Construction of Heat Kernels on Metric Measure Spaces via Multiresolution Analysis

Jun Cao, Alexander Grigor'yan and Liguang Liu

Abstract. Let $(\mathcal{X}, d, \mu)$ be a metric measure space satisfying the volume doubling condition. In this paper, the authors provide a new and direct method of constructing stochastic complete (signed) heat kernels by virtue of the multiresolution analysis (for short, MRA) structure. For any $\beta \in (0, \infty)$, two kinds of applications are given: (i) via taking a non-smooth MRA generated by Haar wavelets, such construction gives rise to a heat kernel satisfying only stable-like upper estimate of index β (no continuity and near-diagonal lower estimate); (ii) via taking a smooth MRA generated by smooth wavelets/splines, such construction gives rise to a signed heat kernel (no positivity) satisfying the stable-like upper estimate of index β, as well as the almost Lipschitz regularity and the near-diagonal lower estimate.

Contents

2020 *Mathematics Subject Classification*. Primary: 35K08; Secondary: 42C40, 46E36.
Key words and phrases. Heat kernel, metric measure space, multiresolution analysis, wavelets.
J. Cao is supported by the National Natural Science Foundation of China (# 11871254); A. Grigor'yan was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Project-ID 317210226 - SFB 1283; L. Liu is supported by the National Natural Science Foundation of China (# 12371102).

1 Introduction

Let $(\mathcal{X}, d)$ be a separable metric space. For any $x \in \mathcal{X}$ and $r \in (0, \infty)$, define the ball

$$B(x, r) := \{y \in \mathcal{X} : d(y, x) < r\}.$$

We always assume that the closure of every open ball is compact. Suppose that on $(\mathcal{X}, d)$ there is a nonnegative Radon measure μ such that $0 < \mu(B(x, r)) < \infty$ for all $x \in \mathcal{X}$ and $r \in (0, \infty)$. The triple $(\mathcal{X}, d, \mu)$ will be referred to as a *metric measure space*.

For any $x, y \in \mathcal{X}$ and $r > 0$, we use the following notation throughout the whole paper:

$$V(x, r) := \mu(B(x, r)) \quad \text{and} \quad V(x, y) := \mu(B(x, d(x, y))).$$

We say that $(\mathcal{X}, d, \mu)$ satisfies the *volume doubling condition* (VD) if there exists $C_D \in (1, \infty)$ such that

$$V(x, 2r) \le C_D V(x, r) \quad \text{for all } x \in \mathcal{X} \text{ and } r \in (0, \infty). \tag{VD}$$

Note that (VD) holds if and only if there exist constants $C'_D \in (1, \infty)$ and $n \in (0, \infty)$ such that for all $x, y \in \mathcal{X}$ and $0 < r \le R < \infty$,

$$\frac{V(x, R)}{V(y, r)} \le C'_D \left(\frac{d(x, y) + R}{r}\right)^n. \tag{1.1}$$

The condition (VD) also implies the following *geometrical doubling property*: there exists an integer $M \in (0, C_D^4]$ (see, e.g. [20, p. 67] or [36, p. 489]) such that:

$$\text{every open ball } B(x, r) \text{ can be covered by at most } M \text{ open balls of radius } r/2. \tag{1.2}$$

Many important underlying spaces, such as the (weighted) Euclidean space $\mathbb{R}^n$, convex unbounded domains in $\mathbb{R}^n$, Riemannian manifolds of nonnegative Ricci curvature, nilpotent Lie groups with polynomial growth and fractals, are all fall into the scope of metric measure spaces fulfilling (VD).

According to [35, Proposition 5.2], if $(\mathcal{X}, d, \mu)$ is connected and satisfies (VD), then it satisfies the following *reverse volume doubling condition* (RVD): there exist constants $C_{RD} \in (0, \infty)$ and $\kappa \in (0, \infty)$ such that

$$\frac{V(x, R)}{V(x, r)} \ge C_{RD} \left(\frac{R}{r}\right)^\kappa \quad \text{for all } x \in \mathcal{X} \text{ and } 0 < r \le R < \operatorname{diam}(\mathcal{X}). \tag{RVD}$$

In other words, (RVD) is a very mild condition, which is occasionally used in analysis on doubling metric measure spaces.

Definition 1.1. A family of jointly measurable function $\{p_t\}_{t>0}$ on $\mathcal{X} \times \mathcal{X}$ is called a *heat kernel*, if it satisfies the following properties for all values of the variables involved:

(P1) *(positivity)* $p_t(x, y) \geq 0$ and $\int_{\mathcal{X}} p_t(x, y)\, d\mu(y) \leq 1$;

(P2) *(symmetry)* $p_t(x, y) = p_t(y, x)$;

(P3) *(semigroup property)* $\int_{\mathcal{X}} p_t(x, z) p_s(z, y)\, d\mu(z) = p_{t+s}(x, y)$;

(P4) *(approximation of identity)* for any $f \in L^2(\mathcal{X})$,

$$\int_{\mathcal{X}} p_t(x, y) f(y)\, d\mu(y) \to f(x) \;\text{ as }\; t \to 0,$$

where the convergence is in $L^2(\mathcal{X}, \mu)$.

The family $\{p_t\}_{t>0}$ is called a *signed heat kernel* if it satisfies (P2)-(P3)-(P4) and the following weaker property (P1′):

(P1′) there exists a positive constant C such that for any $t \in (0, \infty)$ and $x \in \mathcal{X}$,

$$\int_{\mathcal{X}} |p_t(x, y)|\, d\mu(y) \leq C;$$

Given a (signed) heat kernel $\{p_t\}_{t>0}$, if for all $t \in (0, \infty)$ and $x \in \mathcal{X}$,

$$\int_{\mathcal{X}} p_t(x, y)\, d\mu(y) = 1, \tag{1.3}$$

then $\{p_t\}_{t>0}$ is called *stochastically complete*.

Heat kernel is a universal gadget that plays a central role in diverse areas of mathematics and physics; see [4, 11, 23, 32, 40, 51] and the references therein. We remark that there do exist various scenarios where the signed heat kernel appears (see, e.g. [3]). In particular, the heat kernel of the biharmonic operator Δ^2 has infinitely negative-valued points (see, e.g. [29]). Such phenomenon plays an important role in constructing counterexamples for both the Szegö and Boggio-Hadamard conjectures (see, e.g. [53]).

Constructing a heat kernel on different kinds of underlying spaces has been extensively studied in literature; see [14, 30, 32, 46, 48, 50] for the study of heat kernels on manifolds, [5, 9, 33, 49] on graphs, and [6, 8, 42, 43] on fractals. Intuitively, constructing a heat kernel on a metric measure space $\mathcal{X}$ is in some sense equivalent to building a heat conductor in $\mathcal{X}$. By encoding the conduction information from heat kernel, one can explore the underlying geometric structure of $\mathcal{X}$.

Let $\beta \in (0, \infty)$. A heat kernel $\{p_t\}_{t>0}$ is said to satisfy the *two-sided stable-like estimate* $(\mathrm{ULE})_\beta$ provided that

$$p_t(x, y) \simeq \frac{1}{V(x, t^{1/\beta}) + V(x, y)} \left(\frac{t^{1/\beta}}{t^{1/\beta} + d(x, y)} \right)^{\beta} \quad \text{for all } x, y \in \mathcal{X} \text{ and } t \in (0, \infty). \tag{ULE_β}$$

We say that $(\mathrm{UE})_\beta$ (resp. $(\mathrm{LE})_\beta$) is satisfied, if the upper (resp. lower) estimate in $(\mathrm{ULE})_\beta$ holds.

One method of obtaining a heat kernel satisfying $(\mathrm{ULE})_\beta$ is by subordinating a heat kernel $\{p_t\}_{t>0}$ that satisfies the *sub-Gaussian estimate* $(\mathrm{SG})_{d_w}$ of the form

$$p_t(x,y) \asymp \frac{C}{V(x,t^{1/d_w})}\exp\left(-c\left(\frac{d(x,y)}{t^{1/d_w}}\right)^{\frac{d_w}{d_w-1}}\right) \quad \text{for all } x,y\in\mathcal{X} \text{ and } t\in(0,\infty), \qquad (\mathrm{SG})_{d_w}$$

where the sign $\asymp$ means that both $\le$ and $\ge$ hold but with different values of positive constants C, c, and d_w is a parameter from $[2,\infty)$ that is called the *walk dimension* of the heat kernel. Such a method can be found in [31, Section 5.4] when $(\mathcal{X},d,\mu)$ is α-regular for some $\alpha\in(0,\infty)$, that is, $V(x,r)\simeq r^\alpha$ for all $x\in\mathcal{X}$ and $r\in(0,\infty)$. See also [13, Section 4] when $(\mathcal{X},d,\mu)$ is a general metric measure spaces satisfying (VD). From [7], it follows that any heat kernel satisfying $(\mathrm{SG})_{d_w}$ is jointly continuous (it indeed satisfies the Hölder continuity estimate), so does the subordinated heat kernel.

Another widely used method in literature is obtaining a heat kernel from a *regular Dirichlet form* (see, e.g. [28, 34, 35, 15, 16]). Consider the following jump-type bilinear form

$$\mathcal{E}_\beta(f,f) := \int_{\mathcal{X}}\int_{\mathcal{X}} \frac{|f(x)-f(y)|^2}{d(x,y)^\beta}\,\frac{d\mu(y)\,d\mu(x)}{V(x,\,y)}, \tag{1.4}$$

with a natural domain

$$\mathcal{F}_\beta := \left\{f\in L^2(\mathcal{X}) :\ \mathcal{E}_\beta(f,f)<\infty\right\}. \tag{1.5}$$

In the case $\beta\in(0,2)$, since the collection of Lipschitz functions with compact support is dense in $\mathcal{F}_\beta$, it follows that $(\mathcal{E}_\beta, \mathcal{F}_\beta)$ is a regular Dirichlet form. Invoking this fact, on an α-regular space $(\mathcal{X},d,\mu)$, Chen and Kumagai [15] proved that when $\beta\in(0,2)$ the heat kernel $\{p_t\}_{t>0}$ of $(\mathcal{E}_\beta,\mathcal{F}_\beta)$ exists and satisfies $(\mathrm{ULE})_\beta$. Further, we have by [16, Lemma 5.6] that the two-sided estimate $(\mathrm{ULE})_\beta$ of such $\{p_t\}_{t>0}$ ensures a Hölder continuity estimate, but with a very small Hölder exponent.

In general, we follow [13] and define the *critical index* $\beta^\sharp$ that relates to the possible values of β in $(\mathrm{ULE})_\beta$ as follows:

$$\begin{aligned}\beta^\# := \sup\{\beta>0 :\ &\text{there exists a stochastically complete}\\ &\text{continuous heat kernel } \{p_t\}_{t>0} \text{ on } \mathcal{X} \text{ satisfying } (\mathrm{ULE})_\beta\}.\end{aligned} \tag{1.6}$$

Under (VD) and (RVD), we know from [13] that $\beta^\sharp\in[2,\infty)$ and, moreover, for any $\beta\in(0,\beta^\sharp)$, the bilinear form $(\mathcal{E}_\beta, \mathcal{F}_\beta)$ in (1.4)-(1.5) becomes a regular Dirichlet form and there is a stochastically complete continuous heat kernel $\{p_t\}_{t>0}$ satisfying $(\mathrm{ULE})_\beta$ which exists as the transition probability of the Markov process corresponding to $(\mathcal{E}_\beta, \mathcal{F}_\beta)$. This extends the work of Chen and Kumagai [15], which treats only the case $\beta\in(0,2)$. Again, applying [16, Lemma 5.6] (see also [12, Theorem 2.22]) yields that such $\{p_t\}_{t>0}$ satisfies the Hölder regularity estimate.

Under (VD) and (RVD), we know from [13, Theorem 1.3] that the critical index $\beta^\#$ is invariant under *quasi-isometry* of two metric measure spaces, where $(\mathcal{X},d,\mu)$ is quasi-isometric to $(\mathcal{X},d',\mu')$ if and only if $d\simeq d'$ and $\mu\simeq\mu'$. Assuming only (VD), if there is a stochastically complete continuous heat kernel $\{p_t\}_{t>0}$ on $\mathcal{X}$ satisfying $(\mathrm{SG})_{d_w}$, then (see [13, Theorem 1.5])

$$\beta^\sharp = d_w = \beta^*,$$

where β^* is another critical index of *Besov spaces* on $(\mathcal{X},d,\mu)$, defined by

$$\beta^* := \sup\left\{\beta>0 :\ \Lambda^{\beta/2}_{2,\infty}(\mathcal{X}) \text{ is dense in } L^2(\mathcal{X})\right\},$$

with the *Besov space* $\Lambda^{\beta/2}_{2,\infty}(\mathcal{X})$ defined to the collection of all functions $f \in L^2(\mathcal{X})$ such that

$$\|f\|_{\Lambda^{\beta/2}_{2,\infty}(\mathcal{X})} := \|f\|_{L^2(\mathcal{X})} + \left(\sup_{r\in(0,\infty)} \int_{\mathcal{X}} \left(\frac{1}{V(x,r)} \int_{B(x,r)} \frac{|f(x)-f(y)|^2}{r^\beta}\, d\mu(y)\right) d\mu(x)\right)^{1/2} < \infty.$$

In conclusion, the critical index $\beta^{\#}$ is not only an intrinsic value of the underlying space $\mathcal{X}$ but also a good candidate for the walk dimension in future attempts to construct a diffusion process on $\mathcal{X}$.

We are wondering what happens beyond the critical index $\beta^\sharp$. It turns out that $\beta^\sharp$ can be broke through if certain requirements of the heat kernel in (1.6) are sacrificed. The key point for the occurrence of this interesting phenomenon is that, under (VD), the family $\mathcal{D}$ of *dyadic cubes* on $(\mathcal{X}, d, \mu)$ (see Theorem 4.1 below) exists and therefore induces an *ultra-metric* $d_{\mathcal{D}}$ on $\mathcal{X}$ (see Definition 4.4 below). For any given $\beta \in (0,\infty)$, by following the general procedure of heat kernel construction on ultra-metric spaces in [10], the authors in [13] construct a stochastically complete heat kernel $\{p_t^{\mathcal{D}}\}_{t>0}$ via defining

$$p_t^{\mathcal{D}}(x,y) := \int_{d_{\mathcal{D}}(x,y)}^{\infty} \frac{d(\sigma(r))^t}{\mu(\overline{B}_{\mathcal{D}}(x,r))} \qquad \text{for all } t\in(0,\infty) \text{ and } x,y\in\mathcal{X}, \tag{1.7}$$

where

$$\sigma(r) := \exp(-r^{-\beta})$$

and

$$\overline{B}_{\mathcal{D}}(x,r) := \{y\in\mathcal{X} : d_{\mathcal{D}}(y,x) \le r\}.$$

Such $\{p_t^{\mathcal{D}}\}_{t>0}$ satisfies the two-sided stable-like estimate with respect to $d_{\mathcal{D}}$ (see [13, Theorem 5.6]):

$$p_t^{\mathcal{D}}(x,y) \simeq \frac{1}{V(x,\, t^{1/\beta} + d_{\mathcal{D}}(x,y))} \left(\frac{t^{1/\beta}}{t^{1/\beta} + d_{\mathcal{D}}(x,y)}\right)^\beta \qquad \text{for all } t\in(0,\infty) \text{ and } x,y\in\mathcal{X}, \tag{1.8}$$

which implies that $\{p_t^{\mathcal{D}}\}_{t>0}$ satisfies the upper estimate

$$0 \le p_t^{\mathcal{D}}(x,y) \le \frac{C}{V(x,\, t^{1/\beta} + d(x,y))} \left(\frac{t^{1/\beta}}{t^{1/\beta} + d(x,y)}\right)^\beta \qquad \text{for all } t\in(0,\infty) \text{ and } x,y\in\mathcal{X} \tag{1.9}$$

with respect to the original metric d. However, the near-diagonal lower estimate breaks down because $p_t^{\mathcal{D}}(x,y) = 0$ whenever the points x, y are very close to each other but can not be covered by any dyadic cube. As a consequence, there is no Hölder continuity for $\{p_t^{\mathcal{D}}\}_{t>0}$. Further, via considering the adjacent family of dyadic cubes (see [37])

$$\{\mathcal{D}^\tau : \tau = 1, \dots, K\},$$

it turns out that (see [13]) the summation of the heat kernels

$$\left\{\{p_t^{\mathcal{D}^\tau}\}_{t>0} : \tau = 1, \dots, K\right\}$$

forms a family of functions $\{p_t\}_{t>0}$ on $\mathcal{X}\times\mathcal{X}$ which satisfies (ULE)$_\beta$ and all properties of a stochastically complete heat kernel, except the semigroup property.

Let us remark that the heat kernel $\{p_t^{\mathcal{D}}\}_{t>0}$ constructed in (1.7) can be expressed by means of the eigenfunction expansion because by [10, Theorem 3.8] the generator of the corresponding semigroup has a sequence of eigenfunctions that form an orthonormal basis in $L^2(\mathcal{X})$. Denote this sequence by $\{\phi_j\}$ and let the eigenvalue of ϕ_j be λ_j. Then

$$p_t^{\mathcal{D}}(x,y) = \sum_j e^{-t\lambda_j}\phi_j(x)\phi_j(y).$$

It was also shown in [10] that the eigenvalues λ_j are determined by the function σ in (1.7), and the eigenfunctions ϕ_j are similar to Haar functions but determined by pairs of concentric ultra-metric balls.

It is known that, on a metric measure space satisfying (VD), there do exist Haar wavelets (see [41]) and Hölder-continuous wavelets with exponential decay (see [2]), which form an orthonormal basis in $L^2(\mathcal{X})$. Thus, the above discussions inspire us to consider the following question:

> *On a general metric measure space, is it possible to use a certain family of wavelets to construct a heat kernel with "good" estimates?*

Since each family of orthogonal wavelets generates a *multiresolution analysis* (for short, MRA) structure [22, 47], it is more natural to ask the following general question:

> *Given an MRA structure on a metric measure space, is it possible to construct a heat kernel with "good" estimates?*

The main aim of this paper is to address the above questions. Indeed, without referring to the deep theory of Dirichlet forms, we provide a new and direct method of constructing (signed) heat kernels:

– By using only the MRA structure of the underlying space, we construct explicitly stochastically complete (signed) heat kernels (see Theorem 2.6 and Corollary 2.7 below).

– As the first application, taking a non-smooth MRA that is formed by Haar wavelets (see [41]), we show that the corresponding heat kernel $\{p_t\}_{t>0}$ recovers the one $\{p_t^{\mathcal{D}}\}_{t>0}$ given in (1.7), which certainly satisfies stochastic completeness property (1.3) and the stable-like upper estimate (1.9) by terms of [13] (see Theorem 2.8 below).

– As the second application, taking a smooth MRA that is generated by smooth splines/wavelets of Auscher-Hytönen (cf. [2]), we obtain a stochastic complete signed heat kernel $\{p_t\}_{t>0}$ and then prove that it satisfies the stable-like upper estimate, the almost Lipschitz regularity and the near-diagonal lower estimate (see Theorem 2.9 below).

Detailed descriptions of these results are presented in Section 2 below. Before that, based on the previous discussions, we summarize briefly in Table 1 the existence and properties of stochastically complete stable-like heat kernels with index $\beta \in (0, \infty)$.

<table>
<tr><th>Range of β</th><th>Structure used</th><th colspan="2">Heat kernels</th></tr>
<tr><td>$\beta \in (0, \beta^\sharp)$</td><td></td><td colspan="2">exists a jump-type regular Dirichlet form, associated to a stochastic complete positive heat kernel satisfying (ULE)$_\beta$ and Hölder continuity with small Hölder exponent (cf. [13])</td></tr>
<tr><td rowspan="3">$\beta \in (0, \infty)$</td><td>MRA</td><td>exists a stochastic complete signed heat kernel</td><td>no “good” estimates</td></tr>
<tr><td>non-smooth MRA via Haar wavelets</td><td>exists a stochastic complete positive heat kernel satisfying upper estimate (UE)$_\beta$</td><td>no continuity, no near-diagonal lower estimate</td></tr>
<tr><td>smooth MRA via Auscher-Hytönen wavelets/splines</td><td>exists a stochastic complete signed heat kernel satisfying upper estimate (UE)$_\beta$, near-diagonal lower estimate, almost Lipschitz continuity</td><td>no positivity</td></tr>
</table>

Table 1: Good and missing properties of stable-like heat kernels on $\mathcal{X}$ with index β

2 Statement of main results

2.1 Construction of signed heat kernels via MRA

We adopt the following definition of multiresolution analysis from [54, Definition 3.1].

Definition 2.1. Let $\{V_k\}_{k\in\mathbb{Z}}$ be a sequence of closed linear subspaces in $L^2(\mathcal{X})$. Then $\{V_k\}_{k\in\mathbb{Z}}$ is called a *multiresolution analysis* (denoted by MRA for short) in $L^2(\mathcal{X})$ if the following hold:

(i) *(nested property)* for any $k \in \mathbb{Z}$, $V_k \subset V_{k+1}$;

(ii) *(density property)* $\overline{\cup_{k\in\mathbb{Z}} V_k} = L^2(\mathcal{X})$; ;

(iii) *(intersection property)* for some fixed integer $k_0 \in \mathbb{Z}$,

$$\bigcap_{k\in\mathbb{Z}} V_k = \begin{cases} \{0\} & \text{as } \ \mu(\mathcal{X}) = \infty; \\ V_{k_0} & \text{as } \ \mu(\mathcal{X}) < \infty; \end{cases}$$

(iv) *(Riesz basis)* each V_k has a Riesz basis $\{\varphi_{k,\alpha}\}_{\alpha\in\mathcal{A}_k}$, where $\mathcal{A}_k$ is a countable index set. In other words,

$$V_k = \overline{\operatorname{span}\{\varphi_{k,\alpha} : \alpha \in \mathcal{A}_k\}}$$

and, for any sequence $\{\lambda_\alpha\}_{\alpha\in\mathcal{A}_k} \subset \mathbb{C}$ with only finite non-zero elements,

$$\Big\| \sum_{\alpha\in\mathcal{A}_k} \lambda_\alpha \varphi_{k,\alpha} \Big\|_{L^2(\mathcal{X})} \simeq \Big(\sum_{\alpha\in\mathcal{A}_k} |\lambda_\alpha|^2 \Big)^{1/2},$$

where the implicit constant may depend on k but independent of $\{\lambda_\alpha\}_{\alpha\in\mathcal{A}_k}$.

Now we recall some known facts about orthogonal projections in Hilbert spaces (see [57, pp. 81-84]). For any closed linear subspace V of a Hilbert space $\mathcal{H}$, there is a decomposition

$$\mathcal{H} = V \oplus V^{\perp}.$$

In other words, any $x \in \mathcal{H}$ can be uniquely written as $x = v_x + v'_x$, where $v_x \in V$ and $v'_x \in V^{\perp}$. The *orthogonal projector* $\mathbb{P}$ from $\mathcal{H}$ to V is naturally defined by

$$\mathbb{P}x = v_x.$$

Clearly, $\mathbb{P}$ is a linear self-adjoint operator that maps $\mathcal{H}$ onto V and $\mathbb{P}^2 = \mathbb{P}$. Moreover, $x-\mathbb{P}x$ is orthogonal to $\mathbb{P}x$ and

$$\|x\|_{\mathcal{H}}^2 = \|\mathbb{P}x\|_{\mathcal{H}}^2 + \|x - \mathbb{P}x\|_{\mathcal{H}}^2,$$

which implies that $\mathbb{P}$ is a bounded linear operator with operator norm no more than 1. With this in mind, we introduce the following projectors $\{\mathbb{P}_k\}_{k\in\mathbb{Z}}$ and $\{\mathbb{Q}_k\}_{k\in\mathbb{Z}}$.

Definition 2.2. Let $\{V_k\}_{k\in\mathbb{Z}}$ be an MRA in $L^2(\mathcal{X})$ as in Definition 2.1. For any $k \in \mathbb{Z}$, let W_k be the orthogonal complement of V_k in V_{k+1}, that is,

$$W_k := V_{k+1} \ominus V_k. \tag{2.1}$$

For any $k \in \mathbb{Z}$, define $\mathbb{P}_k$ to be the *orthogonal projector* from $L^2(\mathcal{X})$ to the closed linear space V_k, and define $\mathbb{Q}_k$ to be the *orthogonal projector* from $L^2(\mathcal{X})$ to W_k. [1]

Based on the discussions in Remark 3.3 below, if $\{V_k\}_{k\in\mathbb{Z}}$ is an MRA in $L^2(\mathcal{X})$, then the projectors $\mathbb{P}_k$ and $\mathbb{Q}_k$ in Definition 2.2 have integral kernels $\mathbb{P}_k(x,y)$ and $\mathbb{Q}_k(x,y)$, respectively. Thus, it makes sense to introduce the following definitions of admissible MRA and admissible spectrum.

Definition 2.3. A multiresolution analysis $\{V_k\}_{k\in\mathbb{Z}}$ in $L^2(\mathcal{X})$ is called an *admissible MRA* if the family of projectors $\{\mathbb{P}_k\}_{k\in\mathbb{Z}}$, with each $\mathbb{P}_k$ being an orthogonal projector from $L^2(\mathcal{X})$ to V_k, satisfy the following properties:

(A1) there exists a positive constant $C > 0$ such that for any $k \in \mathbb{Z}$ and $x \in \mathcal{X}$,

$$\int_{\mathcal{X}} |\mathbb{P}_k(x,y)|\, d\mu(y) \le C;$$

(A2) for any $k \in \mathbb{Z}$ and $x \in \mathcal{X}$,

$$\int_{\mathcal{X}} \mathbb{P}_k(x,y)\, d\mu(y) = 1;$$

(A3) when $\mu(\mathcal{X}) = \infty$, assume further that for any $x, y \in \mathcal{X}$,

$$\lim_{k\to-\infty} \mathbb{P}_k(x,y) = 0.$$

Definition 2.4. Let $\{V_k\}_{k\in\mathbb{Z}}$ be an MRA in $L^2(\mathcal{X})$ and $k_0 \in \mathbb{Z}$ be the integer given in Definition 2.1(iii). The sequence $\{\lambda_k\}_{k\in\mathbb{Z}}$ is called a family of *admissible spectrum* if the following hold:

[1] Under the case $\mu(\mathcal{X}) < \infty$, we derive from (i) and (iii) of Definition 2.1 that $V_k = V_{k_0}$ for all $k \le k_0$, thereby leading to that $W_k = \{0\}$ and $\mathbb{Q}_k = 0$ whenever $k \le k_0 - 1$.

(a) $\{\lambda_k\}_{k\in\mathbb{Z}} \subset [0,\infty)$ is increasing;

(b) if $\mu(\mathcal{X}) = \infty$, then $\lambda_k \to 0$ as $k \to -\infty$; if $\mu(\mathcal{X}) < \infty$, then $\lambda_k = 0$ for all $k < k_0$;

(c) for any $t \in (0,\infty)$, $\sum_{k=N}^{\infty} e^{-t\lambda_k} \to 0$ as $N \to +\infty$.

Example 2.5. *In this paper, we mainly focus on the following family of spectrum. Let $\delta \in (0,1)$. If $\mu(\mathcal{X}) = \infty$, then we set*

$$\lambda_k := \delta^{-k\beta} \quad \textit{for any } k \in \mathbb{Z}. \tag{2.2}$$

If $\mu(\mathcal{X}) < \infty$, then we set

$$\lambda_k := \begin{cases} \delta^{-k\beta} & \textit{as } \ k \geq k_0; \\ 0 & \textit{as } \ k < k_0. \end{cases} \tag{2.3}$$

No matter $\mu(\mathcal{X})$ is finite or not, one easily verifies that $\{\lambda_k\}_{k\in\mathbb{Z}}$ is admissible.

The first main result of this paper is as follows.

Theorem 2.6. *Let $\{V_k\}_{k\in\mathbb{Z}}$ be an admissible* MRA *in $L^2(\mathcal{X})$ and $\{\lambda_k\}_{k\in\mathbb{Z}}$ be a family of admissible spectrum. For any $t \in (0,\infty)$ and $x, y \in \mathcal{X}$, define*

$$p_t(x,y) := \begin{cases} \displaystyle\sum_{k\in\mathbb{Z}} e^{-t\lambda_k} \mathbb{Q}_k(x,y) & \textit{as } \ \mu(\mathcal{X}) = \infty; \\ \displaystyle e^{-t\lambda_{k_0-1}} \mathbb{P}_{k_0}(x,y) + \sum_{k=k_0}^{\infty} e^{-t\lambda_k} \mathbb{Q}_k(x,y) & \textit{as } \ \mu(\mathcal{X}) < \infty. \end{cases} \tag{2.4}$$

Then, the family $\{p_t\}_{t>0}$ is a stochastic complete signed heat kernel.

We do not know if $\{p_t\}_{t>0}$ in Theorem 2.6 is positive or not. To gain its positivity, we usually need to add more conditions on MRA.

Corollary 2.7. *In addition to the assumptions in Theorem 2.6, assume that for any $k \in \mathbb{Z}$ and $x, y \in \mathcal{X}$,*

$$\mathbb{P}_k(x,y) \geq 0.$$

Then, $\{p_t\}_{t>0}$ in (2.4) is a stochastic complete heat kernel.

The proofs of Theorem 2.6 and Corollary 2.7 are presented in Section 3 (see the much stronger results in Theorems 3.8-3.9-3.10).

As a concluding remark of this subsection, we mention that Theorem 2.6 indicates a fundamental fact that the signed heat kernel and the MRA structure are in some sense equivalent to each other (see, e.g. [19] for the converse direction that heat kernels imply MRA structures).

2.2 Construction of heat kernels via non-smooth MRA

Given a metric measure space $(\mathcal{X}, d, \mu)$ satisfying (VD), there exists a family $\mathcal{D}$ of dyadic cubes (see [39, 37, 2] or Subsection 4.1 below). For any $k \in \mathbb{Z}$, denote by $\mathcal{D}_k$ the family of dyadic cubes in the k-th generation. The diameter of each dyadic cube Q in $\mathcal{D}_k$ is comparable to δ^k, where $\delta \in (0,1)$ is a small enough and fixed number.

With these concepts, we state the second main result of this paper, which serves as an application of Theorem 2.6 and Corollary 2.7.

Theorem 2.8. *Let $(\mathcal{X}, d, \mu)$ be a metric measure space satisfying (VD). For any $k \in \mathbb{Z}$, define*

$$V_k := \overline{\operatorname{span}\left\{\mu(Q)^{-\frac{1}{2}}\mathbf{1}_Q : \ Q \in \mathcal{D}_k\right\}}^{\|\cdot\|_{L^2(\mathcal{X})}}. \tag{2.5}$$

For any $\beta \in (0, \infty)$, let $\{\lambda_k\}_{k\in\mathbb{Z}}$ be the admissible spectrum as in Example 2.5. Then, the sequence $\{V_k\}_{k\in\mathbb{Z}}$ in (2.5) forms an admissible MRA and the heat kernel $\{p_t\}_{t>0}$ in (2.4) coincides exactly with $\{p_t^{\mathcal{D}}\}_{t>0}$ in (1.7). In particular, $\{p_t\}_{t>0}$ is a stochastic complete heat kernel satisfying

$$0 \le p_t(x, y) = p_t^{\mathcal{D}}(x, y) \le \frac{C}{V(x,\ t^{1/\beta} + d(x, y))}\left(\frac{t^{1/\beta}}{t^{1/\beta} + d(x, y)}\right)^{\beta} \tag{2.6}$$

where C is a positive constant independent of $t \in (0, \infty)$ and $x, y \in \mathcal{X}$.

The proof of Theorem 2.8 is presented in Subsection 4.2 below. The preliminary step is show that $\{V_k\}_{k\in\mathbb{Z}}$ given in (2.5) is an admissible MRA so that Corollary 2.7 can be applied to produce a stochastic complete heat kernel. The subsequence argument is to verify that the two heat kernels in (2.4) and (1.7) coincides with each other. Then the upper estimate (2.6) follows directly from the already known estimates in (1.8) and (1.9).

Note that $\{V_k\}_{k\in\mathbb{Z}}$ in (2.5) is called a non-smooth MRA because it is formed by the Haar wavelets (see Remark 4.9 below) that are not continuous on $\mathcal{X}$.

2.3 Construction of signed heat kernels via smooth MRA

In this subsection, we give another application of Theorem 2.6. Again let $(\mathcal{X}, d, \mu)$ be a metric measure space satisfying (VD). Auscher and Hytönen [2] constructed a smooth MRA constituting of smooth *wavelets/splines* (see Subsection 5.2 below). Using this smooth MRA and the admissible spectrum given in Example 2.5, we deduce from Theorem 2.6 a signed heat kernel satisfying stable-like upper estimate, almost Lipschitz-continuity and near-diagonal estimate.

Theorem 2.9. *Let $(\mathcal{X}, d, \mu)$ be a metric measure space satisfying (VD). For any $k \in \mathbb{Z}$, define*

$$V_k := \overline{\operatorname{span}\left\{s_\alpha^k : \ \alpha \in \mathcal{A}_k\right\}}^{\|\cdot\|_{L^2(\mathcal{X})}}, \tag{2.7}$$

where every s_α^k is a spline function located near the dyadic cube Q_α^k and $\mathcal{A}_k$ is a countable index set. For any $\beta \in (0, \infty)$, let $\{\lambda_k\}_{k\in\mathbb{Z}}$ be the admissible spectrum as in Example 2.5. Then, the sequence $\{V_k\}_{k\in\mathbb{Z}}$ in (2.7) forms an admissible MRA and $\{p_t\}_{t>0}$ in (2.4) is a stochastic complete signed heat kernel enjoying the following properties:

(i) *(upper stable-like estimate) there exists a positive constant C such that for any $t \in (0, \infty)$ and $x, y \in \mathcal{X}$,*

$$|p_t(x, y)| \le \frac{C}{V(x,\ t^{1/\beta} + d(x, y))}\left(\frac{t^{1/\beta}}{t^{1/\beta} + d(x, y)}\right)^{\beta};$$

(ii) *(near-diagonal lower estimate) there exist a positive constant c and a small constant $\varepsilon \in (0, 1)$ such that for any $t \in (0, \infty)$ and any $x, y \in \mathcal{X}$ satisfying $d(x, y) \le \varepsilon t^{1/\beta}$,*

$$p_t(x, y) \ge \frac{C}{V(x, t^{1/\beta})};$$

(iii) *(Hölder regularity) there exists a positive constant C such that for any $t > 0$ and x, y, $y' \in \mathcal{X}$ satisfying $d(y, y') \le t^{1/\beta}$,*

$$\left|p_t(x,y) - p_t(x,y')\right| \le C\left(\frac{d(y,y')}{t^{1/\beta}}\right)^{\eta_\delta} \frac{1}{V(x,\, t^{1/\beta} + d(x,y))}\left(\frac{t^{1/\beta}}{t^{1/\beta} + d(x,y)}\right)^{\beta},$$

where $\eta_\delta = 1 - \frac{\log(84M^8)}{\log(1/\delta)}$ tends to 1 as $\delta \to 0$, and $\delta \in (0, 1/(84M^8))$ is the parameter appeared in the construction of dyadic cubes on $\mathcal{X}$ with the integer M being taken as in (1.2).

The proof of Theorem 2.9 is given in Section 5 below (see Theorems 5.6-5.7-5.8-5.9 below). The key ingredients of the proof are as follows. For the sequence $\{V_k\}_{k\in\mathbb{Z}}$ in (2.7), it is an MRA (see Lemma 5.2 below) and the integral kernels of the projectors

$$\mathbb{P}_k : L^2(\mathcal{X}) \to V_k \quad \text{and} \quad \mathbb{Q}_k : L^2(\mathcal{X}) \to W_k = V_{k+1} \ominus V_k$$

have exponential decay off the diagonal and almost Lipschitz continuity (see Lemma 5.3 below). Such properties of $\mathbb{P}_k$ and $\mathbb{Q}_k$ are proved by Auscher and Hytönen [2]. This will not only ensure that $\{V_k\}_{k\in\mathbb{Z}}$ in (2.7) to be an admissible MRA so that Theorem 2.6 can be applied, but also yield the desired estimates in (i)-(ii)-(iii) of Theorem 2.9.

Remark 2.10. Below we give several comments on Theorems 2.8 and 2.9.

(i) Let us remark that in [41, 2] both the Haar wavelets and smooth wavelets are constructed in a quasi-metric measure space satisfying (VD). Due to this reason, in Theorems 2.8 and 2.9, the assumption of d being a metric can be relaxed to a *quasi-metric*, that is, the triangle inequality of d is replaced by the following:

$$d(x,y) \le K\left(d(x,z) + d(z,y)\right),$$

where $K \in [1,\infty)$ is a constant independent of $x, y, z \in \mathcal{X}$. If d is a quasi-metric, then through minor modifications of the current proof of Theorems 4.10-5.7-5.8-5.9, we find that all conclusions of Theorems 2.8-2.9 remain true, except that the Hölder exponent η_δ in Theorem 2.9(iii) will depend on the constant K and it can not approach 1 when $\delta \to 0$.

(ii) A novelty of the method of heat kernel construction in Theorem 2.9 is that, without using the Dirichlet form theory as in [15, 16, 34, 12], we still produce a signed heat kernel that satisfies a number of good properties such as the stable-like upper estimate, near-diagonal lower estimate, stochastic completeness and almost Lipschitz regularity estimate. The price we pay is that for the constructed $\{p_t\}_{t>0}$ there may exist $t \in (0,\infty)$ and $x, y \in \mathcal{X}$ such that

$$p_t(x,y) < 0.$$

But this disadvantage can be neglected in many situations such as analysis of differential operators (see, e.g. [26, 27, 21]), etc.

(iii) For the heat kernel $\{p_t\}_{t>0}$ constructed in Theorem 2.9, it is natural to ask that if one can show the positivity of $p_t(x,y)$ and then obtain a stable-like two-sided estimate $(\mathrm{ULE})_\beta$. Unfortunately, the answer in general is no. Indeed, although each $p_t(x,y)$ has a near-diagonal lower estimate, it may be negative when x and y are far away from each other. The latter can be seen from the representations (3.21) and (3.22) below, together with the fact that $\mathbb{P}_k(x,y)$ may be negative off the diagonal (see Remark 5.4 below for more details). As was pointed out earlier, heat kernels with negative value arise in various scenarios such as integral kernels of semigroups generated by higher order differential operators (see, e.g. [24, 25]).

2.4 Organization of the paper

This paper is organized as follows.

The main aim of Section 3 is to show Theorem 2.6 and Corollary 2.7. Let us be more precise. In Subsection 3.1, we establish some basic properties of the projectors $\{\mathbb{P}_k\}_{k\in\mathbb{Z}}$ and $\{\mathbb{Q}_k\}_{k\in\mathbb{Z}}$ that are defined in Definition 2.2. In Subsection 3.2, we show that (see Theorem 3.7 below) the projectors $\{\mathbb{P}_k\}_{k\in\mathbb{Z}}$ and $\{\mathbb{Q}_k\}_{k\in\mathbb{Z}}$ induce a strongly continuous contractive semigroup $\{P_t\}_{t\geq 0}$ on $L^2(\mathcal{X})$, whose integral kernels are precisely $\{p_t\}_{t>0}$ defined in (2.4). Then, in Subsection 3.3, we show Theorem 2.6 (see Theorems 3.8 and 3.9 below) and a stronger version of Corollary 2.7 (see Theorem 3.10 below). Moreover, in Subsection 3.4, we study the spectrum and functional calculus of the generator of the heat kernel $\{p_t\}_{t>0}$ in (2.4).

Section 4 is devoted to the proof of Theorem 2.8. In Subsection 4.1, we recall the construction of dyadic cubes (see [17, 37]) on a metric measure space satisfying (VD). In Subsection 4.2, we show that $\{V_k\}_{k\in\mathbb{Z}}$ given in (2.5) is an admissible MRA so that Corollary 2.7 can be applied to produce a stochastic complete heat kernel $\{p_t\}_{t>0}$ as given in (2.4). In this case, this heat kernel $\{p_t\}_{t>0}$ coincides exactly with $\{p_t^{\mathcal{D}}\}_{t>0}$ in (1.7). In Subsection 4.3, we give a direct proof of Theorem 2.8. Subsection 4.4 consists of several examples of heat kernels that are constructed by Haar wavelets on various underlying spaces.

Section 5 is devoted to the proof of Theorem 2.9. In Subsection 5.1, we recall the construction of random dyadic cubes on a metric measure space satisfying (VD) (see [2, 39]). Next, in Subsection 5.2, we review Auscher-Hytönen's construction (see [2]) of MRA generated by smooth splines/wavelets, and then we show that such smooth MRA is admissible (see Theorem 5.5 below) in the sense of Definition 2.3. In Subsection 5.3, using this admissible smooth MRA, we follow (2.4) and construct a stochastic complete signed heat kernel $\{p_t\}_{t>0}$ (see (5.1) and Theorem 5.6 below). Further, we prove in Subsection 5.4 that such $\{p_t\}_{t>0}$ satisfies the stable-like upper estimate (see Theorem 5.7). Moreover, in Subsection 5.5, we validate the almost Lipschitz regularity estimate (see Theorem 5.8 below) and the near-diagonal lower estimate (see Theorem 5.9 below). Altogether, we obtain Theorem 2.9.

Notation. Let $\mathbb{N} = \{0, 1, \ldots, \}$, $\mathbb{Z} = \{0, \pm 1, \pm 2, \ldots\}$ and $\mathrm{diam}(\mathcal{X}) := \sup\{d(x, y) : \ x, y \in \mathcal{X}\}$. For any set $E \subset \mathcal{X}$, $\overline{E}$ denotes the closure of E, and $E^{\complement} = \mathcal{X} \setminus E$. The letters C and c are used to denote positive constants that are independent of the variables in question, but may vary at each occurrence. The relation $u \lesssim v$ (resp., $u \gtrsim v$) between functions u and v means that $u \leq Cv$ (resp., $u \geq Cv$) for a positive constant C. We write $u \simeq v$ if $u \lesssim v \lesssim u$.

3 Construction of heat kernels via general MRA

Throughout this section, we always assume that $(\mathcal{X}, d, \mu)$ is a metric measure space endowing with a multiresolution analysis structure $\{V_k\}_{k\in\mathbb{Z}}$ as in Definition 2.1. The conditions (VD) and (RVD) are not used throughout this section.

3.1 Preliminaries on MRA

The following facts of Riesz basis come from [18, Theorems 3.6.2 and 3.6.6, Corollary 3.6.3].

Lemma 3.1. *Let $\{V_k\}_{k\in\mathbb{Z}}$ be an* MRA *in $L^2(\mathcal{X})$. For any $k \in \mathbb{Z}$, if $\{\varphi_{k,\alpha}\}_{\alpha\in\mathcal{A}_k}$ is a Riesz basis of V_k, then there exists a unique dual Riesz basis $\{\widetilde{\varphi}_{k,\alpha}\}_{\alpha\in\mathcal{A}_k}$ of V_k, that is, $\{\widetilde{\varphi}_{k,\alpha}\}_{\alpha\in\mathcal{A}_k}$ is the unique Riesz basis in V_k satisfying that for all $\alpha, \beta \in \mathcal{A}_k$,*

$$\langle \widetilde{\varphi}_{k,\alpha}, \varphi_{k,\beta} \rangle = \delta_{\alpha,\beta} := \begin{cases} 1 & \text{when } \alpha = \beta; \\ 0 & \text{when } \alpha \neq \beta, \end{cases}$$

where $\langle\cdot,\,\cdot\rangle$ *denotes the inner product of* $L^2(\mathcal{X})$. *Moreover, for any* $f\in V_k$,

$$f=\sum_{\alpha\in\mathcal{A}_k}\langle\widetilde{\varphi}_{k,\alpha},\,f\rangle\,\varphi_{k,\alpha}=\sum_{\alpha\in\mathcal{A}_k}\langle\varphi_{k,\alpha},\,f\rangle\,\widetilde{\varphi}_{k,\alpha}\quad\textit{in}\ \ L^2(\mathcal{X}).\tag{3.1}$$

Next, we collect some basic properties of the orthogonal projectors $\{\mathbb{P}_k\}_{k\in\mathbb{Z}}$ and $\{\mathbb{Q}_k\}_{k\in\mathbb{Z}}$ that are defined in Definition 2.2.

Lemma 3.2. *Let* $\{\mathbb{P}_k\}_{k\in\mathbb{Z}}$ *and* $\{\mathbb{Q}_k\}_{k\in\mathbb{Z}}$ *be the orthogonal projectors defined in Definition 2.2. Then, for any* $k\in\mathbb{Z}$, *the following hold:*

(i) *both* $\mathbb{P}_k$ *and* $\mathbb{Q}_k$ *are self-adjoint linear bounded operators on* $L^2(\mathcal{X})$, $\mathbb{P}_k^2=\mathbb{P}_k$ *and* $\mathbb{Q}_k^2=\mathbb{Q}_k$;

(ii) *if* $j\geq k$, *then* $\mathbb{P}_k\mathbb{P}_j=\mathbb{P}_j\mathbb{P}_k=\mathbb{P}_k$ *and* $\mathbb{Q}_j\mathbb{P}_k=0$;

(iii) *if* $j\neq k$, *then* $\mathbb{Q}_j\mathbb{Q}_k=0$;

(iv) *if* V_k *has a Riesz basis* $\{\varphi_{k,\alpha}\}_{\alpha\in\mathcal{A}_k}$ *and a dual Riesz basis* $\{\widetilde{\varphi}_{k,\alpha}\}_{\alpha\in\mathcal{A}_k}$, *then for any* $f\in L^2(\mathcal{X})$,

$$\mathbb{P}_kf=\sum_{\alpha\in\mathcal{A}_k}\langle\widetilde{\varphi}_{k,\alpha},\,f\rangle\,\varphi_{k,\alpha}=\sum_{\alpha\in\mathcal{A}_k}\langle\varphi_{k,\alpha},\,f\rangle\,\widetilde{\varphi}_{k,\alpha}.$$

Proof. Based on the basic facts of orthogonal projections in Hilbert spaces (see [57, pp. 81-84]), we easily obtain (i).

Now, we prove (ii). For any $j\geq k$ and $f\in L^2(\mathcal{X})$, we have by Definition 2.1(i) that $\mathbb{P}_kf\in V_k\subset V_j$, which implies

$$\mathbb{P}_j\mathbb{P}_kf=\mathbb{P}_kf.$$

By this, together with the self-adjoint properties of $\mathbb{P}_k$ and $\mathbb{P}_j$, we see that for any $f,g\in L^2(\mathcal{X})$,

$$\langle\mathbb{P}_k\mathbb{P}_jg,\,f\rangle=\langle g,\,\mathbb{P}_j\mathbb{P}_kf\rangle=\langle g,\,\mathbb{P}_kf\rangle=\langle\mathbb{P}_kg,\,f\rangle,$$

whence

$$\mathbb{P}_k\mathbb{P}_jg=\mathbb{P}_kg.$$

As a consequence, it holds that

$$\mathbb{Q}_j\mathbb{P}_k=(\mathbb{P}_{j+1}-\mathbb{P}_j)\mathbb{P}_k=\mathbb{P}_k-\mathbb{P}_k=0.$$

Thus, we obtain (ii).

To show (iii), by symmetry, we assume without loss of generality that $k<j$. Then the second equality in (ii) implies

$$\mathbb{Q}_j\mathbb{Q}_k=\mathbb{Q}_j\mathbb{P}_{k+1}-\mathbb{Q}_j\mathbb{P}_k=0,$$

as desired.

Finally, we verify (iv). If $f\in L^2(\mathcal{X})$, then $\mathbb{P}_kf\in V_k$ and it can be represented as in (3.1). By this and the self-adjoint property of $\mathbb{P}_k$, together with $\mathbb{P}_k\widetilde{\varphi}_{k,\alpha}=\widetilde{\varphi}_{k,\alpha}$ and $\mathbb{P}_k\varphi_{k,\alpha}=\varphi_{k,\alpha}$, we obtain

$$\mathbb{P}_kf=\sum_{\alpha\in\mathcal{A}_k}\langle\widetilde{\varphi}_{k,\alpha},\,\mathbb{P}_kf\rangle\,\varphi_{k,\alpha}=\sum_{\alpha\in\mathcal{A}_k}\langle\mathbb{P}_k\widetilde{\varphi}_{k,\alpha},\,f\rangle\,\varphi_{k,\alpha}=\sum_{\alpha\in\mathcal{A}_k}\langle\widetilde{\varphi}_{k,\alpha},\,f\rangle\,\varphi_{k,\alpha}$$

and

$$\mathbb{P}_kf=\sum_{\alpha\in\mathcal{A}_k}\langle\varphi_{k,\alpha},\,\mathbb{P}_kf\rangle\,\widetilde{\varphi}_{k,\alpha}=\sum_{\alpha\in\mathcal{A}_k}\langle\mathbb{P}_k\varphi_{k,\alpha},\,f\rangle\,\widetilde{\varphi}_{k,\alpha}=\sum_{\alpha\in\mathcal{A}_k}\langle\varphi_{k,\alpha},\,f\rangle\,\widetilde{\varphi}_{k,\alpha},$$

which proves (iv). $\square$

Remark 3.3. By Definition 2.1 and Lemma 3.1, if $\{V_k\}_{k\in\mathbb{Z}}$ is an MRA in $L^2(\mathcal{X})$, then each V_k has a Riesz basis $\{\varphi_{k,\alpha}\}_{\alpha\in\mathcal{A}_k}$ and a dual Riesz basis $\{\widetilde{\varphi}_{k,\alpha}\}_{\alpha\in\mathcal{A}_k}$. Thus, if $\mathbb{P}_k$ is an orthogonal projector from $L^2(\mathcal{X})$ to V_k, then $\mathbb{P}_k$ is associated to an integral kernel in the following way:

$$\mathbb{P}_k f(x) = \int_{\mathcal{X}} \mathbb{P}_k(x,y) f(y)\,d\mu(y) \qquad \text{for any } \ f\in L^2(\mathcal{X}) \ \text{ and } \ x\in\mathcal{X},$$

where (see Lemma 3.2(iv))

$$\mathbb{P}_k(x,y) = \sum_{\alpha\in\mathcal{A}_k} \varphi_{k,\alpha}(x)\widetilde{\varphi}_{k,\alpha}(y) \qquad \text{for all } k\in\mathbb{Z} \ \text{ and } \ x,y\in\mathcal{X}.$$

For the orthogonal projector $\mathbb{Q}_k$ which maps $L^2(\mathcal{X})$ to $W_k := V_{k+1}\ominus V_k$, since $\mathbb{P}_{k+1}f-\mathbb{P}_k f\in V_{k+1}$ and $\mathbb{P}_{k+1}f-\mathbb{P}_k f\perp V_k$ hold for all $f\in L^2(\mathcal{X})$, it follows that

$$\mathbb{Q}_k = \mathbb{P}_{k+1} - \mathbb{P}_k. \tag{3.2}$$

Consequently, each $\mathbb{Q}_k$ has also an integral kernel $\mathbb{Q}_k(x,y)$ such that

$$\mathbb{Q}_k f(x) = \int_{\mathcal{X}} \mathbb{Q}_k(x,y) f(y)\,d\mu(y) \qquad \text{for all } f\in L^2(\mathcal{X}) \ \text{ and } \ x\in\mathcal{X}$$

and

$$\mathbb{Q}_k(x,y) = \mathbb{P}_{k+1}(x,y) - \mathbb{P}_k(x,y) \qquad \text{for all } k\in\mathbb{Z} \ \text{ and } \ x,y\in\mathcal{X}.$$

Remark 3.4. Applying (ii) and (iii) of Lemma 3.2, together with the self-adjoint property of $\mathbb{Q}_k$, we deduce that for all $f,g\in L^2(\mathcal{X})$,

$$\langle \mathbb{P}_k f,\ \mathbb{Q}_j g\rangle = \langle \mathbb{Q}_j\mathbb{P}_k f,\ g\rangle = 0 \quad \text{as } \ k\le j; \tag{3.3}$$

and

$$\langle \mathbb{Q}_k f,\ \mathbb{Q}_j g\rangle = \langle \mathbb{Q}_j\mathbb{Q}_k f,\ g\rangle = 0 \quad \text{as } \ k\ne j. \tag{3.4}$$

Lemma 3.5. *For the orthogonal projectors $\{\mathbb{P}_k\}_{k\in\mathbb{Z}}$ and $\{\mathbb{Q}_k\}_{k\in\mathbb{Z}}$ that are defined in Definition 2.2, the following hold:*

(i) *no matter $\mu(\mathcal{X})$ is finite or not, $\lim_{k\to+\infty}\mathbb{P}_k = id$ in $L^2(\mathcal{X})$;*

(ii) *if $\mu(\mathcal{X})=\infty$, then $\lim_{k\to-\infty}\mathbb{P}_k = 0$ in $L^2(\mathcal{X})$;*

(iii) *for any $f\in L^2(\mathcal{X})$, we have*

$$f = \begin{cases} \displaystyle\sum_{k\in\mathbb{Z}} \mathbb{Q}_k f & \text{as } \ \mu(\mathcal{X})=\infty; \\ \displaystyle\mathbb{P}_{k_0} f + \sum_{k=k_0}^{\infty} \mathbb{Q}_k f & \text{as } \ \mu(\mathcal{X})<\infty, \end{cases} \tag{3.5}$$

and

$$\|f\|^2_{L^2(\mathcal{X})} = \begin{cases} \displaystyle\sum_{k\in\mathbb{Z}} \|\mathbb{Q}_k f\|^2_{L^2(\mathcal{X})} & \text{as } \ \mu(\mathcal{X})=\infty; \\ \displaystyle\|\mathbb{P}_{k_0} f\|^2_{L^2(\mathcal{X})} + \sum_{k=k_0}^{\infty} \|\mathbb{Q}_k f\|^2_{L^2(\mathcal{X})} & \text{as } \ \mu(\mathcal{X})<\infty, \end{cases} \tag{3.6}$$

where the equality in (3.5) holds in $L^2(\mathcal{X})$, and k_0 is the integer given in Definition 2.1(iii).

Proof. First, we show (i). From Definition 2.1(i), it follows that $\{V_k^\perp\}_{k\in\mathbb{Z}}$ is decreasing. For any $f \in L^2(\mathcal{X})$ and $j \le k$, we know from $L^2(\mathcal{X}) = V_k \oplus V_k^\perp$ that

$$\mathbb{P}_k f - f \in V_k^\perp \subset V_j^\perp.$$

This implies that for all $f \in L^2(\mathcal{X})$ and $g \in \cup_{j\in\mathbb{Z}} V_j$,

$$\lim_{k\to+\infty} \langle \mathbb{P}_k f - f,\ g\rangle = 0.$$

Invoking Definition 2.1(ii) yields that $\lim_{k\to+\infty} \mathbb{P}_k f = f$ in $L^2(\mathcal{X})$. This proves (i).

Now, we assume that $\mu(\mathcal{X}) = \infty$ and show (ii). By the definition of W_k in (2.1) and Definition 2.1(ii), we see that $\overline{\cup_{k\in\mathbb{Z}} W_k} = L^2(\mathcal{X})$. Let $f \in L^2(\mathcal{X})$. If $k \le j$ and $g \in W_j$, then from $\mathbb{P}_k f \in V_k \subset V_j$ we derive

$$\langle \mathbb{P}_k f,\ g\rangle = 0.$$

Therefore, for all $j \in \mathbb{Z}$ and $g \in W_j$ we have

$$\lim_{k\to-\infty} \langle \mathbb{P}_k f,\ g\rangle = 0$$

This, together with the aforementioned fact $\overline{\cup_{k\in\mathbb{Z}} W_k} = L^2(\mathcal{X})$, yields

$$\lim_{k\to-\infty} \langle \mathbb{P}_k f,\ g\rangle = 0 \quad \text{for all } g \in L^2(\mathcal{X}).$$

So, we obtain that $\lim_{k\to-\infty} \mathbb{P}_k f = 0$ in $L^2(\mathcal{X})$.

It remains to show (iii). Let $f \in L^2(\mathcal{X})$. For any $m \le n$, applying $\mathbb{Q}_k = \mathbb{P}_{k+1} - \mathbb{P}_k$ in (3.2) gives

$$\sum_{k=m}^{n} \mathbb{Q}_k = \sum_{k=m}^{n} \mathbb{P}_{k+1} - \sum_{k=m}^{n} \mathbb{P}_k = \mathbb{P}_{n+1} - \mathbb{P}_m.$$

This, along with the just proved (i) and (ii), yields that when $\mu(\mathcal{X}) = \infty$,

$$\sum_{k\in\mathbb{Z}} \mathbb{Q}_k f = \lim_{N\to+\infty} \sum_{k=-N}^{N} \mathbb{Q}_k f = \lim_{N\to+\infty} (\mathbb{P}_{N+1} f - \mathbb{P}_{-N} f) = f \quad \text{in } L^2(\mathcal{X}).$$

Meanwhile, when $\mu(\mathcal{X}) < \infty$,

$$\mathbb{P}_{k_0} f + \sum_{k=k_0}^{\infty} \mathbb{Q}_k f = \mathbb{P}_{k_0} f + \lim_{N\to+\infty} \sum_{k=k_0}^{N} \mathbb{Q}_k f = \lim_{N\to+\infty} \mathbb{P}_{N+1} f = f \quad \text{in } L^2(\mathcal{X}).$$

Thus (3.5) holds in $L^2(\mathcal{X})$. By (3.5), $\mathbb{Q}_k^2 = \mathbb{Q}_k$ and $\mathbb{P}_{k_0}^2 = \mathbb{P}_{k_0}$, we derive that when $\mu(\mathcal{X}) < \infty$,

$$\begin{aligned} \|f\|^2_{L^2(\mathcal{X})} = \left\langle \mathbb{P}_{k_0} f + \sum_{k=k_0}^{\infty} \mathbb{Q}_k f,\ f \right\rangle &= \langle \mathbb{P}_{k_0} f,\ f\rangle + \sum_{k=k_0}^{\infty} \langle \mathbb{Q}_k f,\ f\rangle \\ &= \langle \mathbb{P}_{k_0}^2 f,\ f\rangle + \sum_{k=k_0}^{\infty} \left\langle \mathbb{Q}_k^2 f,\ f \right\rangle \end{aligned}$$

$$= \langle \mathbb{P}_{k_0} f,\ \mathbb{P}_{k_0} f\rangle + \sum_{k=k_0}^{\infty} \langle \mathbb{Q}_k f,\ \mathbb{Q}_k f\rangle$$
$$= \|\mathbb{P}_{k_0} f\|^2_{L^2(X)} + \sum_{k=k_0}^{\infty} \|\mathbb{Q}_k f\|^2_{L^2(X)}.$$

In a similar manner, when $\mu(X) = \infty$, we have

$$\|f\|^2_{L^2(X)} = \left\langle \sum_{k\in\mathbb{Z}} \mathbb{Q}_k f,\ f \right\rangle = \sum_{k\in\mathbb{Z}} \langle \mathbb{Q}_k f,\ f\rangle = \sum_{k\in\mathbb{Z}} \left\langle \mathbb{Q}_k^2 f,\ f \right\rangle = \sum_{k\in\mathbb{Z}} \langle \mathbb{Q}_k f,\ \mathbb{Q}_k f\rangle = \sum_{k\in\mathbb{Z}} \|\mathbb{Q}_k f\|^2_{L^2(X)}.$$

Combining the last two formulae yields (3.6). Altogether, we obtain (iii). □

3.2 Heat semigroups generated by MRA

The aim of this subsection is to show that the projectors $\{\mathbb{P}_k\}_{k\in\mathbb{Z}}$ and $\{\mathbb{Q}_k\}_{k\in\mathbb{Z}}$ that are defined in Definition 2.2 will induce a strongly continuous contractive semigroup $\{P_t\}_{t\geq 0}$ on $L^2(X)$. To achieve this, we begin with the definition of the operators $\{P_t\}_{t\geq 0}$.

Definition 3.6. Given any sequence $\{\lambda_k\}_{k\in\mathbb{Z}} \subset [0,\infty)$, define the family of linear operators $\{P_t\}_{t\geq 0}$ as follows: for any $t \in [0,\infty)$ and $f \in L^2(X)$,

$$P_t f := \begin{cases} \displaystyle\sum_{k\in\mathbb{Z}} e^{-t\lambda_k} \mathbb{Q}_k f & \text{as } \mu(X) = \infty; \\ \displaystyle e^{-t\lambda_{k_0-1}} \mathbb{P}_{k_0} f + \sum_{k=k_0}^{\infty} e^{-t\lambda_k} \mathbb{Q}_k f & \text{as } \mu(X) < \infty, \end{cases} \tag{3.7}$$

where k_0 is the integer given in Definition 2.1(iii), $\{\mathbb{P}_k\}_{k\in\mathbb{Z}}$ and $\{\mathbb{Q}_k\}_{k\in\mathbb{Z}}$ are defined as in Definition 2.2.

Intuitionally, the construction in Definition 3.6 views each subspace W_k (see Definition 2.2) as the eigenspace of the eigenvalue λ_k. Invoking Remark 3.3, we find that $\{p_t\}_{t>0}$ defined in (2.4) are just the integral kernels of the operators $\{P_t\}_{t\geq 0}$ in (3.7).

Theorem 3.7. *The family $\{P_t\}_{t\geq 0}$ in (3.7) is a strongly continuous contractive semigroup on $L^2(X)$, that is, it satisfies the following properties:*

(i) *each P_t is a symmetric linear operator with domain $L^2(X)$;*

(ii) *(contraction property) for any $t \in (0,\infty)$ and $f \in L^2(X)$, $\|P_t f\|_{L^2(X)} \leq \|f\|_{L^2(X)}$;*

(iii) *(semigroup property) for any $t, s \in (0,\infty)$, $P_t P_s = P_{t+s}$;*

(iv) *(strongly continuous property) for any $f \in L^2(X)$, $\lim_{t\downarrow 0} P_t f = f$ in $L^2(X)$.*

Proof. We will prove items (i) through (iv) in the following 4 steps.

Step 1: proof of (i). When $t = 0$, it follows from (3.5) that $P_0 = id$ in $L^2(X)$. Since $\{\mathbb{P}_k\}_{k\in\mathbb{Z}}$ and $\{\mathbb{Q}_k\}_{k\in\mathbb{Z}}$ are symmetric linear operators, so does P_t for each $t \in (0,\infty)$.

Suppose that $t \in (0,\infty)$ and $f \in L^2(X)$. For any $m \leq n$, applying (3.4) yields

$$\left\| \sum_{k=m}^{n} e^{-t\lambda_k} \mathbb{Q}_k f \right\|^2_{L^2(X)} = \sum_{k=m}^{n} \sum_{j=m}^{n} e^{-t\lambda_k} e^{-t\lambda_j} \left\langle \mathbb{Q}_k f,\ \mathbb{Q}_j f \right\rangle$$

$$= \sum_{k=m}^{n} e^{-2t\lambda_k} \langle \mathbb{Q}_k f,\ \mathbb{Q}_k f\rangle$$

$$\leq \sum_{k=m}^{n} \|\mathbb{Q}_k f\|^2_{L^2(\mathcal{X})}. \tag{3.8}$$

By (3.6), (3.8) and $f \in L^2(\mathcal{X})$, we see that both $\{\sum_{k=-n}^{n} e^{-t\lambda_k}\mathbb{Q}_k f\}_{n\in\mathbb{N}}$ (when $\mu(\mathcal{X}) = \infty$) and $\{\sum_{k=k_0}^{n} e^{-t\lambda_k}\mathbb{Q}_k f\}_{n\in\mathbb{N}}$ (when $\mu(\mathcal{X}) < \infty$) are Cauchy sequences in $L^2(\mathcal{X})$. Thus, for any $f \in L^2(\mathcal{X})$, each $P_t f$ in (3.7) is a well-defined $L^2(\mathcal{X})$-function.

Step 2: proof of (ii). Fix $t \in (0, \infty)$ and $f \in L^2(\mathcal{X})$. As was proved in **Step 1**, the series of defining $P_t f$ converges in $L^2(\mathcal{X})$. Thus, when $\mu(\mathcal{X}) = \infty$, applying (3.4) and (3.6) yields

$$\begin{aligned}
\|P_t f\|^2_{L^2(\mathcal{X})} &= \left\langle \sum_{k\in\mathbb{Z}} e^{-t\lambda_k}\mathbb{Q}_k f,\ \sum_{j\in\mathbb{Z}} e^{-t\lambda_j}\mathbb{Q}_j f\right\rangle \\
&= \sum_{k\in\mathbb{Z}}\sum_{j\in\mathbb{Z}} e^{-t\lambda_k}e^{-t\lambda_j}\left\langle \mathbb{Q}_k f,\ \mathbb{Q}_j f\right\rangle \\
&= \sum_{k\in\mathbb{Z}} e^{-2t\lambda_k}\langle \mathbb{Q}_k f,\ \mathbb{Q}_k f\rangle \\
&\leq \sum_{k\in\mathbb{Z}} \|\mathbb{Q}_k(f)\|^2_{L^2(\mathcal{X})} \\
&= \|f\|^2_{L^2(\mathcal{X})}.
\end{aligned} \tag{3.9}$$

In a similar manner, under the case $\mu(\mathcal{X}) < \infty$, we derive from (3.3), (3.4) and (3.6) that

$$\begin{aligned}
\|P_t f\|^2_{L^2(\mathcal{X})} &= \left\langle e^{-t\lambda_{k_0-1}}\mathbb{P}_{k_0} f + \sum_{k=k_0}^{\infty} e^{-t\lambda_k}\mathbb{Q}_k f,\ e^{-t\lambda_{k_0-1}}\mathbb{P}_{k_0} f + \sum_{j=k_0}^{\infty} e^{-t\lambda_j}\mathbb{Q}_j f\right\rangle \\
&= e^{-2t\lambda_{k_0-1}}\langle \mathbb{P}_{k_0} f,\ \mathbb{P}_{k_0} f\rangle + e^{-t\lambda_{k_0-1}}\sum_{j=k_0}^{\infty} e^{-t\lambda_j}\left\langle \mathbb{P}_{k_0} f,\ \mathbb{Q}_j f\right\rangle \\
&\quad + e^{-t\lambda_{k_0-1}}\sum_{k=k_0}^{\infty} e^{-t\lambda_k}\langle \mathbb{Q}_k f,\ \mathbb{P}_{k_0} f\rangle + \sum_{k=k_0}^{\infty}\sum_{j=k_0}^{\infty} e^{-t\lambda_k}e^{-t\lambda_j}\left\langle \mathbb{Q}_k f,\ \mathbb{Q}_j f\right\rangle \\
&= e^{-2t\lambda_{k_0-1}}\langle \mathbb{P}_{k_0} f,\ \mathbb{P}_{k_0} f\rangle + \sum_{k=k_0}^{\infty} e^{-2t\lambda_k}\langle \mathbb{Q}_k f,\ \mathbb{Q}_k f\rangle \\
&\leq \|\mathbb{P}_{k_0}(f)\|^2_{L^2(\mathcal{X})} + \sum_{k=k_0}^{\infty} \|\mathbb{Q}_k(f)\|^2_{L^2(\mathcal{X})} \\
&= \|f\|^2_{L^2(\mathcal{X})},
\end{aligned} \tag{3.10}$$

where in the second equality we used the convergence of $\sum_{k=k_0}^{\infty} e^{-t\lambda_k}\mathbb{Q}_k f$ in $L^2(\mathcal{X})$. Thus, $\{P_t\}_{t\geq 0}$ is contractive.

Step 3: proof of (iii). Let $f \in L^2(\mathcal{X})$ and $s, t \in (0, \infty)$. Again, no matter $\mu(\mathcal{X})$ is finite or not, we know from **Step 1** that the series in defining $P_s f$ converges in $L^2(\mathcal{X})$. From this and the boundedness of

P_t on $L^2(\mathcal{X})$, it follows that when $\mu(\mathcal{X})=\infty$,

$$P_t(P_s f) = P_t\left(\sum_{k\in\mathbb{Z}} e^{-s\lambda_k}\mathbb{Q}_k f\right) = \sum_{k\in\mathbb{Z}} e^{-s\lambda_k} P_t\mathbb{Q}_k f = \sum_{k\in\mathbb{Z}} e^{-s\lambda_k}\left(\sum_{j\in\mathbb{Z}} e^{-t\lambda_j}\mathbb{Q}_j\mathbb{Q}_k f\right), \tag{3.11}$$

and that, when $\mu(\mathcal{X})<\infty$,

$$\begin{aligned} P_t(P_s f) &= P_t\left(e^{-s\lambda_{k_0-1}}\mathbb{P}_{k_0} f + \sum_{k=k_0}^{\infty} e^{-s\lambda_k}\mathbb{Q}_k f\right)\\ &= e^{-s\lambda_{k_0-1}} P_t\mathbb{P}_{k_0} f + \sum_{k=k_0}^{\infty} e^{-s\lambda_k} P_t\mathbb{Q}_k f\\ &= e^{-s\lambda_{k_0-1}}\left(e^{-t\lambda_{k_0-1}}\mathbb{P}_{k_0}\mathbb{P}_{k_0} f + \sum_{j=k_0}^{\infty} e^{-t\lambda_j}\mathbb{Q}_j\mathbb{P}_{k_0} f\right)\\ &\quad + \sum_{k=k_0}^{\infty} e^{-s\lambda_k}\left(e^{-t\lambda_{k_0-1}}\mathbb{P}_{k_0}\mathbb{Q}_k f + \sum_{j=k_0}^{\infty} e^{-s\lambda_j}\mathbb{Q}_j\mathbb{Q}_k f\right). \end{aligned} \tag{3.12}$$

By (ii) and (iii) of Lemma 3.2, we have $\mathbb{Q}_j\mathbb{P}_{k_0}=0$ for all $j\ge k_0$ and $\mathbb{Q}_l\mathbb{Q}_k=0$ for all $l\neq k$. Also, Lemma 3.2(i) implies that $\mathbb{P}_{k_0}^2=\mathbb{P}_{k_0}$ and $\mathbb{Q}_k^2=\mathbb{Q}_k$ for all $k\in\mathbb{Z}$. With these facts, we continue (3.11) and (3.12) respectively as follows: when $\mu(\mathcal{X})=\infty$,

$$P_t(P_s f) = \sum_{k\in\mathbb{Z}} e^{-s\lambda_k}\left(e^{-t\lambda_k}\mathbb{Q}_k\mathbb{Q}_k f\right) = \sum_{k\in\mathbb{Z}} e^{-(s+t)\lambda_k}\mathbb{Q}_k f = P_{t+s} f,$$

and, when $\mu(\mathcal{X})<\infty$,

$$\begin{aligned} P_t(P_s f) &= e^{-s\lambda_{k_0-1}}\left(e^{-t\lambda_{k_0-1}}\mathbb{P}_{k_0}\mathbb{P}_{k_0} f\right) + \sum_{k=k_0}^{\infty} e^{-s\lambda_k}\left(e^{-s\lambda_k}\mathbb{Q}_k\mathbb{Q}_k f\right)\\ &= e^{-(s+t)\lambda_{k_0-1}}\mathbb{P}_{k_0} f + \sum_{k=k_0}^{\infty} e^{-(s+t)\lambda_k}\mathbb{Q}_k f\\ &= P_{t+s} f. \end{aligned}$$

This verifies the semigroup property of $\{P_t\}_{\ge 0}$.

Step 4: proof of (iv). Let $f\in L^2(\mathcal{X})$. Consider first the case $\mu(\mathcal{X})=\infty$. For any $t\in(0,\infty)$, since the series (3.5) and (3.7) both converge in $L^2(\mathcal{X})$, we derive

$$\begin{aligned} \|f-P_t f\|_{L^2(\mathcal{X})}^2 &= \langle f-P_t f,\ f-P_t f\rangle\\ &= \left\langle \sum_{k\in\mathbb{Z}}(1-e^{-t\lambda_k})\mathbb{Q}_k f,\ \sum_{j\in\mathbb{Z}}(1-e^{-t\lambda_j})\mathbb{Q}_j f\right\rangle\\ &= \sum_{k\in\mathbb{Z}}\sum_{j\in\mathbb{Z}}(1-e^{-t\lambda_k})(1-e^{-t\lambda_j})\left\langle \mathbb{Q}_k f,\ \mathbb{Q}_j f\right\rangle\\ &= \sum_{k\in\mathbb{Z}}(1-e^{-t\lambda_k})^2\langle \mathbb{Q}_k f,\ \mathbb{Q}_k f\rangle \end{aligned}$$

$$= \sum_{k\in\mathbb{Z}} (1 - e^{-t\lambda_k})^2 \|\mathbb{Q}_k f\|^2_{L^2(\mathcal{X})},$$

where the penultimate step follows from (3.4). Noting that (3.6) implies $\sum_{k\in\mathbb{Z}} \|\mathbb{Q}_k f\|^2_{L^2(\mathcal{X})} < \infty$, we then derive from that

$$\lim_{t\to 0} \|f - P_t f\|^2_{L^2(\mathcal{X})} = \lim_{t\to 0}\left(\sum_{k\in\mathbb{Z}} (1 - e^{-t\lambda_k})^2 \|\mathbb{Q}_k f\|^2_{L^2(\mathcal{X})}\right) = \sum_{k\in\mathbb{Z}} \lim_{t\to 0}(1 - e^{-t\lambda_k})^2 \|\mathbb{Q}_k f\|^2_{L^2(\mathcal{X})} = 0.$$

The argument for the case $\mu(\mathcal{X}) < \infty$ follows in a similar way. Indeed, if $\mu(\mathcal{X}) < \infty$, then we use (3.3) and (3.4) to write

$$\begin{aligned}
\|f - P_t f\|^2_{L^2(\mathcal{X})} &= \langle f - P_t f,\ f - P_t f\rangle \\
&= \left\langle (1 - e^{-t\lambda_{k_0-1}})\mathbb{P}_{k_0} f + \sum_{k=k_0}^{\infty} (1 - e^{-t\lambda_k})\mathbb{Q}_k f,\ (1 - e^{-t\lambda_{k_0-1}})\mathbb{P}_{k_0} f + \sum_{j=k_0}^{\infty} (1 - e^{-t\lambda_j})\mathbb{Q}_j f \right\rangle \\
&= (1 - e^{-t\lambda_{k_0-1}})\|\mathbb{P}_{k_0} f\|^2_{L^2(\mathcal{X})} + \sum_{k=k_0}^{\infty} (1 - e^{-t\lambda_k})^2 \|\mathbb{Q}_k f\|^2_{L^2(\mathcal{X})},
\end{aligned}$$

which, together with the dominated convergence theorem, again leads to

$$\lim_{t\to 0} \|f - P_t f\|^2_{L^2(\mathcal{X})} = 0.$$

Thus, we obtain the strongly continuous property of $\{P_t\}_{\geq 0}$. □

3.3 Heat kernels generated by MRA

Applying Theorem 3.7, we are about to show Theorem 2.6 and Corollary 2.7 in this subsection. Indeed, Theorem 2.6 follows from the forthcoming Theorems 3.8 and 3.9; Corollary 2.7 is a consequence of a stronger result in Theorem 3.10.

Theorem 3.8. *Let $\{V_k\}_{k\in\mathbb{Z}}$ be an* MRA *in $L^2(\mathcal{X})$. Assume that the following hold:*

(i) *the sequence $\{\lambda_k\}_{k\in\mathbb{Z}} \subset [0,\infty)$ is increasing;*

(ii) *there exists a positive constant $C > 0$ such that for any $k \in \mathbb{Z}$ and $x \in \mathcal{X}$,*

$$\int_{\mathcal{X}} |\mathbb{P}_k(x,y)|\, d\mu(y) \leq C,$$

where $\mathbb{P}_k$ is the orthogonal projector from $L^2(\mathcal{X})$ to V_k.

Then, the family $\{p_t\}_{t>0}$ defined in (2.4) is a signed heat kernel. In particular, if $\mu(\mathcal{X}) = \infty$, then

$$p_t(x,y) = \lim_{N\to+\infty}\left(e^{-t\lambda_N}\mathbb{P}_{N+1}(x,y) + \sum_{k=-N}^{N}\left(e^{-t\lambda_{k-1}} - e^{-t\lambda_k}\right)\mathbb{P}_k(x,y) - e^{-t\lambda_{-N-1}}\mathbb{P}_{-N}(x,y)\right) \tag{3.13}$$

and, if $\mu(\mathcal{X}) < \infty$, then

$$p_t(x,y) = \lim_{N\to\infty}\left(e^{-t\lambda_N}\mathbb{P}_{N+1}(x,y) + \sum_{k=k_0}^{N}\left(e^{-t\lambda_{k-1}} - e^{-t\lambda_k}\right)\mathbb{P}_k(x,y)\right). \tag{3.14}$$

Proof. For any $t \in (0,\infty)$, it can be seen that p_t in (2.4) is the integral kernel of the operator P_t in (3.7). To obtain that $\{p_t\}_{t>0}$ is a signed heat kernel, based on Theorem 3.7 and the definition of a signed heat kernel in Definition 1.1, we only need to verify that

$$\sup_{t\in(0,\infty)} \sup_{x\in\mathcal{X}} \int_{\mathcal{X}} |p_t(x,y)|\, d\mu(y) < \infty. \tag{3.15}$$

Under the case $\mu(\mathcal{X}) = \infty$, we have by (2.4) and (3.2) that

$$\begin{aligned}
p_t(x,y) &= \sum_{k\in\mathbb{Z}} e^{-t\lambda_k}\mathbb{Q}_k(x,y)\\
&= \lim_{N\to+\infty} \sum_{k=-N}^{N} e^{-t\lambda_k}\left(\mathbb{P}_{k+1}(x,y) - \mathbb{P}_k(x,y)\right)\\
&= \lim_{N\to+\infty} \left(\sum_{k=-N+1}^{N+1} e^{-t\lambda_{k-1}}\mathbb{P}_k(x,y) - \sum_{k=-N}^{N} e^{-t\lambda_k}\mathbb{P}_k(x,y)\right)\\
&= \lim_{N\to+\infty} \left(e^{-t\lambda_N}\mathbb{P}_{N+1}(x,y) + \sum_{k=-N}^{N}\left(e^{-t\lambda_{k-1}} - e^{-t\lambda_k}\right)\mathbb{P}_k(x,y) - e^{-t\lambda_{-N-1}}\mathbb{P}_{-N}(x,y)\right).
\end{aligned}$$

This proves (3.13). Moreover, combining with the assumptions (i) and (ii), we further have

$$\begin{aligned}
&\int_{\mathcal{X}} |p_t(x,y)|\, d\mu(y)\\
&\le \limsup_{N\to+\infty} \int_{\mathcal{X}} \left(|\mathbb{P}_{N+1}(x,y)| + \sum_{k=-N}^{N}\left(e^{-t\lambda_{k-1}} - e^{-t\lambda_k}\right)|\mathbb{P}_k(x,y)| + |\mathbb{P}_{-N}(x,y)|\right) d\mu(y)\\
&\le 2C + C\limsup_{N\to\infty} \sum_{k=-N}^{N}\left(e^{-t\lambda_{k-1}} - e^{-t\lambda_k}\right)\\
&= 2C + C\limsup_{N\to\infty}\left(e^{-t\lambda_{-N-1}} - e^{-t\lambda_N}\right)\\
&\le 3C,
\end{aligned}$$

which gives the desired estimate (3.15) under the case $\mu(\mathcal{X}) = \infty$.

Consider now the case $\mu(\mathcal{X}) < \infty$. Via an argument similar to that used in (3.13), we again utilize (2.4) to write

$$\begin{aligned}
p_t(x,y) &= e^{-t\lambda_{k_0-1}}\mathbb{P}_{k_0}(x,y) + \sum_{k=k_0}^{\infty} e^{-t\lambda_k}\mathbb{Q}_k(x,y)\\
&= \lim_{N\to\infty}\left(e^{-t\lambda_N}\mathbb{P}_{N+1}(x,y) + \sum_{k=k_0}^{N}\left(e^{-t\lambda_{k-1}} - e^{-t\lambda_k}\right)\mathbb{P}_k(x,y)\right),
\end{aligned}$$

which proves (3.14). Again, applying the assumptions (i) and (ii) gives

$$\int_{\mathcal{X}} |p_t(x,y)|\, d\mu(y) \le \limsup_{N\to+\infty} \int_{\mathcal{X}} \left(|\mathbb{P}_{N+1}(x,y)| + \sum_{k=k_0}^{N}\left(e^{-t\lambda_{k-1}} - e^{-t\lambda_k}\right)|\mathbb{P}_k(x,y)|\right) d\mu(y)$$

$$\begin{aligned}
&\le C + C \limsup_{N\to\infty} \sum_{k=k_0}^{N} \left(e^{-t\lambda_{k-1}} - e^{-t\lambda_k}\right)\\
&= C + C \limsup_{N\to\infty} \left(e^{-t\lambda_{k_0-1}} - e^{-t\lambda_N}\right)\\
&\le 2C,
\end{aligned}$$

which again implies that (3.15) holds when $\mu(\mathcal{X}) < \infty$. □

In the forthcoming theorem, we use Theorem 3.8 to show Theorem 2.6 by validating the stochastic completeness property of $\{p_t\}_{t>0}$ that is defined in (2.4).

Theorem 3.9. *Let $\{V_k\}_{k\in\mathbb{Z}}$ be an admissible* MRA *in $L^2(\mathcal{X})$ and $\{\lambda_k\}_{k\in\mathbb{Z}} \subset [0,\infty)$ be a family of admissible spectrum. Then, the family $\{p_t\}_{t>0}$ defined in (2.4) is a stochastic complete signed heat kernel.*

Proof. Since $\{V_k\}_{k\in\mathbb{Z}}$ is an admissible MRA and $\{\lambda_k\}_{k\in\mathbb{Z}} \subset [0,\infty)$ is a family of admissible spectrum, it follows from Definitions 2.3 and 2.4 that the hypothesis of Theorem 3.8 are satisfied. Thus, by Theorem 3.8, we know that $\{p_t\}_{t>0}$ is a signed heat kernel.

It remains to validate the stochastic completeness property (1.3). Consider first the case $\mu(\mathcal{X}) = \infty$. Given a large integer $N \in \mathbb{N}$, we utilize (2.4) to write

$$\begin{aligned}
p_t(x,y) &= \sum_{k\in\mathbb{Z}} e^{-t\lambda_k}\mathbb{Q}_k(x,y)\\
&= \sum_{k=-\infty}^{-N-1} e^{-t\lambda_k}\mathbb{Q}_k(x,y) + \sum_{|k|\le N} e^{-t\lambda_k}\mathbb{Q}_k(x,y) + \sum_{k=N+1}^{\infty} e^{-t\lambda_k}\mathbb{Q}_k(x,y). \qquad (3.16)
\end{aligned}$$

From $\mathbb{Q}_k = \mathbb{P}_{k+1} - \mathbb{P}_k$ (see (3.2)) and the condition (A2) of Definition 2.3, it follows that

$$\int_{\mathcal{X}} \left(\sum_{|k|\le N} e^{-t\lambda_k}\mathbb{Q}_k(x,y)\right) d\mu(y) = \sum_{|k|\le N} e^{-t\lambda_k} \int_{\mathcal{X}} \mathbb{Q}_k(x,y)\, d\mu(y) = 0.$$

Moreover, note that the condition (A1) in Definition 2.3 implies

$$\int_{\mathcal{X}} |\mathbb{Q}_k(x,y)|\, d\mu(y) \le \int_{\mathcal{X}} |\mathbb{P}_{k+1}(x,y)|\, d\mu(y) + \int_{\mathcal{X}} |\mathbb{P}_k(x,y)|\, d\mu(y) \le 2C$$

and, hence,

$$\begin{aligned}
\int_{\mathcal{X}} \left|\sum_{k=N+1}^{\infty} e^{-t\lambda_k}\mathbb{Q}_k(x,y)\right| d\mu(y) &\le \sum_{k=N+1}^{\infty} e^{-t\lambda_k} \int_{\mathcal{X}} |\mathbb{Q}_k(x,y)|\, d\mu(y)\\
&\le 2C \sum_{k=N+1}^{\infty} e^{-t\lambda_k},
\end{aligned}$$

which tends to 0 as $N \to +\infty$ by terms of Definition 2.4(c). So, to obtain (1.3), it suffices to validate

$$\lim_{N\to+\infty} \int_{\mathcal{X}} \left(\sum_{k=-\infty}^{-N-1} e^{-t\lambda_k}\mathbb{Q}_k(x,y)\right) d\mu(y) = 1. \qquad (3.17)$$

To prove (3.17), for any $m \le -N-1$, by $\mathbb{Q}_k = \mathbb{P}_{k+1} - \mathbb{P}_k$, we write (see also the proof of (3.13))

$$\begin{aligned}
\sum_{k=m}^{-N-1} e^{-t\lambda_k}\mathbb{Q}_k(x,y) &= \sum_{k=m}^{-N-1} e^{-t\lambda_k}\Big(\mathbb{P}_{k+1}(x,y) - \mathbb{P}_k(x,y)\Big)\\
&= \sum_{k=m+1}^{-N} e^{-t\lambda_{k-1}}\mathbb{P}_k(x,y) - \sum_{k=m}^{-N-1} e^{-t\lambda_k}\mathbb{P}_k(x,y)\\
&= e^{-t\lambda_{-N-1}}\mathbb{P}_{-N}(x,y) + \sum_{k=m}^{-N-1}\Big(e^{-t\lambda_{k-1}} - e^{-t\lambda_k}\Big)\mathbb{P}_k(x,y) - e^{-t\lambda_{m-1}}\mathbb{P}_m(x,y). \qquad (3.18)
\end{aligned}$$

Observe that the condition (A3) of Definition 2.3 implies

$$\lim_{m\to-\infty} e^{-t\lambda_{m-1}}\mathbb{P}_m(x,y) \le \lim_{m\to-\infty} |\mathbb{P}_m(x,y)| = 0.$$

So, letting $m \to -\infty$ in both sides of (3.18) leads to

$$\sum_{k=-\infty}^{-N-1} e^{-t\lambda_k}\mathbb{Q}_k(x,y) = e^{-t\lambda_{-N-1}}\mathbb{P}_{-N}(x,y) + \sum_{k=-\infty}^{-N-1}\Big(e^{-t\lambda_{k-1}} - e^{-t\lambda_k}\Big)\mathbb{P}_k(x,y). \qquad (3.19)$$

For all $k \in \mathbb{Z}$, by Definition 2.4(a) and the fact that $1 - e^{-s} \le s$ for all $s \in (0,\infty)$, we derive

$$\left|e^{-t\lambda_{k-1}} - e^{-t\lambda_k}\right| = e^{-t\lambda_{k-1}}\left|1 - e^{-t(\lambda_k-\lambda_{k-1})}\right| \le 1 - e^{-t(\lambda_k-\lambda_{k-1})} \le t(\lambda_k - \lambda_{k-1}),$$

where we remark that $\lambda_k - \lambda_{k-1} \ge 0$ for all $k \in \mathbb{Z}$. Consequently, the first condition in Definition 2.4(b) implies that when $N \to +\infty$,

$$\sum_{k=-\infty}^{-N-1}\left|e^{-t\lambda_{k-1}} - e^{-t\lambda_k}\right| \le t\sum_{k=-\infty}^{-N-1}(\lambda_k - \lambda_{k-1}) = t\Big(\lambda_{-N-1} - \lim_{m\to-\infty}\lambda_m\Big) \to 0.$$

From this and condition (A1) of Definition 2.3, it follows that

$$\begin{aligned}
\lim_{N\to+\infty}\int_{\mathcal{X}}\left|\sum_{k=-\infty}^{-N-1}\Big(e^{-t\lambda_{k-1}} - e^{-t\lambda_k}\Big)\mathbb{P}_k(x,y)\right| d\mu(y) &\le \lim_{N\to+\infty}\sum_{k=-\infty}^{-N-1}\left|e^{-t\lambda_{k-1}} - e^{-t\lambda_k}\right|\int_{\mathcal{X}}|\mathbb{P}_k(x,y)|\,d\mu(y)\\
&\le C\lim_{N\to+\infty}\sum_{k=-\infty}^{-N-1}\left|e^{-t\lambda_{k-1}} - e^{-t\lambda_k}\right|\\
&= 0.
\end{aligned}$$

Moreover, by condition (A2) of Definition 2.3 and the first condition in Definition 2.4(b), we also obtain

$$\lim_{N\to+\infty}\left(e^{-t\lambda_{-N-1}}\int_{\mathcal{X}}\mathbb{P}_{-N}(x,y)\,d\mu(y)\right) = \lim_{N\to+\infty} e^{-t\lambda_{-N-1}} = 1.$$

Invoking these last two formulae and (3.19), we derive (3.17). This proves (1.3) under the case $\mu(\mathcal{X}) = \infty$.

Now, we are left to validate (1.3) under $\mu(\mathcal{X}) < \infty$. In this case, instead of (3.16), we derive from (2.4) and $\lambda_{k_0-1} = 0$ (see the second condition of Definition 2.4(b)) that

$$p_t(x,y) = \mathbb{P}_{k_0}(x,y) + \sum_{k=k_0}^{N} e^{-t\lambda_k}\mathbb{Q}_k(x,y) + \sum_{k=N+1}^{\infty} e^{-t\lambda_k}\mathbb{Q}_k(x,y).$$

Just like the argument for the case $\mu(\mathcal{X}) = \infty$, we still have

$$\int_{\mathcal{X}} \left(\sum_{k=k_0}^{N} e^{-t\lambda_k} \mathbb{Q}_k(x, y) \right) d\mu(y) = 0 \quad \text{for all } N \in \mathbb{N}$$

and

$$\int_{\mathcal{X}} \left| \sum_{k=N+1}^{\infty} e^{-t\lambda_k} \mathbb{Q}_k(x, y) \right| d\mu(y) \to 0 \quad \text{as } N \to +\infty.$$

These last two formulae, along with condition (A2) of Definition 2.3, again yields (1.3). □

Applying Theorem 3.8, we consider the positivity of $\{p_t\}_{t>0}$ in (2.4).

Theorem 3.10. *In addition to the assumptions of Theorem 3.8, assume that the following hold:*

(a) *for any $k \in \mathbb{Z}$ and x, $y \in \mathcal{X}$, $\mathbb{P}_k(x, y) \geq 0$;*

(b) *when $\mu(\mathcal{X}) = \infty$, assume further that for all $x, y \in \mathcal{X}$, $\lim_{k\to-\infty} \mathbb{P}_k(x, y) = 0$.*

Then, the family $\{p_t\}_{t>0}$ in (2.4) is a heat kernel.

Proof. By Theorem 3.8, we already know that $\{p_t\}_{t>0}$ is a signed heat kernel. To prove that $\{p_t\}_{t>0}$ is a heat kernel, it suffices to validate that, for any $t \in (0, \infty)$ and x, $y \in \mathcal{X}$,

$$p_t(x, y) \geq 0. \tag{3.20}$$

Indeed, when $\mu(\mathcal{X}) = \infty$, we have by (3.13) and the assumption (b) that

$$p_t(x, y) = \lim_{N\to\infty} \left(e^{-t\lambda_N} \mathbb{P}_{N+1}(x, y) + \sum_{k=-N}^{N} \left(e^{-t\lambda_{k-1}} - e^{-t\lambda_k} \right) \mathbb{P}_k(x, y) \right), \tag{3.21}$$

Meanwhile, when $\mu(\mathcal{X}) < \infty$, we know directly from (3.14) that

$$p_t(x, y) = \lim_{N\to\infty} \left(e^{-t\lambda_N} \mathbb{P}_{N+1}(x, y) + \sum_{k=k_0}^{N} \left(e^{-t\lambda_{k-1}} - e^{-t\lambda_k} \right) \mathbb{P}_k(x, y) \right). \tag{3.22}$$

From the assumption (i) of Theorem 3.8, it follows that $e^{-t\lambda_{k-1}} - e^{-t\lambda_k} \geq 0$ for all $k \in \mathbb{Z}$ and $t \in (0, \infty)$. Combining this with (a), we deduce that all the terms in the right hand sides of (3.21) and (3.22) are nonnegative. This implies that (3.20) holds. □

3.4 Generator of the heat kernel

Under the hypothesis of Theorem 3.8, the family $\{p_t\}_{t>0}$ defined in (2.4) is a signed heat kernel. In this subsection, we discuss the spectrum and the functional calculus of the generator of $\{p_t\}_{t>0}$.

The family $\{p_t\}_{t>0}$ defined in (2.4) gives rise to the semigroup $\{P_t\}_{t\geq 0}$ in (3.7). According to [28, p. 19, Lemma 1.3.2(ii)], any strongly continuous symmetric semigroup $\{P_t\}_{t\geq 0}$ corresponds uniquely to a nonnegative self-adjoint densely defined operator $\mathcal{L}$ in $L^2(\mathcal{X})$. Indeed, for any

$$f \in \mathrm{Dom}(\mathcal{L}) := \left\{ f \in L^2(\mathcal{X}) : \|\mathcal{L}f\|_{L^2(\mathcal{X})} < \infty \right\},$$

it holds that

$$\mathcal{L}f = \lim_{t\downarrow 0} \frac{f - P_t f}{t} \quad \text{in } \ L^2(\mathcal{X}). \tag{3.23}$$

The operator $\mathcal{L}$ is called the *generator* of the semigroup $\{P_t\}_{t>0}$ (or the heat kernel $\{p_t\}_{t>0}$). By (3.5), (3.7) and (3.23), we easily deduce

$$\mathcal{L}f = \begin{cases} \sum_{k\in\mathbb{Z}} \lambda_k \mathbb{Q}_k f & \text{as } \ \mu(\mathcal{X}) = \infty; \\ \lambda_{k_0-1}\mathbb{P}_{k_0} f + \sum_{k=k_0}^{\infty} \lambda_k \mathbb{Q}_k f & \text{as } \ \mu(\mathcal{X}) < \infty. \end{cases} \tag{3.24}$$

Moreover, by Lemma 3.2, we argue as in the proofs of (3.9) and (3.10) to derive

$$\|\mathcal{L}f\|^2_{L^2(\mathcal{X})} = \begin{cases} \sum_{k\in\mathbb{Z}} \lambda_k^2 \|\mathbb{Q}_k f\|^2_{L^2(\mathcal{X})} & \text{as } \ \mu(\mathcal{X}) = \infty; \\ \lambda^2_{k_0-1}\|\mathbb{P}_{k_0} f\|^2_{L^2(\mathcal{X})} + \sum_{k=k_0}^{\infty} \lambda_k^2 \|\mathbb{Q}_k f\|^2_{L^2(\mathcal{X})} & \text{as } \ \mu(\mathcal{X}) < \infty. \end{cases}$$

Below we will discuss the spectrum and the functional calculus of the generator $\mathcal{L}$ in (3.24).

Lemma 3.11. *Let $\mathcal{L}$ be the operator as in (3.24) and $\sigma(\mathcal{L})$ be its spectrum. Then*

$$\sigma(\mathcal{L}) = \overline{\{\lambda_k\}_{k\in J_0}}$$

with $J_0 = \mathbb{Z}$ when $\mu(\mathcal{X}) = \infty$ and $J_0 = \{k_0 - 1, k_0, \dots\}$ when $\mu(\mathcal{X}) < \infty$.

Proof. To simplify the notation, let $J = J_0$ when $\mu(\mathcal{X}) = \infty$ and $J = J_0 \setminus \{k_0 - 1\}$ when $\mu(\mathcal{X}) < \infty$. Denote by $\sigma_p(\mathcal{L})$ the point spectrum of $\mathcal{L}$, namely,

$$\sigma_p(\mathcal{L}) := \{\lambda \in \mathbb{C} : \ \lambda \text{ is a eigenvalue of } \mathcal{L}\}.$$

We split the proof into the following 3 steps.

Step 1: proof of $\{\lambda_k\}_{k\in J_0} \subset \sigma_p(\mathcal{L})$. Without loss of generality, we only consider the case $\mu(\mathcal{X}) < \infty$ in this step. Fix $k \in J$. Recall that each closed linear subspace of a separable Hilbert space has a complete countable orthonormal basis (see [57, p. 89, Corollary]). Thus, we respectively denote by $\{\phi_{k_0,\alpha}\}_{\alpha\in\mathscr{A}_{k_0}}$ and $\{\psi_{k,\gamma}\}_{\gamma\in\mathscr{G}_k}$ the orthonormal basis of V_{k_0} and W_k, where $\mathscr{A}_{k_0}$ and $\mathscr{G}_k$ are countable index sets. By the definitions of $\mathbb{P}_{k_0}$ and $\mathbb{Q}_k$, we derive from (3.24) that for any $k \in J$ and $\gamma \in \mathscr{G}_k$,

$$\mathcal{L}\psi_{k,\gamma} = \lambda_k \psi_{k,\alpha}$$

and, for any $\alpha \in \mathscr{A}_{k_0}$,

$$\mathcal{L}\phi_{k_0,\alpha} = \lambda_{k_0-1}\phi_{k_0,\alpha}.$$

Thus, any λ_k with $k \in J_0$ is an eigenvalue of $\mathcal{L}$ and, hence, $\{\lambda_k\}_{k\in J_0} \subset \sigma_p(\mathcal{L})$.

Step 2: proof of $\{\lambda_k\}_{k\in J_0} \supset \sigma_p(\mathcal{L})$. Suppose that $\lambda \in \mathbb{C}$ is an eigenvalue of $\mathcal{L}$, that is, there exists a non-zero function $f \in \mathrm{Dom}(\mathcal{L})$ such that

$$\mathcal{L}(f) = \lambda f.$$

We are going to show that $\lambda \in \{\lambda_k\}_{k\in J_0}$.

Consider the case $\mu(\mathcal{X}) < \infty$. Since $f \in L^2(\mathcal{X})$ and $f \neq 0$, we derive from (3.6) that either $\|\mathbb{P}_{k_0} f\|_{L^2(\mathcal{X})} \neq 0$ or $\|Q_{j_0} f\|_{L^2(\mathcal{X})} \neq 0$ for some $j_0 \geq k_0$. Using (3.5) and (3.24), together with Lemma 3.2, we proceed as the proof of (3.10) to derive

$$
\begin{aligned}
0 = \|\lambda f - \mathcal{L}(f)\|^2_{L^2(\mathcal{X})} &= \left\| (\lambda - \lambda_{k_0-1})\mathbb{P}_{k_0} f + \sum_{j\in\mathbb{Z}} (\lambda - \lambda_j)\mathbb{Q}_j f \right\|^2_{L^2(\mathcal{X})} \\
&= |\lambda - \lambda_{k_0-1}|^2 \|\mathbb{P}_{k_0} f\|^2_{L^2(\mathcal{X})} + \sum_{j\in\mathbb{Z}} |\lambda - \lambda_j|^2 \|\mathbb{Q}_j f\|^2_{L^2(\mathcal{X})}.
\end{aligned}
$$

Thus, we must have with $\lambda = \lambda_{k_0-1}$ or $\lambda = \lambda_{j_0}$, which implies $\lambda \in \{\lambda_k\}_{k\in J_0}$.

When $\mu(\mathcal{X}) = \infty$, since $f \in L^2(\mathcal{X})$ is non-zero, we deduce from (3.6) that $\|Q_{j_0} f\|_{L^2(\mathcal{X})} \neq 0$ for some $j_0 \in \mathbb{Z}$. Applying (3.5), (3.24) and Lemma 3.2, we now follow the proof of (3.9) and obtain

$$
0 = \|\lambda f - \mathcal{L}(f)\|^2_{L^2(\mathcal{X})} = \left\| \sum_{j\in\mathbb{Z}} (\lambda - \lambda_j)\mathbb{Q}_j f \right\|^2_{L^2(\mathcal{X})} = \sum_{j\in\mathbb{Z}} |\lambda - \lambda_j|^2 \|\mathbb{Q}_j f\|^2_{L^2(\mathcal{X})}.
$$

This implies that $\lambda = \lambda_{j_0} \in \{\lambda_k\}_{k\in J_0}$, as desired.

Step 3: proof of $\sigma(\mathcal{L}) = \overline{\{\lambda_k\}_{k\in J_0}}$. Based on **Step 1** and **Step 2**, we have $\{\lambda_k\}_{k\in J_0} = \sigma_p(\mathcal{L})$. Then, by the fact that $\sigma_p(\mathcal{L}) \subset \sigma(\mathcal{L})$, we obtain $\{\lambda_k\}_{k\in J_0} \subset \sigma(\mathcal{L})$, thereby leading to

$$
\overline{\{\lambda_k\}_{k\in J_0}} \subset \sigma(\mathcal{L}),
$$

since $\sigma(\mathcal{L})$ is a closed subset of $\mathbb{C}$. Thus, it remains to show that $\sigma(\mathcal{L}) \subset \overline{\{\lambda_k\}_{k\in J_0}}$. For any $\lambda \in \mathbb{C}\backslash\overline{\{\lambda_k\}_{k\in J_0}}$, we have by **Step 2** that the mapping

$$
\lambda - \mathcal{L}: \ \mathrm{Dom}(\mathcal{L}) \to L^2(\mathcal{X})
$$

is injective. So, it suffices to prove that $\lambda - \mathcal{L}$ is surjective, $(\lambda - \mathcal{L})^{-1}$ exists and bounded on $L^2(\mathcal{X})$.

Since $\lambda \in \mathbb{C} \setminus \overline{\{\lambda_k\}_{k\in J_0}}$, we see that $c := \min_{k\in J_0} |\lambda - \lambda_k|$ exists and $c > 0$. For any $f \in L^2(\mathcal{X})$, define

$$
g := \begin{cases} \displaystyle\sum_{j\in\mathbb{Z}} (\lambda - \lambda_j)^{-1}\mathbb{Q}_j f & \text{as } \mu(\mathcal{X}) = \infty; \\ \displaystyle(\lambda - \lambda_{k_0-1})^{-1}\mathbb{P}_{k_0} f + \sum_{j=k_0}^{\infty} (\lambda - \lambda_j)^{-1}\mathbb{Q}_j f & \text{as } \mu(\mathcal{X}) < \infty. \end{cases}
$$

Applying Lemma 3.2 and proceeding as the proofs of (3.9)-(3.10), we easily deduce that

$$
\begin{aligned}
\|g\|^2_{L^2(\mathcal{X})} &= \begin{cases} \displaystyle\sum_{j\in\mathbb{Z}} |\lambda - \lambda_j|^{-2} \|\mathbb{Q}_j f\|^2_{L^2(\mathcal{X})} & \text{as } \mu(\mathcal{X}) = \infty; \\ \displaystyle|\lambda - \lambda_{k_0-1}|^{-2} \|\mathbb{P}_{k_0} f\|^2_{L^2(\mathcal{X})} + \sum_{j=k_0}^{\infty} |\lambda - \lambda_j|^{-2} \|\mathbb{Q}_j f\|^2_{L^2(\mathcal{X})} & \text{as } \mu(\mathcal{X}) < \infty, \end{cases} \\
&\leq c^{-2} \begin{cases} \displaystyle\sum_{k\in\mathbb{Z}} \|\mathbb{Q}_k f\|^2_{L^2(\mathcal{X})} & \text{as } \mu(\mathcal{X}) = \infty; \\ \displaystyle\left(\|\mathbb{P}_{k_0} f\|^2_{L^2(\mathcal{X})} + \sum_{k=k_0}^{\infty} \|\mathbb{Q}_k f\|^2_{L^2(\mathcal{X})} \right) & \text{as } \mu(\mathcal{X}) < \infty, \end{cases}
\end{aligned}
$$

$$= c^{-2}\|f\|^2_{L^2(\mathcal{X})},$$

where the last step is due to (3.6). Further, observing that the expression of g and Lemma 3.2 imply

$$\mathbb{Q}_k g = (\lambda - \lambda_k)^{-1}\mathbb{Q}_k f \quad \text{for all } k \in J$$

and

$$\mathbb{P}_{k_0} g = (\lambda - \lambda_{k_0})^{-1}\mathbb{P}_{k_0} f,$$

we therefor deduce from (3.5) that

$$\begin{aligned}
(\lambda - \mathcal{L})g &= \begin{cases} \displaystyle\sum_{k\in\mathbb{Z}}(\lambda-\lambda_k)\mathbb{Q}_k g & \text{as } \mu(\mathcal{X}) = \infty; \\ \displaystyle(\lambda - \lambda_{k_0-1})\mathbb{P}_{k_0} g + \sum_{k=k_0}^{\infty}(\lambda-\lambda_k)\mathbb{Q}_k g & \text{as } \mu(\mathcal{X}) < \infty. \end{cases} \\
&= \begin{cases} \displaystyle\sum_{k\in\mathbb{Z}}\mathbb{Q}_k f & \text{as } \mu(\mathcal{X}) = \infty; \\ \displaystyle\mathbb{P}_{k_0} f + \sum_{k=k_0}^{\infty}\mathbb{Q}_k f & \text{as } \mu(\mathcal{X}) < \infty, \end{cases} \\
&= f.
\end{aligned}$$

This proves that $\lambda - \mathcal{L}:\ \mathrm{Dom}(\mathcal{L}) \to L^2(\mathcal{X})$ is surjective and, moreover, its inverse $(\lambda - \mathcal{L})^{-1}$ exists and is bounded on $L^2(\mathcal{X})$. □

Next, we discuss the spectral resolution family of $\mathcal{L}$.

Lemma 3.12. *Let $J_0 = \mathbb{Z}$ when $\mu(\mathcal{X}) = \infty$ and $J_0 = \{k_0 - 1, k_0, \dots\}$ when $\mu(\mathcal{X}) < \infty$. Assume that $\{\lambda_k\}_{k\in J_0} \subset [0,\infty)$ is strictly increasing. For any $\lambda \in \mathbb{R}$, define*

$$E_\lambda = \begin{cases} \displaystyle\sum_{\{k\in\mathbb{Z}:\ \lambda_k\le\lambda\}} \mathbb{Q}_k & \text{if } \lambda > 0 \text{ and } \mu(\mathcal{X}) = \infty; \\ \displaystyle\mathbb{P}_{k_0} + \sum_{\{k\in\{k_0,\dots\}:\ \lambda_k\le\lambda\}} \mathbb{Q}_k & \text{if } \lambda \ge \lambda_{k_0-1} \text{ and } \mu(\mathcal{X}) < \infty; \\ 0 & \text{otherwise}. \end{cases} \tag{3.25}$$

Then $\{E_\lambda\}_{\lambda\in\mathbb{R}}$ is a spectral resolution of the identity in $L^2(\mathcal{X})$, that is, it satisfies the following:

(i) $E_{-\infty} := \lim_{\lambda\to-\infty} E_\lambda = 0$ *in* $L^2(\mathcal{X})$*;*

(ii) $E_{+\infty} := \lim_{\lambda\to+\infty} E_\lambda = id$ *in* $L^2(\mathcal{X})$*;*

(iii) *if $t \le s$ and $f \in L^2(\mathcal{X})$, then $\langle E_t f,\ f\rangle \le \langle E_s f,\ f\rangle$ and $E_t E_s = E_t = E_s E_t$ in $L^2(\mathcal{X})$;*

(iv) *for any $\lambda \in \mathbb{R}$, there is $E_\lambda = E_{\lambda_+} := \lim_{s\downarrow\lambda} E_s$ in $L^2(\mathcal{X})$.*

Proof. Based on Lemma 3.5, all these properties are easy except (iii). We take the case $\mu(\mathcal{X}) = \infty$ to explain (iii); the proof for the case $\mu(\mathcal{X}) < \infty$ follows in a similar manner. Assume without loss of generality that $t > 0$. Then, for any $t \le s$ and $f \in L^2(\mathcal{X})$, we have by Lemma 3.2(i) that

$$\langle E_t f,\ f\rangle = \sum_{\{k\in\mathbb{Z}:\ \lambda_k\le t\}} \langle \mathbb{Q}_k f,\ f\rangle = \sum_{\{k\in\mathbb{Z}:\ \lambda_k\le t\}} \|\mathbb{Q}_k f\|^2_{L^2(\mathcal{X})}$$

$$\leq \sum_{\{k\in\mathbb{Z}:\, \lambda_k\leq s\}} \|\mathbb{Q}_k f\|_{L^2(\mathcal{X})}^2 = \sum_{\{k\in\mathbb{Z}:\, \lambda_k\leq s\}} \langle \mathbb{Q}_k f,\, f\rangle = \langle E_s f,\, f\rangle.$$

Moreover, for any $t \leq s$ and $f \in L^2(\mathcal{X})$, by Lemma 3.2(iii), we obtain

$$E_t E_s f = \sum_{\{k\in\mathbb{Z}:\, \lambda_k\leq t\}} \mathbb{Q}_k E_s f = \sum_{\{k\in\mathbb{Z}:\, \lambda_k\leq t\}} \sum_{\{j\in\mathbb{Z}:\, \lambda_j\leq s\}} \mathbb{Q}_k \mathbb{Q}_j f = \sum_{\{k\in\mathbb{Z}:\, \lambda_k\leq t\}} \mathbb{Q}_j \mathbb{Q}_j f = \sum_{\{k\in\mathbb{Z}:\, \lambda_k\leq t\}} \mathbb{Q}_j f = E_t f.$$

The same argument also yields that $E_s E_t f = E_t f$. This ends the proof. □

Let φ be a real-valued continuous function on $\mathbb{R}$. For any $f \in L^2(\mathcal{X})$, since $\{E_\lambda\}_{\lambda\in\mathbb{R}}$ in (3.25) is a spectral resolution of the identity in $L^2(\mathcal{X})$, we have by the spectral theory (see [57, p. 310, Proposition 2 and Corollary]) that it makes sense to define

$$T_\varphi f := \int_{\mathbb{R}} \varphi(\lambda)\, dE_\lambda f = \lim_{\substack{\alpha\to-\infty\\ \beta\to+\infty}} \int_{(\alpha,\beta]} \varphi(\lambda)\, dE_\lambda f \tag{3.26}$$

whenever the limit exists as an $L^2(\mathcal{X})$-function. Here, the integral $\int_\alpha^\beta \varphi(\lambda)\, dE_\lambda f$ is understood as the limit (again we mean limit in the sense of $L^2(\mathcal{X})$ norm) of Riemannian sum

$$\sum_{j=1}^n \varphi(t_j)\left(E_{s_{j+1}} f - E_{s_j} f\right), \tag{3.27}$$

where $\alpha = s_0 < s_1 < \cdots < s_n = \beta$, $t_j \in (s_{j-1}, s_j]$ and $\delta := \max_j |s_j - s_{j-1}|$ tends to 0. In this sense, the operator $T_\varphi f$ in (3.26) with the domain

$$\mathrm{Dom}(T_\varphi) = \left\{ f \in L^2(\mathcal{X}) :\ \|T_\varphi f\|_{L^2(\mathcal{X})}^2 = \int_{\mathbb{R}} |\varphi(\lambda)|^2\, d\|E_\lambda f\|_{L^2(\mathcal{X})}^2 < \infty \right\}$$

is a self-adjoint operator in $L^2(\mathcal{X})$ (see [57, p. 311, Theorem 2]). Of course, when $f \in \mathrm{Dom}(T_\varphi)$, we can take $t_j = s_j$ in (3.27).

Theorem 3.13. *Let $J_0 = \mathbb{Z}$ when $\mu(\mathcal{X}) = \infty$ and $J = \{k_0 - 1, k_0, \dots\}$ when $\mu(\mathcal{X}) < \infty$. Assume that $\{\lambda_k\}_{k\in J_0} \subset [0,\infty)$ is strictly increasing. If φ is a real-valued continuous function on $\mathbb{R}$, then for any $f \in \mathrm{Dom}(T_\varphi)$, the equality*

$$T_\varphi f = \begin{cases} \displaystyle\sum_{k\in\mathbb{Z}} \varphi(\lambda_k)\mathbb{Q}_k f & \textit{as } \mu(\mathcal{X}) = \infty; \\ \displaystyle\varphi(\lambda_{k_0-1})\mathbb{P}_{k_0} f + \sum_{k=k_0}^\infty \varphi(\lambda_k)\mathbb{Q}_k f & \textit{as } \mu(\mathcal{X}) < \infty, \end{cases} \tag{3.28}$$

holds in $L^2(\mathcal{X})$. Moreover, for any $f \in \mathrm{Dom}(T_\varphi)$, there is

$$\|T_\varphi f\|_{L^2(\mathcal{X})}^2 = \begin{cases} \displaystyle\sum_{k\in\mathbb{Z}} \varphi(\lambda_k)^2 \|\mathbb{Q}_k f\|_{L^2(\mathcal{X})}^2 & \textit{as } \mu(\mathcal{X}) = \infty; \\ \displaystyle\varphi(\lambda_{k_0-1})^2 \|\mathbb{P}_{k_0} f\|_{L^2(\mathcal{X})}^2 + \sum_{k=k_0}^\infty \varphi(\lambda_k)^2 \|\mathbb{Q}_k f\|_{L^2(\mathcal{X})}^2 & \textit{as } \mu(\mathcal{X}) < \infty. \end{cases} \tag{3.29}$$

Proof. Once we have obtained (3.28), by using Lemma 3.2 and proceeding the arguments in the proofs of (3.9)-(3.10), we then obtain (3.29). So, it remains to show (3.28).

Let α and β be as in (3.26). Consider the integral $\int_{(\alpha,\beta]} \varphi(\lambda)\, dE_\lambda f$, which is understood as the limit of the Riemannian sum in (3.27). Our aim is to show that, for any $\epsilon > 0$, there exists some $\delta > 0$ such that whenever $\max_j |s_j - s_{j-1}| < \delta$, the Riemannian sum in (3.27) satisfies the following: when $\mu(\mathcal{X}) = \infty$,

$$\left\| \sum_{j=1}^{n} \varphi(s_j) \left(E_{s_j} f - E_{s_{j-1}} f \right) - \sum_{\{k\in\mathbb{Z}:\, \alpha<\lambda_k\le\beta\}} \varphi(\lambda_k) \mathbb{Q}_k f \right\|_{L^2(\mathcal{X})} < 3\epsilon \|f\|_{L^2(\mathcal{X})}. \tag{3.30}$$

and, when $\mu(\mathcal{X}) < \infty$,

$$\left\| \sum_{j=1}^{n} \varphi(s_j) \left(E_{s_j} f - E_{s_{j-1}} f \right) - \varphi(\lambda_{k_0-1}) \mathbb{P}_{k_0} f - \sum_{\{k\ge k_0:\, \alpha<\lambda_k\le\beta\}} \varphi(\lambda_k) \mathbb{Q}_k f \right\|_{L^2(\mathcal{X})} < 3\epsilon \|f\|_{L^2(\mathcal{X})}. \tag{3.31}$$

Indeed, once we have proved (3.30) and (3.31), then we take the limit in the Riemannian sum in (3.27), thereby deriving

$$\int_\alpha^\beta \phi(\lambda)\, dE_\lambda f = \sum_{\{k\in\mathbb{Z}:\, \alpha<\lambda_k\le\beta\}} \varphi(\lambda_k) \mathbb{Q}_k f \qquad \text{as } \mu(\mathcal{X}) = \infty$$

and

$$\int_\alpha^\beta \phi(\lambda)\, dE_\lambda f = \varphi(\lambda_{k_0-1}) \mathbb{P}_{k_0} f + \sum_{\{k\ge k_0:\, \alpha<\lambda_k\le\beta\}} \varphi(\lambda_k) \mathbb{Q}_k f \qquad \text{as } \mu(\mathcal{X}) < \infty.$$

In both sides of these last two formulae, letting $\alpha \to -\infty$ and $\beta \to +\infty$, we know from (3.26) that (3.28) holds. Therefore, we are left to validate (3.30) and (3.31).

No matter $\mu(\mathcal{X})$ is finite or not, by the fact that $\{\lambda_k\}_{k\in J_0} \subset [0,\infty)$ is strictly increasing, we see that

$$\sigma^* := \lim_{k\to+\infty} \lambda_k$$

may exist as a real positive number or may be $+\infty$. Meanwhile, when $\mu(\mathcal{X}) = \infty$, the limit

$$\sigma_* := \lim_{k\to-\infty} \lambda_k$$

does exist and $0 \le \sigma_* < \sigma^* \le +\infty$. Note that φ is uniformly continuous on $[\alpha,\beta]$, so that for any $\epsilon > 0$ there exists $\delta > 0$ small enough such that when $\max_j |s_j - s_{j-1}| < \delta$, we have

$$\sup_{1\le j\le n} \sup_{s,t\in(s_{j-1},s_j]} |\varphi(t) - \varphi(s)| < \epsilon. \tag{3.32}$$

We may take δ to be small enough so that each interval $(s_{j-1}, s_j]$ contains at most one element of $\{\sigma_*, \sigma^*\} \cup \{\lambda_k\}_{k\in J_0}$. Then, we consider the following five cases:

Case 1: Let $j \in \{1, 2, \dots, n\}$. If $(s_{j-1}, s_j]$ does not contain any element from $\{\lambda_k\}_{k\in J_0}$, then by (3.25), we have

$$E_{s_j} f - E_{s_{j-1}} f = 0. \tag{3.33}$$

Case 2: Let $j \in \{1, 2, \ldots, n\}$. Set $J = J_0 = \mathbb{Z}$ when $\mu(\mathcal{X}) = \infty$ and $J = J_0 \setminus \{k_0 - 1\}$ when $\mu(\mathcal{X}) < \infty$. If $(s_{j-1}, s_j]$ contains exactly one element from $\{\lambda_k\}_{k\in J}$, denoted by λ_{k_j}, then

$$\lambda_{k_j-1} \le s_{j-1} < \lambda_{k_j} \le s_j < \lambda_{k_j+1},$$

so that (3.25) implies

$$E_{s_j} f - E_{s_{j-1}} f = \mathbb{Q}_{k_j} f.$$

Case 3: Consider the case $\mu(\mathcal{X}) < \infty$ and for some $j_0 \in \{1, 2, \ldots, n\}$ the interval $(s_{j_0-1}, s_{j_0}]$ contains λ_{k_0-1}. For any $i \le j_0 - 1$, the interval $(s_{i-1}, s_i]$ does not contain any element of $\{\lambda_k\}_{k\in J}$, so (3.33) implies that

$$E_{s_i} f - E_{s_{i-1}} f = 0 \quad \text{for all } i \le j_0 - 1.$$

Observe that $s_{j_0-1} < \lambda_{k_0-1} \le s_{j_0} < \lambda_{k_0}$. Hence, by (3.25), we have

$$E_{s_{j_0}} f - E_{s_{j_0-1}} f = \mathbb{P}_{k_0} f.$$

This, combined with (3.32), further yields

$$\begin{aligned}
\left\| \sum_{i=1}^{j_0} \varphi(s_i)(E_{s_i} f - E_{s_{i-1}} f) - \varphi(\lambda_{k_0-1}) \mathbb{P}_{k_0} f \right\|_{L^2(\mathcal{X})} &= \|\varphi(s_{j_0})(E_{s_{j_0}} f - E_{s_{j_0-1}} f) - \varphi(\lambda_{k_0-1}) \mathbb{P}_{k_0} f\|_{L^2(\mathcal{X})} \\
&= |\varphi(s_{j_0}) - \varphi(\lambda_{k_0-1})| \, \|\mathbb{P}_{k_0} f\|_{L^2(\mathcal{X})} \\
&\le \epsilon \|\mathbb{P}_{k_0} f\|_{L^2(\mathcal{X})}. \qquad (3.34)
\end{aligned}$$

Case 4: Consider the case $\mu(\mathcal{X}) = \infty$ and that for some $j_0 \in \{1, 2, \ldots, n\}$ the interval $(s_{j_0-1}, s_{j_0}]$ contains σ_*. Noting that $\lambda_k \ge \sigma_* > s_{j_0-1} \ge \alpha$ for all $k \in \mathbb{Z}$, we have by (3.25) that

$$E_{s_i} f - E_{s_{i-1}} f = 0 \quad \text{for all } i \le j_0 - 1$$

and

$$E_{s_{j_0}} f - E_{s_{j_0-1}} f = \sum_{\{k\in\mathbb{Z}:\, \alpha < \lambda_k \le s_{j_0}\}} \mathbb{Q}_k.$$

Further, by (3.32) and Lemma 3.2, we proceed as the proof of (3.9), thereby leading to

$$\begin{aligned}
&\left\| \sum_{i=1}^{j_0} \varphi(s_i)(E_{s_i} f - E_{s_{i-1}} f) - \sum_{\{k\in\mathbb{Z}:\, \alpha<\lambda_k\le s_{j_0}\}} \varphi(\lambda_k)\mathbb{Q}_k f \right\|_{L^2(\mathcal{X})}^2 \\
&= \left\| \varphi(s_{j_0}) \left(E_{s_{j_0}} f - E_{s_{j_0-1}} f \right) - \sum_{\{k\in\mathbb{Z}:\, \alpha<\lambda_k\le s_{j_0}\}} \varphi(\lambda_k)\mathbb{Q}_k f \right\|_{L^2(\mathcal{X})}^2 \\
&= \left\| \sum_{\{k\in\mathbb{Z}:\, \alpha<\lambda_k\le s_{j_0}\}} \left(\varphi(s_{j_0}) - \varphi(\lambda_k) \right) \mathbb{Q}_k f \right\|_{L^2(\mathcal{X})}^2
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{\{k\in\mathbb{Z}:\, \alpha<\lambda_k\le s_{j_0}\}} \left|\varphi(s_{j_0})-\varphi(\lambda_k)\right|^2 \|\mathbb{Q}_k f\|^2_{L^2(\mathcal{X})}\\
&\le \epsilon^2 \sum_{\{k\in\mathbb{Z}:\, \alpha<\lambda_k\le s_{j_0}\}} \|\mathbb{Q}_k f\|^2_{L^2(\mathcal{X})}\\
&\le \epsilon^2 \|f\|^2_{L^2(\mathcal{X})}. \qquad (3.35)
\end{aligned}
$$

Case 5: Assume that σ^* exists and $\sigma^* \in (s_{m_0-1}, s_{m_0}]$ for some $m_0 \in \{1,2,\ldots,n\}$. In this case, $\lambda_k \le \sigma_* \le s_{m_0} \le \beta$ for all $k\in\mathbb{Z}$. Thus, applying (3.25) and (3.5) yields

$$
E_{s_i} f - E_{s_{i-1}} f = 0 \quad \text{for all } i \ge m_0+1
$$

and

$$
E_{s_{m_0}} f - E_{s_{m_0-1}} f = \begin{cases} id - \sum_{\{k\in\mathbb{Z}:\, \lambda_k\le s_{m_0-1}\}} \mathbb{Q}_k = \sum_{\{k\in\mathbb{Z}:\, \beta\ge\lambda_k> s_{m_0-1}\}} \mathbb{Q}_k & \text{as } \mu(\mathcal{X})=\infty;\\ id - \left(\mathbb{P}_{k_0} + \sum_{\{k\ge k_0:\, \lambda_k\le s_{m_0-1}\}} \mathbb{Q}_k\right) = \sum_{\{k\ge k_0:\, \beta\ge\lambda_k> s_{m_0-1}\}} \mathbb{Q}_k & \text{as } \mu(\mathcal{X})<\infty, \end{cases}
$$

which, together with an argument similar to that of (3.35), also implies

$$
\begin{aligned}
&\left\|\sum_{i=m_0}^{n} \varphi(s_i)(E_{s_i} f - E_{s_{i-1}} f) - \sum_{\{k\in J:\, \beta\ge\lambda_k> s_{m_0-1}\}} \varphi(\lambda_k)\mathbb{Q}_k f\right\|^2_{L^2(\mathcal{X})}\\
&= \left\|\varphi(s_{m_0})\left(E_{s_{m_0}} f - E_{s_{m_0-1}} f\right) - \sum_{\{k\in J:\, \beta\ge\lambda_k> s_{m_0-1}\}} \varphi(\lambda_k)\mathbb{Q}_k f\right\|^2_{L^2(\mathcal{X})}\\
&= \left\|\sum_{\{k\in J:\, \beta\ge\lambda_k> s_{m_0-1}\}} \left(\varphi(s_{m_0})-\varphi(\lambda_k)\right)\mathbb{Q}_k f\right\|^2_{L^2(\mathcal{X})}\\
&= \sum_{\{k\in J:\, \beta\ge\lambda_k> s_{m_0-1}\}} \left|\varphi(s_{m_0})-\varphi(\lambda_k)\right|^2 \|\mathbb{Q}_k f\|^2_{L^2(\mathcal{X})}\\
&\le \epsilon^2 \sum_{\{k\in J:\, \beta\ge\lambda_k> s_{m_0-1}\}} \|\mathbb{Q}_k f\|^2_{L^2(\mathcal{X})}\\
&\le \epsilon^2 \|f\|^2_{L^2(\mathcal{X})}. \qquad (3.36)
\end{aligned}
$$

Based on the above arguments in **Case 1** through **Case 5**, let us prove (3.30) and (3.31) in the following two steps.

Step 1: proof of (3.30) under $\mu(\mathcal{X})=\infty$. If all the intervals $\{(s_{j-1}, s_j] : j = 1,\ldots,n\}$ do not contain $\{\sigma_*, \sigma^*\}$, then by **Case 1** and **Case 2**, we obtain

$$
\left\|\sum_{j=1}^{n} \varphi(s_j)\left(E_{s_j} f - E_{s_{j-1}} f\right) - \sum_{\{k\in\mathbb{Z}:\, \alpha<\lambda_k\le\beta\}} \varphi(\lambda_k)\mathbb{Q}_k f\right\|^2_{L^2(\mathcal{X})} = \left\|\sum_{j=1}^{n} \left(\varphi(s_j)-\varphi(\lambda_{k_j})\right)\mathbb{Q}_{k_j} f\right\|^2_{L^2(\mathcal{X})}
$$

Next, by (3.32) and Lemma 3.2, we proceed as the proof of (3.9), thereby obtaining

$$\begin{aligned}\left\|\sum_{j=1}^{n}\left(\varphi(s_j)-\varphi(\lambda_{k_j})\right)\mathbb{Q}_{k_j}f\right\|_{L^2(\mathcal{X})}^2 &= \sum_{j=1}^{n}\left|\varphi(s_j)-\varphi(\lambda_{k_j})\right|^2\left\|\mathbb{Q}_{k_j}f\right\|_{L^2(\mathcal{X})}^2\\ &\le \epsilon^2\sum_{j=1}^{n}\left\|\mathbb{Q}_{k_j}f\right\|_{L^2(\mathcal{X})}^2\\ &\le \epsilon^2\|f\|_{L^2(\mathcal{X})}^2.\end{aligned}$$

Thus, we obtain

$$\left\|\sum_{j=1}^{n}\varphi(s_j)\left(E_{s_j}f-E_{s_{j-1}}f\right)-\sum_{\{k\in\mathbb{Z}:\,\alpha<\lambda_k\le\beta\}}\varphi(\lambda_k)\mathbb{Q}_kf\right\|_{L^2(\mathcal{X})}\le\epsilon\|f\|_{L^2(\mathcal{X})},\tag{3.37}$$

which is just (3.30).

Since now $\mu(\mathcal{X})=\infty$, it might happen that either σ_* or σ^* is contained in some interval $(s_{j-1},s_j]$. In this case, we shall need **Case 4** or **Case 5**. Without loss of generality, we may as well assume that $\sigma_*\in(s_{j_0-1},s_{j_0}]$ and $\sigma^*\in(s_{m_0-1},s_{m_0}]$, where $1\le j_0<m_0\le n$. Then, by an argument similar to that of (3.37), we now have

$$\left\|\sum_{i=j_0+1}^{m_0-1}\varphi(s_i)\left(E_{s_i}f-E_{s_{i-1}}f\right)-\sum_{\{k\in\mathbb{Z}:\,s_{j_0}<\lambda_k\le s_{m_0-1}\}}\varphi(\lambda_k)\mathbb{Q}_kf\right\|_{L^2(\mathcal{X})}\le\epsilon\|f\|_{L^2(\mathcal{X})}.\tag{3.38}$$

This, combined with (3.35)-(3.36) and the Minkowski inequality, further derives

$$\begin{aligned}&\left\|\sum_{i=1}^{n}\varphi(s_i)(E_{s_i}f-E_{s_{i-1}}f)-\sum_{\{k\in\mathbb{Z}:\,\alpha<\lambda_k\le\beta\}}\varphi(\lambda_k)\mathbb{Q}_kf\right\|_{L^2(\mathcal{X})}\\ &\quad\le\left\|\sum_{i=1}^{j_0}\varphi(s_i)(E_{s_i}f-E_{s_{i-1}}f)-\sum_{\{k\in\mathbb{Z}:\,\alpha<\lambda_k\le s_{j_0}\}}\varphi(\lambda_k)\mathbb{Q}_kf\right\|_{L^2(\mathcal{X})}\\ &\qquad+\left\|\sum_{i=j_0+1}^{m_0-1}\varphi(s_i)(E_{s_i}f-E_{s_{i-1}}f)-\sum_{\{k\in\mathbb{Z}:\,s_{j_0}<\lambda_k\le s_{m_0-1}\}}\varphi(\lambda_k)\mathbb{Q}_kf\right\|_{L^2(\mathcal{X})}\\ &\qquad+\left\|\sum_{i=m_0}^{n}\varphi(s_i)(E_{s_i}f-E_{s_{i-1}}f)-\sum_{\{k\in J:\,\beta\ge\lambda_k>s_{m_0-1}\}}\varphi(\lambda_k)\mathbb{Q}_kf\right\|_{L^2(\mathcal{X})}\\ &\quad\le 3\epsilon\|f\|_{L^2(\mathcal{X})},\end{aligned}\tag{3.39}$$

as desired.

Step 2: proof of (3.31) **under** $\mu(\mathcal{X})<\infty$**.** Since $\alpha\to-\infty$, we may assume that $\alpha<\lambda_{k_0-1}$. Suppose that $\lambda_{k_0-1}\in(s_{j_0-1},s_{j_0}]$ for some $j_0\in\{1,2,\dots,n\}$. Thus, the estimate (3.34) in **Case 3** is true.

If for any $j>j_0$ the interval $(s_{j-1},s_j]$ does not contain σ^*, then by **Case 1** and **Case 2**, we argue as the proof of (3.37) and then obtain

$$\left\|\sum_{i=j_0+1}^{n}\varphi(s_i)\left(E_{s_i}f-E_{s_{i-1}}f\right)-\sum_{\{k\ge k_0:\,\alpha<\lambda_k\le\beta\}}\varphi(\lambda_k)\mathbb{Q}_kf\right\|_{L^2(\mathcal{X})}\le\epsilon\|f\|_{L^2(\mathcal{X})},$$

which, together with (3.34) and the Minkowski inequality, yields

$$\begin{aligned}
&\left\|\sum_{j=1}^{n}\varphi(s_i)\left(E_{s_i}f-E_{s_{i-1}}f\right)-\varphi(\lambda_{k_0-1})\mathbb{P}_{k_0}f-\sum_{\{k\in\mathbb{Z}:\,\alpha<\lambda_k\le\beta\}}\varphi(\lambda_k)\mathbb{Q}_kf\right\|_{L^2(\mathcal{X})}\\
&\le\left\|\sum_{i=1}^{j_0}\varphi(s_i)(E_{s_i}f-E_{s_{i-1}}f)-\varphi(\lambda_{k_0-1})\mathbb{P}_{k_0}f\right\|_{L^2(\mathcal{X})}\\
&\quad+\left\|\sum_{i=j_0+1}^{n}\varphi(s_i)\left(E_{s_i}f-E_{s_{i-1}}f\right)-\sum_{\{k\ge k_0:\,\alpha<\lambda_k\le\beta\}}\varphi(\lambda_k)\mathbb{Q}_kf\right\|_{L^2(\mathcal{X})}\\
&\le 2\epsilon\|f\|_{L^2(\mathcal{X})},
\end{aligned}$$

which proves (3.31).

Consider now the case when σ^* exists and $\sigma^*\in(s_{m_0-1},s_{m_0}]$ for some $m_0\in\{j_0+1,\dots,n\}$. Then we have the estimate (3.36) in **Case 5**. Again, by **Case 1** and **Case 2**, we now still have (3.38). Like the estimate of (3.39), but now we apply (3.34), (3.38) and (3.36), then we have

$$\begin{aligned}
&\left\|\sum_{i=1}^{n}\varphi(s_i)\left(E_{s_i}f-E_{s_{i-1}}f\right)-\varphi(\lambda_{k_0-1})\mathbb{P}_{k_0}f-\sum_{\{k\in\mathbb{Z}:\,\alpha<\lambda_k\le\beta\}}\varphi(\lambda_k)\mathbb{Q}_kf\right\|_{L^2(\mathcal{X})}\\
&\quad\le\left\|\sum_{i=1}^{j_0}\varphi(s_i)(E_{s_i}f-E_{s_{i-1}}f)-\varphi(\lambda_{k_0-1})\mathbb{P}_{k_0}f\right\|_{L^2(\mathcal{X})}\\
&\qquad+\left\|\sum_{j=j_0+1}^{m_0-1}\varphi(s_j)\left(E_{s_j}f-E_{s_{j-1}}f\right)-\sum_{\{k\ge k_0:\,\alpha<\lambda_k\le s_{m_0-1}\}}\varphi(\lambda_k)\mathbb{Q}_kf\right\|_{L^2(\mathcal{X})}\\
&\qquad+\left\|\sum_{i=m_0}^{n}\varphi(s_i)(E_{s_i}f-E_{s_{i-1}}f)-\sum_{\{k\in J:\,\beta\ge\lambda_k>s_{m_0-1}\}}\varphi(\lambda_k)\mathbb{Q}_kf\right\|_{L^2(\mathcal{X})}\\
&\quad\le 3\epsilon\|f\|_{L^2(\mathcal{X})}.
\end{aligned}$$

This also proves (3.31).

Summarizing all, we complete the proof of the theorem. □

Corollary 3.14. *Under the hypothesis of Theorem 3.13, let $\{E_\lambda\}_{\lambda\in\mathbb{R}}$ be a spectral resolution of the identity in $L^2(\mathcal{X})$ as defined in (3.25). Then, the operator $\mathcal{L}$ in (3.24) satisfies*

$$\mathcal{L}=\int_{[0,\infty)}\lambda\,dE_\lambda. \tag{3.40}$$

Moreover, for any $t\in(0,\infty)$, the operator P_t defined in (3.7) satisfies

$$P_t=\int_{[0,\infty)}e^{-t\lambda}\,dE_\lambda. \tag{3.41}$$

Proof. In (3.28), taking $\varphi(\lambda)=\lambda$ or $\varphi(\lambda)=e^{-t\lambda}$ yields (3.40) and (3.41), respectively. □

4 Construction of heat kernels via non-smooth MRA

Let $(\mathcal{X}, d, \mu)$ be a metric measure space satisfying (VD). The main aim of this section is to show Theorem 2.8; see Subsections 4.2 and 4.3. Moreover, in Subsection 4.4, we provide examples of heat kernels that are based on Haar wavelets in three different underlying spaces.

4.1 Dyadic cubes

It is known that on the Euclidean space $\mathbb{R}^n$ there is a standard dyadic system

$$\mathcal{D} := \left\{2^{-k}\left([0,1)^n + m\right) :\ k \in \mathbb{Z},\ m \in \mathbb{Z}^n\right\},$$

which enjoys many basic properties. For instance, any two different cubes are either disjoint or one is contained in another, all cubes of the same generation form a partition of $\mathbb{R}^n$, and etc. The first systematic work of constructing a dyadic structure on a doubling metric measure space was due to Christ [17]. Recent constructions of dyadic systems that have been developed in [2, 37, 38, 39] turn out to be very delicate, with the underlying space $(\mathcal{X}, d, \mu)$ satisfying only the geometrical doubling property (1.2) and the metric d being weaken to be a quasi-metric.

Assuming (VD), we review the construction of dyadic cubes from [37, 39]. Fix $\delta \in (0,1)$. Since d is a metric, it follows from [39, Theorem 2.4] that there is a set of *reference dyadic points*

$$\left\{z_\alpha^k :\ k \in \mathbb{Z},\ \alpha \in \mathcal{A}_k\right\}$$

enjoying the following properties:

$$\inf_{\alpha \neq \gamma} d(z_\alpha^k, z_\gamma^k) \geq \delta^k, \qquad \min_{\alpha \in \mathcal{A}_k} d(x, z_\alpha^k) < \delta^k \tag{4.1}$$

and

$$\left\{z_\alpha^k :\ \alpha \in \mathcal{A}_k\right\} \subset \left\{z_\alpha^{k+1} :\ \alpha \in \mathcal{A}_{k+1}\right\} \quad \text{for any}\ \ k \in \mathbb{Z}, \tag{4.2}$$

where $\mathcal{A}_k$ is an index set of countable cardinality. Let $\mathcal{G}$ be the family of parameter pairs from the reference dyadic points

$$\mathcal{G} := \{(k, \alpha) :\ k \in \mathbb{Z},\ \alpha \in \mathcal{A}_k\}. \tag{4.3}$$

For any $k \in \mathbb{Z}$, let

$$r_k \in [\delta^k/4,\ \delta^k/2].$$

According to [39, Lemma 2.8], there is a partial order $\leq$ in $\mathcal{G}$ enjoying the following property: each $(k+1, \gamma)$ satisfies $(k+1, \gamma) \leq (k, \alpha)$ for exactly one (k, α) in such a way that

$$d(z_\gamma^{k+1}, z_\alpha^k) < r_k \quad \Rightarrow \quad (k+1, \gamma) \leq (k, \alpha) \quad \Rightarrow \quad d(z_\gamma^{k+1}, z_\alpha^k) < 4r_k. \tag{4.4}$$

With the partial order $\leq$, it is constructed in [39, Theorem 2.9] a system of dyadic cubes on $(\mathcal{X}, d, \mu)$, by terms of the procedure of the construction in [37, Theorem 2.2, Proposition 2.11]. Indeed, for any $k \in \mathbb{Z}$ and $\alpha \in \mathcal{A}_k$, define the *preliminary dyadic cubes*

$$\widetilde{Q}_\alpha^k := \left\{z_\gamma^l :\ (l, \gamma) \leq (k, \alpha)\right\},$$

and then set

$$\overline{Q}_\alpha^k := \overline{Q_\alpha^k} \quad \text{and} \quad \widetilde{Q}_\alpha^k := \text{interior}\, \overline{Q}_\alpha^k = \left(\bigcup_{\gamma\neq\alpha} \overline{Q_\gamma^k}\right)^{\complement},$$

which are called the *closed* and *open dyadic cubes*, respectively. The forthcoming theorem collects some basic properties of the dyadic cubes on $(\mathcal{X}, d, \mu)$.

Theorem 4.1 ([39, 37])**.** *Let $(\mathcal{X}, d, \mu)$ be a metric measure space satisfying (VD). Let $\delta \in (0, 1/60)$ and M be as in (1.2). Suppose that the reference dyadic points $\{z_\alpha^k : k \in \mathbb{Z}, \alpha \in \mathcal{A}_k\}$ satisfy (4.1) and (4.2). Then there exists a family of Borel sets (called dyadic cubes)*

$$\mathcal{D} := \left\{Q_\alpha^k : k \in \mathbb{Z}, \alpha \in \mathcal{A}_k\right\}$$

such that the following assertions hold:

(i) *for any $k \in \mathbb{Z}$ and $\alpha \in \mathcal{A}_k$, the interior of Q_α^k is $\widetilde{Q}_\alpha^k$, and the closure of Q_α^k is $\overline{Q}_\alpha^k$;*

(ii) *for any $k \in \mathbb{Z}$ and $\alpha \in \mathcal{A}_k$, $\widetilde{Q}_\alpha^k$ and $\overline{Q}_\alpha^k$ are one another's interior and closure;*

(iii) *for any $k \in \mathbb{Z}$, $\{Q_\alpha^k\}_{\alpha\in\mathcal{A}_k}$ are disjoint and $\mathcal{X} = \bigcup_{\alpha\in\mathcal{A}_k} Q_\alpha^k$;*

(iv) *if $j \geq k$, $\alpha \in \mathcal{A}_k$ and $\gamma \in \mathcal{A}_j$, then either $Q_\gamma^j \subset Q_\alpha^k$ or $Q_\gamma^j \cap Q_\alpha^k = \emptyset$;*

(v) *for any $k \in \mathbb{Z}$, $\alpha \in \mathcal{A}_k$ and $j < k$, there exists a unique $\gamma \in \mathcal{A}_j$ such that $Q_\alpha^k \subset Q_\gamma^j$;*

(vi) *for any $k \in \mathbb{Z}$ and $\alpha \in \mathcal{A}_k$, $B(z_\alpha^k, 5^{-1}\delta^k) \subset Q_\alpha^k \subset B(z_\alpha^k, 3\delta^k)$.*

For any $k \in \mathbb{Z}$, denote by $\mathcal{D}_k := \{Q_\alpha^k : \alpha \in \mathcal{A}_k\}$ the set of all dyadic cubes of k-th generation.

Next, we recall the notion of *quadrants*. Being more precisely, for any $Q \in \mathcal{D}$, the set

$$C(Q) := \bigcup_{\substack{Q'\in\mathcal{D} \\ Q'\supseteq Q}} Q' \tag{4.5}$$

is called a *quadrant* of $\mathcal{X}$ containing Q. In other words, $C(Q)$ is the union of all ancestors of Q. According to [1, Lemma 2.2], the quadrants have the following properties.

Lemma 4.2 ([1])**.** *Let $(\mathcal{X}, d, \mu)$ be a metric measure space satisfying (VD). Suppose that $\mathcal{D}$ is a dyadic system as in Theorem 4.1. Then the family of quadrants defined in (4.5) satisfies the following properties:*

(i) *for each quadrant C, the triple (C, d, μ) satisfies (VD);*

(ii) *any two intersecting quadrants coincide;*

(iii) *$\mathcal{X}$ is a disjoint union of finitely many quadrants*;

(iv) *if $\mu(\mathcal{X}) < \infty$ then $\mathcal{X}$ coincides with one quadrant C, where C coincides with some dyadic cube $Q \in \mathcal{D}$;*

(v) *if $\mu(\mathcal{X}) = \infty$ then for every quadrant C we have $\mu(C) = \infty$.*

Remark 4.3. If two points x, y belong to the same dyadic cube then, clearly, they belong to one quadrant. Conversely, if x, y belong to the same quadrant C then there is a dyadic cube containing both x, y.

According to [13, Section 5.2], each dyadic system $\mathcal{D}$ that is stated in Theorem 4.1 gives rise to an ultra-metric $d_{\mathcal{D}}$ on $\mathcal{X}$. A metric d on $\mathcal{X}$ is called an *ultra-metric* if for any $x, y, z \in \mathcal{X}$,

$$d(x, y) \leq \max\{d(x, z), d(z, y)\}.$$

Of course, any ultra-metric is a metric. Usually a metric must take non-negative real values, but we will allow an ultra-metric to take also the value $+\infty$.

Definition 4.4. For any two distinct points $x, y \in \mathcal{X}$ that belong to the same quadrant, denote by $Q_{x,y}$ the smallest dyadic cube from $\mathcal{D}$ containing both x and y; then denote by $k_{x,y}$ the unique integer k such that $Q_{x,y} \in \mathcal{D}_k$. If x, y do not belong to the same quadrant then we set $k_{x,y} = -\infty$. Finally, if $x = y$ then we set $k_{x,y} = +\infty$. For any $x, y \in \mathcal{X}$, set

$$d_{\mathcal{D}}(x, y) := \delta^{k_{x,y}}.$$

In particular, if $x = y$ then $d_{\mathcal{D}}(x, x) = 0$, and if x, y do not belong to one quadrant then $d_{\mathcal{D}}(x, y) = \infty$.

The following results are from [13, Section 5.2].

Proposition 4.5. *Let $(\mathcal{X}, d, \mu)$ be a metric measure space satisfying (VD). Suppose that $d_{\mathcal{D}}$ is defined as in Definition 4.4. Then, the following hold:*

(i) *$d_{\mathcal{D}}$ is an ultra-metric on $\mathcal{X}$.*

(ii) *There exists a positive constant C such that $d(x, y) \leq C d_{\mathcal{D}}(x, y)$ for all $x, y \in \mathcal{X}$.*

(iii) *For any $x \in \mathcal{X}$ and $r \in (0, \infty)$, let*

$$\overline{B}_{\mathcal{D}}(x, r) := \{y \in \mathcal{X} : \ d_{\mathcal{D}}(y, x) \leq r\}.$$

Choose $k \in \mathbb{Z}$ to satisfy $\delta^k \leq r < \delta^{k-1}$ and let $Q_{x,r} \in \mathcal{D}_k$ be the unique dyadic cube of k-th generation that contains x. Then

$$\overline{B}_{\mathcal{D}}(x, r) = Q_{x,r} \qquad \text{and} \qquad \mu(\overline{B}_{\mathcal{D}}(x, r)) \simeq V(x, r).$$

4.2 A stochastic complete heat kernel via non-smooth MRA

Let $(\mathcal{X}, d, \mu)$ be a metric measure space satisfying (VD). With all the notation as in Theorem 4.1, we define for any $k \in \mathbb{Z}$ that

$$V_k := \overline{\operatorname{span}\left\{\mu(Q)^{-\frac{1}{2}}\mathbf{1}_Q : \ Q \in \mathcal{D}_k\right\}}^{\|\cdot\|_{L^2(\mathcal{X})}}. \tag{4.6}$$

Clearly,

$$V_k = \left\{f \in L^2(\mathcal{X}) : \ f \text{ is } \mu\text{-a.e. constant on each } Q \in \mathcal{D}_k\right\}.$$

In particular, if $\mu(\mathcal{X}) < \infty$, then there exists the largest integer, denoted by $m_{\mathcal{X}}$, such that

$$\text{when } k \leq m_{\mathcal{X}}, \text{ the space } \mathcal{X} \text{ itself is the unique dyadic cube in } \mathcal{D}_k, \tag{4.7}$$

thereby leading to that

$$V_k = \left\{ f \in L^2(\mathcal{X}) : \ f \text{ is } \mu\text{-a.e. constant on } \ \mathcal{X} \right\} \quad \text{when } k \le m_{\mathcal{X}}. \tag{4.8}$$

It was proved in [1] that the sequence $\{V_k\}_{k\in\mathbb{Z}}$ in (4.6) forms a standard MRA in $L^2(\mathcal{X})$ in the sense of Definition 2.1, where the Riesz basis of each V_k is given by

$$\left\{ \mu(Q)^{-\frac{1}{2}} \mathbf{1}_Q : \ Q \in \mathcal{D}_k \right\}.$$

For any $k \in \mathbb{Z}$, we denote by W_k the orthogonal complement of V_k in V_{k+1}. Following Definition 2.2, for any $k \in \mathbb{Z}$, denote by $\mathbb{E}_k$ and $\mathbb{D}_k$ the orthogonal projectors from $L^2(\mathcal{X})$ onto V_k and W_k, respectively. As a consequence of Lemma (3.2), for any $k \in \mathbb{Z}$ and $f \in L^2(\mathcal{X})$, we can easily deduce that

$$\mathbb{E}_k f = \sum_{Q\in\mathcal{D}_k} \left(\frac{1}{\mu(Q)} \int_Q f \, d\mu \right) \mathbf{1}_Q,$$

which is also known as the *conditional expectation operator*. Note that for any $k \in \mathbb{Z}$ and $x, y \in \mathcal{X}$,

$$\mathbb{E}_k(x, y) = \sum_{Q\in\mathcal{D}_k} \frac{\mathbf{1}_Q(x)\mathbf{1}_Q(y)}{\mu(Q)} \ge 0. \tag{4.9}$$

Meanwhile, the projector $\mathbb{D}_k$, which is usually called the *martingale difference operator*, satisfies

$$\mathbb{D}_k = \mathbb{E}_{k+1} - \mathbb{E}_k \qquad \text{and} \qquad \mathbb{D}_k(x, y) = \mathbb{E}_{k+1}(x, y) - \mathbb{E}_k(x, y). \tag{4.10}$$

To begin with, we give the following properties of $\mathbb{E}_k$ and $\mathbb{D}_k$ in (4.9) and (4.10).

Lemma 4.6. *Let $(\mathcal{X}, d, \mu)$ be a metric measure space satisfying (VD). Define $\mathbb{E}_k$ and $\mathbb{D}_k$ as in (4.9) and (4.10).*

(i) *If x and y are in different quadrants of $\mathcal{X}$, then $\mathbb{E}_k(x, y) = \mathbb{D}_k(x, y) = 0$ for all $k \in \mathbb{Z}$.*

(ii) *If x and y are in the same quadrants of $\mathcal{X}$, then $\mathbb{E}_k(x, y) = \mathbb{D}_k(x, y) = 0$ for all $k > k_{x,y}$, where $k_{x,y} \in \mathbb{Z}$ is determined by $d_{\mathcal{D}}(x, y)$ as in Definition 4.4.*

(iii) *If $\mu(\mathcal{X}) = \infty$, then for all $x, y \in \mathcal{X}$,*
$$\lim_{k\to-\infty} \mathbb{E}_k(x, y) = 0.$$

(iv) *The MRA determined by $\{V_k\}_{k\in\mathbb{Z}}$ in (4.6) is admissible in the sense of Definition 2.3.*

(v) *There exist positive constants C and C_0, independent of $k \in \mathbb{Z}$ and x, $y \in \mathcal{X}$, such that*

$$\mathbb{E}_k(x, y) + |\mathbb{D}_k(x, y)| \le C \frac{\mathbf{1}_{\{d(x,y) \le C_0 d_{\mathcal{D}}(x,y) \le C_0 \delta^k\}}}{V(x, \delta^k)}.$$

Proof. For all $x, y \in \mathcal{X}$, we recall that

$$\mathbb{E}_k(x, y) = \sum_{Q\in\mathcal{D}_k} \frac{\mathbf{1}_Q(x)\,\mathbf{1}_Q(y)}{\mu(Q)}.$$

If x and y are in two different quadrants of $\mathcal{X}$, there does not exist a dyadic cube Q contains both x and y simultaneously, which induces

$$\mathbb{E}_k(x, y) = \mathbb{D}_k(x, y) = 0$$

and, hence, both (i) and (iii) hold.

Now, suppose that x and y are in the same quadrant. By the definition of $k_{x,y}$, we see that if $k > k_{x,y}$ then any dyadic cube $Q \in \mathcal{D}_k$ can not contain x and y simultaneously, which implies that $\mathbb{E}_k(x, y) = 0$ and $\mathbb{D}_k(x, y) = \mathbb{E}_{k+1}(x, y) - \mathbb{E}_k(x, y) = 0$. This proves (ii).

Still assume that x and y are in the same quadrant. If $k \le k_{x,y}$, then there exists a unique dyadic cube $Q_k \in \mathcal{D}_k$ containing both x and y simultaneously. Then

$$Q_k \supseteq Q_{x,y} \supseteq \{x, y\},$$

where we recall that $Q_{x,y} \in \mathcal{D}_{k_{x,y}}$ and it is the smallest dyadic cube from $\mathcal{D}$ containing both x and y. Invoking the definition of quadrant in (4.5), we find that

$$C(Q_{x,y}) = \bigcup_{\{k\in\mathbb{Z}:\ k\le k_{x,y}\}} Q_k,$$

whose μ-measure is ∞ by terms of Lemma 4.2(v). This further induces that

$$\lim_{k\to-\infty} \mathbb{E}_k(x, y) = \lim_{k\to-\infty} \frac{1}{\mu(Q_k)} = \frac{1}{\mu(C(Q_{x,y}))} = 0.$$

Thus, we obtain (iii). As a consequence of (iii) and Definition 2.3, we easily obtain (iv).

It remains to show (v). By (ii), if x and y are in the same quadrant such that $\mathbb{E}_k(x, y) \neq 0$, then we have not only $k_{x,y} \ge k$ and $d_{\mathcal{D}}(x, y) \le \delta^k$, but also (see Theorem 4.1)

$$d(x, y) \le \operatorname{diam} Q_{x,y} < 6\delta^{k_{x,y}} = 6 d_{\mathcal{D}}(x, y).$$

Moreover, the condition (VD) implies

$$\mu(Q) \approx \mu(B(x, \delta^k)) \approx \mu(B(y, \delta^k)).$$

Therefore,

$$\mathbb{E}_k(x, y) \lesssim \frac{\mathbf{1}_{\{d(x,y)\le 6 d_{\mathcal{D}}(x,y)\le 6\delta^k\}}}{V(x, \delta^k)}.$$

Obviously, the above estimate holds for $|\mathbb{D}_k(x, y)|$ by terms of $\mathbb{D}_k = \mathbb{E}_{k+1} - \mathbb{E}_k$ and (VD). □

Theorem 4.7. *Let $(\mathcal{X}, d, \mu)$ be a metric measure space satisfying (VD). Assume that $\{V_k\}_{k\in\mathbb{Z}}$ is the MRA defined in (4.6). For any $k \in \mathbb{Z}$, let $W_k = V_{k+1} \ominus V_k$ and denote by $\mathbb{D}_k$ the orthogonal projector from $L^2(\mathcal{X})$ to W_k as in (4.10). Fix $\beta \in (0, \infty)$. For any $t \in (0, \infty)$ and $x, y \in \mathcal{X}$, define*

$$p_t(x, y) := \begin{cases} \displaystyle\sum_{k\in\mathbb{Z}} e^{-t\delta^{-\beta k}} \mathbb{D}_k(x, y) & \text{as } \mu(\mathcal{X}) = \infty; \\ \displaystyle\frac{1}{\mu(\mathcal{X})} + \sum_{k=m_{\mathcal{X}}}^{\infty} e^{-t\delta^{-\beta k}} \mathbb{D}_k(x, y) & \text{as } \mu(\mathcal{X}) < \infty, \end{cases} \tag{4.11}$$

where $m_{\mathcal{X}}$ is as in (4.7). Then $\{p_t\}_{t>0}$ is a stochastic complete heat kernel.

Proof. By Lemma 4.6(iv), we know that $\{V_k\}_{k\in\mathbb{Z}}$ in (4.6) is an admissible MRA. Let $\delta \in (0,1)$ be as in Theorem 4.1. Let $\{\lambda_k\}_{k\in\mathbb{Z}}$ be as given in Example 2.5, which is a family of admissible spectrum. Assume for the moment that under the situation $\mu(\mathcal{X}) < \infty$ it holds that

$$\mathbb{E}_{m_{\mathcal{X}}}(x,y) = \frac{1}{\mu(\mathcal{X})} \quad \text{for all } \ x,y \in \mathcal{X}. \tag{4.12}$$

Once we have obtained (4.12), then the two expressions of p_t in (2.4) and (4.11) coincide with each other, which enables us to apply Theorems 3.9 and 3.10 (with k_0 therein taken to be $m_{\mathcal{X}}$ now) to deduce that $\{p_t\}_{t>0}$ in (4.11) is a stochastic complete heat kernel.

It remains to verify (4.12) under $\mu(\mathcal{X}) < \infty$. Indeed, for any $k \le m_{\mathcal{X}}$, we have by (4.7) and (4.8) that the orthogonal projector $\mathbb{E}_k : \ L^2(\mathcal{X}) \to V_k$ fulfills that

$$\mathbb{E}_k f = c_{k,f} \quad \text{for all } f \in L^2(\mathcal{X}),$$

where $c_{k,f}$ is a constant. To determine $c_{k,f}$, observing that $f - \mathbb{E}_k f \perp V_k$ for all $f \in L^2(\mathcal{X})$ (see the discussion before Definition 2.2), we then derive that for any constant $c \in \mathbb{C}$,

$$0 = \langle c,\ f - \mathbb{E}_k f\rangle = \int_{\mathcal{X}} c(f - c_{k,f})\, d\mu = c\left(\int_{\mathcal{X}} f\, d\mu - c_{k,f}\mu(\mathcal{X})\right),$$

which in turn gives

$$\mathbb{E}_k f = c_{k,f} = \frac{1}{\mu(\mathcal{X})} \int_{\mathcal{X}} f\, d\mu \quad \text{for all } f \in L^2(\mathcal{X}).$$

In particular, taking $k = m_{\mathcal{X}}$ implies (4.12). This ends the proof. □

The following lemma gives a representation of $\{p_t\}_{t>0}$ in (4.11) by using $\mathbb{E}_k(x,y)$ and $d_{\mathcal{D}}$.

Lemma 4.8. *Let $(\mathcal{X}, d, \mu)$ be a metric measure space satisfying (VD). Suppose that $\beta \in (0,\infty)$ and $\{p_t\}_{t>0}$ is the heat kernel defined in (4.11). Then, for any $t \in (0,\infty)$ and $x,y \in \mathcal{X}$,*

$$p_t(x,y) = \sum_{k\in\mathbb{Z}} \left(e^{-t\delta^{-(k-1)\beta}} - e^{-t\delta^{-k\beta}}\right) \mathbb{E}_k(x,y) \tag{4.13}$$

and

$$p_t(x,y) = \int_{d_{\mathcal{D}}(x,y)}^{\infty} \frac{de^{-tr^{-\beta}}}{\mu(\overline{B}_{\mathcal{D}}(x,r))} = p_t^{\mathcal{D}}(x,y), \tag{4.14}$$

where $p_t^{\mathcal{D}}$ is defined in (1.7), and we take it for granted that the right hand side integral in (4.14) is 0 when $d_{\mathcal{D}}(x,y) = \infty$.

Proof. Let us first show (4.13) under the case $\mu(\mathcal{X}) = \infty$. For any $t \in (0,\infty)$ and $x,y \in \mathcal{X}$, observing each $\mathbb{E}_k(x,y) \ge 0$ and $\int_{\mathcal{X}} \mathbb{E}_k(x,y)\, d\mu(y) = 1$, we take $\lambda_k = \delta^{-k\beta}$ and $\mathbb{P}_k = \mathbb{E}_k$ in (3.13) and then obtain

$$p_t(x,y) = \lim_{N\to\infty} \left(e^{-t\delta^{-N\beta}}\mathbb{E}_{N+1}(x,y) + \sum_{k=-N}^{N} \left(e^{-t\delta^{-(k-1)\beta}} - e^{-t\delta^{-k\beta}}\right) \mathbb{E}_k(x,y) - e^{-t\delta^{(N+1)\beta}}\mathbb{E}_{-N}(x,y)\right). \tag{4.15}$$

If x and y are in different quadrants of $\mathcal{X}$, then we have by Lemma 4.6(i) that $\mathbb{E}_k(x, y) = 0$ for all $k \in \mathbb{Z}$. If x and y are in the same quadrant of $\mathcal{X}$, then we have by Lemma 4.6(ii) that $\mathbb{E}_k(x, y) = 0$ for all $k > k_{x,y}$. Whatever the cases, for the first term in the right hand side of (4.15), we always have

$$\lim_{N\to\infty}\left(e^{-t\delta^{-N\beta}}\mathbb{E}_{N+1}(x, y)\right) \le \lim_{N\to\infty}\mathbb{E}_{N+1}(x, y) = 0. \tag{4.16}$$

For the third term in the right hand side of (4.15), applying Lemma 4.6(iii) and $\mu(\mathcal{X}) = \infty$ yields

$$\lim_{N\to\infty}\left(e^{-t\delta^{(N+1)\beta}}\mathbb{E}_{-N}(x, y)\right) = \lim_{N\to\infty}\mathbb{E}_{-N}(x, y) = 0. \tag{4.17}$$

Inserting (4.16) and (4.17) into (4.15) yields the desired equality in (4.13) when $\mu(\mathcal{X}) = \infty$.

The proof of (4.13) under $\mu(\mathcal{X}) < \infty$ is similar. In this case, applying (3.14) with $k_0 = m_{\mathcal{X}}$, $\lambda_k = \delta^{-k\beta}$ and $\mathbb{P}_k = \mathbb{E}_k$ therein, we write

$$p_t(x, y) = \lim_{N\to\infty}\left(e^{-t\delta^{-N\beta}}\mathbb{E}_{N+1}(x, y) + \sum_{k=m_{\mathcal{X}}}^{N}\left(e^{-t\delta^{-(k-1)\beta}} - e^{-t\delta^{-k\beta}}\right)\mathbb{E}_k(x, y)\right). \tag{4.18}$$

Observe that (4.16) remains true when $\mu(\mathcal{X}) < \infty$. From this and (4.18), it follows that (4.13) remains true when $\mu(\mathcal{X}) < \infty$.

Finally, let us verify (4.14). Consider first the case when x and y are in the same quadrant but $x \ne y$. With all the notation as in Definition 4.4, we write for some $k_{x,y} \in \mathbb{Z}$ that

$$d_{\mathcal{D}}(x, y) = \delta^{k_{x,y}}.$$

If $k > k_{x,y}$, then by Definition 4.4, any dyadic cube Q in $\mathcal{D}_k$ can not contain x and y simultaneous, which implies

$$\mathbb{E}_k(x, y) = 0.$$

If $k \le k_{x,y}$ and $\delta^k \le r < \delta^{k-1}$, then we use Proposition 4.5(iii) to deduce that

$$Q_{x,r} = \overline{B}_{\mathcal{D}}(x, r),$$

where $Q_{x,r}$ denotes the unique dyadic cube in $\mathcal{D}_k$ that contains x. Since $k \le k_{x,y}$, it follows that $Q_{x,r}$ must contain y, thereby leading to that

$$\mathbb{E}_k(x, y) = \sum_{Q\in\mathcal{D}_k}\frac{\mathbf{1}_Q(x)\,\mathbf{1}_Q(y)}{\mu(Q)} = \frac{1}{\mu(Q_{x,r})} = \frac{1}{\mu(\overline{B}_{\mathcal{D}}(x, r))}.$$

Thus, applying (4.13) yields

$$\begin{aligned} p_t(x, y) &= \sum_{k\le k_{x,y}}\left(e^{-t\delta^{-(k-1)\beta}} - e^{-t\delta^{-k\beta}}\right)\mathbb{E}_k(x, y) \\ &= \sum_{k\le k_{x,y}}\left(\int_{[\delta^k,\,\delta^{k-1})} de^{-tr^{-\beta}}\right)\mathbb{E}_k(x, y) \\ &= \sum_{k\le k_{x,y}}\int_{[\delta^k,\,\delta^{k-1})}\frac{de^{-tr^{-\beta}}}{\mu(\overline{B}_{\mathcal{D}}(x, r))} \end{aligned}$$

$$
\begin{aligned}
&= \int_{d_{\mathcal{D}}(x,y)}^{\infty} \frac{de^{-tr^{-\beta}}}{\mu(\overline{B}_{\mathcal{D}}(x,r))} \\
&= p_t^{\mathcal{D}}(x,y),
\end{aligned}
$$

as desired.

If $x = y$, then $k_{x,y} = +\infty$ and $d_{\mathcal{D}}(x,y) = 0$. In this case, the above arguments remain valid with the summation $\sum_{k \le k_{x,y}}$ therein replaced by $\sum_{k\in\mathbb{Z}}$.

If x and y are in different quadrants, then by Lemma 4.6(i) we know that each $\mathbb{E}_k(x,y) = 0$ so that $p_t^{\mathcal{D}}(x,y) = 0$ by terms of (4.13), which implies that (4.14) holds with both sides being zero.

Summarizing all, we obtain (4.14). □

Remark 4.9. Let M_Q denote the number of children of a dyadic cube $Q \in \mathcal{D}$, which depends on Q but is uniformly bounded (cf. [2, 37]). Let us use the family of Haar functions

$$\left\{h_u^Q : \; Q \in \mathcal{D}, \; u \in \{1, 2, \dots, M_Q - 1\}\right\},$$

that forms an orthogonal basis in $L^2(\mathcal{X})$, which was constructed in [41]. Note that if Q has itself as its only child, then there is no Haar functions associated to Q. For any $f \in L^2(\mathcal{X})$, we have by [41, (4.6)] that each $\mathbb{D}_k f$ has the following decomposition:

$$\mathbb{D}_k f = \sum_{Q\in\mathcal{D}_k} \sum_{u=1}^{M_Q-1} \langle f, h_u^Q\rangle h_u^Q \quad \text{in} \;\; L^2(\mathcal{X}).$$

In other words, for any $k \in \mathbb{Z}$, the MRA $\{V_k\}_{k\in\mathbb{Z}}$ in (4.6) fulfills that

$$W_k = V_{k+1} \ominus V_k = \overline{\operatorname{span}\left\{h_u^Q : \; Q \in \mathcal{D}_k, \; u = \{1, 2, \dots, M_Q - 1\}\right\}}^{\|\cdot\|_{L^2(\mathcal{X})}}.$$

Clearly, the integral kernel $\mathbb{D}_k(x,y)$ of $\mathbb{D}_k$ satisfies

$$\mathbb{D}_k(x,y) = \sum_{Q\in\mathcal{D}_k} \sum_{u=1}^{M_Q-1} h_u^Q(x) h_u^Q(y).$$

Thus, the heat kernel $\{p_t^{\mathcal{D}}\}_{t>0}$ in (4.11) can be expressed via Haar functions as follows:

$$p_t^{\mathcal{D}}(x,y) = \frac{1}{\mu(\mathcal{X})} + \sum_{k\in\mathbb{Z}} \sum_{Q\in\mathcal{D}_k} \sum_{u=1}^{M_Q-1} e^{-t\delta^{-\beta k}} h_u^Q(x) h_u^Q(y). \tag{4.19}$$

Moreover, invoking (2.2) and (2.3), we deduce from (3.24) that the generator of $\{p_t^{\mathcal{D}}\}_{t>0}$, denoted by $\mathcal{L}_{\mathcal{D}}$, has the following expression:

$$\mathcal{L}_{\mathcal{D}} f = \sum_{k\in\mathbb{Z}} \delta^{-k\beta} \mathbb{D}_k f = \sum_{k\in\mathbb{Z}} \sum_{Q\in\mathcal{D}_k} \sum_{u=1}^{M_Q-1} \delta^{-k\beta} \langle f, h_u^Q\rangle h_u^Q$$

with domain

$$\operatorname{Dom}(\mathcal{L}_{\mathcal{D}}) = \left\{ f \in L^2(\mathcal{X}) : \; \|\mathcal{L}_{\mathcal{D}} f\|_{L^2(\mathcal{X})} = \sum_{k\in\mathbb{Z}} \sum_{Q\in\mathcal{D}_k} \sum_{u=1}^{M_Q-1} \delta^{-2k\beta} |\langle f, h_u^Q\rangle|^2 < \infty \right\}.$$

4.3 Stable-like estimates

Due to (4.14), we can apply (1.8)-(1.9) (see [13]) and obtain the following two-sided estimate of the heat kernel $\{p_t^{\mathcal{D}}\}_{t>0}$.

Theorem 4.10. *Let $\beta \in (0,\infty)$ and $(\mathcal{X}, d, \mu)$ be a metric measure space satisfying (VD). Suppose that the two heat kernels $\{p_t\}_{t>0}$ and $\{p_t^{\mathcal{D}}\}_{t>0}$ are defined, respectively, as in (4.11) and (1.7). Then, for any $t \in (0,\infty)$ and $x, y \in \mathcal{X}$,*

$$0 \le p_t(x,y) = p_t^{\mathcal{D}}(x,y) \simeq \frac{1}{V(x,\ t^{1/\beta} + d_{\mathcal{D}}(x,y))} \left(\frac{t^{1/\beta}}{t^{1/\beta} + d_{\mathcal{D}}(x,y)}\right)^{\beta}. \tag{4.20}$$

In particular, there exists a positive constant C such that for any $t \in (0,\infty)$ and $x, y \in \mathcal{X}$,

$$0 \le p_t(x,y) = p_t^{\mathcal{D}}(x,y) \le \frac{C}{V(x,\ t^{1/\beta} + d(x,y))} \left(\frac{t^{1/\beta}}{t^{1/\beta} + d(x,y)}\right)^{\beta}. \tag{4.21}$$

For the completeness of the paper, instead of using (1.8)-(1.9), we will use Lemma 4.6(v) and (4.13) to give a direct proof of the two-sided estimate (4.20) of $\{p_t\}_{t>0}$.

Proof of Theorem 4.10. Note that (4.21) follows directly from (4.20) and the fact $d \lesssim d_{\mathcal{D}}$ (see Proposition 4.5). Thus, we only need to prove (4.20).

By (4.13), the heat kernel $p_t(x,y)$ in (4.11) is nonnegative. If x and y are in different quadrants of $\mathcal{X}$, then $d_{\mathcal{D}}(x,y) = \infty$ and we have by Lemma 4.6 that $\mathbb{E}_k(x,y) = 0$ for all $k \in \mathbb{Z}$, thereby leading to

$$p_t(x,y) = 0$$

and, hence, (4.20) holds. So, it suffices to show (4.20) in the case when x and y are in the same quadrant of $\mathcal{X}$.

Part 1: verifying the upper estimate in (4.20) under the case $\mu(\mathcal{X}) = \infty$. Since x, y are in the same quadrant of $\mathcal{X}$, we let $k_{x,y} \in \mathbb{Z}$ such that $d_{\mathcal{D}}(x,y) = \delta^{k_{x,y}}$. By Lemma 4.6(ii), if $\mathbb{E}_k(x,y) \neq 0$ or $\mathbb{D}_k(x,y) \neq 0$, then

$$k \le k_{x,y}. \tag{4.22}$$

Moreover, by Lemma 4.6(v), for any $k \in \mathbb{Z}$ and $x,\ y \in \mathcal{X}$, there is

$$0 \le \mathbb{E}_k(x,y) \lesssim \frac{\mathbf{1}_{\{d(x,y) \le 2C_0 d_{\mathcal{D}}(x,y) \le 2C_0\delta^k\}}}{V(x,\delta^k)}. \tag{4.23}$$

Let k_t be the unique integer such that

$$\delta^{k_t\beta} \le t < \delta^{(k_t-1)\beta}. \tag{4.24}$$

Since $\mu(\mathcal{X}) = \infty$, we have by (4.11) that

$$\begin{aligned} p_t(x,y) &= \sum_{k \ge k_t} e^{-t\delta^{-\beta k}} \mathbb{D}_k(x,y) + \sum_{k \le k_t - 1} e^{-t\delta^{-\beta k}} \Big(\mathbb{E}_{k+1}(x,y) - \mathbb{E}_k(x,y)\Big) \\ &= \sum_{k \ge k_t} e^{-t\delta^{-\beta k}} \mathbb{D}_k(x,y) + e^{-t\delta^{-\beta(k_t-1)}} \mathbb{E}_{k_t}(x,y) + \sum_{k \le k_t - 1} \Big(e^{-t\delta^{-\beta(k-1)}} - e^{-t\delta^{-\beta k}}\Big) \mathbb{E}_k(x,y) \end{aligned}$$

$$=: Z_1 + Z_2 + Z_3. \tag{4.25}$$

It suffices to consider each Z_i, where $i \in \{1,2,3\}$.

First, we estimate Z_1. If $k \ge k_t$, then by (4.22), we have $d_{\mathcal{D}}(x,y) = \delta^{k_{x,y}} \le \delta^k \le \delta^{k_t} \le t^{1/\beta}$. By this and (4.23), we obtain

$$|Z_1| \le \sum_{\delta^{k\beta} \le t} e^{-t\delta^{-\beta k}} |\mathbb{D}_k(x,y)| \lesssim \sum_{\delta^{k\beta} \le t} e^{-t\delta^{-\beta k}} \frac{\mathbf{1}_{\{d(x,y) \le 2C_0 d_{\mathcal{D}}(x,y) \le 2C_0\delta^k\}}}{V(x,\delta^k)}.$$

Further, since $\delta^k \le t^{1/\beta}$ and $d_{\mathcal{D}}(x,y) \le \delta^k$, we deduce from (VD) that

$$\frac{V(x, t^{1/\beta} + d_{\mathcal{D}}(x,y))}{V(x,\delta^k)} \approx \frac{V(x,t^{1/\beta})}{V(x,\delta^k)} \lesssim \left(\frac{t^{1/\beta}}{\delta^k}\right)^n,$$

thereby leading to

$$\begin{aligned}
|Z_1| &\lesssim \frac{\mathbf{1}_{\{d_{\mathcal{D}}(x,y) \le t^{1/\beta}\}}}{V(x, t^{1/\beta} + d_{\mathcal{D}}(x,y))} \sum_{\delta^{k\beta} \le t} e^{-t\delta^{-\beta k}} \left(\frac{t^{1/\beta}}{\delta^k}\right)^n \\
&\lesssim \frac{\mathbf{1}_{\{d_{\mathcal{D}}(x,y) \le t^{1/\beta}\}}}{V(x, t^{1/\beta} + d_{\mathcal{D}}(x,y))} \\
&\lesssim \frac{1}{V(x,\ t^{1/\beta} + d_{\mathcal{D}}(x,y))} \left(\frac{t^{1/\beta}}{t^{1/\beta} + d_{\mathcal{D}}(x,y)}\right)^\beta,
\end{aligned}$$

as desired.

To estimate Z_2, it follows from (4.22) that if $\mathbb{E}_{k_t}(x,y) \neq 0$ then $d_{\mathcal{D}}(x,y) = \delta^{k_{x,y}} \le \delta^{k_t} \le t^{1/\beta}$. This, combined with (VD) and (4.23), further derives

$$|Z_2| \lesssim e^{-t\delta^{-\beta k_t}} \frac{\mathbf{1}_{\{d_{\mathcal{D}}(x,y) \le t^{1/\beta}\}}}{V(x,\delta^{k_t})} \approx \frac{\mathbf{1}_{\{d_{\mathcal{D}}(x,y) \le t^{1/\beta}\}}}{V(x,t^{1/\beta})} \lesssim \frac{1}{V(x,\ t^{1/\beta} + d_{\mathcal{D}}(x,y))} \left(\frac{t^{1/\beta}}{t^{1/\beta} + d_{\mathcal{D}}(x,y)}\right)^\beta.$$

Finally, we deal with Z_3. For $k \le k_t - 1$, we have $\delta^{k\beta} \ge \delta^{(k_t-1)\beta} > t$, which implies that

$$\left|e^{-t\delta^{-\beta(k-1)}} - e^{-t\delta^{-\beta k}}\right| = e^{-t\delta^{-\beta k}} \left|e^{t\delta^{-\beta k}(1-\delta^\beta)} - 1\right| \approx e^{-t\delta^{-\beta k}} t\delta^{-\beta k}(1-\delta^\beta) \approx \frac{t}{\delta^{k\beta}}.$$

This, combined with (4.23), yields

$$|Z_3| \le \sum_{k \le k_t - 1} \left|e^{-t\delta^{-\beta(k-1)}} - e^{-t\delta^{-\beta k}}\right| \mathbb{E}_k(x,y) \lesssim \sum_{k \le k_t - 1} \frac{t}{\delta^{k\beta}} \frac{\mathbf{1}_{\{d(x,y) \le 2C_0 d_{\mathcal{D}}(x,y) \le 2C_0\delta^k\}}}{V(x,\delta^k)}.$$

Since now we are summing over those k's satisfying both $\delta^k > t^{1/\beta}$ and $\delta^k \ge d_{\mathcal{D}}(x,y)$, we then derive from (VD) that

$$V(x,\ t^{1/\beta} + d_{\mathcal{D}}(x,y)) \lesssim V(x,\delta^k)$$

and, hence,

$$|Z_3| \lesssim \frac{1}{V(x,\ t^{1/\beta} + d_{\mathcal{D}}(x,y))} \sum_{\{k \in \mathbb{Z}:\ \delta^k \ge \max\{t^{1/\beta},\, d_{\mathcal{D}}(x,y)\}\}} \frac{t}{\delta^{k\beta}}$$

$$\lesssim \frac{1}{V(x,\, t^{1/\beta} + d_{\mathcal{D}}(x,y))} \frac{t}{\max\{t^{1/\beta},\, d_{\mathcal{D}}(x,y)\}}$$
$$\approx \frac{1}{V(x,\, t^{1/\beta} + d_{\mathcal{D}}(x,y))} \left(\frac{t^{1/\beta}}{t^{1/\beta} + d_{\mathcal{D}}(x,y)}\right)^{\beta}.$$

Inserting the estimates of Z_1, Z_2 and Z_3 into (4.25), we obtain the desired upper bound for $p_t(x,y)$.

Part 2: verifying the upper estimate in (4.20) under the case $\mu(\mathcal{X}) < \infty$. Let $k_t \in \mathbb{Z}$ be as in (4.24). In the case $k_t \geq m_{\mathcal{X}} + 1$, instead of (4.25), we have

$$\begin{aligned} p_t(x,y) &= \sum_{k \geq k_t} e^{-t\delta^{-\beta k}} \mathbb{D}_k(x,y) + e^{-t\delta^{-\beta(k_t-1)}} \mathbb{E}_{k_t}(x,y) + \sum_{k=m_{\mathcal{X}}}^{k_t-1} \left(e^{-t\delta^{-\beta(k-1)}} - e^{-t\delta^{-\beta k}}\right) \mathbb{D}_k(x,y) \\ &=: \mathrm{Z}_1 + \mathrm{Z}_2 + \mathrm{Z}_3'. \end{aligned}$$

Note that the estimate of $|Z_3|$ remains true for $|Z_3'|$. By this and the estimates of Z_1 and Z_2, we again deduce that p_t satisfies the upper estimate in (4.20) under $\mu(\mathcal{X}) < \infty$ and $k_t \geq m_{\mathcal{X}} + 1$.

If $k_t \leq m_{\mathcal{X}}$, then the formula (4.11) directly gives

$$p_t(x,y) = \frac{1}{\mu(\mathcal{X})} + \sum_{k \geq m_{\mathcal{X}}} e^{-t\delta^{-\beta k}} \mathbb{D}_k(x,y) =: \frac{1}{\mu(\mathcal{X})} + \mathrm{Z}_1'. \tag{4.26}$$

Observing that if $k \geq m_{\mathcal{X}}$ then $k \geq k_t$ and $\delta^{k\beta} \leq t$, we obtain that the estimate of Z_1 also gives

$$|\mathrm{Z}_1'| \lesssim \frac{1}{V(x,\, t^{1/\beta} + d(x,y))} \left(\frac{t^{1/\beta}}{t^{1/\beta} + d(x,y)}\right)^{\beta}.$$

By (4.7) and the fact $t^{1/\beta} \geq \delta^{k_t} \geq \delta^{m_{\mathcal{X}}}$, we have

$$\frac{1}{\mu(\mathcal{X})} \lesssim \frac{1}{V(x,\, \delta^{m_{\mathcal{X}}} + d(x,y))} \left(1 + \frac{d(x,y)}{\delta^{m_{\mathcal{X}}}}\right)^{-\beta} \lesssim \frac{1}{V(x,\, t^{1/\beta} + d(x,y))} \left(\frac{t^{1/\beta}}{t^{1/\beta} + d(x,y)}\right)^{\beta}.$$

Inserting the last two estimates into (4.26), we deduce that p_t satisfies the upper estimate in (4.20) under the situation $\mu(\mathcal{X}) < \infty$ and $k_t \leq m_{\mathcal{X}}$.

Part 3: verifying the lower estimate in (4.20). Suppose that x and y are both contained in the same quadrant of $\mathcal{X}$. Let $Q_{x,y} \in \mathcal{D}$ be the smallest dyadic cube containing both x and y. In other words, any child of $Q_{x,y}$ does not contain x, y simultaneously. Let $k_{x,y} \in \mathbb{Z}$ be the unique integer such that $Q_{x,y} \in \mathcal{D}_{k_{x,y}}$ and, hence,

$$d_{\mathcal{D}}(x,y) = \delta^{k_{x,y}}.$$

For simplicity, set $k_0 = k_{x,y}$. Assume without loss of generality that $Q_{x,y} = Q^{k_0}_{\alpha_{k_0}}$ for some $\alpha_{k_0} \in \mathcal{A}_{k_0}$. For any $k \leq k_0$, applying Theorem 4.1 implies that there exists a unique dyadic cube $Q^k_{\alpha_k} \in \mathcal{D}_k$ such that $x,\, y \in Q^k_{\alpha_k}$ and

$$Q^{k_0}_{\alpha_{k_0}} \subseteq \cdots \subseteq Q^{k+1}_{\alpha_{k+1}} \subseteq Q^k_{\alpha_k} \subseteq \cdots.$$

This in turn gives

$$\mathbb{E}_k(x,y) = \begin{cases} 0 & \text{as } k > k_0; \\ \frac{1}{\mu(Q^k_{\alpha_k})} & \text{as } k \leq k_0. \end{cases}$$

No matter $\mu(\mathcal{X})$ is finite or nor, by (4.13) in Lemma 4.8, we obtain

$$
\begin{aligned}
p_t(x,y) &= \sum_{k\le k_0}\left(e^{-t\delta^{-(k-1)\beta}}-e^{-t\delta^{-k\beta}}\right)\mathbb{E}_k(x,y)\\
&= \sum_{k\le k_0} e^{-t\delta^{-\beta k}}\left(e^{t\delta^{-\beta k}(1-\delta^{\beta})}-1\right)\frac{1}{\mu(Q^k_{\alpha_k})}\\
&\ge \sum_{\{k\in\mathbb{Z}:\,\delta^k\ge\max\{\delta^{k_0},\,t^{1/\beta}\}\}} e^{-t\delta^{-\beta k}}\left(e^{t\delta^{-\beta k}(1-\delta^{\beta})}-1\right)\frac{1}{\mu(Q^k_{\alpha_k})}\\
&\approx \sum_{\{k\in\mathbb{Z}:\,\delta^k\ge\max\{\delta^{k_0},\,t^{1/\beta}\}\}} t\delta^{-\beta k}\frac{1}{\mu(Q^k_{\alpha_k})},
\end{aligned}
\tag{4.27}
$$

where the last step is due to

$$
e^{-t\delta^{-\beta k}}\left(e^{t\delta^{-\beta k}(1-\delta^{\beta})}-1\right)\approx t\delta^{-\beta k}\quad\text{as }\ \delta^k\ge t^{1/\beta}.
$$

Under $\delta^k\ge\max\{\delta^{k_0},\ t^{1/\beta}\}$, from $\{x,y\}\subseteq Q_{x,y}=Q^{k_0}_{\alpha_{k_0}}\subseteq Q^k_{\alpha_k}\in\mathcal{D}_k$ and Theorem 4.1, it follows that

$$
d(x,y)\le \operatorname{diam}(Q_{x,y})\le\operatorname{diam}(Q^k_{\alpha})\le 2C_0\delta^k,
$$

which, along with (VD), yields

$$
\mu(Q^k_{\alpha_k})\le\mu(B(x,2C_0\delta^k))=V(x,2C_0\delta^k)\lesssim\left(\frac{\delta^k}{t^{1/\beta}}\right)^{n_+}V(x,t^{1/\beta}).
$$

With these facts, we continue the estimates of (4.27) by considering the cases $d_{\mathcal{D}}(x,y)=\delta^{k_0}\ge t^{1/\beta}$ and $d_{\mathcal{D}}(x,y)=\delta^{k_0}<t^{1/\beta}$, respectively.

- If $d_{\mathcal{D}}(x,y)=\delta^{k_0}\ge t^{1/\beta}$, then we consider only the term $k=k_0$ in (4.27) and obtain

$$
p_t(x,y)\gtrsim t\delta^{-\beta k_0}\frac{1}{\mu(Q^{k_0}_{\alpha_{k_0}})}\gtrsim t\delta^{-\beta k_0}\frac{1}{V(x,2C_0\delta^{k_0})}\approx\frac{1}{V(x,\ t^{1/\beta}+d_{\mathcal{D}}(x,y))}\left(\frac{t^{1/\beta}}{t^{1/\beta}+d_{\mathcal{D}}(x,y)}\right)^{\beta}
$$

- If $d_{\mathcal{D}}(x,y)=\delta^{k_0}<t^{1/\beta}$, then it follows form (4.27) and (VD) that

$$
\begin{aligned}
p_t(x,y)&\gtrsim\frac{1}{V(x,t^{1/\beta}))}\sum_{\{k\in\mathbb{Z}:\,\delta^k\ge t^{1/\beta}\}}t\delta^{-\beta k}\left(\frac{t^{1/\beta}}{\delta^k}\right)^{n_+}\\
&\approx\frac{1}{V(x,t^{1/\beta}))}\\
&\approx\frac{1}{V(x,\ t^{1/\beta}+d_{\mathcal{D}}(x,y))}\\
&\approx\frac{1}{V(x,\ t^{1/\beta}+d_{\mathcal{D}}(x,y))}\left(\frac{t^{1/\beta}}{t^{1/\beta}+d_{\mathcal{D}}(x,y)}\right)^{\beta}.
\end{aligned}
$$

Combining the arguments in the above two cases we obtain the lower estimate of p_t in (4.20). □

4.4 Examples of heat kernels

4.4.1 A heat kernel on $\mathbb{R}^n$ based on Haar wavelets

Suppose that $\mathcal{X}$ is the n-dimensional Euclidean space $\mathbb{R}^n$, endowed with the classical Euclidean distance and the Lebesgue measure. On $\mathbb{R}^n$ there is a natural dyadic system

$$\mathcal{D}_{\mathbb{R}^n} := \left\{2^{-k}\left([0,1)^n + m\right) : \ k \in \mathbb{Z}, \ m \in \mathbb{Z}^n\right\}.$$

Note that each dyadic cube $Q \in \mathcal{D}_{\mathbb{R}^n}$ has exactly 2^n children of equal sizes. Let

$$h_F := \mathbf{1}_{[0,1)} \quad \text{and} \quad h_M := \mathbf{1}_{[0,1/2)} - \mathbf{1}_{[1/2,1)}.$$

Set

$$G^{n*} := \{F, M\}^n \setminus \{F\}^n.$$

For any $k \in \mathbb{Z}$, $m = (m_1, \cdots, m_n) \in \mathbb{Z}^n$, $G = (G_1, \cdots, G_n) \in G^{n*}$ and $x \in \mathbb{R}^n$, define

$$H_m^G(x) := \prod_{r=1}^{n} h_{G_r}(x_r - m_r)$$

and

$$H_m^{k,G}(x) := 2^{kn/2} H_m^G(2^k x).$$

Then $\{H_m^{k,G} : \ k \in \mathbb{Z}, m \in \mathbb{Z}^n, G \in G^{n*}\}$ form the family of Haar functions on $\mathbb{R}^n$ (see, e.g. [55, Proposition 1.53] for the inhomogeneous case). Based on the previous discussions, there is a heat kernel $\{p_t^{\mathcal{D}_{\mathbb{R}^n}}\}_{t \geq 0}$ such that for all $t \in (0, \infty)$ and almost all $x, y \in \mathcal{X}$,

$$p_t^{\mathcal{D}_{\mathbb{R}^n}}(x, y) = \sum_{k \in \mathbb{Z}} \sum_{m \in \mathbb{Z}^n} \sum_{G \in G^{n*}} e^{-t\delta^{-\beta k}} H_m^{k,G}(x) H_m^{k,G}(y)$$

and

$$p_t^{\mathcal{D}_{\mathbb{R}^n}}(x, y) \approx \frac{t}{(t^{1/\beta} + d_{\mathcal{D}_{\mathbb{R}^n}}(x, y))^{n+\beta}},$$

where $d_{\mathcal{D}_{\mathbb{R}^n}}$ is the ultra-metric induced by the dyadic system $\mathcal{D}_{\mathbb{R}^n}$. In this case, the generator $\mathcal{L}_{\mathcal{D}_{\mathbb{R}^n}}$ is give by

$$\mathcal{L}_{\mathcal{D}_{\mathbb{R}^n}}(f) := \sum_{k \in \mathbb{Z}} \sum_{m \in \mathbb{Z}^n} \sum_{G \in G^{n*}} 2^{k\beta} \langle f, \ H_m^{k,G} \rangle H_m^{k,G},$$

whose domain is

$$\operatorname{Dom}\left(\mathcal{L}_{\mathcal{D}_{\mathbb{R}^n}}\right) := \left\{f \in L^2(\mathbb{R}^n) : \ \mathcal{L}^{\mathcal{D}_{\mathbb{R}^n}}(f) \in L^2(\mathbb{R}^n)\right\}.$$

4.4.2 A heat kernel on $\mathbb{Z}$

Let $\mathcal{X} = \mathbb{Z}$ be the one-dimensional integer lattice endowing with a graph distance d and counting measure μ, namely, for any $m, l \in \mathbb{Z}$ and any set $A \subset \mathbb{Z}$,

$$d(m,l) := |m-l| \quad \text{and} \quad \mu(A) := \#(A).$$

Recall that on the Euclidean space $\mathbb{R}$ there is a standard system of dyadic intervals

$$\mathcal{D}_{\mathbb{R}} := \left\{[2^{-k}m,\ 2^{-k}(m+1)) :\ k \in \mathbb{Z},\ m \in \mathbb{Z}\right\}. \tag{4.28}$$

It is natural to define the dyadic cubes on $\mathbb{Z}$ via

$$\mathcal{D} := \mathbb{Z} \cap \mathcal{D}_{\mathbb{R}} = \left\{I_{k,m} := \mathbb{Z} \cap [2^{-k}m,\ 2^{-k}(m+1)) :\ k \in \mathbb{Z},\ m \in \mathbb{Z}\right\}. \tag{4.29}$$

For any $k \in \mathbb{Z}$, the family of dyadic cubes of the k-th generation is defined by

$$\mathcal{D}_k := \{I_{k,m} :\ m \in \mathbb{Z}\}.$$

If $k = 0$, then $I_{k,m}$ is a isolated point m. If $k > 0$, then every dyadic cube $I_{k,m}$ is either empty or contains at most one point $2^{-k}m$. If $k < 0$, then for any $m \in \mathbb{Z}$ we have

$$I_{k,m} = \left\{2^{-k}m + l :\ l = 0, 1, \ldots, 2^{-k} - 1\right\}.$$

For any $k \in \mathbb{Z}$ and any function $f : \mathbb{Z} \to \mathbb{R}$, we follow (4.9) and set

$$\mathbb{E}_k(f) := \sum_{I \in \mathcal{D}_k} \left(\frac{1}{\mu(I)} \int_I f\, d\mu\right) \mathbf{1}_I$$

and

$$\mathbb{D}_k(f) := \mathbb{E}_{k+1} f - \mathbb{E}_k(f).$$

For any $k \geq 0$ and $i \in \mathbb{Z}$, we find that $I_{k,m} \ni i$ if and only if $i = 2^{-k}m$, which implies that

$$\mathbb{E}_k(f)(i) = f(i)$$

and, hence,

$$\mathbb{D}_k(f) = \mathbb{E}_{k+1}(f) - \mathbb{E}_k(f) = 0 \quad \text{as } k \geq 0.$$

Note that if $I \in \mathcal{D}_k$ then $\mu(I) = 2^{-k}$. So, the integral kernels of $\mathbb{E}_k$ and $\mathbb{D}_k$ are represented as follows: for all $i, j \in \mathbb{Z}$,

$$\mathbb{E}_k(i,j) = 2^k \sum_{I \in \mathcal{D}_k} \mathbf{1}_I(i) \mathbf{1}_I(j)$$

and

$$\mathbb{D}_k(i,j) = \mathbb{E}_{k+1}(i,j) - \mathbb{E}_k(i,j).$$

Then, applying (4.19) yields that the heat kernel associated to $\mathcal{D}$ satisfies that, for any $t > 0$ and $i, j \in \mathbb{Z}$,

$$p_t(i,j) = \sum_{k=-\infty}^{-1} e^{-t 2^{\beta k}} \mathbb{D}_k(i,j)$$

$$
\begin{aligned}
&= \lim_{N\to+\infty} \sum_{k=-N}^{-1} e^{-t2^{\beta k}} \mathbb{D}_k(i,j) \\
&= \lim_{N\to+\infty} \sum_{k=-N}^{-1} e^{-t2^{\beta k}} \left(\mathbb{E}_{k+1}(i,j) - \mathbb{E}_k(i,j)\right) \\
&= \lim_{N\to+\infty} \left(\sum_{k=-N+1}^{0} e^{-t2^{\beta(k-1)}} \mathbb{E}_k(i,j) - \sum_{k=-N}^{-1} e^{-t2^{\beta k}} \mathbb{E}_k(i,j) \right) \\
&= e^{-t2^{-\beta}} \mathbb{E}_0(i,j) + \sum_{k=-\infty}^{-1} \left(e^{-2^{(k-1)\beta}t} - e^{-2^{k\beta}t} \right) \mathbb{E}_k(i,j) - \lim_{N\to+\infty} e^{-t2^{-N\beta}} \mathbb{E}_{-N}(i,j).
\end{aligned}
$$

By the expression of $\mathbb{E}_k(i,j)$, we easily see that

$$
\lim_{N\to+\infty} e^{-t2^{-N\beta}} \mathbb{E}_{-N}(i,j) = 0 \quad \text{and} \quad \mathbb{E}_0(i,j) = \delta_{i,j},
$$

where

$$
\delta_{i,j} := \begin{cases} 1 & \text{as } i = j; \\ 0 & \text{as } i \neq j. \end{cases}
$$

Therefore,

$$
p_t(i,j) = e^{-t2^{-\beta}} \delta_{i,j} + \sum_{k=-\infty}^{-1} 2^k \left(e^{-2^{(k-1)\beta}t} - e^{-2^{k\beta}t} \right) \left(\sum_{I\in\mathcal{D}_k} \mathbf{1}_I(i)\mathbf{1}_I(j) \right). \tag{4.30}
$$

From (4.30), it can be seen that the associated Markov chain jumps only between points $i, j \in \mathbb{Z}$ lying in the same dyadic cubes of $\mathcal{D}$. Moreover, this jump has larger transition probability if there are more generations of dyadic cubes that contain both i, j. For instance, i is more likely to jump to j if they are close to each other. However, $i = 1$ is impossible to jump to $j = -1$, since they can not lie in the same dyadic cube of $\mathcal{D}$.

Now, we can construct an adjacent family of dyadic interval systems on $\mathbb{R}$. For any $\tau \in \{0,1,2\}$, let

$$
\mathcal{D}_{\mathbb{R}}^{\tau} := \bigcup_{k\in\mathbb{Z}} \left\{ 2^{-k} \left([0,1) + m + (-1)^k \tau/3 \right) : \ m \in \mathbb{Z} \right\}.
$$

Note that for $\tau = 0$, $\mathcal{D}_{\mathbb{R}}^0$ reduces to the dyadic cube system $\mathcal{D}_{\mathbb{R}}$ as in (4.28). In view of this and (4.29), we may define the adjacent family $\{\mathcal{D}^\tau\}_{\tau=0}^2$ of dyadic cube systems on $\mathbb{Z}$ by setting for any $\tau \in \{0,1,2\}$,

$$
\mathcal{D}^\tau := \mathbb{Z} \cap \mathcal{D}_{\mathbb{R}}^\tau.
$$

Thus, we can define the heat kernel $p_t^{\mathcal{D}^\tau}(i,j)$ in the way similar to (4.30). For simplicity, when $t = 1$, we write

$$
p^{\mathcal{D}^\tau}(i,j) := p_1^{\mathcal{D}^\tau}(i,j), \tag{4.31}
$$

which represents the transition probability from i to j in a unit time. For any two points $i, j \in \mathbb{Z}$, observe that there exists $\tau \in \{0,1,2\}$ such that both i and j lie in the same dyadic cube of $\mathcal{D}^\tau$.

The existence of adjacent family of dyadic cube systems enable us to define a new kind of random walk on $\mathbb{Z}$, which may be described in the following way that in each step the jumper has a chance to choose a dyadic structure randomly from the adjacent family and then jump by the law of the chosen dyadic structure. To be precise, let $\Omega := \{0, 1, 2\}$ be a sample space equipped with the natural probability $\mathbb{P}$ that for any $\tau \in \Omega$,

$$\mathbb{P}(\omega = \tau) = \frac{1}{3}. \tag{4.32}$$

Now, for any $i, j \in \mathbb{Z}$, define

$$p(i, j) := \sum_{\tau=0}^{2} \mathbb{P}(\omega = \tau)\, p^{\mathcal{D}^\tau}(i, j) = \frac{1}{3} \sum_{\tau=0}^{2} p^{\mathcal{D}^\tau}(i, j), \tag{4.33}$$

where $p^{\mathcal{D}^\tau}(i, j)$ is as in (4.31). By (4.32) and the stochastically completeness of each $p^{\mathcal{D}^\tau}$, we know that the function $p(\cdot, \cdot)$ in (4.33) is a Markov kernel, namely, for any $i \in \mathbb{Z}$,

$$\sum_{j\in\mathbb{Z}} p(i, j) = 1.$$

Recall that each Markov kernel defines a random walk $\{\mathbb{X}_n\}_{n\in\mathbb{Z}_+}$ on $\mathbb{Z}$ (see, e.g. [33]) such that for any $n \in \mathbb{N}$ and $i, j \in \mathbb{Z}$, the n-step transition probability from i to j equals to $p_t(i, j)$ in (4.30) with $t = n$. This is essentially what we have done in [13] by summing up all adjacent heat kernels and Dirichlet forms.

4.4.3 A heat kernel via p-adic MRA

Let $p \in \mathbb{N}$ be a prime number and $\mathbb{Q}_p$ the *p-adic field* defined as the set of all numbers x that can be represented as the series such that

$$x = \sum_{i=\gamma}^{\infty} x_i p^i, \tag{4.34}$$

where $\gamma \in \mathbb{Z}$ and $x_i \in \{0, \dots, p-1\}$ (see, e.g. [56]). By (4.34), it is easy to see each $x \in \mathbb{Q}_p$ has a decomposition

$$x = [x] + \{x\},$$

where $[x] := \sum_{i=0}^{\infty} x_i p^i$ and $\{x\} := \sum_{i=\gamma}^{-1} x_i p^i$ are respectively the integer and fractional parts of x. For any x as in (4.34), its *p-adic norm* is defined by

$$|x|_p := p^{-\gamma}.$$

Thus, for any $x \in \mathbb{Q}_p$ and $y \in \mathbb{Q}_p$, they have the distance

$$d_p(x, y) := |x - y|_p.$$

Such d_p is an ultra metric. Moreover, there exists a Haar measure μ on $\mathbb{Q}_p$ so that $(\mathbb{Q}_p, d_p, \mu)$ is an ultra-metric measure space (see, e.g. [10, Section 2.5] for more details).

Recall in [45] that a function $\phi \in L^2(\mathbb{Q}_p)$ is called a *scaling function* if it satisfies the following refinement equation

$$\phi(\cdot) = \sum_{n\in I_p} \alpha_n \phi(p^{-1} \cdot - n) \tag{4.35}$$

for some $\alpha_n \in \mathbb{C}$ and $I_p := \{x \in \mathbb{Q}_p :\ \{x\} = x\}$. In particular, if we let $B_p(0,1) := \{x \in \mathbb{Q}_p :\ |x|_p \le 1\}$ be the unit closed ball in $\mathbb{Q}_p$, then a possible solution of (4.35) is the following function

$$\phi := \mathbf{1}_{B_p(0,1)}.$$

For any $k \in \mathbb{Z}$, let

$$V_k := \overline{\operatorname{span}\{\phi(p^{-k}\cdot - n) :\ n \in I_p\}}. \tag{4.36}$$

It is proved in [52] that such $\{V_k\}_{k\in\mathbb{Z}}$ forms an MRA as in Definition 2.1. The sequence $\{V_k\}_{k\in\mathbb{Z}}$ defined in (4.36) is called *p-adic Haar* MRA.

Based on [45], the corresponding Haar wavelets are of the form that for any $j \in \{1,\dots,p-1\}$, $k \in \mathbb{Z}$ and $n \in I_p$,

$$\psi_{j;k,n}(x) := p^{-k/2}\exp\left(2\pi i\left\{p^{-1}jx\right\}\right)\phi(p^k x - n),$$

which forms an orthogonal basis of $L^2(\mathbb{Q}_p)$. For any $k \in \mathbb{Z}$, if $\mathbb{Q}_k$ denotes the projector from $L^2(\mathbb{Q}_p)$ to $W_k := V_{k+1} \ominus V_k$, then its integral kernel takes the form of

$$\mathbb{Q}_k(x,y) = \sum_{j\in\{1,\dots,p-1\}}\sum_{n\in I_p}\psi_{j;k,n}(x)\psi_{j;k,n}(y).$$

Let $\beta \in (0,\infty)$. By using (4.11), one can construct a stochastic complete nonnegative heat kernel $\{p_t\}_{t>0}$ such that

$$p_t(x,y) = \sum_{k\in\mathbb{Z}}\sum_{j\in\{1,\dots,p-1\}}\sum_{n\in I_p} e^{-tp^{k\beta}}\psi_{j;k,n}(x)\psi_{j;k,n}(y) \quad \text{for all } t\in(0,\infty) \text{ and } x,y\in\mathcal{X}.$$

In this case, the generator $\mathcal{L}$ of $\{p_t\}_{t>0}$ coincides to the following Vladimirov operator (or, equivalently, Taibleson operator) of p-adic fractional differential D^β defined by for any $\beta > 0$,

$$D^\beta f(x) := \frac{1-p^{-1-\beta}}{p^\beta - 1}\int_{\mathbb{Q}_p}\frac{f(x)-f(y)}{|x-y|_p^{1+\beta}}\,d\mu(y),$$

which is viewed as the corresponding Laplacian in $\mathbb{Q}_p$ (see, e.g. [10, 45]).

5 Construction of heat kernels via smooth MRA

The whole section is devoted to the proof of Theorem 2.9 under (VD). To this end, we will apply Theorem 2.6 to a specific smooth MRA generated by smooth splines from [2].

5.1 Random dyadic cubes

Let us follow [39] and randomize the construction of dyadic cubes in Theorem 4.1 of Section 4.1. We still fix $\delta \in (0,1)$, which is a sufficiently small number (for example, $\delta \le 1/60$). Then there exists a set of reference dyadic points

$$\left\{z_\alpha^k :\ k\in\mathbb{Z},\ \alpha\in\mathcal{A}_k\right\}$$

satisfying (4.1) and (4.2). Let $\mathcal{G}$ be as in (4.3) and $\leq$ the partial order of $\mathcal{G}$ as in (4.4). Denote by

$$\Omega := \{0, 1, 2, \ldots, \lfloor 1/\delta \rfloor\}^{\mathbb{Z}}$$

the *sample space*. Elements of Ω are called the *parameterized points*. A parameterized point $\omega \in \Omega$ is denoted by $\omega := (\omega_k)_{k\in\mathbb{Z}}$, where

$$\omega_k \in \{0, 1, 2, \ldots, \lfloor 1/\delta \rfloor\}.$$

Equip Ω with the natural probability measure $\mathbb{P}_\omega$, which makes all coordinates ω_k independent of each other and uniformly distributed over the finite set $\{0, 1, \ldots, \lfloor 1/\delta \rfloor\}$. In other words, if $a_k : \Omega \to \{0, 1, 2, \ldots, \lfloor 1/\delta \rfloor\}$ is a random variable, with

$$a_k(\{\omega_i\}_{i\in\mathbb{Z}}) = \omega_k,$$

then for every $k \in \mathbb{Z}$ and $T \in \{0, 1, \ldots, \lfloor 1/\delta \rfloor\}$,

$$\mathbb{P}_\omega(a_k = T) = \frac{1}{1 + \lfloor 1/\delta \rfloor}.$$

For any $k \in \mathbb{Z}$, we denote

$$r_k = r_{a_k} := \frac{1}{4}\left(\delta^k + a_k\delta^{k+1}\right).$$

For any $\omega := (\omega_k)_{k\in\mathbb{Z}} \in \Omega$, instead of (4.4), now there is a *new partial order* $\leq_\omega$ which is defined as follows:

$$d(z_\gamma^{k+1}(\omega), z_\alpha^k(\omega)) < r_k \quad \Rightarrow \quad (k+1, \gamma) \leq_\omega (k, \alpha) \quad \Rightarrow \quad d(z_\gamma^{k+1}(\omega), z_\alpha^k(\omega)) < 4r_k.$$

Define the *preliminary*, the *closed*, and the *open random dyadic cubes* as below:

$$\widehat{Q}_\alpha^k(\omega) := \left\{z_\gamma^\ell : (\ell, \gamma) \leq_\omega (k, \alpha)\right\}, \quad \overline{Q}_\alpha^k(\omega) := \overline{\widehat{Q}_\alpha^k(\omega)}, \quad \widetilde{Q}_\alpha^k(\omega) := \text{interior of } \overline{Q}_\alpha^k(\omega).$$

Next, we collect a series of facts on the random dyadic cubes from [39, Theorem 5.2].

Theorem 5.1 ([39]). *Let $(\mathcal{X}, d, \mu)$ be a metric measure space satisfying (VD) and $\delta \in (0, 1/(84M^8))$, where M is as in (1.2). For any fixed parameter*

$$\omega \in \Omega := \{0, 1, 2, \ldots, \lfloor 1/\delta \rfloor\}^{\mathbb{Z}},$$

there exist Borel sets (called random dyadic cubes)

$$\mathscr{D}(\omega) := \left\{Q_\alpha^k(\omega) : k \in \mathbb{Z}, \alpha \in \mathcal{A}_k\right\}$$

satisfying the following properties:

(i) *for any $k \in \mathbb{Z}$ and $\alpha \in \mathcal{A}_k$, the interior of $Q_\alpha^k(\omega)$ is $\widetilde{Q}_\alpha^k(\omega)$, and the closure of $Q_\alpha^k(\omega)$ is $\overline{Q}_\alpha^k(\omega)$;*

(ii) *for any $k \in \mathbb{Z}$ and $\alpha \in \mathcal{A}_k$, $\widetilde{Q}_\alpha^k(\omega)$ and $\overline{Q}_\alpha^k(\omega)$ are one another's interior and closure;*

(iii) *for any $k \in \mathbb{Z}$, $\{Q_\alpha^k(\omega)\}_{\alpha\in\mathcal{A}_k}$ are disjoint and $\mathcal{X} = \bigcup_{\alpha\in\mathcal{A}_k} Q_\alpha^k(\omega)$;*

(iv) *if $j \geq k$, $\alpha \in \mathcal{A}_k$ and $\gamma \in \mathcal{A}_j$, then either $Q_\gamma^j(\omega) \subset Q_\alpha^k(\omega)$ or $Q_\gamma^j(\omega) \cap Q_\alpha^k(\omega) = \emptyset$;*

(v) *for any* $k \in \mathbb{Z}$, $\alpha \in \mathcal{A}_k$ *and* $j < k$, *there exists a unique* $\gamma \in \mathcal{A}_j$ *such that* $Q^k_\alpha(\omega) \subset Q^j_\gamma(\omega)$*;*

(vi) *for any* $k \in \mathbb{Z}$ *and* $\alpha \in \mathcal{A}_k$, $B(z^k_\alpha(\omega), 5^{-1}\delta^k) \subset Q^k_\alpha(\omega) \subset B(z^k_\alpha(\omega), 3\delta^k)$*;*

(vii) *for any* $x \in \mathcal{X}$, $k \in \mathbb{Z}$ *and* $\epsilon \in (0, \infty)$,

$$\mathbb{P}_\omega\left(x \in \bigcup_{\alpha\in\mathcal{A}_k} \partial_\epsilon Q^k_\alpha(\omega)\right) \le \frac{1}{\delta}\epsilon^{\eta_\delta},$$

where $\eta_\delta := 1 - \frac{\log(84M^8)}{\log(1/\delta)}$ *and* $\partial_\epsilon Q^k_\alpha(\omega) := \{y \in \overline{Q}^k_\alpha(\omega) : d(y, \widetilde{Q}^k_\alpha(\omega)^\complement) < \epsilon\delta^k\}$.

5.2 An admissible smooth MRA

Let us adopt all the notation in the previous subsection. For any $k \in \mathbb{Z}$ and $\alpha \in \mathcal{A}_k$, define the *spline function*

$$s^k_\alpha(x) := \mathbb{P}_\omega\left(x \in \overline{Q}^k_\alpha(\omega)\right) \qquad \text{for all } x \in \mathcal{X}.$$

Each spline function s^k_α locates near the dyadic cube $Q^k_\alpha(\omega)$. According to [2, Theorem 3.1] and [39, Corollary 6.13], we have

$$\mathbf{1}_{B(z^k_\alpha,\, 8^{-1}\delta^k)} \le s^k_\alpha \le \mathbf{1}_{B(z^k_\alpha,\, 8^5\delta^k)} \quad \text{and} \quad \sum_{\alpha\in\mathcal{A}_k} s^k_\alpha \equiv 1,$$

and that s^k_α enjoys the following Hölder continuity:

$$|s^k_\alpha(x) - s^k_\alpha(y)| \le C_\delta\left(\frac{d(x,y)}{\delta}\right)^{\eta_\delta} \qquad \text{for all } x,\ y \in \mathcal{X},$$

where $\eta_\delta \in (0,1)$ is the same constant as in Theorem 5.1(vii). Under $\mu(\mathcal{X}) < \infty$, we let $m_{\mathcal{X}}$ be as in (4.7), which implies that $s^k_\alpha \equiv 1$ whenever $k \le m_{\mathcal{X}}$ and $\alpha \in \mathcal{A}_k$. According to [2, Theorem 5.1], the spline functions produce an MRA in the sense of Definition 2.1.

Lemma 5.2 ([2])**.** *Let* $(\mathcal{X}, d, \mu)$ *be a metric measure space satisfying* (VD)*. For any* $k \in \mathbb{Z}$, *let* V_k *be the closed linear span of* $\{s^k_\alpha\}_{\alpha\in\mathcal{A}_k}$ *in* $L^2(\mathcal{X})$*. Then the following hold:*

(i) *for each* $k \in \mathbb{Z}$, $V_k \subset V_{k+1}$*;*

(ii) $\overline{\bigcup_{k\in\mathbb{Z}} V_k} = L^2(\mathcal{X})$*;*

(iii) *when* $\mu(\mathcal{X}) = \infty$, $\bigcap_{k\in\mathbb{Z}} V_k = \{0\}$*;*

(iii)′ *when* $\mu(\mathcal{X}) < \infty$, $\bigcap_{k\in\mathbb{Z}} V_k = V_{m_{\mathcal{X}}} = \{$*constant functions*$\}$, *where* $m_{\mathcal{X}}$ *is as in* (4.7)*;* [2]

(iv) *for each* $k \in \mathbb{Z}$, *with the notation* $\mu^k_\alpha := \mu(B(z^k_\alpha, \delta^k))$, *the functions* $\{s^k_\alpha/\sqrt{\mu^k_\alpha}\}_{\alpha\in\mathcal{A}_k}$ *form a Riesz basis of* V_k, *namely, for any scalar sequence* $\{\lambda_\alpha\}_{\alpha\in\mathcal{A}_k}$,

$$\left\|\sum_{\alpha\in\mathcal{A}_k} \lambda_\alpha s^k_\alpha\right\|_{L^2(\mathcal{X})} \simeq \left(\sum_{\alpha\in\mathcal{A}_k} |\lambda_\alpha|^2\mu^k_\alpha\right)^{1/2},$$

where the implicit positive constants are independent of k *and* $\{\lambda_\alpha\}_{\alpha\in\mathcal{A}_k}$.

[2] Note that in Lemma 5.2 we have $V_k = V_{m_{\mathcal{X}}}$ whenever $k \le m_{\mathcal{X}}$. The value $m_{\mathcal{X}}$ plays the same role as the integer k_0 in Definition 2.1(iii).

Consequently, the sequence $\{V_k\}_{k\in\mathbb{Z}}$ *forms an* MRA *in* $L^2(\mathcal{X})$.

In this and the forthcoming subsections, we will always assume that $\{V_k\}_{k\in\mathbb{Z}}$ is the MRA constructed in Lemma 5.2. For any $k \in \mathbb{Z}$, we still denote by W_k the orthogonal complement of V_k in V_{k+1}. For any $k \in \mathbb{Z}$, denote by $\mathbb{P}_k$ and $\mathbb{Q}_k$ the orthogonal projectors from $L^2(\mathcal{X})$ onto V_k and W_k, respectively. Of course, for any $k \in \mathbb{Z}$, one still has (see (3.2))

$$\mathbb{Q}_k = \mathbb{P}_{k+1} - \mathbb{P}_k.$$

By using a very delicate construction of orthogonal basis[3] of V_k and W_k, Auscher and Hytönen (see [2, Lemma 10.1]) derives the following properties of the integral kernels of $\mathbb{P}_k$ and $\mathbb{Q}_k$.

Lemma 5.3 ([2]). *Let* $(\mathcal{X}, d, \mu)$ *be a metric measure space satisfying* (VD) *and* $\{V_k\}_{k\in\mathbb{Z}}$ *be the MRA constructed in Lemma 5.2. Then, for any* $k \in \mathbb{Z}$*, the orthogonal projectors*

$$\mathbb{P}_k : L^2(\mathcal{X}) \to V_k \quad \text{and} \quad \mathbb{Q}_k : L^2(\mathcal{X}) \to W_k$$

have integral kernels $\mathbb{P}_k(x, y)$ *and* $\mathbb{Q}_k(x, y)$*, which satisfy the following properties:*

(i) **(symmetric)** *for any* $k \in \mathbb{Z}$*, both* $\mathbb{P}_k(x, y)$ *and* $\mathbb{Q}_k(x, y)$ *are symmetric in* x *and* y*.*

(ii) **(exponential decay)** *for any* $k \in \mathbb{Z}$ *and* $x, y \in \mathcal{X}$*,*

$$|\mathbb{P}_k(x, y)| + |\mathbb{Q}_k(x, y)| \le \frac{C}{\sqrt{V(x, \delta^k)V(y, \delta^k)}} \exp\left(-c_* \frac{d(x, y)}{\delta^k}\right),$$

(iii) **(Hölder regularity)** *for any* $k \in \mathbb{Z}$ *and* $x, y, y' \in \mathcal{X}$ *such that* $d(y, y') \le \delta^k$*,*

$$|\mathbb{P}_k(x, y) - \mathbb{P}_k(x, y')| + |\mathbb{Q}_k(x, y) - \mathbb{Q}_k(x, y')| \le \frac{C}{\sqrt{V(x, \delta^k)V(y, \delta^k)}} \left(\frac{d(y, y')}{\delta^k}\right)^{\eta_\delta} \exp\left(-c_* \frac{d(x, y)}{\delta^k}\right);$$

(iv) **(stochastic completeness of** $\mathbb{P}_k$ **and vanishing property of** $\mathbb{Q}_k$**)** *for any* $k \in \mathbb{Z}$ *and* $x \in \mathcal{X}$*,*

$$\int_{\mathcal{X}} \mathbb{P}_k(x, y)\, d\mu(y) = 1 \qquad \text{and} \qquad \int_{\mathcal{X}} \mathbb{Q}_k(x, y)\, d\mu(y) = 0,$$

where δ*,* η_δ *are as in Theorem 5.1 and* C*,* c_* *are positive constants independent of* k, x, y *and* y'*.*

Remark 5.4. Let $\{V_k\}_{k\in\mathbb{Z}}$ be as in Lemma 5.2. We present a three-fold comment of the corresponding projectors $\{\mathbb{P}_k\}_{k\in\mathbb{Z}}$ and $\{\mathbb{Q}_k\}_{k\in\mathbb{Z}}$ in m Lemma 5.3.

(i) According to [2, Section 6], for any $k \in \mathbb{Z}$ and $x, y \in \mathcal{X}$, the integral kernel $\mathbb{P}_k(x, y)$ has the following representation

$$\mathbb{P}_k(x, y) = \sum_{\alpha\in\mathcal{A}_k} \sum_{\gamma\in\mathcal{A}_k} M_k^{-1}(\alpha, \gamma) \frac{s_\alpha^k(x) s_\gamma^k(y)}{\sqrt{V(z_\alpha^k, \delta^k)V(B(z_\gamma^k, \delta^k)}},$$

[3] The orthonormal basis of V_k and W_k constructed in [2] have exponential decay at infinity and Hölder regularity property, which are called *smooth wavelets*.

where $\{M_k^{-1}(\alpha,\gamma)\}_{\alpha,\gamma}$ is the inverse of the infinite matrix $\{M_k(\alpha,\gamma)\}_{\alpha,\gamma}$ with

$$M_k(\alpha,\gamma) := \frac{s_\alpha^k(x)s_\gamma^k(y)}{\sqrt{V(z_\alpha^k,\delta^k)V(z_\gamma^k,\delta^k)}}.$$

By the support conditions of the spline functions, we see that he matrix $\{M_k(\alpha,\gamma)\}_{\alpha,\gamma}$ is *banded*, that is, $M_k(\alpha,\gamma) = 0$ when $d(z_\alpha^k, z_\gamma^k) \geq c\delta^k$, where c is some positive constant independent of α,γ and k. For a banded matrix, its inverse matrix $\{M_k^{-1}(\alpha,\gamma)\}_{\alpha,\gamma}$ may have negative elements off the diagonal (see, e.g. [44]). Thus, $\mathbb{P}_k(x,y)$ may have many negative-valued points when x is far away from y.

(ii) Consider only the case $\mu(\mathcal{X}) = \infty$. For any $k \in \mathbb{Z}$, Auscher and Hytönen [2] constructed an orthonormal basis $\{\psi_\gamma^k : \gamma \in \mathcal{G}_k\}$ of W_k, where $\mathcal{G}_k = \mathcal{A}_{k+1} \setminus \mathcal{A}_k$, such that the following properties hold:

- *(exponential decay)* for any $k \in \mathbb{Z}$, $\gamma \in \mathcal{G}_k$ and $x \in \mathcal{X}$,

$$\left|\psi_\gamma^k(x)\right| \leq \frac{C}{\sqrt{V(z_\gamma^{k+1},\delta^{k+1})}} \exp\left(-c_* \frac{d(z_\gamma^{k+1},x)}{\delta^k}\right);$$

- *(Hölder regularity)* for any $k \in \mathbb{Z}$, $\gamma \in \mathcal{G}_k$ and $x,\ y \in \mathcal{X}$ satisfying $d(x,y) \leq \delta^k$,

$$\left|\psi_\gamma^k(x) - \psi_\gamma^k(y)\right| \leq C\left(\frac{d(x,y)}{\delta^k}\right)^{\eta_\delta} \frac{1}{\sqrt{V(z_\gamma^{k+1},\delta^{k+1})}} \exp\left(-c_* \frac{d(z_\gamma^{k+1},x)}{\delta^k}\right);$$

- *(vanishing mean)* for any $k \in \mathbb{Z}$ and $\gamma \in \mathcal{G}_k$,

$$\int_{\mathcal{X}} \psi_\gamma^k(x)\, d\mu(x) = 0,$$

where δ, η_δ are as in Theorem 5.1 and C, c_* are positive constants independent of k, γ, x and y. In this way, the family $\{\psi_\gamma^k :\ k \in \mathbb{Z},\ \gamma \in \mathcal{G}_k\}$ forms an orthonormal basis of $L^2(\mathcal{X})$, whose elements are called *smooth wavelets*. For any $f \in L^2(\mathcal{X})$, there is the wavelet decomposition formula (see [2]):

$$f = \sum_{k\in\mathbb{Z}} \mathbb{Q}_k f = \sum_{k\in\mathbb{Z}} \sum_{\gamma\in\mathcal{G}_k} \langle f,\, \psi_\gamma^k\rangle \psi_\gamma^k \qquad \text{in } L^2(\mathcal{X}).$$

Note that each ψ_γ^k is associated to a reference dyadic point z_γ^{k+1}. Moreover, ψ_γ^k is located near a random dyadic cube $Q_\gamma^{k+1}(\omega)$ and has exponential decay at infinity. The wavelets $\{\psi_\gamma^k :\ k \in \mathbb{Z},\ \gamma \in \mathcal{G}_k\}$ can be understood as a smooth version of the Haar wavelets (see [41]).

(iii) In view of (ii) and the Hölder regularity of $\mathbb{P}_k(x,y)$ in Lemma 5.3(iii), the multiresolution analysis $\{V_k\}_{k\in\mathbb{Z}}$ in Lemma 5.2 is referred to as a *smooth* MRA.

As a consequences of Lemmas 5.2 and 5.3, we arrive at the following conclusion.

Theorem 5.5. *Let $(\mathcal{X}, d, \mu)$ be a metric measure space satisfying (VD) and $\{V_k\}_{k\in\mathbb{Z}}$ be the* MRA *constructed in Lemma 5.2. Then, $\{V_k\}_{k\in\mathbb{Z}}$ is an admissible* MRA.

Proof. Let $\{\mathbb{P}_k\}_{k\in\mathbb{Z}}$ be as in Lemma 5.3. We are about to validate that the three conditions in Definition 2.3 are all satisfied.

Note that (A2) in Definition 2.3 follows directly from Lemma 5.3. Now, assuming $\mu(\mathcal{X}) = \infty$, we validate (A3) in Definition 2.3. For any $x, y \in \mathcal{X}$, note that Lemma 5.3(ii) yields

$$|\mathbb{P}_k(x, y)| \lesssim \frac{1}{\sqrt{V(x,\ \delta^k)V(y,\ \delta^k)}}.$$

If $k \to -\infty$, then $B(x, \delta^k) \to \mathcal{X}$ and, hence, $V(x,\ \delta^k) \to \mu(\mathcal{X}) = \infty$. In a similar way, we have $V(y,\ \delta^k) \to \mu(\mathcal{X}) = \infty$ as $k \to -\infty$. This leads to

$$\lim_{k\to-\infty} \mathbb{P}_k(x, y) = 0,$$

as desired.

It remains to show (A1) of Definition 2.3. To this end, for any $k \in \mathbb{Z}$ and $x, y \in \mathcal{X}$, we have by (1.1) that

$$\frac{V(x,\ \delta^k + d(x, y))}{\sqrt{V(x,\ \delta^k)V(y,\ \delta^k)}} \le C'_D \left(\frac{\delta^k + 2d(x, y)}{\delta^k}\right)^n \simeq \left(1 + \frac{d(x, y)}{\delta^k}\right)^n,$$

which, along with Lemma 5.3(ii), induces that

$$\begin{aligned}
|\mathbb{P}_k(x, y)| &\lesssim \frac{1}{V(x,\ \delta^k + d(x, y))}\left(1 + \frac{d(x, y)}{\delta^k}\right)^n \exp\left(-c_* \frac{d(x, y)}{\delta^k}\right)\\
&\lesssim \frac{1}{V(x,\ \delta^k + d(x, y))} \exp\left(-c\frac{d(x, y)}{\delta^k}\right)
\end{aligned}$$

holds for some constant $c \in (0, c^*)$. Consequently, we obtain

$$\begin{aligned}
\int_{\mathcal{X}} |\mathbb{P}_k(x, y)|\, d\mu(y) &\lesssim \int_{\mathcal{X}} \frac{1}{V(x,\ \delta^k + d(x, y))} \exp\left(-c\frac{d(x, y)}{\delta^k}\right) d\mu(y)\\
&\simeq \left(\int_{d(x,y)<\delta^k} + \sum_{j=0}^{\infty} \int_{\delta^{k-j}\le d(x,y)<\delta^{k-j-1}}\right) \frac{1}{V(x,\ \delta^k + d(x, y))} \exp\left(-c\frac{d(x, y)}{\delta^k}\right) d\mu(y)\\
&\lesssim \int_{d(x,y)<\delta^k} \frac{1}{V(x,\ \delta^k)}\, d\mu(y) + \sum_{j=0}^{\infty} \int_{\delta^{k-j}\le d(x,y)<\delta^{k-j-1}} \frac{1}{V(x,\ \delta^{k-j})} \exp\left(-c\delta^{-j}\right) d\mu(y)\\
&\lesssim 1 + \sum_{j=0}^{\infty} \exp\left(-c\delta^{-j}\right)\\
&\lesssim 1.
\end{aligned}$$

Thus, the condition (A1) in Definition 2.3 is satisfied. □

5.3 Construction of a stochastic complete signed heat kernel

Since we have already the smooth MRA (see Lemma 5.2), by using the smooth projectors $\{\mathbb{P}_k\}_{k\in\mathbb{Z}}$ and $\{\mathbb{Q}_k\}_{k\in\mathbb{Z}}$ from Lemma 5.3, we follow (2.4) and construct a family of functions $\{p_t\}_{t>0}$ (see (5.1) below). Such $\{p_t\}_{t>0}$ will be exactly the one required in Theorem 2.9.

Theorem 5.6. *Let $(\mathcal{X}, d, \mu)$ be a metric measure space satisfying (VD). Assume that $\{V_k\}_{k\in\mathbb{Z}}$ is the MRA in Lemma 5.2. For any $k \in \mathbb{Z}$, denote by $\mathbb{Q}_k$ the orthogonal projector from $L^2(\mathcal{X})$ to W_k, where $W_k = V_{k+1} \ominus V_k$. Fix $\beta \in (0,\infty)$. For any $t \in (0,\infty)$ and $x, y \in \mathcal{X}$, define*

$$p_t(x,y) := \begin{cases} \sum_{k\in\mathbb{Z}} e^{-t\delta^{-\beta k}} \mathbb{Q}_k(x,y) & \text{as } \mu(\mathcal{X}) = \infty; \\ \frac{1}{\mu(\mathcal{X})} + \sum_{k=m_{\mathcal{X}}}^{\infty} e^{-t\delta^{-\beta k}} \mathbb{Q}_k(x,y) & \text{as } \mu(\mathcal{X}) < \infty, \end{cases} \tag{5.1}$$

where $m_{\mathcal{X}}$ is as in (4.7). Then $\{p_t\}_{t>0}$ is a stochastic complete signed heat kernel.

Proof. As was proved in Theorem 5.5, we know that $\{V_k\}_{k\in\mathbb{Z}}$ in Lemma 5.2 is an admissible MRA. Let $\delta \in (0,1)$ be as in Theorem 5.1 and $\{\lambda_k\}_{k\in\mathbb{Z}}$ be the admissible spectrum given in Example 2.5. The remaining argument runs the same lines as that in the proof of Theorem 4.7. □

5.4 Stable-like upper estimate

In this subsection, we show that $\{p_t\}_{t>0}$ in (5.1) satisfies the stable-like upper estimate as stated in Theorem 2.9.

Theorem 5.7. *Let $(\mathcal{X}, d, \mu)$ be a metric measure space satisfying (VD). Suppose that $\beta \in (0,\infty)$ and $\{p_t\}_{t>0}$ is the signed heat kernel defined in (5.1). Then, there exists a positive constant C such that for all $t \in (0,\infty)$ and $x, y \in \mathcal{X}$,*

$$|p_t(x,y)| \le \frac{C}{V(x,\, t^{1/\beta} + d(x,y))} \left(\frac{t^{1/\beta}}{t^{1/\beta} + d(x,y)} \right)^{\beta}. \tag{5.2}$$

Proof. By symmetry, we may as well assume that $V(x, t^{1/\beta}) \ge V(y, t^{1/\beta})$. Let k_t be the unique integer such that

$$\delta^{k_t - 1} > t^{1/\beta} \ge \delta^{k_t}. \tag{5.3}$$

Consider first the case $\mu(\mathcal{X}) = \infty$. Then by (5.1) and (3.19) (taking $\lambda_k = \delta^{-\beta k}$ and $N = -k_t$ therein), we write

$$\begin{aligned} p_t(x,y) &= \sum_{k\ge k_t} e^{-t\delta^{-\beta k}} \mathbb{Q}_k(x,y) + \sum_{k\le k_t-1} e^{-t\delta^{-\beta k}} \mathbb{Q}_k(x,y) \\ &= \sum_{k\ge k_t} e^{-t\delta^{-\beta k}} \mathbb{Q}_k(x,y) + e^{-t\delta^{-\beta(k_t-1)}} \mathbb{P}_{k_t}(x,y) + \sum_{k\le k_t-1} \left(e^{-t\delta^{-\beta(k-1)}} - e^{-t\delta^{-\beta k}} \right) \mathbb{P}_k(x,y) \\ &=: \mathrm{Y}_1 + \mathrm{Y}_2 + \mathrm{Y}_3. \end{aligned} \tag{5.4}$$

To estimate Y_1, observe that any $k \ge k_t$ satisfies $\delta^{k\beta} \le t$. By this and (1.1), we deduce

$$\frac{V(x,\, t^{1/\beta} + d(x,y))}{V(x,\, \delta^k)} \le C_D' \left(\frac{t^{1/\beta} + d(x,y)}{\delta^k} \right)^n,$$

and

$$\frac{V(x,\, t^{1/\beta} + d(x,y))}{V(y,\, \delta^k)} \le C_D' \left(\frac{t^{1/\beta} + 2d(x,y)}{\delta^k} \right)^n,$$

where the constants C'_D and n are as in (1.1). Further, it follows from Lemma 5.3(ii) that

$$\begin{aligned}
|\mathrm{Y}_1| &\le \sum_{k\ge k_t} e^{-t\delta^{-\beta k}} |\mathbb{Q}_k(x,y)| \\
&\lesssim \sum_{k\ge k_t} e^{-t\delta^{-\beta k}} \frac{1}{\sqrt{V(x,\,\delta^k)\,V(y,\,\delta^k)}} \exp\left(-c_*\frac{d(x,y)}{\delta^k}\right) \\
&\lesssim \frac{1}{V(x,\,t^{1/\beta}+d(x,y))} \sum_{k\ge k_t} e^{-t\delta^{-\beta k}} \left(\frac{t^{1/\beta}+d(x,y)}{\delta^k}\right)^n \exp\left(-c_*\frac{d(x,y)}{\delta^k}\right).
\end{aligned}$$

Observing that

$$\exp\left(-c_*\frac{d(x,y)}{\delta^k}\right) \le \exp\left(-c_*\frac{d(x,y)}{t^{1/\beta}}\right) \quad \text{as } \delta^k \le t^{1/\beta}.$$

and

$$\left(\frac{t^{1/\beta}+d(x,y)}{t^{1/\beta}}\right)^{n+\beta} \exp\left(-c_*\frac{d(x,y)}{t^{1/\beta}}\right) \le \sup_{\tau>0}(1+\tau)^{n+\beta}e^{-c_*\tau} < \infty,$$

we then obtain

$$\begin{aligned}
\left(\frac{t^{1/\beta}+d(x,y)}{\delta^k}\right)^n \exp\left(-c_*\frac{d(x,y)}{\delta^k}\right) &\le \left(\frac{t^{1/\beta}}{\delta^k}\right)^n \left(\frac{t^{1/\beta}+d(x,y)}{t^{1/\beta}}\right)^n \exp\left(-c_*\frac{d(x,y)}{t^{1/\beta}}\right) \\
&\lesssim \left(\frac{t^{1/\beta}}{\delta^k}\right)^n \left(\frac{t^{1/\beta}}{t^{1/\beta}+d(x,y)}\right)^\beta
\end{aligned}$$

Consequently,

$$\begin{aligned}
|\mathrm{Y}_1| &\lesssim \frac{1}{V(x,\,t^{1/\beta}+d(x,y))} \left(\frac{t^{1/\beta}}{t^{1/\beta}+d(x,y)}\right)^\beta \sum_{k\ge k_t} e^{-t\delta^{-\beta k}} \left(\frac{t^{1/\beta}}{\delta^k}\right)^n \\
&\lesssim \frac{1}{V(x,\,t^{1/\beta}+d(x,y))} \left(\frac{t^{1/\beta}}{t^{1/\beta}+d(x,y)}\right)^\beta.
\end{aligned}$$

Now, we consider Y_2. Indeed, by Lemma 5.3 and $\delta^{k_t\beta} \simeq t$, together with (1.1) and the fact

$$\sup_{\tau>0}(1+\tau)^{n+\beta}e^{-c_*\tau} < \infty,$$

we also derive

$$\begin{aligned}
|\mathrm{Y}_2| &\lesssim e^{-t\delta^{-\beta k_t}} \frac{1}{\sqrt{V(x,\,\delta^{k_t})\,V(y,\,\delta^{k_t})}} \exp\left(-c_*\frac{d(x,y)}{\delta^{k_t}}\right) \\
&\lesssim \frac{1}{V(x,\,\delta^{k_t}+d(x,y))} \left(1+\frac{d(x,y)}{\delta^{k_t}}\right)^n \exp\left(-c_*\frac{d(x,y)}{\delta^{k_t+1}}\right) \\
&\lesssim \frac{1}{V(x,\,\delta^{k_t}+d(x,y))} \left(1+\frac{d(x,y)}{\delta^{k_t}}\right)^{-\beta} \\
&\simeq \frac{1}{V(x,\,t^{1/\beta}+d(x,y))} \left(\frac{t^{1/\beta}}{t^{1/\beta}+d(x,y)}\right)^\beta.
\end{aligned}$$

as desired.

It remains to estimate Y_3. If $k \le k_t - 1$, then $\delta^{k\beta} > t$ and, hence,

$$\left| e^{-t\delta^{-\beta(k-1)}} - e^{-t\delta^{-\beta k}} \right| = e^{-t\delta^{-\beta k}} \left| e^{t\delta^{-\beta k}(1-\delta^{\beta})} - 1 \right| \simeq \frac{t}{\delta^{k\beta}}.$$

This, combined with Lemma 5.3(ii), further yields

$$\begin{aligned} |\mathrm{Y}_3| &\le \sum_{k \le k_t - 1} \left| e^{-t\delta^{-\beta(k-1)}} - e^{-t\delta^{-\beta k}} \right| |\mathbb{P}_k(x, y)| \\ &\lesssim \sum_{k \le k_t - 1} \frac{t}{\delta^{k\beta}} \frac{1}{\sqrt{V(x,\, \delta^k)\, V(y,\, \delta^k)}} \exp\left(-c_* \frac{d(x, y)}{\delta^k} \right). \end{aligned}$$

By using (1.1) (when $t^{1/\beta} + d(x, y) \ge \delta^k$), we see that

$$\frac{V(x,\, t^{1/\beta} + d(x, y))}{V(x,\, \delta^k)} \lesssim \left(\frac{t^{1/\beta} + d(x, y)}{\delta^k} \right)^{\sigma}$$

holds uniformly in $\{x, y, t, k\}$, where

$$\sigma = \begin{cases} n & \text{as } t^{1/\beta} + d(x, y) \ge \delta^k; \\ 0 & \text{as } t^{1/\beta} + d(x, y) < \delta^k, \end{cases}$$

where n is as in (1.1). In a similar manner, if we change $V(x,\, \delta^k)$ to $V(y,\, \delta^k)$, then we still have

$$\frac{V(x,\, t^{1/\beta} + d(x, y))}{V(y,\, \delta^k)} \simeq \frac{V(y,\, t^{1/\beta} + d(x, y))}{V(y,\, \delta^k)} \lesssim \left(\frac{t^{1/\beta} + d(x, y)}{\delta^k} \right)^{\sigma}.$$

With these, we continue with the estimate of $|\mathrm{Y}_3|$. If $d(x, y) \le t^{1/\beta}$, then we have by $\delta^{k_t} \simeq t^{1/\beta}$ that

$$\begin{aligned} |\mathrm{Y}_3| &\lesssim \frac{1}{V(x,\, t^{1/\beta} + d(x, y))} \sum_{k \le k_t - 1} \left(\frac{t}{\delta^k} \right)^{\beta} \left(\frac{t^{1/\beta} + d(x, y)}{\delta^k} \right)^{\sigma} \exp\left(-c_* \frac{d(x, y)}{\delta^k} \right) \\ &\lesssim \frac{1}{V(x,\, t^{1/\beta} + d(x, y))} \sum_{k \le k_t - 1} \left(\frac{t}{\delta^k} \right)^{\beta + \sigma} \\ &\simeq \frac{1}{V(x,\, t^{1/\beta} + d(x, y))} \\ &\simeq \frac{1}{V(x,\, t^{1/\beta} + d(x, y))} \left(\frac{t^{1/\beta}}{t^{1/\beta} + d(x, y)} \right)^{\beta}. \end{aligned}$$

If $d(x, y) > t^{1/\beta}$, then

$$\begin{aligned} |\mathrm{Y}_3| &\lesssim \frac{1}{V(x,\, t^{1/\beta} + d(x, y))} \frac{t}{d(x, y)^{\beta}} \sum_{k \le k_t - 1} \left(\frac{d(x, y)}{\delta^k} \right)^{\beta + \sigma} \exp\left(-c_* \frac{d(x, y)}{\delta^k} \right) \\ &\simeq \frac{1}{V(x,\, t^{1/\beta} + d(x, y))} \frac{t}{d(x, y)^{\beta}} \sum_{k \le k_t - 1} \int_{\delta^k}^{\delta^{k-1}} \left(\frac{d(x, y)}{\delta^k} \right)^{\beta + \sigma} \exp\left(-c_* \frac{d(x, y)}{\delta^k} \right) \frac{ds}{s} \end{aligned}$$

$$
\begin{aligned}
&\lesssim \frac{1}{V(x,\, t^{1/\beta}+d(x,y))}\frac{t}{d(x,y)^\beta}\int_{\delta^{k_t-1}}^{\infty}\left(\frac{d(x,y)}{s}\right)^{\beta+\sigma}\exp\left(-c_*\frac{d(x,y)}{s}\right)\frac{ds}{s}\\
&\lesssim \frac{1}{V(x,\, t^{1/\beta}+d(x,y))}\frac{t}{d(x,y)^\beta}\int_{0}^{\infty}\tau^{\beta+\sigma}\exp\left(-c_*\tau\right)\frac{d\tau}{\tau}\\
&\simeq \frac{1}{V(x,\, t^{1/\beta}+d(x,y))}\frac{t}{d(x,y)^\beta}\\
&\simeq \frac{1}{V(x,\, t^{1/\beta}+d(x,y))}\left(\frac{t^{1/\beta}}{t^{1/\beta}+d(x,y)}\right)^{\beta}.
\end{aligned}
$$

Substituting the estimates of Y_1 through Y_3 into (5.4), we derive the desired estimate of $p_t(x,y)$ in (5.2) under $\mu(\mathcal{X})=\infty$.

Now, we consider the case $\mu(\mathcal{X})<\infty$. Let $k_t\in\mathbb{Z}$ be as in (5.3). If $k_t\le m_{\mathcal{X}}$, then the formula (5.1) directly gives

$$
p_t(x,y)=\sum_{k\ge m_{\mathcal{X}}} e^{-t\delta^{-\beta k}}\mathbb{Q}_k(x,y)+\mathbb{P}_{m_{\mathcal{X}}}(x,y)=:\mathrm{Y}_1'+\mathrm{Y}_2'. \tag{5.5}
$$

Since any $k\ge m_{\mathcal{X}}$ also satisfies $\delta^{k\beta}\le t$, we find that the same estimate of Y_1 also implies

$$
|\mathrm{Y}_1'|\lesssim \frac{1}{V(x,\, t^{1/\beta}+d(x,y))}\left(\frac{t^{1/\beta}}{t^{1/\beta}+d(x,y)}\right)^{\beta}.
$$

By (4.7), the fact $t^{1/\beta}\ge\delta^{k_t}\ge\delta^{m_{\mathcal{X}}}$ and that

$$
\mathbb{P}_{m_{\mathcal{X}}}(x,y)=\frac{1}{\mu(\mathcal{X})}\quad \text{for all } x,y\in\mathcal{X},
$$

we have

$$
|\mathrm{Y}_2'|=\frac{1}{\mu(\mathcal{X})}\lesssim\frac{1}{V(x,\,\delta^{m_{\mathcal{X}}}+d(x,y))}\left(1+\frac{d(x,y)}{\delta^{m_{\mathcal{X}}}}\right)^{-\beta}\lesssim\frac{1}{V(x,\, t^{1/\beta}+d(x,y))}\left(\frac{t^{1/\beta}}{t^{1/\beta}+d(x,y)}\right)^{\beta}.
$$

So, inserting the estimates of Y_1' and Y_2' into (5.5), we derive that (5.2) also holds under the situation $\mu(\mathcal{X})<\infty$ and $k_t\le m_{\mathcal{X}}$.

If $k_t\ge m_{\mathcal{X}}+1$, then by (5.1) and (3.18) (taking $\lambda_k=\delta^{-\beta k}$, $m=m_{\mathcal{X}}$ and $N=-k_t$ therein), we write

$$
\begin{aligned}
p_t(x,y)&=\sum_{k\ge k_t} e^{-t\delta^{-\beta k}}\mathbb{Q}_k(x,y)+\sum_{k=m_{\mathcal{X}}}^{k_t-1} e^{-t\delta^{-\beta k}}\mathbb{Q}_k(x,y)+\mathbb{P}_{m_{\mathcal{X}}}(x,y)\\
&=\sum_{k\ge k_t} e^{-t\delta^{-\beta k}}\mathbb{Q}_k(x,y)+e^{-t\delta^{-\beta(k_t-1)}}\mathbb{P}_{k_t}(x,y)+\sum_{k=m_{\mathcal{X}}}^{k_t-1}\left(e^{-t\delta^{-\beta(k-1)}}-e^{-t\delta^{-\beta k}}\right)\mathbb{P}_k(x,y)\\
&=:\mathrm{Y}_1+\mathrm{Y}_2+\mathrm{Y}_3'.
\end{aligned}\tag{5.6}
$$

Note that the estimate of $|Y_3|$ remains true for $|Y_3|$. By this and the estimates of Y_1 and Y_2, we derive from (5.6) the desired upper estimate of $p_t(x,y)$ in (5.2) under $\mu(\mathcal{X})<\infty$ and $k_t\ge m_{\mathcal{X}}+1$.

Altogether, we conclude the proof of Theorem 5.7. □

5.5 Almost Lipschitz regularity and near-diagonal lower estimate

The maim aim of this subsection is to show that $\{p_t\}_{t>0}$ in (5.1) satisfies the almost Lipschitz regularity (see Theorem 5.8 below) and near-diagonal lower estimate (see Theorem 5.9 below) as stated in Theorem 2.9.

Theorem 5.8. *Let $(\mathcal{X}, d, \mu)$ be a metric measure space satisfying (VD). Suppose that $\beta \in (0, \infty)$ and $\{p_t\}_{t>0}$ is the signed heat kernel defined in (5.1). Then, there exists a positive constant C such that for any $t \in (0, \infty)$ and $x, y, y' \in \mathcal{X}$ satisfying $d(y, y') \le t^{1/\beta}$,*

$$|p_t(x, y) - p_t(x, y')| \le C\left(\frac{d(y, y')}{t^{1/\beta}}\right)^{\eta_\delta} \frac{1}{V(x,\ t^{1/\beta} + d(x, y))}\left(\frac{t^{1/\beta}}{t^{1/\beta} + d(x, y)}\right)^{\beta}, \tag{5.7}$$

where $\eta_\delta \in (0, 1)$ is the constant from Theorem 5.1(vii).

Proof. Let $k_t \in \mathbb{Z}$ be the unique integer satisfy (5.3), that is, $\delta^{(k_t-1)\beta} > t \ge \delta^{k_t\beta}$. If $\mu(\mathcal{X}) = \infty$, then by using (5.4), we have

$$\begin{aligned}
|p_t(x, y) - p_t(x, y')| \le{}& \sum_{k\ge k_t} e^{-t\delta^{-\beta k}}\left|\mathbb{Q}_k(x, y) - \mathbb{Q}_k(x, y')\right| \\
&+ e^{-t\delta^{-\beta(k_t-1)}}\left|\mathbb{P}_{k_t}(x, y) - \mathbb{P}_{k_t}(x, y')\right| \\
&+ \sum_{k\le k_t-1}\left(e^{-t\delta^{-\beta(k-1)}} - e^{-t\delta^{-\beta k}}\right)\left|\mathbb{P}_k(x, y) - \mathbb{P}_k(x, y')\right| \\
=:{}& \mathrm{Z}_1 + \mathrm{Z}_2 + \mathrm{Z}_3.
\end{aligned} \tag{5.8}$$

The estimates of $\mathrm{Z}_1, \mathrm{Z}_2$ and Z_3 are similar to that of $\mathrm{Y}_1, \mathrm{Y}_2, \mathrm{Y}_3$ in Theorem 5.7.

First we estimate Z_1. When $d(y, y') < \delta^k$, applying Lemma 5.3(iii) and using (VD), we derive

$$|\mathbb{Q}_k(x, y) - \mathbb{Q}_k(x, y')| \lesssim \left(\frac{d(y, y')}{\delta^k}\right)^{\eta_\delta} \frac{1}{\sqrt{V(x,\ \delta^k)\, V(y,\ \delta^k)}} \exp\left(-c_*\frac{d(x, y)}{\delta^k}\right). \tag{5.9}$$

If $d(y, y') \ge \delta^k$, then applying (VD) and the size condition of $\mathbb{Q}_k$ in Lemma 5.3(i), we derive that

$$\begin{aligned}
|\mathbb{Q}_k(x, y) - \mathbb{Q}_k(x, y')| &\le |\mathbb{Q}_k(x, y)| + |\mathbb{Q}_k(x, y')| \\
&\lesssim \left(\frac{d(y, y')}{\delta^k}\right)^{\eta_\delta} \frac{1}{\sqrt{V(x,\ \delta^k)\, V(y,\ \delta^k)}} \exp\left(-c_*\frac{d(x, y)}{\delta^k}\right) \\
&\quad + \left(\frac{d(y, y')}{\delta^k}\right)^{\eta_\delta} \frac{1}{\sqrt{V(x,\ \delta^k)\, V(y', \delta^k)}} \exp\left(-c_*\frac{d(x, y')}{\delta^k}\right).
\end{aligned} \tag{5.10}$$

When $k \ge k_t$, we observe that $\delta^k \le \delta^{k_t} \le t^{1/\beta}$, which implies

$$\exp\left(-c_*\frac{d(x, y)}{\delta^k}\right) \le \exp\left(-c_*\frac{d(x, y)}{t^{1/\beta}}\right) \quad \text{and} \quad \exp\left(-c_*\frac{d(x, y')}{\delta^k}\right) \le \exp\left(-c_*\frac{d(x, y')}{t^{1/\beta}}\right).$$

Regarding the second inequality in the last formula, since $d(y, y') \le t^{1/\beta}$, it follows that

$$d(x, y) \le d(x, y') + d(y', y) \le d(x, y') + t^{1/\beta},$$

thereby leading to

$$\exp\left(-c\frac{d(x,y')}{t^{1/\beta}}\right) \le \exp\left(-c_*\frac{d(x,y)-t^{1/\beta}}{t^{1/\beta}}\right) = e^{c_*}\exp\left(-c_*\frac{d(x,y)}{t^{1/\beta}}\right).$$

Moreover, by $\delta^k \le t^{1/\beta}$, (1.1) and the fact $d(x,y') \le d(x,y) + d(y,y') \le d(x,y) + t^{1/\beta}$, we have

$$\frac{V(x,\ t^{1/\beta}+d(x,y))}{V(y',\delta^k)} \lesssim \left(\frac{t^{1/\beta}+d(x,y)+d(x,y')}{\delta^k}\right)^n \lesssim \left(\frac{t^{1/\beta}+d(x,y)}{\delta^k}\right)^n.$$

Similarly, we also have

$$\frac{V(x,\ t^{1/\beta}+d(x,y))}{V(x,\ \delta^k)} \lesssim \left(\frac{t^{1/\beta}+d(x,y)}{\delta^k}\right)^n \quad \text{and} \quad \frac{V(x,\ t^{1/\beta}+d(x,y))}{V(y,\ \delta^k)} \lesssim \left(\frac{t^{1/\beta}+d(x,y)}{\delta^k}\right)^n.$$

Invoking these facts, we derive from (5.9) and (5.10) that when $d(y,y') \le t^{1/\beta}$ it always holds

$$\begin{aligned}
&|\mathbb{Q}_k(x,y) - \mathbb{Q}_k(x,y')| \\
&\quad\lesssim \left(\frac{d(y,y')}{\delta^k}\right)^{\eta_\delta} \frac{1}{V(x,\ t^{1/\beta}+d(x,y))}\left(\frac{t^{1/\beta}+d(x,y)}{\delta^k}\right)^n \exp\left(-c_*\frac{d(x,y)}{t^{1/\beta}}\right)\\
&\quad\simeq \left(\frac{d(y,y')}{\delta^k}\right)^{\eta_\delta}\left(\frac{t^{1/\beta}}{\delta^k}\right)^n \frac{1}{V(x,\ t^{1/\beta}+d(x,y))}\left(1+\frac{d(x,y)}{t^{1/\beta}}\right)^n \exp\left(-c_*\frac{d(x,y)}{t^{1/\beta}}\right)\\
&\quad\lesssim \left(\frac{d(y,y')}{\delta^k}\right)^{\eta_\delta}\left(\frac{t^{1/\beta}}{\delta^k}\right)^n \frac{1}{V(x,\ t^{1/\beta}+d(x,y))}\left(1+\frac{d(x,y)}{t^{1/\beta}}\right)^{-\beta}\\
&\quad\simeq \left(\frac{t^{1/\beta}}{\delta^k}\right)^{\eta_\delta+n}\left(\frac{d(y,y')}{t^{1/\beta}}\right)^{\eta_\delta} \frac{1}{V(x,\ t^{1/\beta}+d(x,y))}\left(\frac{t^{1/\beta}}{t^{1/\beta}+d(x,y)}\right)^{\beta},
\end{aligned} \tag{5.11}$$

where the penultimate step is due to the fact that $\sup_{\tau>0}(1+\tau)^{n+\beta}e^{-c_*\tau} < \infty$. As a consequence of (5.11), we arrive at the conclusion

$$\begin{aligned}
\mathrm{Z}_1 &\lesssim \left(\frac{d(y,y')}{t^{1/\beta}}\right)^{\eta_\delta} \frac{1}{V(x,\ t^{1/\beta}+d(x,y))}\left(\frac{t^{1/\beta}}{t^{1/\beta}+d(x,y)}\right)^{\beta}\sum_{k\ge k_t} e^{-t\delta^{-\beta k}}\left(\frac{t^{1/\beta}}{\delta^k}\right)^{\eta_\delta+n}\\
&\lesssim \left(\frac{d(y,y')}{t^{1/\beta}}\right)^{\eta_\delta} \frac{1}{V(x,\ t^{1/\beta}+d(x,y))}\left(\frac{t^{1/\beta}}{t^{1/\beta}+d(x,y)}\right)^{\beta},
\end{aligned}$$

as desired.

The estimate of Z_2 is similar to that of Z_1, but using the size and Hölder regularity conditions of $\mathbb{P}_k$ that was given in Lemma 5.3; we thus omit the details.

To estimate Z_3, we observe that for $k \le k_t - 1$ there is $d(y,y') \le t^{1/\beta} < \delta^{k_t-1} \le \delta^k$, which, combined with Lemma 5.3(ii), leads to

$$|\mathbb{P}_k(x,y) - \mathbb{P}_k(x,y')| \lesssim \left(\frac{d(y,y')}{\delta^k}\right)^{\eta_\delta} \frac{1}{\sqrt{V(x,\ \delta^k)V(y,\ \delta^k)}}\exp\left(-c_*\frac{d(x,y)}{\delta^k}\right).$$

With this, we proceed as the estimate of Y_3 in Theorem 5.7, thereby obtaining

$$|\mathrm{Z}_3| \lesssim \sum_{k\le k_t-1}\left|e^{-t\delta^{-\beta(k-1)}} - e^{-t\delta^{-\beta k}}\right|\left(\frac{d(y,y')}{\delta^k}\right)^{\eta_\delta}\frac{1}{\sqrt{V(x,\ \delta^k)V(y,\ \delta^k)}}\exp\left(-c_*\frac{d(x,y)}{\delta^k}\right)$$

$$\simeq \left(\frac{d(y,y')}{t^{1/\beta}}\right)^{\eta_\delta} \sum_{k\le k_t-1} \frac{t}{\delta^{\beta k}} \frac{1}{\sqrt{V(x,\,\delta^k)V(y,\,\delta^k)}} \exp\left(-c_* \frac{d(x,y)}{\delta^k}\right)$$
$$\lesssim \left(\frac{d(y,y')}{t^{1/\beta}}\right)^{\eta_\delta} \frac{1}{V(x,\,t^{1/\beta}+d(x,y))} \left(\frac{t^{1/\beta}}{t^{1/\beta}+d(x,y)}\right)^{\beta}.$$

Combining the estimates of Z_1, Z_2 and Z_3, we derive from (5.8) that (5.7) holds when $\mu(\mathcal{X}) = \infty$.

Under the case $\mu(\mathcal{X}) < \infty$, we consider the cases $k_t \le m_{\mathcal{X}}$ and $k_t \ge m_{\mathcal{X}} + 1$, respectively. Instead of (5.8), we apply (5.5) and (5.6), thereby writing

$$|p_t(x,y) - p_t(x,y')| \le \sum_{k\ge m_{\mathcal{X}}} e^{-t\delta^{-\beta k}} \left|\mathbb{Q}_k(x,y) - \mathbb{Q}_k(x,y')\right| + \left|\mathbb{P}_{m_{\mathcal{X}}}(x,y) - \mathbb{P}_{m_{\mathcal{X}}}(x,y')\right|$$

and

$$|p_t(x,y) - p_t(x,y')| \le \sum_{k\ge k_t} e^{-t\delta^{-\beta k}} \left|\mathbb{Q}_k(x,y) - \mathbb{Q}_k(x,y')\right| + e^{-t\delta^{-\beta(k_t-1)}} \left|\mathbb{P}_{k_t}(x,y) - \mathbb{P}_{k_t}(x,y')\right|$$
$$+ \sum_{k=m_{\mathcal{X}}}^{k_t-1} \left(e^{-t\delta^{-\beta(k-1)}} - e^{-t\delta^{-\beta k}}\right) \left|\mathbb{P}_k(x,y) - \mathbb{P}_k(x,y')\right|.$$

Running with almost the same lines as that of Z_1 through Z_3, we can estimate the terms in the right hands of the above two formulae, thereby leading to (5.7); the details are omitted. This finishes the proof of Theorem 5.8. □

Based on the Hölder regularity estimate in Theorem 5.8 and the stochastic completeness property in Theorem 5.6, we now give the near-diagonal lower estimate of $\{p_t\}_{t>0}$.

Theorem 5.9. *Let $(\mathcal{X}, d, \mu)$ be a metric measure space satisfying (VD). Suppose that $\beta \in (0,\infty)$ and $\{p_t\}_{t>0}$ is the signed heat kernel defined in (5.1). Then, for any $t \in (0,\infty)$ and $x \in \mathcal{X}$,*

$$p_t(x,x) \ge \frac{C}{V(x,\,t^{1/\beta})}. \tag{5.12}$$

Consequently, there exists a small constant $\varepsilon \in (0,1)$ such that for any $t \in (0,\infty)$ and $x, y \in \mathcal{X}$ satisfying $d(x,y) \le \varepsilon t^{1/\beta}$,

$$p_t(x,y) \ge \frac{C}{V(x,\,t^{1/\beta})}. \tag{5.13}$$

Here, the constants C and ε in (5.12) and (5.13) are positive and independent of t and x, y.

Proof. Once we have proved (5.12), then (5.13) follows from (5.12) and Theorem 5.8, by writing

$$p_t(x,y) \ge p_t(x,x) - |p_t(x,x) - p_t(x,y)|$$

and letting $d(x,y) \le \varepsilon t^{1/\beta}$ with ε being a sufficiently small positive number.

It remains to show (5.12). Let $\lambda \in (1,\infty)$, which will be determined later. For any $t \in (0,\infty)$ and $x \in \mathcal{X}$, by the semigroup property, the Hölder inequality and $\int_{\mathcal{X}} p_t(x,z)\, d\mu(z) = 1$, we have

$$p_{2t}(x,x) = \int_{\mathcal{X}} p_t(x,z) p_t(z,x)\, d\mu(z)$$

$$
\begin{aligned}
&= \int_{\mathcal{X}} p_t(x,z)^2 \, d\mu(z) \\
&\geq \int_{B(x,\lambda t^{1/\beta})} p_t(x,z)^2 \, d\mu(z) \\
&\geq \frac{1}{V(x,\ \lambda t^{1/\beta})} \left(\int_{B(x,\lambda t^{1/\beta})} p_t(x,z) \, d\mu(z) \right)^2 \\
&= \frac{1}{V(x,\ \lambda t^{1/\beta})} \left(1 - \int_{d(z,x) \geq \lambda t^{1/\beta}} p_t(x,z) \, d\mu(z) \right)^2 \\
&= \frac{1}{V(x,\ \lambda t^{1/\beta})} \left(1 - \int_{d(z,x) \geq \lambda t^{1/\beta}} |p_t(x,z)| \, d\mu(z) \right)^2 .
\end{aligned}
$$

Moreover, note that Theorem 5.7 implies

$$
\begin{aligned}
\int_{d(z,x) \geq \lambda t^{1/\beta}} |p_t(x,z)| \, d\mu(z) &\lesssim \int_{d(z,x) \geq \lambda t^{1/\beta}} \frac{1}{V(x,\ t^{1/\beta} + d(x,z))} \left(\frac{t^{1/\beta}}{t^{1/\beta} + d(x,z)} \right)^{\beta} d\mu(z) \\
&\lesssim \sum_{j=1}^{\infty} \int_{2^{j-1}\lambda t^{1/\beta} \leq d(z,x) < 2^j \lambda t^{1/\beta}} \frac{1}{V(x,\ t^{1/\beta} + d(z,\ x))} \left(\frac{t^{1/\beta}}{t^{1/\beta} + d(z,\ x)} \right)^{\beta} d\mu(z) \\
&\lesssim \sum_{j=1}^{\infty} \left(2^j \lambda\right)^{-\beta} \frac{V(x,\ 2^j \lambda t^{1/\beta})}{V(x,\ t^{1/\beta} + 2^{j-1} \lambda t^{1/\beta})} \\
&\lesssim \sum_{j=1}^{\infty} \left(2^j \lambda\right)^{-\beta} \\
&\lesssim \lambda^{-\beta}.
\end{aligned}
$$

In other words, there exists a positive constant C_0 (independent of t, x and λ) such that

$$
\int_{d(z,x) \geq \lambda t^{1/\beta}} |p_t(x,z)| \, d\mu(z) \leq C_0 \lambda^{-\beta}.
$$

Since $\beta > 0$, we choose $\lambda > 1$ large enough so that $C_0 \lambda^{-\beta} < 1/2$, thereby giving

$$
p_{2t}(x,x) \geq \frac{1}{4V(x,\ \lambda t^{1/\beta})} \gtrsim \lambda^{-n} \frac{1}{V(x, t^{1/\beta})}
$$

in terms of (VD). This ends the proof of (5.12). □

Remark 5.10. Theorem 2.9 is an immediate consequence of Theorems 5.6, 5.7, 5.8 and 5.9.

References

[1] H. Aimar, A. Bernardis, and B. Iaffei, *Multiresolution approximations and unconditional bases on weighted Lebesgue spaces on spaces of homogeneous type*, J. Approx. Theory **148** (2007), no. 1, 12–34.

[2] P. Auscher and T. Hytönen, *Orthonormal bases of regular wavelets in spaces of homogeneous type*, Appl. Comput. Harmon. Anal. **34** (2013), no. 2, 266–296; Appl. Comput. Harmon. Anal. **39** (2015), 568–569.

[3] P. Auscher and M. Qafsaoui, *Equivalence between regularity theorems and heat kernel estimates for higher order elliptic operators and systems under divergence form*, J. Funct. Anal. **177** (2000), no. 2, 310–364.

[4] I. G. Avramidi, *Heat Kernel and Quantum Gravity*, Lecture Notes in Physics Monographs, vol. 64, Springer Berlin Heidelberg, Berlin, Heidelberg, 2000.

[5] M. T. Barlow, *Random walks and Heat Kernels on Graphs*, London Mathematical Society Lecture Note Series, vol. 438, Cambridge University Press, Cambridge, 2017.

[6] M. T. Barlow and R. F. Bass, *Brownian motion and harmonic analysis on Sierpinski carpets*, Canad. J. Math. **51** (1999), no. 4, 673–744.

[7] M. T. Barlow, A. Grigor'yan, and T. Kumagai, *On the equivalence of parabolic Harnack inequalities and heat kernel estimates*, J. Math. Soc. Japan **64** (2012), no. 4, 1091–1146.

[8] M. T. Barlow and E. A. Perkins, *Brownian motion on the Sierpiński gasket*, Probab. Theory Related Fields **79** (1988), no. 4, 543–623.

[9] R. F. Bass and D. A. Levin, *Transition probabilities for symmetric jump processes*, Trans. Amer. Math. Soc. **354** (2002), no. 7, 2933–2953.

[10] A. D. Bendikov, A. Grigor'yan, K. Pittè, and V. Vëss, *Isotropic Markov semigroups on ultra-metric spaces*, Uspekhi Mat. Nauk **69** (2014), no. 4(418), 3–102; translation in Russian Math. Surveys **69** (2014), no.4, 589–680.

[11] N. Berline, E. Getzler, and M. Vergne, *Heat Kernels and Dirac Operators*, Grundlehren Text Editions, Springer-Verlag, Berlin, 2004.

[12] J. Cao, A. Grigor'yan, E. Hu, and L. Liu, *Regularity of jump-type Dirichlet forms on metric measure spaces*, (2025), Submitted.

[13] J. Cao, A. Grigor'yan, E. Hu, and L. Liu, *Existence of stable-like heat kernels and stability of the critical index on metric measure spaces*, (2025), Submitted.

[14] I. Chavel, *Eigenvalues in Riemannian Geometry*, Pure and Applied Mathematics, vol. 115, Academic Press, Inc., Orlando, FL, 1984, Including a chapter by Burton Randol, With an appendix by Jozef Dodziuk.

[15] Z.-Q. Chen and T. Kumagai, *Heat kernel estimates for stable-like processes on d-sets*, Stochastic Process. Appl. **108** (2003), no. 1, 27–62.

[16] Z.-Q. Chen, T. Kumagai, and J. Wang, *Stability of Heat Kernel Estimates for Symmetric Non-local Dirichlet Forms*, Mem. Amer. Math. Soc. **271** (2021), no. 1330, v+89.

[17] M. Christ, *A $T(b)$ theorem with remarks on analytic capacity and the Cauchy integral*, Colloq. Math. **60/61** (1990), no. 2, 601–628.

[18] O. Christensen, *An Introduction to Frames and Riesz Bases*, second ed., Applied and Numerical Harmonic Analysis, Birkhäuser/Springer, [Cham], 2016.

[19] R. R. Coifman and M. Maggioni, *Diffusion wavelets*, Appl. Comput. Harmon. Anal. **21** (2006), no. 1, 53–94.

[20] R. R. Coifman and G. Weiss, *Analyse Harmonique Non-commutative sur Certains Espaces Homogènes*, Étude de certaines intégrales singulières, Lecture Notes in Mathematics, vol. Vol. 242, Springer-Verlag, Berlin-New York, 1971.

[21] T. Coulhon, G. Kerkyacharian, and P. Petrushev, *Heat kernel generated frames in the setting of Dirichlet spaces*, J. Fourier Anal. Appl. **18** (2012), no. 5, 995–1066.

[22] I. Daubechies, *Ten Lectures on Wavelets*, CBMS-NSF Regional Conference Series in Applied Mathematics, vol. 61, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992.

[23] E. B. Davies, *Heat Kernels and Spectral Theory*, Cambridge Tracts in Mathematics, vol. 92, Cambridge University Press, Cambridge, 1990.

[24] ———, *Uniformly elliptic operators with measurable coefficients*, J. Funct. Anal. **132** (1995), no. 1, 141–169.

[25] ———, *L^p spectral theory of higher-order elliptic differential operators*, Bull. London Math. Soc. **29** (1997), no. 5, 513–546.

[26] X. T. Duong and D. W. Robinson, *Semigroup kernels, Poisson bounds, and holomorphic functional calculus*, J. Funct. Anal. **142** (1996), no. 1, 89–128.

[27] X. T. Duong and L. Yan, *Duality of Hardy and BMO spaces associated with operators with heat kernel bounds*, J. Amer. Math. Soc. **18** (2005), no. 4, 943–973.

[28] M. Fukushima, Y. Oshima, and M. Takeda, *Dirichlet Forms and Symmetric Markov Processes*, extended ed., De Gruyter Studies in Mathematics, vol. 19, Walter de Gruyter & Co., Berlin, 2011.

[29] F. Gazzola and H.-Christoph Grunau, *Some new properties of biharmonic heat kernels*, Nonlinear Anal. **70** (2009), no. 8, 2965–2973.

[30] A. Grigor'yan, *The heat equation on noncompact Riemannian manifolds*, Mat. Sb. **182** (1991), no. 1, 55–87, translation in Math. USSR-Sb. **72** (1992), no. 1, 47–77.

[31] ———, *Heat kernels and function theory on metric measure spaces*, Heat kernels and analysis on manifolds, graphs, and metric spaces (Paris, 2002), Contemp. Math., vol. 338, Amer. Math. Soc., Providence, RI, 2003, pp. 143–172.

[32] ———, *Heat Kernel and Analysis on Manifolds*, AMS/IP Studies in Advanced Mathematics, vol. 47, American Mathematical Society, Providence, RI; International Press, Boston, MA, 2009.

[33] ———, *Introduction to Analysis on Graphs*, University Lecture Series, vol. 71, American Mathematical Society, Providence, RI, 2018.

[34] A. Grigor'yan, E. Hu, and J. Hu, *Two-sided estimates of heat kernels of jump type Dirichlet forms*, Adv. Math. **330** (2018), 433–515.

[35] A. Grigor'yan and J. Hu, *Upper bounds of heat kernels on doubling spaces*, Mosc. Math. J. **14** (2014), no. 3, 505–563, 641–642.

[36] T. Hytönen, *A framework for non-homogeneous analysis on metric spaces, and the RBMO space of Tolsa*, Publ. Mat. **54** (2010), no. 2, 485–504.

[37] T. Hytönen and A. Kairema, *Systems of dyadic cubes in a doubling metric space*, Colloq. Math. **126** (2012), no. 1, 1–33.

[38] T. Hytönen and H. Martikainen, *Non-homogeneous Tb theorem and random dyadic cubes on metric measure spaces*, J. Geom. Anal. **22** (2012), no. 4, 1071–1107.

[39] T. Hytönen and O. Tapiola, *Almost Lipschitz-continuous wavelets in metric spaces via a new randomization of dyadic cubes*, J. Approx. Theory **185** (2014), 12–30.

[40] J. Jorgenson and S. Lang, *The ubiquitous heat kernel*, Mathematics unlimited—2001 and beyond, Springer, Berlin, 2001, pp. 655–683.

[41] A. Kairema, J. Li, M. C. Pereyra, and L. A. Ward, *Haar bases on quasi-metric measure spaces, and dyadic structure theorems for function spaces on product spaces of homogeneous type*, J. Funct. Anal. **271** (2016), no. 7, 1793–1843.

[42] J. Kigami, *Analysis on fractals*, Cambridge Tracts in Mathematics, vol. 143, Cambridge University Press, Cambridge, 2001.

[43] ———, *Volume Doubling Measures and Heat Kernel Estimates on Self-similar Sets*, Mem. Amer. Math. Soc. **199** (2009), no. 932, viii+94.

[44] E. Kılıç and P. Stanica, *The inverse of banded matrices*, J. Comput. Appl. Math. **237** (2013), no. 1, 126–135.

[45] S. V. Kozyrev, A. Yu. Khrennikov, and V. M. Shelkovich, *p-adic wavelets and their applications*, Proc. Steklov Inst. Math. **285** (2014), no. 1, 157–196, Translation of Tr. Mat. Inst. Steklova **285** (2014), 166–206.

[46] P. Li and S.-T. Yau, *On the parabolic kernel of the Schrödinger operator*, Acta Math. **156** (1986), no. 3-4, 153–201.

[47] Y. Meyer, *Wavelets and Operators*, Cambridge Studies in Advanced Mathematics, vol. 37, Cambridge University Press, Cambridge, 1992, Translated from the 1990 French original by D. H. Salinger.

[48] S. Minakshisundaram, *Eigenfunctions on Riemannian manifolds*, J. Indian Math. Soc. (N.S.) **17** (1953), 159–165.

[49] M. Murugan and L. Saloff-Coste, *Heat kernel estimates for anomalous heavy-tailed random walks*, Ann. Inst. Henri Poincaré Probab. Stat. **55** (2019), no. 2, 697–719.

[50] L. Saloff-Coste, *Aspects of Sobolev-type Inequalities*, London Mathematical Society Lecture Note Series, vol. 289, Cambridge University Press, Cambridge, 2002.

[51] ______, *The heat kernel and its estimates*, Probabilistic approach to geometry, Adv. Stud. Pure Math., vol. 57, Math. Soc. Japan, Tokyo, 2010, pp. 405–436.

[52] V. Shelkovich and M. Skopina, *p-adic Haar multiresolution analysis and pseudo-differential operators*, J. Fourier Anal. Appl. **15** (2009), no. 3, 366–393.

[53] G. Sweers, *When is the first eigenfunction for the clamped plate equation of fixed sign?*, Proceedings of the USA-Chile Workshop on Nonlinear Analysis (Viña del Mar-Valparaiso, 2000), 285–296, Electron. J. Differ. Equ. Conf., vol. 6, Southwest Texas State Univ., San Marcos, TX, 2001.

[54] W. Sweldens, *The lifting scheme: a construction of second generation wavelets*, SIAM J. Math. Anal. **29** (1998), no. 2, 511–546.

[55] H. Triebel, *Theory of Function Spaces. III*, Monographs in Mathematics, vol. 100, Birkhäuser Verlag, Basel, 2006.

[56] V. S. Vladimirov, I. V. Volovich, and E. I. Zelenov, *p-adic Analysis and Mathematical Physics*, Series on Soviet and East European Mathematics, 1. World Scientific Publishing Co., Inc., River Edge, NJ, 1994.

[57] K. Yosida, *Functional Analysis*, sixth ed., Grundlehren der Mathematischen Wissenschaften, vol. 123, Springer-Verlag, Berlin-New York, 1980.

Jun Cao
School of Mathematical Sciences, Zhejiang University of Technology, Hangzhou 310023, People's Republic of China
E-mail: caojun1860@zjut.edu.cn

Alexander Grigor'yan
Department of Mathematics, University of Bielefeld, 33501 Bielefeld, Germany
E-mail: grigor@math.uni-bielefeld.de

Liguang Liu
School of Mathematics, Renmin University of China, Beijing 100872, People's Republic of China
E-mail: liuliguang@ruc.edu.cn