Heat flow regularity, Bismut’s derivative formula, and pathwise Brownian couplings on Riemannian manifolds with Dynkin bounded Ricci curvature

Mathias Braun¹  Batu Güneysu²

January 28, 2020

Abstract We prove that if the Ricci curvature of a geodesically complete Riemannian manifold $X$, endowed with the Riemannian distance $\rho$ and the Riemannian volume measure $m$, is bounded from below by a Dynkin decomposable function $k: X \to \mathbb{R}$, then $X$ is stochastically complete. This assumption on $k$ is satisfied if its negative part $k^-$ belongs to the Kato class of $X$. In addition, given $f \in L^p(X)$ for sufficiently large $p$ in a range depending only on $k^-$, we derive a global Bismut derivative formula for $\nabla P_t f$ for every $t > 0$ along the heat flow $(P_t)_{t \geq 0}$, whose $L^\infty$-Lip-regularization we obtain as a corollary.

Moreover, for such functions $k$, we show that the Ricci curvature of $X$ is bounded from below by $k$ if and only if $X$ supports the $L^1$-gradient estimate w.r.t. $k$, i.e. for every $f \in W^{1,2}(X)$ and $t \geq 0$,

$$|\nabla P_t f| \leq P_t^k |\nabla f|$$

holds $m$-a.e., $(P_t^k)_{t \geq 0}$ being the Schrödinger semigroup in $L^2(X)$ generated by $-(\Delta - k)/2$.

If $k$ is additionally lower semicontinuous, another equivalent characterization of lower Ricci bounds by $k$ is proven to be the existence, given any $x, y \in X$, of a Markovian coupling $(b^x_t, b^y_t)$ of Brownian motions on $X$ starting in $(x, y)$ such that a.s. for every $s, t \in [0, \infty)$ with $s \leq t$, one has the pathwise estimate

$$\rho(b^x_s, b^y_t) \leq e^{-\int_s^t k(b^x_r, b^y_r)/2 \, dr} \rho(b^x_s, b^y_t)$$

involving the “average” $b^k(u, v) := \inf_{\gamma} \int_0^1 k(\gamma_r) \, dr$ of the function $k$ along all minimizing geodesics $\gamma: [0, 1] \to X$ from $u$ to $v$.

Contents

1 Introduction 2
   1.1 Dynkin and Kato decomposable functions ................................. 2
   1.2 Regularity of the heat flow and Bismut’s derivative formula ............. 3
   1.3 Characterizations of Dynkin decomposable lower Ricci bounds .......... 4

¹University of Bonn, Institute for Applied Mathematics, Endenicher Allee 60, 53115 Bonn, Germany, braun@iam.uni-bonn.de. Funded by the European Union through the ERC-AdG “RicciBounds”.

²Humboldt-Universität zu Berlin, Institut für Mathematik, Rudower Chaussee 25, 12489 Berlin, Germany, gueneysu@math.hu-berlin.de.

Acknowledgments. The authors’ collaboration arose from a discussion at the Japanese-German Open Conference on Stochastic Analysis at the University of Fukuoka, Japan, in September 2019. The authors gratefully acknowledge the warm hospitality of this institution.
1 Introduction

Let $(X, g)$ be a smooth, geodesically complete, noncompact, connected Riemannian manifold without boundary. The metric tensor $g$, subsequently abbreviated by $\langle \cdot, \cdot \rangle := g(\cdot, \cdot)$, induces the Riemannian distance function $\rho$ and the Riemannian volume measure $m$. With the usual abuse of notation, the fiberwise norm both on $TX$ and $T^*X$ is $|\cdot| := \langle \cdot, \cdot \rangle^{1/2}$. Without further notice, all mentioned functions and sections are assumed to be real-valued, and all appearing Lebesgue and Sobolev spaces are understood w.r.t. $m$. Let $\nabla$ be the Levi-Civita connection on $X$, giving rise to the Laplace–Beltrami operator $\Delta$.

This paper is devoted to the study of both analytic and probabilistic consequences of the Ricci curvature of $X$ being $m$-a.e. bounded from below by a Borel function $k: X \to \mathbb{R}$ for which $k^-/2$, where $k^-(x) := -\min\{k(x), 0\}$, belongs to the contractive Dynkin class $D(X)$, and equivalent formulations of such bounds on the Ricci tensor $\text{Ric}$. In this context, given any Borel subset $U \subset X$, the statement “$\text{Ric} \geq k$ on $U$” shall always be understood as

$$\text{Ric}(x)(\xi, \xi) \geq k(x) |\xi|^2$$

for $m$-a.e. $x \in U$ and every $\xi \in T_xM$.

To describe our main results, for the moment we assume the reader to be familiar with basic objects such as heat flow, Brownian motion, and Schrödinger semigroups. A detailed account on these as well as basic definitions and notations are collected in Chapter 2.

1.1 Dynkin and Kato decomposable functions

Let us denote by $p \in C^\infty((0, \infty) \times X \times X; (0, \infty))$ the minimal heat kernel on $X$, i.e., the smallest positive fundamental solution to the heat operator $\partial/\partial t - \Delta/2$ [Gri09].

For a historical account on the following definition, we refer to [Gün17, Stu94] as well as the references therein.

Definition 1.1.  (i) The contractive Dynkin class $\mathcal{D}(X)$ of $X$ is the set of all Borel functions $v: X \to \mathbb{R}$ for which there exists $t > 0$ such that

$$\sup_{x \in X} \int_0^t \int_X p_r(x, y) |v(y)| \, dm(y) \, dr < 1.$$
The Kato class $\mathcal{K}(X)$ of $X$ is the set of all Borel functions $v : X \to \mathbb{R}$ such that
\[
\lim_{t \downarrow 0} \sup_{x \in X} \int_{0}^{t} \int_{X} p_{s}(x,y) |v(y)| \, dm(y) \, ds = 0.
\]

(iii) A function $k : X \to \mathbb{R}$ is called Dynkin decomposable (or Kato decomposable, respectively) if it belongs to $L_{\text{loc}}^{1}(X)$ and $v := k^{-1}/2$ belongs to $\mathcal{D}(X)$ (or $\mathcal{K}(X)$, respectively).

In view of
\[
\int_{X} p_{r}(x,y) \, dm(y) \leq 1 \quad \text{for every } x \in X, \ r > 0,
\]
it follows that $L^{\infty}(X) \subset \mathcal{K}(X)$, while $\mathcal{K}(X) \subset \mathcal{D}(X)$ is trivially true, the latter however being a strict inclusion in general [CK09, Example 4.3].

**Remark 1.2.** While $\mathcal{K}(X)$ is a vector space, at the more general Dynkin level it is only true that for every $v \in \mathcal{D}(X)$ there exists $\varepsilon > 0$ such that $\lambda v \in \mathcal{D}(X)$ for every $\lambda \in [-1-\varepsilon,1+\varepsilon]$. The latter, nevertheless, already allows for many $L^{p}$-properties of the associated Schrödinger semigroup and objects related to it, see Theorem 1.4 and Theorem 2.2.

It is important to stress the factor $1/2$ in (iii) which is redundant to demand for Kato decomposability. In our definition of Dynkin decomposable functions, however, it is crucial and mainly introduced for convenience because we will define the heat semigroup and the Schrödinger semigroup w.r.t. $k$ via the generators $-\Delta/2$ and $-(\Delta - k)/2$, respectively.

Both Dynkin and Kato functions have been studied in great detail in the literature, see [AS82, Gim17, SV96] and the references therein. In connection with lower Ricci bounds, Kato functions arise very naturally in different frameworks, see e.g. [GvR19] (which treats molecular Schrödinger operators probabilistically). On the other hand, many results derived from the Kato property do actually only require Dynkin regularity [Gim17]. This motivates the study of basic properties of (smooth) spaces, and objects associated to it, admitting Dynkin decomposable lower bounds on Ric in greater generality. Owing to this enterprise, we regard this article as an initiating point.

### 1.2 Regularity of the heat flow and Bismut’s derivative formula

Since Dynkin and Kato functions are defined in terms of the heat kernel on $X$, one expects that bounds on the Ricci curvature affect the regularity of the associated heat flow or heat semigroup $(P_{t})_{t \geq 0}$ in $L^{2}(X)$ given by $P_{t} := e^{t\Delta/2}$ and yield properties of another object closely connected to $(P_{t})_{t \geq 0}$, namely Brownian motion $b^{x} : [0,\infty) \times \Omega \to X$ on $X$ started at $x \in X$ with explosion time $\zeta^{x}$, defined on a suitable filtered probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

We first prove that if $\text{Ric} \geq k$ on $X$ with some Dynkin decomposable function $k$, then $X$ is stochastically complete, i.e. Brownian motion never leaves $X$ a.s. We also obtain Lipschitz continuity of $P_{t}f$ for every $f \in L^{\infty}(X)$ and $t > 0$ with explicit bounds in terms of $k$ and the $L^{\infty}$-norm of $f$. This regularity is derived through a Bismut formula for $\nabla P_{t}$, $t > 0$. To formulate it, given any $x \in X$, let $\|x\| : [0,\zeta^{x}) \times \Omega \to (b^{x})^{*}(TX)$, i.e. $\|x\| : T_{x}X \to T_{b^{x}X}$ for all $t \in [0,\zeta^{x})$, denote the stochastic parallel transport w.r.t. $\nabla$ along the sample paths of $b^{x}$, let the process $Q^{x} : [0,\zeta^{x}) \times \Omega \to \text{End}(T_{x}X)$ be defined as the unique solution to the pathwise ordinary differential equation
\[
\frac{d}{ds} Q^{x}_{s} = -\frac{1}{2} Q^{x}_{s}(\|x\|)^{-1} \text{Ric}_{b^{x}} \|x\| ds, \quad Q^{x}_{0} = \text{Id}_{T_{x}X}, \tag{1.1}
\]
and let \( w: [0, \zeta] \times \Omega \to T_x X \) denote the anti-development of \( b^x \), a canonically given Brownian motion on the Euclidean space \( T_x X \) (see [Hsu02] for precise definitions).

**Theorem 1.3.** Let \( k: X \to \mathbb{R} \) be a Dynkin decomposable function and assume that \( \text{Ric} \geq k \) on \( X \). Then

(i) \( X \) is stochastically complete, i.e.

\[
\mathbb{P}[\zeta^x = \infty] = 1 \quad \text{for every } x \in X,
\]

(ii) for \( p = \infty \) as well as for every \( p \in (1, \infty) \) for which \( pk/(p-1) \) is still Dynkin decomposable, every \( t > 0 \), every \( f \in L^p(X) \), and for every smooth, compactly supported vector field \( V \) on \( X \), we have Bismut’s derivative formula

\[
\langle \nabla P_t f(x), V(x) \rangle = \frac{1}{t} \mathbb{E} \left[ f(b^x_t) \int_0^t \langle Q_s^x V(x), d\mathbf{w}^x_s \rangle \right] \quad \text{for m-a.e. } x \in X,
\]

where the stochastic integral inside the last expectation is understood in Itô’s sense, and

(iii) for every \( t > 0 \), one has the \( L^\infty \)-Lip-regularization property \( P_t: L^\infty(X) \to \text{Lip}(X) \) with

\[
\text{Lip}(P_t f) \leq \sqrt{8} t^{-1/2} \sup_{x \in X} \mathbb{E} \left[ e^{\int_0^t (k^x - (b^x_t)^2)/2 dr} \right] \|f\|_{L^\infty} \quad \text{for every } f \in L^\infty(X),
\]

where \( \text{Lip}(P_t f) \) is the Lipschitz constant of \( P_t f \) w.r.t. \( \rho \).

In particular, if \( k \) is Kato decomposable, then (ii) holds true for every \( p \in (1, \infty] \).

The proof of (i) be found in Section 3.1 while (ii) and (iii) are studied in Section 3.2.

Each of (i), (ii) and (iii) is known in the restrictive case when the Ricci curvature of \( X \) is bounded from below by a constant number \( K \in \mathbb{R} \). Stochastic completeness in this framework is a classical result by [Yau78]. (Note that a priori, there is no reason to expect stochastic completeness under a Dynkin decomposable lower bound by Grigor’yan’s volume test [Gri09].)

The \( L^\infty \)-Lip-regularization of the heat flow is known in the more general abstract setting of \( \text{RCD}(K, \infty) \) spaces [AGS15]. The Bismut derivative formula has first appeared in [Bis84] in the compact case, and has been extended later on to the case of uniform Ricci bounds [DT01, EL94]. In our situation of Dynkin decomposable lower bounds, technical difficulties arise simply from the fact that, a priori, it is not clear at all that the right hand side of the formula converges. Remarkably, localized versions of the Bismut formula hold without any assumptions on the geometry of the manifold, see e.g. [Tha97, TW98, TW11].

### 1.3 Characterizations of Dynkin decomposable lower Ricci bounds

Our second goal is to provide several equivalent characterizations of lower Ricci bounds by Dynkin decomposable functions \( k \), which we shortly introduce.

The perhaps closest equivalent characterization of \( \text{Ric} \geq k \) on \( X \) is the \( L^1 \)-Bochner inequality, which is related to the Ricci curvature of \( X \) by the well-known Bochner formula

\[
\Delta \frac{|
abla f|^2}{2} = \langle \nabla \Delta f, \nabla f \rangle + |\text{Hess } f|^2 + \text{Ric}(\nabla f, \nabla f) \quad \text{for every } f \in C^\infty(X).
\]
In turn, this is equivalent to certain gradient estimates for \((P_t)_{t \geq 0}\), for which we denote by \((P^k_t)_{t \geq 0}\) the \textit{Schrödinger semigroup} with generator \(-(\Delta - k)/2\), i.e. \(P^k_t := e^{t(\Delta-k)/2}\). This is a strongly continuous semigroup of bounded linear operators in \(L^2(X)\) which is generated by the self-adjoint, lower semibounded Schrödinger operator \(-(\Delta - k)/2\). The latter is defined as the sum of the quadratic form induced by \(-\Delta/2\) and the quadratic form stemming from the maximally defined multiplication operator induced by \(k/2\). As \(k\) is Dynkin decomposable \([SV96, \text{Gün17}]\), \((P^k_t)_{t \geq 0}\) has a pointwise well-defined version that can be expressed via Brownian motion \(b^x\) on \(X\) through the Feynman–Kac formula

\[
P^k_t f(x) = \mathbb{E} \left[ e^{-\int_0^t k(b^x_r)/2 \, dr} f(b^x_t) \mathbbm{1}_{\{t < c^x\}} \right] \quad \text{for every } f \in L^2(X), \, x \in X, \, t \geq 0. \quad (1.4)
\]

This semigroup extends to a strongly continuous semigroup of bounded operators in \(L^p(X)\) for all \(p \in [1, \infty)\), see Theorem \(2.2\).

**Theorem 1.4.** Let \(k : X \to \mathbb{R}\) be a Dynkin decomposable function. Then the following conditions are equivalent:

(i) we have \(\text{Ric} \geq k\) on \(X\),

(ii) the \(L^1\)-Bochner inequality w.r.t. \(k\) is satisfied, i.e.

\[
\Delta|\nabla f| - |\nabla f|^{-1} \langle \nabla \Delta f, \nabla f \rangle \geq k |\nabla f| \quad \text{m-a.e. on } \{|\nabla f| \neq 0\}
\]

for every \(f \in C^\infty_c(X)\), \(\text{(1.5)}\)

(iii) the \(L^1\)-gradient estimate w.r.t. \(k\) holds, i.e.

\[
|\nabla P_t f| \leq P^k_t |\nabla f| \quad \text{m-a.e. for every } f \in W^{1,2}(X), \, t \geq 0. \quad (1.6)
\]

The key implication “(i) \(\implies\) (iii)” has already been derived very carefully in \([\text{Gün17}]\), see Theorem \(2.3\) below, using the concept of Kato–Simon inequalities for covariant Schrödinger semigroups and the Weitzenböck formula. With Theorem \(1.4\), we treat the converse implication, Sections \(4.1\) and \(4.2\) being devoted to the proofs of “(iii) \(\implies\) (ii)” and “(ii) \(\implies\) (i)”, respectively. In the more singular, however slightly different framework of so-called \textit{tamed spaces}, i.e. metric measure spaces admitting a synthetic notion of lower Ricci bounds by a “Kato distribution”, an adapted version of Theorem \(1.4\) is proven in a forthcoming joint work of M. Erbar, C. Rigoni, K.-T. Sturm, and L. Tamanini.

If we additionally assume lower semicontinuity on \(k\), we can also derive a one-to-one connection between lower Ricci bounds by \(k\) and the existence of certain couplings of Brownian motions on \(X\). Here, given \(x, y \in X\), by a \textit{coupling of Brownian motions starting in} \((x, y)\), we understand an \(X \times X\)-valued stochastic process \((b^x, b^y) : [0, \infty) \times \Omega \to X \times X\) on some filtered probability space \((\Omega, \mathcal{F}, \mathbb{P})\) such that \(b^x\) and \(b^y\) are Brownian motions on \(X\) starting in \(x\) and \(y\), respectively. (It will not be necessary to take into account a possible explosion in finite time.) To formulate an appropriate pathwise coupling estimate, let us denote by \(\text{Geo}(X)\) the set of minimizing geodesics \(\gamma : [0, 1] \to X\), and define the lower semicontinuous function \(k : X \times X \to \mathbb{R}\) by

\[
k(u, v) := \inf \left\{ \int_0^1 k(\gamma_r) \, dr : \gamma \in \text{Geo}(X), \, \gamma_0 = u, \, \gamma_1 = v \right\}.
\]
Theorem 1.5. Let $k: X \to \mathbb{R}$ be a Dynkin decomposable and lower semicontinuous function. Then any of the conditions (i), (ii) and (iii) in Theorem 4.1 holds if and only if $X$ obeys the pathwise coupling property w.r.t. $k$, i.e. $X$ is stochastically complete and for every $x,y \in X$, there exists a coupling $(b^x,b^y)$ of Brownian motions on $X$ starting in $(x,y)$ with the Markov property such that a.s., we have

$$\rho(b_t^x,b_t^y) \leq e^{-\int_0^t \frac{\|b_t^x-b_t^y\|^2}{2} dr} \rho(b_0^x,b_0^y) \quad \text{for every } s,t \in [0,\infty) \text{ with } s \leq t,$$

Here and in the sequel, the Markov property for every respective process under consideration is understood w.r.t. its canonically given filtration. The statement of Theorem 1.5 is still true if one does not require the Markov property of $(b^x,b^y)$.

Theorem 1.5 (Chapter 5 being devoted to its proof) as well as the previous Theorem 1.4 extend similar results from [BHS19], where analogous equivalences have been established in the synthetic framework of CD($k,\infty$) spaces with lower semicontinuous, lower bounded variable Ricci bounds $k: X \to \mathbb{R}$ (see also [Stu13]). Let us point out that, in contrast to [BHS19], the coupling technique on manifolds does not require any notion of “Wasserstein contractivity” for the dual heat flow to $(P_t)_{t \geq 0}$ on the space of Borel probability measures on $X$. It is rather provided in a direct way by the Cranston–Kendall coupling method [Cra91, Ken86].

2 Preliminaries

For more details on the facts collected in this chapter, we refer the reader to [Gri09, Gün17, Hsu02, Ros97] and the references therein.

Heat flow on functions By geodesic completeness, the Laplace–Beltrami operator, initially viewed as being defined on $C_c^{\infty}(X)$, has a unique self-adjoint extension in $L^2(X)$ which will again be denoted by $\Delta$. Then the operator $-\Delta/2$ is the generator of the strongly local, regular Dirichlet form $\mathcal{E}: L^2(X) \to [0,\infty]$ given by

$$\mathcal{E}(f) := \frac{1}{2} \int_X |\nabla f|^2 \, dm \quad \text{if } f \in W^{1,2}(X), \quad \mathcal{E}(f) := \infty \quad \text{otherwise.}$$

The heat semigroup or heat flow $(P_t)_{t \geq 0}$ defined in Section 1.2 is directly linked to $\mathcal{E}$ and is a strongly continuous, positivity preserving contraction semigroup of linear operators in $L^2(M)$. For every $t > 0$ and $f \in L^2(X)$, the heat kernel $p$ on $X$ introduced in Section 1.1 gives rise to an integral kernel representing (a version of) $P_t f$ through

$$P_tf(x) := \int_X p_t(x,y) f(y) \, dm(y) \quad \text{for every } x \in X. \quad (2.1)$$

Actually, $(P_t)_{t \geq 0}$ extends to a contraction semigroup of linear operators from $L^p(X)$ into $L^p(X)$ for every $p \in [1,\infty]$ which is strongly continuous on $[0,\infty)$ if $p < \infty$. Moreover, $(2.1)$ is still valid for every $p \in [1,\infty]$ and $f \in L^p(X)$, and for such $f$, $P_tf \in C^{\infty}((0,\infty) \times X)$ solves the heat equation

$$\frac{\partial}{\partial t} P_t f = \frac{1}{2} \Delta P_t f \quad \text{in } (0,\infty) \times X \quad (2.2)$$

with initial condition $f$. In addition, we have $P_tf \in C^{\infty}((0,\infty) \times X)$ if $f$ is also smooth.
Brownian motion Given a locally compact Polish space $Y$ we denote by $C([0, \infty); Y)$ the space of continuous maps $\gamma : [0, \infty) \to Y$, equipped with the topology of locally uniform convergence and the induced Borel $\sigma$-algebra. Let $Y_\partial := Y \cup \{\partial\}$ denote the (essentially unique) one-point compactification of $Y$.

Given a point $x \in X$, any stochastic process $b$ with sample paths in $C([0, \infty); X_\partial)$ which is defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ (i.e. the map $t \mapsto b_t(\omega)$ belongs to $C([0, \infty); X_\partial)$ for all $\omega \in \Omega$) is termed Brownian motion on $X$ starting in $x$ if its law equals the Wiener measure $\mathbb{P}_x$ on $C([0, \infty); X_\partial)$ concentrated at paths starting in $x$. (Usually we want to underline the dependency of $b$ from its starting point $x$, whence we shall often write $b^x$.) Recall that $\mathbb{P}_x$ is the uniquely determined probability measure on $C([0, \infty); X_\partial)$ with $(\text{ev}_0)_x b^x \mathbb{P}_x = \delta_x$ (where $\text{ev}_0(\gamma) := \gamma_0$ is the evaluation map at $0$) and whose transition density is given by the function $q : (0, \infty) \times X_\partial \times X_\partial \to [0, \infty)$ defined by setting, for every $y, y' \in X$,

$$q_t(y, y') = p_t(y, y') \quad q_t(\partial, y') := 0, \quad q_t(\partial, \partial) := 1, \quad q_t(y, \partial) := 1 - \int_X p_t(y, z) \, \text{dm}(z).$$

Now let $\zeta^x := \inf\{t \geq 0 : b^x_t = \partial\}$ denote the explosion time of $b^x$, with the usual convention that $\inf\emptyset := \infty$. Since the Wiener measure is concentrated on paths having $\partial$ as a trap, for every $p \in [1, \infty]$ one has

$$\mathbb{P}_t f(x) = E[f(b^x_t) 1_{\{t < \zeta^x\}}] \quad \text{for every } f \in L^p(X), \ x \in X, \ t \geq 0. \quad (2.3)$$

Therefore, stochastic completeness of $X$ as formulated in $[1, 2]$ is equivalent to

$$\mathbb{P}[t < \zeta^x] = \mathbb{P}_t 1_X(x) = \int_X p_t(x, y) \, \text{dm}(y) = 1 \quad \text{for every } x \in X, \ t > 0. \quad (2.4)$$

If $(\Omega, \mathcal{F}, \mathbb{P})$ is filtered and $b^x$ adapted to the given filtration, then $b^x$ is called an adapted Brownian motion. In this case, $b^x$ is a semimartingale on $X$ in the sense that for every $f \in C^\infty(X)$, the real-valued process $f \circ b^x$ is a semimartingale up to the explosion time $\zeta^x$. The stochastic parallel transport along $b^x$ w.r.t. $\nabla$ started in $x \in X$ constitutes a process $\parallel^{b^x} : [0, \zeta^x) \times \Omega \to (b^x)^s(TX)$, the latter being the pullback bundle of $TX$ along the paths of $b^x$, such that $\parallel^{b^x}_t : T_x X \to T_{b^x_t} X$ is a.s. orthogonal for every $t \in [0, \zeta^x]$.

$L^p$-properties of the Schrödinger semigroup The particular role of Dynkin (decomposable) potentials is due to the effect they have on the behavior of the associated Schrödinger semigroup. They provide uniform spacial bounds for the exponential part of the associated Schrödinger operators at every time $t \geq 0$, which is known as (a consequence of) Khasminskii’s lemma and relies on the Markov property of Brownian motion on $X$.

Lemma 2.1 [Gün17] Lemma VI.8. Let $v \in \mathcal{D}(X)$. Then there exist finite constants $C_1, C_2 \geq 0$ depending only on $||v||$ such that

$$\sup_{x \in X} E\left[ e^{\int_0^t |v(b^x_s)| \, ds} 1_{\{t < \zeta^x\}} \right] \leq C_1 e^{C_2 t} \quad \text{for every } t \geq 0.$$

Theorem 2.2. Let $k : X \to \mathbb{R}$ be a Dynkin decomposable function. Then there exist finite constants $C_1, C_2 \geq 0$ depending only on $k^-$ such that, for every $p \in [1, \infty]$, we have

$$\|P_t f\|_{L^p} \leq C_1 e^{C_2 t} \|f\|_{L^p} \quad \text{for every } f \in L^2(X) \cap L^p(X), \ t \geq 0. \quad (2.5)$$
In particular, for every $p \in [1, \infty]$, $(P^k_t)_{t \geq 0}$ extends to a semigroup of bounded operators from $L^p(X)$ into $L^p(X)$ which indeed satisfies (2.5) for every $f \in L^p(X)$ and, if $p < \infty$, is strongly continuous.

Proof. The idea to prove (2.5) is to use Feynman–Kac’s formula (1.4) together with Lemma 2.1 to show the desired inequality in the cases $p = \infty$ and $p = 1$ (which needs an additional, but elementary exhaustion argument) and to apply Riesz–Thorin’s theorem to extend it to all exponents $p \in [1, \infty]$. See Günther, Theorem IX.2, Corollary IX.4 for details.

The existence of an extension of $(P^k_t)_{t \geq 0}$ to a semigroup of bounded operators from $L^p(X)$ into $L^p(X)$ for every $p \in [1, \infty]$ still satisfying (2.5) is standard by approximation, but we include the argument for the convenience of the reader since similar arguments will appear later (see the proofs of Lemma 3.3 and of (ii) in Theorem 1.3). For $p < \infty$ and $f \in L^p(X)$, for any sequence $(f_n)_{n \in \mathbb{N}}$ in $L^2(X) \cap L^p(X)$ converging to $f$ in $L^p(X)$, (2.5) implies that $(P^k_t f_n)_{n \in \mathbb{N}}$ is a Cauchy sequence in $L^p(X)$. Thus, we define $P^k_t f$ as the $L^p$-limit of the latter sequence as $n \to \infty$, and (2.5) again shows that this definition is independent of the choice of $(f_n)_{n \in \mathbb{N}}$. In the case $p = \infty$, given any reference point $o \in X$, the sequence $(f_n)_{n \in \mathbb{N}}$ defined by $f_n := f 1_{B_n(o)}$ converges pointwise to $f$. By (2.5) and Lemma 2.1, the dominated convergence theorem shows that the pointwise limit $P^k_t f$ of $(P^k_t f_n)_{n \in \mathbb{N}}$ as $n \to \infty$ is well-defined, and again this definition does not depend on the choice of $(f_n)_{n \in \mathbb{N}}$ once it is demanded that $\|f_n\|_{L^\infty} \leq \|f\|_{L^\infty}$. It is clear that both approximation procedures preserve the bound (2.5).

To show strong continuity of $(P^k_t)_{t \geq 0}$ in $L^p(X)$ for $p < \infty$, by approximation and (2.5), it suffices to show continuity of $t \mapsto P^k_t f$ on $[0, \infty)$ in $L^p(X)$ for $f \in L^2(X) \cap L^p(X) \cap L^\infty(X)$. By the semigroup property, we may restrict to continuity at $t = 0$. Given any $x \in X$, note that a.s., we have $\int_0^t k(b^\xi_r) \, dr \to 0$ as $t \downarrow 0$ since $k \in L^1_{\text{loc}}(X)$, and that

$$ \left| e^{-\int_0^t k(b^\xi_r)/2 \, dr} - 1 \right| \leq e^{\int_0^T k(-b^\xi_r)/2 \, dr} + 1 \quad (2.6) $$

for every $t \in [0, T]$ is satisfied a.s. for fixed $T > 0$. Since

$$ \int_X |P^k_t f - P_t f|^p \, dm \leq \int_X \mathbb{E} \left[ \left| e^{-\int_0^T k(b^\xi_r)/2 \, dr} - 1 \right| |f(b^\xi_T)| 1_{\{t < \zeta^x\}} \right]^p \, dm(x), $$

applying the dominated convergence theorem twice using (2.6) as well as Lemma 2.1, we obtain $P^k_t f - P_t f \to 0$ in $L^p(X)$ as $t \downarrow 0$. The result follows immediately by strong continuity of the heat flow $(P_t)_{t \geq 0}$ in $L^p(X)$. \qed

Heat flow on 1-forms In the sequel, Borel equivalence classes of 1-forms on $X$ having a certain regularity $\mathcal{R}$ are denoted by $\Gamma_\mathcal{R}(T^*X)$, and similarly $\Gamma_\mathcal{R}(TX)$ for Borel equivalence classes of vector fields with regularity $\mathcal{R}$. For instance, given $p \in [1, \infty]$, we get the Banach space $\Gamma_{L^p}(T^*X)$ given by all Borel equivalence classes $\omega$ of sections in $T^*X$ such that $|\omega| \in L^p(X)$. Let $\Delta := d^* d + d d^*$ denote the Hodge–de Rham Laplacian. When defined initially on $\Gamma_{C^\infty}(T^*X)$, by geodesic completeness this operator has a unique self-adjoint extension in the Hilbert space $\Gamma_{L^2}(T^*X)$, which will be denoted with the same symbol again. Note our sign convention: $\Delta$ is nonnegative, while $\Delta$ is nonpositive. The heat semigroup $(\overline{P}_t)_{t \geq 0}$ on 1-forms given by $\overline{P}_t := e^{-t\Delta/2}$ in $\Gamma_{L^2}(T^*X)$ is smooth, in the sense for every $\omega \in \Gamma_{L^2}(T^*X)$ one has a jointly smooth representative $\overline{P}_t \omega$ which solves the heat equation

$$ \frac{\partial}{\partial t} \overline{P}_t \omega = -\frac{1}{2} \Delta \overline{P}_t \omega \quad \text{in } (0, \infty) \times X $$
on 1-forms with initial condition $\omega$ (and in $[0, \infty) \times X$ if $\omega$ is also smooth).

On exact forms, $\tilde{P}_t$ can be represented by the heat operator $P_t$; more precisely,

$$\tilde{P}_t df = dP_t f \quad \text{for every } f \in W^{1,2}(X), \ t \geq 0,$$

see [Gün17, Ros97]. Lastly, a key result we shall use at many instances (and which actually constitutes one part of Theorem 1.4) is the following Kato–Simon inequality.

**Theorem 2.3** [Gün17, Theorem VII.8]. Suppose that $\text{Ric} \geq k$ on $X$ for some Dynkin decomposable function $k: X \to \mathbb{R}$. Then

$$|\tilde{P}_t \omega| \leq P_t^k |\omega| \quad \text{m.a.e. for every } \omega \in \Gamma_{L^2}(T^*X), \ t \geq 0.$$

3 Proof of Theorem 1.3

This chapter treats the stochastic completeness of $X$, Bismut’s derivative formula, and the $L^\infty$-Lip-regularization of the heat semigroup $(P_t)_{t \geq 0}$ if we have $\text{Ric} \geq k$ on $X$ for some Dynkin decomposable function $k: X \to \mathbb{R}$. The latter inequality (and regularity of $k$) is assumed throughout this chapter.

3.1 Stochastic completeness

A key tool to prove that Brownian motion stays within $X$ forever a.s. are sequences of first-order cutoff-functions. Their existence is equivalent to the geodesic completeness of $X$.

**Lemma 3.1** [Gün16, Theorem 2.2]. There exists a sequence $(\psi_n)_{n \in \mathbb{N}}$ in $C_c^\infty(X)$ satisfying

(i) $\psi_n(X) \subset [0,1]$ for every $n \in \mathbb{N}$,

(ii) for all compact $K \subset X$, there exists $N \in \mathbb{N}$ such that $\psi_n|_K = 1_K$ for every $n \geq N$, and

(iii) $\|d\psi_n\|_{L^\infty} \to 0$ as $n \to \infty$.

**Proof of (i)** in Theorem 1.3. We are going to show the statement (2.4), i.e. that $P_t 1_X = 1_X$ for every $t > 0$. Let $\phi \in C_c^\infty(X)$, and let $(\psi_n)_{n \in \mathbb{N}}$ be a sequence of first-order cutoff functions provided by Lemma 3.1. Then Theorem 2.3 applied to the 1-form $\omega := d\psi_n$ for every $n \in \mathbb{N}$ and the Feynman–Kac formula (1.4) give

$$\|\tilde{P}_s d\psi_n\|_{L^\infty} \leq \sup_{x \in X} E \left[ e^{-\int_0^s k(b_x^+)^{1/2} \, dr} |d\psi_n(b_x^+)| \mathbb{I}_{\{s < \zeta^x\}} \right] \leq \sup_{x \in X} E \left[ e^{\int_0^s \kappa \, (b_x)^{1/2} \, dr} \mathbb{I}_{\{s < \zeta^x\}} \right] \|d\psi_n\|_{L^\infty} \leq C_1 e^{C_2 t} \|d\psi_n\|_{L^\infty},$$

uniformly in $s \in [0, t]$, where $C_1, C_2 \geq 0$ are finite constants depending only on the negative part of $k$, as provided by Lemma 2.1. Since $P_t \psi_n$ solves the heat equation (2.2), also using Fubini’s theorem, integration by parts as well as the commutation rule (2.7) we arrive at

$$\int_X (P_t \psi_n - \psi_n) \phi \, dm = \frac{1}{2} \int_X \int_0^t \phi \, dP_s \psi_n \, ds \, dm$$

$$= -\frac{1}{2} \int_0^t \int_X \langle d\phi, dP_s \psi_n \rangle \, dm \, ds = -\frac{1}{2} \int_0^t \int_X \langle d\phi, \tilde{P}_s d\psi_n \rangle \, dm \, ds.$$

9
Therefore, we obtain
\[ \left| \int_X (P_t 1_X - 1_X) \phi \right| = \lim_{n \to \infty} \left| \int_X (P_t \psi_n - \psi_n) \phi \right| \]
\[ \leq \limsup_{n \to \infty} \frac{1}{2} \int_0^t \int_X |d\phi| |\bar{P}_s \psi_n| \, dm \, ds \]
\[ \leq \frac{t}{2} C_1 e^{C_2 t} \| \psi_n \|_{L^1} \limsup_{n \to \infty} \| \psi_n \|_{L^\infty} = 0. \]

Since \( \phi \) was arbitrary, this proves the claim. \( \square \)

### 3.2 Bismut’s derivative formula and Lipschitz smoothing property

In view of proving Bismut’s derivative formula and the \( L^\infty \)-Lip-regularization property of \( (P_t)_{t \geq 0} \), for convenience we cite the following version of the Burkholder–Davis–Gundy inequality for \( q \in [1, \infty) \) (although we only need the upper bounds, respectively), which improves the classically known constants to better ones.

**Lemma 3.3** [Ren08 Theorem 2]. Let \( (M_r)_{r \geq 0} \) be a real-valued continuous local martingale with \( M_0 = 0 \), and let \( q \in [1, \infty) \). Then

\[ (8q)^{-q/2} \mathbb{E}[(M_r)^q] \leq \mathbb{E} \left[ \sup_{r \in [0, \tau]} |M_r|^q \right] \leq (8q)^q \mathbb{E}[(M_r)^q] \]

for every stopping time \( \tau \), where \( (|M_r|)_{r \geq 0} \) denotes the quadratic variation process of \( (M_r)_{r \geq 0} \).

Recall the stochastic process \( (Q_s^x)_{s \geq 0} \) defined by (1.1) and taking values in \( T_x X \).

**Lemma 3.3.** Let \( t \geq 0 \) and \( V \in \Gamma_{L^\infty}(TX) \). Then for every \( f \in L^\infty(X) \) and \( x \in X \), the random variable \( f(b_t^x) \int_0^t \langle Q_s^x V(x), d\omega_s^x \rangle \) is integrable. Moreover, for \( p = \infty \) and for every \( p \in (1, \infty) \) for which \( pk/(p - 1) \) is still Dynkin decomposable, the operator \( \mathbb{E}_t^V \) given by

\[ \mathbb{E}_t^V f(x) := \mathbb{E} \left[ f(b_t^x) \int_0^t \langle Q_s^x V(x), d\omega_s^x \rangle \right] \quad \text{for every } f \in L^\infty(X) \cap L^p(X), \ x \in X \quad (3.1) \]

extends to a bounded linear operator from \( L^p(X) \) into \( L^p(X) \), and the representation (3.1) is valid and well-defined for every \( f \in L^p(X) \).

**Proof.** Let \( V \in \Gamma_{L^\infty}(TX) \) and \( f \in L^\infty(X) \), for which we assume without loss of generality that \( \|V\|_{L^\infty} \leq 1 \) and \( \|f\|_{L^\infty} \leq 1 \). Fix \( t \) and \( x \) as above. Given any \( s \geq 0 \), it follows from Gronwall’s inequality and \( \text{Ric} \geq k \) on \( X \) that a.s.,

\[ |Q_s^x| \leq e^{-\int_0^s k(b^x)/2 \, dr}, \quad (3.2) \]

so that for every \( q \in [1, \infty) \), we obtain

\[ \mathbb{E} \left[ \sup_{r \in [0, t]} \left| \int_0^r \langle Q_s^x V(x), d\omega_s^x \rangle \right|^q \right] \leq (8q)^{q/2} \mathbb{E} \left[ \left( \int_0^t |Q_s^x|^2 \, ds \right)^{q/2} \right] \]
\[ \leq (8q)^{q/2} t^{1/2} \sup_{y \in X} \mathbb{E} \left[ e^{\int_0^t q k - (b^y)/2 \, dr} \right] \]
by Lemma 3.2. This inequality for \( q = 1 \) and Lemma 2.1 directly show the claimed integrability of the random variable \( f(b_r^x) \int_0^t \langle Q_s^x V(x), \, dw_r^s \rangle \). These facts also prove that \( E_1Y \) is a bounded linear operator from \( L^\infty(X) \) into \( L^\infty(X) \).

If \( p \in (1, \infty) \) is chosen such that \( pk/(p-1) \) is still Dynkin decomposable, Hölder’s inequality, (3.3) for \( q = p/(p-1) \) and the contractivity of \( (P_t)_{t \geq 0} \) show that there exist finite constants \( C_1, C_2 \geq 0 \) depending only on \( k^- \) and \( p \) such that for every \( f \in L^p(X) \cap L^\infty(X) \),

\[
\| E_1Yf \|_{L^p} \leq C_1 \frac{t}{p} e^{C_2t} \| f \|_{L^p},
\]

and we conclude by an approximation argument as in the proof of Theorem 2.2.

**Proof of (ii) in Theorem 1.3.** Fix \( x \in X \), \( t > 0 \) and \( V \in \Gamma_{C_0^\infty}(TX) \). It clearly suffices to assume that \( \| V \|_{L^\infty} \leq 1 \). We first assume that \( f \in C_0^\infty(X) \). By [DT01], Proposition 3.2], the process \( N^x \) given by

\[
N_r^x := \left\langle Q_r^x(\| x \|)^{-1} \nabla P_{t-r}f(b_r^x), \frac{t-r}{t} V(x) \right\rangle + \frac{1}{t} \int_0^r \langle Q_s^x V(x), \, dw_r^x \rangle,
\]

\( r \in [0, t] \), is a local martingale. We show that under the given assumption on \( k \), this process is even a martingale. Indeed, estimating \( |Q_r^x| \) by (3.2) and using the commutation rule (2.7), Kato–Simon’s inequality from Theorem 2.3 as well as Lemma 2.1 for all \( r \in [0, t] \) one a.s. has

\[
|N_r^x| \leq e^{\int_0^t k^- (b_s^x)/2 \, ds} \left| \int_0^r \langle Q_s^x V(x), \, dw_r^x \rangle \right| + \frac{\| f \|_{L^\infty}}{t} \left| \int_0^r \langle Q_s^x V(x), \, dw_r^x \rangle \right|,
\]

\[
\leq e^{\int_0^t k^- (b_s^x)/2 \, ds} \mathbb{E}_x \left[ e^{-\int_0^t k^- (x_s)/2 \, ds} |df(x_t)| \right] + \frac{\| f \|_{L^\infty}}{t} \left| \int_0^r \langle Q_s^x V(x), \, dw_r^x \rangle \right|,
\]

\[
\leq e^{\int_0^t k^- (b_s^x)/2 \, ds} \mathbb{E}_x \left[ \sup_{y \in X} e^{-\int_0^t k^- (b_s^y)/2 \, ds} \| df \|_{L^\infty} \right] + \frac{\| f \|_{L^\infty}}{t} \left| \int_0^r \langle Q_s^x V(x), \, dw_r^x \rangle \right|,
\]

\[
\leq e^{\int_0^t k^- (b_s)/2 \, ds} C_1 e^{C_2t} \| df \|_{L^\infty} + \frac{\| f \|_{L^\infty}}{t} \left| \int_0^r \langle Q_s^x V(x), \, dw_r^x \rangle \right|.
\]

Here \( x \) denotes Brownian motion on \( X \) starting in \( b_r^x \). It follows that

\[
\mathbb{E} \left[ \sup_{r \in [0, t]} |N_r^x| \right] \leq \mathbb{E} \left[ e^{\int_0^t k^- (b_s^x)/2 \, ds} C_1 e^{C_2t} \| df \|_{L^\infty} + \frac{\| f \|_{L^\infty}}{t} \right] \mathbb{E} \left[ \sup_{r \in [0, t]} \left| \int_0^r \langle Q_s^x V(x), \, dw_r^x \rangle \right| \right].
\]

The first summand is finite by Lemma 2.1, and the second one can be estimated by (3.3) for \( q = 1 \). Again using Lemma 2.1 it follows that \( (N_r^x)_{r \geq 0} \) is a true martingale, and thus

\[
\left\langle \nabla P_t f(x), V(x) \right\rangle = \mathbb{E}[N_0^x] = \mathbb{E}[N_t^x] = \frac{1}{t} \mathbb{E} \left[ f(b_t^x) \int_0^t \langle Q_s^x V(x), \, dw_r^x \rangle \right].
\]

To get rid of the smoothness of \( f \), one argues by approximation, recalling that \( C_0^\infty(X) \) is dense in \( L^p(X) \) for every \( p \in [1, \infty) \). Then one just has to note that, given \( p \in (1, \infty) \) such that \( pk^-/(p-1) \) is Dynkin decomposable, and \( f \in L^p(X) \), it follows from the divergence theorem as well as Lemma 3.3 that both sides of (3.4) are continuous in \( f \) w.r.t. convergence in \( L^p(X) \). The formula for the extremal case \( p = \infty \) then follows from the previous one by an additional approximation procedure through cutoff (compare with the proof of Theorem 2.2), noting that one can always find large enough \( p \in (1, \infty) \) so that \( pk^-/(p-1) \) is Dynkin.
decomposable by Remark 1.2 and that the right-hand side of (3.4) converges pointwise by Remark 3.3 and the dominated convergence theorem.

The claim that Bismut’s formula holds for every \( p \in (1, \infty) \) and \( f \in L^p(X) \) if \( k \) is Kato decomposable simply follows from the fact that, if \( p < \infty \), \( pk/(p - 1) \) is always Dynkin decomposable for such \( k \), see Remark 1.2.

Proof of (iii) in Theorem 1.4. Using Bismut’s formula proved above and Remark 3.3 for \( q = 1 \), for every \( x \in M \), \( t > 0 \) and \( \xi \in T_xM \) with \( |\xi| \leq 1 \), we get

\[
|\langle \nabla P_t f(x), \xi \rangle| \leq \frac{1}{t} \mathbb{E} \left[ \left| \int_0^t \langle Q^x_s \xi, \text{d} \mathcal{W}^x_s \rangle \right| \right] \|f\|_{L^\infty} \leq \sqrt{8} t^{-1/2} \sup_{x \in X} \mathbb{E} \left[ e^{\int_0^t (b^x_r - 1/2) \text{d}r} \right] \|f\|_{L^\infty},
\]

and duality gives

\[
\text{Lip}(P_t f) \leq \sqrt{8} t^{-1/2} \sup_{x \in X} \mathbb{E} \left[ e^{\int_0^t (b^x_r - 1/2) \text{d}r} \right] \|f\|_{L^\infty}.
\]

4 Proof of Theorem 1.4

We turn to equivalent characterizations of Dynkin lower Ricci curvature bounds in terms of functional inequalities. Throughout this chapter, we assume that \( k: X \to \mathbb{R} \) is Dynkin decomposable. Recall that “(i) \( \implies \) (iii)” in Theorem 1.4, i.e. the step from lower Ricci bounds to \( L^1 \)-gradient estimates, is already known thanks to Theorem 2.3.

4.1 From \( L^1 \)-gradient estimates to the \( L^1 \)-Bochner inequality

The proof of the following is quite well-known, see e.g. [BHS19] and the references therein, and extends to our setting.

Proof of “(iii) \( \implies \) (ii)” in Theorem 1.4. Let \( f \in C^\infty_c(X) \), and let \( \phi \in C^\infty_c(X) \) be a nonnegative function with support in \( \{ |\nabla f| \neq 0 \} \). Let \( \varepsilon > 0 \) and define \( \Phi_\varepsilon \in C^\infty([-\varepsilon/2, \infty)) \) by \( \Phi_\varepsilon(r) := (r + \varepsilon)^{1/2} - \varepsilon^{1/2} \). Given a fixed \( T > 0 \), define the functions \( G_\varepsilon, F_\varepsilon: [0, T] \to \mathbb{R} \) by

\[
G_\varepsilon(t) := \int_X \phi \Phi_\varepsilon(|\nabla P_t f|^2) \text{d}m, \quad F_\varepsilon(t) := \int_X \phi \Phi_\varepsilon((P_t^k |\nabla f|^2)^2) \text{d}m.
\]

By the smoothness properties of the heat semigroup and Lebesgue’s theorem, we have \( G_\varepsilon \in C^1([0, T]) \). Moreover, by local absolute continuity of \( t \mapsto P_t^k |\nabla f| \) in \( L^2(X) \) on \( [0, \infty) \) as well as Lemma 2.1, we also get the regularity \( F_\varepsilon \in C^1([0, T]) \). Together with (4.1), from (iii) we obtain \( G'_\varepsilon(0^+) \leq F'_\varepsilon(0^+) \), or equivalently,

\[
\int_X \phi \left( |\nabla f| + \varepsilon \right)^{-1/2} \langle \nabla \Delta f, \nabla f \rangle \text{d}m \leq \int_X \phi \left( |\nabla f| + \varepsilon \right)^{-1/2} |\nabla f| (\Delta - k)|\nabla f| \text{d}m.
\]

Since \( \varepsilon \) and \( \phi \) as above were arbitrary, we directly obtain (1.5). □
4.2 From the $L^1$-Bochner inequality to lower Ricci bounds

As already hinted, the key point in showing the implication “(ii) $\implies$ (i)” in Theorem 1.4 is the well-known Bochner formula (1.3), subject to a clever choice of $f$ as granted by the subsequent lemma, together with the chain rule to deduce $\text{Ric} \geq k$ on $X$.

It is well-known in Riemannian geometry that, given any $x \in X$, there exists an open subset $O_x \subset T_x X$ such that the restriction of the exponential map to $O_x$ provides a diffeomorphism $\exp_x: O_x \to \exp_x(O_x)$. We denote its inverse by $\exp_x^{-1}$.

Lemma 4.1 [RS05] Lemma 3.2. Let $x \in X$ and $\xi \in T_x X$ with unit norm. Let $\mathcal{H} := \{\exp_x \eta: \eta \in T_x X, \langle \eta, \xi \rangle = 0\}$ be the $(\dim(X) - 1)$-dimensional hypersurface in $X$ orthogonal to $\xi$ at $x$. Then there exists an open neighborhood $U \subset \exp_x(O_x)$ of $x$ such that the signed distance function $\rho^{\pm \xi} : U \to \mathbb{R}$ given by

$$\rho^{\pm \xi}(y) := \rho(y, \mathcal{H}) \text{ sgn}(\xi, \exp_x^{-1}(y)),$$ 

obeys

$$\rho^{\pm \xi} \in C^\infty(U), \quad \nabla \rho^{\pm \xi}(x) = \xi, \quad |\nabla \rho^{\pm \xi}(U)| = \{1\}, \quad \text{Hess} \rho^{\pm \xi}(x) = 0.$$

Proof of “(ii) $\implies$ (i)” in Theorem 1.4. Let ${(x_n : n \in \mathbb{N})}$ be a countable dense subset of $X$, and let $(\xi_n)_{n \in \mathbb{N}}$ be a sequence of tangent vectors $\xi_n \in T_{x_n} X$ with unit norm. Let $\mathcal{H}_n$ be the $(\dim(X) - 1)$-dimensional hypersurface and $U_n$ be the open set connected to $x_n$ and $\xi_n$ as provided by Lemma 4.1. Possibly multiplying $\rho^{\pm \xi_n}$ by a cutoff function and shrinking $U_n$ around $x_n$, we can find $f_n \in C^\infty_c(X)$ such that $f_n = \rho^{\pm \xi_n}$ on $U_n$.

Given any $\varepsilon > 0$, by the smoothness of $\text{Ric}$ and $f_n$, there exists $\delta_n > 0$ such that

$$|\text{Ric}_x(\xi, \xi) - \text{Ric}_{x_n}(\xi_n, \xi_n)| \leq \varepsilon, \quad |\langle \nabla \Delta f_n(x_n), \nabla f_n(x_n) \rangle - \langle \nabla \Delta f_n(x), \nabla f_n(x) \rangle| \leq \varepsilon$$

for every $x \in B_{\delta_n}(x_n)$ and every $\xi \in T_x X$ with unit norm. We may assume without loss of generality that $B_{\delta_n}(x_n) \subset U_n$. Therefore, by Bochner’s formula (1.3), the chain rule and Lemma 4.1 at any point $x \in B_{\delta_n}(x_n)$ at which (1.5) holds for $f_n$, we obtain

$$\text{Ric}(x)(\xi, \xi) \geq \text{Ric}(x_n)(\xi_n, \xi_n) - \varepsilon = \text{Ric}(x_n)(\nabla f_n, \nabla f_n) - \varepsilon$$

$$= |\nabla f_n(x_n)| \Delta |\nabla f_n(x_n)| + \langle \nabla |\nabla f_n(x_n)|, \nabla f_n(x_n) \rangle - \langle \nabla \Delta f_n(x_n), \nabla f_n(x_n) \rangle - \varepsilon$$

$$\geq |\nabla f_n(x)| \Delta |\nabla f_n(x)| - \langle \nabla \Delta f_n(x), \nabla f_n(x) \rangle - 2\varepsilon \geq k(x) - 2\varepsilon.$$

It follows that $\text{Ric} \geq k - 2\varepsilon$ on $X$. Since $\varepsilon$ was arbitrary, we obtain the desired claim. \qed

5 Proof of Theorem 1.5

From now on, let $k: X \to \mathbb{R}$ be a Dynkin decomposable and lower semicontinuous function.

For the proof of both implications addressed below we need the following fact, in which Lipschitz continuity on the product manifold $X \times X$ is understood w.r.t. the product metric $\rho_2$ given by $\rho^2_2((x, y), (x', y')) := \rho^2(x, x') + \rho^2(y, y').$

Lemma 5.1. Let $D \subset X$ be a compact subset. Then, in $D \times D$, $k$ is the pointwise limit of a pointwise increasing sequence of functions in $\text{Lip}_b(X \times X)$ which are everywhere not smaller than $\inf k(D \times D)$. In particular, in $D$, $k$ is the pointwise limit of a pointwise increasing sequence of functions in $\text{Lip}_b(X)$ which are everywhere not smaller than $\inf k(D)$. 

13
Proof. Every lower semi-continuous, lower bounded function on $X \times X$ can be approximated pointwise on $X \times X$ by a pointwise increasing sequence of functions in $\text{Lip}_b(X \times X)$ which preserves uniform lower bounds, see [BHS19, Lemma 2.1] and the references therein. If $\tilde{k}$ is not uniformly bounded from below, we apply the previous result to the function $\xi : X \times X \to \mathbb{R}$ given by $\xi(x, y) := \tilde{k}(x, y) \mathbbm{1}_{D \times D}(x, y) + \inf \tilde{k}(D \times D) \mathbbm{1}_{(D \times D)^c}(x, y)$.

The second statement follows by noting that $k(x) = \tilde{k}(x, x)$ for every $x \in X$.

5.1 From lower Ricci bounds to pathwise couplings

We start with the existence of a suitable coupling of Brownian motions under the inequality $\text{Ric} \geq k$ on $X$. (Note that the stochastic completeness of $X$ is already known by Theorem 1.3.) What we only have to show is that the techniques in [Cra91, Ken86] apply, and that the resulting coupling satisfies the desired estimates. See also [Wan94] for a “local” treatise on regular subdomains.

The crucial point is the following explicit Itô expansion from [Cra91], and the stated inequality is derived in [Ken86] using the index lemma. Let us denote by $R$ the Riemannian curvature tensor on $X$, and by $\text{Cut}_x$ the cut-locus of $x \in X$.

Theorem 5.2. For every $x, y \in X$, there exist a coupling $(b^x, b^y)$ of Brownian motions on $X$ starting in $(x, y)$, random Jacobi fields $W^{i,t}$, $i \in \{2, \ldots, \text{dim}(X)\}$, along $\gamma^t$, the unique minimizing unit speed random geodesic from $b^x_t$ to $b^y_t$, and a nondecreasing process $L$ with support contained in $\{t > 0 : b^y_t \in \text{Cut}_{b^x_t}\}$, such that

$$d\rho(b^x_t, b^y_t) = 2 \, d\beta_t + \frac{1}{2} \left[ \int_0^\rho(b^x_t, b^y_t) \sum_{i=2}^{\text{dim}(X)} \left( |\nabla W^{i,t}_s(\dot{\gamma}^t_s)|^2 - \langle R(W^{i,t}_s, \dot{\gamma}^t_s) \dot{\gamma}^t_s, W^{i,t}_s \rangle \right) ds \right] dt - dL_t$$

for every $t < T(b^x, b^y) := \inf \{t \geq 0 : b^x_t = b^y_t\}$, where the $dt$-coefficient is defined to be zero on the support of $L$, and $\beta$ is a Brownian motion on $\mathbb{R}$. In particular, for every $t < T(b^x, b^x)$, we have

$$d\rho(b^x_t, b^y_t) \leq 2 \, d\beta_t - \frac{1}{2} \left[ \int_0^\rho(b^x_t, b^y_t) \text{Ric}(\gamma^t_s)(\dot{\gamma}^t_s, \dot{\gamma}^t_s) ds \right] dt.$$

Moreover, we have $b^x_t = b^y_t$ for every $t \geq T(b^x, b^y)$.

The construction of this coupling is quite time- and space-demanding, whence we only sketch it; details are satisfactorily explained and motivated in [Cra91]. Take a Brownian motion $b^x$ on $X$ starting in $x$. Using Brownian motion on the frame bundle and a coupling by reflection technique, one can define an appropriate process $(b^y_t)_{t \in [0, \sigma(b^y))}$ until it hits the set $\text{Cut}_{b^x_t}$ at time $\sigma(b^y) := \inf \{t > 0 : b^y_t \in \text{Cut}_{b^x_t}\}$. For $t < \sigma(b^y)$, the smoothness of the distance function $\rho$ yields an explicit Itô expansion for the function $\rho(b^x_t, b^y_t)$ in terms of the above Jacobi fields and the velocity fields of $\gamma^t$. Then one has to prove [Cra91, Proposition 1] that $b^y$ can be expanded past the critical time $\sigma(b^y)$, and the only additional effect on $d\rho(b^x_t, b^y_t)$ is caused by the nonincreasing process $-L$. In view of bounding $d\rho(b^x_t, b^y_t)$ from above, this contribution can thus be ignored. Moreover, $(b^x, b^y)$ is a diffusion and therefore a Markov process [Hsu02, Theorem 6.5.1].

We are in a position to provide the final step from lower Ricci bounds to the existence of pathwise couplings. The subsequent lemma from [Shi81] will be useful for this enterprise.
Lemma 5.3. There exists a sequence \((\psi_n)_{n \in \mathbb{N}}\) of functions \(\psi_n \in C^2(\mathbb{R})\) converging pointwise to \(\text{Id}^+\) as \(n \to \infty\), such that for every \(n \in \mathbb{N}\),

(i) \(\psi_n = 0\) on \((-\infty, 0]\),

(ii) \(\psi'_n(\mathbb{R}) \subset [0, 1]\) as well as \(\psi''_n(\mathbb{R}) \subset [0, 2/n]\), and

(iii) \(\psi_n \leq \psi_{n+1}\) on \(\mathbb{R}\).

Proof of the “only if” part in Theorem 1.5. Fix arbitrary points \(x, y \in X\). We claim that the coupling \((b^x, b^y)\) from Theorem 5.2 satisfies the desired estimate, recalling that we already know its Markov property. Indeed, since \(\text{Ric} \geq k\) pointwise on \(X\), we get from Theorem 5.2 and the definition of \(\beta\) that, a.s.,

\[
d\rho(b^x_t, b^y_t) \leq 2 d\beta_t - \frac{1}{2} \left[ \int_0^\infty \rho(b^x_t, b^y_t) k(\gamma_s) \, ds \right] dt \leq 2 d\beta_t - \frac{1}{2} \rho(b^x_t, b^y_t) k(b^x, b^y) dt.
\]

The desired pathwise inequality for this coupling follows from a stochastic Gronwall-type argument shown similarly as for [Shi81] Lemma 3.4 combined with a stopping time argument, which we include for the convenience of the reader. Given any \(R > 0\), let \(\xi \in \text{Lip}_0(X \times X)\) satisfy the inequality \(\xi \leq k\) on \(\overline{B}_R(x) \times \overline{B}_R(y)\), see Lemma 5.1. Denote by \(\tau^x_R\) and \(\tau^y_R\) the first exit times of the marginal processes \(b^x\) and \(b^y\) from \(B_R(x)\) and \(B_R(y)\), respectively. For fixed \(t \geq 0\) and \(s \in [0, t]\), define the stopping time \(\tau_R(t) := \max\{s, \min\{t, \tau^x_R, \tau^y_R\}\}\) and consider the processes \(\rho\), defined for \(u \geq 0\), as well as \(e_s\), and \(y_{s, u}\), defined for \(u \geq s\), given by

\[
\rho_u := \rho(b^x_u, b^y_u), \quad e_{s,u} := e^\int_s^u \xi(b^x_r, b^y_r) \, dr, \quad y_{s,u} := \rho_u e_{s,u} - \rho_s.
\]

Since \(e_{s, u}\) is of locally bounded variation, it follows from the product rule for Itô integrals and Theorem 5.2 that \(d[y_{s,u}]_u = e^\int_s^u \xi(b^x_r, b^y_r) / 2 \, dr \, d\rho_u = 4 e^{\int_s^u \xi / 2} \, du\).

For \((\psi_n)_{n \in \mathbb{N}}\) as in Lemma 5.3, by Itô’s formula and (5.1), a.s. we have

\[
\psi_n(y_{s,\tau(t)}) = \int_s^{\tau(t)} \psi_n(y_{s,u}) \, du + \frac{1}{2} \int_s^{\tau(t)} \psi''_n(y_{s,u}) \, du + \int_s^{\tau(t)} \psi'_n(y_{s,u}) e_{s,u} \, du + 2 \int_s^{\tau(t)} \psi''_n(y_{s,u}) e_{s,u}^2 \, du
\]

\[
\leq 2 \int_s^{\tau(t)} \psi''_n(y_{s,u}) e_{s,u} \, du + \frac{4(t - s)}{n} \, e^{(t - s)\|\xi\|_{L^\infty}}.
\]

Taking expectations, using the optional stopping theorem and sending \(n \to \infty\) implies that \(y_{s,\tau(t)}^+\) has zero mean. Thus, we get that \(\rho_{\tau(t)} \leq \rho_s / e_{s,\tau(t)}\) a.s. Choosing a sequence of \(\ell\) according to Lemma 5.1 which can be chosen without changing \(R\), we get that, a.s.,

\[
\rho(b^x_{\tau_R(t)}, b^y_{\tau_R(t)}) \leq e^{-\int_{\tau_R(t)}^{\tau_R(t)} \xi(b^x_r, b^y_r) / 2} \rho(b^x, b^y).
\]

Finally sending \(R \to \infty\), noting that \(\tau_R(t) \to t\) a.s., and afterwards using the continuity of Brownian sample paths to find an exceptional set independent of \(s\) and \(t\), we obtain the desired pathwise coupling property w.r.t. \(k\).
5.2 From pathwise couplings to the $L^1$-Bochner inequality

The step from the pathwise coupling property w.r.t. $k$ towards (1.5) requires a nontrivial extension of the arguments for [BHS19 Theorem 5.17] (which adapt the duality argument from [Kuw10] to the case of synthetic variable Ricci bounds and make crucial use of uniform lower boundedness of the Ricci curvature) for short times instead of fixed ones. This kind of localization argument was indeed used in [BHS19] in different variants at different instances. For this, the smoothness of Ric, allowing us to bound it locally uniformly from below apart from any information on the relation between Ric and $k$, plays a crucial role.

We shall need the following exit time estimate for Brownian motion.

**Lemma 5.4** [Hsu02 Theorem 3.6.1]. Let $L \geq 1$ be a finite constant such that $\text{Ric} \geq -L^2$ on $B_1(x)$. Let $\tau^x$ be the first exit time of Brownian motion starting in $x$ from the open ball $B_1(x)$, and put $t := 1/8 \dim(X)L$. Then

$$\mathbb{P}[\tau^x \leq t] \leq e^{-L/2}. $$

**Proof of the “if” part in Theorem 1.5.** Let us start with some preparations. Let $f \in C^\infty_c(X)$ and $x \in X$ with $|\nabla f(x)| \neq 0$ be arbitrary, and let $\gamma = \text{Geo}(X)$ start in $x$ with $\rho(x, \gamma_1) \leq 1$. Given any $s \in (0, e^{-1/6}]$, put $L_s := -6 \log s$. The smoothness of $\text{Ric}$ implies the existence of some $\delta \in (0, e^{-1/6}]$ such that for every $s \in (0, \delta)$, one has

$$\text{Ric} \geq -L_s^2 \quad \text{on} \quad \overline{B}_s(x). \quad (5.2)$$

Lower semicontinuity of $k$ yields

$$k \geq K \quad \text{on} \quad \overline{B}_s(x)$$

for some $K \in \mathbb{R}$, whence

$$k \geq K \quad \text{on} \quad D \times D, \quad D := \{ z \in X : z \in \overline{B}_1(\gamma_s) \text{ for some } s \in [0, 1] \}. \quad (5.3)$$

Finally, let $\ell$ be any bounded Lipschitz function on $X \times X$ with $K \leq \ell \leq \bar{k}$ on $D \times D$, see Lemma 5.1. Let us denote by $(b^x, b^{\gamma_s})$ a process starting in $(x, \gamma_s)$ given by the pathwise coupling property w.r.t. $k$. This pair process does still depend on $s$, but we suppress this dependency from the notation. Let $\tau^x$ and $\tau^{\gamma_s}$ denote the first exit times of the marginal Brownian motions $b^x$ and $b^{\gamma_s}$ from $B_1(x)$ and $B_1(\gamma_s)$, respectively. Then for every $s \in [0, 1]$, a.s. we have

$$\rho(b^x_t, b^{\gamma_s}_t) \leq e^{-\int_0^t \ell(b^x_s, b^{\gamma_s}_s) ds} \leq e^{-\int_0^t \ell(b^x_s, b^{\gamma_s}_s) ds} \quad \text{if} \quad t \leq \min\{\tau^x, \tau^{\gamma_s}\}. \quad (5.4)$$

Define $t_s := 1/8 \dim(X)L_s$ and the event $A_s := \{ \tau^x > t_s, \tau^{\gamma_s} > t_s \}$ for $s \in (0, \delta]$. By joint smoothness of the heat semigroup, the time derivative of $|\nabla \mathbb{P}_tf(x)| = (|\nabla \mathbb{P}_tf(x)|^2)^{1/2}$ at $t = 0$ can be written and estimated via

$$\begin{align*}
\frac{1}{2} \langle \nabla f(x), \nabla f(x) \rangle &\leq \limsup_{s \downarrow 0} \frac{1}{t_s} \mathbb{E} \left[ |f(b^x_{t_s}) - f(b^{\gamma_s}_{t_s})| \right] - |\nabla f(x)| \\
&= \limsup_{s \downarrow 0} \frac{1}{t_s} \mathbb{E} \left[ |f(b^x_{t_s}) - f(b^{\gamma_s}_{t_s})| (\mathbb{1}_{A_s} + \mathbb{1}_{A^c_s}) \right] - |\nabla f(x)|.
\end{align*}$$

16
The contribution of $A_s^c$ becomes negligible thanks to

$$
\mathbb{E} \left[ \left| f(b_{t_s}^x) - f(b_{t_s}^{\gamma_x}) \right| \mathbb{1}_{A_s} \right] \leq 2 \|f\|_{L^\infty} \left( \mathbb{P}[\tau^x \leq t_s] + \mathbb{P}[\tau^{\gamma_x} \leq t_s] \right) \leq 4 \|f\|_{L^\infty} s^3
$$

by Lemma 5.4 and (5.2) and since $1/t_s$ only grows logarithmically as $s \downarrow 0$. Thus, we concentrate on the behavior of the integrand on $A_s$, an event which we decompose three further mutually disjoint subsets $V_s := A_s \cap \{ \rho(b_{t_s}^x, b_{t_s}^{\gamma_x}) \geq s^{1/2} \}$, $W_s := A_s \cap \left\{ \int_0^{t_s} \rho(b_{t_s}^x, b_{t_s}^{\gamma_x}) \, dr/t_s \geq s^{1/2} \right\}$, and $U_s := A_s \cap V_s^c \cap W_s^c$. Hence, it remains to estimate these three parts separately.

By (5.3), the contribution of $V_s$ can be bounded via

$$
\mathbb{E} \left[ \frac{|f(b_{t_s}^x) - f(b_{t_s}^{\gamma_x})|}{\rho(b_{t_s}^x, b_{t_s}^{\gamma_x})} \rho(b_{t_s}^x, b_{t_s}^{\gamma_x}) \mathbb{1}_{V_s} \right] 
\leq \|\nabla f\|_{L^\infty} s^{-1/2} \mathbb{E} \left[ \rho(b_{t_s}^x, b_{t_s}^{\gamma_x})^2 \mathbb{1}_{A_s} \right] 
\leq \|\nabla f\|_{L^\infty} s^{3/2} \mathbb{E} \left[ e^{-\int_0^{t_s} \rho(b_{t_s}^x, b_{t_s}^{\gamma_x}) \, dr} \mathbb{1}_{A_s} \right] 
\leq \|\nabla f\|_{L^\infty} s^{3/2} e^{-Kt_s}.
$$

In a similar way, we can control the influence of $W_s$ by

$$
\mathbb{E} \left[ \frac{|f(b_{t_s}^x) - f(b_{t_s}^{\gamma_x})|}{\rho(b_{t_s}^x, b_{t_s}^{\gamma_x})} \rho(b_{t_s}^x, b_{t_s}^{\gamma_x}) \mathbb{1}_{W_s} \right] 
\leq \|\nabla f\|_{L^\infty} s^{-1/2} \int_0^{t_s} \mathbb{E} \left[ \rho(b_{t_s}^x, b_{t_s}^{\gamma_x}) \rho(b_{t_s}^x, b_{t_s}^{\gamma_x}) \mathbb{1}_{A_s} \right] \, dr 
\leq \|\nabla f\|_{L^\infty} s^{3/2} e^{-Kt_s}.
$$

Finally turning to the study of the expectation on $U_s$, it is not difficult to derive from the Lipschitz continuity of $\ell$ and Jensen’s inequality that $\int_0^{t_s} \ell(b_{t_s}^x, b_{t_s}^{\gamma_x}) \, dr \geq \int_0^{t_s} \ell(b_{t_s}^x, b_{t_s}^{\gamma_x}) \, dr - \operatorname{Lip}(\ell) t_s s^{1/2}$ on $W_s^c$, where $\ell \in \operatorname{Lip}_b(X)$ is defined by $\ell(x) := \ell(x, x)$. Together with (5.4) and the definition of $A_s$, we then obtain

$$
\mathbb{E} \left[ \frac{|f(b_{t_s}^x) - f(b_{t_s}^{\gamma_x})|}{\rho(b_{t_s}^x, b_{t_s}^{\gamma_x})} \rho(b_{t_s}^x, b_{t_s}^{\gamma_x}) \mathbb{1}_{U_s} \right] 
\leq s \mathbb{E} \left[ e^{-\int_0^{t_s} \ell(b_{t_s}^x, b_{t_s}^{\gamma_x}) \, dr} e^{\operatorname{Lip}(\ell) t_s s^{1/2}} G_{s} f(b_{t_s}^x) \mathbb{1}_{A_s} \right] 
\leq s \mathbb{E} \left[ e^{-\int_0^{t_s} \ell(x_{t_s}^x, x_{t_s}^{\gamma_x}) \, dr} e^{\operatorname{Lip}(\ell) t_s s^{1/2}} G_{s} f(x_{t_s}^x) \right],
$$

where $G_{s} f(y) := \sup \{|f(y) - f(z)|/\rho(y, z) : z \in B_{s^{1/2}}(y)\}$. In the last step, we switched to a Brownian motion $x^x$ starting in $x$ which is independent of $s$.

Now we paste these three estimates together. Using smoothness and uniform continuity of $f$ (and of $|\nabla f|$ near $x$), we then finally arrive at

$$
\limsup_{s \downarrow 0} \frac{1}{t_s} \mathbb{E} \left[ \int \frac{|f(b_{t_s}^x) - f(b_{t_s}^{\gamma_x})|}{\rho(b_{t_s}^x, b_{t_s}^{\gamma_x})} \rho(b_{t_s}^x, b_{t_s}^{\gamma_x}) \mathbb{1}_{A_s} \right] - |\nabla f(x)|
\leq \limsup_{s \downarrow 0} \frac{1}{t_s} \left[ \mathbb{E} \left[ |f(b_{t_s}^x) - f(b_{t_s}^{\gamma_x})| \mathbb{1}_{A_s} \right] - |\nabla f(x)| \right]
\leq \limsup_{s \downarrow 0} \frac{1}{t_s} \left[ \mathbb{E} \left[ e^{-\int_0^{t_s} \ell(x_{t_s}^x, x_{t_s}^{\gamma_x}) \, dr} e^{\operatorname{Lip}(\ell) t_s s^{1/2}} G_{s} f(x_{t_s}^x) \right] - |\nabla f(x)| \right]
= \frac{1}{2} (\Delta - \ell)|\nabla f(x)|.
$$

Since $\ell$ was arbitrary, we conclude the inequality (1.5) by Lemma 5.1.

\[\square\]
References


