Title
Predicting the Young's modulus of silicate glasses by molecular dynamics simulations and machine learning

Author
Yang, Kai

Publication Date
2020

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Predicting the Young's modulus of silicate glasses by molecular dynamics simulations and machine learning

A thesis submitted in partial satisfaction

or the requirements for the degree Master of Science

in Civil Engineering

by

Kai Yang

2020

ABSTRACT OF THE THESIS

Predicting the Young's modulus of silicate glasses

by molecular dynamics simulations and machine learning

by

Kai Yang

Master of Science in Civil Engineering

University of California, Los Angeles, 2020

Professor Mathieu Bauchy, Chair

Understanding the compositional dependence of properties of silicate glass is critical to design novel glasses for various technology applications. With the development in molecular dynamics simulations and machine learning techniques, a combined and fully computational approach, which is able to reveal the relationship between glass composition and its mechanical properties, can be developed and served as a guide prior to experiments and manufacturing. On one hand, machine learning is a powerful tool to predict the properties based on the existing database. On the other hand, molecular dynamics simulation cannot only produce sufficient data points for machine learning models but also provide a detailed picture of the atomic structure of glasses. This atomic-scale knowledge from molecular dynamics simulation contains an intrinsic relationship between glass compositions and their mechanical properties.

Here, we first use molecular dynamics simulation to generate the dataset for calcium aluminosilicate glasses and apply different machine learning models to predict their Young's modulus using glass compositions in Chapter 1. Next, we apply topological constraint theory to quantify the atomic structures of simulated glasses and use this knowledge to predict Young's modulus for calcium aluminosilicate glass family in Chapter 2. Last, in Chapter 3, we propose a fully analytical model to link the network topology with glass compositions.

The thesis of Kai Yang is approved.

Gaurav Sant

Xiaoyu Zheng

Jaime Marian

Mathieu Bauchy, Committee Chair

University of California, Los Angeles

2020

To my beloved family,

for their love and supports.

To all my friends,

for their encouragement and helps.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

Chapter 1 is a reprint of **K. Yang**, X. Xu, B. Yang, B. Cook, H. Ramos, N. M. A. Krishnan, M. M. Smedskjaer, C. Hoover, and M. Bauchy. Predicting the Young's Modulus of Silicate Glasses using High-Throughput Molecular Dynamics Simulations and Machine Learning. Scientific Reports **9**, 8739 (2019)

Chapter 2 is a reprint of **K. Yang**, B. Yang, X. Xu, C. Hoover, M. M. Smedskjaer, and M. Bauchy. Prediction of the Young's Modulus of Silicate Glasses by Topological Constraint Theory. Journal of Non-Crystalline Solids **514**, 15 (2019).

Chapter 3 is a manuscript under submission process authored by **K. Yang**, Y. Hu, Z. Li, N. M. A. Krishnan, M. M. Smedskjaer, C. Hoover, J. Mauro, G. Sant, and M. Bauchy.

Chapter 1. Predicting the Young's Modulus of Silicate Glasses using High-Throughput Molecular Dynamics Simulations and Machine Learning

1.1 Abstract

The application of machine learning to predict materials' properties usually requires a large number of consistent data for training. However, experimental datasets of high quality are not always available or self-consistent. Here, as an alternative route, we combine machine learning with high-throughput molecular dynamics simulations to predict the Young's modulus of silicate glasses. We demonstrate that this combined approach offers good and reliable predictions over the entire compositional domain. By comparing the performances of select machine learning algorithms, we discuss the nature of the balance between accuracy, simplicity, and interpretability in machine learning.

1.2 Introduction

Improving the mechanical properties of glasses is crucial to address major challenges in energy, communications, and infrastructure [1]. In particular, the stiffness of glass (e.g., its Young's modulus *E*) plays a critical role in flexible substrates and roll-to-roll processing of displays, optical fibres, architectural glazing, ultra-stiff composites, hard discs and surgery equipment, or lightweight construction materials [1–4]. Addressing these challenges requires the discovery of new glass compositions featuring tailored mechanical properties [5,6].

Although the discovery of new materials with enhanced properties is always a difficult task, glassy materials present some unique challenges. First, a glass can be made out of virtually all the elements of the periodic table if quenched fast enough from the liquid state [7]. Second, unlike crystals, glasses are out-of-equilibrium phases and, hence, do not have to obey any fixed stoichiometry [8]. These two unique properties of glass open limitless possibilities for the development of new compositions with enhanced properties—for instance, the total number of possible glass compositions [7] has been estimated to be around 10^{52}! Clearly, only a tiny portion of the compositional envelope accessible to glass has been explored thus far.

The design of new glasses for a targeted application can be formulated as an optimization problem, wherein the composition needs to be optimized to minimize or maximize a cost function (e.g., the Young's modulus) while satisfying some constraints (e.g., ensuring low cost and processability) [9]. Although the vast compositional envelop accessible to glass opens limitless possibilities for compositional tuning, optimization problems in such highly-dimensional spaces are notoriously challenging—which is known as the "curse of dimensionality." Namely, the virtually infinite number of possible glass compositions render largely inefficient traditional discovery methods based on trial-and-error Edisonian approaches [10].

To overcome this challenge, the development of predictive models relating the composition of glasses to their engineering properties is required [9]. Ideally, physics-based models should offer the most robust predictions. In the case of glass stiffness, the Makishima–Mackenzie (MM) model may be the most popular predictive model [11,12]. This approach is essentially an additive model, wherein stiffness is expressed as a linear function of the oxide concentrations. However, such additive models are intrinsically unable to capture any non-linear compositional dependence,

as commonly observed for stiffness [1,5,13]. On the other hand, molecular dynamics (MD) simulations offer a powerful method to compute the stiffness of a given glass [14,15]. However, MD is a brute-force method, that is, it requires (at least) one simulation per glass composition—so that the systematic use of MD to explore the large compositional envelop accessible to glass is not a realistic option.

In turn, machine learning (ML) offers an attractive and pragmatic approach to predict glasses' properties [16]. In contrast with physics-based models, ML-based models are developed by "learning" from existing databases. Although the fact that glass composition can be tuned in a continuous fashion renders glass an ideal material for ML methods, the application of ML to this material has been rather limited thus far [16–20]. This partially comes from the fact that ML methods critically relies on the existence of "useful" data. To be useful, data must be (i) available (i.e., easily accessible), (ii) complete (i.e., with a large range of parameters), (iii) consistent (i.e., obtained with the same testing protocol), (iv) accurate (i.e., to avoid "garbage in, garbage out" models), and (v) representative (i.e., the dataset needs to provide enough information to train the models). Although some glass property databases do exist [21], some inconsistencies in the ways glasses are produced or tested among various groups may render challenging their direct use as training sets for ML methods—or would require some significant efforts in data cleaning and non-biased outlier detection.

To overcome these challenges, we present here a general method wherein high-throughput molecular dynamics simulations are coupled with machine learning methods to predict the relationship between glass composition and stiffness. Specifically, we take the example of the ternary calcium aluminosilicate (CAS) glass system—which is an archetypical model for alkali-

free display glasses [22]—and focus on the prediction of their Young's modulus. We show that our method offers good and reliable predictions of the Young's modulus of CAS glasses over the entire compositional domain. By comparing the performance of select ML algorithms—polynomial regression (PR), LASSO, random forest (RF), and artificial neural network (ANN)—we show that the artificial neural network algorithm offers the highest level of accuracy. Based on these results, we discuss the balance between accuracy, complexity, and interpretability offered by each ML method.

1.3 Methods

1.3.1 **High-throughput molecular dynamics simulations**

To establish our conclusions, we use molecular dynamics simulations to create a database consisting of the Young's modulus values of 231 glasses homogeneously covering the CAS ternary system, with 5% increments in the mol% concentration of the CaO, Al_2O_3, and SiO_2 oxide constituents. At this point, no consideration is made as to whether all these compositions would experimentally exhibit satisfactory glass-forming ability. All the simulations are conducted using the Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) package [23]. Each glass comprises around 3000 atoms. We adopt here the interatomic potential parametrized by Jakse—as it has been found to yield some structural and elastic properties that are in good agreement with experimental data for CAS glasses [24,25]. A cutoff of 8.0 Å is used for the short-range interactions. The Coulombic interactions are calculated by adopting the Fennell damped shifted force model with a damping parameter of 0.25 $Å^{-1}$ and a global cutoff of 8.0 Å [26]. The integration timestep is kept fixed 1.0 fs.

The glass samples are prepared with the conventional melt-quench method as described in the following [27]. First, some atoms are randomly placed in a cubic box using PACKMOL while using a distance cutoff of 2.0 Å between each atom to avoid any unrealistic overlap [28]. These initial configurations are then subjected to an energy minimization, followed by some 100 ps relaxations in the canonical (*NVT*) and isothermal-isobaric (*NPT*) ensembles at 300 K, sequentially. These samples are then fully melted at 3000 K for 100 ps in the *NVT* and, subsequently, *NPT* ensemble (at zero pressure) to ensure the loss of the memory of the initial configurations and to equilibrate the systems. Next, these liquids are cooled from 3000 to 300 K in the *NPT* ensemble at zero pressure with a cooling rate of 1 K/ps. The obtained glass samples are further relaxed at 300 K for 100 ps in the *NPT* ensemble before the stiffness computation. Note that this quenching procedure was slightly adjusted for select compositions. First, a higher initial melting temperature of 5000 K is used for the samples wherein the SiO_2 concentration is larger or equal to 95 mol%—since these glasses exhibit high glass transition temperatures. Second, a faster cooling rate of 100 K/ps is used for the samples wherein the CaO concentration is larger or equal to 90 mol%. Indeed, although the cooling rate can affect the glass stiffness, the use of a higher cooling rate here is necessary as these systems would otherwise tend to crystallize with a cooling rate of 1 K/ps.

The stiffness tensor $C_{\alpha\beta}$ of the equilibrated glasses is then computed by performing a series of 6 deformations (i.e., 3 axial and 3 shear deformations along the 3 axes) and computing the curvature of the potential energy [24,29]:

$$C_{\alpha\beta} = \frac{1}{V} \frac{\partial^2 U}{\partial e_\alpha \partial e_\beta} \quad (1)$$

where V is the volume of the glass, U is the potential energy, e is the strain, and α and β are some indexes representing each Cartesian direction. Note that all of the glass samples are found to be largely isotropic—so that the Young's modulus (E) can be calculated as:

$$E^{-1} = (S_{11} + S_{22} + S_{33})/3 \quad (2)$$

where $S = C^{-1}$ is the compliance matrix [15]. Based on previous results[24], the Jakse forcefield is found to systematically overestimate the Young's modulus of CAS glasses by about 16%—which may be a spurious effect arising from the fast cooling rate used in MD simulations or the parameterization of the forcefield. As such, the computed Young's modulus values are rescaled by this constant factor before serving as a training set for the machine learning models presented in the following.

1.3.2 Machine learning methodology

The 231 Young's modulus values computed by the high-throughput MD simulations serve as a database to infer the relationship between glass composition (x, y) and E in the $(CaO)_x(Al_2O_3)_y(SiO_2)_{1-x-y}$ glass system by ML. In details, we consider x and y to be the only inputs of the model (i.e., we neglect herein the effect of the thermal history of the glasses), whereas E is used as an output. Note that a similar approach can be used to predict the effect of composition on the shear modulus G, bulk modulus K, or Poisson's ratio ν. In the following, we briefly describe our overall ML strategy as well as the different ML algorithms that are considered and compared herein.

To avoid any risk of overfitting, a fraction of the data points is kept fully unknown to the models and is used as a "test set" to *a posteriori* assess the accuracy of each model. However,

isolating a fixed set would further reduce the number of points used for training our models, which can be a serious issue in the case of a small dataset as herein. To overcome this limitation, we adopt here the *k*-fold cross-validation (CV) technique [30]. The CV technique consists in splitting the dataset into *k* smaller sets, wherein the model is trained on "$k - 1$" of the folds and tested on the remaining of the data. The results are then averaged by iteratively using each of the *k* folds. Furthermore, we apply the nested CV to conduct evaluations for different algorithms and models. [31] In detail, we perform a 4-fold outer CV to split the dataset into the training set (75% of data) and test set (25% of data), and use the average scores (i.e., R^2) to compare the performance between different algorithms. Next, to obtain a proper setting for the hyperparameters of each model, we apply 10-fold inner CV within the training set. By taking the advantage of nested CV technique, each data point can be fairly chosen for training, testing and validating.

For optimal predictions, ML models must achieve the best balance between accuracy and simplicity—wherein models that are too simple are usually poorly accurate (i.e., "underfitted"), whereas models that are too complex present the risk of placing too much weight on the noise of the training set and, thereby, often show poor transferability to unknown sets of data (i.e., "overfitted"). Hence, one needs to identify the optimal degree of complexity (e.g., number of terms, number of neurons, etc.) for each model. Here, we optimize the degree of complexity of each model by gradually increasing its complexity and tracking the accuracy of the model prediction for both the training and test sets. Indeed, although the accuracy of the training set prediction typically monotonically increases with increasing model complexity, overfitted models usually manifest themselves by a decrease in the accuracy of the test set prediction.

We herein adopt the polynomial regression (PR), LASSO, random forest (RF) and multilayer perceptron artificial neural network (ANN) algorithms to generate the predictive models. This choice is motivated by the fact that these methods exhibit varying degrees of complexity and interpretability. This allows us to assess the nature of the trade-off between accuracy, simplicity, and interpretability offered by these algorithms.

1.3.4 Accuracy of the models

We assess the accuracy of each model (with different degrees of complexity) by computing the RMSE (root-mean-square error) and R^2 (coefficient of determination) factors. The RMSE factor measures the average Euclidian distance between the predicted and real values as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - Y_i')^2} \quad (3)$$

where Y_i and Y_i' are the predicted and real output values, respectively. The RMSE has the property of being in the same units as the output variables and, hence, can be used to estimate the accuracy of the Young's modulus values predicted by each model (namely, lower RMSE values indicate higher accuracy). Here, we use the RMSE of the training and test sets to determine the optimal degree of complexity for each ML model.

In complement of RMSE, we compute the R^2 factor, which is the percentage of the response variable variation. This factor can be used to quantify how close the data are to the fitted line. R^2 = 1 indicates a perfect prediction, while smaller values indicate less accurate predictions. Here, we

use the R^2 factor to compare the performances of each ML algorithm (once the degree of complexity has been optimized based on the RMSE).

1.4 Results

1.4.1Molecular dynamics simulations

MD simulations are first used to generate a series of 231 glasses that homogeneously span the CAS compositional ternary domain (see Methods section). The Young's modulus (E) of each glass is then computed by MD. We first focus on the compositional dependence of the Young's modulus values predicted by the MD simulations (see Fig. 1.1). Overall, we observe the existence of two main trends: (i) E tends to increase with increasing Al_2O_3 concentration and (ii) E tends to increase with increasing CaO concentration. However, we note that the dependence of E on composition is non-systematic and that CaO and Al_2O_3 have some coupled effects. For example, we find that E increases as the concentration of CaO increases when $[Al_2O_3] = 0$ mol%, whereas E decreases with increasing CaO concentration when $[Al_2O_3] > 40$ mol%. Overall, we find that E exhibits a non-linear dependence on composition—so that one likely cannot rely on simple additive models to predict Young's modulus in the CAS system.

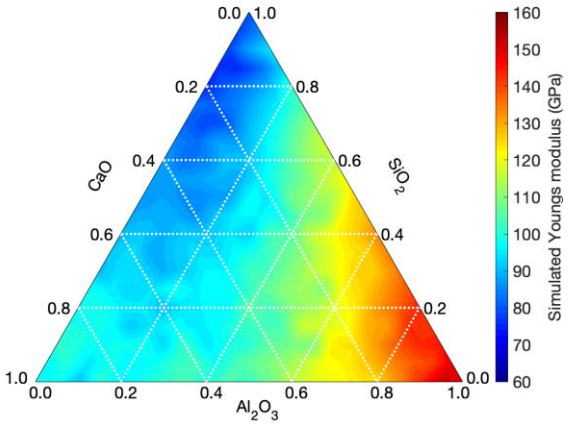

Figure 1. 1 Ternary diagram showing the Young's modulus values predicted by high-throughput molecular dynamics simulations as a function of composition in the $CaO–Al_2O_3–SiO_2$ glass system. This database consists of 231 compositions homogeneously distributed over the entire compositional domain with 5 mol% increments in the oxide concentrations. This database is used as a basis to train the machine learning models presented herein.

1.4.2 Relationship between composition and Young's modulus

We now discuss the nature of the relationship between composition and Young's modulus. In general, the Young's modulus tends to increase with increasing connectivity [4]. To assess whether this trend is here satisfied (and whether it can be used to predict the linkage between composition and *E*), we compute based on the MD simulations the average coordination number <*r*> of the atoms in the network for each glass composition. As shown in Fig. 1.2a, we find that <*r*> increases with increasing CaO and Al_2O_3 concentrations. This arises from that fact that (i) Ca atoms have a large coordination number (around 6), while (ii) the addition of Al atoms tends to increase the degree of polymerization of the glass, i.e., by converting non-briding oxygen (NBO) into bridging oxygen (BO) atoms (we also note the formation of 5- and 6-fold over-coordinated

Al species in Al-rich glasses). Overall, we observe that the ternary plot of <*r*> (Fig. 1.2a) echoes that of *E* (Fig. 1.1), which supports the fact that *E* increases upon increasing network connectivity. Nevertheless, as shown in Fig. 1.2b, we find that, although *E* and <*r*> are indeed positively correlated with each other, the data points are widely spread and the coefficient of determination R^2 only equals 0.623. This indicates that the <*r*> metric alone does not contain enough information to predict *E* and that other effects are not captured by simply considering the connectivity of the network—which renders challenging the development of a robust physics-based predictive model.

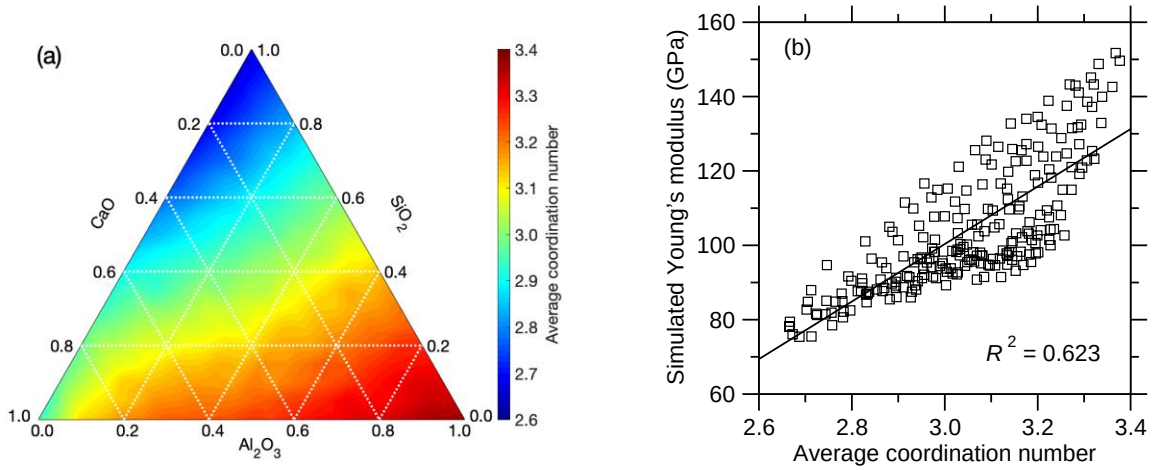

Figure 1. 2 (a) Ternary diagram showing the average atomic coordination number computed by high-throughput molecular dynamics simulations as a function of composition in the $CaO–Al_2O_3–SiO_2$ glass system. (b) Young's modulus computed by molecular dynamics simulations as a function of the average atomic coordination number. The line is a linear fit. The coefficient of determination R^2 indicates the degree of linearity.

We now assess the ability of the popular Makishima–Mackenzie (MM) model to predict the compositional evolution of *E* [11]. The MM model relies on an additive relationship, wherein *E* is expressed as a weighted average of the dissociation energies of each oxide constituent. In details, the Young's modulus *E* is expressed as:

$$E = 83.6\, V_t \sum_{i=1}^{n} X_i G_i \qquad (4)$$

where V_t is the overall packing density of the glass, and X_i and G_i are the concentration and volumic dissociation energy of each oxide constituent i, respectively. Note that the G_i terms are tabulated values, whereas V_t depends on the glass composition and is an explicit input to the model (i.e., the knowledge of the compositional dependence of V_t is a prerequisite to the MM model). To this end, we compute the packing density V_t of each glass based on the MD simulations. Figure 1.3a shows the ternary diagram of the *E* values predicted by the MM model as a function of composition in the CAS glass system. We observe that the MM model properly predicts the increase of *E* with increasing Al_2O_3 concentration, but fails to predict the increase in *E* upon increasing CaO concentration. This is due to fact that the dissociation energy terms associated with the CaO and SiO_2 oxides are close to each other (i.e., 15.5 and 15.4 kcal/cm^3, respectively), whereas that of Al_2O_3 (32 kcal/cm^3) is significantly higher. Overall, we observe that the MM model does not properly predict the non-linear dependence of *E* on composition. This is not surprising as the MM is essentially an additive model (although some level of non-linearity can exist within the V_t term). The MM model also fails to describe any coupling between the effects of CaO and Al_2O_3. Figure 1.3b shows a comparison between the Young's modulus values predicted by the MM model and computed by MD. Overall, we find that, although the MM model offers a fairly good prediction of *E*, the correlation remains poor (with R^2 = 0.556). In addition, we find that the MM model underestimates *E*, especially in the low *E* region (which corresponds to the technologically important low-Al compositional domain wherein glasses exhibit good glass-forming ability). Overall, we note that, although the MM model can be used as a rough guide to infer some compositional trends, it cannot be used to accurately predict *E* in CAS glasses.

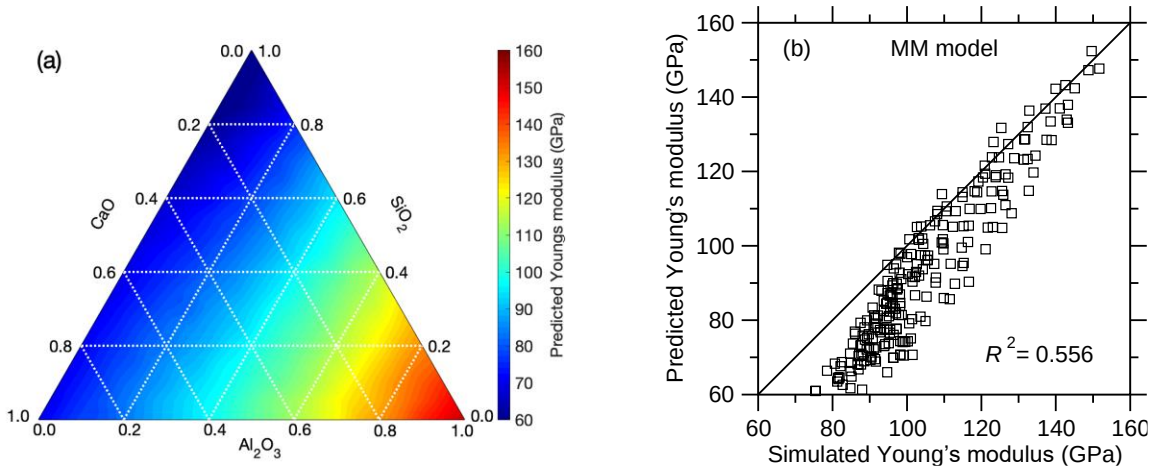

Figure 1. 3 (a) Ternary diagram showing the Young's modulus values E predicted by the Makishima-Mackenzie (MM) model as a function of composition in the $CaO–Al_2O_3–SiO_2$ glass system. (b) Comparison between the Young's modulus values predicted by the MM model and computed by molecular dynamics simulations.

1.4.3 Polynomial regression

The Young's modulus values computed by MD then serve as database to infer the relationship between glass composition (x, y) and E in the $(CaO)_x(Al_2O_3)_y(SiO_2)_{1–x–y}$ glass system by ML (see Methods section). In the following, we compare the performance of select ML algorithms. To this end, a fraction (25%) of the data points is kept fully unknown to the models and is used as a "test set" to *a posteriori* assess the accuracy of each model, whereas the rest of the data (75%) is used as a "training set." The accuracy of the prediction is assessed by calculating the root-mean-square error (RMSE, see Methods section).

We first focus on the outcomes of polynomial regression (PR). Figure 1.4a shows the RMSE offered by polynomial regression for the training and test sets as a function of the maximum polynomial degree considered in the model. As expected, we observe that the RMSE of the training

set decreases upon increasing polynomial degree (i.e., increasing model complexity) and eventually plateaus. This signals that, as the model becomes more complex, it can better interpolate the training set. In contrast, we observe a significant increase in the RMSE when the polynomial degree is equal to 1 or 2—which indicates that, in this domain, the model is underfitted. This confirms again that linear models based on additive relationships are unable to properly describe the linkages between composition and Young's modulus. On the other hand, we observe that the RMSE of the test set initially decreases with increasing polynomial degree, shows a minimum for degree 3, and eventually increases with increasing degree. This demonstrates that the models incorporating some polynomial terms that are strictly larger than 3 are overfitted. This arises from the fact that, in the case of high degrees, the model starts to fit the noise of the training set rather than the "true" overall trend. This exemplifies (i) how the training set allows identifying the minimum level of model complexity that is required to avoid underfitting and (ii) how the test set allows us to track the maximum level of model complexity before overfitting. Overall, the optimal polynomial degree (here found to be 3) manifests itself by a minimum in the RMSE of the test set.

We now focus on assessing the accuracy of the predictions of the best polynomial regression model (i.e., with a maximum polynomial degree of 3). Figure 1.4b shows a comparison between the Young's modulus predicted by the ML model and computed by MD. We find that the R^2 factors for the training and test sets are 0.975 and 0.970, respectively. This indicates that, even in the case of a simple algorithm like polynomial regression, ML offers a good prediction of E based on the simulated results.

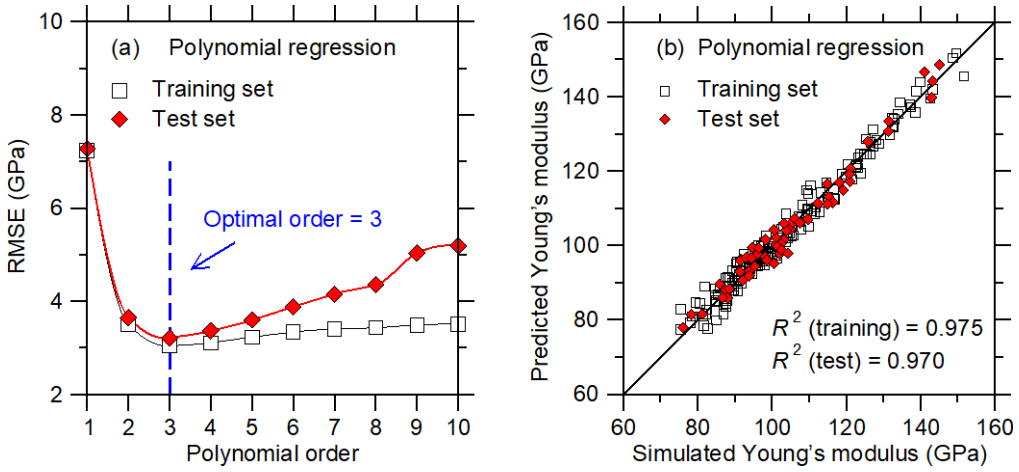

Figure 1. 4 (a) Accuracy (as captured by the RMSE value) of the polynomial regression models as a function of the maximum polynomial degree considered in each model (see Sec. 2b)—as obtained for the training and test set, respectively. The optimal polynomial order is chosen as that for which the RMSE of the test set is minimum. (b) Comparison between the Young's modulus values predicted by polynomial regression (with a degree of 3) and computed by molecular dynamics simulations.

1.4.5 LASSO

We now focus on the outcomes of the LASSO algorithm, which aims to reduce the complexity of the model by placing an extra weight on the model coefficients (see Methods section). Figure 1.5a shows the RMSE offered by LASSO for the training and test sets as a function of the degree of complexity, $-\log(\lambda)$, of the model. In contrast with the outcomes of the polynomial regression, we observe that LASSO does not yield any noticeable overfitting at high model complexity—which would manifest itself by an increase in the RMSE of the test set. This can be understood from the fact that the LASSO algorithm specifically aims to reduce the number of polynomial terms to mitigate the risk of overfitting. Here, since the RMSE of the test set only

shows a plateau with increasing –log(λ), we select the optimal degree of complexity as the one for which the RMSE of the test set becomes less than one standard deviation away from the minimum RMSE (i.e., in the plateau regime).

We now focus on assessing the accuracy of the predictions of the best LASSO model (i.e., with the optimal degree of complexity). Figure 1.5b shows a comparison between the Young's modulus predicted by the ML model and computed by MD. We find that the R^2 factors for the training and test sets are 0.971 and 0.966, respectively. Here, LASSO offers a slight decrease in accuracy as compared to polynomial regression.

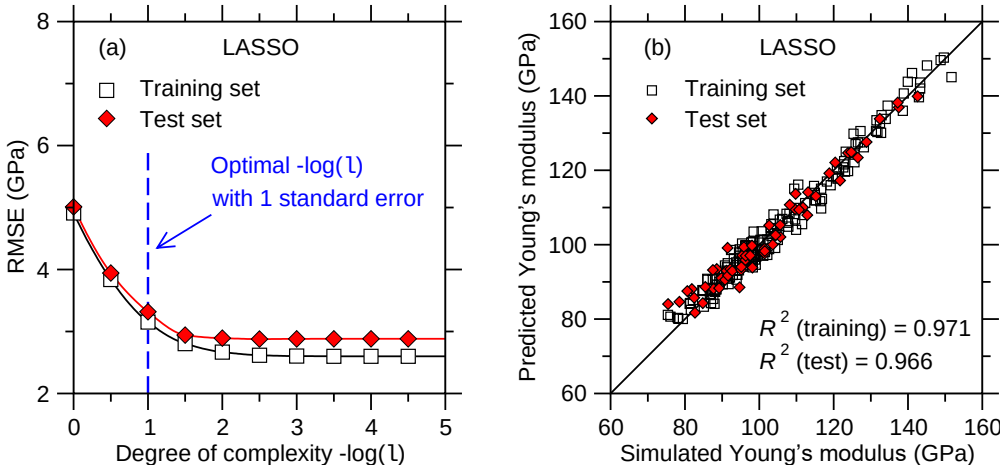

Figure 1. 5 (a) Accuracy (as captured by the RMSE value) of the LASSO models as a function of the degree of complexity (see Methods section)—as obtained for the training and test set, respectively. The optimal degree of complexity is determined as the one for which the RMSE of the test set is one standard deviation away from the minimum RMSE (i.e., in the plateau regime). (b) Comparison between the Young's modulus values predicted by LASSO (with an optimal degree of complexity) and computed by molecular dynamics simulations.

1.4.6 Random forest

We now focus on the outcomes of the RF algorithm. Figure 1.6a shows the RMSE offered by RF for the training and test sets as a function of the number of trees (i.e., which characterizes the complexity of the model). As observed in the case of LASSO, we find that RF does not yield any noticeable overfitting at high model complexity, that is, the RMSE of the test set only plateaus upon increasing number of trees. Here, we select 200 as being the optimal number of trees.

We now focus on assessing the accuracy of the predictions of the best RF model (i.e., with 200 trees). Figure 1.6b shows a comparison between the Young's modulus predicted by the ML model and computed by MD. We find that the R^2 factors for the training and test sets are 0.991 and 0.965, respectively. This suggests that, although RF offers an excellent interpolation of the training set (i.e., with a higher R^2 value than those obtained with the other ML models), its ability to offer a good prediction of the test set is slightly lower than those of the other ML models considered herein.

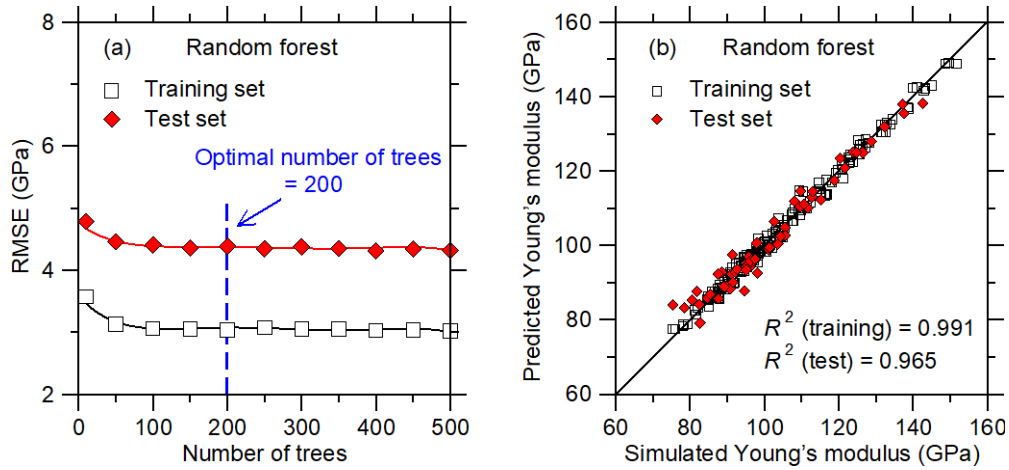

Figure 1. 6 (a) Accuracy (as captured by the RMSE value) of the random forest models as a function of the number of trees considered in each model (see Sec. 2d)—as obtained for the

training and test set, respectively. The optimal number of trees is taken as the threshold at which the RMSE of the test set starts to plateau. (b) Comparison between the Young's modulus values predicted by random forest (with 200 trees) and computed by molecular dynamics simulations.

1.4.7 Artificial neural network.

Finally, we focus on the outcomes of the ANN algorithm. Herein, we adopt a multilayer perceptron (MLP) ANN, which is a class of feedforward neural network containing an input layer, hidden layer, and an output layer, is trained using the back-propagation algorithm. We train ANN models with one hidden layer—which is found to be sufficient considering the nature of the training set. Figure 1.7a shows the RMSE offered by ANN for the training and test sets as a function of the number of neurons (i.e., which characterizes the complexity of the model). Overall, as previously observed in the cases of LASSO and RF, ANN does not yield any noticeable overfitting at high model complexity. Nevertheless, we note that the RMSE of the test set exhibits a slight minimum in the case of 5 neurons, which is the degree of complexity that we adopt herein.

We now focus on assessing the accuracy of the predictions of the best ANN model (with one hidden layer and five neurons). Figure 1.7b shows a comparison between the Young's modulus predicted by the ML model and computed by MD. We find that the R^2 factors for the training and test sets are 0.980 and 0.975, respectively. Overall, we find that the ANN algorithm offers the most accurate model among all the ML methods considered herein—as quantified in terms of the RMSE of the test set.

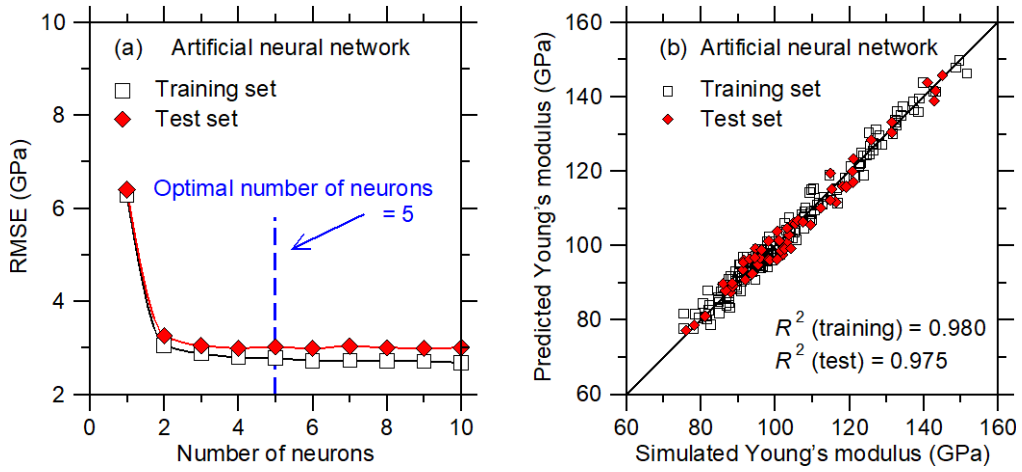

Figure 1. 7 (a) Accuracy (as captured by the RMSE value) of the artificial neural network models as a function of the number of neurons considered in each model (see Methods section)—as obtained for the training and test set, respectively. The optimal number of neurons is determined as that for which the RMSE value of the test set is minimum. (b) Comparison between the Young's modulus values predicted by artificial neural network (with 5 neurons) and computed by molecular dynamics simulations.

1.5 Discussion

We now compare the performance of the different machine learning algorithms used herein. We first focus on the level of accuracy offered by each method. To this end, Table 1.1 presents the coefficient of determination R^2 of each method for the training set (which characterizes the ability of the algorithm to properly interpolate the training data) and test set (which captures the accuracy of the model when predicting unknown data). We first observe that the RF algorithm offers the best interpolation on the training set (i.e., RF shows the highest R^2 for the training set). However, the RF algorithm also yields the lowest level of accuracy for the test set. This suggests that the RF algorithm presents the lowest ability to properly interpolate Young's modulus values in between

two compositions of the training set and/or to offer realistic extrapolations toward the edges of the compositional domain. On the other hand, we note that LASSO offers a lower accuracy in both training and test set, compared to PR. While gaining simplicity from the penalized regression method, LASSO also loses the accuracy (see Tab. 1). Nevertheless, we observe that the artificial neural network algorithm clearly offers the highest level of accuracy among all the models considered herein since it yields the highest R^2 value for the test set.

Table 1. 1 Comparison between the levels of accuracy, complexity, and interpretability offered by the machine learning algorithms used herein, namely, polynomial regression (PR), LASSO, random forest (RF), artificial neural network (ANN). The level of accuracy is described by the coefficient of determination (R^2) for the training and test sets. The complexity is described in parenthesis by the number of non-zero parameters in PR and LASSO, the number of trees in RF, and the product of the number of inputs, neurons, and parameters in each individual neuron in ANN. The "interpretability" describes the degree to which a human can understand the outcome produced by each model.

ML algorithms	Coefficient of determination R^2		Complexity	Interpretability
	Training set	Test set		
PR	0.975	0.970	Low (9)	High
LASSO	0.971	0.966	Low (8)	High
RF	0.991	0.965	High (200)	Intermediate
ANN	0.980	0.975	Intermediate (20)	Low

To further characterize the accuracy offered by each ML algorithm, Figure 1.8 shows the Young's modulus values that are predicted for two series of compositions, namely, (i) $(CaO)_x(Al_2O_3)_{40-x}(SiO_2)_{60}$, wherein the SiO_2 fraction is kept constant and equal to 60 mol% and (ii) $(CaO)_x(Al_2O_3)_x(SiO_2)_{100-2x}$, wherein the CaO/Al_2O_3 molar ratio is kept constant and equal to 1. These two series specifically aim to investigate (i) the effect of the degree of polymerization of the network (i.e., fraction of non-bridging oxygen) and (ii) the effect of network-forming atoms (i.e., Si vs. Al) at constant degree of depolymerization (i.e., in fully charge-compensated glasses). We first note that, in contrast to the other ML methods, RF yields piecewise-constant-shape results, which arises from the fact that the RF method is essentially based on an ensemble of decision trees. In details, the decision tree algorithm works by relying on a binary split, that is, at each node, randomly selected observations are dropped to either the left or right daughter node depending on the values and selected features. Although a single decision tree cannot capture any non-linearity within a dataset, the output of the model is eventually averaged over all its trees—so that an RF model can capture the non-linearity of a set of data by comprising enough trees. Nevertheless, we observe here that the piecewise-constant nature of single decision trees remains encoded in the outcome of this method, which yields non-smooth predictions. This feature of the RF algorithm likely explains its excellent ability to interpolate the training set while offering only a fair prediction of the test set.

We now compare the predictions offered by the PR and LASSO algorithms. Overall, although LASSO yields lower R^2 values on both training and test set compared to PR, we note that LASSO offers an improved prediction of E at the edges of the training set (see Figs. 1.8a and 1.8b). For instance, we note that the PR method predicts an unrealistic slight increase in E in pure SiO_2 (see the right end of Fig. 1.8b). This non-monotonic evolution of E at the edges of the

compositional domain suggests that the PR model might be slightly overfitted. In turn, such behaviour is mitigated by the LASSO algorithm. Finally, we find that ANN offers the best description of the non-linear nature of the data.

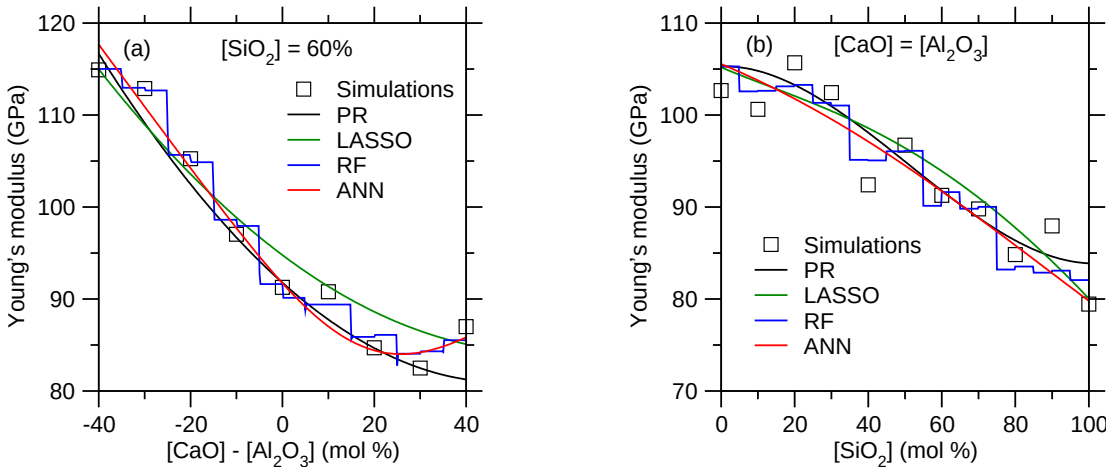

Figure 1. 8 Comparison between the Young's modulus values computed by molecular dynamics simulations and predicted by the polynomial regression (PR), LASSO, random forest (RF), and artificial neural network (ANN) models for the series of compositions (a) $(CaO)_x(Al_2O_3)_{40-x}(SiO_2)_{60}$ and (b) $(CaO)_x(Al_2O_3)_x(SiO_2)_{100-2x}$.

Besides accuracy, it is also desirable for ML-based models to be "simple" (i.e., low complexity) and "interpretable" (i.e., to avoid the use of "black box" models). Unfortunately, a higher level of accuracy often comes at the expense of higher complexity and lower interpretability. Simpler and more interpretable models are usually preferable as (i) simpler models are less likely to overfit small datasets, (ii) simpler models are usually more computationally-efficient, and (iii) more interpretable models are more likely to offer some new insights into the underlying physics governing the relationship between inputs and outputs.

We now discuss the level of complexity/interpretability of the different ML-based models developed herein. The degree of complexity of each of the trained models can be roughly captured by the number of non-zero parameters in PR and LASSO, the number of trees in RF, and the product of the number of inputs, neurons, and parameters in each individual neuron in ANN (i.e., number of weight coefficients and threshold terms to adjust). As presented in Table 1.1, we first note that RF offers a poor balance between accuracy and simplicity (as the number of trees approaches the number of values in the training set). On the other hand, PR and LASSO clearly offer the lowest degree of complexity. The PR and LASSO algorithms also clearly yield the highest level of interpretability thanks to the analytical nature of the inputs/outputs relationship they offer. In details, we note that, LASSO yields a slightly simpler analytic function—with only 8 non-zero terms, vs. 9 non-zero terms for PR, while it also gives a lower level of accuracy. This shows that, by relying on a penalized regression method, LASSO allows us to slightly enhance the level of simplicity of the model. Finally, we note that the increased level of accuracy offered by ANN comes at the cost of higher complexity and lower interpretability, which is a common tradeoff in ML techniques.

We now compare the predictions of the most accurate ML-based model developed herein (i.e., ANN) with the simulated data (i.e., used during the training of the model) and available experimental data (see Fig. 1.9) [13,32–42]. We first note that the experimental data present a higher level of noise as compared to the simulation data. In the present case, these results illustrate the advantage of training ML models based on simulations rather than experimental data. Overall, we observe a good agreement between simulated data, ANN predictions, and experimental data. In contrast, we note that, as mentioned in the Results section, the MM model systematically underestimates E and does not properly capture the non-linear nature of the simulated data.

Combining the results in Fig. 1.8 and Fig. 1.9, we also note that even simple algorithms, e.g., polynomial regression, can capture the non-linearity between composition and Young's modulus and yield some realistic predictions of the Young's modulus values. All the models offer a prediction that is significantly more accurate than that of the MM model. Overall, we find that the ANN model properly captures the non-linear compositional trend of *E* while filtering out the intrinsic noise of the simulation data. These results strongly support the ability of our MD+ML combined method to offer a robust prediction of the stiffness of silicate glasses.

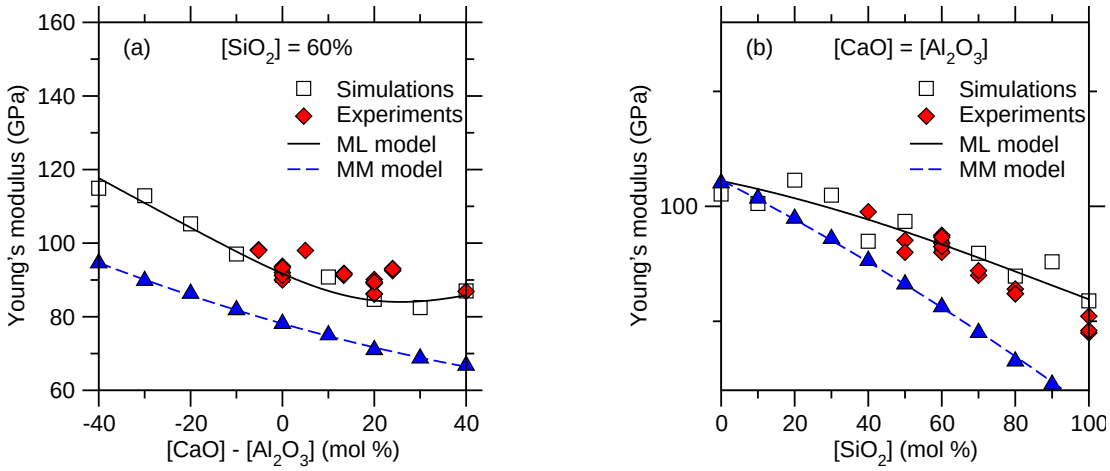

Figure 1. 9 Comparison between the Young's modulus values computed by molecular dynamics simulations, predicted by the artificial neural network model, and predicted by the Makishima-Mackenzie (MM) model for the series of compositions (a) $(CaO)_x(Al_2O_3)_{40-x}(SiO_2)_{60}$ and (b) $(CaO)_x(Al_2O_3)_x(SiO_2)_{100-2x}$. The data are compared with select available experimental data [13,32–42].

Finally, we discuss the advantages of combining ML with high-through MD simulations—rather than directly training ML-based on available experimental data. First, we note

that, although the CAS ternary system may be one of the most studied systems in glass science and engineering, the number of available experimental stiffness data available for this system is fairly limited. Further, most of the data available for this system are clustered in some small regions of the whole compositional domain (namely, pure silica, per-alkaline aluminosilicates, and calcium aluminates glasses). Such clustering of the data is a serious issue as, in turn, available experimental data come with a notable uncertainty—for instance, the Young's modulus of select glasses (at fixed composition) can vary by as much as 20 GPa among different references [21,41]. As such, the combination of a high level of noise and clustering of the data would not allow ML to discriminate the "true" trend of the data from the noise. Finally, we note that conducting MD simulations is obviously faster/cheaper than synthesizing glass samples and measuring their stiffness. In turn, the results presented herein demonstrate that properly conducted MD simulations can offer a quantitative agreement with experimental data and, thereby, offer a desirable alternative to systematic experiments. At the same time, we note some possible limitations for the MD+ML combined method. Since the data-based models are highly sensitive to the nature of data procured, the simulated results have effect on the final models trained from ML algorithms. Which is to say, the gap between simulations and experiments can also cause the gap between predictions from MD+ML models and experimental results. To fill the gaps and limitations of our method, it requires (i) the development of computationally cheap and accurate forcefields to enhance the performance of simulations, and (ii) more experimental data to guide the learning of ML models.

1.6 Conclusions

Overall, these results demonstrate that the combination of high-throughput molecular dynamics simulations and machine learning offers a robust approach to predict the elastic

properties of silicate glasses. Further, our method clearly identifies the optimal level of complexity of each ML-based model, that is, to mitigate the risk of under- or overfitting. Based on these results, we find that the artificial neural network algorithm offers the highest level of accuracy. In contrast, the LASSO algorithm offers a model that features higher simplicity and interpretability—at the expense of a slight decrease in accuracy. The method presented herein is generic and transferable to new properties (e.g., other stiffness metrics) and new systems (e.g., other families of silicate glasses).

1.7 Acknowledgements

This work was supported by the National Science Foundation under Grants No. 1562066, 1762292, and 1826420. M.M.S. acknowledges support from the Independent Research Fund Denmark under Grant No. 8105-00002.

1.8 References

[1] L. Wondraczek, J. C. Mauro, J. Eckert, U. Kühn, J. Horbach, J. Deubener, and T. Rouxel, Advanced Materials **23**, 4578 (2011).

[2] T. Rouxel, in *Challenging Glass: Conference on Architectural and Structural Applications of Glass, Faculty of Architecture, Delft University of Technology, May 2008* (IOS Press, 2008), p. 39.

[3] T. Rouxel, Comptes Rendus Mécanique **334**, 743 (2006).

[4] T. Rouxel, Journal of the American Ceramic Society **90**, 3019 (2007).

[5] J. C. Mauro, C. S. Philip, D. J. Vaughn, and M. S. Pambianchi, International Journal of Applied Glass Science **5**, 2 (2014).

[6] J. C. Mauro and E. D. Zanotto, Int J Appl Glass Sci **5**, 313 (2014).

[7] E. D. Zanotto and F. A. B. Coutinho, Journal of Non-Crystalline Solids **347**, 285 (2004).

[8] A. K. Varshneya, *Fundamentals of Inorganic Glasses* (Academic Press Inc, 1993).

[9] J. C. Mauro, Current Opinion in Solid State and Materials Science **22**, 58 (2018).

[10] H. Liu, T. Du, N. M. A. Krishnan, H. Li, and M. Bauchy, Cement and Concrete Composites (2018).
[11] A. Makishima and J. D. Mackenzie, Journal of Non-Crystalline Solids **12**, 35 (1973).
[12] A. Makishima and J. D. Mackenzie, Journal of Non-Crystalline Solids **17**, 147 (1975).
[13] R. J. Eagan and J. C. Swearekgen, Journal of the American Ceramic Society **61**, 27 (1978).
[14] J. Du, in *Molecular Dynamics Simulations of Disordered Materials: From Network Glasses to Phase-Change Memory Alloys*, edited by C. Massobrio, J. Du, M. Bernasconi, and P. S. Salmon (Springer International Publishing, Cham, 2015), pp. 157–180.
[15] A. Pedone, G. Malavasi, A. N. Cormack, U. Segre, and M. C. Menziani, Chem. Mater. **19**, 3144 (2007).
[16] N. M. Anoop Krishnan, S. Mangalathu, M. M. Smedskjaer, A. Tandia, H. Burton, and M. Bauchy, Journal of Non-Crystalline Solids **487**, 37 (2018).
[17] C. Dreyfus and G. Dreyfus, Journal of Non-Crystalline Solids **318**, 63 (2003).
[18] J. C. Mauro, A. Tandia, K. D. Vargheese, Y. Z. Mauro, and M. M. Smedskjaer, Chem. Mater. **28**, 4267 (2016).
[19] M. C. Onbaşlı, A. Tandia, and J. C. Mauro, Handbook of Materials Modeling 1 (2018).
[20] D. R. Cassar, A. C. P. L. F. de Carvalho, and E. D. Zanotto, Acta Materialia **159**, 249 (2018).
[21] A. I. Priven and O. V. Mazurin, Advanced Materials Research **39–40**, 145 (2008).
[22] A. Ellison and I. A. Cornejo, International Journal of Applied Glass Science **1**, 87 (2010).
[23] S. Plimpton, Journal of Computational Physics **117**, 1 (1995).
[24] M. Bauchy, The Journal of Chemical Physics **141**, 024507 (2014).
[25] M. Bouhadja, N. Jakse, and A. Pasturel, J Chem Phys **138**, 224510 (2013).
[26] C. J. Fennell and J. D. Gezelter, The Journal of Chemical Physics **124**, 234104 (2006).
[27] X. Li, W. Song, K. Yang, N. M. A. Krishnan, B. Wang, M. M. Smedskjaer, J. C. Mauro, G. Sant, M. Balonis, and M. Bauchy, The Journal of Chemical Physics **147**, 074501 (2017).
[28] L. Martínez, R. Andrade, E. G. Birgin, and J. M. Martínez, J Comput Chem **30**, 2157 (2009).
[29] H. Liu, S. Dong, L. Tang, N. M. A. Krishnan, G. Sant, and M. Bauchy, Journal of the Mechanics and Physics of Solids **122**, 555 (2019).

[30] M. Stone, Journal of the Royal Statistical Society. Series B (Methodological) **36**, 111 (1974).

[31] G. C. Cawley and N. L. C. Talbot, J. Mach. Learn. Res. **11**, 2079 (2010).

[32] C. Ecolivet and P. Verdier, Materials Research Bulletin **19**, 227 (1984).

[33] S. Inaba, S. Todaka, Y. Ohta, and K. Morinaga, Journal of the Japan Institute of Metals **64**, 177 (2000).

[34] S. Inaba, S. Oda, and K. Morinaga, Journal of the Japan Institute of Metals **65**, 680 (2001).

[35] C. Weigel, C. Le Losq, R. Vialla, C. Dupas, S. Clément, D. R. Neuville, and B. Rufflé, Journal of Non-Crystalline Solids **447**, 267 (2016).

[36] J. Rocherulle, C. Ecolivet, M. Poulain, P. Verdier, and Y. Laurent, Journal of Non-Crystalline Solids **108**, 187 (1989).

[37] M. Yamane and M. Okuyama, Journal of Non-Crystalline Solids **52**, 217 (1982).

[38] S. Sugimura, S. Inaba, H. Abe, and K. Morinaga, Journal of the Ceramic Society of Japan **110**, 1103 (2002).

[39] T. M. Gross, M. Tomozawa, and A. Koike, Journal of Non-Crystalline Solids **355**, 563 (2009).

[40] I. Yasui and F. Utsuno, in *Computer Aided Innovation of New Materials II*, edited by M. Doyama, J. Kihara, M. Tanaka, and R. Yamamoto (Elsevier, Oxford, 1993), pp. 1539–1544.

[41] L.-G. Hwa, K.-J. Hsieh, and L.-C. Liu, Materials Chemistry and Physics **78**, 105 (2003).

[42] N. P. Bansal and R. H. Doremus, *Handbook of Glass Properties* (Elsevier, 2013).

Chapter 2. Prediction of the Young's Modulus of Silicate Glasses by Topological Constraint Theory

2.1 Abstract

Understanding and predicting the compositional dependence of the stiffness of silicate glasses is key for various technological applications. Here, we a new topological model predicting the Young's modulus of silicate glasses. We show that the Young's modulus is governed by the volumetric density of bond-stretching and bond-bending topological constraints acting in the atomic network. The predicted Young's modulus values offer an excellent agreement with molecular dynamics and experimental data over the entire calcium aluminosilicate ternary domain.

2.2 Introduction

Discovering new glasses with improved mechanical properties is key to address present and future challenges in energy, communication, and infrastructure [1–4]. Among all the mechanical properties that are of interest to glasses, the Young's modulus (E) plays a critical role in the manufacture of glass fibers [5–7]. More generally, the Young's modulus of glasses is an important engineering property for a large range of applications, including flexible substrates and roll-to-roll processing of displays, architectural glazing, ultra-stiff composites, hard discs and surgery equipment, or lightweight construction materials [1,8–10].

Accelerating the discovery of novel glasses with tailored functionalities requires the development of new predictive models that decipher the linkages between glass composition and properties [11]. To this end, several studies have attempted to derive a relationship between glass composition and Young's modulus. Thanks to its elegance and simplicity, the Makishima–

Mackenzie (MM) model may be the most popular model to date [12,13]. This model is based on the idea that the Young's modulus of silicate glasses can be expressed as a linear combination of the dissociation energies of its oxide constituents, normalized by the atomic packing density. Although the predictions offered by the MM model are remarkable accurate considering the simplicity of this model, it is essentially an additive model assuming that the contributions of each oxide to the Young's modulus are proportional to their concentrations. However, the Young's modulus often shows a non-linear dependence on composition, which cannot be captured by purely additive models [14,15]. More generally, the failure of the MM model to properly predict the non-linear relationship between composition and Young's modulus is likely due to the fact that this model does not embed any information about the atomic structure of glasses and the compositional dependence thereof [16].

As an alternative route, topological constraint theory (TCT) offers a promising route to predict the properties of glasses based on the topology of their atomic network [17–24]. TCT reduces complex disordered atomic networks into simpler mechanical trusses, wherein some nodes (the atoms) are connected to each other via some constraints (the chemical bonds). In molecular glasses, such constraints comprise (i) the radial 2-body bond-stretching (BS) constraints that keep the bond lengths fixed around their average values and (ii) the angular 3-body bond-bending (BB) constraints that fix the average values of the interatomic angles. As such, TCT captures the connectivity of the glass network while filtering out second-order structural details that do not significantly affect macroscopic properties. Based on this framework, glasses are classified as flexible, stressed-rigid, or isostatic when the total number of BS and BB constraints per atom n_c is lower, larger, or equal to 3, respectively, which is the number of degrees of freedom per atom.

Within the framework of TCT, glasses can be considered as a network of atoms that are connected to each other via some "small springs" (the interatomic mechanical constraints)—so that the macroscopic stiffness of glasses should be related to the number of interatomic constraints. Indeed, it was shown by Thorpe that the stiffness of model random networks is zero in flexible systems ($n_c < 3$) and then scales with nc in stressed-rigid systems ($n_c > 3$) [25]. A similar relationship was observed in amorphous semiconductors [26–29] and chalcogenide glasses [16]. However, no topological model predicting the stiffness of ionocovalent silicate glasses is available to date.

Here, based on high-throughput molecular dynamics (MD) simulations of calcium aluminosilicate (CAS) glasses, we present a new topological model predicting the compositional dependence of the Young's modulus of silicate glasses. We demonstrate that our topological model offers realistic predictions of Young's modulus values over the entire CAS ternary domain.

2.3 Simulation Methods

To establish our conclusions, we conduct some high-throughput MD simulations of 231 CAS glasses. The chosen compositions homogeneously cover the CAS ternary domain, with 5% increments in the mol% concentration of the CaO, Al_2O_3, and SiO_2 oxide constituents. Note that some of these systems would likely not exhibit satisfactory glass-forming ability. All the simulations are conducted using the Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) package [30]. Each system comprises around 3000 atoms. Since the quality of MD simulations mostly relies on that of the underlying forcefield, we adopt here the interatomic potential parametrized by Jakse [31,32], since this potential has been shown to offer a realistic description of the mechanical properties of CAS glasses [15,33–35]. A cutoff of 8.0 Å is used for the short-range interactions. The Coulombic interactions are calculated by adopting the Fennell

damped shifted force model with a damping parameter of 0.25 $Å^{-1}$ and a global cutoff of 8.0 Å [36]. The integration timestep is 1.0 fs.

The CAS glasses are prepared by quenching liquids, as described in the following [37]. First, some atoms are randomly placed in a cubic box using the PACKMOL package while using a distance cutoff of 2.0 Å between each atom to avoid any unrealistic overlap [38]. These initial configurations are then subjected to an energy minimization, followed by some 100 ps relaxations in the canonical (*NVT*) and isothermal-isobaric (*NPT*) ensembles at 300 K, sequentially. These samples are then melted at 3000 K for 100 ps in the *NVT* and, subsequently, *NPT* ensemble (at zero pressure) to ensure the loss of the memory of the initial configurations and to equilibrate the systems. Next, these liquids are cooled from 3000 to 300 K in the *NPT* ensemble at zero pressure with a cooling rate of 1 K/ps. The obtained glass samples are further relaxed at 300 K for 100 ps in the *NPT* ensemble before the stiffness computation. Note that this quenching procedure is slightly adjusted for select compositions. First, a higher initial melting temperature of 5000 K is used for the samples wherein the SiO_2 concentration is larger or equal to 95 mol%—since these glasses exhibit high glass transition temperatures. Second, a faster cooling rate of 100 K/ps is used for the samples wherein the CaO concentration is larger or equal to 90 mol%. Indeed, although the cooling rate can affect the glass stiffness, the use of a higher cooling rate here is necessary as these systems would otherwise tend to crystallize with a cooling rate of 1 K/ps. Once formed and equilibrated, the glasses are subjected to a series of 6 deformations (i.e., 3 axial and 3 shear deformations along the 3 axes). Their stiffness tensor (and Young's modulus) is computed from the curvature of the potential energy (see Refs. [15,33,39,40] for more details). Based on Ref. [15], the computed Young's modulus values are rescaled by a constant factor (0.86) to enhance the overall agreement with experimental data.

2.4 Topological Model of Young's Modulus

Our topological model is inspired by that developed by Smedskjaer, Mauro, and Yue, wherein hardness is expressed as a linear function of the number of constraints per atom n_c [22,41–43]. Here, since the Young's modulus has the dimension of an energy per unit of volume, we postulate that E can be expressed in terms of the volumic density of the energy created by each constraint. A similar approach was used to refine the original Smedskjaer model to predict hardness [44]. Further, we postulate that the BS and BB constraints do not contribute with equal weight to increasing the Young's modulus, which arises from the fact that BS and BB constraints exhibit different free energies and that different types of constraints may be activated under different loading conditions [23,45,46]. Based on these considerations, we propose the following model:

$$E = \varepsilon_{\mathrm{BS}} n_{\mathrm{BS}} + \varepsilon_{\mathrm{BB}} n_{\mathrm{BB}} \qquad \textbf{(Eq. 1)}$$

where n_{BS} and n_{BB} are the volumetric density of BS and BB constraints, respectively, and $\varepsilon_{\mathrm{BS}}$ and $\varepsilon_{\mathrm{BB}}$ are the typical energies of BS and BB constraints, respectively. Note that this model assumes that a fictitious glass comprising no BS and BB constraints would have a zero Young's modulus.

2.5 Results and Discussion

2.5.1 Constraints enumeration

To assess the validity of our topological model, we first enumerate the number of BS and BB constraints in CAS glasses as a function of composition. In fully-connected covalent glasses, the number of BS constraints created by a given atom is given by $r/2$, where r is the coordination number—where the factor 2 arises from the fact that each BS constraint is shared by two atoms [17]. In turn, the number of BB constraints is usually given by $2r - 3$, which corresponds

to the number of independent angles that need to be fixed to define the angular environment of the atom [17]. However, due to the existence of ionic non-directional bonds, this counting scheme does not always apply to ionocovalent silicate glasses [46]. As such, to avoid relying on any guesses in the enumeration of the constraints, we analyze the structure of the simulated glasses to directly extract the number of BS and BB constraints [19,46,47]. To this end, we compute the coordination number of each atom. We then identify the different types of O species present in the network, namely, (i) bridging-oxygen (BO), i.e., connected to 2 network formers (Si or Al), (ii) non-bridging oxygen (NBO), i.e., connected to only 1 network formers, (iii) "tricluster" oxygen (TO), i.e., connected to 3 network formers [33,48], and (iv) "free oxygen" (FO), i.e., connected to 0 network formers (i.e., only connected to Ca atoms) [33].

Table 2.1 summarizes the average number of BS and BB constraints created by each atomic species. In details, we find that, as expected, Si atoms systematically create 4 BS and 5 BB constraints with their 4 O neighbors—note that, for simplicity, the BS constraints are here fully attributed to the cations. Although a small fraction of over-coordinated Al atoms is found in Al-rich glasses, most of them create 4 BS and BB constraints with their 4 O neighbors. Due to the ionic nature of Ca–O bonds, the constraints enumeration is trickier for Ca atoms. First, these atoms do not form any well-defined angular environment and, as such, do not create any BB constraints [47]. Second, a statistical analysis between the Young's modulus and the partial coordination number of Ca atoms reveals that Ca atoms only create BS constraints with their surrounding NBO and FO atoms [47]. In turn, Ca atoms do not create any constraints with the surrounding BO and TO atoms. This can be understood from the fact that the charge of BO and TO atoms is already fully compensated by those of their surrounding Si and Al neighbors, so that their interaction with Ca atoms is weaker than that between Ca and NBO or FO atoms. Finally, we

find that, as expected, BO create 1 BB constraint, while TO atoms create 3 BB constraints to define their trigonal environment. In contrast, due to the ionic nature of Ca–O bonds, NBO and FO atoms do not create any BB constraint. These inputs then serve to compute the volumetric densities of BS and BB constraints (n_{BS} and n_{BB} in Eq. 1).

Table 2. 1 Summary of the average number of bond-stretching (BS) and bond-bending (BB) constraints created by each atomic species in $(CaO)_x(Al_2O_3)_y(SiO_2)_{1-x-y}$ glasses. Note the BS constraints are here fully attributed to the cations. The quantities $r_{\mathrm{Ca–NBO}}$ and $r_{\mathrm{Ca–FO}}$ refer to the average number of non-bridging oxygen (NBO) and free oxygen (FO) atoms around each Ca atom.

Species	**Fraction**	**BS**	**BB**
Si	$1 - x - y$	4	5
Al	$2y$	4	5
Ca	x	$r_{\mathrm{Ca–NBO}} + r_{\mathrm{Ca–FO}}$	0
O	2 – x + y	-	-
FO		-	0
NBO		-	0
BO		-	1
TO		-	3

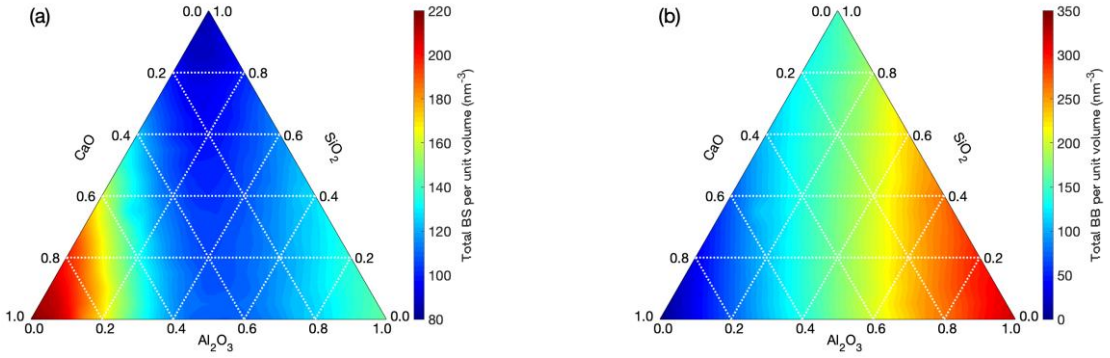

Figure 2. 1 Ternary diagram showing the volumic density of (a) bond-stretching (BS) and (b) bond-bending (BB) constraints as a function of composition in the $CaO–Al_2O_3–SiO_2$ glass system.

Figure 2.1 shows the volumetric densities of BS and BB constraints as a function of composition in the CAS ternary system. Overall, we find that the densities of BS and BB constraints primarily depends on the [CaO] – [Al_2O_3] molar difference. In details, we find that the density of BS constraints is minimum when [CaO] = [Al_2O_3] and increases in the Ca- and Al-rich domains. This arises from the fact that both of these domains exhibit a high average coordination number—i.e., due to the presence of 6-fold coordinated Ca atoms in Ca-rich glasses and TO atoms in Al-rich glasses [15]. On the other hand, the density of BB constraints presents a significantly different compositional dependence as it monotonically with increases [CaO] – [Al_2O_3] molar difference. This arises from the fact that Ca atoms do not create any BB constraints, whereas TO atoms contribute to increasing the number of BB constraints in Al-rich glasses.

2.5.2 Prediction of Young's modulus

We then focus on the compositional dependence of the Young's modulus (*E*) values computed by MD (see Fig. 2.2a). Overall, we observe the existence of two main trends: (i) *E* tends to increase with increasing Al_2O_3 concentration and (ii) *E* tends to increase with increasing CaO

concentration. However, we find that the compositional dependence of *E* is non-monotonic and that CaO and Al_2O_3 exhibit some coupled effects. For example, we find that *E* increases with increases CaO concentration when [Al_2O_3] = 0 mol%, whereas *E* decreases with increasing CaO concentration when [Al_2O_3] > 40 mol%. This highlights the fact that E exhibits a non-linear dependence on composition—so that additive models are unlikely to offer good predictions for this system.

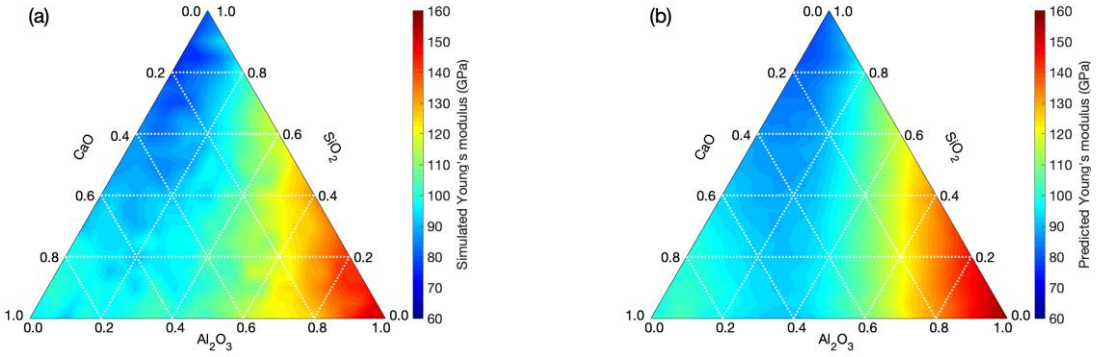

Figure 2. 2 Ternary diagram showing the Young's modulus values (a) computed by high-throughput molecular dynamics and (b) predicted by our topological model as a function of composition in the CaO–Al_2O_3–SiO_2 glass system.

We now assess the validity of our topological model (Eq. 1). To this end, we conduct a polynomial regression using as inputs the volumetric densities of BS and BB constraints (n_{BS} and n_{BB}) shown in Fig. 2.1 and as output the simulated *E* values shows in Fig. 2.2a. This allows us to determine the typical energies of BS and BB constraints ($\varepsilon_{\mathrm{BS}}$ and $\varepsilon_{\mathrm{BB}}$) as fitting parameters. We find $\varepsilon_{\mathrm{BS}} = 2.82$ eV and $\varepsilon_{\mathrm{BB}} = 1.78$ eV. These values have the same order of magnitude as typical interatomic bond energy in silicate glasses. As expected, we find that $\varepsilon_{\mathrm{BS}} > \varepsilon_{\mathrm{BB}}$, in agreement with the fact that the free energy of BS constraints is larger than that of BB constraints [23,46]. A

more detailed polynomial regression using each type of constraints as independent inputs does not significantly improve the quality of the fit and further suggests that all the BS (and BB) constraints contribute to increasing the Young's modulus with a fairly similar energy "weight" ε.

Figure 2.2b shows the *E* values predicted by Eq. 1. Overall, we find that the E values predicted by Eq. 1 agree well with the simulated values (see also Fig. 2.3a), although our topological model tends to slightly unpredict the Young's modulus of select calcium aluminate glasses on the $CaO–Al_2O_3$ joint. Although the simulated values are here used to parameterize the ε_{BS} and ε_{BB} coefficients in Eq.1, it is nevertheless striking that the complex compositional dependence of the Young's modulus of CAS glasses can be well reproduced with only two fitting parameters. We also note that our topological model does not keep the memory of the "noise" present in the MD data, which suggests that the model is not overfitted.

As a final validation of our model, Figs. 2.3a and 2.3b show a comparison between the Young's modulus predictions from our topological model (Eq. 1), the simulation data, and available experimental data [49–60] for two joints, viz., $[SiO_2] = 60\%$ and $[CaO] = [Al_2O_3]$. These two series specifically aim to investigate (i) the effect of the degree of polymerization of the network (i.e., fraction of non-bridging oxygen) and (ii) the effect of network-forming atoms (i.e., Si vs. Al) at constant degree of depolymerization (i.e., in fully charge-compensated glasses). Overall, we observe a good agreement between simulated data, topological predictions, and experimental data. In contrast, we find that the MM model systematically underestimates *E* and does not properly capture the non-linear nature of the Young's modulus data. Overall, these results strongly support the ability of our new topological model to offer reliable predictions of Young's modulus values over the entire CAS ternary domain.

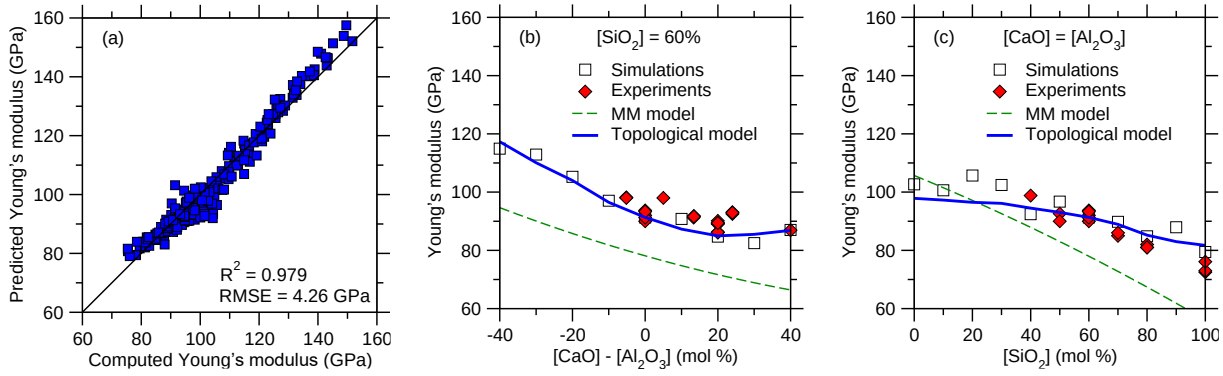

Figure 2. 3 (a) Comparison between the Young's modulus values predicted by our topological model (Eq. 1) and computed by molecular dynamics simulations. We obtain a coefficient of determination R^2 = 0.979 and a root mean squared error (RMSE) of 4.26 GPa. Comparison between the Young's modulus values computed by molecular dynamics simulations, predicted by our topological model, and predicted by the Makishima-Mackenzie (MM) model for the series of compositions (a) $(CaO)_x(Al_2O_3)_{40-x}(SiO_2)_{60}$ and (b) $(CaO)_x(Al_2O_3)_x(SiO_2)_{100-2x}$. The data are compared with select available experimental data [49–60].

2.6 Conclusions

In summary, the results presented herein demonstrate that the Young's modulus of CAS glasses can accurately predict based on the volumetric densities of BS and BB topological constraints. As such, topological constraint theory offers a powerful framework to accelerate the design of new glass formulations with tailored stiffness. The atomistic origin of the energy coefficients ε_{BS} and ε_{BB} and their values for distinct glass families should be investigated in future work.

2.7 Acknowledgements

This work was supported by the National Science Foundation under Grants No. 1562066, 1762292, and 1826420. M.M.S. acknowledges support from the Independent Research Fund Denmark under Grant No. 8105-00002.

2.8 References

[1] L. Wondraczek, J. C. Mauro, J. Eckert, U. Kühn, J. Horbach, J. Deubener, and T. Rouxel, Advanced Materials **23**, 4578 (2011).

[2] J. C. Mauro and E. D. Zanotto, Int J Appl Glass Sci **5**, 313 (2014).

[3] J. C. Mauro, Front. Mater. **1**, 20 (2014).

[4] J. C. Mauro, C. S. Philip, D. J. Vaughn, and M. S. Pambianchi, International Journal of Applied Glass Science **5**, 2 (2014).

[5] D. A. Krohn and A. R. Cooper, Journal of the American Ceramic Society **52**, 661 (1969).

[6] A. R. Cooper and D. A. Krohn, Journal of the American Ceramic Society **52**, 665 (1969).

[7] Z. Tang, N. P. Lower, P. K. Gupta, C. R. Kurkjian, and R. K. Brow, Journal of Non-Crystalline Solids **428**, 98 (2015).

[8] T. Rouxel, in *Challenging Glass: Conference on Architectural and Structural Applications of Glass, Faculty of Architecture, Delft University of Technology, May 2008* (IOS Press, 2008), p. 39.

[9] T. Rouxel, Comptes Rendus Mécanique **334**, 743 (2006).

[10] T. Rouxel, Journal of the American Ceramic Society **90**, 3019 (2007).

[11] J. C. Mauro, Current Opinion in Solid State and Materials Science **22**, 58 (2018).

[12] A. Makishima and J. D. Mackenzie, Journal of Non-Crystalline Solids **12**, 35 (1973).

[13] A. Makishima and J. D. Mackenzie, Journal of Non-Crystalline Solids **17**, 147 (1975).

[14] A. I. Priven, in *Glasses and the Glass Transition* (John Wiley & Sons, Ltd, 2011), pp. 255–309.

[15] K. Yang, X. Xu, B. Yang, B. Cook, H. Ramos, and M. Bauchy, ArXiv:1901.09323 [Cond-Mat] (2019).

[16] M. Plucinski and J. W. Zwanziger, Journal of Non-Crystalline Solids **429**, 20 (2015).

[17] J. C. Phillips, J. Non-Cryst. Solids **34**, 153 (1979).

[18] J. C. Phillips, Journal of Non-Crystalline Solids **43**, 37 (1981).

[19] M. Bauchy, Computational Materials Science **159**, 95 (2019).

[20] J. C. Mauro, Am. Ceram. Soc. Bull. **90**, 31 (2011).

[21] M. Bauchy, American Ceramic Society Bulletin **91**, 34 (2012).

[22] M. M. Smedskjaer, J. C. Mauro, and Y. Yue, Phys. Rev. Lett. **105**, 115503 (2010).

[23] P. K. Gupta and J. C. Mauro, The Journal of Chemical Physics **130**, 094503 (2009).

[24] J. C. Mauro, P. K. Gupta, and R. J. Loucks, The Journal of Chemical Physics **130**, 234503 (2009).

[25] M. F. Thorpe, J. Non-Cryst. Solids **57**, 355 (1983).

[26] G. Bhattarai, S. Dhungana, B. J. Nordell, A. N. Caruso, M. M. Paquette, W. A. Lanford, and S. W. King, Phys. Rev. Materials **2**, 055602 (2018).

[27] Q. Su, S. King, L. Li, T. Wang, J. Gigax, L. Shao, W. A. Lanford, and M. Nastasi, Scripta Materialia **146**, 316 (2018).

[28] S. W. King, L. Ross, and W. A. Lanford, Journal of Non-Crystalline Solids **499**, 252 (2018).

[29] M. M. Paquette, B. J. Nordell, A. N. Caruso, M. Sato, H. Fujiwara, and S. W. King, MRS Bulletin **42**, 39 (2017).

[30] S. Plimpton, Journal of Computational Physics **117**, 1 (1995).

[31] M. Bouhadja, N. Jakse, and A. Pasturel, J Chem Phys **138**, 224510 (2013).

[32] M. Bouhadja, N. Jakse, and A. Pasturel, The Journal of Chemical Physics **140**, 234507 (2014).

[33] M. Bauchy, The Journal of Chemical Physics **141**, 024507 (2014).

[34] B. Wang, Y. Yu, Y. J. Lee, and M. Bauchy, Front. Mater. **2**, 11 (2015).

[35] Y. Yu, B. Wang, Y. J. Lee, and M. Bauchy, in *Symposium UU – Structure-Property Relations in Amorphous Solids* (2015).

[36] C. J. Fennell and J. D. Gezelter, The Journal of Chemical Physics **124**, 234104 (2006).

[37] X. Li, W. Song, K. Yang, N. M. A. Krishnan, B. Wang, M. M. Smedskjaer, J. C. Mauro, G. Sant, M. Balonis, and M. Bauchy, The Journal of Chemical Physics **147**, 074501 (2017).

[38] L. Martínez, R. Andrade, E. G. Birgin, and J. M. Martínez, J Comput Chem **30**, 2157 (2009).

[39] H. Liu, S. Dong, L. Tang, N. M. A. Krishnan, G. Sant, and M. Bauchy, Journal of the Mechanics and Physics of Solids **122**, 555 (2019).

[40] A. Pedone, G. Malavasi, A. N. Cormack, U. Segre, and M. C. Menziani, Chem. Mater. **19**, 3144 (2007).

[41] M. M. Smedskjaer, Front. Mater. **1**, 23 (2014).

[42] M. M. Smedskjaer, J. C. Mauro, S. Sen, and Y. Yue, Chem. Mat. **22**, 5358 (2010).

[43] J. C. Mauro, A. Tandia, K. D. Vargheese, Y. Z. Mauro, and M. M. Smedskjaer, Chem. Mater. **28**, 4267 (2016).

[44] Q. Zheng, Y. Yue, and J. C. Mauro, Appl. Phys. Lett. **111**, 011907 (2017).

[45] M. Bauchy, M. J. A. Qomi, C. Bichara, F.-J. Ulm, and R. J.-M. Pellenq, Phys. Rev. Lett. **114**, 125502 (2015).

[46] M. Bauchy and M. Micoulaut, Journal of Non-Crystalline Solids **357**, 2530 (2011).

[47] M. Bauchy, M. J. Abdolhosseini Qomi, C. Bichara, F.-J. Ulm, and R. J.-M. Pellenq, J. Phys. Chem. C **118**, 12485 (2014).

[48] J. A. Tossell and J. Horbach, J. Phys. Chem. B **109**, 1794 (2005).

[49] R. J. Eagan and J. C. Swearekgen, Journal of the American Ceramic Society **61**, 27 (1978).

[50] C. Ecolivet and P. Verdier, Materials Research Bulletin **19**, 227 (1984).

[51] S. Inaba, S. Todaka, Y. Ohta, and K. Morinaga, Journal of the Japan Institute of Metals **64**, 177 (2000).

[52] S. Inaba, S. Oda, and K. Morinaga, Journal of the Japan Institute of Metals **65**, 680 (2001).

[53] C. Weigel, C. Le Losq, R. Vialla, C. Dupas, S. Clément, D. R. Neuville, and B. Rufflé, Journal of Non-Crystalline Solids **447**, 267 (2016).

[54] J. Rocherulle, C. Ecolivet, M. Poulain, P. Verdier, and Y. Laurent, Journal of Non-Crystalline Solids **108**, 187 (1989).

[55] M. Yamane and M. Okuyama, Journal of Non-Crystalline Solids **52**, 217 (1982).

[56] S. Sugimura, S. Inaba, H. Abe, and K. Morinaga, Journal of the Ceramic Society of Japan **110**, 1103 (2002).

[57] T. M. Gross, M. Tomozawa, and A. Koike, Journal of Non-Crystalline Solids **355**, 563 (2009).

[58] I. Yasui and F. Utsuno, in *Computer Aided Innovation of New Materials II*, edited by M. Doyama, J. Kihara, M. Tanaka, and R. Yamamoto (Elsevier, Oxford, 1993), pp. 1539–1544.

[59] L.-G. Hwa, K.-J. Hsieh, and L.-C. Liu, Materials Chemistry and Physics **78**, 105 (2003).

[60] N. P. Bansal and R. H. Doremus, *Handbook of Glass Properties* (Elsevier, 2013).

Chapter 3. Analytical Model of the Network Topology and Rigidity of Calcium Aluminosilicate Glasses

3.1 Abstract

Topological constraint theory (TCT) has enabled the prediction of various properties of oxide glasses as a function of their composition and structure. However, the robust application of TCT relies on accurate knowledge of the network structure and topology. Here, based on classical molecular dynamics simulations, we derive a fully analytical model describing the topology of the calcium aluminosilicate (CAS) ternary system. This model yields the state of rigidity (flexible, isostatic, or stressed-rigid) of CAS systems as a function of composition and temperature. These results reveal the existence of correlations between network topology and glass-forming ability. This study suggests that glass-forming ability is encoded in the network topology of the liquid state rather than that of the glassy state.

3.2 Introduction

The calcium aluminosilicate (CAS) system is an archetypical model for alkali-free glasses used in display applications [1] and cementitious materials (e.g., cement, fly ash, or slag) [2]. The structure and properties of the CAS ternary system have been extensively studied from experiments [3–6], atomistic simulations [7–10], and machine learning [11–13]. Indeed, this ternary system offers an ideal model to investigate the effects of polymerization (i.e., by varying the Ca/Al ratio) and of the network-forming skeleton (i.e., by varying the Al/Si ratio at fixed Ca/Al).

Since the stoichiometry of CAS and other oxide glasses can be continuously adjusted, there exists a largely untapped opportunity to discover new glass compositions featuring desirable

properties and functionalities [14,15]. However, the vastness of the accessible compositional space renders traditional material discovery approaches (e.g., Edisonian "trial-and-error") inefficient [16]. In that regard, composition-property predictive models can facilitate and rationalize the search for new glasses by targeting searches toward promising compositional domain [17,18].

To this end, topological constraint theory (TCT) has been a key enabler to develop predictive models that relate the composition and structure of glasses to their properties [19,20]. Various TCT-based models have been proposed over the past decades to predict glass-forming ability, glass transition temperature, liquid fragility, hardness, stiffness, dissolution rate, etc. [19,21–28]. The success of TCT is based on the fact that many macroscopic properties of disordered materials primarily depend on the topology of the atomic structure, while other structural details only have a second-order effect [29]. As such, TCT reduces complex disordered atomic networks into simpler structural trusses [22], wherein some nodes (the atoms) are connected to each other by some topological constraints (the chemical bonds). In structural glasses, topological constraints comprise of the radial two-body bond-stretching (BS) and angular three-body bond-bending (BB) constraints. The number of constraints per atom (n_c) then offers a simple, reduced-dimensionality metric that is often correlated with macroscopic properties [22].

Importantly, predictions from TCT critically rely on an accurate knowledge of the glass structure and connectivity, which is the key to enumerate the number of BS and BB constraints created by each type of atoms in the glass network [30]. This is a challenge as the local structure of glasses (and hence, the number of constraints) changes as a function of composition. For instance, CAS glasses exhibit several structural complexities, e.g., over-coordinated Al atoms [5,31–33], tricluster oxygen units [34], free oxygen species [35,36], varying Ca

environments [37], etc. All of these features impact the constraints enumeration and, hence, should be accounted for in robust topological models. Although such information can be accessed by molecular dynamics (MD) simulations of one composition at a time [30], it is not practical to systematically conduct MD simulations over large compositional domains considering their high computational cost. In addition, discrete models (e.g., relying on a finite, discrete number of MD simulations) are not differentiable, that is, they do not allow for the computation of the derivative of the number of constraints per atom with respects to composition [38]. This prevents the use of gradient-based "inverse design" optimization methods (e.g., to pinpoint glasses with minimum, maximum, or tailored rigidity).

Here, to address these challenges, we present a fully analytical model describing the network topology of CAS glasses over the entire CAS ternary domain. This model is informed and validated by a series of classical MD simulations, but, importantly, offers a pathway to continuously predict the properties of CAS glasses as a function of their compositions without the need for any systematic MD simulation.

This paper is organized as follows. Section II describes the MD simulations used in this study. We then introduce our analytical topological model for CAS glasses in Sec. III. In Sec. IV, we discuss the obtained rigidity diagrams of the CAS ternary system. Finally, we establish some conclusions in Sec. V.

3.3 Methods

To establish our analytical model, we first conduct MD simulations of 231 CAS glasses using the Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) package [13,15,16]. The chosen compositions homogeneously cover the entire CAS domain, with 5%

increments in the mol% concentration of the CaO, Al_2O_3 and SiO_2 oxide constituents. Note that some of these CAS glasses do not exhibit satisfactory glass-forming ability in practice, but they all can be generated by MD due to the use of a high cooling rate. Each glass sample comprises around 3000 atoms. Here, we adopt the interatomic potential parametrized by Jakse [8], which has been reported to offer good results in agreement with experiment data [34,40,41] and applied for several previous studies [11,42]. A cutoff of 8.0 Å is used for the short-range interactions. The Coulombic interactions are calculated by adopting the Fennell damped shifted force field model with damping parameter of 0.25 $Å^{-1}$ and a global cutoff of 8.0 Å [43]. We keep the integration timestep as fixed as 1.0 fs.

We generate the CAS glass samples using the conventional melt-quench method, as described in the following [44]. First, atoms are randomly placed in a cubic box using PACKMOL with a distance cutoff of 2.0 Å between each pair of atoms to avoid any unrealistic overlap [45]. These initial configurations are then subjected to an energy minimization process, followed by 100 ps relaxations in the canonical (*NVT*) and isothermal-isobaric (*NPT*) ensembles at 300 K, sequentially. These samples are then fully melted at 3000 K for 100 ps in the *NVT* and, subsequently, *NPT* ensemble (at zero pressure) to ensure the loss of the memory of the initial configurations and to equilibrate the systems. Next, these liquids are cooled from 3000 K to 300 K in the *NPT* ensemble at zero pressure with a cooling rate of 1 K/ps. The obtained glass samples are further relaxed at 300 K for 100 ps in the *NPT* ensemble.

Note that this quenching procedure is slightly adjusted for the following select compositions: (1) a higher initial melting temperature of 5000 K is used for the samples wherein the SiO_2 concentration is larger or equal to 95 mol%—since these glasses exhibit high glass

transition temperatures, and (2) a faster cooling rate of 100 K/ps is used for the samples wherein the CaO concentration is larger or equal to 90 mol%. Indeed, although the cooling rate can affect the glass structure, the use of a higher cooling rate here is necessary as these systems would otherwise tend to crystallize with a cooling rate of 1 K/ps. The coordination number of each atom is computed by enumerating the number of neighbors present in its first coordination shell—wherein the radius cutoff is defined as the minimum after the first peak of the partial pair distribution function (i.e., 2.00, 2.35, and 3.05 Å for Si—O, Al—O, and Ca—O, respectively).

3.4 Topological model of calcium aluminosilicate glasses

3.4.1 Polymerization and depolymerization of the aluminosilicate network

3.4.1.a Effect of Ca and Al atoms on topology

To establish our topological model of CAS glasses, we take as a reference the structure of glassy silica (SiO_2), wherein all the Si and O atoms are 4- and 2-fold coordinated, respectively—that is, all the O act as bridging oxygen (BO) atoms bonded to two network formers [14]. Starting from this reference structure, we then describe the competitive effects of Ca and Al atoms: (i) each Ca atom consumes 2 BOs and, in turn, creates 2 non-bridging oxygen (NBO) atoms [46], whereas, in contrast, (ii) each Al atom consumes 1 NBO and creates 1 BO [14]. These well-known effects arise from the following mechanisms. On one hand, when added to pure SiO_2, Ca^{2+} cations act as network modifiers as they tend to depolymerize the atomic network by breaking some Si–O–Si inter-tetrahedral joints and, in turn, charge-compensating pairs of negatively-charged NBOs [46]. On the other hand, starting from a calcium silicate glass, newly-added Al atoms act as network formers and tend to repolymerize the network by using available Ca^{2+} cations to charge-compensate negatively-charged 4-fold coordinated AlO_4 units [47–50]. This effectively increases the network connectivity since the Ca cations that are used as charge compensators do not create

any NBO any longer. Based on this model, the number of BO (N_{BO}) and NBO (N_{NBO}) are expressed as:

$$N_{\mathrm{BO}} = N_{\mathrm{O}} - 2 \times N_{\mathrm{Ca}} + N_{\mathrm{Al}} \quad \text{(Eq. 1)}$$

$$N_{\mathrm{NBO}} = 2 \times N_{\mathrm{Ca}} - N_{\mathrm{Al}} \quad \text{(Eq. 2)}$$

where N_{BO}, N_{NBO}, N_{O}, N_{Ca}, and N_{Al} are the number of BO, NBO, total oxygen, calcium, and aluminum atoms, respectively. Note that these equations are not valid for all CAS compositions. In detail, mechanism (i) can hold until the added Ca atoms exhaust all the BOs present in the network, that is, for $N_{\mathrm{BO}} \geq 0$, while mechanism (ii) remains possible until the added Al atoms consume all the NBOs present in the network, that is, for $N_{\mathrm{NBO}} \geq 0$. In the following sections, we discuss the cases of the Ca-rich regime (i.e., when mechanism (i) is no longer possible) and Al-rich regime (i.e., when mechanism (ii) breaks down).

3.4.1.b Ca-rich regime: formation of free oxygen atoms

We then focus on the Ca-rich regime (i.e., wherein N_{BO} would become negative assuming that Eq. 1 would continue to hold) and discuss the effect of Ca atoms on the network topology in this compositional domain. In this domain, the network becomes fully depolymerized, that is, there is no remaining BO and all Si and Al polytopes are isolated from each other. In this regime, our MD simulations suggest that the excess of Ca atoms results in the formation of free oxygen (FO) atoms [Fig. 3.2(a)], that is, O atoms that are not connected to any Si or Al network former [42]. This echoes previous experimental findings that also suggested the existence of FOs [35,36,51].

Based on this observation, we describe the number of FO atoms according to the following model. First, we assume that, in this regime, all the Si and Al atoms remain 4-fold coordinated and, hence, they are each surrounded by 4 NBOs. The number of NBOs is then given by:

$$N_{\mathrm{NBO}} = 4 \times (N_{\mathrm{Si}} + N_{\mathrm{Al}}) \quad \text{(Eq. 3)}$$

We then assume that all the remaining O atoms act as FOs, whose number is given by:

$$N_{\mathrm{FO}} = N_{\mathrm{O}} - N_{\mathrm{NBO}} \quad \text{(Eq. 4)}$$

3.4.1.c Al-rich regime: formation of 5-fold aluminum and tricluster oxygen atoms

Next, we focus on the Al-rich regime (i.e., $2N_{\mathrm{Ca}} < N_{\mathrm{Al}}$, wherein N_{NBO} would become negative assuming that Eq. 2 would continue to hold) and discuss how additional Al^{3+} cations impact the atomic structure of CAS glasses. In this domain, the glass compositions are peraluminous, that is, they exhibit an excess number of Al atoms as compared to the ones that are needed to charge-compensate all the calcium atoms in the glass (i.e., $N_{\mathrm{Al}} = 2N_{\mathrm{Ca}}$) [46]. In this regime, all the AlO_4 tetrahedral units cannot be charge-compensated any longer due to the deficit of Ca cations. From this point, two possible mechanisms have been suggested to occur: (i) some Al atoms become overcoordinated (i.e., with a coordination number larger than 4) and (ii) some 3-fold coordinated triclusters oxygen (TO) atoms (i.e., O atoms that are connected to 3 Si or Al network formers) tend to form [34,52,53]. We now discuss these two behaviors.

First, previous experiments [5,31–33] and simulations [11,42] have suggested that some excess Al atoms become overcoordinated, that is, they become 5- or 6-fold coordinated. Neuville *et al.* reported the existence of a small fraction (< 3%) of 6-fold coordinated aluminum (Al^{VI}) at very high Al/Ca ratios, in agreement with our MD simulations [11]. However, here, considering that independently predicting the fractions of both 5- and 6-fold coordinated Al atoms would be ill-defined based on the present assumptions, we assume that the fraction of 6-fold coordinated Al is small enough to be neglected—that is, we assume that, in this regime, all the overcoordinated Al atoms are 5-fold coordinated.

Figure 3. 1 Predicted (from the analytical model) and computed (from molecular dynamics simulations) numbers of 5-fold Al atoms per formula unit of glasses as a function of composition.

shows the composition dependence of the number of predicted and computed 5-fold Al atoms per formula unit of glass for $(CaO)_x(Al_2O_3)_y(SiO_2)_{1-x-y}$, as a function of $[Al_2O_3] - [CaO]$ (since this metric is found to be the most influential overall). As expected, we find that the number of overcoordinated Al remains nearly zero as long as there is no deficit of Ca cations (i.e., $2N_{\mathrm{Ca}} > N_{\mathrm{Al}}$). However, even in this regime, we nevertheless note that our MD simulations suggest the existence of a small fraction (less than 5%) of 5-fold coordinated Al, which echoes previous experiments [5,54]. This small fraction is neglected from our model thereafter. We then find that the computed number of 5-fold coordinated Al atoms scales fairly linearly with the number of excess Al atoms (i.e., $N_{\mathrm{Al}} - 2N_{\mathrm{Ca}}$), which suggests that the number of overcoordinated Al is proportional to the number of excess Al atoms. Based on this observation, we model the number of 5-fold coordinated Al atoms ($N_{\mathrm{Al^V}}$) as:

$$N_{\mathrm{Al^V}} = \alpha(N_{\mathrm{Al}} - 2 \times N_{\mathrm{Ca}}) \qquad \text{(Eq. 5)}$$

where α is an empirical factor that captures the fraction of excess Al that eventually becomes over-coordinated. Here, based on our MD results, we assume that a quarter of the excess Al atoms become overcoordinated (i.e., $\alpha = 1/4$), which yields a good match between MD simulations and the present analytical model (see Figure 3. 1). This specific value of α here is chosen based on the fact that, even though it slightly overestimates the number of 5-fold coordinated Al atoms, it partially compensates for the fact that our model is neglecting 6-fold coordinated Al atoms (see below).

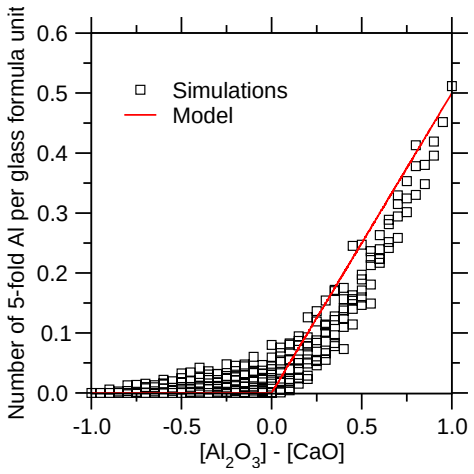

Figure 3. 1 Predicted (from the analytical model) and computed (from molecular dynamics simulations) numbers of 5-fold Al atoms per formula unit of glasses as a function of composition. Note that several CAS glass compositions are associated with the same $[Al_2O_3]$ – [CaO] value.

The fact that the number of 5-fold coordinated Al atoms is lower than the number of excess Al atoms (i.e., that are not charge-compensated by Ca atoms) suggests that there is another charge compensation mechanism at play in the Al-rich regime, namely, through the formation of TO atoms. Indeed, our MD simulations show the existence of such TO species, which echoes previous experiments and simulations [34,52,53]. We model the number of TO atoms as follows.

To determine the fraction of TO atoms in the glass, we first calculate the total excess negative charge of all 4- and 5-fold coordinated Al atoms that are not charge-compensated by Ca atoms and subsequently determine the number of TO atoms that is needed to counterbalance the excess negative charge. We first calculate the excess negative charge of AlO_4 and AlO_5 units by assuming that all the O atoms take the form of BOs in these polytopes (i.e., each O contributes a charge of –1 to the central Al). This analysis yields an overall local charge of –1 and –2 for AlO_4 and AlO_5 units, respectively, so that the total excess negative charge is

$(N_{\mathrm{Al^{IV}}} - 2 \times N_{\mathrm{Ca}}) + 2 \times N_{\mathrm{Al^{V}}}$, where $N_{\mathrm{Al^{IV}}}$ is the number of 4-fold coordinated Al atoms. In turn, replacing a BO by a TO in the Al polytopes increases the local charge by +1/3 since a BO contributes a charge of –1 to the central Al (i.e., –2/2), whereas a TO only contributes a charge of –2/3 since it is shared by three Si or Al polytopes [55]. Altogether, since a TO is connected to 3 distinct polytopes, the overall increase in charge resulting from the transformation of a BO into a TO is equal to +1 (i.e., $3 \times 1/3$). Based on this, the number of TO atoms that is needed to counterbalance the excess negative charge of the Al atoms is then given by:

$$N_{\mathrm{TO}} = N_{\mathrm{Al^{IV}}} + 2 \times N_{\mathrm{Al^{V}}} - 2 \times N_{\mathrm{Ca}} \quad \text{(Eq. 6)}$$

where N_{TO} is the number of TO atoms. Since our model assumes the existence of only 4- and 5-fold coordinated Al atoms in the Al-rich regime, Eq. 6 can be converted into:

$$N_{\mathrm{TO}} = N_{\mathrm{Al}} - 2 \times N_{\mathrm{Ca}} + N_{\mathrm{Al^{V}}} \quad \text{(Eq. 7)}$$

We then assume that all the remaining O atoms act as BOs in Al-rich regime, whose number is given by:

$$N_{\mathrm{BO}} = N_{\mathrm{O}} - N_{\mathrm{TO}} \quad \text{(Eq. 8)}$$

3.4.1.d Comparison between analytical model and MD simulations

To validate the ability of our analytical model to offer accurate predictions of the degree of connectivity in the CAS glasses, we now compare the oxygen species fractions (i.e., in order of increasing connectivity: FO, NBO, BO, and TO) predicted by our model with those computed by MD simulations. **Error! Reference source not found.** 3.1 summarizes the predicted fractions for e ach oxygen species as a function of the molar fractions of CaO and Al_2O_3 (*x* and *y* respectively) following the unit formula $(CaO)_x(Al_2O_3)_y(SiO_2)_{1-x-y}$. Equations are provided for: (1) "*fully-depolymerized regime*," i.e., wherein all BOs are consumed by Ca atoms, so that the glass network

only comprises of FO and NBO atoms (i.e., $(y - x) \leq -2/3$), (2) "*partially-depolymerized regime*," wherein the network only contains NBO and BO atoms (i.e., $-2/3 \leq (y - x) \leq 0$), and (3) "*fully-polymerized regime*," wherein the network only contains BO and TO atoms (i.e., $0 \leq (y - x)$).

Table 3. 1 Summary of the predicted fractions of oxygen species in $(CaO)_x(Al_2O_3)_y(SiO_2)_{1-x-y}$ glasses, where x and y represent the mole percent of [CaO] and [Al_2O_3], respectively. f_{FO}, f_{NBO}, f_{BO}, and f_{TO} are the fractions of free oxygen (FO), non-bridging oxygen (NBO), bridging oxygen (BO), and tricluster oxygen (TO) atoms, respectively. Equations are separated into three distinct compositional regimes as discussed in the text.

	Fully-depolymerized regime	Partially-depolymerized regime	Fully-polymerized regime
Regime condition:	$(y - x) \leq -\frac{2}{3}$	$-\frac{2}{3} \leq (y - x) \leq 0$	$0 \leq (y - x)$
f_{FO}	$\frac{-2 - 3(y - x)}{2 + (y - x)}$	0	0
f_{NBO}	$\frac{4 + 4(y - x)}{2 + (y - x)}$	$\frac{-2(y - x)}{2 + (y - x)}$	0
f_{BO}	0	$\frac{2 + 3(y - x)}{2 + (y - x)}$	$\frac{4 - 3(y - x)}{4 + 2(y - x)}$
f_{TO}	0	0	$\frac{5(y - x)}{4 + 2(y - x)}$

Error! Reference source not found. shows the comparison between predicted (by our a nalytical model) and computed (by MD simulations) fractions for the oxygen species (a) FO, (b) NBO, (c) BO, and (d) TO as a function of $[Al_2O_3]$ – [CaO]. We first note that our analytical model offers good predictions of FO, NBO, BO, and TO over the entire CAS domain, both in terms of trend and magnitude. Then, we also note that our analytical model solely depends on the competition between $[Al_2O_3]$ and [CaO] atoms (i.e., "$y - x$", see Table 3.1) and, hence, does not fully capture the variations in the fractions of oxygen species at a fixed value of "$y - x$." The deviations between model and MD simulations are mostly observed at the vicinity of the transitions between compositional regimes, namely, near the fully-compensated (i.e., $y - x = 0$) and fully-depolymerized domains (i.e., $y - x = -2/3$) Such discrepancies can be attributed to the fact that, for the sake of simplification, our model assumes that only two types of oxygen species can coexist at the same time, while experiments [35,51,53] and MD simulations [34,42] suggest that more than two species can simultaneously coexist in the network. Nevertheless, despite these simplifications, our model offers a realistic description of the overall degree of connectivity of the glass, which is key to analytically predict the number of constraints per atom in the network (see Sec. IV).

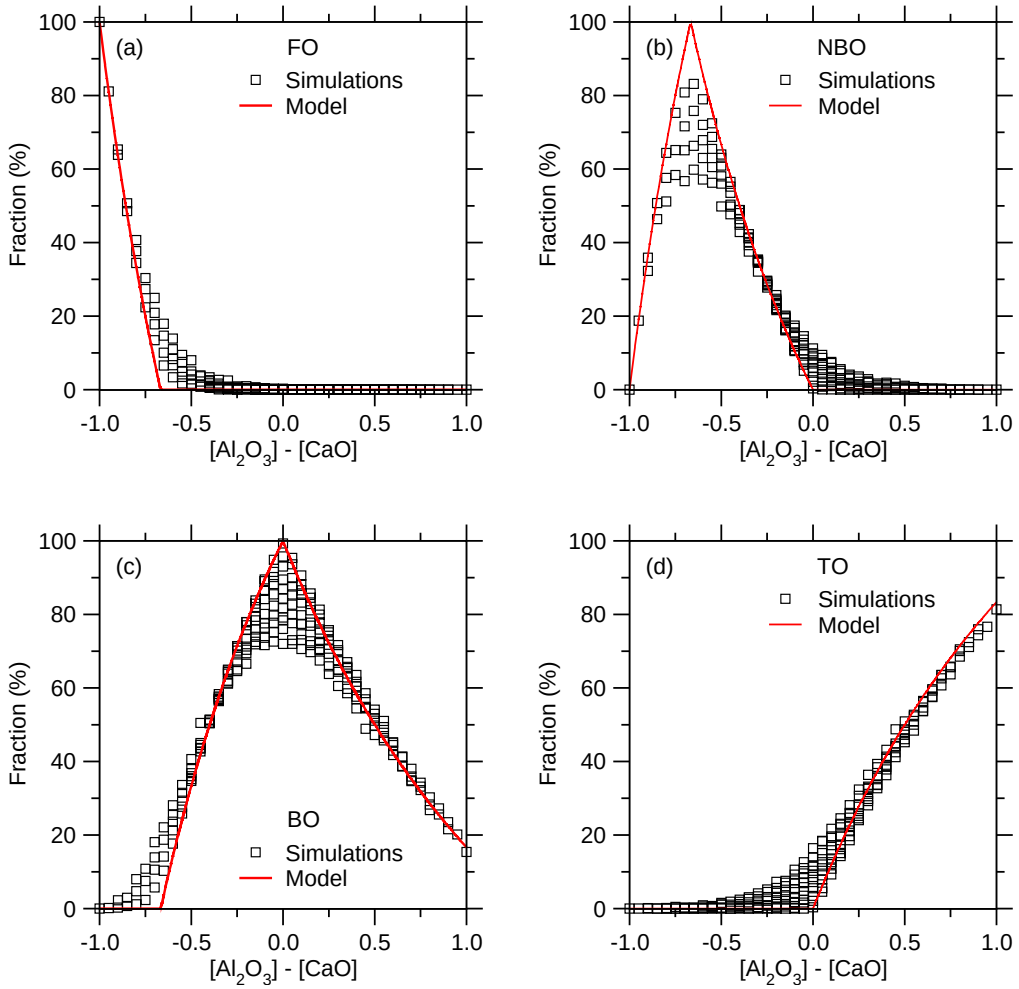

Figure 3. 2 Predicted (from the analytical model) and computed (from molecular dynamics simulations) fractions of each type of oxygen species as a function of composition: (a) free oxygen (FO), (b) non-bridging oxygen (NBO), (c) bridging oxygen (BO), and (d) "tricluster" oxygen (TO).

3.4.2 Connectivity of the network modifiers

We now discuss how the local connectivity of the network-modifying atoms (i.e., Ca atoms) is varying as a function of the glass composition. The analytical model described in the following aims to predict the partial coordination numbers of Ca atoms, that is, the average number of FO, NBO, BO, and TO around each Ca atom—which is critical to inform our topological model (see Sec. IV).

To establish our model of Ca connectivity, we take an alternative viewpoint and start by describing the average number of Ca around each type of O species. Indeed, based on the analysis of our MD simulations, we find that the average number of Ca atoms around FO, NBO, BO, and TO atoms remains fairly constant and barely depends on glass composition. The average FO–Ca, NBO–Ca, BO–Ca, and TO–Ca partial coordination numbers ($\bar{r}_{\mathrm{XO-Ca}}$, wherein XO refers to a given O species) are equal to 5.5, 3.0, 0.78, and 0.12, respectively. This trend (i.e., $\bar{r}_{\mathrm{FO-Ca}} > \bar{r}_{\mathrm{NBO-Ca}} > \bar{r}_{\mathrm{BO-Ca}} > \bar{r}_{\mathrm{TO-Ca}}$) can be understood from the fact that, starting from an isolated FO, the addition of each O–Si or O–Al bond (i) reduces the need for the negative charge of O atoms to be compensated by the nearby presence of a Ca cation and (ii) occupies some space around the central O, which prevents the accumulation of Ca neighbors. These partial coordination numbers can also be understood as a degree of "affinity" between Ca atoms and O species, wherein Ca cations are preferentially located in the vicinity of FO atoms and almost fully avoid TO atoms.

The knowledge of the XO–Ca partial coordination numbers then allows us to determine the Ca–XO coordination numbers ($r_{\mathrm{Ca-XO}}$) by expressing the total number of Ca–XO bonds ($N_{\mathrm{Ca-XO}}$) in two different ways as:

$$N_{\mathrm{Ca-XO}} = N_{\mathrm{Ca}} \times r_{\mathrm{Ca-XO}} = N_{\mathrm{XO}} \times r_{\mathrm{XO-Ca}} \quad \text{(Eq. 9)}$$

where N_{XO} is the number of XO atoms (where XO = FO, NBO, BO, and TO). As an additional refinement of the model, we add a scaling coefficient β to ensure that the total average coordination number of Ca cations (i.e., the sum of all the partial Ca–XO coordination numbers) remains equal to 6 throughout the entire CAS compositional domain [56]:

$$r_{\mathrm{Ca-XO}} = \beta \frac{N_{\mathrm{XO}} \times \bar{r}_{\mathrm{XO-Ca}}}{N_{\mathrm{Ca}}} \quad \text{(Eq. 10)}$$

where β can be expressed as:

$$\beta = \frac{6}{\sum_{\mathrm{XO}}(\bar{r}_{\mathrm{XO-Ca}} \times N_{\mathrm{XO}}) / N_{\mathrm{Ca}}} \quad \text{(Eq. 11)}$$

where the summation spans over all the O species (XO = FO, NBO, BO, and TO).

shows a comparison between the calculated (i.e., by our analytical model) and computed (i.e., by MD simulations) values of the (a) Ca–FO, (b) Ca–NBO, (c) Ca–BO, and (d) Ca–TO partial coordination numbers as a function of $[Al_2O_3]$ – [CaO]. Despite relying on simple assumptions, our model yields realistic predictions, both in terms of trend and magnitude. Similar to **Error! R eference source not found.**, we observe that most of the deviations between our model and MD simulations occur in the vicinity of the transitions between the fully-compensated and fully-depolymerized domains—again a consequence of the fact that our model only assumes that at most two types of oxygen species can coexist at the same time. Finally, we note that the computed Ca–XO partial coordination numbers exhibit some slight variations at fixed value of $[Al_2O_3]$ – [CaO], which is not accounted for by our model. Nevertheless, it is notable that the present analytical model can capture fairly well the non-monotonic and non-linear evolution of the local connectivity around Ca atoms while only relying on simple assumptions regarding the relative affinity of FO, NBO, BO, and FO atoms for Ca cations.

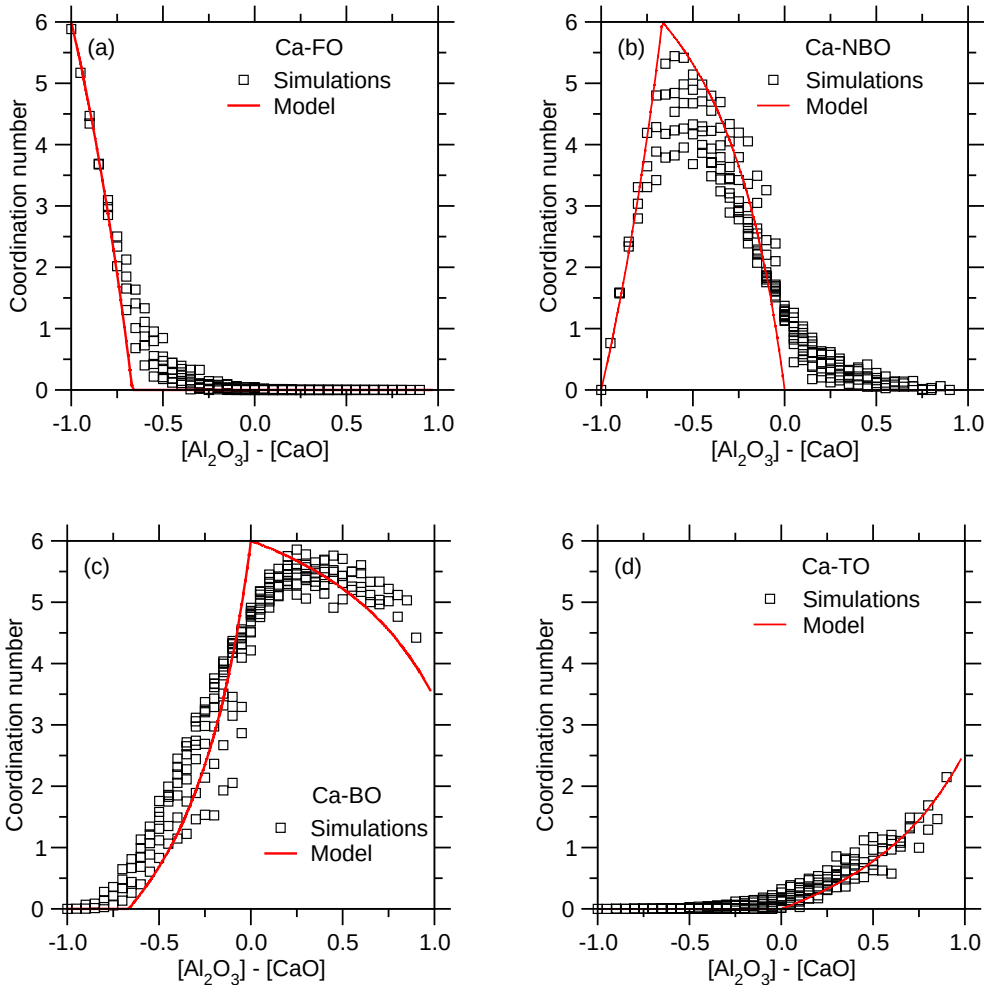

Figure 3. 3 Predicted (from analytical model) and computed (from molecular dynamics simulations) values of the (a) Ca–FO, (b) Ca–NBO, (c) Ca–BO, and (d) Ca–TO partial coordination numbers of Ca atoms as a function of composition.

3.5 Rigidity diagram of the calcium aluminosilicate ternary system

3.5.1 Glassy state

We use the topological inputs presented in Sec. III to determine the state of rigidity of CAS glasses (i.e., flexible, isostatic, or stressed-rigid) as a function of temperature and composition by enumerating the number of radial BS and angular BB constraints per atom based on the

coordination number of each atom, as described in the following. We first focus on the glassy state (i.e., low temperature). Table 3.2 summarizes the average number of BS and BB constraints created by each individual species in the glassy state. Note that, for simplicity, all the BS constraints are fully attributed to the cations (rather than being equally shared by cations and oxygen atoms). As expected, Si atoms systematically create 4 BS constraints with their 4 O neighbors, as well as 5 BB constraints (i.e., the minimum number of independent angles that need to be fixed to define the SiO_4 tetrahedron) [57]. In contrast, Al atoms exhibit a varying coordination number (i.e., 4 or 5 herein), so that the number of BS constraints created by Al atoms is given by their coordination number. Similar to Si atoms, 5 BB constraints are attributed to 4-fold Al atoms. However, we assume that 5-fold coordinated Al atoms do not create any BB constraints since they exhibit an unstable deformed octahedral angular environment [52,58,59]. The enumeration of the topological constraints created by Ca atoms requires more attention considering the ionic nature of Ca–O bonds. First, since ionic Ca–O bonds are non-directional, Ca atoms do not exhibit any well-defined angular environments and, as such, do not create any BB constraints [57]. Second, since BO and TO atoms are already fully charge-compensated by their Si or Al neighbors, we assume that Ca atoms only create strong BS constraints with their surrounding FO and NBO neighbors. Finally, each BO atom creates 1 BB constraint (e.g., to maintain Si/Al–O–Si/Al angles fixed around their average values), while each TO atom creates 3 BB constraints to define its trigonal environment [30]. In contrast, FO and NBO atoms do not create any BB constraint due to the non-directional nature of ionic Ca–O bonds. These inputs then serve to compute the number of BS and BB constraints for each glass composition.

Table 3. 2 Summary of the number of radial bond-stretching (BS) and angular bond-bending (BB) constraints created by each atomic species in calcium aluminosilicate glasses (at low temperature). For Al and O atoms, the BS and BB constraints are distinguished in terms of their coordination numbers. Note that the constraints created by Ca atoms depend on the type of O atoms they are connected to.

	Glassy state	
Species	BS	BB
Si atoms	4	5
Al atoms		
4-fold Al	4	5
5-fold Al	5	0
Ca–O bonds		
Ca–FO	1	/
Ca–NBO	1	/
Ca–BO	0	/
Ca–TO	0	/
O atoms		
FO	/	0
NBO	/	0
BO	/	1
TO	/	3

Figure 3.4 presents the number of radial BS, angular BB, and total (BS+BB) number of constraints per atom (n_c) as a function of $[Al_2O_3] – [CaO]$ for the entire CAS domain. The results obtained from our analytical topological model are compared with those obtained by MD simulations. We first note that the number of BS constraints per atom decreases upon increasing $[Al_2O_3] – [CaO]$,

shows a minimum at [Al_2O_3] = [CaO], and then subsequently increases [Fig. 3.4(a)]. The initial decrease in the number of BS constraints per atom primarily arises from the high number of BS constraints created by Ca atoms at low [Al_2O_3] – [CaO], wherein most of the Ca atoms create 6 BS constraints with their surrounding NBO and FO atoms—thereby resulting in a glass that exhibits a large average coordination number. The average coordination number then decreases as Ca atoms get replaced by Al atoms. However, when [Al_2O_3] > [CaO], the increase in the fraction of 5-fold Al and TO atoms eventually results in an increase in the number of BS constraints per atom upon increasing [Al_2O_3] – [CaO]. We then observe that the number BB constraints per atoms monotonically increases upon increasing [Al_2O_3] – [CaO] [Fig. 3.4(b)]. This can be understood from the fact that atomic species which do not create any BB constraints (i.e., Ca, NBO, and FO) gradually disappear upon increasing [Al_2O_3] – [CaO].

Altogether, we find that the total number of constraints per atom remains fairly constant (around n_c = 3) when [Al_2O_3] < [CaO] and notably increases when [Al_2O_3] > [CaO] [Fig. 3.4(c)]. This indicates that, at [Al_2O_3] < [CaO], the decrease in the number of BS per atom is perfectly balanced by the increase in the number of BB per atom. This behavior agrees with previous findings obtained in densified silicate glasses, wherein the number of BB constraints was found to adapt to the number of BS constraints [60]. This behavior was attributed to a self-organization mechanism, wherein weaker angular constraints form or break in response to variations in the number of stronger radial constraints to ensure that the glass remains isostatic (n_c = 3) [61,62].

Importantly, we find that our analytical topological model matches well with MD results. Note that, although the fractions of oxygen species (Fig. 3.2) and the Ca–O partial coordination numbers (Fig. 3.3) solely depend on [Al_2O_3] – [CaO], the resulting number of constraints also

depends on [SiO_2] (which explains why the outcome of our model is represented as hatched areas in Fig. 3.4 rather than as a single line).

The resulting rigidity diagram for the CAS ternary system is shown in Fig. 3.5(a). We note that CAS glasses are found to be systematically rigid (i.e., isostatic or stressed-rigid, $n_c \geq 3$) irrespectively of composition. Notably, the ternary map of the number of constraints per atom closely mimics that of the Young's modulus of CAS glasses [11], which supports the present constraint enumeration—since the Young's modulus has been shown to depend on the density of BS and BB constraints [42]. The increase in the rigidity of CAS glasses upon increasing [Al_2O_3] – [CaO] also echoes the corresponding increase in glass transition temperature that is observed experimentally [63,64]. Overall, our model offers an accurate description of the atomic structural rigidity for CAS glasses while remaining analytical.

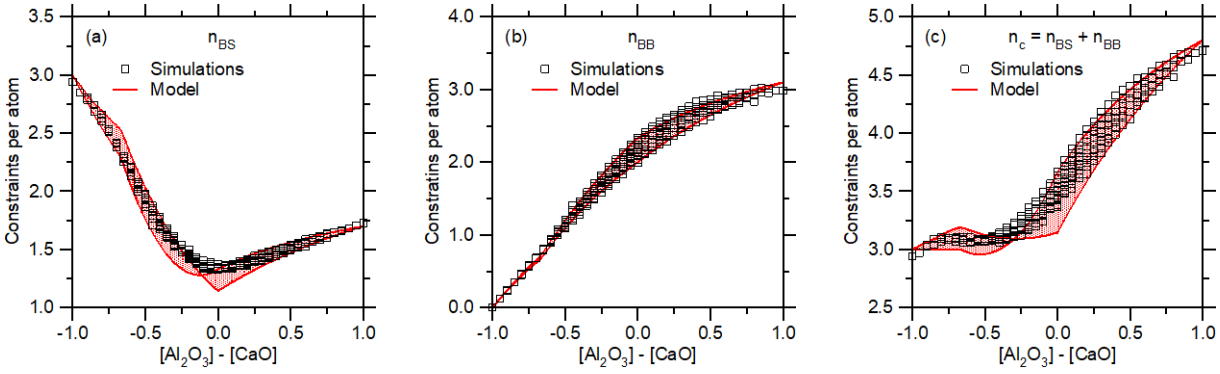

Figure 3. 4 Predicted (from analytical model) and computed (from molecular dynamics simulations) (a) number of bond-stretching (BS), (b) number of bond-bending (BB), and (c) total number of constraints per atom (n_c) as a function of composition.

3.5.2 Liquid state

Next, we discuss how temperature may affect the rigidity diagram of the CAS ternary system. To this end, we adopt temperature-dependent constraint theory, wherein each constraint can be intact (at low temperature, i.e., in the glassy state) or thermally-broken (at high temperature, i.e., in the liquid or supercooled liquid state) based on the competition between the constraints' free energy and available thermal energy [65–67]. In the following, rather than relying on unproven guesses regarding the temperature at which constraints break, we simply discuss how the breaking of each type of constraint affects the rigidity diagram of the CAS ternary system (i.e., the location of the flexible and stressed-rigid domains).

In the following, we discuss the successive effects of the breaking of (a) O–Al–O BB constraints, (b) Ca–FO constraints, and (c) BO and TO BB constraints—wherein these constraints are tentatively ranked from the weakest to the strongest, that, these constraints are ranked in terms of the temperature at which they are expected to break. This ranking is based on the following observations. First, the angular BB constraints associated with 4-fold Al atoms have been suggested by many studies to be notably weaker than those associated with Si atoms [27,28,68–71]. This has been attributed to the fact that the nearby presence of a charge-compensating cation tends to destabilize the angular environment of 4-fold coordinated Al atoms [28]. Second, ionic Ca–FO constraints are expected to exhibit a low bond energy and, hence, to break at low temperature [66]. Finally, the BB constraints of BO and TO atoms (referred to as γ constraints) have been noted to break at low temperature [66,67]. It is worth nothing that this ranking is only tentative, and more work is clearly needed to carefully determine the onset temperature at which each type of constraint breaks (and ranking thereof).

Figure 3.5 shows how the successive breaking of each type of constraint affects the rigidity diagram of the CAS ternary system. We first note that the breaking of the BB of the Al atoms drastically affects the rigidity diagram and results in the appearance of a flexible domain ($n_c < 3$) in the SiO_2-poor and CaO-rich regions [Fig. 3.5(b)]. This suggests that CAS systems may exhibit a composition-driven rigidity transition (i.e., from flexible, $n_c < 3$, to stressed-rigid, $n_c > 3$) in this range of temperature. The subsequent breaking of the Ca–FO BS constraints plays, overall, a more minor effect, but drastically reduces the rigidity of CaO-rich compositions [Fig. 3.5(c)]. Finally, the breaking of the BO and TO BB constraints causes all the compositions to become flexible—at the notable exception of pure SiO_2, which becomes isostatic [Fig. 3.5(d)].

Interestingly, these ternary rigidity diagrams offer useful insights into the possible origin of the glass-forming ability of CAS systems. First, at the highest temperature considered herein [Fig. 3.5(d)], SiO_2 becomes perfectly isostatic ($n_c = 3$). This agrees with the excellent glass-forming ability of SiO_2 [19,30]. It should be noted that SiO_2 exhibits a high glass transition temperature, so that, unlike other silicate glasses, the topological origin of its glass-forming ability should indeed be assessed at high temperature (i.e., wherein the BB constraints of the BO atoms are thermally broken). The state of rigidity of the CAS ternary system at moderate temperature [Figs. 3.5(b) and 3.5(c)] also exhibits some correlations with glass-forming ability [72,73]. First, CaO-rich systems are highly flexible ($n_c < 3$) and, hence—following Zachariasen's viewpoint [74]—do not have the ability to form an extended 3-dimensional rigid network to prevent crystallization. In contrast, Al_2O_3-rich glasses are highly stressed-rigid ($n_c > 3$). This may explain their low glass-forming ability, since the network does not exhibit enough structural flexibility to form a random (non-crystalline) network. The fact that SiO_2-rich calcium silicate glasses are stressed-rigid also echoes the fact that such glasses tend to phase-separate [3,72]. It is

also notable that the position of the isostatic boundary on the $CaO–SiO_2$ joint (around $(CaO)_{50}(SiO_2)_{50}$) corresponds to the compositional domain wherein calcium silicate glasses exhibit maximum glass-forming ability [75]. Altogether, these results suggest that it is the rigidity at finite temperature (i.e., in the supercooled liquid state) rather than at low temperature (i.e., in the glassy state) that governs the glass-forming ability of silicate glasses.

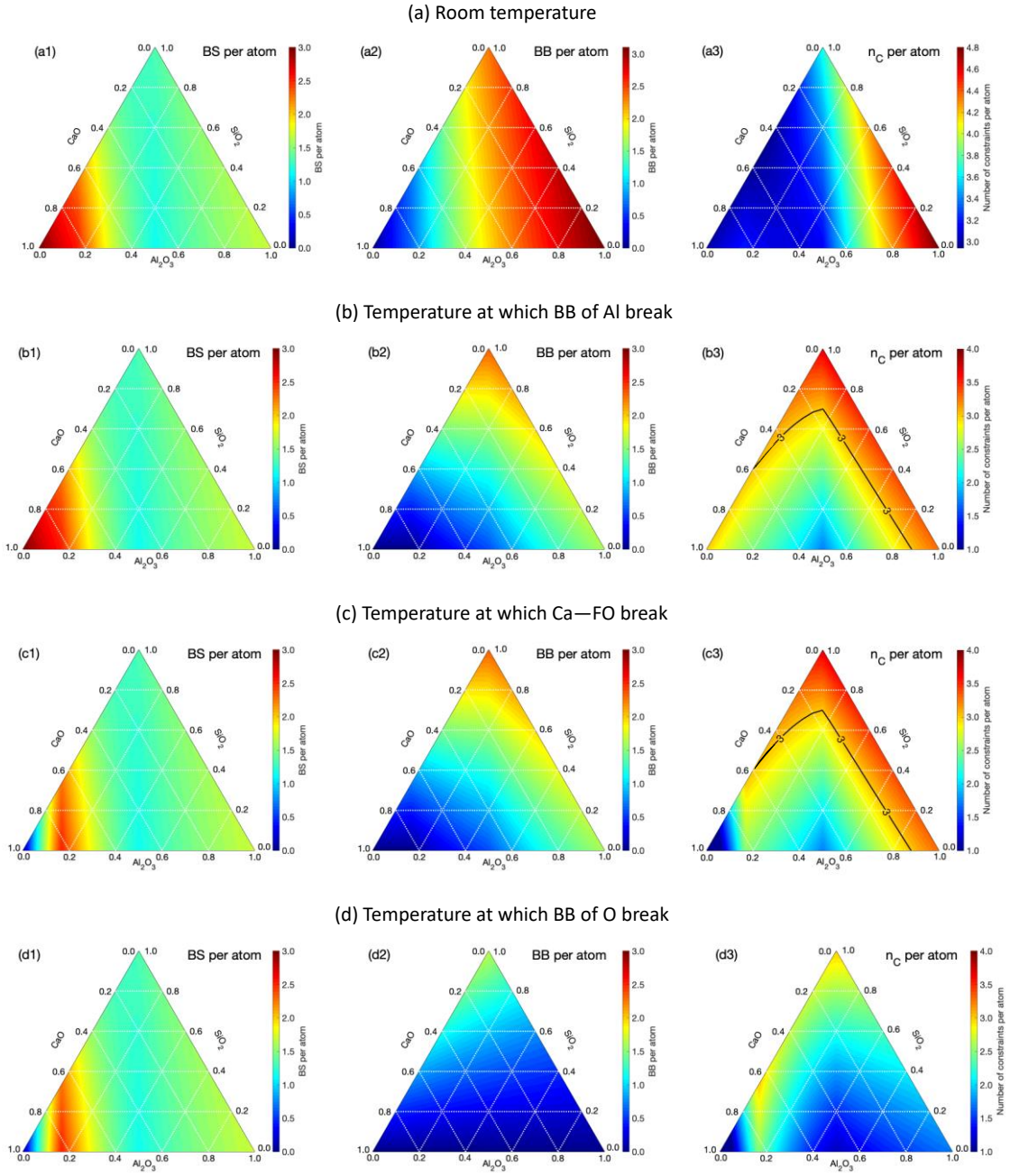

Figure 3. 5 Ternary plots presenting the number of (1) radial bond-stretching (BS), (2) angular bond-bending (BB), and (3) total number of constraints (n_C) per atom predicted by our analytical topological model. The constraints enumeration is conducted at (a) room temperature (glassy state), (b) the temperature at which O–Al–O BB constraints break, (c) the temperature at which Ca–FO

BS constraints break, and (d) the temperature at which BO and TO BB constraints break. The solid black line in the n_c ternary maps (b3 and c3) indicates the location of the flexible-to-rigid transition (i.e., at $n_c = 3$).

3.6 Conclusion

This work establishes a sound model describing the network topology of CAS glasses as a function of their compositions. It is important to note that, although this model is informed and validated by MD simulations, it remains fully analytical and, hence, can bypass MD simulations to offer predictions of the topology of a given CAS glass based on the sole knowledge of its composition. Indeed, this model provides a direct access to the state of rigidity (flexible, isostatic, or stressed-rigid) of CAS glasses as a function of composition and temperature regime (which dictates if a certain constraint species is thermally active or not). The obtained temperature-dependent ternary rigidity diagrams reveal the existence of correlations between network topology and glass-forming ability—wherein highly underconstrained (flexible) and highly overconstrained (stressed-rigid) systems exhibit low glass-forming ability, while, in turn, optimally-constrained glasses (isostatic) feature maximum glass-forming ability. Importantly, this study suggests that glass-forming ability is encoded in the network topology of the liquid state rather than that of the glassy state.

3.7 Acknowledgments

This work was supported by the National Science Foundation (under Grants No. 1762292, 1826420, 1826050, and 1928538) and the Federal Highway Administration (Grant #: 693JJ31950021).

3.8 References

[1] A. Ellison and I. A. Cornejo, International Journal of Applied Glass Science **1**, 87 (2010).

[2] H. F. W. Taylor, *Cement Chemistry* (Thomas Telford, 1997).

[3] C. Huang and E. C. Behrman, Journal of Non-Crystalline Solids **128**, 310 (1991).

[4] L. Cormier, D. R. Neuville, and G. Calas, Journal of Non-Crystalline Solids **274**, 110 (2000).

[5] D. R. Neuville, L. Cormier, and D. Massiot, Chemical Geology **229**, 173 (2006).

[6] L. Hennet, J. W. E. Drewitt, D. R. Neuville, V. Cristiglio, J. Kozaily, S. Brassamin, D. Zanghi, and H. E. Fischer, Journal of Non-Crystalline Solids **451**, 89 (2016).

[7] P. Ganster, M. Benoit, W. Kob, and J.-M. Delaye, J. Chem. Phys. **120**, 10172 (2004).

[8] M. Bouhadja, N. Jakse, and A. Pasturel, J Chem Phys **138**, 224510 (2013).

[9] X. Dai, J. He, J. Bai, Q. Huang, X. Wen, L. Xie, K. Luo, J. Zhang, W. Li, and S. Du, Energy Fuels **30**, 2407 (2016).

[10] H. Jabraoui, M. Badawi, S. Lebègue, and Y. Vaills, Journal of Non-Crystalline Solids **499**, 142 (2018).

[11] K. Yang, X. Xu, B. Yang, B. Cook, H. Ramos, N. M. A. Krishnan, M. M. Smedskjaer, C. Hoover, and M. Bauchy, Scientific Reports **9**, 8739 (2019).

[12] R. Ravinder, K. H. Sridhara, S. Bishnoi, H. S. Grover, M. Bauchy, Jayadeva, H. Kodamana, and N. M. A. Krishnan, Mater. Horiz. (2020).

[13] Y.-J. Hu, G. Zhao, M. Zhang, B. Bin, T. Del Rose, Q. Zhao, Q. Zu, Y. Chen, X. Sun, M. de Jong, and L. Qi, Npj Computational Materials **6**, 1 (2020).

[14] A. K. Varshneya, *Fundamentals of Inorganic Glasses* (Elsevier, 2013).

[15] E. D. Zanotto and F. A. B. Coutinho, Journal of Non-Crystalline Solids **347**, 285 (2004).

[16] H. Liu, T. Du, N. M. A. Krishnan, H. Li, and M. Bauchy, Cement and Concrete Composites **101**, 5 (2019).

[17] J. C. Mauro, Current Opinion in Solid State and Materials Science **22**, 58 (2018).

[18] J. C. Mauro, A. Tandia, K. D. Vargheese, Y. Z. Mauro, and M. M. Smedskjaer, Chem. Mater. **28**, 4267 (2016).

[19] J. C. Phillips, Journal of Non-Crystalline Solids **34**, 153 (1979).

[20] M. F. Thorpe, Journal of Non-Crystalline Solids **57**, 355 (1983).

[21] J. C. Phillips, Journal of Non-Crystalline Solids **43**, 37 (1981).

[22] M. Bauchy, Computational Materials Science **159**, 95 (2019).

[23] J. C. Mauro, A. J. Ellison, D. C. Allan, and M. M. Smedskjaer, International Journal of Applied Glass Science **4**, 408 (2013).

[24] M. M. Smedskjaer, J. C. Mauro, R. E. Youngman, C. L. Hogue, M. Potuzak, and Y. Yue, J. Phys. Chem. B **115**, 12930 (2011).

[25] M. M. Smedskjaer, Front. Mater. **1**, (2014).

[26] M. Bauchy, B. Wang, M. Wang, Y. Yu, M. J. Abdolhosseini Qomi, M. M. Smedskjaer, C. Bichara, F.-J. Ulm, and R. Pellenq, Acta Materialia **121**, 234 (2016).

[27] I. Pignatelli, A. Kumar, M. Bauchy, and G. Sant, Langmuir **32**, 4434 (2016).

[28] H. Liu, T. Zhang, N. M. Anoop Krishnan, M. M. Smedskjaer, J. V. Ryan, S. Gin, and M. Bauchy, Npj Materials Degradation **3**, 1 (2019).

[29] J. C. Mauro, American Ceramic Society Bulletin **90**, 7 (n.d.).

[30] M. Bauchy, 21st Century Nanoscience – A Handbook (2019).

[31] C. Le Losq, D. R. Neuville, P. Florian, G. S. Henderson, and D. Massiot, Geochimica et Cosmochimica Acta **126**, 495 (2014).

[32] P. F. McMillan and R. J. Kirkpatrick, American Mineralogist **77**, 898 (1992).

[33] J. F. Stebbins, S. Kroeker, S. Keun Lee, and T. J. Kiczenski, Journal of Non-Crystalline Solids **275**, 1 (2000).

[34] M. Bauchy, The Journal of Chemical Physics **141**, 024507 (2014).

[35] H. W. Nesbitt, G. M. Bancroft, G. S. Henderson, R. Ho, K. N. Dalby, Y. Huang, and Z. Yan, Journal of Non-Crystalline Solids **357**, 170 (2011).

[36] H. W. Nesbitt, G. M. Bancroft, G. S. Henderson, R. Sawyer, and R. A. Secco, American Mineralogist **100**, 2566 (2015).

[37] L. Cormier, D. Ghaleb, D. R. Neuville, J.-M. Delaye, and G. Calas, Journal of Non-Crystalline Solids **332**, 255 (2003).

[38] S. S. Schoenholz and E. D. Cubuk, ArXiv:1912.04232 [Cond-Mat, Physics:Physics, Stat] (2019).

[39] S. Plimpton, Journal of Computational Physics **117**, 1 (1995).

[40] Y. Yu, B. Wang, Y. J. Lee, and M. Bauchy, MRS Online Proceedings Library Archive **1757**, (2015).

[41] B. Wang, Y. Yu, Y. J. Lee, and M. Bauchy, Front. Mater. **2**, (2015).

[42] K. Yang, B. Yang, X. Xu, C. Hoover, M. M. Smedskjaer, and M. Bauchy, Journal of Non-Crystalline Solids **514**, 15 (2019).

[43] C. J. Fennell and J. D. Gezelter, The Journal of Chemical Physics **124**, 234104 (2006).

[44] X. Li, W. Song, K. Yang, N. M. A. Krishnan, B. Wang, M. M. Smedskjaer, J. C. Mauro, G. Sant, M. Balonis, and M. Bauchy, The Journal of Chemical Physics **147**, 074501 (2017).

[45] L. Martínez, R. Andrade, E. G. Birgin, and J. M. Martínez, J Comput Chem **30**, 2157 (2009).

[46] C. Le Losq, M. R. Cicconi, G. N. Greaves, and D. R. Neuville, in *Springer Handbook of Glass*, edited by J. D. Musgraves, J. Hu, and L. Calvez (Springer International Publishing, Cham, 2019), pp. 441–503.

[47] D. E. Day and G. E. Rindone, Journal of the American Ceramic Society **45**, 579 (1962).

[48] T. D. Taylor and G. E. Rindone, Journal of the American Ceramic Society **53**, 692 (1970).

[49] D. A. McKeown, F. L. Galeener, and G. E. Brown, Journal of Non-Crystalline Solids **68**, 361 (1984).

[50] E. F. Riebling, J. Chem. Phys. **44**, 2857 (1966).

[51] H. Wayne Nesbitt, G. S. Henderson, G. Michael Bancroft, R. Sawyer, and R. A. Secco, Chemical Geology **461**, 13 (2017).

[52] Y. Xiang, J. Du, M. M. Smedskjaer, and J. C. Mauro, J. Chem. Phys. **139**, 044507 (2013).

[53] D. Iuga, C. Morais, Z. Gan, D. R. Neuville, L. Cormier, and D. Massiot, J. Am. Chem. Soc. **127**, 11540 (2005).

[54] S. Takahashi, D. R. Neuville, and H. Takebe, Journal of Non-Crystalline Solids **411**, 5 (2015).

[55] M. J. Toplis, D. B. Dingwell, and T. Lenci, Geochimica et Cosmochimica Acta **61**, 2605 (1997).

[56] M. C. Eckersley, P. H. Gaskell, A. C. Barnes, and P. Chieux, Nature **335**, 525 (1988).

[57] M. Bauchy, M. J. Abdolhosseini Qomi, C. Bichara, F.-J. Ulm, and R. J.-M. Pellenq, J. Phys. Chem. C **118**, 12485 (2014).

[58] J. F. Stebbins, J. Wu, and L. M. Thompson, Chemical Geology **346**, 34 (2013).

[59] C. W. Burnham and M. J. Buerger, Zeitschrift Für Kristallographie - Crystalline Materials **115**, 269 (1961).

[60] M. Bauchy and M. Micoulaut, Nature Communications **6**, 6398 (2015).

[61] M. Bauchy and M. Micoulaut, Phys. Rev. Lett. **110**, 095501 (2013).

[62] M. Bauchy, M. J. A. Qomi, C. Bichara, F.-J. Ulm, and R. J.-M. Pellenq, Phys. Rev. Lett. **114**, 125502 (2015).

[63] T. K. Bechgaard, J. C. Mauro, M. Bauchy, Y. Yue, L. A. Lamberson, L. R. Jensen, and M. M. Smedskjaer, Journal of Non-Crystalline Solids **461**, 24 (2017).

[64] L. Cormier, D. R. Neuville, and G. Calas, Journal of the American Ceramic Society **88**, 2292 (2005).

[65] P. K. Gupta and J. C. Mauro, J. Chem. Phys. **130**, 094503 (2009).

[66] J. C. Mauro, P. K. Gupta, and R. J. Loucks, J. Chem. Phys. **130**, 234503 (2009).

[67] M. Bauchy and M. Micoulaut, Journal of Non-Crystalline Solids **357**, 2530 (2011).

[68] T. Oey, Y.-H. Hsiao, E. Callagon, B. Wang, I. Pignatelli, M. Bauchy, and G. N. Sant, RILEM Technical Letters **2**, 67 (2017).

[69] T. Oey, A. Kumar, I. Pignatelli, Y. Yu, N. Neithalath, J. W. Bullard, M. Bauchy, and G. Sant, Journal of the American Ceramic Society **100**, 5521 (2017).

[70] T. Oey, E. Callagon, G. Falzone, K. Yang, A. Wada, M. Bauchy, J. Bullard, and G. Sant, J AM CERAM SOC (2020).

[71] S. Gin, M. Wang, N. Bisbrouck, M. Taron, X. Lu, L. Deng, F. Angeli, T. Charpentier, J.-M. Delaye, J. Du, and M. Bauchy, Npj Materials Degradation **4**, 1 (2020).

[72] L. Martel, M. Allix, F. Millot, V. Sarou-Kanian, E. Véron, S. Ory, D. Massiot, and M. Deschamps, J. Phys. Chem. C **115**, 18935 (2011).

[73] J. E. Shelby, Journal of the American Ceramic Society **68**, 155 (1985).

[74] W. H. Zachariasen, J. Am. Chem. Soc. **54**, 3841 (1932).

[75] P. Saravanapavan and L. L. Hench, Journal of Non-Crystalline Solids **318**, 1 (2003).