



CHORUS

This is the accepted manuscript made available via CHORUS. The article has been published as:

Efficient math

d -dimensional molecular dynamics simulations for studies of the glass-jamming transition

Robert S. Hoy and Kevin A. Interiano-Alberto

Phys. Rev. E **105**, 055305 — Published 16 May 2022

DOI: [10.1103/PhysRevE.105.055305](https://doi.org/10.1103/PhysRevE.105.055305)

Efficient d -dimensional molecular dynamics **simulations** for studies of the glass-jamming transition

Robert S. Hoy* and Kevin A. Interiano-Alberto

Department of Physics, University of South Florida, Tampa, FL 33620 USA

(Dated: April 28, 2022)

We develop an algorithm suitable for parallel molecular dynamics simulations in d spatial dimensions and describe its implementation in C++. All routines work in arbitrary d ; the maximum simulated d is limited only by available computing resources. These routines include several that are particularly useful for studies of the glass/jamming transition, such as SWAP Monte Carlo and FIRE energy minimization. Scaling of simulation runtimes with the number of particles N and number of simulation threads n_{threads} is comparable to popular MD codes such as LAMMPS. The efficient parallel implementation allows simulation of systems that are much larger than those employed in previous high-dimensional glass-transition studies. As a demonstration of the code's capabilities, we show that supercooled $d = 6$ liquids can possess dynamics that are substantially more heterogeneous and experience a breakdown of the Stokes-Einstein relation that is substantially stronger than previously reported, owing at least in part to the much smaller system sizes employed in earlier simulations.

I. INTRODUCTION

Molecular dynamics (MD) simulations have been an essential part of statistical physicists' toolbox for over 50 years [1–4]. Numerous open-source, highly-optimized multipurpose parallel MD simulation packages are now available [5–10]. These packages are designed to simulate systems embedded in the physically relevant spatial dimensions $2 \leq d \leq 3$; they cannot be used for $d > 3$ simulations without extensive modifications. Accordingly, they employ parallelization algorithms that are efficient in low d , *e.g.* spatial domain decomposition allowing simulation of multibillion-atom systems on distributed-memory supercomputers [11, 12]. These algorithms, however, rapidly become less efficient as d increases. Developing publicly available codes that allow efficient parallel MD simulations of higher-dimensional systems to be performed may be a key step towards answering several open questions in physics, particularly questions related to supercooled liquids and the glass/jamming transition. In this paper, we describe the first such code (**hdMD**) [13]. Then we use it to show that the dynamics of supercooled $d = 6$ liquids at densities above the RFOT dynamical glass transition density ϕ_d [14] can be substantially more heterogeneous than previously reported.

Early simulations of high-dimensional supercooled liquids employing systems of $N = 10^3 - 10^4$ particles indicated that crystallization is strongly suppressed as d increases beyond 3 [15–18]. Recent work, however, has shown that (i) crystallization instabilities in glassforming mixtures often appear only when larger N [19] or SWAP Monte Carlo equilibration [20] are employed, and (ii) hard-sphere crystals are thermodynamically stable for a wide range of packing fractions $\phi < \phi_d$ for all $d \leq 10$ [21]. Both results suggest that the conclusions of many of the

early $d > 3$ studies need to be reexamined using larger systems. Moreover, the N typically employed in recent simulations of higher- d systems have remained small (*e.g.* $N < 10^4$ in Refs. [22–24]), precluding robust investigation of how any static or dynamic length scales that grow substantially as ϕ_d is approached [25–29] depend on d .

For example, simulation studies suggesting that dynamical heterogeneity weakens as d increases [30–32] can be challenged on the grounds that heterogeneous dynamics in these studies are suppressed not (or not primarily) by the increase in spatial dimension, but instead by their use of periodic boundary conditions with simulation cell side lengths L that decrease rapidly with increasing d . This is a reasonable concern given the tendency of the least mobile particles in glassforming liquids to remain near their initial positions after time intervals over which the most mobile particles have hopped by several times their diameter [33, 34] and the fact that L drops as low as ~ 2 particle diameters for the highest studied d [30–32]. Resolving this issue requires performing analogous simulations with larger L , and completing such simulations in a reasonable amount of time requires an efficient parallel implementation of high-dimensional MD.

Simulating supercooled liquids in high d presents several challenges. One of the most difficult involves neighbor-list construction. If one wishes to simulate systems with a given number of particles N , the cell side length $L \sim N^{1/d}$, and thus the number of linked (fixed-size) subcells along each axis of the simulation cell is $n_{\text{cell}} \sim N^{1/d}$. The total number of subcells is $\mathcal{N}_{\text{sc}} = n_{\text{cell}}^d$, and the number of these which must be searched over for each atom during a typical Verlet list (VL) build is 3^d . The total number of atoms in these neighboring cells is $N_{\text{nc}} = (3^d/\mathcal{N}_{\text{sc}})N \equiv (3/n_{\text{cell}})^d N \sim 3^d$, so the total number of pair distances which must be evaluated to rebuild all atoms' Verlet lists is $\mathcal{N}_{\text{VL}} \sim 3^d N$. In other words, the effort required to maintain the neighbor lists increases exponentially with increasing d because building each atom's VL requires searching over a exponen-

* rshoy@usf.edu

tially increasing fraction of the system’s volume.

Avoiding strong finite-size effects requires simulating systems of a minimum spatial extent \mathcal{L} ; typically $\mathcal{L} \gtrsim 10$ particle diameters. If \mathcal{L} is taken to be L , the minimum number of particles simulations must include to avoid such affects is $N_{\min} \sim L^d$. While characteristic packing fractions ϕ scale across dimensions roughly as 2^{-d} or $d2^{-d}$ [35, 36], hypersphere volumes decrease faster than this, so the corresponding particle number densities ρ increase with increasing d . For example, the dynamical-glass-transition number densities for hard spheres, $\rho_d \simeq 1.78 - .5d + .095d^2$ [30], nearly triple as d increases from 3 to 7. Taking $N_{\min} = \rho_d L^d$, the computational effort to build the VL for a d -dimensional supercooled liquid scales roughly as $\rho_d (3L)^d$.

This rapidly becomes prohibitive as d increases, but it can be overcome for intermediate d using efficient parallel algorithms. In this paper, we describe our newly developed `hdMD` code, which employs several such algorithms and is capable of simulating much larger high- d systems than have been studied previously. `hdMD` includes several routines commonly employed in studies of the glass/jamming transition [37–39], such as SWAP Monte Carlo [20, 40], FIRE energy minimization [41, 42], and calculation of the overlap parameter $f_{\text{ov}}(t)$ and van Hove correlation function $G_s(r, t)$. Here we use it to simulate $3 \leq d \leq 6$ supercooled liquids of up to 10^7 particles over short timescales, and 10^5 particles over the much longer timescales typically employed in modern glass-transition studies [20]. The latter simulations show that strongly heterogeneous dynamics [as indicated by large non-Gaussian parameters] can persist to very long times in deeply supercooled liquids for d up to at least 6. By contrasting results for $N = 10^5$ to those for $N = 5000$ liquids at the same ϕ and T , we provide initial evidence that some of the conclusions of Refs. [30–32] – at least in terms of their quantitative details – would have been different had these studies considered larger systems.

The outline of the rest of this paper is as follows. Section II describes `hdMD`’s algorithmic implementation. Section III presents performance benchmarks for simulations at packing fractions $\phi \simeq \phi_d$ for $3 \leq d \leq 6$, focusing on the code’s parallel efficiency and the scaling of simulation runtimes with N and d . Section IV includes an original analysis of the dynamics of supercooled $d = 6$ liquids and serves as a demonstration of `hdMD`’s suitability for state-of-the-art glass-transition-related studies. Finally, in Section V, we discuss our results and conclude. Readers more interested in the code’s performance than its algorithmic implementation are urged to skip to Sec. III, while readers primarily interested in our supercooled-liquid physics results are urged to skip to Sec. IV.

II. ALGORITHMIC IMPLEMENTATION

Most previous simulations of liquids in $d > 3$ have employed hard spheres. This is sensible given that hard-core

interactions dominate liquids’ structure [2] and polydisperse hard-sphere liquids are excellent glassformers [20]. However, hard-particle techniques are inherently limited in the range of physical phenomena they can capture. The lack of finite interparticle forces makes it difficult for them to accurately model the collective rearrangements which increasingly dominate liquids’ relaxation mechanisms as the glass transition is approached [26, 27]. They cannot, for example, capture the long-range elastically-mediated dynamical facilitation that has recently been shown to arise below the mode-coupling temperature T_{MCT} [43]. While attractive forces are well known to exert a strong and nonperturbative influence on 3D glass-forming liquids’ dynamics [44] and to increase their tendency to crystallize [45], the variation of these effects with increasing d remains unexplored and cannot be explored using hard-particle models. Therefore we employ a soft-particle MD approach that can treat both repulsive and attractive forces.

Parallelizing a MD simulation requires devising a method of distributing the computation across its n_{threads} concurrent threads. Modern MD packages, most of which are optimized for large- n_{threads} simulations on distributed-memory machines [5–10], typically parallelize via spatial domain decomposition of simulation cells [11]. In this scheme, different CPU cores “own” geometrically distinct regions, and information that needs to be passed across the boundaries of these regions is typically passed via intercore communication, usually implemented using MPICH. For example, a typical $n_{\text{threads}} = 8$ simulation employing a cubic simulation cell divides it into $2 \times 2 \times 2$ domains, each of which is a cube with $n_{\text{cell}}/2$ subcells along each edge. The intercore communication that is necessary for force evaluation and Verlet list (VL)-building is only required between subcells on the surfaces of these domains [11].

Codes that operate this way work very well for $d \leq 3$. For example, simulations of million-atom liquids have been shown to exhibit nearly-optimal scaling (runtimes $\sim n_{\text{threads}}^{-1}$) for n_{threads} up to $\sim 10^2$ [46]. As d increases, however, the spatial-domain-decomposition method loses its effectiveness for reasons comparable to those outlined above for VL-building. Therefore we respectively employ per-atom, per-atom, and per-subcell parallelization for force evaluation, time integration of equations of motion, and VL-building. In this scheme, each thread is responsible for N/n_{threads} atoms and $\mathcal{N}_{\text{sc}}/n_{\text{threads}}$ subcells rather than a spatial domain, and intercore communication is avoided entirely by using OpenMP rather than MPICH to parallelize the code. While the latter choice limits our code to shared-memory (as opposed to distributed-memory) machines, it makes it far more efficient.

Table I lists `hdMD`’s principal routines. All of these routines work in arbitrary d ; the maximum simulated d is limited only by available computing resources. Our implementation of these routines will be described in detail in the following sections.

TABLE I. Principal routines (C++ functions) included in `hMDM`. Additional standard functions involving I/O and memory allocation are present in the code but are not listed here.

Name	Purpose
<code>main</code>	Control program operation
<code>leapfrog</code>	Integrate Newton's EOM
<code>getforce</code>	Calculate force on atom i
<code>Berendsen</code>	Berendsen thermo/barostat
<code>getthermo</code>	Calculate thermodynamic quantities
<code>getvanHove</code>	Calculate van Hove corr. function
<code>initRMSE</code>	Initialize $\{\vec{r}_i\}$, $\{m_i\}$, $\{\epsilon_i\}$, $\{\sigma_i\}$
<code>setupcell</code>	Initialize simulation cell
<code>needsrebuild</code>	Check if VLs need to be rebuilt
<code>buildneighborlist</code>	Build/rebuild the VLs
<code>FIRE</code>	Control FIRE energy minimization
<code>FIREintegrate</code>	Integrate FIRE EOM
<code>steepest</code>	Steepest-descent energy minimization
<code>buildswapNLs</code>	Generate SWAP-attempt VLs
<code>swapmove</code>	Attempt SWAP moves
<code>writerestart</code>	Write restart files to disk
<code>writecoords</code>	Write $\{\vec{r}_i\}$ and image flags to disk

A. Particle model and initial state generation

Here, for simplicity, we present results for a single hard-sphere-like pair potential, the truncated and shifted Morse potential given by

$$U_a(\epsilon_{ij}, \sigma_{ij}, r_{ij}) = \epsilon_{ij} \left[\exp[-2a(r_{ij} - \sigma_{ij})] - 2 \exp[-a(r_{ij} - \sigma_{ij})] + 1 \right] \quad (1)$$

for $r_{ij} \leq \sigma_{ij}$ and zero for $r_{ij} \geq \sigma_{ij}$, where r_{ij} is the center-to-center distance between particles i and j . We use the standard Lorentz-Berthelot MD mixing rules for particle diameters and force prefactors, i.e. $\sigma_{ij} = (\sigma_i + \sigma_j)/2$ and $\epsilon_{ij} = \sqrt{\epsilon_i \epsilon_j}$. Longer-range interactions including attractive terms can be implemented by adjusting `main()`'s "cut" parameter. The Morse parameter a can be adjusted by editing a single-line in `getforce()`; all results presented below are for $a = 30\tilde{\sigma}^{-1}$. More generally, `hMDM` can be easily generalized to arbitrary radial force laws and mixing rules (incorporating, e.g., nonadditivity [20]) by editing a few lines of code in the `getforce()`, `thermo()`, and `swapmove()` routines. In particular, switching the stock code from the Morse to the Mie potential $U_n(\epsilon_{ij}, \sigma_{ij}, r_{ij}) = \epsilon_{ij} [(\frac{\sigma_{ij}}{r})^{2n} - 2(\frac{\sigma_{ij}}{r})^n]$, which reduces to the widely employed Lennard-Jones/WCA potential for $n = 6$, requires only commenting/uncommenting-out a few lines of code in these routines.

The volume of a d -dimensional spherical particle with diameter σ is

$$v(d, \sigma) = \frac{\pi^{d/2} \sigma^d}{2^d \Gamma(1 + d/2)}, \quad (2)$$

where Γ is the gamma function. For all simulations discussed in Section III, we set particles' mass density $\rho_m = [v(d, \sigma)\sigma^{-d}]^{-1} = \pi^{-d/2} 2^d \Gamma(1 + d/2)$ so particles have mass $m_i = \sigma_i^d$ and typical particles with $\sigma = 1$ have unit mass. Correspondingly, we set $\epsilon_i = \sigma_i^d$ so repulsive interactions also scale with particle volumes [47]. During MD simulation, Newton's equation of motion are integrated using the standard leapfrog algorithm with a timestep $dt = 6\tau/(125a) = .0016\tau$, where $\tau = \sqrt{\tilde{m}\tilde{\sigma}^2/\tilde{\epsilon}}$ is the unit of time. Here \tilde{m} , $\tilde{\sigma}$, and $\tilde{\epsilon}$ are the mass, diameter and force prefactor for typical particles: $\tilde{m} = \tilde{\sigma} = \tilde{\epsilon} = 1$ in dimensionless units.

Polydispersity is one of the most important factors controlling glassforming ability in molecular simulations [48]. We use the particle-diameter distribution

$$P(d, \mathcal{R}, \sigma) = \begin{cases} \frac{(d-1)\sigma^{-d}}{\mathcal{R}^{\frac{d-1}{2d}} - \mathcal{R}^{-\frac{d-1}{2d}}}, & \mathcal{R}^{-\frac{1}{2d}} \leq \sigma \leq \mathcal{R}^{\frac{1}{2d}} \\ 0, & \sigma < \mathcal{R}^{-\frac{1}{2d}} \text{ or } \sigma > \mathcal{R}^{\frac{1}{2d}} \end{cases}, \quad (3)$$

where $\mathcal{R} = (\sigma_{\max}/\sigma_{\min})^d$ is the ratio of maximum to minimum particle volumes. The total volume $\mathcal{V}(\sigma) = NP(d, \mathcal{R}, \sigma)v(d, \sigma)$ occupied by particles of diameter σ is σ -independent over the entire range $\mathcal{R}^{-\frac{1}{2d}} \leq \sigma \leq \mathcal{R}^{\frac{1}{2d}}$; this choice apparently optimizes glass-formability for a wide range of force laws [20]. Preventing crystallization also requires a sufficiently large polydispersity index $\Delta = [\langle \sigma^2 \rangle - \langle \sigma \rangle^2]^{1/2} / \langle \sigma \rangle$, with the minimum Δ decreasing with increasing d [20]. Here we choose $\mathcal{R} = d$, which gives $\Delta = 0.107, 0.099, 0.091, 0.083$ in $d = 3 - 6$. We find that these parameter choices are sufficient to prevent both crystallization and phase separation in all systems discussed below.

To generate N -particle initial states with packing fraction ϕ , we first define the particle diameters $\{\sigma_i\}$ by randomly sampling the distribution $P(d, \mathcal{R}, \sigma)$, using the function `initRMSE()`. This function also sets the particle masses $\{m_i\}$ and force prefactors $\{\epsilon_i\}$. The total volume occupied by these particles is

$$V_{\text{part}} = \sum_{i=1}^N v(d, \sigma_i); \quad (4)$$

thus the total simulation cell volume must be $V = V_{\text{part}}/\phi$. We employ hypercubic simulation cells of this volume (and side length $L = V^{1/d}$), centered at the origin; `initRMSE()` assigns all particles random initial positions $\{\vec{r}_i\}$ within these cells. After initializing the linked-subcell data structures using `setupsupercell()`, we populate the subcells and initialize all particles' Verlet lists using `buildneighborlist()`. These two routines will be described in detail in Section II C.

The randomly generated positions lead to severe particle overlap that must be reduced before the simulation can begin. We accomplish this "pushoff" using a partial FIRE minimization (Section II B). During this minimization and throughout the rest of the simulation, periodic

boundary conditions are applied along all d directions. After the pushoff is complete, we give particles random initial velocities corresponding to the desired target temperature T_{targ} . Alternatively, the abovementioned initial-state-generation-and-pushoff procedure may be skipped by reading initial states (e.g. restart files written during a previous simulation) in from an ASCII file. This is also handled using `initRMSE()`, with the relevant restart file names passed to `hdMD` as command-line arguments as discussed below. Then the linked-subcell data structures are initialized, the subcells are populated and all particles' VLs are initialized as outlined above. Once either of these options is completed, the MD simulation begins.

B. Program control, force evaluation and integration of equations of motion

In this section, we outline `hdMD`'s usage and large-scale structure. Entering the command `./hdMD d N ϕ T n_steps n_threads p resfilename rinitfilename n_msdstart` starts a d -dimensional MD simulation with N particles at packing fraction ϕ and target temperature T . The fifth and sixth command-line parameters specify the number of MD timesteps and the number of parallel OpenMP threads. p is the SWAP attempt fraction, i.e. the fraction of particles for which swaps are attempted every time `swapmove()` gets called. `resfilename` is the name of the file containing the initial particle $\{m_i\}$, $\{\sigma_i\}$, $\{\epsilon_i\}$, positions, velocities, and image flags. Such a file is unnecessary (and `resfilename` should be "null") if these are to be generated as outlined above. Finally, `rinitfilename` is the name of the file containing an *earlier* set of particle positions and image flags to be used as the reference configuration for calculations of particles' mean-squared displacement, etc. as outlined in Section II E. Such a file is unnecessary and `rinitfilename` should be "null" if the initial particle positions are to be used for this reference configuration. In addition to the abovementioned command-line parameters, a number of other runtime parameters are defined in `hdMD`'s `main.cpp` file and summarized in Table II.

Once a simulation is running, it dumps thermodynamic data every n_{thermo} MD timesteps to a file named `thermodata.dd.NN.phi ϕ .pp`, energy-minimization data every n_{min} timesteps to a file named `mindata.dd.NN.phi ϕ .pp` (if `minflag = true`), restart files including all particles $\{m_i\}$, $\{\sigma_i\}$, $\{\epsilon_i\}$, positions, velocities, and image flags every n_{restart} timesteps to files named `restart.dd.NN.phi ϕ .pp.stepstep`, van Hove correlation function data every n_{vhc} timesteps to a file named `vanHoveCorr.dd.NN.phi ϕ .pp` (if `vhcflag = true`), and the particle positions and image flags at step n_{msdstart} to a file named `Rinit.dd.NN.phi ϕ .pp.stepstep` (and also every n_{dump} timesteps to a file named `coords.dd.NN.phi ϕ .pp`). Italicized quantities indicate the numerical values of the various parameters.

TABLE II. Principal user-adjustable parameters defined in `hdMD`. Each of these may be adjusted by editing one line of code in `main()`. Alternatively, they may be straightforwardly converted to command-line parameters.

Name	Definition
n_{thermo}	Steps between <code>getthermo()</code> calls
n_{vhc}	Steps between <code>getvanHove()</code> calls
n_{baro}	Steps between <code>Berendsen()</code> calls
n_{restart}	Steps between <code>writerestart()</code> calls
n_{dump}	Steps between <code>writecoords()</code> calls
n_{bet}	Steps between <code>needsrebuild()</code> calls
n_{min}	Steps between <code>FIRE()</code> calls
n_{swap}	Steps between <code>swapmove()</code> calls
n_{msdstart}	Steps before MSD calculation begins
n_{chunk}	Chunk size for integration of EOM
$T_{\text{targ}}, P_{\text{targ}}$	Target temperature and pressure
$\tau_{\text{temp}}, \tau_{\text{press}}$	Temperature and pressure damping times
$E_{\text{thres}}, F_{\text{thres}}$	Convergence criteria for <code>FIRE()</code>
<code>maxiter</code>	Maximum number of <code>FIRE</code> iterations
Δ_{max}	Maximum particle displacement/timestep
<code>baroflag</code>	True (false) for NPT (NVT) simulations
<code>minflag</code>	True (false) if performing (not performing) periodic energy minimization
<code>vhcflag</code>	True (false) if performing (not performing) van Hove correlation function calculations

After initial states are prepared using either of the two methods outlined above, `hdMD`'s main loop (i.e. time integration for n_{steps} MD timesteps) begins. Before describing the structure of this main loop, however, we discuss some implementation details particular to `hdMD` that we found optimize its performance. First, rather than employing three 2D arrays of format `r[N][d]`, `v[N][d]`, and `f[N][d]` for the positions, velocities and forces as is done in (e.g.) LAMMPS [5], we store all three in a single 1D array (`rvf`) of length $3dN$. Second, for all parallel for loops of length $\sim dN$, we employ the OpenMP scheduling directive `schedule (static, n_chunk)` with $n_{\text{chunk}} = 100$. Both of these choices speed up the code by improving its cache locality; see Section III B.

`hdMD`'s main loop is structured as follows:

- `needsrebuild()` is called once every n_{bet} timesteps to determine whether particles' VLs need to be rebuilt (Section II C).
- On timestep n_{msdstart} , the positions `rinit` that will be used as the reference state for followup calculations of particles' mean-squared displacement, van Hove correlation function, etc. (Section II E) are stored in memory and written to `rinitfile`.
- `getthermo()` is called once every n_{thermo} timesteps.
- If `vhcflag` is set to `true` and `step` $\geq n_{\text{msdstart}}$, `getvanHove()` is called once every n_{vhc} timesteps.

- `writerestart()` is called once every n_{restart} timesteps.
- If `minflag` is set to `true` and `step` $\geq n_{\text{msdstart}}$, `FIRE()` is called once every n_{min} timesteps.
- `swapmove()` is called once every n_{swap} timesteps provided $p > 0$ (Section IID).
- Newton’s equations of motion are integrated forward in time by the increment $\Delta t = n_{\text{bet}} dt$ by calling `leapfrog()` n_{bet} times.

Here `step` indicates the timestep $\#$. n_{bet} is analogous to LAMMPS’ “delay” parameter [5]; large values reduce simulation runtimes slightly but run the risk of insufficiently-frequent VL builds that lead to lost-atom crashes. For the $T = 0.25\tilde{\epsilon}/k_B$ runs described below, we found that $n_{\text{bet}} = 5$ was sufficiently low to ensure stability of the runs. Lower T allow larger n_{bet} values.

Newton’s equation of motion are integrated via the “kick-drift-kick” variant of the leapfrog method [49]. For maximum efficiency, thermostating in NVT simulations (see Section IIE) is implemented as part of `leapfrog()`’s second kick (i.e. its second velocity update); this improves cache locality. Force evaluation within `getforce()` is handled in standard fashion, with one exception; for $d > 5$, the pair distance calculation

$$r_{ij} = \sum_{k=1}^d (\vec{r}_j - \vec{r}_i) \cdot \hat{x}_i, \quad (5)$$

where \hat{x}_i is the unit vector pointing along the i th spatial axis, is truncated before its completion if the partial sum exceeds the cutoff radius.

Efficient energy minimization is implemented using the semi-implicit Euler variant of the “FIRE 2.0” update [42] to the original FIRE algorithm [41]. `hdMD`’s `FIRE()` routine takes two arguments: E_{thres} and F_{thres} . Minimization stops when the average pair energy drops below E_{thres} , the fractional pair energy drop

$$F = \left| \frac{E_{i+1} - E_i}{E_i} \right| \quad (6)$$

is smaller than F_{thres} for ten consecutive iteration steps, or the iteration count reaches `maxiter`. During the abovementioned “pushoff”, complete energy minimization is unnecessary, so we set $E_{\text{thres}} = \epsilon$. `hdMD` also includes `steepest()`, a standard adaptive-timestep gradient-descent energy minimization routine which stops when the average pair energy drops below E_{thres} (its only argument).

C. Cell setup and Verlet-list building

As discussed above, `hdMD`’s simulation cells are hypercubes of side length L , divided into $\mathcal{N}_{\text{sc}} = n_{\text{cell}}^d$ cubic

subcells with side length $L_{\text{sc}} \equiv L/n_{\text{cell}}$. Here

$$n_{\text{cell}} = \text{floor} \left(\frac{L}{\mathcal{R}^{\frac{1}{2d}} \tilde{\sigma} + s} \right), \quad (7)$$

where $\mathcal{R}^{\frac{1}{2d}} \tilde{\sigma}$ is the diameter of the largest particles, s is the skin depth, and `floor[x]` rounds x downward to the nearest integer, e.g. `floor[12.3] = 12`. All particles’ VLs are rebuilt any time `needsrebuild()` finds that at least 0.1% of particles have moved by at least $s/2$ or any particles have moved by at least s since the previous build. Partial VL builds wherein only the VLs of particles that have moved by at least $s/2$ and those in neighboring subcells are rebuilt [50] increased runtimes for simulations like those reported below, so this capability was not included in the final code. We found that $s = .25\tilde{\sigma}$ works well for a wide range of conditions and use this value in all simulations reported below.

The maximum cutoff radius for inclusion in any particle’s VL, $r_c^{\text{max}} = \mathcal{R}^{\frac{1}{2d}} \tilde{\sigma} + s$, is the maximum interaction range, (i.e. the maximum interparticle distance for which nonzero forces can arise) plus s . Thus L_{sc} is equal to or slightly larger than r_c^{max} . Many modern MD codes employ smaller subcells of side length equal to or slightly larger than $r_c^{\text{max}}/2$ and link each subcell to its $5^d - 1$ nearest neighbors (as opposed to its $3^d - 1$ adjacent neighbors as discussed above). Implementing these smaller subcells multiplies the number of pair distances which must be evaluated during VL builds by a factor $\sim (5/6)^d$ and often speeds up VL building [11, 51]. However, it also leads to increased overhead because \mathcal{N}_{sc} increases by a factor $\sim 2^d$. We find that for the short-ranged interactions employed here, the reduction in the number of pair-distance calculations is outweighed by the increased overhead. However, the smaller-subcell approach is likely faster for sufficiently-long-ranged interactions and can be implemented by editing two lines of code in `main()`.

The linked subcells and machinery for VL-building are set up at the beginning of any simulation. Subcells centered at

$$\vec{R} = \sum_{i=1}^d \left(-\frac{1}{2} + \frac{C_i + 1/2}{n_{\text{cell}}} \right) L \hat{x}_i, \quad (8)$$

where C_i is an integer satisfying $0 \leq C_i < n_{\text{cell}}$, and the $-1/2$ term is present because the simulation cell is centered at the origin, are given the index

$$\mathcal{I} = \sum_{i=1}^d C_i n_{\text{cell}}^{d-i}. \quad (9)$$

Thus each subcell has a unique $\mathcal{I} \in \{0, 1, 2, \dots, \mathcal{N}_{\text{sc}} - 1\}$ as well as a unique set of $\{C_i : i = 1, 2, \dots, d\}$. These are defined when `setupsupercell()` is called at the beginning of a simulation.

Building the VL for any given atom requires calculating the distances between it and all other atoms in its subcell as well as the $3^d - 1$ neighboring subcells. This

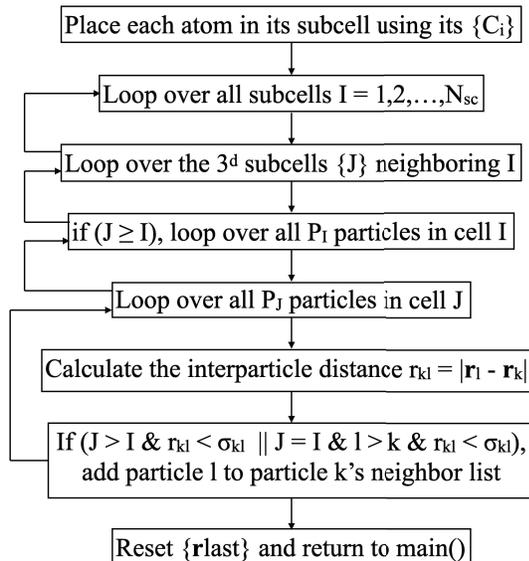


FIG. 1. Basic structure of the `buildneighborlist()` routine. Parallelization is achieved by dividing its outermost loop equally amongst the n_{threads} concurrent threads.

can be accomplished in many different ways. After considerable trial and error, we found that the fastest algorithm is schematically described by Figure 1. The basic features of this algorithm are standard for molecular simulations [4], but we found several details of this approach that are standard and work well for $d \leq 3$ perform poorly in higher d and needed to be modified as detailed below.

First, implementing the outer single loop over all cells $\mathcal{I} = 0, 1, \dots, N_{\text{sc}} - 1$ rather than a d -deep loop over the $\{C_i\}$ both provided a substantial speedup and allowed construction of a `buildneighborlist()` routine in which d is a parameter and thus works for arbitrary $d \geq 2$. A comparable speedup was achieved by implementing a single second-from-outer loop over each subcell’s 3^d linked subcells rather than the more geometrically intuitive approach of a d -deep loop over the linked subcells’ $\Delta C_i = -1, 0, 1$ for each $i = 1, 2, \dots, d$.

Second, while the most intuitively obvious way of handling the linked-subcell structure is to store it in static memory by identifying each subcell’s $3^d - 1$ neighboring subcells at the start of the simulation and including pointers to each neighbor in each subcell’s data structure (e.g. as a private variable in a “subcell” class object in C++) works well for $d \leq 3$, this approach has a substantial memory cost that significantly increases simulation runtimes in larger d . We found that substantially better performance is achieved when each subcell’s neighboring subcells are identified in the first step of the outermost loop in Fig. 1, i.e. each time the VL is built.

Third, we found that implementing the outer double loop over subcells and inner double loop over particles in these cells provides a substantial speedup relative to alternative loop orderings, and this speedup increases rapidly with increasing d . For example, we found that

the runtimes of simulations like the $d = 6$ $\phi = \phi_d$ runs discussed in Sections III-IV are $\sim 40\%$ larger when the order of the second and third loops in Fig. 1 is reversed.

The procedure described above requires $n_{\text{cell}} \geq 3$ to function correctly, and has an $\mathcal{O}(N^2)$ computational cost when $n_{\text{cell}} = 3$. Therefore we also included a more efficient $\mathcal{O}(N^2)$ VL-building routine `buildNsq()` that replaces it whenever $\text{floor}(L[\mathcal{R}^{\frac{1}{2d}}\tilde{\sigma} + s]^{-1}) \leq 3$. `buildNsq()` eliminates the subcells entirely and performs only the bottom three steps depicted in Fig. 1. For the remainder of this paper, we focus on larger systems where $n_{\text{cell}} > 3$.

D. SWAP Monte Carlo

Over the past decade, SWAP Monte Carlo, which speeds equilibration of deeply supercooled liquids by many orders of magnitude by exchanging particles’ diameters [20, 40], has revolutionized simulations of the glass/jamming transition. For example, showing that SWAP allows equilibration of hard-sphere liquids with $\phi > \phi_{\text{MRJ}}$ [52] proved that ϕ_{MRJ} is not the endpoint of the equilibrium-liquid branch of their phase diagram [53]. The algorithm was recently extended to $d > 3$; Ref. [22] showed that while the dynamical speedup it produces decreases exponentially with increasing d , it remains substantial for d as large as 8. Therefore we included SWAP capability in `hdMD`, and describe it here.

Once every n_{swap} timesteps, the `main()` routine calls `swapmove()`. This function controls all aspects of SWAPping. First it generates lists of the atom indices $\{i\}$ and $\{j\}$ for which the swaps $\sigma_i \leftrightarrow \sigma_j$, $\epsilon_i \leftrightarrow \epsilon_j$, and $m_i \leftrightarrow m_j$ will be attempted. Both $\{i\}$ and $\{j\}$ are of length pN (Sec. II B); the included indices are chosen randomly. Next it generates the VLs for these sets using the subroutine `buildNLS()`. This routine operates differently than `buildneighborlist()` for two reasons. First, to properly calculate potential energy changes, *full* VLs (as opposed to the half-VLs outlined in Section II C) are required. Second, there is no reason to loop over all N_{sc} subcells during this process. Instead, the subcell indices $\{\mathcal{I}_i\}$ and $\{\mathcal{I}_j\}$ are identified in advance. This process, as well as the actual population of the VLs $\{\mathcal{L}_i\}$ and $\{\mathcal{L}_j\}$, are parallelized. Specifically, the VL-building loops of length pN over the $\{i\}$ and $\{j\}$ are divided equally amongst the n_{threads} threads.

After the VLs are built, the actual swapping is executed serially, by the master thread. The total potential energy associated with interactions involving particles i and j is initially

$$E_b = \sum_k U_a(\epsilon_{ik}, \sigma_{ik}, r_{ik}) + \sum_l U_a(\epsilon_{jl}, \sigma_{jl}, r_{jl}), \quad (10)$$

where the sums are respectively over the particles $\{k\}$ and $\{l\}$ neighboring them. This energy becomes

$$E_a = \sum_k U_a(\epsilon_{jk}, \sigma_{jk}, r_{ik}) + \sum_l U_a(\epsilon_{il}, \sigma_{il}, r_{jl}). \quad (11)$$

if the swap is accepted, and the particles' velocities are rescaled to conserve kinetic energy. Swaps are accepted or rejected using the standard Metropolis criteria.

E. Temperature and pressure control, thermodynamics metrics

While the results presented below are all from NVT simulations, `hdMD` also includes a Berendsen barostat [54] that allows NPT simulations to be performed. The target pressure P_{targ} and target temperature T_{targ} can be set at the initiation of the MD run, or can be ramped by resetting them within `main()`'s main loop. Every time `Berendsen()` is called, it checks whether n_{cell} (Eq. 7) has changed since the last time it was called, and if it has, calls `setupsupercell()` and then `buildneighborlist()` to repopulate the subcells. As mentioned above, for NVT simulations (i.e. when `baroflag` is set to `false`), Berendsen thermostatting is implemented within `leapfrog()`.

`hdMD`'s `getthermo()` routine calculates standard thermodynamic quantities like the temperature, pair energy E_{pair} from Eq. 1 or its user-defined replacement, pressure P , and average coordination number Z . Calculation and output of these quantities is performed once every n_{thermo} steps. Since `hdMD` will likely be used primarily for glass/jamming-transition-related studies, it also (by default) calculates particles' mean squared displacements $\langle(\Delta\vec{r})^2\rangle$ and mean quartic displacements $\langle(\Delta\vec{r})^4\rangle$ since step n_{msdstart} , the non-Gaussian parameter

$$\alpha_2 = \frac{d\langle(\Delta\vec{r})^4\rangle}{(d+2)\langle(\Delta\vec{r})^2\rangle^2} - 1 \quad (12)$$

and overlap function

$$f_{\text{ov}} = N^{-1} \sum_{i=1}^N \Theta(0.1d\tilde{\sigma} - |\Delta\vec{r}_i|) \quad (13)$$

Eq. 12 is the d -dimensional generalization [30] of the usual 3D expression for α_2 ($3\langle(\Delta\vec{r})^4\rangle/5\langle(\Delta\vec{r})^2\rangle^2 - 1$). In Eq. 13, Θ is the Heaviside step function, and f_{ov} is a simple, commonly used metric for particle relaxation in deeply supercooled liquids. It varies continuously from 1 to zero as particles hop away from their initial positions, and provides roughly the same information as the self-intermediate scattering function $F_s(q, t)$ evaluated at $q = 2\pi/\tilde{\sigma}$ [20]. Finally, `hdMD`'s `getvanHove()` routine calculates the self part of the van Hove correlation function

$$G_s(r, t) = \frac{1}{N} \sum_{i=1}^N \delta(|\vec{r}_i(t) - \vec{r}_i(0)| - r), \quad (14)$$

where $\delta(x)$ is the Dirac delta function. Note that `hdMD`'s main loop is structured in such a way that additional periodic calculations of other thermodynamic quantities can easily be added by the user.

III. PERFORMANCE OF SIMULATIONS AT $\phi \simeq \phi_d$ FOR $3 \leq d \leq 6$

A. Theoretical Background

The dynamical glass transition packing fraction ϕ_d is defined as the packing fraction at which a supercooled d -dimensional hard-sphere liquid's diffusivity would drop to zero *if there were no hopping motion* [14]. Hopping makes the actual packing fraction at which diffusivity drops to zero substantially higher, but examining systems with $\phi \simeq \phi_d$ has proven very fruitful [30, 55] for understanding the ways in which finite- d supercooled liquids and glasses differ from their exactly-solvable, mean-field counterparts [36, 56]. In this section, we examine `hdMD`'s scalability and parallel efficiency for simulations of systems at packing fractions and temperatures that map to ϕ_d but have orders-of-magnitude-larger N than have been employed in previous $d > 3$ studies [15, 17, 18, 21–24, 30–32, 55–58]. We show that its performance (as judged by these metrics) is comparable to those achieved by popular $d = 3$ MD codes [5–10]. Then we examine how `hdMD`'s performance varies with d . We show that in the large- N limit it is nearly optimal, i.e. the runtime per pair distance calculation is nearly d -independent.

Results from simulations employing the Morse potential can be compared to hard-sphere results by considering systems at the same effective packing fraction $\phi_{\text{eff}} = \phi\sigma_{\text{eff}}^d$, where the temperature-dependent effective hard-sphere diameter [2] is

$$\sigma_{\text{eff}}(a, T) = \int_0^{\tilde{\sigma}} [1 - \exp(-U_a(r)/k_B T)] dr. \quad (15)$$

More accurate mapping to hard sphere results can be achieved using a refined version of this method [59], but since the primary focus of this paper is demonstrating the utility of `hdMD` rather than precisely matching hard-sphere results, we use Eq. 15 to estimate the dynamical-glass-transition packing fractions ϕ_d^* for $a = 30$ and $T = 0.25\epsilon/k_B$ (Table III).

TABLE III. Dynamical-glass-transition packing fractions ϕ_d for hard-spheres [30], the associated σ_{eff}^d values for $a = 30$ and $T = 0.25\epsilon/k_B$ (Eq. 15), and the packing fractions $\phi_d^* = [\sigma_{\text{eff}}(30, 0.25\epsilon/k_B)]^{-d}\phi_d$ employed in the simulations discussed in this section.

d	ϕ_d	$[\sigma_{\text{eff}}(30, 0.25\epsilon/k_B)]^d$	ϕ_d^*
3	0.5770	0.9650	0.5980
4	0.4036	0.9536	0.4232
5	0.2683	0.9423	0.2847
6	0.1723	0.9312	0.1850

B. Scaling of runtimes with N , n_{threads} , and d

Figure 2 summarizes **hdMD**'s scalability and parallel efficiency for short (100τ) simulations of supercooled liquids at this temperature and $\phi = \phi_d^*$ for $3 \leq d \leq 6$. All simulations include 10 FIRE energy minimizations and 100 SWAP-MC passes, respectively performed once every 10τ and once every τ after the simulation begins. The FIRE minimizations employ $E_{\text{thres}} = 10^{-25}\bar{\epsilon}$, and the SWAP passes employ $p = 0.1$, which is close to the optimal value [20]. For $a = 30$, $dt = 0.0016\tau$, so there are 625 MD timesteps per τ and thus 62500 total timesteps in each simulation.

Panels (a-b) show how runtimes increase with N for $10^{3.5} \leq N \leq 10^7$ for $3 \leq d \leq 4$ and $10^5 \leq N \leq 10^7$ for $5 \leq d \leq 6$. For $N \leq 10^5$, runtimes scale as N^y with $y \simeq 1.04$, which is very close to the optimal linear scaling. As dN increases beyond $\sim 10^6$, the particles' `rvf` array (Sec. II B) can no longer easily fit within the CPU's L3 cache, and calls to `getforce()` produce more and more cache misses. This worsens the runtime scaling to $1.1 \lesssim y \lesssim 1.15$, which is suboptimal, but only slightly so. Some standard ($d = 3$) MD codes that are optimized for large N (e.g. LAMMPS [5]) periodically re-order particle ids by the particles' positions along one dimension [50, 60]. This improves their large- N scaling by improving cache locality. Since $N = 10^5$ should be large enough for most glass-transition-related studies likely to be conducted in the next few years, we have not yet implemented this feature in **hdMD**, but it can easily be added; we may do so in the near future.

Panel (c) shows how parallel efficiency (PE) decreases as n_{threads} increases for $N = 10^6$ simulations on a typical mid-2010s dual-socket cluster node. For $n_{\text{threads}} = 2$, PE values are $\sim 85\%$. This is typical for numerical applications of OpenMP; PE is never 100% because there is a $\sim 10\mu\text{s}$ overhead associated with parallelizing any for loop. As discussed above, some of the tasks executed during these simulations, e.g. the assignment of particles to subcells and the SWAP moves, are not parallelized or readily parallelizable. Fortunately, PE drops only slowly as n_{threads} increases. As shown in the figure, a very loose lower bound for PE is $(83 - 1.5n_{\text{threads}})\%$, and higher- d simulations have considerably larger PE. For large N , **hdMD**'s efficiency-limiting factor appears to be the memory-bound force-evaluations; increasing n_{threads} increases the rate of cache misses in `getforce()`. Thus $\text{PE}(n_{\text{threads}})$ might also be improved by implementing particle-id reordering. We leave this for future work, but emphasize that since VL-building is less memory-bound than integration of the EOM, $\text{PE}(n_{\text{threads}})$ actually increases with increasing d .

Next we discuss how runtimes for fixed N and n_{threads} vary with d . In general, runtimes for fixed $\hat{\phi} = 2^d\phi$ must increase with d for two reasons. First, the effort required for VL building scales with $\mathcal{N}_{\text{VL}} \sim 3^d$ as discussed above. Second, the size of particles VLs in these runs scales as $\rho_d v(d, \sigma + s)$ (Eq. 2). For these $\phi = \phi_d^*$ runs,

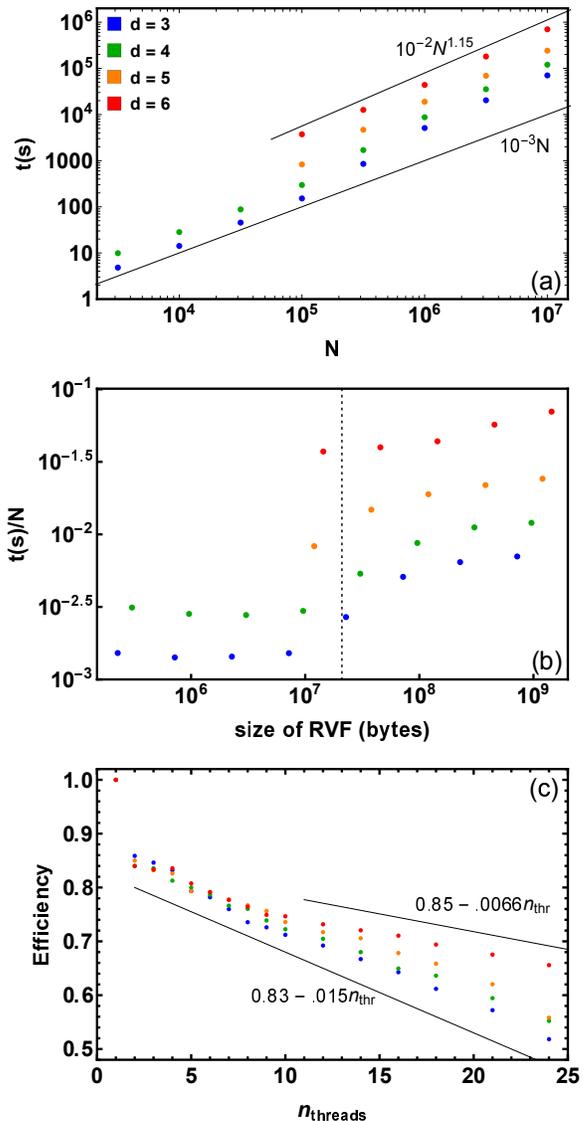


FIG. 2. Scaling of performance with N and n_{threads} for $3 \leq d \leq 6$. Runtimes for panels (a-b) are for $n_{\text{threads}} = 8$ simulations run on an iMac with a single 10-core Intel Core i9 CPU (3.6GHz). Results for $N < 10^5$ and $d > 4$ are not shown here because these N give $n_{\text{cell}} \leq 4$ and the scaling of the time devoted to VL builds is closer to $\mathcal{O}(N^2)$ than $\mathcal{O}(N)$. Runtimes for panel (c) are for $N = 10^6$ simulations on cluster nodes with two 12-core Intel Xeon E5-2650 CPUs (2.2GHz). The larger efficiencies for $n_{\text{threads}} > 12$ probably arise from distributing the computational effort to more than CPU socket. Solid lines in panels (a, c) are guides to the eye, and the dotted vertical line in panel (b) indicates the iMac's L3 cache size (20MB).

$\mathcal{N}_{\text{VL}} = 11.0, 19.4, 31.9, \text{ and } 50.4$ in $d = 3 - 6$. Thus the runtime *per pair distance calculation* in `getforce()`, i.e. $t/[n_{\text{steps}}\mathcal{N}_{\text{VL}}]$, is a good metric for comparing runtimes across different d . Results for all $N \geq 10^5$ systems are shown in Figure 3. For the smaller N , $t/\mathcal{N}_{\text{VL}}$ increases rapidly with d because n_{cell} drops to low values (e.g. $n_{\text{cell}} = 4$ for $N = 10^5$ and $d = 6$) and thus a large

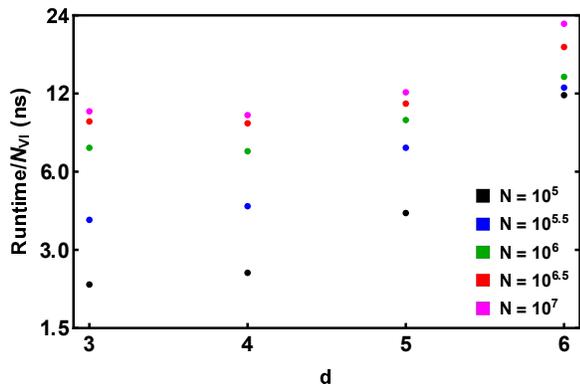


FIG. 3. Scaling of performance with d . Runtimes are for $n_{\text{threads}} = 8$ simulations run on an iMac with a single 10-core Intel Core i9 CPU (3.6GHz).

fraction of the N particles must be searched over during the building of each particle’s VL. For $N \geq 10^6$, however, $t/\mathcal{N}_{\text{VL}}$ increases far slower, only increasing by a factor of ~ 2 as d increases from 3 to 6. This indicates that for large simulations the scaling of **hdMD** runtimes with d is nearly optimal given that pair-distance calculations are the rate-limiting factor.

TABLE IV. Runtime percentages for $n_{\text{threads}} = 8$ simulations with $N = 10^5$ (top rows) and $N = 10^7$ (bottom rows).

d	Leapfrog	VL-building	Minimizations	SWAP	Other
3	83.9	11.1	2.1	1.2	1.7
4	72.0	22.4	1.5	2.9	1.2
5	47.5	45.9	1.0	4.6	1.0
6	26.0	67.2	0.7	5.3	0.7
3	90.6	5.5	1.5	0.6	1.7
4	88.2	7.6	1.5	1.4	1.4
5	72.2	20.1	1.3	3.5	3.0
6	42.8	43.5	0.8	4.2	1.8

The trends shown in Fig. 3 can be understood by examining in greater detail how simulation runtimes are divided among **hdMD**’s various routines. Table IV lists the runtime percentages devoted to leapfrog integration, VL building, energy minimizations, and SWAP for $N = 10^5$ and $N = 10^7$, which are representative of simulations of moderately-sized and very-large systems. For both N , the percentage of simulation runtimes spent in VL-building (leapfrog integration) increases (decreases) rapidly with increasing d . However, the fractional increases/decreases in these percentages as d increases from 3 to 6 are far greater for $N = 10^5$ than for $N = 10^7$. This explains why the runtimes per pair distance calculation increase faster with d when N is smaller and vice versa.

Taken together, the above results show that **hdMD** makes large- N simulations practical in d up to 6, even with modest computational resources. In the following

section, we demonstrate how this feature can be exploited to obtain novel physics results.

IV. HETEROGENEOUS DYNAMICS IN DEEPLY SUPERCOOLED $d = 6$ LIQUIDS

A. Breakdown of the Stokes-Einstein relation

Three recent simulation studies [30–32] of dynamical heterogeneity in supercooled liquids in $3 \leq d \leq 10$ have reported that it weakens with increasing d . However, as mentioned in the Introduction, it may be that the heterogeneous dynamics in these simulations *artificially* weakened with increasing d because they employed small fixed $N < 10^4$ and periodic boundary conditions with $L \sim N^{1/d}$ that may have dropped below the characteristic size \mathcal{L} of the liquids’ spatial heterogeneities [61]. Whether this is so can be determined by simulating liquids with fixed L rather than fixed N over a comparable range of d , taking care that $L > \mathcal{L}$ for all d [62]. As described above, **hdMD**’s efficient parallel implementation makes it well suited to doing so. Here we motivate such studies by showing that $d = 6$ supercooled liquids can be substantially more heterogeneous than previously reported, and provide some initial evidence that the strengths of the Stokes-Einstein-relation (SER) breakdowns reported in Refs. [30–32] were artificially suppressed by the small system sizes they employed.

Figure 4 shows dynamical results for supercooled $d = 6$ liquids with $1.02\phi_d^* \leq \phi \leq 1.05\phi_d^*$ (Section III A, Table III). To obtain a cleaner comparison to the results of Refs. [30–32], all particles in these simulations were given equal masses ($m_i = \tilde{m} = 1$). All systems have $N = 10^5$ and were SWAP-equilibrated at $k_B T = 0.25\tilde{\epsilon}$ for at least $1.75 \times 10^4 \tau$. Following the equilibration runs, SWAP was turned off and systems were evolved forward in time for another $\mathcal{T} = 10^5 \tau$ (also at $k_B T = 0.25\tilde{\epsilon}$). No ensemble or time averaging was performed.

Panel (a) shows particles’ mean squared displacements $\langle [\Delta \vec{r}(t)]^2 \rangle$. The increasing MSD plateau length with increasing ϕ is typical for supercooled liquids [63], as is the gradual increase in the slope $\zeta(t) \equiv d \ln \langle [\Delta \vec{r}(t)]^2 \rangle / d \ln(t)$ towards 1 as systems approach the Brownian-diffusive regime. All systems reach this regime well before $t = \mathcal{T}$, and the values of $\langle [\Delta \vec{r}(\mathcal{T})]^2 \rangle / \tilde{\sigma}^2$ are all well above 1, indicating that typical particles have hopped multiple times by the end of these simulations even for $\phi = 1.05\phi_d^*$.

Panel (b) shows particles’ overlap parameter $f_{\text{ov}}(t)$. Again, all results are typical for supercooled liquids. Following Refs. [20, 32], we define the alpha relaxation times (τ_α) for these liquids using the criterion $f_{\text{ov}}(\tau_\alpha) = e^{-1}$. Numerical results for τ_α , the diffusion coefficients $D = \lim_{t \rightarrow \infty} \langle [\Delta \vec{r}(t)]^2 \rangle / 2dt$, and several related quantities are shown for a wider range of ϕ in Table V. As shown in the inset, the $f_{\text{ov}}(t)$ curves do not collapse when plotted vs. τ/τ_α , indicating that time-density superposition breaks down in these systems. $f_{\text{ov}}(t/\tau)$ decreases faster as ϕ in-

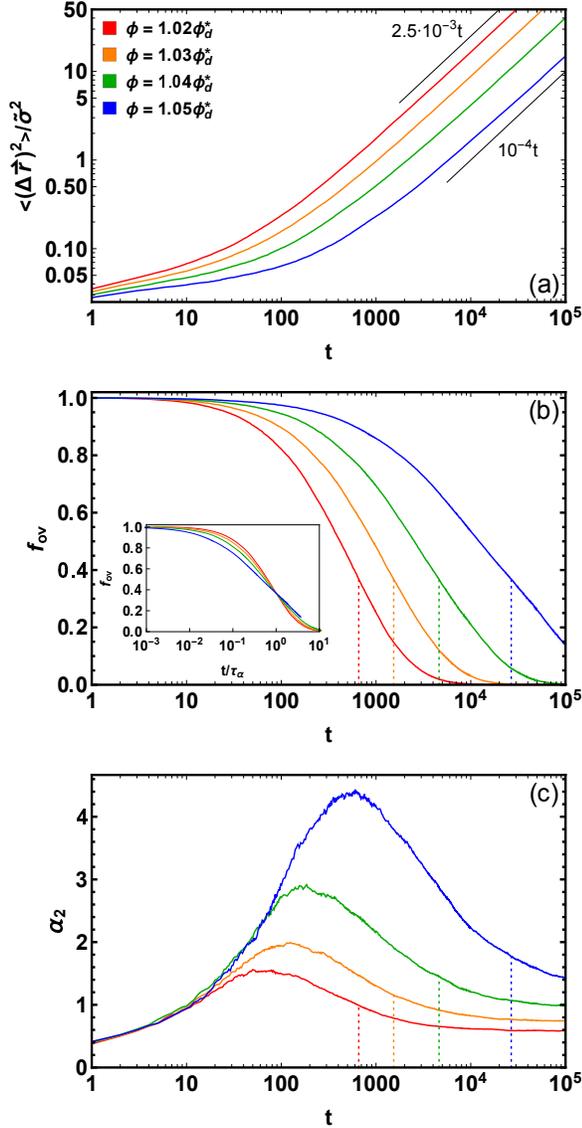


FIG. 4. Dynamics of supercooled $d = 6$ liquids for $1.02\phi_d^* \leq \phi \leq 1.05\phi_d^*$ and $k_B T = 0.25\tilde{\epsilon}$. Panels (a-c) respectively show the mean-squared displacement, overlap parameter f_{ov} (Eq. 13), and non-Gaussian parameter (Eq. 12). Straight gray lines in panel (a) are guides to the eye, and dotted vertical lines in panels (b-c) indicate $t = \tau_\alpha$. All times are given in units of $\tau = \sqrt{\tilde{m}\tilde{\sigma}^2/\tilde{\epsilon}}$ (Sec. II A).

creases because the dynamics of small and large particles have decoupled; larger particles' mobility is decreasing faster with increasing ϕ than that of their smaller counterparts. Such decoupling has long been associated with dynamical heterogeneity [63–65].

Panel (c) shows results for the non-Gaussian parameter $\alpha_2(t)$. As expected from previous $d = 3$ studies [63], results for different ϕ fall on a common curve at small t and exhibit maxima $\alpha_{2,\max}(\phi)$ at times $\tau^*(\phi)$ that increase with ϕ slower than $\tau_\alpha(\phi)$. As expected for relatively-low-temperature liquids, $\alpha_{2,\max} \sim \tau_\alpha^x$ with $x \simeq 0.3$ [66–68].

Intriguingly, the $\alpha_2(\tau_\alpha)$ values increase roughly logarithmically with τ_α as ϕ increases rather than as a power law, i.e. $\alpha_2(\tau_\alpha) \sim \ln(\tau_\alpha)$.

At longer times, rather than trending back to zero as is the case when monodisperse systems are considered or α_2 is calculated using only one component of a bidisperse mixture [30–32, 63], all systems' $\alpha_2(\phi, t)$ decay very slowly (over timescales of order $10^2\tau_\alpha$) towards finite plateau values $\alpha_{2,\infty}$ that increase rapidly with ϕ . Finite $\alpha_{2,\infty}$ are expected since smaller particles have larger diffusion coefficients $D_i = \tilde{D}(\sigma_i)$, where \tilde{D} is an *a priori* unknown function that captures the particle-size dependence of diffusivity. Quantitatively, one expects [69]

$$\alpha_{2,\infty} = \frac{\langle \tilde{D}^2 \rangle}{\langle \tilde{D} \rangle^2} - 1, \quad (16)$$

where

$$\langle \tilde{D}^n \rangle = \int_{\sigma_{\min}}^{\sigma_{\max}} P(\sigma) [\tilde{D}(\sigma)]^n d\sigma. \quad (17)$$

For the particle-size distribution employed here [i.e. $P(\sigma) \equiv P(6, 6, \sigma)$ from Eq. 3], assuming all particles obey the classical relation $\tilde{D}(\sigma) \propto k_B T (\eta\sigma)^{-1}$ [where η is viscosity] predicts $\alpha_{2,\infty} \simeq .047$. Actual $\alpha_{2,\infty}$ values are much larger and increase rapidly with increasing ϕ , consistent with the well-known result that deviations of $\tilde{D}(\sigma)$ from this formula strengthen with increasing ϕ or decreasing T [70]. This contributes to the abovementioned breakdown of t - ϕ superposition. One expects, based on previous studies of polydisperse $d = 3$ supercooled liquids performed as far back as the mid-2000s [71, 72], that it will also contribute to the SER breakdown, indicated by the increase in $D\tau_\alpha$ with increasing ϕ , that occurs as particle motion becomes increasingly hopping-dominated.

While the qualitative behaviors summarized in Figure 4(c) and Table V are unremarkable in and of themselves, they *are* noteworthy because they show in two distinct ways that high- d supercooled liquids can be more heterogeneous than previously reported. First, the $\alpha_{2,\max}$ values are substantially higher than any reported in Refs. [30, 32], neither of which showed any $\alpha_{2,\max} > 1.6$ for $d = 6$ liquids at any T or ϕ . The recently demonstrated one-to-one correspondence between $\alpha_{2,\max}$ and the kinetic fragility m^* [66] implies that these liquids are also more fragile than any $d = 6$ liquids studied in Refs. [30, 32]. Second, they show that the SER violations in these liquids (as quantified via the relation $D\tau_\alpha \sim \tau_\alpha^\omega$) can be much stronger than observed in Refs. [31, 32].

The quantity ω is of particular interest for its ability to shed light on the d -dependence of dynamical heterogeneity in supercooled liquids [62]. D is dominated by the fastest (smallest) particles, while τ_α is primarily set by the slowest (largest) particles [72]. Since τ_α increases with ϕ faster than D decreases, the product $D\tau_\alpha$ increases with both ϕ and τ_α , implying $\omega > 0$. The strength of this effect should decrease with increasing d because particles' cages become more mean-field-like

TABLE V. Measures of mobility and dynamical heterogeneity in supercooled $d = 6$ liquids. τ_α and τ^* are given in units of $\tau = \sqrt{m\tilde{\sigma}^2/\tilde{\epsilon}}$ (Sec. II A). Results for $\phi \leq 1.04\phi_d^*$ are well fit by $D \sim (\tilde{\phi}_d - \phi)^{\gamma_D}$ and $\tau_\alpha \sim (\tilde{\phi}_d - \phi)^{-\gamma_\tau}$, with $\tilde{\phi}_d = 1.068\phi_d^* = .1976$, $\gamma_D \simeq 8/3$, and $\gamma_\tau \simeq 10/3$. $\alpha_{2,\infty}$ values are given only for systems in which $\alpha_2(t)$ has clearly reached its plateau value by $t = 10^5\tau$.

ϕ/ϕ_d^*	$D\tilde{\sigma}^{-2}\tau$	τ_α	$\alpha_{2,\max}$	τ^*	$\alpha_2(\tau_\alpha)$	$\alpha_{2,\infty}$
1.00	4.11×10^{-3}	2.09×10^2	1.06	30	0.735	0.41
1.005	3.40×10^{-3}	2.63×10^2	1.15	32	0.769	0.45
1.01	2.67×10^{-3}	3.58×10^2	1.24	59	0.847	0.49
1.015	2.13×10^{-3}	4.77×10^2	1.40	66	0.898	0.54
1.02	1.62×10^{-3}	6.57×10^2	1.54	75	0.992	0.59
1.025	1.22×10^{-3}	9.92×10^2	1.79	91	1.079	–
1.03	8.72×10^{-4}	1.54×10^3	2.00	127	1.154	–
1.035	6.09×10^{-4}	2.58×10^3	2.40	126	1.321	–
1.04	4.02×10^{-4}	4.63×10^3	2.93	158	1.455	–
1.045	2.52×10^{-4}	1.03×10^4	3.56	295	1.613	–
1.05	1.46×10^{-4}	2.67×10^4	4.44	598	1.770	–

[30]. Mean-field theories predict $D \sim (\phi_d - \phi)^\gamma$ and $\tau_\alpha \sim (\phi_d - \phi)^{-\gamma}$ as ϕ approaches ϕ_d from below, implying $\omega = 0$. Additional theoretical analyses predict that ω should vanish above the upper critical dimension $d_u = 8$ [73]. Numerical results in Refs. [31, 32] were consistent with this hypothesis, and suggested $\omega \sim (d_u - d)$. On the other hand, studies of the mean-field Mari-Kurchan model [55] showed $\omega \simeq .22$ for all $2 \leq d \leq 6$, while a recent study of the kinetically constrained East model showed [74] that ω remains finite for all $d \leq 10$ and may remain finite in the $d \rightarrow \infty$ limit, suggesting that this issue has not yet been resolved.

The $D(\phi)$ and $\tau_\alpha(\phi)$ data shown in Table V are qualitatively consistent with those reported in many previous studies. For ϕ not too close to the packing fraction $\tilde{\phi}_d$ where diffusive motion ceases, $D \sim (\tilde{\phi}_d - \phi)^{\gamma_D}$ and $\tau \sim (\tilde{\phi}_d - \phi)^{-\gamma_\tau}$, where $\tilde{\phi}_d$ is several percent above ϕ_d^* [55] and (in contrast to mean field theories) $\gamma_\tau > \gamma_D$. Thus $D\tau_\alpha \sim (\tilde{\phi}_d - \phi)^{\gamma_D - \gamma_\tau} \sim \tau_\alpha^\omega$ with $\omega = 1 - \gamma_D/\gamma_\tau$. Figure 5 shows $D\tau_\alpha$ vs. τ_α for these systems. Lower-density ($\phi \lesssim 1.02\phi_d^*$) liquids' results fall on a common curve $D\tau_\alpha \sim \tau_\alpha^\omega$ with $\omega \simeq 0.2$. At higher densities, a crossover to a stronger dependence $D\tau_\alpha \sim \tau_\alpha^y$ is observed as systems become sluggish, consistent with Fig. 7(b) of Ref. [31]. Overall, the trends are the same as found in Refs. [31, 32], but the ω value is more than twice as large as in these studies, which respectively found $\omega \simeq .09$ [31] and $\omega \simeq .083$ [32] in comparable $d = 6$ liquids.

There are multiple potential reasons for this difference. For example, here we have employed a moderate-stiffness ($a = 30$) Morse pair potential and moderate temperature. In contrast, Refs. [30, 31] employed hard spheres while Ref. [32] employed soft harmonic spheres at very low T . Thus our liquids experience thermal activation

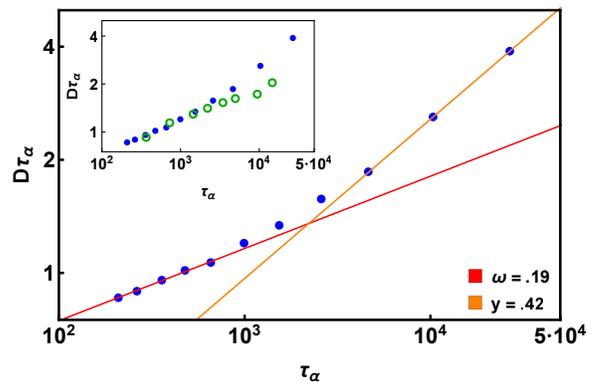


FIG. 5. Breakdown of the Stokes-Einstein relation in supercooled $d = 6$ liquids at $k_B T = 0.25\tilde{\epsilon}$. Symbols show MD data while lines show fits to $D\tau_\alpha \sim \tau_\alpha^\omega$ and $D\tau_\alpha \sim \tau_\alpha^y$. The inset contrasts results for $N = 10^5$ (same symbols shown in the main panel) to $N = 5000$ results for selected $1.01\phi_d^* \leq \phi \leq 1.055\phi_d^*$ (open red circles); all results for $N = 5000$ are averaged over ten independently prepared systems. Note that the range of τ_α depicted here is almost identical to that considered in Ref. [32].

over energy barriers (absent from [30, 31]) and substantially higher mobilities than those of [32].

Another potential reason is that our systems are much larger, with $N = 10^5$ rather than $N = 5000 - 8000$ as was the case in Refs. [30–32]. As discussed above, one expects dynamical heterogeneity to increase with system size. To investigate this possibility, we characterized the dynamics of $N = 5000$ systems at the same $\{\phi\}$ and T . We found that D decreases with ϕ slightly slower in these liquids than in their $N = 10^5$ counterparts, but τ_α increases substantially slower, particularly for $\phi \gtrsim 1.03\phi_d^*$. Results for $D\tau_\alpha$ for these liquids are shown in the inset to Fig. 5. Plainly $D\tau_\alpha$ grows slower with increasing ϕ than in the $N = 10^5$ liquids, and as a consequence, the *apparent* breakdown of the SER is weaker. This difference presumably arises because periodic boundary conditions cap the characteristic size of cooperatively rearranging regions within a model supercooled liquid at L ; the $N = 5000$ liquids' smaller L reduces the characteristic size of their cooperatively rearranging regions and hence their τ_α [28]. If true, this would explain these liquids' delayed crossover to the stronger $D\tau_\alpha \sim \tau_\alpha^y$ scaling.

It is reasonable to suppose that when comparing systems with fixed N and ϕ/ϕ_d^* across multiple d as was done in Refs. [30–32], the decrease in L with increasing d produces a comparable (artificial) reduction in the measured $D\tau_\alpha$ and perhaps also in the inferred ω [75]. As mentioned above, this hypothesis could be tested using simulations where L rather than N is fixed [62]. Our present focus is not to resolve this issue – finite-size effects on the dynamics of supercooled liquids are decidedly nontrivial [61] – but rather to demonstrate that hdMD is well-suited to doing so.

B. Comparison to bidisperse systems

A third potential reason for the larger $\alpha_{2,\max}$ and ω reported above is our use of continuously-polydisperse $P(\sigma)$ (Eq. 3). To investigate this possibility, we repeated the $N = 10^5$ studies highlighted in Figs. 4-5, using the 50:50 1:1.4 bidisperse $P(\sigma)$ employed in Ref. [32] and many other studies of the glass-jamming transition [38]. For maximal consistency with Ref. [32] and other previous studies, we set $\epsilon_{\text{small}} = \epsilon_{\text{large}} = \tilde{\epsilon}$ [47].

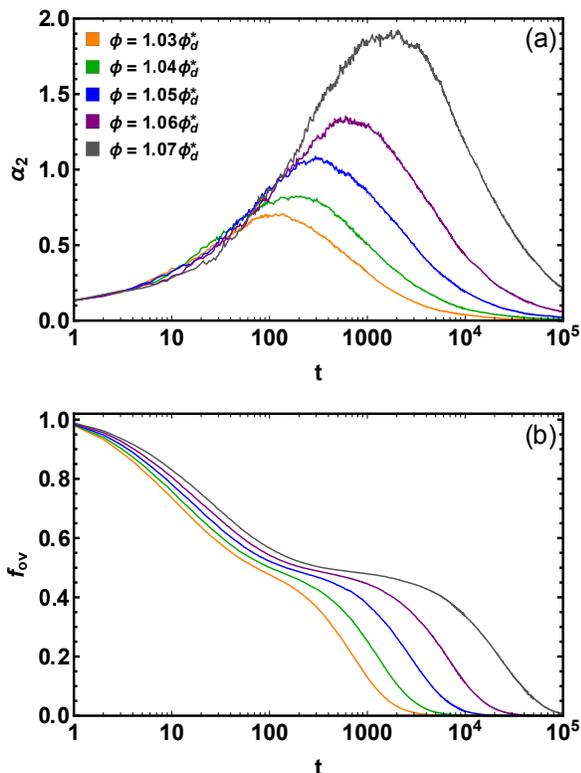


FIG. 6. Heterogeneous dynamics of supercooled 50:50 1:1.4 bidisperse $d = 6$ liquids. Following Refs. [32, 63], we calculated $\alpha_2(t)$ and $f_{\text{ov}}(t)$ separately for small and large particles. Panel (a) shows $\alpha_2(t)$ for the large particles, while panel (b) shows $f_{\text{ov}}(t)$ for all particles.

Figure 6 illustrates two aspects of these bidisperse liquids' heterogeneous dynamics. Panel (a) shows the large particles' $\alpha_2(t)$ for selected ϕ . The finite- $\alpha_{2,\infty}$ plateaus vanish, as expected [63]. Compared to results shown in Fig. 4(c), the $\alpha_{2,\max}$ are lower for systems with comparable τ^* . While they are substantially higher than those reported in Ref. [32], they are comparable to those reported in Ref. [30]. For $\phi \lesssim 1.06\phi_d^*$, if only large particles are used to estimate both D and τ_α , these systems have $D\tau_\alpha \sim \tau_\alpha^\omega$ with $\omega \simeq 0.1$, consistent with Refs. [30–32],

Other metrics, however, indicate that dynamics in these liquids are in fact far more heterogeneous than their $P \propto \sigma^{-d}$ counterparts, as might have been expected from their larger size asymmetry. For example, their $f_{\text{ov}}(t)$ [panel (b)] indicate a decoupling of large and small parti-

cles' dynamics that is much stronger than that shown in Fig. 4(b). These data raise the question: which method of averaging dynamics results from polydisperse supercooled liquids best captures their essential physics?

C. Non-Gaussian particle caging

Closely related to the above discussion is the issue of caging. The probability $P(\vec{r}, t)$ that a particle initially located at the origin is at position \vec{r} at time t is $P(r, t) = G_s(r, t)/A(d, r)$, where

$$A(d, r) = \frac{d\pi^{d/2}r^{d-1}}{\Gamma(1 + d/2)} \quad (18)$$

is the area of a d -dimensional hyperspherical shell of radius r ; here we have assumed isotropy in rewriting $P(\vec{r}, t)$ as $P(r, t)$. Einstein's theory of Brownian motion predicts that $P(r, t)$ is Gaussian, and the central limit theorem requires that it become Gaussian after sufficiently long times. At shorter times, however, $P(r, t)$ is non-Gaussian in a very wide variety of systems, including systems near glass and jamming transitions [33, 34]. Exponential tails of form $P_E(r, t) \propto \exp[-r/\Lambda(t)]$ are universal in systems where particles have hopped a random number of times [34, 76], with $\lambda(t)$ typically growing either as $t^{1/2}$ or as $t^{1/d}$ [34, 76–81]. In a dynamically heterogeneous liquid, these tails correspond to the high-mobility particles.

Quantitatively predicting how $P(r, t)$ varies with ϕ and/or T in arbitrary d is an obvious goal for any theory of liquid-state dynamics. Replica theory [35, 36] and dynamic DFT [82] assume that it is Gaussian. MCT [83, 84], RFOT [85], the nonlinear Langevin equation theory [65, 86], dynamical-facilitation-based theories [87], and CTRW-based theories [34, 76–78] all predict non-Gaussian $P(r, t)$, with varying degrees of success. One might expect from the increase in the number of near neighbors and the simplification of liquids' local structure as d increases [15] that $P(r, t)$ will converge to a Gaussian form even at short times if d is sufficiently large. Ref. [30] showed that in fact no such convergence occurs for $t \lesssim \tau_\alpha$ over the range $3 \leq d \leq 8$, but did not examine any $t \gg \tau_\alpha$. In light of the results presented in Figs. 4-6, it is worthwhile to examine how our larger systems' $P(r, t)$ behave in this long-time limit.

Figure 7 shows $P(r, t)$ for the bidisperse liquids discussed above. Data shown in panel (b) are for a density at which the glassy dynamics are about ten times slower than those illustrated in panel (a). Both panels show the same trends observed in $d = 3$ [34, 79–81]: the $P(r, t)$ are initially dominated by their exponential tails but then slowly cross over towards a Gaussian form as t increases. Two notable features are apparent.

First, contrary to what might be expected in higher- d liquids but consistent with the slow decays of $\alpha_2(t)$ illustrated in Fig. 6(a), substantial exponential tails are evident even for $t = 10\tau_\alpha$. CTRW-based theories of diffusion [76–78] predict that $P(r, t)$ converges to a fully

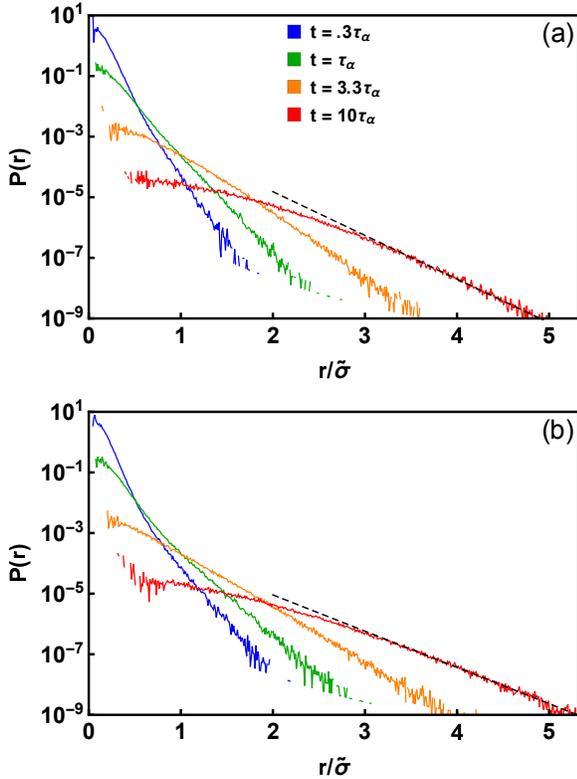


FIG. 7. Non-Gaussian caging in supercooled 50:50 1:1.4 bidisperse $d = 6$ liquids. Panels (a-b) show results for $\phi = 1.03\phi_d^*$ and $\phi = 1.06\phi_d^*$; these systems respectively have $\tau_\alpha \simeq 7.8 \times 10^2$ and $\tau_\alpha = 7 \times 10^3$. The dashed lines show fits to exponential tails with $\lambda = 0.301\bar{\sigma}$ and $\lambda = 0.363\bar{\sigma}$.

Gaussian form only after all (or nearly all) particles have hopped multiple times. For $\phi \gtrsim \phi_d^*$, where mobility is hopping-dominated and times between hops are broadly distributed, this convergence should occur only for $t \gg \tau_\alpha$, independent of d . Second, although $P(r, t)$ for fixed t/τ_α and different ϕ are qualitatively similar, they do not collapse. The exponential-tail lengths $\lambda(t/\tau_\alpha)$ clearly increase faster for $\phi = 1.06\phi_d^*$ than for $\phi = 1.03\phi_d^*$. This result is consistent with the decoupling illustrated in Figs. 4(b) and 6(b); $\lambda(t/\tau_\alpha)$ grows faster in the higher- ϕ system because the disparity in mobility between large and small particles is greater.

Since the slow crossovers to Gaussian $P(r, t)$ have been interpreted [79, 80] as a slow approach to ergodicity, suggesting that they provide a useful metric for understanding how ergodicity breaks, it would be very interesting to quantitatively compare them across multiple d . While we leave this as a challenge for future work, we emphasize here that quantitative analyses of these crossovers are likely to suffer from spurious finite-size effects if the crossovers are not complete by the time the most mobile particles have traveled distances $\gtrsim L$. Taken together, the results shown in Figs. 5 and 7 suggest that avoiding such effects requires $L \gtrsim 5\sigma$, again emphasizing the need for an efficiently parallelized code like `hdMD`.

V. DISCUSSION AND CONCLUSIONS

In this paper, we described a new public-domain, open-source parallel molecular dynamics simulation package (`hdMD`) that is optimized for high spatial dimensions. Four aspects of `hdMD`'s algorithmic implementation differ from those employed in most standard MD codes [5–10]. First, since parallelization of the force evaluations by spatial domain composition works less well in large d than it does in $d \leq 3$ (owing to the larger fraction of any spatial domain that is within a distance $\bar{\sigma} + s$ of its surface), `hdMD` instead employs per-atom parallelization. Second, to further reduce interprocessor communication, `hdMD` uses a shared-memory OpenMP-based parallelization strategy rather than the more commonly employed distributed-memory MPICH-based approach. Third, to avoid the large-for-high- d memory overhead associated with storing pointers to each linked subcell's $3^d - 1$ neighboring subcells, these subcells are instead efficiently identified on the fly each time the VLs are built. Fourth, d is a parameter rather than a fixed quantity in `hdMD`'s various subroutines, all of which have been tested for all $2 \leq d \leq 10$ and (in principle) work in arbitrary d .

`hdMD` is designed for maximum flexibility and extensibility. For example, while above we presented results for a single (stiff repulsive Morse) pair potential, using a different potential or interaction cutoff radius requires only editing a few lines in `getforce()`, `getthermo()`, and `swapmove()`. Incorporating additional diagnostics, e.g. calculation of the self-intermediate scattering function $F_s(q, t)$, is intended to be comparably straightforward. For this reason, `hdMD` is written in “plain vanilla” C++, and software tricks like advanced vectorization techniques, SIMD or AVX intrinsics [88], and GPU offloading of the type used in several popular MD packages [5, 9, 89], all of which can substantially increase a code's speed but often severely reduce its legibility, have not yet implemented. Adding any of these could substantially accelerate the code.

By examining the scalings of parallel simulation runtimes with the number of particles N and the number of simulation threads n_{threads} , we showed that three aspects of `hdMD`'s performance are already nearly optimal. First, the runtimes scale as $t \sim N$ when N is small enough for the particles' position-velocity-force (`rvf`) array to fit in the CPU's L3 cache. Second, for large N , the runtimes per force evaluation increase only slowly with increasing d , e.g. by only a factor of ~ 2 over the range $3 \leq d \leq 6$ for simulations of $N = 10^6$ supercooled liquids at $\phi = \phi_d^*$. Since the computational effort to rebuild all particles' Verlet lists scales as $3^d N$, this small increase is a major strength of the code. Third, `hdMD`'s parallel efficiency is comparable to that of popular public-domain MD codes (at least for selected problems [46, 90]), and actually increases with increasing d owing to its efficiently parallelized VL-building.

The total “size” of each simulation described in Section IV (as defined by the number of particles times the dura-

tion of the simulation) was $N\mathcal{T} = 10^{10}\tau$, making them among the largest $d > 3$ supercooled-liquid simulations ever performed. We found that dynamical heterogeneity in supercooled $d = 6$ liquids can be substantially greater than previously reported [30–32]. In particular, we found that the $D\tau_\alpha \sim \tau_\alpha^\omega$ scaling in continuously-polydisperse systems with $\Delta = .083$ has $\omega \simeq 0.2$, which is about twice the value previously reported [31, 32] for $d = 6$. Simulations of bidisperse systems showed $\omega \simeq 0.1$, but also demonstrated that particle caging can remain substantially non-Gaussian [as indicated by long exponential tails in particles’ displacement-probability distributions $P(r, t)$] for times as large as $10\tau_\alpha$. These dynamics appear to be consistent with recently proposed, CTRW-based theories of diffusion in systems for which the exponential tails of $P(r)$ correspond to particles that have hopped a random number of times [76–78].

We also showed that the crossover to the stronger $D\tau_\alpha \sim \tau_\alpha^\eta$ scaling that occurs as the continuously-polydisperse systems become sluggish [31] occurs at a density that decreases substantially when N is increased from 5000 to 10^5 . This decrease may arise from larger systems’ ability to accommodate larger cooperatively rearranging regions (CRRs). Specifically, our results are consistent with the hypothesis that for $\phi > \phi_c(N)$, τ_α grows faster with ϕ in larger systems of size $N' > N$ than in smaller systems of size $N'' < N$ because the former can accommodate larger CRRs which have a correspondingly larger τ_α [28]. While the validity of this hy-

pothesis can depend on both temperature and the model employed [61], our results nonetheless suggest that the conclusions of many previous studies of supercooled liquids in $d > 3$ which employed fixed $N < 10^4$ and L that decrease as $N^{-1/d}$ (e.g. Refs. [15, 17, 18, 22–24, 30–32, 55–57]) may have been substantially influenced – at least in their quantitative details – by finite-size effects. We have demonstrated that **hdMD** is well-suited to determining whether this is so.

Finally we emphasize that **hdMD** is also well-suited to studying open problems that are less directly related to the glass-jamming transition. For example, studies of melting dynamics across multiple d can shed light on how melting is affected by the symmetries of the crystal lattice and by decorrelation [15] of the liquid state. Previous studies of melting in $d > 4$ (e.g. [21, 91–93]) have all employed $N < 6 \times 10^4$, and most have employed much smaller systems; this has severely limited the accessible size range of any crystal-fluid interfaces. We have used **hdMD** to simulate the (nonequilibrium) melting of an $N = 6.25 \cdot 10^5$ -atom E_7 crystal (the densest lattice in $d = 7$) subjected to a temperature ramp at constant pressure and will report our results elsewhere.

The **hdMD** source code is publicly available and can be downloaded from our group website (<http://labs.cas.usf.edu/softmattertheory/hdMD.html>).

We are grateful to Patrick Charbonneau for numerous helpful discussions. This material is based upon work supported by the National Science Foundation under Grant No. DMR-2026271.

-
- [1] W. G. Hoover and F. H. Ree, “Melting transition and communal entropy for hard spheres,” *J. Chem. Phys.* **49**, 3609 (1968).
- [2] J. D. Weeks, D. Chandler, and H. C. Andersen, “Role of repulsive forces in determining equilibrium structure of simple liquids,” *J. Chem. Phys.* **54**, 5237 (1971).
- [3] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Clarendon Press, Oxford, 1987).
- [4] D. Frenkel and B. Smit, *Understanding Molecular Simulations, 2nd edition* (Academic Press (San Diego), 2002).
- [5] LAMMPS: <https://www.lammps.org/>.
- [6] GROMACS: <https://www.gromacs.org/>.
- [7] NAMD: <https://www.ks.uiuc.edu/Research/namd/>.
- [8] AMBER: <https://ambermd.org/>.
- [9] HOOMD: <http://glotzerlab.engin.umich.edu/hoomd-blue/>.
- [10] RUMD: <http://rumd.org/>.
- [11] S. Plimpton, “Fast parallel algorithms for short-range molecular dynamics,” *J. Comp. Phys.* **117**, 1 (1995).
- [12] J. N. Glosli, K. J. Casperson, J. A. Gunnels, D. F. Richards, R. E. Rudd, and F. H. Streitz, “Extending stability beyond CPU millennium: a micron-scale atomistic simulation of Kelvin-Helmholtz instability,” SC ’07: Proceedings of the 2007 ACM/IEEE Conference on Supercomputing (2007), 10.1145/1362622.1362700.
- [13] The publicly available **pyCudaPacking** package (<https://github.com/SimonsGlass/pyCudaPacking>) implements GPU-parallelized energy minimization, but not standard molecular dynamics routines, for d -dimensional systems. During revision of this manuscript, we learned of a publicly-available, open-source, GPU-parallelized, d -dimensional MD code (<https://doi.org/10.5281/zenodo.6368329>) that has not yet been described or employed in any published studies.
- [14] T. R. Kirkpatrick, D. Thirumalai, and P. G. Wolynes, “Scaling concepts for the dynamics of viscous liquids near an ideal glassy state,” *Phys. Rev. A* **40**, 1045 (1989).
- [15] M. Skoge, A. Donev, F. H. Stillinger, and S. Torquato, “Packing hyperspheres in high-dimensional euclidean spaces,” *Phys. Rev. E* **74**, 041127 (2006).
- [16] R. Brüning, D. A. St-Onge, S. Patterson, and W. Kob, “Glass transitions in one-, two-, three-, and four-dimensional binary Lennard-Jones systems,” *J. Phys. Cond. Matt.* **21**, 035117 (2009).
- [17] J. A. van Meel, B. Charbonneau, A. Fortini, and P. Charbonneau, “Hard-sphere crystallization gets rarer with increasing dimension,” *Phys. Rev. E* **80**, 061110 (2009).
- [18] P. Charbonneau, A. Ikeda, G. Parisi, and F. Zamponi, “Glass transition and random close packing above three dimensions,” *Phys. Rev. Lett.* **107**, 185702 (2011).
- [19] T. S. Ingebrigtsen, J. C. Dyre, T. B. Schröder, and C. P. Royall, “Crystallization instability in glass-forming mix-

- tures,” *Phys. Rev. X* **9**, 031016 (2019).
- [20] A. Ninarello, L. Berthier, and D. Coslovich, “Models and algorithms for the next generation of glass transition studies,” *Phys. Rev. X* **7**, 021039 (2017).
- [21] P. Charbonneau, C. M. Gish, R. S. Hoy, and P. K. Morse, “Thermodynamic stability of hard sphere crystals in dimensions 3 through 10,” *Eur. Phys. Journ. E* **44**, 101 (2021).
- [22] L. Berthier, P. Charbonneau, and J. Kundu, “Bypassing sluggishness: Swap algorithm and glassiness in high dimensions,” *Phys. Rev. E* **99**, 031301 (2019).
- [23] L. Berthier, P. Charbonneau, and J. Kundu, “Finite dimensional vestige of spinodal criticality above the dynamical glass transition,” *Phys. Rev. Lett.* **125**, 108001 (2020).
- [24] P. Charbonneau and P. K. Morse, “Memory formation in jammed hard spheres,” *Phys. Rev. Lett.* **126**, 088001 (2021).
- [25] L. Berthier, G. Biroli, J. P. Bouchaud, M. Cipelletti, D. El Masri, D. L’Hôte, F. Ladieu, and M. Pierno, “Direct experimental evidence of a growing length scale accompanying the glass transition,” *Science* **310**, 1797 (2005).
- [26] W. Kob, C. Donati, S. J. Plimpton, P. H. Poole, and S. C. Glotzer, “Dynamical heterogeneities in a supercooled Lennard-Jones liquid,” *Phys. Rev. Lett.* **79**, 2827 (1997).
- [27] C. Donati, J. F. Douglas, W. Kob, S. J. Plimpton, P. H. Poole, and S. C. Glotzer, “Stringlike cooperative motion in a supercooled liquid,” *Phys. Rev. Lett.* **80**, 2338 (1998).
- [28] F. W. Starr, J. F. Douglas, and S. Sastry, “The relationship of dynamical heterogeneity to the Adam-Gibbs and random first-order transition theories of glass formation,” *J. Chem. Phys.* **138**, 12A541 (2013).
- [29] S. Karmakar, C. Dasgupta, and S. Sastry, “Growing length scales and their relation to timescales in glass-forming liquids,” *Ann. Rev. Cond. Matt. Phys.* **5**, 255 (2014).
- [30] P. Charbonneau, A. Ikeda, G. Parisi, and F. Zamponi, “Dimensional study of the caging order parameter at the glass transition,” *Proc. Nat. Acad. Sci.* **109**, 13839 (2012).
- [31] B. Charbonneau, P. Charbonneau, Y. Jin, G. Parisi, and F. Zamponi, “Dimensional dependence of the Stokes-Einstein relation and its violation,” *J. Chem. Phys.* **139**, 164502 (2013).
- [32] M. Adhikari, S. Karmakar, and S. Sastry, “Spatial dimensionality dependence of heterogeneity, breakdown of the Stokes-Einstein relation, and fragility of a model glass-forming liquid,” *J. Phys. Chem. B* **125**, 10232 (2021).
- [33] E. R. Weeks, J. C. Crocker, A. C. Levitt, A. Schofield, and D. A. Weitz, “Three-dimensional direct imaging of structural relaxation near the colloidal glass transition,” *Science* **287**, 627 (2000).
- [34] P. Chaudhuri, L. Berthier, and W. Kob, “Universal nature of particle displacements close to glass and jamming transitions,” *Phys. Rev. Lett.* **99**, 060604 (2007).
- [35] P. Charbonneau, J. Kurchan, G. Parisi, P. Urbani, and F. Zamponi, “Glass and jamming transitions: From exact results to finite-dimensional descriptions,” *Ann. Rev. Cond. Matt. Phys.* **8**, 265 (2017).
- [36] G. Parisi and F. Zamponi, “Mean-field theory of hard sphere glasses and jamming,” *Rev. Mod. Phys.* **82**, 789 (2010).
- [37] P. G. Debenedetti and F. H. Stillinger, “Supercooled liquids and the glass transition,” *Nature* **410**, 259 (2001).
- [38] A. J. Liu and S. R. Nagel, “The jamming transition and the marginally jammed solid,” *Ann. Rev. Cond. Matt. Phys.* **1**, 347 (2010).
- [39] L. Berthier and G. Biroli, “Theoretical perspective on the glass transition and amorphous materials,” *Rev. Mod. Phys.* **83**, 587 (2011).
- [40] T. S. Grigera and G. Parisi, “Fast Monte Carlo algorithm for supercooled soft spheres,” *Phys. Rev. E* **63**, 045102 (2001).
- [41] E. Bitzek, P. Koskinen, F. Gähler and M. Moseler, and P. Gumbsch, “Structural relaxation made simple,” *Phys. Rev. Lett.* **97**, 170201 (2006).
- [42] J. Guérolé, W. G. Nöhling, A. Vaid, F. Houllé, Z. Xie, A. Prakash, and E. Bitzek, “Assessment and optimization of the fast inertial relaxation engine (FIRE) for energy minimization in atomistic simulations and its implementation in LAMMPS,” *Comp. Mat. Sci.* **175**, 109584 (2020).
- [43] R. N. Chacko, F. P. Landes, G. Biroli, O. Dauchot, A. J. Liu, and D. R. Reichman, “Elastoplasticity mediates dynamical heterogeneity below the mode coupling temperature,” *Phys. Rev. Lett.* **127**, 048002 (2021).
- [44] L. Berthier and G. Tarjus, “Nonperturbative effect of attractive forces in viscous liquids,” *Phys. Rev. Lett.* **103**, 170601 (2009).
- [45] S. Toxvaerd, “Role of the attractive forces in a supercooled liquid,” *Phys. Rev. E* **103**, 022611 (2021).
- [46] S. J. Plimpton and A. P. Thompson, “Computational aspects of many-body potentials,” *MRS Bull.* **37**, 513 (2012).
- [47] The more common σ_i -independent ϵ_i [63] can be implemented by editing a single line in `initRMSE()`.
- [48] S. R. Williams, I. K. Snook, and W. van Meegen, “Molecular dynamics study of the stability of the hard sphere glass,” *Phys. Rev. E* **64**, 021506 (2001).
- [49] https://en.wikipedia.org/wiki/Leapfrog_integration.
- [50] Z. Yao, J.-S. Wang, G.-R. Liu, and M. Cheng, “Improved neighbor list algorithm in molecular simulations using cell decomposition and data sorting method,” *Comp. Phys. Comm* **161**, 27 (2004).
- [51] U. Welling and G. Germano, “Efficiency of linked cell algorithms,” *Comp. Phys. Comm.* **182**, 611 (2011).
- [52] S. Torquato, T. M. Truskett, and P. G. Debenedetti, “Is random close packing of spheres well defined?” *Phys. Rev. Lett.* **84**, 2064 (2000).
- [53] L. Berthier, D. Coslovich, A. Ninarello, and M. Ozawa, “Equilibrium sampling of hard spheres up to the jamming density and beyond,” *Phys. Rev. Lett.* **116**, 238002 (2016).
- [54] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak, “Molecular dynamics with coupling to an external bath,” *J. Chem. Phys.* **81**, 3684 (1984).
- [55] P. Charbonneau, Y. Jin, G. Parisi, and F. Zamponi, “Hopping and the Stokes-Einstein relation breakdown in simple glass formers,” *Proc. Natl. Acad. Sci. U.S.A.* **111**, 15025 (2014).
- [56] P. Charbonneau, J. Kurchan, G. Parisi, P. Urbani, and F. Zamponi, “Fractal free energy landscapes in structural glasses,” *Nat. Comm.* **5**, 3725 (2014).

- [57] L. Lue and M. Bishop, “Molecular dynamics study of the thermodynamics and transport coefficients of hard hyperspheres in six and seven dimensions,” *Phys. Rev. E* **74**, 021201 (2006).
- [58] J. A. van Meel, D. Frenkel, and P. Charbonneau, “Geometrical frustration: A study of four-dimensional hard spheres,” *Phys. Rev. E* **79**, 030201(R) (2009).
- [59] M. Schmiedeberg, T. K. Haxton, S. R. Nagel, and A. J. Liu, “Mapping the glassy dynamics of soft spheres onto hard-sphere behavior,” *Europhys. Lett.* **96**, 36010 (2011).
- [60] S. Meloni, M. Rosati, and L. Colombo, “Efficient particle labeling in atomistic simulations,” *J. Chem. Phys.* **126**, 121102 (2007).
- [61] L. Berthier, G. Biroli, D. Coslovich, W. Kob, and C. Tonelli, “Finite-size effects in the dynamics of glass-forming liquids,” *Phys. Rev. E* **86**, 031502 (2012).
- [62] J. D. Eaves and D. R. Reichman, “Spatial dimension and the dynamics of supercooled liquids,” *Proc. Nat. Acad. Sci.* **106**, 15171 (2009).
- [63] W. Kob and H. C. Andersen, “Testing mode-coupling theory for a supercooled binary Lennard-Jones mixture. i. the van Hove correlation function,” *Phys. Rev. E* **51**, 4626–4641 (1995).
- [64] Y. Ding and A. P. Sokolov, “Breakdown of time-temperature superposition principle and universality of chain dynamics in polymers,” *Macromolecules* **39**, 3322 (2006).
- [65] E. J. Saltzman and K. S. Schweizer, “Non-Gaussian effects, space-time decoupling, and mobility bifurcation in glassy hard-sphere fluids and suspensions,” *Phys. Rev. E* **74**, 061501 (2006).
- [66] L. Wang, N. Xu, W. H. Wang, and P. Guan, “Revealing the link between structural relaxation and dynamic heterogeneity in glass-forming liquids,” *Phys. Rev. Lett.* **120**, 125502 (2018).
- [67] This value $x \simeq 0.3$ is not universal; for example, Ref. [68] found $x \simeq 0.4$.
- [68] U. K. Nandi, W. Kob, and S. M. Bhattacharya, “Connecting real glasses to mean-field models,” *J. Chem. Phys.* **154**, 094506 (2021).
- [69] T. Abete, A. de Candia, E. Del Gado, A. Fierro, and A. Coniglio, “Dynamical heterogeneity in a model for permanent gels: Different behavior of dynamical susceptibilities,” *Phys. Rev. E* **78**, 041404 (2008).
- [70] E. Rössler, “Indications for a change of diffusion mechanism in supercooled liquids,” *Phys. Rev. Lett.* **65**, 1595 (1990).
- [71] R. K. Murarka and B. Bagchi, “Diffusion and viscosity in a supercooled polydisperse system,” *Phys. Rev. E* **67**, 051504 (2003).
- [72] S. K. Kumar, G. Szamel, and J. F. Douglas, “Nature of the breakdown in the Stokes-Einstein relationship in a hard sphere fluid,” *J. Chem. Phys.* **124**, 214501 (2006).
- [73] G. Biroli and J. P. Bouchaud, “Critical fluctuations and breakdown of the Stokes-Einstein relation in the mode-coupling theory of glasses,” *J. Phys. Cond. Matt.* **19**, 205101 (2007).
- [74] S. Kim, D. G. Thorpe, C. Noh, J. P. Garrahan, D. Chandler, and Y.-J. Jung, “Study of the upper-critical dimension of the East model through the breakdown of the Stokes-Einstein relation,” *J. Chem. Phys.* **147**, 084504 (2017).
- [75] Indeed, Ref. [74] implied that this effect may have been the origin of Ref. [31]’s conclusion that $\omega \rightarrow 0$ as $d \rightarrow 8$.
- [76] E. Barkai and S. Burov, “Packets of diffusing particles exhibit universal exponential tails,” *Phys. Rev. Lett.* **124**, 060603 (2020).
- [77] A. V. Chechkin, F. Seno, R. Metzler, and I. M. Sokolov, “Brownian yet non-Gaussian diffusion: From superstatistics to subordination of diffusing diffusivities,” *Phys. Rev. X* **7** (2017).
- [78] J. M. Miotto, S. Pigolotti, A. V. Chechkin, and S. Roldán-Vargas, “Length scales in Brownian yet non-Gaussian dynamics,” *Phys. Rev. E* **11**, 031002 (2021).
- [79] B. Wang, S. M. Anthony, S. C. Bae, and S. Granick, “Anomalous yet Brownian,” *Proc. Nat. Acad. Sci.* **106**, 15160 (2009).
- [80] B. Wang, J. Kuo, S. C. Bae, and S. Granick, “When Brownian diffusion is not Gaussian,” *Nat. Mat.* **11**, 481 (2012).
- [81] J. Guan, B. Wang, and S. Granick, “Even hard-sphere colloidal suspensions display Fickian yet non-Gaussian diffusion,” *ACS Nano* **8**, 3331 (2014).
- [82] T. R. Kirkpatrick and P. G. Wolynes, “Connections between some kinetic and equilibrium theories of the glass transition,” *Phys. Rev. A* **35**, 3072 (1987).
- [83] E. Flenner and G. Szamel, “Relaxation in a glassy binary mixture: Comparison of the mode-coupling theory to a Brownian dynamics simulation,” *Phys. Rev. E* **72**, 031508 (2005).
- [84] B. Schmid and R. Schilling, “Glass transition of hard spheres in high dimensions,” *Phys. Rev. E* **81**, 041502 (2010).
- [85] S. M. Bhattacharyya, B. Bagchi, and P. Wolynes, “Sub-quadratic wavenumber dependence of the structural relaxation of supercooled liquid in the crossover regime,” *J. Chem. Phys.* **132**, 104503 (2010).
- [86] E. J. Saltzman and K. S. Schweizer, “Large-amplitude jumps and non-Gaussian dynamics in highly concentrated hard sphere fluids,” *Phys. Rev. E* **77**, 051504 (2006).
- [87] L. Berthier, D. Chandler, and J. P. Garrahan, “Length scale for the onset of Fickian diffusion in supercooled liquids,” *Europhys. Lett.* **69**, 320 (2005).
- [88] C. E. Leiserson, N. C. Thompson, J. S. Emer, B. C. Kuzmaul, B. W. Lampton, D. Sanchez, and T. B. Schardl, “There’s plenty of room at the top: What will drive computer performance after Moore’s law?” *Science* **368**, 1079 (2020).
- [89] J. A. Anderson, J. Glaser, and S. C. Glotzer, “HOOMD-blue: A Python package for high-performance molecular dynamics and hard particle Monte Carlo simulations,” *Comp. Mat. Sci.* **173**, 109363 (2020).
- [90] J. Glaser, T. D. Nguyen, J. A. Anderson, P. Lui, F. Spiga, J. A. Millian, D. C. Morse, and S. C. Glotzer, “Strong scaling of general-purpose molecular dynamics simulations on GPUs,” *Comp. Phys. Comm.*
- [91] L. Lue, M. Bishop, and P. A. Whitlock, “The fluid to solid phase transition of hard hyperspheres in four and five dimensions,” *J. Chem. Phys.* **132**, 104509 (2010).
- [92] C. D. Estrada and M. Robles, “Fluid-solid transition in hard hypersphere systems,” *J. Chem. Phys.* **134**, 044115 (2011).
- [93] L. Lue, M. Bishop, and P. A. Whitlock, “Molecular dynamics study of six-dimensional hard hypersphere crystals,” *J. Chem. Phys.* **155**, 144502 (2021).