From Harnack inequality to heat kernel estimates on metric measure spaces and applications

Luca Tamanini *

July 16, 2019

Abstract

Aim of this short note is to show that a dimension-free Harnack inequality on an infinitesimally Hilbertian metric measure space where the heat semigroup admits an integral representation in terms of a kernel is sufficient to deduce a sharp upper Gaussian estimate for such kernel. As intermediate step, we prove the local logarithmic Sobolev inequality (known to be equivalent to a lower bound on the Ricci curvature tensor in smooth Riemannian manifolds). Both results are new also in the more regular framework of $\mathsf{RCD}(K,\infty)$ spaces.

Contents

1 Introduction

In a smooth Riemannian manifold (M, g, vol) with Ricci curvature bounded from below it is well-known [23] that the heat kernel r_t, namely the positive fundamental solution of the heat equation defined by the Laplace-Beltrami operator (or, from a probabilistic viewpoint, the transition probability of the Brownian motion), satisfies two-sided Gaussian estimates, which read as

$$\frac{1}{C_1 \mathrm{vol}(B_{\sqrt{t}}(y))} \exp\Big(-\frac{\mathsf{d}^2(x,y)}{(4-\varepsilon)t} - C_2 t\Big) \leq \mathsf{r}_t[x](y) \leq \frac{C_1}{\mathrm{vol}(B_{\sqrt{t}}(y))} \exp\Big(-\frac{\mathsf{d}^2(x,y)}{(4+\varepsilon)t} + C_2 t\Big) \quad (1.1)$$

for every $\varepsilon > 0$, $x, y \in \mathrm{X}$ and $t > 0$, for suitable positive constants $C_1 = C_1(K, N, \delta)$ and $C_2 = C_2(K, N, \delta)$, where K is a lower bound on the Ricci curvature and N an upper bound for the dimension. But because of the great interest in the study of metric (measure) spaces

*Institut für Angewandte Mathematik, Universität Bonn. email: tamanini@iam.uni-bonn.de

(we refer to [17] for an overview on the topic and detailed bibliography) and the essentially metric-measure nature of (1.1), it is rather natural to investigate possible generalizations of these kind of bounds to more abstract settings. For regular, symmetric, strongly local Dirichlet forms on locally compact separable Hausdorff spaces satisfying doubling & Poincaré, this has been achieved by K.-T. Sturm in [24], while more recently R. Jiang, H. Li and H. Zhang studied the problem on finite-dimensional $\mathsf{RCD}^*(K,N)$ spaces [19], which are still locally compact but only locally doubling for $K<0$. This class of metric measure spaces with Ricci curvature bounded from below, introduced in [5], [13] starting from the seminal papers [21], [25], [26], is the natural framework where bounds like (1.1) can be expected to hold, since it enjoys many analytical and geometric properties and the existence of the heat kernel is well understood.

As concerns weighted Riemannian manifolds (M,g,μ), the picture is not so clear, as to the best of our knowledge no lower bounds are known (unless M has finite volume, see [28]), whereas an upper estimate can still be deduced [15], [14] and reads as

$$\mathsf{r}_t[x](y) \leq \frac{1}{\sqrt{\mu(B_{\sqrt{t}}(x))\mu(B_{\sqrt{t}}(y))}} \exp\Big(C_\varepsilon(1+C_K t) - \frac{\mathsf{d}^2(x,y)}{(4+\varepsilon)t}\Big) \tag{1.2}$$

for every $\varepsilon>0$, $x,y\in \mathrm{X}$ and $t>0$; notice however that in general it is not possible to get rid of one of the two volumes on the right-hand side, since M does not need to be doubling. As in the previous discussion, also (1.2) has an essentially metric-measure nature and if on the one hand the non-smooth counterpart of smooth Riemannian manifolds with lower Ricci bounds is given by $\mathsf{RCD}^*(K,N)$ spaces, on the other hand $\mathsf{RCD}(K,\infty)$ spaces represent the natural alternative to weighted manifolds and in this setting (1.2) is still missing; therefore it would be natural, as a first attempt, to prove (1.2) on any $\mathsf{RCD}(K,\infty)$ space.

Yet, after a careful look at [19] one can observe that the key role in the proof of (1.1) is played by the dimension-dependent Harnack inequality [11], [18] and that not surprisingly in [14] the dimension-free Harnack inequality, which is known to hold on $\mathsf{RCD}(K,\infty)$ spaces after [20], is a crucial ingredient as well. For this reason, instead of establishing (1.2) on an $\mathsf{RCD}(K,\infty)$ space, we prefer to work in a more general environment, namely a metric measure space $(\mathrm{X},\mathsf{d},\mathfrak{m})$ supporting a dimension-free Harnack inequality where the heat semigroup admits an integral representation (see Setting 2.1 below). In order to achieve this goal we obtain a local logarithmic Sobolev inequality, that to the best of our knowledge was missing also in the context of $\mathsf{RCD}(K,\infty)$ spaces, and the L^∞-LIP regularization for the heat flow (see Proposition 3.2).

The paper is structured as follows. In Section 2 we recall all the relevant notions, results and bibliographical references related to calculus on metric measure spaces and point out the precise framework we shall work within. In Section 3 we collect some auxiliary results, most notably the local logarithmic Sobolev inequality. Finally, Section 4 is devoted to the proof of the upper Gaussian estimate for the heat kernel and some consequences.

Acknowledgements

The author gratefully acknowledges support by the European Union through the ERC-AdG "RicciBounds" for Prof. K. T. Sturm and would like to thank Prof. M. Gordina for useful suggestions.

2 Preliminaries and setting

By LIP(X) we denote the space of Lipschitz continuous functions and by $C([0,1],\mathrm{X})$ the space of continuous curves with values in the metric space (X,d). For the notion of absolutely continuous curve in a metric space and of metric speed see for instance Section 1.1 in [2]. The collection of absolutely continuous curves on $[0,1]$ is denoted by $AC([0,1],\mathrm{X})$. By $\mathscr{P}(\mathrm{X})$ we denote the space of Borel probability measures on (X,d) and by $\mathscr{P}_2(\mathrm{X}) \subset \mathscr{P}(\mathrm{X})$ the subclass of those with finite second moment.

Let $(\mathrm{X},\mathsf{d},\mathfrak{m})$ be a complete and separable metric measure space endowed with a Borel non-negative measure which is finite on bounded sets.

For the definition of the **Sobolev class** $S^2(\mathrm{X})$ and of **minimal weak upper gradient** $|Df|$ see [4] (and the previous works [9], [22] for alternative - but equivalent - definitions of Sobolev functions). The Sobolev space $W^{1,2}(\mathrm{X})$ is defined as $L^2(\mathrm{X}) \cap S^2(\mathrm{X})$. When endowed with the norm $\|f\|^2_{W^{1,2}} := \|f\|^2_{L^2} + \||Df|\|^2_{L^2}$, $W^{1,2}(\mathrm{X})$ is a Banach space. The **Cheeger energy** is the convex and lower-semicontinuous functional $E : L^2(\mathrm{X}) \to [0,\infty]$ given by

$$E(f) := \begin{cases} \dfrac{1}{2}\displaystyle\int |Df|^2 \,\mathrm{d}\mathfrak{m} & \text{for } f \in W^{1,2}(\mathrm{X}) \\ +\infty & \text{otherwise} \end{cases}$$

$(\mathrm{X},\mathsf{d},\mathfrak{m})$ is **infinitesimally Hilbertian** (see [13]) if $W^{1,2}(\mathrm{X})$ is Hilbert. In this case the **cotangent module** $L^2(T^*\mathrm{X})$ (see [12]) and its dual, the **tangent module** $L^2(T\mathrm{X})$, are canonically isomorphic, the **differential** is a well-defined linear map d from $S^2(\mathrm{X})$ with values in $L^2(T^*\mathrm{X})$ and the isomorphism sends the differential $\mathrm{d}f$ to the gradient ∇f. Furthermore E is a Dirichlet form admitting a carré du champ given by $\langle \nabla f, \nabla g\rangle$, where $\langle\cdot,\cdot\rangle$ is the pointwise scalar product on the Hilbert module $L^2(T\mathrm{X})$. The infinitesimal generator Δ of E, which is a closed self-adjoint linear operator on $L^2(\mathrm{X})$, is called **Laplacian** on $(\mathrm{X},\mathsf{d},\mathfrak{m})$ and its domain denoted by $D(\Delta) \subset W^{1,2}(\mathrm{X})$. A function $f \in W^{1,2}(\mathrm{X})$ belongs to $D(\Delta)$ and $g = \Delta f$ if and only if

$$\int \phi g \,\mathrm{d}\mathfrak{m} = -\int \langle \nabla\phi, \nabla f\rangle \,\mathrm{d}\mathfrak{m}, \qquad \forall \phi \in W^{1,2}(\mathrm{X}).$$

The flow (h_t) associated to E is called **heat flow** (see [4]), and for any $f \in L^2(\mathrm{X})$ the curve $t \mapsto \mathsf{h}_t f \in L^2(\mathrm{X})$ is continuous on $[0,\infty)$, locally absolutely continuous on $(0,\infty)$ and the only solution of

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathsf{h}_t f = \Delta \mathsf{h}_t f, \qquad \mathsf{h}_t f \to f \text{ as } t \downarrow 0.$$

After this preliminary part, we can describe the framework we shall work within throughout this note.

Setting 2.1. $(\mathrm{X},\mathsf{d},\mathfrak{m})$ is a complete and separable metric space equipped with a non-negative Borel measure which is finite on bounded sets and supports the following dimension-free Harnack inequality: for any $p \in (1,\infty)$, $x, y \in \mathrm{X}$ and for any $f \in L^1 \cap L^\infty(\mathrm{X})$ it holds

$$|\mathsf{h}_t f(x)|^p \le (\mathsf{h}_t |f|^p)(y) \exp\Big(\frac{pK\mathsf{d}^2(x,y)}{2(p-1)(e^{2Kt}-1)}\Big) \tag{2.1}$$

with $K \in \mathbb{R}$. We also assume that there exists a function

$$(0,\infty) \times \mathrm{X}^2 \ni (t,x,y) \quad \mapsto \quad \mathsf{r}_t[x](y) \in (0,\infty) \tag{2.2}$$

called heat kernel, such that for every $x \in \mathrm{X}$ and $t > 0$, $\mathsf{r}_t[x]$ is a probability density, and the following identity holds

$$\mathsf{h}_t f(x) = \int f(y)\mathsf{r}_t[x](y)\,\mathrm{d}\mathfrak{m}(y) \qquad \forall t > 0,\, \forall f \in L^2(\mathrm{X}). \tag{2.3}$$

By the arguments in [5] and [1] with slight adaptations, the semigroup property of h_t and the representation formula (2.3) entail that the heat kernel is symmetric, i.e. $\mathsf{r}_t[x](y) = \mathsf{r}_t[y](x)$ $\mathfrak{m} \otimes \mathfrak{m}$-a.e. in X^2, and satisfies the Chapman-Kolmogorov formula

$$\mathsf{r}_{t+s}[x](y) = \int \mathsf{r}_t[x](z)\mathsf{r}_s[z](y)\mathrm{d}\mathfrak{m}(z) \qquad \text{for } \mathfrak{m} \otimes \mathfrak{m}\text{-a.e. } (x,y) \in \mathrm{X}^2,\ \forall t, s \geq 0. \tag{2.4}$$

Moreover, (2.3) and the fact that $\mathsf{r}_t[x]$ is a probability density can also be used to extend the heat flow to $L^1(\mathrm{X})$, show that it is **mass preserving** and satisfies the **maximum principle**, i.e.

$$f \leq c \quad \mathfrak{m} - a.e. \qquad\qquad \Rightarrow \qquad\qquad \mathsf{h}_t f \leq c \quad \mathfrak{m}\text{-a.e.},\ \forall t > 0.$$

Finally, by Proposition 4.1 in [10] the dimension-free Harnack inequality (2.1) implies the strong Feller property for h_t for all $t > 0$, namely if $f \in L^\infty(\mathrm{X})$, then $\mathsf{h}_t f$ is continuous and bounded.

Let us recall that the minimal weak upper gradient is local (i.e. $|Df| = |Dg|$ $\mathfrak{m}$-a.e. on $\{f = g\}$), lower semicontinuous w.r.t. $\mathfrak{m}$-a.e. convergence and that Lipschitz functions with bounded support are dense in $L^p(\mathrm{X})$, $p \in (1, \infty)$ (see [3], where the density is actually proved in $W^{1,p}(\mathrm{X})$).

As regards the properties of the differential, the following calculus rules (see [12] for the proof) will be used extensively without further notice:

$$\begin{aligned}
|\mathrm{d}f| &= |Df| \quad \mathfrak{m}\text{-a.e.} && \forall f \in S^2(\mathrm{X}) \\
\mathrm{d}f &= \mathrm{d}g \qquad \mathfrak{m}\text{-a.e. on } \{f = g\} && \forall f, g \in S^2(\mathrm{X}) \\
\mathrm{d}(\varphi \circ f) &= \varphi' \circ f\, \mathrm{d}f && \forall f \in S^2(\mathrm{X}),\ \varphi : \mathbb{R} \to \mathbb{R} \text{ Lipschitz} \\
\mathrm{d}(fg) &= g\, \mathrm{d}f + f\, \mathrm{d}g && \forall f, g \in L^\infty \cap S^2(\mathrm{X})
\end{aligned}$$

where it is part of the properties the fact that $\varphi \circ f, fg \in S^2(\mathrm{X})$ for φ, f, g as above.

Finally, given $f : \mathrm{X} \to \mathbb{R}$, the **local Lipschitz constant** $\mathrm{lip}(f) : \mathrm{X} \to [0, \infty]$ is defined as 0 on isolated points and otherwise as

$$\mathrm{lip} f(x) := \limsup_{y \to x} \frac{|f(x) - f(y)|}{d(x,y)}.$$

If f is Lipschitz, then its Lipschitz constant is denoted by $\mathrm{Lip}(f)$. It is worth stressing that if f is Lipschitz with $\mathrm{lip}(f) \in L^2(\mathrm{X})$, then $f \in S^2(\mathrm{X})$ with

$$|Df| \leq \mathrm{lip}(f), \quad \mathfrak{m}\text{-}a.e. \tag{2.5}$$

Remark 2.2 (Lipschitz cut-off function)**.** Given a complete and separable metric measure space $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$ equipped with a non-negative Borel measure which is finite on bounded sets, for any $x \in \mathrm{X}$ and $r > 0$ there exists a Lipschitz function $\chi_r : \mathrm{X} \to [0, 1]$ such that $\chi_r \equiv 1$ on $B_r(x)$, $\mathrm{supp}(\chi_r) \subset B_{r+1}(x)$ and $\|\mathrm{lip}(\chi_r)\|_{L^\infty(\mathrm{X})}$ does not depend on r.

Indeed, if $\eta \in \mathrm{LIP}(\mathbb{R})$ with bounded support, $\eta \equiv 1$ on $[0, 1/3]$, $\eta \equiv 0$ on $[2/3, \infty)$ and set $\chi_r := \eta \circ \mathsf{d}(\cdot, B_r(x))$, where $\mathsf{d}(\cdot, B_r(x)) := \inf_{y \in B_r(x)} \mathsf{d}(\cdot, y)$, then it is easy to see that $\chi_r \equiv 1$ on $B_{r+1/3}(x)$, $\mathrm{supp}(\chi_r) \subset B_{r+2/3}(x)$ and is Lipschitz with $\mathrm{lip}(\chi_r) \leq |\eta'(\mathsf{d}(\cdot, B_r(x)))| \leq C$ $\mathfrak{m}$-a.e. with C independent of r. ∎

3 Auxiliary results

In this section we collect all technical results that are required in view of the proof of Theorem 4.1. We begin with a continuity statement for the heat semigroup.

Lemma 3.1. *With the same assumptions and notations as in Setting 2.1, let $f \in L^1 \cap L^\infty(\mathrm{X})$. Then $(\mathsf{h}_t f) \in C([0,\infty), L^p(\mathrm{X}))$ for all $p \in [1,\infty)$.*

In particular, for all $T > 0$, $p \in [1,\infty)$ and $\varepsilon > 0$ there exists a bounded open set B such that

$$\int_{\mathrm{X}\setminus B} |\mathsf{h}_t f|^p \,\mathrm{d}\mathfrak{m} < \varepsilon, \qquad \forall t \in [0,T]. \tag{3.1}$$

proof As a first step, recall that the heat flow satisfies the maximum principle and $(\mathsf{h}_t f) \in C([0,\infty), L^2(\mathrm{X}))$. Therefore, if $p \geq 2$ it is sufficient to observe that

$$\int |\mathsf{h}_s f - \mathsf{h}_t f|^p \mathrm{d}\mathfrak{m} \leq \|\mathsf{h}_s f - \mathsf{h}_t f\|^{p-2}_{L^\infty(\mathrm{X})} \int |\mathsf{h}_s f - \mathsf{h}_t f|^2 \mathrm{d}\mathfrak{m} \leq 2\|f\|^{p-2}_{L^\infty(\mathrm{X})} \|\mathsf{h}_s f - \mathsf{h}_t f\|^2_{L^2(\mathrm{X})}$$

for all $s,t \geq 0$ to deduce the L^p-continuity. For $p = 1$ we rely on Brezis-Lieb's lemma (see [8]); for our purposes it is enough to know that if $(u_n) \subset L^1(\mathrm{X})$ is a bounded sequence with $u_n \to u$ $\mathfrak{m}$-a.e. for some measurable function u, then $u \in L^1(\mathrm{X})$ and

$$\lim_{n\to\infty} \int \Big(|u_n| - |u_n - u|\Big) \mathrm{d}\mathfrak{m} = \int |u| \mathrm{d}\mathfrak{m}. \tag{3.2}$$

Hence pick any $t \geq 0$, any sequence (t_n) converging to it and set $u_n := \mathsf{h}_{t_n} f$. From $(\mathsf{h}_t f) \in C([0,\infty), L^2(\mathrm{X}))$ we know that, up to pass to a subsequence, $u_n \to u = \mathsf{h}_t f$ $\mathfrak{m}$-a.e. and from the mass-preserving property of the heat flow $\|u_n\|_{L^1(\mathrm{X})} = \|u\|_{L^1(\mathrm{X})}$. Therefore (3.2) yields $\|u_n - u\|_{L^1(\mathrm{X})} \to 0$ as $n \to \infty$ and by the arbitrariness of (t_n) we conclude that $(\mathsf{h}_t f) \in C([0,\infty), L^1(\mathrm{X}))$. For $p \in (1,2)$, it is sufficient to notice that, by interpolation, there exists C_p such that

$$\|\mathsf{h}_s f - \mathsf{h}_t f\|_{L^p(\mathrm{X})} \leq C_p \Big(\|\mathsf{h}_s f - \mathsf{h}_t f\|_{L^1(\mathrm{X})} + \|\mathsf{h}_s f - \mathsf{h}_t f\|_{L^2(\mathrm{X})}\Big), \qquad \forall s,t \geq 0.$$

As regards (3.1), we first observe that if we set $u_n := |\mathsf{h}_{t_n} f|^p$ and $u := |\mathsf{h}_t f|^p$ with $t_n \to t$, then the L^p-continuity of the heat flow implies that, up to subsequences, $u_n \to u$ $\mathfrak{m}$-a.e. and $\|u_n\|_{L^1(\mathrm{X})} \to \|u\|_{L^1(\mathrm{X})}$ as $n \to \infty$. Plugging this information into (3.2) we deduce that $(|\mathsf{h}_t f|^p) \in C([0,\infty), L^1(\mathrm{X}))$ and in turn this implies that $K_T := \{|\mathsf{h}_t f|^p \,:\, t \in [0,T]\} \subset L^1(\mathrm{X})$ is compact for all $T \geq 0$. Indeed, if $(u_n) \subset K_T$, then $u_n = |\mathsf{h}_{t_n} f|^p$ for some $t_n \in [0,T]$ and (t_n) admits a convergent subsequence, say $t_{n_k} \to t$, so that $u_{n_k} \to u := |\mathsf{h}_t f|^p \in K_T$ in $L^1(\mathrm{X})$. In particular K_T is weakly compact in $L^1(\mathrm{X})$ and, by the Dunford-Pettis theorem for σ-finite measures, this implies that for every $\varepsilon > 0$ there exists a measurable set B with finite measure such that

$$\int_{\mathrm{X}\setminus B} |\mathsf{h}_t f|^p \,\mathrm{d}\mathfrak{m} < \varepsilon, \qquad \forall t \in [0,T].$$

In order to see that B can be replaced by a bounded set, by the maximum principle and the fact that $f \in L^\infty(\mathrm{X})$ we know that, for some constant $C > 0$, $|\mathsf{h}_t f|^p \leq C$ for all $t \geq 0$ and, since $\mathfrak{m}(B) < \infty$, for all $x \in \mathrm{X}$ there exists $R > 0$ such that $\mathfrak{m}(B \setminus B_R(x)) < \varepsilon/C$. Therefore, fixing $x \in \mathrm{X}$ and setting $B' := B \cap B_R(x)$ we see that

$$\int_{\mathrm{X}\cap B'} |\mathsf{h}_t f|^p \,\mathrm{d}\mathfrak{m} = \int_{\mathrm{X}\setminus B} |\mathsf{h}_t f|^p \,\mathrm{d}\mathfrak{m} + \int_{B\setminus B_R(x)} |\mathsf{h}_t f|^p \,\mathrm{d}\mathfrak{m} < 2\varepsilon,$$

whence the conclusion by the arbitrariness of ε. $\square$

Then we show that the dimension-free Harnack inequality implies the local logarithmic Sobolev inequality, which is the non-smooth counterpart of the main result contained in [6]. On smooth Riemannian manifolds [27] and in the setting of Bakry-Émery Γ-calculus [7], both (2.1) and (3.4) are known to be equivalent to the curvature-dimension condition $\mathsf{CD}(K,\infty)$ ([25], see also Section 4 for the definition), but in the present framework the implication (2.1) $\Rightarrow$ (3.4) was still missing. As a byproduct we improve the strong Feller property of h_t to an L^∞-LIP regularization.

Proposition 3.2. *With the same assumptions and notations as in Setting 2.1,* h_t *maps* $L^\infty(\mathrm{X})$ *into* $\mathrm{LIP}(\mathrm{X})$ *and for any* $f \in L^\infty$ *positive,* $t>0$ *it holds*

$$\mathrm{lip}(\mathsf{h}_t f)^2 \le \frac{2K}{e^{2Kt}-1}\mathsf{h}_t f\big(\mathsf{h}_t(f\log f) - \mathsf{h}_t f\log \mathsf{h}_t f\big) \qquad \textit{pointwise.} \tag{3.3}$$

In addition, for any $f \in L^p(\mathrm{X})$ *positive,* $p\in(1,\infty)$*, and for any* $t>0$ *it holds*

$$|D\mathsf{h}_t f|^2 \le \frac{2K}{e^{2Kt}-1}\mathsf{h}_t f\big(\mathsf{h}_t(f\log f) - \mathsf{h}_t f\log \mathsf{h}_t f\big) \qquad \mathfrak{m}\text{-}a.e. \tag{3.4}$$

proof Let $x\in\mathrm{X}$, $y_n\to x$ and for sake of brevity set $\mathsf{d}_n := \mathsf{d}(x,y_n)$; let us also fix $\delta>0$ and take $f\in L^\infty(\mathrm{X})$ positive (not necessarily in $L^p(\mathrm{X})$). Since the function $\Phi:(0,\infty)\times[0,\infty)\to\mathbb{R}$ defined by

$$\Phi(z,\alpha) := \begin{cases} \dfrac{z^\alpha - 1}{\alpha} & \text{if } \alpha>0 \\ \log z & \text{if } \alpha = 0 \end{cases}$$

is continuous, by the weak Feller property of h_t the function $y\mapsto\Phi(\mathsf{h}_t f(y),\delta\mathsf{d}(x,y))$ is continuous as well, whence

$$\limsup_{n\to\infty}\frac{\mathsf{h}_t f(y_n) - \mathsf{h}_t f(x)}{\mathsf{d}_n} + \delta(\mathsf{h}_t f\log\mathsf{h}_t f)(x) = \limsup_{n\to\infty}\frac{(\mathsf{h}_t f)^{1+\delta\mathsf{d}_n}(y_n) - \mathsf{h}_t f(x)}{\mathsf{d}_n}. \tag{3.5}$$

Moreover by (2.1) we have

$$\begin{aligned}(\mathsf{h}_t f)^{1+\delta\mathsf{d}_n}(y_n) - \mathsf{h}_t f(x) &\le \mathsf{h}_t|f|^{1+\delta\mathsf{d}_n}(x)\exp\Big(\frac{K(1+\delta\mathsf{d}_n)\mathsf{d}_n}{2\delta(e^{2Kt}-1)}\Big) - \mathsf{h}_t f(x) \\ &= \exp\Big(\frac{K(1+\delta\mathsf{d}_n)\mathsf{d}_n}{2\delta(e^{2Kt}-1)}\Big)\big(\mathsf{h}_t|f|^{1+\delta\mathsf{d}_n}(x) - \mathsf{h}_t f(x)\big) + \mathsf{h}_t f(x)\Big(\exp\Big(\frac{K(1+\delta\mathsf{d}_n)\mathsf{d}_n}{2\delta(e^{2Kt}-1)}\Big) - 1\Big)\end{aligned}$$

and on the one hand

$$\begin{aligned}\limsup_{n\to\infty}\exp\Big(\frac{K(1+\delta\mathsf{d}_n)\mathsf{d}_n}{2\delta(e^{2Kt}-1)}\Big)\frac{\mathsf{h}_t|f|^{1+\delta\mathsf{d}_n}(x) - \mathsf{h}_t f(x)}{\mathsf{d}_n} &= \limsup_{n\to\infty}\int\frac{|f|^{1+\delta\mathsf{d}_n} - f}{\mathsf{d}_n}\mathsf{r}_t[x]\mathrm{d}\mathfrak{m} \\ &\le \delta\int f\log f\,\mathsf{r}_t[x]\mathrm{d}\mathfrak{m} = \delta\mathsf{h}_t(f\log f)(x),\end{aligned}$$

where the inequality is motivated by Fatou's lemma (indeed $(f\Phi(f,\delta\mathsf{d}_n))_n$ is bounded from above, since so is f), while on the other hand

$$\limsup_{n\to\infty}\frac{\mathsf{h}_t f(x)}{\mathsf{d}_n}\Big(\exp\Big(\frac{K(1+\delta\mathsf{d}_n)\mathsf{d}_n}{2\delta(e^{2Kt}-1)}\Big) - 1\Big) = \frac{K}{2\delta(e^{2Kt}-1)}\mathsf{h}_t f(x).$$

Plugging these observations into (3.5) yields

$$\limsup_{n\to\infty}\frac{\mathsf{h}_t f(y_n)-\mathsf{h}_t f(x)}{\mathsf{d}_n}\le\delta\mathsf{h}_t(f\log f)(x)-\delta(\mathsf{h}_t f\log\mathsf{h}_t f)(x)+\frac{K}{2\delta(e^{2Kt}-1)}\mathsf{h}_t f(x). \tag{3.6}$$

Next, assume that there exists $c>0$ such that $f\ge c$; since by smoothness $\Phi(z,\alpha)$ converges uniformly to $\log z$ on $[z_0,z_1]$ with $z_0>0$ as $\alpha\downarrow 0$, this implies that $f\Phi(f,\delta\mathsf{d}_n)$ converges uniformly to $f\log f$ as $n\to\infty$. As on the other hand $\mathsf{r}[y_n]\rightharpoonup\mathsf{r}[x]$ in $L^1(\mathrm{X})$ by the strong Feller property of h_t, this yields

$$\lim_{n\to\infty}\frac{\mathsf{h}_t f(y_n)-\mathsf{h}_t|f|^{1+\delta\mathsf{d}_n}(y_n)}{\mathsf{d}_n}=-\delta\mathsf{h}_t(f\log f)(x),$$

whence

$$\limsup_{n\to\infty}\frac{\mathsf{h}_t f(x)-\mathsf{h}_t f(y_n)}{\mathsf{d}_n}-\delta\mathsf{h}_t(f\log f)(x)=\limsup_{n\to\infty}\frac{\mathsf{h}_t f(x)-\mathsf{h}_t|f|^{1+\delta\mathsf{d}_n}(y_n)}{\mathsf{d}_n}.$$

Using again (2.1) we obtain

$$\begin{aligned}\mathsf{h}_t f(x)-\mathsf{h}_t|f|^{1+\delta\mathsf{d}_n}(y_n)&\le\mathsf{h}_t f(x)-(\mathsf{h}_t f)^{1+\delta\mathsf{d}_n}\exp\Big(-\frac{K(1+\delta\mathsf{d}_n)\mathsf{d}_n}{2\delta(e^{2Kt}-1)}\Big)\\ \mathsf{h}_t f(x)\Big(1-\exp\Big(-\frac{K(1+\delta\mathsf{d}_n)\mathsf{d}_n}{2\delta(e^{2Kt}-1)}\Big)\Big)&+\exp\Big(-\frac{K(1+\delta\mathsf{d}_n)\mathsf{d}_n}{2\delta(e^{2Kt}-1)}\Big)\big(\mathsf{h}_t f(x)-\mathsf{h}_t|f|^{1+\delta\mathsf{d}_n}(x)\big)\end{aligned}$$

and on the one hand

$$\limsup_{n\to\infty}\frac{\mathsf{h}_t f(x)}{\mathsf{d}_n}\Big(1-\exp\Big(-\frac{K(1+\delta\mathsf{d}_n)\mathsf{d}_n}{2\delta(e^{2Kt}-1)}\Big)\Big)=\frac{K}{2\delta(e^{2Kt}-1)}\mathsf{h}_t f(x)$$

while on the other hand

$$\limsup_{n\to\infty}\exp\Big(-\frac{K(1+\delta\mathsf{d}_n)\mathsf{d}_n}{2\delta(e^{2Kt}-1)}\Big)\frac{\mathsf{h}_t f(x)-\mathsf{h}_t|f|^{1+\delta\mathsf{d}_n}(x)}{\mathsf{d}_n}=-\delta(\mathsf{h}_t f\log\mathsf{h}_t f)(x),$$

so that

$$\limsup_{n\to\infty}\frac{\mathsf{h}_t f(x)-\mathsf{h}_t f(y_n)}{\mathsf{d}_n}\le\delta\mathsf{h}_t(f\log f)(x)+\frac{K}{2\delta(e^{2Kt}-1)}\mathsf{h}_t f(x)-\delta(\mathsf{h}_t f\log\mathsf{h}_t f)(x). \tag{3.7}$$

The assumption $f\ge c$ for some $c>0$ can now be removed, since if we define $f_k:=f+1/k$ then (3.7) holds for f_k and passing to the limit as $k\to\infty$ yields the validity of (3.7) for f: indeed, the left-hand side is easily seen to be constant w.r.t. k, while on the right-hand side it is sufficient to use the representation formula (2.3) to deduce that $\mathsf{h}_t f_k\to\mathsf{h}_t f$ and $\mathsf{h}_t(f_k\log f_k)\to\mathsf{h}_t(f\log f)$ pointwise.

Combining (3.7) with (3.6) and recalling the definition of $\mathrm{lip}(f)$ imply

$$\mathrm{lip}(\mathsf{h}_t f)\le\delta\mathsf{h}_t(f\log f)-\delta(\mathsf{h}_t f\log\mathsf{h}_t f)+\frac{K}{2\delta(e^{2Kt}-1)}\mathsf{h}_t f$$

pointwise, as the right-hand side is continuous by the strong Feller property of h_t, and it is now sufficient to optimize the right-hand side w.r.t. δ to get (3.3) and the L^∞-LIP regularization for h_t. In order to prove (3.4) it is sufficient to consider (3.3) for $f\in\mathrm{LIP}(\mathrm{X})$ positive with bounded support, recall (2.5) and use the density of Lipschitz functions with bounded support in $L^p(\mathrm{X})$, $p\in(1,\infty)$, together with the lower semicontinuity of the minimal weak upper gradient. $\square$

Finally we prove the integral maximum principle for the heat semigroup. On Riemannian manifolds several different proofs are possible (see for instance [16] and references therein); here we adapt to the metric measure framework the one proposed in [14].

Lemma 3.3 (Integral maximum principle)**.** *With the same assumptions and notations as in Setting 2.1, let $T > 0, p \in (1,\infty), x \in \mathrm{X}$ and set*

$$\xi_t(y) := -\frac{\mathsf{d}^2(x,y)}{2(T-qt)}, \qquad \text{for } y \in \mathrm{X}, t \in [0, T/q),$$

where $q := \frac{p}{2(p-1)}$. Then, for any non-negative $f \in L^p(\mathrm{X})$ and $t \in [0,T/q)$, it holds

$$\int (\mathsf{h}_t f)^p e^{\xi_t} \mathrm{d}\mathfrak{m} \leq \int f^p \exp\Big(-\frac{\mathsf{d}^2(x,\cdot)}{2T}\Big)\mathrm{d}\mathfrak{m}.$$

proof As a preliminary remark, by standard approximation argument it is not restrictive to assume that $f \in L^\infty(\mathrm{X})$ with bounded support: by Lemma 3.1 this implies that $(\mathsf{h}_t f) \in C([0,\infty), L^p(\mathrm{X}))$. Hence for all $t \geq 0$ and any sequence (t_n) converging to t there exists $g \in L^p(\mathrm{X})$ such that, up to pass to a subsequence, $\mathsf{h}_{t_n} f \leq g$ for all $n \in \mathbb{N}$. As $e^{\xi_t} \leq 1$ and ξ_t smoothly depends on $t \in [0, T/q)$, by dominated convergence we deduce that

$$I(t) := \int (\mathsf{h}_t f)^p e^{\xi_t} \mathrm{d}\mathfrak{m}$$

is continuous on $[0, T/q)$. In order to see that it is also locally absolutely continuous on $(0, T/q)$, let $R > 0$ and take a cut-off function χ_R as in Remark 2.2. Notice that $(\mathsf{h}_t f) \in AC_{loc}((0,\infty), L^2(\mathrm{X}))$ and, by the maximum principle, it is uniformly bounded in space and time. As ξ_t smoothly depends on $t \in [0,T/q)$ and $\mathsf{d}^2(x,\cdot) \in W^{1,2}_{loc}(\mathrm{X})$, we deduce that $(0,T/q) \ni t \mapsto \chi_R(\mathsf{h}_t f)^p e^{\xi_t} \in L^2(\mathrm{X})$ is absolutely continuous. In particular, so is $\int \chi_R(\mathsf{h}_t f)^p e^{\xi_t} \mathrm{d}\mathfrak{m}$ and it is then clear that

$$\frac{\mathrm{d}}{\mathrm{d}t} \int \chi_R(\mathsf{h}_t f)^p e^{\xi_t} \mathrm{d}\mathfrak{m} = \int \chi_R\Big(p e^{\xi_t}(\mathsf{h}_t f)^{p-1}\Delta \mathsf{h}_t f - \frac{q\mathsf{d}_x^2}{2(T-qt)^2} e^{\xi_t}(\mathsf{h}_t f)^p\Big)\mathrm{d}\mathfrak{m}, \qquad \text{a.e. } t$$

where $\mathsf{d}_x := \mathsf{d}(x,\cdot)$. As regards the first summand on the right-hand side, using integration by parts and the fact that $|\nabla \mathsf{d}_x| \leq 1$ $\mathfrak{m}$-a.e. we can rewrite it as

$$\begin{aligned}
\int \chi_R e^{\xi_t}(\mathsf{h}_t f)^{p-1}\Delta \mathsf{h}_t f \mathrm{d}\mathfrak{m} &= -\int \langle \nabla(\chi_R e^{\xi_t}(\mathsf{h}_t f)^{p-1}), \nabla \mathsf{h}_t f\rangle \mathrm{d}\mathfrak{m} \\
&= -\int \chi_R\Big((p-1)(\mathsf{h}_t f)^{p-2}|\nabla \mathsf{h}_t f|^2 + (\mathsf{h}_t f)^{p-1}\langle \nabla \xi_t, \nabla \mathsf{h}_t f\rangle\Big)e^{\xi_t}\mathrm{d}\mathfrak{m} \\
&\quad - \int e^{\xi_t}(\mathsf{h}_t f)^{p-1}\langle \nabla \chi_R, \nabla \mathsf{h}_t f\rangle \mathrm{d}\mathfrak{m} \\
&\leq \int \chi_R\Big(-(p-1)(\mathsf{h}_t f)^{p-2}|\nabla \mathsf{h}_t f|^2 + \frac{\mathsf{d}_x}{T-qt}(\mathsf{h}_t f)^{p-1}|\nabla \mathsf{h}_t f|\Big)e^{\xi_t}\mathrm{d}\mathfrak{m} \\
&\quad - \int (\mathsf{h}_t f)^{p-1}\langle \nabla \chi_R, \nabla \mathsf{h}_t f\rangle e^{\xi_t}\mathrm{d}\mathfrak{m}
\end{aligned}$$

so that

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\int \chi_R(\mathsf{h}_t f)^p e^{\xi_t}\mathrm{d}\mathfrak{m} &\leq -p\int \Big((\mathsf{h}_t f)^{p-1}\langle \nabla\chi_R, \nabla \mathsf{h}_t f\rangle + (p-1)\chi_R(\mathsf{h}_t f)^{p-2}|\nabla \mathsf{h}_t f|^2\Big)e^{\xi_t}\mathrm{d}\mathfrak{m} \\
&\quad + \int \chi_R\Big(\frac{p\mathsf{d}_x}{T-qt}(\mathsf{h}_t f)^{p-1}|\nabla \mathsf{h}_t f| - \frac{q\mathsf{d}_x^2}{2(T-qt)^2}(\mathsf{h}_t f)^p\Big)e^{\xi_t}\mathrm{d}\mathfrak{m}.
\end{aligned}$$

The fact that $\mathsf{h}_t f \in L^p(\mathrm{X})$ for all $t \geq 0$, $0 \leq \chi_R \leq 1$ and $\chi_R \to 1$ $\mathfrak{m}$-a.e. as $R \to \infty$ are sufficient to deduce by dominated convergence

$$\lim_{R\to\infty} \int \chi_R (\mathsf{h}_t f)^p e^{\xi_t} \mathrm{d}\mathfrak{m} = \int (\mathsf{h}_t f)^p e^{\xi_t} \mathrm{d}\mathfrak{m}, \qquad \forall t \in (0, T/q).$$

Moreover, since $\mathsf{d}_x^2 e^{\xi_t} \in L^\infty(\mathrm{X})$ uniformly in $t \in [0, T/q)$, from (3.1) we see that for all $\varepsilon > 0$ there exists R sufficiently large such that

$$\int (1 - \chi_R)(\mathsf{h}_t f)^p \mathsf{d}_x^2 e^{\xi_t} \mathrm{d}\mathfrak{m} < \varepsilon, \qquad \forall t \in [0, T/q)$$

whence

$$\lim_{R\to\infty} \int \frac{\chi_R \mathsf{d}_x^2}{(T-qt)^2} (\mathsf{h}_t f)^p e^{\xi_t} \mathrm{d}\mathfrak{m} = \int \frac{\mathsf{d}_x^2}{(T-qt)^2} (\mathsf{h}_t f)^p e^{\xi_t} \mathrm{d}\mathfrak{m} \qquad \text{loc. uniformly in } t \in [0, T/q)$$

by the arbitrariness of ε. Then (3.4) entails that for any $\mathcal{C} \subset (0, T/q)$ compact there exists $C > 0$ such that $|\nabla \mathsf{h}_t f|^2 \leq C \mathsf{h}_t f (\mathsf{h}_t (f \log f) - \mathsf{h}_t f \log \mathsf{h}_t f)$ for all $t \in \mathcal{C}$, whence

$$\begin{aligned}(\mathsf{h}_t f)^{p-2} |\nabla \mathsf{h}_t f|^2 e^{\xi_t} &\leq C (\mathsf{h}_t f)^{p-1} \mathsf{h}_t (f \log f) e^{\xi_t} + C (\mathsf{h}_t f)^p |\log \mathsf{h}_t f| e^{\xi_t} \\ & C (\mathsf{h}_t f)^{p-1} \mathsf{h}_t (f \log f) e^{\xi_t} + C (\mathsf{h}_t f)^{\frac{p+1}{2}} (\mathsf{h}_t f)^{\frac{p-1}{2}} |\log \mathsf{h}_t f| e^{\xi_t} =: g_t\end{aligned}$$

On the one hand $(\mathsf{h}_t f)^{p-1} e^{\xi_t}, (\mathsf{h}_t f)^{(p-1)/2} |\log \mathsf{h}_t f| e^{\xi_t} \in L^\infty(\mathrm{X})$ uniformly in $t \in \mathcal{C}$, since $z \mapsto z^\alpha |\log z|$ is bounded on $[0, \|f\|_{L^\infty(\mathrm{X})}]$ for all $\alpha > 0$ and $(p-1)/2 > 0$; on the other hand (3.1) applies to $\mathsf{h}_t(f \log f)$ and $(\mathsf{h}_t f)^{(p+1)/2}$, since for the former the fact that $f \in L^\infty(\mathrm{X})$ with bounded support entails $f \log f \in L^1 \cap L^\infty(\mathrm{X})$ while for the latter $(p+1)/2 \geq 1$. Thus, arguing as above, by (3.1) we have that for all $\varepsilon > 0$ there exists R large enough so that

$$\begin{aligned}\int (1 - \chi_R)(\mathsf{h}_t f)^{p-2} |\nabla \mathsf{h}_t f|^2 e^{\xi_t} \mathrm{d}\mathfrak{m} &= \int_{\mathrm{X} \setminus B_R(x)} (1 - \chi_R)(\mathsf{h}_t f)^{p-2} |\nabla \mathsf{h}_t f|^2 e^{\xi_t} \mathrm{d}\mathfrak{m} \\ &\leq \int_{\mathrm{X} \setminus B_R(x)} g_t \, \mathrm{d}\mathfrak{m} < \varepsilon, \qquad \forall t \in [0, T/q),\end{aligned}$$

whence

$$\lim_{R\to\infty} \int \chi_R (\mathsf{h}_t f)^{p-2} |\nabla \mathsf{h}_t f|^2 e^{\xi_t} \mathrm{d}\mathfrak{m} = \int (\mathsf{h}_t f)^{p-2} |\nabla \mathsf{h}_t f|^2 e^{\xi_t} \mathrm{d}\mathfrak{m} \qquad \text{loc. uniformly in } t \in (0, T/q).$$

Finally, since $f \log f \leq f \log(f+2)$ and $f \log(f+2) > 0$ as $f > 0$, by (3.4) we have that for any $\mathcal{C} \subset (0, T/q)$ compact there exists $C > 0$ such that

$$\begin{aligned}(\mathsf{h}_t f)^{p-1} |\nabla \mathsf{h}_t f| &\leq C (\mathsf{h}_t f)^{p-1/2} \sqrt{\mathsf{h}_t (f \log(f+2)) - \mathsf{h}_t f \log \mathsf{h}_t f} \\ &\leq C \frac{(\mathsf{h}_t f)^{p-1/2}}{\sqrt{\mathsf{h}_t (f \log(f+2))}} \mathsf{h}_t (f \log(f+2)) + C (\mathsf{h}_t f)^p |\log \mathsf{h}_t f| \\ &\leq \frac{C}{\log 2} (\mathsf{h}_t f)^{p-1} \mathsf{h}_t (f \log(f+2)) + C (\mathsf{h}_t f)^p |\log \mathsf{h}_t f|\end{aligned}$$

for all $t \in \mathcal{C}$, so that arguing as above, using the fact that $\mathsf{d}_x e^{\xi_t} \in L^\infty(\mathrm{X})$ uniformly in $t \in [0, T/q)$ and $|\nabla \chi_R| \in L^\infty(\mathrm{X})$ uniformly in $R > 0$ with $\chi_R, |\nabla \chi_R|$ converging $\mathfrak{m}$-a.e. to 0,1 respectively, we see that

$$\lim_{R\to\infty}\int (\mathsf{h}_t f)^{p-1}\langle\nabla\chi_R,\nabla\mathsf{h}_t f\rangle e^{\xi_t}\mathrm{d}\mathfrak{m}=0,$$
$$\lim_{R\to\infty}\int \chi_R\frac{p\mathsf{d}_x}{T-qt}(\mathsf{h}_t f)^{p-1}|\nabla\mathsf{h}_t f|e^{\xi_t}\mathrm{d}\mathfrak{m}=\int \frac{p\mathsf{d}_x}{T-qt}(\mathsf{h}_t f)^{p-1}|\nabla\mathsf{h}_t f|e^{\xi_t}\mathrm{d}\mathfrak{m},$$

locally uniformly in $t\in(0,T/q)$. We thus obtain that $I\in AC_{loc}((0,T/q))$ with

$$\begin{aligned}I'(t)&\leq\int\Big(-p(p-1)(\mathsf{h}_t f)^{p-2}|\nabla\mathsf{h}_t f|^2+\frac{p\mathsf{d}_x}{T-qt}(\mathsf{h}_t f)^{p-1}|\nabla\mathsf{h}_t f|-\frac{q\mathsf{d}_x^2}{2(T-qt)^2}(\mathsf{h}_t f)^p\Big)e^{\xi_t}\mathrm{d}\mathfrak{m}\\&=-p(p-1)\int(\mathsf{h}_t f)^p\Big(\frac{|\nabla\mathsf{h}_t f|}{\mathsf{h}_t f}-\frac{\mathsf{d}_x}{2(p-1)(T-t)}\Big)^2e^{\xi_t}\mathrm{d}\mathfrak{m}\leq 0\end{aligned}$$

and since I is continuous up to $t=0$, this yields the conclusion. □

4 Main result and applications

We are now in the position to prove the upper Gaussian estimate for the heat kernel, the proof being partially inspired by [15] and [14].

Theorem 4.1 (Gaussian upper bound)**.** *With the same assumptions and notations as in Setting 2.1, for any $\varepsilon>0$ there exist constants $C_\varepsilon>0$ and $C_K\geq 0$ such that, for every $x,y\in\mathrm{X}$ and $t>0$, it holds*

$$\mathsf{r}_t[x](y)\leq\frac{1}{\sqrt{\mathfrak{m}(B_{\sqrt{t}}(x))\mathfrak{m}(B_{\sqrt{t}}(y))}}\exp\Big(C_\varepsilon(1+C_Kt)-\frac{\mathsf{d}^2(x,y)}{(4+\varepsilon)t}\Big).\tag{4.1}$$

If $K\geq 0$, then C_K can be chosen equal to 0.

proof As a first step, for any $x\in\mathrm{X}$ and $t,D>0$ define

$$E_D(x,t):=\int(\mathsf{r}_t[x](y))^2\exp\Big(\frac{\mathsf{d}^2(x,y)}{Dt}\Big)\mathrm{d}\mathfrak{m}(y)\in[0,\infty]$$

and observe that by (2.4) and the triangle inequality

$$\mathsf{r}_t[x](y)=\exp\Big(-\frac{\mathsf{d}^2(x,y)}{2Dt}\Big)\int\mathsf{r}_{t/2}[x](z)\exp\Big(\frac{\mathsf{d}^2(x,z)}{Dt}\Big)\mathsf{r}_{t/2}[z](y)\exp\Big(\frac{\mathsf{d}^2(z,y)}{Dt}\Big)\mathrm{d}\mathfrak{m}(z),$$

whence by the Cauchy-Schwarz inequality

$$\mathsf{r}_t[x](y)\leq\sqrt{E_D(x,t/2)E_D(y,t/2)}\exp\Big(-\frac{\mathsf{d}^2(x,y)}{2Dt}\Big).\tag{4.2}$$

In order to estimate $E_D(x,t/2)$ and $E_D(y,t/2)$, let $f\in L^1\cap L^\infty(\mathrm{X})$ be non-negative, $T>0$ and $p\in(1,2)$, consider the Harnack inequality (2.1) with exponent $2/p$, raise it to the power p, multiply both sides by e^{ξ_t} with ξ_t defined as in Lemma 3.3 and integrate w.r.t. $\mathfrak{m}$ in the y variable to get

$$\int(\mathsf{h}_t f(x))^2e^{\xi_t(y)}\exp\Big(-\frac{pK\mathsf{d}^2(x,y)}{(2-p)(e^{2Kt}-1)}\Big)\mathrm{d}\mathfrak{m}(y)\leq\int(\mathsf{h}_t(f^{2/p})(y))^pe^{\xi_t(y)}\mathrm{d}\mathfrak{m}(y).$$

On the one hand, as $f^{2/p} \in L^p(\mathrm{X})$ Lemma 3.3 can be applied, whence

$$\int (\mathsf{h}_t(f^{2/p}))^p e^{\xi_t} d\mathfrak{m} \leq \int f^2 \exp\Big(-\frac{\mathsf{d}^2(x,\cdot)}{2T}\Big) d\mathfrak{m}, \qquad \forall t \in [0,T/q)$$

with q also defined as in Lemma 3.3. On the other hand

$$\begin{aligned}
&\int (\mathsf{h}_t f(x))^2 e^{\xi_t(y)} \exp\Big(-\frac{pK\mathsf{d}^2(x,y)}{(2-p)(e^{2Kt}-1)}\Big) d\mathfrak{m}(y) \\
&\qquad \geq (\mathsf{h}_t f(x))^2 \int_{B_{\sqrt{2t}}(x)} e^{\xi_t(y)} \exp\Big(-\frac{pK\mathsf{d}^2(x,y)}{(2-p)(e^{2Kt}-1)}\Big) d\mathfrak{m}(y) \\
&\qquad \geq (\mathsf{h}_t f(x))^2 \exp\Big(-\frac{t}{T-qt} - \frac{2pKt}{(2-p)(e^{2Kt}-1)}\Big) \mathfrak{m}(B_{\sqrt{2t}}(x)).
\end{aligned}$$

By combining these inequalities we obtain

$$(\mathsf{h}_t f(x))^2 \leq \frac{1}{\mathfrak{m}(B_{\sqrt{2t}}(x))} \exp\Big(\frac{t}{T-qt} + \frac{2pKt}{(2-p)(e^{2Kt}-1)}\Big) \int f^2 \exp\Big(-\frac{\mathsf{d}^2(x,\cdot)}{2T}\Big) d\mathfrak{m} \tag{4.3}$$

for all $t \in (0,T/q)$ and in particular this is true when f is equal to

$$f_n := (n \wedge \mathsf{r}_t[x]) \exp\Big(n \wedge \frac{\mathsf{d}^2(x,\cdot)}{2T}\Big),$$

where $a \wedge b := \min\{a,b\}$. By the representation formula (2.3) and by the monotone convergence theorem, as the f_n's are monotonically increasing, we deduce that (4.3) actually holds with $f = \mathsf{r}_t[x] e^{\mathsf{d}^2(x,\cdot)/2T}$ and this is equivalent to

$$\int (\mathsf{r}_t[x])^2 \exp\Big(\frac{\mathsf{d}^2(x,\cdot)}{2T}\Big) d\mathfrak{m} \leq \frac{1}{\mathfrak{m}(B_{\sqrt{2t}}(x))} \exp\Big(\frac{t}{T-qt} + \frac{2pKt}{(2-p)(e^{2Kt}-1)}\Big) \tag{4.4}$$

for all $t \in (0,T/q)$. Recalling that $q := \frac{p}{2(p-1)}$ and $p \in (1,2)$ is arbitrary, we observe that for all $D > 2$ there exists p' sufficiently close to 2 such that $D > 2q' > 2$ with $q' = \frac{p'}{2(p'-1)}$. Hence, since also T is arbitrary, if we pick any $t > 0$ and set $T := Dt/2$, in such a way that $t = 2T/D < T/q'$, we deduce that (4.4) holds and

$$\frac{t}{T-q't} = \frac{2}{D-2q'} =: C_D,$$

so that we have just shown that for all $D > 2$ there exists $p' \in (1,2)$ and $C_D > 0$ both depending only on D such that

$$\int (\mathsf{r}_t[x])^2 \exp\Big(\frac{\mathsf{d}^2(x,\cdot)}{Dt}\Big) d\mathfrak{m} \leq \frac{1}{\mathfrak{m}(B_{\sqrt{2t}}(x))} \exp\Big(C_D + \frac{2p'Kt}{(2-p')(e^{2Kt}-1)}\Big), \qquad \forall t > 0.$$

Now observe that

$$\frac{2p'Kt}{(2-p')(e^{2Kt}-1)} \leq C'_D(1+C_K t), \qquad \forall t > 0$$

for suitable constants C'_D, C_K and it is easily seen that C_K can be chosen equal to 0 when $K \geq 0$. Plugging this inequality into the previous one and recalling the definition of $E_D(x,t)$ it follows that for any $D > 2$

$$E_D(x,t) \leq \frac{1}{\mathfrak{m}(B_{\sqrt{2t}}(x))} \exp\Big(C''_D(1+C_K t)\Big), \qquad \forall t > 0$$

for some $C''_D > 0$ depending only on D and combining this inequality with (4.2) the bound (4.1) follows. $\square$

Remark 4.2. In the case of smooth Riemannian manifolds the Varadhan asymptotic formula for short-time behaviour of the heat kernel states that

$$\lim_{t\downarrow 0} t\log \mathsf{r}_t[x](y) = -\frac{\mathsf{d}^2(x,y)}{4}, \qquad x\neq y.$$

Hence (4.1) is sharp for short time. ■

In the same spirit of the remark above we can formulate the following

Corollary 4.3. *With the same assumptions and notations as in Setting 2.1, if we further suppose that for any $x\in \mathrm{X}$ there exist constants $R_x, c_x > 0$, $0<\alpha_x<2$ such that*

$$\mathfrak{m}(B_r(x)) \geq c_x \exp(-1/r^{\alpha_x}), \qquad r<R_x,$$

then the upper Varadhan estimate holds

$$\lim_{t\downarrow 0} t\log \mathsf{r}_t[x](y) \leq -\frac{\mathsf{d}^2(x,y)}{4}$$

for all $x,y\in\mathrm{X}$, $x\neq y$.

As a direct consequence of Proposition 3.2 and of our main theorem we obtain the local logarithmic Sobolev inequality and an upper Gaussian estimate for the heat kernel in $\mathsf{RCD}(K,\infty)$ spaces, which to date were both still missing. For reader's sake, let us recall that $(\mathrm{X},\mathsf{d},\mathfrak{m})$ satisfies the $\mathsf{RCD}(K,\infty)$ condition (see [5]) if it is infinitesimally Hilbertian and the relative entropy functional $\mathrm{Ent}_{\mathfrak{m}} : \mathscr{P}(\mathrm{X}) \to \mathbb{R}\cup\{+\infty\}$ defined as

$$\mathrm{Ent}_{\mathfrak{m}}(\mu) := \begin{cases} \displaystyle\int \rho\log(\rho)\,\mathrm{d}\mathfrak{m} & \text{if } \mu=\rho\mathfrak{m} \text{ and } (\rho\log(\rho))^-\in L^1(\mathrm{X}) \\ +\infty & \text{otherwise} \end{cases}$$

is K-convex in $(\mathscr{P}_2(\mathrm{X}), W_2)$, namely for every $\mu,\nu\in\mathscr{P}_2(\mathrm{X})$ with $\mathrm{Ent}_{\mathfrak{m}}(\mu), \mathrm{Ent}_{\mathfrak{m}}(\nu)<\infty$ there exists a W_2-geodesic (μ_t) with $\mu_0=\mu$, $\mu_1=\nu$ and such that

$$\mathrm{Ent}_{\mathfrak{m}}(\mu_t) \leq (1-t)\mathrm{Ent}_{\mathfrak{m}}(\mu) + t\mathrm{Ent}_{\mathfrak{m}}(\nu) - \frac{K}{2}t(1-t)W_2^2(\mu,\nu), \qquad \forall t\in[0,1].$$

In this class of spaces, H. Li proved that the dimension-free Harnack inequality (2.1) holds [20], so that we can state the following

Corollary 4.4. *Let $(\mathrm{X},\mathsf{d},\mathfrak{m})$ be an $\mathsf{RCD}(K,\infty)$ space with $K\in\mathbb{R}$. Then*

(i) for any $f\in L^p(\mathrm{X})$ positive, $p\in(1,\infty)$, and for any $t>0$ it holds

$$\frac{e^{2Kt}-1}{2K}\frac{|\nabla \mathsf{h}_t f|^2}{\mathsf{h}_t f} \leq \mathsf{h}_t(f\log f) - \mathsf{h}_t f\log \mathsf{h}_t f \qquad \mathfrak{m}\text{-}a.e.$$

(ii) for any $\varepsilon>0$ there exist constants $C_\varepsilon>0$ and $C_K\geq 0$ such that, for every $x,y\in\mathrm{X}$ and $t>0$, it holds

$$\mathsf{r}_t[x](y) \leq \frac{1}{\sqrt{\mathfrak{m}(B_{\sqrt{t}}(x))\mathfrak{m}(B_{\sqrt{t}}(y))}}\exp\Big(C_\varepsilon(1+C_K t) - \frac{\mathsf{d}^2(x,y)}{(4+\varepsilon)t}\Big).$$

If $K\geq 0$, then C_K can be chosen equal to 0.

References

[1] L. Ambrosio, N. Gigli, A. Mondino, and T. Rajala, *Riemannian Ricci curvature lower bounds in metric measure spaces with σ-finite measure*, Trans. Amer. Math. Soc., 367 (2015), pp. 4661–4701.

[2] L. Ambrosio, N. Gigli, and G. Savaré, *Gradient flows in metric spaces and in the space of probability measures*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, second ed., 2008.

[3] ——, *Density of Lipschitz functions and equivalence of weak gradients in metric measure spaces*, Rev. Mat. Iberoam., 29 (2013), pp. 969–996.

[4] ——, *Calculus and heat flow in metric measure spaces and applications to spaces with Ricci bounds from below*, Invent. Math., 195 (2014), pp. 289–391.

[5] ——, *Metric measure spaces with Riemannian Ricci curvature bounded from below*, Duke Math. J., 163 (2014), pp. 1405–1490.

[6] M. Arnaudon, A. Thalmaier, and F.-Y. Wang, *Gradient estimates and Harnack inequalities on non-compact Riemannian manifolds*, Stochastic Processes and their Applications, 119 (2009), pp. 3653–3670.

[7] D. Bakry, I. Gentil, and M. Ledoux, *Analysis and geometry of Markov diffusion operators*, vol. 348 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Springer, Cham, 2014.

[8] H. Brézis and E. Lieb, *A relation between pointwise convergence of functions and convergence of functionals*, Proceedings of the American Mathematical Society, 88 (1983), pp. 486–490.

[9] J. Cheeger, *Differentiability of Lipschitz functions on metric measure spaces*, Geom. Funct. Anal., 9 (1999), pp. 428–517.

[10] G. Da Prato, M. Röckner, and F.-Y. Wang, *Singular stochastic equations on Hilbert spaces: Harnack inequalities for their transition semigroups*, Journal of Functional Analysis, 257 (2009), pp. 992–1017.

[11] N. Garofalo and A. Mondino, *Li-Yau and Harnack type inequalities in* $\mathrm{RCD}^*(K,N)$ *metric measure spaces*, Nonlinear Anal., 95 (2014), pp. 721–734.

[12] N. Gigli, *Nonsmooth differential geometry - an approach tailored for spaces with Ricci curvature bounded from below.* Accepted at Mem. Amer. Math. Soc., arXiv:1407.0809, 2014.

[13] ——, *On the differential structure of metric measure spaces and applications*, Mem. Amer. Math. Soc., 236 (2015), pp. vi+91.

[14] F.-Z. Gong and F.-Y. Wang, *Heat kernel estimates with application to compactness of manifolds*, Quarterly Journal of Mathematics, 52 (2001), pp. 171–180.

[15] A. Grigor'yan, *Heat kernel upper bounds on a complete non-compact manifold*, Revista Matemática Iberoamericana, 10 (1994), pp. 395–452.

[16] ——, *Integral maximum principle and its applications*, Proceedings of the Royal Society of Edinburgh Section A: Mathematics, 124 (1994), pp. 353–362.

[17] J. Heinonen, *Nonsmooth calculus*, Bull. Amer. Math. Soc. (N.S.), 44 (2007), pp. 163–232.

[18] R. Jiang, *The Li-Yau inequality and heat kernels on metric measure spaces*, J. Math. Pures Appl. (9), 104 (2015), pp. 29–57.

[19] R. Jiang, H. Li, and H.-C. Zhang, *Heat Kernel Bounds on Metric Measure Spaces and Some Applications*, Potential Analysis, 44 (2016), pp. 601–627.

[20] H. Li, *Dimension-Free Harnack Inequalities on* $\mathrm{RCD}(K,\infty)$ *Spaces*, Journal of Theoretical Probability, 29 (2016), pp. 1280–1297.

[21] J. Lott and C. Villani, *Ricci curvature for metric-measure spaces via optimal transport*, Ann. of Math. (2), 169 (2009), pp. 903–991.

[22] N. Shanmugalingam, *Newtonian spaces: an extension of Sobolev spaces to metric measure spaces*, Rev. Mat. Iberoamericana, 16 (2000), pp. 243–279.

[23] K.-T. Sturm, *Heat kernel bounds on manifolds.*, Mathematische Annalen, 292 (1992), pp. 149–162.

[24] ——, *Analysis on local Dirichlet spaces. II. Upper Gaussian estimates for the fundamental solutions of parabolic equations*, Osaka J. Math., 32 (1995), pp. 275–312.

[25] ——, *On the geometry of metric measure spaces. I*, Acta Math., 196 (2006), pp. 65–131.

[26] ——, *On the geometry of metric measure spaces. II*, Acta Math., 196 (2006), pp. 133–177.

[27] F.-Y. Wang, *Equivalence of dimension-free Harnack inequality and curvature condition*, Integral Equations and Operator Theory, 48 (2004), pp. 547–552.

[28] ——, *Equivalent semigroup properties for the curvature-dimension condition*, Bulletin des sciences mathematiques, 135 (2011), pp. 803–815.