# A critical stochastic heat equation with long-range noise

September 30, 2025

# Alexander Dunlap,<sup>1</sup> Martin Hairer,<sup>2,3</sup> and Xue-Mei Li<sup>2,3</sup>

- <sup>1</sup> Duke University, USA, email: alexander.dunlap@duke.edu
- $^2$  EPFL, Switzerland, email: {martin.hairer,xue-mei.li}@epfl.ch
- <sup>3</sup> Imperial College London, UK

#### **Abstract**

We consider a semilinear stochastic heat equation in spatial dimension at least 3, forced by a noise that is white in time with a covariance kernel that decays like  $|x|^{-2}$  as  $|x| \to \infty$ . We show that in an appropriate diffusive scaling limit with a logarithmic attenuation of the noise, the pointwise statistics of the solution can be approximated by the solution to a forward-backward stochastic differential equation (FBSDE). The scaling and structure of the problem is similar to that of the two-dimensional stochastic heat equation forced by an approximation of space-time white noise considered by the first author and Gu (*Ann. Probab.*, 2022). However the resulting FBSDE is different due to the long-range correlations of the noise.

# **Contents**

1	Introduction		2	
	1.1	Properties of the forward-backward SDE	5	
	1.2	The multiplicative-noise case and the existence of a phase transition	5	
	1.3	Proof strategy and organization of the paper	6	
	1.4	Notation	7	
	1.5	Acknowledgments	7	
2	The forward-backward SDE		7	
	2.1	Explicit solutions	8	
3	Inte	gral estimates 9		
4	Moı	Moment bound		
5	Tur	ning off the noise	14	
6	Dec	orrelation estimate		
	The approximate decoupling function		21	
	7.1	Definition and regularity	21	
	7.2	Statistics of the averaged fields	23	
8	Analysis of the Markov chain		26	
	8.1	The time scales and definition of the Markov chain	26	
	8.2	Convergence of the Markov chain to the forward-backward SDE	30	
9	Ana	lysis of the second moment in the scalar linear case	31	

#### A Convergence of Markov chains to diffusions

# 1 Introduction

In this paper, we are interested in semilinear stochastic heat equations roughly of the form

$$du_t = \frac{1}{2}\Delta u_t + \sigma(u_t)dW_t, \tag{1.1}$$

where  $(dW_t)$  is a white-in-time noise with spatial covariance given by the Riesz kernel  $|\cdot|^{-2}$ . We note two important features of this stochastic heat equation. First, the correlation function is not compactly supported, and indeed has a rather heavy non-integrable tail, so we consider the noise to have "long-range correlations" like those considered by Gerolla and the second two authors in [13]. Second, this equation is *critical*, in that if for some  $\lambda > 0$  we set  $u^{\lambda}(t,x) = u(\lambda^2 t, \lambda x)$ , then  $u^{\lambda}$  formally satisfies the same equation (1.1). Stochastic heat equations exhibiting this type of formal scale-invariance of have been of significant recent interest in the literature, which we will review below.

Even in the additive and multiplicative cases  $\sigma(u) = \alpha \in \mathbb{R}$  or  $\sigma(u) = \beta u$  for some  $\beta \in \mathbb{R}$ , the equation (1.1) does not yield function-valued solutions in spatial dimension  $d \geq 2$ . Since we do not expect any additional regularity from taking a more general  $\sigma$ , it is not clear how to interpret the nonlinearity  $\sigma(u)$  for general  $\sigma$ , nor (if this quantity can be interpreted) the product with the irregular noise. See [22], discussed further below, for *measure*-valued solutions in the multiplicative case. One way to address this difficulty is to regularize the covariance at some small scale  $\rho > 0$ , and then try to take a limit  $\rho \to 0$ . In this work, we show that if in addition to mollifying the nonlinearity in this way we also attenuate by a logarithmic factor in  $\rho$ , then we can obtain interesting limits of the pointwise values of the solution as  $\rho \downarrow 0$ .

Let us now introduce our setting precisely. Let  $\mathfrak{m}$ ,  $\mathfrak{n}$ ,  $d \in \mathbb{N}$  with  $d \geq 3$ . We study the  $\mathbb{R}^{\mathfrak{m}}$ -valued stochastic heat equation

$$du_t(x) = \frac{1}{2} \Delta u_t(x) dt + \frac{\sigma(u_t(x))}{\sqrt{\log \rho^{-1}}} dW_t^{\rho}(x), \qquad t \in \mathbb{R}, x \in \mathbb{R}^d.$$
 (1.2)

Here,  $\Delta$  denotes the spatial Laplacian and  $\sigma \colon \mathbb{R}^{\mathfrak{m}} \to \mathbb{R}^{\mathfrak{m}} \otimes \mathbb{R}^{\mathfrak{n}}$  is a uniformly Lipschitz nonlinearity, where the space of  $\mathfrak{m} \times \mathfrak{n}$  matrices  $\mathbb{R}^{\mathfrak{m}} \otimes \mathbb{R}^{\mathfrak{n}}$  is equipped with the Frobenius norm  $|\sigma|_F = \sqrt{\sigma \sigma^T}$  (the  $L^2$  norm of the vector of singular values). The  $\mathbb{R}^{\mathfrak{n}}$ -valued noise  $(\mathrm{d}W_t^{\rho}(x))$  is adapted to a temporal filtration  $\{\mathcal{F}_t\}_t$  and has correlation structure

$$\mathbb{E}\left[dW_t^{\rho}(x)dW_{t'}^{\rho}(x')^{\top}\right] = \delta(t - t')R^{\rho}(x - x')\mathrm{Id}_{\mathfrak{n}},\tag{1.3}$$

where we define

$$R(x) = |x|^{-2}$$
 and  $R^{\rho} = G_{2\rho} * R$ . (1.4)

Here,  $G_t(x) = (2\pi t)^{-d/2} \mathrm{e}^{-|x|^2/(2t)}$  is the standard d-dimensional heat kernel and \* denotes spatial convolution. In the sequel, we will use the notation  $\mathcal{G}_t v = G_t * v$  for convolution with the heat kernel. We consider solutions to (1.2) in the mild sense, namely

$$u_t^{\rho}(x) = \mathcal{G}_t u_0(x) + \frac{1}{\sqrt{\log \rho^{-1}}} \int_0^t \mathcal{G}_{t-s} [\sigma(u_s^{\rho}) \, dW_s^{\rho}](x). \tag{1.5}$$

Our aim is to study pointwise limits of (1.5), as  $\rho \downarrow 0$ , in this setting of both *scaling criticality* and *long-range dependence*. Our main result is a close analogue of that in

[12, 10], where the noise covariance is taken to be a scale- $\rho^{1/2}$  approximation of a delta function (in particular the correlation function has compact support). Namely, we show that the law of the solution at any given location asymptotically coincides with the composition of the heat flow and the law of a suitable forward-backward stochastic differential equation (FBSDE). The FBSDE arising in our case is

$$d\Gamma_{a,Q}(q) = J(Q - q, \Gamma_{a,Q}(q))dB(q), \qquad q \in (0,Q);$$
(1.6a)

$$\Gamma_{a,O}(0) = a; \tag{1.6b}$$

$$J(r,b) = \frac{\mathbb{E}[\sigma(\Gamma_{b,r}(r))]}{\sqrt{2(d-2)}}, \qquad r \in (0,Q),$$
(1.6c)

where  $(B(q))_{q \in [0,Q]}$  is a standard  $\mathbb{R}^n$ -valued Brownian motion. We will show in Proposition 2.1 below that, if  $\operatorname{Lip}(\sigma) < \sqrt{2(d-2)}$ , then there is a unique  $J \colon [0,1] \to \mathbb{R}^m \otimes \mathbb{R}^m$  such that the resulting solutions to (1.6a-b) satisfy (1.6c). In other words, the FBSDE (1.6) has a unique solution for  $Q \in [0,1]$ . We note that the diffusivity in (1.6a) can alternatively be written as

$$J(Q - q, \Gamma_{a,Q}(q)) = \frac{\mathbb{E}[\sigma(\Gamma_{b,Q-q}(Q - q))]}{\sqrt{2(d - 2)}} \bigg|_{b = \Gamma_{a,Q}(q)} = \frac{\mathbb{E}[\sigma(\Gamma_{a,Q}(Q)) \mid \Gamma_{a,Q}(q)]}{\sqrt{2(d - 2)}}.$$
(1.7)

This is because the process  $(\Gamma_{a,Q}(q+r))_{r\in[0,Q-q]}$  solves the SDE problem

$$\Gamma_{a,Q}(q+r) = \Gamma_{a,Q}(q) + \int_0^r J(Q-q-s,\Gamma_{a,Q}(q+s)) dB(q+s), \qquad r \in [0, Q-q],$$

which, up to a time translation of B, is the same integral equation as that solved by  $(\Gamma_{b,Q-q}(r))_{r\in[0,Q-q]}$  if we condition on  $\Gamma_{a,Q}(q)=b$  (in the sense of stochastic flows). Therefore, taking r=Q-q, we obtain  $\mathrm{Law}(\Gamma_{a,Q}(Q)\mid\Gamma_{a,Q}(q)=b)=\mathrm{Law}(\Gamma_{b,Q-q}(Q-q))$ , and (1.7) follows. The function J is called the *decoupling function* of the FBSDE. We discuss it in more detail in Section 1.1 below.

We denote by  $\mathcal{U}_{s,t}^{\rho}$  the (nonlinear) propagator of (1.2), so  $\mathcal{U}_{s,t}^{\rho}u_s$  denotes the solution at time t of (1.9) with initial data  $u_s:\mathbb{R}^d\to\mathbb{R}^m$  at time s. Our main theorem states:

**Theorem 1.1.** Suppose that  $\text{Lip}(\sigma) < \sqrt{d-2}$ . Then, for any (possibly random) initial condition  $u_0$ , independent of the noise W, such that

$$\sup_{x\in\mathbb{R}^d}\mathbb{E}|u_0(x)|^{\ell}<\infty$$

for some  $\ell > 2$ , the following holds for any fixed t > 0 and  $x \in \mathbb{R}^d$ ,

$$\lim_{\rho \downarrow 0} W^{2} ((\mathcal{U}_{0,t}^{\rho} u_{0})(x), \Gamma_{\mathcal{G}_{t} u_{0}(x), 1}(1)) = 0, \tag{1.8}$$

where  $W^2$  denotes the Wasserstein-2 distance.

Remark 1.2. While the restriction  $\beta < \sqrt{d-2}$  is used in the proof, it is unlikely that this is optimal. In fact, Theorem 1.5 below strongly suggests that there is no critical value of  $\beta$  in the present case of noise with long-range correlations. Similarly, we expect that the restriction  $\beta < \sqrt{2(d-2)}$  for the well-posedness of (1.6a) is suboptimal as well.

*Remark* 1.3. If *U* solves the SPDE

$$dU_t(x) = \frac{1}{2}\Delta U_t(x)dt + \frac{\sigma(U_t(x))}{\sqrt{\log \rho^{-1}}}dW_t^{-1}(x),$$
(1.9)

then the law of its diffusively scaled solution  $(U_{\rho^{-1}t}(\rho^{-1/2}x))_{t,x}$  is the same as the law of  $(u_t(x))_{t,x}$ , as can be seen by the scaling relations  $\delta(t\rho) = \frac{1}{\rho}\delta(t)$  and  $R_{2\rho}^{\rho}(\frac{x}{\sqrt{\rho}}) = G_2*R$ . Thus, the limit  $\rho \downarrow 0$  of the SPDE (1.2) corresponds to the long-time, large-space behavior of an SPDE with fixed noise covariance. Note however that, in order to obtain pointwise fluctuations of order 1, it is still required to attenuate the noise strength as time is taken to infinity, hence the continued presence of the division by  $\sqrt{\log \rho^{-1}}$  in (1.9).

At a high level, the reason for the appearance of the FBSDE is the same as in [12, 10], and is discussed in more detail in their respective introductions. Roughly speaking, the logarithmic attenuation of the noise strength means the quadratic variation of the stochastic integral appearing in (1.5) has a type of mean-field behavior and thus can be approximated by the solution to a one-dimensional equation. The self-similar structure of the equation allows us to write this one-dimensional equation as an FBSDE.

The key novelty in the present work however, is that the FBSDE [12, (1.5-1.7)] is qualitatively different. In particular, [12, (1.7)] reads

$$J(q, a) = c_1(\mathbb{E}[\sigma^2(\Gamma_{a,q}(q))])^{1/2}, \tag{1.10}$$

rather than  $J(q, a) = c_2 \mathbb{E}[\sigma(\Gamma_{a,q}(q))]$  as in (1.6c). The constants  $c_1, c_2 \in (0, \infty)$  are a matter of normalization and irrelevant for the present discussion.

The reason for the difference in J comes from the covariance structure of the noise. The quadratic variation of the stochastic integral in (1.5) is

$$\frac{1}{\log \rho^{-1}} \int_0^t \iint G_{t-s}(x-y_1) G_{t-s}(x-y_2) R^{\rho}(y_1-y_2) \sigma(u_s(y_1)) \sigma^{\top}(u_s(y_2)) \, \mathrm{d}y_1 \, \mathrm{d}y_2 \, \mathrm{d}s.$$
(1.11)

In the case of compactly supported covariance kernel considered in [12, 10], the quantity  $\sigma(u_s(y_1))\sigma^{\top}(u_s(y_2))$  can be approximated by  $(\sigma\sigma^{\top})(s,u_s((y_1+y_2)/2))$ , since the compactly supported covariance kernel  $R^{\rho}$  forces  $y_1 \approx y_2$ . This leads to the appearance of  $\sigma^2$  in (1.10). On the other hand, in the present setting when  $R^{\rho}$  is long-range, the spatial decorrelation of  $u_s$  implies that the spatial integral in (1.11) more closely resembles  $\mathbb{E}[\sigma(u_s)]\mathbb{E}[\sigma(u_s)]^{\top}$ , which leads to the form of (1.6c). In this discussion, we have glossed over the fact that the quadratic variation in fact has to be estimated "at every scale," and the expectations here really need to be conditioned on the spatial average of  $u_s$  at the present scale, which leads to the conditioning on the right side of (1.7).

This phenomenon is very similar to that observed in comparing the results of [16] with those of [13] on Edwards–Wilkinson fluctuations of stochastic heat equations in settings with supercritical scaling. In [16], compactly supported noise is considered, and the effective noise strength of the resulting Edwards–Wilkinson field involves the microscopic two-point correlation function of  $\sigma$  applied to the stationary solution. (See [16, (1.3)].) In [13], long-range noise is considered, and the effective noise strength is simply the expectation of  $\sigma$  applied to the stationary solution and spatial correlation survives in the limit. (See [13, Theorem 1.3].)

#### 1.1 Properties of the forward-backward SDE

In order to make full use of Theorem 1.1, one of course needs to understand the solutions to the FBSDE (1.6). We refer to [21] for an introduction to the theory of FBSDEs, and [12, 10, 11] for more detailed discussion of an FBSDE quite similar (but, as noted above, not identical) to (1.6).

A standard application of Itô's formula (see e.g. [21, §8.2] or [10, §3]) shows that, at least formally speaking, the decoupling function *J* satisfies the quasilinear PDE

$$\partial_{q} J_{k,\ell}(q,b) = \frac{1}{2} \operatorname{tr}[(JJ^{\top}) \nabla_{b}^{2} J_{k,\ell}](q,b), \quad q > 0, b \in \mathbb{R}^{\mathfrak{m}}, 1 \le k \le \mathfrak{m}, 1 \le \ell \le \mathfrak{n}; \quad (1.12a)$$

$$J(0,b) = \frac{\sigma(b)}{\sqrt{2(d-2)}}, \qquad b \in \mathbb{R}^{\mathfrak{m}}. \quad (1.12b)$$

Here we use the notation  $\nabla_b^2$  for the Hessian. The reason we say this derivation of the PDE problem (1.12) is only formal is that the equation (1.12a) may be degenerate parabolic, as we have not made any assumption to guarantee that  $JJ^{\top}$  has full rank. The interpretation of the problem (1.12) can be made precise as in [10, §3]. We do not go into detail in the present work since we will not need this interpretation here, although it is implicit in proof of Proposition 1.4 below. The PDE formulation has been useful in [11] in studying the FBSDE derived in [12, 10]. We expect that a similar study of (1.12) could be fruitful in obtaining a deeper understanding of the decoupling function in the present setting.

Of course, for most choices of  $\sigma$ , we do not expect to find an explicit solution to the FBDSE (1.6). However, we know of the following family of  $\sigma$ s where the solution is explicit:

**Proposition 1.4.** Suppose that  $\mathfrak{m} = \mathfrak{n} = 1$ ,  $\alpha, \beta \geq 0$  and  $\sigma(b) = \sqrt{\alpha + \beta b^2}$ . The corresponding solution to the FBSDE (1.6) satisfies

$$J(q,b) = \sqrt{\frac{\alpha e^{\beta q/(4(d-2))} + \beta b^2}{2(d-2)}}.$$
(1.13)

See Section 2.1 for a proof of Proposition 1.4, which essentially consists of checking that this choice of J solves (1.12) and justifying the PDE in this setting. Of particular interest are the cases when one of  $\alpha$  and  $\beta$  is 0. When  $\alpha \geq 0$  and  $\beta = 0$ , we are considering the case of additive noise. In this case, the SPDE (1.2) is simple to analyze, as the solution is (conditional on the initial data) a Gaussian process. This is consistent with the fact that, according to (1.13), the decoupling function J is constant in this case, so the solutions to the FBSDE will also be Gaussian.

# 1.2 The multiplicative-noise case and the existence of a phase transition

The case  $\alpha=0$  and  $\beta\geq 0$  in (1.13) corresponds to the multiplicative noise case  $\sigma(u)=\beta u$ , which is of particular interest due to its connection both with directed polymers and with the KPZ equation via the Cole–Hopf transform [3, 14]. The multiplicative problem with compactly supported noise covariance function in d=2 has been studied extensively. Of particular relevance to the present work is [7], in which it was shown that, for  $\beta<\sqrt{2\pi}$ , the limiting pointwise distribution is log-normal. From the point of view of the FBSDE, this is because the solution to the corresponding FBDSE is a deterministic time-change of a geometric Brownian motion; see [12, §1.3]. The log-normal random variable blows up as  $\beta\uparrow\sqrt{2\pi}$ , pointing to the phase transition

at  $\beta = \sqrt{2\pi}$ . At  $\beta = \sqrt{2\pi}$  (in fact in a window of size  $O(1/\sqrt{\log \rho^{-1}})$  around this critical value), a measure-valued process called the *critical stochastic heat flow* has been constructed in [5] and characterized axiomatically in [27]; see also [2, 4, 17].

In our present setting, when  $\alpha=0$  we see from (1.13) that  $J(q,b)=b\sqrt{\beta/(2(d-2))}$ , so that the solution to the FBSDE (1.6a–c) is a geometric Brownian motion. However, unlike in the setting considered in [7], it remains finite for all  $\beta<\infty$ . This suggests that the requirement imposed in [7] that  $\beta$  be smaller than a critical value may be unnecessary in the present setting. As further evidence for this, we prove the following theorem.

**Theorem 1.5.** Let  $\mathfrak{m} = \mathfrak{n} = 1$ ,  $\beta \in (0, \infty)$ ,  $\sigma(u) = \beta u$ , and  $(u_t(x))$  solve (1.2) with initial condition  $u_0 \equiv 1$ . Then, for any  $T \in (0, \infty)$ , we have a constant  $C = C(\beta, T) < \infty$ , independent of  $\rho$ , such that

$$\sup_{t \in [0,T], x \in \mathbb{R}^d} \mathbb{E} |u_t(x)|^2 \le C.$$

In other words, unlike for the problem studied in [2, 7], there is no value of  $\beta$  at which the pointwise second moment of the solution blows up. Thus, if there is to be a phase transition, the behavior must be significantly more subtle.

A similar problem to (1.2) with  $\mathfrak{m}=\mathfrak{n}=1$  and  $\sigma(u)=\beta u$ , but *without* the division by  $\sqrt{\log \rho^{-1}}$ , was studied by Mueller and Tribe in [22]. (See [20] for some limiting behaviors of this process.) In that case, the problem with  $\rho=0$  can be studied directly, and a measure-valued solution is obtained. We can alternatively think of this problem as (1.2) with  $\sigma(u)=u\sqrt{\log \rho^{-1}}$ . Thus, if we believe the conjecture that the analogue of the critical value  $\beta=\sqrt{2\pi}$  from [5] is  $\beta=\infty$  in this setting, then we could see the process constructed in [22] as an analogue of the critical stochastic heat flow.

A discrete version of the multiplicative problem with similar noise correlation decay as  $|x| \to \infty$  but in spatial dimension d=2 has recently been considered in [8]. In that setting, similar log-normal behavior is again observed, but in that case the power of  $\log \rho^{-1}$  in the attenuation is different, and there is a phase transition in  $\beta$ . The study of the semilinear case in that setting is an interesting problem for future work.

#### 1.3 Proof strategy and organization of the paper

The large-scale structure of the proof of Theorem 1.1 is similar to the structure of the proof in [12]. The key tool that allows us to approximate the solution to the (infinite-dimensional) SPDE by the solution to the (finite-dimensional) FBSDE is the technique of turning off the noise on approximate time intervals, which are "long" from the point of view of the diffusive scaling of the heat equation but "short" from the point of view of the effect of the noise on the solution. This means that turning off the noise on these intervals does not affect the solution much but allows the solution to relax enough to become approximately constant on appropriate mesoscopic scales, which allows us to approximate the problem by a finite-dimensional Markov chain. The relevant key estimates are proved in Section 5 and the Markov chain is defined and studied in Section 8.

The remaining step, and the source of key difference between our setting and that of [12, 10], is to determine the form of the FBSDE relation (1.6c). To study this, we must compute the variance of the increments of the approximating Markov chain. The key estimate in this direction is Proposition 7.5 in Section 7.2, which relies on an estimate of the covariance of a function of the solution established in Section 6.

This covariance estimate resembles [13, Lemma 2.13] and is proved using based on Malliavin calculus techniques.

We note that previous works on semilinear stochastic heat equations in the critical dimension at subcritical temperature have proved a wider range of results than we show here, e.g. multipoint statistics [7, 12, 10] and Edwards–Wilkinson behavior for the rescaled and recentered random field [7, 25, 9]. The logarithm of the solution in the linear case has also been studied, and Edwards–Wilkinson behavior for the corresponding field has been proved [6, 15] (see also [14] for a version of this result in the long-range supercritical regime), as well as some mesoscopic averaging results [26]. We believe that analogues of these results should hold in the present setting as well. We have chosen to keep the present paper as simple as possible in order to focus on one of the key novelties in this setting, namely the different form of the FBSDE. We leave the extensions in these various directions to future work (perhaps once the question on the necessity of the Lipschitz bound on  $\sigma$  can be addressed as well).

The paper is organized as follows. In Section 2, a well-posedness theory for the FBSDE (1.6) is established. In Section 3, we establish some results about various integrals that are used in the results throughout the paper. In Section 4, we prove an upper bound on the moments of the solution, adapting the proof of the moment bound in [10]. In Section 5, we implement the "turning off the noise" strategy described above. In Section 6, we prove the necessary bounds on covariances of functions of the solution. In Section 7, we study the approximate decoupling function, and in particular prove the key estimate Proposition 7.5. In Section 8, we define and study the approximating Markov chain and prove Theorem 1.1. Finally, in Section 9, we prove Theorem 1.5.

#### 1.4 Notation

We let  $\mathcal{X}_t^\ell$  denote the space of  $\mathcal{F}_t$ -measurable random fields v such that the norm

$$|||v|||_{\ell} \coloneqq \sup_{x \in \mathbb{R}^d} (\mathbb{E}|v(x)|^{\ell})^{1/\ell}$$
 (1.14)

is finite. We will mostly work with the 2-norm, so we abbreviate  $\| \cdot \| := \| \cdot \|_2$  and  $\mathcal{X}_t := \mathcal{X}_t^2$ . For matrices, we use the notation  $|\cdot|_F$  for the Frobenius norm. Unless otherwise stated, we assume that  $\mathbb{R}^{\mathfrak{m}} \otimes \mathbb{R}^{\mathfrak{n}}$  is equipped with the Frobenius norm. We also use the Japanese bracket notation  $\langle u \rangle = \sqrt{1 + |u|^2}$  for  $u \in \mathbb{R}^k$  (for any  $k \in \mathbb{N}$ ).

#### 1.5 Acknowledgments

A. D. was supported by the National Science Foundation under grants DMS-2002118 and DMS-2346915. X.-M. L. acknowledges support by Swiss National Science Foundation project mint 10000849 and UKRI grant EP/V026100/1.

# 2 The forward-backward SDE

In this section we establish some basic results about the forward-backward SDE. The notation and arguments in this section follow [10].

Suppose that  $g: [0,Q] \times \mathbb{R}^{\mathfrak{m}} \to \mathbb{R}^{\mathfrak{n}} \otimes \mathbb{R}^{\mathfrak{n}}$  is continuous and also that it is Lipschitz in its second argument, uniformly in its first argument. For each  $a \in \mathbb{R}^{\mathfrak{m}}$ , we let  $(\Theta_{a,O}^g)_{q \in [0,Q]}$  solve the SDE

$$d\Theta_{a,O}^{g}(q) = g(Q - q, \Theta_{a,O}^{g}(q))dB(q), \qquad q \in (0,Q);$$
 (2.1a)

$$\Theta_{a,O}^g(0) = a. (2.1b)$$

For a Lipschitz function  $f: \mathbb{R}^m \to \mathbb{R}^m \otimes \mathbb{R}^n$ , we define

$$Q_f g(Q, a) = \mathbb{E}[f(\Theta_{a, Q}^g(Q))], \tag{2.2}$$

and we note that (1.6c) is equivalent to the condition

$$Q_{\sigma/\sqrt{2(d-2)}}J = J. \tag{2.3}$$

We define the function space

$$\mathcal{A}_Q \coloneqq \Big\{ J \in C([0,Q] \times \mathbb{R}^{\mathfrak{m}}; \mathbb{R}^{\mathfrak{m}} \otimes \mathbb{R}^{\mathfrak{n}}) \ : \ \sup_{q \in [0,Q]} \mathrm{Lip}(J(q,\cdot)) < \infty \Big\}.$$

**Proposition 2.1.** Let  $f \in \text{Lip}(\mathbb{R}^m; \mathbb{R}^m \otimes \mathbb{R}^n)$ . For any  $Q \in [0, \text{Lip}(f)^{-2})$ , there is a unique  $J \in \mathcal{A}_Q$  such that  $Q_f J = J$ .

*Proof.* By Jensen's inequality and the convexity of  $|\cdot|_F^2$ , the following analogue of [10, (2.19)] holds for all  $Q \in [0, \operatorname{Lip}(f)^{-2})$  and all  $a \in \mathbb{R}^m$ :

$$|Qg(Q,a)|_{\mathrm{F}}^2 = |\mathbb{E}[\sigma(\Theta_{a,Q}^g(Q))]|_{\mathrm{F}}^2 \le \mathbb{E}|\sigma(\Theta_{a,Q}^g(Q))|_{\mathrm{F}}^2. \tag{2.4}$$

The proof of [10, Proposition 2.5] now works essentially verbatim, with (2.4) replacing [10, (2.25)] and the similarly obvious fact that

$$|\mathbb{E}[\sigma_1] - \mathbb{E}[\sigma_2]|_{\mathrm{F}}^2 \leq \mathbb{E}|\sigma_1 - \sigma_2|_{\mathrm{F}}^2$$

(for any matrix-valued random variables  $\sigma_1, \sigma_2$ ) replacing [10, Proposition A.2].  $\Box$ 

Proposition 2.1 and standard well-posedness results for SDEs immediately imply that if  $\text{Lip}(\sigma) < \sqrt{2(d-2)}$ , then the FBSDE (1.6) has a unique solution.

# 2.1 Explicit solutions

In this section we give a proof of Proposition 1.4 regarding a particular family of explicit solutions to the FBSDE.

*Proof of Proposition 1.4.* Morally, the proof consists of checking that the solution (1.13) satisfies the quasilinear heat equation (1.12). However, since we have not established that PDE problem in this setting, we will work via Itô's formula directly.

Let  $f(q) = \alpha e^{\beta q/(4(d-2))}$ . With this definition and the choice (1.13) of J, the SDE (1.6a–b) becomes

$$\begin{split} \mathrm{d}\Gamma_{a,Q}(q) &= \sqrt{\frac{f(Q-q) + \beta \Gamma_{a,Q}(q)^2}{2(d-2)}} \mathrm{d}B(q), \qquad q \in (0,Q); \\ \Gamma_{a,O}(0) &= a. \end{split}$$

Let

$$N_q(Q, a) := \mathbb{E}[J(Q - q, \Gamma_{a,Q}(q))]. \tag{2.5}$$

We want to apply Itô's formula to  $q \mapsto N_q(Q, a)$ . This is not a problem if  $\alpha > 0$ , since in that case J is smooth. If  $\alpha = 0$ , then J fails to be smooth at (0, 0), but in that case the solution to the SDE will almost surely never reach 0, so we can use an approximation argument analogous to [10, Section 3]. Thus, using that

$$\partial_q J(q,b) = \frac{f'(q)}{\sqrt{2(d-2)(f(q)+\beta b^2)}}$$

Integral estimates 9

and

$$\partial_b^2 J(q, b) = \frac{\beta f(q)}{\sqrt{2(d-2)}(f(q) + \beta b^2)^{3/2}},$$

we conclude that

$$\partial_q N_q(Q, a) = \mathbb{E}\left[\frac{-f'(Q - q) + \beta f(Q - q)/(4(d - 2))}{\sqrt{2(d - 2)(f(Q - q) + \beta \Gamma_{a,Q}(q)^2}}\right] = 0.$$
 (2.6)

This means that

$$\frac{\mathbb{E}[\sigma(\Gamma_{a,Q}(Q))]}{\sqrt{2(d-2)}} \stackrel{\text{(1.13)}}{=} \mathbb{E}[J(0,\Gamma_{a,Q}(Q))] \stackrel{\text{(2.5)}}{=} N_Q(Q,a) \stackrel{\text{(2.6)}}{=} N_0(Q,a) \stackrel{\text{(2.5)}}{=} J(Q,a),$$

which shows that (1.6c) is satisfied, thus completing the proof.

# 3 Integral estimates

In this section we collect some estimates and identities for various integrals that we will use in the sequel. The proofs of these results are somewhat simpler when the Gaussian kernel is used as the mollifier in (1.4), so we stick to this choice for expositional clarity. Indeed, the results of the paper can be extended easily to more general choices of mollifier.

We note that

$$\frac{1}{|x|^2} = \frac{1}{2} (2\pi)^{d/2} \int_0^\infty v^{d/2 - 2} G_v(x) \, \mathrm{d}v, \qquad x \in \mathbb{R}^d \setminus \{0\}, \tag{3.1}$$

so that

$$R^{\rho}(x) = G_{2\rho} * |\cdot|^{-2} = \frac{1}{2} (2\pi)^{d/2} \int_{0}^{\infty} v^{d/2 - 2} G_{\nu + 2\rho}(x) \, \mathrm{d}\nu. \tag{3.2}$$

We will also frequently use the fact that

$$(G_t * R^{\rho})(0) \stackrel{(3.2)}{=} \frac{1}{2} \int_0^{\infty} \frac{v^{d/2 - 2}}{(t + v + \rho)^{d/2}} \, \mathrm{d}v = \frac{1}{(d - 2)(t + 2\rho)},\tag{3.3}$$

and its immediate consequence that

$$\iint R^{\rho}(y_1 - y_2) \prod_{i=1}^{2} G_{T-t}(x - y_i) \, \mathrm{d}y_1 \, \mathrm{d}y_2 = \frac{1}{2(d-2)(T-t+\rho)}.$$
 (3.4)

We will also use the fact that, since  $\log |\cdot|^{-1}$  is a superharmonic function on  $\mathbb{R}^d$  (here we use the assumption that d > 2),

$$(G_t * \log|\cdot|^{-1})(x) \le \log|x|^{-1} \qquad \text{for all } x \in \mathbb{R}^d \setminus \{0\}.$$
 (3.5)

Moreover, we have an absolute constant  $C < \infty$  such that  $R^1(x) \le C|x|^{-2}$  for all  $x \in \mathbb{R}^d \setminus \{0\}$ , which means that

$$R^{\rho}(x) = \rho^{-1}R^{1}(\rho^{-1/2}x) \le C\rho^{-1}|\rho^{-1/2}x|^{-2} = C|x|^{-2}$$
 for all  $\rho > 0$ . (3.6)

We will also need the following estimate.

Integral estimates 10

**Proposition 3.1.** There is a constant C depending only on d such that

$$\int_0^t (G_{2r} * R^{\rho})(x) \, \mathrm{d}r \le C \left( \log \frac{t^{1/2}}{|x|} + 1 + \frac{|x|}{t^{1/2}} \right). \tag{3.7}$$

We note that the estimate (3.7) is quite suboptimal when  $|x| \gg t^{1/2}$ , but in our applications this will not matter since the decay of the heat kernel will dominate in the superdiffusive regime. We will prove Proposition 3.1 as the consequence of a few lemmas.

**Lemma 3.2.** We have, for all  $t, \rho > 0$ , that

$$\int_0^t (G_s * R^{\rho})(x) \, \mathrm{d}s = \frac{1}{d/2 - 1} \left( (G_{2\rho} - G_{t+2\rho}) * \log|\cdot|^{-1} \right)(x). \tag{3.8}$$

Proof. We can compute

$$F(x) := \frac{d-2}{(2\pi)^{d/2}} \int_0^t (G_s * R^{\rho})(x) \, ds$$

$$\stackrel{(3.2)}{=} (d/2 - 1) \int_0^t \int_0^{\infty} v^{d/2 - 2} G_{v+s+2\rho}(x) \, dv \, ds$$

$$= (d/2 - 1) \int_0^{\infty} G_{v+2\rho}(x) \int_0^{t \wedge v} (v - s)^{d/2 - 2} \, ds \, dv$$

$$= \int_0^{\infty} G_{v+2\rho}(x) v^{d/2 - 1} \, dv - \int_t^{\infty} G_{v+2\rho}(x) (v - t)^{d/2 - 1} \, dv$$

$$= \int_0^{\infty} (G_{v+2\rho}(x) - G_{v+t+2\rho}(x)) v^{d/2 - 1} \, dv.$$
(3.10)

Now, from this expression, we can write

$$\nabla F(x) = \int_0^\infty (\nabla G_{\nu+2\rho}(x) - \nabla G_{\nu+t+2\rho}(x)) \nu^{d/2-1} \, d\nu$$
$$= (G_{2\rho} * Q)(x) - (G_{t+2\rho} * Q)(x),$$

where we have defined

$$Q(x) := \int_0^\infty \nabla G_{\nu}(x) \nu^{d/2 - 1} \, d\nu = -x \int_0^\infty \nu^{d/2 - 2} G_{\nu}(x) \, d\nu$$

$$\stackrel{(3.1)}{=} -\frac{2^{1 - d/2} x}{\pi^{d/2} |x|^2} = \frac{2^{1 - d/2}}{\pi^{d/2}} \nabla (\log |\cdot|^{-1})(x).$$

Therefore, we have

$$\nabla F(x) = 2^{1-d/2} \pi^{d/2} \nabla ((G_{2\rho} - G_{t+2\rho}) * \log |\cdot|^{-1})(x).$$

Now it is clear from the definition (3.9) that  $\lim_{|x|\to\infty} F(x) = 0$ , and since we also observe that

$$\lim_{|x| \to \infty} \frac{2^{1-d/2}}{\pi^{d/2}} \left( (G_{2\rho} - G_{t+2\rho}) * \log|\cdot|^{-1} \right) (x) = 0,$$

we see that

$$F(x) = \frac{2^{1-d/2}}{\pi^{d/2}} \left( (G_{2\rho} - G_{t+2\rho}) * \log|\cdot|^{-1} \right) (x).$$

Again recalling (3.9) we obtain (3.8).

Moment bound 11

**Proposition 3.3.** There is a constant C depending only on d such that, for any t > 0 and  $x \in \mathbb{R}^d$ , we have

$$(G_t * \log |\cdot|^{-1})(x) \ge -\frac{1}{2} \log t - C \left(1 + \frac{|x|}{t^{1/2}}\right).$$
 (3.11)

As with Proposition 3.1, this bound is highly suboptimal for  $x \gg t^{1/2}$ , but again we do not need much precision in that regime.

*Proof.* Define  $f(x) := (G_t * \log |\cdot|^{-1})(x)$ . We can compute

$$f(0) = \int G_t(x) \log|x|^{-1} dx = \int G_1(x) \log|t^{1/2}x|^{-1} dx$$
$$= -\frac{1}{2} \log t + \int G_1(x) \log|x|^{-1} dx.$$
(3.12)

We can also compute

$$\nabla f(y) = -\int G_t(y-x) \frac{x}{|x|^2} dx,$$

SO

$$|\nabla f(y)| \le \int \frac{G_t(y-x)}{|x|} dx \le \int \frac{G_t(x)}{|x|} dx = t^{-1/2} \int \frac{G_1(x)}{|x|} dx.$$
 (3.13)

Here the second inequality is by [18]. Combining (3.12) and (3.13), we obtain (3.11).

*Proof of Proposition 3.1.* We can estimate

$$\int_{0}^{t} (G_{2s} * R^{\rho})(x) ds = \frac{1}{2} \int_{0}^{2t} (G_{s} * R^{\rho})(x) ds$$

$$\stackrel{(3.8)}{=} \frac{1}{d-2} \left( (G_{2\rho} - G_{2(t+\rho)}) * \log|\cdot|^{-1} \right)(x)$$

$$\leq \frac{1}{d-2} \left( \log|x|^{-1} + \frac{1}{2} \log t + C \left( 1 + \frac{|x|}{t^{1/2}} \right) \right) \leq C \left( \log \frac{t^{1/2}}{|x|} + 1 + \frac{|x|}{t^{1/2}} \right),$$

with the second inequality by (3.5) and (3.11), and where C has changed in the last inequality.

#### 4 Moment bound

In this section, we prove a bound on the moments of  $u_t(x)$ . The following proposition and its proof are adaptations of [10, Proposition 3.4]. Recall the notation  $\|\cdot\|_{\ell}$  from Section 1.4, as well as the Japanese bracket  $\langle\cdot\rangle$ .

**Proposition 4.1.** Suppose that  $\beta < \sqrt{2(d-2)}$  and  $M \in [0,\infty)$  are such that  $|\sigma(u)|_F^2 \le M + \beta^2 |u|^2$  for all  $u \in \mathbb{R}^m$ . Suppose also that  $(u_t(x))$  solves (1.2). Then, for all T > 0, we have a constant  $C = C(M, \beta, T) \in (1, \infty)$  and an  $\ell_0 = \ell_0(M, \beta, T) > 2$  such that, if  $\ell \in [2, \ell_0]$ ,  $t \in [0, T]$ , and  $\rho \in (0, C^{-1})$ , then

$$|||u_t|||_{\ell} \le C\langle |||u_0|||_{\ell}\rangle.$$
 (4.1)

Moment bound 12

*Proof.* We define the martingale

$$U_{t} = \mathcal{G}_{T-t}u_{t}(x) \stackrel{(1.5)}{=} \mathcal{G}_{T}u_{0}(x) + \frac{1}{\sqrt{\log \rho^{-1}}} \int_{0}^{t} \mathcal{G}_{T-s}[\sigma(u_{s}) dW_{s}^{\rho}](x) ,$$

as well as the differential quadratic variation

$$A_{t} := \frac{\mathrm{d}}{\mathrm{d}t} [U]_{t}$$

$$= \frac{1}{\log \rho^{-1}} \iint R^{\rho} (y_{1} - y_{2}) G_{T-t}(x - y_{1}) G_{T-t}(x - y_{2}) \sigma(u_{t}(y_{1})) \sigma(u_{t}(y_{2}))^{\top} \, \mathrm{d}y_{1} \, \mathrm{d}y_{2}.$$
(4.2)

Here, the notation  $[U]_t$  denotes the matrix whose (i, j) entry is the quadratic covariation of the ith and jth components of  $(U_t)_t$  at time t. We also define  $Y_t = |U_t|^{\ell}$ . By Itô's formula (as in [10, (4.15)]), we obtain

$$\mathrm{d} Y_t = \ell |U_t|^{\ell-2} U_t \cdot \mathrm{d} U_t + \frac{1}{2} \operatorname{tr} \left[ (\ell(\ell-2)|U_t|^{\ell-4} U_t^{\otimes 2} + \ell |U_t|^{\ell-2} \mathrm{Id}_{\mathfrak{m}}) \mathrm{d} [U]_t \right].$$

Integrating and taking expectations, noting that the local martingale term vanishes since it is indeed a martingale as  $U_t$  has all moments bounded uniformly in t in a compact set (for fixed  $\rho$ ), we see that

$$\mathbb{E}|u_{T}(x)|^{\ell} = \mathbb{E}Y_{T}$$

$$\leq |||u_{0}|||_{\ell}^{\ell} + \frac{1}{2}\mathbb{E}\left[\int_{0}^{T} \operatorname{tr}\left[(\ell(\ell-2)|U_{t}|^{\ell-4}U_{t}^{\otimes 2} + \ell|U_{t}|^{\ell-2}\operatorname{Id}_{\mathfrak{m}})\operatorname{d}[U]_{t}\right]\right]$$

$$= |||u_{0}|||_{\ell}^{\ell} + \frac{1}{2}\int_{0}^{T} \operatorname{tr}\mathbb{E}\left[(\ell(\ell-2)|U_{t}|^{\ell-4}U_{t}^{\otimes 2} + \ell|U_{t}|^{\ell-2}\operatorname{Id}_{\mathfrak{m}})A_{t}\right]\operatorname{d}t. \tag{4.3}$$

Now we estimate

$$\operatorname{tr}\left[|U_t|^{\ell-4}U_t^{\otimes 2}A_t\right] \le |U_t|^{\ell-2}\operatorname{tr}A_t,\tag{4.4}$$

so (4.3) becomes

$$\mathbb{E}|u_T(x)|^{\ell} \le |||u_0|||_{\ell}^{\ell} + \frac{1}{2}(\ell - 1) \int_0^T \mathbb{E}[|U_t|^{\ell - 2} \operatorname{tr} A_t] dt.$$
 (4.5)

Thus we want to estimate the quantity

$$\mathbb{E}\left[|U_{t}|^{\ell-2} \operatorname{tr} A_{t}\right] \\
\leq \frac{1}{\log \rho^{-1}} \mathbb{E}\left[|U_{t}|^{\ell-2} \iint R^{\rho} (y_{1} - y_{2}) \prod_{i=1}^{2} (G_{T-t}(x - y_{i}) |\sigma(u_{t}(y_{i}))|_{F} dy_{1} dy_{2}\right], \tag{4.6}$$

where we used Hölder's inequality, the definition (4.2) of  $A_t$ , and the submultiplicativity of the Frobenius norm. The assumption on  $\sigma$ , along with Hölder's inequality as in [10, (4.21)], imply that

$$|\sigma(u)|_{\mathsf{F}}^{\ell} \le 2^{\ell/2-1} [M^{\ell/2} + \beta^{\ell} |u|^{\ell}].$$

Using Young's product inequality and Jensen's inequality, we conclude that for any

Moment bound 13

(deterministic)  $Z \in (0, \infty)$ ,

$$\begin{split} & \mathbb{E} \big[ |U_{t}|^{\ell-2} |\sigma(u_{t}(y_{1}))|_{F} |\sigma(u_{t}(y_{2}))|_{F} \big] \\ & = \mathbb{E} \Big[ \Big( Z^{\frac{\ell-2}{\ell}} |U_{t}|^{\ell-2} \Big) \Big( Z^{-\frac{\ell-2}{2\ell}} |\sigma(u_{t}(y_{2}))|_{F} \Big) \Big( Z^{-\frac{\ell-2}{2\ell}} |\sigma(u_{t}(y_{2}))|_{F} \Big) \Big] \\ & \leq (1 - 2/\ell) Z \mathbb{E} |U_{t}|^{\ell} + \frac{1}{\ell Z^{\frac{\ell-2}{2}}} \sum_{i=1}^{2} \mathbb{E} |\sigma(u_{t}(y_{i}))|_{F}^{\ell} \\ & \leq (1 - 2/\ell) Z |||u_{t}|||_{\ell}^{\ell} + \frac{2^{\ell/2}}{\ell Z^{\frac{\ell-2}{2}}} \Big( M^{\ell/2} + \beta^{\ell} |||u_{t}|||_{\ell}^{\ell} \Big). \end{split}$$

We minimize the right side by taking

$$Z \coloneqq 2^{\frac{\ell-2}{\ell}} \left( \beta^{\ell} + \frac{M^{\ell/2}}{\|\|u_t\|\|_{\ell}^{\ell}} \right)^{2/\ell},$$

which yields

$$\mathbb{E}\left[|U_{t}|^{\ell-2}|\sigma(u_{t}(y_{1}))|_{F}|\sigma(u_{t}(y_{2}))|_{F}\right] \\
\leq (1-2/\ell)2^{(\ell-2)/2} \left(\beta^{\ell} + \frac{M^{\ell/2}}{\|u_{\ell}\|_{\ell}^{\ell}}\right) \||u_{t}\||_{\ell}^{\ell} + \frac{2^{\frac{\ell-2}{\ell}} \left(M^{\ell/2} + \beta^{\ell} \|u_{t}\|_{\ell}^{\ell}\right)}{\ell \left(\beta^{\ell} + \frac{M^{\ell/2}}{\|u_{t}\|_{\ell}^{\ell}}\right)^{(\ell-2)/\ell}} \\
= 2^{\frac{\ell-2}{2}} (1-1/\ell) \||u_{t}\||_{\ell}^{\ell-2} \left(M^{\ell/2} + \beta^{\ell} \||u_{t}\||_{\ell}^{\ell}\right)^{2/\ell} \\
\leq 2^{\frac{\ell-2}{2}} (1-1/\ell) \left(M \||u_{t}\||_{\ell}^{\ell-2} + \beta^{\ell} \||u_{t}\||_{\ell}^{\ell}\right) \\
\leq 2^{1-2/\ell} (1-1/\ell) \left(\frac{2}{\ell} M^{\ell/2} + \left(\beta^{\ell} + 1 - 2/\ell\right) \||u_{t}\||_{\ell}^{\ell}\right) \\
\leq \frac{2}{\ell} M^{\ell/2} + \left(\beta^{\ell} + 1 - 2/\ell\right) \||u_{t}\||_{\ell}^{\ell}. \tag{4.7}$$

Using (4.7) and (3.4) in (4.6), and then substituting into (4.5), we get

$$|||u_T|||_{\ell}^{\ell} \le |||u_0|||_{\ell}^{\ell} + \frac{\ell(\ell-1)}{4(d-2)\log\rho^{-1}} \int_0^T \frac{\frac{2}{\ell}M^{\ell/2} + (\beta^{\ell} + 1 - 2/\ell)||u_t|||_{\ell}^{\ell}}{T - t + \rho} dt.$$
 (4.8)

Defining  $\Upsilon^{\ell} \coloneqq \sup_{t \in [0,T]} |||u_t|||_{\ell}^{\ell}$ , we get

$$\Upsilon^{\ell} \leq \|u_0\|_{\ell}^{\ell} + \frac{\ell(\ell-1)}{4(d-2)} \left( 2M^{\ell/2} / \ell + (\beta^{\ell} + 1 - 2/\ell) \Upsilon^{\ell} \right) \log_{\rho^{-1}} (1 + T\rho^{-1}).$$
(4.9)

Now, since we assumed that  $\beta < \sqrt{2(d-2)}$ , there is an  $\ell_0 > 2$  such that, if  $\ell \in [2, \ell_0]$ , then for sufficiently small  $\rho$  we have

$$\frac{\ell(\ell-1)}{4(d-2)}(\beta^{\ell}+1-2/\ell)\log_{\rho^{-1}}(1+T\rho^{-1})<1,$$

and in this case we can rearrange (4.9) to imply (4.1) (with an appropriate choice of C).

# 5 Turning off the noise

In this section, we implement the "turning off the noise" strategy developed in [12, 10] for the case of compactly supported noise covariance. The estimates we obtain are very similar in form to the corresponding estimates of [12, §4].

First we show that we do not change the solution of the equation much if we "turn off" the noise for a sufficiently short period of time, compared to the time at which we evaluate the solution. In the statement of the proposition below, we note that  $\mathcal{U}_{\tau_2,t}\mathcal{G}_{\tau_2-\tau_1}u_{\tau_1}$  represents the solution of the equation run until time t but with the noise "turned off" on the time interval  $[\tau_1, \tau_2]$ , whereas  $\mathcal{U}_{\tau_1,t}u_{\tau_1}$  represents the solution of the equation run until time t without turning off the noise.

**Proposition 5.1.** For every T > 0 and  $\lambda < \sqrt{2(d-2)}$ , there is a  $C = C(T,\lambda) < \infty$  such that the following holds. Suppose that  $\operatorname{Lip}(\sigma) < \lambda$ . Let  $\tau_1 \in \mathbb{R}$  and let  $u_{\tau_1} \in \mathcal{X}_{\tau_1}$ . Then we have, for  $t > \tau_2 > \tau_1$ , that

$$\|\|\mathcal{U}_{\tau_{2},t}\mathcal{G}_{\tau_{2}-\tau_{1}}u_{\tau_{1}}-\mathcal{U}_{\tau_{1},t}u_{\tau_{1}}\|\|^{2} \leq C\langle \||u_{\tau_{1}}\||\rangle^{2}\frac{\log\left(\frac{t-\tau_{1}+\rho}{t-\tau_{2}+\rho}\right)+1}{\log\rho^{-1}}.$$

*Proof.* Let  $u_t = \mathcal{U}_{\tau_1,t}u_{\tau_1}$  and  $\tilde{u}_t = \mathcal{U}_{\tau_2,t}\mathcal{G}_{\tau_2-\tau_1}u_{\tau_1}$ . We can write

$$\tilde{u}_t(x) = \mathcal{G}_{t-\tau_1} u_{\tau_1}(x) + \frac{1}{\sqrt{\log \rho^{-1}}} \int_{\tau_2}^t \mathcal{G}_{t-s} [\sigma(\tilde{u}_s) dW_s^{\rho}](x)$$

while

$$u_t(x) = \mathcal{G}_{t-\tau_1} u_{\tau_1}(x) + \frac{1}{\sqrt{\log \rho^{-1}}} \int_{\tau_1}^t \mathcal{G}_{t-s} [\sigma(u_t) \,\mathrm{d}W_s^\rho](x).$$

Therefore, we have

$$\begin{split} u_t(x) - \tilde{u}_t(x) &= \frac{1}{\sqrt{\log \rho^{-1}}} \int_{\tau_1}^{\tau_2} \mathcal{G}_{t-s} [\sigma(u_s) \, \mathrm{d}W_s^{\rho}](x) \\ &+ \frac{1}{\sqrt{\log \rho^{-1}}} \int_{\tau_2}^t \mathcal{G}_{t-s} [(\sigma(u_s) - \sigma(\tilde{u}_s)) \, \mathrm{d}W_s^{\rho}](x). \end{split}$$

We can estimate the second moment as

$$\begin{split} \mathbb{E}|u_{t}(x) - \tilde{u}_{t}(x)|^{2} \\ &\leq \frac{1}{\log \rho^{-1}} \int_{\tau_{1}}^{\tau_{2}} \iint R^{\rho}(y_{1} - y_{2}) \mathbb{E}\left[\prod_{i=1}^{2} G_{t-s}(x - y_{i}) |\sigma(u_{s}(y_{i}))|_{F}\right] \mathrm{d}y_{1} \, \mathrm{d}y_{2} \, \mathrm{d}s \\ &+ \frac{1}{\log \rho^{-1}} \int_{\tau_{2}}^{t} \iint R^{\rho}(y_{1} - y_{2}) \mathbb{E}\left[\prod_{i=1}^{2} G_{t-s}(x - y_{i}) |\sigma(u_{s}(y_{i})) - \sigma(\tilde{u}_{s}(y_{i}))|_{F}\right] \\ &\qquad \qquad \qquad \mathrm{d}y_{1} \, \mathrm{d}y_{2} \, \mathrm{d}s \\ &\stackrel{(3.4)}{\leq} \frac{1}{2(d-2) \log \rho^{-1}} \int_{\tau_{1}}^{\tau_{2}} \frac{\||\sigma(u_{s})||^{2}}{t-s+\rho} \, \mathrm{d}s + \frac{\mathrm{Lip}(\sigma)^{2}}{2(d-2) \log \rho^{-1}} \int_{\tau_{2}}^{t} \frac{\||u_{s} - \tilde{u}_{s}||^{2}}{t-s+\rho} \, \mathrm{d}s \\ &\stackrel{(4.1)}{\leq} C \langle \||u_{\tau_{1}}\|| \rangle^{2} \frac{\log \frac{t-\tau_{1}+\rho}{t-\tau_{2}+\rho}}{\log \rho^{-1}} + \frac{\mathrm{Lip}(\sigma)^{2}}{2(d-2) \log \rho^{-1}} \int_{\tau_{2}}^{t} \frac{\||u_{s} - \tilde{u}_{s}||^{2}}{t-s+\rho} \, \mathrm{d}s. \end{split}$$

Then we conclude by [12, Lemma 4.3] (with the appropriate change of constants).

In the next proposition, we show that after the noise has been turned off for some period, the solution is quite flat (due to the smoothing influence of the heat kernel), so we do not lose much if we replace the solution by a constant (its value at a fixed point  $X \in \mathbb{R}^d$ ) at that time. This of course is only true if we evaluate the solution at a point x that is close to X compared to the amount of time elapsed. To state this concisely, for  $X \in \mathbb{R}^d$ , we let  $\mathcal{Z}_X$  be the "freezing" operator that, given a function u defined on  $\mathbb{R}^d$ , returns the constant function on  $\mathbb{R}^d$  with value u(X):

$$(\mathcal{Z}_X u)(x) \equiv u(X). \tag{5.1}$$

In the following statement, we compare two solutions over [0, t] with the noise turned off on  $[\tau_1, \tau_2] \subset [0, t]$ , but one of them additionally has the solution at time  $\tau_2$  replaced by its value at some fixed location X before the evolution of the SPDE continues.

Recall the notation  $\mathcal{X}_r$  from Section 1.4.

**Proposition 5.2.** For every T > 0 and  $\lambda < \sqrt{2(d-2)}$ , there is a  $C = C(T,\lambda) < \infty$  such that the following holds. Suppose that  $\operatorname{Lip}(\sigma) < \lambda$ . Let  $\tau_1 < \tau_2$ ,  $u_{\tau_1} \in \mathcal{X}_{\tau_1}$ , and  $X \in \mathbb{R}^d$ . Then we have, for each  $x \in \mathbb{R}^d$ , that

$$\mathbb{E}|\mathcal{U}_{\tau_2,t}(\mathcal{Z}_X\mathcal{G}_{\tau_2-\tau_1}u_{\tau_1})(x)-\mathcal{U}_{\tau_2,t}(\mathcal{G}_{\tau_2-\tau_1}u_{\tau_1})(x)|^2\leq C(\|u_{\tau_1}\|)^2\frac{t-\tau_2+|x-X|^2}{\tau_2-\tau_1}.$$

*Proof.* Let  $\tilde{u}_t = \mathcal{U}_{\tau_2,t}\mathcal{G}_{\tau_2-\tau_1}u_{\tau_1}$  and  $\overline{u}_t = \mathcal{U}_{\tau_2,t}\mathcal{Z}_X\mathcal{G}_{\tau_2-\tau_1}u_{\tau_1}$ . We can write

$$\begin{split} \tilde{u}_t(x) - \overline{u}_t(x) &= \mathcal{G}_{t-\tau_1} u_{\tau_1}(x) - \mathcal{G}_{\tau_2 - \tau_1} u_{\tau_1}(X) \\ &+ \frac{1}{\sqrt{\log \rho^{-1}}} \int_{\tau_2}^t \int \mathcal{G}_{t-s} [(\sigma(\tilde{u}_s) - \sigma(\overline{u}_s)) \, \mathrm{d}W_s^{\rho}](x). \end{split}$$

This means that, if we define  $f(t, x) = \mathbb{E}|\tilde{u}_t(x) - \overline{u}_t(x)|^2$ , then we have

$$f(t,x) \leq \mathbb{E} |\mathcal{G}_{t-\tau_1} u_{\tau_1}(x) - \mathcal{G}_{\tau_2-\tau_1} u_{\tau_1}(X)|^2$$

$$+ \frac{\operatorname{Lip}(\sigma)^2}{\log \rho^{-1}} \int_{\tau_2}^t \iint R^{\rho} (y_1 - y_2) \Biggl( \prod_{i=1}^2 G_{t-s}(x - y_i) f(s, y_i)^{1/2} \Biggr) dy_1 dy_2 ds.$$

We can estimate

$$\mathbb{E} \Big| \mathcal{G}_{t-\tau_{1}} u_{\tau_{1}}(x) - \mathcal{G}_{\tau_{2}-\tau_{1}} u_{\tau_{1}}(X) \Big|^{2} = \mathbb{E} \Big| \int \Big( G_{t-\tau_{1}}(x-y) - G_{\tau_{2}-\tau_{1}}(X-y) \Big) u_{\tau_{1}}(y) \, \mathrm{d}y \Big|^{2}$$

$$\leq \| u_{\tau_{1}} \| \cdot \| G_{t-\tau_{1}}(x-\cdot) - G_{\tau_{2}-\tau_{1}}(X-\cdot) \|_{L^{1}(\mathbb{R})}^{2}$$

$$\leq \| u_{\tau_{1}} \| \cdot \frac{2(t-\tau_{2}) + |x-X|^{2}}{\tau_{2}-\tau_{1}},$$

with the last inequality by [12, (5.6-5.7)]. We then conclude by Lemma 5.3 below.  $\Box$ 

**Lemma 5.3.** For every  $\omega < 1$ , we have a constant  $C = C(\omega, d) < \infty$  such that, if f is a positive bounded function and  $A_1, A_2 \in (0, \infty)$  are constants such that

$$f(t,x) \le A_1 + A_2|x - X|^2 + \frac{\lambda^2}{\log \rho^{-1}} \int_{\tau_2}^t \iint R^{\rho} (y_1 - y_2) \prod_{i=1}^2 G_{t-s}(x - y_i) f(s, y_i)^{1/2} \, \mathrm{d}y_1 \, \mathrm{d}y_2 \, \mathrm{d}s$$
 (5.2)

and

$$\frac{\lambda^2 \log \frac{t - \tau_2 + \rho}{\rho}}{2(d - 2) \log \rho^{-1}} \le \omega,\tag{5.3}$$

then, for all  $t \ge \tau_2$ , we have

$$f(t,x) \le C(A_1 + A_2(t - \tau_2 + |x - X|^2)). \tag{5.4}$$

*Proof.* Since f is assumed to be bounded, we know that, for some values of  $B_1$  and  $B_2$ ,

$$f(t,x) \le B_1 + B_2|x - X|^2. \tag{5.5}$$

In view of (5.2), the inequality (5.5) implies that

$$f(t,x) \le A_1 + A_2|x - X|^2$$

$$+ \frac{\lambda^2}{\log \rho^{-1}} \int_{\tau_2}^t \iint R^{\rho} (y_1 - y_2) \prod_{i=1}^2 G_{t-s}(x - y_i) (B_1 + B_2|y_i - X|^2)^{1/2} dy_1 dy_2 ds.$$

We make the change of variables

$$w = \frac{1}{2}(y_1 + y_2), \qquad z = y_1 - y_2.$$
 (5.7)

We note that

$$\begin{split} \prod_{i=1}^{2} \left( B_1 + B_2 |y_i - X|^2 \right) &= \left( B_1 + B_2 |w + z/2 - X|^2 \right) \left( B_1 + B_2 |w - z/2 - X|^2 \right) \\ &= B_1^2 + B_1 B_2 \left( |w - X|^2 + \frac{1}{4} |z|^2 \right) \\ &+ B_2^2 \left( \left( |w - X|^2 + \frac{1}{4} |z|^2 \right)^2 - \frac{1}{2} ((w - X) \cdot z)^2 \right) \\ &\leq \left( B_1 + B_2 \left( |w - X|^2 + \frac{1}{4} |z|^2 \right) \right)^2, \end{split}$$

so

$$\prod_{i=1}^{2} \left( B_1 + B_2 |y_i - X|^2 \right)^{1/2} \le B_1 + B_2 \left( |w - X|^2 + \frac{1}{4} |z|^2 \right).$$
(5.8)

Using the change of variables (5.7) and the estimate (5.8) in (5.6), we obtain

$$f(t,x) \leq A_{1} + A_{2}|x - X|^{2}$$

$$+ \frac{\lambda^{2}}{\log \rho^{-1}} \int_{\tau_{2}}^{t} \iint R^{\rho}(z) \left( B_{1} + B_{2} \left( |w - X|^{2} + \frac{1}{4}|z|^{2} \right) \right)$$

$$\times G_{t-s}(x - w - z/2) G_{t-s}(x - w + z/2) dw dz ds$$

$$= A_{1} + A_{2}|x - X|^{2}$$

$$+ \frac{\lambda^{2}}{\log \rho^{-1}} \int_{\tau_{2}}^{t} \iint R^{\rho}(z) \left( B_{1} + B_{2} \left( |w - X|^{2} + \frac{1}{4}|z|^{2} \right) \right)$$

$$\times G_{2(t-s)}(z) G_{\frac{t-s}{2}}(x - w) dw dz ds.$$

$$(5.9)$$

Now we have

$$\frac{\lambda^{2}}{\log \rho^{-1}} \int_{\tau_{2}}^{t} \iint \left( B_{1} + B_{2} |w - X|^{2} \right) R^{\rho}(z) G_{2(t-s)}(z) G_{\frac{t-s}{2}}(x - w) \, dw \, dz \, ds$$

$$\stackrel{(3.3)}{=} \frac{\lambda^{2}}{2(d-2) \log \rho^{-1}} \int_{\tau_{2}}^{t} \frac{B_{1} + B_{2} \int G_{\frac{t-s}{2}}(x - w) |w - X|^{2} \, dw}{t - s + \rho} \, ds$$

$$= \frac{\lambda^{2}}{2(d-2) \log \rho^{-1}} \int_{\tau_{2}}^{t} \frac{B_{1} + B_{2} \left( d(t-s)/2 + |x - X|^{2} \right)}{t - s + \rho} \, ds$$

$$\stackrel{(5.3)}{\leq} \omega \left( B_{1} + B_{2} \left( d(t - \tau_{2})/2 + |x - X|^{2} \right) \right). \tag{5.10}$$

We also have that

$$\frac{\lambda^2}{4\log \rho^{-1}} \int_{\tau_2}^t \iint R^{\rho}(z)|z|^2 G_{2(t-s)}(z) G_{\frac{t-s}{2}}(x-w) \, \mathrm{d}w \, \mathrm{d}z \, \mathrm{d}s \le \frac{C\lambda^2(t-\tau_2)}{\log \rho^{-1}} \qquad (5.11)$$

by (3.6). Using (5.10) and (5.11) in (5.9), we get

$$f(t,x) \le A_1 + (B_1 + B_2 d(t - \tau_2)/2)\omega + CB_2 \frac{\lambda^2 (t - \tau_2)}{\log \rho^{-1}} + (A_2 + \omega B_2)|x - X|^2.$$

Therefore, if we take  $B_2^{(0)} = 0$  and  $B_1^{(0)}$  so large that (5.5) holds with  $B_i \leftarrow B_i^{(0)}$ , and then inductively define

$$B_1^{(n)} = A_1 + \omega \left( B_1^{(n-1)} + B_2^{(n-1)} d(t - \tau_2) / 2 \right) + \frac{C\lambda^2 (t - \tau_2)}{\log \rho^{-1}} B_2^{(n-1)}; \tag{5.12}$$

$$B_2^{(n)} = A_2 + \omega B_2^{(n-1)},\tag{5.13}$$

then we have

$$f(t,x) \le B_1^{(n)} + B_2^{(n)} |x - X|^2$$
 for each  $n \in \mathbb{N}$ .

Now from (5.13) and the assumption that  $\omega$  < 1 we see that

$$\limsup_{n\to\infty} B_2^{(n)} \le \frac{A_2}{1-\omega},$$

and using this in (5.12), we see that

$$\limsup_{n \to \infty} B_1^{(n)} \le \frac{A_1 + \frac{A_2}{1 - \omega} (d/2 + C\lambda^2 / \log \rho^{-1})(t - \tau_2)}{1 - \omega}.$$

The last three displays together imply (5.4).

We summarize the results of this section in the following combined proposition.

**Proposition 5.4.** For every T > 0 and  $\lambda < \sqrt{2(d-2)}$ , there is a  $C = C(T,\lambda) < \infty$  such that the following holds. Suppose that  $\text{Lip}(\sigma) < \lambda$ . Let  $\tau_1 < \tau_2 \in \mathbb{R}$  be such that  $\tau_2 \leq \tau_1 + T$ , and suppose that  $\rho \in (0, C^{-1})$  and  $u_{\tau_1} \in \mathcal{X}_{\tau_1}$ . Then we have

$$\left(\mathbb{E}\left|\mathcal{U}_{\tau_{2},t}\mathcal{Z}_{X}\mathcal{G}_{\tau_{2}-\tau_{1}}u_{\tau_{1}}(x)-\mathcal{U}_{\tau_{1},t}u_{\tau_{1}}(x)\right|^{2}\right)^{1/2} \\
\leq C\langle ||u_{\tau_{1}}|||\rangle \left[\left(\frac{\log\left(\frac{t-\tau_{1}+\rho}{t-\tau_{2}+\rho}\right)+1}{\log\rho^{-1}}\right)^{1/2}+\left(\frac{t-\tau_{2}+|x-X|^{2}}{\tau_{2}-\tau_{1}}\right)^{1/2}\right].$$
(5.14)

*Proof.* This follows from combining Propositions 5.1 and 5.2 using the triangle inequality.  $\Box$ 

#### 6 Decorrelation estimate

Since, unlike [12], we consider noise with long range correlations, the statistics of the solution involve covariances of  $\sigma(u)$  at distant locations, rather than at locations that are very close to one another. The following decorrelation result will be essential for bounding such covariances.

**Proposition 6.1.** For each T > 0, there is a  $C = C(T, \sigma, d) \in (1, \infty)$  such that the following holds. Let  $f \in \text{Lip}(\mathbb{R}^m)$  and let  $u_t$  solve (1.2) with initial data  $u_0 \in \mathcal{X}_0$ . Then we have, for all  $\rho \in (0, C^{-1})$ , that

$$\operatorname{Cov}(f(u_t(x_1)), f(u_t(x_2))) \le \frac{C \operatorname{Lip}(f)^2 \langle ||u_0|||)^2}{\log \rho^{-1}} \left( \log \frac{t^{1/2}}{|x_1 - x_2|} + 1 + \frac{|x_1 - x_2|}{t^{1/2}} \right). \tag{6.1}$$

*Proof.* We assume that  $f: \mathbb{R}^m \to \mathbb{R}$  is smooth; the case of general Lipschitz f follows by approximation. Using the Clark–Ocone formula, we can write

$$f(u_t(x)) - \mathbb{E}[f(u_t(x))] = \int_0^t \int \mathbb{E}[D_{r,z}[f(u_t(x))] \mid \mathcal{F}_s] dW_s^{\rho}(z)$$

$$= \int_0^t \int \mathbb{E}[\nabla f(u_t(x)) \cdot D_{r,z}u_t(x) \mid \mathcal{F}_s] dW_s^{\rho}(z).$$
(6.2)

We then estimate the covariance by

$$\begin{aligned} &\operatorname{Cov}(f(u_{t}(x_{1})), f(u_{t}(x_{2}))) \\ &\leq \int_{0}^{t} \iint R^{\rho}(z_{1} - z_{2}) \mathbb{E} \left[ \prod_{i=1}^{2} \left| \mathbb{E} \left[ \nabla f(u_{t}(x_{i})) \cdot \operatorname{D}_{r, z_{i}} u_{t}(x_{i}) \mid \mathcal{F}_{r} \right] \right| \right] dz_{1} dz_{2} dr \\ &\leq \operatorname{Lip}(f)^{2} \int_{0}^{t} \iint R^{\rho}(z_{1} - z_{2}) \prod_{i=1}^{2} \left( \mathbb{E} \left| \operatorname{D}_{r, z_{i}} u_{t}(x) \right|_{F}^{2} \right)^{1/2} dz_{1} dz_{2} dr. \end{aligned}$$

Applying (6.3) below, we obtain

$$Cov(f(u_{t}(x_{1})), f(u_{t}(x_{2})))$$

$$\leq \frac{C \operatorname{Lip}(f)^{2} \langle ||u_{0}||| \rangle^{2}}{\log \rho^{-1}} \int_{0}^{t} \iint R^{\rho}(z_{1} - z_{2}) \prod_{i=1}^{2} G_{t-r}(x_{i} - z_{i}) dz_{1} dz_{2} dr,$$

so that (6.1) then follows from (3.7).

In the proof of Proposition 6.1, we needed a bound on the Malliavin derivative norm  $\mathbb{E}|D_{r,z_i}u_t(x)|_{\mathbb{F}}^2$ :

**Proposition 6.2.** For each  $T \in (0, \infty)$ ,  $M, \beta \in (0, \infty)$ , and  $\omega < 1$ , there is a constant  $C = C(T, \omega, M, \beta)$  such that the following holds. Suppose that  $|\sigma(u)|_F^2 \leq M + \beta^2 |u|^2$  for all  $u \in \mathbb{R}^m$  and that  $\frac{\operatorname{Lip}(\sigma)^2}{d-2} \log_{\rho^{-1}} (1 + \frac{3}{2}\rho^{-1}T) \leq \omega$ . Let  $u_0 \in \mathcal{X}_0$  and let  $u_t = \mathcal{U}_{0,t}u_0$ . Then we have

$$\mathbb{E}|D_{r,z}u_t(x)|_F^2 \le \frac{C(\|\|u_0\|\|)^2 G_{t-r}^2(x-z)}{\log \rho^{-1}}.$$
(6.3)

*Proof.* Define, for  $t \ge 0$ ,  $u_t^0(x) = u_0(x)$  and, for  $n \ge 1$ ,

$$u_t^n(x) = \mathcal{G}_t u_0(x) + \frac{1}{\sqrt{\log \rho^{-1}}} \int_0^t \mathcal{G}_{t-s} [\sigma(u_s^{n-1}) \, dW_s^{\rho}](x). \tag{6.4}$$

It is standard that  $(u_t^n)$  converges to  $(u_t)$  (for fixed  $\rho > 0$ , in probability, uniformly on compact sets) as  $n \to \infty$ . For 0 < r < t, and  $z, x \in \mathbb{R}^2$ , we can take Malliavin derivatives in (6.4) to obtain

$$\begin{split} \mathbf{D}_{r,z}u_t^n(x) &= \frac{1}{\sqrt{\log \rho^{-1}}}G_{t-r}(x-z)\sigma(u_s^{n-1}(z)) \\ &+ \frac{1}{\sqrt{\log \rho^{-1}}}\int_r^t \int G_{t-s}(x-z)\underline{\nabla}\sigma(u_s^{n-1}(z))\cdot \mathbf{D}_{r,z}u_s^{n-1}(z)\,\mathrm{d}W_s^\rho(z), \end{split}$$

where  $(\nabla \sigma(u_s^{n-1}(z)))$  is an adapted process satisfying

$$\sup_{s,z} \|\nabla \sigma(u_s^{n-1}(z))\|_{\mathbb{F}} \le \text{Lip}(\sigma) \quad \text{almost surely.}$$

In the case when  $\sigma$  is in fact continuously differentiable, then we can indeed take  $\nabla \sigma(u_s^{n-1}(z)) = \nabla \sigma(u_s^{n-1}(z))$ . If this is not the case, then the statement nonetheless holds, as can be seen be a simple adaptation of the proof of [23, Proposition 1.2.4]. That proposition concerns scalar-valued functions and finite-dimensional noise, but the same argument holds for vector-valued functions, and the infinite-dimensional noise we consider can be handled by approximation.

This having been established, we can continue by writing for all  $n \ge 1$  (using also the assumption on  $\sigma$ ),

$$\mathbb{E}|D_{r,z}u_{t}^{n}(x)|_{F}^{2} \leq \frac{M + \beta^{2} \sup_{s \in [0,T]} |||u_{s}^{n-1}|||^{2}}{\log \rho^{-1}} G_{t-r}^{2}(x-z) + \frac{\operatorname{Lip}(\sigma)^{2}}{\log \rho^{-1}} \int_{r}^{t} \iint R^{\rho}(y_{1}-y_{2}) \prod_{i=1}^{2} \left(G_{t-s}(x-y_{i}) \left(\mathbb{E}|D_{r,z}u_{s}^{n-1}(y_{i})|_{F}^{2}\right)^{1/2}\right) dy_{1} dy_{2} ds.$$
(6.5)

Moreover, we have

$$\mathbb{E}|D_{r,z}u_{t}^{0}(x)|_{F}^{2}=0$$

since  $u_t^0(x) = u_0(x)$  is deterministic.

Now we can apply Proposition 6.3 below with

$$\lambda_1 = M + \beta^2 \sup_{s \in [0,T]} |||u_s^{n-1}|||^2$$
 and  $\lambda_2 = \text{Lip}(\sigma)^2$ ,

and time shifted by r, and note that  $\lambda_1 \leq C(\|u_0\|)^2$  by an easy modification of Proposition 4.1 (to apply to the  $u_n$ s rather than to u), to obtain

$$\mathbb{E}|\mathrm{D}_{r,z}u^n_t(x)|_{\mathrm{F}}^2 \leq \frac{C(\|\|u_0\|\|)^2 G_{t-r}^2(x-z)}{\log \rho^{-1}}.$$

Passing to the limit as  $n \to \infty$ , we obtain (6.3).

The following proposition is an analogue of [13, Lemma 2.7] that allows us to analyze the recurrence (6.5).

**Proposition 6.3.** For each  $\omega < 1$  and  $\lambda_1, \lambda_2, T \in (0, \infty)$ , there is a constant  $C = C(\omega, \lambda_1, \lambda_2, T)$  such that the following holds. Suppose that  $g_n : [0, T] \times \mathbb{R}^d \to \mathbb{R}$  are such that

$$g_0(t,x) = 0$$
 for all  $t \in [0,T]$  and  $x \in \mathbb{R}^d$ 

and we have  $\lambda_1, \lambda_2, T \geq 0$  such that

$$g_{n+1}^2(t,x) \leq \frac{1}{\log \rho^{-1}} \left( \lambda_1 G_t^2(x) + \lambda_2 \int_0^t J_t(s,x) [g_n] \, \mathrm{d}s \right) \quad \text{for all } x \in \mathbb{R}^d \text{ and } t \in [0,T],$$

where we have defined

$$J_t^{\rho}(s,x)[g] = \iint R^{\rho}(y_1 - y_2) \prod_{i=1}^{2} (G_{t-s}(x - y_i)g(s,y_i)) \, \mathrm{d}y_1 \, \mathrm{d}y_2.$$

Then, as long as

$$\frac{\lambda_2}{d-2} \cdot \frac{\log(1 + \frac{3}{2}T\rho^{-1})}{\log \rho^{-1}} \le \omega, \tag{6.6}$$

we have

$$g_n(t,x) \le \frac{CG_t(x)}{\log \rho^{-1}}$$
 for all  $x \in \mathbb{R}^d$  and  $t \in [0,T]$ .

Proof. We define

$$f_0(t,x) = 0$$

and

$$f_n^2(t,x) = \frac{1}{\log \rho^{-1}} \left( \lambda_1 G_t^2(x) + \lambda_2 \int_0^t J_t(s,x) [f_{n-1}] \, \mathrm{d}s \right), \tag{6.7}$$

so that  $g_n \leq f_n$  for all n.

We then look for an inductive choice of  $H_n$  with  $H_0 = 0$  and such that

$$f_n^2(t,x) \le \frac{H_n}{\log \rho^{-1}} G_t^2(x),$$
 (6.8)

We have

We use the fact that

$$G_{t-s}(x-y)G_s(y) = G_t(x)G_{\frac{s}{t}(t-s)}\left(y - \frac{s}{t}x\right)$$

to turn this into

$$\int_{0}^{t} J_{t}(s,x)[f_{n}] ds$$

$$\leq \frac{(2\pi)^{d/2} H_{n}}{2 \log \rho^{-1}} G_{t}(x)^{2} \int_{0}^{t} \iint \left( \prod_{i=1}^{2} G_{\frac{s}{t}(t-s)} \left( y_{i} - \frac{s}{t} x \right) \right) \int_{0}^{\infty} v^{d/2 - 2} G_{v+2\rho}(y_{1} - y_{2}) dv dy_{1} dy_{2} ds$$

$$= \frac{(2\pi)^{d/2} H_{n}}{2 \log \rho^{-1}} G_{t}(x)^{2} \int_{0}^{t} \int_{0}^{\infty} v^{d/2 - 2} G_{2\frac{s}{t}(t-s) + v + 2\rho}(0) dv ds$$

$$= \frac{H_{n}}{2 \log \rho^{-1}} G_{t}^{2}(x) \int_{0}^{t} \int_{0}^{\infty} \frac{v^{d/2 - 2}}{(2\frac{s}{t}(t-s) + v + 2\rho)^{d/2}} dv ds$$

$$= \frac{H_{n} G_{t}^{2}(x)}{(d-2) \log \rho^{-1}} \int_{0}^{t} \frac{ds}{2\frac{s}{t}(t-s) + 2\rho}.$$

We estimate the last integral as

$$\int_0^t \frac{\mathrm{d}s}{2\frac{s}{t}(t-s)+2\rho} = \frac{t}{\sqrt{t(t+4\rho)}}\log\frac{t+\sqrt{t(4\rho+t)}+2\rho}{2\rho} \leq \log\frac{3t+2\rho}{2\rho}.$$

Using the last two displays in (6.7), we obtain the bound

$$\begin{split} f_n^2(t,x) &\leq \frac{1}{\log \rho^{-1}} \Biggl( \lambda_1 + \lambda_2 \frac{H_n}{d-2} \frac{\log(1 + \frac{3}{2}t\rho^{-1})}{\log \rho^{-1}} \Biggr) G_t^2(x) \\ &\leq \frac{1}{\log \rho^{-1}} (\lambda_1 + \omega H_n) G_t^2(x) \; . \end{split}$$

It follows that the choice  $H_{n+1} = \lambda_1 + \omega H_n$  allows us to satisfy the bound (6.8) for all n. Since  $H_0 = 0$  and  $\omega < 1$ , we find that  $H_n < \frac{\lambda_1}{1-\omega}$  for all n, completing the proof.  $\square$ 

# 7 The approximate decoupling function

#### 7.1 Definition and regularity

We define the matrix-valued function

$$L_{\rho}(q, a) = \mathbb{E}[\sigma(\mathcal{U}_{0, \rho^{1-q}} a(x))] \qquad \text{for } a \in \mathbb{R}^{m} \text{ and } q \ge 0.$$
 (7.1)

The quantity  $J_{\rho} := L_{\rho}/\sqrt{2(d-2)}$  (see (8.30) below) will be an approximation of the decoupling function J appearing in (1.6c). Establishing this will be the goal of the next two sections. We begin with some temporal and spatial regularity properties for  $L_{\rho}$ . These will in particular give us a compactness statement for  $L_{\rho}$  in Proposition 7.4 below.

**Lemma 7.1.** There is a constant  $C = C(\sigma) \in (1, \infty)$  such that we have, for all  $q \in [0, 1]$  and all  $\rho \in (0, C^{-1})$ , that

$$L_{\rho}(q, a) \le C\langle a \rangle.$$
 (7.2)

*Proof.* This is a consequence of Proposition 4.1 and the fact that  $\sigma$  is Lipschitz.

**Lemma 7.2.** There is a constant  $C \in (1, \infty)$  such that we have, for all  $q_1, q_2 \in [0, 1]$  and all  $\rho \in (0, C^{-1})$ , that

$$|L_{\rho}(q_1, a) - L_{\rho}(q_2, a)|_{\mathbb{F}} \le C \operatorname{Lip}(\sigma) \langle a \rangle \left( |q_2 - q_1|^{1/2} + \frac{1}{\sqrt{\log \rho^{-1}}} \right).$$
 (7.3)

*Proof.* Assume without loss of generality that  $q_2 \ge q_1$ . Then we have

$$\begin{split} |L_{\rho}(q_{2}, a) - L_{\rho}(q_{1}, a)|_{F} &\leq \mathbb{E}|\sigma(\mathcal{U}_{0, \rho^{1-q_{2}}}a(x)) - \sigma(\mathcal{U}_{\rho^{1-q_{2}}-\rho^{1-q_{1}}, \rho^{1-q_{2}}}a(x))|_{F} \\ &\leq \operatorname{Lip}(\sigma)\mathbb{E}|\mathcal{U}_{0, \rho^{1-q_{2}}}a(x) - \mathcal{U}_{\rho^{1-q_{2}}-\rho^{1-q_{1}}, \rho^{1-q_{2}}}a(x)| \\ &\leq \operatorname{Lip}(\sigma) \Big(\mathbb{E}|\mathcal{U}_{0, \rho^{1-q_{2}}}a(x) - \mathcal{U}_{\rho^{1-q_{2}}-\rho^{1-q_{1}}, \rho^{1-q_{2}}}\mathcal{G}_{\rho^{1-q_{1}}}a(x)|^{2}\Big)^{1/2} \\ &\leq C\operatorname{Lip}(\sigma) \langle a \rangle \Bigg(\frac{\log\left(\frac{\rho^{1-q_{2}}+\rho}{\rho^{1-q_{1}}+\rho}\right) + 1}{\log \rho^{-1}}\Bigg)^{1/2} \\ &\leq C\operatorname{Lip}(\sigma) \langle a \rangle \Bigg(|q_{2}-q_{1}|^{1/2} + \frac{1}{\sqrt{\log \rho^{-1}}}\Bigg), \end{split}$$

with the penultimate inequality following from Propositions 4.1 and 5.1.

**Lemma 7.3.** There exists a constant  $C < \infty$  such that, for all  $q \in [0, 1]$  and all  $a_1, a_2 \in \mathbb{R}^m$ ,

$$|L_{\rho}(q, a_1) - L_{\rho}(q, a_2)|_{\mathcal{F}} \le C \operatorname{Lip}(\sigma)|a_1 - a_2|.$$
 (7.4)

Proof. We have

$$|L_{\rho}(q, a_{1}) - L_{\rho}(q, a_{2})|_{F} \leq \mathbb{E}|\sigma(\mathcal{U}_{0, \rho^{1-q}} a_{1}(x)) - \sigma(\mathcal{U}_{0, \rho^{1-q}} a_{2}(x))|_{F}$$

$$\leq \operatorname{Lip}(\sigma)\mathbb{E}|\mathcal{U}_{0, \rho^{1-q}} a_{1}(x) - \mathcal{U}_{0, \rho^{1-q}} a_{2}(x)|. \tag{7.5}$$

Now we perform an  $L^2$  analysis along the lines of [12, Proposition 3.3]. We have, defining  $v_t^{(i)}(x) = \mathcal{U}_{0,\rho^{1-q}}a_i(x)$ , that

$$\begin{split} \mathbb{E}|v_{t}^{(1)}(x) - v_{t}^{(2)}(x)|_{\mathrm{F}}^{2} \\ &\leq |a_{1} - a_{2}|^{2} + \frac{\mathrm{Lip}(\sigma)^{2}}{\log \rho^{-1}} \int_{0}^{\rho^{1-q}} \|v_{s}^{(1)} - v_{s}^{(2)}\|^{2} \iint R^{\rho}(y_{1} - y_{2}) \prod_{i=1}^{2} G(x - y_{1}) \, \mathrm{d}y_{1} \, \mathrm{d}y_{2} \, \mathrm{d}s \\ &\stackrel{(3.3)}{=} |a_{1} - a_{2}|^{2} + \frac{\mathrm{Lip}(\sigma)^{2}}{2(d - 2) \log \rho^{-1}} \int_{0}^{\rho^{1-q}} \frac{\|v_{s}^{(1)} - v_{s}^{(2)}\|^{2}}{\rho^{1-q} - s + \rho} \, \mathrm{d}s. \end{split}$$

Then by [12, Lemma 3.4] (with constants changed), we have

$$|||v_t^{(1)} - v_t^{(2)}|||^2 \le C|a_1 - a_2|^2$$

and using this in (7.5) we obtain (7.4).

The regularity properties of  $L_{\rho}$  just established give us compactness:

**Proposition 7.4.** For any sequence  $\rho_k \downarrow 0$ , there is a subsequence  $\rho_{k_i} \downarrow 0$  and a continuous function  $L \colon [0,1] \times \mathbb{R}^m \to \mathbb{R}^m \otimes \mathbb{R}^n$  such that

$$\lim_{i\to\infty}L_{\rho_{k_i}}|_{[0,1]\times\mathbb{R}^{\mathfrak{m}}}=L,$$

uniformly on compact subsets of  $[0,1] \times \mathbb{R}^{m}$ .

*Proof.* The proof is by compactness as in the proof of [12, Proposition 7.4]. The only difference is that, since we do not assume that  $\sigma(0) = 0$ , we also need the boundedness result Lemma 7.1 to apply Arzelà–Ascoli. Since the proof is otherwise the same, we omit the details.

#### 7.2 Statistics of the averaged fields

In this section we relate the approximate decoupling function  $L_{\rho}$  to spatial averages of the solution run on appropriate time scales. For a random vector  $u \in \mathbb{R}^{\mathfrak{m}}$ , we use the notation  $\operatorname{Var}(u) := \mathbb{E}[u^{\otimes 2}] := \mathbb{E}[uu^{\top}]$ .

**Proposition 7.5.** For each  $T \in (0, \infty)$  and  $\lambda < \sqrt{d-2}$ , we have a constant  $C = C(\sigma, T, \lambda, d) < \infty$  such that, whenever  $\text{Lip}(\sigma) < \lambda$ , then for any  $0 \le \tau_1 \le \tau_2 \le T$ ,  $a \in \mathbb{R}^m$ , and  $x \in \mathbb{R}^d$ , we have

$$\langle a \rangle^{-2} \left| \operatorname{Var}(\mathcal{G}_{\tau_{2} - \tau_{1}} \mathcal{U}_{0, \tau_{1}} a(X)) - \frac{L_{\rho} (1 - \log_{\rho} \tau_{1}, a)^{\otimes 2} \log_{\rho^{-1}} \left( 1 + \frac{\tau_{1}}{2(\tau_{2} - \tau_{1} + \rho)} \right)}{2(d - 2)} \right|_{F}$$

$$\leq C \log_{\rho^{-1}} \frac{\tau_{2} + \rho}{\tau_{2} - \frac{1}{2}\tau_{1} + \rho} + C \left( \frac{1}{\sqrt{\log \rho^{-1}}} + \log_{\rho^{-1}} \frac{\tau_{1}}{\tau_{2} - \tau_{1}} + \frac{\sqrt{\tau_{2}/\tau_{1}}}{\log \rho^{-1}} \right) \log_{\rho^{-1}} \frac{\tau_{2} - \frac{1}{2}\tau_{1}}{\tau_{2} - \tau_{1}}. \tag{7.6}$$

If in addition we assume that  $\tau_2 \leq 2\tau_1$ , then in particular we have the inequality

$$\left| \operatorname{Var}(\mathcal{G}_{\tau_{2}-\tau_{1}}\mathcal{U}_{0,\tau_{1}}a(X)) - \frac{L_{\rho}(1-\log_{\rho}\tau_{1},a)^{\otimes 2}\log_{\rho^{-1}}\left(1+\frac{\tau_{1}}{2(\tau_{2}-\tau_{1}+\rho)}\right)}{2(d-2)} \right|_{F}$$

$$\leq C\langle a \rangle^{2} \left( \frac{1}{\log \rho^{-1}} + \left(\log_{\rho^{-1}}\left(1+\frac{\tau_{1}}{\tau_{2}-\tau_{1}}\right)\right)^{2} \right).$$
(7.7)

*Proof.* Let  $u_s = \mathcal{U}_{0,s}a$ . We have

$$\operatorname{Var}(\mathcal{G}_{\tau_{2}-\tau_{1}}u_{\tau_{1}}(X)) = \operatorname{Var}\left(\frac{1}{\sqrt{\log \rho^{-1}}} \int_{0}^{\tau_{1}} \int G_{\tau_{2}-s}(X-y)\sigma(u_{s}(y)) dW_{s}^{\rho}(y)\right)$$

$$= \int_{0}^{\tau_{1}} A(s) ds, \tag{7.8}$$

where we have defined

$$A(s) := \frac{1}{\log \rho^{-1}} \iint R^{\rho} (y_1 - y_2) \times \mathbb{E} \left[ \left( G_{\tau_2 - s} (X - y_1) \sigma(u_s(y_1)) \right) \left( G_{\tau_2 - s} (X - y_2) \sigma(u_s(y_2)) \right)^{\top} \right] dy_1 dy_2.$$

We now estimate the integral (7.8) in several steps.

*Step 1. The initial layer.* First we deal with the first half of the time interval  $[0, \tau_1]$ , whose contribution we will think of as an error. We have

$$\left| \frac{1}{\log \rho^{-1}} \int_{0}^{\frac{1}{2}\tau_{1}} A(s) \, ds \right|_{F} \stackrel{(4.1)}{\leq} \frac{C\langle a \rangle^{2}}{\log \rho^{-1}} \int_{0}^{\frac{1}{2}\tau_{1}} \iint R^{\rho} (y_{1} - y_{2}) \prod_{i=1}^{2} G_{\tau_{2} - s} (X - y_{i}) \, dy_{1} \, dy_{2} \, ds$$

$$\stackrel{(3.3)}{=} \frac{C\langle a \rangle^{2}}{\log \rho^{-1}} \int_{0}^{\frac{1}{2}\tau_{1}} \frac{ds}{\tau_{2} - s + \rho} = C\langle a \rangle^{2} \log_{\rho^{-1}} \frac{\tau_{2} + \rho}{\tau_{2} - \frac{1}{2}\tau_{1} + \rho}. \tag{7.9}$$

*Step 2. Long-range and short-range terms.* We define, for  $s \in [\frac{1}{2}\tau_1, \tau_1]$ ,

$$A_{\mathbb{E}}(s) := \frac{1}{\log \rho^{-1}} \iint R^{\rho}(y_1 - y_2) \left( G_{\tau_2 - s}(X - y_1) \mathbb{E}[\sigma(u_s(y_1))] \right)$$
(7.10)

$$\times \left(G_{\tau_2-s}(X-y_2)\mathbb{E}\left[\sigma(u_s(y_2))\right]^{\top}\right)\mathrm{d}y_1\,\mathrm{d}y_2$$

$$\stackrel{(3.3)}{=} \frac{\mathbb{E}[\sigma(u_s(X))]^{\otimes 2}}{2(d-2)(\tau_2 - s + \rho)\log \rho^{-1}} \stackrel{(7.1)}{=} \frac{L_\rho(1 - \log_\rho s, a)^{\otimes 2}}{2(d-2)(\tau_2 - s + \rho)\log \rho^{-1}}$$
(7.11)

and

$$A_C(s) := \frac{1}{\log \rho^{-1}} \iint R^{\rho}(y_1 - y_2) \operatorname{Cov}(\sigma(u_s(y_1)), \sigma(u_s(y_2))) \prod_{i=1}^{2} G_{\tau_2 - s}(X - y_i) \, \mathrm{d}y_1 \, \mathrm{d}y_2.$$

$$(7.12)$$

Here, we use the notation  $Cov(u, v) := \mathbb{E}[uv^{\top}] - \mathbb{E}[u]\mathbb{E}[v]^{\top}$ . From these definitions we observe that

$$A(s) = A_{\mathbb{E}}(s) + A_{C}(s).$$
 (7.13)

Step 3. Estimate on  $A_{\mathbb{E}}(s)$ . The main contribution comes from  $A_{\mathbb{E}}(s)$ . From (7.11), we estimate, for all  $s \in [\frac{1}{2}\tau_1, \tau_1]$ , that

$$\left| A_{\mathbb{E}}(s)(\tau_{2} - s + \rho) \log \rho^{-1} - \frac{L_{\rho}(1 - \log_{\rho} \tau_{1}, a)^{\otimes 2}}{2(d - 2)} \right|_{\mathbb{F}} \\
\leq C|L_{\rho}(1 - \log_{\rho} s, a)^{\otimes 2} - L_{\rho}(1 - \log_{\rho} \tau_{1}, a)^{\otimes 2}|_{\mathbb{F}} \\
\stackrel{(7.2)}{\leq} C\langle a\rangle|L_{\rho}(1 - \log_{\rho} s, a) - L_{\rho}(1 - \log_{\rho} \tau_{1}, a)|_{\mathbb{F}} \\
\stackrel{(7.3)}{\leq} C\langle a\rangle^{2} \left(\sqrt{\log_{\rho^{-1}} \frac{\tau_{1}}{s}} + \frac{1}{\sqrt{\log \rho^{-1}}}\right) \leq \frac{C\langle a\rangle^{2}}{\sqrt{\log \rho^{-1}}}, \tag{7.14}$$

with the last inequality by the assumption that  $s \ge \frac{1}{2}\tau_1$ . This means that

$$\left| \int_{\frac{1}{2}\tau_{1}}^{\tau_{1}} A_{\mathbb{E}}(s) \, ds - \frac{L_{\rho}(1 - \log_{\rho} \tau_{1}, a)^{\otimes 2}}{2(d - 2)} \log_{\rho^{-1}} \frac{\tau_{2} - \frac{1}{2}\tau_{1} + \rho}{\tau_{2} - \tau_{1} + \rho} \right|_{F}$$

$$\leq \frac{1}{\log \rho^{-1}} \int_{\frac{1}{2}\tau_{1}}^{\tau_{1}} \frac{\left| A_{\mathbb{E}}(s)(\tau_{2} - s + \rho) \log \rho^{-1} - \frac{L_{\rho}(1 - \log_{\rho} \tau_{1}, a)L_{\rho}(1 - \log_{\rho} \tau_{1}, a)^{\mathsf{T}}}{2(d - 2)} \right|_{F}}{\tau_{2} - s + \rho} \, ds$$

$$\stackrel{(7.14)}{\leq} \frac{C\langle a \rangle^{2}}{(\log \rho^{-1})^{3/2}} \int_{\frac{1}{2}\tau_{1}}^{\tau_{1}} \frac{ds}{\tau_{2} - s + \rho} \leq \frac{C\langle a \rangle^{2}}{(\log \rho^{-1})^{3/2}} \int_{\frac{1}{2}\tau_{1}}^{\tau_{1}} \frac{ds}{\tau_{2} - s}$$

$$= \frac{C\langle a \rangle^{2}}{\sqrt{\log \rho^{-1}}} \log_{\rho^{-1}} \frac{\tau_{2} - \frac{1}{2}\tau_{1}}{\tau_{2} - \tau_{1}}.$$

$$(7.15)$$

Step 4. Estimate on  $A_C(s)$ . The contribution of  $A_C(s)$  is also an error term. We can use

Proposition 6.1 to estimate, for  $s \in [\frac{1}{2}\tau_1, \tau_1]$ ,

$$|A_{C}(s)|_{F} \leq \frac{C \operatorname{Lip}(\sigma)^{2} \langle a \rangle^{2}}{(\log \rho^{-1})^{2}} \iint R^{\rho}(y_{1} - y_{2}) \left( \log \frac{s^{1/2}}{|y_{1} - y_{2}|} + 1 + \frac{|y_{1} - y_{2}|}{s^{1/2}} \right) \cdot \prod_{i=1}^{2} G_{\tau_{2} - s}(X - y_{i}) \, dy_{1} \, dy_{2}$$

$$\cdot \prod_{i=1}^{2} G_{\tau_{2} - s}(X - y_{i}) \, dy_{1} \, dy_{2}$$

$$\cdot \prod_{i=1}^{2} G_{\tau_{2} - s}(X - y_{i}) \, dy_{1} \, dy_{2}$$

$$\cdot \prod_{i=1}^{2} G_{\tau_{2} - s}(X - y_{i}) \, dy_{1} \, dy_{2}$$

$$\cdot \prod_{i=1}^{2} G_{\tau_{2} - s}(X - y_{i}) \, dy_{1} \, dy_{2}$$

$$\cdot \prod_{i=1}^{2} G_{\tau_{2} - s}(X - y_{i}) \, dy_{1} \, dy_{2}$$

$$\cdot \prod_{i=1}^{2} G_{\tau_{2} - s}(X - y_{i}) \, dy_{1} \, dy_{2}$$

$$\cdot \prod_{i=1}^{2} G_{\tau_{2} - s}(X - y_{i}) \, dy_{1} \, dy_{2}$$

$$\cdot \prod_{i=1}^{2} G_{\tau_{2} - s}(X - y_{i}) \, dy_{1} \, dy_{2}$$

$$\cdot \prod_{i=1}^{2} G_{\tau_{2} - s}(X - y_{i}) \, dy_{1} \, dy_{2}$$

$$\cdot \prod_{i=1}^{2} G_{\tau_{2} - s}(X - y_{i}) \, dy_{1} \, dy_{2}$$

$$\cdot \prod_{i=1}^{2} G_{\tau_{2} - s}(X - y_{i}) \, dy_{1} \, dy_{2}$$

$$\cdot \prod_{i=1}^{2} G_{\tau_{2} - s}(X - y_{i}) \, dy_{1} \, dy_{2}$$

$$\cdot \prod_{i=1}^{2} G_{\tau_{2} - s}(X - y_{i}) \, dy_{1} \, dy_{2}$$

$$\cdot \prod_{i=1}^{2} G_{\tau_{2} - s}(X - y_{i}) \, dy_{1} \, dy_{2}$$

$$\cdot \prod_{i=1}^{2} G_{\tau_{2} - s}(X - y_{i}) \, dy_{1} \, dy_{2}$$

$$\cdot \prod_{i=1}^{2} G_{\tau_{2} - s}(X - y_{i}) \, dy_{1} \, dy_{2}$$

$$\cdot \prod_{i=1}^{2} G_{\tau_{2} - s}(X - y_{i}) \, dy_{1} \, dy_{2}$$

$$\cdot \prod_{i=1}^{2} G_{\tau_{2} - s}(X - y_{i}) \, dy_{1} \, dy_{2}$$

$$\cdot \prod_{i=1}^{2} G_{\tau_{2} - s}(X - y_{i}) \, dy_{1} \, dy_{2}$$

$$\cdot \prod_{i=1}^{2} G_{\tau_{2} - s}(X - y_{i}) \, dy_{1} \, dy_{2}$$

$$\cdot \prod_{i=1}^{2} G_{\tau_{2} - s}(X - y_{i}) \, dy_{1} \, dy_{2}$$

$$\cdot \prod_{i=1}^{2} G_{\tau_{2} - s}(X - y_{i}) \, dy_{1} \, dy_{2}$$

$$\cdot \prod_{i=1}^{2} G_{\tau_{2} - s}(X - y_{i}) \, dy_{1} \, dy_{2}$$

$$\cdot \prod_{i=1}^{2} G_{\tau_{2} - s}(X - y_{i}) \, dy_{1} \, dy_{2}$$

$$\cdot \prod_{i=1}^{2} G_{\tau_{2} - s}(X - y_{i}) \, dy_{1} \, dy_{2}$$

$$\cdot \prod_{i=1}^{2} G_{\tau_{2} - s}(X - y_{i}) \, dy_{1} \, dy_{2}$$

$$\cdot \prod_{i=1}^{2} G_{\tau_{2} - s}(X - y_{i}) \, dy_{1} \, dy_{2}$$

$$\cdot \prod_{i=1}^{2} G_{\tau_{2} - s}(X - y_{i}) \, dy_{1} \, dy_{2}$$

$$\cdot \prod_{i=1}^{2} G_{\tau_{2} - s}(X - y_{i}) \, dy_{1} \, dy_{2}$$

$$\cdot \prod_{i=1}^{2} G_{\tau_{2} - s}(X - y_{i}) \, dy_{1} \, dy_{2}$$

$$\cdot \prod_{i=1}^{2} G_{\tau_{2} - s}(X$$

This means that (folding  $Lip(\sigma)$  now into C)

$$\int_{\frac{1}{2}\tau_1}^{\tau_1} A_C(s) \, \mathrm{d}s \le \frac{C\langle a \rangle}{\log \rho^{-1}} \left( 1 + \log \frac{\tau_1}{\tau_2 - \tau_1} + (\tau_2/\tau_1)^{1/2} \right) \log_{\rho^{-1}} \frac{\tau_2 - \frac{1}{2}\tau_1}{\tau_2 - \tau_1}. \tag{7.17}$$

Step 5. Combining the estimates. Combining (7.8), (7.9), (7.13), (7.15) and (7.17), and using the fact that

$$\log_{\rho^{-1}} \frac{\tau_2 - \frac{1}{2}\tau_1 + \rho}{\tau_2 - \tau_1 + \rho} = \log_{\rho^{-1}} \left( 1 + \frac{1}{2} \cdot \frac{\tau_1}{\tau_2 - \tau_1 + \rho} \right),$$

we obtain (7.6). The bound (7.7) is then a simple derivation from (7.6) under the additional assumption  $\tau_2 \le 2\tau_1$ .

We also need a higher moment bound.

**Proposition 7.6.** Let T > 0. There is an  $\ell_0 > 2$  and a  $C = C(\sigma, T) < \infty$  such that, for all  $0 \le \tau_1 \le \tau_2 \le T$  and all  $\ell \in [2, \ell_0]$ , we have

$$\mathbb{E}|\mathcal{G}_{\tau_{2}-\tau_{1}}\mathcal{U}_{0,\tau_{1}}a(X)-a|^{\ell} \leq C\langle a\rangle^{\ell} \left(\log_{\rho^{-1}}\frac{\tau_{2}+\rho}{\tau_{2}-\tau_{1}+\rho}\right)^{\ell/2}.$$
 (7.18)

*Proof.* The proof is quite similar to that of [12, Proposition 7.8]. As in the proof of Proposition 4.1, we note that

$$Z_t := \mathcal{G}_{\tau_2 - t} \mathcal{U}_{0,t} a(X) - a \tag{7.19}$$

is a martingale in t. By the BDG inequality (see e.g. [19, Proposition 4.4]), we have a constant  $C_{\ell} < \infty$  such that

$$\mathbb{E}|Z_t|^{\ell} \le C_{\ell} \mathbb{E}|[Z]_t|^{\ell/2}. \tag{7.20}$$

Let  $u_t = \mathcal{U}_{0,t}a$ . Just as in (4.2), we have

$$[Z]_t = \frac{1}{\log \rho^{-1}} \int_0^t \iint R^{\rho}(y_1 - y_2) \prod_{i=1}^2 (G_{\tau_2 - s}(X - y_i) \sigma(u_s(y_i))) dy_1 dy_2 ds.$$

Therefore, we have by Jensen's inequality that

$$|[Z]_{t}|^{\ell/2} \leq \left(\frac{1}{\log \rho^{-1}} \int_{0}^{t} \iint R^{\rho}(y_{1} - y_{2}) \prod_{i=1}^{2} \left(G_{\tau_{2} - s}(x - y_{i}) |\sigma(u_{s}(y_{i}))|\right) dy_{1} dy_{2} ds\right)^{\ell/2}$$

$$\leq \frac{\left(\int_{t'_{m-1}}^{t} \left(G_{2(\tau_{2} - s)} * R^{\rho}\right)(0) ds\right)^{\ell/2 - 1}}{(\log \rho^{-1})^{\ell/2}}$$

$$\cdot \int_{0}^{t} \iint R^{\rho}(y_{1} - y_{2}) \prod_{i=1}^{2} \left(G_{\tau_{2} - s}(x - y_{i}) |\sigma(u_{s}(y_{i}))|^{\ell/2}\right) dy_{1} dy_{2} ds. \tag{7.21}$$

We have

$$\int_0^t (G_{2(\tau_2 - s)} * R^{\rho})(0) \, \mathrm{d}s \stackrel{(3.3)}{=} \frac{1}{2(d - 2)} \int_0^t \frac{1}{\tau_2 - s + \rho} \, \mathrm{d}s = \frac{\log \frac{\tau_2 + \rho}{\tau_2 - t + \rho}}{2(d - 2)}, \tag{7.22}$$

and also

$$\mathbb{E}\left[\int_{0}^{t} \iint R^{\rho}(y_{1} - y_{2}) \prod_{i=1}^{2} \left(G_{\tau_{2} - s}(x - y_{i}) |\sigma(u_{s}(y_{i}))|^{\ell/2}\right) dy_{1} dy_{2} ds\right]$$

$$\leq C \int_{0}^{t} \langle |||u_{s}|||_{\ell} \rangle^{\ell} \iint R^{\rho}(y_{1} - y_{2}) \prod_{i=1}^{2} G_{\tau_{2} - s}(x - y_{i}) dy_{1} dy_{2} ds$$

$$\leq C \langle a \rangle^{\ell} \log \frac{\tau_{2} + \rho}{\tau_{2} - t + \rho}, \tag{7.23}$$

with the second inequality by Proposition 4.1 and (3.3). Using (7.22) and (7.23) in (7.21), we get

$$\mathbb{E}|[Z]_t|^{\ell/2} \le C\langle a\rangle^{\ell} \left(\frac{\log\frac{\tau_2+\rho}{\tau_2-t+\rho}}{\log\rho^{-1}}\right)^{\ell/2}.$$

Taking  $t = \tau_1$  and recalling (7.19) and (7.20), we obtain (7.18).

# Analysis of the Markov chain

# The time scales and definition of the Markov chain

Now we define a time discretization as in [12]. Throughout this section, we fix  $T \in$  $[0, T_0]$  and  $X \in \mathbb{R}^d$ . We define the small parameters  $\delta_\rho, \gamma_\rho, \eta_\rho$  satisfying the conditions

$$(\log \rho^{-1})^{-1} \ll \gamma_{\rho} \ll \delta_{\rho}^{2} \ll \eta_{\rho} \ll 1, \tag{8.1}$$

$$\delta_{\rho}^{-1} \rho^{\frac{1}{2} \gamma_{\rho}} \ll 1. \tag{8.2}$$

Here we write  $f(\rho) \ll g(\rho)$  to mean that  $\lim_{\rho \to 0} \frac{f(\rho)}{g(\rho)} = 0$ . This means in particular that

$$\rho^{\delta_{\rho}} \ll \rho^{\gamma_{\rho}} \ll 1. \tag{8.3}$$

Now we define

$$s_m \coloneqq \rho^{m\delta_\rho};$$
  $s'_m \coloneqq \rho^{m\delta_\rho + \gamma_\rho};$  (8.4)  
 $t_m \coloneqq T - s_m;$   $t'_m \coloneqq T - s'_m.$  (8.5)

$$t_m \coloneqq T - s_m; \qquad t'_m \coloneqq T - s'_m. \tag{8.5}$$

It follows from (8.1) that

$$t_{m+1} - t'_m \ll t'_m - t_m \ll t_m - t'_{m-1}$$
(8.6)

for each m. We also define

$$M_1(\rho, T) := \lceil \delta_{\rho}^{-1} \log_{\rho} T \rceil - 1, \tag{8.7}$$

$$M_2(\rho) \coloneqq \lfloor \delta_r \rho^{-1} \rfloor. \tag{8.8}$$

We now define a process that solves the SPDE (1.2), but with the noise "turned off" on each time interval  $[t_m, t'_m]$ . Specifically, we define

$$w_t^{(M_1(\rho,T))} \coloneqq \mathcal{G}_t u_0 \tag{8.9}$$

and, for  $m \ge M_1(\rho, T) + 1$ ,

$$Y_m := \mathcal{G}_{t'_m - t_m} w_{t_m}^{(m-1)}(X)$$
 (8.10)

and

$$w_t^{(m)} := \mathcal{U}_{t'_m, t} Y_m = \mathcal{U}_{t'_m, t} \mathcal{Z}_X \mathcal{G}_{t'_m - t_m} w_{t_m}^{(m-1)} \qquad \text{for } t \ge t_{m'}.$$
 (8.11)

In particular, this means that

$$Y_m = \mathcal{G}_{t'_m - t_m} \mathcal{U}_{t'_{m-1}, t_m} Y_{m-1}(X).$$
(8.12)

We need a uniform second moment bound, similar to [12, Lemma 6.3]:

**Proposition 8.1.** For any T > 0, we have a constant  $C = C(T) < \infty$ , independent of  $\rho$ , such that

$$\sup_{m \in [M_1(\rho, T), M_2(\rho)]} \mathbb{E}|Y_m|^2 \le C \langle ||u_0||| \rangle^2.$$
(8.13)

and

$$\sup_{\substack{m \in [M_1(\rho, T), M_2(\rho)] \\ t \in [t'_m, T]}} \mathbb{E} \| w_t^{(m)} \|^2 \le C \langle \| u_0 \| \rangle^2.$$
(8.14)

*Proof.* We have by (8.12) and Proposition 7.6 with  $\ell=2$  that there is a constant  $C=C(\sigma,T)<\infty$ , which we will allow to increase if necessary from line to line throughout this proof, such that

$$\mathbb{E}|Y_{m}|^{2} = \mathbb{E}|Y_{m-1}|^{2} + \mathbb{E}\left[\mathbb{E}|\mathcal{G}_{t'_{m}-t_{m}}\mathcal{U}_{t'_{m-1},t_{m}}Y_{m-1} - Y_{m-1}|^{2} \mid Y_{m-1}]|\right]$$

$$= \mathbb{E}|Y_{m-1}|^{2} + C\mathbb{E}\langle Y_{m-1}\rangle^{2}\log_{\rho^{-1}}\frac{t'_{m} - t'_{m-1} + \rho}{t'_{m} - t_{m} + \rho}$$

$$\leq \mathbb{E}|Y_{m-1}|^{2} + C\mathbb{E}\langle Y_{m-1}\rangle^{2}\log_{\rho^{-1}}\left(1 + \frac{\rho^{\gamma_{\rho}}(\rho^{-\delta_{\rho}} - 1)}{1 - \rho^{\gamma_{\rho}}}\right)$$

$$\leq \mathbb{E}|Y_{m-1}|^{2} + C\delta_{\rho}\mathbb{E}\langle Y_{m-1}\rangle^{2}.$$
(8.15)

This means that

$$\mathbb{E}|Y_m|^2 \le (1 + C\delta_\rho)\mathbb{E}|Y_{m-1}|^2 + C.$$

By induction, this means that for all  $m \in [M_1(\rho, T), M_2(\rho)]$ , we have

$$\mathbb{E}|Y_m|^2 \leq C \langle ||u_0||| \rangle^2 (1 + C\delta_\rho)^{m - M_1(\rho, T)} \leq C \langle ||u_0||| \rangle^2 (1 + C\delta_\rho)^{C\delta_\rho^{-1}} \leq C \langle ||u_0||| \rangle^2.$$

This completes the proof of (8.13), and (8.14) then follows using the definition (8.11) and an application of Proposition 4.1 with  $\ell=2$ .

Now we will show that  $w^{(m)}$  is a good approximation for u at appropriate spacetime points.

**Proposition 8.2.** We have a constant  $C = C(\rho, T) < \infty$  such that, for all  $x \in \mathbb{R}^2$ ,

$$\sup_{\substack{m \in [M_1(\rho,T),M_2(T)] \\ t \in [t_{m+1},T]}} \mathbb{E}|w_t^{(m)}(x) - u_t(x)|^2 \le C \langle ||u_0||| \rangle^2 \Big(|x - X|^2 \rho^{-m\delta_\rho} + o(1)\Big), \quad (8.16)$$

where o(1) denotes a quantity that goes to 0 as  $\rho \downarrow 0$ .

*Proof.* We have, whenever  $M_1(\rho, T) \le m \le M_2(\rho)$  and  $t \in [t_{m+1}, T]$ , that

$$\mathbb{E}|w_{t}^{(m)}(x) - w_{t}^{(m-1)}(x)|^{2} = \mathbb{E}\left|\mathcal{U}_{t'_{m},t}\mathcal{Z}_{X}\mathcal{G}_{t'_{m}-t_{m}}w_{t_{m}}^{(m-1)}(x) - \mathcal{U}_{t_{m},t}w_{t_{m}}^{(m-1)}(x)\right|^{2}$$

$$\stackrel{(5.14)}{\leq} C\langle \|w_{t_{m}}^{(m-1)}\|\rangle^{2}\left(\log_{\rho^{-1}}\frac{t'_{m}-t_{m}+\rho}{t_{m+1}-t'_{m}+\rho} + \frac{1}{\log\rho^{-1}} + \frac{t-t'_{m}+|x-X|^{2}}{t'_{m}-t_{m}}\right). \tag{8.17}$$

We have

$$\|w_{t_{m}}^{(m-1)}\| \le C\langle \|u_{0}\| \rangle$$
 (8.18)

by (8.14). We also have

$$\log_{\rho^{-1}} \frac{t'_m - t_m + \rho}{t_{m+1} - t'_m + \rho} \le \log_{\rho^{-1}} \left( 1 + \frac{1 - \rho^{\gamma_{\rho}}}{\rho^{\gamma_{\rho}} - \rho^{\delta_{\rho}}} \right) \stackrel{(8.3)}{\le} C\gamma_{\rho}, \tag{8.19}$$

as well as

$$\frac{t - t'_m}{t'_m - t_m} \le \frac{T - t'_m}{t'_m - t_m} \le \frac{\rho^{\gamma_\rho}}{1 - \rho^{\gamma_\rho}} \stackrel{(8.3)}{\le} C\rho^{\gamma_\rho}$$
(8.20)

(for sufficiently small  $\rho$ ). Using (8.18–20), as well as the fact that  $t_m' - t_m = \rho^{m\delta_\rho} (1 - \rho^{\gamma_\rho}) \ge C^{-1} \rho^{m\delta_\rho}$  in (8.17), we see that

$$\mathbb{E}|w_t^{(m)}(x) - w_t^{(m-1)}(x)|^2 \le C \langle ||u_0||| \rangle^2 \Big( \gamma_\rho + \rho^{\gamma_\rho} + (\log \rho^{-1})^{-1} + \rho^{-m\delta_\rho} |x - X|^2 \Big).$$

Inductively applying this inequality along with the triangle inequality, we obtain

$$\begin{split} \Big( \mathbb{E}|w_t^{(m)}(x) - u_t(x)|^2 \Big)^2 &\leq \sum_{i=M_1(\rho,T)}^m \Big( \mathbb{E}|w_t^{(i)}(x) - w_t^{(i-1)}(x)|^2 \Big)^{1/2} \\ &\leq C \langle \|u_0\| \rangle \bigg( \delta_\rho^{-1} \gamma_\rho^{1/2} + \delta_\rho^{-1} \rho^{\gamma_\rho/2} + \delta_\rho^{-1} (\log \rho^{-1})^{-1/2} + \frac{|x - X| \rho^{-m\delta_\rho/2}}{1 - \rho^{-\delta_\rho/2}} \bigg). \end{split}$$

The first three terms in brackets go to 0 as  $\rho \downarrow 0$  by (8.1–2), and the denominator  $1 - \rho^{-\delta_{\rho}/2}$  similarly goes to 1 as  $\rho \downarrow 0$ . Hence (8.16) is proved.

**Proposition 8.3.** We have a constant  $C = C(\sigma, T) < \infty$  such that

$$\mathbb{E}|Y_{M_2(\rho)} - u_T(X)|^2 \le C \langle ||u_0||| \rangle^2 o(1), \tag{8.21}$$

where o(1) denotes a quantity that goes to 0 as  $\rho \downarrow 0$ .

*Proof.* By Proposition 8.2, we see that

$$\left(\mathbb{E}|w_T^{(M_2(\rho))}(X) - u_T(X)|^2\right)^{1/2} \le C(\|u_0\|)o(1). \tag{8.22}$$

We also have by Proposition 7.6 (applied with  $\ell=2$  and  $\tau_1=\tau_2=T-t'_{M_2(\rho)}=\rho^{M_2(\rho)\delta_\rho+\gamma_\rho}$ ) that

$$\mathbb{E}|Y_{M_{2}(\rho)} - w_{T}^{(M_{2}(\rho))}(X)|^{2} \leq C \mathbb{E}\langle Y_{M_{2}(\rho)}\rangle^{2} \log_{\rho^{-1}}(1 + \rho^{M_{2}(\rho)\delta_{\rho} + \gamma_{\rho} - 1})$$

Recalling (8.8), we see that  $M_2(\rho)\delta_\rho + \gamma_\rho - 1 \ge \delta_\rho + \gamma_\rho$ , and hence

$$\log_{\rho^{-1}} (1 + \rho^{M_2(\rho)\delta_{\rho} + \gamma_{\rho} - 1)} \le \log_{\rho^{-1}} (1 + \rho^{\delta_{\rho} + \gamma_{\rho}}) \le \frac{C}{\log \rho^{-1}}$$

by (8.3). Together with Proposition 8.1, this implies that

$$\mathbb{E}|Y_{M_2(\rho)} - w_T^{(M_2(\rho))}(X)|^2 \le \frac{C(\|u_0\|)}{\log \rho^{-1}}.$$
(8.23)

Combining (8.22) and (8.23), we obtain (8.21).

**Proposition 8.4.** Let T > 0. There is a constant  $C = C(T, \sigma, d) < \infty$  such that for each  $a \in \mathbb{R}^m$ , we have

$$\sup_{m \in [M_1(\rho,T),M_2(\rho)]} \left| \delta_{\rho}^{-1} \operatorname{Var}(Y_m \mid Y_{m-1} = a) - \frac{L_{\rho}(m\delta_{\rho}, a)^{\otimes 2}}{2(d-2)} \right| \le C\langle a \rangle^2 o(1), \quad (8.24)$$

where o(1) denotes a quantity that goes to 0 as  $\rho \downarrow 0$ .

*Proof.* We recall the recurrence (8.12):

$$Y_m = \mathcal{G}_{t'_m - t_m} \mathcal{U}_{t'_{m-1}, t_m} Y_{m-1}(X).$$

We now apply Proposition 7.5 with

$$\tau_1 = t_m - t'_{m-1} = \rho^{(m-1)\delta_{\rho}} (\rho^{\gamma_{\rho}} - \rho^{\delta_{\rho}})$$
(8.25)

and

$$\tau_2 = t'_m - t'_{m-1} = \rho^{(m-1)\delta_\rho + \gamma_\rho} (1 - \rho^{\delta_\rho}) . \tag{8.26}$$

Since

$$\frac{\tau_2}{\tau_1} = \frac{1 - \rho^{\delta_{\rho}}}{1 - \rho^{\delta_{\rho} - \gamma_{\rho}}} \to 1 \quad \text{as } \rho \downarrow 0,$$

we can use (7.7). We can compute for all  $m \in [M_1(\rho, T), M_2(\rho)]$  that

$$\lim_{\rho \downarrow 0} \delta_{\rho}^{-1} \log_{\rho^{-1}} \left( 1 + \frac{1}{2} \cdot \frac{\tau_{1}}{\tau_{2} - \tau_{1} + \rho} \right) = \lim_{\rho \downarrow 0} \delta_{\rho}^{-1} \log_{\rho^{-1}} \left( 1 + \frac{1}{2} \cdot \frac{\rho^{-\delta_{\rho}} (\rho^{\gamma_{\rho}} - \rho^{\delta_{\rho}})}{1 - \rho^{\gamma_{\rho}} + \rho^{1 - m\delta_{\rho}}} \right) = 1$$

$$(8.27)$$

and

$$\lim_{\rho \downarrow 0} |\log_{\rho} \tau_1 - m\delta_{\rho}| = 0, \tag{8.28}$$

and moreover these limits are uniform in m. These two limits, along with the time-continuity of  $L_{\rho}$  proved in Lemma 7.2, mean that the difference between the left side of (8.24) and  $\delta_{\rho}$  times the left side of (7.7) converges to 0 as  $\rho \downarrow 0$ , again uniformly in m.

To bound the right side of (7.7) given our choices (8.25) and (8.26), we compute (similarly to (8.27))

$$\delta_{\rho}^{-1} \left( \log_{\rho^{-1}} \left( 1 + \frac{\tau_1}{\tau_2 - \tau_1} \right) \right)^2 = \delta_{\rho}^{-1} \left( \log_{\rho^{-1}} \left( 1 + \frac{\rho^{\gamma_{\rho}} - \rho^{\delta_{\rho}}}{\rho^{\delta_{\rho}} (1 - \rho^{\gamma_{\rho}})} \right) \right)^2 \le C \delta_{\rho} \to 0 \quad \text{as } \rho \downarrow 0$$

and also  $(\delta_{\rho} \log \rho^{-1})^{-1} \to 0$  as  $\rho \downarrow 0$  by (8.1). With these estimates on its components, the bound (7.7) leads to the conclusion (8.24).

We also need a higher moment bound, which will be a consequence of Proposition 7.6.

**Proposition 8.5.** There is an  $\ell > 2$  such that, for all  $a \in \mathbb{R}^m$ , we have

$$\sup_{m \in [M_1(\rho, T), M_2(\rho)]} \mathbb{E}[|Y_m - Y_{m-1}|^{\ell} | Y_{m-1} = a] \le C \langle a \rangle^{\ell} \delta_{\rho}^{\ell/2}.$$
(8.29)

*Proof.* By (8.12) and Proposition 7.6, the left side of (8.29) is bounded above by the right side of (7.18) with the choices  $\tau_2 = t_m' - t_{m-1}'$  and  $\tau_1 = t_m - t_{m-1}'$ , namely

$$C\langle a \rangle^{\ell} \left( \log_{\rho^{-1}} \frac{t'_m - t'_{m-1} + \rho}{t'_m - t_m + \rho} \right)^{\ell/2}.$$

We note that

$$\log_{\rho^{-1}} \frac{t'_m - t'_{m-1} + \rho}{t'_m - t_m + \rho} \le \log_{\rho^{-1}} \left( 1 + \frac{\rho^{-\delta_{\rho} + \gamma_{\rho}} - \rho^{\gamma_{\rho}}}{1 - \rho^{\gamma_{\rho}}} \right) \le C\delta_{\rho}$$

for an absolute constant C by (8.1). Thus we obtain (8.29).

# 8.2 Convergence of the Markov chain to the forward-backward SDE

In this section we will use the general convergence criterion Theorem A.1 to show that our Markov chain converges to a diffusion. We will furthermore identify the diffusion as a solution to our forward-backward SDE.

*Proof of Theorem 1.1.* Let

$$J_{\rho} = L_{\rho} / \sqrt{2(d-2)}. \tag{8.30}$$

By Proposition 7.4, we can find a sequence  $\rho_k \downarrow 0$  and a continuous function  $J \colon [0,1] \times \mathbb{R}^m \to \mathbb{R}^m \otimes \mathbb{R}^n$  such that

$$J_{\rho_k} \to J$$
 uniformly on compact subsets of  $[0,1] \times \mathbb{R}^{\mathfrak{m}}$ . (8.31)

For the moment, we fix the choice of subsequence, and all quantities may depend on it. Later, we will show that J in fact does not depend on the choice of subsequence.

For  $Q \in [0, 1]$  and  $a \in \mathbb{R}^m$ , let  $\tilde{\Theta}_{a,Q}^J$  solve the SDE problem

$$d\tilde{\Theta}_{a,Q}^{J}(q) = J(1 - q, \tilde{\Theta}_{a,Q}^{J}(q))dB(q), \qquad q \in (1 - Q, 1];$$
(8.32)

$$\tilde{\Theta}_{a,Q}^{J}(1-Q) = a. \tag{8.33}$$

We note that

$$\tilde{\Theta}_{a,Q}^{J}(1) \stackrel{\text{law}}{=} \Theta_{a,Q}^{J}(Q), \tag{8.34}$$

where  $\Theta_{a,Q}^J$  solves (2.1). Now fix  $Q \in [0,1]$  and  $a \in \mathbb{R}^m$  and let  $T_k = \rho_k^{1-Q}$ . Since we will now simultaneously consider the SPDE with varying choices of  $\rho$ , we will use the notation  $\mathcal{U}_{s,t}^{(\rho)}$  for the propagator to indicate the dependence on  $\rho$ . We want to apply Theorem A.1 (with  $J(q,\cdot)$  replaced by  $J(1-q,\cdot)$ ) to show that, for fixed  $X \in \mathbb{R}^d$ , we have

$$\mathcal{U}_{0,T_k}^{(\rho_k)}a(X) \xrightarrow[k \to \infty]{\text{law}} \tilde{\Theta}_{a,Q}^J(1).$$
 (8.35)

We consider the Markov chain  $(Y_m)$  introduced in Section 8.1 with  $T = T_k$ ,  $\rho = \rho_k$ , and  $u_0 \equiv a$ . We let  $A_1(k) = M_1(\rho_k, T_k) + 1$  and  $A_2(k) = M_2(\rho_k)$ . The definitions (8.9–10) mean that (A.1) is satisfied with  $Z \equiv a$ . The second condition of Theorem A.1 is satisfied by Lemma 7.3 and (8.31). By the definitions (8.7–8), the condition (A.2) is satisfied with  $A_1 = 1 - Q$  and  $A_2 = 1$ . Also, by (8.31) and Proposition 8.4, we see that (A.3) is satisfied, and by Proposition 8.5, we see that (A.4) is satisfied. Therefore, all of the conditions of Theorem A.1 are satisfied, and hence (8.35) holds.

We can now conclude that

$$J(Q, a) \stackrel{\text{(8.31)}}{=} \lim_{k \to \infty} J_{\rho_k}(Q, a) \stackrel{\text{(8.30)}}{=} \frac{\lim_{k \to \infty} \mathbb{E}[\sigma(\mathcal{U}_{0, T_k}^{(\rho_k)} a(X))]}{\sqrt{2(d-2)}} \stackrel{\text{(8.35)}}{=} \frac{\mathbb{E}[\sigma(\tilde{\Theta}_{a, Q}^J(1))]}{\sqrt{2(d-2)}}$$

$$\stackrel{\text{(8.34)}}{=} \frac{\mathbb{E}[\sigma(\Theta_{a, Q}^J(Q))]}{\sqrt{2(d-2)}} \stackrel{\text{(2.2)}}{=} \mathcal{Q}_{\sigma/\sqrt{2(d-2)}} J(Q, a). \tag{8.36}$$

In the third identity we additionally used the fact that the family of random variables  $(\sigma(\mathcal{U}_{0,T_k}^{(\rho_k)}a(X)))_k$  is uniformly integrable by Proposition 4.1 and the assumption that  $\sigma$  is Lipschitz. But (8.36) means that J satisfies (2.3), and hence by Proposition 2.1 that J is uniquely determined, and therefore coincides with the function J coming from the solution of (1.6a–c). In particular, it does not depend on the choice of subsequence, and so in fact (8.31) can be upgraded to convergence as  $\rho \downarrow 0$ .

We can now apply Theorem A.1 again, equipped this time with this stronger notion of convergence. Let  $\rho_k \downarrow 0$  be *any* sequence decreasing to zero. We consider general  $u_0$  satisfying the conditions of the theorem, fix T = t, and let  $A_1(k) = M_1(\rho_k, T)$  and  $A_2(k) = M_2(\rho_k)$ . We note by the definitions (8.9–10) that (again using the Markov chain  $(Y_m)$  from Section 8.1 with these choices), we have

$$Y_{A_1(k)} = \mathcal{G}_{t'_{M_1(\rho_k,T_k)})} u_0(x) \to \mathcal{G}_t u_0(x)$$
 in probability as  $k \to \infty$ 

since  $t'_{M_1(\rho_k,T_k)} \to t$  as  $k \to \infty$  by (8.3–8.5) and (8.7). Thus, using Theorem A.1 in the same way as above with Q=0, we get

$$\mathcal{U}_{0,t}^{(\rho)}u_0(x) \xrightarrow[\rho\downarrow 0]{\text{law}} \Gamma_{\mathcal{G}_t u_0(x),1}(1),$$

where  $\Gamma_{G_t u_0(x),1}$  solves (1.6a–b) with Q=1. We can then upgrade this convergence to the Wasserstein convergence (1.8) using the moment bound in Proposition 4.1.

# 9 Analysis of the second moment in the scalar linear case

In this section we prove Theorem 1.5. The proof is independent of the rest of the paper, and uses different methods. Specifically, since the problem is linear in the initial data, we can write a closed PDE for the two-point correlation function of the solution. A variant of this PDE was studied in [1], and we use an explicit supersolution given there to bound the solution.

*Proof of Theorem 1.5.* As in the theorem statement, we assume that  $\mathfrak{m}=\mathfrak{n}=1$  and  $\sigma(u)=\beta u$  for some  $\beta\in(0,\infty)$ , and let  $(u_t(x))$  solve (1.2) with initial condition  $u_0\equiv 1$ . Fix  $T<\infty$  arbitrarily. We define the function  $k\colon\mathbb{R}^d\to\mathbb{R}$  by

$$k_t(x) \coloneqq \mathbb{E}[u_t(0)u_t(x)]. \tag{9.1}$$

Then, by Itô's formula applied to the SPDE (1.2), the function k satisfies the PDE

$$\partial_t k_t(x) = \left(\Delta + \frac{\beta^2}{\log \rho^{-1}} R^{\rho}(x)\right) k_t(x). \tag{9.2a}$$

$$k_0(x) = 1.$$
 (9.2b)

Let v solve the PDE (9.2a) but with initial condition  $v_0 = \delta_0$  (a delta distribution at the origin). By the symmetry of the right side of (9.2a), we have

$$\mathbb{E}[u_t(x)^2] = k_t(0) = \int_{\mathbb{R}^d} v_t(x) \, \mathrm{d}x.$$
 (9.3)

By the comparison principle, we have

$$v_t(x) \le \exp\left\{\frac{\beta^2 R^{\rho}(0)t}{\log \rho^{-1}}\right\} G_{2t}(x).$$
 (9.4)

In particular, this means that that

$$v_{\rho/2}(x) \le \exp\left\{\frac{\beta^2 R^{\rho}(0)\rho}{2\log \rho^{-1}}\right\} G_{2\rho}(x) \stackrel{(3.3)}{=} \exp\left\{\frac{\beta^2}{4(d-2)\log \rho^{-1}}\right\} G_{\rho}(x) \le 2G_{\rho}(x) \quad (9.5)$$

if  $\rho$  is sufficiently small.

Let  $A < \infty$  be large enough that

$$R^{\rho}(x) \le A|x|^{-2}$$
 for all  $x \in \mathbb{R}^d \setminus \{0\}$  and all  $\rho > 0$ . (9.6)

(That is, let A be the C from (3.6).) Put

$$\alpha_{\rho} = \frac{d-2}{2} \left( 1 - \sqrt{1 - \frac{4A\beta^2}{(d-2)^2 \log \rho^{-1}}} \right)$$
 (9.7)

and define

$$\tilde{v}_t(x) = t^{\alpha_\rho} |x|^{-\alpha_\rho} G_{2t}(x). \tag{9.8}$$

Then we have

$$\partial_t \tilde{v}_t(x) = \left(\Delta + \frac{A\beta^2}{|x|^2 \log \rho^{-1}}\right) \tilde{v}_t(x), \tag{9.9}$$

as computed in [1, §2.1, p. 902]. There is a constant C, independent of  $\rho \in (0, 1]$ , such that  $G_1(x) \le C|x|^{-\alpha_\rho}G_2(x)$ , which means that

$$G_{\rho}(x) = \rho^{-d/2}G_1(\rho^{-1/2}x) \leq C\rho^{\alpha_{\rho}/2 - d/2}|x|^{-\alpha_{\rho}}G_2(\rho^{-1/2}x) = C\rho^{\alpha_{\rho}/2}|x|^{-\alpha_{\rho}}G_{2\rho}(x)$$

for all  $x \in \mathbb{R}^d$ . Using this with (9.5) and (9.8), we see that

$$v_{\rho/2}(x) \le 2C\rho^{-\alpha_{\rho}/2}\tilde{v}_{\rho}(x)$$
 for all  $x \in \mathbb{R}^d$ . (9.10)

The condition (9.6) implies a comparison principle between (9.2a) and (9.9). Using this comparison principle along with the comparison (9.10), we see that

$$v_t(x) \le 2C\rho^{-\alpha_\rho/2}\tilde{v}_{t-\rho/2}(x)$$
 for all  $t \ge \rho/2$  and all  $x \in \mathbb{R}^d$ . (9.11)

In particular, we have

$$\int v_t(x) \, \mathrm{d}x \le 2C \rho^{-\alpha_\rho/2} (t - \rho/2)^{\alpha_\rho} \int |x|^{\alpha_\rho} G_{2t}(x) \, \mathrm{d}x. \tag{9.12}$$

On the right side of (9.12), the integral and the term  $(t - \rho/2)^{\alpha_{\rho}}$  are both bounded above, uniformly in  $\rho \in (0, 1]$  and in  $t \in [\rho/2, T]$ . Also, from (9.7) and the elementary bound  $\sqrt{1-x} \ge 1-x$  for  $x \in [0, 1]$ , we notice that for sufficiently small  $\rho$ , we have

$$\alpha_{\rho} \le \frac{2A\beta^2}{(d-2)\log \rho^{-1}},$$

which means that

$$\rho^{-\alpha_{\rho}/2} = \exp\left\{\frac{\alpha_{\rho}}{2}\log\rho^{-1}\right\} \le \exp\left\{\frac{A\beta^2}{d-2}\right\}$$

is bounded above independently of  $\rho$ . Hence, the right side of (9.12) is bounded above independently of  $\rho$  and  $t \in [\rho/2, T]$ . Thus, by (9.3), the quantity  $\sup_{t \in [\rho/2, T]} \mathbb{E}[u_t(0)^2]$  is bounded uniformly in  $\rho$ .

To complete the proof, we must show that  $\sup_{t \in [0, \rho/2]} \mathbb{E}[u_t(0)^2]$  is bounded uniformly in  $\rho$ . But this is clear from (3.3), (9.3) and (9.4).

# Appendix A Convergence of Markov chains to diffusions

This section is exactly analogous to [12, Appendix A]. We need a multidimensional version of that result, which again is a simple application of the results of [24, Section 11.2].

**Theorem A.1.** Suppose that we are given a sequence  $\delta_k \downarrow 0$ , a sequence of  $\mathbb{R}^{\mathfrak{m}}$ -valued discrete Markov martingales  $(\{Y_m^{(k)}\}_{m\in A_1(k),...,A_2(k)})_{k=1}^{\infty}$ , and a continuous function  $J\colon [A_1,A_2]\times\mathbb{R}^{\mathfrak{m}}\to\mathbb{R}^{\mathfrak{m}}\otimes\mathbb{R}^{\mathfrak{n}}$  satisfying the following conditions:

1. There is an  $\mathbb{R}^{\mathfrak{m}}$ -valued random variable Z such that

$$Y_{A_1(k)}^{(k)} \xrightarrow[law]{k \to \infty} Z$$
 as  $k \to \infty$ . (A.1)

- 2. For each  $q \in [A_1, A_2]$ , the function  $J(q, \cdot) \colon \mathbb{R}^m \to \mathbb{R}^m \otimes \mathbb{R}^n$  is Lipschitz, and the Lipschitz constant is bounded above independent of q.
- 3. We have  $\delta_k m \in [A_1, A_2]$  for all  $k \ge 1$  and  $m = A_1(k), \ldots, A_2(k)$ , and moreover

$$\lim_{k \to \infty} \delta_k A_1(k) = A_1 \qquad and \qquad \lim_{k \to \infty} \delta_k A_2(k) = A_2. \tag{A.2}$$

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

$$\lim_{k \to \infty} \sup_{\substack{|x| \le R \\ A_1(k) \le m < A_2(k)}} \left| \delta_k^{-1} \operatorname{Var}[Y_{m+1}^{(k)} \mid Y_m^{(k)} = x] - J(\delta_k m, x) J(\delta_k m, x)^{\top} \right| = 0.$$
(A.3)

5. There is an  $\ell > 2$  such that, for each  $R < \infty$ , we have

> 2 such that, for each 
$$R < \infty$$
, we have 
$$\sup_{\substack{k < \infty, |x| \le R, \\ A_1(k) \le m < A_2(k)}} \delta_k^{-\ell/2} \mathbb{E}[|Y_{m+1}^{(k)} - Y_m^{(k)}|^{\ell} | Y_m^{(k)} = x] < \infty. \tag{A.4}$$

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

$$\begin{split} \mathrm{d}Y(q) &= J(q,Y(q))\mathrm{d}B(q), \qquad q \in (A_1,A_2); \\ Y(A_1) &= X, \end{split}$$

where B(q) is a standard  $\mathbb{R}^n$ -valued Brownian motion. Then we have

$$Y_{A_2(k)}^{(k)} \xrightarrow[k\to\infty]{\text{law}} Y(A_2).$$

The proof follows that of [12, Theorem A.1], since the needed results from [24] are not specific to the one-dimensional case. We omit the details.

#### References

- [1] Boumediene Abdellaoui, Ireneo Peral, and Ana Primo. Influence of the Hardy potential in a semilinear heat equation. *Proc. Roy. Soc. Edinburgh Sect. A*, 139(5):897–926, 2009. DOI: 10.1017/S0308210508000152.
- [2] Lorenzo Bertini and Nicoletta Cancrini. The two-dimensional stochastic heat equation: renormalizing a multiplicative noise. *J. Phys. A*, 31(2):615–622, 1998. DOI: 10.1088/0305-4470/31/2/019.
- [3] Lorenzo Bertini and Giambattista Giacomin. Stochastic Burgers and KPZ equations from particle systems. *Comm. Math. Phys.*, 183(3):571–607, 1997. DOI: 10.1007/s002200050044.
- [4] 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. DOI: 10.1007/s00220-019-03527-z.
- [5] Francesco Caravenna, Rongfeng Sun, and Nikos Zygouras. The critical 2d stochastic heat flow. *Invent. Math.*, 233(1):325–460, 2023. DOI: 10.1007/s00222-023-01184-7.
- [6] Francesco Caravenna, Rongfeng Sun, and Nikos Zygouras. The two-dimensional KPZ equation in the entire subcritical regime. *Ann. Probab.*, 48(3):1086–1127, 2020. DOI: 10.1214/19-AOP1383.
- [7] Francesco Caravenna, Rongfeng Sun, and Nikos Zygouras. Universality in marginally relevant disordered systems. *Ann. Appl. Probab.*, 27(5):3050–3112, 2017. DOI: 10.1214/17-AAP1276.
- [8] Clément Cosco, Francesca Cottini, and Anna Donadini. A central limit theorem for two-dimensional directed polymers with critical spatial correlation, 2025. arXiv: 2509.16694v1.
- [9] Alexander Dunlap and Cole Graham. Edwards-Wilkinson fluctuations in subcritical 2D stochastic heat equations, 2024. arXiv: 2405.09520v1.

- [10] Alexander Dunlap and Cole Graham. The 2D nonlinear stochastic heat equation: pointwise statistics and the decoupling function, 2023. arXiv: 2308.11850v1.
- [11] Alexander Dunlap and Cole Graham. Uniqueness and root-Lipschitz regularity for a degenerate heat equation. *SIAM J. Math. Anal.*, 57(2):1866–1891, 2025. DOI: 10.1137/23M1596260.
- [12] Alexander Dunlap and Yu Gu. A forward-backward SDE from the 2D nonlinear stochastic heat equation. *Ann. Probab.*, 50(3):1204–1253, 2022. DOI: 10.1214/21-aop1563.
- [13] Luca Gerolla, Martin Hairer, and Xue-Mei Li. Fluctuations of stochastic PDEs with long-range correlations. *Ann. Appl. Probab.*, 35(2):1198–1232, 2025. DOI: 10.1214/24-aap2140.
- [14] Luca Gerolla, Martin Hairer, and Xue-Mei Li. Scaling limit of the KPZ equation with non-integrable spatial correlations, 2024. arXiv: 2407.13215v1.
- [15] Yu Gu. Gaussian fluctuations from the 2D KPZ equation. *Stoch. Partial Differ. Equ. Anal. Comput.*, 8(1):150–185, 2020. DOI: 10.1007/s40072-019-00144-8.
- [16] 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. DOI: 10.1137/19M1296380.
- [17] Yu Gu, Jeremy Quastel, and Li-Cheng Tsai. Moments of the 2D SHE at criticality. *Probab. Math. Phys.*, 2(1):179–219, 2021. DOI: 10.2140/pmp.2021.2.179.
- [18] I. A. Ibragimov. On the composition of unimodal distributions. *Theory Probab. Appl.*, 1(2):255–260, 1956. DOI: 10.1137/1101021.
- [19] Davar Khoshnevisan. *Analysis of stochastic partial differential equations*, vol. 119 of *CBMS Reg. Conf. Ser. Math.* Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 2014. DOI: 10.1090/cbms/119.
- [20] Sefika Kuzgun and Ran Tao. Time-dependent averages of a critical long-range stochastic heat equation, 2024. arXiv: 2411.09058v1.
- [21] Jin Ma and Jiongmin Yong. Forward-backward stochastic differential equations and their applications, vol. 1702 of Lecture Notes in Math. Springer-Verlag, Berlin, 1999. xiv+270.
- [22] Carl Mueller and Roger Tribe. A singular parabolic Anderson model. *Electron. J. Probab.*, 9:no. 5, 98–144, 2004. DOI: 10.1214/EJP.v9-189.
- [23] David Nualart. *The Malliavin calculus and related topics*. Probability and its Applications (New York). Springer-Verlag, Berlin, second edition, 2006. xiv+382.
- [24] Daniel W. Stroock and S. R. Srinivasa Varadhan. *Multidimensional diffusion processes*. Classics in Mathematics. Springer-Verlag, Berlin, 2006. xii+338. Reprint of the 1997 edition.
- [25] Ran Tao. Gaussian fluctuations of a nonlinear stochastic heat equation in dimension two. *Stoch. Partial Differ. Eq. Anal. Comp.*, 12(1):220–246, 2024. DOI: 10.1007/s40072-022-00282-6.
- [26] Ran Tao. Mesoscopic averaging of the two-dimensional KPZ equation. *J. Stat. Phys.*, 191(1):Paper No. 6, 28, 2024. DOI: 10.1007/s10955-023-03222-3.
- [27] Li-Cheng Tsai. Stochastic heat flow by moments, 2024. arXiv: 2410.14657v1.