PATTERN FORMATION AND DYNAMICS IN BACTERIAL CELLS

SRIKANTH SUBRAMANIAN

Marburg 2022

PATTERN FORMATION AND DYNAMICS IN BACTERIAL CELLS

SRIKANTH SUBRAMANIAN

Dissertation

zur Erlangung des Doktorgrades
der Naturwissenschaften
(Dr. rer. nat)

dem Fachbereich Physik
der Philipps-Universität Marburg

vorgelegt von
Srikanth Subramanian
geboren in Kumbakonam, India

Marburg, 2022

angenommen am: 25.01.2023
Tag der mündlichen Prüfung: 13.02.2023
Erstgutachters: Dr. Seán Murray
Zweitgutachters: Prof. Dr. Peter Lenz

ACKNOWLEDGEMENTS

This concludes the challenging but well-run journey of my PhD. I am extremely thankful to all the friends I made along the way. This has been a truly fun, singular experience and will be a significant addition to my life.

Firstly, I thank Seán for being a fantastic collaborator and advisor throughout my PhD. I am thankful that he bravely trusted (tolerated) my decision to pursue experiments in the final year. I thank my group members Ismath, Robin, Lara, Hanno and Seán for creating an agile and fun work environment. I am indebted to Ismath for an introduction to the wet lab and all the hand-holding through the protocols. Additional mentions go to Nicola, Giovanni and Robin. Special thanks go to my Bouldering cult for keeping the competition alive. I thank my friends in the extended Sourjik group family, who were a constant source of support during my time in Marburg. The ritual of workday lunches and weekend dinners with Giovanni, Silvia, Elif, Giacomo, Silvia Gonzalez and others kept me steady through stressful times. Special thanks also go out to the MPI alumni Jannik, Maria, Olli, Joana, Shankar and Delia for all the screaming, parties and celebrations.

I thank Giacomo and Lara for proofreading parts of this thesis and Robin for translations to German. I thank the thesis committee members Victor and Peter for all their feedback over the years. I am also thankful to the Max Planck Institute for Terrestrial Microbiology for offering a welcoming and collaborative atmosphere to an outsider (a physicist). Finally, I am grateful to the city of Marburg, for being a fantastic background character in the last four years.

Moving to a new country and culture is a clear challenge for anyone. I would be remiss if I did not thank the people, culture, restaurants, rivers and gardens in the fantastic city of Munich for welcoming me. I would also like to thank my friends and colleagues from Munich, who made me feel at home in Germany. Thanks to Alex, Ray and Alissa, my time in Munich was a breeze. I thank Kapil and Anchal for their hospitality. I thank my master's supervisors Chase and Karsten, my colleagues Bene, Mareike, Christian and the Statistical Physics group for creating a great working environment that nurtured my interests in Biophysics. I thank my parents and brother for being immensely supportive of my process and choices.

Lastly, I thank Ritu for being strong force in my life guiding me through all important decisions in the last 5 years including the decision to pursue a PhD. While we lived in different cities it never felt like we were apart. Her curiosity, care and love kept me going through difficult times. Our trips to across Europe from the Greek cat islands to the diners

in Strasbourg kept me going. Most importantly, my stay in Kolkata, the food and the stories will be cherished forever. I thank Ritu for tolerating me over these years and for being the kindest person I know.

SUMMARY

Spatio-temporal organisation plays a critical role in all life. More specifically in biological cells, the spatial organisation of key proteins and the chromosome is essential for their function, segregation and faithful inheritance. Within bacterial cells pattern formation appears to play an essential role at different levels. Examples of pattern formation in proteins include pole-to-pole oscillations, self-positioning clusters and protein gradients. Chromosomes on the other hand display an ordered structure with individual domains exhibiting specific spatio-temporal organisation. This work examines the processes determining dynamics and organisation within bacterial cells by combining analytical, computational and experimental approaches. The thesis is split into two distinct parts, one providing new physical insights into pattern formation in general and the other detailing the dynamics of chromosomes.

Reaction-diffusion systems are helpful models in order to study pattern formation in chemical, physical and biological systems. A pattern or Turing state emerges when the spatially homogeneous state becomes unstable to small perturbations. While initially intended for describing pattern formation in biological systems (for example embryogenesis, scale patterning etc.), their practical application has been notoriously difficult. The biggest challenge is our inability to predict in general the steady-state patterns obtained from a given set of parameters. While much is known near the onset (when the system is marginally unstable) of the spatial instability, the mechanisms underlying pattern selection and dynamics away from the onset are much less understood. In the first part of this thesis, we provide physical insight into the dynamics of these patterns and their selection at steady state. We find that peaks in a Turing pattern behave as point sinks, the dynamics of which are determined by the diffusive fluxes into them. As a result, peaks move toward a periodic steady-state configuration that minimizes the mass of the diffusive species. Importantly, we also show that the preferred number of peaks at the final steady state is such that this mass is minimized. Our work presents mass minimization as a general principle for understanding pattern formation in reaction-diffusion systems.

In the second part, we discuss a more biological problem that involves the study of bacterial DNA loci dynamics at short time scales, where we perform polymer simulations, modelling and fluorescent tracking experiments in conjunction. Chromosomal loci in bacterial cells show a robust sub-diffusive scaling of the mean square displacement, $\mathrm{MSD}(\tau) \sim \tau^{\alpha}$, with $\alpha < 0.5$. This is in contrast to scaling predictions from simple polymer models ($\alpha \geq 0.5$). While the motion of the chromosome in a viscoelastic cytoplasm has been proposed as a possible explanation for the difference, recent experiments in compressed cells question

this hypothesis. On the other hand, recent experiments have shown that DNA-bridging Nucleoid Associated Proteins (NAPs) play an important role in chromosome organisation and compaction. Here, using polymer simulations we investigate the role of DNA bridging in determining the dynamics of chromosomal loci. We find that bridging compacts the polymer and reproduces the sub-diffusive elastic dynamics of monomers at timescales shorter than the bridge lifetime. Consistent with this prediction, we measure a higher exponent in a NAP mutant (ΔH-NS) compared to wild-type *E. coli*. Furthermore, bridging can reproduce the rare but ubiquitous rapid movements of chromosomal loci that have been observed in experiments. In our model, the scaling exponent defines a relationship between the abundance of bridges and their lifetime. Using this and the observed mobility of chromosomal loci, we predict a lower bound on the average bridge lifetime of around 5 seconds. We hope that this framework will help guide future model development and understanding of chromosome dynamics.

ZUSSAMENFASUNG

Die Spatio-temporal Organstation ist ein wichtiger Bestandteil alles Lebens. Besonders in biologischen Zellen ist die Organisation essentieller Komponenten wie zum Beispiel von Proteinen und Chromosomen ein wesentlicher Bestandteil jedes lebensfähigen und autonom teilenden Organismus. In Bakterien geschieht dies auf verschiedenen Levels. Beispiele hierfür sind Proteine, die kollektiv von einem Pol der Zelle zum anderen oszillieren oder selbstorganisierende Protein Cluster/Gradienten. Namentlich zeigen Chromosome eine höchst organisierte Struktur, in der verschiedene Domains verschiedene spatio-temporal Charakteristiken besitzen. Diese Dissertation befasst sich mit Prozessen der Organisation in bakteriellen Zellen mit einer Kombination aus analytischer Mathematik, Computer-Simulationen und Experimenten. Diese Dissertation besteht aus zwei Teilen: (1) Neuartige Erkenntnisse im Bereich der Musterformation, (2) Experimentelle Beeinflussung der Dynamik des Chromosoms.

Der erste Teil beschreibt neue Erkenntnisse im Bereich der Musterbildung. Reaktions-Diffusions-Modelle, auch bekannt unter dem Namen Turing-Mechanismus, sind hilfreiche Werkzeuge, um die Formation von Mustern in chemischen, physikalischen oder biologischen Systemen zu erforschen. Muster oder strukturierte Zustände entstehen dann, wenn der räumlich-homogene Zustand eines Systems instabil und anfällig für Störungen wird. In der Vergangenheit wurden diese Modelle genutzt, um biologische Systeme wie zum Beispiel die Embryogenese oder die Formation von Schuppenmuster zu beschreiben, ihre Anwendung blieb jedoch notorisch schwer. Eine der größten Herausforderungen ist es unter verschiedenen Ausgangsbedingungen Stabilitätsmuster vorherzusagen. Über ihre Formation, kurz bevor ein System instabil wird, ist vieles bekannt, doch unser Wissen über die Musterfindung ist stark begrenzt, wenn das System nur geringfügig instabil wird. In diesem Teil der Dissertation geben wir Einblicke in die Formation von diesen Stabilitätsmustern und darüber, wie man sie vorherbestimmen kann. Wir haben herausgefunden, dass Spitzen in Tuning-Mustern sich wie Point-Sinks verhalten, deren Dynamik durch den Strom der diffundierenden Spezies in sie beschrieben werden kann. Dadurch positionieren sich die Spitzen der Stabilitätsmuster an periodischen Positionen, sodass die Masse der diffundierenden Spezies minimiert wird. Darüber hinaus zeigen wir, dass die bevorzugte Anzahl an Spitzen im finalen Stabilitätsmuster diese Masse minimiert. Diese Dissertation präsentiert die Minimierung von Masse als allgemeines Prinzip, um Musterformation in Reaktions-Diffusions-Systemen zu verstehen.

Im zweiten Teil diskutieren wir ein mehr biologisches Problem, welches die Studie der Bewegung von Genlokusse in kurzen Zeitintervallen thematisiert. Wir nutzen dafür Polymer-Simulationen in Kombination mit fluoreszenten Tracking-Experimenten. In bakteriellen Zellen lässt sich Subdiffusion bei der Bewegung von Genlokusse beobachten. Die mittlere quadratische Verschiebung $\mathrm{MSD}(\tau) \sim \tau^{\alpha}$ in lebenden Bakterien wächst langsamer, mit einem von $\alpha < 0.5$, als die Vorhersage von Polymersimulationen (mit einem $\alpha \geq 0.5$). Eine mögliche Erklärung bieten brücken-bauende und nucleoid-assoziierte Proteine (NAP), welche eine wichtige Rolle in der Organisation des Chromosoms spielen. In dieser Arbeit studieren wir, mit Hilfe von Polymersimulationen, welche Rolle die NAP in der subdiffusiven Natur der Genlokusse spielen. Wir haben herausgefunden, dass brücken-bauende NAP den Raum verkleinern, den das Chromosome einnimmt und somit die mittlere quadratische Abweichung reproduzieren, welche wir experimentell beobachtet haben. Wir zeigen, dass ein NAP-Mutant ein höheres α besitzt, was mit unseren Simulationen übereinstimmt. Darüber hinaus kann die Brückenbildung seltene rapide Bewegungen der Genlokusse reproduzieren, die experimentell beobachtet wurden. Mit Hilfe des Parameters, welcher das Verhältnis zwischen der Anzahl von Brücken und ihrer Lebensdauer, in unseren Model, definiert, sind wir in der Lage eine untere Grenze der durchschnittlichen Brückenlebensdauer von 5 Sekunden vorherzusagen.

CONTENTS

Part I

PATTERN SELECTION IN REACTION-DIFFUSION SYSTEMS

1

INTRODUCTION TO PATTERN FORMATION

> "... the totality is not, as it were, a mere heap, but the whole is something besides the parts."
>
> — Aristotle

While the question of the existence of the universe and matter is of interest to philosophers and physicists alike, a more tangible ancillary question of structure is also interesting and fundamental. In other words: why does matter have an interesting structure? [1]. It is not clear yet how fundamental equations that determine the behaviour of matter generate the higher-order complex structures that we observe. This is sometimes referred to as an emergent phenomenon where the properties of a system are not a simple combination of the properties of its components i.e the whole is more than the sum of its parts. These phenomena are observed at length and time scales of the universe down to the scales of humans and microorganisms. At the astronomical scale, observations show that galaxies do not fill up space uniformly like freely diffusing gas molecules but instead are clustered in sheets and lines and leave large voids in between them. This pattern though not a regular structure is a deviation or a statistical anomaly from the naive expectation of randomly distributed points in space. Similarly, on a smaller scale, we observe interesting yet complicated dynamics of weather, oceanic currents and tectonic structures on earth. Furthermore, the diverse forms of self-organised matter and life on earth are also rather surprising. While the exact explanation for this structure is unclear, evidence points to the fact that we inhabit a non-equilibrium universe.

While it is important to explore the diverse landscape of emergent phenomena, the above-described crude description is unsatisfactory. The appearance of such emergent structures becomes apparent on different levels of description or coarse-graining. In this work, we discuss a specific example of emergent phenomena of pattern formation in reaction-diffusion systems. Pattern formation occurs in a huge variety of natural and living systems, from chemical reactions involving cells to environmental patterns. In systems described by reaction-diffusion (RD) equations, the formation of spatially periodic patterns can be explained by the Turing instability, in which patterns emerge due to the presence of two or more interacting components that diffuse (or are transported) at different rates. In his

seminal work "The chemical basis of morphogenesis" in 1952, A.M.Turing proposed a theory based on the observation that, while diffusion itself has a homogenizing effect, coupled with chemical reactions it could lead to the formation of patterns [2]. He speculated that the patterns generated by these systems might suffice to explain biological morphogenesis i.e the emergence of a structure during the growth of an organism. His examples included stripes or spots on animal skin including zebra, tiger, cheetah, giraffe, and fish and the question of morphogenesis of a spherically symmetric zygote into different cells. This proposed mechanism of pattern formation was eventually confirmed through experiments in chemical systems [3,4], while biological systems have been more difficult to understand than originally envisioned.

In the subsequent chapters, we tackle a fundamental question in pattern formation, that of pattern selection. Application of reaction-diffusion or Turing models in biology has been hindered by our inability to predict in general the final pattern obtained at a steady state. Since the equations are highly non-linear, analytical solutions are in general unlikely. While much is known near the onset of spatial instability, the mechanisms underlying pattern selection and dynamics away from the onset are much less understood. Also, most model systems are in general multi-stable, in that more than one pattern is likely at steady state for the same parameters. The mechanism that drives the dynamics and selects the final pattern (the most likely pattern) remains elusive. Hence, a physical understanding of these systems is of critical importance.

We find that peaks in a Turing pattern behave as point sinks, the dynamics of which are determined by the diffusive fluxes into them. As a result, peaks move toward a periodic steady-state configuration that minimizes the mass of the diffusive species. Interestingly, the simpler point sink model was developed initially to model the positioning of plasmids in bacteria. We also show that the preferred number of peaks at the final steady state is such that the mass of the fast-diffusing species is minimized. Our work presents mass minimization as a potential generalisable principle for understanding pattern formation in reaction-diffusion systems far from the onset. This will likely help guide future model creation and curation to describe biological processes.

Publication: Subramanian, S., Murray, S. M. (2021). Pattern selection in reaction diffusion systems. Physical Review E, 103(1), 1–12. https://doi.org/10.1103/PhysRevE.103.012215

1.1 REACTION-DIFFUSION SYSTEMS

In this section, we wish to understand Spatio-temporal pattern formation using a simple mathematical framework of reaction-diffusion models as described by Turing. In equilibrium systems, if external conditions are kept constant, we expect them to relax to a spatially uniform homogeneous state. However, we can tune some parameters or external cues in

small steps until a point when suddenly a non-uniform spatial structure appears. In many examples of pattern formation, the structured state emerges when the spatially homogeneous state becomes unstable to small perturbations.

While multi-component reaction-diffusion systems are also of interest, for much of this work we focus on the simpler case of two-component reaction-diffusion systems. The general equations describing two-component reaction-diffusion systems are given by,

$$\begin{aligned} u_t &= \Gamma f(u,v) + d\nabla^2 u, \\ v_t &= \Gamma g(u,v) + \nabla^2 u, \end{aligned} \tag{1.1}$$

where u and v are two interacting species on a spatially extended system. The relative diffusion co-efficient $d = \frac{D_u}{D_v}$, where $D_u D_v$ are the diffusion constants for each species.

The boundary conditions and initial conditions are given by,

$$(\mathbf{n}.\nabla) \begin{bmatrix} u \\ v \end{bmatrix} = 0; \mathbf{r} \text{ on } \partial B; \text{ given } u(\mathbf{r},0), v(\mathbf{r},0), \tag{1.2}$$

where ∂B is the closed boundary of the RD domain B and $\mathbf{n}$ is the unit outward normal to ∂B. We choose the zero flux condition or Neumann boundary condition as we are interested in spatial patterns with no external input.

Linear stability analysis

Linear stability analysis is a valuable mathematical tool that helps us understand how a non-equilibrium system depends on its parameters. It also enables us to make easily testable predictions and classify these systems based on emergent structures. We first study the stability conditions in the absence of diffusion.

$$\begin{aligned} \frac{du}{dt} &= \Gamma f(u,v), \\ \frac{dv}{dt} &= \Gamma g(u,v). \end{aligned} \tag{1.3}$$

The homogeneous steady state (u_0, v_0) is the solution of,

$$f(u,v) = 0, g(u,v) = 0. \tag{1.4}$$

To study the stability of the fixed points, we perform small perturbations of the homogeneous steady state, $u = u_0 + \epsilon_1$, $v = v_0 + \epsilon_2$ where $\epsilon_1, \epsilon_2 << 1$. The growth of these perturbations that determine the stability of the fixed points is given by

$$\frac{d\boldsymbol{\epsilon}}{dt} = \Gamma \begin{bmatrix} f_u & f_v \\ g_u & g_v \end{bmatrix}_{(u_0,v_0)} ; \boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \end{bmatrix}. \tag{1.5}$$

where $J = \begin{bmatrix} f_u & f_v \\ g_u & g_v \end{bmatrix}_{(u_0,v_0)}$ is the Jacobian evaluated at the fixed point (u_0, v_0). The solution to the above first-order Ordinary Differential Equation (ODE) is given by

$$\boldsymbol{\epsilon} = \mathbf{a}e^{\lambda_+ t} + \mathbf{b}e^{\lambda_- t}. \tag{1.6}$$

λ_+, λ_- are distinct eigenvalues of the Jacobian J. To ensure stability, it is *necessary* that the perturbations do not grow in time. Thus, we have $\mathrm{Re}(\lambda_+), \mathrm{Re}(\lambda_-) < 0$. The eigenvalues of a 2×2 matrix can be explicitly written in terms of its trace (TrJ) and determinant ($DetJ$). The eigenvalue equation is given by,

$$| \Gamma J - \lambda I | = 0, \tag{1.7}$$

and the eigenvalues are,

$$\lambda_\pm = \frac{\Gamma}{2}(TrJ \pm \sqrt{TrJ^2 - 4DetJ}). \tag{1.8}$$

The conditions for stability of the non-linear ODEs then become,

$$TrJ < 0 \implies f_u + g_v < 0. \tag{1.9}$$

$$DetJ > 0 \implies f_u g_v - f_v g_u > 0. \tag{1.10}$$

In the *presence* of diffusion, we study the stability conditions with

$$\frac{\partial \boldsymbol{\epsilon}}{\partial t} = \left(D\nabla^2 + \Gamma J\right)\boldsymbol{\epsilon}, \tag{1.11}$$

where $D = \begin{bmatrix} d & 0 \\ 0 & 1 \end{bmatrix}$ and $J = \begin{bmatrix} f_u & f_v \\ g_u & g_v \end{bmatrix}$. The time-independent solution of the spatial eigenvalue problem is defined by,

$$\nabla^2 \boldsymbol{\epsilon} + \lambda \boldsymbol{\epsilon} = 0, (n.\nabla)\boldsymbol{\epsilon} = 0 \text{ for } \mathbf{r} \text{ on } \partial B, \tag{1.12}$$

where the eigenvalue is denoted by λ. The eigenfunctions of the Laplacian, for any boundary condition on a general, bounded domain $B \subset R^n$, form a countable orthonormal basis for

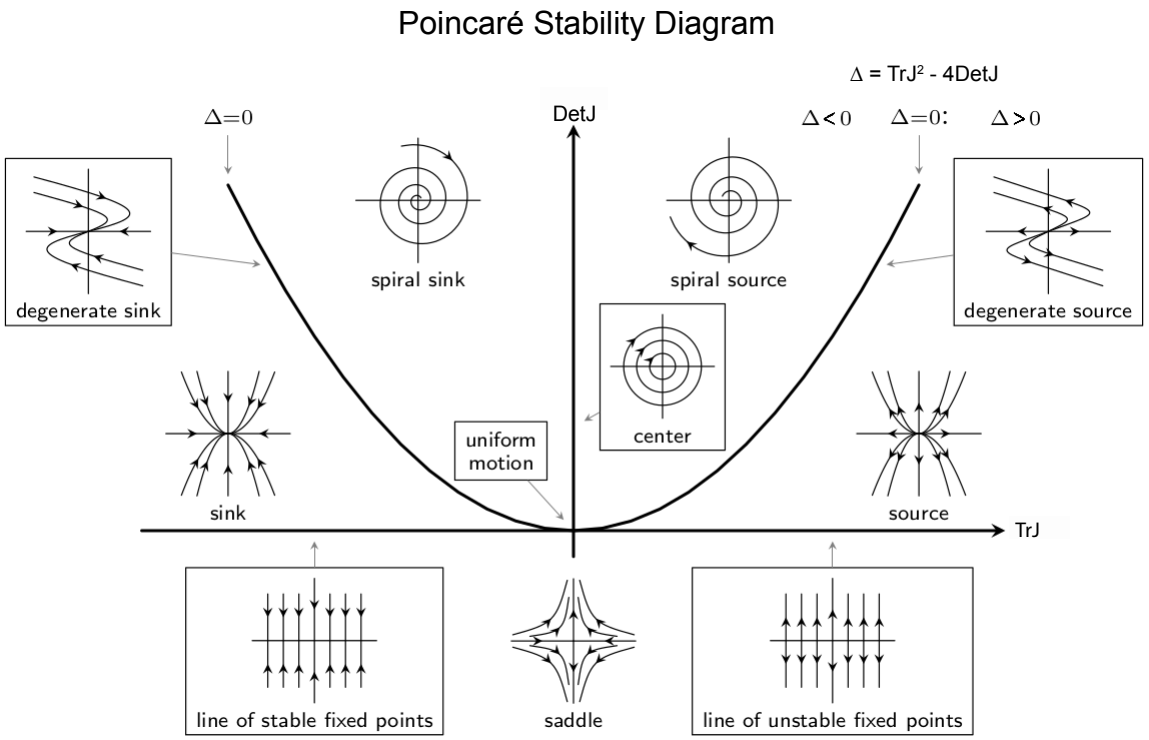

Figure 1: Poincaré diagram for classification of fixed points using Linear Stability Analysis.

the Hilbert space $L^2(B)$. Thus we can express the solution as a linear combination of these eigenfunctions. For example, on a finite 1-D domain say $x \in [0, L]$, and $\lambda > 0$ the solution that satisfies the no flux boundary conditions is, $\epsilon \propto \cos(n\pi x/L)$, where n is an integer. The eigenvalue $k = \frac{n\pi}{L}$ (the wavenumber), can take a set of discrete values on a finite domain. Note that on an infinite domain n would change continuously.

Thus we can express the solution of Eq. (1.11) as a linear combination of these eigenfunctions,

$$\boldsymbol{\epsilon}(\mathbf{r}, t) = \sum_k c_k e^{\lambda t} \boldsymbol{\epsilon}_k(\mathbf{r}), \tag{1.13}$$

where the constants c_k are determined by a Fourier expansion of the initial conditions in terms of $\boldsymbol{\epsilon}_k(r)$. λ is the eigenvalue that determines the temporal growth. Substituting this expression in Eq. (1.11) we get,

$$\begin{aligned} \lambda \boldsymbol{\epsilon}_k &= \Gamma J \boldsymbol{\epsilon}_k + D\nabla^2 \boldsymbol{\epsilon}_k \\ &= \Gamma J \boldsymbol{\epsilon}_k - Dk^2 \boldsymbol{\epsilon}_k \end{aligned} \tag{1.14}$$

The characteristic eigenvalue equation in λ is given by,

$$| \lambda I - \Gamma J + Dk^2 | = 0. \tag{1.15}$$

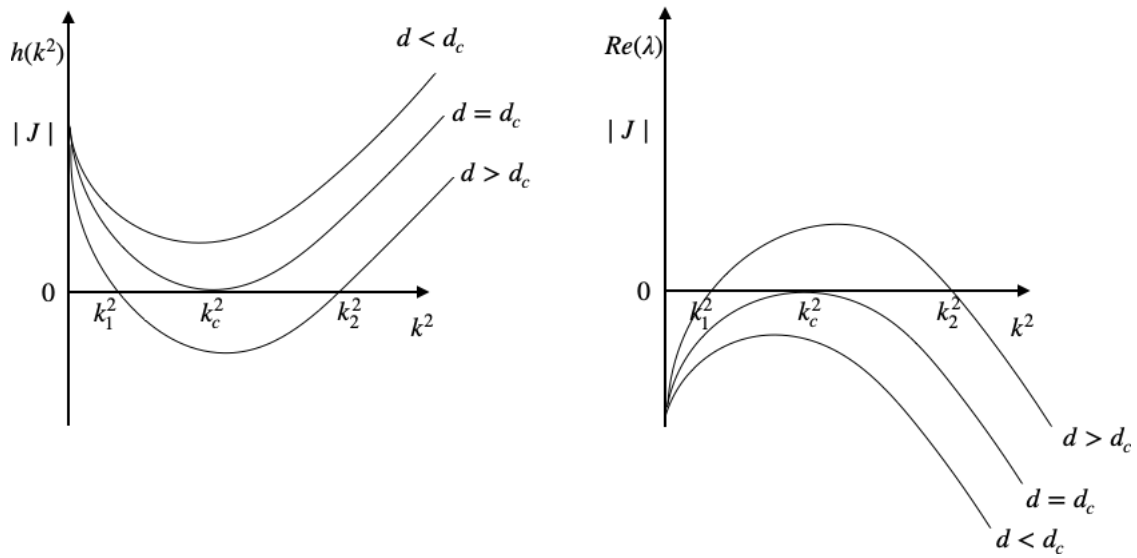

Figure 2: Dispersion relation. **Left**.Plot of $h(k^2)$ defined by eq Eq. (1.16). When d increases beyond the critical value d_c, $h(k^2) < 0$ for a finite $k^2 > 0$, a critical mode k_c is unstable to spatial perturbations. **Right**. The plot of the largest eigenvalue of $\lambda(k^2)$ as a function of k^2. When $d > d_c$ there is a range of k such that $k_1^2 < k^2 < k_2^2$ which are linearly unstable as given by Eq. (1.25).

$$\begin{aligned} \lambda^2 + \lambda\left[k^2(d+1) - \Gamma(f_u + g_v)\right] + h(k^2) = 0 \\ h(k^2) = dk^4 - \Gamma(f_u + dg_v)k^2 + \Gamma^2 \mid J \mid \end{aligned} \tag{1.16}$$

The steady-state (u_0, v_0) is linearly stable if both solutions of Eq. (1.16) have $\mathrm{Re}\lambda(k^2 = 0) < 0$ (where Re represents the real part of the solution). For the steady state to be unstable to *spatial* disturbances we require $\mathrm{Re}\lambda(k) > 0$ for some $k \neq 0$. Since $TrJ < 0$ and $k^2(1+d) > 0$ for all $k \neq 0$ the coefficient of λ, namely,

$$[k^2(1+d) - \Gamma(f_u + g_v)] > 0 \tag{1.17}$$

for $\mathrm{Re}\lambda(k^2)$ can be positive is if $h(k^2) < 0$ for some k. Hence, we require,

$$f_u + dg_v > 0 \implies d > 1 \tag{1.18}$$

since $f_u + g_v < 0$. The *min* of $h(k^2)$ in k^2,

$$k^2 = k^2_{min} = \Gamma\frac{f_u + dg_v}{2d} \tag{1.19}$$

$h(k^2_{min})$ gives the marginal instability curve. We can also tune d and arrive at a critically unstable $h(k^2)$. This defines a critical diffusion coefficient d_c, and critical wavenumber k_c given by,

$$k_c^2 = \Gamma \frac{f_u + d_c g_v}{2d_c} \tag{1.20}$$

Noting that $h(k^2) < 0$ for the system to be unstable w.r.t spatial effects results in,

$$(f_u + dg_v)^2 - 4d(f_u g_v - f_v g_u) > 0 \tag{1.21}$$

Adding on to the conditions from (1.10), the final set of necessary and sufficient conditions for Turing pattern formation then becomes,

$$f_u + dg_v > 0 \implies d > 1, \tag{1.22}$$

$$(f_u + dg_v)^2 - 4d(f_u g_v - f_v g_u) > 0. \tag{1.23}$$

Solving Eq. (1.16) for λ we get,

$$2\lambda = \left[k^2(1+d) - \Gamma(f_u + g_v)\right] \pm \left\{\left[k^2(1+d) - \Gamma(f_u + g_v)\right]^2 - 4h(k^2)\right\} \tag{1.24}$$

To destabilize the steady state w.r.t spatial disturbances we require $\mathrm{Re}\lambda(k^2) > 0$. This implies that $h(k^2) < 0$ for some $k \neq 0$. Now when $h(k^2) < 0$, Eq. (1.24) has solution λ which is positive for the same range of wave numbers that make $h < 0$. With $d > d_c$ the range of unstable wavenumbers $k_1^2 < k^2 < k_2^2$ is obtained from solving $h(k^2) = 0$,

$$\begin{aligned} k_1^2 &= \frac{\Gamma}{2d}\left[(f_u + dg_v) - \{(f_u + dg_v)^2 - 4d \mid J \mid\}^{1/2}\right] < k^2 \\ &< \frac{\Gamma}{2d}\left[(f_u + dg_v) + \{(f_u + dg_v)^2 - 4d \mid J \mid\}^{1/2}\right] < k_2^2 \end{aligned} \tag{1.25}$$

Now, we can find the relation for the fastest growing mode k_m^2, by setting $\frac{d\lambda}{dk^2}|_{k^2=k_m^2} = 0$. Differentiating Eq. (1.16) w.r.t k^2 we get,

$$2\lambda\lambda' + \lambda'\left[k^2(d+1) - \Gamma(f_u + g_v)\right] + \lambda(d+1) + 2dk^2 - \Gamma(f_u + dg_v) = 0 \tag{1.26}$$

as $\lambda' = 0$,

$$\lambda(k_m^2) = \frac{1}{d+1}\left[\Gamma(f_u + d_g v) - 2dk_m^2\right] \tag{1.27}$$

Plugging this back in Eq. (1.16) we find the fastest growing mode,

$$k_m^2 = \frac{\Gamma}{d-1}\left\{(d+1)\left[\frac{-f_v g_u}{d}\right]^{1/2} - f_u + g_v\right\} \tag{1.28}$$

Now consider the solution $\boldsymbol{\epsilon}$ given by (Eq. (1.13)), the dominant contributions at long t are those modes with Re$\lambda(k^2)$ and all other modes decay exponentially to zero. From Eq. (1.25) the solution can now be written as,

$$\boldsymbol{\epsilon}(\boldsymbol{r}, t) \sim \sum_{k_1}^{k_2} c_k e^{\lambda(k^2)t} \epsilon_k(t) \text{ for large } t. \tag{1.29}$$

These linearly unstable eigenfunctions which are growing exponentially with time will eventually be bounded by the non-linear terms in the reaction-diffusion equations and an ultimate steady-state inhomogeneous solution will emerge. While singular perturbation analyses can be done near a bifurcation, a general principle of nonlinear evolution for the finite amplitude steady-state spatial patterns is still an open problem and is of relevance to our work discussed in the later chapters.

1.2 AMPLITUDE EQUATIONS

As discussed above, it is in general unclear how the linearly unstable modes saturate due to nonlinear reaction terms and produce stable inhomogeneous spatial patterns at steady state. In this context, we discuss the amplitude equations which is a useful tool to study spatial and temporal distortions of ideal patterns (stripes, spots, hexagons etc) close to onset and can explain the non-trivial shapes of steady state patterns. This technique employs a slowly varying amplitude function that modulates some stationary periodic pattern of interest. Let us assume a simple case of one-component solution with a spatially dependent complex amplitude $A(\mathbf{x}, t)$ in terms of a perturbation as discussed in the previous section $\boldsymbol{\epsilon} = u - u_0$ of the uniform base state u_0,

$$\boldsymbol{\epsilon}(\mathbf{x}, t) = A(\mathbf{x}, t)\boldsymbol{\epsilon_c}(\mathbf{x})e^{ik_c x} + \text{higher order terms}, \tag{1.30}$$

where we choose our expansion around the critical mode of onset k_c. We present the one-dimensional amplitude equation for type I-s instability without derivation here [5],

$$\tau_0 \frac{\partial A}{\partial t}(x, t) = \varepsilon A + \xi_0^2 \frac{\partial^2 A}{\partial x} - g_0 \mid A \mid^2 A. \tag{1.31}$$

Note that the form of the above equation emerges from the underlying translation and parity symmetry and the expansion around the base state k_c and not from the details of the physical system. ε is a perturbation parameter around the base state, while the constants τ_0, ξ_0^2, g_0 depend on the details of the physical system. The suppressing effect of boundary conditions on patterns at steady state can be hence understood by solving the amplitude equation in Eq. (1.31) at steady state.

As an example, we consider a finite domain with $x \in [0,1]$ with boundary conditions $A(0) = A(L) = 0$. An approximation of the amplitude equation yields the following linear onset solution,

$$A(x) = ae^{i\phi} \sin(\frac{\varepsilon^{1/2}}{\tilde{\zeta}_0} x) \tag{1.32}$$

where a is an unknown constant and ϕ is an arbitrary phase. We can determine a by substituting the solution Eq. (1.32) in Eq. (1.31) and collecting terms of order $\sin x$ and ignoring higher order harmonics. In a finite domain, solutions satisfy boundary conditions if $L = n\pi$ with $n = 1, 2, ...$ We re-write L in terms of a re-scaled system size l, $L = \varepsilon^{1/2} \frac{l}{\tilde{\zeta}_0}$. For a periodic domain, we find that critical onset appears shifted from $\varepsilon = 0$ to

$$\varepsilon_c = \pi^2 \left(\frac{\tilde{\zeta}_0}{l} \right)^2 . \tag{1.33}$$

The above equation is an explicit general calculation of the suppression of onset by finite-size effects.

While amplitude equations provide an interesting general way to address the question of pattern saturation it is important to understand its limitations. These equations are valid only close to the onset of patterning, i.e when only a narrow range of wavenumbers are unstable. Also, it relies on perturbations which serve only as a good approximation of a range of parameters and slow modulation of an ideal sinusoidal pattern.

1.3 Limitations of existing methods

Weakly non-linear approaches like amplitude equations and external constraints like fixed boundary conditions, parameter ramps, template patterns, system geometry and deformable boundaries have been useful in understanding the selection and stability of different fundamental modes in a variety of systems [7–12]. However, these are valid only close to onset (in the vicinity of a bifurcation). Far from the onset, where several modes are linearly unstable, pattern-forming reaction-diffusion systems typically produce patterns with a well-defined periodicity, corresponding to a particular mode number n and its harmonics $2n, 3n, ...$ Fourier decomposition of the observed final pattern clearly shows this underlying periodicity of the final pattern (Fig. 3B and 3C). This is quite unintuitive, as the neighbouring modes $n \pm 1$ have a similar growth rate, which would be expected to lead to aperiodic patterns (Fig. 3A). We will discuss the model used in Fig. 3 in more detail below. This regime is quite relevant as most systems in nature cannot be expected to be close to onset.

We also see that if a system initialized with a mispositioned pattern (and therefore far outside the linear regime), then the peaks in the pattern subsequently move towards an

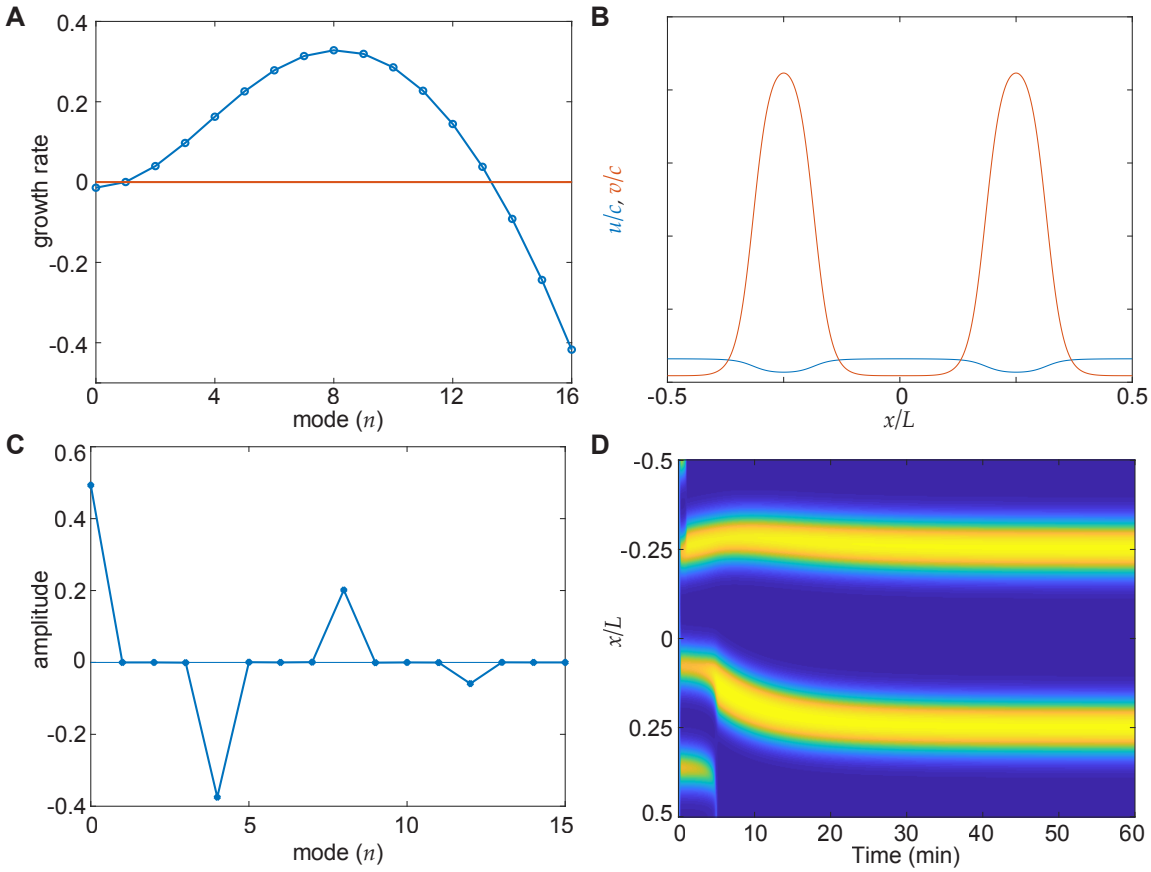

Figure 3: The Turing instability of reaction-diffusion systems [6]. (**A**) The growth rate of different modes for the model in Eq. (1.36) with default parameters and $L = 4$. Note that the growth rate at $n = 0$ is negative – the system is not generically mass-conserving.(**B**) Example of a two-peak steady state pattern. (**C**) The Fourier decomposition of the obtained two-peak pattern shows that it is composed of modes $n = 4, 8, 12, ...$ (**D**) An example kymograph showing pattern development starting from a random perturbation of the uniform state. While the pattern with mode $n = 7$ (3.5 peaks) dominates initially, the pattern coarsens down to two peaks, dominated by mode $n = 4$.

equally separated configuration. This is also clear in models that exhibit coarsening, which we define here as the preference for a steady-state pattern that appears to be lower than that predicted by linear stability. In Fig. 3D we see that a four-peak pattern breaks down and the remaining peaks move to the regularly spaced positions on the domain. From these observations, we conclude that the periodic positioning of peaks is an inherently non-linear effect and not a remnant or a direct consequence of the dominating linear mode of the base state perturbations. While a decomposition to fundamental modes is useful to understand the initial formation of the pattern (starting from a homogeneous steady state), once they saturate into peaks, a different description is required. In this sense, we think of the peaks rather than modes as a more fundamental unit of pattern formation.

In the following work [6], we show that Turing peaks behave like point sinks that move with a velocity proportional to the diffusive flux through them. We find that mass flow through the system is responsible for the regular positioning of the peaks. Diffusion and turnover of the interacting species taken together explain the regular positioning of peaks due to the concept of flux balance [13–15]. The peaks in a mass-conserved reaction-diffusion system do not move, further confirming the link between flux balance and peak movement. Using singular perturbation methods, we solve the system analytically for solutions consisting of well-separated spikes, a regime that occurs in the singular limit $D_v \to 0$ of reaction-diffusion systems [16]. Finally, we find that the regularly positioned configuration minimizes the total mass of the fast diffusing species, i.e the substrate. We then empirically find that this "*minimisation principle*" can be extended to predict not just the positions of the peaks, but also the final number of peaks, even in the presence of coarsening. This is quite important as the amplitude equations and other weakly non-linear approaches are not applicable far from the onset. For the special case of two-component mass-conserved models, Lyapunov functions can be constructed to show that the only non-trivial steady-state solutions consist of a peak or half-peak [17–20]. We show that mass-minimisation also predicts accurately the coarsening down to a single peak in the case of mass-conserved two systems. By comparing the predicted peak number and numerics we uncover the origin of finite size effects observed in reaction-diffusion systems far from the onset. Mass minimisation of the fast diffusing species, therefore, has the potential to be a simple and powerful concept for understanding the pattering in reaction-diffusion systems.

1.4 Illustrative reaction-diffusion model

We introduce an illustrative one-dimensional system, inspired by a recent model of bacterial condensin [21],

$$\partial_t u = D_u \partial_x^2 u - \beta u(u+v)^2 + \gamma v + c\delta - \delta u \tag{1.34a}$$

$$\partial_t v = D_v \partial_x^2 v + \beta u(u+v)^2 - \gamma v - \delta v \,, \tag{1.34b}$$

defined over the domain $[-L/2, L/2]$, with reflective boundary conditions, $\beta, \gamma, c > 0$ and $\delta \geq 0$ and diffusion constant $D_v < D_u$. The model has the form of a mass conserved system with additional terms: a global feed term $c\delta$ and two decay terms δu and δv. This model while structurally similar to some of the classic Turing models such as Brusselator [22] and Schnakenberg [23] models, has some notable properties that we will discuss below. The model does not exhibit oscillatory dynamics.

By writing the source term as $c\delta$, we can change δ, the turnover rate, while leaving the total steady-state concentration c fixed. When $\delta = 0$ and the limit $\delta \to 0$ is well defined

as long as the total initial mass is constrained to be the same as the steady-state mass i.e., $C(0) = c$.

From the above equations, we can easily see that the total mass at steady state is the same for any set of initial conditions i.e.,

$$\int_{-\frac{L}{2}}^{\frac{L}{2}} \bar{u} + \bar{v}\, dx = c, \tag{1.35}$$

where $\bar{u}$ and $\bar{v}$ are the steady-state concentrations.

We implement the following non-dimensionalisation:

$$u \to \frac{u}{c}, \quad v \to \frac{v}{c}, \quad x \to \frac{x}{L}, \quad t \to \frac{D_v}{L^2}t,$$

to obtain

$$\frac{\partial u}{\partial t} = d\frac{\partial^2 u}{\partial x^2} + \Gamma\left(-au(u+v)^2 + v + b(1-u)\right) \tag{1.36a}$$

$$\frac{\partial v}{\partial t} = \frac{\partial^2 v}{\partial x^2} + \Gamma\left(au(u+v)^2 - (1+b)v\right)\ , \tag{1.36b}$$

in terms of the dimensionless variables,

$$a = \frac{\beta c^2}{\gamma}, \quad b = \frac{\delta}{\gamma}, \quad \Gamma = \frac{\gamma L^2}{D_v}, \quad d = \frac{D_u}{D_v}. \tag{1.37}$$

1.5 LINEAR STABILITY ANALYSIS

Let us perform a linear stability analysis on the system. In the absence of diffusion, there is a single fixed point

$$u_0 = \frac{b+1}{a+b+1}, \quad v_0 = \frac{a}{a+b+1}. \tag{1.38}$$

The Jacobian at this point is given by

$$J = \Gamma \begin{bmatrix} f_u & f_v \\ g_u & g_v \end{bmatrix}_{(u_0, v_0)} \tag{1.39}$$

$$= \Gamma \begin{bmatrix} -a - 2au_0 - b & 1 - 2au_0 \\ a + 2au_0 & 2au_0 - 1 - b \end{bmatrix}. \tag{1.40}$$

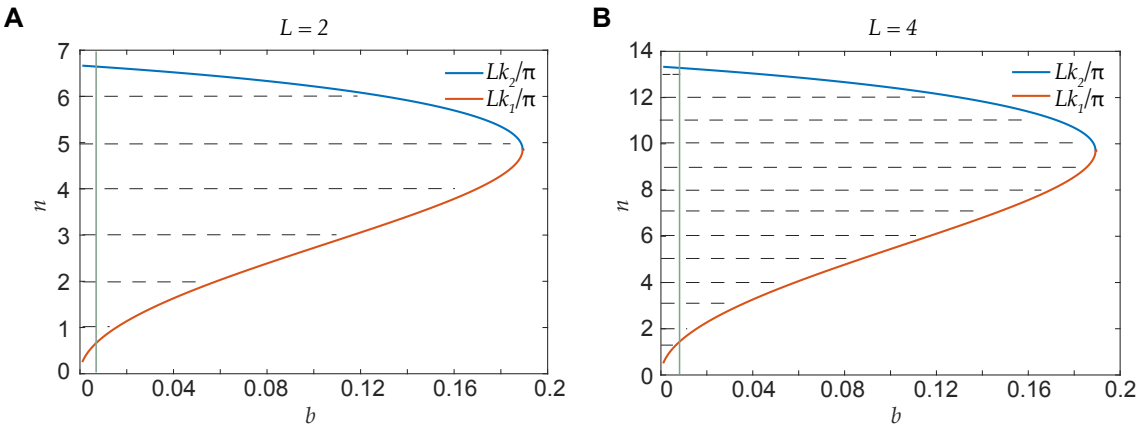

Figure 4: Regions of instability. (**A,B**). We find that a broad range of modes is unstable for different parameters. Plotted is the region of instability (see also Fig 1D in the main text) as a function of b with $a = 3.75$. k_1 and k_2 are the bounds of instability as defined by Eq. (1.45). The horizontal dashed lines indicate the discrete modes that fit in the range. The vertical green line represents our default value of $b = 0.0039$. $L = 2$ and $L = 4$ in (**A**) and (**B**) respectively.

where f and g are the reaction terms in Eq. (1.36a) and Eq. (1.36b) respectively. The trace and determinant of the Jacobian are easily found to be

$$TrJ = \Gamma(-a - 1 - 2b)$$
$$DetJ = \Gamma^2(b\,(a + b + 1)).$$

Since, $TrJ < 0$ and $DetJ > 0$, the homogeneous fixed point is always stable in the absence of diffusion for any set of parameters $a, b, \Gamma > 0$. For the mass-conserved case $b = 0$, the determinant vanishes. This is due to a zero eigenvalue, the eigenvector of which does not obey the conservation condition. The other eigenvector corresponds to the mass-conserving direction $(1, -1)^T$ and has eigenvalue $f_u - f_v = -a < 0$. Hence, the base state is stable for perturbations preserving mass conservation.

Following the standard approach [24], we now consider a spatial perturbation around the uniform state $u = u_0 + \delta u$, $v = v_0 + \delta v$ in presence of diffusion. The evolution of this perturbation $\boldsymbol{w} = \begin{pmatrix} \delta u \\ \delta v \end{pmatrix}$ can be written as

$$\boldsymbol{w}(x,t) = \sum_k \boldsymbol{c}_k \exp^{\lambda_k t} \epsilon_k(x)\ ,$$

where the ϵ_k are the eigenfunctions of the Laplacian

$$\nabla^2 \epsilon_k(x) = -k^2 \epsilon_k(x) \tag{1.41}$$

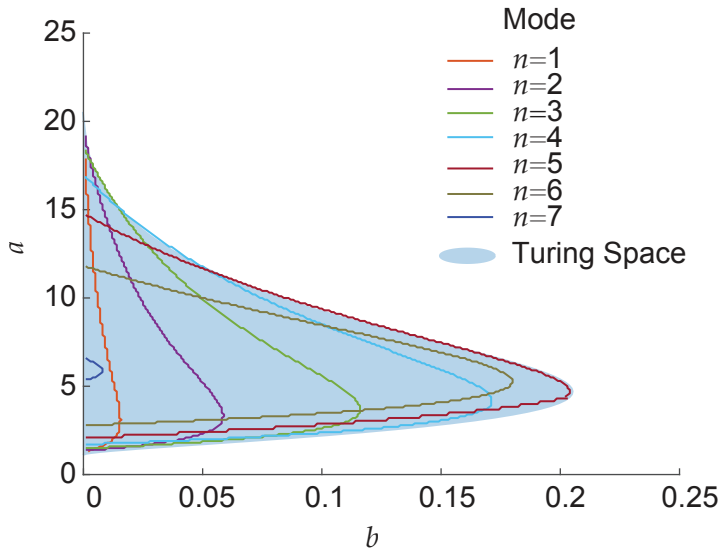

Figure 5: The Turing space of our reaction-diffusion system. The regions of instability of each mode on a domain of length $L = 2$ (the region bounded by the y-axis and the corresponding coloured curve). The blue-shaded region shows the Turing space for an infinite domain.

subject to the given boundary conditions. Inserting this into the equation Eq. (1.36) and linearizing around the homogeneous steady state we obtain the eigenvalue problem

$$\left|J - k^2 D - \lambda_k \mathbf{I}\right| = 0 \quad \text{where } D = \begin{pmatrix} D_u & 0 \\ 0 & D_v \end{pmatrix}. \tag{1.42}$$

The conditions for the growth rate λ_k to be positive for some range of wave numbers $k_1^2 < k^2 < k_2^2$ is given by [24]

$$f_u + dg_v > 2\sqrt{d(f_u g_v - f_v g_u)} \tag{1.43}$$

which we evaluate to find

$$d\frac{b+1}{a+b+1}(a-b-1) - \frac{a^2+b^2+4ab+3a+b}{a+b+1} - 2\sqrt{db(a+b+1)} > 0. \tag{1.44}$$

The inequality Eq. (1.44) relating a, b and d determines the parameter values for which there exists a band of unstable wave numbers i.e. the *Turing space*. As discussed in section 1.1, for a finite domain the eigenvalues of the Laplacian are discrete. In particular, for the domain $[0, L]$ with reflective boundary conditions, the wave number $k = \frac{n\pi}{L}$ for integer n and the eigenfunctions are, labelling now by n, $\epsilon_n = \cos(\frac{n\pi x}{L})$. Since k is discrete the condition Eq. (1.44) is necessary but not sufficient for instability. A discrete mode n with corresponding wave number $k = \frac{n\pi}{L}$ must exist within the range of instability $[k_1, k_2]$. Note

that while Γ does not enter into the condition Eq. (1.44), which, if any, discrete modes fall into the range $[k_1, k_2]$ does depend on Γ. By a straightforward calculation, we find

$$k_1^2 = \frac{\Gamma}{2d}\left[(-a-2au_0-b+d(2a-1-b)) - \left((-a-2au_0-b+d(2a-1-b))^2 - 4db(a+b+1)\right)^{1/2}\right]$$
$$k_2^2 = \frac{\Gamma}{2d}\left[(-a-2au_0-b+d(2a-1-b)) + \left((-a-2au_0-b+d(2a-1-b))^2 - 4db(a+b+1)\right)^{1/2}\right]. \tag{1.45}$$

In Fig. 5, we show that region defined by Eq. (1.44) as well as the parameter regions in which each mode is unstable. Using Eq. (1.28) we can evaluate the fastest growing mode on an infinite domain,

$$k_m^2 = \frac{\Gamma}{d-1}\left[a + (d+1)\sqrt{\frac{a\ (a+3\,b+3)\ (a-b+2\,a\,b-1)}{d\,(a+b+1)^2}} + \frac{4\,a\ (b+1)}{a+b+1} - 1\right]. \tag{1.46}$$

From the above analysis, the fastest growing mode from linear stability predicts the wavelength of the final pattern at steady state. In the following chapter, we compare the results from the numerical simulations of the toy model Eq. (1.34) with predictions from linear stability.

2

EFFECTS OF TURNOVER ON PEAK MOVEMENT

In this chapter, we discuss the preliminary investigations of our model Eq. (1.34) introduced in the previous chapter using numerical simulations. Upon simulating the system, we observed that peaks are initially formed as dictated by linear stability and then coarsening occurs in which peaks break down and the remaining mispositioned peaks subsequently move to regularly positioned configurations (Fig. 3D). This phenomenon has also been observed in the model it is based on [21], and other reaction-diffusion systems and is dubbed competition instability [25–27] (also referred to as interrupted coarsening [28]) in that the final dominant mode has a longer wavelength (lower number of peaks) than predicted by linear stability analysis. It is not evident in models/parameter sets for which the linear prediction holds as in that case, the peaks are created at their steady-state positions. For our default parameter set with $L = 4$ ($\Gamma = 4800$), linear stability predicts (Fig. 3A) that the pattern consists of four peaks (or valleys) (mode $n = 8$) whereas the obtained steady-state pattern most frequently consists of two peaks (mode $n = 4$) (Fig. 3B, 6B). In this study, we

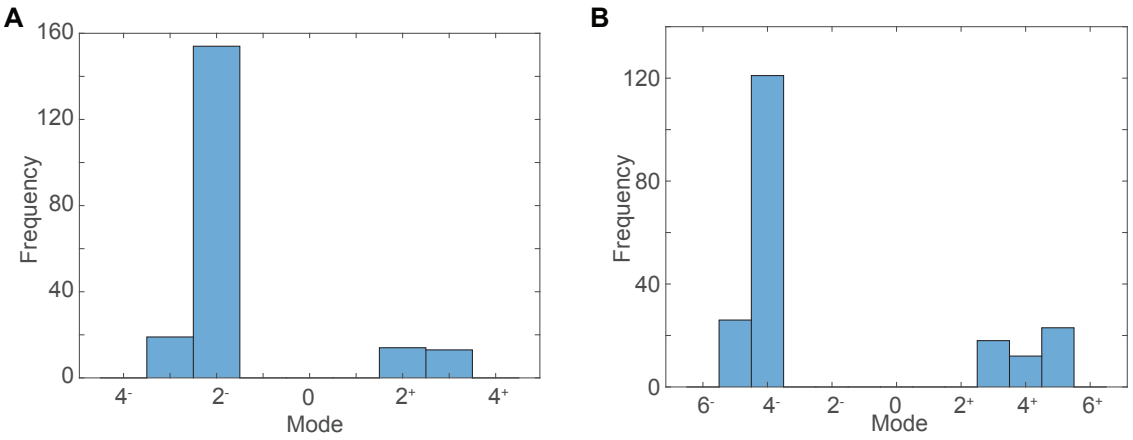

Figure 6: Dominant phase of the steady state pattern (**A,B**). The system was solved for 200 different initial random perturbations of the uniform base state with $L = 2$ (**A**) and $L = 4$ (**B**). The phase of the pattern (the sign of the amplitude of the dominant mode) is indicated by a minus or plus sign. The steady-state pattern is dominated by a mode lower ($n = 2$ and $n = 4$) than what is predicted by the linear stability analysis ($n = 4$ and $n = 8$ respectively). The growth rate of each mode for $L = 4$ is shown in Fig. 3D. In (**A**) and (**B**) we use our default parameter set state in default parameters.

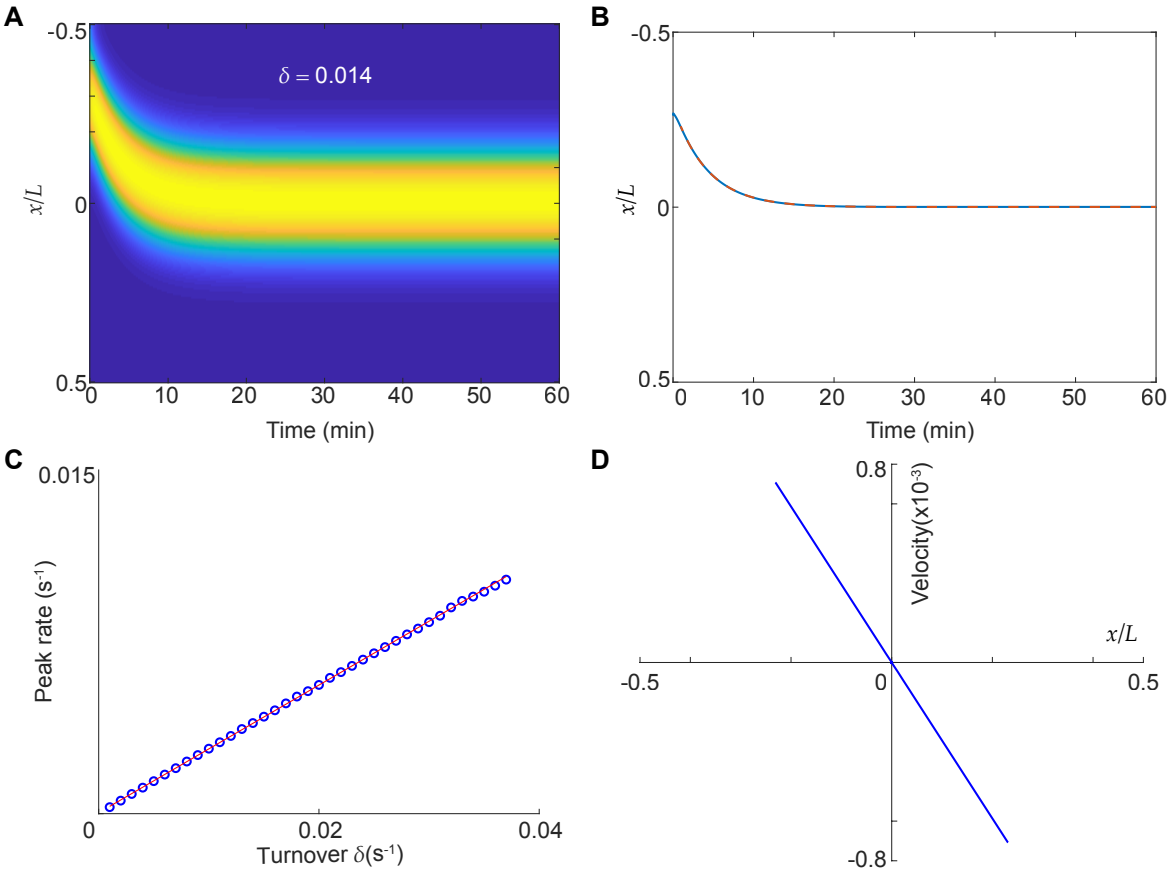

Figure 7: Peak movement and regular positioning depend on flux through the system. (**A**) The system is initialised with a peak away from the mid-domain. The peak subsequently moves to the mid-domain. (**B**) The centroid of the peak (blue line) is plotted as a function of simulation time. The orange dashed line is an exponential fit. (**C**) The rate of movement obtained from fitting the centroid to an exponential as in (**B**) shows a linear dependence on the turnover rate δ. (**D**) Peak velocity is linear in peak position. Default parameters.

mainly focus on the interior peaks on the domain, as the peaks on the boundary do not move (positive amplitudes, Fig. 6A, B).

2.1 PEAK MOVEMENT RATE DEPENDS LINEARLY ON TURNOVER

To examine the movement of peaks in more detail, we focused on the case of a single peak ($n = 2$), typically obtained for $L = 2$ ($\Gamma = 1200$). Examining the movement of the peak (Fig. 7A) for our default parameters, we found that it moves to the mid-domain exponentially in time (Fig. 7B). We found that the peak velocity is linearly proportional to its displacement from the mid-domain (Fig. 7B). This was the case whether the system was initialised with a random perturbation to the homogeneous state or with a single peak pre-formed somewhere on the domain. While peaks moving to respect the symmetry of the system is

perhaps not surprising, it is underappreciated and is relevant for understanding the periodic appearance and positioning of peaks. Indeed, we found that the rate of movement is directly proportional to the turnover rate δ (Fig. 7C) (or equivalently $c\delta$ the flux through the system per unit length) so that in the mass-conserved limit of $\delta \to 0$, peaks do not move (Fig. 8A, B).

Numerical Methods

The simulations were performed on a spatial lattice $x \in [-\frac{L}{2}, \frac{L}{2}]$ and time domain $t \in [0, T]$, where L is the length of the spatial domain and T, the total time. The MATLAB solver *pdepe* was used to solve the time-dependent equation Eq. (1.34). The simulations were performed with the following **default** parameters (unless explicitly stated otherwise):

$$D_u = 0.3, D_v = 0.012, L = 2, c = 300,$$
$$\beta = 1.5 \times 10^{-4}, \gamma = 3.6, \delta = 0.014.$$

the equivalent dimensionless parameters are

$$d = 25, a = 3.75, b = 0.0039, \Gamma = 1200.$$

The simulations were run long enough to obtain the true steady state. The relative and absolute tolerances in the difference between two values of iteration were set to 10^{-6} and 10^{-12} respectively. We used reflective boundary conditions

$$\frac{\partial u}{\partial x}\Big|_{x=-L/2,L/2} = \frac{\partial v}{\partial x}\Big|_{x=-L/2,L/2} = 0.$$

The simulations were initialised with small perturbations from a homogeneous steady state. The results are based on simulations that were run for very long times with very low error tolerances.

2.2 Absence of turnover leads to complete coarsening

Mass-conserving RD models exhibit a complete coarsening process in that the final steady-state pattern is either mono-modal (periodic or reflective) or monotonic (reflective only) depending on the boundary conditions, as has been proved explicitly for several models [17–20]. We find the same coarsening behaviour here (Fig. 7D). In our simulations, we only obtain the half-peak solution for very short domains i.e. when the width of the interface is

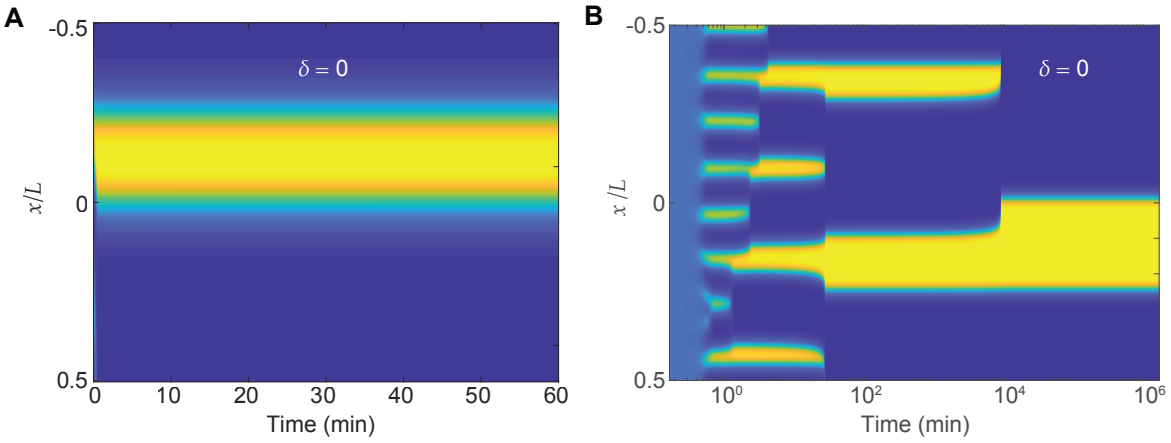

Figure 8: Absence of turnover leads to coarsening.(**A**) A single peak in the mass conserved limit $\delta = 0$ can be positioned anywhere on the domain. No peak movement is observed (**B**) The mass-conserved system exhibits complete coarsening. Irrespective of how many peaks there are initially, the pattern eventually coarsens to a single peak, the position of which depends on which peak of the initial pattern has not coarsened. In (**B**) $\Gamma = 19200$ ($L = 10$).

comparable to the domain length. If the domain length or other parameters are chosen such that there is initially more than one peak then the coarsening process results (eventually) in a single interior peak (Fig. 8A). In this limit, since peaks do not move, the position of this final peak is purely determined by whichever peak of the transient state remains after coarsening i.e. the steady state peak does not respect the symmetry of the system (reflective). We tested these conclusions by initialising the system with a single preformed peak (constructed as a translation of the non-mass conserved steady-state solution). We found that pre-formed peaks are indeed stable solutions that do not move (Fig. 8A). Thus, the mass-conserved case $b = 0$ with reflective boundary conditions has a continuum of single-peak states across the domain, whereas there is at most one unique single-peak solution (regularly positioned) for $b > 0$. This implies that regular positioning is not an intrinsic property of the system determined by its symmetry but rather depends on the parameter b. In the next chapter, we show an analogous simpler system of point sinks exhibiting similar dynamics. These results demonstrate a connection between peak movement towards the regular positioned configuration and the flux of mass through the system. In the successive chapters, we study a simpler system of moving point sinks. We prove analytically that the regular positions balance the flux across the sinks and minimise the mass of the diffusive species. We discuss the similarity of this simpler system to moving Turing peaks and show that they are identical in the regime of low mass flow.

3

MOVEMENT AND REGULAR POSITIONING OF POINT SINKS

In this chapter, we discuss an illustrative toy model involving diffusion and moving point sinks. We will see that this is a simpler and more tractable model to understand the movement of peaks in the more complicated Turing system. We consider the steady-state diffusion equation for a substrate variable $A = A(x)$ over a 1D domain $x \in [-L/2, L/2]$ in the presence of global source and decay terms. There exist n localised point sinks at positions $\boldsymbol{x} = (x_1, \cdots, x_n)$ (each with rate μ):

$$D\frac{d^2A}{dx^2} + c\delta - \delta A - \sum_{i=1}^{n} \mu L \delta(x - x_i)A = 0 \,. \tag{3.1}$$

We impose zero-flux boundary conditions. As before Eq. (1.36a), we write the global source term in terms of the decay rate δ and a concentration c, which is the steady-state concentration in the absence of the point sinks. A simpler system without the decay term and with perfect points sinks (i.e. $\mu \to \infty$) was used by Ietswaart et al. to model the positioning of plasmids within rod-shaped bacterial cells [14]. They found that the flux (gradient differential) across each sink vanishes if and only if the sinks are regularly positioned and, therefore, if sinks were to move up the concentration gradient, they would be regularly positioned. We extend this result to the more complicated case of equation Eq. (3.1) in the presence of decay terms. The decay term introduces an additional diffusive length scale $\sqrt{D/\delta}$, which is the distance that a molecule of A would diffuse (if it does not encounter any point sinks) before it decays. It is small when either diffusion is slow or the decay rate δ is fast.

We can write the solution to Eq. (3.1) as

$$A(x) = c - \sum_i \mu_i' G(x; x_i) \,, \tag{3.2}$$

where $G(x;x_i)$ is the modified Green's function defined by

$$\begin{aligned} -\frac{L^2}{\kappa^2}G_{xx}(x;x_i) + G(x;x_i) &= L\delta(x-x_i) \\ G_x(\pm\frac{L}{2};x_i) = 0, \quad \frac{1}{L}\int_{-\frac{L}{2}}^{\frac{L}{2}} G(x;x_i)dx &= 1\ , \end{aligned} \tag{3.3}$$

where the dimensionless parameter $\kappa = L\sqrt{\frac{\delta}{D}}$ is the ratio of the length of the domain to the length-scale of diffusion. The coefficients $\mu_i' = \mu_i'(\boldsymbol{x})$ are determined by the linear algebraic conditions

$$\begin{aligned} \mu_i' &= \lambda A(x_i) \quad i = 1,\dots,n \\ &= \lambda\left(c - \sum_j \mu_j' G(x_i;x_j)\right)\ , \end{aligned} \tag{3.4}$$

where we define a second dimensionless parameter $\lambda = \frac{\mu}{\delta}$, the ratio of the sink and background decay rates. The quantities μ_i' are directly related to the flux J_i leaving the system via the imperfect sinks,

$$\begin{aligned} J_i &= J_{i+} + J_{i-} \\ &= -D\sum_j \mu_j' \left[G_x(x_i^+;x_j) - G_x(x_i^-;x_j)\right]\ , \end{aligned}$$

where $J_i = |D\frac{dA}{dx}|$ and the $-$ and $+$ subscripts refer to the diffusive flux from the left and right respectively. We also define the flux differential across each sink as

$$\Delta J_i = \frac{1}{2}(J_{i+} - J_{i-}) \tag{3.5}$$

$$= -\frac{D}{2}\sum_j \mu_j' \left[G_x(x_i^+;x_j) + G_x(x_i^-;x_j)\right]\ . \tag{3.6}$$

3.1 FLUXES BALANCE AT REGULAR POSITIONS

To evaluate flux across the sinks we have to solve the coupled Green's function equation Eq. (3.3) and the algebraic equation Eq. (3.2). The explicit form of Green's function is

$$G(x;x_i) = \frac{\kappa}{2}\frac{\cosh\left(\kappa\frac{x+x_i}{L}\right) + \cosh\left(\kappa\frac{|x-x_i|-L}{L}\right)}{\sinh(\kappa)}\ . \tag{3.7}$$

where the second term describes the effect of the point sinks. The derivative $G(x;x_i)$ with respect to x is discontinuous at $x = x_i$

$$G_x(x;x_i) = \begin{cases} \frac{\kappa^2}{2L}\frac{\sinh(\kappa\frac{x+x_i}{L})-\sinh(\kappa\frac{x_i-x-L}{L})}{\sinh(\kappa)} & -\frac{L}{2} \le x < x_i \\ \frac{\kappa^2}{2L}\frac{\sinh(\kappa\frac{x+x_i}{L})+\sinh(\kappa\frac{x-x_i-L}{L})}{\sinh(\kappa)} & x_i < x \le \frac{L}{2}\ . \end{cases}$$

Note the following property

$$G_x(x_i^+;x_j) - G_x(x_i^-;x_j) = -\frac{\kappa^2}{L}\delta_{ij}\ . \tag{3.8}$$

Using this, the flux differential ΔJ_i defined in Eq. (3.5) can be written as

$$\Delta J_i = -D\sum_j \mu_j' \left[G_x(x_i^+;x_j) + \frac{\kappa^2}{2L}\delta_{ij}\right]\ . \tag{3.9}$$

We define the regular position $\bar{x}$ as the equally spaced configuration across the domain,

$$\bar{x}_i = \frac{L}{n}i - \frac{L}{2}(\frac{1}{n}+1), \tag{3.10}$$

Property I

To evaluate the flux differential at regular positions we note the following properties of the Greens functions. Summing over the $G_{ij} = G(\bar{x}_i, \bar{x}_j)$,

$$\boxed{\sum_j G_{ij} = \frac{\kappa}{2}\coth(\frac{\kappa}{2n}) \quad \forall\ i,} \tag{3.11}$$

where we have used the expression for regular positions $\bar{x}$ from Eq. (3.10) and the identity,

$$\sum_{j=1}^{n}\cosh(\frac{\kappa}{n}(j+m)) = \mathrm{csch}(\frac{\kappa}{2n})\sinh(\frac{\kappa}{2})\cosh(\frac{\kappa}{2n}(2m+n+1)).$$

Property II

Since the summation of the Greens Function G over any of its rows or columns is the same Eq. (3.11), the vector of $\hat{\boldsymbol{e}} = (1,1,....1)_n$, is an eigenvector of G. The defining

equations for the $\boldsymbol{\mu'}$ Eq. (3.4), at regular positioning $\boldsymbol{x} = \bar{\boldsymbol{x}}$, we obtain the matrix equation

$$(\lambda G + 1)\boldsymbol{\mu'}(\bar{\boldsymbol{x}}) = \lambda c \hat{\boldsymbol{e}} \,. \tag{3.12}$$

Since $\hat{\boldsymbol{e}}$ is an eigenvector of $\lambda G + 1$, we must have that

$$\boldsymbol{\mu'}(\bar{\boldsymbol{x}}) = C_1 \hat{\boldsymbol{e}} \,,$$

i.e. all the μ_i' are necessarily identical at regular positions or equivalently, the profile of A is symmetric. We can sum over any row and use Eq. (3.11) to find

$$\boxed{\mu_i'(\bar{\boldsymbol{x}}) = \frac{\lambda c}{1 + \lambda \frac{\kappa}{2} \coth\left(\frac{\kappa}{2n}\right)} \,.} \tag{3.13}$$

Property III

We can similarly define a matrix G_x^+ by evaluating the derivative of the Green's function at regular positioning $(G_x^+)_{ij} = G_x(\bar{x}_i^+; \bar{x}_j)$. Summing over the j th column we find

$$\boxed{\sum_j (G_x^+)_{ij} = -\frac{\kappa^2}{2L}} \tag{3.14}$$

in this case by using the identity,

$$\sum_{j=1}^{n} \sinh(a(j+m)) = \operatorname{csch}(\frac{\kappa}{2n}) \sinh(\frac{\kappa}{2}) \sinh(\frac{\kappa}{2n}(2m+n+1)).$$

As in Ietwaart et al., we can determine the configurations for which the flux differentials are all zero. The flux differential across each sink is given by Eq. (3.6). We evaluate this expression for regularly positioned sinks $\boldsymbol{x} = \bar{\boldsymbol{x}}$. Using the fact that all μ_j' are identical for regularly positioned sinks Eq. (3.13) and equation Eq. (3.14), it follows immediately that the flux differentials vanish at regular positions

$$\boxed{\Delta J(x_i) = 0 \quad \forall x_i} \qquad \textbf{Result I} \tag{3.15}$$

While we have shown above that regular positions balance flux differentials across the sinks we need to show that it is indeed a unique configuration. We perform a power series expansion of ΔJ_i in κ and assuming that the configurations are κ-independent (we find no

evidence for the existence of such κ dependent states from numerically solving the system) we have,

$$\mu_i' = \mu_{0i} + \mu_{2i}\kappa^2 + ...$$
$$G(x_i; x_j) = G_0(x_i; x_j) + G_2(x_i; x_j)\kappa^2 + ... ,$$

We first expand μ_i' and $G'(x_i; x_j)$. For the lowest order terms, we find first that $G_0(x_i; x_j) = 1$. Inserting this into the defining equation for the μ_i' (Eq. (3.4)), we have

$$\mu_{0i}' = (c - \sum_j \mu_{0j}')\lambda \quad \forall\, i$$

which has solution,

$$\mu_{0i}' = \mu_0' = \frac{\lambda c}{1 + n\lambda}. \tag{3.16}$$

We then have

$$\begin{aligned}
\frac{\Delta J_i}{\delta L} &= -\frac{D}{\delta L}\sum_j \mu_j' \left[G_x(x_i^+; x_j) + \frac{\kappa^2}{2L}\delta_{ij}\right] \\
&= -\frac{1}{2}\left[\sum_{j=1}^{n} \mu_j' \frac{\sinh(\kappa\frac{x_i + x_j}{L})}{\sinh(\kappa)} + \sum_{j=1}^{i} \mu_j' \frac{\sinh(\kappa\frac{x_i - x_j - L}{L})}{\sinh(\kappa)} - \sum_{j=i+1}^{n} \mu_j' \frac{\sinh(\kappa\frac{x_j - x_i - L}{L})}{\sinh(\kappa)} + \mu_i'\right] \\
&= -\frac{\mu_0'}{2L}\left[\sum_{j=1}^{n}(x_i + x_j) + \sum_{j=1}^{i-1}(x_i - x_j - L) - \sum_{j=i+1}^{n}(x_j - x_i - L)\right] + O(\kappa^2) \\
&= -\frac{n\mu_0'}{L}\left[x_i - \frac{L}{n}i + \frac{L}{2}(\frac{1}{n} + 1)\right] + O(\kappa^2) \ .
\end{aligned} \tag{3.17}$$

Hence, all the flux differentials ΔJ_i vanish for regularly positioned sinks $x_i = \bar{x}_i = \frac{L}{n}i - \frac{L}{2}(\frac{1}{n} + 1)$. Since for any zero of $\Delta J_i(\boldsymbol{x})$, all terms in the κ expansion must vanish independently and as we have shown previously that $\bar{x}$ is a solution to the full equation, it suffices to show uniqueness for the κ^0 term.

3.2 Mass minimisation at regular positions

The total mass (concentration) of $A(x)$ is readily given by integrating equation Eq. (3.2)

$$M(\boldsymbol{x}) = \frac{1}{L}\int_{-\frac{L}{2}}^{\frac{L}{2}} A(x)dx = c - \sum_{i=1}^{n} \mu_i \ . \tag{3.18}$$

We would like to show that the regularly positioned configuration is the unique κ-independent stationary point of M. Hence, we calculate the derivative of M,

$$\frac{\partial}{\partial x_m} M\Big|_{x=\bar{x}} = -\frac{\partial}{\partial x_m} \sum_i \mu_i' \Big|_{x=\bar{x}} .$$

Using Eq. (3.4) ($\mu_i' = \lambda(c - \sum_j \mu_j' G(x_i; x_j))$), we can evaluate the derivative of $\sum_i \mu_i'$ with respect to an arbitrary sink position x_m,

$$\frac{\partial}{\partial x_m} \sum_i \mu_i' = -\lambda \sum_{i,j} \mu_j' G_{x_m}(x_i; x_j) - \lambda \sum_{i,j} G(x_i; x_j) \frac{\partial}{\partial x_m} \mu_j',$$

Evaluating this expression at regular positioning, and defining $C := \sum_j G(\bar{x}_i; \bar{x}_j) = \frac{\kappa}{2} \coth(\frac{\kappa}{2n})$ from equation Eq. (3.11), we obtain

$$(\frac{1}{\lambda} + C) \frac{\partial}{\partial x_m} \sum_i \mu_i' \Big|_{x=\bar{x}} = -\sum_{i,j} \mu_j' G_{x_m}(x_i; x_j) \Big|_{x=\bar{x}} .$$

From Eq. (3.13) we know that all the μ_j' are identical at regular positions. Hence, we just need to evaluate the derivative of the Green function $\sum_{i,j} G_{x_m}(x_i; x_j)\Big|_{x=\bar{x}}$. Inserting the definition of $G(x_i; x_j)$ from equation Eq. (3.7) we have

$$\begin{aligned}
\sum_{i,j} G_{x_m}(x_i; x_j)\Big|_{x=\bar{x}} &= \frac{\kappa}{2\sinh(\kappa)} \sum_{i,j} \frac{\partial}{\partial x_m} \left[\cosh(\kappa \frac{x_i + x_j}{L}) + \cosh(\kappa \frac{|x_i - x_j| - L}{L})\right]\Big|_{x=\bar{x}} \\
&= \frac{\kappa^2}{2L\sinh(\kappa)} \sum_{i,j} \Bigg[\delta_{mi} \left(\sinh(\kappa \frac{x_i + x_j}{L}) + sgn(x_i - x_j) \sinh(\kappa \frac{|x_i - x_j| - L}{L})\right) \\
&\qquad + \delta_{mj} \left(\sinh(\kappa \frac{x_i + x_j}{L}) - sgn(x_i - x_j) \sinh(\kappa \frac{|x_i - x_j| - L}{L})\right)\Bigg]\Big|_{x=\bar{x}} \\
&= \frac{\kappa^2}{L\sinh(\kappa)} \sum_i \left[\sinh(\kappa \frac{x_m + x_i}{L}) + sgn(x_m - x_i) \sinh(\kappa \frac{|x_m - x_i| - L}{L})\right]\Big|_{x=\bar{x}} \\
&= \frac{\kappa^2}{L\sinh(\kappa)} \left[-\sinh(\kappa) - \sinh(-\kappa)\right] = 0
\end{aligned}$$

where the last line follows from noting that the summations are the same as in equation Eq. (3.14) but without the $i = m$ term. Hence, we have shown that regular positioning is a stationary configuration of the total mass

$$\frac{\partial}{\partial x_m} M\Big|_{x=\bar{x}} = 0 .$$

To show that regular positioning is the (κ-independent) unique stationary point, we proceed as in the previous section and perform a power series expansion of M,

$$M = M_0 + M_2\kappa^2 + \cdots \tag{3.19}$$

It then suffices to show uniqueness for the first non-trivial order in the expansion. For the Green's function, we have

$$G_0(x_i; x_j) = 1, \quad G_2(x_i; x_j) = \frac{x_i^2 + x_j^2 - L|x_i - x_j|}{2L^2} + \frac{1}{12}.$$

We already saw that $\mu'_{0i} = \mu'_0 = \frac{\lambda c}{1+n\lambda}$ and hence M_0 is a constant. Inserting these into the equation for μ'_{2i} using Eq. (3.4),

$$\mu'_{2i} = -\lambda \sum_j (\mu'_{2j} + \mu'_{0j} G_2(x_i; x_j)), \tag{3.20}$$

we obtain

$$\begin{aligned} M_2 &= -\sum_i \mu'_{2i} \\ &= \frac{\lambda}{1+n\lambda} \mu'_0 \sum_i \sum_j \left(\frac{x_i^2 + x_j^2 - L|x_i - x_j|}{2L^2} + \frac{1}{12} \right) . \end{aligned} \tag{3.21}$$

The derivative of M_2 is then proportional to

$$\begin{aligned} &\frac{\partial}{\partial x_m} \sum_i \sum_j \left[x_i^2 + x_j^2 - L|x_i - x_j| \right] \\ &= \frac{\partial}{\partial x_m} \left[2n \sum_i x_i^2 - \sum_i \sum_{j \le i} (x_i - x_j) L - \sum_i \sum_{j > i} (x_j - x_i) L \right] \\ &= \frac{\partial}{\partial x_m} \left[2n \sum_i x_i^2 - L(\sum_i i x_i - \sum_i (n-i) x_i - \sum_i \sum_{j \le i} x_j + \sum_i \sum_{j > i} x_j) \right] \\ &= [4n x_m - L(m - (n-m) - (n - m + 1) + m - 1)] \\ &= 4n \left[x_m - \frac{L}{n} m + \frac{L}{2} + \frac{L}{2n}) \right] \end{aligned} \tag{3.22}$$

which vanishes only for the regularly positioned configuration

$$x_m = \bar{x} = \frac{L}{n} m - \frac{L}{2} (\frac{1}{n} + 1). \tag{3.23}$$

Hence, we have shown that regularly positioned sinks are the unique κ-independent configuration for which the total mass, M, is stationary. Based on our numerical results, we assume that this configuration is generically a minimum.

Hence we have shown that the regular positioned configuration $\bar{\boldsymbol{x}}$ is al a stationary point of the mass M, i.e.

$$\boxed{\frac{\partial}{\partial x_i} M(\boldsymbol{x})\Big|_{\boldsymbol{x}=\bar{\boldsymbol{x}}} = 0 \qquad \forall\, i.} \qquad \textbf{Result II} \quad (3.24)$$

Based on our numerical observations, we conclude that this stationary point is generically a minimum.

3.3 MOVING POINT SINKS

Therefore, if we add time dependence to the system by specifying the sink velocities as being proportional to their flux differential. We can make sink movement explicit and given our results above, two natural choices are to take the sink velocities as either directly proportional to the flux differentials ΔJ_i,

$$\boxed{\frac{dx_i}{dt} = \nu \Delta J_i(\boldsymbol{x})} \qquad \textbf{Result III} \qquad (3.25)$$

or to the derivative of the mass $M(\boldsymbol{x})$ with respect to the sink positions

$$\boxed{\frac{dx_i}{dt} = -\frac{n}{2}\nu D \frac{\partial}{\partial x_i} M(\boldsymbol{x})\,,} \qquad \textbf{Result IV} \quad (3.26)$$

where ν is some parameter and the choice of pre-factor is for later convenience (see box below). For the latter choice, the system is analogous to that of n over-damped particles moving in a potential $U(\boldsymbol{x})/k_B T = \frac{n\nu}{2} M(\boldsymbol{x})$. Importantly note that while the velocities in equation Eq. (3.25) are specified in terms of local quantities, in Eq. (3.26), they are specified in terms of the global quantity, $M(\boldsymbol{x})$. Thus if sinks move up the concentration gradient (in the direction of greatest flux) or along the "mass potential", they will be regularly positioned.

Furthermore, the connection between regular positioning and mass minimisation appears to be generalisable and we have observed numerically that it holds for spatial sinks i.e. if the delta function in Eq. (3.1) is replaced by a peak-shaped spatial function such as a Gaussian function or $\mathrm{sech}^2(x)$, then the total mass is minimised when the sink is centred at mid-domain. In either case, the steady-state solution consists of regularly positioned sinks as this is the only configuration for which the fluxes balance and for which the mass is at its unique minimum. This holds as long as $\delta > 0$. For $\delta = 0$, all the velocities vanish identically and the sinks don't move. Hence all sink positions are stable in that limit. This is

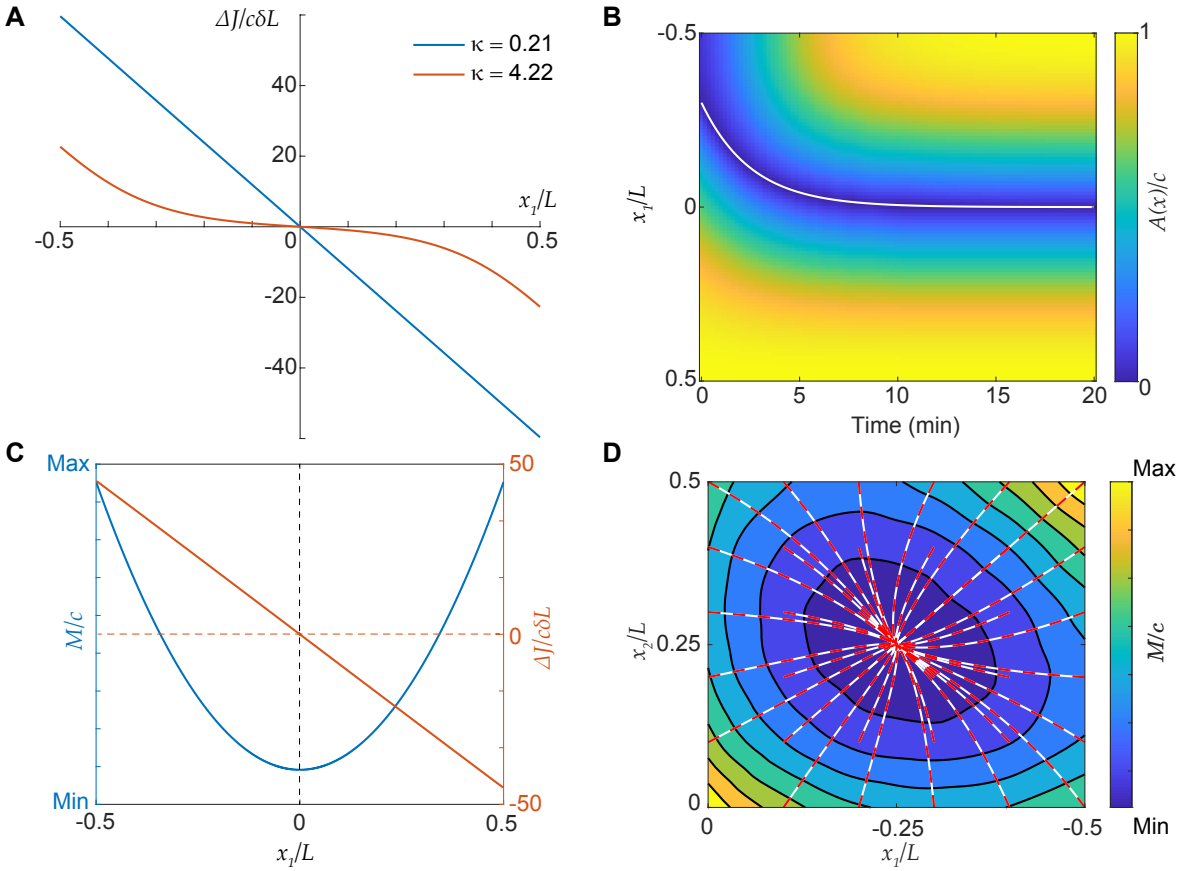

Figure 9: Moving point sinks are regularly positioned and their movement depends on the diffusive length-scale. (**A**) Flux differential across a point sink is calculated analytically as a function of sink position x_1 for two values of κ. (**B**) When ΔJ is linear in x_1 the sink moves exponentially to mid-domain. The flux differential across a point sink is linear in sink position for $\kappa \ll 1$. It vanishes in the middle of the domain. (**C**) Mass M in the system as a function of sink position for a single-point sink is plotted in blue. The Mass is minimal as the sink approaches the middle of the domain. (**D**) Sample trajectories of the two-sink system. White lines are sample trajectories obtained using Eq. (3.25), while the overlaid red dashed lines are trajectories obtained using Eq. (3.26). The coloured contour shows the total mass M as a shown as a function of the sink positions. The minimum occurs at the steady-state configuration (sinks at opposite quarter positions). Parameters: $D = 0.3, \lambda = 166.1, c = 1, L = 1, \nu = 1$. $\kappa = 0.21$ in (B) and (**C**). In (**D**) $L = 2$.

reminiscent of the behaviour that we found in the previous section for the mass-conserved Turing system.

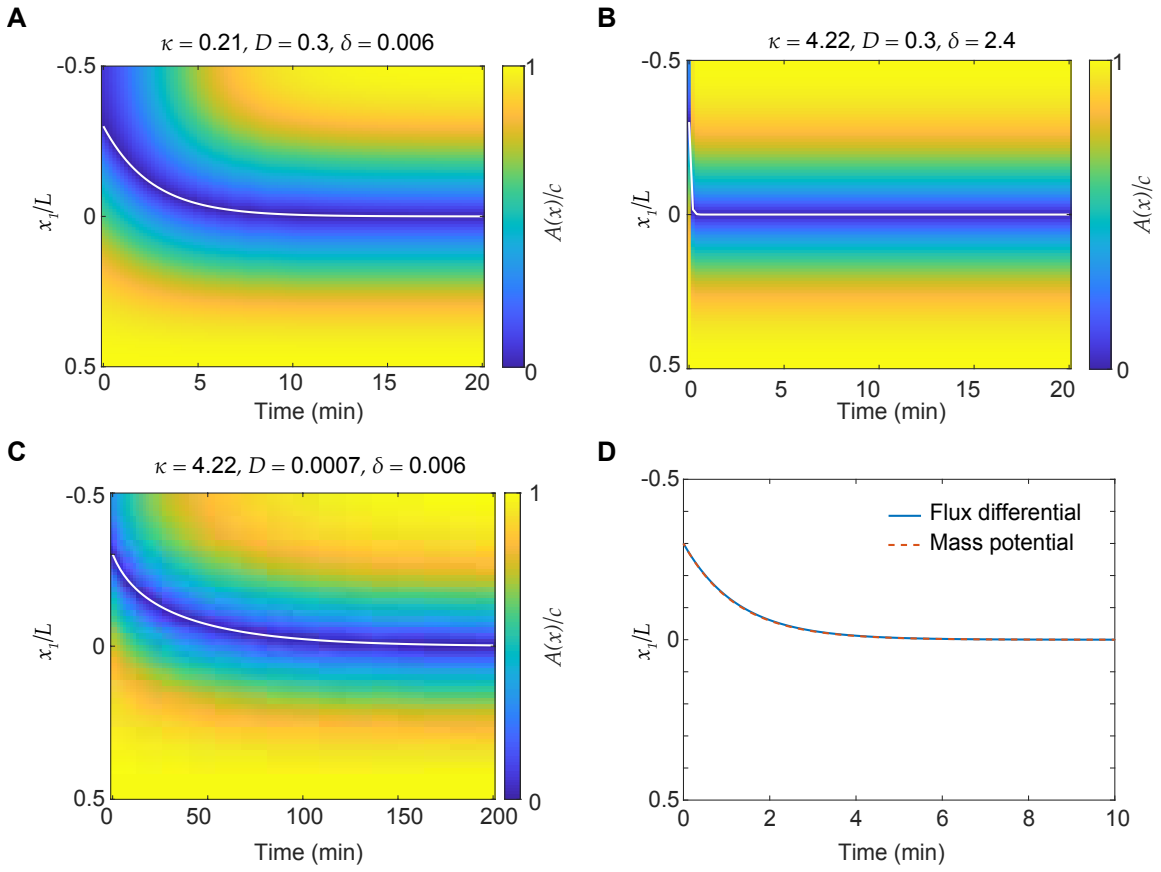

Figure 10: Point sink movement under changes in δ and D. (**A**). Figure 9B reproduced for comparison. (**B**). With all other parameters remaining the same as in (A) κ is increased by changing δ. The sink moves faster to the middle of the domain. This is because the flux through the system is increased. (**C**) Same as (**B**) but κ is changed via D instead of the delta. In this case, the higher value of κ leads to slower movement. This is because the lower diffusive length scale means that there is less flux exiting the system through the sink (and the total flux through the system ($c\delta L$) is unchanged). (**D**). A sample trajectory of the sink (with parameters of **A**) is plotted in blue using Eq. (3.25) (dynamics due to the flux-differential). Overlaid dashed lines are the result of dynamics using Eq. (3.26) in the main text, with the sink moving on a mass potential. Parameters: $\mu = 1, c = 1, \nu = 1$ across all panels. κ and corresponding changes in δ and D are indicated on top of each panel.

Dynamics of n sinks

We consider again the limit $\kappa \ll 1$. Then, the flux differential up to lowest order in κ is given by Eq. (3.17),

$$\frac{\Delta J_i}{c\delta L} = -\frac{n}{L}\frac{\lambda}{1+n\lambda}\left[x_i - \frac{L}{n}i + \frac{L}{2}\left(\frac{1}{n}+1\right)\right] + O(k^2), \tag{3.27}$$

where we have used the expression for the lowest order term of μ' from Eq. (3.16),

$$\mu'_0 = \frac{\lambda c}{1+n\lambda}.$$

The derivative of M up to second order in κ from Eq. (3.22),

$$\begin{aligned}\frac{D}{c\delta L}\frac{\partial M}{\partial x_i} &= \left(\frac{\lambda}{1+n\lambda}\right)^2 \frac{\partial}{\partial x_i}\sum_{i,j} G_2(x_i;x_j) + O(k^4), \\ &= \frac{2n}{L}\left(\frac{\lambda}{1+n\lambda}\right)^2 \left[x_i - \frac{L}{n}i + \frac{L}{2}(\frac{1}{n}+1)\right] + O(k^4)\end{aligned} \tag{3.28}$$

where we have used the second order expansion of $G(x_i;x_j)$ in κ (Eq. (3.22)). Taken together, the above equations imply that in the low mass flow limit $\delta \to 0$ ($\kappa \to 0$, $\lambda \to \infty$)

$$\Delta J_i = -\frac{1}{2} nD \frac{\partial M}{\partial x_i}\,. \tag{3.29}$$

Therefore in that limit, the dynamics of the point sinks are equivalently specified by

$$\frac{dx_1}{dt} = -\nu \frac{n}{2} D \frac{\partial M}{\partial x_i}\,. \tag{3.30}$$

Geometry sensing

Let us consider the case of a single sink, $n = 1$, in more detail. We focus on the regime $\kappa \ll 1$ in which the diffusive length-scale is much longer than the domain size. We expand in κ to find first

$$\frac{\mu'_1}{c} \approx \frac{\lambda}{\lambda+1} - \frac{\lambda^2}{\lambda+1}\left(\frac{x_1^2}{L^2} + \frac{1}{12}\right)\kappa^2 + O(\kappa^4)$$

and then

$$\begin{aligned}\frac{\Delta J_1}{c\delta L} &= -\frac{1}{2}\frac{\mu'_1}{c}\frac{\sinh(2\kappa\frac{x_1}{L})}{\sinh(\kappa)} \\ &\approx -\frac{\lambda}{\lambda+1}\frac{x_1}{L} + O(\kappa^2)\,.\end{aligned} \tag{3.31}$$

If $\kappa \ll 1$, then the flux differential across the sink depends linearly on its relative displacement from the mid-domain. For sufficiently strong sinks ($\lambda \gg 1$), the proportionality factor is linear in δ, just as we observed for the Turing system (Fig. 3B). As κ increases, the flux differential becomes inflected about $x_1 = 0$ (Fig. 9A). Heuristically, if the diffusive length

scale is much shorter than the domain size ($\kappa \gg 1$), then only particles initially created near the sink will fall into it. Therefore the flux differential is only non-zero close to the boundaries (or another sink) resulting in the breakdown of geometry sense. On the other hand, when the diffusive length scale is much longer than the domain size ($\kappa \ll 1$), particles can explore the entire domain before decaying and so the flux differential across the sink reflects its position on the domain, with the fluxes into the sink from either side balancing at mid-domain. The relevance of this interesting aspect of geometry sensing to Turing systems will be made clear later.

If we consider sinks moving on a much slower timescale than that of diffusion, we can use the steady-state solution for $A(x)$ given in equation Eq. (3.2) to solve the dynamic system Eq. (3.25). We find, as expected, that a single sink moves exponentially to the mid-domain (Fig. 9B). We also find that increasing δ, which increases the flux through the system, leads to faster sink movement (Fig. 10) reminiscent of the Turing system (Fig. 7B). On the other hand, if we decrease D, which decreases the diffusive length-scale without affecting the flux through the system, the sink moves more slowly towards the mid-domain (Fig. 10). We also considered the system with two sinks and confirmed that the steady-state solution consists of quarter-positioned sinks, the configuration that minimizes the total mass of A (Fig. 9D).

We can also use the steady-state solution of $A(x)$ to make explicit a correspondence between the two choices for the sink velocities. While the steady states of the two systems are identical, their dynamics are not in general the same. However, we found that when δ is small (or in terms of the dimensionless parameters: $\lambda \gg 1$, $\kappa \ll 1$), the two expressions become equivalent. This equivalence was apparent even for our default parameter set - the sink trajectories arising for either choice were almost identical (Fig. 9D).

4

MOVEMENT OF TURING PEAKS

In the previous chapter, we noticed a striking similarity between moving point-sinks (Fig. 9) and the movement of peaks in a Turing pattern (Fig. 3 and 7). It suggests that the movement and resultant steady-state positions of peaks in a Turing pattern may be due to a dependence of the peak velocity on the flux differential (of the fast species across a peak of the slow species) or due to the total mass of the fast species acting (approximately) as a potential energy surface. Note that we restrict ourselves to Turing patterns consisting only of interior peaks, as boundary peaks are not amenable to a point sink approximation. First, we introduce the following definition of the flux differential into the peak of a single-peak Turing pattern:

$$\Delta J_s(t) = D_u \frac{\int_{-L/2}^{L/2} \frac{\partial u(x,t)}{\partial x} v(x,t) dx}{\int_{-L/2}^{L/2} v(x,t) dx} \,. \tag{4.1}$$

If v is proportional to a Dirac delta function with the property $\lim_{\varepsilon \to 0^+} \int \delta_\varepsilon(x) f(x) dx = f(0)$, an approximation we will use below, this expression reduces to the flux differential of u, defined similarly to Eq. (3.5). We note that a similar expression has already been used to describe the flux into a spatial sink in the context of plasmid positioning [29]. We initialised the system with a single peak and monitored ΔJ_s as a function of the peak position and velocity. We found that similar to the point sinks (Fig. 9C), the flux differential is, away from the domain boundaries, directly proportional to the displacement from the mid-domain position (Fig. 11A). Thus, the peaks move with a velocity proportional to the flux differential.

However, it is not obvious how to extend the definition of the flux differential to patterns with multiple peaks as well as to higher dimensions in which Turing patterns can consist of complex structures such as stripes, spirals and hexagons. The concept of mass minimisation on the other hand is more tractable. To make an analytical comparison of the moving peaks to moving sinks we study the singular spike limit in the Turing system.

4.1 THE SINGULAR LIMIT - SPIKE APPROXIMATION

In the singular limit $D_v \ll D_u$, the peaks in v take the form of narrow spikes or pulses of width $\varepsilon = \mathcal{O}(\sqrt{D_v/\gamma})$ (Fig. S3). This limit allows the use of non-linear analysis methods to study the existence, stability and dynamics of Turing patterns (see [16] for a review). Away from the spike, v is approximately constant with a value v_{out} that is much smaller than u. Here, our goal is simply to derive an approximation for u in this limit by treating the spikes of v as Dirac delta functions as described below.

We look for steady-state solutions consisting of n spikes at positions $x_1, \cdots, x_n$. We assume that the substrate species u changes slowly within each spike and so can be approximated by a constant u_i and within each spike $u_i \ll v$. First, we introduce the inner coordinate (within each spike), $y_i = (x - x_i)/\varepsilon$. We then have the following system for the inner variable $v_i(y)$

$$\frac{D_v}{\varepsilon^2}\frac{d^2 v_i}{dy_i^2} + \beta u_i v_i^2 - (\gamma + \delta) v_i = 0$$
$$v_i \to 0 \quad \text{as} \quad y_i \to \pm\infty \,,$$

which gives

$$v_i = \frac{3}{2}\frac{\gamma+\delta}{\beta u_i}\text{sech}^2\left(\sqrt{\frac{\gamma+\delta}{D_v}}\frac{\varepsilon y_i}{2}\right) . \tag{4.2}$$

In the outer region, each spike is approximated by a weighted Dirac delta function and we, therefore, replace the v and uv^2 terms with Dirac delta functions with weights w_1 and w_2 given by

$$w_{i,1} = \varepsilon \int_{-\infty}^{\infty} v_i(y_i) dy_i = 6\frac{\sqrt{D_v(\gamma+\delta)}}{\beta u_i}$$
$$w_{i,2} = \varepsilon u_i \int_{-\infty}^{\infty} v_i^2(y_i) dy_i = 6\frac{\sqrt{D_v}(\gamma+\delta)^{3/2}}{\beta^2 u_i}$$

respectively, where we have use the expression for $v_i(y_i)$ from Eq. (4.2). Note that since $\mathcal{O}(w_{i,1}) = 1$ (each spike must have finite weight), we find that $\mathcal{O}(u_i) = \varepsilon$ and therefore $\mathcal{O}(v_i) = \varepsilon^{-1}$. Away from the spikes, v is taken to be a constant v_{out}. Therefore from Eq. (1.34b) we have

$$\beta u(u + v_{out})^2 - (\gamma + \delta) v_{out} = 0 \tag{4.3}$$

in the outer region. Given that u and v_{out} must both scale to leading order with a positive power of ε (due to the condition $\frac{1}{L}\int_{-L/2}^{L/2}(u+v)dx = c$), this equation implies that $\mathcal{O}(u^3) = \mathcal{O}(v_{out})$, i.e. in the spike limit $v_{out} \ll u$ in the outer region.

We obtain the outer equation for u by replacing the v terms at the spike with weighted delta functions:

$$D_u \frac{d^2u}{dx^2} - \sum_{i=1}^{n} \left[\beta(2u_i^2 w_{i,1} + w_{i,2}) - \gamma w_{i,1}\right] \delta(x - x_i) \\ -\delta v_{out} + c\delta - \delta u = 0 \,, \tag{4.4}$$

where we used Eq. (4.3) to simplify the outer contribution. We can neglect the u_i^2 term since $\mathcal{O}(u_i^2) = \varepsilon^2$ and the v_{out} term since $\mathcal{O}(v_{out}) < \mathcal{O}(u)$ to arrive at

$$D_u \frac{d^2u}{dx^2} + c\delta - \delta u - \sum_{i=1}^{n} \frac{\rho}{u} L\delta(x - x_i) = 0 \tag{4.5}$$

with $\rho = 6\frac{\sqrt{D_v}}{L}\frac{\delta\sqrt{\gamma+\delta}}{\beta}$. Note the inverse dependence on u in the point sink term (which we call an inverted sink). This form is also obtained for other Turing systems with a uv^2 non-linearity, such as the Schnakenberg and Brusselator models [25, 30] as we show in the next chapter and appendix respectively.

Following the approach of the previous section, the solution to Eq. (4.5) is given by

$$u(x) = c - \sum_i \rho_i' G(x; x_i) \,, \tag{4.6}$$

where the Green's function is defined as for point sinks Eq. (3.7) but in terms of the corresponding dimensionless parameter $\kappa = L\sqrt{\frac{\delta}{D_u}}$, the ratio of the length of the domain to the diffusive length scale of u (henceforth κ replaces b in the set of dimensionless parameters of the system). The coefficients $\rho_i' = \rho_i'(\boldsymbol{x})$ are now determined by the non-linear algebraic system

$$\rho_i' = \sigma \frac{c^2}{u(x_i)} \quad i = 1, \ldots, n \,. \tag{4.7}$$

where $\sigma = \frac{\rho}{c^2\delta} = 6\frac{\sqrt{b+1}}{a\sqrt{\Gamma}}$ is the second dimensionless parameter of Eq. (4.5). The inverse dependence on $u(x_i)$ makes solving this algebraic system challenging. For a general choice of sink positions x_i, there are n coupled quadratic equations in ρ_i', and therefore up to 2^n real solutions. However remarkably, this multiplicity of solutions collapses in the spike limit $\sigma \to 0$, in which the only physical solution is

$$\boxed{\boldsymbol{\rho}' = c\, \boldsymbol{G}^{-1}\hat{\boldsymbol{e}}} \qquad \textbf{Result V} \tag{4.8}$$

where $\boldsymbol{G}_{ij} = G(x_i; x_j)$ and $\hat{\boldsymbol{e}}$ is the column vector with all unit entries as taking the limit $\sigma \to 0$, equation Eq. (4.7) becomes $\rho_i'(c - \sum_j \rho_j' G(x_i; x_j)) = 0$. Since taking any $\rho_i' = 0$ gives a solution of the system without the ith spike, these are unphysical solutions.

This is precisely the same solution obtained in the *perfect sink limit* $\lambda \to \infty$ of the point sink system Eq. (3.2). Thus, in the singular spike limit, steady-state solutions of the Turing system are equivalent to that of a system of perfect sinks, a surprising equivalence is given the inverted pre-factor in Eq. (4.5).

4.2 SPIKE DYNAMICS

We consider the case of a single arbitrarily positioned spike. From Eq. (4.6) the solution is given by

$$u(x) = c - \rho_1' G(x; x_1). \tag{4.9}$$

where, for a given spike position x_1, from Eq. (4.7) there are two solutions for ρ_1'

$$\rho_{1\pm}' = \frac{c}{2G(x_1;x_1)}\left(1 \pm \sqrt{1 - 4\sigma G(x_1;x_1)}\right) \tag{4.10}$$

with

$$G(x_1;x_1) = \frac{\kappa}{2}\left[\frac{\cosh(\frac{2\kappa x_1}{L})}{\sinh(\kappa)} + \coth(\kappa)\right] \tag{4.11}$$

The flux differential across the spike is given by,

$$\begin{aligned}\Delta J_{1\pm} &= -\frac{D}{2}\rho_{1\pm}'\left[G_x(x_1^-;x_1) + G_x(x_1^+;x_1)\right] \\ &= -\frac{1}{2}\delta L \frac{\sinh(2\kappa x_1/L)}{\sinh(\kappa)}\rho_{1\pm}' \,.\end{aligned} \tag{4.12}$$

Non-dimensionalising and Taylor expanding around $\kappa = 0$ we obtain,

$$\frac{\Delta J_{1\pm}}{c\delta L} = -\frac{\left(1 \pm \sqrt{1-4\sigma}\right)}{2}\frac{x_1}{L} + \mathcal{O}(\kappa^2)\,. \tag{4.13}$$

Hence, in the limit $\kappa \ll 1$ we find a similar linear dependence of the flux-differential on the sink position x_1 as for the 'non-inverted' case discussed earlier.
In the limit $\sigma \to 0$, the non-trivial solution of ρ_1' is

$$\rho_1' = \frac{c}{G(x_1;x_1)}. \tag{4.14}$$

The derivative of the mass with respect to the spike position x_1 is,

$$\frac{\partial M}{\partial x_1} = \frac{\partial}{\partial x_1}(c - \rho_1') = \frac{c}{G(x_1;x_1)^2}\frac{\partial G(x_1;x_1)}{\partial x_1}$$
$$\implies \frac{\partial M}{\partial x_1} = \frac{c}{G(x_1;x_1)^2}\frac{\sinh(2\kappa x_1/L)}{\sinh(\kappa)}.$$

Then, in the limit $\sigma \to 0$ and $\kappa \ll 1$ the flux differential and mass derivative become,

$$\frac{\Delta J_1}{c\delta L} = -\frac{1}{2}\frac{\sinh(2\kappa x_1/L)}{\sinh(\kappa)}, \qquad \frac{D}{c\delta L}\frac{\partial M}{\partial x_1} = \frac{\sinh(2\kappa x_1/L)}{\sinh(\kappa)}.$$

Therefore, just like for point sinks, in this limit, the velocity ansatz for a single spike can be expressed equivalently as,

$$\frac{dx_1}{dt} = \nu\Delta J_1 = -\nu\frac{1}{2}D\frac{\partial M}{\partial x_1} \tag{4.15}$$

Note that $\sigma \to 0$ for spikes (inverted sinks) corresponds to $\lambda \to \infty$ for point sinks.

Dynamics of n spikes

Let us consider a solution to the Turing system consisting of n arbitrarily positioned spikes. From Eq. (4.6) and Eq. (4.7) we have,

$$\rho_i' = \sigma\frac{c^2}{c - \sum_j \rho_j' G(x_i;x_j)}, \tag{4.16}$$

which form a set of n algebraic equations dependent on the peak positions. As $\sigma \to 0$, the condition becomes

$$\rho_i'(c - \sum_j \rho_j' G(x_i;x_j)) = 0, \quad \forall i\ .$$

Since $\rho_i' = 0$ is a trivial solution (it corresponds to removing one spike from the system), we must have

$$c - \sum_j \rho_j' G(x_i;x_j) = 0, \quad \forall i\ .$$

Now taking $\kappa \ll 1$ we expand the above equation in powers of κ and collect terms of the lowest order. We have,

$$\rho_{0j}' = \frac{c}{n}, \tag{4.17}$$

where $\rho'_j = \rho'_{0j} + \rho'_{2j}\kappa^2 + O(\kappa^4)$ and $G(x_i; x_j) = G_0(x_i; x_j) + G_2(x_i; x_j)\kappa^2 + O(\kappa^4)$. Next collecting terms of order κ^2 we have,

$$c - \sum_j \rho'_{0j} G_2(x_i; x_j) - \sum_j \rho'_{2j} G_0(x_i; x_j) = 0. \tag{4.18}$$

We see that the above equation is similar to the case of point sinks in the limit $\lambda \gg 1$ and $\kappa \ll 1$. Following similar arguments to that case (with $M = c - \sum_j \rho'_j$ and expansion of M in powers of κ, $M = M_0 + M_2\kappa^2 + ...$) we find that the derivative of the mass with respect to x_i is given by

$$\frac{D}{c\delta L}\frac{\partial M}{\partial x_i} = \frac{2}{n}\frac{\partial}{\partial x_i}\sum_{i,j} G_2(x_i; x_j) + O(\kappa^4), \tag{4.19}$$

$$= \frac{2}{nL}\left[x_i - \frac{L}{n}i + \frac{L}{2}\left(\frac{1}{n} + 1\right)\right] + O(\kappa^4). \tag{4.20}$$

The flux differential across a spike is given by

$$\Delta J_i = -\frac{D}{\delta L}\sum_j \rho'_j \left[G_x(x_i^+; x_j) + \frac{\kappa^2}{2L}\delta_{ij}\right].$$

Expanding the above expression lowest order in κ and following the same steps as for point sinks, we find

$$\frac{\Delta J_i}{c\delta L} = -\frac{1}{L}\left[x_i - \frac{L}{n}i + \frac{L}{2}\left(\frac{1}{n} + 1\right)\right] + O(\kappa^2) \tag{4.21}$$

Hence, in the limit $\sigma \to 0$ and $\kappa \to 0$ the spike velocity can be written as,

$$\frac{dx_i}{dt} = \nu\Delta J_i = -\nu\frac{n}{2}D\frac{\partial M}{\partial x_i}. \tag{4.22}$$

We conclude that while the two definitions for the spike velocity (flux differential and mass potential) are different in general, they become identical in the limit $\sigma \to 0, \kappa \to 0$.

4.3 COMPARISON TO NUMERICAL SIMULATIONS

To compare the movement of peaks in the simulations with the analytical calculations above, we initialise the peaks at different positions (by translation of the steady state pattern) and monitor the evolution of the system. In Fig. 11A (single peak), we calculate the mass of the fast species, $M(t) = \int_{-L/2}^{L/2} u(x,t)dx$, and the flux on the peak ΔJ_s Eq. (4.1) at each time step

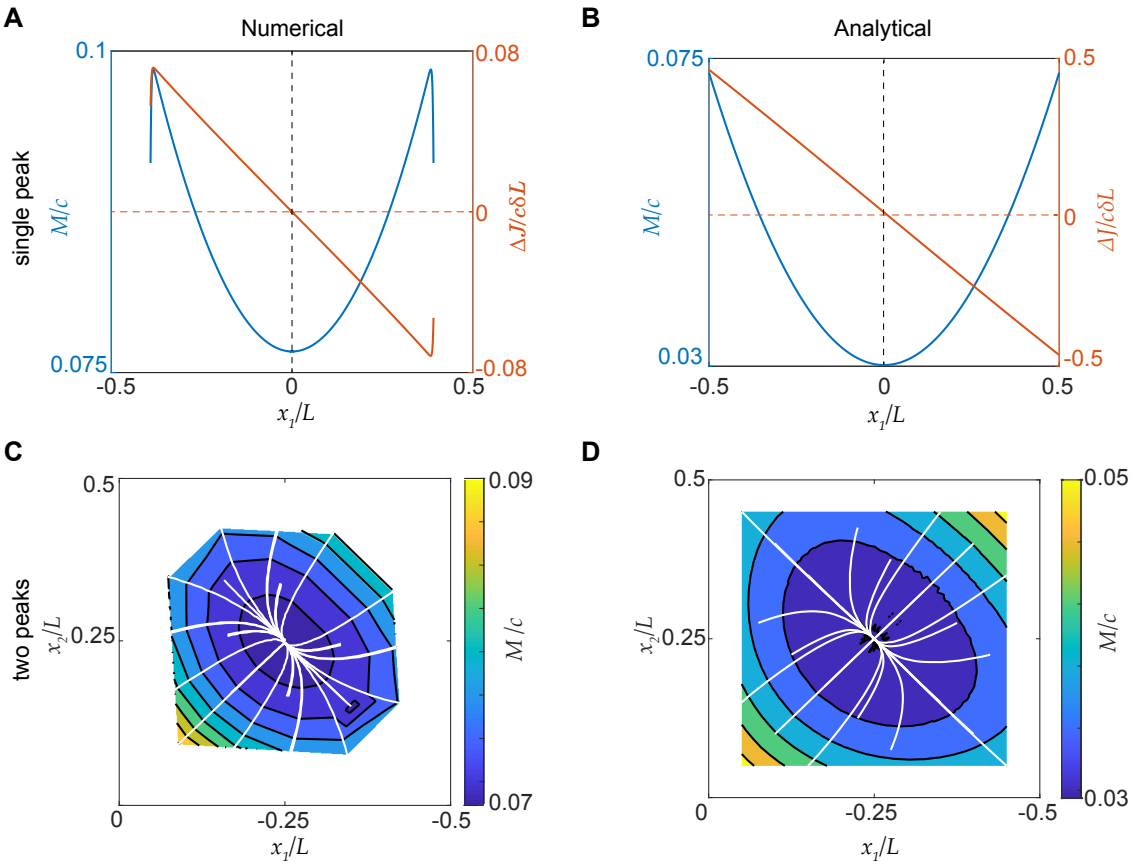

Figure 11: The mass of u is minimised at regular positions. (**A**) Flux differential measured numerically using equation Eq. (4.1) for a single spike (orange) is a linear function of the peak position. The mass of the fast species M (blue) is minimised at the mid-domain. See also Fig. S3 (**B**) The same quantities as in (**A**) but for the analytical expressions from the spike approximation (Eq. (4.13) and $M = c - \hat{\rho}'_{1+}$). (**D**) Mass minimisation for two peaks. Trajectories of two peaks as they move towards opposite quarter positions (white lines). The contours and colour bar represent the mass M interpolated from trajectories. The mass is minimised for regular positioning. (**D**) Same as (**C**) but trajectories obtained from the approximation of peaks as spikes using equation Eq. (4.23). Parameters: $L = 2$ in (**A**) and (**B**), $L = 4$ in (**C**) and (**D**); $D_v = 0.0012$, otherwise default. This gives $\sigma = 0.0146$ in (**B**) and $\sigma = 0.0073$ in (**D**)

as the peak approaches mid-domain. We find that the flux across a single peak is consistent with our numerical observations (Fig. 11B) and just as we found for the non-inverted sinks in the previous section (Fig. 9C). Furthermore in the spike limit, from Eq. (4.13) we find that $\sigma \to 0$, $\Delta J_1 = c\delta x_1$ is linear in δ, consistent with our numerical observations (Fig. 7B). We also find that $M_1 = c - \rho'_+$, the total mass of u, is minimised at mid-domain (Fig. 11B). Note that we only ever observe spikes with large amplitudes, i.e. patterns in which almost all the mass of the system is contained within spikes. So we assume that the low amplitude solution is unphysical for finite σ and not just in the spike limit $\sigma \to 0$ (or possibly unstable in the context of the time-dependent system). However, while the mass and flux differential

displayed very similar qualitative profiles (Fig. 11B), the agreement was not quantitative. In particular, the flux differentials disagree by an order of magnitude. This is likely because the analytic flux differential is defined at the interface region between the inner and outer solutions, which is the precise location where the approximation is not accurate.

To understand the dynamics of two peaks, we generated the contours of M by simulating the Turing system with 200 different initial peak positions (x_1, x_2) (some of which are overlayed in white) and interpolated the mass of u from the trajectories (Fig. 11C). It is important to note that it is not clear how to define the flux differential across each peak in the case of multiple peaks. Similarly, in Fig. 11D, we solved the nonlinear algebraic system in Eq. (4.7) for different sink positions (x_1, x_2). Of the 4 different solutions, we obtain we choose the one with the lowest total mass M. The similarity between the numerical observation (Fig. 11C), the spike approximation in Fig. 11D and the system of point sinks in the previous section (Fig. 9D) is apparent. In all cases, the steady-state solution consists of quarter position peaks/spikes that minimise the mass of u.

Hence, the observation that peaks in a Turing pattern move with a velocity proportional to the flux differential across them (Fig. 7A, B, Fig. 11A, Fig. S3) suggest that the spike approximation can be extended to account for spike movement by specifying the spike velocities as

$$\frac{dx_i}{dt} = \nu \Delta J_i(\boldsymbol{x}) \,, \tag{4.23}$$

where ν is some unknown parameter. By the correspondence with point sinks, this expression becomes mathematically equivalent in the spike limit $\sigma \to 0$ and low mass-flow regime $\kappa \to 0$ (or equivalently $\delta \to 0$) to a description in which the mass M acts as potential and the sinks as over-damped particles as in Eq. (3.26). Together with Eq. (4.7), Eq. (4.23) defines a differential-algebraic system for the dynamics of n spikes. However, not all spike configurations are stable. Based on our numerical simulations, stable solutions consist only of regularly positioned symmetric spike solutions with peaks of the height ($\rho_i' = \rho'$), just as observed for the full Turing system itself. Consistent with this, the flux differentials of these solutions vanish, $\Delta J_i(\bar{\boldsymbol{x}}) = 0$, via the properties of the Green's function, just as for the points sinks of the previous section.

4.4 SUMMARY

To summarise the results so far, we have shown that the movement and positioning of peaks in a Turing pattern are akin to that of a system of moving point sinks. Firstly, the regular, periodic steady-state positions are a result of the flow (creation, diffusion, decay) of mass through the system, and not by some dominant linear mode. The steady-state configuration is the one for which all the flux differentials balance and this is also the configuration that minimises the total mass of the fast species. We found empirically that the movement of

peaks (in the slow species) is well described by the peak velocity being proportional to the flux differential of the fast species across it. Furthermore, in the spike limit, this is equivalent to the total mass of the fast species acting as a potential through which the peaks move as over-damped particles. While this does not imply that the mass or some other function acts as a potential away from this limit (i.e. that $\Delta \boldsymbol{J}$ is a conservative vector field in general), and indeed the trajectories specified by such a relationship are less consistent with the numerical results, the mass is nonetheless minimal at the steady-state configurations (Fig. 4A, C). In the next section, we will see that we can use the steady-state mass to compare the 'energy' of patterns with different numbers of peaks and in this way predict the preferred *number* of peaks at steady state, and not just their positions.

5

MASS MINIMISATION AND PATTERN SELECTION

In this chapter, we address the question of pattern selection in two-component reaction-diffusion systems, i.e starting from homogeneous initial conditions which select the number of peaks (in 1-D) in the final steady state pattern. We have seen that in the mass-conserved limit $\delta \to 0$, the model exhibits a complete coarsening effect in which the only stable patterns consist of a single peak positioned somewhere in the interior of the domain (Fig. 7D) or, on a shorter domain, an interface. As previously observed, the model exhibits incomplete coarsening for small δ . With our default parameters (with $L = 4, \delta = 0.014$), linear stability predicts that mode $n = 8$ (a pattern with four peaks) will dominate (Fig. 3A). While this is true initially, as the pattern evolves, it subsequently coarsens so that we most frequently obtain two peaks (dominated by mode $n = 4$)(Fig. 3B, C, 4D). We never observe a pattern with four peaks as expected by linear stability analysis. This coarsening effect is also referred to as competition instability and has previously been studied in the context of spike solutions [25–27]. To our knowledge, there is currently no general guiding principle to

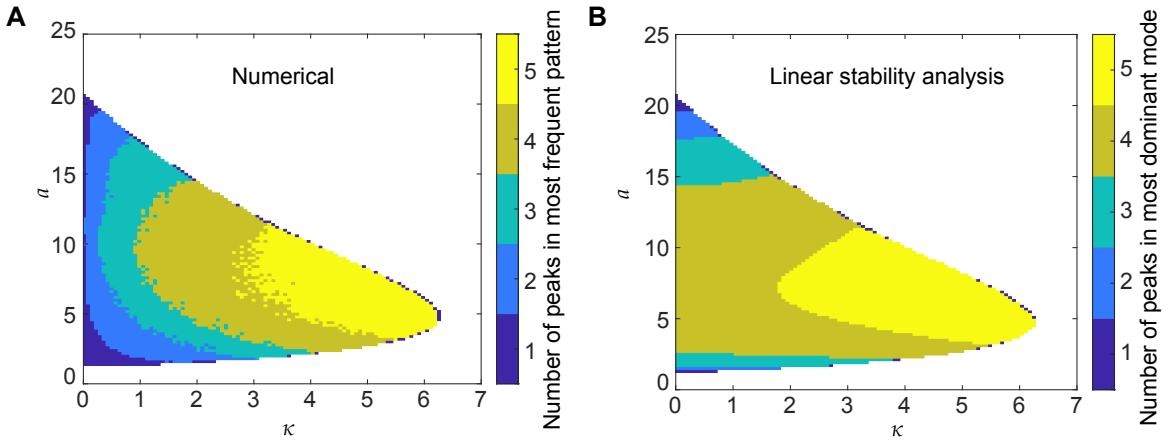

Figure 12: Numerical simulations are inconsistent with linear stability analysis. (**A**) The number of peaks in the most frequent steady-state pattern is plotted as a function of a and κ. Each point represents the most frequent pattern steady-state pattern from 5 simulations. (**B**) The number of peaks in the fastest-growing mode predicted by linear stability analysis is plotted as a function of a and κ.

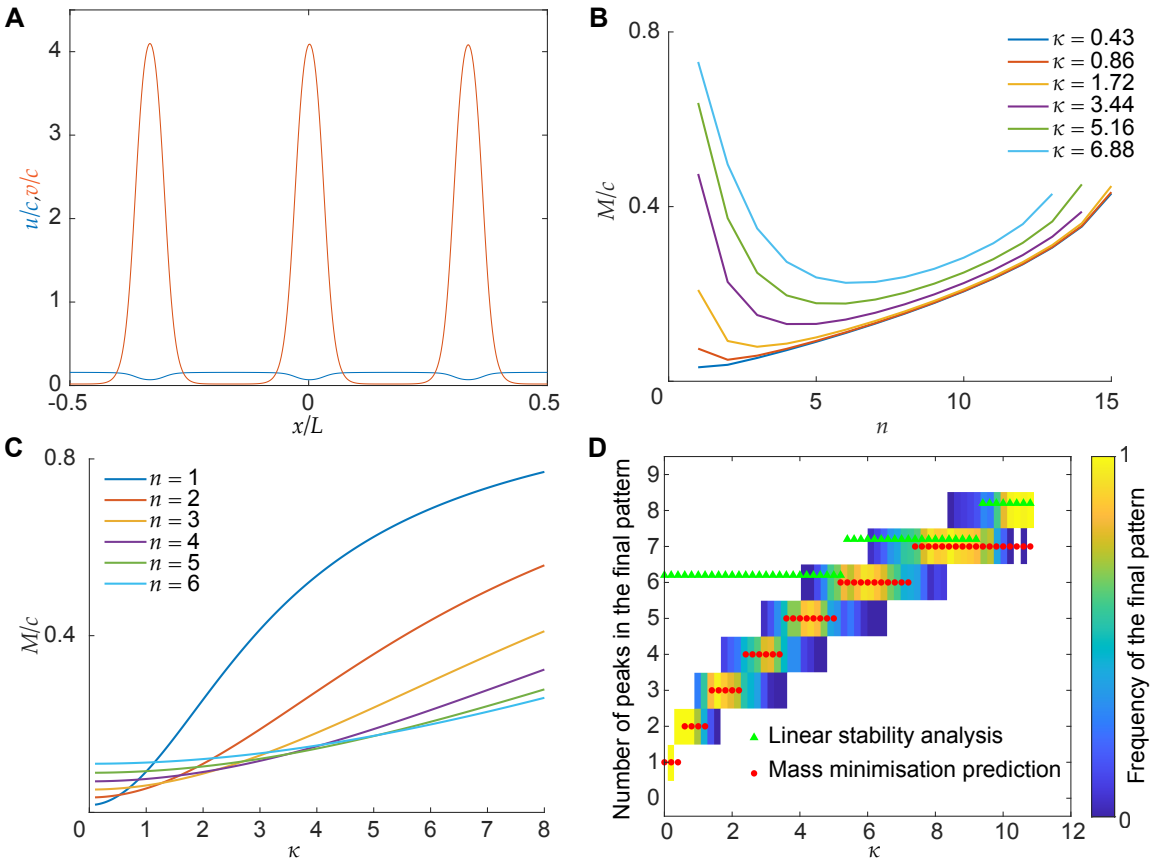

Figure 13: Mass minimisation predicts peak number in the illustrative model. (**A**) An example steady state solution in the spike limit. (**B**) Normalised total mass $M/c = 1 - n\rho'_+/c$ plotted as a function of n. There exists a critical n for which the mass is minimal.(**C**) Normalised total mass $M/c = 1 - n\rho'_+/c$ plotted as a function of κ. As κ increases patterns with a higher number of peaks have a lower mass. (**D**) The numerically obtained distribution of peak number at steady state for different values of κ (colour scale) overlaid with the prediction of the dominant pattern from linear stability (green triangles) and the prediction from mass minimisation (red circles). Mass minimisation correctly predicts the number of peaks at a steady state. Data from 50 simulations for each parameter set. Parameters: Default values as in Figure 3 with $L = 4$ except (**C**)-(**E**) which use $D_v = 0.006$ (to make peaks narrower).

determine which pattern is finally obtained at a steady state (especially in the presence of coarsening). It is important to note that this is a much narrower question than asking which patterns are stable (tells you which patterns are allowed for a particular set of parameters) since Turing systems like our model are generally multi-stable.

It is clear that flow rate δ plays a role in coarsening. We measured the distribution of steady-state patterns obtained for different values of δ (through the dimensionless parameter $\kappa = L\sqrt{\frac{\delta}{D_u}}$) and compared against the prediction of linear instability (Fig 12A, B). We used periodic boundary conditions to avoid peaks on the boundary that are not amenable to the spike approximation. We found that for $\kappa \gtrsim 1$ linearly stability analysis correctly predicts

the dominant mode at steady-state. However, for $\kappa \lesssim 1$, a coarsening process occurs and the steady-state pattern is dominated by a lower mode than that predicted. Our previous observations on the role of the diffusive length scale κ can help explain this disagreement. When the diffusive length scale is longer than the domain size $\kappa \lesssim 1$, all peaks v compete for u molecules created across the domain (the peaks feel each other). Whereas, when the length-scale is short $\kappa \gtrsim 1$, peaks only absorb molecules of u created within a distance given by the diffusive length-scale and therefore compete less or not at all. Also important to note that decreasing δ also decreases the total flux through the system ($c\delta L$) resulting in more competition.

5.1 Minimal mass and peak number selection in our model

We next studied our illustrative model in the spike limit developed in the previous chapter 4.1. We decreased D_v from the default value so that the obtained pattern was reasonably spike-like (Fig. 13A) while at the same time not resulting in a very much enlarged Turing space (since we want to sweep over different values of δ). We considered only symmetric, regularly positioned spike solutions, which are the only observed steady-state solutions numerically. This implies,

$$u(x_i) = c - \sum_j \rho'_j G(x_i; x_j) = c - \rho' \sum_j G(x_i; x_j).$$

We showed in Eq. (3.11) that

$$\sum_j G(x_i; x_j) = \frac{\kappa}{2} \coth(\frac{\kappa}{2n}).$$

As in the case of the point sinks the flux differential across the spikes vanishes for the regularly positioned configuration $\Delta J_i(\bar{x}) = 0$. We then have

$$\rho' = \frac{\sigma c^2}{(c - \rho' \frac{\kappa}{2} \coth \frac{\kappa}{2n})}. \tag{5.1}$$

Solving for ρ', we find two solutions

$$\rho'_{\pm} = \frac{c}{\kappa \coth(\frac{\kappa}{2n})} \left[1 \pm \sqrt{1 - 2\sigma\kappa \coth(\frac{\kappa}{2n})}\right] \tag{5.2}$$

with corresponding total masses

$$M_{\pm} = \frac{1}{L} \int_{-\frac{L}{2}}^{\frac{L}{2}} u dx = c - n\rho'_{\pm}. \tag{5.3}$$

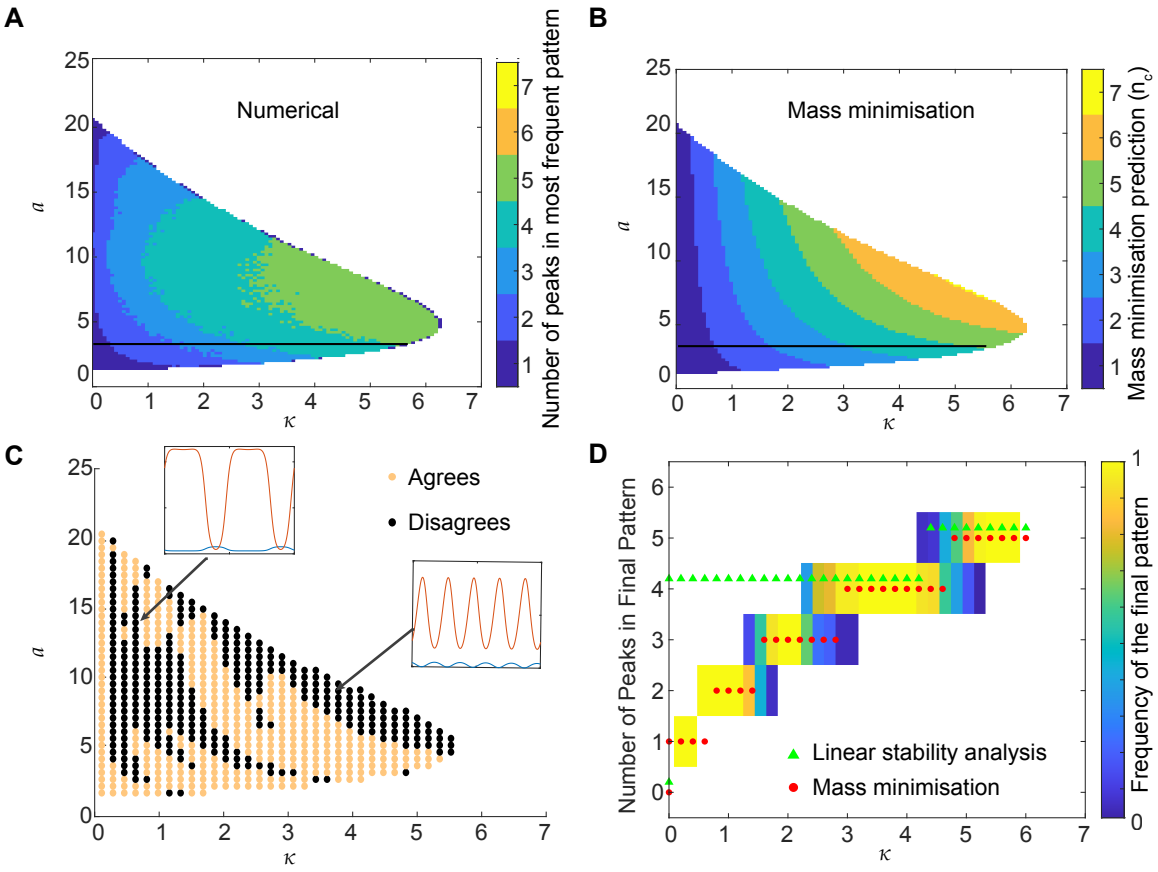

Figure 14: Mass minimisation is surprisingly accurate in the non-spiky regime. (**A**). Figure 12A is reproduced with a changed colorbar for comparison with the predictions from mass minimisation. (**B**). The number of peaks in the pattern with the lowest the total mass M is plotted as a function of a and κ. (**C**). We perform a direct comparison of results between numerical simulations in (**A**) and predictions in (**B**). The prediction is surprisingly accurate (beige dots) for large regions of the Turing space. Though, it is inaccurate (black dots) when $\kappa \gg 1$ or $a \gg 1$, as patterns have a baseline and peaks are now much broader respectively (see inset). (**D**). Same as Figure 13D in the main text but for default parameters (peak like solutions). Mass minimisation correctly predicts the number of peaks at steady state, even though it is less accurate than the spike case. Data from 50 simulations with random perturbations from the homogeneous state for each parameter set. Parameters: Default with $L = 4$. For (**D**) we sweep across κ (black line in (**A**) and (**B**)) for the default value of $a = 3.75$.

As previously discussed, for n spikes, we obtain two possible values of ρ', of which we take the larger, ρ'_+ (the other corresponds to extremely weak spikes i.e. $\rho'_- \approx 0$). This gives a solution $u(x) = c - \rho'_+ \sum_i G(x; \bar{x}_i)$ with mass

$$\boxed{M/c = 1 - n\frac{1 + \sqrt{1 - 2\kappa\sigma \coth(\frac{\kappa}{2n})}}{\kappa \coth(\frac{\kappa}{2n})}}. \tag{5.4}$$

Note that for a real solution we must have $1 > 2\kappa\sigma \coth(\frac{\kappa}{2n})$. Therefore, for a given choice of parameters, there is an upper bound on the number of spikes that a solution can contain. In general, a solution exists for multiple values of n. However, numerically, we observe a very narrow distribution of the number of peaks (Fig. 4, Fig. 13D). We hypothesised that mass minimisation might play a role. Indeed, when we examined the mass M of solutions consisting of different numbers of spikes at their respective steady-state positions, we found that the mass is minimal for a specific number of spikes (Fig. 13D). This could also be easily seen by plotting the mass as a function of κ for different values of n (Fig. 13C). The value of n at the minimum decreases with κ, with a single spike being minimal at $\kappa \to 0$. The curves invert so that as κ is increased multiple spikes produce the lowest mass. Given that we have already shown that the mass of $u(x)$ is minimal at the steady state, we hypothesised that it could also be used to compare solutions with different numbers of peaks and therefore identify a preferred 'minimum energy state.

We compared the number of spikes predicted by this mass minimisation principle (the n that minimises the mass in Eq. (5.4)) against the distribution of patterns obtained numerically (starting from a small random perturbation around the uniform state). We found remarkable agreement (Fig. 13D, red circles). Mass minimisation correctly predicts the most frequent pattern obtained over the entire range of κ, including, most importantly, the regime in which coarsening occurs. There is deviation only at the transition points and close to exiting the Turing regime at high κ. In contrast, as expected, the linear prediction only agrees for the highest values of κ, i.e. close to onset (Fig. 13D, green triangles). Remarkably, the prediction was also reasonably accurate even when the solution is much broader and not spike-like, as for our default parameter set (Fig. 14). Thus, mass minimisation not only explains where the peaks of a Turing pattern are positioned but also *how many* peaks there will be at a steady state after any coarsening has taken place. Importantly, it does so far from the onset and hence outside the region where weakly non-linear approaches such as the amplitude equations are valid.

Numerics with periodic boundary conditions

where we use periodic boundary conditions,

$$u|_{x=-L/2} = u|_{x=L/2}, \quad u'|_{x=-L/2} = u'|_{x=L/2},$$
$$v|_{x=-L/2} = v|_{x=L/2}, \quad v'|_{x=-L/2} = v'|_{x=L/2}.$$

As the *pdepe* does not support such boundary conditions, we used the package *Periodic Reaction-Diffusion PDE solver* [1]. The initial conditions were taken to be random perturba-

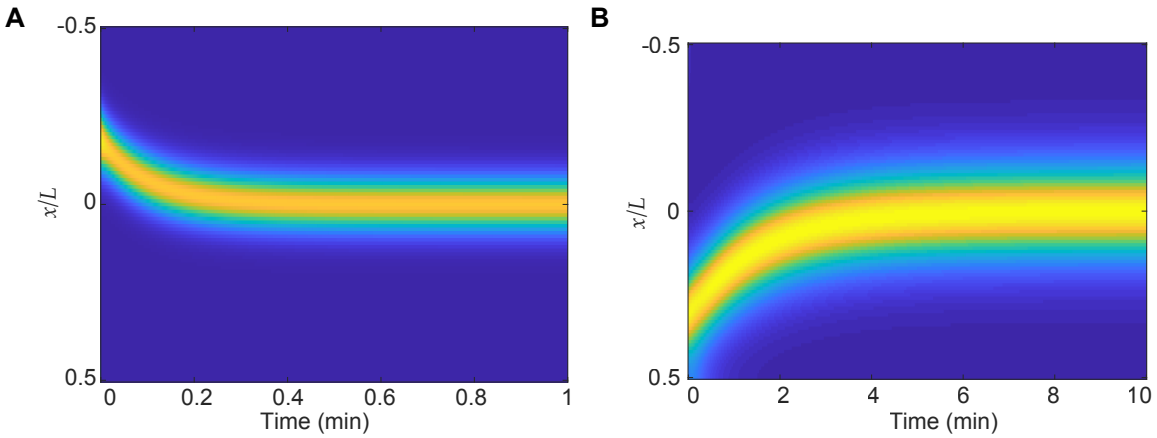

Figure 15: Peak movement in other reaction-diffusion system.
(**A**). Brusselator model. Kymograph of a single peak pattern moving exponentially to the mid-domain. Parameters: $a = 2, b = 0.001, \Gamma = 1500, d = 100$. (**B**). Kymograph of a single peak pattern in Schnakenberg model.

tions around the homogeneous steady state (drawn from a normal distribution with a standard deviation of 1%).

5.2 MASS MINIMISATION IN OTHER RD MODELS

To investigate the generality of these results, we performed a similar analysis of the Brusselator and the Schnakenberg model. Upon numerically simulating these systems with reflective boundary conditions, we also observed a similar movement of peaks to regular positions as was observed in our model (15 A, B). This indicated that turnover might affect peak movement in these systems as well.

The Brusselator model also has the form of a mass-conserved model with additional production and decay terms and a dimensionless parameter b characterising the mass flow through the system (with $b = 0$ being the mass-conserved limit) Eq. (A.1). Following the same approach as above, we derived an expression for the total mass, M, of the fast species of symmetric n-spike solutions

$$\boxed{M = \frac{6n(b+1)^{3/2}}{a\sqrt{\Gamma}} + \frac{\Gamma b}{12d(b+1)n^2}.} \tag{5.5}$$

This mass is again minimised for a particular number $n = n_c$ of spikes and we found this minimum to be an excellent predictor of the final pattern obtained after coarsening (Fig. 17A). Note that the lack of the background decay term in Eq. (A.7) and Eq. (A.14) means

that unlike for Eq. (4.5) and Eq. (3.1), there is no concept of a diffusive length-scale in the Brusselator, or rather, the diffusive length-scale is infinite as decay occurs only through the point sinks.

In our illustrative model Eq. (1.34), the parameter δ (or the dimensionless parameter κ) controlled both the flow of mass through the systems (the limit to a mass conserved model) and the diffusive length-scale. While coarsening increases as the system approach the mass-conserved limit, it only occurs when the diffusion length-scale is longer (κ smaller) than some critical threshold, as this is the point at which peaks begin to compete with each other for the substrate i.e. u. In the Brusselator, since the diffusive length-scale is infinite, peaks always compete and as a result coarsening is observed at all values of δ (compare to Fig. 17A with Figure 13D).

Finally, we performed a similar analysis on the Schnakenberg model. Unlike the previous two, this model does not have a mass-conserved Turing system as a limit. Likely as a result of this, we find no evidence of dynamic coarsening instability. Nevertheless, the final steady-state pattern is often not precisely predicted by linear stability. Applying the spike approximation, we found that the outer equation of this model, and hence the expression for the mass of the fast species, has the same form as for the Brusselator (Eq. (A.20)). On comparing the number of peaks predicted by minimizing the mass and by linear stability analysis, we found that while the two approaches were in broad agreement, mass-minimisation was the better predictor of the obtained pattern.

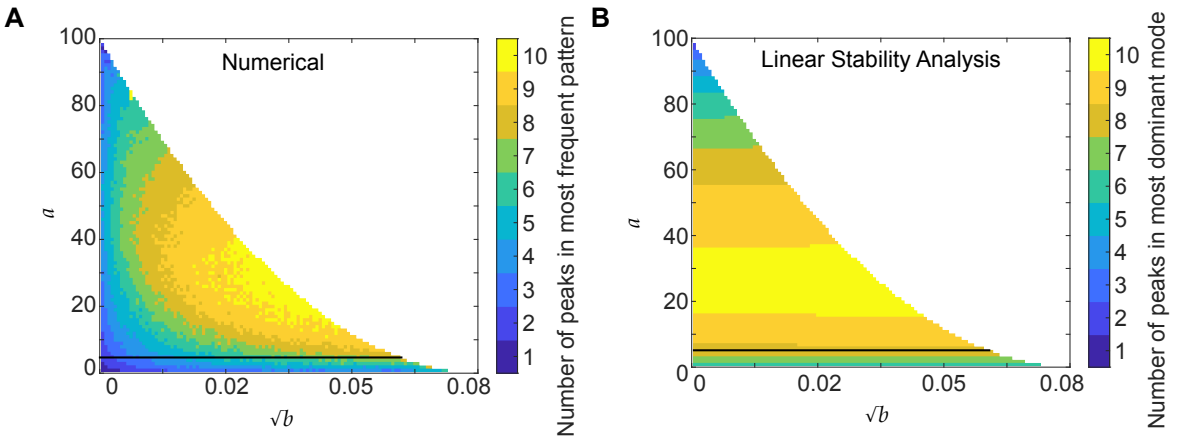

Figure 16: Turing space of Brusselator model. (**A**). The number of peaks in the most frequent steady-state pattern of the Brusselator model is plotted as a function of model parameters a and $\sqrt{b}$ (see Appendix A, B). For each set of parameters, the most frequent pattern was obtained from 10 simulations each initialised with a different random perturbation from the uniform state. The simulations were run for long enough to ensure the steady-state pattern was reached. (**B**). The number of peaks in the mode with the greatest growth rate as predicted by linear stability analysis is plotted as a function of a and $\sqrt{b}$. Coarsening was observed over much of the range of $\sqrt{b}$ and the disparity with the linear stability prediction increase as $b \to 0$. Parameters: $d = 100$ with $D_u = 1, D_v = 0.01$ and $\Gamma = 15000$.

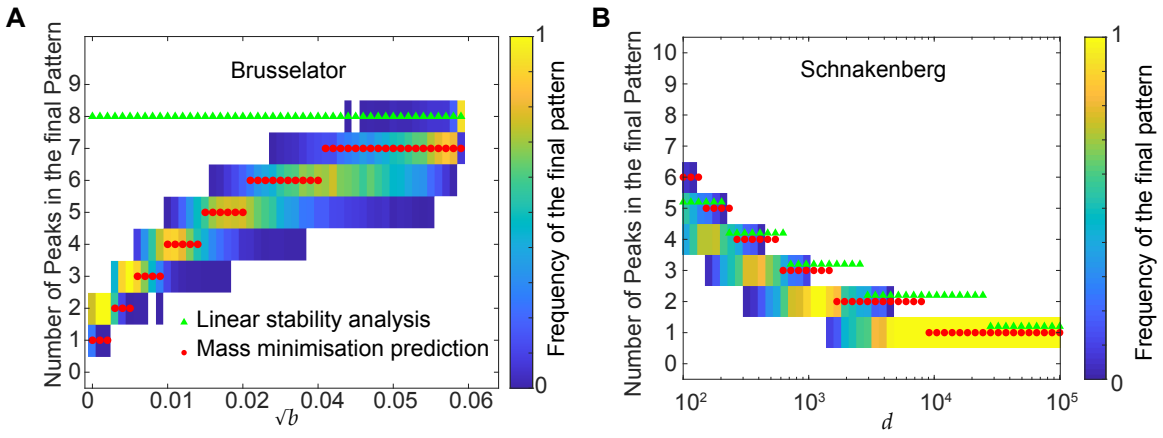

Figure 17: Mass minimisation is good predictor of the final steady state pattern in other models. (**A**) As in Fig. 13D but for the Brusselator model. The numerically obtained distribution of the number of peaks at steady state for different values of $\sqrt{b}$ (colour scale) is overlaid with the prediction of the dominant pattern from linear stability (green triangles) and the prediction from minimisation of the mass of the fast species Eq. (5.5) (red circles). Mass minimisation again correctly predicts the number of peaks at a steady state. Data from 500 simulations for each parameter set. Black line in Figure 16(**A**) and (**B**) shows the range of $\sqrt{b}$ values used within the $(a, \sqrt{b})$ Turing space. Parameters: $d = 100, a = 3.75$, $\Gamma = 15000$. (**B**) Mass minimisation in the Schnakenberg model. Same as A and 13**D** but for the Schnakenberg model. We see that in the limit of well-separated spikes ($d \to \infty$) mass minimization (red circles) is a better predictor of the number of peaks in the final steady state pattern than linear stability analysis (green triangles). Parameters. $a = 1.1, \Gamma = 9000$. Data from 50 simulations with random perturbations from the homogeneous state for each parameter set.

While we can not exclude the possibility that mass-minimization is only predictive for patterns consisting of well-separated peaks as studied here, the consistency of the results over three different systems and surprising agreement even in the non-spiky regime nonetheless hints at a fundamental property of Turing systems and warrants further study.

6

MASS MINIMISATION CAPTURES THE FINITE SIZE EFFECTS FAR FROM ONSET

In this chapter, we briefly discuss the effect of changing domain size on the selection of patterns. As discussed previously, the wavelength of Turing patterns depends on the boundary conditions and the domain size. Below we discuss the appearance of Finite-size effects in our reaction-diffusion system for small values of δ. We show that at low flow, mass minimisation predicts the number of peaks in the final pattern, even in the presence of Finite size effects. Note that this regime is far from the onset, so linear stability analysis fails to predict the steady state. This provides further credence to mass minimisation as a general principle guiding pattern formation in reaction-diffusion systems.

6.1 EFFECTS OF CHANGING DOMAIN SIZE ON STEADY STATE

In a reaction-diffusion, we expect the number of peaks in a pattern to scale linearly with the domain size (the number of peaks doubles with domain length doubling) [1]. At small length-scales, where the wavelength of the pattern is smaller than the size of the domain, this linear scaling breaks down. The effect of discreteness of the lattice is called *Finite size effects*. In the simulations of our Turing model Eq. (1.34), we find that for high values of δ we have a clear linearity between increasing domain length L and the number of peaks in the most frequent final pattern (Fig. 30A, purple line). Interestingly, the domain length L_n at which a 1 peak pattern transitions to a 2 peak pattern increases with decreasing δ. Note that the further transition in peak numbers $(2-3, 3-4, ...)$ scale linearly with increasing L. We wondered if this effect is captured by mass minimisation.

In agreement with the numerical simulations, we find that for the smallest value of $\delta = 0.005$ in Figure 18A, mass minimisation predicts that the number of peaks in the minimal pattern does not double with domain size doubling (Fig. 18B). The mass of the fast species as a function n has a minima at a value much smaller than an integer (Fig. 18B). Hence, the pattern with the minimal mass is predicted to have a *single* peak. As we double the domain size (from $L = 2$ to $L = 4$) we find that the minima is still below $n = 1$ resulting in a single peak prediction for the steady state pattern. Finally, we found good

agreement between the transition lengths L_n observed in the numerical simulations starting from homogeneous conditions and L_m predicted by mass minimisation for different values of δ (Fig. 18C (blue and red dots), D).

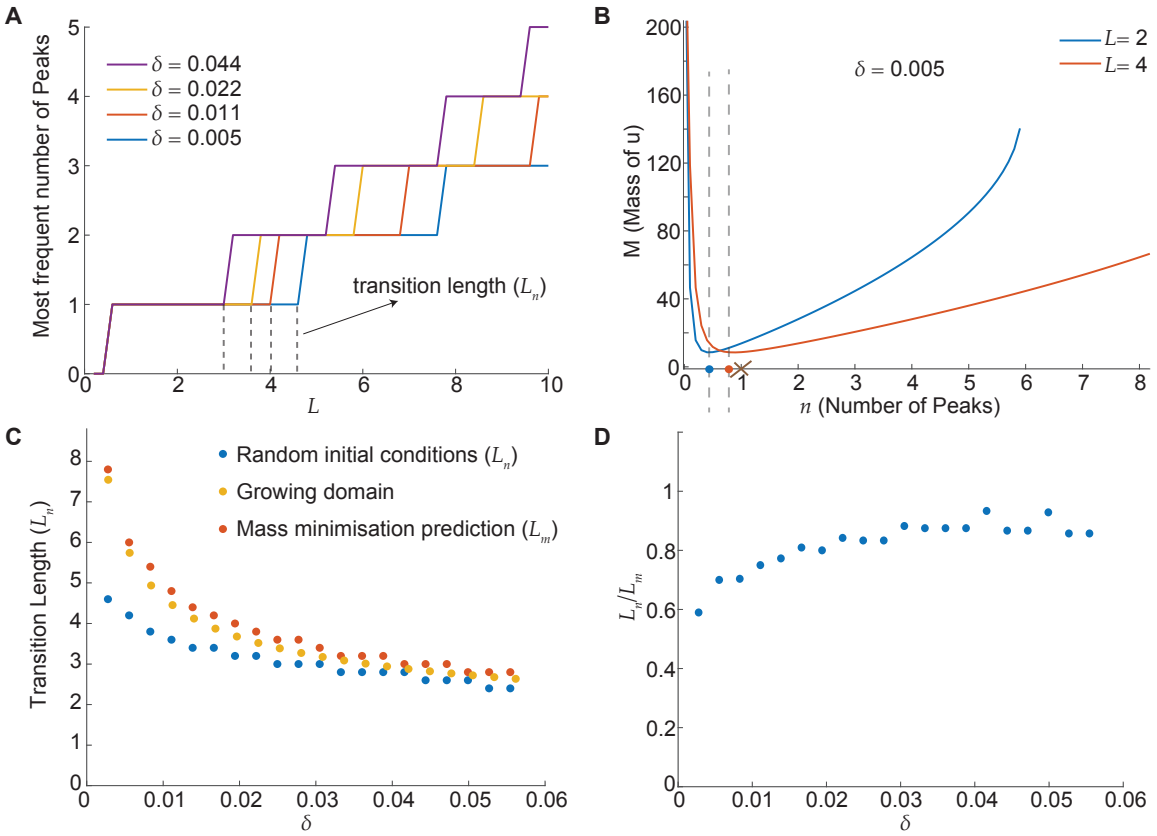

Figure 18: Mass minimisation measures the finite size effects far from the onset. (**A**) The number of peaks in the most frequent final pattern is plotted as function domain length L. As an indication transition from 1 peak to 2 peaks does not scale with domain doubling. As δ is decreased, a single peak doubles at longer lengths (transition length L_n). (**B**) Mass of fast species u plotted as a function of peak number n. For small values of $\delta \ll 1$ doubling the domain size, L does not double the number of peaks n_c due to finite size effects. (**C**) Transition length (L_n) from different methods plotted as function of flow rate δ. Transition length decreases with increasing δ and asymptotically approaches $L = 2$ as expected in the absence of Finite size effects. Also, predictions of transition length from 1 to 2 peaks from mass minimisation L_m match reasonably well with numerical simulations starting from random initial conditions L_m and with on growing domain L_g. (**D**) The ratio of data from (**C**) L_n/L_m shows that mass minimisation is accurate for finite δ. Parameters: Default parameters other than $D_u = 1, D_v = 0.01$.

6.2 COARSENING ON GROWING DOMAINS

In Turing's original hypothesis he envisioned that morphogenesis could occur via two independent processes where reaction-diffusion equations establish spatial patterns on a fast

timescale, while the developmental fate of these patterns is cued by the growth of the cell. This decoupling of timescales allowed us to study pattern formation independently of cell growth. Here, we discuss pattern formation in our RD system Eq. (1.34) on a *slowly* growing domain [31] to illustrate the effect of coarsening in pattern selection. Here we show that mass minimisation also captures the finite size effect in the presence of a growing domain.

To numerically simulate the reaction-diffusion equations on a growing domain, we use the following conservation equation for the evolution of the spatial pattern on Ω_t,

$$\frac{d}{dt}\int_{\Omega_t} c(x,t)dx = \int_{\Omega_t}\left[\frac{\partial c}{\partial t} + \frac{d}{dx}(pc)\right]dx \tag{6.1}$$

where $c = \begin{bmatrix} u \\ v \end{bmatrix}$,$p(x,t)$ is an velocity associated with volumetric growth. Hence, evolution equations for the components in 1-D become,

$$\frac{\partial c}{\partial t} + \frac{d}{dx}(pc) = D\frac{\partial^2 c}{\partial x^2} + R(c) \tag{6.2}$$

where $R(c) = \begin{bmatrix} f(u,v) \\ g(u,v) \end{bmatrix}$ and $D = \begin{bmatrix} d \\ 1 \end{bmatrix}$. The velocity is determined by the local rate of growth,

$$p(x,t) = \frac{\partial x}{\partial t} \tag{6.3}$$

Consider isotropic growth of a 1-D domain is given by,

$$x(t) = Xr(t), \quad r(0) = 1,$$

where X is an initial position marker on the domain, and $r(t)$ is the growth function. The flow is then determined by,

$$p(x,t) = X\dot{r} = x\frac{\dot{r}}{r}\,.$$

Hence, Eq. (6.2) becomes,

$$\frac{\partial c}{\partial t} + \frac{\dot{r}}{r}\left[x\frac{\partial c}{\partial x} + c\right] = D\frac{\partial^2 c}{\partial x^2} + R(c)$$

We transform our system to the unit domain,

$$x' = \frac{x}{L(t)}, t' = t, \tag{6.4}$$

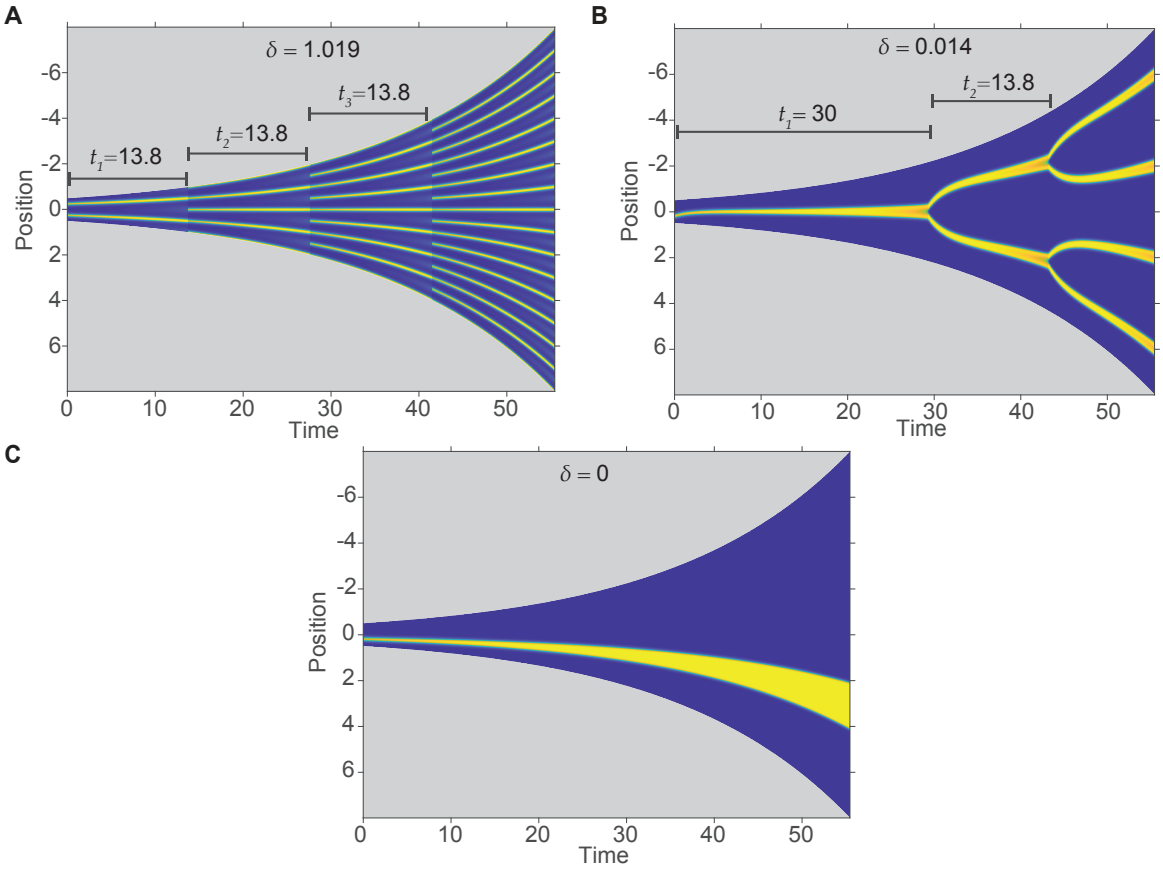

Figure 19: Scaling in a growing domain. A slowly growing domain with a growth rate of $g = 0.05$ is simulated. (**A**) In presence of high flow we are close to onset ($\delta = 1.019$) the patterns double with domain doubling. (**B**) In the presence of low flow ($\delta = 0.014$), the pattern splits eventually. We see that the time for a single peak to split (t_1) is greater than for 2 peaks to split into 4 (t_2), indicating finite size effects for small domain sizes. We call this the quasi-scale invariant regime. (**C**) A single peak is inserted with no flow ($\delta = 0$). The pattern width scales with domain size and does not split as $t \to \infty$.

where the domain length is $L(t) = L_0 r(t)$. We note that,

$$\frac{\partial c}{\partial t'} = \frac{\partial c}{\partial t} + x\frac{\dot{r}}{r}\frac{\partial c}{\partial x}. \tag{6.5}$$

Since r is independent of position the transformation eliminates the advection term. Dropping the primes leads to the transformed equation,

$$\frac{\partial c}{\partial t} = \frac{1}{L(t)^2} D\frac{\partial^2 c}{\partial x^2} + R(c). \tag{6.6}$$

where we have neglected the decay term as for small $\dot{r}$, $c\frac{\dot{r}}{r}$ is much smaller than the linear terms in $R(c)$.

To study the effect of coarsening on a growing domain, we simulated the RD system from our model equations using an exponentially growing domain with $L(t) = L_0 \exp(gt)$ (from Eq. (1.34) and Eq. (6.6)) where we set $g = 0.05$ as the growth rate (we use reflective boundary conditions). Hence, the domain length doubles every $t_d = \frac{\ln(2)}{g} = 13.86$. In the presence of high flow $\delta = 1.109$ (close to the bifurcation point), we find that the number of peaks double exactly with domain size doubling $t_{split} \approx 13.8$ (Fig. 19A). In the absence of flow $\delta = 0$, we find that the initial pattern with a single peak stays locked in its position and its width scales with the length of the domain, resulting in a broad mesa peak at steady state (Fig. 19C). This is also referred to as scale invariance [32], which is observed in many biological systems [33–35]. Our model, offers a simple way of generating scale invariance which has wider consequences in modelling biological systems.

Finally, for our default value of $\delta = 0.014$, we find that the pattern splits as the domain grows and the peaks are at regular positions. We observe finite size effects for smaller domain sizes as the peak doubling from 1 to 2 takes much longer $t_1 = 29$ than the successive doubling from 2 to 4 peaks with $t_2 = 13.8$ (Fig. 19B). Note that doubling time $t_2 \approx t_d \ll t_1$. Similar to the comparisons in the previous section, we wondered whether these Finite size effects observed in the low flow $\delta \ll 1$ regime are explained by mass minimisation. Hence, we performed the growth simulations with periodic boundary conditions. Indeed, we find a striking agreement between the split length L_m predicted by mass minimisation and the split length L_g from the growth simulations (Fig. 19C). This agreement was even better as compared to the numerical simulations starting from random initial conditions. We speculate that mass minimisation might capture the depth of the basin of attraction of the final steady state pattern better as compared to its width.

7

DISCUSSION AND FUTURE DIRECTIONS

> "A pattern exists in some data is *real*, if there is a description of the data that is more efficient than the bit map, whether or not anyone can concoct it."
>
> — Daniel Dennett, *Real Patterns* (The Journal of Philosophy, 1991)

One of the main challenges in understanding pattern-forming Turing systems is the prediction of which pattern will be obtained, not only at onset, i.e. at entry into the parameter space giving patterns but generically for any parameter values. While the dominant mode from linear stability analysis is a good predictor in some regimes, non-linear effects mean that it can be inaccurate on the entire parameter space. Furthermore, as we have argued, linear stability analysis alone cannot explain the periodic nature of final patterns nor the regular positioning of peaks within the domain (with reflective boundary conditions), which occurs dynamically in several settings that are outside of the linear regime (e.g. domain growth, coarsening, initialised peaks). Weakly non-linear approaches, such as the method of amplitude equations, do exist but they are much less useful far from the onset where we still lack a general theory of pattern dynamics and selection.

Here, we have presented evidence that the flow of mass through the system is responsible for the movement and regular positioning of peaks in a Turing pattern. We also showed that the number and positions of peaks at a steady state are such that the mass of the fast species is minimised. This simple principle correctly predicted the preferred steady-state pattern for both our model and the Brusselator even far from the onset and in the presence of coarsening and for the Schnakenberg model, in the well-separated spike limit. We expect that this will guide the development of new non-linear approaches for the study of pattern selection, far from the onset.

The insight came from analysing the behaviour of a diffusive system consisting of point sinks that move with a velocity proportional to the gradient. We showed that the flow of mass through such a system leads to sinks being positioned symmetrically and evenly across the domain (regularly positioned), as this is the unique configuration for which the gradient

across each sink vanishes. We also showed that this configuration uniquely minimises the total mass of the fast species in the system. Consistent with this, we showed that in the low mass-flow regime, the total mass acts as a potential and the sinks as over-damped particles.

We found that the peaks of the Turing system also move toward the regularly positioned configuration with the same dynamics as the moving point sinks and at a rate directly proportional to the rate of mass flow through the system. In doing so, the system minimises the total mass of the fast species u. In the singular limit $D_v \to 0$, in which the peaks of the Turing pattern become narrow point-like spikes [16], an analytical approximation showed that u is indeed described by diffusion in the presence of point sinks but where the Dirac delta function terms have a $1/u$ pre-factor. Nevertheless, in the singular limit, solutions are that of a system of perfect sinks Eq. (4.8). However, for $D_v > 0$ the 'inverted' sink term ($1/u$ coefficient) leads to the total mass of u having a non-trivial dependence on the number of spikes and the rate of mass flow through the system. As a result, there is a critical number of spikes (and hence wavelength) that minimises the mass of u. We, therefore, hypothesised that the mass of u, the fast species, could act, at least approximately, as a multi-well potential i.e. that it could be used to assign, energy to patterns with different numbers of peaks and thereby identify the preferred steady state. In particular, we asked whether it could predict the steady-state pattern selected after the coarsening that occurs in our model. We found that this 'mass minimisation' principle could indeed predict, almost perfectly, the obtained patterns (Fig. 13D) and we confirmed this for another Turing system, the Brusselator (Fig. 17A). It also gave a better prediction of the dominant pattern for the Schnakenberg model, which does not exhibit coarsening (Fig. 17B). Importantly, we are predicting which pattern is selected (most frequently observed) at a steady state and not which patterns are stable. While the latter has already been studied for spike and mesa solutions in the Brusselator (see [16] for a review), to our knowledge much less is known about the former.

The two models with coarsening studied here have a parameter δ controlling the flow of mass through the system and the degree of coarsening. In the mass-conserved limit $\delta = 0$, they display complete coarsening down to a single peak, a seemingly generic property of two-component mass-conserved reaction-diffusion systems that have been proved explicitly for several systems [17–20, 36–38]. This complete coarsening is correctly predicted by mass minimisation in both models for $\delta = 0$ with the mass of the fast species being minimal for only a *single peak*.

We speculate that complete coarsening (the only solution is a single or a half peak on the domain) is due to the absence of mass flow through the system rather than mass-conservation *per se*. Consistent with this, the feed and decay terms in our model and the Brusselator can be replaced with linear couplings to a sufficiently well-mixed third species, i.e. we can embed the open system inside a larger closed system. The result is a three-component mass-conserved model with the same pattern-forming behaviour, including partial (incomplete) coarsening. Indeed, this was the precise form of the previously investigated model [21].

While the entire three-component system is mass-conserved, there is still mass flow through the 'subsystem' of the original two variables (u and v), the rate of which controls the degree of coarsening.

Finally, we have shown that mass minimisation captures the Finite-size effects in reaction-diffusion systems far from the onset. For small values of flow δ, the linearity between domain length doubling and peak doubling break down due to the discreteness of the patterns. We find that the mass minimisation captures the length at which the peaks split both starting from random initial conditions and for patterns on a growing domain.

GENERALITY OF RESULTS

We have shown analytically (in the spiking regime) and numerically that mass minimisation of the fast species predicts the preferred steady state patterns for our illustrative model, the Brusselator and Schnakenberg model. The outer equation for u obtained in the singular limit $D_v \ll D_u$ Eq. (4.5), has the same form in all of the substrate-depletion type models such as the Brusselator (Appendix A), Schnakenberg (Appendix B) and the Gray-Scott model [26,39]. Indeed all these models exhibit the same peak movement towards the regularly positioned configuration (Fig. 15). Substrate-inhibition models that have peaks of the two species overlapping, such as that of Gierer and Meinhardt [40], also exhibit peak movement towards regular positions. However, the outer equation of these models has a point source term rather than a point sink (positive δ-term) [41]. The validity of mass minimisation in such systems remains to be tested.

Do our results hold for regimes outside the well-separated spike limit? While it is currently unclear and testing it would require an analytical description of other solution types (mesas, spots, stripes etc). This is beyond the scope of the current work. However, the following observation suggests to us that mass minimisation of the fast species may have some deeper connection to pattern formation. If we assume the final pattern is selected according to minimising the mass of the fast species, then that mass of the fast species must be minimisable. We can therefore ask what would happen in a model in which the mass of the fast species is a fixed constant at a steady state.

Consider the following class of systems:

$$\begin{aligned} \partial_t u &= D_u \partial_x^2 u - f(u,v) + a - g(u) \\ \partial_t v &= D_v \partial_x^2 v + f(u,v) + b \end{aligned} \tag{7.1}$$

with $D_v < D_u$ and reflective or periodic boundary conditions. Note that at steady state the mass of u is fixed by a measure $g(u)$. We denote the stable fixed point of the homogeneous system by (u_0, v_0). The functions $f(u,v)$ and $g(u)$ are arbitrary apart from the constraint

that $g'(u_0) > 0$. The integral $\int g(\cdot)dx$ then acts as a measure for u and all steady-state solutions $(\bar{u}, \bar{v})$ must have the same mass of u using this measure.

$$\frac{1}{L}\int_{-L/2}^{L/2} g(\bar{u})dx = a + b.$$

The Jacobian at the homogeneous fixed point is given by,

$$J = \begin{bmatrix} -f_u - g_u & -f_v \\ f_u & f_v \end{bmatrix}, \tag{7.2}$$

where f_u, f_v and g_u are evaluated at the (u_0, v_0). If the fixed point is stable we must have

$$\begin{aligned} TrJ &= -f_u - g_u + f_v < 0 \\ DetJ &= -g_u f_v > 0\,. \end{aligned}$$

The latter relation implies that $f_v < 0$. A necessary condition for Turing instability is that

$$-\frac{D_v}{D_u}(f_u + g_u) + f_v > 0. \tag{7.3}$$

However, since $D_v < D_u$, this condition can never be satisfied. Hence, the general system Eq. (7.1) does not admit a Turing instability, consistent with a general role for minimisation of the fast species in Turing pattern formation. Note that there is no such restriction if the decay term is placed only in the equation for v (like for the Brusselator Eq. (A.1) and Schnakenberg models Eq. (B.1)) or in both equations (as in Eq. (1.34) and the Gray-Scott model). Thus, for this general class of models, a Turing instability requires that the mass of the fast species must not be generically a fixed constant at a steady state. Equivalently stated, systems in which mass leaves the system through only the fast species cannot exhibit a Turing instability. While u being minimisable is a fundamental requirement to apply the mass minimisation principle its connection to the existence of patterns is unclear. We speculate that for some classes of models spatial instability of homogeneous state is related to the existence of a stable spatial solution for which the mass of the fast species is lower. As expected by the mass minimisation principle, we numerically find that the steady-state mass of the fast species u both in our model and the Brusselator is lower than its initial mass (u_0). Importantly, this does not depend on a specific parameter regime or solution type. This is an interesting insight that deserves further investigation as it might mean that mass minimisation is not just a useful tool to study pattern selection but is fundamental to pattern formation in reaction-diffusion systems.

Part II

EFFECTS OF DNA BRIDGING ON BACTERIAL CHROMOSOMAL DYNAMICS AND ORGANISATION

8

INTRODUCTION TO BRIDGING DYNAMICS

> "It can scarcely be denied that the supreme goal of all theory is to make the irreducible basic elements as simple and as few as possible without having to surrender the adequate representation of a single datum of experience."
>
> — Albert Einstein (*On the Method of Theoretical Physics* lecture, 1933)

Chromosomes in all organisms are compacted nearly a 1000 times over to fit within cells [42]. Importantly, the chromosome must be packaged in a specific way to be consistent with DNA-related processes like replication, transcription, repair and segregation. Particularly, this problem becomes more complicated inside bacterial cells where replication and segregation occur concomitantly, rather than being temporally separated like in eukaryotes. Recently, microscopy-based approaches have provided greater insight into the temporal dynamics and spatial organisation of the bacterial chromosome. In concert with traditional genetics and biophysical modelling new techniques continue to reveal remarkable insights into the mechanisms involved in chromosome organisation and dynamics.

Time-lapse microscopy has provided insights into the temporal dynamics of chromosomal loci. In slow growing *E. coli* the *ori* (origin of replication) resides near the mid-cell with two chromosomal arms on either side and the *ter* (terminus) localised around the mid-cell in a so-called left-*ori*-right-*ter* configuration [43]. Upon DNA replication and segregation, the *ori*'s move to the quarter positions in the cell and *ter* moves from the poles to mid-cell. The structural maintenance of chromosomes (SMC) complex, MukBEF in *E. coli*, is required to maintain the mid-cell positioning of *ori*, while cells lacking mukB adopt a more polar *ori-ter* configuration [44]. It is not clear whether MukB anchors *ori* proximal regions to mid-cell or the non-polar regions. There is evidence however that MukBEF does indeed associate with *ori*-proximal regions on the chromosome [44, 45], while there is no clear evidence to suggest that SMC associates with the cell membrane. Recent experiments and modelling seem to suggest that SMC promotes a traverse left-*ori*-right pattern by anchoring

origin-proximal regions to mid-cell [46]. While MukBEF is implicated in *ori* positioning, the mechanism of MukBEF positioning was previously unknown. Recently, it was hypothesized that MukBEF forms self-organised clusters positioned at quarter positions of a replicating cell [21]. MukBEF also seems to play a role in the organisation of *ter* macrodomain, as it is excluded at *ter* due to the presence of a protein MatP [47].

On shorter time scales, the diffusive dynamics of chromosomal loci have also been characterised *in vivo* by measuring the scaling exponent α of the mean square displacement $\mathrm{MSD}(\tau) = \langle (\boldsymbol{r}(t+\tau) - \boldsymbol{r}(t))^2 \rangle \sim \tau^{\alpha}$. However, while polymer theory predicts a sub-diffusive scaling exponent of $\alpha = 2\nu/(2\nu+1) \approx 0.54$ for a self-avoiding Rouse polymer ($\nu \approx 0.588$), and $\alpha = 2/3$ for a Zimm polymer in a good or theta solvent, irrespective of chain length and topology [48,49], values measured by tracking experiments are consistently less than these estimates across different species, strains and conditions [50–55]. Fractional Brownian motion (fBm) of chromosomal loci due to the viscoelastic nature of the cytoplasm has been proposed as a possible explanation for this deviation [50,56]. However, this model cannot reproduce the rare but ubiquitous rapid chromosomal movements (RCMs) made by loci [57] and its predictions are inconsistent with a recent study in which compression of the cell was found to only affect the exponent of chromosomal loci and not cytosolic particles [53]. Other mechanisms are therefore required to explain the observed low sub-diffusive exponent.

In this work, we focus on the role of Nucleoid Associated Proteins (NAPs) in determining the chromosomal loci dynamics. NAPs are DNA-binding proteins that condense and organise the bacterial nucleoid through bridging, bending and stiffening the DNA [58–60]. Imaging of fluorescently tagged NAPs which bind the DNA non-specifically, has revealed an interesting twisty longitudinal structure of the nucleoid [61]. Recent work using high-throughput chromosome conformation capture (HiC) has investigated how these proteins affect the contact probability between any two chromosomal loci, with different NAPs found to promote either short or long-range contacts [62]. The obtained two-point contact probabilities have also been used in polymer models to specify an attractive potential between monomers and thereby make predictions about the organisation of the chromosome within the cell [63–66]. However, a bottom-up study of the effect of DNA bridging on bacterial chromosome organisation and dynamics has yet to be performed.

Previous theoretical studies of bridging-like behaviour have been mostly in the context of the networks and gels formed by associative polymers [67–69]. More recently, at the other extreme of high bridge density, computational models have explored how the resulting globular state can explain the organisation of eukaryotic chromatin [70–77]. However, there has been no detailed study of the dynamics induced by bridging in the context of the coiled (non-globular) bacterial chromosome.

Here, in the absence of a dynamical theory, we use polymer simulations to investigate how DNA bridging affects the organisation and dynamics (scaling exponent) of the bacterial chromosome [78]. We confirm that bridging can reduce the scaling exponent of individual

monomers below the classic prediction of polymer theory and we characterise the dependence on both the number of bridges and their lifetime, observing a linear relationship between the number of bridges and the compaction of the polymer. Consistent with these results, we show experimentally in *E. coli* that deleting the NAP H-NS results in an increase in the scaling exponent compared to the wild type. We also find that bridging produces monomer dynamics that display the same rare, rapid movements (RCMs) as have been observed experimentally, i.e. movements inconsistent with fBm. Finally, we use the experimentally observed mobility of loci to fix an internal timescale in our simulations and thereby predict a lower bound for the average bridge lifetime.

Publication: Subramanian, S., Murray, S.M. (2022). DNA bridging explains sub-diffusive movement of chromosomal loci in bacteria. bioRxiv. https://doi.org/10.1101/2022.11.18.517049

8.1 PHYSICS OF POLYMERS

In this chapter, we focus on the underlying theory of polymer dynamics and the implementation of bridging in our simulations. We model a segment of DNA as a self-avoiding linear chain on a periodic cubic lattice (Fig. 20). Before we discuss our simulation framework, we describe a few simple illustrative polymer models in brief. Polymers are random fractals that display self-similar structures at finite length scales [48, 49, 79, 80]. Let us consider a polymer chain made of $N+1$ beads, with no interactions between distant beads (*ideal chain*). The bond vector r_i connects bead a_{i-1} to bead a_i. The *end-to-end vector* is then the sum of all N bond vectors on the chain,

$$\boldsymbol{R_N} = \sum_{i=1}^{N} \boldsymbol{r_i} \tag{8.1}$$

For an ideal chain, it can be shown that the average end-to-end vector is zero,

$$\langle \boldsymbol{R_N} \rangle = 0, \tag{8.2}$$

where the angled brackets denote the ensemble average across different conformations of the polymer. Since there is no preferred direction of bond orientation in this model, the average end-to-end vector is zero. The more interesting, simplest non-zero and often measured

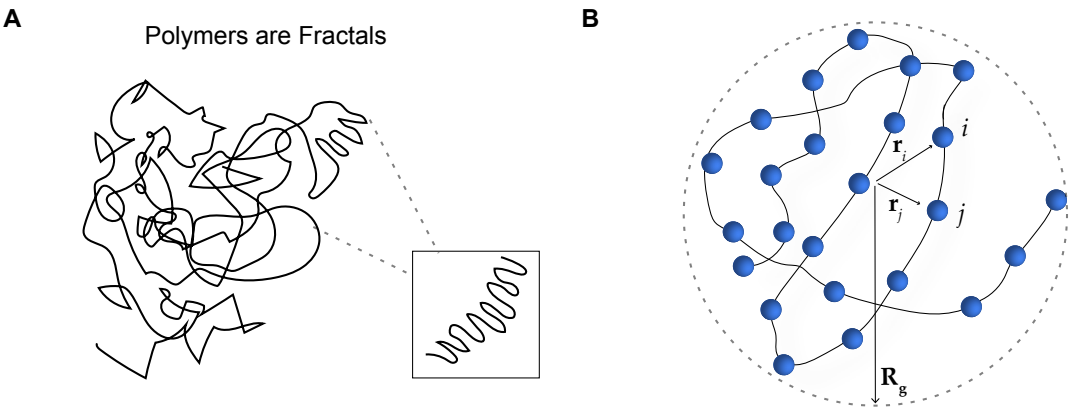

Figure 20: (**A**) Sketch of single polymer. Zoom-in shows the self-similar fractal structure at finite length scales. (**B**) Bead on string model of a polymer, where monomers are represented by blue spheres. R_g represents the radius of gyration of the polymer defined in Eq. (8.5).

quantity to characterize the size of the polymer chain is the mean-squared end-to-end distance,

$$\langle R^2 \rangle \equiv \langle \boldsymbol{R}_N^2 \rangle = \langle \boldsymbol{R}_N \cdot \boldsymbol{R}_N \rangle = \left\langle \left(\sum_{i=1}^{N} \boldsymbol{r_i} \right) \cdot \left(\sum_{j=1}^{N} \boldsymbol{r_i} \right) \right\rangle$$
$$= \sum_{i=1}^{N} \sum_{j=1}^{N} \langle \boldsymbol{r_i} \cdot \boldsymbol{r_j} \rangle.$$

If we consider bonds of equal length $b = \mid \boldsymbol{r_i} \mid$, the dot product can be written as,

$$\boldsymbol{r_i} \cdot \boldsymbol{r_j} = b^2 \cos\theta_{ij}.$$

Hence, the above quantity becomes,

$$\langle R^2 \rangle = b^2 \sum_{i=1}^{N} \sum_{j=1}^{N} \langle \cos\theta_{ij} \rangle.$$

We now consider the freely jointed chain approximation, where the polymer has no correlations between different bond vectors with $\langle \cos\theta_{ij} \rangle = 0$ for $i \neq j$ and the non-zero terms in the double sum for $i = j$ is $\cos\theta_{ij} = 1$. Hence, we get,

$$\langle R^2 \rangle = Nb^2. \tag{8.3}$$

Ideal chains

We present a few important results without derivation [80]. A more general measure of the size to characterize sizes of all kinds of polymers (branched or rings) is the square *radius of gyration* defined as the averaged squared distance between monomers in a given conformation,

$$R_g^2 \equiv \frac{1}{N}\sum_{i=1}^{N}(\boldsymbol{R_i} - \boldsymbol{R_{cm}})^2, \tag{8.4}$$

where R_{cm} is the position vector of the centre of mass,

$$\boldsymbol{R_{cm}} \equiv \frac{1}{N}\sum_{j=1}^{N}\boldsymbol{R}_j. \tag{8.5}$$

For an ideal linear chain, we can show that the *radius of gyration* is related to *end-to-end distance* as,

$$\langle R_g^2 \rangle = \frac{Nl^2}{6} = \frac{\langle R^2 \rangle}{6}. \tag{8.6}$$

The probability distribution of the end-to-end vector of an ideal chain is described by a Gaussian:

$$P_{3D}(N, \boldsymbol{R}) = \left(\frac{3}{2\pi N b^2}\right) \exp\left(-\frac{3\boldsymbol{R^2}}{2Nb^2}\right). \tag{8.7}$$

The free energy of an ideal chain is purely entropic and changes quadratically with the end-to-end vector:

$$F = \frac{3}{2}k_B T \frac{\boldsymbol{R}^2}{Nb^2}. \tag{8.8}$$

Real polymers

In a real polymer, we cannot ignore the interactions between different monomers along the chain. The conformations of a real polymer with repulsive self-interaction are modelled by *Flory theory* [48]. The extent to which a real chain swells up is the balance between energetic and entropic contributions to free energy.

A polymer with N monomers, swollen to size $R > R_0 = bN^{1/2}$, we assume that the monomers are uniformly distributed within the volume R^3. The probability of finding another monomer in the excluded volume v of a given monomer is $v\frac{N}{R^3}$. The energy of exclusion per monomer then is $k_BTv\frac{N}{R^3}$. For the entire chain,

$$F_{\text{int}} \approx k_BTv\frac{N^2}{R^3}. \tag{8.9}$$

The Flory estimate of the entropic cost is given in Eq. (8.8),

$$F_{\text{ent}} \approx k_BT\frac{R^2}{Nb^2}. \tag{8.10}$$

The total free energy of a real chain then is,

$$F = F_{\text{int}} + F_{\text{ent}} \approx k_BT\left(v\frac{N^2}{R^3} + \frac{R^2}{Nb^2}\right). \tag{8.11}$$

Hence, the minimum of free energy gives the optimum size of a real polymer,

$$\begin{aligned}\frac{\partial F}{\partial R} &= 0 = k_BT\left(-3v\frac{N^2}{R_F^4} + 2\frac{R_F}{Nb^2}\right)\\ R_F &\approx v^{1/5}b^{2/5}N^{3/5}.\end{aligned} \tag{8.12}$$

The size of the real polymers is much larger than ideal chains with a similar number of monomers. While the exponent $3/5$ scaling obtained from Flory theory happens to be close to the actual value of ≈ 0.588 obtained from more accurate methods (re-normalization theory, Monte Carlo simulations), it is serendipitous as the overestimation of repulsive energy is cancelled by not accounting for correlations along the chain. Finally, it is interesting to note that Flory's theory predicts a universal power scaling for polymer size:

$$R \sim N^{\nu}. \tag{8.13}$$

8.2 POLYMER DYNAMICS

The random fluctuating motion of small particles in a liquid was observed by Robert Brown and is often dubbed as *Brownian* motion. The displacement in 3D is proportional to the time t, with the coefficient of proportionality being related to diffusion constant D,

$$\langle[\boldsymbol{r}(\boldsymbol{t}) - \boldsymbol{r}(\boldsymbol{0})]^2\rangle = 6Dt. \tag{8.14}$$

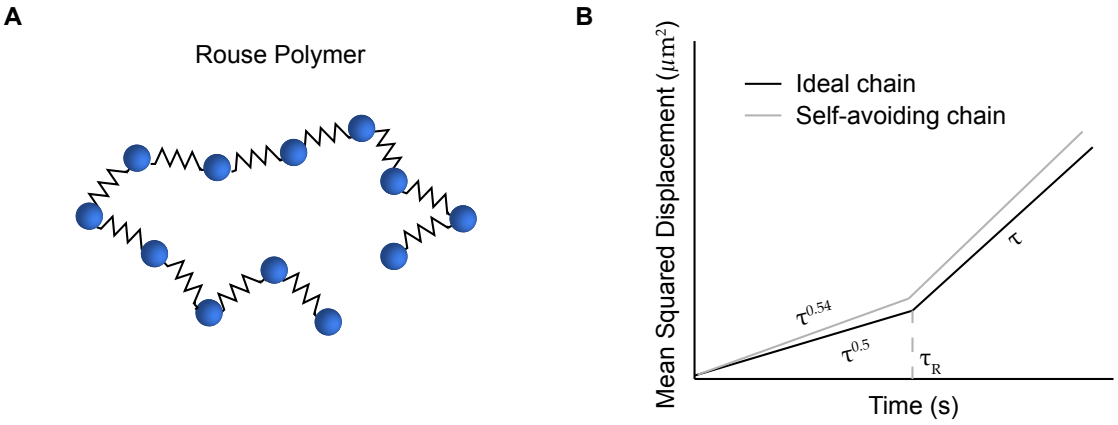

Figure 21: (A) Sketch of Rouse polymer made from beads connected by springs. (b) MSD scaling with delay time τ is plotted. We find $MSD \sim \tau^{0.5}$ and $\tau^{0.54}$ on short timescales $\tau < \tau_R$ for an ideal, self-avoiding chain respectively.

The *Fluctuation dissipation theorem* connects the random Brownian diffusion of particles with the frictional force they experience due to viscosity,

$$D = \frac{k_B T}{\zeta}. \tag{8.15}$$

Rouse Model

We now consider a model of the polymer that is made up of beads connected via springs. Each bead has its own effective independent friction coefficient ζ. The total friction coefficient is given by $\zeta_R = N\zeta$. The diffusion constant of the Rouse polymer is obtained from the Einstein relation,

$$D_R = \frac{k_B T}{N\zeta} \tag{8.16}$$

The *Rouse relaxation* time τ_R, is the time it takes the polymer to diffuse a distance of the order of its size,

$$\tau_R \approx \frac{R^2}{D_R} = \frac{\zeta}{k_B T} N R^2. \tag{8.17}$$

We know from Flory's theory that polymers are fractions with end-to-end distance $R \sim N^{\nu}$. We define a monomer relaxation time as $\tau_0 \approx \frac{\zeta b^2}{k_B T}$, and a power law in the number of monomers in the chain,

$$\tau_R \approx \tau_0 N^{1+2\nu}. \tag{8.18}$$

For an ideal linear chain, we have $\nu = 1/2$ and the Rouse relaxation time is proportional to the number of monomers in the chain,

$$\tau_R \approx \tau_0 N^2. \tag{8.19}$$

Rouse modes

Let us consider the modes of the Rouse model. We know that the longest relaxation mode is that of the entire polymer. Relaxation of the pth mode involves the relaxation of N/p monomers on the polymer and from Eq. (8.18) has the scaling form,

$$\tau_p \approx \tau_0 \left(\frac{N}{p}\right)^{1+2\nu}. \tag{8.20}$$

where $p = 1, 2, ..., N$. Segments of the chain move a distance of the order of their size $b(N/p)^{2\nu}$ during the mode relaxation time τ_p. The mean squared displacement (MSD) of monomer j during a time lag τ_p is given by,

$$\langle [\boldsymbol{r}_j(\tau_p) - \boldsymbol{r}_j(0)]^2 \rangle \approx b^2 \left(\frac{N}{p}\right)^{2\nu} \approx b^2 \left(\frac{\tau_p}{\tau_0}\right)^{\frac{2\nu}{2\nu+1}}. \tag{8.21}$$

Thus on intermediate timescales lower than the Rouse relaxation time τ_R the MSD scales as,

$$\langle [\boldsymbol{r}_j(\tau) - \boldsymbol{r}_j(0)]^2 \rangle \approx b^2 \left(\frac{\tau}{\tau_0}\right)^{\frac{2\nu}{2\nu+1}} \text{ for } \tau_0 < \tau < \tau_R \tag{8.22}$$

At time scales longer than the Rouse time the dynamics of the chain are diffusive. Hence the MSD scales as,

$$MSD \sim \tau^{0.5} \text{ for a free Rouse polymer} \tag{8.23}$$

$$\sim \tau^{0.54} \text{ for a self-avoiding polymer.} \tag{8.24}$$

8.3 POLYMER SIMULATIONS

We used the Bond Fluctuation Method (BFM) [81, 82] to model the polymer dynamics (Fig. 22A). This model is ergodic, allows a large set of bond angles and reproduces the

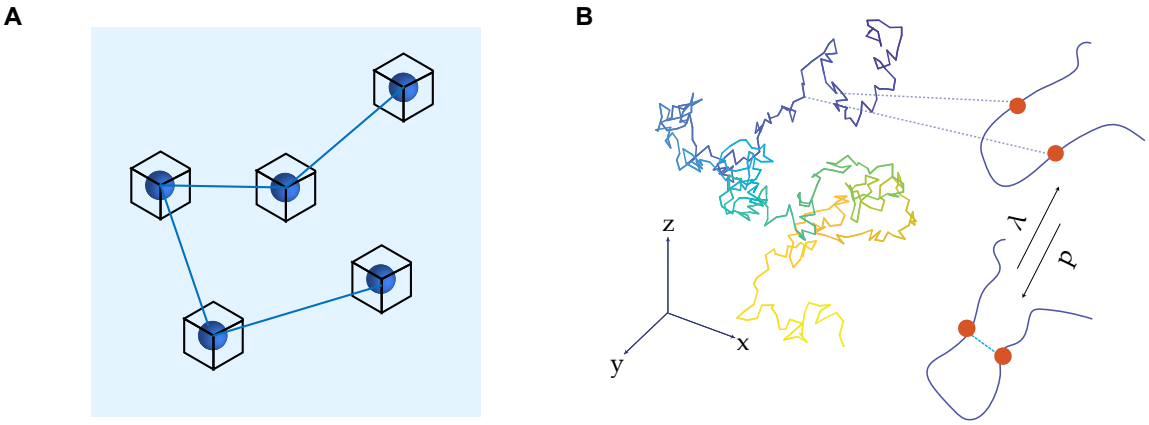

Figure 22: (**A**) Representation of a Bond fluctuation Polymer. (**B**) Snapshot of the simulated polymer ($N = 400$). Bridges form between spatially proximal monomers with a probability p and have a mean lifetime λ.

Rouse polymer dynamics discussed above. This polymer moves only via local jumps of the monomers. The moves do not conserve bond length but preserve the number of bonds. In 3D the monomer is represented by a unit cube and occupies eight lattice sites. Each cube can only host a single monomer to model excluded volume. The bond vectors are typically taken from a set of 108 vectors. The resulting allowed bond lengths are $2, \sqrt{5}, \sqrt{6}, 3, \sqrt{10}$. Furthermore, since bond vector lengths can have any of five values, the volume occupied by the polymer depends on its conformation and, as we will discuss in the next chapter, on the presence of bridges. We simulate a polymer of length $N = 400$, as it is computationally costly to simulate longer polymers. We use a modified version of the open-source LeMonADE software to simulate the BFM polymer and custom-add bridging functionality to it.

Volume Occupied. As each monomer is represented by a cube exclusively occupying 8 vertices, the set of lattice sites occupied by the polymer is the union of $3 \times 3 \times 3$ cubes around the monomers. We define this as the V_{occupied} by the polymer in the box. Based on this measure, we find that our polymer of the length of 400 approximately occupies a total of 8678 lattice sites. Hence, each monomer occupies 21.69 lattice sites on average. We fix the lattice dimension L by using this volume measure to match the volume density of chromosome in the cell ($\sim 1\%$)

$$\rho = \frac{V_{\text{occupied}}}{L^3} = \frac{8678}{L^3} = 1\%. \tag{8.25}$$

This fixes the dimension of the lattice $L = 95$ lattice units. To compare the simulated MSD data with experiments at short time lags we require sufficient spatial resolution at $\sim 0.004 \mu m^2$, the MSD of chromosomal loci at 1 s lag [50, 52] (This will be made clear in later

chapters). We, therefore, fix the lattice spacing to be $h = 0.0056\mu m$. The number of base pairs corresponding to a monomer in our simulations is given by,

$$\frac{\text{base pair}}{\text{monomer}} = \frac{L_g}{V_c}\frac{V_b}{N} \tag{8.26}$$

where $V_c \approx 0.88\mu m^3$ is the volume of the cell, $L_g = 4.5$ Mbp the length of *E. coli* genome, $V_b = L^3h^3$ the volume of the box. Hence, we simulate an 800 kb segment (approximately the size of a macrodomain) of the chromosome with each monomer representing a 2kb segment of the chromosome. We use periodic boundary conditions in our simulations.

Monte Carlo procedure

Each simulation is started from a random conformation of the polymer. Monomer diffusion and bridging are implemented in the following manner,

1. Select a monomer at random and attempt a diffusive move (BFM algorithm).
2. Select a random monomer and if co-localised with another free monomer (distance < 3 lattice units) attempt a bridge with probability p.
3. An un-bridging move is attempted on a random monomer with a probability $\frac{1}{\lambda}$.
4. Repeat.

A set of N moves is defined as a single MCS and we sample every 500 MCS. Note that a bridged monomer can still diffuse as long as the bridge partner is less than 3 lattice units away. We start the Monte Carlo sampling after 5λ MCS to ensure sufficient equilibration of the polymer.

Mean squared displacement

The ensemble-averaged MSD is defined in the following manner,

$$\langle r^2(\tau)\rangle = \frac{1}{N}\frac{1}{T-\tau}\sum_{n=1}^{N}\sum_{t=1}^{T-\tau}\left[\mathbf{r}_n(t+\tau) - \mathbf{r}_n(t)\right]^2. \tag{8.27}$$

We perform a linear fit to the logarithm of the MSD curves,

$$\log(\langle r^2(\tau)\rangle) = \alpha\log(\tau) + D_{\text{apparent}} \tag{8.28}$$

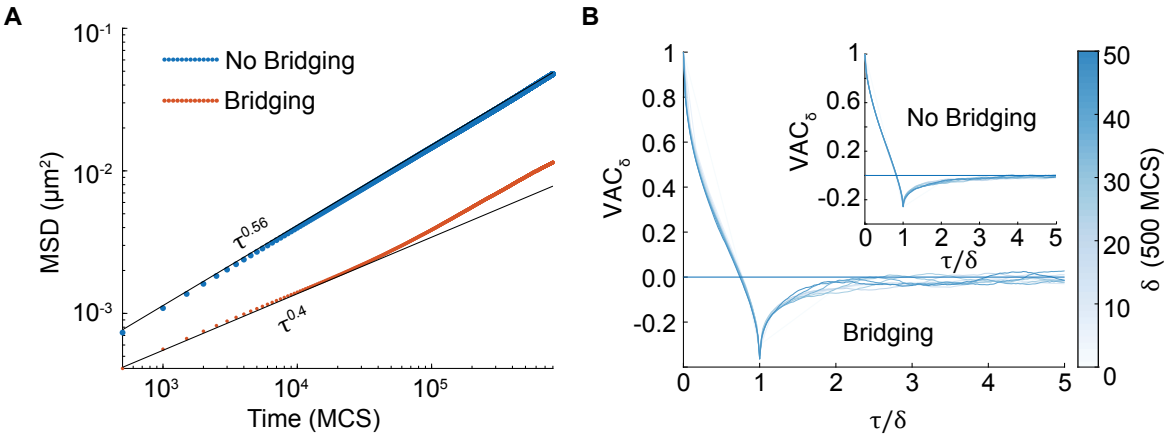

Figure 23: (**A**) The mean squared displacement (MSD) as a function of time lag τ with (orange) and without bridging (blue). Bridging reduces the scaling exponent from $\alpha \approx 0.56$ to $\alpha \approx 0.4$. Bridging parameters: $p = 1.5 \times 10^{-3}, \lambda = 4 \times 10^4$ MCS. Standard error bars are smaller than the thickness of the lines. (**B**) Velocity auto-correlation function (VAC) is negative at short lags and collapses for different windows of τ/δ indicative of a sub-diffusive process.

up to a delay of 20000 MCS and obtain ensemble averaged α. The data in Fig. 24A are obtained from 50 ($\lambda \leq 10^5$), 30 ($\lambda > 10^5$) independent simulations for each parameter. The ensemble-averaged MSD curves are calculated from tracks of every 20th monomer on the polymer ($600 - 1000$ tracks per parameter).

8.4 Bridging reproduces sub-diffusive scaling exponent

We implement bridging between monomers using the Dynamic Loop model [46,70]. Any two non-neighbouring monomers that are a distance of less than 3 lattice units apart can form a bridge between each other with a probability p (Fig. 22B). Bridges dissociate randomly with probability $1/\lambda$, i.e. they have an average lifetime of λ in units of Monte Carlo Steps (MCS). While bridged, monomers can still diffuse on the lattice subject to maintaining a bridge length of less than 3 lattice units. Importantly, each monomer can only form one bridge at a time.

We first confirmed that bridging reduces the scaling exponent of the mean square displacement of a single monomer (23A), as has been previously shown using a similar model in the context of looping of eukaryotic chromatin [70]. We found that bridging indeed lowers the exponent: with a level of bridging that results in 28% of monomers bridged, the exponent decreased from $\alpha \approx 0.56$ (close to the scaling theory prediction of 0.54 Eq. (8.24)) to $\alpha \approx 0.4$,

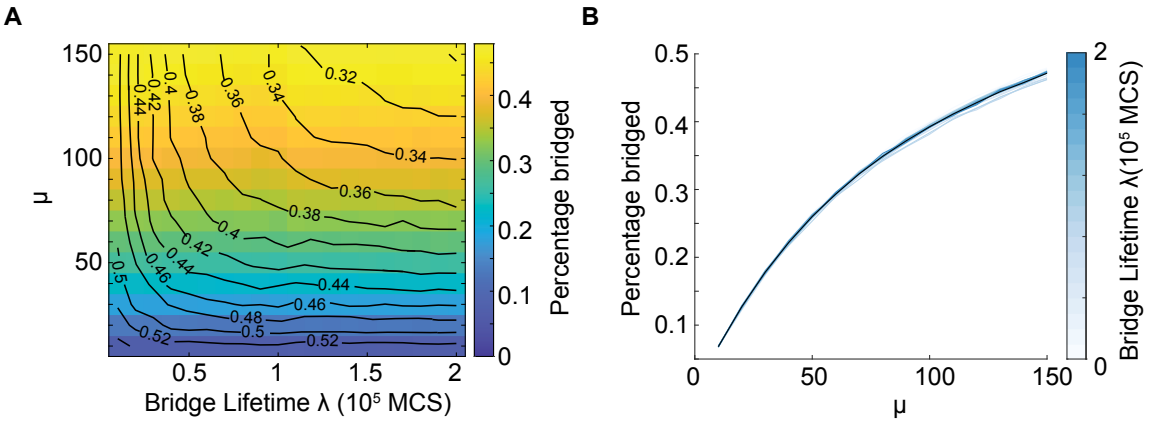

Figure 24: (**A**) Phase diagram of the dynamic bridging model in the $\mu = p\lambda$ and λ space. Note that μ is positively related to the percentage of monomers bridged. Contours indicate a fixed exponent α. **(B)** Percentage of monomers bridged increases with μ and is independent of bridge lifetime.

a value in line with experimental measurements [50]. To examine the dynamics further, we measured the Velocity Auto-Correlation (VAC) function

$$\mathrm{VAC}_\delta(\tau) = \frac{1}{\delta^2}\left\langle (\boldsymbol{r}(\tau+\delta) - \boldsymbol{r}(\tau)) \cdot (\boldsymbol{r}(\delta) - \boldsymbol{r}(0)) \right\rangle \tag{8.29}$$

of individual monomers (with the velocity measured over time points δ MCS apart). We found that bridging did not change the nature of the VAC (Fig. 23B), which remained negative at short time lags with the lowest value at a lag equal to δ. This is indicative of elastic or sub-diffusive dynamics and is consistent with experimental measurements of chromosomal loci [83]. The VAC has previously been used to distinguish between different models of sub-diffusivity like Fractional Brownian Motion (fBm) and Continuous Time Random Walk (CTRW) [50, 56]. However, it is important to note that this nature of the velocity auto-correlation function is true for any dynamics process that has an MSD that scales as a power law (see below).

Velocity auto-correlation function

The velocity auto-correlation function Eq. (8.29) can be expanded as,

$$\mathrm{VAC}_\delta(\tau) = \frac{1}{\delta^2}\Big\langle \boldsymbol{r}(\tau+\delta)\cdot\boldsymbol{r}(\delta) - \boldsymbol{r}(\tau+\delta)\cdot\boldsymbol{r}(0) - \mathbf{r}(\tau)\cdot\boldsymbol{r}(0) + \mathbf{r}(\tau)\cdot\boldsymbol{r}(0)\Big\rangle. \tag{8.30}$$

we know that,

$$\langle \boldsymbol{r}(\tau+\delta)\cdot\boldsymbol{r}(\delta)\rangle = \langle \boldsymbol{r}(\tau+\delta)^2\rangle + \langle \boldsymbol{r}(\delta)^2\rangle - \langle (\boldsymbol{r}(\tau+\delta)-\boldsymbol{r}(\delta))^2\rangle.$$

Hence we can re-write the VAC as,

$$\begin{aligned}\mathrm{VAC}_\delta(\tau) &= -\langle (\boldsymbol{r}(\tau+\delta)-\boldsymbol{r}(\delta))^2\rangle + \langle (\boldsymbol{r}(\tau+\delta)\\ &\quad -\boldsymbol{r}(0))^2\rangle + \langle (\boldsymbol{r}(\tau)-\boldsymbol{r}(\delta))^2\rangle - \langle 0(\boldsymbol{r}(\tau)-\boldsymbol{r}(0))^2\rangle\\ &= D_{app}\left(\mid \tau-\delta \mid^{\alpha} + \mid \tau+\delta \mid^{\alpha} -2 \mid \tau \mid^{\alpha}\right). \end{aligned} \tag{8.31}$$

where we have assumed that the underlying process has an MSD that scales as a power law in time lags. Hence, we find that the normalised VAC has the form,

$$\frac{\mathrm{VAC}_\delta(\tau)}{\mathrm{VAC}_\delta(0)} = \frac{\mid \tau-\delta \mid^{\alpha} + \mid \tau+\delta \mid^{\alpha} -2 \mid \tau \mid^{\alpha}}{2\delta^{\alpha}}. \tag{8.32}$$

This results in a negative peak for all values of δ, indicating a sub-diffusive process. Furthermore, the negative peak is observed at $\tau = \delta$. We find that normalizing time by δ results in a universal curve for all values of δ plotted (Fig. 23B).

Phase diagram of bridging

To systematically examine the effect of bridging, we varied the bridging probability p and average bridge lifetime λ, and measured the percentage of monomers bridged and the resultant scaling exponent α. We found that the percentage of bridged monomers does not depend on the bridge lifetime individually but only on the product $\mu = p\lambda$ (Fig. 24A, B). However, as a dynamical measure, the scaling exponent depends on both μ and λ; the more bridges and the longer their lifetime, the greater the reduction of the scaling exponent. However, short-lived bridges ($\lambda \lesssim 10^4$ MCS) have little effect on the exponent irrespective of their number.

9

BRIDGING COMPACTS THE POLYMER

In this chapter, we study the effects of bridging on polymer conformations. We find that bridging does indeed affect the polymer geometry by compacting it.

9.1 POLYMER SIZE

Examining the effect of bridging on the organization of the polymer, we found a clear effect of compaction (Fig. 25A). This could be quantified using the radius of gyration $\langle R_g^2 \rangle$ defined as in Eq. (8.5) which measures the average size of the polymer. However, we found $\langle R_g^2 \rangle$ to be a relatively noisy measure of polymer size (Fig. 25B), motivating us to look for an alternative. The volume occupied by the polymer (V) as defined in section 8.3 was found to be a much more robust measure (Fig. 25D). In Fig. 25C we plot $\Delta V \equiv \frac{V_{\text{occupied}} - V_{\text{no bridging}}}{V_{\text{no bridging}}}$ as a function of bridges and observe a linear decrease and a collapse across bridge lifetimes. The latter confirms results from previous Brownian dynamics simulations that the polymer relaxation time can be controlled (through the bridge lifetime) independently of the equilibrium structure [84].

Bridging decreases the mesh size of the polymer

The compaction could also be seen by the reduction in the mesh size (from 120 nm without bridging to 85 nm with bridging) (Fig. 26A). Here, we calculate the mesh size ξ from our simulations. The probability of finding a monomer at a distance r and $r + dr$ from another randomly chosen monomer is given by $4\pi\rho r^2 g(r)\, dr$, where $\rho = 0.01$ is the density of the polymer and $g(r)$ is a radial density function. For a semi dilute polymer with $r \gg \xi_c$, $g(r)$ is expected to have the form,

$$g(r) = 1 + \frac{A}{r} exp(-r/\xi_c) \tag{9.1}$$

where $A > 0$ and $\xi_c > 0$. ξ_c is the correlation length of the polymer which is approximately the same as the mesh size ξ for a semi-dilute polymer [85]. We calculate $g(r)$ from our simulations and fit to Eq. (9.1) and find the mesh size (ξ) for different parameters.

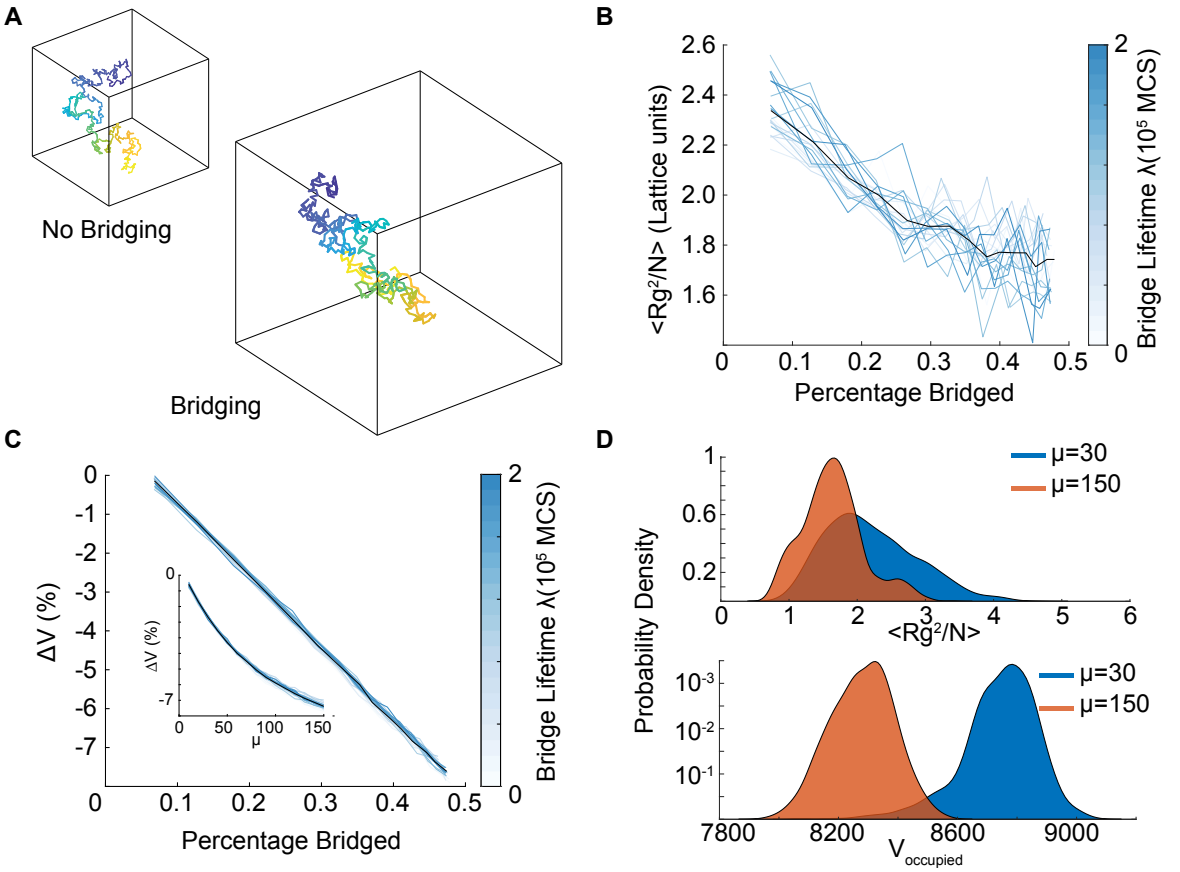

Figure 25: (**A**) Example polymer conformations with and without bridging. **(B)** Radius of Gyration $\langle R_g^2/N \rangle$ as function of Percentage Bridged. It is a noisy measure of the polymer size. **(C)** The change in volume occupied relative to the non-bridging polymer ΔV decreases linearly with bridging. (Inset) The relationship between ΔV and the model parameter μ. The black line indicates the average across bridge lifetimes. **(D)** Distributions of $\langle R_g^2/N \rangle$ are overlapping for different levels of bridging while, distributions of V_{occupied} are significantly separated.

Effects of confinement

Thus far, our simulation parameters were chosen to match the density of DNA within an *E. coli* cell rather than the effect of confinement due to the cell boundaries (average diameter of the polymer relative to the dimension of the box $\frac{2Rg}{L}$) since the former quantity is likely to more strongly affect the probability of bridge formation. The density d scales linearly with number of monomers N, while confinement scales as $R_g \sim N^{\nu}$, where $\nu = 0.588$. If we fix the confinement of the chromosome instead of its density, the general results in the previous chapters remain unaffected, albeit changing the exact values of α for different parameters. As confinement increases, density increases and α decreases even in the absence of bridging up to $\alpha \approx 0.5$ at which point confinement screens the effect of self-avoidance (Fig. 27B) [86].

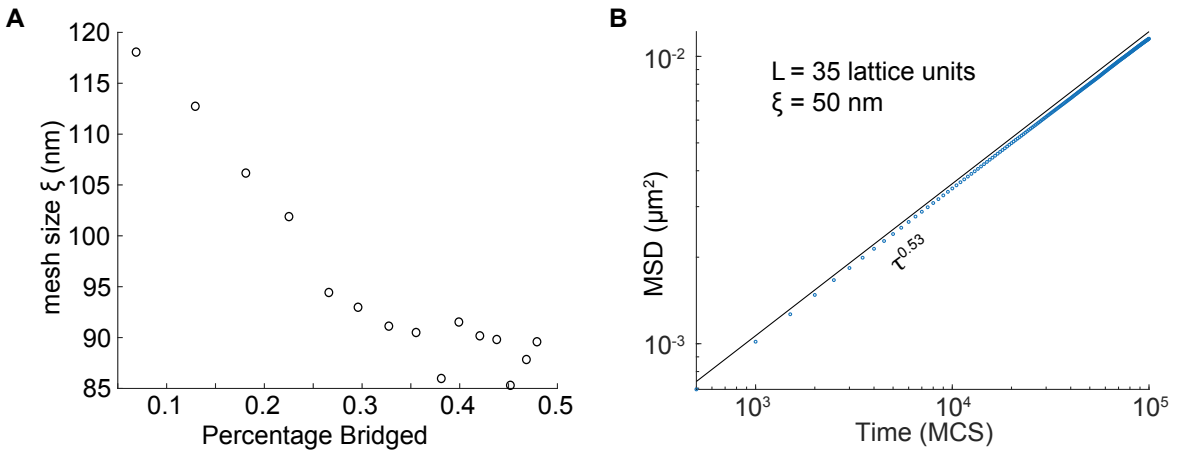

Figure 26: **(A)** Mesh size (ξ) decreases with increasing bridges. **(B)** Confinement affects the scaling exponent of MSD even in the absence of bridging. Parameters.Lattice size $L = 35$, Mesh size $\xi = 50$nm, Polymer length $N = 400$.

Density is the limiting factor for bridging, as a change in density affects the probability of co-localisation for any two monomers. As matching confinement will lead to a higher density of the polymer in the box, we matched the density of the chromosome to model the realistic effects of bridging by NAPs. Recently, it was observed that the unconstrained *E. coli* chromosome is approximately 2.35 times longer than the cell length [54]. By decreasing the lattice size to match the confinement of the *E. coli* chromosome instead of its density we further decrease mesh size and can match the experimental value of 50nm [87] (Fig. 26B). Note that increasing confinement increases the number of bridges and consequently decreases the scaling exponent for the same set of parameters μ, λ.

End-to-end scaling

A swollen or ideal polymer is a fractal in that the end-to-end distance of the sub-segments also scales with the length of the segments as,

$$r(s) \sim s^{\nu}. \tag{9.2}$$

On the other hand, if we consider an equilibrium globule formed by a collapsed linear chain we expect different scaling relations on different length scales. The relation in the above equation is limited by the size of the globule in the following manner [88],

$$r(s) \sim \begin{cases} s^{1/2} & \text{for } s < s^* \sim N^{2/3} \\ N^{1/3} & \text{for } s > s^* \sim N^{2/3} \end{cases}. \tag{9.3}$$

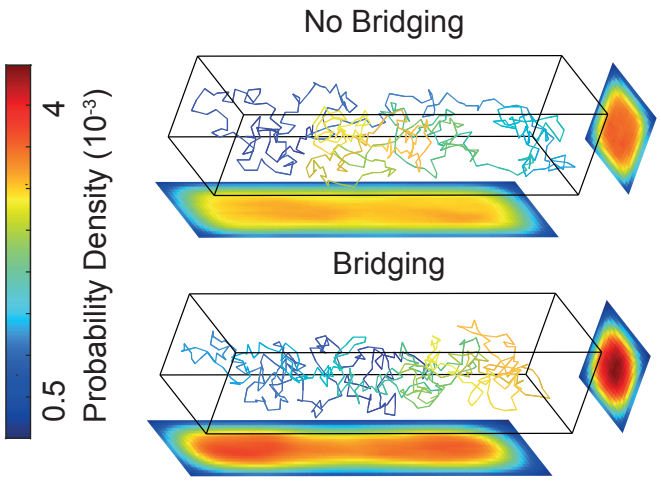

Figure 27: A circular polymer in a cuboid with hard walls matching the confinement of *E.coli* cell. In the presence of bridging the polymer is compacted and has a density that decreases across the cross-section. Lattice dimensions $88 \times 22 \times 22$.

In our simulations, even for the highest values of bridging in our simulations, the polymer does not transition to the globule regime as evidenced in Fig. 28A.

Two-point contact probability or loop factor

For any fractal conformation, we know from the above equations that $r(s) \sim s^{\nu}$. The other end of a segment is at about a distance of $r(s)$ and dispersed over a volume $r^3(s)$ in $3D$. Therefore the probability $P(s)$ to a find another point on the segment is,

$$P(s) \sim r^{-3}(s) \sim s^{-3\nu}. \tag{9.4}$$

This for a Gaussian or ideal coil with $\nu = 1/2$ gives $P(s) \sim s^{-3/2}$ and for a self-avoiding polymer ($\nu = 0.588$) $P(s) \sim s^{-1.76}$ [89]. For a crumpled fractal globule we have $\nu = 1/3$ and hence $P(s) \sim s^{-1}$. While this scaling relation of two-point contact probability shows good agreement with experimental HiC data, for both human and mouse chromosomes it is unfortunately not accurate. The sum $\sum_s P(s)$, which measures the number of neighbours of any given monomer diverges as $P(s) \sim 1/s$. For a N monomer long chain, this would imply that the mean number of spatial neighbours is of $\mathcal{O}(\ln N)$ (diverges as N to infinity). Hence, from a logical standpoint, the assumption of statistical independence of sub-chains cannot be accurate for interacting chains. The exact scaling for an interacting polymer is given by,

$$P(s) \sim s^{-3\nu-\gamma}, \tag{9.5}$$

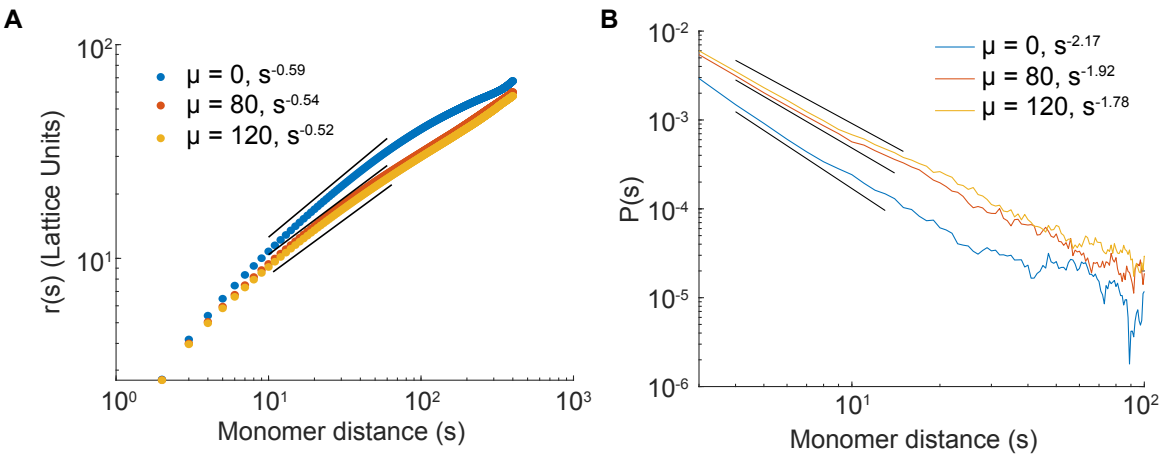

Figure 28: **(A)** End to end distance scales as $\langle R \rangle \sim N^{\nu}$. We observe an exponent $\nu = 0.59$ in the absence of bridging as expected for a self-avoiding polymer. The scaling exponent decreases with increasing bridging. But, we observe that scaling exponent $\nu > 0.5$ for all levels of bridging. **(B)** The contact probability $P(s)$ between monomers is plotted as a function of monomer distance (s). In the absence of bridging the exponent is $P(s) \sim s^{d\nu+\gamma} \approx s^{2.1}$ as expected for a self-avoiding chain. With increasing bridging the exponent increases, but for our level of bridging we never stay below the compact globule regime.

where γ is an independent exponent arising from the accounting of interactions between segments. The theoretical scaling relation for a swollen or self-avoiding chain is given by $P(s) \sim s^{-2.18}$ [90, 91]. We find the exact scaling in our numerical simulations of the polymer

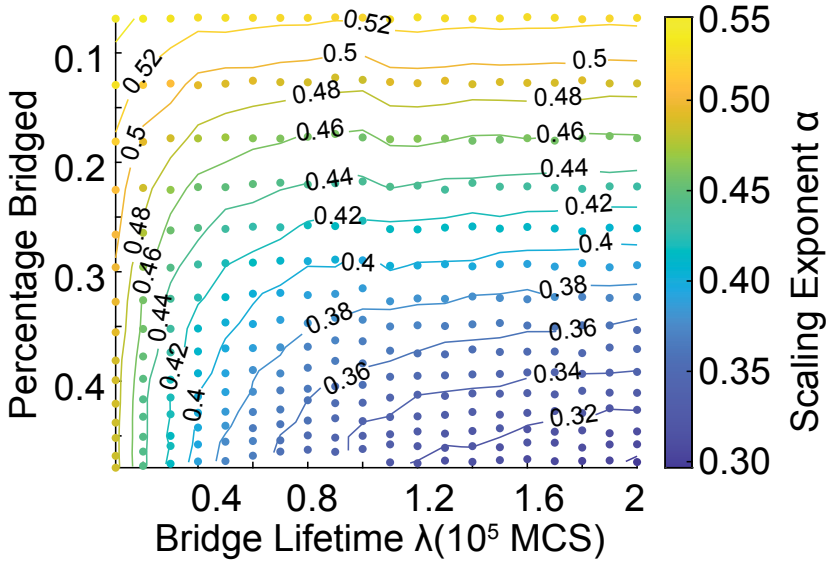

Figure 29: Phase diagram from Figure 24A with μ remapped to the percentage of monomers bridged.

with $P(s) \sim s^{-2.17}$ for a self-avoiding polymer with no bridges (Fig. 28B). As we increase bridging the slope increases but stays below the globule regime $P(s) \sim s^{-1.2}$ for the simulated levels of bridging. The compact globule regime has been studied elsewhere [70, 76, 84].

9.2 REMAPPED PHASE DIAGRAM

To obtain more meaningful physical insights we remapped the phase diagram in terms of the average number (percentage) of monomers bridged rather than the parameter μ (Fig. 29). This makes it clear that, for sufficiently long bridge lifetimes, an increase (decrease) in the number of bridges formed is concomitant with a decrease (increase) of the scaling exponent. As discussed in the previous chapters, Nucleoid Associated Proteins (NAPs) bridge and compact the chromosome. In the next chapter, we present short-timescale loci tracking experiments performed in *E. coli* to verify this hypothesis.

10

NAP MUTANT HAS A HIGHER EXPONENT

As discussed in the previous chapter, according to our model for sufficiently long bridge lifetimes, an increase or decrease in the average number of bridges on the chromosome is concomitant with a decrease or increase of the scaling exponent α. To test this hypothesis, we perform loci tracking experiments in *E. coli* cells. We measured the MSD scaling exponent of a chromosomal locus in a strain lacking the NAP H-NS. We choose this protein as its deletion has a very mild phenotypic effect compared to other NAPs like HU or MukBEF [62]. Following previous work [52], we used fluorescence microscopy and a GFP-ParB/*parS* labelling system to track the *ori* locus of *E. coli* on short timescales.

10.1 LOCI TRACKING EXPERIMENTS

Strains

The *parS*/P1 site from *E.coli* strain RM29 obtained from [92] (originally from [93]) was transduced near the *ori* region into MG1655 WT and ΔH-NS strains, the latter obtained from the Keio collection of the Sourjik lab (MPI Terrestrial Microbiology). GFP-ParB was expressed from the plasmid pALA2705 with no IPTG induction [43, 93, 94]. The strains were

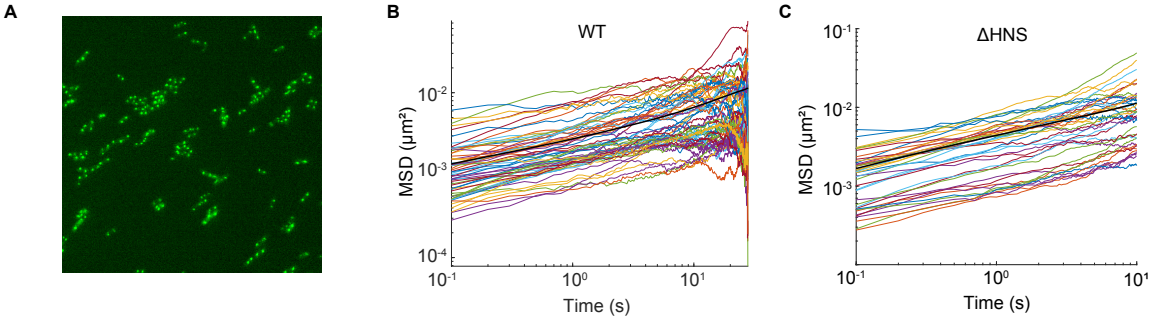

Figure 30: **(A)** Snapshot of microscopy experiment of GFP-ParB/**parS** labelled *ori* loci in *E.coli*. **(B)** Sample MSD curves of individual foci obtained from WT strain (set 1). **(C)** Sample MSD curves of mutant ΔH-NS (set 1). Ensemble averaged MSD is represented by the black lines (**B**, **C**).

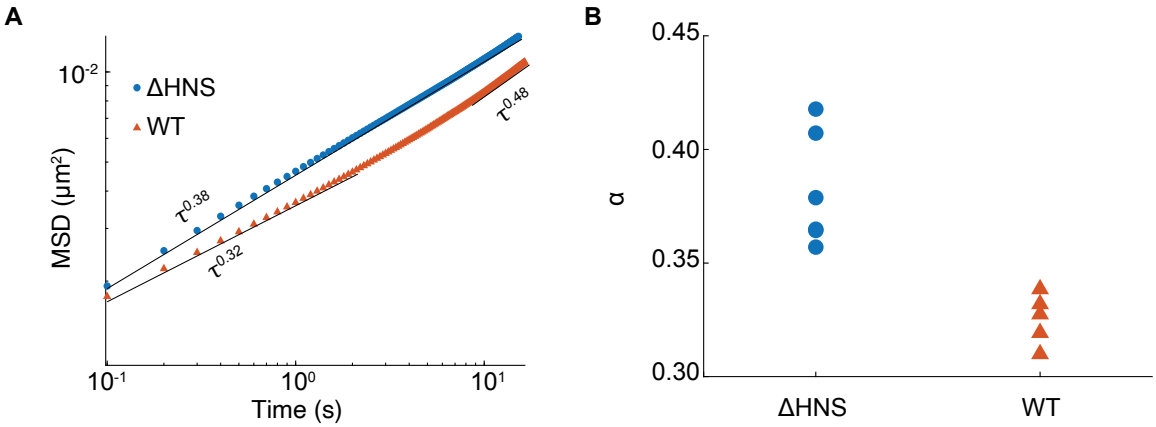

Figure 31: **(A)** Loci tracking experiments. Ensemble averaged MSD curves of wild-type *E.coli* and a strain deleted of the NAP H-NS. The deletion strain has a higher scaling exponent ($\alpha \sim 0.38$) compared to the WT ($\alpha \sim 0.32$). MSD curves are fitted up to a 2 s delay. ΔH-NS: 18089 tracks, WT: 6717 tracks. **(B)** Exponent α from individual replicates.

grown overnight at 30°C in LB medium with appropriate antibiotics (100μg/mL ampicillin). The overnight culture was diluted into media made of M9–Glucose–Casamino acids (as in [52]) and grown to an optical density of $0.1 - 0.2$.

We chose the P1 labelling system in order to compare our results with previous studies [50, 52, 57, 95]. We note that while some differences in the dynamics of the *ter* locus between the ParB labelling systems we use (P1) compared to that of pMT1 have been observed, no substantial differences have been reported for the *ori* locus [43, 55, 94, 96]. This agrees with experiments in our lab studying origin positioning and segregation across many thousands of cell cycles.

Microscopy

1μL of the sample was placed on 1.5% agarose pads (made of the same media as the day culture) and imaged under a Nikon Ti microscope with a 60x/1.4 NA oil objective. The strains were imaged at a constant 30°C. Images were captured on a Hamamatsu CCD camera using NIS-Elements software. Movies were 450 frames long, with 0.1s interval and an exposure time of 100ms (Fig. 30A).

Analysis

We process the raw microscopy images following the procedure and analysis in [52]. Briefly, foci positions were located via two-dimensional fitting of a Gaussian function to the intensity distributions of individual loci. The ensemble-averaged MSD was calculated from pooled

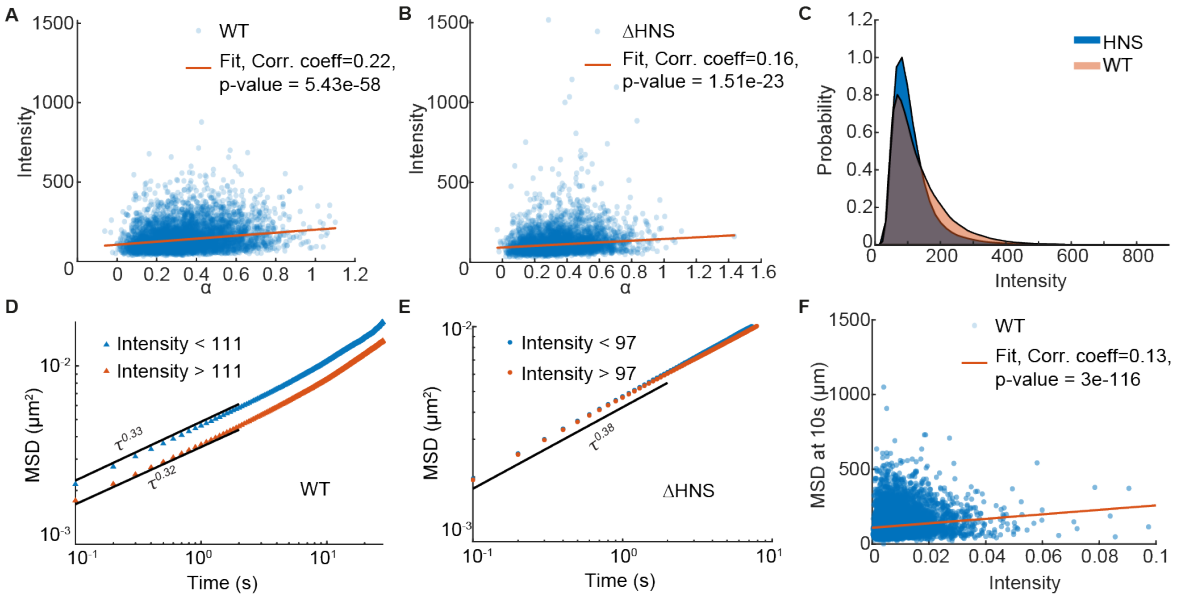

Figure 32: **(A)** Mean intensity of WT tracks plotted versus scaling exponent α fitted to the entire track length. Shows a weak positive correlation. **(B)** Same as in (A) but for the mutant ΔH-NS. **(C)** Intensity distributions of WT and ΔH-NS is plotted. WT distribution has a broader tail. **(D)** Ensemble averaged MSD of WT plotted for two difference sub-populations with intensity $I >$ median(I) and $I <$ median(I). Scaling exponent α shows marginal change while D_{app} decreases with higher intensities [52]. **(E)** Same as in (D) but for the mutant. The sub-populations overlap. **(F)** Intensity of loci in the WT at an MSD of 10s. We find a very weak positive correlation.

trajectories using Eq. (8.27) (Fig. 30B, C). The scaling exponent α was calculated for ensemble-averaged MSD curves by fitting a power law up to 2 s delays. The codes from Javier et.al were used to track the foci, and are available at bacteria-loci-tracker. Custom MATLAB scripts were written to analyse the data.

10.2 ΔH-NS has a higher scaling exponent

The ensemble-averaged MSD of both the wild type and the ΔH-NS strain are shown in Figure 31A. Consistent with our model, we found that the scaling exponent α for the ΔH-NS ($\alpha \sim 0.38$) is greater than that of the wild type $\alpha \sim 0.32$, consistent with a decrease in the number of DNA bridges. While, we observe some variability in the exact value of the exponents between biological replicates (Fig. 31B), the exponent of ΔH-NS is consistently higher than the wild type (see also Table1). We conclude that the bridging of chromosomal DNA by nucleoid-associated proteins affects the nature of chromosome dynamics and can explain why the scaling exponent of chromosomal loci is more sub-diffusive than expected from polymer dynamics alone.

Scaling exponent α													
Strain	set 1	Tracks	set 2	Tracks	set 3	Tracks	set 4	Tracks	set 5	Tracks	set 6	Tracks	All
WT	0.31	3952	0.32	525	0.32	1193	0.34	477	0.33	506	—	—	0.32 ± 0.01
ΔH-NS	0.36	12121	0.36	827	0.42	1647	0.41	1682	0.38	1062	0.36	750	0.38 ± 0.02

Table 1: Scaling exponents α observed in experiments from different sessions.

10.3 INTENSITY OF SPOTS DOES NOT EXPLAIN DIFFERENCES IN SCALING EXPONENTS

We wondered if the difference in MSD scaling exponents between the WT and mutant could be related to the intensity, as it was previously shown that the mobility of loci depends inversely on their intensity [52]. In our experimental data, we find that the WT and mutant have comparable intensity distributions and show a very weak correlation between the loci intensity and the scaling exponent α (Fig. 32A, B). We also found that the intensity distributions of the strains were comparable, while the WT had a slightly fatter tail (Fig. 32C). Comparing the ensemble-averaged MSD of tracks with lower and higher intensity, we found that while intensity affects the apparent diffusion constant D_{app}, it has a marginal effect on the scaling exponent α (Fig. 32D, E). The intensity of loci also shows a very weak correlation with the MSD at 10s (Fig. 32F). Hence, we conclude that the intensity of loci does not explain the differences between the ensemble-averaged MSD exponents of wild type and ΔH-NS.

11

BRIDGING REPRODUCES RAPID CHROMOSOMAL MOVEMENTS

In the previous chapters, we discussed how chromosomal loci move sub-diffusively at short timescales with an MSD scaling exponent consistently less than 0.5. A previous study performing tracking measurements of *ori* loci revealed that a small number of tracks $50-350$ out of $3000-7000$ tracks were abnormally elongated and faster [57]. They further identified that the sub-population of fast-moving trajectories could not be explained by the null phenomenological model of fractional Brownian motion. Instead, these outliers, termed Rapid chromosomal movements (RCMs), were speculated to be due to an active machinery or some stress-relaxation mechanism [57, 97]. To determine if RCMs were also present in our data, we followed the same procedure as Javer *et al.* to identify them.

11.1 PRESENCE OF RCMS IN OUR EXPERIMENTAL DATA

We fit the wild-type ensemble-averaged MSD curves (Fig. 31) to a power law $D_{app}\tau^{\alpha}$ to determine the two parameters of the fBm model [98], namely the apparent diffusion constant D_{app} and the exponent α (Fig. 33A). These parameters were then used to simulate fBm using

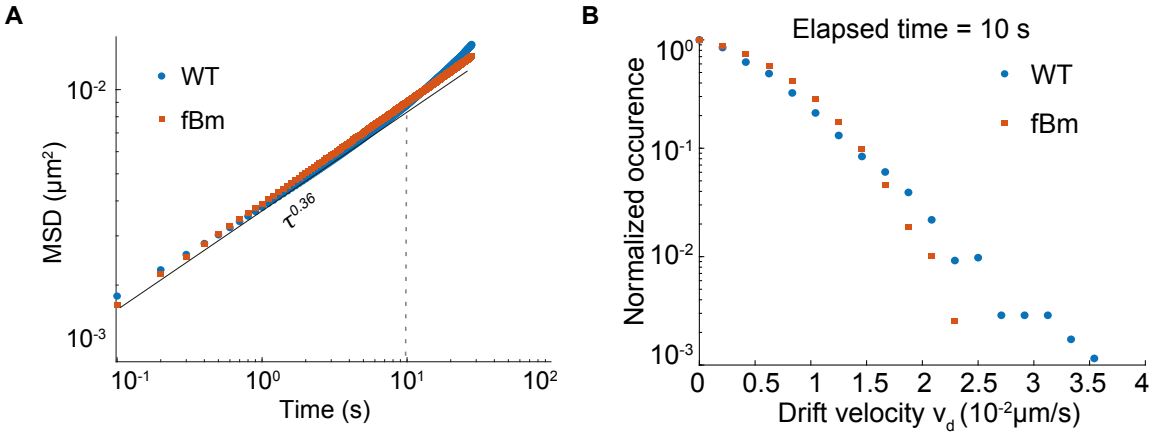

Figure 33: **(A)** Ensemble averaged MSD from experimental data of WT (blue circles) is fitted to fBm (orange squares, Fit $0.0008\tau^{0.36}\mu m^2$, 6717 tracks). **(B)** Drift velocity of wild-type tracks (blue circles) have a much wider tail than from simulations of fBm with model parameters chosen according to the fit in (**A**) (orange squares).

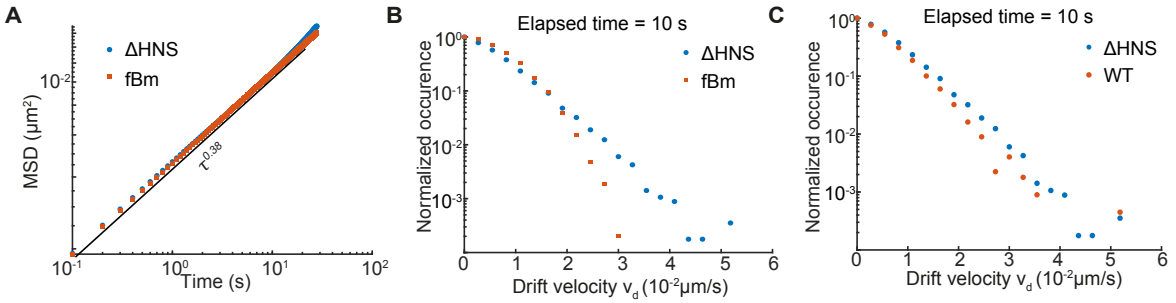

Figure 34: **(A)** Ensemble averaged MSD from FBM simulations overlaid with ΔH-NS data (Fit, $0.0009\tau^{0.38}$). **(B)** Drift velocity distributions v_d from FBM simulations have a smaller tail than ΔH-NS data. **(C)** ΔH-NS has a slightly higher number of faster tracks than the wild type.

the same number and track lengths as the experimental data. We then calculated the drift velocity v_d for each track, defined as the magnitude of the displacement along the major axis between two time points divided by the elapsed time. Similar to Javer *et al.*, we found that the distribution of drift velocities of the experimental data displayed a fatter tail than that of the fBm simulations (Fig. 33B).

Fractional Brownian Motion (fBm)

Fractional Brownian motion is a generalisation of Brownian motion [99]. Unlike classical Brownian motion, the increments of fBm are not constrained to be independent. Let $B(t)$ represent ordinary Brownian motion and H be a parameter satisfying $0 < H < 1$,

$$B_H(t) = B_H(0) + \frac{1}{\Gamma(H+1/2)} \int_{-\infty}^{0} \left[(t-s)^{H-1/2} - (-s)^{H-1/2} \right] dB(s) + \int_{0}^{t} (t-s)^{H-1/2} dB(s). \tag{11.1}$$

The fBm exponent H is called the Hurst index and describes the raggedness of the Brownian paths. Importantly, the process has expectation zero for all t in $[0, T]$ and has the covariance,

$$E[B_H(t)B_H(s)] = \frac{1}{2}\left(\mid t^{2H} \mid + \mid s \mid^{2H} - \mid t-s \mid^{2H} \right). \tag{11.2}$$

where we recover Brownian motion if $H = 1/2$. We simulate this process using Lowen's method [100] which is not only fast $\mathcal{O}(N \log N)$ but also yields exact statistics that precisely match those of fBm.

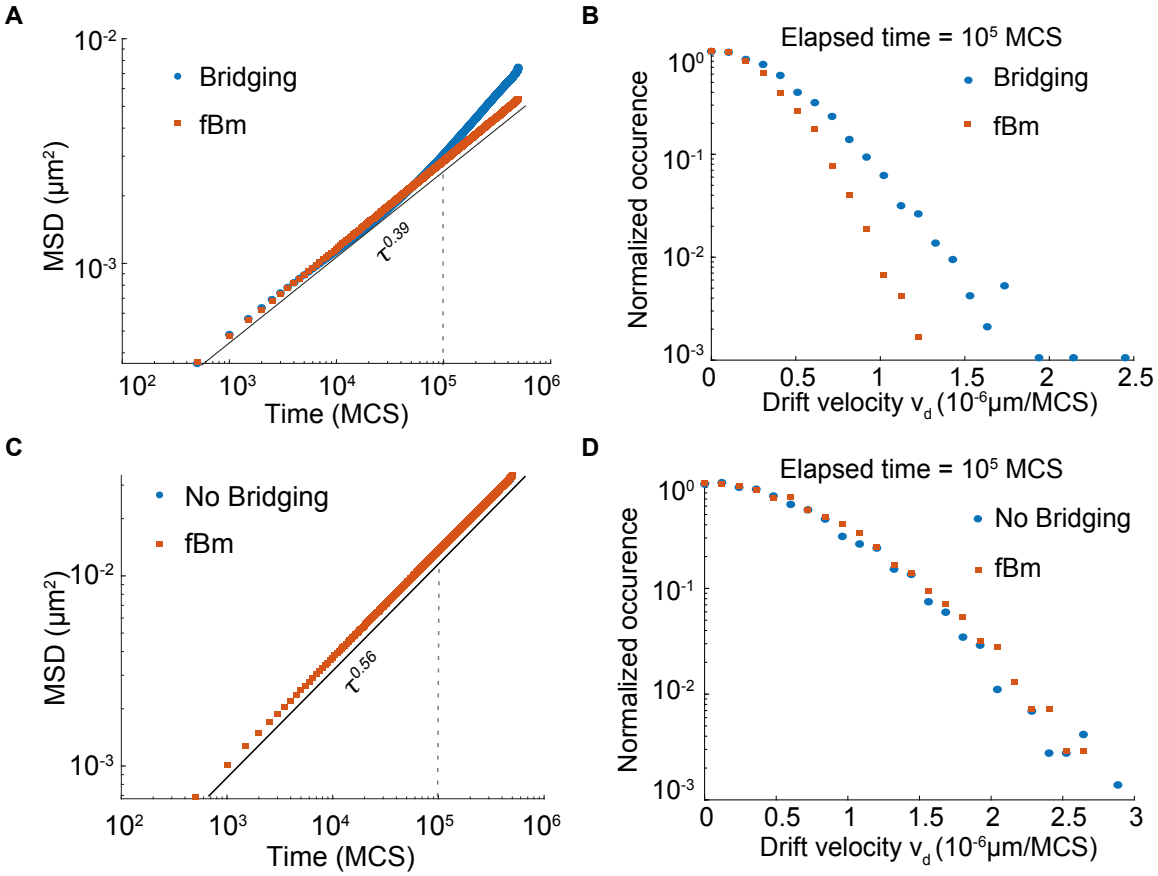

Figure 35: **(A)** Ensemble averaged MSD from polymer simulations fitted with fBm simulations (orange squares, $\mu = 80, \tau = 10^5$ MCS, Fit $0.0009\tau^{0.39}\mu m^2$, 5200 tracks). **(B)** Drift velocity distributions v_d of fBm (orange square) have a smaller distribution than bridging tracks (blue circles). **(C)** Same as (**A**, **B**) but to polymers with no bridging. **(D)** The distributions of v_d overlap.

Similar to the case of the wild type, parameter matching the mutant ΔH-NS MSD to fBm simulations (Fig. 34A) we find that it fails to capture the faster outlier tracks (Fig. 34B). Note that the mismatch in v_d distribution is not dependent on the power law nature of the mutant MSD curve leading to a better fit. This disparity was not dependent on the precise elapsed time used nor on the deviation of the MSD curve from a perfect power law (or the range over which the parameter fitting was performed) as we note that ΔH-NS strain, displays a near-perfect power-law behaviour (Fig. 34A, B). Importantly as predicted by our bridging hypothesis, we find that the mutant has a slightly broader v_d distribution due to the absence of H-NS (Fig. 34C).

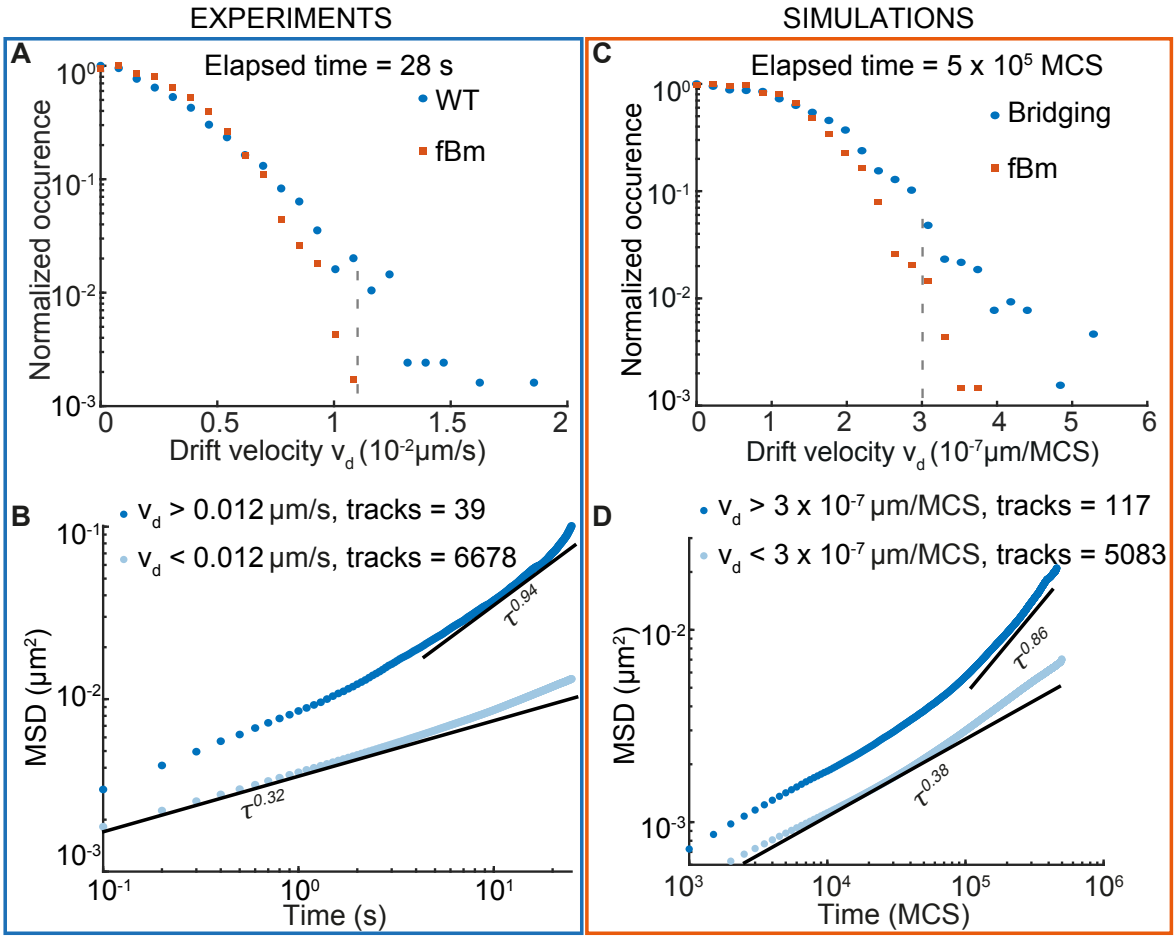

Figure 36: **(A)** Drift velocity distribution comparisons between WT data and fBm measured over 28s (entire track length). **(B)** Ensemble averaged MSD of the wild type strain with tracks selected based on drift velocity v_d (grey dashed line in (**A**) shows the v_d threshold). Tracks with v_d greater than fBm distribution have a transition to higher exponent at longer time lags. (**C**) Same as in (**A**) for the bridging simulations with v_d defined over the entire track length of 5×10^5 MCS. **(D)** We find a similar subset of tracks (split based v_d, grey dashed line in (**C**)) which transition to higher exponent at longer time lags. Parameters $\mu = 80, \lambda = 10^5$ MCS.

11.2 BRIDGING REPRODUCES RCMS

Following the same procedure, we then parameter-matched the fBm model using the ensemble-averaged MSD curves of our simulated monomer trajectories (Fig. 35A). Surprisingly, we found that the bridging simulations produced trajectories with a similar over-representation of high drift velocities compared to the fBm model as seen for the experimental data (Fig. 35B). We could directly attribute this disparity to the effect of bridging since simulations without bridging showed no such disagreement (Fig. 35C, D). Again, these results persisted irrespective of the precise fitting to the MSD curves and the

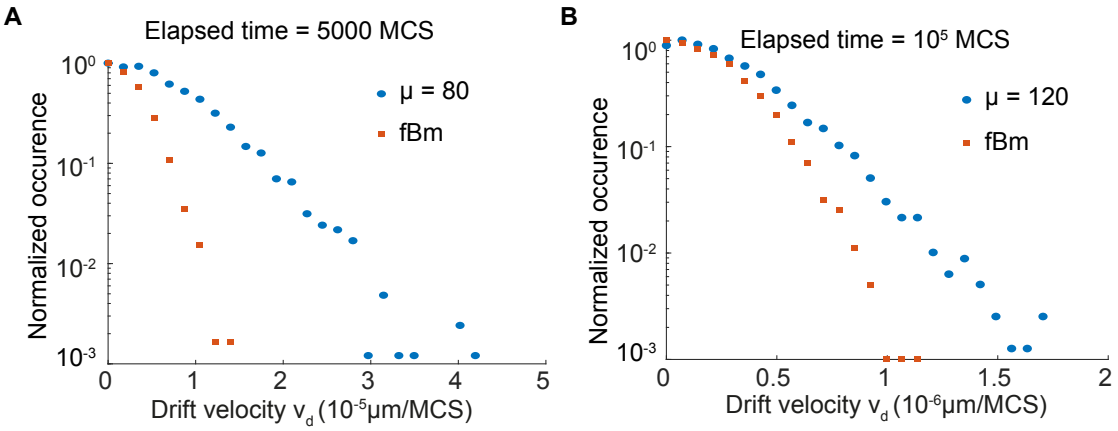

Figure 37: **(A)** Drift velocity distribution comparisons as in Fig. 35, but at shorter elapsed time of 5000 MCS. The differences between bridging simulations and the fBm model are amplified. Parameters. $\mu = 80, \lambda = 10^5$ MCS **(B)** Same as in (**A**) but for a higher level of bridging. The overall tails get smaller and the difference between fBm and bridging simulations increases as compared to Fig. 35B. Parameters. $\mu = 120, \lambda = 80000$ MCS.

elapsed time used (Fig. 33,34). Note that, while bridging produces these outlier movements, overall it slows the dynamics of the polymer and therefore results in lower drift velocities. Consistent with this, the MSD and drift velocities of ΔH-NS were slightly greater than WT (Fig. 31A,34C). In our experimental data, we select a subset of outlier trajectories (39 of 6717 tracks) with drift velocity $v_d > 0.012 \mu m s^{-1}$ (Fig. 36A). The ensemble-averaged MSD curves show a transition to faster dynamics at longer timescales (Fig. 36B) indicating the presence of RCMs [57]. Strikingly, selecting a similar subset of faster tracks (117 of 5083 tracks) with $v_d > 3 \times 10^{-7} \mu m \text{MCS}^{-1}$ not captured by fBm (Fig. 36C) in our bridging simulations, we find similar MSD curves. The tracks with faster v_d transition to a higher exponent at longer time lags (Fig. 36D).

We explain the presence of RCMs in our simulations as being due to the heterogeneity in the bridging state of the tracked monomers. On timescales much longer than the mean bridge lifetime, each segment of the polymer is likely to be bridged for the same percentage of time and the dynamics are therefore relatively homogeneous. However on timescales less than the bridge lifetime, there is greater heterogeneity - some segments will remain bridged throughout, and others will remain unbridged. This results in a corresponding variation in the observed dynamics, with more strongly bridged segments displaying more restricted movements and shifting a proportion (the majority at high bridging) of the drift velocity distribution to lower values, leaving the unbridged segments as outliers. This heterogeneity cannot be captured by the fBm model, which models a single population. This interpretation is consistent with our finding that the fraction of RCMs increases with the degree of bridging and as the drift velocity is measured over shorter elapsed times (Fig. 37).

While we have not been able to quantitatively fit our polymer simulations to the experimental MSD curve and drift velocity distribution due to the increasingly computationally challenging of simulating longer bridge lifetimes and the higher number of simulations required to obtain accurate statistics of the RCMs, we nevertheless conclude that bridging by NAPs provides a potential explanation for both the sub-diffusive scaling of chromosomal loci as well the observed rare rapid chromosomal movements.

12

ESTIMATING BRIDGE LIFETIMES

In our simulations of bridging, we observed that the MSD curve transitions at a longer time lag to the exponent expected in the absence of bridging (≈ 0.56 in our model, lower if we were to account for viscoelasticity) (Fig. 38). Interestingly, we found a similar upward transition in the experimental MSD curve of *ori* (Fig. 31A). This was also observed in previous studies performed at the same (0.1 s) and longer time resolutions (1 s) and was associated with the RCMs discussed in the previous chapter [52, 53, 57]. While the cause of this transition is unclear and confounding effects of photobleaching and statistics cannot be completely discounted, this transition is most apparent for the terminus (*ter*) region in other studies [52, 53, 57]. Since the *ter* macrodomain is affected by NAPs (MukBEF in particular) differently than the rest of chromosome (resulting in a bias towards short-range genomic contacts) [62], this is consistent with the idea that bridging by NAPs is ultimately responsible.

In our simulations, the transition is clearly due to bridging not having an effect on the monomer dynamics at timescales much longer than the bridge lifetime. Indeed, we observe a linear relationship between the transition location (kink) and the bridge lifetime λ (Fig. 38B). We also note that while the transition was visible in the wild-type MSD curve it was invisible in the MSD curve of ΔH-NS strain (at least within the measured range) (Fig. 31A), which could be explained by this strain having a longer average bridge lifetime.

We next wondered whether the location of the transition could be used to infer a bound on the effective bridge lifetime. In this direction, we set the internal timescale of our simulations by matching the MSD of *ori* at a lag of 1 s. In particular, we found an MSD at 1 s of $\approx 0.004\,\mu m^2$ (Fig. 31A), consistent with previous measurements [52]. With a lattice spacing of 0.0056 µm, this MSD was reached at a lag of $\approx 10^5$ MCS (Fig. 31A). We choose this lattice spacing to have a sufficient spatial resolution at this displacement. By assigning 10^5 MCS ≈ 1 s, we can then convert our simulation results from MCS to seconds.

Doing this for all points in the phase diagram we obtain a relationship between two physical measures, the equilibrium percentage of monomers bridged and the bridge lifetime λ in seconds for different values of the scaling exponent (Fig. 38C). We have extended the contours of the fixed exponent to longer lifetimes by hand as it becomes increasingly computationally challenging to access longer bridge lifetimes (in seconds), especially for

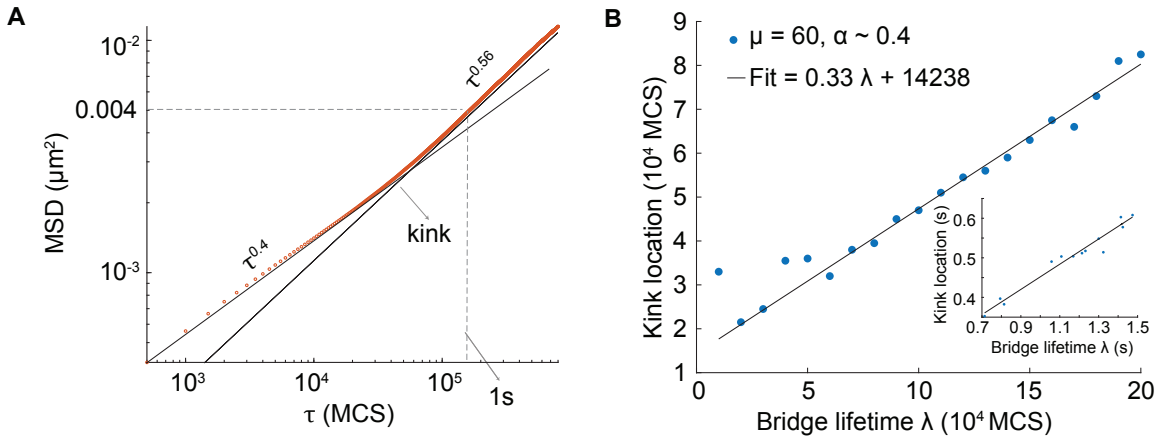

Figure 38: **(A)** Transition to a higher exponent. MSD curve transitions (at the 'kink' point) to a higher exponent $\alpha \sim 0.56$. **(B)** Location of the kink depends linearly on the bridge lifetime (λ). Inset shows the same plot after conversion to seconds.

the lowest exponents, due to the slow dynamics of the polymer (each second corresponds to an increasingly large number of MCS) (Fig. 40A, B). While the scaling exponent can be measured experimentally, the degree of bridging and the effective bridge lifetime are more challenging to quantify. Nevertheless, the relationship between these variables that we have

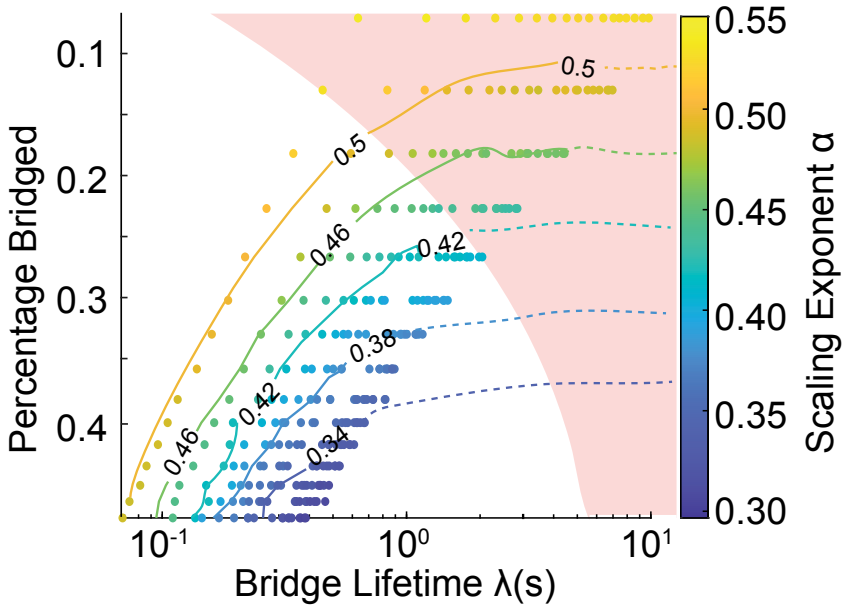

Figure 39: Phase diagram from Figure 29 with Bridge lifetime converted to seconds. The contours of α are overlaid. Dashed lines are estimated projections to higher values of λ. Shading indicates that the kink lies beyond 1 s.

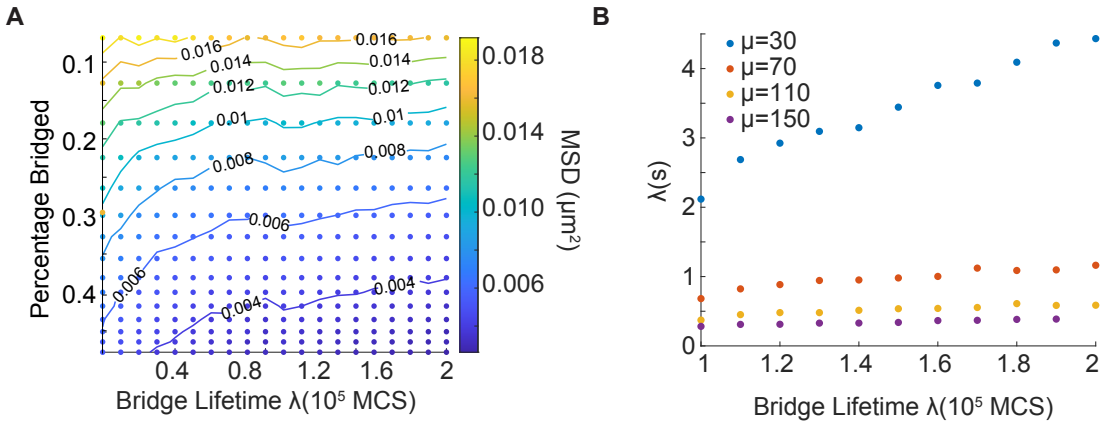

Figure 40: **(A)** Phase diagram of MSD at 25000 MCS for different parameters. MSD value reached at longer time delays decreases with increasing bridging. **(B)** Conversion between bridge lifetime λ in MCS to seconds. λ(s) increases linearly with λ (MCS) but the slope decreases sharply for higher levels of bridging. Hence, it is computationally challenging to access longer bridge lifetimes in seconds for higher levels of bridging.

uncovered here, should be useful in interpreting future experimental results and contributes to our understanding of chromosome dynamics.

Returning to the transition in the MSD curve, we can use the linear relationship to the bridge lifetime (Fig. 38B) and the conversion from MCS to seconds described above to obtain the location of the transition in seconds at each point in the phase diagram. In particular, we can identify the region of the phase diagram in which the transition occurs beyond a delay of 1 s, as seen in our data and other measurements. This results in the shaded region in Figure 39 and provides a lower bound for the bridge lifetime. For $\alpha \sim 0.32$, as observed for the WT in our experiments, we find a lower bound on the effective bridge lifetime of around 5 s, a reasonable estimate given the relatively slow dynamics of chromosomal loci. While measurements of bridge lifetimes of the various NAPs are lacking, estimates for H-NS and HU can be taken from the timescale of their recovery after photobleaching (FRAP) which gave 50 s [101] and 1 s [102] respectively.

13

EFFECT OF BRIDGING ON MACRODOMAIN POSITIONING AND ORGANISATION

In this chapter we discuss how bridging could affect the spatial-temporal organisation of the chromosome within the cell. As discussed in previous chapters, NAPs are highly abundant proteins that play diverse roles as chromosome organisers, transcription factors, and, more generally, involved in DNA mechanisms. A key player in chromosome organisation in *E. Coli* is the SMC complex, MukBEF also found in other gamma-proteobacteria. The absence of MukBEF leads to an increase in the number of anucleate cells and incorrect chromosome organisation. MukBEF appears to co-localize with *ori* at the middle or quarter positions (pre-division) on a cell. A Turing-type mechanism has been proposed [21] to explain MukBEF foci formation and positioning. The model predicts that MukBEF self-organises into foci at regular positions along the long axis of the cell. It was then hypothesized that MukBEF might play a role in *ori* (origin of replication) positioning, resulting in its faithful segregation. From the step-wise velocity of *ori* as a function of its position, it was observed that *ori* moves towards the mid-cell in a directed fashion [15]. Stochastic simulations of the *ori* and polymer simulations of the entire chromosome reveal further insights into *ori*'s movement to mid-cell and its unique relationship with MukBEF.

From Hi-C experiments [62], we find that *ter* (terminus of replication) is relatively decondensed compared to other chromosomal regions. This relative decompaction is hypothesized to arise from the absence of MukBEF at *ter* (which is unloaded by a protein MatP), resulting in a lack of long-range contacts. The emergence of this Hi-C picture is also of interest. Also, at the start of the cell cycle, *ter* moves from the new pole to the middle of the cell albeit for poorly understood reasons. A complete picture of the organisation of chromosomal domains (*ori*, *left*, *ter*, *right*) as a function of cell cycle progression is still elusive.

13.1 *ori* POSITIONING AND DYNAMICS

We performed explicit polymer simulations to see how bridging affects the positioning of *ori* (Fig. 41A). One can see that pure elastic fluctuations in combination with bridging of

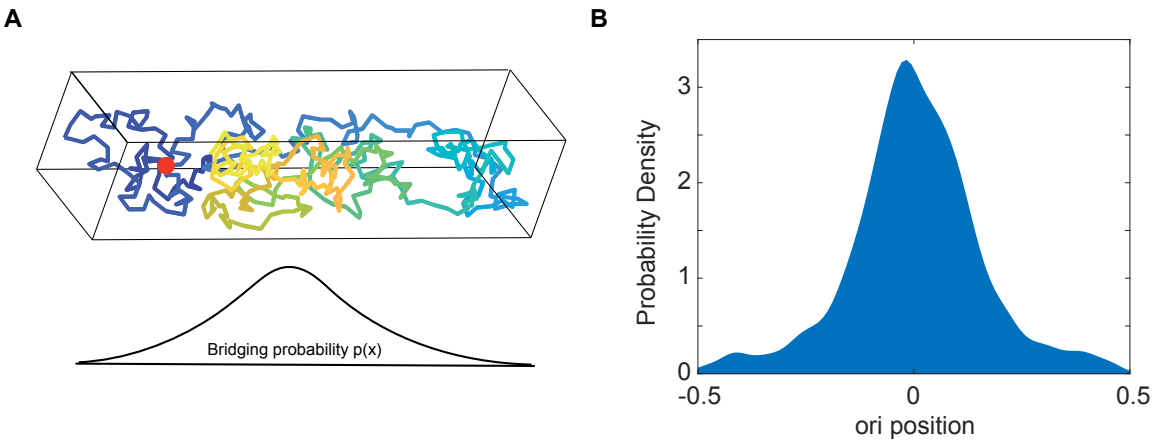

Figure 41: **(A)** Snapshot of polymer conformation with $N = 440$ monomers. The red dot represents the *ori*. **(B)** Probability density of relative *ori* position (from the mid-cell) along the long axis of the cuboid.

monomers due to spatially varying MukBEF concentration, results in a directed motion of *ori* towards mid-cell. We find that the probability density of *ori* is a function of its relative position to the mid-cell (Fig. 41B). As expected, the probability of finding the *ori* at mid-cell is maximal. Next, we ask whether the spatially-varying bridging probability $p(x) = p_{max} \exp(-x^2)$ results in the directed movement of *ori*. We examine the step-wise velocity of *ori* as a function of its position (Fig. 42A) and find that the *ori* indeed experiences a restoring velocity towards mid-cell, i.e., directed movement. We visualize this as follows, on a short timescale, the *ori* fluctuates about its current *home* position. These fluctuations enable the *ori* to locally sample the spatially-varying bridging probability. It is then most likely to form a bridge with another monomer in the position in which the bridging probability is the greatest. The polymer then relaxes, the *ori* is released at a new *home* position and the cycle repeats. Here, we have a model in which the elastic fluctuations of the polymer power the movement of *ori* up the gradient.

The dynamics of *ori* approximated to that of a particle diffusing in a harmonic potential centered at mid-cell. Let the strength of the potential be f with diffusion constant D. Let us consider a harmonic potential $U = \frac{1}{2} f x^2$ over an infinite 1D domain. Given a particle at position x_0, the probability density that it is at position x at time Δt later is [49],

$$p(x, \Delta t \mid x_0) = \sqrt{\frac{f/k_B T}{2\pi S}} \exp\left[-\frac{f/k_B T}{2S} \left(x - x_0 e^{-\Delta t/\tau} \right)^2 \right], \tag{13.1}$$

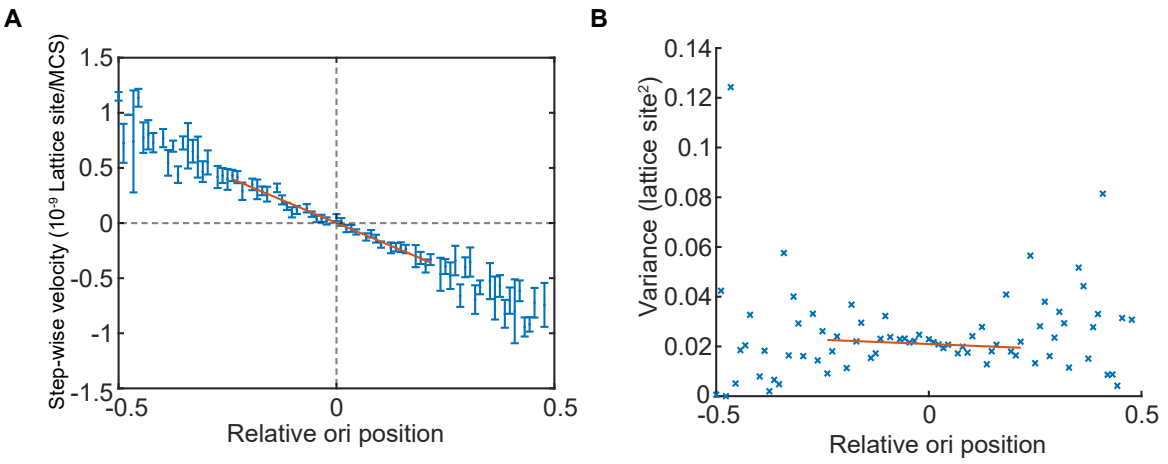

Figure 42: **(A)** The step-wise velocity of *ori* plotted as a function of position, shows that *ori* moves in a directed fashion towards the middle of the cell. Linear fit (orange line) calculates the timescale τ. **(B)** Variance of relative *ori* positions. Linear fit (orange line) defines the diffusion constant D.

where $S = 1 - e^{-2\Delta t/\tau}$ and $\tau = \frac{k_B T}{fD}$. We can calculate the expected value and variance of the step-wise position $(x - x_0)$:

$$E[x - x_0] = \frac{e^{-\frac{\Delta t}{\tau}} - 1}{\Delta t} x_0 \approx -\frac{x_0}{\tau}. \tag{13.2}$$

$$V[x - x_0] = \frac{D\tau}{\Delta t^2}(1 - e^{-\frac{2\Delta t}{\tau}}) \approx \frac{2D}{\Delta t}, \tag{13.3}$$

where the above inequality holds for $\frac{2\Delta t}{\tau} \ll 1$. Importantly, the expected value of step-wise velocity depends linearly on position, while the variance is independent of it. This is observed in our polymer simulations as we find a linear dependence of step-wise velocity in Figure 42A and variance in Figure 42B.

Simulations

Unlike in the previous chapters, in order to understand the positioning of the macrodomains we have to simulate a polymer to represent the entire chromosome rather than a cross-section. Following Hoffmann et.al, we model the *E. coli* chromosome as a polymer and simulate it using the Bond Fluctuation Monte-Carlo method. We consider a polymer of length $N = 440$, with each monomer representing around 10 kb region of the chromosome. This is confined to a cuboid of dimensions $88 \times 22 \times 22$ (this

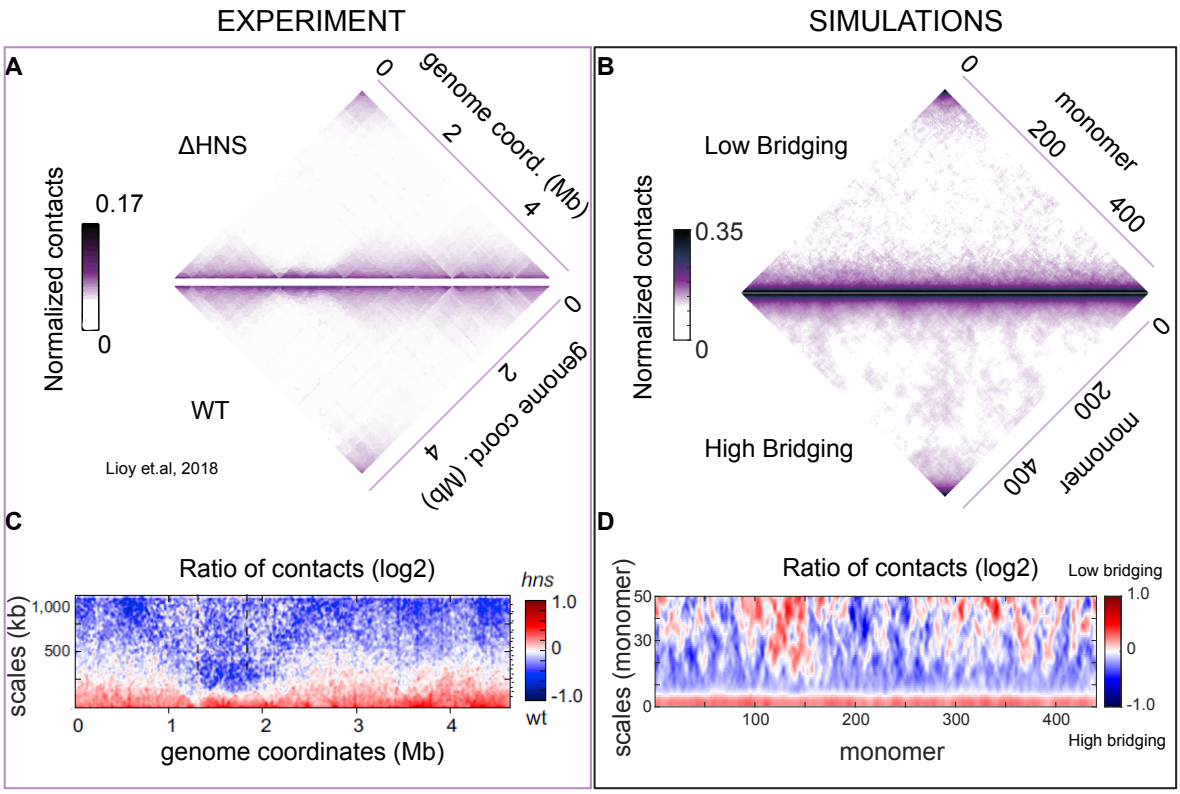

Figure 43: **(A)** HiC maps from experiments comparing mutant ΔH-NS and WT strain [62]. (**B**) HiC maps from simulations comparing contacts of a circular polymer in the absence of bridging and in presence of it. Parameters. Threshold on HiC contacts = 8 lattice units, $\mu = 20, 60$ with $\sim 15, 30\%$ of monomers bridged respectively, bridge lifetime $\lambda = 80000$ MCS. **(C)** Ratio plot of the mutant ΔH-NS and WT contact signals for each 5kb bin along the chromosome. The x-axis indicates the genomic coordinate and the y-axis indicates the genomic distance from the bin. A decrease or increase in contacts in the mutant cells compared with the WT is represented with a blue or red colour, respectively. White indicates no differences between the two conditions.**(D)** Same plot as (**C**) but for the simulated polymers in (**B**) with two levels of bridging.

leads to a system density of about 10%). To this, we add a spatially varying bridging probability (along the long axis) based on a Gaussian centred at mid-cell. Only a single monomer (*ori*) in our polymer is allowed to form bridges with other monomers.

13.2 Effect of bridging on chromosome organisation

Recently, the higher-order organisation of *E. coli* chromosome was explored and revealed several factors affecting its compaction [62]. The analysis using high resolution (5kb) chromosome capture (3C) contact maps of the *E. coli* chromosome reveals an interesting structure. The NAPs HU and MukBEF are observed to promote long-range contacts while

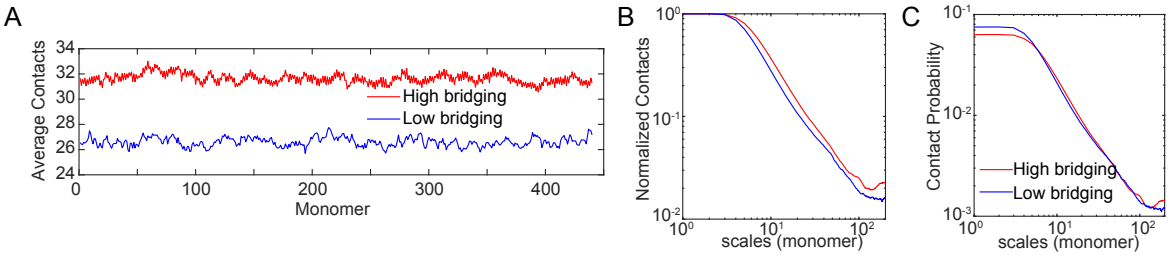

Figure 44: **(A)** Absolute average contacts for each monomer at different levels of bridging (43B). **(B)** Average contacts for different monomer distances (scales) shows that contacts decrease across all scales with bridging.**(C)** Contact probability is plotted for different levels of bridging in 43B. There is relative enrichment of short-range contacts in the low bridging case and hence a relative decrease in long-range contacts.

H-NS acts locally. We attempt to qualitatively compare the contact maps from our polymer simulations with that of the experiments. The experimental HiC maps display enrichment of contacts across nearest neighbours in agreement with the transverse alignment of the *E. coli* chromosome 43A (Lower triangle). The absence of the secondary diagonal reflects the lack of contacts between different chromosomal arms as is observed in *B. subtilis* [103].

Long-range contacts decrease in the absence of bridging

In the absence of a NAP H-NS, which is expected to bridge the DNA uniformly (refer to previous chapters), the Hi-C map superficially appears similar to that of the wild type 43A. We perform polymer simulations of a circular polymer as discussed in the previous section, but now with uniform bridging probability for all monomers. We generate Hi-C maps from simulations with different levels of bridging and find that the overall contact maps remain conserved as in the experiments 43B. Any two monomers are considered to be in contact if their spatial distance is less than an arbitrary threshold of 8 lattice units. We notice that while the threshold affects the exact scales of contacts it does not change the qualitative results. To further investigate the contact maps, we employ a visualization tool called the *scalogram* which represents, for each bin (monomer in the simulations), the cumulative contact frequencies as a function of the genomic distance. A scalogram, therefore, displays the average distribution of contacts for each bin (monomer) with its flanking regions. Performing a ratio of the scalograms or contact signals between the mutant ΔH-NS and WT and taking the logarithm we find significant enrichment in short-range contacts of H-NS binding regions 43C, where blue or red represent a decrease or increase in contacts, while white represents no change. Interestingly, calculating the ratio of contacts in our simulations we find a striking qualitative agreement with experiments. We also find a relative enrichment of short-range contacts and a decrease in long-range contacts. This is

consistent with previous experiments arguing that NAPs promote long-range contacts. This trend was also observed in ratio plots of other NAPs like Fis and HU [62].

SCN Normalisation

Sequential Component Normalization (SCN) is a simple normalization procedure to process the data and allow the generation of a normalized, highly contrasted, chromosomal contact map [104]. This technique corrects for various experimental biases that lead to the over-representation of certain genomic contacts. The normalization gives equal weight to each fragment in the contact map and hence before performing it we remove the low-probability contacts (tails of the Gaussian). The technique belongs to a class of approaches based on the Sinkhorn-Knopp balancing algorithm [105]. We transform a symmetric non-negative matrix of raw contacts $\boldsymbol{C} = (c_{ij}), \boldsymbol{C} \rightarrow R_{+}^{n \times n}$, into a stochastic matrix $\boldsymbol{S} = (s_{ij})$, whose rows and columns are alternatively divided by the euclidean norm to sum to 1 i.e., $\sum_i s_{ij} = \sum_j s_{ij} = 1$. It is an iterative process, that consists of solving,

$$\boldsymbol{S} = \boldsymbol{D_1 C D_2}, \tag{13.4}$$

where D_1 and D_2 are unique up to a scalar factor diagonal matrices obtained from alternatively normalizing columns and rows to 1. The whole process is repeated sequentially until the contact map converges into a symmetric matrix. Usually, a couple of iterations are sufficient to ensure convergence.

While scalograms are very useful in visualizing chromosome organisation, one has to be careful in interpreting the results. To be able to compare the experimental data with each other we have to employ a normalization procedure. Following Lioy et.al, we use the SCN method to generate our HiC maps. Hence, we only capture the relative increase or decrease in chromosomal contacts and not the absolute change.

Experimental ratio plots in ΔH-NS 43C and for similar plots of other NAPs in Lioy et.al, we find a relative enrichment of short-range contacts and a relative decrease in long-range contacts. It is important to remember that in the absence of NAPs bridging, chromosomal contacts are expected to decrease across the chromosome leading to an overall decompaction. We see this clearly in our simulations where the average number of contacts decreases for lower levels of bridging for all monomers and across different scales (44A, B). The relative contacts visualized by the contact probability show that for a polymer with low levels of bridging there is a relative increase at short polymeric scales and a relative decrease at longer scales(44C). This is consistent with the experimental scalogram in Figure 43B.

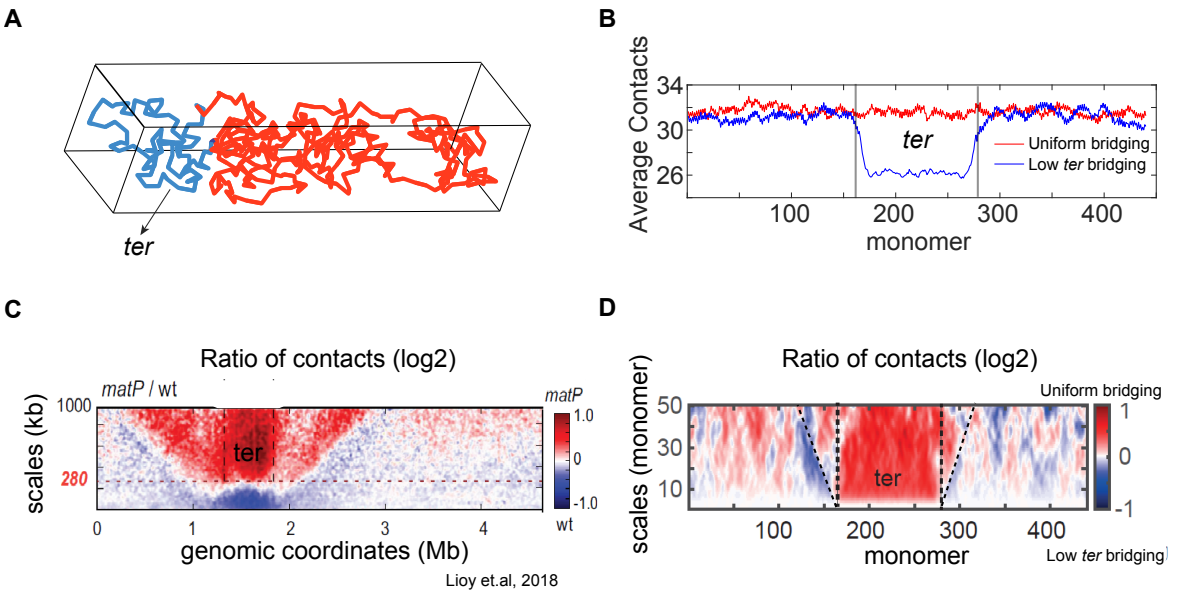

Figure 45: **(A)** Snapshot of polymer simulations with *ter* region labelled in blue. The *ter* region is 110 monomers long. **(B)** Average contacts across monomers for a polymer with uniform bridging and with lower levels bridging at *ter*. Parameters. $\mu = 60$ everywhere, $\mu_t = 20$ at *ter*. $\lambda = \lambda_t = 80000$ MCS everywhere. **(C)** Ratio plot between mutant ΔMatP and wild type as in 43C [62]. (**D**) Ratio plot from polymer simulations between the case of uniform bridging (WT) and lower bridging at *ter* as in (**B**). Shows qualitative agreement with experiments in (**C**).

13.3 *ter* ORGANISATION AND DYNAMICS

Experiments reveal that *ter* macrodomain in *E. coli* exhibits different dynamics and organisation relative to other chromosomal regions. The SMC complex MukBEF is kicked off from *ter* by a protein MatP which binds to specific *matS* sites in *ter* [47]. HiC experiments reveal interesting structures in the absence of MatP where the structure of chromosome contacts was conserved across all genomic regions, except in *ter*, which now appears similar to the rest of the genome 45C [62]. An enrichment in long-range contacts within *ter* (beyond 280 kb) and its flanking regions appeared, in conjunction with a compensatory decrease in contacts under $\sim$ 280kb within *ter*. These results are consistent with recombination assays that show enrichment of long-range contacts at *ter* in the absence of MatP [92].

We wondered if our simple bridging model could capture the interesting structure of the ratio plot between ΔMatP and wild type in Figure 45C. We know that bridging and compaction at *ter* is suppressed by the presence of MatP. Hence, we simulate a polymer with uniform bridging everywhere other than at a subset of monomers ($1/4N = 110$ monomers) (Fig. 45A, B). These monomers represent the *ter* macrodomain which is around 1000 kb long, which is $\sim 1/4$ the size of the *E. coli* chromosome.

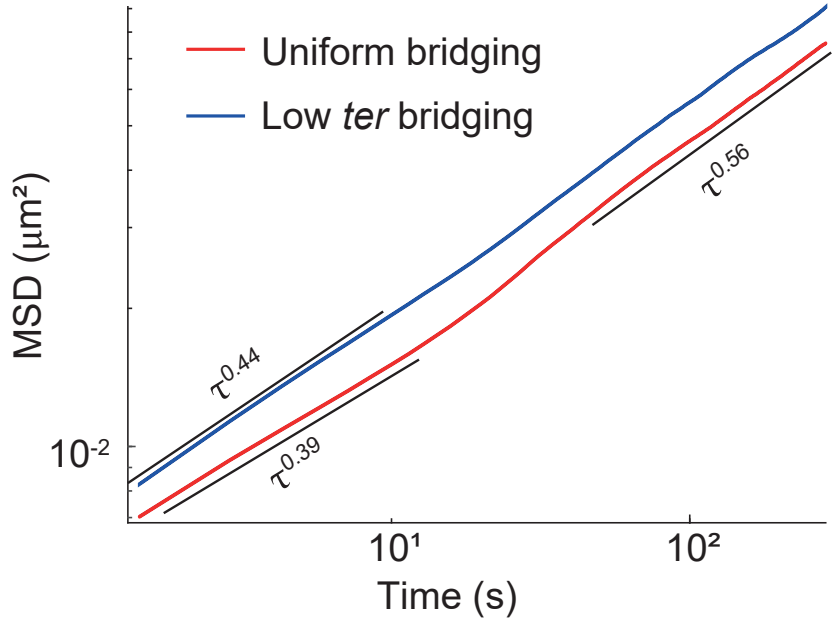

Figure 46: MSD curves, comparing *ter* with rest of the chromosome. *ter* appears more mobile than the rest of the chromosome due to lower bridging than the rest of the chromosome. $\mu = 60, \mu_{ter} = 20, \lambda = \lambda_t = 20000$ MCS.

As expected, unlike the case of uniform bridging, we find a reduction in contacts in the *ter* region (between monomers 165 and 275) of the polymer (Fig. 45B). Interestingly, this was sufficient to reproduce the ratio plots from experiments in Figure 45C, as the MatP mutant shows a relative enrichment only at longer length scales in the *ter* region. Note that in the experimental ratio plot, there exists a diagonal around the *ter*, indicating enrichment of interactions between *ter* and its flanking regions. This was also observed in our polymer simulations with similar diagonal indicating contact with flanking regions around the *ter* (Fig. 45D).

We next wondered how lower levels of bridging affect *ter* dynamics. Absence of bridging results in increased *ter* mobility as seen in the MSD curves in Figure 46. We immediately notice that at short time lags *ter* has a higher scaling exponent ~ 0.44 as compared to other loci with $\alpha \sim 0.39$. This is in contrast to experimental data from Javier et.al, where it was observed that while the scaling exponents don't change significantly between different loci, *ter* has a lower apparent diffusion constant D_{app} than *ori* (MSD curve of *ter* lies below *ori*). While this difference between simulations and experiments could be simply due to the failure of our coarse-grained model to capture *E. coli* chromosome dynamics and organisation simultaneously, other biological effects might be in play. In a temperature-sensitive non-replicating strain, *ter* ($\alpha \sim 0.33$) appears more mobile than *ori* ($\alpha \sim 0.31$) [54].

The post-replicative pairing of sister chromosomes has also been suggested as a possible explanation for the differences in mobility between *ori* and *ter* [55].

The qualitative comparisons in this chapter are crucial first steps to developing a unifying framework to understand chromosome dynamics and organisation in conjunction. In future, we wish to perform more detailed simulations with appropriate coarse-graining to perform quantitative comparisons to experimental data.

14

DISCUSSION AND FUTURE DIRECTIONS

Our results provide insight into the role of DNA bridging by NAPs in determining chromosome dynamics and compaction within bacterial cells. We have shown that bridging, at realistic levels and lifetimes, can explain the sub-diffusive scaling exponent of bacterial chromosomal loci. In particular, our model predicts that a decrease in bridging leads to an increase in the sub-diffusive exponent α and we confirmed this by tracking the *ori* locus of *E. coli* in a strain deleted of the NAP H-NS. We also addressed the upturn in the ensemble-averaged MSD curve seen at long-time lags. Our model displays a similar transition at a timescale of the order of the average bridge lifetime and we obtained a lower bound on the effective bridge lifetime of the *ori* locus of about 5 s. We also show that bridging can qualitatively explain the nature and corresponding changes in chromosomal contact maps in the presence and absence of NAPs. Interestingly, we could reproduce qualitatively the ratio plot of the MatP mutant and the WT.

Bridging can also qualitatively reproduce the rare but ubiquitous rapid chromosomal movements (RCMs) that are observed within experimental trajectories, in contrast to the null phenomenological model of fractional Brownian motion (fBm). The RCMs in our model are due to the rare event of a DNA segment being unbridged by NAPs for long enough that it exhibits an unusually large movement compared to the rest of the slower-moving bridged polymer. This is consistent with the proposal by Javer *et al.* that RCMs arise due to the relaxation of stress caused by the action of bridging proteins and condensins [57, 106]. Therefore, our model explains both the sub-diffusive scaling exponent of chromosomal loci and the observation of RCMs. Furthermore, in contrast to the hypothesis of a viscoelastic cytoplasm [50], bridging is consistent with recent work showing that cell compression lowers the exponent of chromosomal loci but not that of diffusive particles [53]. A lower cell volume increases the density of DNA and therefore increases the rate of bridge formation, lowering the exponent (Fig. 26B). Bridging can also qualitatively reproduce the organisational features of chromosome. In agreement with experiments, in the absence of the uniformly bridging NAP H-NS we find a relative enrichment of chromosomal contacts at short genomic scales and a corresponding decrease at longer scales. Interestingly, we can also reproduce the profile of the contact ratio plot between MatP mutant and WT that only affects the

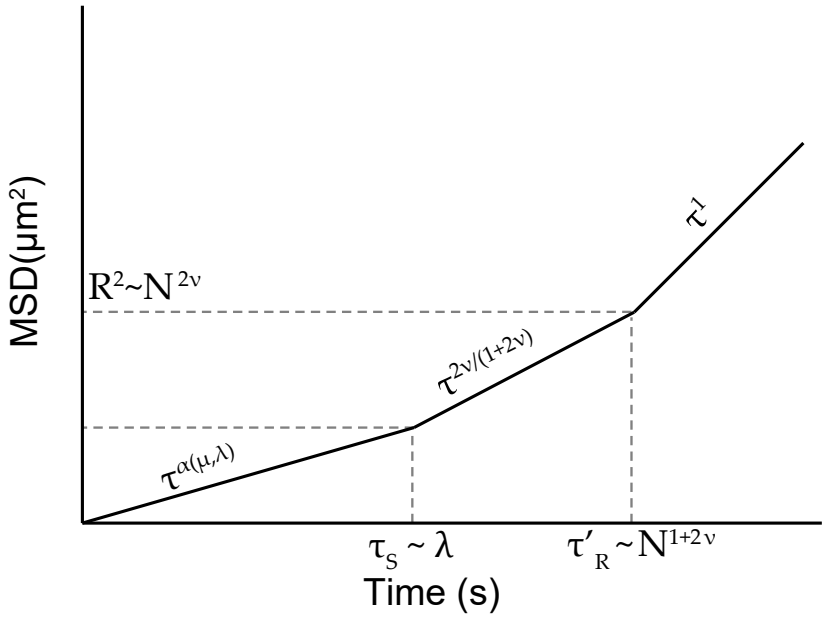

Figure 47: A schematic representation of MSD scaling with delay in the presence of bridging. Unlike a simple rouse polymer, there exists an additional sub-diffusive timescale τ_S at short time lags determined by bridge lifetime.

relative levels bridging at the *ter* macrodomain. More broadly, by characterising the relation between an equilibrium quantity, the number of bridges, and a dynamic quantity, the mean bridge lifetime (λ), our framework provides an intuitive parameter landscape for bacterial chromosome dynamics that will help guide future studies.

We summarize our inference of the empirical scaling behaviour of loci dynamics in Figure 47. At time lags shorter than the average bridge lifetime λ, we find that the scaling exponent depends on the parameters μ and λ (Fig. 47). At intermediate time lags ($\lambda < \tau < \tau'_R$) in our polymer simulations, we find the typical rouse scaling of $\alpha \sim 0.56$ for a monomer on a self-avoiding polymer. We know from polymer theory that τ'_R is the timescale in which the polymer moves a distance R (the end-to-end distance of the entire polymer). The scaling $R \sim N^{\nu}$ and hence, the corresponding MSD at this displacement is $\sim N^{2\nu}$. We know from Eq. (8.18) that $\tau'_R \sim N^{1+2\nu}$. Hence, the slope of the MSD curve for times. At time lags longer than the modified relaxation time τ'_R (discussed below), one expects a scaling exponent $\alpha = 1$, indicative of purely diffusive motion. Finally, we conclude that a new theoretical model describing the interesting scaling dynamics of the bridging polymer is of interest.

A

MASS MINIMISATION AND PATTERN SELECTION IN THE BRUSSELATOR MODEL

We present an analysis of the Brusselator model and show numerically and analytically that most of the features of the exploratory model in Eq. (1.34) still hold. The general spatial version of the Brusselator [22] is described by the following equations,

$$\frac{\partial u}{\partial t} = D_u \frac{\partial^2 u}{\partial x^2} - \beta u v^2 + \gamma v \tag{A.1a}$$

$$\frac{\partial v}{\partial t} = D_v \frac{\partial^2 v}{\partial x^2} + \beta u v^2 - \gamma v + c\delta - \delta v \, . \tag{A.1b}$$

We use reflective boundary conditions at $x = \pm\frac{L}{2}$ and as in Eq. (1.34) write the production term in terms of the degradation rate δ. While the Brusselator also has the form of a mass-conserving Turing system with additional linear terms, here it is the mass of v that is fixed at the steady state,

$$\frac{1}{L}\int \bar{v}\, dx = c \, , \tag{A.2}$$

rather than the total mass. As before, we can change the turnover δ without affecting the steady-state concentration of v. We non-dimensionalise the system by

$$x \to \frac{x}{L}, t \to \frac{D_v t}{L^2}, u \to \frac{u}{c}, v \to \frac{v}{c}$$

to obtain

$$\frac{\partial u}{\partial t} = d\frac{\partial^2 u}{\partial x^2} + \Gamma\Big(- a u v^2 + v\Big) \tag{A.3a}$$

$$\frac{\partial v}{\partial t} = \frac{\partial^2 v}{\partial x^2} + \Gamma\Big(a u v^2 - v + b(1 - v)\Big) \tag{A.3b}$$

where

$$\Gamma = \frac{\gamma L^2}{D_v}, d = \frac{D_u}{D_v}, a = \frac{\beta c^2}{\gamma}, b = \frac{\delta}{\gamma} \, .$$

There is a single fixed point

$$u_0 = \frac{1}{a}, v_0 = 1 \, .$$

The Jacobian is given by,

$$J_{(u_0,v_0)} = \Gamma \begin{bmatrix} -av_0^2 & -2au_0v_0 + 1 \\ av_0^2 & 2au_0v_0 - 1 - b \end{bmatrix} = \Gamma \begin{bmatrix} -a & -1 \\ a & 1 - b \end{bmatrix} \, .$$

The Jacobian (and hence the dispersion relation) becomes independent of b for $b \ll 1$, which is when the flow of mass through the system is low, close to the mass-conserved limit $b = 0$. Hence, we can change b without significantly affecting the linear behaviour of the model. The trace and determinant of the Jacobian are easily found to be

$$\begin{aligned} TrJ &= \Gamma(1 - a - b) \\ DetJ &= \Gamma^2 ab \, . \end{aligned}$$

For the homogeneous fixed point to be stable in the absence of diffusion we must have $TrJ < 0$ and $DetJ > 0$. Hence, we require $a + b > 1$. The Turing condition is given by,

$$d(1 - b) - a - 2\sqrt{dab} > 0.$$

We numerically solve this system, using reflective boundary conditions, by perturbing the homogeneous state as described in numerical methods. Like our illustrative model Eq. (1.36), and every Turing model we are aware of, the peaks of a pattern are periodic and regularly positioned. Furthermore, consistent with our results, a single peak moves exponentially to the mid-domain (Figure S6A). The rate of movement was found to be proportional to b, or equivalently, δ, the turnover rate and at the mass-conserved limit $b = 0$ no peak movement is observed (as was the case for our model).

A.1 SPIKE LIMIT IN THE BRUSSELATOR MODEL

In this section, we develop the spike limit of the Brusselator model [25, 107]. Let us consider the dimensional form in Eq. (A.1). As in the case of our toy model, we consider the limit $D_v \ll Du$, where solutions of v consist of narrow large-amplitude spikes of width $\varepsilon = \mathcal{O}(\sqrt{D_v/\gamma})$. Away from the spikes, v is a spatial constant v_{out}. Since $\int_{-L/2}^{L/2} v dx = c$, inside the spikes we have $v \gg c$ within a spike and $v = v_{out} \ll c$ outside. We search for steady-state solutions consisting of n spikes at positions $x_1, x_2, ..., x_n$ and assume that u changes slowly within each spike and can be approximated by a constant u_i. We introduce

an inner coordinate $y = \varepsilon^{-1}(x - x_i)$, within each spike. The equation for the inner variable $v_i(y)$ becomes,

$$\frac{D_v}{\varepsilon^2}\frac{d^2 v_i}{dy_i^2} + \beta u_i v_i^2 - (\gamma + \delta) v_i = 0$$
$$v_i \to 0 \quad \text{as} \quad y_i \to \pm\infty \,,$$

where we have neglected the constant term since $v_i \gg c$. The solution to this equation is

$$v_i = \frac{3}{2}\frac{\gamma + \delta}{\beta u_i}\text{sech}^2\left(\sqrt{\frac{\gamma + \delta}{D_v}}\frac{\varepsilon y_i}{2}\right) .$$

In the outer region, the equation of each spike is approximated by a Dirac delta function

$$v = v_{out} + \sum_i w_{i,1}\delta(x - x_i) \,, \tag{A.4}$$

where $w_{i,1}$ is the weight

$$w_{i,1} = \varepsilon \int_{-\infty}^{\infty} v_i(y_i) dy_i = 6\frac{\sqrt{D_v(\gamma + \delta)}}{\beta u_i}$$

The ε pre-factor is the spike width. To write the outer equation for u, we also need the weight of the uv^2 term

$$w_{i,2} = \varepsilon u_i \int_{-\infty}^{\infty} v_i^2(y_i) dy_i = 6\frac{\sqrt{D_v}(\gamma + \delta)^{3/2}}{\beta^2 u_i}$$

Since, $\mathcal{O}(w_{i,1}) = 1$ (spikes have finite weight), we find $\mathcal{O}(u_{in}) = \varepsilon$ and $\mathcal{O}(v_{in}) = \varepsilon^{-1}$. Away from the spikes, v is a constant v_{out}. Therefore from Eq. (A.1b) we have,

$$\beta u v_{out}^2 - \gamma v_{out} + c\delta - \delta v_{out} = 0 \tag{A.5}$$

in the outer region. Balance this equation while requiring u and v_{out} not to diverge requires that $\mathcal{O}(u) < \mathcal{O}(v_{out})$. Therefore we can neglect the cubic term in the above equation to obtain

$$v_{out} = \frac{c\delta}{\gamma + \delta}.$$

We obtain the outer equation for u by replacing v with weighted delta functions,

$$\begin{aligned} D_u \frac{d^2u}{dx^2} - \sum_{i=1}^{n} [\beta w_{i,2} - \gamma + w_{i,1}]\delta(x - x_i) \\ + c\delta - \delta v_{out} \\ -\tfrac{L}{2} < x < \tfrac{L}{2}, \;\; u_x(\pm\tfrac{L}{2}) = 0, \end{aligned} \tag{A.6}$$

where we used Eq. (A.5) to simplify contributions from the outer region. Inserting the expressions for each term we arrive at,

$$\begin{aligned} D_u \frac{d^2u}{dx^2} - \frac{c\gamma\delta}{\gamma+\delta} - \sum_{i=1}^{n} \frac{\rho}{u_i} L\delta(x - x_i) = 0 \,, \\ -\tfrac{L}{2} < x < \tfrac{L}{2}, \;\; u_x(\pm\tfrac{L}{2}) = 0 \,. \end{aligned} \tag{A.7}$$

where $\rho = 6\frac{\sqrt{D_v}}{L}\frac{\delta\sqrt{\gamma+\delta}}{\beta}$. Note the above expression for v_{out} is consistent with the condition for the integral of v, $Lv_{out} + \sum_i w_{i,1} = cL$ and the integral of the outer equation.

The solution to the outer equation is

$$u(x) = \bar{u} - \sum_{j=1}^{n} \hat{\rho}_i' \hat{G}(x; x_j), \tag{A.8}$$

where $\bar{u}$ is a constant and the Green's function $\hat{G}(x; x_j)$ is the solution to

$$\begin{aligned} -L^2 \hat{G}_{xx}(x; x_j) + 1 = L\delta(x - x_j), \quad -\tfrac{L}{2} < x < \tfrac{L}{2}, \\ \hat{G}_x(\pm\tfrac{L}{2}) = 0, \quad \frac{1}{L}\int_{-\frac{L}{2}}^{\frac{L}{2}} \hat{G}(x; x_j)dx = 0 \end{aligned} \tag{A.9}$$

given by

$$\hat{G}(x; x_j) = \tfrac{1}{2L^2}(x^2 + x_j^2) - \tfrac{1}{2L}|x - x_j| + \tfrac{1}{12} \tag{A.10}$$

We use hats to distinguish these quantities from those of the model in Eq. (1.36). The coefficients $\hat{\rho}_i' = \hat{\rho}_i'(\mathbf{x})$ and the constant $\bar{u}$ are determined by the non-linear algebraic system

$$\hat{\rho}_i' = \hat{\sigma}\frac{c^2}{u(x_i)} \quad i = 1, ..., n \tag{A.11}$$

where $\hat{\sigma} = \frac{\rho L^2}{c^2 D_u}$ is a dimensionless parameter and the consistency condition (obtained by integrating equation Eq. (A.7)),

$$\sum_i \hat{\rho}_i' = c\eta. \tag{A.12}$$

where $\eta = \frac{\delta}{\gamma+\delta}\frac{\gamma L^2}{D_u}$ is a second dimensionless parameter.

We next consider the spike limit $\hat{\sigma} \to 0$ ($D_v \to 0$). From equation Eq. (A.11) we have

$$\hat{\rho}_i'(\bar{u} - \sum_{j=1}^{n} \hat{\rho}_i' \hat{G}(x_i; x_j)) = 0 \quad \forall \ i.$$

As in the model of the main text, if any $\hat{\rho}_i' = 0$, then the result is a solution to the system with one fewer spike. Hence the only physical n-spike solution is

$$\hat{\boldsymbol{\rho}}' = \bar{u}\hat{\mathbf{G}}^{-1}\hat{\mathbf{e}} \tag{A.13}$$

where $\hat{\mathbf{G}}_{ij} = \hat{G}(x_i; x_j)$ and $\hat{e}$ is the column vector with all unit entries. The constant $\bar{u}$ is determined from equation Eq. (A.12).

A.2 Peak movement and comparison to point sinks

Just as for the model of the main text, the solution obtained above is identical to that of a model of perfect point sinks. Consider the following system of (imperfect) point sinks

$$D\frac{d^2A(x)}{dx^2} + \hat{c} - \sum_{i=1}^{n} \mu L \delta(x - x_i) A(x_i) = 0 \tag{A.14}$$

with reflective boundary conditions at $x = \pm\frac{L}{2}$. This is similar to that of equation Eq. (3.1) but without the background decay (δu term). Indeed it can be obtained from that equation by replacing c by $\hat{c}/\delta$ and taking the limit $\delta \to 0$. The solution to this equation is

$$A(x) = \bar{A} - \sum_{i=1}^{n} \hat{\mu}_i' \hat{G}(x; x_i). \tag{A.15}$$

where $\bar{A}$ is a constant and the Green's function $\hat{G}(x; x_i)$ is the same as that of the Brusselator Eq. (A.10). The coefficients $\hat{\mu}_i' = \hat{\mu}_i'(\mathbf{x})$ are determined by the linear algebraic conditions

$$\hat{\mu}_i' = \hat{\lambda} A(x_i) \quad i = 1, ..., n \tag{A.16}$$

where we have introduced the dimensionless parameter $\hat{\lambda} = \frac{\mu L^2}{D}$. In the perfect sink limit $\hat{\lambda} \to \infty$ the condition for the $\hat{\mu}_i'$'s reduces to

$$\hat{\boldsymbol{\mu}}' = \bar{A}\hat{\mathbf{G}}^{-1}\hat{\mathbf{e}} \tag{A.17}$$

which is precisely what we obtained for the Brusselator model in the spike limit. Thus, steady-state spike solutions of the Brusselator are identical to solutions of a model of perfect point sinks.

In the main text, for the point sink model *with* background decay (δu), we showed a connection (Eq. (3.29)) between the flux differentials into each sink and the derivative of the total mass $M = \frac{1}{L}\int A(x)dx$. It is easily seen that the same relation holds for the model *without* background decay. Replacing c by $\hat{c}/\delta$ and taking the limit $\delta \to 0$, the flux differential across each sink is found from Eq. (3.27) to be

$$\frac{\Delta J_i}{\hat{c}L} = -\frac{1}{L}\left[x_i - \frac{L}{n}i + \frac{L}{2}(\frac{1}{n}+1)\right].$$

Similarly, the derivative of M from Eq. (3.28) is

$$\frac{D}{\hat{c}L}\frac{\partial M}{\partial x_i} = \frac{2}{nL}\left[x_i - \frac{L}{n}i + \frac{L}{2}(\frac{1}{n}+1)\right]$$

Thus

$$\Delta J_i = -\frac{1}{2}nD\frac{\partial M}{\partial x_i}\,. \tag{A.18}$$

Therefore, the dynamics of sinks moving with velocities proportional to the flux differential across them is equivalent to a system of n over-damped particles with the mass M acting as a potential.

A.3 Mass minimisation predicts the pattern obtained after coarsening

In Figure 17, we examine coarsening in the Brusselator model. We use Γ, d, a and b as dimensionless parameters and use periodic boundary conditions to avoid patterns with peaks on the boundary (which are not compatible with the spike approximation). In panels A and B, we compare the number of peaks in the dominant mode as predicted by the linear dispersion relation and the number of peaks in the most frequent pattern as obtained from the numerical simulations. We observe a similar coarsening behaviour as b (or equivalently the turnover rate δ) is decreased, similar to the model of the main text.

To apply our analytical approximation, we consider the situation where spikes are symmetric ($\hat{\rho}'_i = \hat{\rho}'_c$) and regularly positioned,

$$\bar{x}_i = \tfrac{L}{n}i - \tfrac{L}{2}(\tfrac{1}{n}+1)$$

as they are observed to be in a steady state. From Eq. (A.12), we then have

$$\hat{\rho}'_c = \frac{c}{n}\eta \ . \tag{A.19}$$

We determine the constant $\bar{u}$ by evaluating the solution Eq. (A.8) at each spike position $\bar{u}_i$

$$u(\bar{x}_i) = \frac{n\hat{\sigma}}{\eta}c = \bar{u} - \frac{c\eta}{n}\sum_{j=1}^{n}\hat{G}(\bar{x}_i;\bar{x}_j).$$

The sum in the above equation is independent of i since

$$\sum_{j=1}^{n}\hat{G}(\bar{x}_i;\bar{x}_j) = \frac{1}{12n} \ .$$

Thus, we find that the total mass M of the fast species is

$$\begin{aligned} M/c &= \frac{1}{cL}\int_{-\frac{L}{2}}^{\frac{L}{2}} u(x)dx = \frac{1}{cL}\int_{-\frac{L}{2}}^{\frac{L}{2}} \bar{u}\ dx \\ &= \frac{n\hat{\sigma}}{\eta} + \frac{\eta}{12n^2}, \end{aligned} \tag{A.20}$$

where we have used the Green's function property $\int_{-\frac{L}{2}}^{\frac{L}{2}}\hat{G}(x;x_j) = 0$. In terms of the dimensionless parameters set of the first section, the expression for the mass is $M = \frac{6n(b+1)^{3/2}}{a\sqrt{\Gamma}} + \frac{\Gamma b}{12d(b+1)n^2}$. Note that just as for the model in the main text, the mass M is minimal for a critical number of peaks $n = n_c$. In Fig. 17C, we compare the distribution of peaks of the final pattern obtained after coarsening with n_c and the linear prediction. We find that, like in the model of the main text, n_c is an excellent predictor of the number of peaks in the steady-state pattern.

B

MASS MINIMISATION AND PATTERN SELECTION IN SCHNAKENBERG MODEL

The Schnakenberg reaction-diffusion systems are described by the following equations [23] in 1D,

$$\frac{\partial u}{\partial t} = D_u \frac{\partial^2 u}{\partial x^2} + c\delta - \beta u v^2, \tag{B.1a}$$

$$\frac{\partial v}{\partial t} = D_v \frac{\partial^2 v}{\partial x^2} + c_1 - \delta v + \beta u v^2. \tag{B.1b}$$

We set the second feed term $c_1 = 0$, to simplify our parameter space. We perform the following re-scaling of the variables:

$$x \to \frac{x}{L}, \quad t \to \frac{D_v t}{L^2}, \quad u \to \frac{u}{c}, \quad v \to \frac{v}{c},$$
$$\Gamma = \frac{\delta L^2}{D_v}, \quad a = \frac{\beta c^2}{\delta}, \quad d = \frac{D_u}{D_v}.$$

The resulting dimensionless form of the equations is:

$$\frac{\partial u}{\partial t} = d \frac{\partial^2 u}{\partial x^2} + \Gamma(1 - a u v^2), \tag{B.2}$$

$$\frac{\partial v}{\partial t} = \frac{\partial^2 v}{\partial x^2} + \Gamma(-v + a u v^2). \tag{B.3}$$

The fixed points are,

$$u_0 = \frac{1}{a}, v_0 = 1. \tag{B.4}$$

The Jacobian ($J_{(u_0,v_0)}$) is given by,

$$J_{(u_0,v_0)} = \Gamma \begin{bmatrix} -a v_0^2 & -2a u_0 v_0 \\ a v_0^2 & -1 + 2a u_0 v_0 \end{bmatrix} = \Gamma \begin{bmatrix} -a & -2 \\ a & 1 \end{bmatrix} \tag{B.5}$$

The conditions for stability of the base state,

$$TrJ = \Gamma(-a+1) < 0 \implies a > 1$$

$$DetJ = \Gamma^2 a > 0 \, .$$

whereas the additional condition for spatial instability is

$$f_u + dg_v - 2\sqrt{dDetJ} = d - a - 2\sqrt{da} > 0 \tag{B.6}$$

The spike limit approximation can be applied to the Schnakenberg model and follows in a very similar way to the Brusselator model described above. We search for steady-state spike solutions consisting of n spikes at $x_i = x_1, ..., x_n$ and assume that u changes slowly within each spike and can be approximated by a constant u_i. We shift to an inner coordinate $y = \varepsilon^{-1}(x - x_i)$, within each spike. The equation for the inner variable $v_i(y)$ becomes,

$$\begin{aligned} \frac{D_v}{\varepsilon^2}\frac{d^2 v_i}{dy_i^2} + \beta u_i v_i^2 - \delta v_i = 0 \\ v_i \to 0 \quad \text{as} \quad y_i \to \pm\infty \, , \end{aligned}$$

The solution to this equation is

$$v_i = \frac{3}{2}\frac{\delta}{\beta u_i}\text{sech}^2\left(\sqrt{\frac{\delta}{D_v}}\frac{\varepsilon y_i}{2}\right) .$$

In the outer region, the spike is approximated by a Dirac delta function with weight

$$w_i = \varepsilon \int_{-\infty}^{\infty} v(y_i)^2 dy_i = 6\frac{\sqrt{D_v}\delta^{3/2}}{\beta^2 u_i}.$$

We follow the same arguments for the outer equations of u and v as in Appendix C of the main text for the Brusselator model. The outer equation for the fast species u has the form

$$\begin{aligned} D_u \frac{d^2 u}{dx^2} + c\delta - \sum_{i=1}^{n} \frac{\rho_s}{u_i} L\delta(x - x_i) = 0 \, , \\ -L/2 < x < L/2, \quad u_x(\pm L/2) = 0 \, , \end{aligned} \tag{B.7}$$

where $\rho_s = 6\frac{\sqrt{D_v}}{L}\frac{\delta^{3/2}}{\beta}$. We note that this is just Eq. (A.7) in the main text with the term $\frac{c\delta\gamma}{\gamma+\delta}$ replaced by $c\delta$ and ρ by ρ_s. The solution to the above equation is again given by Eq. (A.8) written in terms of the Green's functions defined in Eq. (A.10). Thus, all the comparisons to point sinks hold (namely Eq. (A.17) and Eq. (A.18)).

In the symmetric spike limit (spikes of height and regularly positioned) we find an expression for the mass of the fast species

$$M/c = \frac{n\sigma_s}{\eta_s} + \frac{\eta_s}{12n^2} \tag{B.8}$$

where $\sigma_s = \frac{\rho_s L^2}{c^2 D_u}$ and $\eta_s = L^2 \frac{\delta}{D_u}$.

In Figure 17D, we compare the number of peaks in the final pattern obtained at steady state with the predictions from linear stability analysis and the minimum (in n) of the above mass. Note that unlike the models in the main text, the Schnakenberg model does not have a mass-conserved limit and does not exhibit a dynamic coarsening process. Nonetheless, we find that, like the models in the main text, n_c (the number of peaks for which the mass is minimal) is a better predictor of steady state pattern than linear stability analysis in the well-separated spike limit.

BIBLIOGRAPHY

[1] Michael Cross and Henry Greenside. *Pattern Formation and Dynamics in Nonequilibrium Systems*. Cambridge University Press, 2009.

[2] A. M. Turing. The chemical basis of morphogenesis. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 237(641):37–72, 1952.

[3] BP Belousov. Periodicheski deistvuyushchaya reaktsia i ee mechanism. *Sbornik Referatov po Radiatsionni Meditsine*, 145147, 1958.

[4] Q Ouyang and Harry L Swinney. Transition from a uniform state to hexagonal and striped turing patterns. 352(August):610–612, 1991.

[5] M. C. Cross and P. C. Hohenberg. Pattern formation outside of equilibrium. *Reviews of Modern Physics*, 65(3):851–1112, 1993.

[6] Srikanth Subramanian and Seán M. Murray. Pattern selection in reaction diffusion systems. *Physical Review E*, 103(1):012215, jan 2021.

[7] A. De Wit. Spatial Patterns and Spatiotemporal Dynamics in Chemical Systems. pages 435–513. John Wiley & Sons, Ltd, mar 1999.

[8] B. Peña and C. Pérez-García. Selection and competition of Turing patterns. *EPL (Europhysics Letters)*, 51(3):300, 2000.

[9] B. Peña and C. Pérez-García. Stability of Turing patterns in the Brusselator model. *Physical Review E*, 64(5):056213, oct 2001.

[10] M. E. Cates, D. Marenduzzo, I. Pagonabarraga, and J. Tailleur. Arrested phase separation in reproducing bacteria creates a generic route to pattern formation. *Proceedings of the National Academy of Sciences*, 107(26):11715–11720, jun 2010.

[11] P Arcuri and J D Murray. Pattern sensitivity to boundary and initial conditions in reaction-diffusion models. *Journal of mathematical biology*, 24:141–165, 1986.

[12] Fabian Bergmann, Lisa Rapp, and Walter Zimmermann. Active phase separation: A universal approach. *Physical Review E*, 98(2):020603(R), aug 2018.

[13] Takeshi Sugawara and Kunihiko Kaneko. Chemophoresis as a driving force for intracellular organization: Theory and application to plasmid partitioning. *Biophysics*, 7:77–88, 2011.

[14] Robert Ietswaart, Florian Szardenings, K. Gerdes, and Martin Howard. Competing ParA structures space bacterial plasmids equally over the nucleoid. *PLoS computational biology*, 10(12):e1004009, dec 2014.

[15] Seán M. Murray and Martin Howard. Center Finding in E. coli and the Role of Mathematical Modeling: Past, Present and Future. *Journal of Molecular Biology*, 431(5):928–938, 2019.

[16] Juncheng Wei and Matthias Winter. *Mathematical Aspects of Pattern Formation in Biological Systems*, volume 189 of *Applied Mathematical Sciences*. Springer London, London, 2014.

[17] Mikiya Otsuji, Shuji Ishihara, Carl Co, Kozo Kaibuchi, Atsushi Mochizuki, and Shinya Kuroda. A Mass Conserved Reaction–Diffusion System Captures Properties of Cell Polarity. *PLoS Computational Biology*, 3(6):e108, 2007.

[18] Shuji Ishihara, Mikiya Otsuji, and Atsushi Mochizuki. Transient and steady state of mass-conserved reaction-diffusion systems. *Physical Review E*, 75(1):015203(R), jan 2007.

[19] Yoshihisa Morita and Toshiyuki Ogawa. Stability and bifurcation of nonconstant solutions to a reaction–diffusion system with conservation of mass. *Nonlinearity*, 23(6):1387–1411, jun 2010.

[20] Yoshihisa Morita. Spectrum comparison for a conserved reaction-diffusion system with a variational property. *J. Applied Analysis and Computation*, pages 1–18, 2012.

[21] Seán M. Murray and Victor Sourjik. Self-organization and positioning of bacterial protein clusters. *Nature Physics*, 13(10):1006–1013, jun 2017.

[22] I. Prigogine and R. Lefever. Symmetry breaking instabilities in dissipative systems. II. *The Journal of Chemical Physics*, 48(4):1695–1700, feb 1968.

[23] J. Schnakenberg. Simple chemical reaction systems with limit cycle behaviour. *Journal of Theoretical Biology*, 81(3):389–400, dec 1979.

[24] James D. Murray. *Mathematical Biology II*, volume 18 of *Interdisciplinary Applied Mathematics*. Springer New York, New York, NY, 2003.

[25] JC Tzou, Y Nec, and MJ Ward. The stability of localized spikes for the 1-D Brusselator reaction–diffusion model. *European Journal of Applied Mathematics*, 24(2013):515–564, 2013.

[26] Theodore Kolokolnikov, Michael J. Ward, and Juncheng Wei. The Existence and Stability of Spike Equilibria in the One-Dimensional Gray–Scott Model: The Low Feed-Rate Regime. *Studies in Applied Mathematics*, 115(1):21–71, jul 2005.

[27] Theodore Kolokolnikov, Michael J. Ward, and Juncheng Wei. Slow translational instabilities of spike patterns in the one-dimensional Gray-Scott model. *Interfaces and Free Boundaries*, 8(2):185–222, 2006.

[28] Fridtjof Brauns, Henrik Weyer, Jacob Halatek, Junghoon Yoon, and Erwin Frey. Wavelength selection by interrupted coarsening in reaction-diffusion systems. may 2020.

[29] J.-C. Walter, J. Dorignac, V. Lorman, J. Rech, Jean-Yves Bouet, M. Nollmann, J. Palmeri, A. Parmeggiani, and F. Geniet. Surfing on Protein Waves: Proteophoresis as a Mechanism for Bacterial Genome Partitioning. *Physical Review Letters*, 119(2):028101, jul 2017.

[30] David Iron, Juncheng Wei, and Matthias Winter. Stability analysis of Turing patterns generated by the Schnakenberg model. *Journal of Mathematical Biology*, 49(4):358–390, oct 2004.

[31] Edmund J. Crampin, Eamonn A Gaffney, and Philip K Maini. Reaction and diffusion on growing domains: scenarios for robust pattern formation. *Bulletin of mathematical biology*, 61:1093–1120, 1999.

[32] Jake Cornwall Scoones, Deb Sankar Banerjee, and Shiladitya Banerjee. Size-Regulated Symmetry Breaking in Reaction-Diffusion Models of Developmental Transitions. *Cells*, 9(7):1–11, 2020.

[33] O. Wartlick, P. Mumcu, A. Kicheva, T. Bittig, C. Seum, F. Jülicher, and M. González-Gaitán. Dynamics of Dpp signaling and proliferation control. *Science*, 331(6021):1154–1159, 2011.

[34] Fisun Hamaratoglu, Aitana Morton de Lachapelle, George Pyrowolakis, Sven Bergmann, and Markus Affolter. Dpp signaling activity requires pentagone to scale with tissue size in the growing drosophila wing imaginal disc. *PLoS Biology*, 9(10), 2011.

[35] Danny Ben-Zvi, George Pyrowolakis, Naama Barkai, and Ben Zion Shilo. Expansion-repression mechanism for scaling the Dpp activation gradient in drosophila wing imaginal discs. *Current Biology*, 21(16):1391–1396, 2011.

[36] Alin Pogan and Arnd Scheel. Layers in the Presence of Conservation Laws. *Journal of Dynamics and Differential Equations*, 24(2):249–287, 2012.

[37] Masataka Kuwamura and Yoshihisa Morita. Perturbations and dynamics of reaction-diffusion systems with mass conservation. *Physical Review E*, 92(1):012908, jul 2015.

[38] Shuichi Jimbo and Yoshihisa Morita. Lyapunov function and spectrum comparison for a reaction–diffusion system with mass conservation. *Journal of Differential Equations*, 255(7):1657–1683, oct 2013.

[39] Theodore Kolokolnikov, Michael J. Ward, and Juncheng Wei. The existence and stability of spike equilibria in the one-dimensional Gray–Scott model: The pulse-splitting regime. *Physica D: Nonlinear Phenomena*, 202(3-4):258–293, mar 2005.

[40] A. Gierer and H. Meinhardt. A theory of biological pattern formation. *Kybernetik*, 12(1):30–39, dec 1972.

[41] David Iron, Michael J. Ward, and Juncheng Wei. The stability of spike solutions to the one-dimensional Gierer–Meinhardt model. *Physica D: Nonlinear Phenomena*, 150(1):25–62, 2001.

[42] Anjana Badrinarayanan and Michael T Laub. Bacterial chromosome organization and segregation Anjana. *Annu Rev Cell Dev Biol.*, 31:171–199, 2016.

[43] Henrik J. Nielsen, Jesper R. Ottesen, Brenda Youngren, Stuart J. Austin, and Flemming G. Hansen. The Escherichia coli chromosome is organized with the left and right chromosome arms in separate cell halves. *Molecular Microbiology*, 62(2):331–338, 2006.

[44] Olessia Danilova, Rodrigo Reyes-lamothe, Marina Pinskaya, David Sherratt, and Christophe Possoz. MukB colocalizes with the oriC region and is required for organization of the two Escherichia coli chromosome arms into separate cell halves. 65(August):1485–1492, 2007.

[45] Stephan Gruber and Jeff Errington. Recruitment of Condensin to Replication Origin Regions by ParB / SpoOJ Promotes Chromosome Segregation in B . subtilis. *Cell*, 137(4):685–696, 2009.

[46] Andreas Hofmann, Jarno Mäkelä, David J Sherratt, Dieter Heermann, and Seán M Murray. Self-organised segregation of bacterial chromosomal origins. *eLife*, 8:e46564, August 2019.

[47] Sophie Nolivos, Amy L. Upton, Anjana Badrinarayanan, Julius Müller, Katarzyna Zawadzka, Jakub Wiktor, Amber Gill, Lidia Arciszewska, Emilien Nicolas, and David Sherratt. MatP regulates the coordinated action of topoisomerase IV and MukBEF in chromosome segregation. *Nature Communications*, 7, 2016.

[48] I. Teraoka. *Polymer Solutions: An Introduction to Physical Properties*. Wiley, 2004.

[49] M. Doi, P.A.P.M. Doi, and S.F. Edwards. *The Theory of Polymer Dynamics*. International series of monographs on physics. Clarendon Press, 1986.

[50] Stephanie C. Weber, Andrew J. Spakowitz, and Julie A. Theriot. Bacterial Chromosomal Loci Move Subdiffusively through a Viscoelastic Cytoplasm. *Physical Review Letters*, 104(23):238102, June 2010.

[51] S. C. Weber, A. J. Spakowitz, and J. A. Theriot. Nonthermal ATP-dependent fluctuations contribute to the in vivo motion of chromosomal loci. *Proceedings of the National Academy of Sciences*, 109(19):7338–7343, May 2012.

[52] Avelino Javer, Zhicheng Long, Eileen Nugent, Marco Grisi, Kamin Siriwatwetchakul, Kevin D. Dorfman, Pietro Cicuta, and Marco Cosentino Lagomarsino. Short-time movement of E. coli chromosomal loci depends on coordinate and subcellular localization. *Nature Communications*, 4(1):3003, October 2013.

[53] Shi Yu, Julian Sheats, Pietro Cicuta, Bianca Sclavi, Marco Cosentino Lagomarsino, and Kevin D. Dorfman. Subdiffusion of loci and cytoplasmic particles are different in compressed Escherichia coli cells. *Communications Biology*, 1(1), 2018.

[54] Fabai Wu, Pinaki Swain, Louis Kuijpers, Xuan Zheng, Kevin Felter, Margot Guurink, Jacopo Solari, Suckjoon Jun, Thomas S. Shimizu, Debasish Chaudhuri, Bela Mulder, and Cees Dekker. Cell Boundary Confinement Sets the Size and Position of the E. coli Chromosome. *Current Biology*, 29(13):2131–2144.e4, July 2019.

[55] Estelle Crozat, Catherine Tardin, Maya Salhi, Philippe Rousseau, Armand Lablaine, Tommaso Bertoni, David Holcman, Bianca Sclavi, Pietro Cicuta, and François Cornet. Post-replicative pairing of sister ter regions in Escherichia coli involves multiple activities of MatP. *Nature Communications*, 11(1):3796, December 2020.

[56] Stephanie C. Weber, Julie A. Theriot, and Andrew J. Spakowitz. Subdiffusive motion of a polymer composed of subdiffusive monomers. *Physical Review E*, 82(1):011913, July 2010.

[57] Avelino Javer, Nathan J. Kuwada, Zhicheng Long, Vincenzo G. Benza, Kevin D. Dorfman, Paul A. Wiggins, Pietro Cicuta, and Marco Cosentino Lagomarsino. Persistent super-diffusive motion of Escherichia coli chromosomal loci. *Nature Communications*, 5(1):3854, September 2014.

[58] Joel Stavans and Amos Oppenheim. DNA–protein interactions and bacterial chromosome architecture. *Physical Biology*, 3(4):R1–R10, dec 2006.

[59] Remus T. Dame and Mariliis Tark-Dame. Bacterial chromatin: Converging views at different scales. *Current Opinion in Cell Biology*, 40:60–65, 2016.

[60] Remus T. Dame, Fatema Zahra M. Rashid, and David C. Grainger. Chromosome organization in bacteria: mechanistic insights into genome structure and function. *Nature Reviews Genetics*, 21(4):227–242, 2020.

[61] Jay K. Fisher, Aude Bourniquel, Guillaume Witz, Beth Weiner, Mara Prentiss, and Nancy Kleckner. Four-dimensional imaging of E. coli nucleoid organization and dynamics in living cells. *Cell*, 153(4):882–895, 2013.

[62] Virginia S. Lioy, Axel Cournac, Martial Marbouty, Stéphane Duigou, Julien Mozziconacci, Olivier Espéli, Frédéric Boccard, and Romain Koszul. Multiscale Structuring of the E. coli Chromosome by Nucleoid-Associated and Condensin Proteins. *Cell*, 172(4):771–783.e18, February 2018.

[63] Joris J.B. Messelink, Muriel C.F. van Teeseling, Jacqueline Janssen, Martin Thanbichler, and Chase P. Broedersz. Learning the distribution of single-cell chromosome conformations in bacteria reveals emergent order across genomic scales. *Nature Communications*, 12(1), 2021.

[64] Tejal Agarwal, G. P. Manjunath, Farhat Habib, and Apratim Chatterji. Bacterial chromosome organization. II. Few special cross-links, cell confinement, and molecular crowders play the pivotal roles. *The Journal of Chemical Physics*, 150(14):144909, April 2019.

[65] Abdul Wasim, Ankit Gupta, and Jagannath Mondal. A Hi-C data-integrated model elucidates E. coli chromosome's multiscale organization at various replication stages. *Nucleic Acids Research*, 49(6):3077–3091, 2021.

[66] Palash Bera, Abdul Wasim, and Jagannath Mondal. Hi-C embedded polymer model of Escherichia coli reveals the origin of heterogeneous subdiffusion in chromosomal loci. *Physical Review E*, 105(5):1–13, 2022.

[67] Ludwik Leibler, Michael Rubinstein, and Ralph H. Colby. Dynamics of reversible networks. *Macromolecules*, 24(16):4701–4707, 1991.

[68] Michael Rubinstein and Alexander N. Semenov. Thermoreversible gelation in solutions of associating polymers. 2. linear dynamics. *Macromolecules*, 31(4):1386–1397, 1998.

[69] Michael Rubinstein and Alexander N. Semenov. Dynamics of Entangled Solutions of Associating Polymers. *Macromolecules*, 34(4):1058–1068, February 2001.

[70] Manfred Bohn and Dieter W. Heermann. Diffusion-Driven Looping Provides a Consistent Framework for Chromatin Organization. *PLoS ONE*, 5(8):e12218, August 2010.

[71] Mario Nicodemi and Antonella Prisco. Thermodynamic pathways to genome spatial organization in the cell nucleus. *Biophysical Journal*, 96(6):2168–2177, 2009.

[72] Mariliis Tark-Dame, Hansjoerg Jerabek, Erik M.M. Manders, Dieter W. Heermann, and Roel van Driel. Depletion of the Chromatin Looping Proteins CTCF and Cohesin Causes Chromatin Compaction: Insight into Chromatin Folding by Polymer Modelling. *PLoS Computational Biology*, 10(10), 2014.

[73] Julio Mateos-Langerak, Manfred Bohn, Wim De Leeuw, Osdilly Giromus, Erik M.M. Manders, Pernette J. Verschure, Mireille H.G. Indemans, Hinco J. Gierman, Dieter W. Heermann, Roel Van Driel, and Sandra Goetze. Spatially confined folding of chromatin in the interphase nucleus. *Proceedings of the National Academy of Sciences of the United States of America*, 106(10):3812–3817, 2009.

[74] Mariano Barbieri, Mita Chotalia, James Fraser, Liron-Mark Lavitas, Josée Dostie, Ana Pombo, and Mario Nicodemi. Complexity of chromatin folding is captured by the strings and binders switch model. *Proceedings of the National Academy of Sciences*, 109(40):16173–16178, 2012.

[75] Andrea M. Chiariello, Carlo Annunziatella, Simona Bianco, Andrea Esposito, and Mario Nicodemi. Polymer physics of chromosome large-scale 3D organisation. *Scientific Reports*, 6:1–8, 2016.

[76] Mario Nicodemi and Ana Pombo. Models of chromosome structure. *Current Opinion in Cell Biology*, 28(1):90–95, 2014.

[77] Andreas Hofmann and Dieter W. Heermann. The role of loops on the order of eukaryotes and prokaryotes. *FEBS Letters*, 589(20, Part A):2958–2965, 2015. 3D Genome structure.

[78] Srikanth Subramanian and Sean M. Murray. Dna bridging explains sub-diffusive movement of chromosomal loci in bacteria. *bioRxiv*, 2022.

[79] Pierre-giles De Gennes. Pierre-giles De Gennes - Scaling concepts in polymer physics (1979, Cornell University Press) - libgen.lc.pdf, 1979.

[80] M. Rubinstein and R.H. Colby. *Polymer Physics*. OUP Oxford, 2003.

[81] I. Carmesin and Kurt Kremer. The bond fluctuation method: a new effective algorithm for the dynamics of polymers in all spatial dimensions. *Macromolecules*, 21(9):2819–2823, September 1988.

[82] H. P. Deutsch and K. Binder. Interdiffusion and self-diffusion in polymer mixtures: A Monte Carlo study. *The Journal of Chemical Physics*, 94(3):2294–2304, February 1991.

[83] Stephanie C. Weber, Michael A. Thompson, W. E. Moerner, Andrew J. Spakowitz, and Julie A. Theriot. Analytical tools to distinguish the effects of localization error, con-

finement, and medium elasticity on the velocity autocorrelation function. *Biophysical Journal*, 102(11):2443–2450, 2012.

[84] Charles E. Sing and Alfredo Alexander-Katz. Equilibrium structure and dynamics of self-associating single polymers. *Macromolecules*, 44(17):6962–6971, 2011.

[85] Valerio Sorichetti, Virginie Hugouvieux, and Walter Kob. Determining the Mesh Size of Polymer Solutions via the Pore Size Distribution. *Macromolecules*, 53(7):2568–2581, 2020.

[86] Stephanie C. Weber, Michael A. Thompson, W.E. Moerner, Andrew J. Spakowitz, and Julie A. Theriot. Analytical Tools To Distinguish the Effects of Localization Error, Confinement, and Medium Elasticity on the Velocity Autocorrelation Function. *Biophysical Journal*, 102(11):2443–2450, June 2012.

[87] Yingjie Xiang, Ivan V. Surovtsev, Yunjie Chang, Sander K. Govers, Bradley R. Parry, Jun Liu, and Christine Jacobs-Wagner. Interconnecting solvent quality, transcription, and chromosome folding in Escherichia coli. *Cell*, page S0092867421006899, June 2021.

[88] Mikhail V. Tamm and Kirill Polovnikov. Dynamics of polymers: Classic results and recent developments. *Order, Disorder and Criticality - Advanced Problems of Phase Transition Theory*, 5:113–172, 2017.

[89] Jonathan D. Halverson, Jan Smrek, Kurt Kremer, and Alexander Y. Grosberg. From a melt of rings to chromosome territories: The role of topological constraints in genome folding. *Reports on Progress in Physics*, 77(2), 2014.

[90] J. des Cloizeaux. Short Range Correlation Between Elements of a Long Polymer in a Good Solvent. *Journal de physique Paris*, 41(3):223–238, 1980.

[91] S. Redner. Distribution functions in the interior of polymer chains. *Journal of Physics A: Mathematical and General*, 13(11):3525–3541, 1980.

[92] Romain Mercier, Marie Agnès Petit, Sophie Schbath, Stéphane Robin, Meriem El Karoui, Frédéric Boccard, and Olivier Espéli. The MatP/matS Site-Specific System Organizes the Terminus Region of the E. coli Chromosome into a Macrodomain. *Cell*, 135(3):475–485, 2008.

[93] Yongfang Li, Kirill Sergueev, and Stuart Austin. The segregation of the Escherichia coli origin and terminus of replication. *Molecular Microbiology*, 46(4):985–996, 2002.

[94] Henrik J. Nielsen, Yongfang Li, Brenda Youngren, Flemming G. Hansen, and Stuart Austin. Progressive segregation of the Escherichia coli chromosome. *Molecular Microbiology*, 61(2):383–393, 2006.

[95] Olivier Espeli, Romain Mercier, and Frédéric Boccard. DNA dynamics vary according to macrodomain topography in the *E. coli* chromosome. *Molecular Microbiology*, 68(6):1418–1427, June 2008.

[96] Mathieu Stouf, Jean Christophe Meile, and Francois Cornet. FtsK actively segregates sister chromosomes in Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America*, 110(27):11157–11162, 2013.

[97] Marco Gherardi, Ludovico Calabrese, Mikhail Tamm, and Marco Cosentino Lagomarsino. Model of chromosomal loci dynamics in bacteria as fractional diffusion with intermittent transport. *Physical Review E*, 96(4):042402, oct 2017.

[98] Weihua Deng and Eli Barkai. Ergodic properties of fractional Brownian-Langevin motion. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 79(1):1–7, 2009.

[99] Benoit B. Mandelbrot and John W. Van Ness. Fractional brownian motions, fractional noises and applications. *SIAM Review*, 10(4):422–437, 1968.

[100] Steven B Lowen. Efficient generation of fractional Brownian motion for simulation of infrared focal-plane array calibration drift. *Methodol. Comput. Appl. Probab. 1, No.4, 445-456 (1999).*, 1(4):445–456, 1999.

[101] Mohit Kumar, Mario S. Mommer, and Victor Sourjik. Mobility of cytoplasmic, membrane, and DNA-binding proteins in Escherichia coli. *Biophysical Journal*, 98(4):552–559, 2010.

[102] Kevin Floc'h, Françoise Lacroix, Pascale Servant, Yung Sing Wong, Jean Philippe Kleman, Dominique Bourgeois, and Joanna Timmins. Cell morphology and nucleoid dynamics in dividing Deinococcus radiodurans. *Nature Communications*, 10(1):1–13, 2019.

[103] Tung B.K. Le, Maxim V. Imakaev, Leonid A. Mirny, and Michael T. Laub. High-resolution mapping of the spatial organization of a bacterial chromosome. *Science*, 342(6159):731–734, 2013.

[104] Axel Cournac, Hervé Marie-Nelly, Martial Marbouty, Romain Koszul, and Julien Mozziconacci. Normalization of a chromosomal contact map. *BMC Genomics*, 13(1), 2012.

[105] Richard Sinkhorn and Paul Knopp. Concerning nonnegative matrices and doubly stochastic matrices. *Pacific Journal of Mathematics*, 21(2):343–348, 1967.

[106] Mohan C. Joshi, Aude Bourniquel, Jay Fisher, Brian T. Ho, David Magnan, Nancy Kleckner, and David Bates. Escherichia coli sister chromosome separation includes

an abrupt global transition with concomitant release of late-splitting intersister snaps. *Proceedings of the National Academy of Sciences of the United States of America*, 108(7):2765–2770, 2011.

[107] J. C. Tzou, A. Bayliss, B. J. Matkowsky, and V. A. Volpert. Stationary and slowly moving localised pulses in a singularly perturbed Brusselator model. *European Journal of Applied Mathematics*, 22(05):423–453, oct 2011.

ABBREVATIONS

ODE Ordinary Differential Equation

RD Reaction-diffusion

DNA Deoxyribonucleic acid

NAP Nucleoid Associated Proteins

SMC Structural maintenance of chromosome

HiC Chromosome conformation capture

MSD Mean-squared displacement

MCS Monte Carlo steps

fBm Fractional Brownian motion

CTRW Continous time random walk

VAC Velocity auto-correlation

SAW Self-avoiding walk

RW Random walk

GFP Green Fluorescent Protein

WT Wild Type

RCM Rapid Chromosomal Movement

SCN Sequential Component Normalization

LIST OF FIGURES