Supplemental Information for "Informational active matter"

Bryan VanSaders[1] and Vincenzo Vitelli[1, 2, 3, *]

[1] *James Franck Institute, The University of Chicago, Chicago, Illinois 60637, USA*
[2] *Department of Physics, The University of Chicago, Chicago, Illinois 60637, USA*
[3] *Kadanoff Center for Theoretical Physics, The University of Chicago, Chicago, Illinois 60637, USA*

Contents

In §S1 the definition of informational activity is discussed along with several examples. In §S2 we demonstrate that the protocol used in Molecular Dynamics simulations of the thinker gas does not change the internal energy state of the system. Section §S3 describes how local measurements can drive non-equilibrium dynamics for a two-state Markov model. In §S4 a kinetic gas theory of the dilute pure thinker gas is developed in detail to accompany the abbreviated treatment in the main text methods. Section §S5 contains explicit functional forms for the moments of two thinker gases with special diameter functions. In §S6 the hydrodynamic consequences of the thinker kinetic gas theory are discussed. In §S7 compares pattern resolution for thinkers dropped from a hopper as a function of maximum thinker diameter. Finally, §S8 has descriptions of the two accompaning supplemental videos.

*Electronic address: vitelli@uchicago.edu

S1. DEFINING INFORMATIONAL ACTIVITY

We are concerned here with active media, many-body systems comprised of 'active' subunits. Multiple definitions of 'activity' can be found - here we will define an active subunit as one that has an internal process which biases its microstate transition probabilities (away from detailed balance, i.e. equilibrium). We focus on the distinction between two sub-classes of activity: 'mechanical' and 'informational' activity. Both varieties fulfil the transition biasing definition posed above, however their methods (and dependence on environmental noise) differ. Mechanical activity is familiar from many studies of self-propelled particles [1–13], and its key distinguishing feature is that subunits bias transitions by changing the free energy of the microstate *currently occupied* by the subunit. An iconic example is the active Janus particle, in which a local chemical reaction raises the free energy of the particle's current microstate in such a way that relaxation occurs by forward motion. Informational activity is the complementary strategy - biasing transitions by changing the free energy of *unoccupied* microstates which may become occupied in the future.

Many realistic active systems (particularly biological ones) will be driven by a mix of mechanical and informational activity. In order for informational activity to be possible, three key elements must be present:

1. **Noise**. Forces outside the subunit's control which act upon the d.o.f.s of interest are a requirement for informational activity. As only unoccupied microstates can be biased, informational activity does not directly drive transitions. Instead, some other source of agitation must be present.

2. **Ability to bias unoccupied microstates**. Microstates other than the one currently occupied by the subunit must be biased. Additionally, the applied bias must persist long enough for the subunit to interact with that microstate. Consider an unoccupied microstate, and a subunit about to transition into it. If a bias is applied, and immediately relaxes before the transition occurs, the bias has no effect and the microstate is now occupied.

3. **Measurement and feedback**. The capacity to measure the current microstate of the d.o.f.s of interest is central to informational activity. Unoccupied microstates can only be identified by determining the current subunit microstate. Furthermore, a feedback control loop must exist between the d.o.f.s of interest and the biasing of unoccupied states. If these two processes are decorrelated then biasing becomes merely another random process which is bound to obey detailed balance.

We now examine several key examples using these criteria.

S1.1. The Maxwell demon

In James Clerk Maxwell's iconic paradox [14], a 'demon' operates a trapdoor between two containers full of gas. By allowing only fast particles to cross from the left to right containers, and only slow to pass from right to left, a temperature gradient between the containers is produced. This happens despite the demon carefully opening or closing the trapdoor only when no gas particles are present at the junction between the containers. In this example, noise (in the form of a thermalized gas) plays a key role (requirement 1). The trapdoor fulfils requirement 2 - states where gas molecules occupy the doorway are not biased by closing the door. Rather, the door is closed preemptively to avoid an incoming particle crossing the wrong way. Feedback between the state of incoming gas particles and the state of the trapdoor fulfils requirement 3.

S1.2. Active Brownian particles

Active Brownian particles are self-propelled objects, usually smaller than a micron in size. In practice, they tend to operate by catalyzing a chemical reaction on half their surface, expansion of the reactants pushes them along. Such small objects exist at a scale where room temperature thermal fluctuations are important, satisfying requirement 1. However requirements 2 and 3 are not satisfied - particles have no capacity to measure their current microstate or modulate their chemically-derived force. Nor do they possess the capacity to bias unoccupied states (in the commonly considered overdamped regime).

S1.3. Vicsek-like flockers

The flocking of birds can be represented with a toy model of 'flying spins' in which particles have a polar variable, their heading, which interact similarly to magnetic dipoles [15]. Particles also typically have a fixed velocity magnitude

(imagined to be the result of self-propulsion) which points along their heading. Alignment interactions tend to draw nearby flockers into parallel trajectories. While flockers might be considered to 'measure' the heading of their neighbors, the resulting heading changes are deterministic processes, in which fluctuations play no role. When present, orientational noise is simply a disturbing influence that heading alignments fight against. Flocking models of this type are typically not informational because flockers drive transitions in their heading by forces acting directly on their current state. As such, their active properties are evident even in the absence of noise.

S1.4. RNA polymerase

The process of copying an RNA strand from template DNA is ubiquitous to cellular life, and is carried out by a complex molecular machine known as the RNA polymerase (RNAP). High fidelity [16] is important, as transcription errors could produce proteins that fail to fold, or are ineffective or harmful when folded. Like DNA polymerase, both selectivity and proofreading is employed to achieve high fidelity. Selectivity mechanisms are thought to closely resemble those employed by DNA [17, 18], while proofreading mechanisms are different. In RNAPs the active site where new nucleotides bind can be modulated. During elongation, this site facilitates the addition of incoming bases onto the growing RNA strand. For various reasons, RNAPs may undergo 'pausing' or 'backtracking', in which the progress of elongation is interrupted [19, 20]. Such events have been associated with expression regulation as well as error correction [21–23]. Single-molecule experiments on RNAPs have found evidence to support a 'Brownian ratchet' interpretation of the transcription process, whereby the polymerase moves down the template DNA stochastically and is rectified by the selectivity and proofreading processes [24].

The scale and stochastic behavior of RNAPs fulfill requirement 1 (noisy operation). RNAPs can be interpreted as biasing unoccupied states - specifically the states corresponding to incorrect base additions, which fulfils requirement 2. The ability of RNA polymerase to proofread (and to be selective for the correct base) are a form of measurement, where the conformation of the protein is changed according to an assessment of the template DNA strand or correct/incorrect state of the most recently added RNA base. Selectivity is relative to the pattern of the template DNA, a feedback process, fulfiling requirement 3.

S2. ENERGY CONSERVATION IN MOLECULAR DYNAMICS SIMULATIONS OF THINKERS

The thinker gas diameter changing protocol described in the main text methods is designed to avoid changing the potential or kinetic energy of the system (i.e. exert no mechanical work). As proof of the energy-conserving nature of the thinker gas numerical implementation, a test simulation using a constant-energy ensemble was prepared and the total energy of the system was monitored, shown in fig. S1. The implemented thinker gas algorithm does not inject or remove energy from the system at a rate greater than numerical precision effects under standard fixed-diameter dynamics.

S3. THERMODYNAMICS OF INFORMATION FOR ACTIVE MATTER

To understand the origin of activity in informational systems like the thinker gas, we will begin from basics by first describing the origin of mechanical activity in thermodynamic language. Consider an isothermal system which has had its Helmholtz free energy raised by an amount ΔF (perhaps as the result of an external field applied by a 'controller'). The controller must exert

$$W_{cont}^{in} \geq \Delta F \tag{S1}$$

work to accomplish this change, with saturation of the bound only for reversible processes with no heat transfer. Now that the system is prepared in a high-energy state it can perform work on some external load as it relaxes. This work is similarly

$$\mathcal{W}_{sys}^{out} \geq -\Delta F. \tag{S2}$$

where the sign of $\mathcal{W}_{sys}^{out}$ is negative. The controller exerts work W_{cont}^{in} so as to prepare the system to perform work $-\mathcal{W}_{sys}^{out} \leq W_{cont}^{in}$ on a load. Now consider replacing the controller with an active agent immersed in the system. If $\mathcal{W}_{sys}^{out}$ is used to move the active agent, then the above logic is unchanged except that the fate of $\mathcal{W}_{sys}^{out}$ is to be enventually converted completely to heat in dissipative processes.

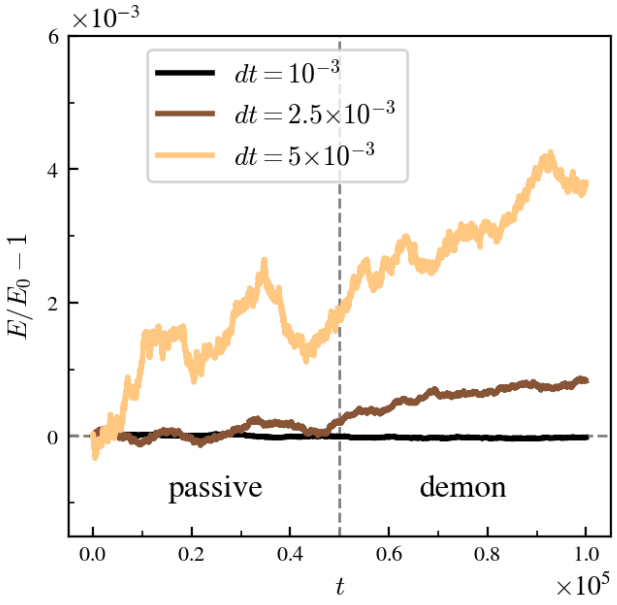

Fig. S1: **Total kinetic and potential energy of a numerically simulated thinker gas for various values of the time step parameter** dt**.** For times less than 5×10^4, the system is evolved under fixed-diameter dynamics with a constant-energy integrator. For $t > 5 \times 10^4$, diameter resizing is enabled. The diameter resizing procedure does not produce a detectable signature in the total energy drift of the system, which is controlled by the choice of dt.

When the active agent has the ability to collect measurements, equation S1 must be ammended with another term accounting for how information changes the entropy of the system [25],

$$W^{in}_{cont} + k_b T I \geq \Delta F. \tag{S3}$$

Where I is the mutual information of a measurement outcome and the microstate of the system. Even in the absence of work exerted by the agent, free energy differences in the system can be created (and therefore used to exert work on loads). Just as in the previous case, $\mathcal{W}^{out}_{sys} \geq -\Delta F$, yielding

$$\mathcal{W}^{out}_{sys} \geq -W^{in}_{cont} - k_b T I = -W^{in}_{cont} + T\Delta S \tag{S4}$$

where $\Delta S = -k_b I$ is the amount that the system's entropy has been reduced by the measurement. The work available to propel the active agent may be non-zero even for agents what exert no forces on the system that surrounds them.

When the action of an informational active agent is repeated (e.g. at intervals Δt), an additional complication arises where each new measurement outcome may be correlated with prior measurements. This dependence can be expressed as [26]:

$$\Delta S_t = -k_b \sum_t I(X_t; m_t | M_{0 \to t-\Delta t}) \tag{S5}$$

where X_t is the microstate measured at time t, m_t is the collected measurement, and $M_{0 \to t-\Delta t}$ is the history of measurements collected at prior times. The gained information is conditional on prior measurement outcomes - entropy is only lowered if a new measurement collects novel information about the system. When the measurements are error-free (i.e. they perfectly report the relevant microscopic information), eq. S5 can be expressed using the Shannon entropy of the measurement sequence:

$$\Delta S_t = -k_b H(M_{0 \to t}) \tag{S6}$$

where the Shannon entropy takes its usual form:

$$H(\mathcal{X}) = -\sum_{x\in\mathcal{X}} P(x)\ ln\ P(x) \tag{S7}$$

The entropy of the measurement history $M_{0\to t}$ is potentially complex to calculate for non-ergodic measurement ensembles. However our focus is on steady states of large collections of informational agents, so we will construct ergodic measurement ensembles with well-defined steady states. In this case the rate of entropy decrease per measurement by agent can be expressed as a limit:

$$\lim_{t\to\infty} \frac{\Delta S_t}{t} = \Delta\bar{S}. \tag{S8}$$

If the system can be represented as a steady Markov process, eq. S5 can be expressed as an entropy rate given by the weighted Shannon entropy of the relevant stochastic matrix:

$$\Delta\bar{S} = -k_b\bar{H}(M) = k_b\sum_{ij}\mu_i P_{ij}\ ln\ P_{ij} \tag{S9}$$

where $\mathbf{P}$ is the stochastic matrix that describes transitions between states [27] and μ is the accompanying stationary distribution.

To organize the distinction between mechanical and informational driving, it is useful to consider a mixed active process with a bound on the work it can exert on the system ($W < \epsilon$). Each time the active process acts, it takes a measurement to check if the action can be taken without violating the bound ϵ. Note that in cases where the mechanical work is capped at some value W_{max} (e.g. the maximum possible output torque of a motor), the bound only becomes relevant if $\epsilon < W_{max}$. Similarly, taking actions may require mechanical work (such as overcoming friction in a mechanism) that has no dependence on the particle or system's state. This kind of work is assumed to be small, or equivalently that the total bound includes an external and internal component, $\epsilon_{total} = \epsilon + \epsilon_{int}$. The 'internal' work bound covers the cost of internal inefficiencies of the active process, such as friction. Here we only consider variations of the 'external' work bound, which constrains the work done by the active process on the particle's, or system's, state. A particle with such a mixed process can be driven at most by

$$\bar{\mathcal{W}} \geq -\langle W\rangle_{(<\epsilon)} + k_b T\Delta\bar{S} \tag{S10}$$

where $\langle W\rangle_{(<\epsilon)}$ is the expected amount of mechanical work per action from actions that obey the bound ϵ. The mechanical contribution is not directly scaled by temperature, whereas the informational part grows (entropy changes due to measurement are always negative) with temperature, indicating that informational active agents grow stronger as fluctuations increase. This dependence on temperature is reminiscent of entropic elasticity [28–31], in which rubber-like materials grow stiffer as temperature is increased. To gain further insight into the work and entropy terms that make up eq. S10, we will now construct a Markov chain model of an active process with two discrete states.

S3.1. Two state Markov chain model of an informational active process

Consider a binary active processes, i.e. one that has only two states. This process is part of a measurement and feedback loop - the active process measures the system and determines if a change of state is warranted. We can describe this situation by focusing on the results of the measurement, the answer to the question "which state should I adopt?". This measurement m is binary valued ($m = 0$ or $m = 1$), and the answer fluctuates due to the noisy system. Measured values of $m = \{0, 1\}$ correspond to control actions, $c = \{0, 1\}$ - e.g. when a measurement $m = 0$ is collected, the control action $c = 0$ is performed on the system. Therefore, measured answers must respect the bound imposed by ϵ. If we represent the dynamics of the measurement variable m with a Markov chain, and focus on the long time limit in which a stationary distribution μ has been reached, we can compute the entropy rate directly using eq. S9. Let us start by specifying a simple but generic form of the transition matrix $\mathbf{P}$,

$$\mathbf{P} = \begin{bmatrix} 1-\alpha\Sigma & \alpha\Sigma \\ \beta\Sigma & 1-\beta\Sigma \end{bmatrix} \tag{S11}$$

where $\alpha\Sigma$ is the probability of the measured variable changing from $m_t = 0$ to $m_{t+\Delta t} = 1$, and $\beta\Sigma$ is the probability of the reverse transition of $m_t = 1$ to $m_{t+\Delta t} = 0$. The terms β and α represent the transition outcomes of some

stochastic process in the surrounding system, while Σ is the likelihood that taking the control action associated with measurement m will respect the work bound ϵ. Said another way, the controller would implement the control action $c = 1$ at a time $t + \Delta t$, given a prior state $m = 0$ at time t, with probability α. However the bound ϵ only permits this with probability Σ. For the transition matrix defined in eq. S11, the stationary distribution is

$$\mu = \left[\frac{\beta}{\beta+\alpha} \quad \frac{\alpha}{\beta+\alpha} \right], \tag{S12}$$

in which Σ has dropped out due to its symmetry.

Take as an instructive example the case of $\alpha = 1/2$, $\beta = 1/2$, $\Sigma = 1$. These parameters describe a delta-correlated random process (coin flips), in which the work bound never restricts the controller, and the entropy rate of this process reduces to $ln\ 2$. Generally, when $\alpha, \beta \neq 1/2$ longer-time correlations exist in the measurement stream - the value of m_{t+dt} becomes conditional on m_t. As we saw before in eq. S5, this reduces the amount of new information in each measurement.

We can gain further understanding by choosing a form for the probability of an action respecting the work bound, Σ. Consider a process where switching from control action $c = 0$ to $c = 1$ (or vice versa) dissipates an amount of work, with a probability distribution given by

$$P(W) = (1 - P_0)\frac{1}{\sigma}\sqrt{\frac{2}{\pi}}e^{-\frac{|W|^2}{\sigma^2}} + P_0\delta(0), \tag{S13}$$

where σ parameterizes the fluctuations of the system that affect expended work, δ is the impulse function ($\int \delta = 1$) and P_0 is the probability that the action costs zero work. The likelihood of this work falling below the bound set by ϵ is the integral:

$$\Sigma = \int_0^\epsilon P(W)dW = (1 - P_0)\ erf\left(\frac{\epsilon}{\sigma}\right) + P_0 \tag{S14}$$

where $erf()$ is the standard error function. We can also now write down the form of the mechanical work term in eq. S10:

$$\langle W \rangle_{(<\epsilon)} = \int_0^\epsilon W(P(W) - P_0\delta(0))dW = \frac{(1 - P_0)\sigma}{\sqrt{2\pi}}\left(1 - e^{-\frac{\epsilon^2}{\sigma^2}}\right) \tag{S15}$$

and so obtain a full expression for the bound of work done by the system per action on the active agent:

$$\bar{\mathcal{W}} \geq -\frac{(1 - P_0)\sigma}{\sqrt{2\pi}}\left(1 - e^{-\frac{\epsilon^2}{\sigma^2}}\right) + k_bT\bar{H}(M). \tag{S16}$$

For $\alpha = \beta = 1/2$, the terms of this equation are plotted in fig. S2 as a function of the ratio of the mechanical work bound to the work fluctuations (ϵ/σ). Note that σ can be a function of the temperature of the system, as well as other parameters of interaction. Consider a case where actions are prohibited by obstacles in the environment - σ will depend upon the fluctuations of those obstacles (i.e. k_bT) as well as parameters such as their mass or mechanical stiffness.

We see that the entropy-like contribution to the non-equilibrium free energy as $\epsilon \to 0$ is:

$$-k_bT\bar{H}(P_0\alpha) = k_bT\left(P_0\alpha\ ln\ P_0\alpha + (1 - P_0\alpha)\ ln\ (1 - P_0\alpha)\right), \tag{S17}$$

which is an entropy associated with the probability that actions will require zero work. Conversely, as $\epsilon \to 0$ the contribution from mechanical work approaches zero. In the limit of large ϵ, the informational contribution approaches $-k_bT\alpha = k_bT\ ln\ 2$, losing any dependence on the statistics of the constraint. The mechanical contribution approaches a value set by the details of σ. In the example shown here, $\sigma = 1$, however arbitrarily large values are possible. Despite this, the mechanical contribution will always equal zero as $\epsilon \to 0$

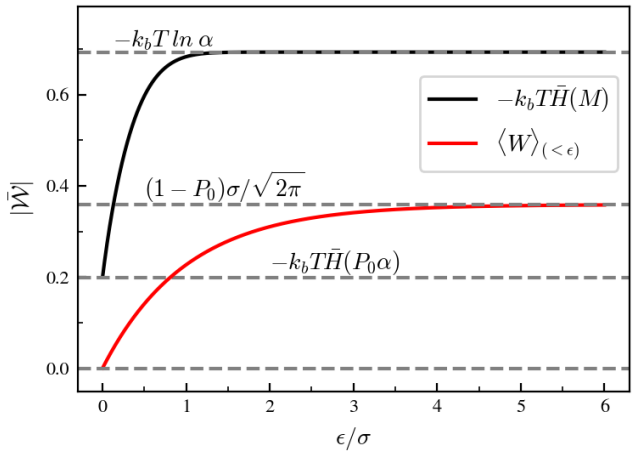

Fig. S2: **Propulsive power bounds for a mixed informational-mechanical Markov process**. A binary variable, updated by a simple Markov process that includes the effect of a mechanical work constraint, ϵ. Here σ is a variance associated with mechanical work exerted by the process, and P_0 is the likelihood of an action requiring no work. The transition probabilities are kept fixed at $\alpha = \beta = 1/2$. The total available power has an informational (black) and mechanical (red) contribution. In the limit of large ϵ/σ the informational process approaches the entropy of unconstrained transitions, and the mechanical work approaches a value controlled by the variation in how much mechanical work actions require (σ). Note that at finite ϵ the mechanical contribution may far exceed the informational contribution depending on the details of σ, however as $\epsilon \to 0$ the mechanical contribution always approaches zero. In contrast, informational driving only approaches zero as $\epsilon \to 0$ if $P_0 = 0$.

S4. THE THINKER GAS VELOCITY DISTRIBUTION FUNCTION

In order to derive an approximate description of the thinker gas velocity distribution function, we will outline key points from the derivation of the Boltzmann equation which describes the evolution of the same distribution for a hard sphere gas. More information on the derivation of the standard Boltzmann equation and its mathematical properties can be found in texts by Cercignani et al [32] and Kardar [33]. From there, the treatment is extended to the case of hard disks with non-constant diameters (which are functions of particle velocity only). We show the form of the collision operator which describes the thinker gas, and demonstrate that this operator has a modified collision invariant expression which includes the diameter function $D(\vec{v})$. We use this property to define a linear system of integral equations which permits solving for the deviation of the thinker gas velocity distribution function from the equilibrium Maxwell-Boltzmann (MB) distribution.

S4.1. Kinetic theory of hard disks

The evolution of a gas in phase space under the influence of interparticle collisions can be described by the BBGKY hierarchy. In this set of coupled differential equations the evolution of N-particle distribution functions (denoted $P^{(N)}$) are dependent upon the dynamics of $N+1$-particle distributions. The full hierarchy is in general too complex for useful prediction, and so it is often truncated to only include one and two-particle distributions. The single-particle distribution function ($P^{(1)}$) is obtained from the full system distribution function of N particles by marginalization:

$$P^{(1)}(\vec{x}_1, \vec{v}_1, t) = \int_{\mathcal{R}^{2N-2}} P(\vec{x}_1, \vec{v}_1, \vec{x}_2, \vec{v}_2, ..., \vec{x}_N, \vec{v}_N, t) d\vec{x}_1 d\vec{v}_1, ...d\vec{x}_N d\vec{v}_N \tag{S18}$$

This is the probability of finding a particle within the interval $[\vec{x}_1 + d\vec{x}, \vec{v}_1 + d\vec{v}_1, t + dt]$. The evolution of $P^{(1)}$ in the absence of collisions proceeds only by the streaming of particles along straight trajectories ($\partial P^{(1)}/\partial t = -\vec{v} \cdot \nabla_{\vec{x}} P^{(1)}$). Collisions between particles introduce a new operator to the equation of motion of the single-particle distribution:

$$\frac{\partial P^{(1)}}{\partial t} + \vec{v} \cdot \nabla_{\vec{x}} P^{(1)} = \mathcal{Q} \tag{S19}$$

$\mathcal{Q}$ describes the rate of change of particle number (or density) within the interval $[\vec{x} + d\vec{x}, \vec{v} + d\vec{v}, t + dt]$ caused by collisions. For hard spheres, collisions occur only when two particles are found in contact. Naturally, $\mathcal{Q}$ must

involve the two particle distribution, $P^{(2)}$ (the probability of finding two particles in a given configuration). If we denote pre-collision velocities for a pair of particles as $\vec{v}$ and $\vec{v}_*$, and their post-collision velocities as $\vec{v}'$ and $\vec{v}'_*$, we can write the collision operator as the difference of two rates, the gain and loss of particles within the interval of interest ($\mathcal{Q} = \mathcal{Q}^+ - \mathcal{Q}^-$). Particles are removed from the interval by collisions (velocities are transformed from $\vec{v}, \vec{v}_*$ to $\vec{v}', \vec{v}'_*$). The likelihood of a collision is the same as the likelihood of finding a two-particle configuration in contact. For particles of fixed diameter (D_0):

$$\mathcal{Q}^- = \frac{1}{2}\int_{\mathcal{R}^2}\int_{\mathcal{S}} P^{(2)}(\vec{x}, \vec{v}, \vec{x} + D_0\hat{n}, \vec{v}_*, t) d\vec{v}_* d\hat{n} \tag{S20}$$

here the variable $\hat{n}$ is a unit vector that lies along the particle contact vector and integration over $d\hat{n}$ occurs on the surface $\mathcal{S}$ of all possible contacts. For a gas in a periodic domain (homogeneous spatial density) the spatial aspect of $P^{(2)}$ can be approximated geometrically. A rectangular area with length $|(\vec{v}_* - \vec{v}) \cdot \hat{n}|dt$ and width $|D_0 d\hat{n}| = D_0$ contains the particles with velocity $\vec{v}_*$ which are capable of striking the first particle. The probability density of finding two particles with velocities $\vec{v}, \vec{v}_*$ in contact can therefore be written as:

$$\mathcal{Q}^- = \frac{1}{2}\int_{\mathcal{R}^2}\int_{\mathcal{S}^-} P^{(2)}(\vec{v}, \vec{v}_*, t)|(\vec{v}_* - \vec{v}) \cdot \hat{n}|D_0 d\vec{v}_* d\hat{n} \tag{S21}$$

where now location variables in $P^{(2)}$ have been dropped due to the assumption of spatial homogeneity. The inner integration is carried out over half the unit disk ($\mathcal{S}^-$, the half for which velocities $\vec{v}$ and $\vec{v}_*$ are approaching each other). This loss term is balanced (at steady state) by a similar term describing the gain of probability density by scattering from other regions of velocity space. The only differences being that the domain $\mathcal{S}^-$ is replaced by $\mathcal{S}^+$, the complementary hemisphere, and the velocities in consideration are $\vec{v}', \vec{v}'_*$, velocities which will undergo a scattering into $\vec{v}, \vec{v}_*$. As in the standard treatment, we will use the 'molecular chaos' assumption to replace two particle distribution terms with the product of single particle terms $P^{(2)}(\vec{v}, \vec{v}_*, t) \approx P^{(1)}(\vec{v}, t)P^{(1)}(\vec{v}_*, t)$. In the standard Boltzmann equation, the gain and loss terms are collected together under common integration limits by exploiting the time reversal symmetry of the collision process. Since we will now expand our treatment to include gases with diameters which are not constant in velocity space, we must examine the reversibility of hard disk collisions more closely.

S4.2. Hard disk collisions with non-constant diameter

We can find the velocities after collision between a pair of hard disks by conservation of momentum and energy. Assuming uniform particle masses, momentum and kinetic energy are conserved through the collision process:

$$\vec{v} + \vec{v}_* = \vec{v}' + \vec{v}'_* \tag{S22}$$

$$|\vec{v}|^2 + |\vec{v}_*|^2 = |\vec{v}'|^2 + |\vec{v}'_*|^2 \tag{S23}$$

This collision process can be viewed as the exchange of velocities projected onto the vector which points from one disk's center to the other. If $\hat{n}$ is a unit vector that points between disk centers at collision contact, then the post-collision velocities can be written as (for disks of equal mass):

$$\vec{v}' = \vec{v} - ((\vec{v} - \vec{v}_*) \cdot \hat{n})\hat{n} \tag{S24}$$

$$\vec{v}'_* = \vec{v}_* + ((\vec{v} - \vec{v}_*) \cdot \hat{n})\hat{n} \tag{S25}$$

For disks of fixed diameter, this scattering process is microreversible; $(\vec{v}, \vec{v}_*)$ map to $(\vec{v}', \vec{v}'_*)$ and $(-\vec{v}', -\vec{v}'_*)$ map to $(-\vec{v}, -\vec{v}_*)$, resulting in a Jacobian for the change of variables equal to -1 [32]. This one-to-one mapping allows all four velocity terms to be collected under the same unit disk integration in the standard Boltzmann collision operator.

When particle diameters are instead permitted to be functions of velocity, the microreversibility of the collision process is broken. We can see this diagrammatically in Fig. S3a. In this representation, we view the two particle collision process as the scattering of a point object with velocity $\vec{v}_* - \vec{v}$ off of a stationary disk with diameter $(D(\vec{v}) + D(\vec{v}_*))/2$. By changing the location of the stationary disk, the mapping of velocities can be preserved (fig. S3b). In a coordinate frame centered on the point of contact, the forward and reverse scattering process are symmetric provided that the coordinates of the stationary particle are scaled by a factor $(D(\vec{v}') + D(\vec{v}'_*))/(D(\vec{v}) + D(\vec{v}_*))$. The collision process can also be represented with the so-called 'impact parameter' (σ) as the integration variable (as opposed to the vector $\hat{n}$). Fig. S3c shows that a rescaling of the impact parameter has an equivalent effect of

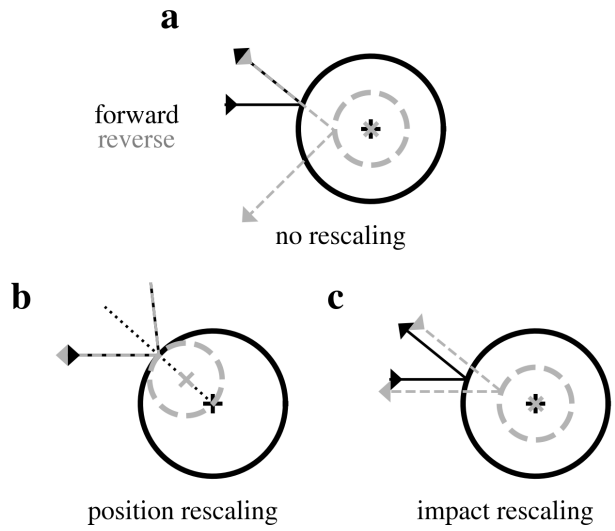

Fig. S3: **The microreversibility of hard disk collisions with velocity-dependent diameters.** Two particles of diameter $D(\vec{v}) = D$ and $D(\vec{v}_*) = D_*$ with velocities $\vec{v}$ and $\vec{v}_*$ are shown as the collision of a point object (with velocity $\vec{v}_* - \vec{v}$) with a stationary disk of diameter $D + D_*$. **a** In the coordinate frame centered on the stationary disk, the forward (black) and reverse (grey) collisions are not symmetric. **b** In the coordinate frame centered on the point of contact forward and reverse collision processes are symmetric, provided that the stationary disk's position is rescaled by a factor $(D' + D'_*)/(D + D_*)$ **c** The symmetry of the collision can also be preserved in the frame centered on the stationary disk by rescaling the impact parameter of the incoming particle.

preserving the velocity mapping. Naturally, the choice of rescaling one particle's position or the other is arbitrary and should produce equivalent results.

This mapping defines the procedure for exchanging unprimed and primed velocity variables in the collision operator of the thinker gas. From now on, we will refer to the single particle velocity distribution function as $f(\vec{x}, \vec{v}, t)$, and suppress the location and time arguments as we are concerned with the homogeneous steady-state solution. For brevity, functions of velocity will have their argument suppressed but retain distinguishing marks (such as $f(\vec{v}'_*) = f'_*$ and $D(\vec{v}'_*) = D'_*$). In this notation, we write the loss term of the thinker gas collision operator as:

$$\mathcal{Q}^- = \frac{1}{2}\int_{\mathcal{R}^2}\int_{\mathcal{S}^-}(D + D_*)ff_*|(\vec{v}_* - \vec{v})\cdot\hat{n}|d\vec{v}_*d\hat{n} \tag{S26}$$

The gain term can be expressed by the exchange of unprimed velocities for primed ones. The term $\vec{v}_* - \vec{v} = \vec{v}'_* - \vec{v}'$ is conserved by the collision process, and so need not be transformed.

$$\mathcal{Q}^+ = \frac{1}{2}\int_{\mathcal{R}^2}\int_{\mathcal{S}^+}(D' + D'_*)f'f'_*|(\vec{v}_* - \vec{v})\cdot\hat{n}|d\vec{v}'_*d\hat{n} \tag{S27}$$

These terms can be collected under a common integration, as the diameter function scaling acts a Jacobian term. The form of $\mathcal{Q}$ shown in the text is achieved by redistribution of diameter terms:

$$\mathcal{Q} = \mathcal{Q}^+ - \mathcal{Q}^- = \frac{1}{2}\int_{\mathcal{R}^2}\int_{\mathcal{S}}(D + D_*)|(\vec{v}_* - \vec{v})\cdot\hat{n}|\left[\frac{D' + D'_*}{D + D_*}f'f'_* - ff_*\right]d\vec{v}_*d\hat{n} \tag{S28}$$

Notice that the resulting collision operator is the sum of two terms that differ only by D and D_*. We will next consider the invariants of the thinker gas collision operator.

S4.3. The thinker gas collision invariant

We will first discuss collision invariants for the standard Boltzmann equation, and then explore the consequences of our modified thinker gas collision operator. If we consider a moment of a distribution function $f(\vec{v}, t)$ which is a solution to the Boltzmann equation with a non-time-varying test function ϕ:

$$\langle\phi\rangle_f = \int_{\mathcal{R}^d} f(\vec{v},t)\phi(\vec{v})d\vec{v} \tag{S29}$$

the evolution of this quantity is:

$$\frac{d}{dt}\langle\phi\rangle_f = \int_{\mathcal{R}^2} \phi(\vec{v})\mathcal{Q}(f)(\vec{v},t)d\vec{v} \tag{S30}$$

The function ϕ can be brought into the integrand of the collision operator since it is not a function of time. As we are now considering an expression integrating over all $\vec{v}$ and $\vec{v}_*$ (and by extension $\vec{v}'$ and $\vec{v}'_*$), the choice of primed and starred variables is entirely arbitrary. A standard procedure is to average over the exchange of velocity variables [32]. We first express the right side of eq. S30 (with $\mathcal{Q}$ appropriate for constant disks, and dropping velocity arguments where obvious) as:

$$\frac{1}{2}\int_{\mathcal{R}^2}\int_{\mathcal{R}^2}\int_{\mathcal{S}} \phi\left(f'f'_* - ff_*\right)|(\vec{v}_* - \vec{v})\cdot\hat{n}|D_0 d\vec{v}_* d\vec{v} d\hat{n} \tag{S31}$$

The exchange of starred and unstarred variables results in no change to the expression, allowing the average over this exchange to be written as:

$$\frac{1}{4}\int_{\mathcal{R}^2}\int_{\mathcal{R}^2}\int_{\mathcal{S}} (\phi+\phi_*)\left(f'f'_* - ff_*\right)|(\vec{v}_* - \vec{v})\cdot\hat{n}|D_0 d\vec{v}_* d\vec{v} d\hat{n} \tag{S32}$$

Exchanging primed variables (which can be done because of the microreversibility of the constant-diameter collision process), results in a change of sign. Therefore the result of averaging over both exchanges is:

$$\frac{1}{8}\int_{\mathcal{R}^2}\int_{\mathcal{R}^2}\int_{\mathcal{S}} (\phi+\phi_*-\phi'-\phi'_*)\left(f'f'_* - ff_*\right)|(\vec{v}_* - \vec{v})\cdot\hat{n}|D_0 d\vec{v}_* d\vec{v} d\hat{n} \tag{S33}$$

If ϕ is a quantity which is conserved through the collision, then $\frac{d}{dt}\langle\phi\rangle_f = 0$. This defines a summational collisional invariant for the system. For the Boltzmann equation, summational invariants can only be linear combinations of the microscopically conserved quantities, $\mathcal{M} = 1$, $\vec{v}$, $|\vec{v}|^2$, which are unchanged during collision by construction [34]. Therefore there exist constants $(a,\vec{b},c) \in \mathcal{R}\times\mathcal{R}^2\times\mathcal{R}$ such that:

$$\mathcal{M}(\vec{v}) = a + \vec{b}\cdot\vec{v} + c|\vec{v}|^2 \tag{S34}$$

The preceding discussion has been for the standard Boltzmann collision operator. We can follow the same logic for the thinker gas collision operator described in eq. S28. If we temporarily shorthand the gain term in the collision operator as $G = (f'f'_*)(D'+D'_*)/(D+D_*)$ and the loss term as $L = (ff_*)$, the result of averaging over primed and unprimed variables for the thinker gas can be written as:

$$\frac{1}{8}\int_{\mathcal{R}^2}\int_{\mathcal{R}^2}\int_{\mathcal{S}} (\phi(D+D_*)+\phi_*(D+D_*)-\phi'(D'+D'_*)-\phi'_*(D'+D'_*))\,(G-L)\,|(\vec{v}_* - \vec{v})\cdot\hat{n}|d\vec{v}_* d\vec{v} d\hat{n} \tag{S35}$$

where the leading term involving the function ϕ can be written as:

$$\phi(D+D_*)+\phi_*(D+D_*)-\phi'(D'+D'_*)-\phi'_*(D'+D'_*) = (\phi D+\phi_* D_* - \phi' D' - \phi'_* D'_*)+(\phi D_* + \phi_* D - \phi' D'_* - \phi'_* D'). \tag{S36}$$

If the entirety of eq. S36 is zero at all points in the velocity plane, then the quantity ϕ is stationary on the thinker distribution. The second term on the r.h.s. of eq. S36 contains only cross terms between starred and unstarred variables, and so all terms within it will integrate to the same value and the total term reduces to zero. An expression for the time evolution of a test function ϕ for the thinker gas can therefore be written with only the first term in the r.h.s. of eq. S36 as:

$$\frac{d}{dt}\langle\phi\rangle_{f_t} = -\frac{1}{2}\int_{\mathcal{R}^2}\int_{\mathcal{R}^2}\int_{\mathcal{S}} D\phi\left(f'f'_*\frac{D'+D'_*}{D+D_*} - ff_*\right)|(\vec{v}_*-\vec{v})\cdot\hat{n}|d\vec{v}_*d\vec{v}d\hat{n} \tag{S37}$$

From this representation, we can define a new test function, $g = D(\vec{v})\phi$. In order for the function ϕ to be stationary under the action of the thinker collision operator, the quantity g must be a collisional invariant (unlike the case for fixed-diameter disks, where ϕ itself must be invariant). For a collisionally invariant g there must exist constants which express it as a linear combination of microscopically conserved quantities: $g = a + \vec{b}\cdot\vec{v} + c|\vec{v}|^2 = D\mathcal{M}$. Note that an invariant test function for the MB distribution does not generically produce an invariant g for the thinker collision operator, due to the diameter rescaling. We will use this relation to define the velocity distribution function of the thinker gas. Beginning with the MB velocity distribution function in 2D:

$$f_m(\vec{v}) = \frac{m\rho}{2\pi kT}e^{-\frac{m|\vec{p}|^2}{2kT}} \tag{S38}$$

where $\vec{p} = \vec{v} - \langle\vec{v}\rangle = \vec{v} - \vec{u}$ is the velocity of the gas in the flow frame. We express the thinker gas distribution function as $f_t = (1+\mathcal{D})f_m$, where $\mathcal{D}$ is a function of velocity only. Since we seek the steady-state form of $\mathcal{D}$, we can treat $(1+\mathcal{D})$ as a non-time-varying test function operating on the MB distribution. We are searching for the form of this test function that has stationary expectation values on the MB velocity distribution under the action of the thinker gas collision operator. By substituting $\langle(1+\mathcal{D})\rangle_{f_m,\,\mathcal{Q}_d}$ for $\langle\phi\rangle_f$ in eq. S37, we see that steady state will be reached when $D(1+\mathcal{D})$ is a summational collision invariant. Finding the form of $\mathcal{D}$ amounts to specifying the four constants a, b_x, b_y, c which define the linear mixing of microscopically conserved quantities. These constants can be found if we require that the first four moments of the thinker gas velocity distribution are equal to the first four moments of the equilibrium MB distribution.

This requirement reflects the fact that thinker gas collisions microscopically conserve number, momentum, and kinetic energy, just as elastic collisions between fixed diameter disks. This requires that any re-sizing of the disks occurs between collisions. From a theoretical point of view, the gas is sparse enough to ignore the cases in which a diameter change post-collision must be delayed until the disks have sufficiently separated. This limitation on the ability of the disks to change their size is likely the most significant contributor to deviations from the theory at high densities or large maximum diameters.

The requirement that the first four moments of the thinker gas distribution are equal to those of the MB and can be written as:

$$\begin{aligned}\int_{\mathcal{R}^2}\mathcal{M}f_m d\vec{v} &= \int_{\mathcal{R}^2}\mathcal{M}(1+\mathcal{D})f_m d\vec{v}\\ 0 &= \int_{\mathcal{R}^2}\mathcal{M}\mathcal{D}f_m d\vec{v}\end{aligned} \tag{S39}$$

which highlights that the expectation value of the deviation of the thinker distribution ($\mathcal{D}$) must be zero for the first four moments. Note that this restriction does not constrain the third-order moments which define the heat flux tensor, or even the individual entries of the pressure tensor (derived from second-order moments). Only the trace of the pressure tensor is fixed relative to the MB distribution.

Using the expression for the thinker gas collision invariant, $g = D(1+\mathcal{D})$, we can trivially reorder to obtain $(1+\mathcal{D}) = gD^{-1}$. As a scalar function which must be greater than zero, division by D presents no problems. In practice, D need not even be non-zero everywhere, but instead only where the probability density of f_m is not infinitesimal. Using this expression for the thinker perturbation, and the notation for expectation value $\langle\phi\rangle_m = \int\phi f_m d\vec{v}$, we can write:

$$\begin{aligned}\langle\mathcal{M}\rangle_m &= \langle\mathcal{M}(1+\mathcal{D})\rangle_m\\ &= \langle\mathcal{M}gD^{-1}\rangle_m\\ &= a\langle\mathcal{M}D^{-1}\rangle_m + b_x\langle\mathcal{M}v_xD^{-1}\rangle_m + b_y\langle\mathcal{M}v_yD^{-1}\rangle_m + c\langle\mathcal{M}|\vec{v}|^2D^{-1}\rangle_m\end{aligned} \tag{S40}$$

For each of the conserved moments ($\mathcal{M} = 1,\ v_x,\ v_y,\ |\vec{v}|^2$) a new expression is found, allowing for the definition of a full rank system of linear equations:

$$\mathcal{A}\cdot G=\mathcal{B} \tag{S41}$$

$$\mathcal{A}=\begin{bmatrix}\langle\frac{1}{D}\rangle_m & \langle\frac{v_x}{D}\rangle_m & \langle\frac{v_y}{D}\rangle_m & \langle\frac{|\vec{v}|^2}{D}\rangle_m\\ \cdot & \langle\frac{v_x^2}{D}\rangle_m & \langle\frac{v_xv_y}{D}\rangle_m & \langle\frac{v_x|\vec{v}|^2}{D}\rangle_m\\ \cdot & \cdot & \langle\frac{v_y^2}{D}\rangle_m & \langle\frac{v_y|\vec{v}|^2}{D}\rangle_m\\ \cdot & \cdot & \cdot & \langle\frac{|\vec{v}|^4}{D}\rangle_m\end{bmatrix} \tag{S42}$$

$$G=\begin{bmatrix}a & b_x & b_y & c\end{bmatrix}^\top \tag{S43}$$

$$\mathcal{B}=\begin{bmatrix}\rho & \rho u_x & \rho u_y & \rho|\vec{u}|^2/2+\rho kT\end{bmatrix}^\top \tag{S44}$$

where $\mathcal{A}$ is symmetric and redundant entries are not shown. Using the solution of this system of equations the thinker gas velocity distribution can be written as:

$$f_t=f_m(1+\mathcal{D})=f_mD^{-1}g=f_mD^{-1}(\mathcal{A}^{-1}\mathcal{B})\cdot G \tag{S45}$$

S4.4. Demonstration of f_t as the steady-state solution to the variable-diameter collision operator

We can show directly that the above procedure for finding a thinker gas distribution function solves $\mathcal{Q}(f_t)=0$ by examining the collision operator written as:

$$\mathcal{Q}=\frac{1}{2}\int_{\mathcal{R}^2}\int_{\mathcal{S}}B(\vec{v}_*,\vec{v},\hat{n})\left[(D'+D'_*)f'f'_*-(D+D_*)ff_*\right]d\vec{v}_*d\hat{n} \tag{S46}$$

in which the cross-section terms have been redistributed and the kernel term has been collected into a function B for brevity. Focusing on the term in brackets (Q), we can replace f with $f_t=D(1+\mathcal{D})f_m$:

$$Q\equiv(D'+D'_*)(1+\mathcal{D}')f'_m(1+\mathcal{D}'_*)f'_{m*}-(D+D_*)(1+\mathcal{D})f_m(1+\mathcal{D}_*)f_{m*} \tag{S47}$$

First, note that the Maxwellian distribution is the exponential of a collision invariant, or:

$$ln\ f_m+ln\ f_{m*}-ln\ f'_m-ln\ f'_{m*}=0 \tag{S48}$$

$$\frac{f_mf_{m*}}{f'_mf'_{m*}}=1 \tag{S49}$$

Equation S49 allows for eq. S47 to be written as:

$$\begin{aligned}Q=\ & f_mf_{m*}\left[D'(1+\mathcal{D}')+\mathcal{D}'_*D'(1+\mathcal{D}')+D'_*(1+\mathcal{D}'_*)+\mathcal{D}'D'_*(1+\mathcal{D}'_*)\right]\\ &+f_mf_{m*}\left[D(1+\mathcal{D})-\mathcal{D}_*D(1+\mathcal{D})-\mathcal{D}_*(1+\mathcal{D}_*)-\mathcal{D}D_*(1+\mathcal{D}_*)\right]\end{aligned} \tag{S50}$$

Recall that the quantity $g=D(1+\mathcal{D})$ is also a collision invariant. We simplify eq. S50 to:

$$Q=f_mf_{m*}\left[g'+g'_*-g-g_*\right]+f_mf_{m*}\left[g'\mathcal{D}'_*+g'_*\mathcal{D}'-g\mathcal{D}_*-g_*\mathcal{D}\right] \tag{S51}$$

The first term in eq. S51 is zero, by the definition of collision invariants. The second term is also equal to zero, because each of the four terms individually integrate to zero. The logic for each term is identical, so we will focus on $f_mf_{m*}g\mathcal{D}_*$:

$$\begin{aligned}q(g,\mathcal{D}_*)\ &\equiv-\tfrac{1}{2}\textstyle\int_{\mathcal{R}^2}\int_{\mathcal{S}}B(\vec{v}_*,\vec{v},\hat{n})f_mf_{m*}g\mathcal{D}_*d\vec{v}_*d\hat{n}\\ &=-\tfrac{1}{2}gf_m\textstyle\int_{\mathcal{R}^2}\int_{\mathcal{S}}B(\vec{v}_*,\vec{v},\hat{n})\mathcal{D}_*f_{m*}d\vec{v}_*d\hat{n}\end{aligned} \tag{S52}$$

Recall that the collision kernel B for hard disk collisions is a linear function of relative velocity. Therefore, eq. S52 is a first-order moment of $\mathcal{D}f_m$, which must equal zero by construction (see eq. S39). The other terms in eq. S51 can all be likewise shown to equal zero by exchange of velocity variables and use of eq. S49.

S4.5. Invertibility of the matrix $\mathcal{A}$

$\mathcal{A}$ can be written as the element-wise expectation over $f_m D^{-1}$ of the symmetric matrix:

$$\mathbb{E}_{f_m D^{-1}} \begin{bmatrix} 1 & v_x & v_y & |\vec{v}|^2 \\ . & v_x^2 & v_x v_y & v_x|\vec{v}|^2 \\ . & . & v_y^2 & v_y|\vec{v}|^2 \\ . & . & . & |\vec{v}|^4 \end{bmatrix} \tag{S53}$$

This matrix can be written as the outer product of two vectors (therefore it has rank one):

$$\begin{bmatrix} 1 & v_x & v_y & |\vec{v}|^2 \end{bmatrix}^\top \begin{bmatrix} 1 & v_x & v_y & |\vec{v}|^2 \end{bmatrix} \tag{S54}$$

Despite this, $\mathcal{A}$ is generically invertible. The element-wise expectation promotes the matrix to full rank. If we temporarily adopt the notation of $\langle\phi\rangle = \int \phi D^{-1} f_m d\vec{v}$, $\mathcal{A}$ can be written as:

$$\begin{bmatrix} \langle 1\rangle & \langle v_x\rangle & \langle v_y\rangle & \langle|\vec{v}|^2\rangle \\ \langle v_x\rangle & \langle v_x^2\rangle & \langle v_x v_y\rangle & \langle v_x|\vec{v}|^2\rangle \\ \langle v_y\rangle & \langle v_x v_y\rangle & \langle v_y^2\rangle & \langle v_y|\vec{v}|^2\rangle \\ \langle|\vec{v}|^2\rangle & \langle v_x|\vec{v}|^2\rangle & \langle v_y|\vec{v}|^2\rangle & \langle|\vec{v}|^4\rangle \end{bmatrix} \tag{S55}$$

The determinant of this matrix can be found by cofactor expansion along the first column. If we examine the determinant of the first of the four 3×3 matrices which make up $\mathcal{A}$, we find that a term composed of the difference of permutations of expectation values is produced:

$$\langle 1\rangle \begin{vmatrix} \langle v_x^2\rangle & \langle v_x v_y\rangle & \langle v_x|\vec{v}|^2\rangle \\ \langle v_x v_y\rangle & \langle v_y^2\rangle & \langle v_y|\vec{v}|^2\rangle \\ \langle v_x|\vec{v}|^2\rangle & \langle v_y|\vec{v}|^2\rangle & \langle|\vec{v}|^4\rangle \end{vmatrix} \tag{S56}$$

$$= \langle 1\rangle \left(\langle v_x^2\rangle\langle v_y^2\rangle - \langle v_x v_y\rangle^2\right)\langle|\vec{v}|^4\rangle$$

$$+\langle 1\rangle \left(\langle v_x v_y\rangle\langle v_x\rangle\langle v_y\rangle + \langle v_x\rangle^2\langle v_y\rangle^2 - \langle v_x^2\rangle\langle v_y\rangle^2 - \langle v_x\rangle^2\langle v_y^2\rangle\right)\langle|\vec{v}|^2\rangle^2 \tag{S57}$$

which will be nonzero due to the non-independence of the terms being exchanged through the expectation operator. All four sub-matrices follow this same pattern, with no requirements on the form of D^{-1}. Provided that $D > 0$ across the velocity plane, A can be numerically inverted without difficulty.

S4.6. Diameter functions approaching zero

The expression for the thinker gas involves the term $1/D$, implying that particle diameters must always remain finite for reasonable solutions to exist. As a demonstration, we can construct a diameter function with a single minimum which is nonzero by a small value ε:

$$D(\vec{v}, \varepsilon) = 1 - e^{-(|\vec{v}|^2/v_{th})^2} + \varepsilon \tag{S58}$$

where v_{th} is the mean thermal speed of particles. In fig. S4 slices through the velocity distribution function for various values of ε are shown. As $\varepsilon \to 0$, a probability density spike develops. This can be understood physically by considering the rate of collision for particles near the minimum. Such particles have greatly reduced collision rates (and so change velocities very slowly) compared to particles at velocities with finite diameters. For a D with a true zero, particles would never leave the point of phase space where collisions could not occur. This region would behave as a particle sink, from which it would be impossible to scatter back out of. Accordingly, the distribution function would become a delta function and the distribution entropy would diverge. The approach to that divergence is shown by the cuts in fig. S4.

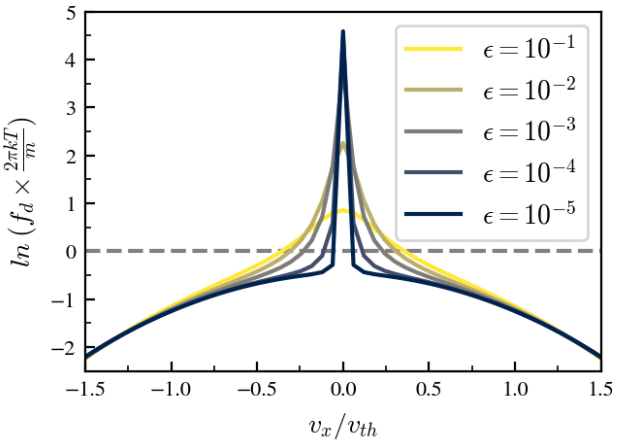

Fig. S4: **Velocity distribution function for thinker gases with a minima approaching zero.** A slice through $v_y = 0$ of the velocity distribution function (f_t) for a diameter function $D = 1 - e^{-|\vec{v}|^2/v_{th}^2} + \varepsilon$. As the minimum of D approaches zero, a density spike in velocity space develops.

S5. EXPLICIT FORM OF $\mathcal{A}$ FOR SELECT $D(\vec{v})$

The matrix $\mathcal{A}$ which defines the thinker gas velocity distribution function involves finding the expectation value (over the MB distribution) of several functions of velocity multiplied with the inverse diameter function (D). For arbitrary D these Gaussian integrals may become too cumbersome to find analytically, and numerical integration is expedient. In this appendix we will present explicit equations for the $\mathcal{A}$ matrices of two special D functions - a diameter step function, and an anisotropic Gaussian diameter function. These matrices may in principle be inverted analytically through a process such as finding the adjugate matrix. The equations for such a process rapidly become unwieldy however, so we will instead use a semi-analytical approach of finding $\mathcal{A}$ exactly, then performing a numerical inverse and estimating small-value derivatives of the relevant constants a, b_x, b_y, and c numerically. This numerical procedure can of course be extended to estimate the values of the $\mathcal{A}$ matrix for diameter functions which do not admit feasible analytic integration.

S5.1. Step function

A convenient diameter function can be written using the Heaviside step function:

$$D(\vec{v}) = D_s + (D_l - D_s)\mathbb{H}\left(-\frac{\vec{v}\cdot\vec{\xi}}{|\vec{v}|}\right) \tag{S59}$$

As we will require D^{-1} in many formulas, it is useful to express this also with the Heaviside function:

$$D^{-1}(\vec{v}) = \frac{1}{D_l} + \left(\frac{1}{D_s} - \frac{1}{D_l}\right)\mathbb{H}\left(-\frac{\vec{v}\cdot\vec{\xi}}{|\vec{v}|}\right) \tag{S60}$$

where D_s and D_l are the small and large diameters respectively, and $\vec{\xi}$ is the normal vector of the large/small decision boundary. We will only consider fixed $\vec{\xi} = \begin{bmatrix}1 & 0\end{bmatrix}$ here. The value of the Heaviside step function at zero will here be taken to be zero (as opposed to 1/2). With this choice of $\vec{\xi}$, integrals over the velocity plane can be split into an integral over the entire plane in addition to an integral over the half plane ($x > 0$). Using well-known expressions for Gaussian integrals of the form $\int x^n \mathcal{N}(x)$ [35], the following components of the $\mathcal{A}$ matrix can be found for a Maxwell-Boltzmann distribution in 2D with bulk velocity $\vec{u}$, thermal energy k_bT and gas density ρ. Where it appears, erf() refers to the standard error function.

$$\mathcal{A}_{11} = \int D^{-1} f_m = \frac{1}{D_l} + \frac{1}{2}\left(\frac{1}{D_s} - \frac{1}{D_l}\right)\left(1 - \mathrm{erf}\left(-u_x\sqrt{\frac{m}{2k_bT}}\right)\right) \tag{S61}$$

$$\mathcal{A}_{12} = \mathcal{A}_{21} = \int v_x D^{-1} f_m = \frac{1}{D_l}u_x + \left(\frac{1}{D_s} - \frac{1}{D_l}\right)\left(\frac{1}{2}u_x\left(\mathrm{erf}\left(-u_x\sqrt{\frac{m}{2k_bT}}\right) - 1\right) - \sqrt{\frac{k_bT}{2\pi m}}e^{-\frac{m}{2k_bT}u_x^2}\right) \tag{S62}$$

$$\mathcal{A}_{13} = \mathcal{A}_{31} = \int v_y D^{-1} f_m = \frac{1}{D_l}u_y + \left(\frac{1}{D_s} - \frac{1}{D_l}\right)\frac{1}{2}u_y\left(1 - \mathrm{erf}\left(-u_x\sqrt{\frac{m}{2k_bT}}\right)\right) \tag{S63}$$

$$\mathcal{A}_{14} = \mathcal{A}_{41} = \int |\vec{v}|^2 D^{-1} f_m = \int v_x^2 D^{-1} f_m + \int v_y^2 D^{-1} f_m = \mathcal{A}_{22} + \mathcal{A}_{33} \tag{S64}$$

$$\begin{aligned}\mathcal{A}_{22} = \int v_x^2 D^{-1} f_m = \quad & \frac{1}{D_l}\left(u_x^2 + \frac{k_bT}{m}\right) + \\ & \frac{1}{2}\left(\frac{1}{D_s} - \frac{1}{D_l}\right)\left(\left(u_x^2 + \frac{k_bT}{m}\right)\left(1 - \mathrm{erf}\left(-u_x\sqrt{\frac{m}{2k_bT}}\right)\right) + u_x\sqrt{\frac{k_bT}{2m\pi}}e^{-\frac{m}{2k_bT}u_x^2}\right)\end{aligned} \tag{S65}$$

$$\mathcal{A}_{33} = \int v_y^2 D^{-1} f_m = \frac{1}{D_l}\left(u_y^2 + \frac{k_bT}{m}\right) + \frac{1}{2}\left(\frac{1}{D_s} - \frac{1}{D_l}\right)\left(u_y^2 + \frac{k_bT}{m}\right)\left(1 - \mathrm{erf}\left(-u_x\sqrt{\frac{m}{2k_bT}}\right)\right) \tag{S66}$$

$$\begin{aligned}\mathcal{A}_{23} = \mathcal{A}_{32} = \int v_x v_y D^{-1} f_m = \quad & \frac{1}{D_l}u_x u_y + \\ & \left(\frac{1}{D_s} - \frac{1}{D_l}\right)\left(\frac{u_x u_y}{2}\left(1 - \mathrm{erf}\left(-u_x\sqrt{\frac{m}{2k_bT}}\right)\right) + u_y\sqrt{\frac{k_bT}{2\pi m}}e^{-\frac{m}{2k_bT}u_x^2}\right)\end{aligned} \tag{S67}$$

$$\mathcal{A}_{24} = \mathcal{A}_{42} = \int v_x|\vec{v}|^2 D^{-1} f_m = \int v_x^3 D^{-1} f_m + \int v_x v_y^2 D^{-1} f_m \tag{S68}$$

$$\begin{aligned}\int v_x^3 D^{-1} f_m = \frac{1}{D_l}u_x\left(u_x^2 + \frac{3k_bT}{m}\right) + \\ \left(\frac{1}{D_s} - \frac{1}{D_l}\right)\left(\frac{u_x}{2}\left(\left(u_x^2 + \frac{3k_bT}{m}\right)\left(1 - \mathrm{erf}\left(-u_x\sqrt{\frac{m}{2k_bT}}\right)\right) + \left(\frac{u_x^2}{2} + \frac{k_bT}{m}\right)\sqrt{\frac{2k_bT}{\pi m}}e^{-u_x^2\frac{m}{2k_bT}}\right)\right)\end{aligned}$$

$$\begin{aligned}\int v_x v_y^2 D^{-1} f_m = \frac{1}{D_l}u_x\left(u_y^2 + \frac{k_bT}{m}\right) + \\ \left(\frac{1}{D_s} - \frac{1}{D_l}\right)\left(u_y^2 + \frac{kT}{m}\right)\left(\frac{1}{2}u_x\left(1 - \mathrm{erf}\left(-u_x\sqrt{\frac{m}{2k_bT}}\right)\right) + \sqrt{\frac{k_bT}{2\pi m}}e^{-\frac{m}{2k_bT}u_x^2}\right)\end{aligned}$$

$$\mathcal{A}_{34} = \mathcal{A}_{43} = \int v_y|\vec{v}|^2 D^{-1} f_m = \int v_y v_x^2 D^{-1} f_m + \int v_y^3 D^{-1} f_m \tag{S69}$$

$$\begin{aligned}\int v_y v_x^2 D^{-1} f_m = \frac{1}{D_l}u_y\left(u_x^2 + \frac{k_bT}{m}\right) + \\ \left(\frac{1}{D_s} - \frac{1}{D_l}\right)\left[\frac{1}{2}u_y\sqrt{\frac{m}{2k_bT}}\left(u_x^2 + \frac{k_bT}{m}\right)\left(1 - \mathrm{erf}\left(-u_x\sqrt{\frac{m}{2k_bT}}\right)\right) + \frac{u_x u_y}{2\sqrt{\pi}}e^{-u_x^2\frac{m}{2k_bT}}\right]\end{aligned}$$

$$\int v_y^3 D^{-1} f_m = u_y\left(u_y^2 + \frac{3k_bT}{m}\right)\left[\frac{1}{D_l} + \frac{1}{2}\left(\frac{1}{D_s} - \frac{1}{D_l}\right)\left(1 - \mathrm{erf}\left(-u_x\sqrt{\frac{m}{2k_bT}}\right)\right)\right]$$

$$\mathcal{A}_{44} = \int |\vec{v}|^4 D^{-1} f_m = \int v_x^4 D^{-1} f_m + 2v_x^2 v_y^2 D^{-1} f_m + v_y^4 D^{-1} f_m \tag{S70}$$

$$\int v_x^4 D^{-1} f_m = \left(u_x^4 + \frac{6k_bT}{m}u_x^2 + \frac{3k_bT^2}{m^2}\right) \times \left[\frac{1}{D_l} + \frac{1}{2}\left(\frac{1}{D_s} - \frac{1}{D_l}\right)\left(1 - \mathrm{erf}\left(-u_x\sqrt{\frac{m}{2k_bT}}\right)\right)\right] + \left(\frac{1}{D_s} - \frac{1}{D_l}\right)\sqrt{\frac{k_bT}{2\pi m}}u_x e^{-u_x\sqrt{\frac{m}{2k_bT}}}\left(u_x^2 + 5\frac{k_bT}{m}\right)$$

$$\int 2v_x^2 v_y^2 D^{-1} f_m = \frac{2}{D_l}\left(u_x^2 + \frac{k_bT}{m}\right)\left(u_y^2 + \frac{k_bT}{m}\right) + \left(\frac{1}{D_s} - \frac{1}{D_l}\right)\left(u_y^2 + \frac{k_bT}{m}\right)\left[\left(u_x^2 + \frac{k_bT}{m}\right)\left(1 - \mathrm{erf}\left(-u_x\sqrt{\frac{m}{2k_bT}}\right)\right) + \sqrt{\frac{2k_bT}{\pi m}}u_x e^{-u_x^2\frac{m}{2k_bT}}\right]$$

$$\int v_y^4 D^{-1} f_m = \left(u_y^4 + \frac{6k_bT}{m}u_y^2 + \frac{3k_bT^2}{m^2}\right)\left[\frac{1}{D_l} + \left(\frac{1}{D_s} - \frac{1}{D_l}\right)\left(1 - \mathrm{erf}\left(-u_x\sqrt{\frac{m}{2k_bT}}\right)\right)\right]$$

S5.2. Anisotropic gaussian diameter function

Another form for D with comparatively simple integrals to compute is an anisotropic Gaussian:

$$D = e^{-\gamma_x\frac{m}{2k_bT}(v_x-u_x)^2} e^{-\gamma_y\frac{m}{2k_bT}(v_y-u_y)^2} \tag{S71}$$

where $\vec{u}$ is again the average velocity of the gas. The relative values of the constants γ_x and γ_x control the degree of anisotropy. This diameter function is particularly convenient because it can be merged with the MB distribution and integrated using the same functional forms:

$$\mathcal{A}_{11} = \int D^{-1} f_m = \frac{1}{\sqrt{1-\gamma_x}\sqrt{1-\gamma_y}} \tag{S72}$$

$$\mathcal{A}_{12} = \mathcal{A}_{21} = \int v_x D^{-1} f_m = \frac{u_x}{\sqrt{1-\gamma_y}\sqrt{1-\gamma_x}} \tag{S73}$$

$$\mathcal{A}_{13} = \mathcal{A}_{31} = \int v_x D^{-1} f_m = \frac{u_y}{\sqrt{1-\gamma_y}\sqrt{1-\gamma_x}} \tag{S74}$$

$$\mathcal{A}_{14} = \mathcal{A}_{41} = \mathcal{A}_{22} + \mathcal{A}_{33} \tag{S75}$$

$$\mathcal{A}_{22} = \int v_x^2 D^{-1} f_m = \frac{1}{\sqrt{1-\gamma_y}\sqrt{1-\gamma_x}}\left(u_x^2 + \frac{k_bT}{m}\frac{1}{(1-\gamma_x)}\right) \tag{S76}$$

$$\mathcal{A}_{33} = \int v_y^2 D^{-1} f_m = \frac{1}{\sqrt{1-\gamma_y}\sqrt{1-\gamma_x}}\left(u_y^2 + \frac{k_bT}{m}\frac{1}{(1-\gamma_y)}\right) \tag{S77}$$

$$\mathcal{A}_{23} = \mathcal{A}_{32} = \int v_x v_y D^{-1} f_m = \frac{u_x u_y}{\sqrt{1-\gamma_x}\sqrt{1-\gamma_y}} \tag{S78}$$

$$\mathcal{A}_{24} = \mathcal{A}_{42} = \int v_x |\vec{v}|^2 D^{-1} f_m = \int v_x^3 D^{-1} f_m + \int v_x v_y^2 D^{-1} f_m \tag{S79}$$

$$\int v_x^3 D^{-1} f_m = \frac{u_x}{\sqrt{1-\gamma_x}\sqrt{1-\gamma_y}}\left(u_x^2 + \frac{3k_bT}{m}\frac{1}{(1-\gamma_x)}\right)$$

$$\int v_x v_y^2 D^{-1} f_m = \frac{u_x}{\sqrt{1-\gamma_x}\sqrt{1-\gamma_y}}\left(u_y^2 + \frac{k_bT}{m}\frac{1}{(1-\gamma_y)}\right)$$

$$\mathcal{A}_{34} = \mathcal{A}_{43} = \int v_y |\vec{v}|^2 D^{-1} f_m = \int v_y v_x^2 D^{-1} f_m + \int v_y^3 D^{-1} f_m \tag{S80}$$

$$\int v_y^3 D^{-1} f_m = \frac{u_y}{\sqrt{1-\gamma_x}\sqrt{1-\gamma_y}} \left(u_y^2 + \frac{3k_bT}{m} \frac{1}{(1-\gamma_y)} \right)$$

$$\int v_y v_x^2 D^{-1} f_m = \frac{u_y}{\sqrt{1-\gamma_x}\sqrt{1-\gamma_y}} \left(u_x^2 + \frac{k_bT}{m} \frac{1}{(1-\gamma_x)} \right)$$

$$\mathcal{A}_{44} = \int v_x^4 D^{-1} f_m + 2 \int v_x^2 v_y^2 D^{-1} f_m + \int v_y^4 D^{-1} f_m \tag{S81}$$

$$\int v_x^4 D^{-1} f_m = \frac{1}{\sqrt{1-\gamma_x}\sqrt{1-\gamma_y}} \left(u_x^4 + u_x^2 \frac{6k_bT}{m} \frac{1}{(1-\gamma_x)} + 3 \left(\frac{k_bT}{m(1-\gamma_x)} \right)^2 \right)$$

$$\int 2v_x^2 v_y^2 D^{-1} f_m = \frac{2}{\sqrt{1-\gamma_x}\sqrt{1-\gamma_y}} \left(u_x^2 + \frac{k_bT}{m} \frac{1}{(1-\gamma_x)} \right) \left(u_y^2 + \frac{k_bT}{m} \frac{1}{(1-\gamma_y)} \right)$$

$$\int v_y^4 D^{-1} f_m = \frac{1}{\sqrt{1-\gamma_x}\sqrt{1-\gamma_y}} \left(u_y^4 + u_y^2 \frac{6k_bT}{m} \frac{1}{(1-\gamma_y)} + 3 \left(\frac{k_bT}{m(1-\gamma_y)} \right)^2 \right)$$

S6. FIRST ORDER HYDRODYNAMICS OF THE THINKER GAS

In this section we comment on the consequences of the thinker gas for first order hydrodynamics by examining the pressure tensor and heat flux. This procedure follows the standard treatment of a passive gas described by the Maxwell-Boltzmann distribution (see for instance, Kardar [33]), but instead utilizes the thinker gas distribution derived in section S4.

Quantities which are conserved microscopically relax at longer timescales than quantities which are not conserved in pair collisions these are denoted 'hydrodynamic' variables and are the main quantities to be described by a hydrodynamic theory. For elastic collisions particle number, components of linear momenta, and kinetic energy are conserved which result in hydrodynamic fields of density, flow velocity, and temperature. By definition, the thinker gas also conserves these quantities (however it has the freedom to redistribute conserved quantities over microscopic degrees of freedom). Regardless of the detailed form, the collision term in equation S30 can be replaced with the left-hand side of the Boltzmann equation, yielding the following expression for the evolution of a test function ϕ:

$$\frac{d}{dt}\langle\phi\rangle = \int_{\mathcal{R}^2} \phi \left[\partial_t + v_i \partial_i + \frac{F_i}{m} \frac{\partial}{\partial v_i} \right] f(\vec{x}, \vec{v}, t) = 0 \tag{S82}$$

where indices i are over Cartesian coordinates. This equation can be manipulated to yield an evolution equation for expectation values of the function ϕ:

$$\partial_t(\rho\langle\phi\rangle) + \partial_i(\rho\langle v_i\phi\rangle) - \rho\langle\partial_t\phi\rangle - \rho\langle v_i\partial_i\phi\rangle - \frac{\rho F_i}{m}\langle\partial\phi/\partial v_i\rangle = 0 \tag{S83}$$

Each of the conserved quantities can be evaluated with this expression to yield familiar hydrodynamic relations. First we define the local velocity, thermal velocity, pressure tensor, rate of strain tensor, and local kinetic energy and heat flux:

$$u_i = \langle v_i \rangle \tag{S84}$$

$$p_i = v_i - u_i \tag{S85}$$

$$P_{ij} = m\rho\langle p_i p_j \rangle \tag{S86}$$

$$u_{ij} = \frac{1}{2}(\partial_i u_i + \partial_j u_j) \tag{S87}$$

$$\epsilon = \langle \frac{1}{2} m |\vec{p}|^2 \rangle = \langle \frac{1}{2} m |\vec{v}|^2 - mv \cdot \vec{u} + \frac{1}{2} m |\vec{u}|^2 \rangle \tag{S88}$$

$$q = \frac{1}{2} m\rho\langle p_i |\vec{p}|^2 \rangle \tag{S89}$$

Hydrodynamic equations pertaining to the conservation of mass, momentum and energy read:

$$\partial_t \rho + \partial_i(\rho u_i) = 0 \tag{S90}$$

$$\partial_t u_i + u_j \partial_j u_i = \frac{F_i}{m} - \frac{1}{m\rho}\partial_j P_{ij} \tag{S91}$$

$$\partial_t \epsilon + u_i \partial_i \epsilon = -\frac{1}{\rho}\partial_i q_i - \frac{1}{\rho} P_{ij} u_{ij} \tag{S92}$$

In order to close this set of equations a form for the pressure tensor and heat flux is needed. These objects are found by taking expectation values of the single-particle probability distribution function, which is in general a function of space, velocity, and time. We will instead follow the standard approach of using a first order approximation of these quantities, found by taking modified expectation values over the homogeneous, steady-state distribution function derived in appendix S4. We begin in the usual way by noting that a mean free path timescale exists, which in the fixed-diameter case is approximately

$$\tau_\times \approx \frac{1}{\rho v_{th} D} \tag{S93}$$

where ρ is the number density of particles, v_{th} is a typical particle speed, and D is a fixed diameter. For the thinker gas, the above expression may be modified by considering that the mean diameter of the gas is the expectation of the diameter function over the velocity distribution:

$$\langle D \rangle_{f_t} = \int D f_t = \int D(1+\mathcal{D}) f_m \tag{S94}$$

For the purposes of this derivation, the mean diameter of the thinker gas will be considered to be fixed, and so $\tau_\times$ is taken to be the same in all directions. If we denote the previously discussed thinker gas velocity distribution for a homogeneous, steady-state gas as f_t^0, we now seek a first-order description of the deviations from this form (denoted f_t^1). We set $f_t^1 = f_t^0(1+h)$, where h describes this small deviation. Furthermore, we linearize the thinker collision operator as:

$$\mathcal{Q}_d^L[f_t^1] \approx -f_t^0 \frac{h}{\tau_\times} \tag{S95}$$

Using the following form of the Boltzmann equation which includes the effects of external forces ($\vec{F}$):

$$\frac{\partial f}{\partial t} + \vec{v} \cdot \nabla_{\vec{x}} f + \frac{\vec{F}}{m} \cdot \nabla_{\vec{v}} f = \mathcal{Q}(f) \tag{S96}$$

$$\mathcal{L}[f] = \mathcal{Q}(f) \tag{S97}$$

where $\mathcal{L}[f] = [\partial/\partial_t + v_i \partial/\partial_i + (F_i/m)\partial/\partial_{v_i}]\, f$ is a linear differential operator. Using the linearized collision operator approximation,

$$\mathcal{L}[f_t^1] \approx -f_t^0 \frac{h}{\tau_\times} \tag{S98}$$

$$h = -\tau_\times \frac{1}{f_t^0} \mathcal{L}[f_t^1] \approx -\tau_\times \mathcal{L}[ln\ f_t^0] \tag{S99}$$

where only the leading term as been retained. We now examine the term $\mathcal{L}[ln\ f_t^0]$:

$$ln\ f_t^0 = ln\ (f_m^0(1+\mathcal{D})) = ln\ (f_m^0 D^{-1} g) \tag{S100}$$

$$= ln\ f_m^0 - ln\ D + ln\ (a + \vec{b}\cdot\vec{v} + c|\vec{v}|^2) \tag{S101}$$

$$\mathcal{L}[ln\ f_t^0] = \mathcal{L}[ln\ f_m^0] - \mathcal{L}[ln\ D] + \mathcal{L}[ln\ (a + \vec{b}\cdot\vec{v} + c|v|^2)] \tag{S102}$$

We can express $h = h_m - h_D + h_g$ to emphasize that three terms contribute to the first order deviations. The first term, h_m, is unchanged from the standard first-order treatment of an equilibrium gas. Since we are only considering time-invariant diameter functions which are functions of velocity only, the second term will be equal to $-h_D = -(F_i/(mD))\partial D/\partial v_i$. Within the bulk of a gas with negligible body forces, this term will not contribute. The final term pertaining to the collision invariant g will have time and space derivatives arising from variations in the fields $T(\vec{x},t)$, $\vec{u}(\vec{x},t)$, and $\rho(\vec{x},t)$, as the values of $a, \vec{b}, c$ are constant for constant $T, \vec{u}, \rho$. We will denote such fields as Φ, and approximate the derivative of the constants $(\partial/\partial\Phi)[a, b_i, c] = [a^\Phi, b_i^\Phi, c^\Phi]$.

If we focus on a single term of the $\mathcal{L}$ operator, $v_i\partial/\partial_i$ (i.e. a steady-state, boundary-less system), and assume only a single field has spatial variation, we can illustrate the contributions from h_g.

$$v_i \frac{\partial ln\ g}{\partial x_i} = \frac{\partial g/\partial r_i}{g} \tag{S103}$$

$$= (a^\Phi + v_j b_j^\Phi + c^\Phi |\vec{v}|^2)\frac{1}{g}\frac{\partial \Phi}{\partial x_i} \tag{S104}$$

The expectation of first order-quantities like the pressure tensor and heat flux can be found by integration over the perturbed zeroth-order distribution:

$$\langle \mathcal{O} \rangle^1 = \int \mathcal{O} f_t^0 (1+h) = \langle \mathcal{O} \rangle^0 + \langle h\mathcal{O} \rangle^0 \tag{S105}$$

$$= \langle \mathcal{O} \rangle^0 + \langle h_m\mathcal{O} \rangle^0 - \langle h_D\mathcal{O} \rangle^0 + \langle h_g\mathcal{O} \rangle^0 \tag{S106}$$

For the pressure tensor ($\mathcal{O} = P_{ij}$) and the heat flux ($\mathcal{O} = q_i$), the contributions from the term $\langle h_m\mathcal{O}\rangle$ are unchanged from the standard treatment of a passive gas, resulting in heat flux terms which relax thermal gradients and off-diagonal pressure terms which relax shear flows. As noted above, the contributions from h_D will often be unimportant, so we focus on the term $\langle h_g\mathcal{O}\rangle$. The following forms of the pressure tensor and heat flux can be found:

$$P_{ij} = m\rho \sum_k \int f_m D^{-1} (a^\Phi + v_l b_l^\Phi + c^\Phi |\vec{v}|^2) \frac{\partial \Phi}{\partial x_k} v_k p_i p_j \tag{S107}$$

$$q_i = \frac{1}{2} m\rho \sum_k \int f_m D^{-1} (a^\Phi + v_l b_l^\Phi + c^\Phi |\vec{v}|^2) \frac{\partial \Phi}{\partial x_k} v_k p_i |\vec{p}|^2 \tag{S108}$$

Several remarks can be made about the first-order thinker corrections to P_{ij} and q_i, however first some comment about velocity reference frames must be made. The velocity, thermal velocity, and local velocity are related by $\vec{p} = \vec{v} - \vec{u}$. Kinetic theory is largely conducted in the $\vec{u}$ co-moving frame, which allows for the definition of familiar pressures, fluxes and temperatures. As the Maxwell-Boltzmann distribution is also defined in this frame, all is well. However for the thinker gas, we may define the particle's function of velocity in a global coordinate frame, or a local coordinate frame ($D(\vec{v})$ or $D(\vec{p})$). In this work we have chosen to use $D(\vec{v})$, except in the case of the anisotropic Gaussian function defined in appendix S5. This is because in any practical realization of a device like the thinkers described here, it is likely that knowledge of velocity in the global coordinate frame will be desirable or even required. In order for the operation of thinkers to induce fluxes towards a goal region, thinkers must be aware of the large-scale coordinate system which locates themselves and the target region. Furthermore, in such a case they are unconcerned with local fluxes relative to nearby flow, but rather absolute fluxes towards or away from the target.

Another consideration is the complexity of estimating the local average velocity of neighboring thinker particles. Simplistic methods of determining position (and hence velocity) such as utilizing a light gradient [36] inform particles of their coordinates in a global frame. Using similar methods to determine relative velocities to neighbors also raises a question of length scale - over what distance, practically, should particles consider their 'local' velocity neighborhood to extend?

Finally, the removal of a global coordinate frame for thinker particles would necessitate each particle to have an orientation vector. As this study concerns isotropic particles, particle orientations will not evolve during simulations unless an orientation-coupling force is defined, in a manner analogous to the Toner-Tu model of flocking [15]. For isotropic thinkers, such a choice would be arbitrary.

The mixed velocity components in eqs. S107 and S108 reflect the different considerations inherent to the operation of thinkers and the definition of the hydrodynamic equations of motion. First let us consider the case of $|\vec{u}| = 0$, in which $\vec{p} = \vec{v}$. In that case, the pressure tensor contains only one even moment of velocity and the term $f_m D^{-1}$. If D and f_m have inversion symmetry about the same point, then only the even term $(\sum_k \int f_m D^{-1} v_l v_k v_i v_j b_l^\Phi \partial\Phi/\partial x_i)$ will contribute to the pressure tensor corrections. However in the case of inversion-symmetric D, the constants b_i are

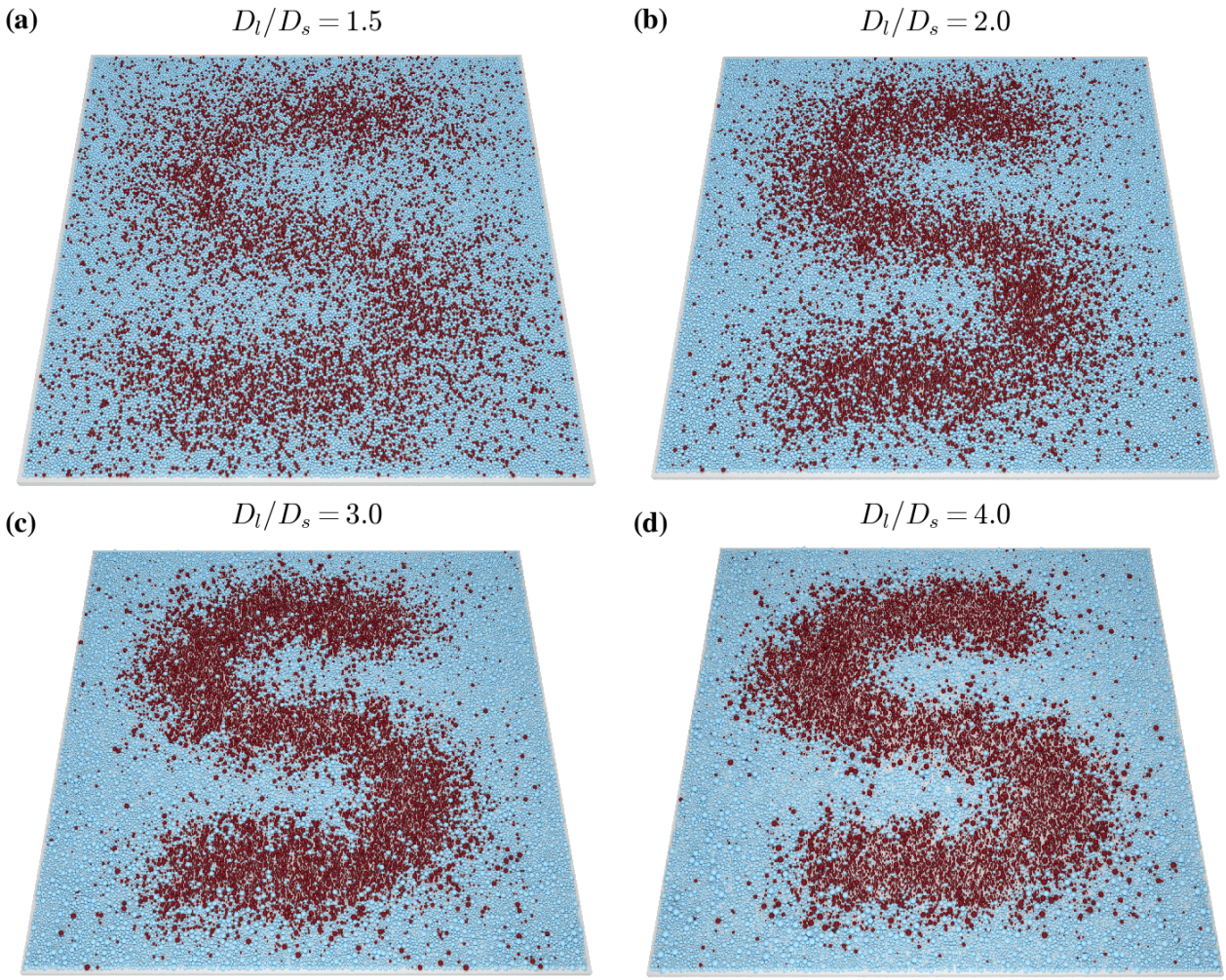

Fig. S5: **Expansion ratio controls pattern fidelity.** Patterns formed by dropping thinker grains with differing large and small diameter ratios. Larger ratios substantially improve pattern fidelity.

zero and no heat fluxes arise in the quiescent fluid. Conversely, in the heat flux tensor all terms except for derivatives of b_i are even moments, and therefore contribute to heat fluxes in the presence of temperature gradients. Fluids with bulk flows ($|\vec{u}| \neq 0$) will have contributions from first order-corrections to both heat fluxes and the pressure tensor.

S7. PATTERN RESOLUTION IN DROPPED THINKER GRAINS

The ability of thinkers to form and maintain patterns was principally a function of the ratio of their large and small diameter states. Figure S5 shows this trend for diameter ratios ranging from $D_l/D_s = 1.5 \rightarrow 4$.

S8. DESCRIPTION OF SUPPLEMENTAL VIDEOS

S8.1. Supplemental video S1

In this video, two species of thinker particles are dropped into a hard-sided box. They fall under a vertical force, while a drag force (linear in velocity) causes them to slowly lose energy and settle into a layer. Feedback control of thinker diameter allows the two species to separate and form a pre-programmed pattern.

S8.2. Supplemental video S2

In this video, two species of thinker particles are continuously agitated while contained in a hard-sided box. Feedback control of thinker diameters allow the two species to form patterns, which are changed over time. The thinkers can therefore write sequences of symbols, as shown here with an equation.

---

[1] L. Zhu, E. Lauga, and L. Brandt, Physics of fluids **24**, 051902 (2012).
[2] P. Romanczuk, M. Bär, W. Ebeling, B. Lindner, and L. Schimansky-Geier, The European Physical Journal Special Topics **202**, 1 (2012).
[3] A. Doostmohammadi, J. Ignés-Mullol, J. M. Yeomans, and F. Sagués, Nature communications **9**, 1 (2018).
[4] T. Sanchez, D. T. N. Chen, S. J. DeCamp, M. Heymann, and Z. Dogic, Nature **491**, 431 (2012), ISSN 1476-4687, URL `http://www.nature.com/articles/nature11591`.
[5] I. Buttinoni, J. Bialké, F. Kümmel, H. Löwen, C. Bechinger, and T. Speck, Physical Review Letters **110**, 238301 (2013), publisher: American Physical Society, URL `https://link.aps.org/doi/10.1103/PhysRevLett.110.238301`.
[6] J. Palacci, S. Sacanna, A. P. Steinberg, D. J. Pine, and P. M. Chaikin, Science **339**, 936 (2013), publisher: American Association for the Advancement of Science, URL `http://www.science.org/doi/full/10.1126/science.1230020`.
[7] C. Bechinger, R. Di Leonardo, H. Löwen, C. Reichhardt, G. Volpe, and G. Volpe, Reviews of Modern Physics **88**, 045006 (2016).
[8] J. Tailleur and M. Cates, Physical review letters **100**, 218103 (2008).
[9] A. Bricard, J.-B. Caussin, N. Desreumaux, O. Dauchot, and D. Bartolo, Nature **503**, 95 (2013).
[10] A. Bricard, J.-B. Caussin, D. Das, C. Savoie, V. Chikkadi, K. Shitara, O. Chepizhko, F. Peruani, D. Saintillan, and D. Bartolo, Nature communications **6**, 1 (2015).
[11] J. Deseigne, O. Dauchot, and H. Chaté, Physical review letters **105**, 098001 (2010).
[12] M. C. Marchetti, J.-F. Joanny, S. Ramaswamy, T. B. Liverpool, J. Prost, M. Rao, and R. A. Simha, Reviews of modern physics **85**, 1143 (2013).
[13] T. H. Tan, A. Mietke, J. Li, Y. Chen, H. Higinbotham, P. J. Foster, S. Gokhale, J. Dunkel, and N. Fakhri, Nature **607**, 287 (2022).
[14] H. Leff and A. F. Rex, *Maxwell's Demon 2 Entropy, Classical and Quantum Information, Computing* (CRC Press, 2002).
[15] J. Toner and Y. Tu, Phys. Rev. E **58**, 4828 (1998).
[16] J. Gelles and R. Landick, Cell **93**, 13 (1998).
[17] J. F. Sydow and P. Cramer, Current opinion in structural biology **19**, 732 (2009).
[18] A. L. Gnatt, P. Cramer, J. Fu, D. A. Bushnell, and R. D. Kornberg, Science **292**, 1876 (2001).
[19] I. Artsimovitch and R. Landick, Proceedings of the National Academy of Sciences **97**, 7090 (2000).
[20] V. R. Tadigotla, D. Ó. Maoiléidigh, A. M. Sengupta, V. Epshtein, R. H. Ebright, E. Nudler, and A. E. Ruckenstein, Proceedings of the National Academy of Sciences **103**, 4439 (2006).
[21] H. Kettenberger, K.-J. Armache, and P. Cramer, Cell **114**, 347 (2003).
[22] D. A. Erie, O. Hajiseyedjavadi, M. C. Young, and P. H. von Hippel, Science **262**, 867 (1993).
[23] M. T. Marr and J. W. Roberts, Molecular cell **6**, 1275 (2000).
[24] E. A. Abbondanzieri, W. J. Greenleaf, J. W. Shaevitz, R. Landick, and S. M. Block, Nature **438**, 460 (2005).
[25] J. M. Parrondo, J. M. Horowitz, and T. Sagawa, Nature physics **11**, 131 (2015).
[26] F. J. Cao and M. Feito, Physical Review E **79**, 041118 (2009).
[27] T. M. Cover and J. A. Thomas, *Elements of information theory* (Wiley-Interscience, 2006).
[28] P. J. Flory and J. Rehner, The Journal of Chemical Physics **11**, 512 (1943), URL `https://doi.org/10.1063/1.1723791`.
[29] L. G. Treloar, *The physics of rubber elasticity* (OUP Oxford, 1975).
[30] K. Wojciechowski and A. Brańka, Physics Letters A **134**, 314 (1989), ISSN 0375-9601, URL `https://www.sciencedirect.com/science/article/pii/0375960189906427`.
[31] B. VanSaders, J. Dshemuchadse, and S. C. Glotzer, Physical Review Materials **2**, 063604 (2018).
[32] C. Cercignani, R. Illner, and M. Pulvirenti, *The Mathematical Theory of Dilute Gases*, Applied Mathematical Sciences (Springer New York, 1994), ISBN 9780387942940.
[33] M. Kardar, *Statistical physics of particles* (Cambridge University Press, 2007).
[34] C. Cercignani, Journal of statistical physics **58**, 817 (1990).
[35] N. E. Korotkov and A. N. Korotkov, *Integrals Related to the Error Function* (CRC Press, 2020).
[36] S. Li, R. Batra, D. Brown, H.-D. Chang, N. Ranganathan, C. Hoberman, D. Rus, and H. Lipson, Nature **567**, 361 (2019).