The Annals of Probability
2019, Vol. 47, No. 2, 989–1035
https://doi.org/10.1214/18-AOP1277

COMPARISON PRINCIPLE FOR STOCHASTIC HEAT EQUATION ON $\mathbb{R}^d$

BY LE CHEN AND JINGYU HUANG

University of Nevada, Las Vegas and University of Utah

We establish the strong comparison principle and strict positivity of solutions to the following nonlinear stochastic heat equation on $\mathbb{R}^d$

$$\left(\frac{\partial}{\partial t} - \frac{1}{2}\Delta\right)u(t,x) = \rho\big(u(t,x)\big)\dot{M}(t,x),$$

for measure-valued initial data, where $\dot{M}$ is a spatially homogeneous Gaussian noise that is white in time and ρ is Lipschitz continuous. These results are obtained under the condition that $\int_{\mathbb{R}^d}(1+|\xi|^2)^{\alpha-1}\hat{f}(\mathrm{d}\xi) < \infty$ for some $\alpha \in (0,1]$, where $\hat{f}$ is the spectral measure of the noise. The weak comparison principle and nonnegativity of solutions to the same equation are obtained under Dalang's condition, that is, $\alpha = 0$. As some intermediate results, we obtain handy upper bounds for $L^p(\Omega)$-moments of $u(t,x)$ for all $p \geq 2$, and also prove that u is a.s. Hölder continuous with order $\alpha - \varepsilon$ in space and $\alpha/2 - \varepsilon$ in time for any small $\varepsilon > 0$.

1. Introduction. In this paper, we study the *sample-path comparison principle*, or simply *comparison principle* of the solutions to the following stochastic heat equation (SHE) with *rough initial conditions*:

$$\begin{cases} \left(\dfrac{\partial}{\partial t} - \dfrac{1}{2}\Delta\right)u(t,x) = \rho\big(u(t,x)\big)\dot{M}(t,x) & x \in \mathbb{R}^d t > 0, \\ u(0,\cdot) = \mu(\cdot). \end{cases} \tag{1.1}$$

In this equation, ρ is assumed to be a globally Lipschitz continuous function. The linear case, that is, $\rho(u) = \lambda u$, is called the *parabolic Anderson model* (PAM) [3]. The noise $\dot{M}$ is a Gaussian noise that is white in time and homogeneously colored in space. Informally,

$$\mathbb{E}\big[\dot{M}(t,x)\dot{M}(s,y)\big] = \delta_0(t-s)f(x-y),$$

where δ_0 is the Dirac delta measure with unit mass at zero and f is a "correlation function" that is, a nonnegative and nonnegative definite function that is not identically zero. The Fourier transform of f is denoted by $\hat{f}$

$$\hat{f}(\xi) = \mathcal{F}f(\xi) = \int_{\mathbb{R}^d} \exp(-i\xi \cdot x) f(x)\,\mathrm{d}x.$$

Received July 2016; revised November 2017.
MSC2010 subject classifications. Primary 60H15; secondary 35R60, 60G60.
Key words and phrases. Stochastic heat equation, parabolic Anderson model, space-time Hölder regularity, spatially homogeneous noise, comparison principle, measure-valued initial data.

In general, $\hat{f}$ is again a nonnegative and nonnegative definite measure, which is usually called the *spectral measure*. The precise meaning of the "rough initial conditions/data" are specified as follows. We first note that by the Jordan decomposition, any signed Borel measure μ can be decomposed as $\mu = \mu_+ - \mu_-$ where $\mu_\pm$ are two nonnegative Borel measures with disjoint support. Denote $|\mu| := \mu_+ + \mu_-$. The rough initial data refers to any signed Borel measure μ such that

$$\int_{\mathbb{R}^d} e^{-a|x|^2} |\mu|(\mathrm{d}x) < +\infty \qquad \text{for all } a > 0, \tag{1.2}$$

where $|x| = \sqrt{x_1^2 + \cdots + x_d^2}$ denotes the Euclidean norm. It is easy to see that condition (1.2) is equivalent to the condition that the solution to the homogeneous equation—$J_0(t, x)$ defined in (1.6) below—exists for all $t > 0$ and $x \in \mathbb{R}^d$.

The comparison principle refers to the property that if two initial conditions are ordered, then the corresponding solutions to the stochastic partial differential equations are also ordered. For any Borel measure μ on $\mathbb{R}^d$, "$\mu \geq 0$" has its obvious meaning that μ is a nonnegative measure and "$\mu > 0$" refers to the fact that $\mu \geq 0$ and μ is nonvanishing, that is, $\mu \neq 0$. Let u_1 and u_2 be two solutions starting from two measures μ_1 and μ_2, respectively. We say that (1.1) satisfies the *weak comparison principle* if $u_1(t, x) \leq u_2(t, x)$ a.s. for all $t > 0$ and $x \in \mathbb{R}^d$ whenever $\mu_1 \leq \mu_2$. Similarly, we say that (1.1) satisfies the *strong comparison principle* if $u_1(t, x) < u_2(t, x)$ for all $t > 0$ and $x \in \mathbb{R}^d$ a.s. whenever $\mu_1 < \mu_2$. Note that when $\rho(u) = \lambda u$, it is relatively easier to establish the weak comparison principle since the solution can be approximated by its regularized version, which admits a Feynman–Kac formula; see [15, 16, 18].

Most strong comparison principles are obtained through Mueller's original work [20], where he proved the case when $d = 1$, $\dot{M}$ is the space-time white noise, $\rho(u) = |u|^\gamma$ (for all $\gamma \leq 1$), and the initial data is a bounded function. In [23], Shiga studied the same equation as that in [20] except that ρ is assumed to be Lipschitz and there can be a drift term. By using concentration of measure arguments for discrete directed polymers in Gaussian environments, Flores established in [19] the strict positivity of solution to 1-d PAM with Dirac delta initial data. Following arguments by Mueller and Shiga, Chen and Kim extended these results in [7] to allow both fractional Laplace operators and rough initial data. Recently, by using paracontrolled distributions, Gubinelli and Perkowski gave an intrinsic proof of the strict positivity; see [14]. Their proof does not depend on the details of noise, though they require the initial data to be a function that is strict positive anywhere.

When $d \geq 2$, in order to study a random field solution, the noise has to be colored in space, where "colored" means "correlated." Equation (1.1) has been much studied since the introduction by Dawson and Salehi [12] as a model for the growth of a population in a random environment. In [10, 11], it is shown that if the initial condition is a bounded function, and under some integrability condition on $\hat{f}$, now called *Dalang's* condition, that is,

$$\Upsilon(\beta) := (2\pi)^{-d} \int_{\mathbb{R}^d} \frac{\hat{f}(\mathrm{d}\xi)}{\beta + |\xi|^2} < +\infty \qquad \text{for some and hence for all } \beta > 0, \tag{1.3}$$

there is a unique random field solution to equation (1.1). This equation has been extensively studied; see, for example, [6, 13, 15, 18]. Recently, Chen and Kim showed that Dalang's condition (1.3) also guarantees an $L^2(\Omega)$-continuous random field solution starting from rough initial conditions; see [8]. To the best of our knowledge, comparison principle in this setting is much less known, though people believe that it is true. In [24], Tessitore and Zabczyk proved the strict positivity for the case when $\hat{f}$ belongs to $L^p(\mathbb{R}^d)$ for some $p \in [1, d/(d-2))$. Clearly, this condition excludes the important Riesz kernel case, that is, $f(x) = |x|^{-\beta}$ with $\beta \in (0, 2 \wedge d)$. Indeed, we will show that under Dalang's condition (1.3), if $\rho(0) = 0$, then the solution $u(t,x)$ starting from any nonnegative rough initial data is a.s. nonnegative for any $t > 0$ and $x \in \mathbb{R}^d$. Moreover, if the nonnegative rough initial data is nonvanishing and f satisfies

$$\int_{\mathbb{R}^d} \frac{\hat{f}(\mathrm{d}\xi)}{(1 + |\xi|^2)^{1-\alpha}} < \infty \qquad \text{for some } \alpha \in (0, 1], \tag{1.4}$$

then we are able to establish the strict positivity of $u(t,x)$ through the following small-ball probability estimate:

$$\mathbb{P}\big(u(t,x) < \varepsilon\big) \le A \exp\big(-A|\log \varepsilon|^{\alpha}\big(\log|\log \varepsilon|\big)^{1+\alpha}\big).$$

Similar small-ball probabilities in various settings can be found in [7, 9, 19, 21]. These nonnegativity statements can be translated into comparison statements by considering $v = u_1 - u_2$.

Condition (1.4) is natural since in a recent paper [6], it is shown that Dalang's condition (1.3) alone cannot guarantee the existence of a continuous version of the solution. There might be solutions that behave so badly that they may hit zero. Whether this phenomenon does happen is still not clear to us and it is left for future exploration. For the moment, we are content with this slightly stronger condition (1.4). Indeed, if the initial condition is a bounded function, Sanz-Solé and Sarrà [22] showed that condition (1.4) guarantees that the solution is a.s. Hölder continuous with order $\alpha - \varepsilon$ in space and $\alpha/2 - \varepsilon$ in time for any small $\varepsilon > 0$. In this paper, we have extended this result for rough initial conditions. The space-time white noise case is proved in [4].

In all these studies, the moment bounds/formulas play an important role. The upper bounds for the second moments under Dalang's condition (1.3) for rough initial conditions is obtained in [8]. In this paper, we extend this bound to obtain similar upper bounds for all pth moments, $p \geq 2$. Using these moments upper bounds, we establish the (weak) comparison principle. Note that similar moment upper bounds have also been recently obtained by the second author [17]. By contrast, the major effort of [8] is to obtain some nontrivial lower bounds for the second moments. Note when $\rho(u) = \lambda u$, the pth moment admits a Feynman–Kac representation, which has been exploited to study the intermittency phenomenon in [15, 16, 18].

In the rest of this Introduction, we will first give the precise definition of the solution and recall the existence/uniqueness result in Section 1.1. The main results are stated in Section 1.2. Then we give an outline of the rest of the paper in Section 1.3.

1.1. *Definition and existence of a solution.* Recall that a *spatially homogeneous Gaussian noise that is white in time* is an $L^2(\Omega)$-valued mean zero Gaussian process on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$

$$\{F(\psi) : \psi \in C_c^\infty([0,\infty) \times \mathbb{R}^d)\},$$

such that

$$\mathbb{E}[F(\psi)F(\phi)] = \int_0^\infty \mathrm{d}s \iint_{\mathbb{R}^{2d}} \psi(s,x)\phi(s,y) f(x-y)\,\mathrm{d}x\,\mathrm{d}y.$$

Let $\mathcal{B}_b(\mathbb{R}^d)$ be the collection of Borel measurable sets with finite Lebesgue measure. As in Dalang–Walsh theory [10, 25], one can extend F to a σ-finite $L^2(\Omega)$-valued martingale measure $B \mapsto F(B)$ defined for $B \in \mathcal{B}_b(\mathbb{R}_+ \times \mathbb{R}^d)$, where $\mathbb{R}_+ := [0,\infty)$. Then define

$$M_t(B) := F([0,t] \times B), \qquad B \in \mathcal{B}_b(\mathbb{R}^d).$$

Let $(\mathcal{F}_t, t \geq 0)$ be the natural filtration generated by $M_\cdot(\cdot)$ and augmented by all $\mathbb{P}$-null sets $\mathcal{N}$ in $\mathcal{F}$, that is,

$$\mathcal{F}_t := \sigma(M_s(A) : 0 \leq s \leq t, A \in \mathcal{B}_b(\mathbb{R}^d)) \vee \mathcal{N}, \qquad t \geq 0.$$

Then for any adapted, jointly measurable (with respect to $\mathcal{B}((0,\infty) \times \mathbb{R}^d) \times \mathcal{F}$) random field $\{X(t,x) : t > 0, x \in \mathbb{R}^d\}$ such that for all integers $p \geq 2$,

$$\int_0^\infty \mathrm{d}s \iint_{\mathbb{R}^{2d}} \mathrm{d}x\,\mathrm{d}y \|X(s,y)X(s,x)\|_{\frac{p}{2}} f(x-y) < \infty,$$

the stochastic integral

$$\int_0^\infty \int_{\mathbb{R}^d} X(s,y) M(\mathrm{d}s, \mathrm{d}y)$$

is well defined in the sense of Dalang–Walsh. Here, we only require the joint-measurability instead of predictability; see Proposition 2.2 in [8] for this case or Proposition 3.1 in [5] for the space-time white noise case. Throughout this paper, $\|\cdot\|_p$ denotes the $L^p(\Omega)$-norm.

The solution to (1.1) is understood in the mild form

$$u(t,x) = J_0(t,x) + \int_0^t \int_{\mathbb{R}^d} G(t-s, x-y)\rho\big(u(s,y)\big)M(\mathrm{d}s,\mathrm{d}y). \tag{1.5}$$

Here, $J_0(t,x)$ denotes the solution to the homogeneous equation

$$J_0(t,x) := \big(\mu * G(t,\cdot)\big)(x) := \int_{\mathbb{R}^d} G(t,x-y)\mu(\mathrm{d}y), \tag{1.6}$$

where

$$G(t,x) = (2\pi t)^{-d/2}\exp\left(-\frac{|x|^2}{2t}\right). \tag{1.7}$$

Denote

$$I(t,x) := \iint_{[0,t]\times\mathbb{R}^d} G(t-s,x-y)\rho\big(u(s,y)\big)M(\mathrm{d}s,\mathrm{d}y).$$

The above stochastic integral is understood in the sense of Walsh [10, 25].

DEFINITION 1.1. A process $u = (u(t,x), (t,x) \in (0,\infty)\times\mathbb{R}^d)$ is called a *random field solution* to (1.1) if:

(1) u is adapted, that is, for all $(t,x) \in (0,\infty)\times\mathbb{R}^d$, $u(t,x)$ is $\mathcal{F}_t$-measurable;
(2) u is jointly measurable with respect to $\mathcal{B}((0,\infty)\times\mathbb{R}^d)\times\mathcal{F}$;
(3) $\|I(t,x)\|_2 < +\infty$ for all $(t,x) \in (0,\infty)\times\mathbb{R}^d$;
(4) I is $L^2(\Omega)$-continuous, that is, the function $(t,x) \mapsto I(t,x)$ mapping $(0,\infty)\times\mathbb{R}^d$ into $L^2(\Omega)$ is continuous;
(5) u satisfies (1.5) a.s., for all $(t,x) \in (0,\infty)\times\mathbb{R}^d$.

Definition 1.1 does not require a random field solution to have a pathwise continuous version. The $L^2(\Omega)$-continuity in condition (4) is a much weaker condition than the condition of having continuous sample path. Actually, one can construct a discontinuous solution as in [6]. On the other hand, from Definition 1.1 one can find sufficient conditions for both the admissible initial data and the admissible correlation function f, which is the content of the following theorem.

THEOREM 1.2 (Theorem 2.4 in [8]). *If the initial data μ satisfies* (1.2), *then under Dalang's condition* (1.3), *SPDE* (1.1) *has a unique* (*in the sense of versions*) *random field solution $\{u(t,x) : t > 0, x \in \mathbb{R}^d\}$ starting from μ. This solution is $L^2(\Omega)$-continuous.*

The existence of the random field solution (except the $L^2(\Omega)$-continuity) has also been obtained recently by the second author in [17]. Note that the $L^2(\Omega)$-continuity that comes with Theorem 1.2 is too weak to be useful in this paper. When we need the pathwise continuity, we will instead work under a stronger condition—(1.4)—than Dalang's condition (1.3).

1.2. *Statements of the main results.* We will prove seven theorems as follows.

THEOREM 1.3 (Weak comparison principle). *Assume that f satisfies Dalang's condition* (1.3). *Let u_1 and u_2 be two solutions to* (1.1) *with the initial measures μ_1 and μ_2 that satisfy* (1.2), *respectively. If $\mu_1 \le \mu_2$, then*

$$\mathbb{P}\big(u_1(t,x) \le u_2(t,x)\big) = 1 \qquad \text{for all } t \ge 0 \text{ and } x \in \mathbb{R}^d. \tag{1.8}$$

Moreover, if the paths of $u_1(t,x)$ and $u_2(t,x)$ are a.s. continuous, then

$$\mathbb{P}\big(u_1(t,x) \le u_2(t,x) \text{ for all } t \ge 0 \text{ and } x \in \mathbb{R}^d\big) = 1. \tag{1.9}$$

If $\rho(0) = 0$, then $u \equiv 0$ is the unique solution to (1.1) starting from $\mu = 0$. Hence, we have the following corollary.

COROLLARY 1.4 (Nonnegativity). *Assume that f satisfies Dalang's condition* (1.3) *and $\rho(0) = 0$. Let u be the solution to* (1.1) *with the initial measure μ that satisfies* (1.2). *If $\mu \ge 0$, then*

$$\mathbb{P}\big(u(t,x) \ge 0\big) = 1 \qquad \text{for all } t \ge 0 \text{ and } x \in \mathbb{R}^d. \tag{1.10}$$

Moreover, if the path of $u(t,x)$ are a.s. continuous, then

$$\mathbb{P}\big(u(t,x) \ge 0 \text{ for all } t \ge 0 \text{ and } x \in \mathbb{R}^d\big) = 1. \tag{1.11}$$

THEOREM 1.5 (Strong comparison principle). *Assume that f satisfies* (1.4) *for some $\alpha \in (0,1]$. Let u_1 and u_2 be two* (*continuous versions of the*) *solutions to* (1.1) *with the initial data μ_1 and μ_2, respectively. Then the fact $\mu_1 < \mu_2$ implies*

$$\mathbb{P}\big(u_1(t,x) < u_2(t,x) \text{ for all } t > 0 \text{ and } x \in \mathbb{R}^d\big) = 1. \tag{1.12}$$

Note that by Theorem 1.8 below, under the assumptions of Theorem 1.5, the solution to (1.1) has a continuous version.

THEOREM 1.6 (Strict positivity). *Assume that f satisfies* (1.4) *for some $\alpha \in (0,1]$ and $\rho(0) = 0$. Let u be the solution to* (1.1) *with initial measure $\mu > 0$ that satisfies* (1.2). *Then for any compact set $K \subset (0,\infty) \times \mathbb{R}^d$, there exists a finite constant $A > 0$ which only depends on K such that for all $\varepsilon > 0$ small enough,*

$$\mathbb{P}\Big(\inf_{(t,x)\in K} u(t,x) < \varepsilon\Big) \le A \exp\big(-A|\log\varepsilon|^{\alpha}\big(\log|\log\varepsilon|\big)^{1+\alpha}\big). \tag{1.13}$$

In order to establish the above results, we need to prove the following four theorems, which are of interest by themselves. The first result is a general moment bound. This provides us with a very handy tool in studying various properties of the solution to (1.1). This result extends the previous work [8] from the two-point correlation function to higher moments. Let $\mathrm{Lip}_\rho > 0$ be the Lipschitz constant for ρ. See Section 2 for the proof.

THEOREM 1.7 (Moment bounds). *Under Dalang's condition* (1.3), *if the initial data μ is a signed measure that satisfies* (1.2), *then the solution u to* (1.1) *for any given $t > 0$ and $x \in \mathbb{R}^d$ is in $L^p(\Omega)$, $p \geq 2$, and*

$$\|u(t,x)\|_p \leq \left[\overline{\varsigma} + \sqrt{2}\big(|\mu| * G(t,\cdot)\big)(x)\right] H(t;\gamma_p)^{1/2}, \tag{1.14}$$

where $\overline{\varsigma} = |\rho(0)|/\operatorname{Lip}_\rho$ and $\gamma_p = 32p\operatorname{Lip}_\rho^2$ and $H(t;\gamma_p)$ is defined in (2.4) *below. Moreover, if for some $\alpha \in (0,1]$ condition* (1.4) *is satisfied, then when $p \geq 2$ is large enough, there exists some constant $C > 0$ such that*

$$\|u(t,x)\|_p \leq C\left[\overline{\varsigma} + \big(|\mu| * G(t,\cdot)\big)(x)\right] \exp\big(C \operatorname{Lip}_\rho^{2/\alpha} p^{1/\alpha} t\big). \tag{1.15}$$

The second result is about the sample-path regularity under (1.4) for rough initial data. This result is used to obtain a large deviation estimates in proving the strong comparison principle. See Section 3 for its proof.

THEOREM 1.8 (Hölder regularity). *Suppose that μ is any measure that satisfies* (1.2) *and f satisfies* (1.4) *for some $\alpha \in (0,1]$. Then the solution to* (1.1) *starting from μ has a version which is a.s. β_1-Hölder continuous in time and β_2-Hölder continuous in space on $(0,\infty) \times \mathbb{R}^d$ for all*

$$\beta_1 \in (0,\alpha/2) \quad \textit{and} \quad \beta_2 \in (0,\alpha).$$

The third theorem consists of two approximation results, which are used to establish the weak comparison principle. The first one says that we can approximate a solution starting from rough initial data by solutions starting from smooth and bounded initial conditions. This result allows us to pass from the weak comparison principle for $L^\infty(\mathbb{R}^d)$-valued initial data to that for rough initial data. In the second approximation, we mollify the noise and establish an uniform $L^2(\Omega)$-limit. See Section 4 for the proof.

THEOREM 1.9 (Two approximations). *Assume that f satisfies Dalang's condition* (1.3).

(1) *Suppose that the initial measure μ satisfies* (1.2). *If u and u_ε are the solutions to* (1.1) *starting from μ and $((\mu\psi_\varepsilon) * G(\varepsilon,\cdot))(x)$, respectively, where*

$$\psi_\varepsilon(x) = \mathbb{1}_{\{|x|\leq 1/\varepsilon\}} + \big(1 + 1/\varepsilon - |x|\big)\mathbb{1}_{\{1/\varepsilon<|x|\leq 1+1/\varepsilon\}}, \tag{1.16}$$

then

$$\lim_{\varepsilon\to 0_+} \|u(t,x) - u_\varepsilon(t,x)\|_2 = 0 \qquad \textit{for all } t > 0 \textit{ and } x \in \mathbb{R}^d.$$

(2) *Let ϕ be any continuous, nonnegative and nonnegative definite function on $\mathbb{R}^d$ with compact support such that $\int_{\mathbb{R}^d} \phi(x)\,\mathrm{d}x = 1$. Let u be the solution to* (1.1)

starting from bounded initial data, that is, $\mu(\mathrm{d}x) = g(x)\,\mathrm{d}x$ *with* $g \in L^\infty(\mathbb{R}^d)$. *If* $\tilde{u}_\varepsilon$ *is the solution to the following mollified equation*

$$\frac{\partial}{\partial t}\tilde{u}_\varepsilon(t,x) = \frac{1}{2}\Delta\tilde{u}_\varepsilon(t,x) + \rho\big(\tilde{u}_\varepsilon(t,x)\big)\dot{M}^\varepsilon(t,x), \tag{1.17}$$

with the same initial condition $\tilde{u}_\varepsilon(0,\cdot) = \mu$ *as* u, *where*

$$M^\varepsilon(\mathrm{d}s,\mathrm{d}x) = \int_{\mathbb{R}^d} \phi_\varepsilon(x-y) M(\mathrm{d}s,\mathrm{d}y)\,\mathrm{d}x, \tag{1.18}$$

and $\phi_\varepsilon(x) = \varepsilon^{-d}\phi(x/\varepsilon)$, *then*

$$\lim_{\varepsilon\to 0_+} \sup_{x\in\mathbb{R}^d} \big\| u(t,x) - \tilde{u}_\varepsilon(t,x)\big\|_2 = 0 \qquad \textit{for all } t > 0. \tag{1.19}$$

REMARK 1.10. One can always find one example of such function ϕ in part (2) of Theorem 1.9, for example, $\phi(x) = \prod_{i=1}^d (1-|x_i|)\mathbb{1}_{\{|x_i|\le 1\}}$ whose Fourier transform is nonnegative: $\hat{\phi}(\xi) = 2^d \prod_{j=1}^d \xi_j^{-2}(1-\cos(\xi_j)) \ge 0$.

The last result shows that the solution $u(t,x)$ to (1.1) converges to its initial data μ weakly as $t \to 0$. This result is used to establish the strong comparison principle for measure-valued initial data given that for function-valued initial data. See Section 5 for the proof. Let $C_c(\mathbb{R}^d)$ be the set of continuous functions with compact support.

THEOREM 1.11. *Under Dalang's condition* (1.3), *if* u *is the solution to* (1.1) *starting from a measure* μ *that satisfies* (1.2), *then, for all* $\phi \in C_c(\mathbb{R}^d)$,

$$\lim_{t\to 0}\int_{\mathbb{R}^d} u(t,x)\phi(x)\,\mathrm{d}x = \int_{\mathbb{R}^d}\phi(x)\mu(\mathrm{d}x) \qquad \textit{in } L^2(\Omega). \tag{1.20}$$

Finally, let us give some more explanations on the reason that we need to work under the stronger condition (1.4) instead of Dalang's condition (1.3). Actually, as one can see, Lemma 7.2 below will play a key role in the proof of the main result—the strong comparison principle. This lemma tells us that for small time step, that is, for large m, with high probability the solution in one time step will not change too much. (Then one can argue using the Markov property that if the initial data is positive somewhere, this property can be propagated to the whole space-time plane.) Hence, this kind of result (Lemma 7.2) has to do with the regularity of the solution. Indeed, the proof of Lemma 7.2, as one can see, consists of an optimization of two competing terms, one from the moment growth rate (Theorem 1.7) and the other from the Hölder continuity (Theorem 1.8). Under (1.3), the dependence on p in (1.14) is implicit, while under (1.4) it becomes explicit, and hence very easy to handle. However, this is not the reason why we assume (1.4). As shown in [6], under (1.3) alone one can construction a densely blow-up solution, that is, for any small time step, the solution may have a drastic change. To avoid such undesirable behavior, one has to use a stronger condition than Dalang's condition (1.3). Condition (1.4) turns out to be both general enough (which can cover the Riesz kernel case) and very convenient, and most of all, it guarantees a pathwise continuous solution.

1.3. *Outline of the paper.* This paper is organized as follows: We first prove the moment bounds, Theorem 1.7, in Section 2. Using these moment bounds, we proceed to establish the Hölder regularity, Theorem 1.8, in Section 3. Then in Section 4, we prove Theorem 1.9 for the two approximations. The weak limit as t goes to zero, that is, Theorem 1.11, is proved in Section 5. With this preparation, we prove the weak comparison principle, Theorem 1.3, in Section 6. Finally, in Section 7 we prove both the strong comparison principle (Theorem 1.5) and the strict positivity (Theorem 1.6). Some technical lemmas are given in the Appendix. Throughout this paper, C will denote a generic constant which may vary at each occurrence.

2. Moment bounds (Proof of Theorem 1.7). We first introduce some notation following [8]. Denote

$$k(t) := \int_{\mathbb{R}^d} f(z)G(t,z)\,\mathrm{d}z. \tag{2.1}$$

By the Fourier transform, this function can be written in the following form:

$$k(t) := (2\pi)^{-d}\int_{\mathbb{R}^d} \hat{f}(\mathrm{d}\xi)\exp\left(-\frac{t|\xi|^2}{2}\right). \tag{2.2}$$

Define $h_0(t) := 1$ and for $n \geq 1$,

$$h_n(t) = \int_0^t \mathrm{d}s\, h_{n-1}(s)k(t-s). \tag{2.3}$$

Let

$$H(t;\gamma) := \sum_{n=0}^{\infty} \gamma^n h_n(t) \qquad \text{for all } \gamma \geq 0. \tag{2.4}$$

This function is defined through the correlation function f. The following lemma tells us that this function has an exponential bound.

LEMMA 2.1 (Lemma 2.5 in [8] or Lemma 3.8 in [2]). *For all $t \geq 0$ and $\gamma \geq 0$, recalling that $\Upsilon(\beta)$ is defined in* (1.3), *it holds that*

$$\limsup_{t\to\infty} \frac{1}{t}\log H(t;\gamma) \leq \inf\left\{\beta > 0 : \Upsilon(2\beta) < \frac{1}{2\gamma}\right\}. \tag{2.5}$$

The following lemma will play a key role in our Picard iteration to obtain the upper bounds for the pth moment. Interested readers may want to compare it with Lemma A.1 below. While Lemma A.1 is appropriate for dealing with the two-point correlation function, the corresponding recursion for the p-point ($p > 2$) correlation function will be much more complicated. Instead if one only needs some upper bounds for the pth moment, the following lemma will do the job.

LEMMA 2.2. *Suppose that μ is a signed measure that satisfies* (1.2) *and recall that $J_0(t,x)$ is the solution to the homogeneous equation* (*see* (1.6)). *If a nonnegative* (*measurable*) *function $g : \mathbb{R}_+ \times \mathbb{R}^d \mapsto \mathbb{R}_+$ satisfies that for all $t > 0$ and $x \in \mathbb{R}^d$,*

$$\int_0^t \mathrm{d}s \iint_{\mathbb{R}^{2d}} G(t-s, x-y_1) G(t-s, x-y_2) \times f(y_1 - y_2) g(s, y_1) g(s, y_2) \,\mathrm{d}y_1 \,\mathrm{d}y_2 < +\infty$$

and

$$\begin{aligned} g(t,x)^2 \le J_0^2(t,x) + \lambda^2 \int_0^t \mathrm{d}s \iint_{\mathbb{R}^{2d}} & G(t-s, x-y_1) G(t-s, x-y_2) \\ & \times f(y_1 - y_2) g(s, y_1) g(s, y_2) \,\mathrm{d}y_1 \,\mathrm{d}y_2, \end{aligned} \tag{2.6}$$

then

$$g(t,x) \le \big(|\mu| * G(t,\cdot)\big)(x) H\big(t; 2\lambda^2\big)^{1/2}. \tag{2.7}$$

PROOF. We prove this lemma using Picard iteration. We need only to prove the case when the inequality in (2.2) is an equality. Let

$$g_0(t,x) = \big(|\mu| * G(t,\cdot)\big)(x),$$

and for $n \ge 1$,

$$\begin{aligned} g_n^2(t,x) = J_0^2(t,x) + \lambda^2 \int_0^t \mathrm{d}s \iint_{\mathbb{R}^{2d}} & G(t-s, x-y_1) G(t-s, x-y_2) \\ & \times g_{n-1}(s, y_1) g_{n-1}(s, y_2) f(y_1 - y_2) \,\mathrm{d}y_1 \,\mathrm{d}y_2. \end{aligned} \tag{2.8}$$

For $\gamma = 2\lambda^2$, we claim that

$$g_n(t,x) \le g_0(t,x) \left(\sum_{i=0}^n \gamma^i h_i(t) \right)^{1/2} \qquad \text{for all } n \ge 0. \tag{2.9}$$

It is clear that (2.9) holds for $n = 0$. Suppose that (2.9) is true for $n \ge 0$. Notice that

$$\begin{aligned} g_{n+1}^2(t,x) &= J_0^2(t,x) + \lambda^2 \int_0^t \iint_{\mathbb{R}^{2d}} G(t-s, x-y_1) G(t-s, x-y_2) f(y_1 - y_2) \\ &\quad \times g_n(s, y_1) g_n(s, y_2) \,\mathrm{d}s \,\mathrm{d}y_1 \,\mathrm{d}y_2 \\ &=: J_0^2(t,x) + \lambda^2 I(t,x). \end{aligned}$$

By the induction assumption,

$$
\begin{aligned}
I(t,x) \le{} & \int_0^t \mathrm{d}s \iint_{\mathbb{R}^{2d}} \mathrm{d}y_1\,\mathrm{d}y_2 f(y_1-y_2) G(t-s,x-y_1)G(t-s,x-y_2) \\
& \times \big|J_0(s,y_1)\big|\big|J_0(s,y_2)\big| \left(\sum_{i=0}^n \gamma^i h_i(s)\right) \\
={} & \int_0^t \mathrm{d}s \iint_{\mathbb{R}^{2d}} \mathrm{d}y_1\,\mathrm{d}y_2 f(y_1-y_2) G(t-s,x-y_1)G(t-s,x-y_2) \\
& \times \iint_{\mathbb{R}^{2d}} |\mu|(\mathrm{d}z_1)|\mu|(\mathrm{d}z_2) G(s,y_1-z_1)G(s,y_2-z_2)\left(\sum_{i=0}^n \gamma^i h_i(s)\right).
\end{aligned}
$$

Because (see [5], Lemma 5.4)

$$
G(s,x)G(t-s,y) = G\left(\frac{s(t-s)}{t}, \frac{sy-(t-s)x}{t}\right) G(t,x+y), \tag{2.10}
$$

we see that

$$
\begin{aligned}
& G(t-s,x-y_1)G(s,y_1-z_1) \\
& \quad = G(t,x-z_1) G\left(\frac{s(t-s)}{t}, y_1-z_1-\frac{s}{t}(x-z_1)\right).
\end{aligned}
$$

Hence,

$$
\begin{aligned}
I(t,x) \le{} & \int_0^t \mathrm{d}s \left(\sum_{i=0}^n \gamma^i h_i(s)\right) \iint_{\mathbb{R}^{2d}} \mathrm{d}y_1\,\mathrm{d}y_2 f(y_1-y_2) \\
& \times G\left(\frac{s(t-s)}{t}, y_1-z_1-\frac{s}{t}(x-z_1)\right) \\
& \times G\left(\frac{s(t-s)}{t}, y_2-z_2-\frac{s}{t}(x-z_2)\right) \\
& \times \iint_{\mathbb{R}^{2d}} |\mu|(\mathrm{d}z_1)|\mu|(\mathrm{d}z_2) G(t,x-z_1)G(t,x-z_2).
\end{aligned}
$$

By the Fourier transform, the above double integral $\mathrm{d}y_1\,\mathrm{d}y_2$ is equal to

$$
(2\pi)^{-d} \int_{\mathbb{R}^d} \hat{f}(\mathrm{d}\xi) \exp\left(i\frac{t-s}{t}(z_1-z_2)\cdot\xi - \frac{s(t-s)}{t}|\xi|^2\right).
$$

Since $\hat{f}$ is nonnegative, this integral is bounded by

$$
(2\pi)^{-d} \int_{\mathbb{R}^d} \hat{f}(\mathrm{d}\xi) \exp\left(-\frac{s(t-s)}{t}|\xi|^2\right).
$$

Hence,

$$
I(t,x) \le g_0^2(t,x) \int_0^t \mathrm{d}s \left(\sum_{i=0}^n \gamma^i h_i(s)\right) (2\pi)^{-d} \int_{\mathbb{R}^d} \hat{f}(\mathrm{d}\xi) \exp\left(-\frac{s(t-s)}{t}|\xi|^2\right).
$$

Then using the fact that $t \to h_i(t)$ is nondecreasing (see Lemma 2.6 in [8]), by Lemma B.1 with $\beta = |\xi|^2/2$, we see that

$$I(t,x) \le 2g_0^2(t,x) \int_0^t \mathrm{d}s \left(\sum_{i=0}^n \gamma^i h_i(s)\right) (2\pi)^{-d} \int_{\mathbb{R}^d} \hat{f}(\mathrm{d}\xi) \exp\left(-\frac{t-s}{2}|\xi|^2\right).$$

Then by (2.2) and (2.3), we see that

$$I(t,x) \le 2g_0^2(t,x) \int_0^t \mathrm{d}s \left(\sum_{i=0}^n \gamma^i h_i(s)\right) k(t-s) = 2g_0^2(t,x) \sum_{i=0}^n \gamma^i h_{i+1}(t).$$

Therefore,

$$g_{n+1}^2(t,x) \le g_0^2(t,x) + 2\lambda^2 g_0^2(t,x) \sum_{i=0}^n \gamma^i h_{i+1}(t) \le J_0^2(t,x) \sum_{i=0}^{n+1} \gamma^i h_i(t).$$

This proves (2.9). Finally,

$$g(t,x) \le \lim_{n\to\infty} g_0(t,x) \left(\sum_{i=0}^n \gamma^i h_i(t)\right)^{1/2} = g_0(t,x) \left(\sum_{i=0}^\infty \gamma^i h_i(t)\right)^{1/2},$$

which completes the proof of Lemma 2.2. □

PROOF OF THEOREM 1.7. The unique solution in $L^2(\Omega)$ has been established in [8]. We will prove the moment bounds in three steps.

Step 1. Now we prove this moment bound using Picard iteration. Let

$$u_0(t,x) = J_0(t,x),$$

and for $n \ge 1$,

$$u_n(t,x) = J_0(t,x) + \int_0^t \int_{\mathbb{R}^d} G(t-s,x-y)\rho\big(u_{n-1}(s,y)\big) M(\mathrm{d}s,\mathrm{d}y). \tag{2.11}$$

Since ρ is Lipschitz, by denoting $\overline{\varsigma} = |\rho(0)|/\operatorname{Lip}_\rho$,

$$\|\rho(X)\|_p \le \operatorname{Lip}_\rho \|\overline{\varsigma} + |X|\|_p \le \operatorname{Lip}_\rho \sqrt{2\big(\overline{\varsigma}^2 + \|X\|_p^2\big)}.$$

Because by the Burkholder–Davis–Gundy inequality and linear growth condition of ρ,

$$\begin{aligned}
&\overline{\varsigma}^2 + \|u_{n+1}(t,x)\|_p^2 \\
&\quad \le \overline{\varsigma}^2 + 2J_0^2(t,x) + 8p \int_0^t \iint_{\mathbb{R}^{2d}} G(t-s,x-y_1) G(t-s,x-y_2) \\
&\qquad \times f(y_1-y_2) \|\rho\big(u_n(s,y_1)\big)\|_p \|\rho\big(u_n(s,y_2)\big)\|_p \,\mathrm{d}s\,\mathrm{d}y_1\,\mathrm{d}y_2 \\
&\quad \le \overline{\varsigma}^2 + 2J_0^2(t,x) + 16p \operatorname{Lip}_\rho^2 \int_0^t \iint_{\mathbb{R}^{2d}} G(t-s,x-y_1) G(t-s,x-y_2) \\
&\qquad \times f(y_1-y_2) \sqrt{\overline{\varsigma}^2 + \|u_n(s,y_1)\|_p^2} \sqrt{\overline{\varsigma}^2 + \|u_n(s,y_2)\|_p^2} \,\mathrm{d}s\,\mathrm{d}y_1\,\mathrm{d}y_2,
\end{aligned}$$

we can apply the same induction arguments as those in the proof of Lemma 2.2 with $\lambda^2 = 16p\,\mathrm{Lip}_\rho^2$ and $g_n(t,x) = \sqrt{\overline{\varsigma}^2 + \|u_n(t,x)\|_p^2}$ and $J_0(t,x)$ replaced by $\overline{\varsigma} + \sqrt{2}J_0(t,x)$ to conclude that for all $n \geq 0$,

(2.12)

$$\begin{aligned}\|u_n(t,x)\|_p &\leq \sqrt{\overline{\varsigma}^2 + \|u_n(t,x)\|_p^2} \\ &\leq \sqrt{2}\big(\overline{\varsigma} + \sqrt{2}\big(|\mu| * G(t,\cdot)\big)(x)\big)\left(\sum_{i=0}^{n}(32p\,\mathrm{Lip}_\rho^2)^i h_i(t)\right)^{1/2}.\end{aligned}$$

Step 2. In this step, we will show that $\{u_n(t,x), n \in \mathbb{N}\}$ defined in (2.11) is a Cauchy sequence in $L^p(\Omega)$. Without loss of generality, we may assume that $\mu \geq 0$, otherwise one may simply replace μ by $|\mu|$ at each occurrence of μ. This will then imply the moment bound in (1.14). Denote

$$F_n(t,x) = \|u_{n+1}(t,x) - u_n(t,x)\|_p.$$

Then

$$\begin{aligned}F_n^2(t,x) \leq 8p\,\mathrm{Lip}_\rho^2 \int_0^t \mathrm{d}s \iint_{\mathbb{R}^{2d}} &\mathrm{d}y_1\,\mathrm{d}y_2 G(t-s,x-y_1)G(t-s,x-y_2) \\ &\times f(y_1-y_2)F_{n-1}(s,y_1)F_{n-1}(s,y_2),\end{aligned}$$

for $n \geq 1$, and

$$\begin{aligned}F_0^2(t,x) &= \|u_1(t,x) - J_0(t,x)\|_p^2 \\ &\leq 8p\,\mathrm{Lip}_\rho^2 \int_0^t \mathrm{d}s \iint_{\mathbb{R}^{2d}} \mathrm{d}y_1\,\mathrm{d}y_2 G(t-s,x-y_1)G(t-s,x-y_2) \\ &\quad \times f(y_1-y_2)J_0(s,y_1)J_0(s,y_2).\end{aligned}$$

Then by setting $F_{-1}(t,x) := J_0(t,x)$ and $\gamma = 16p\,\mathrm{Lip}_\rho^2$, one can apply the same induction arguments in the proof of Lemma 2.2 to conclude that

$$\sum_{n=0}^{\infty} F_n(t,x) \leq J_0(t,x)\left(\sum_{i=0}^{\infty}\gamma^i h_i(t)\right)^{1/2} < \infty.$$

Therefore, $\{u_n(t,x), n \in \mathbb{N}\}$ is a Cauchy sequence in $L^p(\Omega)$ and

$$\begin{aligned}\|u(t,x)\|_p &= \lim_{n\to\infty}\|u_n(t,x)\|_p \\ &\leq \lim_{n\to\infty}\big(\overline{\varsigma} + \sqrt{2}J_0(t,x)\big)\left(\sum_{i=0}^{n}(32p\,\mathrm{Lip}_\rho^2)^i h_i(t)\right)^{1/2} \\ &= \big(\overline{\varsigma} + \sqrt{2}J_0(t,x)\big)H\big(t;32p\,\mathrm{Lip}_\rho^2\big)^{1/2} < \infty.\end{aligned}$$

This proves (1.14).

Step 3. In this step, we will prove (1.15). Notice that in this case for $\beta > 0$,

$$\begin{aligned}\Upsilon(\beta) &= (2\pi)^{-d}\int_{\mathbb{R}^d}\frac{1}{(\beta+|\xi|^2)^\alpha}\frac{\hat{f}(\mathrm{d}\xi)}{(\beta+|\xi|^2)^{1-\alpha}}\\ &\le \frac{C}{\beta^\alpha}\left(\int_{|\xi|\le 1}\frac{\hat{f}(\mathrm{d}\xi)}{\beta^{1-\alpha}}+\int_{|\xi|>1}\frac{\hat{f}(\mathrm{d}\xi)}{|\xi|^{2(1-\alpha)}}\right)\\ &\le C\left(\frac{1}{\beta}+\frac{1}{\beta^\alpha}\right).\end{aligned}$$

From now on fix the constant C on the right-hand side of the above inequalities. If p is large enough such that $32p\,\mathrm{Lip}_\rho^2\,C > 1$, then

$$\begin{aligned}C\left(\frac{1}{\beta}+\frac{1}{\beta^\alpha}\right)\le\frac{1}{32p\,\mathrm{Lip}_\rho^2}\quad &\Longleftarrow\quad \frac{2C}{\beta^\alpha}\le\frac{1}{32p\,\mathrm{Lip}_\rho^2}\\ &\Longleftrightarrow\quad \beta\ge\left(C64p\,\mathrm{Lip}_\rho^2\right)^{1/\alpha}=:\beta_p.\end{aligned}$$

Then an application of Lemma 2.1 shows that

$$\limsup_{t\to\infty}\frac{1}{t}\log H\left(t;32p\,\mathrm{Lip}_\rho^2\right)\le\beta_p.$$

Hence, the function $e^{-\beta_p t}H(t;32p\,\mathrm{Lip}_\rho^2)$ is a continuous function on $[0,\infty]$. Therefore, for some constant $C'>0$, $e^{-\beta_p t}H(t;32p\,\mathrm{Lip}_\rho^2)\le C'$ for all $t\ge 0$. This proves (1.15) and also completes the whole proof of Theorem 1.7. □

3. Hölder regularity (Proof of Theorem 1.8). We first prove the following lemma.

LEMMA 3.1. *For all* $\alpha\in(0,1]$, $x,y\in\mathbb{R}^d$ *and* $t'\ge t>0$, *we have that*

$$\left|G(t,x)-G(t,y)\right|\le\frac{C}{t^{\alpha/2}}\left[G(2t,x)+G(2t,y)\right]|x-y|^\alpha \tag{3.1}$$

and

$$\left|G(t,x)-G\left(t',x\right)\right|\le Ct^{-\alpha/2}G\left(4t',x\right)\left(t'-t\right)^{\alpha/2}. \tag{3.2}$$

PROOF. By the scaling property, for (3.1), it suffices to prove that

$$\left|G(1,x)-G(1,y)\right|\le C\left[G(2,x)+G(2,y)\right]|x-y|^\alpha.$$

We may assume that $|x|\le|y|$. Choosing $\bar{x}\in\mathbb{R}^d$ such that $|\bar{x}|=|x|$ and $y=a\bar{x}$ for some $a\ge 1$, that is, $\bar{x}$, y and the origin are on the same line. By the mean value theorem, for some $c\in[0,1]$ and $\xi=c\bar{x}+(1-c)y$,

$$\begin{aligned}\left|G(1,x)-G(1,y)\right| &= \left|G(1,\bar{x})-G(1,y)\right|\\ &\le G(1,\xi)|\xi||\bar{x}-y|\\ &\le CG(2,\xi)|\bar{x}-y|.\end{aligned}$$

Then by the choice of $\bar{x}$ we see that

$$\begin{aligned} G(2,\xi)|\bar{x}-y| &\le C\big[G(2,\bar{x})+G(2,y)\big]|\bar{x}-y| \\ &\le C\big[G(2,x)+G(2,y)\big]|x-y|. \end{aligned}$$

Therefore,

$$\begin{aligned} &\big|G(1,x)-G(1,y)\big| \\ &\quad = \big|G(1,x)-G(1,y)\big|^{\alpha}\big|G(1,x)-G(1,y)\big|^{1-\alpha} \\ &\quad \le C\big[G(2,x)+G(2,y)\big]^{\alpha}|x-y|^{\alpha}\big|G(2,x)+G(2,y)\big|^{1-\alpha} \\ &\quad = C\big[G(2,x)+G(2,y)\big]|x-y|^{\alpha}, \end{aligned}$$

where we have applied the inequality that is just obtained to the factor $|G(1,x)-G(1,y)|^{\alpha}$ and we have used the fact $0<G(1,x)\le CG(2,x)$ for the other factor. This proves (3.1).

As for (3.2), notice that

$$\begin{aligned} &\big|G(t,x)-G(t',x)\big| \\ &\quad \le (2\pi)^{-d/2}\big|t^{-d/2}-(t')^{-d/2}\big|e^{-\frac{|x|^2}{2t}} + (2\pi)^{-d/2}(t')^{-d/2}\big|e^{-\frac{|x|^2}{2t}}-e^{-\frac{|x|^2}{2t'}}\big| \\ &\quad = t^{d/2}\big|t^{-d/2}-(t')^{-d/2}\big|G(t,x) + (t')^{-d/2}\left|G\left(1,\frac{x}{\sqrt{t}}\right)-G\left(1,\frac{x}{\sqrt{t'}}\right)\right|. \end{aligned}$$

For any $\gamma\in(0,1)$, because $t'>t$,

$$\begin{aligned} \big|t^{-d/2}-(t')^{-d/2}\big| &= \big|t^{-d/2}-(t')^{-d/2}\big|^{1-\gamma}\big|t^{-d/2}-(t')^{-d/2}\big|^{\gamma} \\ &\le C\big[2t^{-d/2}\big]^{1-\gamma}\big[\big(t^{-d/2-1}+(t')^{-d/2-1}\big)|t-t'|\big]^{\gamma} \qquad (3.3) \\ &\le Ct^{-d/2-\gamma}|t-t'|^{\gamma}. \end{aligned}$$

By (3.1), for all $\alpha\in(0,1]$,

$$\begin{aligned} &\left|G\left(1,\frac{x}{\sqrt{t}}\right)-G\left(1,\frac{x}{\sqrt{t'}}\right)\right| \\ &\quad \le C\left[G\left(2,\frac{x}{\sqrt{t}}\right)+G\left(2,\frac{x}{\sqrt{t'}}\right)\right]|x|^{\alpha}\big|t^{-1/2}-(t')^{-1/2}\big|^{\alpha} \\ &\quad \le CG\left(2,\frac{x}{\sqrt{t'}}\right)|x|^{\alpha}\big|t^{-1/2}-(t')^{-1/2}\big|^{\alpha} \\ &\quad = C(t')^{d/2}G(2t',x)|x|^{\alpha}\big|t^{-1/2}-(t')^{-1/2}\big|^{\alpha}. \end{aligned}$$

By the concavity of the square root, we see that

$$\big|t^{-1/2}-(t')^{-1/2}\big| = \frac{\sqrt{t'}-\sqrt{t}}{\sqrt{tt'}} \le \frac{\sqrt{t'-t}}{\sqrt{tt'}}.$$

Hence,

$$\left|G\left(1, \frac{x}{\sqrt{t}}\right) - G\left(1, \frac{x}{\sqrt{t'}}\right)\right| \le C t^{-\alpha/2} (t')^{(d-\alpha)/2} G(2t', x)|x|^{\alpha} (t'-t)^{\alpha/2}$$
$$\le C t^{-\alpha/2} (t')^{d/2} G(4t', x)(t'-t)^{\alpha/2}.$$

The bound in (3.2) is proved by taking $\gamma = \alpha/2$ in (3.3) and using the fact that $G(t,x) \le C G(4t', x)$. This completes the proof of Lemma 3.1. □

PROOF OF THEOREM 1.8. Denote the stochastic integral in (1.5) by $\mathcal{I}(t,x)$. Set $\overline{\varsigma} = |\rho(0)|/\operatorname{Lip}_\rho$. We need only to prove the Hölder regularity for $\mathcal{I}(t,x)$. Fix $n > 1$. For all (t,x) and $(t',x') \in [1/n, n] \times \mathbb{R}^d$ with $t' > t$, we see that

$$\|\mathcal{I}(t,x) - \mathcal{I}(t',x')\|_p^2 \le C I_1(t,x,x') + C I_2(t,t',x') + C I_3(t,t',x'),$$

where

$$I_1(t,x,x') = \int_0^t \mathrm{d}s \iint_{\mathbb{R}^{2d}} \mathrm{d}y_1\, \mathrm{d}y_2\, f(y_1 - y_2)$$
$$\times \left|G(t-s, x-y_1) - G(t-s, x'-y_1)\right| \sqrt{\overline{\varsigma}^2 + \|u(s,y_1)\|_p^2} \tag{3.4}$$
$$\times \left|G(t-s, x-y_2) - G(t-s, x'-y_2)\right| \sqrt{\overline{\varsigma}^2 + \|u(s,y_2)\|_p^2},$$

$$I_2(t,t',x') = \int_0^t \mathrm{d}s \iint_{\mathbb{R}^{2d}} \mathrm{d}y_1\, \mathrm{d}y_2\, f(y_1 - y_2)$$
$$\times \left|G(t-s, x'-y_1) - G(t'-s, x'-y_1)\right| \sqrt{\overline{\varsigma}^2 + \|u(s,y_1)\|_p^2} \tag{3.5}$$
$$\times \left|G(t-s, x'-y_2) - G(t'-s, x'-y_2)\right| \sqrt{\overline{\varsigma}^2 + \|u(s,y_2)\|_p^2}$$

and

$$I_3(t,t',x') = \int_t^{t'} \mathrm{d}s \iint_{\mathbb{R}^{2d}} \mathrm{d}y_1\, \mathrm{d}y_2\, f(y_1 - y_2)$$
$$\times G(t'-s, x'-y_1) \sqrt{\overline{\varsigma}^2 + \|u(s,y_1)\|_p^2} \tag{3.6}$$
$$\times G(t'-s, x'-y_2) \sqrt{\overline{\varsigma}^2 + \|u(s,y_2)\|_p^2}.$$

Note that when $\overline{\varsigma} \ne 0$, from the moment bounds in (1.14), by choosing

$$\widetilde{\mu}(\mathrm{d}x) = \sqrt{2}\mu(\mathrm{d}x) + \overline{\varsigma}\,\mathrm{d}x \quad \text{and} \quad \widetilde{J}_0(t,x) := \sqrt{2} J_0(t,x) + \overline{\varsigma},$$

one can reduce it to the case that $\overline{\varsigma} = 0$, that is, $\rho(0) = 0$. Hence, in the following, we only need to consider the case that $\overline{\varsigma} = 0$. We will study these three increments in three steps.

Step 1. In this step, we study I_1. We apply the moment bound (1.14) to (3.4), it follows that

$$I_1(t,x,x') \le C H(t,\gamma_p) \int_0^t \mathrm{d}s \iint_{\mathbb{R}^{2d}} \mathrm{d}y_1\,\mathrm{d}y_2\, f(y_1-y_2) \\ \times \big|G(t-s,x-y_1) - G(t-s,x'-y_1)\big| \\ \times \big|G(t-s,x-y_2) - G(t-s,x'-y_2)\big| \\ \times \iint_{\mathbb{R}^{2d}} \mu(\mathrm{d}z_1)\mu(\mathrm{d}z_2) G(s,y_1-z_1)G(s,y_2-z_2).$$

Here, we have used the definition of $J_0(t,x)$ and the fact that $H(s,\gamma_p)$ is nondecreasing in s; see Lemma 2.6 in [8]. By Lemma 3.1 and and (2.10), for all $\alpha \in (0,1)$,

$$\big|G(t-s,x-y_1) - G(t-s,x'-y_1)\big| \\ \le C\big[G(2(t-s),x-y_1) + G(2(t-s),x'-y_1)\big] \frac{|x-x'|^\alpha}{(t-s)^{\alpha/2}}$$

and

$$G(s,y_1-z_1)\big|G(t-s,x-y_1) - G(t-s,x'-y_1)\big| \\ \le C G(2s,y_1-z_1)\big[G(2(t-s),x-y_1) + G(2(t-s),x'-y_1)\big] \frac{|x-x'|^\alpha}{(t-s)^{\alpha/2}} \\ = C \frac{|x-x'|^\alpha}{(t-s)^{\alpha/2}} \bigg[G(2t,x-z_1) G\bigg(\frac{2s(t-s)}{t}, y_1 - z_1 - \frac{s}{t}(x-z_1)\bigg) \\ + G(2t,x'-z_1) G\bigg(\frac{2s(t-s)}{t}, y_1 - z_1 - \frac{s}{t}(x'-z_1)\bigg)\bigg].$$

A similar bound holds for the expression with respect to y_2 and z_2. Expanding the product of the two bounds, we will get a sum of four terms,

$$I_1(t,x,x') \le \sum_{k=1}^4 I_{1,k}(t,x,x')$$

where, for example,

$$I_{1,1}(t,x,x') \\ \le C|x-x'|^{2\alpha} \iint_{\mathbb{R}^{2d}} \mu(\mathrm{d}z_1)\mu(\mathrm{d}z_2) \int_0^t \mathrm{d}s \iint_{\mathbb{R}^{2d}} \mathrm{d}y_1\,\mathrm{d}y_2\, f(y_1-y_2) \frac{1}{(t-s)^\alpha} \\ \times G(2t,x-z_1) G\bigg(\frac{2s(t-s)}{t}, y_1 - z_1 - \frac{s}{t}(x-z_1)\bigg) \\ \times G(2t,x'-z_2) G\bigg(\frac{2s(t-s)}{t}, y_2 - z_2 - \frac{s}{t}(x-z_2)\bigg),$$

and similarly for $I_{1,i}$, $i = 2, 3, 4$. Because

$$\big|\mathcal{F}\big[G(t, \cdot + w)\big](\xi)\big| \le \exp\Big(-\frac{t}{2}|\xi|^2\Big) \qquad \text{for all } w \in \mathbb{R}^d,$$

we see that

$$\begin{aligned}
&I_{1,1}(t, x, x') \\
&\quad \le C|x - x'|^{2\alpha} \iint_{\mathbb{R}^{2d}} \mu(\mathrm{d}z_1)\mu(\mathrm{d}z_2) \int_0^t \mathrm{d}s \int_{\mathbb{R}^d} \hat{f}(\mathrm{d}\xi) \frac{1}{(t-s)^\alpha} \\
&\qquad \times G(2t, x - z_1)G(2t, x - z_2) \exp\Big(-\frac{2s(t-s)}{t}|\xi|^2\Big) \\
&\quad = C|x - x'|^{2\alpha} J_0(2t, x) J_0(2t, x') \int_0^t \mathrm{d}s \int_{\mathbb{R}^d} \hat{f}(\mathrm{d}\xi) \frac{\exp(-\frac{2s(t-s)}{t}|\xi|^2)}{(t-s)^\alpha}.
\end{aligned}$$

By Lemma B.1 with $g(s) = s^{-1/\alpha}$ and $\beta = |\xi|^2$ (g is nonincreasing),

$$\begin{aligned}
&\int_0^t \mathrm{d}s \int_{\mathbb{R}^d} \hat{f}(\mathrm{d}\xi) \frac{\exp(-\frac{2s(t-s)}{t}|\xi|^2)}{(t-s)^\alpha} \\
&\qquad \le 2 \int_0^t \mathrm{d}s \int_{\mathbb{R}^d} \hat{f}(\mathrm{d}\xi) \frac{1}{s^\alpha} \exp(-s|\xi|^2) \\
&\qquad \le 2e^t \int_0^t \mathrm{d}s \int_{\mathbb{R}^d} \hat{f}(\mathrm{d}\xi) \frac{1}{s^\alpha} \exp\big(-s(|\xi|^2 + 1)\big) \\
&\qquad \le C \int_{\mathbb{R}^d} \frac{\hat{f}(\mathrm{d}\xi)}{(1 + |\xi|^2)^{1-\alpha}}.
\end{aligned}$$

Hence,

$$I_{1,1}(t, x, x') \le C|x - x'|^{2\alpha} J_0(2t, x) J_0(2t, x') \int_{\mathbb{R}^d} \frac{\hat{f}(\mathrm{d}\xi)}{(1 + |\xi|^2)^{1-\alpha}}.$$

One can obtain similar bounds for all the other three terms. Therefore,

$$I_1(t, x, x') \le C|x - x'|^{2\alpha} \big[J_0(2t, x) + J_0(2t, x')\big]^2 \int_{\mathbb{R}^d} \frac{\hat{f}(\mathrm{d}\xi)}{(1 + |\xi|^2)^{1-\alpha}}.$$

Step 2. Now we consider the time increment I_2. By (1.14),

$$\begin{aligned}
&I_2(t, t', x') \\
&\quad \le C \int_0^t \mathrm{d}s \iint_{\mathbb{R}^{2d}} \mathrm{d}y_1\, \mathrm{d}y_2 \big|G(t - s, x' - y_1) - G(t' - s, x' - y_1)\big| \\
&\qquad \times \big|G(t - s, x' - y_2) - G(t' - s, x' - y_2)\big| f(y_1 - y_2) J_0(s, y_1) J_0(s, y_2) \\
&\quad = C \iint_{\mathbb{R}^d} \mu(\mathrm{d}z_1)\mu(\mathrm{d}z_2) \int_0^t \mathrm{d}s \iint_{\mathbb{R}^{2d}} \mathrm{d}y_1\, \mathrm{d}y_2\, f(y_1 - y_2)
\end{aligned}$$

$$\times G(s, y_1 - z_1)\big|G(t - s, x' - y_1) - G(t' - s, x' - y_1)\big|$$
$$\times G(s, y_2 - z_2)\big|G(t - s, x' - y_2) - G(t' - s, x' - y_2)\big|.$$

Applying (3.2), using the fact that $G(s, y_1 - z_1) \le CG(4s, y_1 - z_1)$ and then applying (2.10), we see that

$$\begin{aligned}
&G(s, y_1 - z_1)\big|G(t - s, x' - y_1) - G(t' - s, x' - y_1)\big| \\
&\quad \le C(t - s)^{-\alpha/2} G(s, y_1 - z_1) G\big(4(t' - s), x' - y_1\big)(t' - t)^{\alpha/2} \\
&\quad \le C(t - s)^{-\alpha/2} G(4s, y_1 - z_1) G\big(4(t' - s), x' - y_1\big)(t' - t)^{\alpha/2} \\
&\quad \le C(t - s)^{-\alpha/2} G\big(4t', x' - z_1\big) \\
&\qquad \times G\bigg(\frac{4s(t' - s)}{t'}, y_1 - z_1 - \frac{s}{t'}(x' - z_1)\bigg)(t' - t)^{\alpha/2}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
&I_2(t, t', x') \\
&\quad \le C(t' - t)^{\alpha} \iint_{\mathbb{R}^{2d}} \mu(\mathrm{d}z_1)\mu(\mathrm{d}z_2) G\big(4t', x' - z_1\big) G\big(4t', x' - z_2\big) \\
&\qquad \times \int_0^t \mathrm{d}s (t - s)^{-\alpha} \iint_{\mathbb{R}^{2d}} \mathrm{d}y_1\, \mathrm{d}y_2 f(y_1 - y_2) \\
&\qquad \times G\bigg(\frac{4s(t' - s)}{t'}, y_1 - z_1 - \frac{s}{t'}(x' - z_1)\bigg) \\
&\qquad \times G\bigg(\frac{4s(t' - s)}{t'}, y_2 - z_2 - \frac{s}{t'}(x' - z_2)\bigg) \\
&\quad \le C(t' - t)^{\alpha} \iint_{\mathbb{R}^{2d}} \mu(\mathrm{d}z_1)\mu(\mathrm{d}z_2) G\big(4t', x' - z_1\big) G\big(4t', x' - z_2\big) \\
&\qquad \times \int_0^t \mathrm{d}s (t - s)^{-\alpha} \int_{\mathbb{R}^d} \hat{f}(\mathrm{d}\xi) \exp\bigg(-\frac{4s(t - s)}{t}|\xi|^2\bigg) \\
&\quad = C(t' - t)^{\alpha} J_0^2\big(4t', x'\big) \int_{\mathbb{R}^d} \hat{f}(\mathrm{d}\xi) \\
&\qquad \times \int_0^t \mathrm{d}s (t - s)^{-\alpha} \exp\bigg(-\frac{4s(t - s)}{t}|\xi|^2\bigg),
\end{aligned}$$

where in the second inequality above we have used the fact that

$$\exp\bigg(-\frac{4s(t' - s)}{t'}|\xi|^2\bigg) \le \exp\bigg(-\frac{4s(t - s)}{t}|\xi|^2\bigg)$$

since $t' \geq t$. By the same arguments as those in Step 1,

$$
\begin{aligned}
I_2(t,t',x') &\leq C(t'-t)^{\alpha} J_0^2(4t',x') \int_{\mathbb{R}^d} \hat{f}(\mathrm{d}\xi) \int_0^t \mathrm{d}s s^{-\alpha} \exp(-2s|\xi|^2) \\
&\leq C(t'-t)^{\alpha} J_0^2(4t',x') \int_{\mathbb{R}^d} \hat{f}(\mathrm{d}\xi) \int_0^t \mathrm{d}s s^{-\alpha} \exp(-2s(1+|\xi|^2)) \\
&\leq C(t'-t)^{\alpha} J_0^2(4t',x') \int_{\mathbb{R}^d} \frac{\hat{f}(\mathrm{d}\xi)}{(1+|\xi|^2)^{1-\alpha}}.
\end{aligned}
$$

Step 3. As for I_3, by the moment bound (1.14) and (2.10),

$$
\begin{aligned}
&I_3(t,t',x') \\
&\quad\leq C \int_t^{t'} \mathrm{d}s \iint_{\mathbb{R}^{2d}} \mathrm{d}y_1\,\mathrm{d}y_2 G(t'-s,x'-y_1)G(t'-s,x'-y_2) \\
&\qquad \times f(y_1-y_2) J_0(s,y_1) J_0(s,y_2) \\
&\quad= C \iint_{\mathbb{R}^d} \mu(\mathrm{d}z_1)\mu(\mathrm{d}z_2) \int_t^{t'} \mathrm{d}s \iint_{\mathbb{R}^{2d}} \mathrm{d}y_1\,\mathrm{d}y_2 f(y_1-y_2) \\
&\qquad \times G(s,y_1-z_1)G(t'-s,x'-y_1)G(s,y_2-z_2)G(t'-s,x'-y_2) \\
&\quad= C \iint_{\mathbb{R}^d} \mu(\mathrm{d}z_1)\mu(\mathrm{d}z_2) G(t',x'-z_1)G(t',x'-z_2) \\
&\qquad \times \int_t^{t'} \mathrm{d}s \iint_{\mathbb{R}^{2d}} \mathrm{d}y_1\,\mathrm{d}y_2 f(y_1-y_2) \\
&\qquad \times G\left(\frac{s(t'-s)}{t'}, y_1-z_1-\frac{s}{t'}(x'-z_1)\right) \\
&\qquad \times G\left(\frac{s(t'-s)}{t'}, y_2-z_2-\frac{s}{t'}(x'-z_2)\right) \\
&\quad\leq C \iint_{\mathbb{R}^d} \mu(\mathrm{d}z_1)\mu(\mathrm{d}z_2) G(t',x'-z_1)G(t',x'-z_2) \int_t^{t'} \mathrm{d}s \int_{\mathbb{R}^d} \hat{f}(\mathrm{d}\xi) \\
&\qquad \times \exp\left(-\frac{s(t'-s)}{t'}|\xi|^2\right) \\
&\quad= C J_0^2(t',x') \int_t^{t'} \mathrm{d}s \int_{\mathbb{R}^d} \hat{f}(\mathrm{d}\xi) \exp\left(-\frac{s(t'-s)}{t'}|\xi|^2\right).
\end{aligned}
$$

Notice that for any $\alpha \in (0,1]$,

$$
\begin{aligned}
&\int_t^{t'} \mathrm{d}s \exp\left(-\frac{s(t'-s)}{t'}|\xi|^2\right) \\
&\qquad \leq \int_t^{t'} \mathrm{d}s \exp\left(-\frac{t(t'-s)}{t'}|\xi|^2\right)
\end{aligned}
$$

$$
\begin{aligned}
&\le \int_t^{t'} \mathrm{d}s \exp\Big(-\frac{t(t'-s)}{t'}(1+|\xi|^2) + \frac{t(t'-t)}{t'}\Big) \\
&\le C \int_t^{t'} \mathrm{d}s \exp\Big(-\frac{t(t'-s)}{t'}(1+|\xi|^2)\Big) \\
&= C \frac{1-\exp(-\frac{t(t'-t)}{t'}(1+|\xi|^2))}{1+|\xi|^2} \\
&\le C \frac{(\frac{t(t'-t)}{t'}(1+|\xi|^2))^\alpha}{1+|\xi|^2} \\
&= \frac{C(t'-t)^\alpha}{(1+|\xi|^2)^{1-\alpha}}.
\end{aligned}
$$

Therefore,

$$
I_3(t,t',x') \le C(t'-t)^\alpha J_0^2(t',x') \int_{\mathbb{R}^d} \frac{\hat{f}(\mathrm{d}\xi)}{(1+|\xi|^2)^{1-\alpha}}. \tag{3.7}
$$

Combining these three cases and applying the Kolmogorov's continuity theorem, we have completed the proof of Theorem 1.8. □

4. One approximation result (Proof of Theorem 1.9).

PROOF OF THEOREM 1.9. (1) By Theorem 1.2, we see that both u and u_ε are well-defined random field solutions to (1.1). Let $v_\varepsilon(t,x) = u_\varepsilon(t,x) - u(t,x)$ and $\tilde{\rho}(v_\varepsilon) := \rho(v_\varepsilon + u) - \rho(u)$. It is clear that $\tilde{\rho}$ is a Lipschitz continuous function satisfying $\tilde{\rho}(0) = 0$ and $\mathrm{Lip}_{\tilde{\rho}} = \mathrm{Lip}_\rho$. Then v_ε is a solution to (1.1) with ρ replaced by $\tilde{\rho}$ starting from $\mu_\varepsilon := ((\mu\psi_\varepsilon) * G(\varepsilon,\cdot))(x) - \mu$. Denote

$$
J_\varepsilon(t,x) = (\mu_\varepsilon * G(t,\cdot))(x) \quad \text{and} \quad g_\varepsilon(t,x,x') = \big|\mathbb{E}[v_\varepsilon(t,x)v_\varepsilon(t,x')]\big|.
$$

Then g satisfies the following integral equation:

$$
\begin{aligned}
g_\varepsilon(t,x,x') \le{}& \big|J_\varepsilon(t,x)J_\varepsilon(t,x')\big| \\
&+ \mathrm{Lip}_\rho^2 \int_0^t \mathrm{d}s \iint_{\mathbb{R}^{2d}} G(t-s,x-y)G(t-s,x'-y') \\
&\times f(y-y')g(s,y,y')\,\mathrm{d}y\,\mathrm{d}y'.
\end{aligned}
$$

By Lemma A.1, we see that

$$
\begin{aligned}
g_\varepsilon(t,x,x') \le{}& \big|J_\varepsilon(t,x)J_\varepsilon(t,x')\big| \\
&+ C \int_0^t \mathrm{d}s \iint_{\mathbb{R}^{2d}} G(t-s,x-y)G(t-s,x'-y') \\
&\times f(y-y')\big|J_\varepsilon(s,y')J_\varepsilon(s,y')\big|\,\mathrm{d}y\,\mathrm{d}y'.
\end{aligned}
$$

Notice that

$$\begin{aligned}\left|J_\varepsilon(t,x)\right| &\le \left[\left(|\mu\psi_\varepsilon| * \left|G(t+\varepsilon,\cdot)-G(t,\cdot)\right|\right)(x) + \left(|\mu\psi_\varepsilon-\mu| * G(t,\cdot)\right)(x)\right]\\ &\le \left[\left(|\mu| * \left|G(t+\varepsilon,\cdot)-G(t,\cdot)\right|\right)(x) + \left(|\mu\psi_\varepsilon-\mu| * G(t,\cdot)\right)(x)\right].\end{aligned}$$

Because for any $\varepsilon \in (0,t)$, $|G(t+\varepsilon,x)-G(t,x)| \le CG(2t,x)$ for all $x\in\mathbb{R}^d$ uniformly in ε, and because $|\mu\psi_\varepsilon-\mu|\le|\mu|$, we see that

$$\left|J_\varepsilon(t,x)\right| \le C\left(|\mu| * G(2t,\cdot)\right)(x) + \left(|\mu| * G(t,\cdot)\right)(x).$$

Then one can apply the dominated convergence theorem twice to conclude that

$$\lim_{\varepsilon\to 0} g_\varepsilon\left(t,x,x'\right)=0,$$

which completes the proof of part (1) of Theorem 1.9.

(2) Since u and $\tilde{u}_\varepsilon$ start from the same initial data, we see that

$$\begin{aligned}&\mathbb{E}\left[\left(u(t,x)-\tilde{u}_\varepsilon(t,x)\right)^2\right]\\ &\quad\le 2\mathbb{E}\left(\int_0^t\int_{\mathbb{R}^d} G(t-s,x-y)\left[\rho\left(u(s,y)\right)-\rho\left(\tilde{u}_\varepsilon(s,y)\right)\right]M(\mathrm{d}s,\mathrm{d}y)\right)^2\\ &\qquad + 2\mathbb{E}\left(\int_0^t\int_{\mathbb{R}^d} G(t-s,x-y)\rho\left(\tilde{u}_\varepsilon(s,y)\right)\left(M(\mathrm{d}s,\mathrm{d}y)-M^\varepsilon(\mathrm{d}s,\mathrm{d}y)\right)\right)^2\\ &\quad:= I_1(t,\varepsilon)+I_2(t,\varepsilon).\end{aligned}$$

For $I_1(t,\varepsilon)$, using the Lipschitz condition on ρ and since the initial condition is bounded, we obtain that

$$\begin{aligned}I_1(t,\varepsilon) &\le C\int_0^t\int_{\mathbb{R}^d} G\left(2(t-s),y\right)f(y)\sup_{z\in\mathbb{R}^d}\mathbb{E}\left[\left(u(s,z)-\tilde{u}_\varepsilon(s,z)\right)^2\right]\mathrm{d}y\,\mathrm{d}s\\ &= C\int_0^t \mathrm{d}s k\left(2(t-s)\right)\sup_{z\in\mathbb{R}^d}\mathbb{E}\left[\left(u(s,z)-\tilde{u}_\varepsilon(s,z)\right)^2\right],\end{aligned}$$

where $k(\cdot)$ function is defined in (2.1). As for $I_2(t,\varepsilon)$, we have that

$$\begin{aligned}&\mathbb{E}\left(\int_0^t\int_{\mathbb{R}^d} G(t-s,x-y)\rho\left(\tilde{u}_\varepsilon(s,y)\right)M(\mathrm{d}s,\mathrm{d}y)\right.\\ &\qquad\left.\times\int_0^t\int_{\mathbb{R}^d} G(t-s,x-y)\rho\left(\tilde{u}_\varepsilon(s,y)\right)M^\varepsilon(\mathrm{d}s,\mathrm{d}y)\right)\\ &\quad=\mathbb{E}\left(\int_0^t\int_{\mathbb{R}^d} G(t-s,x-y)\rho\left(\tilde{u}_\varepsilon(s,y)\right)M(\mathrm{d}s,\mathrm{d}y)\right.\\ &\qquad\left.\times\int_0^t\iint_{\mathbb{R}^{2d}} G(t-s,x-y)\rho\left(\tilde{u}_\varepsilon(s,y)\right)\phi_\varepsilon(y-z)M(\mathrm{d}s,\mathrm{d}z)\,\mathrm{d}y\right)\\ &\quad=\mathbb{E}\left(\int_0^t\int_{\mathbb{R}^d} G(t-s,x-y)\rho\left(\tilde{u}_\varepsilon(s,y)\right)M(\mathrm{d}s,\mathrm{d}y)\right.\end{aligned}$$

$$
\begin{aligned}
&\times \int_0^t \int_{\mathbb{R}^d} \Big(\int_{\mathbb{R}^d} G(t-s, x-y) \rho\big(\tilde{u}_\varepsilon(s,y)\big) \phi_\varepsilon(y-z)\, \mathrm{d}y \Big) M(\mathrm{d}s, \mathrm{d}z) \Big) \\
= {}& \mathbb{E}\Big(\int_0^t \mathrm{d}s \iint_{\mathbb{R}^{2d}} \mathrm{d}y_1\, \mathrm{d}y_2 G(t-s, x-y_1) \rho\big(\tilde{u}_\varepsilon(s,y_1)\big) \\
&\times G(t-s, x-y_2) \rho\big(\tilde{u}_\varepsilon(s,y_2)\big) \\
&\times \int_{\mathbb{R}^d} \mathrm{d}z \phi_\varepsilon(y_2 - z) f(y_1 - z) \Big) \\
= {}& \mathbb{E}\Big(\int_0^t \mathrm{d}s \iint_{\mathbb{R}^{2d}} \mathrm{d}y_1\, \mathrm{d}y_2 f^\varepsilon(y_1 - y_2) G(t-s, x-y_1) \rho\big(\tilde{u}_\varepsilon(s,y_1)\big) \\
&\times G(t-s, x-y_2) \rho\big(\tilde{u}_\varepsilon(s,y_2)\big) \Big),
\end{aligned}
$$

where we have applied the stochastic Fubini theorem and $f^\varepsilon(x) := (\phi_\varepsilon * f)(x)$. In the same way, we can get

$$
\begin{aligned}
&\mathbb{E}\Big[\Big(\int_0^t \int_{\mathbb{R}^d} G(t-s, x-y) \rho\big(\tilde{u}_\varepsilon(s,y)\big) M^\varepsilon(\mathrm{d}s, \mathrm{d}y) \Big)^2 \Big] \\
&\quad = \mathbb{E}\Big(\int_0^t \mathrm{d}s \iint_{\mathbb{R}^{2d}} \mathrm{d}y_1\, \mathrm{d}y_2 G(t-s, x-y_1) \rho\big(\tilde{u}_\varepsilon(s,y_1)\big) \\
&\qquad \times G(t-s, x-y_2) \rho\big(\tilde{u}_\varepsilon(s,y_2)\big) f^{\varepsilon,\varepsilon}(y_1 - y_2) \Big),
\end{aligned}
$$

where $f^{\varepsilon,\varepsilon}(x) := (\phi_\varepsilon * \phi_\varepsilon * f)(x)$. Since ϕ is nonnegative definite, the kernel function $f^{\varepsilon,\varepsilon}$ is nonnegative and nonnegative definite. Moreover, due to

$$
\hat{\phi}_\varepsilon(\xi)^2 = \hat{\phi}(\varepsilon\xi)^2 = \Big| \int_{\mathbb{R}^d} e^{-i\varepsilon\langle \xi, x\rangle} \phi(x)\, \mathrm{d}x \Big|^2 \le \Big(\int_{\mathbb{R}^d} \phi(x)\, \mathrm{d}x \Big)^2 = 1, \tag{4.1}
$$

$f^{\varepsilon,\varepsilon}$ satisfies Dalang's condition (1.3). From the above calculation, we see that the spatial correlation function for the noise M^ε is $f^{\varepsilon,\varepsilon}(x)$. Notice that

$$
\begin{aligned}
k_\varepsilon(t) &:= \int_{\mathbb{R}^d} f^{\varepsilon,\varepsilon}(z) G(t,z)\, \mathrm{d}z \\
&= (2\pi)^{-d} \int_{\mathbb{R}^d} \hat{f}(\mathrm{d}\xi) \hat{\phi}_\varepsilon(\xi)^2 \exp\Big(-\frac{t|\xi|^2}{2} \Big) \\
&\le (2\pi)^{-d} \int_{\mathbb{R}^d} \hat{f}(\mathrm{d}\xi) \exp\Big(-\frac{t|\xi|^2}{2} \Big) = k(t),
\end{aligned}
$$

for all $\varepsilon > 0$. Therefore, by Theorem 1.7,

$$
\sup_{\varepsilon>0} \sup_{(s,x)\in[0,t]\times\mathbb{R}^d} \big\| \tilde{u}_\varepsilon(s,x) \big\|_2 \le \sup_{(s,x)\in[0,t]\times\mathbb{R}^d} \big\| u(s,x) \big\|_2 < \infty.
$$

Thus,

$$
\begin{aligned}
I_2(t,\varepsilon) \le{}& C\int_0^t \iint_{\mathbb{R}^{2d}} G(t-s,x-y)G(t-s,x-z)\\
&\times \left|f(y-z)-2f^{\varepsilon}(y-z)+f^{\varepsilon,\varepsilon}(y-z)\right| \mathrm{d}y\,\mathrm{d}z\,\mathrm{d}s\\
={}& C\int_0^t \int_{\mathbb{R}^d} G\big(2(t-s),y\big)\left|f(y)-2f^{\varepsilon}(y)+f^{\varepsilon,\varepsilon}(y)\right| \mathrm{d}y\,\mathrm{d}s\\
\le{}& C\int_0^t \int_{\mathbb{R}^d} G\big(2(t-s),y\big)\left|f(y)-f^{\varepsilon}(y)\right| \mathrm{d}y\,\mathrm{d}s\\
&+C\int_0^t \int_{\mathbb{R}^d} G\big(2(t-s),y\big)\left|f(y)-f^{\varepsilon,\varepsilon}(y)\right| \mathrm{d}y\,\mathrm{d}s\\
={}& C\int_{\mathbb{R}^d} g\big(2t,|y|\big)\left|f(y)-f^{\varepsilon}(y)\right| \mathrm{d}y\\
&+C\int_{\mathbb{R}^d} g\big(2t,|y|\big)\left|f(y)-f^{\varepsilon,\varepsilon}(y)\right| \mathrm{d}y,
\end{aligned}
$$

where the function $g(t,|x|)$ is defined in Lemma B.4. Because f is nonnegative and

$$
\begin{aligned}
\int_{\mathbb{R}^d} g\big(4t,|y|\big)f(y)\,\mathrm{d}y &= \int_0^t \int_{\mathbb{R}^d} G(4s,y)f(y)\,\mathrm{d}y\,\mathrm{d}s\\
&= \int_0^t k(4s)\,\mathrm{d}s \le h_1(4t) < \infty,
\end{aligned}
$$

part (2) of Lemma B.5 implies that $\lim_{\varepsilon\to 0} I_2(t,\varepsilon)=0$. Hence an application of Gronwall's lemma shows that

$$
\lim_{\varepsilon\to 0_+} \sup_{z\in\mathbb{R}^d} \mathbb{E}\big[\big(u(t,z)-\tilde{u}_{\varepsilon}(t,z)\big)^2\big]=0,
$$

which completes the proof of Theorem 1.9. □

5. A weak limit (Proof of Theorem 1.11).

PROOF OF THEOREM 1.11. Fix $\phi\in C_c(\mathbb{R}^d)$. Let $I(t,x)$ be the stochastic integral part of (1.5). We only need to prove that

$$
\lim_{t\to 0_+} \int_{\mathbb{R}^d} \mathrm{d}x\, I(t,x)\phi(x)=0 \quad \text{in } L^2(\Omega).
$$

Denote $L(t):=\int_{\mathbb{R}} I(t,x)\phi(x)\,\mathrm{d}x$. By the stochastic Fubini theorem (see [25], Theorem 2.6, page 296),

$$
L(t)=\int_0^t \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} \mathrm{d}x\, G(t-s,x-y)\phi(x)\right)\rho\big(u(s,y)\big)M(\mathrm{d}s,\mathrm{d}y).
$$

Hence, by Itô's isometry and the linear growth condition on ρ,

$$\begin{aligned}\mathbb{E}\big[L(t)^2\big] \le \operatorname{Lip}_\rho^2 \int_0^t \mathrm{d}s \iint_{\mathbb{R}^{2d}} \mathrm{d}y_1\, \mathrm{d}y_2\, f(y_1-y_2) \iint_{\mathbb{R}^{2d}} \mathrm{d}x_1\, \mathrm{d}x_2 \\ \times \sqrt{\overline{\varsigma}^2 + \|u(s,y_1)\|_2^2}\, G(t-s,x_1-y_1)\big|\phi(x_1)\big| \\ \times \sqrt{\overline{\varsigma}^2 + \|u(s,y_2)\|_2^2}\, G(t-s,x_2-y_2)\big|\phi(x_2)\big|,\end{aligned}$$

where $\overline{\varsigma} = |\rho(0)|/\operatorname{Lip}_\rho$. Then by the moment bounds (1.14),

$$\begin{aligned}\mathbb{E}\big[L(t)^2\big] \le C \int_0^t \mathrm{d}s \iint_{\mathbb{R}^{2d}} \mathrm{d}y_1\, \mathrm{d}y_2\, f(y_1-y_2) \iint_{\mathbb{R}^{2d}} \mathrm{d}x_1\, \mathrm{d}x_2 \\ \times \sqrt{1 + J_0^2(s,y_1)}\, G(t-s,x_1-y_1)\big|\phi(x_1)\big| \\ \times \sqrt{1 + J_0^2(s,y_2)}\, G(t-s,x_2-y_2)\big|\phi(x_2)\big|.\end{aligned}$$

Assume that $t \le 1/2$. By considering $\mu_*(\mathrm{d}x) = \mu(\mathrm{d}x) + \mathrm{d}x$ and setting $J_*(t,x) = (\mu_* * G(t,\cdot))(x)$, we see that

$$1 + J_0^2(t,x) \le J_*^2(t,x).$$

Because for some constant $C > 0$, $|\phi(x)| \le C G(1,x)$ for all $x \in \mathbb{R}^d$, we can apply the semigroup property to get

$$\begin{aligned}\mathbb{E}\big[L(t)^2\big] \le C \int_0^t \mathrm{d}s \iint_{\mathbb{R}^{2d}} \mathrm{d}y_1\, \mathrm{d}y_2\, f(y_1-y_2) J_*(s,y_1) G(t+1-s,y_1) \\ \times J_*(s,y_2) G(t+1-s,y_2).\end{aligned}$$

Then by a similar argument as those in the proof of Lemma 2.2, we see that

$$\begin{aligned}\mathbb{E}\big[L(t)^2\big] &\le C J_*^2(t+1,x) \int_{\mathbb{R}^d} \hat{f}(\mathrm{d}\xi) \int_0^t \mathrm{d}s \exp\left(-\frac{s(t+1-s)}{t+1}|\xi|^2\right) \\ &\le C J_*^2(t+1,x) \int_{\mathbb{R}^d} \hat{f}(\mathrm{d}\xi) \int_0^t \mathrm{d}s \exp\left(-\frac{s}{2}|\xi|^2\right),\end{aligned}$$

where the last inequality is due to $t \le 1/2$. Since the above double integral is finite for $t = 1/2$, by the dominated convergence theorem, we see that this double integral goes to zero as $t \to 0$. This completes the proof. ☐

6. Weak comparison principle (Proof of Theorem 1.3).

PROOF OF THEOREM 1.3. We begin by noting that (1.9) is an immediate consequence of (1.8). So we only need to prove (1.8). The proof consists of four steps. Both the setup and Steps 1 and 4 of the proof follow the same lines as those in the proof of Theorem 1.1 in [7] with some minor changes. The main difference lies in Step 2 and Step 3.

Now we set up some notation in the proof. We view the $G(t,x)$ as an operator, denoted by $\mathbf{G}(t)$, as follows:

$$\mathbf{G}(t)f(x) := \big(G(t,\cdot) * f\big)(x). \tag{6.1}$$

Let $\mathbf{I}$ be the identity operator: $\mathbf{I}f(x) := (\delta * f)(x) = f(x)$. Set

$$\Delta^{\varepsilon} = \frac{\mathbf{G}(\varepsilon) - \mathbf{I}}{\varepsilon}. \tag{6.2}$$

Let

$$G^{\varepsilon}(t) = \exp\big(t\Delta^{\varepsilon}\big) = e^{-\frac{t}{\varepsilon}} \sum_{n=0}^{\infty} \frac{(t/\varepsilon)^n}{n!} \mathbf{G}(n\varepsilon) := e^{-t/\varepsilon}\mathbf{I} + \mathbf{R}^{\varepsilon}(t), \tag{6.3}$$

where the operator $\mathbf{R}^{\varepsilon}(t)$ has a density, denoted by $R^{\varepsilon}(t,x)$, which is equal to

$$R^{\varepsilon}(t,x) = e^{-t/\varepsilon} \sum_{n=1}^{\infty} \frac{(t/\varepsilon)^n}{n!} G(n\varepsilon, x). \tag{6.4}$$

For $\varepsilon > 0$ and $x \in \mathbb{R}^d$, denote

$$M_x^{\varepsilon}(t) = \int_0^t \int_{\mathbb{R}^d} G(\varepsilon, x-y) M(\mathrm{d}s, \mathrm{d}y) \qquad \text{for } t \geq 0. \tag{6.5}$$

Denote $\dot{M}_x^{\varepsilon}(t) = \frac{\partial}{\partial t} M_x^{\varepsilon}(t)$. Then the quadratic variation of $\mathrm{d}M_x^{\varepsilon}(t)$ is

$$\begin{aligned}
\mathrm{d}\big\langle M_x^{\varepsilon}(t)\big\rangle &= \iint_{\mathbb{R}^{2d}} G(\varepsilon, x-y_1) G(\varepsilon, x-y_2) f(y_1 - y_2)\, \mathrm{d}y_1\, \mathrm{d}y_2\, \mathrm{d}t \\
&= \int_{\mathbb{R}^d} e^{-\varepsilon|\xi|^2} \hat{f}(\mathrm{d}\xi)\, \mathrm{d}t.
\end{aligned}$$

Consider the following stochastic partial differential equation:

$$\begin{cases} \dfrac{\partial}{\partial t} u_{\varepsilon}(t,x) = \Delta^{\varepsilon} u_{\varepsilon}(t,x) + \rho\big(u_{\varepsilon}(t,x)\big)\dot{M}_x^{\varepsilon}(t), & t > 0, x \in \mathbb{R}^d, \\ u_{\varepsilon}(0,x) = \big(\mu * G(\varepsilon,\cdot)\big)(x), & x \in \mathbb{R}^d. \end{cases} \tag{6.6}$$

Since ρ is Lipschitz continuous and Δ^{ε} is a bounded operator, (6.6) has a unique strong solution

$$u_{\varepsilon}(t,x) = \big(\mu * G(\varepsilon,\cdot)\big)(x) + \int_0^t \mathrm{d}s\, \Delta^{\varepsilon} u_{\varepsilon}(s,x) + \int_0^t \rho\big(u_{\varepsilon}(s,x)\big)\, \mathrm{d}M_x^{\varepsilon}(s). \tag{6.7}$$

We proceed the proof in three steps. We fix $t > 0$ and assume that $\varepsilon \in (0, 1 \wedge t)$.

Step 1: Let $u_{\varepsilon,1}(t,x)$ and $u_{\varepsilon,2}(t,x)$ be the solutions to (6.6) with initial data $(\mu_1 * G(\varepsilon,\cdot))(x)$ and $(\mu_2 * G(\varepsilon,\cdot))(x)$, respectively. Following exactly the same lines as those in Step 1 of the proof in [7], we can prove that $v_{\varepsilon}(t,x) := u_{\varepsilon,2}(t,x) - u_{\varepsilon,1}(t,x)$ satisfies

$$\mathbb{P}\big(v_{\varepsilon}(t,x) \geq 0, \text{for every } t > 0 \text{ and } x \in \mathbb{R}^d\big) = 1. \tag{6.8}$$

Actually, one can construct a sequence of $C^2(\mathbb{R})$ functions Ψ_n as in [7] such that

$$\begin{aligned}&\Psi_n(x)\uparrow -(x\wedge 0)=:\Psi(x),\\ &\Psi_n'(x)x\uparrow \Psi(x),\qquad 0\le \Psi_n''(x)x^2\le 2/n.\end{aligned}\tag{6.9}$$

Then we apply Itô's formula to $\Psi_n(v_\varepsilon(t,x))$ and take the expectation on both sides to remove the martingale part. The third property in (6.9) ensures that the quadratic variation part goes to zero as $n\to\infty$. Using the other two properties in (6.9), we see that by passing to the limit, it holds that

$$\mathbb{E}\big[\Psi\big(v_\varepsilon(t,x)\big)\big]\le \frac{1}{\varepsilon}\int_0^t \mathrm{d}s\int_{\mathbb{R}^d}\mathrm{d}y G(\varepsilon,x-y)\mathbb{E}\big[\Psi\big(v_\varepsilon(s,y)\big)\big].$$

Then one can apply Gronwall's lemma to $\sup_{y\in\mathbb{R}^d}\mathbb{E}[\Psi(v_\varepsilon(s,y))]$ to conclude that $\mathbb{E}[\Psi(v_\varepsilon(t,x))]=0$ for all $t>0$ and $x\in\mathbb{R}^d$. This implies (6.8).

Step 2. In this step, we consider the case that the initial condition is bounded nonnegative function, that is, $\mu(\mathrm{d}x)=g(x)\,\mathrm{d}x$ where $g(x)\ge 0$ and $g\in L^\infty(\mathbb{R}^d)$. We also assume that the covariance function f satisfies condition (1.4) with $\alpha=1$, that is,

$$\int_{\mathbb{R}^d}\hat{f}(\mathrm{d}\xi)<\infty.$$

Let $u_\varepsilon(t,x)$ be the solution to (1.1) starting from $u_\varepsilon(0,x):=(\mu * G(\varepsilon,\cdot))(x)$. The aim of this step is to prove

$$\lim_{\varepsilon\to 0}\sup_{x\in\mathbb{R}^d}\big\|u_\varepsilon(t,x)-u(t,x)\big\|_2^2=0\qquad\text{for all } t>0.\tag{6.10}$$

Notice that $u_\varepsilon(t,x)$ can be written in the following mild form using the kernel of $\mathbf{G}^\varepsilon(t)$:

$$\begin{aligned}u_\varepsilon(t,x)&=\big(u_\varepsilon(0,\cdot)*G^\varepsilon(t,\cdot)\big)(x)+\int_0^t e^{-(t-s)/\varepsilon}\rho\big(u_\varepsilon(s,x)\big)\,\mathrm{d}M_x^\varepsilon(s)\\ &\quad+\int_0^t\int_{\mathbb{R}^d}R^\varepsilon(t-s,x-y)\rho\big(u_\varepsilon(s,y)\big)\,\mathrm{d}M_y^\varepsilon(s)\\ &=\big(u_\varepsilon(0,\cdot)*G^\varepsilon(t,\cdot)\big)(x)+\int_0^t e^{-(t-s)/\varepsilon}\rho\big(u_\varepsilon(s,x)\big)\,\mathrm{d}M_x^\varepsilon(s)\\ &\quad+\int_0^t\int_{\mathbb{R}^d}\bigg(\int_{\mathbb{R}^d}R^\varepsilon(t-s,x-z)\rho\big(u_\varepsilon(s,z)\big)G(\varepsilon,y-z)\,\mathrm{d}z\bigg)M(\mathrm{d}s,\mathrm{d}y).\end{aligned}$$

The boundedness of the initial data implies that

$$A_t:=\sup_{\varepsilon\in(0,1]}\sup_{s\in[0,t]}\sup_{x\in\mathbb{R}^d}\big\|u_\varepsilon(s,x)\big\|_2^2\vee\big\|u(s,x)\big\|_2^2<\infty.\tag{6.11}$$

By the assumption on ρ, we have the following estimate:

$$\|u_\varepsilon(t,x)-u(t,x)\|_2^2 \le C\sum_{n=1}^{6} I_n(t,x;\varepsilon),$$

where

$$I_1(t,x;\varepsilon) := \big(\big(u_\varepsilon(0,\cdot)*G^\varepsilon(t,\cdot)\big)(x) - u(0,\cdot)*G(t,\cdot)(x)\big)^2,$$
$$I_2(t,x;\varepsilon) := \int_0^t \mathrm{d}s \int_{\mathbb{R}^d} e^{-\varepsilon|\xi|^2} e^{-\frac{2(t-s)}{\varepsilon}} \hat{f}(\mathrm{d}\xi),$$

and $I_3(t,x;\varepsilon)$, $I_4(t,x;\varepsilon)$, $I_5(t,x;\varepsilon)$, $I_6(t,x;\varepsilon)$ are, respectively, equal to

$$\bigg\|\int_0^t\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} R^\varepsilon(t-s,x-z)\big[\rho\big(u_\varepsilon(s,z)\big)-\rho\big(u(s,z)\big)\big] \times G(\varepsilon,y-z)\,\mathrm{d}z M(\mathrm{d}s,\mathrm{d}y)\bigg\|_2^2,$$

$$\bigg\|\int_0^t\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} R^\varepsilon(t-s,x-z)\big[\rho\big(u(s,z)\big)-\rho\big(u(s,y)\big)\big] \times G(\varepsilon,y-z)\,\mathrm{d}z M(\mathrm{d}s,\mathrm{d}y)\bigg\|_2^2,$$

$$\bigg\|\int_0^t\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \big(R^\varepsilon(t-s,x-z)-G(t-s,x-z)\big)\rho\big(u(s,y)\big) \times G(\varepsilon,y-z)\,\mathrm{d}z M(\mathrm{d}s,\mathrm{d}y)\bigg\|_2^2,$$

$$\bigg\|\int_0^t\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \big(G(t-s,x-y)-G(t-s,x-z)\big)\rho\big(u(s,y)\big) \times G(\varepsilon,y-z)\,\mathrm{d}z M(\mathrm{d}s,\mathrm{d}y)\bigg\|_2^2.$$

Since μ has a bounded density, we see that

$$\begin{aligned} I_1(t,x;\varepsilon) &\le C\big|\big(u_\varepsilon(0,\cdot)*G^\varepsilon(t,\cdot)\big)(x) - \big(u(0,\cdot)*G(t,\cdot)\big)(x)\big| \\ &\le C\big(u_\varepsilon(0,\cdot)*\big|G^\varepsilon(t,\cdot)-G(t,\cdot)\big|\big)(x) \\ &\quad + C\big(u(0,\cdot)*\big|G(t+\varepsilon,\cdot)-G(t,\cdot)\big|\big)(x) \\ &\le C\bigg(e^{-t/\varepsilon} + \int_{\mathbb{R}^d}\big|R^\varepsilon(t,y)-G(t,y)\big|\,dy \\ &\quad + \int_{\mathbb{R}^d}\big|G(t+\varepsilon,y)-G(t,y)\big|\,dy\bigg). \end{aligned}$$

Then by Lemma B.3 and the fact that $\log(1+x) \le \sqrt{x}$, we see that

$$\sup_{x\in\mathbb{R}^d}\sup_{s\in(0,t]} I_1(s,x;\varepsilon) \le C\big(e^{-t/\varepsilon} + \sqrt{\varepsilon/t}\big). \tag{6.12}$$

As for I_2, we see that

$$\begin{aligned} I_2(t,x;\varepsilon) &= \int_{\mathbb{R}^d} e^{-\varepsilon|\xi|^2}\frac{\varepsilon}{2}\big(1-e^{-2t/\varepsilon}\big)\hat{f}(\mathrm{d}\xi) \\ &\le \frac{\varepsilon}{2}\int_{\mathbb{R}^d}\hat{f}(\mathrm{d}\xi) \le C\varepsilon, \end{aligned}$$

which implies that

$$\sup_{x\in\mathbb{R}^d}\sup_{s\in(0,t]} I_2(s,x;\varepsilon) \le C\varepsilon. \tag{6.13}$$

The term I_3 will contribute to the recursion. By (6.11),

$$\begin{aligned} &I_3(t,x;\varepsilon) \\ &\quad \le \mathbb{E}\bigg[\int_0^t \mathrm{d}s \iint_{\mathbb{R}^{2d}} \mathrm{d}y_1\,\mathrm{d}y_2\, f(y_1-y_2) \\ &\qquad \times \int_{\mathbb{R}^d} \mathrm{d}z_1\, R^\varepsilon(t-s,x-z_1)\big[\rho\big(u_\varepsilon(s,z_1)\big)-\rho\big(u(s,z_1)\big)\big]G(\varepsilon,y_1-z_1) \\ &\qquad \times \int_{\mathbb{R}^d} \mathrm{d}z_2\, R^\varepsilon(t-s,x-z_2)\big[\rho\big(u_\varepsilon(s,z_2)\big)-\rho\big(u(s,z_2)\big)\big]G(\varepsilon,y_2-z_2)\bigg] \\ &\quad \le C\int_0^t \mathrm{d}s \sup_{z\in\mathbb{R}^d}\big\|u_\varepsilon(s,z)-u(s,z)\big\|_2^2 \iint_{\mathbb{R}^{2d}} \mathrm{d}y_1\,\mathrm{d}y_2\, f(y_1-y_2) \\ &\qquad \times \big(R^\varepsilon(t-s,\cdot)*G(\varepsilon,\cdot)\big)(x-y_1)\big(R^\varepsilon(t-s,\cdot)*G(\varepsilon,\cdot)\big)(x-y_2) \\ &\quad \le C\int_0^t \mathrm{d}s \sup_{z\in\mathbb{R}^d}\big\|u_\varepsilon(s,z)-u(s,z)\big\|_2^2, \end{aligned}$$

where in the last line we used Lemma B.2. As for I_4,

$$\begin{aligned} &I_4(t,x;\varepsilon) \\ &\quad = \mathbb{E}\bigg[\int_0^t \mathrm{d}s \iint_{\mathbb{R}^{2d}} \mathrm{d}y_1\,\mathrm{d}y_2\, f(y_1-y_2) \\ &\qquad \times \bigg(\int_{\mathbb{R}^d} \mathrm{d}z_1\, R^\varepsilon(t-s,x-z_1)\big[\rho\big(u(s,z_1)\big)-\rho\big(u(s,y_1)\big)\big]G(\varepsilon,y_1-z_1)\bigg) \\ &\qquad \times \bigg(\int_{\mathbb{R}^d} \mathrm{d}z_2\, R^\varepsilon(t-s,x-z_2)\big[\rho\big(u(s,z_2)\big)-\rho\big(u(s,y_2)\big)\big]G(\varepsilon,y_2-z_2)\bigg)\bigg] \\ &\quad \le C\int_0^t \mathrm{d}s \iint_{\mathbb{R}^{2d}} \mathrm{d}y_1\,\mathrm{d}y_2\, f(y_1-y_2) \iint_{\mathbb{R}^{2d}} \mathrm{d}z_1\,\mathrm{d}z_2 \end{aligned}$$

$$\times R^{\varepsilon}(t-s,x-z_1)\|u(s,z_1)-u(s,y_1)\|_2 G(\varepsilon,y_1-z_1)$$
$$\times R^{\varepsilon}(t-s,x-z_2)\|u(s,z_2)-u(s,y_2)\|_2 G(\varepsilon,y_2-z_2).$$

Then by the Hölder continuity of u (see the proof of Theorem 1.8), we have that

$$\begin{aligned} I_4(t,x;\varepsilon) &\le C\int_0^t \mathrm{d}s \iint_{\mathbb{R}^{2d}} \mathrm{d}y_1\,\mathrm{d}y_2\, f(y_1-y_2) \iint_{\mathbb{R}^{2d}} \mathrm{d}z_1\,\mathrm{d}z_2 \\ &\qquad \times R^{\varepsilon}(t-s,x-z_1)|z_1-y_1| G(\varepsilon,y_1-z_1) \\ &\qquad \times R^{\varepsilon}(t-s,x-z_2)|z_2-y_2| G(\varepsilon,y_2-z_2) \\ &\le C\varepsilon\int_0^t \mathrm{d}s \iint_{\mathbb{R}^{2d}} \mathrm{d}y_1\,\mathrm{d}y_2\, f(y_1-y_2) \iint_{\mathbb{R}^{2d}} \mathrm{d}z_1\,\mathrm{d}z_2 \\ &\qquad \times R^{\varepsilon}(t-s,x-z_1)G(2\varepsilon,y_1-z_1)R^{\varepsilon}(t-s,x-z_2)G(2\varepsilon,y_2-z_2) \\ &\le C\varepsilon, \end{aligned}$$

where the last inequality is due to Lemma B.2 and the second inequality is due to the following inequality with $\alpha = 1$:

$$\begin{aligned} &|z_1-y_1|^{\alpha}|z_2-y_2|^{\alpha} G(\varepsilon,y_1-z_1)G(\varepsilon,y_2-z_2) \\ &\qquad \le C\varepsilon^{\alpha} G(2\varepsilon,y_1-z_1)G(2\varepsilon,y_2-z_2), \end{aligned} \tag{6.14}$$

for all $\alpha \in (0,1]$. Hence,

$$\sup_{x\in\mathbb{R}^d}\sup_{s\in[0,t]} I_4(s,x;\varepsilon) \le C\varepsilon. \tag{6.15}$$

Now let us consider I_5,

$$\begin{aligned} &I_5(t,x;\varepsilon) \\ &\quad = \mathbb{E}\bigg[\int_0^t \mathrm{d}s \iint_{\mathbb{R}^{2d}} \mathrm{d}y_1\,\mathrm{d}y_2\, f(y_1-y_2) \\ &\qquad \times \bigg(\int_{\mathbb{R}^d} \mathrm{d}z_1 \big(R^{\varepsilon}(t-s,x-z_1) - G(t-s,x-z_1)\big) \\ &\qquad \times \rho\big(u(s,y_1)\big) G(\varepsilon,y_1-z_1)\bigg) \\ &\qquad \times \bigg(\int_{\mathbb{R}^d} \mathrm{d}z_2 \big(R^{\varepsilon}(t-s,x-z_2) - G(t-s,x-z_2)\big) \\ &\qquad \times \rho\big(u(s,z_2)\big) G(\varepsilon,y_2-z_2)\bigg)\bigg] \\ &\quad \le C\int_0^t \mathrm{d}s \iint_{\mathbb{R}^{2d}} \mathrm{d}y_1\,\mathrm{d}y_2\, f(y_1-y_2) \iint_{\mathbb{R}^{2d}} \mathrm{d}z_1\,\mathrm{d}z_2 \\ &\qquad \times \big|R^{\varepsilon}(t-s,x-z_1) - G(t-s,x-z_1)\big| G(\varepsilon,y_1-z_1) \end{aligned}$$

$$
\begin{aligned}
&\times \left|R^\varepsilon(t-s,x-z_2)-G(t-s,x-z_2)\right|G(\varepsilon,y_2-z_2)\\
&\le C\int_0^t \mathrm{d}s \iint_{\mathbb{R}^{2d}} \mathrm{d}z_1\,\mathrm{d}z_2 \left|R^\varepsilon(s,z_1)-G(s,z_1)\right|\left|R^\varepsilon(s,z_2)-G(s,z_2)\right|\\
&\quad\times \iint_{\mathbb{R}^{2d}} \mathrm{d}y_1\,\mathrm{d}y_2 f(y_1-y_2)G(\varepsilon,y_1-x+z_1)G(\varepsilon,y_2-x+z_2)\\
&= C\int_0^t \mathrm{d}s \iint_{\mathbb{R}^{2d}} \mathrm{d}z_1\,\mathrm{d}z_2 \left|R^\varepsilon(s,z_1)-G(s,z_1)\right|\\
&\quad\times \left|R^\varepsilon(s,z_2)-G(s,z_2)\right| f_{2\varepsilon}(z_1-z_2),
\end{aligned}
$$

where $f_{2\varepsilon}(z)=(f*G(2\varepsilon,\cdot))(z)$. Hence,

$$
\begin{aligned}
I_5(t,x;\varepsilon) \le C\int_0^t \mathrm{d}s \int_{\mathbb{R}^d} &\mathrm{d}z_1\,\mathrm{d}z_2 \left|R^\varepsilon(s,z_1)-G(s,z_1)\right|\\
&\times \int_{\mathbb{R}} \mathrm{d}z_2 \big(R^\varepsilon(s,z_2)+G(s,z_2)\big) f_{2\varepsilon}(z_1-z_2).
\end{aligned}
$$

Notice that by the assumption of f in this step,

$$
\begin{aligned}
&\int_{\mathbb{R}^d} \big(R^\varepsilon(s,z_2)+G(s,z_2)\big) f_{2\varepsilon}(z_1-z_2)\,\mathrm{d}z_2\\
&\quad\le \int_{\mathbb{R}^d} \big(R^\varepsilon(s,z_2)+G(s,z_2)\big) f_{2\varepsilon}(z_2)\,\mathrm{d}z_2\\
&\quad= \int_{\mathbb{R}^d} \left(e^{-s/\varepsilon}\sum_{n=1}^\infty \frac{(s/\varepsilon)^n}{n!} e^{-\frac{n\varepsilon}{2}|\xi|^2} + e^{-\frac{s|\xi|^2}{2}}\right) e^{-\varepsilon|\xi|^2}\hat{f}(\mathrm{d}\xi) \le C.
\end{aligned}
$$

Thus, according to Lemma B.3, we have

$$
I_5(t,x;\varepsilon) \le C\int_0^t \left(e^{-s/\varepsilon}+\frac{\varepsilon^{1/2}}{s^{1/2}}\right) \le C\varepsilon^{1/2}.
$$

Thus,

$$
\sup_{x\in\mathbb{R}^d}\sup_{s\in(0,t]} I_5(s,x;\varepsilon) \le C\varepsilon^{1/2}. \tag{6.16}
$$

Now we study I_6. By Lemma 3.1,

$$
\begin{aligned}
&I_6(t,x;\varepsilon)\\
&\quad= \mathbb{E}\Big[\int_0^t \mathrm{d}s \iint_{\mathbb{R}^{2d}} \mathrm{d}y_1\,\mathrm{d}y_2 f(y_1-y_2)\rho\big(u(s,y_1)\big)\rho\big(u(s,y_2)\big)\\
&\qquad\times \left(\int_{\mathbb{R}^d} \mathrm{d}z_1 \big(G(t-s,x-z_1)-G(t-s,x-y_1)\big)G(\varepsilon,y_1-z_1)\right)\\
&\qquad\times \left(\int_{\mathbb{R}^d} \mathrm{d}z_1 \big(G(t-s,x-z_2)-G(t-s,x-y_2)\big)G(\varepsilon,y_2-z_2)\right)\Big]
\end{aligned}
$$

$$
\begin{aligned}
&\le C\int_0^t \mathrm{d}s \iint_{\mathbb{R}^{2d}} \mathrm{d}y_1\,\mathrm{d}y_2 f(y_1-y_2) \iint_{\mathbb{R}^{2d}} \mathrm{d}z_1\,\mathrm{d}z_2 \\
&\quad \times \left|G(t-s,x-z_1)-G(t-s,x-y_1)\right| G(\varepsilon,y_1-z_1) \\
&\quad \times \left|G(t-s,x-z_2)-G(t-s,x-y_2)\right| G(\varepsilon,y_2-z_2) \\
&\le C\int_0^t \mathrm{d}s \frac{1}{(t-s)^{1/2}} \iint_{\mathbb{R}^{2d}} \mathrm{d}y_1\,\mathrm{d}y_2 f(y_1-y_2) \iint_{\mathbb{R}^{2d}} \mathrm{d}z_1\,\mathrm{d}z_2 \\
&\quad \times |z_1-y_1|^{1/2}\left[G\big(2(t-s),x-z_1\big)+G\big(2(t-s),x-y_1\big)\right] G(\varepsilon,y_1-z_1) \\
&\quad \times |z_2-y_2|^{1/2}\left[G\big(2(t-s),x-z_2\big)+G\big(2(t-s),x-y_2\big)\right] G(\varepsilon,y_2-z_2).
\end{aligned}
$$

Then by (6.14) with $\alpha = 1/2$ and by the semigroup property,

$$
\begin{aligned}
I_6(t,x;\varepsilon) &\le C\varepsilon^{1/2}\int_0^t \mathrm{d}s \frac{1}{s^{1/2}} \iint_{\mathbb{R}^{2d}} \mathrm{d}y_1\,\mathrm{d}y_2 f(y_1-y_2) \iint_{\mathbb{R}^{2d}} \mathrm{d}z_1\,\mathrm{d}z_2 \\
&\quad \times \left[G(2s,x-z_1)+G(2s,x-y_1)\right] G(2\varepsilon,y_1-z_1) \\
&\quad \times \left[G(2s,x-z_2)+G(2s,x-y_2)\right] G(2\varepsilon,y_2-z_2) \\
&= C\varepsilon^{1/2}\int_0^t \mathrm{d}s \frac{1}{s^{1/2}} \iint_{\mathbb{R}^{2d}} \mathrm{d}y_1\,\mathrm{d}y_2 f(y_1-y_2) \\
&\quad \times G\big(2(s+\varepsilon),x-y_1\big) G\big(2(s+\varepsilon),x-y_2\big) \\
&\le C\varepsilon^{1/2}\int_0^\infty \mathrm{d}s \frac{1}{s^{1/2}} \int_{\mathbb{R}^d} e^{-2(s+\varepsilon)(|\xi|^2+1)} \hat{f}(\mathrm{d}\xi) \\
&\le C\varepsilon^{1/2}\int_{\mathbb{R}^d} \frac{\hat{f}(\mathrm{d}\xi)}{(1+|\xi|^2)^{1/2}} \le C\varepsilon^{1/2}.
\end{aligned}
$$

Thus,

$$
\sup_{x\in\mathbb{R}^d}\sup_{s\in[0,t]} I_6(s,x;\varepsilon) \le C\varepsilon^{1/2}. \tag{6.17}
$$

Therefore, by setting

$$
M(t;\varepsilon) := \sup_{y\in\mathbb{R}^d} \left\|u_\varepsilon(t,y)-u(t,y)\right\|_2^2,
$$

we have shown that

$$
M(t;\varepsilon) \le C\big(\varepsilon^{1/2} + e^{-t/\varepsilon} + \sqrt{\varepsilon/t}\big) + C\int_0^t M(s;\varepsilon)\,\mathrm{d}s.
$$

Then an application of Gronwall's lemma shows that

$$
M(t;\varepsilon) \le C\big(\varepsilon^{1/2} + e^{-t/\varepsilon} + \sqrt{\varepsilon/t}\big) + Ce^{Ct}\int_0^t \big(\varepsilon^{1/2} + e^{-s/\varepsilon} + \sqrt{\varepsilon/s}\big)\,\mathrm{d}s \to 0,
$$

as $\varepsilon \to 0$. This proves (6.10).

Step 3. In this step, we still work under the same assumption on the initial condition as in Step 2, that is, $\mu(\mathrm{d}x) = g(x)\,\mathrm{d}x$ with $g \geq 0$ and $g \in L^\infty(\mathbb{R}^d)$, but we assume that the covariance function f satisfies Dalang's condition (1.3). Choose a nonnegative and nonnegative definite function ϕ as in part (2) of Theorem 1.9 (see also Remark 1.10). Let u and u_ε be the solutions to (1.1) and (1.17), respectively, with the same initial data μ. From the proof of part (2) of Theorem 1.9, we see that the spatial covariance function for M^ε is $(f * \phi_\varepsilon * \phi_\varepsilon)(x)$. We claim that $(f * \phi_\varepsilon * \phi_\varepsilon)(x)$ satisfies (1.4) with $\alpha = 1$. Indeed, because $\phi(x) \leq CG(1, x)$, we have that $\phi_\varepsilon(x) \leq CG(\varepsilon^2, x)$ and

$$\begin{aligned}\int_{\mathbb{R}^d} \hat{f}(\mathrm{d}\xi)\hat{\phi}_\varepsilon(\xi)^2 &= C \iint_{\mathbb{R}^{2d}} f(x-y)\phi_\varepsilon(x)\phi_\varepsilon(y)\,\mathrm{d}x\,\mathrm{d}y \\ &\leq C \iint_{\mathbb{R}^{2d}} f(x-y)G\big(\varepsilon^2, x\big)G\big(\varepsilon^2, y\big)\,\mathrm{d}x\,\mathrm{d}y \\ &= C\int_{\mathbb{R}^d} f(y)G\big(2\varepsilon^2, y\big)\,\mathrm{d}y = Ck\big(2\varepsilon^2\big) < \infty,\end{aligned}$$

where $k(\cdot)$ is defined in (2.1). Hence, by Step 2, we see that

$$\mathbb{P}\big(u_\varepsilon(t,x) \geq 0\big) = 1 \qquad \text{for all } t > 0 \text{ and } x \in \mathbb{R}^d.$$

Part (2) of Theorem 1.9 implies that $u_\varepsilon(t,x)$ converges to $u(t,x)$ a.s., for each $t > 0$ and $x \in \mathbb{R}^d$. Therefore,

$$\mathbb{P}\big(u(t,x) \geq 0\big) = 1 \qquad \text{for all } t > 0 \text{ and } x \in \mathbb{R}^d.$$

Finally, suppose that $\mu_i(\mathrm{d}x) = g_i(x)\,\mathrm{d}x$ with $g_i \in L^\infty(\mathbb{R}^d)$, $i = 1, 2$. Let $u_{\varepsilon,i}$ be the solutions of (1.17) driven by M^ε and starting from initial conditions μ_i. If $g_1(x) \leq g_2(x)$ for almost all $x \in \mathbb{R}^d$, then by Step 1, $v_\varepsilon(t,x) := u_{\varepsilon,2}(t,x) - u_{\varepsilon,1}(t,x) \geq 0$ a.s. for all $t > 0$ and $x \in \mathbb{R}^d$. This step implies that $v_\varepsilon(t,x)$ converges to $v(t,x) = u_2(t,x) - u_1(t,x)$ in $L^2(\Omega)$ for all $t > 0$ and $x \in \mathbb{R}^d$. Therefore, $v(t,x)$ is nonnegative a.s., that is,

$$\mathbb{P}\big(u_1(t,x) \leq u_2(t,x)\big) = 1 \qquad \text{for all } t > 0 \text{ and } x \in \mathbb{R}^d.$$

Step 4. Now we assume that the initial data μ_1 and μ_2 are measures that satisfy (1.2). Recall the definition of ψ_ε in (1.16). For $\varepsilon > 0$, let $u_{\varepsilon,i}$, $i = 1, 2$, be the solutions to (1.1) starting from $([\mu_i\psi_\varepsilon] * G(\varepsilon, \cdot))(x)$. Denote $v(t,x) = u_2(t,x) - u_1(t,x)$ and $v_\varepsilon(t,x) = u_{\varepsilon,2}(t,x) - u_{\varepsilon,1}(t,x)$. Because ψ_ε is a continuous function with compact support on $\mathbb{R}$, the initial data for $u_{\varepsilon,i}(t,x)$ are bounded functions. By Step 3, we have that

$$\mathbb{P}\big(v_\varepsilon(t,x) \geq 0\big) = 1 \qquad \text{for all } t > 0, x \in \mathbb{R}^d \text{ and } \varepsilon > 0.$$

Then part (1) of Theorem 1.9 implies that

$$\mathbb{P}\big(v(t,x) \geq 0\big) = 1 \qquad \text{for all } t > 0 \text{ and } x \in \mathbb{R}^d.$$

This completes the whole proof of Theorem 1.3. □

7. Strong comparison principle and strict positivity (Proofs of Theorems 1.5 and 1.6). We need some lemmas. Denote $Q(r) = [-r, r]^d$, that is, a d-dimensional centered cube in $\mathbb{R}^d$ of radius r.

LEMMA 7.1. *Let $\ell > 0$. For all $t > 0$ and $M > 0$, there exists some constants $1 < m_0 = m_0(t, M) < \infty$ and $\gamma > 0$ such that for all $m \geq m_0$, $s \in [\frac{t}{2m}, \frac{t}{m}]$ and $x \in \mathbb{R}^d$,*

$$\big(G(s, \cdot) * \mathbb{1}_{Q(\ell)}\big)(x) \geq \gamma \mathbb{1}_{Q(\ell+\frac{M}{m})}(x). \tag{7.1}$$

PROOF. Since the d-dimensional heat kernel can be factored as a product of one-dimensional heat kernel, so the proof will be parallel with the proof of Lemma 4.1 in [7]. We will not repeat it here. □

LEMMA 7.2. *Let $\ell > 0$, $t > 0$, and $M > 0$. Assume that (1.4) holds for some $\alpha \in (0, 1]$. If $\rho(0) = 0$ and $\mu(\mathrm{d}x) = \mathbb{1}_{Q(\ell)}(x)\,\mathrm{d}x$, then there are some finite constants $\Theta := \Theta(\beta, \mathrm{Lip}_\rho, t) > 0$, $\beta > 0$ and $m_0 > 0$ such that for all $m \geq m_0$,*

$$\mathbb{P}\bigg(u(s, x) \geq \beta \mathbb{1}_{Q(\ell+\frac{M}{m})}(x) \text{ for all } \frac{t}{2m} \leq s \leq \frac{t}{m} \text{ and } x \in \mathbb{R}^d\bigg)$$
$$\geq 1 - \exp\big(-\Theta m^\alpha (\log m)^{1+\alpha}\big).$$

PROOF. This proof follows similar arguments as those in the proof of Lemma 4.3 in [7]. Here, we only give a sketch of it. Denote $S := S_{t,m,\ell,M} := \{(s, y) : \frac{t}{2m} \leq s \leq \frac{t}{m}, y \in Q(\ell + \frac{M}{m})\}$. By Lemma 7.1, for some $\beta > 0$,

$$\big(\mu * G(s, \cdot)\big)(x) \geq 2\beta \mathbb{1}_{Q(\ell+\frac{M}{m})}(x) \qquad \text{for all } s \in \left[\frac{t}{2m}, \frac{t}{m}\right] \text{ and } x \in \mathbb{R}^d. \tag{7.2}$$

Then the stochastic integral part $I(t, x)$ of the mild solution in (1.5) satisfies

$$\mathbb{P}\bigg(u(s, x) < \beta \mathbb{1}_{Q(\ell+\frac{M}{m})} \text{ for some } \frac{t}{2m} \leq s \leq \frac{t}{m} \text{ and } x \in \mathbb{R}^d\bigg)$$
$$\leq \mathbb{P}\big(I(s, x) < -\beta \text{ for some } (s, x) \in S\big)$$
$$\leq \mathbb{P}\Big(\sup_{(s,x)\in S} \big|I(s, x)\big| > \beta\Big) \leq \beta^{-p} \mathbb{E}\Big(\sup_{(s,x)\in S} \big|I(s, x)\big|^p\Big).$$

Denote $\tau = t/m$ and $S' := \{(s, y) : 0 \leq s \leq t/m, |y| \leq \ell + M/m\}$. Using the fact that $I(0, x) \equiv 0$ for all $x \in \mathbb{R}^d$, we see that for all $0 < \eta < 1 - \frac{6d}{\alpha p}$,

$$\mathbb{E}\bigg(\sup_{(s,x)\in S} \bigg|\frac{I(s, x)}{\tau^{\frac{\alpha\eta}{2}}}\bigg|^p\bigg) \leq \mathbb{E}\bigg(\sup_{(s,x),(s',x')\in S'} \bigg|\frac{I(s, x) - I(s', x')}{(|x - x'|^\alpha + |s - s'|^{\alpha/2})^\eta}\bigg|^p\bigg).$$

We are interested in, and hence assume in the following, the case when $p = O([m \log m]^\alpha)$ as $m \to \infty$; see (7.3) below. Since our initial condition is bounded, by (1.15), an application of the Kolmogorov's continuity theorem shows that for large p,

$$\beta^{-p}\mathbb{E}\Big(\sup_{(s,x)\in S}\big|I(s,x)\big|^p\Big) \le C\tau^{\frac{\alpha}{2}p\eta}e^{Cp^{\frac{\alpha+1}{\alpha}}\tau}$$
$$\le C\exp\Big(\frac{1}{2}\alpha p\eta\log(\tau) + Cp^{\frac{\alpha+1}{\alpha}}\tau\Big).$$

Since p is large, we may choose $\eta = 1/2$. Hence, the exponent in the right-hand side of the above inequalities becomes

$$f(p) := \frac{1}{4}\alpha p\log(\tau) + Cp^{\frac{\alpha+1}{\alpha}}\tau.$$

Some elementary calculation shows that $f(p)$ is minimized at

$$p = \Big(\frac{\alpha^2\log(1/\tau)}{4(\alpha+1)C\tau}\Big)^{\alpha} = \Big(\frac{\alpha^2 m\log(m/t)}{4(\alpha+1)Ct}\Big)^{\alpha}. \tag{7.3}$$

Hence, for some positive constants A and Θ,

$$\min_{p\ge 2} f(p) \le f(p') = -\Theta m^{\alpha}\big[\log(m)\big]^{1+\alpha} \qquad \text{with } p' = A\big[m\log(m)\big]^{\alpha}.$$

This completes the proof of Lemma 7.2. □

PROOF OF THEOREM 1.5. This proof follows the same arguments as those in the proof of Theorem 1.3 in [7]. Here, we only give a sketch of the proof. Interested readers are referred to [7] for details.

Let $u(t,x) := u_2(t,x) - u_1(t,x)$ and denote $\tilde{\rho}(u) = \rho(u+u_1) - \rho(u_1)$. Then it is not hard to see that $u(t,x)$ is a solution to (1.1) with the nonlinear function $\tilde{\rho}$ and the initial data $\mu := \mu_2 - \mu_1$. Note that $\tilde{\rho}$ is a Lipschitz continuous function with the same Lipschitz constant as for ρ and $\tilde{\rho}(0) = 0$. For simplicity, we will use ρ instead of $\tilde{\rho}$. By the weak comparison principle, we only need to consider the case when μ has compact support and show that $u(t,x) > 0$ for all $t > 0$ and $x \in \mathbb{R}^d$, a.s.

Case I. We fist assume that $\mu(\mathrm{d}x) = \mathbb{1}_{Q(\ell)}\,\mathrm{d}x$ for some $\ell > 0$. Denote

$$c(m) := \exp\big(-\Theta m^{\alpha}\big[\log(m)\big]^{1+\alpha}\big), \tag{7.4}$$

where Θ is a constant defined in Lemma 7.2. We comment that due to a version mismatch in [7], B_0 should be defined separately, that is,

$$A_k := \Big\{u(s,x) \ge \beta^{k+1}\mathbb{1}_{S_k^m}(x) \text{ for all } s \in \Big[\frac{(2k+1)t}{2m}, \frac{(k+1)t}{m}\Big] \text{ and } x \in \mathbb{R}^d\Big\},$$

for all $k \ge 0$,

$$B_k := \Big\{u(s,x) \ge \beta^{k+1}\mathbb{1}_{S_k^m}(x) \text{ for all } s \in \Big[\frac{kt}{m}, \frac{(2k+1)t}{2m}\Big] \text{ and } x \in \mathbb{R}^d\Big\},$$

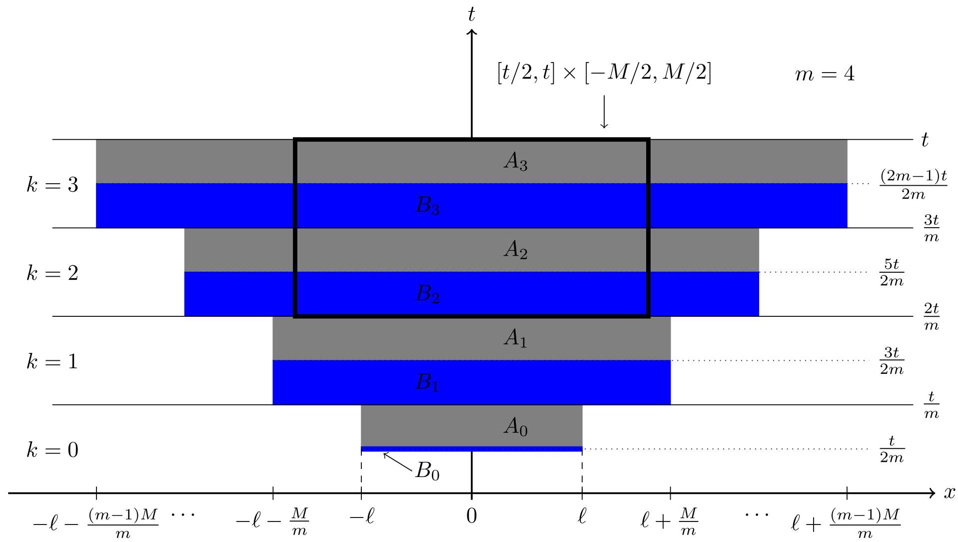

FIG. 1. *Induction schema for the strong comparison principle in the one-spatial dimension case.*

for all $k \geq 1$ and

$$B_0 := \left\{ u\left(\frac{t}{2m}, x\right) \geq \beta \mathbb{1}_{S_0^m}(x) \text{ for all } x \in \mathbb{R}^d \right\},$$

where

$$S_k^m := \left(-\ell - \frac{Mk}{m}, \ell + \frac{Mk}{m} \right).$$

See Figure 1 for an illustration of the schema.

By an argument using the strong Markov property, one can show that

$$\mathbb{P}(A_k \mid \mathcal{F}_{kt/m}) \geq 1 - c(m) \qquad \text{a.s. on } A_{k-1} \text{ for } 0 \leq k \leq m-1,$$

which implies

$$\mathbb{P}(A_k \mid A_{k-1} \cap \cdots \cap A_0) \geq 1 - c(m) \qquad \text{for all } 1 \leq k \leq m-1.$$

Notice that the fact that $A_0 \subseteq B_0$ implies that $\mathbb{P}(B_0) \geq \mathbb{P}(A_0) \geq 1 - c(m)$. By similar arguments as those for A_k, one can show that

$$\mathbb{P}(B_k \mid B_{k-1} \cap \cdots \cap B_0) \geq 1 - c(m) \qquad \text{for all } 1 \leq k \leq m-1.$$

Then

$$
\begin{aligned}
&\mathbb{P}\Big(\bigcap_{0\le k\le m-1}[A_k\cap B_k]\Big)\\
&\quad\ge 1-\Big(1-\mathbb{P}\Big(\bigcap_{0\le k\le m-1}A_k\Big)\Big)-\Big(1-\mathbb{P}\Big(\bigcap_{0\le k\le m-1}B_k\Big)\Big)\\
&\quad\ge \big(1-c(m)\big)^{m-1}\mathbb{P}(A_0)+\big(1-c(m)\big)^{m-1}\mathbb{P}(B_0)-1\\
&\quad\ge 2\big(1-c(m)\big)^m-1.
\end{aligned}
\tag{7.5}
$$

Therefore, for all $t>0$ and $M>0$,

$$
\begin{aligned}
&\mathbb{P}\big(u(s,x)>0 \text{ for all } t/2\le s\le t \text{ and } x\in Q(M/2)\big)\\
&\quad\ge \lim_{m\to\infty}\mathbb{P}\Big(\bigcap_{0\le k\le m-1}[A_k\cap B_k]\Big)\\
&\quad\ge \lim_{m\to\infty}2\big(1-c(m)\big)^m-1=1.
\end{aligned}
$$

Since t and M are arbitrary, this completes the proof for the case when $\mu(\mathrm{d}x)=\mathbb{1}_{Q(\ell)}\,\mathrm{d}x$.

Case II. Now for general initial data μ, we only need to prove that for each $\varepsilon>0$,

$$
\mathbb{P}\big(u(t,x)>0 \text{ for } t\ge\varepsilon \text{ and } x\in\mathbb{R}^d\big)=1. \tag{7.6}
$$

Fix $\varepsilon>0$. Denote $V(t,x):=u(t+\varepsilon,x)$. By the Markov property, $V(t,x)$ solves (1.1) with the time-shifted noise $\dot{M}_\varepsilon(t,x):=\dot{M}(t+\varepsilon,x)$ starting from $V(0,x)=u(\varepsilon,x)$, that is,

$$
\begin{aligned}
V(t,x)&=\big(u(\varepsilon,\circ)*G(t,\cdot)\big)(x)\\
&\quad+\iint_{[0,t]\times\mathbb{R}^d}\rho\big(V(s,y)\big)G(t-s,x-y)M_\varepsilon(\mathrm{d}s,\mathrm{d}y).
\end{aligned}
\tag{7.7}
$$

We first prove by contradiction that

$$
\mathbb{P}\big(u(\varepsilon,x)=0 \text{ for all } x\in\mathbb{R}^d\big)=0. \tag{7.8}
$$

Notice that by Theorem 1.8 the function $x\mapsto u(t,x)$ is Hölder continuous over $\mathbb{R}^d$ a.s. The weak comparison principle (Theorem 1.3) shows that $u(t,x)\ge 0$ a.s. Hence, if (7.8) is not true, then by the Markov property and the strong comparison principle in Case I, at all times $\eta\in[0,\varepsilon]$, with some strict positive probability, $u(\eta,x)=0$ for all $x\in\mathbb{R}^d$, which contradicts Theorem 1.11 as η goes to zero. Therefore, there exists a sample space Ω' with $\mathbb{P}(\Omega')=1$ such that for each $\omega\in\Omega'$, there exists $x\in\mathbb{R}^d$ such that $u(\varepsilon,x,\omega)>0$.

Since $u(\varepsilon, x, \omega)$ is continuous at x, one can find two nonnegative constants $c = c(\omega)$ and $\beta = \beta(\omega)$ such that $u(\varepsilon, y, \omega) \geq \beta \mathbb{1}_{x+Q(c)}(y)$ for all $y \in \mathbb{R}^d$. Then Case I implies that

$$\mathbb{P}\big(V_\omega(t, x) > 0 \text{ for all } t \geq 0 \text{ and } x \in \mathbb{R}^d\big) = 1,$$

where V_ω is the solution to (7.7) starting from $u(\varepsilon, x, \omega)$. Therefore, (7.6) is true. This completes the proof of Theorem 1.5. □

PROOF OF THEOREM 1.6. Following the proof of Theorem 1.5, since K is compact, we can choose $\eta, T, N > 0$ such that $K \subset [\eta, T] \times Q(N)$. Let β, A_k and B_k be as in the proof of Theorem 1.5, we have

$$\begin{aligned}\mathbb{P}\Big(\inf_{(t,x)\in K} u(t, x) < \beta^m\Big) &\leq 1 - \mathbb{P}\Big(\bigcap_{0\leq k\leq m-1} (A_k \cap B_k)\Big) \\ &\leq 2\big[1 - \big(1 - c(m)\big)^m\big],\end{aligned}$$

where $c(m)$ is a positive quantity defined in (7.4). Then we use the fact that $(1 - x)^m \geq 1 - mx$ for all $x > 0$ and $m > 1$ to conclude that for some Θ' slightly bigger than the Θ in (7.4),

$$\mathbb{P}\Big(\inf_{(t,x)\in K} u(t, x) < \beta^m\Big) \leq 2mc(m) \leq \exp\big(-\Theta' m^\alpha (\log m)^{1+\alpha}\big).$$

Finally, by taking $m = |\log \varepsilon|$, we complete the proof of Theorem 1.6. □

APPENDIX A: RECURSION ON THE TWO-POINT CORRELATION

We have encountered two types of recursions. One is (2.6), which is used in the proof of Theorem 1.7; the other is (A.4) below, which is used in the proof of Theorem 1.9. Lemma A.1 below is sharper than Lemma 2.2 and is used in [8] to obtain lower bounds for the second moment.

We need to introduce some notation. For $h, w : \mathbb{R}_+ \times \mathbb{R}^{3d} \mapsto \mathbb{R}$, define the (asymmetric convolution) operation "$\rhd$," which depends on f, as follows:

$$\begin{aligned}&(h \rhd w)\big(t, x, x'; y\big) \\ &\quad := \int_0^t \mathrm{d}s \iint_{\mathbb{R}^{2d}} \mathrm{d}z\, \mathrm{d}z' h\big(t - s, x - z, x' - z'; y - (z - z')\big) \\ &\qquad \times w\big(s, z, z'; y\big) f\big(y - (z - z')\big),\end{aligned} \tag{A.1}$$

or equivalently, by change of variables,

$$\begin{aligned}&(h \rhd w)\big(t, x, x'; y\big) \\ &\quad := \int_0^t \mathrm{d}s \iint_{\mathbb{R}^{2d}} \mathrm{d}z\, \mathrm{d}z' h\big(s, z, z'; y - [(x - z) - (x' - z')]\big) \\ &\qquad \times w\big(t - s, x - z, x' - z'; y\big) f\big(y - [(x - z) - (x' - z')]\big).\end{aligned} \tag{A.2}$$

This operation is associative (see Lemma B.1 in [8])

$$\big((h \rhd w) \rhd v\big)\big(t, x, x'; y\big) = \big(h \rhd (w \rhd v)\big)\big(t, x, x'; y\big).$$

We use the convention that for functions h defined on $\mathbb{R}_+ \times \mathbb{R}^{2d}$, when applying the operation $\rhd$ to h, it is meant for $h'(t, x, x'; y) := h(t, x, x')$.

For $t > 0$ and $x, x', y \in \mathbb{R}^d$, define recursively

$$\mathcal{L}_n\big(t, x, x'; y\big) := \begin{cases} G(t, x)G\big(t, x'\big) & \text{if } n = 0, \\ (\mathcal{L}_0 \rhd \mathcal{L}_{n-1})\big(t, x, x'; y\big) & \text{for } n \geq 1. \end{cases}$$

For $\lambda \in \mathbb{R}$, Lemma 2.7 of [8] ensures that the following series is well defined:

$$\mathcal{K}_\lambda\big(t, x, x'; y\big) := \sum_{n=0}^{\infty} \lambda^{2(n+1)} \mathcal{L}_n\big(t, x, x'; y\big) \leq \mathcal{L}_0\big(t, x, x'\big) H\big(t; 2\lambda^2\big). \tag{A.3}$$

Then the upper bounds for the two-point correlation function in Theorem 2.4 of [8] can be summarized as the following lemma.

LEMMA A.1. *Suppose that* $g : \mathbb{R}_+ \times \mathbb{R}^{2d} \mapsto \mathbb{R}$ *is some measurable function such that* $(\mathcal{L}_0 \rhd |g|)(t, x, x'; 0) < \infty$ *for all* $t > 0$ *and* $x, x' \in \mathbb{R}^d$. *If for some nonnegative function* $J_* : \mathbb{R}_+ \times \mathbb{R}^{2d} \mapsto \mathbb{R}_+$ *and* $\lambda \geq 0$, g *satisfies the following integral inequality*:

$$\begin{aligned} g\big(t, x, x'\big) \leq {} & J_*\big(t, x, x'\big) + \lambda^2 \int_0^t \mathrm{d}s \iint_{\mathbb{R}^{2d}} g(s, y_1, y_2) \\ & \times f(y_1 - y_2) G(t - s, x - y_1) G\big(t - s, x' - y_2\big) \, \mathrm{d}y_1 \, \mathrm{d}y_2, \end{aligned} \tag{A.4}$$

then

$$g\big(t, x, x'\big) \leq J_*\big(t, x, x'\big) + (\mathcal{K}_\lambda \rhd J_*)\big(t, x, x'; 0\big). \tag{A.5}$$

In particular,

$$\begin{aligned} g\big(t, x, x'\big) \leq {} & J_*\big(t, x, x'\big) + H\big(t; 2\lambda^2\big) \int_0^t \mathrm{d}s \iint_{\mathbb{R}^{2d}} J_*(s, y_1, y_2) \\ & \times G(t - s, x - y_1) G\big(t - s, x' - y_2\big) f(y_1 - y_2) \, \mathrm{d}y_1 \, \mathrm{d}y_2. \end{aligned} \tag{A.6}$$

If inequality (A.4) *is an equality, then* (A.5) *is also an equality.*

PROOF. This lemma is proved using Picard iteration. We first prove the case when the inequality (A.4) is an equality. Notice that (A.4) (with inequality replaced by equality) can be written as

$$g\big(t, x, x'\big) = J_*\big(t, x, x'\big) + \lambda^2 (\mathcal{L}_0 \rhd g)\big(t, x, x'; 0\big).$$

Let

$$g_n(t,x,x') = \begin{cases} J_*(t,x,x') & \text{if } n=0, \\ J_*(t,x,x') + \lambda^2(\mathcal{L}_0 \rhd g_{n-1})(t,x,x';0) & \text{for } n \geq 1. \end{cases} \tag{A.7}$$

Then by the associativity of the operator $\rhd$, we see that

$$g_n(t,x,x') = J_*(t,x,x') + \sum_{k=0}^{n-1} \lambda^{2(k+1)}(\mathcal{L}_k \rhd J_*)(t,x,x';0).$$

Therefore,

$$\begin{aligned} g(t,x,x') &= \lim_{n\to\infty} g_n(t,x,x') \\ &= J_*(t,x,x') + \sum_{k=0}^{\infty} \lambda^{2(k+1)}(\mathcal{L}_k \rhd J_*)(t,x,x';0) \\ &= J_*(t,x,x') + (\mathcal{K}_\lambda \rhd J_*)(t,x,x';0) \\ &\leq J_*(t,x,x') + H(t;2\lambda^2)(\mathcal{L}_0 \rhd J_*)(t,x,x';0), \end{aligned}$$

where the last step is due to (A.3). This proves the equality case.

We proceed to prove the inequality case. Let $g_*(t,x,x')$ be the solution to (A.4) with the inequality replaced by equality. Since g satisfies the inequality (A.4), by denoting $F(t,x,x') := g(t,x,x') - g_*(t,x,x')$, we need only show that $F \leq 0$. Notice that

$$F(t,x,x') \leq \lambda^2(\mathcal{L}_0 \rhd F)(t,x,x';0).$$

Apply the asymmetric convolution with respect to $\lambda^2\mathcal{L}_0$ on the both sides of the above inequality to see that

$$\lambda^2(\mathcal{L}_0 \rhd F)(t,x,x';0) \leq \lambda^4(\mathcal{L}_1 \rhd F)(t,x,x';0),$$

where we have used the associativity of $\rhd$ (see Lemma B.1 in [8]). Combining the above two inequalities, we see that

$$F(t,x,x') \leq \lambda^4(\mathcal{L}_1 \rhd F)(t,x,x';0).$$

In this way, one can show by induction that

$$F(t,x,x') \leq \lambda^{2(k+1)}(\mathcal{L}_k \rhd F)(t,x,x';0) \qquad \text{for all } k \in \mathbb{N}.$$

Now we are going to send k to $+\infty$. Because (see Lemma 2.7 of [8])

$$0 \leq \mathcal{L}_k(t,x,x';y) \leq 2^k h_k(t)\mathcal{L}_0(t,x,x'),$$

for all $t > 0$, $x, x', y \in \mathbb{R}^d$ and $k \in \mathbb{N}$, we see that

$$\left|\lambda^{2k}(\mathcal{L}_k \rhd F)(t,x,x';0)\right| \leq (2\lambda^2)^k h_k(t)(\mathcal{L}_0 \rhd |F|)(t,x,x';0). \tag{A.8}$$

By the integrability of g, $(\mathcal{L}_0 \rhd |F|)(t,x,x';0) < \infty$. Lemma 2.1 implies that $H(t;2\lambda^2) = \sum_{k=0}^{\infty}(2\lambda^2)^k h_k(t) < \infty$. Hence, the right-hand side of (A.8) goes to zero as $k \to \infty$. Therefore, $F(t,x,x') \le 0$, which completes the proof. $\square$

APPENDIX B: SOME TECHNICAL LEMMAS

In this section, we list some technical lemmas that are used in the paper.

LEMMA B.1. *If $g(t)$ is a monotone function over $[0,T]$, then for all $\beta > 0$ and $t \in (0,T]$,*

$$\int_0^t g(t-s)\exp\left(-\frac{2\beta s(t-s)}{t}\right)\mathrm{d}s = \int_0^t g(s)\exp\left(-\frac{2\beta s(t-s)}{t}\right)\mathrm{d}s \tag{B.1}$$

$$\le \begin{cases} 2\int_0^t g(s)e^{-\beta(t-s)}\,\mathrm{d}s & \text{if } g \text{ is nondecreasing,} \\ 2\int_0^t g(s)e^{-\beta s}\,\mathrm{d}s & \text{if } g \text{ is nonincreasing.} \end{cases} \tag{B.2}$$

PROOF. Equality (B.1) is clear by change of variables. We first assume that $g(t)$ is nondecreasing in $[0,T]$. Denote the integral by I. Then

$$\begin{aligned} I &= \int_0^{t/2} g(s)\exp\left(-\frac{2\beta s(t-s)}{t}\right)\mathrm{d}s + \int_{t/2}^t g(s)\exp\left(-\frac{2\beta s(t-s)}{t}\right)\mathrm{d}s \\ &\le \int_0^{t/2} g(s)\exp(-\beta s)\,\mathrm{d}s + \int_{t/2}^t g(s)\exp\big(-\beta(t-s)\big)\,\mathrm{d}s \\ &\le \int_{t/2}^t g(t-s)\exp\big(-\beta(t-s)\big)\,\mathrm{d}s + \int_{t/2}^t g(s)\exp\big(-\beta(t-s)\big)\,\mathrm{d}s \\ &\le 2\int_{t/2}^t g(s)\exp\big(-\beta(t-s)\big)\,\mathrm{d}s \\ &\le 2\int_0^t g(s)\exp\big(-\beta(t-s)\big)\,\mathrm{d}s. \end{aligned}$$

If g is nonincreasing in $[0,T]$, we simply replace the above $g(s)$ by $g(t-s)$ thanks to (B.1). This proves Lemma B.1. $\square$

LEMMA B.2. *Let R^ε be defined in (6.4). If f satisfies (1.4) with $\alpha = 1$, then there exists a constant $C > 0$ such that for all $0 \le s, \varepsilon \le t$ and $x \in \mathbb{R}^d$,*

$$\begin{aligned} &\iint_{\mathbb{R}^{2d}} \mathrm{d}y_1\,\mathrm{d}y_2 f(y_1-y_2)\big(R^\varepsilon(t-s,\cdot) * G(\varepsilon,\cdot)\big)(x-y_1) \\ &\quad\times \big(R^\varepsilon(t-s,\cdot) * G(\varepsilon,\cdot)\big)(x-y_2) \le C. \end{aligned}$$

PROOF. Denote the integral by I. Using the Fourier transform we have

$$I \le \int_{\mathbb{R}^d} e^{-\frac{2(t-s)}{\varepsilon}} \sum_{n,m=1}^{\infty} \frac{(\frac{t-s}{\varepsilon})^n}{n!} \frac{(\frac{t-s}{\varepsilon})^m}{m!} e^{-\frac{(n+m)\varepsilon}{2}|\xi|^2} \hat{f}(\mathrm{d}\xi)$$
$$\le C e^{-\frac{2(t-s)}{\varepsilon}} \sum_{n,m=1}^{\infty} \left(\frac{t-s}{\varepsilon}\right)^{m+n} \frac{1}{n!m!}.$$

Letting $n+m=k$ and using the fact that

$$\sum_{n=1}^{k-1} \frac{1}{n!(k-n)!} = \frac{1}{k!}(2^k - 2),$$

we see that the above double sum is equal to

$$\sum_{k=1}^{\infty} \sum_{n=1}^{k-1} \left(\frac{t-s}{\varepsilon}\right)^k \frac{1}{n!(k-n)!} \le \sum_{k=1}^{\infty} \left(\frac{t-s}{\varepsilon}\right)^k \frac{2^k}{k!} \le e^{\frac{2(t-s)}{\varepsilon}} - 1,$$

which proves Lemma B.2. □

LEMMA B.3. *There exists a finite constant $C > 0$ such that*

$$\int_{\mathbb{R}^d} \left| R^{\varepsilon}(t,x) - G(t,x) \right| \mathrm{d}x \le e^{-t/\varepsilon} + C \left(\frac{\varepsilon}{t}\right)^{1/2} \tag{B.3}$$

and

$$\int_{\mathbb{R}^d} \left| G(t+\varepsilon,x) - G(t,x) \right| \mathrm{d}x \le C \log\left(1 + \frac{\varepsilon}{t}\right), \tag{B.4}$$

for all $\varepsilon > 0$ and $t > 0$.

PROOF. Because $|\frac{\partial}{\partial t} G(t,x)| \le C t^{-1} G(2t,x)$, we see that for any t and t' such that $0 < t \le t'$,

$$\int_{\mathbb{R}^d} \left| G(t',x) - G(t,x) \right| \mathrm{d}x \le \int_{\mathbb{R}^d} \mathrm{d}x \int_t^{t'} \mathrm{d}s \left| \frac{\partial}{\partial s} G(s,x) \right|$$
$$\le C \int_{\mathbb{R}^d} \mathrm{d}x \int_t^{t'} \mathrm{d}s s^{-1} G(2s,x)$$
$$\le C \log(t'/t).$$

The rest of the proof will follow exactly the same lines as those in the proof of Lemma 8.2 in [7] and we will not repeat here. □

LEMMA B.4. *The function $g(t,x) := \int_0^t (2\pi s)^{-d/2} \exp(-\frac{x^2}{2s}) \,\mathrm{d}s$, for $t, x \ge 0$, satisfies the following properties*:

(1) *$x \mapsto g(t,x)$ is strictly decreasing functions on $x \in (0,\infty)$.*

(2) *If $d=1$, then $g(t,x)$ does not blow up at $x=0$ and $g(t,x) \le g(t,0) = \sqrt{2t/\pi}$. If $d \ge 2$, then $g(t,x)$ blows up at $x=0$.*

(3) *If $d=1,2$, then for all $\theta>0$ and $t>0$,*

$$\int_{\mathbb{R}^d} g\big(t,|x|\big)^\theta \,\mathrm{d}x < \infty. \tag{B.5}$$

(4) *If $d \ge 3$, then for all $0<\theta<\frac{d}{d-2}$ and $t>0$, (B.5) holds.*

PROOF. (1) It is clear $x \mapsto g(t,x)$ is nonincreasing on $(0,\infty)$ because

$$\frac{\partial}{\partial x} g(t,x) = -\int_0^t (2\pi s)^{-d/2} \frac{x}{s} \exp\left(-\frac{x^2}{2s}\right) \mathrm{d}s < 0 \qquad \text{for } x>0.$$

(2) If $d=1$, then by (1), we see that $g(t,x) \le g(t,0) = \sqrt{2t/\pi}$. By change of variables $z = x^2/(2s)$,

$$g(t,x) = \frac{1}{2\pi^{d/2}} x^{2-d} \int_{\frac{x^2}{2t}}^\infty e^{-z} z^{\frac{d}{2}-2} \,\mathrm{d}z. \tag{B.6}$$

If $d=2$, then the integral in (B.6) blows up as $x \to 0_+$. When $d \ge 3$,

$$g(t,x) \le \frac{1}{2\pi^{d/2}} x^{2-d} \int_0^\infty e^{-z} z^{\frac{d}{2}-2} \,\mathrm{d}z = \frac{\Gamma(d/2-1)}{2\pi^{d/2}} x^{2-d}, \tag{B.7}$$

which blows up as $x \to 0_+$.

(3) If $d=1$, for all $t>0$ and $x \ge 0$,

$$g(t,x) \le \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2t}} \int_0^t \frac{1}{\sqrt{s}} \,\mathrm{d}s = \frac{\sqrt{2t}}{\sqrt{\pi}} e^{-\frac{x^2}{2t}},$$

which shows (B.5) for $d=1$. If $d=2$, then

$$g(t,x) = \frac{1}{2\pi} \int_{x^2/(2t)}^\infty e^{-z} z^{-1} \,\mathrm{d}z.$$

Then by l'Hopital's rule,

$$\lim_{x\to 0_+} \frac{g(t,x)}{\log(1/x)} = \frac{1}{2\pi} \lim_{x \to 0_+} \frac{-e^{-\frac{x^2}{2t}} \frac{2t}{x^2} \frac{x}{t}}{-1/x} = \frac{1}{\pi}.$$

Hence, this case is proved by noting that for $x \ge 1$,

$$g(t,x) = \frac{1}{2\pi} \int_{\frac{x^2}{2t}}^\infty e^{-z} z^{-\frac{3}{2}} \,\mathrm{d}z \le \frac{1}{2\pi} \left(\frac{x^2}{2t}\right)^{-\frac{3}{2}} \int_{\frac{x^2}{2t}}^\infty e^{-z} \,\mathrm{d}z \le \frac{(2t)^{3/2}}{2\pi} e^{-\frac{x^2}{2t}}.$$

(4) For $d \ge 3$, note that there is a constant $C_d > 0$ which only depends on d such that $z^{\frac{d}{2}-2} e^{-z} \le C_d e^{-\frac{z}{2}}$ for all $z \ge 0$. Then for $x \ge 1$,

$$g(t,x) = \frac{1}{2\pi^{d/2}} x^{2-d} \int_{\frac{x^2}{2t}}^\infty e^{-z} z^{\frac{d}{2}-2} \,\mathrm{d}z \le \frac{C_d}{2\pi^{d/2}} \int_{\frac{x^2}{2t}}^\infty e^{-\frac{z}{2}} \,\mathrm{d}z \le \frac{C_d}{\pi^{d/2}} e^{-\frac{x^2}{4t}},$$

this shows that for any $\theta > 0$,

$$\int_{|x|\geq 1} g\big(t, |x|\big)^{\theta} \, \mathrm{d}x < \infty. \tag{B.8}$$

The restriction that $\theta < \frac{d}{d-2}$ comes from the integrability on $|x| \leq 1$, which is clear from the upper bound of $g(t, x)$ in (B.7). □

LEMMA B.5. *Recall the function $g(t, x)$ is defined in Lemma B.4. Let $\psi \in C_c(\mathbb{R}^d)$ be an arbitrary mollifier such that $\int_{\mathbb{R}^d} \psi(x)\,\mathrm{d}x = 1$. Denote $\psi_\varepsilon(x) = \varepsilon^{-d}\psi(x/\varepsilon)$. For each fixed $t > 0$, suppose that $h : \mathbb{R}^d \mapsto \mathbb{R}_+$ is a nonnegative and measurable function such that*

$$\int_{\mathbb{R}^d} h(x) g\big(2t, |x|\big) \, \mathrm{d}x < \infty.$$

Then the following statements hold:

(1) *For any $\eta > 0$, there exists $\phi \in C_c(\mathbb{R}^d)$ such that*

$$\sup_{\varepsilon \in (0, \sqrt{t})} \int_{\mathbb{R}^d} g_\varepsilon\big(t, |x|\big) \big|h(x) - \phi(x)\big| \, \mathrm{d}x < \eta,$$

where $g_\varepsilon(t, |x|) = \int_{\mathbb{R}^d} g(t, |y|)\psi_\varepsilon(x - y)\,\mathrm{d}y$.

(2) *By denoting $h_\varepsilon(x) = (h * \psi_\varepsilon)(x)$, we have that*

$$\lim_{\varepsilon \to 0} \int_{\mathbb{R}^d} g\big(t, |x|\big) \big|h(x) - h_\varepsilon(x)\big| \, \mathrm{d}x = 0.$$

PROOF. Without loss of generality, we may assume that $t = 1$.

(1) Fix $\eta > 0$. It is clear that for some constant $C > 0$, we have

$$\psi(x) \leq CG(1, x) \qquad \text{for all } x \in \mathbb{R}^d.$$

Hence, $\psi_\varepsilon(x) \leq CG(\varepsilon^2, x)$, which implies that

$$\begin{aligned}
g_\varepsilon\big(1, |x|\big) &\leq C \int_{\mathbb{R}^d} \mathrm{d}y G\big(\varepsilon^2, x - y\big) \int_0^1 \mathrm{d}s G(s, y) \\
&= C \int_0^1 \mathrm{d}s G\big(s + \varepsilon^2, x\big) \\
&= C \int_{\varepsilon^2}^{1+\varepsilon^2} \mathrm{d}s G(s, x) \leq C g\big(2, |x|\big),
\end{aligned} \tag{B.9}$$

where the last inequality is due to the definition of $g(t, x)$ and $\varepsilon \in (0, 1)$. Since h is nonnegative, it is known that one can find a monotone nondecreasing sequence

$\{s_j\}$ of simple functions such that $s_j(x) \uparrow h(x)$ pointwise; see, for example, Theorem 1.44 in [1]. Hence, by the dominated convergence theorem,

$$\sup_{\varepsilon\in(0,1)} \int_{\mathbb{R}^d} g_\varepsilon(1,|x|)\big|h(x)-s_j(x)\big|\,\mathrm{d}x < C\int_{\mathbb{R}^d} g(2,|x|)\big|h(x)-s_j(x)\big|\,\mathrm{d}x \to 0$$

as $j\to\infty$. Therefore, for some $s\in\{s_j\}$,

$$\sup_{\varepsilon\in(0,1)} \int_{\mathbb{R}^d} g_\varepsilon(1,|x|)\big|h(x)-s(x)\big|\,\mathrm{d}x \le \eta/2.$$

Now we choose and fix $q>1$ such that

$$C(g,d,q) := \int_{\mathbb{R}^d} g\big(t,|x|\big)^q\,\mathrm{d}x < \infty. \tag{B.10}$$

This is possible thanks to Lemma B.4: $q>1$ can be any number for $d=1,2$ and $q\in(1,\frac{d}{d-2})$ for $d\ge 3$. Since s is a simple function with bounded support, by Lusin's theorem (see, e.g., Theorem 1.42 (f) in [1]) there exists $\phi\in C_c(\mathbb{R}^d)$ such that

$$\big|\phi(x)\big| \le \|s\|_{L^\infty(\mathbb{R}^d)} \qquad \text{for all } x\in\mathbb{R}^d$$

and

$$\mathrm{Vol}\big(\{x\in\mathbb{R}^d : \phi(x)\neq s(x)\}\big) \le \eta^p\big(4C\|s\|_{L^\infty(\mathbb{R}^d)}C(g,d,q)^{1/q}\big)^{-p},$$

where $1/p+1/q=1$ and C is as in (B.9). Thus, using (B.9),

$$\begin{aligned}\sup_{\varepsilon\in(0,1)} &\int_{\mathbb{R}^d} g_\varepsilon(1,|x|)\big|s(x)-\phi(x)\big|\,\mathrm{d}x\\ &\le C\int_{\mathbb{R}^d} g(2,|x|)\big|s(x)-\phi(x)\big|\,\mathrm{d}x\\ &\le 2C\|s\|_{L^\infty(\mathbb{R}^d)}\int_{\mathbb{R}^d} \mathbb{1}_{\{x\in\mathbb{R}^d:\phi(x)\neq s(x)\}} g\big(2,|x|\big)\,\mathrm{d}x\\ &\le 2C\|s\|_{L^\infty(\mathbb{R}^d)}\Big(\int_{\mathbb{R}^d} \mathbb{1}_{\{x\in\mathbb{R}^d:\phi(x)\neq s(x)\}}\,\mathrm{d}x\Big)^{\frac{1}{p}}\Big(\int_{\mathbb{R}^d} g\big(2,|x|\big)^q\,\mathrm{d}x\Big)^{\frac{1}{q}}\\ &\le \frac{\eta}{2}.\end{aligned}$$

(2) For any $\eta>0$, we can write

$$\int_{\mathbb{R}^d}\big|h_\varepsilon(x)-h(x)\big|g\big(1,|x|\big)\,\mathrm{d}x = \int_{\mathbb{R}^d}\bigg|\int_{\mathbb{R}^d}\psi_\varepsilon(x-y)\big[h(y)-h(x)\big]\,\mathrm{d}y\bigg|g\big(1,|x|\big)\,\mathrm{d}x$$

$$
\begin{aligned}
&= \int_{\mathbb{R}^d} \left| \int_{\mathbb{R}^d} \psi_\varepsilon(x-y)\big[h(y)-\phi(y)\big]\,\mathrm{d}y \right| g\big(1,|x|\big)\,\mathrm{d}x \\
&\quad + \int_{\mathbb{R}^d} \left| \int_{\mathbb{R}^d} \psi_\varepsilon(x-y)\big[\phi(y)-\phi(x)\big]\,\mathrm{d}y \right| g\big(1,|x|\big)\,\mathrm{d}x \\
&\quad + \int_{\mathbb{R}^d} \left| \int_{\mathbb{R}^d} \psi_\varepsilon(x-y)\big[\phi(x)-h(x)\big]\,\mathrm{d}y \right| g\big(1,|x|\big)\,\mathrm{d}x \\
&=: I_1 + I_2 + I_3.
\end{aligned}
$$

For I_1, choose $\phi \in C_c(\mathbb{R}^d)$ according to (1), such that $I_1 < \frac{\eta}{3}$. From the proof of (1), it is obvious that with the same choice of ϕ, $I_3 < \frac{\eta}{3}$. For I_2, since ψ is compactly supported, we may choose $\varepsilon_0 > 0$ such that whenever $0 < \varepsilon < \varepsilon_0$, we have $I_2 < \frac{\eta}{3}$ because of the uniform continuity of ϕ. $\square$

Acknowledgments. The authors thank two anonymous referees for their careful reading of the paper and numerous valuable suggestions. Both authors appreciate some stimulating discussions with Davar Khoshnevisan. The first author also thanks Kunwoo Kim for many interesting discussions when both of them were in Utah in the year of 2014.

REFERENCES

[1] ADAMS, R. A. and FOURNIER, J. J. F. (2003). *Sobolev Spaces*, 2nd ed. *Pure and Applied Mathematics (Amsterdam)* **140**. Elsevier/Academic Press, Amsterdam. MR2424078

[2] BALAN, R. M. and CHEN, L. (2018). Parabolic Anderson model with space-time homogeneous Gaussian noise and rough initial condition. *J. Theoret. Probab.* **31** 2216–2265. MR3866613

[3] CARMONA, R. A. and MOLCHANOV, S. A. (1994). Parabolic Anderson problem and intermittency. *Mem. Amer. Math. Soc.* **108** viii+125. MR1185878

[4] CHEN, L. and DALANG, R. C. (2014). Hölder-continuity for the nonlinear stochastic heat equation with rough initial conditions. *Stoch. Partial Differ. Equ., Anal. Computat.* **2** 316–352. MR3255231

[5] CHEN, L. and DALANG, R. C. (2015). Moments and growth indices for the nonlinear stochastic heat equation with rough initial conditions. *Ann. Probab.* **43** 3006–3051. MR3433576

[6] CHEN, L., HUANG, J., KHOSHNEVISAN, D. and KIM, K. (2016). Dense blowup for parabolic SPDEs. Unpublished manuscript. Available at https://arxiv.org/abs/1702.08374.

[7] CHEN, L. and KIM, K. (2017). On comparison principle and strict positivity of solutions to the nonlinear stochastic fractional heat equations. *Ann. Inst. Henri Poincaré Probab. Stat.* **53** 358–388. MR3606745

[8] CHEN, L. and KIM, K. (2019). Nonlinear stochastic heat equation driven by spatially colored noise: Moments and intermittency. *Acta Math. Sci. Ser. B*. To appear. Available at arXiv:1510.06046.

[9] CONUS, D., JOSEPH, M. and KHOSHNEVISAN, D. (2012). Correlation-length bounds, and estimates for intermittent islands in parabolic SPDEs. *Electron. J. Probab.* **17** 15. MR3005720

[10] DALANG, R. C. (1999). Extending the martingale measure stochastic integral with applications to spatially homogeneous s.p.d.e.'s. *Electron. J. Probab.* **4** 29. MR1684157

[11] DALANG, R. C. and QUER-SARDANYONS, L. (2011). Stochastic integrals for spde's: A comparison. *Expo. Math.* **29** 67–109. MR2785545

[12] DAWSON, D. A. and SALEHI, H. (1980). Spatially homogeneous random evolutions. *J. Multivariate Anal.* **10** 141–180. MR0575923

[13] FOONDUN, M. and KHOSHNEVISAN, D. (2013). On the stochastic heat equation with spatially-colored random forcing. *Trans. Amer. Math. Soc.* **365** 409–458. MR2984063

[14] GUBINELLI, M. and PERKOWSKI, N. (2017). KPZ reloaded. *Comm. Math. Phys.* **349** 165–269. MR3592748

[15] HU, Y., HUANG, J. and NUALART, D. (2016). On the intermittency front of stochastic heat equation driven by colored noises. *Electron. Commun. Probab.* **21** 13. MR3485390

[16] HU, Y., HUANG, J., NUALART, D. and TINDEL, S. (2015). Stochastic heat equations with general multiplicative Gaussian noises: Hölder continuity and intermittency. *Electron. J. Probab.* **20** 50. MR3354615

[17] HUANG, J. (2017). On stochastic heat equation with measure initial data. *Electron. Commun. Probab.* **22** 6. MR3685238

[18] HUANG, J., LÊ, K. and NUALART, D. (2017). Large time asymptotics for the parabolic Anderson model driven by spatially correlated noise. *Ann. Inst. Henri Poincaré Probab. Stat.* **53** 1305–1340. MR3689969

[19] MORENO FLORES, G. R. (2014). On the (strict) positivity of solutions of the stochastic heat equation. *Ann. Probab.* **42** 1635–1643. MR3262487

[20] MUELLER, C. (1991). On the support of solutions to the heat equation with noise. *Stoch. Stoch. Rep.* **37** 225–245. MR1149348

[21] MUELLER, C. and NUALART, D. (2008). Regularity of the density for the stochastic heat equation. *Electron. J. Probab.* **13** 2248–2258. MR2469610

[22] SANZ-SOLÉ, M. and SARRÀ, M. (2002). Hölder continuity for the stochastic heat equation with spatially correlated noise. In *Seminar on Stochastic Analysis, Random Fields and Applications, III (Ascona*, 1999). *Progress in Probability* **52** 259–268. Birkhäuser, Basel. MR1958822

[23] SHIGA, T. (1994). Two contrasting properties of solutions for one-dimensional stochastic partial differential equations. *Canad. J. Math.* **46** 415–437. MR1271224

[24] TESSITORE, G. and ZABCZYK, J. (1998). Strict positivity for stochastic heat equations. *Stochastic Process. Appl.* **77** 83–98. MR1644618

[25] WALSH, J. B. (1986). An introduction to stochastic partial differential equations. In *École D'été de Probabilités de Saint-Flour, XIV*—1984. *Lecture Notes in Math.* **1180** 265–439. Springer, Berlin. MR0876085

DEPARTMENT OF MATHEMATICAL SCIENCES
UNIVERSITY OF NEVADA, LAS VEGAS
MAIL CODE: 4020
4505 S. MARYLAND PKWY.
LAS VEGAS, NEVADA 89154-4020
USA
E-MAIL: le.chen@unlv.edu
chenle02@gmail.com
URL: http://lechen.faculty.unlv.edu/

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF UTAH
155 S 1400 E
SALT LAKE CITY, UTAH 84112-0090
USA
E-MAIL: jhuang@math.utah.edu
URL: http://www.math.utah.edu/~jhuang/