Numerical test of the Edwards conjecture shows that all packings become equally probable at jamming

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*

In the late 1980s, Sir Sam Edwards proposed a possible statistical-mechanical framework to describe the properties of disordered granular materials. A key assumption underlying the theory was that all jammed packings are equally likely. In the intervening years it has never been possible to test this bold hypothesis directly. Here we present simulations that provide direct evidence that at the unjamming point, all packings of soft repulsive particles are equally likely, even though generically, jammed packings are not. Typically, jammed granular systems are observed precisely at the unjamming point since grains are not very compressible. Our results therefore support Edwards' original conjecture. We also present evidence that at unjamming the configurational entropy of the system is maximal.

In science, most breakthroughs cannot be derived from known physical laws: they are based on inspired conjectures [1]. Comparison with experiment of the predictions based on such a hypothesis allows us to eliminate conjectures that are clearly wrong. However, there is a distinction between testing the consequences of a conjecture and testing the conjecture itself. A case in point is Edwards' theory of granular media. In the late 1980s, Edwards and Oakeshott [2] proposed that many of the physical properties of granular materials ('powders') could be predicted using a theoretical framework that was based on the assumption that all distinct packings of such a material are equally likely to be observed. The logarithm of the number of such packings was postulated to play the same role as entropy does in Gibbs' statistical-mechanical description of the thermodynamic properties of equilibrium systems. However, statistical-mechanical entropy and granular entropy are very different objects. Until now, the validity of Edwards' hypothesis could not be tested directly – mainly because the number of packings involved is so large that direct enumeration is utterly infeasible – and, as a consequence, the debate about the Edwards hypothesis has focused on its consequences, rather than on its assumptions. Here we present results that show that now, at last, it is possible to test Edwards' hypothesis directly by numerical simulation. Somewhat to our own surprise, we find that the hypothesis appears to be correct precisely at the point where a powder is just at the (un)jamming threshold. However, at higher densities, the hypothesis fails. At the unjamming transition, the configurational entropy of jammed states appears to be at a maximum.

The concept of 'ensembles' plays a key role in equilibrium statistical mechanics, as developed by J. Willard Gibbs, well over a century ago [3]. The crucial assumption that Gibbs made in order to arrive at a tractable theoretical framework to describe the equilibrium properties of gases, liquid and solids was that, at a fixed total energy, every state of the system is equally likely to be observed. The distinction between, say, a liquid at thermal equilibrium and a granular material is that in a liquid, atoms undergo thermal motion whereas in a granular medium (in the absence of outside perturbations) the system is trapped in one of many (very many) local potential energy minima. Gibbsian statistical mechanics cannot be used to describe such a system. The great insight of Edwards was to propose that the collection of all stable packings of a fixed number of particles in a fixed volume might also play the role of an 'ensemble' and that a statistical-mechanics like formalism would result if one assumed that all such packings were equally likely to be observed, once the system had settled into a mechanically stable 'jammed' state. The nature of this ensemble has been the focus of many studies [2, 4–6].

Jamming is ubiquitous and occurs in materials of practical importance, such as foams, colloids and grains when they solidify in the absence of thermal fluctuations. Decompressing such a solid to the point where it can no longer achieve mechanical equilibrium leads to unjamming. Studies of the unjamming transition in systems of particles interacting via soft, repulsive potentials have shown that this transition is characterised by power-law scaling of many physical properties [7–12]. However, both the exact nature of the ensemble of jammed states and the unjamming transition remains unclear.

In this letter, we report a *direct test* of the Edwards conjecture, using a numerical scheme for computing basin volumes of distinct jammed states (energy minima) of N polydisperse, frictionless disks held at a constant packing fraction ϕ. Uniquely, our numerical scheme allows us to compute Ω, the number of distinct jammed states, and the individual probabilities $p_{i\in\{1,\ldots,\Omega\}}$ of each observed packing to occur. Fig. 1a shows a snapshot of a section of the system, consisting of particles with a hard core and a soft shell. We obtain jammed packings by equilibrating a hard disk fluid and inflating the particles instantaneously to obtain the desired soft-disk volume fractions (ϕ), followed by

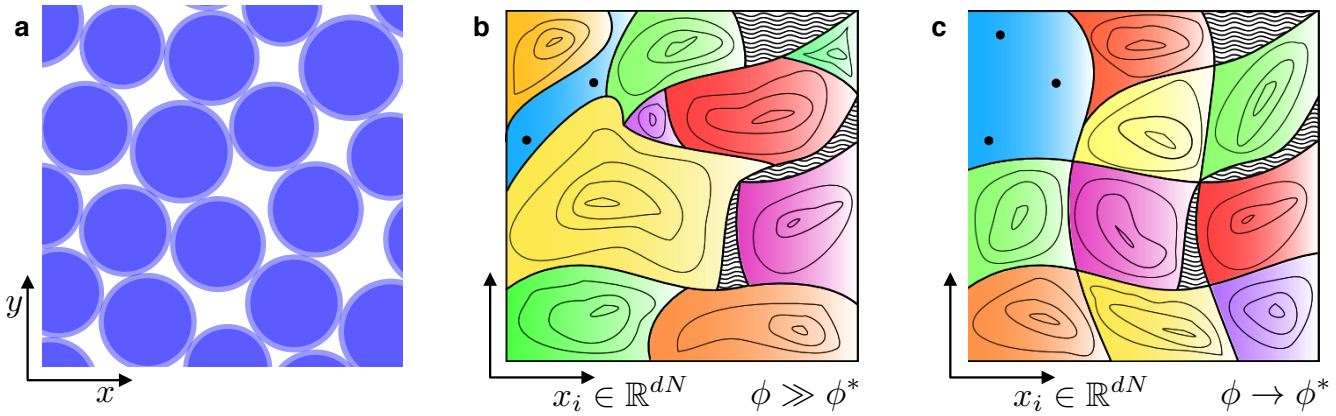


FIG. 1: Snapshot of the system studied and illustration of the associated energy landscape at different volume fractions. (a) Snapshot of a jammed packing of disks with a hard core (dark shaded regions) plus soft repulsive corona (light shaded regions). (b)-(c) Illustration of configurational space for jammed packings. The dashed regions are inaccessible due to hard core overlaps. Single coloured regions with contour lines represent the basins of attraction of distinct minima. The dark blue region with solid dots indicates the coexisting unjammed fluid region and hypothetical marginally stable packings, respectively. The volume occupied by the fluid $V_{\rm unj}$ is significant only for finite size systems at or near unjamming. When $\phi \gg \phi^*$ (b) the distribution of basin volumes is broad but as $\phi \to \phi^*$ (c) the distribution of basin volumes approaches a delta function satisfying Edwards' hypothesis.

energy minimization (see Methods). The minimization procedure finds individual stable packings with a probability p_i proportional to the volume v_i of their basin of attraction. Averages computed using this procedure, represented by $\langle \dots \rangle_{\mathcal{B}}$, would then lead to a bias originating from the different v_i's. Recent advances in numerical methods [13, 15–17] have now enabled direct computation of v_i, and therefore, an unbiased characterization of the phase space. A summary of the technique is provided in Methods.

We report a detailed analysis of the distribution of v_i for a fixed number of disks $N = 64$ (all maximum system sizes in our study were set by the current limits on computing power). We compute v_i using a thermodynamic integration scheme [13, 15–17], and compute the average basin volume $\langle v \rangle(\phi)$. The number of jammed states is, explicitly, $\Omega(\phi) = V_J(\phi)/\langle v \rangle(\phi)$, where $V_J(\phi)$ is the total available phase space volume at a given ϕ. A convenient way to check equiprobability is to compare the Boltzmann entropy $S_B = \ln \Omega - \ln N!$, which counts all packings with the same weight, and the Gibbs entropy $S_G = -\sum_i^{\Omega} p_i \ln p_i - \ln N!$ [18–20]. The Gibbs entropy satisfies $S_G \leq S_B$, saturating the bound when all p_i are equal: $p_{i \in \{1,\dots,\Omega\}} = 1/\Omega$. As shown in Fig. 2a, S_G approaches S_B from below as $\phi \to \phi^{*(S)}_{N=64} \approx 0.82$. Fig. 1b-c schematically illustrates the evolution of the basin volumes as the packing fraction is reduced.

To characterize the distribution of basin volumes, we analyse the statistics of v_i along with the pressure P_i of each packing. It is convenient to study $F_i \equiv -\ln v_i$ as a function of $\Lambda_i \equiv \ln P_i$. As shown in Fig. 2b, we observe a strong correlation between F_i and Λ_i which we quantify by fitting the data to a bivariate Gaussian distribution. The conditional expectation of F given Λ then yields a linear relationship (denoted by solid lines in Fig. 2b) such that $\langle F \rangle_{\mathcal{B}}(\phi; \Lambda) \propto \lambda(\phi)\Lambda$, where $\langle F \rangle_{\mathcal{B}}(\phi; \Lambda)$ represents the average over all basins at a given Λ. Previous studies at higher packing fractions [13] indicate that this relationship is preserved in the thermodynamic limit. Defining $f = F/N$, we have (see Methods for details):

$$\begin{aligned} \langle f \rangle_{\mathcal{B}}(\phi; \Lambda) &= \lambda(\phi)\Lambda + c(\phi) \\ &= \lambda(\phi)\Delta\Lambda + \langle f \rangle_{\mathcal{B}}(\phi) \ , \end{aligned} \tag{1}$$

where $\Delta\Lambda = \Lambda - \langle \Lambda \rangle_{\mathcal{B}}(\phi)$. For Edwards' hypothesis to be valid, we require that in the thermodynamic limit (i) the distribution of volumes approaches a Dirac delta, which follows immediately from the fact that the variance $\sigma_f^2 \sim 1/N$ [16] (see SI) and (ii) F_i needs to be independent of Λ_i, as well as of all other structural observables (see SI), and therefore $\lambda(\phi)$ must necessarily vanish. As can be seen from Fig. 2c-d, within the range of volume fractions studied, $\lambda(\phi)$ decreases but saturates to a minimum as $\phi \to \phi^{*(\lambda)}_{N=64}$. We argue below that the saturation is a finite size effect. An extrapolation using the linear regime in Fig. 2c indicates that $\lambda \to 0$ at packing fraction $\phi^{*(\lambda)}_{N=64} = 0.82 \pm 0.07$,

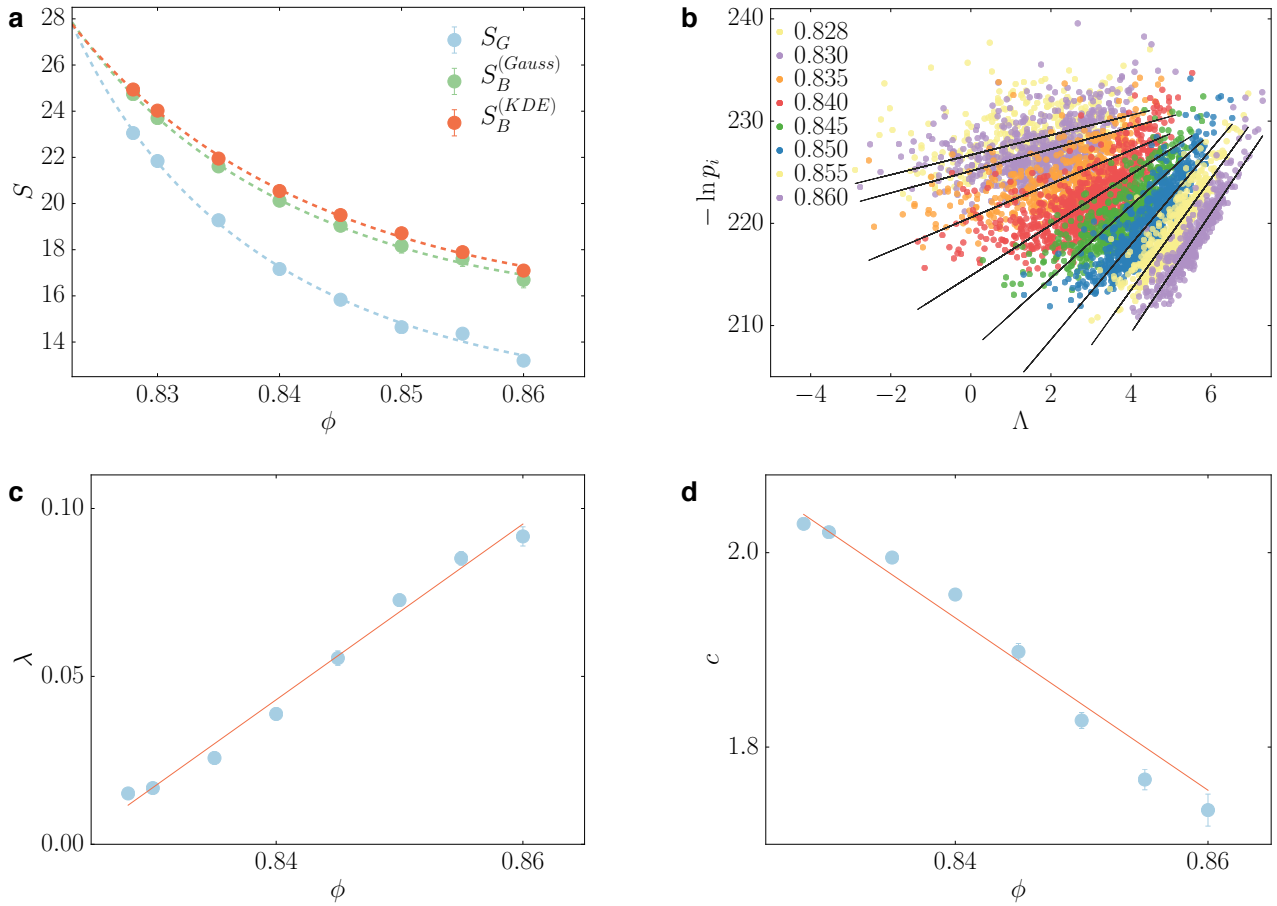


FIG. 2: Numerical results obtained by basin volume calculations for jammed packings of $N = 64$ disks with a hard core and a soft shell. (a) Gibbs entropy S_G and Boltzmann entropy S_B as a function of volume fraction. S_B is computed both parametrically by fitting $\mathcal{B}(f)$ with a generalised Gaussian function ('Gauss') and non-parametrically by computing a Kernel Density Estimate ('KDE') as in Ref. [13]. Dashed curves are a second order polynomial fit. (b) Scatter plot 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 (see Methods) as a function of log-pressure, Λ. Black solid lines are lines of best fit computed by linear minimum mean square error using a robust covariance estimator and bootstrap (see Methods). (c) Slopes $\lambda(\phi)$ and (d) intercepts $c(\phi)$ of linear fits for Eq. 16. Solid lines are lines of best fit and error bars refer to the standard error computed by bootstrap [14].

remarkably close to where our extrapolation yields $S_G = S_B$. The analysis of basin volumes, therefore, strongly suggests that equiprobability is approached only at a characteristic packing fraction and that the vanishing of $\lambda(\phi)$ can be used to estimate the point of equiprobability.

We next show that $\lambda(\phi)$ does indeed tend to zero in the thermodynamic limit. We use the fluctuations σ_f^2, σ_Λ^2, and the covariance $\sigma_{f\Lambda}^2$, obtained from the elements of the covariance matrix $\hat{\sigma} = ((\sigma_f^2, \sigma_{f\Lambda}^2), (\sigma_{f\Lambda}^2, \sigma_\Lambda^2))$ of the joint distribution of f and Λ (see Methods for details), to define λ and c as:

$$\begin{aligned} \lambda(\phi) \equiv & \frac{\sigma_{f\Lambda}^2(\phi)}{\sigma_\Lambda^2(\phi)}, \\ c(\phi) \equiv & \langle f \rangle_{\mathcal{B}}(\phi) - \frac{\sigma_{f\Lambda}^2(\phi)}{\sigma_\Lambda^2(\phi)} \langle \Lambda \rangle_{\mathcal{B}}(\phi). \end{aligned} \tag{2}$$

From Fig. 2b we observe that the decrease of λ is driven by σ_Λ^2 increasing to a maximum, while σ_f^2 and $\sigma_{f\Lambda}^2$ decrease (see Fig. S2 of SI). We expect the main features of these distributions to be preserved as the system size N is increased [13], which suggests that for larger N, where basin volume calculations are still intractable for multiple densities, the

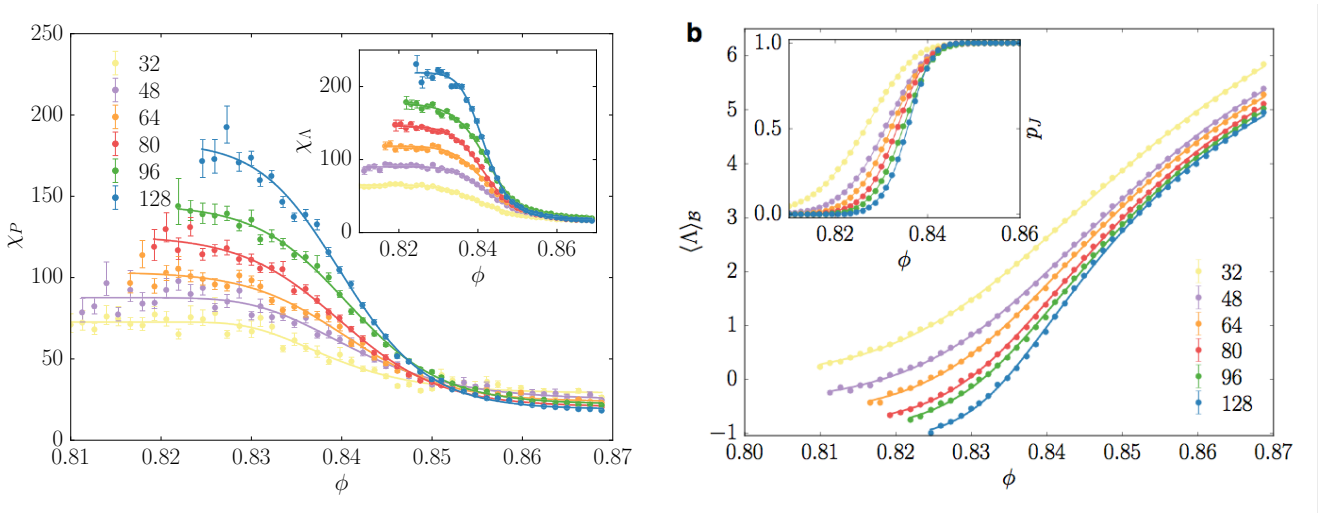


FIG. 3: Finite size scaling analysis. (a) $\chi_P \equiv N\sigma^2(P/\langle P\rangle_{\mathcal{B}})$ and (inset) $\chi_\Lambda \equiv N\sigma^2_\Lambda$, plotted as a function of volume fraction ϕ. By finite size scaling (see SI) we show that the curves diverge in the thermodynamic limit as $\phi \to \phi^{J/*}$, implying $\phi^*_{N\to\infty} = \phi^J_{N\to\infty} = 0.841(3)$, see main text for discussion. For $\phi \gg \phi^J_N$, χ_P approaches a constant value indicating the absence of extensive correlations far from the transition. (b) Observed average log-pressure $\langle\Lambda\rangle_{\mathcal{B}}$ and (inset) probability of obtaining a jammed packings by our protocol, as a function of volume fraction ϕ. By finite size scaling (see SI) we show that $\langle\Lambda\rangle_{\mathcal{B}} \to -\infty$ as $\phi \to \phi^{J_{(\Lambda)}}_{N\to\infty} = 0.841(3)$ and p_J collapses for $\phi \to \phi^{J_{(p_J)}}_{N\to\infty} = 0.844(2)$, thus locating the unjamming point. Error bars, computed by BCa bootstrap [22], refer to 1σ confidence intervals. Solid lines are generalised sigmoid fits of the form $f(\phi) = a - (a-b)/(1+\exp(-w\Delta\phi))^{1/u}$. We only show values of ϕ where the probability of finding a jammed packing is at least 1%, so that the observables are computed over sufficiently large sample sizes.

maximum in σ^2_Λ can be used to identify ϕ^*_N. We have directly measured $\chi_\Lambda = N\sigma^2_\Lambda$ using our sampling scheme – that samples packings with probability proportional to the volume of their basin of attraction – for systems of up to $N = 128$ disks (see inset of Fig. 3a) and finite size scaling indicates that χ_Λ diverges as $\phi \to \phi^{*(\Lambda)}_{N\to\infty} = 0.841(3)$ (see SI). The saturation of λ to a minimum as $\phi \to \phi^*_N$, for small N, is determined by the fact that χ_Λ only diverges in the thermodynamic limit, a detailed discussion is given in Methods.

Interestingly, we find evidence that in the thermodynamic limit, the point of equiprobability $\phi^*_{N\to\infty}$, coincides with the point at which the system unjams, $\phi^J_{N\to\infty}$. We use two characteristics of the unjamming transition to locate $\phi^J_{N\to\infty}$ (i) the average pressure of the packings goes to zero, and therefore $\langle\Lambda\rangle \to -\infty$ (see Fig 3b) and (ii) the probability of finding jammed packings, p_J, goes to zero (see inset of Fig 3b). A scaling analysis indicates that $\langle\Lambda\rangle \to -\infty$ as $\phi^{J_{(\Lambda)}}_{N\to\infty} = 0.841(3)$, and $p_J \to 0$ as $\phi^{J_{(p_J)}}_{N\to\infty} = 0.844(2)$ (see SI). We thus find that $\phi^*_{N\to\infty} = \phi^J_{N\to\infty}$ within numerical error and up to corrections to finite size scaling [21]. Our simulations therefore lead to the surprising conclusion that the Edwards conjecture appears to hold precisely at the (un)jamming transition. We note that our earlier simulations, which were performed at densities much above jamming [13, 16], did not support the equiprobability hypothesis. The earlier simulations were in fact far enough away from unjamming that the emergence of equiprobability at this point could not be anticipated. Our earlier findings, therefore, do not contradict the more recent ones and are completely consistent with these.

Why is χ_Λ related to the unjamming transition? As the particles interact via purely repulsive potentials, the pressure P is strictly positive, which implies that the fluctuations of P have a floor and go to zero at unjamming. The *relative* fluctuations $\chi_P \equiv N\sigma^2\,(P/\langle P\rangle_{\mathcal{B}})$, can be non-zero, and a diverging χ_P would then imply a diverging χ_Λ. Because of the bounded nature of P [23–25], however, χ_P can *only* diverge at the unjamming transition where $\langle P\rangle_{\mathcal{B}} \to 0$ (see Methods). We find that χ_P does diverge (Fig. 3a) and finite size scaling yields $\phi^{*(P)}_{N\to\infty} = 0.841(3)$, in agreement with what has been found for χ_Λ. Returning to the $N = 64$ case that we have analysed using the basin volume statistics, we find that both χ_P and χ_Λ saturate to their maximum values over similar ranges of ϕ and our estimate $\phi^*_{N=64} \approx 0.82$ where $S_G = S_B$ and $\lambda \to 0$, falls in this region. In addition, the average number of contacts $\langle z\rangle_{\mathcal{B}}(\phi^*_{N=64}) = 4.1 \pm 0.2$ is close to the isostatic value $z^{(iso)}_{N=64} \equiv 2d - 2/64 \approx 3.97$ [10] (see SI).

Finally, we note that the states in the generalised Edwards ensemble [5, 23, 26, 27] characterised by ϕ and P have

basin volumes that are similar, if not identical, over the full range of ϕ that we have explored (see scatter plot in Fig. 2b), indicating that equiprobability in the stress-volume ensemble [5, 26] is a more robust formulation of the Edwards hypothesis. This observation is consistent with recent experiments [28].

Although, the equiprobability of jammed states at a given packing fraction was posited by Edwards for jammed packings of hard particles, our analysis shows that for soft particles, the Edwards hypothesis is valid only for the marginally jammed states at $\phi^*_{N\to\infty} = \phi^J_{N\to\infty}$, where the jamming probability vanishes, the entropy is maximised, and relative pressure fluctuations diverge. We have shown not only that there exist a practical 'Edwardsian' packing generation protocol, capable of sampling jammed states equiprobably, but we have uncovered an unexpected property of the energy landscape for this class of systems. At this stage, we cannot establish whether the same considerations are valid in 3D, although the already proven validity of Eq. 16 in 3D would suggest so [13]. The exact value of the entropy at unjamming, whether finite or not, also needs to be elucidated. The implications for 'soft' structural glasses is apparent: at ϕ^J the uniform size of the basins implies that the system, when thermalised, has the same probability of visiting all of its basins of attraction, hence there are no preferred inherent structures. This could be a signature of the hard-sphere transition occurring at the same point [29]. Our approach can, therefore, be extended to spin-glasses and related problems, and it would be clearly very exciting to explore the analogies and differences between 'jamming' in various systems for which the configuration space can break up into many distinct basins.

S.M. acknowledges financial support by the Gates Cambridge Scholarship. K.J.S. acknowledges support by the Swiss National Science Foundation under Grant No. P2EZP2-152188 and No. P300P2-161078. D.F. acknowledges support by EPSRC Programme Grant EP/I001352/1 and EPSRC grant EP/I000844/1. K.R. and B.C. acknowledge the support of NSF-DMR 1409093 and the W. M. Keck Foundation.

S.M, D.F. and B.C. designed the study. S.M. and K.J.S. developed the software and performed the numerical simulations. S.M., K.R., B.C. and D.F. analysed the data and wrote the manuscript.

The data that support the plots within this paper and other findings of this study are available from the corresponding author upon request.

METHODS

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 [13]. 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 [30], to reach the target soft packing fraction $\phi_{\rm SS} > \phi_{\rm HS}$. Each hard sphere is inflated proportionally to its radius, so that the soft sphere radius is given by

$$r_{\rm SS} = \left(\frac{\phi_{\rm SS}}{\phi_{\rm HS}}\right)^{1/d} r_{\rm HS}, \tag{3}$$

where d is the dimensionality of the box (2 in our case), and $r_{\rm SS}$ and $r_{\rm HS}$ 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 [31, 32], 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, $N_{\min} = d(N_{nr} - 1) + 1$ [33], 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 [31].

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 $r_{\rm HS}$, implying a soft core contact distance $r_{\rm SS} = r_{\rm HS}(1+\theta)$, with $\theta = (\phi_{\rm SS}/\phi_{\rm HS})^{1/d} - 1$. We can then write a horizontally shifted hard-sphere plus WCA (HS-WCA) potential as

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

where $\sigma(r_{\rm HS}) = (2\theta + \theta^2) r_{\rm HS}^2 / 2^{1/6}$ guarantees that the potential function and its first derivative go to zero at $r_{\rm SS}$. For computational convenience (avoidance of square-root evaluations), the potential in Eq. 4 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. 4 yields

$$\lim_{r \to r_{\rm SS}} u_{\text{HS-WCA}} = \epsilon \left(\frac{12 r_{\rm SS}}{r_{\rm HS}^2 - r_{\rm SS}^2} \right)^2 (r - r_{\rm SS})^2 + O\left((r - r_{\rm SS})^3 \right), \tag{5}$$

hence, in the limit of no overlap the pair potential is harmonic.

We numerically evaluate this potential, matching the gradient and linearly continuing the function $u_{\text{HS-WCA}}(r)$ for $r \le r_{\rm HS} + \varepsilon$, 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). This inaccessible part of the phase space arises due to hard core constraints and the existence of fluid states (see e.g. [13]). The total phase space volume is equal to $V_{\rm box}^N$. The inaccessible part of this volume arising from the hard core constraints (shown as hatched areas in Fig. 1) is denoted by $V_{\rm HS}$, and the part corresponding to the coexisting unjammed fluid states is denoted by $V_{\rm unj}$ (shown as blue regions with squares in Fig. 1b-c). $V_{\rm unj}$ 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 $V_J = V_{\rm box}^N - V_{\rm HS} - V_{\rm unj}$. In practice we compute V_J using the following equation

$$\ln V_J(N, \phi_{\rm HS}, \phi_{\rm SS}) = N \ln V_{\rm box} - N f_{\rm ex}(\phi_{\rm HS}) + \ln p_J(\phi_{\rm SS}), \tag{6}$$

where $f_{\rm ex}(\phi_{\rm HS})$ 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 [34] as in [13], and $p_J(\phi_{\rm SS})$ is the probability of obtaining a jammed packing at soft volume fraction $\phi_{\rm SS}$ with our protocol, shown in the inset of Fig. 3b.

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

$$p_i = \frac{v_i}{V_J}, \tag{7}$$

where V_J is the total accessible phase space, such that

$$V_J = \sum_{i=1}^{\Omega} v_i. \tag{8}$$

Details of the computation of v_i are discussed in Refs. [13, 17]. To find Ω, we make the simple observation

$$\sum_{i=1}^{\Omega} v_i = \frac{\Omega}{\Omega} \sum_{i=1}^{\Omega} v_i = \Omega \langle v \rangle, \tag{9}$$

from which it follows immediately that

$$\Omega = \frac{V_J}{\langle v \rangle}. \tag{10}$$

The 'Boltzmann-like' entropy, suggested in a similar form by Edwards [2], is then

$$S_B = \ln \Omega - \ln N! \tag{11}$$

where the $\ln N!$ correction ensures that two systems in identical macrostates are in equilibrium under exchange of particles [18–20].

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 ($F = -\ln v_{\rm basin}$). Since the configurations were sampled proportional to the volume of their basin of attraction, we can compute the unbiased distribution as

$$\mathcal{U}(\phi; F) = \mathcal{Q}(\phi)\mathcal{B}(\phi; F)e^F \tag{12}$$

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

$$\mathcal{Q}(\phi) = \left[\int_{F_{\rm min}}^{\infty} {\rm d}F \mathcal{B}(\phi; F) e^F \right]^{-1} = \langle v \rangle(\phi). \tag{13}$$

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. 12 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. [13]. In this work we follow Ref. [13] and fit $\mathcal{B}(\phi; F)$ using both a (parametric) generalised Gaussian model [35], see Eq. S21 of the SI, and a (non-parametric) kernel density estimate (KDE) with Gaussian kernels [36, 37] and bandwidth selection performed by cross validation [13, 38], yielding consistent results in agreement with Ref. [13]. 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 Methods.

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

$$S_G = -\sum_{i=1}^{\Omega} p_i \ln p_i - \ln N! = \sum_{i=1}^{\Omega} [p_i(-\ln v_i)] + \ln V_J - \ln N! = \langle F \rangle_{\mathcal{B}} + \ln V_J - \ln N! \tag{14}$$

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 [13]. 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 of the SI). We therefore expect their joint probability to be well-approximated by a bivariate Gaussian distribution $\mathcal{B}(\phi; F, \Lambda) = \mathcal{N}(\mu, \hat{\sigma})$ [48], with mean $\mu = (\mu_F, \mu_\Lambda)$ and covariance matrix $\hat{\sigma} = ((\sigma_F^2, \sigma_{F\Lambda}^2), (\sigma_{F\Lambda}^2, \sigma_\Lambda^2))$ [39]. This is consistent with the elliptical distribution of points in Fig. 2b. 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 [39]

$$\langle F \rangle_{\mathcal{B}(\mathbf{\Gamma})}(\phi; \Lambda) \equiv \mathbb{E}[F|\phi; \Lambda] = \frac{\sigma_{F\Lambda}^2(\phi)}{\sigma_\Lambda^2(\phi)}(\Lambda - \mu_\Lambda(\phi)) + \mu_F(\phi). \tag{15}$$

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 $\langle F \rangle_{\mathcal{B}(\mathbf{\Gamma})}(\phi; \Lambda) = -\langle \ln v \rangle_{\mathcal{B}(\mathbf{\Gamma})}(\phi; \Lambda) \geq -\ln \langle v \rangle_{\mathcal{B}(\mathbf{\Gamma})}(\phi; \Lambda)$ [40] 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 [13] as

$$\begin{aligned} \langle f \rangle_{\mathcal{B}}(\phi; \Lambda) =& \lambda(\phi)\Lambda + c(\phi) \\ =& \frac{\sigma_{f\Lambda}^2(\phi)}{\sigma_\Lambda^2(\phi)}\Lambda - \frac{\sigma_{f\Lambda}^2(\phi)}{\sigma_\Lambda^2(\phi)}\mu_\Lambda(\phi) + \mu_f(\phi) \\ =& \frac{\sigma_{f\Lambda}^2(\phi)}{\sigma_\Lambda^2(\phi)}\Delta\Lambda + \mu_f(\phi) \end{aligned} \tag{16}$$

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. 16 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

$$\lambda(\phi) \equiv \frac{\sigma_{f\Lambda}^2(\phi)}{\sigma_\Lambda^2(\phi)} \tag{17}$$

$$c(\phi) \equiv \mu_f(\phi) - \frac{\sigma_{f\Lambda}^2(\phi)}{\sigma_\Lambda^2(\phi)}\mu_\Lambda(\phi) \tag{18}$$

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. 4. 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 [41] and MCPELE [42].

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 [43]. 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 [43]). 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 [37, 44] 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 [45]. We use these robust estimates of the location and of the covariance matrix (computed over 1000 bootstrap samples [14]) to fit our observations by linear MMSE [39], see Fig. 2.

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% [37]. 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 [35] and a (non-parametric) kernel density estimate (KDE) with Gaussian kernels [36, 37] and bandwidth selection performed by cross validation [13, 38].

Error analysis

Errors were computed analytically where possible and propagated using the 'uncertainties' Python package [46]. Alternatively, intervals of confidence were computed by bootstrap for the covariance estimation [14] and by BCa bootstrap otherwise using the 'scikit-bootstrap' Python package [22, 47].

* sm958@cam.ac.uk

[1] R. Laughlin, *A Different Universe: Reinventing Physics from the Bottom Down*, Basic Books, 2006.
[2] S. F. Edwards, R. B. S. Oakeshott, "Theory of powders.", Physica A **157**, 1080-1090 (1989).
[3] J. W. Gibbs, *Elementary Principles of Statistical Mechanics*, Charles Scribner's sons, New York, Edward Arnold, London (1902).
[4] S. F. Edwards, D. V. Grinev. "Granular materials: towards the statistical mechanics of jammed configurations.", Adv. Phys. **51**, 1669-1684 (2002).
[5] D. Bi, S. Henkes, K. E. Daniels, B. Chakraborty, "The statistical physics of athermal materials.", Annu. Rev. Condens. Matter Phys. **6**, 63-83 (2015).
[6] A. Baule, F. Morone, H. Hermann, H. A. Makse, "Edwards Statistical Mechanics for Jammed Granular Matter.", arXiv:1602.04369 (2016).
[7] P. Olsson, S. Teitel, "Critical scaling of shear viscosity at the jamming transition.", Phys. Rev. Lett. **99**, 178001 (2007).
[8] M. Wyart, S. R. Nagel, T. A. Witten, "Geometric origin of excess low-frequency vibrational modes in weakly connected amorphous solids.", EPL **72**, 486-492 (2005).

[9] L. E. Silbert, A. J. Liu, S. R. Nagel, "Structural signatures of the unjamming transition at zero temperature.", Phys. Rev. E **73**, 041304 (2006).

[10] C. Goodrich, A. J. Liu, J. P. Sethna, "Scaling ansatz for the jamming transition.", Proc. Natl. Acad. Sci. **113**, 9745-9750 (2016).

[11] K. Ramola, B. Chakraborty, "Disordered contact networks in jammed packings of frictionless disks.", J. Stat. Mech. 114002 (2016).

[12] K. Ramola, B. Chakraborty, "Scaling Theory for the Frictionless Unjamming Transition.", Phys. Rev. Lett. **118**, 138001 (2017).

[13] S. Martiniani, K. J. Schrenk, J. D. Stevenson, D. J. Wales, D. Frenkel, "Turning intractable counting into sampling: Computing the configurational entropy of three-dimensional jammed packings.", Phys. Rev. E **93**, 012906 (2016).

[14] B. Efron, "Bootstrap methods: another look at the jackknife.", Ann. Statist. **7**, 1-26 (1979).

[15] N. Xu, D. Frenkel, A. J. Liu, "Direct determination of the size of basins of attraction of jammed solids.", Phys. Rev. Lett. **106**, 245502 (2011).

[16] D. Asenjo, F. Paillusson, D. Frenkel, "Numerical calculation of granular entropy.", Phys. Rev. Lett. **112**, 098002 (2014).

[17] S. Martiniani, K. J. Schrenk, J. D. Stevenson, D. J. Wales, D. Frenkel, "Structural analysis of high-dimensional basins of attraction.", Phys. Rev. E **94**, 031301 (2016).

[18] R. H. Swendsen, "Statistical mechanics of colloids and Boltzmanns definition of the entropy.", Am. J. Phys. **74**, 187-190 (2006).

[19] D. Frenkel, "Why colloidal systems can be described by statistical mechanics: some not very original comments on the Gibbs paradox.", Mol. Phys. **112**, 2325-2329 (2014).

[20] M. E. Cates, V. N. Manoharan, "Testing the foundations of classical entropy: colloid experiments.", Soft Matter **11**, 6538-6546 (2015).

[21] D. Vagberg, D. Valdez-Balderas, M. A. Moore, P. Olsson, S. Teitel, "Finite-size scaling at the jamming transition: Corrections to scaling and the correlation-length critical exponent.", Phys. Rev. E **83**, 030303 (2011).

[22] B. Efron, "Better bootstrap confidence intervals.", J. Am. Stat. Assoc. **82**, 171-185 (1987).

[23] S. Henkes, B. Chakraborty, "Statistical mechanics framework for static granular matter.", Phys. Rev. E **79**, 061301 (2009).

[24] G. Lois, J. Zhang, T. S. Majmudar, S. Henkes, B. Chakraborty, C. S. OHern, R. P. Behringer, "Stress correlations in granular materials: An entropic formulation.", Phys. Rev. E **80**, 060303 (2009).

[25] B. P. Tighe, *Force distributions and stress response in granular materials*, PhD thesis (2006).

[26] R. Blumenfeld, S. F. Edwards, "On granular stress statistics: Compactivity, angoricity, and some open issues.", J. Phys. Chem. B **113**, 3981-3987 (2009).

[27] S. Henkes, C. S. O'Hern, B. Chakraborty, "Entropy and temperature of a static granular assembly: an ab initio approach.", Phys. Rev. Lett. **99**, 038002 (2007).

[28] J. G. Puckett, K. E. Daniels, "Equilibrating temperaturelike variables in jammed granular subsystems", Phys. Rev. Lett. **110**, 058001 (2013).

[29] P. Charbonneau, J. Kurchan, G. Parisi, P. Urbani, F. Zamponi, "Glass and Jamming Transitions: From Exact Results to Finite-Dimensional Descriptions.", Annu. Rev. Condens. Matter Phys. **8**, 265288 (2017).

[30] J. D. Weeks, D. Chandler, H. C. Andersen, "Role of repulsive forces in determining the equilibrium structure of simple liquids.", J. Chem. Phys. **54**, 5237-5247 (1971).

[31] E. Bitzek, P. Koskinen, F. Ghler, M. Moseler, P. Gumbsch, "Structural relaxation made simple.", Phys. Rev. Lett. **97**, 170201 (2006).

[32] D. Asenjo, J. D. Stevenson, D. J. Wales, D. Frenkel, "Visualizing basins of attraction for different minimization algorithms.", J. Phys. Chem. B **117**, 12717-12723 (2013).

[33] C. P. Goodrich, A. J. Liu, S. R. Nagel, "Finite-size scaling at the jamming transition.", Phys. Rev. Lett. **109**, 095704 (2012).

[34] A. Santos, S. B. Yuste, M. Lopez De Haro, "Equation of state of a multicomponent d-dimensional hard-sphere fluid.", Mol. Phys. **96**, 1-5 (1999).

[35] S. Nadarajah, "A generalized normal distribution.", J. Appl. Stat. **32**, 685-694 (2005).

[36] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, New York, 2009.

[37] Pedregosa et al., "Scikit-learn: Machine learning in Python.", J. Mach. Lern. Res. **12**, 2825-2830 (2011).

[38] A. W. Bowman, "An alternative method of cross-validation for the smoothing of density estimates.", Biometrika **71**, 353-360 (1984).

[39] D. P. Bertsekas, J. N. Tsitsiklis, *Introduction to probability, Vol. 1.* Belmont, MA: Athena Scientific, 2002.

[40] E. W. Weisstein, "Jensen's Inequality" From MathWorld – A Wolfram Web Resource.

[41] J. D. Stevenson *et al.*, https://github.com/pele-python/pele

[42] S. Martiniani *et al.*, https://github.com/pele-python/mcpele

[43] E. M. Knorr, R. T. Ng, "An alternative method of cross-validation for the smoothing of density estimates.", Proceedings of the International Conference on Very Large Data Bases, 392-403 (1998).

[44] P. J. Rousseeuw, K. Van Driessen, "A fast algorithm for the minimum covariance determinant estimator.", Technometrics **41**, 212-223 (1999)

[45] M. Hubert, M. Debruyne, "Minimum covariance determinant.", Wiley Interdiscip. Rev. Comput. Stat. **2**, 36-43 (2010).

[46] E. Lebigot, https://github.com/lebigot/uncertainties

[47] C. Evans, https://github.com/cgevans/scikits-bootstrap

[48] When listing a function's arguments we place parameters that are held constant before the semicolon

Supplementary Information: Numerical test of the Edwards conjecture shows that all packings become equally probable at jamming

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*

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

$$\ln\left(\frac{P_i}{\langle P\rangle}\right) = \ln P_i - \ln\langle P\rangle = \Lambda_i - \ln\langle P\rangle. \tag{S1}$$

Using Jensen's inequality [1], we have the following bound for the first moment of the log-pressures

$$\langle\Lambda\rangle = \langle\ln P\rangle \leq \ln\langle P\rangle. \tag{S2}$$

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

$$\ln P = \ln\langle P\rangle + \left.\frac{d\ln P}{dP}\right|_{P=\langle P\rangle}(P - \langle P\rangle) + ... \tag{S3}$$

We next compute the moments to leading order

$$\begin{aligned}\langle\ln P\rangle &\approx \ln\langle P\rangle,\\ \sigma^2(\ln P) &\approx \frac{\sigma^2(P)}{\langle P\rangle^2} = \sigma^2\left(\frac{P}{\langle P\rangle}\right).\end{aligned} \tag{S4}$$

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

$$0 < P_{\min} \leq P_i \leq P_{\max}. \tag{S5}$$

where $P_{\max}$ and $P_{\min}$ are determined by the packing fraction ϕ, independent of system size, and are physically set limits [2–4]. We can therefore use Popoviciu's inequality on variances, yielding

$$\sigma^2(P) \leq \frac{1}{4}\left(P_{\max} - P_{\min}\right)^2, \tag{S6}$$

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

$$\sigma^2\left(\frac{P}{\langle P\rangle}\right) \leq \frac{1}{4}\frac{\left(P_{\max} - P_{\min}\right)^2}{\langle P\rangle^2}. \tag{S7}$$

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 $p_i = \sum_j \sum_{\alpha=1,2} \mathbf{r}_{i,j}^{\alpha} \mathbf{f}_{i,j}^{\alpha}$ where $\mathbf{r}_{i,j}^{\alpha}$ and $\mathbf{f}_{i,j}^{\alpha}$ are the contact vectors and contact forces, respectively. The variance of $P/\langle P \rangle$ is

$$\sigma^2(P/\langle P \rangle) = \frac{\langle P^2 \rangle - \langle P \rangle^2}{\langle P \rangle^2} = \frac{\sum_{i=1}^{N} \sigma^2(p_i) + \sum_{i \neq j} \mathrm{cov}(p_i, p_j)}{\left(\sum_{i=1}^{N} \langle p_i \rangle\right)^2}. \tag{S8}$$

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

$$\sigma^2(P/\langle P \rangle) = \frac{\sum_{i=1}^{N} \sigma^2(p_i)}{\left(\sum_{i=1}^{N} \langle p_i \rangle\right)^2} \sim \frac{1}{N} \tag{S9}$$

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

$$\sigma^2(\Lambda) \approx \sigma^2(P/\langle P \rangle) \sim 1/L^2, \tag{S10}$$

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

$$\mathrm{cov}(F/N, P/\langle P \rangle) = \frac{\langle FP \rangle - \langle F \rangle \langle P \rangle}{N \langle P \rangle} = \frac{\sum_{i=1}^{N} \mathrm{cov}(f_i, p_i) + \sum_{i \neq j} \mathrm{cov}(f_i, p_j)}{N \sum_{i=1}^{N} \langle p_i \rangle}. \tag{S11}$$

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

$$\mathrm{cov}(F/N, P/\langle P \rangle) = \frac{\sum_{i=1}^{N} \mathrm{cov}(f_i, p_i)}{N \sum_{i=1}^{N} \langle p_i \rangle} \sim \frac{1}{N}, \tag{S12}$$

therefore

$$\mathrm{cov}(f, \Lambda) \approx \mathrm{cov}(f, P/\langle P \rangle) \sim 1/L^2, \tag{S13}$$

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

$$\lambda_{\phi \gg \phi^*} = \frac{\sigma^2(f, \Lambda)}{\sigma^2(\Lambda)} \sim \mathcal{O}(1), \tag{S14}$$

In other words, λ is independent of system size, a fact that has been verified numerically in Ref. [5].

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

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

$$\sigma^2(\Lambda) \sim L^{\gamma/\nu - 2}, \tag{S15}$$

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. S2e. Since we expect this to be generic, we do not expect $L^2\sigma^2(f, \Lambda)$ to diverge but rather that

$$\sigma^2(f, \Lambda) \lesssim L^{-2}, \tag{S16}$$

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

$$\lambda_{\phi \to \phi^*} = \frac{\sigma^2(f, \Lambda)}{\sigma^2(\Lambda)} = 0. \tag{S17}$$

Relation between scaling exponents

Starting from Eq. 1 of 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

$$\sigma_f^2 = \lambda^2\sigma_\Lambda^2 = (\sigma_{f\Lambda}^2)^2/\sigma_\Lambda^2 \tag{S18}$$

By rearranging this expressions we find that

$$(\sigma_{f\Lambda}^2)^2/\sigma_f^2 = \sigma_\Lambda^2 \sim \begin{cases} L^{-2} \text{ for } \phi \gg \phi^* \\ L^{-\zeta} \text{ for } \phi \to \phi^* \end{cases} \tag{S19}$$

where we have defined $\zeta \equiv 2 - \gamma/\nu \approx 1$ as in Eq. (S15) 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

$$2\eta - \vartheta = \zeta \tag{S20}$$

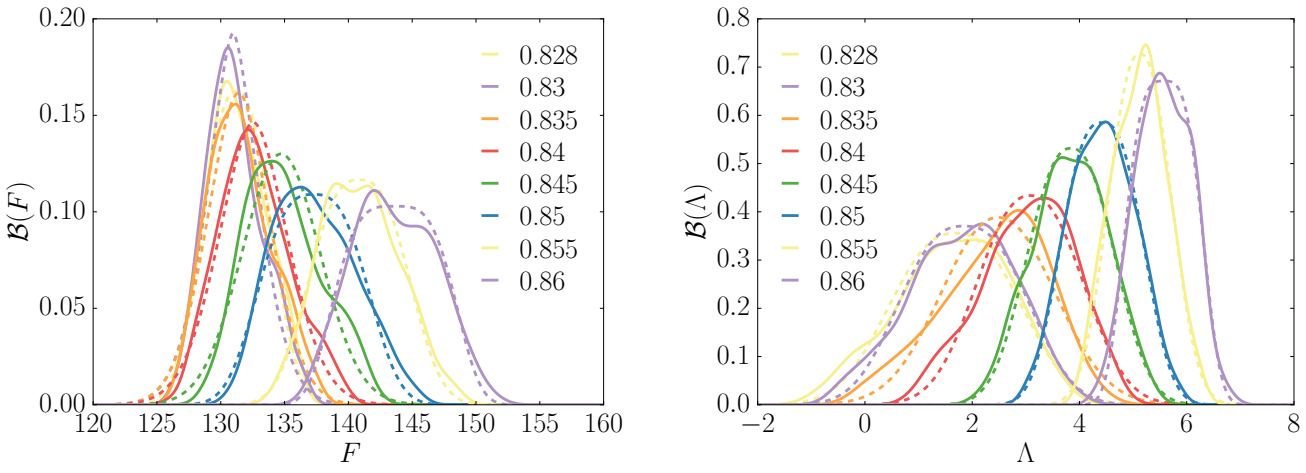
DISTRIBUTION OF BASIN LOG-VOLUMES

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

$$\mathcal{B}(F|\mu_F, \sigma_F, \zeta) \equiv \frac{\zeta}{2\sigma\Gamma(1/\zeta)} \exp\left[-\left(\frac{|F-\mu_F|}{\sigma_F}\right)^\zeta\right], \tag{S21}$$

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. [6] 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

$$e^{-(F-\mu_F)^2/(2\sigma^2)} \sim e^{-N(f-\mu_f)^2} \to \delta(\mu_f) \text{ as } N \to \infty, \tag{S22}$$



(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.

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 $\phi \to \phi^*$.

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 of the main text. 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	$S_B^{(Gauss)}$	$S_B^{(KDE)}$
$\phi^*_{N=64}$	0.82 ± 0.07	0.82	0.82
$\langle z \rangle_{\rm sig}(\phi^*_{N=64})$	4.1 ± 0.2	4.0	4.0

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. S23) of the data in Fig. S3.

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

$$f(a, b, \phi_0, u, w; \phi) = a - \frac{a - b}{(1 + e^{-w(\phi - \phi_0)})^{1/u}} \tag{S23}$$

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 the main text and the finite size scaling section for further discussion).

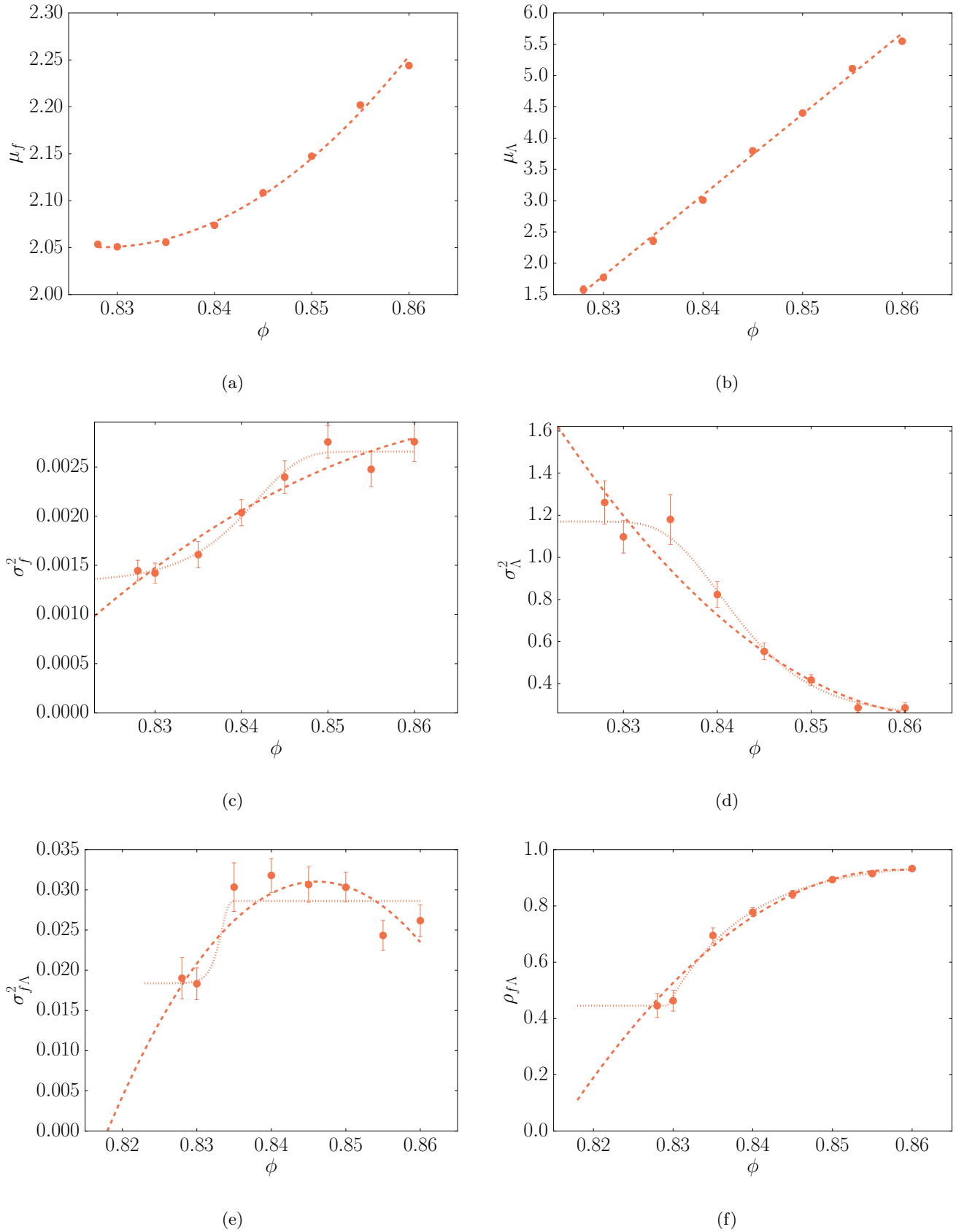


FIG. S2: Moments of the joint distribution $\mathcal{B}(f, \Lambda)$ for jammed packings of $N = 64$ HS-WCA polydisperse disks at various volume fractions $0.828 \leq \phi \leq 0.86$. Elements of the mean $\mu = (\mu_f, \mu_\Lambda)$ are shown in (a) and (b) respectively. Elements of the covariance matrix $\hat{\sigma} = ((\sigma_f^2, \sigma_{f\Lambda}^2), (\sigma_{f\Lambda}^2, \sigma_\Lambda^2))$ 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. S23). Curves of best fit are meant as guide to the eye.

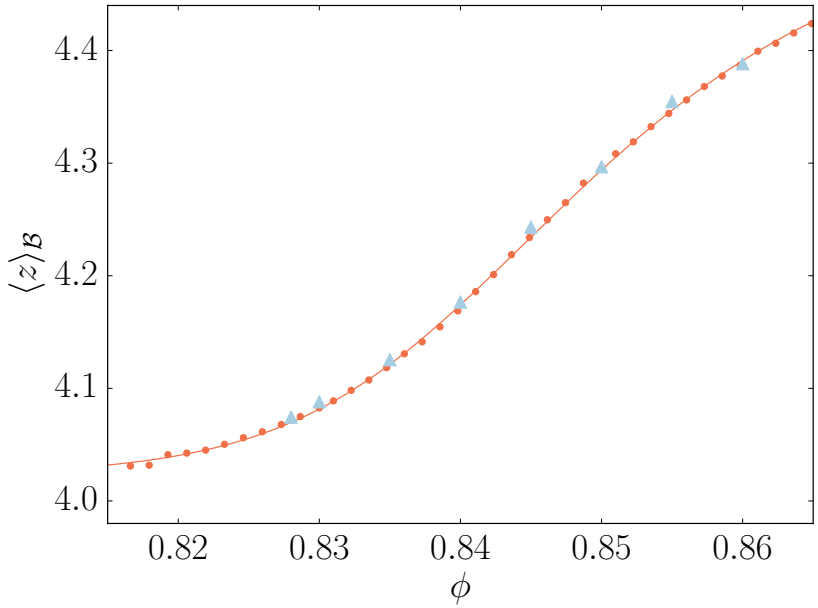


FIG. S3: Observed average contact number for jammed packings of $N = 64$ HS-WCA polydisperse disks. Triangles are estimates from the basin volume measurements datasets. Circles are estimates from the independent measurements used for the finite size scaling analysis. The solid line is a generalised sigmoid fit (Eq. S23) of the latter.

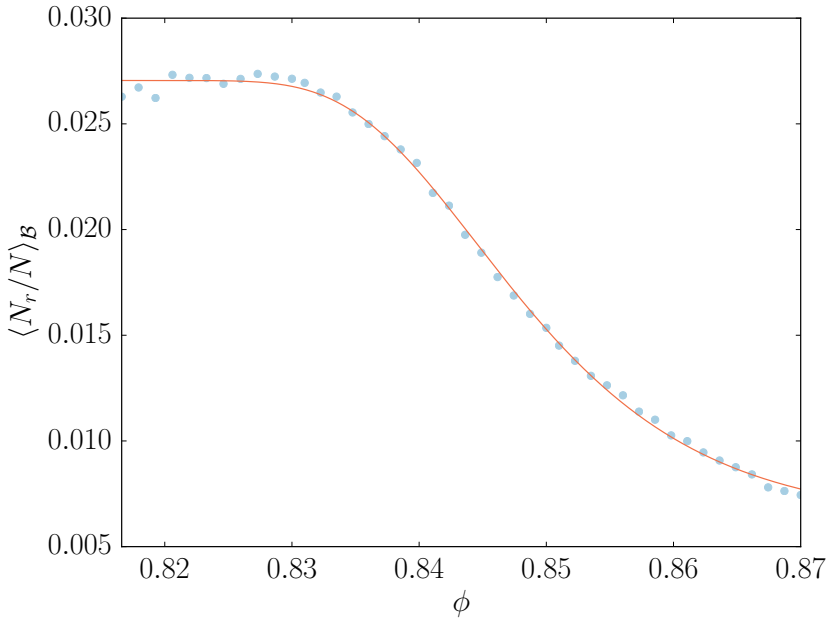


FIG. S4: Fraction of rattlers for jammed packings of $N = 64$ HS-WCA polydisperse disks. We find this tends to a maximum and saturates over the same range of packing fractions where relative pressure fluctuations are maximal, see Fig. 3 of the main text. The solid line is a generalised sigmoid fit, Eq. S23.

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

$$\langle f \rangle_{\mathcal{B}}(\phi; X) = \lambda_X(\phi) \ln X + c_X(\phi). \tag{S24}$$

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 [7]. 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 of the main text, 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 such that the conditional expectation of F given Λ corresponds 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. 17 (Methods).

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 collapsSI 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. [8]. 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 [8]. 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.

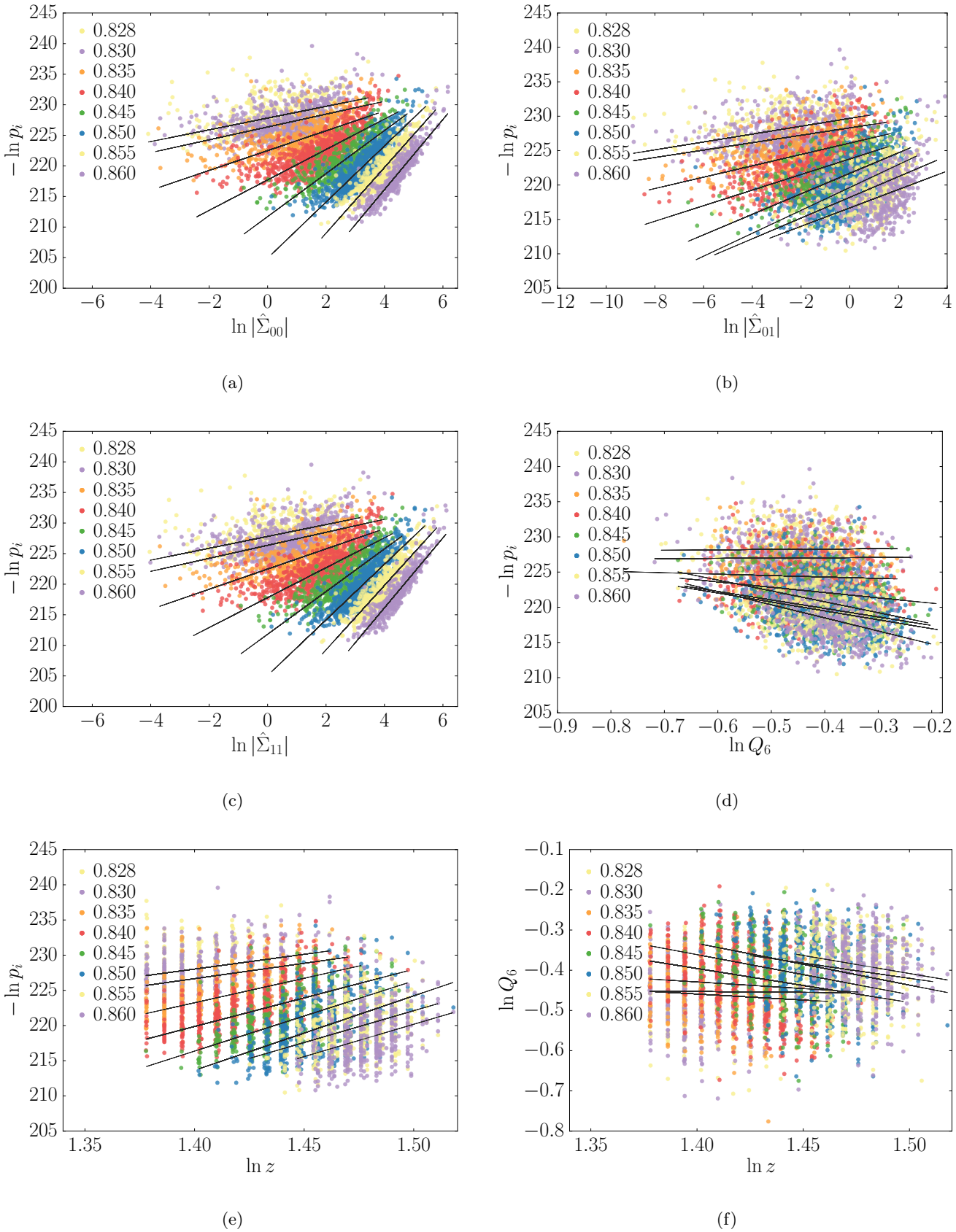
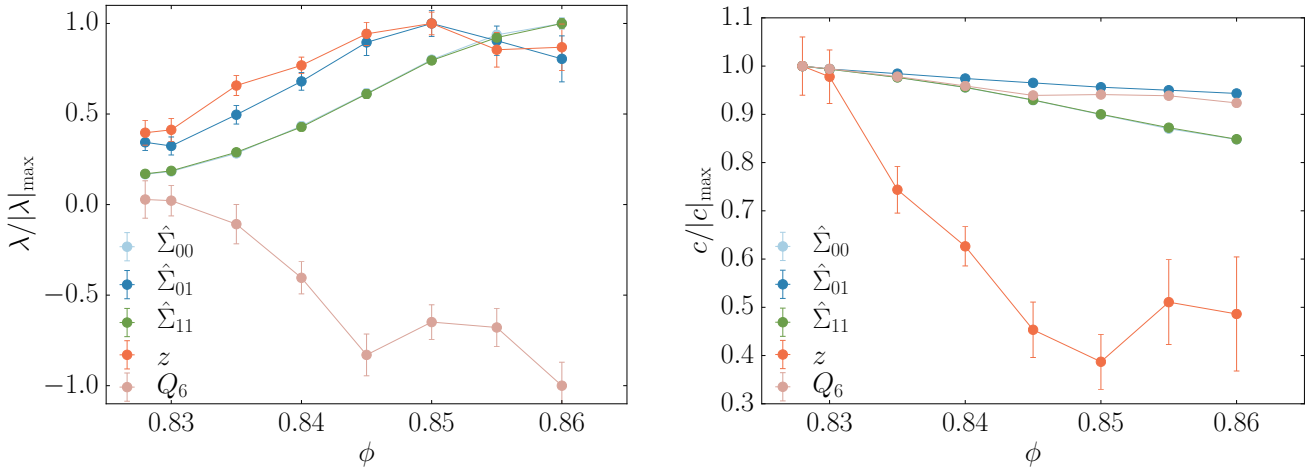
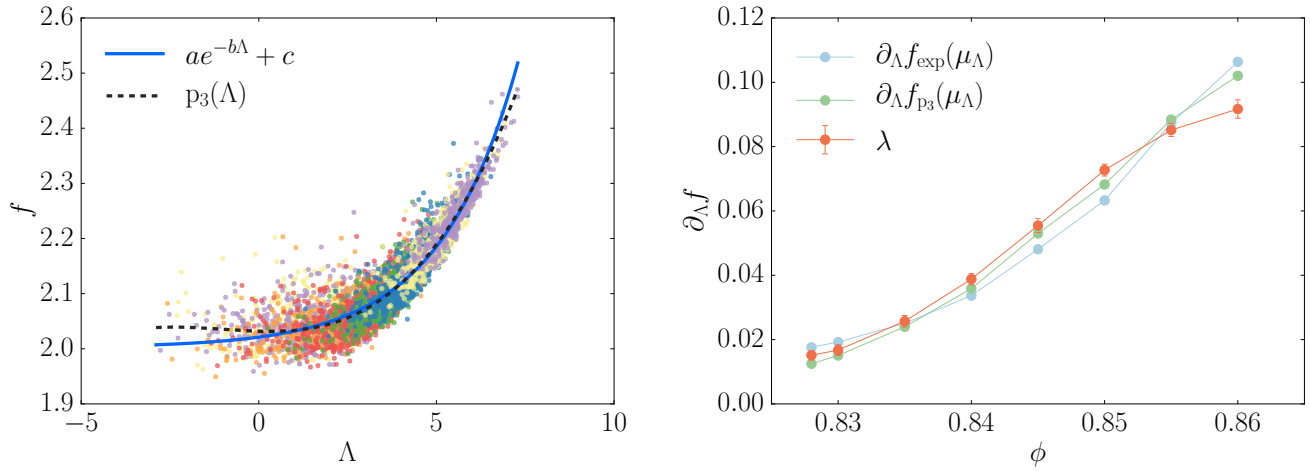


FIG. S5: Scatter plots of the negative log-probability of observing a packing, $-\ln p_i = F_i + \ln V_J(\phi)$, where V_J is the accessible fraction of phase space, as a function of the individual terms of the stress tensor $\hat{\mathbf{\Sigma}}$ (a)-(c), the Q_6 bond-orientational order parameter (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.



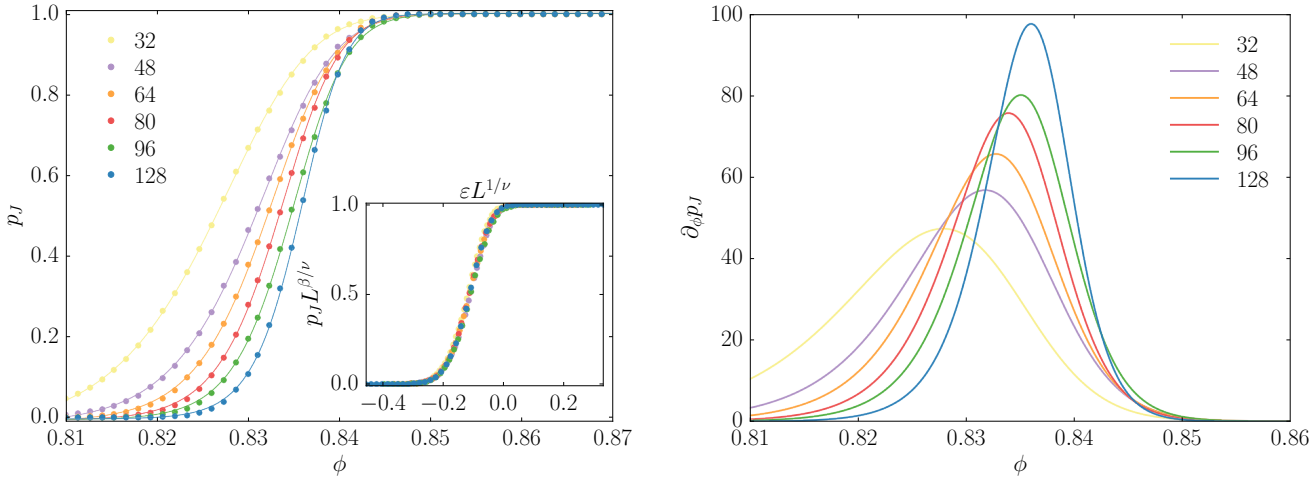
(a) (b)

FIG. S6: Slopes λ_X (a) and intercepts c_X (b) of Eq. S24 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. 17. Solid lines are guide to the eye.



(a) (b)

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. S23. (b) Derivative of the sigmoid fits for p_J for different numbers of disks.

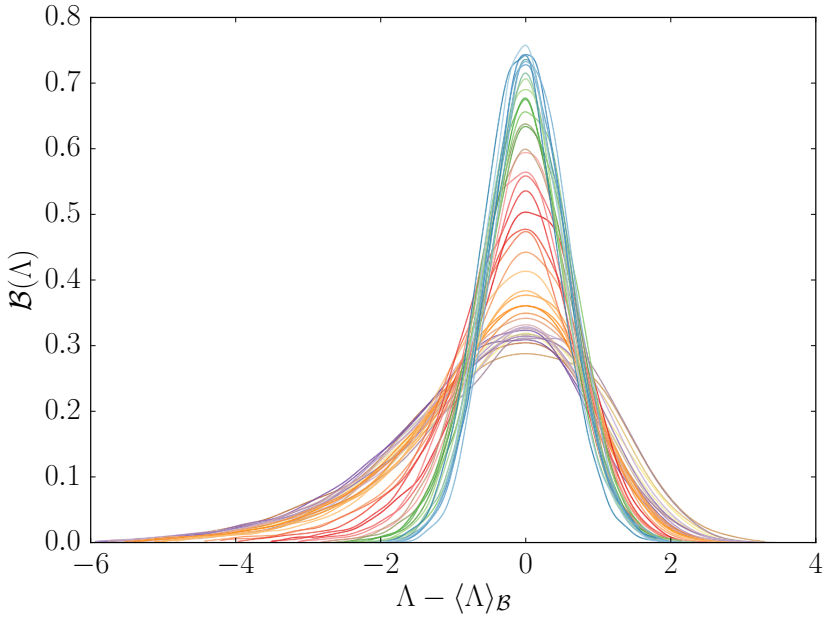
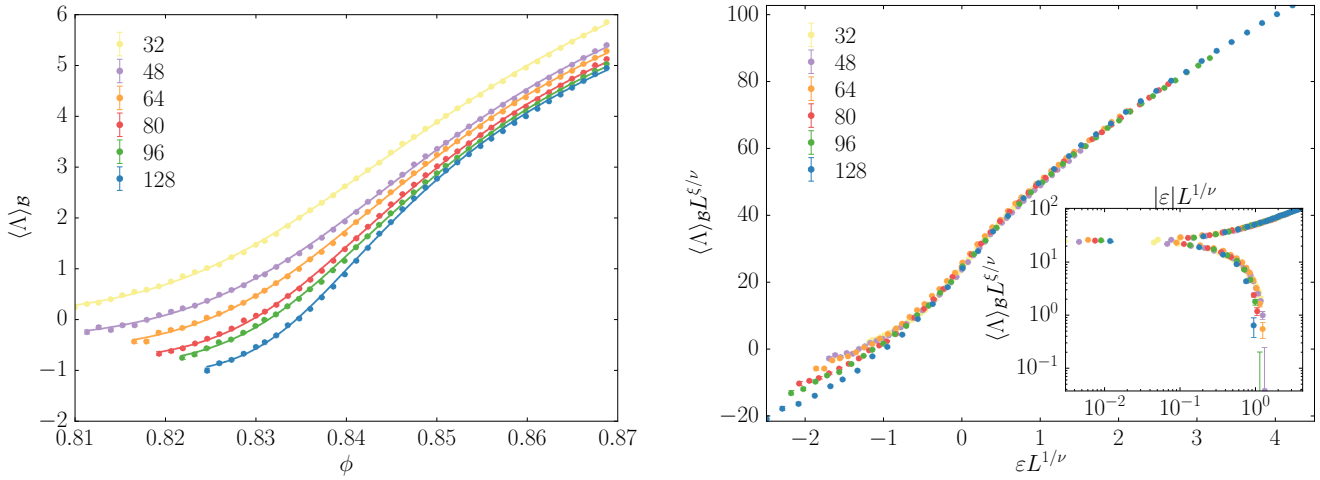
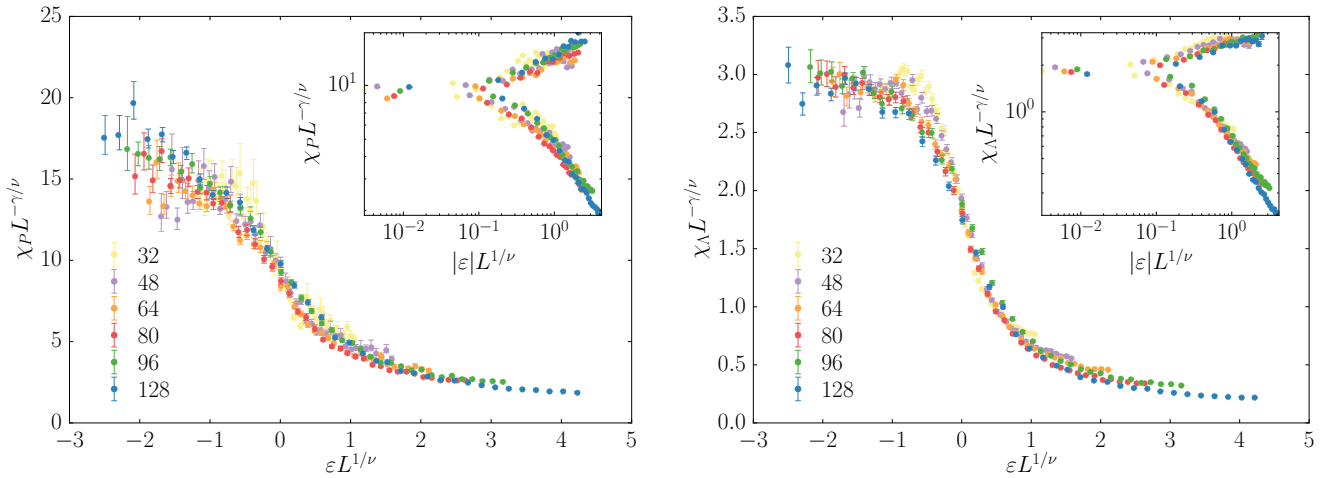


FIG. S9: Log-transformed observed (biased) distribution of pressures for jammed packings of $N = 64$ HS-WCA polydisperse disks, centred around the mean. The variance grows for decreasing volume fractions and becomes more skewed towards low pressures. The overall Gaussian shape is consistent with a log-normal distribution of pressures. Curves are kernel density estimates with Gaussian kernels [9, 10] and bandwidth selection performed by cross validation [5, 11]



(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. S23. Error bars, computed by BCa bootstrap [12], 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] E. W. Weisstein, "Jensen's Inequality" From MathWorld – A Wolfram Web Resource.
[2] S. Henkes, B. Chakraborty, "Statistical mechanics framework for static granular matter.", Phys. Rev. E **79**, 061301 (2009).
[3] G. Lois, J. Zhang, T. S. Majmudar, S. Henkes, B. Chakraborty, C. S. OHern, R. P. Behringer, "Stress correlations in granular materials: An entropic formulation.", Phys. Rev. E **80**, 060303 (2009).
[4] B. P. Tighe, *Force distributions and stress response in granular materials*, PhD thesis (2006).
[5] S. Martiniani, K. J. Schrenk, J. D. Stevenson, D. J. Wales, D. Frenkel, "Turning intractable counting into sampling: Computing the configurational entropy of three-dimensional jammed packings.", Phys. Rev. E **93**, 012906 (2016).
[6] D. Asenjo, J. D. Stevenson, D. J. Wales, D. Frenkel, "Visualizing basins of attraction for different minimization algorithms.", J. Phys. Chem. B **117**, 12717-12723 (2013).
[7] P. J. Steinhardt, D. R. Nelson, M. Ronchetti, "Bond-orientational order in liquids and glasses.", Phys. Rev. B **28**, 784-805 (1983).
[8] D. Vagberg, D. Valdez-Balderas, M. A. Moore, P. Olsson, S. Teitel, "Finite-size scaling at the jamming transition: Corrections to scaling and the correlation-length critical exponent.", Phys. Rev. E **83**, 030303 (2011).
[9] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, New York, 2009.
[10] Pedregosa et al., "Scikit-learn: Machine learning in Python.", J. Mach. Lern. Res. **12**, 2825-2830 (2011).
[11] A. W. Bowman, "An alternative method of cross-validation for the smoothing of density estimates.", Biometrika **71**, 353-360 (1984).
[12] B. Efron, "Better bootstrap confidence intervals.", J. Am. Stat. Assoc. **82**, 171-185 (1987).