Bypassing sluggishness: SWAP algorithm and glassiness in high dimensions

Ludovic Berthier,1 Patrick Charbonneau,2,3 and Joyjit Kundu2
1Laboratoire Charles Coulomb (L2C), University of Montpellier, CNRS, Montpellier, France
2Department of Chemistry, Duke University, Durham, North Carolina 27708, USA
3Department of Physics, Duke University, Durham, North Carolina 27708, USA
(Dated: October 17, 2018)

The recent implementation of a swap Monte Carlo algorithm (SWAP) for polydisperse mixtures fully bypasses computational sluggishness and closes the gap between experimental and simulation timescales in physical dimensions $d = 2$ and 3. Here, we consider suitably optimized systems in $d = 2, 3, \ldots, 8$, to obtain insights into the performance and underlying physics of SWAP. We show that the speedup obtained decays rapidly with increasing the dimension. SWAP nonetheless delays systematically the onset of the activated dynamics by an amount that remains finite in the limit $d \to \infty$. This shows that the glassy dynamics in high dimensions $d > 3$ is now computationally accessible using SWAP, thus opening the door for the systematic consideration of finite-dimensional deviations from the mean-field description.

Introduction – A glass emerges when a supercooled liquid passed its crystallization point becomes so sluggish that it falls out of equilibrium. Upon cooling or increasing packing fraction, the dynamics of glass formers exhibits a marked slowdown beyond the dynamical onset, thus making this outcome inescapable [1, 2]. In mean-field descriptions, the structural relaxation time exhibits a power-law divergence at the dynamical transition [3]. In any finite dimension, although activated processes wash out this transition, the rapid growth of the associated relaxation time nonetheless impedes equilibration of low-temperature or high-density liquids. Standard simulation protocols, in particular, do not easily explore the regime beyond the dynamic transition, because structural relaxation is already too sluggish.

The application of the swap Monte Carlo algorithm (SWAP), which exchanges the identity of pairs of particles, to complex mixtures sidesteps this difficulty [4–6]. By considering systems with, for instance, a continuous size polydispersity one can follow the equilibrium liquid up to unprecedented high packing fractions or low temperatures. Tuning the range and functional form of polydispersity provides systems for which the sampling efficiency of swap moves is maximal within the liquid state, while remaining robust against crystallization and fractionation [7]. For properly chosen polydispersity in $d = 3$ this procedure has recently provided a speedup of at least $10^{10}$ compared to standard dynamics, matching the experimental timescales [7, 8], and in $d = 2$ it has given access to timescales that are truly cosmological [9]. This computational progress has triggered the exploration of new glass physics in computer simulations, notably low-temperature anomalies [10, 11], the Gardner transition [12], the rheology of glasses [13], the extension of the jamming line [14], and the ultrastability of vapor-deposited glasses [15].

The efficiency of SWAP has also triggered theoretical activity aimed at better understanding its physical origin and its physical implications for the glass transition [10]. Ikeda et al. [17] present a replica calculation of a mean-field glass model proposing that SWAP and physical dynamics are ruled by distinct dynamical transitions. A qualitatively similar result is obtained by Szamel who obtains two dynamical transitions for the two dynamics [18]. Brito et al. [19] obtain a similar result, and interpret the dynamical transition as an onset of mechanical rigidity that is again shifted by SWAP. Finally, Berthier et al. [20] argue that the onset of thermal activation past the dynamical transition is also considerably affected by SWAP. There is thus a general consensus that SWAP can delay the dynamical transition by an amount that is system dependent, and can speedup the dynamics even past the avoided dynamical transition.

However, because dynamical transitions are avoided in any finite $d$ [21], other physical processes might also explain the dramatic change in dynamics. In particular, structural imperfections closely tied to local geometry [22], which are putatively important in the dynamics of low-dimensional glass formers, could impact SWAP efficiency. Distinguishing one contribution from the other can be achieved by considering how SWAP performance evolves with increasing $d$. A non-vanishing SWAP efficiency in the limit of $d \to \infty$ or a perturbative correction in $1/d$ would suggest that the mean-field dynamical transition is indeed shifted, while an exponential suppression would suggest that nonperturbative features associated with geometry dominate. Because numerical work on SWAP has thus far only been concerned with physical dimensions, $d = 2$ and 3, distinguishing between these scenarios is not currently possible.

Resolving this question would not only shed light on the physical origin of the glassy slowdown, but help devise novel algorithms that further bypass it. Interestingly, side-stepping the mean-field dynamical threshold could also be key to general algorithmic improvements in hard problems, such as statistical inference, high-dimensional optimization and deep learning [23]. A fundamental grasp of the effectiveness of SWAP dynamics
could thus bolster advances far beyond the problem at hand. More immediately, if one could generically push the current limitations of high \(d\) simulations, crucial questions in glass physics could be tackled \[21, 24, 25\]. In this work, we study the dynamics of suitably optimized polydisperse mixtures of hard spheres in various spatial dimensions, so as to systematically approach the mean-field, \(d \to \infty\) description, and provide microscopic insight into the underlying physics and computational efficiency across a broad range of dimensions.

**Simulation Model**—We consider size polydisperse systems with \(N\) hard spheres in a hypercubic box of constant volume \(V\), under periodic boundary conditions in \(d = 2, 3, \ldots, 8\). The size distribution function has the form, \(P(\sigma) = K/\sigma^3\), with normalization constant \(K\) for \(\sigma \in [\sigma_{\text{min}}, \sigma_{\text{max}}]\), where \(\sigma_{\text{min}}\) and \(\sigma_{\text{max}}\) are the minimum and the maximum diameter values, respectively. The average diameter \(\bar{\sigma} = \int_{\sigma_{\text{min}}}^{\sigma_{\text{max}}} P(\sigma) \sigma d\sigma\) sets the unit of length, and the standard deviation of the size distribution, \(\Delta\), quantifies the degree of polydispersity (see Simulation details and model parameters in [26]). For a fixed \(\Delta\), this specific choice of size distribution function does not significantly affect the system dynamics. Figure 1 which explicitly compares the dynamics at fixed \(\Delta\) and various \(P(\sigma)\) in \(d = 4\), confirms that \(\Delta\) is the most relevant variable. Our analysis should therefore be reasonably independent of the specifics of the model studied.

Standard and SWAP simulations are run for different \(\Delta\) and \(d\). Both dynamical protocols include basic single-particle translational moves along a vector randomly drawn within a \(d\)-dimensional hypercube of side \(\delta t\); SWAP includes additional diameter exchanges between two randomly chosen particles, attempted with probability \(p = 0.2\) (setting \(p = 0\) recovers standard dynamics). While \(0 < p \lesssim 0.2\) monotonically increases sampling efficiency, for \(p \gtrsim 0.2\) efficiency saturates, and hence additional swap moves wastefully slow down simulations \[4\]. For each volume fraction \(\varphi\), the pressure \(P\) is measured using pair correlations \[1, 27\], to compute the unitless reduced pressure, \(Z = \beta P/\rho\), for the number density \(\rho = \varphi/V_0\) with \(V_0\) being the average volume of a \(d\)-dimensional hypersphere.

Equilibration is assessed by the complete decay of the self-part of the particle-scale overlap function

\[
Q(t) = \frac{1}{N} \sum_{i=1}^{N} \Theta(a - |\mathbf{r}_i(t) - \mathbf{r}_i(0)|),
\]

where \(\Theta\) is a step function and \(a = 0.3\bar{\sigma}\) is a microscopic length chosen to be close to the typical particle cage size. The associated structural relaxation time, \(\tau_\alpha\), is defined such that \(Q(\tau_\alpha) = e^{-1}\). We define the relaxation time for both the standard (\(\tau_\alpha^{\text{std}}\)) and SWAP (\(\tau_\alpha^{\text{swap}}\)) dynamics. In all dimensions studied, SWAP equilibrates systems far beyond what is computationally accessible with standard Monte Carlo, and we thus first equilibrate systems using SWAP before measuring properties of the dynamics without it.

**Results**—In physical dimensions, crystallization competes with equilibration of deeply supercooled liquids \[28\]. For instance, for \(\Delta \lesssim 8\%\) in \(d = 3\) crystallization at high \(\varphi\) is unavoidable. For \(d > 3\), by contrast, crystallization does not interfere with the metastable fluid phase even for arbitrarily low \(\Delta\). The nucleation time at finite \(\Delta\) in \(d > 3\) is thus as equally out of computational reach as it is for monodisperse systems (\(\Delta = 0\)) \[30–32\]. In all \(d\), however, size fractionation may take place at high \(\Delta\) and \(\varphi\). In \(d = 3\), fractionation appears at \(\Delta \gtrsim 10\%\), which helps crystallization \[33, 34\]. In practice, this only happens when SWAP is used \[29\], because composition fluctuations leading to fractionation are then much faster. SWAP thus not only accelerates the sampling of the metastable fluid, but also changes the glass-forming ability of the system and forces the use of \(\Delta > 20\%\) in \(d = 3\). In \(d = 4\), by contrast, fractionation only appears at \(\Delta \gtrsim 15\%\) for \(\varphi \gtrsim 0.43\), and is further suppressed at higher \(\Delta\) (see Dynamic and static observables in [26]). For each \(d\), a \(\Delta\) window, within which SWAP efficiency is reasonably good and fractionation (with or without crystallization) does not interfere, can thus be found. Qualitative and even quantitative aspects of the standard Monte Carlo dynamics are otherwise not remarkably affected by changing \(\Delta\), as expected from previous studies of naturally polydisperse systems, such as colloidal suspensions \[35\].

A strong dependence of the SWAP dynamics on \(\Delta\) is observed in the dynamically sluggish regime, beyond the onset of slow diffusion at \(\varphi_0\) (Fig. 2(a)-(d)). As an illustration, we consider the evolution of the SWAP efficiency ratio, \(\tau_\alpha^{\text{std}}/\tau_\alpha^{\text{swap}}\) measured at a fixed \(\tau_\alpha/\tau_0\),
The diffusion of particle diameters becomes slower than the structural relaxation when $\Delta$ is large, because diameter and position dynamics are intimately coupled. Increasing $\Delta$ thus no longer improves SWAP efficiency, and this saturation develops earlier in larger $d$, where the vibrational dynamics (or, loosely speaking, caging) itself occurs over a length-scale decreasing with $d$.

The most remarkable feature of the efficiency results is the weakening of SWAP performance with increasing $d$. Fig. 3(b) shows that the efficiency decays rapidly with increasing $d$ (nearly exponentially, at least up to $d = 8$) for various $\tau_0^{\text{std}}/\tau_0$. The decay of SWAP performance becomes more prominent when estimated beyond the accessible regimes of the standard dynamics, such as where $\tau_0^{\text{std}}/\tau_0 = 5 \times 10^3$—see Fig. 3(c) and (d) and Static observables in [26]).

In order to examine explicitly whether this strong suppression is due to non-perturbative effects of not, we consider how SWAP impacts the avoided mean-field dynamical transition, $\varphi_d$. We estimate $\varphi_d$ for both standard dynamics and SWAP by fitting the growth of the relaxation time to the critical scaling form, $\tau_0 \propto (\varphi_d - \varphi)^{-\gamma}$ [21] (see Mode coupling analysis in [26]). As expected [36], this scaling form captures the data increasingly well as $d$ increases. In $d = 2$, it does not have a good regime of validity, but its validity eventually reaches up to three decades in the computationally accessible regime. We find that $\gamma$ is fairly insensitive to both dimension [21] and polydispersity [37]. Three features of the results are particularly noteworthy. First, collapsing $\tau_0^{\text{std}}/\tau_0$ by rescaling $\varphi/\varphi_d$ clearly reveals that SWAP postpones the putative dynamical transition in all dimensions—Fig. 4(a). Second, while $\varphi_d$ monotonically grows with polydispersity [58], its relative impact, $(\varphi_d^{\text{swap}} - \varphi_d^{\text{std}})/\varphi_d^{\text{std}}$ eventually plateaus on a scale consistent with the estimates for $\Delta_0$—see Fig. 4(b). This suggests that the shift of dynamical transition is directly correlated with the SWAP efficiency, as both quantities evolve similarly with $\Delta$. Third, the plateau height, $A \equiv (\varphi_d^{\text{swap}} - \varphi_d^{\text{std}})/\varphi_d^{\text{std}}$ at the maximum polydispersity considered in Fig. 4(b), decays to a nonzero value ($\approx 0.037$) with correction that scales with dimension as $\sim 1/d$. Our results thus suggest that the gain in SWAP efficiency survives in the limit $d \to \infty$, and that perturbative corrections survive all the way down to $d = 3$, independently of non-perturbative effects.

How can one explain the relatively rapid suppression of swap efficiency despite of the slow decay of the density gap $(\varphi_d^{\text{swap}} - \varphi_d^{\text{std}})/\varphi_d^{\text{std}}$ to a nonzero value? While the relative increase of $\varphi_d$ is qualitatively consistent with mean-field treatments in $d = 3$ [17, 18], the saturation and the asymptotic behavior of the gap with $d$ were not anticipated. Plugging this result into the critical scaling forms $\tau_0 = A(\varphi_d - \varphi)^{-\gamma}$, we obtain an approximate expression for the efficiency ratio

$$\tau_0^{\text{std}}/\tau_0^{\text{swap}} \approx \tau_0^{\text{std}}(h\varphi_d^{\text{std}})^{-\gamma}$$

(2)

For a given value of $\tau_0^{\text{std}}$, the key contribution to the

\[
\begin{array}{c}
\text{FIG. 2. SWAP efficiency, } \tau_0^{\text{std}}/\tau_0^{\text{swap}}, \text{ as a function of the relaxation time of the standard dynamics (representing the sluggishness) for different polydispersities } \Delta \text{ in (a) } d = 3, \text{ (b) } d = 4, \text{ (c) } d = 5, \text{ and (d) } d = 6. \text{ Sluggish dynamics at low } \Delta \text{ cannot be reached in } d = 3 \text{ because crystallization interferes. In all } d, \text{ SWAP performs better as } \Delta \text{ increases, and saturates at larger } \Delta. \\
\end{array}
\]
efficiency gain therefore arises from the term $(\varphi_{\text{std}})\gamma$. Because asymptotically $\varphi_{\text{std}} \sim d^{-2} [21]$, this gain decreases rapidly with increasing $d$—qualitatively consistent with Figs. 3 (b, c) and Fig. 4 (b). Because $\tau_\alpha$ diverges upon approaching $\varphi_d$ in high dimension, however, one should always be able to identify sufficiently sluggish systems for SWAP to speed up sampling. In intermediate dimensions, the approach remains sufficiently productive to obtain equilibrium configurations much beyond the dynamical transition of the standard dynamics. Figure 3 (c) provides a rough estimates of how useful SWAP can be in accessing regimes that are not accessible by the standard dynamics in high dimensions. For instance, in $d = 8$ a speed up of roughly $10^4$ should remain computationally achievable.

**Conclusion** – We have shown that SWAP improves sampling in dimensions $d \geq 2$ by generically delaying the dynamical transition that indicates the emergence of activated dynamics in the standard dynamics. This finding in itself does not directly reveal the microscopic nature (dynamic or thermodynamic) of the standard dynamics in the regime $\varphi_{\text{std}} < \varphi < \varphi_{\text{swap}}$, where SWAP provides most of its dynamic speedup, but offers a platform for assessing this question in the future. Because the gap between the dynamical transition of the standard and the SWAP dynamics remains finite in the limit $d \to \infty$, SWAP can efficiently be used to study pure glass physics in reasonably large dimensions, far from the regime in which significant local structure [22] or orientational ordering [39] might interfere. In other words, although caging imperfections go away exponentially quickly with increasing dimension, SWAP can still break cages in high $d$. Even within this analysis, the two-dimensional speedup is remarkably large, and techniques specifically tailored to identify local structual weaknesses, (e.g., [20, 13]) might thus help obtain additional microscopic insights. More generally, our observations suggest that the standard dynamical transition might not be as strong an algorithmic constraint as previously conceived in problems ranging from physics to information theory. If a proper sampling scheme can be devised and exploited in those problems, some stunning algorithmic advances might thus be within reach.

We thank S. Yaida, M. Ozawa and F. Zamponi for useful discussions. J. K., L. B. and P. C. acknowledge support from the Simons Foundation grant (#454933, Ludovic Berthier, # 454937, Patrick Charbonneau). Most simulations were performed at Duke Compute Cluster (DCC). J.K. thanks Tom Milledge for helping with the
usage of DCC. P.C. and J.K. also thanks Extreme Science and Engineering Discovery Environment (XSEDE), supported by National Science Foundation grant number ACI-1548562, for computer time.

[1] L. Berthier, G. Biroli, J.-P. Bouchaud, L. Cipelletti, and W. van Saarloos, Dynamical Heterogeneities in Glasses, Colloids, and Granular Media (Oxford University Press, 2011).
[2] J. C. Dyre, Rev. Mod. Phys. 78, 953 (2006).
[3] L. Berthier and G. Biroli, Rev. Mod. Phys. 83, 587 (2011).
[4] L. Berthier, D. Coslovich, A. Ninarello, and M. Ozawa, Phys. Rev. Lett. 116, 238002 (2016).
[5] R. Gutiérrez, S. Karmakar, Y. G. Pollack, and I. Procaccia, Europhys. Lett. 111, 56009 (2015).
[6] T. S. Grigera and G. Parisi, Phys. Rev. E(R) 63, 045102 (2001).
[7] A. Ninarello, L. Berthier, and D. Coslovich, Phys. Rev. X 7, 021039 (2017).
[8] L. Berthier, P. Charbonneau, A. Ninarello, M. Ozawa, and S. Yaida, Proc. Natl. Acad. Sci. 114, 11356 (2017).
[9] L. Berthier, P. Charbonneau, A. Ninarello, M. Ozawa, and S. Yaida, arXiv preprint arXiv:1805.09035 (2018).
[10] C. Scalliet, L. Berthier, and F. Zamponi, Phys. Rev. Lett. 119, 205501 (2017).
[11] L. Wang, A. Ninarello, P. Guan, L. Berthier, G. Szamel, and E.flenner, arXiv:1804.08765 (2018).
[12] L. Berthier, P. Charbonneau, Y. Jin, G. Parisi, B. Secane, and F. Zamponi, Proc. Natl. Acad. Sci. 113, 8397 (2016).
[13] M. Ozawa, L. Berthier, G. Biroli, A. Rosso, and G. Tarjus, arXiv preprint arXiv:1803.11502 (2018).
[14] M. Ozawa, L. Berthier, and D. Coslovich, SciPost Phys. 3, 027 (2017).
[15] L. Berthier, P. Charbonneau, E. Flener, and F. Zamponi, Phys. Rev. Lett. 119, 188002 (2017).
[16] M. Wyart and M. E. Cates, Phys. Rev. Lett. 119, 195501 (2017).
[17] H. Ikeda, F. Zamponi, and A. Ikeda, J. Chem. Phys. 147, 234506 (2017).
[18] G. Szamel, arXiv preprint arXiv:1805.02763 (2018).
[19] C. Brito, E. Lerner, and M. Wyart, arXiv preprint arXiv:1801.03796 (2017).
[20] L. Berthier, G. Biroli, J.-P. Bouchaud, and G. Tarjus, arXiv:1805.12378 (2018).
[21] P. Charbonneau, J. Kurchan, G. Parisi, P. Urbani, and F. Zamponi, Annu. Rev. Condens. Matter Phys. 8, 265 (2017).
[22] C. P. Royall, F. Turci, S. Tatsumi, J. Russo, and J. Robinson, J. Phys.: Condens. Matter 30, 363001 (2018).
[23] F. K. L. Zdeborová, Adv. Phys. 65, 453 (2016).
[24] J. D. Eaves and D. R. Reichman, Proc. Natl. Acad. Sci. 106, 15171 (2009).
[25] S. Sengupta, S. Karmakar, C. Dasgupta, and S. Sastry, J. Chem. Phys. 138, 12A548 (2013).
[26] See Supplemental Material for detailed discussions on simulation details and model parameters, dynamic and static observables, and mode coupling analysis.
[27] A. Santos, S. B. Yuste, and M. L. de Haro, J. Chem. Phys. 117, 5785 (2002).
[28] C. Valeriani, E. Sanz, E. Zaccarelli, W. C. K. Poon, M. E. Cates, and P. N. Pusey, J. Phys.: Condens. Matter 23, 194117 (2011).
[29] M. Skoge, A. Donev, F. H. Stillinger, and S. Torquato, Phys. Rev. E 74, 041127 (2006).
[30] J. A. van Meel, B. Charbonneau, A. Fortini, and P. Charbonneau, Phys. Rev. E 80, 061110 (2009).
[31] P. Charbonneau, A. Ikeda, J. A. van Meel, and K. Miyazaki, Phys. Rev. E 81, 040501(R) (2010).
[32] B. A. Lindquist, R. B. Jadrich, and T. M. Truskett, J. Chem. Phys. 148, 191101 (2018).
[33] D. Coslovich, M. Ozawa, and L. Berthier, J. Phys.: Condens. Matter 30, 144004 (2018).
[34] G. L. Hunter and E. R. Weeks, Rep. Prog. Phys. 75, 066501 (2012).
[35] P. Charbonneau, Y. Jin, G. Parisi, and F. Zamponi, Proc. Natl. Acad. Sci. 111, 15025 (2014).
[36] F. Weysser, A. M. Puertas, M. Fuchs, and T. Voigtmann, Phys. Rev. E 82, 011504 (2010).
[37] I. Biazzo, F. Caltagirone, G. Parisi, and F. Zamponi, Phys. Rev. Lett. 102, 195701 (2009).
[38] H. Tong and H. Tanaka, Phys. Rev. X 8, 011041 (2018).
[39] F. Turci, C. P. Royall, and T. Speck, Phys. Rev. X 7, 031028 (2017).
[40] J. E. Hallett, F. Turci, and C. P. Royall, Nat. Comm. 9, 3272 (2018).
[41] S. Schoenholz, E. D. Cubuk, D. M. Sussman, E. Kaxiras, and A. J. Liu, Nat. Phys. 12, 469 (2016).
[42] S. Schoenholz, E. D. Cubuk, E. Kaxiras, and A. J. Liu, Proc. Natl. Acad. Sci. 114, 263 (2017).
Supplemental Information

for

“Bypassing sluggishness: SWAP algorithm and glassiness in high dimensions”

Ludovic Berthier¹, Patrick Charbonneau²,³ and Joyjit Kundu²

¹Laboratoire Charles Coulomb (L2C), University of Montpellier, CNRS, Montpellier, France
²Department of Chemistry, Duke University, Durham, North Carolina 27708, USA
³Department of Physics, Duke University, Durham, North Carolina 27708, USA

SIMULATION DETAILS AND MODEL PARAMETERS

We simulate systems of hard spheres in $d = 2, \ldots, 8$ with continuous size dispersity (polydispersity) using Monte Carlo (MC) simulations with a constant number of particles ($N = 2000$ for $d \leq 6$ and $N = 7000$ in $d = 8$), and a constant volume under periodic boundary conditions. For a given size distribution $P(\sigma)$, the mean diameter $\langle \sigma \rangle$ defines the unit of length and the standard deviation of that distribution, $\Delta = \sqrt{\langle \sigma^2 \rangle - \langle \sigma \rangle^2}$, defines the degree of polydispersity.

We perform both the standard and swap MC (SWAP) dynamics. The standard MC protocol consists solely of translational displacement moves, uniformly drawn over a $d$ dimensional hypercube of side $\ell(d)$. The equilibrium fluid configurations deep inside the glassy regime are obtained using SWAP that involves both local displacements and non-local particle swaps, in which two randomly selected particles exchange their diameters. Particle swaps and displacements are attempted with probability 0.2 and 0.8, respectively, and are accepted if no overlap results. For a given $d$, the value of $\ell$ is chosen such that the relaxation time (in units of MC sweeps) for the standard dynamics is minimal near the dynamical transition (see Fig. S1(a)). A large value of $\ell$ leads to unsuccessful displacement attempts, and a small value is inefficient at sampling the particle cage. We find $\ell \approx 0.110, 0.060, 0.040, 0.033, 0.027, 0.021$ to be optimal in $d = 2, 3, 4, 5, 6, 8$, respectively. These values are robust against the degree of polydispersity. Swap moves attempt to exchange the diameters of two particles with diameter difference $< d\sigma_{\text{tol}}$, which roughly corresponds to the cage diameter. We also optimize the value of $d\sigma_{\text{tol}}$ in different dimensions-- a representative plot for $d = 6$ is shown in Fig. S1(b). We set $d\sigma_{\text{tol}} = 0.10, 0.09, 0.09, 0.04, 0.03, 0.012$ in $d = 2, 3, 4, 5, 6$ and 8 respectively. Please note that SWAP efficiency depends only weakly on $d\sigma_{\text{tol}}$ in the range $0.01 \leq d\sigma_{\text{tol}} \leq 0.10$.

THERMALIZATION: DYNAMIC AND STATIC OBSERVABLES

To ensure thermalization at each state point, system are evolved at least up to $20\tau_{\text{a, swap}}$ (measured from the decay of the overlap function– see Fig. S2), before starting the production run. In order to measure dynamical observables at relatively low packing fractions, production runs lasting at least $40\tau_0$ are used to average over time, while at high densities, 60 replicas are run up to a shorter time, typically $\sim 5\tau_0$, and the overall results are averaged. Typical correlation decays are given in Fig. S2. We consider the dynamics starting from the onset of glassiness, which is detected by the emergence of a inflection point in the mean squared displacement, thus implying non-Fickian diffusion (see Fig. S3). The corresponding $\tau_0$ is estimated by the relaxation time at the onset for both the standard and the swap dynamics. We obtain $\tau_0^{\text{std}} \approx 1400$ and $\tau_0^{\text{swap}} \approx 2400$ in $d = 2$; $\tau_0^{\text{std}} \approx 1800$ and $\tau_0^{\text{swap}} \approx 3000$ in $d = 3$; $\tau_0^{\text{std}} \approx 2700$ and $\tau_0^{\text{swap}} \approx 4200$ in $d = 4$; $\tau_0^{\text{std}} \approx 4000$ and $\tau_0^{\text{swap}} \approx 5600$ in $d = 5$; $\tau_0^{\text{std}} \approx 5000$ and $\tau_0^{\text{swap}} \approx 6500$ in $d = 6$; $\tau_0^{\text{swap}} \approx 6500$ and $\tau_0^{\text{swap}} \approx 8200$ in $d = 8$.

Various static observables, such as the structure factor and the pair correlation function, are used to detect putative (and unwanted) crystallization. An instance of fractionation is shown in Fig. S4. The pressure, $P$, is extracted from the contact value of the pair correlation function properly scaled for a polydisperse system, to calculate the equation of state, $Z(\phi) = P/\rho$, where $\phi = V_d\rho$ is the packing fraction for number density $\rho$ and average sphere volume $V_d$ [1].
FIG. S1. (a) The structural relaxation time, $\tau^\text{std}_\alpha$, for the standard dynamics is minimal at a finite $\ell$, as shown here for $d = 4$, $\Delta = 10\%$, and $\varphi = 0.4015$. (b) The relaxation time $\tau^\text{swap}_\alpha$ for SWAP dynamics is minimal for a finite $\delta\sigma_{\text{tol}}$, as shown here for $d = 6$, $\Delta = 10\%$, and $\varphi = 0.1865$.

Equations of states for different $\Delta$ and $d$ are shown in Fig. S5. In Fig. S6, we show SWAP efficiency as a function of polydispersity for different values of sluggishness, given by $\tau^\text{std}_\alpha$. To estimate the SWAP efficiency beyond the numerically accessible regimes, we use parabolic fitting forms extrapolating the data– see Fig. S7.
FIG. S3. Mean-squared displacement at different densities around the onset of glassiness, at which non-Fickian diffusion sets in. Results are shown here for the standard dynamics in (a) $d = 4$ ($\phi_0 \approx 0.355$), (b) $d = 5$ ($\phi_0 \approx 0.236$), and (c) $d = 6$ ($\phi_0 \approx 0.152$).

FIG. S4. Typical plots of the structure factor, $S(k)$ for (a) $\Delta = 15\%$, and (b) $\Delta = 20\%$ in $d = 4$ at various densities. For $\phi \gtrsim 0.43$, fractionation takes place, as indicated by a low $k$ peak.
FIG. S5. Equations of states in $d = 2, 3, 4, 5, 6,$ and 8.

FIG. S6. The plateau height, which is the maximum efficiency for a given $d$, increases monotonically with $\tau_{\alpha}^{\text{std}}/\tau_0$. Here we specifically consider $\tau_{\alpha}^{\text{std}}/\tau_0 = 5 \times 10^3$, $2 \times 10^3$, and $10^3$, but the dimensional trend is robust against this choice.
FIG. S7. The relaxation time $\tau_\alpha$ versus the reduced pressure $Z$ in different dimensions. The data for $\tau_\alpha$ are extrapolated using the parabolic form: $\tau = \tau_\infty \exp\left[ A(Z - Z_0)^2 \right]$. The vertical line corresponds to $\tau_\text{swap}/\tau_0 = 5 \times 10^3$.

FIG. S8. Critical scaling of the relaxation time $\tau_\alpha^{\text{std}}$ near the dynamical transition in (a) $d = 3$, (b) $d = 4$, (c) $d = 5$, (d) $d = 6$, and (e) $d = 8$.

**MODE-COUPLING ANALYSIS**

In the limit $d \to \infty$, there exists a dynamical critical point, $\varphi_d$, at which the relaxation time $\tau_\alpha$ diverges as $\sim (\varphi_d - \phi)^{-\gamma}$ and the system gets trapped in one of the many metastable minima. This dynamical criticality is avoided in finite dimensions because of competition from activated processes. One can nonetheless fit a critical form to $\tau_\alpha$ versus $\phi$ over a limited regime below $\varphi_d$ to estimate $\varphi_d$ and $\gamma$ (see Fig. S8 and S9). This power-law fit becomes more and more accurate with increasing $d$. Our estimates of $\varphi_d$, and $\gamma$ for both the standard and swap dynamics in different
FIG. S9. Critical scaling of the relaxation time $\tau_{\text{swap}}$ near the dynamical transition in (a) $d = 3$, (b) $d = 4$, (c) $d = 5$, (d) $d = 6$, and (e) $d = 8$.

### TABLE I. The dynamical transition density $\phi_{d}^\text{std}$ and the corresponding critical exponent $\gamma$ for different $\Delta$ in $d = 3, 4, 5, 6$ and 8 for the standard dynamics.

| $d = 3$ | $d = 4$ | $d = 5$ | $d = 6$ | $d = 8$ |
|---------|---------|---------|---------|---------|
| $\Delta$ | $\phi_{d}^\text{std}$ | $\gamma$ | $\Delta$ | $\phi_{d}^\text{std}$ | $\gamma$ | $\Delta$ | $\phi_{d}^\text{std}$ | $\gamma$ | $\Delta$ | $\phi_{d}^\text{std}$ | $\gamma$ |
| 12% | 0.6189 | 2.70 | 1% | 0.4068 | 2.72 | 1% | 0.2711 | 2.72 | 1% | 0.1741 | 2.67 | - | - |
| 17% | 0.6286 | 2.71 | 5% | 0.4200 | 2.64 | 5% | 0.2807 | 2.64 | 5% | 0.1813 | 2.62 | - | - |
| 23% | 0.6399 | 2.71 | 10% | 0.4331 | 2.59 | 10% | 0.2915 | 2.65 | 10% | 0.1899 | 2.62 | 10% | 0.07506 | 2.57 |
| - | - | - | - | - | - | - | - | - | - | - | - | - |

### TABLE II. The dynamical transition density $\phi_{d}^\text{swap}$ and the corresponding critical exponent $\gamma$ for different $\Delta$ in $d = 3, 4, 5, 6$ and 8 for the SWAP.

| $d = 3$ | $d = 4$ | $d = 5$ | $d = 6$ | $d = 8$ |
|---------|---------|---------|---------|---------|
| $\Delta$ | $\phi_{d}^\text{swap}$ | $\gamma$ | $\Delta$ | $\phi_{d}^\text{swap}$ | $\gamma$ | $\Delta$ | $\phi_{d}^\text{swap}$ | $\gamma$ | $\Delta$ | $\phi_{d}^\text{swap}$ | $\gamma$ |
| 12% | 0.6189 | 2.70 | 1% | 0.4068 | 2.72 | 1% | 0.2711 | 2.72 | 1% | 0.1741 | 2.67 | - | - |
| 17% | 0.6286 | 2.71 | 5% | 0.4200 | 2.64 | 5% | 0.2807 | 2.64 | 5% | 0.1813 | 2.62 | - | - |
| 23% | 0.6399 | 2.71 | 10% | 0.4331 | 2.59 | 10% | 0.2915 | 2.65 | 10% | 0.1899 | 2.62 | 10% | 0.07506 | 2.57 |
| - | - | - | - | - | - | - | - | - | - | - | - | - |

dimensions and for different values $\Delta$ are given in Table I and II.

[1] A. Santos and S. B. Yuste, J. Chem. Phys. **123**, 234512 (2005).