Electronic Journal of Probability

Electron. J. Probab. **24** (2019), no. 14, 1–26.
ISSN: 1083-6489 https://doi.org/10.1214/19-EJP279

Can the stochastic wave equation with strong drift hit zero?*

Kevin Lin† Carl Mueller‡

Abstract

We study the stochastic wave equation with multiplicative noise and singular drift:

$$\partial_t u(t,x) = \Delta u(t,x) + u^{-\alpha}(t,x) + g(u(t,x))\dot{W}(t,x)$$

where x lies in the circle $\mathbf{R}/J\mathbf{Z}$ and $u(0,x) > 0$. We show that
(i) If $0 < \alpha < 1$ then with positive probability, $u(t,x) = 0$ for some (t,x).
(ii) If $\alpha > 3$ then with probability one, $u(t,x) \neq 0$ for all (t,x).

Keywords: wave equation; white noise; stochastic partial differential equations.
AMS MSC 2010: Primary 60H15, Secondary 60J45; 35L05.
Submitted to EJP on February 26, 2018, final version accepted on February 11, 2019.
Supersedes arXiv:1802.09487.

1 Introduction

One of the classic questions about stochastic processes is whether they can hit a given set. That is, for a process X_t taking values in a space S, and for $A \subset S$, do we have

$$\mathbf{P}(X_t \in A \text{ for some } t) > 0.$$

For example, consider the Bessel process R_t with parameter n, which satisfies

$$dR = \frac{n-1}{2R}dt + dW$$

where $W(t)$ is a one-dimensional Brownian motion and we assume that $R_0 > 0$. It is well known that if we allow n to take nonnegative real values, then R_t can hit 0 iff $n < 2$. For Markov processes such as R_t, harmonic functions and potential theory are powerful tools which have led to rather complete answers to such questions; see [6] or most other books in Markov processes.

*The second author was supported by a Simons grant.
†Upstart Network, Inc. E-mail: kevin.lin790@gmail.com, https://www.upstart.com
‡Dept. of Mathematics, University of Rochester, Rochester NY 14627
E-mail: carl.e.mueller@rochester.edu, http://www.math.rochester.edu/people/faculty/cmlr

For stochastic partial differential equations (SPDE), potential theory becomes less tractible due to the infinite-dimensional state space of solutions, and hitting questions have not been as thoroughly studied. To be specific, solutions $u(t,x)$ usually depend on a time parameter t and a spatial parameter x. So for a fixed time t, the solution $u(t,x)$ is a function of x, and the state space of the process is an infinite dimensional function space.

Nonetheless, hitting questions have been studied for certain SPDE, see [2, 3, 4, 7, 11] among others. These papers deal with the stochastic heat and wave equations either with no drift or with well behaved drift.

As for SPDE analogues of the Bessel process, the only results known to the authors are in Mueller [8] and Mueller and Pardoux [9]. Here we assume that $u(t,x)$ is scalar valued, and as before $t>0$. But now we let x lie in the unit circle $[0,1]$ with endpoints identified. We also assume that $u(0,x)$ is continuous and strictly positive. Here and throughout the paper we write $\dot{W}(t,x)$ for two-parameter white noise. Suppose u satisfies the following SPDE.

$$\partial_t u(t,x) = \Delta u(t,x) + u^{-\alpha}(t,x) + g(u(t,x))\dot{W}(t,x)$$

where there exist constants $0<c_0<C_0<\infty$ for which $c_0 \le g(u) \le C_0$ for all values of u. Let τ be the first time at which u hits 0, and let $\tau=\infty$ if u does not hit 0. Then $\mathbf{P}(\tau<\infty)>0$ if $\alpha<3$, see [8] Corollary 1.1. Also, $\mathbf{P}(\tau<\infty)=0$ if $\alpha>3$, see Theorem 1 of [MP99].

The situation for vector-valued solutions $u(t,x)$ of the stochastic heat equation is unclear. Indeed, the curve $x \to u(t,x)$ may wind around 0, and perhaps then u will contract to 0 in cases where it would ordinarily stay away from 0.

The purpose of this paper is to study hitting question for the stochastic wave equation with scalar solutions and with strong drift. As is well known, there are crucial differences between the heat and wave equations. For example, the heat equation satisfies a maximum principle while the wave equation does not. The same holds for the comparison principle, which states that if the stochastic heat equation has two solutions with the first solution initially larger than the second, then the first solution will almost surely remain larger than the second as time goes on. So while certain arguments from the heat equation case carry over, new ideas are required.

Here is the setup for our problem. Again, we let $t \ge 0$, and x lies in the circle

$$\mathbf{I} = [0,J]$$

with endpoints identified. We study scalar-valued solutions $u(t,x)$ to the following equation.

$$\begin{aligned}\partial_t^2 u(t,x) &= \Delta u(t,x) + u^{-\alpha}(t,x) + g(u(t,x))W(\dot{t},x) \\ u(0,x) &= u_0(x) \\ \partial_t u(0,x) &= u_1(x).\end{aligned} \tag{1.1}$$

As usual, u and our two-parameter white noise $\dot{W}$ depend on a random parameter ω which we suppress. As for x taking values in higher-dimensional spaces, it is well known that (1.1) is well-posed only in one spatial dimensions. Indeed, in two or more spatial dimensions we would expect that the solution u only exists as a generalized function, but then it is hard to give meaning to nonlinear terms such as $u^{-\alpha}$ or $g(u)$.

Next, we define the first time that u hits 0. Let

$$\tau_\infty = \inf\left\{t>0 : \inf_{0\le s<t}\inf_{x\in\mathbf{I}} u(t,x) = 0\right\}$$

and let $\tau_\infty = \infty$ if the set in the above definition is empty.

Before stating our main theorems, we give some assumptions.

Assumptions

(i) u_0 is Hölder continuous of order $1/2$ on $\mathbf{I}$.

(ii) There exist constants $0 < c_0 < C_0 < \infty$ such that $c_0 \le u_0(x) \le C_0$ for all $x \in \mathbf{I}$.

(iii) u_1 is Hölder continuous of order $1/2$ on $\mathbf{I}$ and hence bounded.

(iv) There exist constants $0 < c_g < C_g < \infty$ such that $c_g \le g(y) \le C_g$ for all $y \in \mathbf{R}$.

Here are our main theorems.

Theorem 1.1. *Suppose that $u(t,x)$ satisfies (1.1), and that the above assumptions hold. Then $\alpha > 3$ implies*

$$\mathbf{P}(\tau_\infty < \infty) = 0.$$

That is, u does not hit 0.

Theorem 1.2. *Suppose that $u(t,x)$ satisfies (1.1), and that above assumptions hold. Then $0 < \alpha < 1$ implies*

$$\mathbf{P}(\tau_\infty < \infty) > 0.$$

That is, u can hit 0.

Here is the plan for the paper. In Section 2 we give a rigorous formulation of (1.1); in particular, the solution is only defined up to the first time t that $u(t,x) = 0$ for some x, since $u^{-\alpha}(t,x)$ blows up there. The same is true for the stochastic heat equation discussed earlier. In Section 3 we prove Theorem 1.1, and in Section 4 we prove Theorem 1.2.

Note the gap between $\alpha < 1$ and $\alpha > 3$. Since there is no comparison principle for the wave equation, we cannot be certain that there exists a critical value α_0 such that u can hit 0 for $\alpha < \alpha_0$ but not for $\alpha > \alpha_0$. We strongly believe in the existence of such a critical value, but we leave the existence and identification of α_0 as an open problem.

2 Technicalities

2.1 Rigorous formulation of the wave SPDE

For the most part we follow Walsh [13] although we could also use the formulation found in Da Prato and Zazbczyk [1].

First we recall the definition the one-dimensional wave kernel on $x \in \mathbf{R}$.

$$S(t,x) = \frac{1}{2}\mathbf{1}(|x| \le t)$$

See [5] for this classical material. If we regard $S(t,x)$ as a Schwartz distribution, then for $t \ge 0$ we can write

$$\partial_t S(t,x) = \frac{1}{2}\delta(x-t) + \frac{1}{2}\delta(x+t).$$

From now on, we interpret such expressions as Schwartz distributions.

Now we switch to the circle $x \in \mathbf{I}$, as defined earlier. It is also a classical result that for $x \in \mathbf{I}$, the wave kernel $S_{\mathbf{I}}$ and its time derivative are given by

$$\begin{aligned} S_{\mathbf{I}}(t,x) &= \sum_{n\in\mathbf{Z}} S(t, x+nJ) \\ \partial_t S_{\mathbf{I}}(t,x) &= \frac{1}{2}\sum_{n\in\mathbf{Z}} \Big(\delta(nJ+x-t) + \delta(nJ+x+t)\Big). \end{aligned}$$

Again, we regard $\partial_t S_{\mathbf{I}}$ as a Schwartz distribution.

Let $w(t,x)$ be the solution of the linear deterministic wave equation on $x \in \mathbf{I}$, with the same initial data as u. That is,

$$\begin{aligned}\partial_t^2 w(t,x) &= \Delta w(t,x) \\ w(0,x) &= u_0(x) \\ \partial_t w(0,x) &= u_1(x)\end{aligned}$$

with periodic boundary conditions, so that

$$\begin{aligned} w(t,x) &= \int_0^J \Big(\partial_t S_{\mathbf{I}}(t,x-y)u_0(y) + S_{\mathbf{I}}(t,x-y)u_1(y)\Big)dy \\ &= \frac{1}{2}\int_0^J \Big(u_0(x-t-y) + u_0(x+t-y) + S_{\mathbf{I}}(t,x-y)u_1(y)\Big)dy \end{aligned}$$

where expressions such as $x-y$ and $x-t-y$ are interpreted using arithmetic modulo J. We note that by Assumptions (i) and (iii), we can conclude that $w(t,x)$ is Hölder continuous of order $1/2$ in (t,x) jointly.

Using Duhamel's principle, if $u^{-\alpha}$ and $g(u(s,y))\dot{W}$ were smooth, we could write

$$\begin{aligned} u(t,x) =& w(t,x) + \int_0^t\int_0^J S_{\mathbf{I}}(t-s,x-y)u(s,y)^{-\alpha}dyds \\ &+ \int_0^t\int_0^J S_{\mathbf{I}}(t-s,x-y)g(u(s,y))W(dyds). \end{aligned} \tag{2.1}$$

If $u^{-\alpha}$ had no singularities, we could use this *mild form* to give rigorous meaning to (1.1), where we define final double integral using Walsh's theory of martingale measures, see [13]. One could also use the Hilbert space theory given in Da Prato and Zabczyk [1].

To deal with the singularity of $u^{-\alpha}$, we use truncation and then take the limit as the truncation is removed. For $N = 1, 2, \dots$ define $u_N(t,x)$ as the solution of

$$\begin{aligned} u_N(t,x) =& w(t,x) + \int_0^t\int_0^J S_{\mathbf{I}}(t-s,x-y)\Big[u_N(s,y)\vee(1/N)\Big]^{-\alpha}dyds \\ &+ \int_0^t\int_0^J S_{\mathbf{I}}(t-s,x-y)g(u_N(s,y))W(dyds). \end{aligned} \tag{2.2}$$

Here $a \vee b = \max(a,b)$. Note that if $\alpha > 0$, then $[u \vee (1/N)]^{-\alpha}$ is a Lipschitz function of u. It is well known that SPDE such as (2.2) with Lipschitz coefficients have unique strong solutions valid for all time, see [13], Chapter III. It follows for each $N = 1, 2, \dots$ that (2.2) has a unique strong solution u_N valid for all $t \geq 0, x \in [0,J]$.

Now let

$$\tau_N = \inf\Big\{t > 0 : \inf_{x\in[0,J]} u_N(t,x) \leq 1/N\Big\}.$$

From the definition, we see that almost surely

$$u_{N_1}(t,x) = u_{N_2}(t,x)$$

for all $t \in [0, \tau_{N_1}\wedge\tau_{N_2})$ and $x \in \mathbf{I}$. Here $a \wedge b = \min(a,b)$. It also follows that $\tau_1 \leq \tau_2 \leq \cdots$ and so we can almost surely define

$$\tau = \sup_N \tau_N.$$

We allow the possibility that $\tau = \infty$. Note that this definition of τ is consistent with the definition given in the introduction.

So, for $t < \tau$ and $x \in \mathbf{I}$, we can define

$$u(t,x) = \lim_{N\to\infty} u_N(t,x)$$

since for $t < \tau$ and $x \in \mathbf{I}$ the sequence $u_1(t,x), u_2(t,x), \ldots$ does not vary with N after a finite number of terms. It follows that $u(t,x)$ satisfies (2.1) for $0 \le t < \tau$.

Finally, we define $u(t,x)$ for all times t by defining

$$u(t,x) = \mathbf{\Delta}$$

for $t \ge \tau$. Here $\mathbf{\Delta}$ is a cemetary state.

2.2 Multi-parameter Girsanov theorem

The proof of Theorem 1.2 is based on Girsanov's theorem for two-parameter white noise. This approach was used earlier in Mueller and Pardoux [9] for the stochastic heat equation, but we need to do some work to adapt the argument to the stochastic wave equation. Girsanov's theorem will allow us to remove the drift from our equation (1.1), at least up to time τ. If this Girsanov transformation gives us an absolutely continuous change of probability measure, then we only need to verify that the stochastic wave equation (1.1) without the drift has a positive probability of hitting 0.

Assume that our white noise $\dot{W}(t,x)$ and hence also $u(t,x)$ is defined on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$. As in Walsh [13], we define $\dot{W}(t,x)$ in terms of a random set function $W(A,\omega)$ on measurable sets $A \subset [0,\infty) \times \mathbf{I}$. Let $(\mathcal{F}_t)_{t\ge 0}$ be the filtration defined by

$$\mathcal{F}_t = \sigma(W(A) : A \subset [0,t] \times \mathbf{I}).$$

Nualart and Pardoux [10] give the following version of Girsanov's theorem.

Theorem 2.1. *Let $T > 0$ be a given constant, and define the probability measure $\mathbf{P}_T$ to be $\mathbf{P}$ restricted to sets in $\mathcal{F}_T$. Suppose that W is a space-time white noise random measure on $[0,T] \times \mathbf{R}$ with respect to $\mathbf{P}_T$, and that $h(t,x)$ is a predictable process such that the exponential process*

$$\mathcal{E}_h(t) = \exp\left(\int_0^t \int_{\mathbf{R}} h(s,y) W(dyds) - \frac{1}{2}\int_0^t \int_{\mathbf{R}} h(s,y)^2 dyds\right)$$

is a martingale for $t \in [0,T]$. Then the measure

$$\tilde{W}(dxdt) = W(dxdt) - h(t,x)\, dxdt \tag{2.3}$$

is a space-time white noise random measure on $[0,T] \times \mathbf{R}$ with respect to the probability measure $\mathbf{Q}_T$, where $\mathbf{Q}_T$ and $\mathbf{P}_T$ are mutually absolutely continuous and

$$d\mathbf{Q}_T = \mathcal{E}_h(T)\, d\mathbf{P}_T. \tag{2.4}$$

We recall Novikov's sufficient condition for $\mathcal{E}_h(t)$ to be a martingale.

Proposition 2.2. *Let $h(t,x)$ be a predictable process with respect to the filtration $(\mathcal{F}_t)_{t\in[0,T]}$. If*

$$\mathbf{E}\left[\exp\left(\frac{1}{2}\int_0^T \int_{\mathbf{R}} h(s,y)^2 dyds\right)\right] < \infty \tag{2.5}$$

then $\mathcal{E}_h(t)$ is a uniformly integrable $\mathcal{F}_t$-martingale for $0 \le t \le T$.

2.3 Hölder continuity of the stochastic convolution

For an almost surely bounded predictable process $\rho(t,x)$, we define the stochastic convolution as follows.

$$N_\rho(t,x) = \int_0^t \int_0^J S_{\mathbf{I}}(t-s, x-y)\rho(s,y)W(dyds).$$

Note that the double integral in (2.1) is equal to $N_{g(u)}(t,x)$ for $t < \tau$. We conveniently define $g(\boldsymbol{\Delta}) = 0$, so that $N_{g(u)}(t,x)$ is defined for all time.

The proofs of both main theorems depend on the Hölder continuity of $N_{g(u)}(t,x)$. Although such results are common in the SPDE literature, unfortunately we could not find the exact result we needed. So for completeness, we state it here.

Theorem 2.3. *Let $\rho(t,x)$ be an almost surely bounded predictable process. For any $T > 0$ and $\beta < 1/2$, there exists a random variable Y with finite expectation, with $\mathbf{E}|Y|$ depending only on β and T, such that*

$$|N_\rho(t+h, x+k) - N_\rho(t,x)| \leq Y\left(h^\beta + k^\beta\right) \tag{2.6}$$

almost surely for all h, k where $t, t+h \in [0,T]$.

We will prove Theorem 2.3 in the appendix.

3 Proof of Theorem 1.1

3.1 Outline and preliminaries

We write the mild solution to (1.1) in the following form:

$$u(t,x) = V_u(t,x) + D_u(t,x) + N_u(t,x) \tag{3.1}$$

where

$$\begin{aligned} V_u(t,x) &= \frac{1}{2}\big(u_0(x+t) + u_0(x-t)\big) + \int_0^J u_1(y) S_{\mathbf{I}}(t, x-y)dy \\ D_u(t,x) &= \int_0^t \int_0^J u(s,y)^{-\alpha} S_{\mathbf{I}}(t-s, x-y)dyds \\ N_u(t,x) &= \int_0^t \int_0^J g(u(s,y)) S_{\mathbf{I}}(t-s, x-y)W(dyds). \end{aligned}$$

We will prove Theorem 1.1 by contradiction. First, we assume that $\tau < \infty$ with positive probability. Then, on the sample paths where this is the case (i.e., all $u(\omega)$ such that $\tau(\omega) < \infty$), we go backwards in time from where u hits zero. The upward drift term $D_u(t,x)$ will then push downwards, since we are going backwards in time. We show that this downward push must overwhelm the modulus of continuity of the $N_u(t,x)$ term, implying the existence of another time $\tau_1 < \tau$ such that u hits zero at τ_1. However, this contradicts the minimality of τ, thus proving the theorem.

3.2 A regularity lemma

Let $\mathbf{A} = \{\tau < \infty\}$. By assumption, $\mathbf{P}(\mathbf{A}) > 0$. We then show the following:

Lemma 3.1. *On the event $\mathbf{A}$, $V_u(t,x) + N_u(t,x)$ is almost surely β-Hölder continuous on $[0,\tau) \times [0,J]$ for any $\beta < 1/2$. The Hölder constant is a random variable depending only on β and ω.*

Proof. Let $\beta < 1/2$ be given. Then by (2.6) we know that $N_u(t,x)$ is almost surely β-Hölder continuous on $[0,\tau) \times [0,J]$, with random Hölder constant Y depending only on β and τ. Since u_1 is continuous on $\mathbf{I}$, the Riemann integral

$$\int_0^J u_1(y) S_{\mathbf{I}}(t, x-y)dy = \int_{x-t}^{x+t} u_1(y)dy$$

is jointly differentiable (and thus β-Hölder continuous) on $(t,x) \in [0,\tau) \times [0,J]$ as well. Finally, from assumption, u_0 is β-Hölder continuous on $[0,J]$, so it follows that $\frac{1}{2}(u_0(x+t) + u_0(x-t))$ is continuous as well.

Thus $V_u(t,x) + N_u(t,x)$ is almost surely β-Hölder continuous on $[0,\tau) \times [0,J]$. As the Hölder constant of V_u depends only on u_0 and u_1, the Hölder constant of $V_u + N_u$ is a random variable depending only on β. □

3.3 The backwards light cone

Given $(t,x) \in \mathbf{R}_+ \times \mathbf{R}$, define the *backwards light cone* as

$$\mathbf{L}(t,x) = \{(s,y) : |x-y| < t-s\}.$$

Note that the light cone cannot include points (s,y) for which $s > t$. It follows that $D_u(t,x)$ can be rewritten as

$$\begin{aligned} D_u(t,x) &= \int_0^t \int_0^J u(s,y)^{-\alpha} S_{\mathbf{I}}(t-s, x-y)dyds \\ &= \int_0^t \int_{\mathbb{R}} u\left(s, y^*\right)^{-\alpha} S(t-s, x-y)dyds \\ &= \iint_{\mathbf{L}(t,x)} u\left(s, y^*\right)^{-\alpha} dyds \end{aligned} \tag{3.2}$$

where

$$y^* = y \mod J. \tag{3.3}$$

and $y^* \in [0,J]$.

Lemma 3.2. *Let $(t,x) \in [0,\tau) \times [0,J]$. Then for any $(s,y) \in \mathbf{L}(t,x)$, we have*

$$D_u(s,y) - D_u(t,x) < 0.$$

Proof. Since $u(s,y) > 0$ on $[0,\tau)$, using (3.2) the result follows from the fact that $\mathbf{L}(s,y) \subsetneq \mathbf{L}(t,x)$. □

3.4 Theorem 1.1, conclusion

Since $\alpha > 3$ by assumption, define $\epsilon \in (0,1/2)$ sufficiently small such that

$$\frac{3-\alpha}{2} + \epsilon(\alpha+1) < 0.$$

Using Lemma 3.1, on the event $\mathbf{A}$ we define Y to be a (random) $1/2 - \epsilon$ Hölder constant of $V_u(t,x) + N_u(t,x)$, depending only on ϵ. By our choice of ϵ, the exponent of R in the expression

$$\frac{\pi Y^{-1-\alpha}}{2^{\alpha+2}} R^{\frac{3-\alpha}{2} + \epsilon(\alpha+1)}$$

is negative. Hence, on $\mathbf{A}$ we can pick a sufficiently small random $R > 0$, depending on ϵ and Y, such that both

$$\frac{\pi Y^{-1-\alpha}}{2^{\alpha+2}} R^{\frac{3-\alpha}{2} + \epsilon(\alpha+1)} > 1 \quad \text{and} \quad R < \frac{\tau}{2}. \tag{3.4}$$

Finally, on $\mathbf{A}$ we pick a random $\delta > 0$ sufficiently small such that both

$$\delta < \min\left(\inf_{x\in[0,J]} u_0(x), YR^{\frac{1}{2}-\epsilon}\right) \tag{3.5}$$

and

$$\tau_\delta = \inf\left\{t > 0 : \inf_{x\in[0,J]} u(t,x) < \delta\right\} > \frac{\tau}{2}, \tag{3.6}$$

which is possible since since $u(t,x)$ is continuous in t for $t < \tau$. Here, τ_δ need not be a stopping time. Note that τ_δ is the first time that $u(t,x)$ reaches δ, and that by continuity of $u(t,x)$ in x, there exists some $x_\delta \in [0,J]$ such that $u(\tau_\delta, x_\delta) = \delta$. We define the differences

$$\begin{aligned}\Delta V(t,x) &= V_u(t,x) - V_u(\tau_\delta, x_\delta)\\ \Delta D(t,x) &= D_u(t,x) - D_u(\tau_\delta, x_\delta)\\ \Delta N(t,x) &= N_u(t,x) - N_u(\tau_\delta, x_\delta)\end{aligned}$$

and for all $(t,x) \in \mathbf{L}(\tau_\delta, x_\delta)$, we decompose

$$\begin{aligned}u(t,x) &= u(t,x) - u(\tau_\delta, x_\delta) + \delta\\ &= \Delta V(t,x) + \Delta D(t,x) + \Delta N(t,x) + \delta.\end{aligned} \tag{3.7}$$

We recall that by construction,

$$\Delta V(t,x) + \Delta N(t,x) < Y\left|(t,x) - (\tau_\delta, x_\delta)\right|^{1/2-\epsilon} \tag{3.8}$$

almost surely on $\mathbf{A}$ with $\mathbf{E}[Y;\mathbf{A}] < \infty$. From Lemma 3.2, we find that

$$\Delta D(t,x) < 0$$

almost surely. Hence, for all $(t,x) \in \mathbf{L}(\tau_\delta, x_\delta)$ we obtain the bound

$$\begin{aligned}u(t,x) &= \Delta V(t,x) + \Delta D(t,x) + \Delta N(t,x) + \delta\\ &< \Delta V(t,x) + \Delta N(t,x) + \delta\\ &< Y\left|(t,x) - (\tau_\delta, x_\delta)\right|^{1/2-\epsilon} + \delta\end{aligned} \tag{3.9}$$

almost surely on $\mathbf{A}$. We define the sector

$$B_R = \left\{(t,x) \in \mathbf{L}(\tau_\delta, x_\delta) : |(\tau_\delta, x_\delta) - (t,x)| \leq R\right\},$$

noting from (3.4) and (3.6) that $t > 0$ on B_R. We then denote the curved part of the boundary of B_R by

$$\partial B_R = \left\{(t,x) \in B_R : |(\tau_\delta, x_\delta) - (t,x)| = R\right\}.$$

Then for all $(t,x) \in \partial B_R$, using (3.2), (3.9), and (3.5) we find that

$$\begin{aligned}\Delta D(t,x) &= -\iint_{\mathbf{L}(\tau_\delta,x_\delta)\setminus\mathbf{L}(t,x)} u(s,y^*)^{-\alpha}\,dyds\\ &\leq -\iint_{B_R} u(s,y^*)^{-\alpha}\,dyds\\ &\leq -|B_R|\left(YR^{\frac{1}{2}-\epsilon} + \delta\right)^{-\alpha}\\ &< -|B_R|\left(2YR^{\frac{1}{2}-\epsilon}\right)^{-\alpha}\\ &= -\frac{\pi R^2}{2^{\alpha+2}} Y^{-\alpha} R^{-\alpha\left(\frac{1}{2}-\epsilon\right)}\\ &= -\frac{\pi Y^{-\alpha}}{2^{\alpha+2}} R^{2-\left(\frac{1}{2}-\epsilon\right)\alpha}\end{aligned} \tag{3.10}$$

on the event $\mathbf{A}$. Recall that on ∂B_R, $|(t,x)-(\tau_\delta,x_\delta)|=R$. Hence from (3.7), (3.8), and (3.10) we find that for all $(t,x)\in\partial B_R$,

$$\begin{aligned} u(t,x) &< YR^{\frac{1}{2}-\epsilon}-\frac{\pi Y^{-\alpha}}{2^{\alpha+2}}R^{2-\left(\frac{1}{2}-\epsilon\right)\alpha} \\ &= YR^{\frac{1}{2}-\epsilon}\left(1-\frac{\pi Y^{-1-\alpha}}{2^{\alpha+2}}R^{2-\left(\frac{1}{2}-\epsilon\right)\alpha-\left(\frac{1}{2}-\epsilon\right)}\right) \\ &= YR^{\frac{1}{2}-\epsilon}\left(1-\frac{\pi Y^{-1-\alpha}}{2^{\alpha+2}}R^{\frac{3-\alpha}{2}+\epsilon(\alpha+1)}\right) \end{aligned} \tag{3.11}$$

almost surely on $\mathbf{A}$. From (3.4) and (3.11) it then follows that $u(t,x)<0$ for all $(t,x)\in\partial B_R$, almost surely on $\mathbf{A}$.

Since $\mathbf{P}(\mathbf{A})>0$ by assumption, the event that $u(t,x)<0$ for all $(t,x)\in\partial B_R$ occurs with positive probability. However, since $R>0$, we know that $t<\tau_\delta<\tau$ for all $(t,x)\in\partial B_R$, which is a contradiction, since τ is defined to be the first hitting time for $u(t,x)\le 0$. Hence we conclude that $\mathbf{P}(\mathbf{A})=0$.

This finishes the proof of theorem 1.1.

4 Proof of Theorem 1.2

4.1 Equation without the drift

Now we use Proposition 2.2 to prove Theorem 1.2. Consider the stochastic wave equation with initial conditions identical to (1.1) but without drift:

$$\begin{aligned} \partial_t^2 v(t,x) &= \Delta v(t,x)+g(v(t,x))\dot{W}(t,x) \\ v(0,x) &= u_0(x) \\ \partial_t v(0,x) &= u_1(x). \end{aligned} \tag{4.1}$$

Here $x\in[0,J]$, as before. Since there are no singular terms in (4.1), we can give this equation rigorous meaning using the mild form:

$$v(t,x)=w(t,x)+\int_0^t\int_0^J S_{\mathbf{I}}(t-s,x-y)g(v(s,y))W(dyds). \tag{4.2}$$

where $w(t,x)$ is as before, the solution to the deterministic wave equation.

First we verify that $v(t,x)$ can hit 0.

Lemma 4.1. *Suppose that $v(t,x)$ is a solution to (4.2). Then*

$$\mathbf{P}(v(t,x)=0 \text{ for some } t>0,\ x\in[0,J])>0.$$

Proof. Let $V(t)=\int_0^J v(t,x)dx$. By the almost sure continuity of $v(t,x)$ (see [13]) Chapter III, it suffices to show that

$$\mathbf{P}\big(V(t)<0\big)>0. \tag{4.3}$$

Since $\int_0^J S_{\mathbf{I}}(t,x-y)dy=t$ by the definition of the one-dimensional wave kernel, and since $\int_0^J \frac{1}{2}\big(u_0(x+t)+u_0(x-t)\big)dx=\int_0^J u_0(x)dx$,

$$V(t)=\int_0^J u_0(x)dx+t\int_0^J u_1(x)dx+\int_0^t\int_0^J (t-s)g(v(s,y))W(dyds).$$

Here we have used the stochastic Fubini theorem (see [13], Theorem 2.6) to change the order of integration in the double integral. Let us define $N_v(t)$ as the double integral:

$$N_v(t) = \int_0^t \int_0^J (t-s) g(v(s,y)) W(dyds).$$

The question would be easy if $g \equiv 1$, as $N_v(t)$ would be a Gaussian variable, with a positive probability of taking values below any desired level. Since this is not necessarily the case, we use another Girsanov transformation to bound $N_v(t)$ by a Gaussian process.

Fix $t > 0$. Choose K sufficiently large so that

$$\frac{c_g J K t^2}{2} - \int_0^J u_0(x)dx - t \int_0^J u_1(x)dx > 0. \tag{4.4}$$

Using Theorem 2.1, we define $\tilde{W}$ as a $\tilde{\mathbf{P}}$ white noise, where $\mathbf{P}$ and $\tilde{\mathbf{P}}$ are equivalent and

$$W(dyds) = \tilde{W}(dyds) - K dyds.$$

Decompose $N_v(t) = N_v^{(1)}(t) - N_v^{(2)}(t)$, where

$$\begin{aligned} N_v^{(1)}(t) &= \int_0^t \int_0^J (t-s) g(v(s,y)) \tilde{W}(dyds) \\ N_v^{(2)}(t) &= \int_0^t \int_0^J (t-s) g(v(s,y)) K dyds. \end{aligned}$$

Since $g(v(s,y))$ is bounded below by $c_g > 0$, we have:

$$N_v^{(2)}(t) \geq \frac{c_g J K t^2}{2}.$$

Hence to show (4.3), it suffices to prove that

$$\mathbf{P}\left(N_v^{(1)}(t) < \frac{c_g J K t^2}{2} - \int_0^J u_0(x)dx - t \int_0^J u_1(x)dx \right) > 0$$

and since $\mathbf{P}$ and $\tilde{\mathbf{P}}$ are equivalent, we can show instead that

$$\tilde{\mathbf{P}}\left(N_v^{(1)}(t) < \frac{c_g J K t^2}{2} - \int_0^J u_0(x)dx - t \int_0^J u_1(x)dx \right) > 0. \tag{4.5}$$

We define the process

$$M_t(r) = \int_0^r \int_0^J (t-s) g(v(s,y)) \tilde{W}(dyds).$$

Since g is bounded, $M_t(r)$ is an $\mathcal{F}_r$-martingale in r, for $r \leq t$. Hence, from Theorem V.1.6 in Revuz and Yor [12], there exists a one-dimensional standard Brownian motion B such that $M_t(r) = B(\tau(r))$, where the time change $\tau(r)$ is given by the predictable process:

$$\begin{aligned} \tau(r) &= \int_0^r \int_0^J (t-s)^2 g^2(v(s,y)) dyds \\ &\leq C_g^2 \int_0^r \int_0^J (t-s)^2 dyds \\ &= \frac{C_g^2 J}{3} t^3 - \frac{C_g^2 J}{3} (t-r)^3. \end{aligned}$$

Then let

$$L = \frac{C_g^2 J}{3} t^3$$

so we have $\tau(t) \leq L$. Using this, we find that:

$$N_v^{(1)}(t) = M_t(t) = B\left(\tau(t)\right) \leq \sup_{0\leq q\leq L} B(q).$$

Due to (4.4), we can use the reflection principle to find that

$$\begin{aligned}
&\tilde{\mathbf{P}}\left(\sup_{0\leq q\leq L} B(q) \geq \frac{c_g J K t^2}{2} - \int_0^J u_0(x)dx - t\int_0^J u_1(x)dx\right) \\
&\leq 2\tilde{\mathbf{P}}\left(B(L) \geq \frac{c_g J K t^2}{2} - \int_0^J u_0(x)dx - t\int_0^J u_1(x)dx\right) \\
&< 1 \qquad\qquad \text{(since } B(L)\sim\mathcal{N}(0,L))
\end{aligned}$$

from which (4.5) follows, and the proof of Lemma 4.1 is complete. □

4.2 Removing the drift term

To finish the proof of Theorem 1.2, it suffices to show that up to the first time τ that u and v hit 0, these two processes induce equivalent probability measures on the canonical paths consisting of continuous functions $f(t,x)$ on $[0,\tau(f)]\times[0,J]$.

Given a (possibly random) function $f:[0,\infty)\times[0,J]\to\mathbf{R}$, define the hitting times

$$\begin{aligned}
\tau^{(f)} &= \inf\left\{t>0 : \inf_{x\in[0,J]} f(t,x)\leq 0\right\} \\
\alpha_m^{(f)} &= \inf\left\{t>0 : \int_0^t\int_0^J f(s,x)^{-2\alpha}dxds > m\right\}
\end{aligned}$$

and for a constant $T>0$, let

$$T_m(f) = \tau^{(f)}\wedge\alpha_m^{(f)}\wedge T.$$

Then define the truncated function $f^{T_m(f)}$ by:

$$f^{T_m(f)}(t,x) = f(t,x)\mathbf{1}_{\{t\leq T_m(f)\}}.$$

Let $\mathbf{P}_u^{T_m(u)}$, $\mathbf{P}_v^{T_m(v)}$ be the measures on path space $\mathcal{C}\left([0,\infty)\times[0,J],\mathbf{R}\right)$ induced by $u^{T_m(u)}(t,x)$, $v^{T_m(v)}(t,x)$ respectively, and let

$$h(r) = \begin{cases}\frac{r^{-\alpha}}{g(r)} & \text{if } r\neq 0\\ 0 & \text{if } r=0.\end{cases} \tag{4.6}$$

We then obtain the following Girsanov transformation:

Lemma 4.2. *For each $m\in\mathbf{N}$, the measures $\mathbf{P}_u^{T_m(u)}$ and $\mathbf{P}_v^{T_m(v)}$ are equivalent, with*

$$\frac{d\mathbf{P}_u^{T_m(u)}}{d\mathbf{P}_v^{T_m(v)}} = \exp\left(\int_0^{T_m(v)}\int_0^J h\left(v(t,x)\right)W(dxdt) - \frac{1}{2}\int_0^{T_m(v)}\int_0^J h\left(v(t,x)\right)^2 dxdt\right).$$

Proof. First, we note that $h(v^{T_m(v)}(t,x))$ satisfies the Novikov condition given in (2.5). Then, define the probability measure $\mathbf{Q}^{T_m(v)}$ by the derivative

$$\begin{aligned}
\frac{d\mathbf{Q}^{T_m(v)}}{d\mathbf{P}_v^{T_m(v)}} &= \exp\left(\int_0^T\int_0^J h\left(v^{T_m(v)}(t,x)\right)W(dxdt) - \frac{1}{2}\int_0^T\int_0^J h\left(v^{T_m(v)}(t,x)\right)^2 dxdt\right) \\
&= \exp\left(\int_0^{T_m(v)}\int_0^J h\left(v(t,x)\right)W(dxdt) - \frac{1}{2}\int_0^{T_m(v)}\int_0^J h\left(v(t,x)\right)^2 dxdt\right).
\end{aligned}$$

Then, from Theorem 2.1, it follows that

$$\tilde{W}(dxdt) = W(dxdt) - h\left(v^{T_m(v)}(t,x)\right) dxdt$$

is a space-time white noise random measure under $\mathbf{Q}^{T_m(v)}$. Note that $\mathbf{Q}^{T_m(v)}$ is the measure on $\mathcal{C}([0,\infty)\times[0,J],\mathbf{R})$ induced by $f^{T_m(f)}(t,x)$ where $f(t,x)$ satisfies

$$\begin{aligned} f(t,x) &= \frac{1}{2}\left(u_0(x+t)+u_0(x-t)\right) + \int_0^J u_1(y) S_{\mathbf{I}}(t,x-y)dy \\ &\quad + \int_0^t \int_0^J g(f(s,y)) S_{\mathbf{I}}(t-s,x-y) W(dyds) \\ &= \frac{1}{2}\left(u_0(x+t)+u_0(x-t)\right) + \int_0^J u_1(y) S_{\mathbf{I}}(t,x-y)dy \\ &\quad + \int_0^t \int_0^J f(s,y)^{-\alpha} S_{\mathbf{I}}(t-s,x-y) dxdt \\ &\quad + \int_0^t \int_0^J g(f(s,y)) S_{\mathbf{I}}(t-s,x-y) \tilde{W}(dyds). \end{aligned}$$

where the last term is a Walsh integral with respect to the underlying measure $\mathbf{Q}^{T_m(v)}$. However, these are just the paths of $u^{T_m(u)}(t,x)$, so the measure $\mathbf{Q}^{T_m(v)} = \mathbf{P}_u^{T_m(u)}$. Then Lemma 4.2 follows. □

Now we wish to apply Lemma 4.2 with h as defined in (4.6). This depends on the finiteness of $\alpha_m(f)$ for some m. Thus 1.2 follows from the following lemma, which we prove by using the regularity of the stochastic wave equation.

Lemma 4.3. *For any constant $T > 0$,*

$$\int_0^{\tau^{(v)}\wedge T} \int_0^J v(t,x)^{-2\alpha} dxdt < \infty \tag{4.7}$$

almost surely.

4.3 Proof of Lemma 4.3

For this entire section, we let $v(t,x)$ be as given in (4.2).

4.3.1 A rectangular grid

For each $K > 0$ define the event:

$$\mathbf{A}(K) := \left\{ \sup_{(t,x)\in[0,T]\times[0,J]} v(t,x) \le K \right\}. \tag{4.8}$$

Since $(t,x) \mapsto v(t,x)$ is almost surely continuous, the above supremum is almost surely finite, so

$$\lim_{K\to\infty} \mathbf{P}\left\{\mathbf{A}(K)\right\} = 1.$$

We split the interval $(0,K]$ into dyadic subintervals

$$(0,K] = \bigcup_{n=0}^{\infty} \left(2^{-n-1}K, 2^{-n}K\right] \tag{4.9}$$

and observe that on the event $\mathbf{A}(K)$,

$$\begin{aligned}\int_0^{\tau^{(v)}\wedge T}&\int_0^J v(t,x)^{-2\alpha}dxdt\\&=\sum_{n=0}^{\infty}\left[\int_0^{\tau^{(v)}\wedge T}\int_0^J v(t,x)^{-2\alpha}1_{\{2^{-n-1}K<v(t,x)\le 2^{-n}K\}}dxdt\right]\\&\le\sum_{n=0}^{\infty}\left[2^{2\alpha(n+1)}K^{-2\alpha}\int_0^{\tau^{(v)}\wedge T}\int_0^J 1_{\{2^{-n-1}K<v(t,x)\le 2^{-n}K\}}dxdt\right]\\&=\sum_{n=0}^{\infty}\Big[2^{2\alpha(n+1)}K^{-2\alpha}\\&\qquad\times\mu\left(\left\{(t,x)\in\left[0,\tau^{(v)}\wedge T\right]\times[0,J]:2^{-n-1}K<v(t,x)\le 2^{-n}K\right\}\right)\Big]\end{aligned}\tag{4.10}$$

where μ denotes Lebesgue measure.

Define a constant $\epsilon>0$ such that

$$0<2\epsilon<1-\alpha$$

and for each $n\in\mathbf{N}$, consider the rectangle

$$D_n=\{(t,x)\in[0,\lambda_n]\times[0,2\lambda_n]\}\tag{4.11}$$

where

$$\lambda_n=2^{-(1-2\epsilon)n}.\tag{4.12}$$

As far as the optimality of this choice of λ_n, the real issue is why the same factor applies to both t and x. This is because the stochastic wave equation with white noise has the same regularity in both t and x, namely it is Hölder $1/2-\varepsilon$. So we do not believe that the result for $\alpha<1$ can be improved by better choosing λ_n.

Next, for each $(t,x)\in D_n$, define the grids of points:

$$\begin{aligned}\Gamma_n(t,x)&=\Big[[0,T]\times[0,J]\Big]\bigcap\left[\bigcup_{k,\ell\in\mathbb{N}}\Big(t+k\lambda_n,x+2\ell\lambda_n\Big)\right]\\\overline{\Gamma}_n(t,x)&=\left[\left[0,\tau^{(v)}\right]\times[0,J]\right]\bigcap\Gamma_n(t,x).\end{aligned}$$

Let $\#$ denote the number of points in a set, and define the strip

$$J_n=\left\{(t,x)\in[0,\lambda_n]\times[0,J]:2^{-n-1}K<v(t,x)\le 2^{-n}K\right\}.$$

Then we have

$$\begin{aligned}\mu&\left(\left\{(t,x)\in\left[0,\tau^{(v)}\wedge T\right]\times[0,J]:2^{-n-1}K<v(t,x)\le 2^{-n}K\right\}\right)\\&\qquad\le\iint_{D_n}\#\left\{(s,y)\in\overline{\Gamma}_n(t,x):v(s,y)\le 2^{-n}K\right\}dxdt+\mu\left(J_n\right).\end{aligned}\tag{4.13}$$

Since $v(t,x)$ is continuous on $[0,T]\times[0,J]$ and $\inf_{x\in[0,J]}u_0(x)>0$, we have that $\mu(J_n)=0$ for sufficiently large random n. Hence,

$$\sum_{n=0}^{\infty}\left[2^{2\alpha(n+1)}K^{-2\alpha}\mu\left(J_n\right)\right]<\infty\tag{4.14}$$

almost surely. We now place a bound on

$$\#\left\{(s,y)\in\overline{\Gamma}_n(t,x):v(s,y)\le 2^{-n}K\right\}$$

in the upcoming lemmas.

4.3.2 The shifted equation

Let (t,x) be an arbitrary point in D_n, as defined in (4.11), and let θ be the time shift operator, defined by $\theta_s W(dxdt) = W(dxd(t+s))$.

Then for given $(s,y) \in \Gamma_n(t,x)$, define

$$s_n^- = s - \lambda_n.$$

Now, we take the approach of considering W as a cylindrical Wiener process, as described in [1]. Furthermore, by Theorem 9.15 on page 256 of Da Prato and Zabczyk [1], there is a version of our solution Φ_t which is a strong Markov process with respect to the Brownian filtration $(\mathcal{F}_t)_{t\geq 0}$.

Using the strong Markov property of solutions, we restart the equation at time s_n^-:

$$\begin{aligned} v(s,y) &= \frac{1}{2}\left(v\left(s_n^-, y+\lambda_n\right) + v\left(s_n^-, y-\lambda_n\right)\right) \\ &\quad + \int_0^J S_{\mathbf{I}}\left(\lambda_n, y-z\right) \frac{\partial v}{\partial t}\left(s_n^-, z\right) dz \\ &\quad + \int_0^{\lambda_n}\int_0^J S_{\mathbf{I}}\left(\lambda_n - r, y-z\right) g\left(v\left(s_n^- + r, z\right)\right) \theta_{s_n^-} W(dzdr). \end{aligned} \tag{4.15}$$

Here, $\frac{\partial v}{\partial t}$ is regarded as a Schwartz distribution.

We analyze (4.15) term by term. Decompose

$$v(s,y) = V_n(s,y) + N_n(s,y) + E_n(s,y)$$

where

$$\begin{aligned} V_n(s,y) &= \frac{1}{2}\left(v\left(s_n^-, y+\lambda_n\right) + v\left(s_n^-, y-\lambda_n\right)\right) + \int_0^J S_{\mathbf{I}}\left(\lambda_n, y-z\right) \frac{\partial v}{\partial t}\left(s_n^-, z\right) dz \\ N_n(s,y) &= \int_0^{\lambda_n}\int_{\{|z-y|\leq\lambda_n\}} S_{\mathbf{I}}\left(\lambda_n - r, y-z\right) g(v(s_n^-, y))\theta_{s_n^-} W(dzdr) \\ E_n(s,y) &= -N_n(s,y) + \int_0^{\lambda_n}\int_0^J S_{\mathbf{I}}(\lambda_n - r, y-z) g(v(s_n^- + r, z))\theta_{s_n^-} W(dzdr). \end{aligned}$$

More specifically,

- First, we take V_n to be the first two terms, representing the contribution to $v(s,y)$ from the shifted initial conditions (both position and velocity).
- Next, we realize the stochastic term as the sum of a conditionally Gaussian term and an error term. The former is the stochastic term integrated over the light cone contained in the square $\{(s,y) + D_n\}$, with the diffusion coefficient g frozen at $v(s_n^-, y)$. We call this term the noise term, N_n.
- Finally, as mentioned above the error term E_n is the difference between the stochastic term of $v(s,y)$ minus the noise term defined above.

As alluded to above, the noise term can be rewritten as:

$$\begin{aligned} N_n(s,y) &= g(v(s_n^-, y))\int_0^{\lambda_n}\int_{\{|z-y|\leq\lambda_n\}} S_{\mathbf{I}}\left(\lambda_n - r, y-z\right)\theta_{s_n^-} W(dzdr) \\ &= g(v(s_n^-, y)) c_n Z \end{aligned} \tag{4.16}$$

where c_n^2 is the quadratic variation of the above double integral and Z is a standard normal random variable. Moreover, for sufficiently small λ_n relative to J, we have:

$$\begin{aligned} c_n^2 &= \int_0^{\lambda_n} \int_{\{|z-y|\le\lambda_n\}} S_{\mathbf{I}}^2(r, y-z) dz dr \\ &= \frac{\lambda_n^2}{4} = 2^{-2(1-2\epsilon)n-2}. \end{aligned} \tag{4.17}$$

4.3.3 A regularity lemma

Now, we find bounds for E_n and N_n by using Hölder continuity of v. Define the events

$$\begin{aligned} \mathbf{B}_n &= \left\{ \sup_{(s,y)\in\Gamma_n(t,x)} |E_n(s,y)| \le 2^{-n} \right\} \\ \mathbf{C}_n &= \left\{ \sup_{(s,y)\in\Gamma_n(t,x)} |N_n(s,y)| \le 2^{-(1-3\epsilon)n} \right\} \\ \mathbf{A}_n &= \mathbf{A}(K) \cap \mathbf{B}_n \cap \mathbf{C}_n. \end{aligned}$$

Then we assert the following:

Lemma 4.4. $\sum_{n=1}^{\infty} \mathbf{P}(\mathbf{B}_n^c) < \infty$ *and* $\sum_{n=1}^{\infty} \mathbf{P}(\mathbf{C}_n^c) < \infty$.

To prove this lemma, we first establish a bound on the error term E_n. Recall its definition:

$$\begin{aligned} E_n &= -N_n(s,y) + \int_0^{\lambda_n} \int_0^J S_{\mathbf{I}}(\lambda_n - r, y-z)\, g(v(s_n^- + r, z))\theta_{s_n^-} W(dzdr) \\ &= -\int_0^{\lambda_n} \int_{\{|z-y|\le\lambda_n\}} S_{\mathbf{I}}(\lambda_n - r, y-z)\, g(v(s_n^-, y))\theta_{s_n^-} W(dzdr) \\ &\quad + \int_0^{\lambda_n} \int_0^J S_{\mathbf{I}}(\lambda_n - r, y-z)\, g(v(s_n^- + r, z))\theta_{s_n^-} W(dzdr) \end{aligned}$$

Note that in the integrals above, $S_{\mathbf{I}}(\lambda_n - r, y-z) = 0$ outside of the light cone $|z-y| \le \lambda_n$. Thus, we restrict the domain of integration of z:

$$\begin{aligned} E_n &= -\int_0^{\lambda_n} \int_{\{|z-y|\le\lambda_n\}} S_{\mathbf{I}}(\lambda_n - r, y-z)\, g(v(s_n^-, y))\theta_{s_n^-} W(dzdr) \\ &\quad + \int_0^{\lambda_n} \int_{\{|z-y|\le\lambda_n\}} S_{\mathbf{I}}(\lambda_n - r, y-z) g(v(s_n^- + r, z))\theta_{s_n^-} W(dzdr) \\ &= \int_0^{\lambda_n} \int_{\{|z-y|\le\lambda_n\}} S_{\mathbf{I}}(\lambda_n - r, y-z) \\ &\qquad \times \left[g(v(s_n^- + r, z)) - g(v(s_n^-, y)) \right] \theta_{s_n^-} W(dzdr) \end{aligned}$$

We define the rectangle

$$\Delta_n(s,y) = \{ r \in \mathbb{R}_+, z \in [0,J] : |r-s| \le \lambda_n, |z-y| \le \lambda_n \}$$

and let p be a positive integer. Then it follows that

$$\begin{aligned} \mathbf{E}\left[E_n^{2p}\right] = \mathbf{E}\bigg(\int_0^{\lambda_n} \int_{\{|z-y|\le\lambda_n\}} & S_{\mathbf{I}}(\lambda_n - r, y-z) \\ & \times \left[g(v(s_n^- + r, z)) - g(v(s_n^-, y)) \right] \theta_{s_n^-} W(dzdr) \bigg)^{2p} \end{aligned}$$

and since the integrand above is continuous in λ_n, we can use the Burkholder-Davis-Gundy inequality to obtain:

$$\begin{aligned}\mathbf{E}\left[E_n^{2p}\right] &\lesssim_p \mathbf{E}\Bigg(\int_{s_n^-}^{s_n^-+\lambda_n}\int_{\{|z-y|\leq\lambda_n\}} S_{\mathbf{I}}^2\left(\lambda_n-\left(r-s_n^-\right),y-z\right) \\ &\qquad\times\left[g(v(r,z))-g(v(s_n^-,y))\right]^2 dzdr\Bigg)^p \\ &\lesssim_p \mathbf{E}\Bigg(\int_{s_n^-}^{s}\int_{\{|z-y|\leq\lambda_n\}} S_{\mathbf{I}}^2\left(s-r,y-z\right) \\ &\qquad\times\left[g(v(r,z))-g(v(s_n^-,y))\right]^2 dzdr\Bigg)^p.\end{aligned}$$

As usual, the notation $a(x)\lesssim_p b(x)$ means that $a(x)\leq C_p b(x)$.

Since g is Lipschitz and since $(s_n^-,y)\in\Delta_n(s,y)$,

$$\begin{aligned}|g(v(r,z))-g(v(s_n^-,y))|^2 &\leq L_g^2\left|v(r,z)-v\left(s_n^-,y\right)\right|^2 \\ &\leq L_g^2\left(2\left|v(r,z)-v(s,y)\right|^2+2\left|v(s,y)-v\left(s_n^-,y\right)\right|^2\right) \\ &\leq 4L_g^2\sup_{(r,z)\in\Delta_n(s,y)}\left|v(r,z)-v(s,y)\right|^2\end{aligned}$$

With this bound, we get:

$$\begin{aligned}\mathbf{E}\left[E_n^{2p}\right] &\lesssim \mathbf{E}\left[\sup_{(r,z)\in\Delta_n(s,y)}\left[\left|v(r,z)-v(s,y)\right|^{2p}\right]\right] \\ &\quad\times\left(\int_{s_n^-}^{s}\int_{\{|z-y|\leq\lambda_n\}} S_{\mathbf{I}}^2\left(s-r,y-z\right)dzdr\right)^p.\end{aligned}\tag{4.18}$$

Recall that $v(s,y)$ is almost surely β-Hölder continuous for any $\beta<\frac{1}{2}$. Setting $\beta=\frac{1}{2}-\frac{1}{2p}$, we obtain

$$\begin{aligned}&\mathbf{E}\left[\sup_{(r,z)\in\Delta_n(s,y)}\left[\left|v(r,z)-v(s,y)\right|^{2p}\right]\right] \\ &\quad\lesssim_{g,p}\sup_{(r,z)\in\Delta_n(s,y)}\left(|r-s|^{\frac{1}{2}-\frac{1}{2p}}+|z-y|^{\frac{1}{2}-\frac{1}{2p}}\right)^{2p} \\ &\quad\lesssim_{g,p}\left(\lambda_n^{\frac{1}{2}-\frac{1}{2p}}\right)^{2p}=\lambda_n^{p-1}=2^{-(1-2\epsilon)n(p-1)}.\end{aligned}\tag{4.19}$$

Recalling (4.12), we then bound the integral:

$$\begin{aligned}&\left(\int_0^{\lambda_n}\int_{\{|z-y|\leq\lambda_n\}} S_{\mathbf{I}}^2(r,y-z)dzdr\right)^p \\ &\qquad=\left(\frac{1}{4}\int_0^{\lambda_n}\int_{\{|z-y|\leq\lambda_n\}}\mathbf{1}_{\{|y-z|<r\}}dzdr\right)^p \\ &\qquad\lesssim\lambda_n^{2p} \\ &\qquad\lesssim 2^{-2(1-2\epsilon)np}\end{aligned}\tag{4.20}$$

so by (4.18), (4.19), and (4.20), we obtain a bound on the error term:

$$\mathbf{E}\left[E_n^{2p}\right]\lesssim_{g,p}2^{-(1-2\epsilon)n(3p-1)}.\tag{4.21}$$

Proof of Lemma 4.4. Recalling that $\#\left\{\Gamma_n(t,x)\right\} \lesssim \lambda_n^{-2} = 2^{(2-4\epsilon)n}$, we find:

$$\begin{aligned}\mathbf{P}\left(\mathbf{B}_n^c\right) &= \mathbf{P}\left\{\sup_{(s,y)\in\Gamma_n(t,x)} |E_n(s,y)| > 2^{-n}\right\} \\ &\le \sum_{(s,y)\in\Gamma_n(t,x)} \mathbf{P}\left\{|E_n(s,y)| > 2^{-n}\right\} \\ &\lesssim 2^{(2-4\epsilon)n}\,\mathbf{P}\left\{|E_n(s,y)| > 2^{-n}\right\}.\end{aligned}$$

By Markov's inequality, we can continue as follows,

$$\mathbf{P}\left(\mathbf{B}_n^c\right) \lesssim 2^{(2-4\epsilon)n+2np}\,\mathbf{E}\left[E_n^{2p}\right]$$

after which we use (4.21) to obtain the bound:

$$\mathbf{P}\left(\mathbf{B}_n^c\right) \lesssim 2^{(3-6\epsilon)n-(1-6\epsilon)np}.$$

Thus, the summation $\sum_{n=1}^{\infty} \mathbf{P}\left(\mathbf{B}_n^c\right)$ converges when $p > \frac{3-6\epsilon}{1-6\epsilon}$.

With a similar decomposition, we obtain:

$$\mathbf{P}\left(\mathbf{C}_n^c\right) \lesssim 2^{(4-4\epsilon)n}\,\mathbf{P}\left\{|N_n(s,y)| > 2^{-(1-3\epsilon)n}\right\}.$$

Recalling (4.16) and (4.17), this implies:

$$\begin{aligned}\mathbf{P}\left(\mathbf{C}_n^c\right) &\lesssim 2^{(4-4\epsilon)n}\,\mathbf{P}\left\{\left|g\left(v\left(s_n^-,y\right)\right)Z\right| > 2^{(1-2\epsilon)n}2^{-(1-3\epsilon)n}\right\} \\ &\lesssim 2^{(4-4\epsilon)n}\,\mathbf{P}\left\{|Z| > C_g^{-1}2^{\epsilon n}\right\}\end{aligned}$$

where Z is a standard normal random variable. Now, we use a standard tail estimate for the normal (often called the Chernoff bound) to conclude

$$\mathbf{P}\left\{|Z| > C_g^{-1}\,2^{\epsilon n}\right\} \le 2\exp\left(-C_g^{-2}\,2^{2\epsilon n-1}\right)$$

and it follows that $\sum_{n=1}^{\infty} \mathbf{P}\left(\mathbf{C}_n^c\right)$ converges. $\square$

4.3.4 A counting lemma

Lemma 4.5. *For each $K > 0$ there exists a constant c_K such that for every $n \in \mathbf{N}$ and $(t,x) \in D_n$,*

$$\mathbf{E}\left[\#\left\{(s,y) \in \overline{\Gamma}_n(t,x) : v(s,y) \le 2^{-n}K\right\};\mathbf{A}_n\right] \le c_K.$$

The proof of Lemma 4.5 will require several preliminary steps.

We fix (t,x) and order the points in $\Gamma_n(t,x)$ lexicographically, calling the ith point (s_i,y_i) for some $i \in \mathcal{I}(t,x) = \{1,2,\ldots,\#\{\Gamma_n(t,x)\}\}$ - i.e., if $i < j$ then $s_i \le s_j$ and if $s_i = s_j$, then $x \le x_i < x_j \mod J$. For given (t,x), we define the set $\Delta^n(s,y)$ as follows:

$$\Delta^n(s,y) = \begin{cases} [0,s_n^-] \times \mathbf{I} & y = x \\ ([0,s_n^-] \times \mathbf{I}) \bigcup ([s_n^-,s] \times [x-\lambda_n, y-\lambda_n]) & y \ne x \end{cases}$$

where the interval $[x-\lambda_n, y-\lambda_n]$ on $\mathbf{I}$ is taken modulo J, wrapping around whenever $x-\lambda_n > y-\lambda_n$. (Note that this is not the same as the previously defined $\Delta_n(s,y)$).

Let $\mathcal{F}_i^n$ be the σ-algebra generated by $\dot{W}$ in the set $\Delta^n(s_i,y_i)$. Then $V_n(s_i,y_i)$ is $\mathcal{F}_i^n$-measurable for all $i \in \mathbb{N}$. Recall from (4.16) that

$$N_n(s_i,y_i) = g\left(v\left((s_i)_n^-, y_i\right)\right)c_n Z_i \tag{4.22}$$

where $c_n = 2^{-(1-2\epsilon)n-1}$ and $Z_i \sim \mathcal{N}(0,1)$ is $\mathcal{F}^n_{i+1}$-measurable but independent of $\mathcal{F}^n_i$.

Let $\mathbf{P}^n_i$ denote the conditional probability with respect to $\mathcal{F}^n_i$ and let $\delta > 1$ be a constant depending only on K. Define

$$\overline{v}_n(s,y) = V_n(s,y) + N_n(s,y).$$

We now prove the following lemma:

Lemma 4.6. *There exists $d_K > 0$ such that for all $i \in \mathcal{I}(t,x)$,*

$$\mathbf{P}^n_i\Big[\overline{v}_n\left(s_i,y_i\right) \le -2^{-n} \,\Big|\, \overline{v}_n\left(s_i,y_i\right) \le 2^{-n}(K+1)\Big] \ge d_K \tag{4.23}$$

almost surely on the event $\{V_n(s_i,y_i) \le \delta 2^{-(1-\epsilon)n}\}$.

Proof. From the definition of conditional probability, the left hand side of (4.23) is:

$$\begin{aligned} H &:= \mathbf{P}^n_i\Big[\overline{v}_n\left(s_i,y_i\right) \le -2^{-n} \,\Big|\, \overline{v}_n\left(s_i,y_i\right) \le 2^{-n}(K+1)\Big] \\ &= \frac{\mathbf{P}^n_i\left\{\overline{v}_n\left(s_i,y_i\right) \le -2^{-n}\right\}}{\mathbf{P}^n_i\left\{\overline{v}_n\left(s_i,y_i\right) \le 2^{-n}(K+1)\right\}} \\ &= \frac{\mathbf{P}^n_i\left\{N_n\left(s_i,y_i\right) \le -2^{-n} - V_n\left(s_i,y_i\right)\right\}}{\mathbf{P}^n_i\left\{N_n\left(s_i,y_i\right) \le 2^{-n}(K+1) - V_n\left(s_i,y_i\right)\right\}}. \end{aligned}$$

Using (4.22), we find that

$$H = \frac{\mathbf{P}^n_i\left\{g\left(v\left(\left(s_i\right)^-_n, y_i\right)\right) c_n Z_i \le -2^{-n} - V_n\left(s_i,y_i\right)\right\}}{\mathbf{P}^n_i\left\{g\left(v\left(\left(s_i\right)^-_n, y_i\right)\right) c_n Z_i \le 2^{-n}(K+1) - V_n\left(s_i,y_i\right)\right\}}$$

and using (4.17), we find that

$$H = \frac{\mathbf{P}^n_i\left\{g\left(v\left(\left(s_i\right)^-_n, y_i\right)\right) Z_i \le -2^{1-2\epsilon n} - 2^{1+(1-2\epsilon)n} V_n\left(s_i,y_i\right)\right\}}{\mathbf{P}^n_i\left\{g\left(v\left(\left(s_i\right)^-_n, y_i\right)\right) Z_i \le 2^{1-2\epsilon n}(K+1) - 2^{1+(1-2\epsilon)n} V_n\left(s_i,y_i\right)\right\}}.$$

Then define $\rho_{n,i} = 2g(v((s_i)^-_n, y_i))^{-1}$. Note that $\rho_{n,i}$ is almost surely bounded above by $2C_g$ and below by $2c_g^{-1} > 0$, both uniformly in n and i. Plugging this into the above equation, we find:

$$H = \frac{\mathbf{P}^n_i\left\{Z_i \le -\rho_{n,i}\left(2^{-2\epsilon n} + 2^{(1-2\epsilon)n} V_n\left(s_i,y_i\right)\right)\right\}}{\mathbf{P}^n_i\left\{Z_i \le -\rho_{n,i}\left(-2^{-2\epsilon n}(K+1) + 2^{(1-2\epsilon)n} V_n\left(s_i,y_i\right)\right)\right\}}. \tag{4.24}$$

Now we examine H in two cases. The first case is on the event

$$\left\{-2^{-2\epsilon n}(K+1) + 2^{(1-2\epsilon)n} V_n\left(s_i,y_i\right) \le 0\right\}. \tag{4.25}$$

Since the denominator in (4.24) is less than or equal to 1, we can bound H below by its numerator:

$$H \ge \mathbf{P}^n_i\left\{Z_i \le -\rho_{n,i}\left(2^{-2\epsilon n} + 2^{(1-2\epsilon)n} V_n\left(s_i,y_i\right)\right)\right\}.$$

Using the decomposition

$$\begin{aligned} &2^{-2\epsilon n} + 2^{(1-2\epsilon)n} V_n\left(s_i,y_i\right) \\ &= \left(2^{-2\epsilon n}(K+2)\right) + \left(-2^{-2\epsilon n}(K+1) + 2^{(1-2\epsilon)n} V_n\left(s_i,y_i\right)\right) \\ &\le \left(2^{-2\epsilon n}(K+2)\right) \end{aligned}$$

and the assumption in (4.25), we find

$$H \geq \mathbf{P}_i^n \left\{ Z_i \leq -\rho_{n,i} 2^{-2\epsilon n}(K+2) \right\}. \tag{4.26}$$

Since $\rho_{n,i} \leq 2C_g$ for all n, we note that for all $K > 0$, $\rho_{n,i}2^{-2\epsilon n}(K+2) \to 0$ as $n \to \infty$. So for sufficiently large n (depending on K), $H \geq 1/3$. Hence in the case given by (4.26), Lemma 4.6 follows.

The second case is on the event

$$\left\{ -2^{-2\epsilon n}(K+1) + 2^{(1-2\epsilon)n} V_n\left(s_i, y_i\right) > 0 \right\}. \tag{4.27}$$

Here, we use the following inequality from Lemma 8 of Mueller and Pardoux [9]: For $a, b > 0$ and Z a standard normal random variable,

$$\frac{\mathbf{P}\{Z > a\}}{\mathbf{P}\{Z > a+b\}} \leq \frac{1}{2\mathbf{P}\{Z > 1\}} \vee \left(1 + \frac{\sqrt{e}}{1-e^{-1}}(a+b)be^{ab+\frac{b^2}{2}} \right).$$

Let $a = \rho_{n,i}(-2^{-2\epsilon n}(K+1) + 2^{(1-2\epsilon)n}V_n(s_i, y_i))$ and $b = \rho_{n,i}2^{-2\epsilon n}(K+2)$. Recalling that from our given conditions, $V_n(s_i, y_i) \leq \delta 2^{-(1-\epsilon)n}$ almost surely, we find that:

$$\begin{aligned}
(a+b)b &= \rho_{n,i}^2 \left(2^{-2\epsilon n} + 2^{(1-2\epsilon)n} V_n\left(s_i, y_i\right) \right) (K+2) 2^{-2\epsilon n} \\
&\leq \rho_{n,i}^2 \left(2^{-2\epsilon n} + 2^{-2\epsilon n}\delta \right) (K+2)2^{-2\epsilon n} \\
&= \rho_{n,i}^2 (\delta+1)(K+2)2^{-4\epsilon n} \\
&\leq \rho_{n,i}^2(\delta+1)(K+2)
\end{aligned}$$

and

$$\begin{aligned}
\left(a + \frac{b}{2} \right) b &= \rho_{n,i}^2 \left(-(0.5K)2^{-2\epsilon n} + 2^{(1-2\epsilon)n} V_n\left(s_i, y_i\right) \right) (K+2)2^{-2\epsilon n} \\
&\leq \rho_{n,i}^2 \left(-(0.5K)2^{-2\epsilon n} + 2^{-2\epsilon n}\delta \right) (K+2)2^{-2\epsilon n} \\
&= \rho_{n,i}^2(\delta - 0.5K)(K+2)2^{-4\epsilon n} \\
&\leq \rho_{n,i}^2(\delta - 0.5)(K+2) \qquad \text{(recalling that } K > 1\text{)}
\end{aligned}$$

almost surely. Using these results with (4.24), we find:

$$\begin{aligned}
H &= \frac{\mathbf{P}\{Z \leq -(a+b)\}}{\mathbf{P}\{Z \leq -a\}} \\
&= \frac{\mathbf{P}\{Z > a+b\}}{\mathbf{P}\{Z > a\}} \\
&\geq (2\mathbf{P}\{Z > 1\}) \wedge \left(1 + \frac{\sqrt{e}}{1-e^{-1}}(a+b)be^{ab+\frac{b^2}{2}} \right)^{-1} \\
&\geq (2\mathbf{P}\{Z > 1\}) \wedge \left(1 + \frac{\sqrt{e}}{1-e^{-1}}\rho_{n,i}^2(\delta+1)(K+2)e^{\rho_{n,i}^2(\delta-0.5)(K+2)} \right)^{-1}.
\end{aligned}$$

Since $\rho_{n,i}$ is almost surely uniformly bounded away from 0 in n and i, there exists $c_{g,\delta} > 0$ such that $\rho_{n,i}^2 \geq 4c_{g,\delta}$. So:

$$\begin{aligned}
&\left(1 + \frac{\sqrt{e}}{1-e^{-1}}\rho_{n,i}^2(\delta+1)(K+2)e^{\rho_{n,i}^2(\delta-0.5)(K+2)} \right)^{-1} \\
&\qquad\qquad \geq c_{g,\delta} \left(c_{g,\delta}^{-1} + \frac{\sqrt{e}}{1-e^{-1}}(\delta+1)(K+2)e^{c_{g,\delta}(\delta-0.5)(K+2)} \right)^{-1}
\end{aligned}$$

and since δ depends only on K, the right hand side above is bounded below by some $\gamma_{K,g} > 0$. Then H is bounded above by

$$d_K = 2\mathbf{P}\{Z > 1\} \wedge \gamma_{K,g} > 0$$

in the case given by (4.27) as well. Hence Lemma 4.6 follows in both cases. □

Proof of Lemma 4.5. Define

$$\xi_n = \mathbf{E}\Big[\#\Big\{(s,y) \in \overline{\Gamma}_n(t,x) : v(s,y) \leq 2^{-n}K\Big\}; \mathbf{A}_n\Big].$$

From the definitions of $\mathbf{A}_n$ and $\overline{\Gamma}$, it follows that ξ_n is bounded by:

$$\begin{aligned}\xi_n \leq \mathbf{E}\Big[\#\Big\{(s,y) \in \overline{\Gamma}_n(t,x) :&0 < v(s,y) \leq 2^{-n}K,\\ &|E_n(s,y)| \leq 2^{-n}, |N_n(s,y)| \leq 2^{-2(1-3\epsilon)n}\Big\}\Big].\end{aligned}$$

Recall that by definition, $\overline{v}_n = V_n(s,y) + N_n(s,y) = v(s,y) - E_n(s,y)$. Thus, we obtain the bound

$$\begin{aligned}\xi_n \leq \mathbf{E}\Big[\#\Big\{(s,y) \in \overline{\Gamma}_n(t,x) :&-2^{-n} < \overline{v}_n(s,y) \leq 2^{-n}(K+1),\\ &|E_n(s,y)| \leq 2^{-n}, |N_n(s,y)| \leq 2^{-2(1-3\epsilon)n}\Big\}\Big].\end{aligned}$$

Note that if $V_n(s,y) + N_n(s,y) \leq 2^{-n}(K+1)$ and $|N_n(s,y)| \leq 2^{-2(1-3\epsilon)n}$, then for some $\delta > 1$ depending on K, ϵ we have $V_n(s,y) \leq \delta 2^{-(1-\epsilon)n}$. So we can write:

$$\xi_n \leq \mathbf{E}\Big[\#\Big\{(s,y) \in \overline{\Gamma}_n(t,x) : -2^{-n} < \overline{v}_n(s,y) \leq 2^{-n}(K+1), V_n(s,y) \leq \delta 2^{-(1-\epsilon)n}\Big\}\Big].$$

Let $\{\sigma_n(k)\}_{k\in\mathbb{N}}$ be the sequence of indices $i \in \mathcal{I}$, in lexicographical order, such that both $\overline{v}_n(s_i,y_i) \leq 2^{-n}(K+1)$ and $V_n(s_i,y_i) \leq \delta 2^{-(1-\epsilon)n}$.

Out of the set of points on Γ_n such that $\overline{v}_n \leq 2^{-n}(K+1)$, one looks at the points where $\overline{v}_n < -2^{-n}$, which would force v to be negative. Thus we define the event

$$\mathbf{D}_k = \left\{\overline{v}_n\left(s_{\sigma_n(k)}, y_{\sigma_n(k)}\right) \leq -2^{-n}\right\}$$

and for $k \in \mathbb{N}$, we define the indicator random variable

$$I_k = 1_{\mathbf{D}_k}.$$

From Lemma 4.6, it is clear that

$$\mathbf{P}\left\{I_1 = 1\right\} \geq d_K$$

and moreover, since V_i and N_i are $\mathcal{F}^n_j$-measurable for all $i < j$, we can also use Lemma 4.6 to find that

$$\mathbf{P}\Big[I_k = 1 \;\Big|\; I_1, \ldots, I_{k-1}\Big] \geq d_K$$

for $k > 1$. Finally, let

$$\overline{\sigma}_n = \inf\left\{k; I_k = 1\right\}.$$

Since $v(s,y) \geq 0$ for all $(s,y) \in \overline{\Gamma}_n$, it follows that $\xi_n \leq \mathbf{E}\overline{\sigma}_n$.

Note that the I_k's are not independent. We couple $\{I_k\}$ with a sequence of independent random variables $\{Y_k\}$ as follows: Let $\{U_k\}_{k\geq 1}$ be a sequence of mutually independent random variables that are globally independent of the I_k's and have uniform law on $[0,1]$. Then define

$$Y_k = \begin{cases} 0 & \text{if } I_k = 0 \text{ or } U_k > d_K/\mathbf{P}\Big[I_k = 1 \,\Big|\, I_1, \dots, I_{k-1}\Big] \\ 1 & \text{if } I_k = 1 \text{ and } U_k \le d_K/\mathbf{P}\Big[I_k = 1 \,\Big|\, I_1, \dots, I_{k-1}\Big] \end{cases}$$

for $k \ge 1$. Then clearly,

$$Y_k \le I_k$$

and for $k > 1$,

$$\begin{aligned} \mathbf{P}\Big[Y_k = 1 \,\Big|\, Y_1, \dots, Y_{k-1}\Big] &= \mathbf{P}\Big[Y_k = 1 \,\Big|\, I_1, \dots, I_{k-1}\Big] \\ &= d_K \end{aligned}$$

so $\{Y_k\}$ is a sequence of i.i.d. random variables. Let $\tilde{\sigma} = \inf\{k; Y_k = 1\}$. Then

$$\begin{aligned} \overline{\sigma}_n &= 1\text{st } k \text{ such that } I_k = 1 \\ \tilde{\sigma} &= 1\text{st } k \text{ such that } Y_k = 1 \end{aligned}$$

and since $Y_k \le I_k$, it follows that $\overline{\sigma}_n \le \tilde{\sigma}$. So

$$\mathbf{E}\overline{\sigma}_n \le \mathbf{E}\tilde{\sigma} = d_K^{-1}. \qquad \square$$

4.4 Lemma 4.3, conclusion

Finally, we cite a measure-theoretic result related to the Borel-Cantelli Lemma:

Lemma 4.7. *Let $\{X_n\}$ be a sequence of $[0,\infty)$-valued random variables, and $\{\mathbf{F}_n\}$ be a sequence of events, such that both:*

$$\sum_{n=0}^{\infty} \mathbf{P}\left(\mathbf{F}_n^c\right) < \infty$$
$$\sum_{n=0}^{\infty} \mathbf{E}\left[X_n; \mathbf{F}_n\right] < \infty$$

Then $\sum_{n=0}^{\infty} X_n < \infty$ almost surely.

Proof. Let $\mathbf{F} = \{\sum_{n=0}^{\infty} X_n = +\infty\}$. Then on the event $\mathbf{F} \cap (\liminf \mathbf{F}_n)$, we have $\sum_{n=0}^{\infty} X_n \mathbf{1}_{\mathbf{F}_n} = +\infty$. So from the second condition, we get $\mathbf{P}(\mathbf{F} \cap \liminf \mathbf{F}_n) = 0$. However, from the first condition and Borel-Cantelli, we find:

$$\mathbf{P}\left(\limsup \mathbf{F}_n^c\right) = 0 \Rightarrow \mathbf{P}\left(\liminf \mathbf{F}_n\right) = 1$$

Implying that $\mathbf{P}(\mathbf{F}) = 0$, which is our desired result. $\square$

Proof of Lemma 4.3. From equations (4.10) and (4.13), we have:

$$\begin{aligned} &1_{A(K)} \int_0^{\tau^{(v)} \wedge T} \int_{\mathbf{I}} v(t,x)^{-2\alpha}\, dx dt \\ &\quad \le 1_{A(K)} \sum_{n=0}^{\infty} \Big[2^{2\alpha(n+1)} K^{-2\alpha} \\ &\qquad \times \mu\left(\left\{(t,x) \in \left[0, \tau^{(v)} \wedge T\right] \times \mathbf{I} : 2^{-n-1}K < v(t,x) \le 2^{-n}K\right\}\right)\Big]. \end{aligned}$$

Now consider the above expression $\mu(\cdots)$. First, we can bound this expression above by dropping the inequality $2^{-n-1}K < v(t,x)$, which enlarges the set under consideration. Secondly, we note that

$$\bigcup_{(t,x)\in D_n} \overline{\Gamma}_n(t,x)$$

covers the entire (t,x)-plane, because $\overline{\Gamma}_n(t,x)$ consists of the corners of rectangles which are translations of D_n, further translated by (t,x). So we can continue as follows.

$$\begin{aligned}
&1_{A(K)}\int_0^{\tau^{(v)}\wedge T}\int_{\mathbf{I}} v(t,x)^{-2\alpha}\,dxdt \\
&\quad\le 1_{A(K)}\sum_{n=0}^{\infty}\Big[2^{2\alpha(n+1)}K^{-2\alpha}\iint_{D_n}\#\left\{(s,y)\in\overline{\Gamma}_n(t,x): v(s,y)\le 2^{-n}K\right\}\,dxdt\Big] \\
&\qquad + 1_{A(K)}\sum_{n=0}^{\infty}\Big[2^{2\alpha(n+1)}K^{-2\alpha}\,\mu\left(J_n\right)\Big].
\end{aligned}$$

Now consider the summation of expectations

$$\begin{aligned}
&\sum_{n=0}^{\infty}\mathbf{E}\Big[1_{A(K)}\,2^{2\alpha(n+1)}K^{-2\alpha}\iint_{D_n}\#\left\{(s,y)\in\overline{\Gamma}_n(t,x): v(s,y)\le 2^{-n}K\right\}\,dxdt; B_n\cap C_n\Big] \\
&\quad=\sum_{n=0}^{\infty}\mathbf{E}\Big[2^{2\alpha(n+1)}K^{-2\alpha}\iint_{D_n}\#\left\{(s,y)\in\overline{\Gamma}_n(t,x): v(s,y)\le 2^{-n}K\right\}\,dxdt; A_n\Big]
\end{aligned}$$

Recalling that $\alpha < 1$, and that D_n was defined in (4.11), we note that:

$$\iint_{D_n} dxdt = 2^{-(2-4\epsilon)n+1}$$

so using Lemma 4.5, we find:

$$\begin{aligned}
&\sum_{n=0}^{\infty}\mathbf{E}\Big[2^{2\alpha(n+1)}K^{-2\alpha}\iint_{D_n}\#\left\{(s,y)\in\overline{\Gamma}_n(t,x): v(s,y)\le 2^{-n}K\right\}\,dxdt; A_n\Big] \\
&\quad=\sum_{n=0}^{\infty}2^{2\alpha(n+1)}K^{-2\alpha}\iint_{D_n}\mathbf{E}\left[\#\left\{(s,y)\in\overline{\Gamma}_n(t,x): v(s,y)\le 2^{-n}K\right\}; A_n\right]\,dxdt \\
&\quad\le\sum_{n=0}^{\infty}2^{2\alpha(n+1)}K^{-2\alpha}c_K\iint_{D_n}dxdt \\
&\quad=\sum_{n=0}^{\infty}c_K K^{-2\alpha}2^{2\alpha+1}2^{(2\alpha-2+4\epsilon)n}
\end{aligned}$$

and since $\alpha < 1-2\epsilon$, the summation converges. Thus from using Lemma 4.4, Lemma 4.7, and (4.14), we have

$$1_{A(K)}\int_0^{\tau^{(v)}\wedge T}\int_0^J v(t,x)^{-2\alpha}dxdt < \infty \tag{4.28}$$

almost surely. Observe that (4.8) and (4.28) imply (4.7) since

$$\begin{aligned}
\int_0^{\tau^{(v)}\wedge T} &\int_{\mathbf{I}} v(t,x)^{-2\alpha} dxdt \\
&= 1_{A(K)} \int_0^{\tau^{(v)}\wedge T} \int_{\mathbf{I}} v(t,x)^{-2\alpha} dxdt + 1_{A(K)^c} \int_0^{\tau^{(v)}\wedge T} \int_{\mathbf{I}} v(t,x)^{-2\alpha} dxdt \\
&\le 1_{A(K)} \int_0^{\tau^{(v)}\wedge T} \int_{\mathbf{I}} v(t,x)^{-2\alpha} dxdt + \int_0^{\tau^{(v)}\wedge T} \int_{\mathbf{I}} K^{-2\alpha} dxdt \\
&< \infty
\end{aligned}$$

almost surely. Indeed, on $A(K)^c$ one knows that $v(t,x) > K$ for $(t,x) \in [0,T]\times[0,J]$, so $v^{-2\alpha}(t,x) \le K^{-2\alpha}$ in this situation. □

A Proof of Theorem 2.3

Proof. For the remainder of this section we will simply write $N(t,x)$ instead of $N_\rho(t,x)$. Fix $T>0$ and consider the space and time differences, given respectively by $|N(t,x+k)-N(t,x)|$ and $|N(t+h,x)-N(t,x)|$. Without loss of generality, let $h,k\in[0,\frac{J}{2}]$.

A.1 Space difference

To bound $N(t,x+k)-N(t,x)$ we first write

$$\begin{aligned}
\Delta_k^p &= \mathbf{E}\big[\,|N(t,x+k)-N(t,x)|^p\,\big] \\
&= \mathbf{E}\left[\left|\int_0^t\int_0^J \rho(s,y)\Big(S_{\mathbf{I}}(t-s,x+k-y)-S_{\mathbf{I}}(t-s,x-y)\Big)W(dyds)\right|^p\right]
\end{aligned}$$

and note that for all r, the following stochastic integral is a martingale over t:

$$\int_0^t\int_0^J \rho(s,y)\Big(S_{\mathbf{I}}(r-s,x+k-y)-S_{\mathbf{I}}(r-s,x-y)\Big)W(dyds).$$

We fix $p>2$ and use Burkholder's inequality to find a constant C_p such that

$$\begin{aligned}
\mathbf{E}&\left[\left|\int_0^t\int_0^J \rho(s,y)\left(S_{\mathbf{I}}(r-s,x+k-y)-S_{\mathbf{I}}(r-s,x-y)\right)W(dyds)\right|^p\right] \\
&\qquad\le C_p\mathbf{E}\left[\left|\int_0^t\int_S |\rho(s,y)|^2\left|S_{\mathbf{I}}(r-s,x+k-y)-S_{\mathbf{I}}(r-s,x-y)\right|^2 dyds\right|^{p/2}\right]
\end{aligned}$$

for all r. As C_p does not depend on r, we can set $r=t$ to obtain

$$\Delta_k^p \le C_p\mathbf{E}\left[\left|\int_0^t\int_0^J |\rho(s,y)|^2\left|S_{\mathbf{I}}(t-s,x+k-y)-S_{\mathbf{I}}(t-s,x-y)\right|^2 dyds\right|^{p/2}\right].$$

Now we use Hölder's Inequality with exponents $\frac{p}{2}$ and $\frac{p}{p-2}$:

$$\begin{aligned}
\Delta_k^p \le& C_p\left(\mathbf{E}\int_0^t\int_0^J \rho(v(s,y)^p dyds\right) \\
&\times\left[\int_0^t\int_0^J \left|S_{\mathbf{I}}(t-s,x+k-y)-S_{\mathbf{I}}(t-s,x-y)\right|^{2p/(p-2)} dyds\right]^{p/2-1}.
\end{aligned}$$

Since ρ is almost surely bounded, the expectation $\mathbf{E}\int_0^t\int_S \rho(s,y)^p dyds$ is bounded by a constant depending on T and p. So we obtain:

$$\begin{aligned}
\Delta_k^p &\lesssim_{p,T} \left[\int_0^t\int_0^J |S_{\mathbf{I}}(t-s,x+k-y)-S_{\mathbf{I}}(t-s,x-y)|^{2p/(p-2)}\,dyds\right]^{p/2-1} \\
&= \left[\int_0^t\int_0^J |S_{\mathbf{I}}(s,x+k-y)-S_{\mathbf{I}}(s,x-y)|^{2p/(p-2)}\,dyds\right]^{p/2-1} \\
&= \left[\int_0^t\int_0^J \left|\sum_{m\in\mathbb{Z}}\frac{1}{2}1_{\{|x+k-y+mJ|<s\}}-\sum_{m\in\mathbb{Z}}\frac{1}{2}1_{\{|x-y+mJ|<s\}}\right|^{2p/(p-2)} dyds\right]^{p/2-1} \\
&= \frac{1}{2^p}\left[\int_0^t\int_0^J \left|\sum_{m\in\mathbb{Z}}\left(1_{\{|x+k-y+mJ|<s\}}-1_{\{|x-y+mJ|<s\}}\right)\right|^{2p/(p-2)} dyds\right]^{p/2-1}.
\end{aligned}$$

Here and in the next few estimates, the infinite sum over $m\in\mathbb{Z}$ is really finite, because the indicator functions will be zero for all but finitely many values of m.

We use the inequality $\left(\sum_{n=1}^N a_n\right)^p \lesssim_{p,N} \sum_{n=1}^N a_n^p$ to get:

$$\begin{aligned}
\Delta_k^p &\lesssim_{p,T} \frac{1}{2^p}\left[\int_0^t\int_0^J \sum_{m\in\mathbb{Z}}\left|1_{\{|x+k-y+mJ|<s\}}-1_{\{|x-y+mJ|<s\}}\right|^{2p/(p-2)} dyds\right]^{p/2-1} \\
&= \frac{1}{2^p}\left[\int_0^t\int_0^J \sum_{m\in\mathbb{Z}}\left|1_{\{|x+k-y+mJ|<s\}}-1_{\{|x-y+mJ|<s\}}\right| dyds\right]^{p/2-1} \\
&= \frac{1}{2^p}\left[\int_0^t\int_{\mathbb{R}} \left|1_{\{|x+k-y|<s\}}-1_{\{|x-y|<s\}}\right| dyds\right]^{p/2-1} \\
&\le \frac{1}{2^p}(2tk)^{p/2-1} \lesssim_{p,T} k^{p/2-1}.
\end{aligned}$$

A.2 Time difference

As in the last section, we can use Burkholder's inequality to find:

$$\begin{aligned}
&\mathbf{E}\left[|N(t+h,x)-N(t,x)|^p\right] \\
&\quad\le C_p\mathbf{E}\left[\left|\int_0^t\int_0^J |\rho(s,y)|^2\,|S_{\mathbf{I}}(t+h-s,x-y)-S_{\mathbf{I}}(t-s,x-y)|^2\,dyds\right|^{p/2}\right] \\
&\qquad\qquad + C_p\mathbf{E}\left[\left|\int_t^{t+h}\int_0^J |\rho(s,y)|^2\,|S_{\mathbf{I}}(t+h-s,x-y)|^2\,dyds\right|^{p/2}\right].
\end{aligned}$$

The first term is handled like the space difference:

$$\begin{aligned}
&\mathbf{E}\left[\left|\int_0^t\int_0^J |\rho(s,y)|^2\,|S_{\mathbf{I}}(t+h-s,x-y)-S_{\mathbf{I}}(t-s,x-y)|^2\,dyds\right|^{p/2}\right] \\
&\quad\lesssim_p \left(\mathbf{E}\int_0^t\int_0^J \rho(s,y)^p dyds\right) \\
&\qquad\times\left[\int_0^t\int_0^J |S_{\mathbf{I}}(t+h-s,x-y)-S_{\mathbf{I}}(t-s,x-y)|^{2p/(p-2)}\,dyds\right]^{p/2-1}.
\end{aligned}$$

Using the inequality $\left(\sum_{n=1}^N a_n\right)^p \lesssim_{p,N} \sum_{n=1}^N a_n^p$, we can continue the above inequality

$$\begin{aligned}
&\lesssim_{p,T} \left[\int_0^t \int_0^J |S_{\mathbf{I}}(s+h, x-y) - S_{\mathbf{I}}(s, x-y)|^{2p/(p-2)}\, dyds\right]^{p/2-1} \\
&= \left[\int_0^t \int_0^J \left|\sum_{m\in\mathbb{Z}} \frac{1}{2} 1_{\{|x-y+mJ|<s+h\}} - \sum_{m\in\mathbb{Z}} \frac{1}{2} 1_{\{|x-y+mJ|<s\}}\right|^{2p/(p-2)} dyds\right]^{p/2-1} \\
&\lesssim_p \frac{1}{2^p} \left[\int_0^t \int_{\mathbf{R}} \left|1_{\{|x-y|<s+h\}} - 1_{\{|x-y|<s\}}\right| dyds\right]^{p/2-1} \\
&\leq \frac{1}{2^p} (2th)^{p/2-1} \lesssim_{p,T} h^{p/2-1}.
\end{aligned}$$

For the second term, we start with Hölder's inequality again:

$$\begin{aligned}
&\mathbf{E}\left[\left|\int_t^{t+h} \int_0^J |\rho(s,y)|^2 \left|S_{\mathbf{I}}(t+h-s, x-y)\right|^2 dyds\right|^{p/2}\right] \\
&\leq C_p \left(\mathbf{E} \int_t^{t+h} \int_0^J \rho(s,y)^p dyds\right) \\
&\quad \times \left[\int_t^{t+h} \int_0^J |S_{\mathbf{I}}(t+h-s, x-y)|^{2p/(p-2)}\, dyds\right]^{p/2-1}.
\end{aligned}$$

Using the inequality $\left(\sum_{n=1}^N a_n\right)^p \lesssim_{p,N} \sum_{n=1}^N a_n^p$ one last time, we continue the inequality:

$$\begin{aligned}
&\lesssim_{p,T} \left[\int_t^{t+h} \int_0^J |G(t+h-s, x-y)|^{2p/(p-2)}\, dyds\right]^{p/2-1} \\
&= \left[\int_0^h \int_0^J |S_{\mathbf{I}}(s, x-y)|^{2p/(p-2)}\, dyds\right]^{p/2-1} \\
&= \left[\int_0^h \int_0^J \left|\frac{1}{2} 1_{\{|x-y|<s\}}\right|^{2p/(p-2)} dyds\right]^{p/2-1} \\
&= \frac{1}{2p} \left(h^2\right)^{p/2-1} \lesssim_p h^{p/2-1}.
\end{aligned}$$

A.3 Conclusion

Putting together the space and time differences, we obtain for h, k:

$$\begin{aligned}
\mathbf{E}\left[|N(t+h, x+k) - N(t,x)|^p\right] &\lesssim_{p,T} h^{p/2-1} + k^{p/2-1} \\
&\lesssim_{p,T} \left(\sqrt{h^2+k^2}\right)^{p/2-1}.
\end{aligned}$$

Finally, we recall Kolmogorov's continuity theorem for multiparameter processes [13], Corollary 1.2.

Theorem A.1 (Kolmogorov)**.** *Let R be a rectangle in $\mathbb{R}^n$ and $\{X_t, t \in R\}$ be a real-valued stochastic process. Suppose there exist a, b, K all positive such that for all $s, t \in R$*

$$\mathbf{E}\,|X_t - X_s|^a \leq K|t-s|^{n+b}.$$

Then,

1. *X has a continuous realization;*
2. *there exist a constant C depending only on a, b, n and a random variable Y such that with probability one,*

$$|X_t - X_s| \le Y|t-s|^{b/a} \qquad (\forall)\, s, t \in R$$

and $\mathbf{E}\left[Y^a\right] \le CK$;

3. *if $\mathbf{E}\left[|X_t|^a\right] < \infty$ for some $t \in R$ then*

$$\mathbf{E}\left[\sup_{t\in R} |X_t|^a\right] < \infty.$$

Setting $a = p$ and $b = p/2 - 3$, we obtain

$$\begin{aligned}|N(t+h, x+k) - N(t, x)| &\le Y\left(\sqrt{h^2+k^2}\right)^{b/a} \\ &\le Y\left(h^{1/2-3/p} + k^{1/2-3/p}\right)\end{aligned}$$

where $\mathbf{E}[Y] < \infty$ and depends only on p and T. Then since $\beta < 1/2$, we can set $p = \frac{6}{1-2\beta}$ and the conclusion follows. $\square$

References

[1] Giuseppe Da Prato and Jerzy Zabczyk, *Stochastic equations in infinite dimensions*, Encyclopedia of Mathematics and its Applications, vol. 44, Cambridge University Press, Cambridge, 1992. MR-1207136

[2] Robert C. Dalang, Davar Khoshnevisan, and Eulalia Nualart, *Hitting probabilities for systems of non-linear stochastic heat equations with additive noise*, ALEA Lat. Am. J. Probab. Math. Stat. **3** (2007), 231–271. MR-2365643

[3] Robert C. Dalang and Marta Sanz-Solé, *Criteria for hitting probabilities with applications to systems of stochastic wave equations*, Bernoulli **16** (2010), no. 4, 1343–1368. MR-2759182

[4] Robert C. Dalang and Marta Sanz-Solé, *Hitting probabilities for nonlinear systems of stochastic waves*, Mem. Amer. Math. Soc. **237** (2015), no. 1120, v+75. MR-3401290

[5] Lawrence C. Evans, *Partial differential equations*, second ed., Graduate Studies in Mathematics, vol. 19, American Mathematical Society, Providence, RI, 2010. MR-2597943

[6] Peter Mörters and Yuval Peres, *Brownian motion*, Cambridge Series in Statistical and Probabilistic Mathematics, vol. 30, Cambridge University Press, Cambridge, 2010, With an appendix by Oded Schramm and Wendelin Werner. MR-2604525

[7] C. Mueller and R. Tribe, *Hitting properties of a random string*, Electron. J. Probab. **7** (2002), no. 10, 29. MR-1902843

[8] Carl Mueller, *Long-time existence for signed solutions of the heat equation with a noise term*, Probab. Theory Related Fields **110** (1998), no. 1, 51–68. MR-1602036

[9] Carl Mueller and Etienne Pardoux, *The critical exponent for a stochastic PDE to hit zero*, Stochastic analysis, control, optimization and applications, Systems Control Found. Appl., Birkhäuser Boston, Boston, MA, 1999, pp. 325–338. MR-1702968

[10] D. Nualart and E. Pardoux, *Markov field properties of solutions of white noise driven quasi-linear parabolic PDEs*, Stochastics Stochastics Rep. **48** (1994), no. 1-2, 17–44. MR-1786190

[11] Eulalia Nualart and Frederi Viens, *The fractional stochastic heat equation on the circle: time regularity and potential theory*, Stochastic Process. Appl. **119** (2009), no. 5, 1505–1540. MR-2513117

[12] Daniel Revuz and Marc Yor, *Continuous martingales and Brownian motion*, third ed., Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 293, Springer-Verlag, Berlin, 1999. MR-1725357

[13] John B. Walsh, *An introduction to stochastic partial differential equations*, École d'été de probabilités de Saint-Flour, XIV—1984, Lecture Notes in Math., vol. 1180, Springer, Berlin, 1986, pp. 265–439. MR-876085