Wildebeest herds on rolling hills: Flocking on arbitrary curved surfaces

Christina L. Hueschen [1,*] Alexander R. Dunn [1] and Rob Phillips [2,3,†]

[1]*Department of Chemical Engineering, Stanford University, Palo Alto, California 94305, USA*
[2]*Department of Physics, California Institute of Technology, Pasadena, California 91125, USA*
[3]*Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, California 91125, USA*

The collective behavior of active agents, whether herds of wildebeest or microscopic actin filaments propelled by molecular motors, is an exciting frontier in biological and soft matter physics. Almost three decades ago, Toner and Tu developed a continuum theory of the collective action of flocks, or herds, that helped launch the modern field of active matter. One challenge faced when applying continuum active matter theories to living phenomena is the complex geometric structure of biological environments. Both macroscopic and microscopic herds move on asymmetric curved surfaces, like undulating grass plains or the surface layers of cells or embryos, which can render problems analytically intractable. In this paper, we present a formulation of the Toner-Tu flocking theory that uses the finite element method to solve the governing equations on arbitrary curved surfaces. First, we test the developed formalism and its numerical implementation in channel flow with scattering obstacles and on cylindrical and spherical surfaces, comparing our results to analytical solutions. We then progress to surfaces with arbitrary curvature, moving beyond previously accessible problems to explore herding behavior on a variety of landscapes. This approach allows the investigation of transients and dynamic solutions not revealed by analytic methods. It also enables versatile incorporation of new geometries and boundary conditions and efficient sweeps of parameter space. Looking forward, the paper presented here lays the groundwork for a dialogue between Toner-Tu theory and data on collective motion in biologically relevant geometries, from drone footage of migrating animal herds to movies of microscopic cytoskeletal flows within cells.

I. INTRODUCTION

The beautiful collective motions of flocking or herding animals have mesmerized the human mind for millennia [1] and inspired the modern field of study of active matter [2–4]. In recent decades, active matter theories have been used to describe collective motion in living systems across nearly a billionfold difference in scales [5], from starling flocks wheeling over Rome at 10 m/s [6], to 50 µm/s flows of cytoplasm in cm-size internodal cells of the algae *Chara* [7], to 0.1 µm/s flows of actin filaments and myosin molecules in 50 µm developing *C. elegans* worm embryos [8]. In these macroscopic and microscopic contexts, beautiful self-organization phenomena often take place on surfaces. The collective motion of flocking sheep or kilometer-wide migrating wildebeest herds are influenced by the hills, valleys, or canyons of the landscape on which they travel. Similarly, the flows of actin and myosin molecules mentioned above occur in thin layers at the surface of the embryo or cell, and are thus constrained by its surface

topology and shape. While modern techniques such as dynamical drone imaging [9] and high-resolution fluorescence microscopy [10,11] enable experimental measurements of these animal, cell, or molecule kinematics, a dialogue between measurement and theory requires predictions that incorporate the complex shapes of the real-world surfaces on which they move.

In this paper, we present a general curved-surface formulation and numerical implementation of a minimalistic continuum theory of flocking or herding active matter with the hope that it will prove useful to others interested in exploring continuum theory predictions on complex geometries. For active matter systems such as bacterial swarms [12], active colloidal fluids [13], self-propelled rods [14], or purified cytoskeletal networks [15], studying the contribution of engineered confinement geometries to emergent patterns has proven a fruitful path [16–20]. Understanding the contribution of *biological* geometries to pattern formation is an exciting direction of growth [21–26]. Indeed, our own interest in solving active matter theories on arbitrary curved surfaces was initially inspired by a need to predict emergent actin polarity patterns at the surface of single-celled *Toxoplasma gondii* parasites, whose gliding motility is driven and directed by this surface actin layer [27]. These cells have a beautiful but complex shape which lacks the symmetries that license traditional analytic approaches.

We focus on a classic continuum active matter model originally developed by Toner and Tu [3,4,28], inspired by the

---

*Corresponding author: chueschen@gmail.com
†Corresponding author: phillips@pboc.caltech.edu

work of Vicsek *et al.* [2], to describe the collective behavior of flocking or herding animals. The Toner-Tu theory helped launch the modern field of active matter [29] and can describe collections of dry, polar, and self-propelled agents at any length scale. We develop a general curved surface framework for the Toner-Tu theory and implement it in the finite element setting, enabling its convenient use on complex surfaces. While flocking theory—and, in particular, the herding of wildebeests—serves as our example in this paper, the general curved surface formulation and finite element method (FEM) approach presented here may prove useful for any continuum active matter theory. There is a rich and beautiful literature associated with formulating field equations on arbitrary curved surfaces [30–36]. It has been said of these formulations, "The complexity of the equations may explain why they are so often written but never solved for arbitrary surfaces." [34]. Because of our interest in collective motions on biological surfaces, we could not afford to avoid solving these equations on such surfaces. We used a finite element approach, which can be flexibly adjusted to an arbitrary choice of geometry and permits an exploration of active, self-organized solutions predicted by the Toner-Tu theory. With our general surface formulation and finite element implementation in hand, we test our framework and approach on cylindrical and spherical geometries for which analytic solutions exist as well. Satisfied by the agreement between our numerical and analytic results, we explore flocking phenomena on a broad collection of surfaces.

The remainder of the paper is organized as follows. In Sec. II, we show how the flocking theory of Toner and Tu can be recast in a general surface form to describe flocking motions on arbitrary curved surfaces. In Sec. III, we construct a finite element surface formulation that respects the low symmetry of realistic curved geometries while permitting numerical analysis of the dynamics. In Sec. IV, we turn to the great wildebeest herds for inspiration and perform simple parameter estimates for Toner-Tu wildebeest herding. We then explore the dimensionless ratios that appear when recasting the Toner-Tu equations in dimensionless form. In Sec. V, with the full surface formulation and its numerical implementation in hand, we explore scaling relationships and changes of dynamical state that arise for herds in channels with scattering obstacles, and we use cylindrical and spherical surfaces to compare our FEM results to corresponding analytic solutions. Our approach also allows the exploration of transients and dynamic solutions not revealed by analytic methods. Finally, in Sec. VI, we playfully use the curved-space formalism and its finite element implementation for case studies of wildebeest herding on landscapes of rolling hills, and we note that this paper serves as a theoretical foundation to solve flocking theories on real-world curved surfaces of biological interest.

II. FLOCKING THEORY FOR ARBITRARY CURVED SURFACES

A. A minimal Toner-Tu theory in the plane

The first step in the development of continuum theories of active matter involves selecting the relevant field variables. Following the classic work of Toner and Tu [3,4] and inspired

by observations of herds of land animals like wildebeests, we consider a two-dimensional density field $\rho(\mathbf{r}, t)$ and a corresponding two-dimensional velocity field $\mathbf{v}(\mathbf{r}, t)$, which captures both the orientation and speed of a polar agent. Throughout this paper, wildebeest herds will serve as inspiration and example, although in a sense the term wildebeest is our shorthand for a self-propelled polar agent at any length scale. We also note that the Toner-Tu model used here describes *dry* systems, in which momentum is not conserved and hydrodynamic coupling between the agents is negligible relative to frictional drag. With variables ρ and $\mathbf{v}$ in hand, our next step is to write the partial differential equations that describe the spatiotemporal evolution of those field variables. The first governing equation is the continuity equation given by

$$\frac{\partial \rho}{\partial t} + \frac{\partial (\rho v_i)}{\partial x_i} = 0, \tag{1}$$

which allows modulations in the density field but enforces mass conservation, forbidding wildebeest birth or death in the midst of herding phenomena. Note that we are using the Einstein summation convention, which tells us to sum over all repeated indices. For example, $\mathbf{a} \cdot \mathbf{b} = \sum_{i=1}^{3} a_i b_i = a_i b_i$. The second governing equation is a minimal representation of the dynamics of the $\mathbf{v}$ field offered by Toner and Tu [3] and is given by

$$\frac{\partial v_i}{\partial t} = [\alpha(\rho - \rho_c) - \beta v_j v_j] v_i - \sigma \frac{\partial \rho}{\partial x_i} + D\nabla^2 v_i - \lambda v_j \frac{\partial v_i}{\partial x_j}, \tag{2}$$

where ρ_c is the critical density above which the herd moves coherently, the ratio of $\alpha(\rho - \rho_c)$ and β sets the wildebeest mean-field speed $v_{\text{pref}} = \sqrt{\alpha(\rho - \rho_c)/\beta}$, the term $\sigma \partial \rho / \partial x_i$ provides an effective pressure, D tunes wildebeest alignment and speed matching with neighbors, and λ tunes velocity self-advection. We can conceptualize each term as an update rule that computes incremental changes in the velocity vector at each point in space and at every step in time. In Fig. 1 and in the remainder of this section, we seek to provide an intuitive interpretation of each term and its contribution to updating the velocity field.

The first term on the right-hand side of Eq. (2), the preferred speed term, pushes the velocity magnitude toward the characteristic speed of a wildebeest, v_{pref} [37]. The second is based on an equation of state originally used by Toner and Tu to relate density and pressure [4]. This pressure term punishes gradients in density, adjusting velocity to flatten the density field. The third, the neighbor coupling term, provides a smoothing or diffusion of velocity (orientation and speed) that reflects coordination between nearby wildebeests. Interestingly, if wildebeests adopt the rule that a given wildebeest averages the difference between its own velocity and that of its neighbors, mathematically, the result is a Laplacian term like that seen in the minimal model [37]. The final term has analogy to the gradient component of the material time derivative in the Navier-Stokes equations. In essence, the velocity field advects itself; wildebeests move along a direction dictated by their orientation and velocity, and they bring that orientation and velocity with them. In the case of pure velocity self-advection, $\lambda = 1$, but this is not necessarily the case for
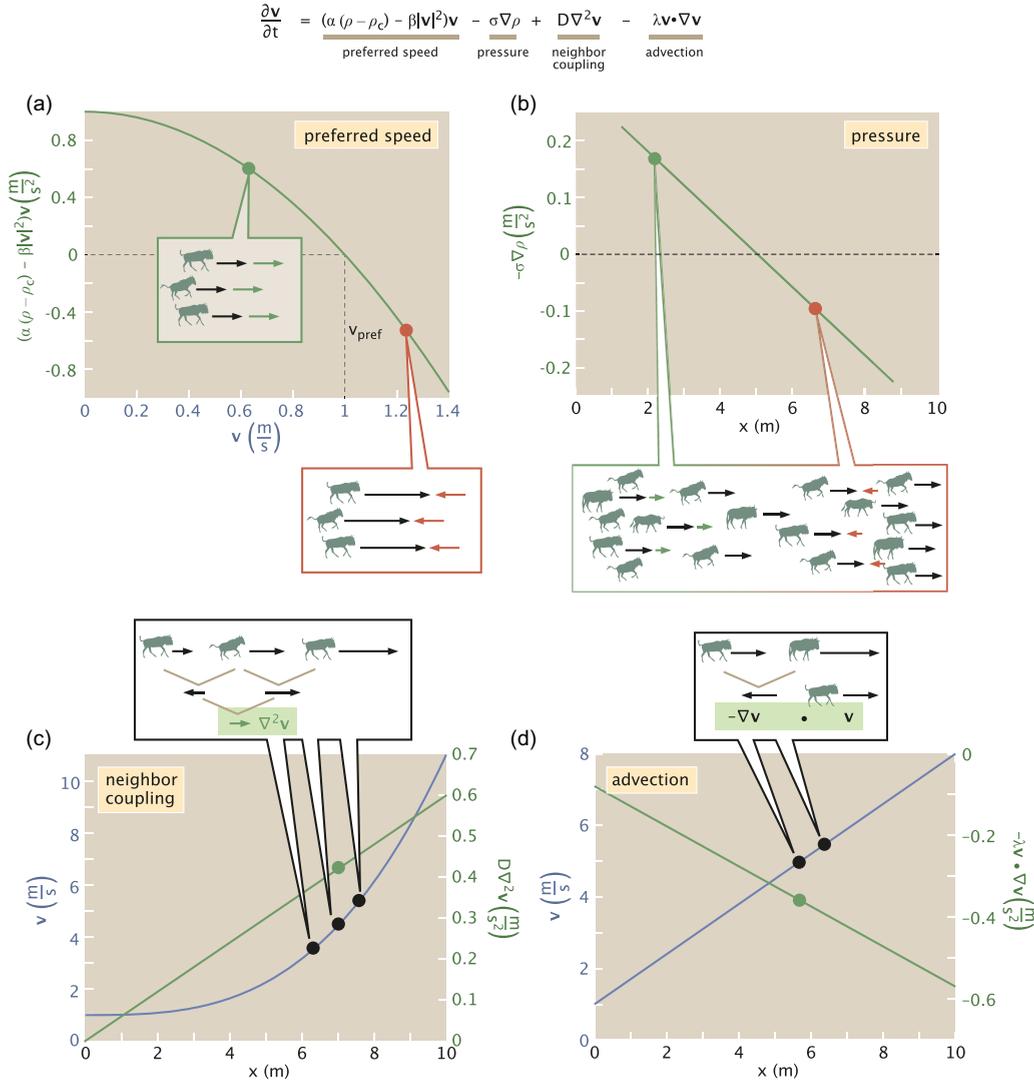
$$\frac{\partial \mathbf{v}}{\partial t} = \underbrace{(\alpha\,(\rho - \rho_c) - \beta|\mathbf{v}|^2)\mathbf{v}}_{\text{preferred speed}} - \underbrace{\sigma\nabla\rho}_{\text{pressure}} + \underbrace{D\nabla^2\mathbf{v}}_{\substack{\text{neighbor}\\\text{coupling}}} - \underbrace{\lambda\mathbf{v}\cdot\nabla\mathbf{v}}_{\text{advection}}$$

FIG. 1. Building intuition for the physics of the velocity field in the Toner-Tu theory. The different terms in the Toner-Tu analysis are represented graphically for a one-dimensional herd. We can think of each term as providing an update to the current velocity. In each case, an example of a simple velocity profile (blue) or density profile [graphic in (b)] is shown alongside the corresponding velocity update (green). (b) The preferred speed term increases the velocity of wildebeests that are moving too slow and decreases the velocity of wildebeests that are moving too fast. (b) The pressure term punishes gradients in density, adjusting velocity to flatten these gradients. Given mass conservation in a finite system, the pressure term alone would lead to a steady state of uniform density $\rho = \rho_0$. (c) The neighbor coupling term captures the velocity adjustment made by a wildebeest to better match its neighbors, smoothing out differences in velocity. A given wildebeest (middle black dot) adjusts its velocity by an amount represented by the green arrow: an average of the difference between its velocity and its two neighbors' velocities. This averaging of two differences is mathematically analogous to taking a local second derivative or local Laplacian, $\nabla^2\mathbf{v}$. (d) Finally, the advection term ensures that the filament velocity field is swept along according to its own velocities, just as fluid velocity is self-advected in the Navier-Stokes equations. Mathematically, the velocity adjustment needed to ensure velocity advection is a function of the spatial gradient of velocity ($-\nabla\mathbf{v}$, how mismatched in velocity nearby wildebeests are) and the velocity itself ($\mathbf{v}$, how fast that mismatch is carried along by the herd).

active flocks [37,38]. In the case of wildebeest herds, we can conceptualize this effect by considering that $\lambda = 1 + \xi$, where ξ reflects a behavioral response to gradients in velocity; for example, wildebeests may resist running quickly into a steep gradient of decreasing velocity, and may slow down. For a full pedagogical derivation of the Toner-Tu model, we recommend a series of lectures by Toner [37] in which he uses symmetry arguments to infer what terms should be kept in a complete continuum description of herding and also provides intuitive arguments about what these terms mean.

B. Formulating the theory for arbitrary curved surfaces

Inspired by a desire to make contact with phenomena of the natural world, like wildebeests navigating an undulating landscape, sheep flocks crossing hilly pastures, and the surfaces flows of flocking actin we study in Ref. [27], we sought to solve the minimal Toner-Tu theory presented above on complex and asymmetric surface geometries. We consider here flocking on nondeformable surfaces, but we refer the interested reader to earlier work on surface hydrodynamics
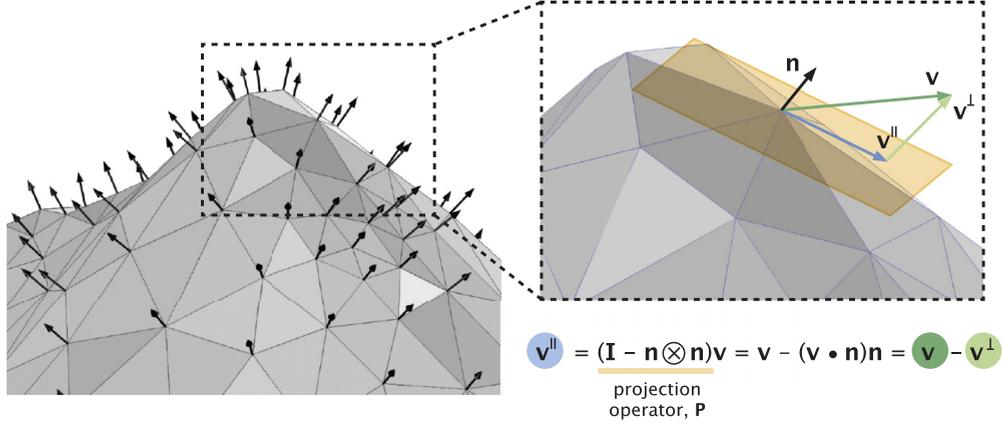
FIG. 2. Illustration of the projection operator essential to the general surface formulation and finite element treatment of the Toner-Tu equations on curved surfaces. The surface of interest is represented by a collection of nodes (vertices of triangles) and a field of surface normal vectors, $\mathbf{n}$ (black arrows). An arbitrary vector $\mathbf{v}$ is projected onto the tangent plane (shown in gold), permitting a decomposition of the form $\mathbf{v} = \mathbf{v}^{\parallel} + \mathbf{v}^{\perp}$.

in the completely general case in which the surface itself can evolve over time [32–34,36]. Predicting flocking or herding behavior on arbitrary surfaces requires us to reformulate the theory in a more general way that accounts for curvature. Instead of basing our formulation on a parameterized surface and the intrinsic differential geometry tools that this approach licenses, as done in Ref. [39], from the outset we have in mind arbitrary surfaces that can be represented using finite element meshes and described by a field of local normal vectors, $\mathbf{n}$.

Our choice and use of this extrinsic differential geometry and FEM approach was aided by the work of many, including Refs. [33–35,40–45] and Chap. 3 of the Supplemental Material for Ref. [46]. The finite element setting permits us the convenience of carrying out the mathematics in the full three-dimensional setting of $\mathbb{R}^3$, while using our knowledge of the normal vectors everywhere on the surface of interest to project our governing equations onto the surface. While velocity is described by the three-dimensional vector $\mathbf{v}$, both $\mathbf{v}$ and the scalar ρ are defined only on the surface. At every point on the surface, derivatives evaluated in the usual $\mathbb{R}^3$ way are projected onto the tangent plane using the local normal. For an insightful description of this extrinsic differential geometry approach to handling curved surfaces and its mathematical equivalence to the intrinsic differential geometry strategy, we recommend Chaps. 22 and 23 of Needham [47]. Central to the extrinsic geometry approach is the projection operator, defined as

$$\text{projection onto tangent plane} = \mathbf{P} = \mathbf{I} - \mathbf{n} \otimes \mathbf{n}, \quad (3)$$

where $\mathbf{I}$ is the identity matrix and $\mathbf{n} \otimes \mathbf{n}$ is the outer product of the surface normal vector as shown in Fig. 2. To make sense of this expression mathematically, we recall that the outer product $\mathbf{n} \otimes \mathbf{n}$ is defined through its action on a vector $\mathbf{v}$ as

$$(\mathbf{n} \otimes \mathbf{n})\mathbf{v} = \mathbf{n}(\mathbf{n} \cdot \mathbf{v}). \quad (4)$$

We can write $\mathbf{v} = \mathbf{v}^{\parallel} + \mathbf{v}^{\perp}$, where $\mathbf{v}^{\parallel}$ is the component of $\mathbf{v}$ in the tangent plane of the surface and $\mathbf{v}^{\perp}$ is normal to the surface. The action of $\mathbf{I} - \mathbf{n} \otimes \mathbf{n}$ on a vector $\mathbf{v}$ is given by

$$(\mathbf{I} - \mathbf{n} \otimes \mathbf{n})\mathbf{v} = \mathbf{v} - \mathbf{n}(\mathbf{n} \cdot \mathbf{v}) = \mathbf{v} - \mathbf{v}^{\perp} = \mathbf{v}^{\parallel}, \quad (5)$$

as illustrated in Fig. 2. We can write the projection operator in component form as

$$P_{ij} = \delta_{ij} - n_i n_j, \quad (6)$$

recalling that the normal vector to the surface is given by $\mathbf{n} = (n_1, n_2, n_3)$, or in full component form as

$$\mathbf{P} = \begin{bmatrix} 1 - n_1^2 & -n_1 n_2 & -n_1 n_3 \\ -n_1 n_2 & 1 - n_2^2 & -n_2 n_3 \\ -n_1 n_3 & -n_2 n_3 & 1 - n_3^2 \end{bmatrix}. \quad (7)$$

Using the projection operator allows us to perform calculations in the ordinary three-dimensional space within which the surface of interest is embedded, but then to pick off only the pieces of the resulting vectors that live within the surface.

In the remainder of this section, we examine each term in the minimal Toner-Tu equations and translate it into its projected form. In Appendix Sec. 1, we define the projected surface form of calculus operators as an additional reference for the reader. First, to modify the preferred speed term to its curved-surface implementation, we note that it is now $\mathbf{v}^{\parallel}$, the in-plane velocity, that has a privileged magnitude. This magnitude is imposed through the condition

$$\text{preferred speed term} = [\alpha(\rho - \rho_c) - \beta v_j^{\parallel} v_j^{\parallel}] v_i^{\parallel}. \quad (8)$$

If $|\mathbf{v}^{\parallel}|$ is either larger or smaller than the privileged value $v_{\text{pref}} = \sqrt{\alpha(\rho - \rho_c)/\beta}$, this term will adjust the velocity towards that steady-state magnitude.

Next, we consider the pressure term, whose physical origin comes from a model of pressure in powers of density of the form introduced by Toner and Tu [3],

$$P(\rho) = \sum_{n=1}^{\infty} \sigma_n (\rho - \rho_0)^n, \quad (9)$$

where ρ_0 is the mean density. In the minimal Toner-Tu theory we adopt here, only the first-order term in that expansion is kept, with the notational simplification that $\sigma_1 = \sigma$,

resulting in

$$\text{pressure term} = -\frac{\partial P}{\partial x_i} = -\sigma \frac{\partial \rho}{\partial x_i}. \tag{10}$$

We note that while ρ_0 is not explicitly present in the gradient of the first-order pressure term, for the case considered here of a finite surface on which total density is conserved, this pressure term effectively maintains a density range centered around the mean density ρ_0 established by our choice of initial condition. The curved-space version of the pressure term requires projecting the full 3D gradient of the density onto the tangent plane, using the projection operator defined in Eq. (3). We follow Jankuhn *et al.* [34] in introducing the notation ∇_Γ for the projected surface gradient operator, where the subscript Γ indicates that the gradient is evaluated on the surface of interest. Mathematically, this amounts to computing

$$\nabla_\Gamma \rho = (\mathbf{I} - \mathbf{n} \otimes \mathbf{n})\nabla \rho, \tag{11}$$

where ∇ is the ordinary, three-dimensional gradient in Cartesian coordinates. This can be rewritten in component form as

$$\begin{aligned}(\nabla_\Gamma \rho)_i &= [(\mathbf{I} - \mathbf{n} \otimes \mathbf{n})\nabla \rho]_i \\ &= (\delta_{ij} - n_i n_j)\frac{\partial \rho}{\partial x_j} \\ &= \frac{\partial \rho}{\partial x_i} - n_i n_j \frac{\partial \rho}{\partial x_j}. \end{aligned} \tag{12}$$

The next term in the Toner-Tu equations that we consider in its curved-space format is the advection term in Eq. (2), namely,

$$\text{advection term} = -\lambda v_j \frac{\partial v_i}{\partial x_j}. \tag{13}$$

In direct notation, the curved-space version of this term has the form

$$\text{advection term} = -\lambda(\mathbf{v}^\| \cdot \nabla_\Gamma)\mathbf{v}^\|, \tag{14}$$

which involves the curved-space version of the velocity gradient tensor. In Appendix Sec. 1, we describe how to compute this tensor, which stated simply is

$$\nabla_\Gamma \mathbf{v}^\| = \mathbf{P}(\nabla \mathbf{v}^\|)\mathbf{P}, \tag{15}$$

where $\mathbf{P}$ is the projection operator defined in Eq. (3). In indicial notation, this leads to the result

$$(\nabla_\Gamma \mathbf{v}^\|)_{ij} = [\mathbf{P}(\nabla \mathbf{v}^\|)\mathbf{P}]_{ij} = P_{ik}\frac{\partial v_k^\|}{\partial x_l}P_{lj}. \tag{16}$$

Invoking the definition of the projection operator from Eq. (6), this expression simplifies to the result

$$\begin{aligned}(\nabla_\Gamma \mathbf{v}^\|)_{ij} &= [\mathbf{P}(\nabla \mathbf{v}^\|)\mathbf{P}]_{ij} \\ &= \left(\frac{\partial v_i^\|}{\partial x_j} - n_l n_j \frac{\partial v_i^\|}{\partial x_l}\right) \\ &\quad - n_i n_k\left(\frac{\partial v_k^\|}{\partial x_j} - n_l n_j \frac{\partial v_k^\|}{\partial x_l}\right). \end{aligned} \tag{17}$$

Thus, the curved-space advection term can be written as

$$\begin{aligned}-\lambda[\mathbf{P}(\nabla \mathbf{v}^\|)\mathbf{P}]_{ij}v_j^\| = &-\lambda\Bigg[\left(\frac{\partial v_i^\|}{\partial x_j} - n_l n_j \frac{\partial v_i^\|}{\partial x_l}\right) \\ &- n_i n_k\left(\frac{\partial v_k^\|}{\partial x_j} - n_l n_j \frac{\partial v_k^\|}{\partial x_l}\right)\Bigg]v_j^\|. \end{aligned} \tag{18}$$

With these examples of tangent-plane calculus established, we now turn to the most tricky of the terms in the Toner-Tu equations, the surface-projected version of the neighbor coupling term, $D\nabla^2\mathbf{v}$. First, we recall that the Laplacian of a vector field in normal 3D Cartesian space is defined as $\nabla^2\mathbf{v} = \nabla \cdot (\nabla\mathbf{v})$. Further, we note that $\nabla\mathbf{v}$ is itself a tensor. As shown by Jankuhn *et al.* [34], the surface-projected version of the Laplacian term (see their Eq. (3.16) for the surface Navier-Stokes equations for the tangential velocity on a stationary surface), is therefore given by

$$\underbrace{\nabla^2\mathbf{v}}_{\text{flat Toner-Tu}} = \underbrace{\text{div}(\nabla\mathbf{v})}_{\text{in the plane}} \longrightarrow \underbrace{\mathbf{P}\text{div}_\Gamma(\mathbf{G}(\mathbf{v}^\|))}_{\text{curved surface}}. \tag{19}$$

To begin to unpack this expression, we note that $\mathbf{P}$ is the projection operator defined in Eq. (3) and that the tensor $\mathbf{G}$ is the surface velocity gradient,

$$\mathbf{G}(\mathbf{v}^\|) = \nabla_\Gamma \mathbf{v}^\|, \tag{20}$$

already presented in Eq. (17) and repeated here in component form as

$$G_{ij} = \left(\frac{\partial v_i^\|}{\partial x_j} - n_l n_j \frac{\partial v_i^\|}{\partial x_l}\right) - n_i n_k\left(\frac{\partial v_k^\|}{\partial x_j} - n_l n_j \frac{\partial v_k^\|}{\partial x_l}\right). \tag{21}$$

We note that Jankuhn *et al.* [34] use a symmetrized version of the velocity gradient since they are deriving the surface versions of the Navier-Stokes equations and are thus taking the divergence of a stress. For the Toner-Tu case of interest here, the psychological viscosity that comes from neighboring wildebeests in the herd comparing their velocities is equivalent to the divergence of the velocity gradient itself. We then invoke the definition of the surface-projected divergence of a tensor $\mathbf{G}$ presented in Appendix A 1, giving rise to a vector of the form

$$\text{div}_\Gamma (\mathbf{G})_l = \frac{\partial G_{lj}}{\partial x_j} - n_j n_k \frac{\partial G_{lj}}{\partial x_k}. \tag{22}$$

We also introduce the shorthand notation for this divergence as

$$\left(\frac{\partial G_{lj}}{\partial x_j}\right)_\Gamma = \frac{\partial G_{lj}}{\partial x_j} - n_j n_k \frac{\partial G_{lj}}{\partial x_k} \tag{23}$$

to simplify some of the complex expressions to follow. Finally, we can write the ith component of the surface version of the Toner-Tu term $D\nabla^2\mathbf{v}$ in indicial notation as

$$D[\mathbf{P}\text{div}_\Gamma(\mathbf{G}(\mathbf{v}^\|))]_i = DP_{il}\left(\frac{\partial G_{lj}}{\partial x_j} - n_j n_k \frac{\partial G_{lj}}{\partial x_k}\right). \tag{24}$$

We now assemble the results of Eqs. (8), (12), (18), and (24) to construct a complete curved-surface formulation of the

minimal Toner-Tu equations. Specifically, we have

$$\frac{\partial v_i^{\parallel}}{\partial t} = [\alpha(\rho - \rho_c) - \beta v_j^{\parallel} v_j^{\parallel}]v_i^{\parallel} - \sigma\left(\frac{\partial \rho}{\partial x_i} - n_i n_j \frac{\partial \rho}{\partial x_j}\right) + DP_{il}\left(\frac{\partial G_{lj}}{\partial x_j} - n_j n_k \frac{\partial G_{lj}}{\partial x_k}\right) - \lambda[\mathbf{P}(\nabla\mathbf{v}^{\parallel})\mathbf{P}]_{ij}v_j^{\parallel}, \tag{25}$$

which can be streamlined to the alternative form

$$\frac{\partial v_i^{\parallel}}{\partial t} = [\alpha(\rho - \rho_c) - \beta v_j^{\parallel} v_j^{\parallel}]v_i^{\parallel} - \sigma\left(\frac{\partial \rho}{\partial x_i}\right)_{\Gamma} + DP_{il}\left(\frac{\partial G_{lj}}{\partial x_j}\right)_{\Gamma}$$
$$- \lambda v_j^{\parallel}\left[\left(\frac{\partial v_i^{\parallel}}{\partial x_j}\right)_{\Gamma} - n_i n_k\left(\frac{\partial v_k^{\parallel}}{\partial x_j}\right)_{\Gamma}\right]. \tag{26}$$

Similarly, the curved-space formulation of the governing equation for density can be written as

$$\frac{\partial \rho}{\partial t} = -\frac{\partial(\rho v_i^{\parallel})}{\partial x_i} + n_i n_j \frac{\partial(\rho v_i^{\parallel})}{\partial x_j} = -\left(\frac{\partial \rho v_i^{\parallel}}{\partial x_i}\right)_{\Gamma}. \tag{27}$$

We now have a complete formulation of our minimal Toner-Tu equations for the general surface context, requiring only a description of the surface in the language of normal vectors. We next turn to the implementation of this general surface formulation in a fashion consistent with finite element treatments on arbitrary surfaces.

III. FORMULATING THE SURFACE TONER-TU EQUATIONS FOR NUMERICAL IMPLEMENTATION

Numerically solving active matter equations on complex surfaces relevant to the living world presents a practical challenge. It precludes the differential geometric formalism used to describe parameterized surfaces, which would lead to equations featuring covariant derivatives [e.g., Eq. (81) in Sec. V B]. In the finite element setting, the surface of interest is represented by a collection of nodes and corresponding surface normals, as shown in Fig. 2. Using these surface normals, we perform surface projections of the full three-space derivatives following Jankuhn *et al.* [34]. In the previous section, we presented a formal statement in Eq. (25) for handling the minimal Toner-Tu model on an arbitrary surface, using these surface normals. We now recast those curved-surface equations once more, in a fashion consonant with a FEM solver. To formulate the equations in an expression convenient for the FEM, we aim to rewrite the Toner-Tu equations in the form

$$\frac{\partial \mathbf{v}}{\partial t} + \nabla \cdot \mathbf{J}^{(v)} = \mathbf{f}^{(v)}. \tag{28}$$

Here, the whole formulation comes down to the definitions of $\mathbf{J}^{(v)}$ and $\mathbf{f}^{(v)}$. The fluxlike quantity $\mathbf{J}^{(v)}$ is a 3×3 matrix defined such that

$$\nabla \cdot \mathbf{J}^{(v)} = \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3}\right)\begin{bmatrix} J_{11} & J_{12} & J_{13} \\ J_{21} & J_{22} & J_{23} \\ J_{31} & J_{32} & J_{33} \end{bmatrix}, \tag{29}$$

recalling that the divergence of a second-rank tensor is a vector. In indicial notation, this can be written as

$$(\nabla \cdot \mathbf{J}^{(v)})_i = \frac{\partial J_{ji}^{(v)}}{\partial x_j}. \tag{30}$$

We need to define the components of the tensor $\mathbf{J}^{(v)}$ such that they yield the correct Toner-Tu terms. As we will see below, $\mathbf{J}^{(v)}$ is not symmetric despite its superficial resemblance to a stress tensor. To set notation and make sure that the strategy is clear, we begin by demonstrating how to implement the flat-space version of the Toner-Tu equations in a form consistent with Eq. (2). If we define the force term $\mathbf{f}^{(v)}$ as

$$f_i^{(v)} = [\alpha(\rho - \rho_c) - \beta v_j v_j]v_i - \sigma\frac{\partial \rho}{\partial x_i} - \lambda v_j\frac{\partial v_i}{\partial x_j}, \tag{31}$$

then the remaining terms are captured if we define the tensor $\mathbf{J}^{(v)}$ as

$$\mathbf{J}^{(v)} = \begin{bmatrix} -D\frac{\partial v_1}{\partial x_1} & -D\frac{\partial v_2}{\partial x_1} & -D\frac{\partial v_3}{\partial x_1} \\ -D\frac{\partial v_1}{\partial x_2} & -D\frac{\partial v_2}{\partial x_2} & -D\frac{\partial v_3}{\partial x_2} \\ -D\frac{\partial v_1}{\partial x_3} & -D\frac{\partial v_2}{\partial x_3} & -D\frac{\partial v_3}{\partial x_3} \end{bmatrix}. \tag{32}$$

Considering the one-component of velocity, we see that

$$(\nabla \cdot \mathbf{J}^{(v)})_1 = \frac{\partial J_{j1}}{\partial x_j} = \frac{\partial J_{11}}{\partial x_1} + \frac{\partial J_{21}}{\partial x_2} + \frac{\partial J_{31}}{\partial x_3}. \tag{33}$$

Plugging in the components of $\mathbf{J}^{(v)}$ gives the result

$$(\nabla \cdot \mathbf{J}^{(v)})_1 = \frac{\partial}{\partial x_1}\left(-D\frac{\partial v_1}{\partial x_1}\right) + \frac{\partial}{\partial x_2}\left(-D\frac{\partial v_1}{\partial x_2}\right) + \frac{\partial}{\partial x_3}\left(-D\frac{\partial v_1}{\partial x_3}\right). \tag{34}$$

Evaluating these derivatives leads to the result

$$(\nabla \cdot \mathbf{J}^{(v)})_1 = -D\frac{\partial^2 v_1}{\partial x_1^2} - D\frac{\partial^2 v_1}{\partial x_2^2} - D\frac{\partial^2 v_1}{\partial x_3^2}, \tag{35}$$

as expected from the original Toner-Tu equations. Combining the contributions from Eqs. (31) and (32) in the form $\partial\mathbf{v}/\partial t + \nabla \cdot \mathbf{J}^{(v)} = \mathbf{f}^{(v)}$, we recover Eq. (2) precisely, as we set out to do. By defining the quantities $\mathbf{J}^{(v)}$ and $\mathbf{f}^{(v)}$, we have successfully reframed the flat-space Toner-Tu equations in a format that will be conveniently implemented in the finite element setting. We now need to tackle the more demanding formulation for an arbitrary curved surface.

Abstractly, our finite element version of the curved-space Toner-Tu equations is written as

$$\frac{\partial \mathbf{v}}{\partial t} = -\operatorname{div}_{\Gamma} \mathbf{J}^{(v)} + \mathbf{f}^{(v)}, \tag{36}$$

where $\operatorname{div}_{\Gamma}$ is the surface projected version of the divergence introduced in Eq. (22). We now introduce the definition

$$J_{ji}^{(v)} = -DP_{il}G_{lj}, \tag{37}$$

where we recall $\mathbf{P}$ from Eq. (3) and $\mathbf{G}$ from Eq. (20). Given this definition, we can attempt to write our Toner-Tu equations as

$$\frac{\partial v_i}{\partial t} = -D\left(\frac{\partial(-P_{il}G_{lj})}{\partial x_j}\right)_{\Gamma} + f_i^{(v)}, \tag{38}$$

where we use the shorthand notation introduced in Eq. (23). This can be rewritten as

$$\frac{\partial v_i}{\partial t} = \underbrace{D\, P_{il}\left(\frac{\partial G_{lj}}{\partial x_j}\right)_\Gamma}_{\mathbf{P}\mathrm{div}_\Gamma(\nabla_\Gamma \mathbf{v}^\|)} + \underbrace{D\left(\frac{\partial P_{il}}{\partial x_j}\right)_\Gamma G_{lj}}_{\text{unwanted term}} + f_i^{(v)}. \quad (39)$$

Unfortunately, as written, these equations contain an extra term that is not present in the Toner-Tu formulation. To remove this unwanted extra term, we must introduce a new term in force $\mathbf{f}^{(v)}$ that subtracts the unwanted term:

$$f_i^{\mathrm{fict}} = -D\left(\frac{\partial P_{il}}{\partial x_j}\right)_\Gamma G_{lj}. \quad (40)$$

We view this as a mathematical trick that allows us to use a formalism convenient for FEM analysis and have not sought to find a physical interpretation of this force in the way that ideas such as the Coriolis force arises in mechanics. In this case, our *real* Toner-Tu equations can be written as

$$\frac{\partial v_i}{\partial t} = \underbrace{D\, P_{il}\left(\frac{\partial G_{lj}}{\partial x_j}\right)_\Gamma}_{\mathbf{P}\mathrm{div}_\Gamma(\nabla_\Gamma \mathbf{v}^\|)} + \underbrace{D\left(\frac{\partial P_{il}}{\partial x_j}\right)_\Gamma G_{lj}}_{\text{unwanted term}} + f_i^{(v)} + f_i^{\mathrm{fict}}. \quad (41)$$

We now have precisely the equations we want, cast in the form we will use in the finite element setting.

As we saw above, the force term $\mathbf{f}^{(v)} + \mathbf{f}^{\mathrm{fict}}$ needs to account for those terms that are not present in $\nabla \cdot \mathbf{J}^{(v)}$ and to subtract terms that are present in $\nabla \cdot \mathbf{J}^{(v)}$ but unwanted. To that end, the one-component of $\mathbf{f}^{(v)} + \mathbf{f}^{\mathrm{fict}}$ takes the form

$$f_1^{(v)} + f_1^{\mathrm{fict}} = (\alpha(\rho - \rho_c) - \beta((v_1^\|)^2 + (v_2^\|)^2 + (v_3^\|)^2))v_1^\| - \sigma\left(\frac{\partial \rho}{\partial x_1}\right)_\Gamma - \lambda\left(v_1^\|\frac{\partial v_1^\|}{\partial x_1} + v_2^\|\frac{\partial v_1^\|}{\partial x_2} + v_3^\|\frac{\partial v_1^\|}{\partial x_3} - v_1^\| n_1 n_1 \frac{\partial v_1^\|}{\partial x_1}\right.$$
$$- v_1^\| n_1 n_2 \frac{\partial v_2^\|}{\partial x_1} - v_1^\| n_1 n_3 \frac{\partial v_3^\|}{\partial x_1} - v_2^\| n_1 n_1 \frac{\partial v_1^\|}{\partial x_2} - v_2^\| n_1 n_2 \frac{\partial v_2^\|}{\partial x_2} - v_2^\| n_1 n_3 \frac{\partial v_3^\|}{\partial x_2} - v_3^\| n_1 n_1 \frac{\partial v_1^\|}{\partial x_3} - v_3^\| n_1 n_2 \frac{\partial v_2^\|}{\partial x_3}$$
$$- v_3^\| n_1 n_3 \frac{\partial v_3^\|}{\partial x_3}\right) - D\left(\left(\frac{\partial P_{11}}{\partial x_1}\right)_\Gamma G_{11} + \left(\frac{\partial P_{11}}{\partial x_2}\right)_\Gamma G_{12} + \left(\frac{\partial P_{11}}{\partial x_3}\right)_\Gamma G_{13}\right) - D\left(\left(\frac{\partial P_{12}}{\partial x_1}\right)_\Gamma G_{21} + \left(\frac{\partial P_{12}}{\partial x_2}\right)_\Gamma G_{22}\right.$$
$$\left. + \left(\frac{\partial P_{12}}{\partial x_3}\right)_\Gamma G_{23}\right) - D\left(\left(\frac{\partial P_{13}}{\partial x_1}\right)_\Gamma G_{31} + \left(\frac{\partial P_{13}}{\partial x_2}\right)_\Gamma G_{32} + \left(\frac{\partial P_{13}}{\partial x_3}\right)_\Gamma G_{33}\right). \quad (42)$$

Note that the α and β terms capture the preferred speed contribution, the σ term captures the pressure, the terms multiplied by λ capture the advection contribution to the Toner-Tu equations, and the final terms involving $\mathbf{P}$ and $\mathbf{G}$ subtract off the fictitious force. We can repeat a similar analysis for the two- and three-components of the force as

$$f_2^{(v)} + f_2^{\mathrm{fict}} = (\alpha(\rho - \rho_c) - \beta((v_1^\|)^2 + (v_2^\|)^2 + (v_3^\|)^2))v_2^\| - \sigma\left(\frac{\partial \rho}{\partial x_2}\right)_\Gamma - \lambda\left(v_1^\|\frac{\partial v_2^\|}{\partial x_1} + v_2^\|\frac{\partial v_2^\|}{\partial x_2} + v_3^\|\frac{\partial v_2^\|}{\partial x_3} - v_1^\| n_2 n_1 \frac{\partial v_1^\|}{\partial x_1}\right.$$
$$- v_1^\| n_2 n_2 \frac{\partial v_2^\|}{\partial x_1} - v_1^\| n_2 n_3 \frac{\partial v_3^\|}{\partial x_1} - v_2^\| n_2 n_1 \frac{\partial v_1^\|}{\partial x_2} - v_2^\| n_2 n_2 \frac{\partial v_2^\|}{\partial x_2} - v_2^\| n_2 n_3 \frac{\partial v_3^\|}{\partial x_2} - v_3^\| n_2 n_1 \frac{\partial v_1^\|}{\partial x_3} - v_3^\| n_2 n_2 \frac{\partial v_2^\|}{\partial x_3}$$
$$- v_3^\| n_2 n_3 \frac{\partial v_3^\|}{\partial x_3}\right) - D\left(\left(\frac{\partial P_{21}}{\partial x_1}\right)_\Gamma G_{11} + \left(\frac{\partial P_{21}}{\partial x_2}\right)_\Gamma G_{12} + \left(\frac{\partial P_{21}}{\partial x_3}\right)_\Gamma G_{13}\right) - D\left(\left(\frac{\partial P_{22}}{\partial x_1}\right)_\Gamma G_{21} + \left(\frac{\partial P_{22}}{\partial x_2}\right)_\Gamma G_{22}\right.$$
$$\left. + \left(\frac{\partial P_{22}}{\partial x_3}\right)_\Gamma G_{23}\right) - D\left(\left(\frac{\partial P_{23}}{\partial x_1}\right)_\Gamma G_{31} + \left(\frac{\partial P_{23}}{\partial x_2}\right)_\Gamma G_{32} + \left(\frac{\partial P_{23}}{\partial x_3}\right)_\Gamma G_{33}\right) \quad (43)$$

and

$$f_3^{(v)} + f_3^{\mathrm{fict}} = (\alpha(\rho - \rho_c) - \beta((v_1^\|)^2 + (v_2^\|)^2 + (v_3^\|)^2))v_3^\| - \sigma\left(\frac{\partial \rho}{\partial x_3}\right)_\Gamma - \lambda\left(v_1^\|\frac{\partial v_3^\|}{\partial x_1} + v_2^\|\frac{\partial v_3^\|}{\partial x_2} + v_3^\|\frac{\partial v_3^\|}{\partial x_3} - v_1^\| n_3 n_1 \frac{\partial v_1^\|}{\partial x_1}\right.$$
$$- v_1^\| n_3 n_2 \frac{\partial v_2^\|}{\partial x_1} - v_1^\| n_3 n_3 \frac{\partial v_3^\|}{\partial x_1} - v_2^\| n_3 n_1 \frac{\partial v_1^\|}{\partial x_2} - v_2^\| n_3 n_2 \frac{\partial v_2^\|}{\partial x_2} - v_2^\| n_3 n_3 \frac{\partial v_3^\|}{\partial x_2} - v_3^\| n_3 n_1 \frac{\partial v_1^\|}{\partial x_3} - v_3^\| n_3 n_2 \frac{\partial v_2^\|}{\partial x_3}$$
$$- v_3^\| n_3 n_3 \frac{\partial v_3^\|}{\partial x_3}\right) - D\left(\left(\frac{\partial P_{31}}{\partial x_1}\right)_\Gamma G_{11} + \left(\frac{\partial P_{31}}{\partial x_2}\right)_\Gamma G_{12} + \left(\frac{\partial P_{31}}{\partial x_3}\right)_\Gamma G_{13}\right) - D\left(\left(\frac{\partial P_{32}}{\partial x_1}\right)_\Gamma G_{21} + \left(\frac{\partial P_{32}}{\partial x_2}\right)_\Gamma G_{22}\right.$$
$$\left. + \left(\frac{\partial P_{32}}{\partial x_3}\right)_\Gamma G_{23}\right) - D\left(\left(\frac{\partial P_{33}}{\partial x_1}\right)_\Gamma G_{31} + \left(\frac{\partial P_{33}}{\partial x_2}\right)_\Gamma G_{32} + \left(\frac{\partial P_{33}}{\partial x_3}\right)_\Gamma G_{33}\right). \quad (44)$$

By setting $\partial \mathbf{v}/\partial t$ equal to the sum of $-\mathrm{div}_\Gamma \mathbf{J}^{(v)}$ and the $\mathbf{f}^{(v)} + \mathbf{f}^{\mathrm{fict}}$ terms, we have fully reproduced the curved-space version of the Toner-Tu equations.

We express the continuity equation in similar form,

$$\frac{\partial \rho}{\partial t} = -\operatorname{div}_\Gamma \mathbf{J}^{(\rho)} + f^{(\rho)}, \tag{45}$$

where $\mathbf{J}^{(\rho)}$ is now a vector and the scalar $f^{(\rho)}$ can be thought of as a source term. We set $\operatorname{div}_\Gamma \mathbf{J}^{(\rho)}$ equal to zero and thus define

$$\begin{aligned}
f^{(\rho)} &= -\operatorname{div}_\Gamma(\rho \mathbf{v}^\parallel) \\
&= -\left(\frac{\partial(\rho v_l^\parallel)}{\partial x_l} - n_l n_k \frac{\partial(\rho v_l^\parallel)}{\partial x_k} \right) \\
&= -\left(\frac{\partial \rho v_1^\parallel}{\partial x_1} \right)_\Gamma - \left(\frac{\partial \rho v_2^\parallel}{\partial x_2} \right)_\Gamma - \left(\frac{\partial \rho v_3^\parallel}{\partial x_3} \right)_\Gamma. \tag{46}
\end{aligned}$$

Altogether, Eqs. (37), (42)–(44), and (46) comprise a complete curved-surface FEM implementation of the minimal Toner-Tu equations and make possible the numerical results presented in the remainder of this paper. We also refer the interested reader to Appendix A 2 for details on our implementation in the specific commercial finite element package COMSOL MULTIPHYSICS. Our code, files, and a tutorial on their use are available in Ref. [48].

IV. PARAMETERS AND DIMENSIONLESS RATIOS FOR THE THEORY

A. Parameter choices for wildebeest herds

To put our finite element formulation into numerical action, we must of course adopt specific values for the parameters ρ_0, ρ_c, α, β, σ, D, and λ that appear in our governing partial differential equations. In this paper, we focus on the macroscopic length scale of animal herds, although we consider the microscopic activity of cytoskeletal proteins elsewhere [27]. While our goal here is to explore the phenomenology of Toner-Tu flocks on curved surfaces in a general way, not to claim an understanding of specific animal behavior, our parameter choices are loosely inspired by migrating wildebeest herds as seen in Fig. 3. By inspecting aerial photographs of wildebeest herds, we estimated an average wildebeest density of

$$\rho_0 = 0.25 \, \text{m}^{-2}. \tag{47}$$

For the critical density above which coordinated herding behavior occurs, we make the estimate

$$\rho_c = 0.05 \, \text{m}^{-2}, \tag{48}$$

based roughly on the observation that for densities much higher than this, the photographed wildebeest have an organized herd structure. We note that it would be very interesting to carefully measure these parameters in the context of herds of wildebeests with a Toner-Tu framework in mind.

The remainder of the parameters are determined in the spirit of exploring how the different terms compete to alter the density and velocity fields. Thus, parameters are chosen to make all terms comparable in magnitude. Here we provide the actual values used in our finite element calculations, fully cognizant that our parameter choices are at best approximate. We begin by estimating the magnitude of $\partial v / \partial t$. We picture

wildebeests circling a small hill in the landscape, moving with a characteristic speed of 1 m/s and taking 20 s to change directions completely. This scenario implies

$$\frac{\partial v}{\partial t} \approx \frac{1 \, \text{m/s} - (-1 \, \text{m/s})}{20 \, \text{s}} \approx 0.1 \, \text{m/s}^2. \tag{49}$$

Using the strategy of balancing the magnitudes of the different terms, we estimate the coefficient for the preferred speed terms by considering that

$$\begin{aligned}
0.1 \, \text{m/s}^2 &= \alpha(\rho_0 - \rho_c)v \\
&\approx \alpha \times (0.25 \, \text{m}^{-2} - 0.05 \, \text{m}^{-2}) \times 1 \, \text{m/s}, \tag{50}
\end{aligned}$$

which leads us to adopt

$$\alpha = 0.5 \, \text{m}^2/\text{s}. \tag{51}$$

For the parameter β, we choose a value that sets the correct mean field speed, v_{pref}, for wildebeests. Thus, we adopt

$$\beta = \frac{\alpha(\rho_0 - \rho_c)}{v_{\text{pref}}^2} \approx \frac{(0.5 \, \text{m}^2/\text{s})(0.2 \, \text{m}^{-2})}{(1 \, \text{m/s})^2} = 0.1 \, \text{s}/\text{m}^2. \tag{52}$$

We can also check independently that the magnitude of the β term is of order 0.1 m/s^2. Indeed,

$$\beta v^3 \approx 0.1 \frac{\text{s}}{\text{m}^2} \times (1 \, \text{m/s})^3 = 0.1 \, \text{m/s}^2. \tag{53}$$

To find the parameter σ, the coefficient of the pressure term, we consider the gradients in density seen at the edge of a wildebeest herd and estimate that ρ drops from $\rho_0 = 0.25 \, \text{m}^{-2}$ to $0 \, \text{m}^{-2}$ over 25 m. Using these numbers implies

$$0.1 \, \text{m/s}^2 = \sigma \frac{\partial \rho}{\partial x} \approx \sigma \frac{0.25 \, \text{m}^{-2}}{25 \, \text{m}}, \tag{54}$$

which leads us to adopt

$$\sigma = 10 \, \text{m}^4/\text{s}^2. \tag{55}$$

We can now apply this thinking to make an estimate of the neighbor coupling coefficient D by using the equality

$$0.1 \, \text{m/s}^2 = D \frac{\partial^2 v}{\partial x^2} \approx D \frac{1 \, \text{m/s}}{(10 \, \text{m})^2}, \tag{56}$$

which leads to the coefficient choice

$$D = 10 \, \text{m}^2/\text{s}. \tag{57}$$

To estimate the magnitude of λ, we imagine that the wildebeest will come to a full stop from a speed of 1 m/s over a distance of roughly 10 m, permitting us to make the correspondence

$$0.1 \, \text{m/s}^2 = \lambda v \frac{\partial v}{\partial x} \approx \lambda \times 1 \, \text{m/s} \times \frac{1 \, \text{m/s}}{10 \, \text{m}}, \tag{58}$$

which implies that

$$\lambda = 1. \tag{59}$$

This estimated set of parameters provides us with a complete description of the Toner-Tu herd in the minimal model we seek to explore. Throughout the remainder of the paper, we consistently use these parameter values. In later sections, in some cases we will go beyond this idealized parameter set to perform parameter sweeps that permit different regimes of behavior to emerge. Unless noted otherwise, time-dependent FEM simulations were initialized with a uniform density field
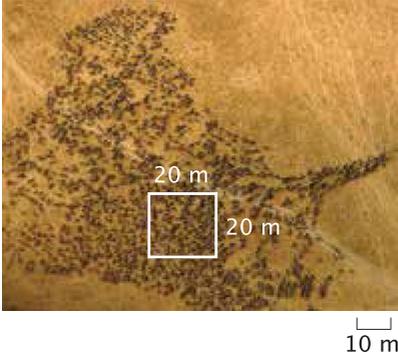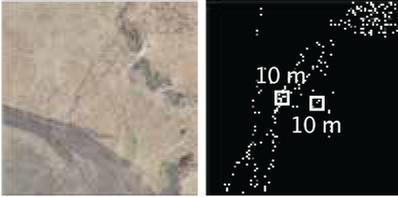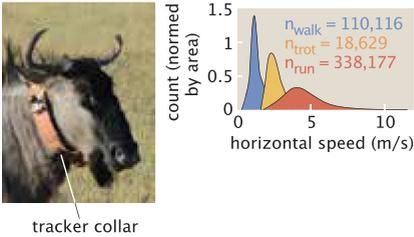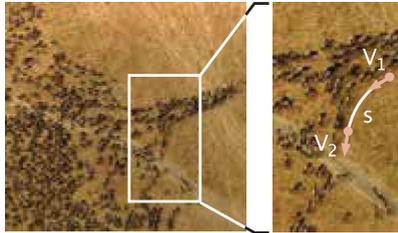
PROPERTY	DATA	ESTIMATED VALUE

mean density, ρ_0

$$\rho_0 \approx \frac{100 \text{ wildebeest}}{400 \text{ m}^2} = \frac{1}{4}\frac{1}{\text{m}^2}$$

critical density, ρ_c

$$\frac{3 \text{ wildebeest}}{100 \text{ m}^2} < \rho_c < \frac{11 \text{ wildebeest}}{100 \text{ m}^2}$$

$$\rho_c \approx \frac{1}{20}\frac{1}{\text{m}^2}$$

speed, v

$n_{\text{walk}} = 110{,}116$
$n_{\text{trot}} = 18{,}629$
$n_{\text{run}} = 338{,}177$

$$\mathbf{v} \approx 1\,\frac{\text{m}}{\text{s}}$$

acceleration, dv/dt

$$\frac{d\mathbf{v}}{dt} = \frac{\mathbf{v}_1 - \mathbf{v}_2}{s/|\mathbf{v}|} \approx 0.1\,\frac{\text{m}}{\text{s}^2}$$

FIG. 3. Estimates for key parameter values from tracking data and images of wildebeest herds. Mean density is estimated in the middle of a herd, while the critical density is estimated as in between the density of a region that does not show flocking structure (right white box) and the density of a region that does (left white box). For this paper, we use a wildebeest walking speed of 1 m/s, inferred from the plot of tracking data adapted from Ref. [49]. The scale of wildebeest acceleration is estimated from images of a turning herd, where a change in the vector **v** of roughly 1 m/s occurs over a timescale s/|**v**|, or roughly 10 s. Aerial photos from drones (mean density and acceleration image) and satellites (critical density image) are from Ref. [50].

$\rho = \rho_0$ and a disordered velocity field, with velocity orientations drawn randomly from a uniform distribution of angles between 0 and 2π and with velocity magnitude $v(0) = 1$ m/s, our estimated characteristic wildebeest speed. We note again that these parameters were chosen to highlight competition between the different terms in the dynamics; they are far from the final word for describing real wildebeest herds.

B. Dimensionless representation of the theory

We next recast the Toner-Tu equations in dimensionless form, rendering the meaning and magnitude of the terms more transparent. For the positional coordinate, we take $x^* = x/L$,

where L is a characteristic length scale in the problem. In this case, we conceptualize L as the length scale of the wildebeest herd. We note that in the problems considered here, the scale of the herd is the same as the scale of the surface landscape on which the herd moves. Similarly, we use a characteristic velocity scale U to define the dimensionless velocity as $v^* = v/U$. We note that U is conceptually related to v_{pref}, but $v_{\text{pref}} = \sqrt{\alpha(\rho - \rho_c)/\beta}$ features a hidden dependence on the density that would complicate our rescaling. Similarly, we can define a dimensionless density $\rho^* = \rho/\rho_c$, using the critical density ρ_c as our scaling variable. Lastly, in light of the definitions above, we can determine a timescale L/U, the

time it takes for a given wildebeest to cross the entire surface landscape. Thus, we define $t^* = t/(L/U)$.

Using the various definitions given above, we can now rewrite the Toner-Tu equations using the dimensionless versions of t, x, ρ, and $\mathbf{v}$ as

$$\frac{U^2}{L}\frac{\partial v_i^*}{\partial t^*} = \alpha\rho_c U(\rho^* - 1)v_i^* - \beta U^3|\mathbf{v}^*|^2 v_i^* - \frac{\sigma\rho_c}{L}\frac{\partial\rho^*}{\partial x_i^*}$$
$$+ \frac{DU}{L^2}\nabla_*^2 v_i^* - \frac{\lambda U^2}{L}v_j^*\frac{\partial v_i^*}{\partial x_j^*}. \tag{60}$$

Dividing everything by U^2/L results in five dimensionless parameters whose magnitudes provide a sense of the relative contributions of the different terms, within the full dynamical equations of the form

$$\frac{\partial v_i^*}{\partial t^*} = \frac{\alpha\rho_c L}{U}(\rho^* - 1)v_i^* - \beta LU|\mathbf{v}^*|^2 v_i^* - \frac{\sigma\rho_c}{U^2}\frac{\partial\rho^*}{\partial x_i^*}$$
$$+ \frac{D}{UL}\nabla_*^2 v_i^* - \lambda v_j^*\frac{\partial v_i^*}{\partial x_j^*}. \tag{61}$$

Using definitions of dimensionless variables given above, we find that the generalized continuity equation takes the form

$$\frac{\partial\rho^*}{\partial t^*} = -\frac{\partial(\rho^* v_i^*)}{\partial x_i^*}. \tag{62}$$

Next, we briefly turn to interpreting the dimensionless ratios that appear when casting the minimal Toner-Tu theory in dimensionless form as exhibited in Eq. (61). The two components of the preferred speed term have dimensionless parameters given by

$$\text{preferred speed } \alpha \text{ term} = \frac{\alpha\rho_c \text{L}}{\text{U}} = \frac{\text{L/U}}{1/(\alpha\rho_c)} = \frac{\text{time for wildebeest to cross herd}}{\text{time for speed to increase to } v_{\text{pref}}} \tag{63}$$

and

$$\text{preferred speed } \beta \text{ term} = \beta LU = \frac{L/U}{1/(\beta U^2)} = \frac{\text{time for wildebeest to cross herd}}{\text{time for speed to decrease to } v_{\text{pref}}}. \tag{64}$$

Each of these terms provides intuition about how quickly the wildebeest will return to their steady-state speed given some perturbation that disturbs them from that value. Recall that L/U (the time for wildebeest to cross herd) is roughly the time it takes for a wildebeest to move a distance equal to the size of the herd or, equivalently, to move across the surface landscape. In other words, L/U is the time for density advection across the herd. The pressure term can be rewritten as

$$\text{pressure term} = \frac{\rho_c}{U^2/\sigma} = \frac{\text{critical density}}{\text{typical density excursion away from } \rho_0}. \tag{65}$$

Increasing σ decreases density variance; in other words, the densities that emerge are within a more narrow range around the mean density ρ_0. The neighbor coupling term that carries out democratic velocity smoothing has the dimensionless prefactor

$$\text{neighbor coupling term} = \frac{D}{UL} = \frac{L/U}{L^2/D} = \frac{\text{time for wildebeest to cross herd}}{\text{time for velocity to diffuse across herd}}, \tag{66}$$

analogous to the Péclet number. The last term can be written as

$$\text{advection term} = \frac{\lambda L/U}{L/U} = \frac{\text{time for velocity advection across herd}}{\text{time for wildebeest to cross herd (density advection)}}. \tag{67}$$

These dimensionless ratios give a sense of how large a contribution a given term will make to the incremental update to $\mathbf{v}(t)$. In a sense, their values capture the relative importance of each term, serving roles analogous to the Reynolds number in thinking about the Navier-Stokes equations and the Péclet number in the context of coupled diffusion-advection problems.

V. SOLVING TONER-TU ON HIGHLY SYMMETRIC GEOMETRIES

In this section, we use the finite element implementation of the curved-space minimal Toner-Tu model to illustrate herding in symmetric geometries in anticipation of the fully general case. Our goals are twofold: to explore the behavior of Toner-Tu herds in classic geometries but also, importantly, to validate our finite element implementation by comparison to analytic solutions. This validation is particularly important in the context of systems as complex as the partial differential equations considered here, which are already subtle in flat space and more demanding yet in their curved-space form. We begin with the planar geometry of a two-dimensional (2D) channel before moving to a 2D channel with an embedded scattering obstacle. Then, we turn to collective herding motions on cylindrical and spherical surfaces, as those two cases do admit analytic solutions that can be directly compared to the curved-space numerical results.

A. Solving the Toner-Tu equations in the plane

One of the classic case studies for traditional fluid mechanics is channel flow, a problem dating back at least to the 19th century. Here, we consider several versions of the channel flow problem in the context of active herding agents, with
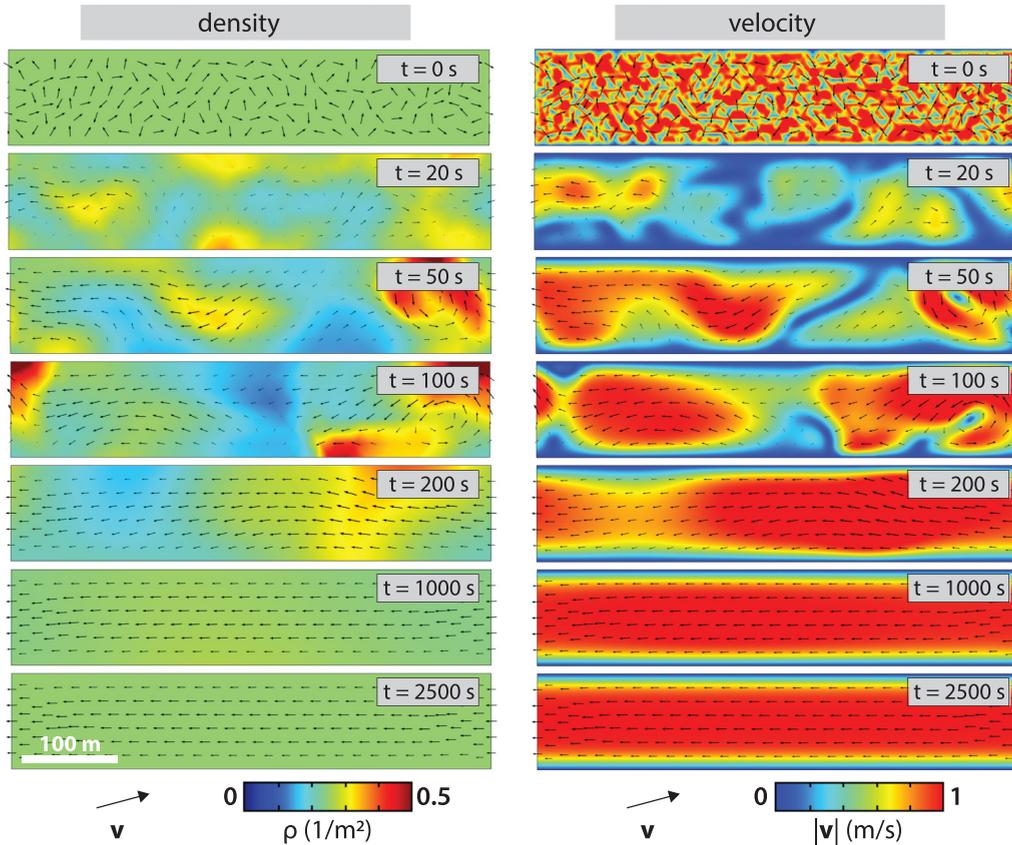
FIG. 4. Toner-Tu herding in a 2D channel with periodic boundary conditions at its ends. The left panel shows the time evolution of the density field (colors) with the velocities superimposed as arrows. The right panel reports on the same simulation but uses a color map to show the speed. In the long-time limit, the density is uniform and the speed is constant throughout the channel except for a boundary layer near the walls. See also video 1 (SM). Here and in all subsequent figures, parameters used are those inferred in Sec. IV unless explicitly stated otherwise.

the recognition that herding animals are sometimes faced with passing through canyons or gorges. In the Disney movie *The Lion King*, Hollywood filmmakers gave their own versions of herding behavior in a scene in which the young lion Simba is trapped in a gorge of stampeding wildebeests as a result of the machinations of his evil uncle Scar. In this section, in a playful frame of mind, we use the minimal Toner-Tu theory in conjunction with the FEM to explore a herd of stampeding wildebeest in a gorge. We first consider a 2D channel, yielding the Toner-Tu version of pipe flow. This problem has been explored experimentally in work on annular channels [12,17,18,51]. We follow this with an analysis of an obstacle in the middle of the channel [20] to show the generality and flexibility of the finite element implementation in action.

Solving the problem of channel flow is a rite of passage in the study of the hydrodynamics of Newtonian fluids. However, for active agents, this simple geometry already reveals interesting phenomenology, as shown in Fig. 4 and video 1 of the Supplemental Material (SM) [52]. As illustrated in Fig. 4, we begin with a uniform density and a randomly oriented velocity field in a 2D channel with periodic boundary conditions at its ends. We then use our finite element implementation of the minimal Toner-Tu theory to explore the transient and steady-state dynamics. Despite starting from an initially disordered velocity field, in the long-time limit the system finds

a steady state with uniform density ρ_0 and speeds that are consistent with the preferred speed term, here equivalent to $|\mathbf{v}| = \sqrt{\alpha(\rho_0 - \rho_c)/\beta}$. Near the edges of the channel, because of the no-slip boundary condition, there is a boundary layer interpolating between the optimal speed at the middle of the channel and the zero velocity at the walls. We note that it is not at all clear that no-slip boundary conditions are the most reasonable biological choice either for animal herds in a gorge or for cytoskeletal filaments in channels, but solving this problem provides a case study for exploring the intuition behind the Toner-Tu theory and our numerical implementation of it.

Tuning the neighbor coupling coefficient D results in different steady-state velocity profiles, $v_{ss}(y)$, as shown in Fig. 5. For comparison to these FEM findings, we now consider several analytical calculations. In the steady state, the velocity field in the minimal Toner-Tu model for this channel geometry is described by

$$D\frac{d^2 v_x(y)}{dy^2} + \left(\alpha(\rho_0 - \rho_c) - \beta v_x^2(y)\right)v_x(y) = 0. \quad (68)$$

We can solve this equation by numerical integration with a custom-written MATLAB code using shooting methods. As shown in Fig. 5, the time-dependent FEM and the shooting method solutions to Eq. (68) give essentially indistinguishable
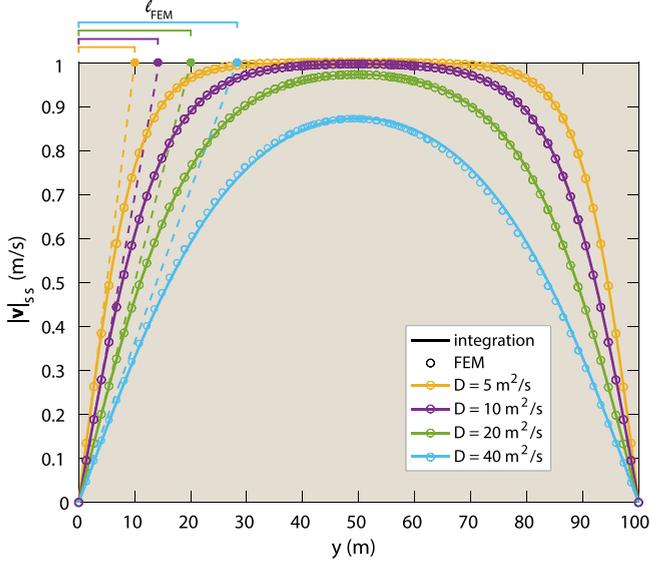
FIG. 5. Steady-state velocity for active agents during channel flow as a function of position y, across the channel. Solutions were obtained both by the time-dependent finite element method (circles) and by numerically integrating the steady-state Toner-Tu equations (solid lines) for the highly symmetric channel geometry. Different choices of the neighbor coupling coefficient D result in boundary layers of different thicknesses. The width ℓ_{FEM} of the boundary layer is estimated by computing the initial slope of the velocity versus position curves and extrapolating out (dashed lines) to the preferred speed ($v_{\text{pref}} = 1$ m/s), as shown in the top left of the figure. In Fig. 6, ℓ_{FEM} is compared to analytical calculations of the boundary layer thickness.

results. This correspondence provides a validation of the FEM approach.

For a deeper understanding of the channel problem, we focus on the velocity boundary layers clearly noticeable in Fig. 5. For sufficiently small values of the neighbor coupling coefficient D, the preferred speed term keeps the steady-state speed at $v_{\text{pref}} = 1$ m/s throughout the channel except for a narrow band near the walls, where the no-slip condition gives rise to a boundary layer. In Fig. 5, we show how the width of the boundary layer in the FEM solution, ℓ_{FEM}, can be estimated using the condition

$$\frac{dv_x(y)}{dy}\ell_{\text{FEM}}\bigg|_{y=0} = v_{\text{pref}}, \tag{69}$$

where $v_{\text{pref}} = \sqrt{\alpha(\rho_0 - \rho_c)/\beta}$ is the preferred speed. We explore the scaling of boundary layer thickness numerically for different values of the neighbor coupling coefficient D using the finite element implementation, and we report these numerically determined boundary layer widths, ℓ_{FEM}, in Fig. 6.

To interpret the length scale of the boundary layer and its scaling with D, we can perform an estimate by introducing the dimensionless variables

$$\bar{y} = \frac{y}{\sqrt{\frac{D}{\alpha(\rho_0 - \rho_c)}}} \tag{70}$$

and

$$\bar{v} = \frac{v_x}{\sqrt{\alpha \frac{(\rho_0 - \rho_c)}{\beta}}}. \tag{71}$$
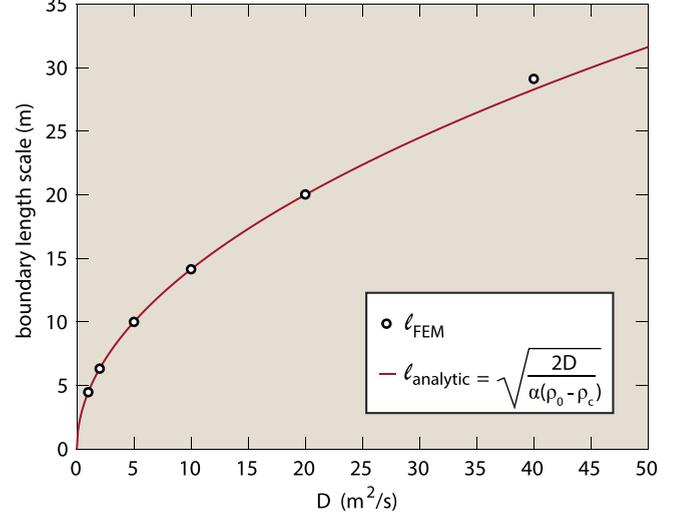

FIG. 6. Boundary layer thickness as a function of the neighbor coupling coefficient, D. A numerical result for the thickness of the boundary layer, ℓ_{FEM}, was estimated by computing dv/dy at the wall and using the condition $(dv/dy)\ell_{\text{FEM}} = v_{\text{pref}}$, where $v_{\text{pref}} = \sqrt{\alpha(\rho_0 - \rho_c)/\beta}$. An analytical result for the thickness of the boundary layer, ℓ_{analytic}, was calculated in Eqs. (74)–(77).

With these dimensionless variables in hand, we can rewrite the steady-state equation for velocity in the channel [Eq. (68)] in dimensionless form as

$$\frac{d^2\bar{v}}{d\bar{y}^2} + (1 - \bar{v}^2)\bar{v} = 0. \tag{72}$$

The act of rendering the equation in dimensionless form immediately presents us with the length scale

$$\ell = \sqrt{\frac{D}{\alpha(\rho_0 - \rho_c)}}, \tag{73}$$

which gives us a sense of the thickness of the boundary layer found by numerical methods in Fig. 5. This scaling relation can be understood more rigorously by considering the problem of the velocity $v_x(y)$ of active agents in a half space for $0 \leqslant y < \infty$. In this case, we have $v_x(0) = 0$ as a no-slip boundary condition and $v_x(\infty) = \sqrt{\alpha(\rho_0 - \rho_c)/\beta}$, implying that far from the wall, the speed achieves the Toner-Tu preferred speed value. In this case, Eq. (72) has the analytic solution $\bar{v}(\bar{y}) = \tanh(\bar{y}/\sqrt{2})$, which can be written in dimensionful form as

$$v_x(y) = \sqrt{\frac{\alpha(\rho_0 - \rho_c)}{\beta}}\tanh\sqrt{\frac{\alpha(\rho_0 - \rho_c)}{2D}}y. \tag{74}$$

For sufficiently small y, we can approximate this by Taylor expanding the tanh function to first order in y, resulting in

$$v_x(y) \approx \frac{\alpha(\rho_0 - \rho_c)}{\sqrt{2\beta D}}y, \tag{75}$$

which implies, in turn, that the thickness of the boundary layer can be estimated using

$$\frac{dv_x}{dy}\bigg|_{y=0}\ell_{\text{analytic}} = \frac{\alpha(\rho_0 - \rho_c)}{\sqrt{2\beta D}}\ell_{\text{analytic}} = \sqrt{\frac{\alpha(\rho_0 - \rho_c)}{\beta}}. \tag{76}$$
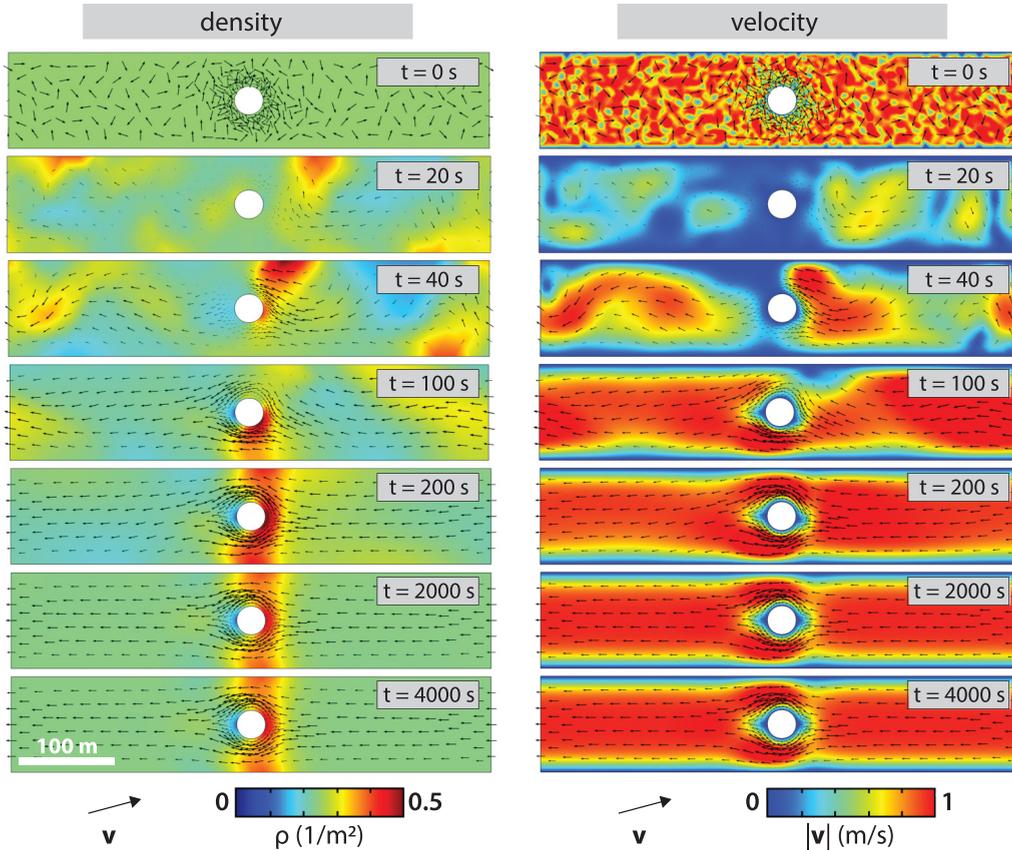
FIG. 7. Solution of the Toner-Tu equations for a herd moving through a channel with a small circular obstacle (15 m obstacle radius; 100 m channel width). The left panel shows the time evolution of the density field (colors) with the velocities superimposed as arrows, from an initial condition of uniform density and a disordered velocity field. The right panel uses a color map to show the speed. For this choice of obstacle size and neighbor coupling coefficient D, a steady-state flow is achieved in which the active agents can reach their preferred speed even in the passage between the obstacle and the wall. See also video 1 (SM).

Solving for the length scale of the boundary layer by projecting the straight line with the correct slope at $y = 0$ out to the saturation speed, as done in Fig. 5 to compute the analogous ℓ_{FEM}, leads to

$$\ell_{\text{analytic}} = \sqrt{\frac{2D}{\alpha(\rho_0 - \rho_c)}}, \tag{77}$$

with the factor of $\sqrt{2}$ arising naturally from the analytic solution for the half space. This analytical calculation for the boundary layer length ℓ_{analytic} jibes exactly with the numerical result ℓ_{FEM}, as seen in Fig. 6, until D reaches values large enough that the half-space assumption of $v_x(\infty) = \sqrt{\alpha(\rho_0 - \rho_c)/\beta}$ becomes invalid.

The problem of Toner-Tu channel flow for wildebeests becomes richer when we introduce a circular obstacle into the channel, as shown in Fig. 7 and video 1 (SM). In the classic example of this problem for a Newtonian fluid, the fluid motion will speed up near the obstacle because of the narrowing of the channel. However, in the case of a Toner-Tu collective, the preferred speed term attempts to maintain all agents at a fixed speed. Thus, if a large obstacle narrows the channel sufficiently, as shown in Fig. 8 and video 1 (SM), most of the herd will reflect off the obstacle rather than squeezing through at a higher velocity. Interestingly, this oscillatory direction reversal at an obstacle was observed in recent experimental and theoretical work, in which rectangular or triangular objects partially blocked the path of an active colloidal fluid confined in a ring-shaped track [20].

To better understand the origins of these reversals of motion, we contrast the Toner-Tu theory with our expectations for Newtonian fluids. Specifically, we note the crucial role of the preferred speed term, which penalizes any active agents with velocity magnitudes other than $v = \sqrt{\alpha(\rho - \rho_c)/\beta}$. We estimate the critical obstacle radius R for a channel of width L by imagining that when the thickness of the passage that the active agents can pass through (i.e., $L/2 - R$) is comparable to the boundary layer thickness, none of the active agents get to move at the optimal speed and as a result will reflect off the obstacle. That condition can be written as

$$\frac{L}{2} - R = 2\sqrt{\frac{D}{\alpha(\rho_0 - \rho_c)}}, \tag{78}$$

where the factor of 2 on the right side captures the idea that if the gap is wide enough, there will be two boundary layers, one at the obstacle and one at the wall. When the width of these *two* boundary layers adds up to the width of the gap, we estimate that reversals will occur. Note that, as written, Eq. (78) features the length scale ℓ that arose from recasting
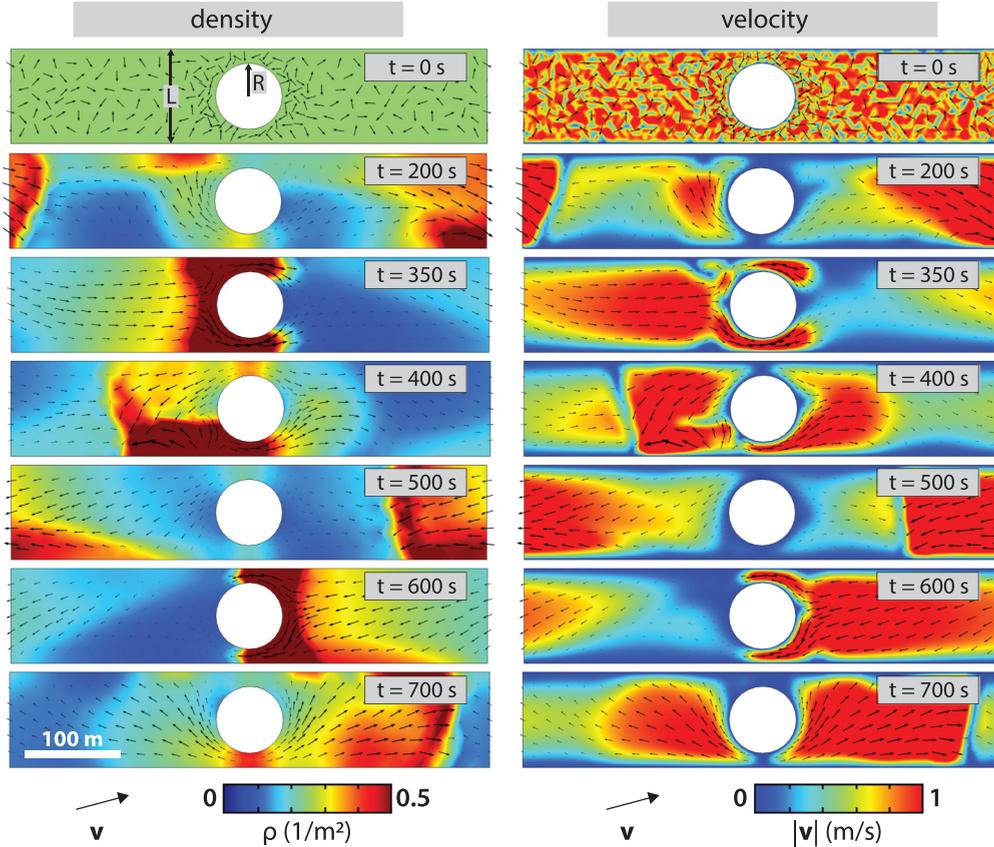
FIG. 8. Solution of the Toner-Tu equations for a herd moving through a channel with a large circular obstacle (30 m obstacle radius; 100 m channel width). The left panel shows the time evolution of the density field (colors) with the velocities superimposed as arrows. The right panel uses a color map to show the speed. For this choice of obstacle size and neighbor coupling coefficient D, no steady state motion is achieved. Most of the herd is reflected off the obstacle at each attempted passage, resulting in repeated reversals of direction. See also video 1 (SM).

the steady-state Toner-Tu equations in dimensionless form as defined in Eqs. (70) and (71), not the length scale ℓ_{analytic} from our analytic solution for the half space.

As seen in Fig. 9, we have explored the phenomenology of the motions in the channel with an obstacle by sweeping through the parameter space of obstacle size R and neighbor

FIG. 9. Phase diagram of herding behavior in a channel with an obstacle. Obstacle radius, R, and neighbor coupling coefficient, D, were systematically varied. Emergent herd dynamics were either unidirectional, reaching a steady-state as in Fig. 7 or displayed motion reversals at the obstacle indefinitely, as in Fig. 8. The transition between these two behaviors appears to occur when the width of the passage between obstacle and channel wall, $L/2 - R$, is comparable to twice the boundary layer length scale ℓ obtained by dimensional analysis in Eq. (73).

coupling coefficient D. Depending upon the choices of these parameters as already seen in Figs. 7 and 8, we find either steady unidirectional motion or dynamical reversals, which continue indefinitely. Interestingly, the boundary in this phase portrait is in very good accord with the estimate of Eq. (78). In light of the apparent success of these case studies on the minimal Toner-Tu model using a finite element implementation for planar geometries, we turn to the application of the minimal Toner-Tu model and its finite element implementation to curved surfaces.
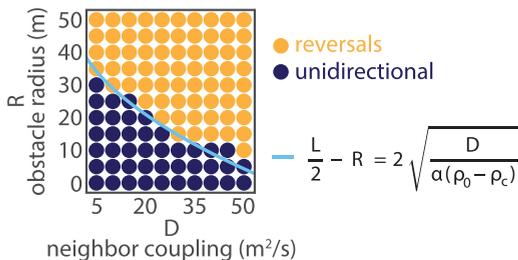
B. Formulating the Toner-Tu equations for parameterized surfaces

We now move to apply and test our general surface formulation (Sec. II) and finite element implementation (Sec. III) of the Toner-Tu equations on curved surfaces. We begin with two highly symmetric surfaces, the cylinder and the sphere, for which it is tractable to derive analytical solutions using intrinsic parametrizations. In the remainder of Sec. V, we will analytically determine dynamical Toner-Tu solutions for velocity on the cylinder and steady-state Toner-Tu solutions for density and velocity on the sphere. These analytical solutions are of great interest to us as test beds for our curved-space finite element formulation of the minimal Toner-Tu theory. We will compare our analytical calculations to results obtained

using the finite element approach, with the ultimate goal of validating its performance.

First, we must formulate the minimal flocking theory in the language of intrinsic differential geometry. We follow the analytical study by Shankar *et al.* [39] of Toner-Tu equations on highly symmetric curved surfaces, in which steady-state solutions for the sphere were considered. Following Shankar *et al.* [39], we need to write the fully covariant form of the dynamical Toner-Tu equations. First, the continuity equation can be written as

$$\partial_t \rho + \nabla_\mu(\rho v^\mu) = 0. \tag{79}$$

This formulation invokes the covariant derivative

$$\nabla_\mu v^\nu = \partial_\mu v^\nu + \Gamma^\nu_{\alpha\mu} v^\alpha, \tag{80}$$

where $\Gamma^\nu_{\alpha\mu}$ are the Christoffel symbols. All of this extra machinery and corresponding notation comes down to the subtlety of getting our derivatives right in the curved-space setting. Then, the dynamical equation for **v** itself is given by

$$\frac{\partial v^\mu}{\partial t} = [\alpha(\rho - \rho_c) - \beta g_{ab} v^a v^b] v^\mu - \sigma \nabla^\mu \rho + D\Delta v^\mu - \lambda v^\nu \nabla_\nu v^\mu, \tag{81}$$

where we introduce the symbol Δ to represent the intrinsic curved-space version of the Laplacian operator. We now put these equations into concrete play in the context of cylinders and spheres, with the ambition of directly comparing analytical results and finite element solutions for the same problems.

C. Solving the Toner-Tu equations on a cylindrical surface

To test our numerical implementation of the Toner-Tu equations in the finite element setting for curved surfaces, we begin with the highly symmetrical case of a cylindrical surface. The geometry of the cylindrical surface is parameterized using $\mathbf{r}(\theta, z) = (R\cos\theta, R\sin\theta, z)$. Using this parametrization, we seek to analyze the steady state and the dynamics of **v** for a simplified version of Eq. (81),

$$\frac{\partial v^\mu}{\partial t} = [\alpha(\rho - \rho_c) - \beta g_{ab} v^a v^b] v^\mu - \sigma \nabla^\mu \rho - \lambda v^\nu \nabla_\nu v^\mu, \tag{82}$$

in which the neighbor coupling term with coefficient D is neglected. In Appendix A 2, we first present a complete derivation of the steady-state solution of this equation for the cylinder, which leads to the result

$$|\mathbf{v}_{ss}|^2 = \frac{\alpha(\rho_0 - \rho_c)}{\beta}, \tag{83}$$

where $\mathbf{v}_{ss}$ is the steady-state velocity field, whose only nonzero component is v^θ. This solution represents a simple circumferential flow around the cylinder coupled to a uniform density field.

We then derive the dynamics of a relaxation to the steady state from an initial condition with the same symmetry of circumferential flow and uniform density, but with an arbitrary initial speed. Given an initial magnitude $v(0)$ and a uniform density ρ, the solution for the time-dependent relaxation is given by

$$|\mathbf{v}|^2 = \frac{\frac{\alpha(\rho_0 - \rho_c)}{\beta}}{1 - \left(1 - \frac{\alpha(\rho_0 - \rho_c)}{\beta v(0)^2}\right) e^{-2\alpha(\rho_0 - \rho_c)t}}. \tag{84}$$

In Fig. 10, we compare the dynamics predicted here analytically to the numerical solution of our curved-surface formulation using the FEM. As with many theoretical analyses, one of the most important aspects of our solutions is understanding how the phenomenology depends upon parameters. Here we see how the parameter α controls the saturation value of $|\mathbf{v}_{ss}|^2$ as well as the timescale $[\tau = 1/2\alpha(\rho_0 - \rho_c)]$ to achieve that saturation. Figure 10 reveals an excellent correspondence between the analytical results and their numerical counterparts as explored using our finite element implementation.

Although it tests a subset of our full surface formulation due to the highly symmetric geometry and initial condition that make this analytic solution tractable, we view the correspondence between the analytical results and the finite element solution as a helpful validation of our formulation and numerical approach, which uses the full projection machinery of the previous sections rather than a knowledge of vector calculus in cylindrical coordinates.

D. Solving the Toner-Tu equations on a spherical surface

We next take up an analysis of the minimal Toner-Tu model on the sphere. Previous works [39,53] described a highly symmetric rotating band solution for flocks on the sphere, in which both the density and velocity depend only upon the polar angle θ at the steady state and not upon the azimuthal angle ϕ. Here we present the analytical solution for density and velocity on the sphere derived by Shankar *et al.* [39], for comparison to our numerical results. For their steady-state analysis, Shankar *et al.* [39] make the approximation that the neighbor coupling term with coefficient D is absent, resulting in a simpler minimal version of the steady-state Toner-Tu equations of the form

$$\lambda v^\nu \nabla_\nu v^\mu = [\alpha(\rho - \rho_c) - \beta g_{ab} v^a v^b] v^\mu - \sigma \nabla^\mu \rho. \tag{85}$$

We note that while Shankar *et al.* chose the field variables ρ and $\mathbf{p} = \rho\mathbf{v}$, we continue in the language of ρ and **v**. In Appendix A 4, following Shankar *et al.*, we present a step-by-step derivation of the steady-state solution of Eq. (85) using the spherical surface parametrization $\mathbf{r}(\theta, \phi) = (R\cos\phi\sin\theta, R\sin\phi\sin\theta, R\cos\theta)$. This analytical calculation leads to the steady-state solution for flock or herd density on the sphere,

$$\rho_{ss}(\theta) = \rho_c + (\rho_0 - \rho_c)A_\eta \sin^\eta \theta, \tag{86}$$

where η is a dimensionless parameter given by

$$\eta = \frac{\lambda\alpha}{\beta\sigma} \tag{87}$$

and the prefactor A_η is defined by

$$A_\eta = 2\Gamma((3 + \eta)/2)/[\sqrt{\pi}\Gamma(1 + \eta/2)], \tag{88}$$

and Γ is the gamma function. Note that the dimensionless parameter η gives us a convenient knob to tune in our
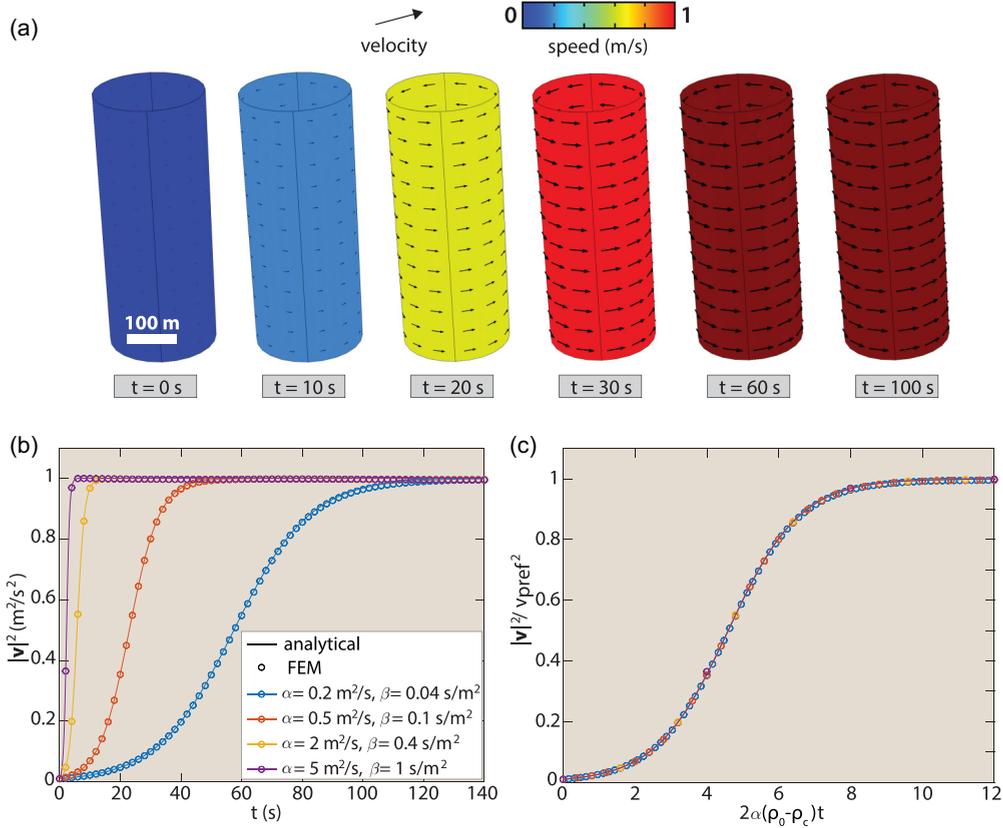
FIG. 10. Dynamics of Toner-Tu flow on a cylinder. (a) The initial condition is a uniform circumferential flow field with a very small initial speed. Over time, the speed relaxes to the steady-state value $|\mathbf{v}_{ss}| = \sqrt{\alpha(\rho_0 - \rho_c)/\beta}$. The timescale for this relaxation is $\tau = 1/(2\alpha(\rho_0 - \rho_c))$. (b) Comparison of analytic and numerical (COMSOL) results for dynamics of $|\mathbf{v}|^2$ for different choices of parameters. The asymptotic values of the curves correspond to the steady-state solution worked out earlier in this section. The ratio α/β is fixed such that the steady-state velocity of 1 m/s is the same for all four cases and $|\mathbf{v}|(0) = 0.1$ m/s. (c) Data collapse in which $|\mathbf{v}|^2$ is plotted against dimensionless time, revealing that all curves are dictated by the same underlying timescale, set by $\alpha(\rho_0 - \rho_c)$.

calculations, which can be compared directly with the numerical results of our finite element calculations as shown in Fig. 11(b). The analytic solution and the result of our finite element treatment of the same problem agree convincingly.

We view this as a crucial validation that explores the parameter dependence of the solution and tests that the finite element treatment of the various curved-space derivatives is done correctly.
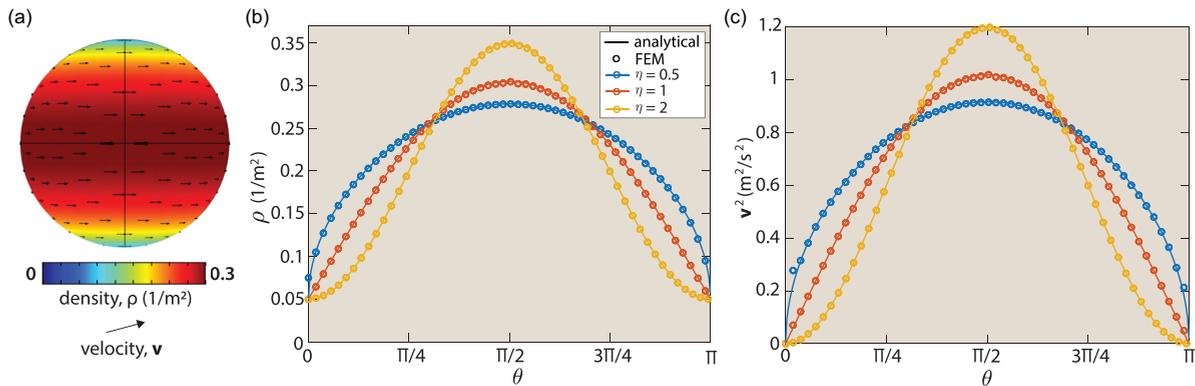

FIG. 11. Comparison of analytic and numerical (COMSOL) solutions to the Toner-Tu equations on a sphere. (a) Example steady-state solution for Toner-Tu herds on the surface of a sphere, for $\eta = 1$. (b) Steady-state herd density on the sphere. Different choices of the dimensionless parameter η elicit different density profiles. (c) Steady-state velocity $\mathbf{v}^2$ for Toner-Tu herds on the surface of a sphere. As in (b), different choices of the dimensionless parameter η result in different profiles. Parameter values used: $\alpha = 2\,\mathrm{m^2/s}$; $\sigma = 40\,\mathrm{m^4/s^2}$, $20\,\mathrm{m^4/s^2}$, or $10\,\mathrm{m^4/s^2}$, respectively; others as in Sec. IV. In both (b) and (c), analytical results (curves) are in good accord with numerical results (open circles) obtained using the finite element method.
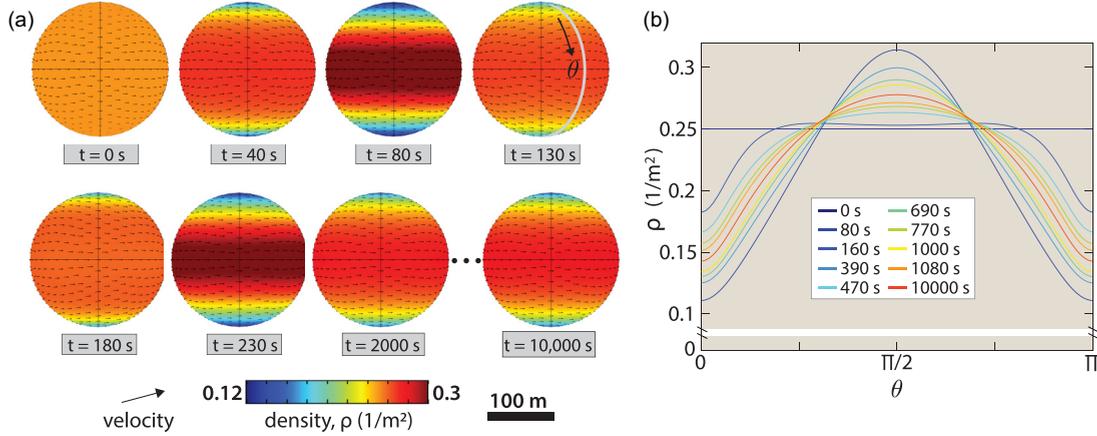
FIG. 12. Toner-Tu dynamics on the sphere. (a) Dynamics of herd density and velocity on the sphere, during the evolution from an initial condition of uniform density and circumferential velocity $v_\phi = v_{\text{pref}} \sin \theta$ toward the steady-state band pattern shown at 2000 s and 10 000 s. In the long-time limit, density oscillations are damped out and a steady state emerges. (b) Density as a function of angular position, showing the damping of density oscillations over time. Parameters as defined in Sec. IV. See also video 2 (SM).

In Appendix A 4, following the work of Shankar *et al.* [39], we also derive an analytical solution for steady-state herd velocity on the sphere as

$$|\mathbf{v}_{ss}|^2 = \frac{\alpha}{\beta}(\rho_0 - \rho_c)A_\eta \sin^\eta \theta. \tag{89}$$

For the specific case of $\eta = 2$, for example, we have

$$A_2 = \frac{2\Gamma\left(\frac{5}{2}\right)}{\sqrt{\pi}\,\Gamma(2)} = \frac{3}{2}. \tag{90}$$

Using this value for A_2, we can write the density and the velocity fields as

$$\rho_{ss}(\theta) = \rho_c + \frac{3}{2}(\rho_0 - \rho_c)\sin^2\theta \tag{91}$$

and

$$|\mathbf{v}_{ss}|^2 = \frac{3}{2}(\rho_0 - \rho_c)\sin^2\theta, \tag{92}$$

respectively. These solutions from the work of Shankar *et al.* [39] provide an important opportunity to test our numerical implementation of the Toner-Tu theory on a curved surface, as shown in Fig. 11(c) for $|\mathbf{v}_{ss}|^2$. As for the density, once again the correspondence between the analytic results and the finite element treatment of the problem is convincing.

VI. WILDEBEEST ON SPHERES AND HILLS: COMPLEX FLOCKING DYNAMICS

In the remainder of the paper, we move beyond situations where analytical solutions are feasible, using the curved-space Toner-Tu formulation and the FEM to explore more complex herding dynamics and herding behaviors on more complex geometries.

A. Toner-Tu dynamics on the sphere

First, we continue with the case study of herding on the sphere but take a step beyond the analytic solution provided in Sec. V. We add back in the analytically intractable neighbor coupling term $D\nabla^2 \mathbf{v}$ and examine its contribution. In other words, we explore the dynamics of density and velocity on

the spherical surface using the full finite element formulation of Sec. III, including all the terms present in Eq. (41). We begin with uniform density and a circumferential velocity field $v_\phi = v_{\text{pref}} \sin \theta$. As seen in Fig. 12 and video 2, this solution develops oscillations in which a nonuniform density moves back and forth in the θ direction, from the north and south poles to the equator. The presence of the neighbor coupling term damps out these oscillations as shown in Fig. 12(b). We qualitatively observe that smaller D corresponds to a longer damping time. In the long-time limit, damping by the neighbor coupling term enables the emergence of a steady state. Figure 13 shows several of these steady-state solutions for different choices of the neighbor coupling coefficient D. We note that at small D, the solution approaches the steady-state analytic solution which is attainable in the absence of this term.

However, the dynamics that can take place on the spherical surface can be significantly richer than the steady-state band solution shown above. To seed the dynamics, we consider an initial condition in which the active agents are confined to an angular wedge between $\phi = 0$ and $\phi = \pi/2$, again with circumferential velocity $v_\phi = v_{\text{pref}} \sin \theta$, as shown in Fig. 14 and video 3 (SM). This asymmetric initial condition can give rise to both dynamic and steady-state solutions, depending on the value of parameters such as ρ_0 or σ. In Fig. 14, different choices of σ lead to different herding behaviors in the long-time limit. In Fig. 14(a), at small σ, a droplet-shaped patch of dense wildebeests circulates around the sphere indefinitely. Tuning σ to increase the magnitude of the pressure term leads instead to the emergence of oscillating density bands, as shown in Fig. 14(b), and then to steady-state density bands, as shown in Fig. 14(c). Note that we have not highlighted the interesting transient dynamics displayed as herds evolve from the wedge initial condition toward the patch and band patterns. For the rotating patch and steady-state band solutions, which reach dynamic or steady-state solutions with mirror symmetry across the equatorial plane ($\theta = \pi/2$), we can plot herd density as a function of ϕ to characterize the shape of the density patch and to further explore the effect of increasing
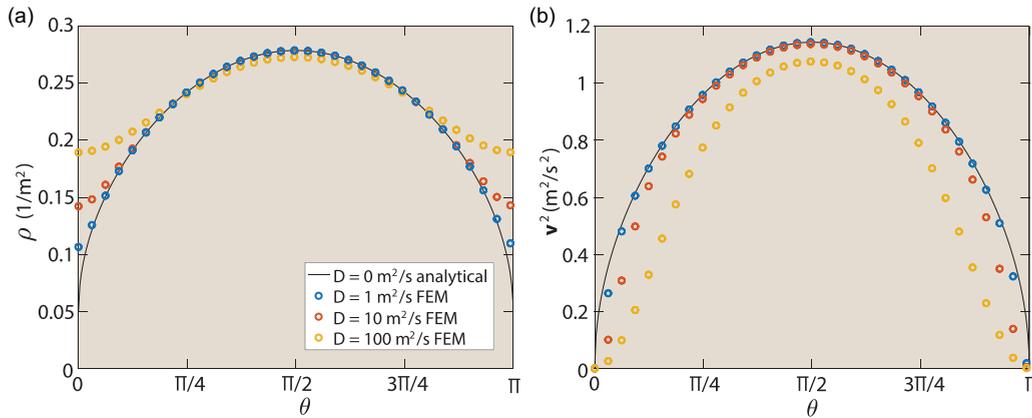
FIG. 13. Steady-state solutions on the sphere including the neighbor coupling term. (a) Steady-state density as a function of angular position θ for different choices of the neighbor coupling coefficient D. An analytic solution in the absence of the D term is plotted for comparison. (b) Steady-state velocity squared as a function of angular position θ for the same parameters values and analytic solution as shown in (a).
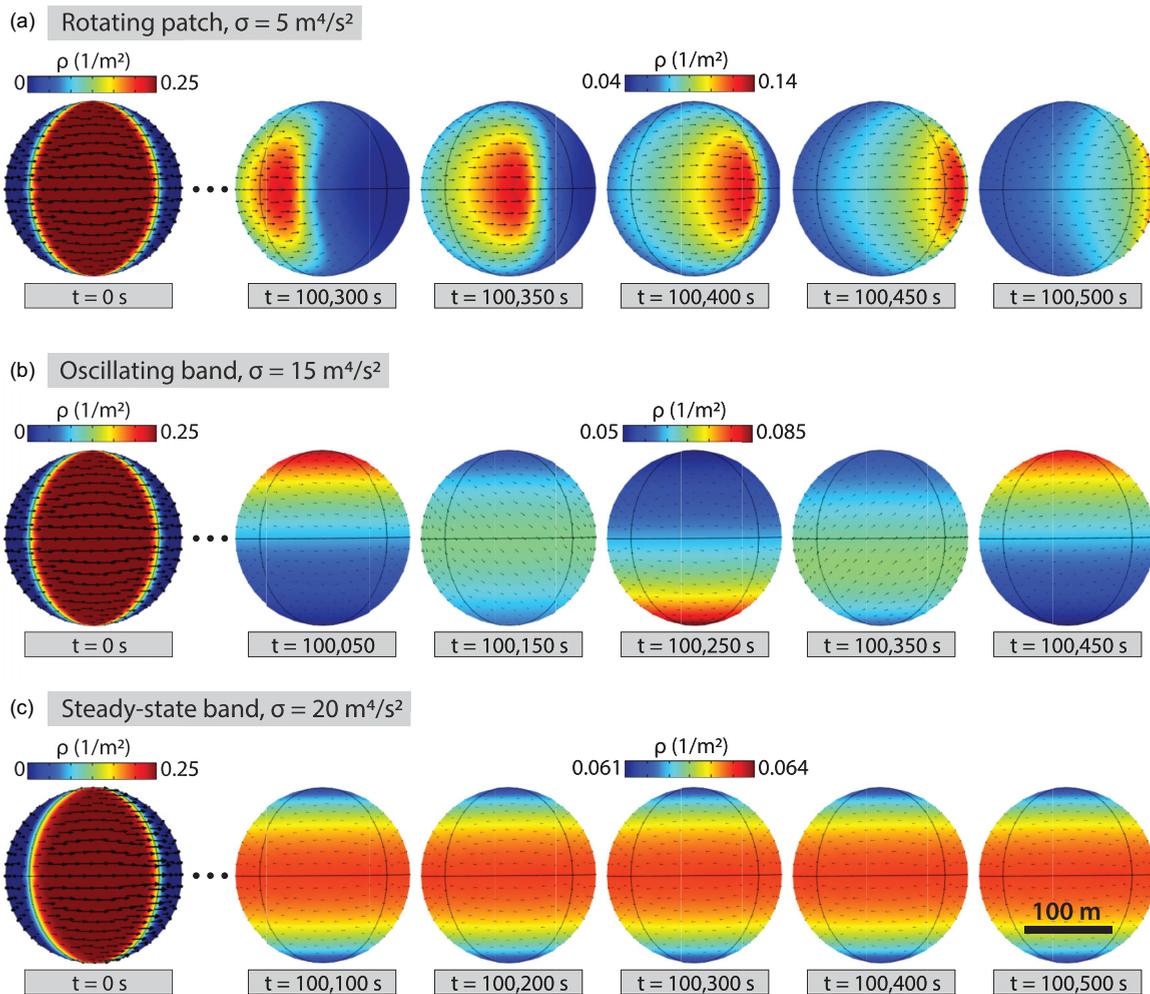

FIG. 14. Herds on the sphere can form rotating patch, oscillating band, and steady-state band patterns. In these examples, herd dynamics are seeded at $t = 0$ s by confining wildebeest to an angular wedge between $\phi = 0$ and $\phi = \pi/2$ and giving them a circumferential velocity $v_\phi = v_{\text{pref}} \sin \theta$. (a) Given a smaller pressure term ($\sigma = 5 \, \text{m}^4/\text{s}^2$), a rotating patch pattern of herd density emerges. (b) A larger pressure term ($\sigma = 15 \, \text{m}^4/\text{s}^2$) leads the density field to exhibit temporal oscillations about the equatorial plane ($\theta = \pi/2$), while remaining axisymmetric in ϕ. (c) Given an ever larger pressure term ($\sigma = 20 \, \text{m}^4/\text{s}^2$), herd density is restricted to a narrow range around ρ_0(0.0625 m^{-2}), and a steady-state banded pattern of density emerges. See also video 3 (SM).
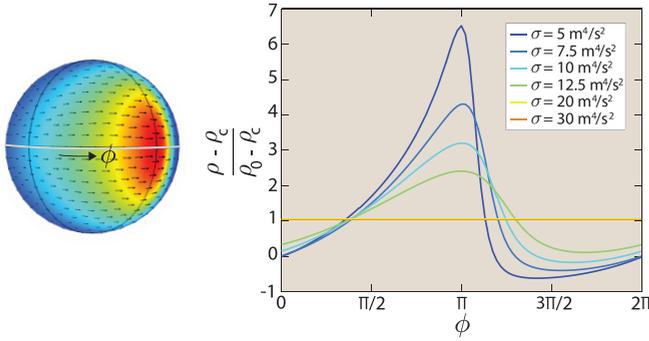
FIG. 15. Increasing the Toner-Tu pressure term drives herd density from patch to band patterns. Plot shows density as a function of ϕ, angular position around the sphere circumference. The simulation was seeded with the initial condition depicted in Fig. 14 for varied σ that lead to rotating patch or steady-state band solutions in the long-time limit. Curves are aligned such that their maxima are at $\phi = \pi$.

the Toner-Tu pressure term. As shown in Fig. 15, increasing σ drives the emergent density pattern from a patch to a band by flattening $\rho(\phi)$ from the interesting asymmetric patch shape observed at $\sigma = 5\,\mathrm{m^4/s^2}$ to the uniform profile observed for $\sigma = 20\,\mathrm{m^4/s^2}$ and $\sigma = 30\,\mathrm{m^4/s^2}$.

B. Herding over a Gaussian hill

Moving beyond the symmetric geometry of the sphere, we next explore the contribution of curvature to steady-state density and velocity for a herd on a racetrack. As shown in Fig. 16, the addition of a Gaussian hill on the racetrack straightaway leads to local changes in herd density. Density is lower on the face of the hill approached by the wildebeests, as they climb and their velocity field diverges, increasing the space between wildebeest neighbors as their paths bend around the hill. Density is higher on the face of the hill that the wildebeest descend, where the velocity field again converges. This curvature-dependent effect is loosely reminiscent of lensing [54,55], but we have not rigorously explored the analogy.

C. Herding on an undulating island landscape

In a final case study, inspired by wildebeests herding on complicated natural landscapes, we consider the motion of a Toner-Tu herd on an island of Gaussian hillsides. We offer this study as a step toward a dialogue with real-world data, for which we hope the approach presented in this paper may prove useful. For this final exploration of how herding behavior is altered by the local topography of a landscape, we develop one additional term for the herding equations that govern the time evolution of wildebeest velocity.

In much the same way that body forces are incorporated into the Navier-Stokes equations, we hypothesize that the topography of the landscape upon which the wildebeest are marauding acts upon them through a gravitational force term of the form

$$\mathbf{f}_g = -\zeta \mathbf{k}, \tag{93}$$

where $\mathbf{k}$ is the unit vector in the z direction (height). We note that the units of ζ are not strictly those of force since the Toner-Tu equations simply have dimensions of acceleration. In the same spirit that we selected the other parameters in Sec. IV to ensure terms of comparable magnitude, we estimate that ζ is on the order of $0.1\,\mathrm{m/s^2}$. For the simulations shown in Fig. 17, we chose $\zeta = 0.05\,\mathrm{m/s^2}$ to prevent the emergence of unphysical negative densities at hilltops. To write the curved-surface contribution of the gravitational force term to the dynamical equations, we carry out the surface projection as usual,

$$\mathbf{P}\mathbf{f}_g = -\zeta \mathbf{P}\mathbf{k} = -\zeta \mathbf{k} + \zeta \mathbf{n}(\mathbf{n} \cdot \mathbf{k}), \tag{94}$$

where we recognize that $(\mathbf{n} \cdot \mathbf{k}) = n_3$. Because of the projection operator, this term affects all three components of velocity. For example, we can write the Toner-Tu equation governing the one-component of velocity, with the addition of the gravitational force, as

$$\frac{\partial v_1^{\parallel}}{\partial t} = [\alpha(\rho - \rho_c) - \beta v_j^{\parallel} v_j^{\parallel}] v_1^{\parallel} - \sigma \left(\frac{\partial \rho}{\partial x_1} \right)_{\Gamma} + D P_{1l} \left(\frac{\partial G_{lj}}{\partial x_j} \right)_{\Gamma}$$
$$- \lambda v_j^{\parallel} \left[\left(\frac{\partial v_1^{\parallel}}{\partial x_j} \right)_{\Gamma} - n_1 n_k \left(\frac{\partial v_k^{\parallel}}{\partial x_j} \right)_{\Gamma} \right] + \zeta n_1 n_3. \tag{95}$$
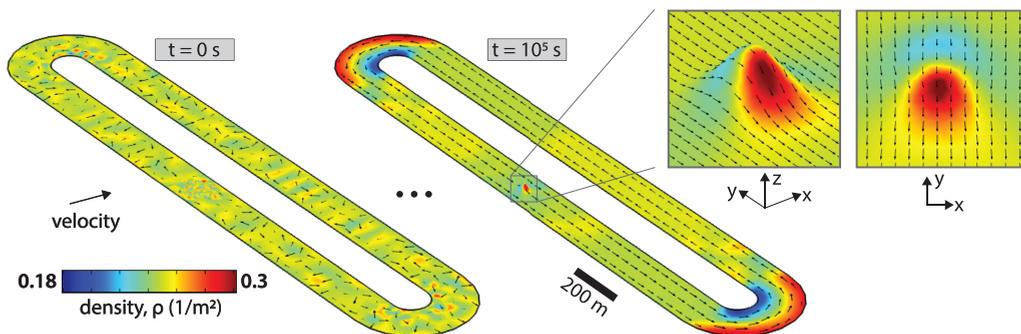

FIG. 16. Curvature induces local density and velocity changes. Steady-state herd density (colors) and velocity (black arrows) for wildebeest traversing a Gaussian hillside embedded in a racetrack straightaway. From an initial condition of random velocity orientation, a steady-state recirculation around the racetrack emerges in the long-time limit. In an effect reminiscent of lensing [54,55], the Gaussian hill's curvature induces low and high density regions.
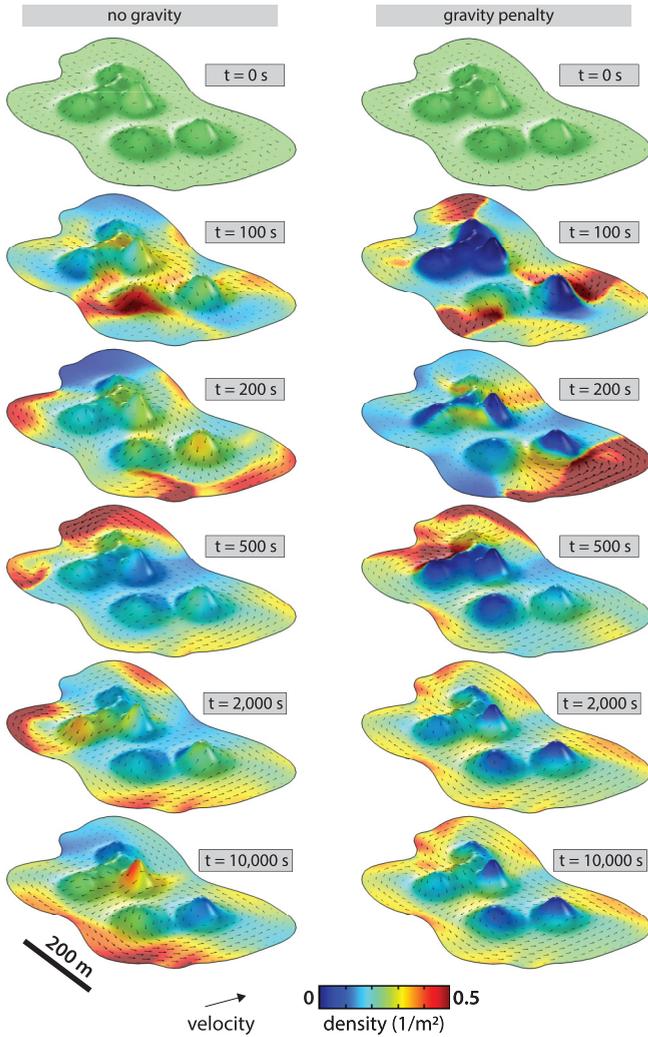
FIG. 17. Dynamics of a Toner-Tu herd on an undulating landscape, initialized with uniform density $\rho_0 = 0.25 \, \text{m}^{-2}$ and randomly oriented velocities. Comparing the left column ($\zeta = 0 \, \text{m/s}^2$) and the right column ($\zeta = 0.05 \, \text{m/s}^2$) highlights the contribution of the gravitational force. In the presence of the gravitational force, the hills act as soft obstacles, favoring herd circulation around them. Other parameters used were presented in Sec. IV. See also video 4 (SM).

As in the channel flow example, where the preferred speed term and the neighbor coupling term compete to modify velocity magnitude near no-slip boundaries, the preferred speed and the gravitational term compete on sloped hillsides.

To explore the dynamics of an active herd on this undulating landscape, we seed the region with an initial distribution of wildebeest at uniform density ρ_0 and with random velocity orientation. From this initial distribution, over time the wildebeests self-organize into a herd that navigates around the complex landscape. In Fig. 17 and video 4 (SM), we show an example of these dynamics and explore the contribution of the gravitational force term. In the presence of the gravitational force, the hills effectively serve as "soft" obstacles, directing the herd to circulate around them. These obstacles can stabilized fixed patterns in the velocity field, as on the right in Fig. 17 and video 4 (SM).

We hope that these playful Toner-Tu solutions on the complex geometry of a hill-studded island show the versatility of the approach presented in this paper. We also hope that they visually illustrate a vision of the dialogue between theory and real-world data towards which we ultimately strive, and towards which the current paper takes one step.

VII. CONCLUSION

The collective motion of animals, cells, and molecules presents a beautiful phenomenon. Toner-Tu flocking theory has proven a versatile mathematical description of such phenomena, helping us to enable an expansion from traditional continuum models of materials into the relatively uncharted waters of the dynamics of living matter. Toner-Tu and other continuum active matter models make possible a dialogue between mathematical prediction and the stunning living world outside our windows. To undertake such a dialogue necessitates solving these equations for real-world geometries and boundary conditions. Flocks of sheep, collectively migrating cells, collectively flowing cytoskeletal filaments—all of these flocking agents move within structured environments that must be accounted for to make a comparison between *in vivo* data and predictions of theoretical models.

Here, we focused on flocking phenomena that occur on curved surfaces, such as wildebeests navigating terrain, or the motility-driving flows of cytoskeletal filaments at the surface of protozoan parasites of particular interest to us [27]. Both the formulation and solution of the Toner-Tu equations is more challenging once our goal is their implementation on curved surfaces. Frequently, surfaces of biological interest are highly asymmetric, rendering analytical approaches challenging or impossible. In this work we sought to present, in accessible form, a general curved-surface formulation of the Toner-Tu equations and its FEM implementation. We hope that our pedagogical derivation of a surface formulation using the tools of extrinsic differential geometry, and our explication of numerically solving it on arbitrary curved surfaces, will prove useful to others interested in a versatile approach for exploring continuum theories on complex shapes. We also sought to demonstrate the equivalence of our approach to analytical ones, for simple geometries where analytical results are licensed. Finally, we sought to harness the versatile power of this approach by exploring its predictions across a range of geometries, boundary conditions, and parameter values. By observing the effect of geometry and parameter choice on herd direction reversal in a channel with a circular obstacle in Fig. 8, for example, we demonstrated the ability of this approach to generate hypotheses that can subsequently be explored analytically, as in Fig. 9. By exploring the local density changes induced by curvature for herds passing over a Gaussian hill in Fig. 16, we highlighted the ability of this approach to generate intriguing observations that generate new questions. By predicting the dynamics of herds on a undulating island landscape in Fig. 17, we took a step towards a vision of predictive dialogue with data from the rich, structured environment in which we live.

Looking forward, we dream of more dialogue between continuum active matter theory and measurements taken in our rich living world, from drone footage of migrating animal

herds to movies of microscopic cytoskeletal flows within embryos and cells. Such a dialogue will enable the determination of absolute parameter values for such systems, which will in turn make possible the contrivance of experiments designed to test the theory. Can the same field theories, indeed, be applied to sheep and to microscopic microtubules? What is the effect of curvature on velocity for different types of agents? We hope that our presentation of the approach taken here will contribute towards efforts to bridge active matter theory and complex real-world shapes, directly testing whether our observations of the living world can be understood and predicted in the language of continuum theories of active collectives.

ACKNOWLEDGMENTS

APPENDIX

1. Calculus on surfaces via projection operators

The formulation of the coupled partial differential equations that describe the herd density and velocity fields required us to perform various derivative operations on surfaces. To do so, we used extrinsic differential geometry to take the tangential calculus approach described pedagogically by Jankuhn *et al.* [34] and Needham [47]. This approach involves defining all necessary differential operators in terms of the projection operator, $\mathbf{P} = \mathbf{I} - \mathbf{n} \otimes \mathbf{n}$ or $P_{ij} = \delta_{ij} - n_i n_j$, instead of working in the language of parameterized surfaces. In this Appendix, we set these tangential calculus operations front and center. We define the tangential curved-space differential operators used in this paper, and we attempt to convey some intuition for what these operations achieve.

a. Gradient of a scalar

Our treatment of Toner-Tu surface problems requires us to compute derivatives in the tangent plane of the surface of interest. For example, we might be interested in the gradient of a scalar field such as the density $\rho(\mathbf{r})$. We can think of the surface gradient as resulting from projecting the full 3D gradient onto the tangent plane, using the projection operator

$\mathbf{P}$ defined in Eq. (3). Thus, we find

$$\mathbf{P}\nabla\rho = (\mathbf{I} - \mathbf{n} \otimes \mathbf{n})\nabla\rho$$
$$= (\delta_{ij} - n_i n_j)\frac{\partial\rho}{\partial x_j} = \frac{\partial\rho}{\partial x_i} - n_i n_j \frac{\partial\rho}{\partial x_j}. \quad \text{(A1)}$$

Hence, throughout the paper, whenever we need to invoke the tangent curved-space gradient operator, we do so in the form

$$(\nabla_\Gamma \rho)_i = (\nabla\rho)_i - n_i n_j \frac{\partial\rho}{\partial x_j} = \frac{\partial\rho}{\partial x_i} - n_i n_j \frac{\partial\rho}{\partial x_j}. \quad \text{(A2)}$$

b. Gradient of a vector

Several terms in the minimal Toner-Tu model involve more complex differential operations. For example, the advection term in the Toner-Tu theory requires us to evaluate the gradients of the velocity field (a vector!). Thus, we also need to formulate the mathematical tools for constructing the gradient of a vector. In direct notation, the tangential curved-space gradient of a vector is

$$\nabla_\Gamma \mathbf{v} = \mathbf{P}(\nabla\mathbf{v})\mathbf{P}, \quad \text{(A3)}$$

which we can write in indicial notation as

$$(\nabla_\Gamma \mathbf{v})_{ij} = [\mathbf{P}(\nabla\mathbf{v})\mathbf{P}]_{ij} = P_{ik}\frac{\partial v_k}{\partial x_l}P_{lj}. \quad \text{(A4)}$$

We can expand this out using the definition of the projection operator, resulting in

$$P_{ik}\frac{\partial v_k}{\partial x_l}P_{lj} = (\delta_{ik} - n_i n_k)\frac{\partial v_k}{\partial x_l}(\delta_{lj} - n_l n_j), \quad \text{(A5)}$$

which simplifies to the form

$$P_{ik}\frac{\partial v_k}{\partial x_l}P_{lj} = \left(\frac{\partial v_i}{\partial x_j} - n_l n_j \frac{\partial v_i}{\partial x_l}\right) - n_i n_k \left(\frac{\partial v_k}{\partial x_j} - n_l n_j \frac{\partial v_k}{\partial x_l}\right), \quad \text{(A6)}$$

an expression that we will see repeatedly. We note that the final result features the projected gradient of the v_i and v_k components of velocity, respectively.

c. Divergence of a vector

Using the definition of the gradient of a vector presented above, we can formally write the tangential curved-space *divergence* of a vector as

$$\text{div}_\Gamma \mathbf{v} = \text{tr}(\nabla_\Gamma \mathbf{v}) = \text{tr}(\mathbf{P}(\nabla\mathbf{v})\mathbf{P}) = \text{tr}(\mathbf{P}(\nabla\mathbf{v})) = \text{tr}((\nabla\mathbf{v})\mathbf{P}). \quad \text{(A7)}$$

The final two expressions follow from the fact that $\text{tr}(\mathbf{ABC}) = \text{tr}(\mathbf{BCA})$ and all other cyclic permutations as well, in conjunction with the fact that $\mathbf{P}^2 = \mathbf{P}$. In component form, this simplifies to

$$\text{div}_\Gamma v_i = \frac{\partial v_i}{\partial x_i} - n_i n_k \frac{\partial v_i}{\partial x_k}, \quad \text{(A8)}$$

an expression we used repeatedly in our derivations of the field equations and in their formulation numerically within COMSOL.

d. Divergence of a tensor

As seen in the main text, the Toner-Tu theory requires us to evaluate the divergence of tensorial quantities on curved surfaces. For example, we invoked the stresslike quantity **J**, which in the full three-dimensional setting we write as

$$\nabla \cdot \mathbf{J} = J_{ji,j}. \tag{A9}$$

This result needs to be amended in the context of an arbitrary surface. To define the divergence of a tensor, we can examine separately the divergence of the component vectors of that tensor, which we obtain by having the tensor act on the unit vectors. More precisely, we work out the divergence of a tensor **A**, following Jankuhn *et al.* [34], as

$$\text{div}_\Gamma \, \mathbf{A} = \left(\text{div}_\Gamma \left(\mathbf{e}_1^T \mathbf{A} \right), \text{div}_\Gamma \left(\mathbf{e}_2^T \mathbf{A} \right), \text{div}_\Gamma \left(\mathbf{e}_3^T \mathbf{A} \right) \right)^T \tag{A10}$$

where we use the definitions

$$\mathbf{e}_1^T = (1, 0, 0), \tag{A11}$$

$$\mathbf{e}_2^T = (0, 1, 0), \tag{A12}$$

and

$$\mathbf{e}_3^T = (0, 0, 1). \tag{A13}$$

We recall that the divergence of a tensor gives rise to a vector, consistent with the multiple components in our expression for $\text{div}_\Gamma \, \mathbf{A}$. Indeed, we can rewrite this expression in component form as

$$(\text{div}_\Gamma \, \mathbf{A})_i = P_{lk} \frac{\partial A_{il}}{\partial x_k} = \frac{\partial A_{ik}}{\partial x_k} - n_l n_k \frac{\partial A_{il}}{\partial x_k}. \tag{A14}$$

e. Laplacian of a vector

As shown by Jankuhn *et al.* [34], the surface-projected version of the Laplacian term (see their Eq. (3.16) for the surface Navier-Stokes equations) is given by

$$\underbrace{\nabla^2 \mathbf{v}}_{\text{in the plane}} = \underbrace{\text{div}(\nabla \mathbf{v})}_{\text{in the plane}} \longrightarrow \underbrace{\mathbf{P} \, \text{div}_\Gamma (\nabla_\Gamma \mathbf{v})}_{\text{on a curved surface}}. \tag{A15}$$

We will now unpack this expression for the curved surface Laplacian, from right to left. We must first compute the surface gradient of the vector **v**, given by the tensor $\mathbf{G} = \nabla_\Gamma \mathbf{v}$. In component form, we write this as

$$\nabla_\Gamma \mathbf{v} = G_{ij} = \left(\frac{\partial v_i}{\partial x_j} - n_l n_j \frac{\partial v_i}{\partial x_l} \right) - n_i n_k \left(\frac{\partial v_k}{\partial x_j} - n_l n_j \frac{\partial v_k}{\partial x_l} \right). \tag{A16}$$

We then invoke the definition of the surface-projected divergence of the tensor **G**:

$$\text{div}_\Gamma (\mathbf{G})_i = P_{lk} \frac{\partial G_{il}}{\partial x_k}. \tag{A17}$$

Using the definition of the projection operator, we can expand this out as

$$P_{lk} \frac{\partial G_{il}}{\partial x_k} = (\delta_{lk} - n_l n_k) \frac{\partial G_{il}}{\partial x_k} = \frac{\partial G_{il}}{\partial x_l} - n_l n_k \frac{\partial G_{il}}{\partial x_k}. \tag{A18}$$

To complete our calculation of the curved-space Laplacian of a vector, we need to compute $\mathbf{P} \, \text{div}_\Gamma (\mathbf{G})$, which we can write

in indicial notation as

$$P_{il} \, \text{div}_\Gamma \, (\mathbf{G})_i = P_{il} \left(\frac{\partial G_{lj}}{\partial x_j} - n_j n_k \frac{\partial G_{lj}}{\partial x_k} \right). \tag{A19}$$

2. Implementation in COMSOL MULTIPHYSICS

Our COMSOL MULTIPHYSICS files are available for download, use, and adaptation at Ref. [48]. They are accompanied by a "How-To" document that provides a tutorial on navigating and using these files.

We hope that others will find our approach user-friendly and adaptable for solving the Toner-Tu equations on other complex curved surfaces, for solving other continuum equations on curved surfaces, or for additional study of the cases presented here. For our FEM calculations, we chose to use the off-the-shelf software COMSOL MULTIPHYSICS for its accessibility and learner-friendly interface. We note with regret, however, that the use of this software and our code requires access to a paid COMSOL MULTIPHYSICS license. If you do not already have a license and are affiliated with an institution, we recommend looking into access options through a shared software library.

Below, for the convenience of COMSOL users among our readers, we highlight some of the COMSOL-specific features and nomenclature that we used.

To solve our custom surface partial differential equations, we used the COMSOL MULTIPHYSICS General Form Boundary PDE interface and took advantage of COMSOL's built-in tangential differentiation operator, dtang(f,x), discussed in detail below. We also made use of the normal vector (nx, ny, nz), a built-in geometric variable. For a thorough and practical introduction to the FEM, we recommend Ref. [56]. For simulations initialized with a disordered velocity field, every node's velocity orientation was drawn randomly from a uniform distribution of angles between 0 and 2π and had magnitude $v(0)$. Unless noted otherwise, as in the case of the cylinder dynamics shown in Fig. 10, $v(0) = v_{\text{pref}} = 1$ m/s. Other parameters are defined and discussed in detail in Sec. IV. On curved surfaces, to avoid the accumulation of out-of-plane components in **v** from numerical error, we implemented a weak constraint of $\mathbf{n} \cdot \mathbf{v} = 0$. Similarly, we implemented a global constraint on the total integrated density ρ on the surface, ensuring it stayed at its initial value. Default COMSOL solvers and settings were used: implicit backward differentiation formula for time stepping and multifrontal massively parallel sparse direct solver for the linear direct spatial solver. Our meshes employed between 1310 and 6525 triangular elements (1310–1674 for the channel with obstacles; 1312 for the cylinder; 6525 for the sphere; 1604 for the racetrack with hill; 2642 for the island) and can be viewed in the source files available at Ref. [48].

We now provide additional details on formulating our custom equations to solve them in COMSOL. The user-defined partial differential equation is written in COMSOL as

$$\frac{\partial \mathbf{v}}{\partial t} + \nabla \cdot \mathbf{\Gamma}^{(v)} = \mathbf{f}^{(\mathbf{v})}. \tag{A20}$$

Here, the whole formulation comes down to the definitions of $\mathbf{\Gamma}^{(v)}$ and $\mathbf{f}^{(\mathbf{v})}$. The fluxlike quantity $\mathbf{\Gamma}^{(v)}$ is a 3×3 matrix that

is defined such that

$$(\boldsymbol{\nabla} \cdot \boldsymbol{\Gamma}^{(v)})_i = \frac{\partial \Gamma_{ji}^{(v)}}{\partial x_j}. \tag{A21}$$

In this section, we use the notations $\boldsymbol{\Gamma}$ and $\mathbf{f}$ to precisely match the notation of COMSOL's General Form PDE interface. We note that in the main body of the paper we referred to $\mathbf{J}^{(v)}$ rather than $\boldsymbol{\Gamma}^{(v)}$ to avoid confusion with subscript Γ, which signifies "surface," following the notation of Jankuhn *et al.* [34] in their description of the tangential calculus. There, we wanted to give readers the opportunity to make contact with Jankuhn *et al.* with minimal difficulty.

COMSOL is constructed to allow us to evaluate what we have earlier called ∇_Γ using an operation known as dtang(f,x_i), which is the tangent plane component of the gradient in the ith direction of the function $f(x_1, x_2, x_3)$. We can write this formally as

$$\begin{aligned} \boldsymbol{\nabla}_\Gamma f &= \boldsymbol{\nabla} f - \mathbf{n}(\mathbf{n} \cdot \boldsymbol{\nabla} f) \\ &= (\text{dtang}(f, x_1), \text{dtang}(f, x_2), \text{dtang}(f, x_3)) \end{aligned} \tag{A22}$$

Thus, these terms can be written in component form as

$$\text{dtang}(f, x_1) = (\boldsymbol{\nabla}_\Gamma f)_1 = \frac{\partial f}{\partial x_1} - n_1 n_j \frac{\partial f}{\partial x_j}. \tag{A23}$$

Specifically, we have for the gradient in the one-direction:

$$\text{dtang}(f, x_1) = \frac{\partial f}{\partial x_1} - n_1 \left(n_1 \frac{\partial f}{\partial x_1} + n_2 \frac{\partial f}{\partial x_2} + n_3 \frac{\partial f}{\partial x_3} \right), \tag{A24}$$

with similar results for the other directions given by

$$\text{dtang}(f, x_2) = \frac{\partial f}{\partial x_2} - n_2 \left(n_1 \frac{\partial f}{\partial x_1} + n_2 \frac{\partial f}{\partial x_2} + n_3 \frac{\partial f}{\partial x_3} \right) \tag{A25}$$

and

$$\text{dtang}(f, x_3) = \frac{\partial f}{\partial x_3} - n_3 \left(n_1 \frac{\partial f}{\partial x_1} + n_2 \frac{\partial f}{\partial x_2} + n_3 \frac{\partial f}{\partial x_3} \right). \tag{A26}$$

To provide detailed examples for those interested in similar FEM implementations, we now translate key expressions from our curved-space Toner-Tu formulation into COMSOL syntax. In Eq. (21), we defined the matrix G_{ij} which can be written in COMSOL syntax as

$$G_{ij} = \text{dtang}(v_i^\parallel, x_j) - n_i n_k \text{dtang}(v_k^\parallel, x_j), \tag{A27}$$

where $v_i^\parallel$ refers to the tangent plane component of the ith component of velocity. For example, in expanded form,

$$\begin{aligned} G_{21} = \text{dtang}(v_2^\parallel, x_1) &- n_2(n_1 \text{dtang}(v_1^\parallel, x_1) \\ &+ n_2 \text{dtang}(v_2^\parallel, x_1) + n_3 \text{dtang}(v_3^\parallel, x_1)). \end{aligned} \tag{A28}$$

In addition, we introduced a force term in Sec. III to recapitulate the Toner-Tu equations. When written in COMSOL format, those terms are of the form

$$\begin{aligned} f_1^{(v)} + f_1^{\text{fict}} &= (\alpha(\rho - \rho_c) - \beta((v_1^\parallel)^2 + (v_2^\parallel)^2 + (v_3^\parallel)^2))v_1^\parallel - \sigma \, \text{dtang}(\rho, x) - \lambda(v_1^\parallel \text{dtang}(v_1^\parallel, x_1) \\ &+ v_2^\parallel \text{dtang}(v_1^\parallel, x_2) + v_3^\parallel \text{dtang}(v_1^\parallel, x_3) - v_1^\parallel n_1 n_1 \text{dtang}(v_1^\parallel, x_1) - v_1^\parallel n_1 n_2 \text{dtang}(v_2^\parallel, x_1) \\ &- v_1^\parallel n_1 n_3 \text{dtang}(v_3^\parallel, x_1) - v_2^\parallel n_1 n_1 \text{dtang}(v_1^\parallel, x_2) - v_2^\parallel n_1 n_2 \text{dtang}(v_2^\parallel, x_2) - v_2^\parallel n_1 n_3 \text{dtang}(v_3^\parallel, x_2) \\ &- v_3^\parallel n_1 n_1 \text{dtang}(v_1^\parallel, x_3) - v_3^\parallel n_1 n_2 \text{dtang}(v_2^\parallel, x_3) - v_3^\parallel n_1 n_3 \text{dtang}(v_3^\parallel, x_3)) - D(\text{dtang}(P_{11}, x_1)G_{11} \\ &+ \text{dtang}(P_{11}, x_2)G_{12} + \text{dtang}(P_{11}, x_3)G_{13}) - D(\text{dtang}(P_{12}, x_1)G_{21} + \text{dtang}(P_{12}, x_2)G_{22} \\ &+ \text{dtang}(P_{12}, x_3)G_{23}) - D(\text{dtang}(P_{13}, x_1)G_{31} + \text{dtang}(P_{13}, x_2)G_{32} + \text{dtang}(P_{13}, x_3)G_{33}), \tag{A29} \end{aligned}$$

for the x_1 component. Note that the α and β terms capture the preferred speed part of Toner-Tu theory, the term involving σ treats the pressure term, the terms preceded by λ capture the advection contribution to the Toner-Tu equations, and the final pieces involving $\mathbf{P}$ and $\mathbf{G}$ subtract off the fictitious force due to the curved-space Laplacian. We have similar equations for the x_2 and x_3 components of the force.

We add here an additional note concerning our implementation of user-defined partial differential equations within COMSOL using $\nabla \cdot \boldsymbol{\Gamma}$, where the tensor $\boldsymbol{\Gamma}$ is a generalized stresslike term. As we saw in the context of the curved-space version of the Laplacian above, we wished to implement an equation of form Eq. (A19) which is *not* in the form of the divergence of a tensor. Thus, we introduced

$$\left(\frac{\partial (P_{il}G_{lj})}{\partial x_j} \right)_\Gamma = \frac{\partial (P_{il}G_{lj})}{\partial x_j} - n_j n_k \frac{\partial (P_{il}G_{lj})}{\partial x_k}. \tag{A30}$$

The right side of this equation asks us to do ordinary calculus according to the product rule, culminating in

$$\begin{aligned} \left(\frac{\partial (P_{il}G_{lj})}{\partial x_j} \right)_\Gamma &= P_{il} \frac{\partial G_{lj}}{\partial x_j} + \frac{\partial P_{il}}{\partial x_j} G_{lj} \\ &- n_j n_k \frac{\partial P_{il}}{\partial x_k} G_{lj} - n_j n_k P_{il} \frac{\partial G_{lj}}{\partial x_k}. \end{aligned} \tag{A31}$$

We can rewrite this as

$$\begin{aligned} \left(\frac{\partial (P_{il}G_{lj})}{\partial x_j} \right)_\Gamma &= P_{il} \left(\frac{\partial G_{lj}}{\partial x_j} - n_j n_k \frac{\partial G_{lj}}{\partial x_k} \right) \\ &+ \left(\frac{\partial P_{il}}{\partial x_j} - n_j n_k \frac{\partial P_{il}}{\partial x_k} \right) G_{lj}. \end{aligned} \tag{A32}$$

This proves the assertion used in the paper that the quantity we want can be written as the divergence of a tensor minus an

unwanted term that we can treat as a force, resulting in

$$P_{il}\left(\frac{\partial G_{lj}}{\partial x_j}\right)_\Gamma = \left(\frac{\partial (P_{il}G_{lj})}{\partial x_j}\right)_\Gamma - \left(\frac{\partial P_{il}}{\partial x_j}\right)_\Gamma G_{lj}. \quad (A33)$$

3. Analytical solutions for Toner-Tu on the cylinder

In this Appendix section, we derive analytical solutions on the cylinder for steady-state Toner-Tu density and velocity, and for velocity dynamics during relaxation to that steady state. These solutions are used for comparison to our FEM results in Sec. V C. To develop analytic intuition on a cylindrical surface to complement our finite element studies, our first step is the parametrization of the cylindrical surface:

$$\mathbf{r}(\theta, z) = (R\cos\theta, R\sin\theta, z). \quad (A34)$$

We have two parameters, $u^1 = \theta$ and $u^2 = z$, which provide an address for every point on the cylindrical surface. That is, every choice of (u^1, u^2) has a corresponding point on the surface. Next we need to find the tangent vectors, which are defined via

$$\mathbf{a}_i = \frac{\partial \mathbf{r}}{\partial u^i}, \quad (A35)$$

where u^i is the coordinate of the surface defined above. Given these definitions, we have

$$\mathbf{a}_\theta = \frac{\partial \mathbf{r}}{\partial \theta} = (-R\sin\theta, R\cos\theta, 0) = R\mathbf{e}_\theta \quad (A36)$$

and

$$\mathbf{a}_z = \frac{\partial \mathbf{r}}{\partial z} = (0, 0, 1) = \mathbf{e}_z \quad (A37)$$

as our two tangent vectors, exactly as expected. Note that these vectors do *not* have the same units. Indeed, this issue of units is revealed in the metric tensor. The components of the metric tensor are obtained by computing the dot product of these tangent vectors and are given as

$$g_{\alpha\beta} \equiv \mathbf{a}_\alpha \cdot \mathbf{a}_\beta. \quad (A38)$$

For the specific case of a cylinder, the metric tensor takes the particularly simple form

$$\mathbf{g} = \begin{bmatrix} R^2 & 0 \\ 0 & 1 \end{bmatrix}, \quad (A39)$$

with much of the simplicity resulting from the fact that the tangent vectors are orthogonal.

With this background, we now want to solve the Toner-Tu equations,

$$\frac{\partial v^\mu}{\partial t} = [\alpha(\rho - \rho_c) - \beta g_{ab}v^a v^b]v^\mu - \sigma \nabla^\mu \rho - \lambda v^\nu \nabla_\nu v^\mu, \quad (A40)$$

for the case of the cylinder. Following the work of Shankar *et al.*, we neglect the diffusive (neighbor coupling) term, both to make an analytical solution tractable and to compare directly with their previous study. This formulation invokes the covariant derivative

$$\nabla_\mu v^\nu = \partial_\mu v^\nu + \Gamma^\nu_{\alpha\mu} v^\alpha, \quad (A41)$$

where $\Gamma^\nu_{\alpha\mu}$ are the Christoffel symbols. All of this extra machinery and corresponding notation comes down to the subtlety of getting our derivatives right in the curved-space setting. We also introduce the symbol Δ to represent the intrinsic curved-space version of the Laplacian. We assume that the Toner-Tu steady state,

$$\lambda v^\nu \nabla_\nu v^\mu = [\alpha(\rho - \rho_c) - \beta g_{ab}v^a v^b]v^\mu - \sigma \nabla^\mu \rho, \quad (A42)$$

has no velocity in the z direction. Equation (A42) corresponds to two equations to solve, one arising from $\mu = z$ and the other arising from $\mu = \theta$. In the steady state, we can evaluate the z equation with the result that

$$\frac{\partial \rho}{\partial z} = 0, \quad (A43)$$

implying a constant density. We can then evaluate the $\mu = \theta$ equation, leading to

$$\alpha(\rho_0 - \rho_c) - \beta R^2 (v^\theta)^2 = 0, \quad (A44)$$

and thus to

$$v^\theta_{ss} = \sqrt{\frac{\alpha(\rho_0 - \rho_c)}{\beta R^2}}. \quad (A45)$$

To find the actual magnitude of the velocity in physical units, we must contract with the metric tensor using

$$|\mathbf{v}_{ss}|^2 = g_{\mu\nu} v^\mu_{ss} v^\nu_{ss}. \quad (A46)$$

Further, since the only nonzero component of the velocity is v^θ_{ss}, this implies that $|\mathbf{v}_{ss}|^2 = g_{\theta\theta} v^\theta_{ss} v^\theta_{ss}$. Given that $g_{\theta\theta} = R^2$, we can now write the velocity as

$$|\mathbf{v}_{ss}|^2 = \frac{\alpha(\rho_0 - \rho_c)}{\beta}. \quad (A47)$$

This solution represents a simple circumferential flow around the cylinder coupled to a uniform density field.

It is one thing to obtain steady-state solutions. It is quite another to work out the time evolution of the solutions. We now revisit the Toner-Tu dynamics on the cylindrical surface for the case in which we prescribe a highly symmetric initial condition with the same symmetry of circumferential flow and uniform density ρ_0, and we work out the relaxation dynamics to the steady state that were shown in Fig. 10. For this geometry, the dynamical equation of interest is given by

$$\frac{\partial v^\theta}{\partial t} = [\alpha(\rho - \rho_c) - \beta R^2 (v^\theta)^2]v^\theta - \frac{\sigma}{R}\frac{\partial \rho}{\partial \theta} - \lambda v^\theta \frac{\partial v^\theta}{\partial \theta}. \quad (A48)$$

For the special case of high symmetry considered here, we know that both v^θ and ρ do *not* depend upon θ. Thus, the dynamical equation simplifies to the form

$$\frac{dv^\theta}{dt} = [\alpha(\rho - \rho_c) - \beta R^2 (v^\theta)^2]v^\theta. \quad (A49)$$

This nonlinear equation is separable resulting in the elementary integral

$$\frac{1}{\beta R^2}\frac{dv^\theta}{v^\theta\left(\frac{\alpha(\rho - \rho_c)}{\beta R^2} - (v^\theta)^2\right)} = dt. \quad (A50)$$

Given an initial magnitude $v(0) = Rv^\theta(0)$ and a uniform density ρ_0, the solution for the time-dependent relaxation is given by

$$|\mathbf{v}|^2 = R^2(v^\theta)^2 = \frac{\frac{\alpha(\rho_0 - \rho_c)}{\beta}}{1 - \left(1 - \frac{\alpha(\rho_0 - \rho_c)}{\beta v(0)^2}\right)e^{-2\alpha(\rho_0 - \rho_c)t}}. \quad \text{(A51)}$$

In Fig. 10, we compared the dynamics predicted here analytically to our FEM solution of the same problem.

4. Analytical solution for Toner-Tu on the sphere

We next take up the analysis of the minimal Toner-Tu model on the sphere. Previous work [39,53] described a highly symmetric rotating band solution for flocks on the sphere, in which both the density and velocity depend only upon the azimuthal angle θ at the steady state. Here, we present the analytical solution for density and velocity on the sphere derived by Shankar *et al.* [39], which is compared to our numerical results in Sec. V. For their steady-state analysis, Shankar *et al.* [39] made the approximation that the neighbor coupling term with coefficient D is absent, resulting in the simpler minimal version of the steady-state Toner-Tu equations of the form

$$\lambda v^\nu \mathbf{\nabla}_\nu v^\mu = [\alpha(\rho - \rho_c) - \beta g_{ab} v^a v^b] v^\mu - \sigma \mathbf{\nabla}^\mu \rho. \quad \text{(A52)}$$

We note that while Shankar *et al.* chose the field variables ρ and $\mathbf{p} = \rho \mathbf{v}$, we continue in the language of ρ and $\mathbf{v}$. To concretely solve these equations, the first step is the parametrization of the spherical surface, resulting in

$$\mathbf{r}(\theta, \phi) = (R \cos \phi \sin \theta, R \sin \phi \sin \theta, R \cos \theta). \quad \text{(A53)}$$

Next we need to find the tangent vectors which are defined via

$$\mathbf{a}_i = \frac{\partial \mathbf{r}}{\partial u^i}, \quad \text{(A54)}$$

where u^i is the coordinate of the surface defined above. Given these definitions, we have

$$\mathbf{a}_\theta = (R \cos \phi \cos \theta, R \sin \phi \cos \theta, -R \sin \theta) \quad \text{(A55)}$$

and

$$\mathbf{a}_\phi = \frac{\partial \mathbf{r}}{\partial \phi} = (-R \sin \phi \sin \theta, R \cos \phi \sin \theta, 0) \quad \text{(A56)}$$

as our two tangent vectors, exactly as expected. The components of the metric tensor are obtained by computing the dot product of these tangent vectors and are given as

$$\mathbf{g} = \begin{bmatrix} R^2 & 0 \\ 0 & R^2 \sin^2 \theta \end{bmatrix}, \quad \text{(A57)}$$

with much of the simplicity resulting from the fact that the basis vectors are orthogonal. In moving back and forth between covariant and contravariant components of our vectors of interest, we also need

$$g^{ab} = (g_{ab})^{-1} = \begin{bmatrix} \frac{1}{R^2} & 0 \\ 0 & \frac{1}{R^2 \sin^2 \theta} \end{bmatrix}. \quad \text{(A58)}$$

Equation (A52) corresponds to two equations we have to solve, one arising from $\mu = \theta$ and the other arising from $\mu = \phi$. To evaluate the covariant derivatives on the surface present in these equations, we will need the Christoffel symbols.

Generically, the Christoffel symbols are defined intuitively to tell us how our basis vectors change when we take small excursions with the parameters that characterize our surface of interest. Specifically, we have

$$\Gamma_{ij}^k = \frac{\partial \mathbf{a}_i}{\partial u^j} \cdot \mathbf{a}^k, \quad \text{(A59)}$$

measuring how an excursion in u^j leads to a change of $\mathbf{a}_i$. For the highly symmetric geometry of the cylinder considered earlier, the Christoffel symbols all vanish. We can see this by recognizing that the z direction and the θ direction are essentially uncoupled. That is, there is no dependence of $\mathbf{a}_\theta$ on the coordinate z and, similarly, there is no dependence of $\mathbf{a}_z$ on θ. Further, $\partial \mathbf{a}_\theta / \partial \theta$ is perpendicular to $\mathbf{a}^\theta$ and hence there is no contribution to the Christoffel symbol.

Evaluation of the covariant derivatives for the sphere requires us to invoke Christoffel symbols which, unlike the cylinder case, do not all vanish. Because of the symmetry of the sphere, only a few of the Christoffel symbols survive. Using the tangent vectors defined above, we have

$$\Gamma_{\phi\phi}^\theta = \frac{\partial \mathbf{a}_\phi}{\partial \phi} \cdot \mathbf{a}^\theta = -\sin \theta \cos \theta \quad \text{(A60)}$$

and

$$\Gamma_{\theta\phi}^\phi = \frac{\partial \mathbf{a}_\theta}{\partial \phi} \cdot \mathbf{a}^\phi = \cot \theta. \quad \text{(A61)}$$

With these geometrical preliminaries settled, we can now turn to the steady-state form of the dynamical equations themselves.

For the case when $\mu = \theta$, the Toner-Tu equations become

$$\lambda v^\nu \mathbf{\nabla}_\nu v^\theta = -\sigma \mathbf{\nabla}^\theta \rho \quad \text{(A62)}$$

since the term $[\alpha(\rho - \rho_c) - \beta g_{ab} v^a v^b] v^\theta$ vanishes because $v^\theta = 0$ as a result of our assumption that the solution only has a ϕ component of velocity. At first, we might be tempted to set the term $\lambda v^\nu \mathbf{\nabla}_\nu v^\theta$ equal to zero, but because of the covariant derivative, there is a contribution due to v_{ss}^ϕ. To see that, we note that the full term $\lambda v^\nu \mathbf{\nabla}_\nu v^\theta$ requires a covariant derivative and is given by

$$\lambda v^\nu \mathbf{\nabla}_\nu v^\theta = \lambda v^\nu \partial_\nu v^\theta + \lambda v^\nu \Gamma_{\alpha\nu}^\theta v^\alpha. \quad \text{(A63)}$$

The first term on the right vanishes because $v^\theta = 0$. The term $\lambda v^\nu \Gamma_{\alpha\nu}^\theta v^\alpha$ results in $\lambda v^\phi \Gamma_{\phi\phi}^\theta v^\phi = -\lambda \sin \theta \cos \theta (v_{ss}^\phi)^2$. A critical point necessary for evaluating $-\sigma \mathbf{\nabla}^\theta \rho$ is the relation between $\mathbf{\nabla}_\mu$ and $\mathbf{\nabla}^\mu$. Specifically, we note that

$$\mathbf{\nabla}^\mu = g^{\mu\nu} \mathbf{\nabla}_\nu. \quad \text{(A64)}$$

In light of this relationship, we have

$$\mathbf{\nabla}^\theta = g^{\theta j} \mathbf{\nabla}_j = g^{\theta\theta} \mathbf{\nabla}_\theta. \quad \text{(A65)}$$

Given that $g^{\theta\theta} = 1/R^2$ and $\mathbf{\nabla}_\theta = \partial/\partial\theta$, we can now write out the steady-state $\mu = \theta$ equation as

$$\lambda \sin \theta \cos \theta (v_{ss}^\phi)^2 = \frac{\sigma}{R^2} \frac{\partial \rho_{ss}}{\partial \theta}. \quad \text{(A66)}$$

For the case when $\mu = \phi$, we have

$$\lambda v^\nu \mathbf{\nabla}_\nu v^\phi = [\alpha(\rho - \rho_c) - \beta g_{ab} v^a v^b] v^\phi, \quad \text{(A67)}$$

where we have dropped the term $\sigma \nabla^\phi \rho$ since, by symmetry, we have *assumed* that the density does not depend upon the coordinate ϕ. The first term vanishes, resulting in

$$v_{ss}^\phi \left[\alpha(\rho_{ss} - \rho_c) - \beta R^2 \sin^2\theta \left(v_{ss}^\phi\right)^2 \right] = 0. \quad (A68)$$

We can now use these two conditions in our two unknowns $\rho_{ss}(\theta)$ and $v_{ss}^\phi(\theta)$ to solve for the steady-state values of density and velocity. Shankar *et al.* [39] began by defining a density $X(\theta)$ with a reference value as

$$X(\theta) = \rho_{ss}(\theta) - \rho_c. \quad (A69)$$

We can use this definition to solve Eq. (A68) for v_{ss}^ϕ as

$$\left(v_{ss}^\phi\right)^2 = \frac{\alpha X}{\beta R^2 \sin^2\theta}. \quad (A70)$$

This can now be substituted into Eq. (A66) to obtain the differential equation:

$$\frac{\sigma}{R^2} \frac{dX}{d\theta} = \frac{\lambda\alpha}{\beta R^2} \frac{\cos\theta}{\sin\theta} X. \quad (A71)$$

Multiplying both sides by R^2/σ results in the emergence of a dimensionless parameter given by

$$\eta = \frac{\lambda\alpha}{\beta\sigma}, \quad (A72)$$

permitting us to rewrite the equation for $X(\theta)$ as

$$\frac{dX}{d\theta} = (\eta\cot\theta)X. \quad (A73)$$

This differential equation is separable and upon integration yields the simple expression

$$\ln X = \eta\ln(\sin\theta) + C. \quad (A74)$$

By exponentiating both sides of the equation, we then find the solution

$$X(\theta) = A\sin^\eta\theta. \quad (A75)$$

The unknown coefficient A can be determined by exploiting the symmetry of the solution that tells us that the maximal density will be at the equator of the sphere, implying

$$X\left(\frac{\pi}{2}\right) = \rho_{ss}\left(\frac{\pi}{2}\right) - \rho_c = A\sin\frac{\pi}{2}, \quad (A76)$$

permitting us then to write the steady-state solutions for the density as

$$\rho_{ss}(\theta) = \rho_c + (\rho_0 - \rho_c)A_\eta \sin^\eta\theta, \quad (A77)$$

where the prefactor A_η is defined by

$$A_\eta = 2\Gamma((3+\eta)/2)/[\sqrt{\pi}\Gamma(1+\eta/2)]. \quad (A78)$$

Note that the dimensionless parameter η gives us a convenient knob to tune in our calculations, which allowed us to directly compare with the numerical results of our finite element calculations in Fig. 11.

We can also work out the steady-state value of the velocity field by invoking the condition

$$\left(v_{ss}^\phi\right)^2 = \frac{\alpha}{\beta R^2 \sin^2\theta} X(\theta). \quad (A79)$$

Recall that to find the actual magnitude of the velocity in physical units, we must contract with the metric tensor using

$$|\mathbf{v}_{ss}|^2 = g_{\mu\nu} v^\mu v^\nu. \quad (A80)$$

Further, since the only nonzero component of the velocity is v_{ss}^ϕ, this implies that $|\mathbf{v}_{ss}|^2 = g_{\phi\phi} v_{ss}^\phi v_{ss}^\phi$ Given that $g_{\phi\phi} = R^2 \sin^2\theta$, we can now write the velocity as

$$|\mathbf{v}_{ss}|^2 = \frac{\alpha}{\beta}(\rho_0 - \rho_c)A_\eta \sin^\eta\theta. \quad (A81)$$

For the specific case of $\eta = 2$, for example, we find

$$A_2 = \frac{2\Gamma\left(\frac{5}{2}\right)}{\sqrt{\pi}\Gamma(2)} = \frac{3}{2}. \quad (A82)$$

Using this value for A_2, we can then write the density and the velocity fields as

$$\rho_{ss}(\theta) = \rho_c + \tfrac{3}{2}(\rho_0 - \rho_c)\sin^2\theta \quad (A83)$$

and

$$|\mathbf{v}_{ss}|^2 = \tfrac{3}{2}(\rho_0 - \rho_c)\sin^2\theta, \quad (A84)$$

respectively. These solutions from the work of Shankar *et al.* [39] provided an opportunity to test our numerical implementation of the Toner-Tu theory on a curved surface using the FEM, as shown in Fig. 11 of Sec. V D.

[1] Pliny, *Natural History, Volume III: Book 10*, Translated by H. Rackham, Loeb Classical Library 353 (Harvard University Press, Cambridge, MA, 1940).

[2] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet, Novel Type of Phase Transition in a System of Self-Driven Particles, Phys. Rev. Lett. **75**, 1226 (1995).

[3] J. Toner and Y. Tu, Long-Range Order in a Two-Dimensional Dynamical XY Model: How Birds Fly Together, Phys. Rev. Lett. **75**, 4326 (1995).

[4] J. Toner and Y. Tu, Flocks, herds, and schools: A quantitative theory of flocking, Phys. Rev. E **58**, 4828 (1998).

[5] M. C. Marchetti, J. F. Joanny, S. Ramaswamy, T. B. Liverpool, J. Prost, M. Rao, and R. A. Simha, Hydrodynamics of soft active matter, Rev. Mod. Phys. **85**, 1143 (2013).

[6] M. Ballerini, N. Cabibbo, R. Candelier, A. Cavagna, E. Cisbani, I. Giardina, V. Lecomte, A. Orlandi, G. Parisi, A. Procaccini, M. Viale, and V. Zdravkovic, Interaction ruling animal collective behavior depends on topological rather than metric distance: Evidence from a field study, Proc. Natl. Acad. Sci. USA **105**, 1232 (2008).

[7] F. G. Woodhouse and R. E. Goldstein, Cytoplasmic streaming in plant cells emerges naturally by microfilament self-organization, Proc. Natl. Acad. Sci. USA **110**, 14132 (2013).

[8] M. Mayer, M. Depken, J. S. Bois, F. Jülicher, and S. W. Grill, Anisotropies in cortical tension reveal the physical basis of polarizing cortical flows, Nature **467**, 617 (2010).

[9] A. Cavagna and I. Giardina, Bird flocks as condensed matter, Annu. Rev. Condens. Matter Phys. **5**, 183 (2014).

[10] S. Münster, A. Jain, A. Mietke, A. Pavlopoulos, S. W. Grill, and P. Tomancak, Attachment of the blastoderm to the vitelline envelope affects gastrulation of insects, Nature **568**, 395 (2019).

[11] L. G. Pimpale, T. C. Middelkoop, A. Mietke, and S. W. Grill, Cell lineage-dependent chiral actomyosin flows drive cellular rearrangements in early *Caenorhabditis elegans* development, eLife **9**, e54930 (2020).

[12] H. Wioland, F. G. Woodhouse, J. Dunkel, and R. E. Goldstein, Ferromagnetic and antiferromagnetic order in bacterial vortex lattices, Nat. Phys. **12**, 341 (2016).

[13] A. Bricard, J. B. Caussin, N. Desreumaux, O. Dauchot, and D. Bartolo, Emergence of macroscopic directed motion in populations of motile colloids, Nature **503**, 95 (2013).

[14] N. Kumar, H. Soni, S. Ramaswamy, and A. K. Sood, Flocking at a distance in active granular matter, Nat. Commun. **5**, 4688 (2014).

[15] T. Sanchez, D. T. N. Chen, S. J. DeCamp, M. Heymann, and Z. Dogic, Spontaneous motion in hierarchically assembled active matter, Nature **491**, 431 (2012).

[16] A. Morin, N. Desreumaux, J. B. Caussin, and D. Bartolo, Distortion and destruction of colloidal flocks in disordered environments, Nat. Phys. **13**, 63 (2017).

[17] A. Souslov, B. C. Van Zuiden, D. Bartolo, and V. Vitelli, Topological sound in active-liquid metamaterials, Nat. Phys. **13**, 1091 (2017).

[18] K.-T. Wu, J. B. Hishamunda, D. T. Chen, S. J. DeCamp, Y.-W. Chang, A. Fernández-Nieves, S. Fraden, and Z. Dogic, Transition from turbulent to coherent flows in confined three-dimensional active fluids, Science **355**, eaal1979 (2017).

[19] S. Shankar, A. Souslov, M. J. Bowick, M. Cristina Marchetti, and V. Vitelli, Topological active matter, Nat. Rev. Phys. **4**, 380 (2022).

[20] B. Zhang, B. Hilton, C. Short, A. Souslov, and A. Snezhko, Oscillatory chiral flows in confined active fluids with obstacles, Phys. Rev. Res. **2**, 043225 (2020).

[21] M. Behrndt, G. Salbreux, P. Campinho, R. Hauschild, F. Oswald, J. Roensch, S. W. Grill, and C.-p. Heisenberg, Forces driving epithelial spreading in zebrafish gastrulation, Science **338**, 257 (2012).

[22] S. J. Streichan, M. F. Lefebvre, N. Noll, E. F. Wieschaus, and B. I. Shraiman, Global morphogenetic flow is accurately predicted by the spatial distribution of myosin motors, eLife **7**, e27454 (2018).

[23] Y. Maroudas-Sacks, L. Garion, L. Shani-Zerbib, A. Livshits, E. Braun, and K. Keren, Topological defects in the nematic order of actin fibres as organization centres of Hydra morphogenesis, Nat. Phys. **17**, 251 (2021).

[24] G. Cicconofri, M. Arroyo, G. Noselli, and A. DeSimone, Morphable structures from unicellular organisms with active, shape-shifting envelopes: Variations on a theme by Gauss, Int. J. Non-Linear Mech. **118**, 103278 (2020).

[25] D. J. G. Pearce, S. Gat, G. Livne, A. Bernheim-Groswasser, and K. Kruse, Programming active metamaterials using topological defects, arXiv:2010.13141.

[26] L. A. Hoffmann, L. N. Carenza, and L. Giomi, Tuneable defect-curvature coupling and topological transitions in active shells, Soft Matter **19**, 3423 (2023).

[27] C. L. Hueschen, L. Segev Zarko, J.-H. Chen, M. A. LeGros, C. A. Larabell, J. C. Boothroyd, R. Phillips, and A. R. Dunn, Emergent actin flows explain diverse parasite gliding modes, bioRxiv:10.1101/2022.06.08.495399.

[28] J. Toner, Y. Tu, and S. Ramaswamy, Hydrodynamics and phases of flocks, Ann. Phys. **318**, 170 (2005).

[29] M. J. Bowick, N. Fakhri, M. C. Marchetti, and S. Ramaswamy, Symmetry, Thermodynamics, and Topology in Active Matter, Phys. Rev. X **12**, 010501 (2022).

[30] L. E. Scriven, Dynamics of a fluid interface: Equation of motion for Newtonian surface fluids, Chem. Eng. Sci. **12**, 98 (1960).

[31] Y. Fily, A. Baskaran, and M. F. Hagan, Active particles on curved surfaces, arXiv:1601.00324.

[32] I. Nitschke, S. Reuter, and A. Voigt, Discrete exterior calculus (DEC) for the surface Navier-Stokes equation, *Advances in Mathematical Fluid Mechanics: Transport Processes at Fluidic Interfaces* (Birkhäuser, Cham, 2017).

[33] A. Sahu, R. A. Sauer, and K. K. Mandadapu, Irreversible thermodynamics of curved lipid membranes, Phys. Rev. E **96**, 042409 (2017).

[34] T. Jankuhn, M. A. Olshanskii, and A. Reusken, Incompressible fluid problems on embedded surfaces: Modeling and variational formulations, Interfaces Free Boundar. **20**, 353 (2018).

[35] A. Mietke, F. Jülicher, and I. F. Sbalzarini, Self-organized shape dynamics of active surfaces, Proc. Natl. Acad. Sci. USA **116**, 29 (2019).

[36] A. Torres-Sanchez, D. Millan, and M. Arroyo, Modelling fluid deformable surfaces with an emphasis on biological interfaces, J. Fluid Mech. **872**, 218 (2019).

[37] J. Toner, Why walking is easier than pointing: Hydrodynamics of dry active matter, *Active Matter and Nonequilibrium Statistical Physics*, Lecture Notes of the Les Houches Summer School, Vol. 112 (Oxford Academic, Oxford, 2022).

[38] L. P. Dadhichi, J. Kethapelli, R. Chajwa, S. Ramaswamy, and A. Maitra, Nonmutual torques and the unimportance of motility for long-range order in two-dimensional flocks, Phys. Rev. E **101**, 052601 (2020).

[39] S. Shankar, M. J. Bowick, and M. C. Marchetti, Topological Sound and Flocking on Curved Surfaces, Phys. Rev. X **7**, 031039 (2017).

[40] A. Mietke, Dynamics of active surfaces, Ph.D. thesis, TU Dresden, 2018, https://www.pks.mpg.de/fileadmin/user_upload/MPIPKS/group_pages/BiologicalPhysics/dissertations2/AlexanderMietke2018.pdf.

[41] M. Nestler, I. Nitschke, S. Praetorius, and A. Voigt, Orientational order on surfaces: The coupling of topology, geometry, and dynamics, J. Nonlinear Sci. **28**, 147 (2018).

[42] T. P. Fries, Higher-order surface FEM for incompressible Navier-Stokes flows on manifolds, Int. J. Numer. Methods Fluids **88**, 55 (2018).

[43] M. Nestler, I. Nitschke, and A. Voigt, A finite element approach for vector- and tensor-valued surface PDEs, J. Comput. Phys. **389**, 48 (2019).

[44] I. Nitschke, S. Reuter, and A. Voigt, Hydrodynamic interactions in polar liquid crystals on evolving surfaces, Phys. Rev. Fluids **4**, 044002 (2019).

[45] P. Rangamani, A. Agrawal, K. K. Mandadapu, G. Oster, and D. J. Steigmann, Interaction between surface shape and intra-surface viscous flow on lipid membranes, Biomech. Model. Mechanobiol. **12**, 833 (2013).

[46] S. C. Takatori and K. K. Mandadapu, Motility-induced buckling and glassy dynamics regulate three-dimensional transitions of bacterial monolayers, arXiv:2003.05618.

[47] T. Needham, *Visual Differential Geometry and Forms* (Princeton University Press, Princeton, NJ, 2020).

[48] https://github.com/RPGroup-PBoC/wildebeest_herds.

[49] N. A. Curtin, H. L. A. Bartlam-Brooks, T. Y. Hubel, J. C. Lowe, A. R. Gardner-Medwin, E. Bennitt, S. J. Amos, M. Lorenc, T. G. West, and A. M. Wilson, Remarkable muscles, remark-able locomotion in desert-dwelling wildebeest, Nature **563**, 393 (2018).

[50] Z. Wu, Counting wildebeest from space using deep learning, Ph.D. thesis, University of Twente, 2021.

[51] A. Morin and D. Bartolo, Flowing Active Liquids in a Pipe: Hysteretic Response of Polar Flocks to External Fields, Phys. Rev. X **8**, 021037 (2018).

[52] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevE.108.024610 for videos.

[53] R. Sknepnek and S. Henkes, Active swarms on a sphere, Phys. Rev. E **91**, 022306 (2015).

[54] R. D. Kamien, D. R. Nelson, C. D. Santangelo, and V. Vitelli, Extrinsic curvature, geometric optics, and lamel-lar order on curved substrates, Phys. Rev. E **80**, 051703 (2009).

[55] R. Green, J. Toner, and V. Vitelli, Geometry of thresholdless active flow in nematic microfluidics, Phys. Rev. Fluids **2**, 104201 (2017).

[56] COMSOL Multiphysics, The Finite Element Method (FEM), Multiphysics Cyclopedia (2017), https://www.comsol.com/multiphysics/finite-element-method.