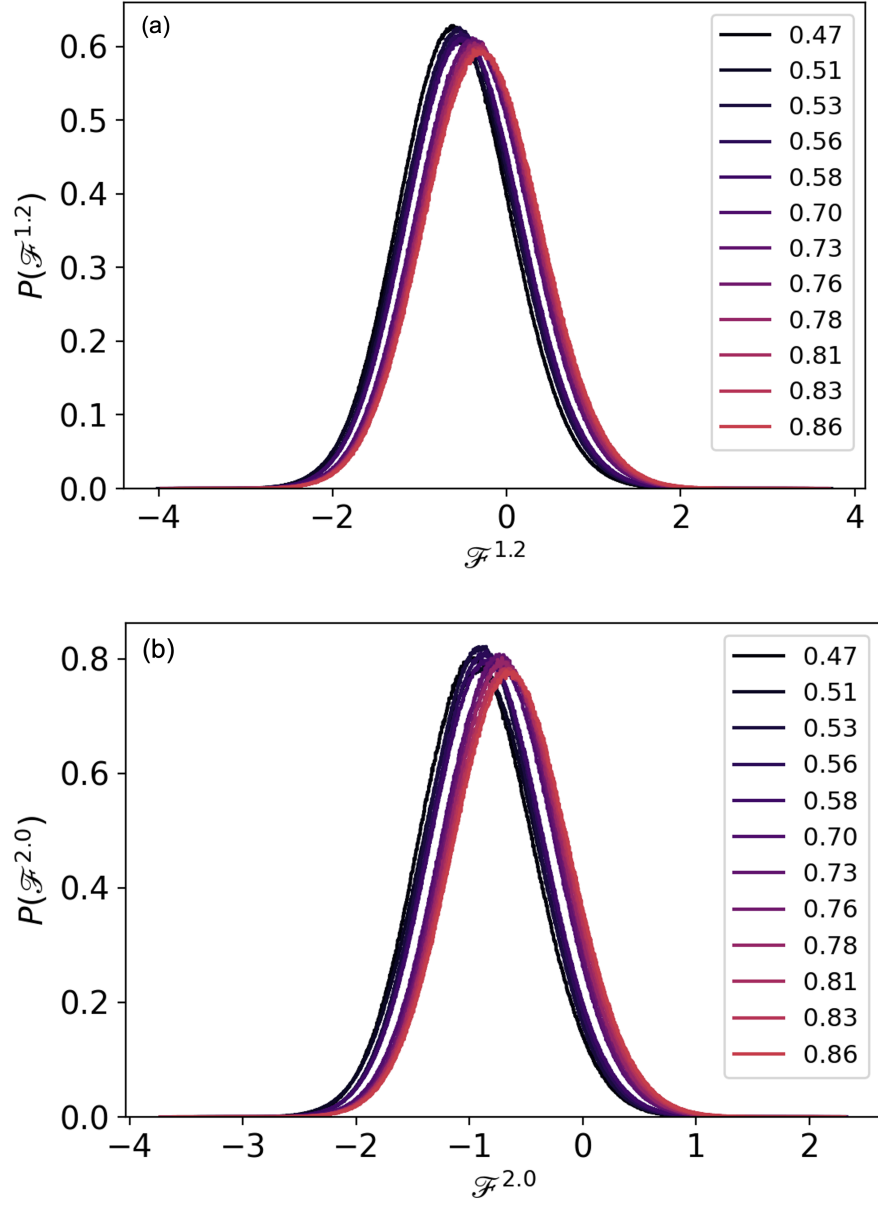


Figure 2.3: **The distribution of fluidity at different test temperatures for two considered training temperatures.** Training at both $T = 1.2$ (a) and $T = 2.0$ (b) the distribution of fluidity is approximately Gaussian. The mean is a monotonic function of the test temperature (legend).

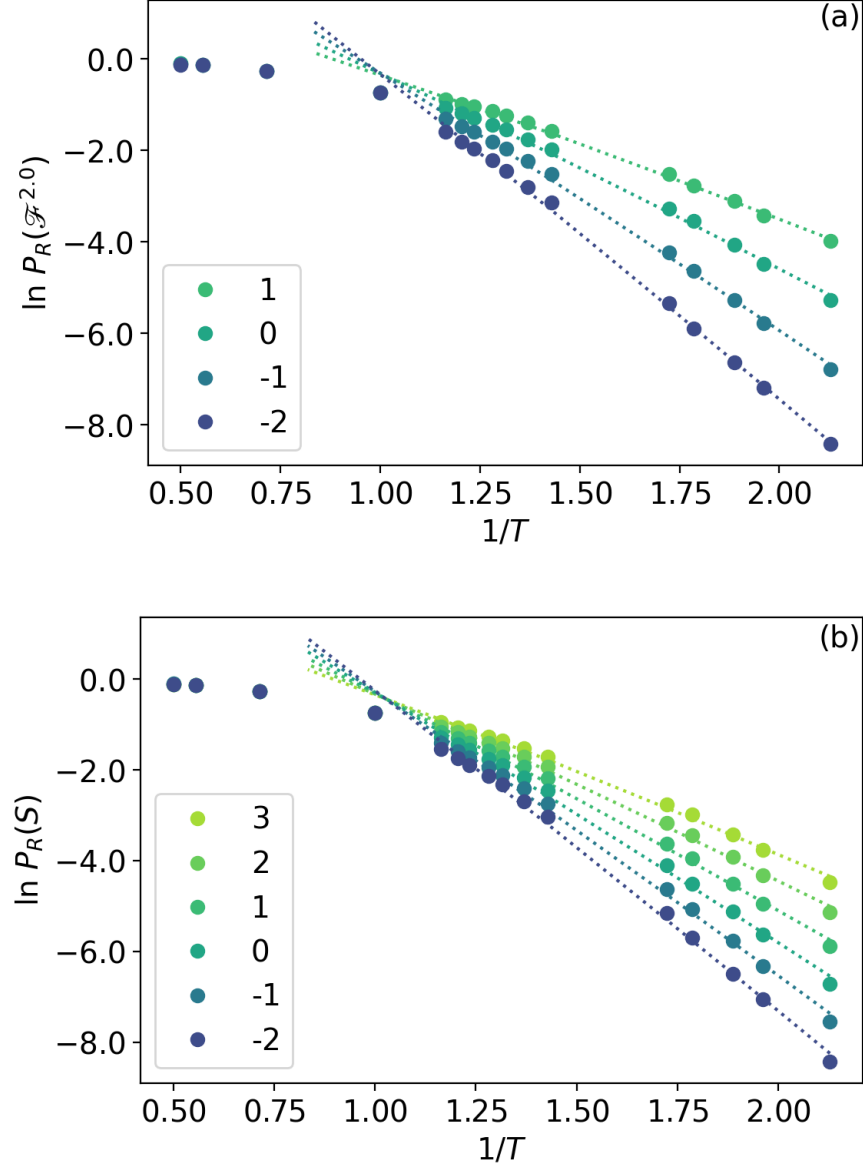


Figure 2.4: **The probability of rearrangement conditioned on fluidity reveals energy barriers below the onset temperature.** (a) The log probability of rearrangement conditioned on $\mathcal{F}^{2.0}$ vs inverse temperature. Point colors correspond to different bins of fluidity, as indicated in the legend. Part (b) shows the same features for rearrangements conditioned on S . In all cases, dotted lines are Kramers-form fits.

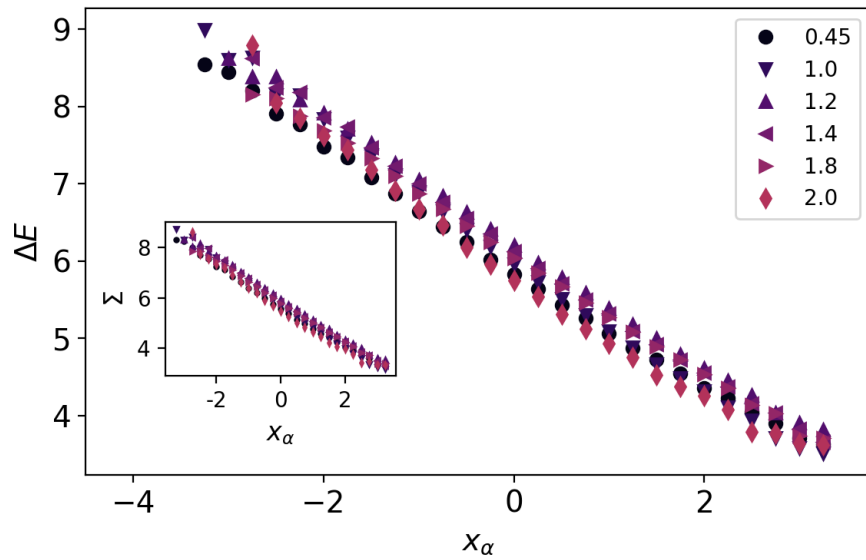


Figure 2.5: **Collapse of inferred landscape features from different training temperatures.** The energy barrier as a function of distance to optimized classifier, ΔE vs. x_α , as inferred from Kramers fits to $P_R(S)$ shows little variation across a wide range of classifier training temperatures. The inset showing the entropic contribution similarly collapses in this representation.

Chapter 3

Machine Learning of Energy Barriers to Rearrangements from Local Structure in Supercooled Liquids

In this chapter, we delve into the physical interpretability of regression models, by which we mean their ability to predict the energy barriers associated with particle rearrangements. Specifically, we aim to map local structural features to the probability of rearrangement, as determined through the iso-configurational ensemble. We depart from the discrete classification approach, which categorizes local structures into two distinct classes, and instead explore a continuous mapping of local structure to a continuous dynamical quantity. We directly compare inferred energy barriers from classification-based model to regression-based models. Our results demonstrate that these regression models can accurately predict energy barriers from structural information. First, we detail the methodology used to calculate the probability of rearrangement for a given local structure within the iso-configurational ensemble. Then,

we present our investigations into how various factors, including model inputs, the choice of regression algorithm, and the method of supervision, influence the physical interpretability of these models and their ability to accurately predict energy barriers. This work is based on [83].

3.1 Introduction

Understanding the physics of glassy systems remains a fascinating challenge. Schematically, they are often described as systems for which their structural properties — as indicated, e.g., by two-point density correlation functions — resemble those of ordinary liquids [67, 84]. And yet, when cooled rapidly into and beyond the supercooled regime they fall out of equilibrium and have a dramatic change in their dynamics [3]. Below a characteristic onset temperature dynamically heterogeneous dynamics sets in, with spatially correlated regions of more quickly and slowly relaxing regions appearing [17, 13, 16], and as the temperature decreases these domains increase in both timescale and spatial extent [85]. The link between structure and dynamics — and particularly the link between the local structures that characterize the arrangement of particles relative to a given particle and the activated “hopping” dynamics of that particle — remains a challenging and active area of research.

In this context, data-driven and machine learning (ML) approaches have been increasingly used to study the correlations between structure and dynamics. These approaches typically work with relatively general descriptors of the local structure, and in this way represent a method complementary to more traditional works that use physical intuition to attempt to write down good low dimensional descriptions of the local environment (such as local free volume [8] or locally favored structures [41]). Several notable studies have pursued this approach, leading to important insights into how structural features do (and do not) impact dynamical behaviors [48, 4, 86, 5, 47,

87, 55, 88, 49, 89, 90, 91]. A challenge inherent in these approaches is that in adopting general, often quite high-dimensional descriptions of local structure, a strong ability to correlate structure and dynamics often comes at the expense of having a physically interpretable order parameter associated with the local structure.

A notable and early exception was the claim that using linear support vector classification (SVC) models could ultimately lead to a prediction of energy barriers to particle-scale rearrangements based on local structure [4]. This SVC-based supervised learning approach introduced an order parameter, “Softness,” derived from the signed distance between a local structure in a high-dimensional feature space and the SVC classifying hyperplane. Softness was shown to map linearly to the energy barrier corresponding to rearrangements associated with that local structure. Despite the many studies that followed, a number of aspects of this result (and the methodology leading to it) remain unclear. For instance, to what extent is the eventual interpretability of the softness model actually physical [92]? Furthermore, to what extent was it informed by (a) the choice of local structural features, (b) the source of the training data, (c) the choice of dynamical label in performing supervised learning with that data, and (d) the choice of machine learning methodology? In our view these open questions both make it harder to develop theories based on softness as an order parameter, and make it harder to extend what one might call the softness methodology to other physical problems.

Some work has begun exploring these issues. Boattini *et al.* showed preliminary data suggesting that using a ridge regression model might lead to physical interpretability in terms of describing energy barriers to particle rearrangements [48]. By binning particles according to their propensity to rearrange, they demonstrated that the probability of particle rearrangement in the Kob-Andersen model shows an Arrhenius dependence on temperature. This Arrhenius temperature dependence is similarly observed when the probability of rearranging is conditioned on softness, but

it is not clear whether these two models (ridge regression and SVC, each using different descriptors of the local structures) infer *the same* energy barriers for similar local structures. Moreover, recent work by Swain *et al.* revealed that in the context of a toy model with known energy barriers, linear SVC models themselves typically do not capture the full underlying physics of the energy barriers they are trying to quantify [92]. Instead, they characteristically underestimate the variance in the distribution of energy barriers. This raises concerns about the accuracy of energy barriers inferred using the linear SVC approach in more complex settings such as the glassy systems currently under consideration.

In this work, we systematically explore the physical interpretability of multiple machine learning methodologies (varying the feature sets used, the ML algorithms employed, and the choice of labeling for the supervised learning) on multiple shared training sets. In order to have statistics for the true probability of particle rearrangement conditioned on different local structures, we make heavy use of the isoconfigurational ensemble [43], in which many random thermal trajectories are launched from the same initial configurational state. We identify a broad range of settings in which very different models and training choices all lead to similarly interpretable features, and that the energy barriers identified by these very different models are often strongly correlated with the predictions of the softness methodology.

Notably, using this isoconfigurational data we are able to compare classification-based and regression-based models directly. This shows that the common interpretation of softness as being trained on data at a single temperature and successfully generalizing to other temperatures neglects an implicit use of data outside the training set. At the same time, by using isoconfigurational runs we are able to systematically compare the extent to which these different approaches underestimate the variance in the true distribution of energy barriers. Using this as a metric for how successfully a methodology is able to capture the underlying physics, our results (a) emphasize

the importance of how model weights need to be varied with test temperature and (b) confirm that the details of local structural features chosen can dramatically affect a model’s ability to capture the relevant physics at the training temperature and generalize across temperatures.

The remainder of our work is structured as follows. In 3.2 we describe our model computational glassformer, our protocol for performing isoconfigurational simulations, the way we choose labels for rearranging and non-rearranging particles, and the set of structural features we consider in our work. In 3.3 we study the energy barriers learned by different ML algorithms on a common isoconfigurational training set, and emphasize differences between regression and classification approaches to inferring energy barriers. In 3.4 we show how the choice of structural features affects the energy barriers learned and the generalization of the physical interpretability associated with different models. In 3.5 we show how the choice of dynamical label used to define a rearranging particle influences these results. Finally, in 3.6 we discuss the implications of our work for the project of learning glassy physics by these means, and discussion potential extensions of this work.

3.2 Methods

3.2.1 Model and Simulations

In this work we focus on simulations of $N = 4096$ particles using the standard Kob-Andersen model [11]: a canonical fragile computational glassformer composed of an 80 : 20 binary mixture of particles interacting via a non-additive Lennard-Jones (LJ) potential truncated at 2.5 times the larger particle diameter. All simulations were done at a particle density of $\rho = 1.2$ in a cubic box with periodic boundary conditions. Throughout we report all quantities in reduced units using the standard LJ convention in which σ_{AA} , the diameter of the large particles, is the base unit of

distance; the interaction parameter ϵ_{AA} is the unit of energy; m , which is set to unity for both particle species, is the unit of mass; and $\tau = \sqrt{\frac{m\sigma_{AA}}{\epsilon_{AA}}}$, is the LJ unit of time. The Boltzmann constant k_B is set to unity and temperature is measured in units of ϵ_{AA}/k_B . We consider a temperatures in the range $T \in [0.45, 0.86]$, for which the system is supercooled and exhibits modest dynamical heterogeneity. The simulations were carried out in the NVT ensemble with a Nosé-Hoover thermostat [74], using the HOOMD-BLUE [93] package. In our analysis, as is standard, we will focus on the large particle species unless otherwise noted.

We make use of some of the configurational and trajectory data in Ref. [63, 94]. We additionally perform a large number of short-duration isoconfigurational simulations [44, 45] in order to quantify the probability of rearranging, p_r , for particles in snapshots drawn from a range of temperatures. Rather than performing first quenching a configuration and then performing simulations with independent realizations of particle velocities consistent with a target temperature, we instead perform isoconfigurational ensemble runs directly from the thermal configurations. That is, we evolve the dynamics of a given thermal configuration multiple times, assigning new velocities drawn from the Maxwell-Boltzmann distribution at the target temperature for each run. The time length of each isoconfigurational evolution was limited to 30τ , and at $t = 25\tau$ we recorded the fraction of runs for which a particle experienced an activated event (as quantified below). We have checked that when the time t is chosen to be a relatively short time scale our results for the probability of rearrangement, p_r , is not qualitatively affected.

In order to collect sufficient statistics, at low temperatures we performed between 10^4 isoconfigurational runs (at higher temperatures) and 10^6 isoconfigurational runs (at lower temperatures) for between three and five snapshots at each target temperature.

3.2.2 Identifying Rearrangements

As noted above, defining the isoconfigurational probability of rearranging on a per-particle level requires a metric to identify activated events from a particle's trajectory. We focus on the common choice of using a thresholded version of the p_{hop} indicator function introduced in Ref. [76]. The indicator function for particle i at time t is defined as

$$p_{hop}(i, t) = \sqrt{\langle (\vec{r}_i(t) - \langle \vec{r}_i \rangle_{\eta_2})^2 \rangle_{\eta_1} \langle (\vec{r}_i(t) - \langle \vec{r}_i \rangle_{\eta_1})^2 \rangle_{\eta_2}}.$$

Here $\vec{r}_i$ is the position of particle i , $\eta_1 = [t - 5, t]$, $\eta_2 = [t, t + 5]$, and $\langle \cdots \rangle_{\eta_i}$ is an average over the designated time interval. This corresponds to a total observation time window of 10τ , which is short compared to the alpha relaxation time in our cold samples but long compared to the duration of a particle rearrangement event.

Consistent with Ref. [4], we use a threshold to define activated events for particle i at time t , choosing $p_{hop}(i, t) > 0.2\sigma_{AA}^2$ as a threshold for a particle rearrangement. In Appendix A.0.2 we explore the effect of alternate dynamical labels for rearranging and non-rearranging particles.

3.2.3 Structural descriptors

Our supervised learning methods require a choice of the parameterization of the local structure around particles in our simulations. We predominantly use a combination of standard (Behler-Parrinello) features to characterize local two-point correlations [50], and bond-orientational order parameters to characterize the local angular environment [37]. Specifically, for particle i we quantify the radial structure via the functions

$$G_X(i; r, \delta) = \sum_{j \in X} \exp \left(\frac{-(r - R_{ij})^2}{2\delta^2} \right). \quad (3.1)$$

Here $X = A, B$ denote the component of the binary mixture being considered, the sum is over particles j of the given component near particle i , and R_{ij} is the dis-

tance between particles i and j . The parameter r is varied to describe local density correlations at different distances, and the parameter δ controls the width of the Gaussian shells. We characterize the local structure with the choice $\delta = 0.2\sigma_{AA}$ and with $0 \leq r < 5\sigma_{AA}$ chosen in increments of $0.1\sigma_{AA}$. In total, this gives 100 structural features describing the radial structure for a particle.

The angular structure function for the particle i is characterized by the functions $Q_l(i; r_{min}, r_{max})$, similar to work done in Ref. [4]. We build this up by first considering

$$\langle Q_{lm}(\vec{r}) \rangle = \frac{1}{N_j} \sum_j Y_{lm}(\vec{R}_{ij}), \quad r_{min} < R_{ij} < r_{max}, \quad (3.2)$$

where Y_{lm} is a standard spherical harmonic. Thus, $\langle Q_{lm}(\vec{r}) \rangle$ is the average of the spherical harmonics contributions for all particle j in a shell near particle i . We then define a rotationally invariant combination of these Q_{lm} as our angular descriptors:

$$Q_l(i; r_{min}, r_{max}) = \left(\frac{4\pi}{2l+1} \sum_{m=-l}^l |\langle Q_{lm}(\vec{r}) \rangle|^2 \right)^{\frac{1}{2}}. \quad (3.3)$$

The inner radius of the shell, r_{min} , is determined from a set. That is, from $r_{min} \in \{1.0\sigma_{AA}, 1.5\sigma_{AA}, 2.0\sigma_{AA}, 2.5\sigma_{AA}, 3.0\sigma_{AA}\}$ and $r_{max} = r_{min} + 0.5\sigma_{AA}$. The l parameter is chosen from the set $l \in \{2, 4, 6, 8, 10, 12, 14\}$. This leads to an additional 35 structural features describing the angular environment of particle i . Thus, the local environment of a particle i is described as a vector, $\vec{F}^i = \{F_1^i, F_2^i, \dots, F_M^i\}$ in a feature space of dimensionality $M = 135$. Unless otherwise noted, we standardize all features [75] so that at the training temperature, each has zero mean and unit variance.

3.2.4 Standard training datasets

For classification tasks (i.e., in computing Softness), we use a balanced training dataset described in Ref. [63]. This consists of the same particle configurations used in that work, with the analysis extended so that the feature vector contains both the radial features computed in the earlier work and the angular features described above. For regression tasks we create a new training dataset composed of the local structures, $\vec{F}$, and the associated probabilities of rearrangement, p_r , for all particles in each of the independent thermal configurations from which isoconfigurational simulations were run. For both regression and classification, we focus on models trained with datasets from the $T = 0.45$ temperature states.

3.3 Inferring energy barriers with different data-driven approaches

3.3.1 Correlating structure with dynamics using linear models

We begin by considering simple regression models that map local structure to the probability of rearrangement. We make the common choice of regressing on the log-odds (logit) of p_r , and focus on linear soft margin Support Vector Regression (SVR) and a Ridge Regression (RR) models. Concretely, for a particle i with probability of rearranging $p_r(i)$ and local structure $\vec{F}^i = \{F_1^i, F_2^i, \dots, F_M^i\}$ we try to learn the bias b and weight vector $\vec{w}$ that best predicts the log-odds:

$$\text{logit } p_r^i \equiv \ln \frac{p_r^i}{1 - p_r^i} = \sum_{\alpha=1}^M w_{\alpha} F_{\alpha}^i + b \quad (3.4)$$

The free parameters for the models (the weights and biases) were optimized using the SCIKIT-LEARN package [77].

The most straightforward measure of the ability of these models to generalize to unseen structure at the training temperature is by testing them on additional data taken at the training temperature. The correlation between the true and predicted log-odds is quantified using Pearson correlation,

$$\rho_{corr} = \frac{\text{cov}(y_{true}, y_{pred})}{\sqrt{\text{var}(y_{true})\text{var}(y_{pred})}}$$

where y_{true} is calculated from the isoconfigurational data and $y_{pred} = \text{logit } p_r^{(model)}$ is the prediction from the given model. That is, we calculate the ratio of the covariance ($\text{cov}(\dots)$) between the true and predicted values and the product of the variance ($\text{var}(\dots)$) of those quantities. We observed average correlations when testing on three new $T = 0.45$ configurations of $\rho_{corr}^{SVR} \approx 0.36$ and $\rho_{corr}^{RR} \approx 0.38$. This value is close to that value reported by [95] for the athermal system they studied.

This relatively low correlation, while suggesting that these models capture some relevant structural information, also indicates that predicting dynamics from local structures is inherently challenging. The limited correlation also implies that while we can proceed with exploring the models' physical interpretation in terms of energy barriers, we should treat these values with caution (as highlighted by recent work on simpler model systems [92]). We next apply the learned models to isoconfigurational data sets generated at higher temperatures. The distribution of predicted p_r for the SVR model is shown in Fig. 3.1 for temperatures ranging from $T = 0.45$ (supercooled) to $T = 0.86$ (comparable to the onset temperature). Comparing this plot to the literature, one notices that the trends in these distributions mirror those of softness quite closely [4]. The distribution of the $\ln$ of p_r is very nearly Gaussian, with a mean that shifts to the right with temperature at a rate comparable to the shift in the

distribution of softness (as shown in Appendix A.0.1).

To quantify more precisely the similarity between an SVR trained on isoconfigurational data and the softness SVM trained on hand-selected examples of “rearranging” and “non-rearranging” particles drawn from the thermal states of molecular-dynamics trajectories, we apply a trained Softness SVM to our isoconfigurational data and compare the distributions. We use the Jensen-Shannon Divergence (JSD), a symmetric version of the Kullback-Leibler Divergence (D): $\text{JSD}(P||Q) = (D(P||M) + D(Q||M)) / 2$, where $M = (P + Q)/2$ is a mixture distribution of P and Q and

$$D(P||Q) = \int_{-\infty}^{\infty} p(x) \log \left(\frac{p(x)}{q(x)} \right) dx.$$

The JSD is bounded between zero and one, and is a measure of the information lost when one distribution is used to approximate another distribution, where a value of zero implies a perfect similarity between the distributions. We find that the distribution of the log-odds of the regression models are extremely similar to the distribution of softness, with a JSD between the softness distribution and the distribution associated with either the SVR or RR models of the order $\mathcal{O}(10^{-2})$. Thus, the distribution of softness and those from the regression models are essentially the same.

3.3.2 Inferring energy barriers

The ability to infer energy barriers to particle rearrangement based on local structure is a key goal of many of these methods, and in the Softness picture this was how the model generalized across different temperatures: binning particles equally far from the classifying hyperplane at different temperatures revealed an Arrhenius form for the probability of rearranging at a given softness, suggesting that the local structure set an energy barrier scale that could be closely correlated with this classification protocol.

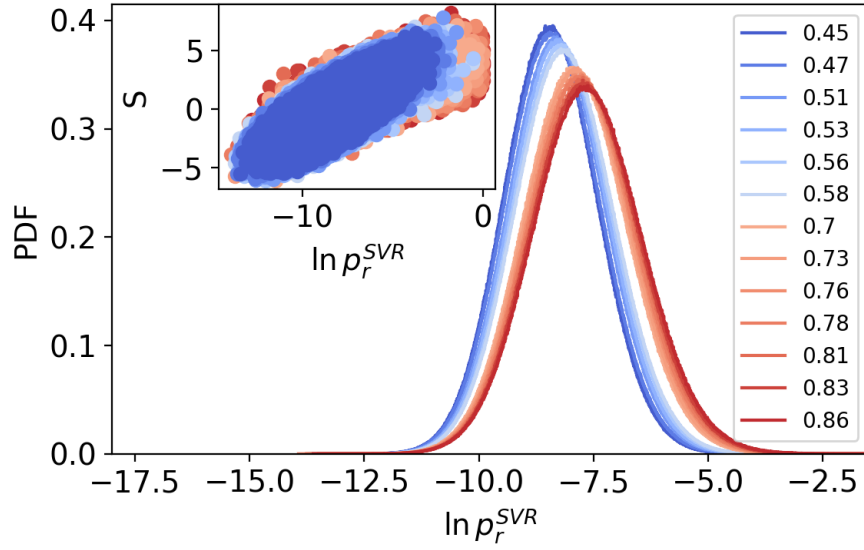


Figure 3.1: **Predictions from the SVR model strongly correlates with Softness.** The distribution of the log of the rearrangement probability predicted by the SVR model follows an approximately Gaussian shape. This distribution shifts progressively to the right as the temperature (legend) increases. The inset scatter plot further demonstrates a correlation between the SVR model predictions and Softness, suggesting that both quantities capture related structural information.

In the context of regression models, there are two different (and not obviously equivalent) ways to study how the predictions generalize across temperatures. The first, most direct, is to simply apply the said regression model to data at different temperatures: because the system is at a different temperature than the training temperature the local structures will on average be different, but it is certainly plausible that over the kind of modestly varying temperature range studied in supercooled KA simulations the distribution of observed structures at the training temperature will be sufficiently broad so as to allow for good generalization. The second is to repeat in the context of the regression models' predictions the Softness protocol for finding energy barriers. To pursue this second avenue, we bin particles by their value of $\ln p_r^{(model)}$ and evaluate the fraction of particles, f , in that bin that rearrange in a short future time window of 2τ . Just as was found in the Softness

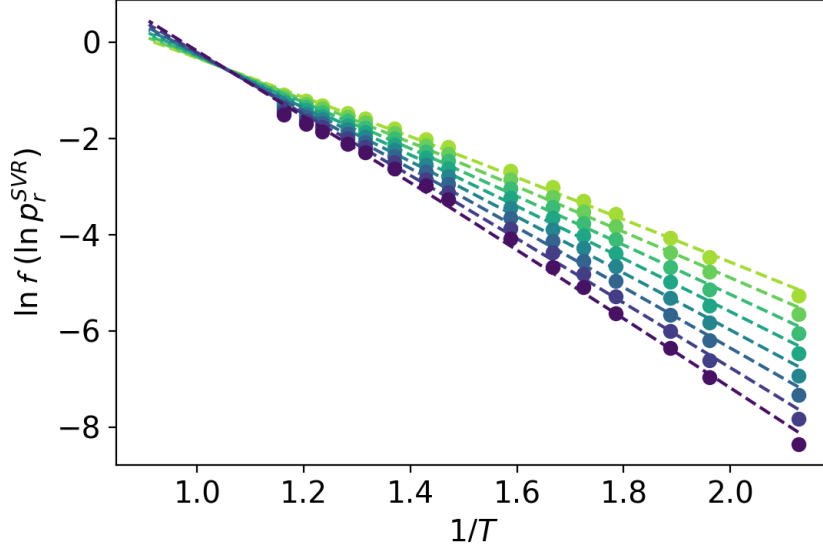


Figure 3.2: **SVR predictions have an Arrhenius form.** Binning particles by $\ln p_r^{(model)}$ and finding the fraction of the binned particles that rearranges at different temperature has an Arrhenius decomposition. The energy barrier, ΔE , and the entropic piece, Σ , are extracted by fitting to these Arrhenius form (dashed lines show a Kramer’s fit). The gradient from light green to dark purple indicates $\ln p_r^{SVR}$ values ranging from -6.5 to -10.

methodology, we find that $\ln f(\ln p_r^{(model)})$ vs. $1/T$ has the typical Arrhenius form: $\ln f(\ln p_r^{(model)}) = \Sigma(\ln p_r^{(model)}) - \Delta E(\ln p_r^{(model)})/T$. This is shown in figure 3.2 for the SVR model, $\ln f$ vs. $1/T$ for different values of $\ln p_r^{SVR}$.

From such Arrhenius forms, we examine the “physical interpretability” of the regression models in terms of the inferred energy barrier, ΔE , by considering the same set of particles (with the same sets of structural descriptors) and comparing the energy barriers that each model “learns”. To facilitate a direct comparison of energy barriers across models (which may exhibit different scales in their predictions), we normalize the output using a Z-score approach, similar to our earlier work [63]. We define the Z-score $x = \vec{w} \cdot \vec{F}/\sigma_{tr}$, where σ_{tr} is the standard deviation of the distribution of $\ln p_r^{(model)}$ for the respective regression model at the training temperature. In the case of softness, σ_{tr} is the standard deviation of the softness distribution at the training

temperature. The actual bias, standard deviations, and mean of the predicted $\ln p_r$ at the training temperature in Table 3.1.

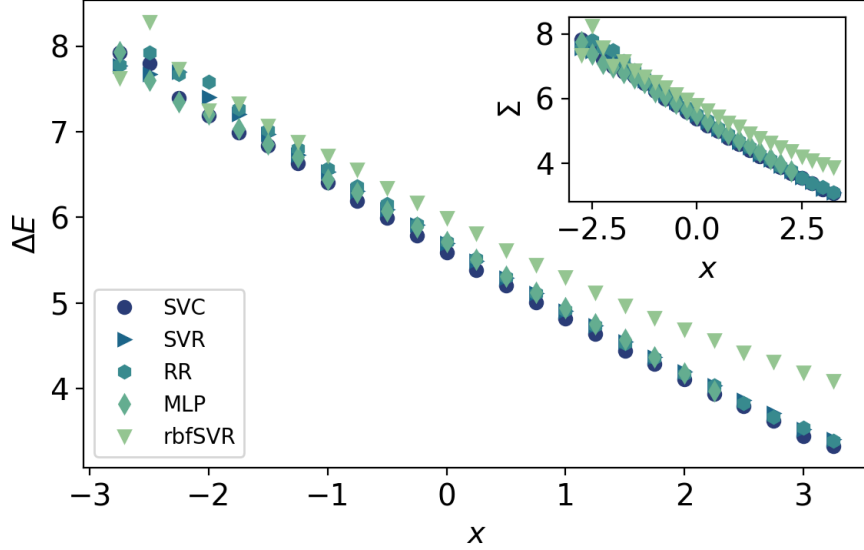


Figure 3.3: **Different models predict similar energy barriers to local rearrangements.** Energy scale and (inset) entropic terms characterizing the local energy barrier to particle rearrangement. A strong correlation between regression models and Softness is observed. Different symbols correspond to different regression and classification models, as noted in the plot legend — the chosen symbol shapes consistently correspond to the indicated model in all future figures.

Figure 3.3 shows that the energy barriers learned by these regression models are highly correlated with the energy barriers identified by softness. Given this high correlation with softness and the relatively low Pearson correlation coefficient that characterize the SVR and RR models on new data at their training temperature, we examine the directions of the SVR and RR regression hyperplanes relative to the Softness classification hyperplane. One would expect that regression and classification hyperplanes would typically be orthogonal, and that (e.g.) the SVR and RR hyperplanes would point in similar directions. As shown in 3.1, the regression hyperplane is *not* completely orthogonal to the classification hyperplane. Furthermore the projection of the SVR and RR hyperplanes is only $\hat{w}_{RR} \cdot \hat{w}_{SVR} \approx 0.4$, indicating

Model	$\hat{w}_S \cdot \hat{w}_{model}$	$\langle \ln p_r \rangle_{tr}$	$b_{\ln}$	σ_{tr}
SVR	0.13264	-8.41951	-8.35800	1.01994
RR	0.75145	-8.33924	-8.28735	0.86906
rbfSVR	–	-8.47193	-8.84931	0.86576
MLP	–	-8.29172	–	0.78949

Table 3.1: **Features of linear and non-linear models** For linear and non-linear models we show (where appropriate) the projection of hyperplanes onto the softness hyperplane, the average of the predicted $\ln p_r$ at the training temperature, the bias, and the standard deviation of the distribution of the predicted $\ln p_r$ at the training temperature.

that there is a substantial range of linear regression directions that are almost equally good at predicting dynamics.

Taken together, these results analyzing performance on and generalization across isoconfigurational datasets indicates that all of the models are capturing only a portion of the underlying physical observables. As suggested in Ref [92], this can be more fully buttressed by quantifying the fraction of the true variance of the distribution of rearranging predicted by each model. We implement the Law of Total Variance to quantify the fraction of the variance of the true distribution of logit p_r^{true} explained by our models. Based on the Law of Total Variance, the fraction of the true variance of Y explained by a random variable X , is the variance of the mean of Y conditioned on X and then normalized by the variance of Y , i.e. $\text{Var}(\langle Y|X \rangle)/\text{Var}(Y)$. Figure 3.4 shows the fraction of the true variance explained by models all trained at $T = 0.45$ when applied to iso-configurational snapshots across a range of temperatures. The fraction of the variance explained by the models is not significant across temperature; it is below a value of 0.5 for across temperatures. This shows that the models, whether classification-based or regression-based, are poorly explaining relevant details of the landscape from structure.

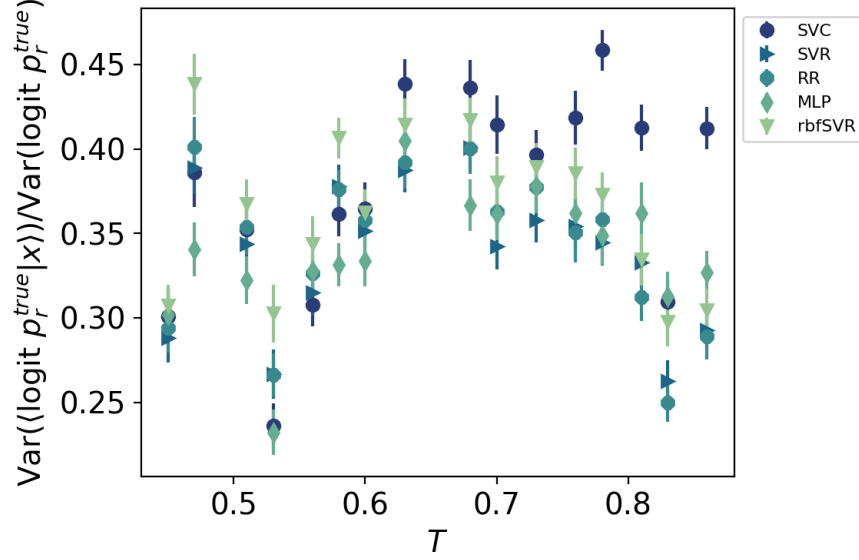


Figure 3.4: **Fraction of the true variance explained by the ML models when applied to iso-configurational snapshots across a range of temperatures.** The explained variance remains lesser than 0.5 across temperatures, indicating that both classification-based and regression-based models poorly capture relevant structural features of the energy landscape. The error bars are standard error on the mean from 40 bins each containing 177 uncorrelated particles.

3.3.3 Correlating structure with dynamics using non-linear models

The original Softness methodology used a support vector classifier with a linear kernel — while it was not a priori clear that the dynamics would be linearly separable in the given feature space, this choice made more straightforward the assignment of a scalar value as the signed distance of a point in feature space to the classification hyperplane. We have seen, though, that using linear methods results in a wide range of regression or classification [63] hyperplanes with similar predictive performance. We thus investigate whether non-linear methods provide better predictive power and / or physical interpretability.

To do so, we focus on two quite different methods: a SVR model using a radial basis function (rbf) kernel for the nonlinear transformation (which we label rbfSVR),

and a multilayer perceptron (MLP) with rectified linear unit activation functions. For the MLP we use architectures that have five hidden layers and a total number of parameters equal to half, the same, or double the number of training examples (results when the number of parameters equals the number of training examples are reported here, and the other results are shown in Appendix A.0.3). For all of these nonlinear models, we use the same structural descriptors and the same training sets as described in Section 3.3.1.

Perhaps surprisingly, we do not find that the predictive capacity of these these nonlinear models is notably different from the linear models reported above. The MLP model learns roughly the same physical quantities as the RR and SVR models (Fig. 3.3), and if anything the rbfSVR model is slightly less sensitive to the details of the energy landscape given the same amount of training data. This may be an indication that the amount of training data we use is sufficient for training linear but not non-linear models.

3.4 Impact of structural descriptors on model performance

When considering the results in Fig. 3.4 (or in the other figures above representing the generalization of these machine learning models), it is unclear how important the specific choice of structural features was. For instance, Boattini *et al.* pointed out that adding features that capture information about the average structure of neighboring particles to a particle’s feature vector improved the performance of ridge regression models [48]. While we have chosen relatively common parameterizations of the two-point and many-body local structural environments, there are many alternate parameterizations that incorporate more or less physical intuition about what is important [41]; it is also clear that we are using a feature set which is far from being an

orthogonal basis set. Thus, in this section we investigate the fraction of true variance learned by a regression model as we vary the number and type of descriptors that define the feature space. To quantify the fraction of the variance learned, we normalize the variance of the distribution of logit p_r^{SVR} at $T = 0.45$ by the variance of the true distribution at that temperature. That is, $\eta = \text{Var}(\text{logit } p_r^{\text{SVR}}) / \text{Var}(\text{logit } p_r^{\text{true}})$. Here we keep the model constant, focusing on the linear SVR model.

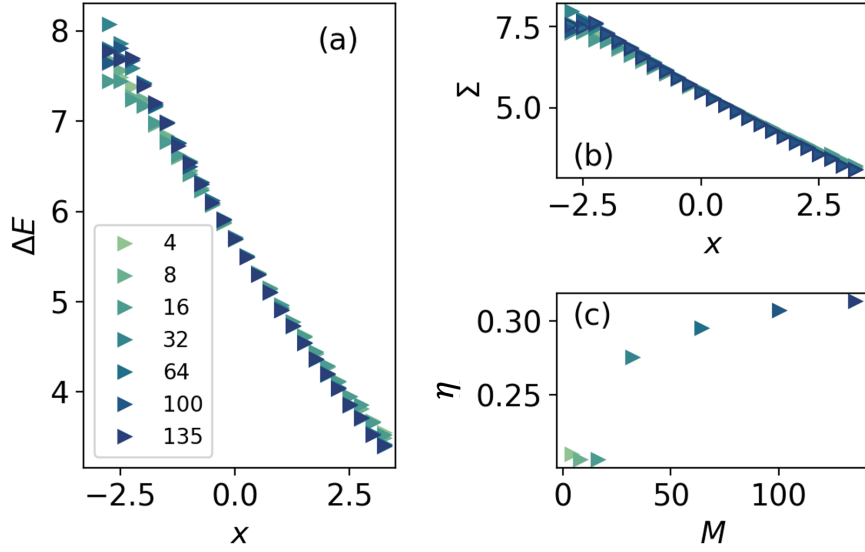


Figure 3.5: **SVR model interpretation as a function of the number of features considered.** Through RFE, the (a) energy barrier height and (b) entropic contribution to the probability of rearranging as a function of the dimension of feature space barely changes. The fraction of the true variance learned by the model (c) increases modestly over this range.

We first vary the dimensionality of the feature space by varying the number of structural descriptors considered. We use the Recursive Feature Elimination (RFE) technique [96], which was also used in Ref. [92] to find the optimal dimensionality of feature space that maximizes the variance of inferred energy barriers in a simple model. We show in Fig. 3.5 how the inferred energy barriers of our particle systems vary with the dimension, M , of the feature space. The trend of $\Delta E(x)$ and $\Sigma(x)$ for $M \leq 16$ is shown in Fig. 3.5(a) and (b), respectively. For these small numbers of

features there are relatively modest changes in the inferred physical quantities, and the fraction of the true variance explained by the model is relatively small (Fig. 3.5(c)). As M increases further, η increases significantly and the inferred energy barriers vary modestly more rapidly as the Z-score-like x is varied. This improvement in the model quickly saturates, though, further emphasizing that the models are in fact inferring a bounded representation of the true distribution of energy barriers.

We next examine the impact of the specific local descriptors used on the regression model’s ability to infer energy barriers. We use the higher-order structural descriptors, $X_i^{(n)}$, proposed by [48]:

$$X_i^{(n)} = \frac{1}{C} \sum_{j:r_{ij} < r_c} e^{-r_{ij}/r_c} X_j^{(n-1)}. \quad (3.5)$$

Here $C = \sum_{j:r_{ij} < r_c} e^{-r_{ij}/r_c}$ and r_c is a cutoff radius whose value is chosen to be the location of the second minimum in the radial distribution function. The zeroth-order descriptors, $X^{(0)}$ are the combination of radial and angular descriptors used in section 3.3. To this we add first- and second-order descriptors (the combination of which were shown to be competitive with more complex graph neural network models [48]), increasing the potential size of the feature space from $M = 135$ to $M = 405$.

As shown in Fig. 3.6(a), including these higher-order features leads to inferred energy barriers that vary less strongly with x than neglecting them. This result is counter-intuitive, but may stem from the more strongly changing typical higher-order features with temperature than is seen in the zeroth-order features. It may also be an indication that we are in a data regime that is too sparse for a three-fold increase in the dimensionality of the feature space, as we have not systematically checked these results as a function of performing more independent isoconfigurational simulations. We do see in Figure 3.6(b) that including up to second-order features does not improve the fraction of variance explained by the linear SVR model.

3.5 Impact of dynamical labels on model performance

The above results involved varying the machine learning methodology used and the set of features used to characterize the local structural environment of a particle. Held fixed was the definition of a “rearrangement.” Previous work showed that continuing to use a thresholded version of p_{hop} to define rearrangements concluded that the dependence of the inferred energy barriers on the p_{hop} threshold amounted to a simple shift in the energy scales but with identical slopes [4]. This was rationalized as indicating that characteristic rearrangement size depended linearly on the cutoff chosen. Other work has indicated that results quite similar to those of the original softness protocol can be obtained not by using p_{hop} as an indicator function, but using thresholds on the total magnitude of particle displacements [63] or on other measures of structural rearrangements [54].

Extending those observations about the relative robustness of inferred energy barriers to the precise choice of dynamical labeling methodology, here we first investigate the impact of defining the probability of rearrangements using not p_{hop} but the cumulative squared displacement (CSD). The CSD captures the cumulative magnitude of particle displacements over a time window, with details in Appendix A.0.2. As indicated in Appendix A.0.2, when using different indicator functions for defining a particle rearrangement, we choose different thresholds so that we approximately match some features of the distribution of p_r in our isoconfigurational simulations. Figure 3.7 shows that, indeed, the details of the inferred energy barriers (and entropic contributions) are extremely highly correlated with those inferred based on dynamical labels derived from p_{hop} .

Another common choice of dynamical label is not to define and then predict “rearrangements” but rather to try to predict *propensity* [46, 44], typically defined

as the isoconfigurational average of the norm of a particle’s displacement after some time window. It is, a priori, unclear how important the use of an inherent state for the isoconfigurational averages are in this protocol, and we thus investigate the Pearson correlation of a linear SVR model regressing on propensity determined from isoconfigurational starting from thermalized configurations. Can such a model predict inherent state propensities? To answer that question, we train a linear SVR model on $t = 25\tau$ from thermal configurations at $T = 0.56$, and compare with the $T = 0.56$ IS propensity data in Ref. [5]. Figure 3.8 shows the Pearson correlation of the predictions from an SVR model trained directly on IS propensities (using the datasets of Ref. [5]) and those of an analogous model trained on our thermal configurations. Although some information has clearly been lost in using thermal configurations, the difference is quantitative rather than qualitative.

3.6 Conclusion

In this chapter we have systematically investigated the degree to which different machine learning methodologies are able to “learn” physically interpretable connections between structure and dynamics in a prototypical glassforming fluid. By using large-scale data from isoconfigurational simulations we are able to compare classification with regression techniques, and have focused on the influence of multiple “researcher degrees of freedom” related to the implementation of the original softness approach. This includes the choice of dynamical label for rearrangement events, the need for inherent state vs thermal snapshots, the feature space used, the choice of classification vs regression, and the importance of linear vs non-linear data-driven models.

Several of our results are consistent with existing studies. Aligned with the arguments and observations of Refs. [92, 97], we find that the dimensionality of the feature space primarily impacts the inferred energy barriers when the variance of the

distribution of distances increases significantly. Many of the other precise details of the feature space — the presence or absence of angular features [4] or the use of higher-order features [48], for instance — has only modest quantitative effects on the inferred local energy landscape. Similarly, the precise choice of dynamical label used to identify rearrangements seems to make only a small quantitative difference as long as roughly equally sparse “rearrangement events” are considered in the training temperature [54, 4].

Most notably, by assessing the true probability of rearrangement through the iso-configurational ensemble p_r , we identify that the machine-learning models do learn details of the energy landscape, however, in a limited way. The machine-learned variables poorly explain significant portion of the distribution of this true probabilities. These results highlight that the complexity in correlating structure with dynamics in supercooled liquids.

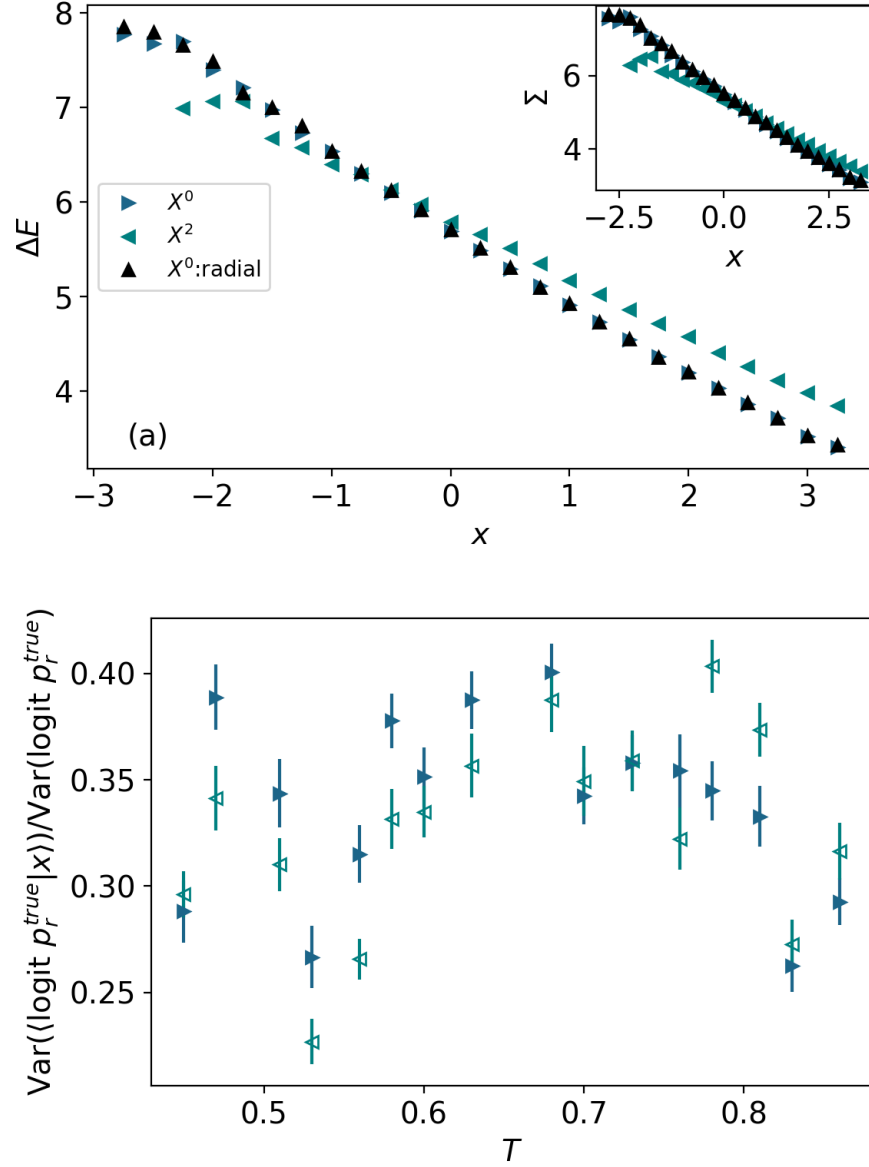


Figure 3.6: **Model predictions as a function of varying the feature space.** (a) The inferred energy barriers and entropic contributions of an SVR model using higher-order features are strongly correlated with those using the standard (Behler-Parrinello) features. (b) Using higher order features does not lead to better estimate of the variance of logit p_r^{true} . The chosen symbol shapes in (b) consistently correspond to the indicated features in (a). Error bars are standard error on the mean from 40 different bins each containing 177 uncorrelated particles.

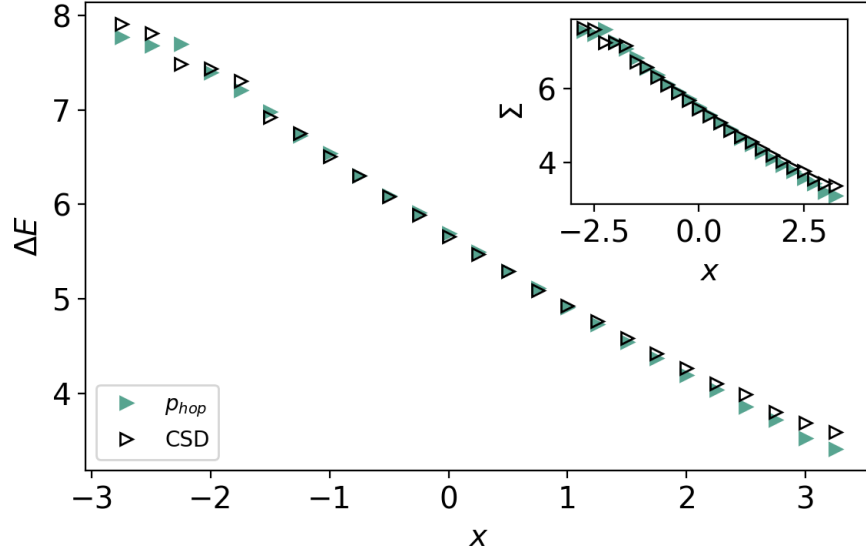


Figure 3.7: **Inferred local landscape features for different choices of dynamical label.** The inferred energy barriers (main plot) and entropic contributions (inset) obtained from an SVR classifier are nearly indistinguishable when training is done with different dynamical labels.

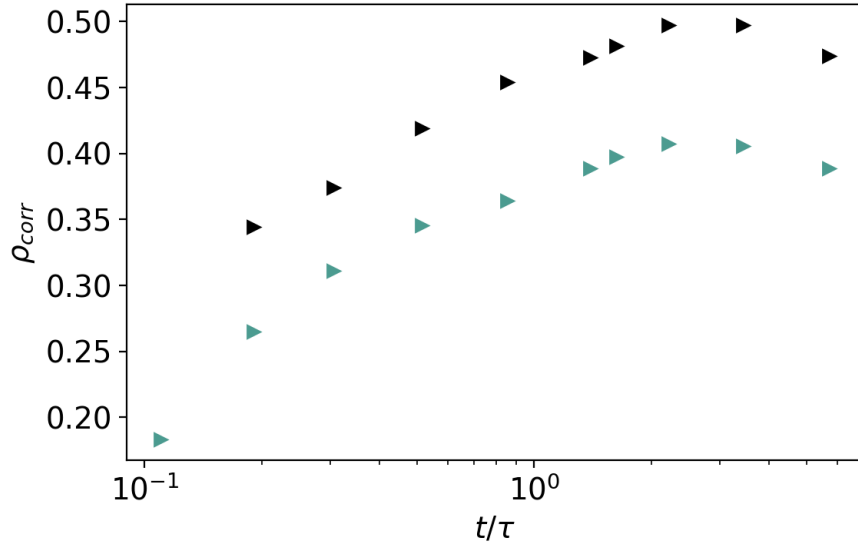


Figure 3.8: **Pearson correlation between true and predicted inherent state propensity.** The Pearson correlation between true and predicted inherent state propensities for an SVR model trained on our thermal configurations (light cyan triangles) and an analogous model trained directly on inherent state data from Ref. [5] (black triangles).

Chapter 4

Ongoing Work on Memory Effect in Supercooled Liquid

This chapter is based on ongoing work. The motivation behind this work is centered on some of our observations of the iso-configurational data attained for analyses on the physical interpretability of machine learning models in Chapter 3. I briefly describe the current paradigm of memory effects in the context of supercooled liquids and glasses, some of our observations of memory in the supercooled phase of Kob-Andersen glass model, and possible directions for the research on correlating structure with dynamics in supercooled liquids.

4.1 Introduction

In Chapter 3, we implemented at least 10000 iso-configurational runs per snapshot from different temperatures and identified the number of times each particle surpassed a rearrangement threshold. The probability that a particle rearranges, p_r , is the number of times such events occur normalized by the total number of iso-configurational runs; this probability was mapped to structure in the last chapter. If we, however, look at the evolution of the average probability for different snapshots, we find that

the average probability evolve towards an asymptotic value. We show the case for three thermal snapshots at $T = 0.86$ in Fig. 4.1. The path through which each snapshot takes is distinct and implies a dependence on initial state despite being at a temperature for which ergodicity holds for the Kob-Andersen glass model.

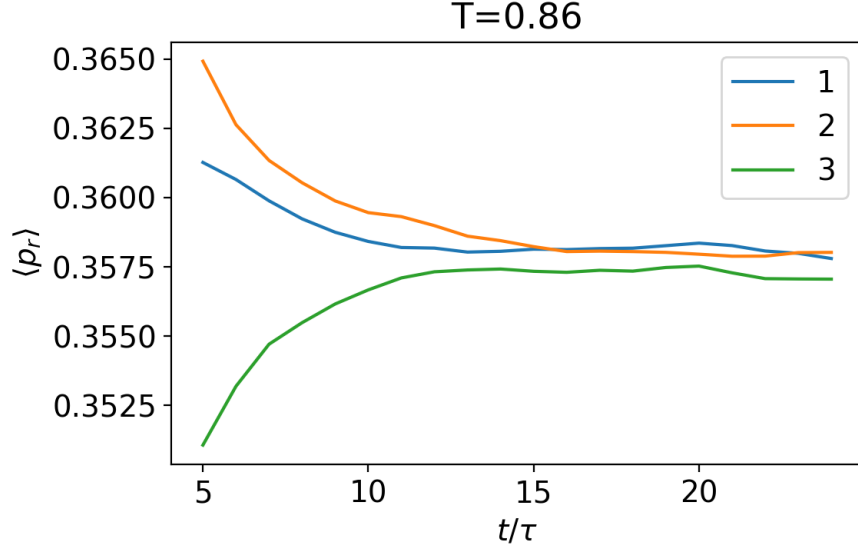


Figure 4.1: **The evolution of the average probability of rearrangement for different snapshots approach a common value at long timescale.** The probability of rearrangement averaged for the big particles for three different thermal snapshots (Legend) evolve in such a way that the path they take depends on the snapshot.

Furthermore, let us consider a collection, H_S , of particle with softness S at time $t = 0$ from a configuration of particles. The average softness for the grouped particles, $\langle H_S(t) \rangle$ evolves toward the system average, as shown in Fig. 4.2 for temperature $T = 0.83$. It is not surprising that the average softness of each bin evolves toward the long-time average; however the timescale at which this happens is seen to be of the order of $10^2 \tau_\alpha$. τ_α at this temperature is of the order of 1τ . In addition, the path each group of particles takes is unique, implying that the state at $t = 0$ influences future states. Thus, the “memory” of the past impacts the present dynamics.

It may be observed that the average softness for each bin exhibits a rapid decay to

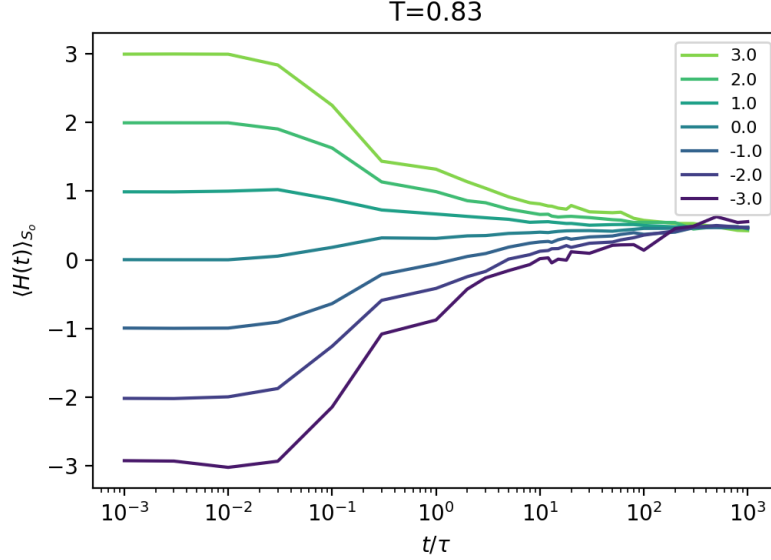


Figure 4.2: **The evolution of the average softness of particles grouped by their softness at time $t = 0$.** The average softness of particles grouped based on their initial softness (Legend) evolve towards the system average. The paths each grouped particles follow towards the long-time depends on their initial value.

less than 60% of its initial value at short timescales. To ascertain whether this decay timescale corresponds to the β -relaxation timescale (τ_β), we perform a normalization procedure. Specifically, we subtract the system-average softness, $\langle S \rangle$, from $\langle H_S(t) \rangle$ and normalize this resulting time-dependent difference by $\langle H_S(t=0) \rangle - \langle S \rangle$. This normalized values are then compared with the self-intermediate scattering function, defined as: $F_s(k, t) = \frac{1}{N} \sum_j^N e^{-i\vec{k} \cdot \Delta \vec{r}_j(t)}$, where $\Delta \vec{r}_j(t)$ represents the displacement of particle j at time t , and $\vec{k}$ is the wavevector corresponding to the first peak of the static structure factor. This comparison allowed us to assess the relationship between the decay of local structural order, as measured by softness, and the β -relaxation dynamics captured by the self-intermediate scattering function. We do this comparison for a temperature of $T = 0.56$.

Figure 4.3 reveals that the timescale associated with the initial decay of structural order, as characterized by softness, is indeed on the order of τ_β . Furthermore, a discernible trend emerges in the evolution of the normalized average softness for the

different particle groups: the soft particle group exhibits a relatively faster decay compared to the hard particle group. However, due to the presence of statistical noise in the data, additional simulations are necessary to improve the averaging and obtain a more precise quantitative measure of this timescale.

Additionally, Figure 4.3 show that the normalized softness for the various grouped particles exhibit slow-relaxation processes that is linear with the logarithm of time. This is reminiscent of the aging and memory in glasses. This linear dependence is unexpected and warrants further analysis on memory effects in supercooled liquids.

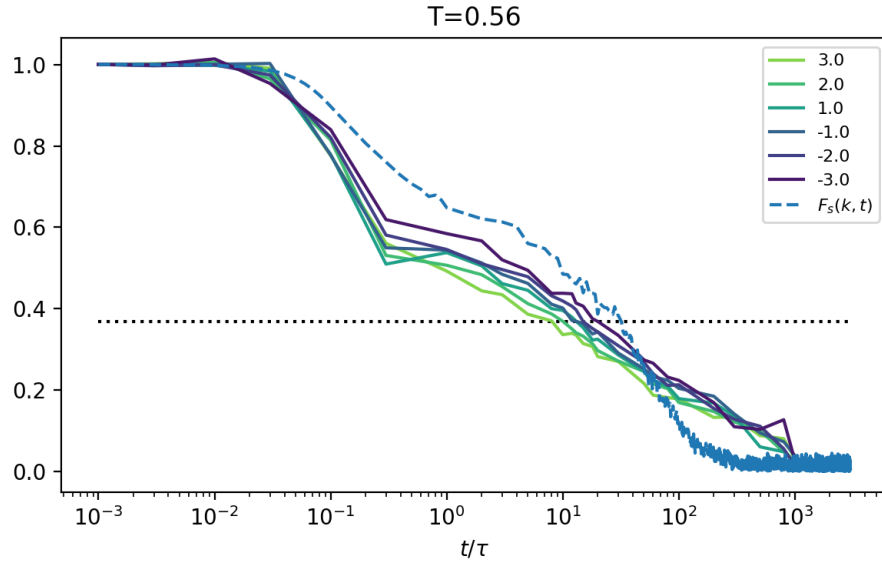


Figure 4.3: **The normalized average softness for different particle groups (binned by initial softness) is compared with the self-intermediate scattering function ($F_s(k, t)$).** The decay of the normalized average for various initial softness (thick lines) at short timescales, on the order of the β -relaxation time (τ_β), suggests a connection between local structural rearrangements and the fast-relaxation process. The dash line represents the $F_s(k, t)$ and the numbers on the legend represents initial softness.

Memory effects manifest in different ways across various forms of matter [98]. In disordered systems, this is the ability to retain information about past states, histories, or external perturbations after the influences are removed; it has been advantageous, technologically, leading to devices for information storage and processing,

energy applications, adaptive and smart applications, to mention but a few. Experiments and simulation studies on spin glasses [99, 100, 101, 102, 103], molecular glasses [104, 105], and polymers [106, 107] have shown that memory effects are present in these systems. By memory effects, we mean by inducing some form of temperature cycle, these studies have shown that these systems remember “something” about their thermal histories [108].

In glasses, memory is associated with aging and rejuvenation. As the glass ages, it sinks into deeper minima in the energy landscape, effectively storing the system’s thermal history. Thermal excitations or external perturbations can cause the system to rejuvenate essentially escaping a minima which effectively leads to a loss (of some) of the thermal histories. However, no case of memory effects in supercooled liquids has been reported to the best of our knowledge. Because supercooled liquids do not age, memory effects are not expected in supercooled liquids.

In this chapter, we show speculatively that supercooled liquids exhibit memory effects by grouping particles by their local structure and watching how the average local structure evolve during a temperature cycle. By observing memory and rejuvenation in supercooled liquids, we bridge the gap between memory effects in supercooled liquids and in glasses essentially improving our understanding of memory effects in thermal disordered systems. Furthermore, by examining memory effects through structure, we show the nuances of structuro-dynamical correlations in these systems.

4.2 Methods

4.2.1 Model and Simulations

To investigate the dynamic properties of our model system, we performed molecular dynamics simulations using the Kob-Andersen binary mixture [11]. Our simulations

comprised $N=4096$ particles, with an 80:20 ratio of A-type to B-type particles, interacting via the standard Lennard-Jones potential truncated at a cutoff radius of $2.5\sigma_{AA}$. The simulations were carried out at a constant particle density of $\rho = 1.2$ within a cubic simulation box, employing periodic boundary conditions to eliminate surface effects. We utilized reduced Lennard-Jones (LJ) units throughout our simulations: distances are measured in units of σ_{AA} , energies in units of ϵ_{AA} , and masses was set to unity for both particle types. Time is expressed in units of $\tau = \sqrt{\frac{m\sigma_{AA}^3}{\epsilon_{AA}}}$. The Boltzmann constant k_B was set to unity, and temperature is reported in units of ϵ_{AA} . The simulations were conducted in the canonical ensemble (NVT) using the Nosé-Hoover thermostat [74] to maintain a constant temperature. We explored the temperature regime for which the system is supercooled and far above the mode-coupling temperature for our simulation. The simulations were performed using the HOOMD-BLUE [93] package.

4.2.2 Temperature Cycle

Following the established simulation protocol, we initiated a temperature cycle by evolving a thermal snapshot at $T = 0.51$ for an equilibration period of 30τ . Subsequently, the temperature was linearly ramped up to $T = 0.56$ over a duration of 10τ . The system was then held at this elevated temperature for another 30τ to allow for equilibration at the new temperature. Finally, the temperature was linearly ramped back down to $T = 0.51$ over another 10τ period, and the simulation was continued at this initial temperature for further analysis. We chose a temperature of $T = 0.51$ because τ_α for this temperature is of the order of 10τ which is comparable to the duration of our heating cycle.

To further investigate the influence of thermal history on the system's dynamics, we plan to perform the aforementioned temperature cycle on a set of iso-configurational snapshots. This will allow us to probe memory effects and the role of initial configura-

tions in the exploration of the energy landscape. Due to the computational demands of such an analysis, we will focus on examining memory effects for particles within a specific configuration, grouped according to their local structural order. This targeted approach will enable us to efficiently investigate the interplay between local structure and the system’s memory of past thermal cycles.

4.3 Results

To investigate the impact of the temperature cycle on local structural order, we analyzed the temporal evolution of particle softness. Particles were binned according to their initial softness, and the average softness within each bin was tracked throughout the simulation. As depicted in Fig. 4.4, we observed a change in the average softness of the binned particles during the temperature cycle, with initially soft and hard particles all becoming softer. However, it is currently unclear whether this behavior is a direct consequence of the temperature cycle or merely an artifact of statistical noise.

To address this ambiguity, we are conducting a comparative analysis by subtracting the average softness observed in a control simulation from that of the temperature-cycled system. We hypothesize that a significant deviation from zero in the softness difference during the temperature cycle, followed by a return to zero, would indicate a memory effect. This analysis is currently underway and will provide further insight into the relationship between thermal history and local structural rearrangement.

4.4 Discussion

Should further analysis, as detailed in Sec. 4.3, confirm the presence of memory effects in our supercooled liquid system, this would establish that such effects are not exclusive to the glassy state. Previous work has demonstrated that even a simple sorting algorithm, designed to mimic thermally activated processes, can exhibit memory

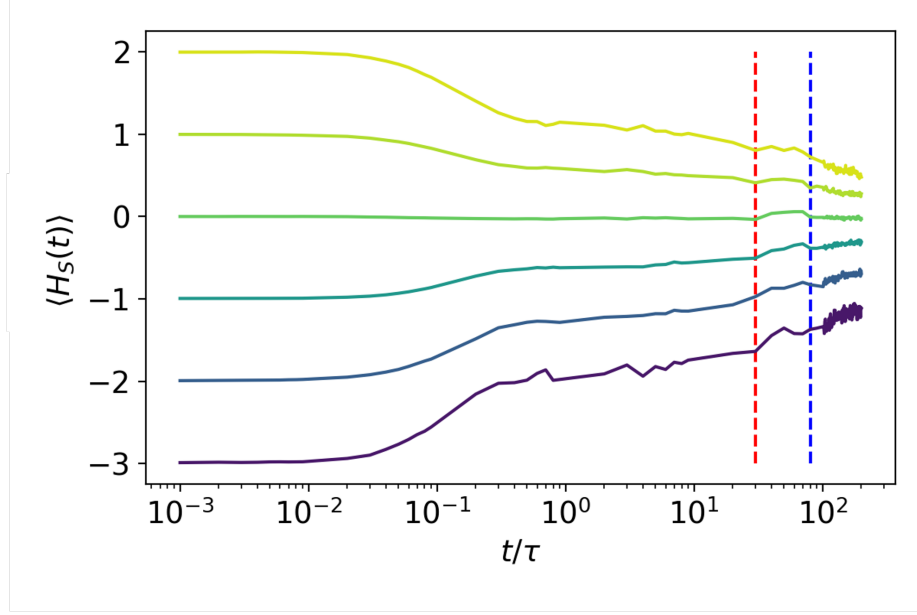


Figure 4.4: **The evolution of average softness for particles grouped by their initial softness.** The red vertical line marks the time at which the temperature began to increase. The blue vertical line marks the time at which temperature is returned back to the initial value. There seem to be a consistent bump in the average structure for each bins during the temperature cycle. At the moment, it is too soon to tell. The bumps seem to show that both soft and hard particles become softer during the temperature cycle. The path of the particles return back to their initial state when the temperature is reverted back to the initial value.

effects when subjected to a thermal cycling protocol [109]. This algorithm, which displays hallmarks of glassy behavior such as aging and rejuvenation, requires the incorporation of a Boltzmann factor—containing an effective temperature and an energy term defined by the difference between adjacent numbers—to weight the cost of swapping nearest-neighbor pairs. Without this thermal activation component, no glassy behavior or memory effects are observed.

The potential observation of memory effects in supercooled liquids could imply that the influence of the energy landscape on the system’s dynamics, which becomes increasingly pronounced with decreasing temperature, can be manifested through these memory effects. Specifically, by grouping particles based on their initial softness, we effectively categorize them according to their initial energy barriers for activated

rearrangements. As the distribution of softness within each group broadens over time, so too do the corresponding energy barriers. This behavior may reflect the exploration of different minima in the energy landscape, akin to aging processes in glassy systems, although the explored minima are not necessarily deeper in our case. However, these remain speculative interpretations that require further investigation and rigorous analysis.

Notwithstanding the need for further analysis, Fig. 4.3 suggests a potential connection between structural relaxation and the β -relaxation timescale (τ_β) in the Kob-Andersen model. It is widely accepted that τ_β represents the timescale for which a particle remains trapped within its cage, undergoing localized vibrational motion [67, 110]. In contrast, the Johari-Goldstein β -relaxation involves cooperative intermolecular rearrangements and is associated with processes occurring on a faster timescale [111, 112, 113, 114, 115, 116]. The observed decay of the normalized average softness at the τ_β timescale appears to support the presence of Johari-Goldstein β -relaxation in this system, suggesting an interplay between local structural rearrangements and the faster β -relaxation processes. Additionally, [117] reported that the fast-relaxation due to fast-moving particles are responsible for the fast-relaxation timescale. Taken together with our preliminary analysis shows the complex relaxation processes that take place in supercooled liquids and glasses.

4.5 Future Directions

An observation of memory effects in supercooled liquids, if such effect is confirmed, could open up several promising avenues for future research. One intriguing direction could be to systematically investigate the dependence of these memory effects on various factors, such as the cooling and heating rates during the temperature cycle, duration of the cooling or heating period, and the composition of the liquid to mention

but a few. This will help to explain the underlying mechanisms responsible for the observed memory effects and their relationship to the dynamics of the supercooled state. Another promising avenue could be to explore the connection between memory effects and other dynamic phenomena in supercooled liquids, such as dynamic heterogeneity and spatially correlated particle motion. This will help us understand how memory effects are manifested in the spatial and temporal organization of the liquid, eventually providing valuable insights into the nature of the glass transition.

Chapter 5

Conclusion

This dissertation has undertaken a comprehensive investigation into the application of machine learning models for predicting and interpreting the structural signatures of dynamics in supercooled liquids. By scrutinizing the relationship between local structure and energy barriers for particle rearrangements, we have significantly advanced our understanding of the physical implications associated with machine-learned order parameters, such as “softness.” The principal findings of this research are summarized below:

1. **Structure-Dynamics Correlations Persist Above the Onset Temperature:** We have provided evidence demonstrating that structural features originating from the high-temperature, diffusive regime possess the capability to predict energy barriers for activated events within the supercooled regime. This observation suggests the existence of shared underlying structural characteristics between diffusive and activated events, implying a continuity of structural influence across temperature regimes.
2. **Various Machine Learning Models Identify Correlated Energy barrier from Structure in Glassy Systems:** Through a comparative analysis of Support Vector Classification (SVC) with regression-based models, including

Ridge Regression, Support Vector Regression, and Multilayer Perceptron, we have shown that these diverse machine learning approaches consistently extract analogous structural signatures relevant to glassy dynamics. This finding reinforces the hypothesis that predictive models are effectively learning complex, high-dimensional structure-dynamics relationships, rather than merely performing arbitrary classification or regression tasks.

3. **Evidence Suggesting Memory Effects in Supercooled Liquids:** Our preliminary investigations have yielded evidence indicating that supercooled liquids retain a memory of past rearrangement events, as evidenced by their response to thermal cycling protocols. This suggests that memory effects, traditionally attributed to the glassy state, may also play a crucial role in the dynamics of supercooled liquids. If substantiated, this discovery could significantly impact our understanding of the glass transition and the emergence of long-lived structural correlations in supercooled systems.

Despite these advances several open questions persist. In Section 5.1 some of these open questions and promising avenues for future research are outlined. A concluding remark is provided in Section 5.2.

5.1 Open Questions and Future Directions

Notwithstanding these advancements, a number of fundamental inquiries persist without definitive resolution. A few of these are:

1. In Chapter 3, we compared various machine learning models and showed that they infer similar physical quantity from structure. However, the work can be expanded to both classification and regression techniques, to a wider range of models, such as deep learning architectures, graph neural networks, and kernel methods. This will provide a more comprehensive understanding of the

strengths and weaknesses of different approaches for correlating structure with dynamics in glassy systems which will uncover new physical insights. Furthermore, while we showed that the precise details of the feature space have a modest impact, there is still room for optimization. Research on more alternative feature engineering approaches, including those based on physical insights or advanced feature selection techniques, is needed to potentially improve model performance and interpretability.

Additionally, the number of iso-configurational runs especially at low temperatures and the number of snapshots used as seeds for the iso-configurational runs were limited. Also, it is not necessarily the case, especially at low temperatures, that the probability of rearrangements we used for training had reached a steady state value. More research is needed on how these factors impact the results in the chapter especially in the aspects of the regression models to accurately explain the variance distribution of rearrangement probabilities across temperatures.

2. In Chapter 2 we showed that the tails of the distribution of diffusive dynamics are informative of activated dynamics in the supercooled regime, it is not clear if this result is unique to Kob-Andersen glass models or universal. An application of the same analysis techniques to other model glass-forming liquids with varying fragility, interaction potentials, and compositions will help assess the generality of the observed structure-dynamics connections and identify potential universal features. Additionally, insights gained from this study may be channeled into existing theoretical frameworks, such as Mode-Coupling Theory (MCT), to improve their predictive capabilities and extend their applicability to higher temperatures. This could involve incorporating structural information into the dynamic descriptions or developing new theoretical approaches that account for the correlation between structure and dynamics across different

temperature regimes.

Furthermore, the insights gained from this chapter facilitate the use of more cost-effective and readily accessible datasets, such as those obtainable from experimental glass-formers like colloids, enabling the application of this methodology to experimental systems. A particularly promising avenue for future research lies in correlating structure with dynamics within these experimental settings. In similar fashion, connections between force chains in granular matter and the machine-learned order parameter can give further insight about correlation in disordered solids.

5.2 Final Remarks

This dissertation has illuminated both the efficacy and the inherent constraints of employing machine learning to decipher the complexities of supercooled liquids. While these models afford substantial predictive power, the attainment of clear physical interpretability remains a critical endeavor. Through a systematic analysis of the structure-dynamics relationship and a rigorous evaluation of how diverse machine learning methodologies capture fundamental physical principles, we have progressed towards a more comprehensive framework for understanding glassy dynamics. Our findings underscore the significant role of structural information, even in seemingly disordered systems, in dictating long-time relaxation behavior. The ongoing refinement of machine learning methodologies, informed by a deep understanding of physical phenomena, holds considerable potential for advancing our comprehension of supercooled liquids and the glass transition, and condensed matter in general.

Appendix A

Supplemental Information to Chapter 3

In this section, I detail the following: Sec. A.0.1 compares the average local structure obtained from various regression models with that derived from softness; Sec. A.0.2 defines the Cumulative Squared Displacement (CSD) function used in Chapter 3; Sec. A.0.3 details the energy barrier and entropic contributions for the multilayer perceptron models under parameter regimes where the total number of training examples is either half or double the total number of model parameters.

A.0.1 Mean trend for softness and log of the probability of rearrangement

In the chapter, we demonstrated that the distribution of $\ln p_r^{SVR}$ is approximately Gaussian, similar to the distribution of softness. Figure A.1 further shows that the mean predictions from the various models, particularly the SVR model, closely follow the trend observed for softness. For fairness in comparison, we plot the projections of the feature vector unto the hyperplane corresponding to the models as a function of temperature.

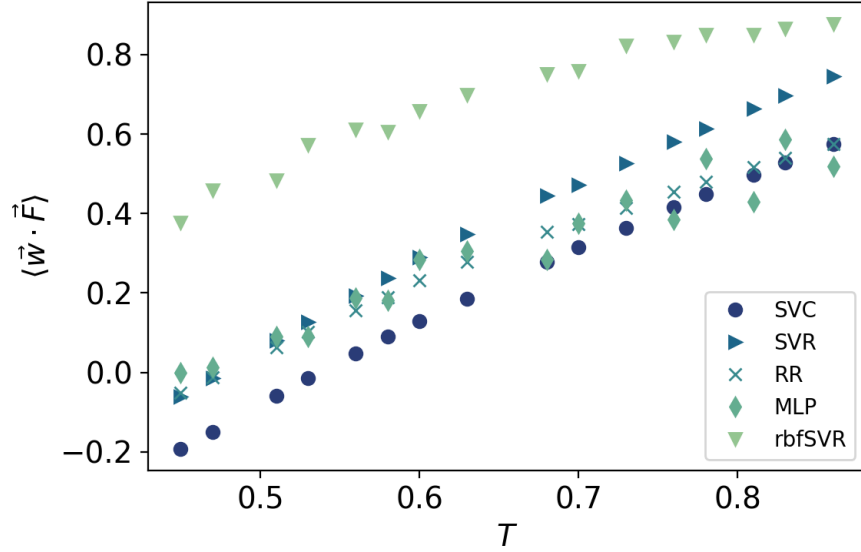


Figure A.1: **Trend of the mean of the distribution from the models as a function of Temperature.** The mean of the outputs from the models increases with temperature. However, the mean from the linearSVR and linearSVC models increases at almost the same rate as a function of temperature.

A.0.2 Definition of Cumulative Squared Displacement (CSD)

In the chapter, we mentioned quantifying dynamics using CSD in addition to p_{hop} . Inspired by the definition of the mean-squared displacement, we defined the cumulative squared displacement for a particle i at time t as

$$\Delta r^2(i, t) = \frac{1}{10} \sum_{t'=t}^{t-9} \|\vec{r}_i(t') - \vec{r}_i(t' - 1)\|^2$$

where $\vec{r}_i$ is the position of particle i . The particle i at time t is said to undergo a rearrangement if $\Delta r^2(i, t) > 0.0966$. This cutoff value is chosen such that the total number of activated events identified matches those identified through p_{hop} for a reference temperature of $T = 0.86$. The choice of cutoff used does not impact the qualitative result in chapter 3.

A.0.3 Influence of the ratio of training examples to model parameters

In the chapter, we presented the inferred energy barrier and the entropic component for the case where the number of training examples matches the total number of parameters in the Multilayer Perceptron (MLP) model, referred to as MLP. Additionally, we examined two other scenarios: one where the number of training examples is half the total number of parameters (MLP_H) and another where it is double (MLP_D). Figure A.2 illustrates that the inferred energy landscape is influenced by the ratio of training examples to the total number of parameters in the MLP models. While the energy barrier and entropic component remain correlated, the MLP_H and MLP_D models demonstrate slightly reduced sensitivity to the finer details of the energy landscape compared to the baseline MLP model.

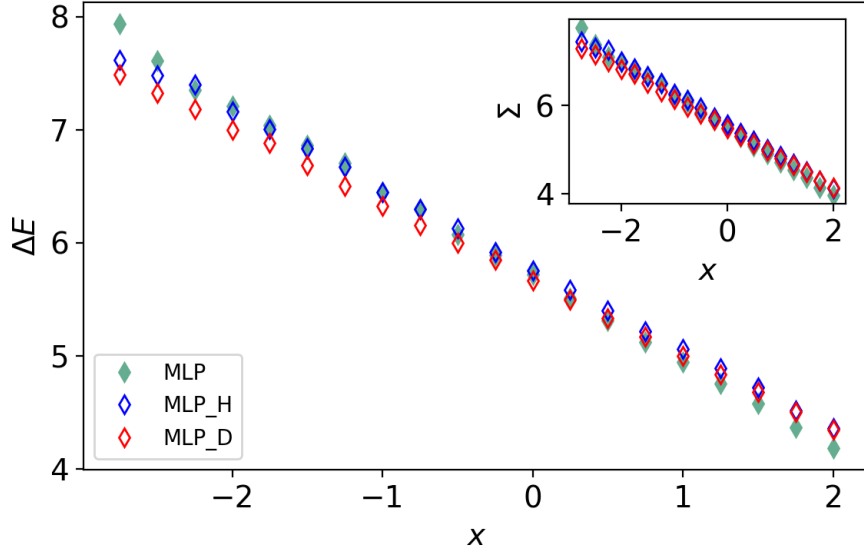


Figure A.2: **Details of the energy landscape inferred by the multilayer perceptron models.** The energy barrier ΔE , and the entropic piece, Σ , are correlated for the different MLP.

Bibliography

- [1] Juan P Garrahan. Dynamic heterogeneity comes to life. *Proceedings of the national academy of sciences*, 108(12):4701–4702, 2011.
- [2] Ludovic Berthier and Gilles Tarjus. The role of attractive forces in viscous liquids. *The Journal of chemical physics*, 134(21), 2011.
- [3] C Austen Angell. Formation of glasses from liquids and biopolymers. *Science*, 267(5206):1924–1935, 1995.
- [4] Samuel S Schoenholz, Ekin D Cubuk, Daniel M Sussman, Efthimios Kaxiras, and Andrea J Liu. A structural approach to relaxation in glassy liquids. *Nature Physics*, 12(5):469–471, 2016.
- [5] Victor Bapst, Thomas Keck, A Grabska-Barwińska, Craig Donner, Ekin Dogus Cubuk, Samuel S Schoenholz, Annette Obika, Alexander WR Nelson, Trevor Back, Demis Hassabis, et al. Unveiling the predictive power of static structure in glassy systems. *Nature physics*, 16(4):448–454, 2020.
- [6] Paul M Chaikin, Tom C Lubensky, and Thomas A Witten. *Principles of condensed matter physics*, volume 10. Cambridge university press Cambridge, 1995.
- [7] Morrel H Cohen and GS Grest. Liquid-glass transition, a free-volume approach. *Physical Review B*, 20(3):1077, 1979.

- [8] David Turnbull and Morrel H Cohen. On the free-volume model of the liquid-glass transition. *The journal of chemical physics*, 52(6):3038–3041, 1970.
- [9] Francis W Starr, Srikanth Sastry, Jack F Douglas, and Sharon C Glotzer. What do we learn from the local geometry of glass-forming liquids? *Physical review letters*, 89(12):125501, 2002.
- [10] Gary S Grest and Morrel H Cohen. Liquids, glasses, and the glass transition: a free-volume approach. *Advances in chemical physics*, pages 455–525, 1981.
- [11] Walter Kob and Hans C Andersen. Testing mode-coupling theory for a supercooled binary lennard-jones mixture i: The van hove correlation function. *Physical Review E*, 51(5):4626, 1995.
- [12] Ulf R Pedersen, Thomas B Schröder, and Jeppe C Dyre. Phase diagram of kob-andersen-type binary lennard-jones mixtures. *Physical review letters*, 120(16):165501, 2018.
- [13] Mark D Ediger. Spatially heterogeneous dynamics in supercooled liquids. *Annual review of physical chemistry*, 51(1):99–128, 2000.
- [14] Ludovic Berthier, Giulio Biroli, Jean-Philippe Bouchaud, Luca Cipelletti, and Wim van Saarloos. *Dynamical heterogeneities in glasses, colloids, and granular media*, volume 150. OUP Oxford, 2011.
- [15] Srikanth Sastry, Pablo G Debenedetti, and Frank H Stillinger. Signatures of distinct dynamical regimes in the energy landscape of a glass-forming liquid. *Nature*, 393(6685):554–557, 1998.
- [16] Ludovic Berthier, Giulio Biroli, Jean-Philippe Bouchaud, and Robert L Jack. Overview of different characterisations of dynamic heterogeneity. *Dynamical heterogeneities in glasses, colloids, and granular media*, 150:68, 2011.

- [17] Walter Kob, Claudio Donati, Steven J Plimpton, Peter H Poole, and Sharon C Glotzer. Dynamical heterogeneities in a supercooled lennard-jones liquid. *Physical review letters*, 79(15):2827, 1997.
- [18] Claudio Donati, Jack F Douglas, Walter Kob, Steven J Plimpton, Peter H Poole, and Sharon C Glotzer. Stringlike cooperative motion in a supercooled liquid. *Physical review letters*, 80(11):2338, 1998.
- [19] Giulio Biroli, Smarajit Karmakar, and Itamar Procaccia. Comparison of static length scales characterizing the glass transition. *Physical review letters*, 111(16):165701, 2013.
- [20] Aaron S Keys, Lester O Hedges, Juan P Garrahan, Sharon C Glotzer, and David Chandler. Excitations are localized and relaxation is hierarchical in glass-forming liquids. *Physical Review X*, 1(2):021013, 2011.
- [21] Yael S Elmatad, David Chandler, and Juan P Garrahan. Corresponding states of structural glass formers. ii. *The Journal of Physical Chemistry B*, 114(51):17113–17119, 2010.
- [22] Simone Ciarella, Rutger A Biezemans, and Liesbeth MC Janssen. Understanding, predicting, and tuning the fragility of vitrimeric polymers. *Proceedings of the National Academy of Sciences*, 116(50):25013–25022, 2019.
- [23] Daniel M Sussman, M Paoluzzi, M Cristina Marchetti, and M Lisa Manning. Anomalous glassy dynamics in simple models of dense biological tissue. *Europhysics Letters*, 121(3):36001, 2018.
- [24] Martin Goldstein. Viscous liquids and the glass transition: a potential energy barrier picture. *The Journal of Chemical Physics*, 51(9):3728–3739, 1969.

- [25] Pablo G Debenedetti and Frank H Stillinger. Supercooled liquids and the glass transition. *Nature*, 410(6825):259–267, 2001.
- [26] Pablo G Debenedetti, Frank H Stillinger, Thomas M Truskett, and Christopher J Roberts. The equation of state of an energy landscape, 1999.
- [27] Vassiliy Lubchenko and Peter G Wolynes. Theory of structural glasses and supercooled liquids. *Annu. Rev. Phys. Chem.*, 58(1):235–266, 2007.
- [28] Giorgio Parisi and Francesco Zamponi. Mean-field theory of hard sphere glasses and jamming. *Reviews of Modern Physics*, 82(1):789, 2010.
- [29] Patrick Charbonneau, Jorge Kurchan, Giorgio Parisi, Pierfrancesco Urbani, and Francesco Zamponi. Glass and jamming transitions: From exact results to finite-dimensional descriptions. *Annual Review of Condensed Matter Physics*, 8(1):265–288, 2017.
- [30] Patrick Charbonneau, Jorge Kurchan, Giorgio Parisi, Pierfrancesco Urbani, and Francesco Zamponi. Fractal free energy landscapes in structural glasses. *Nature communications*, 5(1):3725, 2014.
- [31] Marcus T Cicerone, Paul A Wagner, and MD Ediger. Translational diffusion on heterogeneous lattices: a model for dynamics in glass forming materials. *The Journal of Physical Chemistry B*, 101(43):8727–8734, 1997.
- [32] Daniel Diaz Vela and David S Simmons. The microscopic origins of stretched exponential relaxation in two model glass-forming liquids as probed by simulations in the isoconfigurational ensemble. *The Journal of Chemical Physics*, 153(23), 2020.
- [33] Asaph Widmer-Cooper and Peter Harrowell. Free volume cannot explain the

- spatial heterogeneity of debye–waller factors in a glass-forming binary alloy. *Journal of non-crystalline solids*, 352(42-49):5098–5102, 2006.
- [34] Asaph Widmer-Cooper and Peter Harrowell. On the relationship between structure and dynamics in a supercooled liquid. *Journal of Physics: Condensed Matter*, 17(49):S4025, 2005.
- [35] To Egami, K Maeda, and V Vitek. Structural defects in amorphous solids a computer simulation study. *Philosophical Magazine A*, 41(6):883–901, 1980.
- [36] Xiunan Yang, Rui Liu, Mingcheng Yang, Wei-Hua Wang, and Ke Chen. Structures of local rearrangements in soft colloidal glasses. *Physical review letters*, 116(23):238003, 2016.
- [37] Paul J Steinhardt, David R Nelson, and Marco Ronchetti. Bond-orientational order in liquids and glasses. *Physical Review B*, 28(2):784, 1983.
- [38] Hajime Tanaka, Hua Tong, Rui Shi, and John Russo. Revealing key structural features hidden in liquids and glasses. *Nature Reviews Physics*, 1(5):333–348, 2019.
- [39] Mathieu Leocmach, John Russo, and Hajime Tanaka. Importance of many-body correlations in glass transition: An example from polydisperse hard spheres. *The Journal of chemical physics*, 138(12), 2013.
- [40] Chengjie Xia, Jindong Li, Yixin Cao, Binqun Kou, Xianghui Xiao, Kamel Fezzaa, Tiqiao Xiao, and Yujie Wang. The structural origin of the hard-sphere glass transition in granular packing. *Nature communications*, 6(1):8409, 2015.
- [41] C Patrick Royall and Stephen R Williams. The role of local structure in dynamical arrest. *Physics Reports*, 560:1–75, 2015.

- [42] Glen M Hocky, Daniele Coslovich, Atsushi Ikeda, and David R Reichman. Correlation of local order with particle mobility in supercooled liquids is highly system dependent. *Physical review letters*, 113(15):157801, 2014.
- [43] Asaph Widmer-Cooper, Peter Harrowell, and H Fynewever. How reproducible are dynamic heterogeneities in a supercooled liquid? *Physical review letters*, 93(13):135701, 2004.
- [44] Asaph Widmer-Cooper and Peter Harrowell. On the study of collective dynamics in supercooled liquids through the statistics of the isoconfigurational ensemble. *The Journal of chemical physics*, 126(15), 2007.
- [45] Asaph Widmer-Cooper, Heidi Perry, Peter Harrowell, and David R Reichman. Irreversible reorganization in a supercooled liquid originates from localized soft modes. *Nature Physics*, 4(9):711–715, 2008.
- [46] Ludovic Berthier and Robert L Jack. Structure and dynamics of glass formers: Predictability at large length scales. *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, 76(4):041509, 2007.
- [47] Hayato Shiba, Masatoshi Hanai, Toyotaro Suzumura, and Takashi Shimokawabe. Botan: Bond targeting network for prediction of slow glassy dynamics by machine learning relative motion. *The Journal of Chemical Physics*, 158(8), 2023.
- [48] Emanuele Boattini, Frank Smallenburg, and Laura Filion. Averaging local structure to predict the dynamic propensity in supercooled liquids. *Physical Review Letters*, 127(8):088007, 2021.
- [49] Emanuele Boattini, Susana Marín-Aguilar, Saheli Mitra, Giuseppe Foffi, Frank Smallenburg, and Laura Filion. Autonomously revealing hidden local structures in supercooled liquids. *Nature communications*, 11(1):5479, 2020.

- [50] Jörg Behler and Michele Parrinello. Generalized neural-network representation of high-dimensional potential-energy surfaces. *Physical review letters*, 98(14):146401, 2007.
- [51] Samuel S Schoenholz, Ekin D Cubuk, Efthimios Kaxiras, and Andrea J Liu. Relationship between local structure and relaxation in out-of-equilibrium glassy systems. *Proceedings of the National Academy of Sciences*, 114(2):263–267, 2017.
- [52] Indrajit Tah, Sean A Ridout, and Andrea J Liu. Fragility in glassy liquids: A structural approach based on machine learning. *arXiv preprint arXiv:2205.07187*, 2022.
- [53] Ekin D Cubuk, Andrea J Liu, Efthimios Kaxiras, and Samuel S Schoenholz. Unifying framework for strong and fragile liquids via machine learning: a study of liquid silica. *arXiv preprint arXiv:2008.09681*, 2020.
- [54] Indrajit Tah, Tristan A Sharp, Andrea J Liu, and Daniel M Sussman. Quantifying the link between local structure and cellular rearrangements using information in models of biological tissues. *Soft Matter*, 17(45):10242–10253, 2021.
- [55] Daniel M Sussman, Samuel S Schoenholz, Ekin D Cubuk, and Andrea J Liu. Disconnecting structure and dynamics in glassy thin films. *Proceedings of the National Academy of Sciences*, 114(40):10601–10605, 2017.
- [56] Sean A Ridout and Andrea J Liu. The dynamics of machine-learned” softness” in supercooled liquids describe dynamical heterogeneity. *arXiv preprint arXiv:2406.05868*, 2024.
- [57] Hongyi Xiao, Ge Zhang, Entao Yang, Robert Ivancic, Sean Ridout, Robert Riggelman, Douglas J Durian, and Andrea J Liu. Identifying microscopic fac-

- tors that influence ductility in disordered solids. *Proceedings of the National Academy of Sciences*, 120(42):e2307552120, 2023.
- [58] Cécile Monthus and Jean-Philippe Bouchaud. Models of traps and glass phenomenology. *Journal of Physics A: Mathematical and General*, 29(14):3847, 1996.
- [59] Jeppe C Dyre. Master-equation approach to the glass transition. *Physical review letters*, 58(8):792, 1987.
- [60] Jean-Philippe Bouchaud. Weak ergodicity breaking and aging in disordered systems. *Journal de Physique I*, 2(9):1705–1713, 1992.
- [61] Sean A Ridout, Indrajit Tah, and Andrea J Liu. Building a “trap model” of glassy dynamics from a local structural predictor of rearrangements. *Europhysics Letters*, 144(4):47001, 2023.
- [62] Rinske M Alkemade, Emanuele Boattini, Laura Filion, and Frank Smalenburg. Comparing machine learning techniques for predicting glassy dynamics. *The Journal of Chemical Physics*, 156(20), 2022.
- [63] Tomilola M Obadiya and Daniel M Sussman. Using fluid structures to encode predictions of glassy dynamics. *Physical Review Research*, 5(4):043112, 2023.
- [64] Ludovic Berthier. Dynamic heterogeneity in amorphous materials. *arXiv preprint arXiv:1106.1739*, 2011.
- [65] Atreyee Banerjee, Manoj Kumar Nandi, Srikanth Sastry, and Sarika Maitra Bhattacharyya. Determination of onset temperature from the entropy for fragile to strong liquids. *The Journal of chemical physics*, 147(2):024504, 2017.

- [66] Giampaolo Folena, Silvio Franz, and Federico Ricci-Tersenghi. Rethinking mean-field glassy dynamics and its relation with the energy landscape: The surprising case of the spherical mixed p-spin model. *Physical Review X*, 10(3):031045, 2020.
- [67] Mark D Ediger, C Austen Angell, and Sidney R Nagel. Supercooled liquids and glasses. *The journal of physical chemistry*, 100(31):13200–13212, 1996.
- [68] Takeshi Kawasaki, Takeaki Araki, and Hajime Tanaka. Correlation between dynamic heterogeneity and medium-range order in two-dimensional glass-forming liquids. *Physical review letters*, 99(21):215701, 2007.
- [69] C Patrick Royall, Stephen R Williams, Takehiro Ohtsuka, and Hajime Tanaka. Direct observation of a local structural mechanism for dynamic arrest. *Nature materials*, 7(7):556–561, 2008.
- [70] Robert L Jack, Andrew J Dunleavy, and C Patrick Royall. Information-theoretic measurements of coupling between structure and dynamics in glass formers. *Physical review letters*, 113(9):095703, 2014.
- [71] M Lisa Manning and Andrea J Liu. Vibrational modes identify soft spots in a sheared disordered packing. *Physical Review Letters*, 107(10):108302, 2011.
- [72] Ekin D Cubuk, Samuel Stern Schoenholz, Jennifer M Rieser, Brad Dean Malone, Joerg Rottler, Douglas J Durian, Efthimios Kaxiras, and Andrea J Liu. Identifying structural flow defects in disordered solids using machine-learning methods. *Physical review letters*, 114(10):108001, 2015.
- [73] Gerhard Jung, Giulio Biroli, and Ludovic Berthier. Predicting dynamic heterogeneity in glass-forming liquids by physics-informed machine learning. *arXiv preprint arXiv:2210.16623*, 2022.

- [74] Daan Frenkel and Berend Smit. *Understanding molecular simulation: from algorithms to applications*. Elsevier, 2023.
- [75] Chih-Wei Hsu, Chih-Chung Chang, Chih-Jen Lin, et al. A practical guide to support vector classification, 2003.
- [76] Raphael Candelier, Olivier Dauchot, and Giulio Biroli. Building blocks of dynamical heterogeneities in dense granular media. *Physical review letters*, 102(8):088001, 2009.
- [77] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. Scikit-learn: Machine learning in python. *the Journal of machine Learning research*, 12:2825–2830, 2011.
- [78] Lydéric Bocquet, Annie Colin, and Armand Ajdari. Kinetic theory of plastic flow in soft glassy materials. *Physical review letters*, 103(3):036001, 2009.
- [79] Peter Hänggi, Peter Talkner, and Michal Borkovec. Reaction-rate theory: fifty years after kramers. *Reviews of modern physics*, 62(2):251, 1990.
- [80] Vincent Testard, Ludovic Berthier, and Walter Kob. Intermittent dynamics and logarithmic domain growth during the spinodal decomposition of a glass-forming liquid. *The Journal of chemical physics*, 140(16):164502, 2014.
- [81] Yan-Wei Li, Leon Loh Yeong Wei, Matteo Paoluzzi, and Massimo Pica Ciarrarra. Softness, anomalous dynamics, and fractal-like energy landscape in model cell tissues. *Physical Review E*, 103(2):022607, 2021.
- [82] Diogo EP Pinto, Daniel M Sussman, Margarida M Telo da Gama, and Nuno AM Araújo. Hierarchical structure of the energy landscape in the voronoi model of dense tissue. *Physical Review Research*, 4(2):023187, 2022.

- [83] Tomilola M Obadiya, Claudia Schmidt, Sean A Ridout, and Daniel M Sussman. Physical interpretability of machine learning models when correlating structure with dynamics: learning energy barriers from local structures in supercooled liquids. *In prep for now*.
- [84] Pablo G Debenedetti. Metastable liquids: concepts and principles. 2020.
- [85] Ludovic Berthier and Giulio Biroli. Glasses and aging, a statistical mechanics perspective on., 2009.
- [86] Gerhard Jung, Giulio Biroli, and Ludovic Berthier. Predicting dynamic heterogeneity in glass-forming liquids by physics-inspired machine learning. *Physical Review Letters*, 130(23):238202, 2023.
- [87] Zhao Fan, Jun Ding, and Evan Ma. Machine learning bridges local static structure with multiple properties in metallic glasses. *Materials Today*, 40:48–62, 2020.
- [88] Ryan S DeFever, Colin Targonski, Steven W Hall, Melissa C Smith, and Sapna Sarupria. A generalized deep learning approach for local structure identification in molecular simulations. *Chemical science*, 10(32):7503–7515, 2019.
- [89] P Ronhovde, S Chakrabarty, D Hu, M Sahu, KK Sahu, KF Kelton, NA Mauro, and Z Nussinov. Detecting hidden spatial and spatio-temporal structures in glasses and complex physical systems by multiresolution network clustering. *The European Physical Journal E*, 34:1–24, 2011.
- [90] Daniele Coslovich, Robert L Jack, and Joris Paret. Dimensionality reduction of local structure in glassy binary mixtures. *The Journal of Chemical Physics*, 157(20), 2022.

- [91] Simone Ciarella, Massimiliano Chiappini, Emanuele Boattini, Marjolein Dijkstra, and Liesbeth MC Janssen. Dynamics of supercooled liquids from static averaged quantities using machine learning. *Machine Learning: Science and Technology*, 4(2):025010, 2023.
- [92] Arabind Swain, Sean Alexander Ridout, and Ilya Nemenman. Machine learning that predicts well may not learn the correct physical descriptions of glassy systems. *Physical Review Research*, 6(3):033091, 2024.
- [93] Joshua A Anderson, Jens Glaser, and Sharon C Glotzer. Hoomd-blue: A python package for high-performance molecular dynamics and hard particle monte carlo simulations. *Computational Materials Science*, 173:109363, 2020.
- [94] Tomilola M Obadiya and Daniel M Sussman. Dataset for “does fluid structure encode predictions of glassy dynamics?”. <http://dx.doi.org/110.5281/zenodo.7469766>.
- [95] Jason W Rocks, Sean A Ridout, and Andrea J Liu. Learning-based approach to plasticity in athermal sheared amorphous packings: Improving softness. *APL Materials*, 9(2), 2021.
- [96] Isabelle Guyon, Jason Weston, Stephen Barnhill, and Vladimir Vapnik. Gene selection for cancer classification using support vector machines. *Machine learning*, 46:389–422, 2002.
- [97] Matt Harrington, Andrea J Liu, and Douglas J Durian. Machine learning characterization of structural defects in amorphous packings of dimers and ellipses. *Physical Review E*, 99(2):022903, 2019.
- [98] Nathan C Keim, Joseph D Paulsen, Zorana Zeravcic, Srikanth Sastry, and Sidney R Nagel. Memory formation in matter. *Reviews of Modern Physics*, 91(3):035002, 2019.

- [99] K Jonason, E Vincent, J Hammann, JP Bouchaud, and P Nordblad. Memory and chaos effects in spin glasses. *Physical Review Letters*, 81(15):3243, 1998.
- [100] K Jonason, P Nordblad, E Vincent, J Hammann, and J-P Bouchaud. Memory interference effects in spin glasses. *The European Physical Journal B-Condensed Matter and Complex Systems*, 13(1):99–105, 2000.
- [101] Jean-Philippe Bouchaud, Vincent Dupuis, Jacques Hammann, and Eric Vincent. Separation of time and length scales in spin glasses: Temperature as a microscope. *Physical review B*, 65(2):024439, 2001.
- [102] Marco Baity-Jesi, Enrico Calore, Andres Cruz, Luis Antonio Fernandez, José Miguel Gil-Narvi3n, Antonio Gordillo-Guerrero, David I3niguez, Antonio Lasanta, Andrea Maiorano, Enzo Marinari, et al. The mpemba effect in spin glasses is a persistent memory effect. *Proceedings of the National Academy of Sciences*, 116(31):15350–15355, 2019.
- [103] Marco Baity-Jesi, Enrico Calore, Andrea Cruz, Luis Antonio Fern3ndez, Jos3 Miguel Gil-Narvion, Isidoro Gonz3lez-Adalid Pemart3n, Antonio Gordillo-Guerrero, David I3niguez, Andrea Maiorano, Enzo Marinari, et al. Memory and rejuvenation effects in spin glasses are governed by more than one length scale. *Nature Physics*, 19(7):978–985, 2023.
- [104] H Yardimci and RL Leheny. Memory in an aging molecular glass. *Europhysics Letters*, 62(2):203, 2003.
- [105] Camille Scalliet and Ludovic Berthier. Rejuvenation and memory effects in a structural glass. *Physical review letters*, 122(25):255502, 2019.
- [106] K Fukao and A Sakamoto. Aging phenomena in poly (methyl methacrylate) thin films: memory and rejuvenation effects. *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, 71(4):041803, 2005.

- [107] Ludovic Bellon, Sergio Ciliberto, and Claude Laroche. Advanced memory effects in the aging of a polymer glass. *The European Physical Journal B-Condensed Matter and Complex Systems*, 25:223–231, 2002.
- [108] A Koike and M Tomozawa. Towards the origin of the memory effect in oxide glasses. *Journal of non-crystalline solids*, 354(28):3246–3253, 2008.
- [109] Ling-Nan Zou and Sidney R Nagel. Glassy dynamics in thermally activated list sorting. *Physical review letters*, 104(25):257201, 2010.
- [110] Laurent J Lewis and Göran Wahnström. Molecular-dynamics study of super-cooled ortho-terphenyl. *Physical Review E*, 50(5):3865, 1994.
- [111] KL Ngai and Marian Paluch. Classification of secondary relaxation in glass-formers based on dynamic properties. *The Journal of chemical physics*, 120(2):857–873, 2004.
- [112] Gyan P Johari and Martin Goldstein. Viscous liquids and the glass transition. ii. secondary relaxations in glasses of rigid molecules. *The Journal of chemical physics*, 53(6):2372–2388, 1970.
- [113] Gyan P Johari and Martin Goldstein. Viscous liquids and the glass transition. iii. secondary relaxations in aliphatic alcohols and other nonrigid molecules. *The Journal of Chemical Physics*, 55(9):4245–4252, 1971.
- [114] Gyan P Johari. Glass transition and secondary relaxations in molecular liquids and crystals. *Annals of the New York Academy of Sciences*, 279(1):117–140, 1976.
- [115] GP Johari. Localized molecular motions of β -relaxation and its energy landscape. *Journal of non-crystalline solids*, 307:317–325, 2002.

- [116] Hai Bin Yu, Wei Hua Wang, Hai Yang Bai, and Konrad Samwer. The β -relaxation in metallic glasses. *National Science Review*, 1(3):429–461, 2014.
- [117] Baicheng Mei, Yuyuan Lu, Lijia An, and Zhen-Gang Wang. Two-step relaxation and the breakdown of the stokes-einstein relation in glass-forming liquids. *Physical Review E*, 100(5):052607, 2019.