A forward-backward SDE from the 2D nonlinear stochastic heat equation

Alexander Dunlap* Yu Gu†

October 6, 2021

Abstract

We consider a nonlinear stochastic heat equation in spatial dimension $d = 2$, forced by a white-in-time multiplicative Gaussian noise with spatial correlation length $\varepsilon > 0$ but divided by a factor of $\sqrt{\log \varepsilon^{-1}}$. We impose a condition on the Lipschitz constant of the nonlinearity so that the problem is in the "weak noise" regime. We show that, as $\varepsilon \downarrow 0$, the one-point distribution of the solution converges, with the limit characterized in terms of the solution to a forward-backward stochastic differential equation (FBSDE). We also characterize the limiting multipoint statistics of the solution, when the points are chosen on appropriate scales, in similar terms. Our approach is new even for the linear case, in which the FBSDE can be solved explicitly and we recover results of Caravenna, Sun, and Zygouras (*Ann. Appl. Probab.* 27(5):3050–3112, 2017).

1 Introduction

Fix a Lipschitz function $\sigma : [0,\infty) \to [0,\infty)$ with $\sigma(0) = 0$. Define $\beta = \mathrm{Lip}(\sigma)$. We are interested in the following two-dimensional stochastic heat equation with colored noise of spatial correlation length $\varepsilon > 0$, started at constant initial condition $a \in \mathbf{R}_{\geq 0}$:

$$\mathrm{d}u_{\varepsilon,a}(t,x) = \frac{1}{2}\Delta u_{\varepsilon,a}(t,x)\mathrm{d}t + (\log \varepsilon^{-1})^{-\frac{1}{2}}\sigma(u_{\varepsilon,a}(t,x))\mathrm{d}W^{\varepsilon}(t,x), \quad t > 0, x \in \mathbf{R}^2; \tag{1.1}$$

$$u_{\varepsilon,a}(0,x) = a. \tag{1.2}$$

Here we define $W^{\varepsilon} = G_{\varepsilon^2/2} * W$, where $G_t(x) = \frac{1}{2\pi t}\mathrm{e}^{-|x|^2/(2t)}$ is the two-dimensional heat kernel, $\mathrm{d}W$ is a spacetime white noise, and $*$ denotes convolution in space. The choice of mollifier is not essential, and we restrict to this choice only to simplify some of the computations. The covariance operator of $\mathrm{d}W^{\varepsilon}$ is formally given by

$$\mathbf{E}\mathrm{d}W^{\varepsilon}(t,x)\mathrm{d}W^{\varepsilon}(t',x') = \delta(t-t')G_{\varepsilon^2}(x-x') = \delta(t-t')\tfrac{1}{\varepsilon^2}G_1(\tfrac{x-x'}{\varepsilon}). \tag{1.3}$$

For $\varepsilon > 0$, the well-posedness of the initial value problem (1.1)–(1.2) is well-known (see e.g. [40]), and we consider the mild formulation

$$u_{\varepsilon,a}(t,x) = a + \frac{1}{\sqrt{\log \varepsilon^{-1}}}\int_0^t \int G_{t-s}(x-y)\sigma(u_{\varepsilon,a}(s,y))\,\mathrm{d}W^{\varepsilon}(s,y). \tag{1.4}$$

*Department of Mathematics, Courant Institute of Mathematical Sciences, New York University, New York, NY 10012 USA. alexander.dunlap@cims.nyu.edu. Partially supported by the NSF Mathematical Sciences Postdoctoral Fellowship program under grant no. DMS-2002118, and by NSF grant no. DMS-1910023 and BSF grant no. 2014302 (both awarded to Lenya Ryzhik).

†Department of Mathematics, University of Maryland, College Park, MD 20742 USA. ygu7@umd.edu. Partially supported by the NSF through grant no. DMS-1907928 and CAREER-2042384.

General properties of solutions to the nonlinear stochastic heat equation have previously been studied in general spatial dimensions by many authors. We mention the non-exhaustive list of works [17, 18, 13, 12, 14].

We are interested in taking $\varepsilon \downarrow 0$ and identifying nontrivial limiting behavior for the solutions of (1.1)–(1.2). The linear problem, in which $\sigma(x) = \beta x$, is a particularly important special case. Here it is known that the attenuating factor $(\log \varepsilon^{-1})^{-\frac{1}{2}}$ in (1.1) is required, and that there is phase transition at $\beta = \sqrt{2\pi}$. The subcritical linear problem ($\beta < \sqrt{2\pi}$) was previously studied in [7] (which we will discuss in more detail shortly), while the critical linear problem ($\beta \approx \sqrt{2\pi}$) has been studied in [3, 8, 26, 10]. It is worth mentioning that the notion of "criticality" here is different from the one in [28, Section 8]. In the linear case, the equation is related by the Cole–Hopf transform to the two-dimensional KPZ equation, as considered in [11, 9, 24]. The linear problem also admits a Feynman–Kac formula [2] and thus a connection to directed polymers, with the solution to the SPDE interpreted as the partition function of directed polymers in random environment. The Feynman–Kac representation has proved to be very useful in analyzing properties of the solutions, but is not available in the nonlinear case. In [7], Caravenna, Sun, and Zygouras showed that if $\sigma(x) = \beta x$, $\beta \in (0, \sqrt{2\pi})$, then for any fixed $T > 0$ and $X \in \mathbf{R}^2$, $u_{\varepsilon,a}(T, X)$ converges in distribution as $\varepsilon \downarrow 0$ to a log-normal random variable. Their proof used the Feynman–Kac formula to connect the problem to directed polymers, and then worked to understand a polynomial chaos expansion in great detail.

The goal of the present paper is to study the nonlinear case in which many previously-used tools are not available. We will show in Theorem 1.2 below that if σ is β-Lipschitz, $\beta \in (0, \sqrt{2\pi})$, then $u_{\varepsilon,a}(T, X)$ converges in distribution as $\varepsilon \downarrow 0$. The limit depends on σ and is obtained through the solution of a forward-backward stochastic differential equation. Our method is also new in the linear case. In the nonlinear case, the limit does not seem to be log-normal in general.

Part of the reason we are interested in such a problem comes from the recent progress in proving the Edwards-Wilkinson limit of the KPZ equation [11, 9, 24, 37, 22, 33, 15] in $d \geq 2$. Most of these results rely on the Cole–Hopf transformation which, in some sense, linearizes the problem so that one can focus on studying the linear stochastic heat equation (as in [19, 42, 39, 27, 21, 33, 15]) and how its solution behaves after the logarithmic transformation. For general Hamilton–Jacobi type equations, this linearization does not exist and there are no results of this type. (See a conjecture in [30, p. 5] and some related directions for the anisotropic KPZ equation in [4, 5, 6].) We hope that working on the nonlinear stochastic heat equation can help bridge the difficulty and shed light on other nonlinear problems such as the Hamilton–Jacobi equation. A similar effort in $d \geq 3$ was carried out in [25]. The convergence to Edwards-Wilkinson equation in $d \geq 2$ is as random Schwartz distributions, which, in our case, corresponds to the convergence in distribution of the random variable

$$\sqrt{\log \varepsilon^{-1}} \int [u_{\varepsilon,a}(T, x) - a] g(x) \mathrm{d}x$$

for Schwartz test function g. The limiting marginal distributions of $u_{\varepsilon,a}$ play an important role in passing to the limit of the above random variable, which we will discuss in more detail below in Remark 8.1.

In order to state our main result (Theorem 1.2 below) precisely, we first have to define the limit object. Let $\{B(q)\}_{q \geq 0}$ be a 1D standard Brownian motion with the natural filtration $\{\mathcal{G}_q\}_{q \geq 0}$. We consider the following system of equations, satisfied by $\{\Xi_{a,Q}(\cdot)\}_{a,Q}$, with the parameters $a \geq 0$ and $Q \in [0, 2]$:

$$\mathrm{d}\Xi_{a,Q}(q) = J(Q - q, \Xi_{a,Q}(q)) \mathrm{d}B(q), \qquad q \in (0, Q]; \tag{1.5}$$

$$\Xi_{a,Q}(0) = a; \tag{1.6}$$

$$J(q, b) = \frac{1}{2\sqrt{\pi}} [\mathbf{E}\sigma^2(\Xi_{b,q}(q))]^{1/2}. \tag{1.7}$$

The parameter a plays the role of initial data, Q is the terminal time, and the above equation can be interpreted as follows: for the process started at a with the terminal time Q, to determine the diffusion coefficient at any time $q \in [0, Q]$, we run an independent process, starting from the current position $b = \Xi_{a,Q}(q)$

and with terminal time $Q-q$. The new process at time $Q-q$ is distributed like $\Xi_{b,Q-q}(Q-q)$. Then the square of the diffusion coefficient for the original process, at time q, is given by the expectation of $\frac{1}{4\pi}\sigma^2(\Xi_{b,Q-q}(Q-q))$. We emphasize that a solution to (1.5)–(1.7) consists of both a family of random processes $\{\Xi_{a,Q}(\cdot)\}_{a\ge 0, Q\in[0,2]}$ and also a deterministic function $J:[0,2]\times\mathbf{R}_{\ge 0}\to\mathbf{R}_{\ge 0}$. That is, J is not given as part of the data of the problem but is rather found as part of the solution. Probabilistically, the processes $\Xi_{a,Q}$ are not coupled in any particular way across various choices of a and Q: each $\Xi_{a,Q}$ could be taken to live on a different probability space. However, their *laws* are related through the deterministic function J.

We note that another, equivalent, way to write the system (1.5)–(1.7) is as

$$\mathrm{d}\Xi_{a,Q}(q) = \frac{1}{2\sqrt{\pi}}\left(\mathbf{E}[\sigma^2(\Xi_{a,Q}(Q)) \mid \mathcal{G}_q]\right)^{1/2}\mathrm{d}B(q), \qquad q\in(0,Q]; \tag{1.8}$$

$$\Xi_{a,Q}(0) = a. \tag{1.9}$$

The formulation (1.8)–(1.9) is essentially a forward-backward stochastic differential equation (FBSDE). Fixing $a\ge 0$ and $Q\in[0,2]$, we consider the process $\{(X(q),Y(q),Z(q))\}_{q\in[0,Q]}$, with all components adapted to the filtration $\{\mathcal{G}_q\}_{q\ge 0}$, satisfying the coupled forward-backward stochastic differential equation

$$\mathrm{d}X(q) = \sqrt{Y(q)}\,\mathrm{d}B(q), \qquad X(0) = a, \tag{1.10}$$

$$\mathrm{d}Y(q) = Z(q)\,\mathrm{d}B(q), \qquad Y(Q) = \frac{1}{4\pi}\sigma^2(X(Q)). \tag{1.11}$$

Here the equation for $X(\cdot)$ is forward since the initial condition is given, and the equation for $Y(\cdot)$ is backward since the terminal condition is given. Because Y is supposed to be a martingale with terminal value $\frac{1}{4\pi}\sigma^2(X(Q))$, we actually have $Y(q) = \frac{1}{4\pi}\mathbf{E}[\sigma^2(X(Q)) \mid \mathcal{G}_q]$. As a result, $X(\cdot)$ solves the same equation as $\Xi_{a,Q}(\cdot)$.

In the FBSDE formulation, the auxiliary function J (called a "decoupling function" in the FBSDE literature [34, 35, 23]) is not required, although it can be recovered from (1.8) by (1.7). The formulations (1.8)–(1.9) and (1.5)–(1.7) are equivalent because the law of $\Xi_{a,Q}(Q)$ conditional on $\Xi_{a,Q}(q) = b$ is the same as the law of $\Xi_{b,Q-q}(Q-q)$. We similarly note that a solution to (1.10)–(1.11) will satisfy $Y(q) = J^2(Q-q, X(q))$. The formulation (1.5)–(1.7) turns out to be easier to work with, since one can first solve for the deterministic decoupling function J, and once J is known the problem (1.5)–(1.6) becomes a standard stochastic differential equation. We refer the reader to, for example, [36] for background on FBSDEs. We also point out that the function $J^2(q,b)$ is a viscosity solution to the quasilinear heat equation

$$\partial_q J^2 = \frac{1}{2}J^2\partial_{bb}J^2; \tag{1.12}$$

$$J^2(0,b) = \frac{1}{4\pi}\sigma^2(b), \tag{1.13}$$

as can be seen by an argument similar to that of [36, Section 8.2], using the moment bound in Remark 2.1 below.

The non-Lipschitz dependence of (1.10) on Y, as well as the potentially quadratic growth of σ^2 at infinity, exclude the system (1.10)–(1.11) from the established well-posedness theories for FBSDEs, discussed in [36, 35]. Nonetheless, we can prove the following well-posedness result.

Theorem 1.1. *If $\beta < \sqrt{2\pi}$, then there is a unique continuous function $J:[0,2]\times\mathbf{R}_{\ge 0}\to\mathbf{R}_{\ge 0}$ satisfying the following conditions:*

1. *For each $q\in[0,2]$, $J(q,\cdot)$ is Lipschitz,*

$$J(q,0) = 0, \tag{1.14}$$

and

$$\operatorname{Lip} J(q,\cdot) \le (4\pi/\beta^2 - q)^{-1/2}. \tag{1.15}$$

2. *For each $a \ge 0$ and $Q \in [0,2]$, the solution $\{\Xi_{a,Q}(q)\}_{0\le q\le Q}$ to the problem (1.5)–(1.6) (with this choice of J) satisfies $\frac{1}{2\sqrt{\pi}}(\mathbf{E}\sigma(\Xi_{a,Q}(Q))^2)^{1/2} = J(Q,a)$. In other words, (1.7) is satisfied with $q = Q$ and $b = a$.*

The proof of Theorem 1.1 is given in Section 2. Now that we have established existence and uniqueness of solutions to (1.5)–(1.7), in the sense of Theorem 1.1, we can state our main theorem.

Theorem 1.2. *If $\beta < \sqrt{2\pi}$, then for any $Q \in [0,2]$ and $X \in \mathbf{R}^2$, we have*

$$u_{\varepsilon,a}(\varepsilon^{2-Q}, X) \xrightarrow[\varepsilon\downarrow 0]{\text{law}} \Xi_{a,Q}(Q), \tag{1.16}$$

where $\Xi_{a,Q}$ comes from the solution to (1.5)–(1.7). For any fixed $T > 0$ and $X \in \mathbf{R}^2$ we have

$$u_{\varepsilon,a}(T, X) \xrightarrow[\varepsilon\downarrow 0]{\text{law}} \Xi_{a,2}(2). \tag{1.17}$$

The constant 2 appearing (twice) in (1.17) comes from the fact that, for fixed $T > 0$, the time variables q and t, corresponding to the ODE (1.5) and the PDE (1.1) respectively, are (informally) related by

$$t = T - \varepsilon^q.$$

This is related to the fact that the noise contributes to the solution on this ε-dependent exponential scale, as we discuss more in Sections 1.1 and 1.2 below. The terminal time 2 corresponds to the G_{ε^2} in the correlation function (1.3) for the noise: the mollification cuts off the dynamics below this scale.

Of course, even deterministic ODEs are not generally integrable in elementary terms, so we do not expect to be able to solve the system (1.5)–(1.7) explicitly for general σ. However, in the linear case $\sigma(u) = \beta u$, the system can indeed be solved explicitly. In that case, we recover the log-normal fluctuations proved in [7]. We show how to do this in Section 1.3 below.

The work [7] also dealt with limiting multipoint statistics of solutions to (1.1)–(1.2) with $\sigma(x) = \beta x$. It turns out that $u_{\varepsilon,a}(t_1,x_1)$ and $u_{\varepsilon,a}(t_2,x_2)$ are asymptotically independent if

$$d((\tau_1,x_1),(\tau_2,x_2)) := \max\{|t_1-t_2|^{1/2}, |x_1-x_2|\} \tag{1.18}$$

is of order 1. To see a nontrivial correlation structure, we must put $t_2 = t_1 + \varepsilon^\alpha$ and $x_2 = x_1 + \varepsilon^\beta$ for some $\alpha,\beta > 0$. This situation persists in the nonlinear case, and we can express the limiting joint laws of multiple points separated on these scales by a branching version of the ODE (1.5)–(1.6), as we state in the following theorem. Note that once J has been obtained from the single-point problem (1.5)–(1.7), it is no longer necessary to consider (1.7) in the multipoint problem: J is then simply a fixed deterministic function, depending only on σ.

Theorem 1.3. *Suppose that $\beta < \sqrt{2\pi}$. Let $N \in \mathbf{N}$ and fix N space-time points $(\tau_{\varepsilon,1}, x_{\varepsilon,1}), \dots, (\tau_{\varepsilon,N}, x_{\varepsilon,N}) \in \mathbf{R}_{>0} \times \mathbf{R}^2$, depending on ε. Define the metric d as in (1.18). Suppose that*

$$d_{ij} := 1 - \lim_{\varepsilon\downarrow 0} \log_\varepsilon d((\tau_{\varepsilon,i}, x_{\varepsilon,i}), (\tau_{\varepsilon,j}, x_{\varepsilon,j})) \tag{1.19}$$

exists for all i, j, and suppose that

$$Q := 2 - \lim_{\varepsilon\downarrow 0} \log_\varepsilon \tau_{\varepsilon,j} \tag{1.20}$$

exists, is independent of j, and is at most 2. Define

$$i_q(j) = \min\{i \in \{1,\dots,N\} \,:\, d_{ij} < q\}. \tag{1.21}$$

Let J be as in the solution to (1.5)–(1.7). Let $B_1,\dots,B_N$ be a family of N independent standard Brownian motions. For $a \in \mathbf{R}$, let $(\Gamma_{a,Q,j})_{j=1}^N$ solve the family of SDEs

$$\mathrm{d}\Gamma_{a,Q,j}(q) = J(Q-q,\Gamma_{a,Q,j}(q))\mathrm{d}B_{i_{(Q-q)/2}(j)}(q), \qquad j \in \{1,\dots,N\}; \tag{1.22}$$

$$\Gamma_{a,Q,j}(0) = a. \tag{1.23}$$

Then we have

$$(u_{\varepsilon,a}(\tau_{\varepsilon,j},x_{\varepsilon,j}))_{j=1}^N \xrightarrow[\varepsilon\downarrow 0]{\text{law}} (\Gamma_{a,Q,j}(Q))_{j=1}^N. \tag{1.24}$$

The quantity d_{ij} represents the distance between $(\tau_{\varepsilon,i},x_{\varepsilon,i})$ and $(\tau_{\varepsilon,j},x_{\varepsilon,j})$ on the exponential scale. Of particular note here is the ultrametricity property

$$d_{ik} \le \max\{d_{ij},d_{jk}\} \tag{1.25}$$

for all $i,j,k \in \{1,\dots,N\}$. If one restricts to a single point ($N=1$) then it is of course clear that (1.22)–(1.23) agrees with (1.5)–(1.6). For two points, if we consider $\tau_{\varepsilon,1} = \tau_{\varepsilon,2} = T > 0$ independent of ε and $|x_{\varepsilon,1}-x_{\varepsilon,2}| = \varepsilon^\alpha$ with some $\alpha \in [0,1]$, then $Q=2$, $d_{11}=d_{22}=-\infty$, $d_{12}=1-\alpha$, and it is clear that $\Xi_{a,Q,1}$ is driven by B_1 in $[0,2]$, while $\Xi_{a,Q,2}$ is driven by B_1 in $[0,2\alpha]$ and by B_2 in $[2\alpha,2]$. Two extreme cases are $\alpha=0$ and $\alpha=1$, in which $\Xi_{a,Q,1}$ and $\Xi_{a,Q,2}$ are independent and identical respectively. In the general case, we note that the set $\{i_{(Q-q)/2}(j) : j \in \{1,\dots,N\}\}$ only grows larger as q increases. Therefore, the members of the family of SDEs (1.22)–(1.23) will generally start stuck together and then branch apart at times q such that $1-\frac{q}{2} = d_{ij}$ for some $i,j \in \{1,\dots,N\}$. Thus we obtain a multiscale correlation structure generalizing the one obtained for the linear case in [7, Theorem 2.15 and (2.18)]. In Section 1.3 we show how to recover [7, (2.18)] from Theorem 1.3 in the linear case.

1.1 The exponential time scale

A key feature of the SPDE (1.1)–(1.2) is that, in the subcritical regime $\beta < \sqrt{2\pi}$, it evolves on an exponential time scale, with respect to the strength of the random noise. To see this, consider the following equation in microscopic variables:

$$\mathrm{d}u_a(t,x) = \frac{1}{2}\Delta u_a(t,x)\mathrm{d}t + \delta\,\sigma(u_a(t,x))\mathrm{d}W^1(t,x), \qquad u_a(0,\cdot) \equiv a,$$

with $\mathrm{d}W^1$ the Gaussian noise that is white in time and smooth in space (the spatial covariance function being G_1 by (1.3)), and $\delta > 0$ a fixed small parameter. We are interested in determining the scales on which nontrivial effects from the random noise can be observed. As expected, it depends on the dimension through the integrability of the heat kernel.

In $d=1$, the correct scale turns out to be $(t,x) = (\frac{T}{\delta^4},\frac{X}{\delta^2})$, where (T,X) are the corresponding macroscopic variables as discussed for directed polymers in [1] and for SPDEs in [2, 29]. In $d \ge 3$, if $\delta\beta$ is small enough so that the problem is in the weak disorder regime, one can consider an "arbitrarily long" diffusive scale $(t,x) = (\frac{T}{\varepsilon^2},\frac{X}{\varepsilon})$ with $\varepsilon \to 0$ independent of δ. The $d=2$ case is very special. As observed in [7] for the linear case $\sigma(x) = \beta x$, the second moment $f_a(t) := \mathbf{E}u_a(t,x)^2$ satisfies a closed-form equation

$$f_a(t) = a^2 + \frac{\delta^2\beta^2}{4\pi}\int_0^t \frac{f_a(s)}{t-s+\frac{1}{2}}ds.$$

This is a Volterra equation, and one can easily analyze the asymptotic behavior of $f_a(t)$ for large t and small δ:

$$f_a(t) \approx \frac{a^2}{1 - \frac{\delta^2\beta^2 \log t}{4\pi}}, \quad \text{if } \frac{\delta^2\beta^2}{4\pi}\log t < 1.$$

Due to the dependence on $\log t$, to see a nontrivial evolution, one should consider an exponential time scale and let $t = \mathrm{e}^{Q/\delta^2}$ with $Q \le 2$ (we used Q rather than T as the macroscopic variable here, to emphasize this is on the *exponential* scale). For $\delta = (\log\varepsilon^{-1})^{-\frac{1}{2}}$, this leads to $t = \varepsilon^{-Q}$. On the other hand, by the scaling property of the white noise, one can easily check that, in $d = 2$, we have

$$u_{\varepsilon,a}(\cdot,\cdot) \overset{\text{law}}{=} u_a(\tfrac{\cdot}{\varepsilon^2}, \tfrac{\cdot}{\varepsilon}), \qquad \text{if } \delta = (\log\varepsilon^{-1})^{-\frac{1}{2}}.$$

Thus, $u_{\varepsilon,a}(\varepsilon^{2-Q},0) \overset{\text{law}}{=} u_a(\varepsilon^{-Q},0)$, and from this perspective, it is natural to consider the scaling used in (1.16), which says that for any macroscopic variable $Q \in [0,2]$, we have

$$u_a(\varepsilon^{-Q},0) \xrightarrow[\varepsilon\downarrow 0]{\text{law}} \Xi_{a,Q}(Q).$$

1.2 Sketch of the proof

The proof of Theorem 1.2 begins with a series of approximations of the SPDE (1.1)–(1.2). Fix $T > 0, X \in \mathbf{R}^2$. The underlying phenomenology behind these approximations is that the contribution of the noise $\mathrm{d}W^\varepsilon$ on an interval $[T-\varepsilon^q, T-\varepsilon^{q+\gamma}]$ to the L^2 norm of the solution $u_{\varepsilon,a}(T,X)$ can be bounded from above by $\gamma^{1/2}$. Therefore, we can "turn off" the noise on intervals $[T-\varepsilon^{q_i}, T-\varepsilon^{q_i+\gamma}]$, $i = 1,\dots,M$, and as long as $M\gamma^{1/2} \ll 1$, this will not change $u_{\varepsilon,a}(T,X)$ in the limit. (We describe precisely how we choose these increments at the beginning of Section 6.) For any $A \subset [0,\infty)$, we define $u^A_{\varepsilon,a}$ as the solution to

$$\mathrm{d}u^A_{\varepsilon,a}(t,x) = \frac{1}{2}\Delta u^A_{\varepsilon,a}(t,x)\mathrm{d}t + \frac{\mathbf{1}_{\mathbf{R}\setminus A}(t)}{\sqrt{\log\varepsilon^{-1}}}\sigma(u^A_{\varepsilon,a}(t,x))\mathrm{d}W^\varepsilon(t,x); \tag{1.26}$$

$$u^A_{\varepsilon,a}(0,x) = a. \tag{1.27}$$

This comes from the problem (1.1)–(1.2) by "turning off" the noise on the set A. Section 4 is devoted to bounding the error incurred by turning off the noise on an interval.

Let $\tilde{u}_{\varepsilon,a} = u^A_{\varepsilon,a}$, with $A = \bigcup_{i=1}^M [T-\varepsilon^{q_i}, T-\varepsilon^{q_i+\gamma}]$, denote the solution with the noise turned off in this way. Fix any $i = 1,\dots,M$. Since we expect the problem to have a diffusive scaling, $\tilde{u}_{\varepsilon,a}(T-\varepsilon^{q_i+\gamma},x)$ should contribute to $u_{\varepsilon,a}(T,X)$ only for those x such that $|x-X| \lesssim \varepsilon^{(q_i+\gamma)/2}$. We further choose γ so that $\varepsilon^\gamma \ll 1$. The noise is turned off on the interval $[T-\varepsilon^{q_i}, T-\varepsilon^{q_i+\gamma}]$, so $\tilde{u}_{\varepsilon,a}(T-\varepsilon^{q_i+\gamma},\cdot)$ has been subject to the deterministic heat equation (with no noise) for the last $T-\varepsilon^{q_i+\gamma}-(T-\varepsilon^{q_i}) = \varepsilon^{q_i}(1-\varepsilon^\gamma) \approx \varepsilon^{q_i}$ amount of time, and thus is essentially constant on spatial scales much smaller than $\varepsilon^{q_i/2}$. Thus, since $\varepsilon^\gamma \ll 1$ and thus $\varepsilon^{(q_i+\gamma)/2} \ll \varepsilon^{q_i/2}$, the main contribution of noise up until time $T-\varepsilon^{q_i+\gamma}$ on $u_{\varepsilon,a}(T,X)$ is via the constant $\tilde{u}_{\varepsilon,a}(T-\varepsilon^{q_i+\gamma},X)$. Section 5 is devoted to bounding the error incurred by replacing the field by a (random) constant after the solution has been subject to the deterministic heat equation for some time. In Section 6, we define the time discretization that we use, and then iterate the results of Sections 4 and 5 to bound the total error incurred by this approximation scheme.

Our approximation scheme approximates the solution $u_{\varepsilon,a}(T,X)$ in terms of a scalar-valued Markov chain whose ith value is $\tilde{u}_{\varepsilon,a}(T-\varepsilon^{q_i+\gamma},X)$. (Since the equation starts from constant initial data and we are interested in the marginal distribution, by space-stationarity, the choice of X is arbitrary and plays no role.) This Markov chain, which is also a discrete martingale, will approximate the solution to (1.5)–(1.7). To see why, we note that step $(i+1)$ of the Markov chain is given by solving the original equation (1.1)–(1.2) with the initial condition a equaling to the current value of the Markov chain, which is $\tilde{u}_{\varepsilon,a}(T-\varepsilon^{q_i+\gamma},X)$, on an interval of length $\varepsilon^{q_i+\gamma}-\varepsilon^{q_{i+1}}\approx\varepsilon^{q_i+\gamma}$, and then letting the solution evolve according to the heat equation for time $\varepsilon^{q_{i+1}}-\varepsilon^{q_{i+1}+\gamma}\approx\varepsilon^{q_{i+1}}$. Although it only represents one step of the Markov chain, approximating the solution on these time scales require running another instance of the Markov chain for $M-i$ steps. This is a consequence of the mild solution formula; see Lemma 7.7 below. This corresponds to the $Q-q$ in the argument of J in (1.5). On the other hand, since this only represents one step of the Markov chain, one only needs to understand the variance rather than the complete law in order to compute the diffusivity of the limiting diffusion. Accounting for the averaging from the heat equation (which gives us a factor of q_i-q_{i-1}), it turns out that this variance is approximated by the expression on the right side of (1.7) in the limit. In particular, the fact that only the variance is important is reflected in the fact that an expectation is taken on the right side of (1.7). Making these ideas precise is the main task of Section 7.

The fact that the diffusion coefficient of the limiting SDE can be represented in terms of statistics of the chain itself is of course critical to proving the existence of the limit. The fact that the self-similar structure characterizes the limit is reflected in the fact that the problem (1.5)–(1.7) is well-posed, as stated in Theorem 1.1. This well-posedness allows us to construct the limiting diffusion coefficient and then show that the Markov chain converges to the diffusion using standard techniques. This is the content of Section 8.

We address multipoint statistics, and prove Theorem 1.3, in Section 9. At this stage, since the problem (1.5)–(1.7) has been solved, the function J has been identified. The Markov chains corresponding to multiple points stay together at earlier times, but then eventually branch apart from each other as the remaining time scale approaches the spatial separation of the points. It turns out that once they branch apart, they are completely independent in the limit. This yields the branching diffusion structure (1.22)–(1.23).

1.3 The linear case

In this subsection, we consider the linear case $\sigma(u)=\beta u$ and show that solutions to (1.5)–(1.7) have log-normal one-point statistics, and moreover that we recover the limiting variance [7, (2.18)] obtained in [7, Theorem 2.15]. In this case, the linearity of the problem (1.5)–(1.7) allows us to make the *ansatz* $J(q,b)=b\underline{J}(q)$, with $\underline{J}(q)=J(q,1)$. Then the problem becomes

$$\mathrm{d}\Xi_{a,Q}(q)=\underline{J}(Q-q)\Xi_{a,Q}(q)\mathrm{d}B(q),\qquad q\in[0,Q];\tag{1.28}$$

$$\Xi_{a,Q}(0)=a;\tag{1.29}$$

$$\underline{J}(q)=\frac{\beta}{2\sqrt{\pi}}(\mathbf{E}\Xi_{1,q}(q)^2)^{1/2}.\tag{1.30}$$

We can already see that (up to a time-change determined by $\underline{J}$) the problem (1.28)–(1.29) is solved by a geometric Brownian motion. It turns out that we can compute $\underline{J}$ explicitly. By Itô's formula applied to (1.28) we have

$$\mathrm{d}(\log\Xi_{a,Q})(q)=\underline{J}(Q-q)\mathrm{d}B(q)-\frac{1}{2}\underline{J}(Q-q)^2\mathrm{d}q,\tag{1.31}$$

and hence

$$\Xi_{a,Q}(Q)=a\exp\left\{\int_0^Q\underline{J}(Q-q)\,\mathrm{d}B(q)-\frac{1}{2}\int_0^Q\underline{J}(Q-q)^2\,\mathrm{d}q\right\}.\tag{1.32}$$

Taking $a=1$, substituting (1.32) into (1.30), and computing the expectation, we obtain

$$\underline{J}(Q)^2=\frac{\beta^2}{4\pi}\exp\left\{\int_0^Q\underline{J}(q)^2\,\mathrm{d}q\right\}.$$

Differentiating this expression gives us the differential equation $\frac{\mathrm{d}}{\mathrm{d}Q}\underline{J}(Q)^2 = \underline{J}(Q)^4$. Combining this with the initial condition $\underline{J}(0) = \frac{\beta}{2\sqrt{\pi}}$, which is evident from (1.6) and (1.7), we obtain

$$\underline{J}(Q) = (4\pi/\beta^2 - Q)^{-1/2}. \tag{1.33}$$

Note that the resulting J, given by

$$J(q,b) = \frac{b}{\sqrt{4\pi/\beta^2 - q}}, \tag{1.34}$$

saturates the bound (1.15). Substituting (1.33) into (1.32), we have

$$\begin{aligned} \Xi_{a,Q}(Q) &= a\exp\left\{\int_0^Q \frac{1}{\sqrt{4\pi/\beta^2-(Q-q)}}\,\mathrm{d}B(q) - \frac{1}{2}\int_0^Q \frac{\mathrm{d}q}{4\pi/\beta^2-(Q-q)}\right\} \\ &\overset{\text{law}}{=} a\exp\left\{S - \frac{1}{2}\mathbf{E}S^2\right\}, \end{aligned} \tag{1.35}$$

where $S \sim N(0, \log\frac{4\pi/\beta^2}{4\pi/\beta^2-Q})$. In the case $Q = 2$ and $a = 1$, this agrees with the expression [7, (2.12)].

Now we address the multipoint statistics, i.e. the problem (1.22)–(1.23). As in (1.31), but now knowing (1.33), we have

$$\mathrm{d}(\log\Gamma_{a,Q,j})(q) = \frac{\mathrm{d}B_{i_{(Q-q)/2}(j)}(q)}{\sqrt{4\pi/\beta^2-(Q-q)}} - \frac{\mathrm{d}q}{8\pi/\beta^2-2(Q-q)}.$$

From this linear SDE we see that the family $(\log\Gamma_{a,Q,j}(Q))_{j=1}^N$ is jointly Gaussian. All of the means are equal as

$$\mathbf{E}[\log\Gamma_{a,Q,j}(Q)] = \log a - \frac{1}{2}\int_0^Q \frac{\mathrm{d}q}{4\pi/\beta^2-(Q-q)} = \log a - \frac{1}{2}\log\frac{4\pi/\beta^2}{4\pi/\beta^2-Q}$$

as in (1.35). The covariance structure is given by

$$\begin{aligned} &\operatorname{Cov}(\log\Gamma_{a,Q,i}(Q), \log\Gamma_{a,Q,j}(Q)) \\ &\quad= \int_{\{q\in[0,Q]\,:\,i_{(Q-q)/2}(i)=i_{(Q-q)/2}(j)\}} \frac{\mathrm{d}q}{4\pi/\beta^2-(Q-q)} \\ &\quad= \int_{[0,Q-2d_{ij}\vee 0]} \frac{\mathrm{d}q}{4\pi/\beta^2-(Q-q)} = \log\frac{4\pi/\beta^2-(2d_{ij}\vee 0)\wedge Q}{4\pi/\beta^2-Q}. \end{aligned} \tag{1.36}$$

The second equality is by the ultrametricity property (1.25) of the d_{ij}s. For $Q = 2$, (1.36) is the same as the covariance structure [7, (2.18)] obtained in [7, Theorem 2.15].

2 Proof of Theorem 1.1

In this section we prove Theorem 1.1, establishing the well-posedness of the limiting problem. The analysis here is essentially independent of the rest of the paper.

Proof of Theorem 1.1. If $g : [0,Q]\times\mathbf{R}_{\ge 0} \to \mathbf{R}_{\ge 0}$ is continuous, is Lipschitz in the second variable, and satisfies $g(\cdot,0)\equiv 0$, then for each $a \ge 0$ and $Q \in [0,2]$ we let $\Xi^g_{a,Q}$ solve the problem

$$\mathrm{d}\Xi^g_{a,Q}(q) = g(Q-q, \Xi^g_{a,Q}(q))\mathrm{d}B(q); \tag{2.1}$$

$$\Xi^g_{a,Q}(0) = a. \tag{2.2}$$

It is standard that (2.1)–(2.2) has a unique strong solution with continuous sample paths almost surely, and that this solution is positive with probability 1. (For the last property see e.g. [38, Lemma 2.1].) We write (2.1)–(2.2) in the mild formulation

$$\Xi^g_{a,Q}(q) = a + \int_0^q g(Q-s, \Xi^g_{a,Q}(s))\, \mathrm{d}B_s.$$

Define

$$\mathcal{Q}g(Q,a) = \frac{1}{2\sqrt{\pi}}(\mathbf{E}\sigma(\Xi^g_{a,Q}(Q))^2)^{1/2}.$$

We note that J satisfies the condition 2 in the statement of the theorem if and only if $\mathcal{Q}J = J$. We will show that there is a unique such fixed point J under the additional assumption that condition 1 in the statement of the theorem is satisfied.

To this end, let $\mathcal{X}$ be the Banach space of continuous functions $f : \mathbf{R}_{\geq 0} \to \mathbf{R}$ such that $f(0) = 0$ and the norm

$$\|f\|_{\mathcal{X}} = \sup_{a>0} \frac{|f(a)|}{a}$$

is finite. Let $\mathcal{Y}$ be the Banach space of continuous functions $g : [0,2] \times \mathbf{R}_{\geq 0} \to \mathbf{R}$ such that $g(q,0) = 0$ for all $q \in [0,2]$ and the norm

$$\|g\|_{\mathcal{Y}} = \sup_{\substack{q\in[0,2]\\ a>0}} \mathrm{e}^{-R(\beta)q} \frac{|g(q,a)|}{a} \tag{2.3}$$

is finite, where we have defined

$$R(\beta) = 2\beta^2 \left(\frac{4\pi/\beta^2}{4\pi/\beta^2 - 2}\right)^3. \tag{2.4}$$

Finally, let $\mathcal{Z} \subset \mathcal{Y}$ be the closed subset defined by

$$\mathcal{Z} = \left\{ g \in \mathcal{Y} \ : \ \inf_{a\geq 0} g(q,a) \geq 0 \text{ and } \operatorname{Lip} g(q,\cdot) \leq (4\pi/\beta^2 - q)^{-1/2} \text{ for all } q \in [0,2] \right\}.$$

Thus, we are done if we can show that the map $\mathcal{Q}$ has a unique fixed point in $\mathcal{Z}$, and we will do this by showing that $\mathcal{Q}$ maps $\mathcal{Z}$ into itself and moreover is a contraction on $\mathcal{Z}$.

Step 1: L^2 bound. If $g \in \mathcal{Z}$, by the fact that $g(q,0) = 0$ we have $g(q,x) \leq \operatorname{Lip} g(q,\cdot)x$ for any $x > 0$, so

$$\mathbf{E}\Xi^g_{a,Q}(q)^2 = a^2 + \int_0^q \mathbf{E}g(Q-p, \Xi^g_{a,Q}(p))^2\, \mathrm{d}p \leq a^2 + \int_0^q \frac{\mathbf{E}\Xi^g_{a,Q}(p)^2}{4\pi/\beta^2 - Q + p}\, \mathrm{d}p.$$

By Grönwall's inequality, this means that

$$\mathbf{E}\Xi^g_{a,Q}(q)^2 \leq a^2 \exp\left\{\int_0^q \frac{1}{4\pi/\beta^2 - Q + p}\, \mathrm{d}p\right\} = a^2 \cdot \frac{4\pi/\beta^2 - Q + q}{4\pi/\beta^2 - Q}. \tag{2.5}$$

Step 2: $\mathcal{Q}$ maps $\mathcal{Z}$ to itself. Let $g \in \mathcal{Z}$. It is clear that $\mathcal{Q}g(q,0) = 0$ for all $q \in [0,2]$. It remains to check that $\mathcal{Q}g$ is continuous and $\operatorname{Lip}(\mathcal{Q}g(q,\cdot)) \leq (4\pi/\beta^2 - q)^{-1/2}$ for all $q \in [0,2]$. For the Lipschitz property, we have

$$\begin{aligned} |\mathcal{Q}g(Q,a) - \mathcal{Q}g(Q,b)| &= \frac{1}{2\sqrt{\pi}} \left|(\mathbf{E}\sigma(\Xi^g_{a,Q}(Q))^2)^{1/2} - (\mathbf{E}\sigma(\Xi^g_{b,Q}(Q))^2)^{1/2}\right| \\ &\leq \frac{\beta}{2\sqrt{\pi}} \left(\mathbf{E}[\Xi^g_{a,Q}(Q) - \Xi^g_{b,Q}(Q)]^2\right)^{1/2}. \end{aligned} \tag{2.6}$$

Now we note that, for any $q \le Q$, we have

$$\begin{aligned}\mathbf{E}[\Xi^g_{a,Q}(q) - \Xi^g_{b,Q}(q)]^2 &= (a-b)^2 + \int_0^q \mathbf{E}[g(Q-p, \Xi^g_{a,Q}(p)) - g(Q-p, \Xi^g_{b,Q}(p))]^2\,\mathrm{d}p \\ &\le (a-b)^2 + \int_0^q \mathrm{Lip}(g(Q-p,\cdot))^2 \mathbf{E}[\Xi^g_{a,Q}(p) - \Xi^g_{b,Q}(p)]^2\,\mathrm{d}p.\end{aligned}$$

By Grönwall's inequality, this means that

$$\mathbf{E}[\Xi^g_{a,Q}(q) - \Xi^g_{b,Q}(q)]^2 \le (a-b)^2 \exp\left\{\int_0^q \mathrm{Lip}(g(Q-p,\cdot))^2\,\mathrm{d}s\right\}.$$

Using this in (2.6), we have

$$\begin{aligned}|\mathcal{Q}g(Q,a) - \mathcal{Q}g(Q,b)| &\le \frac{\beta}{2\sqrt{\pi}}|a-b| \exp\left\{\frac{1}{2}\int_0^Q \mathrm{Lip}(g(Q-p,\cdot))^2\,\mathrm{d}p\right\} \\ &= \frac{\beta}{2\sqrt{\pi}}|a-b| \exp\left\{\frac{1}{2}\int_0^Q \mathrm{Lip}(g(p,\cdot))^2\,\mathrm{d}p\right\},\end{aligned}$$

so

$$\mathrm{Lip}(\mathcal{Q}g(Q,\cdot)) \le \frac{\beta}{2\sqrt{\pi}} \exp\left\{\frac{1}{2}\int_0^Q \mathrm{Lip}(g(p,\cdot))^2\,\mathrm{d}p\right\}.$$

Therefore, since

$$\mathrm{Lip}(g(p,\cdot)) \le (4\pi/\beta^2 - p)^{-1/2},$$

we also have

$$\mathrm{Lip}(\mathcal{Q}g(Q,\cdot)) \le \frac{\beta}{2\sqrt{\pi}} \exp\left\{\frac{1}{2}\int_0^Q \frac{1}{4\pi/\beta^2 - p}\,\mathrm{d}p\right\} = (4\pi/\beta^2 - Q)^{-1/2}.$$

Next we show that for each $a > 0$, $\mathcal{Q}g(\cdot, a)$ is continuous on $[0,2]$. The argument is rather standard and similar to the above discussion, so we do not provide all details. Taking $0 \le Q_1 < Q_2 \le 2$, we have

$$|\mathcal{Q}g(Q_1,a) - \mathcal{Q}g(Q_2,a)| \le \frac{\beta}{2\sqrt{\pi}}\left(\mathbf{E}|\Xi^g_{a,Q_1}(Q_1) - \Xi^g_{a,Q_2}(Q_2)|^2\right)^{1/2}.$$

For any $q \le Q_1$, we write the difference as

$$\Xi^g_{a,Q_1}(Q_1) - \Xi^g_{a,Q_2}(Q_2) = \Xi^g_{a,Q_1}(Q_1) - \Xi^g_{a,Q_2}(Q_1) + \Xi^g_{a,Q_2}(Q_1) - \Xi^g_{a,Q_2}(Q_2),$$

and the first term can be estimated as follows: for any $q \le Q_1$,

$$\Xi^g_{a,Q_1}(q) - \Xi^g_{a,Q_2}(q) = \int_0^q g(Q_1 - s, \Xi^g_{a,Q_1}(s))\,\mathrm{d}B_s - \int_0^q g(Q_2 - s, \Xi^g_{a,Q_2}(s))\,\mathrm{d}B_s,$$

which yields

$$\begin{aligned}\mathbf{E}|\Xi^g_{a,Q_1}(q) - \Xi^g_{a,Q_2}(q)|^2 &\le 2\int_0^q \mathbf{E}|g(Q_1 - s, \Xi^g_{a,Q_1}(s)) - g(Q_2 - s, \Xi^g_{a,Q_1}(s))|^2\,\mathrm{d}s \\ &\quad + 2\int_0^q \mathbf{E}|g(Q_2 - s, \Xi^g_{a,Q_1}(s)) - g(Q_2 - s, \Xi^g_{a,Q_2}(s))|^2\,\mathrm{d}s \\ &=: I_1 + I_2.\end{aligned}$$

The term I_2 can be bounded from above by

$$2\int_0^q \mathrm{Lip}(g(Q_2-s,\cdot)^2\mathbf{E}|\Xi^g_{a,Q_1}(s)-\Xi^g_{a,Q_2}(s)|^2\,\mathrm{d}s.$$

For I_1, the integrand

$$\mathbf{E}[|g(Q_1-s,\Xi^g_{a,Q_1}(s))-g(Q_2-s,\Xi^g_{a,Q_1}(s))|^2]$$

is bounded, and converges to zero as $Q_2\to Q_1$ for each s, by the dominated convergence theorem, (2.5) and the fact that g is continuous in the first variable and $g(q,x)\le Cx$ for all $x\ge 0, q\in[0,2]$. Therefore, invoking Grönwall's inequality again, we obtain

$$\mathbf{E}|\Xi^g_{a,Q_1}(Q_1)-\Xi^g_{a,Q_2}(Q_1)|^2\to 0,\qquad \text{as } Q_2\to Q_1.$$

A simpler argument shows that

$$\mathbf{E}|\Xi^g_{a,Q_2}(Q_2)-\Xi^g_{a,Q_2}(Q_1)|^2\to 0,\qquad \text{as } Q_2\to Q_1.$$

Therefore, $\mathcal{Q}g(\cdot,a)$ is continuous, so $\mathcal{Q}$ maps $\mathcal{Z}$ to itself.

Step 3: contraction. Let $g_1,g_2\in\mathcal{Z}$. Then we have

$$\Xi^{g_1}_{a,Q}(q)-\Xi^{g_2}_{a,Q}(q)=\int_0^q[g_1(Q-p,\Xi^{g_1}_{a,Q}(p))-g_2(Q-p,\Xi^{g_2}_{a,Q}(p))]\,\mathrm{d}B(p),$$

so

$$\begin{aligned}\mathbf{E}[\Xi^{g_1}_{a,Q}-\Xi^{g_2}_{a,Q}](q)^2&=\int_0^q\mathbf{E}[g_1(Q-p,\Xi^{g_1}_{a,Q}(p))-g_2(Q-p,\Xi^{g_2}_{a,Q}(p))]^2\,\mathrm{d}p\\&\le 2\int_0^q\left(\|(g_1-g_2)(Q-p,\cdot)\|^2_{\mathcal{X}}\mathbf{E}\Xi^{g_1}_{a,Q}(p)^2+\mathrm{Lip}(g_2(Q-p,\cdot))^2\mathbf{E}[\Xi^{g_1}_{a,Q}-\Xi^{g_2}_{a,Q}](p)^2\right)\mathrm{d}p\\&\le 2\int_0^q\left(\|(g_1-g_2)(Q-p,\cdot)\|^2_{\mathcal{X}}a^2\cdot\frac{4\pi/\beta^2-Q+p}{4\pi/\beta^2-Q}+\frac{\mathbf{E}[\Xi^{g_1}_{a,Q}-\Xi^{g_2}_{a,Q}](p)^2}{4\pi/\beta^2-Q+p}\right)\mathrm{d}p,\end{aligned}$$

with the last inequality by (2.5). By Grönwall's inequality, this means that

$$\begin{aligned}&\mathbf{E}[\Xi^{g_1}_{a,Q}-\Xi^{g_2}_{a,Q}](q)^2\\&\le 2a^2\left(\int_0^q\|(g_1-g_2)(Q-p,\cdot)\|^2_{\mathcal{X}}\frac{4\pi/\beta^2-Q+p}{4\pi/\beta^2-Q}\,\mathrm{d}p\right)\exp\left\{\int_0^q\frac{2}{4\pi/\beta^2-Q+p}\,\mathrm{d}p\right\}\\&=2a^2\cdot\left(\frac{4\pi/\beta^2-Q+q}{4\pi/\beta^2-Q}\right)^2\int_0^q\|(g_1-g_2)(Q-p,\cdot)\|^2_{\mathcal{X}}\frac{4\pi/\beta^2-Q+p}{4\pi/\beta^2-Q}\,\mathrm{d}p.\end{aligned}$$

In particular, we have

$$\mathbf{E}[\Xi^{g_1}_{a,Q}-\Xi^{g_2}_{a,Q}](Q)^2\le 2a^2\cdot\left(\frac{4\pi/\beta^2}{4\pi/\beta^2-Q}\right)^3\int_0^Q\|(g_1-g_2)(p,\cdot)\|^2_{\mathcal{X}}\,\mathrm{d}p.$$

Then we have

$$\begin{aligned}(\mathcal{Q}g_1-\mathcal{Q}g_2)(q,a)^2&=\left|(\mathbf{E}\sigma(\Xi^{g_1}_{a,q}(q))^2)^{1/2}-(\mathbf{E}\sigma(\Xi^{g_2}_{a,q}(q))^2)^{1/2}\right|^2\\&\le\beta^2\mathbf{E}[\Xi^{g_1}_{a,q}(q)-\Xi^{g_2}_{a,q}(q)]^2\\&\le 2a^2\beta^2\left(\frac{4\pi/\beta^2}{4\pi/\beta^2-q}\right)^3\int_0^q\|(g_1-g_2)(p,\cdot)\|^2_{\mathcal{X}}\,\mathrm{d}p.\end{aligned}$$

This implies that, as long as $\beta < \sqrt{2\pi}$, for all $q \in [0,2]$ we have

$$\|(\mathcal{Q}g_1 - \mathcal{Q}g_2)(q,\cdot)\|_{\mathcal{X}}^2 \le 2\beta^2 \left(\frac{4\pi/\beta^2}{4\pi/\beta^2 - q}\right)^3 \int_0^q \|(g_1 - g_2)(p,\cdot)\|_{\mathcal{X}}^2 \,\mathrm{d}p.$$

Therefore,

$$\begin{aligned}\|\mathcal{Q}g_1 - \mathcal{Q}g_2\|_{\mathcal{Y}}^2 &= \sup_{q\in[0,2]} \mathrm{e}^{-2R(\beta)q} \|(\mathcal{Q}g_1 - \mathcal{Q}g_2)(q,\cdot)\|_{\mathcal{X}}^2 \\ &\le \sup_{q\in[0,2]} 2\beta^2 \left(\frac{4\pi/\beta^2}{4\pi/\beta^2 - 2}\right)^3 e^{-2R(\beta)q} \int_0^q \|(g_1 - g_2)(p,\cdot)\|_{\mathcal{X}}^2 \,\mathrm{d}p \\ &\le \beta^2 \left(\frac{4\pi/\beta^2}{4\pi/\beta^2 - 2}\right)^3 \frac{1}{R(\beta)} \|g_1 - g_2\|_{\mathcal{Y}}^2 = \frac{1}{2}\|g_1 - g_2\|_{\mathcal{Y}}^2.\end{aligned}$$

Recall that $R(\beta)$ was defined in (2.4). Therefore, $\mathcal{Q}$ is a contraction on $\mathcal{Z}$ (equipped with the norm inherited from $\mathcal{Y}$) and so $\mathcal{Q}$ admits a unique fixed point in $\mathcal{Z}$, which is what we needed to show. □

Remark 2.1. By the stochastic comparison principle for SDEs [16] and the fact that the geometric Brownian motion (i.e. a log-normal random variable) has finite positive moments of all orders, we see that $\mathbf{E}\Xi_{a,Q}(q)^k < \infty$ for all $k \in [0,\infty)$ as well.

3 Moment bounds

The next several sections will work towards a proof of Theorem 1.2. In order to carry out our analysis, we will need some bounds on the moments of the solutions to (1.1)–(1.2). We establish these in this section. Moment bounds depend crucially on the *subcriticality* of the problem, which for us means $\beta < \sqrt{2\pi}$. We will assume throughout the paper that this is true without further comment. We also now fix a time horizon $T_0 \in [1,\infty)$ which will also remain fixed throughout the paper. Furthermore, fix $\varepsilon_0 \in (0,1]$ so that

$$\frac{\beta^2}{4\pi} \cdot \frac{\log(1 + 2\varepsilon^{-2}T_0)}{\log \varepsilon^{-1}} < 1 \tag{3.1}$$

for all $\varepsilon \in (0,\varepsilon_0]$. The condition that $\beta < \sqrt{2\pi}$ means that such an ε_0 exists. As we are ultimately interested in the limit $\varepsilon \downarrow 0$, the condition (3.1) is simply a convenience so that various quantities are finite. In Definition 3.5 below, we fix a constant $K_0 < \infty$, which depends on β, ε_0, and T_0, and will appear in upper bounds throughout the paper.

Proposition 3.1. *There exist constants $p > 2$ and $K < \infty$ (depending on T_0 and β) so that, for all $\varepsilon \in (0,\varepsilon_0]$, all $a \ge 0$, and all $t \in [0,T_0], x \in \mathbf{R}^2$, we have*

$$\mathbf{E}u_{\varepsilon,a}(t,x)^p \le K^p a^p. \tag{3.2}$$

Proof. Let $v_{\varepsilon,a}$ solve the linear problem given by (1.1)–(1.2) with $\sigma(u) = \beta u$. By [9, (5.11)], for any $p \in [1, 2\pi/\beta^2 + 1)$ we have a constant K so that $\mathbf{E}v_{\varepsilon,a}(t,x)^p \le K^p a^p$. Using the stochastic comparison principle proved in [14, (E-4)], since $\sigma(u) \le \beta u$ for all $u \in [0,\infty)$ we have $\mathbf{E}u_{\varepsilon,a}(t,x)^p \le \mathbf{E}v_{\varepsilon,a}(t,x)^p \le K^p a^p$. By the assumption that $\beta < \sqrt{2\pi}$, we have $2\pi/\beta^2 + 1 > 2$, so we can choose $p > 2$ as required. □

Remark 3.2. The case $p = 2$ in (3.2) is much simpler than the case $p > 2$. Indeed, the $p = 2$ case is a special case of Proposition 3.3 below. On the other hand, the proof of the moment bound for $p > 2$ in [9] for the linear case uses hypercontractivity, and the stochastic comparison principle [14] takes a substantial amount of analysis to prove. Most of the analysis in this paper will be in the L^2 setting, so we will mostly use the $p = 2$ case. However, we will rely on some tightness statements that require a higher moment bound.

The following proposition gives an L^2 bound on the difference of two solutions started at different initial conditions. Recall that $u^A_{\varepsilon,a}$ solves the problem (1.26)–(1.27), with the noise turned off on the set of times A. The problem (1.26)–(1.27) has the mild formulation

$$u^A_{\varepsilon,a}(t,x) = a + \frac{1}{\sqrt{\log\varepsilon^{-1}}}\int_{[0,t]\setminus A}\int G_{t-s}(x-y)\sigma(u^A_{\varepsilon,a}(s,y))\,\mathrm{d}W^\varepsilon(s,y). \tag{3.3}$$

Here and henceforth, when we do not specify the domain of integration for an integral we mean that the integral is taken over all of $\mathbf{R}^2$.

Proposition 3.3. *There exists a constant $K<\infty$ (depending on T_0 and β) so that, for all $\varepsilon\in(0,\varepsilon_0]$, $a_1,a_2\ge 0$, $T\in[0,T_0]$, $x\in\mathbf{R}^2$, and measurable $A\subset[0,\infty)$, we have*

$$\left(\mathbf{E}[u^A_{\varepsilon,a_2}(t,x)-u^A_{\varepsilon,a_1}(t,x)]^2\right)^{1/2} \le K|a_2-a_1|. \tag{3.4}$$

In particular, for any $a>0$,

$$\left(\mathbf{E}u^A_{\varepsilon,a}(t,x)^2\right)^{1/2} \le Ka. \tag{3.5}$$

In fact, (3.4) and (3.5) hold with

$$K = \left(1-\frac{\beta^2}{4\pi}\cdot\frac{\log(1+2\varepsilon^{-2}t)}{\log\varepsilon^{-1}}\right)^{-1/2}. \tag{3.6}$$

Of course, a very important special case is when $A=\emptyset$. Then the bounds (3.4) and (3.5) just involve $u_{\varepsilon,a}$. (In the latter case this of course is a special case of Proposition 3.1.)

Proof. Since (3.5) is just (3.4) with $a_2=a$ and $a_1=0$, it suffices to prove (3.4). Subtracting two copies of (3.3) (with $a=a_1$ and $a=a_2$) and taking second moments, we obtain

$$\begin{aligned}
&\mathbf{E}(u^A_{\varepsilon,a_2}(t,x)-u^A_{\varepsilon,a_1}(t,x))^2\\
&\quad=(a_2-a_1)^2\\
&\qquad+\frac{1}{\log\varepsilon^{-1}}\int_{[0,t]\setminus A}\iint \mathbf{E}\prod_{i=1}^{2}\left([\sigma(u^A_{\varepsilon,a_2}(s,y_i))-\sigma(u^A_{\varepsilon,a_1}(s,y_i))]G_{t-s}(x-y_i)\right)\\
&\qquad\qquad\qquad\cdot G_{\varepsilon^2}(y_1-y_2)\,\mathrm{d}y_1\,\mathrm{d}y_2\,\mathrm{d}s\\
&\quad\le (a_2-a_1)^2+\frac{\beta^2}{2\pi\log\varepsilon^{-1}}\int_0^t\frac{\mathbf{E}|u^A_{\varepsilon,a_2}(s,x)-u^A_{\varepsilon,a_1}(s,x)|^2}{2(t-s)+\varepsilon^2}\,\mathrm{d}s.
\end{aligned}$$

Then (3.4) follows from Lemma 3.4 below. □

It remains to prove the lemma used above, which will also be useful in the future.

Lemma 3.4. *For all $\varepsilon\in(0,\varepsilon_0]$, all $a\ge0$, and all $T\in[0,T_0]$, the following holds. Let $f:[0,T]\to[0,\infty)$ be such that*

$$f(t)\le a^2+\frac{\beta^2}{2\pi\log\varepsilon^{-1}}\int_0^t\frac{f(s)}{2(t-s)+\varepsilon^2}\,\mathrm{d}s$$

for all $t\in[0,T]$. Then, for all $t\in[0,T]$, we have

$$f(t)\le\frac{a^2}{1-\frac{\beta^2}{4\pi}\cdot\frac{\log(1+2\varepsilon^{-2}t)}{\log\varepsilon^{-1}}}.$$

Proof. Define $[0,t]^j_< = \{(s_1,\ldots,s_j) \in [0,t]^j \mid s_1 \le \cdots \le s_j\}$. Then we have

$$
\begin{aligned}
f(t) &\le a^2 \sum_{j=0}^{\infty} \frac{\beta^{2j}}{(4\pi \log \varepsilon^{-1})^j} \int_{[0,t]^j_<} \prod_{k=1}^{j} \frac{1}{s_{k+1} - s_k + \varepsilon^2/2} \, \mathrm{d}s_1 \cdots \mathrm{d}s_j \\
&\le a^2 \sum_{j=0}^{\infty} \frac{\beta^{2j}}{(4\pi \log \varepsilon^{-1})^j} \int_{[0,t]^j} \prod_{k=1}^{j} \frac{1}{r_j + \varepsilon^2/2} \, \mathrm{d}r_1 \cdots \mathrm{d}r_j \\
&= a^2 \sum_{j=0}^{\infty} \frac{\beta^{2j}}{(4\pi \log \varepsilon^{-1})^j} \left(\int_0^t \frac{1}{r + \varepsilon^2/2} \, \mathrm{d}r \right)^j \\
&= a^2 \sum_{j=0}^{\infty} \left(\frac{\beta^2}{4\pi \log \varepsilon^{-1}} \log(1 + 2\varepsilon^{-2} t) \right)^j = \frac{a^2}{1 - \frac{\beta^2}{4\pi} \cdot \frac{\log(1+2\varepsilon^{-2}t)}{\log \varepsilon^{-1}}},
\end{aligned}
\tag{3.7}
$$

where we used (3.1) for the last identity. □

To avoid having to constantly quantify constants, we now fix our essential constant once and for all.

Definition 3.5. Fix

$$
K_0 \ge \sup_{\varepsilon \in (0,\varepsilon_0]} \left(1 - \frac{\beta^2}{4\pi} \cdot \frac{2 + \log(1 + 2\varepsilon^{-2} t)}{\log \varepsilon^{-1}} \right)^{-1/2} \tag{3.8}
$$

large enough so that Propositions 3.1 and 3.3 hold with $K = K_0$.

By (3.6) and the proof of [9, (5.11)], we see that we could take

$$
K_0 = \sup_{\varepsilon \in (0,\varepsilon_0]} \left(1 - \frac{\tilde{\beta}^2}{4\pi} \cdot \frac{2 + \log(1 + 2\varepsilon^{-2} t)}{\log \varepsilon^{-1}} \right)^{-1/2}
$$

for some $\tilde{\beta} \in (\beta, \sqrt{2\pi})$. The precise form of K_0 will not be important for us (although at one point we will directly use the explicit expression (3.6)). The extra summand of 2 in the lower limit condition (3.8) for K_0 (compared to (3.6)) is to allow K_0 to also suffice for bounds in later sections. (See the proofs of Lemmas 4.3 and 5.2 below.)

Now we can bootstrap Proposition 3.3 to obtain a stronger bound on the variance of the solution.

Proposition 3.6. *If $a > 0$, $\varepsilon \in [0, \varepsilon_0)$, and $A \subset [0,\infty)$ is measurable, then*

$$
\left(\mathbf{E}[u^A_{\varepsilon,a}(t,x) - a]^2 \right)^{1/2} \le \frac{\beta a K_0}{2\sqrt{\pi}} \sqrt{\frac{\log(1 + 2\varepsilon^{-2} t)}{\log \varepsilon^{-1}}}. \tag{3.9}
$$

Of course, for t of order 1, the bound (3.9) is redundant to (3.5). It will be used when t is chosen small so that $\log(1 + \varepsilon^{-2} t) \ll \log \varepsilon^{-1}$.

Proof. Similar to the computation in Proposition 3.3, we have

$$
\mathbf{E}[u^A_{\varepsilon,a}(t,x) - a]^2 \le \frac{\beta^2}{2\pi \log \varepsilon^{-1}} \int_0^t \frac{\mathbf{E} u^A_{\varepsilon,a}(s,x)^2}{2(t-s) + \varepsilon^2} \, \mathrm{d}s,
$$

and then (3.9) follows from (3.5). □

4 Turning off the noise on an interval

As discussed in the introduction, an important part of our argument will be turning off the noise in the equation (1.1)–(1.2) for a certain set of times, and comparing the resulting solution to the original solution. In this section we bound the error incurred by this noise shutoff procedure when the noise is shut off on a single interval. In Section 6, we will iterate this procedure to turn off the noise on multiple intervals. For now our goal is to prove the following proposition.

Proposition 4.1. *Let* $A \subset [0,\infty)$ *and suppose that* $\sup A \le \tau_1 \le \tau_2 \le T_0$. *Then for any* $t \in [\tau_2, T_0]$ *and any* $x \in \mathbf{R}^2$ *we have*

$$\mathbf{E}\left(u_{\varepsilon,a}^{A}(t,x) - u_{\varepsilon,a}^{A\cup[\tau_1,\tau_2]}(t,x)\right)^2 \le \frac{K_0^4\beta^2 a^2}{4\pi \log \varepsilon^{-1}}\left(\log\frac{t-\tau_1+\varepsilon^2/2}{t-\tau_2+\varepsilon^2/2} + K_0^2\right).$$

Proof. Subtracting two copies of the mild formulation (3.3) (with the sets A and $A\cup[\tau_1,\tau_2]$ respectively), we have

$$\begin{aligned}
&u_{\varepsilon,a}^{A}(t,x) - u_{\varepsilon,a}^{A\cup[\tau_1,\tau_2]}(t,x)\\
&\quad = \frac{1}{\sqrt{\log\varepsilon^{-1}}}\int_{[0,t]\setminus A}\int G_{t-s}(x-y)\sigma(u_{\varepsilon,a}^{A}(s,y))\,\mathrm{d}W^{\varepsilon}(s,y)\\
&\qquad - \frac{1}{\sqrt{\log\varepsilon^{-1}}}\int_{[0,t]\setminus(A\cup[\tau_1,\tau_2])}\int G_{t-s}(x-y)\sigma(u_{\varepsilon,a}^{A\cup[\tau_1,\tau_2]}(s,y))\,\mathrm{d}W^{\varepsilon}(s,y)\\
&\quad = \frac{1}{\sqrt{\log\varepsilon^{-1}}}\int_{\tau_1}^{\tau_2}\int G_{t-s}(x-y)\sigma(u_{\varepsilon,a}^{A}(s,y))\,\mathrm{d}W^{\varepsilon}(s,y)\\
&\qquad + \frac{1}{\sqrt{\log\varepsilon^{-1}}}\int_{\tau_2}^{t}\int G_{t-s}(x-y)[\sigma(u_{\varepsilon,a}^{A}(s,y)) - \sigma(u_{\varepsilon,a}^{A\cup[\tau_1,\tau_2]}(s,y))]\,\mathrm{d}W^{\varepsilon}(s,y).
\end{aligned}$$

In the second "=" we used that $u_{\varepsilon,a}^{A}(t,x) = u_{\varepsilon,a}^{A\cup[\tau_1,\tau_2]}(t,x)$ whenever $t \le \tau_1$. Taking the second moment, we have for all $t \ge \tau_2$ that

$$\begin{aligned}
&\mathbf{E}\left(u_{\varepsilon,a}^{A}(t,x) - u_{\varepsilon,a}^{A\cup[\tau_1,\tau_2]}(t,x)\right)^2\\
&= \frac{1}{\log\varepsilon^{-1}}\int_{\tau_1}^{\tau_2}\iint G_{\varepsilon^2}(y_1-y_2)\mathbf{E}\prod_{i=1}^{2}\left(G_{t-s}(x-y_i)\sigma(u_{\varepsilon,a}^{A}(s,y_i))\right)\mathrm{d}y_1\,\mathrm{d}y_2\,\mathrm{d}s\\
&\quad + \frac{1}{\log\varepsilon^{-1}}\int_{\tau_2}^{t}\iint \mathbf{E}\prod_{i=1}^{2}\left(G_{t-s}(x-y_i)[\sigma(u_{\varepsilon,a}^{A}(s,y_i)) - \sigma(u_{\varepsilon,a}^{A\cup[\tau_1,\tau_2]}(s,y_i))]\right)\\
&\qquad\qquad \cdot G_{\varepsilon^2}(y_1-y_2)\,\mathrm{d}y_1\,\mathrm{d}y_2\,\mathrm{d}s\\
&\le \frac{\beta^2}{2\pi\log\varepsilon^{-1}}\int_{\tau_1}^{\tau_2}\frac{\mathbf{E}u_{\varepsilon,a}^{A}(s,y)^2}{2(t-s)+\varepsilon^2}\mathrm{d}s + \frac{\beta^2}{2\pi\log\varepsilon^{-1}}\int_{\tau_2}^{t}\frac{\mathbf{E}[u_{\varepsilon,a}^{A}(s,y) - u_{\varepsilon,a}^{A\cup[\tau_1,\tau_2]}(s,y)]^2}{2(t-s)+\varepsilon^2}\mathrm{d}s\\
&\le \frac{\beta^2 a^2 K_0^2}{4\pi\log\varepsilon^{-1}}\log\frac{t-\tau_1+\varepsilon^2/2}{t-\tau_2+\varepsilon^2/2} + \frac{\beta^2}{4\pi\log\varepsilon^{-1}}\int_{\tau_2}^{t}\frac{\mathbf{E}[u_{\varepsilon,a}^{A}(s,y) - u_{\varepsilon,a}^{A\cup[\tau_1,\tau_2]}(s,y)]^2}{t-s+\varepsilon^2/2}\mathrm{d}s. \qquad (4.1)
\end{aligned}$$

In the last inequality we used (3.5). Now if we put

$$f(t) = \mathbf{E}\left(u_{\varepsilon,a}^{A}(\tau_2+t,x) - u_{\varepsilon,a}^{A\cup[\tau_1,\tau_2]}(\tau_2+t,x)\right)^2, \qquad t \ge 0, \qquad (4.2)$$

then (4.1) can be rewritten as

$$f(t) \le \frac{\beta^2 a^2 K_0^2}{4\pi \log \varepsilon^{-1}} \log \frac{t+\tau_2-\tau_1+\varepsilon^2/2}{t+\varepsilon^2/2} + \frac{\beta^2}{4\pi \log \varepsilon^{-1}} \int_0^t \frac{f(s)}{t-s+\varepsilon^2/2} \mathrm{d}s.$$

Now we apply Lemma 4.3 below with $M = (4\pi)^{-1}\beta^2 a^2 K_0^2$ and $r = \tau_2 - \tau_1$. (The requirement that f has a bounded supremum on compact intervals is satisfied by applying Proposition 3.1.) This gives us

$$f(t) \le \frac{K_0^4 \beta^2 a^2}{4\pi \log \varepsilon^{-1}} \left(\log \frac{t+\tau_2-\tau_1+\varepsilon^2/2}{t+\varepsilon^2/2} + K_0^2 \right).$$

Recalling the definition (4.2) completes the proof. □

We will prove Lemma 4.3, which we used in the above proof, shortly. First we need a preliminary lemma.

Lemma 4.2. *For any $t, r, \varepsilon > 0$ we have*

$$\int_0^t \frac{\log \frac{t+r+\varepsilon^2/2}{s+\varepsilon^2/2}}{t-s+\varepsilon^2/2} \mathrm{d}s \le \left(2 + \log(1+2\varepsilon^{-2}t)\right)\left(1 + \log \frac{t+r+\varepsilon^2/2}{t+\varepsilon^2/2}\right).$$

Proof. We write

$$\begin{aligned} \int_0^t \frac{\log \frac{t+r+\varepsilon^2/2}{s+\varepsilon^2/2}}{t-s+\varepsilon^2/2} \mathrm{d}s &= \left(\int_0^{t/2} + \int_{t/2}^t \right) \frac{\log \frac{t+r+\varepsilon^2/2}{s+\varepsilon^2/2}}{t-s+\varepsilon^2/2} \mathrm{d}s \\ &\le \frac{2}{t} \int_0^t \log \frac{t+r+\varepsilon^2/2}{s+\varepsilon^2/2} \mathrm{d}s + \left(\log \frac{t+r+\varepsilon^2/2}{t/2+\varepsilon^2/2} \right) \int_0^t \frac{1}{t-s+\varepsilon^2/2} \mathrm{d}s. \end{aligned} \tag{4.3}$$

Now we have

$$\begin{aligned} \int_0^t \log \frac{t+r+\varepsilon^2/2}{s+\varepsilon^2/2} \mathrm{d}s &= t - \frac{\varepsilon^2}{2} \log \frac{t+\varepsilon^2/2}{\varepsilon^2/2} + t \log \frac{t+r+\varepsilon^2/2}{t+\varepsilon^2/2} \\ &\le t \left(1 + \log \frac{t+r+\varepsilon^2/2}{t+\varepsilon^2/2} \right). \end{aligned} \tag{4.4}$$

Also, we have

$$\log \frac{t+r+\varepsilon^2/2}{t/2+\varepsilon^2/2} = \log \frac{2t+2r+\varepsilon^2}{t+\varepsilon^2} \le \log 2 + \log \frac{t+r+\varepsilon^2/2}{t+\varepsilon^2/2}. \tag{4.5}$$

Using (4.4) and (4.5) in (4.3), we have

$$\begin{aligned} \int_0^t \frac{\log \frac{t+r+\varepsilon^2/2}{s+\varepsilon^2/2}}{t-s+\varepsilon^2/2} \mathrm{d}s &\le 2 + 2\log \frac{t+r+\varepsilon^2/2}{t+\varepsilon^2/2} + \left(\log 2 + \log \frac{t+r+\varepsilon^2/2}{t+\varepsilon^2/2} \right) \log(1+2\varepsilon^{-2}t) \\ &\le \left(2 + \log(1+2\varepsilon^{-2}t)\right)\left(1 + \log \frac{t+r+\varepsilon^2/2}{t+\varepsilon^2/2}\right), \end{aligned} \tag{4.6}$$

which was the claim. □

Lemma 4.3. *Let $\varepsilon \in (0, \varepsilon_0]$ and $M, r > 0$, suppose that f satisfies the bound*

$$f(t) \le \frac{M}{\log \varepsilon^{-1}} \log \frac{t+r+\varepsilon^2/2}{t+\varepsilon^2/2} + \frac{\beta^2}{4\pi \log \varepsilon^{-1}} \int_0^t \frac{f(s)}{t-s+\varepsilon^2/2} \mathrm{d}s \tag{4.7}$$

for all $t \in [0, T_0]$*, and* $\sup_{t\in[0,T_0]} |f(t)| < \infty$*. Then we have*

$$f(t) \le \frac{K_0^2 M}{\log \varepsilon^{-1}} \left(\log \frac{t + r + \varepsilon^2/2}{t + \varepsilon^2/2} + K_0^2 \right) \tag{4.8}$$

for all $t \in [0, T_0]$.

Proof. Suppose that

$$f(t) \le B_1 \log \frac{t + r + \varepsilon^2/2}{t + \varepsilon^2/2} + B_2. \tag{4.9}$$

By assumption, this inequality holds with $B_1 = 0$ and $B_2 = \sup_{t\in[0,T_0]} |f(t)|$. Substituting (4.9) into the r.h.s. of (4.7), we have

$$\begin{aligned} f(t) &\le \frac{M}{\log \varepsilon^{-1}} \log \frac{t + r + \varepsilon^2/2}{t + \varepsilon^2/2} + \frac{\beta^2}{4\pi \log \varepsilon^{-1}} \int_0^t \frac{B_1 \log \frac{s+r+\varepsilon^2/2}{s+\varepsilon^2/2} + B_2}{t - s + \varepsilon^2/2} \, \mathrm{d}s \\ &\le \frac{M}{\log \varepsilon^{-1}} \log \frac{t + r + \varepsilon^2/2}{t + \varepsilon^2/2} + \frac{\beta^2 B_1}{4\pi \log \varepsilon^{-1}} \int_0^t \frac{\log \frac{s+r+\varepsilon^2/2}{s+\varepsilon^2/2}}{t - s + \varepsilon^2/2} \, \mathrm{d}s + \frac{\beta^2 B_2 \log(1 + 2\varepsilon^{-2} T_0)}{4\pi \log \varepsilon^{-1}}. \end{aligned} \tag{4.10}$$

For the middle term of the above inequality, we have

$$\int_0^t \frac{\log \frac{s+r+\varepsilon^2/2}{s+\varepsilon^2/2}}{t - s + \varepsilon^2/2} \, \mathrm{d}s \le \int_0^t \frac{\log \frac{t+r+\varepsilon^2/2}{s+\varepsilon^2/2}}{t - s + \varepsilon^2/2} \, \mathrm{d}s \le \left(2 + \log(1 + 2\varepsilon^{-2} t)\right) \left(1 + \log \frac{t + r + \varepsilon^2/2}{t + \varepsilon^2/2}\right) \tag{4.11}$$

by Lemma 4.2. Substituting (4.11) into (4.10), we have

$$\begin{aligned} f(t) &\le \frac{M}{\log \varepsilon^{-1}} \log \frac{t + r + \varepsilon^2/2}{t + \varepsilon^2/2} + \frac{\beta^2 B_1}{4\pi \log \varepsilon^{-1}} \left(2 + \log(1 + 2\varepsilon^{-2} t)\right) \left(1 + \log \frac{t + r + \varepsilon^2/2}{t + \varepsilon^2/2}\right) \\ &\quad + \frac{\beta^2 B_2 \log(1 + 2\varepsilon^{-2} T_0)}{4\pi \log \varepsilon^{-1}} \\ &= \frac{1}{\log \varepsilon^{-1}} \left(\frac{\beta^2 B_1}{4\pi} \left(2 + \log(1 + 2\varepsilon^{-2} t)\right) + M \right) \log \frac{t + r + \varepsilon^2/2}{t + \varepsilon^2/2} \\ &\quad + \frac{\beta^2 B_1}{4\pi \log \varepsilon^{-1}} \left(2 + \log(1 + 2\varepsilon^{-2} t)\right) + \frac{\beta^2 B_2 \log(1 + 2\varepsilon^{-2} T_0)}{4\pi \log \varepsilon^{-1}} \\ &\le \left((1 - K_0^{-2}) B_1 + \frac{M}{\log \varepsilon^{-1}} \right) \log \frac{t + r + \varepsilon^2/2}{t + \varepsilon^2/2} + B_1 + (1 - K_0^{-2}) B_2, \end{aligned}$$

where in the last inequality we used (3.8). Define $B_1^{(0)} = 0$ and $B_2^{(0)} = \sup_{t\in[0,T_0]} |f(t)|$, so for each $n \ge 0$, (4.9) holds with

$$B_1 = B_1^{(n)} = (1 - K_0^{-2}) B_1^{(n-1)} + \frac{M}{\log \varepsilon^{-1}}, \tag{4.12}$$

$$B_2 = B_2^{(n)} = B_1^{(n-1)} + (1 - K_0^{-2}) B_2^{(n-1)}. \tag{4.13}$$

From (4.12) we conclude that

$$B_1^{(n)} \le \frac{K_0^2 M}{\log \varepsilon^{-1}} \tag{4.14}$$

for all n. Then we have from (4.13) that

$$B_2^{(n)} \le \frac{K_0^2 M}{\log \varepsilon^{-1}} + (1 - K_0^{-2}) B_2^{(n-1)},$$

so

$$\limsup_{n\to\infty} B_2^{(n)} \le \frac{K_0^4 M}{\log \varepsilon^{-1}}. \tag{4.15}$$

Using (4.14) and (4.15) in (4.9), we obtain (4.8). □

5 Replacing a smoothed field with a constant

In Section 4, we estimated the effect on the solution of turning off the noise on a given time interval. In this section we seek a further simplification. After an interval of time in which the noise has been turned off, the resulting solution will have been undergoing nothing more than the deterministic heat equation on that interval. Therefore, it will have been smoothed, with a strength depending on the length of the interval. By restricting our attention to a comparatively small spatial region, we would expect that the solution may be replaced by a constant at the end of this interval. The following proposition is to quantify the induced error when we replace the solution by a (random) constant at the end of each "quiet" interval.

Proposition 5.1. *Let $A \subset [0,\infty)$ be measurable and let $\tau_1 < \tau_2 < T$ be such that $\tau_2 = \sup A$ and $[\tau_1, \tau_2] \subset A$. Fix $X \in \mathbf{R}^2$ and let v solve the problem*

$$\mathrm{d}v(t,x) = \frac{1}{2}\Delta v(t,x)\mathrm{d}t + (\log\varepsilon^{-1})^{-\frac{1}{2}}\sigma(v(t,x))\mathrm{d}W^\varepsilon(t,x), \qquad t > \tau_2, x \in \mathbf{R}^2; \tag{5.1}$$

$$v(\tau_2, x) = u^A_{\varepsilon,a}(\tau_2, X). \tag{5.2}$$

Then we have, for all $t \in [\tau_2, T]$ and $\varepsilon \le e^{-K_0^2}$, that

$$\mathbf{E}(v - u^A_{\varepsilon,a})(t,x)^2 \le K_0^4 a^2 \frac{3(t-\tau_2) + |x - X|^2}{\tau_2 - \tau_1}. \tag{5.3}$$

Proof. We first note that $u^A_{\varepsilon,a}(\tau_2, X) = \int G_{\tau_2-\tau_1}(X-y) u^A_{\varepsilon,a}(\tau_1, y)\mathrm{d}y$, since $u^A_{\varepsilon,a}$ solves the deterministic heat equation in the time interval $[\tau_1, \tau_2]$. Then, we have for any $t > \tau_2$ that

$$\begin{aligned}(v - u^A_{\varepsilon,a})(t,x) &= \int [G_{\tau_2-\tau_1}(X-y) - G_{t-\tau_1}(x-y)] u^A_{\varepsilon,a}(\tau_1, y)\,\mathrm{d}y \\ &\quad + \frac{1}{\sqrt{\log\varepsilon^{-1}}} \int_{\tau_2}^t \int G_{t-s}(x-y)[\sigma(v(s,y)) - \sigma(u^A_{\varepsilon,a}(s,y))]\,\mathrm{d}W^\varepsilon(s,y).\end{aligned}$$

Taking the second moment, we obtain

$$\begin{aligned}&\mathbf{E}(v - u^A_{\varepsilon,a})(t,x)^2 \\ &\le \iint \mathbf{E}\prod_{i=1}^2 \Big([G_{\tau_2-\tau_1}(X-y_i) - G_{t-\tau_1}(x-y_i)] u^A_{\varepsilon,a}(\tau_1, y_i)\Big)\,\mathrm{d}y_1\,\mathrm{d}y_2 \\ &\quad + \frac{\beta^2}{\log\varepsilon^{-1}} \int_{\tau_2}^t \iint G_{\varepsilon^2}(y_1 - y_2)\mathbf{E}\prod_{i=1}^2 \Big(G_{t-s}(x-y_i)|v(s,y_i) - u^A_{\varepsilon,a}(s,y_i)|\Big)\,\mathrm{d}y_1\,\mathrm{d}y_2\,\mathrm{d}s \\ &=: I_1 + I_2.\end{aligned} \tag{5.4}$$

For the first term, we estimate by the Cauchy–Schwarz inequality (on the probability space) that

$$\begin{aligned} I_1 &\le \left(\int |G_{\tau_2-\tau_1}(X-y)-G_{t-\tau_1}(x-y)|\left(\mathbf{E}u^A_{\varepsilon,a}(\tau_1,y)^2\right)^{1/2}\mathrm{d}y\right)^2 \\ &\le K_0^2a^2\|G_{\tau_2-\tau_1}(X-\cdot)-G_{t-\tau_1}(x-\cdot)\|^2_{L^1(\mathbf{R}^2)}, \end{aligned} \tag{5.5}$$

where the second inequality is by (3.5). By Pinsker's inequality (see e.g. [31, Lemma 1.5.3 and Theorem 1.5.4]), we have

$$\|G_{\tau_2-\tau_1}(X-\cdot)-G_{t-\tau_1}(x-\cdot)\|^2_{L^1(\mathbf{R}^2)} \le 2D_{\mathrm{KL}}(G_{t-\tau_1}(x-\cdot)\,\|\,G_{\tau_2-\tau_1}(X-\cdot)), \tag{5.6}$$

where D_{KL} denotes the Kullback–Leibler divergence (also known as the relative entropy). We recall that for two continuous probability distributions F_1 and F_2 on $\mathbf{R}^2$, the Kullback–Leibler divergence is defined as

$$D_{\mathrm{KL}}(F_1\,\|\,F_2)=\int F_1(x)\log\frac{F_1(x)}{F_2(x)}\,\mathrm{d}x.$$

Then we can compute explicitly (see e.g. [31, Theorem 1.8.2]) that

$$\begin{aligned} D_{\mathrm{KL}}(G_{t-\tau_1}(x-\cdot)\,\|\,G_{\tau_2-\tau_1}(X-\cdot)) &= \log\frac{\tau_2-\tau_1}{t-\tau_1}-1+\frac{t-\tau_1}{\tau_2-\tau_1}+\frac{|X-x|^2}{2(\tau_2-\tau_1)} \\ &\le \frac{t-\tau_2+\frac{1}{2}|X-x|^2}{\tau_2-\tau_1}. \end{aligned} \tag{5.7}$$

Substituting (5.7) into (5.6) and then into (5.5), we have

$$I_1 \le \frac{K_0^2a^2}{\tau_2-\tau_1}[2(t-\tau_2)+|X-x|^2]. \tag{5.8}$$

Considering the second term of (5.4), we apply the inequality $|ab|\le\frac{1}{2}(a^2+b^2)$ and use the symmetry in y_1,y_2 to derive

$$\begin{aligned} I_2 &\le \frac{\beta^2}{2\log\varepsilon^{-1}}\sum_{j=1}^2\int_{\tau_2}^t\iint G_{\varepsilon^2}(y_1-y_2)\mathbf{E}|v(s,y_j)-u^A_{\varepsilon,a}(s,y_j)|^2\prod_{i=1}^2 G_{t-s}(x-y_i)\,\mathrm{d}y_1\,\mathrm{d}y_2\,\mathrm{d}s \\ &= \frac{\beta^2}{\log\varepsilon^{-1}}\int_{\tau_2}^t\int G_{t-s+\varepsilon^2}(x-y)G_{t-s}(x-y)\mathbf{E}|v(s,y)-u^A_{\varepsilon,a}(s,y)|^2\,\mathrm{d}y\,\mathrm{d}s. \end{aligned}$$

Recalling the simple fact that in $d=2$,

$$G_{t_1}(\cdot)G_{t_2}(\cdot)=\frac{1}{2\pi(t_1+t_2)}G_{\frac{t_1t_2}{t_1+t_2}}(\cdot), \tag{5.9}$$

for any $t_1,t_2>0$, we further obtain

$$I_2 \le \frac{\beta^2}{4\pi\log\varepsilon^{-1}}\int_{\tau_2}^t\int\frac{1}{t-s+\varepsilon^2/2}G_{\frac{(t-s)(t-s+\varepsilon^2)}{2(t-s)+\varepsilon^2}}(x-y)\mathbf{E}[v(s,y)-u^A_{\varepsilon,a}(s,y)]^2\,\mathrm{d}y\,\mathrm{d}s. \tag{5.10}$$

Using (5.8) and (5.10) in (5.4), we obtain

$$\begin{aligned} &\mathbf{E}(v-u^A_{\varepsilon,a})(t,x)^2 \\ &\quad\le \frac{K_0^2a^2}{\tau_2-\tau_1}[2(t-\tau_2)+|X-x|^2] \\ &\qquad+\frac{\beta^2}{4\pi\log\varepsilon^{-1}}\int_{\tau_2}^t\int\frac{1}{t-s+\varepsilon^2/2}G_{\frac{(t-s)(t-s+\varepsilon^2)}{2(t-s)+\varepsilon^2}}(x-y)\mathbf{E}[v(s,y)-u^A_{\varepsilon,a}(s,y)]^2\,\mathrm{d}y\,\mathrm{d}s. \end{aligned}$$

Thus the hypotheses of Lemma 5.2 below are satisfied with

$$f(t,x)=\mathbf{E}(v-u^A_{\varepsilon,a})(t,x)^2, \qquad A_1=2K_0^2a^2\frac{t-\tau_2}{\tau_2-\tau_1}, \qquad A_2=\frac{K_0^2a^2}{\tau_2-\tau_1},$$

from which we obtain

$$\mathbf{E}(v-u^A_{\varepsilon,a})(t,x)^2\le K_0^4a^2\left(\frac{2(t-\tau_2)}{\tau_2-\tau_1}+\frac{\beta^2K_0^2(t-\tau_2)}{2\pi(\tau_2-\tau_1)\log\varepsilon^{-1}}+\frac{|x-X|^2}{\tau_2-\tau_1}\right),$$

hence (5.3), since we have $K_0^2<\log\varepsilon^{-1}$ by assumption. □

It remains to prove the lemma we used above.

Lemma 5.2. *Suppose that* $0\le\tau\le T\le T_0$, $\sup_{t\in[\tau,T],x\in\mathbf{R}^2}|f(t,x)|<\infty$, *and there exist constants* A_1,A_2 *such that*

$$f(t,x)\le A_1+A_2|x-X|^2+\frac{\beta^2}{4\pi\log\varepsilon^{-1}}\int_\tau^t\int\frac{1}{t-s+\varepsilon^2/2}G_{\frac{(t-s)(t-s+\varepsilon^2)}{2(t-s)+\varepsilon^2}}(x-y)f(s,y)\,\mathrm{d}y\,\mathrm{d}s \tag{5.11}$$

for all $t\in[\tau,T]$ *and all* $x\in\mathbf{R}^2$. *Then, for all* $t\in[\tau,T]$ *and all* $x\in\mathbf{R}^2$, *we have*

$$f(t,x)\le K_0^2\left(A_1+\frac{\beta^2(t-\tau)}{2\pi\log\varepsilon^{-1}}K_0^2A_2+A_2|x-X|^2\right). \tag{5.12}$$

Proof. Suppose that

$$f(t,y)\le B_1+B_2|y-X|^2 \tag{5.13}$$

for all $t\in[\tau,T]$ and all $y\in\mathbf{R}^2$, where $B_1,B_2\ge0$ are constants. Of course this holds for

$$B_1=\sup_{t\in[\tau,T],x\in\mathbf{R}^2}|f(t,x)|, \qquad B_2=0.$$

Assuming (5.13), we compute from (5.11) that

$$f(t,x)\le A_1+A_2|x-X|^2+\frac{\beta^2}{4\pi\log\varepsilon^{-1}}\int_\tau^t\int\frac{1}{t-s+\varepsilon^2/2}G_{\frac{(t-s)(t-s+\varepsilon^2)}{2(t-s)+\varepsilon^2}}(x-y)[B_1+B_2|y-X|^2]\,\mathrm{d}y\,\mathrm{d}s. \tag{5.14}$$

Now we can evaluate the spatial integral by noting that

$$\int G_{\frac{(t-s)(t-s+\varepsilon^2)}{2(t-s)+\varepsilon^2}}(x-y)|y-X|^2\,\mathrm{d}y=\frac{2(t-s)(t-s+\varepsilon^2)}{2(t-s)+\varepsilon^2}+|x-X|^2\le t-s+\varepsilon^2+|x-X|^2.$$

This implies that

$$\begin{aligned}\int_\tau^t\int\frac{G_{\frac{(t-s)(t-s+\varepsilon^2)}{2(t-s)+\varepsilon^2}}(x-y)}{t-s+\varepsilon^2/2}[B_1+B_2|y-X|^2]\,\mathrm{d}y\,\mathrm{d}s&\le\int_\tau^t\frac{B_1+B_2(t-s+\varepsilon^2)+B_2|x-X|^2}{t-s+\varepsilon^2/2}\,\mathrm{d}s\\&\le\left(B_1+B_2|x-X|^2\right)\log\frac{t-\tau+\varepsilon^2/2}{\varepsilon^2/2}+2B_2(t-\tau)\le\left(B_1+B_2|x-X|^2\right)\log(1+2\varepsilon^{-2}T)+2B_2(t-\tau).\end{aligned}$$

Substituting this back into (5.14) and rearranging (also recalling (3.8)), we obtain

$$\begin{aligned}f(t,x)&\le A_1+A_2|x-X|^2+\frac{\beta^2}{4\pi\log\varepsilon^{-1}}\left[\left(B_1+B_2|x-X|^2\right)\log(1+2\varepsilon^{-2}T)+2B_2(t-\tau)\right]\\&\le\left(A_1+(1-K_0^{-2})B_1+\frac{\beta^2(t-\tau)}{2\pi\log\varepsilon^{-1}}B_2\right)+\left(A_2+(1-K_0^{-2})B_2\right)|x-X|^2.\end{aligned} \tag{5.15}$$

Let $B_1^{(0)} = \sup_{t\in[\tau,T],x\in\mathbf{R}^2} |f(t,x)|$, $B_2^{(0)} = 0$, and

$$B_1^{(n)} = A_1 + (1-K_0^{-2})B_1^{(n-1)} + \frac{\beta^2(t-\tau)}{2\pi\log\varepsilon^{-1}}B_2^{(n-1)}, \tag{5.16}$$

$$B_2^{(n)} = A_2 + (1-K_0^{-2})B_2^{(n-1)} \tag{5.17}$$

for each $n \geq 1$. By (5.15) and induction, (5.13) holds with $B_1 = B_1^{(n)}$ and $B_2 = B_2^{(n)}$ for all n. From (5.17) we see that

$$B_2^{(n)} \leq K_0^2 A_2$$

for all n, and thus from (5.16) we obtain

$$\limsup_{n\to\infty} B_1^{(n)} \leq K_0^2\left(A_1 + \frac{\beta^2(t-\tau)}{2\pi\log\varepsilon^{-1}}K_0^2 A_2\right).$$

Using the last two displays in (5.13), we obtain (5.12). □

6 The time discretization and the approximating functions

In this section, we will iterate Propositions 4.1 and Proposition 5.1 on many subintervals of time to construct a discrete Markov chain which approximates the marginal distribution of the solution to the SPDE. First we construct these intervals, which will correspond to our time-discretization scheme.

6.1 The time discretization

Our approximation scheme will ultimately be focused on approximating the distribution of $u_{\varepsilon,a}$ at a single space-time point (T,X). The time intervals of interest thus depend on the terminal time T.

For $\varepsilon \in (0,\varepsilon_0]$, define δ_ε, γ_ε, ζ_ε, and λ_ε such that

$$(\log\varepsilon^{-1})^{-1} \ll \gamma_\varepsilon \ll \delta_\varepsilon^2 \ll \lambda_\varepsilon \ll 1, \tag{6.1}$$

$$\delta_\varepsilon^{-1}\varepsilon^{\frac{1}{2}\gamma_\varepsilon} \ll 1, \tag{6.2}$$

$$(\log\varepsilon^{-1})^{-1} \ll \zeta_\varepsilon \ll 1, \tag{6.3}$$

where the notation $f(\varepsilon) \ll g(\varepsilon)$ means that $f(\varepsilon) \leq g(\varepsilon)$ for all ε and $\lim\limits_{\varepsilon\downarrow 0}\frac{f(\varepsilon)}{g(\varepsilon)} = 0$. To avoid introducing further constants later on, we further assume that

$$\max\{\varepsilon^{\gamma_\varepsilon}, \varepsilon^{\delta_\varepsilon/2}\} \leq 1/2 \tag{6.4}$$

for all $\varepsilon > 0$. The choices of the parameters will become more clear later; see the discussion at the end of this subsection.

Now we define, for $T > \varepsilon^{2-\lambda_\varepsilon}$,

$$s_m = \varepsilon^{m\delta_\varepsilon} \qquad \text{and} \qquad s_m' = \varepsilon^{m\delta_\varepsilon+\gamma_\varepsilon} \tag{6.5}$$

and

$$t_m = T - s_m \qquad \text{and} \qquad t_m' = T - s_m'. \tag{6.6}$$

Note that these quantities all depend on ε, and t_m and t_m' also depend on T, but we suppress this to simplify notations. We note that the time of interest T, unlike the time horizon T_0, is *not* fixed throughout the paper. However, whenever we use t_m and t_m', the T of current interest will be clear from the context.

Define

$$M_1(\varepsilon,T) = \lceil \delta_\varepsilon^{-1} \log_\varepsilon T \rceil - 1, \tag{6.7}$$
$$M_2(\varepsilon) = \lfloor \delta_\varepsilon^{-1}(2-\zeta_\varepsilon) \rfloor. \tag{6.8}$$

Thus $M_1(\varepsilon,T)+1$ is the least integer m so that $t_m \geq 0$, and $M_2(\varepsilon)$ is the greatest integer m so that $s_m \geq \varepsilon^{2-\zeta_\varepsilon}$. For example, for fixed $T>0$ independent of ε, we have for sufficiently small ε that

$$M_1(\varepsilon,T) = \begin{cases} -1, & \text{if } T \geq 1, \\ 0, & \text{if } T \in (0,1). \end{cases}$$

For the discrete time Markov chain to be constructed, the starting point in time will be given by $M_1(\varepsilon,T)$, and the ending point will be given by $M_2(\varepsilon)$. We note for future use that

$$M_2(\varepsilon) - M_1(\varepsilon,T) + 1 \leq \delta_\varepsilon^{-1}(2-\log_\varepsilon T). \tag{6.9}$$

Note that by the assumption of $\delta_\varepsilon > \gamma_\varepsilon$ and $\varepsilon^{\gamma_\varepsilon} < 1$, we have

$$t_{m+1} = T - \varepsilon^{m\delta_\varepsilon+\delta_\varepsilon} > T - \varepsilon^{m\delta_\varepsilon+\gamma_\varepsilon} = t'_m,$$
$$t'_m = T - \varepsilon^{m\delta_\varepsilon+\gamma_\varepsilon} > T - \varepsilon^{m\delta_\varepsilon} = t_m.$$

Thus we can write

$$[t_{M_1(\varepsilon,T)+1}, t'_{M_2(\varepsilon)}] = I_1 \cup I_2, \quad \text{with } I_1 = \bigcup_{m=M_1(\varepsilon,T)+1}^{M_2(\varepsilon)} [t_m, t'_m], \quad I_2 = \bigcup_{m=M_1(\varepsilon,T)+1}^{M_2(\varepsilon)-1} [t'_m, t_{m+1}].$$

To approximate $u_{\varepsilon,a}(T,X)$, we will turn off the noise in I_1, which consists of the "quiet" intervals. For each m, we first solve the deterministic heat equation in the interval $[t_m, t'_m]$. Then we replace the solution at $(t'_m,\cdot)$ by its value at (t'_m, X). In the next "noisy" interval $[t'_m, t_{m+1}]$, we solve the stochastic heat equation with the corresponding "constant" initial data. The error incurred in those "quiet" intervals will be quantified by Proposition 4.1, and is negligible as $\varepsilon \to 0$ by the assumption $\gamma_\varepsilon \ll \delta_\varepsilon^2$. The error incurred by modifying the initial data for those "noisy" intervals will be quantified by Proposition 5.1, and goes to zero by the assumption of $\delta_\varepsilon^{-1}\varepsilon^{\frac{1}{2}\gamma_\varepsilon} \ll 1$. The role of ζ_ε is in (6.8) to provide a small amount of extra separation between the final t_m and the time T, which will be needed for the last step of the approximation; see the proof of Proposition 7.1 below.

In the inequality (6.9), we need $\log_\varepsilon T < 2$ for all $\varepsilon \ll 1$ so that the above construction makes sense with $M_2(\varepsilon) \geq M_1(\varepsilon,T)$, and this prevents us from considering those T of order $O(\varepsilon^2)$. From Proposition 3.6, we already know that, if T is chosen so that $\log(1+2\varepsilon^{-2}T) \ll \log\varepsilon^{-1}$, the random noise plays no role in the short interval $[0,T]$, and we have $u_{\varepsilon,a}(T,x) \to a$ as $\varepsilon \to 0$. Therefore, those small T can be treated separately without constructing the Markov chain. To unify the notations, we use the following conventions:

1. If $T > \varepsilon^{2-\lambda_\varepsilon}$, we have $2 - \log_\varepsilon T \gg \delta_\varepsilon$, and $M_1(\varepsilon,T), M_2(\varepsilon)$ are defined as above.

2. If $T \in [0, \varepsilon^{2-\lambda_\varepsilon}]$, we have $\log(1+2\varepsilon^{-2}T) \ll \log\varepsilon^{-1}$ and hence $u_{\varepsilon,a}(T) \to a$ as $\varepsilon \to 0$, and we simply define $M_1(\varepsilon,T) = M_2(\varepsilon) = 1$.

6.2 The approximating functions

As we have mentioned, our approximation will be focused on a particular terminal space-time point (T,X). So in this section we fix $T \geq 0$, $X \in \mathbf{R}^2$. To define our approximation, we introduce a sequence of functions $\{w^{(m)}\}_{m=M_1(\varepsilon,T),\ldots,M_2(\varepsilon)}$ as follows. Define

$$w^{(M_1(\varepsilon,T))}(t,x) = u_{\varepsilon,a}(t,x), \qquad t \geq 0, x \in \mathbf{R}^2. \tag{6.10}$$

For $m \in \{M_1(\varepsilon,T)+1,\ldots,M_2(\varepsilon)\}$, we then inductively define $\{w^{(m)}(t,x) : t \geq t'_m, x \in \mathbf{R}^2\}$ to be the solution to

$$\mathrm{d}w^{(m)}(t,x) = \frac{1}{2}\Delta w^{(m)}(t,x)\mathrm{d}t + (\log\varepsilon^{-1})^{-\frac{1}{2}}\sigma(w^{(m)}(t,x))\mathrm{d}W^{\varepsilon}(t,x), \qquad t > t'_m, x \in \mathbf{R}^2, \tag{6.11}$$

$$w^{(m)}(t'_m,x) = \int G_{t'_m-t_m}(X-y)w^{(m-1)}(t_m,y)\,\mathrm{d}y, \qquad x \in \mathbf{R}^2. \tag{6.12}$$

Therefore, $w^{(m)}$ solves (1.1)–(1.2) but with constant initial condition at time t'_m. Recall that X is fixed which is our reference spatial point. We note (recalling (6.6) and (6.7)) that (whenever $m \geq t_{M_1(\varepsilon,T)}+1$) we have $t'_m \geq t_m \geq \cdots \geq t'_{M_1(\varepsilon,T)+1} \geq t_{M_1(\varepsilon,T)+1} \geq 0$ and so the initial conditions (6.12) are inductively well-defined. We also emphasize that the function $w^{(m)}$ depends on the parameters ε, a, T, X, and the simplified notation $w^{(m)} = w^{(m)}_{\varepsilon,a,T,X}$ will be used when there is no confusion. We will make the dependence explicit when needed. It is worth mentioning that for those $T \leq \varepsilon^{2-\lambda_\varepsilon}$, we only have one element in the chain which is $w^{(1)} = u_{\varepsilon,a}$.

To compare $u_{\varepsilon,a}$ with $w^{(m)}$, it turns out to be convenient to introduce another sequence of functions $\{\tilde{w}^{(m)}\}_{m=M_1(\varepsilon,T),\ldots,M_2(\varepsilon)}$. Define $\{\tilde{w}^{(m)}(t,x) : t \geq t'_m, x \in \mathbf{R}^2\}$ as the solution to

$$\mathrm{d}\tilde{w}^{(m)}(t,y) = \frac{1}{2}\Delta\tilde{w}^{(m)}(t,x)\mathrm{d}t + \frac{\mathbf{1}_{\mathbf{R}\setminus[t_{m+1},t'_{m+1}]}(t)}{\sqrt{\log\varepsilon^{-1}}}\sigma(\tilde{w}^{(m)}(t,x))\mathrm{d}W^{\varepsilon}(t,x), \quad t > t'_m, x \in \mathbf{R}^2, \tag{6.13}$$

$$\tilde{w}^{(m)}(t'_m,x) = \int G_{t'_m-t_m}(X-y)\tilde{w}^{(m-1)}(t_m,y)\,\mathrm{d}y, \qquad x \in \mathbf{R}^2, m \geq M_1(\varepsilon,T)+1, \tag{6.14}$$

$$\tilde{w}^{(M_1(\varepsilon,T))}(0,x) = a. \tag{6.15}$$

For each $m \geq M_1(\varepsilon,T)+1$, we note that since $\tilde{w}^{(m-1)}$ satisfies the unforced heat equation on the time interval $[t_m,t'_m]$, the initial condition (6.14) can be rewritten as

$$\tilde{w}^{(m)}(t'_m,x) = \tilde{w}^{(m-1)}(t'_m,X). \tag{6.16}$$

We also have the following lemma relating $\tilde{w}^{(m)}$ to $w^{(m)}$.

Lemma 6.1. *For all $m \in \{M_1(\varepsilon,T),\ldots,M_2(\varepsilon)\}$, we have $w^{(m)}(t,x) = \tilde{w}^{(m)}(t,x)$ for all $t \in [t'_m \vee 0, t_{m+1}]$ and all $x \in \mathbf{R}^2$.*

Proof. The proof is by induction on m. For $m = M_1(\varepsilon,T)$, by (6.13), (6.15), and (6.10), we see that $\tilde{w}^{(M_1(\varepsilon,T))} = u_{\varepsilon,a} = w^{(M_1(\varepsilon,T))}$ on $[0,t_{M_1(\varepsilon,T)+1}]\times\mathbf{R}^2$. For the inductive step, if $m \geq M_1(\varepsilon,T)-1$ and we assume that $w^{(m-1)}(t,x) = \tilde{w}^{(m-1)}(t,x)$ for all $(t,x) \in [t'_{m-1}\vee 0, t_m]\times\mathbf{R}^2$, then this in particular means that $w^{(m-1)}(t_m,\cdot) = \tilde{w}^{(m-1)}(t_m,\cdot)$. This means that the initial conditions (6.12) for $w^{(m)}$ and (6.14) for $\tilde{w}^{(m)}$ (both imposed at time t'_m) agree. Since the evolution equations (6.11) and (6.13) also agree on the "noisy" time interval $[t'_m,t_{m+1}]$, this implies that $w^{(m)} = \tilde{w}^{(m)}$ on the time interval $[t'_m,t_{m+1}]$ as well. □

By Lemma 6.1 and (6.16), we see that the initial condition (6.12) is equivalent to

$$w^{(m)}(t'_m,x) = \tilde{w}^{(m-1)}(t'_m,X). \tag{6.17}$$

Thus, for each $m \ge M_1(\varepsilon,T)$, we initiate $w^{(m)}$ and $\tilde{w}^{(m)}$ with the same data at $t = t'_m$, with $w^{(m)}$ solving the original stochastic heat equation for $t > t'_m$ and $\tilde{w}^{(m)}$ solving the equation with the noise turned off in $[t_{m+1},t'_{m+1}]$.

Our goal in this section is to estimate the approximation error $|w^{(m)}(t,x) - u_{\varepsilon,a}(t,x)|$ for $t \ge t'_m$, $x \in \mathbf{R}^2$, and $m \in \{M_1(\varepsilon,T),\dots,M_2(\varepsilon)\}$. By definition, we have $w^{(M_1(\varepsilon,T))} = u_{\varepsilon,a}$, thus by applying triangle inequality it suffices to estimate $w^{(m)} - w^{(m-1)}$for each m. We briefly explain below how it will be achieved, by applying the results from Sections 4 and 5. First, through $\tilde{w}^{(m-1)}$ we can write the difference as

$$w^{(m)}(t,x) - w^{(m-1)}(t,x) = [w^{(m)}(t,x) - \tilde{w}^{(m-1)}(t,x)] + [\tilde{w}^{(m-1)}(t,x) - w^{(m-1)}(t,x)], \qquad t \ge t'_m, x \in \mathbf{R}^2.$$

We bound the two terms separately:

1. For the first error term $w^{(m)}(t,x) - \tilde{w}^{(m-1)}(t,x)$, we recall three facts (i) $w^{(m)}(t'_m,\cdot) = \tilde{w}^{(m-1)}(t'_m,X)$; (ii) $\tilde{w}^{(m-1)}$ solves the deterministic heat equation in the interval $[t_m,t'_m]$; (iii) for $t > t'_m$, $w^{(m)}$ and $\tilde{w}^{(m-1)}$ solve the same stochastic heat equation. Therefore, the difference of $w^{(m)}(t,x)$ from $\tilde{w}^{(m-1)}(t,x)$ only comes from replacing the initial data $\tilde{w}^{(m-1)}(t'_m,\cdot)$ by its value at X, which can be quantified by Proposition 5.1.

2. For the second error term $\tilde{w}^{(m-1)}(t,x) - w^{(m-1)}(t,x)$, we have
$$w^{(m-1)}(t'_{m-1},\cdot) = \tilde{w}^{(m-1)}(t'_{m-1},\cdot) = \tilde{w}^{(m-2)}(t'_{m-1},X).$$
The equations satisfied by $w^{(m-1)}$ and $\tilde{w}^{(m-1)}$ in $t > t'_{m-1}$ are the same except that the noise is turned off in $[t_m,t'_m]$ for $\tilde{w}^{(m-1)}$. Therefore, the error only comes from turning off the noise in $[t_m,t'_m]$. This can be quantified by Proposition 4.1.

The following proposition is the main result of the section.

Proposition 6.2. *Suppose that $(C_\varepsilon)_{\varepsilon>0}$ is an arbitrary sequence of numbers such that $C_\varepsilon \to \infty$ as $\varepsilon \downarrow 0$, and that $c \in [0,1)$ is a fixed constant. Define the set*

$$S_{\varepsilon,T_0,c} := \left\{(T,a,k,t,x) : T \in [0,T_0],\, a>0,\, M_1(\varepsilon,T){+}1 \le k \le M_2(\varepsilon),\, t \in [T - c\varepsilon^{k\delta_\varepsilon+\gamma_\varepsilon},T],\, x \in \mathbf{R}^2\right\}.$$

Then we have

$$\lim_{\varepsilon\downarrow 0} \sup_{(T,a,k,t,x)\in S_{\varepsilon,T_0,c}} \frac{\left(\mathbf{E}(u_{\varepsilon,a} - w^{(k)}_{\varepsilon,a,T,X})(t,x)^2\right)^{1/2}}{a(1+C_\varepsilon \varepsilon^{-k\delta_\varepsilon/2}|x-X|)} = 0. \tag{6.18}$$

In order to prove Proposition 6.2, we need the following second moment bound.

Lemma 6.3. *There is a constant $K_1 < \infty$ so that if $T \in [0,T_0]$, $\varepsilon \in (0,\varepsilon_0]$, $m \in \{M_1(\varepsilon,T),\dots,M_2(\varepsilon)\}$, $a > 0$, then we have for all $x \in \mathbf{R}^2$ that*

$$\mathbf{E} w^{(m)}_{\varepsilon,a,T,X}(t'_m,x)^2 \le K_1^2 a^2. \tag{6.19}$$

Proof. Throughout the proof, we will again use the simplified notation $w^{(m)},\tilde{w}^{(m)}$. Consider a fixed m. For all $t \ge t'_{m-1}$, by the mild formulation of the equation satisfied by $\tilde{w}^{(m-1)}$ and (6.16), we have

$$\begin{aligned}\mathbf{E}\tilde{w}^{(m-1)}(t,X)^2 &= \mathbf{E}\tilde{w}^{(m-2)}(t'_{m-1},X)^2 + \frac{1}{\log\varepsilon^{-1}} \int_{[t'_{m-1},t]\setminus[t_m,t'_m]} \iint G_{t-s}(X-y_1)G_{t-s}(X-y_2)G_{\varepsilon^2}(y_1-y_2) \\ &\qquad \cdot \mathbf{E}\left[\sigma(\tilde{w}^{(m-1)}(s,y_1))\sigma(\tilde{w}^{(m-1)}(s,y_2))\right] \mathrm{d}y_1\,\mathrm{d}y_2\,\mathrm{d}s \\ &\le \mathbf{E}\tilde{w}^{(m-2)}(t'_{m-1},X)^2 + \frac{\beta^2}{2\pi\log\varepsilon^{-1}} \int_{[t'_{m-1},t]\setminus[t_m,t'_m]} \frac{\mathbf{E}[\tilde{w}^{(m-1)}(s,X)^2]}{2(t-s)+\varepsilon^2}\,\mathrm{d}s. \end{aligned} \tag{6.20}$$

Here we used the fact that $\tilde{w}^{(m-1)}$ is stationary in the spatial variable. In particular, we have

$$\mathbf{E}\tilde{w}^{(m-1)}(t,X)^2 \le \mathbf{E}\tilde{w}^{(m-2)}(t'_{m-1},X)^2 + \frac{\beta^2}{2\pi\log\varepsilon^{-1}} \int_{t'_{m-1}}^{t} \frac{\mathbf{E}[\tilde{w}^{(m-1)}(s,X)^2]}{2(t-s)+\varepsilon^2}\,\mathrm{d}s,$$

which by Lemma 3.4 (taking there $f(s) = \mathbf{E}\tilde{w}^{(m-1)}(t'_{m-1}+s,X)^2$, and also using Definition 3.5), implies that

$$\mathbf{E}\tilde{w}^{(m-1)}(t,X)^2 \le K_0^2 \mathbf{E}\tilde{w}^{(m-2)}(t'_{m-1},X)^2.$$

Substituting this back into (6.20), taking $t = t'_m$, and recalling (6.17), we have

$$\begin{aligned}
\mathbf{E}w^{(m)}(t'_m,X)^2 &= \mathbf{E}\tilde{w}^{(m-1)}(t'_m,X)^2 \\
&\le \mathbf{E}\tilde{w}^{(m-2)}(t'_{m-1},X)^2 \left(1 + \frac{K_0^2\beta^2}{4\pi\log\varepsilon^{-1}} \int_{t'_{m-1}}^{t_m} \frac{1}{t'_m - s + \varepsilon^2/2}\,\mathrm{d}s\right) \\
&\le \mathbf{E}\tilde{w}^{(m-2)}(t'_{m-1},X)^2 \left(1 + \frac{K_0^2\beta^2}{4\pi\log\varepsilon^{-1}} \log\frac{t'_m - t'_{m-1}}{t'_m - t_m}\right).
\end{aligned} \tag{6.21}$$

The logarithm can be estimated as

$$\begin{aligned}
\log\frac{t'_m - t'_{m-1}}{t'_m - t_m} &= \log\frac{\varepsilon^{(m-1)\delta_\varepsilon+\gamma_\varepsilon} - \varepsilon^{m\delta_\varepsilon+\gamma_\varepsilon}}{\varepsilon^{m\delta_\varepsilon} - \varepsilon^{m\delta_\varepsilon+\gamma_\varepsilon}} = \log\frac{\varepsilon^{\gamma_\varepsilon-\delta_\varepsilon} - \varepsilon^{\gamma_\varepsilon}}{1-\varepsilon^{\gamma_\varepsilon}} \\
&\le \delta_\varepsilon \log\varepsilon^{-1} + \log\frac{\varepsilon^{\gamma_\varepsilon}}{1-\varepsilon^{\gamma_\varepsilon}} \le \delta_\varepsilon\log\varepsilon^{-1},
\end{aligned}$$

where the last inequality is by (6.4). Substituting this back into (6.21), we have

$$\mathbf{E}w^{(m)}(t'_m,X)^2 = \mathbf{E}\tilde{w}^{(m-1)}(t'_m,X)^2 \le \mathbf{E}\tilde{w}^{(m-2)}(t'_{m-1},X)^2\left(1+\frac{K_0^2\beta^2\delta_\varepsilon}{4\pi}\right).$$

Iterating this and recalling (6.9), we have for all $x \in \mathbf{R}^2$,

$$\begin{aligned}
\mathbf{E}w^{(m)}(t'_m,x)^2 = \mathbf{E}w^m(t'_m,X)^2 &\le K_0^2 a^2 \left(1+\frac{K_0^2\beta^2\delta_\varepsilon}{4\pi}\right)^{m-M_1(\varepsilon,T)} \\
&\le K_0^2a^2 \exp\left\{\frac{\beta^2}{4\pi}K_0^2\delta_\varepsilon(m-M_1(\varepsilon,T))\right\} \\
&\le K_0^2a^2\exp\left\{\frac{2-\log_\varepsilon T}{4\pi}\beta^2K_0^2\right\}
\end{aligned}$$

for all $m \le M_2(\varepsilon)$. The exponential on the right-hand side is uniformly bounded over all $T \le T_0$ and all $\varepsilon \in (0,\varepsilon_0]$, so we obtain (6.19). □

Now we can prove Proposition 6.2.

Proof of Proposition 6.2. For any $t \in [T - c\varepsilon^{m\delta_\varepsilon+\gamma_\varepsilon}, T]$, we clearly have $t \ge t'_m$. By Proposition 5.1 and Lemma 6.3, we have

$$\begin{aligned}
\mathbf{E}(w^{(m)} - \tilde{w}^{(m-1)})(t,x)^2 &\le K_0^4\left(\frac{3(t-t'_m)+|x-X|^2}{t'_m - t_m}\right)\mathbf{E}\tilde{w}^{(m-1)}(t'_{m-1},X)^2 \\
&\le K_0^4K_1^2a^2\left(\frac{3(t-t'_m)+|x-X|^2}{t'_m-t_m}\right).
\end{aligned} \tag{6.22}$$

We have $t'_m - t_m = \varepsilon^{m\delta_\varepsilon}(1-\varepsilon^{\gamma_\varepsilon}) \in \left[\frac{1}{2}\varepsilon^{m\delta_\varepsilon}, \varepsilon^{m\delta_\varepsilon}\right]$ by (6.4), and $t - t'_m \le T - t'_m = \varepsilon^{m\delta_\varepsilon+\gamma_\varepsilon}$ by (6.6), so (6.22) yields

$$\mathbf{E}(w^{(m)} - \tilde{w}^{(m-1)})(t,x)^2 \le K_0^4 K_1^2 a^2 \left(6\varepsilon^{\gamma_\varepsilon} + 2\varepsilon^{-m\delta_\varepsilon}|x-X|^2\right). \tag{6.23}$$

On the other hand, by Proposition 4.1 and Lemma 6.3, we have for all $t \ge t'_m$ that

$$\begin{aligned}\mathbf{E}(\tilde{w}^{(m-1)} - w^{(m-1)})(t,x)^2 &\le \frac{\beta^2 K_0^4}{4\pi\log\varepsilon^{-1}}\left(\log\frac{t-t_m+\varepsilon^2/2}{t-t'_m+\varepsilon^2/2} + K_0^2\right)\mathbf{E}\tilde{w}^{(m-2)}(t'_{m-1},X)^2 \\ &\le \frac{\beta^2 K_0^4 K_1^2 a^2}{4\pi\log\varepsilon^{-1}}\left(\log\frac{t-t_m+\varepsilon^2/2}{t-t'_m+\varepsilon^2/2} + K_0^2\right).\end{aligned} \tag{6.24}$$

We have $t - t_m \le T - t_m = \varepsilon^{m\delta_\varepsilon}$ and $t - t'_m \ge T - c\varepsilon^{m\delta_\varepsilon+\gamma_\varepsilon} - t'_m = (1-c)\varepsilon^{m\delta_\varepsilon+\gamma_\varepsilon}$, so (6.24) gives us

$$\mathbf{E}(\tilde{w}^{(m-1)} - w^{(m-1)})(t,x)^2 \le \frac{\beta^2}{4\pi}K_0^4 K_1^2 a^2\left(\gamma_\varepsilon + \frac{\log\frac{1}{1-c} + K_0^2}{\log\varepsilon^{-1}}\right). \tag{6.25}$$

Iterating (6.23) and (6.25) and using the triangle inequality, we get

$$\begin{aligned}&\left(\mathbf{E}(w^{(k)} - u_{\varepsilon,a})(t,x)^2\right)^{1/2} \\ &\le K_0^2 K_1 a \sum_{m=M_1(\varepsilon,T)+1}^{k}\left[\sqrt{6}\varepsilon^{\gamma_\varepsilon/2} + \sqrt{2}\varepsilon^{-m\delta_\varepsilon/2}|x-X| + \frac{\beta}{2\sqrt{\pi}}\left(\gamma_\varepsilon^{1/2} + \sqrt{\frac{\log\frac{1}{1-c}+K_0^2}{\log\varepsilon^{-1}}}\right)\right] \\ &\le K_0^2 K_1 a\left[(2-\log_\varepsilon T)\delta_\varepsilon^{-1}\left(\sqrt{6}\varepsilon^{\gamma_\varepsilon/2} + \frac{\beta\gamma_\varepsilon^{1/2}}{2\sqrt{\pi}} + \frac{\sqrt{\log\frac{1}{1-c}}+K_0}{\sqrt{\log\varepsilon^{-1}}}\right) + 2^{3/2}|x-X|\varepsilon^{-k\delta_\varepsilon/2}\right],\end{aligned}$$

where in the last inequality we used (6.9) and (6.4). Therefore, we have (with C_ε, as in the statement of the proposition, an arbitrary sequence so that $C_\varepsilon \to \infty$ as $\varepsilon \downarrow 0$)

$$\frac{\left(\mathbf{E}(u_{\varepsilon,a} - w^{(k)})(t,x)^2\right)^{1/2}}{a(1+C_\varepsilon\varepsilon^{-k\delta_\varepsilon/2}|x-X|)} \le K_0^2 K_1\left[(2-\log_\varepsilon T)\delta_\varepsilon^{-1}\left(\sqrt{6}\varepsilon^{\gamma_\varepsilon/2} + \frac{\beta\gamma_\varepsilon^{1/2}}{2\sqrt{\pi}} + \frac{\sqrt{\log\frac{1}{1-c}}+K_0}{\sqrt{\log\varepsilon^{-1}}}\right) + 2^{3/2}C_\varepsilon^{-1}\right].$$

The first summand in the square brackets goes to 0 as $\varepsilon \downarrow 0$ by (6.1) and (6.2), and since we assumed that $C_\varepsilon \to \infty$, we obtain (6.18). ◻

7 The discrete martingale

The key advantage of the approximation carried out in Proposition 6.2 is that we now have an essentially one-dimensional problem. Note from the definitions (6.11)–(6.14), and also (6.17) and the white-in-time nature of the noise, that if we (fix once and for all $X \in \mathbf{R}^2$ and) define

$$\begin{aligned}&Y_{\varepsilon,a,T}(M_1(\varepsilon,T)) = a; \\ &\qquad Y_{\varepsilon,a,T}(m) = w^{(m)}_{\varepsilon,a,T,X}(t'_m,X) = \tilde{w}^{(m-1)}_{\varepsilon,a,T,X}(t'_m,X), \qquad M_1(\varepsilon,T)+1 \le m \le M_2(\varepsilon),\end{aligned}$$

then the process $\{Y_{\varepsilon,a,T}(m)\}_{m=M_1(\varepsilon,T),\ldots,M_2(\varepsilon)}$ is a Markov chain and a martingale (both with respect to its own filtration). The key point is that $w^{(m)}_{\varepsilon,a,T,X}$ evolves with spatially-constant initial condition $Y_{\varepsilon,a,T}(m)$, driven by the noise that is independent of the past. Thus $Y_{\varepsilon,a,T}(m+1)$ depends on the past only via $Y_{\varepsilon,a,T}(m)$. Moreover the expectation of $Y_{\varepsilon,a,T}(m+1)$ conditional on $Y_{\varepsilon,a,T}(m)$ is simply $Y_{\varepsilon,a,T}(m)$ due to the fact that, when starting from constant initial data, the stochastic heat equation (with the noise either on or off) preserves expectations. Recall that in the case of (very small) $T\in[0,\varepsilon^{2-\lambda_\varepsilon}]$, we have defined $M_1(\varepsilon,T)=M_2(\varepsilon)=1$, and in this case we simply let $Y_{\varepsilon,a,T}(M_1(\varepsilon,T))=Y_{\varepsilon,a,T}(M_2(\varepsilon))=u_{\varepsilon,a}(T,X)$.

7.1 Approximating the one-point SPDE solution by the Markov chain

In this section we show that $Y_{\varepsilon,a,T}(M_2(\varepsilon))$, at its terminal time $M_2(\varepsilon)$, is a good approximation for $u_{\varepsilon,a}(T,X)$ (in fact, for $u_{\varepsilon,a}(T,x)$ if x is close to X). Most of the work has already been done in Proposition 6.2.

Proposition 7.1. *We have*

$$\lim_{\varepsilon\downarrow 0}\sup_{\substack{T\in[0,T_0]\\ a>0,x\in\mathbf{R}^2}}\frac{\left(\mathbf{E}(Y_{\varepsilon,a,T}(M_2(\varepsilon))-u_{\varepsilon,a}(T,x))^2\right)^{1/2}}{a(1+\varepsilon^{-1}|x-X|)}=0. \tag{7.1}$$

Proof. By Proposition 6.2 (choosing $C_\varepsilon=\varepsilon^{-\zeta_\varepsilon/2}\to\infty$ by (6.3)), we have

$$\lim_{\varepsilon\downarrow 0}\sup_{\substack{T\in[0,T_0]\\ a>0,x\in\mathbf{R}^2}}\frac{\left(\mathbf{E}(u_{\varepsilon,a}-w^{(M_2(\varepsilon))}_{\varepsilon,a,T,X})(T,x)^2\right)^{1/2}}{a(1+\varepsilon^{-\frac{1}{2}(M_2(\varepsilon)\delta_\varepsilon+\zeta_\varepsilon)}|x-X|)}=0. \tag{7.2}$$

By (6.8), we have $\frac{1}{2}(M_2(\varepsilon)\delta_\varepsilon+\zeta_\varepsilon)\le 1$. Therefore, (7.2) implies that

$$\lim_{\varepsilon\downarrow 0}\sup_{\substack{T\in[0,T_0]\\ a>0,x\in\mathbf{R}^2}}\frac{\left(\mathbf{E}(u_{\varepsilon,a}-w^{(M_2(\varepsilon))}_{\varepsilon,a,T,X})(T,x)^2\right)^{1/2}}{a(1+\varepsilon^{-1}|x-X|)}=0. \tag{7.3}$$

Moreover, by Proposition 3.6, Lemma 6.3 and the fact that $T-t'_{M_2(\varepsilon)}=s'_{M_2(\varepsilon)}<s_{M_2(\varepsilon)}\le\varepsilon^{2-\zeta_\varepsilon-\delta_\varepsilon}$, we have

$$\begin{aligned}
&\frac{1}{a}\left(\mathbf{E}(w^{(M_2(\varepsilon))}_{\varepsilon,a,T,X}(T,x)-Y_{\varepsilon,a,T}(M_2(\varepsilon)))^2\right)^{1/2}\\
&\quad\le\frac{\beta K_0}{2a\sqrt{\pi}}\left(\mathbf{E}w^{(M_2(\varepsilon))}_{\varepsilon,a,T,X}(t'_{M_2(\varepsilon)},X)^2\right)^{1/2}\left(\frac{\log(1+2\varepsilon^{-2}(T-t'_{M_2(\varepsilon)}))}{\log\varepsilon^{-1}}\right)^{1/2}\\
&\quad\le\frac{\beta K_0K_1}{2\sqrt{\pi}}\sqrt{\frac{\log(1+2\varepsilon^{-\zeta_\varepsilon-\delta_\varepsilon})}{\log\varepsilon^{-1}}},
\end{aligned}$$

and the quantity on the right side goes to 0 uniformly in T,a,x by (6.1) and (6.3). This, along with (7.3), implies (7.1). □

The following spatial regularity statement for $u_{\varepsilon,a}(T,\cdot)$ is a consequence of Proposition 7.1, so we record it here for future use.

Corollary 7.2. *We have*

$$\lim_{\varepsilon\downarrow 0}\sup_{\substack{T\in[0,T_0],a>0\\ x_1,x_2\in\mathbf{R}^2}}\frac{\left(\mathbf{E}(u_{\varepsilon,a}(T,x_1)-u_{\varepsilon,a}(T,x_2))^2\right)^{1/2}}{a(1+\varepsilon^{-1}|x_1-x_2|)}=0.$$

Proof. By spatial homogeneity, we can assume that $x_1 = X$. Then the result follows immediately by writing

$$\left(\mathbf{E}(u_{\varepsilon,a}(T,x_1) - u_{\varepsilon,a}(T,x_2))^2\right)^{1/2} \leq \sum_{i=1}^{2} \left(\mathbf{E}(u_{\varepsilon,a}(T,x_i) - Y_{\varepsilon,a,T}(M_2(\varepsilon)))^2\right)^{1/2}$$

and applying Proposition 7.1 to both terms. □

7.2 The martingale differences

The approximation result in Proposition 7.1 motivates us to study the discrete martingale $\{Y_{\varepsilon,a,T}(m)\}_m$. Our ultimate goal will be to show that it approximates a continuous martingale (coming from a solution to (1.5)–(1.7)) as $\varepsilon \downarrow 0$. We will use the martingale problem approach as explained in [41, Section 11.2], and en route it will be important to understand some statistical properties of the increments $Y_{\varepsilon,a,T}(m) - Y_{\varepsilon,a,T}(m-1)$ conditional on $Y_{\varepsilon,a,T}(m-1)$, a task to which we now set ourselves. The first observation is that, due to the independence of $\mathrm{d}W^\varepsilon$ on disjoint time intervals, if we define

$$Z_{\varepsilon,a,m} = \int G_{s_m - s'_m}(X - z) u_{\varepsilon,a}(s'_{m-1} - s_m, z)\,\mathrm{d}z, \qquad M_1(\varepsilon,T) + 1 \leq m \leq M_2(\varepsilon), \tag{7.4}$$

(with s_k, s'_k defined as in (6.5)) then

$$\mathrm{Law}[Y_{\varepsilon,a,T}(m) \mid Y_{\varepsilon,a,T}(m-1) = b] = \mathrm{Law}\, Z_{\varepsilon,b,m}. \tag{7.5}$$

This can be seen by noting that the evolution equations for $u_{\varepsilon,a}$ and $\tilde{w}^{(m)}$ are the same, and that $\tilde{w}^{(m)}$ is started with constant initial condition equal to $Y_{\varepsilon,a,T}(m-1)$.

7.2.1 Martingale difference variances

Our first interest is in the conditional variance $\mathrm{Var}[Y_{\varepsilon,a,T}(m) \mid Y_{\varepsilon,a,T}(m-1) = b] = \mathrm{Var}\, Z_{\varepsilon,b,m}$, and we proceed to study this quantity. The first step is to approximate it by a simpler quantity using the regularity established in Corollary 7.2. An important role will be played by the function $J_\varepsilon : (-\infty, 2 + \log_{\varepsilon^{-1}} T_0] \times \mathbf{R}_{\geq 0} \to \mathbf{R}_{\geq 0}$ defined by

$$J_\varepsilon(q,a) = \frac{1}{2\sqrt{\pi}} (\mathbf{E}\sigma(u_{\varepsilon,a}(\varepsilon^{2-q}, x))^2)^{1/2}. \tag{7.6}$$

As $u_{\varepsilon,a}$ is stationary in the spatial variable, the r.h.s. of (7.6) does not depend on x. Here

$$q \in (-\infty, 2 + \log_{\varepsilon^{-1}} T_0] \text{ corresponds to } \varepsilon^{2-q} \in (0, T_0],$$

i.e., we parameterize J_ε in time on the exponential scale discussed in Section 1.1. This section has two main results. First, we show how to use J_ε to approximate $\mathrm{Var}\, Z_{\varepsilon,b,m}$:

Proposition 7.3. *We have*

$$\lim_{\varepsilon \downarrow 0} \sup_{\substack{T \in [0,T_0], a > 0, \\ M_1(\varepsilon,T)+1 \leq m \leq M_2(\varepsilon)}} a^{-2} \left| \delta_\varepsilon^{-1} \mathrm{Var}\, Z_{\varepsilon,a,m} - J_\varepsilon(2 - (m-1)\delta_\varepsilon, a)^2 \right| = 0. \tag{7.7}$$

Also, we will prove the following compactness result for the family $\{J_\varepsilon\}_\varepsilon$, which will help us in our limit procedure:

Proposition 7.4. *For any sequence $\varepsilon_k \downarrow 0$, there is a subsequence $\varepsilon_{k_\ell} \downarrow 0$ and a continuous function $J : [0,2] \times \mathbf{R}_{\geq 0} \to \mathbf{R}_{\geq 0}$ so that*

$$\lim_{\ell \to \infty} J_{\varepsilon_{k_\ell}}|_{[0,2] \times \mathbf{R}_{\geq 0}} = J \tag{7.8}$$

uniformly on compact subsets of $[0,2] \times \mathbf{R}_{\geq 0}$.

As we assumed that $T_0 \geq 1$, each J_ε is indeed defined on $[0,2]\times\mathbf{R}_{\geq 0}$. We will prove Proposition 7.4 first, since the intermediate results will be useful in the proof of Proposition 7.3. We need two preparatory lemmas, addressing the regularity of J_ε in q and in a. First we address the regularity in q.

Lemma 7.5. *For all $\varepsilon, a > 0$ and $q_1, q_2 \in (-\infty, 2+\log_{\varepsilon^{-1}} T_0]$, we have*

$$|J_\varepsilon(q_2,a) - J_\varepsilon(q_1,a)| \leq \frac{a\beta^2 K_0^2}{4\pi}\left(|q_2-q_1|^{1/2} + K_0(\log\varepsilon^{-1})^{-1/2}\right). \tag{7.9}$$

Proof. Assume $q_1 \leq q_2$. Define $I_\varepsilon = [0, \varepsilon^{2-q_2} - \varepsilon^{2-q_1}]$. We can write

$$\begin{aligned}|J_\varepsilon(q_1,a) - J_\varepsilon(q_2,a)| &= \frac{1}{2\sqrt{\pi}}\left|(\mathbf{E}\sigma(u^{I_\varepsilon}_{\varepsilon,a}(\varepsilon^{2-q_2},x))^2)^{1/2} - (\mathbf{E}\sigma(u_{\varepsilon,a}(\varepsilon^{2-q_2},x))^2)^{1/2}\right| \\ &\leq \frac{\beta}{2\sqrt{\pi}}\left(\mathbf{E}\left(u^{I_\varepsilon}_{\varepsilon,a}(\varepsilon^{2-q_2},x) - u_{\varepsilon,a}(\varepsilon^{2-q_2},x)\right)^2\right)^{1/2}.\end{aligned}$$

In the first equality we used the fact that

$$u^{I_\varepsilon}_{\varepsilon,a}(\varepsilon^{2-q_2},x) \overset{\text{law}}{=} u_{\varepsilon,a}(\varepsilon^{2-q_1},x),$$

where $u^{I_\varepsilon}_{\varepsilon,a}$ is defined as in (1.26)–(1.27), i.e., the noise is turned off in I_ε. Now we apply Proposition 4.1 with $t = \varepsilon^{2-q_2}$, $A = \emptyset$, $\tau_1 = 0$, and $\tau_2 = \varepsilon^{2-q_2} - \varepsilon^{2-q_1}$ to obtain

$$\begin{aligned}|J_\varepsilon(q_1,a) - J_\varepsilon(q_2,a)| &\leq \frac{a\beta^2 K_0^2}{4\pi\sqrt{\log\varepsilon^{-1}}}\left(\sqrt{\log\frac{\varepsilon^{2-q_2}+\varepsilon^2/2}{\varepsilon^{2-q_1}+\varepsilon^2/2}} + K_0\right) &(7.10)\\ &\leq \frac{a\beta^2 K_0^2}{4\pi}\left(|q_2-q_1|^{1/2} + K_0(\log\varepsilon^{-1})^{-1/2}\right), &(7.11)\end{aligned}$$

as claimed. □

Now we address the regularity of J_ε in a. Later on, we will also use the following result to prove that (1.15) is satisfied for the limits of $\{J_\varepsilon\}_\varepsilon$ as $\varepsilon \downarrow 0$. Thus we need the explicit constant in the middle expression of (7.12).

Lemma 7.6. *For all $\varepsilon \in (0,\varepsilon_0]$, $q \in (-\infty, 2+\log_{\varepsilon^{-1}} T_0]$, and $a_1, a_2 \geq 0$, we have*

$$|J_\varepsilon(q,a_2) - J_\varepsilon(q,a_1)| \leq \left(\frac{4\pi}{\beta^2} - \frac{\log(1+2\varepsilon^{-q})}{\log\varepsilon^{-1}}\right)^{-1/2}|a_2-a_1| \leq \frac{\beta K_0}{2\sqrt{\pi}}|a_2-a_1|. \tag{7.12}$$

In particular, for all $a > 0$, we have

$$|J_\varepsilon(q,a)| \leq \frac{\beta a K_0}{2\sqrt{\pi}}. \tag{7.13}$$

Proof. We have

$$\begin{aligned}|J_\varepsilon(q,a_1) - J_\varepsilon(q,a_2)| &= \frac{1}{2\sqrt{\pi}}\left|(\mathbf{E}\sigma(u_{\varepsilon,a_1}(\varepsilon^{2-q},x))^2)^{1/2} - (\mathbf{E}\sigma(u_{\varepsilon,a_2}(\varepsilon^{2-q},x))^2)^{1/2}\right| \\ &\leq \frac{1}{2\sqrt{\pi}}\left(\mathbf{E}[\sigma(u_{\varepsilon,a_1}(\varepsilon^{2-q},x)) - \sigma(u_{\varepsilon,a_2}(\varepsilon^{2-q},x))]^2\right)^{1/2} \\ &\leq \frac{\beta}{2\sqrt{\pi}}\left(\mathbf{E}[u_{\varepsilon,a_1}(\varepsilon^{2-q},x) - u_{\varepsilon,a_2}(\varepsilon^{2-q},x)]^2\right)^{1/2},\end{aligned}$$

and then the first inequality in (7.12) follows from (3.4) with the explicit constant (3.6). The second inequality in (7.12) is then just (3.8). The bound (7.13) comes from (7.12) with $a_2 = a$ and $a_1 = 0$. □

Given the regularity results in Lemmas 7.5 and 7.6, the compactness of the family (J_ε) is straightforward.

Proof of Proposition 7.4. By Lemmas 7.5 and 7.6, along with a simple modification of the Arzelà–Ascoli theorem to account for the second term on the r.h.s. of (7.9) (see e.g. [20, Lemma A.4]), we can extract a suitable subsequence and pass to the limit on any rectangular subset of $[0,2]\times\mathbf{R}_{\ge 0}$ of the form $[0,2]\times[0,M]$, with $M>0$. Sending $M\to\infty$ so that the rectangles exhaust $[0,2]\times\mathbf{R}_{\ge 0}$ and using a diagonalization argument, we obtain the desired limit and convergence (7.8). □

Now we turn to the proof of Proposition 7.3. We first prove the following intermediate result.

Lemma 7.7. *Define*

$$\begin{aligned} V_{\varepsilon,a,m} &= \frac{1}{\log\varepsilon^{-1}}\int_0^{s'_{m-1}-s_m}\frac{J_\varepsilon(2-\log_\varepsilon s,a)^2}{s'_{m-1}-s'_m-s+\varepsilon^2/2}\,\mathrm{d}s \\ &= \frac{1}{2\pi\log\varepsilon^{-1}}\int_0^{s'_{m-1}-s_m}\frac{\mathbf{E}\sigma(u_{\varepsilon,a}(s,X))^2}{2(s'_{m-1}-s'_m-s)+\varepsilon^2}\,\mathrm{d}s. \end{aligned} \tag{7.14}$$

Then we have, for any fixed $T_0<\infty$, that

$$\lim_{\varepsilon\downarrow 0}\sup_{\substack{T\in[0,T_0]\\ M_1(\varepsilon,T)+1\le m\le M_2(\varepsilon)}}\frac{|V_{\varepsilon,a,m}-\operatorname{Var}Z_{\varepsilon,a,m}|}{a^2\delta_\varepsilon}=0. \tag{7.15}$$

Proof. We can first write (recalling (7.4))

$$\begin{aligned} Z_{\varepsilon,a,m} &= a+\frac{1}{\sqrt{\log\varepsilon^{-1}}}\int G_{s_m-s'_m}(X-z)\int_0^{s'_{m-1}-s_m}\int G_{s'_{m-1}-s_m-s}(z-y)\sigma(u_{\varepsilon,a}(s,y))\,\mathrm{d}W^\varepsilon(s,y)\,\mathrm{d}z \\ &= a+\frac{1}{\sqrt{\log\varepsilon^{-1}}}\int_0^{s'_{m-1}-s_m}\int G_{s'_{m-1}-s'_m-s}(X-y)\sigma(u_{\varepsilon,a}(s,y))\,\mathrm{d}W^\varepsilon(s,y). \end{aligned} \tag{7.16}$$

Therefore, we have

$$\operatorname{Var}Z_{\varepsilon,a,m}=\frac{1}{\log\varepsilon^{-1}}\int_0^{s'_{m-1}-s_m}\iint G_{\varepsilon^2}(y_1-y_2)\mathbf{E}\prod_{i=1}^2\Big(G_{s'_{m-1}-s'_m-s}(X-y_i)\sigma(u_{\varepsilon,a}(s,y_i))\Big)\,\mathrm{d}y_1\,\mathrm{d}y_2\,\mathrm{d}s. \tag{7.17}$$

Now we have, by spatial homogeneity, that

$$\mathbf{E}\sigma(u_{\varepsilon,a}(s,y_1))\sigma(u_{\varepsilon,a}(s,y_2))=\mathbf{E}\sigma(u_{\varepsilon,a}(s,X))\sigma(u_{\varepsilon,a}(s,X+y_1-y_2)).$$

We also have (using the Cauchy–Schwarz inequality) that

$$\begin{aligned} &\left|\mathbf{E}\sigma(u_{\varepsilon,a}(s,y_1))\sigma(u_{\varepsilon,a}(s,y_2))-\mathbf{E}\sigma(u_{\varepsilon,a}(s,y_1))^2\right| \\ &\quad\le \mathbf{E}\sigma(u_{\varepsilon,a}(s,y_1))|\sigma(u_{\varepsilon,a}(s,y_1))-\sigma(u_{\varepsilon,a}(s,y_2))| \\ &\quad\le \beta^2\left(\mathbf{E}u_{\varepsilon,a}(s,y_1)^2\right)^{1/2}\left(\mathbf{E}[u_{\varepsilon,a}(s,y_1)-u_{\varepsilon,a}(s,y_2)]^2\right)^{1/2}, \end{aligned}$$

so by (3.5) and Corollary 7.2 we have a function f satisfying $\lim_{\varepsilon\downarrow 0}f(\varepsilon)=0$ and

$$\sup_{\substack{s\in[0,T_0]\\ y_1,y_2\in\mathbf{R}^2}}\left|\mathbf{E}\sigma(u_{\varepsilon,a}(s,y_1))\sigma(u_{\varepsilon,a}(s,y_2))-\mathbf{E}\sigma(u_{\varepsilon,a}(s,y_1))^2\right|\le a^2(1+\varepsilon^{-1}|y_1-y_2|)f(\varepsilon) \tag{7.18}$$

for all $y_1, y_2 \in \mathbf{R}^2$ and all $a \geq 0$. Now we note that

$$\begin{aligned}
&\frac{1}{\log \varepsilon^{-1}} \int_0^{s'_{m-1}-s_m} \iint G_{s'_{m-1}-s'_m-s}(X-y) G_{s'_{m-1}-s'_m-s}(X-y') G_{\varepsilon^2}(y-y') \mathbf{E}\sigma(u_{\varepsilon,a}(s,y))^2 \, \mathrm{d}y \, \mathrm{d}y' \, \mathrm{d}s \\
&\qquad = \frac{1}{\log \varepsilon^{-1}} \int_0^{s'_{m-1}-s_m} \int G_{s'_{m-1}-s'_m-s}(X-y) G_{s'_{m-1}-s'_m-s+\varepsilon^2}(X-y) \mathbf{E}\sigma(u_{\varepsilon,a}(s,y))^2 \, \mathrm{d}y \, \mathrm{d}s \\
&\qquad = \frac{1}{2\pi \log \varepsilon^{-1}} \int_0^{s'_{m-1}-s_m} \frac{\mathbf{E}\sigma(u_{\varepsilon,a}(s,X))^2}{2(s'_{m-1}-s'_m-s)+\varepsilon^2} \, \mathrm{d}s = V_{\varepsilon,a,m}.
\end{aligned} \tag{7.19}$$

where in the second-to-last identity we used spatial homogeneity. Subtracting (7.17) and (7.19) and applying (7.18), we have

$$\begin{aligned}
&|V_{\varepsilon,a,m} - \operatorname{Var} Z_{\varepsilon,a,m}| \\
&\quad \leq \frac{f(\varepsilon)a^2}{\log \varepsilon^{-1}} \int_0^{s'_{m-1}-s_m} \iint \left(\prod_{i=1}^2 G_{s'_{m-1}-s'_m-s}(X-y_i) \right) G_{\varepsilon^2}(y_1-y_2)(1+\varepsilon^{-1}|y_1-y_2|) \, \mathrm{d}y_1 \, \mathrm{d}y_2 \, \mathrm{d}s.
\end{aligned} \tag{7.20}$$

If we define $h(r) = (2\pi)^{-1} e^{-\frac{r^2}{2}}(1+r)$ for $r \geq 0$, then the last double integral is equal to

$$\begin{aligned}
&\iint \left(\prod_{i=1}^2 G_{s'_{m-1}-s'_m-s}(X-y_i) \right) \varepsilon^{-2} h(\varepsilon^{-1}|y_1-y_2|) \, \mathrm{d}y_1 \, \mathrm{d}y_2 \\
&\qquad \leq \left(\int \varepsilon^{-2} h(\varepsilon^{-1}|y|) \, \mathrm{d}y \right) \int G_{s'_{m-1}-s'_m-s}(X-y)^2 \, \mathrm{d}y = \frac{\int h(|y|) \, \mathrm{d}y}{4\pi(s'_{m-1}-s'_m-s)},
\end{aligned}$$

where the inequality is Young's convolution inequality. Substituting this back into (7.20), we have

$$\begin{aligned}
|V_{\varepsilon,a,m} - \operatorname{Var} Z_{\varepsilon,a,m}| &\leq \frac{f(\varepsilon)a^2 \int h(|y|) \, \mathrm{d}y}{4\pi \log \varepsilon^{-1}} \int_0^{s'_{m-1}-s_m} \frac{1}{s'_{m-1}-s'_m-s} \, \mathrm{d}s \\
&= \frac{f(\varepsilon)a^2 \int h(|y|) \, \mathrm{d}y}{4\pi \log \varepsilon^{-1}} \log \frac{s'_{m-1}-s'_m}{s_m-s'_m} = \frac{f(\varepsilon)a^2 \int h(|y|) \, \mathrm{d}y}{4\pi \log \varepsilon^{-1}} \log \frac{\varepsilon^{\gamma_\varepsilon-\delta_\varepsilon}-\varepsilon^{\gamma_\varepsilon}}{1-\varepsilon^{\gamma_\varepsilon}}.
\end{aligned}$$

From this and (6.1) we obtain (7.15). □

In Lemma 7.5 we derived the regularity of J_ε in time (where time is taken on an exponential scale). Since $\log_\varepsilon s$ varies slowly on most of the interval $[0, s'_{m-1} - s_m]$, it should be plausible that we could approximate $J_\varepsilon(2 - \log_\varepsilon s, a)^2$ by

$$J_\varepsilon(2 - \log_\varepsilon s'_{m-1}, a)^2 = J_\varepsilon(2-(m-1)\delta_\varepsilon - \gamma_\varepsilon, a)^2 \approx J_\varepsilon(2-(m-1)\delta_\varepsilon, a)^2$$

in (7.14). Indeed we can, and that is how we will prove Proposition 7.3.

Proof of Proposition 7.3. In light of (7.15), Lemma 7.5 and (7.13), it suffices to show that

$$\lim_{\varepsilon \downarrow 0} \sup_{\substack{T \in [0,T_0] \\ M_1(\varepsilon,T)+1 \leq m \leq M_2(\varepsilon)}} a^{-2} \left| J_\varepsilon(2-(m-1)\delta_\varepsilon - \gamma_\varepsilon, a)^2 - \delta_\varepsilon^{-1} V_{\varepsilon,a,m} \right| = 0.$$

We will compare both $J_\varepsilon(2-(m-1)\delta_\varepsilon - \gamma_\varepsilon, a)^2$ and $\delta_\varepsilon^{-1} V_{\varepsilon,a,m}$ to the intermediate quantity

$$\begin{aligned}
\tilde{V}_{\varepsilon,a,m} &:= \frac{J_\varepsilon(2-(m-1)\delta_\varepsilon - \gamma_\varepsilon, a)^2}{\log \varepsilon^{-1}} \int_0^{s'_{m-1}-s_m} \frac{1}{s'_{m-1}-s'_m-s+\varepsilon^2/2} \, \mathrm{d}s \\
&= \frac{J_\varepsilon(2-(m-1)\delta_\varepsilon - \gamma_\varepsilon, a)^2}{\log \varepsilon^{-1}} \log \frac{\varepsilon^{\gamma_\varepsilon-\delta_\varepsilon} - \varepsilon^{\gamma_\varepsilon} + \varepsilon^{2-m\delta_\varepsilon}/2}{1-\varepsilon^{\gamma_\varepsilon}+\varepsilon^{2-m\delta_\varepsilon}/2}.
\end{aligned}$$

First, we have

$$\begin{aligned}&\left|J_\varepsilon(2-(m-1)\delta_\varepsilon-\gamma_\varepsilon,a)^2-\delta_\varepsilon^{-1}\tilde{V}_{\varepsilon,a,m}\right|\\&\qquad=J_\varepsilon(2-(m-1)\delta_\varepsilon-\gamma_\varepsilon,a)^2\left(1-\frac{1}{\delta_\varepsilon\log\varepsilon^{-1}}\log\frac{\varepsilon^{\gamma_\varepsilon-\delta_\varepsilon}-\varepsilon^{\gamma_\varepsilon}+\varepsilon^{2-m\delta_\varepsilon}/2}{1-\varepsilon^{\gamma_\varepsilon}+\varepsilon^{2-m\delta_\varepsilon}/2}\right),\end{aligned}$$

and from this, (7.13) of Lemma 7.6, (6.1), and (6.8) we have

$$\lim_{\varepsilon\downarrow 0}\sup_{\substack{T\in[0,T_0]\\M_1(\varepsilon,T)+1\le m\le M_2(\varepsilon)}} a^{-2}\left|J_\varepsilon(2-(m-1)\delta_\varepsilon-\gamma_\varepsilon,a)^2-\delta_\varepsilon^{-1}\tilde{V}_{\varepsilon,a,m}\right|=0. \tag{7.21}$$

On the other hand, we have by (7.13) and (7.10) that

$$\begin{aligned}&\delta_\varepsilon^{-1}\left|\tilde{V}_{\varepsilon,a,m}-V_{\varepsilon,a,m}\right|\\&\quad\le\frac{1}{\delta_\varepsilon\log\varepsilon^{-1}}\int_0^{s'_{m-1}-s_m}\frac{\left|J_\varepsilon(2-(m-1)\delta_\varepsilon-\gamma_\varepsilon,a)^2-J_\varepsilon(2-\log_\varepsilon s,a)^2\right|}{s'_{m-1}-s'_m-s+\varepsilon^2/2}\,\mathrm{d}s\\&\quad\le\frac{a^2\beta^3K_0^3}{4\pi^{3/2}\delta_\varepsilon(\log\varepsilon^{-1})^{\frac32}}\int_0^{s'_{m-1}-s_m}\frac{\sqrt{\log\frac{s'_{m-1}+\varepsilon^2/2}{s+\varepsilon^2/2}}+K_0}{s'_{m-1}-s'_m-s+\varepsilon^2/2}\,\mathrm{d}s\\&\quad\le\frac{a^2\beta^3K_0^3}{4\pi^{3/2}\delta_\varepsilon(\log\varepsilon^{-1})^{\frac32}}\left[\int_0^{s'_{m-1}-s'_m}\frac{\log\frac{s'_{m-1}+\varepsilon^2/2}{s+\varepsilon^2/2}}{s'_{m-1}-s'_m-s+\varepsilon^2/2}\,\mathrm{d}s+(1+K_0)(\log 2+\delta_\varepsilon\log\varepsilon^{-1})\right].\end{aligned} \tag{7.22}$$

In the last inequality we used the elementary inequality $\sqrt{a}\le 1+a$ for all $a\ge 0$ as well as the explicit integral computation

$$\begin{aligned}\int_0^{s'_{m-1}-s_m}\frac{\mathrm{d}s}{s'_{m-1}-s'_m-s+\varepsilon^2/2}&=\log\frac{s'_{m-1}-s'_m+\varepsilon^2/2}{s_m-s'_m+\varepsilon^2/2}\le\log\frac{s'_{m-1}-s'_m}{s_m-s'_m}=\log\frac{\varepsilon^{(m-1)\delta_\varepsilon+\gamma_\varepsilon}-\varepsilon^{m\delta_\varepsilon+\gamma_\varepsilon}}{\varepsilon^{m\delta_\varepsilon}-\varepsilon^{m\delta_\varepsilon+\gamma_\varepsilon}}\\&=\log\frac{\varepsilon^{\gamma_\varepsilon-\delta_\varepsilon}-\varepsilon^{\gamma_\varepsilon}}{1-\varepsilon^{\gamma_\varepsilon}}\le\log 2+\delta_\varepsilon\log\varepsilon^{-1},\end{aligned}$$

with the last inequality by (6.4).

For the first term in brackets on the right side of (7.22), we have by Lemma 4.2 (applied with $t=s'_{m-1}-s'_m$ and $r=s'_m$) that

$$\begin{aligned}\int_0^{s'_{m-1}-s'_m}\frac{\log\frac{s'_{m-1}+\varepsilon^2/2}{s+\varepsilon^2/2}}{s'_{m-1}-s'_m-s+\varepsilon^2/2}\,\mathrm{d}s&\le\left(2+\log(1+2\varepsilon^{-2}(s'_{m-1}-s'_m))\right)\left(1+\log\frac{s'_{m-1}+\varepsilon^2/2}{s'_{m-1}-s'_m+\varepsilon^2/2}\right)\\&\le\left(2+\log(1+2\varepsilon^{-2+(m-1)\delta_\varepsilon+\gamma_\varepsilon})\right)\left(1+\log\frac{1}{1-\varepsilon^{\delta_\varepsilon}}\right).\end{aligned}$$

The second bracketed factor goes to 1 as $\varepsilon\downarrow 0$ (recalling (6.1)) while the first factor is bounded by a constant times $\log\varepsilon^{-1}$. Using this in (7.22), we see that there is a constant $C<\infty$ so that

$$a^{-2}\delta_\varepsilon^{-1}\left|\tilde{V}_{\varepsilon,a,m}-V_{\varepsilon,a,m}\right|\le\frac{C}{\delta_\varepsilon(\log\varepsilon^{-1})^{\frac12}}(1+\delta_\varepsilon),$$

and the right side goes to 0 as $\varepsilon\downarrow 0$ (uniformly in a and in $T\in[0,T_0]$) by (6.1). This and (7.21) imply (7.7). □

7.2.2 Higher moments

For tightness purposes, we will also need an upper bound on a higher moment of $Z_{\varepsilon,b,m}$. Let $p>2$ be as in Proposition 3.1.

Proposition 7.8. *We have*

$$\limsup_{\varepsilon\downarrow 0}\sup_{\substack{a>0\\ M_1(\varepsilon,T)+1\le m\le M_2(\varepsilon)}}\frac{\mathbf{E}|Z_{\varepsilon,a,m}|^p}{a^p\delta_\varepsilon^{p/2}}<\infty. \tag{7.23}$$

Proof. Fix ε, m and define the martingale

$$Z(r)=a+\frac{1}{\sqrt{\log\varepsilon^{-1}}}\int_0^r\int G_{s'_{m-1}-s'_m-s}(X-y)\sigma(u_{\varepsilon,a}(s,y))\,\mathrm{d}W^\varepsilon(s,y),\qquad r\ge 0,$$

so by (7.16) we have $Z_{\varepsilon,a,m}=Z(s'_{m-1}-s_m)$. The quadratic variation process is

$$\langle Z\rangle(r)=\frac{1}{\log\varepsilon^{-1}}\int_0^r\iint G_{\varepsilon^2}(y_1-y_2)\prod_{i=1}^2\Big(G_{s'_{m-1}-s'_m-s}(X-y_i)\sigma(u_{\varepsilon,a}(s,y_i))\Big)\,\mathrm{d}y_1\,\mathrm{d}y_2\,\mathrm{d}s.$$

By the Burkholder–Davis–Gundy inequality (see e.g. [32, Proposition 4.4]), we have a constant $C_p<\infty$ so that

$$\mathbf{E}|Z_{\varepsilon,a,m}|^p\le C_p\mathbf{E}[\langle Z\rangle(s'_{m-1}-s_m)]^{p/2}. \tag{7.24}$$

By the inequality

$$|\sigma(u_{\varepsilon,a}(s,y_1))\sigma(u_{\varepsilon,a}(s,y_2))|\le\frac{\beta^2}{2}(u_{\varepsilon,a}(s,y_1)^2+u_{\varepsilon,a}(s,y_2)^2),$$

we can estimate the quadratic variation as

$$\begin{aligned}\langle Z\rangle(r)&\le\frac{\beta^2}{\log\varepsilon^{-1}}\int_0^r\int G_{s'_{m-1}-s'_m-s}(X-y)G_{s'_{m-1}-s'_m-s+\varepsilon^2}(X-y)u_{\varepsilon,a}(s,y)^2\,\mathrm{d}y\,\mathrm{d}s\\ &=\frac{\beta^2}{4\pi\log\varepsilon^{-1}}\int_0^r\frac{1}{s'_{m-1}-s'_m-s+\varepsilon^2/2}\int G_{\frac{(s'_{m-1}-s'_m-s)(s'_{m-1}-s'_m-s+\varepsilon^2)}{2(s'_{m-1}-s'_m-s)+\varepsilon^2}}(X-y)u_{\varepsilon,a}(s,y)^2\,\mathrm{d}y\,\mathrm{d}s,\end{aligned}$$

where we used (5.9) for the above "=". By Jensen's inequality we have

$$\begin{aligned}\langle Z\rangle(r)^{p/2}&\le\frac{\beta^p}{(4\pi\log\varepsilon^{-1})^{p/2}}\left(\int_0^r\frac{1}{s'_{m-1}-s'_m-s+\varepsilon^2/2}\,\mathrm{d}s\right)^{p/2-1}\cdot\\ &\qquad\cdot\int_0^r\int\frac{1}{s'_{m-1}-s'_m-s+\varepsilon^2/2}G_{\frac{(s'_{m-1}-s'_m-s)(s'_{m-1}-s'_m-s+\varepsilon^2)}{2(s'_{m-1}-s'_m-s)+\varepsilon^2}}(X-y)u_{\varepsilon,a}(s,y)^p\,\mathrm{d}y\,\mathrm{d}s\\ &\le\frac{\beta^p}{(4\pi\log\varepsilon^{-1})^{p/2}}\left(\log\frac{s'_{m-1}-s'_m}{s'_{m-1}-s'_m-r}\right)^{p/2-1}\cdot\\ &\qquad\cdot\int_0^r\int\frac{1}{s'_{m-1}-s'_m-s+\varepsilon^2/2}G_{\frac{(s'_{m-1}-s'_m-s)(s'_{m-1}-s'_m-s+\varepsilon^2)}{2(s'_{m-1}-s'_m-s)+\varepsilon^2}}(X-y)u_{\varepsilon,a}(s,y)^p\,\mathrm{d}y\,\mathrm{d}s.\end{aligned}$$

Taking expectations and using spatial homogeneity, we have

$$\mathbf{E}\langle Z\rangle(r)^{p/2}\le\frac{\beta^p}{(4\pi\log\varepsilon^{-1})^{p/2}}\left(\log\frac{s'_{m-1}-s'_m}{s'_{m-1}-s'_m-r}\right)^{p/2}\sup_{s\in[0,r]}\mathbf{E}u_{\varepsilon,a}(s,y)^p.$$

Substituting $r = s'_{m-1} - s_m$ and recalling (7.24) and Proposition 3.1, we have

$$\mathbf{E}|Z_{\varepsilon,a,m}|^p \le \frac{\beta^p C_p K_0^p a^p}{(4\pi \log \varepsilon^{-1})^{p/2}} \left(\log \frac{s'_{m-1} - s'_m}{s_m - s'_m} \right)^{p/2} = \frac{\beta^p C_p K_0^p a^p}{(4\pi \log \varepsilon^{-1})^{p/2}} \left(\log \frac{\varepsilon^{\gamma_\varepsilon - \delta_\varepsilon} - \varepsilon^{\gamma_\varepsilon}}{1 - \varepsilon^{\gamma_\varepsilon}} \right)^{p/2}.$$

From this and (6.1) we see (7.23). □

8 Proof of Theorem 1.2

In this section we complete the proof of Theorem 1.2. The key remaining step is to show the convergence of the Markov chain defined in Section 7 to a continuous diffusion. The technology for doing this is well-known, through the martingale problem of Stroock and Varadhan. We will essentially use [41, Theorem 11.2.3] as a black box, but we state a special case in a form convenient for us in Appendix A.

Proof of Theorem 1.2. Suppose that $\varepsilon_k \downarrow 0$ and $J : [0,2] \times \mathbf{R}_{\ge 0} \to \mathbf{R}_{\ge 0}$ are such that

$$J_{\varepsilon_k}|_{[0,2]\times\mathbf{R}_{\ge 0}} \to J \tag{8.1}$$

uniformly on compact subsets of $[0,2] \times \mathbf{R}_{\ge 0}$. (These are the subsequential limits that are guaranteed to exist by Proposition 7.4.) By Lemma 7.6, this implies in particular that J is uniformly Lipschitz in its second argument. For $Q \in [0,2]$ and $a \ge 0$, we consider the stochastic differential equation

$$\mathrm{d}\tilde{\Xi}^J_{a,Q}(q) = J(2-q, \tilde{\Xi}^J_{a,Q}(q))\mathrm{d}B(q), \qquad q \in (2-Q, 2]; \tag{8.2}$$

$$\tilde{\Xi}^J_{a,Q}(2-Q) = a, \tag{8.3}$$

where B is a standard Brownian motion. Since J is Lipschitz in the spatial variable, the problem (8.2)–(8.3) has a unique strong solution (given Q and J). For the moment, the limit J may depend on the sequence $\{\varepsilon_k\}$, as may the solution to (8.3).

Suppose that $\{Q_\varepsilon \in [0,2]\}_{\varepsilon > 0}$ is such that

$$Q \coloneqq \lim_{\varepsilon \downarrow 0} Q_\varepsilon \tag{8.4}$$

exists. Define $T_{\varepsilon_k} = \varepsilon_k^{2-Q_{\varepsilon_k}}$. We claim that

$$u_{\varepsilon_k,a}(T_{\varepsilon_k}, X) \xrightarrow[k\to\infty]{\text{law}} \tilde{\Xi}^J_{a,Q}(2). \tag{8.5}$$

By Proposition 7.1, it suffices to show that

$$Y_{\varepsilon_k,a,T_{\varepsilon_k}}(M_2(\varepsilon_k)) \xrightarrow[k\to\infty]{\text{law}} \tilde{\Xi}^J_{a,Q}(2). \tag{8.6}$$

We now explain how (8.6) follows from Theorem A.1 with $A_1 = 2-Q$, $A_2 = 2$, and $L(q,b) = J(2-q,b)$. From (6.7)–(6.8) and (8.4) we have $\delta_{\varepsilon_k} M_1(\varepsilon_k, T_{\varepsilon_k}) \to 2-Q$ and $\delta_{\varepsilon_k} M_2(\varepsilon_k) \to 2$ as $k \to \infty$. The condition (A.1) is verified by Proposition 7.3, while the condition (A.2) is verified by Proposition 7.8. Thus Theorem A.1 applies and we obtain (8.6) and thus (8.5).

We note that the family of random variables $\{\sigma(u_{\varepsilon_k,b}(T_{\varepsilon_k}, X))^2\}_{k\ge 1}$ is uniformly integrable by the $p > 2$ moment bound in Proposition 3.1, so from (8.5) we can derive

$$J(Q,a) = \lim_{k\to\infty} J_{\varepsilon_k}(Q,a) = \lim_{k\to\infty} \frac{1}{2\sqrt{\pi}} \left(\mathbf{E}\sigma(u_{\varepsilon_k,a}(T_{\varepsilon_k}, X))^2 \right)^{1/2} = \frac{1}{2\sqrt{\pi}} \left(\mathbf{E}\sigma(\tilde{\Xi}^J_{a,Q}(2)) \right)^{1/2}. \tag{8.7}$$

The problem (8.2), (8.3), (8.7) agrees with the problem (1.5)–(1.7) by the change of variables

$$\Xi_{a,Q}(q) = \tilde{\Xi}^{J}_{a,Q}(q+2-Q). \tag{8.8}$$

Note also that

$$J(Q,0) = \lim_{k\to\infty} J_{\varepsilon_k}(Q,0) = 0,$$

for all $Q \in [0,2]$, and that

$$\operatorname{Lip} J(Q,\cdot) \le \limsup_{k\to\infty} \operatorname{Lip} J_{\varepsilon_k}(Q,\cdot) \le \limsup_{k\to\infty}\left(\frac{4\pi}{\beta^2} - \frac{\log(1+2\varepsilon_k^{-q})}{\log\varepsilon_k^{-1}}\right)^{-1/2} = (4\pi/\beta^2 - q)^{-1/2}$$

by Lemma 7.6. Therefore, J satisfies both conditions of Theorem 1.1, and thus J is uniquely characterized by the properties we have established for it. By Proposition 7.4, this means that in fact

$$\lim_{\varepsilon\downarrow 0} J_\varepsilon|_{[0,2]\times\mathbf{R}_{\ge 0}} = J$$

uniformly on compact subsets of $[0,2]\times\mathbf{R}_{\ge0}$, so the limiting procedure above does not depend on the specific choice of $\{\varepsilon_k\}$. By the same argument as that leading to (8.5), we have

$$u_{\varepsilon,a}(\varepsilon^{2-Q_\varepsilon}, X) \xrightarrow[\varepsilon\to 0]{\text{law}} \tilde{\Xi}^{J}_{a,Q}(2) \overset{\text{law}}{=} \Xi_{a,Q}(Q). \tag{8.9}$$

In particular, for any T independent of ε, taking $Q_\varepsilon = 2 - \log_\varepsilon T \to Q = 2$, we have

$$u_{\varepsilon,a}(T,X) \xrightarrow[\varepsilon\to 0]{\text{law}} \Xi_{a,2}(2),$$

as claimed. □

Remark 8.1. Now we are able to prove the convergence of the variance of the random variable

$$\mathcal{U}_{\varepsilon,a,T}(g) := \sqrt{\log\varepsilon^{-1}}\int [u_{\varepsilon,a}(T,x)-a]g(x)\mathrm{d}x,$$

where $T>0$ and a Schwartz function g are fixed. By the mild formulation (1.4), recalling that $*$ denotes the spatial convolution, we have

$$\begin{aligned}
&\mathbf{E}\mathcal{U}_{\varepsilon,a,T}(g)^2\\
&\quad= \mathbf{E}\left|\int_0^T\int G_{T-s}*g(y)\sigma(u_{\varepsilon,a}(s,y))\mathrm{d}W^\varepsilon(s,y)\right|^2\\
&\quad= \int_0^T\iint G_{T-s}*g(y_1)G_{T-s}*g(y_2)G_{\varepsilon^2}(y_1-y_2)\mathbf{E}\sigma(u_{\varepsilon,a}(s,y_1))\sigma(u_{\varepsilon,a}(s,y_2))\mathrm{d}y_1\mathrm{d}y_2\mathrm{d}s\\
&\quad= \int_0^T\iint G_{T-s}*g(y_1)G_{T-s}*g(y_1+\varepsilon y_2)G_1(y_2)\mathbf{E}\sigma(u_{\varepsilon,a}(s,y_1))\sigma(u_{\varepsilon,a}(s,y_1+\varepsilon y_2))\mathrm{d}y_1\mathrm{d}y_2\mathrm{d}s.
\end{aligned} \tag{8.10}$$

By Theorem 1.2, Corollary 7.2, and Proposition 3.1, we have, for any $s\in(0,T), y_1,y_2\in\mathbf{R}^2$,

$$\mathbf{E}\sigma(u_{\varepsilon,a}(s,y_1))\sigma(u_{\varepsilon,a}(s,y_1+\varepsilon y_2)) \to \mathbf{E}\sigma(\Xi_{a,2}(2))^2, \qquad \text{as } \varepsilon\to 0.$$

Then we pass to the limit in (8.10) to derive

$$\mathbf{E}\mathcal{U}_{\varepsilon,a,T}(g)^2 \to \mathbf{E}\sigma(\Xi_{a,2}(2))^2 \int_0^T \int |G_{T-s} * g(y)|^2 \mathrm{d}y\mathrm{d}s, \tag{8.11}$$

so the variance of $\mathcal{U}_{\varepsilon,a,T}(g)$ converges as $\varepsilon \to 0$. By adopting the approach in [25], one should be able to further prove the convergence

$$\mathcal{U}_{\varepsilon,a,T}(g) \xrightarrow[\varepsilon\downarrow 0]{\text{law}} \mathcal{U}_{a,T}(g) := \int U_a(T,x)g(x)\mathrm{d}x, \tag{8.12}$$

with the random distribution U_a solving the Edwards-Wilkinson equation

$$\mathrm{d}U_a = \frac{1}{2}\Delta U_a \mathrm{d}t + \sqrt{\mathbf{E}\sigma(\Xi_{a,2}(2))^2}\mathrm{d}W(t,x), \qquad U_a(0,x) = 0. \tag{8.13}$$

To avoid further lengthening the paper we do not pursue this direction here.

9 Multipoint statistics

Now we turn our attention to multipoint statistics and work towards proving Theorem 1.3.

9.1 Local-in-space dependence of the solution on the noise

We can interpret Proposition 4.1 of Section 4 as a form of local-in-time dependence of the solution $u_{\varepsilon,a}$ on the noise. In particular, we can turn off the noise in an area temporally distant from where we evaluate the solution without affecting the solution much. To discuss multipoint statistics, we will need a similar property when we turn off the noise in a spatial region that is distant from our point of interest.

For $B \subset \mathbf{R}^2$, let $v^B_{\varepsilon,a}$ solve the problem

$$\mathrm{d}v^B_{\varepsilon,a}(t,x) = \frac{1}{2}\Delta v^B_{\varepsilon,a}(t,x)\mathrm{d}t + (\log\varepsilon^{-1})^{-\frac{1}{2}}\sigma(v^B_{\varepsilon,a}(t,x))\mathrm{d}W^{\varepsilon,B}(t,x); \tag{9.1}$$

$$v^B_{\varepsilon,a}(0,x) = a. \tag{9.2}$$

Here, $W^{\varepsilon,B} = G_{\varepsilon^2/2} * (W\mathbf{1}_B)$. Note that $W^\varepsilon = W^{\varepsilon,B} + W^{\varepsilon,B^c}$, and moreover that $W^{\varepsilon,B}$ and W^{ε,B^c} are independent. Define

$$R^{\varepsilon,B}(x,x') = \int_B G_{\varepsilon^2/2}(x-y)G_{\varepsilon^2/2}(x'-y)\,\mathrm{d}y \tag{9.3}$$

so that, formally,

$$\mathbf{E}\mathrm{d}W^{\varepsilon,B}(t,x)\mathrm{d}W^{\varepsilon,B}(t',x') = \delta(t-t')R^{\varepsilon,B}(x,x').$$

Note that $R^{\varepsilon,B}(x,x') \le G_{\varepsilon^2}(x-x')$ for all $x,x' \in \mathbf{R}^2$. We note that $v^B_{\varepsilon,a}$ has nothing to do with the $v_{\varepsilon,a}$ considered in Section 5.

Our first goal will be an estimate on what happens if we turn off the noise in a half-plane, which we do in Lemma 9.2 below. We then consider complements of rectangles by taking unions of half-planes in Proposition 9.3. First we record a simple moment bound.

Lemma 9.1. *For any $T \in [0,T_0]$ and any $B \subset \mathbf{R}^2$, we have*

$$\sup_{x\in\mathbf{R}^2} \left(\mathbf{E}v^B_{\varepsilon,a}(t,x)^2\right)^{1/2} \le K_0 a. \tag{9.4}$$

Proof. By the mild solution formula and Young's inequality, we have

$$\begin{aligned}\mathbf{E}v^B_{\varepsilon,a}(t,x)^2 &= a^2 + \frac{1}{\log\varepsilon^{-1}}\int_0^t \iint G_{t-s}(x-y_1)G_{t-s}(x-y_2)R^{\varepsilon,B}(y_1,y_2)\cdot \\ &\qquad\qquad \cdot \mathbf{E}[\sigma(v^B_{\varepsilon,a}(t,y_1))\sigma(v^B_{\varepsilon,a}(t,y_2))]\,\mathrm{d}y_1\,\mathrm{d}y_2\,\mathrm{d}s \\ &\le a^2 + \frac{1}{2\log\varepsilon^{-1}}\sum_{i=1}^2\int_0^t \iint G_{t-s}(x-y_1)G_{t-s}(x-y_2)R^{\varepsilon,B}(y_1,y_2)\mathbf{E}\sigma(v^B_{\varepsilon,a}(t,y_i))^2\,\mathrm{d}y_1\,\mathrm{d}y_2\,\mathrm{d}s.\end{aligned}$$

This means that

$$\begin{aligned}\sup_{x\in\mathbf{R}^2}\mathbf{E}v^B_{\varepsilon,a}(t,x)^2 &\le a^2 + \frac{1}{2\log\varepsilon^{-1}}\int_0^t \iint G_{t-s}(x-y_1)G_{t-s}(x-y_2)G_{\varepsilon^2}(y_1-y_2)\cdot \\ &\qquad\qquad \cdot \sup_{x\in\mathbf{R}^2}\mathbf{E}\sigma(v^B_{\varepsilon,a}(t,x))^2\,\mathrm{d}y_1\,\mathrm{d}y_2\,\mathrm{d}s \\ &\le a^2 + \frac{1}{2\pi\log\varepsilon^{-1}}\int_0^t \frac{\sup_{x\in\mathbf{R}^2}\mathbf{E}\sigma(v^B_{\varepsilon,a}(t,x))^2}{2(t-s)+\varepsilon^2}\,\mathrm{d}s,\end{aligned}$$

and (9.4) then follows from Lemma 3.4 (and (3.8)). □

Lemma 9.2. *Let $B\subset\mathbf{R}^2$ and let H be a half-plane in $\mathbf{R}^2$. Then we have, for all $x\notin H$, that*

$$\mathbf{E}(v^B_{\varepsilon,a}-v^{B\setminus H}_{\varepsilon,a})(t,x)^2 \le 5a^2K_0^2\sum_{k=1}^\infty\left(\frac{\beta^2}{4\pi}\frac{\log(1+2\varepsilon^{-2}t)}{\log\varepsilon^{-1}}\right)^k (G_{\frac{1}{2}[t+k\varepsilon^2]}*\mathbf{1}_H)(x). \tag{9.5}$$

Proof. From (9.1)–(9.2) we write the mild solution formula

$$v^B_{\varepsilon,a}(t,x) = a + \frac{1}{\sqrt{\log\varepsilon^{-1}}}\int_0^t\int G_{t-s}(x-y)\sigma(v^B_{\varepsilon,a}(s,y))\,\mathrm{d}W^{\varepsilon,B}(s,y).$$

Subtracting the corresponding expression for $v^{B\setminus H}_{\varepsilon,a}$, we obtain

$$\begin{aligned}&(v^B_{\varepsilon,a}-v^{B\setminus H}_{\varepsilon,a})(t,x) \\ &\quad= \frac{1}{\sqrt{\log\varepsilon^{-1}}}\int_0^t\int G_{t-s}(x-y)[\sigma(v^B_{\varepsilon,a}(s,y))-\sigma(v^{B\setminus H}_{\varepsilon,a}(s,y))]\,\mathrm{d}W^{\varepsilon,B\setminus H}(s,y) \\ &\qquad + \frac{1}{\sqrt{\log\varepsilon^{-1}}}\int_0^t\int G_{t-s}(x-y)\sigma(v^B_{\varepsilon,a}(s,y))\,\mathrm{d}W^{\varepsilon,B\cap H}(s,y).\end{aligned}$$

Taking second moments in this expression, using the independence of $W^{\varepsilon,B\setminus H}$ and $W^{\varepsilon,B\cap H}$, we have

$$\begin{aligned}&\mathbf{E}(v^B_{\varepsilon,a}-v^{B\setminus H}_{\varepsilon,a})(t,x)^2 \\ &\le \frac{\beta^2}{\log\varepsilon^{-1}}\int_0^t\iint R^{\varepsilon,B\setminus H}(y_1,y_2)\prod_{i=1}^2\left(G_{t-s}(x-y_i)\left(\mathbf{E}(v^B_{\varepsilon,a}-v^{B\setminus H}_{\varepsilon,a})(s,y_i)^2\right)^{1/2}\right)\mathrm{d}y_1\,\mathrm{d}y_2\,\mathrm{d}s \\ &\quad + \frac{\beta^2}{\log\varepsilon^{-1}}\int_0^t\iint R^{\varepsilon,B\cap H}(y_1,y_2)\prod_{i=1}^2\left(G_{t-s}(x-y_i)\left(\mathbf{E}v^B_{\varepsilon,a}(s,y_i)^2\right)^{1/2}\right)\mathrm{d}y_1\,\mathrm{d}y_2\,\mathrm{d}s \\ &=: I_1+I_2.\end{aligned} \tag{9.6}$$

For the first term we can estimate

$$\begin{aligned}I_1 &\le \frac{\beta^2}{\log\varepsilon^{-1}}\int_0^t \iint G_{t-s}(x-y_1)G_{t-s}(x-y_2)G_{\varepsilon^2}(y_1-y_2)\prod_{i=1}^2\left(\mathbf{E}(v^B_{\varepsilon,a}-v^{B\setminus H}_{\varepsilon,a})(s,y_i)^2\right)^{1/2}\mathrm{d}y_1\,\mathrm{d}y_2\,\mathrm{d}s\\ &\le \frac{\beta^2}{\log\varepsilon^{-1}}\int_0^t\int G_{t-s}(x-y_1)G_{t-s+\varepsilon^2}(x-y_1)\mathbf{E}(v^B_{\varepsilon,a}-v^{B\setminus H}_{\varepsilon,a})(s,y_1)^2\,\mathrm{d}y_1\,\mathrm{d}s\\ &\le \frac{\beta^2}{4\pi\log\varepsilon^{-1}}\int_0^t\int \frac{G_{\frac{(t-s)(t-s+\varepsilon^2)}{2(t-s)+\varepsilon^2}}(x-y)}{t-s+\varepsilon^2/2}\left(\mathbf{1}_{H^c}(y)\mathbf{E}(v^B_{\varepsilon,a}-v^{B\setminus H}_{\varepsilon,a})(s,y)^2+4K_0^2a^2\mathbf{1}_H(y)\right)\mathrm{d}y\,\mathrm{d}s,\end{aligned}\tag{9.7}$$

where in the last inequality we used (5.9) and Lemma 9.1. For the second term of (9.6) we can estimate

$$\begin{aligned}I_2 &\le \frac{\beta^2a^2K_0^2}{\log\varepsilon^{-1}}\int_0^t\iint G_{t-s}(x-y_1)G_{t-s}(x-y_2)R^{\varepsilon,B\cap H}(y_1,y_2)\,\mathrm{d}y_1\,\mathrm{d}y_2\,\mathrm{d}s\\ &\le \frac{\beta^2a^2K_0^2}{4\pi\log\varepsilon^{-1}}\int_0^t\frac{(G_{\frac12[t-s+\varepsilon^2/2]}*\mathbf{1}_H)(x)}{t-s+\varepsilon^2/2}\,\mathrm{d}s,\end{aligned}\tag{9.8}$$

where in the second inequality we used (5.9). Using (9.7) and (9.8) in (9.6), we have

$$\begin{aligned}&\mathbf{E}(v^B_{\varepsilon,a}-v^{B\setminus H}_{\varepsilon,a})(t,x)^2\\ &\quad\le \frac{\beta^2}{4\pi\log\varepsilon^{-1}}\int_0^t\int\frac{G_{\frac{(t-s)(t-s+\varepsilon^2)}{2(t-s)+\varepsilon^2}}(x-y)}{t-s+\varepsilon^2/2}\mathbf{1}_{H^c}(y)\mathbf{E}(v^B_{\varepsilon,a}-v^{B\setminus H}_{\varepsilon,a})(s,y)^2\,\mathrm{d}y\,\mathrm{d}s\\ &\qquad+\frac{\beta^2a^2K_0^2}{4\pi\log\varepsilon^{-1}}\int_0^t\frac{\left(\left[4G_{\frac{(t-s)(t-s+\varepsilon^2)}{2(t-s)+\varepsilon^2}}+G_{\frac12[t-s+\varepsilon^2/2]}\right]*\mathbf{1}_H\right)(x)}{t-s+\varepsilon^2/2}\,\mathrm{d}s.\end{aligned}\tag{9.9}$$

Now we note that for all $x\notin H$, and all $r>0$, if we let $\omega\ge 0$ be the distance between x and H, then we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}r}(G_r*\mathbf{1}_H)(x)&=\frac{\mathrm{d}}{\mathrm{d}r}\int_\omega^\infty(2\pi r)^{-1/2}\mathrm{e}^{-\xi^2/(2r)}\,\mathrm{d}\xi=\int_\omega^\infty\frac{\partial^2}{\partial\xi^2}(2\pi r)^{-1/2}\mathrm{e}^{-\xi^2/(2r)}\,\mathrm{d}\xi\\ &=-\frac{\partial}{\partial\xi}(2\pi r)^{-1/2}\mathrm{e}^{-\xi^2/(2r)}\bigg|_{\xi=\omega}=(2\pi r)^{-1/2}\frac{\omega}{r}\mathrm{e}^{-\omega^2/(2r)}\ge 0.\end{aligned}\tag{9.10}$$

This means that for all $s\in[0,t]$, we have

$$(G_{\frac12[t-s+\varepsilon^2/2]}*\mathbf{1}_H)(x)\le(G_{\frac12(t+\varepsilon^2)}*\mathbf{1}_H)(x)$$

and similarly

$$(G_{\frac{(t-s)(t-s+\varepsilon^2)}{2(t-s)+\varepsilon^2}}*\mathbf{1}_H)(x)\le(G_{\frac{t(t+\varepsilon^2)}{2t+\varepsilon^2}}*\mathbf{1}_H)(x)\le(G_{\frac12(t+\varepsilon^2)}*\mathbf{1}_H)(x).$$

Using these estimates in (9.9), we see that if we put $f(t,x)=\mathbf{E}(v^B_{\varepsilon,a}-v^{B\setminus H}_{\varepsilon,a})(t,x)^2$, then for all $x\in H^c$ we have

$$\begin{aligned}f(t,x)&\le\frac{\beta^2}{4\pi\log\varepsilon^{-1}}\int_0^t\int\frac{G_{\frac{(t-s)(t-s+\varepsilon^2)}{2(t-s)+\varepsilon^2}}(x-y)}{t-s+\varepsilon^2/2}\mathbf{1}_{H^c}(y)f(s,y)\,\mathrm{d}y\,\mathrm{d}s\\ &\quad+\frac{5}{4\pi}\beta^2a^2K_0^2\left(G_{\frac12[t+\varepsilon^2]}*\mathbf{1}_H\right)(x)\frac{\log(1+2\varepsilon^{-2}t)}{\log\varepsilon^{-1}}\end{aligned}\tag{9.11}$$

Define

$$b^{(k)}(t)=\frac{t}{2}+k\frac{\varepsilon^2}{2}.\tag{9.12}$$

We note that

$$\sup_{s\in[0,t]}\left[b^{(k)}(s)+\frac{(t-s)(t-s+\varepsilon^2/2)}{2(t-s)+\varepsilon^2/2}\right]=\sup_{s\in[0,t]}\left[\frac{s}{2}+k\frac{\varepsilon^2}{2}+\frac{(t-s)(t-s+\varepsilon^2/2)}{2(t-s)+\varepsilon^2/2}\right] \le b^{(k+1)}(t) \tag{9.13}$$

for all $s\in[0,t]$ and all $k\ge 1$. Define

$$B_2^{(k)}=5a^2K_0^2\left(\frac{\beta^2}{4\pi}\frac{\log(1+2\varepsilon^{-2}t)}{\log\varepsilon^{-1}}\right)^k.$$

Suppose that

$$f(t,x)\le B_1^{(n)}+\sum_{k=1}^n B_2^{(k)}(G_{b^{(k)}(t)}*\mathbf{1}_H)(x) \tag{9.14}$$

for all $x\in H^{\mathrm{c}}$. This is automatically true for $n=0$ with $B_1^{(0)}=\|f\|_{L^\infty([0,t]\times\mathbf{R}^2)}$. Then we have from (9.11) that, for all $x\in H^{\mathrm{c}}$,

$$\begin{aligned}
f(t,x) &\le \frac{\beta^2}{4\pi\log\varepsilon^{-1}}\int_0^t\int\frac{G_{\frac{(t-s)(t-s+\varepsilon^2)}{2(t-s)+\varepsilon^2}}(x-y)}{t-s+\varepsilon^2/2}\mathbf{1}_{H^{\mathrm{c}}}(y)\left[B_1^{(n)}+\sum_{k=1}^n B_2^{(k)}(G_{b^{(k)}(s)}*\mathbf{1}_H)(y)\right]\mathrm{d}y\,\mathrm{d}s\\
&\quad+\frac{5}{4\pi}\beta^2a^2K_0^2\left(G_{\frac12[t+\varepsilon^2]}*\mathbf{1}_H\right)(x)\frac{\log(1+2\varepsilon^{-2}t)}{\log\varepsilon^{-1}}\\
&\le \frac{\beta^2B_1^{(n)}}{8\pi}\frac{\log(1+2\varepsilon^{-2}t)}{\log\varepsilon^{-1}}+\frac{\beta^2}{4\pi\log\varepsilon^{-1}}\sum_{k=1}^n B_2^{(k)}\int_0^t\int\frac{1}{t-s+\varepsilon^2/2}(G_{b^{(k)}(s)+\frac{(t-s)(t-s+\varepsilon^2)}{2(t-s)+\varepsilon^2}}*\mathbf{1}_H)(x)\,\mathrm{d}s\\
&\quad+\frac{5}{4\pi}\beta^2a^2K_0^2\left(G_{b^{(1)}(t)}*\mathbf{1}_H\right)(x)\frac{\log(1+2\varepsilon^{-2}t)}{\log\varepsilon^{-1}}\\
&\le \frac{\beta^2B_1^{(n)}}{8\pi}\frac{\log(1+2\varepsilon^{-2}t)}{\log\varepsilon^{-1}}+\frac{\beta^2\log(1+2\varepsilon^{-2}t)}{4\pi\log\varepsilon^{-1}}\sum_{k=1}^n B_2^{(k)}(G_{b^{(k+1)}(t)}*\mathbf{1}_H)(x)+B_2^{(1)}\left(G_{b^{(1)}(t)}*\mathbf{1}_H\right)(x)\\
&= \frac{\beta^2B_1^{(n)}}{8\pi}\frac{\log(1+2\varepsilon^{-2}t)}{\log\varepsilon^{-1}}+\sum_{k=1}^{n+1}B_2^{(k)}(G_{b^{(k)}(t)}*\mathbf{1}_H)(x).
\end{aligned}$$

In the third inequality we used (9.13) and (9.10). By induction, this means that (9.14) holds for all $n\ge 0$, with $B_1^{(n)}=\|f\|_{L^\infty([0,t]\times\mathbf{R}^2)}\left(\frac{\beta^2}{8\pi}\cdot\frac{\log(1+4\varepsilon^{-2}t)}{\log\varepsilon^{-1}}\right)^n\to 0$ as $n\to\infty$. Therefore, we in fact have

$$f(t,x)\le 5a^2K_0^2\sum_{k=1}^\infty\left(\frac{\beta^2}{4\pi}\frac{\log(1+2\varepsilon^{-2}t)}{\log\varepsilon^{-1}}\right)^k(G_{b^{(k)}(t)}*\mathbf{1}_H)(x),$$

which (recalling (9.12)) is (9.5). □

Now we apply Lemma 9.2 four times to bound the effect of turning off the noise outside of a square.

Proposition 9.3. *Suppose that*

$$\lim_{\varepsilon\downarrow 0}\frac{\xi_\varepsilon}{\eta_\varepsilon}=\lim_{\varepsilon\downarrow 0}\frac{t_\varepsilon^{1/2}}{\eta_\varepsilon}=0. \tag{9.15}$$

and

$$\limsup_{\varepsilon\downarrow 0}t_\varepsilon<\infty. \tag{9.16}$$

Let $\square_\varepsilon = [-\eta_\varepsilon, \eta_\varepsilon]^2$. Then we have for all $x \in [-\xi_\varepsilon, \xi_\varepsilon]^2$ that

$$\lim_{\varepsilon\downarrow 0} \mathbf{E}(u_{\varepsilon,a} - v^{\square_\varepsilon}_{\varepsilon,a})(t_\varepsilon, x)^2 = 0. \tag{9.17}$$

Proof. Using Lemma 9.2 four times, we have

$$\mathbf{E}(u_{\varepsilon,a} - v^{\square_\varepsilon}_{\varepsilon,a})(t_\varepsilon, x)^2 \le 5a^2 K_0^2 \sum_{i=1}^{4} \sum_{k=1}^{\infty} c_\varepsilon^k (G_{\frac{1}{2}[t_\varepsilon + k\varepsilon^2]} * \mathbf{1}_{H_i})(x), \tag{9.18}$$

where $H_1, \dots, H_4$ are four half-planes so that $\square_\varepsilon = \bigcap_{i=1}^4 H_i$. Here we have also defined

$$c_\varepsilon = \frac{\beta^2}{4\pi} \frac{\log(1 + \varepsilon^{-2} t_\varepsilon)}{\log \varepsilon^{-1}}.$$

We note that (9.16) and the subcriticality assumption $\beta < \sqrt{2\pi}$ that

$$\limsup_{\varepsilon\downarrow 0} c_\varepsilon < 1. \tag{9.19}$$

Now if $\limsup\limits_{\varepsilon\downarrow 0} \varepsilon^{-2} t_\varepsilon < \infty$, then $c_\varepsilon \to 0$ as $\varepsilon \downarrow 0$, so using the trivial bound $(G_{\frac{1}{2}[t_\varepsilon + k\varepsilon^2]} * \mathbf{1}_{H_i})(x) \le 1$ in (9.18) we get (9.17). Therefore, we can assume that

$$\limsup_{\varepsilon\downarrow 0} \varepsilon^{-2} t_\varepsilon = \infty. \tag{9.20}$$

We break the inner sum in (9.18) into two pieces. First we estimate

$$\sum_{k=\lfloor \varepsilon^{-2} t_\varepsilon \rfloor}^{\infty} c_\varepsilon^k (G_{\frac{1}{2}[t_\varepsilon + k\varepsilon^2]} * \mathbf{1}_{H_i})(x) \le \sum_{k=\lfloor \varepsilon^{-2} t_\varepsilon \rfloor}^{\infty} c_\varepsilon^k = \frac{c_\varepsilon^{\lfloor \varepsilon^{-2} t_\varepsilon \rfloor}}{1 - c_\varepsilon} \to 0$$

as $\varepsilon \downarrow 0$ by (9.19) and (9.20). Then we estimate

$$\sum_{k=1}^{\lfloor \varepsilon^{-2} t_\varepsilon \rfloor} c_\varepsilon^k (G_{\frac{1}{2}[t_\varepsilon + k\varepsilon^2]} * \mathbf{1}_{H_i})(x) \le (G_{t_\varepsilon} * \mathbf{1}_{H_i})(x) \sum_{k=1}^{\infty} c_\varepsilon^k = \frac{c_\varepsilon}{1 - c_\varepsilon} (G_{t_\varepsilon} * \mathbf{1}_{H_i})(x),$$

using the fact that $t_\varepsilon/2 + k\varepsilon^2/2 \le t_\varepsilon$ whenever $k \le \varepsilon^{-2} t_\varepsilon$. Now we have, for $x \in [-\xi_\varepsilon, \xi_\varepsilon]^2$, that

$$\begin{aligned}(G_{t_\varepsilon} * \mathbf{1}_{H_i})(x) \le \frac{1}{\sqrt{2\pi t_\varepsilon}} \int_{\eta_\varepsilon - \xi_\varepsilon}^{\infty} \exp\left\{-\frac{\alpha^2}{2t_\varepsilon}\right\} \mathrm{d}\alpha &\le \frac{1}{\sqrt{2\pi t_\varepsilon}} \int_{\eta_\varepsilon - \xi_\varepsilon}^{\infty} \exp\left\{-\frac{\alpha(\eta_\varepsilon - \xi_\varepsilon)}{2t_\varepsilon}\right\} \mathrm{d}\alpha \\ &= \frac{\sqrt{2t_\varepsilon/\pi}}{\eta_\varepsilon - \xi_\varepsilon} \exp\left\{-\frac{(\eta_\varepsilon - \xi_\varepsilon)^2}{2t_\varepsilon}\right\} \to 0\end{aligned}$$

as $\varepsilon \downarrow 0$ by (9.15). Combining the last three displays and (9.18) gives us (9.17). □

9.2 Proof of Theorem 1.3

We now have the tools we need to prove Theorem 1.3. Throughout this section, our setup is as in the statement of Theorem 1.3. We note in particular that (1.19) implies (with d as in (1.18)) that

$$d((\tau_{\varepsilon,i}, x_{\varepsilon,i}), (\tau_{\varepsilon,j}, x_{\varepsilon,j})) = \varepsilon^{1 - d_{ij} + o(1)}.$$

and (1.20) implies that

$$\tau_{\varepsilon,i} = \varepsilon^{2-Q+o(1)} \tag{9.21}$$

as $\varepsilon \downarrow 0$. Let κ_ε be such that $\kappa_\varepsilon \to 0$ and

$$10\varepsilon^{1-d_{ij}+\kappa_\varepsilon} \le d((\tau_{\varepsilon,i}, x_{\varepsilon,i}), (\tau_{\varepsilon,j}, x_{\varepsilon,j})) \le \frac{1}{2}\varepsilon^{1-d_{ij}-\kappa_\varepsilon}. \tag{9.22}$$

and

$$2\varepsilon^{2-Q+2\kappa_\varepsilon} \le \tau_{\varepsilon,i} \le \varepsilon^{2-Q-2\kappa_\varepsilon}. \tag{9.23}$$

Our first step will apply Proposition 9.3 to show that the values of the solution $u_{\varepsilon,a}$ at distant space-time points are asymptotically independent.

Proposition 9.4. *Let $P_1, \ldots, P_R$ be a partition of $[N]$ so that*

$$d_{ij} \ge Q/2 \iff i \in P_m, j \in P_n, n \ne m. \tag{9.24}$$

Then there is an $\varepsilon_1 \in (0, \varepsilon_0]$ so that if $\varepsilon \in [0, \varepsilon_1)$ then there are independent processes $u_{\varepsilon,a}^{(1)}, \ldots, u_{\varepsilon,a}^{(R)}$ so that $u_{\varepsilon,a}^{(k)} \overset{\text{law}}{=} u_{\varepsilon,a}$ ($k = 1, \ldots, R$), and for each $j \in P_k$ ($k = 1, \ldots, R$), we have

$$\lim_{\varepsilon \downarrow 0} \mathbf{E}(u_{\varepsilon,a}^{(k)}(\tau_{\varepsilon,j}, x_{\varepsilon,j}) - u_{\varepsilon,a}(\tau_{\varepsilon,j}, x_{\varepsilon,j}))^2 = 0. \tag{9.25}$$

Proof. For each $k = 1, \ldots, R$, let i_k be an arbitrary element of P_k. Define

$$D_k \coloneqq \max_{i,j \in P_k} d_{ij} < Q/2, \tag{9.26}$$

with the inequality by (9.24). Define the sets $S_{\varepsilon,k} \subset \mathbf{R} \times \mathbf{R}^2$ by

$$S_{\varepsilon,k} = \left(\tau_{\varepsilon,i_k} + [-\varepsilon^{2-Q+2\kappa_\varepsilon+2\zeta_\varepsilon}, \varepsilon^{2-Q+2\kappa_\varepsilon+2\zeta_\varepsilon}]\right) \times \left(x_{\varepsilon,i_k} + [-\varepsilon^{1-Q/2+\kappa_\varepsilon}, \varepsilon^{1-Q/2+\kappa_\varepsilon}]^2\right).$$

Here κ_ε is as in (9.22)–(9.23) and ζ_ε is as in (6.3).

If $k_1 \ne k_2$, then we have by (9.24) that $d_{i_{k_1} i_{k_2}} \ge Q/2$, so by (1.18) and (9.22) we have

$$\max\{|\tau_{\varepsilon,i_{k_1}} - \tau_{\varepsilon,i_{k_2}}|^{1/2}, |x_{\varepsilon,i_{k_1}} - x_{\varepsilon,i_{k_2}}|\} \ge 10\varepsilon^{1-Q/2+\kappa_\varepsilon}.$$

This means that $\{S_{\varepsilon,1}, \ldots, S_{\varepsilon,R}\}$ forms a pairwise-disjoint family of sets.

Let $A_k = [0, \tau_{\varepsilon,i_{k_1}} - \varepsilon^{2-Q+\kappa_\varepsilon+\zeta_\varepsilon}]$. Define $u_{\varepsilon,a}^{A_k}$ as in (1.26)–(1.27). By Proposition 4.1, we have, for all $j \in P_k$, that

$$\begin{aligned} &\left(\mathbf{E}(u_{\varepsilon,a} - u_{\varepsilon,a}^{A_k})(\tau_{\varepsilon,j}, x_{\varepsilon,j})^2\right)^{1/2} \\ &\qquad \le \frac{\beta a K_0^2}{2\sqrt{\pi \log \varepsilon^{-1}}}\left(K_0 + \sqrt{\log \frac{\tau_{\varepsilon,j} + \varepsilon^2}{\tau_{\varepsilon,j} - \tau_{\varepsilon,i_k} + \varepsilon^{2-Q+2\kappa_\varepsilon+2\zeta_\varepsilon} + \varepsilon^2}}\right). \end{aligned} \tag{9.27}$$

We note (still assuming $j \in P_k$) that

$$|\tau_{\varepsilon,j} - \tau_{\varepsilon,i_k}| \le \frac{1}{4}\varepsilon^{2-2D_k-2\kappa_\varepsilon} \ll \varepsilon^{2-Q+2\kappa_\varepsilon+2\zeta_\varepsilon} \qquad \text{and} \qquad \tau_{\varepsilon,j} \le \varepsilon^{2-Q-2\kappa_\varepsilon} \tag{9.28}$$

by (9.22), (9.23), and (9.26). Thus from (9.27) we obtain a constant C so that

$$\left(\mathbf{E}(u_{\varepsilon,a} - u_{\varepsilon,a}^{A_k})(\tau_{\varepsilon,j}, x_{\varepsilon,j})^2\right)^{1/2} \le \frac{C\beta a K_0^2}{2\sqrt{\pi \log \varepsilon^{-1}}}\left(K_0 + \sqrt{\log \varepsilon^{-4\kappa_\varepsilon - 2\zeta_\varepsilon}}\right) \to 0 \tag{9.29}$$

as $\varepsilon \downarrow 0$ since $\kappa_\varepsilon, \zeta_\varepsilon \to 0$.

Define $\pi_1 : \mathbf{R}\times\mathbf{R}^2 \to \mathbf{R}$ be given by $\pi_1(t,x) = t$ and $\pi_2 : \mathbf{R}\times\mathbf{R}^2 \to \mathbf{R}^2$ be given by $\pi_2(t,x) = x$. Let $\tilde{u}^{(k)}_{\varepsilon,a}$ solve the problem

$$\mathrm{d}\tilde{u}^{(k)}_{\varepsilon,a}(t,x) = \frac{1}{2}\Delta\tilde{u}^{(k)}_{\varepsilon,a}(t,x)\mathrm{d}t + (\log\varepsilon^{-1})^{-\frac{1}{2}}\mathbf{1}_{\pi_1(S_{\varepsilon,k})}(t)\sigma(\tilde{u}^{(k)}_{\varepsilon,a}(t,x))\mathrm{d}W^{\varepsilon,\pi_2(S_{\varepsilon,k})}(t,x); \tag{9.30}$$

$$\tilde{u}^{(k)}_{\varepsilon,a}(0,x) = a. \tag{9.31}$$

This turns off some temporal part of the noise as in (1.26)–(1.27) but also a spatial part of the noise as in (9.1)–(9.2). Since $\{S_{\varepsilon,1},\ldots,S_{\varepsilon,R}\}$ is pairwise-disjoint, the processes $u^{(1)}_{\varepsilon,a},\ldots,u^{(R)}_{\varepsilon,a}$ are independent. We now want to apply (a translated version of) Proposition 9.3 with

$$\xi_\varepsilon = \varepsilon^{1-D_k-\kappa_\varepsilon}, \qquad \eta_\varepsilon = \varepsilon^{1-Q/2+\kappa_\varepsilon}, \qquad t_\varepsilon = \tau_{\varepsilon,j} - \tau_{\varepsilon,i_{k_1}} + \varepsilon^{2-Q+2\kappa_\varepsilon+2\zeta_\varepsilon}.$$

Note that

$$\lim_{\varepsilon\downarrow 0}\frac{\xi_\varepsilon}{\eta_\varepsilon} = \lim_{\varepsilon\downarrow 0}\frac{\varepsilon^{1-D_k-\kappa_\varepsilon}}{\varepsilon^{1-Q/2+\kappa_\varepsilon}} = \lim_{\varepsilon\downarrow 0}\varepsilon^{Q/2-D_k-2\kappa_\varepsilon} = 0$$

since $D_k < Q/2$ and $\kappa_\varepsilon \to 0$, and also that (using these facts along with (6.3) and (9.28)) that

$$\lim_{\varepsilon\downarrow 0}\frac{t_\varepsilon^{1/2}}{\eta_\varepsilon} \le \lim_{\varepsilon\downarrow 0}\frac{(\tau_{\varepsilon,j}-\tau_{\varepsilon,i_{k_1}})^{1/2}}{\varepsilon^{1-Q/2+\kappa_\varepsilon}} + \lim_{\varepsilon\downarrow 0}\frac{\varepsilon^{1-Q/2+\kappa_\varepsilon+\zeta_\varepsilon}}{\varepsilon^{1-Q/2+\kappa_\varepsilon}} \le \lim_{\varepsilon\downarrow 0}\frac{\varepsilon^{1-D_k-\kappa_\varepsilon}}{\varepsilon^{1-Q/2+\kappa_\varepsilon}} + \lim_{\varepsilon\downarrow 0}\varepsilon^{\zeta_\varepsilon} = 0.$$

Therefore, (9.15) is verified, so Proposition 9.3 applies, and we have (combining the result with (9.29)) that

$$\lim_{\varepsilon\downarrow 0}\mathbf{E}(u_{\varepsilon,a} - \tilde{u}^{(k)}_{\varepsilon,a})(\tau_{\varepsilon,i_j}, x_{\varepsilon,i_j})^2 = 0 \tag{9.32}$$

for all $j \in P_k$. Now let $u^{(k)}_{\varepsilon,a}$ solve the problem

$$\begin{aligned}\mathrm{d}u^{(k)}_{\varepsilon,a}(t,x) &= \frac{1}{2}\Delta u^{(k)}_{\varepsilon,a}(t,x)\mathrm{d}t \\ &\quad + (\log\varepsilon^{-1})^{-\frac{1}{2}}\mathbf{1}_{\pi_1(S_{\varepsilon,k})}(t)\sigma(u^{(k)}_{\varepsilon,a}(t,x))\mathrm{d}[W^{\varepsilon,\pi_2(S_{\varepsilon,k})}(t,x) + \tilde{W}^{\varepsilon,\pi_2(S_{\varepsilon,k})^{\mathrm{c}}}] \\ &\quad + (\log\varepsilon^{-1})^{-\frac{1}{2}}\mathbf{1}_{\mathbf{R}\setminus\pi_1(S_{\varepsilon,k})}(t)\sigma(u^{(k)}_{\varepsilon,a}(t,x))\mathrm{d}\tilde{W}^{\varepsilon}(t,x)\end{aligned} \tag{9.33}$$

$$u_{\varepsilon,a}(0,x) = a, \tag{9.34}$$

where $\tilde{W}$ is an independent copy of W (different and independent across different choices of k). Note that $u^{(1)}_{\varepsilon,a},\ldots,u^{(R)}_{\varepsilon,a}$ are independent since the family $\{S_{\varepsilon,1},\ldots,S_{\varepsilon,R}\}$ is disjoint. The pairs $(u_{\varepsilon,a},\tilde{u}^{(k)}_{\varepsilon,a})$ and $(u^{(k)}_{\varepsilon,a},\tilde{u}^{(k)}_{\varepsilon,a})$ have the same joint laws because to go from $u_{\varepsilon,a}$ to $u^{(k)}_{\varepsilon,a}$ we simply replaced a part of the noise (on $S^{\mathrm{c}}_{\varepsilon,k}$) that is independent of $\tilde{u}^{(k)}_{\varepsilon,a}$ (for which the noise on $S^{\mathrm{c}}_{\varepsilon,k}$ is turned off). Therefore, (9.32) also means that

$$\lim_{\varepsilon\downarrow 0}\mathbf{E}(u^{(k)}_{\varepsilon,a} - \tilde{u}^{(k)}_{\varepsilon,a})(\tau_{\varepsilon,i_j}, x_{\varepsilon,i_j})^2 = 0, \tag{9.35}$$

and combining this with (9.32) yields (9.25). □

Now we can prove Theorem 1.3.

Proof of Theorem 1.3. We use induction on N. The base case, (1.24) with $N = 1$, is simply an application of (8.9). Now suppose that $N \ge 2$ and that (1.24) holds for all strictly smaller N. Let

$$q_0 = 2 - 2\max_{i,j\in[N]} d_{i,j}. \tag{9.36}$$

Then we have

$$q < Q-2+q_0 \implies i_{(Q-q)/2}([N]) = \{1\} \tag{9.37}$$

by the definition (1.21). Define

$$m_\varepsilon(q_0) = \max\{M_1(\varepsilon,\tau_{\varepsilon,1}), \lfloor (q_0 - 2\kappa_\varepsilon - 2\gamma_\varepsilon)\delta_\varepsilon^{-1} \rfloor\}, \tag{9.38}$$

recalling the definition (6.7), and also recall the definition (6.11)–(6.12) of $w^{(m)}_{\varepsilon,a,T,X}$. In the case $m_\varepsilon(q_0) = M_0(\varepsilon,\tau_{\varepsilon,1})$, we have

$$u_{\varepsilon,a}(\tau_{\varepsilon,j}, x_{\varepsilon,j}) = w^{(m_\varepsilon(q_0))}_{\varepsilon,a,\tau_{\varepsilon,1},x_{\varepsilon,1}} \tag{9.39}$$

by the definition (6.10). Otherwise, we note using (9.22) that

$$\tau_{\varepsilon,j} \ge \tau_{\varepsilon,1} - \frac{1}{2}\varepsilon^{2-2d_{1j}-2\kappa_\varepsilon} \ge \tau_{\varepsilon,1} - \frac{1}{2}\varepsilon^{q_0-2\kappa_\varepsilon} \ge \tau_{\varepsilon,1} - \frac{1}{2}\varepsilon^{m_\varepsilon(q_0)\delta_\varepsilon+\gamma_\varepsilon}. \tag{9.40}$$

Thus we can apply Proposition 6.2 with $C_\varepsilon = \varepsilon^{-\gamma_\varepsilon/2}$ (recalling (6.1)) and $c = 1/2$, and by (9.40) take $T = \tau_{\varepsilon,1}$, $k = m_\varepsilon(q_0)$, and $t = \tau_{\varepsilon,j}$ in the supremum in (6.18), to obtain

$$\lim_{\varepsilon\downarrow 0} \frac{\left(\mathbf{E}(u_{\varepsilon,a} - w^{(m_\varepsilon(q_0))}_{\varepsilon,a,\tau_{\varepsilon,1},x_{\varepsilon,1}})(\tau_{\varepsilon,j},x_{\varepsilon,j})^2\right)^{1/2}}{a(1+\varepsilon^{-m_\varepsilon(q_0)\delta_\varepsilon/2-\gamma_\varepsilon/2}|x_{\varepsilon,j}-x_{\varepsilon,1}|)} = 0. \tag{9.41}$$

Note that (9.39) implies (9.41) as well, so in fact (9.41) holds unconditionally. On the other hand, we also have, using (9.38), (9.22), (9.36), and (6.1), that

$$\lim_{\varepsilon\downarrow 0} \varepsilon^{-m_\varepsilon(q_0)\delta_\varepsilon/2-\gamma_\varepsilon/2}|x_{\varepsilon,j}-x_{\varepsilon,1}| \le \frac{1}{2}\lim_{\varepsilon\downarrow 0}\varepsilon^{\gamma_\varepsilon/2-q_0/2+1-d_{1j}} \le \frac{1}{2}\lim_{\varepsilon\downarrow 0}\varepsilon^{\gamma_\varepsilon/2} = 0.$$

Combined with (9.41), this means that

$$\lim_{\varepsilon\downarrow 0} a^{-1}\left(\mathbf{E}(u_{\varepsilon,a} - w^{(m_\varepsilon(q_0))}_{\varepsilon,a,\tau_{\varepsilon,1},x_{\varepsilon,1}})(\tau_{\varepsilon,j},x_{\varepsilon,j})^2\right)^{1/2} = 0. \tag{9.42}$$

Now define

$$\ell_\varepsilon = \tau_{\varepsilon,1} - \varepsilon^{m_\varepsilon(q_0)\delta_\varepsilon+\gamma_\varepsilon} \tag{9.43}$$

and

$$w(t,x) = w^{(m_\varepsilon(q_0))}_{\varepsilon,a,\tau_{\varepsilon,1},x_{\varepsilon,1}}(t+\ell_\varepsilon, x_{\varepsilon,1}).$$

Note that if $T = \tau_{\varepsilon,1}$ then $t'_{m_\varepsilon(q_0)} = \tau_{\varepsilon,1} - \ell_\varepsilon$, so $w(0,\cdot)$ is constant in space and $w(0,x) \overset{\text{law}}{=} Y_{\varepsilon,a,\tau_{\varepsilon,1}}(m_\varepsilon(q_0))$. Thus, by applying Theorem A.1 as in the proof of (8.6) (recalling (8.8) and (9.37)), we see that

$$w(0,x) \xrightarrow[\varepsilon\downarrow 0]{\text{law}} \Gamma_{a,Q,1}(Q-(2-q_0)). \tag{9.44}$$

Moreover, w is equal in law to $u_{\varepsilon,b}$, where $b = w(0,x)$ is taken to be independent of the noise driving $u_{\varepsilon,b}$.

Recall the definition (1.21) and let

$$P_k = i^{-1}_{1-q_0/2}(k) = \{j \in [N] \;:\; i_{1-q_0/2}(j) = k\}.$$

Note that $P_1,\ldots,P_N$ form a partition of $[N]$, and by (9.36) this partition is nontrivial. If $i_{1-q_0/2}(j_1) = i_{1-q_0/2}(j_2)$ then $d_{j_1,j_2} < 1-q_0/2$ by the strong triangle inequality (1.25). On the other hand, if $i_{1-q_0/2}(j_1) \ne i_{1-q_0/2}(j_2)$ and $d_{j_1,j_2} < 1-q_0/2$, then we have by (1.25) and (9.36) that

$$d_{i_{1-q_0/2}(j_1),i_{1-q_0/2}(j_2)} \le \max\{d_{i_{q_0}(j_1),j_1}, d_{j_1,j_2}, d_{j_2,i_{q_0}(j_2)}\} < 1-q_0/2,$$

contradicting the definition (1.21). Therefore, we have

$$i_{1-q_0/2}(j_1) = i_{1-q_0/2}(j_2) \iff d_{j_1,j_2} < 1 - q_0/2. \tag{9.45}$$

Furthermore, we note that, for all $j \in P_k$, we have $2d_{j,k} < 2 - q_0$, which means that (recalling (9.43), (9.22), and (9.36)) we have

$$2 - \lim_{\varepsilon \downarrow 0} \log_\varepsilon \left(\tau_{\varepsilon,j} - \ell_\varepsilon\right) = 2 - q_0. \tag{9.46}$$

Comparing this with (1.20), we see that the collection $\{(\tau_{\varepsilon,j} - \ell_\varepsilon, x_{\varepsilon,j})\}_{j \in [N]}$ of space-time points satisfies the hypotheses of the theorem with the same d_{ij}s but with Q replaced by $2 - q_0$. Thus by (9.45), Proposition 9.4 applies and we obtain independent processes $w^{(1)}, \dots, w^{(N)}$, each distributed identically to w, so that, whenever $j \in P_k$, we have

$$\lim_{\varepsilon \downarrow 0} \mathbf{E}(w^{(k)} - w)(\tau_{\varepsilon,j} - \ell_\varepsilon, x_{\varepsilon,j})^2 = 0. \tag{9.47}$$

By the nontriviality of the partition $\{P_1, \dots, P_N\}$ we have $|P_k| < N$ for each k. Therefore, by the inductive hypothesis, we have

$$(w^{(k)}(\tau_{\varepsilon,j} - \ell_\varepsilon, x_{\varepsilon,j}))_{j \in P_k} \xrightarrow[\varepsilon \downarrow 0]{\text{law}} (\Gamma_{b,2-q_0,j}(2 - q_0))_{j \in P_k},$$

with $b = w(0,x)$ independent of the randomness in the processes on the right side. Here we also used that $i_{(2-q_0-q)/2}(j)$ does not change when the minimum in (1.21) is restricted to elements of P_k, since P_k was defined so that this minimum will be an element of P_k anyway. But since the family $(w^{(k)})_{k=1}^N$ is independent, as is the family $((\Gamma_{b,Q-q_0,j}(Q - q_0))_{j \in P_k})_{k=1}^N$, this means that in fact

$$(w^{(k)}(\tau_{\varepsilon,j} - \ell_\varepsilon, x_{\varepsilon,j}))_{j \in 1}^N \xrightarrow[\varepsilon \downarrow 0]{\text{law}} (\Gamma_{b,2-q_0,j}(2 - q_0))_{j=1}^N, \tag{9.48}$$

again with $b = w(0,x)$ independent of the randomness in the processes on the right side. Combining (9.42), (9.44), (9.47), (9.48), and the continuity of the SDE (1.22)–(1.23) with respect to the initial condition, we obtain (1.24). □

A Convergence of discrete Markov martingales to continuous diffusions

For the convenience of readers, we recall in this section a classical result on the convergence of Markov chains to diffusions that is used in the paper. We use the formulation and results given in [41, Section 11.2].

Theorem A.1. *Suppose that we have a sequence of numbers $\delta_k \downarrow 0$, a sequence of discrete Markov martingales $(\{Y_k(m)\}_{m=A_1(k),\dots,A_2(k)})_{k=1}^\infty$, and a continuous function $L : [A_1, A_2] \times \mathbf{R} \to \mathbf{R}$ satisfying the following conditions:*

1. *The sequence of random variables $(Y_k(A_1(k)))$ converges in law to a random variable X as $k \to \infty$.*
2. *For each $q \in [A_1, A_2]$, the function $L(q, \cdot)$ is Lipschitz with the Lipschitz constant bounded above independent of q.*
3. *We have $\delta_k m \in [A_1, A_2]$ for all $k \ge 1$ and $m = A_1(k), \dots, A_2(k)$, and*

$$\lim_{k \to \infty} \delta_k A_1(k) = A_1 \qquad \textit{and} \qquad \lim_{k \to \infty} \delta_k A_2(k) = A_2.$$

4. *For each $R < \infty$, we have*

$$\lim_{k\to\infty} \sup_{\substack{|x|\le R\\ A_1(k)\le m<A_2(k)}} \left|\delta_k^{-1}\operatorname{Var}[Y_k(m+1)\mid Y_k(m)=x]-L(\delta_k m,x)\right|=0. \tag{A.1}$$

5. *There is a $p > 2$ so that, for each $R < \infty$, we have*

$$\sup_{\substack{k<\infty,|x|\le R\\ A_1(k)\le m<A_2(k)}} \delta_k^{-p/2}\mathbf{E}[(Y_k(m+1)-Y_k(m))^p\mid Y_k(m)=x]<\infty. \tag{A.2}$$

Let $(Y(q))_{q\in[A_1,A_2]}$ solve the stochastic differential equation

$$\mathrm{d}Y(q)=L(q,Y(q))\,\mathrm{d}B(q), \qquad q>A_1; \tag{A.3}$$
$$Y(A_1)=X, \tag{A.4}$$

where $B(q)$ is a standard Brownian motion. Then we have

$$Y_k(A_2(k))\xrightarrow[k\to\infty]{\text{law}} Y(A_2). \tag{A.5}$$

Proof. This is essentially an application of [41, Theorem 11.2.3]. Since that theorem is stated in a general form, we provide some details on how to check the conditions. First we note that although [41, Theorem 11.2.3] is stated for time-independent diffusions, it is trivial to add the time-dependence simply by considering the space-time processes of the form $\{(Y_k(m),\delta_k m)\}_{m=A_1(k),\dots,A_2(k)}$. Applying [41, Theorem 11.2.3] requires also knowing that the limiting martingale problem corresponding for (A.3)–(A.4) is well-posed. The SDE (A.3)–(A.4) has pathwise unique solutions by the standard theory and condition 2 in the statement of theorem. This implies that there are unique solutions for the martingale problem by results [44, 43] of Watanabe and Yamada; see [41, Corollary 8.1.6]. Finally, [41, Theorem 11.2.3] is stated for diffusions starting at time 0 and lasting for all time; this can be adapted to our setting (a finite time interval with arbitrary starting time) by shifting time and extending the Markov chains to later times in some arbitrary way.

The quantitative conditions for [41, Theorem 11.2.3] are [41, (11.2.4)–(11.2.6)]. In our setting, [41, (11.2.4)] is a consequence of (A.1) (and the fact that there is no diffusion for the time process). The fact that we have assumed that each $Y_k(\cdot)$ is a martingale means that there is no drift for the space process, and of course the drift condition is satisfied trivially for the time process, so [41, (11.2.5)] is trivial in our setting. Finally, [41, (11.2.6] holds because, by (A.2) and Markov's inequality, we have for any fixed $\kappa > 0$ that

$$\begin{aligned}\frac{1}{\delta_k}\mathbf{P}(|Y_k(m+1)-Y_k(m)|\ge\kappa\mid Y_k(m)=x) &\le \frac{\mathbf{E}[|Y_k(m+1)-Y_k(m)|^p\mid Y_k(m)=x]}{\delta_k\kappa^p}\\ &\le C\delta_k^{p/2-1}\kappa^{-p}\end{aligned}$$

for a constant $C < \infty$, and the last quantity goes to 0 as $k \to \infty$ since $p > 2$ and $\delta_k \downarrow 0$.

Now condition 1 and the proof of [41, Theorem 11.2.3] show that, if we define

$$\overline{Y}_k(A_1+\delta_k[m-A_1(k)])=Y_k(m), \qquad m=A_1(k),\dots,A_2(k),$$

and extend $\overline{Y}_k$ to $[A_1,A_2]$ by linear interpolation (possibly extending it by a constant on the small interval $[A_1+\delta_k(A_2(k)-A_1(k)),A_2]$), then $\overline{Y}_k$ converges to Y in distribution with respect to the uniform topology on continuous functions on $[A_1,A_2]$. Then (A.5) follows. □

References

[1] Tom Alberts, Konstantin Khanin, and Jeremy Quastel. The intermediate disorder regime for directed polymers in dimension $1+1$. *Ann. Probab.*, 42(3):1212–1256, 2014.

[2] Lorenzo Bertini and Nicoletta Cancrini. The stochastic heat equation: Feynman-Kac formula and intermittence. *J. Stat. Phys.*, 78(5-6):1377–1401, 1995.

[3] Lorenzo Bertini and Nicoletta Cancrini. The two-dimensional stochastic heat equation: renormalizing a multiplicative noise. *J. Phys. A*, 31(2):615–622, 1998.

[4] Giuseppe Cannizzaro, Dirk Erhard, and Philipp Schönbauer. 2d anisotropic kpz at stationarity: scaling, tightness and non triviality. *Ann. Probab.*, 49(1):122–156, 2021.

[5] Giuseppe Cannizzaro, Dirk Erhard, and Fabio Toninelli. The stationary AKPZ equation: logarithmic superdiffusivity, arXiv: 2007.12203v3.

[6] Giuseppe Cannizzaro, Dirk Erhard, and Fabio Toninelli. Weak coupling limit of the anisotropic KPZ equation, arXiv: 2108.09046v1.

[7] Francesco Caravenna, Rongfeng Sun, and Nikos Zygouras. Universality in marginally relevant disordered systems. *Ann. Appl. Probab.*, 27(5):3050–3112, 2017.

[8] Francesco Caravenna, Rongfeng Sun, and Nikos Zygouras. On the moments of the $(2+1)$-dimensional directed polymer and stochastic heat equation in the critical window. *Comm. Math. Phys.*, 372(2):385–440, 2019.

[9] Francesco Caravenna, Rongfeng Sun, and Nikos Zygouras. The two-dimensional KPZ equation in the entire subcritical regime. *Ann. Probab.*, 48(3):1086–1127, 2020.

[10] Francesco Caravenna, Rongfeng Sun, and Nikos Zygouras. The critical 2D stochastic heat flow, arXiv: arXiv:2109.03766.

[11] Sourav Chatterjee and Alexander Dunlap. Constructing a solution of the $(2+1)$-dimensional KPZ equation. *Ann. Probab.*, 48(2):1014–1055, 2020.

[12] Le Chen and Jingyu Huang. Comparison principle for stochastic heat equation on $\mathbb{R}^d$. *Ann. Probab.*, 47(2):989–1035, 2019.

[13] Le Chen and Kunwoo Kim. Nonlinear Stochastic Heat Equation Driven by Spatially Colored Noise: Moments and Intermittency. *Acta Math. Sci. Ser. B (Engl. Ed.)*, 39(3):645–668, 2019.

[14] Le Chen and Kunwoo Kim. Stochastic comparisons for stochastic heat equation. *Electron. J. Probab.*, 25:Paper No. 140, 38, 2020.

[15] Clément Cosco, Shuta Nakajima, and Makoto Nakashima. Law of large numbers and fluctuations in the sub-critical and L^2 regions for SHE and KPZ equation in dimension $d \geq 3$, arXiv: 2005.12689v1.

[16] J. Theodore Cox, Klaus Fleischmann, and Andreas Greven. Comparison of interacting diffusions and an application to their ergodic theory. *Probab. Theory Related Fields*, 105(4):513–528, 1996.

[17] Robert C. Dalang. Extending the martingale measure stochastic integral with applications to spatially homogeneous S.P.D.E.'s. *Electron. J. Probab.*, 4:no. 6, 29 pp., 1999.

[18] Robert C. Dalang and Lluís Quer-Sardanyons. Stochastic integrals for spde's: a comparison. *Expo. Math.*, 29(1):67–109, 2011.

[19] Donald A. Dawson and Habib Salehi. Spatially homogeneous random evolutions. *J. Multivariate Anal.*, 10(2):141–180, 1980.

[20] Jian Ding and Alexander Dunlap. Subsequential scaling limits for Liouville graph distance. *Comm. Math. Phys.*, 376(2):1499–1572, 2020.

[21] Alexander Dunlap, Yu Gu, Lenya Ryzhik, and Ofer Zeitouni. The random heat equation in dimensions three and higher: the homogenization viewpoint, arXiv: 1808.07557v2. To appear in *Arch. Rational Mech. Anal.*

[22] Alexander Dunlap, Yu Gu, Lenya Ryzhik, and Ofer Zeitouni. Fluctuations of the solutions to the KPZ equation in dimensions three and higher. *Probab. Theory Related Fields*, 176(3-4):1217–1258, 2020.

[23] Alexander Fromm. *Theory and applications of decoupling fields for forward-backward stochastic differential equations*. PhD thesis, Humboldt-Universität zu Berlin, July 2014.

[24] Yu Gu. Gaussian fluctuations from the 2D KPZ equation. *Stoch. Partial Differ. Equ. Anal. Comput.*, 8(1):150–185, 2020.

[25] Yu Gu and Jiawei Li. Fluctuations of a nonlinear stochastic heat equation in dimensions three and higher. *SIAM J. Math. Anal.*, 52(6):5422–5440, 2020.

[26] Yu Gu, Jeremy Quastel, and Li-Cheng Tsai. Moments of the 2D SHE at criticality. *Probab. Math. Phys.*, 2(1):179–219, 2021.

[27] Yu Gu, Lenya Ryzhik, and Ofer Zeitouni. The Edwards-Wilkinson limit of the random heat equation in dimensions three and higher. *Comm. Math. Phys.*, 363(2):351–388, 2018.

[28] Martin Hairer. A theory of regularity structures. *Invent. Math.*, 198(2):269–504, 2014.

[29] Martin Hairer and Étienne Pardoux. A Wong-Zakai theorem for stochastic PDEs. *J. Math. Soc. Japan*, 67(4):1551–1604, 2015.

[30] Martin Hairer and Jeremy Quastel. A class of growth models rescaling to KPZ. *Forum Math. Pi*, 6:e3, 2018.

[31] Shunsuke Ihara. *Information theory for continuous systems*. World Scientific Publishing Co., Inc., River Edge, NJ, 1993.

[32] Davar Khoshnevisan. *Analysis of stochastic partial differential equations*, volume 119 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 2014.

[33] Dimitris Lygkonis and Nikos Zygouras. Edwards-Wilkinson fluctuations for the directed polymer in the full L^2-regime for dimensions $d \geq 3$, arXiv: 2005.12706. To appear in *Ann. Inst. H. Poincaré Probab. Stat.*

[34] Jin Ma, Philip Protter, and Jiong Min Yong. Solving forward-backward stochastic differential equations explicitly—a four step scheme. *Probab. Theory Related Fields*, 98(3):339–359, 1994.

[35] Jin Ma, Zhen Wu, Detao Zhang, and Jianfeng Zhang. On well-posedness of forward-backward SDEs—a unified approach. *Ann. Appl. Probab.*, 25(4):2168–2214, 2015.

[36] Jin Ma and Jiongmin Yong. *Forward-backward stochastic differential equations and their applications*, volume 1702 of *Lecture Notes in Math.* Springer-Verlag, Berlin, 1999.

[37] Jacques Magnen and Jérémie Unterberger. The scaling limit of the KPZ equation in space dimension 3 and higher. *J. Stat. Phys.*, 171(4):543–598, 2018.

[38] Xuerong Mao. Stochastic stabilization and destabilization. *Systems Control Lett.*, 23(4):279–290, 1994.

[39] Chiranjib Mukherjee, Alexander Shamov, and Ofer Zeitouni. Weak and strong disorder for the stochastic heat equation and continuous directed polymers in $d \geq 3$. *Electron. Commun. Probab.*, 21, 2016.

[40] Szymon Peszat and Jerzy Zabczyk. Stochastic evolution equations with a spatially homogeneous Wiener process. *Stochastic Process. Appl.*, 72(2):187–204, 1997.

[41] Daniel W. Stroock and S. R. Srinivasa Varadhan. *Multidimensional diffusion processes*. Classics in Mathematics. Springer-Verlag, Berlin, 2006. Reprint of the 1997 edition.

[42] Gianmario Tessitore and Jerzy Zabczyk. Invariant measures for stochastic heat equations. *Probab. Math. Statist.*, 18(2, Acta Univ. Wratislav. No. 2111):271–287, 1998.

[43] Shinzo Watanabe and Toshio Yamada. On the uniqueness of solutions of stochastic differential equations. II. *J. Math. Kyoto Univ.*, 11:553–563, 1971.

[44] Toshio Yamada and Shinzo Watanabe. On the uniqueness of solutions of stochastic differential equations. *J. Math. Kyoto Univ.*, 11:155–167, 1971.