Robert C. Dalang
Davar Khoshnevisan
Carl Mueller
David Nualart
Yimin Xiao

A Minicourse on Stochastic Partial Differential Equations

Salt Lake City, Utah, 2006
Editors: D. Khoshnevisan and F. Rassoul-Agha

Springer
Berlin Heidelberg New York
Hong Kong London
Milan Paris Tokyo

Contents

Some Tools and Results for Parabolic Stochastic Partial Differential Equations

Carl Mueller*

Department of Mathematics, University of Rochester,
Rochester, NY 14627,
Email: cmlr@math.rochester.edu
URL: http://www.math.rochester.edu/~cmlr

Summary. These notes give an informal introduction to parabolic stochastic partial differential equations. We emphasize material coming from particle systems, including duality and the Dawson–Watanabe superprocess. We also deal with large deviations and comparison theorems. Applications include blow-up, hitting theorems, and compact support.

1 Introduction

1.1 Outline of the notes

The goal of these notes is to explain a set of ideas in stochastic partial differential equations (SPDE) which I have found useful. The notes are meant for graduate students, so they are not written in a formal style. Sometimes I will explain an idea in a simple case, and leave it to the reader to develop the topic more broadly.

I will begin with some general thoughts on the field. SPDE, and perhaps PDE as well, find their primary motivations in science and engineering. It is best not to think of SPDE as objects in pure mathematics, but as models for physical phenomena. To study SPDE it is not necessary to have a Ph.D. in biology, for example, but it is often helpful to think of an SPDE in terms of a population of organisms. Similar helpful motivations come from other sciences as well. When thinking of the heat equation, with or without noise, it helps to visualize a physical object with varying temperature. It would be foolish to set forth a foundation for the entire field of SPDE, but my goal is to explain some tools which others may find useful.

Both ordinary differential equations (ODE) and partial differential equations (PDE) play a fundamental role in describing reality. However, any model

* Research supported in part by grants from the National Science Foundation and National Security Agency.

of the real world must take into account uncertainty or random fluctuations. It is therefore surprising that while stochastic ODE were studied intensively throughout the twentieth century, SPDE only received attention much later. Some early work stemmed from the Zakai equation in filtering theory, see [32], and on the theoretical side there was the work of Pardoux [31] and Krylov and Rozovskii [17]. Much of this early work centered on foundational questions such as setting up the appropriate function spaces for studying solutions, or using such analytic tools as the method of monotonicity [31]. Later, Walsh [38] introduced the notion of martingale measures as an alternative framework. The diverse origins of SPDE have led to a lively interplay of viewpoints. Some people feel that SPDE should be based on such tools as Sobolev spaces, as is the case for PDE. Others, with a background in probability, feel that an SPDE describes a special kind of stochastic process. Applied mathematicians may feel that the study of SPDE should follow the ideas used in their domain.

By a historical accident, particle systems, which may be considered as discrete SPDE, were studied much earlier than SPDE. Such pioneers as Ted Harris and Frank Spitzer laid the groundwork for this theory. Their research was also influenced by results in percolation, by such mathematicians as Harry Kesten. Particle systems has changed its emphasis over the years, and some of this early work is being forgotten. However, I believe that the main methods of particle systems will always be relevant to SPDE. In particular, I will introduce the method of duality. Unfortunately, there was no time in the course to discuss percolation, which I also believe has fundamental importance for SPDE. Both duality and percolation, as well as many other ideas, are described in more detail in three classics: [21; 22] give detailed technical accounts of the field, and [9] provides a lively intuitive treatment.

Secondly, Watanabe and Dawson found that the scaling limit of critical branching Brownian motions give a fundamentally important model, called the Dawson–Watanabe process or superprocess. Because this model involves independently moving particles, there are powerful tools for studying its behavior, and many of these tools help in the study of SPDE. For example, the heat equation can be thought of as the density of a cloud of Brownian particles. Any SPDE which involves a density of particles can be studied via the Dawson–Watanabe process. There is a huge literature in this area, but two useful surveys are written by Dawson [6] and Perkins [33].

Thirdly, as one might expect, tools from PDE are useful for SPDE. Of course, Sobolev spaces and Hilbert spaces play a role, as in the work of Da Prato and Zabczyk [8] and Krylov [19]. But here I wish to concentrate on qualitative tools, such as the maximum principle. In particular, comparison theorems hold for many SPDE. Given two solutions, suppose that one is initially larger than the other. Then that relationship will continue to hold for later times.

Finally, tools from probability also find applications in SPDE. For example, the theory of large deviations of dynamical systems developed by Wentzell and Freidlin [13] also applies to SPDE. If the noise is small, we can estimate the probability that the solutions of the SPDE and corresponding PDE (without noise) differ by more than a given amount. Unfortunately, I had no time to discuss questions of coupling and invariant measures, which play a large role in the study of the stochastic Navier–Stokes equation.

After introducing these ideas, I give some applications to properties of parabolic SPDE.

1.2 Some comments on the current state of SPDEs

To finish this introduction, let me indicate some of the current directions in SPDE. The subject is expanding in many directions, so all I can do is give a personal view.

I believe that the foundations of SPDE are settled. One can use either Walsh's approach [38] using martingale measures, the Hilbert space approach of Da Prato and Zabczyk [8], or Krylov's L_p theory [19]. The goal should be to study properties of the equations.

Furthermore, the "naive" study of qualitative properties of SPDE, involving the superprocess or simple variants of the heat equation, is also largely finished. I do recommend learning the main ideas of this theory, as well as the foundational approaches mentioned earlier, since they can help us study of more complicated equations.

Luckily, finally, scientists are jumping into the field with a vengeance, and SPDE is expanding chaotically in all directions. I believe that the sciences will continue to provide important SPDE models and conjectures. For example, the reader can consult [7] for polymer models, and [3] for SPDEs in the physics of surfaces. There is broad interest in the stochastic Navier–Stokes equation [23]. Scientists seem to have finally grasped the importance of SPDE models, so the reader should stay alert for new developments.

2 Basic framework

Expanding on Chapter 1, Section 6 of this book, we will mostly consider parabolic equations of the form

$$\begin{aligned} \frac{\partial u}{\partial t} &= \Delta u + a(u) + b(u)\,\dot{W}(t\,,x), \\ u(0\,,x) &= u_0(x). \end{aligned} \tag{1}$$

where $\dot{W}(t\,,x)$ is $d+1$ parameter white noise and $x \in \mathbf{R}^d$. This set-up has been used for a long time, see [38] or [8]. If we expect our noise to arise from many small influences which are independent at different positions in space and time, then white noise is a good model. As with stochastic differential equations (SDE), the solution $u(t\,,x)$ is not differentiable in t or x, so (1) does not make sense as written. In Da Prato and Zabczyk's approach, we regard (1) as a stochastic differential equation (SDE) in function space. In Walsh's theory, which we will adopt here, we regard (1) as a shorthand for the following integral equation, which is often called the mild form of the equation. This is also explained in Chapter 1, Section 6. Of course, Da Prato and Zabczyk's theory can deal with the mild form as well.

$$u(t\,,x) = \int_{\mathbf{R}^d} G(t\,,x-y)u_0(y)\,dy + \int_0^t \int_{\mathbf{R}^d} G(t-s\,,x-y)a(u(s\,,y))\,dy\,ds \\ + \int_0^t \int_{\mathbf{R}^d} G(t-s\,,x-y)b(u(s\,,y))\,W(dy\,ds). \tag{2}$$

Here $G(t\,,x)$ is the heat kernel,

$$G(t\,,x) = \frac{1}{(4\pi t)^{d/2}} \exp\left(-\frac{|x|^2}{4t}\right), \tag{3}$$

and the final integral in (2) is a stochastic integral in the sense of Walsh [38], see Chapter 1 of this book. Since there is a unique time direction t, such integrals can be constructed along the lines of Ito's theory, and their properties are mostly the same as for the Ito integral. Let $\mathscr{F}_t$ denote the σ-field generated by the noise W up to time t. That is, $\mathscr{F}_t$ is generated by the integral

$$\int_0^t \int_{\mathbf{R}^d} g(s\,,y)\,W(dy\,ds), \tag{4}$$

for deterministic function $g \in \mathbf{L}^2(dyds)$. For convenience, when $f(s,y)$ is nonanticipating with respect to $\mathscr{F}_t$ and

$$E\int_0^t \int_{\mathbf{R}^d} G^2(t-s\,,x-y)f^2(s\,,y)\,dy\,ds < \infty, \tag{5}$$

we often write

$$N(t\,,x) = \int_0^t \int_{\mathbf{R}^d} G(t-s\,,x-y)f(s\,,y)\,W(dy\,ds), \tag{6}$$

so that if $f(s\,,y) = b(u(s\,,y))$, then $N(t\,,x)$ is the "noise term" in (2). Then, in particular,

$$E\left[N(t\,,x)^2\right] = \int_0^t \int_{\mathbf{R}^d} E\left[G^2(t-s\,,x-y)f(s\,,y)^2\right]\,dy\,ds. \tag{7}$$

Exercise. For $f(s\,,y) :=$, show that for any $t > 0$ and $x \in \mathbf{R}^d$,

$$E\left[N(t\,,x)^2\right] \qquad \begin{cases} < \infty & \text{if } d = 1, \\ = \infty & \text{if } d > 1. \end{cases} \tag{8}$$

Thus, if $d \geq 2$, (1) is likely to have solutions which are generalized functions, but do not exist as functions in $\mathbf{L}^2(dP)$, for instance. But then nonlinear functions such as $a(u), b(u)$ are hard to define. Since we live in a 3-dimensional world, this situation gives rise to many difficulties. One common solutions is to smooth the noise, perhaps by convolving it with another function of the variable x. Such noise is called colored noise. Another possibility is to replace $\mathbf{R}^d$ by a lattice, and replace the Laplacian by the discrete Laplacian. It would be nice to deal with (1) more directly; maybe you have an idea of how to do this.

Actually, some authors deal with solutions to (1) which are generalized functions over Wiener space. Usually, generalized functions are defined in terms of integrals against a test function. Thus, we would have to define our solution in terms of an integral over all points in the probability space. But in the physical world, we are doomed to experience only a single point in the probability space. Maybe I'm being too pessimistic here, since repeated experiments sample different points in the probability space; readers can form their own conclusion.

Another point is that for x fixed, the process $t \to u(t\,,x)$ is Hölder continuous with parameter $1/4 - \varepsilon$ for every $\varepsilon > 0$, and therefore is not a semimartingale. Therefore there is no Ito's lemma in the usual sense, and this has caused a lot of problems for the theory. However, if $\phi(x)$ is an $\mathbf{L}^2$ function, then

$$X_t = \int_{\mathbf{R}^d} u(t\,,x)\phi(x)\,dx \tag{9}$$

is a semimartingale with quadratic variation

$$\langle X \rangle_t = \int_0^t \int_{\mathbf{R}^d} \left[\int_{\mathbf{R}^d} \phi(x) G(t-s\,,x-y)\,dx \right]^2 b(u(s\,,y))^2 dy\,ds. \tag{10}$$

Note that the inner integral smoothes out the singularity of $G(t-s\,,x-y)$.

Since it doesn't fit in elsewhere in this paper, let me mention a very nice survey by Ferrante and Sanz-Solé [12] which deals with SPDE driven by colored noise. For a colored noise $\dot{F}(t\,,x)$, the covariance in x is not the δ-function as in the case of white noise, so heuristically $E[\dot{F}(t,x)\dot{F}(s,y)] = \delta(t-s)R(x-y)$ for some covariance function R.

Finally two recent books, Chow [4] and Rockner [34], also give nice introductions to SPDE. Both deal with the functional analytic approach similar to Da Prato and Zabczyk [8].

3 Duality

3.1 Definitions

The reader might wonder why I am devoting so much space to duality, since there are only a few papers on SPDE with that word in the title. I firmly believe that many SPDE should be viewed as limits of particle systems, and duality is perhaps the leading method in particle systems. Secondly, the most important tool in superprocesses is the Laplace functional, which is a form of duality. Thirdly, duality is getting relatively little attention, but it remains a powerful tool.

Duality is a relationship between two stochastic processes which allows us to use information from one process in analyzing the other. There are at least two kinds of duality. One involves a duality function, and the other arises from studying the ancestry of particles in a particle system.

I will describe the functional form of duality first; details can be found in [10], pages 188–189. Two processes X_t, Y_t are said to be in duality with respect to a function $H(x\,,y)$ if, for all t in some interval $[0\,,T]$ we have

$$E\Big[H\left(X_t\,,Y_0\right)\Big] = E\Big[H\left(X_0\,,Y_t\right)\Big]. \tag{11}$$

All probabilists know simple examples of duality, although they may not realize it. Let $u(t\,,x)$ satisfy the heat equation

$$\frac{\partial u}{\partial t} = \frac{1}{2}\Delta u, \tag{12}$$

and let B_t be Brownian motion. It is well known that under the appropriate conditions,

$$u(t\,,x) = E_x\Big[u(0\,,B_t)\Big]. \tag{13}$$

Thus, the processes $u(t\,,\cdot)$ and B_t are in duality with respect to the function $H(u\,,B) = u(B)$. Note that deterministic processes such as $u(t\,,x)$ still count as processes.

3.2 Duality for Feller's diffusion

Feller derived the following SDE as a limit for critical birth and death processes; a critical birth and death process has expected offspring size equal to 1, so that the expected number of particles does not change:

$$dX = \sqrt{X}\, dB. \tag{14}$$

Now consider a function $v(t)$ which satisfies

$$v' = -\frac{1}{2}v^2. \tag{15}$$

Explicit solutions $v(t)$ are easy to find. Next, Ito's lemma implies that for $0 < t < T$, if

$$M_t = \exp\left(-X_t v(T-t)\right), \tag{16}$$

then

$$\begin{aligned} dM &= M \cdot (-\sqrt{X}\, dB) + \frac{1}{2} M \cdot X v^2\, dt - M \cdot \frac{1}{2} X v^2\, dt \\ &= M \cdot (-\sqrt{X}\, dB). \end{aligned} \tag{17}$$

Thus M_t is a martingale, and

$$\exp\left(-X_0 v(T)\right) = M_0 = E[M_T] = E\left[\exp\left(-X_T v(0)\right)\right]. \tag{18}$$

So, X_t and $v(t)$ are in duality with respect to the function $H(x\,, v) = \exp(-xv)$. In this case, duality gives us the Laplace transform of X_t. Duality implies that X_t is unique in law, provided X_0 is specified.

Exercise. Explicitly solve for $v(t)$ and hence find the Laplace transform of X_t.

3.3 Duality for the Wright Fisher SDE

Next, let X_t satisfy the following SDE, named after Fisher and Wright. We can think of a population of constant size, consisting of 2 subpopulations whose percentages of the total are X and $1-X$. Due to competition, the population fluctuates randomly, and the variance of the fluctuations are proportional to the number of encounters between the two types. This leads to a standard deviation of $\sqrt{X(1-X)}$, and gives rise to the following SDE.

$$\begin{aligned} dX &= \sqrt{X(1-X)}\, dB \\ X_0 &= x_0 \in [0\,,1]. \end{aligned}$$

For n a nonnegative integer, Ito's formula gives

$$\begin{aligned} dX^n &= nX^{n-1}\sqrt{X(1-X)}\, dB + \frac{n(n-1)}{2} X^{n-2} X(1-X)\, dt \\ &= \text{mart} - \binom{n}{2}(X^n - X^{n-1})\, dt, \end{aligned} \tag{19}$$

where "mart" indicates a martingale term whose expectation is 0 when integrated. The final term in (19) has the intuitive meaning that X^n is replaced by X^{n-1} at rate $\binom{n}{2}$, except that there is a negative sign.

Let N_t be a process taking values in the nonnegative integers, and independent of X_t. We think of N_t as the number of particles present at time t, and let each pair of particles coalesce at rate 1. In other words, N_t is a Markov process such that as $h \to 0$,

$$\begin{aligned} P(N_{t+h} = N_t \mid N_t) &= 1 - \binom{N_t}{2} h + o(h) \\ P(N_{t+h} = N_t - 1 \mid N_t) &= \binom{N_t}{2} h + o(h). \end{aligned} \tag{20}$$

Thus, for $x \in [0\,,1]$, one has

$$
\begin{aligned}
E&\left[x^{N_{t+h}}-x^{N_t}\,\middle|\,N_t\right] \\
&=x^{N_t}\left(-\binom{N_t}{2}h+o(h)\right)+x^{N_t-1}\left(\binom{N_t}{2}h+o(h)\right) \qquad (21)\\
&=-\binom{N_t}{2}\left(x^{N_t}-x^{N_t-1}\right)h+o(h).
\end{aligned}
$$

Exercise. Prove (21).

Note that the last lines in (19) and (21) match. Then X_t, N_t are in duality with respect to the function $H(x\,,n)=x^n$. In other words,

$$
E\left[X_t^{N_0}\right]=E\left[X_0^{N_t}\right]. \qquad (22)
$$

This allows us to compute the moments of X_t in terms of X_0 and N_t.
Exercise. Prove (22).

We can regard duality relations of this kind as a means of calculating moment. Note that since N_t is a nonincreasing process, (22) implies that the n^{th} moment of X_t only depends on X_0^k for $k\le n$. Physicists and others often compute moments by finding systems of differential equations which they solve recursively. These equations are called closed if the derivative of the n^{th} moment only depends on k^{th} moments for $k\le n$.
Exercise. Show that in the above example, the moment equations are closed. Modify this example so that the moment equations are not closed.

If the moment equations are not closed, we may still have a duality relationship, but the process N_t may not be nonincreasing. Thus, even if a system of moment equations is not closed, we may still be able to find a dual process and draw certain conclusions. One may view duality arguments as an attractive packaging for the idea of moment equations.

One important use of duality is the study of uniqueness. Suppose that processes X_t, Y_t are in duality with respect to the function $H(x\,,y)$. This relationship is often enough to show that X_t is unique in law, at least if $X_0=x_0$ is specified. Indeed, let $X_t, \tilde{X}_t$ be two processes with the same initial value x_0, and assume that both $X_t, \tilde{X}_t$ are in dual to Y_t with respect to $H(x\,,y)$. Then

$$
\begin{aligned}
E\Big[H(\tilde{X}_t\,,Y_0)\Big]&=E\Big[H(\tilde{X}_0\,,Y_t)\Big]\\
&=E\Big[H(X_0\,,Y_t)\Big] \qquad (23)\\
&=E\Big[H(X_t\,,Y_0)\Big].
\end{aligned}
$$

If (23) is true for many initial values Y_0, we can often conclude that X_t and $\tilde{X}_t$ are equal in law. Using the Markov property (if X_t is Markov) and repeating this procedure gives us the uniqueness of the finite dimensional distributions. Then the Kolmogorov extension theorem shows that X_t is unique in law.

3.4 The voter model and the contact process

The voter model and contact process were some of the first models studied in particle systems. Although interest has largely shifted to other models, these processes give a useful introduction to duality. For more details, see [9] or [22; 21].

First we describe the voter model. In keeping with the style of these notes, our description is informal. The state space is the set of functions $F : \mathbf{Z}^d \to \{0\,,1\}$, with d fixed. We imagine that there is a voter at each site of $\mathbf{Z}^d$, whose opinion is either 0 (Republican) or 1 (Democrat). Let $\mathbf{E}^d$ denote the set of directed edges connecting nearest neighbors of $\mathbf{Z}^d$. In other words, each nearest neighbor pair $p, q \in \mathbf{Z}^d$ is associated with two edges, which either point from p to q or vice versa. If the edge e points from p to q, we set $e_0 = p$ and $e_1 = q$. To each edge $e \in \mathbf{E}^d$ we associate an independent rate-one Poisson process N^e_t. At the times τ_e of the Poisson process N^e_t, we change the opinion of e_1 to equal the opinion at e_0, that is, we redefine $F_{\tau_e}(e_1)$ to equal $F_{\tau_e}(e_0)$. Another common notation for the voter model involves the set ξ_t of sites $p \in \mathbf{Z}^d$ where voters have opinion 1, i.e., $F_t(p) = 1$. One needs to show that this procedure gives a consistent definition of a stochastic process, and this is done in [22] and other places.

The dual process of the voter model is given in terms of the ancestry of opinions. Fix a site $p \in \mathbf{Z}^d$ and a time $T > 0$. We wish to define a process X^p_t, which traces the history of the opinion found at position p at time T. We will let time t run backwards in the definition of X^p_t, so that t measures the amount of time before T. To be specific, let τ_e be the most recent Poisson time involving an edge e with $e_1 = p$. For $0 \le t < T - \tau_e$ let $X^p_t = p$, and let $X_{T-\tau_e} = e_0$. Now we repeat the construction for the site e_0 and time τ_e, until $t = T$.

Considering our definition of the voter model in terms of Poisson events, we see that X^p_t is a continuous-time nearest neighbor simple random walk on $\mathbf{Z}^d$, with transition rate $2d$. In fact, $\{X^p_t : p \in \mathbf{Z}^d\}$ is a collection of independent coalescing random walks. That is, they evolve independently, but each pair of random walks merge as soon as they land on the same site. One can also view this process in terms of the directed edges. X^p_t starts at p, and moves backwards through time starting at time T. Whenever it encounters an edge e, it moves from e_1 to e_0 if it finds itself at e_1. So, it moves along the edges in the reverse orientation.

Exercise. Convince yourself that our description of the dual for the voter model is correct.

This form of duality can also be expressed in terms of a function $H(A\,,\xi)$. Here, ξ_t is the set of sites with opinion 1 at time 0, and A_t is the set of sites where particles of the coalescing system of random walks are found at time t.

Exercise. What is the duality function H which matches our previous description? The answer is in [9, p. 23].

A beautiful application of duality for the voter model is the study of clustering. Clustering means that for a set of sites $S \subset \mathbf{Z}^d$, if time t is large, then with high probability all of the sites have the same opinion. This will certainly happen if opinions at the sites in S all come from a common ancestor. For simplicity, consider the case of 2 sites, $S = \{x\,, y\}$. Certainly $\xi_t(x) = \xi_t(y)$ if in the dual process, the random walks starting at x and y have coalesced by time t. But in dimensions 1 or 2, this coalescence occurs with probability 1. The probability is less than one in higher dimensions, and it is not hard to show that clustering does not occur in higher dimensions. We leave these details to the reader, who can also consult [9].

3.5 The biased voter model

We might modify the voter model by assuming that one opinion is stronger than another. Using the same notation as for the unbiased voter model, for each directed edge e we construct 2 Poisson processes $N_t^{e,1}, N_t^{e,2}$ with rates λ_1, λ_2 respectively. Recall that the edge e points from the e_0 to e_1. Suppose that at time t there is an event of the Poisson process $N_t^{e,1}$. Then, at time t the point e_1 takes on the opinion at e_0. Secondly, suppose that at time t there is an event of the Poisson process $N_t^{e,2}$. In this case, if e_0 has opinion 1, then e_1 changes its opinion to 1 as well. On the other hand, if e_0 has opinion 0 at time t, then nothing happens. Thus, the opinion 1 is stronger than opinion 0. As of today (January 2006) we could say that 1 means the voter is a democrat.
Exercise. Verify that the ancestry of opinions is as described below, and construct an appropriate dual process.

The path of ancestry of the opinion at position x at time t should go backward in time, and if for an edge e pointing toward x the process $N_t^{e,1}$ has an event, it should follow that edge in the reverse direction to a new site. On the other hand, if $N_t^{e,2}$ has an event, then the ancestral path should split, with one branch staying at the same point, and another branch following the edge e backwards. If at time 0 the cloud of ancestral particles meets a 1, then the original site has opinion 1. Otherwise it has opinion 0. The various branches of the ancestry are meant to sample all possible occurrences of 1.

3.6 The contact process

In the contact process, there are particles at various sites in $\mathbf{Z}^d$. Particles die at exponential rate 1. At rate $\lambda > 0$, they give birth. When a particle at site x gives birth, the new particle chooses a site y at random among the nearest neighbors of x. If site y is already occupied, then the birth is aborted. Otherwise, the new particle moves to position y.
Exercise. Using the preceding ideas, verify that tracing the ancestry of particles gives us another contact process, and hence the contact process is self-dual.

3.7 The Dawson–Watanabe superprocess

The Dawson–Watanabe superprocess, which we will simply call the superprocess, arose as a limit in the theory of spatial branching processes, more specifically in population biology. It is one of the few nonlinear SPDE with solutions in the space of generalized functions or Schwartz distributions. There are several good introductions to superprocesses; [20] is a classic of clear exposition. See also Etheridge's book [11] and the surveys of Dawson [6] and Perkins [33], to name a few. Since there are so many good sources, we will not systematically develop the theory of superprocesses, but rather describe the applications to SPDE.

Here is the intuition. Let μ be a given finite nonnegative measure on $\mathbf{R}^d$; the finiteness condition can be weakened. Fix a natural number m, and let $\{B_t^{(i)}\}_{i=1}^{N(t)}$ be a collection of critical branching Brownian motions taking values in $\mathbf{R}^d$. We assume that the Brownian particles are independent. $N(t)$ is the number of particles existing at time t, and critical branching means that each particle splits in two or dies with equal probability. We assume that the times between branching are independently distributed exponential variables with mean $1/m$. We define a measure-valued process by putting a delta function at the location of each particle, and then dividing by m:

$$X_t^{(m)}(A) = \frac{1}{m}\sum_{i=1}^{N(t)} \delta_{B_t^{(i)}}(A). \tag{24}$$

Here δ_x is the delta measure centered at x. Suppose that $X_0^{(m)}$ converges weakly to μ as $m \to \infty$. The main existence theorem for superprocesses asserts that in the appropriate topology, $X_t^{(m)}$ converges weakly to a limiting process X_t.

The limiting superprocess X_t has many fascinating properties. For example, in $\mathbf{R}^d$ with $d \geq 2$, with probability one X_t is a measure whose support has Hausdorff dimension 2. However, in $\mathbf{R}^2$ the measure X_t is singular with respect to Lebesgue measure. In fact, Perkins has determined the exact Hausdorff measure function, and we can loosely say that the Hausdorff dimension of the support is infinitesimally less than 2, meaning that the exact Hausdorff measure function is x^2 with extra logarithmic terms. These properties and more can be found in [6] and [33].

One important tool is the martingale problem formulation. For appropriate functions φ on $\mathbf{R}^d$, we denote

$$Z_t(\varphi) = X_t(\varphi) - \int_0^t \frac{1}{2} X_s(\Delta\varphi)\, ds. \tag{25}$$

Then $Z_t(\varphi)$ is a continuous martingale with quadratic variation

$$\langle Z \rangle_t = \int_0^t X_s(\varphi^2)\, ds. \tag{26}$$

If ν is a measure, we write $\nu(\varphi) = \int \varphi(x)\,\nu(dx)$. The martingale problem allows us to use Ito calculus. Indeed, since $Z_t(\varphi)$ is a continuous martingale with quadratic variation given by (26), we can use Ito calculus for $Z_t(\varphi)$. But (25) gives $X_t(\varphi)$ in terms of $Z_t(\varphi)$.

Even more important than the martingale problem is the Laplace functional. It is an expansion of the duality relation for Feller's diffusion explained in Subsection 3.2. Suppose that $v(t\,,x)$ satisfies

$$\frac{\partial v}{\partial t} = \frac{1}{2}\Delta v - \frac{1}{2}v^2 \tag{27}$$

We can solve (27) for a wide variety of initial functions $v(0,x)$, but suppose $v(0,x)$ is bounded and nonnegative, say $0 \le v(0,x) \le M$. Then we can replace (27) by

$$\frac{\partial v}{\partial t} = \frac{1}{2}\Delta v - \frac{M}{2}v \tag{28}$$

which is a linear equation and can be solved by standard PDE theory.
Exercise. Show that if $0 \le v(0,x) \le M$, then a solution of (28) also solves (27).

Under the appropriate conditions on $v(0,x)$ and X_0, for $0 \le t \le T$,

$$M_t = \exp\left(-X_t(v(T-t\,,\cdot))\right) \tag{29}$$

is a martingale. For example, we could assume that $v(0,x)$ is nonnegative and bounded.
Exercise. Under the assumption that $v(0,x)$ is nonnegative and bounded, show that M_t is a martingale.

Therefore

$$E\Big[\exp\left(-X_T(v(0\,,\cdot))\right)\Big] = \exp\left(-X_0(v(T\,,\cdot))\right). \tag{30}$$

In other words, X_t and $v(t\,,\cdot)$ are in duality with respect to the function

$$H(X\,,v) = \exp\left(-X(v)\right). = \exp\left(-\int v(x)\,X(dx)\right). \tag{31}$$

Exercise. Verify the above duality relation.

This duality relation gives very useful information about the Laplace transform of X_t, and also proves uniqueness in law for the superprocess.

For $x \in \mathbf{R}^1$, the superprocess has a density $X_t(dx) = u(t\,,x)dx$ which satisfies

$$\frac{\partial u}{\partial t} = \frac{1}{2}\Delta u + \sqrt{u}\,\dot{W}(t\,,x). \tag{32}$$

Thus, we have uniqueness in law for this equation. Almost sure uniqueness is an unsolved problem which has attracted the attention of many of the best probabilists, and I have heard at least two announcements of false proofs. The lack of Ito's lemma hurts us here.

3.8 Branching Brownian motion and a population equation

Consider a population with two types of genes, and let $u(t\,,x)$ be the population density of one type of individual at time t at position x. We assume that individuals perform independent Brownian motions, so the population density might be modeled by the heat equation. Due to mating between individuals, there might be a random contribution to the population density. The variance of this random contribution should be proportional to the product of the two population densities, namely $u(1-u)$. Therefore, its standard deviation should be $\sqrt{u(1-u)}$. This leads us to the following model on $t \geq 0$, $x \in \mathbf{R}$.

$$\begin{aligned} \frac{\partial u}{\partial t} &= \Delta u + \sqrt{u(1-u)}\,\dot{W}(t\,,x) \\ u(0\,,x) &= u_0(x), \end{aligned} \tag{33}$$

where $0 \leq u_0(x) \leq 1$. If there were no dependence on x, this equation would be identical to (19). Using the duality we derived for (19), it is not hard to guess that $u(t\,,x)$ will be dual to a system of Brownian motions $\{B_i(t)\}_{i=1}^{N(t)}$, where each pair of particles $B_i(t), B_j(t)$ coalesce at exponential rate 1, measured with respect to the local time at 0 for the process $B_i(t) - B_j(t)$. To be specific, let τ be an independent exponential variable with parameter 1, and let $\ell_{i,j}(t)$ be the local time at 0 of $B_i(t) - B_j(t)$. If there were no other particles, then the particles $B_i(t), B_j(t)$ would coalesce at time t for which $\ell_{i,j}(t) = \tau$. One has the duality relation

$$H\left(u\,, \{b_i\}_{i=1}^{N}\right) = \prod_{i=1}^{N} \left(1 - u(b_i)\right), \tag{34}$$

so that

$$E\left[\prod_{i=1}^{N(0)} \Big(1 - u(t\,, B_i(0))\Big)\right] = E\left[\prod_{i=1}^{N(t)} \Big(1 - u(0\,, B_i(t))\Big)\right]. \tag{35}$$

See [37] for details. Since we can choose $\{B_i(0)\}_{i=1}^{N(0)}$, this gives us a formula for the moments of $u(t\,,x)$. Notice that the moment equations are closed, since there cannot be more particles at time t than at time 0.

Exercise. Compute the moments of u.

Also, if $u(0\,,x) \approx 1$, then the right side of (35) is close to 0 if there are any particles near x. This gives us a way of relating the size of $u(t\,,x)$ to the probabilities of the Brownian particles.

Among other things, this duality relation gives uniqueness in law for (36). In [27], this duality was used to study the width $D(t)$ of the interface where $0 < u(t\,,x) < 1$, assuming that $u(0\,,x) = \mathbf{1}_{(-\infty,0]}(x)$. This interface was proved to have stochastically finite width, that is

$$\lim_{\lambda\to\infty}\sup_{t\geq 0} P(D(t)>\lambda)=0.$$

At about the same time, and independently, Cox and Durrett [5] proved a similar result for the long-range voter model. They also used duality.

3.9 Branching coalescing Brownian motion and the KPP equation with noise

The KPP equation is one of the simplest equations exhibiting traveling wave solutions.

$$\begin{aligned}\frac{\partial u}{\partial t}&=\Delta u+u(1-u),\\ u(0\,,x)&=u_0(x).\end{aligned} \tag{36}$$

Often, one takes $u_0(x)=\mathbf{1}_{(-\infty,0]}(x)$. One can prove that there is a function $h(x)$ with $\lim_{x\to\infty}h(x)=0$ and $\lim_{x\to-\infty}h(x)=1$, and a function $v(t)$ satisfying $\lim_{t\to\infty}v(t)/t=2$ for which

$$\lim_{t\to\infty}\sup_{x\in\mathbf{R}}|u(t\,,x)-h(x-v(t))|=0. \tag{37}$$

Detailed properties of this equation have been derived by Bramson [1; 2] using the Feynman-Kac formula. Suppose that $\{B_i(t)\}_{i=1}^{\infty}$ is a collection of Brownian motions. At exponential rate 1, each particle splits in two. Then $u(t\,,x)$ and $\{B_i(t)\}_{i=1}^{\infty}$ are in duality, and the duality function $H(\cdot\,,\cdot)$ is the same as in (34). As a thought exercise, the reader may wish to verify that the cloud of particles spread at about the same rate as $u(t\,,x)$. Hint: fix $x_0>0$, and let T be the first time t that $u(t\,,x)=1/2$. Start a single particle at position x_0 at time 0, and write down the duality equation.

If we consider the population model in the previous section, but suppose that one type kills the other, there might be a drift term proportional to the frequency of interactions, which is proportional to $u(1-u)$. This would give us the equation

$$\begin{aligned}\frac{\partial u}{\partial t}&=\Delta u+u(1-u)+\sqrt{u(1-u)}\,\dot{W}(t\,,x),\\ u(0\,,x)&=u_0(x).\end{aligned} \tag{38}$$

Combining the two previous types of duality, we might conjecture that $u(t\,,x)$ is dual to a system of Brownian particles, in which pairs of particles $B_i(t), B_j(t)$ coalesce at rate 1 according to the local time where $B_i(t)-B_j(t)=0$, and each particle splits in two at an exponential rate with parameter 1.

This kind of duality is not easy to work with, but it does prove uniqueness for (38). In [26], the traveling wave behavior for (38) was studied.

4 Large deviations for SPDEs

Roughly speaking, large deviations measures how far the solution of an SPDE can get from the solution of the corresponding PDE. Either noise or time is taken to be small. From another point of view, we might wish to see how large the solution of an SPDE can be. If we know how large the solution of the corresponding PDE is, then large deviations can give an upper bound for the SPDE.

There are many excellent books on large deviations; our goal here is to give an overview with emphasis on intuition. To start at the basic level, we give a tail estimate for a $N(0\,,1)$ random variable Z.

$$P(Z > \lambda) = \int_\lambda^\infty \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx. \tag{39}$$

Note that for $a > 0$ fixed,

$$\lim_{x\to\infty} \frac{\exp\left(-(x+a)^2/2\right)}{\exp\left(-x^2/2\right)} = 0. \tag{40}$$

So it seems reasonable that for large λ, the integral in (39) will be dominated by values of x very close to λ, since the other values of the integrand are very small in comparison. Then as $\lambda \to \infty$,

$$-\log P(Z > \lambda) \sim -\log\left(\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\lambda^2}{2}\right)\right) \sim \frac{\lambda^2}{2}, \tag{41}$$

in the sense that the ratio of the two sides tends to 1.
Exercise. Prove (41).

The large deviations philosophy states that as some parameter tends to 0 or ∞, the probability in question will be determined by a single point in the probability space, along with a small neighborhood around it. In our example, the point would be where $Z = \lambda$.

For SPDE, consider the following setup. Let

$$N(t\,,x) = \int_0^t \int_{\mathbf{R}} G(t-s\,,x-y) f(s\,,y)\, W(dy\,ds), \tag{42}$$

where $f(s\,,y)$ is a predictable random function, with the almost sure bound

$$\sup_{s,y} |f(s\,,y)| \le K. \tag{43}$$

Here is our large deviations theorem.

Theorem 4.1. *Let $M > 0$. There exist constants $C_1, C_2 > 0$ such that for all $T, K, \lambda > 0$,*

$$P\left(\sup_{0\le t\le T} \sup_{|x|\le M} |N(t\,,x)| > \lambda\right) \le C_1 \exp\left(-\frac{C_2\lambda^2}{T^{1/2}K^2}\right). \tag{44}$$

Proof (Theorem 4.1). The reader should go back to Chapter 1, Chapter 4.2 of this book to see the similarities between Theorem 4.1 and Kolmogorov's continuity theorem. Theorem 4.1 can be proved from the Garsia–Rodemich–Rumsey lemma [14], and the proof has much in common with similar ideas from Gaussian processes. We prefer to give a proof from first principles, which duplicates part of the proof of the Garsia–Rodemich–Rumsey lemma.

We need the fact that $N(t\,,x)$ is continuous with probability 1. Actually, this can be deduced from our proof below, but for simplicity, we refer the reader to [38].

Next, observe that by scaling we can assume that $K = 1$. By cutting up the x-axis into intervals of size 1 and adding the corresponding estimates, we can reduce to the case where the supremum over x is taken on the interval $[0\,,1]$.

Furthermore, G and W have the following scaling for $x \in \mathbf{R}$.

$$\begin{aligned} aG(a^2t, ax) &= G(t\,,x), \\ W(d(ax)\,,d(a^2t)) &\overset{\mathscr{D}}{=} a^{3/2}W(dx\,,dt). \end{aligned} \tag{45}$$

By taking $a = t^{-1/2}$ and using the above scaling and setting $t = T$, the reader can verify that we need only prove the theorem for $T = 1$.

To summarize, we must show that if

$$\sup_{t,x} |f(t\,,x)| \leq 1 \quad \text{almost surely,} \tag{46}$$

then there exist constants C_1, C_2 such that for all $\lambda > 0$ we have

$$P\left(\sup_{0\leq t\leq T,}\sup_{0\leq x\leq 1} |N(t\,,x)| > \lambda\right) \leq C_1 \exp\left(-C_2\lambda^2\right). \tag{47}$$

Recall that we have reduced to $0 \leq x \leq 1$ by chopping up the interval $[-M, M]$.

To prove (47), we need the following estimates.

Lemma 4.2. *There exists a constant C such that for all $0 < s < t < 1$ and $x, y \in [-1, 1]$ we have*

$$\begin{aligned} \int_0^t \int_{\mathbf{R}} \left[G_t(x-z) - G_t(y-z)\right]^2\, dz\, ds &\leq C|x-y|, \\ \int_s^t G_{t-r}(z)^2\, dz\, dr &\leq C|t-s|^{1/2}, \\ \int_0^s \left[G_{t-r}(z) - G_{s-r}(z)\right]^2\, dz\, dr &\leq C|t-s|^{1/2}. \end{aligned} \tag{48}$$

The proof of Lemma 4.2 is an exercise in calculus or perhaps real analysis.
Exercise. Verify Lemma 4.2.

The details can also be found in [38]. Observe that Lemma 4.2 has the following corollary.

Corollary 4.3. *Assume that* $|f(t\,,x)| \le 1$ *almost surely for all* t, x. *Then there exist constants* C_1, C_2 *such that for all* $0 < s < t < 1$, $x, y \in [-1\,,1]$, *and* $\lambda > 0$,

$$\begin{aligned} P\left(|N(t\,,x) - N(t\,,y)| > \lambda\right) &\le C_1 \exp\left(-\frac{C_2\lambda^2}{|x-y|}\right), \\ P\left(|N(t\,,x) - N(s\,,x)| > \lambda\right) &\le C_1 \exp\left(-\frac{C_2\lambda^2}{|t-s|^{1/2}}\right). \end{aligned} \tag{49}$$

Proof. We prove only the first assertion of Corollary 4.3, leaving the second assertion to the reader. Let

$$\bar{N}_t(s\,,x) = \int_0^s \int_{\mathbf{R}} G(t-r\,,x-y) f(r\,,y)\, W(dy\,dr), \tag{50}$$

and note that $\bar{N}_t(t, x) = N(t, x)$. In other words, we have frozen the variable t which occurs inside G, in order for the stochastic integral to be a martingale. Let

$$M_s = \bar{N}_t(s\,,x) - \bar{N}_t(s\,,y) \tag{51}$$

and note that $M_t = N(t, x) - N(t, y)$. Thus M_s is a continuous martingale and hence a time-changed Brownian motion, see [36]. By Lemma 4.2, we have

$$\langle M \rangle_t \le C|x-y|. \tag{52}$$

Readers should convince themselves that the time scale of the time changed Brownian motion is given by $\langle M \rangle_t$. In other words, there is a Brownian motion B_t such that $M_t = B_{\langle M \rangle_t}$. Hence by the reflection principle for Brownian motion,

$$\begin{aligned} P(N(t\,,x) - N(t\,,y) > \lambda) &\le P(B_{C|x-y|} > \lambda) \\ &\le C_1 \exp\left(-\frac{C_2\lambda^2}{|x-y|}\right). \end{aligned} \tag{53}$$

The assertion in Corollary 4.3 then follows by making a similar estimate for $P(-N(t\,,x) + N(t\,,y) > \lambda)$. □

Continuing with the proof of Theorem 4.1, we define the grid

$$\mathscr{G}_n = \left\{ \left(\frac{j}{2^{2n}}\,, \frac{k}{2^n} \right) :\ 0 \le j \le 2^{2n},\ 0 \le k \le 2^n \right\}. \tag{54}$$

We write

$$\left(t_j^{(n)}, x_k^{(n)} \right) = \left(\frac{j}{2^{2n}}\,, \frac{k}{2^n} \right). \tag{55}$$

Two points $(t_{j_i}^{(n)}, x_{k_i}^{(n)}) : i = 1, 2$ are nearest neighbors if either

1. $j_1 = j_2$ and $|k_1 - k_2| = 1$, or

2. $|j_1 - j_2| = 1$ and $k_1 = k_2$

Claim 1 *Let $p \in \mathscr{G}_n$ for some n. There exists a path $(0,0) = p_0, p_1, \ldots, p_N = p$ of points in $\mathscr{G}_n$ such that each pair p_{i-1}, p_i are nearest neighbors in some grid $\mathscr{G}_m, m \leq n$. Furthermore, at most 4 such pairs are nearest neighbors in any given grid $\mathscr{G}_m$.*

We will prove a one-dimensional version of the claim for $x \in [0,1]$, and leave the rest to the reader. Let $\mathscr{H}_n = \{k/2^n : k = 0, \ldots, 2^n\}$ denote the dyadic rational numbers of order n lying in $[0\,,1]$.

Claim 2 *Let $x \in \mathscr{H}_n$. There exists a path $0 = p_0, p_1, \ldots, p_N = x$ of points in $\mathscr{H}_n$ such that each pair p_{i-1}, p_i are nearest neighbors in some grid $\mathscr{H}_m, m \leq n$. Furthermore, at most one such pair consists of points which are nearest neighbors in any given grid $\mathscr{H}_m$.*

Proof (Claim 2). Let

$$x = 0.x_1 x_2 \cdots x_n \tag{56}$$

denote the binary expansion of x, that is, its expansion in base 2, and let

$$p_m = 0.x_1 x_2 \cdots x_m. \tag{57}$$

Then Claim 2 follows. Claim 1 is proved using a similar argument, where we write $p = (t, x)$, take the binary expansion of x, and the base 4 expansion of t. $\square$

Exercise. Prove Claim 1.

Next, let $K, \alpha > 0$, and let $A(n\,,\lambda)$ be the event that for all nearest neighbors $p, q \in \mathscr{G}_n$ we have

$$|N(p) - N(q)| \leq \lambda K 2^{-(2-\alpha)n}. \tag{58}$$

By Corollary 4.3, for each pair of nearest neighbors $p, q \in \mathscr{G}_n$, we have

$$P\left(|N(p) - N(q)| > \lambda K 2^{-(2-\alpha)n}\right) \leq C_1 \exp\left(-C_2 \lambda^2 2^{(2-\alpha)n}\right). \tag{59}$$

Since there are 2^{3n} nearest neighbors in $\mathscr{G}_n$, we have

$$\begin{aligned} P(A^c(n, \lambda)) &\leq C_1 2^{3n} \exp\left(-C_2 \lambda^2 2^{(2-\alpha)n}\right) \\ &\leq C_3 \exp\left(-C_4 \lambda^2 2^{(2-\alpha)n}\right), \end{aligned} \tag{60}$$

for appropriate constants C_3, C_4. Let $A(\lambda) = \cup_{n=0}^{\infty} A(n\,,\lambda)$. Summing the previous estimates over n, we get

$$P(A(\lambda)^c) \leq C_1 \exp\left(-C_2 \lambda^2\right), \tag{61}$$

where C_1, C_2 might be different than before. Let $p = (t\,, x)$. If $A(\lambda)$ holds, then using the path $0 = p_0, \ldots, p_N = (t\,, x)$, we have

$$
\begin{aligned}
|N(t\,,x)| &= |N(t\,,x) - N(0\,,0)| \\
&\leq \sum_{j=1}^{N} |N(p_{j-1}) - N(p_k)| \\
&\leq 4 \sum_{j=1}^{\infty} \lambda K 2^{-(2-\alpha)j} \\
&\leq CK\lambda \\
&\leq \lambda,
\end{aligned}
\tag{62}
$$

with the appropriate choice of K. This proves (47) and hence finishes the proof of Theorem 4.1. □

Exercise. Modify the proof of Theorem 4.1 to get the following modulus of continuity for $u(t, x)$. For all $\varepsilon, T, M > 0$ there exists $K = K(\omega, \varepsilon, T, M)$ such that for all $0 \leq s < t \leq T$ and $-M \leq x < y \leq M$, we have

$$
\begin{aligned}
|u(t, x) - u(s, x)| &\leq K|t - s|^{1/4-\varepsilon} \\
|u(t, x) - u(t, y)| &\leq K|x - y|^{1/2-\varepsilon}
\end{aligned}
$$

5 A comparison theorem

The maximum principle is a powerful method for studying elliptic and parabolic equations, and it can be used to prove comparison theorems of the following type. For simplicity, let $\mathscr{C}$ be the circle $[0\,,1]$ with endpoints identified. Consider two solutions $u_1(t\,,x)$ and $u_2(t\,,x)$, $t \geq 0$, $x \in \mathscr{C}$, of the heat equation

$$
\frac{\partial u}{\partial t} = \Delta u. \tag{63}
$$

Here we assume that $u(t\,,x)$ has periodic boundary conditions on $[0, 1]$, that is $u(t\,,0) = u(t\,,1)$ and $u_x(t\,,0) = u_x(t\,,1)$. Suppose that $u_1(0\,,x) \leq u_2(0\,,x)$. Then for every $t > 0$ and $x \in \mathscr{C}$, $u_1(t\,,x) \leq u_2(t\,,x)$. Indeed, let $v(t\,,x) = u_1(t\,,x) - u_2(t\,,x)$. Then $v(0\,,x)$ is nonpositive, and the maximum principle states that since v satisfies the heat equation, its maximum must be attained on the boundary of the domain, namely at $t = 0$. Thus, $v(t\,,x) \leq 0$ for all $t \geq 0$ and $x \in \mathscr{C}$, so $u_1(t\,,x) \leq u_2(t\,,x)$. This argument can be extended to many semilinear heat equations, see [35].

For stochastic equations, comparison principles for finite dimensional diffusions are known, see [16, Theorem 1.1, Chapter VI]. For example, suppose that $a(x), b(x)$ are Lipschitz functions, and that $x_0 \leq y_0$. Suppose that $X(t), Y(t)$ satisfy

$$\begin{aligned} dX &= a(X)\,dt + b(X)\,dB, \\ dY &= a(Y)\,dt + b(Y)\,dB, \\ X_0 &= x_0, \\ Y_0 &= y_0. \end{aligned} \tag{64}$$

Then with probability 1, for all $t \geq 0$ we have $X_t \leq Y_t$.

Theorem 5.1. *Let $a(u), b(u)$ be Lipschitz functions on $\mathbf{R}$, and consider solutions $u_1(t\,,x), u_2(t\,,x)$, $t \geq 0$, $x \in \mathscr{C}$ to the SPDE*

$$\frac{\partial u}{\partial t} = \Delta u + a(u) + b(u)\dot{W}, \tag{65}$$

with $u_1(0\,,x) \leq u_2(0\,,x)$. Then with probability 1, for all $t \geq 0$, $x \in \mathscr{C}$, we have

$$u_1(t\,,x) \leq u_2(t\,,x). \tag{66}$$

Proof. We will only give an outline of the proof, and only for fixed t. A special case is treated in [28, Section 3], but the proof carries over to our situation. For other approaches, see [30] and [37].

All such proofs follow the same strategy: discretize the SPDE and then use the comparison result for diffusions. Fix $N > 0$ and for $k = 1, \ldots, N$ let

$$u_{i,k,N}(0) = N \int_{k/N}^{(k+1)/N} u_i(0\,,x)\,dx, \tag{67}$$

where the interval $[k/N\,,(k+1)/N]$ is taken to be a subset of $\mathscr{C}$. Also, let

$$B_k(t) = N^{1/2} \int_0^t \int_{k/N}^{(k+1)/N} W(dx\,ds), \tag{68}$$

and note that $B_k(t)$ is a standard Brownian motion. Define the operator $\Delta^{(N)}$ by

$$\Delta^{(N)} u_{i,k,N} = N^2 \left(u_{i,k+1,N} - 2u_{i,k,N} + u_{i,k-1,N}\right). \tag{69}$$

In other words, $\Delta^{(N)}$ is the discrete Laplacian. Because of our periodic boundary conditions, we let $u_{i,N,N} = u_{i,0,N}$.

We will construct $u_{i,k,N}(t)$ in stages. For $j/N^2 < t < (j+1)/N^2$ and $j \geq 0$, let $u_{i,k,N}(t)$ satisfy

$$du_{i,k,N} = a(u_{i,k,N})\,dt + N^{1/2} b(u_{i,k,N})\,dB_k. \tag{70}$$

For $t = j/N^2$, let

$$u_{i,k,N}(t) = u_{i,k,N}(t-) + \frac{1}{N^2}\Delta^{(N)} u_{i,k,N}(t-) \tag{71}$$

where the operator $\Delta^{(N)}$ acts on k. Finally, for

$$\left|x - \frac{k}{N^2}\right| < \frac{1}{2N^2} \tag{72}$$

let $v_{i,N}(t,x) = u_{i,k,N}(t)$. It can be shown that for any $T > 0$,

$$\lim_{N\to\infty} E\left[\int_0^T \int_{\mathscr{C}} |u_i(t\,,x) - v_{i,N}(t\,,x)|^2\,d\,xdt\right] = 0. \tag{73}$$

Exercise. Verify (73).

Furthermore, the comparison theorem for diffusions from [16, Theorem 1.1, Chapter VI], and the positivity of the operator $\Delta^{(N)}$ shows that with probability 1, $u_{1,k,N}(t) \le u_{2,k,N}(t)$. See [28, Section 3] for details.

Modulo the missing details, this gives us our comparison theorem. □

6 Applications

We give several applications of the preceding ideas.

6.1 Blow-up

Blow-up in finite time is a well-studied property for PDE, arising in such applications as flame propagation and the shrinkage of an elastic string to a point. In this section we will show that a certain SPDE does not blow up in finite time. The basic idea is to show that if $\sup_x u(t\,,x)$ is large, then it is more likely to decrease than increase. This intuitive idea is implemented by comparison with a random walk whose steps have negative expectation. Probabilities are controlled using large deviations. We express the solution as a sum of two terms, bound one by large deviations, and bound the other by elementary heat kernel estimates.

Consider the following SPDE on the circle $\mathscr{C}$, which we take to be $[0\,,1]$ with endpoints identified. We impose periodic boundary conditions. Let $\gamma \ge 1$. Assume that $u_0(x)$ is a continuous and nonnegative function on $\mathscr{C}$.

$$\begin{aligned} \frac{\partial u}{\partial t} &= \Delta u + u^\gamma \dot{W}, \\ u(0\,,x) &= u_0(x) \ge 0. \end{aligned} \tag{74}$$

For this subsection, let $G(t\,,x\,,y)$ denote the heat kernel on the circle. Equivalently, we could consider $G(t\,,x\,,y)$ to be the heat kernel on $[0,1]$ with periodic boundary conditions.

Exercise Show that

$$G(t\,,x\,,y) = \sum_{n\in\mathbf{Z}} \Big[G(t\,,x-y+2n) + G(t\,,x+y+2n)\Big] \tag{75}$$

where $G(t,x)$ is the heat kernel on $\mathbf{R}$.

As earlier, the rigorous meaning of (74) is given in terms of an integral equation. Since we are concerned about blowup, we will truncate u^γ. Let

$$\begin{aligned} u_N(t\,,x) = &\int_{\mathscr{C}} G(t\,,x\,,y)u_0(y)\,dy \\ &+ \int_0^t \int_{\mathscr{C}} G(t-s\,,x\,,y)(u_N \wedge N)^\gamma(s\,,y)\,W(dy\,ds). \end{aligned} \tag{76}$$

Since $(u_N \wedge N)^\gamma$ is a Lipschitz function of N, the usual theory implies existence and uniqueness for solutions u_N to (76). See [38], Chapter 3, for example. Let τ_N be the first time $t \geq 0$ that $u_N(t\,,x) \geq N$ for some $x \in [0\,,1]$. Then we can construct $u_N : N = 1, 2, \ldots$ on the same probability space, and the limit $\tau_N \uparrow \tau$ exists almost surely. Note that almost surely, $u_m(t\,,x) = u_n(t\,,x)$ as long as $t < \tau_m \wedge \tau_n$. Then we may define $u(t,x)$ for $t < \tau$ by setting $u(t\,,x) = u_n(t,x)$ whenever $t < \tau_n$ for some n, and we see that $\tau \in [0\,,\infty]$ is the time at which $\sup_x u(t\,,x) = \infty$, and $\tau = \infty$ if $u(t\,,x)$ never reaches ∞.

Definition 6.1. *We say that $u(t\,,x)$ blows up in finite time if $\tau < \infty$.*

Then we have

Theorem 6.2. *If $\gamma < 3/2$ then with probability 1, $u(t\,,x)$ does not blow up in finite time.*

We remark that Krylov [18] has a different proof of Theorem 6.2, including generalizations, based on his L_p theory of SPDE.

Proof (Theorem 6.2). First, we claim that

$$U(t) := \int_{\mathscr{C}} u(t\,,x)\,dx \tag{77}$$

is a local martingale. It suffices to show that $U_N(t) := \int_0^1 u_N(t\,,x)\,dx$ is a martingale for each N. Of course

$$\int_{\mathscr{C}} G(t\,,x\,,y)\,dx = 1. \tag{78}$$

Integrating (76) over x and using the stochastic Fubini's lemma [38], we get

$$U_N(t) = \int_{\mathscr{C}} u_0(y)\,dy + \int_0^t \int_{\mathscr{C}} (u_N \wedge N)^\gamma(s\,,y)\,W(dy\,ds), \tag{79}$$

which is a martingale.

Since $U(t)$ is a nonnegative continuous local martingale, it must be bounded, and so

$$U(t) \leq K = K(\omega). \tag{80}$$

for some $K(\omega) < \infty$ almost surely. We would like to study the maximum of $u(t,x)$ over x,

$$M(t) = \sup_{x \in \mathscr{C}} u(t\,,x). \tag{81}$$

Our goal is to show that $M(t)$ does not reach ∞ in finite time, with probability 1. To that end, we define a sequence of stopping times $\tau_n : n \geq 0$ as follows. Let $\tau_0 = 0$, and for simplicity assume that $M(0) = 1$. Given τ_n, let τ_{n+1} be the first time $t > \tau_n$ such that $M(t)$ equals either $2M(\tau_n)$ or $(1/2)M(\tau_n)$. The reader can easily verify that τ_n is well defined for all values of n. Next, we wish to show that if $M(\tau_n)$ is large enough, say $M(\tau_n) > 2^{N_0}$ for some $N_0 > 0$, then

$$P\left(M(\tau_{n+1}) = 2M(\tau_n) \,\middle|\, \mathscr{F}_n\right) < \frac{1}{3}. \tag{82}$$

Assuming (82), we can compare $Z_n = \log_2[M(\tau_n) - N_0]$ to a nearest-neighbor random walk R_n, with $P(R_{n+1} = R_n + 1) = 1/3$. In particular we claim that Z_n, R_n can be constructed on the same probability space such that the following comparison holds. Let σ_n be the first time n such that $Z_n \leq N_0$. Then almost surely, for all $n \leq \sigma_{N_0}$,

$$Z_n \leq R_n.$$

Such a random walk R_n always returns to 0 if $R_0 > 0$.

Exercise. Fill in the details of the above comparison.

Therefore, Z_n either visits the level 2^{N_0} infinitely often, or tends to 0. If $Z_n \to 0$, then clearly $u(t,x)$ does not blow up, either in finite or infinite time. On the other hand, suppose Z_n visits the level 2^{N_0} infinitely often, at the times $\tau_{n_k} : k = 1, 2, \ldots$. We claim that there exists a constant $C > 0$ such that $E[\tau_{n_k+1} - \tau_{n_k} \,|\, \mathscr{F}_n] > C$ for all k. This follows from the strong Markov property of solutions, namely, if $\mathscr{F}_t$ is the σ-field generated by the noise $\dot{W}(s,x)$ for time $s \leq t$, then our solution $u(t,x)$ is a Markov process. It is intuitively clear that the heat equation starts afresh at any given time, with the current solution as new initial value.

Exercise. Prove the Markov property of $u(t,x)$ from (76).

Exercise. Give a full proof of the existence of the constant C mentioned above.

Therefore $\lim_{t\to\infty} \tau_{n_k} = \infty$, and so with probability 1, $u(t\,,x)$ does not blow up in finite time.

Now we turn to the proof of (82). It suffices to consider $u_{2^{m+1}}(t\,,x)$. Suppose that $M(\tau_n) = 2^m$ and let $v(t\,,x) = u_{2^{m+1}}(\tau_n + t, x)$. Using the strong Markov property for $u(t\,,x)$, and the fact that $v(t\,,x) \leq 2^{m+1}$, we have

$$\begin{aligned} v(t\,,x) &= \int_{\mathscr{C}} G(t\,,x\,,y) v(0\,,y)\,dx \\ &\qquad + \int_0^t \int_{\mathscr{C}} G(t-s\,,x\,,y) v(s\,,y)^\gamma W(y\,,s)\,dy\,ds \\ &=: V_1(t\,,x) + V_2(t\,,x). \end{aligned} \tag{83}$$

Note that by the maximum principle, $\sup_x V_1(t,x)$ is nonincreasing. Our goal is to choose a nonrandom time T such that for m large enough,

$$P\left(\sup_{0\le t\le T}\sup_{x\in\mathscr{C}} V_2(t\,,x) > 2^{m/4}\right) < \frac{1}{3} \tag{84}$$

and

$$\sup_{x\in\mathscr{C}} V_1(t\,,x) \le 2^{m/4}. \tag{85}$$

Exercise. Check that (84) and (85) imply (82).

Now for the heat kernel $G(t,x)$ on $x\in\mathbf{R}$,

$$G(t\,,x-y) = \frac{1}{\sqrt{4\pi t}}\exp\left(-\frac{(x-y)^2}{4t}\right) \le Ct^{-1/2} \tag{86}$$

for $C=(4\pi)^{-1/2}$.

Exercise. Use (75) and (86) to show that for some constant $C>0$, the heat kernel $G(t\,,x\,,y)$ on $\mathscr{C}$ satisfies

$$G(t\,,x\,,y) \le Ct^{-1/2} \tag{87}$$

for all $t>0$, $x,y\in\mathscr{C}$.

So for all $x\in\mathscr{C}$ we have

$$V_1(t\,,x) \le Ct^{-1/2}\int_{\mathscr{C}} v(t\,,y)\,dy \le CKt^{-1/2}, \tag{88}$$

where $K=K(\omega)$ was our upper bound for the integral of $u(t\,,x)$ over x.

Now choose K_0 such that if $\mathscr{A}=\{K(\omega)>K_0\}$ then

$$P(\mathscr{A}) < \frac{1}{6}. \tag{89}$$

Also choose

$$T = C^2K_0^2 2^{-m/2-6}. \tag{90}$$

This choice of T gives us (85).

Next we prove (84). Note that

$$u_{2^{m+1}}(t\,,x)^{\gamma} \le C_3 2^{\gamma m} \tag{91}$$

We ask the reader to believe that Theorem 4.1 also holds for equations on $\mathscr{C}$. If this is granted, we have

$$\begin{aligned}
&P\left(\sup_{0\le t\le T}\sup_{x\in\mathscr{C}} V_2(t\,,x) > 2^{m/4}\right)\\
&\qquad\le P(\mathscr{A}^c) + C_0\exp\left(-\frac{C_1(2^{m/4})^2}{T^{1/2}(C_3 2^{\gamma m})^{1/2}}\right)\\
&\qquad\le \frac{1}{6} + C_0\exp\left(-C_4 2^{(m/2)(1-\gamma+(1/2))}\right)
\end{aligned} \tag{92}$$

But if $\gamma<3/2$ then $1-\gamma+(1/2)>0$, and the above probability is less than $1/6$ for $m>m_0$ and m_0 large enough.

This verifies (85), and finishes the proof of Theorem 6.2. □

6.2 Hitting zero

Next we determine the critical drift for an SPDE to hit 0, given that the initial function is strictly positive and bounded away from 0. The argument is similar to the proof of Theorem 6.2. We show that if $\inf_x u(t\,,x)$ is small, then it is more likely to increase than decrease. We express the solution as a sum of two terms, bound one by large deviations, and bound the other by elementary heat kernel estimates. For $t > 0, x \in \mathscr{C}$, we consider solutions $u(t\,,x)$ to

$$\begin{aligned} \frac{\partial u}{\partial t} &= \Delta u + u^{-\alpha} + \dot{W}(t\,,x) \\ u(0\,,x) &= u_0(x) \end{aligned} \tag{93}$$

where $0 < c < u_0(x) < C$ for some constants c, C. The term $u^{-\alpha}$ is singular at $u = 0$, so once again we must restrict to $0 < t < \tau$ where τ is the first time t such that $u(t, x) = 0$. If there is no such time, we let $\tau = \infty$.

Theorem 6.3. *Let $\mathscr{A}$ be the event that $\tau < \infty$, that is, u hits 0 in finite time.*

1. *If $\alpha > 3$ then $P(\mathscr{A}) = 0$.*
2. *If $\alpha < 3$ then $P(\mathscr{A}) > 0$.*

Proof. We will only prove assertion (1). For (2) the reader can consult [25].

As in the previous section, our strategy is to compare $\inf_x \log_2 u(t\,,x)$ to a nearest-neighbor random walk. To verify assertion (1), we need to show that for u small enough, the random walk has a higher probability of moving up than down.

Let $I(t) = \inf_{x\in\mathscr{C}} u(t\,,x)$. We construct a sequence of stopping times τ_n : $n \geq 0$. Let $\tau_0 = 0$. Given τ_n, let τ_{n+1} be the first time $t > \tau_n$ such that $I(t) = 2I(\tau_n)$ or $I(t) = (1/2)I(\tau_n)$. Let $Z_n = \log_2 I(\tau_n)$. Fix n, and let $v(t\,,x) = u(\tau_n + t\,,x)$. Let $\mathscr{A}_n$ be the event that $Z_{n+1} = Z_n - 1$. We claim that for Z_n small enough,

$$P\left(\mathscr{A}_n^c \,\middle|\, \mathscr{F}_{\tau_n}\right) < \frac{1}{3}. \tag{94}$$

First, by the comparison principle of Theorem 5.1, it is enough to show (94) for $v(0\,,x) = I(\tau_n)$. Let $I(\tau_n) = 2^{-m}$. Using the strong Markov property, we can write

$$\begin{aligned} v(t\,,x) &= \int_{\mathscr{C}} G(t\,,x\,,y) v(0\,,y)\,dx \\ &\quad + \int_0^t \int_{\mathscr{C}} G(t-s\,,x\,,y) v(s\,,y)^{-\alpha}\,dy\,ds \\ &\quad + \int_0^t \int_{\mathscr{C}} G(t-s\,,x\,,y) W(dy\,ds)\,dy\,ds \\ &=: 2^{-m} + V_1(t\,,x) + V_2(t\,,x). \end{aligned} \tag{95}$$

Fix $\delta > 0$, and let

$$T = 2^{-4m-2\delta}. \tag{96}$$

Let $\mathscr{B} = \mathscr{B}(T\,,\tau_n)$ be the event that

$$\sup_{0\le t\le T}\sup_{x\in\mathscr{C}} |V_2(t\,,x)| \le 2^{-m-1}. \tag{97}$$

By Theorem 4.1 (for $\mathscr{C}$), we have

$$\begin{aligned} P\left(\mathscr{B}^c\right) &\le C_0 \exp\left(-\frac{C_1(2^{-m-1})^2}{T^{1/2}}\right) \\ &\le C_0 \exp\left(-C_2 2^{\delta m}\right) \\ &< \frac{1}{3}, \end{aligned} \tag{98}$$

if m is large enough.

We claim that on $\mathscr{B}$,

$$\sup_{0<t<T}\sup_{x\in\mathscr{C}} |V_1(t\,,x)| \le 2^{m-1}. \tag{99}$$

Indeed, let $v_m(t\,,x)$ satisfy

$$\begin{aligned} v_m(t\,,x) = &\int_{\mathscr{C}} G(t\,,x\,,y) v(0\,,y)\,dx \\ &+ \int_0^t \int_{\mathscr{C}} G(t-s\,,x\,,y)\left(v_m(s\,,y) \vee 2^{-m-1}\right)^{-\alpha}\,dy\,ds \\ &+ \int_0^t \int_{\mathscr{C}} G(t-s\,,x\,,y)\,W(dy\,ds). \end{aligned} \tag{100}$$

Then $v_m(t\,,x) = v(t\,,x)$ for all $x \in \mathscr{C}$ and for all $0 < t < \sigma_m$, where σ_m is the first time s that

$$\inf_{x\in\mathscr{C}} v_m(s\,,x) \le 2^{-m-1}. \tag{101}$$

However, on the event $\mathscr{B}$, since $v(0\,,x) \equiv 2^{-m}$, we see that

$$\begin{aligned} \inf_{0<t<T}\inf_{x\in\mathscr{C}} v_m(t\,,x) &\ge 2^{-m} - \sup_{0<t<T}\sup_{x\in\mathscr{C}} |V_1(t\,,x)| \\ &\ge 2^{-m-1}, \end{aligned} \tag{102}$$

by the definition of $\mathscr{B}$. Therefore, on $\mathscr{B}$, $v_m(t\,,x) = v(t\,,x)$ for $0 < t < T$ and $x \in \mathscr{C}$. It follows that on $\mathscr{B}$, $v(t\,,x) \ge 2^{-m-1}$ for $0 < t < T$ and $x \in \mathscr{C}$. It follows that on $\mathscr{B}$ and for $0 < t < T$ and $x \in \mathscr{C}$,

$$\begin{aligned} v(t\,,x) &\leq \int_{\mathscr{C}} G(t\,,x\,,y)v(0\,,y)\,dx \\ &\qquad + \int_0^t \int_{\mathscr{C}} G(t-s\,,x\,,y)\left(2^{-m-1}\right)^{-\alpha}\,dy\,ds \\ &\qquad + \int_0^t \int_{\mathscr{C}} G(t-s\,,x\,,y)\,W(dy\,ds) \\ &\leq 2^{-m} + CT2^{m\alpha} + 2^{-m-1} \\ &< 2^{-m+1}. \end{aligned} \tag{103}$$

for m large enough. Thus, if m is large enough,

$$P\left(Z_{n+1} = Z_n + 1 \,\middle|\, \mathscr{F}_n\right) \leq P(\mathscr{B}) < \frac{1}{3}. \tag{104}$$

This proves the comparison of Z_n with a random walk, and finishes the proof of Theorem 6.3. □

Remark 6.4. Recently, Zambotti [39; 40] has found remarkable connections between the problem of hitting 0 and the random string, which is a vector-valued solution of the heat equation with noise.

6.3 Compact support

We all know that nonnegative solutions to the heat equation which are not identically zero have support on all of $\mathbf{R}^d$. It is therefore of great interest that certain stochastic heat equations have solutions with compact support. Intuitively, at the edge of the support the noise term dominates, so it has a chance to push the solution to 0.

Consider solutions $u(t\,,x)$, $0 \leq t < \infty$, $x \in \mathbf{R}$, to

$$\begin{aligned} \frac{\partial u}{\partial t} &= \Delta u + u^{\gamma}\dot{W}(t\,,x) \\ u(0\,,x) &= u_0(x), \end{aligned} \tag{105}$$

where $u_0(x)$ is a continuous nonnegative function of compact support, and

$$\frac{1}{2} < \gamma < 1. \tag{106}$$

By an ingenious argument using duality, Mytnik [29] has proved uniqueness in law. Actually, approximate duality is used, not exact duality. Roughly speaking, the dual process $v(t\,,x)$ satisfies

$$\begin{aligned} \frac{\partial v}{\partial t} &= \Delta v + v^{1/\gamma}\dot{L}(t\,,x) \\ v(0\,,x) &= v_0(x) \geq 0, \end{aligned} \tag{107}$$

where $\dot{L}(t\,,x)$ is a one-sided Lévy noise with positive jumps, of index 2γ. For details, see [29]. Mytnik's duality relation is

$$H(u\,,v) = \exp\left(-\int_{\mathbf{R}} u(x)v(x)\,dx\right), \tag{108}$$

which is one of the standard duality functions. (At least this last part is easy to guess).

Theorem 6.5. *With probability 1, $u(t\,,x)$ has compact support in x for all $t \geq 0$.*

Proof. We will mainly discuss the proof for $\gamma = 1/2$, which is much easier. This argument is essentially due to Iscoe [15]. Let

$$v(x) = \begin{cases} 12(x+R)^{-2} & \text{for } x > -R \\ \infty & \text{for } x \leq -R. \end{cases} \tag{109}$$

and note that for $x > -R$,

$$\Delta v(x) = \frac{1}{2}v(x)^2. \tag{110}$$

Define $0 \cdot \infty = 0$, and let

$$M_t = \exp\left(-\int_{\mathbf{R}} v(x)u(t\,,x)\,dx\right). \tag{111}$$

By Ito's lemma (see (9) and (10)), we have

$$\begin{aligned} dM_t = M_t\bigg(-\int_{x\in\mathbf{R}} \left[v(x)\Delta u(t\,,x) + \frac{1}{2}v(x)^2 u(t\,,x)\right] dx\bigg)\,dt \\ - M_t \int_{x\in\mathbf{R}} v(x)u(t\,,x)^{1/2}\,W(dx\,dt) \qquad (112) \\ = -M_t \int_{x\in\mathbf{R}} v(x)u(t\,,x)^{1/2}\,W(dx\,dt). \end{aligned}$$

Here we have used the martingale problem formulation to do integration by parts, ie to replace $v\Delta u$ by $u\Delta v$. Then we used (110) to substitute for Δv. Actually, to justify this calculation, we must truncate $v(x)$ and let the truncation level tend to ∞. We leave these details to the reader.

Thus M_t is a local martingale. Fix $T > 0$, and let τ be the infimum of times $t \leq T$ for which

$$\int_{-\infty}^{R} u(t\,,x)\,dx > 0 \tag{113}$$

and let $\tau = T$ if there is no such time. If $\tau < T$, we say that u charges $(-\infty\,,R]$ before time T. Since $0 \leq M_t \leq 1$, we can apply the optional sampling theorem to conclude that

$$M_0 = EM_\tau. \tag{114}$$

Let $\mathscr{A}$ be the event that u does not charge $(-\infty\,, R]$ before time T. Note that on $\mathscr{A}^c$ we have $M_\tau = 0$, while on $\mathscr{A}$ we have $M_\tau \leq 1$. Therefore,

$$P(\mathscr{A}) \geq EM_\tau = M_0 = \exp\left(-\int_{\mathbf{R}} v(x)u(0\,,x)\,dx\right) \tag{115}$$

From (110) we see that $v(x) = v(R\,,x) \to 0$ uniformly on compact intervals. Since $u(0\,,x)$ is continuous with compact support, it follows that the right hand side tends to 0 as $R \to \infty$. □

For $1/2 < \gamma < 1$, Mueller and Perkins [24] gave a more complicated argument proving compact support in this situation, but one which gave information which has proved useful for other problems.

Note 6.6. Compact support can also occur in deterministic heat equations such as

$$\frac{\partial u}{\partial t} = \Delta u - u^\rho \tag{116}$$

when $\rho < 1$. Assume that $u(t\,,x)$ is nonnegative. For small values of u, we have $u^\rho \gg u$, so the final term can push the equation to 0. More complicated equations of this type appear in chemical engineering, and the region where $u(t\,,x) = 0$ is called a dead zone. Chemical engineers try to minimize the dead zone, since no reactions take place there, and this leads to inefficient use of the reactor vessel.

6.4 Noncompact support

On the other hand, some stochastic heat equations have solutions whose support is all of $\mathbf{R}^d$. Note that in the following equation, no matter what the size of u, the noise term is comparable to the other terms in the equation.

Theorem 6.7. *Suppose that $u(t\,,x)$, $t \geq 0$, $x \in \mathbf{R}$ satisfies*

$$\begin{aligned} \frac{\partial u}{\partial t} &= \Delta u + u\dot{W}(t,x) \\ u(0\,,x) &= u_0(x), \end{aligned} \tag{117}$$

for $u_0(x)$ continuous, nonnegative, and not identically 0. Then with probability 1, $u(t\,,x) > 0$ for all $t > 0, x \in \mathbf{R}$.

The assertion seems intuitively obvious, since solutions to the heat equation have support on all of $\mathbf{R}$. However, the previous section shows that certain heat equations can have solutions of compact support.

Proof. Working with $x \in \mathbf{R}$ gives rise to technical complications, so let us suppose that $u(t,x) : x \in [-2R\,,2R]$ satisfies

$$\begin{aligned} \frac{\partial u}{\partial t} &= \Delta u + u\dot{W}(t\,,x), \\ u(t\,,-2R) &= u(t\,,2R) = 0, \\ u(0\,,x) &= u_0(x). \end{aligned} \tag{118}$$

As before, we would give rigorous meaning to this equation in terms of the following integral equation.

$$\begin{aligned} u(t\,,x) = &\int_{-2R}^{2R} G_D(t\,,x\,,y)u_0(y)\,dy \\ &+ \int_0^t \int_{-2R}^{2R} G_D(t-s\,,x\,,y)u(s\,,y)\,W(dy\,ds), \end{aligned} \tag{119}$$

where $G_D(t\,,x\,,y)$ is the fundamental solution to the heat equation on $[-2R\,,2R]$ with Dirichlet boundary conditions. It is a standard fact that

$$G_D(t\,,x\,,y) \le G(t\,,x-y). \tag{120}$$

It is not hard to modify the proof of the comparison theorem, Theorem 5.1, to show that the solution to (118) is less than or equal to the solution of (117). From now on, let $u(t\,,x)$ denote the solution to (118). It is enough to show that with probability 1, $u(t\,,x)$ satisfies

$$\operatorname{supp}(u(t\,,\cdot)) \supset [-R\,,R] \tag{121}$$

for all $t > 0$.

Note that the equation is linear, so $v(t\,,x) = cu(t\,,x)$ satisfies the same equation, with different initial conditions, of course. We will subdivide time into stages, and show that with high probability, at each stage the support expands a little more.

Translating $u(0\,,x)$ if necessary, we may assume that $u(0\,,x) \ge \delta\mathbf{1}_{[-a,a]}$ for some $\delta, a > 0$. For simplicity, assume that $a = 1$. Let

$$\mathscr{S}_t = \operatorname{supp}(u(t\,,\cdot)) \tag{122}$$

It suffices to show that for $T, R, \delta > 0$,

$$[-R\,,R] \subset \mathscr{S}_T. \tag{123}$$

For simplicity we will present the proof for $T = 1$ and $R > 2$. Fix $N > 0$ and let $t_k = Tk/N$. Let $\mathscr{A}_k$ be the event that

$$u(t_k\,,x) \ge \delta I_k(x) \tag{124}$$

for some δ, where

$$I_k(x) = \mathbf{1}\left(-1 - \frac{Rk}{N} \le x \le 1 + \frac{Rk}{N}\right). \tag{125}$$

Note that

$$I_N(x) \ge \mathbf{1}_{[-R,R]}(x). \tag{126}$$

We wish to show that for all $\varepsilon > 0$, we can choose N so large that for all $k = 1, \dots, N$

$$P\left(\mathscr{A}_{k+1}^c \,\middle|\, \mathscr{A}_1 \cap \cdots \cap \mathscr{A}_k\right) < \frac{\varepsilon}{N}. \tag{127}$$

Note that $\mathscr{A}_0$ occurs by assumption. It would then follow that

$$\begin{aligned} P\left(\mathscr{A}_N^c\right) &\le \sum_{k=0}^{N-1} P\left(\mathscr{A}_{k+1}^c \,\middle|\, \mathscr{A}_1 \cap \cdots \cap \mathscr{A}_k\right) \\ &\le \varepsilon. \end{aligned} \tag{128}$$

But

$$P\Big(\mathscr{S}_1 \subset [-R\,,R]\Big) \ge P\left(\mathscr{A}_N\right) \ge 1 - \varepsilon. \tag{129}$$

and since ε is arbitrary, we would be done.

Now we turn to the proof of (127). Assuming that $\mathscr{A}_k$ occurs, we have $u(t_k\,,x) \ge \delta I_k(x)$. By the comparison theorem, Theorem 5.1, it is enough to show (127) assuming that $u(t_k, x) = \delta I_k(x)$. Now let

$$v_k(t\,,x) = (1/\delta)u(t_k + t\,,x), \tag{130}$$

so that $v_k(t\,,x)$ satisfies (117), and $v_k(0\,,x) = I_k(x)$. Let

$$\eta(N,k) = \int_{-2R}^{2R} G(t\,,1 + (k+1)/N - y)v_k(0\,,y)\,dy, \tag{131}$$

and let

$$\eta(N) = \inf_{0 \le k \le N} \eta(N,k). \tag{132}$$

We leave it to the reader to show that $\eta(N,k)$ is increasing in k, so that $\eta(N) = \eta(N,0)$ and that

$$\eta := \inf_N \eta(N) > 0. \tag{133}$$

Hint: roughly speaking, the heat kernel spreads a distance $1/N^{1/2}$ in time $1/N$, so it must be close to 1 at distance $1/N$ from the edge of the support of the indicator function.

Next, let $w_k(t\,,x)$ satisfy

$$\begin{aligned} w_k(t\,,x) = \int_{-2R}^{2R} &G_D(t\,,x\,,y)I_k(y)\,dy \\ &+ \int_0^t \int_{-2R}^{2R} G_D(t-s\,,x\,,y)u(s\,,y)\,W(dy\,ds), \end{aligned} \tag{134}$$

and let

$$N_w(t\,,x) = \int_0^t \int_{-2R}^{2R} G_D(t-s\,,x\,,y)u(s\,,y)\,W(dy\,ds). \tag{135}$$

In the same way as for Theorem 4.1, it is not hard to prove that if $0 < M < R$, then there exist constants $C_1, C_2 > 0$ such that for all $T, K, \lambda > 0$,

$$P\left(\sup_{0\le t\le T}\sup_{|x|\le M}|N_w(t\,,x)| > \lambda\right) \le C_1 \exp\left(-\frac{C_2\lambda^2}{T^{1/2}K^2}\right). \tag{136}$$

Now let $\lambda = \eta/2$, and let $T = 1/N$. Thus, given ε, we can choose N so large that the right hand side of (136) is less than ε.

This proves (127), and so finishes the proof of Theorem 6.7. □

References

[1] Maury D. Bramson (1978). Maximal displacement of branching Brownian motion, *Comm. Pure Appl. Math.* **31**, 531–581

[2] Maury Bramson (1983). Convergence of solutions of the Kolmogorov equation to travelling waves, *Mem. Amer. Math. Soc.* **44**(285)

[3] Albert-László Barabási and H. Eugene Stanley (1995). *Fractal Concepts in Surface Growth*, Cambridge University Press, Cambridge

[4] Pao-Liu Chow (2007). *Stochastic partial differential equations.* Chapman & Hall/CRC Applied Mathematics and Nonlinear Science Series. Chapman & Hall/CRC, Boca Raton, FL.

[5] J. T. Cox and R. Durrett (1995). Hybrid zones and voter model interfaces. *Bernoulli*, **1**(4), 343–370

[6] D. A. Dawson (1993). *Measure-valued Markov processes*, In: Lecture Notes in Mathematics **1180**, 1–260, Springer-Verlag, Berlin, Heidelberg, New York

[7] M. Doi and S.F. Edwards (1988). *The Theory of Polymer Dynamics*, Oxford University Press, Oxford

[8] Giuseppe Da Prato and Jerzy Zabczyk (1992). *Stochastic Equations in Infinite Dimensions*, Cambridge University Press, Cambridge

[9] R. Durrett (1988). *Lecture Notes on Particle Systems and Percolation*, Wadsworth and Brooks/Cole, Pacific Grove

[10] S. Ethier and T. Kurtz (1986). *Markov Processes: Characterization and Convergence*, John Wiley & Sons Inc., New York

[11] Alison M. Etheridge (2000). *An Introduction to Superprocesses*, Amer. Math. Soc., Providence

[12] Marco Ferrante and Marta Sanz-Solé (2006). SPDEs with coloured noise: Analytic and stochastic approaches, *ESAIM Probab. Stat.* **10**, 380–405 (electronic)

[13] M. I. Freidlin and A. D. Wentzell (1998). *Random Perturbations of Dynamical Systems*, Springer-Verlag, New York, second edition. Translated from the 1979 Russian original by Joseph Szücs

[14] A. M. Garsia, E. Rodemich, and H. Rumsey, Jr. (1970/1971). A real variable lemma and the continuity of paths of some Gaussian processes. *Indiana Univ. Math. J.* **20**, 565–578

[15] I. Iscoe (1988). On the supports of measure-valued critical branching Brownian motion, *Probab. Th. Rel. Fields* **16**, 200–221

[16] N. Ikeda and S. Watanabe (1989). *Stochastic Differential Equations and Diffusion Processes*, North Holland/Kodansha, Amsterdam, Oxford, New York

[17] N.V. Krylov and B.L. Rozovskii (1982). Stochastic partial differential equations and diffusion processes, *Russian Math. Surveys* **37**(6), 81–105. Originally in Uspekhi Mat. Nauk, **37**:6, 75–95

[18] N.V. Krylov (1996). On L_p-theory of stochastic partial differential equations in the whole space, *SIAM J. Math Anal.* **27**(2), 313–340

[19] Nicolai V. Krylov (2006). On the foundation of the L_p-theory of stochastic partial differential equations, In: *Stochastic Partial Differential Equations and Applications–VII*, *Lect. Notes Pure Appl. Math.* **245**, 179–191, Chapman & Hall/CRC, Boca Raton

[20] Jean-François Le Gall (1999). *Spatial Branching Processes, Random Snakes and Partial Differential Equations*, Lectures in Mathematics ETH Zürich, Birkhäuser-Verlag, Basel

[21] Thomas M. Liggett (1999). *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes*, Springer-Verlag, Berlin

[22] Thomas M. Liggett (2005). *Interacting Particle Systems*, Springer-Verlag, Berlin. Reprint of the 1985 original.

[23] Jonathan C. Mattingly (2003). On recent progress for the stochastic Navier Stokes equations, In: *Journées "Équations aux Dérivées Partielles"*, pages Exp. No. XI, 52. Univ. Nantes, Nantes

[24] C. Mueller and E. Perkins (1992). The compact support property for solutions to the heat equation with noise, *Probab. Th. Rel. Fields* **93**, 325–358

[25] Carl Mueller and Etienne Pardoux (1999). The critical exponent for a stochastic PDE to hit zero, In: *Stochastic Analysis, Control, Optimization and Applications*, 325–338, Birkhäuser Boston, Boston

[26] C. Mueller and R. Sowers (1995). Random traveling waves for the KPP equation with noise, *J. Funct. Anal.* **128**, 439–498

[27] C. Mueller and R. Tribe (1997). Finite width for a random stationary interface, *Electronic J. Prob.* **2**, Paper no. 7, 1–27

[28] C. Mueller (1991). On the support of the heat equation with noise, *Stochastics* **37**(4), 225–246

[29] L. Mytnik (1998). Weak uniqueness for the heat equation with noise, *Ann. Probab.*,**26**(3), 968–984

[30] D. Nualart and É. Pardoux (1992). White noise driven quasilinear SPDEs with reflection, *Probab. Theory Related Fields* **93**(1), 77–89

[31] Étienne Pardoux (1972). Sur des équations aux dérivées partielles stochastiques monotones, *C. R. Acad. Sci. Paris Sér. A-B* **275**, A101–A103

[32] Étienne Pardoux (1991). Filtrage non linéaire et équations aux dérivées partielles stochastiques associées, In: *Lecture Notes in Math.* **1464**, 67–163, Springer, Berlin

[33] Edwin Perkins (2002). Dawson–Watanabe superprocesses and measure-valued diffusions, In: *Lecture Notes in Math.* **1781**, 125–324, Springer, Berlin

[34] Claudia Prévôt and Michael Röckner (2007). *A concise course on stochastic partial differential equations*, volume 1905 of *Lecture Notes in Mathematics*. Springer, Berlin.

[35] Murray H. Protter and Hans F. Weinberger (1984). *Maximum Principles in Differential Equations*, Springer-Verlag, New York

[36] Daniel Revuz and Marc Yor (1999). *Continuous Martingales and Brownian Motion*, Springer-Verlag, Berlin, third edition

[37] T. Shiga (1994). Two contrasting properties of solutions for one-dimensional stochastic partial differential equations, *Can. J. Math* **46**(2), 415–437

[38] J. B. Walsh (1986). *An Introduction to Stochastic Partial Differential Equations*, In: Lecture Notes in Mathematics **1180**, 265–439, Springer-Verlag, Berlin, Heidelberg, New York

[39] Lorenzo Zambotti (2003). Integration by parts on δ-Bessel bridges, $\delta \geq 3$ and related SPDEs, *Ann. Probab.* **31**(1), 323–348

[40] Lorenzo Zambotti (2005). Integration by parts on the law of the reflecting Brownian motion, *J. Funct. Anal.* **223**(1), 147–178

List of Participants

Tom Alberts
New York University
New York, USA
alberts@cims.nyu.edu

Hakima Bessaih
University of Wyoming
Laramie, USA
bessaih@uwyo.edu

Ryan Card
University of Washington
Seattle, USA
card@math.washington.edu

Robert C. Dalang
Ecole Polytechnique Fédérale
Lausanne, Switzerland
robert.dalang@epfl.ch

Sabrina Guettes
University of Wisconsin
Madison, USA
guettes@math.wisc.edu

Gunnar Gunnarsson
University of California
Santa Barbara, USA
gungun@math.ucsb.edu

Samuel Isaacson
University of Utah
Salt Lake City, USA
isaacson@math.utah.edu

Karim Khader
University of Utah
Salt Lake City, USA
khader@math.utah.edu

Davar Khoshnevisan
University of Utah
Salt Lake City, USA
davar@math.utah.edu

An Le
Utah State University
Logan, USA
anle@math.utah.edu

Scott McKinley
Ohio State University
Columbus, USA
mckinley@math.ohio-state.edu

Carl Mueller
University of Rochester
Rochester, USA
cmlr+new@math.rochester.edu

David Nualart
University of Kansas
Lawrence, USA
nualart@math.ku.edu

Chris Orum
Los Alamos National Laboratory
New Mexico, USA
orumc@lanl.gov

Michael Purcell
University of Utah
Salt Lake City, USA
purcell@math.utah.edu

Firas Rassoul-Agha
University of Utah
Salt Lake City, USA
firas@math.utah.edu

Jean-François Renaud
Université de Montreal
Montréal, Canada
renaud@dms.umontreal.ca

Boris Rozovsky
University of Southern California
Los Angeles, USA
rozovski@math.usc.edu

Shang-Yuan Shiu
University of Utah
Salt Lake City, USA
shiu@math.utah.edu

Blerta Shtylla
University of Utah
Salt Lake City, USA
shtyllab@math.utah.edu

Andrejs Treibergs
University of Utah
Salt Lake City, USA
treiberg@math.utah.edu

Dongsheng Wu
Michigan State University
East Lansing, USA
wudongsh@msu.edu

Yimin Xiao
Michigan State University
East Lansing, USA
xiao@stt.msu.edu