RANDOM CONDUCTANCE MODELS WITH STABLE-LIKE JUMPS I: QUENCHED INVARIANCE PRINCIPLE

XIN CHEN TAKASHI KUMAGAI JIAN WANG

Abstract. We study the quenched invariance principle for random conductance models with long range jumps on $\mathbb{Z}^d$, where the transition probability from x to y is in average comparable to $|x-y|^{-(d+\alpha)}$ with $\alpha \in (0,2)$ but possibly degenerate. Under some moment conditions on the conductance, we prove that the scaling limit of the Markov process is a symmetric α-stable Lévy process on $\mathbb{R}^d$. The well-known corrector method in homogenization theory does not seem to work in this setting. Instead, we utilize probabilistic potential theory for the corresponding jump processes. Two essential ingredients of our proof are the tightness estimate and the Hölder regularity of parabolic functions for non-elliptic α-stable-like processes on graphs. Our method is robust enough to apply not only for $\mathbb{Z}^d$ but also for more general graphs whose scaling limits are nice metric measure spaces.

Keywords: random conductance model; long range jump; stable-like process; quenched invariance principle

MSC 2010: 60G51; 60G52; 60J25; 60J75.

1. Introduction and Main Results

Over the last decade, significant progress has been made concerning the quenched invariance principle on random conductance models. A typical and important example is random walk on the infinite cluster of supercritical bond percolation on $\mathbb{Z}^d$. It is shown that the scaling limit of the random walk is a (constant time change of) Brownian motion on $\mathbb{R}^d$ in the quenched sense, namely almost surely with respect to the randomness of the media. See [2, 9, 14, 17, 20, 33, 34, 37] for related progress on this subject and [16, 32] for overall introduction on this area and related topics. Besides i.i.d. nearest-neighbour random conductance models, recently there are great developments on the scaling limit of short range random conductance models on stationary ergodic media (or the media with suitable correlation conditions), see [3, 4, 5, 18, 29, 36] for more details. Here, short range means only finite number of conductances are directly connected to each vertex.

Unlike the short range case, there are only a few results concerning quenched invariance principle for long range random conductance models due to their fundamental technical difficulties. There is a beautiful paper by Crawford and Sly [27] that obtains the quenched invariance principle for random walk on the long range percolation cluster to an isotropic α-stable Lévy process in the range $0 < \alpha < 1$. While [27] proves the invariance principle for a very singular object like the long range percolation, the arguments heavily rely on the special properties (see for instance [13, 15, 26] for related discussions) of the long range percolation and cannot be easily generalized to the setting of general (long range) random conductance models.

In this paper, we will discuss the quenched invariance principle on long range random conductance models. In particular, we consider the case where the conductance between x and y is in average comparable to $|x-y|^{-(d+\alpha)}$ with $\alpha \in (0,2)$ but possibly degenerate. In this setting, there is a significant difficulty in applying classical techniques of homogenization for nearest-neighbour random walk (in random environment) due to the existence of long range

X. Chen: Department of Mathematics, Shanghai Jiao Tong University, 200240 Shanghai, P.R. China. chenxin217@sjtu.edu.cn.

T. Kumagai: Research Institute for Mathematical Sciences, Kyoto University, Kyoto 606-8502, Japan. kumagai@kurims.kyoto-u.ac.jp.

J. Wang: College of Mathematics and Informatics & Fujian Key Laboratory of Mathematical Analysis and Applications, Fujian Normal University, 350007 Fuzhou, P.R. China. jianwang@fjnu.edu.cn.

conductances. To emphasize the novelty of our paper, we first make some remarks. Some more details and technical difficulties of our methods are further discussed in the end of the introduction.

(i) The well known harmonic decomposition method (also called the corrector method in the literature) has been widely used for the nearest-neighbour random walk in random media, see [2, 3, 4, 5, 9, 14, 18, 37]. Because of the lack of L^2 integrability, such method does not work (at least in a straightforward way) for our long range model here.

(ii) Due to singularity in the infinite cluster of long range percolation, [27] established the quenched invariance principle of the associated random walk in the sense of weak convergence on L^q (not the Skorohod topology) and only for the case $0 < \alpha < 1$. In the present paper, we can justify quenched invariance principle of our model under the Skorohod topology for all $\alpha \in (0, 2)$. (To be fair, the long range percolation is "more singular", and it is not included in our conductance model.) Moreover, compared with [22], we can prove the quenched invariance principle for the process with fixed initial point, see e.g. Remark 4.6 below.

(iii) Our approach is to utilize recently developed de Giorgi-Nash-Moser theory for jump processes (see for instance [7, 23, 24, 25]). While detailed heat kernel estimates and Harnack inequalities are established for uniformly elliptic α-stable-like processes, the arguments rely on pointwise estimates of the jumping density (conductance in this setting), which cannot hold in our setting unless we assume uniform ellipticity of conductance. Furthermore, as will be shown in the accompanied paper [19], Harnack inequalities do not hold (even for large enough balls) in general on long range random conductance models. By these reasons, highly non-trivial modifications are required to work on the present random conductance setting. Roughly speaking, in this paper we are concerned with the long rang conductance model with some large scale summable conditions on the conductance, which in some sense can be viewed as a counterpart of the so-called "good ball condition" in [6, 8] to the non-local setting. We believe that our methods are rather robust and could be fundamental tools in exploring scaling limits of random walks on long range random media.

(iv) The advantage of our methods is that they do not use translation invariance of the original graph (we do not use the idea of "the environment viewed from the particle"); hence they are applicable not only for $\mathbb{Z}^d$ but also for more general graphs whose scaling limits are nice metric measure spaces. Even in the setting of $\mathbb{Z}^d$, our results can apply to the case that the conductance is independent but possibly degenerate and not necessarily identically distributed; that is, our results are efficient for some long range random walks on degenerate non-stationary ergodic media. The disadvantage is, since we use the Borel-Cantelli lemma to deduce quenched estimates, the arguments require "strong mixing properties" of the random conductance (see (5.4)–(5.10) below). Hence our method cannot be generalized to general stationary ergodic case on $\mathbb{Z}^d$.

To illustrate our contribution, we present the statement about the quenched invariance principle on a half/quarter space $F := \mathbb{R}_+^{d_1} \times \mathbb{R}^{d_2}$ where $d_1, d_2 \in \mathbb{N} \cup \{0\}$. The readers may refer to Sections 4 and 5 for general results. Let $\mathbb{L} := \mathbb{Z}_+^{d_1} \times \mathbb{Z}^{d_2}$, and let $E_{\mathbb{L}}$ be the set of edges associated with $\mathbb{L}$. Consider a Markov generator

$$L_{\mathbb{L}}^\omega f(x) = \sum_{y \in \mathbb{L}} (f(y) - f(x)) \frac{w_{x,y}(\omega)}{|x-y|^{d+\alpha}}, \quad x \in \mathbb{L}, \tag{1.1}$$

where $d = d_1 + d_2$, $\alpha \in (0, 2)$ and $\{w_{x,y}(\omega) : x, y \in \mathbb{L}\}$ is a sequence of random variables such that $w_{x,y}(\omega) = w_{y,x}(\omega) \geqslant 0$ for all $x \neq y$. We use the convention that $w_{x,x}(\omega) = w_{x,x}^{-1}(\omega) = 0$ for all $x \in \mathbb{L}$. Let $(X_t^\omega)_{t \geqslant 0}$ be the corresponding Markov process. For every $n \geqslant 1$ and $\omega \in \Omega$, we define a process $X_\cdot^{(n),\omega}$ on $V_n = n^{-1}\mathbb{L}$ by $X_t^{(n),\omega} := n^{-1} X_{n^\alpha t}^\omega$ for any $t \geqslant 0$. Let $\mathbb{P}_x^{(n),\omega}$ be the law of $X_\cdot^{(n),\omega}$ with initial point $x \in V_n$. Let $Y := ((Y_t)_{t \geqslant 0}, (\mathbb{P}_x^Y)_{x \in F})$ be a F-valued strong Markov process. We say that the quenched invariance principle holds for $X_\cdot^\omega$ with limit process being Y, if for any $\{x_n \in V_n : n \geqslant 1\}$ such that $\lim_{n\to\infty} x_n = x$ for some $x \in F$, it holds that for $\mathbb{P}$-a.s. $\omega \in \Omega$ and every $T > 0$, $\mathbb{P}^{(n),\omega}_{x_n}$ converges weakly to $\mathbb{P}^Y_x$ on the space of all probability measures on $\mathscr{D}([0,T];F)$, the collection of càdlàg F-valued functions on $[0,T]$ equipped with the Skorohod topology.

Theorem 1.1. *Let $d > 4-2\alpha$. Suppose that $\{w_{x,y} : (x,y) \in E_{\mathbb{L}}\}$ is a sequence of non-negative independent random variables such that $\mathbb{E}w_{x,y} = 1$ for all $x,y \in \mathbb{L}$,*

$$\sup_{x,y\in\mathbb{L},x\neq y} \mathbb{P}\big(w_{x,y} = 0\big) < 2^{-4} \tag{1.2}$$

and

$$\sup_{x,y\in\mathbb{L}} \mathbb{E}[w^{2p}_{x,y}] < \infty, \quad \sup_{x,y\in\mathbb{L}} \mathbb{E}[w^{-2q}_{x,y}\mathbb{1}_{\{w_{x,y}>0\}}] < \infty \tag{1.3}$$

for $p,q \in \mathbb{Z}_+$ with

$$p > \max\big\{(d+2)/d, (d+1)/(2(2-\alpha))\big\}, \quad q > (d+2)/d. \tag{1.4}$$

Then the quenched invariance principle holds for $X^\omega_\cdot$ with the limit process being a symmetric α-stable Lévy process Y on F with jumping measure $|z|^{-d-\alpha}\,dz$.

Remark 1.2. When $\alpha \in (0,1)$, the conclusion still holds true for $d > 2-2\alpha$, if $p > \max\big\{(d+2)/d, (d+1)/(2(1-\alpha))\big\}$ and $q > (d+2)/d$. See Proposition 5.6 for details.

The probability 2^{-4} in (1.2) is far from optimal. In fact, it can be replaced by the critical probability to ensure that condition (4.15) (with $V_n = n^{-1}\mathbb{L}$ and m_n being the counting measure on V_n) holds almost surely. However, we do not know what exact value of this critical probability. We note that the integrability condition (1.4) is far from optimal too, and we also do not even know what could be the optimal integrability condition.

Here is one simple example that satisfies (1.2) and (1.3): for each distinct $x,y \in \mathbb{Z}^d$,

$$\mathbb{P}(w_{x,y} = |x-y|^{\varepsilon}) = (3|x-y|^{2p\varepsilon})^{-1}, \quad \mathbb{P}(w_{x,y} = |x-y|^{-\delta}) = (3|x-y|^{2q\delta})^{-1},$$
$$\mathbb{P}\big(w_{x,y} = 0\big) = 2^{-5}, \quad \mathbb{P}(w_{x,y} = g(x,y)) = 1-(3|x-y|^{2p\varepsilon})^{-1}-(3|x-y|^{2q\delta})^{-1}-2^{-5},$$

where $\varepsilon,\delta > 0$ and $g(x,y)$ are chosen so that $\mathbb{E}w_{x,y} = 1$. (It is easy to see that $c^{-1} \leqslant g(x,y) \leqslant c$ for some constant $c \geqslant 1$.)

In the end of the introduction, let us briefly discuss technical difficulties and the ideas of the proof. There are two essential ingredients in our proof; namely the tightness estimate and the Hölder regularity of parabolic functions for non-elliptic α-stable-like processes on graphs. In order to obtain the former estimate, we first split small jumps and big jumps, which is a standard approach for jump processes, and then change the conductance to the averaged one outside a ball (we call it localization method). By this localization and the on-diagonal heat kernel upper bound (Proposition 2.2), we can apply the so-called Bass-Nash method to control the mean displacement of the process (Proposition 2.3). The tightness estimate (Theorem 3.4) is established by comparing the original process, truncated process and the localized process. We note that when $0 < \alpha < 1$, tightness can be proved in a much simpler way using martingale arguments (Proposition 3.5). The key ingredient for the Hölder regularity of parabolic functions (Theorem 3.8) is to deduce the Krylov-type estimate (Proposition 3.6) that controls the hitting probability to a large set before exiting some parabolic cylinder. Once these estimates are established, we use the arguments in [22] to deduce generalized Mosco convergence, and then obtain the weak convergence (Theorem 4.5).

2. Truncated α-stable-like processes on graphs

In the following few sections, we fix graphs and discuss α-stable-like processes on them. Hence we do not consider randomness of the environment. With a slight abuse of notation, we still use $w_{x,y}$ as the deterministic version. Let $G = (V, E_V)$ be a locally finite and connected graph, where V is the set of vertices, and E_V the set of edges. For any $x \neq y \in V$, we write

$\rho(x, y)$ for the graph distance, i.e., $\rho(x, y)$ is the smallest positive length of a path (that is, a sequence $x_0 = x, x_1, \cdots, x_l = y$ such that $(x_i, x_{i+1}) \in E_V$ for all $0 \leqslant i \leqslant l-1$) joining x and y. Set $\rho(x, x) = 0$ for all $x \in V$. We let $B(x, r) = \{y \in V : \rho(x, y) \leqslant r\}$ denote the ball in graph metric with center $x \in V$ and radius $r > 0$. Let μ be a measure on V such that $\mu_x := \mu(\{x\})$ satisfies for some constant $c_M \geqslant 1$ that

$$c_M^{-1} \leqslant \mu_x \leqslant c_M, \quad x \in V. \tag{2.1}$$

For each $p \in [1, \infty)$, let $L^p(V; \mu) = \{f \in \mathbb{R}^V : \sum_{x \in V} |f(x)|^p \mu_x < \infty\}$, and denote by $\|f\|_p$ the L^p norm of f with respect to μ. Let $L^\infty(V; \mu)$ be the space of bounded measurable functions on V, and let $\|f\|_\infty$ be the L^∞ norm of f. We assume that (G, μ) satisfies the d-set condition with $d > 0$, i.e., there exist $r_G \in [1, \infty]$ and $c_G \geqslant 1$ such that

$$c_G^{-1} r^d \leqslant \mu(B(x, r)) \leqslant c_G r^d, \quad x \in V, 1 \leqslant r < r_G. \tag{2.2}$$

We consider the operator $Lf(x) = \sum_{z \in V} (f(z) - f(x)) \frac{w_{x,z}}{\rho(x,z)^{d+\alpha}} \mu_z$ and the quadratic form

$$D(f, f) = \frac{1}{2} \sum_{x,y \in V} (f(x) - f(y))^2 \frac{w_{x,y}}{\rho(x, y)^{d+\alpha}} \mu_x \mu_y, \quad f \in \mathscr{F} = \{f \in L^2(V; \mu) : D(f, f) < \infty\},$$

where $\alpha \in (0, 2)$ and $\{w_{x,y} : x, y \in V\}$ is a sequence such that $w_{x,x} = 0$ for all $x \in V$, $w_{x,y} \geqslant 0$ and $w_{x,y} = w_{y,x}$ for all $x \neq y$, and

$$\sum_{y \in V} \frac{w_{x,y}}{\rho(x, y)^{d+\alpha}} \mu_y < \infty, \quad x \in V. \tag{2.3}$$

Here by convention we set $0/0 = 0$. According to (the first statement in) [22, Theorem 3.2], $(D, \mathscr{F})$ is a regular symmetric Dirichlet form on $L^2(V; \mu)$. Let $X := (X_t)_{t \geqslant 0}$ be the symmetric Hunt process associated with $(D, \mathscr{F})$. Set $C_{x,y} := w_{x,y}/\rho(x, y)^{d+\alpha}$. Under $\mathbb{P}^x$, $X_0 = x$; then the process X waits for an exponentially distributed random time of parameter $C_x := \sum_{y \in V} C_{x,y} \mu_y$ and jumps to point $y \in V$ with probability $C_{x,y} \mu_y / C_x$; this procedure is then iterated choosing independent hopping times. Such a Markov process is called a variable speed random walk on V.

We write $p(t, x, y)$ for the heat kernel of X on V; that is, the transition density of the process X with respect to μ which is defined by $p(t, x, y) = \mu_y^{-1} \mathbb{P}^x(X_t = y)$.

2.1. **On-diagonal upper estimates for heat kernel.** In this subsection, we are concerned with the truncated Dirichlet form corresponding to $(D, \mathscr{F})$. For fixed $1 \leqslant \delta < r_G$, define the operator $L^\delta f(x) = \sum_{z \in V : \rho(z,x) \leqslant \delta} \big(f(z) - f(x)\big) \frac{w_{z,x}}{\rho(z,x)^{d+\alpha}} \mu_z$. Then, the associated bilinear form is given by

$$D^\delta(f, f) = \frac{1}{2} \sum_{x,y \in V : \rho(x,y) \leqslant \delta} \big(f(x) - f(y)\big)^2 \frac{w_{x,y}}{\rho(x, y)^{d+\alpha}} \mu_x \mu_y.$$

Throughout this part, we always assume that

$$C_{V,\delta} := \sup_{x \in V} \sum_{y \in V : \rho(x,y) > \delta} \frac{w_{x,y}}{\rho(x, y)^{d+\alpha}} \mu_y < \infty. \tag{2.4}$$

By (2.4) and the symmetry of $w_{x,y}$, we can easily see that for all $f \in \mathscr{F}$,

$$D^\delta(f, f) \leqslant D(f, f) \leqslant D^\delta(f, f) + 2 \sum_{x \in V} f(x)^2 \mu_x \sum_{y \in V : \rho(y,x) > \delta} \frac{w_{x,y}}{\rho(x, y)^{d+\alpha}} \mu_y \leqslant D^\delta(f, f) + 2 C_{V,\delta} \|f\|_2^2.$$

Consequently, $(D^\delta, \mathscr{F})$ is also a regular and symmetric Dirichlet form on $L^2(V; \mu)$. Denote by $X^\delta := \big((X_t^\delta)_{t \geqslant 0}, (\mathbb{P}_x)_{x \in V}\big)$ the associated Hunt process, which is called the truncated process associated with X in the literature.

In order to get on-diagonal upper estimates for the heat kernel of the truncated process X^δ, we need the following scaled Poincaré-type inequality. In the following, given a sequence of $w := \{w_{x,y} : x, y \in V\}$, for every $x \in V$ and $r \geqslant 1$, we set $B^w(x, r) := \{z \in B(x, r) : w_{x,z} > 0\}$.

Lemma 2.1. *Suppose that there exist constants $C_1, C_2 > 0$ and $1 \leqslant r_0 < r_G$ such that*

$$\sup_{x\in V} \sum_{y\in B^w(x,r_0)} w_{x,y}^{-1} \leqslant C_1 r_0^d \tag{2.5}$$

and

$$\inf_{x\in V} \mu(B^w(x,r_0)) \geqslant C_2 r_0^d, \tag{2.6}$$

where C_1 and C_2 are independent of r_0 and r_G. Then there is a constant $C_3 > 0$ (also independent of r_0 and r_G) such that for all $x \in V$ and measurable function f on V,

$$\sum_{z\in B(x,r_0)} (f(z)-(f)_{B^w(z,r_0)})^2\mu_z \leqslant C_3 r_0^\alpha \sum_{z\in B(x,r_0),y\in B(x,2r_0)} (f(z)-f(y))^2\frac{w_{z,y}}{\rho(z,y)^{d+\alpha}}\mu_z\mu_y, \tag{2.7}$$

where for $A \subset V$, $(f)_A := \mu(A)^{-1}\sum_{z\in A} f(x)\mu_z$.

Proof. For every $x \in V$ and measurable function f on V, we have

$$\begin{aligned}
&\sum_{z\in B(x,r_0)} (f(z)-(f)_{B^w(z,r_0)})^2\mu_z = \sum_{z\in B(x,r_0)} \Big(\frac{1}{\mu(B^w(z,r_0))}\sum_{y\in B^w(z,r_0)} (f(z)-f(y))\mu_y\Big)^2\mu_z\\
&\leqslant \frac{c_1}{r_0^{2d}} \sum_{z\in B(x,r_0)} \Big[\Big(\sum_{y\in B^w(z,r_0)} (f(z)-f(y))^2\frac{w_{z,y}}{\rho(z,y)^{d+\alpha}}\Big)\Big(\sum_{y\in B^w(z,r_0)} w_{z,y}^{-1}\rho(z,y)^{d+\alpha}\Big)\Big]\\
&\leqslant c_2 r_0^{-d+\alpha}\Big(\sup_{z\in V}\sum_{y\in B^w(z,r_0)} w_{z,y}^{-1}\Big)\Big(\sum_{z\in B(x,r_0),y\in B(x,2r_0)} \big(f(z)-f(y)\big)^2\frac{w_{z,y}}{\rho(z,y)^{d+\alpha}}\Big)\\
&\leqslant c_3 r_0^\alpha \sum_{z\in B(x,r_0),y\in B(x,2r_0)} \big(f(z)-f(y)\big)^2\frac{w_{z,y}}{\rho(z,y)^{d+\alpha}}\mu_z\mu_y,
\end{aligned}$$

where the first inequality follows from (2.1), (2.6) and the Cauchy-Schwarz inequality, in the second inequality we have used the fact that $\rho(z,y) \leqslant r_0$ for every $y \in B^w(z,r_0)$, and the third inequality is due to (2.1) and (2.5). This proves (2.7). □

In the following, we denote by $p^\delta(t,x,y)$ the heat kernel of X^δ.

Proposition 2.2. *Suppose that (2.4) holds, and that there exist constants $\theta \in (0,1)$ and $C_1, C_2 \in (0,\infty)$ (which are independent of δ and r_G) such that for every $\delta^\theta \leqslant r \leqslant \delta$,*

$$\sup_{x\in V} \sum_{y\in B^w(x,r)} w_{x,y}^{-1} \leqslant C_1 r^d, \tag{2.8}$$

$$\inf_{x\in V} \mu\big(B^w(x,r)\big) \geqslant C_2 r^d \tag{2.9}$$

and

$$\sup_{x\in V} \sum_{y\in V:\rho(y,x)\leqslant r} \frac{w_{x,y}}{\rho(x,y)^{d+\alpha-2}} \leqslant C_1 r^{2-\alpha}. \tag{2.10}$$

Then, for each $\theta' \in (\theta,1)$, there is a constant $\delta_0 > 0$ (which only depends on θ' and θ) such that for all $\delta_0 \leqslant \delta < r_G$,

$$p^\delta(t,x,y) \leqslant C_3 t^{-d/\alpha}, \quad \forall\, 2\delta^{\theta'\alpha} \leqslant t \leqslant \delta^\alpha \text{ and } x,y \in V, \tag{2.11}$$

where C_3 is a positive constant independent of δ_0, δ, t, x, y and r_G.

Proof. The proof is partially motivated by that of [6, Propisition 3.1], but some non-trivial modification is required. Without mention, throughout the proof constant c_i will be independent of δ, t, x, y and r_G. Since, by the Cauchy-Schwarz inequality, $p^\delta(t,x,y) \leqslant p^\delta(t,x,x)^{1/2}p^\delta(t,y,y)^{1/2}$ for any $t > 0$ and $x, y \in V$, it suffices to verify (2.11) for the case that $x = y$. The proof is split into three steps.

Step (1): We first note that under (2.4) and (2.10), $\sup_{x\in V}\sum_{y\in V}\frac{w_{x,y}}{\rho(x,y)^{d+\alpha}}\mu_y<\infty$. This along with (the second statement in) [22, Theorem 3.2] yields that the process X^δ is conservative. By [28, Proposition 5 and Theorem 8], we have the following upper bound for $p^\delta(t,x,y)$:

$$p^\delta(t,x_1,x_2)\leqslant \mu_{x_1}^{-1/2}\mu_{x_2}^{-1/2}\inf_{\psi\in L^\infty(V;\mu)}\exp\big(\phi(x_1)-\phi(x_2)+b(\phi)t\big) \tag{2.12}$$

for all $t>0$ and $x_1,x_2\in V$, where

$$b(\phi):=\frac{1}{2}\sup_{x\in V}\sum_{y\in V:\rho(y,x)\leqslant\delta}\frac{w_{x,y}}{\rho(x,y)^{d+\alpha}}\Big(e^{\phi(y)-\phi(x)}+e^{\phi(x)-\phi(y)}-2\Big)\mu_y.$$

For fixed $x_1,x_2\in V$, taking $\phi(x)=\rho(x,x_1)\wedge\rho(x_1,x_2)$ for any $x\in V$, we get that

$$\begin{aligned}b(\phi)&\leqslant\frac{1}{2}\sup_{x\in V}\sum_{y\in V:\rho(y,x)\leqslant\delta}\frac{w_{x,y}}{\rho(x,y)^{d+\alpha}}\Big(e^{\rho(x,y)}+e^{-\rho(x,y)}-2\Big)\mu_y\\&\leqslant\frac{1}{2}\sup_{x\in V}\sum_{y\in V:\rho(y,x)\leqslant\delta}\frac{w_{x,y}}{\rho(y,x)^{d+\alpha}}\rho(x,y)^2e^{\rho(x,y)}\mu_y\\&\leqslant\frac{1}{2}e^\delta\sup_{x\in V}\sum_{y\in V:\rho(y,x)\leqslant\delta}\frac{w_{x,y}}{\rho(x,y)^{d+\alpha-2}}\mu_y\leqslant c_1e^\delta\delta^{2-\alpha}\leqslant 2c_1e^{2\delta},\end{aligned}$$

where in the first inequality above we have used the facts that $s\mapsto e^s+e^{-s}$ is increasing on $[0,\infty)$ and $|\phi(x)-\phi(y)|\leqslant\rho(x,y)$ for all $x,y\in V$, the second inequality is due to the fact that $e^s+e^{-s}-2\leqslant s^2e^s$ for all $s\geqslant 0$, and the fourth inequality follows from (2.10). Combining this with (2.12), we arrive at that for all $t>0$ and $x_1,x_2\in V$,

$$p^\delta(t,x_1,x_2)\leqslant c_M\exp\big(-\rho(x_1,x_2)+2c_1e^{2\delta}t\big). \tag{2.13}$$

Furthermore, it follows from the symmetry of $w_{x,y}$, the fact that $p^\delta(t,x,y)\mu_y\leqslant 1$ for all $t>0$ and $x,y\in V$, (2.10) and (2.13) that for every $x\in V$,

$$\begin{aligned}&\sum_{z,v\in V:\rho(z,v)\leqslant\delta}\big(p^\delta(t,x,z)-p^\delta(t,x,v)\big)^2\frac{w_{z,v}}{\rho(z,v)^{d+\alpha}}\mu_z\mu_v\\&\leqslant\sum_{z,v\in V:\rho(z,v)\leqslant\delta}\big(p^\delta(t,x,z)+p^\delta(t,x,v)\big)^2\frac{w_{z,v}}{\rho(z,v)^{d+\alpha}}\mu_z\mu_v\\&\leqslant 4c_M\sum_{z\in V}p^\delta(t,x,z)\Big(\sup_{z\in V}\sum_{v\in V:\rho(v,z)\leqslant\delta}\frac{w_{z,v}}{\rho(z,v)^{d+\alpha}}\Big)\\&\leqslant 4c_M\sum_{z\in V}p^\delta(t,x,z)\Big(\sup_{z\in V}\sum_{v\in V:\rho(z,v)\leqslant\delta}\frac{w_{z,v}}{\rho(z,v)^{d+\alpha-2}}\Big)\leqslant c_2(\delta,t)\sum_{z\in V}\exp(-\rho(z,x))<\infty,\end{aligned}$$

where in the last inequality we used the fact that

$$\sum_{z\in V}\exp(-\rho(z,x))\leqslant c_M\sum_{r=0}^\infty\sum_{z\in V:\rho(x,z)=r}e^{-r}\mu_z\leqslant c_M\sum_{r=0}^\infty\mu(B(x,r))e^{-r}\leqslant c_Mc_G\sum_{r=1}^\infty r^de^{-r}<\infty.$$

Therefore, according to the Fubini theorem and (2.13), for every $x\in V$,

$$\sum_{z\in V}L^\delta p^\delta(t,x,\cdot)(z)p^\delta(t,x,z)\mu_z=-\frac{1}{2}\sum_{z,v\in V}\big(p^\delta(t,x,z)-p^\delta(t,x,v)\big)^2\frac{w_{z,v}}{\rho(z,v)^{d+\alpha}}\mu_z\mu_v. \tag{2.14}$$

Step (2): Below we fix $x\in V$. Let $f_t(z)=p^\delta(t,x,z)$ and $\psi(t)=p^\delta(2t,x,x)$ for all $z\in V$ and $t\geqslant 0$. Then, $\psi(t)=\sum_{z\in V}f_t(z)^2\mu_z$, and, by (2.14),

$$\psi'(t)=2\sum_{z\in V}\frac{df_t(z)}{dt}f_t(z)\mu_z=2\sum_{z\in V}L^\delta f_t(z)f_t(z)\mu_z=-\sum_{z,y\in V}(f_t(z)-f_t(y))^2\frac{w_{z,y}}{\rho(z,y)^{d+\alpha}}\mu_z\mu_y.$$

Let $\delta^\theta\leqslant r(t)\leqslant\delta$ and $R:=R(\delta)\geqslant 1$ be some constants to be determined later. Suppose that $B(x_i,r(t)/2)$ $(i=1,\cdots,m)$ is the maximal collection of disjoint balls with centers in $B(x,R)$.

Set $B_i = B(x_i, r(t))$ and $B_i^* = B(x_i, 2r(t))$. Then, $B(x,R) \subset \cup_{i=1}^m B_i \subset B(x, R+r(t)) \subset \cup_{i=1}^m B_i^*$; moreover, if $z \in B(x, R+r(t)) \cap B_i^*$ for some $1 \leqslant i \leqslant m$, then $B(x_i, r(t)/2) \subset B(z, 3r(t))$, and so

$$c_3 r(t)^d \geqslant \mu(B(z,3r(t))) \geqslant \sum_{i=1}^m \mathbb{1}_{\{z\in B_i^*\}}\mu(B(x_i, r(t)/2)) \geqslant c_4 r(t)^d |\{i : z \in B_i^*\}|,$$

where in the second inequality we used the fact that $B(x_i, r(t)/2)$, $i = 1, \cdots, m$, are disjoint, and in the first and the last inequality we have used (2.2). Thus, every $z \in B(x, R+r(t))$ is in at most $c_5 := c_3/c_4$ of the ball B_i^* (hence at most c_5 of the ball B_i). In particular,

$$\sum_{i=1}^m \sum_{z\in B_i} = \sum_{i=1}^m \sum_{z\in B(x,R+r(t))} \mathbb{1}_{B_i}(z) = \sum_{z\in B(x,R+r(t))} \sum_{i=1}^m \mathbb{1}_{B_i}(z) \leqslant c_5 \sum_{z\in B(x,R+r(t))} . \tag{2.15}$$

According to (the proof of) Lemma 2.1, (2.8) and (2.9) imply that for every $\delta^\theta \leqslant r \leqslant \delta$, $x \in V$ and measurable function f on V,

$$\sum_{z\in B(x,r)} (f(z) - (f)_{B^w(z,r)})^2 \mu_z \leqslant c_6 r^\alpha \sum_{z\in B(x,r), y\in B(x,2r)} (f(z) - f(y))^2 \frac{w_{z,y}}{\rho(z,y)^{d+\alpha}} \mu_z \mu_y. \tag{2.16}$$

Hence, noticing that $\delta^\theta \leqslant r(t) \leqslant \delta$,

$$\begin{aligned}\sum_{z,y\in V} (f_t(z) - f_t(y))^2 \frac{w_{z,y}}{\rho(z,y)^{d+\alpha}} \mu_z\mu_y &\geqslant \frac{1}{c_5} \sum_{i=1}^m \sum_{z\in B_i} \sum_{y\in B_i^*} (f_t(z) - f_t(y))^2 \frac{w_{z,y}}{\rho(z,y)^{d+\alpha}} \mu_z\mu_y \\ &\geqslant \frac{c_7}{r(t)^\alpha} \Big[\sum_{i=1}^m \sum_{z\in B_i} f_t^2(z)\mu_z - 2\sum_{i=1}^m \sum_{z\in B_i} f_t(z)(f_t)_{B^w(z,r(t))}\mu_z \Big] =: \frac{c_7}{r(t)^\alpha}(I_1 - I_2),\end{aligned}$$

where in the second inequality we have used (2.16).

Furthermore, since $f_t(z)\mu_z \leqslant 1$ for all $z \in V$ and $t > 0$, we have

$$I_1 \geqslant \sum_{z\in \cup_{i=1}^m B_i} f_t^2(z)\mu_z \geqslant \sum_{z\in B(x,R)} f_t^2(z)\mu_z = \psi(t) - \sum_{z\in V:\rho(z,x)>R} f_t^2(z)\mu_z \geqslant \psi(t) - \sum_{z\in V:\rho(z,x)>R} f_t(z).$$

So, by (2.13), we can choose $R := R(\delta) = 2c_1 e^{4\delta}$ such that for all $\delta^{\theta\alpha} \leqslant t \leqslant \delta^\alpha$,

$$\begin{aligned}\sum_{z\in V:\rho(z,x)>R} f_t(z) &\leqslant \sum_{z\in V:\rho(z,x)>2c_1e^{4\delta}} \exp\big(-\rho(z,x) + 2c_1 e^{2\delta}\delta^\alpha\big) \\ &\leqslant c_M \sum_{z\in V:\rho(z,x)>2c_1e^{4\delta}} \exp\big(-\rho(z,x)/2\big)\mu_z \\ &\leqslant c_M \sum_{r=2c_1e^{4\delta}}^\infty \mu(B(x,r))e^{-r/2} \leqslant c_8\delta^{-d} \leqslant c_8 r(t)^{-d},\end{aligned}$$

where the last inequality follows from the fact that $r(t) \leqslant \delta$. On the other hand, due to (2.9) and the fact that $\sum_{z\in V} f_t(z)\mu_z \leqslant 1$ for all $t > 0$,

$$\sup_{z\in V}(f_t)_{B^w(z,r(t))} \leqslant \sup_{z\in V} \mu\big(B^w(z,r(t))\big)^{-1} \cdot \sum_{z\in V} f_t(z)\mu_z \leqslant C_2^{-1} r(t)^{-d}.$$

This along with (2.15) yields that

$$I_2 \leqslant C_2^{-1} r(t)^{-d} \sum_{i=1}^m \sum_{z\in B_i} f_t(z)\mu_z \leqslant C_2^{-1} c_5 r(t)^{-d} \sum_{z\in B(x,R+r(t))} f_t(z)\mu_z \leqslant C_2^{-1} c_5 r(t)^{-d}.$$

Therefore, combining all estimates above, we arrive at that for every $\delta^\theta \leqslant r(t) \leqslant \delta$,

$$\psi'(t) \leqslant -c_9 r(t)^{-\alpha} \left(\psi(t) - c_{10} r(t)^{-d}\right). \tag{2.17}$$

Step (3): For any $\theta' \in (\theta, 1)$ and any $1 \leqslant \delta < r_G$ large enough, we claim that there exists $t_0 \in [\delta^{\theta\alpha}, \delta^{\theta'\alpha}]$ such that

$$\left(\frac{1}{2c_{10}}\psi(t_0)\right)^{-1/d} \geqslant \delta^{\theta}. \tag{2.18}$$

Indeed, suppose that (2.18) does not hold. Then,

$$\left(\frac{1}{2c_{10}}\psi(t)\right)^{-1/d} < \delta^{\theta}, \quad \forall\, \delta^{\theta\alpha} \leqslant t \leqslant \delta^{\theta'\alpha}, \tag{2.19}$$

which means that $\psi(t) \geqslant 2c_{10}\delta^{-d\theta}$ for all $\delta^{\theta\alpha} \leqslant t \leqslant \delta^{\theta'\alpha}$. Hence, taking $r(t) = \delta^{\theta}$ in (2.17), we find that $\psi'(t) \leqslant -2^{-1}c_9\delta^{-\theta\alpha}\psi(t)$ for any $\delta^{\theta\alpha} \leqslant t \leqslant \delta^{\theta'\alpha}$, which along with the fact $\psi(t) \leqslant \mu_x^{-1} \leqslant c_M$ for all $t > 0$ yields that $\psi(t) \leqslant c_M e^{-2^{-1}c_9\delta^{-\theta\alpha}(t-\delta^{\theta\alpha})}$ for any $\delta^{\theta\alpha} \leqslant t \leqslant \delta^{\theta'\alpha}$. In particular, $\psi(\delta^{\theta'\alpha}) \leqslant c_M e^{-2^{-1}c_9\delta^{-\theta\alpha}(\delta^{\theta'\alpha}-\delta^{\theta\alpha})}$. On the other hand, according to (2.19), we have $\psi(\delta^{\theta'\alpha}) \geqslant 2c_{10}\delta^{-d\theta}$. Thus, there is a contradiction between these two inequalities above for δ large enough, and so (2.18) is true.

Next, assume that we can take $1 \leqslant \delta < r_G$ large enough such that (2.18) holds. Since $t \mapsto \psi(t)$ is non-increasing on $(0, \infty)$ and $t_0 \leqslant \delta^{\theta'\alpha}$,

$$\left(\frac{1}{2c_{10}}\psi(t)\right)^{-1/d} \geqslant \delta^{\theta}, \quad \forall\, \delta^{\theta'\alpha} \leqslant t \leqslant \delta^{\alpha}.$$

Let

$$\tilde{t}_0 := \sup\left\{t > 0 : \left(\frac{1}{2c_{10}}\psi(t)\right)^{-1/d} < \delta/2\right\}.$$

By the non-increasing property of ψ on $(0, \infty)$ again, if $\tilde{t}_0 \leqslant \delta^{\theta'\alpha}$, then $\psi(t) \leqslant \psi(\tilde{t}_0) = 2c_{10}(\delta/2)^{-d} \leqslant c_{11}t^{-d/\alpha}$ for any $\delta^{\theta'\alpha} \leqslant t \leqslant \delta^{\alpha}$. This proves (2.11).

When $\tilde{t}_0 > \delta^{\theta'\alpha}$,

$$\delta^{\theta} \leqslant \left(\frac{1}{2c_{10}}\psi(t)\right)^{-1/d} \leqslant \delta/2, \quad \forall\, \delta^{\theta'\alpha} \leqslant t \leqslant \tilde{t}_0.$$

Then, taking $r(t) = \big(\frac{1}{2c_{10}}\psi(t)\big)^{-1/d}$ in (2.17), we have $\psi'(t) \leqslant -c_{12}\psi(t)^{1+d/\alpha}$ for any $\delta^{\theta'\alpha} \leqslant t \leqslant \tilde{t}_0$. Hence, $\psi(s) \leqslant c_{13}\big(s - \delta^{\theta'\alpha} + \psi(\delta^{\theta'\alpha})^{-\alpha/d}\big)^{-d/\alpha} \leqslant c_{14}s^{-d/\alpha}$ for any $2\delta^{\theta'\alpha} \leqslant s \leqslant \tilde{t}_0$. If $\tilde{t}_0 > \delta^{\alpha}$, then (2.11) holds. If $\delta^{\theta'\alpha} < \tilde{t}_0 \leqslant \delta^{\alpha}$, then, for all $\tilde{t}_0 \leqslant s \leqslant \delta^{\alpha}$, $\psi(s) \leqslant \psi(\tilde{t}_0) = 2c_{10}(\delta/2)^{-d} \leqslant c_{15}s^{-d/\alpha}$, so (2.11) also holds. The proof is complete. □

2.2. **Localization method and moment estimates of the truncated process.** In this part, we fix $x_0 \in V$ and $R \geqslant 1$. Define a symmetric regular Dirichlet form $(\hat{D}^{x_0,R}, \hat{\mathscr{F}}^{x_0,R})$ as follows

$$\begin{aligned}\hat{D}^{x_0,R}(f,f) &= \sum_{x,y \in V} \big(f(x) - f(y)\big)^2 \frac{\hat{w}_{x,y}}{\rho(x,y)^{d+\alpha}}\mu_x\mu_y, \quad f \in \hat{\mathscr{F}}^{x_0,R},\\ \hat{\mathscr{F}}^{x_0,R} &= \{f \in L^2(V;\mu) : \hat{D}^{x_0,R}(f,f) < \infty\},\end{aligned}$$

where

$$\hat{w}_{x,y} = \begin{cases} w_{x,y}, & \text{if } x \in B(x_0,R) \text{ or } y \in B(x_0,R),\\ 1, & \text{otherwise.}\end{cases}$$

Note that, according to the definition of $\hat{w}_{x,y}$, for any $x \in V$,

$$\begin{aligned}
\sum_{y\in V}\frac{\hat{w}_{x,y}}{\rho(x,y)^{d+\alpha}} &= \sum_{y\notin B(x_0,R)}\frac{\hat{w}_{x,y}}{\rho(x,y)^{d+\alpha}}+\sum_{y\in B(x_0,R)}\frac{w_{x,y}}{\rho(x,y)^{d+\alpha}}\\
&\leqslant \sup_{z\in B(x_0,R)}\sum_{v\in V}\frac{w_{z,v}}{\rho(z,v)^{d+\alpha}}+\sup_{z\notin B(x_0,R)}\sum_{y\in V:y\neq z}\frac{1}{\rho(z,y)^{d+\alpha}}+\sum_{y\in B(x_0,R)}\frac{w_{x,y}}{\rho(x,y)^{d+\alpha}}\\
&\leqslant \sup_{z\in B(x_0,R)}\sum_{v\in V}\frac{w_{z,v}}{\rho(z,v)^{d+\alpha}}+c_M\sup_{z\notin B(x_0,R)}\sum_{k=1}^{\infty}\sum_{y\in V:2^{k-1}\leqslant\rho(y,z)<2^k}\frac{1}{\rho(y,z)^{d+\alpha}}\mu_y\\
&\quad+\sum_{y\in B(x_0,R)}\left(\sup_{z\in B(x_0,R)}\sum_{v\in V}\frac{w_{z,v}}{\rho(z,v)^{d+\alpha}}\right)\\
&\leqslant \sup_{z\in B(x_0,R)}\sum_{v\in V}\frac{w_{z,v}}{\rho(z,v)^{d+\alpha}}+c_Mc_G\sum_{k=1}^{\infty}\frac{2^{kd}}{2^{(k-1)(d+\alpha)}}+\sum_{y\in B(x_0,R)}\sup_{z\in B(x_0,R)}\sum_{v\in V}\frac{w_{z,v}}{\rho(z,v)^{d+\alpha}}\\
&\leqslant c_1+c_2(1+R^d)\sup_{z\in B(x_0,R)}\left(\sum_{v\in V}\frac{w_{z,v}}{\rho(z,v)^{d+\alpha}}\right)=:C(x_0,R)<\infty,
\end{aligned}\tag{2.20}$$

where (2.3) was used in the fourth inequality. In particular, by (2.20) and (the second statement in) [22, Theorem 3.2], the associated Hunt process $\hat{X}^R := ((\hat{X}^R_t)_{t\geqslant 0}, (\mathbb{P}_x)_{x\in V})$ is conservative. Here and in what follows, we omit the index x_0 for simplicity.

We also consider the following truncated Dirichlet form $(\hat{D}^{x_0,R,R}, \hat{\mathscr{F}}^{x_0,R})$:

$$\hat{D}^{x_0,R,R}(f,f)=\sum_{x,y\in V:\rho(x,y)\leqslant R}\big(f(x)-f(y)\big)^2\frac{\hat{w}_{x,y}}{\rho(x,y)^{d+\alpha}}\mu_x\mu_y,\quad f\in\hat{\mathscr{F}}^{x_0,R}.$$

Let $\hat{X}^{R,R} := ((\hat{X}^{R,R}_t)_{t\geqslant 0}, (\mathbb{P}_x)_{x\in V})$ be the associated Hunt process. In particular, due to (2.20) again, the process $\hat{X}^{R,R}$ is also conservative. Denote by $\hat{p}^R(t,x,y)$ and $\hat{p}^{R,R}(t,x,y)$ heat kernels of the processes $\hat{X}^R$ and $\hat{X}^{R,R}$, respectively.

The following statement is concerned with moment estimates of $\hat{X}^{R,R}$, which are key to yield exit time estimates of the original process X in the next section. We mainly use the method of Bass [12] (see also Barlow [6] and Nash [35]), but some non-trivial modifications are required.

Proposition 2.3. *Suppose that there exist $1 \leqslant R_0 < r_G$ and $\theta \in (0,1)$ such that for every $R_0 < R < r_G$ and $R^\theta \leqslant r \leqslant R$,*

$$\sup_{x\in B(x_0,3R)}\sum_{y\in V:\rho(x,y)\leqslant r}\frac{w_{x,y}}{\rho(x,y)^{d+\alpha-2}}\leqslant C_1r^{2-\alpha},\tag{2.21}$$

$$\inf_{x\in B(x_0,3R)}\mu(B^w(x,r))\geqslant C_2r^d\tag{2.22}$$

and

$$\sup_{x\in B(x_0,3R)}\sum_{y\in B^w(x,r)}w_{x,y}^{-1}\leqslant C_1r^d,\tag{2.23}$$

where C_1 and C_2 are positive constants independent of x_0, R_0, R, r and r_G. Then for every $\theta' \in (\theta,1)$, there exists a constant $R_1 > R_0$ (which depends on θ, θ' and R_0 only) such that for every $R_1 < R < r_G$ and $x \in V$,

$$\mathbb{E}_x\big[\rho\big(\hat{X}^{R,R}_t,x\big)\big]\leqslant C_3R\left(\frac{t}{R^\alpha}\right)^{1/2}\left[1+\log\left(\frac{R^\alpha}{t}\right)\right],\quad\forall\ R^{\theta'\alpha}\leqslant t\leqslant R^\alpha,\tag{2.24}$$

where C_3 is a positive constant independent of x_0, R_1, R, t, x and r_G.

Proof. Throughout the proof, we first suppose that there exist positive constants $c(x_0, R)$ and $\tilde c(x_0, R)$ such that

$$\tilde c(x_0, R) \leqslant \inf_{x,y\in V} \hat w_{x,y} \leqslant \sup_{x,y\in V} \hat w_{x,y} \leqslant c(x_0, R). \tag{2.25}$$

If (2.25) is not satisfied, then, by taking $w^{\varepsilon}_{x,y} := w_{x,y} + \varepsilon$ and then letting $\varepsilon \downarrow 0$, we can prove that (2.24) still holds true. Moreover, all the constants in the proof below are independent of ε unless specifically claimed.

Step (1): By (2.21), (2.22), (2.23) and the definition of $\hat w_{x,y}$, for every $R_0 < R < r_G$ and $R^\theta \leqslant r \leqslant R$,

$$\sup_{x\in V} \sum_{y\in V:\rho(x,y)\leqslant r} \frac{\hat w_{x,y}}{\rho(x, y)^{d+\alpha-2}} \leqslant c_0 r^{2-\alpha}, \tag{2.26}$$

$\inf_{x\in V} \mu(B^{\hat w}(x, r)) \geqslant c_1 r^d$ and $\sup_{x\in V} \sum_{y\in B^{\hat w}(x,r)} \hat w^{-1}_{x,y} \leqslant c_0 r^d$, where $B^{\hat w}(x, r) := \{z \in V : \rho(z, x) \leqslant r,\ \hat w_{z,x} > 0\}$. Let $\theta' \in (\theta, 1)$ and $\theta_0 = (\theta + \theta')/2$. Taking $\rho = R$ in Proposition 2.2, we find that there exists a constant $\tilde R_0 \geqslant R_0$ (which only depends on θ and θ') such that whenever $\tilde R_0 < R < r_G$,

$$\hat p^{R,R}(t, x, y) \leqslant c_2 t^{-d/\alpha}, \quad \forall\ 2R^{\theta_0\alpha} \leqslant t \leqslant R^\alpha,\ x, y \in V. \tag{2.27}$$

For every $t > 0$, we define

$$M(t) = \sum_{y\in V} \rho(x, y)\hat p^{R,R}(t, x, y)\mu_y, \quad Q(t) = -\sum_{y\in V} \hat p^{R,R}(t, x, y) \left[\log \hat p^{R,R}(t, x, y)\right] \mu_y.$$

Below, we fix $x \in V$ and set $f_t(y) = \hat p^{R,R}(t, x, y)$ for all $y \in V$ and $t > 0$.

By (2.25), we can obtain upper and lower bounds for $\hat p^{R,R}(t, x, y)$ (see [28] for upper bounds on graph or [21] for two-sided estimates in the Euclidean space), which yields that

$$\begin{aligned}&\sum_{y,z\in V:\rho(y,z)\leqslant R} |f_t(y) - f_t(z)||\log f_t(y) - \log f_t(z)| \frac{\hat w_{y,z}}{\rho(y, z)^{d+\alpha}}\mu_y\mu_z\\ &\leqslant \sum_{y,z\in V:\rho(y,z)\leqslant R} \big(f_t(y) + f_t(z)\big)\big(|\log f_t(y)| + |\log f_t(z)|\big) \frac{\hat w_{y,z}}{\rho(y, z)^{d+\alpha}}\mu_y\mu_z < \infty.\end{aligned}$$

Thus,

$$\begin{aligned}&-\sum_{y\in V}(\log f_t(y) + 1)\hat L^{R,R} f_t(y)\mu_y\\ &= \frac12 \sum_{y,z\in V:\rho(y,z)\leqslant R} \big(f_t(y) - f_t(z)\big)\big(\log f_t(y) - \log f_t(z)\big) \frac{\hat w_{y,z}}{\rho(y, z)^{d+\alpha}}\mu_y\mu_z,\end{aligned}$$

where $\hat L^{R,R}$ is the generator associated with $(\hat D^{x_0,R,R}, \hat{\mathscr F}^{x_0,R,R})$, i.e.,

$$\hat L^{R,R} f(x) = \sum_{y\in V:\rho(x,y)\leqslant R} (f(y) - f(x)) \frac{\hat w_{x,y}}{\rho(x, y)^{d+\alpha}}\mu_y.$$

Therefore,

$$\begin{aligned}Q'(t) &= -\sum_{y\in V}(\log f_t(y) + 1)\hat L^{R,R} f_t(y)\mu_y\\ &= \frac12 \sum_{y,z\in V:\rho(y,z)\leqslant R} \big(f_t(y) - f_t(z)\big)\big(\log f_t(y) - \log f_t(z)\big) \frac{\hat w_{y,z}}{\rho(y, z)^{d+\alpha}}\mu_y\mu_z \geqslant 0.\end{aligned}$$

In particular, $Q(\cdot)$ is a non-decreasing function on $(0, \infty)$.

On the other hand, for all $\tilde R_0 < R < r_G$, by the Cauchy-Schwarz inequality,

$$M'(t) = \sum_{y\in V} \rho(x, y)\hat L^{R,R} f_t(y)\mu_y$$

$$
\begin{aligned}
&= -\frac{1}{2}\sum_{y,z\in V:\rho(y,z)\leqslant R}\big(\rho(x,y)-\rho(x,z)\big)\big(f_t(y)-f_t(z)\big)\frac{\hat w_{y,z}}{\rho(y,z)^{d+\alpha}}\mu_y\mu_z\\
&\leqslant \left(\frac{1}{4}\sum_{y,z\in V:\rho(y,z)\leqslant R}\big(\rho(x,y)-\rho(x,z)\big)^2\big(f_t(y)+f_t(z)\big)\frac{\hat w_{y,z}}{\rho(y,z)^{d+\alpha}}\mu_y\mu_z\right)^{1/2}\\
&\quad\times\left(\sum_{y,z\in V:\rho(y,z)\leqslant R}\frac{(f_t(y)-f_t(z))^2}{f_t(y)+f_t(z)}\frac{\hat w_{y,z}}{\rho(y,z)^{d+\alpha}}\mu_y\mu_z\right)^{1/2}\\
&\leqslant \left(\frac{c_M}{2}\sup_{z\in V}\sum_{y\in V:\rho(y,z)\leqslant R}\frac{\hat w_{y,z}}{\rho(y,z)^{d+\alpha-2}}\right)^{1/2}\\
&\quad\times\left(\sum_{y,z\in V:\rho(y,z)\leqslant R}\frac{(f_t(y)-f_t(z))^2}{f_t(y)+f_t(z)}\frac{\hat w_{y,z}}{\rho(y,z)^{d+\alpha}}\mu_y\mu_z\right)^{1/2}\\
&\leqslant c_3R^{1-\alpha/2}\left(\sum_{y,z\in V:\rho(y,z)\leqslant R}\frac{(f_t(y)-f_t(z))^2}{f_t(y)+f_t(z)}\frac{\hat w_{y,z}}{\rho(y,z)^{d+\alpha}}\mu_y\mu_z\right)^{1/2},
\end{aligned}
$$

where the equality above follows from the fact

$$\sum_{y,z\in V:\rho(y,z)\leqslant R}|f_t(y)-f_t(z)|\frac{\hat w_{y,z}}{\rho(y,z)^{d+\alpha-1}}<\infty,$$

thank to (2.25) again, in the second inequality we used (2.1) and the fact that $\sum_{z\in V}f_t(z)\mu_z\leqslant 1$ for all $t>0$, and in the last inequality we have used (2.26).

Noting that

$$\frac{(s-t)^2}{s+t}\leqslant (s-t)\big(\log s-\log t\big),\quad s,t>0,$$

we have

$$
\begin{aligned}
&\sum_{y,z\in V:\rho(y,z)\leqslant R}\frac{(f_t(y)-f_t(z))^2}{f_t(y)+f_t(z)}\frac{\hat w_{y,z}}{\rho(y,z)^{d+\alpha}}\mu_y\mu_z\\
&\leqslant \sum_{y,z\in V:\rho(y,z)\leqslant R}\big(f_t(y)-f_t(z)\big)\big(\log f_t(y)-\log f_t(z)\big)\frac{\hat w_{y,z}}{\rho(y,z)^{d+\alpha}}\mu_y\mu_z=2Q'(t).
\end{aligned}
$$

Hence, combining all the estimates above, we arrive at that for all $\tilde R_0<R<r_G$,

$$M'(t)\leqslant \sqrt{2}c_3R^{1-\alpha/2}Q'(t)^{1/2},\ \ \forall\, t>0. \tag{2.28}$$

Step (2): (2.27) yields that for all $\tilde R_0<R<r_G$ and $2R^{\theta_0\alpha}\leqslant t\leqslant R^\alpha$,

$$Q(t)\geqslant -\left(\sum_{y\in V}f_t(y)\right)\log(c_2t^{-d/\alpha})=\frac{d}{\alpha}\log t-c_4,$$

where $c_4>0$ and the conservativeness of $\hat X^{R,R}$ was used in the equality above. Define

$$K(t)=d^{-1}\Big(Q(t)+c_4-\frac{d}{\alpha}\log t\Big),\quad t>0.$$

Obviously, $K(t)\geqslant 0$ for all $t\in[2R^{\theta_0\alpha},R^\alpha]$, and

$$Q'(t)=dK'(t)+\frac{d}{\alpha t},\quad t>0. \tag{2.29}$$

Set $T_0(R) := 0\vee\sup\{t < 2R^{\theta_0\alpha} : K(t) < 0\}$. It is easy to see that $K(t) \geqslant 0$ for all $t \in [T_0(R), R^\alpha]$ and $T_0(R) \leqslant 2R^{\theta_0\alpha}$. By (2.28) and (2.29), we have for all $t \in [T_0(R), R^\alpha]$,

$$\begin{aligned} M(t) &= M(T_0(R)) + \int_{T_0(R)}^t M'(s)\,ds \leqslant M(T_0(R)) + \sqrt{2}c_3R^{1-\alpha/2}\int_{T_0(R)}^t Q'(s)^{1/2}\,ds \\ &= M(T_0(R)) + \sqrt{2}c_3R^{1-\alpha/2}\int_{T_0(R)}^t \Big(dK'(s) + \frac{d}{\alpha s}\Big)^{1/2}\,ds. \end{aligned} \tag{2.30}$$

Note that, by the mean-value theorem, for every $a \in \mathbb{R}$ and $b > 0$ with $a + b \geqslant 0$,

$$(a + b)^{1/2} \leqslant b^{1/2} + a/(2b^{1/2}). \tag{2.31}$$

Then, applying (2.31) in the second term of the right hand side of (2.30) with $a = K'(s)$ and $b = \frac{1}{\alpha s}$, we obtain that for all $t \in [T_0(R), R^\alpha]$,

$$\begin{aligned} M(t) &\leqslant M(T_0(R)) + c_4R^{1-\alpha/2}\int_{T_0(R)}^t s^{-1/2}\,ds + c_5R^{1-\alpha/2}\int_{T_0(R)}^t s^{1/2}K'(s)\,ds \\ &\leqslant M(T_0(R)) + c_6R^{1-\alpha/2}t^{1/2} + c_5R^{1-\alpha/2}\int_{T_0(R)}^t \left[\big(s^{1/2}K(s)\big)' - \frac{s^{-1/2}K(s)}{2}\right] ds \\ &\leqslant M(T_0(R)) + c_6R^{1-\alpha/2}t^{1/2} + c_5R^{1-\alpha/2}t^{1/2}K(t), \end{aligned} \tag{2.32}$$

where the last inequality we used the fact that $K(t) \geqslant 0$ for all $t \in [T_0(R), R^\alpha]$.

Furthermore, suppose that $T_0(R) > 0$. Since $Q'(t) \geqslant 0$, by (2.28) and the Cauchy-Schwarz inequality, we have

$$\begin{aligned} M(T_0(R)) &= \int_0^{T_0(R)} M'(s)\,ds \leqslant \sqrt{2}c_3R^{1-\alpha/2}\int_0^{T_0(R)} Q'(s)^{1/2}\,ds \\ &\leqslant \sqrt{2}c_3R^{1-\alpha/2}T_0(R)^{1/2}\left(\int_0^{T_0(R)} Q'(s)\,ds\right)^{1/2} \\ &\leqslant c_7R^{1-\alpha(1-\theta_0)/2}\big(Q(T_0(R)) - (Q(0)\wedge 0)\big)^{1/2}, \end{aligned}$$

where in the last inequality we have used the fact that $T_0(R) \leqslant 2R^{\theta_0\alpha}$. By the definition of $T_0(R)$, it holds that $K(T_0(R)) = 0$, and so $Q(T_0(R)) = (d/\alpha)\log T_0(R) - c_4 \leqslant c_8(1 + \log R)$, where we have used again $T_0(R) \leqslant 2R^{\theta_0\alpha}$. On the other hand, $Q(0) = \lim_{t\to 0} Q(t) = \log \mu_x \geqslant -\log c_M$. Thus, we can find $R_1 \geqslant 1$ large enough such that for all $R > R_1$ and $t \in [R^{\theta'\alpha}, R^\alpha]$,

$$\begin{aligned} M(T_0(R)) &\leqslant c_9R^{1-\alpha(1-\theta_0)/2}(1 + \log R)^{1/2} = c_9R^{1-\alpha/2}R^{\theta_0\alpha/2}(1 + \log R)^{1/2} \\ &\leqslant c_9R^{1-\alpha/2}R^{\theta'\alpha/2} \leqslant c_9R^{1-\alpha/2}t^{1/2}, \end{aligned}$$

where in the second inequality we used the fact that $\theta_0 \in (\theta, \theta')$, and the last inequality is due to $t \geqslant R^{\theta'\alpha}$. Note that $M(0) = 0$, so the above estimate still holds when $T_0(R) = 0$.

Therefore, combining this with (2.32), we arrive at that for all $t \in [R^{\theta'\alpha}, R^\alpha]$,

$$M(t) \leqslant c_{10}R^{1-\alpha/2}t^{1/2}\big(1 + K(t)\big). \tag{2.33}$$

Step (3): Note that $s(\log s+t) \geqslant -e^{-1-t}$ for all $s > 0$ and $t \in \mathbb{R}$. Then, for every $0 < a \leqslant 2$, $b \in \mathbb{R}$ and $t > 0$,

$$\begin{aligned} -Q(t) + aM(t) + b &= \sum_{y\in V} f_t(y)\big(\log f_t(y) + a\rho(x, y) + b\big)\mu_y \\ &\geqslant -\sum_{y\in V} \exp\big(-1 - a\rho(x, y) - b\big)\mu_y \geqslant -c_{11}e^{-b}a^{-d}, \end{aligned} \tag{2.34}$$

where the equality above follows from the conservativeness of $X^{R,R}$, and in the last inequality we used the fact that

$$\sum_{y\in V} e^{-a\rho(x,y)}\mu_y \leqslant c_M + \sum_{k=1}^{\infty} \sum_{y\in B(x,2^k)\setminus B(x,2^{k-1})} e^{-a2^{k-1}}\mu_y \leqslant c_M + c_G\sum_{k=1}^{\infty} 2^{dk}e^{-a2^{k-1}} \leqslant Ca^{-d}$$

for all $0 < a \leqslant 2$ (see [6, line 6–7 in p. 3056]).

According to (2.27), we could find $R_1 > \tilde{R}_0$ large enough such that for all $R_1 < R < r_G$ and $t \in [R^{\theta'\alpha}, R^\alpha]$,

$$\begin{aligned} M(t) &= \sum_{y\in V}\rho(x,y)f_t(y)\mu_y \geqslant \sum_{y\in V:\rho(x,y)>0} f_t(y)\mu_y = 1 - \mathbb{P}_x\big(\hat{X}_t^{R,R} = x\big) \\ &\geqslant 1 - c_2t^{-d/\alpha} \geqslant 1 - c_2R^{-\theta' d} > 1/2. \end{aligned}$$

Then, choosing $a = 1/M(t)$ and $e^b = M(t)^d = a^{-d}$ in (2.34), we have $-Q(t) + 1 + d\log M(t) \geqslant -c_{11}$, which implies that for all $R_1 < R < r_G$ and $t \in [R^{\theta'\alpha}, R^\alpha]$, $M(t) \geqslant c_{12}\exp(Q(t)/d)$. This along with the definition of $K(t)$ yields that

$$M(t) \geqslant c_{12}\exp(Q(t)/d) \geqslant c_{13}t^{1/\alpha}e^{K(t)}. \tag{2.35}$$

Combining (2.33) with (2.35), we obtain that for all $t \in [R^{\theta'\alpha}, R^\alpha]$, $e^{K(t)} \leqslant c_{14}R^{1-\alpha/2}\big(1 + K(t)\big)t^{1/2-1/\alpha}$, which is equivalent to

$$K(t) \leqslant c_{15}\left[1 + \log\left(\frac{R^\alpha}{t}\right) + \log(1 + K(t))\right].$$

This implies that for all $R_1 < R < r_G$ and $t \in [R^{\theta'\alpha}, R^\alpha]$,

$$K(t) \leqslant c_{16}\left[1 + \log\left(\frac{R^\alpha}{t}\right)\right].$$

The inequality above along with (2.33) further gives us that for all $R_1 < R < r_G$ and $t \in [R^{\theta'\alpha}, R^\alpha]$,

$$M(t) \leqslant c_{17}R^{1-\alpha/2}t^{1/2}\left[1 + \log\left(\frac{R^\alpha}{t}\right)\right] \leqslant c_{18}R\left(\frac{t}{R^\alpha}\right)^{1/2}\left[1 + \log\left(\frac{R^\alpha}{t}\right)\right].$$

The proof is complete. □

3. Stable-like processes on graphs

Let $(D, \mathscr{F})$ be a regular symmetric Dirichlet form on $L^2(V;\mu)$ given in the beginning of Section 2. In particular, we assume that (2.3) holds. Let $X := ((X_t)_{t\geqslant 0}, (\mathbb{P}_x)_{x\in V})$ be the associated symmetric Hunt process associated with $(D, \mathscr{F})$.

3.1. Estimates of exit time.

In order to get estimates of exit time for the process X, we will make full use of results in the previous section. We still adopt notations as before. Fix $x_0 \in V$ and $R \geqslant 1$. According to the definition of $(\hat{D}^{x_0,R}, \hat{\mathscr{F}}^{x_0,R})$, we have

$$\mathbb{P}_{x_0}\big(\tau_{B(x_0,R)} \leqslant t\big) = \mathbb{P}_{x_0}\big(\hat{\tau}^R_{B(x_0,R)} \leqslant t\big), \tag{3.1}$$

where $\tau_A := \inf\{t > 0 : X_t \notin A\}$ and $\hat{\tau}^R_A := \inf\{t \geqslant 0 : \hat{X}^R_t \notin A\}$ for any subset $A \subseteq V$.

In the following, we denote by $(\hat{P}_t^{R,B(x_0,R)})_{t\geqslant 0}$ and $(\hat{P}_t^{R,R,B(x_0,R)})_{t\geqslant 0}$ Dirichlet semigroups of the processes $\hat{X}^R$ and $\hat{X}^{R,R}$ exiting $B(x_0,R)$, respectively. Let $\hat{\tau}^{R,R}_A = \inf\{t \geqslant 0 : \hat{X}^{R,R}_t \notin A\}$ for any $A \subseteq V$.

Lemma 3.1. *For any $f \in L^2(V;\mu)$, $t > 0$ and $x \in B(x_0,R)$,*

$$|\hat{P}_t^{R,R,B(x_0,R)}f(x) - \hat{P}_t^{R,B(x_0,R)}f(x)| \leqslant C_1t\left(\sup_{y\in B(x_0,R)} J(y,R)\right)\left(\sup_{z\in B(x_0,R)}|f(z)|\right), \tag{3.2}$$

where C_1 is a positive constant independent of R and x_0, and

$$J(y,R)=\sum_{z\in V:\rho(y,z)>R}\frac{w_{y,z}}{\rho(y,z)^{d+\alpha}}\mu_z,\quad y\in B(x_0,R). \tag{3.3}$$

In particular, it holds that for any $t>0$ and $x\in B(x_0,R)$,

$$\left|\mathbb{P}_x\big(\hat\tau^{R,R}_{B(x_0,R)}\leqslant t\big)-\mathbb{P}_x\big(\hat\tau^{R}_{B(x_0,R)}\leqslant t\big)\right|\leqslant C_1t\sup_{y\in B(x_0,R)}J(y,R). \tag{3.4}$$

Proof. Let $T_R^R=\inf\{t>0:\rho(\hat X^R_{t-},\hat X^R_t)>R\}$. By (2.20), $\sup_{y\in V}\sum_{z\in V:\rho(z,y)>R}\frac{\hat w_{z,y}}{\rho(z,y)^{d+\alpha}}\mu_z<\infty$. Then, by Meyer's construction of $\hat X^R$ (see [10, Section 3.1]), $\hat X^R_t=\hat X^{R,R}_t$ if $t<T_R^R$. Hence, for any $f\in L^2(V;\mu)$,

$$\begin{aligned}&\left|\hat P_t^{R,R,B(x_0,R)}f(x)-\hat P_t^{R,B(x_0,R)}f(x)\right|\\&=\left|\mathbb{E}_x(f(\hat X^R_t):t\leqslant\hat\tau^R_{B(x_0,R)})-\mathbb{E}_x(f(\hat X^{R,R}_t):t\leqslant\hat\tau^{R,R}_{B(x_0,R)})\right|\\&\leqslant\sup_{z\in B(x_0,R)}|f(z)|\Big[\mathbb{P}_x\big(T^R_R\leqslant t\leqslant\hat\tau^R_{B(x_0,R)}\big)+\mathbb{P}_x\big(T^R_R\leqslant t\leqslant\hat\tau^{R,R}_{B(x_0,R)}\big)\Big]\\&\leqslant 2\left(\sup_{z\in B(x_0,R)}|f(z)|\right)\mathbb{P}_x\big(T^R_R\leqslant t,\hat X^{R,R}_s\in B(x_0,R)\text{ for all }s\in[0,T^R_R]\big).\end{aligned}$$

According to [10, Lemma 3.1(a)],

$$\mathbb{P}_x\Big(T^R_R\in dt\big|\mathscr{F}^{\hat X^{R,R}}\Big)=\hat J(\hat X^{R,R}_t,R)\exp\left(-\int_0^t\hat J(\hat X^{R,R}_s,R)\,ds\right)dt,$$

where $\mathscr{F}^{\hat X^{R,R}}$ denotes the σ-algebra generated by $\hat X^{R,R}$, and

$$\hat J(y,R)=\sum_{z\in V:\rho(y,z)>R}\frac{\hat w_{y,z}}{\rho(y,z)^{d+\alpha}}\mu_z,\quad y\in B(x_0,R).$$

In particular, by the definition of $\hat w_{x,y}$, $J(y,R)=\hat J(y,R)$ for all $y\in B(x_0,R)$. Therefore,

$$\begin{aligned}&\mathbb{P}_x\Big(T^R_R\leqslant t,\hat X^{R,R}_s\in B(x_0,R)\text{ for all }s\in[0,T^R_R]\Big)\\&\leqslant\mathbb{E}_x\left[\int_0^tJ(\hat X^{R,R}_r,R)\exp\left(-\int_0^rJ(\hat X^{R,R}_s,R)\,ds\right)\mathbb{1}_{\{\hat X^{R,R}_s\in B(x_0,R)\text{ for all }s\in[0,r]\}}\,dr\right]\\&\leqslant c_1t\sup_{y\in B(x_0,R)}J(y,R).\end{aligned}$$

Combining all the estimates above, we can obtain (3.2). (3.4) is a direct consequence of (3.2) by taking $f\equiv1$ on $B(x_0,R)$. □

Proposition 3.2. *Assume that for some $\theta\in(0,1)$, there exists $R_0\geqslant1$ such that for every $R_0<R<r_G$ and $R^\theta\leqslant r\leqslant R$, (2.21), (2.22) and (2.23) as well as*

$$\sup_{x\in B(x_0,R)}\sum_{y\in V:\rho(x,y)>R}\frac{w_{x,y}}{\rho(x,y)^{d+\alpha}}\leqslant C_1R^{-\alpha} \tag{3.5}$$

hold, where $C_1>0$ is a constant independent of x_0, R_0, R, r and r_G. Then

(i) *for any $\theta'\in(\theta,1)$, there is a constant $R_1\geqslant1$ (which only depends on θ, θ', R_0 and r_G) such that for every $R_1<R<r_G$,*

$$\mathbb{P}_{x_0}\big(\tau_{B(x_0,R)}\leqslant t\big)\leqslant C_2\left(\frac{t}{R^\alpha}\right)^{1/2}\left[1\vee\log\left(\frac{R^\alpha}{t}\right)\right],\quad t\geqslant R^{\theta'\alpha}, \tag{3.6}$$

where C_2 is a positive constant independent of x_0, R_1, R, t and r_G.

(ii) *for any $\varepsilon > 0$, there is a constant $R_2 \geqslant 1$ (depending on θ, R_0, r_G and ε) such that for all $R_2 < R < r_G$,*

$$\mathbb{P}_{x_0}\big(\tau_{B(x_0,R)} \leqslant t\big) \leqslant \varepsilon + \frac{C_3(\varepsilon)t}{R^\alpha}, \quad t > 0, \tag{3.7}$$

where $C_3(\varepsilon)$ is a positive constant independent of x_0, R_1, R, t and r_G. In particular, the process X is conservative.

Proof. **Step (1):** It immediately follows from (3.5) that

$$\sup_{y \in B(x_0,R)} J(y,R) \leqslant c_1 R^{-\alpha}, \tag{3.8}$$

where $J(y,R)$ is defined by (3.3).

Since (2.21), (2.22) and (2.23) are true, by (2.24), for any $\theta' \in (\theta,1)$, there is a constant $\tilde{R}_1 \geqslant 1$ such that for all $R_1 < R < r_G$ and $x \in V$,

$$\mathbb{E}_x\big[\rho(\hat{X}_t^{R,R},x)\big] \leqslant c_2 R\Big(\frac{t}{R^\alpha}\Big)^{1/2}\left[1+\log\left(\frac{R^\alpha}{t}\right)\right], \quad \forall\, R^{\theta'\alpha} \leqslant t \leqslant R^\alpha.$$

Hence, by the Markov inequality, for all $x \in V$ and $R^{\theta'\alpha} \leqslant t \leqslant R^\alpha/2$,

$$\sup_{s \in [t,2t]} \mathbb{P}_x\Big(\rho\big(\hat{X}_s^{R,R},x\big) > \frac{R}{2}\Big) \leqslant c_3\Big(\frac{t}{R^\alpha}\Big)^{1/2}\left[1+\log\left(\frac{R^\alpha}{t}\right)\right].$$

Therefore, for all $R^{\theta'\alpha} \leqslant t \leqslant R^\alpha/2$,

$$\begin{aligned}
\mathbb{P}_{x_0}\big(\hat{\tau}^{R,R}_{B(x_0,R)} \leqslant t\big) &\leqslant \mathbb{P}_{x_0}\Big(\hat{\tau}^{R,R}_{B(x_0,R)} \leqslant t; \rho\big(\hat{X}^{R,R}_{2t},x_0\big) \leqslant \frac{R}{2}\Big) + \mathbb{P}_{x_0}\Big(\rho\big(\hat{X}^{R,R}_{2t},x_0\big) > \frac{R}{2}\Big)\\
&\leqslant \mathbb{E}_{x_0}\left[\mathbb{1}_{\{\hat{\tau}^{R,R}_{B(x_0,R)} \leqslant t\}} \mathbb{P}_{\hat{X}^{R,R}_{\hat{\tau}^{R,R}_{B(x_0,R)}}}\Big(\rho\big(\hat{X}^{R,R}_{2t-\tau^{R,R}_{B(x_0,R)}},\hat{X}_0^{R,R}\big) > \frac{R}{2}\Big)\right]\\
&\quad + c_3\Big(\frac{t}{R^\alpha}\Big)^{1/2}\left[1+\log\left(\frac{R^\alpha}{t}\right)\right]\\
&\leqslant \sup_{y \in V}\sup_{s \in [t,2t]} \mathbb{P}_y\Big(\rho\big(\hat{X}^{R,R}_s,y\big) > \frac{R}{2}\Big) + c_3\Big(\frac{t}{R^\alpha}\Big)^{1/2}\left[1+\log\left(\frac{R^\alpha}{t}\right)\right]\\
&\leqslant 2c_3\Big(\frac{t}{R^\alpha}\Big)^{1/2}\left[1+\log\left(\frac{R^\alpha}{t}\right)\right].
\end{aligned}$$

Combining this with (3.1), (3.4) and (3.8) yields that for all $\tilde{R}_1 < R < r_G$ and $R^{\theta'\alpha} \leqslant t \leqslant R^\alpha/2$,

$$\mathbb{P}_{x_0}\big(\tau_{B(x_0,R)} \leqslant t\big) \leqslant 2c_3\Big(\frac{t}{R^\alpha}\Big)^{1/2}\left[1+\log\left(\frac{R^\alpha}{t}\right)\right] + \frac{c_4 t}{R^\alpha} \leqslant c_5\Big(\frac{t}{R^\alpha}\Big)^{1/2}\left[1\vee\log\left(\frac{R^\alpha}{t}\right)\right].$$

Thus, (3.6) has been verified for all $R^{\theta'\alpha} \leqslant t \leqslant R^\alpha/2$. When $t > R^\alpha/2$, it holds that

$$\mathbb{P}_{x_0}\big(\tau_{B(x_0,R)} \leqslant t\big) \leqslant 1 \leqslant \Big(\frac{2t}{R^\alpha}\Big)^{1/2}\left[1\vee\log\left(\frac{R^\alpha}{t}\right)\right].$$

Hence we prove (3.6).

Step (2): Fix $\theta' \in (\theta,1)$. By (3.6) and Young's inequality, there is a constant $\tilde{R}_1 \geqslant 1$ such that for every $\tilde{R}_1 < R < r_G$, $t \geqslant R^{\theta'\alpha}$ and $\varepsilon > 0$, $\mathbb{P}_{x_0}\big(\tau_{B(x_0,R)} \leqslant t\big) \leqslant 2^{-1}\varepsilon + c_6(\varepsilon)tR^{-\alpha}$. If $0 < t \leqslant R^{\theta'\alpha}$, then, taking $\tilde{R}_2(\varepsilon) \geqslant \tilde{R}_1$ large enough, we obtain that for all $\tilde{R}_2(\varepsilon) \leqslant R < r_G$, $\mathbb{P}_{x_0}\big(\tau_{B(x_0,R)} \leqslant t\big) \leqslant \mathbb{P}_{x_0}\big(\tau_{B(x_0,R)} \leqslant R^{\theta'\alpha}\big) \leqslant 2^{-1}\varepsilon + c_6(\varepsilon)R^{-(1-\theta')\alpha} \leqslant \varepsilon$. Combining both estimates above together, we know that for all $\tilde{R}_2(\varepsilon) < R < r_G$ and $t > 0$, $\mathbb{P}_{x_0}\big(\tau_{B(x_0,R)} \leqslant t\big) \leqslant \varepsilon + c_7(\varepsilon)tR^{-\alpha}$, which implies that (3.7) holds. □

We are now in a position to present the main result in this subsection. For this, we need the following assumption on $\{w_{x,y} : x, y \in V\}$, which is regarded as the summary of all assumptions in the statements before. For any $x, z \in V$ and $r > 0$, denote $B_z^w(x,r) := \{u \in B(x,r) : w_{u,z} > 0\}$. In particular, $B_x^w(x,r) = B^w(x,r)$.

Assumption (Exi.) *Suppose that for some fixed $\theta \in (0,1)$ and $0 \in V$, there exists a constant $R_0 \geqslant 1$ such that the following hold.*

(i) *For every $R_0 < R < r_G$ and $R^\theta/2 \leqslant r \leqslant 2R$,*

$$\sup_{x\in B(0,6R)} \sum_{y\in V:\rho(x,y)\leqslant r} \frac{w_{x,y}}{\rho(x,y)^{d+\alpha-2}} \leqslant C_1 r^{2-\alpha}, \tag{3.9}$$

$$\mu(B_z^w(x,r)) \geqslant c_0\mu(B(x,r)), \quad x, z \in B(0,6R) \tag{3.10}$$

and

$$\sup_{x\in B(0,6R)} \sum_{y\in B^w(x,c_*r)} w_{x,y}^{-1} \leqslant C_1 r^d, \tag{3.11}$$

where $c_0 > 1/2$ is independent of R_0, R, r, x and z, and $c_* := 8c_G^{2/d}$.

(ii) *For every $R_0 < R < r_G$ and $r \geqslant R^\theta/2$,*

$$\sup_{x\in B(0,6R)} \sum_{y\in V:\rho(x,y)>r} \frac{w_{x,y}}{\rho(x,y)^{d+\alpha}} \leqslant C_1 r^{-\alpha}. \tag{3.12}$$

Here C_1 is a positive constant independent of R_0, R and r_G.

Lemma 3.3. *Let c_* be the constant in Assumption* (**Exi.**)(i). *Under* (3.10) *and* (3.11), *for every $R_0 < R < r_G/(2c_*)$ and $R^\theta/2 \leqslant r \leqslant 2R$,*

$$\inf_{x\in B(0,6R)} \sum_{y\in V:\rho(x,y)>3r} \frac{w_{x,y}}{\rho(x,y)^{d+\alpha}} \geqslant C_2 r^{-\alpha}, \tag{3.13}$$

where $C_2 > 0$ is independent of R_0, R and r_G.

Proof. Noting that $c_* > 4$, for every $x \in V$ and $1 \leqslant r < r_G/c_*$, we have

$$\sum_{y\in V:3r<\rho(x,y)\leqslant c_*r, w_{x,y}>0} \mu_y \geqslant \mu(B^w(x,c_*r)) - \mu(B(x,4r)) \geqslant c_0 c_G^{-1}(c_*r)^d - c_G(4r)^d \geqslant c_1 r^d,$$

where we have used (2.2) and (3.10).

On the other hand, for every $R_0 < R < r_G/(2c_*)$, $x \in B(0,6R)$ and $R^\theta/2 \leqslant r \leqslant 2R$,

$$\begin{aligned}\sum_{y\in V:3r<\rho(x,y)\leqslant c_*r, w_{x,y}>0} \mu_y &\leqslant \Big(\sum_{y\in B^w(x,c^*r)} w_{x,y}^{-1}\mu_y\Big)^{1/2}\Big(\sum_{y\in V:3r<\rho(x,y)\leqslant c_*r} w_{x,y}\mu_y\Big)^{1/2}\\ &\leqslant c_2 r^{d/2}\Big(\sum_{y\in V:3r<\rho(x,y)\leqslant c_*r} w_{x,y}\Big)^{1/2},\end{aligned}$$

where in the first inequality we have applied the Cauchy-Schwarz inequality, and we used (3.11) in the last inequality.

Combining both estimates above together yields that for every $R_0 < R < r_G/(2c_*)$, $x \in B(0,6R)$ and $R^\theta/2 \leqslant r \leqslant 2R$, $\sum_{y\in V:3r<\rho(x,y)\leqslant c_*r} w_{x,y} \geqslant c_3 r^d$, and so

$$\sum_{y\in V:\rho(x,y)>3r} \frac{w_{x,y}}{\rho(x,y)^{d+\alpha}} \geqslant \sum_{y\in V:3r<\rho(x,y)\leqslant c_*r} \frac{w_{x,y}}{\rho(x,y)^{d+\alpha}} \geqslant (c_*r)^{-d-\alpha} \sum_{y\in V:3r<\rho(x,y)\leqslant c_*r} w_{x,y} \geqslant c_4 r^{-\alpha}.$$

Thus, (3.13) is proved. □

Theorem 3.4. *If Assumption* **(Exi.)** *holds with some constant $\theta \in (0,1)$, then, for every $\theta' \in (\theta,1)$, there exist constants $\delta \in (\theta,1)$ and $R_1 \geqslant 1$ such that for all $R_1 < R < r_G/(2c_*)$ and $R^\delta \leqslant r \leqslant R$,*

(1)

$$\sup_{x\in B(0,2R)} \mathbb{P}_x\big(\tau_{B(x,r)} \leqslant C_0 r^\alpha\big) \leqslant \frac{1}{4}, \tag{3.14}$$

where $C_0 > 0$ is a constant independent of R_0, R_1, R and r.

(2)

$$\sup_{x\in B(0,2R)} \mathbb{P}_x\big(\tau_{B(x,r)} \leqslant t\big) \leqslant C_1\Big(\frac{t}{r^\alpha}\Big)^{1/2}\Big[1\vee\log\Big(\frac{r^\alpha}{t}\Big)\Big], \quad t \geqslant r^{\theta'\alpha}, \tag{3.15}$$

and

$$C_2 r^\alpha \leqslant \inf_{x\in B(0,2R)} \mathbb{E}_x\big[\tau_{B(x,r)}\big] \leqslant \sup_{x\in B(0,2R)} \mathbb{E}_x\big[\tau_{B(x,r)}\big] \leqslant C_1 r^\alpha, \tag{3.16}$$

where C_1, C_2 are positive constants independent of R_0, R_1, R, r, t and r_G.

Proof. Suppose that Assumption **(Exi.)** holds with some $\theta \in (0,1)$ and $R_0 \geqslant 1$. Then, for any $\theta < \theta_1 < \theta' < 1$, $R_0 < R < r_G$ and $R^\delta \leqslant s \leqslant R$ with $\delta = \theta/\theta_1$, we know that (2.21), (2.23) and (3.5) hold uniformly (that is, they hold with uniform constants) for every $s^{\theta_1} \leqslant r \leqslant s$ and $x_0 \in B(0,2R)$. Hence, according to (3.6) and (3.7), we obtain that for every $\theta' \in (\theta,1)$, there exists a constant $R_1 \geqslant R_0$ such that for each $R_1 < R < r_G$ and $R^\delta \leqslant r \leqslant R$, (3.15) and

$$\sup_{x\in B(0,2R)} \mathbb{P}_x\big(\tau_{B(x,r)} \leqslant t\big) \leqslant \frac{1}{8} + \frac{c_1 t}{r^\alpha}, \quad \forall\, t > 0 \tag{3.17}$$

hold true. In particular, taking $t = (8c_1)^{-1} r^\alpha$ in (3.17), we get (3.14) immediately.

Let C_0 be the constant in (3.14). For any $R > R_1$, $x \in B(0,2R)$ and $R^\delta \leqslant r \leqslant R$, we have

$$\begin{aligned}\mathbb{E}_x[\tau_{B(x,r)}] &= \int_0^\infty \mathbb{P}_x(\tau_{B(x,r)} > s)\, ds \geqslant \int_0^{C_0 r^\alpha} \mathbb{P}_x(\tau_{B(x,r)} > s)\, ds\\ &\geqslant C_0 r^\alpha \mathbb{P}_x(\tau_{B(x,r)} > C_0 r^\alpha) \geqslant \frac{3C_0 r^\alpha}{4}.\end{aligned}$$

This gives us the first inequality in (3.16). On the other hand, let c_* be the constant in Assumption (**Exi.**)(i). By the Lévy system (see [24, Appendix A]), for any $R_1 < R < r_G/(2c_*)$, $x \in B(0,2R)$ and $R^\delta \leqslant r \leqslant R$,

$$\begin{aligned}1 \geqslant \mathbb{P}_x\big(X_{\tau_{B(x,r)}} \notin B(x,2r)\big) &= \mathbb{E}_x\left[\int_0^{\tau_{B(x,r)}} \sum_{y\in V:\rho(x,y)>2r} \frac{w_{X_s,y}}{\rho(X_s,y)^{d+\alpha}}\mu_y\, ds\right]\\ &\geqslant c_M^{-1}\mathbb{E}_x\left[\int_0^{\tau_{B(x,r)}} \sum_{y\in V:\rho(y,X_s)>3r} \frac{w_{X_s,y}}{\rho(X_s,y)^{d+\alpha}}\, ds\right]\\ &\geqslant c_M^{-1}\left(\inf_{v\in B(0,2R+r)} \sum_{y\in V:\rho(y,v)>3r} \frac{w_{v,y}}{\rho(v,y)^{d+\alpha}}\right)\mathbb{E}_x[\tau_{B(x,r)}] \geqslant c_2 r^{-\alpha}\mathbb{E}_x[\tau_{B(x,r)}],\end{aligned}$$

where in the last inequality we have used (3.13), also thanks to the fact that $\delta = \theta/\theta_1 > \theta$. Thus, we also prove the third inequality in (3.16). □

When $\alpha \in (0,1)$, we can obtain a probability estimate such like (3.7) for the exit time in a more direct way under the following assumption.

Assumption (Exi.') *Suppose that for some fixed $\theta \in (0,1)$ and $0 \in V$, there exists a constant $R_0 \geqslant 1$ such that*

(i) *for every $R_0 < R < r_G$ and $R^\theta/2 \leqslant r \leqslant 2R$,*

$$\sup_{x\in B(0,6R)} \sum_{y\in V:\rho(x,y)\leqslant r} \frac{w_{x,y}}{\rho(x,y)^{d+\alpha-1}} \leqslant C_1 r^{1-\alpha} \tag{3.18}$$

and (3.11) hold.

(ii) (ii) *in Assumption* **(Exi.)** *is satisfied.*

Here C_1 is a positive constant independent of R_0, R and r_G.

Proposition 3.5. *Under* (3.18) *and* (ii) *in Assumption* **(Exi.)**, *there exists a constant $R_1 > R_0$ such that for all $R_1 < R < r_G$, $x \in B(0, 2R)$, $R^\theta \leqslant r \leqslant R$ and $t > 0$,*

$$\mathbb{P}_x(\tau_{B(x,r)} \leqslant t) \leqslant \frac{C_2 t}{r^\alpha}, \tag{3.19}$$

where $C_2 > 0$ is a constant independent of R_1, R, r, x, t and r_G.

Proof. Fix $x \in B(0, 2R)$. Given $f \in C_b^1([0,\infty))$ with $f(0) = 0$ and $f(u) = 1$ for all $u \geqslant 1$, we set $f_{x,r}(z) = f\left(\frac{\rho(z,x)}{r}\right)$ for any $z \in V$ and $r > 0$. For any $r > 0$,

$$\left\{ f_{x,r}(X_t) - f_{x,r}(X_0) - \int_0^t L f_{x,r}(X_s)\, ds, t \geqslant 0 \right\}$$

is a local martingale. Then, for any $t > 0$ and $x \in V$,

$$\mathbb{P}_x(\tau_{B(x,r)} \leqslant t) \leqslant \mathbb{E}_x f_{x,r}(X_{t\wedge\tau_{B(x,r)}}) = \mathbb{E}_x \left[\int_0^{t\wedge\tau_{B(x,r)}} L f_{x,r}(X_s)\, ds \right] \leqslant t \sup_{z \in B(x,r)} L f_{x,r}(z),$$

where we used the fact that $f_{x,r}(x) = 0$ in the equality above.

Furthermore, for any $x \in V$ and $z \in B(x,r)$,

$$\begin{aligned} L f_{x,r}(z) &= \sum_{y \in V} \big(f_{x,r}(y) - f_{x,r}(z)\big) \frac{w_{y,z}}{\rho(z,y)^{d+\alpha}} \mu_y \\ &= \sum_{y \in V: \rho(y,z) \leqslant r} (f_{x,r}(y) - f_{x,r}(z)) \frac{w_{y,z}}{\rho(y,z)^{d+\alpha}} \mu_y \\ &\quad + \sum_{y \in V: \rho(y,z) > r} (f_{x,r}(y) - f_{x,r}(z)) \frac{w_{y,z}}{\rho(y,z)^{d+\alpha}} \mu_y \\ &\leqslant c_1 \left(r^{-1} \sum_{y \in V: \rho(z,y) \leqslant r} \frac{w_{y,z}}{\rho(y,z)^{d+\alpha-1}} + \sum_{y \in V: \rho(z,y) > r} \frac{w_{y,z}}{\rho(y,z)^{d+\alpha}} \right) =: c_1(I_1(z,r) + I_2(z,r)), \end{aligned}$$

where in the first inequality above we have used $|f_{x,r}(y) - f_{x,r}(z)| \leqslant c_1 r^{-1} \rho(y,z)$. According to (3.18) and (3.12), we can find a constant $R_1 \geqslant 1$ such that for all $R_1 < R < r_G$, $x \in B(0, 2R)$ and $R^\theta \leqslant r \leqslant R$, $\sup_{z \in B(x,r)} \big(I_1(z,r) + I_2(z,r)\big) \leqslant c_2 r^{-\alpha}$.

Combining with all estimates above, we prove the desired assertion. □

3.2. **Hölder regularity.** Let $R_+ := (0,\infty)$ and $Z := (Z_t)_{t \geqslant 0} = (U_t, X_t)_{t \geqslant 0}$ be the time-space process such that $U_t = U_0 + t$ for any $t \geqslant 0$. Denote by $\mathbb{P}_{(s,x)}$ the probability of the process Z starting from $(s,x) \in \mathbb{R}_+ \times V$. For any subset $A \subseteq \mathbb{R}_+ \times V$, define $\tau_A = \inf\{s > 0 : Z_s \in A\}$ and $\sigma_A = \inf\{s > 0 : Z_s \in A\}$. For any $t \geqslant 0$, $x \in V$ and $R \geqslant 1$, let $Q(t,x,R) = (t, t + C_0 R^\alpha) \times B(x,R)$ and $d\nu = ds \times d\mu$, where C_0 is the constant in (3.14). In the following, let c_* be the constant in Assumption (**Exi.**)(i).

Proposition 3.6. *If Assumption* **(Exi.)** *holds with some $\theta \in (0,1)$, then there exist constants $\delta \in (\theta, 1)$ and $R_1 \geqslant 1$ such that for any $R_1 < R < r_G/(2c_*)$, $2R^\delta \leqslant r \leqslant R$, $x \in B(0,2R)$, $t \geqslant 0$ and $A \subseteq Q(t,x,r/2)$ with $\frac{\nu(A)}{\nu(Q(t,x,r/2))} \geqslant 1/2$,*

$$\mathbb{P}_{(t,x)}(\sigma_A < \tau_{Q(t,x,r)}) \geqslant C_1, \tag{3.20}$$

where $C_1 \in (0,1)$ is a constant independent of R_1, R, r, t, x and r_G.

Proof. The proof is based on that of [23, Lemma 4.11] with some slight modifications. We write $Q_r = Q(t,x,r)$ for simplicity. Without loss of generality, we may and can assume that

$\mathbb{P}_{(t,x)}(\sigma_A < \tau_{Q_r}) \leqslant 1/4$; otherwise the conclusion holds trivially. Let $T = \sigma_A \wedge \tau_{Q_r}$ and $A_s = \{y \in V : (s,y) \in A\}$ for all $s > 0$. According to the Lévy system,

$$\begin{aligned}\mathbb{P}_{(t,x)}(\sigma_A < \tau_{Q_r}) &\geqslant \mathbb{E}_{(t,x)}\left(\sum_{s\leqslant T}\mathbb{1}_{\{X_s\neq X_{s-},X_s\in A_s\}}\right) = \mathbb{E}_{(t,x)}\left[\int_0^T \sum_{u\in A_s}\frac{w_{X_s,u}}{\rho(X_s,u)^{d+\alpha}}\mu_u\,ds\right]\\ &\geqslant c_M^{-1}\mathbb{E}_{(t,x)}\left[\int_0^{C_0(r/2)^\alpha}\sum_{u\in A_s}\frac{w_{X_s,u}}{\rho(X_s,u)^{d+\alpha}}\,ds; T\geqslant C_0(r/2)^\alpha\right]\\ &\geqslant c_1 r^{-d-\alpha}\left(\inf_{z\in B(x,r)}\int_0^{C_0(r/2)^\alpha}\sum_{u\in A_s}w_{z,u}\,ds\right)\mathbb{P}_{(t,x)}(T\geqslant C_0(r/2)^\alpha),\end{aligned}$$

where in the last inequality we have used fact that $\rho(u,z) \leqslant 2r$ for every $u, z \in B(x,r)$.

Furthermore, according to Theorem 3.4(1), there exist constants $R_1 \geqslant 1$ and $\delta \in (\theta, 1)$ such that for any $R_1 < R < r_G/(2c_*)$, $R^\delta \leqslant r/2 \leqslant R$ and $x \in B(0, 2R)$,

$$\begin{aligned}\mathbb{P}_{(t,x)}\big(T\geqslant C_0(r/2)^\alpha\big) &= \mathbb{P}_{(t,x)}\big(\sigma_A\wedge\tau_{Q_r}\geqslant C_0(r/2)^\alpha\big)\\ &\geqslant 1-\mathbb{P}_{(t,x)}\big(\sigma_A<\tau_{Q_r}\big)-\mathbb{P}_x\big(\tau_{B(x,r)}\leqslant C_0(r/2)^\alpha\big)\geqslant 1-\frac14-\frac14\geqslant\frac12,\end{aligned}$$

where in the first inequality we have used the fact that

$$\mathbb{P}_{(t,x)}\big(\tau_{Q_r}\leqslant C_0(r/2)^\alpha\big)=\mathbb{P}_x\big(\tau_{B(x,r)}\wedge(C_0r^\alpha)\leqslant C_0(r/2)^\alpha\big)=\mathbb{P}_x\big(\tau_{B(x,r)}\leqslant C_0(r/2)^\alpha\big),$$

and the second inequality follows from (3.14).

On the other hand, let $Q_z^w(t,x,r) := (t + C_0r^\alpha) \times B_z^w(x,r)$. Then, for every $R_1 < R < r_G$, $2R^\delta \leqslant r \leqslant R$, $x \in B(0,2R)$ and $z \in B(x,r)$,

$$\begin{aligned}\nu(A\cap Q_z^w(t,x,r/2)) &= \int_0^{C_0(r/2)^\alpha}\sum_{u\in A_s\cap B_z^w(x,r/2)}\mu_u\,ds\\ &\leqslant \Big(\int_0^{C_0(r/2)^\alpha}\sum_{u\in A_s\cap B_z^w(x,r/2)}w_{z,u}^{-1}\mu_u\,ds\Big)^{1/2}\Big(\int_0^{C_0(r/2)^\alpha}\sum_{u\in A_s}w_{z,u}\mu_u\,ds\Big)^{1/2}\\ &\leqslant c_3r^{\alpha/2}\Big(\sum_{u\in B_z^w(x,r)}w_{z,u}^{-1}\Big)^{1/2}\Big(\int_0^{C_0(r/2)^\alpha}\sum_{u\in A_s}w_{z,u}\,ds\Big)^{1/2}\\ &\leqslant c_3r^{\alpha/2}\Big(\sup_{z\in B(0,3R)}\sum_{u\in B^w(z,2r)}w_{z,u}^{-1}\Big)^{1/2}\Big(\int_0^{C_0(r/2)^\alpha}\sum_{u\in A_s}w_{z,u}\,ds\Big)^{1/2}\\ &\leqslant c_4r^{(d+\alpha)/2}\Big(\int_0^{C_0(r/2)^\alpha}\sum_{u\in A_s}w_{z,u}\,ds\Big)^{1/2},\end{aligned}$$

where in the first inequality we have used the Cauchy-Schwarz inequality, the third inequality is due to the fact that $B_z^w(x,r) \subset B^w(z,2r)$ for all $z \in B(x,r)$, and the last inequality follows from (3.11). Note that, by (3.10) and the assumption that $\frac{\nu(A)}{\nu(Q(t,x,r/2))} \geqslant 1/2$, we have $\nu(A \cap Q_z^w(t,x,r/2)) \geqslant \big(1/2+c_0-1\big)\cdot\nu(Q(t,x,r/2)) \geqslant c_5r^{d+\alpha}$. Combing all estimates above yields that for all $R_1 < R < r_G$, $2R^\delta \leqslant r \leqslant R$, $x \in B(0,2R)$ and $z \in B(x,r)$, $\int_0^{C_0(r/2)^\alpha}\sum_{u\in A_s} w_{z,u}\,ds \geqslant c_6r^{d+\alpha}$. According to all the estimates above, we prove the required assertion. □

We also need the following hitting probability estimate.

Lemma 3.7. *Suppose that Assumption* **(Exi.)** *holds with some $\theta \in (0,1)$. Then there are constants $\delta \in (\theta,1)$ and $R_1 \geqslant 1$ such that for every $R_1 < R < r_G/(2c_*)$, $R^\delta \leqslant r \leqslant R$, $x \in B(0,2R)$, $K > 4r$, $t \geqslant 0$ and $z \in B(x,r/2)$,*

$$\mathbb{P}_x(X_{\tau_{Q(t,x,r)}} \notin B(z,K)) \leqslant C_1\left(\frac{r}{K}\right)^\alpha, \tag{3.21}$$

where $C_1 > 0$ is a positive constant independent of R_0, R_1, r, t, x, z and r_G.

Proof. According to the Lévy system, we know that for every $z \in B(x, r/2)$,

$$\begin{aligned}\mathbb{P}_x(X_{\tau_{Q(t,x,r)}} \notin B(z,K)) &= \mathbb{E}_x\left[\int_0^{\tau_{B(x,r)}} \sum_{y \notin B(z,K)} \frac{w_{X_s,y}}{\rho(X_s,y)^{d+\alpha}}\mu_y\,ds\right] \\ &\leqslant c_1 \sup_{u \in B(x,r)} \left(\sum_{y \in V: \rho(u,y) > K-2r} \frac{w_{u,y}}{\rho(u,y)^{d+\alpha}}\right)\mathbb{E}_x[\tau_{B(x,r)}] \\ &\leqslant c_1 \sup_{u \in B(0,2R)} \left(\sum_{y \in V: \rho(u,y) > K/2} \frac{w_{u,y}}{\rho(u,y)^{d+\alpha}}\right)\mathbb{E}_x[\tau_{B(x,r)}].\end{aligned}$$

Note that $K/2 > 2r \geqslant R^\delta$ and $R^\delta \leqslant r \leqslant R$. Then, by (3.12) and (3.16), we can find a constant $R_1 \geqslant 1$ such that for all $R_1 < R < r_G/(2c_*)$ and $x \in B(0, 2R)$,

$$\sup_{u \in B(0,2R)} \left(\sum_{y \in V: \rho(u,y) > K/2} \frac{w_{u,y}}{\rho(u,y)^{d+\alpha}}\right) \leqslant c_2 K^{-\alpha}$$

and $\mathbb{E}_x[\tau_{B(x,r)}] \leqslant c_3 r^\alpha$. Combining with all the estimates above immediately yields (3.21). □

We say that a measurable function $q(t,x)$ on $[0,\infty) \times V$ is parabolic in an open subset A of $[0,\infty) \times V$, if for every relatively compact open subset A_1 of A, $q(t,x) = \mathbb{E}^{(t,x)} q(Z_{\tau_{A_1}})$ for every $(t,x) \in A_1$.

Let $C_0 > 0$ be the constant in (3.14), and θ be the constant in Assumption **(Exi.)**. Set $Q(t_0, x_0, r) = (t_0, t_0 + C_0 r^\alpha) \times B(x_0, R)$.

Theorem 3.8. *Suppose that Assumption* **(Exi.)** *holds with some $\theta \in (0,1)$, and let c_* be the constant in Assumption* (**Exi**.)(i)*. Then, there are constants $R_1 \geqslant 1$ and $\delta \in (\theta, 1)$ such that for all $R_1 < R < r_G/(2c_*)$, $x_0 \in B(0,R)$, $R^\delta \leqslant r \leqslant R$, $t_0 \geqslant 0$ and parabolic function q on $Q(t_0, x_0, 2r)$,*

$$|q(s,x) - q(t,y)| \leqslant C_1 \|q\|_{\infty,r} \left(\frac{|t-s|^{1/\alpha} + \rho(x,y)}{r}\right)^\beta, \tag{3.22}$$

holds for all $(s,x), (t,y) \in Q(t_0, x_0, r)$ such that $(C_0^{-1}|s-t|)^{1/\alpha} + \rho(x,y) \geqslant 2r^\delta$, where $\|q\|_{\infty,r} = \sup_{(s,x) \in [t_0, t_0 + C_0(2r)^\alpha] \times V} q(s,x)$, *and $C_1 > 0$ and $\beta \in (0,1)$ are constants independent of R_0, R_1, x_0, t_0, R, r, s, t, x, y and r_G.*

Remark 3.9. Note that unlike the case of random walk on the supercritical percolation cluster ([11, Proposition 3.2]), in which the Hölder regularity holds for all points in the parabolic cylinder when r is large enough, in the preset setting we can only obtain the Hölder regularity in the region $(C_0^{-1}|s-t|)^{1/\alpha} + \rho(x,y) \geqslant 2r^\delta$ inside the cylinder.

Proof of Theorem 3.8. We mainly follow the argument of [23, Theorem 4.14] with some modification. For simplicity, we assume that $\|q\|_{\infty,r} = 1$ and $q \geqslant 0$. Now, we first show that there are constants $\eta \in (0,1)$, $\delta \in (\sqrt{\delta_0}, 1)$ with $\delta_0 \in (0,1)$ being the constants δ in Theorem 3.4, Proposition 3.6 and Lemma 3.7, $R_1 > R_0$ and $\xi \in (0, (1/4) \wedge \eta^{1/\alpha})$ (which are determined later) such that for any $R_1 < R < r_G/(2c_*)$, $R^\delta \leqslant r \leqslant R$, $k \geqslant 1$ with $\xi^k r \geqslant 2r^\delta$, and any $(\tilde{t}, \tilde{x}) \in Q(t_0, x_0, r)$ with $x_0 \in B(0,R)$ and $t_0 \geqslant 0$,

$$\sup_{Q(\tilde{t},\tilde{x},\xi^k r)} q - \inf_{Q(\tilde{t},\tilde{x},\xi^k r)} q \leqslant \eta^k. \tag{3.23}$$

Let $Q_i = Q(\tilde{t}, \tilde{x}, \xi^i r)$ and $B_i = B(\tilde{x}, \xi^i r)$. Define $a_i = \inf_{Q_i} q$ and $b_i = \sup_{Q_i} q$. Clearly, $b_i - a_i \leqslant \eta^i$ for all $i \leqslant 0$. Suppose that $b_i - a_i \leqslant \eta^i$ for all $i \leqslant k$ with some $k \geqslant 0$. Choose

$z_1, z_2 \in Q_{k+1}$ such that $q(z_1) = b_{k+1}$ and $q(z_2) = a_{k+1}$. Letting $z_1 = (t_1, x_1)$, we define $\tilde{Q}_k = Q(t_1, x_1, \xi^k r)$, $\tilde{Q}_{k+1} = Q(t_1, x_1, \xi^{k+1} r)$ and

$$A_k = \left\{ z \in \tilde{Q}_{k+1} : q(z) \leqslant \frac{a_k + b_k}{2} \right\}.$$

Without of loss of generality, we may and do assume that $\nu(A_k)/\nu(\tilde{Q}_{k+1}) \geqslant 1/2$; otherwise, we will choose $1 - q$ instead of q. We have

$$\begin{aligned} b_{k+1} - a_{k+1} =& q(z_1) - q(z_2) = \mathbb{E}_{z_1}[q(Z_{\sigma_{A_k} \wedge \tau_{\tilde{Q}_k}})] - q(z_2) \\ =& \mathbb{E}_{z_1}\left[q(Z_{\sigma_{A_k} \wedge \tau_{\tilde{Q}_k}}) - q(z_2) : \sigma_{A_k} \leqslant \tau_{\tilde{Q}_k}\right] \\ &+ \mathbb{E}_{z_1}\left[q(Z_{\sigma_{A_k} \wedge \tau_{\tilde{Q}_k}}) - q(z_2) : \sigma_{A_k} > \tau_{\tilde{Q}_k}, X_{\tau_{\tilde{Q}_k}} \in B_{k-1}\right] \\ &+ \sum_{i=1}^{\infty} \mathbb{E}_{z_1}\left[q(Z_{\sigma_{A_k} \wedge \tau_{\tilde{Q}_k}}) - q(z_2) : \sigma_{A_k} > \tau_{\tilde{Q}_k}, X_{\tau_{\tilde{Q}_k}} \in B_{k-i-1} \setminus B_{k-i}\right] \\ =&: I_1 + I_2 + I_3. \end{aligned}$$

It is easy to see that

$$I_1 \leqslant \left(\frac{a_k + b_k}{2} - a_k\right) \mathbb{P}_{z_1}(\sigma_{A_k} \leqslant \tau_{\tilde{Q}_k}) \leqslant \frac{b_k - a_k}{2} p_k \leqslant \frac{\eta^k}{2} p_k = \eta^{k+1}\eta^{-1}\frac{p_k}{2}$$

and $I_2 \leqslant (b_{k-1} - a_{k-1})(1 - p_k) \leqslant \eta^{k-1}(1 - p_k) = \eta^{k+1}\eta^{-2}(1 - p_k)$, where $p_k := \mathbb{P}_{z_1}(\sigma_{A_k} \leqslant \tau_{\tilde{Q}_k}) = \mathbb{P}_{(t_1,x_1)}(\sigma_{A_k} \leqslant \tau_{Q(t_1,x_1,\xi^k r)})$. On the other hand, since $\xi^k r \geqslant 2r^\delta \geqslant 2R^{\delta_0}$, $\tilde{x} \in B(x_1, \xi^{k+1} r) \subset B(x_1, \xi^k r/2)$ and $\xi^{k-i} r > 4\xi^k r$ for $i \geqslant 1$, we can apply (3.21) and obtain that

$$\mathbb{P}_{x_1}(X_{\tau_{\tilde{Q}_k}} \in B_{k-i-1} \setminus B_{k-i}) \leqslant \mathbb{P}_{x_1}\left(X_{\tau_{Q(t_1,x_1,\xi^k r)}} \in B_{k-i}^c\right) \leqslant c_2 \left(\frac{\xi^k r}{\xi^{k-i} r}\right)^\alpha.$$

Thus,

$$\begin{aligned} I_3 &\leqslant \sum_{i=1}^{\infty} (b_{k-i-1} - a_{k-i-1}) \mathbb{P}_{x_1}(X_{\tau_{\tilde{Q}_k}} \in B_{k-i-1} \setminus B_{k-i}) \\ &\leqslant c_2 \sum_{i=1}^{\infty} \eta^{(k-i-1)} \left(\frac{\xi^k r}{\xi^{k-i} r}\right)^\alpha \leqslant \frac{c_2 \eta^{k+1}\eta^{-2}\xi^\alpha}{\eta - \xi^\alpha}. \end{aligned}$$

Note that, since $x_1 \in B(0, 2R)$ and $\xi^k r \geqslant 2r^\delta \geqslant 2R^{\delta_0}$, by (3.20) we have $p_k \geqslant c_3 > 0$. Combining with all the conclusions above, we arrive at that

$$\begin{aligned} b_{k+1} - a_{k+1} &\leqslant \eta^{k+1}\left(\frac{\eta^{-1} p_k}{2} + \eta^{-2}(1 - p_k) + \frac{c_2 \eta^{-2}\xi^\alpha}{\eta - \xi^\alpha}\right) \\ &= \eta^{k+1}\left[\eta^{-2} - \left(\eta^{-2} - \frac{\eta^{-1}}{2}\right) p_k + \frac{c_2 \eta^{-2}\xi^\alpha}{\eta - \xi^\alpha}\right] \\ &\leqslant \eta^{k+1}\left(\eta^{-2}(1 - c_3) + \frac{\eta^{-1} c_3}{2} + \frac{c_2 \eta^{-2}\xi^\alpha}{\eta - \xi^\alpha}\right). \end{aligned}$$

Choosing η close to 1 and then $\xi \in (0, (1/4) \wedge \eta^{1/\alpha})$ close to 0 such that

$$\eta^{-2}(1 - c_3) + \frac{\eta^{-1} c_3}{2} + \frac{c_2 \eta^{-2}\xi^\alpha}{\eta - \xi^\alpha} \leqslant 1,$$

we get $b_{k+1} - a_{k+1} \leqslant \eta_{k+1}$. This proves (3.23).

For any $(s, x), (t, y) \in Q(t_0, x_0, r)$ with $s \leqslant t$ and $(C_0^{-1}|t - s|)^{1/\alpha} + \rho(x, y) \geqslant 2r^\delta$, let k be the smallest integer such that $(C_0^{-1}|s - t|)^{1/\alpha} + \rho(x, y) \geqslant \xi^{k+1} r$. Then, $(C_0^{-1}|s - t|)^{1/\alpha} + \rho(x, y) \leqslant \xi^k r$, and so $\xi^k r \geqslant 2r^\delta$ and $(t,y) \in Q(s,x,\xi^k r)$. According to (3.23), we know that

$$|q(s,x) - q(t,y)| \leqslant \eta^k \leqslant \eta^{-1}\left(\frac{(C_0^{-1}|s-t|)^{1/\alpha} + \rho(x,y)}{r}\right)^{\log_\xi \eta}.$$

The proof is finished. □

Remark 3.10. According to Proposition 3.5, the proof of Theorem 3.4 and the arguments in this subsection, we can obtain that, when $\alpha \in (0,1)$, Theorems 3.4 and 3.8 still hold under assumption **(Exi.')**.

4. Convergence of stable-like processes on metric measure spaces

In this section, we give convergence criteria for stable-like processes on metric measure spaces.

Let (F,ρ,m) be a metric measure space, where (F,ρ) is a locally compact separable and connected metric space, and m is a Radon measure on F. For every $x \in F$ and $r > 0$, let $B_F(x,r) = \{z \in F : \rho(z,x) < r\}$. We always assume the following assumptions on (F,ρ,m).

Assumption (MMS).

(i) *For every $x \in F$ and $r > 0$, the closure of $B_F(x,r)$ is compact, and it holds that $m(\partial(B_F(x,r))) = 0$, where $\partial(B_F(x,r)) = \overline{B_F(x,r)} \backslash B_F(x,r)$.*

(ii) *$\rho : F \times F \to \mathbb{R}_+$ is geodesic, i.e., for any $x,y \in F$, there exists a continuous map $\gamma : [0,\rho(x,y)] \to F$ such that $\gamma(0) = x$, $\gamma(\rho(x,y)) = y$ and $\rho(\gamma(s),\gamma(t)) = t - s$ for all $0 \leqslant s \leqslant t \leqslant \rho(x,y)$.*

(iii) *There exist constants $c_F \geqslant 1$ and $d > 0$ such that*

$$c_F^{-1} r^d \leqslant m(B_F(x,r)) \leqslant c_F r^d, \quad \forall\, x \in F,\ 0 < r < r_F := \sup_{y,z\in F} \rho(y,z). \tag{4.1}$$

The metric measure space (F,ρ,m) will serve as the state space of the stable-like process Y which will be defined later.

According to [22, Theorem 2.1], such a metric measure space is endowed with the following graph approximations.

Lemma 4.1. *Under assumption* **(MMS)**, *F admits a sequence of approximating graphs $\{G_n := (V_n, E_{V_n}), n \geqslant 1\}$ such that the following properties hold.*

(1) *For every $n \geqslant 1$, $V_n \subseteq F$, and (V_n, E_{V_n}) is connected and has uniformly bounded degree. Moreover, $\cup_{n=1}^\infty V_n$ is dense in F.*

(2) *There exist positive constants C_1 and C_2 such that for every $n \geqslant 1$ and $x,y \in V_n$,*

$$\frac{C_1}{n}\rho_n(x,y) \leqslant \rho(x,y) \leqslant \frac{C_2}{n}\rho_n(x,y), \tag{4.2}$$

where ρ_n is the graph distance of (V_n, E_{V_n}).

(3) *For each $n \geqslant 1$, there exist a class of subsets $\{U_n(x) : x \in V_n\}$ of F such that $\bigcup_{x\in V_n} U_n(x) \subset F$, $m\big(U_n(x) \cap U_n(y)\big) = 0$ for $x \neq y$,*

$$V_n \cap \mathit{Int}\, U_n(x) = \{x\},\ \sup\{\rho(y,z) : y,z \in U_n(x)\} \leqslant \frac{C_3}{n}, \quad \forall\, x \in V_n, \tag{4.3}$$

and

$$\frac{C_4}{n^d} \leqslant m\big(U_n(x)\big) \leqslant \frac{C_5}{n^d}, \quad \forall\, n \geqslant 1,\ x \in V_n, \tag{4.4}$$

where $\mathit{Int}\, U_n(x)$ denotes the set of the interior points of $U_n(x)$.

Moreover, for all $r > 0$ and $y \in F$,

$$\lim_{n\to\infty} m\Big(B_F(y,r) \bigcap \big(F \setminus \bigcup_{x\in V_n} U_n(x)\big)\Big) = 0. \tag{4.5}$$

For each $n \geqslant 1$ and $y \in F \setminus \bigcup_{x\in V_n} U_n(x)$, there exists $z \in V_n$ such that $\rho(y,z) \leqslant C_6 n^{-1}$. Here C_i $(i = 3, \cdots, 6)$ are positive constants independent of n.

We will consider stable-like processes on the graphs $\{G_n\}_{n\geqslant 1}$.

4.1. **Stable-like processes on graphs and the metric measure spaces.** We first introduce a class of Dirichlet forms $(D_{V_n}, \mathscr{F}_{V_n})$ on the graph (V_n, E_{V_n}). For any $n \geqslant 1$, define

$$D_{V_n}(f,f) = \frac{1}{2}\sum_{x,y\in V_n}(f(x)-f(y))^2\frac{w_{x,y}^{(n)}}{\rho(x,y)^{d+\alpha}}m_n(x)m_n(y), \quad f\in\mathscr{F}_{V_n},$$
$$\mathscr{F}_{V_n} = \{f\in L^2(V_n;m_n) : D_{V_n}(f,f)<\infty\},$$

where $\alpha\in(0,2)$, $\rho(x,y)$ is the distance function on F, m_n is the measure on V_n defined by

$$m_n(A) := \sum_{x\in A} m\big(U_n(x)\big), \quad \forall\ A\subset V_n,$$

(for simplicity, we write $m_n(x) = m_n(\{x\})$ for all $x\in V_n$), and $\{w_{x,y}^{(n)} : x,y\in V_n\}$ is a sequence satisfying that $w_{x,y}^{(n)}\geqslant 0$ and $w_{x,y}^{(n)} = w_{y,x}^{(n)}$ for all $x\neq y$, and

$$\sum_{y\in V_n}\frac{w_{x,y}^{(n)}}{\rho(x,y)^{d+\alpha}}m_n(y)<\infty, \quad x\in V_n.$$

We note that, in the definition of the Dirichlet form $(D_{V_n}, \mathscr{F}_{V_n})$ we use the metric $\rho(x,y)$ instead of the graph metric $\rho_n(x,y)$ on V_n. According to [22, Theorem 2.1], for any $n\geqslant 1$, $(D_{V_n}, \mathscr{F}_{V_n})$ is a regular Dirichlet form on $L^2(V_n;m_n)$. Let $X^{(n)} := \{(X_t^{(n)})_{t\geqslant 0}, (\mathbb{P}_x)_{x\in V_n}\}$ be the associated symmetric Markov process.

To obtain the weak convergence for $X^{(n)}$, we also introduce a kind of scaling processes associated with $\{X^{(n)}\}_{n\geqslant 1}$. For any $n\geqslant 1$, let $\mathbf{P}_n$ be the projection map from (V_n,ρ) to (V_n,ρ_n) such that $\mathbf{P}_n(x) := x$ for $x\in V_n$. Define a measure $\tilde m_n$ on (V_n,ρ_n) as follows

$$\tilde m_n(A) = n^d m_n\big(\mathbf{P}_n^{-1}(A)\big) = n^d\sum_{x\in\mathbf{P}_n^{-1}(A)} m_n(x), \quad A\subset V_n.$$

For simplicity, $\tilde m_n(x) = \tilde m_n(\{x\})$ for any $x\in V_n$. For any $n\geqslant 1$, we consider the following Dirichlet form: $\big(\tilde D_{V_n}, \tilde{\mathscr{F}}_{V_n}\big)$ on $L^2(V_n;\tilde m_n)$

$$\tilde D_{V_n}(f,f) = \frac{1}{2}\sum_{x,y\in V_n}(f(x)-f(y))^2\frac{\tilde w_{x,y}^{(n)}}{\rho_n(x,y)^{d+\alpha}}\tilde m_n(x)\tilde m_n(y), \quad f\in\tilde{\mathscr{F}}_{V_n},$$
$$\tilde{\mathscr{F}}_{V_n} = \{f\in L^2(V_n;\tilde m_n) : \tilde D_{V_n}(f,f)<\infty\},$$

where

$$\tilde w_{x,y}^{(n)} := w_{x,y}^{(n)}\left(\frac{\rho_n(x,y)}{n\rho(x,y)}\right)^{d+\alpha}, \quad x,y\in V_n.$$

Note that $\tilde D_{V_n}(f,f) = n^{d-\alpha}D_{V_n}(f,f)$ and $\tilde{\mathscr{F}}_{V_n} = \mathscr{F}_{V_n}$. Let $\tilde X^{(n)}$ be the symmetric Markov process associated with $\big(\tilde D_{V_n}, \tilde{\mathscr{F}}_{V_n}\big)$. According to the expressions of $\big(D_{V_n}, \mathscr{F}_{V_n}\big)$ and $\big(\tilde D_{V_n}, \tilde{\mathscr{F}}_{V_n}\big)$, we know that $\big(\mathbf{P}_n(X_t^{(n)})\big)_{t\geqslant 0}$ has the same distribution as $\big(\tilde X_{n^\alpha t}^{(n)}\big)_{t\geqslant 0}$.

As a candidate of the scaling limit of the discrete forms $(D_{V_n}, \mathscr{F}_{V_n})$, we now define a symmetric Dirichlet form $(D_0, \mathscr{F}_0)$ on $L^2(F;m)$ as follows

$$D_0(f,f) = \frac{1}{2}\int_{\{F\times F\setminus \mathrm{diag}\}}\big(f(x)-f(y)\big)^2\frac{c(x,y)}{\rho(x,y)^{d+\alpha}}\,m(dx)\,m(dy), \quad f\in\mathscr{F}_0,$$
$$\mathscr{F}_0 = \{f\in L^2(F;m) : D_0(f,f)<\infty\}, \tag{4.6}$$

where $\alpha\in(0,2)$, $\mathrm{diag} := \{(x,y)\in F\times F : x=y\}$ and $c : F\times F\to(0,\infty)$ is a symmetric continuous function such that $0<c_1\leqslant c(x,y)\leqslant c_2<\infty$ for all $(x,y)\in F\times F\setminus\mathrm{diag}$ and some constants c_1, c_2. According to (4.1) and the fact that $\alpha\in(0,2)$, we have

$$\sup_{x\in F}\int_{F\setminus\{y\in F: y\neq x\}}\big(1\wedge\rho^2(x,y)\big)\frac{c(x,y)}{\rho(x,y)^{d+\alpha}}\,m(dy)$$

$$
\begin{aligned}
&\leqslant \sup_{x\in F}\sum_{k=0}^{\infty}\int_{\{y\in F:2^{-(1+k)}<\rho(y,x)\leqslant 2^{-k}\}}\frac{c(x,y)}{\rho(x,y)^{d+\alpha-2}}\,m(dy)\\
&\quad+\sup_{x\in F}\sum_{k=0}^{\infty}\int_{\{y\in F:2^{k}<\rho(y,x)\leqslant 2^{1+k}\}}\frac{c(x,y)}{\rho(x,y)^{d+\alpha}}\,m(dy)\\
&\leqslant c_2\sup_{x\in F}\left(\sum_{k=0}^{\infty}m(B_F(x,2^{-k}))2^{(d+\alpha-2)(1+k)}+\sum_{k=0}^{\infty}m(B_F(x,2^{1+k}))2^{-(d+\alpha)k}\right)\\
&\leqslant c_3\left(\sum_{k=0}^{\infty}2^{-(2-\alpha)k}+\sum_{k=0}^{\infty}2^{-\alpha k}\right)<\infty.
\end{aligned}
$$

This implies $\mathrm{Lip}_c(F)\subseteq\mathscr{F}_0$, where $\mathrm{Lip}_c(F)$ denotes the space of Lipschitz continuous functions on F with compact support. We also need the following assumption on $(D_0,\mathscr{F}_0)$.

Assumption (Dir.) *$\mathrm{Lip}_c(F)$ is dense in $\mathscr{F}_0$ under the norm* $\|\cdot\|_{D_0,1}:=\big(D_0(\cdot,\cdot)+\|\cdot\|^2_{L^2(F;m)}\big)^{1/2}$.

Therefore, $(D_0,\mathscr{F}_0)$ is a regular Dirichlet form on $L^2(F;m)$, and there exists a strong Markov process $Y:=(Y_t)_{t\geqslant 0}$ associated with $(D_0,\mathscr{F}_0)$. Moreover, by [23, Theorem 1.1] or [24, Theorem 1.2], the process Y has a heat kernel $p^Y:(0,\infty)\times F\times F\to(0,\infty)$, which is jointly continuous. In particular, the process $Y:=\big((Y_t)_{t\geqslant 0},(\mathbb{P}_x^Y)_{x\in F}\big)$ can start from all $x\in F$. The process Y is called a α-stable-like process in the literature, see [23, 24]. Two-sided estimates for heat kernel $p^Y(t,x,y)$ of the process Y have been obtained in [23].

4.2. **Generalized Mosco convergence.** To study the convergence property of process $X^{(n)}$, we will use some results from [22], which are concerned with the generalized Mosco convergence of $X^{(n)}$.

For any $n\geqslant 1$, we define an extension operator $E_n:L^2\big(V_n;m_n\big)\to L^2(F;m)$ as follows

$$
E_n(g)(z)=\begin{cases} g(x), & z\in \mathrm{Int}U_n(x)\text{ for some }x\in V_n,\\ 0, & z\in F\setminus\bigcup_{x\in V_n}U_n(x),\end{cases}\qquad g\in L^2\big(V_n;m_n\big). \tag{4.7}
$$

Note that because $m(\partial U_n(x))=0$ for any $x\in V_n$ by Assumption **(MMS)**(i), there is no need to worry about $E_n(g)$ on $\bigcup_{x\in V_n}\partial U_n(x)$, and the function $E_n(g)$ is a.s. well defined on F. Note also that the definition of the extension operator E_n above is a little different from that in [22], see [22, (2.14)]. Furthermore, we define a projection (restriction) operator $\pi_n:L^2(F;m)\to L^2\big(V_n;m_n\big)$ as follows

$$
\pi_n(f)(x)=m_n(x)^{-1}\int_{U_n(x)}f(z)\,m(dz),\quad x\in V_n,\ f\in L^2\big(F;m\big).
$$

Remark 4.2. As shown in Lemma 4.1, under assumption **(MMS)**, the space F admits a sequence of approximating graphs $\{(V_n,E_{V_n}):n\geqslant 1\}$ enjoying all the properties mentioned in Lemma 4.1. Though these properties are weaker than **(AG.1)**–**(AG.3)** in [22, Theorem 2.1], one can verify that [22, Lemma 4.1] and so [22, Theorem 4.7] still hold with notations above.

For simplicity, we assume that there exists a point $0\in\bigcap_{n=1}^{\infty}V_n$; otherwise, we can take a sequence $\{o_n\}_{n\geqslant 1}$ such that $o_n\in V_n$ for all $n\geqslant 1$ and $\lim_{n\to\infty}o_n$ exists, and then the arguments below still hold true with this limit point $0:=\lim_{n\to\infty}o_n$.

Fix $0\in\cap_{n=1}^{\infty}V_n$. We assume that the following conditions hold for $\{w_{x,y}^{(n)}:x,y\in V_n\}$.

Assumption (Mos.)

(i) *For every $R>0$,*

$$
\lim_{\varepsilon\to 0}\limsup_{n\to\infty}\left[n^{-2d}\sum_{x,y\in B_F(0,R)\cap V_n:0<\rho(x,y)\leqslant\varepsilon}\frac{w_{x,y}^{(n)}}{\rho(x,y)^{d+\alpha-2}}\right]=0 \tag{4.8}
$$

and

$$\lim_{l\to\infty}\limsup_{n\to\infty}\left[n^{-2d}\sum_{x,y\in B_F(0,R)\cap V_n:\rho(x,y)\geqslant l}\frac{w_{x,y}^{(n)}}{\rho(x,y)^{d+\alpha}}\right]=0. \tag{4.9}$$

(ii) *For any sufficiently small $\varepsilon>0$, large $R>0$ and any $f\in \mathrm{Lip}_c(F)$,*

$$\lim_{n\to\infty}\left[n^{-d}\sum_{x\in B_F(0,R)\cap V_n}\left(\sum_{y\in B_F(0,R)\cap V_n:\rho(x,y)>\varepsilon}(f(x)-f(y))\frac{(w_{x,y}^{(n)}-c(x,y))}{\rho(x,y)^{d+\alpha}}m_n(y)\right)^2\right]=0. \tag{4.10}$$

(iii) *For any sufficiently small $\varepsilon>0$, large $R>0$ and any $f\in C_b(B_F(0,R))$,*

$$\lim_{n\to\infty}\sum_{x,y\in B_F(0,R)\cap V_n:\rho(x,y)>\varepsilon}\big(f(x)-f(y)\big)^2\frac{\big(w_{x,y}^{(n)}-c(x,y)\big)}{\rho(x,y)^{d+\alpha}}m_n(x)m_n(y)=0. \tag{4.11}$$

Denote by $(P_t^Y)_{t\geqslant 0}$ the Markov semigroup of the process Y, and denote by $(P_t^{(n)})_{t\geqslant 0}$ the Markov semigroup of the process $X^{(n)}$. We set $\hat P_t^{(n)}f(x)=E_n(P_t^{(n)}(\pi_n(f)))(x)$ for any $f\in L^2(F;m)$.

Proposition 4.3. *Suppose that Assumptions* **(MMS), (Dir.)** *and* **(Mos.)** *hold. Then*

$$\lim_{n\to\infty}\|\hat P_t^{(n)}f-P_t^Yf\|_{L^2(F;m)}=0,\quad f\in L^2(F;m),\ t>0.$$

Proof. It is easy to see that the Dirichlet form $(D_0,\mathscr{F}_0)$ satisfies $(A2)$ in [22, Section 2]. By assumption **(Dir.)** and the continuity of $c(x,y)$, we know that $(A3)^*$ in [22, Section 2] holds true.

Clearly, condition $(A4)^*$ (i) in [22, Section 2] is a direct consequence of (4.8) and (4.9). For any $R,\varepsilon>0$ and $f\in \mathrm{Lip}_c(F)$, define

$$\begin{aligned}
L_{R,\varepsilon}f(x)&=\int_{\{z\in B_F(0,R):\rho(z,x)>\varepsilon\}}(f(z)-f(x))\frac{c(x,z)}{\rho(x,z)^{d+\alpha}}\,m(dz),\quad x\in F,\\
\overline{L_{R,\varepsilon}^n}f(x)&=\sum_{z\in B_F(0,R)\cap V_n:\rho(x,z)>\varepsilon}(f(z)-f(x))\frac{w_{x,z}^{(n)}}{\rho(x,z)^{d+\alpha}}m_n(z),\quad x\in V_n,\\
L_{R,\varepsilon}^nf(x)&=E_n(\overline{L_{R,\varepsilon}^n}f)(x),\quad x\in F.
\end{aligned}$$

Then,

$$\int_{B_F(0,R)}|L_{R,\varepsilon}^nf(x)-L_{R,\varepsilon}f(x)|^2\,m(dx)\leqslant\sum_{i=1}^4 I_{i,n},$$

where

$$\begin{aligned}
I_{1,n}&=2\sum_{x\in B_F(0,R)\cap V_n}\left(\sum_{\substack{y\in B_F(0,R)\cap V_n:\\ \rho(x,y)>\varepsilon}}\big(f(x)-f(y)\big)\frac{(w_{x,y}^{(n)}-c(x,y))}{\rho(x,y)^{d+\alpha}}m_n(y)\right)^2 m_n(x),\\
I_{2,n}&=8\mathrm{osc}_n(f)^2\sum_{x\in B_F(0,R)\cap V_n}\left(\sum_{y\in B_F(0,R)\cap V_n:\rho(x,y)>\varepsilon}\frac{c(x,y)}{\rho(x,y)^{d+\alpha}}m_n(y)\right)^2 m_n(x),\\
I_{3,n}&=8\|f\|_\infty^2\mathrm{osc}_n(c)^2\int_{B_F(0,R)}\left(\int_{B_F(0,R)\cap\{y\in F:\rho(x,y)>\varepsilon\}}\frac{1}{\rho(x,y)^{d+\alpha}}\,m(dy)\right)^2 m(dx),\\
I_{4,n}&=4\|f\|_\infty^2\|c\|_\infty^2\int_{B_F(0,R)\cap(F\setminus\cup_{z\in V_n}U_n(z))}\left(\int_{\substack{B_F(0,R)\cap(F\setminus\cup_{z\in V_n}U_n(z))\\ \cap\{y\in F:\rho(x,y)>\varepsilon\}}}\frac{1}{\rho(x,y)^{d+\alpha}}\,m(dy)\right)^2 m(dx),
\end{aligned}$$

$$\mathrm{osc}_n(f)=\sup_{x\in B_F(0,R)\cap V_n,x_1,x_2\in U_n(x)}|f(x_1)-f(x_2)|,$$

$$\mathrm{osc}_n(c) = \sup_{x,y\in B_F(0,R)\cap V_n, x_1,x_2\in U_n(x), y_1,y_2\in U_n(y)} |c(x_1,y_1) - c(x_2,y_2)|.$$

It follows from (4.4) and (4.10) that $\lim_{n\to\infty} I_{1,n} = 0$. Since $f \in \mathrm{Lip}_c(F)$, $\mathrm{osc}_n(f) \to 0$ as $n\to\infty$. Then, we arrive at

$$\begin{aligned}\limsup_{n\to\infty} I_{2,n} &\leqslant c_1\varepsilon^{-2(d+\alpha)}\big[\limsup_{n\to\infty}\mathrm{osc}_n(f)^2\big]\\ &\quad\times\sup_{n\geqslant 1}\left\{n^{-3d}\sum_{x\in B_F(0,R)\cap V_n}\Big(\sum_{y\in B_F(0,R)\cap V_n:\rho(x,y)>\varepsilon} c(x,y)\Big)^2\right\}\\ &\leqslant c_2(\varepsilon)\big[\limsup_{n\to\infty}\mathrm{osc}_n(f)^2\big] = 0.\end{aligned}$$

By the continuity of $c(x,y)$, it is also easy to see that $\lim_{n\to\infty} I_{3,n} = 0$. Obviously, (4.5) implies that $\lim_{n\to\infty} I_{4,n} = 0$. Therefore, we have

$$\lim_{n\to\infty}\int_{B_F(0,R)} |L^n_{R,\varepsilon} f(x) - L_{R,\varepsilon} f(x)|^2\, m(dx) = 0,$$

which implies that condition $(A4)^*$ (ii) in [22, Section 2] is satisfied.

Similarly, with aid of (4.11), we can claim that condition $(A4)^*$ (iii) in [22, Section 2] is also fulfilled. Therefore, we can verify that all the conditions of $(A4)^*$ in [22, Section 2] hold under assumptions **(MMS), (Dir.)** and **(Mos.)**. Hence, the required assertion follows from [22, Theorem 4.7 and Theorem 8.3]. □

4.3. **Weak convergence.** The main purpose of this subsection is to establish the weak convergence theorem of the law for $X^{(n)}$. For any $T \in (0,\infty]$, denote by $\mathscr{D}([0,T];F)$ the collection of càdlàg F-valued functions on $[0,T]$ equipped with the Skorohod topology. Let $\mathbb{P}^{(n)}_x$ be the law of $X^{(n)}$ with starting point $x \in V_n$. Note that $\mathbb{P}^{(n)}_x$ can be seen as a distribution on $\mathscr{D}([0,T];F)$.

We will make use of scaling processes $\{\tilde X^{(n)}\}_{n\geqslant 1}$ constructed in Subsection 4.1. First, we consider some properties of the space $(V_n, \rho_n, \tilde m_n)$. For any $x\in V_n$ and $r>0$, let $B_{V_n}(x,r) = \{z\in V_n : \rho_n(z,x)\leqslant r\}$.

Lemma 4.4. *Under assumption* **(MMS)***, there are constants $C_0>0$ and $c_V \geqslant 1$ such that for all $n\geqslant 1$,*

$$c_V^{-1} \leqslant \tilde m_n(x) \leqslant c_V, \quad x\in V_n \tag{4.12}$$

and

$$c_V^{-1} r^d \leqslant \tilde m_n(B_{V_n}(x,r)) \leqslant c_V r^d, \quad x\in V_n, 1\leqslant r < C_0 n r_F, \tag{4.13}$$

where r_F is the constant in (4.1).

Proof. By the definition of $\tilde m_n$ and (4.4), (4.12) holds trivially.

Note that, for any $x\in V_n$, $y\in B_F(x,r)\cap V_n$ and $z\in U_n(y)$, by (4.3), we have $\rho(z,x)\leqslant \rho(z,y)+\rho(y,x)\leqslant C_3 n^{-1} + r$, and so $\bigcup_{y\in B_F(x,r)\cap V_n} U_n(y) \subseteq B_F(x, r + C_3 n^{-1})$. Hence, for any $x\in V_n$ and $1\leqslant r < (nr_F - C_3)/C_2$ (where C_2 and C_3 are constants in (4.2) and (4.3)),

$$\begin{aligned}\tilde m_n\big(B_{V_n}(x,r)\big) &= n^d m_n\big(B_{V_n}(x,r)\cap V_n\big) \leqslant n^d m_n\big(B_F(x, C_2 n^{-1} r)\cap V_n\big)\\ &= n^d \sum_{y\in B_F(x,C_2n^{-1}r)\cap V_n} m\big(U_n(y)\big) \leqslant n^d m\big(B_F(x, C_2 n^{-1} r + C_3 n^{-1})\big) \leqslant c_0 r^d,\end{aligned}$$

where in the first inequality we used (4.2), the second inequality is due to the facts that $m(U_n(x)\cap U_n(y)) = 0$ for all $x\neq y$ and $\bigcup_{y\in B_F(x,C_2n^{-1}r)\cap V_n} U_n(y) \subseteq B_F(x, C_2 n^{-1} r + C_3 n^{-1})$ as explained above, and the last inequality follows from (4.1).

On the other hand, for any $z\in B_F(x,r)$, by (3) in Lemma 4.1, there exists $y\in V_n$ such that $\rho(y,z)\leqslant c_0 n^{-1}$ for some constant $c_0>0$, and so $\rho(y,x)\leqslant \rho(z,x)+\rho(z,y)\leqslant r + c_0 n^{-1}$.

This implies that $B_F(x,r) \subset \bigcup_{y\in B_F(x,r+c_0n^{-1})\cap V_n} B_F(y,c_0n^{-1})$. Hence, for $(2(C_1^{-1}c_0)) \vee 1 < r < (nr_F+c_0)/C_1$ (where C_1 is the constant in (4.2)) and $x \in V_n$,

$$\begin{aligned}\tilde m_n\big(B_{V_n}(x,r)\big) &= n^d m_n\big(B_{V_n}(x,r)\big) \geqslant n^d m_n\big(B_F(x,C_1n^{-1}r)\cap V_n\big)\\ &= n^d \sum_{y\in B_F(x,C_1n^{-1}r)\cap V_n} m\big(U_n(y)\big) \geqslant c_1 n^d \sum_{y\in B_F(x,C_1n^{-1}r)\cap V_n} m\big(B_F(y,c_0n^{-1})\big)\\ &\geqslant c_1n^d m\big(B_F(x,C_1n^{-1}r-c_0n^{-1})\big) \geqslant c_2 r^d,\end{aligned}$$

where in the first inequality we used (4.2) again, the second inequality follows from (4.1) and (4.4), the third inequality is due to $\bigcup_{y\in B_F(x,C_1n^{-1}r)\cap V_n} B_F(y,c_0n^{-1}) \supseteq B_F(x,C_1n^{-1}r-c_0n^{-1})$ as claimed before, and in the last one we have used (4.1).

Therefore, combining both estimates above and changing the corresponding constants properly, we prove (4.13). □

By (4.2), for all $n \geqslant 1$, $\sup_{x,y\in V_n} \rho_n(x,y) \leqslant C_1^{-1} n r_F$, where r_F is the constant in (4.1). Below, we let $C_0' = C_1^{-1}$. For any $x,z \in V_n$ and $r>0$, let $B^{w^{(n)}}_{V_n,z}(x,r) = \{y \in B_{V_n}(x,r) : w^{(n)}_{y,z} > 0\}$, and $B^{w^{(n)}}_{V_n}(x,r) = B^{w^{(n)}}_{V_n,x}(x,r)$. We need the following further assumptions on $\{w^{(n)}_{x,y} : x,y\in V_n\}$.

Assumption (Wea.) *Suppose that for some fixed $\theta \in (0,1)$, there exists a constant $R_0 \geqslant 1$ such that*

(i) *For any $n \geqslant 1$, $R_0 < R < C_0' r_F$ and $R^\theta/2 \leqslant r \leqslant 2R$,*

$$\sup_{x\in B_{V_n}(0,6R)} \sum_{y\in V_n:\rho_n(y,x)\leqslant r} \frac{w^{(n)}_{x,y}}{\rho_n(x,y)^{d+\alpha-2}} \leqslant C_3 r^{2-\alpha}, \tag{4.14}$$

$$m_n(B^{w^{(n)}}_{V_n,z}(x,r)) \geqslant c_0 m_n(B_{V_n}(x,r)), \quad x,z \in B_{V_n}(0,6R), \tag{4.15}$$

and

$$\sup_{x\in B_{V_n}(0,6R)\cap V_n} \sum_{y\in V_n:\rho_n(y,x)\leqslant c_* r, w^{(n)}_{x,y}>0} (w^{(n)}_{x,y})^{-1} \leqslant C_3 r^d, \tag{4.16}$$

where $c_0 > 1/2$ is independent of n, R_0, R, r, x, z and r_F, $c_* = 8c_V^{2/d}$ and c_V is the constant in (4.13).

When $\alpha \in (0,1)$, (4.14) *can be replaced by*

$$\sup_{x\in B_{V_n}(0,6R)} \sum_{y\in V_n:\rho_n(y,x)\leqslant r} \frac{w^{(n)}_{x,y}}{\rho_n(x,y)^{d+\alpha-1}} \leqslant C_3 r^{1-\alpha}. \tag{4.17}$$

(ii) *For every $n \geqslant 1$, $R_0 < R < C_0' r_F$ and $r \geqslant R^\theta/2$,*

$$\sup_{x\in B_{V_n}(0,6R)} \sum_{y\in V_n:\rho_n(x,y)> r} \frac{w^{(n)}_{x,y}}{\rho_n(x,y)^{d+\alpha}} \leqslant C_3 r^{-\alpha}. \tag{4.18}$$

Here C_3 is a positive constant independent of n, R_0 and r_F.

The main result of this section is as follows. It is in some sense a generalization of [20, Proposition 2.8]. Indeed, in our case we have the Hölder regularity of parabolic functions only in the region $(C_0^{-1}|s-t|)^{1/\alpha} + \rho(x,y) \geqslant 2r^\delta$ (see Theorem 3.8), hence more careful arguments are required.

Theorem 4.5. *Suppose that Assumptions* **(MMS)**, **(Dir.)**, **(Mos.)** *and* **(Wea.)** *hold. Then, for any $\{x_n \in V_n : n \geqslant 1\}$ such that $\lim_{n\to\infty} x_n = x$ for some $x \in F$, it holds that for every $T>0$, $\mathbb{P}^{(n)}_{x_n}$ converges weakly to $\mathbb{P}^Y_x$ on the space of all probability measures on $\mathscr{D}([0,T];F)$, where $\mathbb{P}^{(n)}_{x_n}$ and $\mathbb{P}^Y_x$ denote the laws of $X^{(n)}_\cdot$ and $Y_\cdot$ on $\mathscr{D}([0,T];F)$, respectively.*

Proof. Throughout the proof, we denote the law of $(X_t^{(n)})_{t\geqslant 0}$ on $\mathscr{D}([0,\infty);F)$ and that of $(\tilde X_t^{(n)})_{t\geqslant 0}$ on $\mathscr{D}([0,\infty);V_n)$ by $\mathbb{P}_\cdot^{(n)}$ and $\tilde{\mathbb{P}}_\cdot^{(n)}$, respectively. Let $X_\cdot^{(n)}$ and $\tilde X_\cdot^{(n)}$ be their associated canonical paths.

Suppose that $\{x_n\in V_n : n\geqslant 1\}$ is a sequence with $\lim_{n\to\infty} x_n = x$ for some $x\in F$.

Step (1): We show that for each fixed $T>0$, $\{\mathbb{P}_{x_n}^{(n)}\}_{n\geqslant 1}$ is tight on $\mathscr{D}([0,T];F)$. To prove the tightness of $\{\mathbb{P}_{x_n}^{(n)}\}_{n\geqslant 1}$, it suffices to verify that

$$\lim_{R\to\infty}\limsup_{n\to\infty}\mathbb{P}_{x_n}^{(n)}\big(\sup_{s\in[0,T]}\rho(0,X_s^{(n)})>R\big)=0, \tag{4.19}$$

and for any sequence of stopping time $\{\tau_n\}_{n\geqslant 1}$ such that $\tau_n\leqslant T$ and any sequence $\{\varepsilon_n\}_{n\geqslant 1}$ with $\lim_{n\to\infty}\varepsilon_n=0$,

$$\limsup_{n\to\infty}\mathbb{P}_{x_n}^{(n)}\Big(\rho\big(X_{\tau_n+\varepsilon_n}^{(n)},X_{\tau_n}^{(n)}\big)>\eta\Big)=0,\quad \eta>0. \tag{4.20}$$

See, e.g., [1, Theorem 1].

When $r_F<\infty$, (4.19) holds trivially. Now, we are going to prove (4.19) for the case that $r_F=\infty$. As we mentioned above, $\big(\mathbf{P}_n(X_t^{(n)})\big)_{t\geqslant 0}$ has the same distribution as $\big(\tilde X_{n^\alpha t}^{(n)}\big)_{t\geqslant 0}$, where $(\tilde X_t^{(n)})_{t\geqslant 0}$ is a strong Markov process generated by the Dirichlet form $(\tilde D_{V_n},\tilde{\mathscr{F}}_{V_n})$. Therefore,

$$\begin{aligned}\mathbb{P}_{x_n}^{(n)}\big(\sup_{s\in[0,T]}\rho(X_s^{(n)},0)>R\big)&=\mathbb{P}_{x_n}^{(n)}\Big(\sup_{s\in[0,T]}\rho\big(\mathbf{P_n}(X_s^{(n)}),0\big)>R\Big)\\&=\tilde{\mathbb{P}}_{x_n}^{(n)}\Big(\sup_{s\in[0,n^\alpha T]}\rho(\tilde X_s^{(n)},0)>R\Big)\\&\leqslant\tilde{\mathbb{P}}_{x_n}^{(n)}\Big(\sup_{s\in[0,n^\alpha T]}\rho_n(\tilde X_s^{(n)},0)>c_1^*nR\Big),\end{aligned} \tag{4.21}$$

where the last inequality follows the fact that $\rho_n(x,y)\geqslant c_1^*n\rho(x,y)$ for all $x,y\in V_n$, thanks to (4.2).

On the other hand, under assumption **(Wea.)**, it is easy to verify that assumption **(Exi.)** (or assumption **(Exi.')** when $\alpha\in(0,1)$) holds on the space $(V_n,\rho_n,\tilde m_n)$ with associated constants independent of n. Combining this fact with (4.12) and (4.13), we can apply Theorem 3.4 (or Remark 3.10) to derive that for any fixed $\theta'\in(\theta,1)$, there exist constants $\delta\in(\theta,1)$ and $R_1\geqslant 1$, such that for all $n\geqslant 1$, $R_1<R<C_0'r_Fn$ and $R^\delta\leqslant r\leqslant R$,

$$\sup_{x\in B_{V_n}(0,2R)\cap V_n}\tilde{\mathbb{P}}_x^{(n)}\big(\tau_{B_{V_n}(0,r)}(\tilde X_\cdot^{(n)})\leqslant t\big)\leqslant c_1\Big(\frac{t}{r^\alpha}\Big)^{1/3},\quad \forall\, t\geqslant r^{\theta'\alpha}, \tag{4.22}$$

where $B_{V_n}(x,r)=\{z\in V_n:\rho_n(z,x)\leqslant r\}$, $\tau_{B_{V_n}(0,r)}(\tilde X_\cdot^{(n)})$ is the first exit time from $B_{V_n}(0,r)$ of the process $\tilde X_\cdot^{(n)}$, and $c_1>0$ is independent of R_1, n, r, R and r_F.

Suppose that $\rho(x_n,0)\leqslant K$ for all $n\geqslant 1$ and some constant $K>0$. Note that, also thanks to (4.2), $\rho_n(x_n,0)\leqslant c_2^*n\rho(x_n,0)\leqslant c_2^*nK$. For every fixed $R>2c_2^*K/c_1^*$ and $T>0$, we have $R_1<c_1^*nR<C_0'nr_F$ (since $r_F=\infty$) and $n^\alpha T>\big(c_1^*nR/2\big)^{\theta'\alpha}$ for n large enough. Thus, by (4.21) and (4.22),

$$\begin{aligned}\mathbb{P}_{x_n}^{(n)}\big(\sup_{s\in[0,T]}\rho(X_s^{(n)},0)>R\big)&\leqslant\tilde{\mathbb{P}}_{x_n}^{(n)}\big(\sup_{s\in[0,n^\alpha T]}\rho_n(\tilde X_s^{(n)},0)>c_1^*nR\big)\\&\leqslant\sup_{z\in B_{V_n}(0,c_2^*nK)\cap V_n}\tilde{\mathbb{P}}_z^{(n)}\big(\tau_{B_{V_n}(0,c_1^*nR)}(\tilde X_\cdot^{(n)})\leqslant n^\alpha T\big)\\&\leqslant\sup_{z\in B_{V_n}(0,c_1^*nR/2)\cap V_n}\tilde{\mathbb{P}}_z^{(n)}\big(\tau_{B_{V_n}(z,c_1^*nR/2)}(\tilde X_\cdot^{(n)})\leqslant n^\alpha T\big)\\&\leqslant c_1\left(\frac{n^\alpha T}{(c_1^*nR/2)^\alpha}\right)^{1/3}=c_2\left(\frac{T}{R^\alpha}\right)^{1/3},\end{aligned}$$

which implies

$$\lim_{R\to\infty}\limsup_{n\to\infty}\mathbb{P}^{(n)}_{x_n}(\sup_{s\in[0,T]}\rho(X^{(n)}_s,0)>R)\leqslant\lim_{R\to\infty}c_2\left(\frac{T}{R^\alpha}\right)^{1/3}=0.$$

This proves (4.19).

Next, let $\{\tau_n\}_{n\geqslant 1}$ be a sequence of stopping time such that $\tau_n\leqslant T$, and $\{\varepsilon_n\}_{n\geqslant 1}$ be a sequence such that $\lim_{n\to\infty}\varepsilon_n=0$. By the strong Markov property, for every $\eta>0$ small enough and $R\geqslant 1$ large enough,

$$\begin{aligned}
&\mathbb{P}^{(n)}_{x_n}\big(\rho(X^{(n)}_{\tau_n+\varepsilon_n},X^{(n)}_{\tau_n})>\eta\big)=\mathbb{E}^{(n)}_{x_n}\big[\mathbb{P}^{(n)}_{X^{(n)}_{\tau_n}}\rho(X^{(n)}_{\varepsilon_n},X^{(n)}_0)>\eta)\big]\\
&\leqslant\sup_{z\in B_F(0,R)\cap V_n}\mathbb{P}^{(n)}_z\big(\rho(X^{(n)}_{\varepsilon_n},X^{(n)}_0)>\eta\big)+\mathbb{P}^{(n)}_{x_n}\big(\sup_{s\in[0,T]}\rho(X^{(n)}_s,0)>R\big)\\
&\leqslant\sup_{z\in B_{V_n}(0,(c_2^*nR)\wedge(C_0'nr_F))\cap V_n}\tilde{\mathbb{P}}^{(n)}_z\big(\rho_n(\tilde X^{(n)}_{n^\alpha\varepsilon_n},\tilde X^{(n)}_0)>c_1^*n\eta\big)+\mathbb{P}^{(n)}_{x_n}\big(\sup_{s\in[0,T]}\rho(X^{(n)}_s,0)>R\big)\\
&\leqslant\sup_{z\in B_{V_n}(0,(c_2^*nR)\wedge(C_0'nr_F))\cap V_n}\tilde{\mathbb{P}}^{(n)}_z\big(\tau_{B_{V_n}(z,c_1^*n\eta)}(\tilde X^{(n)}_\cdot)\leqslant n^\alpha\varepsilon_n\big)+\mathbb{P}^{(n)}_{x_n}\big(\sup_{s\in[0,T]}\rho(X^{(n)}_s,0)>R\big),
\end{aligned}$$

where in the second inequality we have used the fact that $c_1^*n\rho(x,y)\leqslant\rho_n(x,y)\leqslant c_2^*n\rho(x,y)$ for $x,y\in V_n$, due to (4.2). Taking n large enough and η small enough such that $c_2^*nR>R_1$ and $(c_2^*nR)^\delta\leqslant c_1^*n\eta\leqslant(c_2^*nR)\wedge(C_0'nr_F)$. Then, it follows from (4.22) that

$$\begin{aligned}
&\sup_{z\in B_{V_n}(0,(c_2^*nR)\wedge(C_0'nr_F))\cap V_n}\tilde{\mathbb{P}}^{(n)}_z\big(\tau_{B_{V_n}(z,c_1^*n\eta)}(\tilde X^{(n)}_\cdot)\leqslant n^\alpha\varepsilon_n\big)\\
&\leqslant\sup_{z\in B_{V_n}(0,(c_2^*nR)\wedge(C_0'nr_F))\cap V_n}\tilde{\mathbb{P}}^{(n)}_z\big(\tau_{B_{V_n}(z,c_1^*n\eta)}(\tilde X^{(n)}_\cdot)\leqslant(n^\alpha\varepsilon_n)\vee(2c_1^*n\eta)^{\theta'\alpha}\big)\\
&\leqslant c_1\left(\frac{(n^\alpha\varepsilon_n)\vee(2c_1^*n\eta)^{\theta'\alpha}}{(c_1^*n\eta)^\alpha}\right)^{1/3}\leqslant c_3\left((\varepsilon_n\eta^{-\alpha})\vee(n\eta)^{-(1-\theta')\alpha}\right)^{1/3}.
\end{aligned}$$

Combining both estimates above with (4.19), we obtain (4.20).

Step (2): Now it suffices to show that any finite dimensional distribution of $\mathbb{P}^{(n)}_{x_n}$ converges to that of $\mathbb{P}^Y_x$. We first claim that for any fixed $t>0$, $f\in C_\infty(F)\cap L^2(F;m)$ and a sequence $\{z_n:z_n\in V_n\}_{n=1}^\infty$ with $\lim_{n\to\infty}z_n=z\in F$,

$$\lim_{n\to\infty}E_n\big(P^{(n)}_tf\big)(z_n)=P^Y_tf(z),\tag{4.23}$$

where $C_\infty(F)$ denotes the set of continuous functions on F vanishing at infinity.

Indeed, according to assumption **(Mos.)**, Proposition 4.3 and (4.5), there are a subsequence of $\{\hat P^{(n)}_tf:n\geqslant 1\}$ (we still denote it by $\{\hat P^{(n)}_tf:n\geqslant 1\}$ for simplicity) and a sequence $\{y_k\in\cup_{n\geqslant 1}\cup_{x\in V_n}\mathrm{Int}(U_n(x)):k\geqslant 1\}$ such that (i) $y_k\neq z$ and $\lim_{k\to\infty}y_k=z$; (ii) for every $k\geqslant 1$,

$$\lim_{n\to\infty}\hat P^{(n)}_tf(y_k)=P^Y_tf(y_k).\tag{4.24}$$

For every $k\geqslant 1$ and $t>0$, we have

$$\begin{aligned}
&|E_n(P^{(n)}_tf)(z_n)-P^Y_tf(z)|\\
&\leqslant|\hat P^{(n)}_tf(y_k)-P^Y_tf(y_k)|+|\hat P^{(n)}_tf(y_k)-E_n(P^{(n)}_tf)(y_k)|\\
&\quad+|E_n(P^{(n)}_tf)(y_k)-E_n(P^{(n)}_tf)(z_n)|+|P^Y_tf(z)-P^Y_tf(y_k)|\\
&=:|\hat P^{(n)}_tf(y_k)-P^Y_tf(y_k)|+\sum_{i=1}^3J_{i,n,k}.
\end{aligned}\tag{4.25}$$

Recall that $\hat{P}_t^{(n)}f(x) = E_n(P_t^{(n)}(\pi_n(f)))(x)$ for all $x \in F$. By the definition of π_n,

$$\lim_{n\to\infty} \sup_{z\in V_n} |\pi_n(f)(z) - f(z)| = 0$$

for any $f \in C_\infty(F)$. Hence,

$$\begin{aligned}\lim_{n\to\infty} \sup_{k\geqslant 1} J_{1,n,k} &= \lim_{n\to\infty} \sup_{k\geqslant 1} |E_n(P_t^{(n)}(\pi_n(f)))(y_k) - E_n(P_t^{(n)}f)(y_k)| \\ &\leqslant \lim_{n\to\infty} \sup_{z\in V_n} |\pi_n f(z) - f(z)| = 0,\end{aligned}$$

where in the last inequality we used the contractivity of $(P_t^{(n)})_{t\geqslant 1}$ in $L^\infty(V_n; m_n)$.

In the following, for any $x \in F$, let $[x]_n \in V_n$ be such that $x \in U_n([x]_n)$ and $\rho(x, [x]_n) \leqslant c_3^* n^{-1}$, due to (3) in Lemma 4.1. For any $n \geqslant 1$ and $z \in V_n$, noticing that $\big(\tilde{X}_{n^\alpha t}^{(n)}\big)_{t\geqslant 0}$ has the same distribution as $\big(\mathbf{P_n}(X_t^{(n)})\big)_{t\geqslant 0}$,

$$E_n(P_t^{(n)}f)(z) = P_t^{(n)}f([z]_n) = \mathbb{E}_{[z]_n}^{(n)}[f(X_t^{(n)})] = \tilde{\mathbb{E}}_{[z]_n}^{(n)}[f(\tilde{X}_{n^\alpha t}^{(n)})] = \tilde{P}_{n^\alpha t}^{(n)} f([z]_n),$$

where $\tilde{P}_t^{(n)}f(\cdot) := \tilde{\mathbb{E}}_\cdot^{(n)}[f(\tilde{X}_t^{(n)})]$ is the Markov semigroup of $\tilde{X}^{(n)} := (\tilde{X}_t^{(n)})_{t\geqslant 0}$. As mentioned above, due to assumption **(Wea.)** and Lemma 4.4, we can apply Theorem 3.8 (also thanks to Remark 3.10) to obtain that there are constants $\delta \in (\theta, 1)$ and $R_1 \geqslant 1$ such that for all $R_1 < R < C_0' n r_F$, (3.22) holds for every $\{\tilde{X}^{(n)}\}_{n\geqslant 1}$ and with constants independent of n. Let $C_0 > 0$ be the constant in (3.14). For fixed $T > 0$, we define $H_{T,n,f}(s,x) = \tilde{P}_{1+n^\alpha T - s}^{(n)} f(x)$, which is a parabolic function on $Q_{V_n}\big(0, y, (2^{-1}C_0^{-1}n^\alpha T)^{1/\alpha}\big)$ for each $y \in V_n$. Take K large enough such that $K > (2^{-1}C_0^{-1}t)^{1/\alpha}$, $R_1 < nK < C_0' n r_F$ and $z_n \in B_{V_n}(0, nK)$ for all $n \geqslant 1$. According the facts that $y_k \to z$ as $k \to \infty$ and $y_k \neq z$ for all $k \geqslant 1$, for any fixed $t > 0$, we can choose $k \geqslant 1$ large enough such that $0 < \varepsilon_k < \rho(y_k, z) \leqslant (4c_2^*)^{-1}((2^{-1}C_0^{-1}t)^{1/\alpha} \wedge (2^{-1}C_0' r_F))$, where ε_k is a positive constant with $\lim_{k\to\infty} \varepsilon_k = 0$, and $c_2^* > 0$ is the constant such that $\rho_n(x,y) \leqslant c_2^* n^{-1}\rho(x,y)$ for any $x, y \in V_n$. Furthermore, for these fixed k and t, we take n large enough such that $(nK)^\delta \leqslant r_n \leqslant nK$, $\rho(z_n, z) \leqslant (4c_2^*)^{-1} n^{-1} r_n$ and $n\varepsilon_k \geqslant 4(c_1^*)^{-1} r_n^\delta$, where $r_n := (2^{-1}C_0^{-1}n^\alpha t)^{1/\alpha} \wedge (2^{-1}C_0' n r_F)$. Hence,

$$\begin{aligned}\rho_n\big([z_n]_n, [y_k]_n\big) &\geqslant c_1^* n \rho\big([z_n]_n, [y_k]_n\big) \geqslant c_1^* n\big(\rho(z, y_k) - \rho\big(y_k, [y_k]_n\big) - \rho\big(z, [z]_n\big)\big) \\ &\geqslant c_1^* n\varepsilon_k - 2c_1^* c_3^* \geqslant r_n^\delta,\end{aligned}$$

$$\begin{aligned}\rho_n\big([z]_n, [y_k]_n\big) &\leqslant c_2^* n \rho\big([z]_n, [y_k]_n\big) \leqslant c_2^* n\big(\rho\big(z, y_k\big) + \rho\big(z, [z]_n\big) + \rho\big(y_k, [y_k]_n\big)\big) \\ &\leqslant 4^{-1} r_n + 2c_2^* c_3^* \leqslant 2^{-1} r_n\end{aligned}$$

and

$$\begin{aligned}\rho_n\big([z]_n, [z_n]_n\big) &\leqslant c_2^* n \rho([z]_n, [z_n]_n) \leqslant c_2^* n\big(\rho(z, z_n) + \rho(z, [z]_n) + \rho(z_n, [z_n]_n)\big) \\ &\leqslant 4^{-1} r_n + 2c_2^* c_3^* \leqslant 2^{-1} r_n,\end{aligned}$$

where we used the fact that $\rho(y, [y]_n) \leqslant c_3^* n^{-1}$ for all $y \in F$. Note that since $z_n \in V_n$, $[z_n]_n = z_n$. Then as a summary, $(nK)^\delta \leqslant r_n \leqslant nK$, $z_n \in B_{V_n}(0, nK)$, and $[z_n]_n, [y_k]_n \in Q_{V_n}\big(0, [z]_n, r_n\big)$ with $\rho_n\big([z_n]_n, [y_k]_n\big) \geqslant r_n^\delta$. Now, applying (3.22) to the parabolic function $H_{t,n,f}$ on $Q_{V_n}(0, [z]_n, r_n)$, we can obtain that

$$\begin{aligned}&|\tilde{P}_{n^\alpha t}^{(n)} f([y_k]_n) - \tilde{P}_{n^\alpha t}^{(n)} f([z_n]_n)| \\ &= |H_{t,n,f}(1, n[y_k]_n) - H_{t,n,f}(1, n[z_n]_n)| \leqslant c_4 \|\tilde{P}_{n^\alpha t}^{(n)} f\|_\infty \left|\frac{\rho_n\big([y_k]_n, [z_n]_n\big)}{r_n}\right|^\beta \\ &\leqslant c_5(t)\|f\|_\infty \rho\big([y_k]_n, [z_n]_n\big)^\beta \leqslant c_6(t)\|f\|_\infty \big(\rho(y_k, z)^\beta + n^{-\beta}\big).\end{aligned}$$

This yields immediately that

$$\limsup_{n\to\infty} J_{2,n,k} = \limsup_{n\to\infty} |\tilde{P}_{n^\alpha t}^{(n)} f([y_k]_n) - \tilde{P}_{n^\alpha t}^{(n)} f([z_n]_n)|$$

$$\leqslant c_6(t)\limsup_{n\to\infty}\|f\|_\infty\big(\rho(y_k,z_n)^\beta+n^{-\beta}\big)=c_7(t)\|f\|_\infty\rho(y_k,z)^\beta.$$

According to [23, Theorem 4.14], $J_{3,n,k}\leqslant c_8(t)\|f\|_\infty\rho(y_k,z)^\beta$.

Combining all estimates with (4.25) and (4.24), we arrive at that

$$\limsup_{n\to\infty}\big|E_n(P_t^{(n)}f)(z_n)-P_t^Yf(z)\big|\leqslant c_9(t)\|f\|_\infty\rho(y_k,z)^\beta,$$

where $c_9(t)>0$ is independent of k. Note that k is arbitrary, letting $k\to\infty$ in the last inequality, then we prove (4.23). In particular, according to [20, Lemma 2.7], (4.23) implies that for every compact set $K\subseteq F$,

$$\limsup_{n\to\infty}\sup_{x\in K}|E_n(P_t^{(n)}f)(x)-P_t^Yf(x)|=0. \tag{4.26}$$

Next, for any $f_1,f_2\in C_\infty(F)$, $0<s<t\leqslant T$ and any sequence $x_n\in V_n$ with $\lim_{n\to\infty}x_n=x\in F$,

$$\begin{aligned}&\mathbb{E}_{x_n}^{(n)}\big[f_1(X_s^{(n)})f_2(X_t^{(n)})\big]=\mathbb{E}_{x_n}^{(n)}\big[f_1(X_s^{(n)})P_{t-s}^{(n)}f_2(X_s^{(n)})\big]\\&=\mathbb{E}_{x_n}^{(n)}\big[f_1(X_s^{(n)})P_{t-s}^Yf_2(X_s^{(n)})\big]+\mathbb{E}_{x_n}^{(n)}\big[f_1(X_s^{(n)})\big(P_{t-s}^{(n)}f_2(X_s^{(n)})-P_{t-s}^Yf_2(X_s^{(n)})\big)\big]\\&=:J_{1,n}+J_{2,n}.\end{aligned}$$

Set $g(z)=f_1(z)P_{t-s}^Yf_2(z)$. Then $g\in C_\infty(F)$, due to the C_∞-Feller property of the process Y, see [23, Theorem 1.1]. Then, according to (4.23), we have

$$\lim_{n\to\infty}J_{1,n}=\lim_{n\to\infty}P_t^{(n)}g(x_n)=P_s^Yg(x)=\mathbb{E}_x^Y\big[f_1(Y_s)f_2(Y_t)\big].$$

On the other hand, for any $t>0$, $R>2K$ and n large enough,

$$J_{2,n}\leqslant\|f_1\|_\infty\sup_{z\in B_F(0,R)}\big|E_n(P_{t-s}^{(n)}f_2)(z)-P_{t-s}^Yf_2(z)\big|+\|f_1\|_\infty\|f_2\|_\infty\mathbb{P}_{x_n}^{(n)}\big(\sup_{s\in[0,t]}\rho(X_s^{(n)},0)>R\big),$$

By (4.19) and (4.26), we let $n\to\infty$ and then $R\to\infty$ in the last inequality, yielding that $\lim_{n\to\infty}J_{2,n}=0$. Combining all above estimates, we prove that

$$\lim_{n\to\infty}\mathbb{E}_{x_n}^{(n)}\big[f_1(X_s^{(n)})f_2(X_t^{(n)})\big]=\mathbb{E}_x^Y\big[f_1(Y_s)f_2(Y_t)\big].$$

Following the same arguments as above and using the induction procedure, we can obtain from [30, Chapter 3; Proposition 4.4 and Theorem 7.8(b)] that any finite dimensional distribution of $\mathbb{P}_{x_n}^{(n)}$ converges to $\mathbb{P}_x^Y$. The proof is finished. □

Remark 4.6. As shown in the proof of Theorem 4.5 above, the role of adopting the generalized Mosco convergence is to identify the limit process in the L^2 sense. Actually, according to [22, Theorem 5.1], under Assumption **(Mos.)** only, any finite dimensional distribution of $X^{(n)}$ converges to that of Y, when the initial distribution is absolutely continuous with respect to the reference measure m. Thus, Theorem 4.5 improves this weak convergence for any initial distribution. We emphasize that such improvement is highly non-trivial, see [31] for discussions on the uniformly elliptic case by using heat kernel estimates. Here, we will make use of the Hölder regularity of parabolic functions on large scale (Theorem 3.8). This is much weaker than the approach used in [20, Proposition 2.8], where the Hölder regularity of parabolic functions is assumed to be satisfied on the whole space.

5. Random conductance model: quenched invariance principle

We will apply results from Section 4 to study the quenched invariance principle for random conductance models.

5.1. **Quenched invariance principle for stable-like processes on d-sets.** Suppose that (F,ρ,m) is a metric measure space satisfying assumption **(MMS)**. By Lemma 4.1, we have a sequence of graphs with measure $\{(V_n,\rho_n,m_n): n\geqslant 1\}$ that approximate (F,ρ,m). In this part, we further assume the following:

- (i) $\rho(\cdot,\cdot)$ is a metric with dilation; namely, there exists another distance $\bar\rho$ on F such that
  - (i′) for all $x,y\in F$, $C_1\bar\rho(x,y)\leqslant\rho(x,y)\leqslant C_2\bar\rho(x,y)$ holds for some constants $0<C_1<C_2<\infty$.
  - (i″) for each $x,y\in F$, $i\in\{-1,1\}$ and $n\in\mathbb{N}$, there are $x^{(n^i)},y^{(n^i)}\in F$ (we write $n^ix:=x^{(n^i)}, n^iy:=y^{(n^i)}$ for notational simplicity) such that $\bar\rho(n^ix,n^iy)=n^i\bar\rho(x,y)$.
- (ii) There exists $0\in V_1\subset F$ such that $n^i0=0$ for all $i\in\{-1,1\}$ and $n\in\mathbb{N}$.
- (iii) $V_n=n^{-1}V_1:=\{n^{-1}z: z\in V_1\}$, and F is a closure of $\cup_{n\geqslant1}V_n$. Moreover, $nV_1\subset V_1$ and $\mu_n(A)=\mu_1(nA)$ for all $A\subset V_n$ and $n\in\mathbb{N}$, where μ_n denotes the counting measure on V_n.

We note that, due to (4.4), for any $n\in\mathbb{N}$, there exists a measurable function K_n on V_n such that $m_n=n^{-d}K_n\,\mu_n$ and

$$0<C_3\leqslant K_n\leqslant C_4<\infty, \tag{5.1}$$

where μ_n denotes the counting measure on V_n, and C_3,C_4 are constants independent of n.

Remark 5.1. Obviously conditions (i′) and (i″) in assumption (i) above hold true for a bounded Lipschitz domain $F\subset\mathbb{R}^d$. For simplicity, in the arguments below we assume that $\rho(n^ix,n^iy)=n^i\rho(x,y)$ for all $n\in\mathbb{N}$; otherwise, we can express Dirichlet forms $(D^\omega_{V_n},\mathscr{F}^\omega_n)$ and $(D_0,\mathscr{F}_0)$ below with ρ, $w^{(n)}_{x,y}(\omega)$ and $c(x,y)$ replaced by $\bar\rho$, $\bar w^{(n)}_{x,y}(\omega):=\frac{\bar\rho(x,y)^{d+\alpha}}{\rho(x,y)^{d+\alpha}}w^{(n)}_{x,y}(\omega)$ and $\bar c(x,y):=\frac{\bar\rho(x,y)^{d+\alpha}}{\rho(x,y)^{d+\alpha}}c(x,y)$, respectively. Hence, by applying the arguments below for $\bar\rho$, $\bar w^{(n)}_{x,y}(\omega)$ and $\bar c(x,y)$, we can still obtain the quenched invariance principle for $(X^\omega_t)_{t\geqslant0}$.

Let $\{w_{x,y}(\omega): x,y\in V_1\}$ be a sequence of random variables defined on a probability space $(\Omega,\mathscr{F},\mathbb{P})$ such that $w_{x,y}(\omega)=w_{y,x}(\omega)$ and $w_{x,y}(\omega)\geqslant0$ for all $x\neq y\in V_1$. For any $x\in V_n$, $m_n(x):=m_n(\{x\})=n^{-d}K_n(x)$. Define

$$w^{(n)}_{x,y}(\omega)=\frac{K_1(nx)K_1(ny)}{K_n(x)K_n(y)}w_{nx,ny}(\omega). \tag{5.2}$$

We consider the following class of Dirichlet forms

$$D^\omega_{V_n}(f,f)=\frac12\sum_{x,y\in V_n}(f(x)-f(y))^2\frac{w^{(n)}_{x,y}(\omega)}{\rho(x,y)^{d+\alpha}}m_n(x)m_n(y),\quad f\in\mathscr{F}^\omega_n,$$
$$\mathscr{F}^\omega_n=\{f\in L^2(V_n;m_n): D^\omega_{V_n}(f,f)<\infty\}.$$

Let $X^{V_1,\omega}$ be the strong Markov process on V_1 associated with $(D^\omega_{V_1},\mathscr{F}^\omega_1)$. Then, it is easy to show that for a.s. $\omega\in\Omega$, $(D^\omega_{V_n},\mathscr{F}^\omega_n)$ generates a Markov process $X^{(n),\omega}=(X^{(n),\omega}_t)_{t\geqslant0}$ such that $X^{(n),\omega}_t=n^{-1}X^{V_1,\omega}_{n^\alpha t}$ for all $t\geqslant0$. Here and what follows, = means two processes enjoy the same distribution.

Now, consider the Dirichlet form $(D_0,\mathscr{F}_0)$ given by (4.6), i.e.,

$$D_0(f,f)=\frac12\int_{\{F\times F\setminus\mathrm{diag}\}}\big(f(x)-f(y)\big)^2\frac{c(x,y)}{\rho(x,y)^{d+\alpha}}\,m(dx)\,m(dy),\quad f\in\mathscr{F}_0,$$
$$\mathscr{F}_0=\{f\in L^2(F;m): D_0(f,f)<\infty\},$$

where $\alpha\in(0,2)$, $\mathrm{diag}:=\{(x,y)\in F\times F: x=y\}$, and $c: F\times F\to(0,\infty)$ is a symmetric continuous function such that $0<c_1\leqslant c(x,y)\leqslant c_2<\infty$ for all $(x,y)\in F\times F\setminus\mathrm{diag}$ and some constants c_1,c_2. We suppose that assumption **(Dir.)** holds. Let $Y:=((Y_t)_{t\geqslant0},(\mathbb{P}^Y_x)_{x\in F})$ be a α-stable-like process on F.

We next apply Theorem 4.5 to prove the quenched invariance principle for $(X^\omega_t)_{t\geqslant0}$ under some assumptions on $w_{x,y}$. We first assume that the following holds.

Assumption (Den.)

(i) $\mathbb{E}[w_{x,y}] = J_1(x,y)$ *and* $\mathbb{E}[w_{x,y}^{-1}\mathbb{1}_{\{w_{x,y}>0\}}] = J_2(x,y)$ *for any* $x,y \in V_1$*, where* $0 < C_1 < J_i(x,y) < C_2 < \infty$ *for all* $i = 1,2$ *and* $x,y \in V_1$*.*

(ii) *For every compact set* $S \subseteq F$*,*

$$\lim_{n\to\infty}\Big[\sup_{x,y\in S}\Big|J_1(n[x]_n,n[y]_n)\cdot\frac{K_1(n[x]_n)K_1(n[y]_n)}{K_n([x]_n)K_n([y]_n)} - c(x,y)\Big|\Big] = 0, \tag{5.3}$$

where $[x]_n \in V_n$ *is the element such that* $x \in U_n([x]_n)$*.*

Remark 5.2. Obviously when $F = \mathbb{R}^d$, it follows from (5.3) that for any $x \neq y \in \mathbb{R}^d$ and $s \neq 0$, $c(x,y) = c(sx,sy)$, which, along with the fact that $K_n \equiv 1$ for all $n \in \mathbb{N}$ as mentioned in Remark 5.1, implies that the limit process $(Y_t)_{t\geqslant 0}$ satisfies the scaling invariant property as follows

$$\mathbb{P}^Y_{\varepsilon^{-1}x}\left((\varepsilon Y_{t\varepsilon^{-\alpha}})_{t\geqslant 0}\in A\right) = \mathbb{P}^Y_x\left((Y_t)_{t\geqslant 0}\in A\right)$$

for any $x \in \mathbb{R}^d$, $\varepsilon > 0$ and $A \subset \mathscr{D}([0,\infty);\mathbb{R}^d)$.

For $\varepsilon > 0$, $x \in V_1$, $R, r > 0$, $c_0 > 1/2$, $c_0^* \geqslant 2$ and a sequence of bounded functions $\{h_n\}_{n\geqslant 1}$ on $V_1 \times V_1$, define

$$p_1(r,R,\varepsilon) = \mathbb{P}\Big(\Big|\sum_{x,y\in V_1:\rho(0,x)\leqslant R,\rho(x,y)\leqslant r}(w_{x,y}-J_1(x,y))\Big| > \varepsilon r^d R^d\Big),$$

$$p_2(x,r,\varepsilon) = \mathbb{P}\Big(\Big|\sum_{y\in V_1:\rho(x,y)\leqslant r}\big(w_{x,y}-J_1(x,y)\big)\Big| > \varepsilon r^d\Big),$$

$$p_3(x,r,\varepsilon) = \mathbb{P}\Big(\Big|\sum_{y\in V_1:\rho(x,y)\leqslant r}\frac{(w_{x,y}-J_1(x,y))}{\rho(x,y)^{d+\alpha-2}}\Big| > \varepsilon r^{2-\alpha}\Big),$$

$$p_3^*(x,r,\varepsilon) = \mathbb{P}\Big(\Big|\sum_{y\in V_1:\rho(x,y)\leqslant r}\frac{(w_{x,y}-J_1(x,y))}{\rho(x,y)^{d+\alpha-1}}\Big| > \varepsilon r^{1-\alpha}\Big), \quad \alpha\in(0,1),$$

$$p_4(x,r,c_0^*,\varepsilon) = \mathbb{P}\Big(\Big|\sum_{y\in V_1:\rho(x,y)\leqslant c_0^* r}\big(w_{x,y}^{-1}-J_2(x,y)\big)\Big| > \varepsilon_0 r^d\Big),$$

$$p_5^{(n)}(x,R,r,h,\varepsilon) = \mathbb{P}\Big(\Big|\sum_{\substack{y\in B_F(0,nR)\cap V_1:\\ \rho(x,y)\geqslant nr}} h_n(x,y)\frac{(w_{x,y}-J_1(x,y))}{\rho(x,y)^{d+\alpha}}\Big|^2 > \varepsilon(nr)^{-2\alpha}\Big),$$

$$p_6(x,z,r,c_0) = \mathbb{P}\left(\frac{\mu_1\{y\in V_1:\rho(y,x)\leqslant r, w_{y,z}>0\}}{\mu_1\{y\in V_1:\rho(y,x)\leqslant r\}} \leqslant C_4 c_0 C_3^{-1}\right),$$

where $C_3 \leqslant C_4$ are positive constants in (5.1).

Theorem 5.3. *Suppose that assumption* **(Den.)** *holds, and that there exists a constant* $\theta \in (0,1)$ *such that*

(i) *for any* ε_0 *and* ε *small enough, any* N *large enough, and any sequence of bounded function* $\{h_n\}_{n\geqslant 1}$ *on* $V_1 \times V_1$ *with* $\sup_{n\geqslant 1}\|h_n\|_\infty < \infty$*,*

$$\sum_{R=1}^{\infty}\sum_{r=1}^{R} p_1(r,R,\varepsilon_0) < \infty, \tag{5.4}$$

$$\sum_{R=1}^{\infty}\sum_{x\in B_F(0,6R)\cap V_1}\sum_{r=R^\theta/2}^{\infty} p_2(x,r,\varepsilon_0) < \infty, \tag{5.5}$$

and

$$\sum_{n=1}^{\infty}\sum_{x\in B_F(0,nN)\cap V_1} p_5^{(n)}(x,N,\varepsilon,h_n,\varepsilon_0) < \infty. \tag{5.6}$$

(ii) *any ε_0 small enough,*

$$\sum_{R=1}^{\infty} \sum_{x \in B_F(0,6R) \cap V_1} p_3(x, R^\theta, \varepsilon_0) < \infty \tag{5.7}$$

and

$$\sum_{R=1}^{\infty} \sum_{x \in B_F(0,6R) \cap V_1} \sum_{r=R^\theta/2}^{2R} p_4(x, r, c_0^*, \varepsilon_0) < \infty, \tag{5.8}$$

for any fixed $c_0^* \geqslant 0$, as well as

$$\sum_{R=1}^{\infty} \sum_{x,z \in B_F(0,6R) \cap V_1} \sum_{r=R^\theta/2}^{2R} p_6(x, z, r, c_0) < \infty \tag{5.9}$$

for some fixed $c_0 > 1/2$.

When $\alpha \in (0,1)$, (5.7) can be replaced by

$$\sum_{R=1}^{\infty} \sum_{x \in B_F(0,6R) \cap V_1} p_3^*(x, R^\theta, \varepsilon_0) < \infty. \tag{5.10}$$

Then for $\mathbb{P}$-a.s. $\omega \in \Omega$ and any $\{x_n \in V_n : n \geqslant 1\}$ such that $\lim_{n\to\infty} x_n = x$ with some $x \in F$, it holds that for every $T > 0$, $\mathbb{P}_{x_n}^{(n),\omega}$ converges weakly to $\mathbb{P}_x^Y$ on the space of all probability measures on $\mathscr{D}([0,T];F)$, where $\mathbb{P}_{x_n}^{(n),\omega}$ denotes the distribution of process $X_t^{(n),\omega} = n^{-1} X_{n^\alpha t}^{V_1,\omega}$.

Theorem 5.3 immediately holds by applying Theorem 4.5, Lemmas 5.4 and 5.5 below to process $X_t^{(n),\omega}$. From now on, for simplicity we will assume that $K_n \equiv 1$ for all $n \in \mathbb{N}$ (in particular, $C_3 = C_4 = 1$ in (5.1)), since the proof directly works for general case with some mild changes due to the facts that $w_{x,y}^{(n)}(\omega) = \frac{K_1(nx)K_1(ny)}{K_n(x)K_n(y)} w_{nx,ny}(\omega)$ and $C^{-1} \leqslant \frac{K_1(nx)K_1(ny)}{K_n(x)K_n(y)} \leqslant C$ for all $x, y \in V_n$ and $n \in \mathbb{N}$ with some constant $C \geqslant 1$ independent of x, y, n.

Lemma 5.4. *Under assumption* (i) *in Theorem* 5.3, *for $\mathbb{P}$-a.s. $\omega \in \Omega$, Assumption* **(Mos.)** *holds for the conductance $\{w_{x,y}^{(n)}(\omega)\}$.*

Proof. Under (5.4), for any $\varepsilon_0 > 0$,

$$\begin{aligned}
&\sum_{R=1}^{\infty} \mathbb{P}\Big(\bigcup_{r=1}^{R} \Big\{\Big| \sum_{x,y \in V_1 : \rho(0,x) \leqslant R, \rho(x,y) \leqslant r} \big(w_{x,y} - J_1(x,y)\big)\Big| > \varepsilon_0 r^d R^d\Big\}\Big) \\
&\leqslant \sum_{R=1}^{\infty} \sum_{r=1}^{R} \mathbb{P}\Big(\Big| \sum_{x,y \in V_1 : \rho(0,x) \leqslant R, \rho(x,y) \leqslant r} \big(w_{x,y} - J_1(x,y)\big)\Big| > \varepsilon_0 r^d R^d\Big) = \sum_{R=1}^{\infty} \sum_{r=1}^{R} p_1(r, R, \varepsilon_0) < \infty.
\end{aligned}$$

Since $C_1 \leqslant J_1(x,y) \leqslant C_2$ for all $x, y \in V_1$ and some positive constants C_1 and C_2, by the Borel-Cantelli lemma, we know that, for $\mathbb{P}$-a.s. $\omega \in \Omega$, there exists a constant $R_0(\omega) \geqslant 1$ such that for every $R > R_0(\omega)$,

$$c_1 r^d R^d \leqslant \sum_{x,y \in V_1 : \rho(0,x) \leqslant R, \rho(x,y) \leqslant r} w_{x,y}(\omega) \leqslant c_2 r^d R^d, \quad \forall\, 1 \leqslant r \leqslant R,$$

where c_1, c_2 are positive constants independent of ω. Then, for any $0 < 2\eta < N$ and $nN > R_0(\omega)$, we have

$$\begin{aligned}
&n^{-2d} \sum_{x,y \in B_F(0,N) \cap V_n : 0 < \rho(x,y) \leqslant \eta} \frac{w_{nx,ny}(\omega)}{\rho(x,y)^{d+\alpha-2}} \\
&\leqslant n^{-d+\alpha-2} \sum_{k=0}^{[\log(n\eta)/\log 2]+1} \sum_{x,y \in V_1 : \rho(0,x) \leqslant nN \text{ and } 2^k \leqslant \rho(x,y) < 2^{k+1}} \frac{w_{x,y}(\omega)}{\rho(x,y)^{d+\alpha-2}}
\end{aligned}$$

$$\leqslant n^{-d+\alpha-2}\sum_{k=0}^{[\log(n\eta)/\log 2]+1}2^{-k(d+\alpha-2)}\sum_{x,y\in V_1:\rho(0,x)\leqslant nN \text{ and } 2^k\leqslant\rho(x,y)<2^{k+1}}w_{x,y}(\omega)$$

$$\leqslant c_3n^{-d+\alpha-2}\sum_{k=0}^{[\log(n\eta)/\log 2]+1}2^{-k(d+\alpha-2)}2^{(k+1)d}(nN)^d\leqslant c_4N^d\eta^{2-\alpha}.$$

This yields that (4.8) holds for $\mathbb{P}$-a.s. $\omega\in\Omega$.

According to (5.5), for every $\varepsilon_0>0$ small enough,

$$\begin{aligned}&\sum_{R=1}^\infty\mathbb{P}\Big(\bigcup_{x\in B_F(0,6R)\cap V_1}\bigcup_{r=R^\theta/2}^\infty\Big\{\Big|\sum_{y\in V_1:\rho(x,y)\leqslant r}\big(w_{x,y}-J_1(x,y)\big)\Big|>\varepsilon_0r^d\Big\}\Big)\\&\leqslant\sum_{R=1}^\infty\sum_{x\in B_F(0,6R)\cap V_1}\sum_{r=R^\theta/2}^\infty\mathbb{P}\Big(\Big\{\Big|\sum_{y\in V_1:\rho(x,y)\leqslant r}\big(w_{x,y}-J_1(x,y)\big)\Big|>\varepsilon_0r^d\Big\}\Big)\\&\leqslant\sum_{R=1}^\infty\sum_{x\in B_F(0,6R)}\sum_{r=R^\theta/2}^\infty p_2(x,r,\varepsilon_0)<\infty.\end{aligned}$$

Hence, by the Borel-Cantelli lemma, we can find a constant $R_1(\omega)>0$ such that for every $R>R_1(\omega)$, $x\in B_F(0,6R)$ and $r\geqslant R^\theta/2$, $\Big|\sum_{y\in V_1:\rho(x,y)\leqslant r}(w_{x,y}-J_1(x,y))\Big|\leqslant\varepsilon_0r^d$. Due to the fact that $0<C_1\leqslant J_1(x,y)\leqslant C_2<\infty$ for any $x,y\in V_1$ again, we arrive at that for all $R>R_1(\omega)$,

$$c_5r^d\leqslant\sum_{y\in V_1:\rho(x,y)\leqslant r}w_{x,y}\leqslant c_6r^d,\quad\forall\ x\in B_F(0,6R),\ r\geqslant R^\theta/2. \tag{5.11}$$

Therefore, by (5.11), for every $n,j\geqslant1$ large enough such that $2nN>R_1(\omega)$ and $j>N$,

$$\begin{aligned}&n^{-2d}\sum_{x,y\in B_F(0,N)\cap V_n:\rho(x,y)\geqslant j}\frac{w_{nx,ny}(\omega)}{\rho(x,y)^{d+\alpha}}\\&\leqslant n^{-d+\alpha}\sum_{x\in V_1:\rho(0,x)\leqslant nN}\sum_{y\in V_1:\rho(x,y)\geqslant nj}\frac{w_{x,y}(\omega)}{\rho(x,y)^{d+\alpha}}\\&\leqslant n^{-d+\alpha}\sum_{x\in V_1:\rho(0,x)\leqslant nN}\sum_{k=\left[\frac{\log(nj)}{\log 2}\right]}^\infty2^{-k(d+\alpha)}\sum_{y\in V_1:\rho(x,y)\leqslant2^{k+1}}w_{x,y}(\omega)\\&\leqslant c_7n^{-d+\alpha}\sum_{x\in V_1:\rho(0,x)\leqslant nN}\sum_{k=\left[\frac{\log(nj)}{\log 2}\right]}^\infty2^{-k(d+\alpha)}2^{(k+1)d}\leqslant c_8N^dj^{-\alpha}.\end{aligned}$$

Hence, letting $n\to\infty$ first and then $j\to\infty$, we prove that (4.9) holds for $\mathbb{P}$-a.s. $\omega\in\Omega$.

Given $f\in\mathrm{Lip}_c(F)$, let

$$h_n(x,y):=\begin{cases}f(n^{-1}y)-f(n^{-1}x), & n^{-1}x,n^{-1}y\in V_n,\\0, & \text{otherwise.}\end{cases}$$

Applying (5.6) to $h_n(x,y)$ and using the Borel-Cantelli lemma, we know that for any ε and ε_0 small enough, and N large enough, there exists a constant $n_0(\omega)>0$ (which may depend on ε_0, ε, N and f) such that for every $n>n_0(\omega)$ and $x\in B_F(0,nN)$,

$$\Big|\sum_{y\in B_F(0,nN)\cap V_1:\rho(x,y)\geqslant n\varepsilon}\big(f(n^{-1}y)-f(n^{-1}x)\big)\frac{(w_{x,y}(\omega)-J_1(x,y))}{\rho(x,y)^{d+\alpha}}\Big|^2\leqslant\varepsilon_0(n\varepsilon)^{-2\alpha}.$$

Then, for n large enough such that $n\varepsilon > (nN)^\theta$, we have

$$\begin{aligned}
&n^{-d}\sum_{x\in B_F(0,N)\cap V_n}\left(\sum_{\substack{y\in B_F(0,N)\cap V_n:\\ \rho(x,y)>\varepsilon}}\big(f(x)-f(y)\big)\frac{\big(w_{nx,ny}(\omega)-J_1(nx,ny)\big)}{\rho(x,y)^{d+\alpha}}m_n(y)\right)^2\\
&=n^{-d+2\alpha}\sum_{x\in B_F(0,nN)\cap V_1}\left(\sum_{y\in B_F(0,nN)\cap V_1:\rho(x,y)>n\varepsilon}h_n(x,y)\frac{\big(w_{x,y}(\omega)-J_1(x,y)\big)}{\rho(x,y)^{d+\alpha}}\right)^2\\
&\leqslant n^{-d+2\alpha}\sum_{x\in B_F(0,nN)\cap V_1}\varepsilon_0(n\varepsilon)^{-2\alpha}\leqslant c_9N^d\varepsilon^{-2\alpha}\varepsilon_0.
\end{aligned}$$

On the other hand, due to (5.3), we can verify that every fixed $N>0$ and $\varepsilon>0$,

$$\begin{aligned}
&\lim_{n\to\infty}n^{-d}\sum_{x\in B_F(0,N)\cap V_n}\left(\sum_{\substack{y\in B_F(0,N)\cap V_n:\\ \rho(x,y)>\varepsilon}}\big(f(x)-f(y)\big)\frac{\big(J_1(nx,ny)-c(x,y)\big)}{\rho(x,y)^{d+\alpha}}m_n(y)\right)^2\\
&\leqslant 4\|f\|_\infty^2\varepsilon^{-2(d+\alpha)}\lim_{n\to\infty}n^{-3d}\sum_{x\in B_F(0,N)\cap V_n}\left(\sum_{y\in B_F(0,N)\cap V_n:\rho(x,y)>\varepsilon}\big|J_1(nx,ny)-c(x,y)\big|\right)^2\\
&\leqslant c_{10}\|f\|_\infty^2\varepsilon^{-2(d+\alpha)}N^d\lim_{n\to\infty}\left\{n^{-2d}\sum_{x,y\in B_F(0,N)\cap V_n}\big(J_1(nx,ny)-c(x,y)\big)^2\right\}=0.
\end{aligned}$$

Combining two estimates above, we can obtain that (4.10) holds for $\mathbb{P}$-a.s. $\omega\in\Omega$ by first letting $n\to\infty$ and then taking $\varepsilon_0\to 0$.

Since (4.11) can been proved in the similar way, we omit it here. □

Lemma 5.5. *Suppose that condition* (5.5) *and assumption* (ii) *in Theorem* 5.3 *hold. Then for* $\mathbb{P}$*-a.s.* $\omega\in\Omega$*, Assumption* **(Wea.)** *holds for the conductance* $\{w_{x,y}^{(n)}(\omega)\}$.

Proof. First, according to (5.9), the property $\mu_n(A)=\mu_1(nA)$ and the definitions of m_n and $w_{x,y}^{(n)}$, we can easily deduce from the Borel-Cantelli lemma that there is a constant $R_0(\omega)>0$ such that for any $R>R_0(\omega)$ and $R^\theta/2\leqslant r\leqslant R$, (4.15) holds.

By (5.7),

$$\begin{aligned}
&\sum_{R=1}^\infty\mathbb{P}\left(\bigcup_{x\in B_F(0,6R)\cap V_1}\left\{\Big|\sum_{y\in V_1:\rho(x,y)\leqslant R^\theta}\frac{\big(w_{x,y}-J_1(x,y)\big)}{\rho(x,y)^{d+\alpha-2}}\Big|>\varepsilon_0R^{\theta(2-\alpha)}\right\}\right)\\
&\leqslant\sum_{R=1}^\infty\sum_{x\in B_F(0,6R)\cap V_1}\mathbb{P}\left(\Big|\sum_{y\in V_1:\rho(x,y)\leqslant R^\theta}\frac{\big(w_{x,y}-J_1(x,y)\big)}{\rho(x,y)^{d+\alpha-2}}\Big|>\varepsilon_0R^{\theta(2-\alpha)}\right)\\
&=\sum_{R=1}^\infty\sum_{x\in B_F(0,6R)\cap V_1}p_3(x,R^\theta,\varepsilon_0)<\infty.
\end{aligned}$$

Hence, by the Borel-Cantelli lemma, there exists a constant $R_0(\omega)>0$ such that for any $R>R_0(\omega)$,

$$\sum_{y\in V_1:\rho(x,y)\leqslant R^\theta}\frac{w_{x,y}}{\rho(x,y)^{d+\alpha-2}}\leqslant c_1R^{\theta(2-\alpha)},\quad\forall\, x\in B_F(0,6R)\cap V_1. \tag{5.12}$$

Furthermore, using (5.11) and choosing ε_0 small enough and $R_0(\omega)$ large enough, we find that for every $R>R_0(\omega)$,

$$c_2^{-1}r^d\leqslant\sum_{y\in V_1:\rho(x,y)\leqslant r}w_{x,y}\leqslant c_2r^d,\quad\forall\, r>R^\theta/2,\ x\in B_F(0,6R)\cap V_1. \tag{5.13}$$

Combining this with (5.12), we see that for every $R > R_0(w)$, $x \in B_F(0, 6C_2R/n) \cap V_n$ and $R^\theta/2 \leqslant r \leqslant 2R$,

$$\begin{aligned} n^{-(d+\alpha-2)} & \sum_{y\in V_n:\rho(x,y)\leqslant C_2 r/n} \frac{w^{(n)}_{x,y}}{\rho(x,y)^{d+\alpha-2}} \\ \leqslant & \sum_{y\in V_1:\rho(x,y)<R^\theta/2} \frac{w_{x,y}}{\rho(x,y)^{d+\alpha-2}} + \sum_{k=[\log(R^\theta/2)/\log 2]}^{[\log(C_2 r)/\log 2]+1} 2^{-k(d+\alpha-2)} \Big(\sum_{y\in V_1: 2^k<\rho(x,y)\leqslant 2^{k+1}} w_{x,y} \Big) \\ \leqslant & c_4 \Big(R^{\theta(2-\alpha)} + \sum_{k=[\log(R^\theta/2)/\log 2]}^{[\log(C_2 r)/\log 2]+1} 2^{-k(\alpha-2)} \Big) \leqslant c_5 r^{2-\alpha}. \end{aligned}$$

Therefore, (4.14) holds for $\mathbb{P}$-a.s. $\omega \in \Omega$.

Due to (5.13) again, we know that for every $R > R_0(\omega)$, $x \in B_F(0, 6C_2R/n) \cap V_n$ and $r > R^\theta/2$,

$$\begin{aligned} n^{-(d+\alpha)} \sum_{y\in V_n:\rho(x,y)>C_1 r/n} \frac{w^{(n)}_{x,y}}{\rho_n(x,y)^{d+\alpha}} & \leqslant \sum_{k=[\log(C_1 r)/\log 2]}^{\infty} 2^{-k(d+\alpha)} \Big(\sum_{y\in V_1:2^k<\rho(x,y)\leqslant 2^{k+1}} w_{x,y} \Big) \\ & \leqslant c_6 \sum_{k=[\log(C_1 r)/\log 2]}^{\infty} 2^{-k(d+\alpha)} 2^{d(k+1)} \leqslant c_7 r^{-\alpha}, \end{aligned}$$

which implies that (4.18) is satisfied for $\mathbb{P}$-a.s. $\omega \in \Omega$.

Following the arguments above, and using (5.8) and the Borel-Cantelli lemma, we can obtain that (4.16) holds for $\mathbb{P}$-a.s. $\omega \in \Omega$. On the other hand, when $\alpha \in (0,1)$, we can use (5.10) to prove that (4.17) holds for $\mathbb{P}$-a.s. $\omega \in \Omega$. The proof is complete. □

5.2. **Examples.** As an application of Theorem 5.3, we consider three examples. One is a lattice on a half/quarter space, and other two are time-change of stable-like processes and a fractal graph respectively.

5.2.1. *Lattice on a half/quarter space.* Let $F := \mathbb{R}_+^{d_1} \times \mathbb{R}^{d_2}$ with $d_1, d_2 \in \mathbb{N} \cup \{0\}$, and ρ and m be the Euclidean distance and the Lebesgue measure respectively, which clearly satisfy assumption **(MMS)**. Therefore the process Y associated with Dirichlet form $(D_0, \mathscr{F}_0)$ is a reflected stable-like process on F, see e.g. [23]. Obviously $(D_0, \mathscr{F}_0)$ satisfies assumption **(Dir.)**. Here we will take $V_1 = \mathbb{L} := \mathbb{Z}_+^{d_1} \times \mathbb{Z}^{d_2}$, and $K_n \equiv 1$ for all $n \in \mathbb{N}$. Note that the scaling limit of $n^{-1}\mathbb{L}$ is F.

Let $E_{\mathbb{L}}$ be the set of edges associated with $\mathbb{L}$, $\{w_{x,y} : (x,y) \in E_{\mathbb{L}}\}$ be a sequence of non-negative independent random variables, and $(X_t^\omega)_{t\geqslant 0}$ be the Markov process with infinitesimal generator $L_{\mathbb{L}}^\omega$ defined by (1.1). Obviously $(X_t^\omega)_{t\geqslant 0}$ is the symmetric Hunt process associated with the Dirichlet form $(D^\omega_{V_1}, \mathscr{F}^\omega_1)$ with $V_1 = \mathbb{L}$ and $w^{(1)}_{x,y}(\omega) = w_{x,y}(\omega)$.

Proposition 5.6. *Let $d := d_1 + d_2 > 4 - 2\alpha$. Suppose that $\{w_{x,y} : (x,y) \in E_{\mathbb{L}}\}$ is a sequence of non-negative independent random variables satisfying that*

$$\sup_{x,y\in\mathbb{L},x\neq y} \mathbb{P}\big(w_{x,y} = 0\big) < 2^{-4} \tag{5.14}$$

and

$$\sup_{x,y\in\mathbb{L}} \mathbb{E}[w_{x,y}^{2p}] < \infty \text{ and } \sup_{x,y\in\mathbb{L}} \mathbb{E}[w_{x,y}^{-2q}\mathbb{1}_{\{w_{x,y}>0\}}] < \infty \tag{5.15}$$

for $p, q \in \mathbb{Z}_+$ with $p > \max\big\{(d+2)/d, (d+1)/(2(2-\alpha))\big\}$ and $q > (d+2)/d$. If moreover (5.3) holds true, then the quenched invariance principle holds for $X^\omega_\cdot$ with the limit process Y. Moreover, when $\alpha \in (0,1)$, the conclusion still holds true for $d > 2 - 2\alpha$, if $p > \max\big\{(d+1)/(2(1-\alpha)), (d+2)/d\big\}$ and $q > (d+2)/d$.

Proof. According to Theorem 5.3, it suffices to verify (5.4) — (5.10). We first verify (5.9). Recall that in the present setting $K_n \equiv 1$ for all $n \in \mathbb{N}$, and so $C_3 = C_4 = 1$. Suppose that $p_0 := \sup_{x,y\in\mathbb{L},x\neq y} \mathbb{P}\big(w_{x,y} = 0\big) < 2^{-4}$. Denote by $L(x,r) := |\{y \in \mathbb{L} : |y-x| \leqslant r\}|$. Then, for every $r > 0$ and $x, z \in \mathbb{L}$,

$$\begin{aligned}\mathbb{P}\Big(\sum_{y\in\mathbb{L}:|y-x|\leqslant r} \mathbb{1}_{\{w_{z,y}\neq 0\}} \leqslant \frac{3}{4}L(x,r)\Big) &\leqslant \sum_{k=0}^{[\frac{3}{4}L(x,r)]+1} \binom{L(x,r)}{k} p_0^{L(x,r)-k}\\ &\leqslant 2^{L(x,r)} p_0^{[\frac{1}{4}L(x,r)]-1} \left(\left[\frac{3}{4}L(x,r)\right]+1\right) \leqslant c_0 2^{-c_0' r^d},\end{aligned}$$

where in the second inequality we used the fact that $\binom{L(x,r)}{k} \leqslant 2^{L(x,r)}$ for all $0 \leqslant k \leqslant L(x,r)$, and the last inequality follows from $p_0 < 2^{-4}$ and $L(x,r) \asymp r^d$. The estimate above yields that $\sum_{R=1}^{\infty} \sum_{x,z\in B_F(0,6R)\cap V_1} \sum_{r=R^\theta/2}^{2R} p_6(x,z,r,3/4) < \infty$. This is, (5.9) holds with $c_0 = 3/4$.

Recall that, for any independent sequence $\{\xi_n : n \geqslant 1\}$ such that $\mathbb{E}[\xi_n] = 0$ for all $n \geqslant 1$ and $\sup_n \mathbb{E}[|\xi_n|^{2p}] < \infty$ for some $p \in \mathbb{Z}_+$, it holds for every $N \geqslant 1$ that $\mathbb{E}\Big[\big|\sum_{n=1}^{N} \xi_i\big|^{2p}\Big] \leqslant c_1(p)N^p$, where $c_1(p)$ is a constant independent of N. Then, for every $\varepsilon_0 > 0$, R, $r > 0$, $c_0^* \geqslant 2$, $n \geqslant 1$ and a subsequence of bounded measurable function h_n on $\mathbb{L}\times\mathbb{L}$ such that $\sup_{n\geqslant 1} \|h_n\|_\infty < \infty$,

$$p_1(r,R,\varepsilon_0) \leqslant \varepsilon_0^{-2p} R^{-2pd} r^{-2pd} \mathbb{E}\Big[\Big|\sum_{x,y\in\mathbb{L}:|x|\leqslant R,|y-x|\leqslant r} (w_{x,y} - \mathbb{E}[w_{x,y}])\Big|^{2p}\Big] \leqslant c_1(\varepsilon_0) r^{-pd} R^{-pd},$$

$$p_2(x,r,\varepsilon_0) \leqslant \varepsilon_0^{-2p} r^{-2pd} \mathbb{E}\Big[\Big|\sum_{y\in\mathbb{L}:|y-x|\leqslant r} \big(w_{x,y} - \mathbb{E}[w_{x,y}]\big)\Big|^{2p}\Big] \leqslant c_2(\varepsilon_0) r^{-pd},$$

$$p_4(x,r,c_0^*,\varepsilon_0) \leqslant \varepsilon_0^{-2q} r^{-2qd} \mathbb{E}\Big[\Big|\sum_{y\in\mathbb{L}:|y-x|\leqslant c_0^* r} \big(w_{x,y}^{-1} - \mathbb{E}[w_{x,y}^{-1}]\big)\Big|^{2q}\Big] \leqslant c_3(\varepsilon_0, c_0^*) r^{-qd},$$

$$\begin{aligned}p_5^{(n)}(x,N,\varepsilon,h_n,\varepsilon_0) &\leqslant c_4(\varepsilon_0,\varepsilon,\sup_{n\geqslant 1}\|h_n\|_\infty) n^{2\alpha p} \mathbb{E}\left[\left|\sum_{y\in\mathbb{L}:|y-x|\geqslant n\varepsilon,|y|\leqslant nN} \frac{w_{x,y} - \mathbb{E}[w_{x,y}]}{|x-y|^{d+\alpha}}\right|^{2p}\right]\\ &\leqslant c_5(\varepsilon_0,N,\varepsilon,\sup_{n\geqslant 1}\|h_n\|_\infty) n^{2\alpha p} n^{pd} n^{-2p(d+\alpha)} = c_5(\varepsilon_0,N,\varepsilon,\sup_{n\geqslant 1}\|h_n\|_\infty) n^{-pd}.\end{aligned}$$

In the following, we fix $x \in \mathbb{L}$. Let

$$\begin{aligned}S_p(i) &:= \mathbb{E}\left[\left|\sum_{y\in\mathbb{L}:|y-x|\leqslant 2^i} \frac{\big(w_{x,y} - J_1(x,y)\big)}{|x-y|^{d+\alpha-2}}\right|^{2p}\right]\\ &\leqslant c_6 \mathbb{E}\left[\left|\sum_{j=0}^{i} 2^{j(2-d-\alpha)} \sum_{y\in\mathbb{L}:2^{j-1}<|y-x|\leqslant 2^j} \big(w_{x,y} - J_1(x,y)\big)\right|^{2p}\right] =: c_6 \mathbb{E}\left[\left|\sum_{j=1}^{i} a(j)\xi(j)\right|^{2p}\right],\end{aligned}$$

where $a(j) = 2^{j(2-d-\alpha)}$ and $\xi(j) = \sum_{y\in\mathbb{L}:2^{j-1}<|y-x|\leqslant 2^j} \big(w_{x,y} - \mathbb{E}[w_{x,y}]\big)$. Noting that $\mathbb{E}[\xi(j)] = 0$ and $\mathbb{E}[|\xi(j)|^{2p}] \leqslant c_7 2^{jdp}$ for all $j \geqslant 1$, by the independent property of $\{w_{x,y}(\omega)\}$, $\sup_{i\geqslant 1} S_1(i) \leqslant c_6 \sup_{i\geqslant 1}\Big(\sum_{j=1}^{i} a(j)^2 \mathbb{E}[\xi(j)^2]\Big) \leqslant c_8 \sum_{j=1}^{\infty} 2^{j(4-d-2\alpha)} < \infty$, where the last step is due to the fact $d > 4-2\alpha$. Suppose that $\sup_{i\geqslant 1} S_k(i) < \infty$ for every $0 \leqslant k < p$. Then

$$\begin{aligned}S_{k+1}(i) - S_{k+1}(i-1) &= \sum_{l=1}^{k+1} \binom{k+1}{l} a(i)^{2l} \mathbb{E}[\xi(i)^{2l}] S_{k+1-l}(i-1)\\ &\leqslant c_9(k)\big(\sup_{0\leqslant j\leqslant k, i\geqslant 1} S_j(i)\big) 2^{i(4-d-2\alpha)},\end{aligned}$$

which implies $\sup_{i\geqslant 1} S_{k+1}(i) \leqslant c_{10}(k)\sum_{r=1}^{\infty} 2^{i(4-d-2\alpha)} < \infty$. So, by induction, we arrive at that $\sup_{i\geqslant 1} S_p(i) < \infty$. Hence, for every $x \in \mathbb{L}$, $p_3(x, R, \varepsilon_0) \leqslant c_9(\varepsilon_0)R^{-2(2-\alpha)p}$.

Under assumptions of the proposition, we can choose $\theta \in (0, 1)$ (close to 1) such that

$$p > \max\left\{\frac{d+1+\theta}{d\theta}, \frac{d+1}{2\theta(2-\alpha)}\right\} \quad \text{and} \quad q > \frac{d+1+\theta}{d\theta},$$

also thanks to the condition that $d > 4-2\alpha$ again. Then, according to all the estimates above, we know immediately that (5.4) — (5.8) hold for this $\theta \in (0, 1)$ and every sufficiently small $\varepsilon_0 > 0$.

Suppose that $\alpha \in (0, 1)$. If $d > 2 - 2\alpha$, $p > \max\{(d + 1)/(2(1 - \alpha)), (d + 2)/d\}$ and $q > (d + 2)/d$, then we can choose $\theta \in (0, 1)$ (close to 1) such that

$$p > \max\left\{\frac{d+1+\theta}{d\theta}, \frac{d+1}{2\theta(1-\alpha)}\right\} \quad \text{and} \quad q > \frac{d+1+\theta}{d\theta}.$$

Following the argument above, we can prove that (5.4) — (5.6), (5.8) and (5.10) are satisfied. Then, the desired assertion follows from Theorem 5.3 again. The proof is complete. □

Theorem 1.1 is a direct consequence of Proposition 5.6, since (5.3) holds trivially in this setting.

5.2.2. *Time-change of α-stable-like process on $\mathbb{R}^d$.* Let us first fix the triple (F, ρ, m) with $F = \mathbb{R}^d$, ρ being the Euclidean distance and $m(dx) = K(x)\,dx$, where dx denotes the Lebesgue measure on $\mathbb{R}^d$ and K is a continuous function on $\mathbb{R}^d$ satisfying that $0 < C_1 \leqslant K(x) \leqslant C_2 < \infty$ for some constants $C_1 \leqslant C_2$. Then, the process Y associated with the Dirichlet form $(D_0, \mathscr{F}_0)$ given at the beginning of Subsection 5.1 is a time-change of symmetric α-stable process on $\mathbb{R}^d$ with $c(x, y) = K(x)^{-1}K(y)^{-1}$ for $x, y \in \mathbb{R}^d$. It is obvious that $(D_0, \mathscr{F}_0)$ satisfies assumption **(Dir.)**.

Similar to the previous part, we can take $V_1 = \mathbb{Z}^d$, and $m_n = K_n\mu_n$ with μ_n being the counting measure on $n^{-1}\mathbb{Z}^d$ and

$$K_n(x) = n^{-d}\int_{U_n(x)} K(x)\,dx, \quad x \in n^{-1}\mathbb{Z}^d,$$

where $U_n(x) = \Pi_{i=1}^d[x_i, x_i + n^{-1})$ for any $x = (x_1, \cdots, x_d) \in n^{-1}\mathbb{Z}^d$. Let $(X_t^\omega)_{t\geqslant 0}$ be the symmetric Hunt process associated with Dirichlet form $(D^\omega_{V_1}, \mathscr{F}^\omega_1)$ with $V_1 = \mathbb{Z}^d$ and $w^{(1)}_{x,y}(\omega) = w_{x,y}(\omega)$. Note that for any compact set $S \subset \mathbb{R}^d$, $\lim_{n\to\infty}\sup_{x\in S}\left|K_n(n[x]_n) - K(x)\right| = 0$. If $J_1(x, y) = \mathbb{E}[w_{x,y}] = K_1(x)^{-1}K_1(y)^{-1}$ for all $x, y \in \mathbb{Z}^d$, then (5.3) holds true. Hence, following the same arguments in the proof of Proposition 5.6, we can obtain that under assumption (5.15) the quenched invariance principle holds for $(X_t^\omega)_{t\geqslant 0}$ with limiting process Y being a time-change of symmetric α-stable process on $\mathbb{R}^d$.

Remark 5.7. From the example above, we know that to identity the limit process consists of two ingredients. One is to verify locally weak convergence of m_n to m, and the other is to justify convergence of the jumping kernel for the associated Dirichlet form. In fact, by carefully tracking the proof above, we can see that if the measure m_n is replaced by a more general (random) measure which converges locally weakly to m, then the quenched invariance principle still holds with the same limiting process.

5.2.3. *Bounded Lipschitz domain.* In fact, Proposition 5.6 holds not only for a half/quarter space, but also for the closure of a bounded Lipschitz domain in $\mathbb{R}^d$, whose intrinsic distance is equivalent to the Euclidean distance and whose volume growth is with order d. In details, let $F \subset \mathbb{R}^d$ be a closed set such that for any $x, y \in F$ and $r > 0$, $c_1r^d \leqslant m(B_F(x, r)) \leqslant c_2r^d$ and $c_1|x - y| \leqslant \rho_F(x, y) \leqslant c_2|x - y|$, where

$$\rho_F(x, y) := \inf\left\{\int_0^1 |\dot\gamma(s)|\,ds : \gamma \in C^1([0, 1]; F), \gamma(0) = x, \gamma(1) = y\right\}$$

is the intrinsic metric on F, m is the Lebesgue measure, and $B_F(x,r)$ is the ball with respect to ρ_F. For $x=(x_1,\cdots,x_d)\in n^{-1}\mathbb{Z}^d$, set $U_n(x)=\Pi_{i=1}^d[x_i,x_i+n^{-1})$. Note that when F is the closure of a bounded Lipschitz domain, $V_n:=\{n^{-1}\mathbb{Z}^d\cap F: U_n(x)\subset F\}$ satisfies the properties given in Lemma 4.1. Suppose that $\{w_{x,,y}: x,y\in\mathbb{Z}^d\}$ is a sequence of independent random variables satisfying the conditions in Proposition 5.6. Then the conclusion of Proposition 5.6 holds on F. Indeed, in this case, by taking V_n as above, the proofs of Theorem 5.3 and Proposition 5.6 go through without any change (with ρ replaced by ρ_F as explained in Remark 5.1). Note that neither $V_n=n^{-1}V_1$ nor $X_t^{(n),\omega}=n^{-1}X_{n^\alpha t}^{V_1,\omega}$ holds in general in this setting. (However, we can verify that $X_t^{(n),\omega}=n^{-1}X_{n^\alpha t}^{\tilde V_n,\omega}$, where $\tilde V_n:=nV_n\subset nF$.) Note that the proofs do not require these properties, and the integrability condition given for all $x,y\in\mathbb{Z}^d$ is (more than) enough for the estimates in the proofs to hold.

5.2.4. *Fractal graph.* The arguments in Example 5.2.1 work for more general graphs that satisfy (i)–(iv), and that its scaling limit (F,ρ,m) and Dirichlet form which satisfy **(MMS)** and **(Dir.)** respectively as discussed at the beginning of subsection 5.1. In particular, we can prove quenched invariance principle for stable-like processes on various fractal graphs.

Here we introduce the most typical fractal graph; namely the Sierpinski gasket graph. Let $e_0=(0,0,\cdots,0)\in\mathbb{R}^N$, and for $1\leqslant i\leqslant N$, e_i be the unit vector in $\mathbb{R}^N$ whose i-th element is 1. Set $F_i(x)=(x-e_i)/2+e_i$ for $0\leqslant i\leqslant N$. Then, there exists the unique non-void compact set such that $K=\cup_{i=0}^N F_i(K)$; K is called the N-dimensional Sierpinski gasket. Set $F:=\cup_{n=0}^\infty 2^nK$, which is the unbounded Sierpinski gasket. Let

$$V_1=\bigcup_{m=0}^\infty 2^m\Big(\bigcup_{i_1,\cdots,i_m=0}^N F_{i_1}\circ\cdots\circ F_{i_m}(\{e_0,\cdots,e_N\})\Big),\ V_n=2^{-n+1}V_1.$$

(Hence, n^{-1} in the definition of V_n in the previous subsection is now 2^{-n+1}.) The closure of $\cup_{m\geqslant 1}V_m$ is F. F satisfies assumption **(MMS)** with $d=\log(N+1)/\log 2$. We can naturally construct a regular stable-like Dirichlet form satisfying assumption **(Dir.)**. Let $\{w_{x,y}: x,y\in V_1\}$ be a sequence of independent random variables. Then we have Proposition 5.6 with the same proof in this case as well.

Acknowledgements. The research of Xin Chen is supported by National Natural Science Foundation of China (No. 11501361). The research of Takashi Kumagai is supported by the Grant-in-Aid for Scientific Research (A) 25247007 and 17H01093, Japan. The research of Jian Wang is supported by National Natural Science Foundation of China (No. 11522106), Fok Ying Tung Education Foundation (No. 151002) and the Program for Probability and Statistics: Theory and Application (No. IRTL1704).

References

[1] Aldous, D.: Stopping times and tightness, *Ann. Probab.* **6** (1978) 335–340.
[2] Andres, S., Barlow, M.T., Deuschel, J.-D. and Hambly, B.M.: Invariance principle for the random conductance model, *Probab. Theory Relat. Fields* **156** (2013) 535–580.
[3] Andres, S., Deuschel, J.-D. and Slowik, M.: Invariance principle for the random conductance model in a degenerate ergodic environment, *Ann. Probab.* **43** (2015) 1866–1891.
[4] Andres, S., Deuschel, J.-D. and Slowik, M.: Harnack inequalities on weighted graphs and some applications to the random conductance model, *Probab. Theory Relat. Fields* **166** (2017) 713–752.
[5] Andres, S., Chiarini, A., Deuschel, J.-D. and Slowik, M.: Quenched invariance principle for random walks with time-dependent ergodic degenerate weights, *Ann. Probab.* **46** (2018) 302–336.
[6] Barlow, M.T.: Random walks on supercritical percolation clusters, *Ann. Probab.* **32** (2004) 3024–3084.
[7] Barlow, M.T., Bass, R. F. and Kumagai, T.: Parabolic Harnack inequality and heat kernel estimates for random walks with long range jumps, *Math. Z.* **261** (2009) 297–320.
[8] Barlow, M.T. and Chen, X.X.: Gaussian bounds and parabolic Harnack inequality on locally irregular graphs, *Math. Ann.* **366** (2016) 1677–1720.
[9] Barlow, M.T. and Deuschel, J.-D.: Invariance principle for the random conductance model with unbounded conductances, *Ann. Probab.* **38** (2010) 234–276.

[10] Barlow, M.T., Grigor'yan, A. and Kumagai, T.: Heat kernel upper bounds for jump processes and the first exit time, *J. Reine Angew. Math.* **626** (2009) 135–157.
[11] Barlow, M.T. and Hambly. B.M.: Parabolic Harnack inequality and local limit theorem for random walks on percolation clusters, *Electron. J. Probab.* **14** (2009) 1–27.
[12] Bass, R.F.: On Aronsen's upper bounds for heat kernels, *Bull. London Math. Soc.* **34** (2002) 415–419.
[13] Berger, N.: Transience, recurrence and critical behavior for long-range percolation, *Commun. Math. Phys.* **226** (2002) 531–558.
[14] Berger, N. and Biskup, M.: Quenched invariance principle for simple random walk on percolation clusters, *Probab. Theory Relat. Fields* **137** (2007) 83–120.
[15] Biskup, M.: On the scaling of the chemical distance in long range percolation models, *Ann. Probab.* **32** (2004) 2938–2977.
[16] Biskup, M.: Recent progress on the random conductance model, *Probab. Surv.* **8** (2011) 294–373.
[17] Biskup, M and Prescott, T.-M.: Functional CLT for random walk among bounded random conductances, *Electron. J. Probab.* **12** (2007) no. 49, 1323–1348.
[18] Biskup, M. and Rodriguez, P.-F.: Limit theory for random walks in degenerate time-dependent random environments, *J. Funct. Anal.* **274** (2018) 985–1046.
[19] Chen, X., Kumagai, T. and Wang, J.: Random conductance models with stable-like jumps II: heat kernel estimates and Harnack inequalities, Preprint 2018.
[20] Chen, Z.-Q., Croydon, D.A. and Kumagai, T.: Quenched invariance principles for random walks and elliptic diffusions in random media with boundary, *Ann. Probab.* **43** (2015) 1594–1642.
[21] Chen, Z.-Q., Kim, P. and Kumagai, T.: Weighted Poincaré inequality and heat kernel estimates for finite range jump processes, *Math. Ann.* **342** (2008) 833–883.
[22] Chen, Z.-Q., Kim, P. and Kumagai, T.: Discrete approximation of symmetric jump processes on metric measure spaces, *Probab. Theory Relat. Fields* **155** (2013) 703–749.
[23] Chen, Z.-Q. and Kumagai, T.: Heat kernel estimates for stable-like processes on d-sets, *Stoch. Proc. Appl.* **108** (2003) 27–62.
[24] Chen, Z.-Q. and Kumagai, T.: Heat kernel estimates for jump processes of mixed types on metric measure spaces, *Probab. Theory Relat. Fields* **140** (2008) 277–317.
[25] Chen, Z.-Q., Kumagai, T. and Wang, J.: Stability of heat kernel estimates for symmetric non-local Dirichlet forms, to appear in *Memoirs of the AMS*, available at `arXiv:1604.04035`.
[26] Crawford, N. and Sly, A.: Simple random walks on long range percolation cluster I: heat kernel bounds, *Probab. Theory Relat. Fields* **154** (2012) 753–786.
[27] Crawford, N. and Sly, A.: Simple random walks on long range percolation cluster II: scaling limits, *Ann. Probab.* **41** (2013) 445–502.
[28] Davies, E.B.: Large deviations for the heat kernels on graphs, *J. London Math. Soc.* **23** (1993) 65–72.
[29] Deuschel, J., Nguyen, T. and Slowik, M.: Quenched invariance principles for the random conductance model on a random graph with degenerate ergodic weights, *Probab. Theory Relat. Fields* **170** (2018) 363–386.
[30] Ethier, S.N. and Kurtz, T.G.: *Markov Processes: Characterization and Convergence*, John Wiley & Sons, New York, 1986.
[31] Husseini, R. and Kassmann, M.: Markov chain approximations for symmetric jump processes, *Potential Anal.* **27** (2007) 353–380.
[32] Kumagai, T.: Random walks on disordered media and their scaling limits, *Lect. Notes in Math.* **2101**, Ecole d'été de probabilités de Saint-Flour XL–2010, Springer, New York 2014.
[33] Mathieu, P.: Quenched invariance principles for random walks with random conductances, *J. Stat. Phys.* **130** (2008) 1025–1046.
[34] Mathieu, P. and Piatnitski, A.: Quenched invariance principles for random walks on percolation clusters, *Proc. Roy. Soc. A* **463** (2007) 2287–2307.
[35] Nash, J.: Continuity of solutions of parabolic and elliptic equations, *Amer. J. Math.* **80** (1958) 931–954.
[36] Procaccia, E., Rosenthal, R. and Sapozhnikov, A.: Quenched invariance principle for simple random walk on clusters in correlated percolation models, *Probab. Theory Relat. Fields* **166** (2016) 619–657.
[37] Sidoravicius, V. and Sznitman, A.-S.: Quenched invariance principles for walks on clusters of percolation or among random conductances, *Probab. Theory Relat. Fields* **129** (2004) 219–244.