TIMING AND SHAPE OF STOCHASTIC AUTOCATALYTIC BURST FORMATION

Alastair Jamieson-Lane
Department of Mathematics,
University of British Columbia
aja107@math.ubc.ca

Eric N. Cytrynbaum
Department of Mathematics,
University of British Columbia.
cytryn@math.ubc.ca

ABSTRACT

Chemical, physical and ecological systems passing through a saddle-node bifurcation will, momentarily, find themselves balanced at a semi-stable steady state. If perturbed by noise, such systems will escape from the zero-steady state, with escape time sensitive to noise. When the model is extended to include space, this leads to different points in space "escaping from zero" at different times, and uniform initial conditions nucleate into sharp peaks spread randomly across a nearly uniform background, a phenomena closely resembling nucleation during phase transition. We use Large Deviation Theory to determine burst shape and temporal scaling with respect to noise amplitude. These results give a prototype for a particular form of patternless symmetry breaking in the vicinity of a stability boundary, and demonstrates how microscopic noise can lead to macroscopic effects in such a region.

1 Introduction

The universe we live in does not consist of a perfectly uniform cloud of hydrogen gas, nor is the earth a perfect sphere. Clouds condense into droplets of rain [McGraw and Liu(2003)], a single zygote divides and differentiates into a multitude of different cell types, a flipped coin lands either heads, or tails. All around us, symmetry is broken every day.

Depending on the system of interest, symmetry can be broken in either a patterned or unpatterned manner. Pattern formation is ubiquitous in nature, from a Zebra's strips [Turing(1952)] to Tiger bush [Lefever and Lejeune(1997)]. Patterned symmetry breaking in reaction diffusion systems is well understood, and was first studied by Alan Turing in the 1950s [Turing(1952)].

In nature we also observe patternless symmetry breaking, in for example the nucleation of crystals[Erdemir et al.(2009)Erdemir, Lee, and Myerson] or rain drops[McGraw and Liu(2003)]. This patternless symmetry breaking is less studied in a mathematical context, and is the focus of this article.

Our present work was motivated by intriguing experimental results in a recent study of the Min cell division system in E.coli [Vecchiarelli et al.(2016)Vecchiarelli, Li, Mizuuchi, Hwang, Seol, Neuman, and Mizuuchi]. In their study, Vecchiarelli et al. observed that Min proteins demonstrate either patterned or unpatterned symmetry breaking, depending on the concentration conditions. For high concentrations of the relevant Min proteins, spirals and linear waves are observed, with each of these patterns possessing their own well studied mathematical frameworks [Murray(2008)]. For lower concentrations, Vecchiarelli et al's experiment also demonstrates "burst" behavior, whereby the local concentration of Min proteins increases for spatially localized patches of membrane, and then reduces again shortly thereafter (see figure 1). These bursts appear to arise from a relatively homogeneous background, and in contrast to the ordered high concentration patterns described above, they are positioned randomly across the membrane, and appear governed by low level particle scale noise.

In this paper, we demonstrate how dynamical systems in the vicinity of a semi-stable steady state are sensitive to low level white noise, giving rise to similar burst formation behavior. We consider a "canonical" stochastic PDE model of
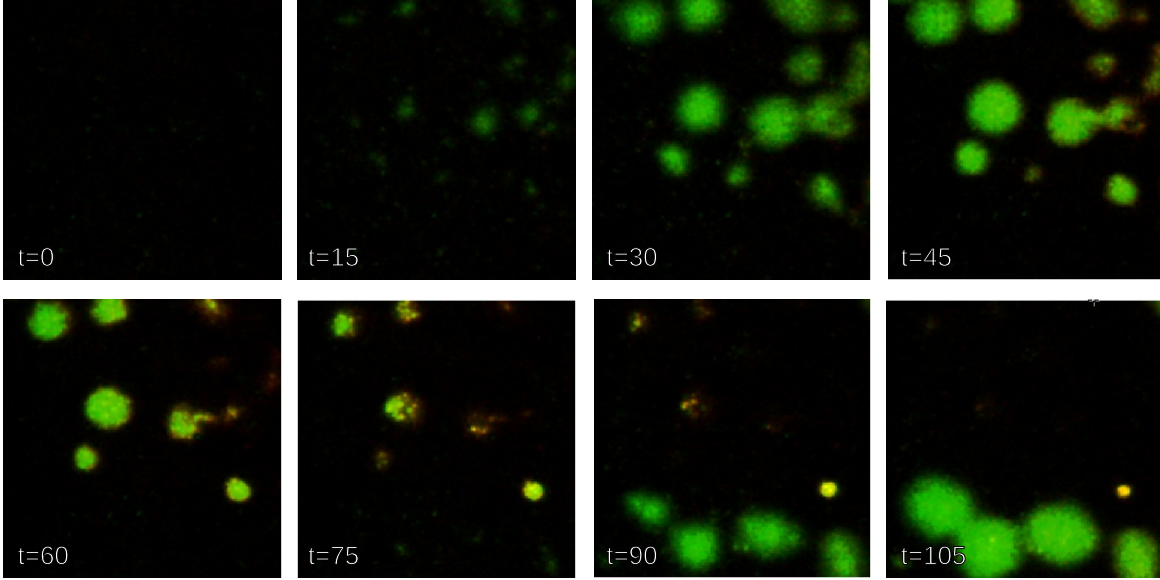
Figure 1: Burst patterning, as observed by Vecchiarelli *et al*. At $t = 0$, the membrane is at a low, almost homogeneous, concentration of particles. This initial condition evolves into bright, high concentration 'bursts' of membrane-bound MinD, which then freeze in place, and fade. While one generation of bursts fades, the next generation of bursts appears (bottom row). These previously unpublished images of the experiments described in [Vecchiarelli et al.(2016)Vecchiarelli, Li, Mizuuchi, Hwang, Seol, Neuman, and Mizuuchi] were kindly provided by A. Vecchiarelli.

such a system

$$u_t = u_{xx} + u^2 + \varepsilon \xi, \tag{1}$$

and use Large Deviation Theory to predict both the time of burst formation, and the shape of the burst formed, in the limit of low amplitude noise.

We consider a particular reaction-diffusion system from the literature in Sect. 2, and use simulations to demonstrate the occurrence of burst formation in appropriate parameter regimes. In Sec. 3, we use Large Deviation Theory to predict the time taken for the stochastic ODE $u_t = u^2 + \varepsilon \xi$ to evolve from $u = 0$ to some fixed value U_f, and determine the most probable path for this occurrence. Sect. 4 extends this work to systems with one spatial dimension via the addition of a diffusion term (as is 1). We use a mixture of numeric and analytic techniques to determine both the shape and time taken for bursts to form. In Sec. 5, we compare our analytic results both to one another, and to the corresponding simulations, and in Sec. 6 we discuss the conceptual difficulties preventing our analysis from being extended to higher dimensions. Finally, in Sec. 7 we summarize the results, and discuss limitations, physical relevance, and future directions for research. The Appendices contain both an introduction to Large Deviation Theory for those unfamiliar with the topic, along with more detailed calculations excluded from the main discussion.

2 Computational Exploration

Before delving into analytical results, we take a brief computational detour to demonstrate the general behaviour of the class of neutrally stable systems that form the focus of this paper.

Consider the following equation due to Mori et al. [Mori et al.(2008)Mori, Jilkine, and Edelstein-Keshet]

$$u_t = D\Delta u + f(u,v) - \delta u \tag{2}$$

$$\text{where } f(u,v) = v\left(k_0 + \frac{\gamma u^2}{K^2 + u^2}\right), \tag{3}$$

over the unit square in $\mathbb{R}^2$ with periodic boundary conditions. For the sake of these simulations, we use $\gamma = 5$, $K = 2$ $\delta = 0.5$ and diffusion coefficient $D = 10^{-4}$.
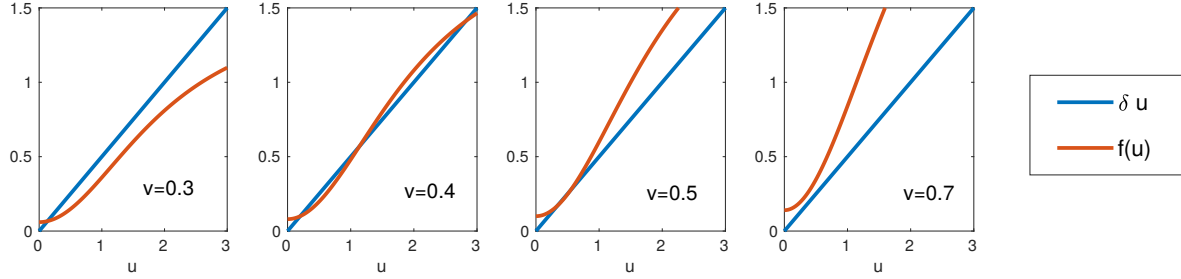
2

Figure 2: As the bulk concentration v is varied, $u_t = f(u,v) - \delta u = v\left(k_0 + \frac{\gamma u^2}{K^2 + u^2}\right) - \delta u$ gains and loses steady states via saddle-node bifurcations. In particular, for $v = 0.3$, only the low u steady state exists. At $v = 0.4$ there are three steady states- stable steady states near $u = 0$ and $u = 3$, with an unstable steady state. At $v \approx 0.5$ the low u steady state is lost, and only the high u steady state remains (outside the bounds of these diagrams)
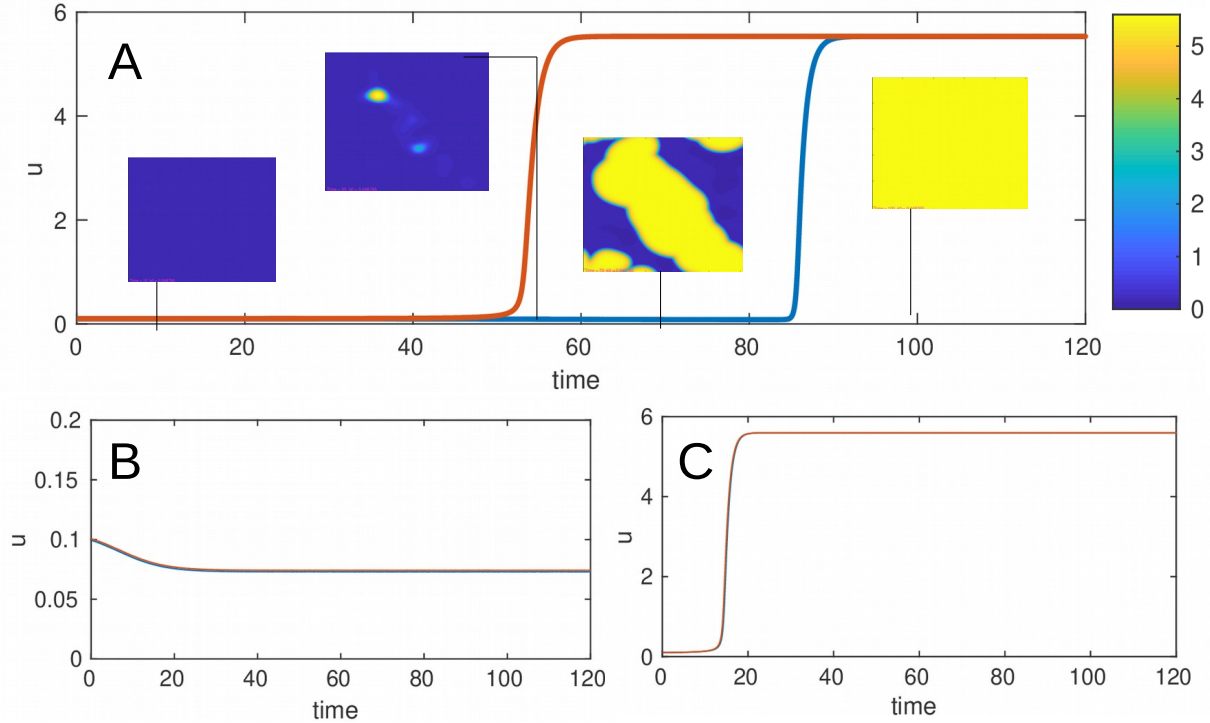

Figure 3: For equation (2) we track the maximum and minimum value of u through time (red and blue, respectively). A) For v selected at the bifurcation point, noise leads to burst formation (see images, inset), leading $\max(u)$ and $\min(u)$ to separate for a time. B) for v selected slightly below the bifurcation value, noise is unable to overcome the energy barrier, and u remains in the vicinity of the low u steady state indefinitely. C) For v slightly above the bifurcation value, u tends towards saturation, but does so homogeneously - $\max(u)$ and $\min(u)$ remain close together throughout the entire process.

Equation (2) was designed to model a system of cell signaling proteins called Rho GTPases. The model was originally used to demonstrate wave pinning, a behavior reminiscent of the "cell polarization" behavior associated with Rho proteins.

In this equation D, k_0, γ, K and δ are reaction and diffusion parameters, u is the concentration of membrane bound Rho protein and v is the concentration of Rho diffusing quickly through the bulk of the cell. As v is varied, the above system gains and loses steady states via a pair of saddle-node bifurcations (see figure 2). Let u_0 and v_0 denote the values of u and v at the second of these saddle-node bifurcations, in which the lower stable steady state is annihilated. In this region $v_0 \approx 0.5$.

Simulating with fixed $v = v_0$, $u(0, \mathbf{x}) = u_0$ and additive noise of the form $\varepsilon\xi$, leads to a long nearly homogeneous period where $u(\mathbf{x}, t) \approx u_0$ for all $\mathbf{x}$, followed by the abrupt formation of several "spikes", which then saturate and spread across the domain. (See figure 3 A)

By contrast, when we select some fixed v greater than v_0 the system remains approximately homogeneous as it tend towards this u-saturated state. For v less than v_0, the system instead approaches some homogeneous low u state. In either of these latter two cases, no significant patterning or heterogeneity is observed. (See figure 3 B & C)

Similar results may be observed for a variety of spatially distributed systems in the vicinity of saddle-node bifurcations. In order to consider this behavior in more generality, we consider a variety of "cannonical" equations, each demonstrating different aspects of the behavior of interest in the vicinity of a generic saddle-node.

3 Escape time from steady states; spaceless case

Let us begin our discussion by considering first the spaceless stochastic differential equation

$$\dot{u} = u^2 + \varepsilon\xi. \tag{4}$$

Here u is our stochastic process of interest, ε is some small noise amplitude parameter, and ξ is assumed to be Gaussian white noise. We assume noise as defined by Walsh [Walsh(1986)], such that the integral over any time interval Ω is given by $\int_\Omega \xi dt = N(0, |\Omega|)$, and integrals over non-intersecting time intervals are independent. The extension to spatio-temporal white noise (also discussed by Walsh) makes use of higher dimensional integration. Equation 4 provides a canonical example of a spaceless noisy system balanced at a saddle node, and provides a toy model on which to build our understanding before moving to the spatial case.

Deterministically, trajectories starting at $u = 0$ remain stationary, and never reach positive U_f. However, any noise added to the system is enough to disturb this equilibrium and allow u to escape. We would like to determine the amount of time it takes our stochastic process u to travel from the semi-stable steady state $u = U_i = 0$ to some positive constant U_f. The appropriate tool for this task is Large Deviation Theory.

Large Deviation Theory (LDT) is a theory used to study unlikely noise driven events in stochastic systems (often, but not always SDEs). For the interested reader, we give a brief introduction to LDT in Appendix A. For those who wish for more detail, we recommend Freidln and Wentzell's "Random Perturbations of Dynamical Systems" [Freidlin and Wentzell(2012)], and Rassoul-agha and Seppalainen's "A Course on Large Deviations With an Introduction to Gibbs Measures"[Rassoul-agha and Seppalainen(2015)]. Here we present only the key results of LDT - those concepts and theorems critical to our present interest.

The central premise of LDT is that when unlikely events occur, they are overwhelmingly likely to occur via a path "close to" the most probable path. The most probable path, $\phi(t)$, is found by minimizing the "normalized action functional"

$$\bar{S}_T(\xi) = \int_0^T \varepsilon^2 \xi^2 dt. \tag{5}$$

For any given SPDE, $\bar{S}_T(\xi)$ can be re-written as a function of ϕ. For example, when studying eq. 4 we can rearrange to isolate $\varepsilon\xi$ and find

$$S_T(\phi) = \int_0^T (\dot{\phi} - \phi^2)^2 dt \tag{6}$$

Minimizing $S_T(\phi)$ allows us to determine the most probably path for a given event, and having identified this part we are able to approximate the probability of an event occurring by a given time using either

$$\lim_{\varepsilon \to 0} \varepsilon^2 \ln \mathbb{E}(\tau) = \min_{T, \phi} \bar{S}_T(\phi), \tag{7}$$

4

to estimate the expected time $\mathbb{E}(\tau)$ or

$$\lim_{\varepsilon \to 0} \varepsilon^2 \ln P(\tau \leq T) = -\min_\phi \bar{S}_T(\phi) \tag{8}$$

to estiamte the probability distribution.

These formula are given as theorem 4.1 and theorem 1.2 (respectively) in chapter 4 of Freidln and Wentzell[Freidlin and Wentzell(2012)].

We will make use of both of these theorems throughout this paper, however, because $\min_{T,\phi} S_T(\phi) = 0$ for many of the examples discussed, eq. 8 will prove to be the more helpful of the two. This is because we are in the slightly unusual position of studying escape from an energy plateau, as opposed to an energy well, as might be more usually studied.

Suppose we wish to study the probability of u passing from $U_i = 0$ to $u \geq U_f$ by time $t = T$, that is to say the "escape from zero" problem for equation 4. In order to use either of the above formula, we must find the minimum of $S_T(\phi)$. By the calculus of variations, any minimizer of $S_T(\phi)$ must satisfy

$$0 = \frac{d\bar{S}_T(\phi)}{d\phi} = -2\ddot{\phi} + 4\phi^3, \tag{9}$$

and hence

$$\int_{U_i}^{\phi(t)} \frac{1}{\sqrt{\psi^4 + C}} d\psi = T - 0. \tag{10}$$

For $C = 0$ we recover the deterministic solution $\phi(t) = 1/(U_i^{-1} - t)$. This solution indicates infinite travel time to U_f in the case where $U_i = 0$ (our case of interest). For $C \neq 0$, the integral can be solved using wolfram|alpha [Wolfram|Alpha(2018a)], determining t in terms of ϕ. Given that the function itself provides limited illumination, and is defined in terms of the ellipticF function (itself defined in terms of an integral), we will refer to solutions of 10 simple as $\tau(\phi)$, the inverse of $\phi(t)$.

In order to match our assumed boundary conditions, we must pick C such that $\tau(U_f) = T$. This is not generically an easy problem to solve exactly, but for large U_f it can be well approximated. In order to determine C, in this limit, we first note $\tau_\infty = \lim_{\phi \to \infty} \tau(\phi) = \frac{4\Gamma(5/4)^2}{\sqrt{\pi}C^{1/4}}$ (as given by wolfram|alpha [Wolfram|Alpha(2018b)]). Here Γ is the Gamma function [Artin(2015)], and takes the value $\Gamma(5/4) \approx 0.9064$. The deterministic time taken to get from $\phi = U_f$ to $\phi = \infty$ is $1/U_f$, thus $\tau_\infty \approx T + 1/U_f$. Rearranging gives

$$C \approx \left(\frac{4\Gamma(5/4)^2}{\sqrt{\pi}(T + 1/U_f)} \right)^4. \tag{11}$$

Numerical experimentation indicates that, for this choice of C, $|\tau(U_f) - T|/T < 10^{-6}$ whenever $T, U_f \geq 4$, indicating that our approximation of C is very good.

Having determined C, we are able to plot $t = \tau(\phi)$ for arbitrary U_f and T (see figure 4, Left).

In order to invoke either eq. 7 or eq. 8, we must must determine $S_T(\phi)$ along this action minimizing path. Starting from the definition of $S_T(\phi)$ (equation 6), it is possible to show that in our case

$$S_T(\phi) = \frac{2}{3} \left[\phi^3 (\sqrt{1 + C\phi^{-4}} - 1) \right]_0^{U_f} + \frac{1}{3} CT, \tag{12}$$

$$\approx \frac{C}{3U_f} - 0 + \frac{1}{3} CT, \quad \text{for large } U_f. \tag{13}$$

Details are given in appendix B. Remembering from equation (11) that $C = O(T^{-4})$ we find $S_T(\phi) = O(T^{-3})$.

This analytic result agrees with numeric results found by approximating $S_T(\phi)$ directly for a variety of different T values.

Because $S_T(\phi) \to 0$ as $T \to \infty$, we see that eq. 7 gives no information. This is a reflection of the fact that ϕ need not overcome any energy *barrier* in its path from 0 to U_f, but merely an energy plateau.

We can however apply equation 8 to find

$$\lim_{\varepsilon \to 0} \varepsilon^2 \ln[P(\tau \leq T)] = -\min_\phi \bar{S}_T(\phi) = \frac{1}{3} CT, \tag{14}$$

$$\Rightarrow P(\tau \leq T) \sim \exp \left[-kT^{-3} \varepsilon^{-2} \right], \tag{15}$$
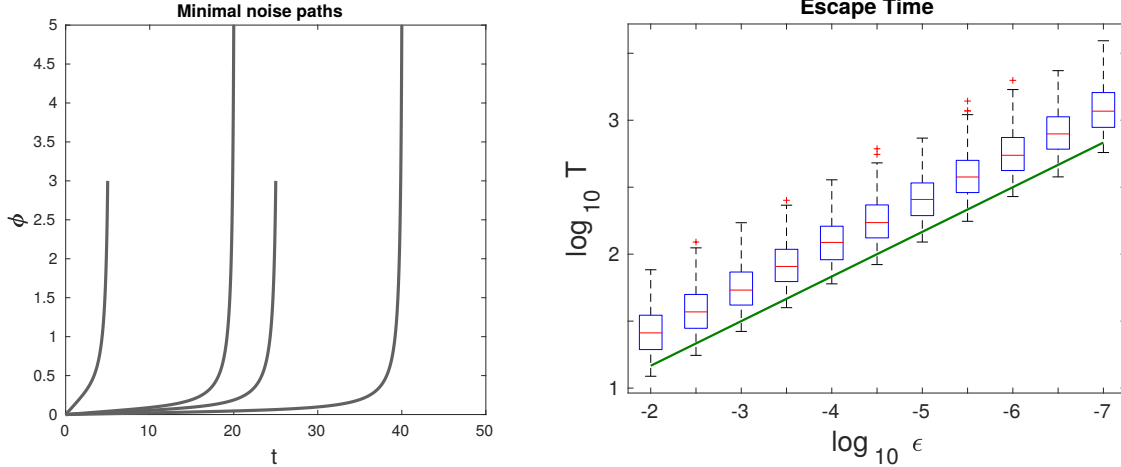
Figure 4: (Left) Some optimal ϕ found for four particular boundary conditions, $\tau(U_f) = T$, using $\tau(\phi)$, as given by wolfram|alpha [Wolfram|Alpha(2018a)]. In all cases, $\tau(U_f)$ matches the requested T to an exceptionally high degree of accuracy. (Right) Passage time from $U_i = 0$ to $U_f = 1$ vs ε for $u_t = u^2 + \varepsilon\xi$ (equation 4). For each ε value we use 4000 simulations. Boxplots indicate distribution of passage times. The solid green line indicates the predicted scaling of passage time $T = O(\varepsilon^{-2/3})$. The slope from our theory matches simulation results. The vertical shift between theory and simulation is a reflection of the unknown coefficient in the scaling relationship.

for small ε.

$$\mathbb{E}(\tau) = \int_0^\infty P(\tau > T)dt \approx \int_0^\infty 1 - \exp\left[-kT^{-3}\varepsilon^{-2}\right]dT, \tag{16}$$

$$= O(\varepsilon^{-2/3}) \tag{17}$$

Hence we find that time taken for a solution to $u_t = u^2 + \varepsilon\xi$ to escape from zero and approach infinity is predicted to scale like $O(\varepsilon^{-2/3})$. The majority of this time is spent close to $u = 0$, and the escape time behavior of the system is dominated by the behavior of the system in this region; the exact value of U_f has negligible impact, so long as U_f is selected such that $U_f \gg O(\varepsilon^{2/3})$. Simulation of equation 4 for a variety of ε agrees with these asymptotic results (see figure 4, Right).

3.1 Systems near a saddle-node bifurcation

The above work describes the behaviour of the system precisely balanced *at* a saddle-node bifurcation. For the sake of completeness, we might also consider systems in the vicinity of such a bifurcation.

The system

$$\dot{u} = u^2 + \gamma^2 + \varepsilon\xi \tag{18}$$

is a canonical example of a system which, for small γ, has just lost its steady states. In the deterministic limit $\varepsilon \to 0$, this system admits solutions of the form $T = \left[\tan^{-1}(u/\gamma)/\gamma\right]_{U_i}^{U_f}$, which gives travel times of order $O(\gamma^{-1})$.

The system

$$\dot{u} = u^2 - \gamma^2 + \varepsilon\xi \tag{19}$$

is a canonical example of a system with a stable/unstable pair of steady states. Assuming in this case that $U_i = -\gamma$ (the stable steady state), and $U_f > \gamma$, it is possible to show that the action minimizing path obeys $\dot{\phi}^2 = (\phi^2 - \gamma^2)^2 + C^2$. In the limit of large T, we can select $C \to 0$, and find $\bar{S}_T(\phi) \to 16\gamma^3/3$.

By eq. 7 we thus have $\mathbb{E}(\tau) \sim \exp[\varepsilon^{-2}16\gamma^3/3]$.

In addition to the above static results, the behavior of systems passing through saddle node bifurcations (either with or without noise), has also been studied. A review of this "delayed bifurcation theory" is given by Christian Kuehn [Kuehn(2011)]. For a more detailed introduction, see Berglund and Gentz [Berglund and Gentz(2006)]. We now move on to consider the previously unexplored spatially distributed problem.

4 The Spatially Distributed Problem

Now that we have built up our understanding using the spaceless model, let us turn our attention to the spatially distributed model. Our primary interest in the spatially distributed case is considering the behavior of systems at a saddle-node bifurcation. Spatial systems of this form can be represented via the canonical equation

$$u_t = u_{xx} + u^2 + \varepsilon \xi. \tag{20}$$

Here, both u and ξ are functions of x and t. We assume that u starts at the steady state; $u(x,0) = 0$. Here ξ represents our white noise term, although it may be thought of as a forcing function, with an associated energy $\int \xi^2 dx dt$.

We wish to describe the behavior of the system as it 'escapes' from the steady state at zero. We assume (without loss of generality) that u takes its maximum at $x = 0$, and ask "for a given time T, what is the most probable path such that $u(0,T) = U_f$, where $u(x,t) < U_f$ for all $t < T$?"

For the sake of comparison, we will also consider escape from a linear unstable stead state,

$$u_t = u_{xx} + u + \varepsilon \xi. \tag{21}$$

While less mathematically interesting, this linear escape problem gives us an analytically tractable case to explore, along with something to compare our non-linear results to. We begin with this simpler case.

4.1 Linearly Unstable Case

In what follows we present only the most important results, and the conceptually important steps leading to these results. Details can be found in Appendix C. In order to understand spike formation in equation (21), we must minimize the associated "action functional":

$$S_T(u) = \int_0^T \int_{-\infty}^{\infty} (u_t - u_{xx} - u)^2 dx dt, \tag{22}$$

subject to the conditions $u(x,0) = 0$ and $u(0,T) = U_f$. As per standard functional analysis techniques, this minima can be found when the functional derivative $\frac{dS_T(u)}{du}(x,t) = 0$. This condition *can* be written as a single differential equation with double the number of derivatives with respect to each variable, but is more conveniently written as a coupled *pair* of PDEs:

$$
\begin{aligned}
u_t &= u_{xx} + u + \varepsilon \xi, \\
u(x,0) &= 0, \\
-\xi_t &= \xi_{xx} + \xi, \\
\xi(x,T) &= \alpha \delta(x).
\end{aligned}
\tag{23}
$$

This system admits the explicit solution:

$$\xi(x,t) = \alpha e^{T-t} N(x, T-t), \tag{24}$$

$$u(x,t) = \int_0^t \alpha \varepsilon e^{T+t-2\tau} N(x, T+t-2\tau) d\tau, \tag{25}$$

where $N(x,\gamma) = \frac{1}{\sqrt{4\pi\gamma}} \exp\left[\frac{-x^2}{4\gamma}\right]$, and $\alpha = \frac{4U_f}{\varepsilon erfi(\sqrt{2T})}$.

The associated action functional can be shown to be $S_T(u) = 4U_f^2/erfi(2T) = U_f^2 O(e^{-2T})$.

Further analysis and discussion of these results we postpone until section 5.1.

4.2 Semi-stable Case

We now consider the non-linear case, (equation 20). As previously, in order to determine the dynamics of spike formation, we concern ourselves primarily with determining the most probable path to spike formation. We present here only the conceptually important milestones in our calculations. Details are similar to those given in Appendix C.

$$\text{Minimizing } S_T(u) = \int_0^T \int_{-\infty}^{\infty} (u_t - u_{xx} - u^2)^2 dx dt \tag{26}$$

subject to the constraints $u(x,0) = 0$ and $u(0,T) = U_f$, we find that the minimizer $u(x,t)$ satisfies

$$
\begin{aligned}
u_t &= u_{xx} + u^2 + \varepsilon\xi, \\
u(x,0) &= 0, \\
-\xi_t &= \xi_{xx} + 2u\xi, \\
\xi(x,T) &= \alpha\delta(x).
\end{aligned}
\tag{27}
$$

Unlike the linear case, where ξ could be solved independently of u, and then used to determine u explicitly, here no such analytic solution is available. ξ and u are inextricably coupled. Solving 27 numerically in either direction in time requires that we solve the ill posed backwards heat equation, either for u or ξ, depending on which direction we solve in. Creating a full spatial-temporal mesh and solving for both ξ and u simultaneously is possible, but quickly becomes computationally expensive for finer meshes. We can, however, solve iteratively. This is done by initially assuming $u_0(x,t) = 0$, and solving

$$-\hat{\xi}_t = \hat{\xi}_{xx} + 2u_0\hat{\xi}$$

with the boundary condition $\hat{\xi}(x,T) = \delta(x)$, to determine $\hat{\xi}_1$. We use $\alpha\hat{\xi}_1$ as a forcing term to find u_1, and u_1 to determine $\hat{\xi}_2$ and so on, until $|S(u_i) - S(u_{i+1})| < 0.0005|S(u_i) + S(u_{i+1})|$. (See figure 5). Using this iteration scheme, we are able to determine u, ξ and $S_T(u)$ for the action minimizing u profile. The value of α in each iteration step is found using a bisection method in which α is reduced if $u(0,T) > U_f$ and increased if $u(0,T) < U_f$.
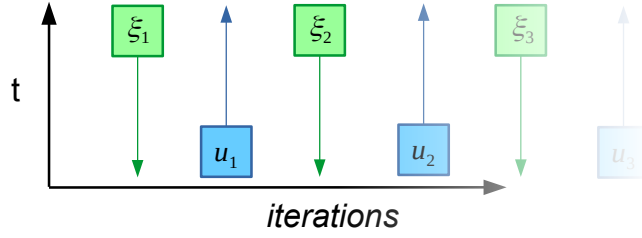

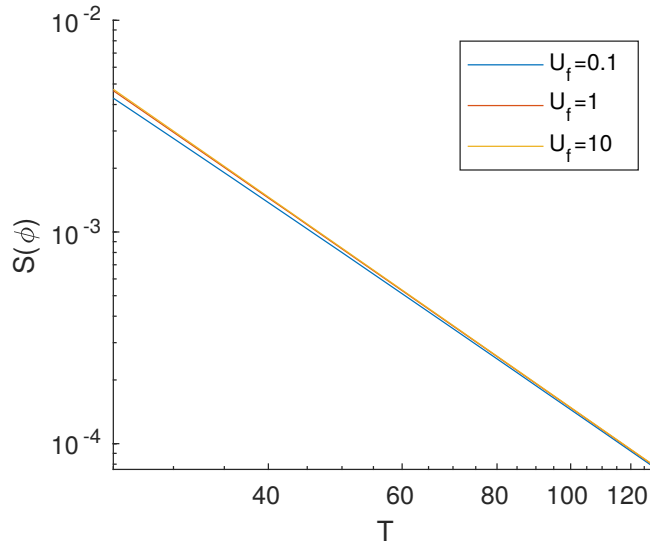Figure 5: Schematic view of our iteration scheme, as used to solve the system of equations (27).

Figure 6: Numerical calculation of the action cost $S_T(\phi)$ to reach $u(0,T) = U_f$ by time $t = T$ for equation (20), the spatial non-linear case. We consider $U_f = 0.1, 1$ and 10; the resulting differences are negligible. The resulting lines have mean slope $-2.465, -2.495 -2.498$ in log-log space, which we take to be a good approximation of $5/2$. Note that this $5/2$ result is explicitly an approximation based on numeric results, and it is possible (though presumably not likely) that the exact asymptotic result is some other value.
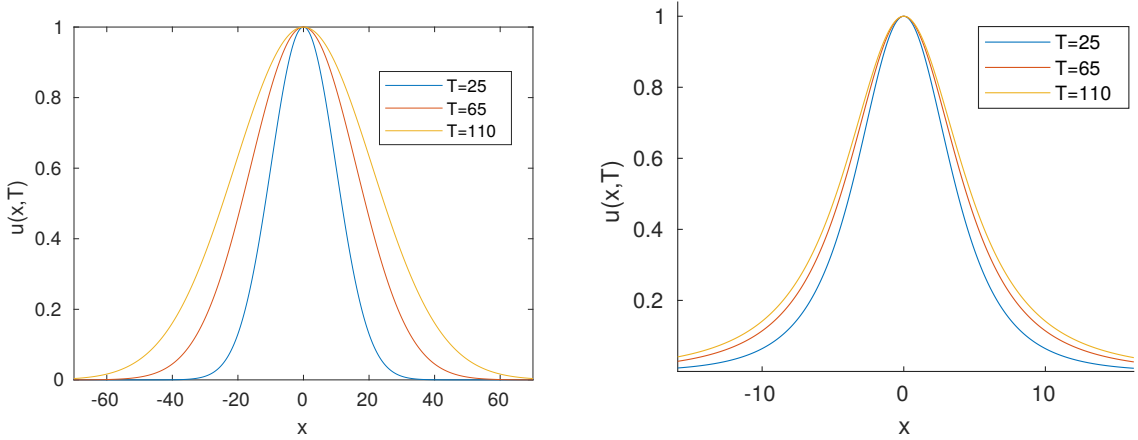
Figure 7: (Left) In the case of linear instability (equation 21) The profile of $u(x,T)$ is well approximated by $N(x,2T)$. The width of our peak varies significantly depending on T. (Right) In the case of quadratic instability (equation 20), as might be observed in the vicinity of a saddle-node bifurcation, the profile of $u(x,T)$ is relatively insensitive to T, and appears to approach some limiting "canonical burst" as T is made large.

5 Comparisons of results

5.1 Comparison Between Results for Linear and Non-Linear Case

Now that we have determined the action minimizing profile in both the linear and non-linear case (equations (21) and (20), respectively), let us now compare these two models.

To begin, we would like to determine the time till burst formation in each model. For the sake of notation, let us define $m(T) = \min_\phi S_T(\phi)$. While under normal circumstances we would find the expected escape time by invoking equation 7, in all cases studied here, $m(T) \to 0$ as $T \to \infty$. Hence equation 7 provides us with no information, and we instead rely on equation 8.

Regardless of the particular system under study, when $m(T) \gg \varepsilon^2$ equation 8 implies $-\ln P(\tau \leq T) \gg 1$ and hence $P(\tau \leq T) \approx 0$. This in turn implies that the probability *density* of burst formation for such values of T is low. Similarly, when $m(T) \ll \varepsilon^2$ we have $-\ln P(\tau \leq T) \ll 1$ and hence $P(\tau \leq T) \approx 1$, and once again the probability of escape is negligible, as escape has almost certainly already occurred. Escape can only have non-negligible probability density when $\ln P(\tau \leq T)$ is neither too big nor too small, namely when $m(T)$ and ε^2 are similar orders of magnitude.

This implies that for the linear model escape is predicted when $\varepsilon^2 \approx \bar{S}_T = U_f^2 O(e^{-2T})$, or equivalently when $T \sim -\log(\varepsilon)$. The time of escape changes little, even as ε is varied over several orders of magnitude. In the nonlinear case we observe that $m(T) = O(T^{-5/2})$ as $T \to \infty$ (see figure 6). By the above arguments, we predict $T \sim \varepsilon^{-4/5}$; the time till burst formation is *sensitive* to the amplitude of the noise driving the system. Table 1 summarises these timing results, and compares to the previously discussed spaceless models.

Far more interesting than the difference in escape times for the linear and quadratic case is the substantial difference in the profile of our action-minimizing u. In the linear case, $u(x,t)$ was found explicitly (equation 25). In this case $u(x,T)$ is well approximated by a corresponding normal distribution $N(x,2T)$, and becomes wider over time (see figure 5.1, upper). In contrast, in the quadratic case, the final width of u increases at most very slowly, and appears to approach a limiting distribution as $T \to \infty$ (see figure 5.1, lower).

As a final point of comparison, we consider the reaction of the two systems to changes in U_f. As might be expected, the *shape* of the action-minimising u profile is unaffected by changes in U_f for the linear system. Doubling U_f for a given T doubles both ξ and u, but does not change the shape of either. In contrast, in the non-linear case, the profile of $u(x,t)$ is sensitive to changes in U_f, with larger U_f leading to a sharper "spike" solution, and smaller U_f leading to soft "bumps" (see figure 8).
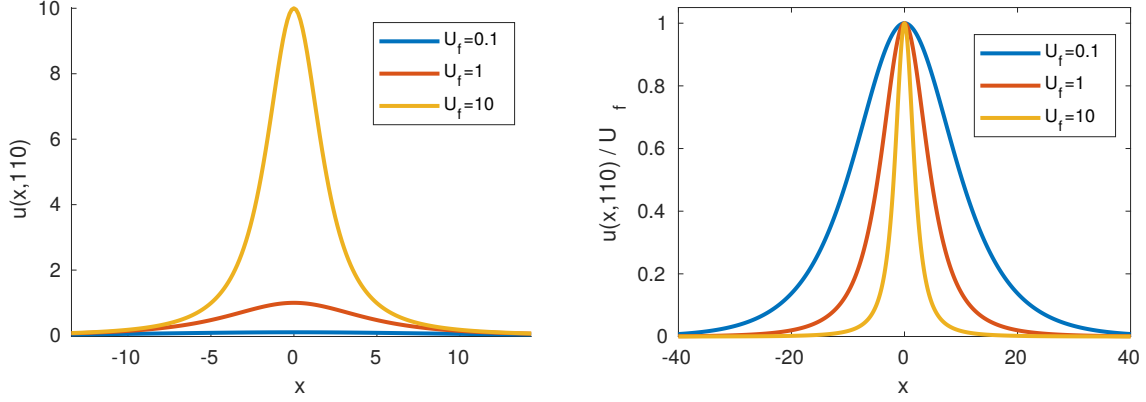
Figure 8: (Left) For fixed $T = 110$, we find $u(x,t)$ for a variety of U_f. As U_f is increased, $u(x,T)$ increases monotonically, however the most significant gains happen near $x = 0$, and hence, the profile becomes more 'spike' like, as can be seen in the amplitude normalized plots (Right)

Equation	$m(T)$	$\mathbb{E}(T) \sim$
$u_t = u^2 + \varepsilon\xi$	$O(T^{-3})$	$\varepsilon^{-2/3}$
$u_t = u^2 + \gamma^2 + \varepsilon\xi$	0	γ^{-1}
$u_t = u^2 - \gamma^2 + \varepsilon\xi$	$16\gamma^3/3$	$\exp[\varepsilon^{-2}16\gamma^3/3]$
$u_t = \Delta u + u + \varepsilon\xi$	$U_f O(e^{-2T})$	$\log(\varepsilon)$
$u_t = \Delta u + u^2 + \varepsilon\xi$	$O(T^{-5/2})$	$\varepsilon^{-4/5}$

Table 1: Comparison of action functionals and approximate escape times.

5.2 Comparison of Analytic Results to Direct Simulations

As a final check on the above analytic results, we compare to simulations of the corresponding systems.

In the linear case (equation 21), we find that the time until $\max|u(x,t)| = U_f = 1$ scales like $-\log(\varepsilon)$, as predicted by theory (see figure 9, left). In the non-linear case (equation 20), burst time is reasonably approximated by $\varepsilon^{-4/5}$ as predicted by theory, but the agreement is weaker than might be hoped (see figure 9, right). It is unclear if this disagreement stems from the limitations of our simulations, or the numeric-analytic arguments proposed earlier.

Comparison of shape in the linear case finds reasonable agreement in the vicinity of the primary burst itself, however because a large number of bursts form simultaneously, the final profile of u ends up being an overlapping combination of many different such bursts (see figure 10,left).

In contrast, comparison between the predicted and observed shape of bursts shows strong agreement in the non-linear case, even for relatively large values of ε (see figure 10,right). Bursts spend a greater portion of their development time close to zero in the non-linear case, and hence, although many bursts may begin growing at roughly the same time, the second and third place bursts will have only a minor effect on the overall profile of $u(x,T)$.

The code for all simulations can be found in the supplementary materials.

6 Difficulties in 2D, and higher dimensions

Given that the universe we occupy is not one dimensional, it would be beneficial to extend the above results to higher dimensions, preferably three dimensions, although for certain contexts, such as the binding of proteins to a cell membrane[Vecchiarelli et al.(2016)Vecchiarelli, Li, Mizuuchi, Hwang, Seol, Neuman, and Mizuuchi], two dimensions would suffice.

Unfortunately, the properties of white noise preclude this avenue of investigation; as noted by Ryser et al [Ryser et al.(2012)Ryser, Nigam, and Tupper], equations of the form $u_t = \Delta u + \xi$ do not remain well posed in dimension two or higher. The smoothing effects of diffusion are insufficient to restrain the irregularity produced by space-time white noise.
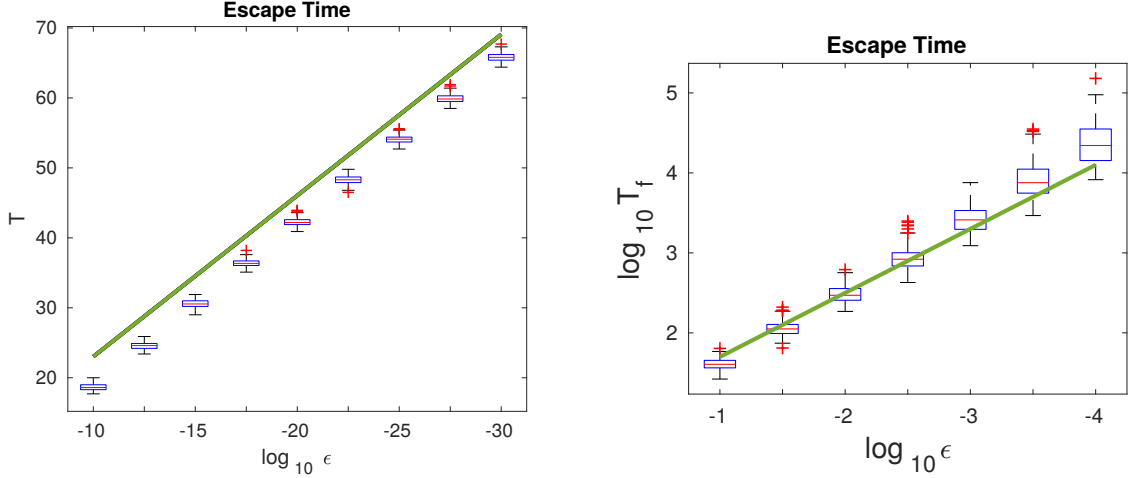
Figure 9: (Left) Passage times from $u(x,0) = 0$ to $\max |u(x,t)| = 1$, for equation 21. For each ε value we use 300 simulations. Box plots indicate simulated passage times, while the solid green line indicates the predicted scaling $T \approx -log(\varepsilon)$. Given that theory only guarantees this approximation correct up to an order of magnitude, the observed agreement between theory and simulation is gratifying. The vertical shift between theory and simulation is a reflection of the unknown coefficient in the scaling relationship. (Right) Similar simulations for equation 20. For each ε value we use 250 simulations. Box plots indicate simulated passage times, while the solid green line indicates the predicted scaling $\log(T) \approx -4/5 \log(\varepsilon)$. While predictions are reasonable, linear regression indicates that the best fitting slope is -0.924344 ± 0.003856, however even this assertion must be treated with some skepticism- as the graph does not appear to obey a linear relationship for small values of ε. Given that the $\log(T) \approx -4/5 \log(\varepsilon)$ prediction is based on extrapolation from numerical results, it is unclear whether the discrepancy comes from inaccuracies in our simulation method or is a result of genuine limitations on the $-4/5$ relation itself.

This result can be seen through a variety of lenses, depending on ones outlook. Ryser et al, describe u in terms of a sum of Fourier modes, and show

$$\mathbb{E}||u||^2 = \sum \varepsilon(1 - e^{-2(1+|\mathbf{k}|^2)t})/2(1 + |\mathbf{k}|^2).$$

In one dimension, when $\mathbf{k} \in \mathbb{Z}$ this sum converges, however for two or more dimensions, $\mathbf{k} \in \mathbb{Z}^d$ and the sum is unbounded for all $t > 0$. This indicates that u is not well defined. Simulations in 2d, such as those presented in section 5.1 function for fixed dx, but invariably break down as $dx \to 0$.

In terms of the methodology used in this paper, the degeneracy of higher dimensions manifests as a break down of various integrals. Issues arise regardless of what equation is used, but the problem is most easily illustrated by considering the linear case. Consider the solution $u(x,t)$ of eq. 21 given in eq. 25. The parameter α of eq. 25 is determined by considering the boundary condition $u(0,T) = U_f$,

$$U_f = \frac{\alpha\varepsilon}{2} \int_0^{2T} e^{\gamma} N(0, \gamma) d\gamma,$$

where $\gamma = T + t - 2\tau$.

In 1D $N(\mathbf{x}, \gamma) = (4\pi\gamma)^{-1/2} \exp\left[-x^2/4\gamma\right]$, and hence $U_f = \frac{\alpha\varepsilon}{4\sqrt{\pi}} \int e^{\gamma} \gamma^{-1/2} e^{-0/\gamma} d\gamma$. Because the integral converges to a finite number, we can rearrange to determine α and go on to find $S(u)$ explicitly for any U_f.

By contrast, in 2D we have $N(\mathbf{x}, \gamma) = (4\pi\gamma)^{-1} \exp\left[-||\mathbf{x}||^2/4\gamma\right]$, and hence, $U_f = O(\alpha\varepsilon) \int e^{\gamma} \gamma^{-1} e^{-0/\gamma} dx$. The integral is unbounded, and thus we are able to select α arbitrarily small. This in turn allows both T and $S(u)$ to be selected arbitrarily close to zero, regardless of U_f. Physically, we are picking ξ to be a spike that is sufficiently narrow so as to ensure $S(u)$ arbitrarily small, yet also sufficiently high so as to force u from zero to U_f in time T. This sleight of hand is impossible in one dimensional systems, but for all higher dimensions it is the "action minimizing" strategy.

While here we have described the difficulties of integration in the linear case, the problems described extends to all equations that include noise terms in 2D, and become progressively worse for higher dimensions. They are a symptom, rather than the root cause of the problems of noise in higher dimensions.
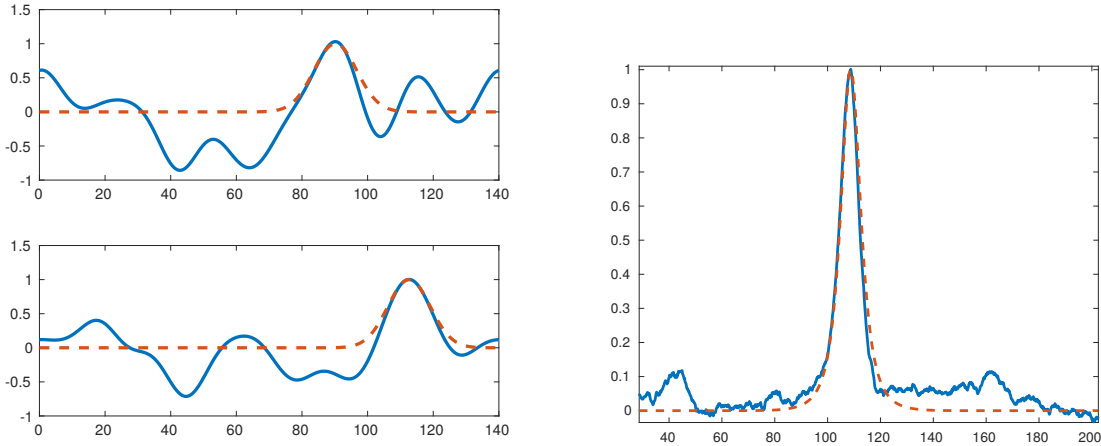
Figure 10: (Left) Final profiles of $u(x,T)$ (solid line), along with the accompanying "ideal burst" (dashed line) for two simulations of equation 21. (Right) Final profile of $u(x,T)$ (solid line), along with the accompanying "ideal burst" for a single simulation of equation (20) with $\varepsilon = 0.1$. In both cases, no fitting was used beyond a horizontal shift moving the peak of our analytic solution in line with the peak realized in simulation. Peak shape is determined entirely by the asymptotic arguments previously described.

7 Conclusions, Discussion and Future work

In this article we have presented a number of prototypical models exploring the behavior of systems starting at unstable and semi-stable steady states.

We have demonstrated both analytically and through simulations that under the influence of spatio-temporal white noise, systems of the form $u_t = u_{xx} + u^2 + \varepsilon\xi$ develop well defined "spikes" in finite time. This phenomena can be seen as delayed bifurcation [Kuehn(2011)] in a spatially distributed system, and also as a compliment to non-linear blow up phenomena, in which we consider instead the system's "escape from zero", as opposed to the more typically studied "approach to infinity" [Vázquez and Galaktionov(2002)].

The symmetry breaking observed here does not rely on multiple chemical species, nor on contrasts in diffusion rate, as might be expected for the more classical Turing type pattern formation.

We were able to determine both the "energy cost" of spike formation, along with the shape of the most probable spike, and showed that in the limit of large time, spike formation requires infinitesimally small energy. The sensitivity of the system in the vicinity of its semi-stable steady state allows noise of amplitude $\varepsilon \ll 1$ to have macroscopic effects in $O(\varepsilon^{-4/5})$ time. We also explored similar results both in the spaceless case, and in the case where our $u = 0$ state is linearly unstable as opposed to semi-stable.

In terms of physical relevance, the systems studied can be best thought of as a prototype for chemical reaction-diffusion systems in the vicinity of a saddle-node bifurcation. Our study demonstrates the breakdown of a homogeneous state, and can be seen as a first step towards understanding systems such as the Min system studied by Vecchiarelli et al [Vecchiarelli et al.(2016)Vecchiarelli, Li, Mizuuchi, Hwang, Seol, Neuman, and Mizuuchi], in which changes in the bulk concentration of a protein pushes the system through a bifurcation boundary and leads to the formation of membrane bound "bursts".

There are a number of further questions which must be answered before the work here can be applied to any experimental context. First and foremost, the work here considers a system perfectly balanced at the saddle-node; $u_t = u^2$. No physical system however is ever so perfectly balanced, and so determining the robustness of these results for a system that is *passing through* such a saddle node bifurcation is critical to our understanding of the relevance of these results. Further in order for the results here to prove useful for experimental scientists, time must be invested in investigating the mapping between experimental observations, and the associated reaction and diffusion parameters implied; for example, does the relationship between spike width and height give a reliable signature that can be used to determine system parameters?

Finally, our work here alludes (in passing) to potentially deeper questions in chemistry; namely the difficulties in standard representations of noise when in higher dimensions. Such questions must be answered, or at the very least sidestepped, before the phenomena observed here can be sensibly applied to the multi dimensional systems ubiquitous in the real world.

8 Code

The Code used in this project is available on github at alastair-JL/StochasticBurst.

9 Acknowledgements

A Introduction to Large Deviation Theory

Here we give a very brief introduction to the principle ideas and techniques used in Large Deviation Theory (LDT), and used throughout this paper. In this appendix we present LDT as it applies to discrete time Stochastic processes. The continuous time case is conceptually similarly. Our goal here is to build intuition rather than mathematical rigor.

As a concrete example with which to frame our discussion, consider the Discrete Orenstien-Ulembeck process[Larralde(2004)]

$$u_i = 0.9u_{i-1} + \varepsilon \xi_i, \tag{28}$$

$$u_0 = 0. \tag{29}$$

Here ξ_i is assumed to be a normal random variable with mean zero and variance one, such that each ξ_i is independent. We assume $0 < \varepsilon \ll 1$.

Suppose we wish to know the first time that $u_i > 1$. At first glance, given $O(\varepsilon)$ noise, and the decay term of our OU process, $u_i > 1$ seems unlikely to occur. That said, as we take $i \to \infty$, unlikely, even exceedingly unlikely events should occur *eventually*. LDT concerns itself with such questions as "how long will it take for X to occur?" and "given that X occurs, what path is the system most likely to take in order to get there?". In our particular case X is "$u_i > 1$".

The total probability of our event X is equal to the integral over the probability density of all paths leading to that event. This integral can be formulated either in terms of the paths of the stochastic process $\mathbf{u}$, or in terms of the underlying noise $\boldsymbol{\xi}$.

$$P(X) = \int_{u \in X} p_u(\mathbf{u}) d\mathbf{u} = \int_{\xi \in X} p_\xi(\boldsymbol{\xi}) d\boldsymbol{\xi}. \tag{30}$$

The probability density function for a noise vector $\boldsymbol{\xi}$ of length N is

$$p_\xi(\boldsymbol{\xi}) = (2\pi)^{-N/2} e^{-\sum \xi_i^2/2}. \tag{31}$$

Unfortunately, even with this well defined probability density, the boundary of the integral $\int_{\xi \in X} p_\xi(\boldsymbol{\xi}) d\boldsymbol{\xi}$ is generically complicated enough so as to prevent us from evaluating $P(X)$ directly (the OU process being a notable exception).

In order to avoid this complex integration boundary, it is useful to transform our integral back into a form dependent on $\mathbf{u}$. To achieve this we rearrange our recurrence relation (equation 29), and find

$$\xi_i = \frac{u_i - 0.9u_{i-1}}{\varepsilon}. \tag{32}$$

Using this one-to-one correspondence between $\mathbf{u}$ and $\boldsymbol{\xi}$ along with the change of random variables formula, we find

$$P(X) = \int_{u \in X} p_u(\mathbf{u}) d\mathbf{u} = \int_{\xi \in X} p_\xi(\boldsymbol{\xi}) d\boldsymbol{\xi} \tag{33}$$

$$= (2\pi)^{-N/2} \int_{\xi \in X} \exp\left[-\sum \xi_i^2/2\right] d\boldsymbol{\xi} \tag{34}$$

$$= (2\pi)^{-N/2} \int_{u \in X} \exp\left[-\sum (u_i - 0.9u_{i-1})^2/2\varepsilon^2\right] d\mathbf{u} \tag{35}$$

$$= (2\pi)^{-N/2} \int_{u \in X} e^{-S(\mathbf{u})/\varepsilon^2} d\mathbf{u}. \tag{36}$$

Here $S(\mathbf{u})$ is said to be the "normalized action functional" of our problem. In the particular case discussed here $S(\mathbf{u}) = \sum(u_i - 0.9u_{i-1})^2/2$. In general $S(\mathbf{u})$ depends both on the equation governing a stochastic process, and the particular form of the noise generating it. $S(\mathbf{u})$ can be thought of as a measure of the "total improbability" associated with a given path, and is associated with the amount of "energy" that noise must pour into the system in order to cause a particular path to occur. When dealing with continuous systems, S is defined as an integral over the square of noise, rather than a sum.

At this stage, in order to determine the probability of our event, we need to evaluate $\int_{u \in X} e^{-S(\mathbf{u})/\varepsilon^2} d\mathbf{u}$. Typically, this integral can not be evaluated exactly, but it can be well approximated via Laplace's Principle [Laplace(1986), Olivieri and Vares(2005)].

Laplace's Principle states that:

$$\int_{u \in X} e^{-S(\mathbf{u})/\varepsilon^2} d\mathbf{u} \sim \exp\left[-\min S(\mathbf{u})/\varepsilon^2\right], \tag{37}$$

where here we minimize $S(\mathbf{u})$ over all $\mathbf{u} \in X$. Laplace's principle is based on the idea that for integrals of the from $\int_{u \in X} e^{-S(\mathbf{u})/\varepsilon^2} d\mathbf{u}$, the overwhelming majority of all probability mass is concentrated in the vicinity of $\min S(\mathbf{u})$ whenever $\varepsilon \ll 1$, (see figure 11). From the point of view of stochastic processes, what Laplace's Principle is effectively stating is that if an improbable event does occur, the observed path is overwhelmingly likely to be 'close' to the most probable path.
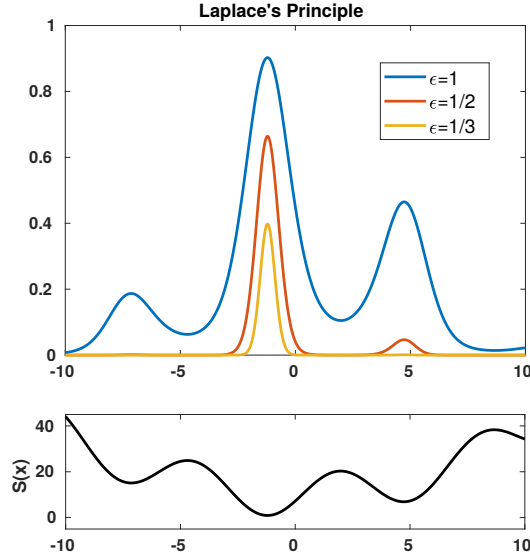

Figure 11: For some arbitrary function $S(x)$, as $\varepsilon \to 0$ the vast majority of the integral $\int e^{-S(x)/\varepsilon^2} dx$ can be found in a narrow window near the minimum of $S(x)$. As a result $\int e^{-S(x)/\varepsilon^2} dx$ scales like $e^{-S(x_0)/\varepsilon^2}$. In this particular example, for $\varepsilon = 1$ probability mass is spread across a number of separate "peaks" (local minima of $S(x)$). For $\varepsilon = 1/3$ by contrast, only the global minima of $S(x)$ has non-negligible mass associated with it. While varying ε does change the width of our global peak, this effect is negligible compared to the change in height resulting from $e^{-S(x_0)/\varepsilon^2}$.

Let us return to the original question proposed at the start of this section: how long does it take before the Discrete Orenstien-Ulembeck process defined in equation 29 exceeds one for the first time?

Suppose we wish to determine the probability that $u_N \geq 1$, for some particular N, assumed to be large. By Laplace's Principle, we must thus minimize $S(\mathbf{u}) = \sum(u_i - 0.9u_{i-1})^2/2$. Because any $u_N > 1$ will only increase $S(\mathbf{u})$ we can assume $u_N = 1$. In order to minimize, we require that $u_N = 1, u_0 = 0$ and $\frac{dS}{du_i} = 0$ for all i. As $N \to \infty$, solutions to the above can be well approximated by $u_i = 0.9^{N-i}$. Substituting back into the definition gives $S(\mathbf{u}) = 0.095$, and hence $P(u_N \geq 1) \approx \exp[-0.095/\varepsilon^2]$. If this is the probability of success for any particular large N, then we infer that the expected time until $u_i > 1$ will scale such that $\mathbb{E}(\tau) = O(e^{0.095/\varepsilon^2})$.

Freidln and Wentzell[Freidlin and Wentzell(2012)] provide the general formulation for the above two results. They state (chapter 4, theorem 1.2) that for small epsilon the probability of a particular rare event occurring by a given time

14

$P(\tau < T)$ is governed by:

$$\lim_{\varepsilon \to 0} \varepsilon^2 \ln P(\tau \le T) = -\min_{\phi} S_T(\phi). \tag{38}$$

Here, as is customary, ϕ denotes a *particular* trajectory of u and $S_T(\phi)$ indicates that we are minimising over all trajectories ϕ such that our rare event occurs by time T. This theorem is referred to as equation (8) in the main text.

B   Derivation of analytic expression for $\bar{S}_T$ in the spaceless case.

Here we present the derivation of equations 13. We begin by taking the definition of $\bar{S}_T$ and expanding. From eq. 9 we can show $\dot{\phi} = \sqrt{\phi^4 + C}$ and hence:

$$\bar{S}_T(\phi) = \int_0^T \dot{\phi}^2 - 2\dot{\phi}\phi^2 + \phi^4 dt \tag{39}$$

$$= \int_0^T \dot{\phi}^2 - 2\dot{\phi}\phi^2 + \phi^4 + C - C dt \tag{40}$$

$$= 2\int_0^T \dot{\phi}\sqrt{\phi^4 + C} - \dot{\phi}\phi^2 dt - \int_0^T C dt \tag{41}$$

$$= 2\int_0^{U_f} \sqrt{\phi^4 + C} d\phi - 2\int_0^{U_f} \phi^2 d\phi - \int_0^T C dt. \tag{42}$$

The problematic term here is the first integral. Applying integration by parts gives

$$\int_0^{U_f} \sqrt{\phi^4 + C} d\phi = \left[\phi\sqrt{\phi^4 + C}\right]_0^{U_f} - \int_0^{U_f} \phi \frac{4\phi^3}{2\sqrt{\phi^4 + C}} d\phi, \tag{43}$$

$$= \left[\phi\sqrt{\phi^4 + C}\right]_0^{U_f} - \int_0^{U_f} 2\frac{\phi^4 + C - C}{\sqrt{\phi^4 + C}} d\phi. \tag{44}$$

$$\int_0^{U_f} 3\sqrt{\phi^4 + C} d\phi = \left[\phi\sqrt{\phi^4 + C}\right]_0^{U_f} + \int_0^{U_f} \frac{2C}{\sqrt{\phi^4 + C}} d\phi. \tag{45}$$

By eq. 10 this last term is $2CT$. Substituting this result back in gives.

$$\bar{S}_T(\phi) = \frac{2}{3}\left[\phi\sqrt{\phi^4 + C}\right]_0^{U_f} + \frac{4}{3}CT - 2\int_0^{U_f} \phi^2 d\phi - \int_0^T C dt, \tag{46}$$

$$= \frac{2}{3}\left[\phi^3\sqrt{1 + C\phi^{-4}}\right]_0^{U_f} - \frac{2}{3}\left[\phi^3\right]_0^{U_f} + \frac{1}{3}CT, \tag{47}$$

$$= \frac{2}{3}\left[\phi^3\left(\sqrt{1 + C\phi^{-4}} - 1\right)\right]_0^{U_f} + \frac{1}{3}CT, \tag{48}$$

hence recovering equation 13.

At no point in this derivation have we used the approximation of C given in eq 11, hence this is an analytic result.

C   The linearly unstable case; full calculation

Here we provide the detailed algebra suppressed in section 4.1. Consider the minimization problem:

$$\text{Given } u_t = u_{xx} + u + \varepsilon\xi,$$
$$u(x,0) = 0,$$
$$\text{minimise } S(u(\xi)) = \int_0^T \int_{-\infty}^{\infty} \varepsilon^2 \xi^2 dx dt$$
$$\text{such that } u(0,T) = U_f. \tag{49}$$

This is equivelent to equation (26) of the main text.

If we consider ξ as a form of forcing function, then the above asks for the lowest energy forcing required to lift $u(0,T)$ to height U_f. Thinking of ξ as a form of noise we are asking for the *most probable* noise. For the rest of this appendix we will think of ξ in terms of energy input, although we will keep the white noise formulation in mind. Please note that while both u and ξ are stochastic processes, in a slight abuse of notation we also use the symbols to indicate the optimal path subject to the above constraints.

In order to find our minima, we first re-write S in terms of u, and then take the functional derivative:

$$S = \int_0^T \int_{-\infty}^{\infty} \varepsilon^2 \xi^2 dx dt = \int_0^T \int_{-\infty}^{\infty} [u_t - u_{xx} - u^2]^2 dx dt. \tag{50}$$

From the definition of functional derivatives we have

$$\iint \frac{\delta S}{\delta u} \psi dx dt = \lim_{h \to 0} \frac{S(u + h\psi) - S(u)}{h}. \tag{51}$$

Combining and taking limits gives

$$\iint \frac{\delta S}{\delta u} \psi dx dt = 2 \iint (\psi_t - \psi_{xx} - \psi)(u_t - u_{xx} - u) dx dt. \tag{52}$$

Remembering that $\varepsilon \xi = u_t - u_{xx} - u$ gives

$$\iint \frac{\delta S}{\delta u} \psi dx dt = 2\varepsilon \iint \psi_t \xi - \psi_{xx} \xi - \psi \xi dx dt. \tag{53}$$

Now using integration by parts on each ψ term, as appropriate, we find

$$\iint \frac{\delta S}{\delta u} \psi dx dt = 2\varepsilon \int [\psi \xi]_{t=0}^{t=T} dx + 2\varepsilon \int [\psi \xi_x - \psi_x \xi]_{x=-\infty}^{x=\infty} dt$$
$$+ 2\varepsilon \iint -\psi \xi_t - \psi \xi_{xx} - \psi \xi dx dt. \tag{54}$$

Because we seek to minimise $\int \int \xi^2 dx dt$, we can safely demand that $\xi, \xi_x \to 0$ as $x \to \pm \infty$ and thus $[\psi \xi_x - \psi_x \xi]_{-\infty}^{\infty} = 0$. We are left with the potentially problematic $\int [\psi \xi]_0^T \xi dx$ term, however, since u is fixed at zero for $t = 0$ and at U_f for $x = 0, t = T$ we know ψ (our perturbation to u) is zero in these location. Further, we can assume that $\xi = 0$ for $t = T, x \neq 0$, as any non-zero forcing at these locations would increase S, but have no effect on $u(0,T)$. Hence $\int [\psi \xi]_0^T \xi dx = 0$, and so

$$\frac{\delta S}{\delta u} = 0 = -\xi_t - \xi_{xx} - \xi. \tag{55}$$

The above can be stated as a fourth order differential equation purely in terms of u, but this provides little illumination. Instead we leave it as a backwards heat equation, ready to be coupled to 49.

As argued previously, for our optimal solution $\xi(x,T) = 0$ whenever $x \neq 0$, as non-zero values of ξ in these locations increase $S(\xi)$ but have no effect on $u(0,T)$. In order for ξ to have any effect on u it must have non-zero total mass, and since $\int -\xi_t dx = \int \xi dx$ we know that it must have non-zero total mass at time T. Hence $\xi(x,T) = \alpha \delta(x)$ for some α.

Combining the above, the minimal noise carrying u from 0 to $u(0,T) = U_f$ must solve:

$$\begin{aligned} u_t &= u_{xx} + u + \varepsilon \xi, \\ u(x,0) &= 0, \\ -\xi_t &= \xi_{xx} + \xi, \\ \xi(x,T) &= \alpha \delta(x). \end{aligned} \tag{56}$$

This recovers equation (27) of the main text.

Solving directly gives

$$\xi(x,t) = \alpha e^{T-t} N(x, T-t) \tag{57}$$

$$u(x,t) = \int_0^t \alpha \varepsilon e^{T+t-2\tau} N(x, T+t-2\tau) d\tau \tag{58}$$

Where

$$N(x, \gamma) = \frac{1}{\sqrt{4\pi\gamma}} \exp\left[\frac{-x^2}{4\gamma}\right] \tag{59}$$

The noise scaling α can be found using our boundary condition:

$$U_f = u(x,t) \tag{60}$$

$$= \int_0^T \alpha \varepsilon e^{2T-2\tau} N(0, 2T - 2\tau) d\tau \tag{61}$$

$$= \alpha \varepsilon \int_{2T}^0 e^{\gamma} \frac{1}{\sqrt{4\pi\gamma}} \frac{d\gamma}{-2} \tag{62}$$

$$= \frac{\alpha \varepsilon}{4\sqrt{\pi}} \int_0^{2T} \gamma^{-1/2} e^{\gamma} d\gamma \tag{63}$$

$$= \frac{\alpha \varepsilon}{4} erfi(\sqrt{2T}) = \frac{\alpha \varepsilon}{4} O(e^{2T}) \tag{64}$$

In order to find the normalized action functional we find:

$$S(\xi) = \iint \varepsilon^2 \xi^2 dx dt, \tag{65}$$

$$= \varepsilon^2 \alpha^2 \iint e^{2T-2t} N(x, T-t) N(x, T-t) dx dt, \tag{66}$$

$$= \varepsilon^2 \alpha^2 \iint e^{2T-2t} \frac{N(x, [T-t]/2)}{\sqrt{8\pi(T-t))}} dx dt. \tag{67}$$

Remembering that $\int N(x, \gamma) dx = 1$ for any γ, followed by the substitution $2(T-t) = \gamma$

$$= \frac{\varepsilon^2 \alpha^2}{4\sqrt{\pi}} \int_0^{2T} \gamma^{-1/2} e^{\gamma} d\gamma. \tag{68}$$

$$= \varepsilon \alpha U_f = 4U_f^2 / erfi(2T) = U_f^2 O(e^{-2T}) \tag{69}$$

Escape is predicted to occur when $\varepsilon^{-1} = O(e^{2T})$. For all but the smallest ε values, this can be considered to take place in $O(1)$ time.

Conceptually similar calculations can be used to get from equation (26) to (27) in the non-linear case.

References

[Artin(2015)] E. Artin. *The Gamma Function*. Courier Dover Publications, Jan. 2015. ISBN 978-0-486-80300-5. Google-Books-ID: c3R2BgAAQBAJ.

[Berglund and Gentz(2006)] N. Berglund and B. Gentz. *Noise-Induced Phenomena in Slow-Fast Dynamical Systems: A Sample-Paths Approach*. Probability and Its Applications. Springer-Verlag, London, 2006. ISBN 978-1-84628-038-2. URL https://www.springer.com/gp/book/9781846280382.

[Erdemir et al.(2009)Erdemir, Lee, and Myerson] D. Erdemir, A. Y. Lee, and A. S. Myerson. Nucleation of Crystals from Solution: Classical and Two-Step Models. *Accounts of Chemical Research*, 42(5):621–629, May 2009. ISSN 0001-4842. doi: 10.1021/ar800217x. URL https://doi.org/10.1021/ar800217x.

[Freidlin and Wentzell(2012)] M. I. Freidlin and A. D. Wentzell. *Random Perturbations of Dynamical Systems*. Grundlehren der mathematischen Wissenschaften. Springer-Verlag, Berlin Heidelberg, 3 edition, 2012. ISBN 978-3-642-25846-6. URL //www.springer.com/gp/book/9783642258466.

[Kuehn(2011)] C. Kuehn. A mathematical framework for critical transitions: Bifurcations, fast–slow systems and stochastic dynamics. *Physica D: Nonlinear Phenomena*, 240(12):1020–1035, June 2011. ISSN 0167-2789. doi: 10.1016/j.physd.2011.02.012. URL `http://www.sciencedirect.com/science/article/pii/S0167278911000443`.

[Laplace(1986)] P. S. Laplace. Memoir on the Probability of the Causes of Events. *Statistical Science*, 1(3):364–378, 1986. ISSN 0883-4237. URL `http://www.jstor.org/stable/2245476`.

[Larralde(2004)] H. Larralde. A first passage time distribution for a discrete version of the Ornstein–Uhlenbeck process. *Journal of Physics A: Mathematical and General*, 37(12):3759, 2004. ISSN 0305-4470. doi: 10.1088/0305-4470/37/12/003. URL `http://stacks.iop.org/0305-4470/37/i=12/a=003`.

[Lefever and Lejeune(1997)] R. Lefever and O. Lejeune. On the Origin of Tiger Bush. *Bulletin of Mathematical Biology*, 59:263–294, Mar. 1997. doi: 10.1007/BF02462004.

[McGraw and Liu(2003)] R. McGraw and Y. Liu. Kinetic Potential and Barrier Crossing: A Model for Warm Cloud Drizzle Formation. *Physical Review Letters*, 90(1):018501, Jan. 2003. doi: 10.1103/PhysRevLett.90.018501. URL `https://link.aps.org/doi/10.1103/PhysRevLett.90.018501`.

[Mori et al.(2008)Mori, Jilkine, and Edelstein-Keshet] Y. Mori, A. Jilkine, and L. Edelstein-Keshet. Wave-Pinning and Cell Polarity from a Bistable Reaction-Diffusion System. *Biophysical Journal*, 94(9):3684–3697, May 2008. ISSN 0006-3495. doi: 10.1529/biophysj.107.120824. URL `http://www.sciencedirect.com/science/article/pii/S0006349508704442`.

[Murray(2008)] J. D. Murray. *Mathematical Biology II: Spatial Models and Biomedical Applications*. Springer, New York, 3rd corrected ed. 2003. corr. 3rd printing 2008 edition, Aug. 2008. ISBN 978-0-387-95228-4.

[Olivieri and Vares(2005)] E. Olivieri and M. E. Vares. *Large Deviations and Metastability*. Cambridge University Press, Feb. 2005. ISBN 978-0-521-59163-8.

[Rassoul-agha and Seppalainen(2015)] F. Rassoul-agha and T. Seppalainen. *A Course on Large Deviations With an Introduction to Gibbs Measures*. Amer Mathematical Society, Providence, Rhode Island, Apr. 2015. ISBN 978-0-8218-7578-0.

[Ryser et al.(2012)Ryser, Nigam, and Tupper] M. D. Ryser, N. Nigam, and P. F. Tupper. On the well-posedness of the stochastic Allen-Cahn equation in two dimensions. *Journal of Computational Physics*, 231(6):2537–2550, Mar. 2012. ISSN 00219991. doi: 10.1016/j.jcp.2011.12.002. URL `http://arxiv.org/abs/1104.0720`. arXiv: 1104.0720.

[Turing(1952)] 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. ISSN 0080-4622. URL `http://www.jstor.org/stable/92463`.

[Vecchiarelli et al.(2016)Vecchiarelli, Li, Mizuuchi, Hwang, Seol, Neuman, and Mizuuchi] A. G. Vecchiarelli, M. Li, M. Mizuuchi, L. C. Hwang, Y. Seol, K. C. Neuman, and K. Mizuuchi. Membrane-bound MinDE complex acts as a toggle switch that drives Min oscillation coupled to cytoplasmic depletion of MinD. *Proceedings of the National Academy of Sciences of the United States of America*, 113(11):E1479–E1488, Mar. 2016. ISSN 0027-8424. doi: 10.1073/pnas.1600644113. URL `http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4801307/`.

[Vázquez and Galaktionov(2002)] J.-L. Vázquez and V. Galaktionov. The problem Of blow-up in nonlinear parabolic equations. *Discrete and Continuous Dynamical Systems*, 8(2):399–433, Jan. 2002. ISSN 1078-0947. doi: 10.3934/dcds.2002.8.399. URL `http://www.aimsciences.org/journals/displayArticles.jsp?paperID=297`.

[Walsh(1986)] J. Walsh. An Introduction to Stochastic Partial-Differential Equations. *Lecture Notes in Mathematics*, 1180:265–437, 1986. ISSN 0075-8434. WOS:A1986D006200003.

[Wolfram|Alpha(2018a)] Wolfram|Alpha. Wolfram|Alpha: Calculating tau of phi, Oct. 2018a. URL `https://www.wolframalpha.com/input/?i=int+(x%5E4%2BC)%5E-0.5+dx`.

[Wolfram|Alpha(2018b)] Wolfram|Alpha. Wolfram|Alpha: Infinite Burst Time, Oct. 2018b. URL `https://www.wolframalpha.com/input/?i=int_0%5Einf+(x%5E4%2BC)%5E-0.5+dx`.