A self-exciting point process to study multicellular spatial signaling patterns

Archit Verma[a,1], Siddhartha G. Jena[b,1], Danielle R. Isakov[b], Kazuhiro Aoki[c,d,e], Jared E. Toettcher[b], and Barbara E. Engelhardt[f,g,2]

[a]Department of Chemical and Biological Engineering, Princeton University, Princeton, NJ 08544; [b]Department of Molecular Biology, Princeton University, Princeton, NJ 08544; [c]National Institute of Basic Biology, National Institutes of Natural Sciences, Okazaki 444-8585, Japan; [d]Exploratory Research Center on Life and Living Systems, National Institutes of Natural Sciences, Okazaki 444-8787, Japan; [e]International Research Collaboration Center, National Institutes of Natural Sciences, Tokyo 105-0001, Japan; [f]Department of Computer Science, Princeton University, Princeton, NJ 08540; and [g]Center for Statistics and Machine Learning, Princeton University, Princeton, NJ 08540

Edited by Kathryn Roeder, Carnegie Mellon University, Pittsburgh, PA, and approved June 10, 2021 (received for review January 14, 2021)

Multicellular organisms rely on spatial signaling among cells to drive their organization, development, and response to stimuli. Several models have been proposed to capture the behavior of spatial signaling in multicellular systems, but existing approaches fail to capture both the autonomous behavior of single cells and the interactions of a cell with its neighbors simultaneously. We propose a spatiotemporal model of dynamic cell signaling based on Hawkes processes—self-exciting point processes—that model the signaling processes within a cell and spatial couplings between cells. With this cellular point process (CPP), we capture both the single-cell pathway activation rate and the magnitude and duration of signaling between cells relative to their spatial location. Furthermore, our model captures tissues composed of heterogeneous cell types with different bursting rates and signaling behaviors across multiple signaling proteins. We apply our model to epithelial cell systems that exhibit a range of autonomous and spatial signaling behaviors basally and under pharmacological exposure. Our model identifies known drug-induced signaling deficits, characterizes signaling changes across a wound front, and generalizes to multichannel observations.

point process | Hawkes process | keratinocytes | kinase networks | cell signaling

Complex life is largely characterized by multicellular structures (1). Classical multicellular processes such as the patterning of cells within a tissue and the precise spatial arrangement of tissues within an organ are the product of different gene expression programs organized carefully over space and time (2). These different programs emerge from both intracellular pathways governing gene and protein expression on a single-cell level and the intercellular signaling that allows cells near one another to interact. Understanding how these networks are regulated as well as the factors leading to their dysfunction is a topic of active research (3).

Intracellular signaling is a term used to describe information-carrying modifications of proteins in a single cell. One example of this is the extracellular signal-regulated kinase (Erk), which is activated by phosphorylation in response to changes in the cell's environment. This pathway is also called the Ras/Erk signaling pathway since signaling originates at the membrane protein Ras (rat sarcoma). Signaling proteins can then operate on downstream effectors such as transcription factors that regulate gene expression.

Intercellular signaling specifically involves signaling as a result of an input delivered by a neighboring cell. Often, this involves the release of ligands from one cell that bind to receptors on a neighboring cell and cause a change in behavior of the neighbor cell. Both intra- and intercellular signaling may make use of the same signaling protein. For example, Erk can be activated by the presence of growth factors in the surrounding media or upon cleavage and binding of growth factors from an adjacent cell.

Nearly every cell in a physiological context is simultaneously processing information about its own state (intracellular) as well as the states of cells around it (intercellular). Therefore, these two modes of communication may interact in complex and unexpected ways, especially when they make use of the same signaling proteins. Decoupling their relative effects on cell state is challenging and often requires invasive perturbations such as pharmacological inhibitors that may have unforeseen consequences on cell or tissue health. Nevertheless, estimating the relative contributions of intrinsic cellular behavior and extrinsic spatial signaling is an important goal for understanding multicellular systems. This is particularly true with the advent of cellular imaging modalities that allow us to visualize signaling behaviors in single cells, heterogeneous multicellular ensembles, and even in vivo tissue (4).

Here, we focus on a case of one signaling pathway being used to convey information about both intra- and intercellular cell states. The mammalian Ras/Erk pathway has been found to display transient "pulses" consisting of pathway activation followed by rapid deactivation in a range of epithelial cell types (5, 6). These pulses can be modulated by environmental context such as the presence of certain growth factors, as well as

Significance

Cells are under constant pressure to integrate information from both their environment and internal cellular processes. However, these effects often use the same signaling pathways, making autonomous and coupled signaling difficult to decouple from one another. Here, we present a statistical modeling framework, the cellular point process (CPP), that decouples these two modes of signaling using videos of living, actively signaling cells as input. Our model reveals modulation of autonomous and coupled signaling parameters in a number of contexts ranging from pharmacological treatment to wound healing that were previously unavailable. The CPP enhances our understanding of cellular information processing and can be extended to a wide range of systems.

Author contributions: A.V., S.G.J., J.E.T., and B.E.E. designed research; A.V., S.G.J., D.R.I., and K.A. performed research; A.V., S.G.J., D.R.I., and K.A. contributed new reagents/analytic tools; A.V. and S.G.J. analyzed data; and A.V., S.G.J., J.E.T., and B.E.E. wrote the paper.

Competing interest statement: B.E.E. is on the Scientific Advisory Board of Freenome, Celsius Therapeutics, and Creyon Bio.

This article is a PNAS Direct Submission.

[1]A.V. and S.G.J. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: bee@princeton.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2026123118/-/DCSupplemental.

Published August 6, 2021.

physical perturbations such as a wound (4). Moreover, cells have been found to "transmit" pulses of activity from cell to cell, and to pulse autonomously, suggesting that the Ras/Erk pathway is involved in both intra- and intercellular signaling (5, 6). Since epithelial tissues largely derive their functionality from intercellular communication and multicellular behavior, which regulates their differentiation and growth, it is likely that Ras/Erk pulses in epithelia are not simply an epiphenomenon but rather functionally linked to tissue-level dynamics, for instance, to the ability of a cell to leave its stem cell niche and undergo subsequent differentiation (7).

Models have been proposed that approximate spatial signaling patterns to simulate signaling behaviors of cells. Dynamical models based on diffusion processes describe the behavior of biological systems spatially by approximating the tissue as a continuum (8, 9). Cellular automata models equip each member of a discrete set of agents with a primitive set of directives and then observe how the resulting system of agents evolves across time (10). To our knowledge, inferring signaling parameters for these models from observations is rare, limiting their applicability to observational data from experimental systems.

In this paper, we introduce a statistical approach to modeling the pulsing times of each cell in a neighborhood as a function of their spatial organization. We treat collections of signaling pulses as realizations of a point process—a stochastic model of events over time or space (11). Self-exciting point processes, known as Hawkes processes, have been successfully used to model social media interactions (12, 13), financial time series (14, 15), neuron spike trains (16), and events along DNA sequences (17, 18), as well as a range of other time-varying or space-varying processes (19, 20). These processes allow us to explicitly tease apart the rate of cell pulsing and the influence of a pulsing event at one cell on the probability of a pulsing event in that same cell and in neighboring cells as a function of distance. Previous work has mostly focused on learning the connectivity of a network given data; we are interested in learning the strength of connections based on a given spatial network.

Our model—the cellular point process (CPP)—quantitatively estimates the base rate of pulsing, the rate of intracellular signaling, and the strength of intercellular signaling using experimental data. The CPP is an adaptation of the original Hawkes process that limits effective signaling to cells that are within a cutoff distance from one another. The CPP model parameters are estimated by maximum-likelihood methods using data that capture pulse times and spatial coordinates for each cell annotated in an imaging experiment. These types of experiments are becoming increasingly tractable in many laboratory environments (6). We demonstrate a correlation between the duration (i.e., number of signaling events) and scale (i.e., number of cells) quantified in an experiment and the accuracy of the intra- and intercellular signaling rates inferred by the CPP. This suggests that our model can be applied to a wide range of imaging data to deconvolve pulsing rates due to intra- and intercellular signaling patterns.

We validate the CPP's ability to estimate the relative contributions of intra- and intercellular signaling on pulsing rates in simulated data where these contributions are known. We then analyze mouse epidermal stem cells, or keratinocytes, which display naturally occurring Ras/Erk dynamics in and ex vivo (4, 6). We quantify the decrease in spatial signaling when these cells are treated with a known cell-signaling inhibitor compared to untreated cells. Then, we examine and disentangle the inter- and intracellular contributions of a variety of pharmacological kinase inhibitors on Erk bursting dynamics in keratinocytes. Next, we study the contributions of inter- and intracellular signaling on the response of Madin-Darby canine kidney (MDCK) cells to an acute wounding event, finding that both factors change as a function of distance away from the wound. Finally, we demonstrate that the CPP model estimates how multiple reporter channels interact with each other across cells.

Results

The CPP model treats each peak in a cell as an event in a self-exciting point process, where events in one cell influence the likelihood of an event occurring in the future in that cell and in neighboring cells (Fig. 1). The CPP model estimates five parameters from a list of events in different cells with known spatial organization:

- μ, the baseline, autonomous frequency of events;
- a, the strength of signal each event emits to the cell neighborhood (higher a indicates stronger signaling effects);
- a_{self}, the strength of signal each event emits to self-excite the cell of the event (higher a_{self} indicates stronger intracellular effects);
- b, the variance of a log-normal kernel that defines the effect of any event on the conditional intensity of its neighbors (higher b corresponds to larger variance in time between events); and
- b_{self}, the variance of a log-normal kernel that defines the effect of any event on the conditional intensity of its own cell.

We first validate the CPP model's ability to identify spatial and autonomous signaling on simulated data where these factors are known a priori. We then turn to an experimentally tractable system of mouse epidermal stem cells, or keratinocytes, that display naturally occurring Ras/Erk dynamics in and ex vivo (4, 6). We estimate the natural intercellular and intracellular signaling effects in these cells and validate the CPP model by quantifying decreases in spatial signaling when cells are treated with a known signaling inhibitor. Next, we estimate the spatial and autonomous signaling effects of a variety of drugs on keratinocytes from a prior assay (6). We also quantify the spatial and autonomous signaling effects of three drugs on cell behavior during wound healing stratified by distance from the wound (21). Finally, we demonstrate that the model can estimate parameters from more complex histories with multiple channels per cell using data with two fluorescence channels from mouse keratinocytes.

Spatial Point Processes Estimate Parameters from Simulated Data. We first verify that the model accurately estimates parameters from simulated data. Observations were simulated from the generative model across a range of observed cells and total number of observed events. We find that the CPP is able to accurately estimate the parameters of the generative model with low normalized mean square error (NMSE) (Table 1 and Fig. 2; see *Materials and Methods* for equations). Control estimates of the parameters, described in *Materials and Methods*, all perform worse than our estimates in NMSE.

The accuracy of the parameter estimates varies over the number of cells and number of events observed. The estimates for the parameters degrade as the number of peaks decreases (Fig. 2 *B*, *E*, *H*, *K*, and *N*). Estimation of parameters a and b improves substantially with more cells (Fig. 2 *K* and *N*). The confidence intervals are mainly a function of the number of peaks observed, particularly for a and b (Fig. 2 *L* and *O*). The accuracy of estimates also depends on the true parameter values. The autonomous parameter μ and intercellular signaling duration parameter b tend to be underestimated when their true value is relatively large (Fig. 2 *A* and *M*). Estimation for intracellular signaling duration parameters a_{self} and b_{self} are less accurate than the other parameters (see Table 3), so we ascribe less significance to this estimate in later sections. Nevertheless, the small errors of the parameter estimates demonstrate that the CPP model accurately deconvolves signaling parameters from experimental data. The data that we collected have around 150 cells and 1,000 peaks,

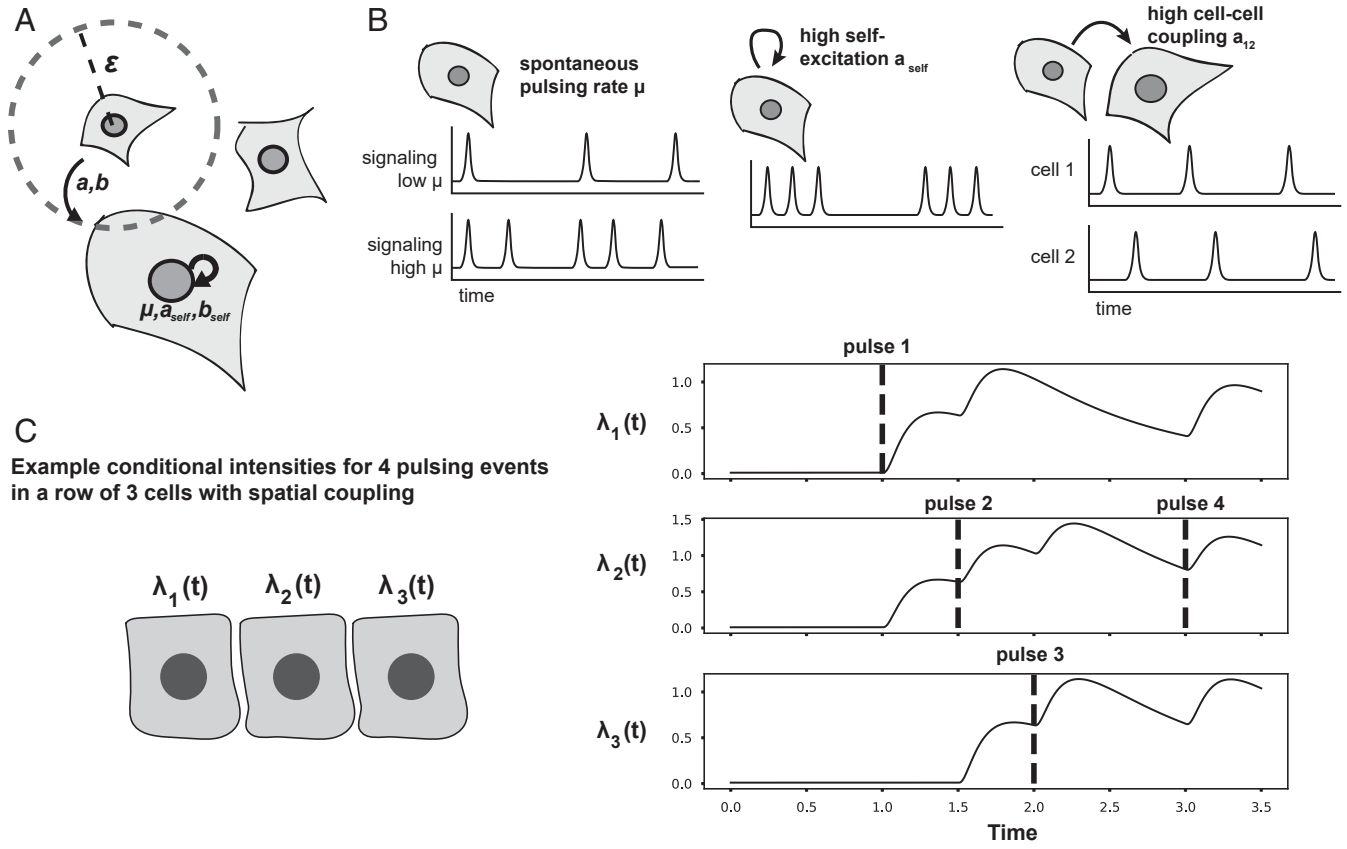

Fig. 1. (*A*) Three cells demonstrating autonomous pulsing μ, self-signaling effects parameterized by a_{self}, b_{self}, and cell-proximity specific effects within some radius ϵ, with magnitude of those effects parameterized by a, b. (*B*) Effect of parameters in *A* on signaling pulses in cells. (*C*) Example conditional intensity ($\lambda(t)$) plots for four pulses across three cells, where cell 1's initial pulse increases the expected number of pulses for cell 3 and itself through spatial coupling. High intensity means more expected pulses.

a region where CPP's estimates for each parameter are close to the true values in simulations.

Spatial Point Processes Capture the Effect of Pharmacological Treatments on Keratinocyte Behavior. Next, we evaluate the CPP model's ability to identify changes in signaling behavior across pharmacological treatments of cells. Mouse basal keratinocytes are epidermal stem cells that form monolayers in and ex vivo. Studies have shown dynamic signaling behavior in the Ras/Erk pathway of these cells linked to spatial patterning (5). TNF-α protease inhibitor (TAPI-1) is a matrix metalloproteinase inhibitor (22) that prevents spatial signaling and Erk activation between cells (5). We used a KTR fluorescent marker (23) to image Erk concentration over time across sheets of keratinocyte cells dosed with 5, 10, and 20 μM of TAPI-1, as well as a control group of untreated cells.

Table 1. NMSE and normalized standard deviation of NMSE (NSTD) of CPP-estimated parameters and a control model to ground-truth parameters from simulated data

Parameter	NMSE (NSTD)	Control NSME (NSTD)
μ	0.09 (0.30)	1,226.51 (33.00)
a	0.22 (0.34)	79.87 (8.33)
a_{self}	0.65 (0.67)	19,914.99 (181.23)
b	0.05 (0.15)	0.08 (0.28)
b_{self}	0.35 (0.23)	0.49 (0.27)

Our CPP model is able to quantitatively capture the change in keratinocyte signaling behavior as a function of TAPI-1 concentration (Fig. 3). We find that the estimated autonomous $\hat{\mu}$ parameter and strength of spatial signaling $\hat{a}$ decrease with increased TAPI-1 concentration (Table 2). We also observe that the self-exciting parameter a_{self} does decrease but only by about 20%. We would not expect TAPI-1 to change the self-excitation signaling of keratinocytes. The kernel parameters $\hat{b}$ and $\hat{b}_{self}$ increase with increased TAPI-1 concentration, representing an increase in the time between peaks (Table 2). We note that the values for $\hat{b}$ and $\hat{b}_{self}$ are the same; this is due to the inference methods. We estimate b_{self} as a multiplier of b, initialized at one. The nearly identical values indicate that the gradient updates for the b_{self} multiplier were small given these data. This finding highlights an important advantage of our model. In contrast to simpler approaches, such as the Ising model, the CPP allows for calculating an explicit term for memory, or self-excitation—i.e., the propensity for a cell to change state as a function of state changes that have occurred in its recent history.

Pairwise nonparametric Mann–Whitney U tests between TAPI-1 concentrations show substantial decreases in $\hat{\mu}$ between 0 and 5 μM ($U = 0$, $P \leq 0.01$) and 10 and 20 μM ($U = 0$, $P \leq 0.005$), and also show decreases in a between 0 and 5 μM ($U = 5$, $P \leq 0.01$; Fig. 3 *A* and *B*). We also find a decrease in signaling associated with an increase in TAPI-1 dose using the likelihoods of the model. To do this, we compare the likelihood of the estimated model to a model where all pulses are due to autonomous signaling parameter μ; in other words,

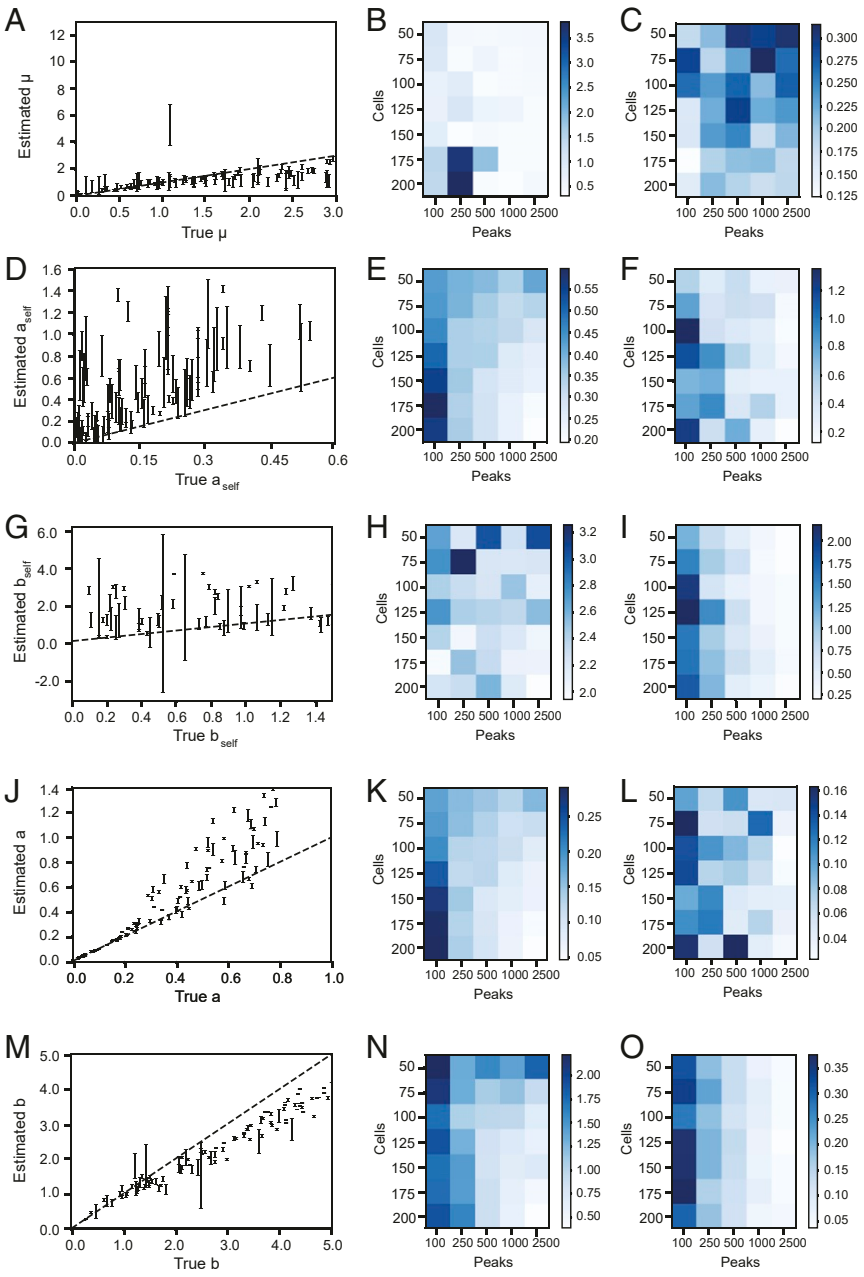

Fig. 2. (*Left*) Scatter plots of true versus estimated (*A*) μ, (*D*) a_{self}, (*G*) b_{self}, (*J*) *a*, and (*M*) *b*. Dot color represents the number of cells and dot size represents the number of events for that simulation. (*Middle*) Heatmap of mean squared error (MSE) of estimated (*B*) μ, (*E*) a_{self}, (*H*) b_{self}, (*K*) *a*, and (*N*) *b* for different simulation parameters for number of cells and number of peaks. (*Right*) Heatmap of confidence intervals (CI) of estimated (*C*) μ, (*F*) a_{self}, (*I*) b_{self}, (*L*) *a*, and (*O*) *b* for different simulation parameters for number of cells and number of peaks.

$a = 0$. The estimated $\hat{\mu}$ is the ratio of $\frac{\text{number of peaks}}{\text{number of cells} \times \text{total time}}$. When we take the difference of the log-likelihoods, we observe that the difference between CPP and a control model decreases as TAPI-1 concentration increases (Fig. 3*E*), including a substantial decrease between 0 and 5 μM ($U = 3, P \leq 0.01$), although the difference always remains positive. This means that the CPP model is better at explaining the data relative to a fully autonomous model at low TAPI-1 concentrations. Alternatively, we calculate the contribution of the spatial, self-exciting, and autonomous influences at each peak. We observe that the average percentage contribution from spatial signaling across peaks decreases as TAPI-1 concentration increases (Fig. 3*D*). The ability of the CPP model to quantify this biological inhibition demonstrates its ability to analyze spatial signaling in complex systems.

We extend this analysis to evaluate the effects of a variety of drugs on spatial signaling in keratinocytes. Recent work (6) sought to quantify the effects of over 400 receptor tyrosine kinase inhibitors (RTKi) on endogenous keratinocyte Ras/Erk dynamics. We fitted our CPP model on time series from treated cells from this study, with experiments across 432 drugs and 18 control dimethylsulfoxide (DMSO) samples, to determine whether the autonomous or spatial components, or both, are substantially affected by targeted RTK inhibition.

The original analysis of these data (6) divided drugs into three categories, which also took into account the "set point"

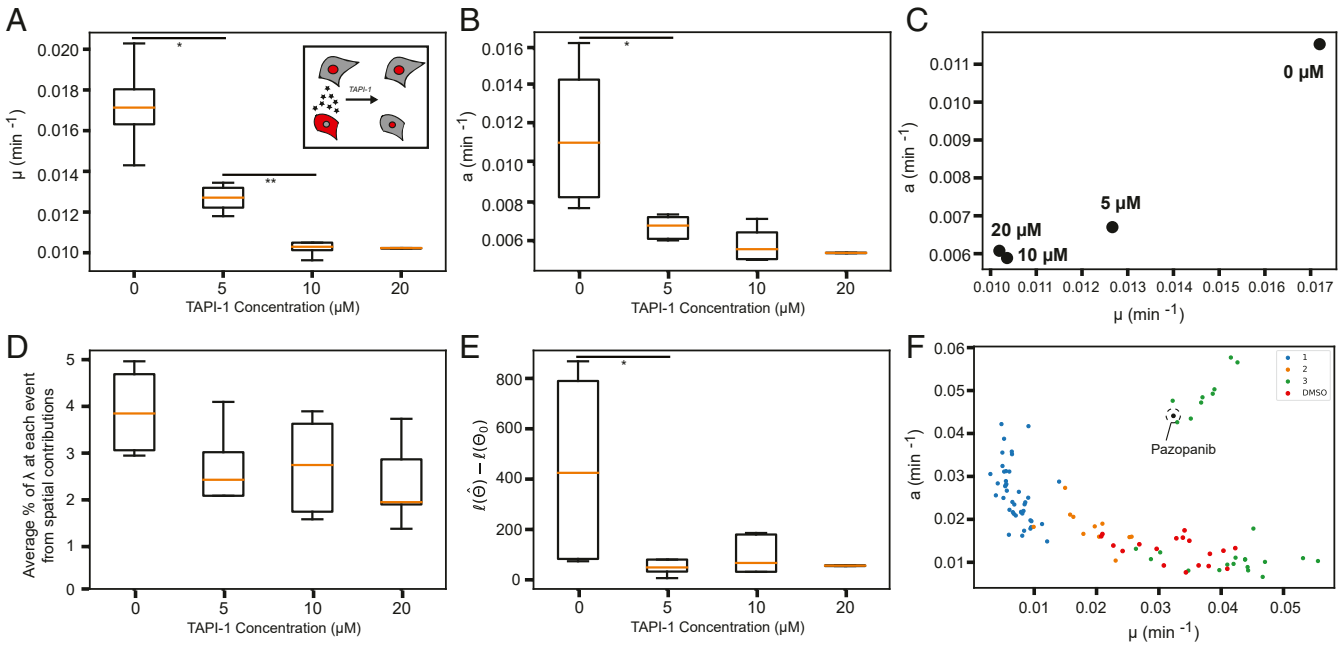

Fig. 3. (*A* and *B*) Estimated parameters μ and a for keratinocytes as a function of TAPI-1 concentration. Black points represent individual wells. Schematic in *A*, *Inset* shows the effect of TAPI-1 on intercellular signaling. (*C*) Scatter plot of average estimates of μ versus a for each TAPI-1 concentration. (*D*) Average contribution of spatial signaling to pulses in each TAPI-1 concentration. (*E*) Difference in unnormalized log-likelihood of spatial model versus a spatial-free control model for all TAPI-1 concentrations and replicates. (*F*) Class 1, 2, and 3 drugs (6) on a plot of μ versus a. Statistical significance is calculated using a Mann–Whitney *U* test, with $*P < 0.01$ and $**P < 0.005$.

or baseline level of Erk activity. Class 1 drugs reduced Erk activity to extremely low levels, corresponding to μ close to 0. Class 2 drugs, on the other hand, increased Erk activity to a high constant level. Since a pulsing event is defined as a local maximum in the Erk activity trace, these cells demonstrated few pulsing events since the pathway could likely not be activated beyond this high level. Since our model does not capture mean Erk activity, but rather the events where activity changes, we expect these drugs to have lower autonomous pulsing parameter μ in the CPP model. Finally, class 3 drugs increased Erk activity over time by increasing the pulse frequency (6), which corresponds to a higher estimated μ value in our model.

The CPP model finds differences across these three classes and in comparison to untreated DMSO controls (Fig. 3*F*). We find that class 1 ($\hat{\mu} = 0.025\ \text{min}^{-1}$) and 2 ($\hat{\mu} = 0.013\ \text{min}^{-1}$) drugs have lower mean autonomous activity $\hat{\mu}$ than DMSO controls ($\hat{\mu} = 0.03\ \text{min}^{-1}$), using a t test to compare classes to DMSO (t statistic $= -21.7$, $P \leq 2.2 \times 10^{-16}$ for class 1 drugs and t statistic $= -5.63$, $P \leq 5.7 \times 10^{-6}$ for class 2 drugs). Class 3 drugs have higher autonomous activity $\hat{\mu}$ ($\hat{\mu} = 0.04\ \text{min}^{-1}$, t statistic $= 3.69$, $P \leq 0.0006$). While class 1 and 2 drugs have spatial signaling parameter $\hat{a}$ in a close range (approximately 0.01 to $0.04\ \text{min}^{-1}$), we find that class 3 drugs diverge between low signaling, $\hat{a} < 0.02\ \text{min}^{-1}$, and high signaling behavior ($\hat{a} > 0.04\ \text{min}^{-1}$). An interesting example is Pazopanib ($\hat{a} = 0.044\ \text{min}^{-1}$), a class 3 drug that increases signaling relative to DMSO (mean $\hat{a} = 0.013\ \text{min}^{-1}$) that is known to target the membrane proteins such as RIPK1 and VEGFR (6, 24). Interactions with membrane proteins would be expected to modulate intercellular signaling. We note that the confidence intervals are large relative to the parameter estimates. The average 95% confidence interval for $\hat{\mu}$ is $\pm 0.020\ \text{min}^{-1}$ and the average 95% confidence interval is ± 0.197 for $\hat{a}_{self}\ \text{min}^{-1}$. Thus, these results should be interpreted cautiously and replications are required for each treatment to draw stronger conclusions.

While estimates for autonomous parameter $\hat{\mu}$ and signaling strength parameter $\hat{a}$ both decrease as TAPI-1 dose concentration increased, in the drug screen increasing $\hat{\mu}$ mostly corresponds to decreasing $\hat{a}$. The divergence in class 3 drugs, however, indicates that both parameters are necessary to fully characterize the signaling system. One would expect that the addition of TAPI-1 would decrease the pulsing of the high-$\hat{a}$ class 3 drugs but would leave the pulsing in low-$\hat{a}$ class 3 drugs mostly unaltered. The estimation of these distinct signaling parameters adds nuance to our understanding of the cell response to various pharmacological agents beyond basic statistics such as frequency and duration of pulses.

CPP Quantifies Trends in Cell Signaling in a Wound Healing Context. MDCK cells move to close an artificially inflicted wound in vitro while expressing a live cell reporter of Ras/Erk activity

Table 2. Mean estimated parameters and standard error of the mean (in parentheses) for each TAPI-1 dose concentration

Concentration	μ (10^{-2} min^{-1})	a (10^{-2} min^{-1})	a_{self} (10^{-2}min^{-1})	b (10^{-2} min)	b_{self} (10^{-2} min)
No treatment	1.72 (0.12)	1.15 (0.20)	0.83 (0.10)	0.29 (0.17)	0.29 (0.17)
5 μM	1.27 (0.03)	0.67 (0.03)	0.78 (0.07)	0.66 (0.18)	0.66 (0.17)
10 μM	1.03 (0.02)	0.59 (0.04)	0.71 (0.03)	0.57 (0.20)	0.57 (0.20)
20 μM	1.02 (0.02)	0.61 (0.07)	0.68 (0.01)	0.69 (0.17)	0.69 (0.17)

(21). Experiments were performed in the presence of DMSO (control), TAPI-1, or Trametinib (MEK inhibitor). Notably, as can be seen at the 6-h timepoint, ERK activity was high in the cells closest to the wound (the wound front) in both DMSO and TAPI-1 conditions, but this activity was abrogated in the presence of pathway inhibition by Trametinib. On the other hand, ERK activity was heterogeneous in the submarginal cells behind the wound front in control conditions, but much lower in the presence of TAPI-1 and similarly low in Trametinib conditions. We analyzed these spatiotemporal data using TrackMate to obtain pulse information and cell positions from the cell images over time. We then fitted our CPP model with these data. Cells were binned into their relative distance from the wound edge; the total width of the cell sheet from the inner edge of the field of view to the wound was split into 10 bins, and the CPP model was fitted for each bin (Fig. 4 *A*–*E*). We note for clarity that our kernel still implements the distance cutoff ϵ and that the relative distance from the wound edge is a separate parameter.

The estimated value of autonomous pulsing parameter $\hat{\mu}$ peaked next to the site of the wound in control (DMSO-treated) cells (Fig. 4*D*). On the other hand, the addition of TAPI-1 prewounding resulted in a much lower autonomous pulsing rate relative to DMSO in this region (Wilcoxon signed-rank statistic = 0, $P \leq 0.008$; red curve, Fig. 4*D*). Trametinib, an Erk inhibitor, decreased autonomous pulsing even further relative to DMSO, as would be expected (Wilcoxon signed-rank statistic = 0, $P \leq 0.008$). The estimated value of intercellular signaling strength parameter $\hat{a}$ also decreased from DMSO to TAPI-1 (Wilcoxon signed-rank statistic = 2, $P \leq 0.023$) and decreased further still in Trametinib-treated cells (Wilcoxon signed-rank statistic = 0, $P \leq 0.008$; Fig. 4*E*). Taken together, these results suggest, in line with previous work (21), that both cell-autonomous and cell-to-cell signaling effects occur with specific spatial organization in response to a wound and can be abrogated to different extents through pharmacological inhibition. This suggests that both cell-autonomous and cell-to-cell Ras/Erk signaling may be important factors in allowing MDCK cells to close a wound. This may be why TAPI-1–treated and Trametinib-treated cells fail to fully heal a wound over the full 12-h time course (Fig. 4 *B* and *C*), where the control wound is fully healed at the 12-h timepoint.

CPP Quantifies Signaling-to-Gene Relationships Using Multichannel Learning. The Ras/Erk pathway is responsible for immediate activation of a family of genes called immediate early genes (IEGs; Fig. 5*A*). We leveraged a system that allowed us to engineer mouse keratinocytes to express a destabilized green fluorescent protein (dGFP) with a half-life of ~1 h, under the control of the minimal promoter of the IEG fibroblast osteogenic sarcoma (Fos) (Fig. 5*B*). These cells were imaged for 24 h and

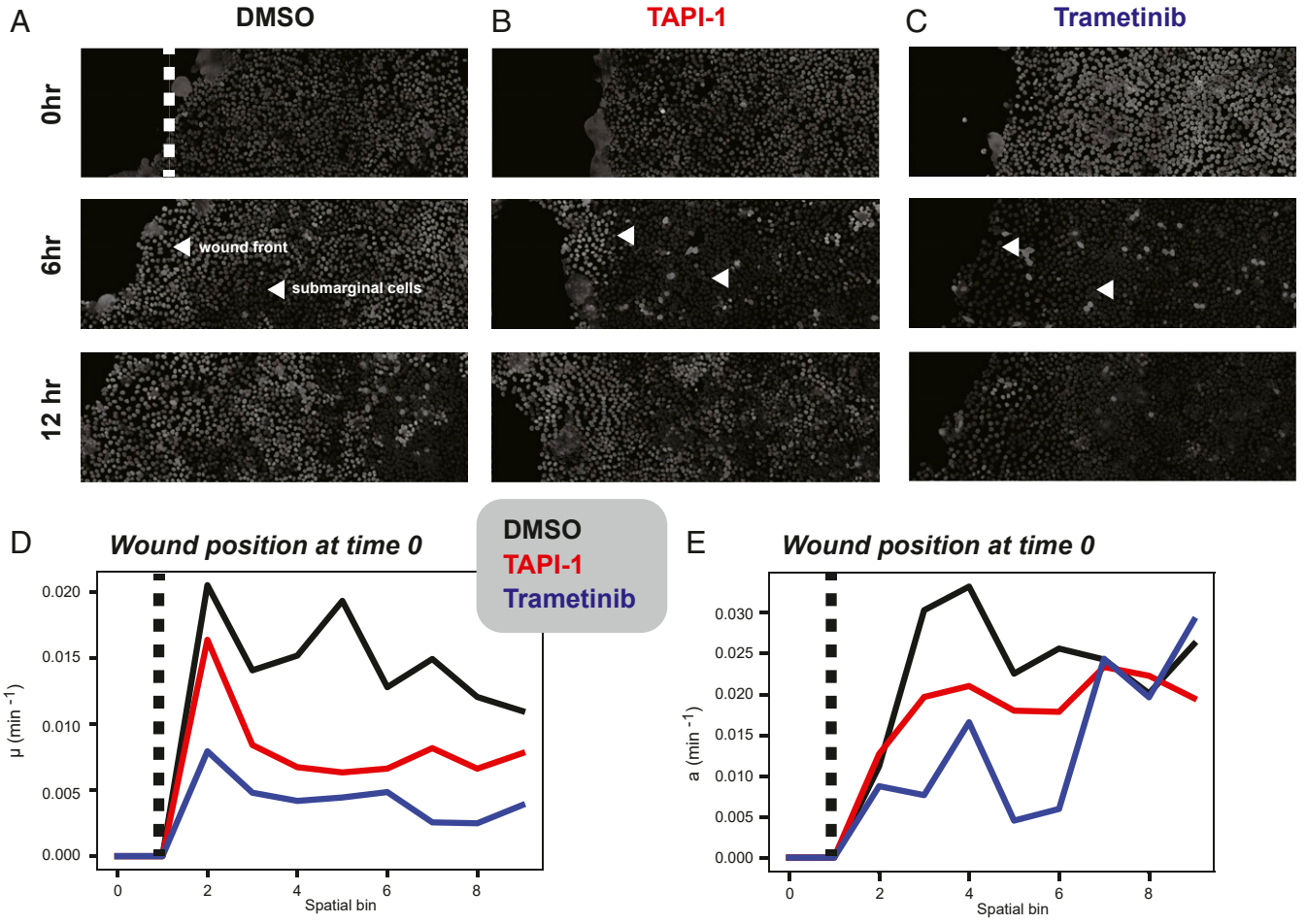

Fig. 4. (*A*–*C*) Time course data of MDCK cells moving to heal an inflicted wound at 0, 6, and 12 h postwounding. Dark cells denote low Erk activity, while light cells denote high Erk activity. A representative wound is marked by a dashed line in *A*, *Top*, and wound front cells/submarginal cells are marked by white arrows at the 6-h timepoints in *A*–*C*, *Middle*. (*D*) x binning captures cells at different distances away from the site of a wound. μ is estimated for bins at different distances away from the wound. (*E*) *a* estimates for spatial bins at different distances away from the wound. For both *D* and *E*, black lines represent DMSO cells, the red lines represent TAPI-1–treated cells, and the blue lines represent Trametinib-treated cells.

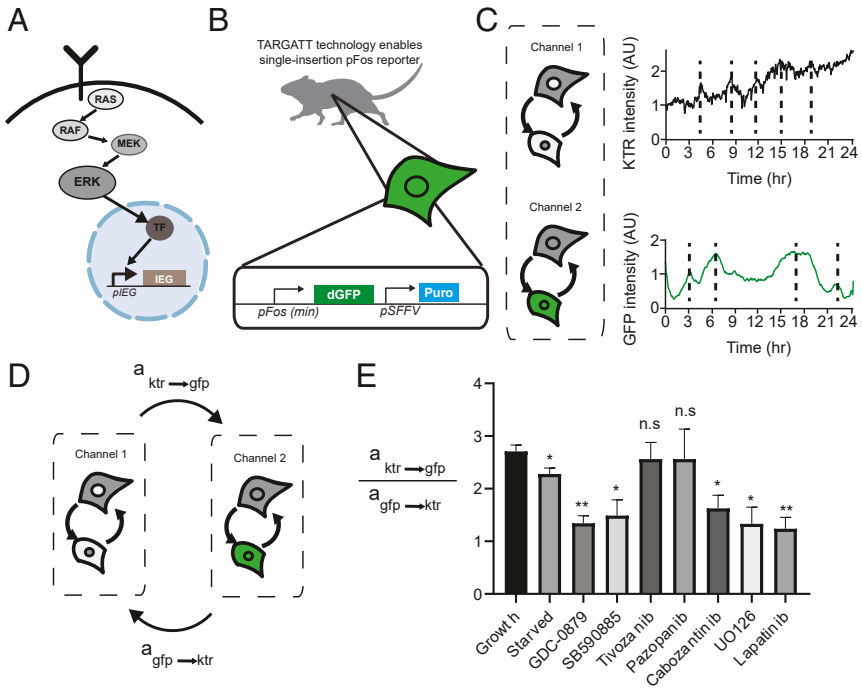

Fig. 5. (*A*) The Ras/Erk pathway results in the transcription of canonical immediate early genes. (*B*) TARGATT technology enables insertion of a Fos minimal promoter (pFos) into a single site in the genome, allowing for quantitative monitoring of pFos activation. (*C*) KTR and dGFP readout can be measured in the same cell. (*D*) CPP quantifies cross-channel directed signaling through $a_{ktr \to gfp}$ and $a_{gfp \to ktr}$. (*E*) Pharmacological inputs modulate the strength of the association between Ras/Erk signaling and Fos gene expression dynamics. Statistical significance was calculated using Student's *t* test, with $*P < 0.05$ and $**P < 0.005$.

analyzed using CPP. CPP allowed us to quantify the signaling between channels, Erk and Fos, with no prior information that Erk affects transcription of Fos. To measure this interchannel signaling, we estimated the intercellular signaling parameters $a_{ktr \to gfp}$ and $a_{gfp \to ktr}$ using the multivariate CPP model and took the ratio of the two values for each field of cells imaged (Fig. 5*D*). We were able to recapitulate the strong directed signaling of Erk to Fos transcription, as seen by the ratios of 3 and 2 for cells in growth and starved media, respectively (Fig. 5*E*). On the other hand, testing several drugs from a previously published keratinocyte drug screen (6) showed a range of signaling behaviors (Fig. 5*E*). Erk inhibition (UO126; Lapatinib) showed similar signaling in the presence of Erk-activating drugs GDC-0879 and SB590885, perhaps because the signaling between pulses of Erk signaling and pulses of GFP accumulation decreases when both are constantly on or constantly inhibited. On the other hand, Tivozanib and Pazopanib, which increase pulsing frequency μ (6), maintain a similar level of signaling and gene expression to those of cells grown in standard or starved conditions (Fig. 5*E*).

Discussion

In this paper, we present a spatiotemporal model, the CPP, to capture pulsatile cell-signaling events based on self-exciting Hawkes processes. Applying this model to processed cell-imaging data across time, we estimate model parameters that quantify the strength of spatial and autonomous signaling in a multicellular system, even in the context of heterogeneous cell types, multiple signaling channels, or environmental conditions. We use these parameters to quantitatively compare systems, e.g., pre- and postexposure to pathway-targeting drugs. We validate our model on simulated data and demonstrate its ability to capture known inhibition effects of TAPI-1 on spatial Erk coupling in keratinocytes. We then use the CPP model to interrogate the effects of different drug treatments on keratinocytes and are able to replicate the known effects on expression of three classes of drugs and extend knowledge of the effects of the drugs to cell signaling. The estimation of intercellular signaling parameters and autonomous pulsing parameters adds to our understanding of keratinocyte drug response. Finally, we use the CPP model to capture heterogeneous cell-signaling behaviors across distance from the wound frontier in response to wound healing.

The CPP model leaves room for further development. Point processes are known to be brittle models that have poor estimates under model misspecification (25). Under conditions where this kernel is inappropriate or the nature of spatial interaction is different, modifications and extensions would be necessary. Signaling mechanisms may have refractory periods, the time after which the likelihood of an event is depressed, leading to a different kernel with repressive properties. The kernel and the relationship between distance and signaling strength might be better modeled nonparametrically to account for differences in cell size, shape, and imaging scales. Currently, the model assumes that interactions are local and symmetric. However, systems such as wound healing may demonstrate global and directional behavior. The model of spatial interaction could be modified to vary over the region. Alternatively, the intensity of interaction could be a function of the vector distance between positions rather than a scalar distance. For longer histories of observations, we would also be interested in allowing these parameters to vary over time, e.g., to detect possible switch points between local and global signaling regimes.

The probabilistic aspects of the model could also be expanded to make the model fully Bayesian. Priors may be placed on the parameters to learn a maximum a posteriori estimate or the posterior distribution over parameters. Hierarchical models may be developed to estimate smoothly varying model parameters across the space instead of binning cells by distance to the wound. Hierarchical structure could be used to learn shared parameters

from multiple observations of the same condition such as repeats of the DMSO controls. Regardless, self-exciting point processes, and the CPP specifically, represent a powerful class of stochastic models that can be used to accurately and robustly quantify the spatial components of multicellular dynamics.

Materials and Methods

The CPP Model. Point processes are probabilistic models of events in some mathematical space, generally used to model event occurrences across space or time (or both) (19, 20, 26). A point process can be defined by the conditional intensity function, $\lambda(\cdot)$, which represents the expected infinitesimal rate at which events occur (27). We first describe a one-dimensional point process that starts at $t=0$. This point process produces N events over an interval of time $\delta t > 0$. At any moment $t > 0$, there is a history of previous events, H_t, that consists of the times, $\{\tau_1, \tau_2, \ldots, \tau_i < t\}$, of each previous event. The conditional intensity function is then defined as

$$\lambda(t) = \lim_{\delta t \to 0} \frac{\mathbb{E}[N(t, t+\delta t)|H_t]}{\delta t},$$

where δt represents a nonnegative interval of time.

A simple point process (28, 29), such as a Poisson process, may have a constant conditional intensity over time:

$$\lambda(t) = \mu.$$

In this simple Poisson process, the expected number of events is $\mu \Delta t$ over time Δt. However, biological processes are generally nonstationary—the expected number of events changes as a function of time. This nonstationary behavior comes from self-excitation, meaning an event makes future events more likely for a period. This phenomenon can be caused by a variety of mechanisms but is broadly referred to as positive feedback (30, 31). The conditional intensity function can be altered to represent such nonstationary behavior.

Hawkes processes (27, 32) model self-excitation with a conditional intensity dependent on history:

$$\lambda(t) = \mu(t) + \sum_{i:\tau_i<t} \nu(t-\tau_i),$$

where $\mu(t)$ represents an autonomous underlying rate of events, τ_i is the time of event $i \in H_t$, and ν is a kernel function defining the association between previous events and the conditional intensity.

In the multivariate case (13, 15, 33), where more than one set of events is observed simultaneously, the conditional intensity of one dimension λ_j is a function of the history in K dimensions:

$$\lambda_j(t) = \mu_j(t) + \sum_{k=1}^{K} \sum_{i:\tau_i<t} \nu_{jk}(t-\tau_i).$$

In this model, events in one variable may influence the conditional intensity function of other variables through the kernel function $\nu_{jk}(\cdot)$.

To model cells from a homogeneous population recorded on a two-dimensional plane, we treat each cell as a different variable in a multivariate point process. We assume that the autonomous component $\mu_j(t)$ is a positive constant $\mu > 0$ over time and across cells. Based on observations from prior work (6), the time kernel ν_{jk} is assumed to have a log-normal shape with zero mean and variance b_{jk}:

$$\nu_{jk}(\Delta t) = \frac{1}{b_{jk}\Delta t\sqrt{2\pi}} \exp\left(-\frac{(\ln \Delta t)^2}{2b_{jk}^2}\right).$$

We assume a constant $b_{jk} = b$ for all $j \neq k$ that represents intercell signaling and $b_{jk} = b_{self}$ for $j = k$ that represents intracell excitation. Intuitively, when a cell pulses, the conditional intensity of pulsing in its neighbors increases until peaking at time $\exp(-b^2)$, after which the conditional intensity decreases back toward the baseline μ. Biologically, this captures the expected delay between subsequent signaling events.

We make a_{jk} a function of the distance between cells to capture the spatial nature of cell–cell interactions:

$$a_{jk} = \begin{cases} a_{self} & d_{jk} = 0 \\ a & 0 < d_{jk} < \epsilon \\ 0 & d_{jk} > \epsilon, \end{cases}$$

where d_{jk} represents the distance between cells, ϵ is a radius inside which signaling is possible, a is a positive constant that quantifies the strength of cell–cell interactions, and a_{self} is a positive constant for the magnitude of intracell self-excitation.

For any individual cell j among K cells, the full CPP model is defined as

$$\lambda_j(t) = \mu + \sum_{k=1}^{K} \sum_{i:\tau_i<t} \frac{a_{jk}}{b_{jk}(t-\tau_i)\sqrt{2\pi}} \exp\left(-\frac{\left(\ln (t-\tau_i)\right)^2}{2b_{jk}^2}\right).$$

We can generalize the CPP model to account for multiple types of observations per cell, such as different fluorescence channels corresponding to different components of a protein-signaling network.

In this case, each channel and protein–protein interaction has a different set of model parameters. Given L channels capturing different proteins, and letting ℓ represent a particular channel, there is the following:

- $\mu_\ell > 0$ for each channel, representing L parameters;
- $a_{\ell,\ell'} > 0$ for each channel pair, representing L^2 parameters;
- $a_{self,\ell} > 0$ for each channel, representing L parameters;
- $b_{\ell,\ell'} > 0$ for each channel pair, representing L^2 parameters; and
- $b_{self,\ell} > 0$ for each channel, representing L parameters.

Thus, for any channel ℓ_j for cell j, the conditional intensity according to the CPP is

$$\lambda_{\ell_j}(t) = \mu_\ell + \sum_{k=1}^{K} \sum_{\ell}^{L} \sum_{i:\tau_i<t} \frac{a_{jk\ell}}{b_{jk\ell}(t-\tau_i)\sqrt{2\pi}} e^{\left(-\frac{(\ln (t-\tau_j))^2}{2b_{jk\ell}^2}\right)}$$

$$a_{jk\ell} = \begin{cases} a_{self,\ell} & d_{jk} = 0 \\ a_{\ell,\ell} & 0 < d_{jk} < \epsilon \\ 0 & d_{jk} > \epsilon. \end{cases}$$

$$b_{jk\ell} = \begin{cases} b_{self,\ell} & d_{jk} = 0 \\ b_{\ell,\ell} & d_{jk} > 0. \end{cases}$$

Relationship to Previous Models. A number of related models have been used to address biological patterning in the past. Here, we focus chiefly on the two-dimensional (2D) Ising lattice model and the Kuramoto oscillator system, as these are the closest in context to the model that we propose.

The 2D Ising model is used to describe a two-dimensional array of spins, each of which can be in one of two discrete states (spin up or spin down). This model has been implemented in creating discrete patterns in the study of reaction–diffusion systems with coupled agents (34). The operator function that gives the energy of the system as a function of the spin states σ of all of the constituent particles is called the Hamiltonian:

$$H(\sigma) = -\mu \sum_j h_j \sigma_j - \sum_{<i,j>} J_{i,j}\sigma_i\sigma_j,$$

where the first term $\mu \sum_j h_j \sigma_j$ is the effect of an external magnetic field h on a particle's spin state, weighted by μ, and the second term $-\sum_{<i,j>} J_{i,j}\sigma_i\sigma_j$ is the coupling term J between spins.

In our model, we can think of the coupling terms as similar to those presented in our model; i.e., cells interact with other cells within a small neighborhood around them. The magnetic-field term is usually held as constant over the system and is similar to our term μ that represents the probability of randomly pulsing in a cell-autonomous fashion. The Hamiltonian for the 2D Ising model is therefore similar in spirit to our model as described.

However, as many studies have shown positive feedback effects that give rise to self-excitation processes in cells, this model does not capture a reasonable range of cell-signaling behavior. Additionally, upon examining sheets of cells experiencing Ras/Erk pulses, we rarely see "stable states" of cells that are constantly on or off, in contrast to the steady-state behaviors often seen in Ising model simulations.

A second model used for coupled cells, and in particular cells capable of oscillations, is the Kuramoto model (35). The Kuramoto model treats cells as entities having an intrinsic oscillator "frequency" from a continuous range of values. Cells are also coupled to some number N of adjacent cells, with a "coupling constant" K:

$$\frac{d\theta_i}{dt} = \mu_i + \frac{K}{N} \sum_{j=1}^{N} \sin(\theta_j - \theta_i).$$

Here, the intrinsic frequency contribution μ_i describes the cell-intrinsic change in oscillator phase θ_i. This is akin to our μ parameter. In this function, the contribution of adjacent cell states to the state of cell i is distinct from the contribution of the autonomous behavior of cell i. Since the state of each cell is in the space of oscillator frequencies, this parameter is forced to oscillate through the sine of the difference in frequencies rather than parameterizing a random process, making the effects of neighboring cells periodic and deterministic instead of stochastic. This deterministic function precludes transient signaling events from occurring. The Kuramoto model, when simulated for long time courses, has been shown to relax into smooth patterns of phases with clear boundaries between regions of opposite phase (35); however, this is not the signaling pattern that we are trying to capture in pulsatile cells.

While our CPP model incorporates some of the elements of these two related models, it is more suited to capture the signaling patterns that we observe in experimental data. In the design of the CPP model, we make use of two terms that describe cell-autonomous and cell-to-cell signaling contributions. In all three cases, cell-to-cell signaling terms are applied over a small region around the cell in question, for example, in the eight-cell neighborhood of a square in a 2D lattice. The range of states that a cell can occupy differs among the three models. While the Ising model and our CPP model capture binary states (on or off), cells in the Kuramoto oscillator model may take values over a range of oscillator frequencies. These values, however, have the downside of being modeled deterministically, without allowing for the possibility of stochastic events.

The CPP model makes use of discrete states—more specifically, we model the bursting state of a cell as being on or off with respect to a specific protein. This point process approach does not allow for different amplitudes of signaling, but does take advantage of a framework that allows for phenomena such as positive feedback that create self-excitatory pulse trains of signaling pathway activation. Moreover, the structure of the signaling term in the CPP model allows us to take into account the history of a cell's signaling state, which is not a part of the Ising or Kuramoto models and allows for a more explicit treatment of self-excitatory processes in biological systems.

Both the Ising and Kuramoto models approach a "steady-state" pattern of phases or spins as time goes to infinity. However, active behavior and constant emergence of nonstationary fluctuations in a population limit the applicability of these models to data. A stochastic, self-exciting, and history-dependent model such as the CPP described here better captures these behaviors.

Inference for the CPP. The conditional intensity $\lambda_j(t)$ of the CPP has six parameters to be inferred: μ, a, a_{self}, b, b_{self}, and ϵ. Biologically, Erk signaling is regulated across cells by interactions between membrane proteins, limiting the signal to neighboring cells. Since the (x, y) spatial coordinates of each cell in our experimental data represent the cell center, we set ϵ to 60 pixels, corresponding to roughly 84 μm, unless otherwise stated. This value was obtained through visual inspection of the raw imaging data. We noticed that cells on average had five neighbors in direct contact with them, forming a local neighborhood of cells around each cell. We then chose ϵ such that the average cell had five neighbors. Of course, this parameter is calibrated to the cell distributions observed in our experiments; in future applications it will need to be adapted for images with different resolutions and cell geometries.

We optimize the remaining parameters by maximizing the log-likelihood of the observations. Given a history H_t with N events $\{\tau_1, \tau_2, \ldots, \tau_n\}$ in time interval $[0, T]$, the log likelihood is (12)

$$\mathcal{L} = \log \frac{\prod_{i=1}^{N} \lambda(\tau_i)}{\exp \int_0^T \lambda(t)dt} = \sum_{i=1}^{N} \log \lambda(\tau_i) - \int_0^T \lambda(t)dt.$$

We maximize this likelihood using automatic differentiation from PyTorch (36). We use the PyTorch stochastic gradient descent (SGD) optimizer (torch.optim.SGD) to minimize the negative log likelihood, with stopping criteria when a local minimum has been reached (the negative log likelihood at iteration i is greater than the negative log likelihood at iteration $i-1$) or when the absolute change in log likelihood between iterations is less than 0.001%. We take simultaneous gradient steps for all of the parameters with a learning rate equal to 1×10^{-4}, and we do not use momentum, dampening, or weight decay. We initialize all parameters with a value of 1.

Confidence Interval Estimation. It has been demonstrated for both temporal and spatiotemporal point processes that the covariance converges to the inverse of the expected Fisher information matrix as $T \to \infty$ (37–39). We use an existing estimator of the asymptotic covariance (38):

$$\hat{\Sigma} = \left(\sum_{i=1}^{n} \frac{\Delta(s_i, t_i)}{\lambda(s_i, t_i)} \right)^{-1}$$

$$\Delta_{kj}(s, t) = \frac{\dot{\lambda_k}(s, t)\dot{\lambda_j}(s, t)}{\lambda(s, t)},$$

where $i \in [1, N]$ indexes each peak, (s_i, t_i) is the likelihood of each peak at time t_i and location s_i, and $\dot{\lambda_k}$ is the partial derivative of λ_i with respect to k. Under the assumption of asymptotic normalcy, the 95% confidence interval for a parameter k is $1.96\sqrt{\hat{\Sigma}_{kk}}$. We implement an estimator of the Fisher information and confidence intervals in PyTorch using automatic differentiation to calculate λ_k for each parameter at each peak.

Simulations. We simulate histories of peaks in one channel from the generative model to test whether inference accurately estimates the true parameters. We first generate 100 sets of $\{\mu, a, a_{self}, b, b_{self}\}$. Each parameter is selected from a uniform distribution with set minimums and maximums (Table 3). The maximum distance of spatial interactions, ϵ, is kept constant at 60 pixels, corresponding to roughly 84 μm.

For each parameter set, for every combination for cells in $[50, 75, 100, 125, 150, 175, 200]$ and number of peaks in $[100, 250, 500, 1{,}000, 2{,}500]$, a history is simulated. Cell positions are drawn from a uniform distribution between (0, 0) and (x_{max}, y_{max}). To match data collected from microscopy, $x_{max} = y_{max}$ to create a square region. The maximum coordinates as well as ϵ are set such that the expected number of neighbors, $n_{cells} \times \frac{\epsilon^2}{x_{max}^2} = 5$, is similar to data from Goglia et al. (6). This corresponds to an intercellular distance of approximately 30 to 50 μm. For each history, we fit the model until convergence. We evaluate the goodness of fit by calculating the NMSE, $\text{NMSE}_q = \frac{1}{N} \sum_{n=1}^{N} \frac{(q_n - \hat{q_n})^2}{(q_{max} - q_{min})^2}$, where q represents some model parameter, between estimated and true parameter values.

We compare the NMSE of CPP parameter estimates to a simple control estimate for each parameter. The naive estimator of μ is the total number of peaks across cells, P, divided by the number of cells, C, times maximum time, T:

$$\hat{\mu} = \frac{P}{C \times T}.$$

The control a_{self} for a simulation is the average over each cell of the peaks in a cell (P_c) per unit time minus the true autonomous pulsing rate μ, intuitively the rate of excess peaks above μ in a cell:

$$\hat{a_{self}} = \frac{1}{C} \sum_{c=1}^{C} \frac{P_c}{T} - \mu.$$

The control a is similarly the average of excess peaks in each cell divided by the number of neighbors (n_c) a cell has:

$$\hat{a} = \frac{1}{C} \sum_{c=1}^{C} \left(\frac{P_c}{T} - \mu \right) / n_c.$$

We include only cells with neighbors for this calculation.

The control b_{self} is calculated from the average time between two sequential peaks from the same cell, across all pairs of same-cell sequential peaks in the simulation $(\bar{\Delta t})$. If $\bar{\Delta t} < 1$, we consider the average the mode of the log-normal distribution and estimate $b_{self} = \sqrt{-\log(\bar{\Delta t})}$. If $\bar{\Delta t} \geq 1$, we consider the mean of the log-normal distribution and estimate $b_{self} = \sqrt{2\log(\bar{\Delta t})}$. Similarly, the control b is calculated by taking the average time between every pair of peaks in cells that are considered neighbors and transformed to b by the same rules as for b_{self}.

Table 3. Parameter range for simulating a spatial point process

Parameter	Minimum	Maximum
μ	0.01	3
a	0.01	0.8
a_{self}	$0.01a$	$0.8a$
b	0.01	5
b_{self}	$0.01b$	$1.5b$

Experimental Methods.

Cell culture and generation of transgenic cell lines. Dorsal epidermal keratinocytes derived from CD1 mice and stably expressing a lentivirally delivered histone H2B-RFP and ErkKTR-BFP (6) were cultured as described previously (40). Briefly, keratinocytes were grown in complete low-calcium (50 mM) growth media ("E media" supplemented with 15% serum and 0.05 mM Ca^{2+}) in Nunclon flasks with filter caps (Thermo-Fisher) and were maintained in a humidified incubator at 37 °C with 5% CO_2. Cell passage number was kept below 30. Keratinocyte media was prepared as per prior work (40).

To create pFos-GFP–expressing cells, dorsal epidermal keratinocytes were derived from TARGATT mice containing a safe harbor locus with an attB insertion site (Applied Stem Cell). A vector containing the minimal Fos promoter driving a destabilized GFP and a CMV promoter driving a hygromycin resistance gene was constructed using infusion cloning and flanked with attP sites for insertion into the attB sites in TARGATT keratinocytes. Keratinocytes were cotransfected with this reporter plasmid as well as a plasmid encoding the phiC integrase driven by a CMV promoter, which, when expressed, completed the integration of the reporter construct into the safe harbor locus.

Cells were selected for expression with hygromycin (Sigma Aldrich). Prior to imaging experiments, cells were transduced with lentiviral vectors encoding a H2B-RFP marker, as well as with ErkKTR-iRFP.

Imaging experiments were performed in 96-well black-walled, 0.17-mm-high performance glass-bottom plates (Cellvis). For plating cells, wells were pretreated with a solution of 10 mg/mL bovine plasma fibronectin (Thermo Fisher) solubilized in phosphate-buffered saline (PBS) to support cell adherence. Two days before imaging, keratinocytes were seeded at approximately 96, 000 cells per well in 100 μL of low-calcium E media (in a 96-well plate). Glass-bottom plates were briefly centrifuged at 800 rpm to ensure even plating distribution, and cells were allowed to adhere overnight. Twenty-four hours before imaging, wells were washed two to three times with PBS to remove nonadherent cells and were shifted to high-calcium (1.5 mM $CaCl_2$) complete E media to promote epithelial monolayer formation. For experiments in growth factor-free (starvation) media, cells were washed once with PBS and shifted to high-calcium P media (Dulbecco's Modified Eagle Medium [DMEM] containing only pH buffer, penicillin/streptomycin, and 1.5 mM $CaCl_2$) 8 h before imaging. To prevent evaporation during time-lapse imaging, a 50-mL layer of mineral oil was added to the top of each well immediately before imaging.

Imaging was performed on a Nikon Eclipse Ti confocal microscope, with a Yokogawa CSU-X1 spinning disk; a Prior Proscan III motorized stage; an Agilent MLC 400B laser launch containing 405-, 488-, 561-, and 650-nm lasers; and a cooled iXon DU897 EMCCD camera, and fitted with an environmental chamber to ensure cells were kept at 37 °C and 5% CO_2 during imaging. All images were captured with a 20× air objective and were collected at intervals of 3 min. Each frame was associated with a specific time point, with accuracy to the thousandth of a minute.

For TAPI-1 experiments, drug was obtained from SelleckChem and diluted to 10× the relevant concentrations in DMSO. A total of 11 μL of drug was added to 100 μL of cells in 96-well plates immediately before imaging. For drug treatment experiments (Fig. 4), drugs were added to a final concentration of 2.5 μM.

For MDCK wound healing experiments, MDCK cells were maintained in minimal essential medium (MEM) (ThermoFisher Scientific; 10370-021) supplemented with 10% fetal bovine serum (FBS) (Sigma; 172012-500 ML), 1× Glutamax (ThermoFisher; 35050-061), and 1 mM sodium pyruvate (ThermoFisher; 11360070), in a 5% CO_2 humidified incubator at 37 °C. For time-lapse imaging, MDCK cells were plated on 35-mm glass-base dishes (Asahi Techno Glass). Before time-lapse imaging, the medium was replaced with FluoroBrite (Invitrogen) supplemented with 5% FBS and 1× Glutamax.

For the generation of MDCK cell lines stably expressing the Forster Resonance Energy Transfer (FRET) biosensor, a PiggyBac transposon system was used (5, 41). The pPBbsr-based FRET biosensor and pCMV-mPBase (neo-) encoding the piggyBac transposase were cotransfected into MDCK cells using an Amaxa nucleofector system (Lonza) at a ratio of 4:1. The cells were selected with 10 mg/mL of blasticidin S for at least 10 d. Single-cell clones expressing the biosensor were further isolated by limited dilution.

MDCK cells (4×10^5 cells) were plated on 35-mm glass-based dishes. Two days after seeding, confluent cells were scratched with a 200-μL pipette tip to establish the wound. Just after scratching, the media and dislodged cells were aspirated and replaced by FluoroBrite with 5% FBS and 1× Glutamax. Immediately after replacing the media, the cells were imaged with an epifluorescence wide-field microscope. The cells were imaged every 3 min for 12 h. DMSO, 100 nM Trametinib, or 10 nM TAPI-1 was added 2 h after starting time-lapse imaging.

TAPI Dose Response. A number of publications on the phenomenon of spatially coupled Ras/Erk pulses have noted that the matrix metalloprotease inhibitor TAPI-1 is capable of reducing the extent of cell-to-cell signaling in pulsatile activity (5, 21). The drug inhibits the cleavage and release of ligands that activate the Ras/Erk pathway in adjacent cells, a process called juxtacrine signaling (42). This body of work suggests that TAPI-1 specifically inhibits intercellular, but not intracellular, signaling pulses.

Previous work on TAPI-1 as an inhibitor of spatial signaling in pulsatile activity has been limited to analyses of approximately 5 to 10 cells at a time and focuses on isolated instances of cells losing spatial coupling upon TAPI-1 addition, rather than on a population-level response. We treated keratinocytes with a range of TAPI-1 doses, at 5, 10, and 20 μM, and imaged cells from the point of TAPI-1 exposure. An untreated well to which only the solvent DMSO had been added was imaged as a vehicle control, since TAPI-1 was solubilized in DMSO prior to addition to the well. DMSO has not been found to affect Ras/Erk activity dynamics (6). Imaged cells were incubated in growth factor-free media, and cells were imaged every 3 min for 12 h after the addition of TAPI-1. We noticed that cells went through a period of deactivation after the addition of the drug after which pulsing resumed; to remove this from our analysis, time series were truncated to the last 6 h of imaging. Time-series measurements were converted to a series of peaks for each cell as described previously (6). The model was fitted for each well separately until convergence.

Estimating the Effects of Different Drugs on Keratinocyte Signaling. We next fitted the model to data from prior work (6), in which keratinocytes were treated with various RTKis, which target proteins upstream of endogenous Erk activity. The data consist of 450 wells with 432 different drug treatments and 18 DMSO vehicle controls (which contain no inhibitor) (*SI Appendix*, Table 1). Imaged cells were incubated in growth factor-free media, and RTKi was added 30 min prior to imaging. Cells were imaged every 3 min for 12 h after addition of RTKi. Time-series measurements were converted into peaks for each cell as described previously (6). The model was fitted to data from each well separately until convergence. Due to differences in spatial organization across wells, the signaling radius ϵ was set independently for each well such that each cell had on average five neighbors.

MDCK Wound Healing. Extensive prior work has been done on the association between Ras/Erk pathway activity and cell proliferation and migration, events that are critical for regeneration and wound healing. In light of this, we demonstrated the use of live-cell Ras/Erk activity reporters in combination with our model to characterize the behavior of the Ras/Erk pathway in response to an acute wounding event. Since a wound has a particular spatial location relative to different cells, we used our model to quantify signaling rates at various distances from the wound. To do this, we collected data on a large sheet of MDCK cells, which are widely used for studies of collective cell motility. Cells expressing the EKAREV Erk activity reporter (43) were established using a piggyBac transposon system (41, 44) and sorted to ensure uniform expression of the reporter construct. For wound healing assay experiments, a wound was inflicted on cells by scratching a pipette across a confluent layer of cells, and the sheet of cells was imaged every 3 min for 12 h. Nuclei were segmented, and Erk activity was measured for each cell over time using the cell-tracking software TrackMate (6). Due to cell movement over the course of the experiment, the field of cells was split into 10 bins according to each cell's distance from the wound edge along the x axis at the start of the experiment, immediately after the wound was inflicted. Our model was fitted until convergence to each bin, consisting of the cells present in that spatial bin at the first time point, over the duration of the wound healing process. As a control, we also binned cells along the y axis, to ensure that these 10 bins result in identical estimated signaling behaviors since these bins run parallel to the wound. Data collected in the presence of the matrix metalloprotease inhibitor TAPI-1 and the MEK inhibitor Trametinib were also processed and analyzed in the same manner.

Analyzing Behavior in Multiple Channels. As described earlier, the CPP model makes it possible to examine couplings between separate channels, for example, to analyze separate components of a signaling network. The Ras/Erk pathway has a well-defined set of target genes, called IEGs, that respond acutely and rapidly to Ras/Erk stimulation. We engineered mouse keratinocytes to express a dGFP with a half-life of ∼1 h, under the control of the minimal promoter of the IEG Fos. Using these cells, we could measure

Erk-KTR as well as dGFP across 24 h in the same cells. Time-series measurements were converted into a series of peaks for each channel in each cell (6). CPP was fit run until convergence for each experiment to estimate the μ, a, a_{self}, b, and b_{self} terms for each channel and cross-channel interaction.

Data Availability. All code is publicly available at the GitHub repository (https://github.com/architverma1/CPP) and videos and data have been deposited in Dropbox (https://www.dropbox.com/sh/ctrb51chmkyfhlt/AABH1A1jBFrVSahljz7VSbWaa?dl=0).

ACKNOWLEDGMENTS. We would like to thank Alexander Goglia for kindly providing data and annotations from his experimental drug screen (6). S.G.J. was supported by NIH Ruth Kirschstein fellowship F31AR075398 and NIH Training Grant T32GM007388. A.V. and B.E.E. were supported by the NIH National Heart, Lung, and Blood Institute R01 HL133218 and NSF CAREER AWD1005627. J.E.T. was supported by NIH Grant DP2EB024247 and NSF CAREER Award 1750663. B.E.E. and J.E.T. gratefully acknowledge financial support from the Schmidt DataX Fund at Princeton University made possible through a major gift from the Schmidt Futures Foundation.

1. J. T. Bonner, The origins of multicellularity. *Integr. Biol. Issues News Rev.* **1**, 27–36 (1998).
2. L. G. Morelli, K. Uriu, S. Ares, A. C. Oates, Computational approaches to developmental patterning. *Science* **336**, 187–191 (2012).
3. C. S. Greene *et al.*, Understanding multicellular function and disease with human tissue-specific networks. *Nat. Genet.* **47**, 569–576 (2015).
4. T. Hiratsuka *et al.*, Intercellular propagation of extracellular signal-regulated kinase activation revealed by in vivo imaging of mouse skin. *Elife* **4**, e05178 (2015).
5. K. Aoki *et al.*, Stochastic Erk activation induced by noise and cell-to-cell propagation regulates cell density-dependent proliferation. *Mol. Cell* **52**, 529–540 (2013).
6. A. G. Goglia *et al.*, A live-cell screen for altered Erk dynamics reveals principles of proliferative control. *Cell Syst.* **10**, 240–253 (2020).
7. T. Hiratsuka, I. Bordeu, G. Pruessner, F. M. Watt, Regulation of Erk basal and pulsatile activity control proliferation and exit from the stem cell compartment in mammalian epidermis. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 17796–17807 (2020).
8. A. N. Kolmogorov, A study of the equation of diffusion with increase in the quantity of matter, and its application to a biological problem. *Moscow Univ. Bull. Math.* **1**, 1–25 (1937).
9. A. M. Turing, The chemical basis of morphogenesis. *Philos. Trans. R. Soc. Lond. Ser. B* **237**, 37–72 (1952).
10. D. G. Mallet, L. G. De Pillis, A cellular automata model of tumor–immune system interactions. *J. Theor. Biol.* **239**, 334–350 (2006).
11. J. Scholefield *et al.*, Super-resolution microscopy reveals a preformed NEMO lattice structure that is collapsed in incontinentia pigmenti. *Nat. Commun.* **7**, 12629 (2016).
12. K. Zhou, H. Zha, L. Song, "Learning social infectivity in sparse low-rank networks using multi-dimensional Hawkes processes" in *Artificial Intelligence and Statistics* (2013), pp. 641–649.
13. H. Mei, J. M. Eisner, "The neural Hawkes process: A neurally self-modulating multivariate point process" in *Advances in Neural Information Processing Systems* (2017), pp. 6754–6764.
14. E. Bacry, I. Mastromatteo, J.-F. Muzy, Hawkes processes in finance. *Market Microstructure and Liquidity* **1**, 1550005 (2015).
15. P. Embrechts, T. Liniger, L. Lin, Multivariate Hawkes processes: An application to financial data. *J. Appl. Probab.* **48**, 367–378 (2011).
16. F. Gerhard, M. Deger, W. Truccolo, On the stability and dynamics of stochastic spiking neuron models: Nonlinear Hawkes process and point process GLMs. *PLoS Comput. Biol.* **13**, e1005390 (2017).
17. P. Reynaud-Bouret *et al.*, Adaptive estimation for Hawkes processes; application to genome analysis. *Ann. Stat.* **38**, 2781–2822 (2010).
18. L. Carstensen, A. Sandelin, O. Winther, N. R. Hansen, Multivariate Hawkes process models of the occurrence of regulatory elements. *BMC Bioinf.* **11**, 456 (2010).
19. P. J. Laub, T. Taimre, P. K. Pollett, Hawkes processes. arXiv:1507.02822 (10 July 2015).
20. P. Diggle, *Statistical Analysis of Spatial and Spatio-Temporal Point Patterns* (Chapman & Hall, 2013).
21. K. Aoki *et al.*, Propagating wave of Erk activation orients collective cell migration. *Dev. Cell* **43**, 305–317 (2017).
22. W. G. Stetler-Stevenson, Tissue inhibitors of metalloproteinases in cell signaling: Metalloproteinase-independent biological activities. *Sci. Signal.* **1**, re6 (2008).
23. S. Regot, J. J. Hughey, B. T. Bajar, S. Carrasco, M. W. Covert, High-sensitivity measurements of multiple kinase activities in live single cells. *Cell* **157**, 1724–1734 (2014).
24. A. Fauster *et al.* A cellular screen identifies ponatinib and pazopanib as inhibitors of necroptosis. *Cell Death Dis.* **6**, e1767 (2015).
25. A. Reinhart, J. Greenhouse, Self-exciting point processes with spatial covariates: Modelling the dynamics of crime. *J. R. Stat. Soc. Ser. C Appl. Stat.* **67**, 1305–1329 (2018).
26. L. Von Bortkiewicz, *Das Gesetz Der Kleinen Zahlen* (BG Teubner, 1898).
27. F. Paik Schoenberg, *Introduction to Point Processes* (Wiley Encyclopedia of Operations Research and Management Science, 2010).
28. S. D. Poisson, *Recherches sur La Probabilité des Jugements en Matière Criminelle et en Matière Civile* (Bachelier, 1837).
29. A. K. Erlang, Sandsynlighedsregning og telefonsamtaler [probability calculation and telephone conversations]. *Nyt Tidsskrift Matematik* **20**, 33–39 (1909).
30. S. Boykevisch *et al.*, Regulation of Ras signaling dynamics by SOS-mediated positive feedback. *Curr. Biol.* **16**, 2173–2179 (2006).
31. G. Hornung, N. Barkai, Noise propagation and signaling sensitivity in biological networks: A role for positive feedback. *PLoS Comput. Biol.* **4**, e8 (2008).
32. A. G. Hawkes, Spectra of some self-exciting and mutually exciting point processes. *Biometrika* **58**, 83–90 (1971).
33. Y. Yang, J. Etesami, N. He, N. Kiyavash, "Online learning for multivariate Hawkes processes" in *Advances in Neural Information Processing Systems* (2017), pp. 4937–4946.
34. M. Merle, L. Messio, J. Mozziconacci, Turing-like patterns in an asymmetric dynamic Ising model. *Phys. Rev. E* **100**, 042111 (2019).
35. M. Breakspear, S. Heitmann, A. Daffertshofer, Generative models of cortical oscillations: Neurobiological implications of the Kuramoto model. *Front. Hum. Neurosci.* **4**, 190 (2010).
36. A. Paszke *et al.*, "PyTorch: An imperative style, high-performance deep learning library" in *Advances in Neural Information Processing Systems*, H. Wallach *et al.*, Eds. (Chapman & Hall, 2019), vol. 32, pp. 8024–8035.
37. Y. Ogata, The asymptotic behaviour of maximum likelihood estimators for stationary point processes. *Ann. Inst. Stat. Math.* **30**, 243–261 (1978).
38. S. L. Rathbun, Asymptotic properties of the maximum likelihood estimator for spatio-temporal point processes. *J. Stat. Plann. Inference* **51**, 55–74 (1996).
39. A. Reinhart *et al.*, Rejoinder: A review of self-exciting spatio-temporal point processes and their applications. *Stat. Sci.* **33**, 330–333 (2018).
40. J. A. Nowak, E. Fuchs, Isolation and culture of epithelial stem cells. *Methods Mol Biol.* **482**, 215–232 (2009).
41. K. Yusa, L. Zhou, M. A. Li, A. Bradley, N. L. Craig, A hyperactive piggybac transposase for mammalian applications. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 1531–1536 (2011).
42. A. B. Singh, R. C. Harris, Autocrine, paracrine and juxtacrine signaling by EGFR ligands. *Cell. Signal.* **17**, 1183–1193 (2005).
43. N. Komatsu *et al.*, Development of an optimized backbone of FRET biosensors for kinases and GTPases. *Mol. Biol. Cell* **22**, 4647–4656 (2011).
44. X. Li *et al.*, piggyBac transposase tools for genome engineering. *Proc. Natl. Acad. Sci. U.S.A.* **110**, E2279–E2287 (2013).