



## BIOPHYSICS

# Scale-free correlations and potential criticality in weakly ordered populations of brain cancer cells

Kevin B. Wood<sup>1,2\*</sup>, Andrea Comba<sup>3,4</sup>, Sebastien Motsch<sup>5</sup>, Tomás S. Grigera<sup>6,7,8,9,10</sup>, Pedro R. Lowenstein<sup>3,4,11,12\*</sup>

Collective behavior spans several orders of magnitude of biological organization, from cell colonies to flocks of birds. We used time-resolved tracking of individual glioblastoma cells to investigate collective motion in an *ex vivo* model of glioblastoma. At the population level, glioblastoma cells display weakly polarized motion in the (directional) velocities of single cells. Unexpectedly, fluctuations in velocities are correlated over distances many times the size of a cell. Correlation lengths scale linearly with the maximum end-to-end length of the population, indicating that they are scale-free and lack a characteristic decay scale other than the size of the system. Last, a data-driven maximum entropy model captures statistical features of the experimental data with only two free parameters: the effective length scale ( $n_c$ ) and strength ( $J$ ) of local pairwise interactions between tumor cells. These results show that glioblastoma assemblies exhibit scale-free correlations in the absence of polarization, suggesting that they may be poised near a critical point.

## INTRODUCTION

Glioblastoma is the most aggressive brain cancer and remains the cancer with worst prognosis and shortest life expectancy. The standard of care treatment consists of resective surgery, radiotherapy, and chemotherapy. Long-term survival, nevertheless, has remained stagnant over the past 30 years despite major research efforts and multiple clinical trials (1, 2). Even when the glioblastoma tumor is resected, it always recurs, usually within a 1- to 2-cm margin of the original resection cavity. The tumor only rarely metastasizes to distant organs but invades the surrounding normal brain, destroying normal brain areas and disrupting brain function (3). Glioblastoma is the sole cancer that kills by direct invasion of surrounding normal brain tissue rather than through metastasis of distant organs. Understanding the mechanisms of glioblastoma growth and invasion are thus paramount to any future successful treatment of this disease.

Collective motion in brain tumors has been studied *in vitro* and *in vivo*. However, our understanding of the mesoscale organization of brain tumor collective motion has not been studied in sufficient detail. In particular, collective motility requires close behavioral coordination between individual cells (4–7). Such coordination could be implemented either by direct cell-to-cell contact, by close but indirect contact, or through long-range information exchange between cells (8, 9). The potential existence of such long-range

communication might allow tumor cells to respond quickly to a number of insults and thus provide robustness to tumor growth and progression. However, it is not clear whether such long-range correlations exist in brain tumors, and if so, whether these large-scale patterns actually arise from local interactions between nearby cells.

Collective motion arises in a large range of biological systems, from flocks of birds (10–14) to schools of fish (15, 16), from single cells (17, 18) and insects (19–23) to populations of mammals (24, 25), and has been intensely studied for decades in both natural systems (26–28) and robotics (29). Centralized, top-down mechanisms, for example, the presence of one or more “leaders” that dictate the behavior of the others (30, 31), could drive large-scale order, leading to synchronized motion between different cells or organisms. On the other hand, populations may exhibit emergent self-organization driven by local interactions between cells (26–28), leading to signatures of collective behavior even when motion appears only weakly polarized. Despite the apparent lack of order, self-organized systems can respond coherently to external perturbations, for example, the presence of a predator in animal flocks, because local interactions facilitate an exquisite sensitivity to the changing environment, effectively propagating information from one region of the flock to another in a manner unseen in top-down organizational schemes (10, 32, 33).

One clear signature of collective behavior is the presence of scale-free correlations, which have been observed in a number of biological systems (32, 34), including many exhibiting collective motion (10, 13). Correlations are considered scale-free when they lack a characteristic decay scale other than the size of the system. For example, in starlings, scale-free correlations in directional velocity occur in highly polarized flocks, where the distribution of velocity angles is extremely narrow (10, 12). Similar scale-free correlations arise in physical systems with a continuous symmetry, in this case, the rotational invariance of the velocity angle, where low-energy fluctuations give rise to Goldstone modes such as long-wavelength spin waves (35). By contrast, long-range correlations can also arise in the absence of order, for example, swarming midges exhibit scale-

<sup>1</sup>Department of Biophysics, University of Michigan, Ann Arbor, MI, USA.

<sup>2</sup>Department of Physics, University of Michigan, Ann Arbor, MI, USA.

<sup>3</sup>Department of Neurosurgery, University of Michigan, Ann Arbor, MI, USA.

<sup>4</sup>Rogel Cancer Center, University of Michigan Medical School, Ann Arbor, MI, USA.

<sup>5</sup>School of Mathematical and Statistical Sciences, Arizona State University, Tempe, AZ, USA.

<sup>6</sup>Instituto de Física de Líquidos y Sistemas Biológicos (IFLySiB), Buenos Aires, Argentina.

<sup>7</sup>Universidad Nacional de La Plata, La Plata, Buenos Aires, Argentina.

<sup>8</sup>CONICET, Godoy Cruz, Buenos Aires, Argentina.

<sup>9</sup>Departamento de Física, Universidad Nacional de La Plata, La Plata, Buenos Aires, Argentina.

<sup>10</sup>Istituto dei Sistemi Complessi, Consiglio Nazionale delle Ricerche, Rome, Italy.

<sup>11</sup>Department of Cell and Developmental Biology, University of Michigan, Ann Arbor, MI, USA.

<sup>12</sup>Department of Biomedical Engineering, University of Michigan, Ann Arbor, MI, USA.

\*Corresponding author. Email: kbwood@umich.edu (K.B.W.); pedrol@umich.edu (P.R.L.)

free correlations despite showing only weak levels of directional polarization (21). In physics, similar phenomena can occur when system parameters are tuned to a so-called critical point, where fluctuations become correlated across length scales and the system becomes exquisitely sensitive to small perturbations (35). Out of equilibrium, scale-free correlations can also appear as a consequence of a conservation law (36) or through a feedback mechanism (34, 37, 38). Recent work has argued for the existence of criticality in a wide range of biological systems (32, 39), including synchronization in neural ensembles (40, 41), brain activity (42), long-range speed correlations in starling flocks (13), the dynamics of biochemical networks (43) and protein folding (44), and the scale-free correlations in swarming midges (21). Independently of their origin, long-range correlations could potentially serve different biological functions. For example, in flocks of birds, collective behavior may have evolved to promote lossless information flow throughout the group as they move through space, while swarms of midges might use collective behavior to stabilize their behavior against environmental perturbations, such as a predator.

In this work, we quantitatively characterize collective motion in an *ex vivo* experimental model of glioblastoma using time-resolved tracking of individual glioma cells. We found that populations of glioma cells exhibit collective motion characterized by weak polarization in the (directional) velocities of single cells, suggesting, on the surface, that glioma cells may be moving largely independently. However, by examining populations of different sizes, we observed correlated fluctuations on scales many times the size of a single cell, with characteristic (correlation) lengths that increase linearly with the total system size, showing that they are scale-free. Last, we show that statistical features of the populations are well captured with a simple maximum entropy model, which reduces to the classical XY model with nonlocal coupling, which describes the shape of the directional velocity distributions and the existence of long-range correlations. Our results indicate that beneath a weakly ordered facade, brain tumor assemblies exhibit scale-free features of collective behavior on scales of millimeters or more.

## RESULTS

We analyzed glioma cell dynamics in a previously developed *ex vivo*, explant model derived from the orthotopic implantation of genetically engineered NPA–green fluorescent protein–positive (NPA–GFP<sup>+</sup>) cells, in which we can track the location and velocity of fluorescently labeled individual glioma cells for up to 2 or 3 days (Materials and Methods). Briefly, this experimental model enables *de novo* the induction of glioma tumors through the injection of different plasmids encoding (i) driver genes found in human gliomas and (ii) genes encoding for luminescent and fluorescent markers in postnatal day 1 (P01) wild-type C57BL/6 mice. When animals became symptomatic, tumors were removed, and neurosphere cell cultures were established as described earlier (45–47). Cells from these neurosphere cultures can then be implanted into adult C57BL/6 mice to reliably generate tumors (see Materials and Methods for details on the establishment of NPA tumors). Specifically, we implanted NPA cells into the brains of adult C57BL/6 mice (4, 45, 46). Animals were euthanized at day 19 after tumor implantation to generate explant brain tumor slices for time-lapse imaging on a laser scanning confocal microscope equipped with a tissue culture incubation chamber. We studied a collection of 12

different glioblastoma populations drawn from eight different explants: four from the inner tumor and four from the outer tumor region bordering the normal surrounding brain. We subdivided each explant into regions based on a previously developed classification scheme for glioblastoma cells (4). While this scheme was originally developed to classify subpopulations of cells based on histological and statistical properties of the cell orientations, in the present context, the subdivision scheme can be viewed as a box-like approximation for estimating finite-size scaling in the experimental system.

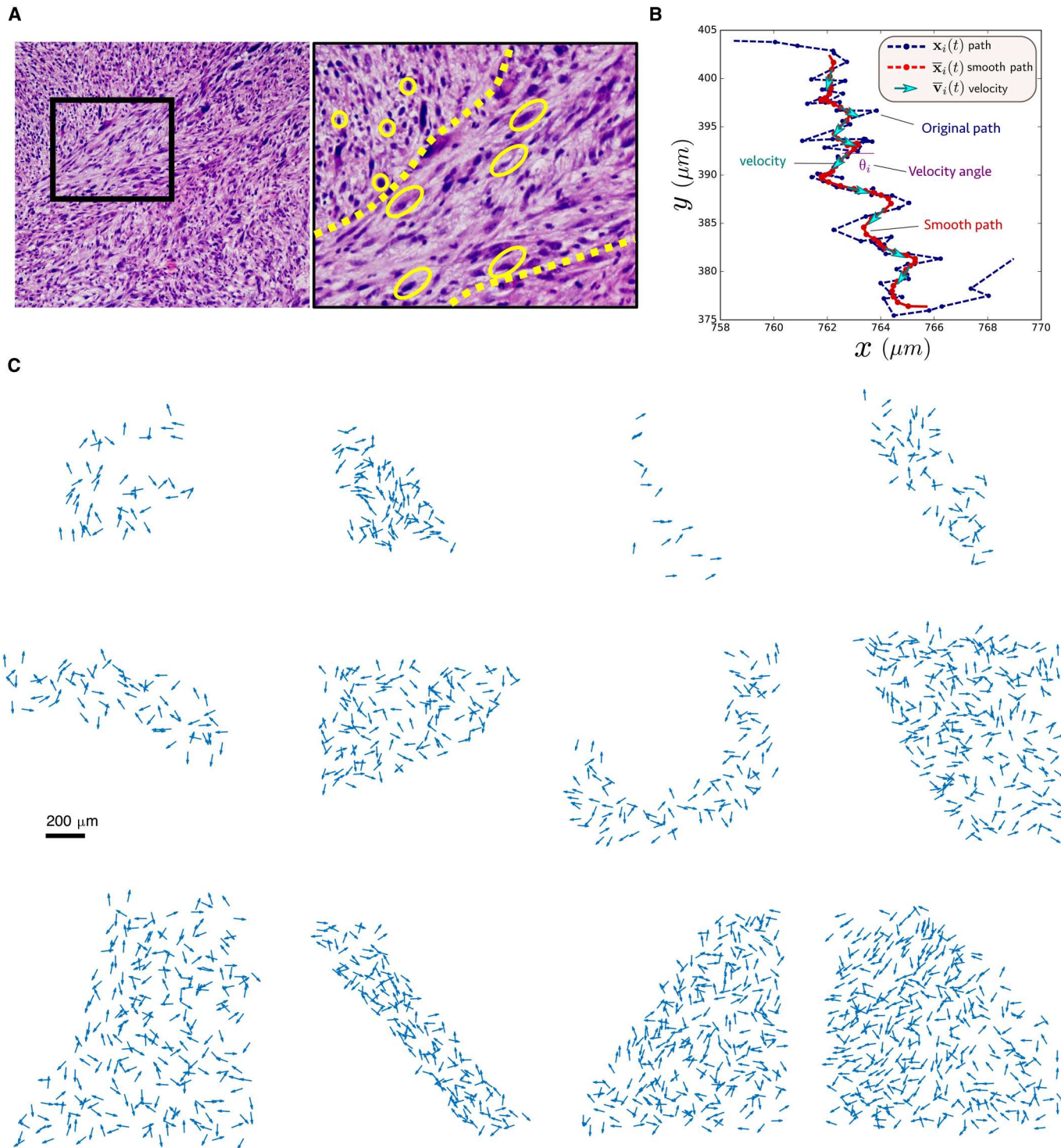
## Weakly ordered directional motion in glioblastoma populations

To quantify the population-level motion in the tumor, we estimated the position and velocity of each cell using semiautomated image analysis (Fig. 1; Materials and Methods). At each time point, the population is described by a set of normalized (unit) velocity vectors  $\{\mathbf{s}_i(t)\}$  and a corresponding set of position vectors  $\{\mathbf{x}_i(t)\}$ , one for each cell  $i = 1, 2, \dots, N$ . To quantify the degree of global ordering in the population, we calculated the polarization  $S(t)$ , which is defined as the magnitude of the population's mean velocity;  $S = 1$  when all cells move in the same direction, while  $S = 0$  if the velocities are uniformly distributed in all directions. We found that glioblastoma populations are weakly polarized, often exhibiting levels of polarization comparable to, or only slightly exceeding, those in size-matched populations with velocities randomly drawn from a uniform distribution (compare black and red curves; Fig. 2 and fig. S1). This weak polarization corresponds to broad but typically unimodal distributions of directional velocity (Fig. 2 and fig. S2). In contrast to starling flocks, which are highly polarized (10, 12), glioblastoma populations show minimal levels of polarization and therefore appear disordered at the population level. Such weak polarization could be evidence that cells move largely independently, with little functional coupling between cells; on the other hand, weak polarization, alone, is insufficient to rule out collective behavior. Populations weakly ordered on the global scale have been shown to exhibit features of long-range collective behavior in a number of biological contexts, including swarming behavior of midges (48) and synchronous firing activity in neural populations (49, 50). In these cases, fluctuations can be correlated over long spatial distances, even when ordering (i.e., polarization) is weak.

## Glioblastoma populations exhibit scale-free correlations in velocity fluctuations

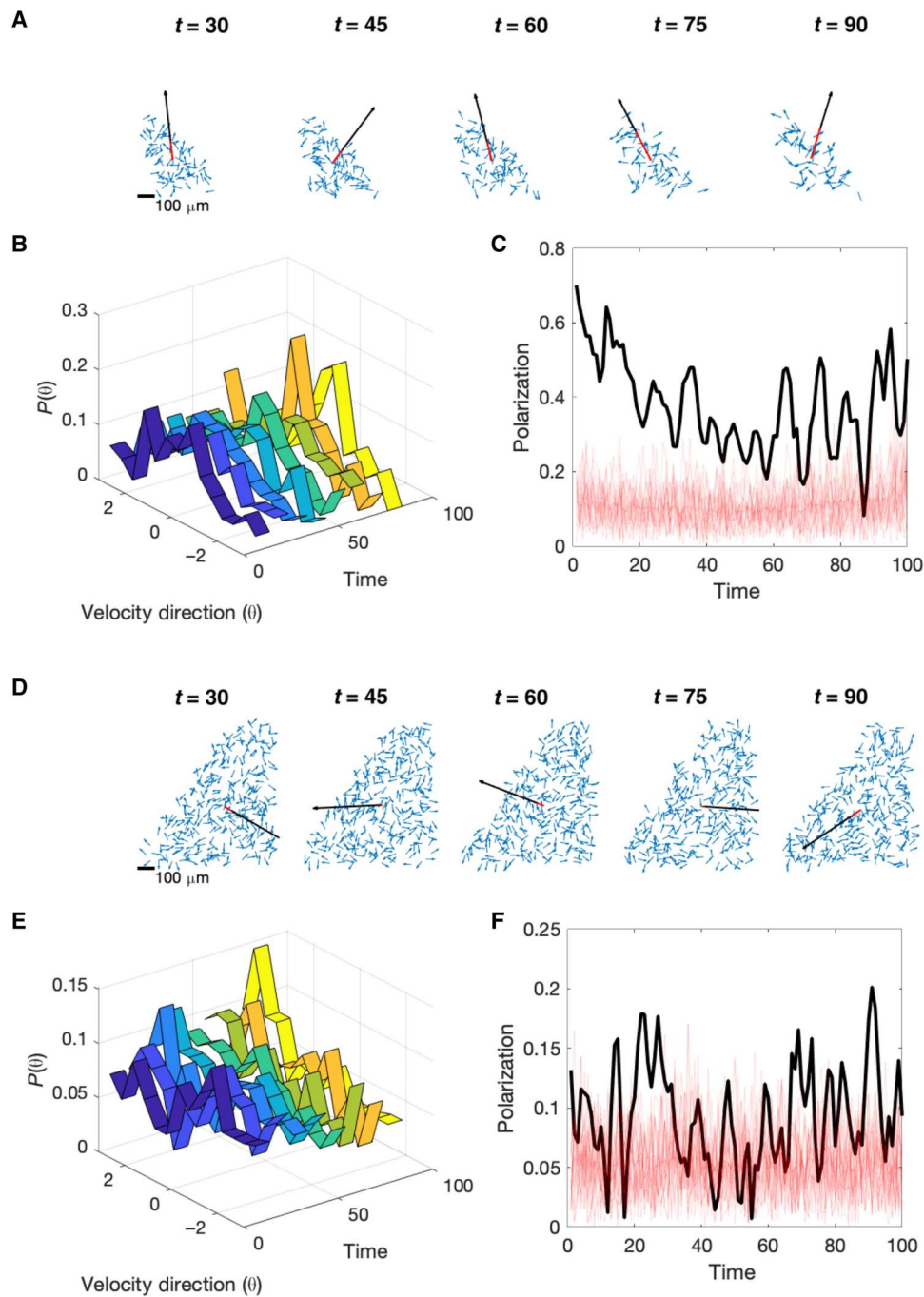
To determine whether glioma populations exhibit correlated spatial fluctuations, we calculated correlations between directional velocity for cells separated by a given distance ( $r$ ; in micrometers) for populations with different (time-averaged) spatial sizes ( $L$ ), where the size of the system is defined as the largest linear dimension of the population, that is, the maximum separation between any pair of cells. Correlation functions are defined by  $C(r) = \langle \delta \mathbf{s}_i \cdot \delta \mathbf{s}_j \rangle_r$ , where  $\delta \mathbf{s}_i \equiv \mathbf{s}_i - \langle \mathbf{s} \rangle$  is the velocity of cell  $i$  in the moving reference frame where the population average velocity ( $\langle \mathbf{s} \rangle$ ) has been subtracted out. Angle brackets  $\langle \rangle_r$  indicate an average taken over all cells separated by distance  $r$ .

We found that all populations exhibit weak local correlations over tens of micrometers (Fig. 3). To characterize the decay of these correlations over space, we computed  $r_0$  [i.e., the point at which  $C(r)$  first crosses 0] for populations of different sizes. We

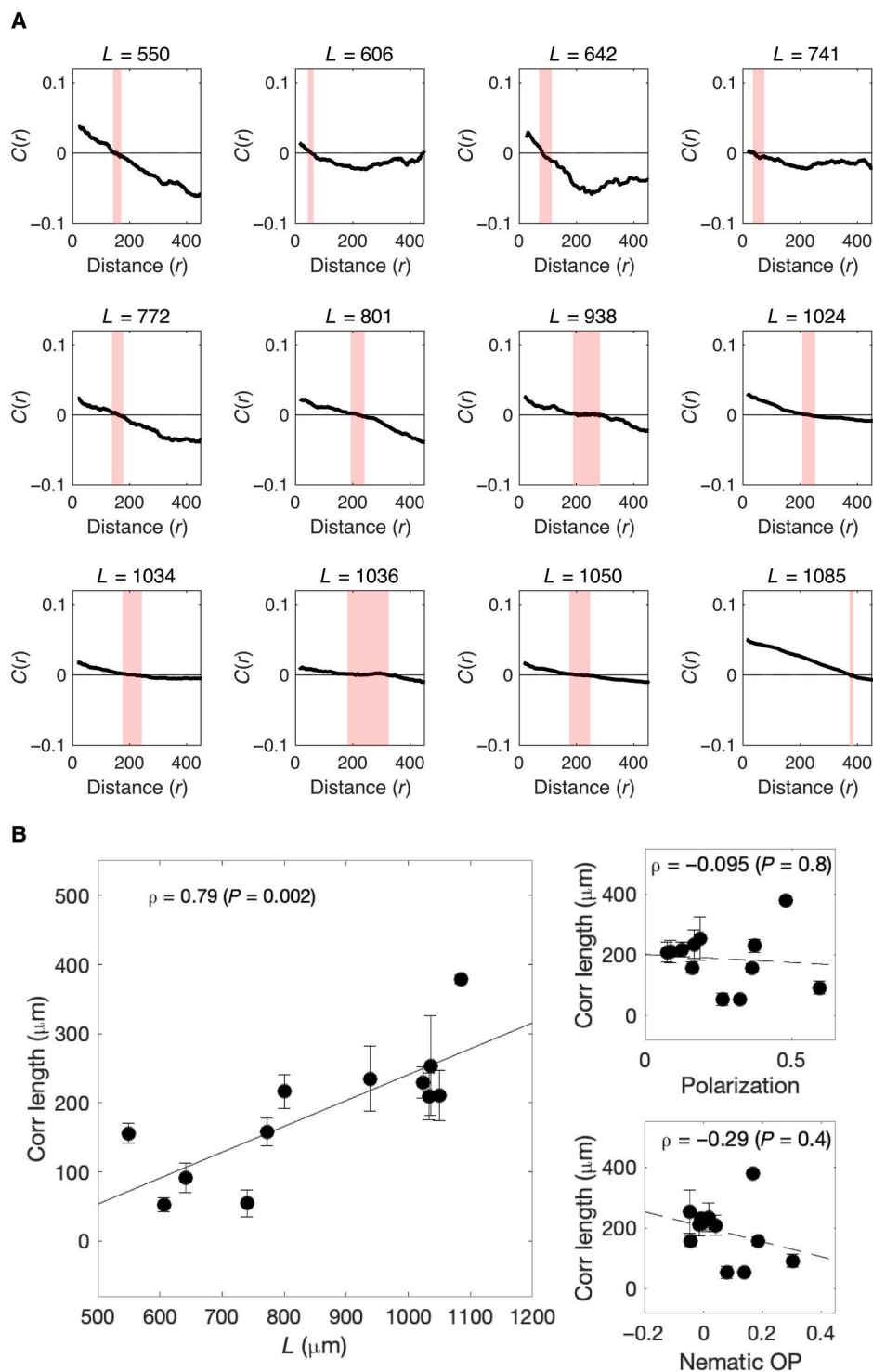


**Fig. 1. Tracking cell velocities in glioblastoma.** (A) Representative image of a 5- $\mu\text{m}$  hematoxylin and eosin–stained sections from an NPA mouse glioma tumor. The left image has been taken at lower power, and the black box is shown at higher power on the right. Oblong yellow outlines indicate elongated cells within an oncostream, which refers to a bundle of spindle-like cells that move throughout brain tumors (4, 7). Round yellow outlines indicate round cells located outside the oncostream. (B) Example of path obtained with ImageJ (blue dashed line) and the filtered path (red) obtained after filtering. The smooth path allows us to estimate the velocity of the cell at each time step. (C) Snapshot of unit velocity vectors in 12 different populations of different sizes.

Downloaded from <https://www.science.org> on September 18, 2024



**Fig. 2. Cell velocities exhibit weak directional order.** (A) Snapshot of unit velocity vectors (blue) and average velocity vector (red) at different time points. The black arrow is a reference vector indicating full alignment (polarization = 1); length of red vector relative to black vector indicates polarization. (B) Distribution of velocity directions over time (dark blue, early times; yellow, late times). (C) Polarization over time for glioblastoma populations (black) and for 25 simulated datasets (size-matched to actual data) but with velocity directions drawn from a uniform distribution (light red; mean over all datasets in dark red). Polarization is defined by  $S = \frac{1}{N} \sum_i \mathbf{s}_i$ , where  $|\mathbf{s}_i| \equiv 1$  and angle brackets indicate an average over all cells. (D to F) Identical to (A to C) but for a second dataset. See the Supplementary Materials for all datasets.



**Fig. 3. Velocity fluctuations are correlated over a length scale that depends on system size. (A)** Correlations between directional velocity for cells separated by a given distance ( $r$ ; in micrometers) for populations with different (average) spatial sizes ( $L$ ). Correlation functions are defined by  $C(r) = \langle \delta s_i \cdot \delta s_j \rangle_r$ , where  $\delta s_i \equiv s_i - \langle s \rangle$  is the velocity of cell  $i$  in the moving reference frame where the population average velocity ( $\langle s \rangle$ ) has been subtracted out. Angle brackets  $\langle \rangle_r$  indicate an average taken over all cells separated by distance  $r$ . Black curves: time-averaged correlations in a given population; red dashed line: estimated correlation length  $r_0$ , which corresponds to the crossover point  $C(r_0) = 0$ . **(B)** Main panel: Correlation length ( $r_0$ ) for subpopulations of different sizes (black circles; error bars represent uncertainty in estimate of crossover point). The solid line is best-fit line (for visualization). Right: Correlation length versus polarization (top) and nematic (bottom) order parameters for the populations in (A). The dashed line is best-fit line. Nematic order parameter is defined as  $\text{Nematic OP} = 2\langle \cos^2 \theta \rangle - 1$ , where  $\theta$  defines the orientation of the mean spin. Insets show Pearson correlation coefficient ( $\rho$ ) and corresponding  $P$  value.

found that  $r_0$  depends approximately linearly on the spatial size ( $L$ ) of each population. This linear relation is expected to hold when the correlation length is much larger than the size of the system [i.e., when the correlated fluctuations are scale-free (51)], and  $r_0$  can be interpreted as a correlation length scale; in contrast, when the correlation length is less than the system size,  $r_0$  grows as  $\log L$ . We note that similar scaling is observed when system size is defined using other metrics (fig. S4) and for different levels of smoothing (fig. S3). By contrast, order parameters, including both polarization and an order parameter for nematic ordering in 2D, are not correlated with the size of the system (Fig. 3).

While the scaling analysis above suggests that correlations in velocity fluctuations are scale-free, it is limited to the relatively small number of sample populations, whose sizes span approximately a factor of two, and correlation lengths (strictly,  $r_0$ ) estimated from these samples are inherently noisy. To complement this analysis, we performed a “box scaling” analysis (see Materials and Methods), a powerful approach to finite-size scaling built on subsampling data in randomly chosen spatial boxes of different sizes (52). To do so, we randomly selected many subpopulations of cells from all datasets using square boxes of a fixed width  $W$  placed randomly (with uniform distribution) within the field of view. Data from boxes of a given width were then combined and used to estimate correlation functions and crossover points  $r_0$  for these different-sized subsystems. This approach is advantageous because (i) repeated subsampling allows us to estimate  $C(r)$  and  $r_0$  more precisely by dividing data into a large number of smaller samples, (ii) the analysis does not depend on a priori division of datasets into subpopulations based on histological or other characteristics, and (iii) it allows us to explore scaling behavior over a wider range of system sizes. Previous studies have shown that the approach provides an excellent approximation to a full scaling analysis in both classical statistical physics models and real datasets, even when a wide range of system sizes is not available (52).

The box scaling analysis reveals a clear dependence of  $C(r)$  on the box size chosen for sampling (Fig. 4, top), and the characteristic length  $r_0$  depends linearly on the system size (Fig. 4, left). For comparison, we also show the plot on a log-linear scale (Fig. 4, right), which allows us to rule out logarithmic scaling with system size ( $r_0 \sim \log L$ ) that would indicate a correlation length smaller than the system size (52). The results are similar if we use alternative measures of system size such as the linear length of the boxes used for scaling (fig. S6). As with the original datasets, the box scaling data exhibits only weak ordering as measured by polarization or nematic order parameters (fig. S6). Together, these findings indicate that glioma populations exhibit collective behavior, specifically, scale-free correlations, in the absence of strong ordering.

### A data-driven maximum entropy model captures statistical features of glioma migration

To develop a minimal effective model of glioma collective motion, we used the experimental velocity data to parameterize a simple maximum entropy model (Fig. 5A). Maximum entropy methods have been widely applied to model biological phenomena (32, 53), including the collective firing activity of neurons (49, 50); the flocking behavior of birds (12, 13); and correlations in antibody diversity (12), drug interactions (54), or sequence motifs in biological polymers (55). Maximum entropy approaches are closely connected to classical “inverse problems” in statistical physics, which involve

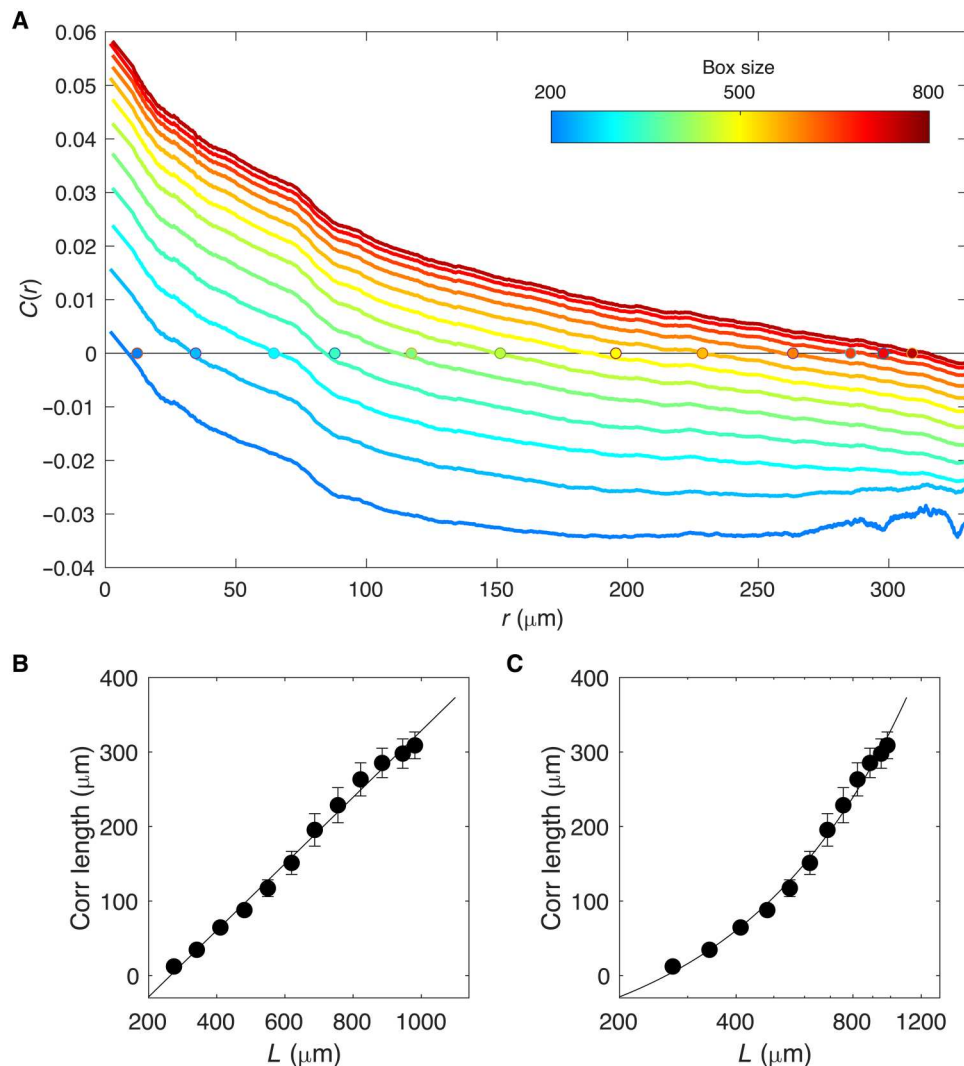
estimating (often unobservable) microscopic parameters from a set of macroscopic observables (56–59). The maximum entropy approach allows one to incorporate a specific set of experimental observables into a model but, in a strict statistical sense, does not introduce additional structure. At its core, a maximum entropy model is consistent with a defined set of measurements but is otherwise as “unbiased” as possible. In the case of tumor cell populations, we would like to determine whether a minimal model that incorporates pairwise interactions of a fixed length scale is sufficient to reproduce the large-scale order observed in experiments. Following (12), we developed a maximum entropy model consistent with the local correlation structure measured in experiments, a quantity that we refer to as  $C_{\text{int}}$ .  $C_{\text{int}}$  is a single number that describes the average correlation of a cell’s velocity with that of its  $n_c$  closest neighbors; if cells tend to be directionally aligned with their neighbors, then  $C_{\text{int}}$  approaches 1, while  $C_{\text{int}}$  is 0 if directional velocities are random.

The maximum entropy model consistent with  $C_{\text{int}}$  is formally identical to a nonlocally coupled version of the XY model, which was originally developed in statistical physics to describe systems such as superconductors characterized by a continuous (planar) symmetry (35). The model contains two free parameters: the length scale over which cells interact ( $n_c$ ) and the strength of those interactions ( $J$ ). Our goal is to determine whether such a minimal model, containing only two scalar parameters, can reproduce measured features of the cell ensembles.

The model itself is probabilistic, that is, it specifies the probability of observing a particular set of velocities given any specific choice of the parameters  $J$  and  $n_c$ . In physics parlance,  $J$  and  $n_c$  are “microscopic parameters” that describe physical interactions between the constituent parts of the system (e.g., a parameter analogous to  $J$  is used to describe the interactions between individual magnetic dipoles in a ferromagnet). Particular choices of these parameters give rise to different types of “macroscopic” behavior, quantities, like polarization, that reflect measurable properties of the system as whole. The mapping from microscopic parameters ( $J, n_c$ ) to macroscopic observables (polarization and correlation functions) is often not straightforward, although statistical physics provides a number of tools for calculating the latter from the former.

We estimated parameters ( $J, n_c$ ) for each population at each time point using a spin-wave approximation to calculate the partition function  $Z$  (see Materials and Methods). We note that while the spin-wave approximation is strictly valid only in highly polarized systems, we found that the model with these parameter estimates qualitatively captures many features of our data. Different populations are characterized by widely variable (time-averaged) values of  $n_c$ , with length scales ranging from tens of cells, in some cases, to hundreds of cells in others, nearly the size of the entire population (Fig. 5B). To compare the model with experimental data, we used Monte Carlo simulations to estimate the distribution of velocity angles (Fig. 5C) and correlation functions (Fig. 5D) for the model with experimentally determined values of  $J$  and  $n_c$ . Despite the simplicity of the model, it captures qualitative features of both the angular distribution and the correlation functions for nearly all populations without additional tuning of free parameters (Fig. 5; see also figs. S7 and S8).

Because the model suggests that cells are coupled over large length scales, with  $n_c$  sometimes approaching the size of the population, we also consider a mean field (all-to-all coupled) version of



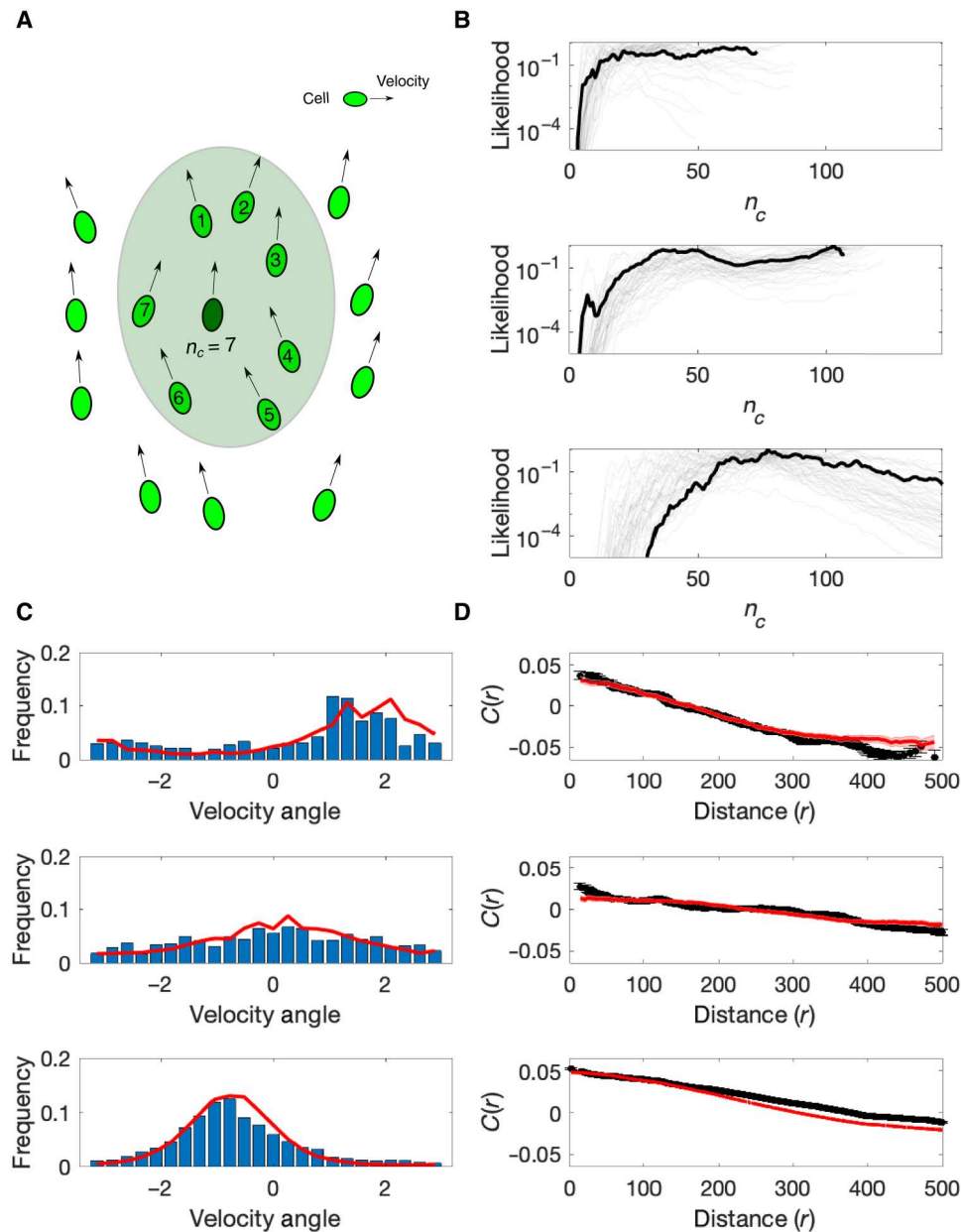
**Fig. 4. Box scaling indicates correlations in velocity fluctuations are scale-free.** (A) Correlations between directional velocity for cells separated by a given distance ( $r$ ; in micrometers) for subsampled populations in boxes of different sizes [ranging from 200  $\mu\text{m}$  by 200  $\mu\text{m}$  (blue) to 800  $\mu\text{m}$  by 800  $\mu\text{m}$  (red)]. Filled circles: estimated correlation length  $r_0$ , which corresponds to the crossover point  $C(r_0) = 0$ . (B) Correlation length ( $r_0$ ) for subsampled populations of different sizes (black circles) on linear-linear (B) and log-linear (C) scales. Error bars represent uncertainty in estimate of crossover point. Solid line is best-fit line (for visualization).

the model. In this case, the model is characterized by a single parameter ( $J$ ), and estimating  $J$  from experimental data reduces to a classic inverse problem in statistical physics. The advantage of this simpler model is that parameter estimation no longer relies on the spin-wave approximation, and instead, the maximum likelihood estimate of  $J$  can be calculated analytically in terms of the experimental observable  $C_{\text{int}}$  (Material and Methods). Somewhat unexpectedly, the mean field model captures the angular distribution of velocities nearly as well as the full model (fig. S9). As a numerical control, we confirmed that such agreement does not occur in randomly generated datasets of size-matched populations, confirming that the qualitative agreement between the model and experimental data is unlikely to arise from finite-size statistical effects (fig. S11).

### Effective model of glioma dynamics is poised at a critical point

Glioma cells exhibit scale-free correlations in the absence of polarization. One explanation for this observation is that the system is poised near a critical point. We therefore asked whether the effective maximum entropy model is also poised near a critical point: that is, whether the experimentally derived parameters ( $J$ ,  $n_c$ ) correspond to a critical point of a model. The absence of criticality in such a model may point to alternative explanations for the experimental observations.

To characterize the phase diagram for the maximum entropy model (i.e., an  $XY$  model with nonlocal coupling determined by  $J$  and  $n_c$ ), we used Monte Carlo simulations to generate representative velocity distributions for different values of  $n_c$  and  $J$ . For each parameter pair, we calculated the polarization order parameter (the mean polarization across trials), the generalized susceptibility  $\chi$ , which corresponds (up to a proportionality constant) to the

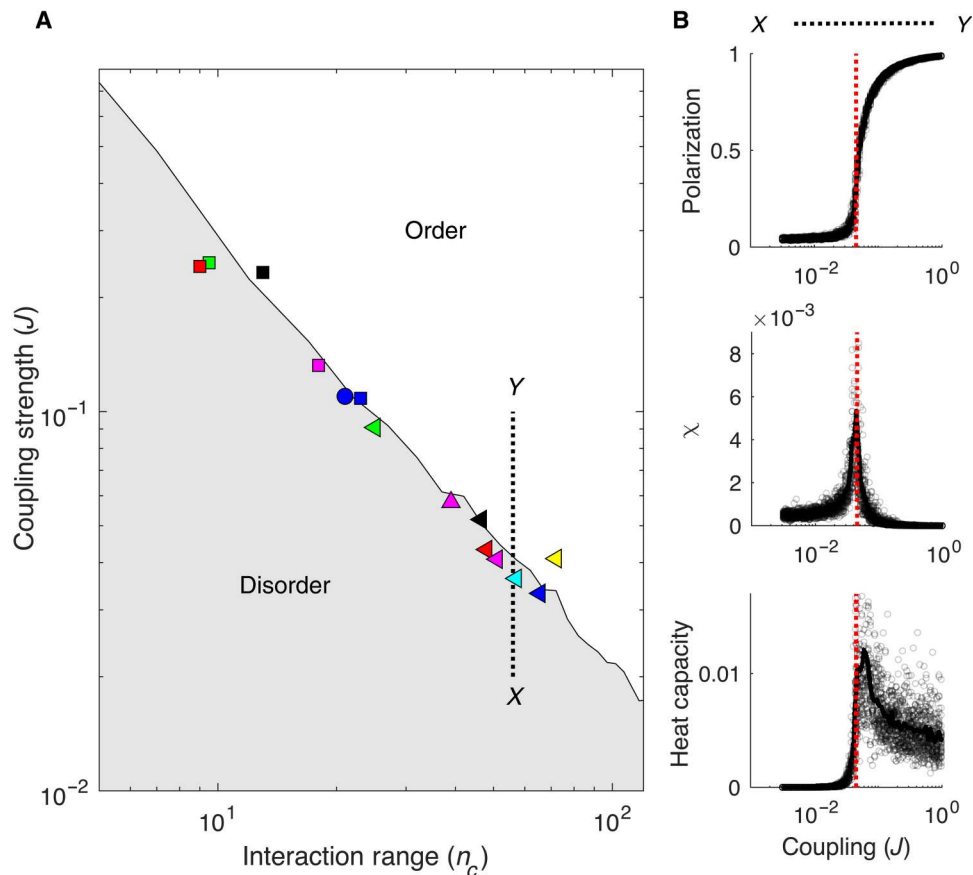


**Fig. 5. Maximum entropy model captures statistical features of glioblastoma populations.** (A) Moving cells (ovals, with arrows representing velocity) interact pairwise with the closest  $n_c$  neighbors in their local community. Shaded region indicates range of interaction for a single focal cell (dark green) for an illustrative value of  $n_c = 7$ . A coupling parameter  $J$  indicates the strength of interaction between pairs of cells, with larger  $J$  favoring stronger directional alignment. (B) Likelihood (normalized) for the interaction range ( $n_c$ ) for three different populations; light curves show individual time points, and dark curves are averages over time. (C) Histograms of velocity direction (angle) for experiment (blue) and the maximum entropy model (red) for the same three populations as in (B). (D) Correlation functions calculated from experimental data (black) and for the maximum entropy model (red) for the same three populations as in (B). See also figs. S7 and S8 for full results on all populations.

variance of the polarization, and the heat capacity, which measures fluctuations in the effective energy. We then estimated the critical surface in the  $(J, n_c)$  plane to be the curve in parameter space that corresponds to a peak in  $\chi$ . We also verified that the system exhibits sharply increasing polarization and a peak in the generalized heat capacity at the critical surface. For all simulations, cell positions (and, therefore, cell density) were taken from a representative experimental dataset.

The phase diagram (Fig. 6A) is divided by the critical surface into an ordered region (upper right) and disordered region (lower left, gray). To investigate whether the experimental populations are poised near the critical surface, we plotted the estimated parameter values ( $J$  and  $n_c$ ) for each population (triangles, squares, and circles) on the phase diagram. We find that all populations are characterized by parameter pairs that lie close to the critical surface. Modulating the parameters, for example, by simulating a system with a fixed (experimentally determined) value of  $n_c$  but  $J$  values that differ from





**Fig. 6. Experimental estimates of cell-cell coupling parameters are poised near critical point of nonlocally coupled XY model.** (A) Phase diagram shows regions of disorder (gray) and order (white) for different values of the coupling strength ( $J$ ) and interaction range ( $n_c$ ). Phase boundary was estimated from Monte Carlo simulations of the maximum entropy model (Materials and Methods). Markers (squares, triangles, and circles) show estimates of  $J$  and  $n_c$  from different glioblastoma populations. The dashed black line shows an example trajectory through parameter space, in this case, from a point  $X$  in the disordered region to a point  $Y$  in the ordered region, that crosses the critical surface by increasing  $J$  at a fixed value of  $n_c$ . (B) Polarization (top), generalized susceptibility  $\chi$  (middle), and generalized heat capacity (bottom) calculated from Monte Carlo simulations of the model as parameters are varied along the dashed trajectory connecting points  $X$  and  $Y$  in (A). Gray circles represent Monte Carlo simulations across a range of coupling strengths ( $J$ ) at fixed  $n_c = 58$  (corresponding to the estimated  $n_c$  for one representative dataset); solid lines are averages over a total of 100 simulations. The dashed red line indicates estimated value of coupling  $J$  for one particular dataset [(green triangle in (A))] that lies near the critical surface. Susceptibility ( $\chi$ ) and heat capacity are defined as the variance of polarization and energy, respectively, across independent realizations.

subcritical to supercritical (as indicated by the dashed line connecting points  $X$  and  $Y$ ; Fig. 6A), indicates that the parameters estimated from experiments occur near a critical point, where polarization begins to rapidly increase and  $\chi$  and the heat capacity peak (Fig. 6B).

We stress, however, that criticality in this model alone is insufficient to infer criticality in the glioma populations, both because of technical limitations in distinguishing critical and subcritical systems (see the Supplementary Materials) and, more importantly, because this model may not correspond to the precise physical mechanisms underlying the motion. Nevertheless, the maximum entropy model accurately captures qualitative behavior of the glioma populations and, in doing so, provides a parsimonious effective model for the data. While the overall data suggest that glioma tumors exhibit scale-free correlations in the absence of ordering, which is compatible with criticality, further experiments will be needed to determine brain tumor criticality beyond doubt.

## DISCUSSION

Our data suggest that the brain tumors exhibit features of large scale collective behavior in the absence of strong order. One potential explanation is that glioma populations are poised near a critical point, that is, on the boundary between an ordered and disordered phase. A growing number of studies suggest that biological systems may operate at or near criticality (13, 32, 34, 40–44, 60–62). However, scale-free correlations can also be found in some systems that are noncritical. For example, scale-free correlations have been observed in strongly ordered animal flocks (12, 26), and similar behavior is expected in equilibrium systems with a continuous symmetry (63). This explanation does not apply to glioblastoma cells, however, because they are largely disordered (the order parameter of the sub-systems is rather small; Fig. 2).

Out of equilibrium, it is possible to build models that display long-range, scale-free correlations without a tuning parameter, for example, in systems with a conservation law (36) or in self-organized criticality (SOC) models (34, 37, 38), where a feedback mechanism keeps the system at the edge between a state with high but

fluctuating local activity and an absorbing inactive state. In the case of glioma cells, we cannot see such SOC-like mechanisms obviously at work, although these or other mechanisms might ultimately underlie the scale-free behavior that we observe.

What is the potential advantage of such “disordered” collective behavior? Past work suggests that these properties confer a number of potential biological advantages, including increased sensitivity to environmental stimuli, a large dynamic range, and improved capacity to store and process information (39). In the case of glioma cells, these populations may be able to rapidly access a larger set of behaviors when faced with growth or chemotherapy insults. Strongly ordered motion imposes a rigid structure that limits the sensitivity of the population to new environmental insults, such as attacks from the immune system. On the other hand, cells that move entirely independently and lack the ability to move collectively, might represent cells best situated to invade surrounding normal tissue and could also facilitate immune escape. The ideal tumor state may therefore be one that exhibits features of both order and disorder, where tumors may be poised to modulate their capacity for motility, as dictated by clues of the tumor microenvironment.

It is important to keep in mind several limitations of our study, both technical and conceptual. First, the experiments are performed in an *ex vivo* model, which allowed us to measure cell velocities and track movement for up to 72 hours. As always, however, one should use caution when extrapolating from laboratory results to *in vivo* dynamics, although our recent data have demonstrated similar dynamics using multiphoton intravital imaging of glioma (4). In addition, our use of a maximum entropy approach comes with important caveats. While the model captures many statistical properties of the underlying experimental data, it should be viewed as an effective model, not necessarily a true physical representation of the underlying system. For example, the model suggests that pairwise interactions between cells are long-range, often extending across tens of cells or more, but there is no guarantee that the true biological interactions, whose mechanisms are, at this stage, largely unknown, extend over this scale. Instead, long-range interactions in the model represent effective interactions that appear artificially long range because (for example) the local cell dynamics have not reached equilibrium on the time scales of our measurements (64, 65), meaning that the effective interactions represent many fast physical interactions between different pairs of cells. It is also worth stressing that the finding of criticality in the data-driven model is not, on its own, proof of criticality in the glioblastoma populations [it is well known that quantitative maximum entropy predictions do not necessarily reflect similar features in the underlying system (39, 66–71)]. In any case, our primary claim, that fluctuations are scale-free, is based on direct the observation of correlations of the experimental system not on any particular model. We stress that these long-range correlations do not imply direct long-range physical cell-cell interactions.

In our glioblastoma models, growth is not constrained by the pattern of brain anatomy [i.e., white matter bundles and white matter/brain matter borders (72)], as growing rodent tumors replace brain tissue. Thus, in our model, the formation of oncostreams does not depend on the underlying brain structure (73); it is unclear whether they depend on the mechanics of surrounding tissue (74, 75). The role of blood vessels in glioblastoma growth and response to treatment, as well as their potential role in oncostream dynamics, remains to be determined (76, 77).

There are also a number of technical limitations that bear mention. We focused on a model with topological (rather than metric) connections between cells. For these populations, where the densities of cells are typically similar across populations, we expect that both approaches would yield similar results, although imposing a different coupling structure could improve the agreement between model and experiment (10). In addition, we have used free boundary conditions, partly because confining the boundary cells (using fixed boundary conditions) would, in some cases, reduce the number of cells for analysis. Previous work in starling flocks showed that these boundary conditions can make a difference in the inferred values of  $J$  and  $n_c$ , although their product remains largely constant (12). Correcting for these boundary conditions might provide a more accurate fit to, for example, simulated systems with known control parameters. In this case, however, our goal was not to precisely estimate experimental parameters, and because the model captures experimental results reasonably well, we have not further investigated these issues.

Our results raise a number of open questions for future work. From a physics perspective, more detailed physical models may shed additional light on quantitative features of glioma migration that are not captured by the simple model used here. For example, we recently showed that a model of collective motion that incorporates cell shape can produce a rich collection of phenomena that includes nematic ordering and qualitatively distinct types of disordered motion (7) similar to that seen in glioma populations (4). Perhaps, our results imply the existence of (effective) long-range communication within glioma tumors, even if the physical basis for this communication is not yet understood. Previous work indicates that natural insect swarms exhibit long-range correlations (21), but these correlations seem to appear in response to different environmental perturbations in laboratory populations (23, 78). Similarly, it may be possible to investigate these mechanisms in glioma populations using controlled laboratory conditions that strip away complexities of the *ex vivo* tumor environment. From a biological perspective, several mechanisms exist that could play a role in transmitting signals across large portions of glioma (79). For example, cytokines or neurotransmitters could be released by glioma cells, diffuse throughout the tumor tissue, and thus alter the behavior of distantly located cells. Equally, networks of microtubes connecting glioma cells have recently been described and shown to provide signals to other cells via calcium ion transients (80). So far, only shorter distance communication has been shown, but if microtubes are truly functional across larger distances, then they could be an anatomo-physiological substrate of long-range communication described herein. Furthermore, criticality could help understand and explore the recalcitrant robustness of these tumors when exposed to therapeutic modalities such as x-rays and chemotherapy. Criticality could play a role in glioma cells responding to resective surgery; for example, no matter how much tumor is resected, cells located within 1 to 2 cm from the resection cavity eventually reconstitute the tumor. Thus, it might be possible that low-density cells might need to grow to a certain population size before long-range communication can be exploited to support cell replication and especially cell invasion. Outcomes in the response to chemo- and radiotherapy could be tested experimentally by exposing brain slices to these treatments and then evaluating collective cell motility using *ex vivo* or *in vivo* imaging (4). Last, a role for the recently discovered brain innervation of tumors

(80) could be to directly connect distant parts of the tumor, i.e., tumor cells could signal to innervating neurons that, through neural networks, would then signal back to distant tumor regions. We believe that the physical and mathematical modeling of glioblastomas has the potential to inspire further understandings of glioblastoma biology and potentially unique treatments aimed at destabilizing collective behavior.

## MATERIALS AND METHODS

### Time-lapse confocal imaging in explant brain glioma slice culture model

To analyze glioma cell dynamics, we used a brain tumor explant model for ex vivo imaging. In vivo experimentation was performed following the procedures approved by the Institutional Animal Care and Use Committee (IACUC) at the University of Michigan. Glioma tumor cells were obtained from a genetically engineered mouse glioma model generated in our laboratory as previously described (Sleeping Beauty Transposase System) (45–47). This model enables de novo generation of glioma tumors through the injection of different plasmids encoding the genes of interest in postnatal day 1 (P01) wild-type C57BL/6 mice. The plasmid sequences used to generate these tumors were (i) the transposase and luciferase enzyme expression (pT2C-LucPGK-SB100X), (ii) the NRAS gene expression (pT2CAG-NRAS-G12V), (iii) the short hairpin for down-regulation of p53 protein (pT2-shp53-GFP4), and (iv) the short hairpin for ATRX (pT2-shATRx-GFP4). These glioma tumor cells named NPA exhibit the overexpression of the NRAS protein, the knock-down of p53, and the down-regulation of ATRX. shp53 and shATRx plasmids also contain the gene for the expression of enhanced GFP (EGFP). These signaling pathways are altered in human gliomas. Cells were generated from genetically engineered glioma tumors and cultured in Dulbecco's minimum essential medium (DMEM)/F12 media supplemented with 2% B-27, 1% N-2, 1% of penicillin-streptomycin, 0.2% of Normocin, and growth factors human Fibroblast Growth Factor (hFGF) and human Epidermal Growth Factor (hEGF at 20 ng/ml and maintained in a 37°C incubator supplied with 95% air and 5% CO<sub>2</sub>. Tumors were induced by intracranial implantation of  $3 \times 10^4$  NPA cells (NRAGV12, shp53, and shATRx) glioma tumor cells in C57BL/6 mice. Mice were held in a pathogen-free, humidity- and temperature-controlled vivarium on a 12-hour light/12-hour dark cycles with free access to food and water following. Before implantation, mice were anesthetized using an intraperitoneal injection of the anesthetics ketamine (120 mg/kg) and dexmedetomidine (0.5 mg/kg). Following anesthesia, carprofen (5.0 mg/kg) was administered subcutaneously for pain management. The skull of the mouse was then immobilized in a stereotactic device. A burrhole was made using a 0.45-mm drill bit at coordinates corresponding to the striatum (0.5 mm anterior and 2 mm lateral from bregma). Cells were injected with a Hamilton syringe at a dorsoventral position of 3.5 mm ventral into the striatum. After injection, the incision was sutured and immediately following surgery, the animals were recovered from anesthesia using atipamezole via intraperitoneal injection (1.0 mg/kg) to reverse the effect of dexmedetomidine. A single subcutaneous injection of buprenorphine (0.01 mg/kg, subcutaneous) was administered as postoperative pain relief. Sutures were removed 10 days after surgery (4, 45).

Tumors were allowed to grow for 19 to 21 days. At 19 to 21 days after implantation, animals were euthanized to generate the brain tumor slice explants for imaging. To generate brain tumor explants, brains were embedded in 4% low melting temperature agarose and kept on ice until solidification. Embedded brains were then immersed in ice-cold and oxygenated DMEM without phenol red and then sectioned in a Leica VT100S vibratome (Leica, Buffalo Grove, IL). Brain tumor sections (300 μm thick) were transferred to laminin-coated cell culture insert (Millipore Sigma, USA) placed into a 27-mm-diameter dish (Thermo Scientific) with DMEM/F12 media supplemented with 25% fetal bovine serum and penicillin-streptomycin. All steps were performed under sterile conditions in a BSL2 laminar flow hood. Tumor slices were then maintained in a cell culture incubator at 37°C with a 5% CO<sub>2</sub> atmosphere. After 6 to 18 hours, medium was replaced with DMEM/F12 media supplemented with 2% B-27, 1% N<sub>2</sub>, 0.2% Normocin, penicillin-streptomycin (10,000 U/ml), and growth factors EGF and FGF (20 ng/ml). Tumor explants were then transferred to the incubator chamber of a single-photon laser scanning confocal microscope model LSM 880 (Carl Zeiss, Jena, Germany). For tumor imaging, the incubation chamber of the microscope was maintained at 37°C and 5% CO<sub>2</sub>. Images were acquired in a time-lapse frame of 10 min for 100 to 300 cycles. Raw data from movies used to perform the analysis described throughout this manuscript were originally described in (4).

### Image analysis

To track the evolution of the cells, we use the software Fiji with the plugin TrackMate. We use as parameters for the cell size (called "blob") 20 μm and a threshold of 1 together with the DoG method (difference of Gaussian detectors). Each experiment gives several paths denoted by  $\mathbf{x}_i(t)$  where  $i$  is an index for the cell and  $t$  represents the time. The paths are however erratic; thus, we apply a filter to smooth the trajectories over time (see Fig. 1). As a filter, we use a Gaussian kernel with SD  $\sigma^2 = 2$  and a stencil of 9 points

$$\bar{\mathbf{x}}_i(t) = \sum_{k=-4}^4 \phi_k \mathbf{x}_i(t - k\Delta t), \phi_k = C \cdot e^{-k^2/4} \quad (1)$$

where  $\bar{\mathbf{x}}_i$  is the smooth trajectory,  $\Delta t = 10$  mn is the time step between successive image, and  $C$  is such that  $\sum_{k=-4}^4 \phi_k = 1$ . From the smooth trajectories  $\bar{\mathbf{x}}_i(t)$ , we then estimate the velocities of the cells  $\bar{\mathbf{v}}_i(t)$  using a finite difference

$$\bar{\mathbf{v}}_i(t) = \frac{\bar{\mathbf{x}}_i(t + \Delta t) - \bar{\mathbf{x}}_i(t - \Delta t)}{2\Delta t} \quad (2)$$

For each velocity vector  $\mathbf{v}_i(t)$ , we estimate a corresponding velocity angle  $\theta_i(t) \in [0, 2\pi]$  (see Fig. 1). For subsequent analysis, all velocity vectors are normalized and referred to as  $\mathbf{s}_i$ .

### Classification flock-stream-swarm

We subdivide glioma populations into connected subpopulations, which we refer to as flocks, streams, or swarms [see fig. S14 (left)]. This empirical classification scheme was developed in a previous work (4) to segment populations into subgroups with similar histological and statistical properties, as described briefly below. In the present context, these classifications are not particularly important, and instead, one may view this as a systematic way of

subdividing populations for finite-size scaling, which is necessary because the number of experimental populations is limited [see (52) for further discussion of these issues in limited datasets].

To classify in which category an experiment is (i.e., flock, stream, or swarm), we analyze the collection of velocity angles  $\{\theta_n\}_{n=1..N}$  (where  $N$  is the sample size). We determine three densities for each pattern [see fig. S14 (right)]

$$\rho_{\text{flock}} = \text{“Gaussiandistribution”} \tag{3}$$

$$\rho_{\text{stream}} = \text{“symmetrizedGaussiandistribution”} \tag{4}$$

$$\rho_{\text{swarm}} = \text{“Constant”} \tag{5}$$

We then evaluate the likelihood of the velocity angles  $\{\theta_n\}_{n=1..N}$  given each distribution, for instance

$$\mathcal{L}_{\text{flock}} = \prod_{n=1}^N \rho_{\text{flock}}(\theta_n) \tag{6}$$

Last, we compare the three likelihoods (i.e.,  $\mathcal{L}_{\text{flock}}$ ,  $\mathcal{L}_{\text{stream}}$ , and  $\mathcal{L}_{\text{swarm}}$ ) and select the pattern with the highest likelihood.

### Calculating correlation functions

The correlation function is given by  $C(r) = \langle \delta s_i \cdot \delta s_j \rangle_r$ , where  $\delta s_i \equiv s_i - \langle s \rangle$  is the velocity of cell  $i$  in the moving reference frame where the population average velocity ( $\langle s \rangle$ ) has been subtracted out. Angle brackets  $\langle \rangle_r$  indicate an average taken over all cells separated by distance  $r$ . In practice, we calculated correlation functions by first calculating pairwise distances and dot products ( $\delta s_i \cdot \delta s_j$ ) for all pairs of cells in an image at a given time. We repeated this process for all images in a time series and combined the data from all time points into a single pair of lists (distances and corresponding dot products, ordered according to distance). The correlation functions are then calculated by smoothing using a moving average filter with a centered window of  $\pm 100 \mu\text{m}$ . Error bars/shading represent  $\pm 1$  SEM within each window. Correlation functions (and the system-size scaling of correlations length) are qualitatively similar (although more or less noisy) for a range of window sizes from  $\pm 25$  up to  $\pm 100 \mu\text{m}$  (fig. S3).

### Estimating system size ( $L$ )

To estimate the size  $L$  of each population, we calculated the maximum pairwise separation [ $L(t) \equiv d_{\text{max}}$ ] between any two cells at each time point. The distance  $L(t)$  will fluctuate over time as the cells move. We define the size of the system as the mean value of this separation distance over the entire time series [ $L \equiv \langle L(t) \rangle_t$ ]. We found qualitatively similar results (i.e., scaling of correlation length with system size) if we alternatively defined the system size at each time point as  $L(t) = d_{\text{con}}$ , where  $d_{\text{con}} = A^{1/2}$  and  $A$  is the area of the convex hull of all cell positions S5.

### Box scaling analysis

To perform the box scaling analysis, we combined the data from all samples into a single dataset consisting of approximately 400 single-time snapshots of approximately  $800 \mu\text{m}$  by  $800 \mu\text{m}$ , each containing position and velocity information for hundreds of cells. For a given box size  $W$ , we selected all cells within a randomly positioned

square box of length  $W$  from each snapshot. We then calculated the separation distances and dot products ( $\delta s_i \cdot \delta s_j$ ) for all pairs of cells within the box. The box sampling was repeated a total of 30 times for each snapshot, each with a newly positioned box, and the pairwise distance and dot product data were then combined to estimate a single correlation function  $C_w(r)$  for systems in boxes of length  $W$ . We calculated these correlation functions for box sizes  $W$  ranging from approximately 200 to  $800 \mu\text{m}$  and then used them to estimate a single characteristic length  $r_0$  for each box size, as described above in the section “Calculating correlation functions” (this time using a moving average filter with a centered window of  $\pm 50 \mu\text{m}$ ).

### Nonlocally coupled XY model as a data-constrained maximum entropy model

To model collective motion in tumor cells, we used a maximum entropy framework. Maximum entropy models are required to match certain features of the data (in this case, the observed correlation  $C_{\text{int}}$ , see below) but contain minimal additional statistical structure. Similar models have been previously used to describe firing patterns in neurons, immune system dynamics, interactions between antibiotics, and collective motion in starlings. Following (12), we constrain the model to match the scalar correlation over local neighborhoods of size  $n_c$ , which is given by

$$C_{\text{int}} = \frac{1}{N} \sum_{i=1}^N \frac{1}{n_c} \sum_{j \in n_i^c} \bar{s}_i \cdot \bar{s}_j \tag{7}$$

where  $N$  is the total number of cells,  $n_c$  is the integer size of the local neighborhood, and  $\bar{s}_i$  is the unit velocity vector describing the motion of cell  $i$ . The maximum entropy model consistent with this constraint is then given by (12)

$$P(\{\bar{s}_i\}) = \frac{1}{Z(J, n_c)} \exp \left( \frac{J}{2N} \sum_{i=1}^N \sum_{j \in n_i^c} \bar{s}_i \cdot \bar{s}_j \right) \tag{8}$$

where  $P(\{\bar{s}_i\})$  is the distribution over all configurations and  $Z(J, n_c)$  is the partition function (i.e., the normalization constant). To fully specify the model, we must choose values of the parameters  $J$  and  $n_c$  such that the model reproduces the observed value of  $C_{\text{int}}$ , a process that is equivalent to maximizing the likelihood that the model generates the configuration observed in a given snapshot of the flock.

To estimate model parameters, we first calculate the experimental correlation  $C_{\text{int}}^{\text{exp}}$  (where the superscript “exp” indicates that this is the value observed in the experiment) for a single snapshot of the population. This experimental value must match the value of  $C_{\text{int}}$  produced by the model, which provides an explicit data-dependent constraint between the parameters  $J$  and  $n_c$

$$\frac{1}{J} = \frac{n_c}{2N} (1 - C_{\text{int}}^{\text{exp}}) \tag{9}$$

We then determine the remaining free parameter,  $n_c$ , by numerically maximizing the log likelihood of the data given the model, which can be written as a function of the both  $C_{\text{int}}^{\text{exp}}$  (a measured quantity) and  $Z(J, n_c)$  (the partition function for the model).

Directly calculating the partition function  $Z(J, n_c)$  is computationally expensive, even for this simple model. One option to simplify the calculation is to use a spin-wave approximation, which provides an analytical expression for  $Z(J, n_c)$  in terms of eigenvalues

of a matrix  $A$  that describes the neighborhood structure of the population (i.e., which cells are in the local neighborhood, of size  $n_c$ , of each cell). In the spin-wave approximation, the partition function reduces to (12)

$$\log Z(J, n_c) = - \sum_{k>1} \log \left( \frac{J \lambda_k}{N} \right) + \frac{J n_c}{2} \quad (10)$$

where  $\lambda_k$  are the eigenvalues of the matrix  $A = \delta_{ij} \sum_k n_{ik} - n_{ij}$ , where  $n_{ij}$  is 1 if cell  $i$  is in the local neighborhood of cell  $j$  and vice versa,  $1/2$  if cell  $i$  is in the local neighborhood of cell  $j$  but  $j$  is not in the local neighborhood of cell  $i$  (or vice versa), and 0 otherwise. Note that the spin-wave approximation is strictly valid only in highly polarized populations, as it relies on neglecting higher-order terms in an expansion velocity components perpendicular to the direction of polarization. In this work, we used the spin-wave approximation to provide first-pass estimates of the parameters  $J$  and  $n_c$ . While cell motion is not typically highly polarized in our dataset and the spin-wave approximation is therefore not strictly valid, we found that the parameters estimated from this approach do lead to strong agreement with the data.

### Estimating parameters from simulated data

To probe the reliability of parameters estimated with this spin-wave approximation, we used Monte Carlo simulations to produce artificial datasets, snapshots of velocity configurations for populations of cells whose positions are identical to those in the experimental data, drawn from the model with specified values of  $J$  and  $n_c$ . We then estimated the parameters directly from the simulated data using either (i) the spin-wave approximation or (ii) a direct least squares fitting to observed distribution  $P(\{\theta_i\})$ , where  $\{\theta_i\}$  is the configuration of velocity angles in the population.

### Mean field XY model

The maximum entropy model has two free parameters,  $J$  and  $n_c$ , that are inferred from experimental data. In the limit  $n_c \rightarrow N - 1 \approx N$  (all-to-all coupling), this model reduces to a mean field version of the 2D XY model with only a single free parameter,  $J$ . Estimating this parameter from data reduces to a classic inverse problem in statistical physics, with the log likelihood  $L(J)$  of the observed configurations taking the form

$$L(J) = \frac{1}{2} J n_c C_{\text{int}}^{\text{exp}} - \log Z(J) \quad (11)$$

where  $C_{\text{int}}^{\text{exp}}$  is now the average pairwise correlation taken over the entire population. Maximizing the likelihood is equivalent to requiring that the correlations measured experimentally ( $C_{\text{int}}^{\text{exp}}$ ) match those from the mean field model, which reduces to a classical inverse problem. In this case, we can write down an explicit expression for  $J$  in terms of the experimental measurable

$$J = \frac{2}{(1 - C_{\text{int}}^{\text{exp}})} \quad (12)$$

### Estimating critical parameters in systems with long-range coupling

Equation 12 highlights an important caveat of our approach when the range of coupling  $n_c$  is comparable to the total size of the system ( $N$ ). The mean field XY model undergoes a phase transition at a

critical value of  $J_c = 2$ . Equation 12 therefore indicates that observed data characterized by  $C_{\text{int}}^{\text{exp}} \approx 0$  would appear to be at a critical point of the mean field model. In practice,  $C_{\text{int}}^{\text{exp}} \approx 0$  would also be expected for completely disordered populations, that is, for populations where the velocity direction is drawn from a uniform distribution. As a result, maximum likelihood estimates that indicate  $J \approx J_c$  are not, alone, sufficient evidence of criticality, as they cannot distinguish systems in the disordered region ( $J < J_c$ ) from those at criticality. To quantitatively characterize these limitations for datasets comparable in size to our experimental data, we used Monte Carlo simulations to generate artificial datasets representing cell populations globally coupled with different values of  $J$ . We then calculated the maximum likelihood estimates of  $J$  from those in silico datasets (just as we did with experimental data). As expected, we found that estimates of  $J$  hover around the critical value  $J_c = 2$  for simulated systems at or below the critical point (fig. S11). On the other hand, estimates of  $J$  are accurate for systems slightly above the critical point.

### Monte Carlo simulations and phase diagram

To characterize the phase diagram for the nonlocal XY model, we used Monte Carlo simulations to generate representative velocity distributions for different values of  $n_c$  and  $J$ . For each parameter pair, we calculated the mean polarization across trials and the generalized susceptibility  $\chi$ , which corresponds (up to a proportionality constant) to the variance of the polarization across trials. We then estimated the critical surface to be the curve in parameter space that corresponds to a peak in  $\chi$ . We also verified that the system exhibits sharply increasing polarization and a peak in the generalized heat capacity at the critical surface. For all simulations, cell positions (and, therefore, cell density) were taken from a representative experimental dataset.

### Comparing velocity angle distributions between the model and data

To compare experimental results with results from the effective model, we calculated histograms of velocity angle. The model is symmetric under rotation and cannot provide information about the specific velocity direction, that is, we can globally rotate all velocities by an arbitrary angle. To compare histograms between the model and the data or, for example, to combine data from multiple independent simulations (model) or multiple time points (experiment), we first calculated the two frequency histograms to be compared. Then, we rotated all velocities in one population by an angle  $\theta$ , and we tuned  $\theta$  to achieve maximal alignment between the distributions (i.e., minimal difference in a least squares sense). Because the histograms themselves are noisy, this process could, in principle, lead to apparent similarities between distributions with fundamentally different shapes. Therefore, as a control, we simulated size-matched populations where velocity angles were drawn from a uniform distribution, and we then used this alignment process to compare these (nominally) uniform histograms with the histograms from the maximum entropy model. We confirmed that this alignment process does not yield substantial agreement between these different distributions (fig. S11).

## Supplementary Materials

## This PDF file includes:

Supplementary materials. S1 to S14

[View/request a protocol for this paper from Bio-protocol.](#)

## REFERENCES AND NOTES

- D. N. Louis, A. Perry, P. Wesseling, D. J. Brat, I. A. Cree, D. Figarella-Branger, C. Hawkins, H. K. Ng, S. M. Pfister, G. Reifenberger, R. Soffietti, A. von Deimling, D. W. Ellison, The 2021 WHO classification of tumors of the central nervous system: A summary. *Neuro Oncol.* **23**, 1231–1251 (2021).
- M. Ceccarelli, F. P. Barthel, T. M. Malta, T. S. Sabedot, S. R. Salama, B. A. Murray, O. Morozova, Y. Newton, A. Radenbaugh, S. M. Pagnotta, S. Anjum, J. Wang, G. Manyam, P. Zoppoli, S. Ling, A. A. Rao, M. Grifford, A. D. Cherniack, H. Zhang, L. Poisson, C. G. Carloti Jr., D. P. da Cunha Tirapelli, A. Rao, T. Mikkelsen, C. C. Lau, W. K. A. Yung, R. Rabadan, J. Huse, D. J. Brat, N. L. Lehman, J. S. Barnholtz-Sloan, S. Zheng, K. Hess, G. Rao, M. Meyerson, R. Beroukhim, L. Cooper, R. Akbani, M. Wrensch, D. Haussler, K. D. Aldape, P. W. Laird, D. H. Gutmann; TCGA Research Network, H. Nounshahr, A. Iavarone, R. G. W. Verhaak, Molecular profiling reveals biologically discrete subsets and pathways of progression in diffuse glioma. *Cell* **164**, 550–563 (2016).
- P. G. Gritsenko, O. Ilna, P. Friedl, Interstitial guidance of cancer invasion. *J. Pathol.* **226**, 185–199 (2012).
- A. Comba, S. Motsch, P. J. Dunn, T. C. Hollon, A. E. Argento, D. B. Zmler, P. E. Kish, A. Kahana, C. G. Kleer, M. G. Castro, P. R. Lowenstein, Spatiotemporal analysis of glioma heterogeneity reveals COL1A1 as an actionable target to disrupt tumor progression. *Nat. Commun.* **13**, 3606 (2022).
- M. Alieva, V. Leidgens, M. J. Riemenschneider, C. A. Klein, P. Hau, J. van Rheenen, Intravital imaging of glioma border morphology reveals distinctive cellular dynamics and contribution to tumor cell invasion. *Sci. Rep.* **9**, 2054 (2019).
- A. Haeger, M. Krause, K. Wolf, P. Friedl, Cell jamming: Collective invasion of mesenchymal tumor cells imposed by tissue confinement. *Biochim. Biophys. Acta* **1840**, 2386–2395 (2014).
- S. Jamous, A. Comba, P. R. Lowenstein, S. Motsch, Self-organization in brain tumors: How cell morphology and cell density influence glioma pattern formation. *PLOS Comput. Biol.* **16**, e1007611 (2020).
- P. G. Gritsenko, N. Atlay, C. E. J. Dieteren, A. C. Navis, J. H. Venhuizen, C. Veelken, D. Schubert, A. Acker-Palmer, B. A. Westerman, T. Wurdinger, W. Leenders, P. Wesseling, H. G. Stunnenberg, P. Friedl, p120-catenin-dependent collective brain infiltration by glioma cell networks. *Nat. Cell Biol.* **22**, 97–107 (2020).
- C. J. Liu, G. A. Shamsan, T. Akkin, D. J. Odde, Glioma cell migration dynamics in brain tissue assessed by multimodal optical imaging. *Biophys. J.* **117**, 1179–1188 (2019).
- A. Cavagna, A. Cimarelli, I. Giardina, G. Parisi, R. Santagati, F. Stefanini, M. Viale, Scale-free correlations in starling flocks. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 11865–11870 (2010).
- G. F. Young, L. Scardovi, A. Cavagna, I. Giardina, N. E. Leonard, Starling flock networks manage uncertainty in consensus at low cost. *PLOS Comput. Biol.* **9**, e1002894 (2013).
- W. Bialek, A. Cavagna, I. Giardina, T. Mora, E. Silvestri, M. Viale, A. M. Walczak, Statistical mechanics for natural flocks of birds. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 4786–4791 (2012).
- W. Bialek, A. Cavagna, I. Giardina, T. Mora, O. Pohl, E. Silvestri, M. Viale, A. M. Walczak, Social interactions dominate speed control in poising natural flocks near criticality. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 7212–7217 (2014).
- H. Ling, G. E. McIvor, K. van der Vaart, R. T. Vaughan, A. Thornton, N. T. Ouellette, Local interactions and their group-level consequences in flocking jackdaws. *Proc. Royal Soc. B* **286**, 20190865 (2019).
- U. Lopez, J. Gautrais, I. D. Couzin, G. Theraulaz, From behavioural analyses to models of collective motion in fish schools. *Interface Focus* **2**, 693–707 (2012).
- C. Hemelrijk, D. Reid, H. Hildenbrandt, J. Padding, The increased efficiency of fish swimming in a school. *Fish Fisheries* **16**, 511–521 (2015).
- A. Sokolov, I. S. Aranson, Physical properties of collective motion in suspensions of bacteria. *Phys. Rev. Lett.* **109**, 248109 (2012).
- H.-P. Zhang, A. Be'er, E.-L. Florin, H. L. Swinney, Collective motion and density fluctuations in bacterial colonies. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 13626–13630 (2010).
- O. Feinerman, I. Pinkovizky, A. Gelblum, E. Fonio, N. S. Gov, The physics of cooperative transport in groups of ants. *Nat. Phys.* **14**, 683–693 (2018).
- J. Buhl, D. J. Sumpter, I. D. Couzin, J. J. Hale, E. Despland, E. R. Miller, S. J. Simpson, From disorder to order in marching locusts. *Science* **312**, 1402–1406 (2006).
- A. Attanasi, A. Cavagna, L. Del Castello, I. Giardina, S. Melillo, L. Parisi, O. Pohl, B. Rossaro, E. Shen, E. Silvestri, M. Viale, Finite-size scaling as a way to probe near-criticality in natural swarms. *Phys. Rev. Lett.* **113**, 238102 (2014).
- N. Ouellette, A physics perspective on collective animal behavior. *Phys. Biol.* **19**, 021004 (2022).
- K. van der Vaart, M. Sinhuber, A. M. Reynolds, N. T. Ouellette, Environmental perturbations induce correlations in midge swarms. *J. R. Soc. Interface* **17**, 20200018 (2020).
- N. Bain, D. Bartolo, Dynamic response and hydrodynamics of polarized crowds. *Science* **363**, 46–49 (2019).
- A. Bottinelli, D. T. Sumpter, J. L. Silverberg, Emergent structural mechanisms for high-density collective motion inspired by human crowds. *Phys. Rev. Lett.* **117**, 228301 (2016).
- A. Cavagna, I. Giardina, Large-scale behaviour in animal groups. *Behav. Processes* **84**, 653–656 (2010).
- T. Vicsek, A. Zafeiris, Collective motion. *Phys. Rep.* **517**, 71–140 (2012).
- E. Méhes, T. Vicsek, Collective motion of cells: From experiments to models. *Integrative Biol.* **6**, 831–854 (2014).
- G. Vásárhelyi, C. Virágh, G. Somorjai, T. Nepusz, A. E. Eiben, T. Vicsek, Optimized flocking of autonomous drones in confined environments. *Sci. Robot.* **3**, eaat3536 (2018).
- L. Qin, D. Yang, W. Yi, H. Cao, G. Xiao, Roles of leader and follower cells in collective cell migration. *Mol. Biol. Cell* **32**, 1267–1272 (2021).
- J. Zhang, K. F. Goliwas, W. Wang, P. V. Tafalele, F. Bordeleau, C. A. Reinhart-King, Energetic regulation of coordinated leader-follower dynamics during collective invasion of breast cancer cells. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 7867–7872 (2019).
- T. Mora, W. Bialek, Are biological systems poised at criticality? *J. Stat. Phys.* **144**, 268–302 (2011).
- H. Trenchard, M. Perc, Energy saving mechanisms, collective behavior and the variation range hypothesis in biological systems: A review. *Biosystems* **147**, 40–66 (2016).
- P. Bak, M. Paczusi, Complexity, contingency, and criticality. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 6689–6696 (1995).
- P. M. Chaikin, T. C. Lubensky, T. A. Witten, *Principles of condensed matter physics*, vol. 10 (Cambridge university press Cambridge, 2012).
- G. Grinstein, D.-H. Lee, S. Sachdev, Conservation laws, anisotropy, and “self-organized criticality” in noisy nonequilibrium systems. *Phys. Rev. Lett.* **64**, 1927–1930 (1990).
- N. W. Watkins, G. Pruessner, S. C. Chapman, N. B. Crosby, H. J. Jensen, 25 years of self-organized criticality: Concepts and controversies. *Space Sci. Rev.* **198**, 3–44 (2016).
- V. Buendía, S. Di Santo, J. A. Bonachela, M. A. Muñoz, Feedback mechanisms for self-organization to the edge of a phase transition. *Front. Phys.* **8**, 333 (2020).
- M. A. Muñoz, Colloquium: Criticality and dynamical scaling in living systems. *Rev. Mod. Phys.* **90**, 031001 (2018).
- G. Tkačik, T. Mora, O. Marre, D. Amodei, S. E. Palmer, M. J. Berry, W. Bialek, Thermodynamics and signatures of criticality in a network of neurons. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 11508–11513 (2015).
- M. Rubinov, O. Sporns, J.-P. Thivierge, M. Breakspear, Neurobiologically realistic determinants of self-organized criticality in networks of spiking neurons. *PLOS Comput. Biol.* **7**, e1002038 (2011).
- D. Fraiman, D. R. Chialvo, What kind of noise is brain noise: Anomalous scaling behavior of the resting brain activity fluctuations. *Front. Physiol.* **3**, 307 (2012).
- B. C. Daniels, H. Kim, D. Moore, S. Zhou, H. B. Smith, B. Karas, S. A. Kauffman, S. I. Walker, Criticality distinguishes the ensemble of biological regulatory networks. *Phys. Rev. Lett.* **121**, 138102 (2018).
- Q.-Y. Tang, Y.-Y. Zhang, J. Wang, W. Wang, D. R. Chialvo, Critical fluctuations in the native state of proteins. *Phys. Rev. Lett.* **118**, 088102 (2017).
- A. Comba, P. J. Dunn, A. E. Argento, P. Kadiyala, M. Ventosa, P. Patel, D. B. Zmler, F. J. Nunez, L. Zhao, M. G. Castro, Fyn tyrosine kinase, a downstream target of receptor tyrosine kinases, modulates antiglioma immune responses. *Neuro Oncol.* **22**, 806–818 (2020).
- F. J. Núñez, F. M. Mendez, P. Kadiyala, M. S. Alghamri, M. G. Savelieff, M. B. Garcia-Fabiani, S. Haase, C. Koschmann, A.-A. Calinescu, N. Kamran, M. Saxena, R. Patel, S. Carney, M. Z. Guo, M. Edwards, M. Ljungman, T. Qin, M. A. Sartor, R. Tagett, S. Venneti, J. Brosnan-Cashman, A. Meeker, V. Gorbunova, L. Zhao, D. M. Kremer, L. Zhang, C. A. Lyssiotis, L. Jones, C. J. Herting, J. L. Ross, D. Hambarzumyan, S. Hervey-Jumper, M. E. Figueroa, P. R. Lowenstein, M. G. Castro, Idh1-r132h acts as a tumor suppressor in glioma via epigenetic up-regulation of the dna damage response. *Sci. Transl. Med.* **11**, eaaq1427 (2019).
- C. Koschmann, A. A. Calinescu, F. J. Nunez, A. Mackay, J. Fazal-Salom, D. Thomas, F. Mendez, N. Kamran, M. Dzaman, L. Mulpuri, J. Krasinkiewicz, R. Doherty, R. Lemons, J. A. Brosnan-Cashman, Y. Li, S. Roh, L. Zhao, H. Appelman, D. Ferguson, V. Gorbunova, A. Meeker, C. Jones, P. R. Lowenstein, M. G. Castro, Atrx loss promotes tumor growth and impairs non-homologous end joining dna repair in glioma. *Sci. Transl. Med.* **8**, 328ra28 (2016).
- T. Niizato, H. Murakami, Y.-P. Gunji, Emergence of the scale-invariant proportion in a flock from the metric-topological interaction. *Biosystems* **119**, 62–68 (2014).

49. E. Schneidman, M. J. Berry, R. Segev, W. Bialek, Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature* **440**, 1007–1012 (2006).
50. J. Shlens, G. D. Field, J. L. Gauthier, M. I. Grivich, D. Petrusca, A. Sher, A. M. Litke, E. Chichilnisky, The structure of multi-neuron firing patterns in primate retina. *J. Neurosci.* **26**, 8254–8266 (2006).
51. A. Cavagna, I. Giardina, T. S. Grigera, The physics of flocking: Correlation as a compass from experiments to theory. *Phys. Rep.* **728**, 1–62 (2018).
52. D. A. Martin, T. L. Ribeiro, S. A. Cannas, T. S. Grigera, D. Plenz, D. R. Chialvo, Box scaling as a proxy of finite size correlations. *Sci. Rep.* **11**, 1–9 (2021).
53. J. N. Kapur, *Maximum-entropy models in science and engineering* (John Wiley & Sons, 1989).
54. K. Wood, S. Nishida, E. D. Sontag, P. Cluzel, Mechanism-independent method for predicting response to multidrug combinations in bacteria. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 12254–12259 (2012).
55. G. Yeo, C. B. Burge, Maximum entropy modeling of short sequence motifs with applications to rna splicing signals. *J. Comput. Biol.* **11**, 377–394 (2004).
56. E. T. Jaynes, Information theory and statistical mechanics. *Phys. Rev.* **106**, 620–630 (1957).
57. H. C. Nguyen, R. Zecchina, J. Berg, Inverse statistical problems: From the inverse Ising problem to data science. *Adv. Phys.* **66**, 197–261 (2017).
58. M. Castellana, W. Bialek, Inverse spin glass and related maximum entropy problems. *Phys. Rev. Lett.* **113**, 117204 (2014).
59. R. R. Stein, D. S. Marks, C. Sander, Inferring pairwise interactions from biological data using maximum-entropy probability models. *PLoS Comput. Biol.* **11**, e1004182 (2015).
60. J. Hidalgo, J. Grilli, S. Suweis, M. A. Munoz, J. R. Banavar, A. Maritan, Information-based fitness and the emergence of criticality in living systems. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 10095–10100 (2014).
61. J. Humplik, G. Tkačik, Probabilistic models for neural populations that naturally capture global coupling and criticality. *PLoS Comput. Biol.* **13**, e1005763 (2017).
62. N. Stollenwerk, V. Jansen, *Population Biology and Criticality* (World Scientific, 2010).
63. N. Goldenfeld, *Lectures on phase transitions and the renormalization group* (CRC Press, 2018).
64. T. Mora, A. M. Walczak, L. Del Castello, F. Ginelli, S. Melillo, L. Parisi, M. Viale, A. Cavagna, I. Giardina, Local equilibrium in bird flocks. *Nat. Phys.* **12**, 1153–1157 (2016).
65. A. Cavagna, I. Giardina, F. Ginelli, T. Mora, D. Piovani, R. Tavarone, A. M. Walczak, Dynamical maximum entropy approach to flocking. *Phys. Rev. E* **89**, 042707 (2014).
66. Y. Roudi, S. Nirenberg, P. E. Latham, Pairwise maximum entropy models for studying large biological systems: When they can work and when they can't. *PLoS Comput. Biol.* **5**, e1000380 (2009).
67. L. Merchan, I. Nemenman, On the sufficiency of pairwise interactions in maximum entropy models of networks. *J. Stat. Phys.* **162**, 1294–1308 (2016).
68. I. Mastromatteo, M. Marsili, On the criticality of inferred models. *J. Stat. Mech.* **2011**, P10012 (2011).
69. D. J. Schwab, I. Nemenman, P. Mehta, Zipf's law and criticality in multivariate data without fine-tuning. *Phys. Rev. Lett.* **113**, 068102 (2014).
70. M. Nonnenmacher, C. Behrens, P. Berens, M. Bethge, J. H. Macke, Signatures of criticality arise from random subsampling in simple population models. *PLoS Comput. Biol.* **13**, e1005718 (2017).
71. R. J. Cubero, J. Jo, M. Marsili, Y. Roudi, J. Song, Statistical criticality arises in most informative representations. *J. Stat. Mech.* **2019**, 063402 (2019).
72. K. Painter, T. Hillen, Mathematical modelling of glioma growth: The use of diffusion tensor imaging (dti) data to predict the anisotropic pathways of cancer invasion. *J. Theor. Biol.* **323**, 25–39 (2013).
73. G. J. Baker, V. N. Yadav, S. Motsch, C. Koschmann, A.-A. Calinescu, Y. Mineharu, S. I. Camelo-Piragua, D. Orringer, S. Bannykh, W. S. Nichols, A. C. deCarvalho, T. Mikkelsen, M. G. Castro, P. R. Lowenstein, Mechanisms of glioma formation: Iterative perivascular glioma growth and invasion leads to tumor progression, vegf-independent vascularization, and resistance to antiangiogenic therapy. *Neoplasia* **16**, 543–561 (2014).
74. D. A. Hormuth, J. A. Weis, S. L. Barnes, M. I. Miga, E. C. Rericha, V. Quaranta, T. E. Yankeelov, A mechanically coupled reaction–diffusion model that incorporates intra-tumoural heterogeneity to predict in vivo glioma growth. *J. R. Soc. Interface* **14**, 20161010 (2017).
75. D. A. Hormuth II, S. L. Eldridge, J. A. Weis, M. I. Miga, T. E. Yankeelov, Mechanically coupled reaction–diffusion model to predict glioma growth: Methodological details. *Methods Mol. Biol.*, 225–241 (2018).
76. L. Hutchinson, H.-J. Mueller, E. Gaffney, P. Maini, J. Wagg, A. Phipps, C. Boetsch, H. Byrne, B. Ribba, Modeling longitudinal preclinical tumor size data to identify transient dynamics in tumor response to antiangiogenic Drugs. *Pharmacology* **5**, 636–645 (2016).
77. M. Khasraw, M. S. Ameratunga, R. Grant, H. Wheeler, N. Pavlakis, Antiangiogenic therapy for high-grade glioma. *Cochrane Database Syst. Rev.*, CD008218 (2014).
78. R. Ni, N. Ouellette, Velocity correlations in laboratory insect swarms. *Eur. Phys. J. Spec. Top.* **224**, 3271–3277 (2015).
79. M. Osswald, E. Jung, F. Sahn, G. Solecki, V. Venkataramani, J. Blaes, S. Weil, H. Horstmann, B. Wiestler, M. Syed, L. Huang, M. Ratliff, K. Karimian Jazi, F. T. Kurz, T. Schmenger, D. Lemke, M. Gömmel, M. Pauli, Y. Liao, P. Häring, S. Pusch, V. Herl, C. Steinhäuser, D. Krunic, M. Jarahian, H. Miletic, A. S. Berghoff, O. Griesbeck, G. Kalamakis, O. Garaschuk, M. Preusser, S. Weiss, H. Liu, S. Heiland, M. Platten, P. E. Huber, T. Kuner, A. von Deimling, W. Wick, F. Winkler, Brain tumour cells interconnect to a functional and resistant network. *Nature* **528**, 93–98 (2015).
80. V. Venkataramani, D. I. Tanev, C. Strahle, A. Studier-Fischer, L. Fankhauser, T. Kessler, C. Körber, M. Kardorff, M. Ratliff, R. Xie, H. Horstmann, M. Messer, S. P. Paik, J. Knabbe, F. Sahn, F. T. Kurz, A. A. Acikgöz, F. Herrmannsdörfer, A. Agarwal, D. E. Bergles, A. Chalmers, H. Miletic, S. Turcan, C. Mawrin, D. Hänggi, H. K. Liu, W. Wick, F. Winkler, T. Kuner, Glutamate-ergic synaptic input to glioma cells drives brain tumour progression. *Nature* **573**, 532–538 (2019).

#### Acknowledgments

**Funding:** This work was supported by the NIH/National Institute of Neurological Disorders and Stroke (NIH/NINDS) grants R01-NS105556, R01-NS122536, R01-NS124167, and R21-NS123879-01; NIH/NINDS grants R01NS122234, R01NS127378, and NIH/NCI R01-CA243916 to P.R.L.; Rogel Cancer Center, the Department of Neurosurgery; the Pediatric Brain Tumor Foundation, Leah's Happy Hearts Foundation (G013908), Ian's Friends Foundation (IFF) (G024230), Chad Tough Foundation (G023419), and Smiles for Sophie Forever Foundation to P.R.L.; NIH R35GM124875 to K.W.; Health and Human Services, NIH, UL1 TR002240 to Michigan Institute for Clinical and Health Research (MICH), Postdoctoral Translational Scholars Program (PTSP), Project F049768 to A.C.; and the NSF DMS-2206330 to S.M. **Author contributions:** Conceptualization: K.B.W. and P.R.L. Methodology: K.B.W., A.C., T.S.G., S.M., and P.R.L. Investigation: K.B.W. and P.R.L. Data analysis and interpretation: K.B.W., A.C., T.S.G., S.M., and P.R.L. Writing (original draft): K.B.W. and P.R.L. Writing (review and editing): K.B.W., A.C., T.S.G., S.M., and P.R.L. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. All data from the time-lapse confocal imaging (movies) used in this study are available online in the Supplementary Materials of the paper by Comba *et al.* (4).

Submitted 9 November 2022

Accepted 24 May 2023

Published 28 June 2023

10.1126/sciadv.adf7170