HAL open science

Localized bounds on log-derivatives of the heat kernel on incomplete Riemannian manifolds

Robert W. Neel, Ludovic Sacchelli

▶ To cite this version:

Robert W. Neel, Ludovic Sacchelli. Localized bounds on log-derivatives of the heat kernel on incomplete Riemannian manifolds. 2022. hal-03906996

HAL Id: hal-03906996

https://inria.hal.science/hal-03906996v1

Preprint submitted on 19 Dec 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

LOCALIZED BOUNDS ON LOG-DERIVATIVES OF THE HEAT KERNEL ON INCOMPLETE RIEMANNIAN MANIFOLDS

ROBERT W. NEEL AND LUDOVIC SACCHELLI

Abstract. Bounds on the logarithmic derivatives of the heat kernel on a compact Riemannian manifolds have been long known, and were recently extended, for the log-gradient and log-Hessian, to general complete Riemannian manifolds. Here, we further extend these bounds to incomplete Riemannan manifolds under the least restrictive condition on the distance to infinity available, for derivatives of all orders. Moreover, we consider not only the usual heat kernel associated to the Laplace-Beltrami operator, but we also allow the addition of a conservative vector field. We show that these bounds are sharp in general, even for compact manifolds, and we discuss the difficulties that arise when the operator incorporates non-conservative vector fields or when the Riemannian structure is weakened to a sub-Riemannian structure.

1. Introduction

Uniform bounds for the logarithmic derivatives of the heat kernel on a Riemannian manifold, in small time, have been studied since the 90s. In particular, improving on earlier work of Sheu [25], Hsu [14] and Stroock-Turetsky [26] independently proved that, on a compact (smooth) Riemannian manifold M, there exist two families of positive constants C_N and D_N such that

$$\left|\nabla_x^N \log p_t(x,y)\right| \leq C_N \left[\frac{d(x,y)}{t} + \frac{1}{\sqrt{t}}\right]^N, \tag{1}$$

and also

$$\left|\nabla_x^N p_t(x,y)\right| \leq D_N \left[\frac{d(x,y)}{t} + \frac{1}{\sqrt{t}}\right]^N p_t(x,y),$$

both holding for $(t,x,y) \in (0,1] \times M \times M$. (These two families of bounds are widely-known to be equivalent, though, for completeness, we provide the details below.) Here $p_t(x,y)$ is the usual heat kernel (corresponding to the Laplace-Beltrami operator), ∇^N represents Nth-order covariant differentiation with $|\cdot|$ the (pointwise) norm on N-tensors induced by the Riemannian mertric, and the subscript x indicates that the covariant derivative acts on the first spatial variable (that is, the x-variable). In this case, p_t is symmetric, so the result also holds for y-derivatives, but the difference in technique for studying the two spatial variables will play a prominent role in what follows, so we emphasize where our derivatives fall throughout.

The work of Sheu, Hsu, and Stroock-Turetsky naturally applies to x-derivatives because it considers perturbations of the entire (random) path induced by perturbing the starting point of the diffusion. However, a perturbation of the entire path is in principle a global object, making localization, and thus the extension of these results to non-compact manifolds, non-trivial (at least those without bounded geometry). Indeed, see [12] for later work in this spirit, looking at more general maps between vector bundles, but still over a compact Riemannian manifold. Recently, Chen-Li-Wu [5] extended (1) for $N = 1, 2$ to general complete Riemannian manifolds (that is, complete manifolds with no additional assumptions on the geometry).

Interest in such estimates comes from a few sources. For example, control of the logarithmic gradient (with respect to x) of $p_t(x,y)$ is central to the question of whether the Brownian bridge process is a semi-martingale up to its terminal time (which is essentially a question of the integrability of the drift

2010 *Mathematics Subject Classification.* Primary 58J65; Secondary 58J35 58K55 35K08.

Key words and phrases. heat kernel bounds, logarithmic derivatives, small-time asymptotics, incomplete manifolds.

term, given in terms of the logarithmic derivative). While it has long been known that the Brownian bridge is a semi-martingale on a compact Riemannian manifold (see [15]), extending this result to general complete manifolds was only accomplished a few years ago by Güneysu [11], which required sufficient localization of the bound on the logarithmic gradient of the heal kernel. More generally, control of the logarithmic gradient and logarithmic Hessian of the heat kernel, on compact manifolds (or complete manifolds with bounded geometry), has been used in various aspects of stochastic analysis on loop space; we refer to the introduction of [5] for an overview of the literature.

The primary goal of the present paper is to further extend (1) to possibly-incomplete Riemannian manifolds under the weakest condition on the distance to infinity available, for all values of N. (This obviously includes all complete manifolds, with no additional conditions.) Moreover, we consider not only the usual heat kernel associated to the Laplace-Beltrami operator, but we allow the addition of a (smooth) conservative vector field. That is, if we let Δ_{LB} be the Laplace-Beltrami operator, we consider the diffusion operator $\Delta = \Delta_{LB} + Z$ where Z is the gradient vector field of some smooth function on M. The point is that Z being conservative is exactly the condition that there exists a (smooth) symmetrizing reference measure, namely a volume measure ν on M such that the heat kernel is symmetric when written with respect to ν. (See Section 2 for more details on the heat kernel with respect to an arbitrary volume.) More precisely, our main result is the following.

Theorem 1. *Let M be a (possible incomplete) Riemannian manifold, let $\Delta = \Delta_{LB} + Z$, where Δ_{LB} is the Lapace-Beltrami operator and Z is a smooth vector field, and let $p_t(x,y)$ be the heat kernel for Δ (with respect to the Riemannian volume measure). Suppose that Z is conservative (of course, this includes the case $Z \equiv 0$), and let $K \subset M \times M$ be compact such that there exists $\varepsilon > 0$ so that, for all $(x,y) \in K$, $\{z : d(x,z) + d(z,y) < d(x,y) + \varepsilon\}$ has compact closure. Then there exist sequences of constants C_N and D_N such that, for any $N \geq 1$,*

$$\left|\nabla_x^N \log p_t(x,y)\right| \leq C_N \left[\frac{d(x,y)}{t} + \frac{1}{\sqrt{t}}\right]^N$$

$$\text{and } \left|\nabla_x^N p_t(x,y)\right| \leq D_N \left[\frac{d(x,y)}{t} + \frac{1}{\sqrt{t}}\right]^N p_t(x,y)$$

for all $t \in (0,1]$ and $(x,y) \in K$.

The underlying logic of our proof is that y-derivatives are fundamentally easier to localize, because the Markov property applies forward in time, meaning perturbing the endpoint of a (random) path only requires perturbing the path in a neighborhood of the endpoint, as opposed to perturbing the entire path. The existence of a symmetrizing measure then allows us to transfer bounds on y-derivatives to x-derivatives, which, as indicated above, are generally of more interest.

We complement this bound with a natural improvement away from the cut locus. Let $\mathrm{Cut}(M) \subset M \times M$ be the set of (x,y) such that y is in the cut locus of x, or equivalently, x is in the cut locus of y (the relationship is symmetric).

Theorem 2. *Let everything be as in Theorem 1, with the additional condition that $K \cap \mathrm{Cut}(M) = \emptyset$. Then, for any $N \geq 1$,*

$$-4t\nabla_x^N \log p_t(x,y) - \nabla_x^N d^2(x,y) = O(t) \quad \text{as } t \searrow 0,$$

uniformly over $(x,y) \in K$.

A version of this result (with the $O(t)$ replaced by $o(1)$) was proven on a compact Riemannian manifold in the case when $Z \equiv 0$ in [22] (see also [9] for a variant when $N = 2$) and then extended to general complete manifolds in [5]. We are able to localize it and thus extend it to incomplete M just as above.

Of course, these two results can be combined in various ways. To give one example, we have the following.

Corollary 3. *Let everything be as in Theorem 1. Then there is a positive constant c and a sequence of constants C_N (not necessarily the same as in Theorem 1) such that,*

$$|\nabla \log p_t(x,y)| \leq \begin{cases} C_1\left(\frac{d(x,y)}{t}+1\right) & \text{for } d(x,y)\leq c \\ \frac{C_1}{t} & \text{for } d(x,y)>c \end{cases},$$

$$\text{for any } N\geq 2, \quad |\nabla^N \log p_t(x,y)| \leq \begin{cases} \frac{C_N}{t} & \text{for } d(x,y)\leq c \\ \frac{C_N}{t^N} & \text{for } d(x,y)>c \end{cases}$$

for all $(t,x,y)\in(0,1]\times K$.

Remark 4. We have stated Theorems 1 and 2 and Corollary 3 for x-derivatives, as emphasized, but they hold for y-derivatives as well, which is a central feature of the proofs.

In Section 2, we describe the geometric context and notation in more detail, including the stochastic aspect. We also recall previous estimates of Léandre [17] and Bailleul-Norris [1], proven in the generality of sub-Riemannian manifolds, which underlie our localization procedure. In Section 3, we prove the basic localization results for heat kernel derivatives. In this section, we work in greater generality than just Riemannian manifolds, since the method applies to any diffusion satisfying certain small-time asymptotics, and thus is of independent interest. In particular, the results of this section apply on sub-Riemannian manifolds as well. In Section 4, we apply the general localization results to prove Theorems 1 and 2. Finally, in Section 5, we discuss a few complementary issues. First, we note that the the upper bound given in Theorem 1 is sharp in the power of t for all N, which follows from a recently proven result in [24] giving the coefficient of $1/t^N$ in terms of the joint cumulant of geometrically relevant random variables. We illustrate this in the most important case, that of $N=2$, by treating the explicit example of the circle "by hand." Second, we discuss the general (Riemannian) non-compact case, in which no symmetry assumptions are placed on $\Delta+Z$. In this case, while one can still localize y-derivatives, it is unclear whether or not the corresponding result for x-derivatives holds. This is further related to allowing Δ to also include a (smooth) potential. Finally, we observe that in the sub-Riemannian situation, even for a symmetric operator on a compact manifold, no estimate of the type (1) is known, because uniform control of the heat kernel, at the level of precision of (1), near the diagonal is currently unavailable. This is essentially because the diagonal, on a true sub-Riemannian manifold (that is, one which is not Riemannian) is always non-strictly abnormal, and behavior of the heat kernel at abnormals is poorly understood.

As alluded to above, many of the localization ideas we employ also apply to sub-Riemannian manifolds, and this direction is pursued in [24]. Moreover, we make use of existing estimates on sub-Riemannian manifolds, though often only because they accommodate the case when $Z\neq 0$ without additional justification. For example, (1) on a compact M for non-zero Z is established in [24]. Nonetheless, the eventual results in the Riemannian case are significantly stronger, admit shorter proofs, and draw from different literature, making it best to treat them separately from the sub-Riemannian case. For those wishing to avoid any reference to sub-Riemannian results, we point out that the necessary estimates to prove a weaker version of Theorem 1, following the approach in this paper, were essentially already present in the 90s literature on Riemannian manifolds, as discussed in Remark 10.

2. Riemannian background

For a (possibly incomplete) Riemannian manifold M, we let Δ_{LB} be the Laplace-Beltrami operator, and we consider the diffusion operator $\Delta=\Delta_{LB}+Z$ where Z is a smooth vector field. We let X_t be the diffusion generated by Δ; in particular, X_t is a Brownian motion (run at $\sqrt{2}$ times the normal speed) with drift. Of course, it is possible that X_t explodes in finite time, and we understand that the process is killed upon explosion.

If we let μ be the Riemannian volume measure, then X_t has a smooth transition function $p_t(x,y) \in C^\infty((0,\infty)\times M\times M)$ with respect to μ, that is, for any Borel set $A \subset M$

$$\mathbb{P}(X_t \in A | X_0 = x) = \int_{y\in A} p_t(x,y)\, d\mu(y).$$

(Note that the integral over all of M is 1 if and only if the process does not explode, otherwise the integral will be strictly less than 1.) This transition function is also the associated (minimal) heat kernel for Δ (or fundamental solution to the heat equation), meaning that

$$\partial_t p_t(x,y) = \Delta_x p_t(x,y) \quad \text{and} \quad \lim_{t\searrow 0} p_t(x,y) = \delta_x(y)$$

where Δ_x denotes Δ acting on the x-variable (that is, on the first spatial variable) and δ_x is the usual point mass at x. One consequence is that, for any smooth, compactly supported function ϕ on M, the corresponding solution to the heat equation with initial values given by ϕ is $\phi_t(x) = \int_{y\in M} p_t(x,y)\phi(y)\, d\mu(y)$. (See [23] and [7] for classical treatments of the existence and basic properties of the heat kernel on incomplete Riemannian manifolds, and [10] for a modern survey in the context of weighted Riemannian manifolds.)

Note that the right-hand side of (1) is bounded from below by a positive constant for (t,x,y) in any compact subset of $(0,\infty)\times M\times M$, so, for any $t_0 \in (0,1)$ the estimate holds on any compact subset of $M\times M$ for all $t \in [t_0, 1]$ simply by smoothness of the heat kernel. Thus establishing (1) for $t \in (0,1]$ is really a matter of the small-time asymptotics of the heat kernel. Said differently, once we have shown (1) holds uniformly for $t \in (0,t_0)$, we can extend it to $t \in (0,1]$ at the cost of possibly adjusting the constants C_N. Our basic approach is to combine a localization result for the heat kernel itself with a global result on y-derivatives of the heat kernel in order to localize these derivatives. To provide context for the argument, we first introduce the necessary notation and describe the known estimates in the Riemannian case.

We will frequently restrict or extend the domain of the heat kernel, and for any open set U, we let $p_t^U(x,y)$ be the heat kernel on U (which should be understood with Dirichlet boundary conditions– that is, it is the transition density for the process killed upon explosion, or equivalently, upon the first exit from U). Note that since we are not assuming that our Riemannian manifolds are complete, there is no real distinction between a Riemannian manifold and an open subset of a Riemannian manifold. For a closed set A, we let $p_t(x,A,y) = p_t(x,y) - p_t^{A^c}(x,y)$, where $p_t^{A^c}(x,y)$ is extended to be zero for either x or y inside of A (that is, outside of A^c). In particular, $p_t(x,A,y)$ gives the contribution to $p_t(x,y)$ from paths that pass through A. Further, we define

$$d(x,A,y) = \inf\{d(x,z) + d(z,y) : z \in A\},$$

which gives the distance from x to y via paths that pass through A. We will also consider the heat kernel with respect to different reference measures, and we will write $p_t^\nu(x,y)$ for the heat kernel with respect to a smooth volume measure ν (meaning a measure with a smooth, positive density in any system of local cooridnates) when there is a chance of confusion. This means that, for any Borel set $A \subset M$

$$\mathbb{P}(X_t \in A | X_0 = x) = \int_{y\in A} p_t^\nu(x,y)\, d\nu(y),$$

and thus $d\nu = \frac{d\nu}{d\mu} d\mu$ for a smooth, positive Radon-Nikodym derivative $\frac{d\nu}{d\mu}$.

Assume that M a possibly incomplete Riemannian manifold and Δ is symmetric, with respect to some smooth reference volume ν. Then Theorem 1.1 of Bailleul-Norris [1] gives that (here the heat kernel is understood with respect to ν)

$$\limsup_{t\searrow 0} 4t \log p_t(x,A,y) \le -d^2(x,A,y) \tag{2}$$

uniformly for x and y in any compact subset of $M\setminus\partial A$. Further, under the same assumptions, Theorem 1.2 of [1] says that

$$\lim_{t\searrow 0} 4t \log p_t(x,y) = -d^2(x,y) \tag{3}$$

uniformly for (x, y) in any compact subset of $M \times M$.

Next, consider a Riemannian metric on $\mathbb{R}^n$ given by a global orthonormal frame $Z_1, \ldots, Z_n$, where the Z_i are smooth, bounded vector fields with bounded derivatives of all orders. Further, suppose that if we write Δ as $\sum_{i=1}^n Z_i^2 + Z_0$, then Z_0 is also a smooth, bounded vector field with bounded derivatives of all orders. (We don't pursue the question of when a metric admits such a representation, since eventually we localize the following estimate to a small enough neighborhood of a point.) In this case, Léandre [17] proves that, for any multi-index α,

$$\limsup_{t \searrow 0} 4t \log \left(\left| \partial_y^\alpha p_t(x, y) \right| \right) \leq -d^2(x, y), \tag{4}$$

uniformly for (x, y) in any compact subset of $\mathbb{R}^n \times \mathbb{R}^n$. Here $d(x, y)$ is the Riemannian distance, the partial derivatives are understood with respect to the standard Cartesian coordinates, and the result holds for any smooth reference volume on $\mathbb{R}^n$ (which is clear via the same logic as in Lemma 6 below). Note that this result doesn't require any symmetry of Δ.

We continue with a simple consequence of stochastic calculus for exit times of balls on M.

Lemma 5. *Let A be a compact subset of a (possibly incomplete) Riemannian manifold M, let a be a positive constant, and let σ_a be the first exit time of the diffusion X_t (determined by having infinitesimal generator Δ) from the open ball of radius a around $X_0 = x$ (that is, the first time the diffusion moves a distance a from its starting point, which never happens if the diffusion explodes first). Then there exists $T > 0$ such that*

$$\mathbb{P}\left(\sigma_a < T | X_0 = x\right) < \frac{1}{2}$$

for any $x \in A$.

Proof. Since σ_a is non-decreasing in a, without loss of generality we can assume that a is small enough so that an a-neighborhood of A, briefly A^a, is relatively compact, and that a is less than the injectivity radius at all $x \in A$. This follows from the fact that the injectivity radius is continuous and everywhere positive, and thus is bounded from below on a compact. (The continuity of the injectivity radius appears widely known; for example, it is asserted without proof in [4]. For a recent proof, one can see Section 10.8 of [3].) Thus the squared distance function is smooth on $A \times A^a$.

In particular, for any $x \in A$, let a be fixed as above, and let r be the distance from x, and $r_t = r(X_t)$ be the usual radial process, where $X_0 = x$ so that $r_0 = 0$. Then Itô's lemma implies there exists some (standard one-dimensional) Brownian motion W_t such that r_t^2 satisfies the SDE (written in integral form)

$$r_{t \wedge \sigma_a}^2 = \int_0^{t \wedge \sigma_a} 2 r_s \, dW_s + \int_0^{t \wedge \sigma_a} \frac{1}{2} \Delta\left(r_s^2\right) \, ds.$$

Now by smoothness of the squared distance, there exists some $\alpha > 0$ such that the last integral on the right-hand side is bounded from above by $\alpha(t \wedge \sigma_a)$, and thus also by αt, uniformly in x. (Indeed, it is well known that α can be given in terms of a lower bound on Ricci curvature and an upper bound on the length of Z over A^a.) Moreover, the first integral on the right-hand side is a martingale started from 0, and its quadratic variation is bounded from above by $4a^2 t$, uniformly in x. Thus, by Chebyshev's inequality, it converges to 0 in probability as $t \searrow 0$, uniformly in x. On the other hand, we have $a^2 \mathbb{P}^x(\sigma_a < t) \leq r_{t \wedge \sigma_a}^2$. Putting this together, we have

$$\mathbb{P}^x(\sigma_a < t) \leq \frac{1}{a^2} \left(\int_0^{t \wedge \sigma_a} 2 r_s \, dW_s + \alpha t \right),$$

where the right-hand side goes to 0 in probability with t (for any fixed $a > 0$), uniformly over $x \in A$. The lemma follows. □

3. Localizing derivatives of the heat kernel

For this section, we work in a more general situation than the rest of the paper. In particular, let M be a smooth manifold (of dimension n) with a smooth volume ν. We suppose that we have a diffusion X_t on M (that may explode in finite time with positive probability) with a smooth transition density

$p_t(x,y)$ with respect to ν; that is, started from any $x \in M$, X_t almost surely has continuous paths, satisfies the strong Markov property (for stopping times with respect to the filtration it generates), and for any Borel $A \subset M$,

$$\mathbb{P}\left(X_t \in A | X_0 = x\right) = \int_{y \in A} p_t(x,y)\, d\nu(y).$$

We also assume that $d(\cdot,\cdot)$ is a metric on M which is compatible with the smooth manifold topology (so it is also continuous). Finally, we assume that the diffusion satisfies the conclusion of Lemma 5 (where balls are understood with respect to the metric d); that is, we assume that the probability of the diffusion from x leaving an open ball of radius $a > 0$ around x before time T can be made less than $1/2$ by choosing T small enough, uniformly over all x in a compact. (Recall that the hitting time of a closed set is a stopping time for the filtration generated by X_t.) The point is that the basic localization of y-derivatives can be established for such a $p_t(x,y)$ using only the type of small-time estimates described in the previous section, and no other aspects of the Riemannian structure.

We begin with some preliminary observations, essentially amounting to smooth calculus, which we collect in the following lemma. For any system of (smooth) coordinates $u_1, \dots, u_n$ on an open set $U \subset M$ with compact closure, we say that the coordinates are extendable if they can be extended to a neighborhood of the closure of U. Note that if we have any two such extendable coordinate systems on U, the transition function between them is smooth on a neighborhood of the closure of U.

Lemma 6. *Let M and $p_t(x,y)$ be as above. Let $B \subset M$ be an open ball with compact closure that admits (smooth) coordinates $u_1, \dots, u_n$ that extend to some neighborhood of $\overline{B}$, and let $K \subset M \times M$ be compact, such that $\pi_2(K)$, the projection of K onto the second factor (respectively, $\pi_1(K)$, the projection onto the first factor) is contained in B. (Below we understand partial derivatives ∂^α on B to be with respect to the u_i.)*

Then there exists a sequence of positive constants C_N such that, for any $N \geq 1$,

$$|\partial^\alpha \log p_t^\nu(x,y)| \leq C_N \left[\frac{d(x,y)}{t} + \frac{1}{\sqrt{t}}\right]^N \quad \text{for any multi-index } \alpha \text{ of weight } N, \tag{5}$$

uniformly for $(x,y) \in K$, with the derivatives acting in the y-variable (respectively, the x-variable), if and only if there exists a sequence of positive constants D_N such that, for any $N \geq 1$,

$$|\partial^\alpha p_t^\nu(x,y)| \leq D_N \left[\frac{d(x,y)}{t} + \frac{1}{\sqrt{t}}\right]^N p_t^\nu(x,y) \quad \text{for any multi-index } \alpha \text{ of weight } N, \tag{6}$$

uniformly for $(x,y) \in K$, with the derivatives acting in the y-variable (respectively, the x-variable).

Further, if either of (5) and (6) hold (with the derivatives acting either on x or on y) with respect to some smooth volume ν, then it holds with respect to any smooth volume (that is, for $p_t^{\nu'}$ where ν' is another smooth volume) with possibly different sequences of constants C_N and D_N.

Proof. Assume (5) holds (for all N), with the partials on either y or x. For convenience, let $\ell_t(x,y) = \log p_t^\nu(x,y)$. Then we have

$$\partial^\alpha p_t^\nu(x,y) = \partial^{n_1}\partial^{n_2}\cdots\partial^{n_N} e^{\ell_t(x,y)}$$

where the n_i correspond to the indices in α, and $N = |\alpha|$. Here the partials are understood to act either all on y or all on x. Faà di Bruno's formula gives

$$\partial^\alpha p_t^\nu(x,y) = \sum_{\pi \in \Pi} e^{\ell_t(x,y)} \prod_{B \in \pi} \partial^B \ell_t(x,y) = p_t^\nu(x,y) \sum_{\pi \in \Pi} \prod_{B \in \pi} \partial^B \ell_t(x,y)$$

where the sum is over all partitions π of $\{1, \dots, N\}$, the product is over all blocks in the partition π, and ∂^B is understood to mean $\partial^{n_{i_1}} \cdots \partial^{n_{i_{|B|}}}$ where $i_1, \dots, i_{|B|}$ are the elements of B and $|B|$ is the cardinality of the block B. Applying (5) and using that the sum of the $|B|$ over all $B \in \pi$ is necessarily N, we have

$$|\partial^\alpha p_t^\nu(x,y)| \leq p_t^\nu(x,y) \sum_{\pi \in \Pi} \prod_{B \in \pi} C_{|B|} \left[\frac{d(x,y)}{t} + \frac{1}{\sqrt{t}}\right]^{|B|} = p_t^\nu(x,y) \left[\frac{d(x,y)}{t} + \frac{1}{\sqrt{t}}\right]^N \left(\sum_{\pi \in \Pi} \prod_{B \in \pi} C_{|B|}\right),$$

and thus (6) holds with $D'_N = \sum_{\pi\in\Pi}\prod_{B\in\pi} C_{|B|}$.

The other direction is analogous. Assuming that (6) holds (for all N), and using the same notation as above, we have

$$\partial^\alpha \log p_t^\nu(x,y) = \sum_{\pi\in\Pi} \frac{(-1)^{|\pi|-1}(|\pi|-1)!}{(p_t^\nu(x,y))^{|\pi|}} \prod_{B\in\pi} \partial^B p_t^\nu(x,y),$$

so that

$$\begin{aligned}|\partial^\alpha \log p_t^\nu(x,y)| &\le \sum_{\pi\in\Pi} \frac{(-1)^{|\pi|-1}(|\pi|-1)!}{(p_t^\nu(x,y))^{|\pi|}} \prod_{B\in\pi} p_t^\nu(x,y) D_{|B|}\left[\frac{d(x,y)}{t}+\frac{1}{\sqrt{t}}\right]^{|B|} \\ &= \left[\frac{d(x,y)}{t}+\frac{1}{\sqrt{t}}\right]^N \left(\sum_{\pi\in\Pi}(-1)^{|\pi|-1}(|\pi|-1)!\prod_{B\in\pi} D_{|B|}\right).\end{aligned}$$

Thus (5) holds with $C_N = \sum_{\pi\in\Pi}(-1)^{|\pi|-1}(|\pi|-1)!\prod_{B\in\pi} D_{|B|}$. Again, the computation is valid with ∂^α acting on either coordinate.

Next, assume that one, and thus both, of (5) and (6) hold with derivatives on the y-variable. Then if ν' is any other smooth volume, the Radon-Nikodym derivative $\frac{d\nu}{d\nu'}$ is smooth. The product rule gives

$$\partial_y^\alpha p_t^{\nu'}(x,y) = \partial_y^{n_1}\partial_y^{n_2}\cdots\partial_y^{n_N}\left(p_t^\nu(x,y)\frac{d\nu}{d\nu'}(y)\right) = \sum_{A\subset 1,\dots,N}\left(\partial_y^A p_t^\nu(x,y)\right)\left(\partial_y^{A^c}(y)\frac{d\nu}{d\nu'}\right)$$

where the sum is over all subsets of $\{1,\dots,N\}$ and for each term, the partial derivatives are divided into A and its complement (with $\partial_y^\emptyset f(y) = f(y)$). Since we work on a compact K, each factor $\partial_y^{A^c}(y)\frac{d\nu}{d\nu'}$ is bounded on K. Then using (6) on each factor $\partial_y^A p_t^\nu(x,y)$, since $|A|\le|\alpha| = N$, we see that (6) holds for $\partial_y^\alpha p_t^{\nu'}(x,y)$ with possibly different constants. By the above, (5) also holds for $\partial_y^\alpha p_t^{\nu'}(x,y)$, with possibly different constants. Finally, if one, and thus both, of (5) and (6) hold with derivatives on the x-variable, then

$$\partial_x^\alpha p_t^{\nu'}(x,y) = \left(\partial_x^\alpha p_t^\nu(x,y)\right)\frac{d\nu}{d\nu'}(y),$$

and the desired result again follows by smoothness and compactness. □

Next, we show that estimates on the y-derivatives of the type in (4) can be localized to a neighborhood of the endpoint y.

Lemma 7. *For M and $p_t(x,y)$ as described above, suppose that, for some $\eta>0$, B, B' and B'' are concentric open balls of radii $\eta/2$, $(3/2)\eta$, and $(7/2)\eta$, respectively, and that B'' has compact closure. Suppose further that we have an extendable coordinate system on B'' and that, for any multi-index α, we have*

$$\limsup_{t\searrow 0} 4t\log\left(\left|\partial_y^\alpha p_t(x,y)\right|\right) \le -d^2(x,y), \tag{7}$$

uniformly for x and y in $\overline{B''}$ (where the partial derivatives are understood with respect to this coordinate system). Then, for any multi-index α,

$$\limsup_{t\searrow 0} 4t\log\left(\left|\partial_y^\alpha p_t^{B''}(x,y)\right|\right) \le -d^2(x,y) \tag{8}$$

uniformly over $(x,y)\in\overline{B'}\times B$.

Proof. For $(x,y)\in\overline{B'}\times B$ we let μ_h^x be the (spacetime) hitting measure of $\partial B''$ for the diffusion started from x; that is, μ_h^x is a (sub-) probability measure on $\partial B''\times(0,\infty)$ that describes the first place and time the diffusion from x hits $\partial B''$. Then the strong Markov property implies we have the path decomposition

$$p_t^{B''}(x,y) = p_t(x,y) - \int_{(z,\tau)\in\partial B''\times(0,t)} p_{t-\tau}(z,y)\,d\mu_h^x. \tag{9}$$

Turning our attention to derivatives, since p_t is smooth in y for all positive times, we have

$$\partial_y^i \int_{(z,\tau)\in\partial B''\times(0,t)} p_{t-\tau}(z,y)\,d\mu_h^x = \int_{(z,\tau)\in\partial B''\times(0,t)} \partial_y^i p_{t-\tau}(z,y)\,d\mu_h^x$$

for any i. Iterating this, we have that, for any multi-index α,

$$\partial_y^\alpha \int_{(z,\tau)\in\partial B''\times(0,t)} p_{t-\tau}(z,y)\,d\mu_h^x = \int_{(z,\tau)\in\partial B''\times(0,t)} \partial_y^\alpha p_{t-\tau}(z,y)\,d\mu_h^x, \tag{10}$$

for all $y \in B$ and any x.

From (7) and the construction of B'', for any $\delta > 0$, the integrand in the right-hand side of (10) has absolute value bounded by $e^{-(3\eta)^2+\delta}$ once t is small enough, uniformly in $(z,y) \in \partial B'' \times B$. Using this, and applying ∂_y^α to both sides of (9), we find that, for any multi-index α,

$$\left|\partial_y^\alpha p_t^{B''}(x,y) - \partial_y^\alpha p_t(x,y)\right| \le e^{-(3\eta)^2+\delta}$$

for all $(x,y) \in \overline{B'} \times B$, if t is small enough. On the other hand, $\partial_y^\alpha p_t(x,y)$ term satisfies (7) with $d^2(x,y) \le (2\eta)^2$ for $(x,y) \in \overline{B'} \times B$, so its contribution to the bound on $\partial_y^\alpha p_t^{B''}(x,y)$ is more significant by an exponential factor (which is the underlying idea of localizing the estimate). At any rate, it follows that

$$\limsup_{t\searrow 0} 4t \log\left(\left|\partial_y^\alpha p_t^{B''}(x,y)\right|\right) \le -d^2(x,y)$$

uniformly over $(x,y) \in \overline{B'} \times B$, and the result is proven. □

Finally, we show how to combine this local result with a localization result for the heat kernel itself to localize derivatives.

Lemma 8. *In the situation described above, suppose that, $K_0 \subset U_0 \subset K_1 \subset U_1 \subset M$ where K_0 and K_1 are compact and U_0 and U_1 are open with compact closure, and suppose that, for some $\varepsilon > 0$, the heat kernel on M satisfies the estimate*

$$\limsup_{t\searrow 0} 4t \log p_t\left(x, U_1^c, y\right) \le -\left(d(x,y)+\varepsilon\right)^2 \tag{11}$$

uniformly for x and y in K_1.

Suppose further that there exists $\eta > 0$ and $y_0 \in K_0$, such that, if we consider the concentric balls B, B', B'' centered at y_0 (as in Lemma 7), the closure of B'' is contained in U_0, and the conclusion of Lemma 7 holds (including the existence of an appropriate extendable coordinate system). Then for any multi-index α,

$$\limsup_{t\searrow 0} 4t \log\left(\left|\partial_y^\alpha p_t(x, U_1^c, y)\right|\right) \le \left(d(x,y)+\varepsilon-3\eta\right)^2 \tag{12}$$

uniformly over $x \in K_0$ and $y \in B$.

Proof. Let $y_0 \in K_0$ and the corresponding concentric balls B, B', and B'' be as in the lemma. Let σ be the first hitting time (of the diffusion X_t) of ∂U_1, let τ_1 be the first hitting time of $\partial B'$ after σ, and let τ_1' be the first hitting time of $\partial B''$ after τ_1. Then, for $i > 1$, we iteratively define τ_i to be the first hitting time of $\partial B'$ after τ_{i-1}' and τ_i' the first hitting time of $\partial B''$ after τ_i. In the notation of (2), we have

$$\partial_y^\alpha p_t(x, U_1^c, y) = \partial_y^\alpha p_t(x,y) - \partial_y^\alpha p_t^{U_1}(x,y).$$

We understand $p_t^{B''}(x,y)$ to be zero whenever $x \notin B''$ or $t < 0$, and we let μ_i^x be the (spacetime) distribution of (τ_i, X_{τ_i}) when the diffusion is started from x, so that in particular, μ_i^x is a sub-probability measure on $[0,\infty) \times \partial B'$. Then we have the path decomposition

$$p_t(x, U_1^c, y) = \sum_{i=1}^\infty \int_{(z,\tau)\in\partial B'\times(0,t)} p_{t-\tau}^{B''}(z,y)\;d\mu_i^x, \tag{13}$$

which is valid for any $x \in K_0$ and $y \in B$.

We now bound the distributions of the τ_i, starting with τ_1. For any small enough δ, let V be a δ-neighborhood of $\partial B'$. By the conclusion of Lemma 5, which we assume holds, there is a t_0, such that, for any $z \in K_1$, the probability that the diffusion, started from z, leaves the ball of radius δ centered at z before t_0 is less than $1/2$. It follows that, for any $t < t_0$, for any $x \in K_0$,

$$\frac{1}{2}\mathbb{P}^x\left(\tau_1 \le t\right) \le \int_V p_t\left(x, U^c, z\right)\, d\nu(z).$$

On the other hand, by the triangle inequality and the definitions of U_1, σ, and τ_1, we have $d(x, U_1^c, z) \ge d\left(x, \partial B'\right) + \varepsilon - \delta$ for all $z \in V$ and $x \in K_1$. Note that $\overline{B''}$ is a compact subset of U_0, and thus for small enough δ, (11) holds for $y \in V$. So for all small enough t,

$$\int_V p_t(x,z)\, d\nu(z) \le \nu(V) \exp\left[-\frac{\left(d(x,\partial B') + \varepsilon - 2\delta\right)^2}{4t}\right],$$

uniformly in $x \in K_0$. Since $\nu(V)$ is positive for all $\delta > 0$, we conclude that

$$\limsup_{t \searrow 0} 4t \log \mathbb{P}^x\left(\tau_1 \le t\right) \le -\left(d(x,\partial B') + \varepsilon\right)^2 \tag{14}$$

uniformly over $x \in K_0$.

By the strong Markov property, the distribution of $\tau_i - \tau'_{i-1}$ depends only on $X_{\tau'_{i-1}}$; in particular, it doesn't depend on i, and similarly for the distribution of $\tau'_i - \tau_i$. Moreover, $\tau_i - \tau'_{i-1}$ is conditionally independent of all of the previous increments $\tau_j - \tau'_{j-1}$ and $\tau'_{j-1} - \tau_{j-1}$ for $j < i$ given $X_{\tau'_{i-1}}$, and similarly for $\tau'_i - \tau_i$. Now, in order for τ_i to be less than t, $\tau_i - \tau'_{i-1}$ and $\tau'_{i-1} - \tau_{i-1}$ must both be less than t. Further, $\tau_i - \tau'_{i-1}$ can only be less than t if the shifted diffusion $\{X_{\tau'_{i-1}+t}, t \ge 0\}$ moves a distance 2η from its starting point (at $X_{\tau'_{i-1}} \in \partial B''$) in time t or less, and similarly, $\tau'_{i-1} - \tau_{i-1}$ can only be less than or equal to t if it moves a distance 2η in time t or less. By independence and the conclusion of Lemma 5, it follows that there exists t_0 such that, for $t < t_0$ and any $i \ge 2$,

$$\mathbb{P}^x\left(\tau_i - \tau_1 < t\right) < \left(\frac{1}{4}\right)^{i-1}$$

for all $x \in K_0$.

Combining this with (14) and again using the independence, we see that, for any $\delta > 0$ there exists $t_0 > 0$ such that

$$\mathbb{P}^x\left(\tau_i \le t\right) \le \left(\frac{1}{4}\right)^{i-1} \exp\left[-\frac{\left(d(x,\partial B') + \varepsilon - \delta\right)^2}{4t}\right]$$

whenever $0 < t < t_0$, for all $x \in K_0$ and all $i = 1, 2, 3, \ldots$. Returning to the spacetime hitting measures μ_i^x in (13), we see that if we consider their restrictions to $[0,t]$ (for $t < t_0$), their masses are absolutely summable, by the above and the summability of the geometric series. More precisely, if we let $\mu^{x,t}$ denote their sum, then it is a finite measure on $\partial B' \times [0,t]$ with total mass bounded from above by

$$F(t) = \frac{1}{2} \exp\left[-\frac{\left(d(x,\partial B') + \varepsilon - \delta\right)^2}{4t}\right] \tag{15}$$

for any $x \in K_0$. That is, if $F^{x,t_0}(t)$ is the distribution function of the time-marginal of μ^{x,t_0}, viewed as a measure on $[0,t_0]$ indexed by x, then $F(t)$ is an upper bound on $F^{x,t_0}(t)$ (for all $t \in [0,t_0]$) which holds uniformly in x. Combining this with (13) and the smoothness of $p_t^{B''}(z,y)$ in $y \in B$ for fixed $t > 0$ (as elaborated on in the proof of Lemma 7), the dominated convergence theorem allows us to show that

$$\partial_y^\alpha p_t(x, U_1^c, y) = \int_{(z,\tau) \in \partial B' \times (0,t)} \partial_y^\alpha p_{t-\tau}^{B''}(z,y)\, d\mu^{x,t}.$$

Using (8), we see that, for any δ, for all small enough t (uniformly in $x \in K_0$ and $y \in B$),

$$\begin{aligned}\left|\partial_y^\alpha p_t(x, U_1^c, y)\right| &\leq \int_{(z,\tau)\in\partial B'\times(0,t)} \exp\left[-\frac{(d(\partial B', y)-\delta)^2}{4(t-\tau)}\right] d\mu^{x,t} \\ &= \int_{\tau\in(0,t)} \exp\left[-\frac{(d(\partial B', y)-\delta)^2}{4(t-\tau)}\right] dF^{x,t_0}(\tau),\end{aligned}$$

where we've used the fact that the integrand doesn't depend on $z \in \partial B'$ to reduce to the time-marginal of $\mu^{x,t}$. Because the integrand is monotone decreasing, we get an upper bound from the upper bound (15) on the distribution function, so that, for some $C > 0$

$$\begin{aligned}&\left|\partial_y^\alpha p_t(x, U_1^c, y)\right| \leq \int_{\tau\in\times(0,t)} \exp\left[-\frac{(d(\partial B', y)-\delta)^2}{4(t-\tau)}\right] dF(\tau) \\ &= \int_{\tau\in(0,t)} \exp\left[-\frac{(d(\partial B', y)-\delta)^2}{4(t-\tau)}\right] \exp\left[-\frac{(d(x,\partial B')+\varepsilon-\delta)^2}{4t}\right] \frac{(d(x,\partial B')+\varepsilon-\delta)^2}{8t^2}\, d\tau \\ &\leq C\int_{\tau\in(0,t)} \exp\left[-\frac{(d(\partial B', y)-\delta)^2}{4(t-\tau)}\right] \exp\left[-\frac{(d(x,\partial B')+\varepsilon-2\delta)^2}{4t}\right] d\tau,\end{aligned}$$

where, in the final line, we've absorbed the $((d(x,\partial B')+\varepsilon-\delta)^2)/(8t^2)$ into the second exponential at the cost of doubling δ and multiplying by a constant.

Note that, for any constants $a, b, t > 0$ the maximum of

$$-\frac{a^2}{4(t-\tau)} - \frac{b^2}{4\tau} \quad \text{for } \tau \in (0,t)$$

is $-(a+b)^2/(4t)$ (achieved at $\tau = t/(1+(a/b))$). Then, applying this to the integrand above and using that $\delta > 0$ is arbitrary to clean up the constants, we have that, for any $\delta > 0$, there exists $t_0 > 0$ such that, for all $t \in (0, t_0)$, all $y \in B$, and all $x \in K_0$,

$$\left|\partial_y^\alpha p_t(x, U_1^c, y)\right| \leq \exp\left[-\frac{(d(x,\partial B')+d(\partial B', y)+\varepsilon-\delta)^2}{4t}\right].$$

Since the diameter of B' is at most 3η, we conclude that, for every α,

$$\limsup_{t\searrow 0} 4t\log\left(\left|\partial_y^\alpha p_t(x, U_1^c, y)\right|\right) \leq (d(x,y)+\varepsilon-3\eta-\delta)^2$$

uniformly over $(x, y) \in K_0 \times B$. Since δ is arbitrary, we can send it to zero, completing the proof. □

4. Proving the Riemannian bounds

We now return to the situation of Theorems 1 and 2; namely, where M is a (possibly incomplete) Riemannian manifold, $\Delta = \Delta_{LB} + Z$ where Z is conservative, $d(\cdot,\cdot)$ is the Riemannian distance, and $p_t(x, y)$ is the corresponding heat kernel with respect to the Riemannian volume. This is a special case of the situation considered in the previous section, so those results apply here.

As is clear from the previous two sections, we mostly work with partial derivatives in coordinates. But since we eventually want to give an estimate for covariant derivatives, we now relate the two. Since covariant differentiation is tensorial in the vector fields along which we differentiate, it's essentailly enough to look at the coordinate vector fields.

Lemma 9. *Let $B \subset M$ be an open ball in M with compact closure and extendable coordinates $\partial^1, \ldots, \partial^n$, and suppose that f is a smooth function on a neighborhood of $\overline{B}$. Then if $I = (i_1, \ldots, i_N)$ is a sequence of indices from $1, \ldots, n$, we have*

$$\nabla^N_{\partial^I} f(x) = \nabla^N_{\partial^{i_1},\ldots,\partial^{i_N}} f(x) = \partial^{i_1}\cdots\partial^{i_N} f(x) + \sum_{J:J\subsetneq I} F^{I,J}(x)\partial^J f(x)$$

where the sum is over subsequences J of I which are neither empty nor all of I, and each $F^{I,J}$ is a smooth function on a neighborhood of $\overline{B}$ depending on the Christoffel symbols (of the Levi-Civita connection on M).

Further, there exists $c > 0$ (depending only on N, B, and the choice of coordinates), such that, for any collection $V_1, \dots, V_N$ of (smooth) vector fields of length 1 or less,

$$\nabla^N_{V_1,\dots,V_N} f(x) = \sum_{I \in \{1,\dots,n\}^N} a_I(x) \nabla^N_{\partial^I} f(x),$$

where the a_I are smooth functions on a neighborhood of $\overline{B}$, indexed by all sequences of indices I and depending on the V_i, such that $|a_I(x)| < c$ for all I and all x in a neighborhood of $\overline{B}$.

Proof. The first result is trivial for $N = 1$ and well known for $N = 2$, where the formula for the $(0, 2)$-version of the Hessian in coordinates is

$$\nabla^2_{\partial_j,\partial_i} f = \partial_j \partial_i f - \sum_{k=1}^n \Gamma^k_{ji} \partial_k f.$$

We now proceed by induction, assuming the result for any $\nabla^N_{\partial^I}$. Since $\nabla^N_{\partial^I}$ is a $(0, N)$-tensor, we compute its covariant derivative as

$$\nabla_{\partial_j} \nabla^N_{\partial^I} f = \partial_j \nabla^N_{\partial^I} f - \nabla^N_{\nabla_{\partial_j}\partial^{i_1},\partial^{i_2}\dots,\partial^{i_N}} f - \dots - \nabla^N_{\partial^{i_1}\dots,\partial^{i_{N-1}},\nabla_{\partial_j}\partial^{i_N}} f$$

Then using the definition of the Christoffel symbols and multilinearity, we have

$$\nabla_{\partial_j} \nabla^N_{\partial^I} f = \partial_j \nabla^N_{\partial^I} f - \sum_{k=1}^n \left[\Gamma^k_{ji_1} \nabla^N_{\partial^k,\partial^{i_2}\dots,\partial^{i_N}} f - \dots - \Gamma^k_{ji_N} \nabla^N_{\partial^{i_1}\dots,\partial^{i_{N-1}},\partial^k} f \right].$$

Using the induction assumption on all Nth-order covariant derivatives on the right-hand side and basic calculus, the result follows for $\nabla_{\partial_j} \nabla^N_{\partial^I} f$. But since all $N + 1$st-order covariant derivatives under consideration are of this form for some j and I, the first claim follows.

For the second, since the metric in the coordinate system is smooth and uniformly positive definite, any unit vector can be written as a linear combination of coordinate vector fields with uniformly bounded coefficients. Then the claim follows by linear algebra. □

Proof of Theorem 1. Note that the condition on K, namely that $\{z : d(x, z) + d(z, y) < d(x, y) + \varepsilon\}$ has compact closure for all $(x, y) \in K$, is symmetric in x and y, that is, it is preserved under reflection across the diagonal in $M \times M$. Therefore, we can enlarge K by taking the union with its reflection to get a symmetric set which satisfies the same condition, and obviously it's enough to prove the theorem for this (potentially) larger set. Thus, without loss of generality, we assume that K is symmetric for the rest of the proof.

Because Z is conservative, there exists a smooth measure ν such that $p^\nu_t(x, y)$ is symmetric. To elaborate, recall that $\Delta_{LB} = \operatorname{div}_\mu \circ \operatorname{grad}$, where div_μ is the divergence with respect to the Riemannian volume μ and grad is the Riemannian gradient. Any other smooth volume ν can be written as $e^f \mu$ for some smooth f, and then we can write our operator in divergence form with respect to ν as

$$\Delta = \operatorname{div}_\nu \circ \operatorname{grad} - \operatorname{grad} f + Z,$$

where $\operatorname{grad} f$ is a vector field acting on smooth functions. Since $\operatorname{div}_\nu \circ \operatorname{grad}$ is symmetric with respect to ν and $Z - \operatorname{grad} f$ is only when it is identically 0, we see that Δ admits a symmetrizing volume if and only if $Z = \operatorname{grad} f$ for some smooth f, in which case the symmetrizing measure is $\nu = e^f \mu$ (unique up to constant rescaling). Thus, we now let ν be this measure for the given Z, and we note that p^ν_t satisfies the assumptions of the previous section and is also symmetric.

Let y_0 be any point on M. Then there exists some η such that the concentric balls B, B', and B'' around y_0 are such that B'' has compact closure and is contained in a single extendable coordinate system. Using a bump function, B'' can be isometrically included in a Riemannian manifold M' diffeomorphic to $\mathbb{R}^n$ with a Riemannian metric and smooth volume that satisfy the assumptions under which Léandre proved (4). In particular, (7) is satisfied for $p^{\nu,M'}_t(x, y)$ uniformly for (x, y) in any

compact subset of $M' \times M'$, and thus for x and y in $\overline{B''}$. It follows, by Lemma 7, that, for any multi-index α,

$$\limsup_{t \searrow 0} 4t \log \left(\left| \partial_y^\alpha p_t^{\nu, B''}(x,y) \right| \right) \leq -d^2(x,y) \tag{16}$$

uniformly over $(x,y) \in \overline{B'} \times B$. But since $p_t^{\nu,B''}(x,y)$ only depends on (the Riemannian structure and the volume on) B'', it this holds for $p_t^{\nu,B''}(x,y)$ with B'' viewed as an open ball in the original M as well. Also note that, for fixed y_0, if (16) holds for some η, then it also holds for any smaller η.

Now take any $(x_0, y_0) \in K$. By the triangle inequality and continuity of the distance, there is some $\delta > 0$ such that, if $N_{x_0}(\delta)$ is an open δ-ball around x_0 and $N_{y_0}(\delta)$ is an open δ-ball around y_0 (here we use N instead of B for these balls to avoid notational confusion with B, B', and B''), then

$$U_1 = \bigcup_{x \in N_{x_0}(\delta), y \in N_{y_0}(\delta)} \left\{ z : d(x,z) + d(z,y) < d(x,y) + \frac{\varepsilon}{2} \right\}$$

is open with compact closure (because it is a subset of $\{z : d(x_0,z) + d(z,y_0) < d(x_0,y_0) + \varepsilon\}$ for small enough δ). Now let

$$\begin{aligned} K_0 &= \overline{N_{x_0}(\delta/2)} \cup \overline{N_{y_0}(\delta/2)}, \\ U_0 &= N_{x_0}(\delta) \cup N_{y_0}(\delta), \\ \text{and} \quad K_1 &= \overline{U_0}. \end{aligned}$$

Then $K_0 \subset U_0 \subset K_1 \subset U_1 \subset M$ where K_0 and K_1 are compact and U_0 and U_1 are open with compact closure. Further, again by the triangle inequality, we have that there exists $\varepsilon' > 0$ such that

$$d\left(x, U_1^c, y\right) > d(x,y) + \varepsilon' \tag{17}$$

for all x and y in K_1. Because Δ is symmetric with respect to ν and K_1 is a compact subset of U_1, we can apply the result of Bailleul-Norris, so that (2) yields

$$\limsup_{t \searrow 0} 4t \log p_t^\nu(x, U_1^c, y) \leq -\left(d(x,y) + \varepsilon'\right)^2$$

uniformly for x and y in K_1. After possibly shrinking η, we have that $\overline{B''} \subset U_0$, which means we've shown that all of the assumptions of Lemma 8 are satisfied. Further, after possibly shrinking η again, we have $3\eta < \varepsilon'/2$. We conclude that, for any multi-index α,

$$\limsup_{t \searrow 0} 4t \log \left(\left| \partial_y^\alpha p_t^\nu(x, U_1^c, y) \right| \right) \leq -\left(d(x,y) + \varepsilon' - 3\eta\right)^2 < -\left(d(x,y) + \frac{\varepsilon'}{2} \right)^2 \tag{18}$$

uniformly over $x \in K_0$ and $y \in B$.

Because U_1 has compact closure, we can isometrically include it in a compact Riemannian manifold M' (say, by a smooth doubling construction), equipped with a measure that extends ν and a conservative vector field that extends Z; we just extend the smooth f in $\nu = e^f \mu$, and the associated heat kernel also remains symmetric. Moreover, we can choose M' such that (17) holds with respect to the distance on M' as well. (In particular, we can use a bump function to conformally rescale the metric outside of U_1 to make paths from x to y that leave U_1 "too long," and for all $(x,y) \in K_1$ by smoothness and compactness.) Then the argument leading to (18) applies to $p_t^{\nu,M'}(x, U_1^c, y)$ as well (and perhaps more easily, since M' is compact). Because $p_t^\nu(x,y) = p_t^{\nu,U_1}(x,y) + p_t^\nu(x, U_1^c, y)$, $p_t^{\nu,M'}(x,y) = p_t^{\nu,U_1}(x,y) + p_t^{\nu,M'}(x, U_1^c, y)$, and $p_t^{\nu,U_1}(x,y)$ is the same whether it's viewed as a subset of M or of M', we conclude that, for any $\delta > 0$ and any multi-index α, there exists $t_0 > 0$, such that

$$\left| \partial_y^\alpha p_t^\nu(x,y) - \partial_y^\alpha p_t^{\nu,M'}(x,y) \right| \leq \exp\left[-\frac{\left(d(x,y) + \frac{\varepsilon'}{2}\right)^2 - \delta}{4t} \right]$$

for all $x \in K_0$, $y \in B$, and $t \in (0, t_0]$. By Lemma 9, for each N, there is some $t_0 > 0$ such that

$$\left|\nabla^N_{V_1,\dots,V_N} p^\nu_t(x,y) - \nabla^N_{V_1,\dots,V_N} p^{\nu,M'}_t(x,y)\right| \le \exp\left[-\frac{\left(d(x,y)+\frac{\varepsilon'}{2}\right)^2 - \delta}{4t}\right] \tag{19}$$

for any collection of unit vector fields $V_1, \dots, V_N$ (acting on y) on B, and any $x \in K_0$, $y \in B$, and $t \in (0, t_0]$ (perhaps after decreasing δ).

Moreover, because M' is compact and $p^{\nu,M'}_t$ is symmetric (so that we can exchange x and y derivatives) we know that $p^{\nu,M'}_t$ satisfies

$$\left|\nabla^N_y p^{\nu,M'}_t(x,y)\right| \le D_N \left[\frac{d(x,y)}{t} + \frac{1}{\sqrt{t}}\right]^N p^{\nu,M'}_t(x,y) \tag{20}$$

for $t \in (0,1]$, uniformly over $(x,y) \in M'$, for some sequence of constants. Since both $p^{\nu,M'}_t$ and p^ν_t satisfy (2) and (3), we see that they also satisfy (19) with $N = 0$ (that is, such a bound holds for the heat kernels themselves, not just their derivatives). It follows that

$$4t \log p^{\nu,M'}_t(x,y) \to -d^2(x,y) \quad \text{and} \quad \frac{p^{\nu,M'}_t(x,y)}{p^\nu_t(x,y)} \to 1 \quad \text{as } t \searrow 0,$$

uniformly for $x \in K_0$ and $y \in B$. In light of this and (19) (recall that δ can be chosen as small as needed), we can replace $p^{\nu,M'}_t$ with p^ν_t on both sides of (20) at the cost of possibly adjusting the D_N. We conclude that, for any N, there is a constant D_N and a $t_0 > 0$ such that

$$\left|\nabla^N_y p^\nu_t(x,y)\right| \le D_N \left[\frac{d(x,y)}{t} + \frac{1}{\sqrt{t}}\right]^N p^\nu_t(x,y) \tag{21}$$

for all $t \in (0, t_0)$, $x \in K_0$, and $y \in B$.

Because x is in the interior of K_0 and B is open, and (x,y) is an arbitrary point in K, we've proven that any point in K has an open neighborhood on which (21) holds. By compactness of K, this means that we can find D_N and t_0 such that (21) holds on all of K. Further, as already noted, we are free to let $t_0 = 1$ by smoothness, at the cost of again adjusting the D_N. Now by symmetry (of both K and the heat kernel with respect to ν) we can replace the x-derivatives by y-derivatives, at which point we use Lemma 6 to replace p^ν_t with p_t (since we no longer need the symmetry) and then again to deduce (5) from (6). This proves the theorem. □

Proof of Theorem 2. As in the proof of Theorem 1, we assume K is symmetric (recall that $\mathrm{Cut}(M)$ is symmetric, so this is compatible with the assumptions), and we construct $K_0 \subset U_0 \subset K_1 \subset U_1 \subset M$ in the same way.

The difference with the proof of Theorem 1 is that now we use bump function to include U_1 in a manifold M' for which the heat kernel satisfies the following expansion: there exists some $t_0 > 0$ and a function $c_0(x,y)$, smooth on a neighborhood of K such that $c_0 > 0$ on K and the heat kernel satisfies the expansion

$$p^{\nu,M'}_t(x,y) = \frac{1}{t^{n/2}} e^{-\frac{d^2(x,y)}{4t}} \left(c_0(x,y) + r(t,x,y)\right) \tag{22}$$

where the remainder $r(t,x,y)$ satisfies the bound

$$\sup_{(t,x,y)\in(0,t_0]\times K} \left|\partial^\alpha_x \partial^\beta_y \partial^k_t r(t,x,y)\right| < \infty$$

for all multi-indices α and β and all $k = 0, 1, 2, \dots$. That this is possible follows from the well-known paper of Ben Arous [2], although one should account for any topological difficulties in the inclusion, as done in [24]. Alternatively, Ludewig [21] proves that (22) holds for compact M' (or even a general complete M'), when one considers the specialization from his vector bundle context to the scalar Laplacian. (Actually, both authors give an expansion to any number of terms, but we only need the leading coefficient c_0 at present.)

Either way, we have such an M' so that (19) holds. Since the right-hand side of (19) is exponentially small compared to the remainder in (22), we have that (22) holds for p^ν at the cost of altering the remainder but still having it satisfy the the same bound. Thus

$$\log p_t^\nu(x,y) = -\frac{n}{2}\log t - \frac{d^2(x,y)}{4t} + \log\left(c_0(x,y) + r(t,x,y)\right).$$

Taking derivatives yields

$$-4t\nabla_y^N \log p_t^\nu(x,y) - \nabla_y^N d^2(x,y) = -4t\nabla_y^N \log\left(c_0(x,y) + r(t,x,y)\right).$$

Now since $c_0 > 0$ and r has uniformly bounded derivatives, $\nabla_y^N \log\left(c_0(x,y) + r(t,x,y)\right)$ is uniformly bounded for $(t,x,y) \in (0,t_0] \times K$ and thus the right-hand side of the above is uniformly $O(t)$. By symmetry, we have that

$$-4t\nabla_x^N \log p_t^\nu(x,y) - \nabla_x^N d^2(x,y) = O(t) \quad \text{as } t \searrow 0$$

uniformly in $(x,y) \in K$. Because

$$\log p_t(x,y) = \log p_t^\nu(x,y) + \log \frac{d\nu}{d\mu}(y),$$

applying $-4t\nabla_x^N$ to both sides gives the theorem.

Finally, in regard to Remark 4, we note that

$$4t\nabla_y^N \log p_t(x,y) - 4t\nabla_y^N \log p_t^\nu(x,y) = 4t\nabla_y^N \log \frac{d\nu}{d\mu}(y).$$

Since $\frac{d\nu}{d\mu}(y)$ is smooth and positive for $y \in \pi_2(K)$, the right-hand side is $O(t)$ uniformly, just as we saw for $\nabla_y^N \log\left(c_0(x,y) + r(t,x,y)\right)$ above. This shows that Theorem 2 holds for either x or y derivatives, as desired. □

Remark 10. As mentioned, weaker versions of these theorems can be proven using results on Riemannian Brownian motion from the 90s. In particular, assume that, in Theorems 1 and 2, $Z \equiv 0$ and K satisfies the stronger assumption that there exists $\varepsilon > 0$ such that, for all $(x,y) \in K$, $d(x,\infty) + d(y,\infty) > d(x,y) + \varepsilon$, where

$$d(x,A) = \inf\{d(x,z) : z \in A\}$$
$$\text{and} \quad d(x,\infty) = \sup\{d(x,A) : A \text{ closed and } M \setminus A \text{ relatively compact}\}.$$

In this case, we can use that (1) itself holds on compact manifolds, along with Lemmas 6 and 9 to move between partial derivatives and covariant derivatives as well as between logarithmic derivatives and "regular" derivatives of p_t, in place of Léandre's result (4). That gives the necessary input for Lemma 7. Moreover, a pair of papers of Hsu [13, 16] can be used in place of the results of Bailleul-Norris (2) and (3), giving the necessary input for Lemma 8 with a larger U_1 than before, reflecting the weaker localization condition for K.

Given this, the proofs of Theorems 1 and 2 proceed in essentially the same way. In Theorem 1 we localize (1) as proven in Hsu [14] and Stroock-Turetsky [26] , with no need to refer to [24] for the case of non-zero Z, and similarly in Theorem 2, we can localize the results of [22] without using the expansion (22) to handle the case of non-zero Z, although at the cost of weakening the $O(t)$ rate of convergence.

5. Sharpness of the upper bound and explicit examples

In [24], an expression for the leading (that is, the $1/t^N$ order) term in the asymptotics of an Nth-order derivative of $\log p_t(x,y)$ is given. Briefly, if Γ is the set of midpoints of minimal geodesics from x to y, there is a probability measure m_0 on Γ such that, for any smooth vector fields $Z_1, \dots, Z^N$ in a neighborhood of y,

$$\lim_{t\searrow 0} t^N Z_y^N \cdots Z_y^1 \log p_t(x,y) = \left(-\frac{d(x,y)}{2}\right)^N \kappa^{m_0}\left(Z_y^1 d(\cdot,y), \dots, Z_y^N d(\cdot,y)\right). \tag{23}$$

where κ^{m_0} is the joint cumulant with respect to m_0 and the $Z^i_y d(\cdot, y)$ are viewed as random variables on Γ. Note that if $(x, y) \notin \mathrm{Cut}(M)$, then Γ is a single point, m_0 is a point mass, and any cumulant of order 2 or more is 0, which is consistent with Theorem 2. However, since this cumulant will generally be non-zero for $(x, y) \in \mathrm{Cut}(M)$, it's clear that the upper bound in (1) is sharp (that is, the constant is not explicitly given, but $1/t^N$ is the correct order of blow up) even on a compact Riemannian manifold (so this is not a matter of completeness or incompleteness). To underscore this, and also to illustrate the pointwise result just mentioned, we treat the simple example of the circle explicitly.

If we consider $\mathbb{S}^1 \cong \mathbb{R}/\mathbb{Z}$ with the standard heat kernel (that is, with $Z \equiv 0$), then for any x and y (understood mod 1), we have

$$p_t^{\mathbb{S}^1}(x, y) = \frac{1}{\sqrt{4\pi t}} \sum_{n=-\infty}^{\infty} e^{-\frac{(y-x+n)^2}{4t}}.$$

Here the symmetry between x and y is explicit, and one can take spatial derivatives on either variable. By translation invariance, without loss of generality we can let $x = 0$, in the sense that it is the image of 0 under the quotient map, in which case the cut locus of x is the point $y = \frac{1}{2}$, again in the sense that it is the image of $\frac{1}{2}$ under the quotient map. More geometrically, y is the antipodal point of x and there are two minimizing geodesics from x to y, corresponding to the two ways to go halfway around the circle.

To compute the log-Hessian of $p_t(0, 1/2)$, let $y = \frac{1}{2} + \alpha$, for small α (which is trivially normal coordinates centered at $\frac{1}{2}$). Then explicit computation gives that,

$$\partial_\alpha p_t^{\mathbb{S}^1}\left(0, \frac{1}{2} + \alpha\right) = \frac{1}{\sqrt{4\pi t}} \sum_{n=-\infty}^{\infty} -\frac{\frac{1}{2} + \alpha + n}{2t} e^{-\frac{\left(\frac{1}{2}+\alpha+n\right)^2}{4t}} \quad \text{and}$$

$$\partial_\alpha^2 p_t^{\mathbb{S}^1}\left(0, \frac{1}{2} + \alpha\right) = \frac{1}{\sqrt{4\pi t}} \sum_{n=-\infty}^{\infty} \left[-\frac{1}{2t} + \frac{\left(\frac{1}{2} + \alpha + n\right)^2}{4t^2}\right] e^{-\frac{\left(\frac{1}{2}+\alpha+n\right)^2}{4t}}.$$

Letting $\alpha = 0$, we see that only two terms from the sum contribute (up to an exponentially small error), corresponding to the two minimizing geodesics. In particular, for any $\varepsilon > 0$, we have

$$p_t^{\mathbb{S}^1}\left(0, \frac{1}{2}\right) = \frac{1}{\sqrt{4\pi t}} \left[2e^{-\frac{1}{16t}}\right] + O\left(e^{-\frac{9-\varepsilon}{16t}}\right),$$

$$\partial_\alpha p_t^{\mathbb{S}^1}\left(0, \frac{1}{2}\right) = 0 \quad \text{and}$$

$$\partial_\alpha^2 p_t^{\mathbb{S}^1}\left(0, \frac{1}{2}\right) = \frac{1}{\sqrt{4\pi t}} \left[-\frac{1}{2t} + \frac{1}{16t^2}\right] 2e^{-\frac{1}{16t}} + O\left(e^{-\frac{9-\varepsilon}{16t}}\right).$$

Then since $\partial_y^2 \log p_t(x, y) = \frac{\partial_y^2 p_t(x,y)}{p_t(x,y)} - \left(\frac{\partial_y p_t(x,y)}{p_t(x,y)}\right)^2$, we have

$$\partial_x^2 \log p_t^{\mathbb{S}^1}\left(0, \frac{1}{2}\right) = \partial_y^2 \log p_t^{\mathbb{S}^1}\left(0, \frac{1}{2}\right) = -\frac{1}{2t} + \frac{1}{16t^2} + O\left(e^{-\frac{8-\varepsilon}{16t}}\right)$$

for any $\varepsilon > 0$.

We note that here we have computed the logarithmic Hessian of the heat kernel directly (that is, by elementary methods, independent of any of the preceding results of this paper). In reference to (23), note that, by symmetry, m_0 assigns probability 1/2 each to the 2 midpoints of the 2 minimizing geodesics from 0 to 1/2, and thus the random variable $\partial_\alpha d\left(\cdot, \frac{1}{2}\right)$ under m_0 takes values ± 1 each with probability 1/2. So the variance (note that the joint cumulant of 2 random variables is their covariance) in (23) is 1, and we have

$$\lim_{t \searrow 0} t^2 \nabla_x^2 \log p_t^{\mathbb{S}^1}\left(0, \frac{1}{2}\right) = \frac{d^2\left(0, \frac{1}{2}\right)}{4} = \frac{1}{16}$$

using $\partial_\alpha \otimes \partial_\alpha$ as a basis for bilinear forms on $T_x M$, in agreement with the above.

Also, note that this is not special to this simple example. Any time (on any Riemannian manifold and for any x and y points such that the conditions of Theorem 1 hold) there are exactly 2 non-conjugate minimizing geodesics between x and y, one will have similar behavior for the logarithmic Hessian.

6. The non-symmetric and sub-Riemannian cases

Now suppose Z is such that Δ no longer admits a symmetrizing measure. Then y-derivatives of the heat kernel can still be localized, and thus are seen to satisfy the same estimates as above, under either the condition on the distance to infinity in Theorem 1 (if the asymmetry is "not too bad" in the sense of the sector condition of [1]) or the condition described in Remark 10 (in general). However, without the symmetry, y-derivatives cannot be transferred to x-derivatives. Indeed, when exchanging the role of the variables in general, one needs to take the adjoint of the operator (this is the relationship between the forward and backward Kolmogorov equations). The point is that the adjoint may have a potential term (that is, 0th-order term), which is not covered by the preceding arguments and is less easy to localize.

To quickly illustrate the point, we let (u, v) be Cartesian coordinates on $\mathbb{R}^2$. If we consider $\Delta = \partial_u^2 + \partial_v^2 + a(u, v)\partial_u + b(u, v)\partial_v$ then the (formal) adjoint, with respect to the standard Euclidean area, is $\Delta^* = \partial_u^2 + \partial_v^2 - (u, v)\partial_u - b(x, y)\partial_v - (\partial_u a)(u, v) - (\partial_v b)(u, v)$. Thus we now have a potential, and to find a stochastic representation for the heat kernel, corresponding to Δ^*, we should look at the Feynman-Kac formula

$$p_t(x, y) = \mathbb{E}^x \left[\exp \left(\int_0^t (\partial_u a + \partial_v b)(X_s)\, ds \right) \right] \tag{24}$$

where X_t is the diffusion with generator $\partial_u^2 + \partial_v^2 - (u, v)\partial_u - b(x, y)\partial_v$ started from x. On a non-compact manifold, the integrand on the right-hand side need not be bounded, so even if the probability of a path leaving some open U and eventually arriving at y can be estimated to be small, as above, the contribution of such a path to p_t is weighted by the Feynman-Kac factor, which depends on the entire trajectory and which we have no a priori control over. Similarly, trying to treat the potential via an h-transform runs into the problem that it's not clear in general why the heat equation for the adjoint should admit a positive solution. Indeed, it is not clear that we should expect (1) to hold on a non-compact Riemannian manifold without some additional assumptions to control the potential. In that spirit, one can see [20] for recent results in this direction on complete (but not necessarily compact) Riemannian manifolds under additional geometric assumptions (for example, curvature bounds, the existence of a pole, etc.). (Note also that the case where the path integral in (24) only depends on the endpoints and thus can be localized is exactly the case when Z admits a potential, and this is directly related to the density of the symmetrizing measure.)

In a different direction, suppose $p_t(x, y)$ is the heat kernel associated to a sub-Laplacian on a sub-Riemannian manifold M. Then the localization for p_t and its y-derivatives still works (and thus under a symmetry assumption, it also works for x-derivatives), but we don't have a global estimate to localize. That is, no estimates comparable to (1) are known to hold for compact sub-Riemannian manifolds (or for structures on $\mathbb{R}^n$ given by bounded vector fields with bounded derivatives of all orders, as in the work of Léandre and Ben Arous used above). More precisely, such bounds are known to hold away from abnormal minimizers (see [24]), but on a properly sub-Riemannian manifold (that is, not one that happens to be Riemannian), the diagonal is always abnormal, and thus one does not have appropriate estimates on any K that intersects the diagonal. (Note that more than the asymptotic results of [6] are needed in the present context.) Derivative bounds for the special case of H-type groups were established in [8] and [18], including bounds for the log-gradient of the heat kernel in [18], but this comes nowhere near the complete generality we have in the Riemannian situation.

That said, note that the semi-martingale property for the bridge process on compact, step-2 sub-Riemannian manifolds was established in [19], so there is hope for future advances.

7. Acknowledgements

This work was partially supported by a grant from the Simons Foundation (#524713 to Robert Neel).

References

[1] Ismael Bailleul and James Norris, *Diffusion in small time in incomplete sub-Riemannian manifolds*, Anal. PDE **15** (2022), no. 1, 63–84. MR 4395153

[2] G. Ben Arous, *Développement asymptotique du noyau de la chaleur hypoelliptique hors du cut-locus*, Ann. Sci. École Norm. Sup. (4) **21** (1988), no. 3, 307–331. MR MR974408 (89k:60087)

[3] Nicolas Boumal, *An introduction to optimization on smooth manifolds*, To appear with Cambridge University Press, Jun 2022.

[4] Isaac Chavel, *Riemannian geometry*, second ed., Cambridge Studies in Advanced Mathematics, vol. 98, Cambridge University Press, Cambridge, 2006, A modern introduction. MR 2229062

[5] Xin Chen, Xue-Mei Li, and Bo Wu, *Logarithmic heat kernels: estimates without curvature restrictions*, preprint, arXiv:2106.02746 (2021).

[6] Yves Colin de Verdière, Luc Hillairet, and Emmanuel Trélat, *Small-time asymptotics of hypoelliptic heat kernels near the diagonal, nilpotentization and related results*, Ann. H. Lebesgue **4** (2021), 897–971. MR 4315774

[7] Jozef Dodziuk, *Maximum principle for parabolic inequalities and the heat flow on open manifolds*, Indiana Univ. Math. J. **32** (1983), no. 5, 703–716. MR 711862

[8] Nathaniel Eldredge, *Precise estimates for the subelliptic heat kernel on H-type groups*, J. Math. Pures Appl. (9) **92** (2009), no. 1, 52–85. MR 2541147

[9] Fu-Zhou Gong and Zhi-Ming Ma, *The log-Sobolev inequality on loop space over a compact Riemannian manifold*, J. Funct. Anal. **157** (1998), no. 2, 599–623. MR 1638269

[10] Alexander Grigor'yan, *Heat kernels on weighted manifolds and applications*, The ubiquitous heat kernel, Contemp. Math., vol. 398, Amer. Math. Soc., Providence, RI, 2006, pp. 93–191. MR 2218016

[11] Batu Güneysu, *On the semimartingale property of Brownian bridges on complete manifolds*, ALEA Lat. Am. J. Probab. Math. Stat. **16** (2019), no. 1, 15–31. MR 3903023

[12] Elton Hsu and Zhenan Wang, *Bismut's gradient formula for vector bundles*, preprint, arXiv:1601.02285 (2016).

[13] Elton P. Hsu, *On the principle of not feeling the boundary for diffusion processes*, J. London Math. Soc. (2) **51** (1995), no. 2, 373–382. MR 1325580

[14] ———, *Estimates of derivatives of the heat kernel on a compact Riemannian manifold*, Proc. Amer. Math. Soc. **127** (1999), no. 12, 3739–3744. MR 1618694

[15] ———, *Stochastic analysis on manifolds*, Graduate Studies in Mathematics, vol. 38, American Mathematical Society, Providence, RI, 2002. MR 1882015

[16] Pei Hsu, *Heat kernel on noncomplete manifolds*, Indiana Univ. Math. J. **39** (1990), no. 2, 431–442. MR 1089046

[17] Rémi Léandre, *Majoration en temps petit de la densité d'une diffusion dégénérée*, Probab. Theory Related Fields **74** (1987), no. 2, 289–294. MR 871256 (88c:60144)

[18] Hong-Quan Li, *Estimations optimales du noyau de la chaleur sur les groupes de type Heisenberg*, J. Reine Angew. Math. **646** (2010), 195–233. MR 2719560

[19] Xue-Mei Li, *On hypoelliptic bridge*, Electron. Commun. Probab. **21** (2016), Paper No. 24, 12. MR 3485393

[20] ———, *Hessian formulas and estimates for parabolic Schrödinger operators*, J. Stoch. Anal. **2** (2021), no. 3, Art. 7, 53. MR 4304478

[21] Matthias Ludewig, *Strong short-time asymptotics and convolution approximation of the heat kernel*, Ann. Global Anal. Geom. **55** (2019), no. 2, 371–394. MR 3923544

[22] Paul Malliavin and Daniel W. Stroock, *Short time behavior of the heat kernel and its logarithmic derivatives*, J. Differential Geom. **44** (1996), no. 3, 550–570. MR 1431005

[23] H. P. McKean, Jr., *Stochastic integrals*, Probability and Mathematical Statistics, No. 5, Academic Press, New York-London, 1969. MR 0247684

[24] Robert Neel and Ludovic Sacchelli, *Uniform, localized asymptotics for sub-riemannian heat kernels and diffusions*, preprint, arXiv:2012.12888 (2020).

[25] Shuenn Jyi Sheu, *Some estimates of the transition density of a nondegenerate diffusion Markov process*, Ann. Probab. **19** (1991), no. 2, 538–561. MR 1106275

[26] Daniel W. Stroock and James Turetsky, *Upper bounds on derivatives of the logarithm of the heat kernel*, Comm. Anal. Geom. **6** (1998), no. 4, 669–685. MR 1664888

Department of Mathematics, Lehigh University, Bethlehem, Pennsylvania, USA
E-mail address: robert.neel@lehigh.edu

Université Côte d'Azur, Inria, CNRS, LJAD, France
E-mail address: ludovic.sacchelli@inria.fr