

Insertion space in repulsive active matter

Luke K. Davis^{1,2,3,*} and Karel Proesmans⁴¹*School of Mathematics and Maxwell Institute for Mathematical Sciences,
University of Edinburgh, Edinburgh EH9 3FD, United Kingdom*²*Higgs Centre for Theoretical Physics, University of Edinburgh, Edinburgh EH9 3FD, United Kingdom*³*Isaac Newton Institute for Mathematical Sciences, University of Cambridge, Cambridge CB3 0EH, United Kingdom*⁴*Niels Bohr International Academy, Niels Bohr Institute, Blegdamsvej 17, DK-2100 Copenhagen, Denmark*

(Received 9 September 2025; revised 13 April 2026; accepted 9 June 2026; published 17 July 2026)

For equilibrium hard spheres, the statistical geometry of the insertion space, the room to accommodate another sphere, relates exactly to the equation of state. We begin to extend this idea to active matter, analyzing the insertion space for repulsive active particles in one and two dimensions using both on- and off-lattice models. In one dimension, we derive closed-form expressions for the mean insertion cavity size, cavity number, and total insertion volume, all in excellent agreement with simulations. Strikingly, activity increases the total insertion volume and tends to keep the insertion space more connected. We also find that insertion space metrics contain signatures of collective phase behaviors occurring at previously predicted packing fractions. Taken together, our work provides the first quantitative foundation for the statistical geometry of active matter.

DOI: [10.1103/n9s4-ykn2](https://doi.org/10.1103/n9s4-ykn2)

Insertion is one of the most basic actions in the natural and human worlds: it is the act of storing physical objects in space. The practical difficulty of insertion lies in the objects taking up a limited amount of space and accounting for the many ways to arrange the objects (Fig. 1). Since the work of B. Widom (1960s), only insertion into a medium in equilibrium with its environment has been considered [1,2], but what happens if the insertion medium is, instead, motile and far from equilibrium? Understanding insertion has important practical ramifications regarding the storage of motile units, such as self-propelled colloids [3], robots [4], and biological cells [5], which could form the next generation of engines [6]. Here, we provide the first answer to this question in the context of active matter: a broad class of nonequilibrium systems whose constituents turn over energy to fuel their dynamics [7].

Active constituents interact in many ways, with volume exclusion (a form of repulsion) being the most relevant for insertion [8–10]. An important phenomenon that arises for active—as opposed to passive—repulsive particles is motility-induced phase separation (MIPS): the formation of dense clusters of particles surrounded by a dilute vapor [11]. The nonequilibrium driving and the many-body repulsive interactions make understanding the insertability of active matter difficult.

It has been shown that some frameworks and ideas from equilibrium statistical mechanics, i.e., of passive matter, can

be leveraged to understand the thermodynamics of active matter [12]. Motivated by these successes, we here begin to investigate an unexplored avenue: the relationship between thermodynamics and insertion space (IS), or, in general, statistical geometry, which was established for equilibrium hard spheres [13–20]. Remarkably, there exist *exact* relations, largely attributed to R. J. Speedy, connecting both the equilibrium compressibility $Z = \beta PV/N$ and, separately, the total chemical potential μ to the averaged insertion space, i.e., the room available to place another sphere averaged over all configurations [13,14] (Fig. 1):

$$Z_G := 1 + \frac{\sigma \langle S_I \rangle}{2d \langle V_I \rangle}, \quad \mu_G := -\ln \frac{\langle V_I \rangle}{V}, \quad (1)$$

$$\underbrace{Z_G = Z \equiv \beta PV/N, \quad \mu_G = \mu_{\text{exc}} = \beta \mu - \ln \frac{N \Lambda^d}{V}}_{\text{At equilibrium}}, \quad (2)$$

where $\beta = (k_B T)^{-1}$ is the inverse temperature, P is the bulk pressure, V is the total volume of the space, N is the number of constituents, σ is the particle diameter (or repulsion range), d is the spatial dimension, Λ is the thermal de Broglie wavelength, $\langle V_I \rangle$ is the averaged (over all values in all configurations) total insertion volume, and $\langle S_I \rangle$ its averaged concomitant surface. The insertion volume is expressed in terms of the averaged (over all values across all configurations) number of cavities and cavity size as $\langle V_I \rangle = \langle N_\Delta \rangle \langle \Delta \rangle$ [16,18].

The appeal of the above relations is that—at equilibrium—each of the left-hand sides of Eq. (1) has clear thermodynamic interpretation (2), while the right-hand sides of Eq. (1) are purely geometric. However, it is not straightforward how one should view the above relations in light of active matter: systems far from equilibrium. Already, it has been shown that

*Contact author: luke.davis@ed.ac.uk

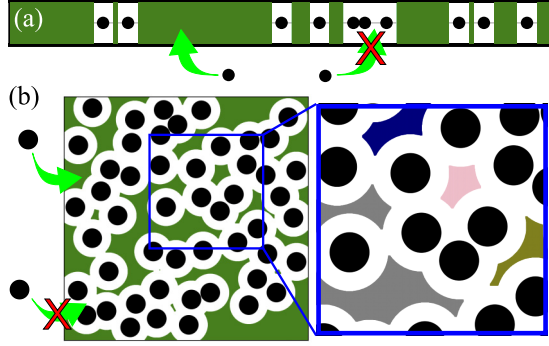


FIG. 1. (a) Visualized insertion cavities for a configuration of hard spheres in $d = 1$. (b) Example insertion cavities for a configuration of hard disks in a closed box. Disconnected cavities identified in our algorithm (right).

an equation of state, Z , with the equilibrium interpretation, is the exception and not the rule for active systems [21]. Thus, if Eq. (2) were to somehow hold, the interpretation of the thermodynamic objects must change. Though, before this, a systematic investigation into the geometric objects in Eq. (1) for active systems needs to be done. An important anchoring point is that the determination and interpretation of the IS quantities are purely geometric and remain concrete far away from equilibrium. Here, we determine the IS and its statistics for repulsive active matter models in low dimensions.

First, in $d = 1$, the averaged total number of cavities $\langle N_\Delta \rangle$ can be written in terms of P_C , the contact probability of two particles, as

$$\langle N_\Delta \rangle = N(1 - P_C), \quad (3)$$

with $\langle N_\Delta \rangle = N$ in the dilute limit (packing fraction $\eta = N\sigma/V \rightarrow 0$) and $\langle N_\Delta \rangle = 0$ at full packing ($\eta \rightarrow 1$). The averaged total insertion volume is, as a function of P_C ,

$$\langle V_I \rangle = V + N\sigma(P_C - 2), \quad (4)$$

with $\langle V_I \rangle = V - N(2\sigma)$ in the dilute case, where 2σ is the volume excluded by one particle, and $\langle V_I \rangle = 0$ at full packing. Then, based on simple manipulations of Eqs. (3) and (4), the dimensionless averaged cavity size $\langle \Delta \rangle / \sigma \equiv \langle \tilde{\Delta} \rangle$ is given as

$$\langle \tilde{\Delta} \rangle = \frac{\langle V_I \rangle}{\sigma \langle N_\Delta \rangle} = \frac{1 + \eta(P_C - 2)}{\eta(1 - P_C)}. \quad (5)$$

The above exact relations, expressing the insertion observables in terms of P_C , to the best of our knowledge appear to be new and are consistent with the expressions in Ref. [14].

In one dimension, $\langle S_I \rangle = 2\langle N_\Delta \rangle$ exactly, implying that the geometric compressibility Z_G (1) can be simplified as $Z_G = 1 + \langle \tilde{\Delta} \rangle^{-1}$ [14]. The above observables for the equilibrium one-dimensional hard-sphere model, i.e., the Tonks gas [8], are given exactly as [8, 15, 16]

$$\begin{aligned} \langle \tilde{\Delta} \rangle_0 &= \frac{(1 - \eta)}{\eta}, & \langle N_\Delta \rangle_0 &= N \exp\left(-\frac{1}{\langle \tilde{\Delta} \rangle_0}\right), \\ \langle V_I \rangle_0 &= V(1 - \eta) \exp\left(-\frac{1}{\langle \tilde{\Delta} \rangle_0}\right), \end{aligned} \quad (6)$$

with a corollary that $P_C^{\text{eq}} = 1 - \exp(-\langle \tilde{\Delta} \rangle_0^{-1})$.

To make analytical progress for the active model, we consider repulsive run-and-tumble particles (RTPs) in $d = 1$. The model consists of N RTPs, of diameter σ , on a lattice of $n = V/\sigma \in \mathbb{Z}^+$ sites. In this model, the run lengths k_i , for particle i , are exponentially distributed with rate parameter $K^{-1} \geq 0$, and during the runs the particles have a constant self-propulsion speed v_0 [22] (see Supplemental Material [23] for all model details). We note that $K \sim l_p/\sigma$ is to be interpreted as some persistence length divided by the particle size. At equilibrium ($K \rightarrow 0$), the model reduces to the simple symmetric exclusion process [24]. The RTP particle model can be mapped to a mass transfer model with fluctuating directed bonds [as is done in the Dandekar-Chakraborti-Rajesh (DCR) model [25]], which we exploit here. The neighboring probability P_C can be broken down as $4P_C = p_{\rightarrow, \leftarrow} + 2p_{\rightarrow, \rightarrow} + p_{\leftarrow, \leftarrow}$, where $p_{\rightarrow, \leftarrow}$ is the probability that two particles are next to each other with the left one moving right and the right one moving left, $p_{\rightarrow, \rightarrow} \equiv p_{\leftarrow, \leftarrow}$ is the probability that two particles are next to each other where both particles move in the same direction, and $p_{\leftarrow, \rightarrow}$ is the probability to be next to each other with the left particle moving left and the right particle moving right. The expressions for $p_{\rightarrow, \leftarrow}$, $p_{\rightarrow, \rightarrow}$, and $p_{\leftarrow, \rightarrow}$ are difficult to obtain exactly, and so we resort to a mean-field approximation (as is done in Ref. [25]; see Supplemental Material for details [23]). For weak activity, i.e., $K \lesssim 1$, the mean-field P_C reads [up to $\mathcal{O}(K^3)$] as

$$P_C \underset{K \lesssim 1}{=} \eta + \frac{\eta(\eta - 1)^2 K}{2} + \frac{(\eta - 1)^2 \eta (4\eta^2 - 7\eta + 1) K^2}{8}. \quad (7)$$

For the high-activity (large persistence) regime, naively expanding P_C for small $K^{-1} \ll 1$ is pathological (as noted in Ref. [25]). The main reason for this is that the mean-field treatment assumes bonds (orientations between close particles) are independent and thus cannot—by design—represent the long, strongly correlated domains of like orientation that arise at large persistence. A better high-activity starting point is the coalescence-fragmentation picture in the DCR model (Sec. IV of Ref. [25]), which focuses on the interfaces of comoving particle clusters. In this picture, in the steady state, there is a balance in the rates of clusters joining and breaking, resulting in the density of interfacial bonds (jammed interfaces) going as $\zeta^* = \sqrt{K^{-1} \langle \tilde{\Delta} \rangle_0}$ and the typical unoccupied cavity size going as $K\zeta^*$. Then, the contact probability is simply $P_C \underset{K \gg 1}{=} 1 - \zeta^*$. In the full packing regime, both the weak- and high-activity limits obey $P_C(\eta \rightarrow 1) = 1$, yet in the dilute regime $P_C(\eta \rightarrow 0) \sim 0$ for $K \rightarrow 0$, while $P_C(\eta \rightarrow 0) \sim \mathcal{O}(1)$ for $K \rightarrow \infty$ (high persistence limit taken first).

The dimensionless averaged cavity size (5) in the presence of weak and strong activity is then, respectively,

$$\begin{aligned} \langle \tilde{\Delta} \rangle \underset{K \lesssim 1}{=} & \langle \tilde{\Delta} \rangle_0 + \frac{K}{2}(1 - \eta) - \frac{K^2(1 - \eta)(1 - 5\eta + 2\eta^2)}{8}, \\ \langle \tilde{\Delta} \rangle \underset{K \gg 1}{=} & K\zeta^* - 1 \equiv \sqrt{K \langle \tilde{\Delta} \rangle_0} - 1, \end{aligned} \quad (8)$$

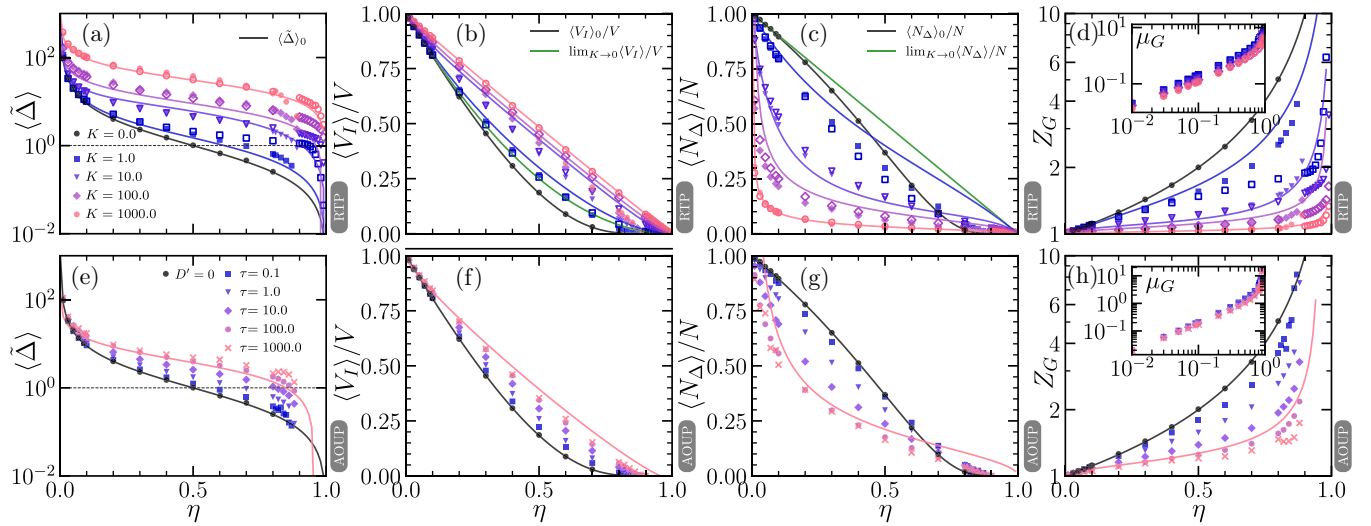


FIG. 2. Insertion space for repulsive active particles in $d = 1$. (a) Averaged normalized cavity size as a function of packing fraction η , as determined from off-lattice (filled points) and on-lattice (unfilled points) simulations of RTPs at various K . Analytical results are solid lines; for $K \leq 10.0$, the lines are the weak-activity mean-field expressions, while for $K > 10.0$ we show the high-activity expressions [Eqs. (8)–(10)]. Exact passive expressions are from Eq. (6) and are shown along with strict passive simulations ($K = 0$). (b) The averaged total insertion volume divided by the system volume. (c) Averaged number of insertion cavities per particle as a function of η . (d) The geometric compressibility, Z_G , as a function of η . Inset is the geometric chemical potential μ_G against η . (e)–(h) The same quantities as for panels (a)–(d) but for the AOUP (off-lattice) model. High-activity analytics (for $\tau = 1000$) are Eqs. (8)–(10), with $K \simeq \sqrt{D'\tau}/\sigma$. Parameters (RTPs): $N = 1024$, $\mu = 1$ (off-lattice), $\epsilon = 10$ (off-lattice), and $\sigma = 1$; AOUPs: $N = 1024$, $\mu = 1$, $\epsilon = 10$, $D' = \frac{1}{2}$, and $\sigma = 1$.

with $\langle N_{\Delta} \rangle$ and $\langle V_I \rangle$, also given, respectively, as

$$\begin{aligned} \frac{\langle N_{\Delta} \rangle}{N} &\underset{K \lesssim 1}{=} (1 - \eta) - \frac{K\eta(1 - \eta)^2}{2} \left(1 - \frac{K(1 - 7\eta + 4\eta^2)}{4} \right), \\ \frac{\langle N_{\Delta} \rangle}{N} &\underset{K \gg 1}{=} \zeta^* \equiv \sqrt{K^{-1} \langle \tilde{\Delta} \rangle_0}, \end{aligned} \quad (9)$$

and

$$\begin{aligned} \frac{\langle V_I \rangle}{V} &\underset{K \lesssim 1}{=} (1 - \eta)^2 \left(1 + \frac{K\eta}{2} + \frac{K^2\eta(1 - 7\eta + 4\eta^2)}{8} \right), \\ \frac{\langle V_I \rangle}{V} &\underset{K \gg 1}{=} 1 - \eta(1 + \zeta^*) \equiv 1 - \eta \left(1 + \sqrt{K^{-1} \langle \tilde{\Delta} \rangle_0} \right). \end{aligned} \quad (10)$$

These relations have three strong physical implications: (1) the relations (8) show that—to leading order—the averaged cavity size is always higher at higher activity, (2) relations (9) imply that the presence of activity tends to reduce, relative to the passive case, the total number of insertion cavities, and (3) relations (10) imply that the presence of activity increases, relative to equilibrium, the averaged total insertion space across the entire packing fraction. Another way to think of this is through the increased probability of insertion $\mathcal{P}_I \equiv \langle V_I(K, \eta) \rangle / V$ [14]. Moreover, in the limit of extreme activity ($K \rightarrow \infty$), we obtain a rather simple relationship $\langle V_I \rangle / V = 1 - \eta$, which is even simpler than both the exact and mean-field-limit equilibrium expressions. Due to $\langle \tilde{\Delta} \rangle > \langle \tilde{\Delta} \rangle_0$, for $K > 0$, we predict a lower Z_G across the full range of η . Interestingly, in the limit $\lim_{K \rightarrow \infty} Z_G = 1$, which is equivalent to the compressibility for an equilibrium ideal gas. Additionally, we predict a lower geometric chemical potential μ_G on the basis that $\langle V_I(K, \eta) \rangle > \langle V_I(\eta) \rangle_0 \forall K, \eta$. These insights and the relations [Eqs. (8)–(10)] are our main theoretical results.

To test our theoretical predictions [Eqs. (8)–(10)], we perform on- and off-lattice numerical simulations of repulsive RTPs in $d = 1$ (Fig. 2). The implementation of the on-lattice model consists of RTPs hopping between adjacent sites and tumbling at exponentially sampled rates with rate parameter K^{-1} , with a successful move to a new site only if that site is unoccupied. The off-lattice model has essentially the same ingredients, with the exception of continuous space and strict volume exclusion being modeled by a repulsive Weeks-Chandler-Andersen (WCA) potential (see Supplemental Material [23]). In the simulations, for a given configuration ω , the IS can be quantified through the insertion profile $\rho_I(x; \omega) := 1 - \Theta(\sum_{i=1}^N \Theta(\sigma - |x - x_i|))$, where $\Theta(\dots)$ is the Heaviside function. By assuming that particles are ordered on the line as $i = 1, \dots, N$, the insertion cavity $\tilde{\Delta}_{i,i+1}(\omega)$ between two neighboring particles i and $i + 1 \bmod N$, the instantaneous number of insertion cavities per particle $N_{\Delta}(\omega)/N$, and the total insertion space $V_I(\omega)/V$ are exact (see Supplemental Material [23]).

Our numerical results are, on the whole, in very good agreement with the theoretical mean-field predictions (Fig. 2). For $\langle \tilde{\Delta} \rangle$, the agreement between the on- and off-lattice simulations is excellent for $\eta \lesssim 0.5$ but worsens for both higher η and lower K [Fig. 2(a)]. The reason for the former is that at higher packing, the off-lattice simulations are able to better sample smaller (than $\sigma = 1$) insertion cavities, which is difficult in the on-lattice simulations with a single site per particle. For the diminishing discrepancy at higher K : for a particular η , the increase in activity increases the mean cavity size, which tends to be higher than σ and which the on-lattice simulations are better able to sample; however, as the system becomes

more dense, the average cavity size will naturally start to shrink, where, again, the on-lattice simulations will struggle to sample these lower values. As expected, as the activity weakens, with $K \rightarrow 0$, the data approach the known passive result $\langle \tilde{\Delta} \rangle_0$ (6) [see also leading order term of weak-activity expression (8)].

Gratifyingly, the numbers for $\langle V_I \rangle / V$ [Fig. 2(b)] at high activity align excellently with our prediction that $\langle V_I \rangle / V = 1 - \eta$. We note that the passive limit of Eq. (10) is in much better agreement for the $K = 1.0$ data than the full weak-activity expression [up to $\mathcal{O}(K^3)$], this suggests that (at $K = 1$) all K terms should approximately cancel to leave $\langle V_I \rangle / V = (1 - \eta)^2$. For the averaged number of insertion cavities $\langle N_\Delta \rangle / N$ [Fig. 2(c)], we note that the excellent agreement between the simulations and the mean-field prediction, which worsens for lower K . Interestingly, at higher activity, e.g., $K = 1000$, there is a rapid decline of the number of insertion cavities at small η , which then flattens out. The interpretation of this is that the insertion space tries to be as connected (in the percolation sense) as possible. We find that both the geometric compressibility and geometric chemical potential are lower for higher K [Fig. 2(d)]. Finally, we find that, out of equilibrium, care must be taken with the averaging, where naive averages of N_Δ and Δ are no longer statistically independent (see Supplemental Material [23]).

We next wondered whether our insights carried over to a different model of repulsive active matter. Thus, we performed off-lattice numerical simulations of repulsive active Ornstein-Uhlenbeck particles (AOUPS) in one dimension. To tune activity, we varied the persistence time of the OU process τ , with passivity taken at $\tau \rightarrow 0$. Overall, we find that the insertion space in the AOUP case closely follows that for RTPs: the average insertion cavity and averaged total insertion volume are always higher for increasing activity, while the averaged number of cavities, geometric compressibility, and geometric chemical potential tend to decrease with activity (see Fig. 2). For the AOUP model, a typical persistence length is $\ell_p = v_{\text{rms}} \tau$ with $v_{\text{rms}} = \sqrt{\langle v^2 \rangle} = \sqrt{D'/\tau}$, so that $\ell_p = \sqrt{D'\tau}$ and one can define $K = \sqrt{D'\tau}/\sigma$. To our surprise, we find that the high-activity expressions of the IS observables of $\langle \tilde{\Delta} \rangle$ [Eqs. (8)–(10)] with the replacement $K = \sqrt{D'\tau}/\sigma$, derived for an on-lattice RTP model, quantitatively describe the off-lattice AOUP numerical data for $\tau = 1000$ almost as well as the off-lattice RTP data.

Thus far, we have only considered one spatial dimension. We now determine how the insertion space changes for repulsive active particles in $d = 2$. Analytical results in higher dimensions present significant challenges, even in equilibrium [14,26]. The mathematical challenge is that, unlike in $d = 1$, different particles can form the boundary of a single cavity, and thus cavities cannot be labeled by particle indices consistently between different configurations. One certain difference in higher dimensions is that the number of cavities per particle approaches (in the dilute limit) $\langle N_\Delta \rangle \rightarrow 1$, in contrast with $\langle N_\Delta \rangle \rightarrow N$ for $d = 1$ [17]. Another fact is that $\langle S_I \rangle / \langle N_\Delta \rangle$, the typical surface area of the insertion cavity, is no longer just a constant.

As there are no analytical relations (on the lattice) in $d = 2$, we here only perform off-lattice simulations [27]. In higher

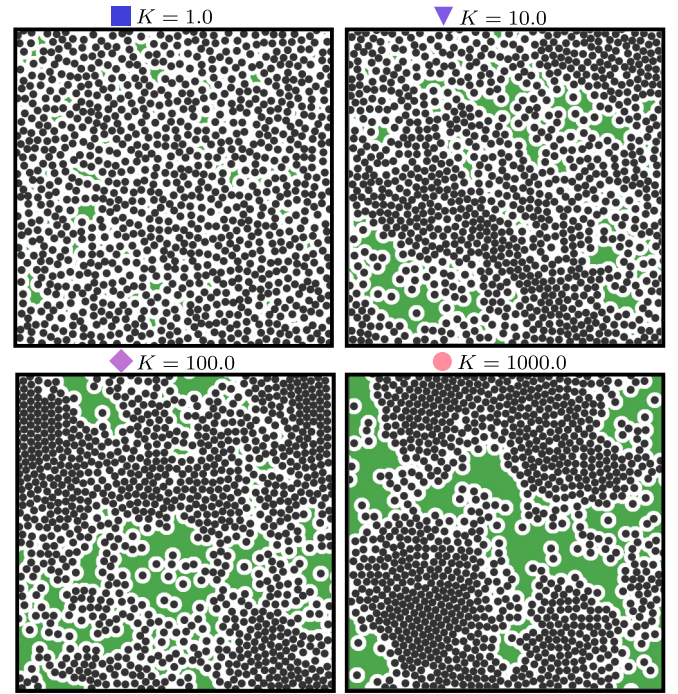


FIG. 3. Snapshots from $d = 2$ off-lattice simulations of $N = 500$ RTPs and insertion cavities for varying K at $\eta = 0.4$.

dimensions, the determination of the IS is nontrivial. Here, we develop a conceptually simple and efficient voxel-based algorithm to determine the IS quantities, which is straightforward to parallelize and also to extend to higher dimensions (see Supplemental Material [23] for the algorithm and benchmark test). In short, our algorithm partitions the space into voxels of side length $l \ll V^{1/2}$ and removes voxels inside the exclusion zones of particles. If implemented naively (comparing each voxel with every particle), the algorithm is incredibly slow, to avoid this, we only check voxels surrounding the particles and leverage the symmetry of the d -dimensional cube, greatly improving the computational complexity and attainable resolution. While exact algorithms based on voronoi constructions exist [18,26,28], they are not as easy to implement and to extend to higher dimensions [29].

Crucially, the off-lattice simulations of RTPs in $d = 2$ still show that the total insertion space (V_I) remains higher and tends to be more connected as compared with equilibrium (Figs. 3 and 4). Interestingly, we find that, at high activities $K \geq 100$, $\langle \tilde{\Delta} \rangle$ no longer remains strictly monotonic (as in $d = 1$) with a wide shelf between $\eta = 0.5$ – 0.75 [Fig. 4(a)]. As expected, our close-to-equilibrium data $K \approx 1$ are in excellent agreement with previous IS quantifications in equilibrium hard-disk simulations [17]. We observe a tendency for lower averaged number of insertion cavities at higher K , with a $\gtrsim$ twofold decrease in the maximum at our highest explored activity [Fig. 4(b)]. Interestingly, the location of the maximum of $\langle N_\Delta \rangle$ is approximately the same as in the equilibrium simulations $\eta = 0.3$ – 0.4 [17]. The similar η values for the peaks of the number of cavities are likely due to the underlying percolation transition of the hard-disk model, where the number of disconnected insertion cavities and particle clusters become equal, though we do not analyze that here [17,34]. We note that for small packing fractions, $\eta < 0.1$, the averaged

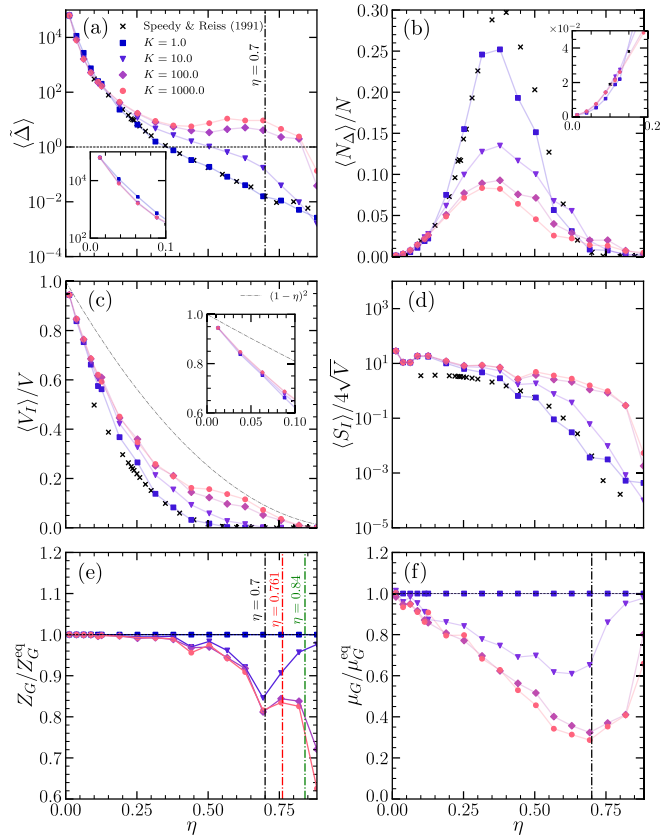


FIG. 4. Quantification of the insertion space for off-lattice simulations of $N = 1000$ repulsive RTPs in $d = 2$. (a) Averaged insertion cavity size as a function of η . Hard-disk equilibrium data from Ref. [17] are shown as reference. Inset shows a close-up for small packing fractions. (b) Number of insertion cavities per particle. (c) Averaged total insertion volume. (d) The boundary of the insertion volume. (e) Z_G (1) as compared to passivity. The packing fraction $\eta = 0.7$ is within 5% of the freezing transition [30] and also the location of liquid-hexatic coexistence at low Pe [31], $\eta = 0.761$ is the packing fraction at which the hard-disk fluid freezes [17,32], $\eta = 0.84$ is the predicted packing fraction at which the metastable fluid to maximally random jammed state transition occurs [33]. (f) The geometric chemical potential compared to passivity.

cavity size is slightly smaller for increased activity yet the averaged number of cavities is slightly higher [Figs. 4(a) and 4(b)], which, nonetheless, results in higher total insertion volume, for higher activities, across the entire range of packing fractions [Fig. 4(c)].

For the boundary of the insertion space $\langle S_I \rangle$, we find it to be a monotonically decreasing function of η (due to periodic boundaries) and generally higher for higher activity, though some saturation appears for very high activities $K \geq 100$ [Fig. 4(d)]. For the geometric compressibility Z_G , we find

that the equilibrium Z_G^{eq} well approximates the active data for low packing $\eta < 0.4$, but then Z_G/Z_G^{eq} starts to rapidly deviate to lower values [Fig. 4(e)]. Interestingly, close to known values of phase transitions (freezing [30] and liquid-hexatic coexistence [31]), $\eta \gtrsim 0.7$, we find that the slope of Z_G/Z_G^{eq} abruptly changes sign. Only for the high-activity data does Z_G/Z_G^{eq} again change sign, at $\eta \approx 0.8$. We also observe a sign change in the slope of the geometric chemical potential μ_G/μ_G^{eq} [Fig. 4(f)].

In this letter, we have made the first analytical and numerical investigations into the insertion space of volume-excluding (repulsive) active matter. We show that, across the entire range of packing fractions, the total insertion space is *always* higher in the presence of activity, however, weak or strong. This fact is borne out in analytics and simulations for two different well-studied particle models of active matter: RTPs and AOUPs. We expect our findings to be true for different models of repulsive active matter. We have discovered that activity tends to reduce the total number of insertion cavities, implying that the insertion space remains more connected. Importantly, our results and predictions regarding the insertion space of active matter are clearly experimentally testable and could be tested in systems of monodisperse active colloids. Future work and challenges include a systematic investigation comparing mechanical definitions of compressibility, pressure, and chemical potential in active systems [21,35–38]. Further, it is not exactly clear how the increased (active) insertion volume should relate to mechanical explanations of MIPS: the slowing down of particles with density results in instabilities that favor clustering [11]. Other questions are: how is the insertion space related to both the degree of time irreversibility and the entropy production [39]? Can the active insertion space reveal the thermodynamic phases of active matter [40]? Remaining technical challenges are to find exact, as well as more direct perturbative, relations for the insertion space. We also wonder how the insertion space behaves in the context of other active matter systems, such as nonreciprocal active matter [41], and how it could inform the control of active matter [42].

We thank Prashant Singh, Mike Cates, and Rob Jack for early discussions. L.K.D. acknowledges the Flora Philip Fellowship at the University of Edinburgh and funding from the Isaac Newton Institute for Mathematical Sciences Postdoctoral Research Fellowship (EPSRC Grant No. EP/V521929/1). L.K.D. thanks the Rosenfeld Foundation for supporting a visit to the NBI, Copenhagen. L.K.D. acknowledges the use of the Edinburgh Compute and Data Facility (ECDF).

Data availability. The data that support the findings of this article are not publicly available. The data are available from the authors upon reasonable request.

- [1] B. Widom, Some topics in the theory of fluids, *J. Chem. Phys.* **39**, 2808 (1963).
- [2] R. J. Speedy, Accurate theory of the hard sphere fluid, *J. Chem. Soc. Faraday Trans. 2* **73**, 714 (1977).

- [3] K. J. M. Bishop, S. L. Biswal, and B. Bharti, Active colloids as models, materials, and machines, *Annu. Rev. Chem. Biomol. Eng.* **14**, 1 (2023).
- [4] A. Barona Balda, A. Argun, A. Callegari, and G. Volpe, Playing with active matter, *Am. J. Phys.* **92**, 847 (2024).

- [5] D. Needleman and Z. Dogic, Active matter at the interface between materials science and cell biology, *Nat. Rev. Mater.* **2**, 17048 (2017).
- [6] P. Pietzonka, E. Fodor, C. Lohrmann, M. E. Cates, and U. Seifert, Autonomous engines driven by active matter: Energetics and design principles, *Phys. Rev. X* **9**, 041032 (2019).
- [7] M. C. Marchetti, J. F. Joanny, S. Ramaswamy, T. B. Liverpool, J. Prost, M. Rao, and R. A. Simha, Hydrodynamics of soft active matter, *Rev. Mod. Phys.* **85**, 1143 (2013).
- [8] L. Tonks, The complete equation of state of one, two and three-dimensional gases of hard elastic spheres, *Phys. Rev.* **50**, 955 (1936).
- [9] T. A. de Pirey, G. Lozano, and F. van Wijland, Active hard spheres in infinitely many dimensions, *Phys. Rev. Lett.* **123**, 260602 (2019).
- [10] C. P. Royall, P. Charbonneau, M. Dijkstra, J. Russo, F. Smallenburg, T. Speck, and C. Valeriani, Colloidal hard spheres: Triumphs, challenges, and mysteries, *Rev. Mod. Phys.* **96**, 045003 (2024).
- [11] M. E. Cates and J. Tailleur, Motility-induced phase separation, *Annu. Rev. Condens. Matter Phys.* **6**, 219 (2015).
- [12] E. Fodor, C. Nardini, M. E. Cates, J. Tailleur, P. Visco, and F. van Wijland, How far from equilibrium is active matter? *Phys. Rev. Lett.* **117**, 038103 (2016).
- [13] R. J. Speedy, Statistical geometry of hard-sphere systems, *J. Chem. Soc. Faraday Trans. 2* **76**, 693 (1980).
- [14] R. J. Speedy, Cavities and free volume in hard-disc and hard-sphere systems, *J. Chem. Soc. Faraday Trans. 2* **77**, 329 (1981).
- [15] R. J. Speedy, Cavities in one-dimensional fluids, *J. Chem. Soc. Faraday Trans. 2* **78**, 1645 (1982).
- [16] R. J. Speedy and H. Reiss, Cavities in the hard sphere fluid and crystal and the equation of state, *Mol. Phys.* **72**, 999 (1991).
- [17] R. J. Speedy and H. Reiss, A computer simulation study of cavities in the hard disc fluid and crystal, *Mol. Phys.* **72**, 1015 (1991).
- [18] S. Sastry, T. M. Truskett, P. G. Debenedetti, S. Torquato, and F. H. Stillinger, Free volume in the hard sphere liquid, *Mol. Phys.* **95**, 289 (1998).
- [19] P. G. Debenedetti and T. M. Truskett, The statistical geometry of voids in liquids, *Fluid Phase Equilib.* **158–160**, 549 (1999).
- [20] R. K. Bowles and D. S. Corti, Statistical geometry of hard sphere systems: Exact relations for first-order phase transitions in multicomponent systems, *Mol. Phys.* **98**, 429 (2000).
- [21] A. P. Solon, Y. Fily, A. Baskaran, M. E. Cates, Y. Kafri, M. Kardar, and J. Tailleur, Pressure is not a state function for generic active fluids, *Nat. Phys.* **11**, 673 (2015).
- [22] F. Mori, P. Le Doussal, S. N. Majumdar, and G. Schehr, Universal properties of a run-and-tumble particle in arbitrary dimension, *Phys. Rev. E* **102**, 042133 (2020).
- [23] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/n9s4-ykn2> for the technical model details and mathematical details for analytical relations and algorithms; Supplementary figures containing data or information that are not essential but provide sanity checks or transparency in the methodology. Citations [1,2] in Supplemental Materials are [25,43], respectively.
- [24] B. Derrida, Microscopic versus macroscopic approaches to non-equilibrium systems, *J. Stat. Mech.: Theory Exp.* (2011) P01030.
- [25] R. Dandekar, S. Chakraborti, and R. Rajesh, Hard core run and tumble particles on a one-dimensional lattice, *Phys. Rev. E* **102**, 062111 (2020).
- [26] S. Sastry, D. S. Corti, P. G. Debenedetti, and F. H. Stillinger, Statistical geometry of particle packings. I. Algorithm for exact determination of connectivity, volume, and surface areas of void space in monodisperse and polydisperse sphere packings, *Phys. Rev. E* **56**, 5524 (1997).
- [27] A straightforward way of numerically implementing a two-dimensional repulsive RTP (lattice) model is the four-direction RTP model introduced in Ref. [44].
- [28] E. Specht, A precise algorithm to detect voids in polydisperse circle packings, *Proc. R. Soc. A* **471**, 20150421 (2015).
- [29] R. A. Dwyer, The expected number of k -faces of a Voronoi diagram, *Comput. Math. Appl.* **26**, 13 (1993).
- [30] T. M. Truskett, S. Torquato, S. Sastry, P. G. Debenedetti, and F. H. Stillinger, Structural precursor to freezing in the hard-disk and hard-sphere systems, *Phys. Rev. E* **58**, 3083 (1998).
- [31] P. Digregorio, D. Levis, A. Suma, L. F. Cugliandolo, G. Gonnella, and I. Pagonabarraga, 2D melting and motility induced phase separation in active Brownian hard disks and dumbbells, *J. Phys.: Conf. Ser.* **1163**, 012073 (2019).
- [32] B. J. Alder, W. G. Hoover, and T. E. Wainwright, Cooperative motion of hard disks leading to melting, *Phys. Rev. Lett.* **11**, 241 (1963).
- [33] X. Xu and S. A. Rice, Maximally random jamming of one-component and binary hard-disk fluids in two dimensions, *Phys. Rev. E* **83**, 021120 (2011).
- [34] W. G. Hoover, N. E. Hoover, and K. Hanson, Exact hard-disk free volumes, *J. Chem. Phys.* **70**, 1837 (1979).
- [35] S. C. Takatori and J. F. Brady, Towards a thermodynamics of active matter, *Phys. Rev. E* **91**, 032117 (2015).
- [36] P. Gaspard and R. Kapral, Active matter, microreversibility, and thermodynamics, *Research* **2020**, 9739231 (2020).
- [37] S. Cameron, M. Mosayebi, R. Bennett, and T. B. Liverpool, Equation of state for active matter, *Phys. Rev. E* **108**, 014608 (2023).
- [38] E. Schiltz-Rouse, H. Row, and S. A. Mallory, Kinetic temperature and pressure of an active Tonks gas, *Phys. Rev. E* **108**, 064601 (2023).
- [39] J. O'Byrne, Y. Kafri, J. Tailleur, and F. van Wijland, Time irreversibility in active matter, from micro to macro, *Nat. Rev. Phys.* **4**, 167 (2022).
- [40] J. U. Klamser, S. C. Kapfer, and W. Krauth, Thermodynamic phases in two-dimensional active matter, *Nat. Commun.* **9**, 5045 (2018).
- [41] S. H. L. Klapp, Non-reciprocal interaction for living matter, *Nat. Nanotechnol.* **18**, 8 (2023).
- [42] L. K. Davis, K. Proesmans, and E. Fodor, Active matter under control: Insights from response theory, *Phys. Rev. X* **14**, 011012 (2024).
- [43] E. W. Weisstein, Circle-circle intersection, MathWorld—A Wolfram Resource, <https://mathworld.wolfram.com/Circle-CircleIntersection.html>.
- [44] T. Pietrangeli, C. Ybert, C. Cottin-Bizonne, and F. Datchevy, Optimal run-and-tumble in slit-like confinement, *Phys. Rev. Res.* **6**, 023028 (2024).