Supplementary Information: Are some packings more equal than others? A direct test of the Edwards conjecture

Stefano Martiniani,[1, *] K. Julian Schrenk,[1] Kabir Ramola,[2] Bulbul Chakraborty,[2] and Daan Frenkel[1]

[1] *Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge, CB2 1EW, UK*

[2] *Martin Fisher School of Physics, Brandeis University, Waltham, MA 02454, USA*

PACKING PREPARATION PROTOCOL

In this section we describe the algorithm that we have used to sample the phase space of jammed packings. This procedure samples each configuration proportional to the volume of its basin of attraction.

Hard sphere fluid sampling

We start by equilibrating a fluid of N hard disks (that serve as the cores of the particles with soft outer shells) at a volume fraction ϕ_{HS} in a square box with periodic boundary conditions. The particle radii are sampled from a truncated Gaussian distribution with mean $\mu = 1$ and standard deviation $\sigma = 0.1$. We achieve equilibration by performing standard Markov Chain Monte Carlo (MCMC) simulations consisting of single particle displacements and particle-particle swaps, as in [1]. To assure statistical independence, we draw fluid configurations every n_{MC} steps, where n_{MC} is (pre-computed by averaging over multiple simulations) the total number of MCMC steps necessary for each individual particle to diffuse at least a distance equal to the largest diameter in the system.

Soft shells and minimization

We next take each equilibrated hard disk fluid configuration and inflate the particles (instantaneously) with a WCA-like soft outer shell [2], to reach the target soft packing fraction $\phi_{\mathrm{SS}} > \phi_{\mathrm{HS}}$. Each hard sphere is inflated proportionally to its radius, so that the soft sphere radius is given by

$$r_{\mathrm{SS}} = \left(\frac{\phi_{\mathrm{SS}}}{\phi_{\mathrm{HS}}}\right)^{1/d} r_{\mathrm{HS}}, \tag{S1}$$

where d is the dimensionality of the box (2 in our case), and r_{SS} and r_{HS} are the soft and hard sphere radii respectively. Clearly, this procedure does not change the polydispersity of the sample. The radii are identical across volume fractions and system sizes, and the hard disk fluid density is chosen so that the radius ratio of hard to soft disks is $(0.88/0.7)^{1/2} \approx 1.121$.

Next, particle inflation is followed by energy minimization using FIRE [3, 4], to produce mechanically stable packings at the desired soft volume fractions ϕ. This protocol has the advantage of generating packings sampled proportionally to the volume of their basin of attraction. In our simulations, we considered all mechanically stable packings, irrespective of the number of 'rattlers'. To guarantee mechanical stability we required that the total number of contacts is sufficient for the bulk modulus to be strictly positive, $N_{\min} = d(N_{nr} - 1) + 1$ [5], where N_{nr} is the number of non-rattlers and d the dimensionality of the system.

Our implementation of FIRE enforces a maximum step size (set to be equal to the soft shell thickness) and forbids uphill steps by taking one step back every time the energy increases (and restarts the minimizer in the same fashion as the original FIRE implementation). We use a maximum time step $\Delta t_{\max} = 1$, although the maximum step size is directly controlled in our implementation. All other parameters are set as in the original implementation [3].

HS-WCA POTENTIAL

We define the WCA-like potential around a hard core as follows: consider two spherical particles with a distance between the hard cores r_{HS}, implying a soft core contact distance $r_{\mathrm{SS}} = r_{\mathrm{HS}}(1+\theta)$, with $\theta = (\phi_{\mathrm{SS}}/\phi_{\mathrm{HS}})^{1/d} - 1$. We can then write a horizontally shifted hard-sphere plus WCA (HS-WCA) potential as

$$u_{\text{HS-WCA}}(r) = \begin{cases} \infty & r \leq r_{\text{HS}}, \\ 4\epsilon \left[\left(\dfrac{\sigma(r_{\text{HS}})}{r^2 - r_{\text{HS}}^2} \right)^{12} \right. & \\ \left. - \left(\dfrac{\sigma(r_{\text{HS}})}{r^2 - r_{\text{HS}}^2} \right)^{6} \right] + \epsilon & r_{\text{HS}} < r < r_{\text{SS}}, \\ 0 & r \geq r_{\text{SS}} \end{cases} \tag{S2}$$

where $\sigma(r_{\text{HS}}) = (2\theta + \theta^2) r_{\text{HS}}^2 / 2^{1/6}$ guarantees that the potential function and its first derivative go to zero at r_{SS}. For computational convenience (avoidance of square-root evaluations), the potential in Eq. S2 differs from the WCA form in that the inter-particle distance in the denominator of the WCA potential has been replaced with a difference of squares.

A power series expansion of Eq. S2 yields

$$\lim_{r \to r_{\text{SS}}} u_{\text{HS-WCA}} = \epsilon \left(\frac{12 r_{\text{SS}}}{r_{\text{HS}}^2 - r_{\text{SS}}^2} \right)^2 (r - r_{\text{SS}})^2 + O\left((r - r_{\text{SS}})^3 \right), \tag{S3}$$

hence, in the limit of no overlap the interaction is harmonic.

We numerically evaluate this potential, matching the gradient and linearly continuing the function $u_{\text{HS-WCA}}(r)$ for $r \leq r_{\text{HS}} + \varepsilon$, where $\varepsilon > 0$ is an arbitrary small constant, such that minimization is still meaningful if hard core overlaps do occur.

Our choice of potential is based on the fact that (i) the hard cores greatly reduce the amount of configurational space to explore, replacing expensive energy minimizations (to test whether the random walker has stepped outside the basin) with fast hard-core overlap rejections, and (ii) the hard cores exclude high-energy minima (jammed packings) that are not 'hard-sphere-like'.

TOTAL ACCESSIBLE VOLUME

The basins of attraction of energy minima tile the "accessible" phase space (schematically shown in Fig. 1b-c of the main text). This inaccessible part of the phase space arises due to hard core constraints and the existence of fluid states (see e.g. [1]). The total phase space volume is equal to V_{box}^N. The inaccessible part of this volume arising from the hard core constraints (shown as hatched areas in Fig. 1 of main text) is denoted by V_{HS}, and the part corresponding to the coexisting unjammed fluid states is denoted by V_{unj} (shown as blue regions with squares in Fig. 1 of main text). V_{unj} is significant only for finite size systems at or near unjamming. We denote the space tiled by the basins of mechanically stable jammed packings by V_J. We then have $V_J = V_{\text{box}}^N - V_{\text{HS}} - V_{\text{unj}}$. In practice we compute V_J using the following equation

$$\ln V_J(N, \phi_{\text{HS}}, \phi_{\text{SS}}) = N \ln V_{\text{box}} - N f_{\text{ex}}(\phi_{\text{HS}}) + \ln p_J(\phi_{\text{SS}}), \tag{S4}$$

where $f_{\text{ex}}(\phi_{\text{HS}})$ is the excess free energy, *i.e.* the difference in free energy between the hard sphere fluid and the ideal gas, computed from the Santos-Yuste-Haro (eSYH) equation of state [6] as in [1], and $p_J(\phi_{\text{SS}})$ is the probability of obtaining a jammed packing at soft volume fraction ϕ_{SS} with our protocol, shown in Fig. S8.

COUNTING BY SAMPLING

We briefly review our approach to computing the number Ω of distinct jammed packings for a system of N soft disks at volume fraction ϕ. We prepare packings by the protocol described above, that generates jammed structures (energy minima) with probability p_i proportional to the volume of their basin of attraction v_i. We define the probability of sampling the i-th packing as

$$p_i = \frac{v_i}{V_J}, \tag{S5}$$

where V_J is the total accessible phase space, such that

$$V_J = \sum_{i=1}^{\Omega} v_i. \tag{S6}$$

Details of the computation of v_i are discussed in Refs. [1, 7]. To find Ω, we make the simple observation

$$\sum_{i=1}^{\Omega} v_i = \frac{\Omega}{\Omega} \sum_{i=1}^{\Omega} v_i = \Omega \langle v \rangle, \tag{S7}$$

from which it follows immediately that

$$\Omega = \frac{V_J}{\langle v \rangle}. \tag{S8}$$

The 'Boltzmann-like' entropy, suggested in a similar form by Edwards [8], is then

$$S_B = \ln \Omega - \ln N! \tag{S9}$$

where the $\ln N!$ correction ensures that two systems in identical macrostates are in equilibrium under exchange of particles [9–11].

Note that $\langle v \rangle$ is the *unbiased* average basin volume (the mean of the unbiased distribution of volumes). We distinguish between the biased, $\mathcal{B}(\phi; F)$ (as sampled by the packing protocol), and the unbiased, $\mathcal{U}(\phi; F)$, basin log-volumes distributions ($F = -\ln v_{\text{basin}}$). Since the configurations were sampled proportional to the volume of their basin of attraction, we can compute the unbiased distribution as

$$\mathcal{U}(\phi; F) = \mathcal{Q}(\phi) \mathcal{B}(\phi; F) e^F \tag{S10}$$

where $\mathcal{Q}(\phi)$ is the normalisation constant, such that

$$\mathcal{Q}(\phi) = \left[\int_{F_{\min}}^{\infty} \mathrm{d}F \mathcal{B}(\phi; F) e^F \right]^{-1} = \langle v \rangle(\phi). \tag{S11}$$

Since small basins are much more numerous than large ones, and grossly under-sampled, it is not sufficient to perform a weighted average of the sampled basin volumes. Instead, to overcome this problem, one can fit the biased measured basin log-volumes distribution $\mathcal{B}(\phi; F)$ with an analytical (or at least numerically integrable) distribution, and perform the unbiasing via Eq. S10 on the best fitting distribution. Different approaches to modelling this distribution give rise to somewhat different analysis methods, which all yield consistent results as shown in Ref. [1]. In this work we follow Ref. [1] and fit $\mathcal{B}(\phi; F)$ using both a (parametric) generalised Gaussian model [12], see Eq. S37, and a (non-parametric) kernel density estimate (KDE) with Gaussian kernels [13, 14] and bandwidth selection performed by cross validation [1, 15], yielding consistent results in agreement with Ref. [1]. Before performing the fit we remove outliers from the free energy distribution in an unsupervised manner, as discussed in the "Data Analysis" section of the SI.

No such additional steps are needed to compute the 'Gibbs-like' version of the configurational entropy, in fact

$$S_G = -\sum_{i=1}^{\Omega} p_i \ln p_i - \ln N! = \sum_{i=1}^{\Omega} [p_i(-\ln v_i)] + \ln V_J - \ln N! = \langle F \rangle_{\mathcal{B}} + \ln V_J - \ln N! \tag{S12}$$

is simply the arithmetic average of the *observed* volumes: The sample mean of $F = -\ln v_{\text{basin}}$ is already correctly weighted because our packing generation protocol generates packings with probability p_i.

POWER-LAW BETWEEN PRESSURE AND BASIN VOLUME

A power-law relationship between the volume of the basin of attraction of a jammed packing and its pressure was first reported in [1]. In what follows we provide insight into this expression on the basis of this work's findings. We observe that distributions of basin negative log-volumes, $F = -\ln v_{\text{basin}}$, and log-pressures, $\Lambda = \ln P$, are approximately normally distributed (see Fig. S1 and S9). We therefore expect their joint probability to be well-approximated by a bivariate Gaussian distribution $\mathcal{B}(\phi; F, \Lambda) = \mathcal{N}(\mu, \hat{\sigma})$ [32], with mean $\mu = (\mu_F, \mu_\Lambda)$ and covariance matrix $\hat{\sigma} = ((\sigma_F^2, \sigma_{F\Lambda}^2), (\sigma_{F\Lambda}^2, \sigma_\Lambda^2))$ [16]. This is consistent with the elliptical distribution of points in Fig. 2b of the main text. For a given random variable X, with an (observed/biased) marginal distribution $\mathcal{B}(X)$, the mean is given by $\mu_X(\phi) = \langle X \rangle_{\mathcal{B}} = \int X \mathcal{B}(\phi; X)\, \mathrm{d}X$. Similarly, the (biased) conditional expectation of F for a given Λ is then [16]

$$\langle F \rangle_{\mathcal{B}(\mathbf{\Gamma})}(\phi; \Lambda) \equiv \mathbb{E}[F|\phi; \Lambda] = \frac{\sigma_{F\Lambda}^2(\phi)}{\sigma_\Lambda^2(\phi)}(\Lambda - \mu_\Lambda(\phi)) + \mu_F(\phi). \tag{S13}$$

This is simply the linear minimum mean square error (MMSE) regression estimator for F, *i.e.* the linear estimator $\hat{Y}(X) = aX + b$ that minimizes $\mathbb{E}[(Y - \hat{Y}(X))^2]$. The expectation of the dimensionless free energy $\langle F \rangle_{\mathcal{B}(\mathbf{\Gamma})}(\phi; \Lambda) = -\langle \ln v \rangle_{\mathcal{B}(\mathbf{\Gamma})}(\phi; \Lambda) \geq -\ln \langle v \rangle_{\mathcal{B}(\mathbf{\Gamma})}(\phi; \Lambda)$ [17] is the average basin negative log-volume at volume fraction ϕ and log-pressure Λ. Here the average is also taken over all other relevant, but unknown, order parameters $\mathbf{\Gamma}$, such that $\langle F \rangle_{\mathcal{B}(\mathbf{\Gamma})}(\phi; \Lambda) = \int \mathrm{d}\mathbf{\Gamma} \mathcal{B}(\mathbf{\Gamma}) F(\phi; \mathbf{\Gamma}, \Lambda)$. In other words, we write the expectation of F at a given pressure as the (biased) average over the unspecified order parameters $\mathbf{\Gamma}$. An example of such a parameter would be some topological variable that makes certain topologies more probable than others. Note that $F(\phi, \Lambda; \mathbf{\Gamma})$ is narrowly distributed around $\mathbb{E}[F|\phi; \Lambda]$. To simplify the notation we write $\langle F \rangle_{\mathcal{B}(\mathbf{\Gamma})}(\phi; \Lambda) \equiv \langle F \rangle_{\mathcal{B}}(\phi; \Lambda)$. We can thus rewrite the power-law reported in [1] as

$$\begin{aligned}\langle f \rangle_{\mathcal{B}}(\phi; \Lambda) =& \lambda(\phi)\Lambda + c(\phi) \\ =& \frac{\sigma_{f\Lambda}^2(\phi)}{\sigma_\Lambda^2(\phi)}\Lambda - \frac{\sigma_{f\Lambda}^2(\phi)}{\sigma_\Lambda^2(\phi)}\mu_\Lambda(\phi) + \mu_f(\phi) \\ =& \frac{\sigma_{f\Lambda}^2(\phi)}{\sigma_\Lambda^2(\phi)}\Delta\Lambda + \mu_f(\phi)\end{aligned} \tag{S14}$$

where $f = F/N$ is the basin negative log-volume per particle and $\lambda \equiv 1/\kappa$ is the slope of the power-law relation, which depends crucially on the packing fraction ϕ. The last equality in Eq. S14 highlights how $\lambda(\phi)$ controls the contributions of the fluctuations of the log-pressures $\Delta\Lambda \equiv \Lambda - \mu_\Lambda(\phi)$ to changes in the basin negative log-volume. Note that one can rewrite the ratio of fluctuations as $\sigma_{f\Lambda}^2/\sigma_\Lambda^2 = \rho_{f\Lambda}\sigma_f/\sigma_\Lambda$ where $\rho_{f\Lambda} = \sigma_{f\Lambda}^2/(\sigma_f\sigma_\Lambda)$ is the linear correlation coefficient of f and Λ. Finally, we can gain further insight into the power-law dependence by noting that

$$\lambda(\phi) \equiv \frac{\sigma_{f\Lambda}^2(\phi)}{\sigma_\Lambda^2(\phi)} \tag{S15}$$

$$c(\phi) \equiv \mu_f(\phi) - \frac{\sigma_{f\Lambda}^2(\phi)}{\sigma_\Lambda^2(\phi)}\mu_\Lambda(\phi) \tag{S16}$$

VARIANCE OF RELATIVE PRESSURES

In this section we relate the statistics of the log-pressures of the packings to the *relative* pressures. For a given N and ϕ, with the set of pressures $\{P_i\}$, the log-pressures are given by $\Lambda_i \equiv \ln P_i$ and the relative pressures are $P_i/\langle P \rangle$. The two quantities are then simply related as

$$\ln\left(\frac{P_i}{\langle P \rangle}\right) = \ln P_i - \ln\langle P \rangle = \Lambda_i - \ln\langle P \rangle. \tag{S17}$$

Using Jensen's inequality [17], we have the following bound for the first moment of the log-pressures

$$\langle \Lambda \rangle = \langle \ln P \rangle \leq \ln\langle P \rangle. \tag{S18}$$

Therefore $\langle P \rangle \to 0$ implies $\langle \Lambda \rangle \to -\infty$. In order to relate the means and the variances of P_i and $\Lambda_i \equiv \ln P_i$, we perform the Taylor expansion

$$\ln P = \ln\langle P \rangle + \left.\frac{d \ln P}{dP}\right|_{P=\langle P \rangle}(P - \langle P \rangle) + ... \tag{S19}$$

We next compute the moments to leading order

$$\begin{aligned}\langle \ln P \rangle &\approx \ln\langle P \rangle, \\ \sigma^2(\ln P) &\approx \frac{\sigma^2(P)}{\langle P \rangle^2} = \sigma^2\left(\frac{P}{\langle P \rangle}\right).\end{aligned} \tag{S20}$$

Thus, to first order, the variance of the log-pressures is equal to the variance of the relative pressures.

Bounds

For a fixed N and ϕ, the pressures of the individual packings are bounded as

$$0 < P_{\mathrm{min}} \leq P_i \leq P_{\mathrm{max}}. \tag{S21}$$

where P_{max} and P_{min} are determined by the packing fraction ϕ, independent of system size, and are physically set limits [18–20]. We can therefore use Popoviciu's inequality on variances, yielding

$$\sigma^2(P) \leq \frac{1}{4}\left(P_{\mathrm{max}} - P_{\mathrm{min}}\right)^2, \tag{S22}$$

which is also bounded. The relative pressure fluctuations are therefore bounded as

$$\sigma^2\left(\frac{P}{\langle P\rangle}\right) \leq \frac{1}{4}\frac{\left(P_{\mathrm{max}} - P_{\mathrm{min}}\right)^2}{\langle P\rangle^2}. \tag{S23}$$

We thus find that the variance of the relative pressures $\sigma^2\left(\frac{P}{\langle P\rangle}\right)$ can diverge only when $\langle P\rangle \to 0$, which is precisely where the unjamming transition occurs.

Scaling at $\phi \gg \phi^*$

First we show that away from a critical point, the relative pressure fluctuations scale as $1/L^2$, where $L = \sqrt{N}$ and N is the number of particles. The internal Virial is defined by $P = \sum_{i=1,N} p_i$, where p_i is the particle level "pressure" given by $p_i = \sum_j \sum_{\alpha=1,2} \mathbf{r}^{\alpha}_{i,j}\mathbf{f}^{\alpha}_{i,j}$ where $\mathbf{r}^{\alpha}_{i,j}$ and $\mathbf{f}^{\alpha}_{i,j}$ are the contact vectors and contact forces, respectively. The variance of $P/\langle P\rangle$ is

$$\sigma^2(P/\langle P\rangle) = \frac{\langle P^2\rangle - \langle P\rangle^2}{\langle P\rangle^2} = \frac{\sum_{i=1}^N \sigma^2(p_i) + \sum_{i\neq j}\mathrm{cov}(p_i,p_j)}{\left(\sum_{i=1}^N \langle p_i\rangle\right)^2}. \tag{S24}$$

When away from a critical point we expect $\sum_{i\neq j}\mathrm{cov}(p_i,p_j)$ to scale subextensively, and the variance of relative pressure fluctuations to be

$$\sigma^2(P/\langle P\rangle) = \frac{\sum_{i=1}^N \sigma^2(p_i)}{\left(\sum_{i=1}^N \langle p_i\rangle\right)^2} \sim \frac{1}{N} \tag{S25}$$

hence the relative pressure fluctuations away from the critical point will scale as $1/N = 1/L^2$, and

$$\sigma^2(\Lambda) \approx \sigma^2(P/\langle P\rangle) \sim 1/L^2, \tag{S26}$$

as can be verified in Fig. 3a of the main text.

Second, we analyse the covariance of the basin negative log-volume per particle ($F/N = -\ln(v_{\mathrm{basin}})/N$) and the relative pressure fluctuations, $\mathrm{cov}(F/N, P/\langle P\rangle)$. Similarly to the internal Virial we define the particle level basin negative log-volume as $F = \sum_{i=1,N} f_i$. Then we have

$$\mathrm{cov}(F/N, P/\langle P\rangle) = \frac{\langle FP\rangle - \langle F\rangle\langle P\rangle}{N\langle P\rangle} = \frac{\sum_{i=1}^N \mathrm{cov}(f_i,p_i) + \sum_{i\neq j}\mathrm{cov}(f_i,p_j)}{N\sum_{i=1}^N \langle p_i\rangle}. \tag{S27}$$

From the power-law relation between F and Λ we know that away from the critical point $\mathrm{cov}(f_i,p_i) > 0$ and we expect $\sum_{i\neq j}\mathrm{cov}(f_i,p_j)$ to scale subextensively in this region, hence

$$\mathrm{cov}(F/N, P/\langle P\rangle) = \frac{\sum_{i=1}^N \mathrm{cov}(f_i,p_i)}{N\sum_{i=1}^N \langle p_i\rangle} \sim \frac{1}{N}, \tag{S28}$$

therefore

$$\mathrm{cov}(f,\Lambda) \approx \mathrm{cov}(f, P/\langle P\rangle) \sim 1/L^2, \tag{S29}$$

where $f = F/N$. We can thus conclude that the slope of the power-law relation is

$$\lambda_{\phi\gg\phi^*} = \frac{\sigma^2(f,\Lambda)}{\sigma^2(\Lambda)} \sim \mathcal{O}(1), \tag{S30}$$

In other words, λ is independent of system size, a fact that has been verified numerically in Ref. [1].

Scaling as $\phi \to \phi^*$

Near the critical point, as $\phi \to \phi^*$, the variance of Λ follows the scaling form

$$\sigma^2(\Lambda) \sim L^{\gamma/\nu-2}, \tag{S31}$$

with $\gamma/\nu \approx 1$ as found by finite size scaling, shown in Fig. S11. While we do not have a finite size scaling collapse for the covariance $\sigma^2(f,\Lambda)$, due to the high computational cost of performing the basin volume calculations for multiple system sizes, we do observe that for $N = 64$ the covariance decreases with respect to the "background" $1/L^2$ fluctuations as $\phi \to \phi^*$, see Fig. S2. Hence, we do not expect $\sigma^2(f,\Lambda)$ to diverge but rather that

$$\sigma^2(f,\Lambda) \lesssim L^{-2}, \tag{S32}$$

Hence in the limit $\phi \to \phi^*$ we expect that the slope of the power-law relation will be

$$\lambda_{\phi\to\phi^*} = \frac{\sigma^2(f,\Lambda)}{\sigma^2(\Lambda)} = 0. \tag{S33}$$

Relation between scaling exponents

Starting from Eq. S14 (Eq. 1 in the main text), we use the fact that $\sigma^2(aX \pm bY) = a^2\sigma^2(X) + b^2\sigma^2(Y) \pm 2ab\,\mathrm{cov}(X,Y)$ to compute the variance of $f = F/N$ to find

$$\sigma_f^2 = \lambda^2\sigma_\Lambda^2 = (\sigma_{f\Lambda}^2)^2/\sigma_\Lambda^2 \tag{S34}$$

By rearranging this expressions we find that

$$(\sigma_{f\Lambda}^2)^2/\sigma_f^2 = \sigma_\Lambda^2 \sim \begin{cases} L^{-2} \text{ for } \phi \gg \phi^* \\ L^{-\zeta} \text{ for } \phi \to \phi^* \end{cases} \tag{S35}$$

where we have defined $\zeta \equiv 2 - \gamma/\nu \approx 1$ as in Eq. (S31) and as found by finite size scaling, shown in Fig. S11. For $\phi \to \phi^*$, by assuming scalings $\sigma_{f\Lambda}^2 \sim L^{-\eta}$ and $\sigma_f^2 \sim L^{-\vartheta}$, we find the following relation between scaling exponents

$$2\eta - \vartheta = \zeta \tag{S36}$$

DISTRIBUTION OF BASIN LOG-VOLUMES

The distributions of basin negative log-volumes, shown in Fig. S1, are well represented by a three-parameter generalised Gaussian distribution

$$\mathcal{B}(F|\mu_F,\sigma_F,\zeta) \equiv \frac{\zeta}{2\sigma\Gamma(1/\zeta)} \exp\left[-\left(\frac{|F-\mu_F|}{\sigma_F}\right)^\zeta\right], \tag{S37}$$

where $\Gamma(x)$ is the gamma function, σ_F is the scale parameter, ζ is the shape parameter and μ_F is the mean log-volume with variance $\sigma^2\Gamma(3/\zeta)/\Gamma(1/\zeta)$. In Ref. [4] it is shown that in the limit $N \to \infty$ the shape parameter approaches that of a standard Gaussian distribution, $\zeta = 2$. Since $\sigma_F^2 \sim N$ and $\mu_F \sim N$, we have that

$$e^{-(F-\mu_F)^2/(2\sigma^2)} \sim e^{-N(f-\mu_f)^2} \to \delta(\mu_f) \text{ as } N \to \infty, \tag{S38}$$

where δ is the Dirac delta function and $f = F/N$. The distribution of basin volumes thus becomes infinitely narrow in the thermodynamic limit. However, this is not sufficient for the Edwards conjecture to be correct, in fact we also require that the basin volumes are uncorrelated with respect to any structural observables in this limit. In this manuscript we argue that this occurs only as $\phi \to \phi^*$.

DATA ANALYSIS

Reduced units

While presenting data from our computations, we express pressure and volume in reduced units as P/P^* and v/v^* respectively. The unit of volume is given by $v^* \equiv \pi\langle r_{HS}^2\rangle$, where $\langle r_{HS}^2\rangle$ is the mean squared hard sphere radius. The unit of pressure is then $P^* \equiv \epsilon/v^*$, where ϵ is the stiffness of the soft-sphere potential, defined in Eq. S2. The pressure is computed as $P = \mathrm{Tr}(\hat{\mathbf{\Sigma}})/2V_{\mathrm{box}}$ where $\hat{\mathbf{\Sigma}}$ is the Virial stress tensor and V_{box} the volume of the enclosing box.

Summary of calculations

For the basin volume calculations we consider systems of $N = 64$ disks sampled at a range of 8 volume fractions $0.828 \leq \phi \leq 0.86$ and for each ϕ we measure the basin volume for about $365 < M < 770$ samples.

For the finite size scaling analysis of the relative pressure fluctuations we study system sizes $N = 32, 48, 64, 80, 96, 128$ for 48 volume fractions in the range $0.81 \leq \phi \leq 0.87$. For each system size we generate up to 10^5 hard disk fluid configurations and compute the pressure for between approximately 10^3 and 10^4 jammed packings (depending on the probability of obtaining a jammed packing at each volume fraction).

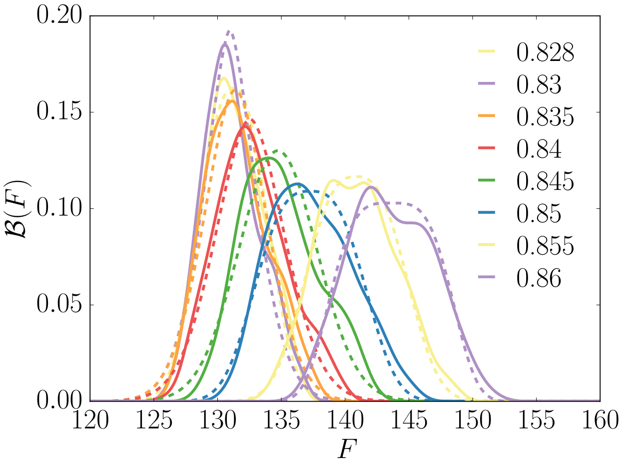
Simulations were performed using the open source libraries PELE [21] and MCPELE [22].

Outlier detection and robust covariance estimation

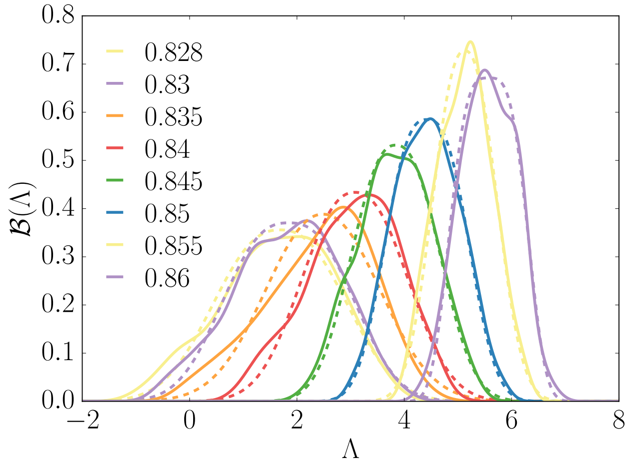
Before manipulating the raw data we remove outliers from the joint distribution $\mathcal{B}(f, \Lambda)$ following the distance-based outlier removal method introduced by Knorr and Ng [23]. This is applied in turn to each dimension, such that we choose to keep only those points for which at least $R = 0.5$ of the remaining data set is within $D = 4\sigma$ (compared to the much stricter $R = 0.9988$, $D = 0.13\sigma$ required to exclude any points further than $|\mu - 3\sigma|$ for normally distributed data [23]). On our datasets we find that this procedure removes typically none and at most 0.8% of all data points.

Mean and covariance estimates of $\mathcal{B}(f, \Lambda)$ are computed using a robust covariance estimator, namely the Minimum Covariance Determinant (MCD) estimator [14, 24] with support fraction $h/n_{\mathrm{samples}} = 0.99$. The MCD estimator defines μ_{MCD}, the mean of the h observations for which the determinant of the covariance matrix is minimal, and $\hat{\sigma}_{\mathrm{MCD}}$, the corresponding covariance matrix [25]. We use these robust estimates of the location and of the covariance matrix (computed over 1000 bootstrap samples [26]) to fit our observations by linear MMSE [16], see Fig. 2 of the main text.

Before fitting $\mathcal{B}(f)$ (required to compute Ω), we perform an additional step of outlier detection based on an elliptic (Gaussian) envelope criterion constructed using the MCD estimator. We assume a support fraction $h/n_{\mathrm{samples}} = 0.99$ and a contamination equal to 10% [14]. We compute S_G and S_B from the resulting datasets. The procedure is strictly unsupervised and allows us to achieve robust fits despite the small sample sizes. We fit $\mathcal{B}(f)$ using both a (parametric) generalised Gaussian model [12] and a (non-parametric) kernel density estimate (KDE) with Gaussian kernels [13, 14] and bandwidth selection performed by cross validation [1, 15].



(a)



(b)

FIG. S1: Observed distribution of the basin log-volume F (a) and log-pressure Λ (b) for jammed packings of $N = 64$ HS-WCA polydisperse disks at various volume fractions $0.828 \leq \phi \leq 0.86$. Solid lines are Kernel Density Estimates and dashed lines are generalised Gaussian fits.

Error analysis

Errors were computed analytically where possible and propagated using the 'uncertainties' Python package [27]. Alternatively, intervals of confidence were computed by bootstrap for the covariance estimation [26] and by BCa bootstrap otherwise using the 'scikit-bootstrap' Python package [28, 29].

RESULTS

Distributions of F and Λ

In Fig. S1 we show the biased distributions $\mathcal{B}(F)$ and $\mathcal{B}(\Lambda)$ of the basin negative log-volumes and log-pressures, which are the marginal distributions of the joint distribution $\mathcal{B}(F, \Lambda)$ shown in Fig. 2b. In Fig. S2 we plot the moments of $\mathcal{B}(F, \Lambda)$, namely the elements of the mean $\mu = (\mu_f, \mu_\Lambda)$ and the elements of the covariance matrix $\hat{\sigma} = ((\sigma_f^2, \sigma_{f\Lambda}^2), (\sigma_{f\Lambda}^2, \sigma_\Lambda^2))$, as well as the linear correlation coefficient $\rho_{f\Lambda} = \sigma_{f\Lambda}^2/(\sigma_f \sigma_\Lambda)$.

Estimates of the equiprobability density $\phi^*_{N=64}$

We summarise the estimated values for $\phi^*_{N=64}$ in Table S1

	λ	$S_B^{(Gauss)}$	$S_B^{(KDE)}$
$\phi^*_{N=64}$	0.824 ± 0.070	0.823	0.823
$\langle z \rangle_{\mathrm{sig}}(\phi^*_{N=64})$	4.050 ± 0.24	4.048	4.048

TABLE S1: Predicted values of $\phi^*_{N=64}$ obtained from the linear extrapolation of $\lambda \to 0$ and from the point of intersection of the Gibbs entropy S_G with the Boltzmann entropy S_B, computed both parametrically by fitting $\mathcal{B}(f)$ with a generalised Gaussian function ('Gauss') and non-parametrically by computing a Kernel Density Estimate ('KDE') of the distribution. The corresponding average contact number has been computed using a sigmoid fit (Eq. S39) of the data in Fig. S3.

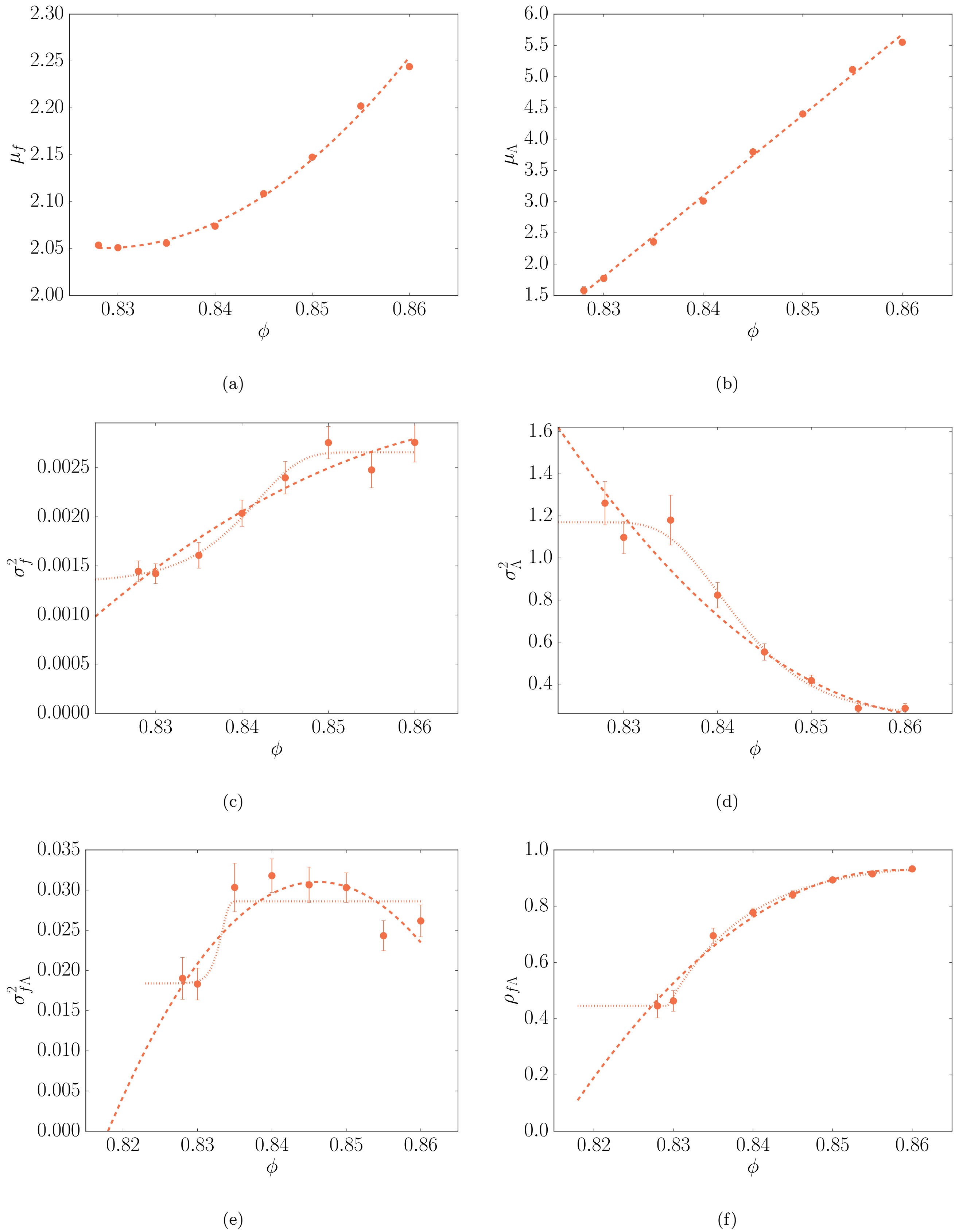


FIG. S2: Moments of the joint distribution $\mathcal{B}(f,\Lambda)$ for jammed packings of $N = 64$ HS-WCA polydisperse disks at various volume fractions $0.828 \leq \phi \leq 0.86$. Elements of the mean $\mu = (\mu_f, \mu_\Lambda)$ are shown in (a) and (b) respectively. Elements of the covariance matrix $\hat{\sigma} = ((\sigma_f^2, \sigma_{f\Lambda}^2), (\sigma_{f\Lambda}^2, \sigma_\Lambda^2))$ are shown in (c)-(e). The linear correlation coefficient $\rho_{f\Lambda} = \sigma_{f\Lambda}^2/(\sigma_f \sigma_\Lambda)$ is shown in (f). All values are computed by the MCD estimator with 0.99 support fraction over 1000 bootstrap samples. Error bars are standard errors computed by bootstrap. Dashed lines are second order polynomial fits and dotted lines are sigmoid fits (Eq. S39). Curves of best fit are meant as guide to the eye.

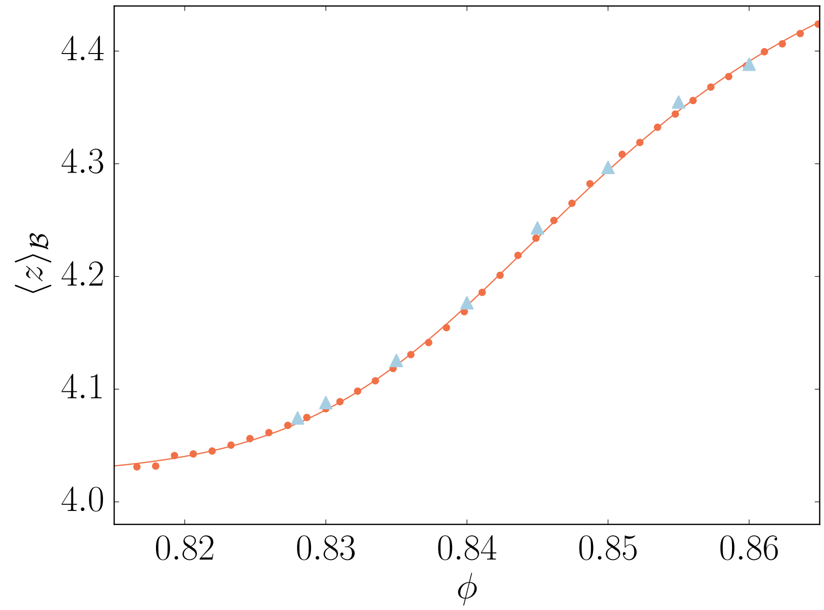


FIG. S3: Observed average contact number for jammed packings of $N = 64$ HS-WCA polydisperse disks. Triangles are estimates from the basin volume measurements datasets. Circles are estimates from the independent measurements used for the finite size scaling analysis. The solid line is a generalised sigmoid fit (Eq. S39) of the latter.

Contact number

The mean contact number is plotted as a function of volume fraction in Fig. S3. The data are fitted with a generalised sigmoid function of the form

$$f(a, b, \phi_0, u, w; \phi) = a - \frac{a-b}{(1+e^{-w(\phi-\phi_0)})^{1/u}} \tag{S39}$$

In Fig. S4 we also show the fraction of rattlers at different packing fractions for $N = 64$ disks. We note that the fraction of rattlers is maximal and saturates over the same range of densities as where the relative pressure fluctuations are maximal and saturate (see Fig. 3 of main text and finite size scaling section for further discussion).

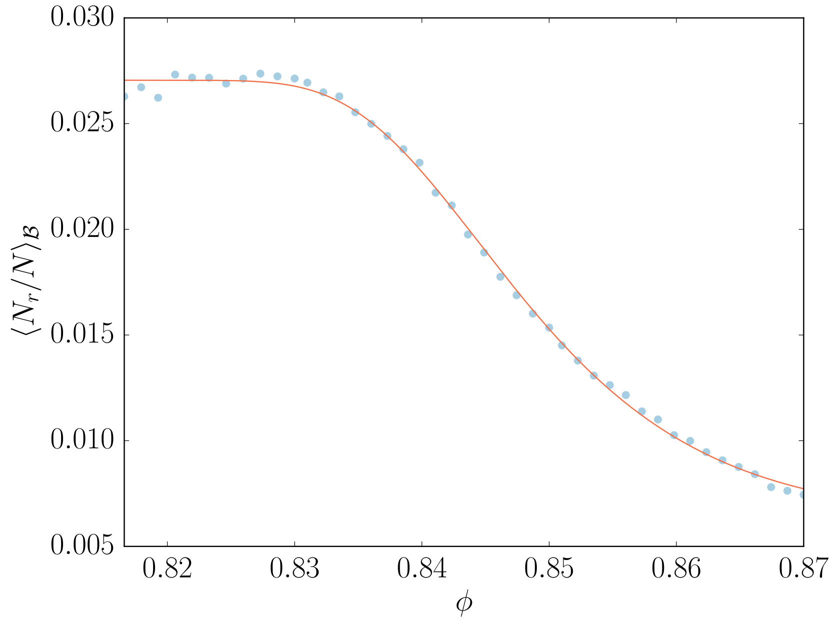


FIG. S4: Fraction of rattlers for jammed packings of $N = 64$ HS-WCA polydisperse disks. We find this tends to a maximum and saturates over the same range of packing fractions where relative pressure fluctuations are maximal, see Fig. 3 of main text. The solid line is a generalised sigmoid fit, Eq. S39.

Correlations with structural parameters

We analyse the correlation of the basin negative log-volume with a number of structural parameters other than the pressure $P = \mathrm{Tr}(\hat{\mathbf{\Sigma}})/(dL^d)$, where $\hat{\mathbf{\Sigma}}$ is the stress tensor and $d = 2$, discussed in detail in the main text (see Fig. 2).

For all observables X we assume a linear correlation defined analogously to Eq. 1, namely

$$\langle f \rangle_{\mathcal{B}}(\phi; X) = \lambda_X(\phi) \ln X + c_X(\phi). \tag{S40}$$

We perform the analysis for the individual elements of the stress tensor $\hat{\mathbf{\Sigma}}_{ij}$, the average contact number z and the Q_6 bond-orientational order parameter [30]. Scatter plots with bootstrapped linear MMSE fits are shown in Fig. S5, and the fitted parameters are plotted as a function of volume fraction in Fig. S6. The results are qualitatively similar to those obtained for the pressure in that λ_X is decreasing towards 0 as $\phi \to \phi^*$, indicating that the basin volumes decorrelate from X in this limit. As explained in the main text, this is a necessary condition for the equiprobability of jammed states.

In Fig. S5d, we observe that λ_{Q_6} becomes precisely zero at the lowest volume fractions while for larger volume fractions $\lambda_{Q_6} < 0$, implying that larger basins correspond to more ordered structures. At the same time we note from Figs. S5e-f that larger volumes correspond on average to lower average contact numbers (z) and that z and Q_6 are (therefore) negatively correlated.

Global model of $\mathcal{B}(F, \Lambda)$

In Fig. 2b we fit the joint probability $\mathcal{B}(\phi; F, \Lambda)$ by linear MMSE, or in other words we fit the data with a bivariate Guassian with first principal component corresponding to the linear fit. We compute linear MMSE fits for each volume fraction ϕ independently, as the probabilities of the packings $(-\ln p_i = F_i + \ln V_J(\phi))$ are obtained by subtracting different normalization constants for each ϕ (accessible volume $V_J(\phi)$). However, since $\mu_f(\phi) \equiv \langle F/N \rangle_{\mathcal{B}}(\phi)$ and $\mu_\Lambda \equiv \langle \Lambda \rangle_{\mathcal{B}}(\phi)$ are slowly varying (see Fig. S2a-b), we attempt to fit to the full distribution of $\mathcal{B}(f, \Lambda)$ (for all ϕ at once) by an exponential function of the form $a\exp(-b\lambda) + c$, and a third order polynomial $\mathrm{p}_3(\Lambda)$. Fits are shown in Fig. S7a, showing an evident decay of the correlations between pressure and basin volume. Evaluating the derivative of these global fits at each $\mu_\Lambda(\phi)$ we find that they are in excellent agreement with the estimates of λ obtained by linear MMSE, see Eq. S15.

Finite size scaling

In order to locate the unjamming transition, we compute the probability of obtaining jammed packings as a function of volume fraction ϕ. A finite size scaling collapse for $p_J L^{\beta/\nu}$ vs. $L^{1/\nu}\left(\phi/\phi^J_{N\to\infty} - 1\right)$, shown in Fig. S8, yields critical exponents $\nu \approx 1$, $\beta = 0$ and critical volume fraction $\phi^{J_{(p_J)}}_{N\to\infty} = 0.844(2)$, in agreement with Vagberg et al. [31]. We obtain an independent estimate of the unjamming transition by locating the point where the average pressure goes to zero and therefore $\langle \Lambda \rangle_{\mathcal{B}} \to -\infty$. KDE distributions for Λ are shown in Fig. S9. The average log-pressure is shown in Fig. S10a and a finite size scaling collapse for $\langle \Lambda \rangle_{\mathcal{B}} L^{\xi/\nu}$ vs. $L^{1/\nu}\left(\phi/\phi^J_{N\to\infty} - 1\right)$, shown in Fig. S10b, yields $\nu = 0.50(5)$, $\xi = 0.62(3)$ and critical volume fraction $\phi^{J_{(\Lambda)}}_{N\to\infty} = 0.841(3)$.

We then analyse the relative pressure fluctuations $\chi_P = N\sigma^2(P/\langle P \rangle_{\mathcal{B}})$ and the log-pressure fluctuations $\chi_\Lambda = N\sigma^2_\Lambda$. A scaling collapse for different system sizes of $\chi_P L^{-\gamma/\nu}$ vs. $L^{1/\nu}\left(\phi/\phi^J_{N\to\infty} - 1\right)$ with $L = N^{1/d}$, shown in Fig. S11a, yields $\nu = 0.5(3)$, $\gamma = 0.47(5)$ and $\phi^{*(P)}_{N\to\infty} = 0.841(3)$. An analogous scaling collapse of $\chi_\Lambda L^{-\gamma/\nu}$ vs. $L^{1/\nu}\left(\phi/\phi^*_{N\to\infty} - 1\right)$, shown in Fig. S11b, yields $\nu = 0.5(3)$, $\gamma = 0.89(5)$ and $\phi^{*(\Lambda)}_{N\to\infty} = 0.841(3)$.

Together these results lead us to conclude that the point of equiprobability $\phi^*_{N\to\infty}$ coincides with the unjamming point $\phi^J_{N\to\infty}$, to within numerical error and finite size corrections that we do not take into account. Note that the precise numerical value of ν varies through the literature and has been shown to depend on the quantity being observed, and also crucially on finite size corrections to scaling [31]. In this work we have not attempted to establish ν definitely, nor elucidate its origin with respect to the diverging correlation length(s) that might be involved.

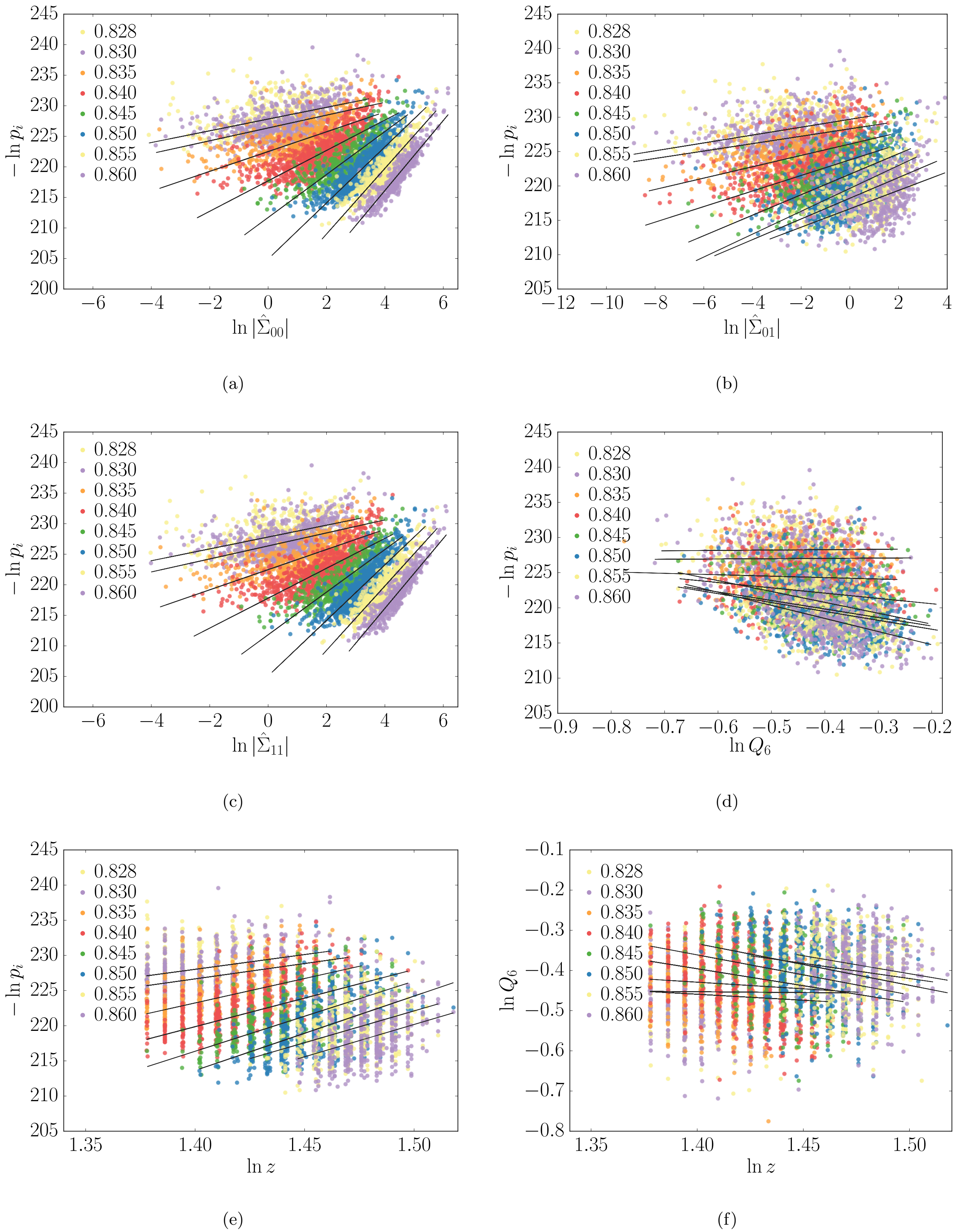
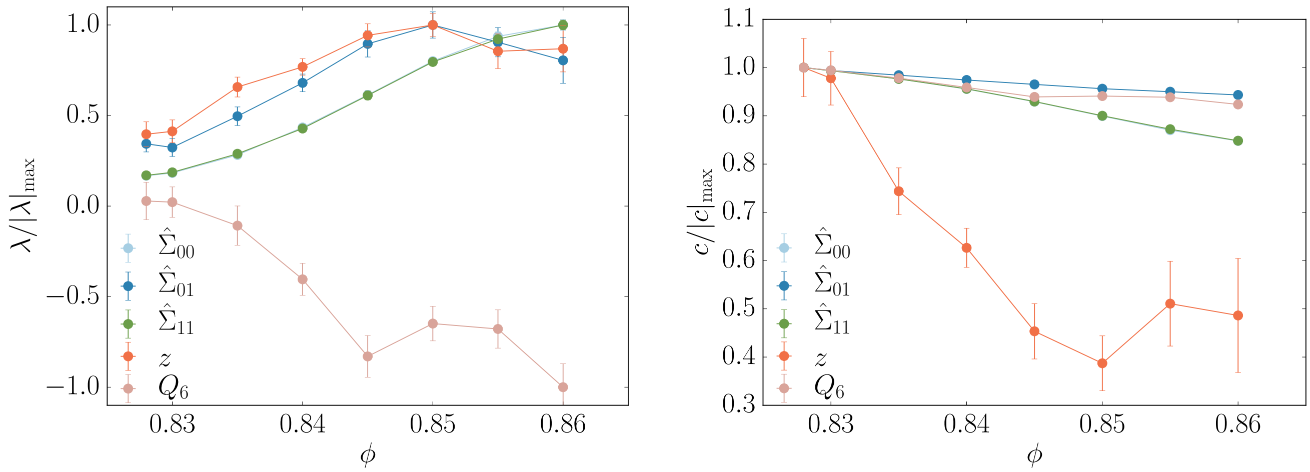
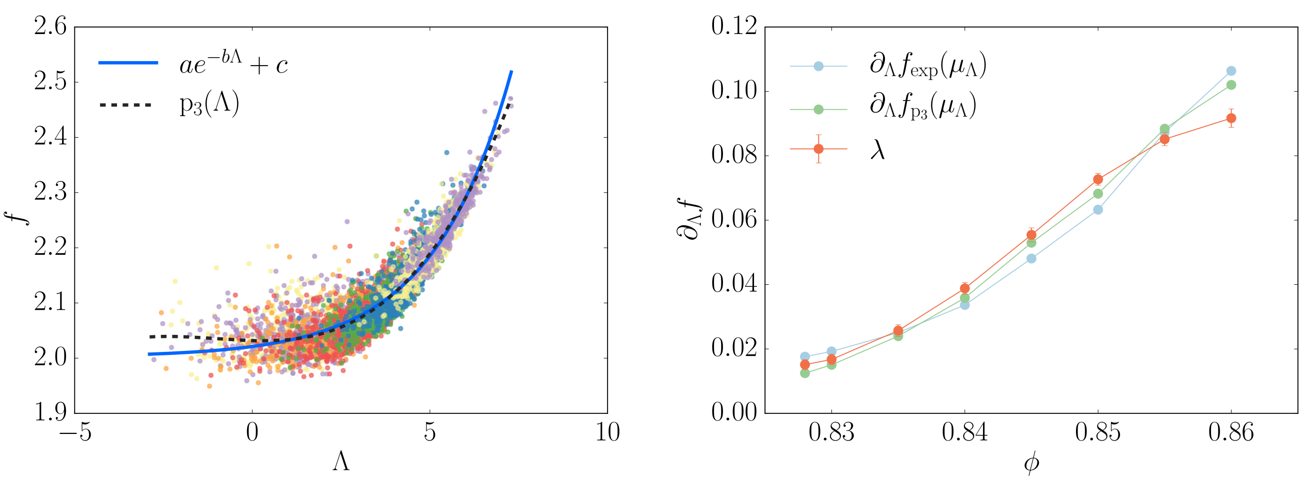


FIG. S5: Scatter plots of the negative log-probability of observing a packing, $-\ln p_i = F_i + \ln V_J(\phi)$, where V_J is the accessible fraction of phase space, as a function of the individual terms of the stress tensor $\hat{\mathbf{\Sigma}}$ (a)-(c), the Q_6 bond-orientational order parameter (d) and the average contact number z (e). The scatter plot in (f) shows the Q_6 bond-orientational order parameter as a function of the average contact number z. Black solid lines are lines of best fit computed by bootstrapped linear MMSE using a robust covariance estimator.



(a) (b)

FIG. S6: Slopes λ_X (a) and intercepts c_X (b) of Eq. S40 for the individual components of the the stress tensor $\hat{\mathbf{\Sigma}}$, the Q_6 bond-orientational order parameter, and the average contact number z. Estimates were obtained by bootstrapped linear MMSE fits using a robust covariance estimator and error bars refers to the standard error computed by bootstrap. Solid lines are guide to the eye.



(a) (b)

FIG. S7: (a) Global fit of $\mathcal{B}(f,\Lambda)$ by an exponential function of the form $a\exp(-b\lambda)+c$, and a third order polynomial $\mathrm{p}_3(\Lambda)$. (b) First derivative of the fits evaluated at mean log-pressure $\mu_\Lambda(\phi)$ are in excellent agreement with the estimates of λ obtained by linear MMSE, see Eq. S15. Solid lines are guide to the eye.

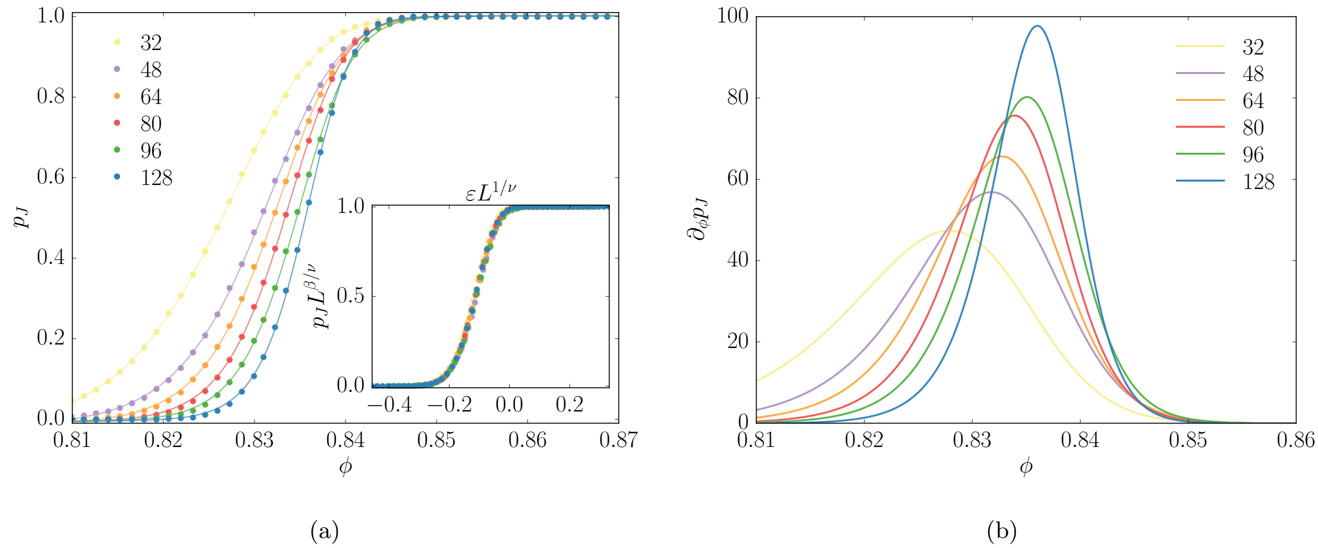


FIG. S8: (a) Probability of obtaining a jammed packing p_J by our preparation protocol for $N = 32$ to 128 HS-WCA polydisperse disks as a function of volume fraction. Inset: Scaling collapse for $p_J L^{\beta/\nu}$ vs. $L^{1/\nu}\left(\phi/\phi^J_{N\to\infty} - 1\right)$, with $L = N^{1/d}$, yields critical exponents $\nu \approx 1$, $\beta = 0$ and critical volume fraction $\phi^{J_{(p_J)}}_{N\to\infty} = 0.844(2)$. Circles are observed data and solid lines correspond to sigmoid fits, Eq. S39. (b) Derivative of the sigmoid fits for p_J for different numbers of disks.

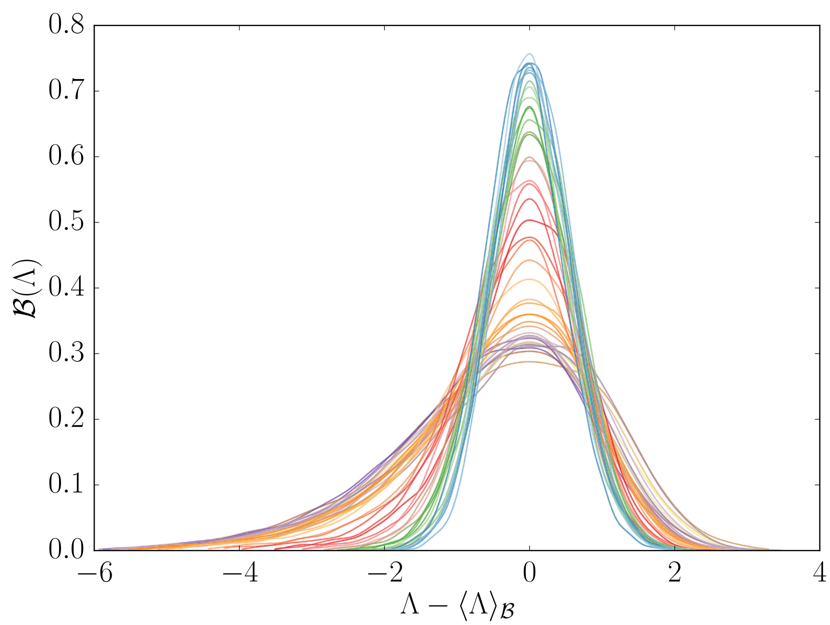
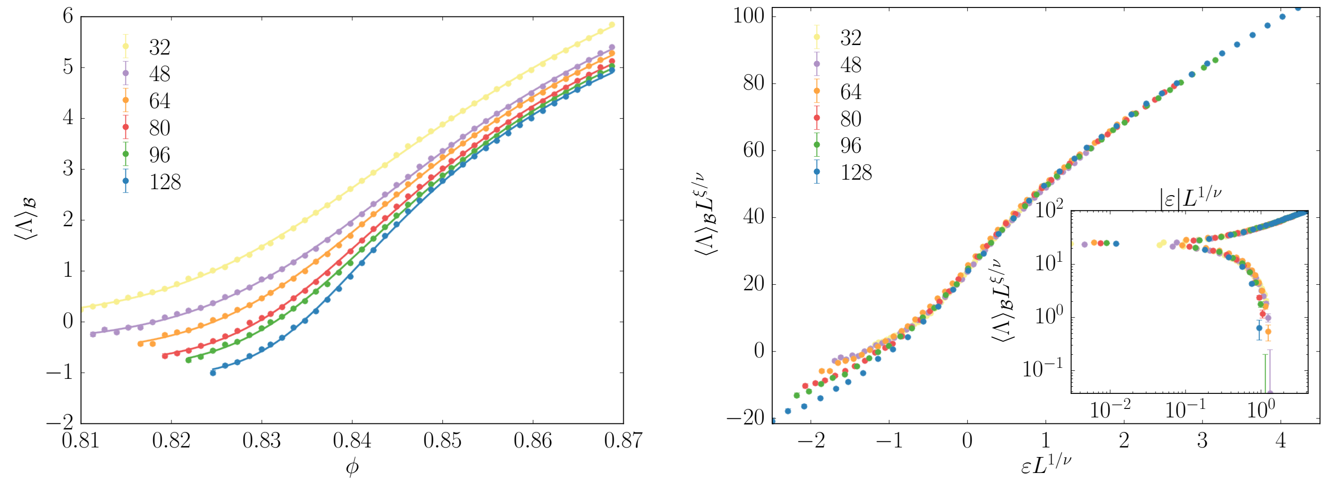
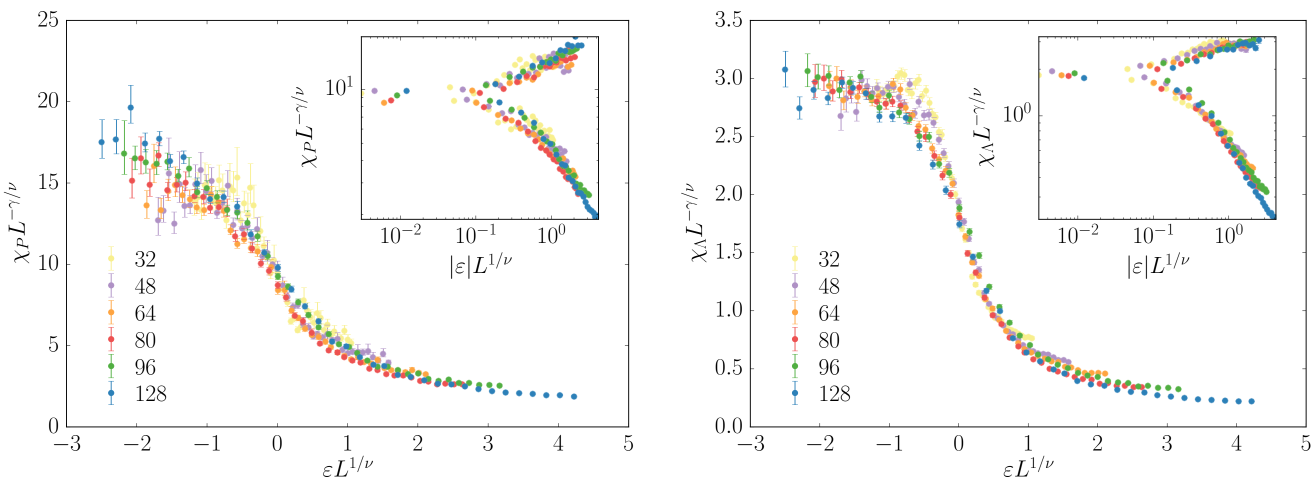


FIG. S9: Log-transformed observed (biased) distribution of pressures for jammed packings of $N = 64$ HS-WCA polydisperse disks, centred around the mean. The variance grows for decreasing volume fractions and becomes more skewed towards low pressures. The overall Gaussian shape is consistent with a log-normal distribution of pressures. Curves are kernel density estimates with Gaussian kernels [13, 14] and bandwidth selection performed by cross validation [1, 15]



(a) (b)

FIG. S10: (a) Average log-pressure $\langle\Lambda\rangle_{\mathcal{B}}$ for N HS-WCA polydisperse disks. (b) Scaling collapse for $\langle\Lambda\rangle_{\mathcal{B}} L^{\xi/\nu}$ vs. $L^{1/\nu}\left(\phi/\phi^{J}_{N\to\infty} - 1\right)$, with $L = N^{1/d}$. The estimated critical exponents are $\nu = 0.50(5)$ and $\xi = 0.62(3)$, and the critical volume fraction $\phi^{J_{(\Lambda)}}_{N\to\infty} = 0.841(3)$. Inset: A logarithmic plot of the same data. Circles are observed data and solid lines are sigmoid fits, Eq. S39. Error bars, computed by BCa bootstrap [28], refer to 1σ confidence intervals.



(a) (b)

FIG. S11: (a) Data collapse from finite size scaling analysis of the variance of the relative pressures. The plot shows $\chi_P L^{-\gamma/\nu}$ vs. $L^{1/\nu}\left(\phi/\phi^{J}_{N\to\infty} - 1\right)$, with $L = N^{1/d}$. The estimated critical exponents are $\nu = 0.5(3)$ and $\gamma = 0.47(5)$, and the critical volume fraction is $\phi^{*(P)}_{N\to\infty} = 0.841(3)$. (b) Scaling collapse of the variance of the log-pressures. The plot shows $\chi_\Lambda L^{-\gamma/\nu}$ vs. $L^{1/\nu}\left(\phi/\phi^{*}_{N\to\infty} - 1\right)$. The estimated critical exponents are $\nu = 0.5(3)$ and $\gamma = 0.89(5)$, and the critical volume fraction is $\phi^{*(\Lambda)}_{N\to\infty} = 0.841(3)$. Error bars, computed by BCa bootstrap, refer to 1σ confidence intervals.

* sm958@cam.ac.uk

[1] S. Martiniani, K. J. Schrenk, J. D. Stevenson, D. J. Wales, D. Frenkel, Phys. Rev. E **93**, 012906 (2016).
[2] J. D. Weeks, D. Chandler, H. C. Andersen, J. Chem. Phys. **54**, 5237 (1971).
[3] E. Bitzek, P. Koskinen, F. Ghler, M. Moseler, P. Gumbsch, Phys. Rev. Lett. **97**, 170201 (2006).
[4] D. Asenjo, J. D. Stevenson, D. J. Wales, D. Frenkel, J. Phys. Chem. B **117**, 12717 (2013).
[5] C. P. Goodrich, A. J. Liu, S. R. Nagel, Phys. Rev. Lett. **109**, 95704 (2012).
[6] A. Santos, S. B. Yuste, M. Lopez De Haro, Mol. Phys. **96**, 1 (1999).
[7] S. Martiniani, K. J. Schrenk, J. D. Stevenson, D. J. Wales, D. Frenkel, Phys. Rev. E **94**, 031301 (2016).
[8] S. F. Edwards, R. B. S. Oakeshott, Physica A **157**, 180 (1991).
[9] R. H. Swendsen, Am. J. Phys. **74**, 187 (2006).
[10] D. Frenkel, Mol. Phys. **112**, 2325 (2014).
[11] M. E. Cates, V. N. Manoharan, Soft Matter **11**, 6538 (2015).
[12] S. Nadarajah, J. Appl. Stat. **32**, 7, 685 (2005).
[13] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, New York, 2009.
[14] Pedregosa et al., J. Mach. Lern. Res. **12**, 2825 (2011).
[15] A. W. Bowman, Biometrika **71**, 353 (1984.)
[16] D. P. Bertsekas, J. N. Tsitsiklis, *Introduction to probability, Vol. 1.* Belmont, MA: Athena Scientific, 2002.
[17] E. W. Weisstein, "Jensen's Inequality" From MathWorld – A Wolfram Web Resource.
[18] S. Henkes, B. Chakraborty, Phys. Rev. E **79**, 061301 (2009).
[19] G. Lois, J. Zhang, T. S. Majmudar, S. Henkes, B. Chakraborty, C. S. OHern, R. P. Behringer, Phys. Rev. E **80**, 060303 (2009).
[20] B. P. Tighe, *Force distributions and stress response in granular materials*, PhD thesis (2006).
[21] J. D. Stevenson, V. Rühle, S. Martiniani, K. J. Schrenk, https://github.com/pele-python/pele
[22] S. Martiniani, J. D. Stevenson, K. J. Schrenk, https://github.com/pele-python/mcpele
[23] E. M. Knorr, R. T. Ng, Proceedings of the International Conference on Very Large Data Bases, 392 (1998).
[24] P. J. Rousseeuw, K. Van Driessen, Technometrics **41**, 3 (1999)
[25] M. Hubert, M. Debruyne, Wiley Interdiscip. Rev. Comput. Stat. **2**, 36 (2010).
[26] B. Efron, Ann. Statist. **7**, 1 (1979).
[27] E. Lebigot, https://github.com/lebigot/uncertainties
[28] B. Efron, J. Am. Stat. Assoc. **82**, 397 (1987).
[29] C. Evans, https://github.com/cgevans/scikits-bootstrap
[30] P. J. Steinhardt, D. R. Nelson, M. Ronchetti, Phys. Rev. B **28**, 784 (1983)
[31] D. Vagberg, D. Valdez-Balderas, M. A. Moore., P. Olsson, S. Teitel, Phys. Rev. E **83**, 030303 (2011).
[32] When listing a function's arguments we place parameters that are held constant before the semicolon