Besov class via heat semigroup on Dirichlet spaces II: BV functions and Gaussian heat kernel estimates

Patricia Alonso-Ruiz, Fabrice Baudoin[*], Li Chen, Luke Rogers, Nageswari Shanmugalingam[†], Alexander Teplyaev

Abstract

We introduce the class of bounded variation (BV) functions in a general framework of strictly local Dirichlet spaces with doubling measure. Under the 2-Poincaré inequality and a weak Bakry-Émery curvature type condition, this BV class is identified with the heat semigroup based Besov class $\mathbf{B}^{1,1/2}(X)$ that was introduced in our previous paper. Assuming furthermore a strong Bakry-Émery curvature type condition, we prove that for $p > 1$, the Sobolev class $W^{1,p}(X)$ can be identified with $\mathbf{B}^{p,1/2}(X)$. Consequences of those identifications in terms of isoperimetric and Sobolev inequalities with sharp exponents are given.

Contents

[*]Supported in part by NSF grant DMS-1660031 and the Simons Foundation
[†]Partially supported by NSF (U.S.A.) grant DMS-1800161

1 Introduction

In metric measure spaces X that are highly path-connected, the theory of Sobolev classes based on upper gradients provides an approach to calculus using a derivative structure that is strongly local. One can construct an analog of $|\nabla f|$ (called a weak upper gradient) when f is a measurable function on the metric space; $|\nabla f|$ satisfies a variant of the fundamental theorem of calculus along most rectifiable curves in X, and has the property that if f is constant on a Borel set $E \subset X$, then $|\nabla f| = 0$ almost everywhere in E, see [39]. Should the measure on X be doubling and X be sufficiently well-connected by paths that X supports a 1-Poincaré inequality controlling f by $|\nabla f|$, the notion of functions of bounded variation (BV) as constructed by [72] leads to a fruitful exploration of geometry of X in terms of BV functions and sets of finite perimeter (sets whose characteristic functions are BV functions).

However, there are doubling metric measure spaces that are well-connected enough to support a p-Poincaré inequality for some $p > 1$ but not a 1-Poincaré inequality, see for example [68]. Moreover, there are many metric measure spaces, including fractals like the Sierpiński gasket, that have a doubling measure but support no Poincaré inequality of this type; some of these spaces are even quasiconvex. In such settings, the theory of BV functions based on upper gradients as in [6, 72] is not very productive. On the other hand, the theory of Dirichlet forms is well-developed in many such spaces, including both certain types of fractal spaces and some general settings like connected metric measure spaces supporting a p-Poincaré inequality for some $p > 1$, see for example [23, 48, 54, 55, 57, 60, 61, 74, 76, 79–81], which is far from an exhaustive list of the literature. In such situations, the theory of BV functions has not been well-explored. In [61] the Sobolev type spaces constructed using Dirichlet forms were shown to be the same as those constructed using upper gradients if the metric space supports a 2-Poincaré inequality. From [23] it follows that in a doubling metric measure space supporting a p-Poincaré inequality for some $1 \le p \le 2$, there is a Dirichlet form that is compatible with the upper gradient Sobolev class structure, see for example [60].

The goal of this paper is to develop a theory of a BV class of functions from Dirichlet forms in a specific setting: that of a locally compact metric space X, equipped with a doubling Radon measure μ and a strictly local Dirichlet form $\mathcal{E}$ on $L^2(X)$. We propose a notion of BV functions and prove fundamental properties, including the Radon measure property of the BV energy seminorm, the notion of sets of finite perimeter, and a co-area formula (see Theorem 3.11) connecting sets of finite perimeter to BV energy.

In this paper we will also compare the notion of BV functions developed here to Besov classes derived from the heat semigroup and from purely metric notions. The heat-semigroup based Besov classes have already been considered in some specific settings, for example in [75], and are investigated in a very general setting in the first paper in this series [3]. The paper [75] also considered a notion of metric-based Besov classes under the assumption that the measure is Ahlfors regular; our metric notion of Besov classes will not assume that the underlying measure μ is Ahlfors regular, but we will assume that it is doubling with respect to the intrinsic metric $d_{\mathcal{E}}$ of the Dirichlet form, see (1), which is well-defined as $\mathcal{E}$ is regular. We prove that under a weak Bakry-Émery curvature condition, the heat semigroup BV class we construct is the same as the metric-based Besov class $\mathbf{B}^{1,1/2}(X)$. One of the key tools used there is the co-area formula Theorem 3.11. Finally, we will establish Sobolev-type and isoperimetric inequalities that parallel the classical Sobolev embedding theorems associated with the classical Sobolev and BV classes as in [1, 70].

The tools of heat semigroup based Besov spaces needed in this paper were developed in the first paper in this series [3], but the present paper can largely be read independently from [3]. The results of this paper depend heavily on the assumption that $\mathcal{E}$ be not strictly local, but in the third paper [4] in this series we develop two types of Besov classes (one based on the heat semigroup and the other based on a metric) as possible substitutes for BV functions.

The structure of this paper is as follows. In Section 2 we give a description of the background notions and discuss related results from some of the existing literature on Dirichlet forms, though in the interests of brevity our discussion is far from exhaustive.

The definition of BV class of functions is given in Section 3, where we also give a proof that the BV energy density $\|Du\|$ of $u \in BV(X)$ is a Radon measure on X, see Theorem 3.7. The main tool used in the proof of this theorem is a characterization of Radon measures due to De Giorgi and Letta. We then establish a co-area formula for BV functions.

Section 4 is the heart of the paper. We begin by comparing the heat semigroup-based Besov class $\mathbf{B}^{p,\alpha/2}(X)$ introduced in [3] with a more classical Besov class $B^{\alpha}_{p,\infty}(X)$ that was defined in [34] and is based on the intrinsic metric $d = d_{\mathcal{E}}$ rather than the heat semigroup P_t. Under our standing hypotheses (μ is doubling and supports a 2-Poincaré inequality) we show that $\mathbf{B}^{p,\alpha/2}(X)$ coincides with $B^{\alpha}_{p,\infty}(X)$. It should be noted that this differs from the correspondence of metric and heat semigroup based Besov class established in [75] in that the latter assumes μ is Ahlfors regular. We then compare the class $BV(X)$ to the heat semi-group Besov class $\mathbf{B}^{1,1/2}(X)$ and show these coincide under the additional hypothesis that $\mathcal{E}$ supports a weak Bakry-Émery curvature condition, see Theorem 4.4. We also explore a connection between the co-dimension 1 Hausdorff measure of the regular boundary $\partial^{r_0}_{\alpha} E$ of a set E of finite perimeter (meaning $\mathbf{1}_E \in BV(X)$) to its perimeter measure $\|D\mathbf{1}_E\|(X) =: P(E, X)$, see Proposition 4.7. In the last part of Section 4 we show that if X supports a strong Bakry-Émery curvature condition and $p > 1$, then the heat semigroup-based Besov class $\mathbf{B}^{p,1/2}(X)$ coincides with the Sobolev space $W^{1,p}(X)$, see Theorems 4.9, 4.11 and 4.17.

We conclude the paper in Section 5 with a discussion of Sobolev type embedding theorems for Besov and BV spaces in the context of strictly local Dirichlet forms satisfying the weak Bakry-Émery estimate.

2 Preliminaries

2.1 Strictly local Dirichlet spaces, doubling measures, and our standing assumptions

Throughout the paper, let X be a locally compact metric space equipped with a Radon measure μ supported on X. Let $(\mathcal{E}, \mathcal{F} = \mathbf{dom}(\mathcal{E}))$ be a Dirichlet form on X, meaning it is a densely defined, closed, symmetric and Markovian form on $L^2(X)$. The book [32] is a classical reference on the theory of Dirichlet forms. We also refer to the foundational papers by K.T. Sturm [82–84].

We denote by $C_c(X)$ the vector space of all continuous functions with compact support in X and $C_0(X)$ its closure with respect to the supremum norm. A core for $(X, \mu, \mathcal{E}, \mathcal{F})$ is a subset $\mathcal{C}$ of $C_c(X) \cap \mathcal{F}$ which is dense in $C_c(X)$ in the supremum norm and dense in $\mathcal{F}$ in the norm

$$\left(\|f\|^2_{L^2(X)} + \mathcal{E}(f, f) \right)^{1/2}.$$

The Dirichlet form $\mathcal{E}$ is called regular if it admits a core. It is called *strongly local* if for any two functions $u, v \in \mathcal{F}$ with compact supports such that u is constant in a neighborhood of the support of v, we have $\mathcal{E}(u, v) = 0$ (see Page 6 of [32]).

Throughout this paper, we assume that $(\mathcal{E}, \mathcal{F})$ is a strongly local regular Dirichlet form on $L^2(X)$. Since $\mathcal{E}$ is regular, for every $u, v \in \mathcal{F} \cap L^\infty(X)$, we can define the energy measure $\Gamma(u, v)$

through the formula

$$\int_X \phi \, d\Gamma(u,v) = \frac{1}{2}[\mathcal{E}(\phi u, v) + \mathcal{E}(\phi v, u) - \mathcal{E}(\phi, uv)], \quad \phi \in \mathcal{F} \cap C_c(X).$$

Then $\Gamma(u,v)$ can be extended to all $u,v \in \mathcal{F}$ by truncation (see [24, Theorem 4.3.11]). According to Beurling and Deny [19], one has then for $u,v \in \mathcal{F}$

$$\mathcal{E}(u,v) = \int_X d\Gamma(u,v)$$

and $\Gamma(u,v)$ is a signed Radon measure.

Definition 2.1. *Observe that the energy measures $\Gamma(u,v)$ inherit a strong locality property from $\mathcal{E}$, namely that $\mathbf{1}_U d\Gamma(u,v) = 0$ for any open subset $U \subset X$ and $u,v \in \mathcal{F}$ such that u is a constant on U. One can then extend Γ to $\mathcal{F}_{\text{loc}}(X)$ defined as*

$$\mathcal{F}_{\text{loc}}(X) = \{u \in L^2_{\text{loc}}(X) : \forall \text{ compact } K \subset X, \exists v \in \mathcal{F} \text{ such that } u = v|_K \text{ a.e.}\}.$$

We will still denote this extension by Γ. For later use we collect some properties of this extension (see for instance [32, Section 3.2] and also [83, Section 4]).

- *Strong locality.* For all $u,v \in \mathcal{F}_{\text{loc}}(X)$ and all open subset $U \subset X$ on which u is a constant
$$\mathbf{1}_U d\Gamma(u,v) = 0.$$
- *Leibniz and chain rules.* For all $u \in \mathcal{F}_{\text{loc}}(X), v \in \mathcal{F}_{\text{loc}}(X) \cap L^\infty_{\text{loc}}(X)$, $w \in \mathcal{F}_{\text{loc}}(X)$ and $\eta \in C^1(\mathbb{R})$,
$$\begin{aligned} d\Gamma(uv, w) &= u d\Gamma(v,w) + v d\Gamma(u,w), \\ d\Gamma(\eta(u), v) &= \eta'(u) d\Gamma(u,v). \end{aligned}$$

With respect to $\mathcal{E}$ we can define the following *intrinsic metric* $d_{\mathcal{E}}$ on X by

$$d_{\mathcal{E}}(x,y) = \sup\{u(x) - u(y) \,:\, u \in \mathcal{F} \cap C_0(X) \text{ and } d\Gamma(u,u) \le d\mu\}. \tag{1}$$

Here the condition $d\Gamma(u,u) \le d\mu$ means that $\Gamma(u,u)$ is absolutely continuous with respect to μ with Radon-Nikodym derivative bounded by 1. The term "intrinsic metric" is potentially misleading because in general there is no reason why $d_{\mathcal{E}}$ is a metric on X (it could be infinite for a given pair of points x,y or zero for some distinct pair of points), however in this paper we will work in a standard setting in which it is a metric. The following definition is from [67, and references therein], which is based on the classical papers [20–22, 82–84]).

Definition 2.2. *A strongly local regular Dirichlet space is called strictly local if $d_{\mathcal{E}}$ is a metric on X and the topology induced by $d_{\mathcal{E}}$ coincides with the topology on X.*

We will assume strict locality throughout the paper.

Example 2.3. In the context of a complete metric measure space (X,d,μ) supporting a 2-Poincaré inequality and where μ is doubling, one can construct a Dirichlet form $\mathcal{E}$ with domain $N^{1,2}(X)$ by using a choice of a Cheeger differential structure as in [23]. This Dirichlet form is then strictly local and the intrinsic distance $d_{\mathcal{E}}$ is bi-Lipschitz equivalent to the original metric d. We refer to [69] and the references therein for further details. This framework encompasses for instance the one of Riemannian manifolds with non-negative Ricci curvature and the one of doubling sub-Riemannian spaces supporting a 2-Poincaré inequality.

Example 2.4. In the context of fractals, strictly local Dirichlet forms appear in [5, 21, 22, 31, 42, 49, 50, 53, 56, 63, 71, 85, 86] and play an important role in analysis of first-order derivatives in these settings. Whether every local Dirichlet form admits a change of measure under which it becomes strictly local is an open question, though some natural conditions for this are discussed in [40, 43], where it is also proved that Γ is the norm of a well defined gradient that may be extended to measurable 1-forms, see [41]. Without giving details of this analysis, we mention that existence of a suitable collection of finite (Dirichlet) energy coordinate functions, which depend only on the Dirichlet form $\mathcal{E}$, is essentially equivalent to the existence of a measure which is compatible with an intrinsic distance. In particular, [40] proves existence of a measure which is compatible with an intrinsic distance for any local resistance form in the sense of Kigami [53, 55, 57, 58]. Thus, any fractal space with a local resistance form has an intrinsic metric and is a strictly local Dirichlet form for an appropriate choice of the measure.

Now suppose in addition to strict locality we know that open balls have compact closures and that $(X, d_{\mathcal{E}})$ is complete. In this setting we may apply [83, Lemma 1, Lemma 1′] to obtain that the distance function $\varphi_x : y \mapsto d_{\mathcal{E}}(x, y)$ on X is in $\mathcal{F}_{\text{loc}}(X) \cap \mathcal{C}$ and $d\Gamma(\varphi_x, \varphi_x) \le d\mu$. Then cut-off functions on intrinsic balls $B(x, r)$ of the form

$$\varphi_{x,r} : y \mapsto (r - d_{\mathcal{E}}(x, y))_+$$

are also in $\mathcal{F}_{\text{loc}}(X) \cap \mathcal{C}$ and $d\Gamma(\varphi_{x,r}, \varphi_{x,r}) \le d\mu$ (for all $r > 0$ and $x \in X$). The following lemma will be useful.

Lemma 2.5. *Let $f : X \to \mathbb{R}$ be locally Lipschitz continuous with respect to $d_{\mathcal{E}}$. Then $f \in \mathcal{F}_{\text{loc}}(X)$ with $\Gamma(f, f) \ll \mu$. If f is locally K-Lipschitz, then $\Gamma(f, f) \le K^2\,\mu$.*

Proof. Let Q be a countable dense subset of X. Let $U \subset X$ be a bounded open set and let $\{q_i\}_{i\in I\subset\mathbb{N}}$ be an enumeration of $Q \cap U$. Note that $Q \cap U$ is dense in U. For each $i \in I$ let $\psi_i(x) = d_{\mathcal{E}}(x, q_i)$. Then as explained above, $\psi_i \in \mathcal{F}(U)$ with $\Gamma(\psi_i, \psi_i) \le \mu$. For $j \in I$ set

$$f_j(x) := \inf\{f(q_i) + K\,\psi_i(x) \,:\, i \in I \text{ with } i \le j\},$$

where $K \ge 0$ is the Lipschitz constant of f in U. The above functions are inspired by the proof of the McShane extension theorem (see for example [38]). By the lattice properties of Dirichlet forms, it is seen that each $f_j \in \mathcal{F}(U)$ with $d\Gamma(f_j, f_j) \le K^2\,d\mu$. Here, by the lattice property, we mean that if $u, v \in \mathcal{F}$, then $w_1 = \min\{u, v\}$ and $w_2 = \max\{u, v\}$ are also in $\mathcal{F}$ with $\Gamma(w_1, w_1) = \mathbf{1}_{\{u>v\}}\Gamma(v, v) + \mathbf{1}_{\{u\le v\}}\Gamma(u, u)$. Furthermore, f_j are K-Lipschitz in U with $f_j(q_i) = f(q_i)$ for $i \in I$ with $i \le j$. We can see that $f_j \to f$ monotonically and hence (as f and f_j are bounded in U because U is bounded) $f_j \to f$ in $L^2(U)$, with $d\Gamma(f, f)/d\mu \le K^2$ on U. ☐

At many places in the paper we will need to approximate using locally Lipschitz functions and use locally Lipschitz cutoffs, so will assume density of these functions in $L^1(X)$. Now we come to the final assumption which will be made throughout the paper, namely that μ is volume doubling.

Definition 2.6. *We say that the metric measure space $(X, d_{\mathcal{E}}, \mu)$ satisfies the volume doubling property if there exists a constant $C > 0$ such that for every $x \in X$ and $r > 0$,*

$$\mu(B(x, 2r)) \le C\,\mu(B(x, r)).$$

It follows from the doubling property of μ (see [38]) that there is a constant $0 < Q < \infty$ and $C \ge 1$ such that whenever $0 < r \le R$ and $x \in X$, we have

$$\frac{\mu(B(x, R))}{\mu(B(x, r))} \le C \left(\frac{R}{r}\right)^Q. \tag{2}$$

Another well-known consequence of the doubling property is the availability of a maximal ε-separated covering. Let $U \subset X$ be a non-empty subset and let $\varepsilon > 0$. Then there exists a family of balls $\{B_i^\varepsilon = B(x_i^\varepsilon, \varepsilon)\}_i$ such that

- The collection $\{\frac{1}{2}B_i^\varepsilon\}_i$ is a maximal pairwise disjoint family of balls with radius $\varepsilon/2$;
- The collection $\{B_i^\varepsilon\}_i$ covers U, that is, $U = \bigcup_i B_i^\varepsilon$;
- There exists $K \in \mathbb{N}$ such that each point $x \in X$ is contained in at most K balls from the family $\{2B_i^\varepsilon\}_i$.

Moreover if μ has the volume doubling property and $(X, d_{\mathcal{E}})$ is complete then it also follows that closed and bounded subsets of X are compact.

We now summarize the assumptions that will be in force throughout the paper, some of the above conclusions that will be used without further comment, and the notation for the upper gradient we frequently use.

Assumption 2.7.

- *The Dirichlet space $(X, \mu, \mathcal{E}, \mathcal{F})$ is strictly local, so $\mathcal{E}$ is strongly local and regular and $d_{\mathcal{E}}$ is a metric on X that induces the topology on X;*
- *the metric space $(X, d_{\mathcal{E}})$ is complete;*
- *μ is volume doubling;*
- *closed bounded subsets of $(X, d_{\mathcal{E}})$ are compact (this is a consequence of the preceeding two assumptions);*
- *locally Lipschitz functions are dense in $L^1(X)$;*
- *if $\Gamma(f, f)$ is absolutely continuous with respect to μ, as is the case for locally Lipschitz functions, then $|\nabla f|$ is the square root of its Radon-Nikodym derivative, so $\Gamma(f, f) = |\nabla f|^2 d\mu$.*

It should be noted that with the exception of some parts of Section 3, we will typically also assume existence of a 2-Poincaré inequality, which is discussed next.

2.2 The 2-Poincaré inequality

Let $(X, \mu, \mathcal{E}, \mathcal{F})$ be a strictly local regular Dirichlet space as in Section 2.1.

Definition 2.8. *We say that $(X, \mu, \mathcal{E}, \mathcal{F})$ supports the 2-Poincaré inequality if there are constants $C > 0$ and $\lambda \geq 1$ such that whenever B is a ball in X (with respect to the metric $d_{\mathcal{E}}$) and $u \in \mathcal{F}$, we have*

$$\frac{1}{\mu(B)} \int_B |u - u_B| \, d\mu \leq C \, \mathrm{rad}\,(B) \left(\frac{1}{\mu(\lambda B)} \int_{\lambda B} d\Gamma(u, u) \right)^{1/2}.$$

Remark 2.9. *The 2-Poincaré inequality does not need $\mathcal{E}$ to be strictly local, but it does need it to be regular, in order for the measure $\Gamma(u, u)$ representing the Dirichlet energy of $u \in \mathcal{F}$ to exist, see [32] for more details. However we will always be considering strictly local forms.*

Example 2.10. Examples of strictly local Dirichlet spaces $(X, \mu, \mathcal{E}, \mathcal{F})$ that satisfy the volume doubling property and support the 2-Poincaré inequality include:

- Complete Riemannian manifolds with non-negative Ricci curvature or more generally $\mathrm{RCD}(0,\infty)$ spaces in the sense of Ambrosio-Gigli-Savaré [8],
- Carnot groups and other complete sub-Riemannian manifolds satisfying a generalized curvature dimension inequality (see [12, 17]),
- Doubling metric measure spaces that support a 2-Poincaré inequality with respect to the upper gradient structure of Heinonen and Koskela (see [39, 62, 63]).
- Metric graphs with bounded geometry (see [36]).

When the 2-Poincaré inequality is satisfied, a standard argument due to Semmes tells us that locally Lipschitz continuous functions form a dense subclass of $\mathcal{F}$, where $\mathcal{F}$ is equipped with the norm

$$\|u\|_{\mathcal{F}} := \|u\|_{L^2(X)} + \sqrt{\mathcal{E}(u,u)},$$

see for example [39, Theorem 8.2.1]. Moreover, by [61], we know that if the 2-Poincaré inequality is satisfied and μ is doubling then the Newton-Sobolev class (based on upper gradients, see [39]) is the same as the class $\mathcal{F}$, with comparable energy seminorms.

The next lemma is used to define a *length of the gradient* in the current setting and shows that the Dirichlet form admits a carré du champ operator.

Lemma 2.11. *Suppose that $(X, \mu, \mathcal{E}, \mathcal{F})$ satisfies the doubling property and supports the 2-Poincaré inequality. Then for all $u \in \mathcal{F}$, we have $d\Gamma(u,u) \ll \mu$ and we set $|\nabla u|^2$ to be the Radon-Nikodym derivative $\frac{d\Gamma(u,u)}{d\mu}$.*

Proof. Let $u \in \mathcal{F}$. Fix $\varepsilon > 0$. Let $\{B_i^\varepsilon = B(x_i^\varepsilon, \varepsilon)\}_i$ be a maximal ε-separated covering of X such that the family $\{2B_i^\varepsilon\}_i$ has bounded overlap property. Let φ_i^ε be a (C/ε)-Lipschitz partition of unity subordinated to this cover: that is, $0 \le \varphi_i^\varepsilon \le 1$ on X, $\sum_i \varphi_i^\varepsilon = 1$ on X, and $\varphi_i^\varepsilon = 0$ in $X \setminus B_i^\varepsilon$. We then set

$$u_\varepsilon := \sum_i u_{B_i^\varepsilon}\, \varphi_i^\varepsilon,$$

where $u_{B_i^\varepsilon} = \fint_{B_i^\varepsilon} u\, d\mu$. Then as each φ_i^ε is Lipschitz, we know that u_ε is locally Lipschitz and hence is in $\mathcal{F}_{\mathrm{loc}}(X)$. Indeed, for $x, y \in B_j^\varepsilon$ we see that from the 2-Poincaré inequality

$$\begin{aligned}
|u_\varepsilon(x) - u_\varepsilon(y)| &\le \sum_{i: 2B_i^\varepsilon \cap 2B_j^\varepsilon \neq \emptyset} |u_{B_i^\varepsilon} - u_{B_j^\varepsilon}||\varphi_i^\varepsilon(x) - \varphi_i^\varepsilon(y)| \\
&\le \frac{C\, d(x,y)}{\varepsilon} \sum_{i: 2B_i^\varepsilon \cap 2B_j^\varepsilon \neq \emptyset} \left(\fint_{B_i^\varepsilon} \fint_{B(x,2\varepsilon)} |u(y) - u(x)|^2 \, d\mu(y)\, d\mu(x) \right)^{1/2} \\
&\le \frac{C\, d(x,y)}{\varepsilon} \left(\fint_{6B_j^\varepsilon} |u(y) - u_{6B_j^\varepsilon}|^2 \, d\mu(y) \right)^{1/2} \\
&\le C\, d(x,y) \left(\fint_{6\lambda B_j^\varepsilon} d\Gamma(u,u) \right)^{1/2}.
\end{aligned}$$

It follows from Lemma 2.5 that $\Gamma(u_\varepsilon, u_\varepsilon) \ll \mu$ and the Radon-Nikodym measure is denoted by $|\nabla u_\varepsilon|^2$. Moreover, we also have on B_i^ε that

$$d\Gamma(u_\varepsilon, u_\varepsilon) \le C \left(\fint_{6\lambda B_i^\varepsilon} d\Gamma(u,u) \right) d\mu.$$

This yields

$$\int_X |\nabla u_\varepsilon|^2 d\mu = \mathcal{E}(u_\varepsilon, u_\varepsilon) \le \sum_i \int_{B_i} d\Gamma(u_\varepsilon, u_\varepsilon) \le C \sum_i \mu(B_i) \fint_{6\lambda B_i^\varepsilon} d\Gamma(u,u) \le C\mathcal{E}(u,u). \quad (3)$$

In the last inequality above we used the fact that μ is a doubling measure.

In a similar manner,

$$|u(x) - u_\varepsilon(x)| \le \sum_i |u(x) - u_{B_i^\varepsilon}| \, |\varphi_i^\varepsilon(x)| \le \sum_i |u(x) - u_{B_i^\varepsilon}| \, \mathbf{1}_{B_i^\varepsilon}(x).$$

Notice that the above sum has at most K terms due to the finite overlap property. Hence by the 2-Poincaré inequality

$$\int_X |u(x) - u_\varepsilon(x)|^2 \, d\mu(x) \le C \sum_i \int_{B_i^\varepsilon} |u(x) - u_{B_i^\varepsilon}|^2 d\mu \le C \sum_i \varepsilon^2 \int_{\lambda B_i^\varepsilon} d\Gamma(u,u) \le C\varepsilon^2 \int_X d\Gamma(u,u).$$

that is, $u_\varepsilon \to u$ in $L^2(X)$ as $\varepsilon \to 0^+$.

Take a sequence $\varepsilon_n \to 0^+$. From (3) and the reflexivity of $L^2(X)$, there exists a subsequence of $\{|\nabla u_{\varepsilon_n}|\}_n$ that is weakly convergent in $L^2(X)$. By Mazur's lemma, a sequence of convex combinations of u_{ε_n} (still denoted by $\{u_{\varepsilon_n}\}$) converges in the norm $\|\cdot\|_{\mathcal{F}}$. Since u_{ε_n} converges to u in $L^2(X)$ we see that $|\nabla u_{\varepsilon_n}|$ converges in $L^2(X)$ and denote the limit by $|\nabla u|$. At the same time, $\mathcal{E}(u_\varepsilon, u_\varepsilon)$ converges to $\mathcal{E}(u,u)$. We conclude that $\frac{d\Gamma(u,u)}{d\mu} = |\nabla u|^2$.

□

Definition 2.12. *Let $1 \le p < \infty$. We say that $(X, \mu, \mathcal{E}, \mathcal{F})$ supports a p-Poincaré inequality if there are constants $C > 0$ and $\lambda \ge 1$ such that whenever B is a ball in X (with respect to the metric $d_{\mathcal{E}}$) and $u \in \mathcal{F}$, we have*

$$\frac{1}{\mu(B)} \int_B |u - u_B| \, d\mu \le C \, \mathrm{rad}(B) \left(\frac{1}{\mu(\lambda B)} \int_{\lambda B} |\nabla u|^p \, d\mu \right)^{1/p}.$$

Of course, the p-Poincaré inequality for any $p \ne 2$ does not make sense if $\mathcal{E}$ does not satisfy the condition of strict locality. The requirement that $\mathcal{E}$ supports a 1-Poincaré inequality is a significantly stronger requirement than supporting a 2-Poincaré inequality.

Much of the current theory on functions of bounded variation in the metric setting requires a 1-Poincaré inequality. In this paper we will *not* require that $(X, \mu, \mathcal{E}, \mathcal{F})$ supports a 1-Poincaré inequality but only the weaker 2-Poincaré inequality. However in some of our analysis we will need an additional requirement called the weak Bakry-Émery curvature condition.

2.3 Sobolev classes $W^{1,p}(X)$

The theory of Sobolev spaces was first advanced in order to prove solvability of certain PDEs, see for example [30, 70]. When X is a Riemannian manifold, a function $f \in L^p(X)$ is said to be in the Sobolev class $W^{1,p}(X)$ if its distributional derivative is given by a vector-valued function $\nabla f \in L^p(X : \mathbb{R}^n)$. Extensions of this idea to sub-Riemannian spaces have been considered in [33]. However, in more general metric spaces where the distributional theory of derivatives (which relies on integration by parts) is unavailable, an alternate notion of derivatives needs to be found. Indeed, we do not need an alternative to ∇f, as long as we have a substitute for $|\nabla f|$. For metric spaces X, Lipschitz functions $f : X \to \mathbb{R}$ have a natural such alternative, $\mathrm{Lip} f$, given by

$$\mathrm{Lip} f(x) := \limsup_{r \to 0^+} \sup_{y \in B(x,r)} \frac{|f(y) - f(x)|}{r}.$$

Other notions such as upper gradients and Hajłasz gradients play this substitute role well, see for example [39]. In the current paper we consider another possible notion of $|\nabla f|$ which has a more natural affinity to the heat semigroup and the Dirichlet form, as in Lemma 2.11. So, in this paper, our definition of $W^{1,p}(X)$, $p \geq 1$ is the following:

$$W^{1,p}(X) = \{u \in L^p(X) \cap \mathcal{F}_{\mathrm{loc}}(X) \,:\, \Gamma(u,u) \ll \mu, |\nabla u| \in L^p(X)\} \,. \tag{4}$$

The norm on $W^{1,p}(X)$ is then given by

$$\|u\|_{W^{1,p}(X)} = \|u\|_{L^p(X)} + \||\nabla u|\|_{L^p(X)}.$$

Note, in particular, that $W^{1,2}(X) = \mathcal{F}$. In the context of Sobolev spaces, Besov function classes arise naturally in two ways. Given a Sobolev class $W^{1,p}(\mathbb{R}^{n+1})$ and a bi-Lipschitz embedding of $\mathbb{R}^n$ into $\mathbb{R}^{n+1}$, there is a natural trace of functions in $W^{1,p}(\mathbb{R}^{n+1})$ to the embedded surface, and this trace belongs to a Besov class, see for example [46,47]. Besov classes also arise via real interpolations of $L^p(\mathbb{R}^n)$ and $W^{1,p}(\mathbb{R}^n)$, see for example [18,87]. In the present paper we will relate Sobolev classes $W^{1,p}(X)$ to two types of Besov classes defined in our previous paper [3], see Theorems 4.11, 4.17, and 4.9. One of these types of Besov classes is defined from the heat semigroup, while the other uses only the metric structure of X. We note that previous metric characterizations of Sobolev spaces in the presence of doubling and 2-Poincaré have been studied in [27].

2.4 Bakry-Émery curvature conditions

Let $\{P_t\}_{t\in[0,\infty)}$ denote the self-adjoint semigroup of contractions on $L^2(X,\mu)$ associated with the Dirichlet space $(X,\mu,\mathcal{E},\mathcal{F})$ and L the infinitesimal generator of $\{P_t\}_{t\in[0,\infty)}$. The semigroup $\{P_t\}_{t\in[0,\infty)}$ is referred to as the heat semigroup on $(X,\mu,\mathcal{E},\mathcal{F})$. For classical properties of $\{P_t\}_{t\in[0,\infty)}$, we refer to Section 2.2 in [3]. It is known that that doubling property together with the 2-Poincaré inequality imply that the semigroup $\{P_t\}$ is conservative, i.e. $P_t 1 = 1$.

The work of Sturm [82, 84] (see Saloff-Coste [77] and Grigor'yan [35] for earlier results on Riemannian manifolds) tells us that doubling property together with the 2-Poincaré inequality are equivalent to the property that the heat semigroup P_t admits a heat kernel function $p_t(x,y)$ on $[0,\infty)\times X\times X$ for which there are constants $c_1, c_2, C > 0$ such that whenever $t > 0$ and $x, y \in X$,

$$\frac{1}{C}\,\frac{e^{-c_1 d(x,y)^2/t}}{\sqrt{\mu(B(x,\sqrt{t}))\mu(B(y,\sqrt{t}))}} \leq p_t(x,y) \leq C\,\frac{e^{-c_2 d(x,y)^2/t}}{\sqrt{\mu(B(x,\sqrt{t}))\mu(B(y,\sqrt{t}))}}. \tag{5}$$

The above inequalities are called *Gaussian bounds* for the heat kernel. Due to the doubling property, one can equivalently rewrite the Gaussian bounds as:

$$\frac{1}{C}\,\frac{e^{-c_1 d(x,y)^2/t}}{\mu(B(x,\sqrt{t}))} \leq p_t(x,y) \leq C\,\frac{e^{-c_2 d(x,y)^2/t}}{\mu(B(x,\sqrt{t}))}, \tag{6}$$

for some different constants $c_1, c_2, C > 0$. The combination of the doubling property and the 2-Poincaré inequality also implies the following Hölder regularity of the heat kernel

$$|p_t(x,y) - p_t(z,y)| \leq \left(\frac{d(x,z)}{\sqrt{t}}\right)^{\alpha} \frac{C}{\mu(B(y,\sqrt{t}))},$$

for some $C > 0$, $\alpha \in (0,1)$, and all $x, y, z \in X$ (see for instance [78]). In some parts of this paper, we need a stronger condition than Hölder regularity for the heat kernel, in which case we will use the following uniform Lipschitz continuity property.

Definition 2.13. *We say that the Dirichlet metric space* $(X, \mathcal{E}, d_{\mathcal{E}}, \mu)$ *satisfies a weak Bakry-Émery curvature condition if, whenever* $u \in \mathcal{F} \cap L^\infty(X)$ *and* $t > 0$*,*

$$\||\nabla P_t u|\|^2_{L^\infty(X)} \leq \frac{C}{t} \|u\|^2_{L^\infty(X)}. \tag{7}$$

We refer to (7) as a weak Bakry-Émery curvature condition because, in many settings, its validity is related to the existence of curvature lower bounds on the underlying space.

Example 2.14. The weak Bakry-Émery curvature condition is satisfied in the following examples:

- Complete Riemannian manifolds with non-negative Ricci curvature and more generally, the $RCD(0, +\infty)$ spaces (see [45]).
- Carnot groups (see [13])
- Complete sub-Riemannian manifolds with generalized non-negative Ricci curvature (see [12, 17])
- On non-compact metric graphs with finite number of edges, the weak weak Bakry-Émery curvature condition has been proved to hold for $t \in (0, 1]$ (see [16, Theorem 5.4]), and is conjectured to be true for all t. If the graph is moreover compact, the weak Bakry-Émery estimate holds for every $t > 0$ [16, Theorem 5.4]) .

Several statements equivalent to the weak Bakry-Émery curvature condition are given in [25, Theorem 1.2]. There are some metric measure spaces equipped with a doubling measure supporting a 2-Poincaré inequality but without the above weak Bakry-Émery condition, see for example [60]. For instance, it should be noted, that in the setting of complete sub-Riemannian manifolds with generalized non-negative Ricci curvature in the sense of [15], while the weak Bakry-Émery curvature condition is known to be satisfied (see [12, 17]), the 1-Poincaré inequality is so far not known to hold, though the 2-Poincaré inequality is known to be always satisfied, see [14].

We will also sometimes need a stronger condition than (7).

Definition 2.15. *We say that the Dirichlet metric space* $(X, \mathcal{E}, d_{\mathcal{E}}, \mu)$ *satisfies a strong Bakry-Émery curvature condition if there exists a constant* $C > 0$ *such that for every* $u \in \mathcal{F}$ *and* $t \geq 0$ *we have* μ *a.e.*

$$|\nabla P_t u| \leq C P_t |\nabla u|. \tag{8}$$

The strong Bakry-Émery curvature condition implies the weak one, as is demonstrated in the proof of Theorem 3.3 in [16]. Examples where the strong Bakry-Émery estimate is satisfied include: Riemannian manifolds with non negative Ricci curvature and more generally $RCD(0, +\infty)$ spaces (in that case $C = 1$, see [29]), some metric graphs like the Walsh spider (see [16, Example 5.1] and also [16, Theorem 5.4])), the Heisenberg group and more generally H-type groups (see [10, 28]).

3 BV class and co-area formula

In this section we use the Dirichlet form and the associated family $\Gamma(\cdot, \cdot)$ of measures to construct a BV class of functions on X. To do so, we only need μ to be a doubling measure on X for $d_{\mathcal{E}}$ and the class of locally Lipschitz functions to be dense in $L^1(X)$. So in this section we do *not* need the 2-Poincaré inequality nor do we need the weak Bakry-Émery curvature condition. In the second part of the section we prove a co-area formula for BV functions; such a co-area formula is highly useful in understanding the structure of BV functions, and underscores the importance of studying sets of finite perimeter (sets whose characteristic functions are BV functions).

3.1 BV class

We set the *core* of the Dirichlet form, $\mathcal{C}(X)$, to be the class of all $f \in \mathcal{F}_{\rm loc}(X) \cap C(X)$ such that $\Gamma(f,f) \ll \mu$ and recall that the Sobolev class $W^{1,1}(X)$ is the class of all $f \in \mathcal{F}_{\rm loc}(X) \cap L^1(X)$ for which $\Gamma(f,f) \ll \mu$ and $|\nabla f| \in L^1(X)$ (see Definition (4)).

Definition 3.1. *We say that $u \in L^1(X)$ is in $BV(X)$ if there is a sequence of local Lipschitz functions $u_k \in L^1(X)$ such that $u_k \to u$ in $L^1(X)$ and*

$$\liminf_{k\to\infty} \int_X |\nabla u_k|\, d\mu < \infty.$$

We note that if the Dirichlet form supports a 1-Poincaré inequality, then the Sobolev space $W^{1,1}(X)$ is a subspace of $BV(X)$.

Definition 3.2. *For $u \in BV(X)$ and open sets $U \subset X$, we set*

$$\|Du\|(U) = \inf_{u_k \in \mathcal{C}(U), u_k \to u \text{ in } L^1(U)} \liminf_{k\to\infty} \int_U |\nabla u_k|\, d\mu.$$

We will see in the next part of this section that $\|Du\|$ can be extended from the collection of open sets to the collection of all Borel sets as a Radon measure, see Definition 3.5.

Lemma 3.3. *If $u, v \in BV(X)$ and η is a Lipschitz continuous function on X with $0 \le \eta \le 1$ on X, then $\eta u + (1-\eta)v \in BV(X)$ with*

$$\|D(\eta u + (1-\eta)v)\|(X) \le \|Du\|(X) + \|Dv\|(X) + \int_X |u-v|\, |\nabla \eta|\, d\mu.$$

Proof. From Lemma 2.5 we already know that such η are in $\mathcal{F}_{\rm loc}(X)$ with $|\nabla\eta| \in L^\infty(X)$. From the definition, we can choose sequences $u_k, v_k \in L^1(X)$ of locally Lipschitz functions on X such that $u_k \to u$ and $v_k \to v$ in $L^1(X)$ and $\int_X |\nabla u_k|\, d\mu \to \|Du\|(X)$ and $\int_X |\nabla v_k|\, d\mu \to \|Dv\|(X)$ as $k \to \infty$. Now an application of the Leibniz rule to the functions $\eta u_k + (1-\eta)v_k$ tells us that

$$\begin{aligned}\|D(\eta u + (1-\eta)v)\|(X) &\le \liminf_{k\to\infty} \int_X |\nabla[\eta u_k + (1-\eta)v_k]|\, d\mu \\ &\le \liminf_{k\to\infty} \left(\int_X \eta |\nabla u_k|\, d\mu + \int_X (1-\eta)|\nabla v_k|\, d\mu + \int_X |u_k - v_k|\, |\nabla \eta|\, d\mu \right).\end{aligned}$$

Now using $0 \le \eta \le 1$ and $u_k - v_k \to u - v$ in $L^1(X)$ we obtain the required inequality. ☐

We now establish some elementary properties of $\|Du\|$.

Lemma 3.4. *Let U and V be two open subsets of X. If $u \in BV(X)$, then*

1. $\|Du\|(\emptyset) = 0$,
2. $\|Du\|(U) \le \|Du\|(V)$ *if* $U \subset V$,
3. $\|Du\|(\bigcup_i U_i) = \sum_i \|Du\|(U_i)$ *if* $\{U_i\}_i$ *is a pairwise disjoint subfamily of open subsets of* X.

Proof. We will only prove the third property here, as the other two are quite direct consequences of the definition of $\|Du\|$. Since any function $f \in \mathcal{F}(\bigcup_i U_i)$ has restrictions $u_i = f|_{U_i} \in \mathcal{F}(U_i)$ with $\int_{\bigcup_i U_i} |\nabla f|\, d\mu = \sum_i \int_{U_i} |\nabla u_i|\, d\mu$, it follows that

$$\|Du\|(\bigcup_i U_i) \geq \sum_i \|Du\|(U_i).$$

In the above we also used the fact that as f gets closer to u in the $L^1(\bigcup_i U_i)$ sense, u_i gets closer to u in the $L^1(U_i)$ sense.

To prove the reverse inequality, for $\varepsilon > 0$ we can choose locally Lipschitz continuous $u_i \in \mathcal{F}(U_i)$ for each i such that

$$\int_{U_i} |u - u_i|\, d\mu < 2^{-i-2}\varepsilon$$

and

$$\int_{U_i} |\nabla u_i|\, d\mu < \|Du\|(U_i) + 2^{-i-2}\varepsilon.$$

Now the function $f_\varepsilon = \sum_i u_i \mathbf{1}_{U_i}$ is in $\mathcal{F}(\bigcup_i U_i)$ because the U_i are pairwise disjoint open sets, and $\mathcal{E}$ is local. Therefore

$$\int_{\bigcup_i U_i} |u - f_\varepsilon|\, d\mu \leq \sum_i \int_{U_i} |u - u_i|\, d\mu \leq \frac{\varepsilon}{2}$$

and

$$\int_{\bigcup_i U_i} |\nabla f_\varepsilon|\, d\mu = \sum_i \int_{U_i} |\nabla u_i|\, d\mu \leq \frac{\varepsilon}{2} + \sum_i \|Df\|(U_i).$$

From the first of the above two inequalities it follows that $\lim_{\varepsilon\to 0^+} f_\varepsilon = u$ in $L^1(\bigcup_i U_i)$, and therefore

$$\|Du\|(\bigcup_i U_i) \leq \liminf_{\varepsilon\to 0^+} \left(\frac{\varepsilon}{2} + \sum_i \|Du\|(U_i)\right) = \sum_i \|Du\|(U_i). \qquad \square$$

We use the above definition of $\|Du\|$ on open sets to consider the following Caratheodory construction.

Definition 3.5. *For $A \subset X$, we set*

$$\|Du\|^*(A) := \inf\{\|Du\|(O) \,:\, O \text{ is an open subset of } X, A \subset O\}.$$

By the second property listed in the above lemma, we note that if A is an open subset of X, then $\|Du\|^*(A) = \|Du\|(A)$. With this observation, we re-name $\|Du\|^*(A)$ as $\|Du\|(A)$ even when A is not open.

We end this section by proving that $\|Du\|$, as constructed above, is a Radon measure on X. The idea of the proof is from [72]. The principal tool used in the proof is the following lemma due to De Giorgi and Letta [26, Theorem 5.1], see also [7, Theorem 1.53].

Lemma 3.6 ([26, Theorem 5.1])**.** *If ν is a non-negative function on the class of all open subsets of X such that for open sets U_1, U_2*

1. $\nu(\emptyset) = 0$,
2. *if* $U_1 \subset U_2$ *then* $\nu(U_1) \leq \nu(U_2)$,
3. $\nu(U_1 \cup U_2) \leq \nu(U_1) + \nu(U_2)$,

4. *if $U_1 \cap U_2$ is empty then $\nu(U_1 \cup U_2) = \nu(U_1) + \nu(U_2)$,*

5. $\nu(U_1) = \sup\{\nu(V) : V$ *is bounded and open in X with $\overline{V} \subset U_1$*$\}$.

Then the Carathéodory extension of ν to all subsets of X is a Borel regular outer measure on X.

Theorem 3.7. *If $f \in BV(X)$, then $\|Df\|$ is a Radon outer measure on X and the restriction of $\|Df\|$ to the Borel sigma algebra is a Radon measure which is the weak limit of $\|Du_k\|$ for some sequence u_k of locally Lipschitz functions in $L^1(X)$ such that $u_k \to f$ in $L^1(X)$.*

Proof. For simplicity of notation we will assume that X is itself bounded. Thanks to the lemma of De Giorgi and Letta (Lemma 3.6), it suffices to verify that $\|Du\|$ satisfies the five conditions set forth in Lemma 3.6. By Lemma 3.4, we know that $\|Du\|$ satisfies Conditions 1, 2 and 4. Thus it suffices for us to verify Condition 3 and Condition 5. We will first show the validity of Condition 5, and use it (or rather, its proof) to show that Condition 3 holds. We will do so for bounded open subsets of X. A simple modification (by truncating U_δ by balls) would complete the proof for unbounded sets and we leave this part of the extension to the interested reader.

Proof of Condition 5: From the monotonicity condition 2, it suffices to prove that

$$\|Df\|(U) \le \sup\{\|Df\|(V) : V \text{ is open in } X, \overline{V} \text{ is a compact subset of } U\}.$$

For $\delta > 0$ we set

$$U_\delta = \{x \in U : \operatorname{dist}(x, X \setminus U) > \delta\}.$$

For $0 < \delta_1 < \delta_2 < \operatorname{diam}(U)/2$, let $V = U_{\delta_1}$ and $W = U \setminus \overline{U_{\delta_2}}$. Then V and W are open subsets of U, and the closure of V is a compact subset of U (recall that we assume from Assumption 2.7 that X is complete, and hence as μ is doubling with respect to $d_{\mathcal{E}}$ we know that closed and bounded subsets of X are compact). Note also that $U = V \cup W$ and that $\partial V \cap \partial W$ is empty. Thus we can find a Lipschitz function η on U that can be used as a "needle and thread" to stitch Sobolev functions on V to Sobolev functions on W to obtain a Sobolev function on U as follows: take η with $0 \le \eta \le 1$ on U, $\eta = 1$ on $V \setminus W = \overline{U_{\delta_2}}$, $\eta = 0$ on $W \setminus V = U \setminus U_{\delta_1}$, and

$$\operatorname{Lip} \eta \le \frac{2}{\delta_2 - \delta_1} \mathbf{1}_{V \cap W}.$$

Now, for $v \in \mathcal{F}(V)$ and $w \in \mathcal{F}(W)$ we set $u = \eta v + (1-\eta) w$. As we have the Leibniz rule (see [83]), we can see that $u \in \mathcal{F}(U)$ and

$$\int_U |\nabla u| \, d\mu \le \int_V |\nabla v| \, d\mu + \int_W |\nabla w| \, d\mu + \frac{2}{\delta_2 - \delta_1} \int_{V \cap W} |v - w| \, d\mu. \tag{9}$$

Furthermore, whenever $h \in L^1(U)$, we can write $h = \eta h + (1-\eta) h$ to see that

$$\int_U |u - h| \, d\mu \le \int_V |v - h| \, d\mu + \int_W |w - h| \, d\mu. \tag{10}$$

Now, we take v_k from $\mathcal{F}(V)$ such that $v_k \to f$ in $L^1(V)$ and $\lim_{k\to\infty} \int_V |\nabla v_k| \, d\mu = \|Df\|(V)$, and take $w_k \in \mathcal{F}(W)$ analogously. We then follow through by stitching together v_k and w_k into the function u_k as prescribed above. By (10) with $h = f$, we have that

$$\int_U |f - u_k| \, d\mu \le \int_V |v_k - f| \, d\mu + \int_W |w_k - f| \, d\mu \to 0 \text{ as } k \to \infty.$$

It follows from (9) and the fact $\int_{V\cap W} |v_k - w_k|\, d\mu \to 0$ as $k \to \infty$ that

$$\|Df\|(U) \le \liminf_{k\to\infty} \int_U |\nabla u_k|\, d\mu \le \|Df\|(V) + \|Df\|(W).$$

Remembering again that the closure of V is a compact subset of U, we see that

$$\|Df\|(U) \le \sup\{\|Df\|(V) \,:\, V \text{ is open in } X, \overline{V} \text{ is compact subset of } U\} + \|Df\|(U \setminus \overline{U_{\delta_2}}).$$

So now it suffices to prove that

$$\lim_{\delta\to 0^+} \|Df\|(U \setminus \overline{U_\delta}) = 0. \tag{11}$$

To prove this, we note first that the above limit exists as $\|Df\|(U \setminus \overline{U_\delta})$ decreases as δ decreases. We fix a strictly monotone decreasing sequence of real numbers δ_k with $\lim_{k\to\infty} \delta_k = 0$, and for $k \ge 2$ we set $V_k := U_{\delta_{2k+3}} \setminus \overline{U_{\delta_{2k}}}$. Observe that each of $\{V_{2k}\}_k$ and $\{V_{2k+1}\}_k$ is a pairwise disjoint families open subsets of X

By Lemma 3.4, we know that

$$\sum_{k=1}^{\infty} \|Df\|(V_{2k}) = \|Df\|(\bigcup_{k\ge 1} V_{2k}) \le \|Df\|(U) < \infty,$$

and

$$\sum_{k=1}^{\infty} \|Df\|(V_{2k+1}) = \|Df\|(\bigcup_{k\ge 1} V_{2k+1}) \le \|Df\|(U) < \infty.$$

It follows that for $\varepsilon > 0$ there is some positive integer $k_\varepsilon \ge 2$ such that

$$\sum_{k=k_\varepsilon}^{\infty} \|Df\|(V_{2k}) + \sum_{k=k_\varepsilon}^{\infty} \|Df\|(V_{2k+1}) < \varepsilon.$$

Now we stitch together approximations on V_{2k} to approximations on V_{2k+1}, and from there to V_{2k+2} and so on. For each k we choose a "stitching function" η_k as a Lipschitz function on $\bigcup_{j=k_\varepsilon}^{k+1} V_j$ such that $0 \le \eta_k \le 1$, with $\eta_k = 1$ on $V_k \setminus V_{k-1}$, $\eta_k = 0$ on $\bigcup_{j=k_\varepsilon}^{k-1} V_j \setminus V_k$, and $|\nabla \eta_k| \le C_k \mathbf{1}_{V_k \cap V_{k-1}}$.

Next, for each k we can find $v_{k,j} \in \mathcal{F}(V_k)$ such that

$$\int_{V_k} |v_{k,j} - f|\, d\mu \le \frac{2^{-k-j}}{3(1 + C_k)}$$

and

$$\int_{V_k} |\nabla v_{k,j}|\, d\mu \le \|Df\|(V_k) + 2^{-j-k}.$$

We now inductively stitch the functions together. To do so, we first fix $i \in \mathbb{N}$.

Starting with $k = k_\varepsilon$, we stitch $u_{k,i}$ to $u_{k+1,i}$ using $\eta_{k+1} = \eta_{k_\varepsilon+1}$ to obtain $w_{i,k} \in \mathcal{F}(V_{k_\varepsilon} \cup V_{k_\varepsilon+1})$ so that we have

$$\int_{V_{k_\varepsilon} \cup V_{k_\varepsilon+1}} |w_{i,k} - f|\, d\mu \le \frac{2^{-i-k_\varepsilon}}{1 + C_{k_\varepsilon+1}}$$

and

$$\int_{V_{k_\varepsilon} \cup V_{k_\varepsilon+1}} |\nabla w_{i,k}|\, d\mu \le \sum_{j=k_\varepsilon}^{k_\varepsilon+1} \|Df\|(V_j) + 2^{1-i-k_\varepsilon}.$$

Suppose now that for some $k \in \mathbb{N}$ with $k \geq k_\varepsilon + 1$ we have constructed $w_{i,k} \in \mathcal{F}(\bigcup_{j=k_\varepsilon}^k V_j)$ such that

$$\int_{\bigcup_{j=k_\varepsilon}^k V_j} |w_{i,k} - f| \, d\mu \leq \sum_{k=k_\varepsilon}^{k} \frac{2^{-i-j}}{1+C_j}$$

and

$$\int_{\bigcup_{j=k_\varepsilon}^k V_j} |\nabla w_{i,k}| \, d\mu \leq \sum_{j=k_\varepsilon}^{k} \left(\|Df\|(V_j) + 2^{1-i-j} \right).$$

Then we stitch $u_{k+1,i}$ to $w_{i,k}$ using η_{k+1} to obtain $w_{i,k+1}$ satisfying inequalities analogous to the above two. Note that $w_{i,k+1} = w_{i,k-1}$ on V_{k-1} for $k \geq k_\varepsilon + 2$. Thus, in the limit, we obtain a function $w_i = \lim_{k\to\infty} w_{i,k} \in \mathcal{F}(\bigcup_{k=k_\varepsilon}^\infty V_k)$ satisfying

$$\int_{\bigcup_{j=k_\varepsilon}^\infty V_j} |w_i - f| \, d\mu \leq \sum_{k=k_\varepsilon}^{k} \frac{2^{-i-j}}{1+C_j} < 2^{1-i},$$

$$\int_{\bigcup_{j=k_\varepsilon}^\infty V_j} |\nabla w_i| \, d\mu \leq \sum_{j=k_\varepsilon}^{\infty} \|Df\|(V_j) + 2^{2-i} < \varepsilon + 2^{2-i}.$$

From the first of the above two inequalities, we see that $w_i \to f$ in $L^1(\bigcup_{j=k_\varepsilon}^\infty V_j)$ as $i \to \infty$, and so from the second of the above two inequalities we obtain

$$\|Df\|(\bigcup_{j=k_\varepsilon}^{\infty} V_j) = \|Df\|(U \setminus \overline{U_{\delta_{k_\varepsilon}}}) \leq \liminf_{i\to\infty} \int_{\bigcup_{j=k_\varepsilon}^\infty V_j} |\nabla w_i| \, d\mu \leq \varepsilon.$$

The last inequality above tells us that the claim we set out to prove, namely

$$\lim_{\delta\to 0^+} \|Df\|(U \setminus \overline{U_\delta}) = 0.$$

This completes the proof of Condition 5.

Proof of Condition 3: By Condition 5, which was proved above, for each $\varepsilon > 0$ we can find relatively compact open subsets $U_1' \Subset U_1$ and $U_2' \Subset U_2$ such that $\|Df\|(U_1 \cup U_2) \leq \|Df\|(U_1' \cup U_2') + \varepsilon$. We then choose a Lipschitz "stitching function" η on X such that $0 \leq \eta \leq 1$ on X, $\eta = 1$ on U_1', $\eta = 0$ on $X \setminus U_1$, and

$$|\nabla \eta| \leq \frac{1}{C_{U_1,U_1'}} \mathbf{1}_{U_1 \setminus U_1'}.$$

For $u_1 \in \mathcal{F}(U_1)$ and $u_2 \in \mathcal{F}(U_2)$, we obtain the stitched function $w = \eta u_1 + (1-\eta) u_2$ and note that $w \in \mathcal{F}(U_1' \cup U_2')$. Observe that we cannot in general have $w \in \mathcal{F}(U_1 \cup U_2)$, as w is not defined in $U_1 \setminus (U_1' \cup U_2)$ because $1 - \eta$ is non-vanishing and u_2 is not defined there. Then we have

$$\int_{U_1' \cup U_2'} |\nabla w| \, d\mu \leq \int_{U_1} |\nabla u_1| \, d\mu + \int_{U_2} |\nabla u_2 \, d\mu + \frac{1}{C_{U_1,U_1'}} \int_{U_1 \cap U_2} |u_1 - u_2| \, d\mu$$

and

$$\int_{U_1' \cup U_2'} |w - f| \, d\mu \leq \int_{U_1} |u_1 - f| \, d\mu + \int_{U_2} |u_2 - f| \, d\mu.$$

As before, choosing $u_{1,k}$ to be the optimal approximating sequence for f on U_1 and $u_{2,k}$ correspondingly for f on U_2, we see from the first of the above two inequalities that the stitched sequence w_k approximates f on $U_1' \cup U_2'$. Therefore we obtain

$$\|Df\|(U_1 \cup U_2) \leq \varepsilon + \|Df\|(U_1' \cup U_2') \leq \varepsilon + \liminf_{k\to\infty} \int_{U_1 \cup U_2} |\nabla w_k| \, d\mu \leq \|Df\|(U_1) + \|Df\|(U_2) + \varepsilon.$$

Letting $\varepsilon \to 0$ now gives the desired Condition 3.

Proof of weak convergence: Consider an optimal sequence for the convergence from Definition 3.2, that is, u_k is a sequence of locally Lipschitz functions on X such that $u_k \to u$ in $L^1(X)$ and $\lim_k \int_X |\nabla u_k| \, d\mu = \|Du\|(X)$. Then for each open set $U \subset X$ we have that $\|Du\|(U) \leq \liminf_k \int_U |\nabla u_k| \, d\mu$. We can choose a subsequence of u_k so that $|\nabla u_k| \, d\mu$ has a weak limit, say ν, on X.

Now by weak limits, we have that $\nu(X) = \|Du\|(X)$ and for open sets $U \subset X$ for which $\nu(\partial U) = 0$ we have that $\nu(U) = \lim_k \int_U |\nabla u_k| \, d\mu \geq \|Du\|(U)$. Then

$$\|Du\|(X \setminus U) = \|Du\|(X) - \|Du\|(U) \geq \nu(X \setminus U).$$

On the other hand, we can approximate $X \setminus U$ by open sets W_j with $\nu(\partial W_j) = 0$ and $X \setminus U \subset W_j$, that is, $\|Du\|(X \setminus U) = \lim_j \|Du\|(W_j)$; the above then tell us that $\nu(W_j) \geq \|Du\|(W_j)$, and hence $\nu(X \setminus U) \geq \|Du\|(X \setminus U)$. Thus we obtain $\|Du\|(X \setminus U) = \nu(X \setminus U)$, and therefore $\|Du\|(U) = \nu(U)$ because $\|Du\|(X) = \nu(X)$. Thus we conclude that whenever $U \subset X$ is an open set with $\nu(\partial U) = 0$, we have $\|Du\|(U) = \nu(U)$. Since every open set can be approximated by open sets O with $\nu(\partial O) = 0$, we have $\|Du\| = \nu$. ☐

Example 3.8. In the context of a doubling metric measure space (X, d, μ) supporting a 2-Poincaré inequality, where the Dirichlet form is given in terms of a Cheeger differential structure (see Example 2.3), the construction of $BV(X)$ and $\|Df\|$ is due to M. Miranda [72]. When applied to Riemannian or sub-Riemannian spaces, it yields the usual notion of variation (see [72]).

Example 3.9. There is a large class of fractal examples [54, 55, 65, 86] with resistance forms $\mathcal{E}$, a so-called Kusuoka measure μ, and a base of open sets O with finite boundaries, such that $1_O \in BV(X)$ and $\|D1_O\|$ is absolutely continuous with respect to the counting measure on ∂O. Among these examples, the most notable are the Sierpinski gasket in harmonic coordinates [31, 49, 50, 56, 63, 71, 85], fractal quantum graphs [5] and diamond fractals [2, and references therein]. In particular, on diamond fractals [2] provides explicit formulas for the heat kernel, which allow for many computations relevant to our paper. On the Sierpinski gasket [85, Proposition 4.14] shows how to make computations at the dense set of junction points. One might expect that if $u \in BV(X)$ then, following [41, 42], Du could be defined as a vector valued Borel measure, however the details of this construction are outside of the scope of this article. The long term motivation for this type of analysis comes from stochastic PDEs, see [11, 44, 51, 52, 73] and the references therein.

3.2 Co-area formula

The goal of this subsection is to prove a co-area formula that connects the BV energy seminorm of a BV function with the perimeter measure of its super-level sets.

Definition 3.10. *A function u is said to be in $BV_{\text{loc}}(X)$ if for each bounded open set $U \subset X$ there is a compactly supported Lipschitz function η_U on X such that $\eta_U = 1$ on U and $\eta_U \, u \in BV(X)$. We say that a measurable set $E \subset X$ is of* finite perimeter *if $\mathbf{1}_E \in BV_{\text{loc}}(X)$ with $\|D\mathbf{1}_E\|(X) < \infty$. For any Borel set $A \subset X$, we denote by $P(E, A) := \|D\mathbf{1}_E\|(A)$ the perimeter measure of E.*

Theorem 3.11. *The co-area formula holds true, that is, for Borel sets $A \subset X$ and $u \in L^1_{\mathrm{loc}}(X)$,*

$$\|Du\|(A) = \int_{\mathbb{R}} P(\{u > s\}, A)\, ds.$$

Proof. We first prove the formula for open sets A.

Suppose $u \in BV_{\mathrm{loc}}(X)$ with $\|Du\|(A) < \infty$. For $s \in \mathbb{R}$ let

$$E_s := \{x \in X \,:\, u(x) > s\}.$$

The set E_s is denoted by the abbreviation $\{u > s\}$ in the statement of the theorem. Consider the function $m : \mathbb{R} \to \mathbb{R}$ given by

$$m(t) = \|Du\|(A \cap E_t).$$

Then m is a monotone decreasing function, and hence is differentiable almost everywhere. Let $t \in \mathbb{R}$ such that $m'(t)$ exists. Then

$$m'(t) = \lim_{h \to 0^+} \frac{\|Du\|(A \cap E_t \setminus E_{t+h})}{h}.$$

Note that the functions

$$u_{t,h} := \frac{\max\{t, \min\{t+h, u\}\} - t}{h}$$

converge in $L^1(X)$ to $\mathbf{1}_{E_t}$ as $h \to 0^+$. Using the fact that A is open, it follows that

$$P(E_t, A) \le \liminf_{h \to 0^+} \|Du_{t,h}\|(A) = \liminf_{h \to 0^+} \frac{\|Du\|(A \cap E_t \setminus E_{t+h})}{h} = m'(t).$$

Note also that by this lower semicontinuity of BV energy, $t \mapsto P(E_t, A)$ is a lower semicontinuous function, and hence is measurable; and as it is non-negative, we can talk about its integral, whether that integral is finite or not. Therefore, by the fundamental theorem of calculus for monotone functions,

$$\int_{\mathbb{R}} P(E_t, A)\, dt \le \int_{\mathbb{R}} m'(t)\, dt \le \lim_{s,\tau \to \infty} (m(s) - m(-\tau)) = \|Du\|(A).$$

The above in particular tells us that if $u \in BV_{\mathrm{loc}}(X)$ then almost all of its superlevel sets E_t have finite perimeter. If u is not a BV function on A, then $\|Du\|(A) = \infty$, and hence we also have

$$\int_{\mathbb{R}} P(E_t, A)\, dt \le \|Du\|(A). \tag{12}$$

In particular, it also follows that $\int_{\mathbb{R}} P(E_t, A)\, dt < \infty$ if $u \in BV(X)$.

We continue to assume that A is open, and prove the reverse of the above inequality. If $\int_{\mathbb{R}} P(E_t, A)\, dt = \infty$, then trivially

$$\|Du\|(A) \le \int_{\mathbb{R}} P(E_t, A)\, dt.$$

So we may assume without loss of generality that $\int_{\mathbb{R}} P(E_t, A)\, dt$ is finite. Note also by the Markovian property of Dirichlet forms, filtered down to the level of the measure $|\nabla f|$, we have that

$$\|Du\|(A) = \lim_{s,\tau \to \infty} \|Du_{s,\tau}\|(A),$$

where $u_{s,\tau} = \max\{-\tau, \min\{u,s\}\}$. So without loss of generality we may assume that $a \le u \le b$ for some finite $a, b \in \mathbb{R}$. For positive integers k we can divide $[0,1]$ into k equal sub-intervals $[t_i, t_{i+1}]$, $i = 0, \cdots, k-1$ with $t_{i+1} - t_i = 1/k$. Then we can find $\rho_{k,i} \in (t_i, t_{i+1})$ such that

$$\frac{1}{k} P(E_{\rho_{k,i}}, A) \le \int_{t_i}^{t_{i+1}} P(E_s, A)\, ds.$$

We set

$$u_k = \frac{1}{k} \sum_{j=1}^{k} \mathbf{1}_{E_{\rho_{k,i}}}.$$

Then as $|u_k - u| \le 1/k$ on X, we have that $u_k \to u$ in $L^1(A)$ as $k \to \infty$, and so

$$\|Du\|(A) \le \liminf_{k\to\infty} \|Du_k\|(A) = \liminf_{k\to\infty} \sum_{j=1}^{k} \frac{1}{k} P(E_{\rho_{k,i}}, A) \le \int_0^1 P(E_s, A)\, ds. \tag{13}$$

Note now that by the proofs of inequalities (12) and (13), if A is an open set then $u \in BV(A)$ if and only if $\int_{\mathbb{R}} P(E_t, A)\, dt$ is finite.

Finally, we remove the requirement that A be open. By the above comment, it suffices to prove this for the case that $u \in BV(X)$. In this case, the maps $A \mapsto \|Du\|(A)$ and $A \mapsto \int_{\mathbb{R}} P(E_t, A)\, dt$ are both Radon measures on X that agree on open subsets of X (that is, they are equal for open A). Hence it follows that they agree on Borel subsets of X. This completes the proof of the coarea formula. □

4 BV, Sobolev and heat semigroup-based Besov classes

Throughout the section, let $(X, \mu, \mathcal{E}, d_{\mathcal{E}}, \mathcal{F})$ be a strictly local regular Dirichlet space that satisfies the general assumptions of Section 2, the doubling property and the 2-Poincaré inequality. We stress that the 1-Poincaré inequality is not assumed.

4.1 Heat semigroup-based Besov classes

We first turn our attention to the study of Besov classes. In [3], we defined the heat semigroup-based Besov classes. Our basic definition of the Besov seminorm is the following:

Definition 4.1 ([3]). *Let $p \ge 1$ and $\alpha \ge 0$. For $f \in L^p(X)$, we define the Besov seminorm:*

$$\|f\|_{p,\alpha} = \sup_{t>0} t^{-\alpha} \left(\int_X \int_X p_t(x,y) |f(x) - f(y)|^p d\mu(x) d\mu(y) \right)^{1/p},$$

and the Besov spaces

$$\mathbf{B}^{p,\alpha}(X) = \{ f \in L^p(X) \,:\, \|f\|_{p,\alpha} < +\infty \}. \tag{14}$$

The norm on $\mathbf{B}^{p,\alpha}(X)$ is defined as:

$$\|f\|_{\mathbf{B}^{p,\alpha}(X)} = \|f\|_{L^p(X)} + \|f\|_{p,\alpha}.$$

It is proved in Proposition 4.14 and Corollary 4.16 of [3] that $\mathbf{B}^{p,\alpha}(X)$ is a Banach space for $p \ge 1$ and that it is reflexive for $p > 1$. In this section, we compare the spaces $\mathbf{B}^{p,\alpha}(X)$ to more classical notions of Besov classes that have previously been considered in the metric setting.

We recall the following definition from [34]. For $0 \le \alpha < \infty$, $1 \le p < \infty$ and $p < q \le \infty$, let $B^\alpha_{p,q}(X)$ be the collection of functions $u \in L^p(X)$ for which, if $q < \infty$

$$\|u\|_{B^\alpha_{p,q}(X)} := \left(\int_0^\infty \left(\int_X \int_{B(x,t)} \frac{|u(y)-u(x)|^p}{t^{\alpha p} \mu(B(x,t))} \, d\mu(y) \, d\mu(x) \right)^{q/p} \frac{dt}{t} \right)^{1/q} < \infty \tag{15}$$

and in the case $q = \infty$

$$\|u\|_{B^\alpha_{p,\infty}(X)} := \sup_{t>0} \left(\int_X \int_{B(x,t)} \frac{|u(y)-u(x)|^p}{t^{\alpha p} \mu(B(x,t))} \, d\mu(y) \, d\mu(x) \right)^{1/p} < \infty. \tag{16}$$

Proposition 4.2. *For $1 \le p < \infty$ and $0 < \alpha < \infty$ we have*

$$\mathbf{B}^{p,\alpha/2}(X) = B^\alpha_{p,\infty}(X),$$

with equivalent seminorms.

Proof. Since μ is doubling and supports a 2-Poincaré inequality, we have the Gaussian double bound (6) for $p_t(x,y)$. Hence if $u \in \mathbf{B}^{p,\alpha}(X)$, we then must have

$$\begin{aligned}
\|u\|^p_{p,\alpha/2} &\ge C^{-1} \sup_{t>0} \int_X \int_X \frac{|u(y)-u(x)|^p}{t^{\alpha p/2}} \frac{e^{-c\, d(x,y)^2/t}}{\mu(B(x,\sqrt{t}))} \, d\mu(y) \, d\mu(x) \\
&\ge C^{-1} \sup_{\sqrt{t}>0} \int_X \int_{B(x,\sqrt{t})} \frac{|u(y)-u(x)|^p}{t^{\alpha p/2}} \frac{e^{-c\, d(x,y)^2/t}}{\mu(B(x,\sqrt{t}))} \, d\mu(y) \, d\mu(x) \\
&\ge C^{-1} \sup_{\sqrt{t}>0} \int_X \int_{B(x,\sqrt{t})} \frac{|u(y)-u(x)|^p}{t^{\alpha p/2} \, \mu(B(x,\sqrt{t}))} \, d\mu(y) \, d\mu(x) \\
&= C^{-1} \|u\|^p_{B^\alpha_{p,\infty}(X)},
\end{aligned}$$

and from this it follows that $\mathbf{B}^{p,\alpha/2}(X)$ embeds boundedly into $B^\alpha_{p,\infty}(X)$.

Now we focus on proving the converse embedding. From (2) and (6), we have

$$\begin{aligned}
&\frac{1}{t^{\alpha p/2}} \int_X \int_X |u(y)-u(x)|^p \, p_t(x,y) \, d\mu(y) \, d\mu(x) \\
&\le \frac{C}{t^{\alpha p/2}} \int_X \sum_{i=-\infty}^{\infty} \int_{B(x,2^i\sqrt{t}) \setminus B(x,2^{i-1}\sqrt{t})} \frac{|u(y)-u(x)|^p \, e^{-c4^i}}{\mu(B(x,\sqrt{t}))} \, d\mu(y) \, d\mu(x) \\
&\le \frac{C}{t^{\alpha p/2}} \int_X \sum_{i=-\infty}^{\infty} \int_{B(x,2^i\sqrt{t})} \frac{|u(y)-u(x)|^p \, e^{-c4^i}}{\mu(B(x,2^i\sqrt{t}))} \frac{\mu(B(x,2^i\sqrt{t}))}{\mu(B(x,\sqrt{t}))} \, d\mu(y) \, d\mu(x) \\
&\le \frac{C}{t^{\alpha p/2}} \sum_{i=-\infty}^{\infty} e^{-c4^i} \max\{1, 2^{iQ}\} \, (2^i\sqrt{t})^{\alpha p} \int_X \int_{B(x,2^i\sqrt{t})} \frac{|u(y)-u(x)|^p}{(2^i\sqrt{t})^{\alpha p} \, \mu(B(x,2^i\sqrt{t}))} \, d\mu(y) \, d\mu(x) \\
&\le C \, \|u\|^p_{B^\alpha_{p,\infty}(X)} \sum_{i=-\infty}^{\infty} e^{-c4^i} \, 2^{i\alpha p} \max\{1, 2^{iQ}\}.
\end{aligned}$$

Since

$$\sum_{i=-\infty}^{\infty} e^{-c4^i} \, 2^{i\alpha p} \max\{1, 2^{iQ}\} \le \sum_{i\in\mathbb{N}} e^{-c4^i} \, 2^{i(\alpha p + Q)} + \sum_{i=0}^{\infty} 2^{-i\alpha p} < \infty,$$

the desired bound follows. ∎

4.2 Under the weak Bakry-Émery condition, $\mathbf{B}^{1,1/2}(X) = BV(X)$

Recall from Definition 2.1 that $u \in \mathcal{F}_{\rm loc}(X)$ if for each ball B in X there is a compactly supported Lipschitz function φ with $\varphi = 1$ on B such that $u\varphi \in \mathcal{F}$; in this case we can set $|\nabla u| = |\nabla(u\varphi)|$ in B, thanks to the strict locality property of $\mathcal{E}$.

Lemma 4.3. *Suppose that the weak Bakry-Émery condition (7) holds. Then for $u \in \mathcal{F} \cap W^{1,1}(X)$, we have that*

$$\|P_t u - u\|_{L^1(X)} \le C\sqrt{t} \int_X |\nabla u|\, d\mu.$$

Hence, if $u \in BV(X)$, then

$$\|P_t u - u\|_{L^1(X)} \le C\sqrt{t}\, \|Du\|(X).$$

Proof. To see the first part of the claim, we note that for each $x \in X$ and $s > 0$, $\frac{\partial}{\partial s} P_s u(x)$ exists, and so by the fundamental theorem of calculus, for $0 < \tau < t$ and $x \in X$,

$$P_t u(x) - P_\tau u(x) = \int_\tau^t \frac{\partial}{\partial s} P_s u(x)\, ds.$$

Thus for each compactly supported function $\varphi \in \mathcal{F} \cap L^\infty(X)$, by the facts that $P_t u$ satisfies the heat equation and that P_s is a symmetric operator for each $s > 0$,

$$\begin{aligned}
\left| \int_X \varphi(x)[P_t u(x) - P_\tau u(x)]\, d\mu(x) \right| &= \left| -\int_X \int_\tau^t \varphi(x) \frac{\partial}{\partial s} P_s u(x)\, ds\, d\mu(x) \right| \\
&= \left| \int_\tau^t \int_X d\Gamma(\varphi, P_s u)(x)\, ds \right| \\
&= \left| \int_\tau^t \int_X d\Gamma(P_s\varphi, u)(x)\, ds \right| \\
&\le \int_\tau^t \int_X |\nabla P_s \varphi|\, |\nabla u|\, d\mu\, ds \\
&\le \||\nabla P_s\varphi|\|_{L^\infty(X)} \int_\tau^t \int_X |\nabla u|\, ds\, d\mu.
\end{aligned}$$

An application of (7) gives

$$\begin{aligned}
\left| \int_X \varphi(x)[P_t u(x) - P_\tau u(x)]\, d\mu(x) \right| &\le \frac{C}{\sqrt{t}} \|\varphi\|_{L^\infty(X)} \int_X \int_\tau^t |\nabla u|\, ds\, d\mu \\
&= C\, \frac{t-\tau}{\sqrt{t}}\, \|\varphi\|_{L^\infty(X)} \int_X |\nabla u|\, d\mu.
\end{aligned}$$

As the above holds for all compactly supported $\varphi \in \mathcal{F} \cap L^\infty(X)$, we obtain

$$\|P_t u - P_\tau u\|_{L^1(X)} \le C\, \frac{t-\tau}{\sqrt{t}} \int_X |\nabla u|\, d\mu.$$

Now by the fact that by the fact that $\{P_t\}_{t>0}$ has an extension as a contraction semigroup to $L^1(X)$ such that $P_\tau u \to u$ as $\tau \to 0^+$ in $L^1(X)$, (see [3, Section 2.2]), we have

$$\|P_t u - u\|_{L^1(X)} \le C\sqrt{t} \int_X |\nabla u|\, d\mu.$$

Finally, if $u \in BV(X)$, then we can find a sequence $u_k \in \mathcal{F} \cap W^{1,1}(X)$ such that $u_k \to u$ in $L^1(X)$ and $\lim_{k\to\infty} \int_X |\nabla u_k|\, d\mu = \|Du\|(X)$. By the contraction property of P_t on $L^1(X)$, we have

$$\begin{aligned}\|P_t u - u\|_{L^1(X)} &\le \|P_t(u-u_k)\|_{L^1(X)} + \|P_t u_k - u_k\|_{L^1(X)} + \|u_k - u\|_{L^1(X)} \\ &\le C\|u-u_k\|_{L^1(X)} + C\sqrt{t}\int_X |\nabla u_k|\, d\mu + \|u-u_k\|_{L^1(X)}.\end{aligned}$$

Letting $k \to \infty$ concludes the proof. □

Note from the results of [69, Theorem 4.1] that if the measure μ is doubling and supports a 1-Poincaré inequality, then a measurable set $E \subset X$ is in the BV class if

$$\liminf_{t\to 0^+} \frac{1}{\sqrt{t}} \int_{E^{\sqrt{t}}\setminus E} P_t \mathbf{1}_E \, d\mu < \infty.$$

Here $E^\varepsilon = \bigcup_{x\in E} B(x,\varepsilon)$. Note that by the symmetry and conservativeness of the operator P_t,

$$\begin{aligned}\int_X |P_t\mathbf{1}_E - \mathbf{1}_E|\, d\mu &= \int_E (1 - P_t\mathbf{1}_E)\, d\mu + \int_{X\setminus E} P_t\mathbf{1}_E\, d\mu \\ &= \int_X \mathbf{1}_E(1 - P_t\mathbf{1}_E)\, d\mu + \int_{X\setminus E} P_t\mathbf{1}_E\, d\mu \\ &= \int_X (P_t\mathbf{1}_E)\,\mathbf{1}_{X\setminus E}\, d\mu + \int_{X\setminus E} P_t\mathbf{1}_E\, d\mu = 2\int_{X\setminus E} P_t\mathbf{1}_E\, d\mu.\end{aligned}$$

Therefore,

$$\int_{E^{\sqrt{t}}\setminus E} P_t\mathbf{1}_E\, d\mu \le \int_{X\setminus E} P_t\mathbf{1}_E\, d\mu = \frac{1}{2}\|P_t\mathbf{1}_E - \mathbf{1}_E\|_{L^1(X)}.$$

Thus if μ is doubling and supports a 1-Poincaré inequality, and in addition

$$\sup_{t>0} \frac{1}{\sqrt{t}} \|P_t\mathbf{1}_E - \mathbf{1}_E\|_{L^1(X)} < \infty,$$

then E is of finite perimeter. In our framework, those results coming from [69] can not be used, since we do not assume the 1-Poincaré inequality. Instead we prove the following theorem, which is the main result of the section.

Theorem 4.4. *If the weak Bakry-Émery condition* (7) *holds, then* $\mathbf{B}^{1,1/2}(X) = BV(X)$ *with comparable seminorms. Moreover, there exist constants* $c, C > 0$ *such that for every* $u \in BV(X)$,

$$c \limsup_{s\to 0} s^{-1/2} \int_X P_s(|u - u(y)|)(y) d\mu(y) \le \|Du\|(X) \le C \liminf_{s\to 0} s^{-1/2} \int_X P_s(|u-u(y)|)(y) d\mu(y).$$

Proof. First we assume that $u \in BV(X)$. Then we know that for almost every $t \in \mathbb{R}$ the set E_t is of finite perimeter, where

$$E_t = \{x \in X \,:\, u(x) > t\},$$

and by the co-area formula for BV functions (see Theorem 3.11),

$$\|Du\|(X) = \int_{\mathbb{R}} \|D\mathbf{1}_{E_t}\|(X)\, dt.$$

For such t, by Lemma 4.3 we know that

$$\sup_{s>0} \frac{1}{\sqrt{s}} \int_X |P_s \mathbf{1}_{E_t}(x) - \mathbf{1}_{E_t}(x)|\, d\mu(x) \le C\, \|D\mathbf{1}_{E_t}\|(X).$$

Now, setting $A = \{(x,y) \in X \times X \,:\, u(x) < u(y)\}$, we have for $s > 0$,

$$\begin{aligned}
\int_X \int_X & p_s(x,y)|u(x) - u(y)|\, d\mu(x)d\mu(y) \\
&= 2\int_A p_s(x,y)|u(x)-u(y)|\, d\mu(x)d\mu(y) \\
&= 2\int_A \int_{u(x)}^{u(y)} p_s(x,y)\, dt\, d\mu(x)d\mu(y) \\
&= 2\int_X \int_X \int_{\mathbb{R}} \mathbf{1}_{[u(x),u(y))}(t)\, \mathbf{1}_A(x,y)\, p_s(x,y)\, dt\, d\mu(x)d\mu(y) \\
&= 2\int_{\mathbb{R}} \int_X \int_X \mathbf{1}_{E_t}(y)[1-\mathbf{1}_{E_t}(x)]\, p_s(x,y)\, d\mu(x)d\mu(y)\, dt \\
&= 2\int_{\mathbb{R}} \int_X P_s\mathbf{1}_{E_t}(x)[1-\mathbf{1}_{E_t}(x)]\, d\mu(x)\, dt \\
&= 2\int_{\mathbb{R}} \int_{X\setminus E_t} P_s\mathbf{1}_{E_t}(x)\, d\mu(x)\, dt.
\end{aligned}$$

Observe that

$$\int_{X\setminus E_t} P_s\mathbf{1}_{E_t}(x)\, d\mu(x) = \int_{X\setminus E_t} |P_s\mathbf{1}_{E_t}(x) - \mathbf{1}_{E_t}(x)|\, d\mu(x) \le \int_X |P_s\mathbf{1}_{E_t}(x) - \mathbf{1}_{E_t}(x)|\, d\mu(x).$$

Therefore we obtain

$$\int_X \int_X p_s(x,y)\ |u(x)-u(y)|\, d\mu(x)d\mu(y) \le 2\int_{\mathbb{R}} \|P_s\mathbf{1}_{E_t} - \mathbf{1}_{E_t}\|_{L^1(X)}\, dt.$$

An application of Lemma 4.3 now gives

$$\int_X \int_X p_s(x,y)\ |u(x)-u(y)|\, d\mu(x)d\mu(y) \le C\sqrt{s}\ \int_{\mathbb{R}} \|D\mathbf{1}_{E_t}\|(X)\, dt,$$

whence with the help of the co-area formula we obtain

$$\|u\|_{1,1/2} \le C\, \|Du\|(X),$$

that is, $u \in \mathbf{B}^{1,1/2}(X)$. Thus $BV(X) \subset \mathbf{B}^{1,1/2}(X)$ boundedly.

Now we show that $\mathbf{B}^{1,1/2}(X) \subset BV(X)$. This inclusion holds even when $\mathcal{E}$ does not support a Bakry-Émery curvature condition; only a 2-Poincaré inequality and the doubling condition on μ are needed. Suppose that $u \in \mathbf{B}^{1,1/2}(X)$. Then there is some $C \ge 0$ such that for each $t > 0$,

$$\int_X \int_X p_t(x,y)|u(y)-u(x)|\, d\mu(y)\, d\mu(x) \le C\sqrt{t}.$$

By (6), we have a Gaussian lower bound for the heat kernel:

$$p_t(x,y) \ge \frac{e^{-c\, d(x,y)^2/t}}{C\mu(B(x,\sqrt{t}))}.$$

Let $C_0 = \|u\|_{1,1/2}$. Therefore, setting $\Delta_\varepsilon = \{(x,y) \in X : d(x,y) < \varepsilon\}$ for some $\varepsilon > 0$, we get

$$\begin{aligned} C_0 \sqrt{t} &\geq \int_X \int_X \frac{e^{-c\, d(x,y)^2/t}}{C\mu(B(x,\sqrt{t}))} |u(y) - u(x)|\, d\mu(y)\, d\mu(x) \\ &\geq \iint_{\Delta_\varepsilon} \frac{e^{-c\, d(x,y)^2/t}}{C\mu(B(x,\sqrt{t}))} |u(y) - u(x)|\, d\mu(x) d\mu(y) \\ &\geq \frac{e^{-c\varepsilon^2/t}}{C} \iint_{\Delta_\varepsilon} \frac{|u(y) - u(x)|}{\mu(B(x,\sqrt{t}))}\, d\mu(x) d\mu(y). \end{aligned}$$

With the choice of $\varepsilon = \sqrt{t}$, we now get

$$C_0\, \varepsilon \geq \frac{1}{C} \iint_{\Delta_\varepsilon} \frac{|u(y) - u(x)|}{\mu(B(x,\varepsilon))}\, d\mu(x) d\mu(y).$$

It follows that

$$\liminf_{\varepsilon \to 0^+} \frac{1}{\varepsilon} \iint_{\Delta_\varepsilon} \frac{|u(y) - u(x)|}{\mu(B(x,\varepsilon))}\, d\mu(x) d\mu(y) \leq C\, C_0 < \infty. \tag{17}$$

Now an argument as in the second half of the proof of [69, Theorem 3.1] tells us that $u \in BV(X)$. We point out here that although Theorem 3.1 in [69] assumes that X supports a 1-Poincaré inequality, the second part of the proof there does not need this assumption. In fact, the argument using discrete convolution there is valid also in our setting. It is this second part of the proof that we referred to above. We then obtain

$$\|Du\|(X) \leq \liminf_{\varepsilon \to 0^+} \frac{1}{\varepsilon} \iint_{\Delta_\varepsilon} \frac{|u(y) - u(x)|}{\mu(B(x,\varepsilon))}\, d\mu(x) d\mu(y) \leq \|u\|_{1,1/2}. \qquad \square$$

Remark 4.5. *As a byproduct of this proof, we also obtain that there exists a constant $C > 0$ such that for every $u \in BV(X)$,*

$$\sup_{\varepsilon > 0} \frac{1}{\varepsilon} \iint_{\Delta_\varepsilon} \frac{|u(y) - u(x)|}{\mu(B(x,\varepsilon))}\, d\mu(x) d\mu(y) \leq \liminf_{\varepsilon \to 0^+} \frac{C}{\varepsilon} \iint_{\Delta_\varepsilon} \frac{|u(y) - u(x)|}{\mu(B(x,\varepsilon))}\, d\mu(x) d\mu(y)$$

because both sides are comparable to $\|Du\|(X)$. Indeed, the fact that $\|Du\|(X)$ is dominated by the right hand side is directly from Theorem 4.4, which, together with Proposition 4.2 (the metric characterization of Besov spaces), implies that the left hand side can be bounded by $\|Du\|(X)$. This property of the metric measure space (X, d, μ) can be viewed as an interesting consequence of the weak Bakry-Émery estimate.

Remark 4.6. *Another application of Proposition 4.2 is the following. It is in general not true that if $\|Du\|(X) = 0$ then u is constant almost everywhere in X, even if X is connected. Should X support a 1-Poincaré inequality, it follows immediately that if $\|Du\|(X) = 0$ then u is constant. We can use the above proposition to show that even if we do not have 1-Poincaré inequality, if X supports the Bakry-Émery curvature condition (7), then*

$$\sup_{t>0} \int_X \int_{B(x,t)} \frac{|u(x) - u(y)|}{t\mu(B(x,t))}\, d\mu(y)\, d\mu(x) \simeq \|Du\|(X),$$

and hence if $\|Du\|(X) = 0$ then u is constant.

4.3 Sets of finite perimeter

We introduce some notions from the paper [6] of Ambrosio, Miranda and Pallara. Given $A \subset X$ we set

$$\mathcal{H}(A) := \lim_{\varepsilon \to 0^+} \inf \left\{ \sum_i \frac{\mu(B_i)}{\mathrm{rad}(B_i)} \;:\; A \subset \bigcup_i B_i, \text{ and } \forall i, \; \mathrm{rad}(B_i) < \varepsilon \right\}.$$

It is known, see [59, Proposition 6.3], even without the assumption that X supports a 2-Poincaré inequality, that if $\mathcal{H}(\partial E) < \infty$, then E is of finite perimeter.

Now let $E \subset X$ be a set of finite perimeter and define the measure-theoretic boundary by

$$\partial_m E = \left\{ x \in X : \limsup_{r \to 0^+} \frac{\mu(B(x,r) \cap E)}{\mu(B(x,r))} > 0, \; \limsup_{r \to 0^+} \frac{\mu(B(x,r) \setminus E)}{\mu(B(x,r))} > 0 \right\}.$$

For $\alpha \in (0, 1/2)$, define also

$$\partial_\alpha E = \left\{ x \in X : \liminf_{r \to 0^+} \min \left\{ \frac{\mu(B(x,r) \cap E)}{\mu(B(x,r))}, \, \frac{\mu(B(x,r) \setminus E)}{\mu(B(x,r))} \right\} \geq \alpha \right\}.$$

If X supports a 1-Poincaré inequality then, by the results of [6, Theorem 4.4], there is a number $\gamma \in (0, 1/2]$ such that $\mathcal{H}(\partial_m E \backslash \partial_\gamma E) = 0$, where γ depends solely on the doubling and the 1-Poincaré constants. The same result also tells us that if E is of finite perimeter then $\mathcal{H}(\partial_m E) \simeq P(E, X)$.

We are not assuming X supports a 1-Poincaré inequality, but only that μ is doubling and X supports a 2-Poincaré inequality. In this setting we instead consider for $r_0 > 0$ and $0 < \alpha \leq 1/2$ the quantity

$$\partial_\alpha^{r_0} E = \left\{ x \in X : \min \left\{ \frac{\mu(B(x,r) \cap E)}{\mu(B(x,r))}, \, \frac{\mu(B(x,r) \setminus E)}{\mu(B(x,r))} \right\} > \alpha \text{ for all } 0 < r \leq r_0 \right\}.$$

Observe that $\partial_m E = \bigcup_{0<\alpha<1} \bigcup_{0<r_0<1} \partial_\alpha^{r_0}(E)$ and the union can be made countable by taking α and r_0 to be rational numbers.

Proposition 4.7. *Suppose that $E \subset X$ with $\|\mathbf{1}_E\|_{\mathbf{B}^{1,1/2}(X)} < \infty$. Then for all $r_0 > 0$ and $0 < \alpha < 1$,*

$$\mathcal{H}(\partial_\alpha^{r_0} E) \leq \frac{C}{\alpha} P(E, X).$$

Consequently, $\mathcal{H}(\partial_\alpha E) \leq \frac{C}{\alpha} P(E, X)$ and $\mathcal{H}|_{\partial_m E}$ is a σ-finite measure.

Proof. Examining the proof of Theorem 4.4 we see that even without the Bakry-Émery condition (7), if $\mathbf{1}_E \in \mathbf{B}^{1,1/2}(X)$ then $\mathbf{1}_E \in BV(X)$. By the definition of $\mathbf{B}^{1,1/2}(X)$, we know that

$$\sup_{t>0} \frac{1}{\sqrt{t}} \int_X \int_X p_t(x,y) |\mathbf{1}_E(x) - \mathbf{1}_E(y)| \, d\mu(x) \, d\mu(y) \leq C \, P(E, X).$$

Fix $t < (r_0/3)^2$. Let $\{B_i\}_i$ be a maximal $\sqrt{t}$-separated covering of $\partial_\alpha^{r_0} E$ such that the balls $5B_i$ have a bounded overlap (see Section 2.2). Then by the doubling property of μ and by the Gaussian lower bound for $p_t(x,y)$ in (5),

$$\begin{aligned} C \sqrt{t} \, P(E, X) &\geq \sum_i \int_{B_i \cap E} \int_{B_i \setminus E} p_t(x,y) \, d\mu(x) \, d\mu(y) \\ &\geq C^{-1} \sum_i \int_{B_i \cap E} \int_{B_i \setminus E} \frac{e^{-C}}{\mu(B_i)} \, d\mu(x) \, d\mu(y) \end{aligned}$$

$$\geq C^{-1} \sum_i \frac{\mu(B_i \cap E)\, \mu(B_i \setminus E)\, \mu(B_i)}{\mu(B_i)^2}.$$

In the above computations, C stands for various generic constants that depend only on the doubling and Poincaré constants of the space, and the value of C could change at each occurrence. Note that at least one of $\mu(B_i \cap E)/\mu(B_i)$ and $\mu(B_i \setminus E)/\mu(B_i)$ is larger than 1/2. Now by the definition of $\partial_\alpha^{r_0} E$ we obtain

$$C\, P(E, X) \geq \alpha \sum_i \frac{\mu(B_i)}{\sqrt{t}}.$$

Since $\sqrt{t}$ is the radius of each B_i, we get

$$C\, P(E, X) \geq \alpha \limsup_{t \to 0^+} \sum_i \frac{\mu(B_i)}{\sqrt{t}} \geq \alpha\, \mathcal{H}(\partial_\alpha^{r_0} E).$$

Recall that

$$\partial_m E = \bigcup_{\alpha \in (0,1) \cap \mathbb{Q}} \ \bigcup_{r_0 \in (0,1) \cap \mathbb{Q}} \partial_\alpha^{r_0}(E).$$

This yields that $\mathcal{H}|_{\partial_m E}$ is a σ-finite measure.

If $0 < r_1 < r_0$, then

$$\partial_\alpha^{r_0} E \subset \partial_\alpha^{r_1} E \subset \partial_m E.$$

Observing that

$$\partial_\alpha E = \bigcup_{0 < r_0 < 1} \partial_\alpha^{r_0} E = \bigcup_{(0,1) \cap \mathbb{Q}} \partial_\alpha^{r_0} E,$$

we now see by the continuity of measure that if the sets $\partial_\alpha^{r_0} E$ are Borel sets, then

$$\mathcal{H}(\partial_\alpha E) \leq \frac{C}{\alpha} P(E, X).$$

To see that $\partial_\alpha^{r_0} E$ is a Borel set we argue as follows. Recall that we assume μ to be Borel regular. Therefore, given a μ-measurable set E and $r > 0$, the function

$$x \mapsto \mu(B(x, r) \cap E)$$

is lower semicontinuous, and so the map

$$\varphi_{E,r}(x) := \frac{\mu(B(x, r) \cap E)}{\mu(B(x, r))}$$

is a Borel function. Hence the function $\Phi_{E,r_0} := \inf_{r \in \mathbb{Q} \cap (0, r_0]} \varphi_{E,r}$ is also a Borel function, and hence

$$\partial_\alpha^{r_0} E = \{x \in X \,:\, \Phi_{E,r_0}(x) > \alpha\} \cap \{x \in X \,:\, \Phi_{X \setminus E, r_0}(x) > \alpha\}$$

is a Borel set. □

Proposition 4.7 gives us a way to control, from above, the $\mathcal{H}$-measure of $\partial_m E$ for a set E of finite perimeter. This should be contrasted with the following lower bound on the co-dimension 1 Minkowski measure of ∂E. For a set $A \subset X$, the co-dimension 1-Minkowski measure of A is defined to be

$$\mathcal{M}_{-1}(A) := \liminf_{\varepsilon \to 0^+} \frac{\mu(A_\varepsilon)}{\varepsilon},$$

where $A_\varepsilon = \bigcup_{x \in A} B(x, \varepsilon)$.

Proposition 4.8. *Assuming the weak Bakry-Émery condition* (7) *we have for a set E of finite perimeter that*

$$P(E,X) \leq C\mathcal{M}_{-1}(\partial E)$$

Proof. Observe that $|\mathbf{1}_E(x) - \mathbf{1}_E(y)|$ is bounded by 1 and is zero if $y \in B(x,t)$ and x is not in the $2t$ neighborhood $(\partial E)_{2t}$ of the boundary. Using the doubling property of μ we immediately deduce

$$\liminf_{t\to 0^+} \int_X \int_{B(x,t)} \frac{|\mathbf{1}_E(x) - \mathbf{1}_E(y)|}{t\mu(B(x,t))}\, d\mu(y)\, d\mu(x) \leq C\mathcal{M}_{-1}(\partial E).$$

Now recall from Remark 4.5 that under the weak Bakry-Émery assumption the above limit inferior is comparable to the supremum, which is the $B^1_{1,\infty}$ norm defined in (16), because both are comparable to the perimeter measure $P(E,X) = \|D\mathbf{1}_E(X)\|$. ☐

4.4 Under the strong Bakry-Émery condition, $\mathbf{B}^{p,1/2}(X) = W^{1,p}(X)$ for $p > 1$

In this section we compare the Besov and Sobolev seminorms for $p > 1$. The case $p = 1$ was studied in detail in Section 4.2. Our main theorem in this section is the following:

Theorem 4.9. *Suppose that the strong Bakry-Émery condition* (8) *holds. Then, for every $p > 1$, $\mathbf{B}^{p,1/2}(X) = W^{1,p}(X)$ with comparable norms.*

We will divide the proof of Theorem 4.9 in two parts. In the first part, Theorem 4.11, we prove that $\mathbf{B}^{p,1/2}(X) \subset W^{1,p}(X)$. As we will see, this inclusion does not require the strong Bakry-Émery condition (8). In the second part, Theorem 4.17 we will prove the inclusion $W^{1,p}(X) \subset \mathbf{B}^{p,1/2}(X)$ and, to this end, will use the strong Bakry-Émery condition. Before turning to the proof, we point out the following corollary regarding the Riesz transform.

Corollary 4.10. *Suppose that the strong Bakry-Émery condition* (8) *holds. Let $p > 1$. Then for any $f \in \mathbf{B}^{p,1/2}(X) \cap \mathcal{F}$,*

$$\|f\|_{p,1/2} \simeq \|\sqrt{-L}f\|_{L^p(X)}.$$

Consequently, $\mathbf{B}^{p,1/2}(X) = \mathcal{L}^{1/2}_p$, where $\mathcal{L}^{1/2}_p$ is the domain of the operator $\sqrt{-L}$ in $L^p(X)$ (see [3, Section 4.6] *for the definition).*

Proof. In view of Theorem 4.9, we have that for any $f \in \mathbf{B}^{p,1/2}(X)$

$$\|f\|_{p,1/2} \simeq \||\nabla f|\|_{L^p(X)}.$$

On the other hand, it follows from [9, Theorem 1.4] that for any $f \in \mathcal{L}^{1/2}_p$,

$$\|\sqrt{-L}f\|_{L^p(X)} \simeq \||\nabla f|\|_{L^p(X)}.$$

We conclude the proof by combining the above two facts. ☐

4.4.1 $\mathbf{B}^{p,1/2}(X) \subset W^{1,p}(X)$

Theorem 4.11. *Let $p > 1$. There exists a constant $C > 0$ such that for every $u \in \mathbf{B}^{p,1/2}(X)$,*

$$\||\nabla u|\|_{L^p(X)} \leq C\|u\|_{p,1/2}.$$

Proof. Let $u \in \mathbf{B}^{p,1/2}(X)$. Then from Proposition 4.2, we see that for each $\varepsilon > 0$,

$$\frac{1}{\varepsilon^p} \iint_{\Delta_\varepsilon} \frac{|u(x)-u(y)|^p}{\mu(B(x,\varepsilon))}\, d\mu(y)\, d\mu(x) \le \|u\|_{p,1/2}^p < \infty.$$

Fix $\varepsilon > 0$. As in the proof of Lemma 2.11, let $\{B_i^\varepsilon = B(x_i^\varepsilon, \varepsilon)\}_i$ be a maximal ε-separated covering and $\{\varphi_i^\varepsilon\}_i$ be a (C/ε)-Lipschitz partition of unity subordinated to this covering. We also set

$$u_\varepsilon := \sum_i u_{B_i^\varepsilon}\, \varphi_i^\varepsilon.$$

Then u_ε is locally Lipschitz and hence is in $\mathcal{F}_{\mathrm{loc}}(X)$. Indeed, for $x, y \in B_j^\varepsilon$ we see that

$$\begin{aligned} |u_\varepsilon(x) - u_\varepsilon(y)| &\le \sum_{i:2B_i^\varepsilon \cap 2B_j^\varepsilon \neq \emptyset} |u_{B_i^\varepsilon} - u_{B_j^\varepsilon}||\varphi_i^\varepsilon(x) - \varphi_i^\varepsilon(y)| \\ &\le \frac{C\, d(x,y)}{\varepsilon} \sum_{i:2B_i^\varepsilon \cap 2B_j^\varepsilon \neq \emptyset} \left(\fint_{B_i^\varepsilon} \fint_{B(x,2\varepsilon)} |u(y)-u(x)|^p\, d\mu(y)\, d\mu(x) \right)^{1/p}. \end{aligned}$$

Therefore, by Lemma 2.5, we see that

$$\begin{aligned} |\nabla u_\varepsilon| &\le \frac{C}{\varepsilon} \sum_{i:2B_i^\varepsilon \cap 2B_j^\varepsilon \neq \emptyset} \left(\fint_{B_i^\varepsilon} \fint_{B(x,2\varepsilon)} |u(y)-u(x)|^p\, d\mu(y)\, d\mu(x) \right)^{1/p} \\ &\le C \left(\fint_{2B_j^\varepsilon} \fint_{B(x,2\varepsilon)} \frac{|u(y)-u(x)|^p}{\varepsilon^p}\, d\mu(y)\, d\mu(x) \right)^{1/p}, \end{aligned}$$

and so by the bounded overlap property of the collection $2B_j^\varepsilon$,

$$\begin{aligned} \int_X |\nabla u_\varepsilon|^p\, d\mu &\le \sum_j \int_{B_j^\varepsilon} |\nabla u_\varepsilon|^p\, d\mu \\ &\le C \sum_j \int_{2B_j^\varepsilon} \fint_{B(x,2\varepsilon)} \frac{|u(y)-u(x)|^p}{\varepsilon^p}\, d\mu(y)\, d\mu(x) \\ &\le C \int_X \fint_{B(x,2\varepsilon)} \frac{|u(y)-u(x)|^p}{\varepsilon^p}\, d\mu(y)\, d\mu(x) \\ &\le C\, \frac{1}{\varepsilon^p} \int_{\Delta_{2\varepsilon}} \frac{|u(x)-u(y)|^p}{\mu(B(x,\varepsilon))}\, d\mu(y)\, d\mu(x) \le C\, \|u\|_{p,1/2}^p. \end{aligned}$$

Hence we have

$$\sup_{\varepsilon>0} \int_X |\nabla u_\varepsilon|^p\, d\mu \le C\, \|u\|_{p,1/2}^p. \tag{18}$$

In a similar manner, we can also show that

$$\int_X |u_\varepsilon(x) - u(x)|^p\, d\mu(x) \le C\varepsilon^p \int_{\Delta_{2\varepsilon}} \frac{|u(x)-u(y)|^p}{\varepsilon^p\, \mu(B(x,\varepsilon))}\, d\mu(y)\, d\mu(x) \le C\, \varepsilon^p\, \|u\|_{p,1/2}^p,$$

that is, $u_\varepsilon \to u$ in $L^p(X)$ as $\varepsilon \to 0^+$.

Take a sequence $\varepsilon_n \to 0^+$. From (18) and the reflexivity of $L^p(X)$, there exists a subsequence of $\{\nabla u_{\varepsilon_n}\}_n$ that is weakly convergent in $L^p(X)$. By Mazur's lemma, a sequence of convex combinations of u_{ε_n} converges in the norm of $W^{1,p}(X)$. Since it converges to u in $L^p(X)$, we conclude that $u \in W^{1,p}(X)$ and hence

$$\||\nabla u|\|_{L^p(X)} \le C\|u\|_{p,1/2}.$$

□

4.4.2 $W^{1,p}(X) \subset \mathbf{B}^{p,1/2}(X)$

We now turn to the proof of the upper bound for the Besov seminorm in terms of the Sobolev seminorm and assume that the strong Bakry-Émery condition (8) holds.

A first important corollary of the strong Bakry-Émery estimate is the following Hamilton's type gradient estimate for the heat kernel. This type of estimate is well-known on Riemannian manifolds with non-negative Ricci curvature (see for instance [64]), but is new in our general framework.

Theorem 4.12. *There exists a constant $C > 0$ such that for every $t > 0$, $x, y \in X$,*

$$|\nabla_x \ln p_t(x,y)|^2 \leq \frac{C}{t}\left(1 + \frac{d(x,y)^2}{t}\right).$$

Proof. The proof proceeds in two steps.

Step 1: We first collect a gradient bound for the heat kernel. Observe that (8) implies a weaker L^2 version as follows

$$|\nabla P_t u|^2 \leq C P_t(|\nabla u|^2),$$

and hence the following pointwise heat kernel gradient bound (see [9, Lemma 3.3]) holds:

$$|\nabla_x p_t(x,y)| \leq \frac{C}{\sqrt{t}} \frac{e^{-cd(x,y)^2/t}}{\sqrt{\mu(B(x,\sqrt{t}))\mu(B(y,\sqrt{t}))}}.$$

In particular, we note that $|\nabla_x p_t(x,\cdot)| \in L^p(X)$ for every $p \geq 1$.

Step 2: In the second step, we prove a reverse log-Sobolev inequality for the heat kernel. Let $\tau, \varepsilon > 0$ and $x \in X$ be fixed. We denote $u = p_\tau(x,\cdot) + \varepsilon$. One has, from the chain rule for strictly local forms [32, Lemma 3.2.5],

$$\begin{aligned}
P_t(u \ln u) - P_t u \ln P_t u &= \int_0^t \partial_s \left(P_s(P_{t-s}u \ln P_{t-s}u)\right) ds \\
&= \int_0^t LP_s(P_{t-s}u \ln P_{t-s}u) - P_s(LP_{t-s}u \ln P_{t-s}u) - P_s(LP_{t-s}u) ds \\
&= \int_0^t P_s(L(P_{t-s}u \ln P_{t-s}u)) - P_s(LP_{t-s}u \ln P_{t-s}u) - P_s(LP_{t-s}u) ds \\
&= \int_0^t P_s\left[L(P_{t-s}u \ln P_{t-s}u)) - LP_{t-s}u \ln P_{t-s}u - LP_{t-s}u\right] ds \\
&= \int_0^t 2P_s\left(\frac{|\nabla P_{t-s}u|^2}{P_{t-s}u}\right) ds, \qquad (19)
\end{aligned}$$

where the above computations may be justified by using the Gaussian heat kernel estimates for the heat kernel and the Gaussian upper bound for the gradient of the heat kernel obtained in Step 1. In particular, we point out that the commutation $LP_s(P_{t-s}u \ln P_{t-s}u) = P_s(L(P_{t-s}u \ln P_{t-s}u))$ is justified by noting that $P_{t-s}u \ln P_{t-s}u - \varepsilon \ln \varepsilon$ is in the domain of L in $L^2(X,\mu)$. Here, L is the infinitesimal generator (the Laplacian operator) associated with $\mathcal{E}$.

Using the Cauchy-Schwarz inequality in the form $P_s\left(\frac{f^2}{g}\right) \geq \frac{(P_s f)^2}{P_s g}$ and then the strong Bakry-Émery estimate, we obtain from (19)

$$P_t(u \ln u) - P_t u \ln P_t u \geq 2\int_0^t \frac{(P_s|\nabla P_{t-s}u|)^2}{P_s(P_{t-s}u)} ds$$

$$\geq \frac{1}{C}\frac{1}{P_t u}\int_0^t |\nabla P_t u|^2 ds$$
$$\geq \frac{t}{C}\frac{1}{P_t u}|\nabla P_t u|^2.$$

Coming back to the definition of u, noting that $P_t p_\tau(x,\cdot) = p_{t+\tau}(x,\cdot)$ and applying the previous inequality with $t = \tau$, we may set $M_t(x) = \sup_{y\in X} p_t(x,y)$ and bound the $P_t(u \ln u)$ term by $(P_t u)\ln(M_t + \epsilon)$ to deduce

$$|\nabla_y \ln(p_{2t}(x,y)+\varepsilon)|^2 \leq \frac{C}{t} P_t \left[\ln\left(\frac{M_t(x)+\varepsilon}{p_{2t}(x,\cdot)+\varepsilon} \right) \right](y).$$

By letting $\varepsilon \to 0$ and using the Gaussian heat kernel estimate, one concludes

$$|\nabla_y \ln p_{2t}(x,y)|^2 \leq \frac{C}{t}\left(1 + \frac{d(x,y)^2}{t}\right)$$

Our desired inequality follows by rescaling t, adjusting the constant C and using the symmetry of $p_t(x,y)$ in x and y. □

Corollary 4.13. *Let $p > 1$. There exists a constant $C > 0$ such that for every $u \in L^p(X)$,*

$$|\nabla P_t u| \leq \frac{C}{\sqrt{t}}(P_t|u|^p)^{1/p}.$$

Proof. Let $p > 1$, q be the conjugate exponent and $u \in L^p(X)$. One has from Hölder's inequality

$$\begin{aligned}
|\nabla P_t u|(x) &\leq \int_X |\nabla_x p_t(x,y)||u(y)|d\mu(y) \\
&\leq \left(\int_X \frac{|\nabla_x p_t(x,y)|^q}{p_t(x,y)^{q/p}} d\mu(y) \right)^{1/q} (P_t|u|^p)^{1/p} \\
&\leq \left(\int_X |\nabla_x \ln p_t(x,y)|^q p_t(x,y) d\mu(y) \right)^{1/q} (P_t|u|^p)^{1/p}.
\end{aligned}$$

The proof follows then from Theorem 4.12 and the Gaussian upper bound for the heat kernel. □

Note that by integrating over X the previous proposition immediately yields:

Lemma 4.14. *Let $p > 1$. There exists a constant $C > 0$ such that for every $u \in L^p(X)$*

$$\| |\nabla P_t u| \|^2_{L^p(X)} \leq \frac{C}{t}\|u\|^2_{L^p(X)}.$$

From this estimate we obtain the following result.

Lemma 4.15. *Let $p > 1$. There exists a constant $C > 0$ such that for every $u \in L^p(X) \cap \mathcal{F}$ with $|\nabla u| \in L^p(X)$*

$$\|P_t u - u\|_{L^p(X)} \leq C\sqrt{t}\| |\nabla u| \|_{L^p(X)}$$

Proof. With the previous lemma in hand, the proof is similar to the one in Lemma 4.3, with φ in $\mathcal{F} \cap L^q(X)$ and compactly supported, where $p^{-1} + q^{-1} = 1$. As compactly supported functions in $\mathcal{F} \cap L^q(X)$ form a dense subclass of $L^q(X)$ we recover the L^p-norm of $P_t u - u$ by taking the supremum over all such φ with $\int_X |\varphi|^q \, d\mu \leq 1$. □

Lemma 4.16. *Let $p > 1$, then for every $u \in L^p(X) \cap \mathcal{F}$ with $|\nabla u| \in L^p(X)$*

$$\left(\int_X \int_X |P_t u(x) - u(y)|^p p_t(x,y) d\mu(x) d\mu(y)\right)^{1/p} \le C\sqrt{t} \| |\nabla u| \|_{L^p(X)}.$$

Proof. Let $u \in L^p(X)$ and $t > 0$ be fixed in the above argument. By an application of Fubini's theorem we have

$$\left(\int_X \int_X |P_t u(x) - u(y)|^p p_t(x,y) d\mu(x) d\mu(y)\right)^{1/p} = \left(\int_X P_t(|P_t u(x) - u|^p)(x) d\mu(x)\right)^{1/p}.$$

The main idea now is to adapt the proof of [10, Theorem 6.2]. As above, let q be the conjugate of p. Let $x \in X$ be fixed. Let g be a function in $L^\infty(X)$ such that $P_t(|g|^q)(x) \le 1$.

We first note that from the chain rule:

$$\begin{aligned}
&\partial_s \left[P_s((P_{t-s}u)(P_{t-s}g))(x)\right] \\
=&LP_s((P_{t-s}u)(P_{t-s}g))(x) - P_s((LP_{t-s}u)(P_{t-s}g))(x) - P_s((P_{t-s}u)(LP_{t-s}g))(x) \\
=&P_s(L(P_{t-s}u)(P_{t-s}g))(x) - P_s((LP_{t-s}u)(P_{t-s}g))(x) - P_s((P_{t-s}u)(LP_{t-s}g))(x) \\
=&2P_s(\Gamma(P_{t-s}u, P_{t-s}g)).
\end{aligned}$$

Therefore we have

$$\begin{aligned}
P_t((u - P_t u(x))g)(x) &= P_t(ug)(x) - P_t u(x) P_t g(x) \\
&= \int_0^t \partial_s \left[P_s((P_{t-s}u)(P_{t-s}g))(x)\right] ds \\
&= 2\int_0^t P_s\left(\Gamma(P_{t-s}u, P_{t-s}g)\right)(x) ds \\
&\le 2\int_0^t P_s\left(|\nabla P_{t-s}u||\nabla P_{t-s}g|\right)(x) ds \\
&\le 2\int_0^t P_s\left(|\nabla P_{t-s}u|^p\right)^{1/p}(x) P_s\left(|\nabla P_{t-s}g|^q\right)^{1/q}(x) ds.
\end{aligned}$$

Now from the strong Bakry-Émery estimate and Hölder's inequality we have

$$P_s\left(|\nabla P_{t-s}u|^p\right)^{1/p}(x) \le C P_s\left(P_{t-s}(|\nabla u|^p)\right)^{1/p}(x) = C P_t(|\nabla u|^p)^{1/p}(x).$$

On the other hand, Corollary 4.13 gives

$$|\nabla P_{t-s}g|^q \le \frac{C}{(t-s)^{q/2}} P_{t-s}(|g|^q).$$

Thus,

$$P_s\left(|\nabla P_{t-s}g|^q\right)^{1/q}(x) \le \frac{C}{(t-s)^{1/2}} P_t(|g|^q)^{1/q}(x) \le \frac{C}{(t-s)^{1/2}}.$$

One concludes

$$P_t((u - P_t u(x))g)(x) \le C\sqrt{t} P_t(|\nabla u|^p)^{1/p}(x).$$

Thus by $L^p - L^q$ duality in $(X, p_t(x,y)\mu(dy))$, one concludes

$$P_t(|u - P_t u(x)|^p)(x)^{1/p} \le C\sqrt{t} P_t(|\nabla u|^p)^{1/p}(x)$$

and finishes the proof by integration over X. □

We are finally in a position to prove the inclusion of the Sobolev space $W^{1,p}(X)$ into the Besov class $\mathbf{B}^{p,1/2}$, which in turn completes the proof of Theorem 4.9, which is the main result of this section.

Theorem 4.17. *Let $p>1$. There exists a constant $C>0$ such that for every $u \in W^{1,p}(X)$,*

$$\|u\|_{p,1/2} \le C \||\nabla u|\|_{L^p(X)}.$$

Proof. We first assume $u \in L^p(X) \cap \mathcal{F}$ with $|\nabla u| \in L^p(X)$. One has

$$\begin{aligned}
&\left(\int_X \int_X |u(x)-u(y)|^p p_t(x,y) d\mu(x) d\mu(y)\right)^{1/p} \\
&\le \left(\int_X \int_X |u(x)-P_t u(x)|^p p_t(x,y) d\mu(x) d\mu(y)\right)^{1/p} + \left(\int_X \int_X |P_t u(x)-u(y)|^p p_t(x,y) d\mu(x) d\mu(y)\right)^{1/p} \\
&\le \|P_t u - u\|_{L^p(X)} + \left(\int_X \int_X |P_t u(x)-u(y)|^p p_t(x,y) d\mu(x) d\mu(y)\right)^{1/p} \\
&\le 2C\sqrt{t} \||\nabla u|\|_{L^p(X)},
\end{aligned}$$

where in the last step we applied Lemma 4.15 to the first term and Lemma 4.16 to the second term. Thus

$$\|u\|_{p,1/2} \le C \||\nabla u|\|_{L^p(X)}.$$

Now let $u \in W^{1,p}(X)$ and choose an increasing sequence of functions $\phi_n \in C^\infty([0,\infty))$ such that $\phi_n \equiv 1$ on $[0,n]$, $\phi_n \equiv 0$ outside $[0,2n]$, and $|\phi_n'| \le \frac{2}{n}$. Let $x_0 \in X$. If $h_n(x) = \phi_n(d(x_0,x))$ then $h_n u \in \mathcal{F}$, $h_n \nearrow 1$ on X as $n \to \infty$, and $\||\nabla(h_n u)|\|_{L^p(X)} \to \||\nabla u|\|_{L^p(X)}$. Taking the limit in the inequality

$$\|h_n u\|_{p,1/2} \le C \||\nabla(h_n u)|\|_{L^p(X)}$$

yields the result. □

4.5 Continuity of P_t in the Besov spaces and critical exponents

We first note the following continuity property of P_t in the Besov spaces.

Proposition 4.18. *Suppose that the strong Bakry-Émery condition (8) holds. Let $p>1$. There exists a constant $C_p>0$ such that for every $f \in L^p(X,\mu)$ and $t>0$*

$$\|P_t f\|_{p,1/2} \le \frac{C_p}{t^{1/2}} \|f\|_{L^p(X)}.$$

Proof. This is a consequence of Lemma 4.14 and Theorem 4.17. □

Remark 4.19. *The above result is true without the strong Bakry-Émery condition for $1<p\le 2$ on very general Dirichlet spaces, see [3, Theorem 5.1].*

For $p \ge 1$, as in [3], we define the L^p Besov density critical exponent $\alpha_p^*(X)$ and triviality critical exponent $\alpha_p^\#(X)$ as follows:

$$\alpha_p^*(X) = \sup\{\alpha > 0 \,:\, \mathbf{B}^{p,\alpha}(X) \text{ is dense in } L^p(X)\},$$

$$\alpha_p^\#(X) = \sup\{\alpha > 0 \,:\, \mathbf{B}^{p,\alpha}(X) \text{ contains non-constant functions}\}.$$

Theorem 4.20. *Suppose that the weak Bakry-Émery condition (7) holds, then for $1 \le p \le 2$,*

$$\alpha_p^*(X) = \alpha_p^\#(X) = \frac{1}{2}.$$

Furthermore, if the strong Bakry-Émery condition (8) holds, then for every $p > 2$,

$$\alpha_p^*(X) = \alpha_p^\#(X) = \frac{1}{2}.$$

Proof. Assume that the weak Bakry-Émery condition (7) holds and begin with the case $p = 1$. Let $f \in \mathbf{B}^{1,\alpha}(X)$ with $\alpha > 1/2$. Since $\mathbf{B}^{1,\alpha}(X) \subset \mathbf{B}^{1,1/2}(X) = BV(X)$, we deduce that f is a BV function. Now since $f \in \mathbf{B}^{1,\alpha}(X)$, one has for every $t > 0$,

$$\int_X \int_X p_t(x,y)|f(x) - f(y)|d\mu(x)d\mu(y) \le t^\alpha \|f\|_{1,\alpha}.$$

By using the gaussian heat kernel lower bound we obtain

$$\liminf_{\varepsilon \to 0^+} \frac{1}{\varepsilon} \iint_{\Delta_\varepsilon} \frac{|f(y) - f(x)|}{\mu(B(x,\varepsilon))} d\mu(x)d\mu(y) = 0,$$

so $\|Df\|(X) = 0$, and from Remark 4.6 one gets that f is constant. It follows that $\alpha_1^\#(X) \le 1/2$. On the other hand, from Corollary 4.8 in [3], $\mathbf{B}^{1,1/2}(X)$ is dense in $L^1(X)$, so $\alpha_1^*(X) = \alpha_1^\#(X) = \frac{1}{2}$. From Proposition 5.6 in [3], one has:

1. Both $p \mapsto \alpha_p^*(X)$ and $p \mapsto \alpha_p^\#(X)$ are non-increasing;
2. For $1 \le p \le 2$ we have $\alpha_p^\#(X) \ge \alpha_p^*(X) \ge \frac{1}{2}$.

Therefore, for $1 \le p \le 2$ we also have $\alpha_p^*(X) = \alpha_p^\#(X) = \frac{1}{2}$.

Now let $p > 2$ and assume the strong Bakry-Émery condition (8). In that case, according to Proposition 4.18, for every $f \in L^p(X)$ and $t > 0$ one has $P_t f \in \mathbf{B}^{p,1/2}(X)$. Thus, $\mathbf{B}^{p,1/2}(X)$ is dense in $L^p(X)$ by strong continuity of the semigroup P_t in $L^p(X)$. Hence $\alpha_p^*(X) \ge 1/2$. Using again the fact that both $p \mapsto \alpha_p^*(X)$ and $p \mapsto \alpha_p^\#(X)$ are non-increasing and moreover that $\alpha_2^*(X) = \alpha_2^\#(X) = \frac{1}{2}$, one concludes that for every $p > 2$, $\alpha_p^*(X) = \alpha_p^\#(X) = \frac{1}{2}$. □

5 Sobolev and isoperimetric inequalities

Combining the conclusions in this paper with the results in [3, Section 6], we immediately obtain the following results that generalize the Sobolev embedding theorems from the classical Euclidean setting (see for example [70]) and metric upper gradient setting (see for example [39] and [37]) to the setting of Dirichlet forms and BV functions.

The following proposition is a weak-type version of the standard Sobolev embedding theorem. It gives weak-L^q control of the Besov function f, with q the Sobolev conjugate of p, and can therefore be used to control the L^s-norm of f in terms of the Besov norm of f when $1 \le s < pQ/(Q-p)$.

Proposition 5.1. *If the volume growth condition $\mu(B(x,r)) \ge C_1 r^Q$, $r \ge 0$, is satisfied for some $Q > 0$ then one has the following weak type Besov space embedding. Let $0 < \delta < Q$ and $1 \le p < \frac{Q}{\delta}$. Then there exists a constant $C_{p,\delta} > 0$ such that for every $f \in \mathbf{B}^{p,\delta/2}(X)$,*

$$\sup_{s \ge 0} s\mu\left(\{x \in X, |f(x)| \ge s\}\right)^{\frac{1}{q}} \le C_{p,\delta} \sup_{r>0} \frac{1}{r^{\delta+Q/p}} \left(\iint_{\{(x,y)\in X\times X | d(x,y)<r\}} |f(x) - f(y)|^p \, d\mu(x)\, d\mu(y) \right)^{1/p}$$

where $q = \frac{pQ}{Q-p\delta}$*. Furthermore, for every* $0 < \delta < Q$*, there exists a constant* $C_{\mathrm{iso},\delta}$ *such that for every measurable* $E \subset X$*,* $\mu(E) < +\infty$*,*

$$\mu(E)^{\frac{Q-\delta}{Q}} \le C_{\mathrm{iso},\delta} \sup_{r>0} \frac{1}{r^{\delta+Q}} (\mu \otimes \mu) \{(x,y) \in E \times E^c \,:\, d(x,y) \le r\}$$

Proof. From the heat kernel upper bound (5), the volume growth condition $\mu(B(x,r)) \ge C_1 r^Q$, $r \ge 0$, implies the ultracontractive estimate

$$p_t(x,y) \le \frac{C}{t^{Q/2}}. \tag{20}$$

We are therefore in the framework of Theorem 6.1 in [3], from which one obtains that there is a constant $C_{p,\delta} > 0$ such that for every $f \in \mathbf{B}^{p,\delta/2}(X)$,

$$\sup_{s \ge 0} s\, \mu\left(\{x \in X \,:\, |f(x)| \ge s\}\right)^{\frac{1}{q}} \le C_{p,\delta} \|f\|_{p,\delta/2}$$

where $q = \frac{pQ}{Q-p\delta}$. The conclusion follows from Theorem 4.2. □

In Euclidean space there is a standard method for using the above weak-type Sobolev embedding to obtain the usual Sobolev embedding theorem, in which the weak-L^q control of f is replaced by the strong-L^q control. However this approach uses locality properties which need not be valid for the Besov seminorm $\|\cdot\|_{p,\alpha}$. We direct the interested reader to [37] for more details on this topic.

The one circumstance we have investigated in which the Besov seminorm has a locality property arose in Theorem 4.4, see also Remark 4.5, for the space $\mathbf{B}^{1,1/2}$ under the assumption of a weak Bakry-Émery estimate, in which case we had $\mathbf{B}^{1,1/2} = BV(X)$. This locality property lets us obtain a standard Sobolev embedding in which the L^q norm is controlled by the BV norm. We may view this as an extension of known results on Riemannian manifolds with non-negative Ricci curvature (see Theorem 8.4 in [66]) or on Carnot groups (see [88]) to our metric measure Dirichlet setting under the further hypothesis that there is a weak Bakry-Émery estimate.

Theorem 5.2. *Suppose that the weak Bakry-Émery estimate (7) is satisfied. If the volume growth condition* $\mu(B(x,r)) \ge C_1 r^Q$*,* $r \ge 0$*, is satisfied for some* $Q > 0$*, then there exists a constant* $C_2 > 0$ *such that for every* $f \in BV(X)$*,*

$$\|f\|_{L^q(X)} \le C_2 \|Df\|(X)$$

where $q = \frac{Q}{Q-1}$*. In particular, if* E *is a set with finite perimeter in* X*, then*

$$\mu(E)^{\frac{Q-1}{Q}} \le C_2 P(E,X).$$

Proof. Observe that as in the above proof, the heat kernel satisfies the ultracontractive estimate (20). From Theorem 4.4 we have

$$\|f\|_{1,1/2} \le C \liminf_{s \to 0} s^{-1/2} \int_X P_s(|f - f(y)|)(y) d\mu(y).$$

This verifies a condition denoted by $(P_{1,1/2})$ in Definition 6.7 of [3]), putting us in the framework of [3, Theorem 6.9] with $p=1$, $\alpha = 1/2$ and $\beta = Q/2$. Notice also that $\|f\|_{1,1/2} \le C\|Df\|(X)$ from Theorem 4.4, so we have

$$\|f\|_{L^q(X)} \le C\|f\|_{1,1/2} \le C_2 \|Df\|(X),$$

where $q = \frac{Q}{Q-1}$. Taking $f = \mathbf{1}_E$ then yields

$$\mu(E)^{\frac{Q-1}{Q}} \le C_2 P(E,X).$$

□

References

[1] Robert A. Adams and John J. F. Fournier. *Sobolev spaces*, volume 140 of *Pure and Applied Mathematics (Amsterdam)*. Elsevier/Academic Press, Amsterdam, second edition, 2003. 2

[2] P. Alonso-Ruiz. Explicit formulas for heat kernels on diamond fractals. *Communications in Mathematical Physics*, 364(3):1305–1326, Dec 2018. 16

[3] P. Alonso-Ruiz, F. Baudoin, L. Chen, L. Rogers, N. Shanmugalingam, and A. Teplyaev. Besov class via heat semigroup on Dirichlet spaces I: Sobolev type inequalities. *preprint*, 2018. 2, 3, 9, 18, 20, 26, 31, 32, 33

[4] P. Alonso-Ruiz, F. Baudoin, L. Chen, L. Rogers, N. Shanmugalingam, and A. Teplyaev. Besov class via heat semigroup on Dirichlet spaces III: BV functions and sub-Gaussian heat kernel estimates. *preprint*, 2018. 3

[5] Patricia Alonso-Ruiz, Daniel J. Kelleher, and Alexander Teplyaev. Energy and Laplacian on Hanoi-type fractal quantum graphs. *J. Phys. A*, 49(16):165206, 36, 2016. 5, 16

[6] L. Ambrosio, M. Miranda, Jr., and D. Pallara. Special functions of bounded variation in doubling metric measure spaces. In *Calculus of variations: topics from the mathematical heritage of E. De Giorgi*, volume 14 of *Quad. Mat.*, pages 1–45. Dept. Math., Seconda Univ. Napoli, Caserta, 2004. 2, 24

[7] Luigi Ambrosio, Nicola Fusco, and Diego Pallara. *Functions of bounded variation and free discontinuity problems*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 2000. 12

[8] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. Metric measure spaces with Riemannian Ricci curvature bounded from below. *Duke Math. J.*, 163(7):1405–1490, 2014. 7

[9] Pascal Auscher, Thierry Coulhon, Xuan Thinh Duong, and Steve Hofmann. Riesz transform on manifolds and heat kernel regularity. *Ann. Sci. École Norm. Sup. (4)*, 37(6):911–957, 2004. 26, 28

[10] Dominique Bakry, Fabrice Baudoin, Michel Bonnefont, and Djalil Chafaï. On gradient bounds for the heat kernel on the Heisenberg group. *J. Funct. Anal.*, 255(8):1905–1938, 2008. 10, 30

[11] Viorel Barbu and Michael Röckner. Stochastic variational inequalities and applications to the total variation flow perturbed by linear multiplicative noise. *Archive for Rational Mechanics and Analysis*, 209(3):797–834, Sep 2013. 16

[12] Fabrice Baudoin and Michel Bonnefont. Log-Sobolev inequalities for subelliptic operators satisfying a generalized curvature dimension inequality. *J. Funct. Anal.*, 262(6):2646–2676, 2012. 7, 10

[13] Fabrice Baudoin and Michel Bonnefont. Reverse Poincaré inequalities, isoperimetry, and Riesz transforms in Carnot groups. *Nonlinear Anal.*, 131:48–59, 2016. 10

[14] Fabrice Baudoin, Michel Bonnefont, and Nicola Garofalo. A sub-Riemannian curvature-dimension inequality, volume doubling property and the Poincaré inequality. *Math. Ann.*, 358(3-4):833–860, 2014. 10

[15] Fabrice Baudoin and Nicola Garofalo. Curvature-dimension inequalities and Ricci lower bounds for sub-Riemannian manifolds with transverse symmetries. *J. Eur. Math. Soc. (JEMS)*, 19(1):151–219, 2017. 10

[16] Fabrice Baudoin and Daniel J Kelleher. Differential one-forms on Dirichlet spaces and Bakry-Emery estimates on metric graphs. *arXiv:1604.02520, to apear in Trans. Amer. Math. Soc.*, pages 1–42, 2017. 10

[17] Fabrice Baudoin and Bumsik Kim. Sobolev, Poincaré, and isoperimetric inequalities for subelliptic diffusion operators satisfying a generalized curvature dimension inequality. *Rev. Mat. Iberoam.*, 30(1):109–131, 2014. 7, 10

[18] Colin Bennett and Robert Sharpley. *Interpolation of operators*, volume 129 of *Pure and Applied Mathematics*. Academic Press, Inc., Boston, MA, 1988. 9

[19] A. Beurling and J. Deny. Espaces de Dirichlet. I. Le cas élémentaire. *Acta Math.*, 99:203–224, 1958. 4

[20] Marco Biroli and Umberto Mosco. Sobolev and isoperimetric inequalities for Dirichlet forms on homogeneous spaces. *Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl.*, 6(1):37–44, 1995. 4

[21] Marco Biroli and Umberto Mosco. Sobolev inequalities on homogeneous spaces. *Potential Anal.*, 4(4):311–324, 1995. Potential theory and degenerate partial differential operators (Parma). 4, 5

[22] Marco Biroli and Umberto Mosco. Kato space for Dirichlet forms. *Potential Anal.*, 10(4):327–345, 1999. 4, 5

[23] Jeff Cheeger. Differentiability of lipschitz functions on metric measure spaces. *Geom. Funct. Anal.*, 9:428–517, 1999. 2, 4

[24] Zhen-Qing Chen and Masatoshi Fukushima. *Symmetric Markov processes, time change, and boundary theory*, volume 35 of *London Mathematical Society Monographs Series*. Princeton University Press, Princeton, NJ, 2012. 4

[25] Thierry Coulhon, Renjin Jiang, Pekka Koskela, and Adam Sikora. Gradient estimates for heat kernels and harmonic functions. *arXiv:1703.02152*, 2017. 10

[26] E. De Giorgi and G. Letta. Une notion générale de convergence faible pour des fonctions croissantes d'ensemble. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 4(1):61–99, 1977. 12

[27] Simone Di Marino and Marco Squassina. New characterizations of sobolev metric spaces, year = 2018. *Journal of Functional Analysis*. 9

[28] Nathaniel Eldredge. Gradient estimates for the subelliptic heat kernel on H-type groups. *J. Funct. Anal.*, 258(2):504–533, 2010. 10

[29] Matthias Erbar, Kazumasa Kuwada, and Karl-Theodor Sturm. On the equivalence of the entropic curvature-dimension condition and Bochner's inequality on metric measure spaces. *Invent. Math.*, 201(3):993–1071, 2015. 10

[30] Lawrence C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010. 8

[31] Daniel Fontaine, Thomas Smith, and Alexander Teplyaev. Resistance of random Sierpiński gaskets. In *Quantum graphs and their applications*, volume 415 of *Contemp. Math.*, pages 121–136. Amer. Math. Soc., Providence, RI, 2006. 5, 16

[32] Masatoshi Fukushima, Yoichi Oshima, and Masayoshi Takeda. *Dirichlet forms and symmetric Markov processes*, volume 19 of *De Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, extended edition, 2011. 3, 4, 6, 28

[33] Nicola Garofalo and Duy-Minh Nhieu. Isoperimetric and Sobolev inequalities for Carnot-Carathéodory spaces and the existence of minimal surfaces. *Comm. Pure Appl. Math.*, 49(10):1081–1144, 1996. 8

[34] Amiran Gogatishvili, Pekka Koskela, and Nageswari Shanmugalingam. Interpolation properties of Besov spaces defined on metric spaces. *Math. Nachr.*, 283(2):215–231, 2010. 3, 19

[35] A. A. Grigor'yan. The heat equation on noncompact Riemannian manifolds. *Mat. Sb.*, 182(1):55–87, 1991. 9

[36] Sebastian Haeseler. Heat kernel estimates and related inequalities on metric graphs, year = 2011. *preprint*. 7

[37] Piotr Hajł asz and Pekka Koskela. Sobolev met Poincaré. *Mem. Amer. Math. Soc.*, 145(688):x+101, 2000. 32, 33

[38] Juha Heinonen. *Lectures on analysis on metric spaces*. Universitext. Springer-Verlag, New York, 2001. 5

[39] Juha Heinonen, Pekka Koskela, Nageswari Shanmugalingam, and Jeremy T. Tyson. *Sobolev spaces on metric measure spaces*, volume 27 of *New Mathematical Monographs*. Cambridge University Press, Cambridge, 2015. An approach based on upper gradients. 2, 7, 9, 32

[40] Michael Hinz, Daniel J. Kelleher, and Alexander Teplyaev. Metrics and spectral triples for Dirichlet and resistance forms. *J. Noncommut. Geom.*, 9(2):359–390, 2015. 5

[41] Michael Hinz, Michael Röckner, and Alexander Teplyaev. Vector analysis for Dirichlet forms and quasilinear PDE and SPDE on metric measure spaces. *Stochastic Process. Appl.*, 123(12):4373–4406, 2013. 5, 16

[42] Michael Hinz and Alexander Teplyaev. Dirac and magnetic Schrödinger operators on fractals. *J. Funct. Anal.*, 265(11):2830–2854, 2013. 5, 16

[43] Michael Hinz and Alexander Teplyaev. Vector analysis on fractals and applications. In *Fractal geometry and dynamical systems in pure and applied mathematics. II. Fractals in applied mathematics*, volume 601 of *Contemp. Math.*, pages 147–163. Amer. Math. Soc., Providence, RI, 2013. 5

[44] Michael Hinz and Martina Zähle. Semigroups, potential spaces and applications to (S)PDE. *Potential Anal.*, 36(3):483–515, 2012. 16

[45] Renjin Jiang. The Li-Yau inequality and heat kernels on metric measure spaces. *J. Math. Pures Appl. (9)*, 104(1):29–57, 2015. 10

[46] A. Jonsson and H. Wallin. A trace theorem for generalized Besov spaces with three indexes. In *Fourier analysis and approximation theory (Proc. Colloq., Budapest, 1976), Vol. I*, volume 19 of *Colloq. Math. Soc. János Bolyai*, pages 429–449. North-Holland, Amsterdam-New York, 1978. 9

[47] A. Jonsson and H. Wallin. A Whitney extension theorem in L_p and Besov spaces. *Ann. Inst. Fourier (Grenoble)*, 28(1):vi, 139–192, 1978. 9

[48] Naotaka Kajino. Spectral asymptotics for Laplacians on self-similar sets. *J. Funct. Anal.*, 258(4):1310–1360, 2010. 2

[49] Naotaka Kajino. Heat kernel asymptotics for the measurable Riemannian structure on the Sierpinski gasket. *Potential Anal.*, 36(1):67–115, 2012. 5, 16

[50] Naotaka Kajino. Analysis and geometry of the measurable Riemannian structure on the Sierpiński gasket. In *Fractal geometry and dynamical systems in pure and applied mathematics. I. Fractals in pure mathematics*, volume 600 of *Contemp. Math.*, pages 91–133. Amer. Math. Soc., Providence, RI, 2013. 5, 16

[51] Davar Khoshnevisan, Kunwoo Kim, and Yimin Xiao. Intermittency and multifractality: a case study via parabolic stochastic PDEs. *Ann. Probab.*, 45(6A):3697–3751, 2017. 16

[52] Davar Khoshnevisan, Kunwoo Kim, and Yimin Xiao. A macroscopic multifractal analysis of parabolic stochastic PDEs. *Comm. Math. Phys.*, 360(1):307–346, 2018. 16

[53] J. Kigami. Harmonic metric and Dirichlet form on the Sierpiński gasket. In *Asymptotic problems in probability theory: stochastic models and diffusions on fractals (Sanda/Kyoto, 1990)*, volume 283 of *Pitman Res. Notes Math. Ser.*, pages 201–218. Longman Sci. Tech., Harlow, 1993. 5

[54] Jun Kigami. Harmonic calculus on p.c.f. self-similar sets. *Trans. Amer. Math. Soc.*, 335(2):721–755, 1993. 2, 16

[55] Jun Kigami. *Analysis on fractals*, volume 143 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 2001. 2, 5, 16

[56] Jun Kigami. Measurable Riemannian geometry on the Sierpinski gasket: the Kusuoka measure and the Gaussian heat kernel estimate. *Math. Ann.*, 340(4):781–804, 2008. 5, 16

[57] Jun Kigami. Volume doubling measures and heat kernel estimates on self-similar sets. *Mem. Amer. Math. Soc.*, 199(932):viii+94, 2009. 2, 5

[58] Jun Kigami. Resistance forms, quasisymmetric maps and heat kernel estimates. *Mem. Amer. Math. Soc.*, 216(1015):vi+132, 2012. 5

[59] Juha Kinnunen, Riikka Korte, Nageswari Shanmugalingam, and Heli Tuominen. A characterization of Newtonian functions with zero boundary values. *Calc. Var. Partial Differential Equations*, 43(3-4):507–528, 2012. 24

[60] Pekka Koskela, Kai Rajala, and Nageswari Shanmugalingam. Lipschitz continuity of Cheeger-harmonic functions in metric measure spaces. *J. Funct. Anal.*, 202(1):147–173, 2003. 2, 10

[61] Pekka Koskela, Nageswari Shanmugalingam, and Jeremy T. Tyson. Dirichlet forms, Poincaré inequalities, and the Sobolev spaces of Korevaar and Schoen. *Potential Anal.*, 21(3):241–262, 2004. 2, 7

[62] Pekka Koskela, Nageswari Shanmugalingam, and Yuan Zhou. Geometry and analysis of Dirichlet forms (II). *J. Funct. Anal.*, 267(7):2437–2477, 2014. 7

[63] Pekka Koskela and Yuan Zhou. Geometry and analysis of Dirichlet forms. *Adv. Math.*, 231(5):2755–2801, 2012. 5, 7, 16

[64] Brett L. Kotschwar. Hamilton's gradient estimate for the heat kernel on complete manifolds. *Proc. Amer. Math. Soc.*, 135(9):3013–3019, 2007. 28

[65] Shigeo Kusuoka. Dirichlet forms on fractals and products of random matrices. *Publ. Res. Inst. Math. Sci.*, 25(4):659–680, 1989. 16

[66] Michel Ledoux. Isoperimetry and Gaussian analysis. In *Lectures on probability theory and statistics (Saint-Flour, 1994)*, volume 1648 of *Lecture Notes in Math.*, pages 165–294. Springer, Berlin, 1996. 33

[67] Daniel Lenz, Peter Stollmann, and Ivan Veselić. Generalized eigenfunctions and spectral theory for strongly local Dirichlet forms. In *Spectral theory and analysis*, volume 214 of *Oper. Theory Adv. Appl.*, pages 83–106. Birkhäuser/Springer Basel AG, Basel, 2011. 4

[68] John Mackay, Jeremy T. Tyson, and Kevin Wildrick. Modulus and poincaré inequalities on non-self-similar sierpiński carpets. *Geom. Funct. Anal.*, 23:985–1034, 2013. 2

[69] Niko Marola, Michele Miranda, Jr., and Nageswari Shanmugalingam. Characterizations of sets of finite perimeter using heat kernels in metric spaces. *Potential Anal.*, 45(4):609–633, 2016. 4, 21, 23

[70] Vladimir Maz'ya. *Sobolev spaces; with applications to elliptic partial differential equations*. Springer Verlag Berlin Heidelberg, Grundlehren der mathematischen Wissenschaften **342**, 2011. 2, 8, 32

[71] Robert Meyers, Robert S. Strichartz, and Alexander Teplyaev. Dirichlet forms on the Sierpiński gasket. *Pacific J. Math.*, 217(1):149–174, 2004. 5, 16

[72] Michele Miranda, Jr. Functions of bounded variation on "good" metric spaces. *J. Math. Pures Appl. (9)*, 82(8):975–1004, 2003. 2, 12, 16

[73] Ionuţ Munteanu and Michael Röckner. Total variation flow perturbed by gradient linear multiplicative noise. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.*, 21(1):1850003, 28, 2018. 16

[74] Kasso A. Okoudjou, Luke G. Rogers, and Robert S. Strichartz. Szegö limit theorems on the Sierpiński gasket. *J. Fourier Anal. Appl.*, 16(3):434–447, 2010. 2

[75] Katarzyna Pietrushka-Pałuba. Heat kernel characterisation of besov-lipschitz spaces on metric measure spaces. *Manuscripta Math.*, pages 131–199, 2010. 2, 3

[76] Luke G. Rogers and Alexander Teplyaev. Laplacians on the basilica Julia sets. *Commun. Pure Appl. Anal.*, 9(1):211–231, 2010. 2

[77] Laurent Saloff-Coste. A note on Poincaré, Sobolev, and Harnack inequalities. *Internat. Math. Res. Notices*, (2):27–38, 1992. 9

[78] Laurent Saloff-Coste. *Aspects of Sobolev-type inequalities*, volume 289 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 2002. 9

[79] Robert S. Strichartz. Fractafolds based on the Sierpiński gasket and their spectra. *Trans. Amer. Math. Soc.*, 355(10):4019–4043, 2003. 2

[80] Robert S. Strichartz. Function spaces on fractals. *J. Funct. Anal.*, 198(1):43–83, 2003. 2

[81] Robert S. Strichartz and Alexander Teplyaev. Spectral analysis on infinite Sierpiński fractafolds. *J. Anal. Math.*, 116:255–297, 2012. 2

[82] K. T. Sturm. Analysis on local Dirichlet spaces. III. The parabolic Harnack inequality. *J. Math. Pures Appl. (9)*, 75(3):273–297, 1996. 3, 4, 9

[83] Karl-Theodor Sturm. Analysis on local Dirichlet spaces. I. Recurrence, conservativeness and L^p-Liouville properties. *J. Reine Angew. Math.*, 456:173–196, 1994. 3, 4, 5, 13

[84] Karl-Theodor Sturm. Analysis on local Dirichlet spaces. II. Upper Gaussian estimates for the fundamental solutions of parabolic equations. *Osaka J. Math.*, 32(2):275–312, 1995. 3, 4, 9

[85] Alexander Teplyaev. Energy and Laplacian on the Sierpiński gasket. In *Fractal geometry and applications: a jubilee of Benoît Mandelbrot. Part 1*, volume 72 of *Proc. Sympos. Pure Math.*, pages 131–154. Amer. Math. Soc., Providence, RI, 2004. 5, 16

[86] Alexander Teplyaev. Harmonic coordinates on fractals with finitely ramified cell structure. *Canad. J. Math.*, 60(2):457–480, 2008. 5, 16

[87] H. Triebel. *Interpolation theory, function spaces, differential operators*. VEB Deutscher Verlag der Wissenschaften, Berlin, 1978. 9

[88] N. Th. Varopoulos. Small time Gaussian estimates of heat diffusion kernels. I. The semigroup technique. *Bull. Sci. Math.*, 113(3):253–277, 1989. 33