Data-Driven Theory Reveals Protrusion and Polarity Interactions Governing Collision Behavior of Distinct Motile Cells

Tom Brandstätter,[1,2,*] Emily Brieger,[3,*] David B. Brückner,[4,5] Georg Ladurner,[3,6] Joachim O. Rädler,[3] and Chase P. Broedersz[1,2,†]

[1]*Department of Physics and Astronomy, Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands*
[2]*Arnold-Sommerfeld-Center for Theoretical Physics, Ludwig-Maximilians-Universität München, Theresienstr. 37, 80333 Munich, Germany*
[3]*Faculty of Physics and Center for NanoScience, Ludwig-Maximilians-Universität München, Geschwister-Scholl-Platz 1, 80539 Munich, Germany*
[4]*Institute of Science and Technology Austria, Am Campus 1, 3400 Klosterneuburg, Austria*
[5]*Biozentrum and Department of Physics, University of Basel, Basel, Switzerland*
[6]*Center for Medical Physics and Biomedical Engineering, Währinger Gürtel 18-20, 1090 Vienna, Austria*

(Received 9 August 2024; accepted 1 August 2025; published 26 August 2025)

The migration behavior of colliding cells is critically determined by transient contact interactions. During these interactions, the motility machinery, including the front-rear polarization of the cell, dynamically responds to surface protein-mediated transmission of forces and biochemical signals between cells. While biomolecular details of such contact interactions are increasingly well understood, it remains unclear what biophysical interaction mechanisms govern the cell-level dynamics of colliding cells and how these mechanisms vary across cell types. Here we develop a phenomenological theory based on 14 candidate contact-interaction mechanisms coupling cell position, protrusion, and polarity. Using high-throughput micropattern experiments, we detect which of these phenomenological contact interactions captures the interaction behaviors of cells. We find that various cell types—ranging from mesenchymal to epithelial cells—are accurately captured by a single model with only two interaction mechanisms: polarity-protrusion coupling and polarity-polarity coupling. Remarkably, the qualitatively different interaction behaviors of distinct cells, as well as cells subject to molecular perturbations of surface protein-mediated signaling, can all be quantitatively captured by varying the strength and sign of the polarity-polarity coupling mechanism. Altogether, our data-driven phenomenological theory of cell-cell interactions reveals polarity-polarity coupling as a versatile and general contact-interaction mechanism, which may underlie diverse collective migration behaviors of motile cells.

DOI: 10.1103/3hhj-rt1n

I. INTRODUCTION

Contact interactions between cells control the coordinated migration of tissues during fundamental physiological processes [1] in development, health, and disease [2–4]. Already at the level of two motile cells, contact interactions can lead to dramatic changes in the trajectories of cells after they collide [5–8]. Such interaction behavior is not merely the result of physical forces arising, for example, from membrane adhesion or deformations of the colliding cells. Instead, contact interactions trigger a dynamical response in the intracellular biomolecular motility machinery [9–11]. This machinery involves establishing a front-rear polarization [12,13], as well as active cytoskeleton contraction and protrusion formation, enabling cells to self-propel [14]. Contact-interactions can modify cell polarity via biochemical signaling between cells [9,15], reflecting the active, responsive, and adaptive nature of cell motility and cell-cell interactions. Revealing the dominant interaction mechanisms that steer cell polarity and cell migration in response to cell-cell contacts is critical for a general understanding of the dynamics of interacting cells.

The interaction behavior of migrating cells in various processes is highly diverse. During wound healing [3], for instance, motile epithelial cells follow each other after making contact [16,17]. By contrast, several developmental processes rely on cells retreating from each other after forming cell-cell contacts [18–21], termed contact inhibition of locomotion (CIL) [6,22]. Many details of the biomolecular machinery underlying these and other interaction behaviors [23,24] are now becoming increasingly clear. For instance, force transmission between epithelial cells is enabled by E-cadherin-mediated junctions [25,26], while receptor-ligand Eph-ephrin interactions [15,21,27–29] enable cell-cell recognition during CIL. Both of these surface proteins enable cells to regulate their polarity machinery in response to cell-cell contacts [11,29,30]. Such molecular pathways vary between different cell types and can change, for instance, during the

*These authors contributed equally to this work.
†Contact author: c.p.broedersz@vu.nl

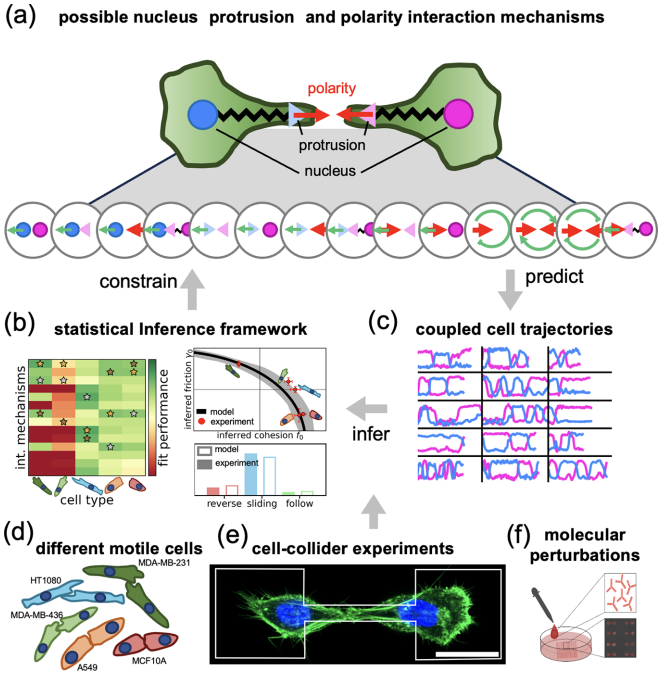

FIG. 1. Schematic of our data-driven theoretical approach. (a) Phenomenological interaction mechanisms coupling the nuclei, protrusions, and polarities of two interacting cells. We show a detailed analysis of these mechanisms in Fig. 2. (b) Statistical inference of our theory using the experimental trajectory data of each cell type. We quantify the experimental dynamics by inferring effective interaction terms that describe the nucleus dynamics of our experimental data (red points) and the phenomenological model (curves) [8]. In addition, we quantify collision behavior. The heatmap shows the best-fit performance of each interaction mechanism for the nucleus dynamics and collision behavior of each cell type. (c) We experimentally observe and predict the trajectories of the nuclei of the two cells, which quantifies cell-level interaction behavior. (d)–(f) Our experimental setup consists of cell-collider experiments, which we perform on dumbbell-shaped micropatterns. Scale bar: 25 μm, nuclei are stained in blue, actin is stained in green (e). In addition to considering different motile cells ranging from mesenchymal cancer cells to epithelial cells (d), we also consider various molecular perturbations of the surface protein-driven interaction machinery (f).

epithelial-to-mesenchymal transition (EMT) [31,32]. However, it remains a major challenge to understand how contact interactions are controlled by the molecular machinery in various cells types.

The complexity of the motility and cell-cell interaction machinery makes it difficult to gain mechanistic understanding of contact interactions. Nevertheless, biophysical models can give mechanistic insight into single-cell motility [14,33–36] and how force transmission and cell-cell recognition give rise to the dynamics of interacting cells [5,17,23,37–48]. However, these bottom-up models are commonly tailored to understand specific cell types in concrete settings, making it difficult to achieve a unifying conceptual understanding of contact interactions. Furthermore, biophysical models for contact interactions are rarely systematically constrained on quantitative experimental data [49]. High-throughput experimental micropattern assays [7,8,40,48] enabled the application of data-driven inference approaches to learn dynamical descriptions of migrating and interacting cells directly from large datasets of experimental cell trajectories [8,49–51]. In previous work, we inferred the dynamics of the cell nuclei of interacting motile cells on micropatterns. The interaction behaviors of these cells could be captured by effective cohesion and friction interactions between the cell nucleus positions and velocities [8]. Importantly however, this approach considered neither the protrusion nor the polarity dynamics, which both play central biophysical roles in cell-cell interactions. Thus, it remains unclear what interaction mechanisms involving cell shape and the polarity-driven motility and interaction machinery govern the cell-level dynamics of interacting cells and how such mechanisms vary across cell types.

Here we develop a theory for contact-interaction mechanisms for a range of motile mesenchymal and epithelial cell types (Fig. 1). Using phenomenology based on symmetry and simplicity arguments, we derive 14 possible interaction mechanisms coupling the nucleus, protrusion, cell deformation, and polarity of pairs of cells [Fig. 1(a)]. We determine which of these mechanisms control interaction behavior of cells by analyzing roughly 50 000 hours of cell nucleus trajectories [Fig. 1(c)], which is a substantially larger dataset compared to our previous work [8]. To make use of the information contained within interacting cell trajectories, we develop a statistical inference framework accounting for both the long-time collision behavior and short-time accelerations of interacting cells [Fig. 1(b)]. In phenomenological approaches, it is common that many terms that are allowed by symmetry are necessary to capture experimental observations [52,53]. Remarkably, we discover here that, from the 14 possible mechanisms, the interaction behavior of all our cell types and various molecular perturbations [Figs. 1(d)–1(f)] are governed by only two interaction mechanisms controlling cell polarity dynamics. Additionally, only a single interaction strength needs to be varied to capture the various interaction behaviors of the different cell types. Thus, our results reveal a surprisingly simple, yet general quantitative theory of contact interactions, extending the general understanding of biophysical mechanisms of single-cell migration [14] to interacting cells.

II. RESULTS

A. Theory of contact-interaction mechanisms

To develop a dynamical theory of contact interactions, we employ a minimal large-scale description where the position of the cell's nucleus $\mathbf{x}_n(t)$ and the position of the geometric center of newly formed protrusions $\mathbf{x}_p(t)$ are the relevant positional degrees of freedom of a migrating cell [Fig. 2(a)]. In a coarse-grained sense, this simple choice faithfully captures the cell's position, as well as the elongated and dynamic shape of migrating cells as tested in single-cell migration experiments [33,55,56] and as observed more broadly during cell-cell interactions [5–8]. To complete our description, we include an internal degree of freedom $\mathbf{P}(t)$ to capture cell polarity. The polarity describes the cell's anisotropic organization of the cytoskeletal motility machinery, distinguishing front and rear [12]. This minimal level of description has

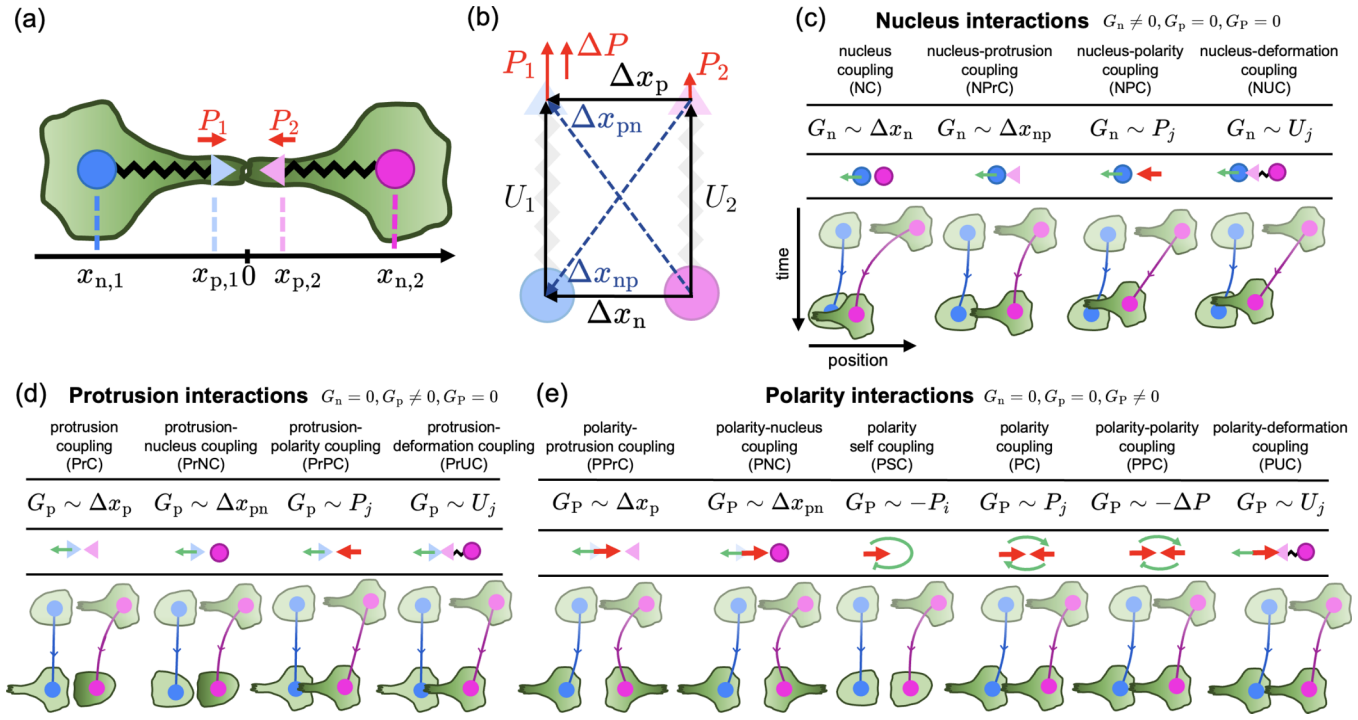

FIG. 2. Phenomenological contact-interaction mechanisms. (a) To describe the interaction behavior of a whole cell, we include three dynamical degrees of freedom in our model: the position of the nucleus x_n, the position of the protrusion x_p, and the polarity of a cell P. The polarity determines a self-propulsion force, giving rise to cell migration (Appendix B 1). (b) Schematic representation of all linear vectorial quantities in our model for both cell 1 and cell 2. $\Delta P = P_i - P_j$ is the difference between polarities and $U_i = x_{p,i} - x_{n,i}$ is the deformation vector of the cells. (c)–(e) Possible candidate cell-cell interaction mechanisms acting on the nucleus (c), on the protrusion (d) and on the polarity of cells (e). In all panels (c)–(e) the upper row of the table shows for each mechanism the corresponding interaction term and to what linear vectorial coupling this term is proportional. The middle row shows a schematic representation of that coupling at the associated interaction strength ε being positive. In these schematics, the green arrows always indicate the dynamics arising from the mechanisms for the "blue" cell in panel (b). The lower row of the tables shows then a schematic of the expected long-time scale interaction behaviors of both cells for one specific collision scenario. The different shapes of the cells indicate whether cells are elongated (nuclei and protrusions are sufficiently separated), and the green color gradient indicates the polarity of the cells. All mechanisms are described in detail in Sec. 4.1 in [54].

proven adequate to describe single-cell migration [33,34], thereby providing a foundation to develop a general theory for how contact interactions determine cell protrusions, polarity, and the dynamics of migrating cells.

We define the dynamics of interacting cells by equations of motion for the positional degrees of freedom together with a stochastic description of cell polarity. For simplicity, we consider interaction behavior of cells in one dimension, and write for cell i

$$\zeta_n \dot{x}_{n,i} = F_n(x_{n,i}, x_{p,i}) + G_n(\Delta x, U, P, \Delta P), \quad (1)$$

$$\zeta_p \dot{x}_{p,i} = F_p(x_{n,i}, x_{p,i}) + P_i(t) + G_p(\Delta x, U, P, \Delta P), \quad (2)$$

$$\dot{P}_i = F_P(x_{p,i}, P_i) + G_P(\Delta x, U, P, \Delta P) + \sigma \eta_i(t). \quad (3)$$

Here ζ_n and ζ_p are friction coefficients effectively capturing the dissipative dynamics of the nucleus and the protrusion due to the binding-unbinding dynamics of adhesions with the substrate [57,58]. Furthermore, the polarity directly drives the protrusion [59], and the functions F_n, F_p, and F_P are one-body terms describing the single-cell behavior including a mechanical coupling between nucleus and protrusion, as well as a response of the cells to their microenvironment [33] (Appendix B 1). These single-cell terms can be derived from a microscopic theory of cell migration [13,34,36] and are constrained by experimental data of single migrating cells [33] (Appendix B 2). In this model, stochasticity arises in the polarity dynamics. For simplicity, we model this stochasticity by Gaussian white noise $\eta_i(t)$ with amplitude σ, which is sufficient to capture single-cell behavior on the dumbbell-shaped micropattern [33]. Taken together, by including both protrusion, cell deformation, and polarity in our description of a migrating cell, we go beyond previous models [8,39] of the dynamics of the cell as a whole, allowing us to investigate how contact-interaction mechanisms couple to the protrusion and polarity.

The central goal of our data-driven theoretical approach is to determine whether the description in Eqs. (1)–(3) is adequate for interacting cell behavior and to learn the two-body terms G_n, G_p, and G_P. These terms encode nucleus, protrusion, and polarity interaction mechanisms and thus may depend on the distances Δx between nucleus and protrusion, cell deformation $U = x_p - x_n$, polarity P, and polarity differences $\Delta P = P_i - P_j$ [Fig. 2(b)]. Deriving these terms rigorously from detailed biophysical and biochemical signaling processes and mechanical couplings between cells is currently unfeasible due to the large number of possibly involved molecular mechanisms.

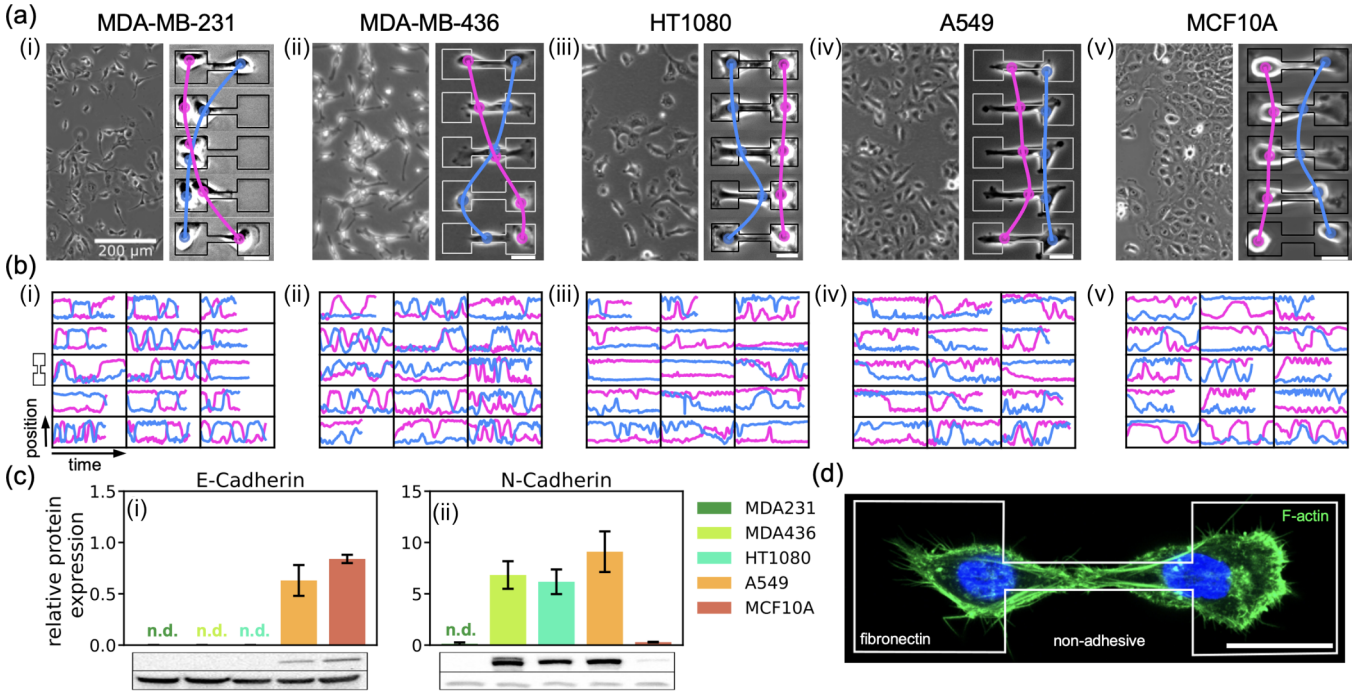

FIG. 3. Studying transient cell collisions on a minimal cell collider. (a) Brightfield microscopy images of (i) MDA-MB-231 cells, (ii) MDA-MB-436 cells, (iii) HT1080 cells, (iv) A549 cells, and (v) MCF10A cells. Left panels show multicellular tissues, scale bar: 200 μm throughout. Right panels show time series of brightfield images of two cells colliding repeatedly while hopping between the islands. Scale bar: 25 μm. (b) Small selection of nucleus trajectories for the five different cell lines. For all panels (i)–(v), horizontal time axis ranges from 0 h to 25 h and vertical position axis ranges from 0 μm to 60 μm. In total, for MDA-MB-231 cells, we consider $N = 185$, for MDA-MB-436 cells $N = 102$, for HT1080 cells $N = 87$, for A549 cells $N = 100$, and for MCF10A cells $N = 251$ cell pairs. Refer to Table 2 in [54] for a full overview of the data. (c) Quantitative western blot analysis of protein expression levels of E-cadherin and N-cadherin in the different cell lines. We show one blot for each protein with its corresponding loading control (β-actin) (Appendix A 6). Error bars indicate the error of the mean (s.e.m.) of triplicate measurements. (d) Two migrating MCF10A cells (F-actin stained in green, nucleus in blue) are confined on a dumbbell-shaped micropattern (white outline), which consists of two islands connected by a bridge. The micropattern is coated with fibronectin and is surrounded by a nonadhesive PLL-PEG layer. Scale bar: 25 μm.

Here we propose a phenomenological approach to construct possible interactions: In an unbiased way, we determine possible candidate cell-cell interaction mechanisms arising from the lowest-order linear couplings of the degrees of freedom that behave like vectors in arbitrary dimensions, obeying rotational and translational symmetry [Fig. 2(b); Appendix B 3]. Specifically, we construct each interaction mechanism as a single term proportional to one of the nine possible linear vectorial couplings [Fig. 2(b)], which contributes to either of the three interaction terms G_n, G_p, or G_P. In this way we derive in total 27 phenomenological cell-cell interaction mechanisms. However, by requiring that interaction mechanisms cannot introduce couplings of the degrees of freedom of the same cell that are already introduced by the single-cell terms (Appendix B 3), we further reduce the set of interactions to 14 candidate mechanisms, providing a broader class of possible mechanisms than those previously considered [37].

Each interaction mechanism is characterized by an interaction strength ε, setting the sign and magnitude of the interaction. In addition, we assume that each interaction decays exponentially with the intercellular distances $|\Delta x|$ on a characteristic interaction range r. Thus, while the direction of the interaction mechanism is determined by the linear vectorial couplings, the spatial structure is nonlinear in the intercellular distances $|\Delta x|$ to limit the interaction range. Altogether, these assumptions lead us to a set of interaction mechanisms ranging from couplings of the nuclei and protrusions via repulsion or attraction to various interactions affecting cell polarity such as alignment and antialignment [Figs. 2(c)–2(e)].

Phenomenologically, all 14 candidate cell-cell interaction mechanisms may contribute to the behavior of interacting cells. Additionally, distinct cell types may exhibit different cell-cell interactions, further exacerbating the challenge to reveal the cell-cell interaction mechanisms underlying behavior. Therefore, we require strong quantitative model constraints from experiments to discover what interaction mechanisms are relevant to describe interaction behavior across a broad range of cell types.

B. Cell-cell collision experiments reveal diversity of interaction behaviors

To detect contact interactions in experiments, we use a high-throughput assay to study the dynamics of homotypic pairs of interacting motile cells. Specifically, we employ a dumbbell-shaped micropattern as a minimal cell collider [8] [Fig. 3(d); Appendix A]. This geometry effectively confines cells in one dimension, isolates cell pairs, and generates

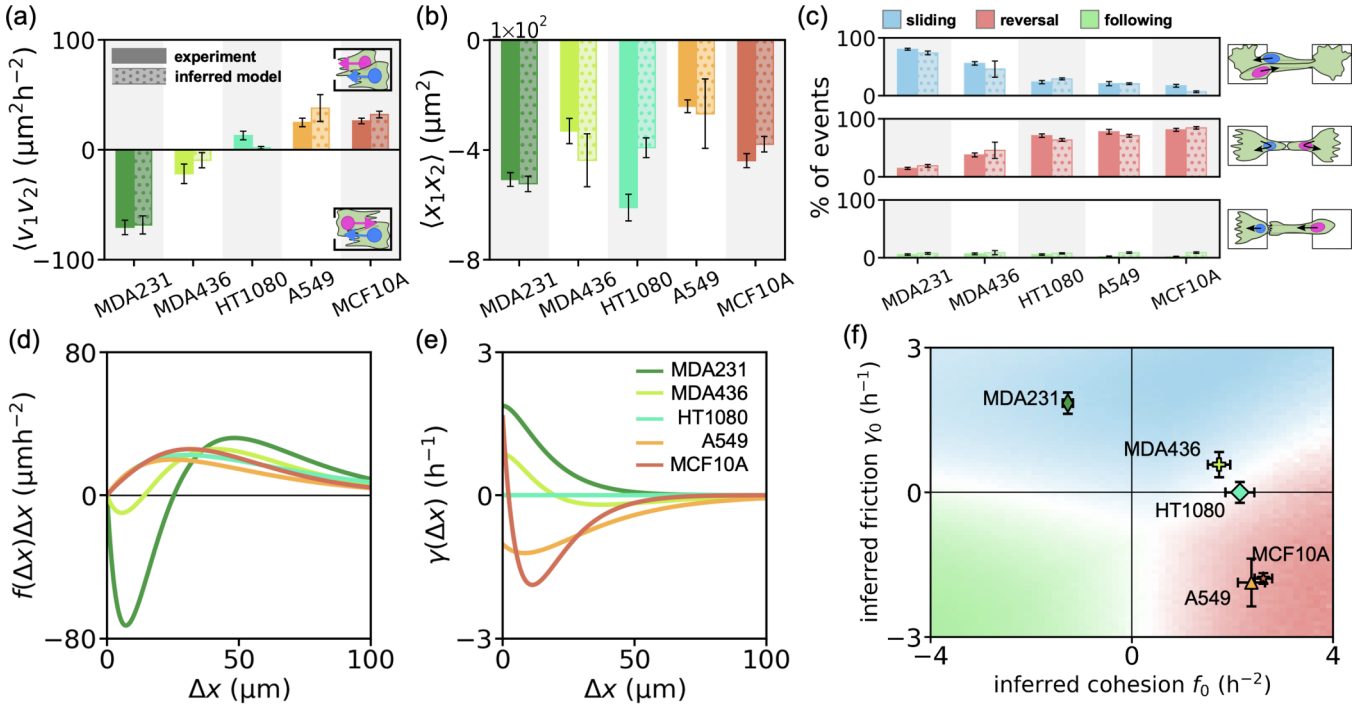

FIG. 4. Quantifying interaction behavior using data-driven inference. (a) Instantaneous velocity cross-correlation $C_V(|t-t'|=0) = \langle v_1 v_2 \rangle_{\text{same}}$ between the two cells when they occupy the same island. Solid bars show experimental results; dotted bars show the prediction of the inferred underdamped description. For panels (a)–(c), error bars show the error of the mean (s.e.m.) obtained from bootstrapping the experimental data. (b) Instantaneous position cross-correlation $C_X(|t-t'|=0) = \langle x_1 x_2 \rangle$. $x=0$ is defined as the center of the dumbbell-shaped micropattern [Fig. 2(a)]. Consequently, negative values indicate cells being on average separated on the two different islands. (c) Percentages of the different collision events for the different cell lines. For all cell lines, differences between inferred model and experiment are usually of the order of the experimental error bars. (d) Inferred cohesion interactions $f(\Delta x)\Delta x$ for the five different cell lines. (e) Inferred friction interactions $\gamma(\Delta x)\Delta v$ of the various cell lines. (f) Interaction behavior space spanned by the amplitudes of the inferred cohesion and friction interactions, f_0 and γ_0. Colors show the dominant collision event predicted by numerical simulations of Eq. (4) with varying f_0 and γ_0 (Appendix B 6). Throughout this figure, results for MCF10A and MDA-MB-231 cells have been obtained from data adapted from [8]. Number of cell pairs used is given in the caption of Fig. 3.

many cell-cell collision events. To capture a variety of cell-cell interactions across cell types in parallel, we consider a range of distinct motile cells from human tissue. We study the two metastatic breast cancer cell lines MDA-MB-231 and MDA-MB-436, the fibrosarcoma cell line HT-1080, the nonmetastatic lung cancer cell line A549, and the epithelial noncancerous breast cell line MCF10A [Fig. 3(a)]. These cell lines express different levels of adhesion proteins and exhibit distinct collective behaviors *in vitro*: MDA-MB-231 cells express neither the cell adhesion molecule E-cadherin nor N-cadherin [Fig. 3(c)] and do not form confluent monolayers. Instead, there are large gaps between individual cells [Fig. 3(a) (i)][60]. The breast cancer cells MDA-MB-436 and fibrosarcoma cells express N-cadherin [Fig. 3(c) (ii)] and also do not form monolayers [Fig. 3(a) (ii), (iii)]. These features are characteristic of a mesenchymal phenotype [61]. In contrast, MCF10A and A549 cells express E-cadherin [Fig. 3(c) (i)] and form confluent monolayers [Fig. 3(a) (iv), (v)], characteristic of an epithelial phenotype [61].

All tested cell lines are motile on our micropattern and repeatedly collide with each other [Fig. 3(a); Movie 1 [54]]. However, the collision behavior of cells is variable with marked differences across the five cell lines. To quantify the behavior of pairs of colliding cells in our experiments, we use a low-dimensional representation of these interaction dynamics by tracking the 1D motion of the nucleus of both cells over time [Fig. 3(b)]. This approach yields a large ensemble of experimental trajectory data reflecting the cell-level dynamics of interacting cells.

To characterize the interaction dynamics of distinct cell types, we use various interaction behavior statistics, computed from the tracked nucleus positions x and numerically estimated nucleus velocities v [8]. First, we quantify how cells coordinate their behaviors in close proximity using a velocity cross-correlation function of two cells occupying the same island: $C_V(|t-t'|) = \langle v_1(t) v_2(t') \rangle_{\text{same}}$ (Appendix B 4). The breast cancer and fibrosarcoma cell lines exhibit negative or very small instantaneous velocity correlations, while the epithelial cells exhibit positive instantaneous velocity correlations [Fig. 4(a)]. Furthermore, we determine a position correlation function $C_X(|t-t'|) = \langle x_1(t) x_2(t') \rangle$. All cell types exhibit negative instantaneous position correlations [Fig. 4(b)], indicating mutual exclusion behavior. To gain insight into how cells navigate each other on longer timescales, we analyze cell-cell collisions [8]. We detect sliding events (two cells interchange positions), reversal events (at least one cell retracts), and following events (both cells transition simultaneously) (Appendix B 4). The two breast cancer cell

lines mostly exhibit sliding behavior, while the epithelial and fibrosarcoma cells mostly exhibit reversal behavior [solid bars in Fig. 4(c)]. Altogether, these results show different dominant interaction behaviors across our cell types with marked differences between epithelial and cancer cell lines.

C. Data-driven inference additionally constrains cell-cell interactions

In addition to the long-time scale interaction behavior of colliding cells, we can constrain our theory for contact interactions on the detailed short-time scale dynamics of the nucleus, which are readily obtained from our experiments [8]. In this way we capitalize on the information contained in the experimental trajectory data over a broad range of timescales. Importantly, the order of the differential equation used to describe the dynamics of interacting cells depends on which degrees of freedom are explicitly included. In contrast to the general overdamped (first-order) model for the coupled dynamics of the nucleus, protrusion, and polarity in Eqs. (1)–(3), a more coarse-grained dynamical description where only the nucleus position x_i and velocity v_i appear explicitly as dynamical degrees of freedom are underdamped (second-order) [49,55,62]. Because a mapping between the overdamped and underdamped descriptions exists by rewriting multiple coupled first-order equations in terms of a higher-order equation (Appendix C 1), we can make progress in constraining our theory of contact interactions by studying the experimentally more accessible underdamped nucleus dynamics.

To capture the motion of the cell nuclei of interacting cells, we assume coupled differential equations of the form [8]

$$\frac{dv_i}{dt} = F(x_i, v_i) + f(\Delta x)\Delta x + \gamma(\Delta x)\Delta v + \sigma\eta_i(t). \quad (4)$$

Here $F(x_i, v_i)$ describes single-cell behavior within the confining geometry, and $\sigma\eta_i(t)$ represents a Gaussian white noise with amplitude σ. Of particular importance here are the cohesion interaction term $f(\Delta x)\Delta x$ and friction interaction term $\gamma(\Delta x)\Delta v$, which capture how the cell nuclei deterministically accelerate depending on the cells' relative separation Δx and velocity Δv. The functions $f(\Delta x)$ and $\gamma(\Delta x)$ encode the sign and spatial structure of the interaction terms and are not *a priori* specified, and $f(\Delta x) > 0$ implies effective repulsion, while negative values indicate effective attraction. In contrast, $\gamma(\Delta x) < 0$ indicates effective cell-cell friction, while positive values indicate effective antifriction. The latter case describes cells accelerating to increase their relative velocity in contrast to slowing down as expected from friction interactions.

To infer the underdamped interaction terms $f(\Delta x)\Delta x$ and $\gamma(\Delta x)\Delta v$ directly from our experimental data, we employ a stochastic inference framework called underdamped Langevin inference (ULI) [50] (Appendix B 5). In fact, our overdamped interaction mechanisms can be mapped onto these underdamped nucleus interaction terms by the aforementioned mapping between overdamped and underdamped descriptions (Appendix C 1). Thus, learning the underdamped interaction terms $f(\Delta x)\Delta x$ and $\gamma(\Delta x)\Delta v$ partially constrains the overdamped interaction terms G_n, G_p, G_P.

Interestingly, we infer qualitatively similar single-cell terms, $F(x_i, v_i)$, across our cell types (Fig. 3 in [54]). By contrast, inferred interaction terms vary strongly: The nuclei of breast cancer cells exhibit short-range attraction and pronounced antifriction dynamics [8] [Figs. 4(d) and 4(e)]. The nuclei of fibrosarcoma cells show repulsive dynamics, but exhibit no detectable friction interaction terms. For the epithelial cell lines, we find repulsion and friction nucleus dynamics. For all cell lines, the learned equations of motion accurately predict long-time scale behavior statistics [dotted bars in Figs. 4(a)–4(c) and Fig. 6 in [54]].

To visualize the inferred dynamics, we plot the inferred amplitudes f_0 and γ_0 of the cohesion and friction interaction terms of the various cells [Fig. 4(f)]. These amplitudes are defined via low-dimensional approximations of the full underdamped interaction terms in Figs. 4(d) and 4(e): $f(\Delta x) \sim f_0 e^{-|\Delta x|/r}$ and $\gamma(\Delta x) \sim \gamma_0 e^{-|\Delta x|/r}$ inferred from the experimental trajectory data (Appendix B 6). The five different cell lines occupy different regions in this interaction behavior space (IBS) with an apparent correlation between inferred effective cohesion and friction interaction terms for the various cell types. Together, this inference procedure reveals a variety of different dynamics ranging from attraction and antifriction of breast cancer cells to repulsion and friction of epithelial cells. This brings us to our key question: How do the diverse dynamics of pairs of motile cells relate to the candidate interaction mechanisms in our overdamped model [Eqs. (1)–(3)]?

D. Connecting cell-cell interaction mechanisms to whole cell-level dynamics

To address this question, we analyze the underdamped nucleus dynamics of colliding cells emerging from our nucleus, protrusion, and polarity interaction mechanisms in our overdamped model. In other words we will numerically perform the aforementioned mapping of the overdamped interaction mechanisms onto the underdamped interactions visualized in the IBS [Fig. 4(f)]. To this end, we numerically simulate cell trajectories using our overdamped model [Fig. 5(a)] and infer the underdamped effective interactions arising from each individual candidate interaction mechanism. Specifically, fixing the interaction range r, varying the amplitude ε of each interaction mechanism, and employing our inference procedure yields a numerically calculated mapping $f_0(\varepsilon)$ and $\gamma_0(\varepsilon)$, which we depict in the IBS [Fig. 5(b)].

Intuitively, varying the nucleus coupling strength ε_{NC} and thus tuning between nucleus attraction and repulsion corresponds mainly to varying the underdamped cohesion parameter f_0 [red curve in Fig. 5(b)]. By contrast, interactions between cell protrusions and polarities give rise to more complicated underdamped dynamics, including friction or antifriction [gray curves in Fig. 5(b)]. Interestingly, polarity-polarity coupling (PPC) can qualitatively predict the experimentally observed correlations between inferred cohesion and friction interactions when tuning between antialignment ($\varepsilon_{PPC} < 0$) and alignment ($\varepsilon_{PPC} > 0$) [blue curve in Fig. 5(b)]. Furthermore, with PPC, we can capture a switch from anticorrelated sliding to correlated reversal behavior when tuning the interaction strength [Figs. 5(c) and 5(e) and Figs. 9(c) and 9(e); Movie 5 in [54]], as experimentally observed [Figs. 4(a) and 4(c)]. These results show that inferring the underdamped nucleus dynamics can provide strong

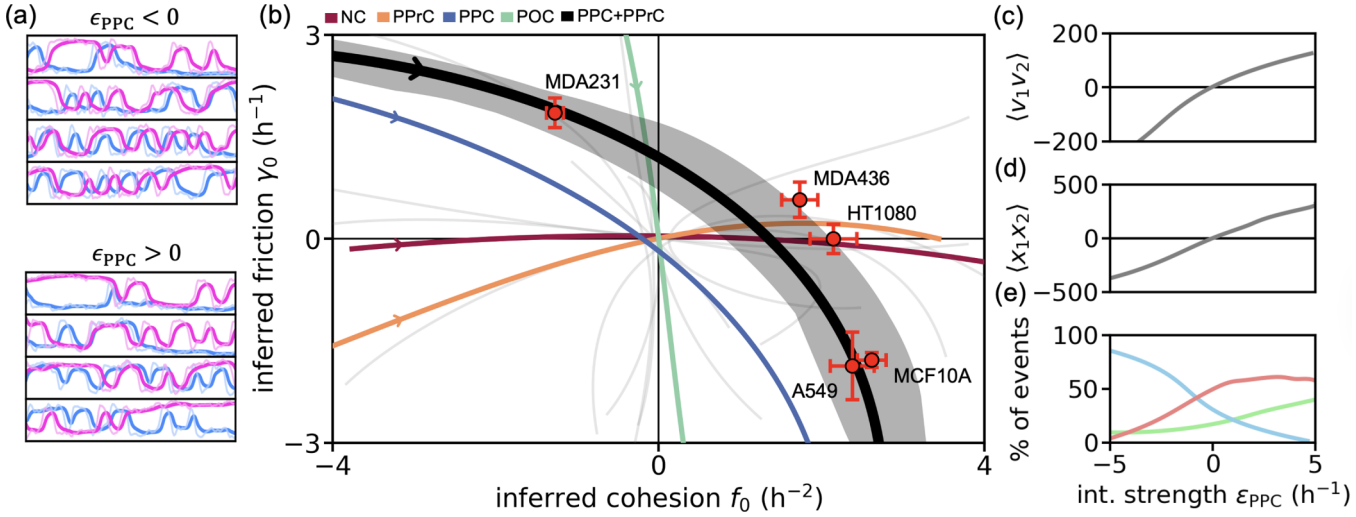

FIG. 5. Visualization of underdamped dynamics of cell-cell interaction mechanisms. (a) Nucleus and protrusion trajectories obtained from simulating our model for contact interactions [Eqs. (1)–(3)] with polarity-polarity coupling interactions for two values of the interaction strength ε_{PPC}. (b) Colored and gray lines show a mapping $[f_0(\varepsilon), \gamma_0(\varepsilon)]$ predicted by the 14 candidate cell-cell interactions while varying the associated interaction strength ε and fixing the interaction range to a value of $r = 15$ μm. The choice of r does not greatly affect these curves (Fig. 11 in [54]). We highlight four different candidate models by colored lines and refer to Fig. 10 in [54] for a complete overview of the mappings of all candidate interactions. Little arrowheads are located at $\varepsilon < 0$ and show the direction in which we increase ε. For all curves, $\varepsilon \approx 0$ is located at the coordinate center of the IBS. Note that these curves are not used in the overdamped inference procedure but only visualize the different dynamics of different interaction mechanisms. Black curve shows the model prediction of PPC + PPrC that best fits the experimental data of the five different cell lines. Shaded black region around that curve indicates the spread of the model result for 20 best parameter combinations (Appendix B 8). Red symbols indicate experimental result. Error bars show the bootstrapped error of the mean of the inferred cohesion and friction coefficients. Number of cell pairs used to obtain these points is given in the caption of Fig. 2. (c)–(e) Behavior statistics predicted by our mechanistic model with the candidate cell-cell interaction PPC. Panel (c) shows the instantaneous velocity correlations, panel (d) shows instantaneous position cross-correlation between the two cells, and panel (e) shows the collision statistics. Curves in panels (b)–(e) are slightly smoothed by fitting splines.

quantitative constraints on our phenomenological nucleus, protrusion, and polarity interaction mechanisms. However, in general, this simple low-dimensional description does not uniquely map back onto the overdamped cell-interaction mechanisms. Thus, to fully constrain the interaction mechanisms, we need to take both the detailed underdamped nucleus dynamics [Figs. 4(d) and 4(e)] and the full behavior statistics [Figs. 4(a)–4(c)] into account.

E. Polarity-polarity coupling tunes between the behavior of various cell types

To quantitatively detect and determine cell-cell interaction mechanisms, we next conduct statistical inference of our complete overdamped phenomenological model [Eqs. (1)–(3)]. The main idea of this inference procedure is to learn which interaction mechanism best captures the experimental underdamped interaction terms [Eq. (4); Figs. 4(d) and 4(e)] and the various long-time scale interaction behavior statistics of each cell type [Figs. 4(a)–4(c)]. To quantify how well each interaction mechanism captures the experimental data, we perform a fit. In this fit, we vary the interaction strength ε and range r of the individual interaction mechanisms, summarizing the set of parameters as Θ. The goal of the fit is to quantify and minimize the difference between the inferred underdamped interactions, as well as between the nucleus velocity correlations and collision statistics of model and experiment using an R^2 value (Appendix B 7). Afterwards, we validate this fit by predicting the position correlation function not used in the fit (Appendix C 4). Finally, we select the best-performing interaction mechanisms based on their fit performance and their predictive power.

Initially, we take a minimal approach and allow only a single possible interaction mechanism. Furthermore, we first consider an individual fit, allowing the interaction mechanism and all parameters to vary between cell types. This inference procedure shows that we can confidently rule out many interaction mechanisms for all cell types [Fig. 6(a)]. Further, we find that epithelial and breast cancer cells are best captured by mechanisms like PPC, PUC, PC, or PrUC that allow either antialignment or alignment between the polarities or deformation vectors of the cells [Figs. 2(d) and 2(e)]. Fibrosarcoma (HT1080) cells are best described by mechanisms that couple either the polarity or the protrusion to the position of the other cell, reminiscent of CIL [Fig. 6(a)].

Next, we consider a global fit, where we use the same interaction mechanism for all cell types, allowing only the interaction strength ε to vary (Appendix B 7). Interestingly, the global fit reveals that both PPC and PUC best capture the dynamics of all cell types, while only varying their coupling strength ε between cells [Fig. 6(b); Figs. 17 and 18 in [54]]. However, both PPC and PUC alone do not quantitatively reproduce the mutual exclusion behavior, the percentage of following events, and the effective repulsion interactions of epithelial and breast cancer cells [Figs. 5(b), 5(d), and 5(e); Figs. 17 and 18 in [54]].

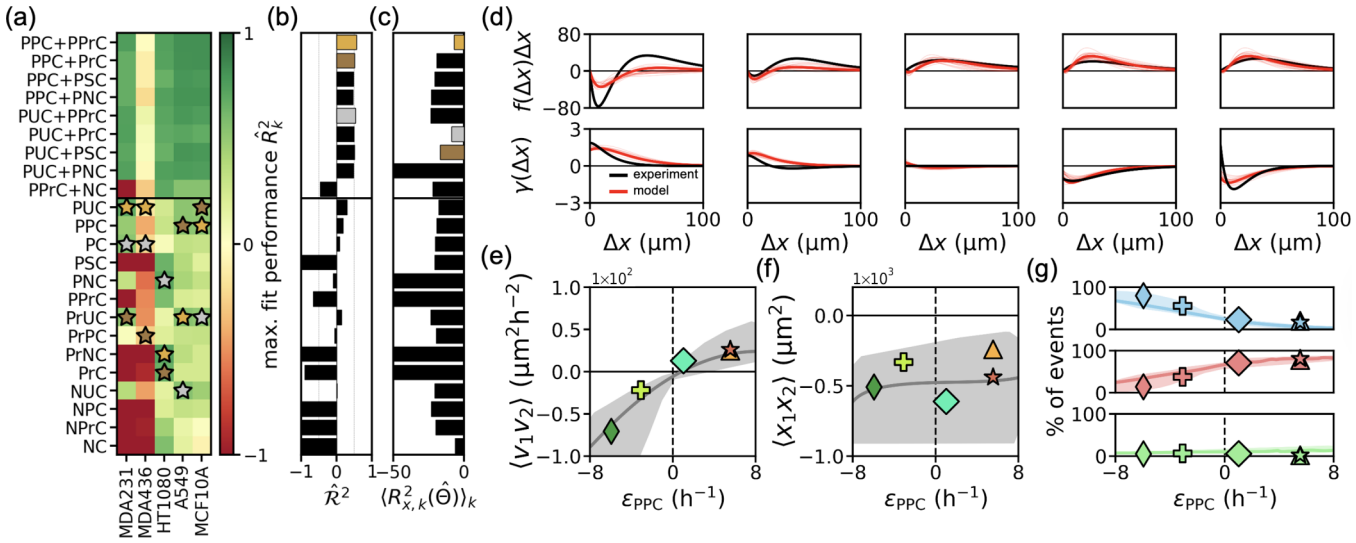

FIG. 6. Overdamped statistical inference of interaction mechanisms from experimental data. (a) Maximum fit performance as quantified by the maximum coefficient of determination (COD) $\hat{R}^2_k$ for each interaction mechanism and each cell type (index k) at the best parameter combination $\hat{\Theta}$ (Appendix B 7). R^2_k is equal to 1 when the model perfectly fits the data and can be negative, indicating bad model performance [Eq. (B10)]. The shown R^2_k is an average over the R^2 values of the differences between model and experiment calculated for the inferred underdamped interactions, the velocity correlation function and the collision statistics, and is then maximized in the fit. After this fit, we predict the position correlation function in panel (c) to validate the fit. Gold, silver, and bronze stars indicate the three best-fitting interaction mechanisms for each cell type. We show multiple combinations of PPC and PUC with mechanisms that well fit the fibrosarcoma HT1080 cells. (b) Maximum fit performance $\hat{\mathcal{R}}^2$ of the global fit (Appendix B 7). The three models giving the highest $\hat{\mathcal{R}}^2$ for all cell types are highlighted in gold, silver, and bronze. (c) Quantification of the goodness of the prediction of the position correlation function of all models. We take the best-fitting parameter combination $\hat{\Theta}$ for each candidate interaction and average R^2 of the position correlation over all cell types to obtain $\langle R^2_{x,k}(\hat{\Theta})\rangle_k$. Again the best performing models are marked with gold, silver, and bronze. (d) Inferred underdamped cohesion and friction interactions for both experiment (black) and best-fitting model (red) for all the various cell lines. Model result is obtained from the best-fitting candidate interaction PPC + PPrC. Thin red lines indicate model results for the 20 best-fitting parameter combinations. (e)–(g) Comparison of behavior statistics between experiment (symbols) and model (curves). Experimental results are plotted at the best-fitting coupling strength ε_{PPC} for each cell type to facilitate comparison between model and experiment. Panels (e) and (f) show the instantaneous velocity and position correlations, respectively. Panel (g) shows percentages of sliding (blue), reversal (red), and following (green) events. Curves in panels (e) and (g) are fits; the gray curve in panel (f) is a prediction as quantified by the gold bar in panel (c). Number of experimental cell pairs used in the fitting procedure is given in the caption of Fig. 3.

To fully capture the dynamics of epithelial and breast cancer cells, additional interaction mechanisms are required. Indeed, combining PPC and PUC with mechanisms that couple the polarity or protrusion to the position of the other cell provides an improved individual and global fit to these cells. From these combinations, PPC and PUC both together with PPrC best capture the experimental interaction behavior across cells [Fig. 6(a)]. However, PPC together with PPrC better captures both the underdamped interactions and the experimental behavior statistics of all cell types [Figs. 6(d)–6(g) and Fig. 10; Movie 6 in [54]]. In particular, PPC + PPrC best predicts the position correlation function, which we have not used in the fit [Fig. 6(c), and Fig. 12]. PPrC with $\varepsilon_{PPrC} > 0$ adds an additional repulsive component to the dynamics, induces mutual exclusion behavior, and suppresses following behavior [Fig. 13 in [54]]. Remarkably, the combined fit of PPC and PPrC is successful even if we allow only the strength of the polarity coupling ε_{PPC} to vary between different cell types [Fig. 6(b), and Fig. 9].

Taken together, we find that the two breast cancer cell lines exhibit antialignment with $\varepsilon_{PPC} < 0$. By contrast, the fibrosarcoma cell line exhibits very weak alignment interactions and its dynamics are dominated by PPrC. Finally, the epithelial cells exhibit alignment interactions with $\varepsilon_{PPC} > 0$, yielding correlated velocities and dominant reversal behavior (Fig. 12 in [54]). These results do not depend sensitively on the details of the overdamped statistical inference procedure (Fig. 16 in [54]), on the single-cell behavior (Fig. 11), or the chosen dumbbell-shaped geometry (Fig. 13). Furthermore, considering examples of different pairs of interactions [Fig. 6(a)] or more than two (Fig. 19 in [54]) does not provide a better description of the dynamics of these cell lines. Taken together, PPC and PPrC enable cells to adapt their polarity relative to both polarity and position of the other cell, providing a large range of possible behavioral responses to cell-cell collisions. Thus, with these interaction mechanisms, we can quantitatively capture a wide variety of interaction behaviors across a broad range of distinct cell types.

F. Modeling molecular perturbations of cell-cell interaction pathways

We hypothesize that the varying polarity-polarity coupling strength across our cell types originates in different molecular expression profiles of several relevant proteins [Fig. 3(c)]. To test this, we disrupt specific cell-cell interaction

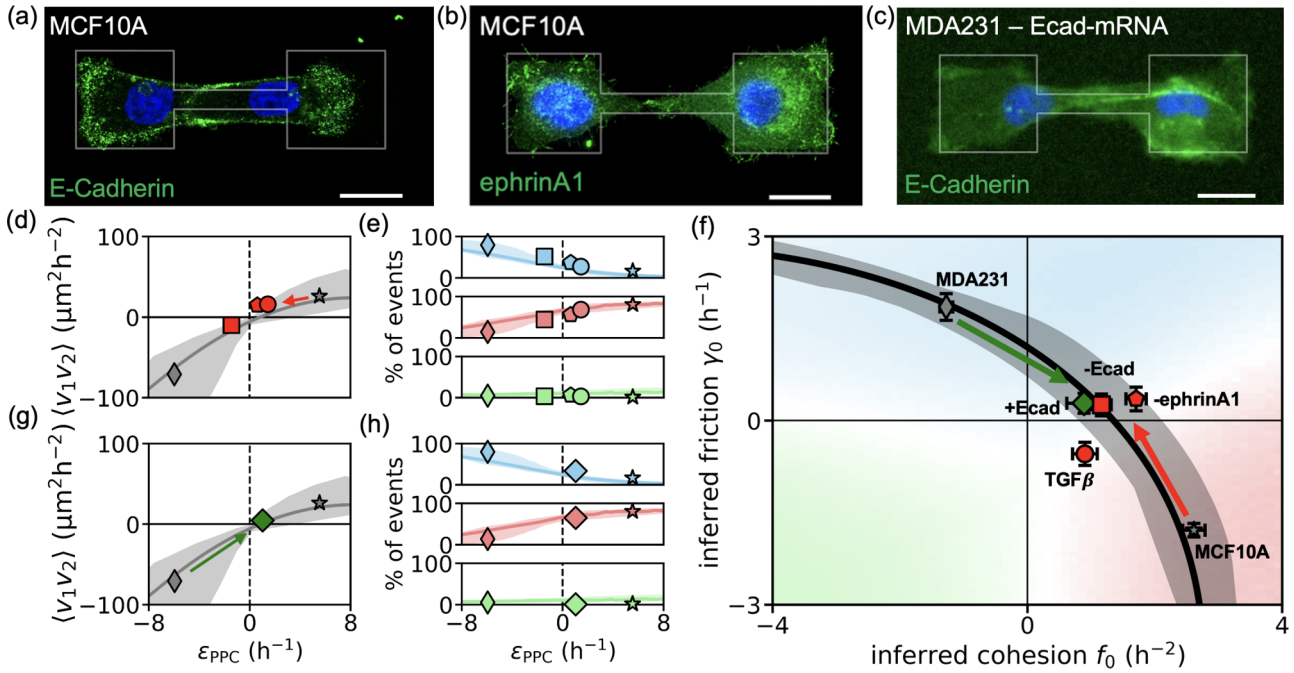

FIG. 7. Capturing molecular perturbations with interaction mechanisms. (a), (b) Fluorescence images of E-cadherin expression (a) and ephrin A1 expression (b) in MCF10A cells. (c) E-cadherin expression in transfected MDA-MB-231 cells while confined on the dumbbell-shaped micropattern. Nuclei stained in blue and scale bar for all panels (a)–(c): 20 μm. (d), (g) Instantaneous velocity cross-correlation between two cells when they occupy the same island. Throughout, the solid lines indicates our model result obtained from simulating the phenomenological model with PPC and PPrC varying only $\varepsilon_{\mathrm{PPC}}$ but using the parameters obtained from the global fit to the various cell lines as shown in Fig. 6. Symbols show experimental results of the various perturbed cell lines plotted at the best-fitting interaction strength $\varepsilon_{\mathrm{PPC}}$. (e), (h) Percentages of the three different collision events for the five different cell lines. Again, symbols show experimental results and solid line shows model result. (f) Interaction behavior space shows the change of the dynamics of the MCF10A and MDA-MB-231 cell lines due to various molecular perturbations. Solid black curve indicates model result from the best global fit of the experimental dynamics of the various cell lines as presented in Fig. 5(c). In total, for wild-type MDA-MB-231 cells, we consider $N = 185$, for MDA-MB-231 + Ecad cells, $N = 63$, for MCF10A-Ecad cells $N = 89$, for MCF10A-ephrin A1 cells $N = 103$, for MCF10A TGFβ cells $N = 101$, and for wild-type MCF10A cells $N = 251$ cell pairs.

pathways through molecular perturbations, which we would then expect to tune the polarity-polarity coupling strength. Specifically, we use antibodies to block E-cadherin bonds that are established upon contact between two epithelial MCF10A cells [Fig. 7(a)]. These bonds are known to be adhesive and are crucial for collective cell migration of epithelial tissues [17]. In addition, we consider perturbations of Eph-ephrin interactions, which directly regulate cell polarity through up- or downregulation of polarity cues such as Cdc42 or RhoA [29]. Epithelial MCF10A cells express the ligand ephrinA1 [Fig. 7(b)], which we block using antibodies. Furthermore, we investigate the behavioral shift in MCF10A cells undergoing an epithelial-mesenchymal transition (EMT). This transition changes the coordinated interaction behavior of epithelial cells to that of more individually migrating cells [16,31,63]. Finally, we induce E-Cadherin in the breast cancer cell line MDA-MB-231 via mRNA transfection [Fig. 7(c); Secs. 2.4 and 3.2 in [54]]. Importantly, all perturbations do not greatly affect single-cell behavior (Fig. 4 in [54]), and the resulting interaction dynamics are insensitive to the used transfection protocol (Fig. 5 in [54]).

In all these perturbation experiments (Movies 3 and 4 in [54]), we can capture the changing interaction dynamics with our phenomenological model and the combination of polarity-polarity coupling and polarity-protrusion repulsion. Specifically, we observe that blocking of E-cadherin and Eph-ephrin pathways, as well as EMT in epithelial MCF10A cells, reduces the polarity-polarity coupling strength $\varepsilon_{\mathrm{PPC}}$, thereby inhibiting polarity alignment interactions between MCF10A cells. Consequently, these perturbations reduce positive velocity correlations [Fig. 7(d)], the number of reversal events observed in MCF10A cells [Fig. 7(e)], and the amplitudes of the inferred underdamped cohesion and friction interactions [Fig. 7(f)]. Similarly, our model captures the changing dynamics upon transfecting MDA-MB-231 breast cancer cells with E-cadherin by a reduction of polarity-polarity coupling [Fig. 7(f)], indicating that E-cadherin can inhibit anti-alignment interactions in these cells. The E-cadherin transfected cells no longer exhibit significant velocity correlations [Fig. 7(g)] and fewer sliding events [Fig. 7(h)]. Taken together, our overdamped model Eqs. (1)–(3) captures the changing dynamics of cell-cell interactions upon molecular perturbations of the underlying interaction pathways by tuning the strength of the polarity-polarity coupling mechanism.

III. DISCUSSION

Our data-driven theoretical approach reveals that from a set of 14 phenomenologically possible interaction mechanisms

only polarity-polarity coupling (PPC) and polarity-protrusion coupling (PPrC) dominate, suggesting that these mechanisms are conserved mechanisms underlying contact interactions of a broad range of cell types. This insight provides a step towards extending the successful generalization of the biophysical principles of single-cell migration [14,33–36] to interacting cells. We speculate that evolution has selected from the many possibilities couplings of cell polarity dynamics to both the position and polarity of other cells as versatile, effective, and sufficient interaction modes for a cell to sense the location and the motility of other cells.

With our approach, we identify several interaction modes underlying the collision behavior of mesenchymal and epithelial cells. First, we identify antialignment interactions between cell polarity as a novel interaction mode, which promotes sliding behavior of breast cancer cells. This interaction appears to be reminiscent of contact enhancement of locomotion [23] and has been associated with doublet rotations of MDCK cells [37,64]. We show that activating expression of E-cadherin in breast cancer cells inhibits antialignment in cancerous MDA-MB-231 cells, which may be interpreted as a partial reversion of their mesenchymal phenotype [65]. However, further experimental work is needed to uncover the underlying biomolecular interaction mechanisms.

Second, we find polarity alignment, which promotes correlated reversal behavior of epithelial cells. This interaction mode may be effectively capturing mechanisms of epithelial tissue migration [17,40,66–68], where so-called E-cadherin fingers have been associated with signaling the state of polarity to other cells [11]. Polar order between neighbors could arise from such signaling or may be a consequence of couplings of mechanical stress to cell polarity [17,66,69]. Such mechanotransduction [70] has been observed in epithelial cells [71,72] and blocking the function of the key epithelial cell-cell adhesion molecule E-cadherin [30,63,73] reveals that mechanical coupling between cells is indeed critical to establish polarity alignment.

Finally, we find PPrC, which causes polarities to grow away from the other cell, similar to contact inhibition of locomotion [9,27]. We show that this interaction mode is essential for capturing the mutual exclusion behavior of all our cells, which also exhibit similar levels of this interaction. This observation may be related to CIL being observed in a wide range of different cell types [18,28,44,63].

Polarity-polarity coupling has been invoked in phenomenological models for active matter [74] and captures the tissue-scale flocking behavior of epithelial MDCK cells on 1D tracks [40,48,68]. We showed in an unbiased way that such polarity interactions accurately capture the collision dynamics of epithelial and breast cancer cells on a two-cell level. In fact, in Appendix C 6, we demonstrate that our model for contact interactions constrained on data from pairs of colliding cells can be readily generalized to an unconstrained 2D collective setting: Polarity-polarity coupling in combination with polarity-protrusion coupling with similar parameters as those learned from two colliding cells can qualitatively capture the ordered collective front migration behavior of epithelial cells [Figs. 14(e) and 14(f)], as observed in experiments [60]. We qualitatively capture that cell speeds are highest at the leading edge of the spreading epithelial monolayer [Fig. 14(h)], as measured in experiments [39]. With small adjustments to the parameters, necessary possibly due to different experimental conditions, we even quantitatively capture the distribution of velocities of collectively moving epithelial cells [Figs. 14(e), 14(f), and 14(h)]. These results indicate that although model parameters may vary between experiments, cell-cell interactions learned from transient pairwise collisions can be generalized to collective settings and thus may underlie collective cell migration in small migrating epithelial cell clusters [40–42,48], 2D epithelial tissues [3,66,75,76], or collective rotations of 3D organoids [77].

In contrast to the ordered collective behavior of epithelial cells, the loss of polar order is associated with cancer progression and cancer cells exhibit nonaligned disordered motion in 2D sheets [32,60]. In Appendix C 6, we show that the absence of large-scale order in collective front migration experiments with breast cancer cells [60] can be captured by the combination of polarity antialignment and polarity-protrusion coupling, which we infer from the two-cell collision dynamics of these cells [Figs. 14(d) and 14(f)]. These results indicate that polarity antialignment could potentially lead to disordered cell spreading [23] or fluid-like disordered collective behavior [78–80]. Taken together, our phenomenological theory could give insight into how contact interactions at the cellular level control collective cell migration of epithelial and mesenchymal cells.

ACKNOWLEDGMENTS

We thank Johannes Flommersfeld, Bram Hoogland, and Ricard Alert for helpful discussions. We thank Gerlinde Schwake for producing the E-cadherin mRNA. This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), Project-ID 201269156 - SFB 1032 (Project B01 and B12).

T.B. and E.B. contributed equally to this work. E.B., T.B., and J.O.R. designed experiments; E.B. and G.L. performed experiments; T.B., D.B.B., and E.B. analyzed data; T.B. and D.B. developed the model and performed simulations and statistical inference; T.B., E.B., D.B.B., J.O.R., and C.P.B. interpreted the results and wrote the manuscript. J.O.R. and C.P.B. conceived this study and led the research. All authors edited and approved the manuscript.

DATA AVAILABILITY

The data that support the findings of this article are not publicly available upon publication because it is not technically feasible and/or the cost of preparing, depositing, and hosting the data would be prohibitive within the terms of this research project. The data are available from the authors upon reasonable request.

APPENDIX A: EXPERIMENTAL METHODS

1. Micropatterning and sample preparation

For the cell collision experiments, we employ a micropattern structure with a dumbbell-shaped design with two 35 × 35 μm squares connected by a 40 × 7 μm bridge [8]. For micropatterning, we employ a photopatterning technique using

the PRIMO module (Alvéole). For background passivation of the μ-dish (ibidi), a drop of 0.01% (w/v) PLL (Sigma-Aldrich) is added and incubated for 30 min. Afterwards, the sample is rinsed with HEPES buffer (Sigma-Aldrich) and 100 mg/ml PEG-SVA (LaysanBio) is incubated for 1 h at room temperature. The passivated dish is photopatterned by employing the PRIMO module mounted on an automated inverted microscope (Nikon Ti Eclipse). The photoinitiator PLPP (Alvéole) is added to the dish. The dumbbell-shaped pattern was placed on top of the dish via the Leonardo software (Alvéole) and illuminated with UV light with a dose of 15 mJ/mm^2. The dish is washed with milliQ water and rehydrated with PBS for 5 min followed by an incubation with 20 µg/ml of labeled Fibronectin-Alexa647 (YO-protein, ThermoFisher) for 15 min at room temperature. For the antibody-blocking experiment fibronectin micropatterns are made by microplasma-initiated protein patterning as described in [55].

2. Cell culture

In this study we analyzed the behavior of the cell lines MDA-MB-231, MDA-MB436, HT-1080, A549, and MCF10A. The individual culture condition are shown in Table 1 in [54]. Cells are grown at 37 °C in an atmosphere containing 5% CO_2. For passaging, cells are being washed and treated with accutase for 5 min. The cell solution is centrifuged at 800 r.c.f. for 3 min, and the cells are resuspended in medium. Approximately 10 000 cells are added into the micropatterned μ-dish and left to adhere for up to 4 h in the incubator. After this incubation period, the medium is exchanged for phenol-red-free medium, and 25 nmol Hoechst 33342 (Invitrogen) is added to stain the nuclei when needed.

3. Inhibitors and antibody blocking

In order to block the function of E-cadherin or ephrin-A1, blocking antibody CD324 (functional grade, Invitrogen) or anti-ephrin-A1 antibody (Invitrogen, ThermoFisher) was added to the dish after cells adhered to the pattern at a concentration of 5 µg/ml and 1 µg/ml, respectively. After an incubation period of 1 h, time-lapse measurements were started. EMT was induced by 10 ng/ml treatment with TGFβ (ThermoFisher) for up to 7 days.

4. Transfection

Prior to the transfection, cells were seeded in an μ-dish (ibidi) without patterning. At 90% confluency the cells were transfected with LipofectamineTM 2000 (Invitrogen, Germany) complexes containing E-cadherin mRNA (Sec. 2.4 in [54]). First, 2 µl Lipofectamine 2000 reagent is mixed with 398 µl OptiMEM (Invitrogen, Germany) and incubated for 5 min at room temperature. Simultaneously, 2 µl mRNA (173 ng/µl) is diluted in 198 µl OptiMEM. The mRNA mix is added onto 200 µl of the Lipofectamine 2000 dilution and incubated for 20 min at room temperature. Cells are washed once with OptiMEM, and then the lipoplexes are added to the dish. After 1 h, cells are washed again and reincubated with normal medium. For control experiments, lipoplexes were created with either GFP-mRNA or substituted with milliQ water at the same ratio. For more details see Sec. 2.4 in [54]. Note that we consider the control experiments with only GFP transfection and show that our results do not depend sensitively on the transfection protocol (Fig. 5 in [54]).

5. Microscopy and cell tracking

All measurements are performed in time-lapse mode for 48 h on a Nikon Eclipse Ti microscope or on a Nikon Eclipse Ti2 microscope using a 10× objective. The samples are kept in a heated chamber (Okolab) at 37 °C at 5% CO2 throughout the measurements. Every 10 min a bright-field image and a fluorescence image of the stained nuclei are acquired. Cell tracking of the Nuclei is performed using TrackPy. Track length varied between 8 h and 48 h. For more details see Sec. 2.3 in [54]. To track MCF10A cells on the unconfined 2D pattern (Sec. 2.5 in [54]), we use LIM tracker [81].

6. Western blots

Cells were harvested and lysed in RIPA lysis buffer supplemented with 1M PMSF and a protease inhibitor cocktail (ThermoFisher). Lysates were centrifuged at 14 000 g for 20min at 4 °C. Supernatant was transferred and stored at −80 °C. Protein concentration were determined by Bradford assay, and an equal amount of protein was loaded onto the precast SDS-gel (BioRad). Proteins were separated by gel electrophoresis. The transfer was performed on Immuno-Blot polyvinylidene difluoride (PVDF) membranes (BioRad) with the Transblot turbo transfer system (BioRad) for 7 min. After blocking for 1 h with 5% nonfat dried milk (ThermoFisher) in PBS 0.1% Tween 20 (Roth) the membranes were probed with primary antibodies mouse anti-ECAD (ThermoFisher) (1:1000) and rabbit anti-NCAD (ThermoFisher) (1:2000) overnight at 4 °C. Secondary antibodies HRP-anti-mouse IgG (1:10 000) and HRP-anti-rabbit IgG (1:10 000) were incubated on the membrane for 1 h at room temperature. Development was performed using Pierce western ECL substrate (ThermoFisher) and the ChemiDoc Imaging System (BioRad). The intensity of the band was quantified via densitometry using ImageJ, and the protein amount was normalized to a β-actin loading control on the same membrane.

7. Immunohistochemistry

The cells were fixed after the experiment with 4% PFA for 10 min. Cells were washed three times with PBS. Afterwards the cells were permeabilized for 5 min in 0.1% Triton-X-100 solution (Sigma) and blocked for 1 h in cold 4% BSA (ThermoFisher). The cells were rinsed once with cold 1% BSA. The excess liquid was removed, and the cells were subjected to primary antibodies mouse anti-ECAD (ThermoFisher) (1:100) and rabbit anti-ephrinA1 (ThermoFisher) (1:100) diluted in 1% BSA overnight at 4 °C. Three washes with 1% BSA were carried out before the adding of the secondary antibody AlexaFluor 488 goat antimouse (ThermoFisher) (1:1000) or Alexa 647 goat antirabbit (ThermoFisher) (1:1000) for 1 h in the dark. The solution was then removed and washed three times with PBS. F-actin staining was done with rhodamine-phalloidin (ThermoFisher, 1:1000).

Cells were imaged on a Confocal LSM980 microscope using the airy scan mode with a 40× water immersion objective.

APPENDIX B: THEORETICAL METHODS

1. Model development

Based on previous work [38,39,46,82,83], we employ a generalized active particle model to describe pairs of interacting cells. This model aims to describe the large scale cell-level dynamics in terms of three minimal key degrees of freedom. First, we describe the nucleus, marking the COM of the cells. Second, we include the position of the leading protrusion, allowing us to capture the dynamic and elongated shape of cells. Third, we capture cell polarity as an internal stochastic degree of freedom, providing the active pushing force onto the leading protrusion. To this end, we generalize a previously employed mechanistic model for single migrating cells in confinement [33]:

$$\zeta_n \dot{x}_{n,i} = -\frac{k}{\gamma(x_{n,i})}(x_{n,i} - x_{p,i}) + G_n(\Delta x, U, P, \Delta P), \quad \text{(B1)}$$

$$\zeta_p \dot{x}_{p,i} = k(x_{n,i} - x_{p,i}) + P_i(t) + G_p(\Delta x, U, P, \Delta P) + F_{\text{boundary}}(x_{p,i}), \quad \text{(B2)}$$

$$\dot{P}_i = -\alpha(x_{p,i})P_i - \beta P_i^3 + G_P(\Delta x, U, P, \Delta P) + \sigma\eta_i(t). \quad \text{(B3)}$$

The first terms on the right-hand side of Eqs. (B1) and (B2) are defined as F_n and F_p in Eqs. (1) and (2). These single-cell terms describe how the nucleus and the protrusion of the same cell are coupled by a linear spring. Furthermore, $\gamma(x_{n,i})$ is a dimensionless rescaling factor of the friction coefficient ζ_n that depends on the position of the cell nucleus with a minimum when the cell nucleus is on the bridge [Fig. 8(a)]. This models the reduced adhesive area accessible to the cell on the bridge. These two features have been shown to faithfully capture confined cell migration on our dumbbell-shaped micropattern [33]. In agreement with previous work [33], we choose

$$\gamma(x_{n,i}) = \frac{1-\gamma_{\min}}{2}\left[1 - \cos\left(\frac{x_{n,i}\pi}{L_{\text{system}}}\right)\right] + \gamma_{\min}, \quad \text{(B4)}$$

where L_{system} is half of the size of the dumbbell-shaped micropattern and $\gamma_{\min}$ is the minimum rescaling factor reached while the nucleus is on the bridge [Fig. 8(a)]. The cell's protrusion is confined on the micropattern by a boundary force $F_{\text{boundary}}(x_{p,i})$, which represents a soft repulsive force at the boundaries of the micropattern. In this model, the polarity of the cell is guided by the geometry of the micropattern [33,34]. Specifically, the first terms on the right-hand side of Eq. (B3) is defined as F_P in Eq. (3). It describes that the polarity switches from a negative feedback loop on the island to a positive feedback loop on the bridge of the micropattern, causing the polarity to grow into the bridge as cells transitions between the islands [33]. This is implemented by the function $\alpha(x_{p,i})$, which switches sign dependent on $x_{p,i}$:

$$\alpha(x_{p,i}) = -\frac{\alpha_0 - \alpha_{\min}}{2}\cos\left(\frac{x_{p,i}\pi}{L_{\text{sys}}}\right) + \frac{\alpha_{\min} + \alpha_0}{2}, \quad \text{(B5)}$$

where $\alpha(x_{p,i}) = \alpha_0$ on the island and $\alpha(x_{p,i}) = \alpha_{\min} < 0$ on the bridge [Fig. 8(b)]. The higher order term $-\beta P_i^3$

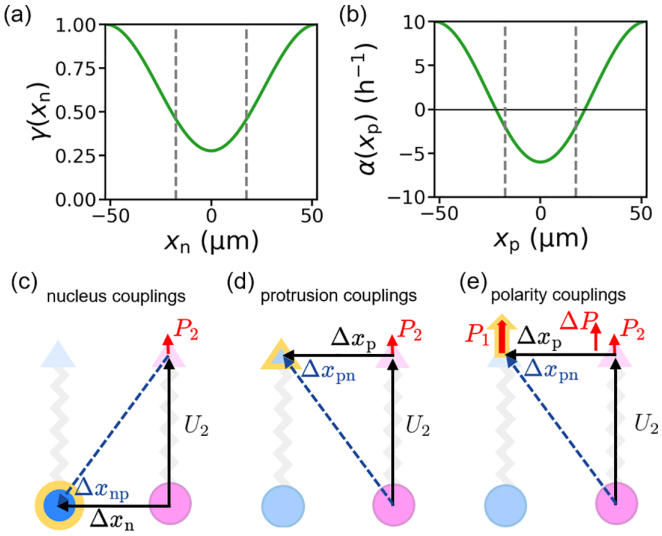

FIG. 8. Model details (a) Rescaling factor $\gamma(x_n)$ used in the model. Vertical dashed lines indicate the boundary of the two islands and the bridge of the dumbbell-shaped micropattern. (b) Geometry-dependent growth rate of the polarity $\alpha(x_p)$. (c)–(e) Schematic of all couplings between the nucleus of cell i and cell j (c), between the protrusion of cell i and cell j (d), and between the polarity of cell i and DOFs of cell j (e). With gold color we highlight the degree of freedom of cell 1 that gets coupled to the other cell.

prevents unbound growth of the polarity and gives rise to a preferred polarity when $\alpha(x_{p,i})$ is negative. The interaction terms G_n, G_p, and G_P encode phenomenological interaction mechanisms, which we derive in an unbiased way. The term $\sigma\eta_i(t)$ encodes uncorrelated Gaussian white noise, modeling the inherent stochasticity of cell polarity and cell migration. For details of the numerical implementation refer to Sec. 4.2 in [54].

2. Single-cell parameters

Throughout, we make use of a set of parameters ($\sigma = 100\ \mu\text{m h}^{-\frac{1}{2}}$, $k_n = k/\zeta_n = 0.6\ \text{h}^{-1}$, $k_p = k/\zeta_p = 1.2\ \text{h}^{-1}$, $\beta = 0.0001\ \mu\text{m}^{-2}\,\text{h}$, $\alpha_0 = 10\ \text{h}^{-1}$, $\alpha_{\min} = -6\ \text{h}^{-1}$, $\gamma_{\min} \approx 0.25$, $L_{\text{system}} = 52.5\ \mu\text{m}$) which has been constrained for MDA-MB-231 cells using a data-driven inference approach and best describes the detailed dynamics of the hopping behavior of single MDA-MB-231 cells on the dumbbell-shaped micropattern used here [33]. In Appendix C 3 we describe how these parameters vary across cell types and investigate the role of single-cell aspects on the interaction dynamics of cells.

3. Phenomenological derivation of interaction mechanisms

To construct phenomenological interaction mechanisms, we first assume that our model obeys rotational and translational symmetry. Thus, cell-cell interactions in our model can not depend on the absolute position of cells and can depend only on quantities in our model that behave as vectors in arbitrary dimensions. These quantities are the intercellular distances $\Delta x_n = x_{n,i} - x_{n,j}$, $\Delta x_{np} = x_{n,i} - x_{p,j}$, $\Delta x_p = x_{p,i} - x_{p,j}$, $\Delta x_{pn} = x_{p,i} - x_{n,j}$, the two deformation vectors of the cells $U_i = x_{p,i} - x_{n,i}$, $U_j = x_{p,j} - x_{n,j}$, and the polarities P_i,

P_j [Fig. 2(b)]. We further consider differences between these vectorial quantities only if the vectorial quantities themselves are not already differences of the degrees of freedom. This adds the potentially biophysically relevant polarity difference $\Delta P = P_i - P_j$. Further, we employ a minimal approach and propose that interaction terms are linear in the vectorial quantities and impose a characteristic interaction range by assuming that interactions exponentially decay on a length scale r. Thus, we can use the following functional dependencies to construct interaction mechanisms:

$$\varepsilon e^{-|\Delta x|/r}\Delta x, \tag{B6}$$

$$\varepsilon e^{-|\Delta x|/r}U, \tag{B7}$$

$$\varepsilon e^{-|\Delta x|/r}P, \tag{B8}$$

$$\varepsilon e^{-|\Delta x|/r}\Delta P. \tag{B9}$$

Here ε is the interaction strength in our phenomenological model and Δx is either Δx_{n}, Δx_{np}, Δx_{p} or Δx_{pn}. Also P and U are either equal to P_i or P_j and U_i or U_j, respectively. To construct interaction mechanisms and encode them in G_{n}, G_{p}, and G_{P}, we need to specify the allowed couplings and thus the allowed dependencies of the interaction terms. We assume that interactions acting on the positional degrees of freedom (G_{n}, G_{p}) do not introduce additional couplings between the degrees of freedom of the same cells that are not already introduced by the single-cell terms. In practice, this means that G_{n} can depend only on Δx_{n}, Δx_{np}, U_j, or P_j (and not on the vectorial quantities that involve the protrusion position or polarity of cell i). Thus, we couple the nucleus of cell i to the nucleus, protrusion, deformation, and polarity of cell j [Fig. 8(c)]. Likewise, G_{p} depends only on Δx_{p}, Δx_{pn}, U_j, or P_j coupling the protrusion of cell i to the nucleus, protrusion, deformation, and polarity of cell j [Fig. 8(d)]. For the polarity couplings encoded in G_{P} we do not make assumption (iv). This is because the polarity is an internal degree of freedom. Thus, in order to fulfill the assumption that interactions decay in space, we have to allow G_{P} to depend on the relative position of cells even though this makes G_{P} dependent on the DOFs of cell i. This means that G_{P} may depend on Δx_{p} or Δx_{pn}, U_j, P_i, P_j, ΔP, coupling the polarity of cell i to the nucleus, protrusion, deformation, and polarity of cell j as well as to its own polarity [Fig. 8(e)]. These assumptions give rise to 14 different interaction mechanisms, which we show in (Fig. 2) and describe in detail in Sec. 4.1 in [54]. Note that for each candidate mechanism, we consider only one of the three interaction terms G_{n}, G_{p}, G_{P} to be nonzero. Only when combining the mechanisms can two of the terms be nonzero.

4. Behavior statistics

To quantify interaction behavior of pairs of motile cells, we compute several behavior statistics introduced in the literature [8]. First, we find the correlation functions $C_V(|t - t'|) = \langle v_1(t)v_2(t')\rangle_{\mathrm{same}}$ and $C_X(|t - t'|) = \langle x_1(t)x_2(t')\rangle$. For both these correlation functions, we define $t' = t + \Delta$ and then average the products of the velocities or positions of the two cells over time t and all pairs of cells. For C_V, we condition our averaging on both cells being located on the same island. For easy visualization, we show in the main text figures the instantaneous correlations ($|t - t'| = 0$) and show the full result in Fig. 6 in [54]. Furthermore, we detect collision events in our cell trajectories by defining a threshold $\Delta x < \Delta x_c$, where Δx is the intercellular distance. Then we analyze the trajectories within a time frame of dT following the first time when $\Delta x < \Delta x_c$. If cells simultaneously transition from one island to the other, we detect a following event. If cells switch positions at least once during dT, we detect a sliding event. Finally, if cells do not switch positions during dT, we detect a reversal event. The collision statistics is robust against the choice of parameters of the detection algorithm [8].

5. Data-driven inference

We perform a data-driven inference approach called underdamped Langevin inference (ULI) to learn effective dynamical interaction terms that govern the short time scale dynamics of the nuclei of interacting cells [8,49,50]. These dynamics implicitly capture the dynamics of protrusion and polarity [49]. This approach involves proposing coupled underdamped Langevin equations for the nucleus trajectories $x(t)$ of the two cells [8] [Eq. (4)]. The key idea of this approach is to estimate from the experimentally measured nucleus trajectory $x(t)$ both the instantaneous velocity $\hat{v}(t)$ and the instantaneous acceleration $\hat{a}(t)$, while considering errors introduced by the discrete sampling and possible measurement errors. Then, assuming that the Langevin equation can capture the dynamics of these cells, we can rigorously infer the deterministic single-cell term $F(x_i, v_i)$ and the interaction terms $f(\Delta x)\Delta x$ and $\gamma(\Delta x)\Delta v$ on the right-hand side of Eq. (4) using stochastic estimators [50]. Briefly, this is done by fitting the experimental accelerations, measured over the phase space of our system, by the deterministic terms expanded in sets of basis functions. Throughout, we expand the interaction terms into exponentials of the form $b_n(|\Delta x|) = e^{-|\Delta x|/nr}$, where $n = 1, \ldots, N$. Thus, for instance $f(\Delta x) \approx \sum_n u_n b_n(|\Delta x|)$ and the coefficients u_n get estimated rigorously from the trajectory data [50]. This procedure is robust against the choice of basis function [8], but the experimental data are best captured at certain values of N. For MCF10A, MDA231, and MDA436 cells, we use $N = 3$ and $r = 20$ μm. For A549 cells we use $N = 2$ and $r = 25$ μm, and for HT1080 cells, we use $N = 1$ and $r = 30$ μm. Note that for HT1080 cells, we found that the best Langevin equation is one without the term $\gamma(\Delta x)\Delta v$.

6. Interaction behavior space

For the position of the experimental cell lines in the interaction behavior space (IBS), we perform underdamped Langevin inference defined in Appendix B 5 with $N = 1$ and $r = 30$ μm. Thus, we define $f(\Delta x) = f_0 e^{-\Delta x/r}$ and $\gamma(\Delta x) = \gamma_0 e^{-\Delta x/r}$ and infer the amplitudes f_0 and γ_0 using ULI. Finally, to predict the interaction behavior space (IBS), we fix $N = 1$ and $r = 30$ μm and vary manually f_0 and γ_0 for $f(\Delta x)$ and $\gamma(\Delta x)$, respectively. For each parameter combination, we simulate trajectories and find the dominant collision behavior, which we depict as color map in [Fig. 4(g)].

7. Overdamped statistical inference procedure

To perform our statistical inference procedure of our overdamped model to the experimental data, we perform a fitting procedure. Specifically, we sweep over the interaction strength ε and interaction range r of each candidate interaction mechanism as well as for each pair of mechanisms. We call the set of parameters Θ. For each parameter combination, we predict the velocity cross correlation $C_V(|t-t'|)$, the position cross-correlation $C_X(|t-t'|)$, and find the collision distribution. Note that in the fitting procedure, we consider only $|t-t'|<\tau$ to avoid fitting the correlations that are close to zero in the experiment. For C_V, we choose $\tau = 0.5\,\mathrm{h}$ and for C_X we choose $\tau = 10\,\mathrm{h}$. This choice does not greatly influence the final results of the inference procedure described in Figs. 6(a) and 6(b) as shown in Fig. 16 in [54]. Furthermore, we perform our data-driven inference method to find the underdamped cohesion $f(\Delta x)\Delta x$ and friction interactions $\gamma(\Delta x)\Delta v$ from the simulated data. For each of these five statistics, we compute the so-called coefficient of determination (COD) defined by

$$R^2 = 1 - SS_{\mathrm{res}}/SS_{\mathrm{tot}}, \tag{B10}$$

where SS_{res} is the sum of squared deviations between model and experiment, $SS_{\mathrm{res}} = \sum_i (y_i - \hat{y}_i)^2$, with y_i being the experimental result and $\hat{y}_i$ being the model result for the five different statistics. The index i runs over the discrete values of the numerically computed statistics. $SS_{\mathrm{tot}} = \sum_i (y_i - \overline{y})^2$ is proportional to the variance of the experimental data as $\overline{y}$ is the mean of the experimental data. Consequently, R^2 is equal to 1 if the model perfectly captures the experimental data but equal to 0 if the model merely captures the mean of the experimental data. R^2 is negative when the model performs even worse than that. Varying all parameters Θ in the model gives the COD as a function of the model parameters for each candidate cell-cell interaction, for each cell type and for each statistics, $R^2_{j,k}(\Theta)$. Here j indicates the statistics of which we want to assess the goodness of fit. Thus, $R^2_{d,k}(\Theta)$ indicates the COD of the model fit to the underdamped interactions, while $R^2_{v,k}(\Theta)$ and $R^2_{x,k}(\Theta)$ are the COD of capturing the correlation functions of the velocities and positions, respectively. $R^2_{c,k}(\Theta)$ is the COD of the collision statistics. The index k indicates different cell types.

A simultaneous fit to all five statistics is challenging given that our models do not well predict the position correlation function $C_X(|t-t'|)$. Thus, the COD of the position correlation is strongly negative, rendering an average of the R^2 values over all statistics biased (Fig. 12). Thus, throughout, we average R^2 over the underdamped interactions, the velocity correlation function and the collision statistics, excluding the position correlation:

$$R^2_k(\Theta) = \tfrac{1}{2}R^2_{d,k}(\Theta) + \tfrac{1}{4}R^2_{v,k}(\Theta) + \tfrac{1}{4}R^2_{c,k}(\Theta). \tag{B11}$$

However, we later predict this position correlation using the model coming out of the fitting procedure (Fig. 12). Further, the weight we put on each statistics in this average does not greatly affect the final results described in the main text as shown in Fig. 16 in [54]. Note that the inferred dynamics and behavior statistics act effectively as independent degrees of freedom given our overdamped model Fig. 15 in [54], making it possible to constrain our overdamped model using the underdamped dynamics. For the individual fit, we maximize the average $R^2_k(\Theta)$ for each cell type individually. This gives us optimal parameters $\hat{\Theta}_k$ for each cell type with maximum COD $\hat{R}^2_k$. We show the maximum COD $\hat{R}^2_k$ for each cell type and each candidate interaction mechanism in Fig. 6(a). For the global fit, we only allow the interaction strength ε to vary between cell types, while fixing all other parameters. Thus, we maximize R^2 averaged over all statistics and cell types, giving us a single optimal value $\hat{r}$, but varying interaction strengths for each cell type $\hat{\varepsilon}$. We show the maximum global COD $\hat{\mathcal{R}}^2$ for each candidate interaction in Fig. 6(b). Throughout, we adjust the bounds of the parameters ε and r such that we at least include a local maximum of $R^2(\Theta)$ (Fig. 14 in [54]). For more details on the global fit, refer to Sec. 5.1 in [54].

8. Statistical analysis

Throughout, error bars indicate the standard error of the mean (s.e.m.) which we obtain for experimental data by bootstrapping. For the error bars of the statistics predicted by the inferred Langevin equation as shown by dotted bars in Figs. 4(a)–4(c), we again bootstrap the experimental data, perform the inference procedure and resimulate the inferred equation to predict these statistics. The error bars then again represent the error of the mean across these bootstrapped subsets of the data. For the phenomenological model, we indicate the variability of the predicted behavior statistics and inferred underdamped interaction terms by showing the spread of the 20 best-fitting parameter combinations. For model results, we typically simulate $N = 300$ trajectories. The number of experimental trajectories is given in Table 2 in [54].

APPENDIX C: SUPPLEMENTAL THEORY RESULTS

1. Mapping overdamped to underdamped interaction dynamics

A key challenge of our approach is to constrain the dynamics of nucleus, protrusion, and polarity using stochastic trajectories of the cell positions. On the one hand, we do this by comparing various long-time scale interaction behavior statistics computed from the trajectories of the cells. In addition, we consider an underdamped level of description of the cells, which takes into account the detailed short-time scale dynamics of the position x and velocity v of interacting cells. We do this for several reasons: (1) The dynamics of interacting cells on such a level of description has been studied before [8], and we aim to connect our overdamped model to these underdamped dynamics. (2) The underdamped dynamics provides additional constraints on the short-time scale dynamics of our overdamped model (Fig. 15 in [54]). (3) We can visualize the dynamics of our model in a low-dimensional way. In this section we theoretically show that our overdamped model can be mapped to such an underdamped level of description, providing further insight and justification of using the underdamped dynamics of the cell as a model constraint and as a visualization tool.

We consider our overdamped model in the following general form with friction coefficients set to 1 for

simplicity:

$$\dot{x}_{\mathrm{n},i} = F_{\mathrm{n}}(x_{\mathrm{n},i}, x_{\mathrm{p},i}) + G_{\mathrm{n}}(\Delta x, U, P), \tag{C1}$$

$$\dot{x}_{\mathrm{p},i} = F_{\mathrm{p}}(x_{\mathrm{n},i}, x_{\mathrm{p},i}) + P_i(t) + G_{\mathrm{p}}(\Delta x, U, P), \tag{C2}$$

$$\dot{P}_i = F_{\mathrm{P}}(x_{\mathrm{p},i}, P_i) + G_{\mathrm{P}}(\Delta x, U, P, \Delta P) + \sigma\eta_i(t). \tag{C3}$$

To map these dynamics to an underdamped level of description, we need to take a time derivative of Eq. (C1) and eliminate the protrusion coordinate of cells. This way, we introduce the velocity of the nucleus as a degree of freedom, effectively capturing the dynamics of the cell protrusion. We write $v_{\mathrm{n},i} = \dot{x}_{\mathrm{n},i}$ and find

$$\dot{v}_{\mathrm{n},i} = \frac{\partial F_{\mathrm{n}}}{\partial x_{\mathrm{n},i}} v_{\mathrm{n},i} + \frac{\partial F_{\mathrm{n}}}{\partial x_{\mathrm{p},i}} \dot{x}_{\mathrm{p},i} + \frac{d}{dt} G_{\mathrm{n}}(\Delta x, U, P), \tag{C4}$$

$$\dot{v}_{\mathrm{n},i} = \frac{\partial F_{\mathrm{n}}}{\partial x_{\mathrm{n},i}} v_{\mathrm{n},i} + \frac{\partial F_{\mathrm{n}}}{\partial x_{\mathrm{p},i}} [F_{\mathrm{p}}(x_{\mathrm{n},i}, x_{\mathrm{p},i}) + P_i(t) + G_{\mathrm{p}}(\Delta x, U, P)] + \frac{d}{dt} G_{\mathrm{n}}(\Delta x, U, P). \tag{C5}$$

Here $x_{\mathrm{p},i}$ has to be found through inverting Eq. (C1), which can give complicated functions $x_{\mathrm{p},i}(x_{\mathrm{n},i}, v_{\mathrm{n},i}, \Delta x_{\mathrm{n}}, \Delta v_{\mathrm{n}})$ as $x_{\mathrm{p},i}$ depends on $U = x_{\mathrm{p}} - x_{\mathrm{n}}$ and the intercellular distances Δx through Eq. (C1). In addition, we are still left with the "unobserved" polarity:

$$\dot{P}_i = F_{\mathrm{P}}(x_{\mathrm{p},i}, P_i) + G_{\mathrm{P}}(\Delta x, U, P, \Delta P) + \sigma\eta_i(t). \tag{C6}$$

The fact that the polarity remains unobserved makes an analytic mapping only possible if we could solve Eq. (C6) and analytically find the time evolution of the conditional probability distribution of P. In general, this is not possible and beyond the scope of this study. Nevertheless, we can define a mapping to the underdamped level of description through numerical methods. Consider the underdamped equation of motion of the form

$$\frac{dv_i}{dt} = F(x_i, v_i) + G(\Delta x, \Delta v) + \sigma\eta_i(t). \tag{C7}$$

Then we define the single-cell term as a conditional average $F(x_i, v_i) = \langle \dot{v}_{\mathrm{n},i} | x_{\mathrm{n},i}, v_{\mathrm{n},i} \rangle$ and get [33]

$$F(x_i, v_i) = \frac{\partial F_{\mathrm{n}}}{\partial x_{\mathrm{n},i}} v_{\mathrm{n},i}|_{\Delta x_{\mathrm{n}}=0, \Delta v_{\mathrm{n}}=0} + \frac{\partial F_{\mathrm{n}}}{\partial x_{\mathrm{p},i}} F_{\mathrm{p}}(x_{\mathrm{n},i}, x_{\mathrm{p},i})|_{\Delta x_{\mathrm{n}}=0, \Delta v_{\mathrm{n}}=0} + \frac{\partial F_{\mathrm{n}}}{\partial x_{\mathrm{p},i}}|_{\Delta x_{\mathrm{n}}=0, \Delta v_{\mathrm{n}}=0} \langle P_i(t) | x_{\mathrm{n},i}, v_{\mathrm{n},i} \rangle. \tag{C8}$$

Here $\langle P_i(t) | x_{\mathrm{n},i}, v_{\mathrm{n},i} \rangle$ is a conditional average and needs to be numerically estimated. Furthermore, note that $x_{\mathrm{p},i}$ needs to be eliminated but depends on the intercellular distances and relative velocities. Thus, when eliminating $x_{\mathrm{p},i}$ in Eq. (C8), terms that depend on Δx_{n} or Δv_{n} need to be ignored, which we indicate by $\ldots|_{\Delta x_{\mathrm{n}}=0, \Delta v_{\mathrm{n}}=0}$. For the interactions, we define $G(\Delta x, \Delta v) = \langle \dot{v}_{\mathrm{n},i} | \Delta x_{\mathrm{n}}, \Delta v_{\mathrm{n}} \rangle$. A deterministic form for $G(\Delta x, \Delta v)$ can be written down in a simple way when G_{n} and G_{p} do not depend on the unobserved polarity (which is true for 12/14 interaction mechanisms). We find the expression

$$G(\Delta x, \Delta v) = \frac{d}{dt} G_{\mathrm{n}}(\Delta x, U) + \frac{\partial F_{\mathrm{n}}}{\partial x_{\mathrm{p},i}} G_{\mathrm{p}}(\Delta x, U) + \langle P_i(t) | \Delta x_{\mathrm{n}}, \Delta v_{\mathrm{n}} \rangle + \frac{\partial F_{\mathrm{n}}}{\partial x_{\mathrm{p},i}} F_{\mathrm{p}}(x_{\mathrm{n},i}, x_{\mathrm{p},i})|_{\Delta x_{\mathrm{n}} \neq 0, \Delta v_{\mathrm{n}} \neq 0}, \tag{C9}$$

where again $x_{\mathrm{p},i}$ needs to be expressed in terms of $x_{\mathrm{n},i}$, $v_{\mathrm{n},i}$, Δx_{n}, and Δv_{n}. The term $\langle P_i(t) | \Delta x_{\mathrm{n}}, \Delta v_{\mathrm{n}} \rangle$ represents a conditional average and captures effectively the polarity interactions G_{P}. The last term arises due to $x_{\mathrm{p},i}$ being a function of the intercellular distance and relative velocities. Thus, terms that depend on Δx_{n} or Δv_{n} may arise when eliminating $x_{\mathrm{p},i}$ in F_{p} in Eq. (C8) and have to be considered in Eq. (C9) instead, which we indicate by $\ldots|_{\Delta x_{\mathrm{n}} \neq 0, \Delta v_{\mathrm{n}} \neq 0}$. Taken together, these calculations demonstrate that a mapping through numerically solving Eqs. (C8) and (C9) from the overdamped to an underdamped level of description is generally possible. This implies that the underdamped dynamics of the nucleus coordinate can provide additional strong constraints on the form of the overdamped equations. Implicitly, we perform such a mapping when numerically simulating our overdamped (1)–(3) and inferring the underdamped nucleus interactions (Fig. 5). Finally, note that even though such a mapping is possible, the two levels of description are not equivalent given that the overdamped model has an additional degree of freedom. To fully capture these overdamped dynamics, a third degree of freedom in the underdamped description such as a derivative of the acceleration or a memory term would be needed. This additional complexity is beyond the scope of our approach given that such an equation cannot be currently inferred from experimental data. In the following, we give two examples where even analytic mapping is possible, illustrating how overdamped interaction terms can be mapped to underdamped interaction terms.

a. Example: Linear spring and nucleus repulsion

To illustrate the presented mapping, we consider a simple example model where the nucleus and protrusion are coupled by a linear spring and cell nuclei are coupled through repulsion or attraction interactions with strength ε (Sec. B 3).

$$\dot{x}_{\mathrm{n},i} = -k(x_{\mathrm{n},i} - x_{\mathrm{p},i}) + \varepsilon e^{-|\Delta x_{\mathrm{n}}|/r} \Delta x_{\mathrm{n}}, \tag{C10}$$

$$\dot{x}_{\mathrm{p},i} = -k(x_{\mathrm{p},i} - x_{\mathrm{n},i}) + P_i(t), \tag{C11}$$

$$\dot{P}_i = -\frac{1}{\tau} P_i + \sigma\eta_i(t). \tag{C12}$$

For illustrative purposes, we enable an analytic mapping and consider that polarity exhibits an Ornstein-Uhlenbeck process with $\langle P\rangle = 0$ and $\langle P(t)P(t')\rangle = \delta(t-t')$. Inverting Eq. (C10) yields

$$x_{\mathrm{p},i} = \frac{1}{k}(v_{\mathrm{n},i} + kx_{\mathrm{n},i} - \varepsilon e^{-|\Delta x_{\mathrm{n}}|/r}\Delta x_{\mathrm{n}}). \tag{C13}$$

Computing the single-cell term, we find

$$F(x_i, v_i) = -2kv_{\mathrm{n},i}. \tag{C14}$$

Here we ignored a term $k\varepsilon e^{-|\Delta x_{\mathrm{n}}|/r}\Delta x_{\mathrm{n}}$, illustrating that interaction terms may arise from Eq. (C8) as discussed above. Computing interaction terms through Eq. (C9), we find

$$G(\Delta x, \Delta v) = \frac{d}{dt}\varepsilon e^{-|\Delta x_{\mathrm{n}}|/r}\Delta x_{\mathrm{n}} + k\varepsilon e^{-|\Delta x_{\mathrm{n}}|/r}\Delta x_{\mathrm{n}}, \tag{C15}$$

where the latter term comes from the fourth term in Eq. (C9). If we write $G(\Delta x, \Delta v) = f(\Delta x)\Delta x + \gamma(\Delta x)\Delta v$, this mapping gives

$$f(\Delta x) = k\varepsilon e^{-|\Delta x_{\mathrm{n}}|/r}, \tag{C16}$$

$$\gamma(\Delta x) = \varepsilon\left(1 - \frac{|\Delta x_{\mathrm{n}}|}{r}\right)e^{-|\Delta x_{\mathrm{n}}|/r}, \tag{C17}$$

where we evaluated the total time derivative of the nucleus coupling term. Thus, on an underdamped level of description, *overdamped* nucleus coupling interactions give rise to both a rescaled *underdamped* repulsion term and an *underdamped* friction term. The sign of the latter depends on the intercellular distance compared to the interaction range of the overdamped nucleus repulsion interaction.

b. Example: Polarity-polarity coupling interactions

A key interaction mechanism in our overdamped model is polarity-polarity coupling (PPC). This interaction captures the dynamics of all cell types (Fig. 6). Thus, in a second example, we show how overdamped PPC gives rise to underdamped interaction dynamics of the cell nucleus. Because polarity is an unobserved degree of freedom in the mapping outlined above, we consider our model in a limit where nucleus and protrusion positions are equal and polarity directly drives the nucleus position of the cell. In this limit an analytic mapping becomes possible. In this model, we consider both nucleus coupling and polarity-polarity coupling. Thus, we write two coupled stochastic equations of motion:

$$\dot{x}_{\mathrm{n},i} = P_i(t) + g(|\Delta x_{\mathrm{n}}|)\Delta x_{\mathrm{n}}, \tag{C18}$$

$$\dot{P}_i = -\frac{1}{\tau}P_i + h(|\Delta x_{\mathrm{n}}|)\Delta P + \sigma\eta_i(t). \tag{C19}$$

Here the functions $g(|\Delta x_{\mathrm{n}}|)$ and $h(|\Delta x_{\mathrm{n}}|)$ encode the spatial structure of the overdamped interactions and could be simple exponential decays as in the previous example. Again, for simplicity, we consider that the polarity exhibits an Ornstein-Uhlenbeck process with $\langle P\rangle = 0$ and $\langle P(t)P(t')\rangle = \delta(t-t')$ in the absence of interactions. First, inverting Eq. (C18) gives

$$P_i = v_{\mathrm{n},i} - g(|\Delta x_{\mathrm{n}}|)\Delta x_{\mathrm{n}}, \tag{C20}$$

from which we find

$$\Delta P = P_i - P_j = \Delta v_{\mathrm{n}} - 2g(|\Delta x_{\mathrm{n}}|)\Delta x_{\mathrm{n}}. \tag{C21}$$

TABLE I. Optimal parameters $\hat{\Theta}$ for the various different cell types obtained from a global fit (first four rows) and individual fit (last four rows). For the global fit, all parameters except ϵ_{PPC} are constant across cell types (Appendix B 7).

	MDA-MB-231	MDA-MB-436	HT1080	A549	MCF10A
$\varepsilon_{\mathrm{PPC}}$	$-5.9\ \mathrm{h}^{-1}$	$-3.1\ \mathrm{h}^{-1}$	$1\ \mathrm{h}^{-1}$	$5.5\ \mathrm{h}^{-1}$	$5.5\ \mathrm{h}^{-1}$
r_{PPC}	5 μm	5 μm	5 μm	5 μm	5 μm
$\varepsilon_{\mathrm{PPrC}}$	$3.3\ \mathrm{h}^{-2}$	$3.3\ \mathrm{h}^{-2}$	$3.3\ \mathrm{h}^{-2}$	$3.3\ \mathrm{h}^{-2}$	$3.3\mathrm{h}^{-2}$
r_{PPrC}	31.4 μm	31.4 μm	31.4 μm	31.4 μm	31.4 μm
$\varepsilon_{\mathrm{PPC}}$	$-6.8\ \mathrm{h}^{-1}$	$-3.9\ \mathrm{h}^{-1}$	$1\ \mathrm{h}^{-1}$	$1.4\ \mathrm{h}^{-1}$	$2.3\ \mathrm{h}^{-1}$
r_{PPC}	5 μm	5 μm	5 μm	20 μm	12.8 μm
$\varepsilon_{\mathrm{PPrC}}$	$5\ \mathrm{h}^{-2}$	$8.3\ \mathrm{h}^{-2}$	$3.3\ \mathrm{h}^{-2}$	$3.3\ \mathrm{h}^{-2}$	$3.3\mathrm{h}^{-2}$
r_{PPrC}	27.1 μm	22.8 μm	31.4 μm	27.1 μm	31.4 μm

The factor 2 comes from a necessary flip of the sign of Δx_{n} if we consider two different cells i and j. Then taking a time derivative of Eq. (C18), we find

$$\dot{v}_{\mathrm{n},i} = \dot{P}_i(t) + \frac{d}{dt}[g(|\Delta x_{\mathrm{n}}|)\Delta x_{\mathrm{n}}] \tag{C22}$$

$$= -\frac{1}{\tau}P_i + h(|\Delta x_{\mathrm{n}}|)\Delta P + \frac{d}{dt}[g(|\Delta x_{\mathrm{n}}|)\Delta x_{\mathrm{n}}] + \sigma\eta_i(t). \tag{C23}$$

Eliminating P from the equation, we can write

$$\dot{v}_{\mathrm{n},i} = -\frac{1}{\tau}v_i + f(|\Delta x|)\Delta x + \gamma(|\Delta x|)\Delta v + \sigma\eta_i(t) \tag{C24}$$

with $f(|\Delta x|)$ and $\gamma(|\Delta x|)$ both depending on combinations of $g(|\Delta x_{\mathrm{n}}|)$ and $h(|\Delta x_{\mathrm{n}}|)$. Both previous examples illustrate how overdamped interactions such as NC or PPC can be mapped to both underdamped cohesion and friction interactions. However, to enable the analytic mapping in these examples, we simplified our system by eliminating one of the three degrees of freedom. In general, the mapping is more complicated and requires numerically solving Eqs. (C8) and (C9).

2. Overdamped inference results

In this section, we show the full results of the overdamped inference procedure of which we have shown a summary in Fig. 6. Specifically, for the best-fitting candidate interaction mechanisms PPC + PPrC, we show a full comparison of the fitted model to the experimental data using the global fit (Fig. 9) and the individual fit (Fig. 10). Both these fits provide a good description of the underdamped interactions, collision statistics, and position and velocity correlation functions of the experimental data, also at longer timescales [Figs. 9(b) and 9(c); Figs. 10(b) and 10(c)]. Note that the individual fit yields a slightly better model for MDA-MB-231 cells, better capturing the underdamped interaction terms. Table I shows the parameters that our global and individual fitting procedures yield for the combination of PPC + PPrC. These parameters exhibit great similarities between global and individual fits, confirming the validity of the global fit.

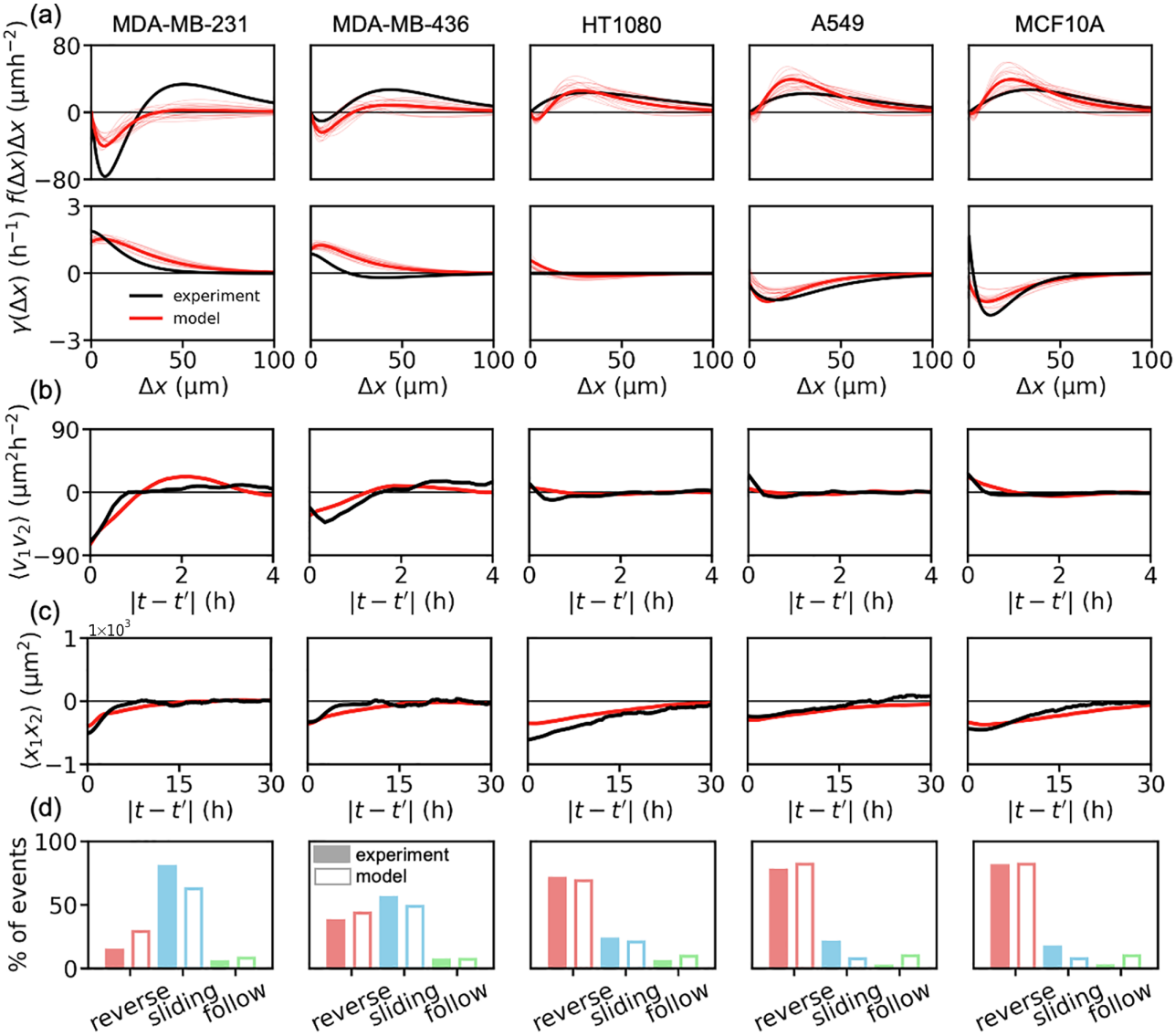

FIG. 9. Results of the global fit of PPC + PPrC to the experimental data. (a) Underdamped cell-cell interactions for both experiment (black) and best-fitting model (red). Upper row shows the underdamped cohesion interactions, while lower row shows the inferred friction interactions. (b)–(d) All behavior statistics for both experiment (black, solid bars) and model (red, empty bars). (b) The velocity correlation function $\langle v_1(t)v_2(t')\rangle_{\text{same}}$ when cells are on the same island. (c) The position correlation function $\langle x_1(t)x_2(t')\rangle$. (d) The collision statistics for all different cell types.

3. The role of single-cell behavior

In the absence of cell-cell interactions, our model is consistent with the mechanistic model for single-cell migration of MDA-MB-231 cells used in [33]. Given that we here study multiple different cell lines, it is a key question whether the single-cell dynamics of our model has a significant impact on the results about cell-cell interactions presented in the main text. To answer this question, we first determine the range of single-cell parameters that is relevant to describe single-cell behavior of a cell type other than MDA-MB-231 cells. MCF10A cells are a prime candidate as they show a lower probability to be on the bridge and a significantly higher survival probability during their stochastic hopping process [Fig. 3(d) in [54]]. Additionally, we find that MCF10A cells have the greatest differences in their behavior statistics compared to MDA-MB-231 cells [Figs. 4(a)–4(c)]. To capture the different single-cell features of MCF10A cells, we vary key single-cell parameters in our overdamped model. Specifically, the reduced spring constant $k_p = k/\zeta_p$ [Eqs. (2) and (B2)] determines how the protrusion of a cell responds to the nucleus position. We find that k_p influences the probability distribution of the nucleus positions and sets the survival probability of single cells [Figs. 11(a) and 11(b)]. Decreasing k_p from 1.2 h^{-1} for MDA-MB-231 cells to a value of around 0.1 h^{-1} allows us to capture the reduced probability of cells being on the bridge and enables us to capture the higher survival probabilities of MCF10A cells. Importantly, this different single-cell behavior does not greatly influence the inferred underdamped cell-cell interactions and behavior statistics predicted by our overdamped model [Figs. 11(c) and 11(g)]. Nevertheless, we repeat our quantitative fitting procedure with the different single-cell parameters for MCF10A

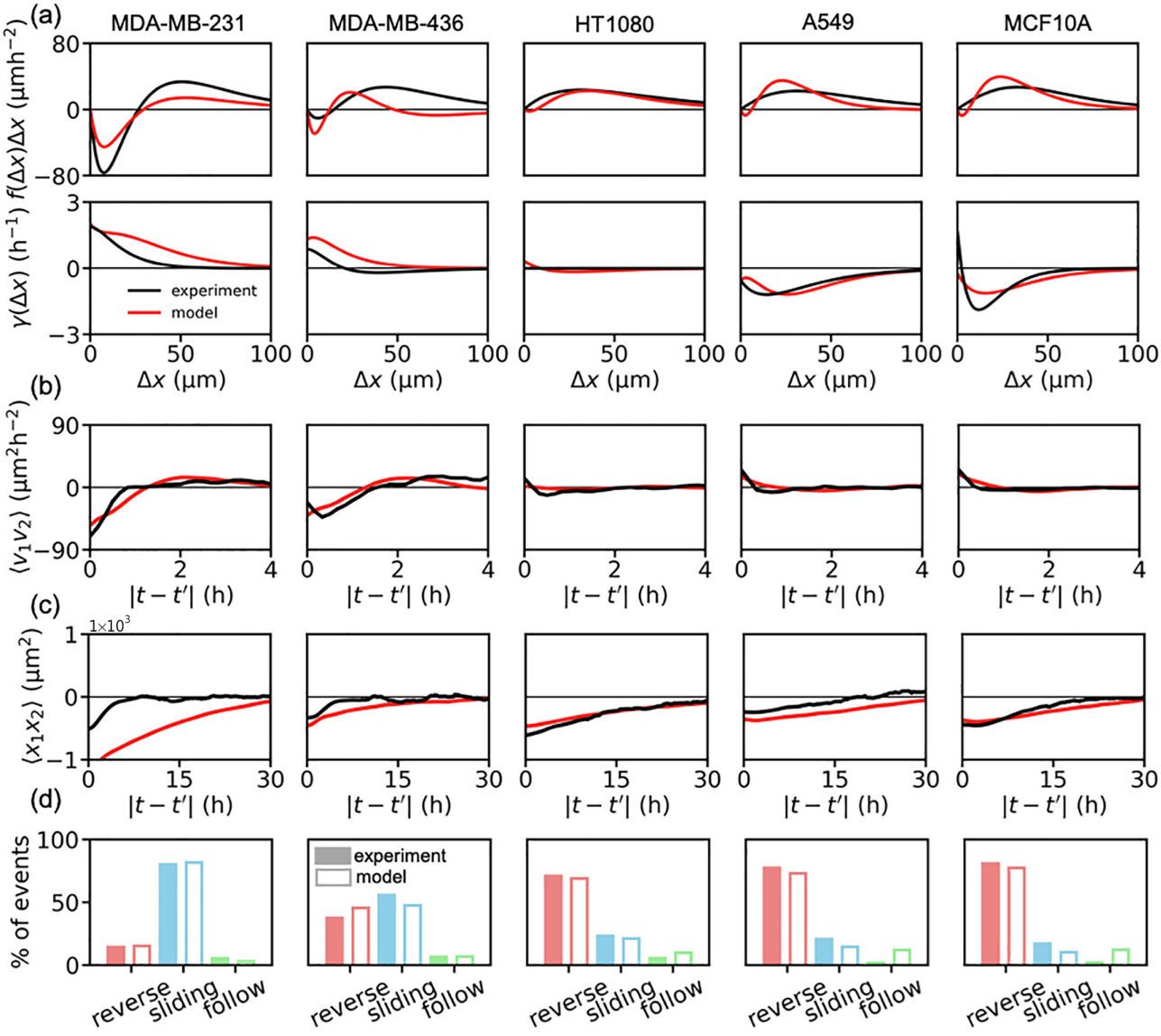

FIG. 10. Results of individual fit of PPC + PPrC to the experimental data. (a) Underdamped cell-cell interactions for both experiment (black) and best-fitting model (red). Upper row shows the underdamped cohesion interactions, while lower row shows the inferred friction interactions. (b)–(d) All behavior statistics for both experiment (black, solid bars) and model (red, empty bars). (b) The velocity correlation function $\langle v_1(t)v_2(t')\rangle_{\text{same}}$ when cells are on the same island. (c) The position correlation function $\langle x_1(t)x_2(t')\rangle$. (d) The collision statistics for all different cell types.

cells instead of the single-cell parameters of the MDA-MB-231 cells. We find that the best-fitting interaction parameters in our overdamped model with PPC + PPrC are only mildly affected by the different single-cell behavior (Tables I and II).

TABLE II. Optimal parameters $\hat{\Theta}$ for the various different cell types obtained from a global fit with single-cell parameters of MCF10A.

	MDA-MB-231	MDA-MB-436	HT1080	A549	MCF10A
ε_{PPC}	-5.9 h^{-1}	-3.1 h^{-1}	1 h^{-1}	5.5 h^{-1}	5.5 h^{-1}
r_{PPC}	8.8 μm	8.8 μm	8.8 μm	8.8 μm	8.8 μm
$\varepsilon_{\text{PPrC}}$	8.3 h^{-2}	8.3 h^{-2}	8.3 h^{-2}	8.3 h^{-2}	8.3h^{-2}
r_{PPrC}	22.9 μm	22.9 μm	22.9 μm	22.9 μm	22.9 μm

Thus, for simplicity, we keep the single-cell behavior constant throughout this study, knowing that the single-cell behavior does not qualitatively influence the dynamics emerging from cell-cell interactions.

4. Prediction of position correlations

In this section we describe to what extend our candidate cell-cell interactions are able to correctly predict the position correlation function $C_X(|t - t'|)$ which has not been used in the fitting procedure to constrain our candidate interactions. To this end, we investigate $R^2_{x,k}(\Theta)$, which quantifies how well our model captures the position correlation function $C_X(|t - t'|)$ for a specific cell type, a specific candidate cell-cell interaction, and a specific parameter combination Θ. Thus, of particular interest is the quantity $R^2_{x,k}(\Theta = \hat{\Theta})$, where $\hat{\Theta}$ is the best-fitting parameter combina-

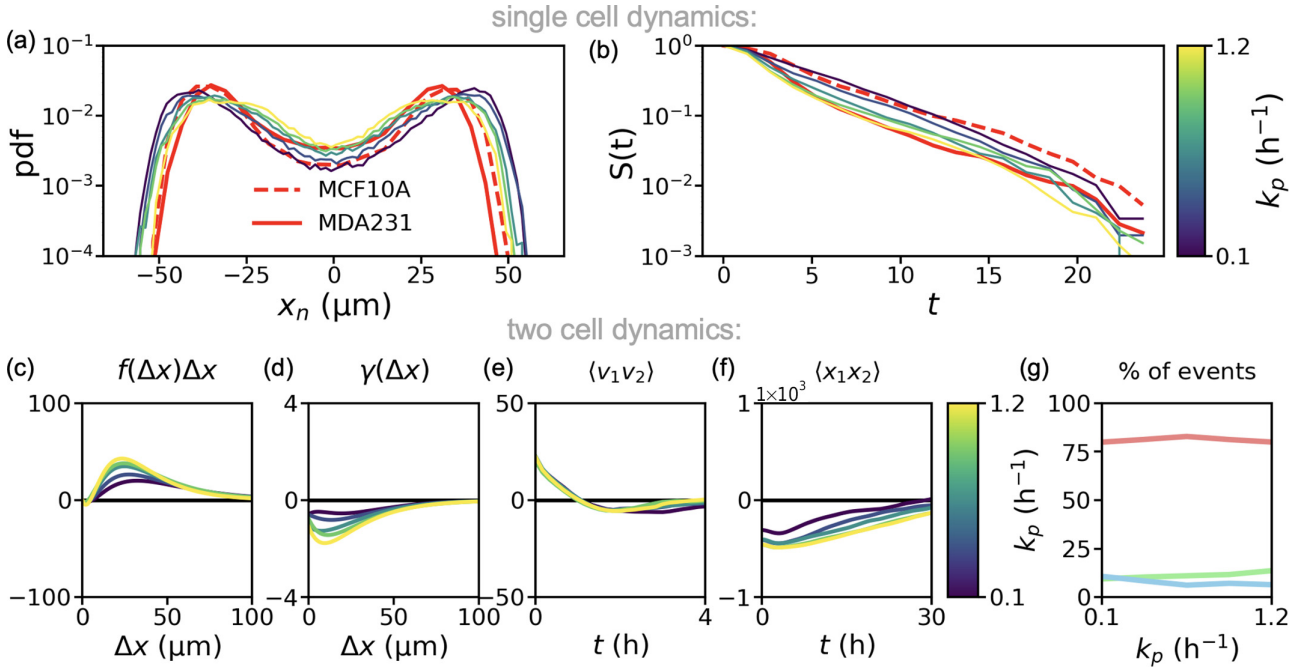

FIG. 11. Influence of single-cell dynamics on two cell dynamics. (a) Distribution of nucleus positions for both model and experiment (red) for model and experiments of a single cell on the dumbbell-shaped micropattern. Red solid line shows the experimental result of MDA-MB-231 cells, and red dashed line indicates MCF10A cells. Colored lines indicate model results and color bar for varying k_p is located next to panel (b). (b) Survival probability of both model (colored) and experiment (red) for single cells on the pattern. Legend of panel (a) applies also here. (c)-(g) Model predictions for the two-cell dynamics dependent on the single-cell parameter k_p. (c) The underdamped repulsion interactions; (d) the underdamped friction interactions; (e) the velocity cross-correlation of two cells on the same island; (f) the position correlations of the two cells. Colorbar in panel (f) applies to panels (c)–(f). Panel (g) shows the percentage of each type of collision event dependent on k_p.

tion of the individual fit as described in Appendix B 7. We find that the best-fitting models for each candidate interaction in general do not well capture the position correlation function [Figs. 12(a) and 12(b)]. In fact, $R^2_{x,k}(\Theta = \hat{\Theta})$ is consistently negative showing that our models struggle to capture these correlations. However, note that combining candidate interactions in general does improve the prediction of the position correlation functions. Specifically, the best working candidate interactions coming out of our inference procedure PPC + PPrC do provide the best prediction of the position correlation function at its optimal individual and global interaction parameters [Figs. 12(b) and 12(c)], outperforming PUC + PPrC [Figs. 12(d) and 12(e)]. These results show that we can use the predicted position correlation function to further constrain our candidate interaction mechanisms.

5. Polarity-polarity coupling on rectangular patterns

To test whether insights of our overdamped model sensitively depend on the chosen dumbbell-shaped geometry of our micropattern, we investigate cell-cell interactions on rectangular micropatterns [Fig. 13(e)]. Specifically, we consider the epithelial cell line MCF10A and the breast cancer cell line MDA-MB-231 and track cell nuclei [Fig. 13(f)]. We first infer the underdamped cell-cell interactions for these two cell types on the rectangular pattern and compare our results to the inferred underdamped cell-cell interactions on the dumbbell-shaped micropattern. For MCF10A cells, we find that these cells obey qualitatively similar underdamped cell-cell interactions on both geometries [black and gray curves in Figs. 13(h) and 13(i)]. However, the underdamped cohesion interactions are weaker on the rectangular pattern. Furthermore, in agreement with previous work [8], MDA-MB-231 cells exhibit quantitatively similar underdamped cell-cell interactions [black and gray curves in Figs. 13(k) and 13(l)]. These results show that the geometry of the micropattern has only a minor impact on the inferred underdamped cell-cell interactions.

To further test whether our overdamped model can also capture the two-cell dynamics on a rectangular pattern, we perform numerical simulations of our overdamped model on this alternative geometry. Specifically, we consider the best-fitting candidate cell-cell interaction PPC + PPrC and use for both MCF10A and MDA-MB-231 cells the best interaction parameters obtained from our global fit on the dumbbell pattern. Importantly, in order to capture the dynamics on a different geometry, we have to adjust the single-cell aspects of our overdamped model: We remove the geometry dependence of the single-cell terms in our overdamped model, which implemented the adaptation of the cell's migratory machinery to the confinement of the bridge in our dumbbell-shaped micropattern [Eq. (B2)]. Specifically, we set $\gamma(x_n) = 1$ and $\alpha(x_p) = \alpha_0$. We analyze single cells migrating on rectangular pattern to constrain single-cell parameters in our model [Figs. 13(a)–13(d)]. We find that $\alpha_0 = -0.3$ is negative such that cells remain polarized on the rectangular pattern. To better capture experimental trajectories of single cells on rectangular

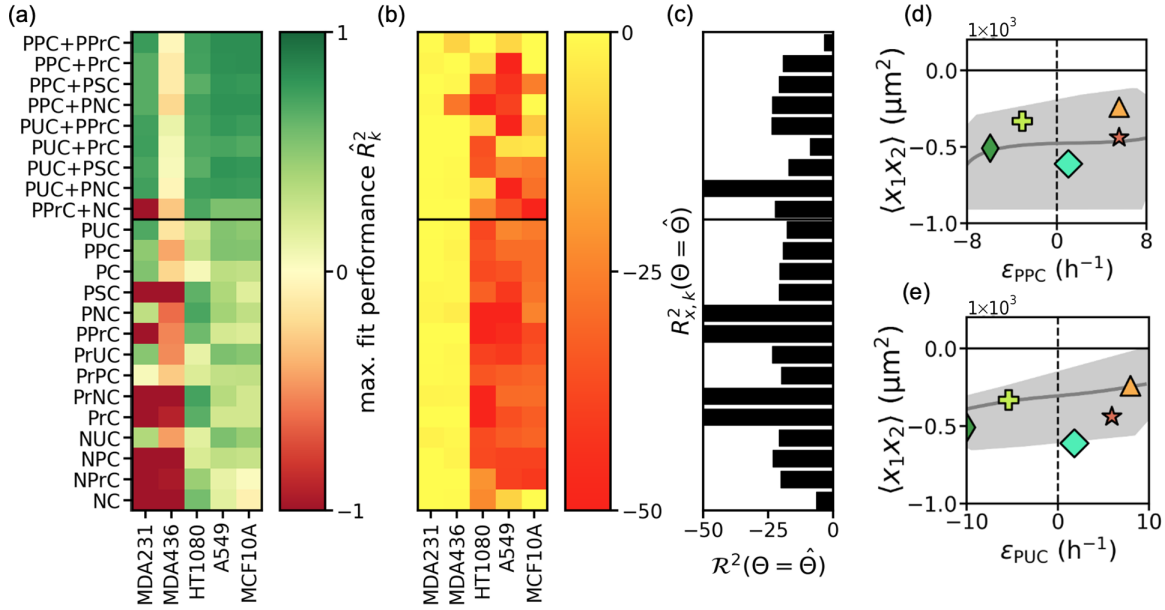

FIG. 12. Prediction of position correlation function (a) Maximum fit performance as quantified by the maximum coefficient of determination $\hat{R}^2$ for each interaction mechanism using an individual fit. (b) Coefficient of determination of the position correlation function evaluated at the best-fitting parameters $\hat{\Theta}$ obtained from our individual fit $R^2_{x,k}(\Theta=\hat{\Theta})$. (c) Coefficient of determination of the position correlation function evaluated at the best-fitting parameters $\hat{\Theta}$ obtained from our global fit. (d), (e) Instantaneous position correlations as predicted by the two models PPC + PPrC (d) and PUC + PPrC (e). Solid lines show model result, symbols show experimental results as also presented in Fig. 6(f).

patterns, we also change the values of σ and k_n. Specifically, for MCF10A cells, we choose $\sigma = 40$ and $k_n = 1.2$. These two changes are necessary to capture the high persistence and frequent turnaround events of our MCF10A cells on the rectangular pattern [Fig. 13(a)] as also quantified by a velocity autocorrelation function $\Phi(t) = \langle v(t')v(t+t')\rangle_{t'}$ [Fig. 13(b)]. Physically, increasing k_n means that we couple the nucleus closer to the protrusion. This change may be in agreement with the observation of reduced elongation of the cells on rectangular patterns compared to the dumbbell-shaped patterns [Fig. 13(e)]. On the rectangular patterns, we observe that when cells hit one of the ends of the rectangular pattern, they polarize away from the end and turn around. We implement this by including a boundary condition where the polarity P of a cell gets flipped to $-P$ as soon as cells hit the boundary of our pattern, preserving the state of polarity [56]. For MDA-MB-231 cells, we increase σ to $\sigma = 150$ to capture the higher stochasticity of the single-cell behavior of these cells on the rectangular pattern [Figs. 13(c) and 13(d)].

Having constrained the single-cell parameters, we simulate our overdamped model on a rectangular pattern. We find that using the interaction parameters best describing the dynamics on the dumbbell-shaped geometry also yields a good description of the dynamics on the rectangular geometry. Specifically, we can qualitatively capture the flocking and frequent reversal behavior of MCF10A cells on the rectangular pattern [Figs. 13(f) and 13(g)]. Furthermore, our overdamped model qualitatively captures the inferred interactions [Figs. 13(h) and 13(i)] and the positive velocity correlations of MCF10A cells on the rectangular pattern [Fig. 13(j)]. Moreover, our model captures the frequent sliding events of MDA-MB-231 cells [Figs. 13(f) and 13(g)] as well as qualitatively the inferred interactions and cross-correlations of these cells [Figs. 13(k) and 13(m)]. Taken together, these results show that our overdamped model can be generalized to a different, less constrained geometry. Thus, the identified polarity-polarity coupling interactions do not depend sensitively on the specific dumbbell-shaped geometry used in the main text.

6. Predicting collective front migration of epithelial and breast cancer cells

To show the generality of our overdamped model and the inferred dominant polarity-polarity coupling mechanism, we predict collective cell migration in a physiologically relevant migration scenario: collective front migration during wound healing. Specifically, similar to our study, a previous study [60] considered epithelial MCF10A cells and cancerous MDA-MB-231 cells, of which large monolayers were initially confined to a circular region and subsequently released to invade into the surrounding. Their experiments and detailed quantification of the collective migration behavior showed that MCF10A cells exhibit ordered and steady collective front migration. In contrast, MDA-MB-231 cells exhibit no significant front migration with highly disordered motion of individual cells [60]. In this section, we generalize our model and predict the collective front migration of MCF10A and MDA-MB-231 cells with our overdamped model and best-fitting interaction mechanisms PPC + PPrC learned from two-cell collisions.

To enable a prediction of collective cell migration behavior in an unconfined 2D environment, we first generalize the single-cell dynamics of our overdamped model. Thus, we remove the geometry dependence of the friction coefficient and polarity feedback [Figs. 8(a) and 8(b); Appendix B 1]. We perform numerical simulations of our model in the absence of

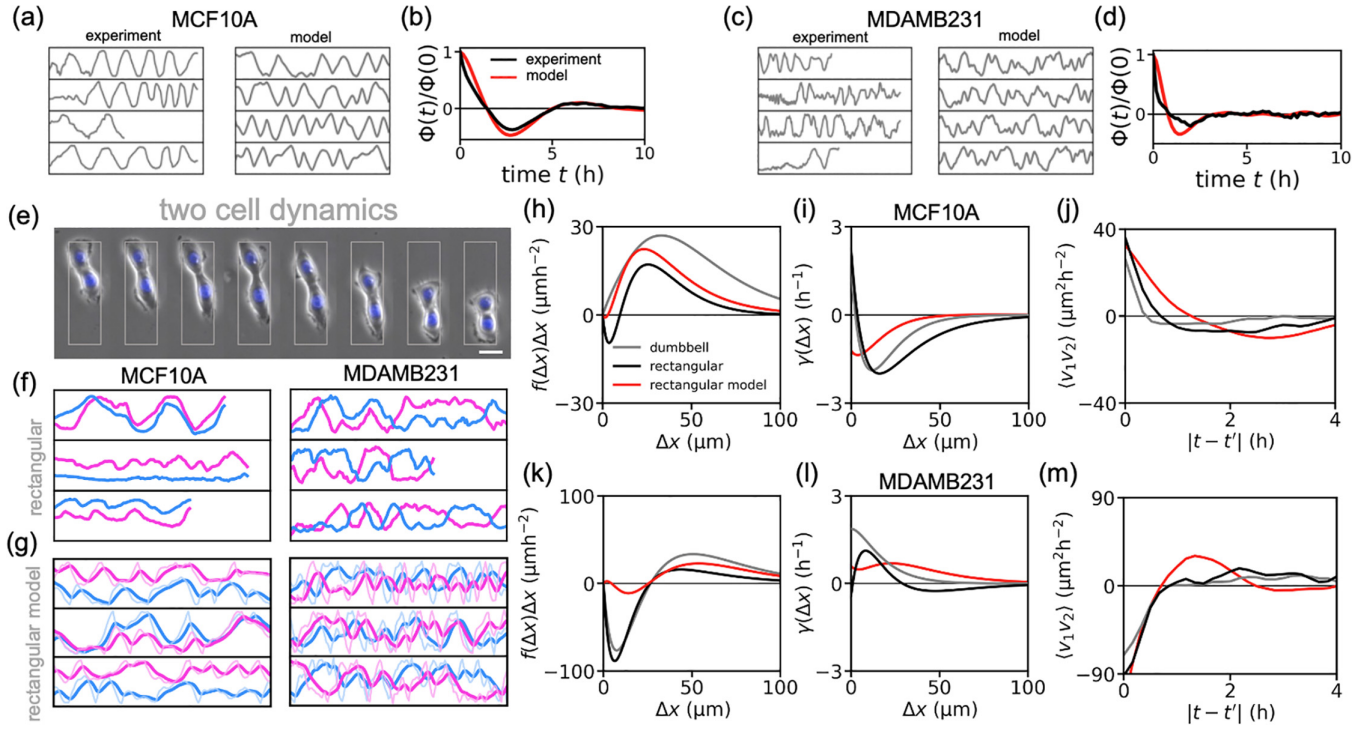

FIG. 13. Dynamics of PPC + PPrC on rectangular patterns. (a) Sample of cell trajectories of a single MCF10A cell on a rectangular pattern for both experiment and model. (b) Velocity autocorrelation function of MCF10A cells on rectangular pattern for both experiment (black) and model (red). (c) Sample of cell trajectories of a single MDA-MB-231 cell on a rectangular pattern for both experiment and model. (d) Velocity autocorrelation function of MDA-MB-231 cells on rectangular pattern for both experiment (black) and model (red). (e) Time series of brightfield images of two MCF10A cells interacting on a rectangular micropattern. Scale bar: 25 μm. (f) Sample of cell trajectories obtained from our experiments on a rectangular pattern for both MCF10A and MDA-MB-231 cells. (g) Sample of trajectories predicted by our overdamped model with PPC + PPrC as cell-cell interaction. (h)–(j) Inferred underdamped cohesion interactions (h), inferred underdamped friction interactions (i), and velocity cross-correlation function of two MCF10A cells for the dumbbell pattern (gray), the rectangular pattern (j), and the overdamped model fitted to the rectangular dynamics (red). (k)–(m) Inferred underdamped cohesion interactions (k), inferred underdamped friction interactions (l), and velocity cross-correlation function of two MDA-MB-231 cells for the dumbbell pattern (gray), the rectangular pattern (m), and the overdamped model fitted to the rectangular dynamics (red).

any interactions between cells to find experimentally realistic single-cell parameters. This requires analyzing experimental trajectory data of single cells migrating on an unconfined 2D substrate. For MDA-MB-231 cells, these data have been published in [55], for MCF10A cells we perform experiments with single MCF10A cells on a 2D substrate [Sec. 2.5 in [54]]. For both cell types, we compute a velocity autocorrelation function defined as $\Phi(t) = \langle v(t')v(t+t')\rangle_{t'}$, which quantifies the persistence of cell motion. In addition, we compute the distribution of the speeds of the cells. These are common quantities to characterize single-cell migration behavior [62]. By sweeping over single-cell parameters, we identify experimentally relevant parameter regimes in which we capture approximately the exponential decay of the velocity autocorrelation function and the mean speed of both single MCF10A and MDA-MB-231 cells [Figs. 14(a) and 14(b)].

To predict collective front migration behavior, we consider a rectangular confinement with one open face and extending walls [Fig. 14(c)]. This geometry approximates a small section of the large circular geometry considered in [60], while allowing us to simulate less cells in total. The boundary is implemented by reflective boundary conditions on both the cell positions and cell polarities. We consider a rectangular geometry with a large width of 1000 μm to avoid effects of the boundary condition. In our model for collective front migration, we implement the 2D generalization of the interaction mechanisms PPC and PPrC. The parameters of these interaction mechanisms are set to the best-fitting parameters of these mechanisms inferred from the two-cell collision dynamics using our global fit (Fig. 6 and Table I). To create realistic cell monolayers, we implement cell division. After initializing 500 cells and a short prerun, we record the collective cell migration behavior of both MCF10A and MDA-MB-231 cells and quantify their collective front migration [Figs. 14(d) and 14(e)].

For MDA-MB-231 cells, we implement polarity antialignment ($\varepsilon_{\mathrm{PPC}} = -6\ \mathrm{h}^{-1}$, $r_{\mathrm{PPC}} = 5$ μm) and polarity-protrusion repulsion ($\varepsilon_{\mathrm{PPrC}} = 3.3\ \mathrm{h}^{-2}$, $r_{\mathrm{PPrC}} = 31$ μm) as determined by our overdamped inference procedure presented in the main text (Fig. 6). For these cells, we find inefficient front migration, with the front of the monolayer sometimes even retracting [inset of Fig. 14(d)]. Furthermore, cell velocities are highly disordered [Fig. 14(d)] with no dominant instantaneous velocity direction into the empty region [green curve Fig. 14(f)]. This behavior is qualitatively similar to the absence of front migration of MDA-MB-231 cells in [60]. In contrast, for MCF10A cells, our two-cell inference procedure uncovers polarity alignment interactions ($\varepsilon_{\mathrm{PPC}} = 6\ \mathrm{h}^{-1}$,

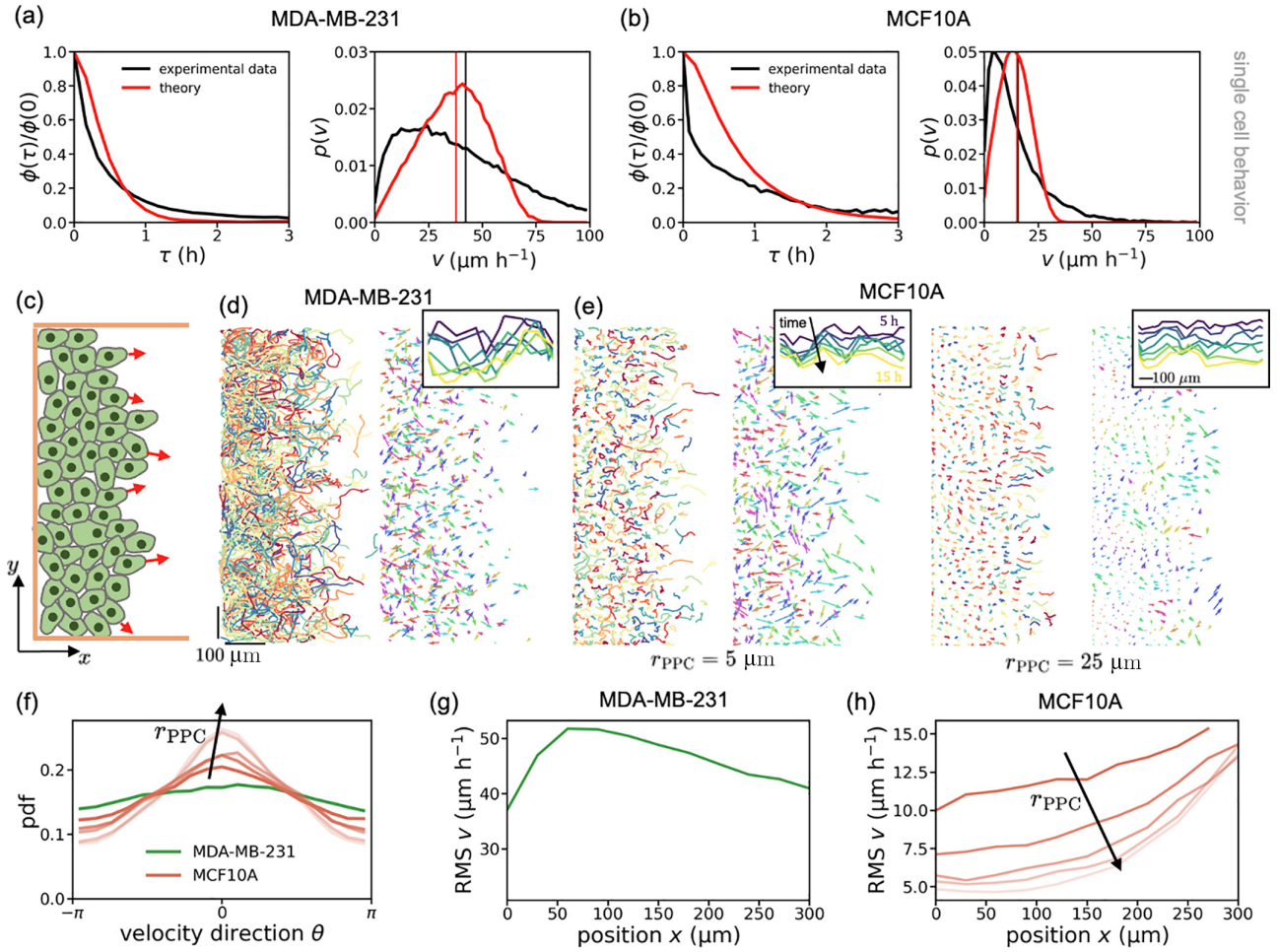

FIG. 14. Predicting collective front migration of epithelial and breast cancer cells. (a), (b) Single-cell behavior of freely migrating single MDA-MB-231 cells (a) and MCF10A cells (b). Left panel shows the velocity autocorrelation function $\Phi(t)/\Phi(0)$ for experimental data (black) and prediction from our theory (red). Right panel shows the distribution of speeds in the experiment (black) and in our theoretical prediction (red) over time and different single cells. Vertical lines indicate the mean speed. (c) Wound healing setup that we use to test the predictive power of our overdamped model learned in two-cell collision experiments. Red arrows indicate cell velocity of a few chosen cells. (d), (e) Predicted collective dynamics of MDA-MB-231 cells (d) and MCF10A cells (e). We show the predicted cell trajectories (left) and the instantaneous velocity field at a time point late in the simulation (right). Inset shows the collective front migration behavior. Scale bar and timescales indicated in the inset of panel (e) apply to all simulations in panels (d) and (e). (f) Distribution of velocity directions for the predicted collective dynamics. Here $\Theta = 0$ indicates motion along the x axis in positive direction. Green lines indicate the prediction for MDA-MB-231 cells, and red lines indicate the prediction for MCF10A cells. Opacity of the red lines indicate the value of the interaction range r_{PPC} with increasing values resulting in more faded lines. (g), (h) Root-mean-squared (RMS) velocity as a function of the x position along the direction of the spreading monolayer for MDA-MB-231 cells (g) and MCF10A cells (h). Again opacity of lines in panel (h) indicate the value of the interaction range r_{PPC}.

$r_{PPC} = 5$ μm) and polarity-protrusion repulsion ($\varepsilon_{PPrC} = 3.3$ h^{-2}, $r_{PPrC} = 31$ μm). Simulating their collective behavior, we find rather steady front migration with cell velocities predominantly directed towards the empty region [Figs. 14(e) and 14(f)]. Thus, qualitatively, we capture the peaked distribution of velocity directions observed in [60] [Fig. 14(f)]. Increasing the interaction range r_{PPC} of PPC from 5 μm to 25 μm, which are both reasonable interaction ranges, yields more efficient front migration [insets of Fig. 14(e)] and more order in the velocity field [Figs. 14(e) and 14(f)]. This higher degree of order is better in agreement with the experimentally observed front migration behavior [60]. Finally, we predict the spatial structure of the velocity field expected for these cell types. For MDA-MB-231 cells, we observe high RMS (root-mean-squared) velocities in the bulk of the monolayer [Fig. 14(g)]. This high motility may be in agreement with the fluid phenotype observed in cancer cell monolayers [79]. For MCF10A cells, we observe significantly lower velocities inside the bulk compared to the migrating front [Fig. 14(h)]. This, spatial structure of the velocity field is in agreement with the velocity distribution of expanding monolayers of epithelial MDCK cells [39]. The slowdown in the bulk is stronger with increasing interaction range of PPC [Fig. 14(h)]. Taken together, the interaction mechanisms learned from confined two-cell collisions provides a good qualitative description of experimental collective front migration of epithelial and

breast cancer cells. These results indicate that the form of the interaction terms and to some extent also the inferred interaction parameters can be generalized to collective and less confined migration scenarios.

[1] A. Haeger, K. Wolf, M. M. Zegers, and P. Friedl, Collective cell migration: Guidance principles and hierarchies, Trends Cell Biol. **25**, 556 (2015).

[2] E. Scarpa and R. Mayor, Collective cell migration in development, J. Cell Biol. **212**, 143 (2016).

[3] M. Poujade, E. Grasland-Mongrain, A. Hertzog, J. Jouanneau, P. Chavrier, B. Ladoux, A. Buguin, and P. Silberzan, Collective migration of an epithelial monolayer in response to a model wound, Proc. Natl. Acad. Sci. USA **104**, 15988 (2007).

[4] P. Friedl and D. Gilmour, Collective cell migration in morphogenesis, regeneration and cancer, Nat. Rev. Mol. Cell Biol. **10**, 445 (2009).

[5] S. Vedel, S. Tay, D. M. Johnston, H. Bruus, and S. R. Quake, Migration of cells in a social context, Proc. Natl. Acad. Sci. USA **110**, 129 (2013).

[6] M. Abercrombie, Contact inhibition and malignancy, Nature (London) **281**, 259 (1979).

[7] E. Scarpa, A. Roycroft, E. Theveneau, E. Terriac, M. Piel, R. Mayor, E. Scarpa, A. Roycroft, E. Theveneau, E. Terriac, M. Piel, and R. Mayor, A novel method to study contact inhibition of locomotion using micropatterned substrates, Biol. Open **5**, 1553 (2016).

[8] D. B. Brückner, N. Arlt, A. Fink, P. Ronceray, J. O. Rädler, and C. P. Broedersz, Learning the dynamics of cell–cell interactions in confined cell migration, Proc. Natl. Acad. Sci. USA **118**, e2016602118 (2021).

[9] A. Roycroft and R. Mayor, Molecular basis of contact inhibition of locomotion, Cell. Mol. Life Sci. **73**, 1119 (2016).

[10] C. Collins and W. J. Nelson, Running with neighbors: Coordinating cell migration and cell-cell adhesion, Curr. Opin. Cell Biol. **36**, 62 (2015).

[11] B. Ladoux and R. M. Mège, Mechanobiology of collective cell behaviours, Nat. Rev. Mol. Cell Biol. **18**, 743 (2017).

[12] A. J. Ridley, Life at the leading edge, Cell **145**, 1012 (2011).

[13] W. J. Rappel and L. Edelstein-Keshet, Mechanisms of cell polarization, Curr. Opin. Syst. Biol. **3**, 43 (2017).

[14] G. Danuser, J. Allard, and A. Mogilner, Mathematical modeling of eukaryotic cell migration: Insights beyond experiments, Annu. Rev. Cell Dev. Biol. **29**, 501 (2013).

[15] S. Kadir, J. W. Astin, L. Tahtamouni, P. Martin, and C. D. Nobes, Microtubule remodelling is required for the front-rear polarity switch during contact inhibition of locomotion, J. Cell Sci. **124**, 2642 (2011).

[16] E. Theveneau and R. Mayor, Collective cell migration of epithelial and mesenchymal cells, Cell. Mol. Life Sci. **70**, 3481 (2013).

[17] S. Jain, V. M. Cachoux, G. H. Narayana, S. de Beco, J. D'Alessandro, V. Cellerin, T. Chen, M. L. Heuzé, P. Marcq, R. M. Mège, A. J. Kabla, C. T. Lim, and B. Ladoux, The role of single-cell mechanical behaviour and polarity in driving collective cell migration, Nat. Phys. **16**, 802 (2020).

[18] C. Carmona-Fontaine, H. K. Matthews, S. Kuriyama, M. Moreno, G. A. Dunn, M. Parsons, C. D. Stern, and R. Mayor, Contact inhibition of locomotion *in vivo* controls neural crest directional migration, Nature (London) **456**, 957 (2008).

[19] E. Theveneau, L. Marchant, S. Kuriyama, M. Gull, B. Moepps, M. Parsons, and R. Mayor, Collective Chemotaxis Requires Contact-Dependent Cell Polarity, Develop. Cell **19**, 39 (2010).

[20] J. R. Davis, C. Y. Huang, J. Zanet, S. Harrison, E. Rosten, S. Cox, D. Y. Soong, G. A. Dunn, and B. M. Stramer, Emergence of embryonic pattern through contact inhibition of locomotion, Development **139**, 4555 (2012).

[21] V. Villar-Cerviño, M. Molano-Mazón, T. Catchpole, M. Valdeolmillos, M. Henkemeyer, L. M. Martínez, V. Borrell, and O. Marín, Contact Repulsion Controls the Dispersion and Final Distribution of Cajal-Retzius cells, Neuron **77**, 457 (2013).

[22] R. Mayor and C. Carmona-Fontaine, Keeping in touch with contact inhibition of locomotion, Trends Cell Biol. **20**, 319 (2010).

[23] J. D'alessandro, A. P. Solon, Y. Hayakawa, C. Anjard, F. Detcheverry, J. P. Rieu, and C. Rivière, Contact enhancement of locomotion in spreading cell colonies, Nat. Phys. **13**, 999 (2017).

[24] M. Hayakawa, T. Hiraiwa, Y. Wada, H. Kuwayama, and T. Shibata, Polar pattern formation induced by contact following locomotion in a multicellular system, eLife **9**, e53609 (2020).

[25] A. M. Mendonsa, T. Y. Na, and B. M. Gumbiner, E-cadherin in contact inhibition and cancer, Oncogene **37**, 4769 (2018).

[26] J. L. Maître and C. P. Heisenberg, Three functions of cadherins in cell adhesion, Curr. Biol. **23**, R626 (2013).

[27] B. Stramer and R. Mayor, Mechanisms and in vivo functions of contact inhibition of locomotion, Nat. Rev. Mol. Cell Biol. **18**, 43 (2017).

[28] J. W. Astin, J. Batson, S. Kadir, J. Charlet, R. A. Persad, D. Gillatt, J. D. Oxley, and C. D. Nobes, Competition amongst Eph receptors regulates contact inhibition of locomotion and invasiveness in prostate cancer cells, Nat. Cell Biol. **12**, 1194 (2010).

[29] N. K. Noren and E. B. Pasquale, Eph receptor-ephrin bidirectional signals that target Ras and Rho proteins, Cell. Signal. **16**, 655 (2004).

[30] R. A. Desai, L. Gao, S. Raghavan, W. F. Liu, and C. S. Chen, Cell polarity triggered by cell-cell adhesion via E-cadherin, J. Cell Sci. **122**, 905 (2009).

[31] E. Scarpa, A. Szabó, A. Bibonne, E. Theveneau, M. Parsons, and R. Mayor, Cadherin switch during EMT in neural crest cells leads to contact inhibition of locomotion via repolarization of forces, Develop. Cell **34**, 421 (2015).

[32] I. Y. Wong, S. Javaid, E. A. Wong, S. Perk, D. A. Haber, M. Toner, and D. Irimia, Collective and individual migration following the epithelial–mesenchymal transition, Nat. Mater. **13**, 1063 (2014).

[33] D. B. Brückner, M. Schmitt, A. Fink, G. Ladurner, J. Flommersfeld, N. Arlt, E. Hannezo, J. O. Rädler, and C. P. Broedersz, Geometry adaptation of protrusion and polarity dynamics in confined cell migration, Phys. Rev. X **12**, 031041 (2022).

[34] J. Flommersfeld, S. Stöberl, O. Shah, J. O. Rädler, and C. P. Broedersz, Geometry-sensitive protrusion growth di-

rects confined cell migration, Phys. Rev. Lett. **132**, 098401 (2024).
[35] L. Edelstein-Keshet, W. R. Holmes, M. Zajac, and M. Dutot, From simple to detailed models for cell polarization, Phil. Trans. R. Soc. B **368**, 20130003 (2013).
[36] P. Sens, Stick-slip model for actin-driven cell protrusions, cell polarization, and crawling, Proc. Natl. Acad. Sci. USA **117**, 24670 (2020).
[37] B. A. Camley, Y. Zhang, Y. Zhao, B. Li, E. Ben-Jacob, H. Levine, and W. J. Rappel, Polarity mechanisms such as contact inhibition of locomotion regulate persistent rotational motion of mammalian cells on micropatterns, Proc. Natl. Acad. Sci. USA **111**, 14770 (2014).
[38] B. Smeets, R. Alert, J. Pešek, I. Pagonabarraga, H. Ramon, and R. Vincent, Emergent structures and dynamics of cell colonies by contact inhibition of locomotion, Proc. Natl. Acad. Sci. USA **113**, 14621 (2016).
[39] N. Sepúlveda, L. Petitjean, O. Cochet, E. Grasland-Mongrain, P. Silberzan, and V. Hakim, Collective cell motion in an epithelial sheet can be quantitatively described by a stochastic interacting particle model, PLoS Comput. Biol. **9**, e1002944 (2013).
[40] J. E. Ron, J. d'Alessandro, V. Cellerin, R. Voituriez, B. Ladoux, and N. S. Gov, Polarization and motility of one-dimensional multi-cellular trains, Biophys. J. **122**, 4598 (2023).
[41] K. Copenhagen, G. Malet-Engra, W. Yu, G. Scita, N. Gov, and A. Gopinathan, Frustration-induced phases in migrating cell clusters, Sci. Adv. **4**, eaar8483 (2018).
[42] E. Vercruysse, D. B. Brückner, M. Gómez-González, A. Remson, M. Luciano, Y. Kalukula, L. Rossetti, X. Trepat, E. Hannezo, and S. Gabriele, Geometry-driven migration efficiency of autonomous epithelial cell clusters, Nat. Phys. **20**, 1492 (2024).
[43] D. A. Kulawiak, B. A. Camley, and W. J. Rappel, Modeling contact inhibition of locomotion of colliding cells migrating on micropatterned substrates, PLoS Comput. Biol. **12**, e1005239 (2016).
[44] R. A. Desai, S. B. Gopal, S. Chen, and C. S. Chen, Contact inhibition of locomotion probabilities drive solitary versus collective cell migration, J. R. Soc. Interface **10**, 20130717 (2013).
[45] T. Zisis, D. B. Brückner, T. Brandstätter, W. X. Siow, J. d'Alessandro, A. M. Vollmar, C. P. Broedersz, and S. Zahler, Disentangling cadherin-mediated cell-cell interactions in collective cancer cell migration, Biophys. J. **121**, 44 (2022).
[46] R. Alert and X. Trepat, Physical models of collective cell migration, Annu. Rev. Condens. Matter Phys. **11**, 77 (2020).
[47] J. Löber, F. Ziebert, and I. S. Aranson, Collisions of deformable cells lead to collective migration, Sci. Rep. **5**, 9172 (2015).
[48] T. Bertrand, J. D'Alessandro, A. Maitra, S. Jain, B. Mercier, R. M. Mège, B. Ladoux, and R. Voituriez, Clustering and ordering in cell assemblies with generic asymmetric aligning interactions, Phys. Rev. Res. **6**, 023022 (2024).
[49] D. B. Brückner and C. P. Broedersz, Learning dynamical models of single and collective cell migration: A review, Rep. Prog. Phys. **87**, 056601 (2024).
[50] D. B. Brückner, P. Ronceray, and C. P. Broedersz, Inferring the dynamics of underdamped stochastic systems, Phys. Rev. Lett. **125**, 058103 (2020).
[51] E. Agliari, P. J. Sáez, A. Barra, M. Piel, P. Vargas, and M. Castellana, A statistical inference approach to reconstruct intercellular interactions in cell migration experiments, Sci. Adv. **6**, eaay2103 (2020).
[52] C. Joshi, S. Ray, L. M. Lemma, M. Varghese, G. Sharp, Z. Dogic, A. Baskaran, and M. F. Hagan, Data-driven discovery of active nematic hydrodynamics, Phys. Rev. Lett. **129**, 258001 (2022).
[53] N. Romeo, A. D. Hastewell, A. Mietke, and J. Dunkel, Learning developmental mode dynamics from single-cell trajectories, eLife **10**, e68679 (2021).
[54] See Supplemental Material at http://link.aps.org/supplemental/10.1103/3hhj-rt1n for additional experimental analysis and theoretical results, which also contains Refs. [8,33,48,55,81].
[55] D. B. Brückner, A. Fink, C. Schreiber, P. J. Röttgermann, J. O. Rädler, and C. P. Broedersz, Stochastic nonlinear dynamics of confined cell migration in two-state systems, Nat. Phys. **15**, 595 (2019).
[56] Y. Kalukula, M. Luciano, G. Charras, D. B. Brückner, and S. Gabriele, The actin cortex acts as a mechanical memory of morphology in confined migrating cells, bioRxiv:2024.08.05.606589.
[57] J. E. Ron, P. Monzo, N. C. Gauthier, R. Voituriez, and N. S. Gov, One-dimensional cell motility patterns, Phys. Rev. Res. **2**, 033237 (2020).
[58] I. C. Fortunato, D. B. Brückner, S. Grosser, L. Rossetti, M. Bosch-Padrós, J. Trebicka, P. Roca-Cusachs, R. Sunyer, E. Hannezo, and X. Trepat, Single cell migration along and against confined haptotactic gradients, bioRxiv (2024).
[59] T. D. Pollard, The cytoskeleton, cellular motility and the reductionist agenda, Nature (London) **422**, 741 (2003).
[60] R. M. Lee, M. I. Vitolo, W. Losert, and S. S. Martin, Distinct roles of tumor associated mutations in collective cell migration, Sci. Rep. **11**, 10291 (2021).
[61] T. Blick, E. Widodo, H. Hugo, M. Waltham, M. E. Lenburg, R. M. Neve, and E. W. Thompson, Epithelial mesenchymal transition traits in human breast cancer cell lines, Clin. Exp. Metastasis **25**, 629 (2008).
[62] D. Selmeczi, S. Mosler, P. H. Hagedorn, N. B. Larsen, and H. Flyvbjerg, Cell motility as persistent random motion: Theories from experiments, Biophys. J. **89**, 912 (2005).
[63] D. F. Milano, N. A. Ngai, S. K. Muthuswamy, and A. R. Asthagiri, Regulators of metastasis modulate the migratory response to cell contact under spatial confinement, Biophys. J. **110**, 1886 (2016).
[64] L. Lu, T. Guyomar, Q. Vagne, R. Berthoz, A. Torres-Sánchez, M. Lieb, C. Martin-Lemaitre, K. van Unen, A. Honigmann, O. Pertz, D. Riveline, and G. Salbreux, Polarity-driven three-dimensional spontaneous rotation of a cell doublet, Nat. Phys. **20**, 1194 (2024).
[65] Y. L. Chao, C. R. Shepard, and A. Wells, Breast carcinoma cells re-express E-cadherin during mesenchymal to epithelial reverting transition, Mol. Cancer **9**, 179 (2010).
[66] X. Serra-Picamal, V. Conte, R. Vincent, E. Anon, D. T. Tambe, E. Bazellieres, J. P. Butler, J. J. Fredberg, and X. Trepat, Mechanical waves during tissue expansion, Nat. Phys. **8**, 628 (2012).
[67] D. T. Tambe, C. Corey Hardin, T. E. Angelini, K. Rajendran, C. Y. Park, X. Serra-Picamal, E. H. Zhou, M. H. Zaman, J. P. Butler, D. A. Weitz, J. J. Fredberg, and X. Trepat, Collective cell guidance by cooperative intercellular forces, Nat. Mater. **10**, 469 (2011).

[68] S. Lo Vecchio, O. Pertz, M. Szopos, L. Navoret, and D. Riveline, Spontaneous rotations in epithelia as an interplay between cell polarity and boundaries, Nat. Phys. **20**, 322 (2024).

[69] M. Basan, J. Elgeti, E. Hannezo, W. J. Rappel, and H. Levine, Alignment of cellular motility forces with tissue flow as a mechanism for efficient wound healing, Proc. Natl. Acad. Sci. USA **110**, 2452 (2013).

[70] B. D. Hoffman, C. Grashoff, and M. A. Schwartz, Dynamic molecular processes mediate cellular mechanotransduction, Nature (London) **475**, 316 (2011).

[71] G. F. Weber, M. A. Bjerke, and D. W. DeSimone, A mechanoresponsive cadherin-keratin complex directs polarized protrusive behavior and collective cell migration, Develop. Cell **22**, 104 (2012).

[72] B. R. Acharya, A. Nestor-Bergmann, X. Liang, S. Gupta, K. Duszyc, E. Gauquelin, G. A. Gomez, S. Budnar, P. Marcq, O. E. Jensen, Z. Bryant, and A. S. Yap, A mechanosensitive RhoA pathway that protects epithelia against acute tensile stress, Develop. Cell **47**, 439 (2018).

[73] M. Nakagawa, M. Fukata, M. Yamaga, N. Itoh, and K. Kaibuchi, Recruitment and activation of Rac1 by the formation of E-cadherin-mediated cell-cell adhesion sites, J. Cell Sci. **114**, 1829 (2001).

[74] T. Vicsek, A. Czirok, 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).

[75] 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).

[76] G. Peyret, R. Mueller, J. d'Alessandro, S. Begnaud, P. Marcq, R. M. Mège, J. M. Yeomans, A. Doostmohammadi, and B. Ladoux, Sustained oscillations of epithelial cell sheets, Biophys. J. **117**, 464 (2019).

[77] T. Brandstätter, D. B. Brückner, Y. L. Han, R. Alert, M. Guo, and C. P. Broedersz, Curvature induces active velocity waves in rotating multicellular spheroids, Nat. Commun. **14**, 1643 (2023).

[78] C. Malinverno, S. Corallino, F. Giavazzi, M. Bergert, Q. Li, M. Leoni, A. Disanza, E. Frittoli, A. Oldani, E. Martini, T. Lendenmann, G. Deflorian, G. V. Beznoussenko, D. Poulikakos, K. H. Ong, M. Uroz, X. Trepat, D. Parazzoli, P. Maiuri, W. Yu *et al.*, Endocytic reawakening of motility in jammed epithelia, Nat. Mater. **16**, 587 (2017).

[79] W. Kang, J. Ferruzzi, C. P. Spatarelu, Y. L. Han, Y. Sharma, S. A. Koehler, J. A. Mitchel, A. Khan, J. P. Butler, D. Roblyer, M. H. Zaman, J. A. Park, M. Guo, Z. Chen, A. F. Pegoraro, and J. J. Fredberg, A novel jamming phase diagram links tumor invasion to non-equilibrium phase separation, iScience **24**, 103252 (2021).

[80] A. Palamidessi, C. Malinverno, E. Frittoli, S. Corallino, E. Barbieri, S. Sigismund, G. V. Beznoussenko, E. Martini, M. Garre, I. Ferrara, C. Tripodo, F. Ascione, E. A. Cavalcanti-Adam, Q. Li, P. P. Di Fiore, D. Parazzoli, F. Giavazzi, R. Cerbino, and G. Scita, Unjamming overcomes kinetic and proliferation arrest in terminally differentiated cells and promotes collective motility of carcinoma, Nat. Mater. **18**, 1252 (2019).

[81] H. Aragaki, K. Ogoh, Y. Kondo, and K. Aoki, LIM Tracker: a software package for cell tracking and analysis with advanced interactivity, Sci. Rep. **12**, 2702 (2023).

[82] A. Szabó, M. Melchionda, G. Nastasi, M. L. Woods, S. Campo, R. Perris, and R. Mayor, In vivo confinement promotes collective migration of neural crest cells, J. Cell Biol. **213**, 543 (2016).

[83] J. M. Belmonte, G. L. Thomas, L. G. Brunnet, R. M. De Almeida, and H. Chaté, Self-propelled particle model for cell-sorting phenomena, Phys. Rev. Lett. **100**, 248702 (2008).