The Annals of Probability
2012, Vol. 40, No. 3, 1212–1284
DOI: 10.1214/11-AOP645
© Institute of Mathematical Statistics, 2012

TWO-SIDED ESTIMATES OF HEAT KERNELS ON METRIC MEASURE SPACES

BY ALEXANDER GRIGOR'YAN[1] AND ANDRAS TELCS[2]

University of Bielefeld and Budapest University of Technology and Economics

We prove equivalent conditions for two-sided sub-Gaussian estimates of heat kernels on metric measure spaces.

CONTENTS

Received May 2010; revised January 2011.

[1]Supported in part by SFB 701 of the German Research Council.

[2]Supported in part by a visiting grant of SFB 701 of the German Research Council.

MSC2010 subject classifications. Primary 47D07; secondary 35K08, 28A80.

Key words and phrases. Heat kernel, heat semigroup, sub-Gaussian estimates, fractals.

1. Introduction.

1.1. *Historical background.* The notion of heat kernel has a long history. The oldest and the best-known heat kernel is the Gauss–Weierstrass function

$$p_t(x,y) = \frac{1}{(4\pi t)^{n/2}} \exp\left(-\frac{|x-y|^2}{4t}\right),$$

where $t > 0$ and $x, y \in \mathbb{R}^n$, which is a fundamental solution of the heat equation

$$\frac{\partial u}{\partial t} = \Delta u, \tag{1.1}$$

where Δ is the Laplace operator in $\mathbb{R}^n$. A more general parabolic equation $\frac{\partial u}{\partial t} = Lu$, where

$$L = \sum_{i,j=1}^{n} \frac{\partial}{\partial x_i}\left(a_{ij}(x)\frac{\partial}{\partial x_j}\right)$$

is a uniformly elliptic operator with measurable coefficients $a_{ij} = a_{ji}$, has also a positive fundamental solution $p_t(x,y)$, and the latter admits the Gaussian bounds

$$p_t(x,y) \asymp \frac{C}{t^{n/2}} \exp\left(-\frac{|x-y|^2}{ct}\right), \tag{1.2}$$

where the sign $\asymp$ means that both $\le$ and $\ge$ are true, but the positive constants c and C may be different for upper and lower bounds. The estimate (1.2) was proved by Aronson [1] using the parabolic Harnack inequality of Moser [54].

The next chapter in the history of heat kernels was opened in differential geometry. Consider the heat equation (1.1) on a Riemannian manifold M, where Δ is now the Laplace–Beltrami operator on M. The heat kernel $p_t(x,y)$ is defined as the minimal positive fundamental solution of (1.1), which always exists and is a smooth nonnegative function of t, x, y; cf. [13, 28, 57]. The question of estimating the heat kernel on Riemannian manifolds was addressed by many authors (see, e.g., [17, 28, 52, 59]). Apart from obvious analytic and geometric motivation, a strong interest to heat kernel estimates persists in stochastic analysis because the heat kernel coincides with the transition density of Brownian motion on M generated by the Laplace–Beltrami operator.

One of the most powerful estimates of heat kernels was proved by Li and Yau [51]: if M is a complete Riemannian manifold of nonnegative Ricci curvature, then

$$p_t(x,y) \asymp \frac{C}{V(x,\sqrt{t})} \exp\left(-\frac{d^2(x,y)}{ct}\right), \tag{1.3}$$

where $d(x,y)$ is the geodesic distance on M, and $V(x,r)$ is the Riemannian volume of the geodesic ball $B(x,r) = \{y \in M : d(x,y) < r\}$. Similar estimates were obtained by Gushchin with coauthors [38, 39] for certain unbounded domains in $\mathbb{R}^n$ with the Neumann boundary condition.

An interesting question is what minimal geometric assumptions imply (1.3). The upper bound in (1.3) is know to be equivalent to a certain *Faber–Krahn*-type inequality (see Section 3.3). The geometric background of the lower bound in (1.3) is more complicated and is closely related to the Harnack inequalities. In fact, the full estimate (1.3) is equivalent, on one hand, to the parabolic Harnack inequality of Moser (see [18]), and, on the other hand, to the conjunction of the volume doubling property and the Poincaré inequality (see [22, 55]). For a more detailed account of heat kernel bounds on manifolds we refer the reader to the books and surveys [13, 17, 25, 27, 28, 42, 57, 59].

New dimensions in the history of heat kernels were literally discovered in analysis on fractals. Fractals are typically subsets of $\mathbb{R}^n$ with certain self-similarity properties, like the Sierpinski gasket (SG) or the Sierpinski carpet (SC). One makes a fractal into a metric measure space by choosing appropriately a metric d (e.g., the extrinsic metric from the ambient $\mathbb{R}^n$) and a measure μ (usually the Hausdorff measure). The next crucial step is introduction of a strongly local regular *Dirichlet form* on a fractal, that is, an analogy of the Dirichlet integral $\int |\nabla f|^2$ on manifolds, which is equivalent to construction of Brownian motion on the fractal in question; cf. [20]. This step is highly nontrivial and its implementation depends on a particular class of fractals. On SG Brownian motion was constructed by Goldstein [21] and Kusuoka [49], on SC, by Barlow and Bass [3]. Kigami [43, 44] introduced a class of *post-critically finite* (p.c.f.) fractals, containing SG, and constructed the Dirichlet form on such a fractal as a scaled limit of the discrete Dirichlet forms on the graph approximations.

A strongly local regular Dirichlet form canonically leads to the notion of the heat semigroup and the heat kernel, where the latter can be defined either as the integral kernel of the heat semigroup or as the transition density of Brownian motion. Surprisingly enough, the Dirichlet forms on many families of fractals admit continuous heat kernels that satisfy the *sub-Gaussian* estimates

$$p_t(x,y) \asymp \frac{C}{t^{\alpha/\beta}} \exp\left(-c\left(\frac{d^\beta(x,y)}{t}\right)^{1/(\beta-1)}\right), \tag{1.4}$$

where $\alpha > 0$ and $\beta > 1$ are two parameters that come from the geometric properties of the underlying fractal. Estimate (1.4) was proved by Barlow and Perkins [12] on SG, by Kumagai [48] on nested fractals, by Fitzsimmons, Hambly and Kumagai [19] on affine nested fractals and by Barlow and Bass on SC [4] and on generalized Sierpinski carpets [5] (see also [2, 7, 44, 47, 50]). In fact, α is the Hausdorff dimension of the space, while β is a new quantity that is called the *walk dimension* and that can be characterized either in terms of the exit time of Brownian motion from balls or as the critical exponent of a family of Besov function spaces on the fractal (cf. [2, 26, 29, 32]).

1.2. *Description of the results*. The purpose of this paper is to find convenient equivalent conditions for sub-Gaussian estimates of the heat kernels on abstract metric measure spaces. Let (M,d) be a locally compact separable metric space, μ be a Radon measure on M with full support and $(\mathcal{E},\mathcal{F})$ be a strongly local regular Dirichlet form on M (see Section 2.1 for the details). We are interested in the conditions that ensure the existence of the heat kernel $p_t(x,y)$ as a measurable or continuous function of x, y, and the estimates of the following type

$$p_t(x,y) \asymp \frac{C}{V(x,\mathcal{R}(t))} \exp\left(-ct\Phi\left(c\frac{d(x,y)}{t}\right)\right), \tag{1.5}$$

where $V(x,r) = \mu(B(x,r))$ and $\mathcal{R}(t)$, $\Phi(s)$ are some nonnegative increasing functions on $[0,\infty)$. For example, (1.3) has the form (1.5) with $\mathcal{R}(t) = \sqrt{t}$ and $\Phi(s) = s^2$, while (1.4) has the form (1.5) with $\mathcal{R}(t) = t^{1/\beta}$ and $\Phi(s) = s^{\beta/(\beta-1)}$ [assuming that[3] $V(x,r) \simeq r^\alpha$, which, in fact, follows from (1.4)].

To describe the results of the paper, let us introduce some hypotheses. First, we assume that the metric space (M,d) is unbounded and that all metric balls are precompact[4] (although these assumptions are needed only for a part of the results). Next, define the following conditions:

• the volume doubling property (*VD*): there is a constant C such that

$$V(x,2r) \le CV(x,r) \tag{VD}$$

for all $x \in M$ and $r > 0$;

• the elliptic Harnack inequality (H): there is a constant C such that, for any nonnegative harmonic function u in any ball $B(x,r) \subset M$,

$$\operatorname*{esup}_{B(x,r/2)} u \le C \operatorname*{einf}_{B(x,r/2)} u, \tag{H}$$

where esup and einf are the essential supremum and infimum, respectively, (see Section 3.2 for more details);

• the estimate of the mean exit time (E_F),

$$\mathbb{E}_x \tau_{B(x,r)} \simeq F(r), \tag{E_F}$$

where $\tau_{B(x,r)}$ is the first exist time from ball $B(x,r)$ of the associated diffusion process, started at the center x, and $F(r)$ is a given function with a certain regularity (see Section 3.3 for more details). A typical example is $F(r) = r^\beta$ for some constant $\beta > 1$.

The conditions $(H) + (VD) + (E_F)$ are known to be true on p.c.f. fractals (see [40, 44]) as well as on generalized Sierpinski carpets (see [5, 8]) so that our results

[3]The sign $\simeq$ means that the ratio of both sides is bounded between two positive constants.

[4]The precompactness of balls implies that (M,d) is a complete metric spaces. The following partial converse is also true: if (M,d) is complete and the volume doubling property (*VD*) holds, then all balls are precompact. However, since we do not always assume (*VD*), we make an independent assumption of precompactness of the balls.

apply to such fractals. Another situation where $(H)+(VD)+(E_F)$ are satisfied is the setting of *resistance forms* introduced by Kigami [46]. A resistance form is a specific Dirichlet form that corresponds to a strongly recurrent Brownian motion. Kigami showed that, in the setting of resistance forms on self-similar sets, (VD) alone implies (H) and (E_F) with $F(r)=r^{\beta}$, for a suitable choice of a distance function. Examples with more general functions $F(r)$ appear in [11] and [58].

Let us emphasize in this connection that our results do not depend on the recurrence or transience hypotheses and apply to both cases, which partly explains the complexity of the proofs. A transient case occurs, for example, for some generalized Sierpinski carpets. Another point worth mentioning is that we do not assume specific properties of the metric d such as being geodesic; the latter is quite a common assumption in the fractal literature. This level of generality enables applications to resistance forms where the distance function is usually the resistance metric that is not geodesic.

Our first main result, which is stated in Theorem 5.15 and which, in fact, is a combination of Theorems 3.11, 4.2, 5.11, 5.14, says the following: if the hypotheses $(VD)+(H)+(E_F)$ are satisfied, then the heat kernel $p_t(x,y)$ exists, is Hölder continuous in x,y and satisfies the following upper estimate:

$$(UE) \qquad p_t(x,y) \le \frac{C}{V(x,\mathcal{R}(t))}\exp\left(-\frac{1}{2}t\Phi\left(c\frac{d(x,y)}{t}\right)\right),$$

where $\mathcal{R}=F^{-1}$ and

$$\Phi(s) := \sup_{r>0}\left\{\frac{s}{r}-\frac{1}{F(r)}\right\},$$

and the *near-diagonal* lower estimate

$$(NLE) \qquad p_t(x,y) \ge \frac{c}{V(x,\mathcal{R}(t))} \qquad \text{provided } d(x,y)\le \eta\mathcal{R}(t),$$

where $\eta>0$ is a small enough constant. Furthermore, assuming that (VD) holds a priori, we have the equivalence[5]

$$(1.6) \qquad (UE)+(NLE) \Leftrightarrow (H)+(E_F)$$

(Theorem 7.4).

[5]For comparison, let us observe that, under the same standing assumptions, it was proved in [10] that

$$(UE)+(NLE)\Leftrightarrow (PHI_F),$$

where (PHI_F) stands for the *parabolic* Harnack inequality for caloric functions. Hence, we see that the "difference" between (PHI_F) and (H) is the condition (E_F), that in particular provides a necessary space/time scaling for (PHI_F).

For example, if $F(r) = r^\beta$ for some $\beta > 1$, then $\mathcal{R}(t) = t^{1/\beta}$ and $\Phi(s) = \text{const}\, s^{\beta/(\beta-1)}$. Hence, (*UE*) and (*NLE*) become as follows:

$$p_t(x, y) \le \frac{C}{V(x, t^{1/\beta})} \exp\left(-c\left(\frac{d^\beta(x, y)}{t}\right)^{1/(\beta-1)}\right) \tag{1.7}$$

and

$$p_t(x, y) \ge \frac{c}{V(x, t^{1/\beta})} \qquad \text{provided } d(x, y) \le \eta t^{1/\beta}.$$

It is desirable to have a lower bound of $p_t(x, y)$ for all x, y that would match the upper bound (1.7). However, such a lower bound fails in general. The reason for that is the lack of *chaining properties* of the distance function, where by chaining properties we loosely mean a possibility to connect any two points $x, y \in M$ by a chain of balls of controllable radii so that the number of balls in this chain is also under control. More precisely, this property can be stated in terms of the modified distance $d_\varepsilon(x, y)$ where $\varepsilon > 0$ is a parameter. The exact definition of d_ε is given in Section 6.1, where it is also shown that

$$d_\varepsilon(x, y) \simeq \varepsilon N_\varepsilon(x, y),$$

where $N_\varepsilon(x, y)$ is the smallest number of balls in a chain of balls of radii ε connecting x and y. As ε goes to 0, $d_\varepsilon(x, y)$ increases and can go to ∞ or even become equal to ∞. If the distance function d is geodesic then $d_\varepsilon \equiv d$, which corresponds to the best possible chaining property. In general, the rate of growth of $d_\varepsilon(x, y)$ as $\varepsilon \to 0$ can be regarded as a quantitative description of the chaining properties of d. For this part of our work, we assume that

$$\frac{F(\varepsilon)}{\varepsilon} d_\varepsilon(x, y) \to 0 \qquad \text{as } \varepsilon \to 0, \tag{1.8}$$

which allows to define a function $\varepsilon(t, x, y)$ from the identity

$$\frac{F(\varepsilon)}{\varepsilon} d_\varepsilon(x, y) = t. \tag{1.9}$$

Our second main result states the following: if (1.8) and (*VD*) + (*H*) + (E_F) are satisfied, then

$$p_t(x, y) \asymp \frac{C}{V(x, \mathcal{R}(t))} \exp\left(-ct\Phi\left(c\frac{d_\varepsilon(x, y)}{t}\right)\right) \tag{1.10}$$

$$\asymp \frac{C}{V(x, \mathcal{R}(t))} \exp(-cN_\varepsilon), \tag{1.11}$$

where $\varepsilon = \varepsilon(ct, x, y)$ (Theorem 6.5). For example, the above hypotheses and, hence, the estimates (1.10) and (1.11) hold on connected p.c.f. fractals endowed with resistance distance, where one has $V(x, r) \simeq r^\alpha$ and $F(r) = r^{\alpha+1}$ for some constant α. The estimate (1.11) on p.c.f. fractals was first proved by Hambly and

Kumagai [40]. In fact, we use the argument from [40] to verify our hypotheses (see Remark 6.8).

Note that the dependence on t, x, y in the estimates (1.10) and (1.11) in very implicit and is hidden in $\varepsilon(ct, x, y)$. One can loosely interpret the use of this function in (1.10) and (1.11) as follows. In order to find a most probable path for Brownian motion to go from x to y in time t, one determines the optimal size $\varepsilon = \varepsilon(ct, x, y)$ of balls and then the optimal chain of balls of radii ε connecting x and y, and this chain provides an optimal route between x and y. This phenomenon was discovered by Hambly and Kumagai in the setting of p.c.f. fractals, where they used instead of balls the construction cells of the fractal. As it follows from our results, this phenomenon is generic and independent of self-similar structures.

If the distance function satisfies the *chain condition* $d_\varepsilon \le Cd$, which is stronger than (1.8), then one can replace in (1.10) d_ε by d and obtain (1.5) (Corollary 6.11). In fact, in this case we have the equivalence

$$(VD) + (H) + (E_F) \Leftrightarrow (1.5) \tag{1.12}$$

(Corollary 7.6).

In the setting of random walks on infinite graphs, the equivalence (1.12) was proved by the authors in [36, 37]. Of course, in this case all the conditions have to be adjusted to the discrete setting.

For the sake of applications (cf., e.g, [8]), it is desirable to replace the probabilistic condition (E_F) in all the above results by an analytic condition, namely, by a certain estimate of the capacity between two concentric balls. This type of result requires different techniques and will be treated elsewhere.

1.3. *Structure of the paper and interconnection of the results.* In Section 2 we revise the basic properties of the heat semigroups and heat kernels and prove the criterion for the existence of the heat kernel in terms of local ultracontractivity of the heat semigroup (Theorem 2.12).

In Section 3 we prove two preparatory results:

(1) $(VD) + (H) + (E_F) \Rightarrow (FK)$ where (FK) stands for a certain *Faber–Krahn inequality*, which provides a lower bound for the bottom eigenvalue in any bounded open set $\Omega \subset M$ via its measure (Theorem 3.11). In turn, (FK) implies the local ultracontractivity of the heat semigroup, which by Theorem 2.12 ensures the existence of the heat kernel.

(2) (E_F) implies the following estimate of the tail of the exit time from balls:

$$\mathbb{P}_x\big(\tau_{B(x,R)} \le t\big) \le C \exp\left(-t\Phi\left(c\frac{R}{t}\right)\right) \tag{1.13}$$

(Theorem 3.15).

In Section 4 we prove the upper estimate of the heat kernel, more precisely, the implication

$$(VD) + (FK) + (E_F) \Rightarrow (UE)$$

(Theorem 4.2). The main difficulty lies already in the proof of the diagonal upper bound

$$(DUE) \qquad p_t(x,x) \le \frac{C}{V(x,\mathcal{R}(t))}.$$

Using (*FK*), we obtain first some diagonal upper bound for the Dirichlet heat kernels in balls, and then use Kigami's iteration argument and (1.13) to pass to (*DUE*). The latter argument is borrowed from [31]. The full upper estimate (*UE*) follows from (*DUE*) and (1.13).

In Section 5 we prove the lower bounds of the heat kernel. The diagonal lower bound

$$(DLE) \qquad p_t(x,x) \ge \frac{C}{V(x,\mathcal{R}(t))}$$

follows directly from (1.13) (Lemma 5.13). To obtain the near diagonal lower estimate (*NLE*), one estimates from above the difference

$$|p_t(x,x) - p_t(x,y)|, \tag{1.14}$$

where y is close to x, which requires the following two ingredients:

(1) the oscillation inequalities that are consequences of the elliptic Harnack inequality (H) (Lemma 5.2 and Proposition 5.3);

(2) the upper estimate of the time derivative $\partial_t p_t(x,y)$ (Corollary 5.7).

Combining them with (*UE*), one obtains an upper bound for (1.14), which together with (*DLE*) yields (*NLE*) (Theorem 5.14).

The same method gives also the Hölder continuity of the heat kernel (Theorem 5.11).

In Section 6 we prove two-sided estimates (1.10) and (1.11) (Theorem 6.5). For the upper bound, we basically repeat the proof of (*UE*) by tracing the use of the distance function d and replacing it by d_ε. The lower bound for large $d(x,y)$ is obtained from (*NLE*) by a standard chaining argument using the semigroup property of the heat kernel and the chaining property of the distance function.

In Section 7 we prove the converse Theorem 7.4, which essentially consists of the equivalence (1.6).

NOTATION 1. We use the letters C, c, C', c' etc. to denote positive constant whose value is unimportant and can change at each occurrence. Note that the value of such constants in the conclusions depend on the values of the constants in the hypotheses (and, perhaps, on some other explicit parameters). In this sense, all our results are quantitative.

The relation $f \simeq g$ means that $C^{-1}g \le f \le Cg$ for some positive constant C and for a specified range of the arguments of functions f and g. The relation $f \asymp g$ means that both inequalities $f \le g$ and $f \ge g$ hold but possibly with different values of constants c, C that are involved in the expressions f and/or g.

2. Heat semigroups and heat kernels.

2.1. *Basic setup.* Throughout the paper, we assume that (M, d) is a locally compact separable metric space, and μ is a Radon measure on M with full support. As usual, denote by $L^q(M)$ where $q \in [1, +\infty]$ the Lebesgue function space with respect measure μ, and by $\|\cdot\|_q$ the norm in $L^q(M)$. The inner product in $L^2(M)$ is denoted by $(\cdot, \cdot)$. All functions on M are supposed to be real valued. Denote by $C_0(M)$ the space of all continuous functions on M with compact supports, equipped with the sup-norm.

Let $(\mathcal{E}, \mathcal{F})$ be Dirichlet form in $L^2(M)$. This means that $\mathcal{F}$ is a dense subspace of $L^2(M)$, and $\mathcal{E}(f, g)$ is a bilinear, nonnegative definite, closed[6] form defined for functions $f, g \in \mathcal{F}$, which satisfies, in addition, the Markovian property.[7] The Dirichlet form $(\mathcal{E}, \mathcal{F})$ is called *regular* if $\mathcal{F} \cap C_0(M)$ is dense both in $\mathcal{F}$ and in $C_0(M)$. The Dirichlet form is called *strongly local* if $\mathcal{E}(f, g) = 0$ for all functions $f, g \in \mathcal{F}$ such that g has a compact support and $f \equiv \text{const}$ in a neighborhood of $\operatorname{supp} g$. In this paper, we assume by default that $(\mathcal{E}, \mathcal{F})$ is a regular, strongly local Dirichlet form. A general theory of Dirichlet forms can be found in [20].

Let $\mathcal{L}$ be the generator of $(\mathcal{E}, \mathcal{F})$; that is, $\mathcal{L}$ is a self-adjoint nonnegative definite operator in $L^2(M)$ with the domain $\operatorname{dom}(\mathcal{L})$ that is a dense subset of $\mathcal{F}$ and such that, for all $f \in \operatorname{dom}(\mathcal{L})$ and $g \in \mathcal{F}$

$$\mathcal{E}(f, g) = (\mathcal{L}f, g).$$

The associated *heat semigroup*

$$P_t = e^{-t\mathcal{L}}, \qquad t \ge 0,$$

is a family of bounded self-adjoint operators in $L^2(M)$. The Markovian properties allow the extension of P_t to a bounded operator in $L^q(M)$, with the norm ≤ 1, for any $q \in [1, +\infty]$.

Denote by $\mathcal{B}(M)$ the class of all Borel functions on M, by $\mathcal{B}_b$ the class of bounded Borel functions, by $\mathcal{B}_+(M)$ the class of nonnegative Borel functions and by $\mathcal{B}L^q(M)$ the class of Borel functions that belong to $L^q(M)$.

---

[6]The form $(\mathcal{E}, \mathcal{F})$ is called closed if $\mathcal{F}$ is a Hilbert space with respect to the following inner product:

$$\mathcal{E}_1(f, g) = \mathcal{E}(f, g) + (f, g).$$

[7]The Markovian property (which could be also called the Beurling–Deny property) means that if $f \in \mathcal{F}$, then also the function $\hat{f} = f_+ \wedge 1$ belongs to $\mathcal{F}$ and $\mathcal{E}(\hat{f}, \hat{f}) \le \mathcal{E}(f, f)$.

By [20], Theorem 7.2.1, for any local Dirichlet form, there exists a diffusion process $\{\{X_t\}_{t\geq 0}, \{\mathbb{P}_x\}_{x\in M\setminus \mathcal{N}_0}\}$ with the initial point x outside some properly exceptional set[8] $\mathcal{N}_0 \subset M$, which is associated with the heat semigroup $\{P_t\}$ as follows: for any $f \in \mathcal{B}L^q(M)$, $1 \leq q \leq \infty$,

$$\mathbb{E}_x f(X_t) = P_t f(x) \qquad \text{for } \mu\text{-a.a. } x \in M. \tag{2.1}$$

Consider the family of operators $\{\mathcal{P}_t\}_{t\geq 0}$ defined by

$$\mathcal{P}_t f(x) := \mathbb{E}_x f(X_t), \qquad x \in M \setminus \mathcal{N}_0, \tag{2.2}$$

for all functions $f \in \mathcal{B}_b(M)$ (if X_t has a finite lifetime, then f is to be extended by 0 at the cemetery). The function $\mathcal{P}_t f(x)$ is a bounded Borel function on $M \setminus \mathcal{N}_0$. It is convenient to extend it to all $x \in M$ by setting

$$\mathcal{P}_t f(x) = 0, \qquad x \in \mathcal{N}_0, \tag{2.3}$$

so that $\mathcal{P}_t$ can be considered as an operator in $\mathcal{B}_b(M)$. Obviously, $\mathcal{P}_t f \geq 0$ if $f \geq 0$ and $\mathcal{P}_t 1 \leq 1$. Moreover, the family $\{\mathcal{P}_t\}_{t\geq 0}$ satisfies the semigroup identity

$$\mathcal{P}_t \mathcal{P}_s = \mathcal{P}_{t+s}.$$

Indeed, if $x \in M \setminus \mathcal{N}_0$, then we have by the Markov property, for any $f \in \mathcal{B}_b(M)$,

$$\mathcal{P}_{t+s} f(x) = \mathbb{E}_x(f(X_{t+s})) = \mathbb{E}_x(\mathbb{E}_{X_t}(f(X_s))) = \mathbb{E}_x(\mathcal{P}_s f(X_t)) = \mathcal{P}_t(\mathcal{P}_s f)(x),$$

where we have used that $X_t \in M \setminus \mathcal{N}_0$ with $\mathbb{P}_x$-probability 1. If $x \in \mathcal{N}_0$, then we have again

$$\mathcal{P}_{t+s} f(x) = \mathcal{P}_t(\mathcal{P}_s f)(x),$$

because the both sides are 0.

By considering increasing sequences of bounded functions, one extends the definition of $\mathcal{P}_t f$ to all $f \in \mathcal{B}_+(M)$ so that the defining identities (2.2) and (2.3) remain valid also for $f \in \mathcal{B}_+(M)$ [allowing value $+\infty$ for $\mathcal{P}_t f(x)$]. For a signed function $f \in \mathcal{B}(M)$, define $\mathcal{P}_t f$ by

$$\mathcal{P}_t f(x) = \mathcal{P}_t(f_+)(x) - \mathcal{P}_t(f_-)(x),$$

provided at least one of the functions $\mathcal{P}_t(f_+)$, $\mathcal{P}_t(f_-)$ is finite. Obviously, identities (2.2), (2.3) are satisfied for such functions as well.

If follows from the comparison of (2.1) and (2.2) that, for all $f \in \mathcal{B}L^q(M)$,

$$\mathcal{P}_t f(x) = P_t f(x) \qquad \text{for } \mu\text{-a.a. } x \in M.$$

---

[8] A set $\mathcal{N} \subset M$ is called properly exceptional if it is Borel, $\mu(\mathcal{N}) = 0$ and

$$\mathbb{P}_x(X_t \in \mathcal{N} \text{ for some } t \geq 0) = 0$$

for all $x \in M \setminus \mathcal{N}$ (see [20], page 134 and Theorem 4.1.1 on page 137).

It particular, $\mathcal{P}_t f$ is finite almost everywhere.

The set of the above assumptions will be referred to as the *basic hypotheses*, and they are assumed by default in all parts of this paper. Sometimes we need also the following property.

DEFINITION 2.1. The Dirichlet form $(\mathcal{E}, \mathcal{F})$ is called *conservative* (or *stochastically complete*) if $\mathcal{P}_t 1 \equiv 1$ for all $t > 0$.

EXAMPLE 2.2. Let M be a connected Riemannian manifold, d be the geodesic distance on M, μ be the Riemannian volume. Define the Sobolev space

$$W^1 = \{f \in L^2(M) : \nabla f \in L^2(M)\},$$

where ∇f is the Riemannian gradient of f understood in the weak sense. For all $f, g \in W^1$, one defines the energy form

$$\mathcal{E}(f, g) = \int_M (\nabla f, \nabla g)\, d\mu.$$

Let $\mathcal{F}$ be the closure of $C_0^\infty(M)$ in W^1. Then $(\mathcal{E}, \mathcal{F})$ is a regular strongly local Dirichlet form in $L^2(M)$.

2.2. *The heat kernel and the transition semigroup.*

DEFINITION 2.3. The *heat kernel* (or the *transition density*) of the transition semigroup $\{\mathcal{P}_t\}$ is a function $p_t(x, y)$ defined for all $t > 0$ and $x, y \in D := M \setminus \mathcal{N}$, where $\mathcal{N}$ is a properly exceptional set containing $\mathcal{N}_0$, and such that the following properties are satisfied:

(1) for any $t > 0$, the function $p_t(x, y)$ is measurable jointly in x, y;
(2) for all $f \in \mathcal{B}_+(M)$, $t > 0$ and $x \in D$,

$$\mathcal{P}_t f(x) = \int_D p_t(x, y) f(y)\, d\mu(y); \tag{2.4}$$

(3) for all $t > 0$ and $x, y \in D$,

$$p_t(x, y) = p_t(y, x); \tag{2.5}$$

(4) for all $t, s > 0$ and $x, y \in D$,

$$p_{t+s}(x, y) = \int_D p_t(x, z) p_s(z, y)\, d\mu(z). \tag{2.6}$$

The set D is called the domain of the heat kernel.

Let us extend $p_t(x, y)$ to all $x, y \in M$ by setting $p_t(x, y) = 0$ if x or y is outside D. Then (2.5) and (2.6) hold for all $x, y \in M$, and the domain of integration in (2.4) and (2.6) can be extended to M. The existence of the heat kernel allows us to extend the definition of $\mathcal{P}_t f$ to all measurable functions f by choosing a Borel measurable version of f and noticing that the integral (2.4) does not change if function f is changed on a set of measure 0.

It follows from (2.1) and (2.4) that, for any $f \in L^2(M)$,

$$P_t f(x) = \int_M p_t(x, y) f(y)\, d\mu(y) \tag{2.7}$$

for all $t > 0$ and μ-a.a. $x \in M$. A measurable function $p_t(x, y)$ that satisfies (2.7) is called the heat kernel of the semigroup P_t. It is well known that the heat kernel of P_t satisfies (2.5) and (2.6) although for *almost* all $x, y \in M$ (see [31], Section 3.3).

Hence, the relation between the heat kernels of $\mathcal{P}_t$ and P_t is as follows: the former is defined as a pointwise function of x, y, while the latter is defined almost everywhere, and the former is a pointwise realization of the latter, where the defining identities (2.4), (2.5), (2.7) must be satisfied pointwise. In this paper the heat kernel is understood exclusively in the sense of Definition 2.3.

The existence of the heat kernel is not obvious at all and will be given a special treatment. Those who are interested in the settings where the pointwise existence of the heat kernel is known otherwise, can skip the rest of this section and go to Section 3.

LEMMA 2.4. *Let p_t be the heat kernel of $\mathcal{P}_t$.*

(a) *The function $p_t(x, \cdot)$ belongs to $\mathcal{B}L^2(M)$ for all $t > 0$ and $x \in M$.*
(b) *For all $t > 0$, $x, y \in M$, we have $p_t(x, y) \geq 0$ and*

$$\int_M p_t(x, z)\, d\mu(z) \leq 1. \tag{2.8}$$

Consequently, $p_t(x, \cdot) \in \mathcal{B}L^1(M)$.

(c) *If q_t is another heat kernel, then $p_t = q_t$ in the common part of their domains.*

PROOF. (a) Set $f = p_{t/2}(x, \cdot)$ and observe that, by (2.5) and (2.6),

$$p_t(x, y) = \int_M p_{t/2}(x, \cdot) p_{t/2}(y, \cdot)\, d\mu = \mathcal{P}_{t/2} f(y) \tag{2.9}$$

for all $t > 0$ and $x, y \in D$. Since $\mathcal{P}_{t/2} f$ is a Borel function, we obtain that $p_t(x, \cdot)$ is Borel. The latter is true also if $x \in \mathcal{N}$ since in this case $p_t(x, \cdot) = 0$. Setting in (2.9) $x = y$, we obtain

$$\int_M p_{t/2}(x.\cdot)^2\, d\mu = p_t(x, x) < \infty, \tag{2.10}$$

whence $p_{t/2}(x, \cdot) \in L^2(M)$.

(b) By (2.2), (2.3) we have $\mathcal{P}_t f(x) \geq 0$ for all $t > 0$, $x \in M$ provided $f \geq 0$. Setting $f = [p_t(x, \cdot)]_-$, we obtain

$$0 \leq \mathcal{P}_t f(x) = \int_M p_t(x, \cdot)[p_t(x, \cdot)]_- \, d\mu = -\int_M [p_t(x, \cdot)]_-^2 \, d\mu,$$

whence it follows that $[p_t(x, \cdot)]_- = 0$ a.e., that is, $p_t(x, \cdot) \geq 0$ a.e. on M. It follows from (2.9) that, for all $x, y \in M$,

$$p_t(x, y) = \int_M p_{t/2}(x, \cdot) p_{t/2}(y, \cdot) \, d\mu \geq 0.$$

Inequality (2.8) is trivial if $x \in \mathcal{N}$, and if $x \in D$ then it follows from

$$\int_M p_t(x, \cdot) \, d\mu = \mathcal{P}_t 1(x) = \mathbb{E}_x 1 \leq 1.$$

(c) Let D be the intersection of the domains of p_t and q_t. For all $f \in \mathcal{B}_+(M)$ and $t > 0$, $x \in D$, we have

$$\int_D p_t(x, \cdot) f \, d\mu = \mathcal{P}_t f(x) = \int_D q_t(x, \cdot) f \, d\mu.$$

Applying this identity to function $f = p_t(y, \cdot)$ where $y \in D$, and using (2.9), we obtain

$$p_{2t}(x, y) = \int_D q_t(x, \cdot) p_t(y, \cdot) \, d\mu.$$

Similarly, we have

$$q_{2t}(x, y) = \int_D p_t(y, \cdot) q_t(x, \cdot) \, d\mu,$$

whence $p_{2t}(x, y) = q_{2t}(x, y)$. $\square$

Following [20], page 67, a sequence $\{F_n\}_{n=1}^{\infty}$ of subsets of M will be called a *regular nest* if:

(1) each F_n is closed;
(2) $F_n \subset F_{n+1}$ for all $n \geq 1$;
(3) $\operatorname{Cap}(M \setminus F_n) \to 0$ as $n \to \infty$ (see [20] for the definition of capacity);
(4) measure $\mu|_{F_n}$ has full support in F_n (in the induced topology of F_n).

DEFINITION 2.5. A set $\mathcal{N} \subset M$ is called *truly exceptional* if:

(1) $\mathcal{N}$ is properly exceptional;
(2) $\mathcal{N} \supset \mathcal{N}_0$;
(3) there is a regular nest $\{F_n\}$ in M such that $M \setminus \mathcal{N} = \bigcup_{n=1}^{\infty} F_n$ and that the function $\mathcal{P}_t f|_{F_n}$ is continuous for all $f \in \mathcal{B}L^1(M)$, $t > 0$, and $n \in \mathbb{N}$.

The conditions under which a truly exceptional set exists, will be discussed later on. Let us mention some important consequences of the existence of such a set.

LEMMA 2.6. *Let $\mathcal{N}$ be a truly exceptional set. If, for some $f \in \mathcal{B}L^1(M)$, $t > 0$, and for an upper semicontinuous function $\varphi : M \to (-\infty, +\infty]$, the inequality*

$$\mathcal{P}_t f(x) \le \varphi(x)$$

holds for μ-a.a. $x \in M$, then it is true for all $x \in M \setminus \mathcal{N}$. Similarly, if $\psi : M \to [-\infty, +\infty)$ is a lower semicontinuous function and

$$\mathcal{P}_t f(x) \ge \psi(x)$$

holds for μ-a.a. $x \in M$, then it is true for all $x \in M \setminus \mathcal{N}$.

PROOF. This proof is essentially the same as in [20], Theorem 2.1.2(ii). Assume that $\mathcal{P}_t f(x_0) > \varphi(x_0)$ for some $x_0 \in M \setminus \mathcal{N}$. By Definition 2.5, x_0 belongs to one of the sets F_n. Since $\mathcal{P}_t f|_{F_n}$ is continuous and, hence, $(\mathcal{P}_t f - \varphi)|_{F_n}$ is lower semicontinuous, the condition $(\mathcal{P}_t f - \varphi)(x_0) > 0$ implies that $(\mathcal{P}_t f - \varphi)(x) > 0$ for all x in some open neighborhood U of x_0 in F_n. Since measure μ has full support in F_n, we have $\mu(U) > 0$ so that $\mathcal{P}_t f(x) > \varphi(x)$ in a set of positive measure, that contradicts the hypothesis.

The second claim follows from the first one with $\varphi = -\psi$. $\square$

Denote by $\operatorname{esup}_A f$ the μ-essential supremum of a function f on a set $A \subset M$, and by $\operatorname{einf}_A f$—the μ-essential infimum.

COROLLARY 2.7. *Let $\mathcal{N}$ be a truly exceptional set. Then, for any $f \in \mathcal{B}L^1(M)$, $t > 0$, and an open set $X \subset M$,*

$$\operatorname*{esup}_X \mathcal{P}_t f = \sup_{X \setminus \mathcal{N}} \mathcal{P}_t f \quad \text{and} \quad \operatorname*{einf}_X \mathcal{P}_t f = \inf_{X \setminus \mathcal{N}} \mathcal{P}_t f. \tag{2.11}$$

PROOF. Function

$$\varphi(x) = \begin{cases} \operatorname*{esup}_X \mathcal{P}_t f, & x \in X, \\ +\infty, & x \notin X, \end{cases}$$

is upper semicontinuous. Since $\mathcal{P}_t f(x) \le \varphi(x)$ for μ-a.a. $x \in M$, we conclude by Lemma 2.6 that this inequality is true for all $x \in M \setminus \mathcal{N}$, whence

$$\sup_{X \setminus \mathcal{N}} \mathcal{P}_t f \le \operatorname*{esup}_X \mathcal{P}_t f.$$

The opposite inequality follows trivially from the definition of the essential supremum.

The second identity in (2.11) follows from the first one by changing f to $-f$. $\square$

Note that if $p_t(x,y)$ is the heat kernel with domain $D = M \setminus \mathcal{N}$, then we have by (2.6) that, for all $x, y \in D$, $0 < s < t$,

$$p_t(x,y) = \mathcal{P}_s f(x), \tag{2.12}$$

where $f = p_{t-s}(\cdot, y)$. Hence, if $\mathcal{N}$ is truly exceptional, then the claims of Lemma 2.6 and Corollary 2.7 apply to function $p_t(x,y)$ in place of $\mathcal{P}_t f(x)$, with any fixed $y \in D$.

LEMMA 2.8. *Let $p_t(x,y)$ be the heat kernel with the domain $D = M \setminus \mathcal{N}$ such that $\mathcal{N}$ is a truly exceptional set. Let $\varphi : D \times D \to [0, +\infty]$ be an upper semicontinuous function and $\psi : D \times D \to [0, +\infty)$ be a lower semicontinuous function. If, for some fixed $t > 0$, the following inequality*:

$$\psi(x,y) \le p_t(x,y) \le \varphi(x,y) \tag{2.13}$$

holds for $\mu \times \mu$-almost all $x, y \in D$, then (2.13) *holds for all $x, y \in D$.*

This lemma is a generalization of [9], Lemma 2.2, and the proof follows the argument in [9].

PROOF OF LEMMA 2.8. Consider the set

$$D' = \{y \in D \colon \text{(2.13) holds for } \mu\text{-a.a. } x \in D\}.$$

If $y \in D'$ then applying Lemma 2.6 to the function $p_t(\cdot, y)$, we obtain that (2.13) holds for all $x \in D$.

Now fix $x \in D$. Since by Fubini's theorem $\mu(D \setminus D') = 0$, (2.13) holds for μ-a.a. $y \in M$. Applying Lemma 2.6 to the function $p_t(x, \cdot)$, we conclude that (2.13) holds for all $y \in D$. $\square$

COROLLARY 2.9. *Under the hypotheses of Lemma 2.8, if X, Y are two open subsets of M then*

$$\operatorname*{esup}_{\substack{x \in X \\ y \in Y}} p_t(x,y) = \sup_{\substack{x \in X \setminus \mathcal{N} \\ y \in Y \setminus \mathcal{N}}} p_t(x,y) \tag{2.14}$$

and

$$\operatorname*{einf}_{\substack{x \in X \\ y \in Y}} p_t(x,y) = \inf_{\substack{x \in X \setminus \mathcal{N} \\ y \in Y \setminus \mathcal{N}}} p_t(x,y). \tag{2.15}$$

PROOF. This follows from Lemma 2.8 with functions

$$\varphi(x,y) = \begin{cases} \text{const}, & x \in X, y \in Y, \\ +\infty, & \text{otherwise}, \end{cases}$$

and

$$\psi(x,y)=\begin{cases}\text{const}, & x\in X, y\in Y,\\ 0, & \text{otherwise.}\end{cases}\qquad\square$$

In conclusion of this section, let us state a result that ensures the existence of the heat kernel outside a truly exceptional set.

THEOREM 2.10 ([6], Theorem 2.1). *Assume that there is a positive left-continuous function $\gamma(t)$ such that for all $f\in L^1\cap L^2(M)$ and $t>0$,*

$$\|P_t f\|_\infty \le \gamma(t)\|f\|_1. \tag{2.16}$$

Then the transition semigroup $\mathcal{P}_t$ possesses the heat kernel $p_t(x,y)$ with domain $D=M\setminus\mathcal{N}$ for some truly exceptional set $\mathcal{N}$, and $p_t(x,y)\le\gamma(t)$ for all $x,y\in D$ and $t>0$.

If the semigroup $\{P_t\}$ satisfies (2.16), then it is called *ultracontractive* (cf. [17]). It was proved in [6] that the ultracontractivity implies the existence of a function $p_t(x,y)$ that satisfies all the requirements of Definition 2.3 except for the joint measurability in x,y. Let us prove the latter so that $p_t(x,y)$ is indeed the heat kernel in our strict sense. Given that $p_t(x,y)$ satisfies conditions (2)–(4) of Definition 2.3, we see that the statement of Lemma 2.4 remains true because the proof of that lemma does not use the joint measurability. In particular, for any $t>0$, $x\in D$, the function $p_t(x,\cdot)$ is in $L^2(M)$. Also, the mapping $x\mapsto p_t(x,\cdot)$ is weakly measurable as a mapping from D to $L^2(M)$ because for any $f\in L^2(M)$, the function $x\mapsto(p_{t,x},f)=\mathcal{P}_t f\ (x)$ is measurable. Since $L^2(M)$ is separable, by Pettis's measurability theorem (see [60], Chapter V, Section 4) the mapping $x\mapsto p_t(x,\cdot)$ is strongly measurable in $L^2(M)$. It follows that the function

$$p_{2t}(x,y)=(p_t(x,\cdot),p_t(y,\cdot))$$

is jointly measurable in $x,y\in D$ as the composition of two strongly measurable mappings $D\to L^2(M)$ and a continuous mapping $f,g\mapsto(f,g)$.

2.3. *Restricted heat semigroup and local ultracontractivity.* Any open subset Ω of M can be considered as a metric measure space (Ω,d,μ). Let us identify $L^2(\Omega)$ as a subspace in $L^2(M)$ by extending functions outside Ω by 0. Define $\mathcal{F}(\Omega)$ as the closure of $\mathcal{F}\cap C_0(\Omega)$ in $\mathcal{F}$ so that $\mathcal{F}(\Omega)$ is a subspace of both $\mathcal{F}$ and $L^2(\Omega)$. Then $(\mathcal{E},\mathcal{F}(\Omega))$ is a regular strongly local Dirichlet form in $L^2(\Omega)$, which is called the restriction of $(\mathcal{E},\mathcal{F})$ to Ω. Let $\mathcal{L}^\Omega$ be the generator of the form $(\mathcal{E},\mathcal{F}(\Omega))$ and $P_t^\Omega=e^{-t\mathcal{L}^\Omega}$, $t\ge0$, be the *restricted* heat semigroup.

Define the *first exit time* from Ω by

$$\tau_\Omega=\inf\{t>0: X_t\notin\Omega\}.$$

The diffusion process associated with the restricted Dirichlet form, can be canonically obtained from $\{X_t\}$ by killing the latter outside Ω, that is, by restricting the life time of X_t by τ_Ω (see [20]). It follows that the transition operator $\mathcal{P}_t^\Omega$ of the killed diffusion is given by

$$\mathcal{P}_t^\Omega f(x) = \mathbb{E}_x\big(\mathbf{1}_{\{t<\tau_\Omega\}} f(X_t)\big) \qquad \text{for all } x \in \Omega \setminus \mathcal{N}_0, \tag{2.17}$$

for all $f \in \mathcal{B}_+(\Omega)$. Then $\mathcal{P}_t^\Omega f$ is defined for f from other function classes in the same way as $\mathcal{P}_t$. Also, extend $\mathcal{P}_t^\Omega f(x)$ to all $x \in \Omega$ by setting it to be 0 if $x \in \mathcal{N}_0$.

DEFINITION 2.11. We say that the semigroup P_t is *locally ultracontractive* if the restricted heat semigroup P_t^B is ultracontractive for any metric ball B of (M, d).

THEOREM 2.12. *Let the semigroup P_t be locally ultracontractive. Then the following is true.*

(a) *There exists a properly exceptional set $\mathcal{N} \subset M$ such that, for any open subset $\Omega \subset M$, the semigroup $\mathcal{P}_t^\Omega$ possesses the heat kernel $p_t^\Omega(x, y)$ with the domain $\Omega \setminus \mathcal{N}$.*

(b) *If $\Omega_1 \subset \Omega_2$ are open subsets of M, then $p_t^{\Omega_1}(x, y) \le p_t^{\Omega_2}(x, y)$ for all $t > 0$, $x, y \in \Omega_1 \setminus \mathcal{N}$.*

(c) *If $\{\Omega_k\}_{k=1}^\infty$ is an increasing sequence of open subsets of M and $\Omega = \bigcup_k \Omega_k$, then $p_t^{\Omega_k}(x, y) \to p_t^\Omega(x, y)$ as $k \to \infty$ for all $t > 0$, $x, y \in \Omega \setminus \mathcal{N}$.*

(d) *Set $D = M \setminus \mathcal{N}$. Let $\varphi(x, y): D \times D \to [0, +\infty]$ be an upper semicontinuous function such that, for some open set $\Omega \subset M$ and for some $t > 0$,*

$$p_t^\Omega(x, y) \le \varphi(x, y) \tag{2.18}$$

for almost all $x, y \in \Omega$. Then (2.18) *holds for all $x, y \in \Omega \setminus \mathcal{N}$.*

For simplicity of notation, set $p_t^\Omega(x, y)$ to be 0 for all x, y outside Ω (which, however, does not mean the extension of the domain of p_t^Ω).

PROOF OF THEOREM 2.12. (a) Since the metric space (M, d) is separable, there is a countable family of balls that form a base. Let $\mathcal{U}$ be the family of all finite unions of such balls so that $\mathcal{U}$ is countable and any open set $\Omega \subset M$ can be represented as an increasing union of sets of $\mathcal{U}$. Since any set $U \in \mathcal{U}$ is contained in a metric ball, the semigroup P_t^U is dominated by P_t^B and, hence, is ultracontractive. By Theorem 2.10, there is a truly exceptional set $\mathcal{N}_U \subset U$ such that the $\mathcal{P}_t^U$ has the heat kernel p_t^U in the domain $U \setminus \mathcal{N}_U$. Since the family $\mathcal{U}$ is countable, the set

$$\mathcal{N} = \bigcup_{U \in \mathcal{U}} \mathcal{N}_U \tag{2.19}$$

is properly exceptional.

Let us first show that if U_1, U_2 are the sets from $\mathcal{U}$ and $U_1 \subset U_2$, then

$$p_t^{U_1}(x,y) \le p_t^{U_2}(x,y) \qquad \text{for all } t>0, x,y \in U_1 \setminus \mathcal{N}. \tag{2.20}$$

It follows from (2.17) that, for any $f \in \mathcal{B}_+(U_1)$,

$$\mathcal{P}_t^{U_1} f(x) \le \mathcal{P}_t^{U_2} f(x) \qquad \text{for all } t>0 \text{ and } x \in U_1,$$

that is,

$$\int_{U_1} p_t^{U_1}(x,\cdot) f \, d\mu \le \int_{U_2} p_t^{U_2}(x,\cdot) f \, d\mu. \tag{2.21}$$

Setting here $f = P_t^{U_1}(y,\cdot)$ where $y \in U_1 \setminus \mathcal{N}$, we obtain

$$p_{2t}^{U_1}(x,y) \le \int_{U_1} p_t^{U_2}(x,\cdot) p_t^{U_1}(y,\cdot)\, d\mu.$$

Setting in (2.21) $f = P_t^{U_2}(y,\cdot)$, we obtain

$$\int_{U_1} p_t^{U_1}(x,\cdot) P_t^{U_2}(y,\cdot)\, d\mu \le p_{2t}^{U_2}(x,y).$$

Combining the above two lines gives (2.20).

Let Ω be any open subset of M and $\{U_n\}_{n=1}^{\infty}$ be an increasing sequence of sets from $\mathcal{U}$ such that $\Omega = \bigcup_{n=1}^{\infty} U_n$. Let us set

$$p_t^{\Omega}(x,y) = \lim_{n\to\infty} p_t^{U_n}(x,y) \qquad \text{for all } t>0 \text{ and } x,y \in \Omega \setminus \mathcal{N}. \tag{2.22}$$

This limit exists (finite or infinite) by the monotonicity of the sequence $\{p_t^{U_n}(x,y)\}$. It follows from (2.17) that, for any $f \in \mathcal{B}_+(\Omega)$,

$$\mathcal{P}_t^{U_n} f(x) \uparrow \mathcal{P}_t^{\Omega} f(x) \qquad \text{for all } t>0 \text{ and } x \in \Omega \setminus \mathcal{N}.$$

By the monotone convergence theorem, we obtain

$$\mathcal{P}_t^{U_n} f(x) = \int_{\Omega} p_t^{U_n}(x,y) f(y)\, d\mu(y) \to \int_{\Omega} p_t^{\Omega}(x,y) f(y)\, d\mu(y)$$

for all $t>0$ and $x \in \Omega \setminus \mathcal{N}$. Comparing the above two lines, we obtain

$$\mathcal{P}_t^{\Omega} f(x) = \int_{\Omega} p_t^{\Omega}(x,y) f(y)\, d\mu(y) \qquad \text{for all } t>0 \text{ and } x \in \Omega \setminus \mathcal{N}.$$

The symmetry of $p_t^{\Omega}(x,y)$ is obvious from (2.22), and the semigroup property of p_t^{Ω} follows from that of $p_t^{U_n}$ by the monotone convergence theorem. Note that p_t^{Ω} does not depend on the choice of $\{U_n\}$ by the uniqueness of the heat kernel (Lemma 2.4).

(b) For two arbitrary open sets $\Omega_1 \subset \Omega_2$ let $\{U_n\}_{n=1}^{\infty}$ and $\{W_n\}_{n=1}^{\infty}$ be increasing sequences of sets from $\mathcal{U}$ that exhaust Ω_1 and Ω_2, respectively. Set $V_n = U_n \cup W_n$

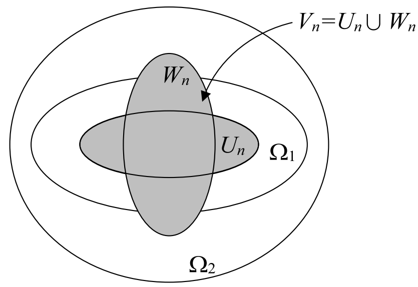

FIG. 1. *Sets* U_n, W_n, V_n.

so that $V_n \in \mathcal{U}$ and Ω_2 is the increasing union of sets V_n (see Figure 1). Then $U_n \subset V_n$ and, hence, $p_t^{U_n} \leq p_t^{V_n}$, which implies as $n \to \infty$ that $p_t^{\Omega_1} \leq p_t^{\Omega_2}$.

(c) Let $\{\Omega_k\}_{k=1}^{\infty}$ be an increasing sequence of open sets whose union is Ω. Let $\{U_n^{(k)}\}_{n=1}^{\infty}$ be an increasing sequence of sets from $\mathcal{U}$ that exhausts Ω_k. As in the previous argument, we can replace $U_n^{(2)}$ by $V_n^{(2)} = U_n^{(1)} \cup U_n^{(2)}$ so that $U_n^{(1)} \subset V_n^{(2)}$. Rename $V_n^{(2)}$ back to $U_n^{(2)}$, and assume in the sequel that $U_n^{(1)} \subset U_n^{(2)}$. Similarly, replace $U_n^{(3)}$ by $U_n^{(1)} \cup U_n^{(2)} \cup U_n^{(3)}$ and assume in the sequel that $U_n^{(2)} \subset U_n^{(3)}$. Arguing by induction, we redefine the double sequence $U_n^{(k)}$ in the way that it is monotone increasing not only in n but also in k. Then we claim that

$$\Omega = \bigcup_{m=1}^{\infty} U_m^{(m)}.$$

Indeed, if $x \in \Omega$, then $x \in \Omega_k$ for some k and, hence, $x \in U_n^{(k)}$ for some n, which implies $x \in U_m^{(m)}$ for $m = \max(k, n)$. Finally, we have $p_t^{\Omega} \geq p_t^{\Omega_m}$ and

$$p_t^{\Omega} = \lim_{m \to \infty} p_t^{U_m^{(m)}} \leq \lim_{m \to \infty} p^{\Omega_m},$$

whence it follows that

$$p_t^{\Omega} = \lim_{m \to \infty} p^{\Omega_m}.$$

(d) Let $U \in \mathcal{U}$ be subset of Ω. Then the semigroup P_t^U is ultracontractive and possesses the heat kernel p_t^U with the domain $U \setminus \mathcal{N}_U$ where $\mathcal{N}_U$ is a truly exceptional set as in part (a). Note that $\mathcal{N}_U \subset \mathcal{N}$. Since $p_t^U \leq p_t^{\Omega}$ in $U \setminus \mathcal{N}$, we obtain by hypothesis that

$$p_t^U(x, y) \leq \varphi(x, y)$$

for almost all $x, y \in U$. By Lemma 2.8, we conclude that this inequality is true for all $x, y \in U \setminus \mathcal{N}$. Exhausting Ω be a sequence of subsets $U \in \mathcal{U}$ and using (2.22), we obtain (2.18). $\square$

3. Some preparatory results.

3.1. *Green operator.* A priori we assume here only the basic hypotheses. All necessary additional assumptions are explicitly stated. The main result of this section is Theorem 3.11.

Given an open set $\Omega \subset M$, define the Green operator G^Ω first for all $f \in \mathcal{B}_+(\Omega)$ by

$$G^\Omega f(x) = \int_0^\infty \mathcal{P}_t^\Omega f(x)\,dt \tag{3.1}$$

for all $x \in M \setminus \mathcal{N}_0$, where we admit infinite values of the integral. If $f \in \mathcal{B}(\Omega)$ and $G^\Omega |f| < \infty$, then $G^\Omega f$ is also defined by

$$G^\Omega f = G^\Omega f_+ - G^\Omega f_-.$$

LEMMA 3.1. *We have, for any open $\Omega \subset M$ and all $f \in \mathcal{B}_+(\Omega)$,*

$$G^\Omega f(x) = \mathbb{E}_x\left(\int_0^{\tau_\Omega} f(X_t)\,dt\right) \tag{3.2}$$

for any $x \in \Omega \setminus \mathcal{N}_0$. In particular,

$$G^\Omega 1(x) = \mathbb{E}_x \tau_\Omega. \tag{3.3}$$

PROOF. Indeed, integrating (2.17) in t, we obtain

$$\begin{aligned} G^\Omega f(x) &= \int_0^\infty \mathcal{P}_t^\Omega f(x)\,dt \\ &= \int_0^\infty \mathbb{E}_x\big(\mathbf{1}_{\{t<\tau_\Omega\}} f(X_t)\big)\,dt \\ &= \mathbb{E}_x \int_0^\infty \big(\mathbf{1}_{\{t<\tau_\Omega\}} f(X_t)\big)\,dt \\ &= \mathbb{E}_x\left(\int_0^{\tau_\Omega} f(X_t)\,dt\right). \end{aligned}$$

Obviously, (3.3) follows from (3.2) for $f \equiv 1$. □

Denote by $\lambda_{\min}(\Omega)$ the bottom of the spectrum of $\mathcal{L}^\Omega$ in $L^2(\Omega)$, that is,

$$\lambda_{\min}(\Omega) := \inf \operatorname{spec} \mathcal{L}^\Omega = \inf_{f \in \mathcal{F}(\Omega)\setminus\{0\}} \frac{\mathcal{E}(f,f)}{(f,f)}. \tag{3.4}$$

For any open set $\Omega \subset M$, we will consider the *mean exit time* $\mathbb{E}_x \tau_\Omega$ from Ω as a function of $x \in \Omega \setminus \mathcal{N}_0$. Also, set

$$\widetilde{E}(\Omega) := \operatorname*{esup}_{x\in\Omega} \mathbb{E}_x \tau_\Omega. \tag{3.5}$$

LEMMA 3.2. *If $\widetilde{E}(\Omega) < \infty$, then G^{Ω} is a bounded operator on $\mathcal{B}_b(\Omega)$, and it uniquely extends to each of the spaces $L^{\infty}(\Omega)$, $L^1(\Omega)$, $L^2(\Omega)$, with the following norm estimates*:

$$\|G^{\Omega}\|_{L^{\infty}\to L^{\infty}} \le \widetilde{E}(\Omega), \tag{3.6}$$

$$\|G^{\Omega}\|_{L^{1}\to L^{1}} \le \widetilde{E}(\Omega), \tag{3.7}$$

$$\|G^{\Omega}\|_{L^{2}\to L^{2}} \le \widetilde{E}(\Omega). \tag{3.8}$$

Moreover,

$$\lambda_{\min}(\Omega)^{-1} \le \widetilde{E}(\Omega), \tag{3.9}$$

and G^{Ω} is the inverse in $L^2(\Omega)$ to the operator $\mathcal{L}^{\Omega}$.

PROOF. It follows from (3.3) that

$$\|G^{\Omega}1\|_{\infty} = \widetilde{E}(\Omega), \tag{3.10}$$

which implies that for any $f \in \mathcal{B}_b(\Omega)$,

$$\|G^{\Omega} f\|_{\infty} \le \widetilde{E}(\Omega)\|f\|_{\infty}.$$

Hence, G^{Ω} can be considered as a bounded operator in L^{∞} with the norm estimate (3.6).

Estimate (3.7) follows from (3.6) by duality. Indeed, for any two functions $f, h \in \mathcal{B}_+(\Omega)$, we have

$$\int_{\Omega} (G^{\Omega} f) h \, d\mu = \int_{\Omega} f G^{\Omega} h \, d\mu, \tag{3.11}$$

which follows from (3.1) and the symmetry of $\mathcal{P}_t^{\Omega}$. By linearity, (3.11) extends to all $f, h \in \mathcal{B}_b(\Omega)$. Then, for any $f \in C_0(\Omega)$, we have

$$\begin{aligned}
\|G^{\Omega} f\|_1 &= \sup_{h \in \mathcal{B}_b(\Omega)\setminus\{0\}} \frac{\int_{\Omega} (G^{\Omega} f) h \, d\mu}{\|h\|_{\infty}} \\
&= \sup_{h \in \mathcal{B}_b(\Omega)\setminus\{0\}} \frac{\int_{\Omega} f G^{\Omega} h \, d\mu}{\|h\|_{\infty}} \\
&\le \sup_{h \in \mathcal{B}_b(\Omega)\setminus\{0\}} \frac{\|G^{\Omega} h\|_{\infty} \|f\|_1}{\|h\|_{\infty}} \\
&\le \widetilde{E}(\Omega)\|f\|_1,
\end{aligned}$$

whence it follows that G^{Ω} uniquely extends to a bounded operator in L^1 with the norm estimate (3.7).

The estimate (3.8) [as well as a similar estimate for $\|G\|_{L^p \to L^p}$ for any $p \in (1, \infty)$] follows from (3.6) and (3.7) by the Riesz–Thorin interpolation theorem.

To prove (3.9), let us consider the following "cut-down" version of the Green operator:

$$G_T^\Omega f = \int_0^T \mathcal{P}_t^\Omega f\, dt,$$

where $T \in (0, +\infty)$. The same argument as above shows that G_T^Ω can be considered as an operator in L^2 with the same norm bound

$$\|G_T^\Omega\|_{L^2\to L^2} \le \widetilde{E}(\Omega).$$

On the other hand, using the spectral resolution $\{E_\lambda\}_{\lambda\ge 0}$ of the generator $\mathcal{L}^\Omega$, we obtain, for any $f \in C_0(\Omega)$,

$$\begin{aligned} G_T^\Omega f &= \int_0^T \left(\int_0^\infty e^{-\lambda t}\, dE_\lambda f\right) dt \\ &= \int_0^\infty \left(\int_0^T e^{-\lambda t}\, dt\right) dE_\lambda f \\ &= \int_0^\infty \varphi_T(\lambda)\, dE_\lambda f \\ &= \varphi_T(\mathcal{L}^\Omega) f, \end{aligned} \tag{3.12}$$

where

$$\varphi_T(\lambda) = \int_0^T e^{-\lambda t}\, dt = \frac{1-e^{-T\lambda}}{\lambda}.$$

Since φ_T is a bounded function on $[0, +\infty)$, the operator $\varphi_T(\mathcal{L}^\Omega)$ is a bounded operator in L^2. By the spectral mapping theorem, we obtain

$$\begin{aligned} \sup \varphi_T(\operatorname{spec} \mathcal{L}^\Omega) &= \sup \operatorname{spec} \varphi_T(\mathcal{L}^\Omega) \\ &= \|\varphi_T(\mathcal{L}^\Omega)\|_{L^2\to L^2} \\ &= \|G_T^\Omega\|_{L^2\to L^2} \\ &\le \widetilde{E}(\Omega). \end{aligned}$$

On the other hand, since $\varphi_T(\lambda)$ is decreasing in λ,

$$\sup \varphi_T(\operatorname{spec} \mathcal{L}^\Omega) = \varphi_T(\lambda_{\min}(\Omega)),$$

whence

$$\varphi_T(\lambda_{\min}(\Omega)) \le \widetilde{E}(\Omega).$$

By letting $T \to \infty$ and observing that $\varphi_T(\lambda) \to \frac{1}{\lambda}$, we obtain

$$\lambda_{\min}(\Omega)^{-1} \le \widetilde{E}(\Omega),$$

which in particular means that $\lambda_{\min}(\Omega) > 0$. Consequently, the operator $\mathcal{L}^\Omega$ has a bounded inverse. Passing in (3.12) to the limit as $T \to \infty$, we obtain $G^\Omega = (\mathcal{L}^\Omega)^{-1}$. $\square$

3.2. *Harmonic functions and Harnack inequality.* Let Ω be an open subset of M.

DEFINITION 3.3. We say that a function $u \in \mathcal{F}$ is *harmonic* in Ω if

$$\mathcal{E}(u, v) = 0 \qquad \text{for any } v \in \mathcal{F}(\Omega).$$

LEMMA 3.4. *Let Ω be an open subset of M such that $\widetilde{E}(\Omega) < \infty$, and let U be an open subset of Ω.*

(a) *For any $f \in L^2(\Omega \setminus U)$, the function $G^\Omega f$ is harmonic in U.*
(b) *For any $f \in L^2(\Omega)$, the function $G^\Omega f - G^U f$ is harmonic in U.*

REMARK 3.5. If $f \in L^2(\Omega)$, then $G^U f$ is defined as the extension of $G^U(f|_U)$ to Ω by setting it to be equal to 0 in $\Omega \setminus U$.

PROOF OF LEMMA 3.4. (a) Set $u = G^\Omega f$. To prove that u is harmonic in U, we need to show that $\mathcal{E}(u, v) = 0$, for any $v \in \mathcal{F}(U)$. Since by Lemma 3.2 $G^\Omega = (\mathcal{L}^\Omega)^{-1}$, we have $u \in \operatorname{dom}(\mathcal{L}^\Omega)$. Therefore, by the definition of $\mathcal{L}^\Omega$,

$$\mathcal{E}(u, v) = (\mathcal{L}^\Omega u, v) = (f, v) = 0.$$

(b) Set $u = G^\Omega f - G^U f$. Any function $v \in \mathcal{F}(U)$ can be considered as an element of $\mathcal{F}(\Omega)$ by setting it to be 0 in $\Omega \setminus U$. Then both u and v are in $\mathcal{F}(\Omega)$ whence

$$\begin{aligned}\mathcal{E}(u, v) &= \mathcal{E}(G^\Omega f, v) - \mathcal{E}(G^U f, v)\\ &= (f, v)_{L^2(\Omega)} - (f, v)_{L^2(U)}\\ &= 0.\end{aligned}$$

□

Denote by

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

the open metric ball of radius $r > 0$ centered at a point $x \in M$, and set

$$V(x, r) = \mu(B(x, r)).$$

That μ has full support implies $V(x, r) > 0$ whenever $r > 0$. Whenever we use the function $V(x, r)$, we always assume that

$$V(x, r) < \infty \qquad \text{for all } x \in M \text{ and } r > 0.$$

For example, this condition is automatically satisfied if all balls are precompact. However, we do not assume precompactness of all balls unless otherwise explicitly stated.

DEFINITION 3.6. We say that the *elliptic Harnack inequality* (H) holds on M, if there exist constants $C > 1$ and $\delta \in (0, 1)$ such that, for any ball $B(x, r)$ in M and for any function $u \in \mathcal{F}$ that is nonnegative and harmonic $B(x, r)$,

$$(H) \qquad \operatorname*{esup}_{B(x,\delta r)} u \le C \operatorname*{einf}_{B(x,\delta r)} u.$$

DEFINITION 3.7. We say that the *volume doubling* property (*VD*) holds if there exists a constant C such that, for all $x \in M$ and $r > 0$

$$(VD) \qquad V(x, 2r) \le C V(x, r).$$

It is known that (*VD*) implies that, for all $x, y \in M$ and $0 < r < R$,

$$\frac{V(x, R)}{V(y, r)} \le C\left(\frac{R + d(x, y)}{r}\right)^{\alpha} \tag{3.13}$$

for some $\alpha > 0$ (see [31]).

LEMMA 3.8. *Assume that* (*VD*) + (H) *hold. Let* Ω *be an open subset of* M *such that* $\widetilde{E}(\Omega) < \infty$, *and let* $B = B(x, r)$ *be a ball contained in* Ω.

(a) *For any nonnegative function* $\varphi \in L^1(\Omega \setminus B)$,

$$\operatorname*{esup}_{B(x,\delta r)} G^{\Omega}\varphi \le C \frac{\widetilde{E}(\Omega)}{V(x, r)} \|\varphi\|_1. \tag{3.14}$$

(b) *For and any nonnegative function* $\varphi \in L^1(\Omega)$,

$$\operatorname*{esup}_{B(x,\delta r)} (G^{\Omega}\varphi - G^{B}\varphi) \le \frac{C\widetilde{E}(\Omega)}{V(x, r)} \|\varphi\|_1. \tag{3.15}$$

PROOF. (a) Since identity (3.14) survives monotone increasing limits of sequences of functions φ, it suffices to prove (3.14) for any nonnegative function $\varphi \in L^1 \cap L^2(\Omega \setminus B)$. Then, by Lemma 3.4, the function $u = G^{\Omega}\varphi$ is harmonic in $B(x, r)$. Since $u \ge 0$, we can use the Harnack inequality (H) in ball B, which yields

$$\begin{aligned} \operatorname*{esup}_{B(x,\delta r)} u(x) &\le C \operatorname*{einf}_{B(x,\delta r)} u \le \frac{C}{V(x, r)} \|u\|_1 \\ &\le \frac{C}{V(x, r)} \|G^{\Omega}\|_{L^1 \to L^1} \|\varphi\|_1 \\ &\le \frac{C\widetilde{E}(\Omega)}{V(x, r)} \|\varphi\|_1. \end{aligned} \tag{3.16}$$

(b) Assume first that $\varphi \in L^1 \cap L^2(\Omega)$. By Lemma 3.4, the function $u = G^{\Omega}\varphi - G^B\varphi$ is harmonic in $B(x,r)$. Since $u \geq 0$, applying for this function argument (3.16), we obtain (3.15). An arbitrary nonnegative function $\varphi \in L^1(\Omega)$ can be represented as a sum in $L^1(\Omega)$

$$\varphi = \sum_{k=0}^{\infty} \varphi_k,$$

where $\varphi_k := (\varphi - k)_+ \wedge 1 \in L^1 \cap L^\infty(\Omega)$. Applying (3.15) to each φ_k and summing up, we obtain (3.15) for φ. $\square$

3.3. *Faber–Krahn inequality and mean exit time*. A classical theorem of Faber and Krahn says that for any bounded open set $\Omega \subset \mathbb{R}^n$,

$$\lambda_{\min}(\Omega) \geq \lambda_{\min}(B),$$

where B is a ball in $\mathbb{R}^n$ of the same volume as Ω. If the radius of B is r, then $\lambda_{\min}(B) = \frac{c}{r^2}$ where c is a positive constant depending only on n, which implies that

$$\lambda_{\min}(\Omega) \geq c\mu(\Omega)^{-2/n}; \tag{3.17}$$

cf. [13, 14]. We refer to lower estimates of $\lambda_{\min}(\Omega)$ via a function of $\mu(\Omega)$ as *Faber–Krahn inequalities*. A more general type of a Faber–Krahn inequality holds on a complete n-dimensional Riemannian manifold M of nonnegative Ricci curvature: for any bounded open set $\Omega \subset M$ and for any ball B of radius r containing Ω,

$$\lambda_{\min}(\Omega) \geq \frac{c}{r^2}\left(\frac{\mu(B)}{\mu(\Omega)}\right)^{\nu}, \tag{3.18}$$

where $\nu = 2/n$ and $c = c(n) > 0$ (see [22]). Note that (3.17) follows from (3.18) (apart from the sharp value of the constant c) because in $\mathbb{R}^n$ we have $\mu(B) = \operatorname{const} r^n$.

It was proved in [23] that, on any complete Riemannian manifold,

$$(3.18) \Leftrightarrow (VD) + (UE),$$

where (UE) is here the upper bound of the heat kernel in the Li–Yau estimate (1.3). In Section 4 we will derive a general upper bound (UE) from a set of hypotheses containing a suitable version of (3.18). In this section, we will deduce a Faber–Krahn inequality from the main hypotheses.

We fix from now on a function $F : (0,\infty) \to (0,\infty)$ that is a continuous increasing bijection of $(0,\infty)$ onto itself, such that, for all $0 < r \leq R$,

$$C^{-1}\left(\frac{R}{r}\right)^{\beta} \leq \frac{F(R)}{F(r)} \leq C\left(\frac{R}{r}\right)^{\beta'} \tag{3.19}$$

for some constants $1 < \beta \le \beta'$, $C > 1$. In the sequel we will use the inverse function $\mathcal{R} = F^{-1}$. It follows from (3.19) that

$$C^{-1}\left(\frac{T}{t}\right)^{1/\beta'} \le \frac{\mathcal{R}(T)}{\mathcal{R}(t)} \le C\left(\frac{T}{t}\right)^{1/\beta} \tag{3.20}$$

for all $0 < t \le T$.

DEFINITION 3.9. We say that the Faber–Krahn inequality (FK) holds if, for any ball B in M of radius r and any open set $\Omega \subset B$,

$$\lambda_{\min}(\Omega) \ge \frac{c}{F(r)}\left(\frac{\mu(B)}{\mu(\Omega)}\right)^{\nu} \tag{FK}$$

with some positive constants c, ν.

DEFINITION 3.10. We say that the mean exit time estimate (E_F) holds if, for all $x \in M \setminus \mathcal{N}_0$ and $r > 0$,

$$C^{-1}F(r) \le \mathbb{E}_x\tau_{B(x,r)} \le CF(r) \tag{E_F}$$

with some constant $C > 1$.

We denote by ($E_F\le$) and ($E_F\ge$) the upper and lower bounds of $\mathbb{E}_x\tau_{B(x,r)}$ in (E_F), respectively.

THEOREM 3.11. *The hypotheses* (VD) + (H) + ($E_F\le$) *imply* (FK).

PROOF. We have by (3.9) and (3.3)

$$\lambda_{\min}(\Omega)^{-1} \le \widetilde{E}(\Omega) = \operatorname*{esup}_{x\in\Omega} G^{\Omega}1_{\Omega}. \tag{3.21}$$

It will be convenient to rename R to $R/2$ and let the original ball B be $B(z, R/2)$ and $\Omega \subset B(z, R/2)$. Fix a point $x \in \Omega$ so that $\Omega \subset B(x, R)$, consider a numerical sequence $R_k = \delta^k R$, $k = 0, 1, 2, \ldots,$ where δ is the parameter from (H), and the balls $B_k = B(x, R_k)$. We have

$$G^{\Omega}1_{\Omega} \le G^{B_0}1_{\Omega} = \sum_{k=0}^{n-1}(G^{B_k} - G^{B_{k+1}})1_{\Omega} + G^{B_n}1_{\Omega},$$

where n is to be chosen (see Figure 2), whence

$$\operatorname*{esup}_{B_{n+1}} G^{\Omega}1_{\Omega} \le \sum_{k=0}^{n-1} \operatorname*{esup}_{B_{k+2}}(G^{B_k} - G^{B_{k+1}})1_{\Omega} + \operatorname*{esup}_{B_n} G^{B_n}1_{\Omega}.$$

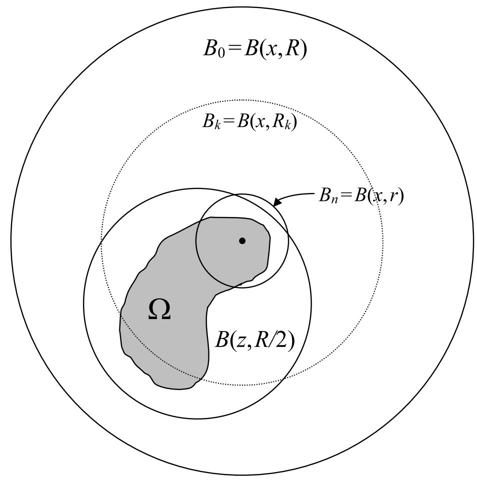

FIG. 2. *Balls B_k.*

Setting $V(r) = V(x,r)$ and using $\widetilde{E}(B_k) \le F(R_k)$, we obtain, by Lemma 3.8,

$$\operatorname*{esup}_{B_{k+2}}(G^{B_k} - G^{B_{k+1}})1_\Omega \le \frac{CF(R_k)}{V(R_k)}\mu(\Omega).$$

Also, by (3.10),

$$\operatorname*{esup}_{B_n} G^{B_n} 1_\Omega \le \operatorname*{esup}_{B_n} G^{B_n} 1 = \widetilde{E}(B_n) \le CF(R_n).$$

Hence, collecting together the previous lines, we obtain

$$\operatorname*{esup}_{B_{n+1}} G^\Omega 1_\Omega \le C \sum_{k=0}^{n-1} \frac{F(R_k)}{V(R_k)}\mu(\Omega) + CF(R_n).$$

Choose any $\nu \in (0, 1)$ so that $\nu < \beta/\alpha$. Using (3.19), (3.13) and the monotonicity of $V(s)$, we obtain

$$\begin{aligned}
\sum_{k=0}^{n-1} \frac{F(R_k)}{V(R_k)} &= \frac{F(R)}{V(R_n)^{1-\nu}V(R)^\nu} \sum_{k=0}^{n-1} \frac{F(R_k)}{F(R)} \left(\frac{V(R)}{V(R_k)}\right)^\nu \left(\frac{V(R_n)}{V(R_k)}\right)^{1-\nu} \\
&\le \frac{CF(R)}{V(R_n)^{1-\nu}V(R)^\nu} \sum_{k=0}^{n-1} \left(\frac{R_k}{R}\right)^\beta \left(\frac{R}{R_k}\right)^{\alpha\nu} \\
&= \frac{CF(R)}{V(R_n)^{1-\nu}V(R)^\nu} \sum_{k=0}^{n-1} \delta^{k(\beta-\alpha\nu)} \\
&\le \frac{CF(R)}{V(R_n)^{1-\nu}V(R)^\nu}.
\end{aligned}$$

Now choose n from the condition

$$V(R_{n+1}) < \mu(\Omega) \le V(R_n),$$

and set $r = R_n$. We obtain then

$$\begin{aligned}\operatorname*{esup}_{B(x,\delta r)} G^{\Omega}1_{\Omega} &\le C\frac{F(R)}{V(r)^{1-\nu}V(R)^{\nu}}\mu(\Omega) + CF(r)\\ &\le CF(R)\left(\frac{V(r)}{V(R)}\right)^{\nu} + CF(r).\end{aligned} \tag{3.22}$$

Using again (3.13), (3.19) and $\alpha\nu < \beta$, we obtain

$$\frac{F(r)}{F(R)} \le C\left(\frac{r}{R}\right)^{\beta} \le C\left(\frac{r}{R}\right)^{a\nu} \le C\left(\frac{V(r)}{V(R)}\right)^{\nu},$$

which implies that the second term in (3.22) can be absorbed by the first one, thus giving

$$\operatorname*{esup}_{B(x,\delta r)} G^{\Omega}1_{\Omega} \le CF(R)\left(\frac{V(r)}{V(R)}\right)^{\nu} \le CF(R)\left(\frac{\mu(\Omega)}{V(R)}\right)^{\nu}.$$

Since the point $x \in \Omega$ was arbitrary, covering Ω by a countable family of balls like $B(x,\delta r)$, we obtain

$$\operatorname*{esup}_{\Omega} G^{\Omega}1_{\Omega} \le CF(R)\left(\frac{\mu(\Omega)}{V(R)}\right)^{\nu},$$

which together with (3.21) finishes the proof. $\square$

3.4. *Estimates of the exit time.* Our main result in this section is Theorem 3.15 saying that the condition (E_F) implies a certain upper bound for the tail $\mathbb{P}_x(\tau_B \le t)$ of the exit time from balls. The results of this type go back to Barlow [2], Theorem 3.11. Here we give a self-contained proof in the present setting, which is based on the ideas of [2]. An alternative analytic approach can be found in [31].

For any open set $\Omega \subset M$, set

$$\overline{E}(\Omega) = \sup_{\Omega\setminus\mathcal{N}_0} \mathbb{E}_x\tau_{\Omega}. \tag{3.23}$$

LEMMA 3.12. *For any open $\Omega \subset M$ such that $\overline{E}(\Omega) < \infty$, we have, for all $t > 0$ and $x \in \Omega \setminus \mathcal{N}_0$,*

$$\mathbb{P}_x(\tau_{\Omega} < t) \le 1 - \frac{\mathbb{E}_x(\tau_{\Omega})}{\overline{E}(\Omega)} + \frac{t}{\overline{E}(\Omega)}. \tag{3.24}$$

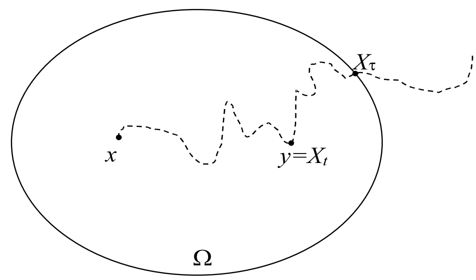

FIG. 3. *Illustration to the proof of Lemma 3.12.*

PROOF. Denote $\tau = \tau_\Omega$, and observe that

$$\tau \le t + (\tau - t)\mathbf{1}_{\{\tau \ge t\}} = t + (\tau \circ \Theta_t)\mathbf{1}_{\{\tau \ge t\}},$$

where Θ_t is the time shift of trajectories. Using the Markov property, we obtain, for any $x \in \Omega \setminus \mathcal{N}_0$,

$$\mathbb{E}_x \tau \le t + \mathbb{E}_x\big((\tau \circ \Theta_t)\mathbf{1}_{\{\tau \ge t\}}\big) = t + \mathbb{E}_x\big(\mathbb{E}_{X_t}(\tau)\mathbf{1}_{\{\tau \ge t\}}\big),$$

whence

$$\mathbb{E}_x \tau \le t + \mathbb{P}_x(\tau \ge t) \sup_{y \in \Omega \setminus \mathcal{N}_0} \mathbb{E}_y \tau = t + \mathbb{P}_x(\tau \ge t)\overline{E}(\Omega)$$

(see Figure 3), and (3.24) follows. □

LEMMA 3.13. *Assume that the condition* (E_F) *is satisfied. Then there are constants* $\varepsilon, \sigma > 0$ *such that, for all* $x \in M \setminus \mathcal{N}_0$, $R > 0$, *and* $\lambda \ge \frac{\sigma}{F(R)}$,

$$\mathbb{E}_x\big(e^{-\lambda \tau_{B(x,R)}}\big) \le 1 - \varepsilon. \tag{3.25}$$

PROOF. Denoting $B = B(x, R)$ and using Lemma 3.12, we have, for any $t > 0$,

$$\begin{aligned}
\mathbb{E}_x(e^{-\lambda \tau_B}) &= \mathbb{E}_x\big(e^{-\lambda \tau_B}\mathbf{1}_{\{\tau_B < t\}}\big) + \mathbb{E}_x\big(e^{-\lambda \tau_B}\mathbf{1}_{\{\tau_B \ge t\}}\big) \\
&\le \mathbb{P}_x(\tau_B < t) + e^{-\lambda t} \\
&\le 1 - \frac{\mathbb{E}_x \tau_B}{\overline{E}(B)} + \frac{t}{\overline{E}(B)} + e^{-\lambda t}.
\end{aligned}$$

The condition (E_F) implies that

$$\overline{E}(B) = \sup_{z \in B(x,R) \setminus \mathcal{N}_0} \mathbb{E}_z \tau_{B(x,R)} \le \sup_{z \in M \setminus \mathcal{N}_0} \mathbb{E}_z \tau_{B(z,2R)} \le C F(2R),$$

whence

$$\overline{E}(B) \le C \mathbb{E}_x \tau_B. \tag{3.26}$$

Using these two estimates of $\overline{E}(B)$, we obtain

$$\mathbb{E}_x(e^{-\lambda\tau_B}) \le 1 - \frac{1}{C} + \frac{Ct}{F(R)} + e^{-\lambda t}.$$

Setting $\varepsilon = \frac{1}{3C}$ and choosing $t = \frac{\varepsilon}{C}F(R)$, we obtain

$$\mathbb{E}_x(e^{-\lambda\tau_B}) \le 1 - 3\varepsilon + \varepsilon + e^{-\lambda t}.$$

If also $e^{-\lambda t} \le \varepsilon$, then we obtain (3.25). Clearly, the former condition will be satisfied provided

$$\lambda \ge \frac{\log(1/\varepsilon)}{t} = \frac{(C/\varepsilon)\log(1/\varepsilon)}{F(R)},$$

which finishes the proof. □

LEMMA 3.14. *Assume that the condition* (E_F) *is satisfied. Then there exists constant* $\gamma > 0$ *such that, for all precompact balls* $B(x, R)$ *with* $x \in M \setminus \mathcal{N}_0$ *and for all* $\lambda > 0$,

$$\mathbb{E}_x\big(e^{-\lambda\tau_{B(x,R)}}\big) \le C\exp\left(-\gamma\frac{R}{\mathcal{R}(1/\lambda)}\right), \tag{3.27}$$

where $\mathcal{R} = F^{-1}$.

PROOF. Rename the center x of the ball to z so that the letter x will be used to denote a variable point. Fix some $0 < r < R$ to be specified later, and set $n = [\frac{R}{r}]$. Set also $\tau = \tau_{B(z,R)}$,

$$u(x) = \mathbb{E}_x(e^{-\lambda\tau})$$

and

$$m_k = \sup_{\overline{B}(z,kr)\setminus\mathcal{N}_0} u,$$

where $k = 1, 2, \dots, n$. Note that all m_k are bounded by 1. Choose $0 < \varepsilon' < \varepsilon$ where ε is the constant from Lemma 3.13, and let x_k be a point in $\overline{B}(z, kr) \setminus \mathcal{N}_0$ for which

$$(1 - \varepsilon')m_k \le u(x_k) \le m_k.$$

Fix $k \le n - 1$, observe that

$$B(x_k, r) \subset B\big(z, (k+1)r\big) \subset B(z, R)$$

and consider the following function in $B(x_k, r)$:

$$v_k(x) = \mathbb{E}_x(e^{-\lambda\tau_k}),$$

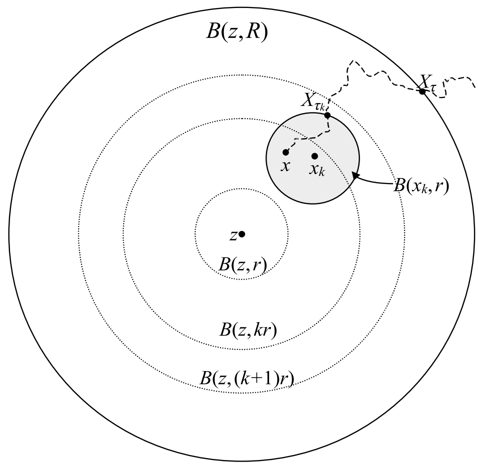

FIG. 4. *Exit times from $B(x_k, r)$ and $B(z, R)$.*

where $\tau_k = \tau_{B(x_k,r)}$ (see Figure 4). Since the ball $B(x_k, r)$ is precompact, we have $X_{\tau_k} \in \overline{B}(x_k, r)$ (while for noncompact balls the exit point could have been at the cemetery).

Let us show that, for all $x \in B(x_k, r) \setminus \mathcal{N}_0$,

$$u(x) \le v_k(x) \sup_{\overline{B}(x_k,r)\setminus\mathcal{N}_0} u. \tag{3.28}$$

Indeed, we have by the strong Markov property

$$\begin{aligned} u(x) &= \mathbb{E}_x(e^{-\lambda\tau_k}) = \mathbb{E}_x\big(e^{-\lambda\tau_k} e^{-\lambda(\tau-\tau_k)}\big) \\ &= \mathbb{E}_x\big(e^{-\lambda\tau_k}(e^{-\lambda\tau} \circ \Theta_{\tau_k})\big) \\ &= \mathbb{E}_x(e^{-\lambda\tau_k}\mathbb{E}_{X_{\tau_k}}(e^{-\lambda\tau})) \\ &= \mathbb{E}_x(e^{-\lambda\tau_k}u(X_{\tau_k})) \\ &\le \mathbb{E}_x(e^{-\lambda\tau_k}) \sup_{\overline{B}(x_k,r)\setminus\mathcal{N}_0} u, \end{aligned}$$

which proves (3.28). It follows from (3.28) that

$$u(x_k) \le v_k(x_k) \sup_{\overline{B}(z,(k+1)r)\setminus\mathcal{N}_0} u = v_k(x_k) m_{k+1},$$

whence

$$(1-\varepsilon')m_k \le v_k(x_k) m_{k+1}.$$

By Lemma 3.13, if

$$\lambda \ge \frac{\sigma}{F(r)}, \tag{3.29}$$

then $v_k(x_k) \leq 1 - \varepsilon$. Therefore, under hypothesis (3.29), we have

$$(1 - \varepsilon')m_k \leq (1 - \varepsilon)m_{k+1},$$

whence it follows by iteration that

$$u(z) \leq m_1 \leq \left(\frac{1 - \varepsilon}{1 - \varepsilon'}\right)^{n-1} m_n \leq C \exp\left(-c\frac{R}{r}\right), \tag{3.30}$$

where in the last inequality we have used that $n \geq \frac{R}{r} - 1$ and $c := \log \frac{1-\varepsilon'}{1-\varepsilon} > 0$.

Condition (3.29) can be satisfied by choosing

$$r = \mathcal{R}\left(\frac{\sigma}{\lambda}\right). \tag{3.31}$$

This value of r is legitimate only if $r < R$, that is, if

$$R > \mathcal{R}\left(\frac{\sigma}{\lambda}\right). \tag{3.32}$$

If (3.32) is not fulfilled, then (3.27) is trivially satisfied by choosing the constant C large enough. Assuming that (3.32) is satisfied and defining r by (3.31) we obtain from (3.30) that

$$u(z) \leq C \exp\left(-c\frac{R}{\mathcal{R}(\sigma/\lambda)}\right),$$

whence (3.27) follows. $\square$

THEOREM 3.15. *Assume that* (E_F) *holds. Then, for any precompact ball* $B(x, R)$ *with* $x \in M \setminus \mathcal{N}_0$ *and for any* $t > 0$,

$$\mathbb{P}_x\big(\tau_{B(x,R)} \leq t\big) \leq C \exp(-\Phi(\gamma R, t)), \tag{3.33}$$

where $\gamma > 0$ *is the constant from Lemma* 3.14 *and*

$$\Phi(R, t) = \sup_{r>0}\left\{\frac{R}{r} - \frac{t}{F(r)}\right\}. \tag{3.34}$$

Changing the variable r in (3.34), we obtain the following equivalent definitions of Φ:

$$\Phi(R, t) = \sup_{\xi>0}\left\{\frac{R}{\mathcal{R}(\xi)} - \frac{t}{\xi}\right\} = \sup_{\lambda>0}\left\{\frac{R}{\mathcal{R}(1/\lambda)} - \lambda t\right\}, \tag{3.35}$$

where $\mathcal{R} = F^{-1}$.

PROOF OF THEOREM 3.15. Denoting $B = B(x, R)$ and using Lemma 3.14, we obtain that, for any $\lambda > 0$,

$$\begin{aligned} \mathbb{P}_x(\tau_B \leq t) &= \mathbb{P}_x(e^{-\lambda\tau_B} \geq e^{-\lambda t}) \\ &\leq e^{\lambda t}\mathbb{E}_x(e^{-\lambda\tau_B}) \\ &\leq C \exp\left(-\gamma\frac{R}{\mathcal{R}(1/\lambda)} + \lambda t\right). \end{aligned} \tag{3.36}$$

Taking the supremum in λ and using (3.35), we obtain (3.33). $\square$

REMARK 3.16. It is clear from (3.34) that function $\Phi(R,t)$ is increasing in R and decreasing in t. Also, we have, for any constants $a,b>0$,

$$\Phi(aR,bt)=ab\Phi\left(\frac{R}{b},\frac{t}{a}\right). \tag{3.37}$$

In particular, it follows that

$$\Phi(R,t)=t\Phi\left(\frac{R}{t},1\right)=t\Phi\left(\frac{R}{t}\right), \tag{3.38}$$

where

$$\Phi(s):=\Phi(s,1)=\sup_{r>0}\left\{\frac{s}{r}-\frac{1}{F(r)}\right\}. \tag{3.39}$$

Hence, (3.33) can be written also in the form

$$\mathbb{P}_x\left(\tau_{B(x,R)}\le t\right)\le C\exp\left(-t\Phi\left(\gamma\frac{R}{t}\right)\right). \tag{3.40}$$

Clearly, $\Phi(0)=0$. Let us show that $0<\Phi(s)<\infty$ for all $s>0$. Since

$$\lim_{r\to\infty}\left(\frac{s}{r}-\frac{1}{F(r)}\right)=0,$$

we see from (3.39) that $\Phi(s)\ge 0$. It follows from (3.19) and $\beta>1$ that

$$\lim_{r\to 0}\frac{r}{F(r)}=\infty \quad \text{and} \quad \lim_{r\to+\infty}\frac{r}{F(r)}=0. \tag{3.41}$$

Given $s>0$, choose r so big that $\frac{r}{F(r)}<s$ [such r exists by the second condition in (3.41)]. Then

$$\Phi(s)\ge\frac{s}{r}-\frac{1}{F(r)}>0.$$

In order to prove that $\Phi(s)<\infty$, it suffices to show that

$$\lim_{r\to 0}\left(\frac{s}{r}-\frac{1}{F(r)}\right)\le 0.$$

Indeed, if r is sufficiently small, then by the first condition in (3.41), $\frac{r}{F(r)}>s$ whence $\frac{s}{r}<\frac{1}{F(r)}$.

Another useful property of function $\Phi(s)$ is the inequality

$$\Phi(as)\ge a\Phi(s) \qquad \text{for all } s\ge 0 \text{ and } a\ge 1. \tag{3.42}$$

Indeed, we have for any $r>0$

$$\frac{as}{r}-\frac{1}{F(r)}\ge a\left(\frac{s}{r}-\frac{1}{F(r)}\right),$$

whence (3.42) follows by taking sup in r.

EXAMPLE 3.17. If $F(r)$ is differentiable then the supremum in (3.39) is attained at the value of r that solves the equation

$$\frac{r^2 F'(r)}{F^2(r)} = s.$$

For example, $F(r) = r^\beta$ then we obtain $r = (\frac{\beta}{s})^{1/(\beta-1)}$ whence $\Phi(s) = cs^{\beta/(\beta-1)}$ and

$$\Phi(R,t) = c\left(\frac{R^\beta}{t}\right)^{1/(\beta-1)}.$$

Consequently, (3.33) becomes

$$\mathbb{P}_x\left(\tau_{B(x,R)} \le t\right) \le C \exp\left(-c\left(\frac{R^\beta}{t}\right)^{1/(\beta-1)}\right).$$

EXAMPLE 3.18. Consider the following example of function F:

$$F(r) = \begin{cases} r^{\beta_1}, & r < 1, \\ r^{\beta_2}, & r \ge 1, \end{cases} \tag{3.43}$$

where $\beta_1, \beta_2 > 1$. It is easy to see that (3.19) is satisfied with $\beta = \beta_1 \wedge \beta_2$ and $\beta' = \beta_1 \vee \beta_2$. Similarly to the previous example, one obtains that

$$\Phi(s) \simeq \begin{cases} s^{\beta_1/(\beta_1-1)}, & s > 1, \\ s^{\beta_2/(\beta_2-1)}, & s \le 1, \end{cases} \tag{3.44}$$

so that (3.33) becomes

$$\mathbb{P}_x\left(\tau_{B(x,R)} \le t\right) \le C \begin{cases} \exp\left(-c\left(\frac{R^{\beta_1}}{t}\right)^{1/(\beta_1-1)}\right), & t < R, \\ \exp\left(-c\left(\frac{R^{\beta_2}}{t}\right)^{1/(\beta_2-1)}\right), & t \ge R. \end{cases}$$

LEMMA 3.19. *The function $\Phi(R,t)$ satisfies the following inequality*:

$$\Phi(R,t) \ge c \min\left\{\left(\frac{F(R)}{t}\right)^{1/(\beta'-1)}, \left(\frac{F(R)}{t}\right)^{1/(\beta-1)}\right\} \tag{3.45}$$

for all $R, t > 0$.

PROOF. By (3.34), we have, for any $r > 0$,

$$\Phi(R,t) \ge \frac{R}{r} - \frac{t}{F(r)}.$$

We claim that there exists $r > 0$ such that

$$\frac{t}{F(r)} = \frac{1}{2}\frac{R}{r}. \tag{3.46}$$

Indeed, the function $\frac{F(r)}{r}$ is continuous on $(0, +\infty)$, tends to 0 as $r \to 0$ and to ∞ as $r \to \infty$ so that $\frac{F(r)}{r}$ takes all positive values, whence the claim follows. With the value of r as in (3.46), we have

$$\Phi(R, t) \geq \frac{t}{F(r)}. \tag{3.47}$$

If $r \leq R$ then using the left-hand side inequality of (3.19), we obtain

$$\frac{R}{r} \geq c\left(\frac{F(R)}{F(r)}\right)^{1/\beta},$$

which together with (3.46) yields

$$F(r) \leq C\left(\frac{t^\beta}{F(R)}\right)^{1/(\beta-1)}.$$

Substituting into (3.47), we obtain

$$\Phi(R, t) \geq c\left(\frac{F(R)}{t}\right)^{1/(\beta-1)}.$$

Similarly, it $r > R$ then using the right-hand side inequality in (3.19) we obtain

$$\frac{R}{r} \geq c\left(\frac{F(R)}{F(r)}\right)^{1/\beta'},$$

whence it follows that

$$\Phi(R, t) \geq c\left(\frac{F(R)}{t}\right)^{1/(\beta'-1)}. \qquad \square$$

COROLLARY 3.20. *Under the hypotheses of Theorem* 3.15, *we have, for any* $x \in M \setminus \mathcal{N}_0$, $R > 0$, $t > 0$,

$$\mathbb{P}_x\left(\tau_{B(x,R)} \leq t\right) \leq C \exp\left(-c\left(\frac{F(R)}{t}\right)^{1/(\beta'-1)}\right). \tag{3.48}$$

PROOF. Indeed, if $\frac{F(R)}{t} \geq 1$, then (3.48) follows from Theorem 3.15, Lemma 3.19 and (3.19). If $\frac{F(R)}{t} < 1$, then (3.48) is trivial. $\square$

4. Upper bounds of heat kernel. The following result will be used in the proof of Theorem 4.2 below.

PROPOSITION 4.1 ([31], Lemma 5.5). *Let U be an open subset of M, and assume that, for any nonempty open set $\Omega \subset U$,*

$$\lambda_{\min}(\Omega) \geq a\mu(\Omega)^{-\nu}$$

for some positive constants a,ν. Then the semigroup $\{P_t^B\}$ is ultracontractive with the following estimate:

$$\|P_t^B f\|_\infty \le C(at)^{-1/\nu}\|f\|_1 \tag{4.1}$$

for any $f\in L^1(B)$.

The next theorem provides pointwise upper bounds for the heat kernel.

THEOREM 4.2. *If the conditions $(VD)+(FK)+(E_F)$ are satisfied and all balls are precompact then the heat kernel exists with the domain $M\setminus\mathcal{N}$ for some properly exceptional set $\mathcal{N}$, and satisfies the upper bound*

$$p_t(x,y)\le \frac{C}{V(x,\mathcal{R}(t))}\exp\left(-\frac{1}{2}\Phi(cd(x,y),t)\right) \tag{UE}$$

for all $t>0$ and $x,y\in M\setminus\mathcal{N}$, where $\mathcal{R}=F^{-1}$ and Φ is defined by (3.34).

REMARK 4.3. As it follows from Theorem 3.11, the hypotheses $(VD)+(FK)+(E_F)$ here can be replaced by $(VD)+(H)+(E_F)$. Also, using (3.38), one can write (UE) in the form

$$p_t(x,y)\le \frac{C}{V(x,\mathcal{R}(t))}\exp\left(-\frac{1}{2}t\Phi\left(c\frac{d(x,y)}{t}\right)\right)$$

as it was stated in the Introduction.

REMARK 4.4. A version of Theorem 4.2 was proved by Kigami [45] under additional assumptions that the heat kernel is a priori continuous and ultracontractive, and using instead of (FK) a local Nash inequality. In the case $F(r)=r^\beta$, another version of Theorem 4.2 was proved in [31], where the upper bound (UE) was understood for *almost all* x,y. The proof below uses a combination of techniques from [31] and [45].

EXAMPLE 4.5. If function $F(r)$ is given by (3.43) as in Example 3.18, then $\Phi(s)$ is given by (3.44) and (UE) becomes

$$p_t(x,y)\le \frac{C}{V(x,\mathcal{R}(t))}\begin{cases}\exp\left(-c\left(\frac{r^{\beta_1}}{t}\right)^{1/(\beta_1-1)}\right), & t<r,\\ \exp\left(-c\left(\frac{r^{\beta_2}}{t}\right)^{1/(\beta_2-1)}\right), & t\ge r,\end{cases}$$

where $r=d(x,y)$.

PROOF OF THEOREM 4.2. The hypothesis (FK) can be stated as follows: for any ball $B=B(x,r)$ where $x\in M$ and $r>0$, and for any nonempty open set $\Omega\subset B$, we have

$$\lambda_{\min}(\Omega)\ge a(B)\mu(\Omega)^{-\nu}, \tag{4.2}$$

where

$$a(B) = \frac{c}{F(r)}\mu(B)^{\nu} \tag{4.3}$$

and ν, c are positive constants. Hence, (FK) implies by Proposition 4.1 that

$$\|P_t^B f\|_{L^1 \to L^\infty} \le \frac{C}{(a(B)t)^{1/\nu}}. \tag{4.4}$$

In particular, the semigroup $\{P_t^B\}$ is ultracontractive and $\{P_t\}$ is locally ultracontractive. By Theorem 2.12, there exists a properly exceptional set $\mathcal{N} \subset M$ (containing $\mathcal{N}_0$) such that, for any open subset $\Omega \subset M$, the semigroup $\mathcal{P}_t^{\Omega}$ possesses the heat kernel $p_t^{\Omega}(x, y)$ with the domain $\Omega \setminus \mathcal{N}$. Fix this set $\mathcal{N}$ for what follows. By Theorem 2.10, (4.3) and (4.4) imply the following estimate:

$$p_t^B(x, y) \le \frac{C}{(a(B)t)^{1/\nu}} = \frac{C}{\mu(B)}\left(\frac{F(r)}{t}\right)^{1/\nu} \tag{4.5}$$

for any ball B of radius r, and for all $t > 0$, $x, y \in B \setminus \mathcal{N}$.

Our next step is to prove the on-diagonal estimate

$$p_t(x, x) \le \frac{C}{V(x, \mathcal{R}(t))} \tag{DUE}$$

for all $x \in M \setminus \mathcal{N}$ and $t > 0$. To understand the difficulties, let us first consider a particular case when the volume function satisfies the following estimate:

$$V(x, R) \simeq F(R)^{1/\nu} \tag{4.6}$$

for all $x \in M$ and $R > 0$, where ν is the exponent in (FK) [e.g., (4.6) holds, if $V(x, R) \simeq R^{\alpha}$, $F(R) = R^{\beta}$ and $\nu = \beta/\alpha$]. In this case, the value $F(R)$ in (FK) cancels our, and we obtain

$$\lambda_{\min}(\Omega) \ge c\mu(\Omega)^{-\nu}. \tag{4.7}$$

Hence, by Proposition 4.1, the semigroup $\{P_t\}$ is ultracontractive, and by Theorem 2.10 we obtain the estimate

$$p_t(x, x) \le Ct^{-1/\nu} \tag{4.8}$$

for all $x \in M \setminus \mathcal{N}$ and $t > 0$. Observing that

$$V(x, \mathcal{R}(t)) \simeq F(\mathcal{R}(t))^{1/\nu} = t^{1/\nu},$$

we see that (4.8) is equivalent to (DUE). Although this argument works only under restriction (4.6), it has an advantage that it can be localized as follows. Assuming that (4.6) is satisfied for all $R < R_0$ with some fixed constant R_0, (4.7) is satisfied for all open sets Ω with a bounded value of $\mu(\Omega)$, and (E_F) is satisfied for all balls with a bounded value of the radius, one can prove in the same way that (4.8) is true for $t < t_0$ for some $t_0 > 0$. The proof below does not allow such a localization in the general case.

In the general case, without the hypotheses (4.6), the heat semigroup $\{P_t\}$ is not necessarily ultracontractive, which requires other tools for obtaining (*DUE*). In the case of Riemannian manifolds, one can obtain (*DUE*) from (*FK*) using a certain mean value inequality (see [15, 25]) but this method heavily relies on a specific property of the distance function that $|\nabla d| \le 1$, which is not available in our generality. We will use Kigami's iteration argument that allows us to obtain (*DUE*) from (4.5) using, in addition, the hypothesis (E_F). This argument is presented in an abstract form in [31], Lemma 5.6, that says the following. Assume that the following two conditions are satisfied:

(1) for any ball $B = B(x, r)$,

$$\operatorname*{esup}_B p_t^B \le \Psi_t(x, r), \tag{4.9}$$

where function $\Psi_t(x, t)$ satisfies certain conditions[9];

(2) for all $x \in M \setminus \mathcal{N}_0$, $t > 0$, and $r \ge \varphi(t)$,

$$\mathbb{P}_x(\tau_B \le t) \le \varepsilon, \tag{4.10}$$

where $\varepsilon > 0$ is a sufficiently small[10] constant and φ is a positive increasing function on $(0, +\infty)$ such that

$$\int_0 \varphi(t)\,\frac{dt}{t} < \infty. \tag{4.11}$$

Then the heat kernel on M satisfies the estimate

$$\operatorname*{esup}_{B(x,\varphi(t))} p_t \le C\Psi_t(x, \varphi(t)). \tag{4.12}$$

Obviously, (4.5) implies (4.9) with the function

$$\Psi_t(x, r) = \frac{C}{V(x, r)}\left(\frac{F(r)}{t}\right)^{1/\nu}. \tag{4.13}$$

By Corollary 3.20, (E_F) implies (3.48), which means that (4.10) is satisfied provided $Ct \le F(r)$ for a sufficiently large constant C; hence, the function $\varphi(t)$ can be chosen as follows:

$$\varphi(t) = \mathcal{R}(Ct),$$

---

[9]Function $\Psi_t(x, r)$ should be monotone decreasing in t and should satisfy the following doubling condition: if $r \le r' \le 2r$ and $t' \ge t/2$, then

$$\Psi_{t'}(x, r') \le K\Psi_t(x, r)$$

for some constant K. This is obviously satisfied for the function Ψ given by (4.13).

[10]More precisely, this means that $\varepsilon \le \frac{1}{2K}$ where K is the constant from the conditions for Ψ.

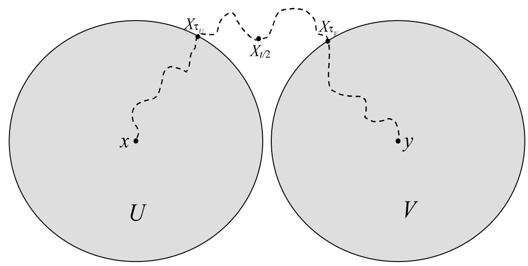

FIG. 5. *Illustration to* (4.14).

which clearly satisfies (4.11) [indeed, by (3.20), we have $\mathcal{R}(t) \le Ct^{1/\beta'}$ for all $0 < t < 1$, whence (4.11) follows].

By (4.12), we obtain

$$\operatorname*{esup}_{B(x,\varphi(t))} p_t \le C\Psi_t(x, \mathcal{R}(Ct)) \le \frac{C}{V(x, \mathcal{R}(Ct))} \le \frac{C}{V(x, \mathcal{R}(t))},$$

where we have also used (3.20) and (3.13). By Theorem 2.12(d), we can replace here esup p_t by sup p_t outside $\mathcal{N}$, whence (*DUE*) follows.

Now we prove the full upper estimate (*UE*). Fix two disjoint open subsets U, V of M and use the following inequality proved in [9], Lemma 2.1: for all functions $f, g \in \mathcal{B}_+(M)$,

$$\begin{aligned}(\mathcal{P}_t f, g) \le {} & \big(\mathbb{E}.\big(\mathbf{1}_{\{\tau_U \le t/2\}} \mathcal{P}_{t-\tau_U} f(X_{\tau_U})\big), g\big) \\ & + \big(\mathbb{E}.\big(\mathbf{1}_{\{\tau_V \le t/2\}} \mathcal{P}_{t-\tau_V} g(X_{\tau_V})\big), f\big)\end{aligned} \tag{4.14}$$

(see Figure 5).

Assume in addition that $f \in \mathcal{B}L^1(V)$ and $g \in \mathcal{B}L^1(U)$. Then, under the condition $\tau_U \le t/2$, we have

$$\mathcal{P}_{t-\tau_U} f(X_{\tau_U}) = \int_{V \setminus \mathcal{N}} p_{t-\tau_U}(X_{\tau_U}, y) f(y)\, d\mu(y) \le S\|f\|_1$$

almost surely, where

$$S := \sup_{t/2 \le s \le t} \sup_{\substack{u \in \overline{U} \setminus \mathcal{N} \\ v \in \overline{V} \setminus \mathcal{N}}} p_s(u, v). \tag{4.15}$$

Here we have used that $X_{\tau_U} \in \overline{U} \setminus \mathcal{N}$ almost surely, which is due to the fact that $\{X_t\}$ is a diffusion and the set $\mathcal{N}$ is properly exceptional. It follows that

$$\big(\mathbb{E}.\big(\mathbf{1}_{\{\tau \le t/2\}} \mathcal{P}_{t-\tau} f(X_\tau)\big), g\big) \le S\|f\|_1 \int_U \mathbb{P}_x\Big(\tau_U \le \frac{t}{2}\Big) g(x)\, d\mu(x).$$

Estimating similarly the second term in (4.14), we obtain from (4.14)

$$\begin{aligned}\int_U \int_V p_t(x, y) f(y) g(x)\, d\mu(x)\, d\mu(y) &\\ \leq S\|f\|_1 \int_U \mathbb{P}_x\left(\tau_U \leq \frac{t}{2}\right) g(x)\, d\mu(x) &\\ + S\|g\|_1 \int_V \mathbb{P}_y\left(\tau_V \leq \frac{t}{2}\right) f(y)\, d\mu(y).&\end{aligned}$$

By [31], Lemma 3.4, we conclude that, for μ-a.a. $x \in V$ and $y \in U$,

$$p_t(x, y) \leq S\mathbb{P}_x\left(\tau_V \leq \frac{t}{2}\right) + S\mathbb{P}_y\left(\tau_U \leq \frac{t}{2}\right). \tag{4.16}$$

A slightly different inequality (i.e., also enough for our purposes) was proved in [33]. For the case of heat kernels on Riemannian manifolds, (4.16) was proved in [35], Lemma 3.3.

By Theorem 3.15, we have

$$\mathbb{P}_x\left(\tau_{B(x,R)} \leq t\right) \leq C \exp(-\Phi(\gamma R, t)) \tag{4.17}$$

for all $x \in M \setminus \mathcal{N}$ and $t, R > 0$. Let

$$V_R = \{x \in V : d(x, V^c) > R\}.$$

Then, for any $x \in V_R \setminus \mathcal{N}$, we obtain by (4.17)

$$\mathbb{P}_x\left(\tau_V \leq \frac{t}{2}\right) \leq \mathbb{P}_x\left(\tau_{B(x,R)} \leq t\right) \leq C \exp(-\Phi(\gamma R, t)).$$

Using a similar estimate for $y \in U_R$, we obtain from (4.16) that, for μ-a.a. $x \in V_R$ and $y \in U_R$,

$$p_t(x, y) \leq CS \exp(-\Phi(\gamma R, t)). \tag{4.18}$$

Since the right-hand side here is a constant in x, y, we conclude by Theorem 2.12(d) that (4.18) holds for all $x \in V_R \setminus \mathcal{N}$ and $y \in U_R \setminus \mathcal{N}$.

Now fix two distinct points $x, y \in M \setminus \mathcal{N}$, set

$$R = \tfrac{1}{4} d(x, y) \tag{4.19}$$

and observe that the balls $V = B(x, 2R)$ and $U = B(y, 2R)$ are disjoint. Since $x \in V_R$ and $y \in U_R$, we conclude that (4.18) is satisfied for these points x, y with the above value of R. Let us estimate the quantity S defined by (4.15). Using the semigroup property and (*DUE*), we obtain, for all $u, v \in M \setminus \mathcal{N}$,

$$p_s(u, v) \leq \sqrt{p_s(u, u) p_s(v, v)} \leq \frac{C}{\sqrt{V(u, \mathcal{R}(s)) V(v, \mathcal{R}(s))}}.$$

Observe that by (3.13), for all $z \in M$,

$$\frac{V(z, \mathcal{R}(s))}{V(x, \mathcal{R}(s))} \le C\left(1 + \frac{d(x,z)}{\mathcal{R}(s)}\right)^{\alpha}. \tag{4.20}$$

Applying this for $z = u \in \overline{U}$ and $z = v \in \overline{V}$ so that $d(x,u) \le 2R$ and $d(x,v) \le 6R$, and substituting to the above estimate of $p_s(u,v)$ we obtain

$$p_s(u,v) \le \frac{C}{V(x, \mathcal{R}(s))}\left(1 + \frac{R}{\mathcal{R}(s)}\right)^{\alpha}.$$

Using that $s \in [t/2, t]$ as well as (3.20) and (3.13), we obtain

$$S \le \frac{C}{V(x, \mathcal{R}(t))}\left(1 + \frac{R}{\mathcal{R}(t)}\right)^{\alpha},$$

which together with (4.18) yields

$$p_t(x,y) \le \frac{C}{V(x, \mathcal{R}(t))}\left(1 + \frac{R}{\mathcal{R}(t)}\right)^{\alpha} \exp(-\Phi(\gamma R, t)). \tag{4.21}$$

On the other hand, we have by (3.35)

$$\Phi(R,t) = \sup_{\xi > 0}\left\{\frac{R}{\mathcal{R}(\xi)} - \frac{t}{\xi}\right\} \ge \frac{R}{\mathcal{R}(t)} - 1,$$

where we have chosen $\xi = t$. Using the elementary estimate

$$1 + z \le \frac{1}{a}\exp(az), \qquad z > 0, 0 < a \le 1,$$

and its consequence

$$2 + z \le \frac{2}{a}\exp(az),$$

we obtain

$$1 + \frac{R}{\mathcal{R}(t)} \le 2 + \Phi(R,t) \le \frac{2}{a}\exp\left(\frac{a}{\gamma}\Phi(\gamma R, t)\right),$$

whence

$$\left(1 + \frac{R}{\mathcal{R}(t)}\right)^{\alpha} \le \left(\frac{2}{a}\right)^{\alpha}\exp\left(\frac{\alpha a}{\gamma}\Phi(\gamma R, t)\right). \tag{4.22}$$

Choosing a small enough and substituting this estimate to (4.21), we obtain (UE). $\square$

REMARK 4.6. It is desirable to have a localized version of Theorem 4.2 when the hypotheses are assumed for balls of bounded radii and the conclusions are proved for a bounded range of time. As was already mentioned in the proof, Kigami's argument requires the ultracontractivity of P_t^B for *all* balls, and (E_F)

should also be satisfied for all balls, because, loosely speaking, one deals with the estimate of $p_t^{B_{k+1}} - p_t^{B_k}$ for an exhausting sequence of balls $\{B_k\}$ (see [31] or [45]). As we will see in Section 7.2, the hypotheses $(VD) + (H) + (E_F)$ for all balls imply that the space (M, d) is unbounded. Note that all other arguments used in this paper do admit localization.

5. Lower bounds of heat kernel.

5.1. *Oscillation inequalities.* The Harnack inequality (H) is a standing assumption in this subsection. The main result is Proposition 5.3 that is heavily based on the oscillation inequality of Lemma 5.2. The latter is considered as a standard consequence of (H), but we still provide a full proof to emphasize the use of the precompactness of balls.

LEMMA 5.1. *Let B be a precompact ball of radius r in M. If $u \in \mathcal{F}$ is harmonic in B, and if $u \geq a$ in B for some real constant a, then*

$$\operatorname*{esup}_{\delta B}(u - a) \leq C \operatorname*{einf}_{\delta B}(u - a), \tag{5.1}$$

where C and δ are the same constants as in (H).

PROOF. Let ψ be a cutoff function of B in M, that is, $\psi \in \mathcal{F} \cap C_0(M)$ and $\psi \equiv 1$ in an open neighborhood of $\overline{B}$. The function $v = u - a\psi$ belongs to $\mathcal{F}$ and is equal to $u - a$ in B. Let us show that v is harmonic in B. Indeed, for any $\varphi \in \mathcal{F}(B)$, we have

$$\mathcal{E}(v, \varphi) = \mathcal{E}(u - a\psi, \varphi) = \mathcal{E}(u, \varphi) - a\mathcal{E}(\psi, \varphi) = 0, \tag{5.2}$$

because $\mathcal{E}(u, \varphi) = 0$ by the harmonicity of u in B, and $\mathcal{E}(\psi, \varphi) = 0$ by the strong locality as $\psi \equiv 1$ in a neighborhood of $\operatorname{supp} \varphi$. Applying (H) to v, we obtain (5.1). $\square$

For any function f on any set $S \subset M$, define the essential oscillation of f in S by

$$\operatorname*{eosc}_S f = \operatorname*{esup}_S f - \operatorname*{einf}_S f.$$

The following statement is well known for functions in $\mathbb{R}^n$ (see, e.g., [53] and [56], Lemma 2.3.2).

LEMMA 5.2. *There exists $\theta > 0$ such that, for any precompact ball $B(x, r) \subset M$, for any nonnegative harmonic function u in $B(x, r)$, and any $0 < \rho \leq r$,*

$$\operatorname*{eosc}_{B(x,\rho)} u \leq 2\left(\frac{\rho}{r}\right)^\theta \operatorname*{eosc}_{B(x,r)} u, \tag{5.3}$$

where θ is a positive constant that depends on the constants in (H).

PROOF. Fix $x \in M$, and write for simplicity $B_r = B(x,r)$. Consider first the case when $\rho = \delta r$ [where δ is a parameter from (H)] and set

$$a = \operatorname*{esup}_{B_r} u, \qquad b = \operatorname*{einf}_{B_r} u$$

and

$$a' = \operatorname*{esup}_{B_\rho} u, \qquad b' = \operatorname*{einf}_{B_\rho} u.$$

Clearly, $b \le b' \le a' \le a$. By Lemma 5.1, we have

$$\operatorname*{esup}_{B_\rho}(u-b) \le C \operatorname*{einf}_{B_\rho}(u-b), \tag{5.4}$$

that is,

$$a' - b \le C(b' - b).$$

Similarly, applying Lemma 5.1 to function $-u$, we obtain

$$\operatorname*{esup}_{B_\rho}(a-u) \le C \operatorname*{einf}_{B_\rho}(a-u),$$

whence

$$a - b' \le C(a - a').$$

Adding up the two inequalities yields

$$(a-b) + (a'-b') \le C(a-b) - C(a'-b'),$$

whence

$$a' - b' \le \frac{C-1}{C+1}(a-b),$$

that is,

$$\operatorname*{eosc}_{B_{\delta r}} u \le \gamma \operatorname*{eosc}_{B_r} u, \tag{5.5}$$

where $\gamma := \frac{C-1}{C+1} < 1$. For an arbitrary $\rho \le r$, find a nonnegative integer k such that

$$\delta^{k+1} r < \rho \le \delta^k r.$$

Iterating (5.5), we obtain

$$\begin{aligned}\operatorname*{eosc}_{B_\rho} u \le \operatorname*{eosc}_{B_{\delta^k r}} u \le \gamma^k \operatorname*{eosc}_{B_r} u \le \gamma^{\log(r/\rho)\log(1/\delta)-1} \operatorname*{eosc}_{B_r} u \\ = \frac{1}{\gamma}\left(\frac{\rho}{r}\right)^{(\log(1/\gamma))/(\log(1/\delta))} \operatorname*{eosc}_{B_r} u.\end{aligned}$$

Note that the constant C in (5.4) can be assumed to be big enough, say $C > 3$. Then $\gamma > 1/2$ and (5.3) follows from the previous line with $\theta = (\log(1/\gamma))/(\log(1/\delta))$. $\square$

PROPOSITION 5.3. *Let Ω be an open subset of M such that $\widetilde{E}(\Omega) < \infty$. Fix a function $f \in \mathcal{B}_b(\Omega)$, and set $u = G^{\Omega} f$. Then, for any precompact ball $B(x,r) \subset \Omega$ and all $\rho \in (0,r]$,*

$$\operatorname*{eosc}_{B(x,\rho)} u \leq 2\widetilde{E}(x,r) \operatorname*{esup}_{B(x,r)} |f| + 4\left(\frac{\rho}{r}\right)^{\theta} \operatorname*{esup}_{B(x,r)} |u|, \tag{5.6}$$

where θ is the same constant as in Lemma 5.2 and

$$\widetilde{E}(x,r) := \widetilde{E}(B(x,r)).$$

PROOF. Write for simplicity $B_r := B(x,r)$. Let us first prove that if $f \geq 0$, then

$$\operatorname*{eosc}_{B_\rho} u \leq \widetilde{E}(x,r) \operatorname*{esup}_{B_r} f + 2\left(\frac{\rho}{r}\right)^{\theta} \operatorname*{esup}_{B_r} u. \tag{5.7}$$

By Lemma 3.2, we have for the function $v = G^{B_r} f$ that

$$\operatorname*{eosc}_{B_\rho} v \leq \operatorname*{esup}_{B_r} v \leq \widetilde{E}(x,r) \operatorname*{esup}_{B_r} f.$$

The function $w := u - v$ is harmonic in B_r by Lemma 3.4 and nonnegative by Theorem 2.12(b). By Lemma 5.2, we obtain

$$\operatorname*{eosc}_{B_\rho} w \leq 2\left(\frac{\rho}{r}\right)^{\theta} \operatorname*{eosc}_{B_r} w \leq 2\left(\frac{\rho}{r}\right)^{\theta} \operatorname*{esup}_{B_r} u.$$

Since $u = v + w$, (5.7) follows by adding up the two previous lines.

For a signed function f, write $f = f_+ - f_-$ and set

$$\overline{u} := G^{\Omega} f_+ \quad \text{and} \quad \underline{u} := G^{\Omega} f_-.$$

Then $\overline{u}$ and $\underline{u}$ are nonnegative and $u = \overline{u} - \underline{u}$, whence it follows that

$$\operatorname{eosc} u = \operatorname{eosc}(\overline{u} - \underline{u}) \leq \operatorname{eosc} \overline{u} + \operatorname{eosc} \underline{u}.$$

Applying (5.7) separately to $\overline{u}$ and $\underline{u}$ and adding up the inequalities, we obtain (5.6). □

5.2. *Time derivative.* In this section, we assume only the basic hypotheses. If $f \in L^2(M)$, then, for any $t > 0$, the function $u_t = P_t f$ is in $\operatorname{dom}(\mathcal{L})$ and satisfies the heat equation

$$\partial_t u_t + \mathcal{L} u = 0, \tag{5.8}$$

where $\partial_t u_t$ is the strong derivative in $L^2(M)$ of the mapping $t \mapsto u_t$; cf. [30] and [28], Section 4.3. The argument in the next lemma is well known in the context of the semigroup theory (see [16, 17]), and we reproduce it here for the sake of completeness.

LEMMA 5.4. *For any $f \in L^2(M)$ and all $0 \le s < t$, we have*

$$\|\partial_t u_t\|_2 \le \frac{1}{t-s}\|u_s\|_2, \tag{5.9}$$

where $u_t = \mathcal{P}_t f$.

PROOF. Let $\{E_\lambda\}_{\lambda \ge 0}$ be spectral resolution in $L^2(M)$ of the operator $\mathcal{L}$. Then we have

$$u_t = e^{-t\mathcal{L}} f = \int_0^\infty e^{-t\lambda}\, dE_\lambda f,$$

$$\partial_t u_t = -\mathcal{L}e^{-t\mathcal{L}} f = \int_0^\infty (-\lambda) e^{-t\lambda}\, dE_\lambda f,$$

whence

$$\|u_t\|_2^2 = \int_0^\infty e^{-2t\lambda}\, d\|E_\lambda f\|^2,$$

$$\|\partial_t u_t\|_2^2 = \int_0^\infty \lambda^2 e^{-2t\lambda}\, d\|E_\lambda f\|^2.$$

Since

$$\lambda^2 e^{-2t\lambda} = \big(\lambda e^{-(t-s)\lambda}\big)^2 e^{-2s\lambda} \le \frac{1}{(t-s)^2} e^{-2s\lambda},$$

we obtain

$$\|\partial_t u_t\|_2^2 \le \frac{1}{(t-s)^2} \int_0^\infty e^{-2s\lambda}\, d\|E_\lambda f\|^2 = \frac{1}{(t-s)^2}\|u_s\|_2^2,$$

which was to be proved. □

In the rest of this section, assume that $p_t(x, y)$ is the heat kernel with domain D.

COROLLARY 5.5. *For any $t > 0$ and $y \in D$, the function $t \mapsto p_t(\cdot, y)$ is strongly differentiable in $L^2(M)$ and, for all $0 < s < t$,*

$$\|\partial_t p_t(\cdot, y)\|_2 \le \frac{1}{t-s}\sqrt{p_{2s}(y, y)}.$$

PROOF. Setting $f = p_\varepsilon(\cdot, y)$ for some $\varepsilon > 0$ and using (2.12), we obtain that the function

$$u_t = \mathcal{P}_t f = p_{t+\varepsilon}(\cdot, y)$$

satisfies (5.9), that is,

$$\|\partial_t p_{t+\varepsilon}(\cdot, y)\|_2 \le \frac{1}{t-s}\|p_{s+\varepsilon}(\cdot, y)\|_2 = \frac{1}{t-s}\sqrt{p_{2(s+\varepsilon)}(y, y)}.$$

Renaming $t+\varepsilon$ by t and $s+\varepsilon$ by s, we finish the proof. $\square$

Set

$$p'_t(x,y) \equiv \partial_t p_t(x,y),$$

where the strong derivative ∂_t is taken with respect to the first variable x. Hence, for any $y \in D$ and $t>0$, $p'_t(x,y)$ is an L^2-function of x.

LEMMA 5.6. *For all* $0<s<t$ *and* $y \in D$, *we have*

$$p'_t(\cdot,y) = \mathcal{P}_s p'_{t-s}(\cdot,y). \tag{5.10}$$

PROOF. Indeed, we have by (2.12)

$$p_t(\cdot,y) = \mathcal{P}_s p_{t-s}(\cdot,y).$$

Since $\mathcal{P}_s$ is a bounded operator in L^2, it commutes with the operator ∂_t of strong derivation. Applying the latter to the both sides of the above identity, we obtain (5.10). $\square$

COROLLARY 5.7. *For all* $t>0$, $y \in D$ *and* μ-*a.a.* $x \in D$,

$$|\partial_t p_t(x,y)| \le \frac{2}{t}\sqrt{p_{t/2}(x,x)p_{t/2}(y,y)}. \tag{5.11}$$

PROOF. By Lemma 5.6, we have, for all $y \in D$ and μ-a.a. $x \in D$,

$$p'_t(x,y) = (p_s(x,\cdot), p'_{t-s}(\cdot,y)),$$

whence by Corollary 5.5,

$$|p'_t(x,y)| \le \|p_s(x,\cdot)\|_2 \|p'_{t-s}(\cdot,y)\|_2 \le \frac{1}{t-s-r}\sqrt{p_{2s}(x,x)p_{2r}(y,y)}$$

for any $0<r<t-s$. Choosing $s=r=t/4$, we finish the proof of (5.11). $\square$

REMARK 5.8. It follows easily from the identity

$$p_t(x,y) = (p_s(x,\cdot), p_{t-s}(\cdot,y)),$$

that the function $t \mapsto p_t(x,y)$ is differentiable for all fixed $x,y \in D$ and

$$\frac{\partial}{\partial t}p_t(x,y) = (p_s(x,\cdot), \partial_t p_{t-s}(\cdot,y)) = (p_s(x,\cdot), q_{t-s}(\cdot,y)).$$

Arguing as in the previous proof, we obtain

$$\left|\frac{\partial}{\partial t}p_t(x,y)\right| \le \frac{2}{t}\sqrt{p_{t/2}(x,x)p_{t/2}(y,y)}$$

for all $t>0$ and $x,y \in D$. However, for applications we need estimate (5.11) for the strong derivative $\partial_t p_t$ rather than for the partial derivative $\frac{\partial}{\partial t}p_t(x,y)$.

LEMMA 5.9. *If Ω is an open subset of M and if $\widetilde{E}(\Omega) < \infty$, then, for all $t > 0$ and $z \in M$, the function $u_t := p_t^{\Omega}(\cdot, z)$ satisfies in $(0, +\infty) \times \Omega$ the equation*

$$G^{\Omega}(\partial_t u_t) + u_t = 0.$$

PROOF. By Lemma 3.2, the Green operator G^{Ω} is a bounded operator in $L^2(\Omega)$, and G^{Ω} is the inverse operator to $\mathcal{L}^{\Omega}$. Since the function u_t satisfies the equation $\partial_t u_t + \mathcal{L}^{\Omega} u_t = 0$, applying G^{Ω} proves the claim. $\square$

5.3. *The Hölder continuity.* In this subsection we use the hypotheses $(VD) + (H) + (E_F \leq)$. As it follows from Theorem 3.11 and Proposition 4.1, under these hypotheses the heat semigroup $\{P_t\}$ is locally ultracontractive. Hence, by Theorem 2.12, for any open set $\Omega \subset M$, the heat kernel p_t^{Ω} exists with the domain $\Omega \setminus \mathcal{N}$ where $\mathcal{N} \subset M$ is a fixed properly exceptional set; cf. the beginning of the proof of Theorem 4.2.

LEMMA 5.10. *Let the hypotheses $(VD) + (H) + (E_F \leq)$ be satisfied, and let Ω be an open subset of M. Fix $t > 0$, $0 < \rho \leq \mathcal{R}(t)$, and set*

$$r = (\mathcal{R}(t)^{\beta} \rho^{\theta})^{1/(\beta+\theta)}, \tag{5.12}$$

where β is the exponent from (3.19), and θ is the constant from Lemma 5.2. Fix also a point $x \in \Omega \setminus \mathcal{N}$ and assume that the ball $B(x, r)$ is precompact, and its closure is contained in Ω. Then

$$\operatorname*{osc}_{y \in B(x,\rho)\setminus\mathcal{N}} p_t^{\Omega}(x, y) \leq C \left(\frac{\rho}{\mathcal{R}(t)} \right)^{\Theta} \sup_{y \in B(x,r)\setminus\mathcal{N}} p_{t/2}^{\Omega}(y, y), \tag{5.13}$$

where $\Theta = \frac{\beta\theta}{\beta+\theta}$ and C depends on the constants in $(E_F \leq)$ and (3.19).

PROOF. By construction in the proof of Theorem 2.12, the heat kernel p_t^{Ω} is obtained as a monotone increasing limit of $p_t^{U_n}$ as $n \to \infty$ where $\{U_n\}$ is an exhaustion of Ω by sets U_n that are finite union of balls from a countable base and the convergence is pointwise in $\Omega \setminus \mathcal{N}$. Suppose for a moment that we have proved (5.13) for U_n instead of Ω, that is,

$$\begin{aligned} &\sup_{y \in B(x,\rho)\setminus\mathcal{N}} p_t^{U_n}(x, y) \\ &\quad \leq \inf_{y \in B(x,\rho)\setminus\mathcal{N}} p_t^{U_n}(x, y) + C \left(\frac{\rho}{\mathcal{R}(t)} \right)^{\Theta} \sup_{y \in B(x,r)\setminus\mathcal{N}} p_{t/2}^{U_n}(y, y) \end{aligned} \tag{5.14}$$

[note that if n is large enough, then $B(x, r) \Subset U_n$]. Replacing on the right-hand side of (5.14) $p_t^{U_n}$ by a larger value p_t^{Ω} and letting $n \to \infty$ on the left-hand side, we obtain (5.13).

To prove (5.14), rename for simplicity U_n into U and recall that, by construction in the proof of Theorem 2.12, the domain of p_t^U is $U \setminus \mathcal{N}_U$ where $\mathcal{N}_U$ is a truly exceptional set in U, that is contained in $\mathcal{N}$. It follows from Corollary 2.7, that, for any $x \in U \setminus \mathcal{N}$

$$\sup_{y \in B(x,\rho)\setminus\mathcal{N}} p_t^U(x,y) = \operatorname*{esup}_{y \in B(x,\rho)} p_t^U(x,y)$$

and a similar identity for inf and einf. Hence, it suffices to prove that

$$\operatorname*{eosc}_{y \in B(x,\rho)} p_t^U(x,y) \le C\left(\frac{\rho}{\mathcal{R}(t)}\right)^{\Theta} A,$$

where

$$A = \sup_{y \in B(x,r)\setminus\mathcal{N}} p_{t/2}^U(y,y).$$

Set

$$u(y) = p_t^U(x,y) \quad \text{and} \quad f(y) = \partial_t p_t^U(x,y),$$

where ∂_t is the strong derivative in $L^2(U)$ with respect to the variable y. Applying Corollary 5.7 to the heat kernel p_t^U, we obtain, for μ-a.a. $y \in B(x,r)$,

$$|f(y)| \le \frac{2}{t}\sqrt{p_{t/2}^U(x,x)\,p_{t/2}^U(y,y)} \le \frac{2}{t}A.$$

By Lemma 5.9, we have $u = -G^U f$. Since for all $y \in B(x,r) \setminus \mathcal{N}$

$$u(y) \le \sqrt{p_{t/2}^U(x,x)\,p_{t/2}^U(y,y)} \le A$$

and $\rho \le r$, we obtain by Proposition 5.3 and ($E_F\le$) that

$$\begin{aligned}\operatorname*{eosc}_{B(x,\rho)} u &\le 2\widetilde{E}(x,r) \operatorname*{esup}_{B(x,r)} |f| + 4\left(\frac{\rho}{r}\right)^{\theta} \operatorname*{esup}_{B(x,r)} |u| \\ &\le C F(r)\frac{A}{t} + 4\left(\frac{\rho}{r}\right)^{\theta} A.\end{aligned}$$

Since $r \le \mathcal{R}(t)$, we have by (3.19)

$$F(r) \le C\left(\frac{r}{\mathcal{R}(t)}\right)^{\beta} F(\mathcal{R}(t)) = C\left(\frac{r}{\mathcal{R}(t)}\right)^{\beta} t,$$

whence it follows that

$$\operatorname*{eosc}_{B(x,\rho)} u \le C\left(\left(\frac{r}{\mathcal{R}(t)}\right)^{\beta} + \left(\frac{\rho}{r}\right)^{\theta}\right)A.$$

Note that this inequality is true for any r such that $B(x,r) \Subset U$ and $\rho \le r \le \mathcal{R}(t)$. Choosing $r = (\mathcal{R}(t)^{\beta}\rho^{\theta})^{1/(\beta+\theta)}$, we obtain (5.13). $\square$

THEOREM 5.11. *Let the hypotheses* $(VD)+(H)+(E_F)$ *be satisfied. Then, for any open set* $\Omega \subset M$, *the heat kernel* $p_t^{\Omega}(x,y)$ *is Hölder continuous in* x *and* y *in* $\Omega \setminus \mathcal{N}$.

PROOF. Fix $x \in \Omega \setminus \mathcal{N}$, $t>0$, and choose $\rho>0$ so small that $B(x,r) \Subset \Omega$ where $r = r(t,\rho)$ is defined by (5.12). Using Theorem 4.2, (VD), and (3.19), we obtain that, for any $y \in B(x,r)\setminus \mathcal{N}$,

$$\begin{aligned} p_{t/2}^{\Omega}(y,y) \le p_{t/2}(y,y) &\le \frac{C}{V(y,\mathcal{R}(t))} \\ &= \frac{C}{V(x,\mathcal{R}(t))}\frac{V(x,\mathcal{R}(t))}{V(y,\mathcal{R}(t))} \\ &\le \frac{C}{V(x,\mathcal{R}(t))}\left(1+\frac{d(x,y)}{\mathcal{R}(t)}\right)^{\alpha} \\ &\le \frac{C}{V(x,\mathcal{R}(t))}, \end{aligned}$$

where we have used that $d(x,y) < r \le \mathcal{R}(t)$. Therefore, by Lemma 5.10,

$$\operatorname*{osc}_{y\in B(x,\rho)\setminus\mathcal{N}} p_t^{\Omega}(x,y) \le \left(\frac{\rho}{\mathcal{R}(t)}\right)^{\Theta} \frac{C}{V(x,\mathcal{R}(t))}. \tag{5.15}$$

In particular, if $y \in \Omega\setminus\mathcal{N}$ is close enough to x, then we have

$$|p_t^{\Omega}(x,x) - p_t^{\Omega}(x,y)| \le \left(\frac{d(x,y)}{\mathcal{R}(t)}\right)^{\Theta} \frac{C}{V(x,\mathcal{R}(t))},$$

which means that $p_t^{\Omega}(x,\cdot)$ is Hölder continuous in $\Omega\setminus\mathcal{N}$. $\square$

COROLLARY 5.12. *Let the hypotheses* $(VD)+(H)+(E_F)$ *be satisfied, and let* $B(x,R)$ *be a precompact ball, such that* $x \in M\setminus\mathcal{N}$. *Then for all* ρ *and* t, *such that*

$$0 < \rho \le \mathcal{R}(t) < R, \tag{5.16}$$

and for all $y \in B(x,\rho)\setminus\mathcal{N}$, *the following estimate holds*:

$$\left|p_t^{B(x,R)}(x,x) - p_t^{B(x,R)}(x,y)\right| \le \left(\frac{\rho}{\mathcal{R}(t)}\right)^{\Theta} \frac{C}{V(x,\mathcal{R}(t))}. \tag{5.17}$$

PROOF. Set $\Omega = B(x,R)$. Then the condition $B(x,r) \Subset \Omega$ from the previous proof is satisfied because $r \le \mathcal{R}(t) < R$ by (5.12) and (5.16). Hence, (5.17) follows from (5.15). $\square$

5.4. *Proof of the lower bounds.*

LEMMA 5.13. *Assume that* (VD) + (E_F) *are satisfied. Then there exists* $\varepsilon > 0$ *such that, for all precompact balls* $B(x, R)$ *with* $x \in M \setminus \mathcal{N}$ *and for all* $0 < t \leq \varepsilon F(R)$,

$$p_t^{B(x,R)}(x, x) \geq \frac{c}{V(x, \mathcal{R}(t))}. \tag{5.18}$$

PROOF. Choose r from the condition $t = \varepsilon F(r)$ which implies $R \geq r$ and, hence, $p_t^{B(x,R)} \geq p_t^{B(x,r)}$. Hence, it suffices to prove that

$$p_t^{B(x,r)}(x, x) \geq \frac{c}{V(x, \mathcal{R}(t))}.$$

Setting $B = B(x, r)$, we have by (2.17)

$$\int_{B\setminus\mathcal{N}} p_t^B(x, y)\, d\mu(y) = \mathcal{P}_t^B 1 = \mathbb{P}_x(t < \tau_B) = 1 - \mathbb{P}_x(\tau_B \leq t).$$

By Corollary 3.20, we obtain

$$\mathbb{P}_x(\tau_B \leq t) \leq C \exp\left(-c\left(\frac{F(r)}{t}\right)^{1/(\beta'-1)}\right) = C \exp\bigl(-c\varepsilon^{-1/(\beta'-1)}\bigr),$$

whence it follows that, for small enough $\varepsilon \in (0, 1)$,

$$\int_B p_t^B(x, y)\, d\mu(y) \geq \frac{1}{2}.$$

Therefore,

$$\begin{aligned} p_{2t}^B(x, x) &= \int_{B\setminus\mathcal{N}} p_t^B(x, y)^2\, d\mu(y) \\ &\geq \frac{1}{\mu(B)}\left(\int_{B\setminus\mathcal{N}} p_t^B(x, y)\, d\mu(y)\right)^2 \\ &\geq \frac{1/4}{V(x, \mathcal{R}(t/\varepsilon))}, \end{aligned}$$

where we have used that $r = \mathcal{R}(t/\varepsilon)$. Finally, using (3.20) and (3.13) we obtain (5.18). □

THEOREM 5.14. *Assume that the hypotheses* (VD) + (H) + (E_F) *are satisfied, and all metric balls in* (M, d) *are precompact. Then there exist* $\varepsilon, \eta > 0$ *such that*

$$p_t^{B(x,R)}(x, y) \geq \frac{c}{V(x, \mathcal{R}(t))} \tag{5.19}$$

for all $R > 0$, $0 < t \leq \varepsilon F(R)$ *and* $x, y \in M \setminus \mathcal{N}$, *provided*

$$d(x, y) \leq \eta \mathcal{R}(t) \tag{5.20}$$

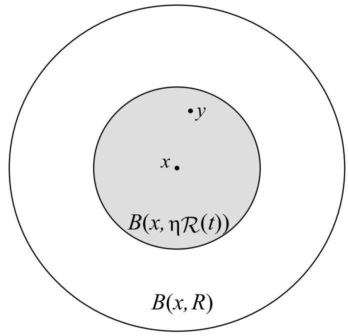

FIG. 6. *Illustration to Theorem* 5.14.

(*see Figure* 6). *Consequently*, *the following inequality*:

(*NLE*)
$$p_t(x, y) \geq \frac{c}{V(x, \mathcal{R}(t))}$$

holds for all $t > 0$ *and* $x, y \in M \setminus \mathcal{N}$ *satisfying* (5.20).

PROOF. Obviously, (*NLE*) follows from (5.19) by letting $R \to \infty$, so that it suffices to prove (5.19). Let ρ and t be such that

$$0 < \rho \leq \eta \mathcal{R}(t) \quad \text{and} \quad t \leq \varepsilon F(R),$$

where $\varepsilon \in (0, 1)$ is the constant from Lemma 5.13, and $\eta \in (0, 1)$ is to be defined below. Then the hypotheses of Lemma 5.13 and Corollary 5.12 are satisfied. Writing for simplicity $B = B(x, R)$, we obtain by (5.19) and (5.17) that, for any $y \in B(x, \rho) \setminus \mathcal{N}$,

$$\begin{aligned} p_t^B(x, y) &\geq p_t^B(x, x) - |p_t^B(x, x) - p_t^B(x, y)| \\ &\geq \frac{c}{V(x, \mathcal{R}(t))} - \left(\frac{\rho}{\mathcal{R}(t)}\right)^{\Theta} \frac{C}{V(x, \mathcal{R}(t))} \\ &\geq \frac{c - C\eta^{\Theta}}{V(x, \mathcal{R}(t))}. \end{aligned}$$

Choosing η sufficiently small, we obtain (5.19). $\square$

Combining Theorems 4.2, 5.11 and 5.14, we obtain the main result:

THEOREM 5.15. *If the hypotheses* (*VD*) + (*H*) + (E_F) *are satisfied and all metric balls are precompact*, *then the heat kernel exists*, *is Hölder continuous in* x, y, *and satisfies* (*UE*) *and* (*NLE*).

EXAMPLE 5.16. Under the hypotheses of Theorem 5.15, assume that the volume function $V(x,r)$ satisfies the uniform estimate

$$V(x,r) \simeq r^{\alpha}$$

with some $\alpha > 0$, and function F be as follows:

$$F(r) = \begin{cases} r^{\beta_1}, & r < 1, \\ r^{\beta_2}, & r \geq 1, \end{cases} \tag{5.21}$$

where $\beta_1 > \beta_2 > 1$. Then

$$\mathcal{R}(t) = \begin{cases} t^{1/\beta_1}, & t < 1, \\ t^{1/\beta_2}, & t \geq 1, \end{cases}$$

and the heat kernel satisfies the estimate

$$p_t(x,y) \leq \frac{C}{V(x,\mathcal{R}(t))} \simeq C \begin{cases} t^{-\alpha/\beta_1}, & t < 1, \\ t^{-\alpha/\beta_2}, & t \geq 1. \end{cases}$$

It follows that

$$p_t(x,y) \leq C t^{-\alpha/\beta} \tag{5.22}$$

for any β from the interval $\beta_2 < \beta < \beta_1$. Let us verify that the following upper bound fails:

$$p_t(x,y) \leq C t^{-\alpha/\beta} \exp\left(-\left(\frac{r^{\beta}}{Ct}\right)^{1/(\beta-1)}\right), \tag{5.23}$$

where $r = d(x,y)$. Indeed, by (*NLE*) we have

$$p_t(x,y) \geq \frac{c}{V(x,\mathcal{R}(t))}$$

provided $r \leq \eta \mathcal{R}(t)$. Assuming that $t < 1$ and setting $r = \eta\mathcal{R}(t) = \eta t^{1/\beta_1}$ we obtain[11]

$$p_t(x,y) \geq \frac{c}{t^{\alpha/\beta_1}}, \tag{5.24}$$

while it follows from (5.23) that

$$p_t(x,y) \leq \frac{C}{t^{\alpha/\beta}} \exp\bigl(-c(t^{\beta/\beta_1 - 1})^{1/(\beta-1)}\bigr). \tag{5.25}$$

Since $\beta/\beta_1 < 1$, the exponent of t under the exponential is negative so that the right-hand side of (5.25) becomes as $t \to 0$ much smaller than that of (5.24), which is a contradiction.

[11]The existence of a couple x, y with a prescribed distance $r = d(x,y)$ can be guaranteed, provided the space (M,d) is connected.

Another way to see a contradiction is to observe that (5.23) implies (E_F) with function $F(r) \simeq r^\beta$ (cf. [31, 45]), which is incompatible with (E_F) with function (5.21) [although this argument requires the conservativeness of $(\mathcal{E}, \mathcal{F})$].

The conclusion is that in general (5.22) does not imply (5.23). For comparison, let us note that if $\beta = 2$ and the underlying space is a Riemannian manifold, then (5.22) does imply (5.23); cf. [24].

6. Matching upper and lower bounds.

6.1. *Distance d_ε.*

DEFINITION 6.1. We say that a sequence $\{x_i\}_{i=0}^N$ of points in M is an ε-*chain* between points $x, y \in M$ if

$$x_0 = x, \qquad x_N = y \quad \text{and} \quad d(x_i, x_{i-1}) < \varepsilon \qquad \text{for all } i = 1, 2, \dots, N.$$

One can view an ε-chain as a sequence of *chained* balls $\{B_i\}_{i=0}^N$ of radii ε, that connect x and y; that is, the center of B_0 is x, the center of B_N is y, and the center of B_i is contained in B_{i-1} for any $i = 1, \dots, N$ (see Figure 7).

DEFINITION 6.2. For any $\varepsilon > 0$ and all $x, y \in M$, define

$$d_\varepsilon(x, y) = \inf_{\{x_i\} \text{ is } \varepsilon\text{-chain}} \sum_{i=1}^N d(x_i, x_{i-1}), \tag{6.1}$$

where the infimum is taken over all ε-chains $\{x_i\}_{i=0}^N$ between x, y with arbitrary N.

It is obvious that $d_\varepsilon(x, y)$ is a decreasing left-continuous function of ε and

$$d_\varepsilon(x, y) \geq d(x, y). \tag{6.2}$$

Furthermore,

$$\varepsilon > d(x, y) \quad \Rightarrow \quad d_\varepsilon(x, y) = d(x, y). \tag{6.3}$$

It is clear that d_ε is an extended metric in the sense that d_ε satisfies all the axioms of a metric except for finiteness. If an ε-chain exists for any two points x, y, then $d_\varepsilon(x, y) < \infty$, and hence d_ε is a true metric.

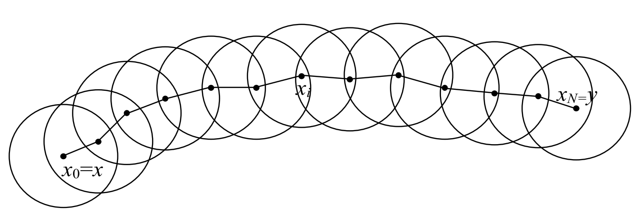

FIG. 7. *An ε-chain connecting x and y.*

LEMMA 6.3. *If* $0 < d_\varepsilon(x, y) < \infty$ *for some* $x, y \in M$ *and* $\varepsilon > 0$, *then there exists an* ε*-chain* $\{x_i\}_{i=0}^N$ *between* x, y *such that*

$$N \le 9\left\lceil \frac{d_\varepsilon(x, y)}{\varepsilon} \right\rceil. \tag{6.4}$$

Here $\lceil\cdot\rceil$ stands for the least integer upper bound of the argument. It follows from (6.1) by the triangle inequality that always

$$N \ge \left\lceil \frac{d_\varepsilon(x, y)}{\varepsilon} \right\rceil.$$

Hence, denoting by $N_\varepsilon(x, y)$ the minimal value of N for which there exists an ε chain $\{x_i\}_{i=0}^N$ between x and y, we obtain

$$N_\varepsilon(x, y) \simeq \left\lceil \frac{d_\varepsilon(x, y)}{\varepsilon} \right\rceil. \tag{6.5}$$

The number $N_\varepsilon(x, y)$ can be also viewed as the minimal number in a sequence of chained balls of radii ε connecting x and y.

PROOF OF LEMMA 6.3. If $d_\varepsilon(x, y) < \varepsilon$, then also $d(x, y) < \varepsilon$, and hence $\{x, y\}$ is an ε-chain with $N = 1$. Assume $d_\varepsilon(x, y) \ge \varepsilon$, and let $\{x_i\}_{i=0}^n$ be a ε-chain between x, y, such that

$$\sum_{i=1}^n d(x_i, x_{i-1}) \le 2d_\varepsilon(x, y), \tag{6.6}$$

which exists by hypothesis. Set $r_i = d(x_i, x_{i-1})$. Then (6.6) implies

$$\#\{i : r_i \ge \varepsilon/2\} \le \frac{4d_\varepsilon(x, y)}{\varepsilon},$$

whence

$$\#\{i : r_i < \varepsilon/2\} \ge n - \frac{4d_\varepsilon(x, y)}{\varepsilon}.$$

If $n > 9\lceil\frac{d_\varepsilon(x,y)}{\varepsilon}\rceil$, then $n > 9$ and $n > 9\frac{d_\varepsilon(x,y)}{\varepsilon}$, whence it follows that

$$\#\{i : r_i < \varepsilon/2\} > \frac{5n}{9} > \frac{n+1}{2}.$$

Hence, there is an index i such that both r_{i-1} and r_i are smaller than $\varepsilon/2$. This implies that $d(x_{i-1}, x_{i+1}) < \varepsilon$ so that by removing the point x_i from the chain we still have an ε-chain. Continuing this way, we finally obtain an ε-chain satisfying (6.4). $\square$

6.2. *Two-sided estimates of the heat kernel.* If $x \neq y$, then it follows from (6.2) and (3.19) that

$$\frac{F(\varepsilon)}{\varepsilon} d_\varepsilon(x, y) \to \infty \qquad \text{as } \varepsilon \to \infty. \tag{6.7}$$

In this section, we make an additional assumption that, for all $x, y \in M$,

$$\frac{F(\varepsilon)}{\varepsilon} d_\varepsilon(x, y) \to 0 \qquad \text{as } \varepsilon \to 0. \tag{6.8}$$

In particular, (6.8) implies the finiteness of d_ε for all $\varepsilon > 0$. Define the function $\varepsilon(t, x, y)$ as follows:

$$\varepsilon(t, x, y) = \sup\left\{\varepsilon > 0 : \frac{F(\varepsilon)}{\varepsilon} d_\varepsilon(x, y) \leq t\right\}. \tag{6.9}$$

If $x = y$, then $\varepsilon(t, x, x) = \infty$. If $x \neq y$, then it follows from (6.8) and (6.7) that $0 < \varepsilon(t, x, y) < \infty$.

LEMMA 6.4. *If* (6.8) *is satisfied, then the function* $\varepsilon(t, x, y)$ *satisfies the identity*

$$\frac{F(\varepsilon)}{\varepsilon} d_\varepsilon(x, y) = t \tag{6.10}$$

for all distinct $x, y \in M$ *and* $t > 0$.

PROOF. Since the function $F(\varepsilon)$ is continuous and $d_\varepsilon(x, y)$ is left-continuous in ε, we have

$$\frac{F(\varepsilon)}{\varepsilon} d_\varepsilon(x, y) \leq t.$$

Assume from the contrary that

$$\frac{F(\varepsilon)}{\varepsilon} d_\varepsilon(x, y) < t,$$

and note that, for any $\varepsilon' > \varepsilon$, we have by definition of ε that

$$\frac{F(\varepsilon')}{\varepsilon'} d_{\varepsilon'}(x, y) > t. \tag{6.11}$$

On the other hand, $d_{\varepsilon'}(x, y) \leq d_\varepsilon(x, y)$ and

$$\frac{F(\varepsilon')}{\varepsilon'} \to \frac{F(\varepsilon)}{\varepsilon} \qquad \text{as } \varepsilon' \to \varepsilon + .$$

Hence,

$$\limsup_{\varepsilon' \to \varepsilon+} \frac{F(\varepsilon')}{\varepsilon'} \leq \frac{F(\varepsilon)}{\varepsilon} d_\varepsilon(x, y) < t,$$

which contradicts (6.11). $\square$

THEOREM 6.5. *Let all metric balls be precompact. Let the hypotheses* (VD)$+$ (E_F) $+$ (H) *and* (6.8) *be satisfied, and let* $\varepsilon(t,x,y)$ *be the function from* (6.9). *Then, for all* $x, y \in M \setminus \mathcal{N}$ *and* $t > 0$, *we have*

$$p_t(x,y) \asymp \frac{C}{V(x,\mathcal{R}(t))} \exp(-c\Phi(cd_\varepsilon(x,y),t)), \tag{6.12}$$

where $\varepsilon = \varepsilon(\kappa t, x, y)$ *and* $\kappa = 8$ *for the upper bound in* (6.12) *while* κ *is a small enough positive constant for the lower bound.*

The proof of Theorem 6.5 is preceded by a lemma.

LEMMA 6.6. *For all distinct* $x, y \in M$ *and* $t > 0$, *we have*

$$\Phi(cd_\varepsilon(x,y),t) \le \frac{d_\varepsilon(x,y)}{\varepsilon} \le \Phi(Cd_\varepsilon(x,y),t), \tag{6.13}$$

where $\varepsilon = \varepsilon(t,x,y)$.

PROOF. Let us first show that, for all $\varepsilon > 0$ and some $c \in (0,1)$,

$$\Phi\left(c\frac{\varepsilon}{F(\varepsilon)}\right) \le \frac{1}{F(\varepsilon)} \le \Phi\left(2\frac{\varepsilon}{F(\varepsilon)}\right). \tag{6.14}$$

By (3.39), we have, for all $r > 0$,

$$\Phi\left(\frac{2\varepsilon}{F(\varepsilon)}\right) \ge \frac{2\varepsilon}{F(\varepsilon)r} - \frac{1}{F(r)}.$$

Choosing $r = \varepsilon$ we obtain the right-hand side inequality in (6.14). By (3.39), the left-hand side inequality in (6.14) is equivalent to

$$\frac{c\varepsilon}{F(\varepsilon)r} - \frac{1}{F(r)} \le \frac{1}{F(\varepsilon)} \qquad \text{for all } r > 0,$$

that is, to

$$\frac{F(\varepsilon)}{F(r)} \ge \frac{c\varepsilon}{r} - 1.$$

If $r \ge \varepsilon$, then this is trivially satisfied provided $c \le 1$. If $r < \varepsilon$, then by (3.19) we have

$$\frac{F(\varepsilon)}{F(r)} \ge c\left(\frac{\varepsilon}{r}\right)^\beta \ge c\frac{\varepsilon}{r},$$

which proves the previous inequality and, hence, (6.14).

Putting in (6.14) $\varepsilon = \varepsilon(t,x,y)$ and using $\frac{\varepsilon}{F(\varepsilon)} = \frac{d_\varepsilon(x,y)}{t}$, which is true by Lemma 6.4, we obtain

$$\Phi\left(c\frac{d_\varepsilon(x,y)}{t}\right) \le \frac{d_\varepsilon(x,y)}{\varepsilon t} \le \Phi\left(2\frac{d_\varepsilon(x,y)}{t}\right),$$

whence (6.13) follows. $\square$

PROOF OF THEOREM 6.5. If $x = y$, then $d_\varepsilon(x, y) = 0$, and (6.12) follows from Theorems 4.2 and 5.14. Assume in the sequel that $x \neq y$. Let us first prove the lower bound in (6.12), that is,

$$p_t(x, y) \geq \frac{c}{V(x, \mathcal{R}(t))} \exp(-C\Phi(Cd_\varepsilon(x, y), t)). \tag{6.15}$$

By Theorem 5.14, we have

$$p_t(x, y) \geq \frac{c}{V(x, \mathcal{R}(t))} \tag{NLE}$$

provided

$$d(x, y) \leq \eta \mathcal{R}(t) \tag{6.16}$$

for some $\eta > 0$. Set $\varepsilon = \varepsilon(\kappa t, x, y)$ where $\kappa \in (0, 1)$ will be chosen later.

Consider first the case $\varepsilon \geq d_\varepsilon(x, y)$. By (6.13), we have

$$\Phi(cd_\varepsilon(x, y), \kappa t) \leq 1.$$

Applying (3.45) with $R = cd_\varepsilon(x, y)$, we obtain

$$F(cd_\varepsilon(x, y)) \leq C\kappa t,$$

whence by (3.20)

$$d_\varepsilon(x, y) \leq c^{-1}\mathcal{R}(C\kappa t) \leq \eta \mathcal{R}(t),$$

provided κ is small enough. Since $d(x, y) \leq d_\varepsilon(x, y)$, we see that the condition (6.16) is satisfied and, hence, (6.15) follows from (*NLE*).

Assume now that $\varepsilon < d_\varepsilon(x, y)$. By Lemma 6.3, there is an ε-chain $\{x_i\}_{i=1}^N$ connecting x and y and such that

$$N \simeq \frac{d_\varepsilon(x, y)}{\varepsilon}. \tag{6.17}$$

By (2.6), we have

$$\begin{aligned} p_t(x, y) &= \int_M \cdots \int_M p_{t/N}(x, z_1) p_{t/N}(z_1, z_2) \cdots \\ &\qquad \times p_{t/N}(z_{N-1}, y)\, dz_1 \cdots dz_{N-1} \\ &\geq \int_{B_1} \cdots \int_{B_{N-1}} p_{t/N}(z_0, z_1) p_{t/N}(z_1, z_2) \cdots \\ &\qquad \times p_{t/N}(z_{N-1}, z_N)\, dz_1 \cdots dz_{N-1}, \end{aligned} \tag{6.18}$$

where $z_0 = x$, $z_N = y$, $B_i = B(x_i, \varepsilon)$. We will estimate $p_{t/N}(z_i, z_{i+1})$ from below by means of (*NLE*). For that, we need to verify the condition

$$d(z_i, z_{i+1}) \leq \eta \mathcal{R}(t/N).$$

By (6.10), we have

$$\frac{F(\varepsilon)}{\varepsilon} = \frac{\kappa t}{d_\varepsilon(x, y)}. \tag{6.19}$$

It follows from (6.19) and (6.17) that

$$F(\varepsilon) \simeq \frac{\kappa t}{N},$$

whence by (3.20)

$$\varepsilon \simeq \mathcal{R}\left(\frac{\kappa t}{N}\right).$$

Clearly, if κ is small enough, then

$$\varepsilon \leq \frac{\eta}{3}\mathcal{R}\left(\frac{t}{N}\right). \tag{6.20}$$

Since in (6.18) $z_i \in B(x_i, \varepsilon)$ and $d(x_i, x_{i+1}) \leq \varepsilon$, it follows from (6.20) that

$$d(z_i, z_{i+1}) \leq 3\varepsilon \leq \eta\mathcal{R}(t/N).$$

Hence, by (*NLE*) and (3.13),

$$p_{t/N}(z_i, z_{i+1}) \geq \frac{c}{V(z_i, \mathcal{R}(t/N))} \geq \frac{c}{V(x_i, \mathcal{R}(t/N))} \geq \frac{c}{V(x_i, \varepsilon)}.$$

Therefore, (6.18) implies

$$\begin{aligned} p_t(x, y) &\geq \frac{c}{V(x, \mathcal{R}(t/N))} \prod_{i=1}^{N-1} \frac{c}{V(x_i, \varepsilon)} V(x_i, \varepsilon) \\ &\geq \frac{c^{-N}}{V(x, \mathcal{R}(t/N))} \\ &\geq \frac{\exp(-CN)}{V(x, \mathcal{R}(t))} \\ &\geq \frac{\exp(-C(d_\varepsilon(x, y))/\varepsilon)}{V(x, \mathcal{R}(t))}. \end{aligned} \tag{6.21}$$

By Lemma 6.6, we have

$$\frac{d_\varepsilon(x, y)}{\varepsilon} \leq \Phi(C d_\varepsilon(x, y), \kappa t) = \kappa\Phi\left(\frac{C}{\kappa} d_\varepsilon(x, y), t\right).$$

Substituting into (6.21), we obtain (6.15).

To prove the upper bound in (6.12), we basically repeat the proof of Theorem 4.2 with d being replaced by d_ε for an appropriate ε. Fix some $\varepsilon > 0$ and denote by $B_\varepsilon(x, r)$ the ball in the metric d_ε. It follows from (6.3) that

$$B_\varepsilon(x, r) = B(x, r) \qquad \text{for all } r \leq \varepsilon, \tag{6.22}$$

which allows to modify Lemma 3.14 as follows: for all $x \in M \setminus \mathcal{N}_0$ and $R > 0$

$$\mathbb{E}_x\left(e^{-\lambda \tau_{B_\varepsilon(x,R)}}\right) \le C \exp\left(-c\frac{R}{r}\right), \tag{6.23}$$

provided the values of parameters r and λ satisfy the conditions

$$0 < r \le \varepsilon \quad \text{and} \quad \lambda \ge \frac{\sigma}{F(r)}. \tag{6.24}$$

Indeed, (6.23) is analogous to estimate (3.30) from the proof of Lemma 3.14 for d-balls, which was proved using $\lambda \ge \frac{\sigma}{F(r)}$. To repeat the proof for the metric d_ε, we need the precompactness of d_ε-balls, that follows from (6.2), and the condition (E_F) for d_ε-balls of radii $\le r$, that follows from (6.22), provided $r \le \varepsilon$.

Consequently, the statement of Theorem 3.15 is modified as follows: for all $x \in M \setminus \mathcal{N}_0$ and $R, t > 0$,

$$\mathbb{P}_x\left(\tau_{B_\varepsilon(x,R)} \le t\right) \le C \exp(-c\Phi_\varepsilon(cR, t)), \tag{6.25}$$

where

$$\Phi_\varepsilon(R, t) := \sup_{0<r\le\varepsilon} \left\{\frac{R}{r} - \frac{t}{F(r)}\right\}. \tag{6.26}$$

Indeed, arguing as in (3.36) and using (6.23) we obtain that, under the assumptions of (6.24),

$$\mathbb{P}_x\left(\tau_{B_\varepsilon(x,R)} \le t\right) \le C \exp\left(-c\frac{R}{r} + \lambda t\right).$$

Setting here $\lambda = \sigma/F(r)$ yields

$$\mathbb{P}_x\left(\tau_{B_\varepsilon(x,R)} \le t\right) \le C \exp\left(-\left(c\frac{R}{r} - \frac{\sigma t}{F(r)}\right)\right). \tag{6.27}$$

Finally, minimizing the right-hand side of (6.27) in $r \le \varepsilon$, we obtain (6.25).

Let us show that if

$$t \le \frac{1}{2}\frac{F(\varepsilon)}{\varepsilon} R, \tag{6.28}$$

then

$$\Phi(R, t) \le 2\Phi_\varepsilon(R, t). \tag{6.29}$$

We have

$$\sup_{r>\varepsilon}\left\{\frac{R}{r} - \frac{t}{F(r)}\right\} \le \frac{R}{\varepsilon},$$

whereas

$$\sup_{0<r\le\varepsilon}\left\{\frac{R}{r} - \frac{t}{F(r)}\right\} \ge \frac{R}{\varepsilon} - \frac{t}{F(\varepsilon)} \ge \frac{R}{\varepsilon} - \frac{1}{2}\frac{R}{\varepsilon} = \frac{1}{2}\frac{R}{\varepsilon}.$$

It follows that

$$\Phi(R,t)=\sup_{r>0}\left\{\frac{R}{r}-\frac{t}{F(r)}\right\}\le 2\sup_{0<r\le\varepsilon}\left\{\frac{R}{r}-\frac{t}{F(r)}\right\},$$

which proves (6.29). Hence, we can rewrite (6.25) in the form

$$\mathbb{P}_x\big(\tau_{B_\varepsilon(x,R)}\le t\big)\le C\exp(-c\Phi(cR,t)), \tag{6.30}$$

provided the relation (6.28) between ε, t, R is satisfied.

As in the last part of the proof of Theorem 4.2, we apply (6.30) with $R=\frac{1}{4}d_\varepsilon(x,y)$ for fixed $x,y\in M\setminus\mathcal{N}$. Note that in (4.20) $d(x,z)$ can be replaced by a larger value $d_\varepsilon(x,z)$. The rest of the argument goes through unchanged, and we obtain

$$p_t(x,y)\le\frac{C}{V(x,\mathcal{R}(t))}\exp(-c\Phi(cd_\varepsilon(x,y),t)), \tag{6.31}$$

provided

$$t\le\frac{1}{8}\frac{F(\varepsilon)}{\varepsilon}d_\varepsilon(x,y). \tag{6.32}$$

By (6.10), condition (6.32) can be satisfied by setting $\varepsilon=\varepsilon(8t,x,y)$. $\square$

COROLLARY 6.7. *Under the hypotheses of Theorem 6.5, we have*

$$p_t(x,y)\asymp\frac{C}{V(x,\mathcal{R}(t))}\exp\left(-c\frac{d_\varepsilon(x,y)}{\varepsilon}\right) \tag{6.33}$$

$$\asymp\frac{C}{V(x,\mathcal{R}(t))}\exp(-cN_\varepsilon(x,y)), \tag{6.34}$$

where $\varepsilon=\varepsilon(\kappa t,x,y)$ and κ is a large enough constant for the upper bound and a small enough positive constant for the lower bound.

PROOF. The lower bound in (6.33) follows from (6.21), while the upper bound follows from (6.31) and

$$\frac{d_\varepsilon(x,y)}{\varepsilon}\le\Phi(Cd_\varepsilon(x,y),\kappa t)=\kappa\Phi\left(\frac{C}{\kappa}d_\varepsilon(x,y),t\right),$$

provided κ is chosen large enough to ensure $C/\kappa\le c$ where c is the constant from (6.31). Estimate (6.34) follows then from (6.5). $\square$

REMARK 6.8. A good example to illustrate Theorem 6.5 and Corollary 6.7 is the class of post critically finite (p.c.f.) fractals, where the heat kernel estimate (6.34) was proved by Hambly and Kumagai [40]. Without going into the details of [40], let us mention that $d(x,y)$ is the resistance metric on such a fractal M, and μ is the Hausdorff measure of M of dimension $\alpha:=\dim_H M$. One has in this setting $V(x,r) \simeq r^\alpha$, in particular, (VD) is satisfied. Hambly and Kumagai proved that (E_F) is satisfied with $F(r) = r^\beta$ where $\beta = \alpha + 1$; cf. [40], Theorem 3.8. Condition (6.8) follows from the estimate

$$N_\varepsilon(x,y) \le C\left(\frac{d(x,y)}{\varepsilon}\right)^\gamma, \tag{6.35}$$

proved in [40], Lemma 3.3, with $\gamma = \beta/2$, as (6.35) implies that

$$\frac{F(\varepsilon)}{\varepsilon} d_\varepsilon(x,y) \le C\varepsilon^\beta N_\varepsilon(x,y) \le C d(x,y)^\gamma \varepsilon^{\beta-\gamma} \to 0 \qquad \text{as } \varepsilon \to 0.$$

The Harnack inequality (H) on p.c.f. fractals was proved by Kigami [44]. Hence, Corollary 6.7 applies and gives on unbounded p.c.f. fractals estimate (6.34). The same estimate was proved in [40], Theorem 1.1, for bounded p.c.f. fractals using a different method.

Note that (VD) implies (6.35) with $\gamma = \alpha$ [where α comes from (3.13)], provided all balls in M are connected. Indeed, (VD) implies by the classical ball covering argument that any ball of radius r can be covered by at most $C(\frac{r}{\varepsilon})^\alpha$ balls of radii $\varepsilon \in (0,r)$. Consequently, any point $y \in B(x,r)$ can be connected to x by a chain of ε-balls containing at most $C(\frac{r}{\varepsilon})^\alpha$ balls. Taking $r \simeq d(x,y)$ we obtain (6.35) with $\gamma = \alpha$. Therefore, hypothesis (6.8) is satisfied automatically for $F(r) = r^\beta$ with $\beta > \alpha$.

Estimate (6.34) means that the diffusion process goes from x to y in time t in the following way. The process first "computes" the value[12] of ε as a function of t, x, y, then "detects" a shortest chain of ε-balls connecting x and y and finally goes along that chain (see Figure 8).

This phenomenon was first observed by Hambly and Kumagai on p.c.f. fractals, but it seems to be generic. Hence, to obtain matching upper and lower bounds, one needs, in addition to the usual hypotheses, also the information encoded in the function $N_\varepsilon(x,y)$, namely, the graph distance between x and y on any ε-net approximation of M.

6.3. *Chain condition.* The statement of Theorem 6.5 can be simplified if the metric space (M,d) possesses an additional property as follows.

[12] For example, in the above setting, when (6.35) is satisfied with $\gamma < \beta$, we obtain from (6.10) $\varepsilon^\beta N_\varepsilon \simeq t$ whence

$$\varepsilon^\beta \left(\frac{d(x,y)}{\varepsilon}\right)^\gamma \ge ct$$

and

$$\varepsilon \ge c\left(\frac{t}{d(x,y)^\gamma}\right)^{1/(\beta-\gamma)}.$$

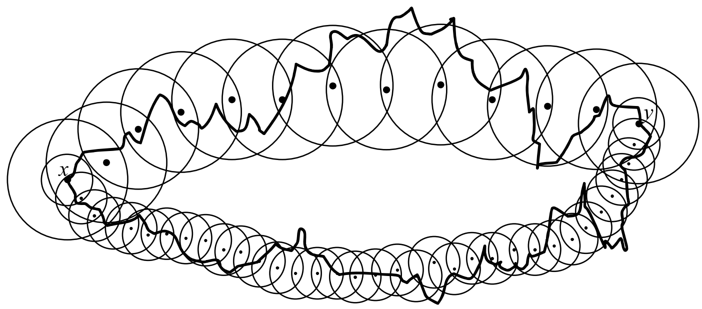

FIG. 8. *Two shortest chains of ε-balls for two distinct values of ε provide different routes for the diffusion from x to y for two distinct values of time t.*

DEFINITION 6.9. We say that a metric space (M, d) satisfies the *chain condition* if there exists a constant $C \geq 1$ such that, for any positive integer n and for all $x, y \in M$, there is a sequence $\{x_k\}_{k=0}^n$ of points in M such that $x_0 = x$, $x_n = y$ and

$$d(x_{k-1}, x_k) \leq C\frac{d(x, y)}{n} \qquad \text{for all } k = 1, \dots, n.$$

For example, any geodesic metric satisfies the chain condition.

LEMMA 6.10. *If (M, d) satisfies the chain condition, then $d_\varepsilon \leq Cd$ for any $\varepsilon > 0$.*

PROOF. Indeed, fix $\varepsilon > 0$ and two distinct points $x, y \in M$, and choose n so big that $C\frac{d(x,y)}{n} < \varepsilon$. Let $\{x_k\}_{k=0}^n$ be a sequence from the chain condition. Then it is also an ε-chain, whence

$$d_\varepsilon(x, y) \leq \sum_{k=1}^n d(x_{k-1}, x_k) \leq Cd(x, y),$$

which was to be proved. □

COROLLARY 6.11. *Let the metric space (M, d) satisfy the chain condition, and let all metric balls be precompact. If the hypotheses (VD) + (E_F) + (H) are satisfied, then, for all $x, y \in M \setminus \mathcal{N}$ and $t > 0$,*

$$p_t(x, y) \asymp \frac{C}{V(x, \mathcal{R}(t))} \exp\left(-ct\Phi\left(c\frac{d(x, y)}{t}\right)\right), \tag{6.36}$$

where $\Phi(s)$ is defined by (3.39).

PROOF. Since by Lemma 6.10 $d_\varepsilon \leq Cd$, condition (6.8) is obviously satisfied. Since $d_\varepsilon \simeq d$, we can replace in (6.12) d_ε by d, which together with (3.38) yields (6.36). □

REMARK 6.12. Obviously, estimate (6.36) (should it be true) implies (*UE*). We claim that (6.36) implies also (*NLE*); moreover, the parameter η in (*NLE*) can be chosen to be arbitrarily large, say $\eta > 1$. Indeed, we need to show that if

$$d(x, y) \le \eta \mathcal{R}(t),$$

where η is a (large) given constant, then

$$t\Phi\left(c\frac{d(x, y)}{t}\right) \le \text{const},$$

which amounts to

$$\Phi\left(\eta\frac{\mathcal{R}(t)}{t}\right) \le \frac{\text{const}}{t}, \tag{6.37}$$

where we have renamed $c\eta$ to η. Indeed, by (3.39) we have

$$\Phi(s) = \sup_{\xi>0}\left\{\frac{s}{\mathcal{R}(\xi)} - \frac{1}{\xi}\right\}$$

so that (6.37) is equivalent to

$$\frac{\eta\mathcal{R}(t)}{\mathcal{R}(\xi)} \le \frac{t}{\xi} + \text{const}. \tag{6.38}$$

If $\xi \le t$, then by (3.20)

$$\frac{\mathcal{R}(t)}{\mathcal{R}(\xi)} \le C\left(\frac{t}{\xi}\right)^{1/\beta}.$$

If the ratio $\frac{t}{\xi}$ is large enough then, using $1/\beta < 1$, we obtain that

$$\eta C\left(\frac{t}{\xi}\right)^{1/\beta} \le \frac{t}{\xi},$$

whence (6.38) follows. If $\frac{t}{\xi}$ is bounded by a constant, say $\frac{t}{\xi} \le C'$ (which includes also the case $\xi > t$), then by (3.20)

$$\frac{\eta\mathcal{R}(t)}{\mathcal{R}(\xi)} \le \eta\frac{\mathcal{R}(C'\xi)}{\mathcal{R}(\xi)} \le \text{const},$$

whence (6.37) follows again.

7. Consequences of heat kernel bounds.

7.1. *Harmonic function and the Dirichlet problem.* We assume only basic hypotheses in this subsection. Moreover, we use neither the locality of $(\mathcal{E}, \mathcal{F})$ nor the existence of the process $\{X_t\}$. We state and prove some basic properties of the Dirichlet problem in the abstract setting, that will be used in the proof of Theorem 7.4.

Fix an open set $\Omega \subset M$ such that $\lambda_{\min}(\Omega) > 0$, and consider the following weak Dirichlet problem in Ω: given a function $f \in \mathcal{F}$, find a function $u \in \mathcal{F}$ such that

$$\begin{cases} u \text{ is harmonic in } \Omega, \\ u = f \bmod \mathcal{F}(\Omega), \end{cases} \tag{7.1}$$

where the second condition is a weak boundary condition and means that $u - f \in \mathcal{F}(\Omega)$.

LEMMA 7.1. (a) *For any $f \in \mathcal{F}$, problem* (7.1) *has a unique solution u.*

(b) *If u solves* (7.1) *and $w \in \mathcal{F}$ is another function such that $w = f \bmod \mathcal{F}(\Omega)$, then $\mathcal{E}(u) \le \mathcal{E}(w)$. Moreover, the identity $\mathcal{E}(u) = \mathcal{E}(w)$ holds if and only if $u = w$.*

PROOF. (a) The condition $\lambda_{\min}(\Omega) > 0$ implies that

$$\mathcal{E}(\varphi) \simeq \mathcal{E}(\varphi) + \|\varphi\|_2^2$$

for all $\varphi \in \mathcal{F}(\Omega)$. Hence, $\mathcal{F}(\Omega)$ is a Hilbert space also with respect to the inner product $\mathcal{E}(\varphi, \psi)$. The harmonicity of u in (7.1) means that

$$\mathcal{E}(u, \varphi) = 0 \qquad \text{for all } \varphi \in \mathcal{F}(\Omega). \tag{7.2}$$

Equivalently, this means for the function $v = f - u$ that

$$\mathcal{E}(v, \varphi) = \mathcal{E}(f, \varphi) \qquad \text{for all } \varphi \in \mathcal{F}(\Omega). \tag{7.3}$$

Since $\mathcal{E}(f, \varphi) \le \mathcal{E}(f)^{1/2}\mathcal{E}(\varphi)^{1/2}$, the functional $\varphi \mapsto \mathcal{E}(f, \varphi)$ is a bounded linear functional in $\mathcal{F}(\Omega)$, and equation (7.3) has a unique solution $v \in \mathcal{F}(\Omega)$ by the Riesz representation theorem. Then $u = f - v$ is a unique solution of (7.1).

(b) Setting $\varphi = w - u$ and noticing that $\varphi \in \mathcal{F}(\Omega)$, we obtain using (7.2)

$$\mathcal{E}(w) = \mathcal{E}(u + \varphi) = \mathcal{E}(u) + 2\mathcal{E}(u, \varphi) + \mathcal{E}(\varphi) = \mathcal{E}(u) + \mathcal{E}(\varphi).$$

Hence, $\mathcal{E}(u) \le \mathcal{E}(w)$, and the equality is attained when $\mathcal{E}(\varphi) = 0$, that is, when $\varphi = 0$. $\square$

In what follows, denote by R the resolvent operator of (7.1), that is, $u = Rf$. Obviously, R is a linear operator in $\mathcal{F}$. Since by Lemma 7.1 $\mathcal{E}(Rf) \le \mathcal{E}(f)$, we see that the norm of the operator R in $\mathcal{F}$ is bounded by 1.

LEMMA 7.2. (a) *If $f \le g$, then $Rf \le Rg$. In particular, if $f \ge 0$, then $Rf \ge 0$.*

(b) *If $0 \le f \le 1$, then also $0 \le Rf \le 1$.*

(c) *If $\{f_n\}_{n=1}^{\infty}$ is an increasing sequence from $\mathcal{F}$ and $f_n \xrightarrow{\mathcal{F}} f$ as $n \to \infty$, then $Rf_n \to Rf$ a.e. in Ω as $n \to \infty$.*

PROOF. (a) The function $u = Rf - Rg$ is harmonic in B and satisfies the boundary condition $u \le 0 \bmod \mathcal{F}(\Omega)$. By [30], Lemma 4.4, the latter condition implies $u_+ \in \mathcal{F}(\Omega)$. Substituting $\varphi = u_+$ into (7.2), we obtain $\mathcal{E}(u, u_+) = 0$. On the other hand, by [30], Lemma 4.3, $\mathcal{E}(u, u_+) \ge \mathcal{E}(u_+)$, whence it follows that $\mathcal{E}(u_+) = 0$ and, hence, $u_+ = 0$. Consequently, $u \le 0$ and $Rf \le Rg$.

(b) Set $u = Rf$ and $w = u \wedge 1$ so that $u, w \in \mathcal{F}$ and $\mathcal{E}(w) \le \mathcal{E}(u)$. Setting $\varphi = u - f$ and $\psi = w - f$, we see that $\varphi \in \mathcal{F}(\Omega)$, $\psi \in \mathcal{F}$ and $\psi \le \varphi$. By [30], Lemma 4.4, we conclude that $\psi_+ \in \mathcal{F}(\Omega)$. On the other hand, we have $\psi_- = \varphi_- \in \mathcal{F}(\Omega)$ whence $\psi \in \mathcal{F}(\Omega)$. It follows that $w = f \bmod \mathcal{F}(\Omega)$. By Lemma 7.1 we conclude that $\mathcal{E}(u) \le \mathcal{E}(w)$. Since the opposite inequality is true by the definition of a Dirichlet form, we see that $\mathcal{E}(w) = \mathcal{E}(u)$. It follows from Lemma 7.1 that $w = u$, which implies $u \le 1$.

(c) Since R is a bounded operator in $\mathcal{F}$, we see that $Rf_n \overset{\mathcal{F}}{\to} Rf$ as $n \to \infty$. It follows that also $Rf_n \overset{L^2(\Omega)}{\to} Rf$. Then there is a subsequence of $\{Rf_n\}$ that converges to Rf almost everywhere in Ω. Finally, since the sequence $\{Rf_n\}$ is monotone increasing, the entire sequence $\{Rf_n\}$ also converges to Rf almost everywhere in Ω. $\square$

7.2. *Some consequences of the main hypotheses*. The next lemma states useful consequences of the main hypotheses and motivates the statement of Theorem 7.4 below. It is also used in the proof of Corollary 7.5.

LEMMA 7.3. *Let all metric balls be precompact. Then the following implications are true*:

(a) (H) *implies that the metric space* (M, d) *is connected*;
(b) (E_F) *implies that the Dirichlet form* $(\mathcal{E}, \mathcal{F})$ *is conservative*;
(c) (E_F) *implies that* $\operatorname{diam} M = \infty$.

PROOF. (a) Assume that (M, d) is disconnected, and let Ω be a nonempty open subset of M such that Ω^c is also nonempty and open. There is a big enough ball $B \subset M$ such that the intersections of δB both with Ω and Ω^c are nonempty, where δ is the parameter from (H). Since $\overline{B} \cap \Omega$ is a compact set, there is a cutoff function u of $\overline{B} \cap \Omega$ in Ω; that is, $u \in \mathcal{F} \cap C_0(\Omega)$ and $u \equiv 1$ in a neighborhood of $\overline{B} \cap \Omega$. Obviously, $u \equiv 0$ in Ω^c. We claim that u is harmonic in B. Indeed, for every function $v \in \mathcal{F} \cap C_0(B)$, we have $uv \in \mathcal{F} \cap C_0(B \cap \Omega)$ and

$$\mathcal{E}(u, v) = \mathcal{E}(u, uv) + \mathcal{E}(u, v - uv).$$

Since $\operatorname{supp}(uv) \subset \overline{B} \cap \Omega$ and, hence, $u \equiv 1$ in a neighborhood of $\operatorname{supp}(uv)$, we obtain by the strong locality of $(\mathcal{E}, \mathcal{F})$ that $\mathcal{E}(u, uv) = 0$. Since

$$\operatorname{supp}\bigl(v(1-u)\bigr) \subset \overline{B} \cap (\overline{B} \cap \Omega)^c = \overline{B} \cap \Omega^c$$

and $u = 0$ in Ω^c, it follows that $\mathcal{E}(u, v - uv) = 0$. Hence $\mathcal{E}(u, v) = 0$ and u is a nonnegative harmonic function in B. However, the function u does not satisfy (H) because u takes in δB the values 1 and 0.

(b) By Corollary 3.20, (E_F) implies that

$$\mathbb{P}_x\big(\tau_{B(x,R)} \le t\big) \le C \exp\biggl(-c\biggl(\frac{F(R)}{t}\biggr)^{1/(\beta'-1)}\biggr)$$

for any $x \in M \setminus \mathcal{N}_0$, $R > 0$, $t > 0$. Using this estimate and (2.17), we obtain

$$\begin{aligned}\mathcal{P}_t 1(x) &\ge \mathcal{P}_t^{B(x,R)} 1(x)\\ &= \mathbb{P}_x\big(\tau_{B(x,R)} > t\big)\\ &\ge 1 - C \exp\biggl(-c\biggl(\frac{F(R)}{t}\biggr)^{1/(\beta'-1)}\biggr).\end{aligned}$$

As $R \to \infty$, we see that $\mathcal{P}_t 1(x) \ge 1$, which proves the stochastic completeness.

(c) If $\operatorname{diam} M = R < \infty$, then $M = B_R$ so that the exit time from B_R is ∞ and ($E_F \le$) fails. $\square$

7.3. *The converse theorem.* In the next statement, we use weaker versions of (UE) and (NLE) that will be denoted by (UE_{weak}) and (NLE_{weak}). Namely, in each of these conditions we assume that the heat kernel exists as a measurable integral kernel of the heat semigroup $\{P_t\}$ and satisfies the estimates (UE) and (NLE) for all $t > 0$ and for *almost* all $x, y \in M$. Note that unlike the conditions (UE) and (NLE), their weak versions do not use the diffusion process $\{X_t\}$.

THEOREM 7.4. *Assume that all metric balls are precompact and* $\operatorname{diam} M = \infty$. *Then the following sets of conditions are equivalent*:

(i) $(VD) + (H) + (E_F)$;

(ii) $(VD) + (UE) + (NLE)$, *and the heat kernel is Hölder continuous outside a properly exceptional set*;

(iii) $(VD) + (UE_{\text{weak}}) + (NLE_{\text{weak}})$.

Note that, by Lemma 7.3, (i) implies that $\operatorname{diam} M = \infty$. However, neither of conditions (ii) or (iii) implies that M is unbounded because (ii) is satisfied on any compact Riemannian manifold.

PROOF OF THEOREM 7.4. The implication (i) $\Rightarrow$ (ii) is contained in Theorem 5.15, and the implication (ii) $\Rightarrow$ (iii) is trivial. In what follows we prove the implication (iii) $\Rightarrow$ (i).

Assuming (iii), let us first show that M is connected. Indeed, let M split into a disjoint union of two nonempty open sets Ω_1 and Ω_2. By the continuity of the paths of $\{X_t\}$, we have $p_t(x, y) = 0$ for all $t > 0$ and $x \in \Omega_1 \setminus \mathcal{N}$, $y \in \Omega_2 \setminus \mathcal{N}$, whereas by

(*NLE*) we have $p_t(x,y) > 0$ whenever $t > \eta^{-1}d(x,y)$. This contradictions proves the connectedness of M. By [31], Corollary 5.3, (*VD*), the connectedness, and the unboundedness of M imply the *reverse volume doubling* (*RVD*); that is, the following inequality holds:

$$(RVD) \qquad \frac{V(x,R)}{V(x,r)} \geq c\left(\frac{R}{r}\right)^{\alpha'},$$

which holds for all $x \in M$, $0 < r \leq R$, with some positive constants c, α'. By [31], Theorem 2.2 and Section 6.4 (see also [45]), (*VD*) + (*RVD*) + (UE_{weak}) imply ($E_F \leq$).[13]

Let us now prove ($E_F \geq$), that is,

$$\int_0^\infty \mathcal{P}_t^{B(x,R)} 1(x)\,dt \geq cF(R) \tag{7.4}$$

for all $x \in M \setminus \mathcal{N}$ and $R > 0$, where $\mathcal{N}$ is a properly exceptional set. It suffices to prove that there is a constant $\zeta > 0$ such that, for any ball $B = B(x_0, R)$,

$$\int_0^\infty P_t^B 1\,dt \geq cF(R) \qquad \text{a.e. in } \zeta B. \tag{7.5}$$

Indeed, the function

$$u = \int_0^\infty \mathcal{P}_t^B 1\,dt = G^B 1$$

is quasi-continuous by [20], Theorem 4.2.3. By [31], Proposition 6.1, if $u(x) \geq a$ for almost all $x \in \Omega$, where a is a constant and Ω is an open set, then $u(x) \geq a$ for all $x \in \Omega \setminus \mathcal{N}$ where $\mathcal{N}$ is a properly exceptional set. Hence, (7.5) implies that

$$\int_0^\infty \mathcal{P}_t^B 1(x)\,dt \geq cF(R) \qquad \text{for all } x \in \zeta B \setminus \mathcal{N} \tag{7.6}$$

for some properly exceptional set $\mathcal{N} = \mathcal{N}_B$. Taking the union of such sets $\mathcal{N}_B$ where B varies over a countable family S of all balls with rational radii and whose centers form a dense subset of M, we obtain a properly exceptional set $\mathcal{N}$ such that (7.6) holds for any ball $B \in S$. Approximating any ball B from inside by balls of the family S, we obtain (7.6) for all balls, which implies (7.4).

Now let us prove (7.5). By the comparison principle of [31], Proposition 4.7 (see also [30], Lemma 4.18), we have, for any nonnegative function $f \in L^2 \cap L^\infty(B)$,

$$P_t f(x) \leq P_t^B f(x) + \sup_{s\in(0,t]} \operatorname*{esup}_{y \in B\setminus(1/2)B} P_s f(y) \tag{7.7}$$

[13] Note that (*RVD*) is essential for ($E_F \leq$) (see [31], Theorem 2.2). In fact, it was shown in [31] and [45] that (*VD*) + (*RVD*) + (UE_{weak}) imply also ($E_F \geq$) provided the Dirichlet form is conservative. In our setting the conservativeness of the Dirichlet form can also be proved but a direct proof of ($E_F \geq$) is shorter.

for almost all $x \in B$. Let ζ be a small positive constant to be specified below, and set $f = \mathbf{1}_{\zeta B}$. It follows from (NLE_{weak}) and (3.13) that

$$p_t(x,z) \ge \frac{c}{V(x_0, \mathcal{R}(t))} \qquad \text{for a.a. } x, z \in B\left(x_0, \frac{1}{2}\eta\mathcal{R}(t)\right), \tag{7.8}$$

provided $0 < t \le \varepsilon F(R)$. The initial value of ε is given by the condition (NLE_{weak}) but we are going to further reduce this value of ε in the course of the proof. Assume that t varies in the following interval:

$$\tfrac{1}{2}\varepsilon F(R) \le t \le \varepsilon F(R). \tag{7.9}$$

The left-hand side inequality in (7.9) implies by (3.19) that

$$R \le C\left(\frac{1}{\varepsilon}\right)^{1/\beta}\mathcal{R}(t). \tag{7.10}$$

Chose ζ from the identity

$$\zeta C\left(\frac{1}{\varepsilon}\right)^{1/\beta} = \frac{1}{2}\eta \tag{7.11}$$

so that (7.10) implies

$$B(x_0, \zeta R) \subset B\big(x_0, \tfrac{1}{2}\eta\mathcal{R}(t)\big).$$

Integrating (7.8) over $B(x_0, \zeta R)$ and using (VD) and (7.11), we obtain

$$\begin{aligned} P_t f(x) &= \int_{B(x_0,\zeta R)} p_t(x,z)\,d\mu(z) \\ &\ge \frac{cV(x_0,\zeta R)}{V(x_0,\mathcal{R}(t))} \\ &\ge c\zeta^{\alpha} \\ &= c'\varepsilon^{\alpha/\beta} \end{aligned} \tag{7.12}$$

for almost all $x \in B(x_0, \zeta R)$. On the other hand, for almost all $y \in B \setminus \frac{1}{2}B$, we have by ($UE_{\text{weak}}$) and Lemma 3.19

$$\begin{aligned} P_s f(y) &= \int_{B(x_0,\zeta R)} p_s(y,z)\,d\mu(z) \\ &\le C\frac{V(x_0,R)}{V(y,\mathcal{R}(s))}\exp\left(-c\left(\frac{F(R)}{s}\right)^{1/(\beta'-1)}\right), \end{aligned}$$

where we have used that $d(y,z) \simeq R$ and $s \le t < F(R)$. Using (3.13) and (3.19) we obtain

$$\frac{V(x_0,R)}{V(y,\mathcal{R}(s))} \le C\left(\frac{R}{\mathcal{R}(s)}\right)^{\alpha} \le C'\left(\frac{F(R)}{s}\right)^{\alpha/\beta}.$$

Finally, it follows from (7.9) and $s \le t$ that $\frac{F(R)}{s} \ge \frac{1}{\varepsilon}$ whence

$$P_s f(y) \le C\left(\frac{1}{\varepsilon}\right)^{\alpha/\beta} \exp\left(-c\left(\frac{1}{\varepsilon}\right)^{1/(\beta'-1)}\right) \tag{7.13}$$

for almost all $y \in B \setminus \frac{1}{2}B$. Combining (7.7), (7.12) and (7.13), we obtain, for almost all $x \in B(x_0, \zeta R)$,

$$\begin{aligned} P_t^B f(x) &\ge P_t f(x) - \sup_{s\in(0,t]} \operatorname*{esup}_{B\setminus K} P_s f \\ &\ge c'\varepsilon^{\alpha/\beta} - C\left(\frac{1}{\varepsilon}\right)^{\alpha/\beta} \exp\left(-c\left(\frac{1}{\varepsilon}\right)^{1/(\beta'-1)}\right) \\ &\ge \frac{1}{2}c'\varepsilon^{\alpha/\beta}, \end{aligned}$$

provided ε is chosen small enough. The path $t \mapsto P_t^B f$ is a continuous path in $L^2(B)$ and, hence, can be integrated in t. It follows from the previous inequality that

$$\int_0^\infty P_t^B 1\, dt \ge \int_{(1/2)\varepsilon F(R)}^{\varepsilon F(R)} P_t^B f\, dt \ge c\varepsilon^{\alpha/\beta+1} F(R),$$

which finishes the proof of $(E_F \ge)$.

We are left to prove that (iii) $\Rightarrow$ (H). By [10], Theorem 3.1 (see also [18] and [41], Theorem 5.3), $(VD) + (UE_{\text{weak}}) + (NLE_{\text{weak}})$ imply the *parabolic* Harnack inequality for bounded caloric function and, hence, the Harnack inequality (H) for bounded harmonic functions (note that this result uses the precompactness of the balls). We still have to obtain (H) for all nonnegative harmonic functions. Note that by [31], Theorem 2.1,

$$(VD) + (RVD) + (UE_{\text{weak}}) \Rightarrow (FK).$$

In particular, for any ball B, we have $\lambda_{\min}(B) > 0$. Given a function $u \in \mathcal{F}$ that is nonnegative and harmonic in a ball $B \subset M$, set $f_n = u \wedge n$ for any $n \in \mathbb{N}$, and denote by u_n the solution of the Dirichlet problem

$$\begin{cases} u_n \text{ is harmonic in } B, \\ u_n = f_n \operatorname{mod} \mathcal{F}(B); \end{cases}$$

cf. Section 7.1. Since $0 \le f_n \le n$, we have also $0 \le u_n \le n$. Since the sequence $\{f_n\}$ increases and $f_n \xrightarrow{\mathcal{F}} u$ (cf. [20], Theorem 1.4.2), it follows by Lemma 7.2 that $u_n \to u$ almost everywhere in B. Each function u_n is bounded and, hence, satisfies the Harnack inequality in B, that is,

$$\operatorname*{esup}_{\delta B} u_n \le C \operatorname*{einf}_{\delta B} u_n.$$

Replacing in the right-hand side u_n by a larger function u and passing to the limit in the left-hand side as $n \to \infty$, we obtain the same inequality for u, which was to be proved. $\square$

COROLLARY 7.5. *Assume that all metric balls are precompact*, diam $M = \infty$, *and the Dirichlet form* $(\mathcal{E}, \mathcal{F})$ *is conservative. Then the following sets of conditions are equivalent*:

(i) $(VD) + (H) + (UE_{\mathrm{weak}})$;
(ii) $(VD) + (UE) + (NLE)$.

PROOF. In the view of Theorem 7.4, it suffices to prove that (i) $\Rightarrow$ (E_F). By Lemma 7.3, (H) implies the connectedness of M. By [31], Corollary 5.3, $(VD) \Rightarrow (RVD)$ provided M is connected and unbounded, which is the case now. By [31], Theorem 2.2, the conservativeness and $(VD) + (RVD) + (UE_{\mathrm{weak}})$ imply (E_F). $\square$

Many equivalent conditions for (UE_{weak}) were proved in [31] under the standing assumptions $(VD) + (RVD)$ and the conservativeness of $(\mathcal{E}, \mathcal{F})$. Of course, each of these conditions can replace (UE_{weak}) in the statement of Corollary 7.5.

COROLLARY 7.6. *Assume that all metric balls are precompact*, diam $M = \infty$, *and* (M, d) *satisfies the chain condition. Then the following two sets of conditions are equivalent*:

(i) $(VD) + (H) + (E_F)$;
(ii) *The heat kernel exists and satisfies the two-sided estimate* (6.36).

PROOF. The implication (i) $\Rightarrow$ (ii) is contained in Corollary 6.11. Let us prove the implication (ii) $\Rightarrow$ (i). Estimate (6.36) implies (UE) as well as (NLE) with any value of η, in particular, $\eta > 1$; cf. Remark 6.12. By [34], Lemma 4.1, (NLE) with $\eta > 1$ implies (VD). Finally, by Theorem 7.4, we obtain $(H) + (E_F)$. $\square$

APPENDIX: LIST OF CONDITIONS

We briefly list the lettered conditions used in this paper with references to the appropriate places in the main body.

(H) $\operatorname{esup}_{B(x,\delta r)} u \le C \operatorname{einf}_{B(x,\delta r)} u$ (Section 3.2);
(VD) $V(x, 2r) \le C V(x, r)$ (Section 3.2);
(E_F) $\mathbb{E}_x \tau_{B(x,r)} \simeq F(r)$ (Section 3.3);
(FK) $\lambda_{\min}(\Omega) \ge \frac{c}{F(R)} (\frac{\mu(B)}{\mu(\Omega)})^\nu$ (Section 3.3);
(UE) $p_t(x, y) \le \frac{C}{V(x,\mathcal{R}(t))} \exp(-\frac{1}{2}\Phi(cd(x, y), t))$ (Section 4);
(NLE) $p_t(x, y) \ge \frac{c}{V(x,\mathcal{R}(t))}$ provided $d(x, y) \le \eta\mathcal{R}(t)$ (Section 5.4);
(RVD) $\frac{V(x,R)}{V(x,r)} \ge c(\frac{R}{r})^{\alpha'}$ (Section 7.2).

Acknowledgments. The authors thank Martin Barlow, Jiaxin Hu, Jun Kigami, Takashi Kumagai and Alexander Teplyaev for valuable conversations on the topics of this paper. The authors are grateful to the unnamed referees for careful reading of the manuscript and for the useful remarks.

REFERENCES

[1] ARONSON, D. G. (1968). Non-negative solutions of linear parabolic equations. *Ann. Sc. Norm. Super. Pisa* (3) **22** 607–694. MR0435594

[2] BARLOW, M. T. (1998). Diffusions on fractals. In *Lectures on Probability Theory and Statistics* (*Saint-Flour*, 1995). *Lecture Notes in Math.* **1690** 1–121. Springer, Berlin. MR1668115

[3] BARLOW, M. T. and BASS, R. F. (1989). The construction of Brownian motion on the Sierpiński carpet. *Ann. Inst. Henri Poincaré Probab. Stat.* **25** 225–257. MR1023950

[4] BARLOW, M. T. and BASS, R. F. (1992). Transition densities for Brownian motion on the Sierpiński carpet. *Probab. Theory Related Fields* **91** 307–330. MR1151799

[5] BARLOW, M. T. and BASS, R. F. (1999). Brownian motion and harmonic analysis on Sierpinski carpets. *Canad. J. Math.* **51** 673–744. MR1701339

[6] BARLOW, M. T., BASS, R. F., CHEN, Z.-Q. and KASSMANN, M. (2009). Non-local Dirichlet forms and symmetric jump processes. *Trans. Amer. Math. Soc.* **361** 1963–1999. MR2465826

[7] BARLOW, M. T., BASS, R. F. and KUMAGAI, T. (2006). Stability of parabolic Harnack inequalities on metric measure spaces. *J. Math. Soc. Japan* **58** 485–519. MR2228569

[8] BARLOW, M. T., BASS, R. F., KUMAGAI, T. and TEPLYAEV, A. (2010). Uniqueness of Brownian motion on Sierpiński carpets. *J. Eur. Math. Soc.* (*JEMS*) **12** 655–701. MR2639315

[9] BARLOW, M. T., GRIGOR'YAN, A. and KUMAGAI, T. (2009). Heat kernel upper bounds for jump processes and the first exit time. *J. Reine Angew. Math.* **626** 135–157. MR2492992

[10] BARLOW, M. T., GRIGOR'YAN, A. and KUMAGAI, T. (2012). On the equivalence of parabolic Harnack inequalities and heat kernel estimates. *J. Math. Soc. Japan.* To appear.

[11] BARLOW, M. T. and HAMBLY, B. M. (1997). Transition density estimates for Brownian motion on scale irregular Sierpinski gaskets. *Ann. Inst. Henri Poincaré Probab. Stat.* **33** 531–557. MR1473565

[12] BARLOW, M. T. and PERKINS, E. A. (1988). Brownian motion on the Sierpiński gasket. *Probab. Theory Related Fields* **79** 543–623. MR0966175

[13] CHAVEL, I. (1984). *Eigenvalues in Riemannian Geometry*. *Pure and Applied Mathematics* **115**. Academic Press, Orlando, FL. MR0768584

[14] CHAVEL, I. (2001). *Isoperimetric Inequalities*: *Differential Geometric and Analytic Perspectives*. *Cambridge Tracts in Mathematics* **145**. Cambridge Univ. Press, Cambridge. MR1849187

[15] COULHON, T. and GRIGORYAN, A. (1998). Random walks on graphs with regular volume growth. *Geom. Funct. Anal.* **8** 656–701. MR1633979

[16] DAVIES, E. B. (1980). *One-Parameter Semigroups*. *London Mathematical Society Monographs* **15**. Academic Press [Harcourt Brace Jovanovich Publishers], London. MR0591851

[17] DAVIES, E. B. (1989). *Heat Kernels and Spectral Theory*. *Cambridge Tracts in Mathematics* **92**. Cambridge Univ. Press, Cambridge. MR0990239

[18] FABES, E. B. and STROOCK, D. W. (1986). A new proof of Moser's parabolic Harnack inequality using the old ideas of Nash. *Arch. Ration. Mech. Anal.* **96** 327–338. MR0855753

[19] FITZSIMMONS, P. J., HAMBLY, B. M. and KUMAGAI, T. (1994). Transition density estimates for Brownian motion on affine nested fractals. *Comm. Math. Phys.* **165** 595–620. MR1301625

[20] FUKUSHIMA, M., OSHIMA, Y. and TAKEDA, M. (1994). *Dirichlet Forms and Symmetric Markov Processes*. *de Gruyter Studies in Mathematics* **19**. de Gruyter, Berlin. MR1303354

[21] GOLDSTEIN, S. (1987). Random walks and diffusions on fractals. In *Percolation Theory and Ergodic Theory of Infinite Particle Systems* (*Minneapolis*, *Minn*., 1984–1985) (H. Kesten, ed.). *The IMA Volumes in Mathematics and its Applications* **8** 121–129. Springer, New York. MR0894545

[22] GRIGOR'YAN, A. (1991). The heat equation on noncompact Riemannian manifolds. *Mat*. *Sb*. **182** 55–87. MR1098839

[23] GRIGOR'YAN, A. (1994). Heat kernel upper bounds on a complete non-compact manifold. *Rev*. *Mat*. *Iberoam*. **10** 395–452. MR1286481

[24] GRIGOR'YAN, A. (1997). Gaussian upper bounds for the heat kernel on arbitrary manifolds. *J*. *Differential Geom*. **45** 33–52. MR1443330

[25] GRIGOR'YAN, A. (1999). Estimates of heat kernels on Riemannian manifolds. In *Spectral Theory and Geometry* (*Edinburgh*, 1998). *London Mathematical Society Lecture Note Series* **273** 140–225. Cambridge Univ. Press, Cambridge. MR1736868

[26] GRIGOR'YAN, A. (2003). Heat kernels and function theory on metric measure spaces. In *Heat Kernels and Analysis on Manifolds*, *Graphs*, *and Metric Spaces* (*Paris*, 2002). *Contemporary Mathematics* **338** 143–172. Amer. Math. Soc., Providence, RI. MR2039954

[27] GRIGOR'YAN, A. (2006). Heat kernels on weighted manifolds and applications. In *The Ubiquitous Heat Kernel*. *Contemporary Mathematics* **398** 93–191. Amer. Math. Soc., Providence, RI. MR2218016

[28] GRIGOR'YAN, A. (2009). *Heat Kernel and Analysis on Manifolds*. *AMS/IP Studies in Advanced Mathematics* **47**. Amer. Math. Soc., Providence, RI. MR2569498

[29] GRIGOR'YAN, A. (2010). Heat kernels on metric measure spaces with regular volume growth. In *Handbook of Geometric Analysis*, *No. 2* (L. Ji, P. Li, R. Schoen and L. Simon, eds.). *Advanced Lectures in Math.* **13** 1–60. International Press, Somerville, MA. MR2743439

[30] GRIGOR'YAN, A. and HU, J. (2008). Off-diagonal upper estimates for the heat kernel of the Dirichlet forms on metric spaces. *Invent*. *Math*. **174** 81–126. MR2430977

[31] GRIGOR'YAN, A. and HU, J. (2009). Upper bounds of heat kernels on doubling spaces. Preprint.

[32] GRIGOR'YAN, A., HU, J. and LAU, K.-S. (2003). Heat kernels on metric measure spaces and an application to semilinear elliptic equations. *Trans*. *Amer. Math*. *Soc*. **355** 2065–2095 (electronic). MR1953538

[33] GRIGOR'YAN, A., HU, J. and LAU, K.-S. (2010). Comparison inequalities for heat semigroups and heat kernels on metric measure spaces. *J. Funct*. *Anal*. **259** 2613–2641. MR2679020

[34] GRIGOR'YAN, A., HU, J. and LAU, K. S. (2009). Heat kernels on metric spaces with doubling measure. In *Fractal Geometry and Stochastics IV* (Ch. Brandt, P. Mörters and M. Zähle, eds.). *Progress in Probability* **61** 3–44. Birkhäuser, Basel. MR2762672

[35] GRIGOR'YAN, A. and SALOFF-COSTE, L. (2009). Heat kernel on manifolds with ends. *Ann*. *Inst*. *Fourier* (*Grenoble*) **59** 1917–1997. MR2573194

[36] GRIGOR'YAN, A. and TELCS, A. (2001). Sub-Gaussian estimates of heat kernels on infinite graphs. *Duke Math*. *J*. **109** 451–510. MR1853353

[37] GRIGOR'YAN, A. and TELCS, A. (2002). Harnack inequalities and sub-Gaussian estimates for random walks. *Math*. *Ann*. **324** 521–556. MR1938457

[38] GUSHCHIN, A. K. (1982). Uniform stabilization of solutions of the second mixed problem for a parabolic equation. *Mat*. *Sb*. (*N.S.*) **119** 451–508, 590. MR0682495

[39] GUSHCHIN, A. K., MIKHAĬLOV, V. P. and MIKHAĬLOV, Y. A. (1985). Uniform stabilization of the solution of the second mixed problem for a second-order parabolic equation. *Mat*. *Sb.* (*N.S.*) **128** 147–168, 287. MR0809483

[40] HAMBLY, B. M. and KUMAGAI, T. (1999). Transition density estimates for diffusion processes on post critically finite self-similar fractals. *Proc*. *Lond*. *Math*. *Soc*. (3) **78** 431–458. MR1665249

[41] HEBISCH, W. and SALOFF-COSTE, L. (2001). On the relation between elliptic and parabolic Harnack inequalities. *Ann. Inst. Fourier* (*Grenoble*) **51** 1437–1481. MR1860672

[42] JORGENSON, J. and LYNNE, W., eds. (2006). *The Ubiquitous Heat Kernel. Contemporary Mathematics* **398**. Amer. Math. Soc., Providence, RI. MR2218505

[43] KIGAMI, J. (1993). Harmonic calculus on p.c.f. self-similar sets. *Trans. Amer. Math. Soc.* **335** 721–755. MR1076617

[44] KIGAMI, J. (2001). *Analysis on Fractals. Cambridge Tracts in Mathematics* **143**. Cambridge Univ. Press, Cambridge. MR1840042

[45] KIGAMI, J. (2004). Local Nash inequality and inhomogeneity of heat kernels. *Proc. Lond. Math. Soc.* (3) **89** 525–544. MR2078700

[46] KIGAMI, J. (2009). Resistance forms, quasisymmetric maps and heat kernel estimates. Preprint.

[47] KIGAMI, J. (2009). Volume doubling measures and heat kernel estimates on self-similar sets. *Mem. Amer. Math. Soc.* **199** viii+94. MR2512802

[48] KUMAGAI, T. (1993). Estimates of transition densities for Brownian motion on nested fractals. *Probab. Theory Related Fields* **96** 205–224. MR1227032

[49] KUSUOKA, S. (1987). A diffusion process on a fractal. In *Probabilistic Methods in Mathematical Physics* (*Katata/Kyoto*, 1985) (K. Ito and N. Ikeda, eds.) 251–274. Academic Press, Boston, MA. MR0933827

[50] KUSUOKA, S. and YIN, Z. X. (1992). Dirichlet forms on fractals: Poincaré constant and resistance. *Probab. Theory Related Fields* **93** 169–196. MR1176724

[51] LI, P. and YAU, S.-T. (1986). On the parabolic kernel of the Schrödinger operator. *Acta Math.* **156** 153–201. MR0834612

[52] MOLČANOV, S. A. (1975). Diffusion processes and Riemannian geometry. *Uspehi Mat. Nauk* **30** 3–59. MR0413289

[53] MOSER, J. (1961). On Harnack's theorem for elliptic differential equations. *Comm. Pure Appl. Math.* **14** 577–591. MR0159138

[54] MOSER, J. (1964). A Harnack inequality for parabolic differential equations. *Comm. Pure Appl. Math.* **17** 101–134. MR0159139

[55] SALOFF-COSTE, L. (1992). A note on Poincaré, Sobolev, and Harnack inequalities. *Int. Math. Res. Not. IMRN* **2** 27–38. MR1150597

[56] SALOFF-COSTE, L. (2002). *Aspects of Sobolev-Type Inequalities. London Mathematical Society Lecture Note Series* **289**. Cambridge Univ. Press, Cambridge. MR1872526

[57] SCHOEN, R. and YAU, S. T. (1994). *Lectures on Differential Geometry*. International Press, Cambridge, MA. MR1333601

[58] TELCS, A. (2006). *The Art of Random Walks. Lecture Notes in Math.* **1885**. Springer, Berlin. MR2240535

[59] VAROPOULOS, N. T., SALOFF-COSTE, L. and COULHON, T. (1992). *Analysis and Geometry on Groups. Cambridge Tracts in Mathematics* **100**. Cambridge Univ. Press, Cambridge. MR1218884

[60] YOSIDA, K. (1980). *Functional Analysis*, 6th ed. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **123**. Springer, Berlin. MR0617913

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF BIELEFELD
33501 BIELEFELD
GERMANY
E-MAIL: grigor@math.uni-bielefeld.de

DEPARTMENT OF COMPUTER SCIENCES
AND INFORMATION THEORY
BUDAPEST UNIVERSITY OF TECHNOLOGY
AND ECONOMICS
MAGYAR TUDÓSOK KÖRÚTJA 2
H-1117, BUDAPEST
HUNGARY
E-MAIL: telcs@szit.bme.hu