Discrete Approximation of Symmetric Jump Processes on Metric Measure Spaces

Zhen-Qing Chen[*] **Panki Kim**[†] and **Takashi Kumagai**[‡]

(December 25, 2011)

Abstract

In this paper we give general criteria on tightness and weak convergence of discrete Markov chains to symmetric jump processes on metric measure spaces under mild conditions. As an application, we investigate discrete approximation for a large class of symmetric jump processes. We also discuss some application of our results to the scaling limit of random walk in random conductance.

AMS 2010 Mathematics Subject Classification: Primary 60B10, 60J25; Secondary 60J35, 60G52, 60J75.

Keywords and phrases: weak convergence, Mosco convergence, tightness, Skorohod space, Dirichlet form, random conductance, jump process, symmetric jump process

1 Introduction

This paper is concerned with the following two questions.

(Q1) Given a symmetric Hunt process X on $\mathbb{R}^d$, can it be approximated by a sequence of symmetric Markov chains $X^{(k)}$ on $k^{-1}\mathbb{Z}^d$?

(Q2) For a sequence of $\{X^{(k)}; k \geq 1\}$ of symmetric Markov chains on $k^{-1}\mathbb{Z}^d$, when does $X^{(k)}$ converge weakly to a 'nice' Hunt process X on $\mathbb{R}^d$ as $k \to \infty$?

In this paper, we address these two questions for symmetric processes X of pure jump type on a general metric space E.

[*]Research partially supported by NSF Grants DMS-0906743 and DMR-1035196.

[†]Research supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education, Science and Technology(0409-20110087)

[‡]Research partially supported by the Grant-in-Aid for Scientific Research (B) 22340017.

Let us briefly mention some work on these problems when X is a diffusion. When X is a diffusion corresponding to an operator in non-divergence form, these problems were studied, for example, in the book of Stroock-Varadhan ([28, Chapter 11]) by solving the corresponding martingale problem. When X is a symmetric diffusion corresponding to a uniformly elliptic divergence form operator, **(Q1)** is solved completely by Stroock-Zheng [29]. Let $X_t^{(k)}$ be a continuous time symmetric Markov chain on $k^{-1}\mathbb{Z}^d$ with conductances $\mathcal{C}^{(k)}(x,y)$; This means that $X^{(k)}$ stays at a state x for an exponential length of time with parameter $\mathcal{C}^{(k)}(x) := \sum_{z\neq x} \mathcal{C}^{(k)}(x,z)$ and then jumps to the next state y with probability $\mathcal{C}^{(k)}(x,y)/\mathcal{C}^{(k)}(x)$. In [29], they also answered **(Q2)** when $\mathcal{C}^{(k)}(\cdot,\cdot)$ is of finite range (i.e. $\mathcal{C}^{(k)}(x,y) = 0$ if $|x-y| \geq R_0/k$ for some $R_0 > 0$) and has certain uniform regularity. The core of their paper is to establish a discrete version of the De Giorgi-Moser-Nash theory. Recently, in [3], the main results in [29] are extended in two ways: Markov chains with unbounded range were allowed and the strong uniform regularity conditions on conductances in [29] are weakened. This was further extended in [4] so that the limiting process X had a continuous part and a jump part. For both [3, 4], a crucial step is to obtain a priori estimate of the solution of the heat equation, which can be derived thanks to the recent developments of the De Giorgi-Moser-Nash theory for jump processes. When X is reflected Brownian motion on a domain, **(Q1)** was solved in [5] by a completely different method, without using a priori estimates on the transition function of the Markov processes. The methodology of [5] is a Dirichlet form based approach.

Now consider the case where X is a symmetric Hunt process of pure jump. Let $(\mathcal{E},\mathcal{F})$ be its associated symmetric Dirichlet form on $L^2(\mathbb{R}^d;m)$, where m is a Radon measure on $\mathbb{R}^d$ and

$$\begin{aligned} \mathcal{F} &:= \left\{ u \in L^2(\mathbb{R}^d;m) : \int_{\mathbb{R}^d\times\mathbb{R}^d\setminus\widehat{d}} (u(x)-u(y))^2 J(dx,dy) < \infty \right\}, \qquad (1.1) \\ \mathcal{E}(u,v) &:= \frac{1}{2}\int_{\mathbb{R}^d\times\mathbb{R}^d\setminus\widehat{d}} (u(x)-u(y))(v(x)-v(y))J(dx,dy) \qquad \text{for } u,v\in\mathcal{F}. \end{aligned}$$

Here $\widehat{d}$ is the diagonal set in $\mathbb{R}^d\times\mathbb{R}^d$, $J(\cdot,\cdot)$ is a measure on $\mathbb{R}^d\times\mathbb{R}^d$ such that $J(A,B)=J(B,A)$. The paper [15] considered **(Q1)**–**(Q2)** when $J(dx,dy) = j(x,y)dxdy$, $j(x,y) \asymp |x-y|^{-d-\alpha}$ for some $0<\alpha<2$ and $m(dx)=dx$. (Here and in the following, $f \asymp g$ means that there are $c_1,c_2>0$ so that $c_1 g(x) \leq f(x) \leq c_2 g(x)$ in the common domain of definition for f and g.) This is extended in [2] to more general Dirichlet form $(\mathcal{E},\mathcal{F})$. Again, for both [2, 15], the crucial point is to obtain a priori Hölder estimate of the solution of the heat equation. However for general symmetric Markov processes, obtaining good a priori estimate for their transition densities is impossible. Indeed, even in the case $c_1|x-y|^{-d-\alpha_1} \leq j(x,y) \leq c_2|x-y|^{-d-\alpha_2}$ for $|x-y|<1$ where $\alpha_1<\alpha_2$, one can construct an example where there is a bounded harmonic function that is not continuous (see [1, Theorem 1.9]).

In this paper, we will answer **(Q1)** affirmatively for a very general class of symmetric Markov processes whose associated Dirichlet forms are of the form (1.1) (see Theorem 6.1), and give answer to **(Q2)** when $X^{(k)}$ and X satisfy either conditions **(A1)**–**(A4)** or conditions **(A1)**–**(A2)** and **(A3)***–**(A4)*** in Section 2 (see Theorem 2.2). Our approach does not rely on the a priori estimate of the heat kernel, instead we adapt the ideas of [5] and use the Lyons-Zheng decomposition to obtain tightness (Proposition 3.4). The drawback is we can only obtain tightness when the initial distribution is absolutely continuous with respect to the reference measure. Note that when we have a priori estimate of the heat kernel (such as examples discussed in [2, 15]), we can obtain tightness for any initial distributions. To show finite dimensional distribution convergence, we establish the Mosco convergence, which is equivalent to strong convergence of the semigroups (Theorems 4.5 and 4.7). We will obtain these results on a large class of metric measure spaces with volume doubling property.

It is important and useful if we can obtain **(Q2)** in such a way that is applicable to prove convergence of Markov chains on some random media. In order to establish such results, we need to relax the assumption for $X^{(k)}$. In Theorem 4.7, we prove the Mosco convergence under a milder condition on $X^{(k)}$ and a stronger condition on X. Then the following example can be handled. Let $\{\xi_{xy}\}_{x,y\in\mathbb{Z}^d, x\neq y}$ be a sequence of i.i.d. non-negative real-valued random variables on a probability space $(\Omega, \mathcal{A}, \mathbf{P})$ with $\mathbf{E}[\xi_{x,y}] = 1$ and Var $(\xi_{x,y}) < \infty$. Let $d \geq 2$, $0 < \alpha < 2$ and

$$\mathcal{C}(x,y) = \xi_{xy}|x-y|^{-d-\alpha}, \qquad x, y \in \mathbb{Z}^d$$

be the random conductance. Let $X^{(1)}$ be the corresponding Markov chain on $\mathbb{Z}^d$ with this conductance. Then we can prove that $X^{(k)}_t = k^{-1}X^{(1)}_{k^\alpha t}$ converges weakly to (a constant time change of) symmetric α-stable process on $\mathbb{R}^d$ equipped with convergence-in-measure topology $\mathbf{P}$-a.s. (see Proposition 7.1(i), and see the paragraph after Theorem 2.2 for the definition of convergence-in-measure topology). Moreover, if we further assume that $0 \leq \xi_{xy} \leq C_1$ for some deterministic constant $C_1 > 0$, we can prove that $X^{(k)}_t$ converges weakly to symmetric α-stable process on $\mathbb{R}^d$ equipped with the Skorohod topology $\mathbf{P}$-a.s. (see Proposition 7.1(ii)).

The rest of the paper is organized as follows. In Section 2, we present the framework of the base metric measure space E and present a graph approximation result. We then give the precise statements of two main weak convergence results of this paper, and the conditions under which these results hold. The proof of these two theorems will be given in Section 5. It is standard that weak convergence of stochastic processes is established through two steps: tightness and convergence of finite dimensional distributions. How to carry out these steps varies from problems to problems and they can be very challenging tasks. In Section 3, we establish tightness results for a family of Markov chains $X^{(k)}$ on the approximating graphs in the space $\mathbb{D}[0,1]$ of right continuous functions having left limits equipped either with the Skorohod topology or with the convergence-in-measure topology. The latter topology is also called pseudo-path topology in literature and is weaker than the Skorohod topology. In Section 4, we give sufficient conditions for finite dimensional distribution convergence of $X^{(k)}$ to X, through Mosco convergence method. Note the state spaces of $\{X^{(k)}, k \geq 1\}$ are changing. So we need an extension of the Mosco convergence introduced in [26]. A full proof of the generalized Mosco convergence can be found in Appendix at the end of this paper. In Section 6, we investigate the discrete approximation of X. Applications of our main results to random walk in random conductance models are given in Section 7.

For technical convenience, we will often consider stochastic processes whose initial distribution is a finite measure, not necessarily normalized to have total mass 1, for example, $\varphi(x)m(dx)$ where φ is bounded function with compact support. Translating our results to the usual probabilistic setting is straightforward and is left to the reader.

Throughout paper, we use ":=" to denote a definition, which is read as "is defined to be". The letter c, with or without subscripts, signifies a constant whose value is unimportant and which may change from location to location, even within a line. We will use ∂ to denote the cemetery point and for every function f, we extend its definition to ∂ by setting $f(\partial) = 0$. For a locally compact metric space E, we use $E_\partial := E \cup \{\partial\}$ to denote the one-point compactification of E. For a metric space E, we use $C(E)$ to denote the space of continuous functions on E and $\mathrm{Lip}(E)$ the space of Lipschitz continuous functions on E. For any collection of numerical functions $\mathcal{H}$, $\mathcal{H}^+$ denotes the set of nonnegative functions in $\mathcal{H}$, $\mathcal{H}_b$ denotes the set of bounded functions in $\mathcal{H}$ and $\mathcal{H}_c$ denotes the set of functions in $\mathcal{H}$ with compact support. Moreover, we denote $\mathcal{H}_c^+ := \mathcal{H}^+ \cap \mathcal{H}_c$ and $\mathcal{H}_b^+ := \mathcal{H}^+ \cap \mathcal{H}_b$. For any topological space W and any subset $I \subset [0, \infty)$, we denote

$$\mathbb{D}_W I := \big\{ f : I \to W \,\big|\, f \text{ is right continuous having left limits.} \big\}. \tag{1.2}$$

We will use $\#S$ is the cardinality of a set S.

2 Statement of main results

2.1 Discrete approximation of state spaces

Let (E, ρ, m) be a metric measure space, where (E, ρ) is a locally compact separable connected metric space and m is a Radon measure on E with $V(x, r) := m(B(x, r)) \in (0, \infty)$ and $m(\partial B(x, r)) = 0$ for each $r > 0$ and $x \in E$. Here and in the sequel, $B(x, r)$ denotes the open ball of radius r centered at x, and $\partial B(x, r) = \overline{B(x, r)} \setminus B(x, r)$. The metric measure space (E, ρ, m) will serve as the state space of our jump processes X. We assume the following:

(MMS.1) The closure of $B(x, r)$ is compact for every $x \in E$ and $r > 0$.

(MMS.2) ρ is geodesic, that is, for any two points $x, y \in E$, there exists a continuous map $\gamma : [0, \rho(x, y)] \to E$ such that $\gamma(0) = x$, $\gamma(\rho(x, y)) = y$ and $\rho(\gamma(s), \gamma(t)) = t - s$ for all $0 \le s \le t \le \rho(x, y)$.

(MMS.3) (E, ρ, m) satisfies volume doubling property (VD for short), that is,

$$\text{there is a constant } C_* > 0 \text{ such that } V(x, 2r) \le C_* V(x, r) \quad \text{for every } x \in E \text{ and } r > 0.$$

Fix some $x_0 \in E$. Condition **(MMS.3)** in particular implies that

$$V(x_0, 2^n) \le C_*^n V(x_0, 1) = (2^n)^{\log_2 C_*} V(x_0, 1) \qquad \text{for every } n \ge 1.$$

So there are constants $c_0 = c_0(x_0) > 0$ and $d_0 > 0$ such that

$$V(x_0, r) \le c_0 r^{d_0} \qquad \text{for every } r \ge 1. \tag{2.1}$$

It follows then

$$\begin{aligned}\int_E e^{-\lambda\rho(x,x_0)} m(dx) &= \int_0^\infty e^{-\lambda r} d(V(B(x_0,r)) = \lambda \int_0^\infty V(B(x_0,r))\, e^{-\lambda r}\, dr \\ &\le c\,\lambda \left(1+\int_1^\infty r^{d_0}\, e^{-\lambda r} dr\right) < \infty. \end{aligned} \tag{2.2}$$

Property (2.2) will imply that the jump process X under consideration in this paper is conservative under the assumption (5.1) (see one line after (5.1)).

To study discrete approximation of X, we first need to have a discrete approximation of the state space E.

Consider approximating graphs $\{(V_k, \Xi_k),\ k \in \mathbb{N}\}$ of E with the graph distance ρ_k and the associated partition $\{U_k(x), x \in V_k;\ k \in \mathbb{N}\}$ that satisfies the following properties. Here V_k is the set of vertices and Ξ_k is the set of edges of the graph (V_k, Ξ_k).

(AG.1) (V_k, Ξ_k) is connected and has uniformly bounded degree.

(AG.2) $V_k \subset E$, $\cup_{k=1}^\infty V_k$ is dense in E and

$$\frac{C_1}{k}\rho_k(x,y) \le \rho(x,y) \le \frac{C_2}{k}\rho_k(x,y) \qquad \text{for every } x, y \in V_k. \tag{2.3}$$

(AG.3) For each $k \ge 1$, $\cup_{x\in V_k} U_k(x) = E$, $m(U_k(x) \cap U_k(y)) = 0$ for $x \ne y$, and

$$\sup\{\rho(\xi,\eta) : \xi, \eta \in U_k(x)\} \le C_3/k. \tag{2.4}$$

Moreover, for each $x \in V_k$, $V_k \cap \operatorname{Int} U_k(x) = \{x\}$, and we have

$$C_4\, m(U_k(x)) \le V(x, 1/k) \le C_5\, m(U_k(x)). \tag{2.5}$$

Here $\mathrm{Int}U_k(x)$ denotes the set of the interior points of $U_k(x)$.

Theorem 2.1 *Suppose that (E, ρ, m) is a metric measure space satisfying conditions* **(MMS.1)**–**(MMS.3)**. *Then E admits approximating graphs $\{(V_k, \Xi_k), k \ge 1\}$ and associated partitions $\{U_k(x), x \in V_k; k \ge 1\}$ that satisfy the properties* **(AG.1)**–**(AG.3)**.

The proof of Theorems 2.1 will be given in Section 3.

2.2 Random walk on graphs and its weak limit

For the remainder of this paper, we assume that (E, ρ, m) is a metric measure space satisfying conditions **(MMS.1)**–**(MMS.3)** and that $\{(V_k, \Xi_k), k \ge 1\}$, with the graph distance ρ_k, are approximating graphs with associated partitions $\{U_k(x), x \in V_k; k \ge 1\}$ satisfying **(AG.1)**–**(AG.3)**.

Let m_k be the measure defined on V_k by

$$m_k(A) = \sum_{y\in A} m(U_k(y)) \qquad \text{for } A \subset V_k. \tag{2.6}$$

For $y \in V_k$, $m_k(\{y\})$ will simply be denoted by $m_k(y)$.

For $k \in \mathbb{N}$, let $\{j^{(k)}(x,y),\ x,y \in V_k\}$ be a family of non-negative functions defined on the graph (V_k, Ξ_k) such that $j^{(k)}(x,x)=0$, $j^{(k)}(x,y)=j^{(k)}(y,x)$ for $x,y \in V_k$ and

$$\sum_{y\in V_k} j^{(k)}(x,y)m_k(y) < \infty \qquad \text{for every } x \in V_k. \tag{2.7}$$

Then $\{\mathcal{C}^{(k)}(x,y) := m_k(x)j^{(k)}(x,y)m_k(y),\ x,y \in V_k\}$ form a family of conductance defined on the graph (V_k, Ξ_k). Note that in contrast with notations in some literatures on graphs, here the set Ξ_k of edges only gives the topological structure of the graph and has nothing to do with the conductances; that is, Ξ_k can be different from the bond set $\{(x,y) : \mathcal{C}^{(k)}(x,y) > 0\}$. Note also that the graph with vertices V_k and bonds $\{(x,y) : \mathcal{C}^{(k)}(x,y) > 0\}$ could be disconnected. We consider the following quadratic form $(\mathcal{E}^{(k)}, \mathcal{F}^{(k)})$:

$$\begin{aligned}
\mathcal{F}^{(k)} &:= \Big\{u \in L^2(V_k; m_k)\,;\ \sum_{x,y\in V_k} (u(x)-u(y))^2 j^{(k)}(x,y)m_k(x)m_k(y) < \infty\Big\} \\
\mathcal{E}^{(k)}(u,v) &:= \frac{1}{2}\sum_{x,y\in V_k} (u(x)-u(y))(v(x)-v(y))j^{(k)}(x,y)m_k(x)m_k(y) \quad \text{for } u,v \in \mathcal{F}^{(k)}.
\end{aligned} \tag{2.8}$$

It is easy to check that $(\mathcal{F}^{(k)}, \mathcal{E}^{(k)})$ is a regular Dirichlet form on $L^2(V_k; m_k)$ (see Theorem 3.2). Let $X^{(k)} = (X^{(k)}_t, \mathbb{P}^{(k)}_x, x \in V_k)$ be the continuous time strong Markov process on V_k associated with the Dirichlet form $(\mathcal{F}^{(k)}, \mathcal{E}^{(k)})$. The process $X^{(k)}$ is sometimes called the continuous time random walk on V_k with conductance $\mathcal{C}^{(k)}$. We are interested in when and to which process $X^{(k)}$ converges weakly.

For notational convenience, let us fix some $x_0 \in E$ and, for $r > 0$, denote $B(x_0, r)$ by B_r. Note that by assumption **(MMS.1)**, $\overline{B_r}$ is compact for every $r > 0$.

Consider the following conditions:

(A1). There is $k_0 \geq 1$ so that for every integer $j \geq 1$,

$$\sup_{k\geq k_0}\ \sup_{x\in \overline{B}_j\cap V_k} \sum_{y\in V_k} j^{(k)}(x,y)\Big(\frac{\rho_k(x,y)}{k} \wedge 1\Big)^2 m_k(y) < \infty \tag{2.9}$$

and

$$\sup_{k\geq k_0}\ \sup_{x\in (B_{j+2})^c\cap V_k} \sum_{y\in B_j\cap V_k} j^{(k)}(x,y)m_k(y) < \infty. \tag{2.10}$$

(A2). *For m-a.e. $x \in E$, $j(x,\cdot)$ is a positive measure on $E\setminus\{x\}$ such that the following holds:*

(i) *For any $\varepsilon > 0$, $x \mapsto j(x, E\setminus B(x,\varepsilon))$ is locally integrable with respect to m.*

(ii) *For any non-negative Borel measurable functions u, v,*

$$\int_E u(x)(jv)(x)m(dx) = \int_E (ju)(x)v(x)m(dx)\ (\leq \infty).$$

Here $ju(x) := \int_{E\setminus\{x\}} u(y)j(x,dy)$.

(iii) *For any compact set K,*

$$\sup_{x\in K}\int_E (\rho(x,y)\wedge 1)^2 j(x,dy) < \infty. \tag{2.11}$$

Denote by $\widehat{d}$ the diagonal set in $E\times E$. The kernel j then determines a positive symmetric Radon measure $J(dx,dy)$ on $E\times E\setminus\widehat{d}$ by

$$\int_{E\times E\setminus\widehat{d}} f(x,y)J(dx,dy) = \int_E\left(\int_E f(x,y)j(x,dy)\right)m(dx) \quad \text{for } f\in C_c(E\times E\setminus\widehat{d}).$$

Define a bilinear form $(\mathcal{E},\mathcal{F})$ on $L^2(E;m)$ as follows:

$$\begin{aligned}\mathcal{F} &:= \left\{u\in L^2(E;m): \int_{E\times E\setminus\widehat{d}}(u(x)-u(y))^2 J(dx,dy)<\infty\right\}, \quad (2.12)\\ \mathcal{E}(u,v) &:= \frac{1}{2}\int_{E\times E\setminus\widehat{d}}(u(x)-u(y))(v(x)-v(y))J(dx,dy) \qquad \text{for } u,v\in\mathcal{F}.\end{aligned}$$

Under condition **(A2)**, it can be shown (see Lemma 4.2) that $\mathrm{Lip}_c(E)\subset\mathcal{F}$. We now introduce condition **(A3)**.

(A3). $\mathrm{Lip}_c(E)$ *is dense in* $(\mathcal{F},\mathcal{E}(\cdot,\cdot)+\|\cdot\|_2^2)$.

Under conditions **(A2)** and **(A3)**, by [9, Propostion 2.2] and its proof, $(\mathcal{E},\mathcal{F})$ is a regular Dirichlet form on $L^2(E;m)$. Denote by $X=\{X_t,t\geq 0,\mathbb{P}_x,x\in E\}$ the symmetric Hunt process on E associated with $(\mathcal{E},\mathcal{F})$.

To state condition **(A4)**, we need the following. First we define the restriction operator $\pi_k: L^2(E;m)\to L^2(V_k;m_k)$ and the extension operator $E_k: L^2(V_k;m_k)\to L^2(E;m)$ as follows:

$$\pi_k f(x) = \frac{1}{m_k(x)}\int_{U_k(x)} f(y)m(dy) \qquad \text{for } f\in L^2(E;m) \text{ and } x\in V_k, \tag{2.13}$$

$$E_k g(z) = g(x) \qquad \text{for } g\in L^2(V_k;m_k) \text{ and } z\in \mathrm{Int}\,U_k(x) \text{ with } x\in V_k. \tag{2.14}$$

For each $k\geq 1$, whenever needed, we extend the definition of $j^{(k)}(x,y)$ on $V_k\times V_k$ to $E\times E$ by taking

$$j^{(k)}(z,w) = \begin{cases} j^{(k)}(x,y) & \text{when } z\in\mathrm{Int}\,U_k(x) \text{ and } w\in\mathrm{Int}\,U_k(y) \text{ for some } x,y\in V_k,\\ 0 & \text{elsewhere.}\end{cases} \tag{2.15}$$

Next we will use the following definition for the remainder of this paper. For $k,j\geq 1$ and $\delta>0$, define for function $f:E\to\mathbb{R}$,

$$\overline{\mathcal{E}}^{(k)}_{j,\delta}(f,f) := \frac{1}{2}\int\int_{\{(z,w)\in B_j\times B_j:\,\rho(z,w)>\delta\}}(f(w)-f(z))^2 j^{(k)}(w,z)m(dw)m(dz), \tag{2.16}$$

and

$$\mathcal{E}_{j,\delta}(f,f) := \frac{1}{2}\int\int_{\{(z,w)\in B_j\times B_j:\,\rho(z,w)>\delta\}}(f(w)-f(z))^2 J(dw,dz). \tag{2.17}$$

Now we can state the following condition.

(A4). (i) *For any compact subset $K \subset E$,*

$$\lim_{\eta\to 0}\limsup_{k\to\infty}\int\int_{\{(x,y)\in K\times K:\rho(x,y)\le\eta\}}\rho(x,y)^2 j^{(k)}(x,y)m(dx)m(dy)=0, \tag{2.18}$$

$$\lim_{j\to\infty}\limsup_{k\to\infty}\int_K\int_{(B_j)^c} j^{(k)}(x,y)m(dx)m(dy)=0. \tag{2.19}$$

(ii) *For every $\varepsilon>0$, there exists $N>0$ such that for every $k\ge i\ge N$ and $f\in L^2(V_i;m_i)$,*

$$\mathcal{E}^{(k)}(\pi_k E_i f,\,\pi_k E_i f)^{1/2}\le \mathcal{E}^{(i)}(f,f)^{1/2}+\varepsilon.$$

(iii) *For any sufficiently small $\delta>0$ and large $j\in\mathbb{N}$,*

$$\lim_{k\to\infty}\overline{\mathcal{E}}^{(k)}_{j,\delta}(f,f)=\mathcal{E}_{j,\delta}(f,f)\qquad \textit{for every } f\in \mathrm{Lip}_c(E). \tag{2.20}$$

We will also consider in this paper the following alternative conditions to **(A3)** and **(A4)**. First, for $u\in L^2(B_j;m)$, let

$$\overline{\mathcal{L}}^{(k)}_{j,\delta}u(x) \;:=\; \int_{B_j}(u(y)-u(x))j^{(k)}(x,y)\mathbf{1}_{\{\rho(x,y)>\delta\}}m(dy)\qquad \text{for } x\in B_j, \tag{2.21}$$

$$\mathcal{L}_{j,\delta}u(x) \;:=\; \int_{B_j}(u(y)-u(x))\mathbf{1}_{\{\rho(x,y)>\delta\}}j(x,dy)\qquad \text{for } x\in B_j. \tag{2.22}$$

(A3)*. *Condition* **(A3)** *holds and $\mathcal{L}_{j,\delta}f$ is continuous for all $f\in\mathrm{Lip}_c(E)$.*

(A4)*. (i) *Same as* **(A4)**(i).

(ii) *For any sufficiently small $\delta>0$ and large $j\in\mathbb{N}$,*

$$\lim_{k\to\infty}\int_{B_j}(\overline{\mathcal{L}}^{(k)}_{j,\delta}f(x))^2m(dx)=\int_{B_j}(\mathcal{L}_{j,\delta}f(x))^2m(dx),\quad \forall f\in\mathrm{Lip}_c(E).$$

(iii) *For any sufficiently small $\delta>0$ and large $j\in\mathbb{N}$,*

$$\lim_{k\to\infty}\overline{\mathcal{E}}^{(k)}_{j,\delta}(f,f)=\mathcal{E}_{j,\delta}(f,f)\qquad \textit{for every } f\in C_b(B_j).$$

Note that, by the polarization identity, **(A4)*** (iii) is equivalent to

$$\lim_{k\to\infty}\overline{\mathcal{E}}^{(k)}_{j,\delta}(f,g)=\mathcal{E}_{j,\delta}(f,g)\qquad \text{for every } f,g\in C_b(B_j). \tag{2.23}$$

For every function $\varphi\in C_c^+(E)$, we define measures

$$\mathbb{P}^{(k)}_{\varphi}(\,\cdot\,)\;:=\;\sum_{x\in V_k}\mathbb{P}^{(k)}_x(\,\cdot\,)\varphi(x)m_k(x)\qquad\text{and}\qquad \mathbb{P}_{\varphi}(\,\cdot\,)\;:=\;\int_E\mathbb{P}_x(\,\cdot\,)\varphi(x)m(dx). \tag{2.24}$$

The following are two of the main results of this paper.

Theorem 2.2 *Assume that* **(A1)**–**(A2)** *hold and that the symmetric Hunt process X on E associated with $(\mathcal{E}, \mathcal{F})$ is conservative. Assume further that either* **(A3)**–**(A4)** *hold, or* **(A3)***–**(A4)*** *hold. Then, for any $\varphi \in C_c^+(E)$, the symmetric Hunt process $\{(X^{(k)}, \mathbb{P}_\varphi^{(k)});\ k \geq 1\}$ on V_k associated with $(\mathcal{E}^{(k)}, \mathcal{F}^{(k)})$ converges weakly to $(X, \mathbb{P}_\varphi)$ on $\mathbb{D}_{E_\partial}[0,1]$ equipped with the Skorohod topology.*

In some of the applications, tightness in the space $\mathbb{D}_{E_\partial}[0,1]$ equipped with Skorohod topology is very difficult to establish, if not impossible. So we need a weaker topology on spaces $\mathbb{D}_{E_\partial}[0,1]$ and $\mathbb{D}_E[0,1]$, namely, the convergence-in-measure topology. This topology was introduced in [12], which is also called pseudo-path topology in literature, see [25, Lemma 1].

Let λ be the Lebesgue measure on $[0,1]$. For a E_∂-valued Borel function w on $[0,1]$, the pseudo-path of w is a probability law on $[0,1] \times E_\partial$: the image measure of λ under the mapping $t \mapsto (t, w(t))$. Denote by Ψ the mapping which associates to a path w its pseudo-path, which identifies two paths if and only if they are equal λ-a.e. on $[0,1]$. In particular, Ψ is one-to-one on $\mathbb{D}_{E_\partial}[0,1]$ and embeds it into the compact space of all probability measures on the compact space $[0,1] \times E_\partial$. Meyer gave the name of the pseudo-path topology to the induced topology on $\mathbb{D}_{E_\partial}[0,1]$. (See [12, chapter IV, n 40-46] for more details.) [25, Theorem 5] tells us that if the law of $\{X^{(k)}, k \geq 1\}$ is tight in $\mathbb{D}_{E_\partial}[0,1]$ equipped with pseudo-path topology, then there is a subsequence $\{n_k\}$ and a subset A of $[0,1]$ having zero Lebesgue measure so that $X^{(n_k)}$ convergence in finite dimensional distribution on $[0,1] \setminus A$.

Tightness of stochastic processes on $\mathbb{D}_{E_\partial}[0,1]$ (respectively, on $\mathbb{D}_E[0,1]$) equipped with the convergence-in-measure topology is closely related to the number of crossing between two disjoint sets by the stochastic processes (see [25]). The latter has been investigated in [7, 23].

Theorem 2.3 *Assume that either* (2.9) *of* **(A1)** *and* **(A2)**–**(A4)** *hold, or* **(A.2)**, **(A.3)*** *and* **(A.4)*** *hold. Then for every $\varphi \in C_c^+(E)$, $\{(X^{(k)}, \mathbb{P}_\varphi^{(k)});\ k \geq 1\}$ converges weakly to $(X, \mathbb{P}_\varphi)$ on $\mathbb{D}_{E_\partial}[0,1]$ equipped with the convergence-in-measure topology.*

The proofs of Theorems 2.2 and 2.3 will be given in Section 5.

3 Tightness

Before we go to tightness results, let's first give a proof of Theorem 2.1, which gives the discrete approximation of state space. We need the following 'nice' open covering of E (see, for example [21, Lemma 3.1], for a proof).

Lemma 3.1 *Suppose (E, ρ, m) is a metric measure space satisfying conditions* **(MMS.1)**–**(MMS.3)**. *Then there exist integers $N_0, L_0 \geq 1$ that depend only on the constant C_* in* **(MMS.3)** *such that for each $r > 0$ there exists an open covering $\{B(x_i, r),\ i \geq 1\}$ of E with the following property:*

- *No point in E is contained in more than N_0 of the balls $\{B(x_i, r),\ i \in \mathbb{N}\}$.*
- *$\{B(x_i, r/2),\ i \in \mathbb{N}\}$ are disjoint.*

- *For each $x \in E$, the number of balls $B(x_i, r)$ which intersects with $B(x, 2r)$ is bounded by L_0.*

Proof of Theorem 2.1. Let $V^{(r)} = \{x_i, i \geq 1\}$, where $\{x_i, i \geq 1\}$ are given in Lemma 3.1. We say two distinct $x, y \in V^{(r)}$ are connected by a *bond* (which we will denote as $\{x, y\} \in \Xi^{(r)}$) if $\rho(x, y) < 3r$. In this way, we can define a graph $(V^{(r)}, \Xi^{(r)})$ of bounded degree. We also define $\{U_{(r)}(x)\}_{x \in V^{(r)}}$, an associated partition of E, as follows; $U_{(r)}(x_1) = \overline{B(x_1, r)}$ and $U_{(r)}(x_k) = \overline{B(x_k, r)} \setminus \cup_{i=1}^{k-1} B(x_i, r)$ for $k \geq 2$. Clearly, $c_1 V(x_i, r) \leq m(U_{(r)}(x_i)) \leq V(x_i, r)$ and $U_{(r)}(x_i) \cap U_{(r)}(x_j) \subset \cup_{k=1}^{j} \partial B(x_k, r)$ for $i < j$. The definition of $(V^{(r)}, \Xi^{(r)})$ and partition $\{U_{(r)}(x),\, x \in V^{(r)}\}$ depends on the choice of the open covering of E (and its labeling). In the following, for each $r > 0$, we choose one open covering with the above mentioned property and fix the graph $(V^{(r)}, \Xi^{(r)})$ and a partition $\{U_{(r)}(x),\, x \in V^{(r)}\}$. For each sequence (r_m) which converges to zero, the set $\cup_m V^{(r_m)}$ is dense in E. Note that since ρ is geodesic, for each $x \in V^{(r)}$, there exists $y \in V^{(r)} \setminus \{x\}$ such that $y \in B(x, 2r)$. So $(V^{(r)}, \Xi^{(r)})$ is connected. Further, $(V^{(r)}, \Xi^{(r)})$ has bounded degree, i.e. $\sup_{x \in V^{(r)}} \sharp\{y \in V^{(r)} : \{x, y\} \in \Xi^{(r)}\} < \infty$. Let $\rho^{(r)}$ be the graph distance of $(V^{(r)}, \Xi^{(r)})$; then

$$\frac{r}{2}\, \rho^{(r)}(x, y) \leq \rho(x, y) < 3r \rho^{(r)}(x, y) \qquad \text{for } x, y \in V^{(r)}. \tag{3.1}$$

Clearly, this holds if $\{x, y\} \in \Xi^{(r)}$. In general, the second inequality of (3.1) clearly holds and the first inequality can be verified as follows. Let γ be a geodesic connecting x and y. Set $k = [1 + r^{-1}\rho(x, y)]$, the largest integer not exceeding $1 + r^{-1}\rho(x, y)$. Let $\{y_i, 0 \leq i \leq k\}$ be equally spaced points on γ so that $\rho(y_{i-1}, y_i) = \rho(x, y)/k < r$ for $k = 1, \cdots, k$ with $y_0 = x$ and $y_k = y$. For each $1 \leq i \leq k = 1$, there is some $x_i \in V^{(r)}$ so that $y_i \in B(x_i, r_i)$ (we take $x_0 = y_0 = x$ and $x_k = y_k = y$). By the triangle inequality,

$$\rho(x_{i-1}, x_i) \leq \rho(x_{i-1}, y_{i-1}) + \rho(y_{i-1}, y_i) + \rho(y_i, x_i) < 3r \qquad \text{for } i = 1, \cdots, k.$$

This shows that $\rho^{(r)}(x, y) \leq k \leq 2\rho(x, y)/r$, establishing the first inequality in (3.1). Let $V_k := V^{(1/k)}$, $\Xi_k := \Xi^{(1/k)}$, $\rho_k := \rho^{(1/k)}$ and $U_k(x) := U_{(1/k)}(x)$. It is now easy to verify that (V_k, Ξ_k, ρ_k) together with $\{U_k(x),\, x \in V_k\}$ satisfies **(AG.1)**–**(AG.3)**. □

Recall that (E, ρ, m) is a metric measure space satisfying conditions **(MMS.1)**–**(MMS.3)** and that $\{(V_k, \Xi_k), k \geq 1\}$, with the graph distance ρ_k, are approximating graphs with associated partitions $\{U_k(x), x \in V_k; k \geq 1\}$ satisfying **(AG.1)**–**(AG.3)**.

We now investigate the tightness of the continuous time random walks on graphs V_k. Recall that m_k is defined in (2.6), $m_k(y) = m_k(\{y\})$ and the Dirichlet form $(\mathcal{E}^{(k)}, \mathcal{F}^{(k)})$ defined in (2.8).

Theorem 3.2 *$(\mathcal{E}^{(k)}, \mathcal{F}^{(k)})$ is a regular Dirichlet form on $L^2(V_k; m_k)$ with $C_c(V_k) \subset \mathcal{F}^{(k)}$. If*

$$\sup_{x \in V_k} \sum_{y \in V_k} j^{(k)}(x, y) m_k(y) < \infty, \tag{3.2}$$

then the symmetric Hunt process $X^{(k)}$ on V_k associated with the regular Dirichlet form $(\mathcal{E}^{(k)}, \mathcal{F}^{(k)})$ is conservative.

Proof. For $f \in C_c(V_k)$, let K denote its support (note that K is a finite set). Then by (2.5) and (2.7),

$$\begin{aligned}
\mathcal{E}^{(k)}(f,f) &= \frac{1}{2}\sum_{x,y\in K}(f(x)-f(y))^2 j^{(k)}(x,y)m_k(x)m_k(y) \\
&\quad + \sum_{x\in K} f(x)^2 \left(\sum_{y\in K^c} j^{(k)}(x,y)m_k(y)\right) m_k(x) \\
&\le 3\|f\|_\infty^2 \sum_{x\in K}\left(\sum_{y\in V_k} j^{(k)}(x,y)m_k(y)\right) m_k(x) \\
&\le 3\|f\|_\infty^2 m_k(K)\left(\max_{x\in K}\sum_{y\in V_k} j^{(k)}(x,y)m_k(y)\right) < \infty.
\end{aligned}$$

This shows that $f \in \mathcal{F}^{(k)}$ and so $C_c(V_k) \subset \mathcal{F}^{(k)}$. Let K_j be an increasing sequence of compact (or equivalently, finite) subsets of V_k with $\cup_{j\ge 1}K_j = V_k$. For every $u \in \mathcal{F}^{(k)}_b$, define $u_j = u - ((-1/j)\vee u))\wedge(1/j)$. By [14, Theorem 1.4.2(iv)], u_j is $\mathcal{E}^{(k)}_1$-convergent to u where $\mathcal{E}^{(k)}_1(\cdot,\cdot) := \mathcal{E}^{(k)}(\cdot,\cdot)+\|\cdot\|_2^2$. Since $u \in L^2(V_k;m_k)$, $\mathrm{supp}[u_j] \subset \{x\in V_k : |u(x)| > 1/j\}$ is a finite set. Consequently $u_j \in C_c(V_k)$ and so $(\mathcal{E}^{(k)},\mathcal{F}^{(k)})$ is a regular Dirichlet form on $L^2(V_k;m_k)$. Thus there is an associated m_k-symmetric Hunt process $X^{(k)}$ on V_k.

To prove the second claim of Theorem 3.2, we will use [24, Theorem 3.1]. Note that ρ_k is a discrete metric on V_k, so in view of (2.5), condition (3.2) is equivalent to having

$$\sup_{x\in V_k}\sum_{y\in V_k}(\rho_k(x,y)^2\wedge 1)j^{(k)}(x,y)m_k(y) < \infty.$$

Thus, since $j^{(k)}(x,y)$ is symmetric, under the assumption (3.2), [24, Condition (C)] holds. Thus, to apply [24, Theorem 3.1] to deduce the conservativeness of $X^{(k)}$, we only need to check that $x \to e^{-\lambda\rho_k(x,x_0)} \in L^1(V_k;m_k)$ for some $x_0 \in V_k$.

Fix some $x_0 \in V_k$. Note that for $r > 0$, by **(AG.1)**–**(AG.3)** and (2.1)

$$\begin{aligned}
m_k(B(x_0,r)) &:= \sum_{y\in V_k,\rho_k(x_0,y)\le r} m_k(y) \;=\; m\Big(\bigcup_{y\in V_k,\rho_k(x_0,y)\le r} U_k(y)\Big) \\
&\le m\left(B(x_0, C_2 r + C_3)\right) \;\le\; c(r+1)^{d_0}.
\end{aligned}$$

Thus for every $\lambda > 0$,

$$\begin{aligned}
\int_{V_k} e^{-\lambda\rho_k(x,x_0)} m_k(dx) &= \int_0^\infty e^{-\lambda r} d(m_k(B(x_0,r)) \;=\; \lambda\int_0^\infty m_k(B(x_0,r))\, e^{-\lambda r}\, dr \\
&\le c\,\lambda\left(\int_0^\infty (1+r)^{d_0}\, e^{-\lambda r} dr\right) < \infty.
\end{aligned}$$

So we conclude from [24, Theorem 3.1] that under the condition (3.2), $X^{(k)}$ is conservative. $\square$

Recall that $\mathbb{P}^{(k)}_\varphi$ and $\mathbb{P}_\varphi$ is defined for every positive function $\varphi \in C_c(E)$ in (2.24).

Lemma 3.3 *Assume condition* **(A1)** *holds. Then for every $g \in \mathrm{Lip}_c(E)$, there exists a positive constant c such that for every $k \geq k_0$ and $0 \leq s < t < \infty$,*

$$\int_s^t \sum_{y \in V_k} (g(X_u^{(k)}) - g(y))^2 j^{(k)}(X_u^{(k)}, y) m_k(y) du \;\leq\; c(t-s).$$

Proof. Let Λ be the Lipschitz constant of g. There is an integer $j \geq 1$ so that the topological support K of g is contained in ball B_j centered at x_0 with radius j. Let $K_1 := \overline{B_{j+1}}$ and $K_2 := \overline{B_{j+3}}$. By (2.3) and (2.9)–(2.10),

$$\begin{aligned}
&\sup_{x \in V_k} \sum_{y \in V_k} (g(x) - g(y))^2 j^{(k)}(x, y) m_k(y) \\
= \;& \sup_{x \in V_k} \left(\sum_{y \in K_1^c \cap V_k} g(x)^2 j^{(k)}(x, y) m_k(y) + \sum_{y \in K_1 \cap V_k} (g(x) - g(y))^2 j^{(k)}(x, y) m_k(y) \right) \\
\leq \;& \|g\|_\infty^2 \sup_{x \in K \cap V_k} \sum_{y \in K_1^c \cap V_k} j^{(k)}(x, y) m_k(y) + \sup_{x \in K_2^c \cap V_k} \sum_{y \in K_1 \cap V_k} g(y)^2 j^{(k)}(x, y) m_k(y) \\
& + \sup_{x \in K_2 \cap V_k} \sum_{y \in K_1 \cap V_k} \left(\Lambda^2 \rho(x, y)^2 \wedge 4\|g\|_\infty^2 \right) j^{(k)}(x, y) m_k(y) \\
\leq \;& c_1 \|g\|_\infty^2 + c_2 \sup_{x \in K_2 \cap V_k} \sum_{y \in V_k} \left(\frac{\rho_k(x, y)}{k} \wedge 1 \right)^2 j^{(k)}(x, y) m_k(y) \leq c_3,
\end{aligned}$$

where c_1, c_2 and c_3 are positive constants independent of $k \geq k_0$. The conclusion of the lemma follows directly from the above inequality. □

Recall that E_∂ is the one-point compactification of E and the space $\mathbb{D}$ of right continuous functions having left limits is defined in (1.2). Clearly $X^{(k)} \in \mathbb{D}_{E_\partial}[0, \infty)$.

Since $\mathrm{Lip}_c^+(E) = \{f \in \mathrm{Lip}_c(E) : f \geq 0\}$ separates points of E, using Stone-Weierstrass theorem, it is easy to check that $\mathrm{Lip}_c^+(E)$ is a dense subset of $C_\infty^+(E)$ (space of non-negative continuous functions on E that vanishes at infinity).

Proposition 3.4 *Assume* **(A1)** *holds and let $\zeta^{(k)}$ denote the lifetime of the process $X^{(k)}$. Then, for any $\varphi \in C_c^+(E)$, $T > 0$, $m \geq 1$ and $\{g_1, \cdots, g_m\} \subset \mathrm{Lip}_c^+(E)$, the laws of $\left\{(g_1, \cdots, g_m)(X^{(k)})\right\}_{k \geq 1}$ on $\{\zeta^{(k)} > T\}$ with initial distribution $\varphi(x) m_k(dx)$ is tight in $\mathbb{D}_{\mathbb{R}^m}[0, T]$ equipped with the Skorohod topology. Moreover, the laws of $\left\{X_t^{(k)}, t \in [0, T]\right\}$ on $\{\zeta^{(k)} > T\}$ with initial distribution $\varphi(x) m_k(dx)$ is tight in $\mathbb{D}_{E_\partial}[0, T]$ equipped with the Skorohod topology.*

Proof. Without loss of generality, we assume that $m = 1$, $T = 1$ and $g = g_1$. We first show that $\left\{\left(g(X^{(k)}), \mathbb{P}_\varphi^{(k)}\right);\, k \geq 1\right\}$ is relatively compact in $\mathbb{D}_{\mathbb{R}}[0, 1]$ equipped with the Skorohod topology.

Given $t > 0$ and a path $\omega \in \mathbb{D}_E[0, 1]$, the time reversal operator r_t is defined by

$$r_t(\omega)(s) := \begin{cases} \omega((t-s)-), & \text{if } \; 0 \leq s \leq t, \\ \omega(0) & \text{if } s \geq t. \end{cases}$$

Here for $r > 0$, $\omega(r-) := \lim_{s\uparrow r} \omega(s)$ is the left limit at r and we use the convention that $\omega(0-) := \omega(0)$

Since $f|_{V_k} \in \mathcal{F}^{(k)}$ for every $f \in \mathrm{Lip}_c(E)$, by the same argument as that for [6, (2.3)] (see also [8]), we have the following forward-backward martingale decomposition of $f(X_t^{(k)})$ for every $f \in \mathrm{Lip}_c(E)$; There exists a martingale $M^{k,f}$ such that on $\{\zeta^{(k)} > 1\}$,

$$f(X_t^{(k)}) - f(X_0^{(k)}) = \frac{1}{2} M_t^{k,f} - \frac{1}{2}(M_1^{k,f} - M_{(1-t)-}^{k,f}) \circ r_1, \quad t \in [0,1]. \tag{3.3}$$

By [8, Proposition 2.8], for each $M^{k,f}$, there exists the continuous predictable quadratic variation process $\langle M^{k,f}\rangle_t$. Note that (for example, see [14, page 214])

$$\langle M^{k,f}\rangle_t - \langle M^{k,f}\rangle_s = \int_s^t \sum_{y \in V_k} \big(f(X_u^{(k)}) - f(y)\big)^2 j^{(k)}\big(X_u^{(k)}, y\big) m_k(y) du.$$

Thus by Lemma 3.3 and [16, Proposition VI.3.26], $\{\langle M^{k,f}\rangle_t\}_{k\geq 1}$ is C-tight in $\mathbb{D}_{\mathbb{R}}[0,1]$ equipped with the Skorohod topology, i.e., $\{\langle M^{k,f}\rangle_t\}_{k\geq 1}$ is tight in $\mathbb{D}_{\mathbb{R}}[0,1]$ equipped with the Skorohod topology and all limit points of $\{\langle M^{k,f}\rangle_t\}_{k\geq 1}$ are laws of continuous processes. As m_k converges weakly to m, by [16, Theorem VI.4.13] the laws of $\{M^{k,f}\}_{k\geq 1}$ is tight in $\mathbb{D}_{\mathbb{R}}[0,1]$ with the initial distribution $\mathbb{P}_h^{(k)}$ for every $h \in \mathrm{Lip}_c^+(E)$. Thus the laws of $\{M^{k,f}, \mu_{h_1,h_2}^{(k)}\}_{k\geq 1}$ is tight in the sense of Skorohod topology on $\mathbb{D}_{\mathbb{R}}[0,1]$ for every $h_1, h_2 \in \mathrm{Lip}_c^+(E)$ where

$$\mu_{h_1,h_2}^{(k)}(A) := \mathbb{E}\left[h_1(X_0^{(k)}(\omega))\mathbf{1}_A(\omega) h_2(X_1^{(k)}(\omega));\, \zeta^{(k)} > 1\right], \quad \forall A \in \mathcal{B}(\mathbb{D}_E[0,1]).$$

Note that for every $A \in \mathcal{B}(\mathbb{D}_E[0,1])$,

$$\begin{aligned}
\mu_{h_1,h_2}^{(k)}(A \circ r_1) &= \mathbb{E}\left[h_1(X_0^{(k)}(\omega))\mathbf{1}_A \circ r_1(\omega) h_2(X_1^{(k)}(\omega));\, \zeta^{(k)} > 1\right] \\
&= \mathbb{E}\left[h_2(X_0^{(k)}(\omega))\mathbf{1}_A(\omega) h_1(X_1^{(k)}(\omega));\, \zeta^{(k)} > 1\right] \\
&= \mu_{h_2,h_1}^{(k)}(A).
\end{aligned}$$

Thus the laws of $\{M^{k,f}, \mu_{h_1,h_2}^{(k)}\}_{k\geq 1}$ is the same as the laws of $\{M^{k,f} \circ r_1, \mu_{h_2,h_1}^{(k)}\}_{k\geq 1}$ and so the laws of $\{M^{k,f} \circ r_1, \mu_{h_1,h_2}^{(k)}\}_{k\geq 1}$ is tight in the sense of Skorohod topology on $\mathbb{D}_{\mathbb{R}}[0,1]$ for every $h_1, h_2 \in \mathrm{Lip}_c(E)$, too. So the laws of $\big\{M^{k,f}, \mu_{\varphi,f}^{(k)}\big\}_{k\geq 1}$ and the laws of $\big\{M^{k,f} \circ r_1, \mu_{\varphi,f}^{(k)}\big\}_{k\geq 1}$ are tight. Since the laws of $\big\{g(X^{(k)}), \mathbb{P}_\varphi^{(k)}\big\}_{k\geq 1}$ restricted to $\{\zeta^{(k)} > 1\}$ are the same as $\big\{g(X^{(k)}), \mu_{\varphi,g}^{(k)}\big\}_{k\geq 1}$ in $\mathbb{D}_{\mathbb{R}}[0,1]$, by (3.3) $\big\{g(X^{(k)}),\, \mathbb{P}_\varphi^{(k)}\big\}_{k\geq 1}$ restricted to $\{\zeta^{(k)} > 1\}$ is tight (and so relatively compact) in the sense of Skorohod topology on $\mathbb{D}_{\mathbb{R}}[0,1]$.

Since E_∂ is compact and the linear span of $\mathrm{Lip}_c^+(E)$ and constants is a dense subset in $C(E_\partial)$ equipped with uniform topology, we conclude from [13, Theorem 3.9.1 and Corollary 3.9.3] that the laws of $\big\{X_t^{(k)}, t \in [0,1]\big\}$ on $\{\zeta^{(k)} > 1\}$ with initial distribution $\varphi(x)m_k(dx)$ is tight in $\mathbb{D}_{E_\partial}[0,1]$ equipped with the Skorohod topology. $\square$

4 Mosco convergence

We will show in Section 5 that symmetric continuous time random walk $X^{(k)}$ with conductance $\mathcal{C}^{(k)}$ converges to the symmetric Hunt processes X associated with $(\mathcal{E}, \mathcal{F})$ in the sense of finite dimensional distributions (Theorem 5.1). One way to establish this is to show that corresponding Dirichlet form converges in the sense of Mosco, a concept introduced in [26]. In [26], a symmetric bilinear form $a(u, u)$ defined on a linear subspace $\mathcal{D}[a]$ of a Hilbert space $\mathcal{H}$ is extended to the whole space $\mathcal{H}$ by defining $a(u, u) = \infty$ for every $u \in \mathcal{H} \setminus \mathcal{D}[a]$. We will use this extension throughout this paper. In [26], Mosco showed that the Mosco convergence of a sequence of densely defined symmetric closed forms defined on the same Hilbert space is equivalent to the convergence of the sequence of semigroups in strong operator sense. However, in many cases, semigroups and their associated closed forms may live on different Hilbert spaces. Fortunately, the Mosco convergence theory can be extended to cover these cases of varying state spaces. Theorem 8.3 in the Appendix, which was obtained in [17] and [18, Theorem 2.5], gives one such extension. See [22] for another extension.

In this section, we establish the Mosco convergence of $(\mathcal{E}^{(k)}, \mathcal{F}^{(k)})$ in the sense of Definition 8.1 under two sets of conditions. We first prove some basic facts on the restriction and extension operators.

Recall the restriction operator $\pi_k : L^2(E; m) \to L^2(V_k; m_k)$ and the extension operator $E_k : L^2(V_k; m_k) \to L^2(E; m)$ defined in (2.13) and (2.14), respectively. Let $\langle \cdot, \cdot \rangle_k$ (resp. $\langle \cdot, \cdot \rangle$) be the inner product in Hilbert space $L^2(V_k; m_k)$ (resp. $L^2(E; m)$) and $\| \cdot \|_{k,p}$ (resp. $\| \cdot \|_p$) be the L^p-norm of $L^p(V_k; m_k)$ (resp. $L^p(E; m)$).

Lemma 4.1 (i) *π_k is a bounded operator from $L^2(E; m)$ to $L^2(V_k; m_k)$ with $\sup_{k \geq 1} \|\pi_k\| \leq 1$, where $\|\pi_k\|$ is the operator norm of π_k. Further, $\lim_{k\to\infty} \|\pi_k f\|_{k,2} = \|f\|_2$ for every $f \in L^2(E; m)$.*

(ii) *For each $f_k \in L^2(V_k; m_k)$, we have the following;*

$$\pi_k E_k f_k = f_k \qquad m\text{-a.e.}, \tag{4.1}$$

$$\langle \pi_k g, f_k \rangle_k = \langle g, E_k f_k \rangle \qquad \text{for every } g \in L^2(E; m). \tag{4.2}$$

(iii) *For every $f \in L^2(V_k; m_k)$, $E_k f \in L^2(E; m)$ and $\|E_k f\|_2^2 = \|E_k(f^2)\|_1 = \|f\|_{k,2}^2$.*

(iv) *For every $f \in L^2(E; m)$, $E_k \pi_k f$ converges strongly to f in $L^2(E; m)$.*

(v) *Suppose $f \in C_c(E)$. Let $f_k := f|_{V_k} \in L^2(V_k; m_k)$. Then $E_k f_k$ converges strongly to f in $L^2(E; m)$.*

Proof. (i) By the Cauchy-Schwarz inequality,

$$\begin{aligned} \|\pi_k f\|_{k,2}^2 &= \sum_{x\in V_k} m_k(x) \left(\frac{1}{m_k(x)} \int_{U_k(x)} f(y)m(dy) \right)^2 \\ &\le \sum_{x\in V_k} \frac{m_k(x)}{m_k(x)} \int_{U_k(x)} f(y)^2 m(dy) \;=\; \|f\|_2^2. \end{aligned} \tag{4.3}$$

Moreover, by the uniform continuity, we easily see from (4.3) that $\lim_{k\to\infty} \|\pi_k f\|_{k,2}^2 = \|f\|_2^2$ for $f \in C_c(E)$. As $C_c(E)$ is dense in $L^2(E;m)$ and $\|\pi_k\| \le 1$, we have $\lim_{k\to\infty} \|\pi_k f\|_{k,2}^2 = \|f\|_2^2$ for $f \in L^2(E;m)$.

(ii) (4.1) is clear from the definitions of π_k and E_k. The left hand side of (4.2) is

$$\sum_{x\in V_k} \frac{1}{m_k(x)} \int_{U_k(x)} g(y)m(dy) f_k(x) m_k(x).$$

By Fubini's theorem, the above is equal to

$$\int_E \sum_{x\in V_k} f_k(x) g(y) \mathbf{1}_{U_k(x)}(y) m(dy) \;=\; \langle E_k f_k, g\rangle.$$

(iii) Note that, since $m(U_k(x)\cap U_k(y)) = 0$ for $x \ne y$, we have for $f \in L^2(V_k; m_k)$

$$\begin{aligned} \|E_k f\|_2^2 &= \int_E \Big(\sum_{x\in V_k} f(x) \mathbf{1}_{U_k(x)}(y) \Big)^2 m(dy) \\ &= \int_E \sum_{x\in V_k} f(x)^2 \mathbf{1}_{U_k(x)}(y) m(dy) \;=\; \|E_k(f^2)\|_1. \end{aligned}$$

Moreover, by Fubini's theorem,

$$\int_E \sum_{x\in V_k} f(x)^2 \mathbf{1}_{U_k(x)}(y) m(dy) \;=\; \sum_{x\in V_k} f(x)^2 m_k(x) = \|f\|_{k,2}^2.$$

(iv) First assume that $f \in C_c(E)$. Let $K := \{x \in E : \rho(x, \mathrm{supp}[f]) \le 1\}$. By the Cauchy-Schwarz inequality, for sufficiently large $k \ge 1$,

$$\begin{aligned} \|E_k f - f\|_2^2 &= \int_K |E_k f_k(x) - f(x)|^2 \, m(dx) \\ &\le \sum_{z\in V_k\cap K} \int_{U_k(z)} \left(\frac{1}{m_k(z)} \int_{U_k(z)} (f(y) - f(x)) m(dy) \right)^2 m(dx) \\ &\le \sum_{z\in V_k\cap K} \frac{1}{m_k(z)} \int_{U_k(z)\times U_k(z)} (f(y) - f(x))^2 m(dy) m(dx), \end{aligned}$$

which, by the uniform continuity of $f \in C_c(E)$, tends to zero as $k \to \infty$. That is, for $f \in C_c(E)$, $E_k \pi_k f$ converges strongly in $L^2(E;m)$ to f. Since by (i) and (iii),

$$\|E_k \pi_k f\|_2 = \|\pi_k f\|_{k,2} \le \|f\|_2 \qquad \text{for } f \in L^2(E;m)$$

and that $C_c(E)$ is dense in $L^2(E;m)$, we conclude that for every $f \in L^2(E;m)$, $E_k \pi_k f$ converges strongly in $L^2(E;m)$ to f.

(v) Let $K := \{x \in E : \rho(x, \mathrm{supp}[f]) \le 1\}$. Then for k sufficiently large,

$$\int_E |E_k f_k(x) - f(x)|^2 m(dx) = \int_K |E_k f_k(x) - f(x)|^2 m(dx),$$

which goes to zero by the uniform continuity of f. □

4.1 Mosco convergence under conditions (2.9) and (A2)–(A4)

We start with

Lemma 4.2 *Under the condition* **(A2)**, $\mathrm{Lip}_c(E) \subset \mathcal{F}$.

Proof. Let $u \in \mathrm{Lip}_c(E)$. Clearly it is $L^2(E;m)$-integrable. Denote by Λ the Lipschitz constant of u and $K := \mathrm{supp}[u]$. Then by the symmetry of $j(x, dy)$,

$$\begin{aligned} \mathcal{E}(u,u) &\le \int_K \left(\int_{E \setminus \{x\}} (u(x) - u(y))^2 j(x, dy) \right) m(dx) \\ &\le \int_K \left(\int_E \left(\Lambda^2 \rho(x,y)^2 \mathbf{1}_{\{\rho(x,y) \le 1\}} + 4\|u\|_\infty^2 \mathbf{1}_{\{\rho(x,y) > 1\}} \right) j(x, dy) \right) m(dx) \\ &\le c\, m(K) \sup_{x \in K} \int_E \left(\rho(x,y)^2 \wedge 1 \right) j(x, dy), \end{aligned}$$

which is finite by condition (2.11). This proves that $u \in \mathcal{F}$. □

Lemma 4.2 in particular implies that $\mathcal{F}$ is a dense linear subspace of $L^2(E;m)$. It is easy to check by using Fatou's lemma that $(\mathcal{E}, \mathcal{F})$ is a Dirichlet form on $L^2(E;m)$ (cf. [14, Example 1.2.4]).

Recall that we have fixed some $x_0 \in E$ and $B_r = B(x_0, r)$, and that quadratic forms $\overline{\mathcal{E}}^{(k)}_{j,\delta}$ and $\mathcal{E}_{j,\delta}$ are defined in (2.16) and (2.17), respectively. Recall also that the definition of $j^{(k)}$ has been extended to be defined on $E \times E$ by (2.15). For $f : E \to \mathbb{R}$, we define

$$\overline{\mathcal{E}}^{(k)}(f,f) := \frac{1}{2} \int_{E \times E} (f(w) - f(z))^2 j^{(k)}(w,z) m(dw) m(dz).$$

Note that for function f on V_k, $(E_k f(z) - E_k f(w))^2 = (f(x) - f(y))^2$ where $x, y \in V_k$ with $z \in U_k(x), w \in U_k(y)$. Thus

$$\overline{\mathcal{E}}^{(k)}(E_k u,\, E_k u) \;=\; \mathcal{E}^{(k)}(u,u), \qquad \text{for all } u \in \mathcal{F}^{(k)} \tag{4.4}$$

Remark 4.3 It follows from (2.11) of **(A2)** that for every compact subset $K \subset E$

$$\lim_{\eta \to 0} \int_{\{(x,y) \in K \times E : \rho(x,y) \le \eta\}} \rho(x,y)^2 j(x, dy) m(dx) = 0. \tag{4.5}$$

□

Lemma 4.4 *Suppose the conditions* **(A2)**, **(A3)** *and* **(A4)** (i)(iii) *hold, then for every* $f \in \mathrm{Lip}_c(E)$, $\lim_{k\to\infty} \mathcal{E}^{(k)}(\pi_k f, \pi_k f) = \mathcal{E}(f, f)$.

Proof. First, note that by (4.4), $\mathcal{E}^{(k)}(\pi_k f, \pi_k f) = \overline{\mathcal{E}}^{(k)}(E_k\pi_k f,\, E_k\pi_k f)$.

Fix $f \in \mathrm{Lip}_c(E)$ and let K be the support of f, $K_1 := \{x \in E : \rho(x, K) \le 1\}$ and $M_f := \sup_{x\in E} |f(x)|$. Then, by (2.19) and the symmetry of $j^{(k)}$ for each $\varepsilon > 0$, there exists j_0 such that the following holds for $j \ge j_0$,

$$
\begin{aligned}
&\limsup_{k\to\infty} \frac{1}{2}\int\int_{(B_j\times B_j)^c} (E_k\pi_k f(x) - E_k\pi_k f(y))^2 j^{(k)}(x,y)m(dx)m(dy)\\
\le\ & (2M_f)^2 \limsup_{k\to\infty} \int_K \int_{(B_j)^c} j^{(k)}(x,y)m(dx)m(dy) < \varepsilon.
\end{aligned}
$$

Similarly, using (2.11) and choosing j_0 larger if necessary, we have

$$
\frac{1}{2}\int\int_{(B_j\times B_j)^c} (f(x) - f(y))^2 J(dx, dy) < \varepsilon.
$$

Since $f \in \mathrm{Lip}_c(E)$ is Lipschitz continuous, using **(AG.2)**, **(AG.3)**, (2.18) and (4.5) and arguing similarly, we have

$$
\limsup_{k\to\infty} \frac{1}{2}\int\int_{\{(x,y)\in K_1\times K_1:\rho(x,y)\le\delta\}} (E_k\pi_k f(x) - E_k\pi_k f(y))^2 j^{(k)}(x,y)m(dx)m(dy) < \varepsilon
$$

and

$$
\frac{1}{2}\int\int_{\{(x,y)\in K_1\times K_1:\rho(x,y)\le\delta\}} (f(x) - f(y))^2 J(dx, dy) < \varepsilon
$$

for all $\delta \in (0, 1)$. Thus, it is enough to show the following for any sufficiently small δ and large j:

$$
\lim_{k\to\infty} \overline{\mathcal{E}}^{(k)}_{j,\delta}(E_k\pi_k f,\, E_k\pi_k f) = \mathcal{E}_{j,\delta}(f, f).
$$

By the symmetry of $\mathcal{E}^{(k)}_{j,\delta}$ and Lemma 4.1(iv),

$$
\begin{aligned}
&\lim_{k\to\infty} \left|\overline{\mathcal{E}}^{(k)}_{j,\delta}(E_k\pi_k f, E_k\pi_k f)^{1/2} - \overline{\mathcal{E}}^{(k)}_{j,\delta}(f, f)^{1/2}\right|\\
\le\ & \lim_{k\to\infty} \overline{\mathcal{E}}^{(k)}_{j,\delta}(f - E_k\pi_k f,\, f - E_k\pi_k f)^{1/2}\\
=\ & \lim_{k\to\infty} \left(\frac{1}{2}\int_{B_j\times B_j} \big((f - E_k\pi_k f)(x) - (f - E_k\pi_k f)(y)\big)^2 j^{(k)}(x,y)\mathbf{1}_{\{\rho(x,y)>\delta\}}m(dx)m(dy)\right)^{1/2}\\
\le\ & \lim_{k\to\infty} \left(\int_{B_j} (f(x) - E_k\pi_k f(x))^2 \left(\int_{B_j} j^{(k)}(x,y)\mathbf{1}_{\{\rho(x,y)>\delta\}}m(dy)\right) m(dx)\right)^{1/2}\\
\le\ & \lim_{k\to\infty} c(j,\delta)\, \|f - E_k\pi_k f\|_2 = 0.
\end{aligned}
$$

Hence we have

$$
\lim_{k\to\infty} \overline{\mathcal{E}}^{(k)}_{j,\delta}(E_k\pi_k f, E_k\pi_k f) = \lim_{k\to\infty} \overline{\mathcal{E}}^{(k)}_{j,\delta}(f, f).
$$

On the other hand, by (2.20), $\lim_{k\to\infty} \overline{\mathcal{E}}^{(k)}_{j,\delta}(f,f) = \mathcal{E}_{j,\delta}(f,f)$. This completes the proof of the Lemma. $\square$

Theorem 4.5 *Suppose the conditions* (2.9) *of* **(A1)** *and* **(A2)**–**(A4)** *hold, then* $(\mathcal{E}^{(k)}, \mathcal{F}^{(k)})$ *is Mosco convergent to* $(\mathcal{E}, \mathcal{F})$ *in the generalized sense of Definition 8.1.*

Proof. Take $\mathcal{D} = \mathrm{Lip}_c(E)$ in Lemma 8.2. Then, by our assumption **(A3)** and Lemmas 4.2, 4.4, 8.2, we only need to check condition (i) in Definition 8.1.

It is enough to consider sequences $\{u_k\}_{k\geq 1} \subset L^2(V_k; m_k)$ such that $E_k u_k$ converges weakly to $u \in L^2(E; m)$ and $\liminf_{k\to\infty} \mathcal{E}^{(k)}(u_k, u_k) < \infty$. Taking a subsequence if necessary, we may and do assume that $\lim_{k\to\infty} \mathcal{E}^{(k)}(u_k, u_k)$ exists and is finite, and that

$$\sup_{k\geq 1} \left(\mathcal{E}^{(k)}(u_k, u_k) + \sum_{x\in V_k} u_k(x)^2 m_k(x) \right) < \infty. \tag{4.6}$$

So in particular, $u_k \in \mathcal{F}^{(k)}$ for every $k \geq 1$. By uniform boundedness principle, $\{E_k u_k;\, k \geq 1\}$ is a bounded sequence on $L^2(E; m)$.

By the Banach-Saks theorem, taking a subsequence if necessary, $v_k := \frac{1}{k}\sum_{i=1}^k E_i u_i$ converges to some v_∞ in $L^2(E; m)$. Since $E_k u_k$ converges weakly to u in $L^2(E; m)$, v_∞ must be u m-a.e. on E.

Fix an integer $j \geq 1$ and $\delta > 0$. For $\varepsilon > 0$, let $f \in \mathrm{Lip}_c(E)$ such that $\|u - f\|_2 \leq \varepsilon / \sqrt{2a_{j,\delta}}$, where

$$a_{j,\delta} := \max\left\{ \sup_{k\geq k_0} \sup_{x\in B_j} \int_E j^{(k)}(x,y)\, \mathbf{1}_{\{\rho(x,y)>\delta\}}\, m(dy),\ \sup_{z\in B_j} \int_E \mathbf{1}_{\{\rho(x,y)>\delta\}}\, j(z, dw) \right\},$$

which is finite by (2.9) of **(A1)** and **(A2)**(iii). Observe that by **(A4)**(iii)

$$\begin{aligned}
&\limsup_{k\to\infty} \left| \overline{\mathcal{E}}^{(k)}_{j,\delta}(v_k, v_k)^{1/2} - \mathcal{E}_{j,\delta}(f,f)^{1/2} \right| \\
\leq\ & \limsup_{k\to\infty} \left| \overline{\mathcal{E}}^{(k)}_{j,\delta}(v_k, v_k)^{1/2} - \overline{\mathcal{E}}^{(k)}_{j,\delta}(f,f)^{1/2} \right| \\
\leq\ & \limsup_{k\to\infty} \overline{\mathcal{E}}^{(k)}_{j,\delta}(v_k - f, v_k - f)^{1/2} \\
\leq\ & \limsup_{k\to\infty} \left(2\int_{B_j} (v_k(x) - f(x))^2 \left(\int_E j^{(k)}(x,y) \mathbf{1}_{\{\rho(x,y)>\delta\}} m(dy) \right) m(dx) \right)^{1/2} \\
\leq\ & \limsup_{k\to\infty} \sqrt{2a_{j,\delta}}\, \|v_k - f\|_2 = \sqrt{2a_{j,\delta}}\, \|u - f\|_2 < \varepsilon.
\end{aligned}$$

Similarly, we have

$$\left| \mathcal{E}_{j,\delta}(f,f)^{1/2} - \mathcal{E}_{j,\delta}(u,u)^{1/2} \right| \leq \mathcal{E}_{j,\delta}(f-u, f-u)^{1/2} \leq \sqrt{2a_{j,\delta}}\, \|f - u\|_2 < \varepsilon.$$

Thus we have

$$\liminf_{k\to\infty} \overline{\mathcal{E}}^{(k)}_{j,\delta}(v_k, v_k)^{1/2} \geq \mathcal{E}_{j,\delta}(f, f)^{1/2} - \varepsilon \geq \mathcal{E}_{j,\delta}(u, u)^{1/2} - 2\varepsilon. \tag{4.7}$$

Observe that for $k_0 \leq n \leq k$,

$$\begin{aligned}
&\overline{\mathcal{E}}^{(k)}_{j,\delta}(v_n, v_n)^{1/2} - \mathcal{E}_{j,\delta}(f, f)^{1/2} \\
\leq\ & \left|\overline{\mathcal{E}}^{(k)}_{j,\delta}(v_n, v_n)^{1/2} - \overline{\mathcal{E}}^{(k)}_{j,\delta}(f, f)^{1/2}\right| \\
\leq\ & \overline{\mathcal{E}}^{(k)}_{j,\delta}(v_n - f, v_n - f)^{1/2} \\
\leq\ & \left(2\int_{B_j} (v_n(x) - f(x))^2 \left(\int_E j^{(k)}(x, y)\mathbf{1}_{\{\rho(x,y)>\delta\}} m(dy)\right) m(dx)\right)^{1/2} \\
\leq\ & \sqrt{2a_{j,\delta}}\, \|v_n - f\|_2.
\end{aligned}$$

Thus

$$\lim_{n\to\infty} \sup_{k\geq n} \overline{\mathcal{E}}^{(k)}_{j,\delta}(v_n, v_n)^{1/2} \leq \mathcal{E}_{j,\delta}(f, f)^{1/2} + \sqrt{2a_{j,\delta}}\, \|u - f\|_2 \leq \mathcal{E}_{j,\delta}(f, f)^{1/2} + \varepsilon < \infty.$$

By condition **(A4)**(ii) and the above, there exists $N > 0$ such that for every $k \geq i \geq N$,

$$\mathcal{E}^{(k)}(\pi_k E_i u_i,\, \pi_k E_i u_i)^{1/2} \leq \mathcal{E}^{(i)}(u_i, u_i)^{1/2} + \varepsilon; \tag{4.8}$$

and

$$\sup_{m\geq N} \overline{\mathcal{E}}^{(m)}_{j,\delta}(v_N, v_N)^{1/2} < \infty. \tag{4.9}$$

Since, for $k > N$

$$\begin{aligned}
\overline{\mathcal{E}}^{(k)}_{j,\delta}(v_k, v_k)^{1/2} &= \overline{\mathcal{E}}^{(k)}_{j,\delta}\Big(\frac{1}{k}\sum_{i=1}^k E_i u_i,\, \frac{1}{k}\sum_{i=1}^k E_i u_i\Big)^{1/2} \\
&= \overline{\mathcal{E}}^{(k)}_{j,\delta}\Big(\frac{1}{k}\sum_{i=1}^N E_i u_i + \frac{1}{k}\sum_{i=N+1}^k E_i u_i,\, \frac{1}{k}\sum_{i=1}^N E_i u_i + \frac{1}{k}\sum_{i=N+1}^k E_i u_i\Big)^{1/2} \\
&\leq \frac{N}{k}\overline{\mathcal{E}}^{(k)}_{j,\delta}(v_N, v_N)^{1/2} + \frac{1}{k}\sum_{i=N+1}^k \overline{\mathcal{E}}^{(k)}_{j,\delta}\big(E_i u_i,\, E_i u_i\big)^{1/2} \\
&\leq \frac{N}{k}\left(\sup_{m\geq N} \overline{\mathcal{E}}^{(m)}_{j,\delta}(v_N, v_N)^{1/2}\right) + \frac{1}{k}\sum_{i=N+1}^k \mathcal{E}^{(k)}\big(\pi_k E_i u_i,\, \pi_k E_i u_i\big)^{1/2}
\end{aligned}$$

by (4.8)–(4.9),

$$\begin{aligned}
\liminf_{k\to\infty} \overline{\mathcal{E}}^{(k)}_{j,\delta}(v_k, v_k)^{1/2} &\leq \liminf_{k\to\infty} \frac{1}{k}\left(\sum_{i=N+1}^k \mathcal{E}^{(i)}\big(u_i,\, u_i\big)^{1/2}\right) + \varepsilon \\
&\leq \lim_{k\to\infty} \mathcal{E}^{(k)}(u_k,\, u_k)^{1/2} + \varepsilon.
\end{aligned}$$

Now from (4.7), we have

$$\mathcal{E}_{j,\delta}(u,u)^{1/2} \leq \varliminf_{k\to\infty} \mathcal{E}^{(k)}(u_k,\, u_k)^{1/2} + 3\varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, we have

$$\mathcal{E}_{j,\delta}(u,u) \leq \varliminf_{k\to\infty} \mathcal{E}^{(k)}(u_k,\, u_k).$$

By first letting $j \to \infty$ and then $\delta \to 0$, one has $\lim_{k\to\infty} \mathcal{E}^{(k)}(u_k,\, u_k) \geq \mathcal{E}(u,u)$, which completes the proof of the theorem. □

4.2 Mosco convergence under conditions (A2), (A3)* and (A4)*

In this subsection, we present Mosco convergence under the assumptions **(A2)**, **(A3)*** and **(A4)***. We do *not* assume **(A1)** in this subsection. Recall that operators $\overline{\mathcal{L}}^{(k)}_{j,\delta}$ and $\mathcal{L}_{j,\delta}$ are defined in (2.21) and (2.22), respectively. Observe that

$$\overline{\mathcal{E}}^{(k)}_{j,\delta}(u,v) = -(u, \overline{\mathcal{L}}^{(k)}_{j,\delta} v)_{2,B_j} \qquad \text{and} \qquad \mathcal{E}_{j,\delta}(u,v) = -(u, \mathcal{L}_{j,\delta} v)_{2,B_j},$$

where $(u,v)_{2,B_j} = \int_{B_j} u(x)v(x)m(dx)$ and $\overline{\mathcal{E}}^{(k)}_{j,\delta}(u,v)$ and $\mathcal{E}_{j,\delta}(u,v)$ are defined in (2.16) and (2.17) respectively.

In this subsection, we assume conditions **(A2)**, **(A3)*** and **(A4)*** hold. Let

$$K_{j,\delta} := \sup_{x\in B_j} \int_{B_j} \mathbf{1}_{\{\rho(x,y)>\delta\}} j(x,dy),$$

which is finite due to **(A2)**. Also, let $\|\cdot\|_{2,B_j}$ be the L^2-norm on B_j. We then have the following basic estimates.

Lemma 4.6 *The following holds for any $\delta > 0$ and $j \in \mathbb{N}$.*
(i) $\mathcal{E}_{j,\delta}(u,u) \leq K_{j,\delta}\|u\|^2_{2,B_j}$ *for all* $u \in L^2(B_j;m)$. *In particular,* $\mathcal{E}_{j,\delta}(u,u) < \infty$ *for all* $u \in L^2(B_j;m)$.
(ii) $\|\mathcal{L}_{j,\delta}u\|^2_{2,B_j} \leq K_{j,\delta}\mathcal{E}_{j,\delta}(u,u)$ *for all* $u \in L^2(B_j;m)$.
(iii) $\lim_{k\to\infty} \|(\mathcal{L}_{j,\delta} - \overline{\mathcal{L}}^{(k)}_{j,\delta})f\|_{2,B_j} = 0$ *for all* $f \in \mathrm{Lip}_c(E)$.

Proof. (i) For $u \in L^2(B_j;m)$, we have

$$\begin{aligned}
\mathcal{E}_{j,\delta}(u,u) &= \frac{1}{2}\int_{B_j}\int_{B_j} (u(x)-u(y))^2 j(x,y)\mathbf{1}_{\{\rho(x,y)>\delta\}} dxdy \\
&\leq \|u\|^2_{2,B_j} \sup_{x\in B_j} \int_{B_j} j(x,y)\mathbf{1}_{\{\rho(x,y)>\delta\}} dy \leq K_{j,\delta}\|u\|^2_{2,B_j}.
\end{aligned}$$

(ii) As in (i), $\mathcal{E}_{j,\delta}(u,u) < \infty$ for $u \in L^2(B_j;m)$. So, using the Cauchy-Schwarz inequality, we have

$$\begin{aligned}
\|\mathcal{L}_{j,\delta}u\|^2_{2,B_j} &= \int_{B_j}\Big(\int_{B_j} (u(y)-u(x))\mathbf{1}_{\{\rho(x,y)>\delta\}} j(x,dy)\Big)^2 m(dx) \\
&\leq \int_{B_j}\Big(\int_{B_j} (u(x)-u(y))^2\mathbf{1}_{\{\rho(x,y)>\delta\}} j(x,dy) \cdot \int_{B_j} \mathbf{1}_{\{\rho(x,y)>\delta\}} j(x,dy)\Big) m(dx) \\
&\leq K_{j,\delta}\mathcal{E}_{j,\delta}(u,u).
\end{aligned}$$

(iii) Using the second half of **(A3)*** and **(A4)***(ii)(iii) (and (2.23)), we have

$$\|(\mathcal{L}_{j,\delta} - \overline{\mathcal{L}}_{j,\delta}^{(k)})f\|_{2,B_j}^2 = \|\mathcal{L}_{j,\delta} f\|_{2,B_j}^2 + \|\overline{\mathcal{L}}_{j,\delta}^{(k)} f\|_{2,B_j}^2 - 2\overline{\mathcal{E}}_{j,\delta}^{(k)}(\mathcal{L}_{j,\delta} f, f) \to 0.$$

□

We now prove the Mosco convergence that corresponds to Theorem 4.5. Recall that we do not assume **(A1)** in this subsection.

Theorem 4.7 *Under the assumptions* **(A2)**, **(A3)*** *and* **(A4)***, $(\mathcal{E}^{(k)}, \mathcal{F}^{(k)})$ *is Mosco convergent to* $(\mathcal{E}, \mathcal{F})$ *in the generalized sense of Definition 8.1.*

Proof. Since Lemma 4.4 works in this setting, as before, we only need to check condition (i) in Definition 8.1. Also, as in the proof of Theorem 4.5, we may assume $\{E_k u_k;\, k \geq 1\}$ is a bounded sequence on $L^2(E;m)$ that converges weakly to $u \in L^2(E;m)$, $\lim_{k\to\infty} \mathcal{E}^{(k)}(u_k, u_k) < \infty$, and (4.6) holds

Fix j large and $\delta > 0$ small then take positive $\varepsilon < \mathcal{E}_{j,\delta}(u,u)$. In the following, we simply write $(\cdot,\cdot)$, $\|\cdot\|_2$ for inner product and L^2-norm on B_j and use $L^2 = L^2(B_j;m)$. For $u \in L^2$ which is the weak limit of $E_k u_k$, take $f \in \mathrm{Lip}_c(E)$ so that $\mathcal{E}_{j,\delta}(u-f, u-f) + \|u-f\|_2^2 < \varepsilon$ (note that by Lemma 4.6(i), it is enough to take $\|u-f\|_2^2$ small). First, note that

$$\lim_{k\to\infty} (E_k u_k, (\mathcal{L}_{j,\delta} - \overline{\mathcal{L}}_{j,\delta}^{(k)})f) = 0, \tag{4.10}$$

where u_k, u and f are as above. Indeed, using Lemma 4.6(iii),

$$|(E_k u_k, (\mathcal{L}_{j,\delta} - \overline{\mathcal{L}}_{j,\delta}^{(k)})f)| \leq \|E_k u_k\|_2 \|(\mathcal{L}_{j,\delta} - \overline{\mathcal{L}}_{j,\delta}^{(k)})f\|_2 \leq \left(\sup_k \|E_k u_k\|_2\right) \|(\mathcal{L}_{j,\delta} - \overline{\mathcal{L}}_{j,\delta}^{(k)})f\|_2 \to 0.$$

Now

$$\begin{aligned} |\overline{\mathcal{E}}_{j,\delta}^{(k)}(E_k u_k, f) - \mathcal{E}_{j,\delta}(f,f)| &= |(f, \mathcal{L}_{j,\delta} f) - (E_k u_k, \overline{\mathcal{L}}_{j,\delta}^{(k)} f)| \\ &\leq |(E_k u_k, (\mathcal{L}_{j,\delta} - \overline{\mathcal{L}}_{j,\delta}^{(k)})f)| + |(E_k u_k - u, \mathcal{L}_{j,\delta} f)| + |(u-f, \mathcal{L}_{j,\delta} f)|. \end{aligned}$$

Using (4.10), the first term of the last line goes to zero and since $\{E_k u_k\}$ converges weakly to u, the second term goes to zero as $k \to \infty$ (note that $\mathcal{L}_{j,\delta} f \in L^2$ due to Lemma 4.6(i)(ii)). Further, there exists a $C = C(j,\delta,u) > 0$ such that

$$\begin{aligned} |(u-f, \mathcal{L}_{j,\delta} f)| &\leq \|u-f\|_2 \|\mathcal{L}_{j,\delta} f\|_2 \ \leq\ \|u-f\|_2 (\|\mathcal{L}_{j,\delta}(u-f)\|_2 + \|\mathcal{L}_{j,\delta} u\|_2) \\ &\leq \|u-f\|_2 (K_{j,\delta} \|u-f\|_2 + \|\mathcal{L}_{j,\delta} u\|_2) \ \leq\ C\varepsilon^{1/2}, \end{aligned}$$

where Lemma 4.6(i),(ii) are used in the third inequality.

Thus, using the Cauchy-Schwarz inequality, we have

$$\begin{aligned} \mathcal{E}_{j,\delta}(f,f) &\leq \limsup_{k\to\infty} |\overline{\mathcal{E}}_{j,\delta}^{(k)}(E_k u_k, f)| + C\varepsilon^{1/2} \\ &\leq \lim_{k\to\infty} \left(\overline{\mathcal{E}}_{j,\delta}^{(k)}(E_k u_k, E_k u_k)^{1/2} \mathcal{E}_{j,\delta}^{(k)}(f,f)^{1/2}\right) + C\varepsilon^{1/2} \\ &= \lim_{k\to\infty} \overline{\mathcal{E}}_{j,\delta}^{(k)}(E_k u_k, E_k u_k)^{1/2} \overline{\mathcal{E}}_{j,\delta}(f,f)^{1/2} + C\varepsilon^{1/2} \end{aligned}$$

where the last equality is due to **(A4)*** (iii). Since $\varepsilon < \mathcal{E}_{j,\delta}(u,u)$, by a rearrangement, we obtain

$$\begin{aligned}\mathcal{E}_{j,\delta}(u,u)^{1/2} &\leq \mathcal{E}_{j,\delta}(f,f)^{1/2} + \varepsilon^{1/2} \leq \lim_{k\to\infty} \overline{\mathcal{E}}^{(k)}_{j,\delta}(E_k u_k, E_k u_k)^{1/2} + C\frac{\varepsilon^{1/2}}{\mathcal{E}_{j,\delta}(f,f)^{1/2}} + \varepsilon^{1/2} \\ &\leq \lim_{k\to\infty} \overline{\mathcal{E}}^{(k)}(E_k u_k, E_k u_k)^{1/2} + C\frac{\varepsilon^{1/2}}{\mathcal{E}_{j,\delta}(u,u)^{1/2} - \varepsilon^{1/2}} + \varepsilon^{1/2}.\end{aligned}$$

Taking $\varepsilon \to 0$ and then $j \to \infty$ and $\delta \to 0$, we obtain the desired inequality. □

Remark 4.8 The second assumption in **(A3)*** is used only in the proof of Lemma 4.6(iii). Thus if we strengthen **(A4)*** (iii) further by assuming instead

$$\lim_{k\to\infty} \overline{\mathcal{E}}^{(k)}_{j,\delta}(f,f) = \mathcal{E}_{j,\delta}(f,f) \qquad \text{for every bounded measurable function } f \text{ on } B_j.$$

Then we can remove the second assumption in **(A3)***. Note that $\mathcal{L}_{j,\delta} f$ is bounded on B_j for each $f \in \mathrm{Lip}_c(E)$ by (2.11).

5 Proofs of Main Results

In this section, we give the proof of the main results of this paper.

Under conditions **(A2)** and **(A3)**, by [9, Propostion 2.2] and its proof, $(\mathcal{E}, \mathcal{F})$ is a regular Dirichlet form on $L^2(E;m)$. Let $X = \{X_t, t \geq 0, \mathbb{P}_x, x \in E\}$ be the symmetric Hunt process associated with $(\mathcal{E}, \mathcal{F})$ on E and recall that $X^{(k)}$ is symmetric continuous time random walk on V_k with conductance $\mathcal{C}^{(k)}(x,y) = m_k(x) j^{(k)}(x,y) m_k(y)$.

In the next theorem we show that $X^{(k)}$ converges to X in the sense of finite dimensional distributions under the assumption that X is conservative; that is, X has infinite lifetime $\mathbb{P}_x$-a.s. for $\mathcal{E}$-q.e. $x \in E$. We remark here, that, if

$$\sup_{x\in E} \int_E (\rho(x,y)^2 \wedge 1) j(x,dy) < \infty, \tag{5.1}$$

then we have by (2.2) and [24, Theorem 3.1] that the process X is conservative.

Throughout this section, $X^{(k)}$ and X are the symmetric Hunt processes associated with $(\mathcal{E}^{(k)}, \mathcal{F}^{(k)})$ and $(\mathcal{E}, \mathcal{F})$, respectively.

Theorem 5.1 *Assume that* **(A2)** *holds and that X is conservative. Assume further that either* (2.9) *of* **(A1)**, **(A3)**–**(A4)** *hold, or* **(A3)***–**(A4)*** *hold. Suppose φ is in $C_c^+(E)$. Then $\{X^{(k)}\}_{k\geq 1}$ with initial distribution $\mathbb{P}^{(k)}_\varphi$ converge to X with initial distribution $\mathbb{P}_\varphi$ in the finite dimensional sense.*

Proof. Without loss of generality, we assume $\int_E \varphi(x) m(dx) = 1$. Let $P_t f(x) := \mathbb{E}_x[f(X_t)]$ and $P^{(k)}_t g(x) := \mathbb{E}^{(k)}_x[g(X^{(k)}_t)]$ be the contraction semigroups on $L^2(E;m)$ and $L^2(V_k;m_k)$ respectively.

By Theorems 4.5, 4.7 and 8.3, $E_k P_t^{(k)} \pi_k$ converges to P_t strongly in $L^2(E;m)$. For any $l \geq 1$, $\{h_1, \cdots, h_l\} \subset L_b^2(E;m)$ and $0 \leq t_1 < t_2 < \cdots < t_l$, we have by Lemma 4.1 and the Markov property of $X^{(k)}$ and X that

$$\lim_{k\to\infty} \mathbb{E}^{(k)}_{\varphi\cdot m_k}\left[\pi_k h_1(X^{(k)}_{t_1})\cdots \pi_k h_l(X^{(k)}_{t_l})\right] = \mathbb{E}_{\varphi\cdot m}\left[h_1(X_{t_1})\cdots h_l(X_{t_l})\right]. \tag{5.2}$$

We fix $l \geq 1$. Since X is conservative, for any $\varepsilon > 0$, there is ball $B = B(x_0, r)$ so that $\mathbb{P}_{\varphi\cdot m}(X_{t_j} \in B) > 1-\varepsilon$ for every $j \in \{1, \ldots l\}$. By the strong L^2-convergence of $E_k P_{t_j}^{(k)} \pi_k \mathbf{1}_B$ to $P_{t_j}\mathbf{1}_B$ in $L^2(E;m)$, we have

$$\lim_{k\to\infty} \mathbb{P}^{(k)}_{\varphi\cdot m_k}\left(X^{(k)}_{t_j} \in B\right) > 1-\varepsilon \quad \text{ for every } j \in \{1, \ldots l\}. \tag{5.3}$$

For any $\{f_1, \cdots, f_l\} \subset C_b(E)$, since $E_k\pi_k f_j$ converges uniformly to f_j on $\overline{B}$, from (5.2) we have

$$\begin{aligned}
&\lim_{k\to\infty} \mathbb{E}^{(k)}_{\varphi\cdot m_k}\left[f_1(X^{(k)}_{t_1})\cdots f_l(X^{(k)}_{t_l}) : \cap_{j=1}^l \{X^{(k)}_{t_j} \in B\}\right] \\
=\;& \lim_{k\to\infty} \mathbb{E}^{(k)}_{\varphi\cdot m_k}\left[\pi_k(f_1\mathbf{1}_B)(X^{(k)}_{t_1})\cdots \pi_k(f_l\mathbf{1}_B)(X^{(k)}_{t_l})\right] \\
=\;& \mathbb{E}_{\varphi\cdot m}\left[(f_1\mathbf{1}_B)(X_{t_1})\cdots (f_j\mathbf{1}_B)(X_{t_j})\right] \\
=\;& \mathbb{E}_{\varphi\cdot m}\left[f_1(X_{t_1})\cdots f_j(X_{t_j}) : \cap_{j=1}^l \{X_{t_j} \in B\}\right].
\end{aligned} \tag{5.4}$$

We deduce the finite-dimensional convergence from (5.3) and (5.4). □

Definition 5.2 ([13]) *A collection of function $S \subset C_b(E)$ is said to strongly separate points if for every $x \in E$ and $\delta > 0$, there exists a finite set $\{h_1, \cdots, h_l\} \subset S$ such that*

$$\inf_{y:\rho(y,x)\geq\delta} \max_{1\leq i\leq l} |h_i(y) - h_i(x)| > 0.$$

We can easily check that $\mathrm{Lip}_c^+(E)$ strongly separates points in E.

Proof of Theorem 2.2. First, note that, by Proposition 3.4, for every $T > 0$ and any $m \geq 1$ and $\{g_1, \cdots, g_m\} \subset \mathrm{Lip}_c^+(E)$, $\left\{(g_1, \cdots, g_m)(X^{(k)})\right\}_{k\geq 1}$ restricted to $\{\zeta^{(k)} > T\}$ is tight in the Skorohod space $\mathbb{D}_{\mathbb{R}^m}[0,T]$ with the initial distribution $\mathbb{P}^{(k)}_\varphi$. Since X is conservative, by (5.3), for every $\varepsilon > 0$,

$$\lim_{k\to\infty} \mathbb{P}^{(k)}_{\varphi\cdot m_k}\left(\zeta^{(k)} > T\right) > 1-\varepsilon.$$

So it follows from [16, Theorem VI.3.21], $\left\{(g_1, \cdots, g_m)(X^{(k)})\right\}_{k\geq 1}$ is tight in the Skorohod space $\mathbb{D}_{\mathbb{R}^m}[0,T]$ with the initial distribution $\mathbb{P}^{(k)}_\varphi$. This together with Theorem 5.1 implies the weak convergence of $\left\{(g_1, \cdots, g_m)(X^{(k)})\right\}_{k\geq 1}$ with initial distribution $\mathbb{P}^{(k)}_\varphi$ to $(g_1, \cdots, g_m)(X)$ with initial distribution $\mathbb{P}_\varphi$. Since $\mathrm{Lip}_c^+(E)$ strongly separates points in E, we have the desired result by [13, Corollary 3.9.2]. □

We now turn our attention to the weak convergence of $\{X^{(k)}, k \geq 1\}$ under the convergence-in-measure (or pseudo-path) topology.

Proposition 5.3 *Assume that* **(A.2)**, **(A.3)** *and* **(A.4)**(i)(iii) *hold. Then for every $\varphi \in C_c^+(E)$, the law $\{\mathbb{P}_\varphi^{(k)}, k \geq 1\}$ is tight on $\mathbb{D}_{E_\partial}[0,1]$ equipped with the convergence-in-measure topology.*

Proof. Let D_1 and D_2 be two relatively compact open subsets in E with disjoint closure. By **(A.3)**, there is some $f \in \mathrm{Lip}_c(E) \subset \mathcal{F}$ so that $f = 1$ in an open neighborhood of $\overline{D}_2$ and $f = 0$ in an open neighborhood of $\overline{D}_1$. Then for k sufficiently large, $\pi_k f = 1$ on $V_k \cap D_2$ and $\pi_k f = 0$ on $V_k \cap D_1$. Let $N^{(k)}$ be the number of crossings by $X^{(k)}$ from D_1 into D_2. By [7, Theorem in page 69], if $g \in \mathcal{F}^{(k)}$ such that $g = 1$ on $D_2 \cap V_k$ and $g = 0$ on $D_1 \cap V_k$, then

$$\mathbb{E}^{(k)}_{\varphi \cdot m_k}[N^{(k)}] \leq 2\|\varphi\|_\infty \, \mathcal{E}^{(k)}(g,g). \tag{5.5}$$

It follows from Lemma 4.4 that

$$\sup_{k \geq 1} \mathbb{E}^{(k)}_{\varphi \cdot m_k}[N^{(k)}] < \infty.$$

Since the above holds for every pair of relatively compact open subsets in E with disjoint closure, we conclude by [25, Theorem 2] and a diagonal selection procedure that the law $\{\mathbb{P}_\varphi^{(k)}, k \geq 1\}$ is tight on $\mathbb{D}_{E_\partial}[0,1]$ equipped with the convergence-in-measure topology. □

Proof of Theorem 2.3. First, note that conditions **(A.3)*** and **(A.4)*** are stronger than conditions **(A.3)** and **(A.4)**(i)(iii). So, by Proposition 5.3, for any subsequence $\{n_k; k \geq 1\}$, there exists a sub-subsequence $\{n'_k; k \geq 1\}$ such that $\{(X^{(n'_k)}, \mathbb{P}_\varphi^{(n'_k)}); k \geq 1\}$ converges weakly on $\mathbb{D}_{E_\partial}[0,1]$ equipped with the convergence-in-measure topology to a law of say $\tilde{\mathbb{P}}$. Thus by [25, Theorem 5], we may assume without loss of generality that there is a subset $A \subset [0,1]$ of zero Lebesgue measure so that $\{(X^{(n'_k)}, \mathbb{P}_\varphi^{(n'_k)}); k \geq 1\}$ converges in finite dimension over the time interval $[0,1] \setminus A$ to that of $\tilde{\mathbb{P}}$. Let $P_t f(x) := \mathbb{E}_x[f(X_t)]$ and $P_t^{(k)} g(x) := \mathbb{E}_x^{(k)}[g(X_t^{(k)})]$. By Theorem 4.5 or Theorem 4.7, we know that $(\mathcal{E}^{(k)}, \mathcal{F}^{(k)})$ is Mosco convergent to $(\mathcal{E}, \mathcal{F})$. So by Theorem 8.3 (ii), $E_k P_t^{(k)} \pi_k f$ converges to $P_t f$ in $L^2(E;m)$. This implies by the Markov property that, for any $l \geq 1$, $\{h_1, \cdots, h_l\} \subset C_c(E)$ and $0 \leq t_1 < t_2 < \cdots < t_l$,

$$\lim_{k\to\infty} \mathbb{E}^{(k)}_{\varphi \cdot m_k}\left[\pi_k h_1(X^{(k)}_{t_1}) \cdots \pi_k h_l(X^{(k)}_{t_l})\right] = \mathbb{E}_{\varphi \cdot m}\left[h_1(X_{t_1}) \cdots h_l(X_{t_l})\right].$$

Thus the finite dimensional distribution under $\tilde{\mathbb{P}}$ over the time interval $[0,1] \setminus A$ is the same as that of $(X, \mathbb{P}_\varphi)$. Since both laws $\tilde{\mathbb{P}}$ and $\mathbb{P}_\varphi$ are carried on $\mathbb{D}_{E_\partial}[0,1]$, it follows that $\tilde{\mathbb{P}}$ has the same distribution as the law of $(X, \mathbb{P}_\varphi)$. Since this holds for any subsequence $\{n_k; k \geq 1\}$, we obtain the desired result. □

6 Discrete approximation

In this section, we give a general criteria for the approximation of symmetric pure-jump processes on metric measure spaces.

We introduce a condition on our approximating graphs.

(AG.4) There exists $n_0 \geq 1$ such that for every $j > n \geq n_0$ and $x \in V_{2^j}$, there is some $y \in V_{2^n}$ so that $U_{2^j}(x) \subset U_{2^n}(y)$.

Recall that conditions **(AG1)**–**(AG.3)** are given in Section 2. When $E = \mathbb{R}^d$, the following sequence of approximating graphs $\{(V_k, \Xi_k); k \geq 1\}$ and associated partitions $\{U_k(x), x \in V_k; k \geq 1\}$ satisfy conditions **(AG.1)**–**(AG.4)**:

$$V_k = k^{-1}\mathbb{Z}^d, \qquad (x, y) \in \Xi_k \quad \text{if and only if} \quad x, y \in k^{-1}\mathbb{Z}^d \text{ with } \|x - y\| = k^{-1}, \tag{6.1}$$

and

$$U_k(x) = \prod_{i=1}^d [x_i - (2k)^{-1}, x_i + (2k)^{-1}] \qquad \text{for } x = (x_1, \cdots, x_d) \in V_k. \tag{6.2}$$

Condition **(AG.4)** is needed only in this section. Recall that $B_r = B(x_0, r)$ for $r > 0$.

Theorem 6.1 *Let $j(x, y)$ be a non-negative measurable symmetric function on $E \times E$ such that*

$$j(x, y) \leq M_0 < \infty \qquad \text{for every } x, y \in E \text{ with } \rho(x, y) \geq 1$$

and for every compact set $K \subset E$,

$$\lim_{j \to \infty} \sup_{x \in K} j(x, (B_j)^c) = 0.$$

Assume that the Dirichlet form $(\mathcal{E}, \mathcal{F})$ determined by the jumping kernel $j(x, dy) := j(x, y)m(dy)$ satisfies the conditions **(A2)**–**(A3)** *and that the symmetric Hunt process X associated with the regular Dirichlet form $(\mathcal{E}, \mathcal{F})$ on $L^2(E; m)$ is conservative. Let $\{(V_{2^k}, \Xi_{2^k}); k \geq 1\}$ be approximating graphs of E and $\{U_{2^k}(x), x \in V_{2^k}; k \geq 1\}$ be the associated partitions satisfying* **(AG.1)**–**(AG.4)**. *Let*

$$j^{(2^k)}(x, y) := \mathbf{1}_{\{\rho_{2^k}(x,y) \geq 4C_3/C_1\}} \frac{1}{m_{2^k}(x) m_{2^k}(y)} \int_{U_{2^k}(x)} j(\xi, U_{2^k}(y)) m(d\xi) \qquad \text{for } x, y \in V_{2^k}, \tag{6.3}$$

where $m_{2^k}(x) := m(U_{2^k}(x))$ and C_1, C_3 are given in (2.3), (2.4). Then $(\mathcal{E}^{(2^k)}, \mathcal{F}^{(2^k)})$ as defined in (2.8) is a regular Dirichlet form on $L^2(V_{2^k}; m_{2^k})$. Let $X^{(2^k)}$ be its associated continuous time Markov chain on V_{2^k}. Then, for any positive function $\varphi \in C_c^+(E)$, $\{(X^{(2^k)}, \mathbb{P}_\varphi^{(2^k)}); k \geq 1\}$ converges weakly as $k \to \infty$ to $(X, \mathbb{P}_\varphi)$ on $\mathbb{D}_{E_\partial}[0, 1]$ equipped with the Skorohod topology.

Proof. For notational simplicity, in this proof we write k for 2^k. In view of Theorem 2.2, it is enough to show **(A1)** and **(A4)** hold. For $\rho_k(x, y) \geq 4C_3/C_1$ and $\xi \in U_k(x), \eta \in U_k(y)$, we have by (2.3)–(2.4) and the triangle inequality that $\rho(x, y) \geq C_1 \rho_k(x, y)/k \geq 4C_3/k$,

$$|\rho(\xi, \eta) - \rho(x, y)| \leq \rho(x, \xi) + \rho(\eta, y) \leq C_3/k + C_3/k = 2C_3/k \tag{6.4}$$

and so

$$\frac{C_1}{2} \frac{\rho_k(x, y)}{k} \leq \rho(x, y)/2 \leq \rho(\xi, \eta) \leq 3\rho(x, y)/2 \leq \frac{3C_2}{2} \frac{\rho_k(x, y)}{k}. \tag{6.5}$$

Take a compact set $K \subset E$ and $K_1 := \{x \in E : \rho(x, K) \le 1\}$. Then by (6.5)

$$
\begin{aligned}
& \sup_{k \in \mathbb{N}} \sup_{x \in K \cap V_k} \sum_{y \in V_k} j^{(k)}(x,y) \Big(\frac{\rho_k(x,y)}{k} \wedge 1 \Big)^2 m_k(y) \\
= & \sup_{k \in \mathbb{N}} \sup_{x \in K \cap V_k} \sum_{y \in V_k} \Big(\frac{\rho_k(x,y)}{k} \wedge 1 \Big)^2 \mathbf{1}_{\{\rho_k(x,y) \ge 4C_3/C_1\}} \frac{1}{m_k(x)} \int_{U_k(x)} j(\xi, U_k(y)) m(d\xi) \\
= & \sup_{k \in \mathbb{N}} \sup_{x \in K \cap V_k} \sum_{y \in V_k} \frac{1}{m_k(x)} \int_{U_k(x)} \int_{U_k(y)} \Big(\frac{\rho_k(x,y)}{k} \wedge 1 \Big)^2 \mathbf{1}_{\{\rho_k(x,y) \ge 4C_3/C_1\}} j(\xi, d\eta) m(d\xi) \\
\le & \ c \sup_{k \in \mathbb{N}} \sup_{x \in K \cap V_k} \sum_{y \in V_k} \frac{1}{m_k(x)} \int_{U_k(x)} \Big(\sup_{\xi \in U_k(x)} \int_{U_k(y)} (\rho(\xi,\eta)^2 \wedge 1) j(\xi, d\eta) \Big) m(d\xi) \\
\le & \ c \sup_{k \in \mathbb{N}} \sup_{\xi \in K_1} \sum_{y \in V_k} \int_{U_k(y)} (\rho(\xi,\eta)^2 \wedge 1) j(\xi, d\eta) \\
\le & \ c \sup_{\xi \in K_1} \int_E (\rho(\xi,\eta)^2 \wedge 1) j(\xi, d\eta) \le C_K
\end{aligned}
$$

by **(A2)** (iii). This proves (2.9) of **(A1)**.

By (6.4), for $k \ge 2C_3$ and $x, y \in V_k$ with $\rho_k(x,y) \ge 2$,

$$
\rho(\xi,\eta) \ge \rho(x,y) - 2C_3/k \ge 1 \qquad \text{for } \xi \in U_k(x) \text{ and } \eta \in U_k(y).
$$

So for each $k \ge 2C_3$, $j \ge 1$ and $x \in \overline{B}_j \cap V_k$, $y \in (B_{j+2})^c \cap V_k$,

$$
j^{(k)}(x,y) \le \frac{1}{m_k(x) m_k(y)} \int_{U_k(x) \times U_k(y)} j(\xi,\eta) m(d\xi) m(d\eta) \le M,
$$

which establishes (2.10) of **(A1)**.

By the definition of $j^{(k)}(\cdot,\cdot)$, (2.18) clearly holds. For any compact set $K \subset E$ with $K_1 := \{x \in E : \rho(x, K) \le 1\}$, we have

$$
\lim_{j \to \infty} \sup_{k \ge 1} \sup_{x \in K} \int_{(B_j)^c} j^{(k)}(x,y) m(dy) \quad \le \quad \lim_{j \to \infty} \sup_{x \in K_1} \int_{(B_j)^c} j(\xi, y) m(dy) = 0,
$$

so (2.19) holds.

On the other hand by **(A1)**, for any $f \in L^2_b(E,m)$ with $\|f\|_\infty \le M_1$, $j \ge 1$ and $\delta > 0$,

$$
\begin{aligned}
& \Big| \overline{\mathcal{E}}^{(k)}_{j,\delta}(f,f)^{1/2} - \overline{\mathcal{E}}^{(k)}_{j,\delta}(E_k \pi_k f, E_k \pi_k f)^{1/2} \Big| \\
\le & \ \overline{\mathcal{E}}^{(k)}_{j,\delta}(E_k \pi_k f - f, E_k \pi_k f - f)^{1/2} \\
\le & \ \Big(2 \int_{B_j} (f(x) - E_k \pi_k f(x))^2 \Big(\int_{B_j} j^{(k)}(x,y) \mathbf{1}_{\{\rho(x,y) > \delta\}} m(dy) \Big) m(dy) \Big)^{1/2} \\
\le & \ c(j,\delta) \, \|f - E_k \pi_k f\|_2,
\end{aligned}
\tag{6.6}
$$

which goes to 0 as $k \to \infty$ by Lemma 4.1(iv). Note that for large k and small δ,

$$
\begin{aligned}
&\overline{\mathcal{E}}^{(k)}_{j,\delta}(E_k\pi_k f, E_k\pi_k f) = \int_{B_j\times B_j} (E_k\pi_k f(x) - E_k\pi_k f(y))^2 j^{(k)}(x,y)\mathbf{1}_{\{\rho(x,y)>\delta\}} m(dx)m(dy) \\
= \;& \frac{1}{2}\sum_{(z,w)\in V_k\times V_k} (\pi_k f(z) - \pi_k f(w))^2 \,\frac{1}{m_k(z)m_k(w)}\int_{U_k(z)} j(\xi, U_k(w))m(d\xi) \times \\
&\times \int_{(B_j\times B_j)\cap(U_k(z)\times U_k(w))} \mathbf{1}_{\{\rho(x,y)>\delta\}} m(dx)m(dy) \qquad (6.7)
\end{aligned}
$$

and

$$
\begin{aligned}
&\mathcal{E}_{j,\delta}(E_k\pi_k f, E_k\pi_k f) = \frac{1}{2}\int_{B_j\times B_j} (E_k\pi_k f(x) - E_k\pi_k f(y))^2 j(x,y)\mathbf{1}_{\{\rho(x,y)>\delta\}} m(dx)m(dy) \\
= \;& \frac{1}{2}\sum_{(z,w)\in V_k\times V_k} (\pi_k f(z) - \pi_k f(w))^2 \int_{(B_j\times B_j)\cap(U_k(z)\times U_k(w))} j(x,y)\mathbf{1}_{\{\rho(x,y)>\delta\}} m(dx)m(dy). \qquad (6.8)
\end{aligned}
$$

Since, except the case $\rho(x,y)$ is small and y is near the boundary of B_j, the summands in (6.7) and (6.8) are the same, it is easy to see that there exists $k_0 = k_0(\delta) > 0$ and $c > 0$ such that for $k \ge k_0$,

$$
\begin{aligned}
&\left|\overline{\mathcal{E}}^{(k)}_{j,\delta}(E_k\pi_k f, E_k\pi_k f) - \mathcal{E}_{j,\delta}(E_k\pi_k f, E_k\pi_k f)\right| \\
\le \;& 2\int_{B_j\times\{j-c\frac{1}{k}<\rho(y,x_0)<j+c\frac{1}{k}\}} (E_k\pi_k f(x) - E_k\pi_k f(y))^2 j(x,y)\mathbf{1}_{\{\rho(x,y)>\delta-c\frac{1}{k}\}} m(dx)m(dy) \\
&+ \int_{B_{j+1}\times B_{j+1}} (E_k\pi_k f(x) - E_k\pi_k f(y))^2 j(x,y)\mathbf{1}_{\{\delta+c\frac{1}{k}>\rho(x,y)>\delta-c\frac{1}{k}\}} m(dx)m(dy) \\
\le \;& 2(2M_1)^2\int_{B_j\times\{j-c\frac{1}{k}<\rho(y,x_0)<j+c\frac{1}{k}\}} j(x,y)\mathbf{1}_{\{\rho(x,y)>\delta-c\frac{1}{k}\}} m(dx)m(dy) \\
&+(2M_1)^2\int_{B_{j+1}\times B_{j+1}} j(x,y)\mathbf{1}_{\{\delta+c\frac{1}{k}>\rho(x,y)>\delta-c\frac{1}{k}\}} m(dx)m(dy),
\end{aligned}
$$

which goes to zero as k goes to ∞. Therefore

$$
\begin{aligned}
&\lim_{k\to\infty}\left|\overline{\mathcal{E}}^{(k)}_{j,\delta}(E_k\pi_k f, E_k\pi_k f) - \mathcal{E}_{j,\delta}(f,f)\right| \\
\le \;& c\lim_{k\to\infty}\left|\overline{\mathcal{E}}^{(k)}_{j,\delta}(E_k\pi_k f, E_k\pi_k f)^{1/2} - \mathcal{E}_{j,\delta}(f,f)^{1/2}\right| \\
\le \;& c\lim_{k\to\infty}\left|\mathcal{E}_{j,\delta}(E_k\pi_k f, E_k\pi_k f)^{1/2} - \mathcal{E}_{j,\delta}(f,f)^{1/2}\right| \\
\le \;& c\lim_{k\to\infty}\mathcal{E}_{j,\delta}(E_k\pi_k f - f, E_k\pi_k f - f)^{1/2} \\
\le \;& c\lim_{k\to\infty}\left(\int_{B_j\times B_j}\left((f - E_k\pi_k f)(x) - (f - E_k\pi_k)f(y)\right)^2 j(x,y)\mathbf{1}_{\{\rho(x,y)>\delta\}} m(dx)m(dy)\right)^{1/2} \\
\le \;& c\lim_{k\to\infty}\left(\int_{B_j}(f(x) - E_k\pi_k f(x))^2\left(\int_{B_j} j(x,y)\mathbf{1}_{\{\rho(x,y)>\delta\}} m(dy\right) m(dx)\right)^{1/2} \\
\le \;& c\lim_{k\to\infty}\|f - E_k\pi_k f\|_2 = 0 \qquad (6.9)
\end{aligned}
$$

where $c = c(M_1, j, \delta, f) > 0$. This combined with (6.6) shows that $\lim_{k\to\infty} \overline{\mathcal{E}}^{(k)}_{j,\delta}(f,f) = \mathcal{E}_{j,\delta}(f,f)$ for any $f \in L^2_b(E;m)$. The monotonicity property of **(A4)**(ii) (with 2^k instead of k) is an immediate consequence of **(AG.4)** and (6.3). So we have established **(A4)**. □

Remark 6.2 For any $f \in L^2_b(E;m)$ with $\|f\|_\infty \le M_1$, $j \ge 1$ and $\delta > 0$, computing similarly to (6.9), we have

$$\left| \|\overline{\mathcal{L}}^{(k)}_{j,\delta} f\|_{2,B_j} - \|\overline{\mathcal{L}}^{(k)}_{j,\delta} E_k\pi_k f\|_{2,B_j} \right| \le \left| \|\overline{\mathcal{L}}^{(k)}_{j,\delta} (f - E_k\pi_k f)\|_{2,B_j} \right| \le c(j,\delta)\, \|f - E_k\pi_k f\|_2,$$

which goes to 0 as $k \to \infty$ by Lemma (4.1) (iv). Moreover, by Lemma 4.6 (i) (ii),

$$\begin{aligned} \lim_{k\to\infty} \left| \|\mathcal{L}_{j,\delta} f\|_{2,B_j} - \|\mathcal{L}_{j,\delta} E_k\pi_k f\|_{2,B_j} \right| &\le \left| \|\mathcal{L}_{j,\delta}(f - E_k\pi_k f)\|_{2,B_j} \right| \\ &\le \lim_{k\to\infty} c(j,\delta)\, \|f - E_k\pi_k f\|_2 = 0. \end{aligned}$$

Thus, to show **(A4)*** (ii), it is enough to show that

$$\limsup_{k\to\infty} \left| \|\overline{\mathcal{L}}^{(k)}_{j,\delta} E_k\pi_k f\|^2_{2,B_j} - \|\mathcal{L}_{j,\delta} E_k\pi_k f\|^2_{2,B_j} \right| = 0. \tag{6.10}$$

Remark 6.3 Note that Theorem 6.1 is applicable to the example discussed after (1.1), namely when $E = \mathbb{R}^d$, $m(dx) = dx$ and $J(dx,dy) = j(x,y)dxdy$ where j is a symmetric measurable function on $\mathbb{R}^d \times \mathbb{R}^d$ so that

$$c_1|x-y|^{-d-\alpha_1} \le j(x,y) \le c_2|x-y|^{-d-\alpha_2} \qquad \text{when } |x-y| \le 1 \tag{6.11}$$

for some $0 < \alpha_1 \le \alpha_2 < 2$ and that j is bounded on $\{|x-y| > 1\}$ with

$$\sup_{x\in\mathbb{R}^d} \int_{B(x,1)^c} |x-y|^\gamma j(x,y) m(dy) < \infty \qquad \text{for some } \gamma > 0. \tag{6.12}$$

In this case, one may take $\{(V_k, \Xi_k); k \ge 1\}$ and $\{U_k(x), x \in V_k; k \ge 1\}$ as in (6.1) and (6.2).

We may also apply Theorem 6.1 in more general metric measure spaces (for instance in a subclass of spaces discussed in [9]). Here we give one simple example which is the 2-dimensional Sierpinski gasket. Let $a_0 = (0,0)$, $a_1 = (1,0)$ and $a_2 = (\frac{1}{2}, \frac{\sqrt{3}}{2})$, and let $F_0 = \{a_0, a_1, a_2\}$. Define inductively

$$F_{n+1} = F_n \cup (2^n a_1 + F_n) \cup (2^n a_2 + F_n), \qquad n = 0, 1, 2, \cdots,$$

where we write $a + A = \{a + x : x \in A\}$. Let $V_0 = \cup_{n=0}^\infty F_n$ and define $V_{2^k} = 2^{-k} V_0$. Then $E := \overline{\cup_{k=0}^\infty V_{2^k}}$ is the 2-dimensional Sierpinski gasket having Hausdorff dimension $d = \log 3/\log 2$. Let $\rho(\cdot,\cdot)$ be the geodesic distance function on E and m the d-dimensional Hausdorff measure on E. Then (E, ρ, m) satisfies **(MMS.1)**–**(MMS.3)**. Define Ξ_{2^k} by $(x,y) \in \Xi_{2^k}$ if and only if $x, y \in V_{2^k}$ with $\rho(x,y) = 2^{-k}$. For each $x \in V_{2^k}$, set $U_{2^k}(x) = \{y \in E : \rho(x,y) \le 2^{-k}\}$. Then, $\{(V_{2^k}, \Xi_{2^k}); k \ge 1\}$ and $\{U_{2^k}(x), x \in V_{2^k}; k \ge 1\}$ satisfies **(AG.1)**–**(AG.4)**. Now consider the Dirichlet form (2.12) with $J(dx,dy) = j(x,y)m(dx)m(dy)$, where j a symmetric Borel measurable function that satisfies the conditions (6.11)–(6.12) with $|x-y|$ and $\mathbb{R}^d$ being replaced by $\rho(x,y)$ and E, respectively. (We remark here that the geodesic distance ρ on E is in fact comparable to the Euclidean distance on E.) Then $(\mathcal{E}, \mathcal{F})$ satisfies the conditions **(A2)**–**(A3)** and it is conservative since (5.1) is satisfied. Thus, the conclusion of Theorem 6.1 holds for this example as well.

7 Application to random walk in random conductance

In this section, we present application of Theorem 4.7 to the scaling limit of some random walk in random conductance.

Throughout this subsection, $E = \mathbb{R}^d$ and m is d-dimensional Lebesgue measure. Let $V_k = k^{-1}\mathbb{Z}^d$ and $m_k(x) = k^{-d}$ for every $x \in V_k$. Let $j(x,y)$ be a symmetric non-negative continuous function of x and y on $\mathbb{R}^d \times \mathbb{R}^d \setminus \widehat{d}$ such that there exist $\alpha, \beta \in (0,2)$, $\alpha > \beta$ and positive constants κ_1, κ_2 such that

$$\kappa_1 |y-x|^{-d-\beta} \le j(x,y) \le \kappa_2 |y-x|^{-d-\alpha} \quad \text{for } |y-x|<1 \tag{7.1}$$

and

$$\sup_{\substack{(x,y)\in\mathbb{R}^d\times\mathbb{R}^d \\ |y-x|\ge 1}} j(x,y) \le \kappa_0 < \infty \quad \text{and} \quad \sup_{x\in\mathbb{R}^d} \int_{B(x,1)^c} j(x,y) m(dy) < \infty. \tag{7.2}$$

Let $(\mathcal{E},\mathcal{F})$ be the Dirichlet form defined by (2.12) with $J(dx,dy) = j(x,y)m(dx)m(dy)$, where the jumping kernel $j(x,y)$ is given by (7.1)–(7.2). Finally we assume **(A3)** holds, i.e., $\mathrm{Lip}_c(E)$ is dense in $(\mathcal{F}, \mathcal{E}(\cdot,\cdot) + \|\cdot\|_2^2)$. Then, by [9, Propostion 2.2] and its proof, the Dirichlet form $(\mathcal{E},\mathcal{F})$ is regular on $\mathbb{R}^d$ and so it associates a Hunt process X starting from quasi-everywhere on $L^2(\mathbb{R}^d;m)$. Moreover X is conservative since (5.1) is satisfied.

Proposition 7.1 (i) *Suppose $d \ge 2$. Let $\{\xi_{x,y}\}_{x,y\in\mathbb{Z}^d, x\neq y}$ a sequence of i.i.d. non-negative real-valued random variables defined on a probability space $(\Omega, \mathcal{A}, \mathbf{P})$ with $\mathbf{E}[\xi_{x,y}] = 1$ and Var $(\xi_{x,y}) < \infty$. Set*

$$j^{(k)}(x,y) := \xi_{kx,ky} j(x,y) \qquad \text{for } x,y \in V_k. \tag{7.3}$$

Let $(\mathcal{E}^{(k)}, \mathcal{F}^{(k)})$ be the Dirichlet form on $L^2(V_k;m_k)$ defined by (2.8) with $j^{(k)}(x,y)$ in (7.3) and $X^{(k)}$ be the continuous-time Markov chain associated with the regular Dirichlet form $(\mathcal{E}^{(k)}, \mathcal{F}^{(k)}$ on $L^2(V_k;m_k)$. Let X be the Hunt process corresponding to $(\mathcal{E},\mathcal{F})$ which is defined by (2.12) with $J(dx,dy) = j(x,y)m(dx)m(dy)$ where $j(x,y)$ defined in (7.1)–(7.2). Let $\{T_t^{(k)}, t \ge 0\}$ and $\{T_t, t\ge 0\}$ be the transition semigroups of $X^{(k)}$ and X, respectively. Then for each $t \ge 1$, as $k \to \infty$, $E_k T_t^{(k)} \pi_k \to T_t$ strongly in $L^2(\mathbb{R}^d;m)$ $\mathbf{P}$-a.s. and the convergence is uniform in any finite interval of $t \ge 0$. Moreover, for every $\varphi \in C_c^+(E)$, $(X^{(k)}, \mathbb{P}_\varphi^{(k)})$ converges weakly to $(X, \mathbb{P}_\varphi)$ on $\mathbb{D}_{E_\partial}[0,1]$ equipped with convergence-in-measure topology $\mathbf{P}$-a.s..

(ii) *Assume further that $0 \le \xi_{x,y} \le C$ $\mathbf{P}$-a.s. for some deterministic constant $C > 0$. Then for any $\varphi \in C_c^+(E)$, $\{(X^{(k)}, \mathbb{P}_\varphi^{(k)});\ k \ge 1\}$ converges weakly to $(X, \mathbb{P}_\varphi)$ on $\mathbb{D}_{E_\partial}[0,1]$ equipped with the Skorohod topology $\mathbf{P}$-a.s..*

Proof. (i) Note first that since, by (7.2)

$$\mathbf{E}\left[\sum_{y\in V_k} j^{(k)}(x,y) m_k(y)\right] \le \kappa_2 \sum_{y\in V_k, |x-y|<1} k^{-d}|x-y|^{-d-\alpha} + \sum_{y\in V_k, |x-y|\ge 1} k^{-d} j(x,y) < \infty,$$

we have $\sum_{y\in V_k} j^{(k)}(x,y)m_k(y) < \infty$ $\mathbf{P}$-a.s., so (2.7) holds. Thus, by Theorem 3.2, $(\mathcal{E}^{(k)}, \mathcal{F}^{(k)})$ is a regular Dirichlet form. In order to prove the first assertion of (i), by Theorem 4.7, Theorem 5.1 and

Theorem 8.3, it is enough to prove **(A2)**, **(A3)*** and **(A4)*** **P**-a.s.. Recall that we assume **(A3)**. Moreover, the second half of **(A3)*** is true by the continuity of $j(x,y)$. Furthermore, by symmetry of $j(x,y)$ and (7.1)–(7.2), one can easily see that **(A2)** is true. So, we will prove **(A4)*** below.

We first show (2.18). Let $\eta \le 1$. Note that, by (7.1)

$$\int\int_{\{(x,y)\in K\times K:|x-y|\le\eta\}} |x-y|^2 j^{(k)}(x,y)m(dx)m(dy) \le \kappa_2 k^{-2d} \sum_{\substack{x,y\in V_k\cap K\\ |x-y|\le\eta}} \frac{|x-y|^2\xi_{kx,ky}}{|x-y|^{d+\alpha}} =: \kappa_2 k^{-2d} H_k.$$

Since $|x-y|\ge k^{-1}$ when $x\neq y$, setting $2-\alpha=\varepsilon$,

$$\mathrm{Var}\ (H_k) = \sum_{\substack{x,y\in V_k\cap K\\ |x-y|\le\eta}} |x-y|^{2(2-d-\alpha)}\mathrm{Var}\ (\xi_{kx,ky}) \le c_1 k^{3d}\cdot k^{-2d} \sum_{\substack{x,y\in V_k\cap K\\ |x-y|\le\eta}} |x-y|^{-d+2\varepsilon} \le c_2 k^{3d} m(K)\eta^{\varepsilon}.$$

So,

$$\mathbf{P}\left(k^{-2d}\left|H_k-\mathbf{E}[H_k]\right|\ge \eta^{\varepsilon/2}\right) \le \kappa_2^2\frac{\mathrm{Var}\ (H_k)}{k^{4d}\eta^{\varepsilon}} \le \frac{c_3}{k^d},$$

and using the Borel-Cantelli Lemma, we have $\limsup_{k\to\infty} k^{-2d}|H_k-\mathbf{E}[H_k]|\le\eta^{\varepsilon/2}$ **P**-a.s., so

$$\lim_{\eta\to 0}\limsup_{k\to\infty} k^{-2d}|H_k-\mathbf{E}[H_k]| = 0.$$

On the other hand, by (7.1)

$$\begin{aligned}
\limsup_{k\to\infty} k^{-2d}\mathbf{E}[H_k] &\le \kappa_2 \limsup_{k\to\infty} k^{-2d} \sum_{\substack{x,y\in V_k\cap K\\ |x-y|\le\eta}} |x-y|^{(2-d-\alpha)}\mathbf{E}[\xi_{kx,ky}] \\
&= \kappa_2 \limsup_{k\to\infty} k^{-2d} \sum_{\substack{x,y\in V_k\cap K\\ |x-y|\le\eta}} |x-y|^{2-d-\alpha} \le c\, m(K)\, \eta^{(2-\alpha)/2},
\end{aligned}$$

which vanishes when $\eta\to 0$, so we obtain (2.18) **P**-a.s..

We next show (2.19). Note that

$$\int_K\int_{(B_j)^c} j^{(k)}(x,y)m(dx)m(dy) = k^{-2d}\sum_{y\in V_k\cap K}\ \sum_{x\in V_k\cap(B_j)^c} \xi_{kx,ky} j(x,y) =: k^{-2d}H'_k.$$

Then, for $j\ge j_0$ where $K\subset B_{j_0-1}$, by (7.2) we have

$$\begin{aligned}
k^{-2d}\mathrm{Var}\ (H'_k) &= k^{-2d}\sum_{\substack{x\in V_k\cap(B_j)^c\\ y\in V_k\cap K}} \mathrm{Var}\ (\xi_{kx,ky}) j(x,y)^2 \\
&\le ck^{-2d}\sum_{\substack{x\in V_k\cap(B_j)^c\\ y\in V_k\cap K}} j(x,y) \le c\, k^{-2d}\sum_{\substack{x\in V_k:|x-y|>j-j_0\\ y\in V_k\cap K}} j(x,y) =: c\, a_j^k.
\end{aligned}$$

Thus,

$$\mathbf{P}\left(k^{-2d}(a_j^k)^{-1/2}\left|H'_k-\mathbf{E}[H'_k]\right|\ge 1\right) \le \frac{\mathrm{Var}\ (H'_k)}{k^{4d}a_j^k} \le \frac{c}{k^{2d}},$$

and using the Borel-Cantelli Lemma, we have $\limsup_{k\to\infty} k^{-2d}(a_j^k)^{-1/2}|H'_k - \mathbf{E}[H'_k]| \le 1$ $\mathbf{P}$-a.s.. Since a_j^k converges to

$$a_j := \int_K \int_{\{|x-y|>j-j_0\}} j(x,y)m(dx)m(dy) \in (0,\infty)$$

by continuity of $j(x,y)$ and (7.2), we have

$$\lim_{j\to\infty}\limsup_{k\to\infty} k^{-2d}|H'_k - \mathbf{E}[H'_k]| = \Big(\limsup_{k\to\infty} k^{-2d}(a_j^k)^{-1/2}|H'_k - \mathbf{E}[H'_k]|\Big)\lim_{j\to\infty}\sqrt{a_j} \le \lim_{j\to\infty}\sqrt{a_j} = 0.$$

In the last equality above, we have used (7.2). On the other hand, by similar computation we have

$$\lim_{j\to\infty}\limsup_{k\to\infty} k^{-2d}\mathbf{E}[H'_k] \le c\lim_{j\to\infty} a_j = 0$$

We have proved (2.19).

For the remainder part of the proof, we fix $\delta, j > 0$. We now show **(A4)*** (iii).

Let h be a bounded and continuous function in $B_j \times B_j$. By the continuity and boundedness of $h(x,y)$ and $j(x,y)$ on $B_j \times B_j \setminus \widehat{d}$, we have

$$\lim_{k\to\infty} k^{-2d} \sum_{\substack{x,y\in V_k\cap B_j\\ |x-y|>\delta}} h(x,y)j(x,y) = \int_{B_j\times B_j} h(x,y)\,\mathbf{1}_{\{|x-y|>\delta\}}\,j(x,y)m(dx)m(dy), \tag{7.4}$$

so it is enough to show

$$\lim_{k\to\infty} k^{-2d} \sum_{\substack{x,y\in V_k\cap B_j\\ |x-y|>\delta}} h(x,y)(\xi_{kx,ky}-1)j(x,y) = 0 \quad \mathbf{P}\text{-a.s..} \tag{7.5}$$

Using (7.1)–(7.2), we have,

$$\begin{aligned}
&\mathbf{P}\Big(k^{-2d}\Big|\sum_{\substack{x,y\in V_k\cap B_j\\ |x-y|>\delta}} h(x,y)(\xi_{kx,ky}-1)j(x,y)\Big| > \varepsilon^{1/2}\Big)\\
\le\ & c_1\frac{1}{k^{4d}\varepsilon}\mathrm{Var}\,\Big(\sum_{\substack{x,y\in V_k\cap B_j\\ |x-y|>\delta}} h(x,y)(\xi_{kx,ky}-1)j(x,y)\Big)\\
\le\ & c_2\frac{1}{k^{2d}\varepsilon}\mathrm{Var}\ (\xi_{kx,ky})\Big(\frac{1}{k^{2d}}\sum_{\substack{x,y\in V_k\cap B_j\\ |x-y|>\delta}} h(x,y)^2|x-y|^{-2d-2\alpha}\Big) \le \frac{c_{\delta,j}}{k^{2d}\varepsilon},
\end{aligned}$$

so using the Borel-Cantelli Lemma, computing similarly as before, we obtain (7.5).

Lastly, we show **(A4)*** (ii). Fix $f \in \mathrm{Lip}_c(E)$ and let $\|\cdot\|_2$ be L^2-norm on B_j. Note that

$$\begin{aligned}
\overline{\mathcal{L}}^{(k)}_{j,\delta} f(x) &= \frac{1}{k^d}\sum_{\substack{y\in V_k\cap B_j\\ |x-y|>\delta}} (f(y)-f(x))j(x,y) + \frac{1}{k^d}\sum_{\substack{y\in V_k\cap B_j\\ |x-y|>\delta}} (\xi_{kx,ky}-1)(f(y)-f(x))j(x,y)\\
&=: I_1^{(k)}(x) + I_2^{(k)}(x).
\end{aligned}$$

One can easily see that $\|I_1^{(k)} - \mathcal{L}_{j,\delta} f\|_2 \to 0$ as $k \to \infty$. Indeed, by the continuity and boundedness of j and f, it is clear that $\lim_{k\to\infty} I_1^{(k)}(x) = \mathcal{L}_{j,\delta} f(x)$ for all x and $|I_1^{(k)}(x)| \le C$ for large C. Thus the bounded convergence theorem can be applied. So all we need is to show $\|I_2^{(k)}\|_2 \to 0$ $\mathbf{P}$-a.s. as $k \to \infty$. Since

$$\begin{aligned}
\mathbf{E}[\|I_2^{(k)}\|_2^2] &= k^{-2d}\mathbf{E}\Big[\int_{B_j}\Big(\sum_{\substack{y\in V_k\cap B_j\\ |x-y|>\delta}} (\xi_{kx,ky}-1)(f(y)-f(x))j(x,y)\Big)^2 m(dx)\Big] \\
&= k^{-2d}\int_{B_j}\sum_{\substack{y\in V_k\cap B_j\\ |x-y|>\delta}} (f(x)-f(y))^2 \mathrm{Var}\ (\xi_{kx,ky}) j(x,y)^2 m(dx) \\
&= ck^{-d}\int_{B_j}\sum_{\substack{y\in V_k\cap B_j\\ |x-y|>\delta}} (f(x)-f(y))^2 j(x,y)^2 m_k(y) m(dx) \ \le\ c_{f,\delta,j} k^{-d},
\end{aligned}$$

computing similarly as before,

$$\mathbf{P}(\|I_2^{(k)}\|_2^2 > \varepsilon) \le \varepsilon^{-1}\mathbf{E}[\|I_2^{(k)}\|_2^2] \le \frac{c_{f,\delta,j}}{\varepsilon k^d}. \tag{7.6}$$

So using the Borel-Cantelli Lemma, $\|I_2^{(k)}\|_2 \to 0$ $\mathbf{P}$-a.s. for $d \ge 2$. The weak convergence follows from Theorem 2.3.

(ii) Using (7.1)–(7.2), it is easy to show that **(A1)** holds $\mathbf{P}$-a.s., and X is conservative. Thus, by Theorems 2.2 and 4.7, we obtain the desired result. □

More concretely, we have the following example.

Example 7.2 Let $\phi : (0,\infty) \to (0,\infty)$ be a strictly increasing, continuous function such that $\phi(0) = 0$ and for all $0 < r < R < \infty$,

$$c_1\left(\frac{R}{r}\right)^{\alpha_1} \le \frac{\phi(R)}{\phi(r)} \ \le\ c_2\left(\frac{R}{r}\right)^{\alpha_2} \qquad \text{and} \qquad \int_0^r \frac{s}{\phi(s)}\,ds \le c_3\,\frac{r^2}{\phi(r)}.$$

Here $0 < \alpha_1 \le \alpha_2 \le 2$. Assume that there exists $\psi : (0,\infty) \to (0,\infty)$ a strictly increasing, continuous function with $\psi(0) = 0$ such that

$$\lim_{k\to\infty}\frac{\phi(k)}{\phi(kr)} = \frac{1}{\psi(r)} \qquad \text{for every } r > 0. \tag{7.7}$$

(i) Let $\{\xi_{xy}\}_{x,y\in\mathbb{Z}^d, x\neq y}$ be i.i.d. on $(\Omega,\mathcal{F},\mathbf{P})$ such that $0 \le \xi_{xy}$, $\mathbf{E}[\xi_{xy}] = 1$ and Var $(\xi_{xy}) < \infty$. Let $j_\xi(x,y) = \frac{\xi_{xy}}{|x-y|^d\phi(|x-y|)}$ for $x, y \in \mathbb{Z}^d$, and define instead of (7.3),

$$j^{(k)}(x,y) := k^d\phi(k) j_\xi(kx,ky) = \frac{\xi_{kx,ky}\phi(k)}{|x-y|^d\phi(k|x-y|)} \qquad \text{for } x,y\in V_k.$$

Then the claim of Proposition 7.1(i) holds, where $X_t^{(k)} := k^{-1}X^{(1)}_{\phi(k)t}$ and X is the Hunt process where the jump kernel of the Dirichlet form is $j(x,y) = (|x-y|^d\psi(|x-y|))^{-1}$.

(ii) Assume further that $0 \le \xi_{xy} \le C_1$ for some deterministic constant $C_1 > 0$. Then the claim of Proposition 7.1(ii) holds.

Proof. The proof of Proposition 7.1 works line by line by plugging $\frac{\phi(k)}{|x-y|^d \phi(k|x-y|)}$ into $j(x,y)$. Note that instead of (7.4), the following holds by using (7.7),

$$\lim_{k\to\infty} k^{-2d} \sum_{\substack{x,y\in V_k\cap B_j\\ |x-y|>\delta}} h(x,y)\frac{\phi(k)}{|x-y|^d\phi(k|x-y|)} = \int_{B_j\times B_j} h(x,y)\,\frac{\mathbf{1}_{\{|x-y|>\delta\}}}{|x-y|^d\psi(|x-y|)} m(dx)m(dy).$$

Given this equality, we can obtain **(A4)*** (iii) by the same way as that of Proposition 7.1. □

Remark 7.3 (i) In Theorem 7.1, we assumed that $d \ge 2$. The above proof of Theorem 7.1 works for $d = 1$ except that the right hand side of (7.6) is no longer summable. We can however obtain the corresponding results (strong convergence of the semigroup and weak convergence) in dimension 1 for any subsequence $\{n_k\}$ such that $\sum_k 1/n_k < \infty$.

(ii) The most typical case in the Example 7.2 is to take $\phi(r) = r^\alpha$. Then $X^{(k)}_t = k^{-1}X^{(1)}_{k^\alpha t}$. Thus Proposition 7.1 says that, if $d \ge 2$, $0 \le \xi_{x,y}$, $\mathbf{E}[\xi_{x,y}] = 1$ and Var $(\xi_{x,y}) < \infty$, then for any positive function $\varphi \in C_c(E)$, $\{(k^{-1}X^{(1)}_{k^\alpha t}, \mathbb{P}^{(k)}_\varphi);\, k \ge 1\}$ converges weakly to $(X, \mathbb{P}_\varphi)$ on $\mathbb{D}_{E_\partial}[0,1]$ equipped with the convergence-in-measure topology $\mathbf{P}$-a.s., which in particular implies the finite dimensional convergence. Assume further that $0 \le \xi_{x,y} \le C$ $\mathbf{P}$-a.s. Then $\{(k^{-1}X^{(1)}_{k^\alpha t}, \mathbb{P}^{(k)}_\varphi);\, k \ge 1\}$ converges weakly to $(X, \mathbb{P}_\varphi)$ on $\mathbb{D}_{E_\partial}[0,1]$ equipped with the Skorohod topology $\mathbf{P}$-a.s..

(iii) As mentioned in the introduction, one cannot obtain the a priori Hölder estimates of caloric functions in general (see [1, Theorem 1.9]).

(iv) It would be very nice if one can prove the Mosco convergence for random walk on long range percolation. Unfortunately, **(A4)***(ii) does not hold for the corresponding generator, so we cannot apply Theorem 4.7 to this model. We note that the heat kernel upper bound is obtained recently in [10] for simple random walk on the infinite cluster of supercritical long range percolation on $\mathbb{Z}^d$, where the probability that two vertices x, y are connected behaves asymptotically as $\|x-y\|^{-s}$ for $s \in (d, (d+2)\wedge 2d)$. Further, it is proved in [11] that the scaling limit of the simple random walk converges to an $(s-d)$-stable process for $s \in (d, d+1)$.

8 Appendix

This appendix contains several equivalence conditions for generalized Mosco convergence that was first obtained in [18, Theorem 2.5] (appeared earlier in the second author's thesis [17]). In fact, a similar and more general form of such equivalence conditions for generalized Mosco convergence was discussed in [22] independently. Since we are using a minor modified version of [18, Theorem 2.5] and only the proof of (i) $\Longrightarrow$ (iv) is given in [18], we give full details for readers' convenience. We believe that, even if the version in [22] is quite general, our version in this paper is simple, and it is applicable to many cases.

For $k \geq 1$, $(\mathcal{H}_k, \langle \cdot, \cdot \rangle_k)$ and $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ are Hilbert spaces with the corresponding norms $\| \cdot \|_k$ and $\| \cdot \|$. Suppose that $(a^{(k)}, \mathcal{D}(a^{(k)}))$ and $(a, \mathcal{D}(a))$ are densely defined closed symmetric bilinear forms on $\mathcal{H}^{(k)}$ and $\mathcal{H}$, respectively. We extend the definition of $a^{(k)}(u,u)$ to every $u \in \mathcal{H}^{(k)}$ by defining $a^{(k)}(u,u) = \infty$ for $u \in \mathcal{H}_k \setminus \mathcal{D}[a^{(k)}]$. Similar extension is done for a as well.

We assume throughout this section that for each $k \geq 1$, there is a bounded linear operator $E_k : \mathcal{H}_k \to \mathcal{H}$ such that $\pi_k := E_k^*$ is a left inverse of E_k, that is,

$$\langle \pi_k f, f_k \rangle_k = \langle f, E_k f_k \rangle \quad \text{and} \quad \pi_k E_k f_k = f_k \qquad \text{for every } f \in \mathcal{H}, f_k \in \mathcal{H}_k. \tag{8.1}$$

Moreover we assume that $\pi_k : \mathcal{H} \to \mathcal{H}_k$ satisfies the following two conditions

$$\sup_{k \geq 1} \|\pi_k\| < \infty, \tag{8.2}$$

where $\|\pi_k\|$ denotes the operator norm of π_k, and

$$\lim_{k \to \infty} \|\pi_k f\|_k = \|f\| \qquad \text{for every } f \in \mathcal{H}, \tag{8.3}$$

Let $\|E_k\|$ denote the operator norm of $E_k : \mathcal{H}^{(k)} \to \mathcal{H}$. Note that $\langle E_k f_k, E_k g_k \rangle = \langle f_k, g_k \rangle_k$ for every $f_k, g_k \in \mathcal{H}_k$, $k \geq 1$ and so clearly

$$\|E_k\| \equiv 1 \quad \text{and} \quad \|E_k f_k\| = \|f_k\|_k \quad \text{for every } f_k \in \mathcal{H}_k, \quad k \geq 1. \tag{8.4}$$

Definition 8.1 Under the above setting, we say that the closed bilinear form a^k is Mosco-convergent to a in the generalized sense if

(i) If $v_k \in \mathcal{H}_k$, $u \in \mathcal{H}$ and $E_k v_k \to u$ weakly in $\mathcal{H}$, then

$$\liminf_{k \to \infty} a^{(k)}(v_k, v_k) \geq a(u,u).$$

(ii) For every $u \in \mathcal{H}$, there exists $u_k \in \mathcal{H}_k$ such that $f \in \mathcal{H}$ $E_k u_k \to u$ strongly in $\mathcal{H}$ and

$$\limsup_{k \to \infty} a^{(k)}(u_k, u_k) \leq a(u,u).$$

Before we prove several equivalence conditions for generalized Mosco convergence, we give the following lemma which is useful in establishing the Mosco convergence. Even though the next lemma is essentially same as [20, Lemma 2.8], we give the proof for the completeness.

Lemma 8.2 *Under the above setting, $a^{(k)}$ is Mosco convergent to a in the generalized sense of Definition 8.1 if Definition 8.1(i) holds and in addition the following hold:*

(1) There exists a set $\mathcal{D} \subset \mathcal{H}$ which is dense in $(\mathcal{D}[a], a + \| \cdot \|^2)$.

(2) $\pi_k(\phi) \in \mathcal{D}[a^{(k)}]$ for every $\phi \in \mathcal{D}$.

(3) For every $\phi \in \mathcal{D}$,

$$\limsup_{k\to\infty} a^{(k)}(\pi_k\phi, \pi_k\phi) = a(\phi, \phi).$$

Proof. Fix $u \in \mathcal{H}$ with $a(u,u) < \infty$ and, using the assumption (1), choose $g_k \in \mathcal{D}$ such that $g_k \to u$ in $\mathcal{H}$ and

$$\lim_{k\to\infty} a(g_k, g_k) = a(u,u). \tag{8.5}$$

Note that

$$\lim_{k\to\infty} \|g_k\| = \|u\| \tag{8.6}$$

and, by Definition 8.1(i) and the assumptions (2)–(3),

$$\lim_{m\to\infty} a^{(m)}(\pi_m g_k, \pi_m g_k) = a(g_k, g_k), \quad \text{for all } k \geq 1. \tag{8.7}$$

Using (8.3) and (8.7), recursively we choose $n_k > n_{k-1}$ with $n_0 = 0$ and $k \geq 1$ such that

$$\big|\|\pi_m g_k\|_m - \|g_k\|\big| < \frac{1}{k} \quad \text{and} \quad \big|a^{(m)}(\pi_m g_k, \pi_m g_k) - a(g_k, g_k)\big| < \frac{1}{k} \quad \text{for all } m \geq n_k. \tag{8.8}$$

Define

$$u_m = \begin{cases} \pi_m g_k & \text{if } k \geq 2 \text{ and } n_k \leq m < n_{k+1}, \\ \pi_m g_1 & \text{if } 1 \leq m < n_2. \end{cases}$$

Then by (8.5), (8.6) and (8.8), we have

$$\lim_{m\to\infty} a^{(m)}(u_m, u_m) = a(u,u) \quad \text{and} \quad \lim_{m\to\infty} \|E_m u_m\| = \|u_m\|_m = \|u\|.$$

Moreover, using (8.1)–(8.2) we have that for every $h \in \mathcal{H}$ and $n_k \leq m < n_{k+1}, k \geq 2$,

$$\big|\langle E_m u_m, h\rangle - \langle u, h\rangle\big| = \big|\langle \pi_m(g_k - u), \pi_m h\rangle_m\big| \leq c\|g_k - u\|\|h\|,$$

which goes to zero as $m \to 0$ since $g_k \to u$ in $\mathcal{H}$. Thus $E_m u_m$ to u strongly in $\mathcal{H}$. □

Let $\{T_t^{(k)}, t \geq 1\}$ and $\{G_\lambda^{(k)}, \lambda > 0\}$ be the strongly continuous symmetric contraction semigroup and the resolvent associated with $(a^{(k)}, \mathcal{D}(a^{(k)}))$. The infinitesimal generator of $\{T_t^{(k)}, t \geq 1\}$ (equivalently, of $(a^{(k)}, \mathcal{D}(a^{(k)}))$) will be denoted by A^k. Similarly, the semigroup, resolvent and infinitesimal generator associated with $(a, \mathcal{D}(a))$ will be denoted by $\{T_t, t \geq 0\}$, $\{G_\lambda, \lambda > 0\}$ and A respectively.

Theorem 8.3 *Under the above setting, the followings are equivalent.*

(i) $a^{(k)}$ *is Mosco-convergent to* a *in the generalized sense;*

(ii) $E_k T_t^{(k)} \pi_k \to T_t$ *strongly in* $\mathcal{H}$ *and the convergence is uniform in any finite interval of* $t \geq 0$*;*

(iii) *For each* $f \in \mathcal{C}_0$*, there exists* $\{f_k\}_{k\geq 1}$ *such that* $f_k \in \mathcal{D}[\mathcal{A}^{(k)}]$*,* $E_k f_k \to f$ *and* $E_k \mathcal{A}^{(k)} f_k \to \mathcal{A} f$ *in* $\mathcal{H}$*;*

(iv) $E_k G_\lambda^{(k)} \pi_k \to G_\lambda$ *strongly in* $\mathcal{H}$ *for every* $\lambda > 0$.

Proof. Let $M_0 := \sup_{k\geq 1} \|\pi_k\|$. Note that, by polarization identity and (8.3), we have

$$\lim_{k\to\infty} \langle \pi_k u, \pi_k v\rangle_k = \langle u, v\rangle, \quad \text{for all } u, v \in \mathcal{H}. \tag{8.9}$$

By (8.1)-(8.4), we see that for every $f \in \mathcal{H}$ and $f_k \in \mathcal{H}_k$,

$$\begin{aligned} \lim_{k\to\infty} \|f_k - \pi_k f\|_k^2 &= \lim_{k\to\infty} \left(\|f_k\|_k^2 - 2\langle f_k, \pi_k f\rangle_k + \|\pi_k f\|_k^2\right) \\ &= \lim_{k\to\infty} \left(\|E_k f_k\|^2 - 2\langle E_k f_k, f\rangle + \|f\|^2\right) = \lim_{k\to\infty} \|E_k f_k - f\|^2. \end{aligned}$$

Therefore

$$\lim_{k\to\infty} \left\|T_t^{(k)} \pi_k f - \pi_k T_t f\right\|_k = \lim_{k\to\infty} \left\|E_k T_t^{(k)} \pi_k f - T_t f\right\| \tag{8.10}$$

for every $f \in \mathcal{H}$ and

$$\lim_{k\to\infty} \left\|G_\lambda^{(k)} \pi_k f - \pi_k G_\lambda f\right\|_k = \lim_{k\to\infty} \left\|E_k G_\lambda^{(k)} \pi_k f - G_\lambda f\right\| \tag{8.11}$$

for every $f \in \mathcal{H}$ and $\lambda > 0$.

(ii) $\Longleftrightarrow$ (iii) : It is a special case of [13, Theorem 1.6.1].

(ii) $\Longleftrightarrow$ (iv) : This can be proved using similar argument in the proof of [27, Theorem 3.4.2 and Lemma 3.4.1]. We give a sketch here. Similar to [27, Lemma 3.4.1], one can check the following

$$E_k G_\lambda^{(k)} \left(\pi_k T_t - T_t^{(k)} \pi_k\right) G_\lambda f = \int_0^t E_k T_{t-s}^{(k)} \left(\pi_k G_\lambda^{(k)} - G_\lambda \pi_k\right) T_s f ds \tag{8.12}$$

for $f \in \mathcal{H}$ and $\lambda > 0$. We first prove that (iv) implies (ii).

(ii) $\Longleftarrow$ (iv) : We assume (iv) is true. Fix $\lambda > 0$ and $T > 0$, If $f \in \mathcal{H}$ and $0 \leq t \leq T$,

$$\begin{aligned} & \left\|\left(E_k T_t^{(k)} \pi_k - T_t\right) G_\lambda f\right\| \\ \leq\ & \left\|E_k T_t^{(k)} \left(\pi_k G_\lambda - G_\lambda^{(k)} \pi_k\right) f\right\| + \left\|E_k G_\lambda^{(k)} \left(T_t^{(k)} \pi_k - \pi_k T(t)\right) f\right\| + \left\|\left(E_k G_\lambda^{(k)} \pi_k - G_\lambda\right) T_t f\right\| \\ =\ & I_1 + I_2 + I_3. \end{aligned}$$

$I_1 + I_3$ goes to 0 uniformly on $[0, t]$ as $k \to \infty$ by (iv) and (8.11). If $f \in \mathcal{D}[A]$, the domain of A, there exists $g \in \mathcal{H}$ such that $f = G_\lambda g$. Since

$$\left\|E_k T_{t-s}^{(k)} \left(\pi_k G_\lambda T_s - G_\lambda^{(k)} \pi_k T_s\right) g\right\| \leq M_0 \|G_\lambda T_s g\| + \left\|G_\lambda^{(k)} \pi_k T_s g\right\| \leq \frac{2M_0}{\lambda} \|g\|,$$

by (8.12) and Lebesgue's dominated convergence theorem, we have

$$\begin{aligned} I_2 &\leq \int_0^t \left\|E_k T_{t-s}^{(k)} \left(\pi_k G_\lambda T_s - G_\lambda^{(k)} \pi_k T_s\right) g\right\| ds \\ &\leq \int_0^t \left\|\pi_k G_\lambda T_s g - G_\lambda^{(k)} \pi_k T_s g\right\|_k ds \to 0 \end{aligned}$$

uniformly on $[0,T]$ as $k\to\infty$ by (iv) and (8.11). Since A is densely defined, the above implies that (ii) is true.

(ii) $\Longrightarrow$ (iv): Assume now that (ii) holds. Then for $\lambda>0$ and $f\in\mathcal{H}$,

$$\left\| E_k G_\lambda^{(k)}\pi_k f - G_\lambda f\right\| \le \int_0^\infty e^{-\lambda t}\left\|\left(E_k T_t^{(k)}\pi_k - T_t\right) f\right\| dt \to 0 \qquad \text{as } k\to\infty.$$

(iv) $\Longrightarrow$ (i) : Let

$$a_\lambda(u,v) := \lambda\,\langle u-\lambda G_\lambda u,\, v\rangle \qquad \text{for } u,v\in\mathcal{H}$$

and

$$a_\lambda^{(k)}(u_k,v_k) := \lambda\big\langle u_k-\lambda G_\lambda^k u_k,\, v_k\big\rangle_k \quad \text{for } u_k,v_k\in\mathcal{H}_k.$$

It is well known that $a_\lambda(u,u)$ and $a_\lambda^{(k)}(u_k,u_k)$ are non-decreasing, and $\lim_{\lambda\to\infty} a_\lambda(u,u)=a(u,u)$ and $\lim_{\lambda\to\infty} a_\lambda^{(k)}(u_k,u_k)=a^k(u_k,u_k)$ for every $u\in\mathcal{H}$ and $u_k\in\mathcal{H}_k$.

Assume (iv) is true. By (8.11) and (8.3),

$$\lim_{k\to\infty}\left\|\left(G_\lambda^k\pi_k-\pi_k G_\lambda\right) f\right\|_k = \lim_{k\to\infty}\left\|E_k G_\lambda^k\pi_k f - G_\lambda f\right\| = 0, \quad \lim_{k\to\infty}\left\|G_\lambda^k \pi_k f\right\|_k = \|G_\lambda f\| \tag{8.13}$$

for every $f\in\mathcal{H}$ and $\lambda>0$. Since

$$\begin{aligned}
&|\,\lambda\langle \pi_k u-\lambda G_\lambda^k\pi_k u,\,\pi_k u\rangle_k - \lambda\langle u-\lambda G_\lambda u,\, u\rangle\,| \\
\le\;& \lambda^2\,\|\,(G_\lambda^k\pi_k-\pi_k G_\lambda)u\,\|_k\,\|\,\pi_k u\,\|_k + \lambda\,|\,\langle \pi_k(u-\lambda G_\lambda u),\,\pi_k u\rangle_k - \langle u-\lambda G_\lambda u,\, u\rangle\,|,
\end{aligned}$$

by (8.2), (8.9) and (8.13) we have

$$\lim_{k\to\infty} a_\lambda^{(k)}(\pi_k u,\,\pi_k u) \;=\; a_\lambda(u,u) \qquad \text{for } \lambda>0. \tag{8.14}$$

Suppose $v_k\in\mathcal{H}_k$, $u\in\mathcal{H}$ and $E_k v_k$ converges weakly to u in $\mathcal{H}$. By (8.1) and (8.9)

$$\lim_{k\to\infty}|\langle v_k-\pi_k u,\,\pi_k g\rangle_k| \qquad \text{for every } g\in\mathcal{H}. \tag{8.15}$$

We also have

$$\lim_{k\to\infty}\langle v_k,\,\pi_k u\rangle_k \;=\; \|u\|^2, \qquad \sup_{k\ge 1}\|v_k\|_k<\infty \qquad\text{and}\qquad \liminf_{k\to\infty}\|v_k\|_k \;\ge\; \|u\|.$$

Note that

$$a^k(v_k,v_k)\ge a_\lambda^{(k)}(v_k,\;v_k)\ge a_\lambda^{(k)}(\pi_k u,\pi_k u)+2\lambda\langle \pi_k u-\lambda G_\lambda^k\pi_k u,\, v_k-\pi_k u\rangle_k.$$

Since, by (iv) and (8.15),

$$\begin{aligned}
|\,\langle \pi_k u-\lambda G_\lambda^k\pi_k u,\, v_k-\pi_k u\rangle_k\,| \;\le\;& |\,\langle \pi_k u,\, v_k-\pi_k u\rangle_k\,| \\
&+\lambda\,|\,\langle \pi_k G_\lambda u,\, v_k-\pi_k u\rangle_k\,| \\
&+\lambda\,\|\,G_\lambda^k\pi_k u-\pi_k G_\lambda u\,\|_k\,(\|\,v_k\,\|_k+\|\,\pi_k u\,\|_k),
\end{aligned}$$

goes to 0 as $k \to \infty$, we have by (8.14),

$$\liminf_{k\to\infty} a^k(v_k, v_k) \geq \liminf_{k\to\infty} a_\lambda^{(k)}(v_k,\ v_k) \geq \liminf_{k\to\infty} a_\lambda^{(k)}(\pi_k u, \pi_k u) = a_\lambda(u, u).$$

Letting $\lambda \to \infty$, we obtain

$$\liminf_{k\to\infty} a^k(v_k,\ v_k)\ \geq\ a(u,\ u).$$

Now we suppose $u \in \mathcal{D}[a]$ and show (ii) in Definition 8.1. First note that, by (iv),

$$\lim_{\lambda\to\infty} \lambda \lim_{k\to\infty} E_k G_\lambda^k \pi_k u = \lim_{\lambda\to\infty} \lambda G_\lambda u = u, \qquad \text{in } \mathcal{H}.$$

Thus, by (8.14) and the monotonicity of $a_\lambda^{(k)}$, we can choose an increasing sequence $\{\lambda_k\}_{k\geq 1}$ such that

$$\lim_{k\to\infty} \lambda_k = \infty, \quad \lim_{k\to\infty} \lambda_k E_k G_{\lambda_k}^k \pi_k u = u \text{ in } \mathcal{H} \qquad \text{and} \qquad \lim_{k\to\infty} a_{(\lambda_k)}^k(\pi_k u,\ \pi_k u)\ \leq\ a(u,\ u) < \infty.$$

For $k \geq 1$, let $u_k := \lambda_k G_{\lambda_k}^k \pi_k u \in \mathcal{H}_k$ and note that $E_k u_k \to u$ in $\mathcal{H}$. Since

$$a_{(\lambda_k)}^k(\pi_k u,\ \pi_k u) = a^k(u_k, u_k) + \lambda_k \|u_k - \pi_k u\|_k^2 = a^k(u_k,\ u_k) + \lambda_k \|E_k u_k - u\|^2,$$

we conclude that

$$a(u,\ u)\ \geq\ \limsup_{k\to\infty} a^k(u_k,\ u_k).$$

(i) $\Longrightarrow$ (iv) : Suppose (i) is true. Fix $\lambda > 0$ and assume $f \in \mathcal{H}$. Since

$$\sup_{k\geq 1} \|E_k G_\lambda^{(k)} \pi_k\| \leq \frac{M_0}{\lambda} < \infty,$$

there exists a subsequence of $\left\{E_k G_\lambda^{(k)} \pi_k f\right\}_{k\geq 1}$, still denoted $\left\{E_k G_\lambda^{(k)} \pi_k f\right\}_{k\geq 1}$, such that $E_k G_\lambda^{(k)} \pi_k f$ converges weakly in $\mathcal{H}$ to some $\tilde{u}$ in $\mathcal{H}$. So by Definition 8.1(i)

$$\liminf_{k\to\infty} \left(a^{(k)}(G_\lambda^{(k)} \pi_k f,\ G_\lambda^{(k)} \pi_k f) + \lambda \left\| G_\lambda^{(k)} \pi_k f \right\|_k^2 \right)\ \geq\ a(\tilde{u}, \tilde{u}) + \lambda \|\tilde{u}\|^2 . \tag{8.16}$$

By (8.1) and (8.16),

$$\begin{aligned}
& a(\tilde{u}, \tilde{u}) + \lambda \|\tilde{u}\|^2 - 2\langle f, \tilde{u}\rangle \\
\leq\ & \liminf_{k\to\infty} \left(a^{(k)}(G_\lambda^{(k)} \pi_k f, G_\lambda^{(k)} \pi_k f) + \lambda \left\| G_\lambda^{(k)} \pi_k f \right\|_k^2 \right) - 2 \lim_{k\to\infty} \langle f, E_k G_\lambda^{(k)} \pi_k f\rangle \\
\leq\ & \liminf_{k\to\infty} \left(a^{(k)}(G_\lambda^{(k)} \pi_k f, G_\lambda^{(k)} \pi_k f) + \lambda \left\| G_\lambda^{(k)} \pi_k f \right\|_k^2 - 2\langle \pi_k f, G_\lambda^{(k)} \pi_k f\rangle_k \right) \qquad (8.17) \\
\leq\ & \limsup_{k\to\infty} \left(a^{(k)}(G_\lambda^{(k)} \pi_k f, G_\lambda^{(k)} \pi_k f) + \lambda \left\| G_\lambda^{(k)} \pi_k f \right\|_k^2 - 2\langle \pi_k f, G_\lambda^{(k)} \pi_k f\rangle_k \right). \qquad (8.18)
\end{aligned}$$

For arbitrary $v \in \mathcal{H}$, by Definition 8.1(ii), there exist $v_k \in \mathcal{H}_k$ such that

$$\lim_{k\to\infty} \|v_k\|_k\ =\ \|v\|, \quad \lim_{k\to\infty} \langle E_k v_k, f\rangle = \langle v, f\rangle \quad \text{and} \quad \limsup_{k\to\infty} a^{(k)}(u_k, u_k) \leq a(u, u). \tag{8.19}$$

Since $G_\lambda^{(k)}\pi_k f$ is the unique minimizer of $a^{(k)}(\,\cdot\,,\,\cdot\,)+\lambda\,\|\cdot\|_k^2-2\langle\pi_k f,\,\cdot\,\rangle_k$ over $\mathcal{H}_k$ for each $k\geq 1$, (8.18) is less than or equals to

$$\limsup_{k\to\infty} a^{(k)}(v_k,v_k)+\lambda\limsup_{k\to\infty}\|v_k\|_k^2-2\liminf_{k\to\infty}\langle\pi_k f,v_k\rangle_k.$$

By (8.19), the above is less than or equals to $a(v,v)+\lambda\,\|v\|^2-2\langle f,v\rangle$. Therefore $\tilde u=G_\lambda f$ because $G_\lambda f$ is the unique minimizer of $a(\,\cdot\,,\,\cdot\,)+\lambda\,\|\cdot\|^2-2\langle f,\,\cdot\,\rangle$ over $\mathcal{H}$.

On the other hand, by (i) there exists $w_k\in\mathcal{H}_k$ such that

$$\lim_{k\to\infty}\|E_k w_k-G_\lambda f\|=0\quad\text{and}\quad\lim_{k\to\infty}a^{(k)}(w_k,\ w_k)=a(G_\lambda f,\,G_\lambda f).$$

So by (8.16), the second equation above and the unique minimizer argument used above, we have

$$\begin{aligned}
&\lambda\limsup_{k\to\infty}\left\|G_\lambda^{(k)}\pi_k f-\frac{\pi_k f}{\lambda}\right\|_k^2\\
\leq\ &\limsup_{k\to\infty}\left(a^{(k)}(w_k,w_k)-a^{(k)}(G_\lambda^{(k)}\pi_k f,\ G_\lambda^{(k)}\pi_k f)+\lambda\left\|w_k-\frac{\pi_k f}{\lambda}\right\|_k^2\right)\\
\leq\ &\limsup_{k\to\infty}a^{(k)}(w_k,w_k)-\liminf_{k\to\infty}a^{(k)}(G_\lambda^{(k)}\pi_k f,\ G_\lambda^{(k)}\pi_k f)+\lambda\limsup_{k\to\infty}\left\|w_k-\frac{\pi_k f}{\lambda}\right\|_k^2\\
\leq\ &\lambda\limsup_{k\to\infty}\left\|w_k-\frac{\pi_k f}{\lambda}\right\|_k^2.
\end{aligned}$$

Combining the above inequality with

$$\lim_{k\to\infty}\langle G_\lambda^{(k)}\pi_k f,\,\pi_k f\rangle_k\ =\ \langle G_\lambda f,f\rangle\quad\text{and}\quad\lim_{k\to\infty}|\langle\pi_k f,w_k\rangle_k-\langle f,G_\lambda f\rangle|\ =\ 0,$$

we obtain

$$\begin{aligned}
&\limsup_{k\to\infty}\|G_\lambda^{(k)}\pi_k f\|_k\\
=\ &\limsup_{k\to\infty}\left\|G_\lambda^{(k)}\pi_k f-\frac{\pi_k f}{\lambda}\right\|_k^2+2\lim_{k\to\infty}\langle G_\lambda^{(k)}\pi_k f,\,\pi_k f\rangle_k-\lim_{k\to\infty}\left\|\frac{\pi_k f}{\lambda}\right\|_k^2\\
\leq\ &\limsup_{k\to\infty}\left\|w_k-\frac{\pi_k f}{\lambda}\right\|_k^2+2\lim_{k\to\infty}\langle G_\lambda^{(k)}\pi_k f,\,\pi_k f\rangle_k-\lim_{k\to\infty}\left\|\frac{\pi_k f}{\lambda}\right\|_k^2\\
\leq\ &\limsup_{k\to\infty}\left\|w_k-\frac{\pi_k f}{\lambda}\right\|_k^2+2\lim_{k\to\infty}\langle w_k,\,\pi_k f\rangle_k-\lim_{k\to\infty}\left\|\frac{\pi_k f}{\lambda}\right\|_k^2\\
=\ &\limsup_{k\to\infty}\left\|(w_k-\frac{\pi_k f}{\lambda})+\frac{\pi_k f}{\lambda}\right\|_k^2\ =\ \limsup_{k\to\infty}\|w_k\|_k.
\end{aligned}$$

Therefore

$$\limsup_{k\to\infty}\|E_k G_\lambda^{(k)}\pi_k f\|\ =\ \limsup_{k\to\infty}\|G_\lambda^{(k)}\pi_k f\|_k\ \leq\ \limsup_{k\to\infty}\|E_k w_k\|\ =\ \|G_\lambda f\|$$

and we conclude that, for every $f\in\mathcal{H}$, $E_k G_\lambda^{(k)}\pi_k f$ converges to $G_\lambda f$ in $\mathcal{H}$. □

References

[1] M.T. Barlow, R.F. Bass, Z.-Q. Chen and M. Kassmann. Non-local Dirichlet forms and symmetric jump processes. *Trans. Amer. Math. Soc.* **361** (2009), 1963–1999.

[2] R. F. Bass, M. Kassmann and T. Kumagai. Symmetric jump processes: localization, heat kernels, and convergence. *Ann. Inst. Henri. Poincaré Probab. Statist.* **46** (2010), 59–71.

[3] R. F. Bass and T. Kumagai. Symmetric Markov chains on $\mathbb{Z}^d$ with unbounded range. *Trans. Amer. Math. Soc.* **360** (2008), 2041–2075.

[4] R.F. Bass, T. Kumagai and T. Uemura. Convergence of symmetric Markov chains on $\mathbb{Z}^d$. *Probab. Theory Relat. Fields* **148** (2010), 107–140.

[5] K. Burdzy and Z.-Q. Chen. Weak convergence of reflecting Brownian motions. *Elect. Comm. Probab.* **3** (1998), 29–33.

[6] K. Burdzy and Z.-Q. Chen. Discrete approximations to reflected Brownian motion. *Ann. Probab.* **36** (2008), 698–727.

[7] Z.-Q. Chen, P. J. Fitzsimmons, R. Song. Crossing estimates for symmetric Markov processes. *Probab. Theory Relat. Fields* **120** (2001), 68–84.

[8] Z.-Q. Chen, P. J. Fitzsimmons, K. Kuwae and T.-S. Zhang. Stochastic calculus for symmetric Markov processes. *Ann. Probab.* **36** (2008), 931–970.

[9] Z.-Q. Chen and T. Kumagai. Heat kernel estimates for jump processes of mixed types on metric measure spaces. *Probab. Theory Relat. Fields* **140** (2008), 277–317.

[10] N. Crawford and A. Sly. Simple random walks on long range percolation clusters I: heat kernel bounds. *Probab. Theory Relat. Fields*, to appear.

[11] N. Crawford and A. Sly. Simple random walks on long range percolation clusters II: scaling limits. *Preprint*.

[12] D. Dellacherie and P. A. Meyer. *Probabilités et Potentie, I-IV*. Hermann, Paris, 1975.

[13] S. N. Ethier and T. G. Kurtz. *Markov Processes : Characterization and Convergence*. John Wiley & Sons, New York, 1986.

[14] M. Fukushima, Y. Oshima and M. Takeda. *Dirichlet Forms and Symmetric Markov Processes.* Walter de Gruyter, Berlin, 1994.

[15] R. Husseini and M. Kassmann. Markov Chain approximations for symmetric jump processes. *Potential Anal.* **27** (2007), 353–380.

[16] J. Jacod and A.N. Shiryaev. *Limit Theorems for Stochastic Processes.* Springer-Verlag, 1987.

[17] P. Kim, *Potential Theory for Stable Processes*. PhD Thesis, University of Washington (2004).

[18] P. Kim. Weak convergence of censored and reflected stable processes. *Stochastic Process. Appl.* **116** (2006), 1792–1814.

[19] P. Kim, R. Song and Z. Vondraček, Boundary Harnack principle for subordinate Brownian motion. *Stochastic Process. Appl.* **119** (2009), 1601–1631.

[20] A. V. Kolesnikov. Mosco convergence of Dirichlet forms in infinite dimensions with changing reference measures. *J. Funct. Anal.* **230** (2006), 382–418.

[21] T. Kumagai, K.-T. Sturm. Construction of diffusion processes on fractals, d-sets, and general metric measure spaces. *J. Math. Kyoto Univ.* **45** (2005), 307–327.

[22] K. Kuwae and T. Shioya. Convergence of spectral structures: a functional analytic theory and its applications to spectral geometry. *Comm. Anal. Geom.* **11** (2003), 599–673.

[23] T. Lyons and W. Zheng. A crossing estimate for the canonical process on a Dirichlet space and a tightness result. *Astérisque* **157–158** (1988), 249–271.

[24] J. Masamune and T. Uemura. Conservation property of symmetric jump processes. *Ann. Inst. Henri. Poincaré Probab. Statist.*, **47(3)** (2011), 650–662.

[25] P.A. Meyer and W. Zheng. Tightness criteria for laws of semimartingales. *Ann. Inst. Henri. Poincaré Probab. Statist.* **20** (1984), 353–372.

[26] U. Mosco. Composite media and asymptotic Dirichlet forms. *J. Funct. Anal.* **123** (1994), 368–421.

[27] A. Pazy. *Semigroups of Linear Operators and Applications to Partial Differential Equations*. Springer-Verlag, New York, 1983.

[28] D.W. Stroock and S.R.S. Varadhan. *Multidimensional Diffusion Processes*. Springer-Verlag, Berlin, 1979.

[29] D.W. Stroock and W. Zheng. Markov chain approximations to symmetric diffusions. *Ann. Inst. Henri. Poincaré Probab. Statist.* **33** (1997), 619–649.

Zhen-Qing Chen
Department of Mathematics, University of Washington, Seattle, WA 98195, USA
E-mail: zqchen@uw.edu

Panki Kim
Department of Mathematical Sciences and Research Institute of Mathematics,

Seoul National University, Building 27, 1 Gwanak-ro, Gwanak-gu, Seoul 151-747, Republic of Korea
E-mail: pkim@snu.ac.kr

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