Stochastic disks that roll

Miranda Holmes-Cerfon[*]

Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, New York 10012, USA

We study a model of rolling particles subject to stochastic fluctuations, which may be relevant in systems of nano- or microscale particles where rolling is an approximation for strong static friction. We consider the simplest possible nontrivial system: a linear polymer of three disks constrained to remain in contact and immersed in an equilibrium heat bath so the internal angle of the polymer changes due to stochastic fluctuations. We compare two cases: one where the disks can slide relative to each other and the other where they are constrained to roll, like gears. Starting from the Langevin equations with arbitrary linear velocity constraints, we use formal homogenization theory to derive the overdamped equations that describe the process in configuration space only. The resulting dynamics have the formal structure of a Brownian motion on a Riemannian or sub-Riemannian manifold, depending on if the velocity constraints are holonomic or nonholonomic. We use this to compute the trimer's equilibrium distribution with and without the rolling constraints. Surprisingly, the two distributions are different. We suggest two possible interpretations of this result: either (i) dry friction (or other dissipative, nonequilibrium forces) changes basic thermodynamic quantities like the free energy of a system, a statement that could be tested experimentally, or (ii) as a lesson in modeling rolling or friction more generally as a velocity constraint when stochastic fluctuations are present. In the latter case, we speculate there could be a "roughness" entropy whose inclusion as an effective force could compensate the constraint and preserve classical Boltzmann statistics. Regardless of the interpretation, our calculation shows the word "rolling" must be used with care when stochastic fluctuations are present.

I. INTRODUCTION

Particles that live on the nano- or microscale commonly have short-ranged interactions, so their surfaces come close enough that surface frictional effects may be important. For example, recent experiments and simulations have shown that tangential frictional forces between rough and otherwise stochastic particles are probably the origin of the shear-thickening behavior of many materials [1,2]. Other studies demonstrate that sticky tethers attached to particle surfaces can change their dynamics [3,4]. Since one promising method of creating colloids with programmable interactions is to coat them with strands of DNA [5–8], which could impede their relative sliding, this could have major implications for their assembly pathways and hence structures that can be formed by self-assembly. On these scales it is extremely difficult to measure the particles' rotational degrees of freedom, so one must resort to indirect methods to determine whether tangential frictional forces are present [9,10]. Therefore, it would be highly desirable to find a simpler way to quantify these forces, via macroscopic measurements of spatial positions only.

While the macroscopic effect of dry friction has been studied in detail in granular systems [11–16], whose components are large and typically athermal, it has rarely been considered for small particles subject to thermal fluctuations, except in simple one-dimensional models [17–21]. A starting point would be to ignore the details of the friction, which are not well understood [22,23], and consider the limit of infinite friction: stochastic particles that roll relative to each other when they are in contact. Rolling has been studied in nonstochastic systems and is known to produce a wealth of counterintuitive phenomena: a spinning top spontaneously reverses its direction, a golf ball pops out of a hole without

hitting the bottom, a dropped quarter spins infinitely quickly in finite time [24–26]. Collectively, rolling particles have different phase behaviors than those that slide [27]. Yet despite their intriguing dynamics, rolling has been considered in stochastic settings only for simple systems such as a rolling ball or sled [28–30], or as a noisy relaxation of the rolling constraint itself [31].

This paper studies a natural model of stochastic, rolling particles, with the aim of determining how rolling could affect quantities that are macroscopically measurable. It considers a system whose dynamics can be worked out explicitly: a polymer of three two-dimensional disks that are constrained to roll relative to each other, like gears. Unlike traditional gears, however, the disks can change their relative positions in space. We start with the Langevin equations for the stochastic dynamics combined with velocity constraints to model perfect rolling, and from this calculate the equilibrium distribution of the internal angle of the trimer. Surprisingly, the distributions are different depending on if the velocity constraints are included or not. If this is an accurate model of stochastic particles interacting with infinitely strong friction, it suggests that even finite friction could change the free energies of a system of particles. Such a result can only hold if the friction force causes the system to deviate from the predictions of classical statistical mechanics, but would be possible to test experimentally via macroscale measurements.

An outline is as follows. Section II describes the setup and notation, including the full Langevin equations and specific forms of the constraints for arbitrary collections of disks. Section III describes the overdamped Langevin dynamics. Section IV derives the equilibrium distributions for a trimer of disks both with and without rolling constraints. Section V discusses the results in a physical context. Section VI concludes and speculates how this might apply to spheres, whose configuration space is geometrically fundamentally different.

---

[*]holmes@cims.nyu.edu

II. SETUP

We represent the disks as a vector $x = (x^{(1)}, x^{(2)}, x^{(3)}, \theta_1, \theta_2, \theta_3) \in \mathbb{R}^9$, where each disk has three coordinates representing the center of mass $x^{(i)} = (x_i, y_i)$ and the overall internal rotation θ_i relative to a fixed, external coordinate system. We will call $\{x^{(i)}\}_{i=1}^3$ the "position" variables because they describe the disks' overall positions in space, and we will call $\{\theta_i\}_{i=1}^3$ the "spin" variables because they describe how much each disk has internally rotated, or spun about an axis, like a gear fixed in place. The spin variables are the ones that are usually not accessible by macroscopic measurements. All vectors in this paper are column vectors, though we write them inline for readability. The disks are identical with unit diameters, and pairs (1,2) and (2,3) are in contact. For each such pair (i,j) there are two possible constraints: one requires the disks to be a fixed distance apart so they are exactly touching, and another requires the points in contact to move with the same relative velocity. These each imply a constraint on the velocities (not momenta), as

$$(x^{(i)} - x^{(j)})(\dot{x}^{(i)} - \dot{x}^{(j)}) = 0, \tag{1}$$

$$(x^{(i)} - x^{(j)})^\perp (\dot{x}^{(i)} - \dot{x}^{(j)}) = \tfrac{1}{2}(\dot{\theta}_i + \dot{\theta}_j). \tag{2}$$

We write $(u,v)^\perp = (-v,u)$. The second constraint comes from noting the velocity on disk i of the point in contact with disk j is $\dot{x}^{(i)} + \dot{\theta}_i \frac{1}{2}(x^{(j)} - x^{(i)})^\perp$, and considering the component of relative velocity that is perpendicular to $x^{(i)} - x^{(j)}$, since the component parallel to it is accounted for by the first constraint. We call Eq. (1) the "bond constraints" and Eq. (2) the "rolling constraints." In addition, we constrain the center of mass to the origin. The complete set of constraints can be written as

$$C(x)\dot{x} = 0, \tag{3}$$

where $C(x) \in \mathbb{R}^{m \times n}$ is a matrix whose rows are the coefficients multiplying velocities in Eqs. (1) and (2). Here $m = 6$ is the number of constraints, and $n = 9$ is the number of configuration space variables.

We suppose the potential energy of the system is a smooth function $U(x)$, and the disks are immersed in a fluid or other medium that provides a white noise forcing to the momentum and a viscous damping that is linear in velocity. We use the Langevin equations to model the dynamics and assume the friction tensor $\Gamma(x)$ and forcing tensor $\sigma(x)$ satisfy a fluctuation-dissipation relation $\sigma\sigma^T = 2\beta^{-1}\Gamma$, where $\beta = (k_B T)^{-1}$ is the inverse of temperature T times the Boltzmann constant. This ensures that the invariant measure for the unconstrained system is the Boltzmann distribution: $e^{-\beta U(x)} e^{-\frac{\beta}{2}\dot{x}^T M \dot{x}}$. Here M is the mass matrix, which is diagonal with entries equal to either the mass or moment of inertia of a disk.

It should be noted that the Langevin equations do not correctly describe the velocity correlations of particles immersed in a fluid, even for large particles, because the correlation times for the particles' velocities and for the fluid's momentum fluctuations are of the same magnitude [32,33]. However, they do capture the correct static thermodynamic behavior (at least without constraints), and they do lead to the

correct overdamped equations. Since our goal is to obtain an overdamped equation which involves restrictions on velocities, and not to correctly describe the velocity correlations induced by hydrodynamic interactions, we proceed with the Langevin equations as a starting point from which we can impose velocity constraints via mechanical principles.

To account for constraints, we apply d'Alembert's principle, or the principle of virtual work. This requires that the constraints do no work in a "virtual" move, namely one which holds all variables fixed and takes a step in a tangent direction consistent with the constraints. The constraints must therefore be imposed by forces perpendicular to the allowable tangent directions, which can be done using Lagrange multipliers [34]. This is the only principle available for arbitrary linear velocity constraints, since variational principles are only valid when the constraints are known to be holonomic [35]. The constrained Langevin equations are

$$M\ddot{x} + \Gamma\dot{x} = -\nabla U(x) + \sigma\eta(t) - C^T\lambda, \tag{4}$$

combined with the constraints Eq. (3). In the above, $\eta(t)$ is a n-dimensional white noise, and $\lambda \in \mathbb{R}^n$ are the Lagrange multipliers that ensure the constraints are satisfied. The product $\sigma(x)\eta$ can be interpreted in either the Itô or the Stratonovich sense, since $\sigma(x)$ does not depend on $\dot{x}$ so it is of bounded variation. The mass can be removed from the equations by changing to mass-scaled variables (see Appendix A, or Ref. [36]), so hereafter we set $M = I$. This is not a nondimensionalization, but simply a convenient change of variables, which can be inverted to put the mass back in at any step in the subsequent analysis when desired.

III. OVERDAMPED DYNAMICS

The particles we aim to model have very short correlation times for momentum, so they are effectively modeled by the overdamped Langevin equations, which describe the dynamics of the system in configuration space only. We derive these overdamped equations by considering the limit of large viscous friction and long timescales. In this section we sketch the results; the detailed calculations are shown in Appendices B and C.

First, we write Eq. (4) explicitly. The Lagrange multipliers can be computed analytically by taking the time derivative of Eq. (3) and substituting for $\ddot{x}$ from Eq. (4) (see Appendix B). The resulting equations are

$$\ddot{x} + P\Gamma\dot{x} = -P\nabla U(x) + P\sigma\eta(t) + C^T G^{-1}\nabla C(\dot{x},\dot{x}). \tag{5}$$

The matrix $P(x)$ is an orthogonal projection onto the complement to the row space of $C(x)$, and $G(x)$ is the Gram matrix of the constraints:

$$P = I - C^T G^{-1} C, \quad G = CC^T. \tag{6}$$

Here and hereafter I is an identity matrix with dimensions correct for the context. The final term in Eq. (5) is a vector with components $(\nabla C(x)(\dot{x},\dot{x}))_i = \sum_{i,j} \frac{\partial c_i}{\partial x_j} \dot{x}_i \dot{x}_j$, and represents the extra acceleration due to the curvature of the constraints. This shows the constrained dynamics are given by projecting the original momentum equation onto the subspace of allowed, unconstrained directions, plus a curvature-driven acceleration term [37]. Again, it doesn't matter if the projected force

$P(x)\sigma(x)\eta(t)$ is interpreted in the Itô or Stratonovich sense because x is of bounded variation.

Next, we consider the overdamped limit by letting $\Gamma \rightarrow \Gamma/\epsilon$ and $t \rightarrow t/\epsilon$, and performing formal homogenization on the generator of Eq. (5) [38]. This is a standard technique to obtain the overdamped Langevin equations asymptotically; the novelty here is the arbitrary linear velocity constraints. The result is weakly equivalent to the stochastic process $x(t)$ that solves the Itô equation (see Appendix C for details):

$$\dot{x} = -\Gamma_P^{\dagger} \nabla U - \beta^{-1} \sum_{ij} P_{ij} \partial_j (\Gamma_P^{\dagger})_{ki} + \bar{\sigma}_P \eta. \qquad (7)$$

Here $\Gamma_P = P\Gamma P$, and $\Gamma_P^{\dagger}$ is its Moore-Penrose pseudoinverse [39] (see, e.g., Ref. [40]). The matrix $\bar{\sigma}_P$ is any matrix such that $\bar{\sigma}_P \bar{\sigma}_P^T = 2\beta^{-1}\Gamma_P^{\dagger}$.

To highlight the fundamental ideas we will analyze Eq. (7) in the simplest possible setup: constant friction and no long-range potential energy, so $\Gamma = I$ and $U(x) = 0$. In this case $\Gamma_P^{\dagger} = P$, so after a change of time scale $t \rightarrow t/(2\beta)$ Eq. (7) becomes

$$\dot{x} = P(x) \circ \eta, \qquad (8)$$

where $\circ$ denotes the Stratonovich product. The process $x(t)$ looks locally like a Brownian motion that can only move in a subspace of its ambient space.

This process is well-understood mathematically when the velocity constraints are holonomic, meaning they imply an equal number of constraints in configuration space, i.e., on the variables contained in x. In this case the process is constrained to remain on a manifold $\mathcal{M}$ in configuration space, whose dimension equals the rank of $P(x)$ (which is an orthogonal projection matrix onto the tangent space to $\mathcal{M}$ at each point x). The process is actually a Brownian motion on $\mathcal{M}$, which is by definition a process whose generator is the Laplace-Beltrami operator on $\mathcal{M}$ [41,42]. Briefly, to see why, note that the generator of Eq. (8) is

$$\mathcal{L}u = \tfrac{1}{2} \operatorname{Tr}[P\nabla(P\nabla u)], \qquad (9)$$

where $\operatorname{Tr}(AB) = A{:}B = \sum_{i,j} A_{ij}B_{ij}$ for matrices A,B, and ∇ applied to a vector expands each element into a row so that $(\nabla v)_{ij} = \partial_j v_i$. Here u may be thought of as a function on $\mathcal{M}$, even though the gradient operator in $\mathcal{L}$ acts on all directions in the ambient space. Then, $P\nabla u$ can be shown to equal $\operatorname{grad} u$ where grad is the gradient operator on $\mathcal{M}$, and $\operatorname{Tr}(P\nabla v)$ can be shown to equal $\operatorname{div} v$, where div is the divergence operator on $\mathcal{M}$ (see, e.g., Refs. [36,37,43]). Therefore, for holonomic constraints $\mathcal{L}u = \tfrac{1}{2}\operatorname{div}\operatorname{grad} u$, which is the Laplace-Beltrami operator on $\mathcal{M}$.

For nonholonomic constraints there is no similarly canonical interpretation of $\mathcal{L}$ of the form Eq. (9) at the present time, as we discuss briefly in the conclusion.

For disks, the bond constraints Eq. (1) are holonomic since they imply the distances between pairs in contact are conserved, which is a constraint in configuration space. The constraints on the center of mass are also holonomic. The rolling constraints Eq. (2) are not immediately seen to be holonomic, since they cannot be integrated in time directly. Although one can show they are indeed holonomic as we discuss briefly in Sec. IV B 2 and Appendix E, we will proceed without this

knowledge, to show that one can still work with Eq. (8) without knowing the geometric structure of the constraints.

IV. EQUILIBRIUM DISTRIBUTION

A. Result

Next we ask what is the equilibrium distribution for a trimer in position space with and without the rolling constraints. We will show these have densities proportional to, respectively,

$$\pi_{\text{slide}}(\omega) \propto \sqrt{1 + 2\sin^2\omega}\sqrt{1 + 2\cos^2\omega},$$
$$\pi_{\text{roll}}(\omega) \propto \sqrt{5 + 2\sin^2\omega}\sqrt{13 + 2\cos^2\omega}, \qquad (10)$$

where 2ω is the internal angle of the trimer. The domain is $\omega \in [\pi/6, 5\pi/6]$ if the spheres cannot interpenetrate, and $\omega \in [0,\pi]$ if they can (as is allowed in simulations). These calculations are performed for the simplest setup described by Eq. (8), but we expect them to be valid in more general settings [with a suitable modification to account for the potential energy $U(x)$].

Figure 2 plots the two distributions. The rolling constraints favor more open configurations than purely bond constraints. This figure also plots the empirical histograms obtained by numerically simulating the Langevin Eqs. (4) directly (see Appendix D for methods); the agreement verifies our calculations. The small discrepancies are thought to arise partly from statistical fluctuations, and, in the case of rolling constraints, because the numerical method does not conserve the additional implied constraints in configuration space (see Appendix E).

The distributions above are for "hard" constraints, i.e., the constraints are satisfied exactly. In a physical system constraints are often an approximation for a concentration of probability near a lower-dimensional manifold, but the system can wiggle around near this: the constraints are "soft." This happens, for example, when constraints of the form $q_i(x) = z$ (where i indexes the constraints) are imposed by a stiff potential energy, such as $U(x) = \epsilon^{-1}|q_i(x) - z|^2$ with $\epsilon \ll 1$. This wiggle room changes the equilibrium density, and in the limit of infinite stiffness it is not the same as imposing hard constraints; this is the well-known "paradox" of hard versus soft constraints in statistical mechanics that has been discussed many times in the literature [e.g., 44,45]. The distributions for infinitely stiff soft constraints can be obtained from those for hard ones and we will show they are

$$\pi_{\text{slide,vibr}}(\omega) \propto 1,$$
$$\pi_{\text{roll,vibr}}(\omega) \propto \sqrt{\frac{5 + 2\sin^2\omega}{1 + 2\sin^2\omega}}\sqrt{\frac{13 + 2\cos^2\omega}{1 + 2\cos^2\omega}}. \qquad (11)$$

These are the distributions one would typically compare to experimentally; for example, similarly obtained distributions accurately predict the equilibrium probabilities of colloidal clusters [46]. The distribution when disks can slide is constant (see Fig. 2), as one would expect since each outer disk should be uniformly distributed on the surface of the central disk.
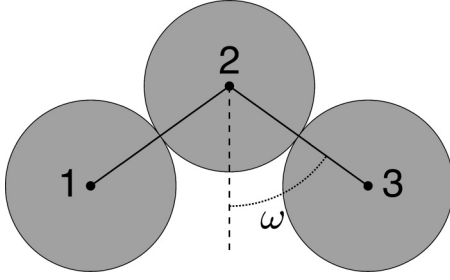
FIG. 1. A trimer of disks, constrained to preserve the distances between disk pairs 1-2, 2-3, and sometimes constrained to roll when pairs are in contact. This setup illustrates the parametrization in Eq. (13).

B. Derivation

In this section we show Eqs. (10) and (11) explicitly. This section is technical and not essential to understanding the subsequent discussion.

Our strategy will be to parametrize the position degrees of freedom of the cluster explicitly to remove the bond and center of mass constraints, write the equations in these variables, and finally solve the Fokker-Planck equation by direct calculation. This is a brute-force approach yet it still gives insight into the geometry and mechanics of the constraints, by explicitly identifying the linear subspaces involved in setting the dynamics.

1. Hard constraints

Let $\omega \in [0,\pi]$ (or $[\pi/6,5\pi/6]$) be half the internal angle, measured underneath the line 1-2-3 when disk 2 has been rotated to lie on the y axis, and let $\phi \in [0,2\pi]$ be the overall rotation of the cluster. See Fig. 1 for an illustration. Let the position variables be

$$\bar{x}(\phi,\omega) = R(\phi)\bar{x}_0, \tag{12}$$

where

$$\bar{x}_0 = \left(-\sin\omega, -\tfrac{1}{3}\cos\omega, 0, \tfrac{2}{3}\cos\omega, \sin\omega, -\tfrac{1}{3}\cos\omega \right), \tag{13}$$

and $R(\phi)$ is a 6×6 block diagonal matrix, whose blocks are 2×2 matrices that rotate each point $x^{(i)}$ by an angle ϕ about the origin [47]. The full cluster is parameterized by

$$x = (\bar{x},\theta) = (R(\phi)\bar{x}_0,\theta_1,\theta_2,\theta_3). \tag{14}$$

This preserves the bond and center of mass constraints so they can be removed from the rows of $C(x)$, which form the projection $P(x) \in \mathbb{R}^{9 \times 9}$.

We now perform this change of variables in Eq. (8), to write the dynamics in terms of the new variables $y = (\omega,\phi,\theta_1,\theta_2,\theta_3)$. Let $s = (\omega,\phi)$ be the position variables and let $\theta = (\theta_1,\theta_2,\theta_3)$ be the spin variables. Let us define the following matrices:

$$
\begin{aligned}
S &= \nabla_s x = \left(\tfrac{\partial x}{\partial \omega}, \tfrac{\partial x}{\partial \phi} \right) & &\in \mathbb{R}^{9 \times 2}, \\
T &= \nabla_\theta x = \left(\tfrac{\partial x}{\partial \theta_1}, \tfrac{\partial x}{\partial \theta_2}, \tfrac{\partial x}{\partial \theta_3} \right) & &\in \mathbb{R}^{9 \times 3}, \\
Y &= \nabla_y x = (S \quad T) & &\in \mathbb{R}^{9 \times 5}, \\
Q &= S^T S = \begin{pmatrix} K^2 & 0 \\ 0 & L^2 \end{pmatrix} & &\in \mathbb{R}^{2 \times 2}.
\end{aligned} \tag{15}
$$

Here 0 is the matrix of zeros with appropriate dimensions, and the diagonal elements of Q are

$$K^2(\omega) = \left| \frac{\partial \bar{x}}{\partial \omega} \right|^2 = \frac{2}{3} + \frac{4}{3}\cos^2\omega, \tag{16}$$

$$L^2(\omega) = \left| \frac{\partial \bar{x}}{\partial \phi} \right|^2 = \frac{2}{3} + \frac{4}{3}\sin^2\omega. \tag{17}$$
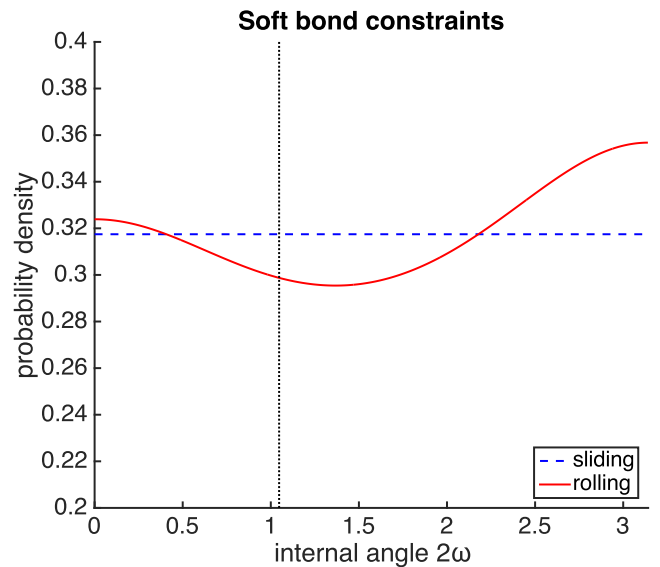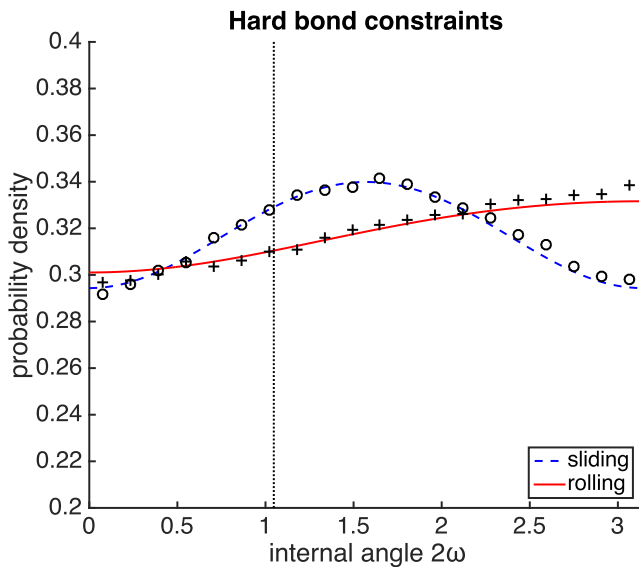

FIG. 2. Equilibrium probability densities for a trimer where disks may slide (blue dashed lines) and where they are constrained to roll (red solid lines), as a function of internal angle 2ω, in radians. Left plot is for bond constraints imposed exactly, right plot is for bond constraints imposed with a stiff spring. Markers are the empirical densities obtained by numerically simulating the Langevin Eqs. (4). The vertical dotted line shows where disks 1 and 3 overlap, so a physical density should be truncated here.

Note that $L^2(\omega)$ is the two-dimensional moment of inertia of the cluster.

We can use the regular chain rule of calculus on Eq. (8), since this is in Stratonovich form. This gives

$$\frac{dx}{dt} = \sum_{i=1}^{5} \frac{\partial x}{\partial y_i} \frac{dy_i}{dt} = Y\frac{dy}{dt} = P(x) \circ \eta. \qquad (18)$$

Multiplying by $(Y^T Y)^{-1} Y^T$ gives an equation for $\dot{y}$. Note that $Y^T Y = \begin{pmatrix} Q & 0 \\ 0 & I \end{pmatrix}$, since $S^T T = 0$, $T^T S = 0$, and $T^T T = I$. Separating the equations for the position and spin variables separately gives

$$\dot{s} = Q^{-1}(PS)^T \circ \eta, \qquad (19)$$

$$\dot{\theta} = (PT)^T \circ \eta. \qquad (20)$$

Here $\eta \in \mathbb{R}^9$ is the same white noise for each.

Notice that Eq. (19) for the position variables does not depend on θ_i, because P, S, Q are independent of θ_i. (The spin variables, however, do depend on the positions.) Therefore, we can analyze it independently, to compute the equilibrium distribution in these variables only.

First consider the equilibrium density for Eq. (19) *without* the rolling constraints, so that $P = I$. One strategy would be to compute the matrix elements in Eq. (19) directly and solve the stationary Fokker-Planck equation, as we will do when rolling constraints are included. However, it is simpler to proceed geometrically, and recognize that, based on Eq. (8) and the subsequent discussion, Eq. (19) is a parameterized version of a Brownian motion on a manifold. (Note that S has zeros in the entries corresponding to the spin variables so these components of η do not contribute.) This manifold (call it $\mathcal{N}$) is the set of accessible configurations in position space when internal rotations are ignored. It can be embedded in the full configuration space by setting the spin variables to fixed constants, for example, as $\mathcal{N} = \{x : \theta_1 = \theta_2 = \theta_3 = 0\}$. This embedding respects the inner product inherited from the ambient space, so the columns of S form a basis for tangent vectors to $\mathcal{N}$ and Q is the metric tensor on $\mathcal{N}$ in the variables (ω, ϕ) [48]. The equilibrium density is the surface measure on $\mathcal{N}$, which in these variables is $\pi_{\text{slide}}(\omega, \phi) = |Q|^{1/2}$. Result Eq. (10) follows from Eqs. (15)–(17).

Next consider the invariant measure *with* the rolling constraints. The stationary probability density $\pi_{\text{roll}}(\omega, \phi)$ for Eq. (19) solves the stationary Fokker-Planck equation,

$$\sum_{i,j=1}^{2} \partial_i(c_{ij}\pi_{\text{roll}} + d_{ij}\partial_j\pi_{\text{roll}}) = 0, \qquad (21)$$

where $c_{ij} = b_i\partial_j b_j$, $d_{ij} = b_i b_j$, and b_i is the ith row of the matrix $B = Q^{-1}(PS)^T \in \mathbb{R}^{2\times 9}$. Here $\partial_1 = \partial_\omega$, $\partial_2 = \partial_\phi$. The boundary condition is the one that conserves probability: a no-flux boundary condition in ω that requires $\sum_j(c_{1j}\pi_{\text{roll}} + d_{1j}\partial_j\pi_{\text{roll}}) = 0$ at $\omega = \pi/6, 5\pi/6$ (or $\omega = 0, \pi$), and a periodic

boundary condition in ϕ. To determine b_i we first compute an orthonormal basis of P, as

$$t_r = (0, \dots, 0, 1, -1, 1)/\sqrt{3},$$

$$t_\omega = \left(\frac{\partial\bar{x}}{\partial\omega}, -2, 0, 2\right)/\sqrt{K^2 + 8}, \qquad (22)$$

$$t_\phi = \left(\frac{\partial\bar{x}}{\partial\phi}, \frac{2}{3}, \frac{4}{3}, \frac{2}{3}\right)/\sqrt{L^2 + \frac{8}{3}}.$$

These are obtained as follows: t_r is the motion obtained by fixing the positions of the disks and only letting them spin; we call this "pure spinning." For t_ω we prescribe the first six components to be $\frac{\partial\bar{x}}{\partial\omega}$ and solve the two linear Eqs. (2) for $\dot{\theta}$. There is a one-parameter family of solutions $\dot{\theta} = (-2, 0, 2) + \dot{\theta}_2(-1, 1, -1)$. We choose the one that minimizes $|\dot{\theta}|^2$ or equivalently that is perpendicular to t_r. For t_ϕ we similarly fix the first six components to be $\frac{\partial\bar{x}}{\partial\phi}$ and solve for $\dot{\theta}$. The solutions are $\dot{\theta} = (0, 2, 0) + \dot{\theta}_1(1, -1, 1)$, and we choose the one with minimum L_2-norm. Each set of solutions for $\dot{\theta}$ has physical meaning since they tell us how the disks must spin, like gears, to produce a desired motion of the cluster in position space. They are each equal to a fixed combination of spins ($\dot{\theta} = (-2, 0, 2)$ to change the internal angle, and $\dot{\theta} = (0, 2, 0)$ to rotate the cluster overall), plus an arbitrary multiple of the pure spinning motion t_r.

We project each column of S using Eq. (22) to find $P\frac{\partial x}{\partial\omega} = \frac{K^2}{\sqrt{K^2+8}}t_\omega$, $P\frac{\partial x}{\partial\phi} = \frac{L^2}{\sqrt{L^2+8}}t_\phi$, so

$$b_1 = \frac{t_\omega^T}{\sqrt{K^2 + 8}}, \quad b_2 = \frac{t_\phi^T}{\sqrt{L^2 + 8/3}}. \qquad (23)$$

From this, computing the c_{ij}, d_{ij} in Eq. (21) is a matter of algebra. We eventually write Eq. (21) as

$$\partial_1\big[(\alpha_1\alpha_1' - LL'\alpha_1^2\alpha_2^2)\pi_{\text{roll}} + \alpha_1^2\partial_1\pi_{\text{roll}}\big]$$
$$+ \partial_2\big(\alpha_2^2\partial_2\pi_{\text{roll}}\big) = 0, \qquad (24)$$

where $'$ denotes a derivative with respect to ω, and

$$\alpha_1(\omega) = |b_1| = (K^2(\omega) + 8)^{-1/2},$$
$$\alpha_2(\omega) = |b_2| = (L^2(\omega) + 8/3)^{-1/2}. \qquad (25)$$

A solution that is independent of ϕ is $\pi_{\text{roll}}(\omega) \propto (\alpha_1\alpha_2)^{-1} = (K^2(\omega) + 8)^{1/2}(L^2(\omega) + 8/3)^{1/2}$. One can check this satisfies the boundary conditions, so Eq. (10) holds, as claimed.

2. Soft constraints

Given constraints of the form $q_i(x) = z$ (where i indexes the constraints), which are imposed by a stiff potential energy, such as $U(x) = \epsilon^{-1}|q_i(x) - z|^2$ with $\epsilon \ll 1$, we can obtain the distribution for infinitely stiff soft constraints from that for hard ones. This is done by multiplying the distribution for hard constraints by a factor of $|A|^{-1/2}$, where A is the Gram matrix of $\{\nabla q_i\}$ evaluated at $q_i(x) = z$ [36,37,44].

If we assume the bond-distance constraints are imposed softly by spring-like forces so that $q_1(x) = |x^{(1)} - x^{(2)}|$, $q_2(x) = |x^{(2)} - x^{(3)}|$, and $z = 1$, then one can calculate using Eq. (13) that $|A|^{1/2} \propto (1 + 2\cos^2\omega)^{-1/2}(1 + 2\sin^2\omega)^{-1/2}$.

Including this factor in Eq. (10) shows the probabilities including these vibrational modes are given by Eq. (11).

To impose the rolling constraints softly, they must be holonomic, meaning they imply two additional constraints in configuration space only. This is the case when the rows of $C(x)$ are each a perfect gradient, but it can also hold when some nonlinear combinations of the rows are. It turns out that although each individual rolling constraint is not a perfect gradient, they are still holonomic after multiplying by a suitable integrating factor (Appendix E). Therefore, with knowledge of these additional constraints one could write down the equilibrium density π_{roll} immediately in the same way we did for π_{slide}. Such an approach may be able to consider larger, more general collections of disks.

We do not attempt to impose the rolling constraints softly here, for at least two reasons. One, because it is not clear whether the additional conserved quantities in configuration space come from a stiff potential that is the origin of the friction force, or whether they are accidents of our two-dimensional geometry; this probably depends on the details of how the friction comes about. Two, because there are infinitely many functions q_i that have the same level set and we currently have no physical principle with which to choose one.

V. DISCUSSION

A. Physical interpretation

It is surprising that the equilibrium distributions for the trimer with and without rolling constraints are different, since according to classical statistical mechanics, if there is no external force on a system then each outer disk should be uniformly distributed on the central one, so the internal angle distribution is flat as for π_{slide} in Eq. (11). What then are we to make of this result? Two interpretations are suggested here.

First, one can take this example as a lesson in imposing constraints in a statistical mechanical system, even when these constraints are effective models for mechanical systems and the forces that impose the constraints are derived in an analogous manner to the mechanical system. Similar to the much-discussed difference between hard and soft constraints in configuration space, one must even be careful when imposing constraints on velocities, seemingly innocuous because it is not immediately obvious that these should affect distributions in configuration space.

Nevertheless, it is often useful to model systems using constraints—it removes fast, often unnecessary degrees of freedom, and also reduces the dimensionality, making numerical and analytical descriptions more tractable [see, e.g., Ref. [49]). In this first interpretation where we assume the classical statistical mechanical result holds, then this would imply a sort of "roughness" entropy associated with the velocity constraint, which would provide an additional force that would counteract the effect of the constraint and keep the equilibrium angle distribution constant. Such a roughness entropy would be similar in spirit to a vibrational entropy, but different in form because it should not necessarily be possible to obtain it as a harmonic expansion of a function of variables in configuration space only. Indeed, the constraint that models a sphere rolling on a plane is nonholonomic [50,51], so

any jiggling about the constraint cannot depend only on the location and overall rotation of the sphere. Even for a pair of disks, one may wish to allow irreversible, nonharmonic slippage about their points of contact.

To see why this suggestion is plausible, imagine the following: three gears on a slippery plane, subject to stochastic fluctuations (such as by vibrations or fluctuations from the surrounding medium), whose centers are bound by elastic spring forces as for the trimer. The gears must roll in order to change their internal angle because the teeth are long and the spring forces strong. The teeth of the gears must have small gaps between them if the setup is to have nonzero probability, and the tangential rattling of the gears within these gaps could provide the conjectured roughness entropy in the limit as the teeth becomes smaller and closer together. A similar argument would hold for particles with rough surfaces, where asperities may interlock like gears with randomly spaced teeth. The jiggling of the disks about their points of contact are coupled to the configuration space variables, since depending on the configuration (the angle of the trimer) there could be larger or smaller infinitesimal displacements available. An intriguing possibility is that this collective jiggling could result in a roughness entropy that causes the angle distribution to deviate from a constant, or even the distribution with rolling constraints derived in this paper, since even in a classical equilibrium system it could be the case that the limiting entropy depends on the way in which the limit is obtained, i.e., whether one considers regularly spaced identical gear teeth, randomly spaced teeth with random heights, or some other pattern. The author is not aware of results showing the free energy of a collection of hard particles is a continuous function of their shape.

Second, and perhaps more interestingly, is the literal interpretation of the result, which would imply that particles interacting with friction that creates rolling, have different free energies than those without. This is only possible if friction causes the system to deviate from classical statistical mechanics, which is possible if it involves nonconservative forces or kinetic effects. Dry friction is known to be a complicated, time-dependent, nonequilibrium phenomenon [52,53] that takes energy and dissipates it into heat or sound, via processes ranging from, among others, van der Waals interactions, capillary bridges, covalent bonds, plastic and elastic deformations of the bodies, fracture, wear, and quantum mechanical interactions; it is remarkable that it is so well modeled by the Coulomb interaction law across a vast range of scales [23,54,55]. Yet this Coulomb interaction law involves an intrinsically nonlinear response to applied forcing and therefore is difficult to reconcile with the conditions of the fluctuation-dissipation theorem. It is not so implausible that such a dissipative force could cause a system of particles to deviate from the predictions of classical statistical mechanics; indeed, such deviations are observed in the widely studied area of active particles, where active forcing due to internal motors, chemotaxis, external magnetic fields, and such forces push a system out of equilibrium (see, e.g., Refs. [56–58]). An active component might even be able to create a dissipative force that mimics the effects of rolling. For example, a popular method to create a reversible interaction between colloids is to coat them with strands of sticky DNA, which acts like velcro when

the colloids are close enough. Certain kinds of DNA must consume fuel in order to create an effective colloid-colloid interaction, which pushes the system out of equilibrium, and could arguably cause the colloids to roll preferentially [59]. Relatedly, colloidal particles of many different kinds are being synthesized and simulated where rotational degrees of freedom are actively forced, including particles that actively rotate [60] and look like gears [61], for which this study may provide fundamental and preliminary intuition into a system with a rich and not very well understood phase space.

If Eq. (11) does describe the equilibrium angle distribution of a collection of particles interacting with very strong dry friction or other similar nonequilibrium dissipative forces, then it provides a method to determine experimentally whether friction is present for a certain type of particle: one can construct a trimer that stays connected for long enough to generate sufficient statistics of the internal angle, and then compare the distributions. For example, the probability of a rolling cluster having angle greater than $\omega_c = 2.2$ (where the two densities cross) is 0.48, while that for a sliding cluster is 0.45; measuring $P(\omega > \omega_c)$ could be one way to compare the distributions. Conversely, given a system where strong friction is present, measuring the angle distribution of a trimer (or other cluster of disks or spheres) could be one way to verify whether friction changes its free energy. It is worth noting that gears have been used as the basis for mechanical metamaterials [62], and if these systems are made on smaller scales where thermal effects are relevant, then they could be used to test (or possibly implement) the predictions in this paper, at least for certain kinds of classically imposed velocity constraints.

B. Mathematical interpretation

Even at the mathematical level, it is perhaps surprising that the equilibrium distributions for sliding and rolling disks are different, since rolling constraints do not change the accessible configurations in position space. Some insight into the mathematical reason for why comes from imagining how the constraints alter the amount of white-noise forcing that is projected onto the position variables, producing observable motion. The white noise acts equally in all directions in the subspace spanned by the columns of $P(x)$, but the forcing we observe in the position variables depends on the projection of the noise to the subspace $\mathcal{N} = \{x : \theta_1 = \theta_2 = \theta_3 = 0\}$. The magnitude of this observed forcing depends on the angles between the two subspaces, which varies with x. A stochastic process with no drift spends more time in regions where it diffuses more slowly, so the equilibrium distribution changes accordingly.

As a side note, we can determine the specific magnitude of this projection from the calculations in Sec. IV B. The subspace spanned by the columns of $P(x)$ has an orthonormal basis contained in the columns of $E = (t_\omega, t_\phi, t_r) \in \mathbb{R}^{9 \times 3}$, where the vectors t_i are defined in Eq. (22). The subspace $\mathcal{N}$ where only position variables vary has an orthonormal basis contained in the columns of $F = (K^{-1} \frac{\partial x}{\partial \omega}, L^{-1} \frac{\partial x}{\partial \phi}) \in \mathbb{R}^{9 \times 2}$, where ω, ϕ are defined in Eq. (13) and $K(\omega), L(\omega)$ are defined in Eqs. (16) and (17). The element of area on one subspace changes magnitude when projected to the other subspace by an amount equal to $|E^T F| = |F^T E|$ [63], where the

determinant applied to a rectangular matrix is the product of its singular values. We can calculate this determinant to be $(K^2 + 8)^{-1/2}(L^2 + 8/3)^{-1/2}$, which is consistent with the equilibrium distribution Eq. (10) and also reminiscent of Eq. (25).

Physically, these calculations tell us how much forcing is absorbed by the spinning of the gears, and how much produces observable motion in the internal angle 2ω or the overall rotation ϕ. For example, consider how the cluster might change the angle ω by some small amount $\Delta\omega$. This requires a change in positions with magnitude $|\Delta\bar{x}| \approx K(\omega)|\Delta\omega|$. When disks can slide, all the white noise forcing may be applied to change the angle so the timescale for this change to happen is roughly $\Delta t \approx K^2(\omega)|\Delta\omega|^2/2$. However, if the disks must roll, then Eq. (22) shows that it takes a constant amount of spinning to change the angle by some small amount $\Delta\omega$. This spinning has magnitude $|(-2,0,2)| = \sqrt{8}$ so it absorbs a constant amount of forcing, producing a timescale of $\Delta t \approx (K^2(\omega) + 8)|\Delta\omega|^2/2$. The difference with the sliding case arises because to change ω, the gears must spin in a way that is not proportional to how much they move in position space.

VI. OUTLOOK AND CONCLUSION

We have derived a set of overdamped Langevin equations for systems with linear velocity constraints. We applied this to a trimer of disks whose internal angle can change, and derived the equilibrium distribution in two cases: one where the disks can slide against each other, the other where they must roll. The two distributions are different, which shows that rolling dynamics modeled as velocity constraints can change even such basic things as the free energy of a system.

Whether this model is physically valid depends on the details of how the friction force imposing the velocity constraint arises, a question we do not attempt to answer here since dry friction is a complicated and not fully understood phenomenon. For it to create the demonstrated free-energy difference, the friction must be a nonequilibrium force, to push the system away from the classical Boltzmann equilibrium. Regardless of whether or not it is, this example is a useful lesson in modeling statistical mechanical systems by imposing constraints: even if the constraints act on the velocities, they can still have a fundamental effect on positions. We discussed how in a classical system in equilibrium we would expect the distributions with and without rolling constraints to be the same, and we conjectured that there may be a form of entropy, a "roughness" entropy, associated with the rolling constraints that models the infinitesimal jiggling and slippage of the disks about their points of contact as they roll around each other. Such an entropy would be similar in spirit to a vibrational entropy but structurally different, since it would be associated with the dynamical degrees of freedom and not purely with the locational ones.

We suggested ways to test the predictions of this model via experiments on clusters of colloidal particles, or with a system of gears on a vibrating table, where macroscale measurements like the internal configuration of a cluster may help determine microscale interactions. Experiments that measure the effect of friction on the steady-state properties in any system with stochastic fluctuations would be

valuable, because it is clear we do not have an adequate understanding of this phenomenon, which is becoming increasingly important in soft-matter and other mesoscale systems.

Our model has also suggested problems where new mathematical developments could help shed light on physical systems. Our derivation of the overdamped Langevin equations is valid for arbitrary linear velocity constraints, both holonomic and nonholonomic. In the former case the overdamped equations describe a Brownian motion on a manifold, whose equilibrium distribution is the surface measure on the manifold, but in the latter there is no such interpretation. While it turns out that disks in the plane are holonomic, a cluster of spheres should be nonholonomic: it can access a space that is higher-dimensional than the space along which it is constrained to move. This should be true because a single sphere rolling on a plane is nonholonomic [50,51]. Geometrically, it lives on a sub-Riemannian manifold [64–66], an object that has been little studied in the physics literature. In this case there is no general method to determine the equilibrium distribution of Eq. (8), since there is no canonical volume form (surface measure) on a sub-Riemannian manifold [67]. It is difficult to even identify a Laplacian, since it is not clear which volume form to use to define the divergence operator, though some recent progress has been made in comparing different choices [67,68]. One could probably work out the equilibrium distribution for individual cases directly as we have done in this paper, but the delicacy of parameterizing $SO(3)$ requires separate treatment. Extending this study to spheres would not only potentially provide an experimental method to determine whether friction is present, but would also bring insight into the physics of stochastic, nonholonomic systems, which have rarely been considered.

APPENDIX A: MASS-SCALED COORDINATES

We show how the mass matrix M can be eliminated from Eq. (4) by a suitable change of variables. This is not a nondimensionalization and the mass still appears implicitly in the new variables. Let $y = M\dot{x}$, $\tilde{x} = M^{1/2}x$, $\tilde{y} = M^{-1/2}y = M^{1/2}\dot{x}$. Then Eq. (4) becomes

$$\dot{\tilde{x}} = \tilde{y},$$

$$\dot{\tilde{y}} + \tilde{\Gamma}\dot{\tilde{x}} = \nabla_{\tilde{x}}\tilde{U}(\tilde{x})dt + \tilde{\sigma}\eta(t) - \tilde{C}^T\lambda, \quad \text{(A1)}$$

and the constraints become

$$\tilde{C}(\tilde{x})\dot{\tilde{y}} = 0. \quad \text{(A2)}$$

Here

$$\tilde{\Gamma}(\tilde{x}) = M^{-1/2}\Gamma(M^{-1/2}\tilde{x})M^{-1/2},$$

$$\tilde{\sigma}(\tilde{x}) = M^{-1/2}\sigma(M^{-1/2}\tilde{x}),$$

$$\tilde{U}(\tilde{x}) = U(M^{-1/2}\tilde{x}),$$

$$\tilde{C}(\tilde{x}) = C(M^{-1/2}\tilde{x})M^{-1/2}.$$

The friction and forcing remain in fluctuation-dissipation balance. Equations (A1) and (A2) have exactly the same structure as Eqs. (4) and (3), respectively, so hereafter we work in these mass-scaled coordinates and remove the tildes.

APPENDIX B: SOLVING FOR THE LAGRANGE MULTIPLIERS

The time derivative of Eq. (3) is

$$C(x)\ddot{x} + \nabla C(x)(\dot{x},\dot{x}) = 0, \quad \text{(B1)}$$

where the second term is a vector with components $(\nabla C(x)(\dot{x},\dot{x}))_i = \sum_{i,j}\frac{\partial c_i}{\partial x_j}\dot{x}_i\dot{x}_j$. Substituting for $\ddot{x}$ from Eq. (4) gives

$$C^T\lambda = -P^\perp\Gamma\dot{x} + P^\perp\sigma\eta + P^\perp\nabla U(x)$$
$$+ C^T G^{-1}\nabla C(\dot{x},\dot{x}), \quad \text{(B2)}$$

where $P^\perp(x)$ is the projection matrix onto the row space of $C(x)$, and $G(x)$ is the Gram matrix. Specifically,

$$P^\perp = C^T G^{-1}C, \qquad G = CC^T. \quad \text{(B3)}$$

One can check that $(P^\perp)^2 = P^\perp$, and $(P^\perp)^T = P^\perp$ so it is orthogonal. Substituting for λ in Eq. (4) gives

$$\ddot{x} + P\Gamma\dot{x} = -P\nabla U(x) + P\sigma\eta + C^T G^{-1}\nabla C(\dot{x},\dot{x}). \quad \text{(B4)}$$

Here $P(x) = I - P^\perp(x)$ is the projection of the velocities onto the tangent space to the manifold in phase space satisfying the constraints. One can check that $P^T = P$ so it is an orthogonal projection.

APPENDIX C: DERIVATION OF THE OVERDAMPED DYNAMICS

In this section we derive the equations for the dynamics in configuration space (x variables only) when viscous friction is large, and over long timescales. Let $y = \dot{x}$, let $\Gamma \to \Gamma/\epsilon$, and let $t \to t/\epsilon$, with $\epsilon \ll 1$. The equations become

$$\dot{x} = \frac{Py}{\epsilon},$$

$$\dot{y} = -\frac{\Gamma_P y}{\epsilon^2} + \frac{\sigma_P}{\epsilon}\eta - \frac{P\nabla U}{\epsilon} + \frac{C^T G^{-1}\nabla C(y,y)}{\epsilon}. \quad \text{(C1)}$$

We have defined $\Gamma_P = P\Gamma P$ and $\sigma_P = P\sigma$. These terms remain in fluctuation-dissipation balance, and Γ_P is symmetric. We can replace y with Py, since the dynamics preserves the constraint $C(x)y = 0$.

The backward equation for Eq. (C1) is

$$\frac{\partial\phi}{\partial t} = \frac{\mathcal{L}_0\phi}{\epsilon^2} + \frac{\mathcal{L}_1\phi}{\epsilon}, \quad \text{(C2)}$$

where

$$\mathcal{L}_0 = -\Gamma_P y \cdot \nabla_y + \beta^{-1}\Gamma_P : \nabla_y^2,$$

$$\mathcal{L}_1 = Py \cdot \nabla_x - P\nabla U \cdot \nabla_y + C^T G^{-1}\nabla C(y,y) \cdot \nabla_y.$$

We write ∇_x, ∇_y for the gradient acting only on the x, y variables, respectively.

We formally expand the solution to Eq. (C2) as $\phi = \phi_0 + \epsilon\phi_1 + \epsilon^2\phi_2 + \ldots$, and collect terms of the same order. The leading order equation is $\mathcal{L}_0\phi_0 = 0$. Since $\mathcal{L}_0$ acts only on the y variables, we must have

$$\phi_0(x,y,t) = \phi_0(x,t). \tag{C3}$$

The next-order equation is $-\mathcal{L}_0\phi_1 = \mathcal{L}_1\phi_0$. Since $\mathcal{L}_0$ is linear in y, this is straightforward to solve, as

$$\phi_1 = \Gamma_P^\dagger y \cdot \nabla_x \phi_0, \tag{C4}$$

where $\Gamma_P^\dagger$ is the Moore-Penrose pseudoinverse of Γ_P. To check this, we calculate

$$-\mathcal{L}_0\phi_1 = \Gamma_P y \cdot \nabla_y(\Gamma_P^\dagger y \cdot \nabla_x \phi_0) = \Gamma_P y \cdot \Gamma_P^\dagger \nabla_x \phi_0$$
$$= y^T \Gamma_P \Gamma_P^\dagger \nabla_x \phi_0 = y^T P \nabla_x \phi_0$$
$$= Py \cdot \nabla_x \phi_0 = \mathcal{L}_1\phi_0,$$

where we have used the fact that $\Gamma_P^T = \Gamma_P$, and $\Gamma_P\Gamma_P^\dagger$ is an orthogonal projection onto the column space of Γ_P [40], so it equals P.

The final equation is $-\mathcal{L}_0\phi_2 = -\frac{\partial\phi_0}{\partial t} + \mathcal{L}_1\phi_1$. By the Fredholm alternative, a solution exists only if the inner product with any element in the null space of $\mathcal{L}_0^*$ is zero. This gives the solvability condition

$$\int \pi(y)\left(-\frac{\partial\phi_0}{\partial t} + \mathcal{L}_1\phi_1\right)dy = 0, \tag{C5}$$

where $\pi(y)$ is any solution to $\mathcal{L}_0^*\pi(y) = 0$. When the integral above is explicitly evaluated, the fast variables y are eliminated and we obtain an evolution equation for ϕ_0 in the slow variables x.

To calculate this integral explicitly, we first find π, which is the equilibrium distribution for the velocities (the fast variables) when the positions and spins (the slow variables) are held constant. The adjoint of $\mathcal{L}_0$ is

$$\mathcal{L}_0^*\rho = \beta^{-1}\nabla_y \cdot [e^{-\frac{1}{2}\beta|y|^2}\Gamma_P\nabla_y(e^{\frac{1}{2}\beta|y|^2}\rho)]. \tag{C6}$$

We have used the fact that Γ_P is independent of y, to pull it out of the inner gradient. It is clear that the invariant measure is

$$\pi(y) = Z^{-1}e^{-\frac{1}{2}\beta|y|^2}\sigma_{\Sigma_x}(dy)$$
$$= Z^{-1}e^{-\frac{1}{2}\beta|y|^2}\delta(C(x)y)|G|^{1/2}, \tag{C7}$$

where $\sigma_{\Sigma_x}(dy)$ is the surface measure on the linear subspace $\Sigma_x(y) \equiv \{y : C(x)y = 0\}$, and Z is a normalization constant to ensure that $\int \pi(y)dy = 1$. The density must be restricted to Σ_x since the dynamics remain on this subspace. We used the co-area formula $\sigma_{\Sigma_x}(dy) = \delta(C(x)y)|G|^{1/2}dy$ to write Eq. (C7) in both mathematicians' and physicists' notation. The matrix G was defined in Eq. (B3).

Next, we evaluate each of the terms in Eq. (C5). We have $\int \pi(y)\frac{\partial\phi_0}{\partial t}dy = \frac{\partial\phi_0}{\partial t}$. The other terms are

$$\mathcal{L}_1\phi_1 = \underbrace{Py \cdot \nabla_x(y^T\Gamma_P^\dagger\nabla_x\phi_0)}_{\text{term 1}} - \underbrace{P\nabla_x U \cdot \Gamma_P^\dagger\nabla_x\phi_0}_{\text{term 2}}$$
$$+ \underbrace{C^T G^{-1}\nabla_x C(y,y) \cdot \Gamma_P^\dagger\nabla_x\phi_0}_{\text{term 3}}. \tag{C8}$$

Let's evaluate the integral of $\pi(y)$ over each of the terms in turn. We will make use of the following fact:

$$\int \pi(y)y_i y_j dy = \beta^{-1}P_{ij}. \tag{C9}$$

To show this, consider an orthonormal basis $\{e_i\}_{i=1}^d$ of the column space of $P(x)$, and let $z_i = y \cdot e_i$ be the variables lying along these directions. Then

$$\int yy^T e^{-\frac{1}{2}\beta|y|^2}\delta(C(x)y)|G|^{1/2}dy$$

$$= \int Py(Py)^T e^{-\frac{1}{2}\beta|y|^2}\delta(C(x)y)|G|^{1/2}dy$$

$$= \sum_{k,l} e_k e_l^T \int z_k z_l e^{-\frac{1}{2}\beta|z|^2}dz = \sum_{k,l}\beta^{-1}\delta_{kl}e_k e_l^T$$

$$= \sum_k e_k e_k^T = P.$$

We use Eq. (C9) to calculate the integral of term 1:

$$\int \pi(y)y_k y_j\partial_k[(\Gamma_P^\dagger)_{ij}\partial_i\phi_0]dy = \beta^{-1}P_{kj}\partial_k[(\Gamma_P^\dagger)_{ij}\partial_i\phi_0]$$
$$= \beta^{-1}\text{Tr}[P\nabla(\Gamma_P^\dagger\nabla\phi_0)],$$

where $(\nabla v)_{jk} = \partial_k v_j$. The subscript x is removed on the final gradient, since it is no longer needed.

The integral of term 2 is $\Gamma_P^\dagger P\nabla_x U \cdot \nabla_x\phi_0 = \Gamma_P^\dagger\nabla_x U \cdot \nabla_x\phi_0$, since there are no terms containing y. This uses the fact that $\Gamma_P^\dagger P = \Gamma_P^\dagger\Gamma_P\Gamma_P^\dagger = \Gamma_P^\dagger$, by the properties of the pseudoinverse.

Term 3 can be written as $(\nabla_x\phi_0)^T\Gamma_P^\dagger C^T G^{-1}\nabla C(y,y)$. But $\Gamma_P^\dagger C^T = 0$, since $\Gamma_P^\dagger C^T = \Gamma_P^\dagger\Gamma_P\Gamma_P^\dagger C^T = \Gamma_P^\dagger PC^T = 0$, using the properties of the pseudoinverse and the fact that the columns of C^T are orthogonal to P. Therefore, this term equals 0.

Putting this together gives the following evolution equation for ϕ_0:

$$\frac{\partial\phi_0}{\partial t} = \Gamma_P^\dagger\nabla U \cdot \nabla\phi_0 + \beta^{-1}\text{Tr}[P\nabla(\Gamma_P^\dagger\nabla\phi_0)]. \tag{C10}$$

APPENDIX D: NUMERICALLY SIMULATING THE LANGEVIN EQUATIONS

We numerically simulated the Langevin Eqs. (4) by writing this second-order equation as two first-order equations for the positions and spins $q = x$ and momenta $p = m\dot{x}$. We used a mass m and friction coefficient γ that were the same for all variables. We alternated updates of q, p by cycling through the following four steps:

(1) Update q by increment $\Delta q = p/m \, \Delta t$;

(2) Project q to manifold where bond-distance constraints Eq. (1) are exactly satisfied (the projection method was the same as that used in [49]);

(3) Update p by increment $\Delta p = -(\gamma/m)p\Delta t + \sigma\sqrt{\Delta t}\, N$, where $N \in \mathbb{R}^9$ is a vector of independent standard normal random variables;

(4) Project $p + \Delta p$ to space of allowed velocities [this is done by multiplying by matrix P defined in Eq. (6)].

Apart from the projection step, this is exactly an Euler-Maruyama method so is expected to be weakly first-order accurate [69]. We did not include a nonoverlap condition for the disks, though this is easily accounted for *a posteriori* by truncating the histogram. The parameters used were $m = 0.1$, $\gamma = 1$, $\sigma = 1$. We set $\Delta t = 5 \times 10^{-3}$ for the sliding simulations, and $\Delta t = 1 \times 10^{-4}$ for the rolling ones. A finer time step was needed for the rolling simulations to get good agreement with the theory, presumably because the simulations do not conserve the additional implied conserved quantities in configuration space Eq. (E3) (see Appendix E.) The total time each simulation was run for was $T_{\max} = 10^5$ for sliding disks and $T_{\max} = 1.8 \times 10^5$ for rolling disks. We needed to run the rolling simulations longer than the sliding ones to converge to the equilibrium distribution, because the effective diffusion coefficient in angle space is smaller.

APPENDIX E: GEOMETRY OF THE ROLLING TRIMER'S CONFIGURATION SPACE

To calculate the trimer's equilibrium distribution, we did not need to know the geometric structure of its configuration space—neither the constants of integration nor the dimension of the manifold on which it lives. The calculation was possible because of the symmetries that let us project the dynamics to a lower-dimensional manifold without losing information, and on this lower-dimensional manifold the trimer had no constraints. Nevertheless, this geometric structure is an interesting mechanics problem in itself.

Let us count degrees of freedom: we began with five parameters to describe the configuration space, and two constraints, so there is a three-dimensional space along which the cluster can move (the "horizontal space"). What is the actual dimension of the space in which it lives?

This can be understood by calculating iterated Lie brackets of the horizontal space. It is simplest to do this in the parameterized space, in which an orthogonal basis of horizontal tangent vectors (proportional to t_ω, t_ϕ, t_r) is

$$T_\omega = (1, 0, -2, 0, 2),$$
$$T_\phi = (0, 3, 2, 4, 2), \qquad \text{(E1)}$$
$$T_r = (0, 0, 1, -1, 1).$$

(See subsection *a* for an explanation.) In this parametrization the horizontal space is a single, constant plane; clearly all Lie

brackets give 0. Therefore, by the Frobenius theorem [70] the trimer lives on a three-dimensional manifold, so there are two extra conserved quantities. One can check that a basis for the normal space is

$$N_1 = (-4, 0, -1, 0, 1), \quad N_2 = (0, -4, 1, 2, 1). \qquad \text{(E2)}$$

These are the gradients of the following scalar functions:

$$Q_1 = -4\omega - \theta_1 + \theta_3, \quad Q_2 = -4\phi + \theta_1 + 2\theta_2 + \theta_3. \qquad \text{(E3)}$$

It is these functions (or any nonlinear function of them) that are conserved by the dynamics with rolling constraints. From these one could calculate the equilibrium distribution directly.

1. Tangent map

The horizontal vector Eqs. (22) come from considering the *tangent map* induced by a smooth map $f : M \to N$ from one manifold M to another manifold N. Recall that a tangent vector at point p can be thought of as an equivalence class of curves $[c]$, where the equivalence relation is $c(t) \equiv d(t)$ if $c'(0) = d'(0)$ and $c(0) = d(0) = p$ [48]. Then, f induces a natural linear map between tangent spaces, $df_p : T_p M \to T_p N$, defined by

$$df_p[c'(0)] = \left.\frac{d}{dt}\right|_{t=0} f[c(t)].$$

If we have a description of the manifolds in the variables $x \in X$, $y \in Y$, where X, Y are subsets of suitable spaces, and if we have a map $f : X \to Y$, then the tangent map is

$$df_p[c'(0)] = \nabla f\, c'(0)$$

$$= \begin{pmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} & \cdots & \frac{\partial x_1}{\partial y_n} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} & \cdots & \frac{\partial x_2}{\partial y_n} \\ \vdots & & & \vdots \end{pmatrix} \begin{pmatrix} c'_1(0) \\ c'_2(0) \\ \cdots \\ c'_n(0) \end{pmatrix}.$$

Let M be the manifold of accessible configurations parameterized by the variables (ω, ϕ, θ), and let N be the same manifold described by the variables x. We have an explicit mapping $f : M \to N$, given by Eq. (13) and the subsequent inline equations. The Jacobian of this mapping in block form is

$$\nabla f = \begin{pmatrix} \frac{\partial \bar{x}}{\partial \omega} & \frac{\partial \bar{x}}{\partial \phi} & 0 \\ 0 & 0 & I_3 \end{pmatrix}, \qquad \text{(E4)}$$

where I_3 is the 3×3 identity matrix and 0 is a matrix of zeros with dimensions correct for the context. From this, we can see that

$$\nabla f\, T_\omega \propto t_\omega, \quad \nabla f\, T_\phi \propto t_\phi, \quad \nabla f\, T_r \propto t_r.$$

[1] N. Y. C. Lin, B. M. Guy, M. Hermes, C. Ness, J. Sun, W. C. K. Poon, and I. Cohen, Phys. Rev. Lett. **115**, 228304 (2015).

[2] R. Mari, R. Seto, J. F. Morris, and M. M. Denn, Proc. Natl. Acad. Sci. USA **112**, 15326 (2015).

[3] M. Mani, A. Gopinath, and L. Mahadevan, Phys. Rev. Lett. **108**, 226104 (2012).

[4] S. Sircar, J. G. Younger, and D. M. Bortz, J. Biol. Dyn. **9**, 79 (2014).

[5] R. Dreyfus, M. E. Leunissen, R. Sha, A. V. Tkachenko, N. C. Seeman, D. J. Pine, and P. M. Chaikin, Phys. Rev. Lett. **102**, 048301 (2009).

[6] R. Macfarlane, B. Lee, M. Jones, N. Harris, G. Schatz, and C. Mirkin, Science **334**, 204 (2011).

[7] W. B. Rogers and J. C. Crocker, Proc. Natl. Acad. Sci. USA **108**, 15687 (2011).

[8] W. B. Rogers and V. N. Manoharan, Science **347**, 639 (2015).

[9] I. C. Jenkins, M. T. Casey, J. T. McGinley, J. C. Crocker, and T. Sinno, Proc. Natl. Acad. Sci. USA **111**, 4803 (2014).

[10] T. Still, C. P. Goodrich, K. Chen, P. J. Yunker, S. Schoenholz, A. J. Liu, and A. G. Yodh, Phys. Rev. E **89**, 012301 (2014).

[11] N. Rivier, J. Non-Cryst. Solids **352**, 4505 (2006).

[12] A. Taboada, N. Estrada, and F. Radjaï, Phys. Rev. Lett. **97**, 098302 (2006).

[13] E. Somfai, M. van Hecke, W. G. Ellenbroek, K. Shundyak, and W. van Saarloos, Phys. Rev. E **75**, 020301 (2007).

[14] F. Radjai and V. Richefeu, Philos. Trans. R. Soc. A: Math. Phys. Eng. Sci. **367**, 5123 (2009).

[15] A. J. Liu and S. R. Nagel, Ann. Rev. Condensed Matter Phys. **1**, 347 (2010).

[16] N. Estrada, E. Azéma, F. Radjaï, and A. Taboada, Phys. Rev. E **84**, 011306 (2011).

[17] P.-G. de Gennes, J. Stat. Phys. **119**, 953 (2005).

[18] H. Hayakawa, Physica D: Nonlin. Phenom. **205**, 48 (2005).

[19] H. Touchette, E. Van der Straeten, and W. Just, J. Phys. A: Math. Theor. **43**, 445002 (2010).

[20] A. M. Menzel and N. Goldenfeld, Phys. Rev. E **84**, 011122 (2011).

[21] P. S. Goohpattader, S. Mettu, and M. K. Chaudhury, Eur. Phys. J. E **34**, 120 (2011).

[22] G. Reiter, A. L. Demirel, and S. Granick, Science **263**, 1741 (1994).

[23] A. Vanossi, N. Manini, M. Urbakh, S. Zapperi, and E. Tosatti, Rev. Mod. Phys. **85**, 529 (2013).

[24] M. Gualtieri, T. Tokieda, L. Advis-Gaete, B. Carry, E. Reffet, and C. Guthmann, Am. J. Phys. **74**, 497 (2006).

[25] T. Tokieda, Amer. Math. Monthly **120**, 265 (2013).

[26] N. M. Bou-Rabee, J. E. Marsden, and L. A. Romero, SIAM Rev. **50**, 325 (2008).

[27] B. Kim and V. Putkaradze, Phys. Rev. Lett. **105**, 244302 (2010).

[28] N. K. Moshchuk and I. N. Sinitsyn, J. Appl. Math. Mech. **54**, 174 (1990).

[29] S. Hochgerner, Rep. Math. Phys. **66**, 385 (2010).

[30] G. Marchegiani and F. Marchesoni, J. Chem. Phys. **143**, 184901 (2015).

[31] F. Gay-Balmaz and V. Putkaradze, J. Nonlinear Sci. **26**, 1571 (2016).

[32] E. J. Hinch, J. Fluid Mech. **72**, 499 (1975).

[33] J.-N. Roux, Physica A: Stat. Mech. Appl. **188**, 526 (1992).

[34] L. D. Landau and E. M. Lifshitz, *Mechanics* (Butterworth-Heinemann, London, 1976).

[35] M. R. Flannery, Am. J. Phys. **73**, 265 (2005).

[36] T. Lelievre, G. Stoltz, and M. Rousset, *Free Energy Computations: A Mathematical Perspective* (Imperial College Press, London, 2010).

[37] G. Ciccotti, T. Lelièvre, and E. Vanden-Eijnden, Commun. Pure Appl. Math. **61**, 371 (2007).

[38] G. A. Pavliotis and A. Stuart, *Multiscale Methods: Averaging and Homogenization* (Springer, Berlin, 2008).

[39] Given an $m \times n$ matrix A, the Moore-Penrose pseudoinverse $A^\dagger$ is the unique matrix that satisfies (i) $A A^\dagger A = A$, (ii) $A^\dagger A A^\dagger = A^\dagger$, (iii) $(A A^\dagger)^T = A A^\dagger$, and (iv) $(A^\dagger A)^T = A^\dagger A$.

[40] G. Strang, *Linear Algebra and Its Applications*, 3rd ed. (Brooks Cole, New York, 1988).

[41] N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes* (Elsevier, New York, 1981).

[42] P. Hsu, Contemp. Math. **73**, 95 (1988).

[43] C. Hartmann and C. Schütte, Commun. Math. Sci. **3**, 1 (2005).

[44] M. Fixman, Proc. Natl. Acad. Sci. USA **71**, 3050 (1974).

[45] E. J. Hinch, J. Fluid Mech. **271**, 219 (1994).

[46] R. W. Perry, M. C. Holmes-Cerfon, M. P. Brenner, and V. N. Manoharan, Phys. Rev. Lett. **114**, 228301 (2015).

[47] Specifically, the blocks are $R_1(\phi) = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}$.

[48] W. Kühnel, *Differential Geometry. Student Mathematical Library* (American Mathematical Society, Providence, Rhode Island, 2002), Vol. 16.

[49] M. Holmes-Cerfon, S. J. Gortler, and M. P. Brenner, Proc. Natl. Acad. Sci. USA **110**, E5 (2013).

[50] B. D. Johnson, Amer. Math. Monthly **114**, 500 (2007).

[51] A. M. Bloch, J. Ballieul, P. Crouch, and J. E. Marsden, *Nonholonomic Mechanics and Control, vol. 24, Interdisciplinary Applied Mathematics* (Springer-Verlag, Berlin, 2003).

[52] O. Ben-David, S. M. Rubinstein, and J. Fineberg, Nature **463**, 76 (2010).

[53] Q. Li, T. E. Tullis, D. Goldsby, and R. W. Carpick, Nature **480**, 233 (2011).

[54] M. Nosonovsky, Entropy **12**, 1345 (2010).

[55] Y. Rezek, Entropy **12**, 1885 (2010).

[56] S. Ramaswamy, Ann. Rev. Condensed Matter Phys. **1**, 323 (2010).

[57] J. Yan, M. Bloom, S. C. Bae, E. Luijten, and S. Granick, Nature **491**, 578 (2012).

[58] J. Palacci, S. Sacanna, A. P. Steinberg, D. J. Pine, and P. M. Chaikin, Science **339**, 936 (2013).

[59] D. Y. Zhang and G. Seelig, Nature Chem. **3**, 103 (2011).

[60] J. Yan, S. C. Bae, and S. Granick, Soft Matter **11**, 147 (2015).

[61] N. H. P. Nguyen, D. Klotsa, M. Engel, and S. C. Glotzer, Phys. Rev. Lett. **112**, 075701 (2014).

[62] A. S. Meeussen, J. Paulose, and V. Vitelli, arXiv:1602.08769 (2016).

[63] A. Björck and G. H. Golub, Math. Comp. **27**, 579 (1973).

[64] R. Montgomery, *A Tour of Subriemannian Geometries, Their Geodesics and Applications*, Mathematical Surveys and Monographs, Vol. 91 (American Mathematical Society, Providence, Rhode Island, 2006).

[65] L. Capogna, D. Danielli, S. D. Pauls, and J. Tyson, *An Introduction to the Heisenberg Group and the Sub-Riemannian Isoperimetric Problem* (Springer Science & Business Media, Berlin, 2007).

[66] M. Gromov, *Metric Structures for Riemannian and non-Riemannian Spaces*, English ed., Modern Birkhäuser Classics, Vol. 152 (Birkhäuser Boston, Inc., Boston, MA, 2007).

[67] D. Barilari and L. Rizzi, Anal. Geom. Metric Spaces (2013).

[68] M. Gordina and T. Laetsch, Potential Anal. **44**, 811 (2016).

[69] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*, Applications of Mathematics (New York), Vol. 23 (Springer-Verlag, Berlin/Heidelberg, 1992).

[70] J. M. Lee, *Manifolds and Differential Geometry*, Graduate Studies in Mathematics Vol. 107 (American Mathematical Society, Providence, Rhode Island, 2009).