Supplementary Information

1 Numerical Method

In this section, we present an algorithm to address the following optimization problem

$$\begin{aligned} &\text{miminize } \sum_i (\boldsymbol{r}_i - \boldsymbol{\zeta}_i)^2, \\ &\text{subject to } -(\boldsymbol{r}_i - \boldsymbol{r}_j)^2 + \sigma^2 \le 0 \text{ for all } i < j. \end{aligned} \tag{S1}$$

for a range of packing fractions ϕ. We establish our system within a box of size L, subject to periodic boundary conditions (PBC). To uphold these conditions, we ensure the center of mass coincides with the box center, i.e., $\sum_i \boldsymbol{r}_i = 0$. Our proposed algorithm unfolds in several steps:

1. Begin with $\phi = \phi_m$, where the packing configurations represent maximum close packing – either FCC or HCP for 3D, and Hexagonal packing for 2D. Generate N quenched points $\boldsymbol{\zeta}$ randomly within the box. Construct a complete bipartite graph $K_{N,N}$ between the particles and quenched points, with each edge assigned a weight $w_{ij} = (\boldsymbol{r}_i - \zeta_j)^2$. Apply the assignment algorithm proposed by Jonker and Volgenant [1] to determine the optimal match that minimizes the objective function in (S1). Because of PBC, a global translation $\boldsymbol{\zeta} \to \boldsymbol{\zeta} + \boldsymbol{r}_0$ with a constant $\boldsymbol{r}_0$ is permissible. This can be addressed by shifting the center of mass of $\boldsymbol{\zeta}$ to zero.

2. Decrease packing fraction: Shift ϕ to $\phi - \Delta\phi$ (with a typical choice of $\Delta\phi = 0.1$) by correspondingly reducing the particle diameter σ, and find the new minimum. We start with an admissible packing with the optimal assignment, i.e., the best label matches between the particles and quenched random points such that $\boldsymbol{r}_{ij} \cdot \boldsymbol{\zeta}_{ij} \le 0$. This can be achieved by the assignment algorithm in Step 1). We then implement the interior-point method to optimize the following auxiliary free energy:

$$f(\boldsymbol{R}, \sigma) \equiv \sum_i (\boldsymbol{r}_i - \boldsymbol{\zeta}_i)^2 - t \sum_{i<j} U(r_{ij}/\sigma), \tag{S2}$$

   where $\boldsymbol{R} \equiv \{\boldsymbol{r}_i\}$, $\boldsymbol{r}_{ij} \equiv \boldsymbol{r}_i - \boldsymbol{r}_j$, and $r_{ij} = |\boldsymbol{r}_{ij}|$. Here, $U(x) = \ln(x^2 - 1)$, aan auxiliary entropy represented by a logarithmic barrier function that forces the search step to fall inside the hard-sphere constraints, and the parameter t serves as the auxiliary temperature. Readers should not confuse these auxiliary parameters with the physical temperature and free energy

defined in Eq. (1). The interior-point method commences with a relatively high t value, which is iteratively halved until it reaches a sufficiently small value, typically $t = 10^{-10}$. In this sense, the barrier method is similar to simulated annealing, where we start with a sufficiently large auxiliary temperature t and cool down the system by reducing the auxiliary temperature until reaching a minimum. Note that the interior-point method is insensitive to initial conditions, allowing us to use an initial interior point from the previous larger packing fractions, which is admissible with the optimal assignment having been solved, to save computational time. Alternatively, we can start from any admissible packing, followed by the assignment algorithm. We observe numerically that different choices of initial conditions lead to the same optimal packing.

3. For a given t, the optimization (S2) is solved using the Limited-memory Broyden–Fletcher –Goldfarb–Shanno (L-BFGS) algorithm [2]. To avoid stagnation at a local minimum during the optimization process, we maintain a watch on particle matches to ensure $\boldsymbol{r}_{ij} \cdot \boldsymbol{\zeta}_{ij} \leq 0$. If not, we exchange the labels i and j to achieve a lower energy.

4. Repeat Step 2 iteratively across the entire range of packing fractions.

The forces exerted on the particles, and consequently, the reduced pressure can be evaluated from the optimal configuration. This can be elucidated by observing that, for a fixed parameter t, the optimal packing $\boldsymbol{R}^*$ of Eq. (S2) requires the gradient $\boldsymbol{g}(r, \sigma) \equiv \nabla_{\boldsymbol{R}} f(\boldsymbol{R}, \sigma)$ vanish, i.e., $\boldsymbol{g}(\boldsymbol{R}^*, \sigma) = 0$, leading to

$$2(\boldsymbol{r}_i - \zeta_i) = t \sum_j \nabla_i U(r_{ij}/\sigma) = \frac{t}{\sigma} \sum_j U'(r_{ij}/\sigma) \hat{\boldsymbol{n}}_{ij}, \tag{S3}$$

where $\hat{\boldsymbol{n}}_{ij} \equiv \boldsymbol{r}_{ij}/r_{ij}$. From this, we can derive the force $\boldsymbol{f}_i$ acting on particle i as

$$\boldsymbol{f}_i = -2(\boldsymbol{r}_i - \zeta_i) = \sum_j \boldsymbol{f}_{ij}. \tag{S4}$$

Here, $\boldsymbol{f}_{ij}$ is the force of contact between particles i and j, as

$$\boldsymbol{f}_{ij} = -f_{ij} \hat{\boldsymbol{n}}_{ij}, \tag{S5}$$

with its magnitude

$$f_{ij} \equiv \frac{t}{\sigma} U'(r_{ij}/\sigma). \tag{S6}$$

In the limit as t approaches 0, only neighboring contacts where $r_{ij} \to \sigma$ remain non-zero. Hence, Eq. (S6) provides a method to compute the pairwise contact force f_{ij} for the optimization algorithm [3].

2 MSDs for $\phi = \phi_m$

Figure S1 investigate the finite scaling of MSD as a function of the system size N. Our results show that, for $d < 3$, $\Delta(\phi_m)$ diverges with increasing N, indicating a delocalized phase. In particular,

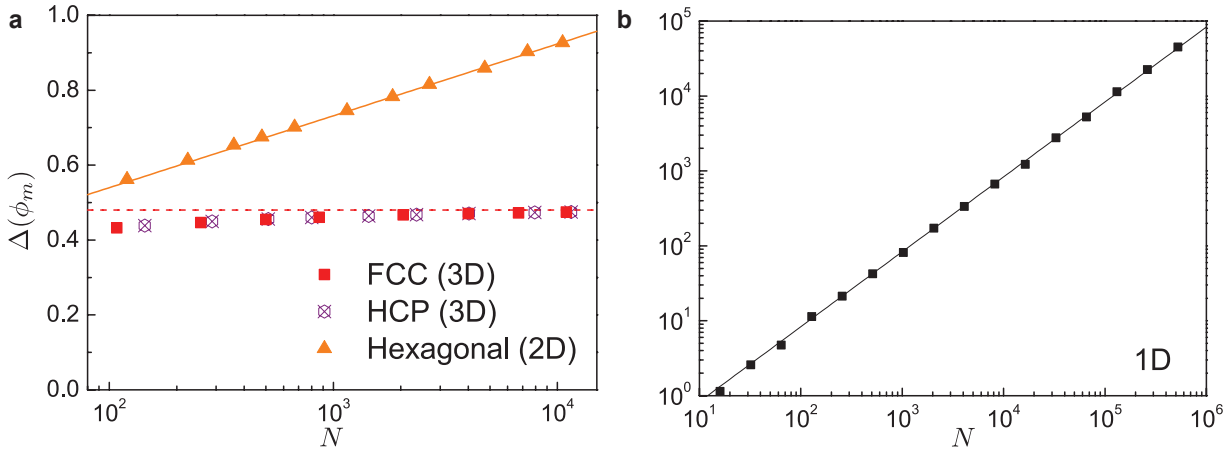

Figure S1: **Maximum close packing.** The mean square displacement $\Delta(\phi_m)$ as a function of system size N, for (a) FCC (red, solid squares) and HCP (blue, crossed squares) sphere packings, and two-dimensional hexagonal packing (orange, solid triangles), and for (b) one-dimensional packing (black, solid squares). The orange and black lines represent $\phi_m \sim \ln N$ for 2D, and $\phi_m \sim N/12$ for 1D, respectively.

for $d = 1$, the MSD scales linearly with size, $\Delta(\phi_m) \sim N/12$ for the periodic boundary condition and $\sim N/6$ for the open boundary condition. For two dimensions, our numerical results suggest a logarithmic divergence, i.e., $\Delta(\phi_m) \sim \ln N$. In contrast, $\Delta(\phi_m)$ is finite in three dimensions for both HCP and FCC packings. This result is rather surprising because one would expect a large configurational change from the most ordered packings (FCC/HCP) to the most disordered ordered ones (Poisson). However, we find the opposite: a bounded displacement is sufficient to transform one to the other, implying that all three-dimensional packings are close to each other. It appears that $\Delta \sim (\phi_m - \phi)^{2-d} + O(1)$, suggesting that $d = 2$ is the lower critical dimension of the delocalization, which appears linked to the jamming transition.

3 Local rearrangement

In this section, we derive the relation between the reduced pressure π and the change of MSD $\mathrm{d}\Delta/\mathrm{d}\phi^{-1}$ under a local rearrangement. By local rearrangement, we mean that under a small change of diameter σ and thereby the packing fraction ϕ, the optimal packing configuration only deforms slightly.

Starting with a fixed diameter σ and the corresponding optimal packing $\boldsymbol{R}^*$, we expand $f(\boldsymbol{R}, \sigma)$ around $\boldsymbol{R}^*$,

$$f(\boldsymbol{R}, \sigma) = f(\sigma) + \boldsymbol{g}(\boldsymbol{R}^*, \sigma) \cdot \delta\boldsymbol{R}^* + \frac{1}{2}(\delta\boldsymbol{R}^*)^T \boldsymbol{H}(\boldsymbol{R}^*, \sigma)\delta\boldsymbol{R}^* + O((\delta\boldsymbol{R}^*)^3), \tag{S7}$$

where $f(\sigma) \equiv f(\boldsymbol{R}^*, \sigma)$, the gradient $\boldsymbol{g}(\boldsymbol{R}, \sigma) \equiv \nabla_{\boldsymbol{R}} f(\boldsymbol{R}, \sigma)$ satisfies $\boldsymbol{g}(\boldsymbol{R}^*, \sigma) = 0$, and the Hessian $\boldsymbol{H}(\boldsymbol{R}, \sigma) \equiv \nabla_{\boldsymbol{R}} \otimes \nabla_{\boldsymbol{R}} f(\boldsymbol{R}, \sigma)$.

Considering a small perturbation $\sigma \to \sigma + \delta\sigma$, and $\boldsymbol{R}^* \to \boldsymbol{R}^* + \delta\boldsymbol{R}^*$, where $\delta\boldsymbol{R}^*$ only undergoes a local rearrangement and is of the same order as $\delta\sigma$, the new minimum is $f(\sigma + \delta\sigma) \equiv f(\boldsymbol{R}^* + \delta\boldsymbol{R}^*, \sigma + \delta\sigma)$. Expanding f around $(\boldsymbol{R}^*, \sigma)$, we have

$$f(\sigma+\delta\sigma) = f(\sigma)+\partial_\sigma f(\boldsymbol{R}^*,\sigma)\delta\sigma+\partial_\sigma \boldsymbol{g}(\boldsymbol{R}^*,\sigma)\cdot\delta\boldsymbol{R}^*\delta\sigma+\frac{1}{2}\left(\partial_\sigma^2 f(\boldsymbol{R}^*,\sigma)\delta\sigma^2 + (\delta\boldsymbol{R}^*)^T\boldsymbol{H}(\boldsymbol{R}^*,\sigma)\delta\boldsymbol{R}^*\right)+O(\delta\sigma^3) \tag{S8}$$

Minimizing $f(\boldsymbol{R}^* + \delta\boldsymbol{R}^*, \sigma + \delta\sigma)$ requires $\partial_{\delta\boldsymbol{R}^*} f(\boldsymbol{R}^* + \delta\boldsymbol{R}^*, \sigma + \delta\sigma) = 0$, leading to

$$\delta\boldsymbol{R}^* = -\boldsymbol{H}^{-1}(\boldsymbol{R}^*,\sigma)\partial_\sigma \boldsymbol{g}(\boldsymbol{R}^*,\sigma)\delta\sigma \tag{S9}$$

Substituting into Eq. (S8), we obtain

$$f(\sigma + \delta\sigma) = f(\sigma) + \frac{\mathrm{d}f(\sigma)}{\mathrm{d}\sigma}\delta\sigma + \frac{1}{2}\frac{\mathrm{d}^2 f(\sigma)}{\mathrm{d}\sigma^2}\delta\sigma^2 + O(\delta\sigma^3), \tag{S10}$$

where

$$\frac{\mathrm{d}f(\sigma)}{\mathrm{d}\sigma} = \partial_\sigma f(\boldsymbol{R}^*,\sigma), \tag{S11a}$$

$$\frac{\mathrm{d}^2 f(\sigma)}{\mathrm{d}\sigma^2} = \partial_\sigma^2 f(\sigma,\boldsymbol{R}^*) - \partial_\sigma \boldsymbol{g}(\boldsymbol{R}^*,\sigma)^T h(\boldsymbol{R}^*,\sigma)^{-1}\partial_\sigma \boldsymbol{g}(\boldsymbol{R}^*,\sigma). \tag{S11b}$$

Note that from Eq. (S2), we have

$$\boldsymbol{R}\cdot\nabla_{\boldsymbol{R}} f(\boldsymbol{R},\sigma) = \boldsymbol{R}\cdot\boldsymbol{F} - \sigma\partial_\sigma f(\boldsymbol{R},\sigma), \tag{S12}$$

where $\boldsymbol{F} \equiv \{\boldsymbol{f}_i\}$. The validity of this identity is independent of the choices of t and U functions, and thus holds also in the optimization limit $t \to 0$. Setting $\boldsymbol{R} = \boldsymbol{R}^*$ and $\boldsymbol{F} = \boldsymbol{F}^*$ and taking the limit $t \to 0$, we have $\lim_{t\to 0} f(\boldsymbol{R}^*,\sigma) = N\Delta(\sigma)$. Substituting Eq. (S11a) into Eq. (S12) and noting that the gradient $\nabla_{\boldsymbol{R}} f(\boldsymbol{R},\sigma)$ vanishes at $\boldsymbol{R}^*$, we obtain

$$\frac{1}{N}\boldsymbol{R}^*\cdot\boldsymbol{F}^* = \sigma\frac{\mathrm{d}\Delta}{\mathrm{d}\sigma} = \phi d\frac{\mathrm{d}\Delta}{\mathrm{d}\phi}. \tag{S13}$$

Recall the definition of the reduced pressure

$$\pi \equiv \frac{\phi}{dN}\sum_{i<j}\boldsymbol{r}_{ij}\cdot\boldsymbol{f}_{ij} = \frac{\phi}{dN}\boldsymbol{R}^*\cdot\boldsymbol{F}^*, \tag{S14}$$

where the second equality arises from the Virial theorem, $\boldsymbol{R}^*\cdot\boldsymbol{F}^* = \sum_i \boldsymbol{r}_i\cdot\boldsymbol{f}_i = \frac{1}{2}\sum_{ij}\left(\boldsymbol{r}_i\cdot\boldsymbol{f}_{ij} + \boldsymbol{r}_j\cdot\boldsymbol{f}_{ji}\right) = \sum_{i<j}\boldsymbol{r}_{ij}\cdot\boldsymbol{f}_{ij}$. Substituting into Eq. (S13), we obtain the relation between π and Δ for the local rearrangement

$$\pi = -\frac{\mathrm{d}\Delta}{\mathrm{d}\phi^{-1}}. \tag{S15}$$

This completes our proof.

4 Jamming transition

In this section, we show the condition under which local rearrangement is valid and its association with the jamming transition. The validity of Taylor expansion in Eq. (S10) requires a well-defined second derivative

$$\sigma^2 \frac{\mathrm{d}^2 f(\sigma)}{\mathrm{d}\sigma^2} = \sigma^2 \partial_\sigma^2 f(\sigma, \boldsymbol{R}^*) - \boldsymbol{G}^T \boldsymbol{H}^{-1} \boldsymbol{G}. \tag{S16}$$

where $\boldsymbol{G} \equiv \sigma \partial_\sigma \boldsymbol{g}(\boldsymbol{R}^*, \sigma)$. By respectively applying $\nabla_{\boldsymbol{R}}$ and ∂_σ to both sides of Eq. (S12), we obtain

$$\boldsymbol{H}\boldsymbol{R}^* = (\boldsymbol{F}^* - 2\boldsymbol{R}^*) - \boldsymbol{G}, \tag{S17a}$$

$$\boldsymbol{R}^* \cdot \boldsymbol{G} = -\boldsymbol{R}^* \cdot \boldsymbol{F}^* - \sigma^2 \partial_\sigma^2 f(\sigma, \boldsymbol{R}^*). \tag{S17b}$$

Upon integrating these results, we find

$$\sigma^2 \frac{\mathrm{d}^2 f(\sigma)}{\mathrm{d}\sigma^2} = (2\boldsymbol{R}^* - \boldsymbol{F}^*) \cdot \boldsymbol{q} - \boldsymbol{R}^* \cdot \boldsymbol{F}^*, \tag{S18}$$

where $\boldsymbol{q} \equiv \boldsymbol{H}^{-1}\boldsymbol{G}$. One can write Eq. (S18) as

$$\sigma^2 \frac{\mathrm{d}^2 f(\sigma)}{\mathrm{d}\sigma^2} = \sum_{i<j} (2N^{-1}\boldsymbol{r}_{ij} - \boldsymbol{f}_{ij}) \cdot \boldsymbol{q}_{ij} - \boldsymbol{f}_{ij} \cdot \boldsymbol{r}_{ij}, \tag{S19}$$

where $\boldsymbol{q}_{ij} \equiv \boldsymbol{q}_i - \boldsymbol{q}_j$.

Upon taking the limit as $t \to 0$, Eq. (S6) implies that $U' \sim t^{-1}$ as $r_{ij} \to \sigma$, ensuring that $\boldsymbol{F}^*$ is finite. Conversely, both $\boldsymbol{G}$ and $\boldsymbol{H}$ diverge because they involve U'', which exhibits a stronger divergence as $r_{ij} \to \sigma$. To have a finite $f''(\sigma)$, we require

$$\boldsymbol{q}^* \equiv \lim_{t\to 0} \boldsymbol{q} = \lim_{t\to 0} \boldsymbol{H}^{-1}\boldsymbol{G}, \tag{S20}$$

to be well-defined. Indeed, Equation (S9) suggests

$$\frac{\delta \boldsymbol{R}^*}{\delta \sigma} = -\sigma^{-1}\boldsymbol{q}^*. \tag{S21}$$

Therefore, the local rearrangement requires $\boldsymbol{q}$ to be finite. To illustrate this more clearly, we find

$$\boldsymbol{G}_i = -\boldsymbol{f}_i - \frac{t}{\sigma^2} \sum_j U''(x_{ij}) \boldsymbol{r}_{ij} \tag{S22}$$

where $x_{ij} \equiv r_{ij}/\sigma$. Similarly, we find

$$\boldsymbol{H}_{ij} = \frac{t}{\sigma^2} \left(-U''(x_{ij}) \hat{\boldsymbol{n}}_{ij} \otimes \hat{\boldsymbol{n}}_{ij} + x_{ij}^{-1} U'(x_{ij}) (\hat{\boldsymbol{n}}_{ij} \otimes \hat{\boldsymbol{n}}_{ij} - I_d) \right) \tag{S23}$$

is applicable for off-diagonal elements $i \neq j$, where I_d is the d-dimensional identity matrix, and

$$\boldsymbol{H}_{ii} = 2I_d - \sum_{j\neq i} \boldsymbol{H}_{ij} \tag{S24}$$

is used for diagonal elements. We assume that U'' diverges as $(t\epsilon(t))^{-1}$ for small t, where the function $\epsilon(t)$ captures the asymptotic behavior such that $-\lim_{t\to 0}\frac{\epsilon(t)t}{\sigma^2}U''(x_{ij}) = u_{ij}$ is finite for neighboring contacts. For instance, for the logarithmic barrier function $U(x) = \ln(x-1)$, $U'(x) = \frac{1}{x-1}$ and $-U''(x) = \frac{1}{(x-1)^2}$, Eq. (S6) implies that only the neighboring contact $r_{ij} = \sigma + \frac{1}{f_{ij}}t + O(t^2)$ contributes. Additionally, setting $\epsilon(t) = t$, we find that $-\lim_{t\to 0}\epsilon(t)\frac{t}{\sigma^2}U''(x_{ij}) = f_{ij}^2 = u_{ij}$. The argument below will be maintained for a general U form, as the specific choice of the barrier function does not alter the result. Equation (S20) is consequently transformed into

$$\boldsymbol{H}^*\boldsymbol{q}^* = \boldsymbol{G}^*, \tag{S25}$$

where $\boldsymbol{H}^*_{ij} \equiv \sigma^{-2}u_{ij}\boldsymbol{r}_{ij}\otimes\boldsymbol{r}_{ij}$ with $\boldsymbol{H}^*_{ii} \equiv -\sum_{j\neq i}\boldsymbol{H}^*_{ij}$, and $\boldsymbol{G}^*_i \equiv \sum_j u_{ij}\boldsymbol{r}_{ij}$. We expand

$$\boldsymbol{H}^* = \sum_{ij} h_{ij}|\boldsymbol{v}^{(ij)}\rangle\langle\boldsymbol{v}^{(ij)}|, \tag{S26}$$

in terms of the basis $\{\boldsymbol{v}^{(ij)}\}$, where the coefficient $h_{ij} \equiv -\sigma^{-2}u_{ij}$, and $\boldsymbol{v}^{(ij)} = (\boldsymbol{e}^{(i)} - \boldsymbol{e}^{(j)})\otimes(\boldsymbol{r}_i - \boldsymbol{r}_j)$ are dN-dimensional vectors. Here, $\boldsymbol{e}^{(i)}$ is the N-dimensional vector with a unit at position i and zeros elsewhere. The number of bases $\{\boldsymbol{v}^{(ij)}\}$ is equivalent to the number of contacts $zN/2$, with z representing the kissing number.

The basis set $\{\boldsymbol{v}^{(ij)}\}$ is not orthogonal and possesses a d-dimensional non-trivial kernel $\boldsymbol{T} = \sum_i \boldsymbol{e}^{(i)}\otimes\boldsymbol{r}_0$ with an arbitrary d-dimensional vector $\boldsymbol{r}_0$, such that $\boldsymbol{T}\cdot\boldsymbol{v}^{(ij)} = 0$. This property corresponds to the global translational invariance of the PBC, i.e., $\boldsymbol{r}_i \to \boldsymbol{r}_i + \boldsymbol{r}_0$. Consequently, $\boldsymbol{H}^*$ has a maximum rank $d(N-1)$ and is not directly invertible. However, Eq. (S25) does offer a solution since $\boldsymbol{G}^*$ resides outside of the null space, i.e., $\boldsymbol{G}^*\cdot\boldsymbol{T} = 0$. In fact, it encompasses a d-dimensional solution space, as if $\boldsymbol{q}^*$ serves as a solution, then $\boldsymbol{q}^* + \boldsymbol{T}$ is also a valid solution. This ambiguity, however, does not affect Eq. (S19) as $\boldsymbol{q}_{ij}$ is invariant under the global translation.

Eq. (S25) allows solutions when the rank of $\boldsymbol{H}^*$ reaches its maximum value $d(N-1)$. Given that there are $zN/2$ bases, it requires $zN \geq 2d(N-1)$. In the thermodynamic limit of $N\to\infty$, we recover the Maxwell counting argument

$$z \geq 2d, \tag{S27}$$

which provides a necessary condition for the validity of the local rearrangement. To ensure its sufficiency, we need the rank of these $zN/2$ base vectors equal to $d(N-1)$, a condition that depends on the configuration $\boldsymbol{R}^*$. Our numerical results, however, indicate that this condition is likely met, at least statistically. Thus, the isostatic condition corresponds to the critical value $z_c = 2d$ at the jamming transition. Any value below this threshold triggers a global rearrangement of the system, leading to its unjammed state.

References

[1] Roy Jonker and Anton Volgenant. A shortest augmenting path algorithm for dense and sparse linear assignment problems. *Computing*, 38(4):325–340, 1987.

[2] Dong C Liu and Jorge Nocedal. On the limited memory bfgs method for large scale optimization. *Mathematical programming*, 45(1-3):503–528, 1989.

[3] Claudia Artiaco, Rafael Díaz Hernández Rojas, Giorgio Parisi, and Federico Ricci-Tersenghi. Hard-sphere jamming through the lens of linear optimization. *Physical Review E*, 106(5):055310, 2022.