Expectations and Martingales in Metric Spaces

T. Christiansen and K.T. Sturm

Abstract. We develop a theory of martingales with values in metric spaces of nonpositive curvature. Our main results state existence of (continuous-time) martingales $(X_t)_{0\leq t\leq T}$ with given terminal data X_T. These processes will be constructed via time discretization. The notion of discrete-time martingale is based on the concept of iterated conditional barycenters. Moreover, in more specific cases we present martingale characterizations.

Introduction

In terms of Ito's calculus, the martingale theory for time-continuous processes in $\mathbb{R}^n$ could be successfully generalized to the case of Riemannian manifolds. Moreover, this definition turned out to be equivalent to Darling's characterization in terms of convex functions, being a localized version of Jensen's inequality (cf. [Dar82]). In singular metric spaces one does not have a smooth calculus, so the former definition cannot be extended. The latter, can easily be extended to define ('weak') martingales with values in metric spaces. However, for general metric spaces the main question remains to prove existence and uniqueness of martingales with given terminal data. So far, existence results are only known in manifolds and trees, and typically the technique of approximation of iterated barycenters is used (cf. [Ken90], [Pic91], [Arn95], [Pic04]).

In this paper we use the above technique in order to develop a theory of ('strong') martingales, based on time discretization and the concept of iterated barycenters in spaces of nonpositive curvature (NPC spaces). The more elementary theory of time-discrete martingales with values NPC spaces was already developed in [Stu02]. In order to define martingales for continuous-time filtrations $(\mathcal{F}_t)_{0\leq t\leq T}$, we fix a refining sequence $\Delta^n = \{t_0^n, t_1^n, \ldots, t_{k(n)}^n\}$ of partitions of $[0,T]$ with their mesh converging to 0. A process $(X_t)_{0\leq t\leq T}$ will be called *martingale* if it is the limit of a sequence of processes $(X^n_{t^n_k})_{k=1,\ldots,k(n)}$ which are martingales with respect to the discrete-time filtrations $(\mathcal{F}_{t^n_k})_{k=1,\ldots,k(n)}$.

The main results of this paper will be two existence results for martingales $(X_t)_{0\leq t\leq T}$ with given terminal data X_T (Thm. 2.5 and Thm. 2.14). Moreover, we give a characterization of martingales in terms of a 'quadratic variation process' (Thm. 3.2). As a consequence, if N is a Riemannian manifold, a continuous integrable semimartingale is a martingale if and only if it is a $\nabla-$martingale in the classical sense. Moreover, on trees our definition is shown to be equivalent to the one of Picard ([Pic04]).

1 NPC spaces

1.1 Barycenters and Expectations

Metric spaces with $Curv \leq \kappa$ in the sense of Alexandrov are defined by comparing geodesic triangles in the two-dimensional Model space with constant curvature κ. For more details the reader is referred to [BH99]. Here we are interested in the case $Curv \leq 0$, or in other words, nonpositive curvature (NPC). Calculating Euclidean distances yields the following rigorous

Definition 1.1 *A complete separable geodesic space* (N,d) *is called* NPC-space *if*

$$d^2(z,\gamma(t)) \leq (1-t)d^2(z,\gamma(0)) + td^2(z,\gamma(1)) - t(1-t)d^2(\gamma(0),\gamma(1)) \tag{1}$$

for any $z \in N$*, any geodesic* $\gamma : [0,1] \to N$ *and any* $t \in [0,1]$.

A consequence of this definition is Reshetniak´s *Quadruple inequality* (cf [Jos97]) which we quote for later use. Namely, for all $x_1, x_2, x_3, x_4 \in N$,

$$d^2(x_1,x_3) + d^2(x_2,x_4) \leq d^2(x_2,x_3)d^2(x_1,x_4) + 2d(x_1,x_2)d(x_3,x_4) \tag{2}$$

Example 1.2 The following classes of spaces are NPC spaces:

- Complete Riemannian manifolds with nonpositive sectional curvature
- Hilbert spaces
- Bruhat Tits buildings, in particular metric trees

Moreover, if M is an NPC space and $(\Omega, \mathcal{F}, P)$ is a probability space, then $L^2(\mathcal{F}, N)$ is an NPC space, too (cf [Jos97]).

Let (N,d) be a complete separable metric space. Denote by $\mathcal{P}(N)$ the set of all probability measures on $(N, \mathcal{B}(N))$ and by $\mathcal{P}^\theta(N)$ the set of $p \in \mathcal{P}(N)$ with $\int d^\theta(x,y)p(dy) < \infty$ for some (and by the triangle inequality all) $x \in N$. On $\mathcal{P}^1(N)$, we define the (L^1-) *Wasserstein-distance* by

$$d^W(p,q) := \inf \left\{ \int_{N^2} d(x,y)\mu(d(x,y)) : \mu \in \mathcal{M}(p,q) \right\}, \tag{3}$$

where $\mathcal{M}(p,q)$ is the set of all *couplings* of p and q, i.e. all probability measures $\mu \in \mathcal{P}(N^2)$ whose marginals are p and q, i.e.

$$\mu(A \times N) = p(A) \quad \text{and} \quad \mu(N \times A) = q(A) \qquad (\forall A \in \mathcal{B}(N)).$$

Definition 1.3 *A* <u>*barycenter map*</u> *is a map* $b : \mathcal{P}^1(N) \to N$ *satisfying*

(i) $b(\delta_x) = x$ *for all* $x \in N$

(ii) $d(b(p), b(q)) \leq d^W(p,q)$ *for all* $p, q \in \mathcal{P}^1(N)$

In NPC spaces, there is a canonical barycenter map in the following sense: From (1) follows that for $z \in N$ the function

$$f^z(y) := d^2(z,y) \tag{4}$$

is strictly convex. Thus, a barycenter can be defined as the minimizer of the mean squared distance in the spirit of C.F. Gauß. For details and proofs of the following Proposition we refer to [Stu02] and [EF01].

Proposition 1.4 *Let* N *be an NPC space. Then there is a unique barycenter map* $b : \mathcal{P}^1(N) \to N$ *such that for all* $p \in \mathcal{P}^2(N)$

$$\int d^2(x, b(p))p(dx) \leq \int d^2(x,z)p(dx)$$

for all $z \in N$*. This map is called* canonical barycenter *and has the following properties:*

(i) (**Variance inequality**) *For all $p \in \mathcal{P}^2(N)$ and $z \in N$,*

$$\int d^2(x,z)p(dx) \geq \int d^2(x,b(p))p(dx) + d^2(b(p),z). \tag{5}$$

(ii) (**Jensen's inequality**) *For all $p \in \mathcal{P}^1(N)$ and all lower semicontinuous convex fuctions $\varphi : N \to \mathbb{R}$,*

$$\varphi(b(p)) \leq \int \varphi(x)p(dx).$$

Remark 1.5 (i) If N is a Cartan-Hadamard-manifold, then the above barycenter is the exponential barycenter (or Karcher´s mean), cf e.g. [ÉM91]. In general Riemannian manifolds, there need not exist a barycenter map anymore and things become more involved, cf [Ken92].
(ii) Other barycenter maps in metric spaces were constructed in [LT91] and [ESH99].

Let $(\Omega, \mathcal{F}, P)$ be a probability space and $X \in L^1(\mathcal{F}, N)$, the space of all $\mathcal{F}-$measurable random maps such that $P \circ X^{-1} \in \mathcal{P}^1(N)$. Let $\mathcal{G} \subset \mathcal{F}$ be a sub-σ-algebra and let $P_{X|\mathcal{G}} : \Omega \to \mathcal{P}(N)$ be the regular conditional probability for X given $\mathcal{G}$ which exists and is a.s. unique because N is separable. Then

$$\mathbf{E}^{\mathcal{G}}[X] := \mathbf{E}[X|\mathcal{G}] := b \circ P_{X|\mathcal{G}}$$

is called the *conditional expectation* of X, conditioned on $\mathcal{G}$.
$\mathbf{E}^{\mathcal{G}}[X] : (\Omega, \mathcal{G}) \to (N, \mathcal{B}(N))$ is measurable, which can be shown by approximating X by a sequence of random variables with finite range, or directly by showing that $P_{X|\mathcal{G}} : (\Omega, \mathcal{G}) \to (\mathcal{P}(N), \mathcal{B}(\mathcal{P}(N)))$ is measurable, where $\mathcal{B}(\mathcal{P}(N))$ is the Borel $\sigma-$algebra induced by the topology of the Wasserstein distance.
In particular, if $\mathcal{G} = \{\emptyset, \Omega\}$, then

$$\mathbf{E}[X] := \mathbf{E}^{\mathcal{G}}[X]$$

is a constant and is called the *expectation* of X.

Example 1.6 Let (M, ρ) be a separable metric space. Let $(p_t)_{t>0}$ be a Markovian transition function on M and $(\Omega, X_t, \mathcal{F}_t, P^x)$ the corresponding Markov process with canonical filtration. For a measurable map $f : M \to N$ such that $p_t(x) \circ f^{-1} \in \mathcal{P}^1(N)$ for all t and x (where $p_t(x)$ is regarded as a probability measure on M), we define the *nonlinear Markov operator* $P_t f : M \to N$ by

$$P_t f(x) := b(p_t(x) \circ f^{-1}). \tag{6}$$

If we put $Y_t := f(X_t)$ then for all $x \in M$

$$\mathbf{P}^x_{Y_{s+t}|\mathcal{F}_s}(\omega) = p_t(X_s(\omega)) \circ f^{-1}$$

and hence $\mathbf{E}^x[Y_{s+t}|\mathcal{F}_s] = P_t f(X_s)$.

1.2 Martingales in Discrete Time

We will now come to the notion of martingales. Let $(\Omega, (\mathcal{F}_n)_{n\in\mathbb{N}}, \mathcal{F}, P)$ be a filtered probability space, $m \in \mathbb{N}$ and $X \in L^1(\mathcal{F}_m, N)$. Unfortunately, the conditional expectation is in general not projective, i.e. for $k \leq l \leq m$ the classical identity

$$\mathbf{E}^{\mathcal{F}_k}[X] = \mathbf{E}^{\mathcal{F}_k}\mathbf{E}^{\mathcal{F}_l}[X]$$

does not hold in general (c.f. [Stu02], Example 3.2.).
However, we can define the discrete *filtered conditional expectation* (short: *FCE*) of X w.r.t. $(\mathcal{F}_n)_{n\in\mathbb{N}}$ by

$$\mathbf{E}^{(\mathcal{F}_n)_{n\geq k}}[X] := \mathbf{E}[X|(\mathcal{F}_n)_{n\geq k}] := \begin{cases} \mathbf{E}^{\mathcal{F}_k}\mathbf{E}^{\mathcal{F}_{k+1}}\dots\mathbf{E}^{\mathcal{F}_{m-1}}[X] & \text{if } k < m \\ X & \text{if } k \geq m. \end{cases}$$

Clearly,

$$\mathbf{E}^{(\mathcal{F}_n)_{n\geq k}}[X] = \mathbf{E}^{(\mathcal{F}_n)_{n\geq k}}[\mathbf{E}^{(\mathcal{F}_n)_{n\geq l}}[X]] \qquad \text{for all } k \leq l$$

or in words, the discrete FCE is projective.
We will call an adapted process $(X_n)_{n\in\mathbb{N}}$ such that $X_n \in L^1(\mathcal{F}_n, N)$ for all n a *martingale* if $\mathbf{E}^{(\mathcal{F}_n)_{n\geq k}}[X_l] = X_k$ for all $k \leq l$, or equivalently, if $\mathbf{E}^{\mathcal{F}_k}[X_{k+1}] = X_k$ for all $k \in \mathbb{N}$.

As we have seen, the canonical barycenter enjoys certain properties, in particular the variance inequality and Jensen's inequality. In [Stu02] was developed a discrete-time martingale theory. We will shortly quote some results, which can be derived from Proposition 1.4:

Proposition 1.7 *Let $(X_n)_{n\in\mathbb{N}_0}$ be a martingale. Then*

(i) $(\varphi(X_n))_{n\in\mathbb{N}}$ *is a submartingale for all lower semicontinuous convex functions* $\varphi : N \to \mathbb{R}$ *such that* $\varphi(X_n) \in L^1$ *for all* n.

(ii) Let $X_n \in L^2(\mathcal{F}_n, N)$ *for all* n. *Define*

$$V_n := \sum_{k=1}^{n} \mathbf{E}^{\mathcal{F}_{k-1}}[d^2(X_{k-1}, X_k)].$$

Then $f^z(X_n) - V_n := d^2(z, X_n) - V_n$ *is a submartingale for all* $z \in N$. □

1.3 Martingales in Continuous Time

Let $0 \leq s < t \leq \infty$ and $(\Omega, (\mathcal{F}_\tau)_{s\leq\tau\leq t}, \mathcal{F}, P)$ be a filtered probability space and $\xi \in L^1(\mathcal{F}_t, N)$. In order to define FCE in continuous time, we take a sequence of partitions of $[s,t]$ with their mesh converging to 0 and consider the limit of the discrete FCE, provided it exists.
In order to formulate this rigorously, we need some notation. A partition of $[0,\infty[$ is a set $\Delta = \{t_k : k \in \mathbb{N}\}$ such that $t_0 = 0$, $t_k < t_{k+1}$ and $t_k \to \infty$ as $k \to \infty$. The mesh of Δ is defined by

$$\|\Delta\| := \sup_{t_k\in\Delta} |t_{k+1} - t_k|.$$

For the sequel we fix a sequence $(\Delta^n)_{n\in\mathbb{N}}$ of partitions of $[0,\infty[$ such that $\Delta^n \subset \Delta^{n+1}$ and $\|\Delta^n\| \to 0$ and put $\mathbb{T} := \bigcup \Delta^n$. $\mathbb{T}$ is a dense subset of $[0,\infty[$.

Let $n \in \mathbb{N}$ and $s,t \in \Delta^n$ such that $s < t$. Then $\Delta^n \cap [s,t] = \{t_0, \dots, t_m\}$ with $s = t_0 < t_1 < \dots < t_m = t$. For $\xi \in L^1(\mathcal{F}_t, N)$ we define

$$\xi^n_{t_k} := \mathbf{E}^{\Delta^n}_k[\xi] := \mathbf{E}^{\mathcal{F}_{t_k}}\mathbf{E}^{\mathcal{F}_{t_{k+1}}}\dots\mathbf{E}^{\mathcal{F}_{t_{m-1}}}[\xi], \qquad k = 0\dots m-1$$

and the elementary process

$$\xi^n_\tau := \xi^n_{t_k} \qquad \text{for } \tau \in [t_k, t_{k+1}[. \tag{7}$$

Note that $(\xi^n_{t_k})_{0\leq k\leq m}$ is the martingale w.r.t. the discrete-time filtration $(\mathcal{F}_{t_k})_{0\leq k\leq m}$ with endpoint $\xi^n_m = \xi$. If ξ^n_s converges to some (ξ_s) in L^1, then ξ_s is called the (continuous-time) *FCE* of ξ w.r.t $(\Delta^n)_{n\in\mathbb{N}}$, conditioned on $\mathcal{F}_s$ and we write

$$\mathbf{E}^{(\mathcal{F}_\tau)_{\tau\geq s}}[\xi] := \xi_s.$$

Now we can introduce the notion of a martingale.

Definition 1.8 *Let $(\Omega,(\mathcal{F}_t)_{t\in\mathbb{T}},\mathcal{F},P)$ be a filtered probability space and $X=(X_t)_{t\in\mathbb{T}}$ be a process such that $X_t\in L^1(\mathcal{F}_t,N)$ for all $t\in\mathbb{T}$. X is called a* martingale *w.r.t. $(\Delta^n)_{n\in\mathbb{N}}$ if for all $s,t\in\mathbb{T}$ with $s\leq t$, $\mathbf{E}^{(\mathcal{F}_\tau)_{\tau\geq s}}[X_t]$ exists and is equal to X_s.*

Theorem 1.9 *Let $X=(X_t)_{t\in\mathbb{T}}$ be a martingale. Then $(\varphi(X_t))_{t\in\mathbb{T}}$ is a submartingale for all lower semicontinuous convex functions $\varphi : N\to\mathbb{R}$ such that $\varphi(X_t)\in L^1$ for all t.*

Proof : Let $s,t\in\mathbb{T}$ with $s<t$. Put $\xi := X_t$. We use the notation of (7). Then $\xi^n_s\to X_s$ in L^1, and by choosing a subsequence we can assume that $\xi^n_s\to X_s$ $P-$a.s. Now $\varphi(\xi^n_s)\leq \mathbf{E}^{\mathcal{F}_s}[\varphi(X_t)]$ for all n, $P-$a.s. Due to the lower semicontinuity of φ we have

$$\varphi(X_s)\leq \liminf_{n\to\infty}\varphi(\xi^n_s)\leq \mathbf{E}^{\mathcal{F}_s}[\varphi(X_t)] \qquad \text{P-a.s.}\square$$

Another feature of a martingale is that it 'respects' the product structure of NPC spaces. For instance, let $(N_1,d_1),(N_2,d_2)$ be two NPC spaces. On $N_1\times N_2$ define the product distance by $d^2((x_1,x_2),(y_1,y_2)) := d_1^2(x_1,y_1)+d_2^2(x_2,y_2)$. Then $N_1\times N_2$ is again an NPC space, cf [Jos97].

Proposition 1.10 *Let N_1,N_2 be two NPC spaces. Let $(X^1_t)_{t\in\mathbb{T}}$ and $(X^2_t)_{t\in\mathbb{T}}$ be two adapted processes in N_1 and N_2, respectively. Then (X^1,X^2) is a martingale in $N_1\times N_2$ if and only if X^i is a martingale in N_i for $i=1,2$.*

Proof : The definition of the canonical barycenter implies that if $p_i\in\mathcal{P}^2(N_i)$, then $b(p) = (b(p_1),b(p_2))$ for any coupling p of p_1 and p_2. Since $\mathcal{P}^2(N_i)$ is dense in $\mathcal{P}^1(N_i)$, this is also true for $p_i\in\mathcal{P}^1(N_i)$. So if $X^i\in L^1(\mathcal{F},N_i)$, then $\mathbf{E}^{\mathcal{G}}[(X^1,X^2)] = (\mathbf{E}^{\mathcal{G}}[X^1],\mathbf{E}^{\mathcal{G}}[X^2])$ and consequently the assertion holds for time-discrete martingales. Thus by approximation the Proposition is proved. $\square$

It is known that the distance function is convex on $N\times N$, cf [Jos97]. So combining the above proposition and Theorem 1.9 immediately yields

Corollary 1.11 *Let $(X_t)_{t\in\mathbb{T}}$ and $(Y_t)_{t\in\mathbb{T}}$ be two martingales w.r.t. the same filtration and sequence of partitions. Then the distance process $(d(X_t,Y_t))_{t\in\mathbb{T}}$ is a submartingale.*

In particular, if $(X_t)_{t\in\mathbb{T}}$ is a martingale, then $(d(X_t,z))_{t\in\mathbb{T}}$ is a submartingale for all $z\in N$. Thus X is a martingale in the sense of [Dos62]. For these 'martingales', Doss proved a convergence theorem; for a proof see e.g. [Stu02]).

Recall that a complete metric space is called *proper* if all closed balls are compact. For instance, an NPC space is proper if and only if it is locally compact (cf. [Bal95]).

Proposition 1.12 (Convergence Theorem) *Let $(\Omega,\mathcal{F},(\mathcal{F}_n)_{n\in\mathbb{N}},P)$ be a filtered probability space and N a proper metric space. Let $(X_n)_{n\in\mathbb{N}}$ be an adapted process such that $d(z,X)$ is a submartingale with $\sup_{n\in\mathbb{N}}\mathbf{E}[d(z,X_n)]<\infty$ for all $z\in N$. Then there is an $\mathcal{F}_\infty-$ measurable map $X_\infty:\Omega\to N$ such that*

$$\lim_{n\to\infty} X_n = X_\infty \qquad P\text{-a.s.}$$

If $d(X,z)$ is uniformly $p-$integrable, then we also have convergence in L^p.

Remark 1.13 (i) One can prove a corresponding backward martingale convergence theorem, i.e. for decreasing filtrations.

(ii) The convergence theorem immediately implies an analogous result for contiuous-time processes $(X_t)_{t\in\mathbb{T}}$ and hence we have the following

Corollary 1.14 *Let N be a proper NPC space and let $(X_t)_{t\in\mathbb{T}}$ be a martingale such that $\sup_{t\in\mathbb{T}}\mathbf{E}[d(z,X_t)]<\infty$ for all $z\in N$. Then there is an $\mathcal{F}_\infty-$ measurable map $X_\infty:\Omega\to N$ such that*

$$\lim_{t\to\infty} X_t = X_\infty \qquad P\text{-}a.s.$$

If $d(X,z)$ is uniformly $p-$integrable, then we also have convergence in L^p.

2 Existence of FCE and martingales

Given a filtered probability space $(\Omega,\mathcal{F}_t,\mathcal{F},P)$, we want to formulate conditions under which we can prove existence of continuous-time FCE's, or equivalently, of martingales with prescribed limit. For similar results in the case of Riemannian manifolds, see e.g. [Ken90], [Pic91] and [Arn95].

2.1 A coupling condition

For simplicity, we will only consider dyadic partitions. More precisely, let $\Delta^n := \{k2^{-n} : k\in\mathbb{N}\}$ and $\mathbb{T} := \bigcup\Delta^n$, the set of nonnegative dyadic numbers.
Let $(\Omega,(\mathcal{F}_t)_{t\in\mathbb{T}},\mathcal{F},P)$ be a filtered probability space such that Ω is a Polish space and $\mathcal{F}\subset\mathcal{B}(\Omega)$. Then for all $s\in\mathbb{T}$, the regular conditional probability of P given $\mathcal{F}_s$, denoted by

$$Q_s(\omega) := P_{id|\mathcal{F}_s}(\omega)\in\mathcal{P}(\Omega), \tag{8}$$

exists and is unique. Let $(\rho_t)_{t\in\mathbb{T}}$ be a family of symmetric nonnegative functions on $\Omega\times\Omega$. Assume that each of the ρ_s is a.s. separable in the sense that there is a countable set $\Omega_1\subset\Omega$ such that for all $s\in\mathbb{T}$ and almost all $\omega\in\Omega$

$$\inf\{\rho_s(\omega,\tilde\omega):\tilde\omega\in\Omega_1\}=0. \tag{9}$$

Moreover, assume that for all $t,s\in\mathbb{T}$ with $s\le t$

$$\rho_t^W(Q_s(\omega_1),Q_s(\omega_2))\le\rho_s(\omega_1,\omega_2) \tag{10}$$

where ρ_t^W is defined as in (3). Note that since ρ_t need not be a metric on Ω, ρ_t^W will not be a metric in general.

Let now N be an NPC space and put

$$\mathcal{L}_N := \{Y:\Omega\times\mathbb{T}\to N : \mathbf{E}[d(Y_s,z)]<\infty \text{ for all } s\in\mathbb{T} \text{ and } z\in N\}.$$

For $Y\in\mathcal{L}_N$ we define a new process $P_tY\in\mathcal{L}_N$ by

$$P_tY(\omega,s) := \mathbf{E}^{\mathcal{F}_s}[Y_{t+s}](\omega) = b(Q_s(\omega)\circ Y_{t+s}^{-1}). \tag{11}$$

Lemma 2.1 *Let $Y \in \mathcal{L}_N$ be a process such that for all $s \in \mathbb{T}$ and almost all ω_1, ω_2*

$$d(Y_s(\omega_1), Y_s(\omega_2)) \leq C\rho_s(\omega_1, \omega_2). \tag{12}$$

Then

$$d((P_t^n Y)_s(\omega_1), (P_t^n Y)_s(\omega_2)) \leq C\rho_s(\omega_1, \omega_2).$$

Proof : Let $n = 1$. From (12) and (10) it follows that

$$\begin{aligned} d^W(Q_s(\omega_1) \circ Y_{t+s}^{-1}, Q_s(\omega_2) \circ Y_{t+s}^{-1}) &\leq C\rho_{s+t}^W(Q_s(\omega_1), Q_s(\omega_2)) \\ &\leq C\rho_s(\omega_1, \omega_2) \end{aligned}$$

and hence by the barycenter contraction property we derive the claim for $n = 1$. For arbitrary n, this can be iterated. $\square$

A process Y satisfying (12) is a kind of 'Lipschitz continuous' map (regarding the ρ_t as a family of pseudometrics on Ω). We set

$$Lip(\Omega, N) := \{Y \in \mathcal{L}_N : Y \text{ satisfies (12) }\}$$

Theorem 2.2 *Let N be a proper NPC space. Let $(\Omega, (\mathcal{F}_t)_{t\in\mathbb{T}}, \mathcal{F}, P)$ be a filtered probability space satisfying (10). Then there is a subsequence n_k such that for all $Y \in \mathcal{L}_N$ satisfying (12), all $s, t \in \mathbb{T}$ and almost all $\omega \in \Omega$*

$$P_t^* Y(\omega, s) := \lim_{k\to\infty} P_{\delta_k}^{t/\delta_k} Y(\omega, s)$$

exists, where $\delta_k := 2^{-n_k}$. The family $(P_t^*)_{t\geq 0}$ is a semigroup acting on $Lip(\Omega, N)$. Moreover, for any t, the process $((P_t Y)_s)_{s\in\mathbb{T}}$ is a martingale.

Proof : First note that by Corollary 1.11,

$$d(P_t^n Y(\omega, s), P_t^n \tilde{Y}(\omega, s)) \leq \mathbf{E}^{\mathcal{F}_s}[d(Y_{s+nt}, \tilde{Y}_{s+nt})](\omega) \tag{13}$$

for almost all ω. Thus there is a set $\Omega_0 \subset \Omega$ with $P(\Omega_0) = 1$ such that (10) and (13) hold pointwise for all $s, t \in \mathbb{T}$, $n \in \mathbb{N}$ and $\omega \in \Omega_0$.
Fix $t \in \mathbb{T}$. Let $s \in \mathbb{T}$. and $\omega \in \Omega_0$. Put $z_n(\omega, s, Y) := P_{2^{-n}}^{t2^n} Y(\omega, s) \in N$, where n is assumed to be large enough such that $t2^n \in \mathbb{N}$. Let $z \in N$. From (13), applied to Y and the constant process $\tilde{Y}(\omega, s) \equiv z$, follows that

$$d(z_n(\omega, s, Y), z) \leq \mathbf{E}^{\mathcal{F}_s}[d(Y_{s+t}, z)](\omega) < \infty$$

for all n. In other words, all $z_n(\omega, s, Y)$ are contained in a closed ball, which is compact by assumption. Thus there is a subsequence (n_k) such that $z_{n_k}(\omega, s, Y)$ converges. By passing to another subsequence, again denoted by n_k, $z_{n_k}(\omega, s, Y)$ converges for all $Y \in \mathcal{L}_N^0$ and $s \in \mathbb{T}$.
Now since $L^1(\Omega, N)$ is separable, there is a countable set $\mathcal{L}_N^0 \subset \mathcal{L}_N$ such that for all $s \in \mathbb{T}$

$$\inf\{\mathbf{E}d(Y_s, \tilde{Y}_s) : \tilde{Y} \in \mathcal{L}_N^0\} = 0.$$

and hence by (13), a standard $\epsilon/3-$argument yields that $z_{n_k}(\omega, s, Y)$ converges for all $Y \in \mathcal{L}_N$ and $s \in \mathbb{T}$. Finally, since Ω_1 is countable, we find a subsequence, again denoted by (n_k), such that $z_{n_k}(\tilde{\omega}, s, Y)$ converges for all $\tilde{\omega} \in \Omega_1$, $Y \in \mathcal{L}_N$ and $s \in \mathbb{T}$. By Lemma 2.1 we have

$$d(P_{\delta_k}^{t/\delta_k} Y(\omega_1, s), P_{\delta_k}^{t/\delta_k} Y(\omega_2, s)) \leq C\rho_s(\omega_1, \omega_2)$$

for all $k \in \mathbb{N}$ and $\omega_1, \omega_2 \in \Omega_0$. Again an $\epsilon/3-$argument yields that $P_{\delta_k}^{t/\delta_k} Y(\omega, s)$ converges for all $\omega \in \Omega_0$, $Y \in \mathcal{L}_N$ and $s \in \mathbb{T}$.
Now for any $t \in \mathbb{T}$, $P_t Y$ is a martingale by construction. The semigroup property follows from the projectivity of the filtered conditional expectation.$\square$

Example 2.3 The previous Theorem becomes a bit clearer when the filtered probability space comes from a Markov process. For instance, consider the situation of Example 1.6. Assume that there exists a $\kappa \in \mathbb{R}$ such that

$$\rho^W(p_t(x), p_t(y)) \leq e^{\kappa t} \rho(x, y) \qquad \forall x, y \in M, \forall t > 0 \tag{14}$$

Now put $\rho_t(\omega, \tilde{\omega}) := e^{-\kappa t} \rho(X_t(\omega), X_t(\tilde{\omega}))$. If the Markov semigroup is regular enough (e.g. if X has continuous paths), then Ω can be chosen to be a polish space such that also (9) is fulfilled. Since X is a Markov process, we have $Q_s(\omega) \circ X_{t+s}^{-1} = p_t(X_s(\omega), \cdot)$ and hence (10) holds. Moreover, for all $f \in Lip(M, N)$ the process $Y_t := f(X_t)$ satisfies (12). Thus by Theorem 2.2, for all $t \in \mathbb{T}$ and $x \in M$, $P_t^* Y$ exists under P^x,the probability measure for the Markov process starting at $x \in M$. If we set $p_t^* f(x) := (P_t^* Y)_0$ (under P^x), then $(p_t^*)_{t \geq 0}$ is a semigroup acting on $Lip(M, N)$ which is called the nonlinear semigroup associated with p_t. In [Stu05] it is studied in great detail. Geometrically, condition (14) can be regarded as a kind of lower curvature bound. For instance, in [vRS04] was shown that if (M, ρ) is a Riemannian manifold and p_t the heat kernel on M, then (14) holds with κ if and only if $Ric_M \geq -\kappa$.

With the same technique we can prove the existence of martingales with a given terminal value Z. More precisely, fix $T \in \mathbb{T}$ and let $Z \in L^1(\mathcal{F}_T, N)$ such that there is some $C > 0$ such that for almost all $\omega_1, \omega_2 \in \Omega$,

$$d(Z(\omega_1), Z(\omega_2)) \leq C \rho_T(\omega_1, \omega_2) \tag{15}$$

Lemma 2.4 *Let $Z \in L^1(\mathcal{F}_T, N)$ satisfying (15). Put $Y_t := Z$ if $t \geq T$ and $Y_t \equiv y_0 \in N$ if $t < T$. Then for all $s \in \mathbb{T}$ with $s \leq T$ and all $t \in \mathbb{T}$, $n \in \mathbb{N}$ and almost all $\omega_1, \omega_2 \in \Omega$,*

$$d((P_t^n Y)_s(\omega_1), (P_t^n Y)_s(\omega_2)) \leq \rho_s(\omega_1, \omega_2).$$

Proof : Again we first prove the assertion for $n = 1$. Let $s, t \in \mathbb{T}$ such that $s \leq T$. If $s + t < T$, then $(P_t Y)_s \equiv y_0$. Moreover, if $s + t \geq T$, then

$$\begin{aligned} d^W(Q_s(\omega_1) \circ Y_{t+s}^{-1}, Q_s(\omega_2) \circ Y_{t+s}^{-1}) &\leq C \rho_T^W(Q_s(\omega_1), Q_s(\omega_2)) \\ &\leq C \rho_s(\omega_1, \omega_2). \end{aligned}$$

An iteration in n now proves the Lemma. $\square$

Theorem 2.5 *Under the assumptions of Theorem 2.2, there is a subsequence such that for all $T \in \mathbb{T}$, all $s \in \mathbb{T}$ with $s \leq T$ and all $Z \in L^1(\mathcal{F}_T, N)$ satisfying (15), $\mathbf{E}^{(\mathcal{F}_\tau)_{\tau \geq s}}[Z]$ exists w.r.t. the sequence of partitions Δ^{δ_k}. In other words, for any such Z there is martingale $(X_s)_{s \in \mathbb{T} \cap [0,T]}$ with $X_T = Z$.*

Proof : Put $Y_t := Z$ if $t \geq T$ and $Y_t \equiv y_0 \in N$ if $t < T$. We can now imitate the proof of Theorem 2.2 except that we use Lemma 2.4 instead of Lemma 2.1. $\square$

Example 2.6 (i) Let $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, P)$ be the canonical Wiener space on $\mathbb{R}^d$. Fix $T \in \mathbb{R}_+$ and let $X_t(\omega) := \omega(t \wedge T)$ be the stopped process. Put

$$\rho_t(\omega_1, \omega_2) := \begin{cases} u\left(\frac{\|X_t(\omega_1) - X_t(\omega_2)\|}{2\sqrt{T-t}}\right) & \text{if } t < T \\ 1 & \text{if } t \geq T. \end{cases} \tag{16}$$

where $\|\cdot\|$ denotes the Euclidean norm and $u(r) := \frac{2}{\sqrt{\pi}} \int_0^r e^{-y^2/2} dy$.
Then condition (15) is equivalent to

$$Z \in L^\infty(\mathcal{F}_T, N).$$

Moreover, (10) is satisfied for all $s \le t \in \mathbb{R}_+$. Indeed, define metrics d_t on $\mathbb{R}^k$ by $d_t(x_1, x_2) := u(\frac{\|x_1 - x_2\|}{2t})$ for $t > 0$ and $d_t(x_1, x_2) := 1$ for $t \le 0$. Then according to [Stu05], Ex. 4.6. (i), equation (22)

$$d_s^W(p_t(x_1), p_t(x_2)) \le d_{s+t}(x_1, x_2)$$

for all $s, t \in \mathbb{R}_+$ and $x_1, x_2 \in \mathbb{R}^d$. Moreover, $\rho_t(\omega_1, \omega_2) = d_{T-t}(X_t(\omega_1), X_t(\omega_2))$ and hence for all $s, t \in \mathbb{T}$ with $s < T$

$$\begin{aligned} \rho_{s+t}^W(Q_s(\omega_1), Q_s(\omega_2)) &= d_{T-(s+t)}^w(Q_s(\omega_1) \circ X_{s+t}^{-1}, Q_s(\omega_2) \circ X_{s+t}^{-1}) \\ &= d_{T-(s+t)}^w(p_t(X_s(\omega_1)), p_t(X_s(\omega_2))) \\ &\le d_{T-s}(X_s(\omega_1), X_s(\omega_2)) = \rho_s(\omega_1, \omega_2). \end{aligned}$$

Note that (10) is trivially satisfied for $s \ge T$. Thus by Theorem 2.5, for all $Z \in L^\infty(\mathcal{F}_T, N)$ there exists a martingale X with $X_T = Z$.

(ii) Analogous arguments apply if $(\Omega, (\mathcal{F}_t)_{t \ge 0}, \mathcal{F}, P)$ is the standard path space for a Lévy process on $\mathbb{R}^d$ with generator $A = -\Psi(-\frac{1}{2}\Delta)$, where Ψ is some Lévy function on $\mathbb{R}_+$ with $\Psi(0) = 0$. In this case, the pseudo metrics ρ_t will be chosen as $\rho(\omega_1, \omega_2) := 2u(T - t, \frac{\|X_t(\omega_1), X_t(\omega_2)\|}{2})$, where $u(s, r) = \int_{\mathbb{R}^d} \mathbf{1}_{[0,r]}(y_1) k_s(y) dy$ and the transition kernel $(k_s)_{s \ge 0}$ is given by its Fourier transform: $\hat{k}_s(y) = \exp(-s\Psi(\|y\|^2/2)$, cf [Stu05] Ex. 4.6. (iii).

Remark 2.7 (i) Although in the two theorems of this section N was supposed to be an NPC space with the canonical barycenter, in the proofs we only used the contraction property of this barycenter, cf Definition 1.3. Thus these Theorems are also valid if N is more generally a proper metric space with a barycenter map $b : \mathcal{P}^1(N) \to N$.

(ii) Picard ([Pic04]) uses similar techniques in order to prove the existence of martingales with prescribed terminal value of the form $Y = f(X_T)$, where X is a Markov process on a metric space M satisfying a certain coupling condition and $f : M \to N$ is uniformly continuous and bounded. Although Picard stated his result only for trees, the techniques also apply in our context in order to prove martingales along subsequences in general NPC spaces (or proper metric spaces with barycenter).

Lemma 2.8 *Let $(M, \mathcal{F})$, $(N, \mathcal{G})$ be measurable spaces such that M is a polish space and $\mathcal{F} \subset \mathcal{B}(M)$. Let $f : M \to N$ and $h : N \times N \to \mathbb{R}$ be measurable maps. Define $h_f(x, y) := h(f(x), f(y))$. Then for all $p^1, p^2 \in \mathcal{P}(M)$ such that h_f is intergrable for some $\mu \in \mathcal{M}(p^1, p^2)$ we have*

$$h_f^W(p^1, p^2) = h^W(p^1 \circ f^{-1}, p^2 \circ f^{-1}).$$

Proof : Let $\mu \in \mathcal{M}(p^1, p^2)$. Then $\mu \circ (f^{-1}, f^{-1}) \in \mathcal{M}(p^1 \circ f^{-1}, p^2 \circ f^{-1})$ and

$$\int_{N \times N} h(y_1, y_2)[\mu \circ (f^{-1}, f^{-1})](d(y_1, y_2)) = \int_{M \times M} h_f(x_1, x_2)\mu(d(x_1, x_2))$$

which implies that $h_f^W(p^1, p^2) \ge h^W(p^1 \circ f^{-1}, p^2 \circ f^{-1})$.
For the other inequality put $\mathcal{A} := f^{-1}(\mathcal{G})$. Then for $i = 1, 2$, the regular conditional probability

$p^i_{id|\mathcal{A}}$ exists, since M is polish. Let $A \in \mathcal{B}(M)$. Since $x \mapsto p^i_{id|\mathcal{A}}(x, A)$ is $\mathcal{A}-$measurable, there is a $\mathcal{G}-$measurable map $K^i_A : N \to [0, 1]$ such that $p^i_{id|\mathcal{A}}(x, A) = K_A(f(x))$ for all $x \in M$. Note that K^i is a Markov kernel from $(N, \mathcal{G})$ to Ω. Let now $\nu \in \mathcal{M}(p^1 \circ f^{-1}, p^2 \circ f^{-1})$. We define a probability measure μ on $M \times M$ by

$$\mu(A_1 \times A_2) := \int_{N\times N} K^1_{A_1}(y_1)K^2_{A_2}(y_2)\nu(d(y_1, y_2)).$$

It is easy to see that $\mu \in \mathcal{M}(p^1, p^2)$. Moreover, since $K^i_{f^{-1}(B)}(y) = \mathbf{1}_B(y)$ for $(p^i \circ f^{-1})-$ almost all $y \in N$ and all $B \in \mathcal{G}$, it follows that $\mu = \nu \circ (f^{-1}, f^{-1})$ and hence

$$\int_{M\times M} h(f(x_1), f(x_2))\mu(d(x_1, x_2)) = \int_{N\times N} h(y_1, y_2)\nu(d(y_1, y_2)).$$

Thus, $h^W_f(p^1, p^2) \le h^W(p^1 \circ f^{-1}, p^2 \circ f^{-1})$. $\square$

2.2 Lower Curvature Bounds

Analogously to the case of upper curvature bounds, lower curvature bounds will be defined by comparing triangles. Let us briefly sketch the definition. Let $z \in N$ and $\gamma : [a, b] \to N$ be a unit-speed geodesic. z and γ span a triangle. Let $\overline{z}$ and $\overline{\gamma}$ be a comparison triangle in $\mathbb{H}_\kappa$ (the Hyperbolic plane of constant curvature $-\kappa$), i.e. $d(\gamma(i), z) = d(\overline{\gamma}(i), \overline{z})$, $i = a, b$ and $d(\gamma(a), \gamma(b)) = d(\overline{\gamma}(a), \overline{\gamma}(b))$ (such a comparison triangle always exists, cf [BH99]). Then $Curv(N) \ge -\kappa$ means nothing else but $d(\gamma(t), z) \ge d(\overline{\gamma}(t), \overline{z})$ for all $z \in N$, all geodesics $\gamma : [a, b] \to N$ and $t \in [a, b]$.
In the above situation define a function $g_{a,b} : [a, b] \to \mathbb{R}$ by $g_{a,b}(t) := d^2(\overline{\gamma}(t), \overline{z})$ (here the distance is taken in $\mathbb{H}_\kappa$). We have $g_{a,b}(t) \le f^z(\gamma(t))$ for all $t \in [a, b]$ with equality if $t = a$ and $t = b$. Moreover, it follows from Riemannian comparison theorems (c.f. [Jos02], Thm. 4.6.1) that for $R > 0$ there is a $C = C(R)$ such that whenever $\gamma([a, b]) \subset B_R(z)$, then

$$0 \le \frac{1}{2} g''_{a,b}(t) - 1 \le C d^2(\overline{\gamma}(t), \overline{z}) \le C d^2(\gamma(t), z). \tag{17}$$

In particular, $g''_{a,b}$ is uniformly bounded for all geodesics which are contained in $B_R(z)$.

Lemma 2.9 *Let $\varphi : \mathbb{R} \to \mathbb{R}$ be a convex function such that for all a, b with $a < b$ there is a smooth function $g_{a,b} : [a, b] \to \mathbb{R}$ such that $g_{a,b}(t) \le \varphi(t)$ for all $t \in [a, b]$ and $g_{a,b}(a) = \varphi(a)$, $g_{a,b}(b) = \varphi(b)$.*
Let I be an open interval and $c : I \to [0, \infty[$ such that $|g''_{a,b}(t)| \le c(t)$ for all $a, b \in I$ and all $t \in [a, b]$. Moreover, let

$$D := \sup\{|g'''_{a,b}(t)| : a, b \in I; \quad t \in [a, b]\} < \infty.$$

Then φ is differentiable in I and we have for all $s, t \in I$

$$|\varphi(s) - \varphi(t) - \varphi'(t)(s - t)| \le \frac{1}{2} c(t)(t - s)^2 + D|t - s|^3$$

Proof : Let $\epsilon > 0$. We can assume that $0 \in I$ and, by adding an affine function if necessary, that $\varphi(0) = 0$ and $\varphi(t) \ge 0$ for all $t \in [-\epsilon, \epsilon]$. We show that φ is differentiable in 0. Since φ is

convex, the one-sided derivatives $\varphi'^{+}(0)$ and $\varphi'^{-}(0)$ exist and $a := \varphi'^{+} - \varphi'^{-} \geq 0$. Thus φ is differentiable in 0 if and only if $a = 0$.
Assume that $a > 0$. Then $0 = \varphi(0) \leq \frac{1}{2}(\varphi(-\epsilon) + \varphi(\epsilon)) - \frac{a}{2}\epsilon$. But $g_{-\epsilon,\epsilon}(0) \leq 0$ and hence there is a $\xi \in]-\epsilon, \epsilon[$ such that $g'_{-\epsilon,\epsilon}(\xi) = 0$. Thus by the Taylor formula

$$\varphi(\epsilon) \leq g_{-\epsilon,\epsilon}(\epsilon) - g_{-\epsilon,\epsilon}(\xi) \leq \frac{1}{2}c(\xi)(\epsilon - \xi)^2 \leq 2C\epsilon^2.$$

The same holds for $-\epsilon$. Letting $\epsilon \to 0$ yields a contradiction. Hence, $a = 0$ and so φ is differentiable in 0.
Now we prove the second claim. Let $t = 0$. Again we can add an affine function and can hence assume that $\varphi(0) = 0$ and $\varphi'(0) = 0$. Let $s \in I$. Then Taylor's formula yields

$$\begin{aligned}\varphi(s) = g_{0,\epsilon}(s) - g_{0,\epsilon}(0) &\leq g'_{0,\epsilon}(0)s + \frac{1}{2}c(0)s^2 + \frac{D}{3}|s|^3 \\ &\leq \frac{1}{2}c(0)s^2 + \frac{D}{3}|s|^3\end{aligned}$$

because $g'_{0,\epsilon}(0)s \leq 0$. □

Recall that the functions f^z from (4) are convex. Hence we can define $df^z_x(y) := \lim_{t\to 0} \frac{1}{t}(f^z(\gamma(t)) - f^z(x))$, where $\gamma : [0,1] \to N$ is the unique geodesic from x to y.

Corollary 2.10 *Let N be a geodesically complete NPC space of lower bounded curvature on all balls, i.e. for all $z \in N$ and all $R > 0$ there is a $\kappa > 0$ such that $Curv(B_R(z)) \geq -\kappa$. Let $z \in N$. Then f^z, is differentiable along geodesics, i.e. for all geodesics $\gamma : \mathbb{R} \to N$ the map $f^z \circ \gamma$ is differentiable. Moreover, for all $z_0 \in N$ and all $R > 0$ there is a $C > 0$ such that for all $x, y, z \in B_R(z_0)$*

$$f^z(y) - f^z(x) - df^z_x(y) - d^2(x,y) \leq Cd^2(x,z)d^2(x,y) + Cd^3(x,y)$$

Proof : Let $x, y, z \in B_R(z_0)$. Let $\gamma : \mathbb{R} \to N$ be the (unit-speed) geodesic with $\gamma(0) = x$ and $\gamma(d(x,y)) = y$. Let $I := \gamma^{-1}(B_R(z_0))$. Then we can apply the preceding Lemma to $\varphi := f^z \circ \gamma$. Note that $df^z_x(y) = \varphi'(0)d(x,y)$. Moreover, by (17) we can choose $c(t) = Cd^2(z, \gamma(t)) + 1$. Putting this together proves the Corollary. □

Lemma 2.11 *Let N be an NPC space of lower bounded curvature on all balls. Let $X \in L^2(\mathcal{F})$. Then for all $z \in N$*

(i)

$$\mathbf{E}^{\mathcal{G}}[df^Z_{\mathbf{E}^{\mathcal{G}}[X]}(X)] = 0 \qquad P\text{-}a.s.$$

(ii) Let $z_0 \in N$ and $R > 0$ such that $X(\Omega) \subset B_R(z_0)$. Then there is a $C > 0$ such that for all $z \in B_R(z_0)$

$$\begin{aligned}\mathbf{E}^{\mathcal{G}}[d^2(X,z) - &d^2(\mathbf{E}^{\mathcal{G}}[X], z) - d^2(X, \mathbf{E}^{\mathcal{G}}[X])] \\ &\leq C\mathbf{E}^{\mathcal{G}}[d^2(\mathbf{E}^{\mathcal{G}}[X], z)d^2(X, \mathbf{E}^{\mathcal{G}}[X]) + Cd^3(X, \mathbf{E}^{\mathcal{G}}[X])]\end{aligned}$$

Proof : Fix $z \in N$. For $y \in N$ let $\gamma_y : [0,1] \to N$ be the geodesic from y to z. Since N is geodesically complete, we can extend it to a geodesic $\gamma : \mathbb{R} \to N$. Moreover, the map $y \mapsto \gamma_y(t)$ is continuous for all $t \in \mathbb{R}$.
Now put $Y := \mathbf{E}^{\mathcal{G}}[X]$. Then $\gamma_Y(t) \in L^2(\mathcal{G}, N)$ for all t, hence γ_Y is a geodesic in $L^2(\mathcal{G}, N)$.

Since $t \mapsto d^2(X(\omega), \gamma_{Y(\omega)}(t))$ is differentiable for all ω by Corollary 2.10 (i), so is the map $t \mapsto \varphi_A(t) := \int_A d^2(X(\omega), \gamma_{Y(\omega)}(t))P(d\omega)$ for all $A \in \mathcal{G}$. Moreover, since $\gamma_Y(0) = \mathbf{E}^{\mathcal{G}}[X]$, 0 is the minimizer of φ_A and hence

$$\int_A df_Y^z(X)dP = \int_A df_Y^X(z)dP = \varphi'_A(0) = 0$$

for all $A \in \mathcal{G}$, proving (i).
(ii) follows from (i) and Corollary 2.10. □

For the rest of this section let $(\Delta^n)_{n\in\mathbb{N}}$ be a refining sequence of partitions such that the mesh converges to 0 as n tends to infinity. Put $\mathbb{T} := \bigcup_{n\in\mathbb{N}} \Delta^n$. The next Lemma will give a sufficient condition for the existence of continuous-time FCE's. Again we will use the notation of (7) and define

$$v_t^{n,m} := \sum_{t_k\in\Delta^n} \mathbf{E}d^2(\xi^m_{t_k\wedge t}, \xi^m_{t_{k+1}\wedge t}). \tag{18}$$

Let $n \leq m$. Let $\Delta^n = \{t_k : k = 0, \dots K^n\}$. Then an iterated application of the Variance Inequality yields that for all $Z \in L^2(\mathcal{F}_{t_{k-1}})$

$$\mathbf{E}d^2(\xi^{\Delta^m}_{t_k}, Z) - \mathbf{E}d^2(\xi^{\Delta^m}_{t_{k-1}}, Z) \geq v^{m,m}_{t_k} - v^{m,m}_{t_{k-1}}. \tag{19}$$

Lemma 2.12 *Let $s < t$. If*

$$\limsup_{n\to\infty} \limsup_{m\to\infty} v_t^{n,m} - v_s^{n,m} - (v_t^{m,m} - v_s^{m,m}) \leq 0,$$

then $\xi^n_s \to \mathbf{E}^{(\mathcal{F}_\tau)_{\tau\geq s}}[\xi]$ in L^2.

Proof : Let $\Delta^n = \{t_k : k = 0, \dots K\}$. Let $k \in \{0, \dots K\}$. Then (19) implies that

$$\mathbf{E}d^2(\xi^n_{t_{k-1}}, \xi^m_{t_{k-1}}) \leq \mathbf{E}d^2(\xi^n_{t_{k-1}}, \xi^m_{t_k}) - (v^{m,m}_{t_k} - V^{m,m}_{t_{k-1}})$$

and

$$\mathbf{E}d^2(\xi^n_{t_{k-1}}, \xi^m_{t_{k-1}}) \leq \mathbf{E}d^2(\xi^n_{t_k}, \xi^m_{t_{k-1}}) - \mathbf{E}d^2(\xi^n_{t_{k-1}}, \xi^n_{t_k}).$$

Adding up the two inequalities and applying the quadruple inequality (2) yields

$$\begin{aligned}
2\mathbf{E}d^2(\xi^n_{t_{k-1}}, \xi^m_{t_{k-1}}) &\leq \mathbf{E}d^2(\xi^n_{t_{k-1}}, \xi^m_{t_k}) + \mathbf{E}d^2(\xi^n_{t_k}, \xi^m_{t_{k-1}}) \\
&\qquad - \mathbf{E}d^2(\xi^n_{t_{k-1}}, \xi^n_{t_k}) - (v^{m,m}_{t_k} - v^{m,m}_{t_{k-1}}) \\
&\leq \mathbf{E}d^2(\xi^n_{t_k}, \xi^m_{t_k}) + \mathbf{E}d^2(\xi^n_{t_{k-1}}, \xi^m_{t_{k-1}}) \\
&\qquad + \mathbf{E}d^2(\xi^m_{t_{k-1}}, \xi^m_{t_k}) - (v^{m,m}_{t_k} - v^{m,m}_{t_{k-1}})
\end{aligned}$$

and hence

$$\mathbf{E}d^2(\xi^n_{t_{k-1}}, \xi^m_{t_{k-1}}) \leq \mathbf{E}d^2(\xi^n_{t_k}, \xi^m_{t_k}) + \mathbf{E}d^2(\xi^m_{t_{k-1}}, \xi^m_{t_k}) - (v^{m,m}_{t_k} - v^{m,m}_{t_{k-1}})$$

By iteration we get for all $n \leq m$

$$\mathbf{E}d^2(\xi^n_s, \xi^m_s) \leq v_t^{n,m} - v_s^{n,m} - (v_t^{m,m} - v_s^{m,m}).$$

Thus, by assumption, ξ^n_s is a Cauchy sequence for all $s \in \mathbb{T}$, converging to some $\xi_s \in L^2(\mathcal{F}_s)$ which is, by definition, equal to $\mathbf{E}^{(\mathcal{F}_\tau)_{\tau\geq s}}[\xi]$. □

Definition 2.13 *Let $\xi_t \in L^2(\mathcal{F}_t)$. ξ is called* regular *for $(\Delta^n)_{n\in\mathbb{N}}$ if the increments of v^m, m_t are finally controlled by a continuous function, i.e. there is a continuous function v_t such that for all $s, t \in \mathbb{T}$,*

$$\limsup_{m\to\infty} v_t^{m,m} - v_s^{m,m} \leq v_t - v_s$$

Theorem 2.14 *Let N be a geodesically complete NPC space of lower bounded curvature on all balls. Let $(\Omega, (\mathcal{F}_t)_{0\leq t\leq T}, \mathcal{F}, P)$ be a filtered probability space. Let $t \in \mathbb{T}$ and $\xi \in L^\infty(\mathcal{F}_t, N)$ be regular for $(\Delta^n)_{n\in\mathbb{N}}$. Then $\mathbf{E}^{(\mathcal{F}_\tau)_{\tau\geq s}}[\xi]$ exists for all $s \in \mathbb{T}$.*

Proof : Recall the notations of Lemma 2.12. Let $n \leq m$. Let $\Delta^n \cap [s,t] = \{t_k : k = 0, \dots K\}$ and $\Delta^m \cap [s,t] = \{s_l : l = 0, \dots L\}$.

$$\begin{aligned}
v_{t_k}^{n,m} - v_{t_{k-1}}^{n,m} &= \mathbf{E}d^2(\xi_{t_k}^m, \xi_{t_{k-1}}^m) = \sum_{t_{k-1}<s_l\leq t_k} \mathbf{E}d^2(\xi_{s_l}^m, \xi_{t_{k-1}}^m) - \mathbf{E}d^2(\xi_{s_{l-1}}^m, \xi_{t_{k-1}}^m) \\
&\leq \sum_{t_{k-1}<s_l\leq t_k} \mathbf{E}d^2(\xi_{s_l}^m, \xi_{s_{l-1}}^m) + C \sum_{t_{k-1}<s_l\leq t_k} \mathbf{E}d^2(\xi_{s_l}^m, \xi_{t_{k-1}}^m)\mathbf{E}d^2(\xi_{s_l}^m, \xi_{s_{l-1}}^m) \\
&\quad + C \sum_{t_{k-1}<s_l\leq t_k} \mathbf{E}[d^3(\xi_{s_l}^m, \xi_{s_{l-1}}^m)] \\
&= V_{t_k}^{m,m} - V_{t_{k-1}}^{m,m} + C \sum_{t_{k-1}<s_l\leq t_k} \mathbf{E}d^2(\xi_{s_l}^m, \xi_{t_{k-1}}^m)\mathbf{E}d^2(\xi_{s_l}^m, \xi_{s_{l-1}}^m) \\
&\quad + C \sum_{s_l} \mathbf{E}[d^3(\xi_{s_l}^m, \xi_{s_{l-1}}^m)] \\
&\leq \tilde{C}(V_{t_k}^{m,m} - V_{t_{k-1}}^{m,m}).
\end{aligned}$$

Since ξ is regular, $\epsilon^n := \sup_{t_k\in\Delta^n, m\geq n} d^2(\xi_{t_k}^m, \xi_{t_{k-1}}^m)$ tends to 0 as $n \to \infty$. So, looking again at the first inequality, we have

$$\begin{aligned}
V_t^{n,m} - V_s^{n,m} \leq V_t^{m,m} - V_s^{m,m} &+ C \sum_{t_k} \sum_{t_{k-1}<s_l\leq t_k} \mathbf{E}d^2(\xi_{s_l}^m, \xi_{t_{k-1}}^m)\mathbf{E}d^2(\xi_{s_l}^m, \xi_{s_{l-1}}^m) \\
&+ C \sum_{s_l} \mathbf{E}[d^3(\xi_{s_l}^m, \xi_{s_{l-1}}^m)]
\end{aligned}$$

The second sum tends to 0 for $n \to \infty$ by the preceding considerations. Clearly, the third sum goes to 0, too. Hence the assumptions of Lemma 2.12 are satisfied and it follows that $\mathbf{E}^{(\mathcal{F}_\tau)_{\tau\geq s}}[\xi]$ exists. $\square$

3 Characterization of martingales

The next Theorem gives a characterization of martingales in terms of their 'quadratic variation'. The prove will use similar techniques as those in Lemma 2.12. Again let $(\Delta^n)_{n\in\mathbb{N}}$ be a refining sequence of partitions such that the mesh converges to 0 as n tends to infinity. Put $\mathbb{T} := \bigcup_{n\in\mathbb{N}} \Delta^n$.

Definition 3.1 *We say that a process $(X_t)_{t\in\mathbb{T}}$ has a quadratic variation if there is a nondecreasing process $(\langle X\rangle_t)_{t\in\mathbb{T}}$ such that for all $t \in \mathbb{T}$, $X_t \in L^2(\mathcal{F}_t)$ and*

$$V_t^n := \sum_{t_k\in\Delta^n} \mathbf{E}^{\mathcal{F}_{t_k}}[d^2(X_{t_k\wedge t}, X_{t_{k+1}\wedge t})] \to \langle X\rangle_t$$

in L^1 as $n \to \infty$.

Theorem 3.2 *Let N be a separable NPC space and $(\Omega, (\mathcal{F}_t)_{t\in\mathbb{T}}, \mathcal{F}, P)$ be a filtered probability space. Let $(X_t)_{t\in\mathbb{T}}$ be an adapted process with quadratic variation $\langle X\rangle$. Then X is a martingale if and only if $d^2(X_t, z) - \langle X\rangle_t$ is a submartingale for all $z \in N$.*

Proof : The 'only if'-implication follows from Proposition 1.7.
For the 'if'-implication we first remark that since $d^2(X_t, z) - \langle X\rangle_t$ is a submartingale for all $z \in N$ and N is separable, it follows that for all $s < t$ and $Z \in L^2(\mathcal{F}_s, N)$

$$\mathbf{E}^{\mathcal{F}_s}[d^2(X_t, Z) - d^2(X_s, Z) - (\langle X\rangle_t - \langle X\rangle_s)] \geq 0 \tag{20}$$

Let $s, t \in \mathbb{T}$ with $s < t$. Put $\xi := X_t$. We have to prove that $\xi^n_s \to X_s$. Let $\Delta^n \cap [s,t] = \{t_0, \dots, t_m\}$ with $s = t_0 < t_1 < \dots < t_m = t$. Using the notation of (7), we have for $k = 1, \dots m$

$$d^2(\xi^n_{t_{k-1}}, X_{t_{k-1}}) \leq \mathbf{E}^{\mathcal{F}_{t_{k-1}}}[d^2(\xi^n_{t_{k-1}}, X_{t_k}) - (\langle X\rangle_{t_k} - \langle X\rangle_{t_{k-1}})]$$

by (20), and the variance inequality yields

$$d^2(\xi^n_{t_{k-1}}, X_{t_{k-1}}) \leq \mathbf{E}^{\mathcal{F}_{t_{k-1}}}[d^2(\xi^n_{t_k}, X_{t_{k-1}}) - d^2(\xi^n_{t_{k-1}}, \xi^n_{t_k})]$$

Adding up the two inequalities and applying the quadruple inequality (2) yields

$$\begin{aligned} 2d^2(\xi^n_{t_{k-1}}, X_{t_{k-1}}) &\leq \mathbf{E}^{\mathcal{F}_{t_{k-1}}}[d^2(\xi^n_{t_{k-1}}, X_{t_k}) + d^2(\xi^n_{t_k}, X_{t_{k-1}}) \\ &\qquad - d^2(\xi^n_{t_{k-1}}, \xi^n_{t_k}) - (\langle X\rangle_{t_k} - \langle X\rangle_{t_{k-1}})] \\ &\leq \mathbf{E}^{\mathcal{F}_{t_{k-1}}}[d^2(\xi^n_{t_k}, X_{t_k}) + d^2(\xi^n_{t_{k-1}}, X_{t_{k-1}}) \\ &\qquad + d^2(X_{t_k}, X_{t_{k-1}}) - (\langle X\rangle_{t_k} - \langle X\rangle_{t_{k-1}})] \end{aligned}$$

and hence

$$d^2(\xi^n_{t_{k-1}}, X_{t_{k-1}}) \leq \mathbf{E}^{\mathcal{F}_{t_{k-1}}}[d^2(\xi^n_{t_k}, X_{t_k}) + (V^n_{t_k} - V^n_{t_{k-1}}) - (\langle X\rangle_{t_k} - \langle X\rangle_{t_{k-1}})]. \tag{21}$$

Iterating this yields

$$\mathbf{E}[d^2(\xi^n_s, X_s)] \leq \mathbf{E}[(V^n_t - V^n_s) - (\langle X\rangle_t - \langle X\rangle_s)]$$

while the right-hand side tends to 0 as n tends to infinity. □

Remark 3.3 (i) From (21) follows that the process $S_k := d^2(\xi^n_{t_k}, X_{t_k}) + V^n_{t_k} - \langle X\rangle_{t_k}$ is a submartingale w.r.t. the filtration $(\mathcal{F}_{t_k})_{0\leq k\leq m}$. Let $\epsilon > 0$. Then

$$\begin{aligned} P(\sup_{0\leq k\leq m} d^2(\xi^n_{t_k}, X_{t_k}) > \epsilon) &\leq P(\sup_{0\leq k\leq m} S_{t_k} > \epsilon) + P(\sup_{0\leq k\leq m} |V^n_{t_k} - \langle X\rangle_{t_k}| > \epsilon) \\ &\leq \frac{1}{\epsilon}\mathbf{E}[|V^n_t - \langle X\rangle_t|] + P(\sup_{0\leq k\leq m} |V^n_{t_k} - \langle X\rangle_{t_k}| > \epsilon) \end{aligned}$$

where the last inequality follows from Doob's inequality. In particular, if $V^n \to \langle X\rangle$ locally uniformly in L^1, then $\xi^n \to X$ locally uniformly in L^2.

(ii) If N is a Riemannian manifold, then Theorem 3.2 yields that every continuous $\nabla-$martingale is (locally) a martingale in our sense. Together with Corollary 1.9 we deduce that a continuous semimartingale X such that $X_t \in L^2(\mathcal{F}_t, N)$ is a $\nabla-$martingale if and only if it is a martingale. □

3.1 Martingales in Stars

The simplest example of a singular NPC space is the n-star (or n-pod), a particular example of a tree. In this section we will use Theorem 3.2 to show that our notion of martingales and Darling's characterization are basically the same for continuous processes.

A star is obtained by gluing n copies of $[0,\infty[$ at 0. Let $R_i := [0,\infty[\times\{i\}$ and $N := (\bigcup_i R_i)_{/\sim}$, where $(\xi,i) \sim (\eta,j)$ if and only if $\xi = \eta = 0$. The intrinsic distance is given by

$$d((\xi,i),(\eta,j)) := \begin{cases} |\xi-\eta| & \text{if } i=j \\ |\xi+\eta| & \text{if } i\neq j. \end{cases}$$

(N,d) is an NPC space. For $x\in N$ we define the $i-$th projection $x^i\in[0,\infty[$ by

$$x^i := \begin{cases} |x| & \text{if } x\in R^i \\ 0 & \text{else} \end{cases}$$

For a function $f: N\to\mathbb{R}$ we define functions $f_i: [0,\infty[\to\mathbb{R}$ by $f_i(\xi) := f((\xi,i))$. Then we have the decomposition

$$f(x) = \sum_i (f_i(x^i) - f(0)). \tag{22}$$

A stochastic calculus on stars was developed in [Pic04], and sufficient conditions for the existence of weak martingales are given. For details we refer to that paper. Here we will briefly quote some notations and results from there.

Let $(\Omega,(\mathcal{F}_t)_{t\geq 0},\mathcal{F},P)$ be a filtered probability space and $X=(X_t)_{t\geq 0}$ be a continuous adapted process. X will be called a *semimartingale* if $\varphi(X)$ is a semimartingale for every convex function $\varphi: N\to\mathbb{R}$. In this case the calculus of local times (c.f. [RY99]) yields

$$\begin{aligned}\varphi(X_t)-\varphi(X_s) &= \sum_i \int_s^t \mathbf{1}_{\{X^i_\tau>0\}} d\varphi_i(X^i_\tau) + \frac{1}{2}\sum_i \varphi_i'^{+}(0)(L^i_t - L^i_s) \\ &= \sum_i \Big[\int_s^t \mathbf{1}_{\{X^i_\tau>0\}} \varphi_i'^{-}(X^i_\tau)dX^i_\tau + \frac{1}{2}\int_{]0,\infty[} (L^{i,a}_t - L^{i,a}_s)\varphi_i''(da)\Big] \\ &\quad + \frac{1}{2}\sum_i \varphi_i'^{+}(0)(L^i_t - L^i_s)\end{aligned}$$

where $L^{i,a}$ denotes the local time at a of the $[0,\infty[-$valued process X^i and $\varphi_i'^{+}(t)$ (resp. $\varphi_i'^{-}(t)$) is the right (resp. left) hand side derivative of φ in t.

Picard defines a (local) martingale to be process such that $\varphi(X)$ is a local submartingale for all convex functions $\varphi: N\to\mathbb{R}$. We will refer to these processes as *weak martingales.* The following Proposition from [Pic04] gives a characterization of weak martingales in terms of a stochastic calculus. We denote with Σ^+ the set of nonnegative local submartingales $Y=M+A$ such that A only increases at $\{Y=0\}$, i.e. $\int_0^t \mathbf{1}_{\{Y_s\neq 0\}}dA_s = 0$.

Proposition 3.4 *Let $(\Omega,(\mathcal{F}_t)_{t\geq 0},\mathcal{F},P)$ be a filtered probability space such and let X be a continuous adapted process. Then X is a weak martingale if and only if $X^i\in\Sigma^+$ for all i and the local times L^i of X^i at 0 satisfy*

$$dL^i_t/dL_t \leq 1/2 \tag{23}$$

with $L_t := \sum_i L^i_t$.

Lemma 3.5 *Let (X_t) be a continuous semimartingale such that $X_t \in L^2(\mathcal{F}_t, N)$ for all t. Put $\langle X\rangle_t := \sum \langle X^i\rangle_t$, where $\langle X^i\rangle$ is the quadratic variation of the $[0,\infty)$- valued semimartingale X^i. Then*

$$\sum_{t_k\in\Delta} \mathbf{E} d^2(X_{t_k}, X_{t_{k+1}}) \to \langle X\rangle_t \qquad \textit{for } \|\Delta\| \to 0.$$

If X is a weak martingale, then for all $s \geq 0$ and all $Z \in L^2(\mathcal{F}_s)$ the process

$$(d^2(Z, X_t) - \langle X\rangle_t)_{t\geq s}$$

is a submartingale.

Proof : We apply the above decomposition to $\varphi(y) := f^{X_s}(y) := d^2(X_s, y)$ and get

$$\begin{aligned} d^2(X_s, X_t) = \sum_i [\int_s^t \mathbf{1}_{\{X_\tau^i>0\}} (f^{X_s})_i'^-(X_\tau^i)dX_\tau^i + (\langle X^i\rangle_t - \langle X^i\rangle_s)] \\ + \frac{1}{2}\sum_i (f^{X_s})_i'^+(0)(L_t^i - L_s^i) \end{aligned}$$

and for a partition

$$\begin{aligned} \sum_{t_k\in\Delta} d^2(X_{t_k}, X_{t_{k+1}}) &= \sum_{t_k\in\Delta}\sum_i \int_{t_k}^{t_{k+1}} \mathbf{1}_{\{X_\tau^i>0\}} (f^{X_{t_k}})_i'^-(X_\tau^i)dX_\tau^i \\ &\quad + \sum_i (\langle X^i\rangle_t] \\ &\quad + \frac{1}{2}\sum_{t_k\in\Delta}\sum_i (f^{X_{t_k}})_i'^+(0)(L_{t_{k+1}}^i - L_{t_k}^i) \\ &= \sum_i [\int_0^t H_\tau^i dX_\tau^i + \int_0^t G_\tau^i dL_\tau^i] + \langle X\rangle_t \end{aligned}$$

with

$$H_\tau^i := \sum_{t_k\in\Delta} \mathbf{1}_{\{X_\tau^i>0\}} (f^{X_{t_k}})_i'^-(X_\tau^i)\mathbf{1}_{]t_k,t_{k+1}]}(\tau)$$

and

$$G_\tau^i := \sum_{t_k\in\Delta} (f^{X_{t_k}})_i'^+(0)\mathbf{1}_{]t_k,t_{k+1}]}(\tau).$$

Since X_t is continuous, H_t tends to 0 uniformly on $[0,T]$ a.s. and so does $\int_0^t H_\tau^i dX_\tau^i$. Moreover, $|G_t^i| = 2|X_t|$ and hence

$$|\int_0^t G_\tau^i dL_\tau^i| \leq 2\int_0^t |X_\tau| dL_\tau = 0$$

which proves the first claim.
For the second assertion we choose $\varphi := f^Z$ which yields

$$\begin{aligned} f^Z(X_t) - f^Z(X_s) = \sum_i [\int_s^t \mathbf{1}_{\{X_\tau^i>0\}} (f^Z)_i'^-(X_\tau^i)dX_\tau^i + \frac{1}{2}(f^Z)_i'^+(0)(L_t^i - L_s^i)] \\ + \langle X\rangle_t - \langle X\rangle_s. \end{aligned}$$

Now since X is a weak martingale, $X^i \in \Sigma^+$ and hence $\int \mathbf{1}_{\{X^i_\tau > 0\}} (f^Z)'^-_i (X^i_\tau) dX^i_\tau$ is a real-valued martingale for all i, which yields

$$\mathbf{E}^{\mathcal{F}_s}[f^Z(X_t) - f^Z(X_s)] = \mathbf{E}^{\mathcal{F}_s}[\frac{1}{2}\sum_i (f^Z)'^+_i(0)(L^i_t - L^i_s) + \langle X\rangle_t - \langle X\rangle_s].$$

Now $|(f^Z)'^+_i(0)| = |(f^Z)'^+_j(0)|$. Moreover, $(f^Z)'^+_j(0) < 0$ only on $\{Z \in R_j\}$. Thus (23) yields that

$$\frac{1}{2}\sum_i (f^Z)'^+_i(0)(L^i_t - L^i_s)(\omega) \geq 0$$

for all $\omega \in \{Z \in R_j\}$. Since j is arbitrary, we get

$$\mathbf{E}^{\mathcal{F}_s}[f^Z(X_t) - f^Z(X_s)] \geq \mathbf{E}^{\mathcal{F}_s}[\langle X\rangle_t - \langle X\rangle_s].\square$$

Theorem 3.6 *Let $(\Omega, (\mathcal{F}_t)_{t\geq 0}, \mathcal{F}, P)$ be a filtered probability space such and let X be a continuous adapted process such that $X_t \in L^2(\mathcal{F}_t, N)$ for all t. Then the following are equivalent:*

(i) *$\varphi(X)$ is a submartingale for any convex function s.t. $\varphi(X_t) \in L^1(\mathcal{F}_t, \mathbb{R})$ for all t.*

(ii) *X is a martingale w.r.t. all sequences (Δ^n) of refining partitions of $[0, \infty[$ such that the mesh goes to 0.*

(iii) *X is a martingale w.r.t. one sequence (Δ^n) of refining partitions of $[0, \infty[$ such that the mesh goes to 0.*

Proof : $(i) \Rightarrow (ii)$ by the preceding Lemma and Theorem 3.2.
$(iii) \Rightarrow (i)$: Let $\mathbb{T} := \bigcup_{n\in\mathbb{N}} \Delta^n$. Then (iii) yields that $\varphi(X_t) \leq \mathbf{E}^{F_s}[\varphi(X_t)]$ for all $s, t \in \mathbb{T}$ with $s \leq t$. By continuity of X we can extend this to all $s \leq t \in [0, \infty[$. $\square$

References

[Arn95] Marc Arnaudon, *Barycentres convexes et approximations des martingales continues dans les variétés*, Séminaire de Probabilités, XXIX, Lecture Notes in Math., vol. 1613, Springer, Berlin, 1995, pp. 70–85.

[Bal95] Werner Ballmann, *Lectures on spaces of nonpositive curvature*, Birkhäuser Verlag, Basel, 1995.

[BH99] Martin R. Bridson and André Haefliger, *Metric spaces of non-positive curvature*, Springer-Verlag, Berlin, 1999.

[Dar82] R. W. R. Darling, *Martingales in manifolds—definition, examples, and behaviour under maps*, Seminar on Probability, XVI, Supplement, Lecture Notes in Math., vol. 921, Springer, Berlin, 1982, pp. 217–236.

[Dos62] Shafik Doss, *Moyennes conditionnelles et martingales dans un espace métrique*, C. R. Acad. Sci. Paris **254** (1962), 3630–3632.

[EF01] J. Eells and B. Fuglede, *Harmonic maps between Riemannian polyhedra*, Cambridge University Press, Cambridge, 2001.

[ÉM91] Michel Émery and G. Mokobodzki, *Sur le barycentre d'une probabilité dans une variété*, Séminaire de Probabilités, XXV, Lecture Notes in Math., vol. 1485, Springer, Berlin, 1991, pp. 220–233.

[ESH99] Aziz Es-Sahib and Henri Heinich, *Barycentre canonique pour un espace métrique à courbure négative*, 355–370.

[Jos97] Jürgen Jost, *Nonpositive curvature: geometric and analytic aspects*, Birkhäuser Verlag, Basel, 1997.

[Jos02] ———, *Riemannian geometry and geometric analysis*, third ed., Universitext, Springer-Verlag, Berlin, 2002.

[Ken90] Wilfrid S. Kendall, *Probability, convexity, and harmonic maps with small image. I. Uniqueness and fine existence*, Proc. London Math. Soc. (3) **61** (1990), no. 2, 371–406.

[Ken92] ———, *The propeller: A counterexample to a conjectured criterion for the existence of certain convex functions*, J. London Math. Soc. **46** (1992), no. 2, 364–374.

[LT91] Michel Ledoux and Michel Talagrand, *Probability in Banach spaces*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3), vol. 23, Springer-Verlag, Berlin, 1991.

[Pic91] Jean Picard, *Martingales on Riemannian manifolds with prescribed limit*, J. Funct. Anal. **99** (1991), no. 2, 223–261.

[Pic04] ———, *Stochastic calculus and martingales on trees*, Preprint (2004).

[RY99] Daniel Revuz and Marc Yor, *Continuous martingales and Brownian motion*, third ed., Springer-Verlag, Berlin, 1999.

[Stu02] Karl-Theodor Sturm, *Nonlinear martingale theory for processes with values in metric spaces of nonpositive curvature*, Ann. Probab. **30** (2002), no. 3, 1195–1222.

[Stu05] K. T. Sturm, *A semigroup approach to harmonic maps*, Potential Anal. **23** (2005), 225–277.

[vRS04] M. K. von Renesse and K. T. Sturm, *Transport inequalities, gradient estimates, entropy and ricci curvature*, Comm. Pure Appl.Math. **58(7)** (2004), 923–940.