TWO CONTRASTING PROPERTIES OF PARABOLIC SPDE ON METRIC MEASURE SPACES

LOUIS FAN, ZHENYAO SUN, AND JOHNNY YANG

Abstract. We consider a family of parabolic stochastic partial differential equations on metric measure spaces $(\mathbb{X}, d, m)$ equipped with a self-adjoint operator $\mathcal{L}$ on $L^2(\mathbb{X}, m)$ that generates a symmetric Markov process:

$$\partial_t u(t,x) = \mathcal{L}u(t,x) + b(u(t,x)) + \sigma(u_t(x))\dot{W}(t,x), \quad t > 0,\, x \in \mathbb{X},$$

where $\dot{W}$ is a Gaussian noise. Under certain regularity assumptions, we show 1) the solution has compact support if the noise term coefficient is degenerate and the initial condition decays sufficiently fast. 2) The strong comparison principle holds when the coefficients are Lipschitz.

1. Introduction

In this article, we focus on the parabolic-type stochastic partial differential equations (SPDEs) on general metric measure spaces. Let $(\mathbb{X}, d)$ be a locally compact separable metric space and m is a Radon measure on $\mathbb{X}$ with full support. Suppose $\mathcal{L}$ is a self-adjoint operator on $L^2(\mathbb{X}, m)$ that generates a $\mathbb{X}$-valued Markov process. We study the equation

$$\partial_t u(t,x) = \mathcal{L}u(t,x) + b(t,x,u(t,x)) + \sigma(t,x,u(t,x))\dot{W}(t,x), \quad t > 0,\, x \in \mathbb{X}. \tag{1.1}$$

for some measurable (random) functions b, $\sigma : \mathbb{R}_+ \times \mathbb{X} \times \mathbb{R} \to \mathbb{R}$ and a Gaussian noise $\dot{W}$.

In the case where $(\mathbb{X}, d) = (\mathbb{R}, |\cdot|)$ is the one dimensional Euclidean space, m is the Lebesgue's measure, $\mathcal{L} = \Delta$ is the usual Laplacian, $\dot{W}$ is the space time white noise and b, σ are deterministic functions and are independent of t and x. Then (1.1) with initial condition f is viewed as a short hand notation for the integral equation: for $t > 0$ and $x, y \in \mathbb{R}$,

$$u(t,x) = \int_{\mathbb{R}} p_t(x,y) f(y)\mathrm{d}y + \int_0^t \int_{\mathbb{R}} p_t(x,y) b(u(s,y))\mathrm{d}y\mathrm{d}s + \int_0^t \int_{\mathbb{R}} p_t(x,y)\sigma(u(s,y))W(\mathrm{d}y, \mathrm{d}s),$$

which is also called a mild solution, where for $t > 0$, $x, y \in \mathbb{R}$,

$$p_t(x,y) = \frac{1}{\sqrt{4\pi t}} e^{-\frac{|x-y|^2}{4t}}.$$

However, let us recall that the theory of stochastic integration with respect to a martingale measures in Walsh's notes [56] was constructed on Lusin spaces. On one hand, the Walsh's mild solutions on $\mathbb{R}$ involves only the heat kernel of Δ. On the other hand, by the theory of Dirichlet forms in [25], any m-symmetric strong Feller process on $\mathbb{X}$ admits a heat kernel, which has a corresponding $L^2(\mathbb{X}, m)$ generator $\mathcal{L}$. Therefore, it is natural to consider SPDEs with the operator $\mathcal{L}$ in the place of Δ on $\mathbb{X}$.

Indeed, there are various articles published in the past decade for SPDEs on general metric measure spaces. Besides the abstract functional analytic formulations ([21]), the earliest results traces back to the Dirichlet form approaches in [36], and the semi-group approaches with heat kernel estimates in [37, 40]. The well-posedness and regularity of solutions to parabolic-type SPDEs on post critical finite fractals were studied in [32, 33, 35]. There, the authors also investigated intermittency and invariant measures of solutions. The stochastic wave equations on fractals were studied in [34]. Cerrai and Freidlin initiated

2020 *Mathematics Subject Classification.* TBD.
Key words and phrases. TBD.

the study of SPDEs on metric graphs in [11, 12] via a scaling limit of narrowing of 2 dimensional domains (see also [14, 13]). Later, Fan showed in [23] one can obtain solutions to stochastic FKPP equations on metric graphs by a scaling limit of a system of interacting particle systems. Recently in [8], the authors obtained estimates on the intermittency properties of the parabolic Anderson model on a class of bounded recurrent metric measure spaces, which includes compact metric graphs and bounded Sierpiński gasket. We should also mention the following articles for singular SPDEs on non-flat spaces: [2, 50, 31].

Our main concerns in this article are the different comparison principles of (1.1). To better illustrate our point, let us go back and consider the one dimensional case. Mueller in [51] showed that, for the one dimensional heat equation with space time white noise, if

$$\sigma(u) = |u|^{\alpha}, \text{ for some } \alpha \geq 1 \quad \text{ and } \quad b(u) = 0, \quad \text{ for all } u \in \mathbb{R},$$

then the up to some (random) explosion time τ_σ, the solution u to the one-dimensional stochastic heat equation on $\mathbb{R}$, satisfies $\mathbb{P}(u_t(x) > 0, \text{ for all } t \in (0, \tau_\sigma), x \in \mathbb{R})$. Iscoe showed in [39] for stochastic heat equation on $\mathbb{R}$ with $b(u) = 0$ and $\sigma(u) = |u|^{\frac{1}{2}}$, if u is a solution to (1.1) so that u_0 has compact support, then almost surely for all $t \geq 0$, $u_t(\cdot)$ has compact support on $\mathbb{R}$. Both results were generalized by Shiga in [55], where he showed when b and σ are uniformly Lipschitz and $\dot{W}$ is the space time white noise, then the strong comparison principle holds, i.e. $\mathbb{P}(u_1(t,x) > u_2(t,x), \text{ for all } t > 0, x \in \mathbb{R}) = 1$ whenever $u_1(0,\cdot) - u_2(0,\cdot)$ is a non-negative and non-trivial function on $\mathbb{R}$. On the other hand, if for some $C > 0$,

$$|b(u)| \leq C|u| \text{ and } \sigma(0) = 0 \text{ with } \frac{1}{C}\sqrt{|u|} \leq |\sigma(u)| \text{ for all } u \in \mathbb{R}, \tag{1.2}$$

then the compact support property holds, that is, $\mathbb{P}(u_t(\cdot) \in \mathcal{C}_c^+(\mathbb{R}), \text{ for all } t > 0) = 1$ whenever $u_0 \in \mathcal{C}_c^+(\mathbb{R})$, where $\mathcal{C}_c^+(\mathbb{R})$ denote the collection of non-negative compactly supported functions on $\mathbb{R}$.

Since the well-posedness of (1.1) is already established in the above mentioned articles on some general metric measure spaces. It is then nature to ask if and when the strong comparison principle and compact support property hold for (1.1). To the best of our knowledge, neither properties was studied in such general settings before, not even the weak comparison principle (see Corollary 5.3). In this article, we answer both questions by providing easy-to-check sufficient conditions for both properties. We also give various concrete examples in Section 4.

Under certain regularity assumptions, our main results can be roughly stated as follows (see Theorem 3.1 and Theorem 3.6):

Informally Theorem (Compact support property)**.** *If the Markov process generated by $\mathcal{L}$ behaves like diffusion processes, then* (1.1) *has compact support property if condition* (1.2) *holds for some $C > 0$.*

Informally Theorem (Strong comparison principle)**.** *If the Markov process generated by $\mathcal{L}$ behaves either like a diffusion or stable process. Then the strong comparison principle holds for* (1.1) *provided b and σ are globally Lipschitz.*

Our results opens up many new and interesting questions. For example, consider the equation super-Brownian density on the unbounded Sierpiński gasket or a regular metric tree:

$$\partial_t u_t(x) = \mathcal{L}u_t(x) + \sqrt{u_t(x)}\dot{W}(x,t).$$

Since the compact support property holds (see Section 4), one may now ask how fast the support of the solution grow conditioned on non-extinction. On the other hand, the strong comparison principle opens up the possibility of solving the KPZ equation on fractals via Cole-Hopf transformation.

Let us briefly discuss our geometric settings (see Section 2 for more details). Our only explicit requirement is that, the metric space $(\mathbb{X}, d)$ have a metric entropy bounds (see (2.1)) which supports a Kolmogorov continuity type lemma. However, one quickly realizes that many implicit requirements of $\mathbb{X}$ are hidden in the heat kernel (see for example [26]). Hence, instead of making explicit assumptions on the geometry of $(\mathbb{X}, d, m)$ and the heat kernel bounds, we only make implicit assumptions of interactions between the space, heat kernels and the Gaussian noise. The advantage of such approach is that, it helps us to identify rather weak sufficient conditions for our results. The down side is that, each condition needs to be verified separately for different cases. However, We will demonstrate in Section 4.4 on how to verify our assumptions in terms of heat kernel bounds. Also, examples given in Section 4 are explicit space and with heat kernel bounds, includes metric graphs and certain fractals like Sierpiński gaske.

This paper is organized as follows. In section 2 we will introduce the ingredients of the equation (1.1): metric measure space, generator of Markov processes and Gaussian noises. We also give our regularity assumptions at the end of Section 2. We state our main results on the comparison principles in Sections 3, where we also provide easy-to-check conditions for the key requirements. Section 5 contains the well-posedness of (1.1) as well weak comparison principle and weak existence. Section 6 - 8 contains the proofs to the results listed. Section 9 contains the proof of miscellaneous results we used.

2. Setups

Throughout our presentations, let $(\mathbb{X}, d)$ be a locally compact separable complete metric space, and let m be a Radon measure on $\mathbb{X}$ with full support. Such a triple $(\mathbb{X}, d, m)$ is referred as a *metric measure space*. For $x \in \mathbb{X}$ and $r > 0$, denote by $B(x, r) = B_r(x) := \{y \in \mathbb{X} : d(x, y) < r\}$ the open ball with radius r centered at x.

Let $D(B(x, R), d; \varepsilon)$ be the smallest number of open balls of radius $\varepsilon > 0$ required to cover $B(x, R)$. We require $(\mathbb{X}, d)$ satisfies for some $b > 0$ and $R \geq 1$ that

$$D(B(x, R), d; \varepsilon) \lesssim \varepsilon^b \quad \text{uniformly in } x \in \mathbb{X}, \text{ and } 0 < \varepsilon \leq 1. \tag{2.1}$$

Here, for real-valued functions f and g defined on an abstract product space $X \times Y$, we say for every $y \in Y$, $f(x, y) \lesssim g(x, y)$ uniformly in $x \in X$, if for every $y \in Y$ there exists a constant $C(y) > 0$ such that for every $x \in X$, $f(x, y) \leq C(y)g(x, y)$. If $f \lesssim g$ and $g \lesssim f$, then we write $f \asymp g$.

In this paper, we assume $\mathcal{L}$ is the generator of a $\mathbb{X}$-valued Feller process, $(Y_t)_{t\geq 0}$, with transition function $(P_t)_{t\geq 0}$. To be more precise, we denote by Π_x the probability under which $Y_0 = x$ for every $x \in \mathbb{X}$. Let $\mathcal{C}_0(\mathbb{X})$ be the closure of the family of continuous functions with compact support on $\mathbb{X}$, $\mathcal{C}_c(\mathbb{X})$, in supremum norm. Then for any $f \in \mathcal{C}_0(\mathbb{X})$ and $x \in \mathbb{X}$, $P_t f \in \mathcal{C}_0(\mathbb{X})$ and $\lim_{t\to 0^+} P_t f(x) = f(x)$, where

$$P_t f(x) := \int_{\mathbb{X}} f(y) P_t(x, \mathrm{d}y) = \Pi_x [f(Y_t)], \quad f \in \mathcal{C}_c(\mathbb{X}), \, x \in \mathbb{X}.$$

$\mathcal{L}$ is defined via

$$\mathcal{L} f = \lim_{t \downarrow 0} \frac{P_t f - f}{t}, \quad \text{for } f \in \mathcal{C}_0(\mathbb{X}),$$

whenever the limit exists in supremum norm. We assume also that there is a measurable function $(p_t(x, y), \, t > 0, \, x, y \in \mathbb{X})$ that is symmetric in (x, y), for which we will call it the **heat kernel** of $\mathcal{L}$, so that for every $x \in \mathbb{X}$,

$$P_t f(x) = \Pi_x [f(Y_t)] = \int_{\mathbb{X}} p_t(x, y) f(y) m(\mathrm{d}y), \quad \text{for } f \in \mathcal{B}_b^+(\mathbb{X}). \tag{2.2}$$

Here, $\mathcal{B}_b^+(\mathbb{X})$ denote the collection of bounded non-negative measurable functions on $\mathbb{X}$.

Let $\dot{W}$ be a Gaussian noise on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ that satisfies the usual hypothesis, we will call $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}, W)$ a stochastic basis. Let b and σ be measurable functions on the product space $\mathbb{R}_+ \times \mathbb{X} \times \mathbb{R} \times \Omega$ w.r.t. the predictable σ-field generated by the filtration $(\mathcal{F}_t)_{t\geq 0}$. For the simplicity of notations, we sometimes write $b(t, x, u)$ and $\sigma(t, x, u)$ for $b(t, x, u, \omega)$ and $\sigma(t, x, u, \omega)$. We consider the equation

$$\partial_t u_t(x) = \mathcal{L} u_t(x) + b(t, x, u_t(x)) + \sigma(t, x, u_t(x)) \dot{W}(t, x), \quad t \in \mathbb{R}_+, \, x \in \mathbb{X}, \tag{2.3}$$

where $\mathbb{R}_+ := [0, \infty)$. Equation (2.3) is only formal and it is interpreted in the following sense.

Definition 2.1. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}, W)$ be a stochastic basis, a stochastic process $(u_t(x))_{t\geq 0, x\in\mathbb{X}}$ is called a **mild solution** to (2.3) with initial value $(u_0(x))_{x\in\mathbb{X}}$ if $(u_t(x))_{t\geq 0, x\in\mathbb{X}}$ is a predictable process so that, for each $(t,x) \in \mathbb{R}_+ \times \mathbb{X}$, almost surely,

$$u_t(x) = P_t u_0(x) + \int_0^t \int_{\mathbb{X}} p_{t-s}(x,y) b(s,y,u_s(y)) m(\mathrm{d}y)\mathrm{d}s + \iint_0^t p_{t-s}(x,y)\sigma(s,y,u_s(y)) W(\mathrm{d}s,\mathrm{d}y). \tag{2.4}$$

We now give rigorous meaning to the stochastic integral term in (2.4). Let K be a symmetric non-negative definite signed Radon measure on $\mathbb{X}^2$. Let Hilbert space H be the collection of measurable function f on $\mathbb{X}$ such that $\|f\|_H^2 := \int_{\mathbb{X}^2} f(x)f(y)K(\mathrm{d}x,\mathrm{d}y)$ is finite. Let $W = (W_t(f))_{f\in H, t\geq 0}$ be an adapted cylindrical Wiener process on H with covariance structure

$$\mathbb{E}[W_t(f)W_s(g)] = (t \wedge s)\langle f, g\rangle_H,$$

on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ satisfying the usual hypothesis. We call K the covariance measure of W and $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}, W)$ a stochastic basis. When $K(dxdy) = \delta_x(dy)m(dx)$, we call $\dot{W} := \partial_t W$ a space-time white noise. The stochastic integral w.r.t. W is given as in [56]. In particular, for any predictable H-valued process $(g_s)_{s\geq 0}$ such that almost surely $\int_0^t \|g_s\|_H^2 ds < \infty$ for every $t \geq 0$, the stochastic integral $\int_0^\cdot \int_{\mathbb{X}} g_s(y) W(ds,dy)$ is a continuous local martingale with quadratic variation $\int_0^\cdot \|g_s\|_H^2 ds$.

We note that the solutions the equation (2.3) might not live in $\mathcal{C}_b(\mathbb{X})$ (e.g. the Parabolic Anderson model), where $\mathcal{C}_b$ denote the collection of bounded continuous functions. Hence we will introduce spaces of continuous functions with weights similar to [55], where the weight function therein are $\{e^{\lambda|\cdot|} : \lambda \in \mathbb{R}\}$. In the following, we denote $\mathcal{C}^+(\mathbb{X})$ to be the collection non-negative continuous functions on $\mathbb{X}$.

Definition 2.2. We say h is a *weight* on a metric space $(\mathbb{X}, d)$ if $h \in \mathcal{C}^+(\mathbb{X})$ with $h \geq 1$, where $f \geq g$ means $f(x) \geq g(x)$ for all $x \in \mathbb{X}$. We say a weight h on $\mathbb{X}$ is *good* if:

(1) (Fast growth) For some $\lambda_0 > 0$ and sequence $(x_n)_{n\in\mathbb{N}}$ in $\mathbb{X}$ with $\mathbb{X} = \cup_{n\in\mathbb{N}} B(x_n, 1)$,

$$\sum_{n\in\mathbb{N}} h(x_n)^{-\lambda} < \infty, \qquad \text{for all } \lambda \geq \lambda_0. \tag{2.5}$$

(2) (Sufficiently regular)

$$\sup_{x\in\mathbb{X}} \frac{\sup_{y\in B(x,1)} h(y)}{\inf_{y\in B(x,1)} h(y)} < \infty. \tag{2.6}$$

Let h be a weight, we introduce the following weighted function spaces: for $\lambda \in \mathbb{R}$ and $f \in \mathcal{C}(\mathbb{X})$, denote $\|f\|_{(\lambda)} := \sup_{x\in\mathbb{X}} |f(x)|h(x)^{-\lambda}$, where the dependence on h will be suppressed for the simplicity of notations. Denote by

$$\mathcal{C}_{\text{tem}} := \left\{ f \in \mathcal{C}(\mathbb{X}) : \|f\|_{(\lambda)} < \infty, \text{ for all } \lambda > 0 \right\}, \text{ and } \mathcal{C}_{\text{rap}} := \left\{ f \in \mathcal{C}(\mathbb{X}) : \|f\|_{(-\lambda)} < \infty, \text{ for all } \lambda > 0 \right\},$$

the space of functions on $\mathbb{X}$ with temporal growth, and rapidly decline, respectively. The topology on $\mathcal{C}_{\text{tem}}$, and $\mathcal{C}_{\text{rap}}$, is generated by the family of norms $\{\|\cdot\|_{(\lambda)} : \lambda > 0\}$, and $\{\|\cdot\|_{(\lambda)} : \lambda < 0\}$, in the sense of [54, Theorem 1.37], respectively. It can be shown that they are Polish.

For any non-negative and non-negative definition symmetric measure μ on $\mathbb{X}^2$, we denote for measurable function f on $\mathbb{X}$ that

$$\|f\|_\mu^2 := \int_{\mathbb{X}} \int_{\mathbb{X}} |f(x)f(y)| \mu(\mathrm{d}x,\mathrm{d}y),$$

whenever the integral makes sense.

Assumption 2.3. There exists a non-negative and non-negative definite symmetric measure $\widetilde{K}$ on $\mathbb{X} \times \mathbb{X}$ so that $|K| \leq \widetilde{K}$ so that the heat kernel p of $\mathcal{L}$ satisfies the following conditions:

(H$_1$) There are ξ_1, $\xi_2 \in (0,1]$, such that for each $T>0$, the following inequality holds uniformly for all $t \le t'$ in $[0,T]$ and $x,y \in \mathbb{X}$ with $d(x,y) \le 1$,

$$\int_0^{t'} \left(\int_{\mathbb{X}} |p_{t'-s}(x,z) - p_{t-s}(y,z)|\, m(\mathrm{d}z) + \|p_{t'-s}(x,\cdot) - p_{t-s}(y,\cdot)\|_{\widetilde{K}}^2 \right) \mathrm{d}s \lesssim \left(|t'-t|^{\xi_1} + d(x,y)^{\xi_2} \right).$$

(H$_2$) For each $\delta \in (0,1)$,

$$\lim_{\varepsilon \downarrow 0} \sup_{x \in \mathbb{X}} \left\| p_\epsilon(x,\cdot) 1_{B(x,\delta)^c}(\cdot) \right\|_{\widetilde{K}}^2 = 0. \tag{2.7}$$

(H$_3$) There exists $\beta \in (0,1)$ and a non-decreasing, continuous function $\gamma : \mathbb{R}_+ \to \mathbb{R}_+$ with $\gamma(0)=0$ so that for each $T>0$, uniformly for every $x \in \mathbb{X}$ and $t,s \in (0,T]$,

$$\int_{\mathbb{X}} |p_t(x,y) - p_s(x,y)|\, m(\mathrm{d}y) + \|p_t(x,\cdot) - p_s(x,\cdot)\|_{\widetilde{K}}^2 \lesssim \frac{\gamma(|t-s|)}{(t \wedge s)^\beta}.$$

Remark 2.4. We will see in Proposition 4.5 and Lemma 4.11 that the inequality (2.7) is not restrictive.

Assumption 2.5. There exists a good weight h on $(\mathbb{X}, d)$ such that

(W$_1$) The map $t \mapsto P_t f$ for $t \in \mathbb{R}_+$ is continuous in $\mathcal{C}_{rap}$ for $f \in \mathcal{C}_{\mathrm{rap}}$ and in $\mathcal{C}_{\mathrm{tem}}$ if $f \in \mathcal{C}_{\mathrm{tem}}$.

(W$_2$) There is $\alpha \in (0,1)$ so that for K and $\widetilde{K}$ in Assumption 2.3 and for each $\lambda \in \mathbb{R}$,

$$\begin{aligned} &\sup_{t \in [0,T]} \sup_{x \in \mathbb{X}} h(x)^{-\lambda} \int_{\mathbb{X}} p_t(x,y) h(y)^\lambda m(\mathrm{d}y) < \infty, \quad \text{for all } T>0, \\ &\sup_{x \in \mathbb{X}} h(x)^{-2\lambda} \left\| p_t(x,\cdot) h^\lambda(\cdot) \right\|_{\widetilde{K}}^2 \lesssim t^{-\alpha} \vee 1, \quad \text{uniformly in } t>0, \end{aligned} \tag{2.8}$$

where $a \vee b := \max\{a,b\}$ for $a,b \in \mathbb{R}$.

Assumption 2.6. There exists a good weight h on $(\mathbb{X}, d)$ such that The map $t \mapsto P_t f$ for $t \in \mathbb{R}_+$ is continuous in $\mathcal{C}_{tem}$ if $f \in \mathcal{C}_{tem}$, and there is $\alpha \in (0,1)$ so that inequalities in (2.8) holds for $\lambda \in [0,1]$.

Remark 2.7. Let us comment on the above assumptions. (H$_1$) and (W$_1$) are for the continuity of sample paths via Kolmogorov-type lemma; condition (W$_2$) is for controlling the growth of solutions and the second inequality in (2.8) determines the resolvability of equation (2.3); conditions (H$_2$) and (H$_3$) are needed for the non-negativity of solutions in Theorem 5.2.

Remark 2.8. Feller property, condition (H$_2$) and condition (H$_3$) are not needed for the well-posedness of (2.3) (Theorem 5.1).

3. Main results

In this section, we will state our main results on compact support property and strong comparison principle of the equation (2.3). We will also give easy-to-check sufficient condition for the key assumption in terms of a heat kernel bound in Section 4.1 Proposition 4.1 and Proposition 4.3.

Theorem 3.1 (Compact support property)**.** *Let* $(\mathbb{X}, d, m)$ *be a metric measure space that satisfies* (2.1)*,* W *has covariance measure* K *and suppose Assumption 2.6 hold.*

(1) Suppose condition (H$_1$) holds for some non-negative and non-negative definite measure $\widetilde{K}$ *on* $\mathbb{X}^2$ *with* $\widetilde{K} \ge |K|$*, and there exists* $L \ge 0$ *so that almost surely for all* $(t,x,u) \in \mathbb{R}_+ \times \mathbb{X} \times \mathbb{R}$

$$|b(t,x,u)| + |\sigma(t,x,u)| \le L\,(1+|u|). \tag{3.1}$$

Then for any initial valued $f \in \mathcal{C}_{tem}$*, there exists a* $\mathcal{C}_{tem}$*-valued solution* u *to* (2.3) *with* $u_0 = f$ *on a suitable stochastic basis* $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \ge 0}, \mathbb{P}, W)$*.*

(2) If, in addition, Assumption 2.3 holds and $b(t,x,0) \ge 0$*,* $\sigma(t,x,0) = 0$ *almost surely for all* $(t,x) \in \mathbb{R}_+ \times \mathbb{X}$*, then for any* $f \in \mathcal{C}_{tem}^+$*, there is a* $\mathcal{C}_{tem}^+$*-valued solution* u *to* (2.3) *with* $u_0 = f$ *on a suitable stochastic basis* $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \ge 0}, \mathbb{P})$*.*

(3) *Assume, in addition, Assumption 2.5 holds and there exists $\theta \in (0,1)$ so that for any $M > 0$, we have almost surely that*

$$|b(t,x,u)| \lesssim u \quad \text{and} \quad |\sigma(t,x,u)| \lesssim u^\theta \qquad \text{uniformly in } (t,x,u) \in \mathbb{R}_+ \times \mathbb{X} \times [0,M]. \tag{3.2}$$

Then for $f \in \mathcal{C}^+_{rap}$, any $\mathcal{C}^+_{tem}$-valued solution u to (2.3) *with $u_0 = f$ is a $\mathcal{C}^+_{rap}$-valued solution.*

(4) *Finally, if, in addition, $\dot{W}$ is the space time white noise and the following conditions holds*
   (a) *For any $M > 0$, it holds almost surely that*

$$|u|^{\frac{1}{2}} \lesssim |\sigma(t,x,u)| \qquad \text{uniformly in } (t,x,u) \in \mathbb{R}_+ \times \mathbb{X} \times [0,M]. \tag{3.3}$$

   (b) *Recall $(Y_t)_{t\geq 0}$ and $\{\Pi_x\}_{x\in\mathbb{X}}$ in* (2.2)*. For any $T \geq 0$, there exists constants $C_1 = C_1(T) > 0$, $p = p(T) \geq 3$, $\varepsilon = \varepsilon(T) > 0$ such that*

$$\Pi_x\left[\sup_{0\leq r\leq t} d(x,Y_r)^p\right] \leq C_1 t^{2+\varepsilon}, \qquad \text{for all } x \in \mathbb{X},\ t \in [0,T]. \tag{3.4}$$

Then the compact support property holds, i.e., if $f \in \mathcal{C}^+_{rap}$, then any $\mathcal{C}^+_{tem}$-valued solution u to (2.3) *with $u_0 = f$ satisfies $\mathbb{P}(u_t \in \mathcal{C}^+_c(\mathbb{X})) = 1$ for each $t > 0$.*

Remark 3.2. The proof of Theorem 3.1 is given in Section 6. A sufficient condition for (3.4) in terms of heat kernel upper bound is given in Proposition 4.1.

Remark 3.3. The initial condition need not to be compactly supported and $\sigma(u) = |u|^\gamma$ satisfies the assumption of Theorem 3.1 for $\gamma \in (0,1/2]$.

Remark 3.4. It can be seen in the proof of Theorem 3.1 that, if $(\mathbb{X},d,m)$ satisfies (2.1) and Assumption 2.3 and 2.5 holds. If b and σ satisfies (3.3) and inequality (3.4) holds. Then for any $\mathcal{C}^+_{rap}$-valued solution u to (2.3), we have $\mathbb{P}\left(u_t \in \mathcal{C}_c(\mathbb{X})\right) = 1$ for all $t > 0$.

Remark 3.5. Our conditions (3.3) and (3.2) are slightly different from that of Shiga's in [55, Theorem 1.1]. In there, the $\sigma \in \mathcal{C}(\mathbb{R})$ with $\sigma(0) = 0$ and $|\sigma(u)| \leq C(1+|u|)$ for some $C > 0$. In addition, for each $A > 0$, there exists $c_A > 0$ so that $c_A|u|^{\frac{1}{2}} \leq |\sigma(u)|$ for $|u| \leq A$. However, in Shiga's proof, the upper bound $|\sigma(u)| \lesssim |u| + |u|^\theta$ for some $\theta \in (0,1/2]$ is implicitly used to construct a $\mathcal{C}^+_{rap}$-valued solution in the beginning of [55, page 422]. We note that there exists functions σ that satisfies the conditions of [55, Theorem 1.1] but does not satisfies the upper bound (e.g. functions behave like $(\ln(|u|^{-1}))^{-1}$ around 0). Hence the upper bound on σ as in (3.2) is needed here as well as in [55, Theorem 1.1].

Theorem 3.6 (Strong comparison principle)**.** *Let $(\mathbb{X},d,m)$ be a metric measure space that satisfies* (2.1)*, $\mathcal{L}$ satisfies Assumption 2.3 and Assumption 2.6 holds with K being the covariance measure of W.*

(1) *Suppose there is $L > 0$ so that almost surely for all $(t,x) \in \mathbb{R}_+ \times \mathbb{X}$ and $u,v \in \mathbb{R}$,*

$$|b(t,x,u) - b(t,x,v)| + |\sigma(t,x,u) - \sigma(t,x,v)| \leq L|u-v|. \tag{3.5}$$

Then for any given stochastic basis and any $u_0 \in \mathcal{C}_{tem}$, there exists a solution to (2.3) *with initial value u_0.*

(2) *Suppose in addition that for each $a,M,t > 0$, $x_0 \in \mathbb{X}$, there is $N_0 = N_0(a,M,t,x_0) \in \mathbb{N}$ such that for all sufficiently large $Q > 0$,*

$$\inf_{s\in\left[\frac{t}{Q},\frac{2t}{Q}\right]} \inf_{x\in B(x_0,a+M/Q)} \int_{B(x_0,a)} p_s(x,y)m(\mathrm{d}y) > \frac{1}{N_0}. \tag{3.6}$$

Let u_1 and u_2 be $\mathcal{C}_{tem}$-valued solutions of (2.3) *with the initial conditions $u_i(0) = f_i \in \mathcal{C}_{tem}$ for $i = 1,2$ with $f_1(x) \geq f_2(x)$ for all $x \in \mathbb{X}$ and $f_1(x) > f_2(x)$ for some $x \in \mathbb{X}$. Then*

$$\mathbb{P}\left(u_1(t,x) > u_2(t,x) \text{ for every } t > 0,\ x \in \mathbb{X}\right) = 1.$$

(3) In particular, if b and σ satisfies (3.5), *and almost surely for all $(t,x) \in \mathbb{R}_+ \times \mathbb{X}$*

$$b(t,x,0) \geq 0, \quad \text{and} \quad \sigma(t,x,0) = 0. \tag{3.7}$$

If $u_0 \in \mathcal{C}_{tem}^{+}$ and $u_0(x) > 0$ for some $x \in \mathbb{X}$, then the solution u to (2.3) *with initial value u_0 satisfies $\mathbb{P}(u_t(x) > 0,\ \text{for all } t > 0,\ x \in \mathbb{X}) = 1$.*

Remark 3.7. The proof to Theorem 3.6 is given in Section 7, and a sufficient condition for (3.6) in terms of heat kernel lower bound is given in Proposition 4.3.

4. Examples

In this section, we will first provide sufficient conditions for the key assumptions of Theorem 3.1 and Theorem 3.6 in terms of heat kernel bounds in Section 4.1. Then we will give several examples of operators $\mathcal{L}$ and metric metric measure spaces including (fractional) Laplacian on Euclidean spaces, Kirchhoff's Laplacian on metric graphs with bounded geometry, the generator of Brownian motions on fractals and random trees. The readers who are interested in the specific examples can skip Section 4.1.

4.1. **Sufficient conditions for key conditions.** We first give sufficient conditions for the key assumptions (3.4) in Theorem 3.1, (3.6) in Theorem 3.6 and (2.7) in Assumption 2.3.

First, we give a sufficient condition for (3.4) in terms of heat kernel upper bounds.

Proposition 4.1. *Let $(\mathbb{X}, d, m)$ be a metric measure space and $V : \mathbb{R}_+ \to [1,\infty)$ is a non-decreasing function. Suppose there is an $d_h \geq 1$ so that uniformly in $r > 0$,*

$$m(B(x,r)) \lesssim V(r) r^{d_h}.$$

Let $(Y_t)_{t\geq 0}$ be a strong Markov process with almost sure continuous sample paths on $\mathbb{X}$ and $(p_t(x,y),\ t > 0,\ x,y \in \mathbb{X})$ be as in (2.2). *Suppose the following conditions holds:*

(1) There is a $d_w > 1$ and a non-increasing function $\Phi : \mathbb{R}_+ \times \mathbb{R}_+$ so that for each $T > 0$,

$$p_t(x,y) \lesssim t^{-\frac{d_h}{d_w}} \Phi\left(\frac{d(x,y)}{t^{\frac{1}{d_w}}}\right), \quad \text{uniformly in } t \in [0,T],\ x,y \in \mathbb{X}. \tag{4.1}$$

(2) There is $p > \max\{2d_w, 3\}$ so that for each $t > 0$,

$$\int_0^\infty V(ts) s^{d_h+p} \Phi(s) \mathrm{d}s < \infty, \tag{4.2}$$

$$\lim_{r\to\infty} t^{-\frac{d_h}{d_w}} \Phi\left(\frac{r}{t^{1/d_w}}\right) V(r) r^{d_h+p} = 0. \tag{4.3}$$

Then the process $(Y_t)_{t\geq 0}$ satisfies (3.4).

Remark 4.2. In particular, if there are $d_w > 1$, $d_h \geq 1$, $C \geq 0$ so that for each $T > 0$, there exists $c = c(T) > 0$ with

$$m(B(x,r)) \lesssim r^{d_h} \exp(Cr), \quad \text{uniformly in } r > 0,$$

$$p_t(x,y) \lesssim t^{-\frac{d_h}{d_w}} \exp\left(-c\left(\frac{d(x,y)}{t^{\frac{1}{d_w}}}\right)^{\frac{d_w}{d_w-1}}\right), \quad \text{uniformly in } t \in (0,T],\ x,y \in \mathbb{X}.$$

Then the associated process Y satisfies (3.4).

We give an sufficient condition for the inequality (3.6) in terms of heat kernel lower bounds.

Proposition 4.3. *Suppose $(\mathbb{X}, d)$ is a length space and m is a measure on $\mathbb{X}$ satisfies for some $d_h > 0$,*

$$r^{d_h} \lesssim m(B(x,r)), \quad \text{uniformly for all } x \in \mathbb{X}, \quad 0 < r < \mathit{diam}(\mathbb{X}), \tag{4.4}$$

where $\mathit{diam}(\mathbb{X}) = \sup_{x,y\in\mathbb{X}} d(x,y) \in (0,\infty]$. Let $\Phi : [0,\infty) \to (0,\infty)$ be a strictly positive and decreasing function so that for some constant $c > 0$,

$$\Phi(ca)\Phi(cb) \lesssim \Phi(a+b), \quad \text{uniformly for } a,b \in \mathbb{R}_+. \tag{4.5}$$

Then $(p_t(x,y),\, t>0,\, x,y\in\mathbb{X})$ *satisfies* (3.6) *provided that there is* $d_w>1$ *so that for each* $T>0$*, there are* $c_T, C_T>0$ *so that*

$$C_T t^{-\frac{d_h}{d_w}}\Phi\left(c_T\frac{d(x,y)}{t^{\frac{1}{d_w}}}\right)\le p_t(x,y),\qquad \text{for all } x,y\in\mathbb{X},\, t\in[0,T]. \tag{4.6}$$

Remark 4.4. $\Phi(r)=(1+r)^{-(d_h+d_w)}$ and $\Phi(r)=\exp(-r^{\frac{d_w}{d_w-1}})$ for some $d_w>1$, $d_h\ge 1$ satisfies the inequality (4.5).

Proposition 4.5. *Suppose* $(p_t(x,y),\, t>0,\, x,y\in\mathbb{X})$ *satisfies the two sided (sub) Gaussian bounds, that is, there exists* $d_w\ge 2$ *and* $d_h\ge 1$ *and* $c>0$ *so that*

$$p_t(x,y)\asymp t^{-d_h/d_w}\Phi\left(cd(x,y)/t^{\frac{1}{d_w}}\right),\qquad \text{uniformly in } t\in(0,1],\, x,y\in\mathbb{X}, \tag{4.7}$$

where $\Phi(r):=\exp(-r^{\frac{d_w}{d_w-1}})$*, and the constant* c *might be different for upper and lower bound. Then* (2.7) *holds if the second inequality in* (2.8) *holds for* $h\equiv 1$ *on* $\mathbb{X}$*.*

Remark 4.6. It is known that if a complete metric measure space $(\mathbb{X},d,m)$ is Ahlfors regular, i.e. there is $s>0$ so that $m(B(x,r))\asymp r^s$ uniformly in $x\in\mathbb{X}$ and $r\le \operatorname{diam}(\mathbb{X})$, where $\operatorname{diam}(\mathbb{X}):=\sup_{x,y\in\mathbb{X}} d(x,y)\in(0,\infty]$, then $(\mathbb{X},d)$ satisfies (2.1) (c.f. [48]).

4.2. **Euclidean spaces.** The simplest example that fits our framework are the n-dimensional Euclidean spaces $\mathbb{R}^n$ for $n\ge 1$, equipped with the Euclidean distance and Lebesgue measure. Let Δ be the standard Laplacian on $\mathbb{R}^n$, it is well known the associated heat kernel is

$$p_t(x,y)=\frac{1}{(4\pi t)^{\frac{n}{2}}}e^{-\frac{|x-y|^2}{4t}},\quad t>0,\, x,y\in\mathbb{R}^n.$$

Suppose W is a Gaussian noise whose covariance measure admits a density, $(k(x,y),\, x,y\in\mathbb{R}^n)$, with respect to the Lebesgue's measure on $\mathbb{R}^n\times\mathbb{R}^n$ so that for some $\alpha\in(0,2\wedge d)$,

$$|k(x,y)|\lesssim |x-y|^{-\alpha}+1.$$

Then Assumption 2.3 and 2.6 holds with $h(x)=e^{|x|}$ (c.f. [53]), where the strong comparison principle was shown in [15] with measure-valued initial values. When $n=1$, then Assumption 2.5 holds for $h(x)=e^{|x|}$ and $\dot W$ being the space time white noise, where the compact support property and strong comparison principle was first shown in [55].

Another example on $\mathbb{R}$ that fits our framework is the fractional Laplacian:

$$(-\Delta)^{\frac{\alpha}{2}},\quad \alpha\in(1,2),$$

whose heat kernel satisfies

$$p_t(x,y)\asymp t^{-\frac{1}{\alpha}}\left(1+\frac{|x-y|}{t^{\frac{1}{\alpha}}}\right)^{-(1+\alpha)},\quad \text{uniformly in } t>0,\, x,y\in\mathbb{R}.$$

Then it is known that Assumptions 2.3 and 2.6 hold if $\dot W$ is the space time white noise and $h(x)=(1+|x|)^{\frac{1}{2}}$. The strong comparison principle is for fractional Parabolic Anderson model on $\mathbb{R}$ was shown in [16].

4.3. **Metric graphs.** Roughly speaking, a metric graph is a topological graph whose edges quipped with one dimensional Euclidean metric (see [23]). Now suppose the metric graph (Γ,d) has bounded geometry in the sense of [30], and m is the one-dimensional Hausdorff measure and Δ is the Kirchhoff's Laplacian on (Γ,d), it was shown in [30] that there is a diffusion on Γ associated with Δ whose transition density satisfies the (short-time) two-sided Gaussian estimates, and is jointly Hölder continuous. Thus, if $\dot W$ is the space time white noise on (Γ,m), then Δ satisfies Assumption 2.3 and $h(x)=\exp(d(x,x_0))$ for any $x_0\in\Gamma$ satisfies Assumption 2.5.

Example 4.7 (Super-Brownian motion density on regular trees). let (Γ, d) be a regular metric tree and m is the one-dimensional Hausdorff measure on Γ. Consider the following equation with Kirchhoff's boundary condition:

$$\partial_t u_t(x) = \Delta u_t(x) + \sqrt{u_t(x)}\dot{W}, \quad (t,x) \in \mathbb{R}_+ \times \Gamma.$$

Suppose $u_0 \in \mathcal{C}_c^+(\Gamma)$ then by Theorem 3.1 and proposition 4.1, there is a solution u to above equation with initial value u_0 and

$$\mathbb{P}\left(u_t \in \mathcal{C}_c^+ \text{ for every } t > 0\right) = 1.$$

Example 4.8 (Parabolic Anderson model on regular trees). Under the setting of Example 4.7, consider the following equation with Kirchhoff's boundary condition:

$$\partial_t u_t(x) = \Delta u_t(x) + u_t(x)\dot{W}, \quad (t,x) \in \mathbb{R}_+ \times \Gamma.$$

Suppose $u_0 \in \mathcal{C}_{tem}^+(\Gamma)$ and $u_0(x) > 0$ for some $x \in \Gamma$, then by Theorem 3.6 and Proposition 4.3, there is a unique solution u to the above equation with initial value u_0 and

$$\mathbb{P}\left(u_t(x) > 0, \ \text{for all } t > 0, \, x \in \Gamma\right) = 1.$$

4.4. **Fractals.** Another class of examples that fits our settings is a class of fractals which includes the Sierpiński gasket (bounded and unbounded). For heat kernel and function theories on fractals, we refer the readers to [3] and [42]. Without going into details of analysis on fractals, we remark that in many cases, the fractal M equipped with a metric d and a measure μ is d_h-Ahlfors regular, where d_h is the Hausdorff dimension of M, and hence (M,d) satisfies (2.1). In addition, it carries a nature Laplacian operator $\mathcal{L}$, which admits a symmetric heat kernel satisfies for some $2 \le d_w$ and $d_h < d_w$

$$p_t(x,y) \asymp t^{-\frac{d_h}{d_w}} \Phi\left(\frac{d(x,y)}{t^{\frac{1}{d_w}}}\right), \qquad \text{uniformly in } t \in (0,\infty), \, x, y \in M \tag{4.8}$$

where $\Phi(r) = \exp\left(-cr^{\frac{d_w}{d_w - 1}}\right)$ for some $c > 0$. It can be shown (see [28, Theorem 7.4] or [4, Thoerem 3.1]) that the heat kernel is jointly Hölder continuous and Assumptions 2.3 and 2.5 are satisfied for $\dot{W}$ being the space time white noise and weight $h(x) = e^{d(x_0,x)}$ for some fixed point $x_0 \in M$. In addition, the heat kernel satisfies (4.8) also satisfies the conditions in Proposition 4.1 and Proposition 4.3

For example, it is known that if M is the unbounded Sierpiński gasket in $\mathbb{R}^n$, then carries a heat kernel satisfies (4.8) with $d_h = \log_2(n+1)$ and $d_w = \log_2(n+3)$ (see [5, 3]). In addition, there is a small $\theta > 0$ so that uniformly in $t > 0$ and $z, x, y \in M$ with $d(x,y) \le t^{\frac{1}{d_w}}$,

$$|p_t(z,x) - p_t(z,y)| \lesssim \left(\frac{d(x,y)}{t^{\frac{1}{d_w}}}\right)^{\theta} t^{-\frac{d_h}{d_w}} \Phi\left(\frac{d(x,y)}{t^{\frac{1}{d_w}}}\right). \tag{4.9}$$

By [22, Corrollary 5],

$$|\partial_t p_t(x,y)| \lesssim t^{-1} t^{-\frac{d_h}{d_w}} \Phi\left(\frac{d(x,y)}{t^{\frac{1}{d_w}}}\right), \qquad \text{uniformly in } t > 0, \, x, \, y \in M. \tag{4.10}$$

Example 4.9 (Super-Brownian motion density on Sierpiński gasket). Suppose $n \in \mathbb{N}$ and (M, d, μ) is the unbounded Sierpeński gasket in $\mathbb{R}^n$ and let $\mathcal{L}$ be the generator of the Brownian motion on (M, d, μ) and consider the following equation:

$$\partial_t u_t(x) = \mathcal{L} u_t(x) + \sqrt{u_t(x)}\dot{W}(x,t).$$

Suppose $u_0 \in \mathcal{C}_c^+(M)$, then by Theorem 3.1 and Proposition 4.1, there is a solution u with initial value u_0, such that

$$\mathbb{P}\left(u_t \in \mathcal{C}_c^+ \text{ for every } t > 0\right) = 1.$$

Example 4.10 (Parabolic Anderson model on Sierpiński gasket). Under the setting of Example 4.9, if $u_0 \in \mathcal{C}_c^+(M)$ is not identically zero, then by Theorem 3.6 and Proposition 4.3, there is a unique solution u to the following equation with initial value u_0,

$$\partial_t u_t(x) = \mathcal{L}u_t(x) + u_t(x)\dot{W}(x,t),$$

and

$$\mathbb{P}\left(u_t(x) > 0 \text{ for every } t > 0,\, x \in \mathbb{X}\right) = 1.$$

Similar to the case of $\mathbb{R}$, the fractional powers of $\mathcal{L}$ on the the Sierpiński gasket also fits in our framework and satisfies Assumption 2.3 with the space time white noise and weight $h(x) = (1+d(x_0,x))^{\frac{1}{2}}$ satisfies Assumption 2.6. We refer the readers to [18, 19, 17] for more details about heat kerenl associated with fractional powers of diffusion operators on non-smooth spaces.

The following result shows the metric measure spaces that carries a sub-Gaussian heat kernel fits our framework. The proof of which is given in the Appendix.

Lemma 4.11. *Suppose (M,d) is a length space and μ is a Radon measure on M with full support. Assume the heat kernel of $\mathcal{L}$ satisfies the inequality (4.8) for $d_w \geq 2$ and $d_w > d_h \geq 1$, Hölder continuity (4.9) and is stochastic completeness. If we fix an arbitrary $x_0 \in M$ and let $h(x) = (1 + d(x_0,x))^{\frac{1}{2}}$ for $x \in M$. Then the assumptions of Theorem 3.6 are satisfied for the space time white noise $\dot{W}$ and if we replace $\mathcal{L}$ by $\mathcal{L}^\delta$ for $\delta \in (d_h/d_w, 1)$.*

Lemma 4.11, together with Theorem 3.6 and Theorem 5.1 implies the following Proposition

Proposition 4.12 ((Fractional) Parabolic Anderson model)**.** *Under the setting of Lemma 4.11 and let $\delta \in (d_w/d_h, 1]$. Suppose $u_0 \in \mathcal{C}_c^+(M)$ is not identically zero, then there is a unique $\mathcal{C}_{tem}^+(M)$-valued solution u to the following equation with initial value u_0,*

$$\partial_t u_t(x) = -(-\mathcal{L})^\delta u_t(x) + u_t(x)\dot{W}(x,t),$$

and $\mathbb{P}\left(u_t(x) > 0, \text{ for all } (t,x) \in (0,\infty) \times \mathbb{X}\right) = 1$, where $-(-\mathcal{L})^1 := \mathcal{L}$.

Example 4.13 (Fractional Parabolic Anderson model on the Sierpeński gasket). Given any integer $n \geq 2$, let (M,d,μ) be the unbounded Sierpeński gasket in $\mathbb{R}^n$ and $\mathcal{L}$ be the generator of the Brownian motion on (M,d,μ). Let $1 > \delta > \log_2(n+1)/\log_2(n+3)$ and $u_0 \in \mathcal{C}_c^+(M)$ that is not identically zero. Then there is a unique solution u to

$$\partial_t u_t(x) = -(-\mathcal{L})^\delta u_t(x) + u_t(x)\dot{W}(x,t),$$

where $\dot{W}$ is the space-time white noise. In addition, $\mathbb{P}\left(u_t(x) > 0, \text{ for every } t > 0,\, x \in M\right) = 1$.

4.5. **Random spaces with volume fluctuations.** We demonstrate in this section that the well-posedness of (2.3) without any regularity requirements needs minimum assumptions in general.

With the theory of Dirichlet form, one can construct diffusion processes on certain random metric measure spaces and obtain heat kernel bounds. In [20], Croydon showed that there is a Brownian motion on continuum random trees and gave the two-side heat kernel estimates. In particular, if we let $\mathcal{I}$ be the continuum random tree and $d_{\mathcal{I}}$ and μ be the natural metric and volume measure on $\mathcal{I}$ constructed in [20]. He showed there exists random constants C_2, C_3, θ, $T > 0$ so that for all $t \in [0,T]$ and $\sigma, \sigma' \in \mathcal{I}$,

$$p_t(\sigma,\sigma') \leq C_2 t^{-\frac{2}{3}} \left(\ln_1 t^{-1}\right)^{\frac{1}{3}} \exp\left(-C_3 \left(\frac{d_{\mathcal{I}}(\sigma,\sigma')^3}{t}\right)^{1/2} \ln_1\left(\frac{d_{\mathcal{I}}}{t}\right)^{-\theta}\right).$$

It can be seen by the semi-group property that the heat kernel $(p_t(\sigma,\sigma'),\, t > 0,\, \sigma,\, \sigma' \in \mathcal{I})$ satisfies conditions (2.8) with $\dot{W}$ being space time white noise independent of $\mathcal{I}$, $h \equiv 1$ on $\mathcal{I}$ and $\lambda = 0$ for some finite random time horizon $T > 0$ almost surely. Hence one can solve (2.3) with Lipschitz coefficients almost surely on $(\mathcal{I}, d_{\mathcal{I}}, \mu)$ with $\mathcal{L}$ being the generator of the Brownian motion on $\mathcal{I}$ (see Theorem 5.1).

And other example covered by Theorem 5.1 is the the Liouville Brownian motion. One can also obtain explicit heat kernel upper bound and assumptions of 5.1 are fulfilled with appropriate Gaussian noise and weight $h \equiv 1$. For details about Liouville Brownian motion, and its heat kernel, see [1, 49, 9].

5. Existence, uniqueness and weak comparison principle

We summarize several facts about SPDEs on general metric measure spaces. Even though the well posedness and regularities of solutions to equation (2.3) was studied in various settings before, none of them is sufficient for our purpose since we need controls on the spatial growth of the solution. Hence, we provide several existence and uniqueness results in the functions spaces $\mathcal{C}_{\mathrm{tem}}$ and $\mathcal{C}_{\mathrm{rap}}$.

Theorem 5.1 (Well-posedness). *Let $(\mathbb{X}, d, m)$ be a metric measure space, $\lambda \in \mathbb{R}$ and h be a weight on $\mathbb{X}$. Suppose the b and σ satisfies* (3.5) *and the inequalities* (2.8) *hold for some $\widetilde{K}$ that dominates $|K|$, where K is the covariance measure of W. Then for any $u_0 \in \mathcal{C}(\mathbb{X})$ with $\|u_0\|_{(\lambda)} < \infty$, the equation* (2.3) *has a unique mild solution u (up to modifications), and for all $T > 0$, $p \geq 2$,*

$$\sup_{t\in[0,T]} \sup_{x\in\mathbb{X}} h(x)^{-\lambda}\mathbb{E}\left[|u_t(x)|^p\right] < \infty.$$

Suppose in addition $(\mathbb{X}, d, m)$ satisfies (2.1)*, $\mathcal{L}$ satisfies (H$_1$) of Assumption 2.3 and h satisfies Assumption 2.6. Then for any initial value $u_0 \in \mathcal{C}_{tem}$, the solution to the equation* (2.3) *u satisfies $\mathbb{P}(u_t(\cdot) \in \mathcal{C}_{tem},\ \text{for all } t \geq 0) = 1$.*

Theorem 5.2 (Non-negativity). *Suppose $(\mathbb{X}, d, m)$ satisfies* (2.1)*, $\mathcal{L}$ satisfies Assumption 2.3 and Assumption 2.6 holds for some $\widetilde{K}$ that dominates $|K|$, where K is the covariance measure of W. Suppose b, σ satisfies* (3.5) *and* (3.7)*. Let u be the unique continuous $\mathcal{C}_{tem}$-valued solution of the equation* (2.3) *with initial value $u_0 \in \mathcal{C}^+_{tem}$. Then*

$$\mathbb{P}\left(u_t \geq 0 \text{ for all } t \geq 0\right) = 1.$$

Corollary 5.3 (Weak comparison principle). *Under the assumptions of Theorem 5.2, let σ and b_i $(i = 1, 2)$ satisfy the inequality* (3.5) *and u_i be the $\mathcal{C}_{tem}$-valued solution of the SPDE* (2.3) *with coefficients b_i, σ and initial values $f_i \in \mathcal{C}_{tem}$ for $i = 1, 2$ respectively. Suppose almost surely for all $(t, x) \in \mathbb{R}_+ \times \mathbb{X}$,*

$$b_1(t, x, u) \geq b_2(t, x, u) \text{ for all } u \in \mathbb{R} \quad \text{and} \quad f_1 \geq f_2.$$

Then $\mathbb{P}\big(u_1(t, \cdot) \geq u_2(t, \cdot) \text{ for all } t \geq 0\big) = 1$.

Theorem 5.4 (Weak existence). *Under the settings of Theorem 5.2 and suppose there is $L > 0$ so that almost surely for all $(t, x) \in \mathbb{R}_+ \times \mathbb{X}$ and $b(t, x, \cdot)$ and $\sigma(t, x, \cdot)$ are continuous functions on $\mathbb{R}$ and satisfies* (3.1)*. Then for every $u_0 \in \mathcal{C}_{tem}$, there is $\mathcal{C}_{tem}$-valued solution u of the equation* (2.3) *on a suitable stochastic basis. If in addition u_0 is non-negative, then the the solution is $\mathcal{C}^+_{tem}$-valued.*

Theorem 5.5 ($\mathcal{C}^+_{rap}$-valued solutions). *Under the settings of Theorem 5.2 and assume in addition Assumption 2.5 holds. If there exists $C > 0$ and $\theta \in (0, 1)$ so that almost surely for all $(t, x) \in \mathbb{R}_+ \times \mathbb{X}$*

$$|b(t, x, u)| \leq C|u| \quad \text{and} \quad |\sigma(t, x, u)| \leq C(|u| + |u|^\theta). \tag{5.1}$$

If u is a $\mathcal{C}^+_{tem}$-valued solution to the SPDE (2.3) *with initial value $u_0 \in \mathcal{C}^+_{rap}$, then u is a $\mathcal{C}^+_{rap}$-valued solution. In addition, for all $\lambda, T > 0$ and $p \geq 1$,*

$$\sup_{t\in[0,T]} \sup_{x\in\mathbb{X}} h(x)^\lambda \left\|u_t(x)\right\|_{L^p(\mathbb{P})} < \infty.$$

6. Proof of Theorem 3.1

Note that part (1) and 2 are shown in Theorem (5.1) and Theorem 5.2 respectively. Part (3) is a consequence of Theorem 5.5 with the observation that if b and σ satisfies both (3.1) and (3.2), then they also satisfies (5.1). Hence, we will only prove part (4) in this section.

Recall $(Y_t)_{t\geq 0}$, $(\Pi_x)_{x\in\mathbb{X}}$ and $(p_t(x, y),\ t > 0,\ x, y \in \mathbb{X})$ in (2.2). We denote

$$P_t(x, \mathrm{d}y) = p_t(x, y)m(\mathrm{d}y), \quad t > 0, x, y \in \mathbb{X}.$$

One ingredient for the proof is the theory of a class of measure-valued processes known as the superprocesses. For our purpose, we only need to consider a family of Y-superprocesses with critical binary branching. The laws of the processes in this family can be indexed by a single parameter $\beta > 0$ known as the branching rate. Let us give a description of this process. For any $\beta > 0$ and bounded non-negative Borel function f on $\mathbb{X}$, there exists a unique locally bounded non-negative Borel function $(V_t^\beta f(x) : t \geq 0, x \in \mathbb{X})$ such that

$$V_t^\beta f(x) + \beta \int_0^t \mathrm{d}s \int_{\mathbb{X}} \big(V_{t-s}^\beta f(y)\big)^2 P_s(x, \mathrm{d}y) = P_t f(x), \quad t \geq 0, x \in \mathbb{X}. \tag{6.1}$$

The Polish space of all finite Borel measures on $\mathbb{X}$, equipped with the topology of weak convergence, is denoted by $\mathcal{M} := \mathcal{M}(\mathbb{X})$. It is known that there exists an $\mathcal{M}$-valued (conservative) Borel right process $(X_t)_{t\geq 0}$ with transition semigroup $(Q_t^\beta)_{t\geq 0}$ such that for each $\mu \in \mathcal{M}(\mathbb{X}), t \geq 0$ and bounded non-negative Borel function f on $\mathbb{X}$,

$$\int_{\mathcal{M}} \exp\Big\{-\int_{\mathbb{X}} f(x)\eta(\mathrm{d}x)\Big\} Q_t^\beta(\mu, \mathrm{d}\eta) = \exp\Big\{-\int_{\mathbb{X}} V_t^\beta f(x)\mu(\mathrm{d}x)\Big\}. \tag{6.2}$$

We call the process $(X_t)_{t\geq 0}$ a critical binary Y-superprocess with branching rate β. It can be verified that the (non-linear) operators $(V_t^\beta : t \geq 0)$ satisfies the semigroup property: $V_{t+s}^\beta f = V_t^\beta V_s^\beta f$ for every $t, s \geq 0$ and bounded non-negative Borel function f on $\mathbb{X}$. We call $(V_t^\beta : t \geq 0)$ the cumulant semigroup of the critical binary Y-superprocess with branching rate β. We refer our readers to [47] for more details.

Proposition 6.1. *Suppose that* (3.4) *holds, $\beta > 0$ and denote by $(V_t^\beta : t \geq 0)$ the cumulant semigroup of the critical binary Y-superprocess with branching rate β. Let $x_0 \in \mathbb{X}$ be arbitrarily but fixed, then for any $t > 0$ and $x \in \mathbb{X}$, the map*

$$(\theta, n) \mapsto V_t^\beta \left(\theta \mathbf{1}_{B(x_0,n)^c}\right)(x), \quad (\theta, n) \in \mathbb{R}_+^2,$$

is non-decreasing in θ, non-increasing in n, and uniformly bounded by $1/(\beta t)$. Furthermore, for every $t > 0$ and $x \in \mathbb{X}$, we have the iterated limit

$$\lim_{n\uparrow\infty} \lim_{\theta\uparrow\infty} V_t^\beta \big(\theta \mathbf{1}_{B(x_0,n)^c}\big)(x) = 0.$$

The proof of Proposition 6.1 relies on Le Gall's Brownian snake, which we postpone. We also need the following estimates for the solutions of SPDE (2.3), which is a consequence of Theorem 5.5 that requires $(\mathbb{X}, m, d)$ satisfies (2.1), Assumptions 2.3 and 2.5 and inequalities in (3.3) hold.

Lemma 6.2. *Suppose the assumptions of Theorem 3.1 hold. Let u be a $\mathcal{C}_{tem}^+$-valued solution to the equation* (2.3) *for a initial value $u_0 \in \mathcal{C}_{rap}^+$ on a given stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, W)$. Then for every $T \geq 0$,*

$$\sup_{t\in[0,T]} \int_{\mathbb{X}} \mathbb{E}[u_t(x)] m(\mathrm{d}x) + \int_0^t \int_{\mathbb{X}} \mathbb{E}[u_s(y)] m(\mathrm{d}y) \,\mathrm{d}s < \infty.$$

We also need the following proposition, whose proof is will be given at the end of this section.

Proposition 6.3. *Let $t > 0$, $f \in L^\infty(\mathbb{X})$ and $\Phi = (\Phi_r(y) : r \in [0,t], y \in \mathbb{X}) \in L^\infty([0,t] \times \mathbb{X})$. For every $s \in [0,t]$ and $y \in \mathbb{X}$, define $\Phi_t(y) := f(y)$ and*

$$\Phi_s(y) := \int_{\mathbb{X}} p_{t-s}(y, x) f(x) m(\mathrm{d}x) - \int_s^t \int_{\mathbb{X}} p_{r-s}(y, z) \Phi_r(z) \mathrm{m}(dz) dr. \tag{6.3}$$

Then $(\Phi_s(y) : s \in [0,t], y \in \mathbb{X}) \in L^\infty([0,t] \times \mathbb{X})$ and $\Phi_s \in L^\infty(\mathbb{X})$ for every $s \in [0,t]$.

Furthermore, if u is a solution to the equation (2.3) *with initial value $u_0 \in \mathcal{C}_{rap}^+$ on a stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, W)$, then almost surely for every $r \in [0,t]$,*

$$\begin{aligned}\int_{\mathbb{X}} u_r(x)\Phi_r(x)m(\mathrm{d}x) - \int_{\mathbb{X}} u_0(x)\Phi_0(x)m(\mathrm{d}x) = \int_0^r \int_{\mathbb{X}} &\Phi_s(y)u_s(y)\mathrm{d}s\, m(\mathrm{d}y) \\ &+ \int_0^r \int_{\mathbb{X}} \Phi_s(y)b(s,y,u_s(y))\mathrm{d}s\, m(\mathrm{d}y) \\ &+ \int_0^r \int_{\mathbb{X}} \Phi_s(y)\sigma(s,y,u_s(y))W(\mathrm{d}y,\mathrm{d}s).\end{aligned} \tag{6.4}$$

Remark 6.4. We see by Lemma 6.2 that the integrals in (6.4) are all well-defined.

Remark 6.5. Fix an arbitrary $\beta > 0$ and a bounded non-negative Borel function f on $\mathbb{X}$. Note that (6.1) implies

$$V_{t-s}^{\beta} f(y) + \beta \int_s^t \int_{\mathbb{X}} \left(V_{t-r}^{\beta} f(z)\right)^2 P_{r-s}(y,\mathrm{d}z)\mathrm{d}r = P_{t-s}f(y)$$

for every $0 \leq s \leq t < \infty$, and $y \in \mathbb{X}$. Define

$$\Phi_s(y) := V_{t-s}f(y), \quad \text{and} \quad \Phi_s(y) := \beta\left(V_{t-s}f(y)\right)^2, \quad s \in [0,t], y \in \mathbb{X}.$$

Then, (6.3) holds for $t > 0$, $s \in [0,t]$ and $y \in \mathbb{X}$.

Let us explain that Theorem 3.1 is a consequence of Propositions 6.1 and 6.3.

Proof of Theorem 3.1. Let us fix an arbitrary time $t \geq 0$, an arbitrary branching rate $\beta > 0$ and an arbitrary bounded non-negative Borel function f on $\mathbb{X}$. For $f \in L^\infty(\mathbb{X})$, define $(\Phi_s(y) : s \in [0,t), y \in \mathbb{X})$ and $(\Phi_r(z) : r \in [0,t], z \in \mathbb{X})$ according to Remark 6.5 which depends on β implicitly.

Consider the process

$$Z_s := \int_{\mathbb{X}} u_s(y)\Phi_s(y)m(\mathrm{d}y), \quad s \in [0,t].$$

By Proposition 6.3, $(Z_s : s \in [0,t])$ is a semi-martingale whose decomposition is as in (6.4).

Let us also fix some large number $M > 0$ and set $\tau = \tau_M = \inf\{t \geq 0, \sup_{x\in\mathbb{X}} |u_t(x)| \geq K\}$. Since u is a continuous $\mathcal{C}_{rap}^+$-valued solution, it holds $\tau_M \to \infty$ and $M \to \infty$, $\mathbb{P}$-almost surely. Here, we make the constants in (3.3) explicit: assume for all $T > 0$, there are positive constants $C_4 = C_4(T)$ and $C_5 = C_5(T,M)$ so that almost surely for all $(t,x,u) \in [0,T] \times \mathbb{X} \times \mathbb{R}_+$,

$$|b(t,x,u)| \leq C_4|u| \quad \text{and} \quad C_5(u \wedge M)^{1/2} \leq \sigma(t,x,u\wedge M).$$

We set $\widetilde{Z}_s := \exp\{-e^{-C_4 s}Z_s\}e^{-C_4 s}$ for $s > 0$, then by Ito's formula, we have almost surely for every $s \in [0,t]$,

$$\begin{aligned}\exp\{-e^{-C_4 s\wedge\tau}Z_{s\wedge\tau}\} - \exp\{-Z_0\} &= C_4 \int_0^{s\wedge\tau} \widetilde{Z}_r Z_r \mathrm{d}r - \int_0^s \widetilde{Z}_r \mathrm{d}Z_r + \frac{1}{2}\int_0^{s\wedge\tau} \widetilde{Z}_r e^{-C_4 r}\mathrm{d}\langle Z_\cdot\rangle_r \\ &= \int_0^{s\wedge\tau} \int_{\mathbb{X}} \widetilde{Z}_r \Phi_r(y)\left(C_4 u_r(y) - b(u_r(y))\right) m(\mathrm{d}y)\,\mathrm{d}r \\ &\quad + \frac{1}{2}\int_0^{s\wedge\tau} \int_{\mathbb{X}} \widetilde{Z}_r e^{-C_4 r}\Phi_r(y)^2 \sigma(r,y,u_r(y))^2 m(\mathrm{d}y)\mathrm{d}r \\ &\quad - \int_0^{s\wedge\tau} \int_{\mathbb{X}} \widetilde{Z}_r \Phi_r(y) u_r(y) m(\mathrm{d}y)\mathrm{d}r \\ &\quad - \int_0^{s\wedge\tau} \int_{\mathbb{X}} \widetilde{Z}_r \Phi_r(y)\sigma(r,y,u_r(y))W(\mathrm{d}y,\mathrm{d}r).\end{aligned} \tag{6.5}$$

From (3.3), we know that the first term on the right hand side of (6.5) is greater than 0. Therefore, for all $0 \le s \le \tau$,

$$\exp\{-e^{-C_4 s} Z_s\} - \exp\{-Z_0\} \ge \int_0^s \int_{\mathbb{X}} \widetilde{Z}_r \Theta(r,y) m(\mathrm{d}y)\mathrm{d}r - \int_0^s \int_{\mathbb{X}} \widetilde{Z}_r \Phi_r(y)\sigma(r,y,u_r(y)) W(\mathrm{d}y,\mathrm{d}r),$$

where $\Theta(r,y) := \frac{1}{2}e^{-C_4 r}\Phi_r(y)^2\sigma(r,y,u_r(y))^2 - \Phi_r(y)u_r(y)$.

Recall that the branching rate β is arbitrary. Let us take $\beta = \beta(t)$ small enough so that $\frac{C_5^2}{2}e^{-C_4 t} \ge \beta$. Now from (3.3), we have almost surely for every $r \in [0, \tau \wedge t]$ and $y \in \mathbb{X}$,

$$\frac{1}{2}e^{-C_4 r}\Phi_r(y)^2\sigma(r,y,u_r(y))^2 \ge \beta\Phi_r(y)^2 u_r(y) = \Phi_r(y)u_r(y)$$

which implies that $\Theta(r,y) \ge 0$. Then we have $\mathbb{E}\left[\exp\left(-e^{-C_4 s\wedge\tau}Z_{s\wedge\tau}\right) - e^{-Z_0}\right] \ge 0$ for $s \in [0,t]$.

Note that all metric space satisfies (2.1) also satisfies the Hiene-Borel property, and since the test function f can be arbitrary, we may take $f = \theta\mathbf{1}_{B(x_0,n)^c}$ with arbitrary $\theta > 0$, $x_0 \in \mathbb{X}$ and $n > 0$ and get for $t > 0$

$$\begin{aligned}
\mathbb{P}(\mathrm{supp}(u_{t\wedge\tau}) \text{ is a compact}) =& \mathbb{P}(\mathrm{supp}(u_{t\wedge\tau}) \text{ is a } d\text{-bounded}) \\
=& \lim_{n\uparrow\infty}\mathbb{P}\left(\int_{B(x_0,n)^c} u_{t\wedge\tau}(x)m(\mathrm{d}x) = 0\right) \\
=& \lim_{n\uparrow\infty}\lim_{\theta\uparrow\infty}\mathbb{E}\left(\exp\left(-e^{-C_4 t\wedge\tau}\theta\int_{\mathbb{X}}\mathbf{1}_{B(x_0,n)^c}(x)u_{t\wedge\tau}(x)m(\mathrm{d}x)\right)\right) \\
\ge& \lim_{n\uparrow\infty}\lim_{\theta\uparrow\infty}\mathbb{E}\left(\exp\left(-\int_{\mathbb{X}} V^{\beta}_{t\wedge\tau}(\theta\mathbf{1}_{B(x_0,n)^c})(x)u_0(x)m(\mathrm{d}x)\right)\right) \\
=& 1,
\end{aligned}$$

where the last inequality is given by Proposition 6.1. Since it holds for arbitrarily large $M > 0$, the desired result follows. ☐

We now turn to the proof of Proposition 6.1 and Lemma 6.3. Fix arbitrary $\beta > 0$ and $T > 0$, and let $C_1 := C_1(T) > 0$, $p := p(T) \ge 3$ and $\varepsilon := \varepsilon(T) > 0$ be given as in (3.4). Define a new family of transition probability kernels on E by

$$\widetilde{P}_{r,t} = P_{t\wedge T - r\wedge T}, \quad 0 \le r \le t < \infty.$$

It is clear that the semi-group property still hold:

$$\widetilde{P}_{r,s}\widetilde{P}_{s,t} = \widetilde{P}_{r,t}, \quad 0 \le r \le s \le t < \infty.$$

For each $r \ge 0$ and $z \in \mathbb{X}$, define a probability measure $\widetilde{\Pi}_{r,z}$ on the Wiener space $\mathcal{C}([r,\infty);\mathbb{X})$ such that it is the law of the process $(Y_{t\wedge T - r\wedge T} : t \in [r,\infty))$ under Π_z. Denote by $(\widetilde{Y}_{r,t} : t \ge r)$ the canonical process of $\mathcal{C}([r,\infty);\mathbb{X})$ under $(\widetilde{\Pi}_{r,z})_{z\in\mathbb{X}}$ for each $r \ge 0$. Then, it is clear that, for $r \ge 0$ and $z \in \mathbb{X}$, the process $(\widetilde{Y}_{r,t} : t \ge 0)$ under $\Pi_{r,z}$ is a time-inhomogeneous Markov process, initiated at time r and location z, with respect to the transition kernels $(\widetilde{P}_{s,t} : 0 \le s \le t < \infty)$. For every $z \in \mathbb{X}$ and $0 \le r \le t < \infty$, we have from (3.4)

$$\begin{aligned}
\widetilde{\Pi}_{r,z}\left[\sup_{r\le s\le t} d(z,\widetilde{Y}_{r,s})^p\right] &= \Pi_z\left[\sup_{r\le s\le t} d(z, Y_{s\wedge T - r\wedge T})^p\right] \\
&= \Pi_z\left[\sup_{0\le s\le t\wedge T - r\wedge T} d(z,Y_s)^p\right] \\
&\le C_1|t\wedge T - r\wedge T|^{2+\varepsilon} \\
&\le C_1|t-r|^{2+\varepsilon}.
\end{aligned} \tag{6.6}$$

This allows us to define a $\widetilde{Y}$-Brownian snake in the sense of [43, Theorem 1.1] noticing that (6.6) is exactly the [43, Assumption (A)].

Let us provide more details for this $\widetilde{Y}$-Brownian snake. Denote by

$$\mathcal{W} := \bigcup_{t\geq 0} \mathcal{C}([0,t];\mathbb{X})$$

the space of $\mathbb{X}$-valued continuous finite paths. For every $t \geq 0$ and $w \in \mathcal{C}([0,t];\mathbb{X})$, we call $\zeta(w) := t$ the lift-time of w. Define a metric δ on $\mathcal{W}$ such that for every w and w' in $\mathcal{W}$,

$$\delta(w,w') := \sup_{s\geq 0} d\Big(w\big(s\wedge\zeta(w)\big), w'\big(s\wedge\zeta(w')\big)\Big) + |\zeta(w)-\zeta(w')|.$$

It is known that $\mathcal{W}$, equipped with the metric δ, is a Polish space. The $\widetilde{Y}$-Brownian snake $(W_t)_{t\geq 0}$ is a continuous time-homogeneous strong Markov process taking values in $\mathcal{W}$. In particular, $(\zeta(W_t) : t \geq 0)$ is a reflected Brownian motion. The transition semigroup of the $\widetilde{Y}$-Brownian snake is given as in [43, P. 29]. For our purpose, we do not need the precise form of this semigroup. Rather, we want to present a connection between the $\widetilde{Y}$-Brownian snakes and the $\widetilde{Y}$-superprocesses.

Similar to the Y-superprocess mentioned in the beginning of this Section, we can construct the $\widetilde{Y}$-superprocess with only a slightly difference due to the fact that $\widetilde{Y}$ is time-inhomogeneous. For any $t \geq 0$ and bounded non-negative Borel function f on $\mathbb{X}$, there exists a unique locally-bounded non-negative Borel function $(\widetilde{V}^{\beta}_{r,t}f(x) : r \in [0,t], x \in \mathbb{X})$ such that

$$\widetilde{V}^{\beta}_{r,t}f(x) + \beta\int_r^t \mathrm{d}s\int_{\mathbb{X}} \big(\widetilde{V}^{\beta}_{s,t}f(y)\big)^2 \widetilde{P}_{r,s}(x,\mathrm{d}y) = \widetilde{P}_{r,t}f(x), \quad r\in[0,t], x\in\mathbb{X}.$$

It is known that there exists a family of probability transition kernels $(\widetilde{Q}^{\beta}_{r,t} : 0 \leq r \leq t < \infty)$ on $\mathcal{M}(\mathbb{X})$, satisfying the semi-group property:

$$\widetilde{Q}^{\beta}_{r,s}\widetilde{Q}^{\beta}_{s,t} = \widetilde{Q}^{\beta}_{r,t}, \quad 0 \leq r \leq s \leq t < \infty,$$

such that for each $0 \leq r \leq t < \infty, \mu \in \mathcal{M}$ and bounded non-negative Borel function f on $\mathbb{X}$,

$$\int_{\mathcal{M}} \exp\Big\{-\int_{\mathbb{X}} f(x)\eta(\mathrm{d}x)\Big\}\widetilde{Q}^{\beta}_{r,t}(\mu,\mathrm{d}\eta) = \exp\Big\{-\int_{\mathbb{X}} \widetilde{V}^{\beta}_{r,t}f(x)\mu(\mathrm{d}x)\Big\}. \tag{6.7}$$

We say a process is a critical binary $\widetilde{Y}$-superprocess with branching rate β, if it is an $\mathcal{M}$-valued time-inhomogeneous Markov process with transition kernels $(\widetilde{Q}^{\beta}_{r,t} : 0 \leq r \leq t < \infty)$.

Remark 6.6. Let f be a bounded non-negative Borel function on $\mathbb{X}$. From (6.1), we have for $x \in \mathbb{X}$

$$V^{\beta}_{t-r}f(x) + \beta\int_0^{t-r}\mathrm{d}s\int_{\mathbb{X}}\big(V^{\beta}_{t-r-s}f(y)\big)^2 P_s(x,\mathrm{d}y) = P_{t-r}f(x),$$

which can be rewritten as

$$V^{\beta}_{t-r}f(x) + \beta\int_r^{t}\mathrm{d}s\int_{\mathbb{X}}\big(V^{\beta}_{t-s}f(y)\big)^2 P_{s-r}(x,\mathrm{d}y) = P_{t-r}f(x)$$

for every $0 \leq r \leq t < \infty$ and $x \in \mathbb{X}$. Notice that if $0 \leq r \leq t \leq T$, then $\widetilde{P}_{r,t} = P_{t-r}$. So we actually have

$$V^{\beta}_{t-r}f(x) + \beta\int_r^{t}\mathrm{d}s\int_{\mathbb{X}}\big(V^{\beta}_{t-s}f(y)\big)^2 \widetilde{P}_{r,s}(x,\mathrm{d}y) = \widetilde{P}_{r,t}f(x),$$

which implies that $V^{\beta}_{t-r}f(x) = \widetilde{V}^{\beta}_{r,t}f(x)$, for every $0 \leq r \leq t \leq T$ and $x \in \mathbb{X}$.

Lemma 6.7 ([43, Theorem 2.1 & P. 36])**.** *Let $x \in \mathbb{X}$. Suppose that $(W_t : t \geq 0)$ is a $\widetilde{Y}$-Brownian snake with initial values W_0 satisfying that $\zeta(W_0) = 0$ and $W_0(0) = x$. Let $(L^z_t : t \geq 0, z \geq 0)$ be the non-negative continuous random field such that, for each $t \geq 0$ and $z \geq 0$, L^z_t is the local time of the reflected Brownian motion $(\zeta(W_r) : r \geq 0)$ accumulated at value z up to time t. Define $\sigma := \inf\{t \geq 0 : L^0_t \geq 2\beta\}$*

and an $\mathcal{M}$-valued process $(\xi_z : z \geq 0)$ such that for every $z \geq 0$, almost surely for every non-negative bounded measurable function φ on $\mathbb{X}$,

$$\int_{\mathbb{X}} \varphi(x)\xi_z(\mathrm{d}x) = \frac{1}{2\beta}\int_0^{\sigma} \varphi(W_t(z))\mathrm{d}L_t^z.$$

Then $(\xi_z : z \geq 0)$ is a critical binary $\widetilde{Y}$-superprocess with branching rate β and initial value δ_x. Moreover,

$$\bigcup_{z\geq 0} \operatorname{supp}(\xi_z) = \{W_t(\zeta(W_t)) : 0 \leq t \leq \sigma\}. \tag{6.8}$$

Proof of Proposition 6.1. For a given $t > 0$, $x \in \mathbb{X}$, $\theta \in (0,\infty)$ and $n \in [0,\infty)$, replacing μ by Dirac measure δ_x and f by $\theta \mathbf{1}_{B(x_0,n)^c}$ in (6.2), we obtain formula

$$V_t^{\beta}\big(\theta\mathbf{1}_{B(x_0,n)^c}\big)(x) = -\log\left(\int_{\mathcal{M}} \exp\Big\{-\int_{\mathbb{X}} \theta\mathbf{1}_{B(x_0,n)^c}(x)\eta(\mathrm{d}x)\Big\}Q_t^{\beta}(\delta_x,\mathrm{d}\eta)\right)$$

which says that the quantity $V_t^{\beta}\big(\theta\mathbf{1}_{B(x_0,n)^c}\big)(x)$ is monotonically non-decreasing in θ and non-increasing in n. We can verify for each $t > 0$, $x \in \mathbb{X}$ and $\theta \in (0,\infty)$,

$$\begin{aligned} V_t^{\beta}\big(\theta\mathbf{1}_{B(x_0,n)^c}\big)(x) &\leq V_t^{\beta}(\theta\mathbf{1}_{\mathbb{X}})(x) \leq \lim_{\theta\uparrow\infty} V_t^{\beta}(\theta\mathbf{1}_{\mathbb{X}})(x) \\ = -\log\int_{\mathcal{M}} \mathbf{1}_{\{\eta(\mathbb{X})=0\}}Q_t^{\beta}(\delta_x,\mathrm{d}\eta) &= -\log \mathbf{P}_{\delta_x}(X_t(\mathbb{X})=0). \end{aligned}$$

Notice that, the process of the total population $(X_t(\mathbb{X}))_{t\geq 0}$ is a critical binary continuous-state branching process with branching rate β, whose extinction probability is given by

$$\mathbf{P}_{\delta_x}(X_t(\mathbb{X})=0) = \exp\left(-\frac{1}{\beta t}\right), \quad t > 0.$$

Therefore we have a uniform bound

$$\sup_{\theta\in(0,\infty),n\in[0,\infty),x\in\mathbb{X}} V_t\big(\theta\mathbf{1}_{B(x_0,n)^c}\big)(x) \leq \frac{1}{\beta t}, \quad t > 0.$$

Let us now fix an arbitrary $x \in \mathbb{X}$. Let the $\widetilde{Y}$-Brownian snake $(W_t : t \geq 0)$ and the $\widetilde{Y}$-superprocess $(\xi_z : z \geq 0)$ be given as in Lemma 6.7. Since $(W_t : t \geq 0)$ is a $\mathcal{W}$-valued continuous process, we can verify that $(W_t(\zeta(W_t)) : t \geq 0)$ is a continuous $\mathbb{X}$-valued process. Therefore, the right hand side of (6.8) is a compact subset of $\mathbb{X}$. In particular, the random measure ξ_t has compact support for every $t \geq 0$ almost surely.

Similar to the last step in the proof of Theorem 3.1, we can verify via Heine–Borel that for every $t > 0$,

$$\begin{aligned} 1 &= \mathbb{P}\left(\operatorname{Supp}\xi_t \text{ is a compact subset of } \mathbb{X}\right) \\ &= \mathbb{P}\left(\operatorname{Supp}\xi_t \text{ is a } d\text{-bounded subset of } \mathbb{X}\right) \\ &= \lim_{n\uparrow\infty} \mathbb{P}\left(\int_{\mathbb{X}} \mathbf{1}_{B(x_0,n)^c}(x)\xi_t(\mathrm{d}x) = 0\right) \\ &= \lim_{n\uparrow\infty}\lim_{\theta\uparrow\infty} \mathbb{P}\left(\exp\left(-\theta\int_{\mathbb{X}} \mathbf{1}_{B(x_0,n)^c}(x)\xi_t(\mathrm{d}x)\right)\right) \\ &\overset{(6.7)}{=} \lim_{n\uparrow\infty}\lim_{\theta\uparrow\infty} \exp\left(-\widetilde{V}_{0,t}^{\beta}\big(\theta\mathbf{1}_{B(x_0,n)^c}\big)(x)\right). \end{aligned}$$

This says that

$$\lim_{n\uparrow\infty}\lim_{\theta\uparrow\infty} \widetilde{V}_{0,t}^{\beta}\big(\theta\mathbf{1}_{B(x_0,n)^c}\big)(x) = 0.$$

From Remark 6.6, we can verify that the desired result holds provided $t \leq T$. Recalling that T is arbitrary, we are done. $\square$

Proof of Proposition 6.3. Recall for $t > 0$ that $(\Phi_r(x),\, r \in [0,t],\, x \in \mathbb{X})$ in (6.3). For solution u to (2.3), consider for $r \in [0,t]$

$$\begin{aligned}\int_{\mathbb{X}} \Phi_r(x) u_r(x) m(\mathrm{d}x) = &\int_{\mathbb{X}} \int_{\mathbb{X}} p_r(x,y) u_0(y) m(\mathrm{d}y)\, \Phi_r(y) m(\mathrm{d}x) \\ &+ \int_{\mathbb{X}} \int_0^r \int_{\mathbb{X}} p_{r-s}(x,y) b(u_s(y)) m(\mathrm{d}y)\, \mathrm{d}s\, \Phi_r(x) m(\mathrm{d}x) \\ &+ \int_{\mathbb{X}} \int_0^r \int_{\mathbb{X}} p_{r-s}(x,y) \sigma(u_s(y)) W(\mathrm{d}y, \mathrm{d}s)\, \Phi_r(x) m(\mathrm{d}x).\end{aligned}$$

We will call the three terms on the right hand side $I_1(r), I_2(r)$ and $I_3(r)$ respectively. Note by the semi-group property and Fubini's, we see for all $\tau \in (0,t]$ and $r \in [0,t)$,

$$\int_{\mathbb{X}} p_\tau(x,y) \Phi_r(x) \mathrm{d}x = \int_{\mathbb{X}} p_{t-r+\tau}(y,z) f(z) \mathrm{d}z - \int_r^t \int_{\mathbb{X}} p_{a-r+\tau}(y,z) \Phi_a(z) \mathrm{d}z\, \mathrm{d}a.$$

Since $u_0 \in \mathcal{C}^+_{rap}$, we consider first by Fubini's that

$$\begin{aligned} I_1(r) &= \int_{\mathbb{X}} u_0(y) \int_{\mathbb{X}} p_t(y,z) f(z) m(\mathrm{d}z) m(\mathrm{d}y) - \int \int_r^t \int_{\mathbb{X}} p_a(y,z) \Phi_a(z) g(y) m(\mathrm{d}z)\, \mathrm{d}a\, m(\mathrm{d}y) \\ &= \int_{\mathbb{X}} u_0(y) \Phi_0(y) m(\mathrm{d}y) + \int_0^r \int_{\mathbb{X}} P_a u_0(z) \Phi_a(z) m(\mathrm{d}z)\, \mathrm{d}a.\end{aligned}$$

Similarly, by Lemma 6.2 and Fubini's, we have the following almost sure equality

$$\begin{aligned} I_2(r) &= \int_0^r \int_{\mathbb{X}} \int_{\mathbb{X}} p_{r-s}(x,y) \Phi_r(x)\, m(\mathrm{d}x)\, b(u_s(y)) m(\mathrm{d}y)\, \mathrm{d}s \\ &= \int_0^r \int_{\mathbb{X}} b(u_s(y)) \left(\int p_{t-s}(y,z) f(z) m(\mathrm{d}z) - \int_r^t \int_{\mathbb{X}} p_{a-s}(y,z) \Phi_a(z) m(\mathrm{d}z)\, \mathrm{d}a \right) m(\mathrm{d}y)\, \mathrm{d}s \\ &= \int_0^r \int_{\mathbb{X}} b(u_s(y)) \left(\Phi_s(y) + \int_s^r \int_{\mathbb{X}} p_{a-s}(y,z) \Phi_a(z) m(\mathrm{d}z)\, \mathrm{d}a \right) m(\mathrm{d}y)\, \mathrm{d}s.\end{aligned}$$

Finally, by Lemma 6.2 and stochastic Fubini's, Theorem 8.2 (which easily extends to σ-compact case), we see almost surely for each $r \in [0,t)$ that

$$I_3(r) = \int_0^r \int_{\mathbb{X}} \sigma(u_s(y)) \left(\Phi_s(y) + \int_s^r \int_{\mathbb{X}} p_{a-s}(y,z) \Phi_a(z) m(\mathrm{d}z)\, \mathrm{d}a \right) W(\mathrm{d}y, \mathrm{d}s).$$

Argue similarly as above, we see

$$\begin{aligned} &\int_0^r \int_{\mathbb{X}} b(u_s(y)) \int_s^r \int_{\mathbb{X}} p_{a-s}(y,z) \Phi_a(z)\, m(\mathrm{d}z)\, \mathrm{d}s\, m(\mathrm{d}y)\, \mathrm{d}s \\ = &\int_0^r \int_{\mathbb{X}} \Phi_a(z) \left(\int_0^a p_{a-s}(y,z) b(u_s(y)) m(\mathrm{d}y)\, \mathrm{d}s \right) m(\mathrm{d}z)\, \mathrm{d}s\end{aligned}$$

and

$$\begin{aligned} &\int_0^r \int_{\mathbb{X}} \sigma(u_s(y)) \int_s^r \int_{\mathbb{X}} p_{a-s}(y,z) \Phi_a(z) m(\mathrm{d}z)\, \mathrm{d}a W(\mathrm{d}y, \mathrm{d}s) \\ = &\int_0^r \int_{\mathbb{X}} \left(\int_0^a \int_{\mathbb{X}} p_{a-s}(z,y) \sigma(u_s(y)) W(\mathrm{d}y, \mathrm{d}s) \right) \Phi_a(z) m(\mathrm{d}z)\, \mathrm{d}a.\end{aligned}$$

Collecting all the terms above to see the desired equality. □

To summarize the proof of Theorem 3.1, we used the theory of Y-supper-processes to construct a time-dependent (non-linear) integral equation (6.1) and used Le Gall's Brownian snake and showed in Proposition 6.1 that the solution to (6.1) have similar properties to that of the one-dimensional ODE of Iscoe used by Shiga in [55]. Finally, the proof of Theorem 3.1 is done by integrating the solution of

(2.3) against a time-dependent test function as in (6.1) combining with the martingale representation in Proposition 6.3.

7. Proof of Theorem 3.6

Since Assumption 2.3 and Assumption 2.6 hold, the existence of a unique $\mathcal{C}^+_{tem}$-valued solution is then given shown in Theorem 5.1, we will only prove the strong comparison principle.

Our proof of uses the methods in [51], which is based on a large deviation estimates for stochastic integrals with respect to a space-time white noise on $\mathbb{R}_+ \times \mathbb{R}$. We will generalize the result to Gaussian noise on metric measure spaces satisfies (2.1).

We first observe for u_1 and u_2 as in Theorem 3.6, denote $u_t(x) := u_1(t,x) - u_2(t,x)$ and $f(x) := u_1(0,x) - u_2(0,x)$ for $t > 0$ and $x \in \mathbb{X}$. Then u solves the integral equation

$$u_t(x) = P_t f(x) + \int_0^t \int_{\mathbb{X}} p_t(x,y)\widetilde{b}(t,y,u_t(y))m(\mathrm{d}y)\mathrm{d}s + \int_0^t \int_{\mathbb{X}} p_t(x,y)\widetilde{\sigma}(t,x,u_t(x))W(\mathrm{d}y,\mathrm{d}s),$$

for $t > 0$, $x \in \mathbb{X}$ and $u \in \mathbb{R}$, where

$$\widetilde{b}(t,x,u) := b(t,x,u(t,x) + u_2(t,x)) - b(t,x,u_2(t,x)), \quad \widetilde{\sigma}(t,x,u) := \sigma(t,x,u + u_2(t,x)) - \sigma(t,x,u_2(t,x)).$$

Since b and σ satisfy (3.5), so does $\widetilde{b}$ and $\widetilde{\sigma}$, and almost surely for all $t > 0$, $x \in \mathbb{X}$ and $u \in \mathbb{R}$,

$$|\widetilde{b}(t,x,u)| + |\widetilde{\sigma}(t,x,u)| \le L|u|.$$

Therefore, by Corollary 5.3, it is enough to assume $b(t,x,u)$ is identically zero for all $(t,x,u) \in \mathbb{R}_+ \times \mathbb{X} \times \mathbb{R}$ and show for σ satisfies (3.5) and $\sigma(t,x,0) = 0$ almost surely for all $(t,x) \in \mathbb{R}_+ \times \mathbb{X}$, if u is the solution to the following integral equation

$$u_t(x) = P_t f(x) + \int_0^t \int_{\mathbb{X}} p_{t-s}(x,y)\sigma(s,y,u_s(y))W(\mathrm{d}y\,\mathrm{d}s), \quad (t,x) \in \mathbb{R}_+ \times \mathbb{X}, \tag{7.1}$$

with $f \in \mathcal{C}^+_c$ and $f(x) > 0$ for some $x \in \mathbb{X}$, then $\mathbb{P}\left(u_t(x) > 0, \text{ for every } t > 0,\, x \in \mathbb{X}\right) = 1$.

The key estimate we will be using is the following, similar to that of [51], which will allow us to have tail probability estimates of stochastic integrals globally. We will postpone its proof to the end of this section.

Lemma 7.1. *Under the assumptions of Theorem 3.6 with $\lambda_0 > 0$ as in (2.5) and $\alpha \in (0,1)$ as in (2.8), let $\lambda > 0$ and $\eta : \mathbb{R}_+ \times \mathbb{X} \to \mathbb{R}_+$ be a deterministic measurable function. Supposed for each $T > 0$, the following inequality holds uniformly in $(t,x) \in [0,T] \times \mathbb{X}$*

$$\sup_{s,s' \in [0,t],\, y,y' \in B(x,1)} \frac{\int_0^{s \vee s'} \|(p_{s-\tau}(y,\cdot) - p_{s'-\tau}(y',\cdot))\,\eta(\tau,\cdot)\|^2_{\widetilde{K}}\,\mathrm{d}\tau}{d(y,y')^{\xi_1} + |s-s'|^{\xi_2}} \lesssim h(x)^{2\lambda - \lambda_0} t^{1-\alpha}.$$

Suppose $(f(t,x),\, (t,x) \in \mathbb{R}_+ \times \mathbb{X})$ is a predictable random field so that $|f(t,x)| \le \eta(t,x)$ almost surely for all $(t,x) \in [0,T] \times \mathbb{X}$. If we let

$$N(t,x) := \int_0^t \int_{\mathbb{X}} p_{t-s}(x,y) f(s,y) W(\mathrm{d}y\,\mathrm{d}s), \quad (t,x) \in \mathbb{R}_+ \times \mathbb{X},$$

then for all $T > 0$, there exists $C_6 = C_6(T,\lambda) > 0$ so that

$$\mathbb{P}\left(\sup_{(s,x) \in [0,t] \times \mathbb{X}} |N(s,x)| h(x)^{-\lambda} > \varepsilon\right) \lesssim \varepsilon^{-2} t^{1-\alpha} \exp\left(-\frac{C_6}{2t^{1-\alpha}}\varepsilon^2\right), \quad \textit{uniformly in } t \in (0,T],\, \varepsilon \in (0,1).$$

Next we state an elementary lemma that can be verified with straight forward computations.

Lemma 7.2. *Under the assumptions of Theorem 3.6, there exists $\xi > 0$ so that for each $\lambda \in [0, 1/2]$ and $T > 0$, the following inequality holds uniformly for $x, x' \in \mathbb{X}$ with $d(x, x') \leq 1$ and $0 \leq t \leq t' \leq T$,*

$$\int_0^{t'} \left\| (p_{t-s}(x, \cdot) - p_{t'-s}(x', \cdot)) h(\cdot)^{\lambda} \right\|_{\widetilde{K}}^2 \mathrm{d}s \lesssim t'^{1-\alpha} \left(|t - t'|^{\xi} + d(x, x')^{\xi} \right) h(x)^{2\lambda},$$

where $\alpha \in (0, 1)$ is as in (2.8).

Let u be the solution to the integral equation (7.1) and set for $(t, x) \in \mathbb{R}_+ \times \mathbb{X}$

$$Z(t, x) = \int_0^t \int_{\mathbb{X}} p_{t-s}(x, y) \sigma(s, y, u_s(y)) W(\mathrm{d}y, \mathrm{d}s). \tag{7.2}$$

Lemma 7.3. *Under the assumptions of Theorem 3.6 and suppose u is a $\mathcal{C}^+_{tem}$-valued solution to the integral equation* (7.1) *with initial value $0 \leq f(x) \leq \beta$ for all $x \in \mathbb{X}$ and some $\beta > 0$. For every $\lambda > \frac{\lambda_0}{2}$, there exist a constant $C_7 = C_7(\lambda, L) > 0$ so that uniformly for $\varepsilon \in (0, 1)$ and $T \in (0, 1]$,*

$$\mathbb{P}\left(\sup_{(t,x) \in (0,T] \times \mathbb{X}} |Z(t, x)| \, h(x)^{-\lambda} \geq \varepsilon\beta \right) \lesssim \varepsilon^{-2} T^{1-\alpha} \exp\left(-\frac{C_7}{2T^{1-\alpha}} \varepsilon^2 \right)$$

where $\alpha \in (0, 1)$ is as in (2.8).

Proof. Let $\lambda \in (0, 1/2]$ be arbitrary but fixed through out this proof. Since $h(x) \geq 1$ for all $x \in \mathbb{X}$, we see for all $\lambda \in [0, 1]$ and $(t, x) \in \mathbb{R}_+ \times \mathbb{X}$ that

$$\int_{\mathbb{X}} p_t(x, y) f(y) m(\mathrm{d}y) \leq \beta \leq \beta h(x)^{\lambda}.$$

Let

$$\tau := \inf \left\{ t > 0 : u_t(x) \geq 2\beta h(x)^{\lambda}, \ \text{ for some } x \in \mathbb{X} \right\},$$

and

$$M(t, x) := Z(t \wedge \tau, x) = \int_0^t \int_{\mathbb{X}} 1_{s<\tau} p_{t-s}(x, y) \sigma(s, y, u_s(y)) W(\mathrm{d}y, \mathrm{d}s).$$

Note that for $\varepsilon \in (0, 1)$, on the event $A_\varepsilon := \left\{ |M(t, x)| \leq \varepsilon\beta h(x)^{\lambda}, \ \text{ for all } (t, x) \in [0, 1] \times \mathbb{X} \right\}$, we see

$$\begin{aligned} u_{t\wedge\tau}(x) &\leq \int_{\mathbb{X}} p_{t\wedge\tau}(x, y) f(y) m(\mathrm{d}y) + |Z(t \wedge \tau, x)| \\ &\leq \beta h(x)^{\lambda} + |M(t, x)| \\ &\leq \beta (1 + \varepsilon) h(x)^{\lambda} \\ &< 2\beta h(x)^{\lambda}. \end{aligned}$$

Therefore $t \wedge \tau < \tau$ on the event A_ε, meaning $t < \tau$ by continuity of u. Hence

$$|Z(t, x)| = |M(t, x)| \leq \varepsilon\beta h(x)^{\lambda}, \qquad \text{for all } (t, x) \in [0, 1] \times \mathbb{X} \text{ on } A_\varepsilon.$$

By (3.5), we see $1_{s<\tau} \sigma(u_s(x)) \leq 2L\beta h(x)^{\lambda}$ almost surely for all $(s, x) \in [0, 1] \times \mathbb{X}$. Hence by Lemma 7.2 and Lemma 7.1,

$$\begin{aligned} \mathbb{P}\left(\sup_{(t,x) \in [0,T] \times \mathbb{X}} |Z(t, x)| h(x)^{-(\lambda + \frac{\lambda_0}{2})} > \varepsilon\beta \right) &\leq \mathbb{P}\left(A_\varepsilon^c \right) \\ &= \mathbb{P}\left(\sup_{(t,x) \in [0,T] \times \mathbb{X}} |M(t, x)| h(x)^{-(\lambda + \frac{\lambda_0}{2})} > \varepsilon\beta \right) \\ &\lesssim \varepsilon^{-2} T^{1-\alpha} \exp\left(-\frac{C_7}{2T^{1-\alpha}} \varepsilon^2 \right), \end{aligned}$$

uniformly for $\varepsilon \in (0,1)$, $T \in [0,1]$ and $\beta > 0$ for some C_7 depend only on λ and L, which is the desired result. $\square$

Proof of Theorem 3.6. Let u be the solution to the integral equation (7.1) with initial value $f \in \mathcal{C}^+_{temp}$ and $f(x) > 0$ for some $x \in \mathbb{X}$. By Corollary 5.3, it is enough to assume for some $x_0 \in \mathbb{X}$ and $a > 0$ that

$$2\beta \geq f(x) \geq \beta 1_{B(x_0,a)}(x), \qquad \text{for all } x \in \mathbb{X}.$$

Recall Z in (7.2), let λ_0 be as in (2.5) and fix some $\lambda > \frac{\lambda_0}{2}$. Let $t > 0$, by Lemma 7.3, there are constants $C_8, C_9 > 0$, depending on λ, so that for all $1 > \varepsilon > 0$, $Q \geq 2t$,

$$\mathbb{P}\left(|Z(s,x)| > \varepsilon\beta h(x)^\lambda \text{ for some } (s,x) \in \left[0, \frac{2t}{Q}\right] \times \mathbb{X}\right) \leq C_8 \varepsilon^{-2} \left(\frac{t}{Q}\right)^{1-\alpha} \exp\left(-C_9\varepsilon^2 \left(\frac{Q}{t}\right)^{1-\alpha}\right) \tag{7.3}$$

Let $M > 0$ be arbitrary but fixed and $Q \in \mathbb{N}$ be large enough so that (3.6) holds for some $N_0 = N_0(x_0, M, t, a)$, then

$$\inf_{s\in\left[\frac{t}{Q},\frac{2t}{Q}\right]} \inf_{x\in B(x_0,a+M)} P_s f(x) \geq \frac{\beta}{N_0}.$$

Since for each $t \geq 0$ and $x \in \mathbb{X}$, $u_t(x) = P_t f(x) + Z(t,x)$ almost surely, it follows from Lemma 7.3 that, if we set $H(\lambda, M) = \sup_{x\in B(x_0,a+M)} h(x)^\lambda$, then

$$\begin{aligned}
&\mathbb{P}\left(u(s,x) < \frac{\beta}{2N_0} \text{ for some } (s,x) \in \left[\frac{t}{Q}, \frac{2t}{Q}\right] \times B(x_0, a + M/Q)\right) \\
\leq\ &\mathbb{P}\left(|Z(s,x)| > \frac{\beta}{2N_0} \text{ for some } (s,x) \in \left[\frac{t}{Q}, \frac{2t}{Q}\right] \times B(x_0, a + M/Q)\right) \\
\leq\ &\mathbb{P}\left(|Z(s,x)| > \frac{\beta}{2N_0}\frac{h(x)^\lambda}{H(\lambda,M)} \text{ for some } (s,x) \in \left[\frac{t}{Q}, \frac{2t}{Q}\right] \times B(x_0, a + M/Q)\right) \\
\leq\ &\mathbb{P}\left(|Z(s,x)| > \frac{\beta}{2N_0}\frac{h(x)^\lambda}{H(\lambda,M)} \text{ for some } (s,x) \in \left[\frac{t}{Q}, \frac{2t}{Q}\right] \times \mathbb{X}\right) \\
\overset{(7.3)}{\leq}\ &C_8 N_0^2 H(\lambda,M)^2 \left(\frac{t}{Q}\right)^{1-\alpha} \exp\left(-C_9 N_0^{-2} H(\lambda,M)^{-2} \left(\frac{Q}{t}\right)^{1-\alpha}\right)
\end{aligned}$$

In other words, If we let $H := N_0 H(\lambda, M)$, then for all large enough Q,

$$\begin{aligned}
&\mathbb{P}\left(u_s(x) \geq \frac{\beta}{2N_0} 1_{B(x_0,a+M/Q)}(x) \text{ for all } (s,x) \in \left[\frac{t}{Q}, \frac{2t}{Q}\right] \times \mathbb{X}\right) \\
\geq\ &1 - C_8 H^2 \left(\frac{t}{Q}\right)^{1-\alpha} \exp\left(-C_9 H^{-2} \left(\frac{Q}{t}\right)^{1-\alpha}\right) > 0.
\end{aligned} \tag{7.4}$$

Now define for $1 \leq k \leq Q$,

$$\mathcal{A}_k := \left\{u(s,x) \geq \beta \frac{1}{(2N_0)^k} 1_{B(x_0,a+kM/Q)}(x), \text{ for all } (s,x) \in \left[\frac{kt}{Q}, \frac{(k+1)t}{Q}\right] \times \mathbb{X}\right\}.$$

For $k = 2, \ldots, Q$, note that on $\mathcal{A}_{k-1}$,

$$u\left(\frac{tk}{Q}, x\right) \geq \beta \frac{1}{(2N_0)^k} 1_{B(x_0,a+kM/Q)}(x), \qquad \text{for all } x \in \mathbb{X}.$$

Hence by Markov property, one can repeat the above argument for (7.4) to see that

$$\mathbb{P}\left(\mathcal{A}_k \middle| \bigcap_{i=1}^{k-1} \mathcal{A}_i\right) = \mathbb{P}(\mathcal{A}_k | \mathcal{A}_{k-1}) \geq 1 - C_8 H^2 \left(\frac{t}{Q}\right)^{1-\alpha} \exp\left(-C_9 16^{-1} H^{-2} \left(\frac{Q}{t}\right)^{1-\alpha}\right).$$

This implies

$$\begin{aligned}\mathbb{P}\left(u_s(x)>0 \text{ on } \left[\frac{t}{Q},t\right]\times B(x_0,a+M)\right) &\geq \mathbb{P}\left(\bigcap_{k=1}^{Q}\mathcal{A}_k\right)\\ &\geq \left(1-C_8H^2\left(\frac{t}{Q}\right)^{1-\alpha}\exp\left(-C_9 16^{-1}H^{-2}\left(\frac{Q}{t}\right)^{1-\alpha}\right)\right)^Q.\end{aligned}$$

We may take $Q\to\infty$ to see

$$\mathbb{P}\left(u_s(x)>0 \text{ on } (0,t)\times B(x_0,a+M)\right)=1.$$

Take $t,M\uparrow\infty$ to see the desired result. □

Now we turn to the proof of Lemma 7.1. We first prove a local version of the Lemma that is similar to that of [52, Theorem 4.2] (also see [6, Lemma 9.1]). The proof relies on the chaining technique of balls in metric spaces that can be found, for example, in [46] and [41, Chapter 2.4].

Theorem 7.4. *Suppose $(\mathbb{X},d,m)$ is a metric measure space and $\dot W$ is a Gaussian noise with covariance measure K so that $|K|\leq\widetilde K$ for some non-negative and non-negative definite symmetric measure $\widetilde K$ on $\mathbb{X}^2$. Let $T,R>0$, $x_0\in\mathbb{X}$ and assume the following assumptions hold:*

(1) There is some $\beta_0>0$ so that $D(B(x_0,R),d;\varepsilon)\lesssim\varepsilon^{-\beta_0}$ uniformly for $\varepsilon\in(0,1]$, where D is defined as in (2.1).

(2) $g_{t,x}(s,y):=g(t,s,x,y)$ on $\mathbb{R}_+\times\mathbb{R}_+\times\mathbb{X}\times\mathbb{X}$ and $\eta(s,y)$ on $\mathbb{R}_+\times\mathbb{X}$ are non-negative deterministic real-valued functions. There are $\xi_1,\xi_2\in(0,1]$ so that

$$B:=\sup_{t,t'\in[0,T],\,x,x'\in B(x_0,R)}\frac{\int_0^\infty\|(g_{t,x}(s,\cdot)-g_{t',x'}(s,\cdot))\,\eta(s,\cdot)\|_{\widetilde K}^2\,\mathrm{d}s}{\left(\frac{d(x,x')}{2R}\right)^{\xi_1}+\left|\frac{t-t'}{T}\right|^{\xi_2}}<\infty.$$

(3) σ is a predictable random field on $\mathbb{R}_+\times\mathbb{X}$ so that

$$|\sigma(t,x)|\leq\eta(t,x),\quad \textit{almost surely for all } (t,x)\in\mathbb{R}_+\times\mathbb{X}.$$

For $(t,x)\in[0,T]\times B(x_0,R)$, let

$$N(t,x):=\int_0^\infty\int_{\mathbb{X}}g(t,s,x,y)\sigma(s,y)W(\mathrm{d}s,\mathrm{d}y). \tag{7.5}$$

Then there is a joint continuous modification $\widetilde N(t,x)$ of $N(t,x)$ on $[0,T]\times B(x_0,R)$, so that for $c:=\sqrt{8}\sum_{n=1}^\infty\sqrt{n}2^{-n/2}$, the following inequality holds uniformly in $\lambda>0$

$$\mathbb{P}\left(\sup_{(t,x),(t',x')\in[0,T]\times B(x_0,R)}\left|\widetilde N(t,x)-\widetilde N(t',x')\right|>\lambda\right)\lesssim\exp\left(-\frac{\lambda^2}{2c^2B}\right). \tag{7.6}$$

We start with the following lemma.

Lemma 7.5. *Under the assumptions and notations of Theorem 7.4, for all $0\leq s\leq t\leq T$, $x,y\in B(x_0,R)$, and $\lambda>0$,*

$$\mathbb{P}\left(|N(t,x)-N(t',x')|\geq\sqrt{B}\lambda\right)\leq 4\exp\left(-\frac{\lambda^2}{2\left(\left(\frac{d(x,x')}{2R}\right)^{\xi_1}+\left|\frac{|t-t'|}{T}\right|^{\xi_2}\right)}\right).$$

Proof. For $t\in[0,T]$ and $x,x'\in\mathbb{X}$, we define for $\tau\in[0,T]$

$$N_{t,x}(\tau):=\int_0^\tau\int_{\mathbb{X}}g(t,s,x,y)\sigma(s,y)W(\mathrm{d}y\,\mathrm{d}s)$$
$$M_{(t,t'),(x,x')}(\tau):=N_{t,x}(\tau)-N_{t',x'}(\tau).$$

Then $\{M_{(t,t'),(x,x')}(\tau),\, \tau \in [0,T]\}$ is a continuous martingale for each $t,t' \in [0,T]$ and $x,x' \in \mathbb{X}$, and it holds almost surely for all $\tau \in [0,T]$ that

$$\begin{aligned}\left\langle M_{(t,t'),(x,x')}\right\rangle_\tau &\le \int_0^\tau \|(g_{t,x}(s,\cdot) - g_{t',x'}(s,\cdot))\,\sigma(s,\cdot)\|^2_{\widetilde{K}}\,\mathrm{d}s \\ &\le \left(\left(\frac{d(x,x')}{2R}\right)^{\xi_1} + \left|\frac{t-t'}{T}\right|^{\xi_2}\right) B,\end{aligned}$$

and hence exists a Brownian motion $(B_t)_{t\ge 0}$ such that $M_{(t,t'),(x,x')}(\tau) = B_{\langle M_{(t,t'),(x,x')}(\tau)\rangle}$ almost surely for each $\tau \in [0,T]$. By the reflection principle of Brownian motion, we see for each $t,t' \in [0,T]$ and $x,x' \in \mathbb{X}$,

$$\begin{aligned}\mathbb{P}\left(\left|M_{(t,t'),(x,x')}(\tau)\right| \ge \sqrt{B}\lambda\right) &= \mathbb{P}\left(\left|B_{\langle M_{(t,t'),(x,x')}\rangle_\tau}\right| \ge \lambda\right) \\ &\le \mathbb{P}\left(\sup_{0\le s\le \left(\left(\frac{d(x,x')}{2R}\right)^{\xi_1}+\left|\frac{|t-t'|}{T}\right|^{\xi_2}\right)B} |B_s| \ge \sqrt{B}\lambda\right) \\ &\le 4\exp\left(-\frac{\lambda^2}{2\left(\left(\frac{d(x,x')}{2R}\right)^{\xi_1}+\left|\frac{|t-t'|}{T}\right|^{\xi_2}\right)}\right).\end{aligned}$$

Where we used the standard Gaussian tail bound: if X is a centered Gaussian random variable with variance $\sigma > 0$, then

$$\mathbb{P}(|X| > \lambda) \le 2e^{-\frac{\lambda^2}{2\sigma}}.$$

□

Proof of Theorem 7.4. Define the $\mathcal{T} := [0,T] \times B(x_0,R)$ and the metric $d_{\mathcal{T}}$ on $\mathcal{T}$ by

$$d_{\mathcal{T}}((s,x),(s',x)) := \frac{1}{2}\left(\left(\frac{d(x,x')}{2R}\right)^{\xi_1} + \left|\frac{t-t'}{T}\right|^{\xi_2}\right), \quad (s,x),\,(s',x') \in \mathcal{T}.$$

Then $(\mathcal{T}, d_{\mathcal{T}})$ is a metric space with diameter at most 1. One can verify that for $\beta = \frac{1}{\xi_2} + \frac{\beta_0}{\xi_1}$, there is $C_{10} = C_{10}(\beta) > 0$

$$D(\mathcal{T}, d_{\mathcal{T}}; \varepsilon) \le C_{10}\varepsilon^{\beta}, \quad 0 < \varepsilon < 1. \tag{7.7}$$

Denote $B_{\mathcal{T}}(t,r)$ as the open ball in $\mathcal{T}$ centered at $t \in \mathcal{T}$ with radius $r > 0$ under the metric $d_{\mathcal{T}}$. Pick any $t_0 \in \mathcal{T}$ and let $\mathcal{T}_0 = \{t_0\}$. For $n \ge 1$, we can find $\mathcal{T}_n := \{x_i\}_{i\in I_n}$ with $|I_n| \le C_{10}2^{\beta n}$, where $|I_n|$ denote the cardinality of I_n, and

$$\mathcal{T} \subset \bigcup_{i\in I_n} B_{\mathcal{T}}\left(x_i, \frac{1}{2^n}\right).$$

Finally, we let $\mathcal{T}_\infty := \bigcup_{n\ge 0}\mathcal{T}_n$ which is dense in $\mathcal{T}$.

Now we define the chaining map $h_n : \mathcal{T}_n \to \mathcal{T}_{n-1}$ for $n \ge 1$. By construction, for any $n \ge 1$ and $t \in \mathcal{T}_n$, there is $t_{n-1} \in \mathcal{T}_{n-1}$ so that $d_{\mathcal{A}}(t_n, t_{n-1}) \le 2^{-n+1}$. For each $t \in \mathcal{T}_n$, we choose exactly one such $t' \in \mathcal{T}_{n-1}$ and define $h_n(t) := t'$ and set $h_0(t_0) = t_0$.

For $\rho > 2^{2\beta}$ and define the events

$$A_\rho := \bigcap_{n\ge 1}\{|N(t) - N(h_n(t))| \le c_n(\rho), \quad \text{for all } t \in \mathcal{T}_n\}, \quad \text{where } c_n(\rho) = \sqrt{Bn\ln(\rho)2^{-(n-1)}}.$$

Then by Lemma 7.5,

$$\begin{aligned}
\mathbb{P}\left(A_\rho^c\right) &\leq \sum_{n=1}^{\infty}\sum_{t\in\mathcal{T}_n}\mathbb{P}\left(|N(t)-N(h_n(t))|>c_n(\rho)\right)\\
&\leq 4\sum_{n=1}^{\infty}\sum_{t\in\mathcal{T}_n}\exp\left(-\frac{c_n(\rho)^2}{2d_{\mathcal{T}}(t,h_n(t))B}\right)\\
&\overset{(7.7)}{\leq} 4C_{10}\sum_{n=1}^{\infty}2^{\beta n}\exp\left(-n\ln(\rho)/2\right)\\
&= 4C_{10}\sum_{n=1}^{\infty}\left(2^{\beta}\rho^{-\frac{1}{2}}\right)^n\\
&= \frac{4C_{10}2^{\beta}}{\rho^{\frac{1}{2}}-2^{\beta}}
\end{aligned}$$

On the even A_ρ, for any $n\geq 1$ and $t\in\mathcal{T}_n$,

$$\begin{aligned}
|N(t)-N(t_0)| &\leq \sum_{i=1}^{n}|N(h_i\circ\cdots\circ h_n(t))-N(h_{i+1}\circ\cdots\circ h_{i-1}(t))|\\
&\leq \sqrt{2B\ln(\rho)}\sum_{i=1}^{\infty}\sqrt{n2^{-n}}.
\end{aligned}$$

Hence

$$\sup_{t,t'\in\mathcal{T}_\infty}|N(t)-N(t')|\leq\sqrt{8B\ln(\rho)}\sum_{n\geq 1}\sqrt{n}2^{-n/2}.$$

Now for any $\rho>2^{2\beta+1}$, let $\lambda>0$ be that $\exp\left(\lambda^2\right)=\rho$, and $c:=\sqrt{8}\sum_{i=1}^{\infty}\sqrt{n}2^{-n/2}$, then

$$\mathbb{P}\left(\sup_{t,t'\in\mathcal{T}_\infty}|N(t)-N(t')|>c\lambda\sqrt{B}\right)\leq\mathbb{P}\left(A^c\right)\leq\frac{4C_{10}2^{\beta}}{\rho^{\frac{1}{2}}-2^{\beta}}\leq\frac{8C_{10}2^{\beta}}{\rho^{\frac{1}{2}}}=\frac{8C_{10}2^{\beta}}{\exp\left(\lambda^2/2\right)}.$$

For $\rho\in\left[0,2^{2\beta+1}\right]$, we see $\exp\left(-\frac{\lambda^2}{2}\right)\geq 2^{-\beta-1}$. Hence $8\max\{C_{10},1\}\,2^{\beta}\exp\left(-\lambda^2/2\right)\geq 1$. So we get for all $\lambda>0$

$$\mathbb{P}\left(\sup_{t,t'\in\mathcal{T}_\infty}|N(t)-N(t')|>c\lambda\sqrt{B}\right)\leq\frac{8\max\{C_{10},1\}\,2^{\beta}}{\exp\left(\lambda^2/2\right)},$$

which yields the desired inequality. Finally, the existence of continuous modification can be check by [41, Theorem 2.3.1] combined with the estimates in Lemma 7.5. $\square$

We will also need the following elementary result:

Lemma 7.6. *Suppose $y>0$ and $x\geq 1$, then*

$$\exp\left(-yx\right)\leq\frac{\exp\left(-\frac{y}{2}\right)}{xy}.$$

Proof. We first show $\ln(xy)-y(x-1)<0$ for all $y>0$ and $x\geq 2$. Observe for each $x\geq 2$, the function $f_x(y):=\ln(xy)-y(x-1)$ is increasing for $0<y\leq\frac{1}{x-1}$ and decreasing when $y>\frac{1}{x-1}$. Hence

$$\sup_{x\geq 2}\sup_{y>0}\left(\ln(xy)-y(x-1)\right)=\sup_{x\geq 2}\ln\left(\frac{x}{x-1}\right)-1\leq\ln(2)-1<0.$$

Therefore, $xy\exp(-xy+y)=\exp\left(\ln(xy)-y(x-1)\right)<1,$ for all $y>0$, $x\geq 2$. Multiply both sides by $\exp(-y)$ to get

$$xy\exp(-xy)<\exp(-y),\qquad\text{for all } y>0,\ x\geq 2.$$

This implies

$$xy\exp\left(-\frac{y}{2}2x\right)\le\exp\left(-\frac{y}{2}\right),\qquad\text{for all } y>0, x\ge 1.$$

□

We are now ready to prove Lemma 7.1, in fact, we will prove the following more general result that implies Lemma 7.1 when combined with Lemma 7.2.

Lemma 7.7. *Suppose the conditions of Theorem 7.4 hold,* $(\mathbb{X},d)$ *satisfies* (2.1) *and* $h:\mathbb{X}\mapsto[1,\infty)$ *is a good weight. Let* $\lambda_0>0$ *be as in* (2.5) *and* $\lambda>0$*. Assume there is* $\alpha>0$ *so that for each* $T>0$*, the following inequality holds uniformly in* $(t,x)\in[0,T]\times\mathbb{X}$*,*

$$\sup_{s,s'\in[0,t],\,y,y'\in B(x,1)}\frac{\int_0^\infty\left\|\left(g_{s,x}(\tau,\cdot)-g_{s',x'}(\tau,\cdot)\right)\eta(\tau,\cdot)\right\|_{\widetilde{K}}^2\,\mathrm{d}s}{\left(\frac{d(x,x')}{2R}\right)^{\xi_1}+\left|\frac{t-t'}{T}\right|^{\xi_2}}\lesssim h(x)^{2\lambda-\lambda_0}t^\alpha. \tag{7.8}$$

Let $N(t,x)$ *be as in* (7.5)*, then there exists* $C_{11}=C_{11}(T,\lambda)>0$ *so that uniformly for* $t\in[0,T]$ *and* $\varepsilon\in(0,1)$*,*

$$\mathbb{P}\left(\sup_{(s,x)\in[0,t]\times\mathbb{X}}|N(s,x)|h(x)^{-\lambda}>\varepsilon\right)\lesssim\varepsilon^{-2}t^\alpha\exp\left(-\frac{C_{11}}{2t^\alpha}\varepsilon^2\right).$$

Proof. We see from condition (2.6) that for each $\lambda\ge 0$, there is $C_\lambda>0$ so that for all and $x\in\mathbb{X}$,

$$C_\lambda h(x)^{-\lambda}\ge\sup_{z\in B(x,1)}h(z)^{-\lambda}\ge h(y)^{-\lambda},\qquad\text{for all } y\in B(x,1). \tag{7.9}$$

Let λ,λ_0 be as in (7.8), then by Theorem 7.4, we see there is $C_{11}=C(\lambda,T)>0$ so that uniformly in $\varepsilon\in(0,1)$, $(t,x)\in[0,T]\times\mathbb{X}$,

$$\begin{aligned}\mathbb{P}\left(\sup_{(s,y)\in[0,t]\times B(x,1)}|N(s,y)|h(y)^{-\lambda}>\varepsilon\right)&\overset{(7.9)}{\le}\mathbb{P}\left(\sup_{(s,y)\in[0,t]\times B(x,1)}|N(s,y)|\cdot C_\lambda h(x)^{-\lambda}>\varepsilon\right)\\&=\mathbb{P}\left(\sup_{(s,x)\in[0,t]\times B(x,1)}|N(s,x)|>C_\lambda^{-1}h(x)^{\lambda}\varepsilon\right)\\&\overset{(7.6)}{\lesssim}\exp\left(-C_{11}\frac{h(x)^{\lambda_0}}{t^\alpha}\varepsilon^2\right)\\&\le\varepsilon^{-2}t^\alpha\exp\left(-\frac{C_{11}}{2t^\alpha}\varepsilon^2\right)h(x)^{-\lambda_0},\end{aligned}$$

where we used Lemma 7.6 and the fact $h(x)\ge 1$ for all $x\in\mathbb{X}$ in the last inequality. Now let $\{x_n\}_{n\in\mathbb{N}}\subset\mathbb{X}$ so that (2.5) holds. Then we see uniformly in $t\in[0,T]$ and $\varepsilon\in(0,1)$

$$\begin{aligned}\mathbb{P}\left(\sup_{(s,x)\in[0,t]\times\mathbb{X}}|N(s,x)|h(x)^{-\lambda}>\varepsilon\right)&=\mathbb{P}\left(\bigcup_{n\in\mathbb{N}}\left\{\sup_{(s,x)\in[0,t]\times B(x_n,1)}|N(s,x)|h(x)^{-\lambda}>\varepsilon\right\}\right)\\&\le\sum_{n\in\mathbb{N}}\mathbb{P}\left(\sup_{(s,x)\in[0,t]\times B(x_n,1)}|N(s,x)|h(x)^{-\lambda}>\varepsilon\right)\\&\le\varepsilon^{-2}t^\alpha\sum_{n\in\mathbb{N}}\exp\left(-\frac{C_{11}}{2t^\alpha}\varepsilon^2\right)h(x_n)^{-\lambda_0}\\&\lesssim\varepsilon^{-2}t^\alpha\exp\left(-\frac{C_{11}}{2t^\alpha}\varepsilon^2\right).\end{aligned}$$

□

To summarize the proof of Theorem 3.6, we first used a chaining technique to prove a Mueller-type large deviation estimates for the stochastic integral on compact subsets of $\mathbb{X}$. Then we used properties of a good weight to control the spatial growth of Z in (7.5). By using condition (3.6), we can estimate the difference between $P_t f$ and the stochastic integral and show solution is strictly positive on any compact subset of $\mathbb{R}_+ \times \mathbb{X}$ with probability one via Markov property. Strong comparison principle is then obtained by observing those compact subsets can be arbitrarily large.

8. Proofs of results in Section 5

The results stated in Section 5 can be obtained by more or less standard arguments. One thing that we needs to be careful is the choice of the weight in relation to the geometry of the metric space $(\mathbb{X}, d)$.

8.1. **Some preparation lemmas.**

We prepare some elementary lemmas in this section for future uses.

Lemma 8.1. *Let $\mathcal{X}$ be a measurable space and μ be a non-negative, non-negative definite and symmetric Radon measure on $\mathcal{X} \times \mathcal{X}$. Then for all functions f, g on $\mathbb{X}$, we have*

$$\|f+g\|_\mu \le \|g\|_\mu + \|f\|_\mu\,, \qquad \int_{\mathbb{X}^2} |f(x)g(y)|\mu(\mathrm{d}x,\mathrm{d}y) \le \|f\|_\mu \|g\|_\mu \tag{8.1}$$

whenever the right hand sides of (8.1) *is finite.*

Theorem 8.2 (Stochastic Fubini)**.** *Let $(G, \mathcal{B}(G), \nu)$ be a finite measure space and W is a Gaussian noise with covariance measure K on $\mathbb{X}^2$ and $|K| \le \widetilde{K}$ for some symmetric, non-negative and non-negative definite measure $\widetilde{K}$ on $\mathbb{X}^2$. Let $(f(x,t,\omega,\lambda)$, $x \in \mathbb{X}$, $t \ge 0$, $\omega \in \Omega$, $\lambda \in G)$ be a measurable function. Suppose for some $T > 0$,*

$$\mathbb{E}\left[\int_{[0,T]\times G} \|f(\cdot,t,\lambda)\|^2_{\widetilde{K}}\, \mathrm{d}t\, \nu(\mathrm{d}\lambda)\right] < \infty,$$

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

$$\int_G \int_0^t \int_{\mathbb{X}} f(x,s,\lambda) W(\mathrm{d}x,\mathrm{d}s)\nu(\mathrm{d}\lambda) = \int_0^t \int_{\mathbb{X}} \int_G f(x,s,\lambda)\nu(\mathrm{d}\lambda) W(\mathrm{d}x,\mathrm{d}s), \quad \mathbb{P}-a.s.$$

Lemma 8.3. *Let W be as in Lemma 8.2 and $\Phi : \mathbb{R}^+ \times \mathbb{X} \times \Omega \to \mathbb{R}$ be $\mathcal{F}_t$-predictable, then for all $p > 0$, there is $C_p \in \mathbb{R}^+$ depending only on p such that*

$$\mathbb{E}\left[\left(\int_0^t \int_{\mathbb{X}} \Phi(s,x) W(\mathrm{d}s,\mathrm{d}x)\right)^{2p}\right] \le C_p \mathbb{E}\left[\left(\int_0^t \|\Phi(s,\cdot)\|^2_{\widetilde{K}}\, \mathrm{d}s\right)^p\right], \quad t \ge 0,$$

whenever the integral make sense.

Lemma 8.3 combined with the Cauchy–Schwarz inequality and Minkowski's inequality gives us the following elementary inequality.

Corollary 8.4. *Under the setting of Lemma 8.3, for any $\Phi : \mathbb{R}_+ \times \mathbb{X} \times \Omega \to \mathbb{R}$ that is $\mathcal{F}_t$-predictable, then for all $p \ge 2$, there is $C_p \in \mathbb{R}_+$ depend only on p so that*

$$\left\|\int_0^t \int_{\mathbb{X}} \Phi(s,x) W(\mathrm{d}s,\mathrm{d}x)\right\|_{L^p(\mathbb{P})} \le C_p \left(\int_0^t \left\| \|\Phi(s,\cdot)\|_{L^p(\mathbb{P})}\right\|^2_{\widetilde{K}} ds\right)^{\frac12}, \quad t \ge 0,$$

whenever the integral makes sense.

Lemma 8.5. *Let $T > 0$, $\mathcal{X}$ be a measurable space and $(\mu(t,x,\mathrm{d}y,\mathrm{d}z)$, $t \in [0,T]$, $x,y,z \in \mathbb{X})$ be an integral kernel from $\mathcal{X}^2$ to $[0,T] \times \mathcal{X}$ such that for each $(t,x) \in [0,T] \times \mathcal{X}$, $\mu(t,x,\mathrm{d}y,\mathrm{d}z)$ is a symmetric, non-negative and non-negative definite measure on $\mathcal{X}^2$. Suppose $h : \mathcal{X} \to [1,\infty)$ satisfies for some $\alpha \in (0,1)$ and $C > 0$ that*

$$\int_{\mathcal{X}} \int_{\mathcal{X}} h(y)^{\pm\frac12} h(z)^{\pm\frac12} \mu(s,x,\mathrm{d}y,\mathrm{d}z) \le C\, s^{-\alpha} h(x)^{\pm 1}, \quad \text{for all } (t,x) \in [0,T] \times \mathcal{X}. \tag{8.2}$$

Let $u : [0,T] \times \mathcal{X} \to \mathbb{R}_+$ be a measurable function that satisfies

$$u_t(x) \le h(x)^{-1} + \int_0^t \int_{\mathcal{X}} \int_{\mathcal{X}} u_s(y)^{\frac{1}{2}} u_s(z)^{\frac{1}{2}} \mu(t-s, x, \mathrm{d}y, \mathrm{d}z)\mathrm{d}s, \quad (t,x) \in [0,T] \times \mathcal{X}. \tag{8.3}$$

Then

$$\sup_{0 \le t \le T} \sup_{x \in \mathcal{X}} u_t(x) h(x) < \infty.$$

Lemma 8.6. *Let $(\mathbb{X}, d)$ be a metric space with satisfies* (2.1) *and h is a good weight on $\mathbb{X}$. Let $\{X_n\}_{n\in\mathbb{N}}$ is a family of random fields on $\mathbb{R}_+ \times \mathbb{X}$ so that*

$$\sup_{n\in\mathbb{N}} \sup_{x\in\mathbb{X}} \mathbb{E}\left[|X_n(0,x)|^p\right] h(x)^{-\lambda} < \infty.$$

If there exists $\xi_1, \xi_2 \in (0,1]$ so that for each $T > 0$, $\lambda \in (0,1]$ and $p \ge 2$, the following inequality holds uniformly for $0 \le t, t' \le T$, $x \in \mathbb{X}$ and $y_1, y_2 \in B(x,1)$

$$\sup_{n\in\mathbb{N}} \mathbb{E}\left[|X_n(t,y_1) - X_n(t',y_2)|^p\right] \lesssim \left(|t'-t|^{\xi_1} + d(y_1,y_2)^{\xi_2}\right)^p h(x)^{\lambda}.$$

Then $\{X_n\}_n$ is tight on $C(\mathbb{R}_+; \mathcal{C}_{tem})$. In particular, $X_n \in C(\mathbb{R}_+; \mathcal{C}_{tem})$ almost surely for all $n \ge 1$.

Lemma 8.7. *Under the assumptions of Lemma 8.6, if for al $\lambda > 0$, $p \ge 1$,*

$$\sup_{x\in\mathbb{X}} h(x)^{\lambda} \mathbb{E}\left[|X(0,x)|^p\right] < \infty,$$

and for each $T > 0$, $\lambda > 0$ and $p \ge 1$, the following inequality holds uniformly for $0 \le t, t' \le T$, $x \in \mathbb{X}$ and $y_1, y_2 \in B(x,1)$,

$$\mathbb{E}\left[|X(t,y_1) - X(t',y_2)|^p\right] \lesssim \left(|t'-t|^{\xi_1} + d(x,x')^{\xi_2}\right)^p h(x)^{-\lambda}.$$

Then $\mathbb{P}\left(X_t(\cdot) \in \mathcal{C}_{rap} \text{ for all } t \in \mathbb{R}_+\right) = 1$.

Lemma 8.8. *Let $(\mathbb{X}, d)$ be a Polish space and h is a good weight on it. Then $\mathcal{C}_0(\mathbb{X})$ is dense in $\mathcal{C}_{tem}$, i.e., for each $\lambda > 0$ and $f \in \mathcal{C}_{tem}$, there exists a sequence $\{f_n\}_{n\in\mathbb{N}} \subset \mathcal{C}_0(\mathbb{X})$ so that $\|f_n - f\|_{(\lambda)}$ converges to zero as $n \to \infty$.*

Lemma 8.1 can be found in [56, Exercise 2.4] and Theorem 8.2 in [56, Theorem 2.6]. Lemma 8.3 is an direct application of the Burkholder-Davis-Gundy inequality. The proof of Lemma 8.7 requires only a minor modification in the proof of Lemma 8.6. So we only provide proof for Lemma 8.5, Lemma 8.6 and Lemma 8.8 in Section 9.

8.2. **Proofs of Theorem 5.1.** Let $h : \mathbb{X} \to [1,\infty)$ be a continuous function. For $\lambda \in \mathbb{R}$, $A, T > 0$ and $p \ge 1$, we define the Banach space

$$\mathbb{M}^A_{\lambda,p}(T) := \Bigg\{ u(t,x) \text{ predictable random field on } [0,T] \times \mathbb{X} : \\ \|u\|_{\mathbb{M}^A_{\lambda,p}(T)} := \sup_{t\in[0,T],\, x\in\mathbb{X}} e^{-As} h(x)^{-\lambda} \|u(t,x)\|_{L^p(\mathbb{P})} < \infty \Bigg\}.$$

Lemma 8.9. *Let f, g be predictable random fields on $\mathbb{R}_+ \times \mathbb{X}$, and u_0 is a random field on $\mathbb{X}$. Suppose* (2.8) *holds for some $\lambda \in \mathbb{R}$ and set*

$$\psi(t,x) := \int_{\mathbb{X}} p_t(x,y) u_0(y)\mathrm{d}y + \int_0^t \int_{\mathbb{X}} p_{t-s}(x,y) f(s,y)\mathrm{d}y\mathrm{d}s + \int_0^t \int_{\mathbb{X}} p_{t-s}(x,y) g(s,y) W(\mathrm{d}y\,\mathrm{d}s).$$

Let α be as in (2.8)*, then for each $p \ge 2$ and $T > 0$, the following inequality holds uniformly in $A \ge 1$ and $t \in [0,T]$,*

$$\|\psi\|_{\mathbb{M}^A_{\lambda,p}(t)} \lesssim A^{\frac{\alpha-1}{2}} \left(\|f\|_{\mathbb{M}^A_{\lambda,p}(T)} + \|g\|_{\mathbb{M}^A_{\lambda,p}(T)} \right) + \|u_0\|_{\mathbb{M}^A_{\lambda,p}(T)},$$

whenever the right hand side is finite.

Proof. We see by Lemma 8.3 that

$$\|\psi(t,x)\|_{L^p(\mathbb{P})}^p \leq \mathbb{E}\left[\left(\int_{\mathbb{X}} p_t(x,y)\,|u_0(y)|\,m(\mathrm{d}y)\right)^p\right] + \mathbb{E}\left[\left|\int_0^t \int_{\mathbb{X}} p_{t-s}(x,y)\,|f(s,y)|\,m(\mathrm{d}y)\mathrm{d}s\right|^p\right]$$
$$+ \left(\int_0^t \left\|p_{t-s}(x,\cdot)\,\|g(s,\cdot)\|_{L^p(\mathbb{P})}\right\|_{\widetilde{K}}^2\right)^{\frac{p}{2}}. \tag{8.4}$$

We will call the terms on the right-hand side of (8.4) $I_i(t,x)$ for $i=1,2,3$ respectively. Note (2.8) asserts $I_1(t,x) \lesssim \|u_0\|_{\mathbb{M}_{\lambda,p}^A(T)}^p e^{pAt} h(x)^{\lambda p}$ uniformly in $(t,x) \in [0,T]\times\mathbb{X}$. By Minkowski's inequality, we have uniformly in $(t,x)\in[0,T]\times\mathbb{X}$ that

$$\begin{aligned}
I_2(t,x) &\leq \left(\int_0^t \int_{\mathbb{X}} p_{t-s}(x,y)\,\|f(s,y)\|_{L^p(\mathbb{P})}\,m(\mathrm{d}y)\mathrm{d}s\right)^p \\
&\leq \left(\int_0^t e^{As}\int_{\mathbb{X}} p_{t-s}(x,y)h(y)^{\lambda} \sup_{\tau\in[0,s]}\sup_{x\in\mathbb{X}}\left(e^{-A\tau}h(y)^{-\lambda}\,\|f(\tau,y)\|_{L^p(\mathbb{P})}\right) m(\mathrm{d}y)\mathrm{d}s\right)^p \\
&\overset{(2.8)}{\lesssim} h(x)^{p\lambda}\left(\int_0^t e^{As}\,\|f\|_{\mathbb{M}_{\lambda,p}^A(s)}\,\mathrm{d}s\right)^p \\
&\lesssim h(x)^{p\lambda}\left(\int_0^t e^{As}(t-s)^{-\alpha}\,\|f\|_{\mathbb{M}_{\lambda,p}^A(s)}\,\mathrm{d}s\right)^p.
\end{aligned}$$

Similarly,

$$\begin{aligned}
I_3(t,x) &= \left(\int_0^t e^{2As}e^{-2As}\left\|p_{t-s}(x,\cdot)h^{\lambda}(\cdot)h^{-\lambda}(\cdot)\,\|g(s,\cdot)\|_{L^p(\mathbb{P})}\right\|_{\widetilde{K}}^2 \mathrm{d}s\right)^{\frac{p}{2}} \\
&\leq \left(\int_0^t e^{2As}\left\|p_{t-s}(x,\cdot)h^{\lambda}\right\|_{\widetilde{K}}^2 \mathrm{d}s\right)^{\frac{p}{2}} \|g\|_{\mathbb{M}_{\lambda,p}^A(T)} \\
&\overset{(2.8)}{\lesssim} h^{p\lambda}(x)\left(\int_0^t e^{2As}(t-s)^{-\alpha}\mathrm{d}s\right)^{\frac{p}{2}} \|g\|_{\mathbb{M}_{\lambda,p}^A(T)}.
\end{aligned}$$

Note for every $\alpha\in[0,1)$,

$$\int_0^t e^{2As}(t-s)^{-\alpha}\mathrm{d}s \lesssim A^{\alpha-1}e^{2At}, \quad \text{uniformly in } A,\ t>0.$$

Since $\alpha\in(0,1)$, we have for $A\geq 1$, and $t\in[0,T]$,

$$\|\psi\|_{\mathbb{M}_{\lambda,p}^A(t)} \lesssim A^{\frac{\alpha-1}{2}}\left(\|f\|_{\mathbb{M}_{\lambda,p}^A(T)} + \|g\|_{\mathbb{M}_{\lambda,p}^A(T)}\right) + \|u_0\|_{\mathbb{M}_{\lambda,p}^A(T)},$$

which is the desired result. □

In what follows, we suppress the dependent of the coefficients σ and b on $t\in\mathbb{R}_+$ and $x\in\mathbb{X}$ and write

$$b(u_t(x)) := b(t,x,u_t(x)), \quad \sigma(u_t(x)) := \sigma(t,x,u_t(x)). \tag{8.5}$$

We set

$$\mathbb{M}_{\lambda} := \bigcup_{T,\lambda>0,\,A,p\geq 2} \mathbb{M}_{\lambda,p}^A(T),$$

and Φ be a map on $\mathbb{M}_{\lambda}$ defined for $(t,x)\in\mathbb{R}_+\times\mathbb{X}$ as

$$\Phi u(t,x) := \int_{\mathbb{X}} p_t(x,y)u_0(y)m(\mathrm{d}y) + \int_0^t\int_{\mathbb{X}} p_{t-s}(x,y)b(u(s,y))m(\mathrm{d}y)\mathrm{d}s$$

$$+\int_0^t \int_{\mathbb{X}} p_{t-s}(x,y)\sigma(u(s,y))W(\mathrm{d}y\,\mathrm{d}s).$$

Furthermore, we will call the three terms on the right hand side $\Phi_0 u$, $\Phi_1 u$ and $\Phi_2 u$ respectively. The following proposition proves the first part of Theorem 5.1.

Proposition 8.10. *Suppose $\lambda \geq 0$, h is a weight on $\mathbb{X}$ so that the inequalities* (2.8) *hold. If b, σ satisfies the Lipschitz condition* (3.5) *and the (random) initial condition $u_0 \in \mathcal{F}_0$ satisfies*

$$\sup_{x\in\mathbb{X}} h(x)^\lambda \|u_0(x)\|_{L^p} < \infty, \qquad \textit{for all } p \geq 1.$$

Then there is a unique solution (up to modification) to the equation (2.3) *in $\mathbb{M}_\lambda$ with initial value u_0.*

Proof. We first note by a standard approximation method, we can show $(\Phi u(t,x),\, t>0,\, x\in\mathbb{X})$ is a predictable random field for $u \in \mathbb{M}_\lambda$. By (3.5) and Lemma 8.9, for each $T>0$, $p\geq 2$ the following inequality holds uniformly for $A \geq 1$ and $u \in \mathbb{M}^A_{\lambda,p}(T)$:

$$\|\Phi u\|_{\mathbb{M}^A_{\lambda,p}(T)} \lesssim A^{(\alpha-1)/2} L \left(\|1+u\|_{\mathbb{M}^A_{\lambda,p}(T)} + \|1+u\|_{\mathbb{M}^A_{\lambda,p}(T)} \right) + \|u_0\|_{\mathbb{M}^A_{\lambda,p}(T)},$$

$$\|\Phi u - \Phi v\|_{\mathbb{M}^A_{\lambda,p}}(T) \lesssim A^{(\alpha-1)/2} L \|u-v\|_{\mathbb{M}^A_{\lambda,p}(T)}.$$

Hence Φ maps from $\mathbb{M}^A_{\lambda,p}(T)$ to itself for each $A\geq 1$, and is a contraction for large enough A. Then Banach contraction theorem asserts that there is a unique solution to the equation (2.3) in $\mathbb{M}^A_{\lambda,p}(T)$ for all large enough A. We conclude the proof by noting that $p\geq 2$ and $T>0$ can be arbitrarily large. ☐

The following proposition proves the second part of Theorem 5.1.

Proposition 8.11. *Suppose $(\mathbb{X}, d, m)$ satisfies* (2.1)*, $\mathcal{L}$ satisfies (H_1) of Assumption 2.3 and h satisfies Assumption 2.5. Then the solution u to the equation* (2.3) *with initial value $u_0 \in \mathcal{C}_{tem}$, given by Proposition 8.10, satisfies*

$$\mathbb{P}\left(u \in C(\mathbb{R}_+; \mathcal{C}_{tem})\right) = 1.$$

Proof. Let u be the solution to (2.3) given by Proposition 8.10, since $\Phi_0 u \in \mathcal{C}_{tem}$ is assumed by (W_1), we only need to worry about $\Phi_i u$ for $i=1,2$. Denote $\mathbb{M}_{\lambda,p} = \mathbb{M}^0_{\lambda,p}$ and observe by Proposition 8.10 $\Phi u \in \cup_{\lambda>0}\mathbb{M}_\lambda$. Let $\lambda>0$ and $p>\max\{1,\lambda/2\}$, then for any $T>0$ we have by Minkowski's and Hölder's inequality that uniformly for all $0\leq t,t'\leq T$, $x,x'\in\mathbb{X}$ with $d(x,x')\leq 1$,

$$\begin{aligned}
\|\Phi_1 u(t,x) - \Phi_1 u(t',x')\|_{L^p(\mathbb{P})} &\leq \int_0^{t'} \int_{\mathbb{X}} |p_{t'-s}(x,y) - p_{t-s}(x',y)| \, \|b(u(s,y))\|_{L^p(\mathbb{P})} \, m(\mathrm{d}y)\,\mathrm{d}s \\
&\leq \|b(u)\|_{\mathbb{M}_{\lambda/p,p}(T)} \left(\int_0^{t'} |p_{t'-s}(x,y) - p_{t-s}(x',y)| \, m(\mathrm{d}y)\mathrm{d}s \right)^{\frac{1}{2}} \\
&\quad \times \left(\int_0^{t'} \int_{\mathbb{X}} \left(p_{t'-s}(x,y) + p_{t-s}(x',y)\right) h(y)^{2\lambda/p} m(\mathrm{d}y)\,\mathrm{d}s \right)^{\frac{1}{2}} \\
&\overset{(H_1)}{\lesssim} \|b(u)\|_{\mathbb{M}_{\lambda/p,p}(T)} \left(|t'-t|^{\xi_1} + d(x,x')^{\xi_2} \right)^{\frac{1}{2}} \left(h^{\lambda/p}(x) + h^{\lambda/p}(x') \right).
\end{aligned}$$

By Lemma 8.1, Lemma 8.3 and Minkowski's inequality, a similar calculations as in the proof of Lemma 8.9 shows that uniformly for $0\leq t,\, t'\leq T$ and $x,x'\in\mathbb{X}$ with $d(x,x')\leq 1$,

$$\begin{aligned}
\|\Phi_2 u(t,x) - \Phi_2 u(t',x')\|_{L^p(\mathbb{P})} \lesssim \Bigg(\int_0^{t'} \int_{\mathbb{X}} \int_{\mathbb{X}} &|p_{t-s}(x,y) - p_{t'-s}(x',y)| \\
&\times |p_{t-s}(x,z) - p_{t'-s}(x',z)|
\end{aligned}$$

$$\times h^{\lambda/p}(y)h^{\lambda/p}(z)\,\widetilde{K}(\mathrm{d}z,\mathrm{d}y)\mathrm{d}s\Bigg)^{\frac{1}{2}}\|a(u)\|_{\mathbb{M}_{\lambda,p}(T)}$$

$$\lesssim A_1(t,t',x,x')\times A_2(t,t',x,x')\,\|a(u)\|_{\mathbb{M}_{\lambda,p}(T)}$$

where

$$A_1(t,t',x,x')=\left(\int_0^{t'}\|p_{t-s}(x,\cdot)-p_{t'-s}(x',\cdot)\|_{\widetilde{K}}^2\,\mathrm{d}s\right)^{\frac{1}{4}}\overset{(\mathrm{H}_1)}{\lesssim}\left(|t-t'|^{\xi_1/4}+d(x,x')^{\xi_2/4}\right)$$

and

$$A_2(t,t',x,x')=\left(\int_0^{t'}\left\|(p_{t'-s}(x',\cdot)+p_{t-s}(x,\cdot))\,h^{2\lambda/p}\right\|_{\widetilde{K}}^2\mathrm{d}s\right)^{\frac{1}{4}}\overset{(2.8)}{\lesssim}\left(h^{\lambda/p}(x)+h^{\lambda/p}(x')\right),$$

holds uniformly in $0\le t,\,t'\le T$ and $x,x'\in\mathbb{X}$ with $d(x,x')\le 1$. Therefore,

$$\sum_{i=1}^{2}\|\Phi_i(u)(t,x)-\Phi_i(u)(t',x')\|_{L^p(\mathbb{P})}^p\lesssim\left(h(x)^\lambda+h(x')^\lambda\right)\left(|t-t'|^{\frac{\xi_1}{4}}+d(x,x')^{\frac{\xi_2}{4}}\right)^p$$

$$\overset{(2.6)}{\lesssim}h(x)^\lambda\left(|t-t'|^{\frac{\xi_1}{4}}+d(x,x')^{\frac{\xi_2}{4}}\right)^p.$$

Since p can be arbitrarily large, [46, Theorem 11.2] ensures that $\Phi_1u+\Phi_2u$ is jointly continuous. Also, by Lemma 8.6, we see $u\in\mathcal{C}(\mathbb{R}_+;\mathcal{C}_{tem})$ almost surely if h is a good weight. □

8.3. **Proof of Theorem 5.2.** We will need the following continuity with respect to initial condition type result, whose proof will be given at the end of this section.

Proposition 8.12. *Under the setting of Theorem 5.2, suppose $f\in\mathcal{C}_{tem}$ and $\{f_n\}_{n\in\mathbb{N}}\subset\mathcal{C}_{tem}$ is a sequence so that $\lim_{n\to\infty}\|f_n-f\|_\lambda=0$ for all $\lambda>0$. Let u^n be the solution to the equation with initial value f_n* (2.3) *and u solves* (2.3) *with initial value f, then for each $p\ge 2$, we have*

$$\lim_{n\to\infty}\sup_{t\in[0,T]}\sup_{x\in\mathbb{X}}h^{-\lambda}(x)\,\|u_s^n(x)-u_s(x)\|_{L^p(\mathbb{P})}=0,\qquad \textit{for each } T>0.$$

We note by Proposition 8.12 and Lemma 8.8, it is enough to assume $u_0\in\mathcal{C}_0(\mathbb{X})$. For $\varepsilon>0$, we let $\mathcal{L}_\varepsilon:=\frac{P_\varepsilon-I}{\varepsilon}$ be the approximation of $\mathcal{L}$, $P_\varepsilon(t)$ and $p_\varepsilon(t,x,y)$ are the semi-group and the corresponding integral kernel generated by $\mathcal{L}_\varepsilon$, namely, we set

$$P_\varepsilon(t)=\exp\left(t\mathcal{L}_\varepsilon\right)=e^{-t/\varepsilon}\sum_{n=0}^{\infty}\frac{(t/\varepsilon)^n}{n!}P(n\varepsilon):=e^{-t/\varepsilon}+R_\varepsilon(t),\tag{8.6}$$

where for $t>0$, $R_\varepsilon(t)$ admits a density ($R_\varepsilon(t,x,y)$, $t>0$, $x,y\in\mathbb{X}$):

$$R_\varepsilon(t,x,y)=e^{-t/\varepsilon}\sum_{n=1}^{\infty}\frac{(t/\varepsilon)^n}{n!}p(n\varepsilon,x,y).$$

We will need the following estimates, whose proof is in the appendix.

Lemma 8.13. *Under the assumptions of Theorem 5.2, let $T>0$, then*

(1) let $\alpha\in(0,1)$ be as in 2.8, then

$$\sup_{x\in\mathbb{X}}\|R_\varepsilon(t,x,\cdot)\|_{\widetilde{K}}^2\lesssim t^{-\alpha}\vee 1.$$

(2) For each $t>0$,

$$\lim_{\varepsilon\downarrow 0}\sup_{x\in\mathbb{X}}\int_0^t\left\|\int_{\mathbb{X}}R_\varepsilon(s,x,y)p_\varepsilon(y,\cdot)m(\mathrm{d}y)-p(s,x,\cdot)\right\|_{\widetilde{K}}^2\mathrm{d}s=0.\tag{8.7}$$

(3) *There is $c_T \in (0,1]$ and a continuous non-decreasing function on $\mathbb{R}_+$: $\varepsilon \mapsto \gamma_\varepsilon$ with $\gamma_0 = 0$, so that for all $t \in (0,T]$, $\varepsilon > 0$,*

$$\int_{\mathbb{X}} |R_\varepsilon(t,x,y) - p(t,x,y)|\, m(\mathrm{d}y) \lesssim_T e^{-c_T \frac{t}{\varepsilon}} + \frac{\gamma_\varepsilon}{t^\beta}, \tag{8.8}$$

where β is as in (H_3).

Proof of Theorem 5.2. For $\varepsilon > 0$, define $W_x^\varepsilon(t)$ for $t \ge 0, x \in \mathbb{X}$ by

$$W_x^\varepsilon(t) = \int_{\mathbb{X}} p_\varepsilon(x,y) W([0,t], \mathrm{d}y). \tag{8.9}$$

For each $x \in \mathbb{X}$, $(W_x^\varepsilon(t),\, t \ge 0)$ is a (scaled) one-dimensional Brownian motion:

$$\left\langle \int_0^\cdot W_x^\varepsilon(\mathrm{d}s) \right\rangle_t = \left\langle \int_0^\cdot \int_{\mathbb{X}} p_\varepsilon(x,y) W(\mathrm{d}y, \mathrm{d}s) \right\rangle_t = t \int_{\mathbb{X}} \int_{\mathbb{X}} p_\varepsilon(x,y)\, p_\varepsilon(x,z) K(\mathrm{d}y, \mathrm{d}z). \tag{8.10}$$

One can view $W_x^\varepsilon(t)$ as a martingale measure $W^\varepsilon(x,s)$ with the following covariance relation:

$$\mathbb{E}\left[W^\varepsilon(x,t) W^\varepsilon(x',t')\right] = \delta(t-t') k_\varepsilon(x,x'), \quad x, x' \in \mathbb{X},$$

where $k_\varepsilon(x,x') := \int_{\mathbb{X}} \int_{\mathbb{X}} p_\varepsilon(x,y) p_\varepsilon(x,z) K(\mathrm{d}y, \mathrm{d}z)$ for $x, x' \in \mathbb{X}$.

Following the convention in (8.5), let $\varepsilon > 0$ and let $\mathcal{L}_\varepsilon, P_\varepsilon$ and R_ε be defined as in (8.6) and consider the following integral equation:

$$u_\varepsilon(t,x) = f(x) + \int_0^t \mathcal{L}_\varepsilon u_\varepsilon(s,x) + b(u_\varepsilon(s,x)) \mathrm{d}s + \int_0^t \sigma(u_\varepsilon(s,x)) \mathrm{d}W_x^\varepsilon(s). \tag{8.11}$$

It can be verified by stochastic Fubini's that

$$\begin{aligned} u_\varepsilon(t,x) = P_\varepsilon(t) f(x) + &\int_0^t \int_{\mathbb{X}} p_\varepsilon(t-s,x,y) b(u_\varepsilon(s,y)) m(\mathrm{d}y)\, \mathrm{d}s \\ &+ \int_{\mathbb{X}} \int_0^t p_\varepsilon(t-s,x,y) \sigma(u_\varepsilon(s,y)) \mathrm{d}W_y^\varepsilon(s)\, m(\mathrm{d}y). \end{aligned}$$

The well-posedness of (8.11) can be shown in a similar procedure as before. It can be checked that

$$\sup_{0<\varepsilon\le 1} \sup_{0\le t\le T} \sup_{x\in\mathbb{X}} \mathbb{E}\left[|u_\varepsilon(t,x)|^2\right] < \infty. \tag{8.12}$$

Furthermore, by (8.9) and (8.11), for each $x \in \mathbb{X}$, $u_\varepsilon(t,x)$ is a continuous semi-martingale in t. Let $\varphi(u) = -u \wedge 0 = u^-$, $u \in \mathbb{R}$, which is convex. By Ito-Mayers-Tanaka's formula ([45, Theorem 9.6]),

$$\begin{aligned} \mathbb{E}[\varphi(u_\varepsilon(t,x))] = \varphi(f(x)) &- \mathbb{E}\left[\int_0^t I_{u_\varepsilon(s,x)\le 0} \left(\mathcal{L}_\varepsilon u_\varepsilon(s,x) + b(u_\varepsilon(s,x))\right) \mathrm{d}s\right] \\ &- \mathbb{E}\left[\int_0^t I_{u_\varepsilon(s,x)\le 0}\, \sigma(u_\varepsilon(s,x)) \mathrm{d}W_x^\varepsilon(s) - L_x^0(t)\right] \end{aligned}$$

where L_x^0 is the local time of $u_\varepsilon(\cdot,x)$ at zero, which is non-negative. Now, Liptchiz condition and (3.7) asserts $b(u) \ge -L|u|$, and by the expression for $\mathcal{L}_\varepsilon$ we get the second term (including the minus sign) is no bigger than

$$\begin{aligned} &\mathbb{E}\left[\int_0^t \frac{1}{\varepsilon} \mathbf{1}_{u_\varepsilon(s,x)\le 0} u_\varepsilon(s,x) - \frac{1}{\varepsilon} \mathbf{1}_{u_\varepsilon(s,x)\le 0} P_\varepsilon u(s)(x) + L\, \mathbf{1}_{u_\varepsilon(s,x)\le 0} |u_\varepsilon(s,x)| \mathrm{d}s\right] \\ &\le (L + 1/\varepsilon) \int_0^t \mathbb{E}[\varphi(u_\varepsilon(s,x))]\, \mathrm{d}s - \frac{1}{\varepsilon} \int_0^t \int_{\mathbb{X}} p(\varepsilon,x,y) \mathbb{E}\left[\mathbf{1}_{u_\varepsilon(s,x)\le 0} u_\varepsilon(s,y)\right] m(\mathrm{d}y)\, \mathrm{d}s \end{aligned}$$

where $\mathbf{1}_A(\cdot)$ denote the indicator function on the set A. Note $\varphi(f) \equiv 0$ on $\mathbb{X}$ since $f \ge 0$. By [44, Proposition 9.9], we have almost surely that

$$L_x^0(t) = \lim_{\delta\downarrow 0} \frac{1}{2\delta} \int_0^t 1_{\{|u_\varepsilon(s,x)|\le\delta\}} \mathrm{d}\left\langle u_\varepsilon(\cdot,x)\right\rangle_s$$

$$
\begin{aligned}
&= \lim_{\delta\downarrow 0} \frac{1}{2\delta}\int_0^t 1_{\{|u_\varepsilon(s,x)|\le\delta\}}\sigma\big(u_\varepsilon(x,s)\big)^2 \mathrm{d}\,\langle W_x^\varepsilon(\cdot)\rangle_s \\
&\overset{(3.7)}{\le} \lim_{\delta\downarrow 0} \frac{1}{2\delta}\int_0^t 1_{\{|u_\varepsilon(s,x)|\le\delta\}} \mathrm{d}\,\langle W_x^\varepsilon(\cdot)\rangle_s \\
&\overset{(8.10)}{\le} k_\varepsilon(x,x)\lim_{\delta\downarrow 0} \frac{1}{2\delta}\int_0^t 1_{\{|u_\varepsilon(s,x)|\le\delta\}}\delta^2 \mathrm{d}s \\
&= 0.
\end{aligned}
$$

Hence

$$
\mathbb{E}[\varphi(u_\varepsilon(t,x))] \le (L_T + 1/\varepsilon)\int_0^t \mathbb{E}\left[\varphi(u_\varepsilon(s,x))\right]\mathrm{d}s + \ (1/\varepsilon)\int_0^t\int_{\mathbb{X}} p(\varepsilon,x,y)\mathbb{E}\left[\varphi(u_\varepsilon(s,y)\right] m(\mathrm{d}y)\,\mathrm{d}s.
$$

Since $p_t(x,\cdot)$ is a (sub)probability density, by Gronwall's lemma to see $\sup_{x\in\mathbb{X}}\mathbb{E}\left[\varphi(u_\varepsilon(t,x))\right]=0$ for all $t\ge 0$, which means for each $(t,x)\in\mathbb{R}_+\times\mathbb{X}$, $u_\varepsilon(t,x)\ge 0$ almost surely.

Let u_t be the unique solution to the equation (2.3), we will show $u_\varepsilon \to u$ in an appropriate sense as $\varepsilon\downarrow 0$. Consider

$$
\begin{aligned}
\big|p(t-s,x,y)b\big(u(s,y)\big) - p_\varepsilon(t-s,x,y)b\big(u_\varepsilon(s,y)\big)\big| &\le e^{-(t-s)/\varepsilon}\left|b\Big(u_\varepsilon(s,y)\Big)\right| \\
&+ \left|R_\varepsilon(t-s,x,y)b\Big(u(s,y)\Big) - p(t-s,x,y)b\Big(u(s,y)\Big)\right| \\
&+ \left|R_\varepsilon(t-s,x,y)b\Big(u(s,y)\Big) - R_\varepsilon(t-s,x,y)b\Big(u_\varepsilon(s,y)\Big)\right|.
\end{aligned}
$$

Also, by stochastic Fubini's,

$$
\begin{aligned}
&\int_{\mathbb{X}}\int_0^t R_\varepsilon(t-s,x,y)\sigma\big(u_\varepsilon(s,y)\big)\mathrm{d}W_y^\varepsilon(s)\,m(\mathrm{d}y) - \int_0^t\int_{\mathbb{X}} p(t-s,x,z)\sigma\big(u(s,z)\big)W(\mathrm{d}z,\mathrm{d}s) \\
&= \int_0^t\int_{\mathbb{X}}\int_{\mathbb{X}} R_\varepsilon(t-s,x,y)\sigma\big(u_\varepsilon(s,y)\big)p_\varepsilon(y,z)\,m(\mathrm{d}y)\,W(\mathrm{d}z,\mathrm{d}s) - \int_0^t\int_{\mathbb{X}} p(t-s,x,y)\sigma\big(u(s,y)\big)W(\mathrm{d}y,\mathrm{d}s) \\
&= \int_0^t\int_{\mathbb{X}}\int_{\mathbb{X}} R_\varepsilon(t-s,x,y)\big(\sigma\big(u_\varepsilon(s,y)\big) - \sigma\big(u(s,y)\big)\big)p_\varepsilon(y,z)\,m(\mathrm{d}y)\,W(\mathrm{d}z,\mathrm{d}s) \\
&\quad + \int_0^t\int_{\mathbb{X}}\int_{\mathbb{X}} R_\varepsilon(t-s,x,y)\big(\sigma\big(u(s,y)\big) - \sigma\big(u(s,z)\big)\big)p_\varepsilon(y,z)\,m(\mathrm{d}y)\,W(\mathrm{d}z,\mathrm{d}s) \\
&\quad + \int_0^t\int_{\mathbb{X}}\int_{\mathbb{X}} R_\varepsilon(t-s,x,y)\sigma\big(u(s,z)\big)p_\varepsilon(y,z)\,\mathrm{d}y - p(t-s,x,z)\sigma\big(u(s,z)\big)\,W(\mathrm{d}z,\mathrm{d}s).
\end{aligned}
$$

We obtain uniformly in $(t,x)\in[0,T]\times\mathbb{X}$,

$$
\begin{aligned}
\mathbb{E}\left[|u_\varepsilon(t,x)-u(t,x)|^2\right] &\lesssim |P_\varepsilon(t)f(x) - P(t)f(x)|^2 + \mathbb{E}\left[\left|\int_0^t e^{-(t-s)/\varepsilon}b(u_\varepsilon(s,x))\,\mathrm{d}s\right|^2\right] \\
&+ \mathbb{E}\left[\left|\int_0^t\int_{\mathbb{X}} R_\varepsilon(t-s,x,y)\Big(b(u_\varepsilon(s,y)) - b(u(s,y))\Big)m(\mathrm{d}y)\,\mathrm{d}s\right|^2\right] \\
&+ \mathbb{E}\left[\left|\int_0^t\int_{\mathbb{X}}\Big(R_\varepsilon(t-s,x,y) - p(t-s,x,y)\Big)b(u(s,y))\,\mathrm{d}s\,m(\mathrm{d}y)\right|^2\right] \\
&+ \int_0^t e^{-2(t-s)/\varepsilon}\sup_{x\in\mathbb{X}}\mathbb{E}\left[a(u_\varepsilon(s,x))^2\right]\varepsilon^{-\alpha}\mathrm{d}s \\
&+ \int_0^t \mathbb{E}\left[\left\|\int_{\mathbb{X}} R_\varepsilon(t-s,x,y)\left(\sigma(u_\varepsilon(s,z)) - \sigma(u(s,z))\right)p_\varepsilon(y,\cdot)m(\mathrm{d}y)\right\|_{\widetilde{K}}^2\right]\mathrm{d}s
\end{aligned}
$$

$$
\begin{aligned}
&+\int_0^t \mathbb{E}\left[\left\|\int_{\mathbb{X}} R_\varepsilon(t-s,x,y)\left|\sigma(u(s,y))-\sigma(u(s,\cdot))\right| p_\varepsilon(y,\cdot)m(\mathrm{d}y)\right\|_{\widetilde{K}}^2\right]\mathrm{d}s\\
&+\int_0^t \mathbb{E}\left[\left\|\left(\int_{\mathbb{X}} R_\varepsilon(t-s,x,y)p_\varepsilon(y,\cdot)m(\mathrm{d}y)-p(t-s,x,\cdot)\right)\sigma(u(s,\cdot))\right\|_{\widetilde{K}}^2\right]\mathrm{d}s\\
:=&\sum_{i=1}^8 J_i(\varepsilon,t,x).
\end{aligned}
$$

Using (8.8) of Lemma (8.13) and Jessen's inequality to see uniformly in $\varepsilon>0$, $t\in[0,T]$

$$
\sup_{x\in\mathbb{X}} J_1(\varepsilon,t,x)\lesssim e^{-c_T\frac{t}{\varepsilon}}+\frac{\gamma_\varepsilon}{t^\beta},\quad \lim_{\varepsilon\downarrow 0}\sup_{t\in[0,T]}\sup_{x\in\mathbb{X}} J_4(\varepsilon,t,x)=0,
$$

where β is as in (H_3). Also, by (3.5), (3.7) and (8.12), we have

$$
\lim_{\varepsilon\downarrow 0}\sup_{t\in[0,T]}\sup_{x\in\mathbb{X}} J_2(\varepsilon,t,x)+J_5(\varepsilon,t,x)=0.
$$

By (8.12) and Lemma 8.13,

$$
\lim_{\varepsilon\downarrow 0}\sup_{t\in[0,T]}\sup_{x\in\mathbb{X}} J_8(\varepsilon,t,x)\lesssim \lim_{\varepsilon\downarrow 0}\sup_{x\in\mathbb{X}}\int_0^T\left\|\int_{\mathbb{X}} R_\varepsilon(t-s,x,y)p_\varepsilon(y,\cdot)m(\mathrm{d}y)-p(t-s,x,\cdot)\right\|_{\widetilde{K}}^2\mathrm{d}s\overset{(8.7)}{=}0.
$$

Let $U_\varepsilon(t)=\sup_{x\in\mathbb{X}}\mathbb{E}\left[|u_\varepsilon(t,x)-u(t,x)|^2\right]$, by (3.5), we see that uniformly in $\varepsilon\in(0,1)$ and $(t,x)\in[0,T]\times\mathbb{X}$ that

$$
\begin{aligned}
J_3(\varepsilon,t,x)&\lesssim\int_0^t\left(\int_{\mathbb{X}} R_\varepsilon(t-s,x,y)m(\mathrm{d}y)\right)^2 U_\varepsilon(s)\mathrm{d}s\\
&\le\int_0^t U_\varepsilon(s)\mathrm{d}s\\
&\lesssim\int_0^t (t-s)^{-\alpha}U_\varepsilon(s)\mathrm{d}s.
\end{aligned}
$$

Similarly, uniformly in $\varepsilon\in(0,1)$ and $(t,x)\in[0,T]\times\mathbb{X}$,

$$
J_6(\varepsilon,t,x)\lesssim\int_0^t (t-s)^{-\alpha}U_\varepsilon(s)\mathrm{d}s.
$$

Now we consider J_7, by Minkowski's inequality we see for each $\varepsilon\in(0,1)$ and $(t,x)\in[0,T]\times\mathbb{X}$ that

$$
J_7(\varepsilon,t,x)\le\int_0^t\left\|\int_{\mathbb{X}} R_\varepsilon(t-s,x,y)\left\|\sigma(u(s,y))-\sigma(u(s,\cdot))\right\|_{L^2(\mathbb{P})} p_\varepsilon(y,\cdot)m(\mathrm{d}y)\right\|_{\widetilde{K}}^2 ds.
$$

Observe from the proof of Theorem 5.1, there exists $\xi\in(0,1]$ so that uniformly for $s\in[0,T]$ and $y,z\in\mathbb{X}$

$$
\|\sigma(u(s,y))-\sigma(u(s,z))\|_{L^2(\mathbb{P})}\overset{(3.5)}{\lesssim}|P_s u_0(y)-P_s u_0(z)|+d(y,z)^\xi\wedge 1.
$$

Since $u_0\in\mathcal{C}_0(\mathbb{X})$, by Feller property, we see that $P_t u_0(\cdot)$ is uniformly continuous on $\mathbb{X}$, and hence for each $\delta>0$, there is $r_\delta>0$ so that

$$
\|\sigma(u(s,y))-\sigma(u(s,z))\|_{L^2(\mathbb{P})}\le\delta,\qquad\text{for all } d(y,z)\le r_\delta,\ \text{and } s\in[0,T].
$$

Therefore, it holds uniformly in $z\in\mathbb{X}$ that

$$
\begin{aligned}
&\int_{\mathbb{X}} R_\varepsilon(t-s,x,y)\left\|\sigma(u(s,y))-\sigma(u(s,z))\right\|_{L^2(\mathbb{P})} p_\varepsilon(y,z)m(\mathrm{d}y)\\
\le&\left(\delta\int_{\mathbb{X}}+\int_{d(z,y)>r_\delta}\right)R_\epsilon(t-s,x,y)p_\epsilon(y,z)m(\mathrm{d}y)
\end{aligned}
$$

$$\leq \delta R_\epsilon(t-s+\epsilon,x,z) + \int_{\mathbb{X}} 1_{d(y,z)>r_\delta} R_\epsilon(t-s,x,y) p_\epsilon(y,z) m(\mathrm{d}y).$$

Hence, for each $T>0$, there is $C=C(T)>0$ so that for all $\epsilon\in(0,1)$ and $(t,x)\in(0,T]\times\mathbb{X}$,

$$J_7(\epsilon,t,x) \leq C\int_0^t J_{7,1}(\epsilon,t-s,x) + J_{7,2}(\epsilon,t-s,x)\mathrm{d}s$$

where

$$J_{7,1}(\epsilon,t-s,x) := \delta \left\| R_\epsilon(t-s,y,\cdot)\right\|_{\widetilde{K}}^2 \overset{(8.8)}{\leq} C\delta\left((t-s)^{-\alpha}\vee 1\right),$$

and the constant in the inequality does not depend on δ. Also

$$\begin{aligned}
J_{7,2}(\epsilon,t-s,x) :&= \left\|\int_{\mathbb{X}} 1_{d(y,z)>r_\delta} R_\epsilon(t-s,x,y)p_\epsilon(y,\cdot)m(\mathrm{d}y)\right\|_{\widetilde{K}}^2 \\
&= \int_{\mathbb{X}^2}\int_{\mathbb{X}^2} R_\epsilon(t-s,x,y_1)R_\epsilon(t-s,x,y_2)p_\epsilon(y_1,z_1)p_\epsilon(y_2,z_2) \\
&\qquad\qquad \times \mathbf{1}_{d(y_1,z_1)>r_\delta}\mathbf{1}_{d(y_2,z_2)>r_\delta} m(\mathrm{d}y_1)\, m(\mathrm{d}y_2)\, \widetilde{K}(\mathrm{d}z_1,\mathrm{d}z_2) \\
&= \int_{\mathbb{X}^2} R_\epsilon(t-s,x,y_1)R_\epsilon(t-s,x,y_2)\Bigg(\int_{\mathbb{X}^2}\mathbf{1}_{B(y_1,r_\delta)^c}(z_1)p_\epsilon(y_1,z_1) \\
&\qquad\qquad \times \mathbf{1}_{B(y_2,r_\delta)^c}(z_2)p_\epsilon(y_2,z_2)\widetilde{K}(\mathrm{d}z_1,\mathrm{d}z_2)\Bigg) m(\mathrm{d}y_1)m(\mathrm{d}y_2) \\
&\leq \int_{\mathbb{X}^2} R_\epsilon(t-s,x,y_1)R_\epsilon(t-s,x,y_2) \\
&\qquad\qquad \times \left\|p_\epsilon(y_1,\cdot)1_{B(y_1,r_\delta)^c}\right\|_{\widetilde{K}}\left\|p_\epsilon(y_2,\cdot)1_{B(y_2,r_\delta)^c}\right\|_{\widetilde{K}} m(\mathrm{d}y_1)m(\mathrm{d}y_2) \\
&\leq \sup_{x\in\mathbb{X}}\left\|p_\epsilon(x,\cdot)\mathbf{1}_{B(x,r_\delta)}(\cdot)\right\|_{\widetilde{K}}^2
\end{aligned}$$

Where we used Fubini's in the first and third equality, Lemma 8.1 for the first inequality. Hence by (H_2) we see uniformly in $\epsilon\in(0,1]$,

$$\begin{aligned}
\lim_{\epsilon\downarrow 0}\sup_{t\in[0,T]}\sup_{x\in\mathbb{X}} J_7(\epsilon,t,x) &\leq C\left(\delta + \lim_{\varepsilon\downarrow 0}\sup_{y\in\mathbb{X}}\left\|p_\epsilon(y,\cdot)\mathbf{1}_{B(y,r_\delta)^c}\right\|_{\widetilde{K}}^2\right) \\
&\overset{(2.7)}{\leq} C\delta.
\end{aligned}$$

Since $\delta>0$ is arbitrary and the constants does not depend on δ, we see $\lim_{\varepsilon\downarrow 0}\sup_{x\in\mathbb{X}} J_7(\varepsilon,t,x)=0$.

Collecting all terms, we have for all $t\in[0,T]$,

$$U_\varepsilon(t) \lesssim \int_0^t (t-s)^{-\alpha}U_\varepsilon(s)\mathrm{d}s + H_\varepsilon(t) + e^{-c_T\frac{t}{\varepsilon}} + \frac{\gamma_\varepsilon}{t^\beta}$$

for some $H_\varepsilon(t)$ with $\lim_{\varepsilon\downarrow 0}\sup_{t\in[0,T]} H_\varepsilon(t)=0$. By Grönwall's Lemma, we see uniformly for $t\in(0,T]$ that

$$U_\varepsilon(t) \lesssim H_\varepsilon(t) + e^{-c_T\frac{t}{\varepsilon}} + \frac{\gamma_\varepsilon}{t^\beta} + \int_0^t \left(H_\varepsilon(s) + e^{-c_T\frac{s}{\varepsilon}} + \frac{\gamma_\varepsilon}{s^\beta}\right)(t-s)^{-\alpha}\exp\left(\int_s^t (t-\tau)^{-\alpha}\mathrm{d}\tau\right)\mathrm{d}s$$

which converges to zero for each $t>0$ by the dominated convergence theorem. By continuity of sample paths, the non-negativity of u is inherited from u_ε. □

Proof of Proposition 8.12. Fix a $T>0$ and set $v_n(t,x) := |u_t^n(x) - u_t(x)|$ for $(t,x)\in(0,T]\times\mathbb{X}$, consider by Minkowski's inequality and Corollary 8.4 that for $p\geq 2$

$$\|v_n(t,x)\|_{L^p(\mathbb{P})} \lesssim \int_{\mathbb{X}} p_t(x,y)|f_n(y)-f(y)|m(\mathrm{d}y) + \int_0^t\int_{\mathbb{X}} p_t(x,y)\left\|v_n(s,y)\right\|_{L^p(\mathbb{P})} m(\mathrm{d}y)\mathrm{d}s$$

$$+\left(\int_0^t \left\|p_{t-s}(x,\cdot)\left\|v_n(s,\cdot)\right\|_{L^p(\mathbb{P})}\right\|_{\widetilde{K}}^2\right)^{\frac{1}{2}},$$

we shall call the right hand side of the inequality $I_i(t,x)$ for $i=0,1,2$ respectively. For $\lambda\in(0,1/2]$, we see by Assumption 2.6, it holds uniformly in $(t,x)\in[0,T]\times\mathbb{X}$ that

$$\begin{aligned}I_0(t,x)&\le\int_{\mathbb{X}}p_t(x,y)h(y)^\lambda m(\mathrm{d}y)\left\|f_n-f\right\|_{(\lambda)}\\&\lesssim h(x)^\lambda\left\|f_n-f\right\|_{(\lambda)}.\end{aligned}$$

Similarly,

$$\begin{aligned}I_1(t,x)&\le\int_0^t\int_{\mathbb{X}}p_{t-s}(x,y)h(y)^\lambda\sup_{\tau\in[0,s]}\left\|\left\|u_n(\tau,\cdot)\right\|_{L^p(\mathbb{P})}\right\|_{(\lambda)}\mathrm{d}s\\&\lesssim h(x)^\lambda\int_0^t\sup_{\tau\in[0,s]}\left\|\left\|u_n(\tau,\cdot)\right\|_{L^p(\mathbb{P})}\right\|_{(\lambda)}\mathrm{d}s.\end{aligned}$$

Finally,

$$\begin{aligned}I_2(t,x)&\lesssim\left(\int_0^t\left\|p_{t-s}(x,\cdot)h(\cdot)^\lambda\right\|_{\widetilde{K}}^2\sup_{\tau\in[0,s]}\left\|\left\|v_n(\tau,\cdot)\right\|_{L^p(\mathbb{P})}\right\|_{(\lambda)}^2\mathrm{d}s\right)^{\frac{p}{2}}\\&\lesssim h(x)^\lambda\left(\int_0^t(t-s)^{-\alpha}\sup_{\tau\in[0,s]}\left\|\left\|v_n(\tau,\cdot)\right\|_{L^p(\mathbb{P})}\right\|_{(\lambda)}^2\mathrm{d}s\right)^{\frac{1}{2}}.\end{aligned}$$

Hence, if we set $U_n(t):=\sup_{s\in[0,t]}\left\|\left\|v_n(s,\cdot)\right\|_{L^p(\mathbb{P})}\right\|_{(\lambda)}$, then we see the following inequality holds uniformly in $t\in[0,T]$:

$$U_n(t)^2\lesssim\left\|f_n-f\right\|_{(\lambda)}^2+\int_0^t U_n(s)^2\mathrm{d}s+\int_0^t(t-s)^{-a}U_n(s)^2\mathrm{d}s.$$

By Grönwall's inequality, we obtain uniformly in $t\in[0,T]$ that

$$U_n(t)^2\le\left\|f_n-f\right\|_{(\lambda)}^2\left(1+\int_0^t\left(1+(s-\tau)^{-\alpha}\right)\exp\left(\int_0^s(1+(t-s)^{-\alpha}\mathrm{d}\tau\right)\mathrm{d}s\right)$$

which converges to zero uniformly in $t\in[0,T]$ as $n\to\infty$. □

8.4. **Proof of Theorem 5.4 and 5.5.**

Proof of Theorem 5.4. We precede as [55], let $\{\sigma_n\}_n$ and $\{b_n\}_n$ be two sequence of global Lipchitz functions that uniformly approximate σ and b respectively with almost surely that

$$\begin{gathered}|\sigma_n(t,x,u)|+|b_n(t,x,u)|\le L(1+|u|),\quad u\in\mathbb{R},\ \text{ for all } n\ge1,\\a_n(t,x,0)=0,\ b_n(t,x,0)\ge0,\quad n\ge1.\end{gathered}$$

Hence by Theorem 5.2 there exists a unique non-negative $\mathcal{C}_{tem}$-valued solution for each n, $u_n(t)$ with continuous sample paths. Therefore, let

$$H_{n,\lambda}^{2p}(t)=\sup_{x\in\mathbb{X}}h(x)^{-\lambda}\mathbb{E}\left[|u_n(t,x)|^{2p}\right],$$

we may obtain the following inequality by a similar calculation to the proof of Proposition 8.11 that for each $T>0$, $p\ge2$ and $\lambda\in(0,1/2)$, there exists $C_{T,\lambda,p}>0$ which is independent of $n\ge1$ so that

$$H_{n,\lambda}^{2p}(t)\le C_{T,\lambda,p}\int_0^t\left(1+(t-s)^{-\alpha}\right)H_{n,\lambda}^{2p}(s)\mathrm{d}s,\quad t\in[0,T].$$

Apply Gronwall's lemma to see for all $T>0$, there is $C'_{T,\lambda,p}$ independent of n with

$$\sup_{0\le t\le T} H^{2p}_{n,\lambda}(t) < C'_{T,\lambda,p} \qquad \text{for all } n\ge 1,$$

which gives tightness to $\{u_n(t,x)\}_n$ for each $(t,x)\in[0,\infty)\times\mathbb{X}$. Now argue in the same line as in the proof of Theorem 5.1 to get the moment bound as in Lemma 8.6. One then can check that any limit point of the probability measures on induced by u_n's solves the desired SPDE. □

Proof of Theorem 5.5. Let u be a non-negative $\mathcal{C}_{tem}$ valued solution the the SPDE (2.3) with initial condition $u(0)=u_0\in\mathcal{C}^+_{rap}$, then $u_t\ge 0$ almost surely for all $t\ge 0$ by Theorem 5.2. Using condition (5.1) and (2.8) to see for all $\lambda>0$ and $t\in[0,T]$, there is a $C_{T,\lambda}>0$ and

$$\mathbb{E}\left[u(t,x)\right] \le C_{T,\lambda} h(x)^{-\lambda} + \int_0^t \int_{\mathbb{X}} p(t-s,x,y)\mathbb{E}\left[u(s,y)\right] m(\mathrm{d}y)\mathrm{d}s$$

By Lemma 8.5 and (2.8), we see for all $\lambda>0$,

$$\sup_{t\in[0,T]}\sup_{x\in\mathbb{X}} \mathbb{E}\left[u(t,x)\right] h(x)^{\lambda} < \infty. \tag{8.13}$$

For each $\lambda>0$ and $p\ge 1/2$, we see by Jessen's inequality and Corollary 8.4 that uniformly for $(t,x)\in[0,T]\times\mathbb{X}$

$$\begin{aligned}
\mathbb{E}\left[|u_t(x)|^{2p}\right] &\lesssim \int_{\mathbb{X}} p_t(x,y)\,|u_0(y)|^{2p}\, m(\mathrm{d}y) + \int_0^t\int_{\mathbb{X}} p_{t-s}(x,y)\, \|b(u_s(y))\|^{2p}_{L^{2p}(\mathbb{P})}\, m(\mathrm{d}y)\mathrm{d}s \\
&\quad + \int_0^t \left\| p_t(x,\cdot)\, \|\sigma(u_s(\cdot))\|^{p}_{L^{2p}(\mathbb{P})}\right\|^2_{\widetilde{K}} \mathrm{d}s \\
&\overset{(5.1)}{\lesssim} h(x)^{-\lambda} + \int_0^t\int_{\mathbb{X}} p_{t-s}(x,y)\,\|u_s(y)\|^{2p}_{L^{2p}(\mathbb{P})}\, m(\mathrm{d}y)\mathrm{d}s \\
&\quad + \int_0^t \left\| p_{t-s}(x,\cdot)\left(\|u_s(\cdot)\|^{p}_{L^{2p}(\mathbb{P})} + \|u_s(\cdot)\|^{p\theta}_{L^{2p\theta}(\mathbb{P})}\right)\right\|^2_{\widetilde{K}} \mathrm{d}s \\
&\overset{(8.1)}{\lesssim} h(x)^{-\lambda} + \int_0^t\int_{\mathbb{X}} p_{t-s}(x,y)\,\|u_s(y)\|^{2p}_{L^{2p}(\mathbb{P})}\, m(\mathrm{d}y)\mathrm{d}s \\
&\quad + \int_0^t \left\| p_{t-s}(x,\cdot)\,\|u_s(\cdot)\|^{p}_{L^{2p}(\mathbb{P})}\right\|^2_{\widetilde{K}} + \left\| p_{t-s}(x,\cdot)\,\|u_s(\cdot)\|^{p\theta}_{L^{2p\theta}(\mathbb{P})}\right\|^2_{\widetilde{K}} \mathrm{d}s
\end{aligned}$$

By (8.13) and Lemma 8.5 again, we see for $p=\frac{1}{2\theta}>\frac{1}{2}$,

$$\sup_{t\in[0,T]}\sup_{x\in\mathbb{X}} h^{\lambda}(x)\mathbb{E}\left[|u(t,x)|^{2p}\right] < \infty. \tag{8.14}$$

Now argue inductively, i.e. if (8.14) holds for $p=\frac{1}{2\theta^n}$, then it holds for $p=\frac{1}{2\theta^{n+1}}$, we see (8.14) holds for all $p\ge\frac{1}{2}$ and $\lambda>0$. Finally, by Lemma 8.7 and mimicking the proof of Proposition 8.11, we see if h is a good weight, then $\mathbb{P}\left(u\in C(\mathbb{R}_+;\mathcal{C}_{rap})\right)=1$. □

9. Appendix

We first give the proofs of Proposition 4.1 and Proposition 4.3. We will need the following generalized reflection principle for the proof of Proposition 4.1 which is analogous to [7, Theorem 2.7].

Lemma 9.1. *Let $(\mathbb{X},d)$ be a Polish space and $(Y_t)_{t\ge0}$ is a $\mathbb{X}$-valued strong Markov process with almost sure continuous sample paths, whose law is denoted as $(\Pi_x)_{x\in\mathbb{X}}$. Then for any $t>0$,*

$$\sup_{x\in\mathbb{X}} \Pi_x\left(\sup_{0\le s\le t} d(Y_s,x)\ge\lambda\right) \le 2\sup_{y\in\mathbb{X},0\le s\le t} \Pi_y\left(d(Y_s,y)\ge\frac{\lambda}{2}\right). \tag{9.1}$$

Proof. For $\lambda > 0$, let $\tau_\lambda := \inf\{t > 0 : d(Y_t, Y_0) \geq \lambda\}$ which is a stopping time with respect to the natural filtration of Y by continuity of sample paths. We first note that

$$\{\tau_\lambda < \infty\} \subset A_\lambda, \quad \text{where } A_\lambda := \left\{d(Y_t, Y_{\tau_\lambda}) \geq \frac{\lambda}{2}\right\} \bigcup \left\{d(Y_t, Y_x) \geq \frac{\lambda}{2}\right\}. \tag{9.2}$$

To see this, we observe by triangular inequality that on the event $\{\tau_\lambda < \infty\} \cap A_\lambda^c$, we have

$$\lambda = d(Y_0, Y_{\tau_\lambda}) \leq d(Y_t, Y_0) + d(Y_t, Y_{\tau_\lambda}) < \lambda$$

where the first equality follows from the continuity of sample paths and $\tau_\lambda < \infty$. This implies $\{\tau_\lambda < \infty\} \cap A_\lambda^c = \emptyset$ and (9.2) holds. We see by strong Markov property that for all $\lambda > 0$ and $(t, x) \in \mathbb{R}_+ \times \mathbb{X}$,

$$\begin{aligned}
\Pi_x\left(\sup_{0 \leq s \leq t} d(Y_s, x) \geq \lambda\right) &= \Pi_x(\tau_\lambda \leq t) \\
&\overset{(9.2)}{=} \Pi_x\left(\tau_\lambda \leq t,\, \left\{d(Y_t, x) \geq \frac{\lambda}{2}\right\} \cup \left\{d(Y_{\tau_\lambda}, Y_t) \geq \frac{\lambda}{2}\right\}\right) \\
&\leq \Pi_x\left(d(Y_t, x) \geq \frac{\lambda}{2}\right) + \Pi_x\left(\tau_\lambda \leq t,\, d(Y_{\tau_\lambda}, Y_t) \geq \frac{\lambda}{2}\right) \\
&\leq \Pi_x\left(d(Y_t, x) \geq \frac{\lambda}{2}\right) + \Pi_x\left(d(Y_{\tau_\lambda \wedge t}, Y_t) \geq \frac{\lambda}{2}\right) \\
&= \Pi_x\left(d(Y_t, x) \geq \frac{\lambda}{2}\right) + \Pi_x\left[\Pi_{Y_{\tau_\lambda \wedge t}}\left(d(Y_{t - \tau_\lambda \wedge t}, Y_0) \geq \frac{\lambda}{2}\right)\right] \\
&= \Pi_x\left(d(Y_t, x) \geq \frac{\lambda}{2}\right) + \int_0^t \Pi_x\left[\Pi_{Y_s}\left(d(Y_{t-s}, Y_0) \geq \frac{\lambda}{2}\right); \tau_\lambda \in \mathrm{d}s\right] \\
&\leq \Pi_x\left(d(Y_t, x) \geq \frac{\lambda}{2}\right) + \sup_{y \in \mathbb{X}, 0 \leq s \leq t} \Pi_y\left(d(Y_s, Y_0) \geq \frac{\lambda}{2}\right) \\
&\leq 2 \sup_{y \in \mathbb{X}, 0 \leq s \leq t} \Pi_y\left(d(Y_s, Y_0) \geq \frac{\lambda}{2}\right).
\end{aligned}$$

□

Proof of Proposition 4.1. Fix any $T > 0$ and let $t \in [0, T]$. Using monotonicity of Φ and V to get uniformly in $(t, x) \in [0, T] \times \mathbb{X}$ that

$$\begin{aligned}
\Pi_x\left(d(Y_t, x) \geq \frac{\lambda}{2}\right) &\overset{(4.1)}{\lesssim} t^{-\frac{d_h}{d_w}} \int_{B(x, \lambda/2)^c} \Phi\left(\frac{d(x,y)}{t^{\frac{1}{d_w}}}\right) m(\mathrm{d}y) \\
&= t^{-\frac{d_h}{d_w}} \sum_{n=0}^{\infty} \int_{B(x, 2^{n+1}\lambda/2) \setminus B(x, 2^n \lambda/2)} \Phi\left(\frac{d(x,y)}{t^{\frac{1}{d_w}}}\right) m(\mathrm{d}y) \\
&\leq \sum_{n=0}^{\infty} m(B(x, 2^{n+1}\lambda/2) t^{-\frac{d_h}{d_w}} \Phi\left(\frac{2^n \lambda}{2t^{\frac{1}{d_w}}}\right) \\
&\lesssim \sum_{n=0}^{\infty} V(2^n \lambda) \left(\frac{2^n \lambda}{t^{\frac{1}{d_w}}}\right)^{\alpha} \Phi\left(\frac{2^{n+1}\lambda}{4t^{\frac{1}{d_w}}}\right) \frac{2^n}{2^{n+1}} \\
&\leq \int_1^{\infty} V(\lambda r) \left(\frac{\lambda r}{t^{\frac{1}{d_w}}}\right)^{\alpha} \Phi\left(\frac{\lambda r}{4t^{\frac{1}{d_w}}}\right) \frac{dr}{r} \\
&= \int_{\lambda/(4t^{\frac{1}{d_w}})}^{\infty} V\left(4t^{\frac{1}{d_w}} r\right) (4r)^{\alpha} \Phi(r) \frac{dr}{r}. \\
&\lesssim \int_{\lambda/(4t^{\frac{1}{d_w}})}^{\infty} V\left(4T^{\frac{1}{d_w}} r\right) r^{\alpha} \Phi(r) \frac{dr}{r},
\end{aligned}$$

which is finite by (4.2). For $t > 0$ and $\lambda \geq 0$, denote

$$D_t(\lambda) := \int_{\lambda/(4t^{\frac{1}{d_w}})}^{\infty} V\left(4T^{\frac{1}{d_w}} r\right) r^{d_h} \Phi(r) \frac{dr}{r},$$

and observe $D_\lambda(t)$ is non-decreasing in $t \in (0, T]$ for each $\lambda \geq 0$. Furthermore, let $p > \max\{3, 2d_w\}$ as in the assumption, we see by L'Hôpital's rule and (4.3) that for each $t \in [0, T]$,

$$\lim_{\lambda \to \infty} D_t(\lambda) \lambda^p = 0, \tag{9.3}$$

$$\frac{d}{d\lambda} D_t(\lambda) = -V\left((T/t)^{\frac{1}{d_w}} \lambda\right) \left(\frac{\lambda}{t^{\frac{1}{d_w}}}\right)^{d_h} \Phi\left(\frac{\lambda}{4t^{\frac{1}{d_w}}}\right) \frac{1}{\lambda}. \tag{9.4}$$

By Lemma 9.1 and monotonicity,

$$\begin{aligned}
\Pi_x\left[\sup_{0 \leq s \leq t} d(Y_t, x)^p\right] &= \int_0^\infty \Pi_x\left(\sup_{0 \leq s \leq t} d(Y_t, x) \geq \lambda\right) p\lambda^{p-1} d\lambda \\
&\overset{(9.1)}{\leq} 2 \int_0^\infty \sup_{x \in \mathbb{X}} \sup_{0 \leq s \leq t} \Pi_x\left(d(X_s, x) \geq \frac{\lambda}{2}\right) p\lambda^{p-1} d\lambda \\
&\leq 2 \int_0^\infty D_t(\lambda) p\lambda^{p-1} d\lambda \\
&= 2 \int_0^\infty V\left((T/t)^{\frac{1}{d_w}} \lambda\right) \left(\frac{\lambda}{4t^{\frac{1}{d_w}}}\right)^{d_h} \Phi\left(\frac{\lambda}{4t^{\frac{1}{d_w}}}\right) \lambda^p \frac{d\lambda}{\lambda}
\end{aligned}$$

where we used integration by parts with (9.4) and (9.3) in the last equality. We let $u = t^{-1/\beta}\lambda/4$ and perform a substitution, we see uniformly in $t \in (0, T]$,

$$\int_0^\infty V\left((T/t)^{\frac{1}{d_w}} \lambda\right) \left(\frac{\lambda}{4t^{\frac{1}{d_w}}}\right)^{d_h} \Phi\left(\frac{\lambda}{4t^{\frac{1}{d_w}}}\right) \lambda^p \frac{d\lambda}{\lambda} \lesssim t^{\frac{p}{d_w}} \int_0^\infty V\left(4T^{\frac{1}{d_w}} u\right) u^{d_h} \Phi(u) u^p \frac{du}{u}.$$

Hence we obtain that uniformly in $(t, x) \in [0, T] \times \mathbb{X}$,

$$\Pi_x\left[\sup_{0 \leq s \leq t} d(Y_t, x)^p\right] \lesssim t^{\frac{p}{d_w}} \int_0^\infty V\left(4T^{\frac{1}{d_w}} u\right) u^{d_h + p} \Phi(u) \frac{du}{u}.$$

Observe that the right hand side is independent of $x \in \mathbb{X}$ and $p/d_w > 2$, so we may use (4.2) to obtain the desired result. □

Proof of Proposition 4.3. Let $x_0 \in \mathbb{X}$ and $t, a, M > 0$ be arbitrary but fixed and $Q > 0$ be sufficiently large. Since $\beta > 1$, we see by [10, Exercise 2.4.11] that for each $x \in B(x_0, a + M/Q)$ there is $x' \in B(x_0, a - (M/Q)^{\frac{1}{d_w}})$ so that $d(x, x') \leq 2(\frac{M}{Q} + (M/Q)^{\frac{1}{d_w}})$. By triangular inequality and monotonicity of Φ, we see uniformly for $s \in \left[\frac{t}{Q}, \frac{2t}{Q}\right]$ and $x \in B(x_0, a + M/Q)$ that,

$$\begin{aligned}
\int_{B(x_0,a)} p_s(x, y) m(\mathrm{d}y) &\overset{(4.6)}{\gtrsim} \int_{B(x_0,a)} s^{-\frac{d_h}{d_w}} \Phi\left(\frac{d(x, y)}{s^{\frac{1}{d_w}}}\right) m(\mathrm{d}y) \\
&\geq \int_{B(x_0,a)} s^{-\frac{d_h}{d_w}} \Phi\left(\frac{d(x, x') + d(x', y)}{s^{\frac{1}{d_w}}}\right) m(\mathrm{d}y) \\
&\overset{(4.5)}{\gtrsim} \Phi\left(c\frac{d(x, x')}{s^{\frac{1}{d_w}}}\right) \int_{B\left(x', (\frac{M}{Q})^{\frac{1}{d_w}}\right)} s^{-\frac{d_h}{d_w}} \Phi\left(c\frac{d(x', y)}{s^{\frac{1}{d_w}}}\right) m(\mathrm{d}y) \\
&\overset{(4.4)}{\gtrsim} \Phi\left(2c\frac{\frac{M}{Q} + \left(\frac{M}{Q}\right)^{\frac{1}{d_w}}}{s^{\frac{1}{d_w}}}\right) s^{-\frac{d_h}{d_w}} \Phi\left(c\left(\frac{M}{Qs}\right)^{\frac{1}{d_w}}\right) \left(\frac{M}{Q}\right)^{\frac{d_h}{d_w}},
\end{aligned}$$

which is independent of $x \in B(x_0, a-(M/Q))$. Using the monitonicity again and the fact that $d_w > 1$, we obtain

$$\begin{aligned}\int_{B(x_0,a)} p_s(x,y)m(\mathrm{d}y) &\gtrsim \Phi\left(2ct^{\frac{1}{d_w}}\left(MQ^{\frac{1}{d_w}-1}+M^{\frac{1}{d_w}}\right)\right)\Phi\left(c\left(\frac{M}{t}\right)^{\frac{1}{d_w}}\right)\left(\frac{M}{t}\right)^{\frac{d_h}{d_w}}\\ &\geq \Phi\left(2ct^{\frac{1}{d_w}}\left(M+M^{\frac{1}{d_w}}\right)\right)\Phi\left(c\left(\frac{M}{t}\right)^{\frac{1}{d_w}}\right)\left(\frac{M}{t}\right)^{\frac{d_h}{d_w}}\end{aligned}$$

which is independent of Q and holds uniformly for $s \in [t/Q, 2t/Q]$ and $x \in B(x_0, a+M/Q)$. □

Proof of Proposition 4.5. We will assume for simplicity that equality holds in (4.7) with $c = 1$. Observe for each $\delta > 0$, there exists $C > 0$ so that the $\frac{d(x,y)}{\delta} \geq \mathbf{1}_{d(x,y)>\delta}$ for all $x, y \in \mathbb{X}$. Observe also that uniformly in $t \in (0,1]$ and $r > 0$,

$$r^{d_w}\Phi(r) \lesssim \exp\left(\left(1-2^{\frac{1-d_w}{d_w}}\right)r^{\frac{d_w}{d_w-1}}\right)\Phi(r) \leq \Phi(r/2).$$

Hence uniformly in $\epsilon \in (0,1]$ and $x, y \in \mathbb{X}$,

$$1_{d(x,y)>\delta}p_\epsilon(x,y) \leq \delta^{-d_w}\epsilon\,\epsilon^{-\frac{d_h}{d_w}}\left(\frac{d(x,y)}{\epsilon^{\frac{1}{d_w}}}\right)^{d_w}\Phi\left(d(x,y)/\epsilon^{\frac{1}{d_w}}\right) \lesssim \delta^{-d_w}\epsilon p_{2^{d_w}\epsilon}(x,y).$$

Hence, if we let $\alpha \in (0,1)$ be as in (2.8) we see uniformly in $\epsilon > 0$ and $x \in \mathbb{X}$,

$$\left\|p_\epsilon(x,\cdot)1_{B(x,\delta)^c}\right\|^2_{\widetilde{K}} \lesssim \delta^{-d_w}\epsilon\left\|p_{2^{d_w}\epsilon}(x,\cdot)\right\|^2_{\widetilde{K}} \overset{(2.8)}{\lesssim} \delta^{-d_w}\epsilon^{1-\alpha}.$$

Now (2.7) can be seen when taking supreme over $x \in \mathbb{X}$ on the left hand side and take $\epsilon \downarrow 0$. □

Proof of Lemma 4.11. The condition (2.1) follows from Remark 4.6. Fix an $\delta \in (d_h/d_w, 1)$ and denote by $(p_t^\delta(x,y) : t > 0,\ x,\ y \in M)$ the heat kernel generated by $\mathcal{L}^\delta$. It is known that there is family $(\eta_t^\delta)_{t\geq 0}$ of non-negative continuous functions on $(0,\infty)$ so that $\exp(-t\lambda^\delta) = \int_0^\infty \eta_t^\delta(s)e^{-s\lambda}ds$ for all $\lambda \geq 0$, and for $t > 0$, $x, y \in M$,

$$p_t^\delta(x,y) = \int_0^\infty \eta_t^\delta(s)p_s(x,y)ds. \tag{9.5}$$

η_t has the scaling property

$$\eta_t^\delta(s) = \frac{1}{t^{\frac{1}{\delta}}}\eta_1^\delta\left(\frac{s}{t^{\frac{1}{\delta}}}\right), \quad t, s > 0. \tag{9.6}$$

Let $\Phi_\delta(r) = (1+r)^{-(d_h+\delta d_w)}$ for $r > 0$, it was shown in [26] that uniformly in $t > 0$ and $x,\ y \in M$,

$$p_t^\delta(x,y) \asymp t^{-\frac{d_h}{d_w\delta}}\Phi_\delta\left(\frac{d(x,y)}{t^{\frac{1}{\delta d_w}}}\right),$$

and uniformly in $t > 0$ and $x,\ y,\ z \in M$ with $d(x,y) \leq t^{\frac{1}{\delta d_w}}$,

$$\left|p_t^\delta(z,x) - p_t^\delta(z,y)\right| \lesssim \left(\frac{d(x,y)}{t^{\frac{1}{\delta d_w}}}\right)^\theta\Phi_\delta\left(\frac{d(x,z)}{t^{\frac{1}{\delta d_w}}}\right). \tag{9.7}$$

We will first show (2.7) holds when $\widetilde{K}(\mathrm{d}z,\mathrm{d}y) = K(\mathrm{d}z,\mathrm{d}y) = \delta_{y=z}(\mathrm{d}z)m(\mathrm{d}y)$. We see for each $r > 0$, it holds uniformly in $\epsilon > 0$ and $x \in M$ that

$$\begin{aligned}\left\|p_\epsilon^\delta(x,\cdot)1_{B(x,\delta)^c}\right\|^2_{\widetilde{K}} &= \int_{d(x,y)>r} p_\epsilon^\delta(x,y)p_\epsilon^\delta(x,y)m(\mathrm{d}y)\\ &\lesssim \int_r^\infty \epsilon^{-\frac{2d_h}{\delta d_w}}\left(1+\frac{s}{\epsilon^{\frac{1}{\delta d_w}}}\right)^{-2(d_h+\delta d_w)} s^{d_h}\frac{ds}{s}\end{aligned}$$

$$\lesssim \int_r^\infty \epsilon^{-\frac{2d_h}{\delta d_w}} \left(1 + \frac{s}{\epsilon^{\frac{1}{\delta d_w}}}\right)^{-2(d_h+\delta d_w)} s^{2d_h} \frac{ds}{s}$$
$$\le \int_{r/\epsilon^{1/d_w}} (1+s)^{-2\delta d_w} \frac{ds}{s},$$

which converges to zero as $\epsilon \downarrow 0$.

Next, we show conditions (H$_3$) and (H$_1$) are satisfied. By (9.7), we see uniformly for $t > 0$ and $d(x,y) \le 1$,

$$\int_0^t \int_M \left|p_s^\delta(x,z) - p_s^\delta(y,z)\right| m(\mathrm{d}z)\mathrm{d}s \lesssim 2d(x,y)^{\delta d_w} + d(x,y)^\theta \int_{d(x,y)^{\delta d_w}}^t s^{-\frac{\theta}{\delta d_w}} \mathrm{d}s$$
$$\lesssim t^{1-\frac{\theta}{\delta d_w}} d(x,y)^\theta, \tag{9.8}$$

where we used the fact $\delta d_w > d_h \ge 1 \ge \theta > 0$. Similarly, by the semi-group property

$$\int_0^t \int_{\mathbb{X}} \left(p_s^\delta(x,z) - p_s^\delta(y,z)\right)^2 m(\mathrm{d}z)\mathrm{d}s = \int_0^t p_{2s}^\delta(x,x) - 2p_{2s}^\delta(x,y) + p_{2s}^\delta(y,y)\mathrm{d}s$$
$$\lesssim t^{1-\frac{\theta}{\delta d_w}} d(x,y)^\theta. \tag{9.9}$$

By [29, Proposition 2.3] for $t > 0$ and $x,\, y \in M$ that

$$\left|\partial_t p_t^\delta(x,y)\right| \lesssim t^{-\frac{d_h}{\delta d_w}-1}. \tag{9.10}$$

Hence for $\tau \ge t > 0$ and $x \in \mathbb{X}$, we have

$$\int_{\mathbb{X}} (p_t^\delta(x,y) - p_\tau^\delta(x,y))^2 \mu(dy) = p_{2t}^\delta(x,x) - 2p_{t+\tau}^\delta(x,x) + p_{2\tau}^\delta(x,x)$$
$$\le \int_{2t}^{2\tau} \left|\partial_s p_s(x,x)\right| \mathrm{d}s$$
$$\overset{(9.10)}{\lesssim} t^{-\frac{d_h}{\delta d_w}} - \tau^{-\frac{d_h}{\delta d_w}}$$
$$\lesssim t^{-1+\varepsilon_1} |\tau - t|^{\varepsilon_2} \tag{9.11}$$

for some small $\varepsilon_1, \varepsilon_2 > 0$, where we used Hölder's inequality in the last line.

$$\left|p_t^\delta(x,y) - p_\tau^\delta(x,y)\right| \lesssim \frac{|\tau - t|^{\frac{d_h}{\delta d_w}}}{t^{\frac{d_h}{\delta d_w}}} t^{-\frac{d_h}{\delta d_w}} \Phi_\delta\left(\frac{d(x,y)}{t^{\frac{1}{\delta d_w}}}\right). \tag{9.12}$$

Indeed, let $\tau \ge t > 0$, we may use (9.5), (9.6) and a change of variable to see

$$\left|p_t^\delta(x,y) - p_\tau^\delta(x,y)\right| = \left|\int_0^\infty t^{-\frac{1}{\delta}} \eta_1\left(\frac{s}{t^{\frac{1}{\delta}}}\right) p_s(x,y) - \tau^{-\frac{1}{\delta}} \eta_1\left(\frac{s}{\tau^{\frac{1}{\delta}}}\right) p_s(x,y) ds\right|$$
$$\le t^{-\frac{1}{\delta}} \int_0^\infty \eta_1^\delta\left(\frac{s}{t^{\frac{1}{\delta}}}\right) \left|p_s(x,y) - p_{\left(\frac{\tau}{t}\right)^{\frac{1}{\delta}} s}(x,y)\right| ds.$$

Using (4.10) and monotonicity of Φ to see that

$$\left|p_s(x,y) - p_{\left(\frac{\tau}{t}\right)^{\frac{1}{\delta}} s}(x,y)\right| \le \int_s^{\left(\frac{\tau}{t}\right)^{\frac{1}{\delta}} s} \tau^{-\frac{d_h}{d_w}} \Phi\left(\frac{d(x,y)}{\tau^{\frac{1}{d_w}}}\right) \frac{d\tau}{\tau}$$
$$\lesssim s^{-\frac{d_h}{d_w}} \left(1 - \left(\frac{\tau}{t}\right)^{-\frac{d_h}{\delta d_w}}\right) \Phi\left(\frac{d(x,y)}{t^{\frac{1}{d_w}}}\right).$$

Hence, by [26, Lemma 5.4],

$$\begin{aligned}\left|p_t^\delta(x,y)-p_\tau^\delta(x,y)\right| &\lesssim \left(1-\left(\frac{t}{\tau}\right)^{\frac{d_h}{\delta d_w}}\right)\int_0^\infty \eta_t^\delta(s)s^{-\frac{d_h}{d_w}}\Phi\left(\frac{d(x,y)}{s^{\frac{1}{d_w}}}\right)ds\\ &\lesssim \left(1-\left(\frac{t}{\tau}\right)^{\frac{d_h}{\delta d_w}}\right)t^{-\frac{d_h}{\delta d_w}}\Phi_\delta\left(\frac{d(x,y)}{t^{\frac{1}{\delta d_w}}}\right).\end{aligned}$$

Now use the fact that $\frac{d_h}{d_w}<\delta<1$ and Hölder's inequality to see (9.12) holds, which implies that

$$\int_{\mathbb{X}}|p_t^\delta(x,y)-p_\tau^\delta(x,y)|\mu(dy)\lesssim\frac{|\tau-t|^{\frac{d_h}{\delta d_w}}}{t^{\frac{d_h}{\delta d_w}}},$$

combined with (9.8) (9.9) and (9.11) to see (H_3) and (H_1) hold.

For (2.8), we observe first by triangular inequality that for all $x_0,x,y\in\mathbb{X}$,

$$1+d(x_0,y)\le 1+d(x_0,x)+d(x,y)\le(1+d(x_0,x))\,(1+d(x_0,y))$$

let $\lambda\in[0,1]$, then uniformly for $x\in M$,

$$\begin{aligned}\int_{\mathbb{X}}p_t^\delta(x,y)(1+d(x_0,y))^\lambda m(dy) &\lesssim (1+d(x_0,x))\,t^{-\frac{d_h}{\delta d_w}}\int_{\mathbb{X}}\Phi_\delta\left(\frac{d(x,y)}{t^{\frac{1}{\delta d_w}}}\right)(1+d(x,y))^\lambda m(dy)\\ &\lesssim (1+d(x_0,x))^\lambda.\end{aligned}$$

On the other hand, if we let $\alpha:=\frac{d_h}{\delta d_w}\in(0,1)$, then for all $\lambda\in[0,1/2]$, the following inequality holds uniformly in $x\in\mathbb{X}$,

$$\begin{aligned}\int_{\mathbb{X}}p_t^\delta(x,y)^2(1+d(x_0,y))^{2\lambda}m(\mathrm{d}y) &\lesssim t^{-\alpha}\int_{\mathbb{X}}p_t^\delta(x,y)(1+d(x_0,y))^{2\lambda}m(\mathrm{d}y)\\ &\lesssim t^{-\alpha}\left(1+d(x_0,x)\right)^{2\lambda}.\end{aligned}$$

Hence we see (2.8) holds for $\lambda\in[0,1]$ with $h(x):=(1+d(x_0,x))^{\frac{1}{2}}$.

For condition (W_1), we note by the semi-group property, it is enough to verify $\lim_{t\downarrow 0}\left\|P_t^\delta f-f\right\|_{(\lambda)}=0$ for each $f\in\mathcal{C}_{tem}$ and arbitrarily small $\lambda>0$, where $P_t^\delta f(x):=\int_M p_t^\delta(x,y)f(y)m(\mathrm{d}y)$ for $(t,x)\in\mathbb{R}_+\times M$ is the semi-group associated with p^δ. Denote by $h(x):=(1+d(x_0,x))^{\frac{1}{2}}$ for $x\in\mathbb{X}$, we have by triangular inequality and the fact that $f\in\mathcal{C}_{tem}$, for $\lambda\in(0,1)$

$$\begin{aligned}h^{-\lambda}(x)\left|P_t^\delta f(x)-f(x)\right| &= h^{-\lambda}(x)\left|\int_M p_t^\delta(x,y)\left(h^\lambda(y)\frac{f(y)}{h^\lambda(y)}-h(x)\frac{f(x)}{h^\lambda(x)}m(\mathrm{d}y)\right)\right|\\ &\lesssim_\lambda \int_M p_t(x,y)|h^\lambda(x)-h^\lambda(y)|m(\mathrm{d}y)+\int_M p_t(x,y)\left|\frac{f(y)}{h(y)}-\frac{f(x)}{h(x)}\right|m(\mathrm{d}y).\end{aligned}$$

We will call the two terms on the right hand side $I_1(t,x)$ and $I_2(t,x)$ respectively. Let $\varepsilon>0$, then triangular inequality and convexity asserts that for $\lambda\in(0,\delta d_w)$ and $t\in(0,1]$,

$$\begin{aligned}I_1(t,x) &\le \int_M p_t^\delta(x,y)d(x,y)^{\frac{\lambda}{2}}m(\mathrm{d}y)\\ &\lesssim \varepsilon^{\frac{\lambda}{2}}\int_{B_\varepsilon(x)}p_t^\delta(x,y)m(\mathrm{d}y)+\int_{B_\varepsilon(x)^c}t^{-\frac{d_h}{\delta d_w}}\left(1+\frac{d(x,y)}{t^{\frac{1}{\delta d_w}}}\right)^{-(d_h+\delta d_w)}d(x,y)^{\frac{\lambda}{2}}m(\mathrm{d}y)\\ &\le \varepsilon^{\frac{\lambda}{2}}+\left(1+\frac{\varepsilon}{t^{\frac{1}{\delta d_w}}}\right)^{\delta d_w/2}\int_M t^{-\frac{d_h}{\delta d_w}}\left(1+\frac{d(x,y)}{t^{\frac{1}{\delta d_w}}}\right)^{-(d_h+(\delta d_w-\lambda)/2)}m(\mathrm{d}y).\end{aligned}$$

By Lemma [27, Lemma 2.1], we see

$$\limsup_{t\downarrow 0}\sup_{x\in M} I_1(t,x)\lesssim \varepsilon^{\frac{\lambda}{2}}. \tag{9.13}$$

For I_2, we note the function $M\ni x\mapsto \frac{f(x)}{h(x)}$ is bounded continuous and vanish at infinity, and hence is uniformly continuous. Hence by a similar argument, we may choose $r>0$ small enough so that uniformly in $d(x,y)<r$, $\left|\frac{f(x)}{h(x)}-\frac{f(y)}{h(y)}\right|<\varepsilon^{\frac{\lambda}{2}}$

$$\limsup_{t\downarrow 0}\sup_{x\in M} I_2(t,x)\lesssim \varepsilon^{\frac{\lambda}{2}}+\limsup_{t\downarrow 0}\sup_{x\in M}\int_{B_r(x)^c} p_t(x,y)m(\mathrm{d}y)=\varepsilon^{\frac{\lambda}{2}}.$$

Combined with (9.13) to see that for all sufficiently small $\lambda>0$,

$$\limsup_{t\downarrow 0}\left\|P_t^\delta f-f\right\|_{(\lambda)}\lesssim \varepsilon^{\frac{\lambda}{2}}.$$

Since $\varepsilon>0$ is arbitrary, we see (W_1) holds. □

Proof of Lemma 8.5. For $0\le s<t\le T$ and $x\in\mathcal{X}$, denote $\mu_{t-s,x}(\mathrm{d}y,\mathrm{d}z):=\mu(t-s,x,\mathrm{d}y,\mathrm{d}z)$ and set $I_0(t):=1$ and $I_n(t):=\int_0^t(t-s)^{-\alpha}I_{n-1}(s)\mathrm{d}s$ for $n\in\mathbb{N}$ and $t\in[0,T]$. Consider for $t\in[0,T]$ and $x\in\mathcal{X}$,

$$\begin{aligned}
u_t(x) &\overset{(8.3)}{\le} h^{-1}(x)+\int_0^t\left\|h^{\frac12}\right\|^2_{\mu_{t-s,x}}\mathrm{d}s\\
&\overset{(8.2)}{\le} h^{-1}(x)+C\int_0^t(t-s)^{-\alpha}h(x)\mathrm{d}s\\
&\le h^{-1}(x)+CI_1(t)h(x).
\end{aligned}$$

Now we assume $u_t(x)\le h^{-1}(x)\sum_{k=0}^{n-1}C^kI_k(t)+C^nI_n(t)h(x)$ for some $n\ge 1$. Then by replacing $u_s(\cdot)$ for $s\in[0,T]$ by the above bound in (8.3), we see uniformly in $(t,x)\in[0,T]\times\mathcal{X}$ that

$$\begin{aligned}
u_t(x) &\le h^{-1}(x)+\int_0^t\left\|\left(h^{-1}\left(\sum_{k=0}^{n-1}I_k(s)C^k\right)+C^n\,I_n(s)h\right)^{\frac12}\right\|^2_{\mu_{t-s,x}}\mathrm{d}s\\
&\le h^{-1}(x)+\int_0^t\left\|h^{-\frac12}\left(\sum_{k=0}^{n-1}I_k(s)C^k\right)^{\frac12}+C^{n/2}\,(I_n(s))^{\frac12}\,h^{\frac12}\right\|^2_{\mu_{t-s,x}}\mathrm{d}s\\
&\overset{(8.1)}{\lesssim} h^{-1}(x)+\int_0^t\left(\sum_{k=0}^{n-1}I_k(s)C^k\right)\left\|h^{-\frac12}\right\|^2_{\mu_{t-s,x}}+C^nI_n(s)\left\|h^{\frac12}\right\|^2_{\mu_{t-s,x}}\mathrm{d}s\\
&\overset{(8.2)}{\le} h^{-1}(x)+h^{-1}(x)\sum_{k=1}^{n}I_{\widetilde{K}}(t)C^k+h(x)C^{n+1}\,I_{n+1}(t)\\
&= h^{-1}(x)\left(\sum_{k=0}^{n}I_k(t)\,C^k\right)+h(x)C^{n+1}I_{n+1}(t).
\end{aligned}$$

Therefore, we may take $n\to\infty$ and see that for each $t\in[0,T]$ and $x\in\mathcal{X}$,

$$u_t(x)\le h^{-1}(x)\left(\sum_{n=0}^{\infty}I_n(t)C^k\right)+h(x)\limsup_{n\to\infty}C^nI_n(t).$$

By [56, Lemma 3.3] we see for each $C>0$, $\sup_{t\in[0,T]}\sum_{n=0}^{\infty}I_n(t)C^n<\infty$ and $\sup_{t\in[0,T]}\lim_{n\to\infty}C^nI_n(t)=0$, which implies the desired inequality. □

We will now work towards the proof of Lemma 8.6 and we will break down its proof into several lemmas. We first observe the following simple fact:

Lemma 9.2. *Let $(\mathbb{X}, d)$ be a metric space with infinite diameter and h is a good weight on it. Then for each $x_0 \in \mathbb{X}$,*

$$\lim_{R\to\infty} \inf_{x\in B(x_0,R)^c} h(x) = \infty. \tag{9.14}$$

Proof. Suppose by contradiction that the left hand side of (9.14) is bounded by some $M \geq 1$ for some $x_0 \in \mathbb{X}$. Let $\lambda_0 > 0$ and $\{x_n\}_{n\in\mathbb{N}} \subset \mathbb{X}$ so that condition (2.5) holds. Then one can find a sequence of points $\{y_k\}_{k\in\mathbb{N}} \subset \mathbb{X}$ so that $y_k = x_{n_k}$ for some sub-sequence $\{n_k\}_{k\in\mathbb{N}}$ of $\mathbb{N}$. Then by (2.6), there is some $C \geq 1$ so that for each $\lambda \geq \lambda_0$, and $k \in \mathbb{N}$,

$$h(x_{n_k})^{-\lambda} \geq C^{-\lambda} \sup_{y\in B_1(x_{n_k})} h(y)^{-\lambda} \geq C^{-\lambda} h(y_k)^{-\lambda} \geq (CM)^{-\lambda},$$

which is a contradiction to (2.5). □

The following lemma is similar to [38, Lemma 1.9.3]. Note the topology in $\mathcal{C}_{tem}$ is given by the following metric: for $f, g \in \mathcal{C}_{tem}$, define $d_{tem}(f, g)$ as

$$d_{tem}(f,g) := \sum_{n\in\mathbb{N}} 2^{-n} \min\left\{ \|f-g\|_{(\frac{1}{n})}, 1 \right\}.$$

Lemma 9.3 (Arzela-Ascoli type lemma)**.** *Let $(\mathbb{X}, d)$ be a complete metric space so that each closed ball is compact, and let h be a good weight on $(\mathbb{X}, d)$. Suppose $\mathcal{K} \subset \mathcal{C}(\mathbb{R}^+; \mathcal{C}_{tem})$ is closed under d_{tem} and*

- *for each λ, $T > 0$, there is $C_{\lambda,T} > 0$ so that*

$$\sup_{f\in K} \sup_{t\in[0,T]} \|f\|_{(\lambda)} \leq C_{\lambda,T}. \tag{9.15}$$

- *the set of functions in $\mathcal{K}$ are equicontinuous at each point of $\mathbb{R}_+ \times \mathbb{X}$.*

Then $\mathcal{K}$ is a compact subset of $\mathcal{C}(\mathbb{R}^+; \mathcal{C}_{tem})$ in the compact open topology, i.e. for any sequence $\{f_n\}_{n\in\mathbb{N}} \subset \mathcal{K}$, there is a convergence subsequence in the given metric.

Proof. Let $\{f_n\}_n \subset \mathcal{K}$ be a sequence. By Arzela-Ascoli (c.f. [24, Theorem 4.44]), there is a subsequence $\{f_{n_r}\}_{r\in\mathbb{N}}$ that converges to some function $f \in \mathcal{C}([0,T]\times\mathbb{X})$ on each compact set of $[0,T]\times\mathbb{X}$ uniformly. Hence if $(\mathbb{X}, d)$ is compact, any weight is a good weight and $\mathcal{C}_{tem} = \mathcal{C}$ and we are done at this point. So we assume $(\mathbb{X}, d)$ has infinite diameter. It remains to show for each $T > 0$,

$$\lim_{r\to\infty} \sup_{t\in[0,T]} d_{tem}\left(f_{n_r}(t,\cdot), f(t,\cdot)\right) = 0.$$

Step 1: Consider for λ, T, $R > 0$ and $x_0 \in \mathbb{X}$ that

$$\begin{aligned}\sup_{t\leq T} \sup_{x\in B(x_0,R)} h(x)^{-\lambda} |f(t,x)| &\leq \sup_{t\in[0,T]} \sup_{x\in B(x_0,R)} |f(t,x) - f_{n_r}(t,x)|\, h(x)^{-\lambda} + \sup_{t\in[0,T]} \|f_{n_r}(t,\cdot)\|_{(\lambda)} \\ &\overset{(9.15)}{\leq} \sup_{t\in[0,T]} \sup_{x\in B(x_0,R)} |f(t,x) - f_{n_r}(t,x)|\, h(x)^{-\lambda} + C_{\lambda,T}\end{aligned}$$

for any $r \in \mathbb{N}$. We may take $r \uparrow \infty$ to see the first term vanishes, then we can take $R \uparrow \infty$ to conclude that $\sup_{0\leq t\leq T} \|f(t,\cdot)\|_{(\lambda)} < C_{\lambda,T}$.

Step 2: We let $N \in \mathbb{N}$ be large, we see by monotonicity of $\|\cdot\|_{(\lambda)}$ that

$$\begin{aligned}\sup_{0\leq t\leq T} d_{tem}(f_{n_r}(t,\cdot), f(t,\cdot)) &= \sup_{0\leq t\leq T} \sum_{n=1}^{\infty} 2^{-n} \left(\|f(t,\cdot) - f_{n_r}(t,\cdot)\|_{(1/n)} \wedge 1 \right) \\ &\leq \sup_{0\leq t\leq T} \sum_{n=1}^{N} 2^{-n} \left(\|f(t,\cdot) - f_{n_r}(t,\cdot)\|_{(1/n)} \right) + \sum_{n=N+1}^{\infty} 2^{-n}\end{aligned}$$

$$\leq \sup_{0\leq t\leq T} \|f(t,\cdot) - f_{r_n}(t,\cdot)\|_{(1/N)} + 2^{-N}.$$

Let $t \in [0,T]$ and consider

$$\begin{aligned}
\|f(t,\cdot) - f_{r_n}(t,\cdot)\|_{(1/N)} &= \left\|f(t,\cdot) - 1_{B(x_0,R)}(\cdot)f(t,\cdot) + 1_{B(x_0,R)}(\cdot)f(t,\cdot) - f_{n_r}(t,\cdot)\right\|_{(1/N)} \\
&\leq \sup_{x\in B(x_0,R)^c} f(t,x)h(x)^{-\frac{1}{N}} + \sup_{x\in B(x_0,R)} |f(t,x) - f_{n_r}(t,x)|\, h(x)^{-\frac{1}{N}} \\
&\qquad + \sup_{x\in B(x_0,R)^c} |f_{n_r}(t,x)| h(x)^{-\frac{1}{N}} \\
&\leq \|f(t,\cdot)\|_{(\frac{1}{2N})} \sup_{x\in B(x_0,R)^c} h(x)^{-\frac{1}{2N}} + \sup_{x\in B(x_0,R)} |f(t,x) - f_{n_r}(t,x)|\, h(x)^{-\frac{1}{N}} \\
&\qquad + \|f_{n_r}(t,\cdot)\|_{(\frac{1}{2N})} \sup_{x\in B(x_0,R)^c} h(x)^{-\frac{1}{2N}} \\
&\overset{(9.15)}{\leq} 2C_{T,\frac{1}{2N}} \sup_{x\in B(x_0,R)^c} h(x)^{-\frac{1}{2N}} + \sup_{x\in B(x_0,R)} |f(t,x) - f_{n_r}(t,x)|\,.
\end{aligned}$$

Hence,

$$\sup_{0\leq t\leq T} d_{tem}(f_{n_r}, f) \leq 2C_{T,\frac{1}{2N}} \sup_{x\in B(x_0,R)^c} h(x)^{-\frac{1}{2N}} + \sup_{0\leq t\leq T} \sup_{x\in B(x_0,R)} |f(t,x) - f_{n_r}(t,x)| + 2^{-N}.$$

We first take $n_r \to \infty$ on both sides to see

$$\limsup_{r\to\infty} \sup_{0\leq t\leq T} d_{tem}(f_{n_r}(t,\cdot), f(t,\cdot)) \leq 2C_{T,\frac{1}{2N}} \sup_{x\in B(x_0,R)^c} h(x)^{-\frac{1}{2N}} + 2^{-N}.$$

Then take $R \to \infty$ so that the first term on the right hand side disappears due to Lemma 9.2. Finally, take $N \to \infty$ to see f_{n_r} converges to f in the corresponding norm in $\mathcal{C}([0,T];\mathcal{C}_{tem})$. $\square$

Lemma 9.4. *Suppose assumptions in Lemma 9.3 hold, then a sequence of probability measure, $\{\mathbb{P}_n\}_{n\in\mathbb{N}}$, on $\mathcal{C}(\mathbb{R}_+;\mathcal{C}_{tem})$ is tight if there is $x_0 \in \mathbb{X}$ so that for every λ, T, ε, $R > 0$, the following conditions holds:*

$$\lim_{N\uparrow\infty} \sup_{n\in\mathbb{N}} \mathbb{P}_n\left(\sup_{t\in[0,T]} \sup_{x\in\mathbb{X}} |X(t,x)|\, h(x)^{-\lambda} > N\right) = 0 \tag{9.16}$$

$$\lim_{\delta\downarrow 0} \sup_{n\in\mathbb{N}} \mathbb{P}_n\left(\sup_{t,t'\in[0,T]:\, |t'-t|\leq\delta}\ \sup_{x,x'\in B(x_0,R):\, d(x,x')<\delta} |X(t,x) - X(t',x')|\, h(x)^{-\lambda} > \varepsilon\right) = 0. \tag{9.17}$$

Proof. Let $a > 0$ be small, then for any T, R, k, $j \in \mathbb{N}$, we may choose $N^a_{j,k,T,R} \in \mathbb{N}$ large enough and $\delta^a_{j,k,T,R} > 0$ small enough so that

$$\sup_{n\in\mathbb{N}} \mathbb{P}_n\left(\sup_{t\in[0,T]} \sup_{x\in\mathbb{X}} |X(t,x)|\, h(x)^{-\frac{1}{k}} > N^a_{j,k,T,R}\right) \leq \frac{a}{2^{k+T+R+j}}$$

$$\sup_{n\in\mathbb{N}} \mathbb{P}_n\left(\sup_{t,t'\in[0,T]:|t'-t|\leq\delta^a_{j,k,T,R}}\ \sup_{x,x'\in B(x_0,R):d(x,x')<\delta^a_{j,k,T,R}} |X(t,x) - X(t',x')|\, h(x)^{-\frac{1}{k}} > \frac{1}{j}\right) < \frac{a}{2^{k+T+R+j}}.$$

We define $A^a_{j,k,T,R} \subset \mathcal{C}(\mathbb{R}_+;\mathcal{C}_{tem})$ by

$$\begin{aligned}
A^a_{j,k,T,R} := \Bigg\{ f \in&\mathcal{C}\left(\mathbb{R}_+;\mathcal{C}_{tem}\right) : \sup_{t\in[0,T]} \|f(t,\cdot)\|_{(1/k)}\, h(x)^{-\frac{1}{k}} \leq N^a_{j,k,T,R}, \\
&\sup_{t,t'\in[0,T]:\, |t'-t|\leq\delta^a_{j,k,T,R}}\ \sup_{x,x'\in B(x_0,R):\, d(x,x')<\delta^a_{j,k,T,R}} |f(t,x) - f(t',x')|\, h(x)^{-\frac{1}{k}} \leq \frac{1}{j} \Bigg\},
\end{aligned}$$

and set

$$A_a := \bigcap_{j,k,T,R\in\mathbb{N}} A^a_{j,k,T,R},$$

which is compact subset of $\mathcal{C}(\mathbb{R}_+;\mathcal{C}_{tem})$ for each $a>0$ by Lemma 9.3, and

$$\mathbb{P}_n\left((A_a)^c\right) \le 2a\sum_{j\in\mathbb{N}}\sum_{k\in\mathbb{N}}\sum_{R\in\mathbb{N}}\sum_{T\in\mathbb{N}} 2^{-(k+T+R+j)} = 2a.$$

Since we may choose a to be arbitrarily small, this shows tightness. □

Proof of Lemma 8.6. We will prove tightness by verifying the conditions in Lemma 9.4. Let $p, T \ge 1$ and $m \le T$ be an integer. Then for all $\lambda > 0$, there is a $C_{T,p,\lambda} > 0$ so that for each $x \in \mathbb{X}$,

$$\begin{aligned}\mathbb{E}\left[|X_n(m,x)|^p\right] &\le \mathbb{E}\left[|X_n(0,x)|^p\right] + \sum_{k=1}^m \mathbb{E}\left[|X_n(k-1,x) - X_n(k,x)|^p\right]\\ &\le C_{T,p,\lambda} h(x)^\lambda.\end{aligned}$$

Let $\mathcal{S} := \mathbb{R}_+ \times \mathbb{X}$ and $d_{\mathcal{S}}$ be the metric on $\mathcal{S}$ defined via

$$d_{\mathcal{S}}((t,x),(t',y)) := |t-t'|^{\xi_1} + d(x,y)^{\xi_2}.$$

One can verify that $(\mathcal{S}, d_{\mathcal{S}})$ is a complete metric space and if we let $B_{\mathcal{S}}(s,r)$ be the open ball in $\mathcal{S}$ with radius r centered at $s \in \mathcal{S}$ under metric $d_{\mathcal{S}}$, then we see

$$D(B_{\mathcal{S}}(s,1), d_{\mathcal{S}}; \varepsilon) \lesssim \varepsilon^{\frac{1}{\xi_1}+\frac{\beta}{\xi_2}}, \qquad \text{uniformly in } s\in\mathcal{S},\ \varepsilon\in(0,1].$$

Now by [46, Theorem 11.1 and p302], for sufficiently large $p \ge 1$, we have uniformly in $x \in \mathbb{X}$ and $n \in \mathbb{N}$ that

$$\mathbb{E}\left[\sup_{t,t'\in[0,T],\,|t-t'|\le 1}\ \sup_{y,y'\in B(x,1)} |X_n(t,y) - X_n(t',y')|^p\right] \lesssim h(x)^\lambda. \tag{9.18}$$

Therefore, we see uniformly in $x \in \mathbb{X}$ and for integers $m \le T$ and $n \in \mathbb{N}$ that

$$\begin{aligned}\mathbb{E}\left[\sup_{y\in B(x,1),\, s\in[m,m+1]} |X_n(s,y)|^p\right] &\le \mathbb{E}\left[|X_n(m,x)|^p\right] + \mathbb{E}\left[\sup_{y\in B(x,1),\, s\in[m,m+1]} |X_n(s,y) - X_n(m,x)|^p\right]\\ &\lesssim h(x)^\lambda. \end{aligned}\tag{9.19}$$

Let $\{x_i\}_{i\in\mathbb{N}} \subset \mathbb{X}$ be a sequence so that (2.5) holds. By Markov's inequality we see that uniformly in $A > 0$,

$$\begin{aligned}\mathbb{P}\left(\sup_{s\le T}\|X_n(s)\|_{(\lambda)} \ge A\right) &\le \sum_{0\le m\le T}\sum_{i=1}^\infty \mathbb{P}\left(\sup_{t\in[m,m+1]}\ \sup_{x\in B(x_i,1)} |X_n(s,x)|^p\, h(x)^{-\lambda p} \ge A^p\right)\\ &\le \sum_{0\le m\le T}\sum_{i=1}^\infty \mathbb{P}\left(\sup_{t\in[m,m+1]}\ \sup_{x\in B(x_i,1)} |X_n(s,x)|^p \ge A^p \inf_{x\in B(x_i,1)} h(x)^{\lambda p}\right)\\ &\overset{(9.19)}{\lesssim} \sum_{i=1}^\infty \frac{1}{A^p} h(x_i)^\lambda \sup_{x\in B(x_i,1)} h(x)^{-\lambda p}\\ &\overset{(2.6)}{\lesssim} \frac{1}{A^p}\sum_{i=1}^\infty h(x_i)^{-\lambda p+\lambda}.\end{aligned}$$

By (2.5), we may choose p to be large enough so that the sum is finite. This verifies condition (9.16) in Lemma 9.4.

Let $x_0 \in \mathbb{X}$ and $N > 0$ and let $\{B(x_i, 1/2)\}_{i\in I}$ be a finite open cover of $B(x_0, N)$ for some finite index set I. Observe if $x, x' \in B(x_0, N)$ with $d(x,x') < \frac{1}{2}$, then $x, x' \in B(x_i, 1)$ for some $i \in I$. We again choose

$\alpha \in (\frac{1}{\xi_1}+\frac{\beta}{\xi_2}, p/2)$ as above and use Markov inequality to obtain that uniformly for $\delta \in (0,1/2)$, $\varepsilon \in (0,1)$ and $n \in \mathbb{N}$,

$$
\begin{aligned}
&\mathbb{P}\left(\sup_{t,t'\in[0,T]:\,|t'-t|\le\delta}\ \sup_{x,x'\in B(x_0,N):\,d(x,x')<\delta} |X_n(t,x)-X_n(t',x')|\,h(x)^{-\lambda} > \varepsilon\right)\\
\le&\mathbb{P}\left(\sup_{t,t'\in[0,T]:\,|t'-t|\le\delta}\ \sup_{x,x'\in B(x_0,N):\,d(x,x')<\delta} \frac{|X_n(t,x)-X_n(t',x')|}{\left(|t-t'|^{\xi_1}+d(x,x')^{\xi_2}\right)^{1-\frac{2\alpha}{p}}}h^{-\lambda}(x) > \frac{\varepsilon}{\left(\delta^{\xi_1}+\delta^{\xi_2}\right)^{1-\frac{2\alpha}{p}}}\right)\\
\le&\sum_{0\le m\le T}\sum_{i\in I}\mathbb{P}\left(\sup_{t,t'\in[m,m+1]}\ \sup_{x,x'\in B(x_i,1)} \frac{|X_n(t,x)-X_n(t',x')|}{\left(|t-t'|^{\xi_1}+d(x,x')^{\xi_2}\right)^{1-\frac{2\alpha}{p}}}h^{-\lambda}(x) > \frac{\varepsilon}{\left(\delta^{\xi_1}+\delta^{\xi_2}\right)^{1-\frac{2\alpha}{p}}}\right)\\
\overset{(9.18)}{\lesssim}&\sum_{1\le m\le T}\sum_{i\in I}\frac{\left(\delta^{\xi_1}+\delta^{\xi_2}\right)^{p-2\alpha}}{\varepsilon^p},
\end{aligned}
$$

Since I is finite and $p>2\alpha$, we see (9.17) of Lemma 9.4 holds for each $\lambda, T, N, \varepsilon \ge 1$. Therefore, $\{X_n\}_{n\ge1}$ is tight in $\mathcal{C}(\mathbb{R}_+;\mathcal{C}_{tem})$. ☐

Proof of Lemma 8.8. The statement is automatically true if $(\mathbb{X},d)$ is compact, hence we assume diameter of $(\mathbb{X},d)$ is infinite. First, let us fix an arbitrary reference point $x_0\in\mathbb{X}$ and define for each $n\in\mathbb{N}$ the function $\mathcal{X}_n$ on $\mathbb{X}$, so that $\mathcal{X}_n\in\mathcal{C}_c(\mathbb{X})$ and

$$
\mathcal{X}_n(x)=\begin{cases}1 & \text{if } x\in B(x_0,n)\\ 0 & \text{if } x\in B(x_0,n+1)^c.\end{cases}
$$

Now let $f\in\mathcal{C}_{tem}$ and assume without loss of generality that $f(x)\ge0$ for all $x\in\mathbb{X}$. Define $f_n(x):=\mathcal{X}_n(x)f(x)$ for $x\in\mathbb{X}$, clearly $f_n\in\mathcal{C}_c(\mathbb{X})$. Let $\lambda>0$ and consider

$$
\|f_n-f\|_{(\lambda)}=\sup_{x\in\mathbb{X}}h(x)^{-\lambda}|f(x)-f_n(x)|=\sup_{x\in\mathbb{X}}\left|f(x)h(x)^{-\lambda}-f_n(x)h^{-\lambda}(x)\right|.
$$

Observe that the function $(f(x)h(x)^{-\lambda},\ x\in\mathbb{X})$ is an element of $C_0(\mathbb{X})$ and so is the sequence of functions $f_n(x)h^{-\lambda}(x)=f(x)h(x)^{-\lambda}\mathcal{X}_n(x)$, which clearly converges in supreme norm to $f(x)h(x)^{-\lambda}$. ☐

Proof of Lemma 8.13. (1) Let $t>0$, $\varepsilon\in(0,1]$ and $x\in\mathbb{X}$, consider by Jessen's inequality that

$$
\begin{aligned}
\|R_\varepsilon(t,x,\cdot)\|^2_{\widetilde{K}} &\overset{(8.1)}{\le}\left(e^{-t/\varepsilon}\sum_{n=1}^\infty\frac{(t/\varepsilon)^n}{n!}\|p(n\varepsilon,x,\cdot)\|_{\widetilde{K}}\right)^2\\
&\le e^{-t/\varepsilon}\sum_{n=1}^\infty\frac{(t/\varepsilon)^n}{n!}\|p(n\varepsilon,x,\cdot)\|^2_{\widetilde{K}}\\
&\overset{(2.8)}{\le} e^{-t/\varepsilon}\sum_{n=1}^\infty\frac{(t/\varepsilon)^n}{n!}(n\varepsilon)^{-\alpha}+1.
\end{aligned}
$$

Using Hölder's inequality with exponent $\frac{1}{\alpha}$ to see that

$$
e^{-t/\varepsilon}\sum_{n=1}^\infty\frac{(t/\varepsilon)^n}{n!}(n\varepsilon)^{-\alpha}\le 2\left(e^{-t/\varepsilon}\sum_{n=1}^\infty\frac{(t/\varepsilon)^{n+1}}{(n+1)!}t^{-1}\right)^{-\alpha}\le 2t^{-\alpha}.
$$

This implies $\|R_\varepsilon(t,x,\cdot)\|^2_{\widetilde{K}}\lesssim t^{-\alpha}\vee 1$ uniformly in $(t,x)\in\mathbb{R}_+\times\mathbb{X}$, which is the desired inequality.

(2) The objective is to show

$$
\lim_{\varepsilon\downarrow0}\sup_{x\in\mathbb{X}}\int_0^t\left\|\int_{\mathbb{X}}R_\varepsilon(t-s,x,y)p_\varepsilon(y,\cdot)m(\mathrm{d}y)-p(t-s,x,\cdot)\right\|^2_{\widetilde{K}}\mathrm{d}s\to0. \tag{9.20}
$$

Consider for $t>0$, we have uniformly in $x\in\mathbb{X}$ that

$$\begin{aligned}
&\int_0^t \left\|\int_{\mathbb{X}} R_\varepsilon(t-s,x,y)p_\varepsilon(y,\cdot)m(\mathrm{d}y)-p(t-s,x,\cdot)\right\|_{\widetilde{K}}^2 \mathrm{d}s\\
&\overset{(2.8)}{\lesssim} \int_0^t e^{-2(t-s)/\varepsilon}\left\|\sum_{n=1}^{\infty}\frac{(t/\varepsilon)^n}{n!}\left(p((n+1)\varepsilon,x,\cdot)-p(t-s,x,\cdot)\right)\right\|_{\widetilde{K}}^2 \mathrm{d}s\\
&\quad+\int_0^t e^{-(t-s)/\varepsilon}(t-s)^{-\alpha}\mathrm{d}s\\
&\le \int_0^t e^{-(t-s)/\varepsilon}\sum_{n=1}^{\infty}\frac{\left(\frac{t-s}{\varepsilon}\right)^n}{n!}\|p(\varepsilon(n+1),x,\cdot)-p(t-s,x,\cdot)\|_{\widetilde{K}}^2\,\mathrm{d}s+\int_0^t e^{-s/\varepsilon}s^{-\alpha}\mathrm{d}s,
\end{aligned}$$

where we used Jessen's inequality and Lemma 8.1 in the last line.

Now let Π be a Poisson point process on $\mathbb{R}_+$ with Lebesgue's' measures as its intensity measure. Let $\Pi(a):=\Pi([0,a])$ for $a\ge 0$, then the integrand in the first term in the last inequality above is bounded above by

$$\mathbb{E}\left[\left\|p\left(\varepsilon\Pi\left(\frac{t-s}{\varepsilon}\right)+\varepsilon,x,\cdot\right)-p(t-s,x,\cdot)\right\|_{\widetilde{K}}^2\right].$$

Let $A_{\varepsilon,t-s,T}:=\left\{\varepsilon\Pi\left(\frac{t-s}{\varepsilon}\right)<T\right\}$, by ($\mathrm{H}_3$), we see that uniformly in $0\le s<t\le T$ and $\varepsilon\in(0,1)$

$$\begin{aligned}
&\sup_x \mathbb{E}\left[\left\|p\left(\varepsilon\Pi\left(\frac{t-s}{\varepsilon}\right)+\varepsilon,x,\cdot\right)-p(t-s,x,\cdot)\right\|_{\widetilde{K}}^2 \mathbf{1}_{(A_{\varepsilon,t-s,T})^c}\right]\\
&\lesssim e^{-(t-s)/\varepsilon}\sum_{\varepsilon n\ge T}\frac{\left(\frac{t-s}{\varepsilon}\right)^n}{n!}\left(\max\left\{(\varepsilon(n+1))^{-\alpha},1\right\}+(t-s)^{-\alpha}\right) \qquad (9.21)
\end{aligned}$$

which converges to zero as $\varepsilon\downarrow 0$ for all $T\ge t>s>0$ and is dominated by $(t-s)^{-\alpha}\vee 1$ up to a multiplicative constant. On the other hand, uniformly in $\varepsilon\in(0,1)$, $0\le s<t\le T$ and $x\in\mathbb{X}$,

$$\begin{aligned}
&\mathbb{E}\left[\left\|p\left(\varepsilon\Pi\left(\frac{t-s}{\varepsilon}\right)+\varepsilon,x,\cdot\right)-p(t-s,x,\cdot)\right\|_{\widetilde{K}}^2 \mathbf{1}_{A_{\varepsilon,t-s,T}}\right]\\
&\overset{(H_3)}{\lesssim} \mathbb{E}\left[\frac{\gamma\left(\left|\varepsilon\Pi\left(\frac{t-s}{\varepsilon}\right)+\varepsilon-(t-s)\right|\right)}{\left(\varepsilon\left(1+\Pi\left(\frac{t-s}{\varepsilon}\right)\right)\wedge(t-s)\right)^{\beta}}\mathbf{1}_{A_{\varepsilon,t-s,T}}\right]\\
&\lesssim\left(\mathbb{E}\left[\frac{\gamma\left(\left|\varepsilon\Pi\left(\frac{t-s}{\varepsilon}\right)+\varepsilon-(t-s)\right|\right)}{\varepsilon^{\beta}\left(1+\Pi\left(\frac{t-s}{\varepsilon}\right)\right)^{\beta}}\mathbf{1}_{A_{\varepsilon,t-s,T}}\right]+\mathbb{E}\left[\frac{\gamma\left(\left|\varepsilon\Pi\left(\frac{t-s}{\varepsilon}\right)+\varepsilon-(t-s)\right|\right)}{(t-s)^{\beta}}\mathbf{1}_{A_{\varepsilon,t-s,T}}\right]\right)
\end{aligned}$$

The second term converges to zero for all $0\le s<t$ and is dominated by $(t-s)^{-\beta}$. For the first term, apply Hölder inequality for large enough $p\ge 1$ so that $\beta\frac{p}{p-1}<1$ to get

$$\begin{aligned}
\mathbb{E}\left[\frac{\gamma\left(\left|\varepsilon\Pi\left(\frac{t-s}{\varepsilon}\right)+\varepsilon-(t-s)\right|\right)}{\left(\varepsilon+\varepsilon\Pi\left(\frac{t-s}{\varepsilon}\right)\right)^{\beta}}\mathbf{1}_{A_{\varepsilon,t-s,T}}\right]&\le\left\|\gamma\left(\left|\varepsilon\Pi\left(\frac{t-s}{\varepsilon}\right)+\varepsilon-(t-s)\right|\right)\mathbf{1}_{A_{\varepsilon,t-s,T}}\right\|_{L^p(\mathbb{P})} &(9.22)\\
&\quad\times\mathbb{E}\left[\left(\varepsilon+\varepsilon\Pi\left(\frac{t-s}{\varepsilon}\right)\right)^{-\frac{\beta p}{p-1}}\right]^{\frac{p-1}{p}}. &(9.23)
\end{aligned}$$

The right hand side term in (9.22) is bounded above by $\gamma(2T)$ and converges to zero as $\varepsilon \downarrow 0$ for all $t, s \leq T$. By Jessen's inequality we see the term in (9.23) is bounded above by

$$\begin{aligned}
\mathbb{E}\left[\left(\varepsilon + \varepsilon\Pi\left(\frac{t-s}{\varepsilon}\right)\right)^{-1}\right]^{-\beta} &\leq \left(e^{-(t-s)/\varepsilon}\sum_{n=0}^{\infty}\frac{\left(\frac{t-s}{\varepsilon}\right)^n}{n!}\frac{1}{\varepsilon(n+1)}\right)^{-\beta} \\
&= \left(e^{-(t-s)/\varepsilon}\sum_{n=0}^{\infty}\frac{1}{t-s}\frac{\left(\frac{t-s}{\varepsilon}\right)^{n+1}}{(n+1)!}\right)^{-\beta} \\
&= \left(\frac{1}{t-s}e^{-(t-s)/\varepsilon}\sum_{n=1}^{\infty}\frac{\left(\frac{t-s}{\varepsilon}\right)^n}{n!}\right)^{-\beta} \\
&\leq (t-s)^{-\beta} \qquad (9.24)
\end{aligned}$$

Combining (9.21) - (9.24), by the dominated convergence theorem, we see (9.20) holds.

For (3), we see for $T > 0$ that

$$\int_{\mathbb{X}} |R_\varepsilon(t,x,y) - p_t(x,y)|\,\mathrm{d}y \leq e^{-\frac{t}{\varepsilon}}\left(1 + \sum_{n\varepsilon \leq T+1}\frac{(t/\varepsilon)^n}{n!}\int_{\mathbb{X}}|p_{n\varepsilon}(x,y) - p_t(x,y)|\,\mathrm{d}y + 2\sum_{n\varepsilon > T+1}\frac{(t/\varepsilon)^n}{n!}\right).$$

Let $\psi(\lambda) = (\lambda+1)\ln(\lambda+1) - \lambda$ and $c_T = \min\left\{\psi\left(\frac{T+1}{T} - 1\right), 1\right\}$, by Bennett's inequality for Poisson random variables, we see for all $t \in (0, T]$,

$$e^{-t/\varepsilon}\sum_{n\varepsilon > T+1}\frac{(t/\varepsilon)^n}{n!} \leq \exp\left(-\frac{t}{\varepsilon}\psi\left(\frac{T+1}{t} - 1\right)\right) \leq \exp\left(-c_T\frac{t}{\varepsilon}\right).$$

Let $p \geq 1$ and $q := \frac{p}{p-1}$, by (H_3), we may choose p large so that $q\beta < 1$, and obtain by Hölder's inequality that

$$\begin{aligned}
e^{-t/\varepsilon}\sum_{n\varepsilon \leq T+1}\frac{(t/\varepsilon)^n}{n!}\int_{\mathbb{X}}|p_{n\varepsilon}(x,y) - p_t(x,y)|\,\mathrm{d}y &\lesssim_T e^{-t/\varepsilon}\sum_{n\varepsilon \leq T+1}\frac{(t/\varepsilon)^n}{n!}\left(\frac{1}{(n\varepsilon)^\beta} + \frac{1}{t^\beta}\right)\gamma\left(|n\varepsilon - t|\right) \\
&\leq \gamma_\varepsilon\left(t^{-\beta} + I(t)\right)
\end{aligned}$$

where $\gamma_\varepsilon = \left(e^{-t/\varepsilon}\sum_{n\varepsilon \leq T+1}\frac{(t/\varepsilon)^n}{n!}\gamma(|n\varepsilon - t|)\right)^{\frac{1}{p}}$ and

$$\begin{aligned}
I(t) &= \left(e^{-t/\varepsilon}\sum_{n\varepsilon \leq T+1}\frac{(t/\varepsilon)^n}{n!}\frac{1}{(n\varepsilon)^{q\beta}}\right)^{\frac{1}{q}} \\
&\leq \left(e^{-t/\varepsilon}\sum_{n=1}^{\infty}\frac{(t/\varepsilon)^n}{n!}\frac{1}{n\varepsilon}\right)^{\beta} \\
&\leq \frac{2}{t^\beta}\left(e^{-t/\varepsilon}\sum_{n=0}^{\infty}\frac{(t/\varepsilon)^n}{n!}\right)^{\beta} \\
&= \frac{2}{t^\beta},
\end{aligned}$$

where we used Jensen's inequality in the second line. Combining all above, we see uniformly in $(t, x) \in [0, T] \times \mathbb{X}$,

$$\int_{\mathbb{X}} |R_\varepsilon(t,x,y) - p_t(x,y)|\,\mathrm{d}y \lesssim e^{-c_T\frac{t}{\varepsilon}} + \gamma_\varepsilon t^{-\beta}.$$

Finally, by the dominated convergence theorem, we see $\lim_{\varepsilon \downarrow 0}\gamma_\varepsilon = 0$, which proves the statement. $\square$

References

[1] Sebastian Andres and Naotaka Kajino. Continuity and estimates of the liouville heat kernel with applications to spectral dimensions. *Probability Theory and Related Fields*, 166:713–752, 2016.

[2] Ismaël Bailleul and Frederic Bernicot. Heat semigroup and singular pdes. *Journal of Functional Analysis*, 270(9):3344–3452, 2016.

[3] Martin T Barlow. Diffusions on fractals. *Lectures on Probability Theory and Statistics: Ecole d'Eté de Probabilités de Saint-Flour XXV—1995*, pages 1–121, 2006.

[4] Martin T Barlow, Alexander Grigor'yan, and Takashi Kumagai. On the equivalence of parabolic harnack inequalities and heat kernel estimates. *Journal of the Mathematical Society of Japan*, 64(4):1091–1146, 2012.

[5] Martin T Barlow and Edwin A Perkins. Brownian motion on the sierpinski gasket. *Probability theory and related fields*, 79(4):543–623, 1988.

[6] Clayton Barnes, Leonid Mytnik, and Zhenyao Sun. Effect of small noise on the speed of reaction-diffusion equations with non-lipschitz drift. *arXiv preprint arXiv:2107.09377*, 2021.

[7] Richard F Bass and Pei Hsu. Some potential theory for reflecting brownian motion in holder and lipschitz domains. *The Annals of Probability*, pages 486–508, 1991.

[8] Fabrice Baudoin, Li Chen, Che-Hung Huang, Cheng Ouyang, Samy Tindel, and Jing Wang. Parabolic anderson model in bounded domains of recurrent metric measure spaces. *arXiv preprint arXiv:2401.01797*, 2024.

[9] Nathanaël Berestycki and Ewain Gwynne. Random walks on mated-crt planar maps and liouville brownian motion. *Communications in Mathematical Physics*, 395(2):773–857, 2022.

[10] D Burago. A course in metric geometry. *American Mathematical Society*, 2001.

[11] Sandra Cerrai and Mark Freidlin. Spdes on narrow domains and on graphs: an asymptotic approach. 2017.

[12] Sandra Cerrai and Mark Freidlin. Fast flow asymptotics for stochastic incompressible viscous fluids in r 2 and spdes on graphs. *Probability Theory and Related Fields*, 173(1):491–535, 2019.

[13] Sandra Cerrai and Wen-Tai Hsu. Spdes on narrow channels and graphs: convergence and large deviations in case of non smooth noise. *arXiv preprint arXiv:2403.13493*, 2024.

[14] Sandra Cerrai and Guangyu Xi. Incompressible viscous fluids in r 2 and spdes on graphs, in presence of fast advection and non smooth noise. In *Annales de l'Institut Henri Poincare (B) Probabilites et statistiques*, volume 57, pages 1636–1664. Institut Henri Poincaré, 2021.

[15] Le Chen and Jingyu Huang. Comparison principle for stochastic heat equation on $\diagdown^d$. *The Annals of Probability*, 47(2):989–1035, 2019.

[16] Le Chen and Kunwoo Kim. On comparison principle and strict positivity of solutions to the nonlinear stochastic fractional heat equations. *Annales de l'Institut Henri Poincaré-Probabilités et Statistiques*, 53(1):358–388, 2017.

[17] Zhen-Qing Chen, Panki Kim, and Renming Song. Heat kernel estimates for the dirichlet fractional laplacian. *Journal of the European Mathematical Society*, 12(5):1307–1329, 2010.

[18] Zhen-Qing Chen and Takashi Kumagai. Heat kernel estimates for stable-like processes on d-sets. *Stochastic Processes and their applications*, 108(1):27–62, 2003.

[19] Zhen-Qing Chen and Takashi Kumagai. Heat kernel estimates for jump processes of mixed types on metric measure spaces. *Probability theory and related fields*, 140(1):277–317, 2008.

[20] David A Croydon. Volume growth and heat kernel estimates for the continuum random tree. *Probability Theory and Related Fields*, 140(1):207–238, 2008.

[21] Giuseppe Da Prato and Jerzy Zabczyk. *Stochastic equations in infinite dimensions*, volume 152. Cambridge university press, 2014.

[22] Edward B Davies. Non-gaussian aspects of heat kernel behaviour. *Journal of the London Mathematical Society*, 55(1):105–125, 1997.

[23] Wai-Tong Louis Fan. Stochastic pdes on graphs as scaling limits of discrete interacting systems. *Bernoulli*, 27(3):1899–1941, 2021.

[24] Gerald B Folland. *Real analysis: modern techniques and their applications*, volume 40. John Wiley & Sons, 1999.

[25] Masatoshi Fukushima, Yoichi Oshima, and Masayoshi Takeda. *Dirichlet forms and symmetric Markov processes*, volume 19. Walter de Gruyter, 2011.

[26] Alexander Grigor'yan. Heat kernels and function theory on metric measure spaces. *Contemporary Mathematics*, 338:143–172, 2003.

[27] Alexander Grigor'yan and Liguang Liu. Heat kernel and lipschitz–besov spaces. In *Forum Mathematicum*, volume 27, pages 3567–3613. De Gruyter, 2015.

[28] Alexander Grigor'yan and Andras Telcs. Two-sided estimates of heat kernels on metric measure spaces. *The Annals of Probability*, pages 1212–1284, 2012.

[29] Alexander Grigor'yan, Jiaxin Hu, and Ka-Sing Lau. Heat kernels on metric measure spaces. In *Geometry and Analysis of Fractals: Hong Kong, December 2012*, pages 147–207. Springer, 2014.

[30] Sebastian Haeseler. Heat kernel estimates and related inequalities on metric graphs. *arXiv preprint arXiv:1101.3010*, 2011.

[31] Martin Hairer and Harprit Singh. Regularity structures on manifolds and vector bundles. *arXiv preprint arXiv:2308.05049*, 2023.
[32] Ben Hambly and Weiye Yang. Continuous random field solutions to parabolic spdes on pcf fractals. *arXiv preprint arXiv:1709.00916*, 2017.
[33] Ben Hambly and Weiye Yang. Existence and space-time regularity for stochastic heat equations on pcf fractals. *Electron. J. Probab*, 23(22):1–30, 2018.
[34] Ben Hambly and Weiye Yang. The damped stochastic wave equation on post-critically finite fractals. In *Analysis, Probability and Mathematical Physics on Fractals*, pages 521–556. World Scientific, 2020.
[35] Michael Hinz and Melissa Meinert. Approximation of partial differential equations on compact resistance spaces. *Calculus of Variations and Partial Differential Equations*, 61(1):19, 2022.
[36] Michael Hinz, Michael Röckner, and Alexander Teplyaev. Vector analysis for dirichlet forms and quasilinear pde and spde on metric measure spaces. *Stochastic Processes and their Applications*, 123(12):4373–4406, 2013.
[37] Michael Hinz and Martina Zähle. Semigroups, potential spaces and applications to (s) pde. *Potential Analysis*, 36:483–515, 2012.
[38] Paul Robert Horridge. *Methods of proving uniqueness of stationary distributions for stochastic PDEs*. PhD thesis, University of Warwick, 2001.
[39] Ian Iscoe. On the supports of measure-valued critical branching brownian motion. *The Annals of Probability*, pages 200–221, 1988.
[40] Elena Issoglio and Martina Zähle. Regularity of the solutions to spdes in metric measure spaces. *Stochastic Partial Differential Equations: Analysis and Computations*, 3:272–289, 2015.
[41] Davar Khoshnevisan. *Multiparameter Processes: an introduction to random fields*. Springer Science & Business Media, 2002.
[42] Jun Kigami. *Analysis on fractals*. Number 143. Cambridge University Press, 2001.
[43] Jean-François Le Gall. A class of path-valued markov processes and its applications to superprocesses. *Probability Theory and Related Fields*, 95:25–46, 1993.
[44] Jean-François Le Gall. A probabilistic poisson representation for positive solutions of δu= u2 in a planar domain. *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, 50(1):69–103, 1997.
[45] Jean-François Le Gall. *Brownian motion, martingales, and stochastic calculus*. Springer, 2016.
[46] Michel Ledoux and Michel Talagrand. *Probability in Banach Spaces: isoperimetry and processes*. Springer Science & Business Media, 2013.
[47] Z. Li. *Measure-valued branching Markov processes*. Probability and its Applications (New York). Springer, Heidelberg.
[48] Jouni Luukkainen and Eero Saksman. Every complete doubling metric space carries a doubling measure. *Proceedings of the American Mathematical Society*, 126(2):531–534, 1998.
[49] Pascal Maillard, Rémi Rhodes, Vincent Vargas, and Ofer Zeitouni. Liouville heat kernel: regularity and bounds. 2016.
[50] Avi Mayorcas and Harprit Singh. Singular spdes on homogeneous lie groups. *arXiv preprint arXiv:2301.05121*, 2023.
[51] Carl Mueller. On the support of solutions to the heat equation with noise. *Stochastics: An International Journal of Probability and Stochastic Processes*, 37(4):225–245, 1991.
[52] Carl Mueller. Some tools and results for parabolic stochastic partial differential equations. *A minicourse on stochastic partial differential equations*, pages 111–144, 2009.
[53] Leonid Mytnik, Edwin Perkins, and Anja Sturm. On pathwise uniqueness for stochastic heat equations with non-lipschitz coefficients. 2006.
[54] Walter Rudin. Functional analysis, mcgrawhill. *Inc, New York*, 45(46):4, 1991.
[55] Tokuzo Shiga. Two contrasting properties of solutions for one-dimensional stochastic partial differential equations. *Canadian Journal of Mathematics*, 46(2):415–437, 1994.
[56] John B. Walsh. An introduction to stochastic partial differential equations. In *École d'été de probabilités de Saint-Flour, XIV—1984*, volume 1180 of *Lecture Notes in Math.*, pages 265–439. Springer, Berlin, 1986.

University of Indianna
Email address, L. Fan: xxx@xxx.edu
Email address, J. Yang: xxx@xxx.edu

School of Mathematics and Statistics, Beijing Institute of Technology, Beijing 100081, China
Email address, Z. Sun: zhenyao.sun@gmail.com