Models and Algorithms Towards the Metallic Glasses

Vitrification at Deep Supercooled Region

Rui Su$^1$, Xuefeng Zhang$^1$, Pengfei Guan$^{1,2}$

$^1$Institute of Advanced Magnetic Materials, College of Materials & Environmental Engineering, Hangzhou Dianzi University, Hangzhou 310018, P. R. China

$^2$Beijing Computational Science Research Center, Beijing 100193, P. R. China

*Corresponding authors:

X. F. Zhang (zhang@hdu.edu.cn)

P. F. Guan (pguan@csrc.ac.cn)
Abstract

Understanding and tuning of the glass forming ability of metallic glasses requires atomic-level knowledge of structure evolutions of supercooled liquids at low temperatures. However, the accurate description of inter-atomic interactions and the dramatically increasing relaxing time in the deep supercooled region hinders a realistic description of local structure ordering, which leaves the microscopic structure evolution in the deep supercooled region rarely touched so far. We present a comprehensive analysis of realistic metallic glasses systems to demonstrate that both challenges could be tackled efficiently and simultaneously through combination of accurate neural network based inter-atomic potential and an efficient hybrid Monte Carlo algorithm incorporating swaps of atom pairs. We compare the performance of our algorithm with ordinary molecular dynamics and show that it effectively samples the phase space and generates realistic metallic glasses samples over unprecedented range of glass stabilities, which opens the door for accurate atomic-level understanding of the structure, physical properties, and moreover, the glass-forming ability of laboratory-generated metallic glasses systems.
1. Introduction

Metallic glasses (MGs) have shown great potentials for a broad range of applications as structural [1,2] or functional [3,4] materials. Like other glasses, MGs are often fabricated by rapidly quenching the high temperature melts to avoid the crystallization [5]. However, the atomic-level process of the vitrification of MGs from their parent supercooled liquids is far from clear. So far, many experiments show that the two-body structural correlations show no qualitative changes during the vitrification process [6]. Due to limited spatial and temporal resolutions of current experimental techniques, revealing the microscopic structural motifs and their evolution in the deep supercooled region of metallic glasses still faces major challenges and only became possible recently [7–9]. In contrast, computational simulations [10–14] provide effective means of exploring the structure evolution during the dynamical arresting of metallic glasses liquids, while major challenges persist due to the complexity of the multi-element interactions in common MGs and the size and timescale limitations due to limited computation resources. Thus, an insurmountable gap between the simulated and realistic metallic glasses exists and the impacts of such gap on the theoretical results can hardly be speculated.

Realistic MGs are often composed of multiple metal or metalloid elements where the local chemical bonding can be complex. *Ab initio* molecular dynamics (AIMD) have been widely used to describe the complex atom-level interactions in such realistic MGs systems. However, AIMD encounters major difficulties of small timescale (~$ps$ or a quenching rate of $\sim 10^{13} K/s$) and system size (~100 atoms), which prohibits the developing of the short-to-medium order in such simulations and introduces large statistical fluctuations in the simulation results. Additionally, constant pressure ($NpT$) simulation of MGs is expensive due to the requiring of large basis cutoff to counteract the artificial Pulay stresses [15]. Thus, people resort to classical molecular dynamics simulations with the embedded-atom method (EAM) potentials. While the EAM based classical MD allows much longer (100 – 1,000 ns) and larger (1,000 – 1,000,000 atoms) simulations of MGs, the reliability of EAM is particularly
important for comparing the simulation results with experimental ones. For example, people recently found that the EAM potential of binary Fe-P system, which has been applied to several simulations of the cavitation processes in the MGs, undergoes room temperature spinodal decomposition that seriously affects its fracture behaviors [16]. Another example is that several EAM potentials of Cu-Zr binary system overestimates the phase stability of the Laves phases, which strongly promotes the formation of Laves phase in the undercooled Cu-Zr liquids [17]. The inaccuracies of EAM potential come from two sources: firstly, EAM potentials are usually fitted against a small set of \textit{ab initio} calculated configurations, which must be carefully selected to cover the convincing systems; secondly, EAM model does not well describe the angular interactions, which is important for MGs containing metalloid or non-metal elements.

Recently, machine-learning potentials [18–20] (MLPs) trained against extensive \textit{ab initio} datasets have been applied to MGs which allows them to be simulated at the accuracy close to \textit{ab initio} simulations. As \textit{there is no free lunch}, the transferability of MLPs strongly depends on the \textit{ab initio} training database and the computational complexity is significantly larger than EAM potentials, which makes the application of MLPs on the studying the vitrification process a challenging task.

In this work, we describe an effective workflow to construct accurate and efficient neural network inter-atomic potentials (NNPs) that accurately models many structural and thermodynamical properties of MGs. The NNPs, when combined with a novel hybrid Monte Carlo method that include atom swaps (SHMC), allow us to thermalize MG models at the deep supercooled region, which allows at least five orders of magnitudes gain in the structure equilibrium time and the preparation of stable MG samples close to laboratory-made ones. Our method thus opens the door for accurate atomic-level understanding of the vitrification process of realistic MGs at the deep supercooled temperature region.
2. Methods

2.1 Neural network inter-atomic potential

The neural network inter-atomic potential is a pure mathematical potential model that maps the local chemical environment of each atom to per-atom energies and forces. To this end, fixed-length descriptor vectors are computed from local structure of atoms and passed through a feed-forward neural network for obtaining per-atom energies and forces. On the space of structural descriptors that NNP was trained, the structure energy and forces can be accurately predicted as the reference density functional theory (DFT) calculations, but at a much-reduced computation effort. We are aware that there have been many efforts on developing codes that implement NNP models for diverse chemical systems during recent years \[19,21,22\]. However, we choose to develop the specific software implementation of NNP that we called NNAP, which is designed from scratch to be user friendly and specially optimized for describing the multi-component MGs.

Like most NNPs, the potential energy of an atom configuration $\sigma$ in the NNAP implementation is expressed as the summation of per-atom energies:

$$E(\sigma) = \sum_{i} E_{\text{atom}}(\sigma_i^{R_c})$$

(1)

Where $\sigma_i^{R_c}$ defines a local environment of atom $i$ within a radial cutoff of $R_c$.

The local environment $\sigma_i^{R_c}$ cannot be directly used as neural network input layer since it does not have a fixed length. Thus, kind of function that map $\sigma_i^{R_c}$ to a fixed-length vector is required. Additionally, the function form must ensure that the atomic energy is invariant under rotation, translation, and permutation of equivalent atoms. Several types of such functions have been proposed in literatures \[19,23–26\]. In this work, we employ a self-developed descriptor which we call it as “Spherical Chebyshev” (SC) basis. The reason for developing a new descriptor root back to our practical need for simulating the multi-component metallic glass
systems: Firstly, metallic glass is usually a multi-component system that contains much more complex local structure motifs than their crystal counterparts. Secondly, to obtain realistic glass structure, long quenching time in the molecular dynamics (MD) simulations is required. Thus, the new descriptor should have systematically adjustable parameters like SOAP [26,27] (soft overlap of atom positions) to minimize the parameter tuning work (which is important in the Behler-Parrinello (BP) descriptor based NNPs [28]) and a simpler math form for fast computation as most CPU time are spent on descriptor evaluations. Ideally, the size of descriptor should not depend on the number of chemical species, since most real-life metallic glasses contain more than three chemical species.

To introduce the SC basis, we first consider the weighted and truncated local atom density function:

\[ \rho(r) = \sum_{i \neq j} w(t_j) \delta(r - r_{ij}) f_c(r_{ij}) \] (2)

where \( w(t_j) \) is a weight parameter of chemical specie \( t_j \), \( \delta(r - r_{ij}) \) is the Dirac’s delta function, \( f_c(r) \) is radial cutoff function. \( \rho(r) \) is then expanded in the spherical coordinates:

\[ c_{nlm} = \sum_{i \neq j} w(t_j) R_n(r_{ij}) Y_{lm}(\theta, \phi) f_c(r_{ij}) \] (3)

The radial expansion function: \( R_n(r_{ij}) \) takes the form of the linear transformed first-kind Chebyshev functions:

\[ R_n(r_{ij}) = T_n \left( 1 - \frac{2r_{ij}}{r_c} \right) \] (4)

The linear transform is constructed so that the \( R_n(r_{ij}) \) is orthogonal in range \([0, R_c]\) and the radial function is monolithic increasing as \( r_{ij} \to 0 \). While the expanding coefficients \( c_{nlm} \) are complex numbers and not rotational invariant, the power spectrum:

\[ P_{nl} = \frac{4\pi}{2l + 1} \sum_{|m| \leq l} c_{nlm}^* c_{nlm} \] (5)
is real and rotational invariant, which is suitable for using as descriptors.

For efficient evaluation of \( c_{nlm} \), the symmetry relation:

\[
Y_{lm}(\theta, \phi) = Y_{l-m}^*(\theta, \phi)
\]

is employed which nearly halves the time of evaluation the per-atom descriptor vectors and their derivatives.

While Eq. 2-5 already allow us to define a basis for multi-component systems if the weighting function \( w(t_j) \) is properly selected [29], we found the using of a dual basis [24] for the multicomponent system is beneficial since the weight function can be “automatically” learned from training data without tuning the form of weight functions. For the dual basis representation, we first calculate the unweighted power spectrum: \( P_{nl}^s \), where \( w(t_j) \equiv 1 \), then the weighted power spectrum \( P_{nl}^c \) is calculated by using a simple weight function:

\[
w(t_j) = (-1)^{t_j-1} t_j, \quad t_j = 1, 2, \ldots, n_t
\]

where \( n_t \) is the number of chemical species in system. It is easily to conclude that the descriptor size is independent on \( n_t \) for \( n_t \geq 2 \).

For truncating descriptor to zero at the spherical cutoff \( r_c \), the cutoff function \( f_c(r) \) is used. Currently, two kinds of cutoff functions can be used in the NNAP code. The first one is:

\[
f_c(r) = \begin{cases} 
\exp \left( \frac{\gamma (r/r_c)^2}{(r/r_c)^2 - 1} \right), & r < r_c \\
0, & r \geq r_c
\end{cases}
\]

Where \( \gamma \) is a positive adjustable parameter for fine tuning the shape of cutoff functions. The second one is a simple parameter-free polynomial cutoff function:

\[
f_c(r) = \begin{cases} 
(1 - (r/r_c)^2)^4, & r < r_c \\
0, & r \geq r_c
\end{cases}
\]

One can easily prove that the cutoff functions 1 and 2, in together with their high-order derivatives up to \( \infty \) and three orders are differentiable and equals to zero at \( r_c \), respectively.
Thus, both cutoff functions ensure smoothly changes of atomic forces at cutoff when atom come into or out of the cutoff radius, which is important for reserving energy stability in MD simulations.

**Fig. 1** Illustration of NNP construction workflow for MGs in NNAP. $S_i$ and $C_i$ refers to the structure-only and the chemical-weighted part of descriptors, respectively.

**Fig. 1** illustrates the basic workflow for NNP construction procedure for MGs. The whole workflow is designed as a two-stage one containing bootstrapping and refining stages with aims to enhance sample diversity of the final training dataset and minimizing time-consuming AIMD samplings.

In the bootstrapping stage, we first construct preliminary database containing perturbated crystal phases and liquid configurations from high-temperature AIMD simulations. The CUR algorithm [28] is used to select a subset of AIMD trajectory that minimize the configuration correlations. The preliminary database is then evaluated by the density functional theory (DFT) codes and send for a $k$-fold cross validation training which give the initial $k$-NNPs ensemble. The $k$-NNPs ensemble is then iteratively optimized with an active-learning strategy: firstly, we define the ensemble error for a configuration as:
\[ \epsilon = \sqrt{\frac{1}{k} \sum_{i=1}^{k} (E_i - \bar{E})^2} \]  

(9)

where \( E_i \) and \( \bar{E} \) are per-atom potential energy from \( i \)-th NNP and ensemble average.; secondly, NNP-MD simulations are performed to locate configurations with large ensemble errors employing an online active-learning algorithm (Algorithm 1), which enhances the DFT training database. In practice, several iterations for each chemical composition are required until stable MD simulations can be performed.

**Algorithm-1:** Online active selection of training samples

| Input: | Initial configuration: \( P_0 \), temperature: \( T \), number of MD steps \( N \), number of data collection steps \( N_{collect} \), max ensemble error: \( \epsilon_{max} \), minimal sampling interval: \( N_{min} \), minimal sampling threshold: \( Z_{min} \). |
| Output: | Collection of sampled MD configurations: \( S \) |

Set \( S \leftarrow {} \)

Collection of ensemble error: \( \epsilon \leftarrow {} \)

Last MD step of collection: \( N_{last} \leftarrow 0 \)

for \( i = 1, 2, ..., N \) do

Take one MD step: \( P_{i-1} \rightarrow P_i \)

Compute current ensemble error: \( \epsilon_i \)

if \( \epsilon_i > \epsilon_{max} \): stop

else: add \( \epsilon_i \) to \( \epsilon \)

if \( i > N_{collect} \):

Compute current sample score: \( Z_{score} = (\epsilon_i - \text{mean}(\epsilon))/\text{std}(\epsilon) \)

if \( |Z_{score}| > Z_{min} \&\& i - N_{last} > N_{min} \):

Append current configuration \( P_i \) to \( S \)

\( N_{last} = i \)

return \( S \)

In the refining stage, we employ different methods to explore the current NNP potential energy landscape. Currently, we employ multiple methods including random structure search (RSS), MD, SHMC, basin hopping (BH), activation relaxation technique *nouveau* (ARTn) for sampling different parts of PEL. The RSS is used to enhance sampling in ordered crystal phase space which has been previously used for MLP training [30]. The MD quenching is used to sample instantaneous and inherent configurations with quenching rate larger than \( 10^{10} K/s \). The SHMC is used to obtain the deep supercooled MG samples with effective quench rate close
to laboratory made samples, a detailed description of SHMC will be given in following text. BH runs are mainly used for MG clusters which samples PEL minima around the global minimum [31]. ARTn runs provides the saddle point samples of MGs which is important for understanding the MGs relaxations [32]. In practice, we employ ARTn at last when the PEL is already well refined by other methods and the saddle points are close to the real ones. For each part of refining, we iteratively perform training and DFT validation cycles until the prediction error is comparable with training error.

2.2 Hybrid Monte-Carlo simulations with atom swapping (SHMC)

Swapping Monte Carlo simulations have shown great ability to reach equilibrium at deep supercooled temperature region for poly-dispersed model glass systems [33]. Given that most MGs systems have three or more chemical elements, it is straightforward to think that the same method might be appliable for NNP based simulations for realistic MGs. However, the direct application of swapping MC to MGs face major obstacles. Standard MC simulation consists with many sweeps where in each sweep atoms are randomly moved in sequential or random, which hinders efficient parallel simulations. For NNP simulations, the computation of the energy change due to single atom MC move requires re-evaluation of \( N \) times atom structure descriptor and the NN forwarding, where \( N \approx 50 - 100 \) closing to the average number of per-atom neighbors. As a single MC sweep requires \( 3N_{\text{atoms}} \) moves, the NNP-MC simulation for medium-to-large MG samples containing several thousand of atoms is impractical. To enhance parallel efficiency, the hybrid MD/MC method that combines MD with atom swap MC has been suggested [34]. However, careful selection of simulation parameters is required so that the resulting sampling still follows the canonical probability distribution.

In this work, we suggest combining the hybrid Monte Carlo (HMC) method [35,36] with atom swapping which can be effectively parallelized over hundreds of CPU cores. Instead of local moves in standard MC sweeps, HMC make global update of atom positions by taking short-time MD steps with initial atom velocities drawn from a Gaussian distribution at temperature \( T \). The MC update is accepted with a transition probability:
where \( x; v \) and \( x'; v' \) are successive points in the phase space, \( \beta = 1/k_B T \) is the Boltzmann factor and \( \Delta H \) is the change of Hamiltonian (total energy) due to the MD update. The detailed balance of such MD update is ensured if the numerical integration of the motion equations is time-reversible and area-preserving [36]. In this work, we use the leap-frog method for integrating the motion equations. We note that since the detailed balance is obeyed, the whole simulation can be viewed as a Markov process that the canonical probability distribution of the system is solely determined by Eq. 10, which does not depend on the MD parameters. Unlike model glass systems, simulations of MGs are usually performed under the isothermal isobaric (NpT) ensemble. While generalizing current MD schemes to NpT ensemble is possible, we use an additional volume-scaling MC step [37] to maintain system pressure \( P \):\[
P_{\text{acc}}(V_0 \rightarrow V) = \min\{1, \exp\left[-\beta (\delta U + P (V - V_0) - N\beta^{-1} \ln(V/V_0))\right]\}
\] (11)

Atom swapping MC is performed by exchanging chemical symbols between two atoms. Here, two methods can be used: The first one (method I) is randomly selecting two atoms and exchanging their symbols if they are in different types. Another more complex way (method II) is randomly choosing two chemical elements A and B of current composition at first, and randomly choosing one atom of element A and B in the configuration. Both methods meet the detailed balance condition and are implemented in NNAP. While method I is used in swap MC of model systems [33,34], we found the method II results in higher acceptance ratio in practical MGs simulation as it is insensitive to the chemical composition of MGs.

Since the potential energy changes due to atom swapping is evaluated many times, full evaluation of the system potential energy at each swap step is unacceptable. Luckily, full evaluation of potential energy is not necessary for large systems since only part of per-atom descriptors changes due to swapping. Assuming the maximum number of neighbors is \( N_{\text{neigh}} \), the maximum number of per-atom energy evaluations \( N_{\text{eval}} = \min\{N_{\text{atoms}}, 2 \times (N_{\text{neigh}} + 1)\} \) if the neighbors of two atoms do not overlap. Thus, swap steps only require nearly constant
computing time. To further reduce the computing cost, the per-atom local potential energies are calculated and saved in memory before swapping, which are updated in case of successive swaps. This optimization saves one energy evaluation for each swap attempt and halves the number of potential energy evaluations in the atom swapping step.

In the current implementation, one MC sweep of the SHMC simulation consists of three sequential parts: MD, volume scaling (only for the $NpT$ runs) and atom swapping. The number of atom-swapping attempts is determined by a single parameter $r_{\text{swap}}$ as $N_{\text{swap}} = r_{\text{swap}} \times N_{\text{atoms}}$.

2.3 Details of NNP training of Zr-Cu-Al system

**Fig. 2** Composition of DFT training database for Zr-Cu-Al MGs. a: Bar plot showing the number of samples with different generating methods; b: Ternary plot of sampled chemical compositions in the full Zr-Cu-Al chemical space.

In this work, we select the simplest ternary Zr-Cu-Al system as representative MGs, which has been extensively investigated in the last decade [38–41], for illustrating the NNP-SHMC method. The training dataset is generated with the workflow described in Sec. 2.1. **Fig. 2** shows the dataset compositions of generating methods and chemical compositions, respectively. In summary, the DFT training database contains 36,187 configurations that corresponds to 1,794,343 atom environments. **Fig. 2** shows the compositions of the database. The initial part corresponds to the stage I in our workflow and others corresponds the state II. To enhance the
transferability of the final NNPs, we generate Zr-Cu-Al samples in the whole chemical space (Fig. 2b).

All DFT calculations were performed by the Vienna \textit{ab initio} simulation package [42] (VASP). The electron-nuclear interaction is described by the all-electron projector augmented wave [43] (PAW) method. The PBE [44] exchange-correlation functional is used to describe the electron-electron correlation energies. A planewave cutoff of 420 eV is used for expanding the wave functions. A uniform K-point mesh with minimal spacing 0.3 Å$^{-1}$ is used for reciprocal space integrations of periodic structures, otherwise a single $\Gamma$ point is used. The DFT energy convergence criterion is set to $1 \times 10^{-5}$ eV.

The SC basis requires three parameters: $n_{\text{max}}$, $l_{\text{max}}$ and $r_\text{c}$ for each specie. In this work, all species share the same basis setup of: $n_{\text{max}} = 5$, $l_{\text{max}} = 6$ and $r_\text{c} = 6.5$ Å. The polynomial cutoff function (\textbf{Eq. 8}) is employed. For each specie, we employ a three-hidden-layer NN architecture of $84 \times 16 \times 16 \times 16 \times 1$, which results 5,763 degrees of freedom in total. The activation function is set to the sigmoid linear unit (SiLU) function: $\text{SiLU}(x) = x/(1 + e^{-x})$. Rather than achieve maximum accuracy, the hyper-parameters in NNP training are manually optimized for balancing the MD performance and fitting accuracy in this work. All trainings were performed by the LBFGS-B [45] optimizer implemented in NNAP and a $\ell_2$-regulation factor $\lambda = 0.002$ is used to reduce overfitting. For the bootstrap stage, a 10-fold cross-validation run is performed to generate the NNP ensembles. Since our database contains many atoms, training against both of energies and atom forces is impractical due to the huge computation cost [46], and we use only potential energies for the NNP trainings.

3. Results

3.1 Testing on the Zr-Cu-Al NNP

\textbf{Fig. 3} shows the prediction results of selected Zr-Cu-Al NNP on the testing set. The NNP achieves energy and force mean absolute errors (MAE) of 5.09 meV/atom and 0.15 eV/Å on the testing test, respectively. In comparison, the energy and force MAE on the training set is
4.42 meV/atom and 0.149 eV/Å in respective, which shows no overfitting.

![Fig. 3](image_url) Energies (a) and force components (b) fitting errors on the testing set. The perfect fitting: $y = x$ is shown as dashed lines. Data points in (b) are colored by their local number densities.

We compare the Zr-Cu composition diagrams obtained from different potential models in Fig. 4 to validate the NNP performances on the crystalline phases. The full results for all investigated crystalline phases are listed in Table S1. The NNP results are compared with the ones from widely used Zr-Cu-(Al) EAM potentials: MSK [47], MKOSYP [48], MSZWH [17] and CMS [40] (abbreviated by author names in respective). Among all of the investigated models, our NNP model gives almost identical predictions of Zr-Cu phase energies as DFT, which a MAE of ~2 meV/atom (see Table S1). Surprisingly, none of previous EAM models can fully predict the correct Zr-Cu phase stabilities. The CMS potential, as well as MSK and MKOSYP ones (not shown), is well known [17,49] to incorrectly predict the ZrCu and ZrCu$_2$ phases as stable ones. While the recently developed MSZWH potential gives correct phase boundaries, the derivation to DFT results is very large for the Cu-rich Zr$_{14}$Cu$_{51}$ phase which owns a complex crystal structure that is hard to be described by the simplified EAM model. Importantly, none but the NNP model can predict the correct energy sequence of ZrCu$_2$ metastable phases (Fig. 4b). The ZrCu$_2$ composition is a well-known glass forming liquid [50] in the Cu-Zr phase diagram. Thus, the energy sequence of metastable phases is very important to study the vitrification and crystallization processes. As shown in Fig. 4b, all EAM models
predict some (C14 or C15) of the Laves phases as the most stable ZrCu$_2$ phase, while DFT and NNP show that the $\sigma$ phase is the most stable one. From the view of local structures, the Laves phases (C14, C36 and C15) all correspond to Cu-centered icosahedrons while the $\sigma$ phase corresponds to distorted bcc-like structure and does not contain any icosahedrons. Thus, the EAM potentials might overestimate the stability of such icosahedrons in simulations. However, the detailed effects of different potentials on the GFAs and crystallization paths of ZrCu$_2$ are clearly out of the scope of this work and will be presented in future.

**Fig. 4** Cu-Zr composition diagrams (a) and decomposition energies $E_d$ (b) of Zr$_2$Cu phases as obtained from varies potentials. For the CMS and MSZWH potentials, only stable phase boundaries are shown in (a). The ZrCu$_2$ decomposition energy is defined as: $E_d \equiv (13E_{ZrCu_2} - 2E_{Zr_3Cu_8} - E_{Zr_7Cu_{12}})/39$.

The B2-ZrCu phase corresponds to another glass forming composition [50] on the phase diagram and consists as the parent structure for Zr-Cu-Al MGs. The B2-ZrCu has been thoroughly investigated for the thermodynamical evolution during the crystallization [39,51]. Previously modeling [52] has shown that the melting temperature of the B2 phase is strongly underestimated by some EAM models. In **Fig. 5**, we presented the NNP-predicted free energy curve and melting temperature, which determine the thermodynamical stable region of the B2-ZrCu phase. The NNP, EAM and experimental results are summarized in **Table 1**. In this work, the free energies are computed by the Frenkel-Ladd method [53] at various temperatures and
the melting temperature is calculated through the “interface pinning” method [54,55]. The details on the free energy and melting temperature calculations are presented in the supplemental materials. Of all examined potentials, only NNP and MSZWH can describe the meta-stability of the B2 phase. However, the actual numbers of the MSZWH predictions departure far away from the experimental measurements [39]. In comparison, NNP predictions are in good agreement with the experiments, which shows that the NNP provide us accurate predictions on the Cu-Zr phase stabilities.

![Free energy difference plot showing the thermodynamically stable region of B2 phase.](image)

**Fig. 5** Free energy difference plot showing the thermodynamically stable region of B2 phase. The red dots show the calculated points, and the blue line shows the quadratic fitting.

| Table 1 | Phase decomposition temperature $T_d$ and melting temperature $T_m$ for B2-ZrCu. |
|---------|----------------------------------------------------------------------------------|
|         | Exp.     | NNP     | CMS    | MSZWH  | MKOSYP | MSK    |
| $T_d$ (K) | 985-1003 | 1090    | -      | 190    | -      | -      |
| $T_m$ (K) | 1209-1229 | 1197    | 1347   | 989    | 1353   | 1741   |

NNP-calculated pair correlation functions $g(r)$ and structure factors $s(q)$ for the Zr$_{54}$Cu$_{46}$ and Zr$_{46}$Cu$_{46}$Al$_8$ systems are used for validating the NNP performance on the structure properties of MGs. **Fig. 6a** shows the calculated $g(r)$ in comparison with the DFT-MD results. The DFT results are obtained for a 100-atom configuration with a time duration of 15 ps, while the NNP results are obtained for a 1000-atom configuration with a time duration of
200 ps. As shown in Fig. 6a, the NNP results are in good agreements with the DFT ones for both Zr_{54}Cu_{46} and Zr_{46}Cu_{46}Al_{8}. Experimentally, $s(q)$, rather than $g(r)$, are obtained from scattering experiments. In Fig. 6b we compare calculated $s(q)$ curves to the measured results [56,57]. For the direct comparison to the Zr_{46}Cu_{46}Al_{8} as-cast sample, we use the SHMC-generated 1000-atom MG sample for MD simulation. This sample currently corresponds the most stable one that can be prepared by the NNP potential. More details on the SHMC simulations will be presented in the following text. In both compositions, the results compare well to the measurements, while slightly underestimated first-peak strength can be observed. Given that the correspondence to DFT results is good, the minor differences might come from the sample-size differences between experiments and simulations. However, the splitting of second $s(q)$ peak in Zr_{46}Cu_{46}Al_{8} is well reproduced by NNP simulation. It has been shown that the splitting of the second peak corresponds to development of medium-range order in MGs [41], which hints that our NNP also provides good description on the medium-range order of Zr-Cu-Al MGs.

**Fig. 6** Structure properties for Zr-Cu-Al MGs. (a): Pair correlation functions $g(r)$ calculated by NNP (solid blue lines) and DFT (open red circles), respectively. (b): Static structure factor $s(q)$ obtained by NNP-MD (solid blue lines) and experimental measurements (open red circles), respectively.
Transferability of trained NNPs has been under debating since the formulation does not incorporate physical information which might result unphysical behaviors under some limiting situations. Pun et. al. has shown that some Al-NNPs shown unphysical energy-volume curves when the crystals are highly compressed, and they suggest that a physical-informed NNP which could solve such problem [58]. While their method offers new insights on this problem, we found that our NNP does not suffer from such disadvantages. As shown in Fig. 7, the rigid E-V curves for four crystalline phases are plotted against the per-atom volume in a range of (0, 150). It can be clearly seen that the energy curves are very smooth and the divergence of energy near \( V \to 0 \) is obvious. Thus, our NNP should be expected to have the same level of transferability as the physical potentials like EAMs.

![Fig. 7 Energy-volume curves in the full range of crystal volumes for B2-ZrCu (red solid lines), bcc-Zr (blue solid lines), fcc-Cu (purple solid lines) and fcc-Al (green solid lines). The inset plot shows E-V curves in the extended energy range.](image)

When applying NNP to the real-world MD simulations, the major difficulty is the computation complexity of its math forms. In our NNP model, profiling shows that most CPU instructions are spent on the evaluation of the descriptor and their derivatives. In general, the
computation complexity of descriptor computation scales linearly as $n_{\text{max}}$ and quadratically as $l_{\text{max}}$. For current Zr-Cu-Al NNP, we measured the NNP-MD performance as $\sim 0.19 \text{ ms/timestep/core}$ on an Intel(R) Xeon(R) Gold 6140 CPU. In comparison, we also test the BP-NNP performance as implemented in the AENET package [18], where the MD performance is $\sim 2.0 \text{ ms/timestep/core}$ with a published BP parameters set [18] but adjust the cutoff to the same one of our NNP. All testing are performed through the LAMMPS [59] interfaces to the underlying NNP codes. Since the BP-NNP is trained on the same database with a smaller descriptor size ($N = 72$) and neural network ($72 \times 10 \times 10 \times 1$), our NNP implementation is at least one order of magnitude faster than the previous BP-NNP implementation. However, the EAM potential shows a performance of $\sim 0.002 \text{ ms/timestep/core}$, which is about two orders faster than NNP. Fortunately, the NNP parallel scaling is often better than EAM for small-to-medium systems due to the constant communication time, which makes up for some total wall time.

3.2 Well-annealed Zr-Cu-Al samples with NNP-SHMC

As have been discussed in above, NNP-MD is about two orders slower than the widely used EAM potential. Thus, NNP-MD can only afford quenching rates as low as $10^{10} K/\text{s}$ for the MD quenching. For example, a $10^{10} K/\text{s}$ quenching run costs 41 hours with 180 CPU cores for a 1,000-atom $\text{Zr}_{46}\text{Cu}_{46}\text{Al}_{8}$ configuration. Such a limitation on the MG sample stability hinders the preparation of stable MG samples which are critical to understanding the structure and mechanical properties evolutions during the vitrification of MGs.

In Sec. 2.2, we describe the SHMC method which has the great ability for obtaining stable MG samples. The SHMC method is implemented as a standalone part of the NNAP code. Here, we use a 1000-atom $\text{Zr}_{46}\text{Cu}_{46}\text{Al}_{8}$ MG configuration to put the NNP-SHMC method into real-world tests. We measure the $\alpha$-relaxation time $\tau_\alpha$ with the self-intermediate-scattering function (SISF) $F_3(q, t)$:
\[ F_s(q, t) = \frac{1}{N} \sum_{j} e^{i q [r_j(t) - r_j(0)]} \]

where \( q = 2.67 \text{ Å}^{-1} \) is used and the \( \alpha \)-relaxation time is defined as \( F_s(q, \tau_\alpha) = e^{-1} \).

SHMC simulations requires three main parameters: \( N_{MD}, t_{MD} \) and \( r_{swap} \), which defines the number of MD steps, timestep size and number of swap attempts at each MC sweep. As we have discussed in Sec. 2.2, any combination of parameters would satisfy the detailed balance condition and the final equilibrium is not altered by those parameters, but the structural relaxation time \( \tau_\alpha \) would be strongly affected by the detailed parameter values and must be optimized before performing SHMC simulations.

**Fig. 8** Optimization of SHMC parameters. (a): Optimization of \( t_{MD} \) and \( N_{MD} \). (b): Optimization of \( r_{swap} \). Intersect plots show SISF plots of different parameter sets, respectively. The optimized SHMC parameter sets are labeled by red arrows.

**Fig. 8** shows our optimization procedure for the SHMC parameters. We firstly fix the \( r_{swap} = 0 \) and optimize the MD parameters as shown in **Fig. 8a**. Unlike MD simulation, very large timestep can be used in SHMC since the numerical stability is not important here. However, the acceptance ratio decreases significantly as \( t_{MD} \) and \( N_{MD} \) grows up, and the
maximum performance of SHMC shows up at some specific values. Here, we can see that $t_{MD} = 16 \text{ fs}$ approximately corresponds to the timestep size of smallest relaxation time, which is independent to the selection of $N_{MD}$ values. On the other hand, increasing $N_{MD}$ from 4 to 8 always decreases the relaxation time. However, we stop at $N_{MD} = 8$ without further increasing its value, as the relaxation time is already smaller than the MD one and increasing $N_{MD}$ only offers minor decrease of $\tau_\alpha$ but significantly increases the computation time of each MC sweep. Secondly, we optimize the ratio of swap attempts $r_{swap}$. $r_{swap}$ determines the number of swap attempts by $N_{swap} = r_{swap} \cdot N_{atoms}$. As shown in Fig. 8b, increasing $r_{swap}$ to 0.3 significantly reduces $\tau_\alpha$, while further increasing does not help much. For current simulation temperature $T = 950 \text{ K}$, introducing atom swapping results a four-fold acceleration in comparison to the swapping-free run. As we will show in the following text, the acceleration ratio is temperature-dependent but the optimization of $r_{ratio}$ is not significantly affected by the simulation temperature $T$ and $T = 950 \text{ K}$ is almost arbitrarily selected here.

![Fig. 9 (a): Structure relaxation time $\tau_\alpha$ with inverse temperature $1000/T$ for MD simulations (red open squares) and SHMC ones with $r_{swap} = 0$ (red solid triangles) and 0.3 (blue balls), respectively. An effective time scale $t_0 = 25 \text{ fs}$ is used to scale the SHMC runs.](image)
Vogel–Fulcher–Tammann (VFT) fittings for MD (red solid line) and SHMC (blue solid line) runs are presented to estimate the glass transition temperatures corresponding to $\tau_\alpha \approx 100 \, s$. (b): Snapshot of SHMC-equilibrated sample at $T = 740 \, K$.

We analyzed the structure relaxation time $\tau_\alpha$ evolutions for different simulation protocols as shown in Fig. 9a. Extrapolation of the Vogel–Fulcher–Tammann (VFT) fitting on the MD results to $\tau_\alpha = 100 \, ps$ predicts the glass-transition temperature $T_g = 705.8 \, K$, which is in excellent agreement with experiment result that $T_g^{exp} \approx 709 \, K$ [57]. As SHMC are performed by consequent MC sweeps, effective time scale must be estimated for direct comparison to the MD results. As shown in Fig. 9a, an effective time scale $t_0 = 25 \, ps$ for each MC sweep can be obtained that shows nicely scaling collapse between the MD and swapping-free SHMC results. By introducing atom swapping, the glass-transition temperature is reduced by about 97 K, and the equilibration processes of MG samples are accelerated at low temperatures. The acceleration ratio is found to be strongly temperature dependent. For the lowest investigated temperature $T = 740 \, K$, the acceleration ratio is about $10^5$ folds. Consider that the NNP-SHMC run at 740 K is performed for $2 \times 10^6$ MC sweeps that cost about 5 days using 108 CPU cores, the MD preparation of the MG samples with the same stability will cost about 1,370 years that greatly exceeds current computer power. As shown in Fig. 9b, no crystallization can be observed even at such a large energy stability.

The convention of the glass physics community is using the “parent temperature” $T_p$, where the glass samples are prepared, to characterize the glass stability of samples [33,60,61]. However, material scientists prefer to use the quenching rate $Q$ which connect more closely with experiments. Since the time in a SHMC run is not well defined, we estimate the effective quenching rates of MG samples by comparing their inherent structure energies $E_{IS}$ against MD quenching runs. The inherent structure energy $E_{IS}$ has been previously used to define the absolute “glass stability” or “annealing degree” [62] which also well correlates with the parent temperature $T_p$. The MD and SHMC results are presented in Fig. 10. Since the analytical relation between $E_{IS}$ and $Q$ is not known, we use two functional forms for extrapolating the
effective $Q$ values. Firstly, we use the simplest logarithmic extrapolation. The extrapolation is only performed for the last three points as the MD runs at higher quenching rates breaks the logarithmic relations. Alternatively, the power-law relation as suggested in reference [63] is used to make use of full MD data points. We can see that both extrapolations differ greatly at lower $Q$ values. For the sample equilibrated at $T_p = 740$ K, logarithmic extrapolation yields an effective $Q_e \approx 10^7$ K/s while the power-law extrapolation yields $Q_e \approx 10^4$ K/s. However, power-law extrapolation cannot estimate $Q_e$ for the SHMC quenched sample that corresponds to the most stable MG sample in this work, as the power-law function is always larger than the sample’s $E_{IS}$. The departure of the two extrapolations comes from the lack of MD reference data at lower quenching rate which cannot be obtained due to the wall time limit for nowadays computers. However, the simple logarithmic extrapolation should still provide the safe estimation of the upper limit of $Q_e$ values. Note that for the most stable MG sample, the $Q_e$ is estimated to be $\sim 1 \times 10^5$ K/s, which is smaller than the maximum experimental quenching rate ($\sim 10^6$ K/s) of the melt-spinning MG ribbons [64]. Thus, our NNP-SHMC method allows us to prepare the MG samples at the DFT-level quality that are close to laboratory-made MG samples.

Fig. 10 Estimating effective quenching rates for SHMC runs. Red balls present the inherent
structure energies $E_{IS}$ for samples from five MD quenching rates: $10^{14}, 10^{13}, 10^{12}, 10^{11}$ and $10^{10} \text{K/s}$. Error bars are calculated from ten independent MD runs for all but the $10^{10} \text{K/s}$ one, which is calculated from five independent runs. Black solid line shows the logarithmic extrapolation of the lower part of MD data. Green solid line shows the power-law extrapolation. Extrapolated points for the equilibrated sample at $T_p = 740 \text{K}$ and the SHMC quenched one are shown as blue filled diamonds. SHMC quenching is performed from 740 K to 300 K at $P = 0 \text{GPa}$ with $10^6$ MC sweeps.

4. **Discussion and perspectives**

Accurate atom models of MGs at deep supercooled temperature region is the key to understand the structural origin of the dynamical arresting [6,11]. Due to the limitation of DFT calculations, EAM based potentials are developed that contribute most of our current knowledge of the vitrification process. Developing accurate and transferable EAM potentials requires hard working on optimize the model parameters and is error-prone to the selection of reference datasets [17]. In this work, we suggest and detailly analyze the novel NNP-SHMC method to generated accurate and well-annealed MG samples. With a simple and computational efficient descriptor form, we achieve accurate and efficient NNP modeling of the representative Zr-Cu-Al MGs with a diverse DFT training database that expands extended energy and composition space. The accuracy of NNP is illustrated by comparing many thermodynamical and structural properties of Zr-Cu-Al crystalline and glass phases against experimental results. With the powerful SHMC simulation, we show that the NNP can be used to achieve equilibrium of a 1,000-atom MG within several days closing to $T_g$, which might need thousands of years for conventional MD simulations. The most stable MG sample shows an effective quenching rate that is comparable with the laboratory-made MG samples, which allows direct comparison of NNP-generated samples with future experiments.

In continuing work, it will be interesting to obtain larger MG samples with NNP-SHMC as the size-effect would be large for understanding the mechanical properties of MGs. Our preliminary work shows that an 8,000-atom sample of similar stability can be obtained within
a month which is well within the nowadays computer resource limit. Such a large MG samples would show less size-effects and suited for most research purposes. On the other side, our NNP specially contains clusters results in the DFT database, which would help to understand the surface dynamics of MG surfaces which could be problematic for previous EAM models. Extending NNP-SHMC to other multi-component MGs or high entropy alloys (HEA) is also straightforward and should helps to clarify the solidification processes of MGs or HEAs. Finally, as NNP-SHMC generate samples closing to the DFT reference, it could provide configurations for electron structure calculations which might help to understand the chemical bonding changes in the deep supercooled region. Thus, we expect many related works to be published that extend people’s knowledges in this old yet important area.

Acknowledgements

This work is supported by the NSF of China (Grant Nos. 5211101002 and U1930402) and the National Key R&D Program of China (Grant No. 2017YFA0303400). R.S. acknowledges the Young Scientists Fund of the National Natural Science Foundation of China (No. 51801046). P.F.G. and R.S. acknowledge the computational support from the Beijing Computational Science Research Center (CSRC). Additional computational resources from the Institute of Advanced Magnetic Materials of the Hangzhou Dianzi University are also gratefully acknowledged.
References

[1] M. Chen, *Mechanical Behavior of Metallic Glasses: Microscopic Understanding of Strength and Ductility*, Annu. Rev. Mater. Res. 38, 445 (2008).

[2] M. D. Demetriou, M. E. Launey, G. Garrett, J. P. Schramm, D. C. Hofmann, W. L. Johnson, and R. O. Ritchie, *A Damage-Tolerant Glass*, Nat. Mater. 10, 123 (2011).

[3] M. Zhao, K. Abe, S. Yamaura, Y. Yamamoto, and N. Asao, *Fabrication of Pd–Ni–P Metallic Glass Nanoparticles and Their Application as Highly Durable Catalysts in Methanol Electro-Oxidation*, Chem. Mater. 26, 1056 (2014).

[4] Y. C. Hu et al., *A Highly Efficient and Self-Stabilizing Metallic-Glass Catalyst for Electrochemical Hydrogen Generation*, Adv. Mater. 28, 10293 (2016).

[5] W. Klement, R. H. Willens, and P. Duwez, *Non-Crystalline Structure in Solidified Gold–Silicon Alloys*, Nature 187, 869 (1960).

[6] C. P. Royall and S. R. Williams, *The Role of Local Structure in Dynamical Arrest*, Phys. Rep. 560, 1 (2015).

[7] H. W. Sheng, W. K. Luo, F. M. Alamgir, J. M. Bai, and E. Ma, *Atomic Packing and Short-to-Medium-Range Order in Metallic Glasses*, Nature 439, 419 (2006).

[8] A. Hirata, L. J. Kang, T. Fujita, B. Klumov, K. Matsue, M. Kotani, A. R. Yavari, and M. W. Chen, *Geometric Frustration of Icosahedron in Metallic Glasses*, Science 341, 376 (2013).

[9] A. Hirata, P. Guan, T. Fujita, Y. Hirotsu, A. Inoue, A. R. Yavari, T. Sakurai, and M. Chen, *Direct Observation of Local Atomic Order in a Metallic Glass*, Nat. Mater. 10, 28 (2011).

[10] P. F. Guan, T. Fujita, A. Hirata, Y. H. Liu, and M. W. Chen, *Structural Origins of the Excellent Glass Forming Ability of Pd$_{40}$Ni$_{40}$P$_{20}$*, Phys. Rev. Lett. 108, 175501 (2012).

[11] Y. C. Hu, F. X. Li, M. Z. Li, H. Y. Bai, and W. H. Wang, *Five-Fold Symmetry as Indicator of Dynamic Arrest in Metallic Glass-Forming Liquids*, Nat. Commun. 6, 8310 (2015).

[12] B. Wang, B. S. Shang, X. Q. Gao, W. H. Wang, H. Y. Bai, M. X. Pan, and P. F. Guan, *Understanding Atomic-Scale Features of Low Temperature-Relaxation Dynamics in Metallic Glasses*, J. Phys. Chem. Lett. 7, 4945 (2016).

[13] B. Xu, M. L. Falk, J. F. Li, and L. T. Kong, *Predicting Shear Transformation Events in Metallic Glasses*, Phys. Rev. Lett. 120, 125503 (2018).

[14] Y. C. Hu, Y. W. Li, Y. Yang, P. F. Guan, H. Y. Bai, and W. H. Wang, *Configuration Correlation Governs Slow Dynamics of Supercooled Metallic Liquids*, Proc. Natl. Acad. Sci. 115, 6375 (2018).

[15] G. P. Francis and M. C. Payne, *Finite Basis Set Corrections to Total Energy*
Pseudopotential Calculations, J. Phys. Condens. Matter 2, 4395 (1990).

[16] Y. He, P. Yi, and M. L. Falk, Critical Analysis of an FeP Empirical Potential Employed to Study the Fracture of Metallic Glasses, Phys. Rev. Lett. 122, 035501 (2019).

[17] M. I. Mendelev, Y. Sun, F. Zhang, C. Z. Wang, and K. M. Ho, Development of a Semi-Empirical Potential Suitable for Molecular Dynamics Simulation of Vitrification in Cu-Zr Alloys, J. Chem. Phys. 151, 214502 (2019).

[18] N. Artrith and A. Urban, An Implementation of Artificial Neural-Network Potentials for Atomic Materials Simulations: Performance for TiO2, Comput. Mater. Sci. 114, 135 (2016).

[19] L. Zhang, J. Han, H. Wang, R. Car, and W. E, Deep Potential Molecular Dynamics: A Scalable Model with the Accuracy of Quantum Mechanics, Phys. Rev. Lett. 120, 143001 (2018).

[20] T. Wen et al., Development of a Deep Machine Learning Interatomic Potential for Metalloid-Containing Pd-Si Compounds, Phys. Rev. B 100, 174101 (2019).

[21] J. Behler and M. Parrinello, Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces, Phys. Rev. Lett. 98, 146401 (2007).

[22] H. Yanxon, D. Zagaceta, B. C. Wood, and Q. Zhu, Neural Networks Potential from the Bispectrum Component: A Case Study on Crystalline Silicon, J. Chem. Phys. 153, 054118 (2020).

[23] J. Behler, Atom-Centered Symmetry Functions for Constructing High-Dimensional Neural Network Potentials, J. Chem. Phys. 134, 074106 (2011).

[24] N. Artrith, A. Urban, and G. Ceder, Efficient and Accurate Machine-Learning Interpolation of Atomic Energies in Compositions with Many Species, Phys. Rev. B 96, 014112 (2017).

[25] M. Gastegger, L. Schwiedrzik, M. Bittermann, F. Berzsenyi, and P. Marquetand, WACSF—Weighted Atom-Centered Symmetry Functions as Descriptors in Machine Learning Potentials, J. Chem. Phys. 148, 241709 (2018).

[26] A. P. Bartók, R. Kondor, and G. Csányi, On Representing Chemical Environments, Phys. Rev. B 87, 184115 (2013).

[27] A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi, Gaussian Approximation Potentials: The Accuracy of Quantum Mechanics, without the Electrons, Phys. Rev. Lett. 104, 136403 (2010).

[28] G. Imbalzano, A. Anelli, D. Giofré, S. Klees, J. Behler, and M. Ceriotti, Automatic Selection of Atomic Fingerprints and Reference Configurations for Machine-Learning Potentials, J. Chem. Phys. 148, 241730 (2018).
[29] M. J. Willatt, F. Musil, and M. Ceriotti, Feature Optimization for Atomistic Machine Learning Yields a Data-Driven Construction of the Periodic Table of the Elements, Phys. Chem. Chem. Phys. 20, 29661 (2018).

[30] V. L. Deringer, C. J. Pickard, and G. Csányi, Data-Driven Learning of Total and Local Energies in Elemental Boron, Phys. Rev. Lett. 120, 156001 (2018).

[31] D. J. Wales and J. P. K. Doye, Global Optimization by Basin-Hopping and the Lowest Energy Structures of Lennard-Jones Clusters Containing up to 110 Atoms, J. Phys. Chem. A 101, 5111 (1997).

[32] Y. Fan, T. Iwashita, and T. Egami, How Thermally Activated Deformation Starts in Metallic Glass, Nat. Commun. 5, 5083 (2014).

[33] A. Ninarello, L. Berthier, and D. Coslovich, Models and Algorithms for the next Generation of Glass Transition Studies, Phys. Rev. X 7, 021039 (2017).

[34] L. Berthier, E. Flenner, C. J. Fullerton, C. Scalliet, and M. Singh, Efficient Swap Algorithms for Molecular Dynamics Simulations of Equilibrium Supercooled Liquids, J. Stat. Mech. Theory Exp. 2019, 064004 (2019).

[35] M. E. Clamp, P. G. Baker, C. J. Stirling, and A. Brass, Hybrid Monte Carlo: An Efficient Algorithm for Condensed Matter Simulation, J. Comput. Chem. 15, 838 (1994).

[36] B. Mehlig, D. W. Heermann, and B. M. Forrest, Hybrid Monte Carlo Method for Condensed-Matter Systems, Phys. Rev. B 45, 679 (1992).

[37] M. P. Allen and D. J. Tildesley, Computer Simulation of Liquids, Second edition (Oxford University Press, Oxford, United Kingdom, 2017).

[38] J. Ding, Y.-Q. Cheng, and E. Ma, Full Icosahedra Dominate Local Order in Cu64Zr34 Metallic Glass and Supercooled Liquid, Acta Mater. 69, 343 (2014).

[39] S. H. Zhou and R. E. Napolitano, Phase Stability for the Cu–Zr System: First-Principles, Experiments and Solution-Based Modeling, Acta Mater. 58, 2186 (2010).

[40] Y. Q. Cheng, E. Ma, and H. W. Sheng, Atomic Level Structure in Multicomponent Bulk Metallic Glass, Phys. Rev. Lett. 102, 245501 (2009).

[41] Y. Q. Cheng and E. Ma, Atomic-Level Structure and Structure–Property Relationship in Metallic Glasses, Prog. Mater. Sci. 56, 379 (2011).

[42] G. Kresse and D. Joubert, From Ultrasoft Pseudopotentials to the Projector Augmented-Wave Method, Phys. Rev. B 59, 1758 (1999).

[43] P. E. Blöchl, Projector Augmented-Wave Method, Phys. Rev. B 50, 17953 (1994).

[44] J. P. Perdew, K. Burke, and M. Ernzerhof, Generalized Gradient Approximation Made Simple, Phys. Rev. Lett. 77, 3865 (1996).
[45] C. Zhu, R. H. Byrd, P. Lu, and J. Nocedal, Algorithm 778: L-BFGS-B: Fortran Subroutines for Large-Scale Bound-Constrained Optimization, ACM Trans. Math. Softw. 23, 550 (1997).

[46] A. M. Cooper, J. Kästner, A. Urban, and N. Artrith, Efficient Training of ANN Potentials by Including Atomic Forces via Taylor Expansion and Application to Water and a Transition-Metal Oxide, Npj Comput. Mater. 6, 54 (2020).

[47] M. I. Mendelev, D. J. Sordelet, and M. J. Kramer, Using Atomistic Computer Simulations to Analyze X-Ray Diffraction Data from Metallic Glasses, J. Appl. Phys. 102, 043501 (2007).

[48] M. I. Mendelev, M. J. Kramer, R. T. Ott, D. J. Sordelet, D. Yagodin, and P. Popel, Development of Suitable Interatomic Potentials for Simulation of Liquid and Amorphous Cu–Zr Alloys, Philos. Mag. 89, 967 (2009).

[49] C. Tang and P. Harrowell, Predicting the Solid State Phase Diagram for Glass-Forming Alloys of Copper and Zirconium, J. Phys. Condens. Matter 24, 245102 (2012).

[50] Y. Li, Q. Guo, J. A. Kalb, and C. V. Thompson, Matching Glass-Forming Ability with the Density of the Amorphous Phase, Science 322, 1816 (2008).

[51] W. H. Wang, J. J. Lewandowski, and A. L. Greer, Understanding the Glass-Forming Ability of Cu<sub>50</sub>Zr<sub>50</sub> Alloys in Terms of a Metastable Eutectic, J. Mater. Res. 20, 2307 (2005).

[52] K. G. S. H. Gunawardana, S. R. Wilson, M. I. Mendelev, and X. Song, Theoretical Calculation of the Melting Curve of Cu-Zr Binary Alloys, Phys. Rev. E 90, 052403 (2014).

[53] R. Freitas, M. Asta, and M. de Koning, Nonequilibrium Free-Energy Calculation of Solids Using LAMMPS, Comput. Mater. Sci. 112, 333 (2016).

[54] U. R. Pedersen, Direct Calculation of the Solid-Liquid Gibbs Free Energy Difference in a Single Equilibrium Simulation, J. Chem. Phys. 139, 104102 (2013).

[55] U. R. Pedersen, F. Hummel, G. Kresse, G. Kahl, and C. Dellago, Computing Gibbs Free Energy Differences by Interface Pinning, Phys. Rev. B 88, 094101 (2013).

[56] V. Wessels et al., Rapid Chemical and Topological Ordering in Supercooled Liquid Cu<sub>46</sub>Zr<sub>54</sub>, Phys. Rev. B 83, 094116 (2011).

[57] X. Wu, S. Lan, X. Wei, J. Zhou, Z. Lu, J. D. Almer, and X.-L. Wang, Elucidating the Nature of Crystallization Kinetics in Zr46Cu46Al8 Metallic Glass through Simultaneous WAXS/SAXS Measurements, Appl. Phys. Lett. 114, 211903 (2019).

[58] G. P. P. Pun, R. Batra, R. Ramprasad, and Y. Mishin, Physically Informed Artificial Neural Networks for Atomistic Modeling of Materials, Nat. Commun. 10, 1 (2019).

[59] A. P. Thompson et al., LAMMPS - a Flexible Simulation Tool for Particle-Based Materials
Modeling at the Atomic, Meso, and Continuum Scales, Comput. Phys. Commun. 271, 108171 (2022).

[60] L. Wang, A. Ninarello, P. Guan, L. Berthier, G. Szamel, and E. Flenner, Low-Frequency Vibrational Modes of Stable Glasses, Nat. Commun. 10, 26 (2019).

[61] A. Shakerpoor, E. Flenner, and G. Szamel, Stability Dependence of Local Structural Heterogeneities of Stable Amorphous Solids, Soft Matter 16, 914 (2020).

[62] T. S. Grigera, V. Martin-Mayor, G. Parisi, and P. Verrocchio, Phonon Interpretation of the ‘Boson Peak’ in Supercooled Liquids, Nature 422, 289 (2003).

[63] K. Vollmayr, W. Kob, and K. Binder, How Do the Properties of a Glass Depend on the Cooling Rate? A Computer Simulation Study of a Lennard-Jones System, J. Chem. Phys. 105, 4714 (1996).

[64] V. I. Tkatch, A. I. Limanovskii, S. N. Denisenko, and S. G. Rassolov, The Effect of the Melt-Spinning Processing Parameters on the Rate of Cooling, Mater. Sci. Eng. A 323, 91 (2002).
Supplementary materials of

Models and Algorithms Towards the Metallic Glasses Properties at Deep Supercooled Region

Rui Su\textsuperscript{1}, Xuefeng Zhang\textsuperscript{1}, Pengfei Guan\textsuperscript{2,1}

\textsuperscript{1}Institute of Advanced Magnetic Materials, College of Materials & Environmental Engineering, Hangzhou Dianzi University, Hangzhou 310018, P. R. China

\textsuperscript{2}Beijing Computational Science Research Center, Beijing 100193, P. R. China

*Corresponding authors:

X. F. Zhang (zhang@hdu.edu.cn)

P. F. Guan (pguan@csrc.ac.cn)
1. Calculation of crystalline free energies

We calculate the Gibbs free energies at p = 0 GPa at different temperatures by the Frenkel-Ladd method in the non-equilibrium thermodynamical integration (TI) approach. In this approach, the system Helmholtz free energy is calculated as:

\[ F_0(N, V, T) = F_E(N, V, T) + \Delta F(N, V, T) \]  

(S1)

where \( F_E(N, V, T) \) is the free energy of the referencing Einstein solid and the free energy difference \( \Delta F(N, V, T) \) between current and reference Hamiltonian is calculated as:

\[ \Delta F = \int_0^1 d\lambda \left( \frac{\partial H}{\partial \lambda} \right)_\lambda = \frac{1}{2} (W_{0\rightarrow E}^{\text{irr}} - W_{E\rightarrow 0}^{\text{irr}}) \]  

(S2)

where \( \langle \cdots \rangle \) denotes for ensemble averaging, 0 and E stand for initial NNP Hamiltonian and final Einstein solid one, in respective.

In the TI calculation, \( \langle \partial H / \partial \lambda \rangle_\lambda \) is calculated as the system Hamiltonian changes from 0 to E by varying the time-dependent parameter \( \lambda \):

\[ H(\lambda) = \lambda H(0) + (1 - \lambda)H(E) \]  

(S3)

where \( \lambda \) changes from 0 to 1 in the switching time \( t_s \). In this work, the switching function: \( \lambda(t) = t^5(70t^4 - 315t^3 + 540t^2 - 420t + 126) \) as suggests by Freitas et. al.[S1] is used.

The irreversible work along the TI path is calculated as:

\[ W_{0\rightarrow E}^{\text{irr}} = \int_0^{t_s} dt \frac{d\lambda}{dt} \left( \frac{\partial H}{\partial \lambda} \right)_{\Gamma(t)} \]  

(S4)

where \( \Gamma(t) \) denotes for the phase-space trajectory produced by the simulation. In the non-equilibrium TI, the reversible part of work (free energy difference) is estimated by canceling the irreversible part as:

\[ \Delta F = \frac{1}{2} (W_{0\rightarrow E}^{\text{irr}} - W_{E\rightarrow 0}^{\text{irr}}) \]  

(S5)

The absolute free energy of the Einstein solid is:
\[ F_E(N, V, T) = 3Nk_B T \ln \left( \frac{\hbar \omega}{k_B T} \right) \]  

(S6)

where the oscillator frequency \( \omega \) is obtained from spring constant: \( k = m \omega^2 \). The mean squared displacement \( \langle \Delta r^2 \rangle \) is used to calculate the spring constant \( k \):

\[ k = \frac{3k_B T}{\langle \Delta r^2 \rangle} \]  

(S7)

To obtain Gibbs free energy, the system is first equilibrated at external pressure \( P \). Then the Helmholtz free energy \( F \) is evaluated at fixed volume. The Gibbs free energy is obtained as \( G = F + PV \).

In this work, we calculate free energies of crystalline phases as following: Firstly, the system is equilibrated for 50 ps at given temperature and pressure in a \( NpT \) ensemble and the spring constant \( k \) is estimated. Secondly, we fix the box and vary \( \lambda \) to perform forward and backward TI runs to calculate the free energy difference \( \Delta F \). Finally, the Gibbs free energy is obtained by adding the contribution from Einstein solid. For both forward and backward TI runs, we first equilibrate system for \( t_{eq} = 10 \text{ps} \) and switch \( \lambda \) with a switching time \( t_s = 50 \text{ps} \). Supercells containing 2,000, 1,800 and 3264 atoms are used for ZrCu, Zr\(_2\)Cu and Zr\(_7\)Cu\(_{10}\), respectively.
2. Calculation of the melting temperatures of B2-ZrCu

The melting temperature of B2-ZrCu is calculated by the “interface pinning” (IP) method\[\text{S2}\] at $P = 0 \text{ GPa}$. IP method hold the solid-liquid interface at $T$ by adding a harmonic bias term to the potential energy:

$$U'(R) = U(R) + \frac{\kappa}{2} (Q(R) - a)^2$$  \hspace{1cm} (S8)

where $Q(R)$ is a global order parameter that depends on the number of crystalline atoms. $\kappa$ is the harmonic sprint constant. $a$ corresponds to the “interface position” in the 1D space of order parameter $Q(R)$.

The chemical potential difference between solid and liquid phases is calculated as:

$$\Delta \mu = -\kappa \langle (Q)' - a \rangle \Delta Q / N$$  \hspace{1cm} (S9)

The entropy difference $\Delta s$ can be calculated from basic thermodynamic relation:

$$\Delta s = \frac{\Delta u + p \Delta v - \Delta \mu}{T}$$  \hspace{1cm} (S10)

where $\Delta \mu$ and $\Delta v$ are differences of internal energy and specific volume between crystalline and liquid phases, respectively. From a starting guess $T_0$, the melting temperature $T$ estimation is made through Newton’s iteration:

$$T^{i+1} = T^i + \frac{\Delta \mu}{\Delta s}$$  \hspace{1cm} (S11)

The iteration terminated when $T$ variance is less than statistical error.

In this work, we use the translation order parameter\[\text{S2}\] $\rho_k$ as $Q(R)$:

$$Q(R) \equiv |\rho_k| = \left| \sum_{j=1}^{N} \exp(-i\mathbf{k} \cdot \mathbf{r}_j) \right|$$ \hspace{1cm} (S12)

For the B2-CuZr, we construct a $8 \times 8 \times 20$ supercell elongating along the $z$ axis. $\rho_k$ is evaluated using the wave vector $\mathbf{k} = (2\pi n_x / X, 0, 0)$ where $n_x = 16$. The biased harmonic
potential is applied using $\kappa = 4.0$ and $a = 26$. The solid-liquid interface is equilibrated using a $Np_zT$ ensemble for 4 ns and the last 200 ps is used to calculate the ensemble averaged $\langle Q \rangle'$. Fig. S1 shows a snapshot of equilibrated solid-liquid interface at $T_m$ using the NNP potential.

Fig. S1 MD snapshot of equilibrated solid-liquid interface of B2-ZrCu. The Cu and Zr atoms are shown as red and blue atoms, respectively.
Table S1 Formation enthalpies ($\Delta H_f$) of different crystalline phases.

| Formula | Phase | $\Delta H_{f,0K}$ (eV/atom) | DFT  | NNP  | CMS  | MSZWH | MKOSYP | MSK  |
|---------|-------|-----------------------------|------|------|------|--------|--------|------|
| Zr      | hcp   | 0                           | 0    | 0    | 0    | 0      | 0      | 0    |
| Zr      | bcc   | 0.084                       | 0.0855| 0.0237| 0.0517| 0.0517| 0.0517| 0    |
| Zr$_2$Cu| C11$_b$| -0.1428                     | -0.1453| -0.1146| -0.1465| -0.0935| -0.0237| 0    |
| ZrCu    | B2    | -0.1052                     | -0.1034| -0.1552| -0.134 | -0.15  | -0.1391| 0    |
| Zr$_7$Cu$_{10}$ | $\phi$ | -0.175                     | -0.1735| -0.1074| -0.1651| -0.1144| -0.0949| 0    |
| ZrCu$_2$| $\sigma$ | -0.129                  | -0.1304| -0.0547| -0.0857| -0.0683| 0.0223| 0    |
| ZrCu$_2$| C14   | -0.1172                     | -0.1154| -0.1522| -0.1664| -0.1504| -0.1849| 0    |
| ZrCu$_2$| C36   | -0.0903                     | -0.0868| -0.1467| -0.1611| -0.1511| -0.1958| 0    |
| ZrCu$_2$| C15   | -0.0662                     | -0.0625| -0.1411| -0.1345| -0.1517| -0.2065| 0    |
| Zr$_3$Cu$_8$ | $\delta$ | -0.1754                | -0.1769| -0.1177| -0.1726| -0.0893| -0.0945| 0    |
| Zr$_{14}$Cu$_{51}$ | $\beta$ | -0.1713               | -0.1778| -0.0788| -0.1393| -0.0868| -0.0689| 0    |
| ZrCu$_5$| C15$_b$ | -0.1266                   | -0.1266| -0.0289| -0.0989| -0.034 | -0.0522| 0    |
| Cu      | fcc   | 0                           | 0    | 0    | 0    | 0      | 0      | 0    |

MAE       | -    | -     | 0.0023 | 0.0631 | 0.0335 | 0.0628 | 0.0898 | 0  |

AlCu      | mp-2500$^b$ | -0.215 | -0.2122| -0.2458| -    | -    | -    | 0  |
Al$_2$Cu  | mp-998$^b$  | -0.1559| -0.1523| -0.2467| -    | -    | -    | 0  |
Al        | fcc          | 0      | 0      | 0      | -    | -    | -    | 0  |
ZrAl      | mp-11233$^b$| -0.4717| -0.4783| -0.4717| -    | -    | -    | 0  |
Zr$_2$Al  | mp-2557$^b$ | -0.3607| -0.3645| -0.4043| -    | -    | -    | 0  |
Zr$_3$Al  | mp-1471$^b$ | -0.3026| -0.2938| -0.3118| -    | -    | -    | 0  |
ZrAlCu$_2$| mp-3736$^b$ | -0.3619| -0.3687| -0.3211| -    | -    | -    | 0  |

MAE       | -    | -     | 0.0034 | 0.0535 | -    | -    | -    | 0  |

(a): $\Delta H_f(Zr_nCu_mAl_k) = (E - mE_{Zr} - nE_{Cu} - kE_{Al}) / (n + m + k)$
(b): Materials project IDs.
References

[S1] R. Freitas, M. Asta, and M. de Koning, *Nonequilibrium Free-Energy Calculation of Solids Using LAMMPS*, Comput. Mater. Sci. **112**, 333 (2016).

[S2] U. R. Pedersen, *Direct Calculation of the Solid-Liquid Gibbs Free Energy Difference in a Single Equilibrium Simulation*, J. Chem. Phys. **139**, 104102 (2013).