Chemical bonding in metallic glasses from machine learning and crystal orbital Hamilton population

Ary R. Ferreira *

Department of Physics, Universidade Federal de São Carlos (UFSCar), 13565–905, São Carlos, SP, Brazil

(Received 30 April 2020; revised 21 July 2020; accepted 20 October 2020; published 9 November 2020)

The chemistry (composition and bonding information) of metallic glasses (MGs) is at least as important as structural topology for understanding their properties and production/processing peculiarities. This paper reports a machine learning (ML)-based approach that brings an unprecedented "big picture" view of chemical bond strengths in MGs of a prototypical alloy system. The connection between electronic structure and chemical bonding is given by crystal orbital Hamilton population (COHP) analysis; within the framework of density functional theory (DFT). The stated comprehensive overview is made possible through a combination of: efficient quantitative estimate of bond strengths supplied by COHP analysis, representative statistics regarding structure in terms of atomic configurations achieved with classical molecular dynamics simulations, and the smooth overlap of atomic positions (SOAP) descriptor. The study is supplemented by an application of that ML model under the scope of mechanical loading in which the resulting overview of chemical bond strengths revealed a chemical/structural heterogeneity that is in line with the tendency to *bond exchange* verified for atomic local environments. The encouraging results pave the way towards alternative approaches applicable in plenty of other contexts in which atom categorization (from the perspective of chemical bonds) plays a key role.

DOI: 10.1103/PhysRevMaterials.4.113603

I. INTRODUCTION

Since early reports of glassy alloys, over almost 60 years ago [1], their importance within the broad scope of technological developments of metallic materials has grown noticeably, albeit such pertinence has apparently reached a threshold in the last decade [2]. Metallic glasses (MGs) are amorphous alloys that exhibit a glass transition and are notorious for their extreme hardness and strength; thus they became obvious candidates for structural applications, while other relevant usage proposals have already been put forward in domains like biomedicine, nanotechnology, and energy [3,4]. Yet, at the present time, research and development (R&D) activities related to this class of advanced materials are still facing challenging issues such as critical casting thickness or brittleness of some nominal compositions found to be good glass formers. Nevertheless, recent advances in basic and applied research have provided innovative strategies for designing new MGs, with the focus on enhanced mechanical, chemical, and magnetic properties. For example, 3D printing (or additive manufacturing) is emerging as a promising alternative for the fabrication of Fe-based MGs parts to be employed as magnetic shielding or transformer (electrical device) core laminations [5,6]. Another prospect, relevant for energy applications, is the synthesis of MGs with large specific surface areas, found to be efficient for ultrafast hydrogen uptake [4].

Over the last decades, technological developments in characterization techniques were crucial to promote insights into the structure of multicomponent MGs within the full range of length scales. Down to the atomistic scale, the lack of long-range crystalline order (often accompanied by nontrivial chemistry) makes it a hard task to uncover mechanisms underlying transformations occurring in two essential contexts: (i) one is the production process, which requires mastery of the key factors (or formation mechanisms) that determine glass-forming ability (GFA); (ii) the other scenario is application and concerning their most prominent usability as structural materials [7], recent efforts have been focused on unveiling defects likely to determine mechanical behavior, in close analogy to well-known plastic flow mechanisms existing in crystalline systems. Particularly in this latter context, recent basic research has been conducted into atomic-scale characterization of structural heterogeneity [8] and dynamic 'defects' from the perspective of *flow units* [9]. This is an aspect that reflects space-time heterogeneity in MGs and a technique that has been increasingly used to study the associated relaxation dynamics is dynamical mechanical analysis (DMA, aka dynamical mechanical spectroscopy), due to its high sensitivity in detecting atomic rearrangements [10–14].

Still regarding structural characterization, large-scale molecular dynamics (MD) simulations have proven effective in providing trustworthy structural models of MGs, able to reproduce elementary properties like density, glass transition temperature [15], and representative statistics regarding atomic configurations [16]. Nevertheless, the quest for efficient strategies for handling realistic structural models (that often contain thousands of atoms) in such simulations is a problematic and topical issue [17], even for the prediction of essential "static" features like short- and medium-range orders (SRO and MRO). In fact, this is a constraint that imposes a critical limitation on the use of *ab initio* quantum mechanical MD simulations for that end, and the solution is nothing new:

* ary.ferreira@df.ufscar.br

 ©2020 American Physical Society

a multiscale approach, in which the role of electrons on interatomic interactions is abstracted and described in terms of the so-called interaction models (or interatomic potentials) employed in the well-known classical MD (CMD) simulations.

The central drawback of this approach, however, is that so many additional properties relying strongly on an accurate description of the material's electronic structure become simply not accessible in CMD simulations. In other words, having a thorough atomistic (structural topology) insight into a certain material may not be enough to reveal all the technologically relevant phenomena. A good example from recent literature is the key role played by local homopolar bonds on the structural stabilization of specific sites and how it impacts the resistance drift of amorphous phase-change materials [18].

This paper reports an unprecedented "big picture" view of chemical bond strengths in MGs of the Zr-Cu-Al (ZCA) alloy system. The link between electronic structure and first-principles chemical bonding information is given by density functional theory (DFT) [19,20] and crystal orbital Hamilton population (COHP) analysis [21], whereas representative statistics is attained by applying a machine learning (ML) approach to realistic structural models generated by CMD simulations. Under the specific scope of mechanical loading, the resulting overview of chemical bond strengths revealed a chemical/structural heterogeneity that is quite in line with the tendency to *bond exchange* verified for atomic local environments in the chosen alloy model system.

II. THEORY AND COMPUTATIONAL DETAILS

A. The machine learning-based approach

The ZCA system was selected as a prototypical alloy for application of the proposed ML-based protocol because of its practicality. First of all, its corresponding MGs are conventional model systems extensively studied due to their high GFA [22–24], and there are plenty of experimental and theoretical works available in the literature covering different topics on them. Moreover, the computational modeling process is made easy for this alloy given the availability of an interatomic potential [16] that has been widely used for years, providing valuable theoretical support to experimental studies until recently [25,26] (see Sec. S1 in the Supplemental Material [27]). For the goals of the present study, it certainly provides the required plausible description of "static" bulk SRO and MRO in these MGs with affordable 10 000-atoms cells [16]—naturally, assuming a homogeneous distribution of constituent elements in a glassy structure free of nanocrystals [4,28].

The precise motivation for a ML-based approach here is to enable the prediction of bond strengths between atom pairs with the DFT accuracy in 10 000-atoms cells of the ZCA MGs derived from CMD simulations. Such cells are expected to supply representative statistics regarding SRO (chemical/structural local environments) in these systems. However, it is manifest that the corresponding electronic structure quantum mechanical simulations are unfeasible and unnecessary. This is where the proposed ML model comes onto the scene, by learning bond strengths from a database of interactions (DBIs).

Here, it is important to point a proper definition of chemical bond strength within the scope of this work. In fact, it can be seen as one more quantum mechanical concept associated to bond order, i.e., the stability of a chemical bond indicated by the electron density preferably distributed within a region between the related pair of atoms rather than closer to the individual corresponding sites. Its origins lie in the linear combination of atomic orbitals (LCAO) molecular overlap population Mulliken formalism [29], followed by its extension to solids (periodic extended systems) dubbed crystal orbital overlap population (COOP) analysis [30], which was built on top of a particular crystal orbital (CO) scheme combined with an extended Hückel method. In a simple CO-tight-binding language, for two sites A and B with corresponding COs ϕ_ν and ϕ_μ, the elements of the overlap population matrix in a closed shell system are

$$P_{AB} = 2\sum_{\nu}^{A}\sum_{\mu}^{B} c_\nu^* c_\mu S_{\nu\mu},$$

with $c_\nu^* c_\mu$ and $S_{\nu\mu}$ the elements of the density matrix and the overlap matrix, respectively. So, as in any CO approach, the COOP method is based on the density of states (DOS) curve (i.e., on the number of electronic states over the energy scale [31]) and, as an energy-resolved bonding descriptor, it provides information about the chemical bonding from a weighted DOS curve obtained by the product of $S_{\nu\mu}$ and DOS matrix elements in the applicable energy ranges.

The resulting COOP curve constitutes an electron number partitioning scheme and therefore carries features that allow the use of its integral up to the Fermi level (ICOOP) in close analogy to bond order as an index of bond strength; since the computed quantities $\Re\{c_\nu^* c_\mu S_{\nu\mu}\}$ point to bonding (positive), nonbonding (zero), and antibonding (negative) contributions [21]. The COHP method (proposed in that same Ref. [21]) used in the present work is nothing but an alternative approach, in which the DOS curve is weighted by the Hamiltonian matrix elements ($H_{\nu\mu} = \langle\phi_\nu|\hat{H}|\phi_\mu\rangle$) instead of the corresponding $S_{\nu\mu}$ employed in COOP. So, whereas the latter is an electron-partitioning scheme, the former is an energy-partitioning scheme that minimizes the drawback caused by the basis-set dependence of the overlap integral and is very well suitable for first-principles DFT electronic structure simulations, but both are energy-resolved bonding descriptors.

Moving to the strategy for the generation of the database of structures to train the ML model (the aforementioned DBIs), it was grounded on the assumption that the referred statistical representativeness existing in a 10 000-atoms cell of a given nominal composition (NC) of the ZCA alloy can be attained with a corresponding ensemble of smaller 100-atoms cells. As described in details in Sec. II B, all these cells were obtained by CMD simulated cooling from the melt, and the stated statistical equivalence was indeed achieved (see Supplemental Material [27]). Moreover, in order to draw a parallel with experiments reported in the literature [28], the set of four NCs $(Zr_{0.5}Cu_{0.5})_{100-x}Al_x$ (with $x = 2$, 6, 8, and 10) has been selected.

Hence, a set of per-type DBIs has been created for all possible interaction types (ITs): Al-Al, Cu-Al, Cu-Cu, Zr-Al,

TABLE I. Overview in numbers of the per-type DBIs created for the MGs using the 100-atoms cells generated following the quenching protocol described in the text and whose amounts are indicated in the last row. All cells are available in the Supplemental Material [27] as *extended XYZ* files.

	$Zr_{49}Cu_{49}Al_2$	$Zr_{47}Cu_{47}Al_6$	$Zr_{45}Cu_{45}Al_{10}$
Al-Al	27	64	552
Cu-Al	5080	1775	17 073
Cu-Cu	57 230	6311	32 859
Zr-Al	7011	2497	22 938
Zr-Cu	169 487	19 153	99 015
Zr-Zr	76 919	8558	43 547
100-atoms cells	488	58	326

Zr-Cu, and Zr-Zr. Table I brings out an overview in numbers of these databases and details on their construction are provided in Sec. II B, including strategies to diversify the collection of chemical environments and to ensure transferability among different NCs. Here, it has to be pointed that the small number of Al-Al interactions shown in Table I is intrinsic to the material. It follows upon alloying of Al and has already been described in the literature as a *solute-solute avoidance* effect [32,33]. As discussed in the Supplemental Material [27], concerning the above-mentioned transferability, this effect was not a complicating factor when further applying the ML model, and a single Al-Al DBI was set up with all interactions of this type listed in Table I.

Recalling that each individual interaction in the set of DBIs has an associated bond strength value derived from COHP analysis, precisely, this value is assumed to be the additive inverse of the integral of the COHP curve up to the Fermi level (−ICOHP). The minus sign is conventionally included to make it compatible with the ICOOP counterpart, since the system undergoes a lowering of its energy when there are bonding contributions [21]. Hence, in the COHP curve, the product of $H_{v\mu}$ and the corresponding DOS matrix elements point to bonding (negative), nonbonding (zero), and antibonding (positive) values.

Next, it is appropriate to introduce a key "ingredient" of the proposed ML model, whose versatility also allowed the assessment of the convergence of SRO statistics in the 100-atoms cells in advance. It refers to the mathematical descriptor of each chemical environment around an individual atom, necessary to measure the "distance" (or dissimilarity) between atomic environments. The descriptor adopted in the present work is the smooth overlap of atomic positions (SOAP) [34]. The SOAP fingerprints describe the chemical/structural local environments ensuring, in a natural way, invariance to the basic symmetries operations: rotation, reflection, translation, and permutation of atoms of the same species. In short, it describes the atomic neighborhood by expanding it in a basis composed of spherical harmonics and a set of orthogonal radial basis functions. The derived rotationally invariant power spectrum yield elements that are collected into a unit-length vector $\mathbf{q}$ [35]. So, in practice, the normalized similarity between such SOAP vectors computed for two atoms, i and j, is given by their dot product, $\mathbf{q}_i.\mathbf{q}_j$.

Turning to the specification of the ML model, it is important to remark that its implementation is not intended to predict atomic (or per-atom) scalar quantities. The −ICOHP is a scalar property that is associated to atom pairs and, since the proposed ML model is founded upon a Gaussian process regression (GPR) framework, the function that measures the similarity between two chemical bonds [the *kernel* function $k(B_m, B_n)$] has to take into account two basic features: bond distances and the individual SOAP vectors of the atoms involved. Thus, the covariance (or *kernel*) function was defined as a squared exponential weighted by the normalized similarities given by the referred SOAP vectors, as following

$$k(B_m, B_n) = \exp\left(-\frac{\left(d_{ij}^m - d_{ij}^n\right)^2}{2\theta^2}\right)\left(\frac{\left(\mathbf{q}_i^m.\mathbf{q}_i^n + \mathbf{q}_j^m.\mathbf{q}_j^n + \mathbf{q}_i^m.\mathbf{q}_j^n + \mathbf{q}_j^m.\mathbf{q}_i^n\right)}{4}\right). \quad (1)$$

With d_{ij}^m the distance between atoms i and j in the chemical bond B_m, and $\mathbf{q}_i^m$ and $\mathbf{q}_j^m$ their corresponding SOAP vectors. The adjustable scaling parameter θ defines the ML model's behavior and sets the characteristic length scale of the GPR.

Finally, using the *kernel* function in equation (1), the ML model is able to predict the −ICOHP value associated to an arbitrary bond B according to

$$\mathrm{ICOHP}^{\mathrm{ML}}(B) = \sum_{m=1}^{N} \alpha_m k(B, B_m)^{\zeta}, \quad (2)$$

with the hyperparameter $\zeta = 1$ establishing a linear *kernel* and N the size of the training DBI at issue containing the set of reference chemical bonds $\{B_m\}_{m=1}^N$ (or training set) with their corresponding *ab initio* −ICOHP values. By inverting the $N \times N$ *kernel* matrix $\mathbf{K}$, whose elements $K_{mn} = k(B_m, B_n)$ are defined with the training set, the per-interaction weights α_m in equation (2) are computed as following

$$\alpha_m = \sum_{n=1}^{N} \{\mathbf{K}^{\zeta} + [(\sigma^2\gamma)\mathbf{1}]\}_{mn}^{-1} \mathrm{ICOHP}^{\mathrm{DFT}}(B_n), \quad (3)$$

with σ^2 the standard deviation of the −ICOHP values in the training set $\{B_n\}_{n=1}^N$, γ a regularization adjustable parameter, $\mathbf{1}$ the $N \times N$ unit matrix, and $\mathrm{ICOHP}^{\mathrm{DFT}}(B_n)$ the −ICOHP value of the reference chemical bond B_n computed from first-principles DFT.

B. Computational details

All the CMD simulations were carried out using the velocity-Verlet integrator as implemented in the LAMMPS package [36] (release 16 Feb 2016). The embedded atom model (EAM) interatomic potential developed and

properly tested by Cheng *et al.* [16,37] (see Sec. S1 in the Supplemental Material [27]) was used with a cutoff radius of 6.5 Å to describe the interatomic forces in all MGs. A common quenching protocol was used for the generation of all 10 000-atoms and 100-atoms cells introduced in Sec. II A. Firstly, an initial configuration was set by randomly positioning the atoms in a cubic supercell, whose initial volume was estimated for each NC by initially assuming a dense sphere packing weighted with the Zr, Cu, and Al atomic radii, with a length in excess of 10% added to each cell vector. In order to avoid superposition of atoms, a conjugate gradient minimization on the random initial structure was executed with a stop criterion defined by a force threshold of 10^{-8} eV/Å. Next, the system was thermalized at 2000 K in the isothermal-isobaric (*NPT*) ensemble for 2 ns [in the time evolution of the CMD simulation (Δt)]. With a time step of 2 fs (adopted in all runs), the Nosé-Hoover thermostat was used with a dump coefficient of 0.2 ps, whereas the barostat was set to zero pressure with a dump coefficient of 2 ps. Subsequently, the system was cooled to 300 K with a minimal and feasible rate of 8.5×10^9 K/s, and finally, the glassy structure was allowed to relax at 300 K for $\Delta t = 2$ ns.

As pointed out in Sec. II A, some of the 100-atoms cells listed in Table I were submitted on demand to DFT first-principles calculations to set up the DBIs for the ML model. For the purpose of diversifying the collection of chemical environments, each one of those selected cubic supercells derived from CMD simulations was submitted to 3D geometrical transformations which have generated 14 new structures, namely, shearing along the x, y, and z axes (six new structures), compression and tension along the x, y, and z axes (six new structures), and isotropic compression and tension (two new structures).

All the referred DFT electronic structure simulations were performed using the QUANTUM ESPRESSO [38,39] (QE) open-source software suite version 6.2.0, with plane-wave (PW) basis sets and projector augmented waves (PAW) [40] datasets from the PSLibrary project version 1.0.0 [41]. The Perdew-Burke-Ernzerhof (PBE) [42] generalized gradient approximation was used to describe the exchange-correlation functional in all computations, remarking that the above-mentioned EAM interatomic potential [16,37] used in the CMD simulations was also parametrized from results of DFT-PBE calculations. For all 100-atoms cells, the PW basis set was truncated with a kinetic energy cutoff of 70 Ry and the Monkhorst-Pack procedure [43] was used to determine the k-points disposition in the first Brillouin zone from a $2 \times 2 \times 2$ sampling (corresponding to a density of k points of about 0.04 in all structures). A Fermi-Dirac probability distribution was used as a smearing function to set the occupations of energy levels, with a common broadening parameter $k_bT = 8$ mRy.

The ICOHP method has been introduced in Sec. II A and was earlier described in Ref. [21] as a scheme suitable for first-principles DFT calculations. However, further development was required to make it compatible with currently predominant and numerically efficient PAW-based computations with PW basis sets, what was achieved with the projected COHP (pCOHP) approach [44,45] which is implemented in the LOBSTER code [46,47] version 3.0.0 employed in the present work. Apropos, this implementation has already proven to be effective when using PAW data generated by QE [48], which are required for projections from delocalized wave functions onto local auxiliary basis. For the referred projections, a default local basis set provided by Bunge *et al.* [49] was adopted with the following set of local orbitals for each atomic specie: Al ($3s\,3p_x\,3p_y\,3p_z$), Cu ($3s\,4s\,3p_x\,3p_y\,3p_z\,3d_{xy}\,3d_{yz}\,3d_{z^2}\,3d_{xz}\,3d_{x^2-y^2}$), and Zr ($4s\,5s\,4p_x\,4p_y\,4p_z\,4d_{xy}\,4d_{yz}\,4d_{z^2}\,4d_{xz}\,4d_{x^2-y^2}$). Here, it is important to report that the high quality and reliability of these projections is reflected by the small values of the charge spilling (no higher than 1.6%) in all ICOHP calculations with 100-atoms cells reported in this work—in passing, absolute total spilling did not exceed 4.8%.

The SOAP descriptors were generated using the QUIP package [50], with a cutoff of 3.75 Å for the definition of the range of each chemical environment around atoms, including all elements in its composition. This cutoff value was based on partial pair distribution functions that are well known for the MGs of the ZCA alloy [16]. The spherical harmonics basis band limit and the number of radial basis functions were set to 6 and 8, respectively; all the remaining parameters kept with their corresponding default values.

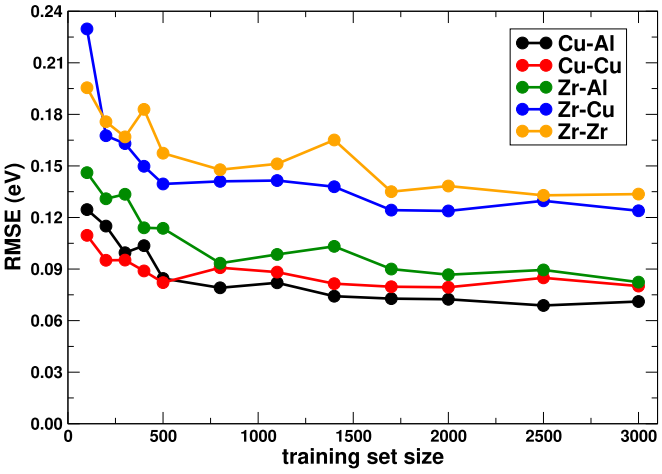

FIG. 1. Root-mean-square errors (RMSE) calculated from different Gaussian process regressions made for each interaction type in the nominal composition $Zr_{45}Cu_{45}Al_{10}$ assuming distinct training set sizes. The ML model parameters were set as $\zeta = 1$, $\theta = 0.5$, and $\gamma = 0.010$, and a fixed testing set size of 1000 interactions was taken.

III. RESULTS AND DISCUSSION

A. Predictive power of the ML model

The predictive power of the ML model given by equation (2) was evaluated individually for each interaction type (IT) by studying its convergence with respect to the training set size, for a fixed testing set size. Due to its completeness (see Table I), the set of DBIs created for the NC $Zr_{45}Cu_{45}Al_{10}$ was selected for initial tests, for which the ML model parameters were arbitrarily set as $\zeta = 1$, $\theta = 0.5$, and $\gamma = 0.010$ [see also equations (1) and (3)]. The corresponding results are presented in Fig. 1, from which it is possible to see that the corresponding root-mean-square errors (RMSE) are dependent on the IT but, in general, they are fairly converged

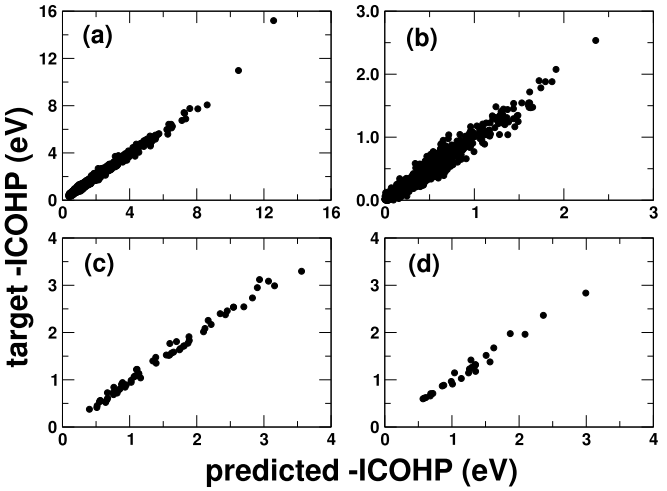

FIG. 2. Scatter plots of results from some of the Gaussian process regressions listed in Table S1: (a) Zr-Zr in $Zr_{45}Cu_{45}Al_{10}$, (b) Cu-Cu in $Zr_{45}Cu_{45}Al_{10}$, (c) Al-Al in $Zr_{47}Cu_{47}Al_6$, and (d) Al-Al in $Zr_{49}Cu_{49}Al_2$.

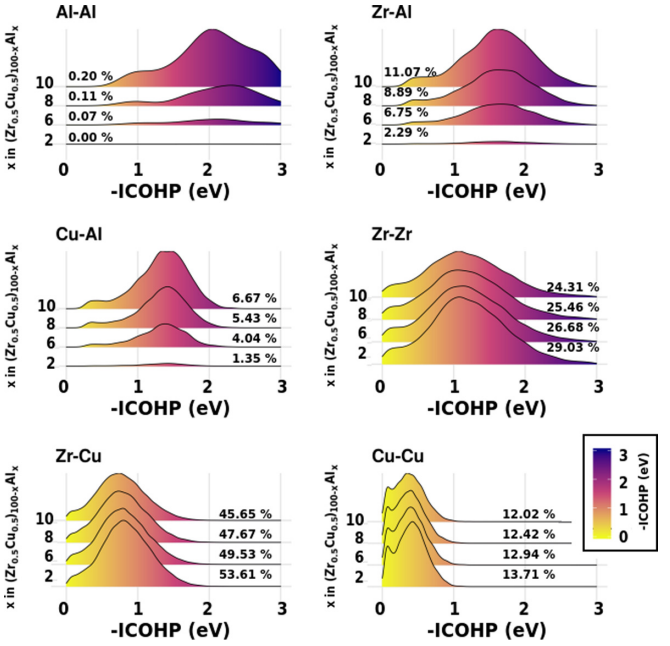

FIG. 3. Distributions of −ICOHP values predicted by the ML model for all interatomic interactions existing in 10 000-atoms cells of the series of nominal compositions (NC) $(Zr_{0.5}Cu_{0.5})_{100-x}Al_x$ (with $x = 2$, 6, 8, and 10). The fractions of each interaction type (IT) in each NC are indicated as percentages.

with small training set sizes (from about 500 interactions). Furthermore, given the range of −ICOHP values computed for the MGs under study listed in Table S1, the individual RMSE values calculated for each IT show that the ML model produces small prediction errors. Remarking that each point in the plots of Fig. 1 corresponds to a GPR in which the training and testing sets were randomly created from the DBIs listed in Table I.

Table S1 provides further statistics from a set of GPRs equivalent to those shown in Fig. 1, however adopting rather large training and testing sets. As already mentioned, the lack of Al-Al interactions is intrinsic to the ZCA alloy and made it impossible to carry out the same thorough tests for that specific IT. Nevertheless, it can be checked in Table S1 that the corresponding RMSE values also point to small prediction errors. With respect to the other two NCs $Zr_{47}Cu_{47}Al_6$ and $Zr_{49}Cu_{49}Al_2$, it can be verified in Tables S2 and S3 that similar RMSE values were found for their correlated DBIs. Hence, since RMSE values are scale dependent, the overall predictive power of the proposed ML model can be said satisfactory within the context of this study. The scatter plots of some particular ITs listed in Table S1 are shown in Fig. 2, and the complete list is available in the Supplemental Material [27] (Figs. S13 to S30).

B. Application of the ML model

1. Static structures at room conditions

Once the validation is complete, the ML-based approach was used to predict bond strengths between atom pairs in the four 10 000-atoms cells created for each NC of the ZCA alloy. Based on the convergence tests and transferability of individual DBIs explained in the Supplemental Material [27], a minimal and feasible training set ($\{B_n\}_{n=1}^{N}$) with 600 interactions for each IT has been set by merging equally the corresponding DBIs listed in Table I. The referred 10 000-atoms cells are able to provide the required statistical representativeness regarding chemical environments in these MGs (see Sec. S1 in the Supplemental Material [27]) and they were generated with CMD simulations following the protocol described in Sec. II B.

Above all, this work is aimed at delivering an approach able to promote insights into the interplay of chemical/structural and dynamical inhomogeneities existing in MGs in different contexts—to be explored in future studies. The most promising applications comprise research focused on topical conceptions in which atom categorization based on the strength of chemical bonds plays a key role, like the aforementioned atomic models in terms of *flow units* [9] or *tightly bonded clusters* [51], *weakest configurations* [52], *bond exchange* [53] processes, among others. Therefore, instead of going straight to the issue of how −ICOHP values are 3D distributed within the 10 000-atoms cells, this work will be limited to the assessment of the propensity for *bond exchange* of individual atomic local environments under the scope of mechanical loading, and then, evaluate its connection to bond strength and atomic mobility using ordinary descriptive statistics. However, before doing so, it is worth looking at a simpler per-bond analysis as depicted in Fig. 3.

Most importantly, it should be emphasized that the set of results in Fig. 3 represents an unprecedented "big picture" view of bond strengths between atoms in these MGs derived from quantum mechanics. With respect to chemical bonding information, that comprehensive overview is arguably more complete and revealing than previous reports with similar aims, including quantum chemistry calculations with localized basis sets for rather small isolated atomic clusters [54,55] or even *ab initio* MD simulations carried out with periodic boundary conditions but for cells containing a few hundred interatomic interactions [16].

In fact, the profiles of the distributions in Fig. 3 do not exhibit significant changes upon alloying of Al and, as expected, only their respective fractions vary distinctively in each IT (indicated as percentages). Yet, they suggest that atoms can be grouped according to their chemical bonding situations, pointing to the possibility of drawing strategies to segregate atoms in a 3D fashion based on a given reference −ICOHP value. This is reasonable indeed, and one can see that Cu-Cu interactions show up as the weakest ones, together with a non-negligible fraction of Zr-Cu and Zr-Zr bonds, all having −ICOHP values below a value around 0.7 eV. Conversely, substantial amounts of relatively strong Zr-Cu and Zr-Zr bonds are also predicted by the ML model.

One further remark concerning the Al alloying effect is that despite the negligible number of Al-Al bonds—related to the already introduced *solute-solute avoidance* effect [32,33]—a fair amount of strong Zr-Al and Cu-Al interactions with $-\text{ICOHP} \gtrapprox 1.0$ eV rise. Based on the percentages of ITs pointed out in Fig. 3, it can be inferred that these two latter are replacing the weaker Zr-Zr and Zr-Cu bonds; given the small variation of the weakest Cu-Cu bonds. This is likely to be the main chemical effect of Al alloying, which is reflected in the quite distinctive mechanical behavior upon mechanical loading reported for the two extreme NCs reported in the literature [28]: Whereas the MG $Zr_{45}Cu_{45}Al_{10}$ was found to be very brittle, the lowest Al-content counterpart $Zr_{49}Cu_{49}Al_2$ presented a large plasticity that has been assigned to localized nanocrystallization in the structure upon mechanical loading. This is the idea that will be further explored in the next section through a straightforward assessment of the propensity for *bond exchange* of individual atomic local environments.

2. *Structures subjected to uniaxial compression*

So far, the distributions in Fig. 3 do not express any information concerning the collective role of the set of bonds existing in each individual local environment (LE, i.e., the neighborhood surrounding a central atom). Moreover, the respective bond strengths were extracted out of snapshots of static structures equilibrated at room environment conditions, which is not enough to assess the propensity for *bond exchange*, which is a dynamical aspect of the MGs under study. So, in order to account these factors under the specific scope of uniaxial compression, a set of additional CMD simulations have been planned.

Stressing that the ultimate goal of these extra simulations is to show that there is a consistency between the chemical bonding heterogeneity disclosed by the ML model (as shown in Fig. 3), and the time evolution of the structural topology of individual LEs upon mechanical loading—as described by the Newton's equations and underlying interaction model (the EAM potential). It is important to bare in mind that these are two distinct aspects of the chemistry and dynamics, respectively, that are being revealed in this work through two completely independent computational approaches, whose complementarity has a great potential to bring insights into the synergic role played by chemical/structural and dynamical heterogeneities in MGs.

It was introduced in Sec. II A that the set of four NCs $(Zr_{0.5}Cu_{0.5})_{100-x}Al_x$ (with $x = 2$, 6, 8, and 10) has been selected aiming at drawing a parallel with experimental results reported in the literature [28], and the choice of that experimental study was motivated by two key aspects. First of all, the samples were subjected to uniaxial compression, a mechanical load test with a constrained geometry that simplifies its computational modeling, providing reliable results for the evaluation of the time evolution of LEs, as sought in this work. Additionally, the authors resorted to transmission electron microscopy and 3D atom probe tomography to characterize the microstructure of the samples, what provided the proper experimental backing to describe them as homogeneous amorphous alloys (i.e., free of nanocrystals)—at least before compression. This is quite relevant regarding computational modeling, since that is the precise microstructure represented by the 10 000-atoms cells used as structural models in the present work.

So, the corresponding CMD simulations were carried out for all NCs (with a rather tight time step of 1 fs) and they are quite consistent; further technical details are provided in the Supplemental Material [27] in a comprehensive way and complementary to those introduced in Sec. II B. However here it is worth commenting that MD simulations of uniaxial compression require proper approximation to describe temperature and pressure dissipation. In this work, the equations of motion proposed by Melchionna *et al.* [56] have been employed with periodic boundary conditions in the *NPT* ensemble, in which temperature is controlled with the Nosé-Hoover thermostat, whereas pressure is regulated by decoupling the boundary in the loading direction from the *NPT* equations governing the other two orthogonal directions. This approach has already proven effective elsewhere [57] and was fairly satisfactory for the purposes of the present study.

Furthermore, for the sake of transparency, there is another technical aspect of these CMD simulations that deserves a critical remark here in the main text, and it is not commonly discussed in the literature. The accurate atomistic simulation of the stress-strain (SS) curves measured in the experiments reported by Kumar *et al.* [28] requires structural models able to describe concomitantly two key facets: the role of surface effects on the brittle behavior of the MG $Zr_{45}Cu_{45}Al_{10}$ and the localized nanocrystallization (particles with some nanometers) taking place within shear bands, as well as its relationship with the extended plasticity reported for the MG $Zr_{49}Cu_{49}Al_2$.

The solution could be to increase the structural models up to some dozens (or even few hundreds) of millions of atoms or, in a much more consistent approach, resort to a multiscale-based strategy. For instance, one could derive an upscaled peridynamic model to describe the brittle behavior of the former MG or parametrize a constitutive model to describe the SS curve of the latter MG accounting for localized nanocrystallization. However, this is definitely out of the scope of the present work and unnecessary for its aims. That issue is well discussed in Sec. S2 in the Supplemental Material [27]. In short, the SS curves depicted in Fig. 4 fairly reproduce the linear elastic behavior of the MGs; however, the portions that correspond to yield strength (strain $\approx$5%) and further plastic deformation (strain >5%) bear two common artifactual features that will certainly not affect further discussion within the context of the aims of the present study. Nevertheless, from here onward, most of the analyses will be restricted to

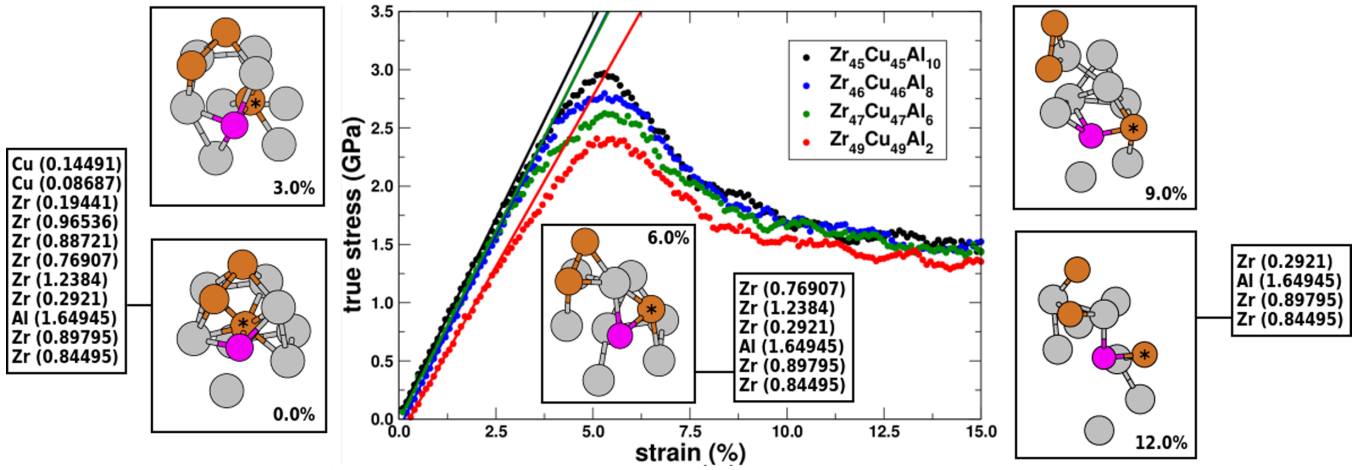

FIG. 4. Stress-strain curves simulated at 300 K with a compression rate of 1×10^7 s^{-1} for the MGs studied in this work. The solid lines denote the elastic behavior (see also Fig. S32). The time evolution of a specific Cu-centered (marked with an asterisk) nonpersistent local environment (NPLE, see text) extracted out of the structural model of the MG $Zr_{45}Cu_{45}Al_{10}$ is depicted, whose snapshots are labeled with the corresponding strain value. The −ICOHP values (in eV) of the first neighbors computed for the 0% strained structure are listed. These same values are repeated (not recalculated) for the persistently bonded atoms in the 6% and 12% strained NPLEs. Zr, Cu, and Al atoms are shown in gray, brown, and pink colors, respectively.

strain $\lessapprox$6%, although the extended segments of the SS curves will be carefully taken into account for examining particular aspects of the LEs, which are unrelated to the macromechanics of the corresponding MGs.

Returning to the topic of the time evolution of LEs in the MGs upon mechanical loading, in addition to the SS curves, Fig. 4 also brings an example of a particular Cu nonpersistent local environment (NPLE) extracted out of the structural model of the MG $Zr_{45}Cu_{45}Al_{10}$, whose central atom is marked with an asterisk. First of all, in the terminology of this work, a NPLE is a LE whose neighborhood has changed permanently over the mechanical loading—i.e., through some kind of migration or rearrangement mechanisms, it has underwent *bond exchanges* and some of its neighbor atoms have been replaced. Such changes were monitored in all CMD simulations for the two extreme NCs $Zr_{45}Cu_{45}Al_{10}$ and $Zr_{49}Cu_{49}Al_2$, and the referred example Cu-centered NPLE has lost five of its eleven neighbors at 6% strain. Pointing that such NPLE broken bonds will be referred to as NPLE-BBs in this work, whereas, in contrast, those remaining persistent bonds will be dubbed NPLE-PBs.

Additionally, it is self-evident that, for illustrative purposes, new neighbors are not shown in Fig. 4, and the neighbor's −ICOHP values predicted by the ML model for the nonstrained structure (those from Fig. 3) are shown, primarily, to serve as a label for identifying the corresponding atoms in further strain values—i.e., they have not been recalculated in the new strained structures. Moreover, in order to circumvent thermal fluctuation effects when monitoring neighbors, an additional 0.50 Å has been added to the cutoff value of 3.75 Å introduced in Sec. II B, and this is the only arbitrary parameter used in this procedure.

The Cu-centered NPLE in Fig. 4 provides just an extreme example of a set of *bond exchange* processes in which the central atom has lost 64% of its original neighbors at ("spurious") 12% strain. From that figure, it can be seen that the LE of this atom has indeed underwent severe changes over compression, and, as expected, this is a Cu atom whose bonds are overall weak and hence representative of the corresponding distributions depicted in Fig. 3. The complete list of bond strengths together with the respective bond distances of this particular NPLE is available in Table S4, where it is noticeable that, incidentally, all the broken bonds have an associated −ICOHP value that is less than 1.0 eV. Naturally there are exceptions, since the dynamics of such *bond exchange* processes is influenced by multiple factors like the number of nearest neighbors, relative distribution of the respective bond strengths around the coordination sphere, vibration fluctuations, and related entropic effects, among others.

So, in order to provide a complementarity support to that line of reasoning, a second contrast example has been extracted out of the same 10 000-atoms cell of the MG $Zr_{45}Cu_{45}Al_{10}$. It refers to a Zr-centered NPLE that has exchanged only 26% of its bonds at 12% strain. The associated bonding data is also listed in Table S4 and the snapshots of its time evolution over compression are depicted in Fig. S37. That second example of NPLE reinforces the idea that all those factors influencing the dynamics of *bond exchange* processes mentioned in the last paragraph must be taken into account if one aims at specifying a type of local descriptor, able to allow sound segregation of atoms in MGs. In fact, atom categorization is a very promising application of the ML-based approach proposed in this work, capable of opening up new prospects and of playing a valuable role in further studies focused on current conceptions regarding chemical/structural and dynamical inhomogeneities existing in these materials other than *bond exchange* [53] processes. Nevertheless, the development of such referred descriptors definitely goes beyond the goal of the present work.

Moving to descriptive statistics of the whole 10 000-atoms cells of the two extreme NCs $Zr_{45}Cu_{45}Al_{10}$ and $Zr_{49}Cu_{49}Al_2$, before all, the reader is referred to Figs. S38 and S39 to see that extreme cases like the Cu-centered NPLE shown in Fig. 4 are rare in the CMD simulations. At 6% strain, most

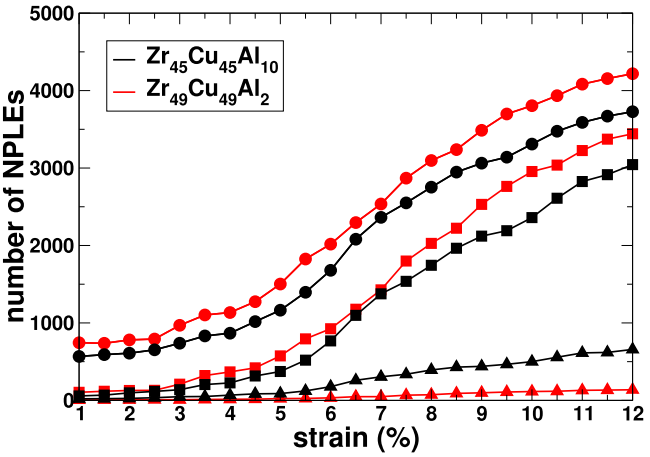

FIG. 5. Time evolution over compression of the number of non-persistent local environments (NPLEs, see text) centered at Zr (•), Cu (■), and Al (▲) atoms, counted for the two extreme NCs $Zr_{45}Cu_{45}Al_{10}$ and $Zr_{49}Cu_{49}Al_2$.

of the NPLEs have no more than 10% of broken bonds—i.e., NPLEs-BBs $\lesssim$10% for most of the cases and that is a quantity that indicates the extent of *bond exchange* in individual NPLEs. Also from Figs. S38 and S39, it is possible to verify that even at ("spurious") 12% strain, NPLEs-BBs $\gtrsim$50% are unusual.

With respect to the amounts of NPLEs, Fig. 5 brings the corresponding per-specie time evolution over compression. It goes without saying that those bare numbers reflect the stoichiometries of each NC, but it is clear that Zr atoms show higher propensity of bond breaking as already reported in the literature [53]. Moreover, it is possible to see in Table S5 that at 6% strain, the amounts of Zr- and Cu-centered NPLEs are relatively greater in the lowest Al-content NC $Zr_{49}Cu_{49}Al_2$ (in percentage terms regarding individual stoichiometries). For instance, whereas 41% of the Zr atoms are NPLEs in that latter NC, this number drops to 37% in the highest Al-content MG $Zr_{45}Cu_{45}Al_{10}$, and in the case of Cu atoms, that reduction is from 19% to 17%, respectively.

From the counting results in Fig. 5, it can be said, in principle, that the propensities for *bond exchange* of solvent Zr and Cu atoms—in the CMD simulations carried out in the present work—decrease upon alloying of Al, and this deduction has to be linked to the drastic change in the experimental mechanical behavior of the same MGs reported by Kumar *et al.* [28]. Indeed, one can see in Fig. 4 that the onset of plastic deformation takes place earlier in the MG $Zr_{49}Cu_{49}Al_2$, which is the plastic NC in Ref. [28]. Moreover, it is also noticeable in Fig. 4 that the deviation of the linear elastic behavior occurs first and is more pronounced in that lowest Al-content NC than in the MG $Zr_{45}Cu_{45}Al_{10}$ (the one found to be brittle in Ref. [28]). This is of course in line with the common perception that local bonds in that latter NC are less susceptible to be broken, since they become more strong, and maybe more covalent in addition to ionic contribution (i.e., more unidirectional and less isotropic than the metallic counterparts).

In fact, none of this is new. However, the actual missing information regarding bond strengths is the semiquantitative quantum mechanical analysis provided by the COHP method [21], which is being enabled by the ML model proposed in this work. Naturally, before the universe of thousands of predicted −ICOHP values depicted in Fig. 3, the analysis of individual cases like the Cu- and Zr-centered NPLEs discussed above will not offer the sought assessment of the propensity for *bond exchange* under the scope of mechanical loading. So in order to accomplish this, in a word, Fig. 6 brings a combination of the bonding information from Fig. 3 with the outcomes of the CMD simulations depicted in Figs. 4 and 5.

The histograms in Fig. 6 represent the distributions of mean −ICOHP values computed for the zero strain structures (those from Fig. 3) for the two extreme NCs $Zr_{45}Cu_{45}Al_{10}$ and $Zr_{49}Cu_{49}Al_2$. However, in order to keep track on *bond exchange* processes upon mechanical loading (as done for the Cu-centered NPLE in Fig. 4), all the histograms correspond to LEs existing in the 6% strained structures. So, those per-bond −ICOHP values from Fig. 3 have been cast into a per-atom (or per-LE) representation, dividing the neighbors of each central atom in three distinct groups, from which the referred mean bond strengths have been computed.

The former group comprises the neighbors of atoms whose LEs have been preserved throughout the entire uniaxial compression in the CMD simulations (up to 12%); those are labeled PLEs (an abbreviation for persistent LEs) and they make up the histograms in Figs. 6(a) and 6(d). The other two groups comprehend the already introduced NPLEs; however, they are evaluated using two distinct sets of histograms for the same entries, namely: the persistent bonds (NPLEs-PBs) and the broken bonds (NPLEs-BBs)—it refers to the histograms in Figs. 6(b), 6(c) 6(e), and 6(f). For example, the same Cu-centered NPLE depicted in Fig. 4 at 6% strain is an entry of the corresponding histograms in Figs. 6(b) and 6(c), with the corresponding arithmetic mean −ICOHP values averaged over its six PBs and its five BBs, respectively.

Back to the matter of the drastic change in the mechanical behavior of the MGs $Zr_{49}Cu_{49}Al_2$ (plastic) and $Zr_{45}Cu_{45}Al_{10}$ (brittle) reported by Kumar *et al.* [28], this experimental outcome can be revisited in the light of the histograms from Fig. 6. First of all, they show that in all LEs (persistent or not) the means computed for persistent bond strengths (PLEs and NPLEs-PBs) are very dependent on the chemical specie of the central atom, regardless of the composition of its first coordination shell (i.e., the chemical species of neighbor atoms). Moreover, the profiles of the corresponding distributions also do not depend on the NC of the studied MGs. As can be seen, for all Zr-centered PLEs and NPLEs-PBs, the mean −ICOHP values are distributed around about 1.1 eV, whereas the corresponding distributions for Cu- and Al-centered LEs are centered around about 0.8 eV and 1.5 eV, respectively. Moreover, it is noticeable that the range of mean −ICOHP values covered by those histograms for Zr- and Cu-centered LEs (between 0.5 and 1.3 eV) cover the same range of most of the moderately stronger Zr-Zr and Zr-Cu bonds as shown in Fig. 3, which allows pointing 0.8 eV as a rough minimal value for what can be stated as a persistent bond in these materials.

Still regarding persistent bonds, the differences concerning the Al content of the two ZCA alloys lie mostly in the frequencies of observations, markedly those associated to Zr-centered NPLEs-PBs—the gray histograms in Figs. 6(b) and 6(e). Pointing that these histograms are not normalized, as

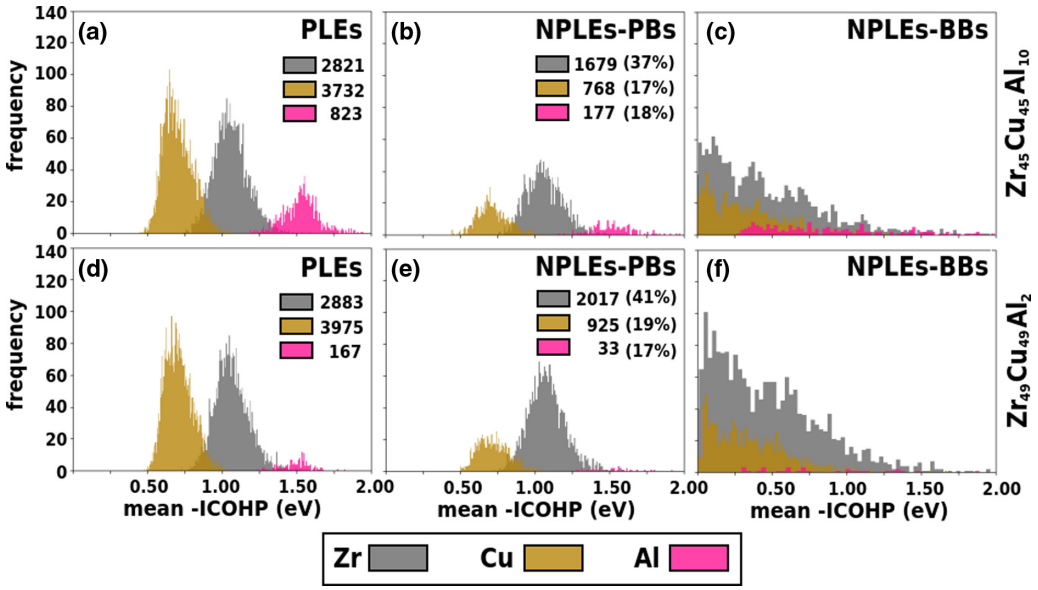

FIG. 6. Distributions of mean −ICOHP values of persistent bonds (PBs) and broken bonds (BBs) in persistent local environments (PLEs) and in nonpersistent local environments (NPLEs, see text) existing in the 10 000-atoms cells of the two extreme NCs $Zr_{45}Cu_{45}Al_{10}$ and $Zr_{49}Cu_{49}Al_2$. The number of bins in each histogram was calculated from Sturges's formula and arbitrarily multiplied by 18, whereas the corresponding colors are defined according to the species of central atoms. The total numbers of PLEs and NPLEs are shown, with the percentages of NPLEs also provided between brackets.

well as the counting results in Fig. 5. Despite this, the referred difference is still evident, which suggests that the drop from 41% to 37% of Zr-centered NPLEs caused by alloying of Al is a key factor related to the embrittlement of the MG $Zr_{45}Cu_{45}Al_{10}$, with respect to the lowest Al-content counterpart. However, especially concerning the plasticity of the MG $Zr_{49}Cu_{49}Al_2$, the role of Cu-centered NPLEs shall not be ruled out, since the associated PBs are weaker, with mean −ICOHP values distributed around about 0.8 eV.

Within this context, it is possible to show that the computed mean −ICOHP values are also correlated with the nature of the *bond exchange* processes and propose that such processes in Zr-centered NPLEs are more likely to be associated to local rearrangements with low mobility—just as depicted in Fig. S37 for the example Zr-centered NPLE. On the other hand, *bond exchange* processes going on in Cu-centered NPLEs could be said more likely to be related to a higher mobility. Without claiming to make quantitative predictions on the rheology of these MGs, the mean-squared displacements perpendicular to the loading direction (MSD_{YZ}) averaged over atoms of the same specie were also computed. These bare results are shown in Figs. S40 and S41, and they can be used to gain a qualitative view of atomic mobility upon uniaxial compression up to 5% strain. In short, it can be seen that the overall atomic mobility in the MG $Zr_{49}Cu_{49}Al_2$ is higher. However, in both NCs, Zr and Al atoms present equivalent mobility, whereas Cu atoms present greater and discernible mobility, especially in that lowest Al-content MG.

Finally, with respect to the NPLEs-BBs, one can see in Figs. 6(c) and 6(f) that they are not very dependent on the chemical specie of the central atom, concerning the ranges of the associated mean −ICOHP values, which go from 0.0 up to 0.7 eV for Cu-centered NPLEs, and are a little more extended up to about 1.0 eV for the Zr-centered NPLEs. An additional remark that is valid for all histograms in Fig. 6 is that the coordination numbers of Zr atoms are in general higher (about 18) than in Cu and Al (about 12). So, it is expected that the frequencies of broken bonds in Zr-centered LEs will be higher as well. Nevertheless, it is clear that the above-mentioned ranges of −ICOHP values verified for NPLEs-BBs match those pointed in Fig. 3 as the weakest bonds in the structures of the studied ZCA alloys (all the Cu-Cu interactions and a small fraction of the Zr-Zr and Zr-Cu bonds).

A final note, also related to the histograms shown in Fig. 6, concerns the use of the arithmetic mean −ICOHP values averaged over individual groups of neighbours in each LE. The mean has just been taken as an index able to yield a range of −ICOHP values more compatible with those computed for individual bonds shown in Fig. 3. The corresponding sums (as shown in Figs. S42 to S47) is an alternative option that could be used to discuss the outcomes of the CMD simulations in terms of chemical bonding, for instance. Moreover, in order to provide an insight into the dispersion of the −ICOHP values in each LE used to compute the associated arithmetic means, the corresponding standard deviations are also available in Figs. S48 to S53.

IV. SUMMARY AND CONCLUSIONS

This work has introduced a machine learning (ML)-based approach that brings to research in the field of metallic glasses (MGs) a feasible solution for the daunting task of extracting quantum chemical information from realistic structural models of such highly complex systems—normally containing thousands of atoms. The quantum mechanical bonding descriptors are computed with the crystal orbital Hamilton

population (COHP) method, with the accuracy of first-principles density functional theory calculations and used as a measure of bond strengths. The ML model is founded upon a Gaussian process regression framework and the smooth overlap of atomic positions (SOAP) descriptor for atomic local environments (LEs).

The ML-based approach has proven effective when applied to MGs of a prototypical alloy system, providing, firstly, an unprecedented "big picture" view of bond strengths between atoms in their structures. Next, it has been employed under the specific scope of mechanical loading, aiming at looking for insights into the drastic change in the experimental mechanical behavior upon alloying of Al reported in the literature [28]. Using ordinary descriptive statistics, the resulting overview of chemical bond strengths revealed a chemical/structural heterogeneity that is quite in line with the propensity for *bond exchange* processes verified for different types of LEs in the structures of the studied MGs. Additionally, it also enabled the assignment of such processes to migration and rearrangement mechanisms, based on identified differentiated atomic mobilities.

It is important to point out that bond strength is a key and enabling element for the development of methods for short-range order identification and atom categorization in MGs. In fact, introducing chemical bonding theory into that specific scope is not trivial, but it can have the power to bring a complementary chemical sense to those already established tools like Voronoi polyhedrons and common neighbor analysis (CNA) [58], which are based purely on structural topology. Moreover, in addition to bond strengths (−ICOHP values), there are other quantum-mechanical indicators of bonding in materials that can be learned [59] in the same way. Also, the versatility of the SOAP descriptor allows complementarity converse approaches to recover detailed information regarding electronic structure, which can be interfaced with the proposed ML model without much effort. For example, the detailed features (bonding, nonbonding, and antibonding) of the full COHP curve of a given LE in a realistic structural model used in a classical molecular dynamics simulation can be attained by looking for the best matching LE in the database of interactions.

Regarding the application of the ML model presented in this work, although rather restrained—but quite consistent—the exposed intermix of chemical/structural and dynamical inhomogeneities existing in the studied MGs certainly paves the way towards innovative approaches from the perspective of chemical bonds. There are promising applications in plenty of other contexts in which atom categorization based on the strength of chemical bonds plays a key role. Among those, it is possible to highlight: the interpretation of dynamical mechanical analysis experimental results [8], for unveiling atomic-scale mechanisms related to phenomena like viscoelastic behavior and internal friction, and also studies of shear bands [60] for gaining insights into, for instance, the role of the nonpersistent local environments in the MG $Zr_{49}Cu_{49}Al_2$ on nucleation of nanocrystals formed upon deformation, as reported in the literature [28].

ACKNOWLEDGMENTS

The author is especially thankful to Prof. Gábor Csányi from University of Cambridge, UK for providing valuable guidelines and suggestions on the implementation of the ML model. Many thanks for fruitful discussions to Prof. Jichao Qiao from Northwestern Polytechnical University, China. The author also acknowledges a fellowship from the São Paulo State Research Foundation (FAPESP; Grant No. 2016/12319–0) and thanks UFSCar and Prof. José P. Rino for support. This work used the Petaflop computing facilities and associated support services of SDumont, provided by the Brazilian National Laboratory for Scientific Computing (LNCC) in Petrópolis, RJ. It also has to be pointed out that about 20% of the CMD simulations were carried out using the high-performance computing facilities and associated support services of the National Center for High Performance Computing in São Paulo (CENAPAD-SP), which is also acknowledged.

[1] W. Klement, R. H. Willens, and P. Duwez, Nature (London) **187**, 869 (1960).
[2] D. H. Milanez, L. I. L. Faria, D. R. Leiva, C. S. Kiminami, and W. J. Botta, J. Alloys Compd. **716**, 330 (2017).
[3] M. Chen, NPG Asia Mater. **3**, 82 (2011).
[4] W. Jiao, P. Liu, H. Lin, W. Zhou, Z. Wang, T. Fujita, A. Hirata, H.-W. Li, and M. Chen, Chem. Mat. **29**, 4478 (2017).
[5] Z. Mahbooba, L. Thorsson, M. Unosson, P. Skoglund, H. West, T. Horn, C. Rock, E. Vogli, and O. Harrysson, Appl. Mater. Today **11**, 264 (2018).
[6] N. Li, J. Zhang, W. Xing, D. Ouyang, and L. Liu, Mater. Design **143**, 285 (2018).
[7] J. J. Kruzic, Adv. Eng. Mater. **18**, 1308 (2016).
[8] J. C. Qiao, Q. Wang, J. M. Pelletier, H. Kato, R. Casalini, D. Crespo, E. Pineda, Y. Yao, and Y. Yang, Prog. Mater. Sci. **104**, 250 (2019).
[9] Z. Wang and W.-H. Wang, Natl. Sci. Rev. **6**, 304 (2018).
[10] H. B. Yu, K. Samwer, W. H. Wang, and H. Y. Bai, Nat. Commun. **4**, 2204 (2013).
[11] H. B. Yu, W.-H. Wang, H. Y. Bai, and K. Samwer, Natl. Sci. Rev. **1**, 429 (2014).
[12] J. C. Qiao and J. M. Pelletier, J. Mater. Sci. Technol. **30**, 523 (2014).
[13] J. C. Qiao, Q. Wang, D. Crespo, Y. Yang, and J. M. Pelletier, Chin. Phys. B **26**, 016402 (2017).
[14] H. B. Yu, R. Richert, and K. Samwer, Sci. Adv. **3**, e1701577 (2017).
[15] Y. Yokoyama, T. Yamasaki, P. K. Liaw, and A. Inoue, Mater. Trans. **48**, 1846 (2007).
[16] Y. Q. Cheng, E. Ma, and H. W. Sheng, Phys. Rev. Lett. **102**, 245501 (2009).
[17] A. P. Bartók, S. De, C. Poelking, N. Bernstein, J. R. Kermode, G. Csányi, and M. Ceriotti, Sci. Adv. **3**, e1701816 (2017).

[18] V. L. Deringer, W. Zhang, M. Lumeij, S. Maintz, M. Wuttig, R. Mazzarello, and R. Dronskowski, Angew. Chem. Int. Ed. **53**, 10817 (2014).

[19] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).

[20] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[21] R. Dronskowski and P. E. Blöchl, J. Phys. Chem. **97**, 8617 (1993).

[22] S. Pauly, G. Liu, S. Gorantla, G. Wang, U. Kühn, D. H. Kim, and J. Eckert, Acta Mater. **58**, 4883 (2010).

[23] S. Pauly, G. Liu, G. Wang, U. Kühn, N. Mattern, and J. Eckert, Acta Mater. **57**, 5445 (2009).

[24] N. S. Barekar, S. Pauly, R. B. Kumar, U. Kühn, B. K. Dhindaw, and J. Eckert, Mater. Sci. and Eng. **527**, 5867 (2010).

[25] X. H. Wang, A. Inoue, J. F. Zhao, F. L. Kongc, S. L. Zhu, I. Kaban, M. Stoica, S. Oswald, C. Fan, E. Shalaan, F. Al-Marzouki, J. Eckert, F. X. Yin, and Q. Li, J. Alloys Compd. **739**, 1104 (2018).

[26] T. C. Pekin, J. Ding, C. Gammer, B. Ozdol, C. Ophus, M. Asta, R. O. Ritchie, and A. M. Minor, Nat. commun. **10**, 2445 (2019).

[27] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevMaterials.4.113603 for: (i) overview of all structural models, (ii) strategy for the generation of the database to train the ML model, (iii) full set of 100-atoms cells making up the database as *extended XYZ* files, (iv) details on the simulations of uniaxial compression, (v) complementary scatter plots, histograms, and distributions of −ICOHP values predicted by the ML model for additional nominal compositions.

[28] G. Kumar, T. Ohkubo, T. Mukai, and K. Hono, Scripta Mater. **57**, 173 (2007).

[29] R. S. Mulliken, J. Chem. Phys. **23**, 1841 (1955).

[30] T. Hughbanks and R. Hoffmann, J. Amer. Chem. Soc. **105**, 3528 (1983).

[31] Ľubomír Benco, Solid State Comm. **94**, 861 (1995).

[32] H. W. Sheng, W. K. Luo, F. M. Alamgir, J. M. Bai, and E. Ma, Nature (London) **439**, 419 (2006).

[33] C. C. Yuan, F. Yang, X. K. Xi, C. L. Shi, D. Holland-Moritz, M. Z. Li, F. Hu, B. L. Shen, X. L. Wang, A. Meyer, and W. H. Wang, Mater. Today **32**, 26 (2019).

[34] A. P. Bartók, R. Kondor, and G. Csányi, Phys. Rev. B **87**, 184115 (2013).

[35] S. De, A. P. Bartók, G. Csányi, and M. Ceriotti, Phys. Chem. Chem. Phys. **18**, 13754 (2016).

[36] S. Plimpton, J. Comp. Phys. **117**, 1 (1995).

[37] "Eam potentials," https://sites.google.com/site/eampotentials/Home/ZrCuAl, accessed: 2020-10-21.

[38] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. D. Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari, and R. M. Wentzcovitch, J. Phys.: Condens. Matter **21**, 395502 (2009).

[39] P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M. B. Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni, N. Colonna, I. Carnimeo, A. D. Corso, S. de Gironcoli, P. Delugas, R. A. DiStasio, A. Ferretti, A. Floris, G. Fratesi, G. Fugallo, R. Gebauer, U. Gerstmann, F. Giustino, T. Gorni, J. Jia, M. Kawamura, H.-Y. Ko, A. Kokalj, E. Küçükbenli, M. Lazzeri, M. Marsili, N. Marzari, F. Mauri, N. L. Nguyen, H.-V. Nguyen, A. O. de-la Roza, L. Paulatto, S. Poncé, D. Rocca, R. Sabatini, B. Santra, M. Schlipf, A. P. Seitsonen, A. Smogunov, I. Timrov, T. Thonhauser, P. Umari, N. Vast, X. Wu, and S. Baroni, J. Phys.: Condens. Matter **29**, 465901 (2017).

[40] P. E. Blöchl, Phys. Rev. B **50**, 17953 (1994).

[41] A. D. Corso, Comp. Mater. Sci. **95**, 337 (2014).

[42] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[43] H. J. Monkhorst and J. D. Pack, Phys. Rev. B **13**, 5188 (1976).

[44] V. L. Deringer, A. L. Tchougréeff, and R. Dronskowski, J. Phys. Chem. A **115**, 5461 (2011).

[45] S. Maintz, V. L. Deringer, A. L. Tchougréeff, and R. Dronskowski, J. Comp. Chem. **34**, 2557 (2013).

[46] S. Maintz, V. L. Deringer, A. L. Tchougréeff, and R. Dronskowski, J. Comp. Chem. **37**, 1030 (2016).

[47] S. Maintz, M. Esser, and R. Dronskowski, Acta Phys. Pol. B **47**, 1165 (2016).

[48] R. Nelson, P. M. Konze, and R. Dronskowski, J. Phys. Chem. A **121**, 7778 (2017).

[49] C. F. Bunge, J. A. Barrientos, and A. Bunge, Data Nucl. Data Tables **53**, 113 (1993).

[50] "libatoms/quip molecular dynamics framework," https://github.com/libAtoms/QUIP, accessed: 2020-10-21.

[51] C. Fan, P. K. Liaw, and C. T. Liu, Intermetallics **17**, 86 (2009).

[52] M. C. Li, M. Q. Jiang, G. Ding, Z. H. Peng, F. Jiang, L. He, and J. Sun, J. Non-Cryst. Solids **468**, 52 (2017).

[53] W. Jiao, X. L. Wang, S. Lan, S. P. Pan, and Z. P. Lu, Appl. Phys. Lett. **106**, 061910 (2015).

[54] C. E. Lekka and G. A. Evangelakis, Scripta Mater. **61**, 974 (2009).

[55] C. E. Lekka, G. Bokas, G. A. Almyras, D. G. Papageorgiou, and G. A. Evangelakis, J. Alloys Compd. **536**, S65 (2012).

[56] S. Melchionna, G. Ciccotti, and B. L. Holian, Mol. Phys. **78**, 533 (1993).

[57] M. A. Tschopp and D. L. McDowell, J. Mech. Phys. Solids **56**, 1806 (2008).

[58] X. Yue, A. Inoue, C.-T. Liu, and C. Fan, Materials Research **20**, 326 (2017).

[59] J.-Y. Raty, M. Schumacher, P. Golub, V. L. Deringer, C. Gatti, and M. Wuttig, Adv. Mater. **31**, 1806280 (2019).

[60] A. L. Greer, Y. Q. Cheng, and E. Ma, Mater. Sci. Eng., R **74**, 71 (2013).