Martingales and diffusions, limit theory and statistical inference
van Zanten, J.H.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
In this chapter we study continuous local martingales. There are several excellent textbooks treating the general theory of continuous semimartingales. Let us mention in particular the books by Kallenberg (1997), Karatzas and Shreve (1991), Revuz and Yor (1991) and Rogers and Williams (1987). We assume that the reader is familiar with the classical theory of continuous semimartingales that is treated in those textbooks.

The present chapter deals with aspects of martingale theory that are somewhat less classical. We focus in particular on recently developed results that have important applications in diffusion theory and in asymptotic statistics. Examples of applications of the material can be found throughout part I of this thesis. In part II we will consider the extension of some of the results to the realm of martingales with jumps.

1.1 Continuous local martingales

The most important example of a continuous martingale is of course Brownian motion. If $W$ is a Brownian motion and $(\tau_t)_{t \geq 0}$ is an increasing family of finite stopping times (a so-called time-change) such that the function $t \mapsto \tau_t$ is almost surely continuous, then the optional sampling theorem shows that the time-changed Brownian motion $M_t = W_{\tau_t}$ is a continuous local martingale with quadratic variation process $\langle M \rangle_t = \tau_t$. The well-known Dambis-Dubins-Schwarz theorem states that the converse is also true: every continuous local
Continuous martingales

The only requirement is that the underlying probability space is rich enough.

If the latter requirement is not met, the underlying probability space has to be enlarged. As in Revuz and Yor (1991), we call the filtered probability space \((\tilde{\Omega}, (\tilde{\mathcal{F}}_t), \tilde{\mathcal{F}}, \tilde{P})\) an *enlargement* of the filtered probability space \((\Omega, (\mathcal{F}_t), \mathcal{F}, P)\) if there exists a map \(\pi\) from \(\tilde{\Omega}\) onto \(\Omega\) such that \(\pi^{-1}(\mathcal{F}_t) \subseteq \tilde{\mathcal{F}}_t\) for every \(t\) and \(P = \tilde{P} \circ \pi^{-1}\). A process \(X\) defined on \(\Omega\) may then be viewed as defined on \(\tilde{\Omega}\) by putting \(X(\tilde{\omega}) = X(\omega)\) if \(\pi(\tilde{\omega}) = \omega\). With this notion of enlargement at our disposal we can state the Dambis-Dubins-Schwarz time-change theorem.

**Theorem 1.1.1 (Dambis-Dubins-Schwarz).** Let \(M\) be a continuous local martingale on a filtered probability space \((\Omega, \mathcal{F}, (\mathcal{F}_t), P)\). For every \(t \geq 0\), define \(\tau_t = \inf\{s \geq 0 : \langle M \rangle_s > t\}\) and \(\mathcal{G}_t = \mathcal{F}_{\tau_t}\). If \(\langle M \rangle_\infty = \infty\) almost surely, then there exists a \((\mathcal{G}_t)\)-Brownian motion \(W\) such that \(M_t = W_{\langle M \rangle_t}\) for all \(t \geq 0\). If \(\langle M \rangle_\infty < \infty\) with positive probability, then there exists an enlargement of \((\Omega, \mathcal{F}, (\mathcal{G}_t), P)\), supporting a Brownian motion \(\tilde{W}\), such that \(M_t = \tilde{W}_{\langle M \rangle_t}\) for all \(t \geq 0\).

**Proof.** See for instance Karatzas and Shreve (1991), section 3.4.B or Revuz and Yor (1991), section V.1.

The time-change theorem plays an important role in this chapter. It allows us to reduce a problem concerning general continuous local martingales to the corresponding problem for Brownian motion.

### 1.2 An exponential inequality

In this section we present the first application of the Dambis-Dubins-Schwarz theorem. We prove a Bernstein-type inequality for continuous local martingales. This result will be the starting point of our treatment of entropy methods for martingales in section 1.5. In chapter 5 we treat Bernstein-type inequalities for martingales with jumps.

The random time \(\tau\) in the statement of theorem 1.2.1 below is arbitrary, it is not necessarily a stopping time. This improvement of the classical
exponential inequality for continuous local martingales is a consequence of the time-change method. Our line of reasoning is similar to that of Barlow et al. (1986), who proved Burkholder-Davis-Gundy-type inequalities for continuous local martingales at arbitrary random times.

**Theorem 1.2.1.** Let $M$ be a continuous local martingale and let $\tau$ be a nonnegative random variable defined on the same probability space as $M$. Then we have

$$P\left(\sup_{t \leq \tau} |M_t| \geq z, \langle M \rangle_\tau \leq L\right) \leq 2e^{-\frac{1}{2} \frac{z^2}{L}}$$

for all $z, L > 0$.

**Proof.** By theorem 1.1.1, the underlying probability space can be enlarged to ensure that there exists a Brownian motion $W$ such that $M_t = W_{\langle M \rangle_t}$ for all $t \geq 0$. It then holds that

$$\sup_{t \leq \tau} M_t = \sup_{t \leq \langle M \rangle_\tau} W_t.$$  

Clearly, this implies that

$$P\left(\sup_{t \leq \tau} M_t \geq z, \langle M \rangle_\tau \leq L\right) \leq P\left(\sup_{t \leq L} W_t \geq z\right).$$

The distribution of the running maximum of Brownian motion is well-known, see for instance Karatzas and Shreve (1991), section 2.8.A. The random variable $\sup_{t \leq L} W_t$ has density

$$x \mapsto \frac{2}{\sqrt{2\pi L}} e^{-\frac{1}{2} \frac{x^2}{L}}$$

with respect to the Lebesgue measure on $\mathbb{R}^+$. Using the Gaussian tail bound

$$\frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{1}{2} u^2} du \leq \frac{1}{2} e^{-\frac{1}{2} x^2}, \quad x > 0$$

we thus find that

$$P\left(\sup_{t \leq L} W_t \geq z\right) = \frac{2}{\sqrt{2\pi L}} \int_x^\infty e^{-\frac{1}{2} \frac{x^2}{L}} \, dx$$

$$= \frac{2}{\sqrt{2\pi}} \int_{x/\sqrt{L}}^\infty e^{-\frac{1}{2} y^2} \, dy$$

$$\leq e^{-\frac{1}{2} \frac{z^2}{L}}.$$
Consequently, we have the bound

$$P\left( \sup_{t \leq \tau} M_t \geq z, \langle M \rangle_{\tau} \leq L \right) \leq e^{-\frac{1}{2} \frac{z^2}{L}}$$

for all \( z, L > 0 \). Since the quadratic variation processes of \( M \) and \( -M \) are equal, the same bound holds with \( -M \) in the place of \( M \). Combination of the inequality for \( M \) and the inequality for \( -M \) yields the assertion of the theorem.

\[ \square \]

### 1.3 Nested Brownian motions

The Dambis-Dubins-Schwarz theorem is also useful tool in the study of the central limit problem for continuous local martingales. In the next section we will prove a central limit theorem for \( d \)-dimensional continuous local martingales. If \( M \) is such a martingale, we will assume that there exist deterministic normalizing matrices \( K_t \) such that \( K_t \rightarrow 0 \) as \( t \rightarrow \infty \) and

$$K_t \langle M \rangle_t K_t^T \xrightarrow{P} \eta \eta^T,$$

where \( \eta \) is some random matrix. It will turn out that under this condition, the normalized martingale \( K_t M_t \) converges weakly to a mixture of normals. The degree of difficulty of this problem depends mainly on whether or not the matrix \( \eta \) is random. To be able to handle the case of a random \( \eta \) we treat so-called nested sequences of Brownian motions in this section.

Before we explain this in more detail, let us introduce some notation. By \( C[0, \infty) \) we denote the space of all continuous functions \( f : [0, \infty) \rightarrow \mathbb{R} \), endowed with the local uniform metric

$$d(f, g) = \sum_{n=1}^{\infty} 2^{-n} \left( \max_{t \leq n} |f(t) - g(t)| \land 1 \right).$$

Recall that under this metric, \( C[0, \infty) \) is a Polish space. We can view one-dimensional, continuous random processes as random elements of \( C[0, \infty) \), i.e. as measurable maps on the underlying probability space, with values in \( C[0, \infty) \). Weak convergence and convergence in probability in Polish spaces
are denoted by the symbols $\Rightarrow$ and $\rightarrow^P$, respectively. We refer to Billingsley (1968) for the basic theory of weak convergence in Polish spaces.

Now let $M$ be a one-dimensional continuous local martingale and suppose that there exist deterministic normalizing constants $(k_t)_{t\geq 0}$ such that as $t \to \infty$, we have $k_t \to \infty$ and

$$\frac{\langle M \rangle_t}{k_t} \rightarrow^P \eta$$

where $\eta$ is a nonnegative, real number. Let $W$ be the Brownian motion corresponding to $M$ via theorem 1.1.1. For each $t \geq 0$ define the process $W_t$ by putting $W_t = W_{k_t s}/\sqrt{k_t}$, for all $s \geq 0$. Then the well-known scaling property of Brownian motion implies that each process $W_t$ is again a Brownian motion. We have $M_t = W_{\langle M \rangle_t}$, hence for all $t \geq 0$ it holds that

$$\frac{M_t}{\sqrt{k_t}} = W_t \frac{\langle M \rangle_t}{k_t}.$$  

Each $W_t$ is a Brownian motion, so we trivially have the weak convergence $W_t \Rightarrow B$ in $C[0, \infty)$, where $B$ is a Brownian motion. Since $\eta$ is deterministic, Slutsky's lemma gives the implication

$$\left\{ W_t \Rightarrow B, \quad \frac{\langle M \rangle_t}{k_t} \rightarrow^P \eta \right\} \text{ implies } \left\{ \left( W_t, \frac{\langle M \rangle_t}{k_t} \right) \Rightarrow (B, \eta) \right\}. \tag{1.2}$$

By the continuous mapping theorem it thus follows that for any continuous map $\phi$ on $C[0, \infty) \times \mathbb{R}^+$, we have

$$\phi \left( W_t, \frac{\langle M \rangle_t}{k_t} \right) \Rightarrow \phi(B, \eta). \tag{1.3}$$

In the special case of the map $\phi : C[0, \infty) \times \mathbb{R}^+ \to \mathbb{R}$ defined by $\phi(f, x) = f(x)$ the left-hand side is equal to $M_t/\sqrt{k_t}$ and the right-hand side equals $B_\eta$, so (1.3) yields

$$\frac{M_t}{\sqrt{k_t}} \Rightarrow N(0, \eta).$$

So if $\eta$ in (1.1) is deterministic, the time-change device already gives us a desired result, a central limit theorem for the normalized martingale $M_t/\sqrt{k_t}$. But when $\eta$ is random, the matter is more complicated. Then we can not apply Slutsky's lemma to justify the implication (1.2). Below we will prove that thanks to the special nesting relation between the Brownian
motions $W^t$, they are asymptotically independent of $\langle M \rangle_t/k_t$. This means that in the case of a random $\eta$ the implication (1.2) also holds, with $B$ a Brownian motion that is independent of $\eta$. As a result, $M_t/\sqrt{k_t}$ will then converge weakly to a mixture of normals.

The following theorem states that nested Brownian motions are asymptotically independent of any other random element. The nesting condition is formulated in terms of the corresponding filtrations.

**Theorem 1.3.1.** Let $W^n = (W^n_t, \mathcal{F}^n_t : t \geq 0)$ be a sequence of Brownian motions on a common probability space $(\Omega, \mathcal{F}, P)$. Suppose that for every $n$, there exists an $(\mathcal{F}^n_t)$-stopping time $\tau_n$ such that

(i) $\tau_n \overset{P}{\to} 0$,

(ii) $\mathcal{F}^n_{\tau_n} \subseteq \mathcal{F}^{n+1}_{\tau_{n+1}}$ for every $n$

(iii) $\bigvee_{n=1}^{\infty} \mathcal{F}^n_{\tau_n} = \bigvee_{n=1}^{\infty} \mathcal{F}^\infty_{\tau_n}$.

Then for every random element $X$ on $(\Omega, \mathcal{F}, P)$, with values in a Polish space $(\mathcal{X}, \mathcal{B}^{(\mathcal{X})})$, we have $(W^n, X) \Rightarrow (W, X)$ in $C[0, \infty) \times \mathcal{X}$, where $W$ is a Brownian motion that is independent of $X$.

**Proof.** For every $n$, define the process $V^n$ by

$$V^n_t = W^n_{\tau_{n+t}} - W^n_{\tau_n}.$$ 

Then we can write

$$W^n - V^n = \phi(W^n, \tau_n) - \psi(W^n, \tau_n), \quad (1.4)$$

where $\phi$ and $\psi$ are the continuous maps given by

$$\phi : C[0, \infty) \times [0, \infty) \to \mathbb{R}, \quad \phi(f, t) = f(t),$$

$$\psi : C[0, \infty) \times [0, \infty) \to C[0, \infty), \quad \psi(f, t) = f(t + \cdot) - f(\cdot).$$

Of course, the processes $W^n$ converge weakly to a Brownian motion $W$. Together with assumption (i) this implies that $(W^n, \tau_n) \Rightarrow (W, 0)$. Hence, using (1.4) and the continuous mapping theorem we see that $W^n - V^n \Rightarrow 0$. It thus suffices to show that $(V^n, X) \Rightarrow (W, X)$, where $W$ is a Brownian motion that is independent of $X$.  


1.3 Nested Brownian motions

We will show that for all $W$-continuity sets $A \in \mathcal{B}(C[0,\infty))$ and $X$-continuity sets $B \in \mathcal{B}(\mathcal{X})$, we have

$$P(V^n \in A, X \in B) \to P(W \in A)P(X \in B)$$

(this is sufficient, see theorem 3.1 of Billingsley (1968)). The fact that $W^n - V^n \Rightarrow 0$ implies in particular that $V^n$ converges weakly to a Brownian motion. Hence, by the portmanteau theorem, we have

$$P(V^n \in A) \to P(W \in A)$$

for all $W$-continuity sets $A \in \mathcal{B}(C[0,\infty))$. In view of the inequality

$$|P(V^n \in A, X \in B) - P(W \in A)P(X \in B)| \leq$$

$$|P(V^n \in A, X \in B) - P(V^n \in A)P(X \in B)| +$$

$$|P(V^n \in A)P(X \in B) - P(W \in A)P(X \in B)|$$

it thus remains to show that $|P(V^n \in A, X \in B) - P(V^n \in A)P(X \in B)| \to 0$.

For notational convenience, put $\mathcal{G} = \bigvee_{n=1}^{\infty} \mathcal{F}_n$. From assumptions (ii) and (iii) it follows, by the martingale convergence theorem, that for all $B \in \mathcal{B}(\mathcal{X})$

$$P(X \in B \mid \mathcal{F}_n) \overset{L^1}{\to} P(X \in B \mid \mathcal{G}).$$

Consequently, we have for all $A \in \mathcal{B}(C[0,\infty))$ and $B \in \mathcal{B}(\mathcal{X})$

$$\left| E[1_{\{V^n \in A\}}P(X \in B \mid \mathcal{F}_n)] - E[1_{\{V^n \in A\}}P(X \in B \mid \mathcal{G})] \right| \to 0.$$

By the strong Markov property of Brownian motion, $V^n$ is independent of $\mathcal{F}_n$. This implies that the first expectation in the preceding display is equal to $P(V^n \in A)P(X \in B)$. The $\mathcal{G}$-measurability of $V^n$ implies that the second expectation is equal to $P(V^n \in A, X \in B)$.

In chapter 6 we will need the following corollary of theorem 1.3.1.

**Corollary 1.3.2.** Let $W$ be a Brownian motion and suppose that we have positive numbers $(k_t)_{t \geq 0}$ that increase to infinity as $t \to \infty$. For every $t \geq 0$, define the Brownian motion $W^t$ by $W^t_t = W_t + \sqrt{k_t}$. Then for every random element $X$ on the same probability space, with values in a Polish space $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$, we have $(W^t, X) \Rightarrow (B, X)$ in $C[0,\infty) \times \mathcal{X}$, where $B$ is a Brownian motion that is independent of $X$. 


Proof. Let \( a_n \) be an arbitrary sequence that increases to infinity. If \( W \) is a Brownian motion with respect to the filtration \((\mathcal{F}_t)\), the process \( W^{a_n} \) defined in the statement of the corollary is a Brownian motion with respect to \((\mathcal{F}^n_k) = (\mathcal{F}_{k_n}^n)\). For every \( n \), define the deterministic stopping time \( \tau_n = (k_{a_n})^{-1/2} \). Clearly, the filtrations \((\mathcal{F}^n_k)\) and the stopping times \( \tau_n \) satisfy the conditions of the preceding theorem. We may thus conclude that \((W^{a_n}, X) \Rightarrow (B, X)\), where \( B \) is independent of \( X \). Since the sequence \( a_n \) is arbitrary, the proof is completed.

\[ \square \]

### 1.4 Central limit theorems

Using the time-change theorem and the results of the preceding sections we can prove a number of useful central limit theorems for continuous local martingales. We begin with a simple one that will be used in chapter 3. It is just a special case of the general limit theorems for semimartingales that can be found for instance in Jacod and Shiryaev (1987) or Liptser and Shiryaev (1989). As usual, we denote the \( d \)-dimensional normal distribution with mean 0 and covariance matrix \( \Sigma \) by \( N_d(0, \Sigma) \).

**Theorem 1.4.1.** For every \( t \geq 0 \) and \( i = 1, \ldots, d \) let \( M^{t,i} \) be a continuous local martingale. Suppose that for \( i, j = 1, \ldots, d \) there exist deterministic numbers \( \Sigma_{i,j} \) such that

\[
\langle M^{t,i}, M^{t,j} \rangle_1 \overset{P}{\rightarrow} \Sigma_{i,j}
\]

as \( t \to \infty \). Then \((M^{t,1}, \ldots, M^{t,d}) \Rightarrow N_d(0, \Sigma)\), where \( \Sigma \) is the \( d \times d \) matrix with elements \( \Sigma_{i,j} \).

**Proof.** By the Cramér-Wold device, it suffices to consider the case \( d = 1 \). By theorem 1.1.1 we can enlarge the underlying probability space to ensure that there exist Brownian motions \( W^t \) such that \( M^{t,1} = W^t_{(M^{t,1})_1} \). By the assumption on the quadratic variation processes and Slutsky’s lemma we have \((W^t, \langle M^{t,1} \rangle_1) \Rightarrow (W, \Sigma_{1,1})\), where \( W \) is a Brownian motion. So by the continuous mapping theorem \( M^{t,1} \Rightarrow W_{\Sigma_{1,1}} \). This completes the proof, since \( W_{\Sigma_{1,1}} \) has a \( N(0, \Sigma_{1,1}) \)-distribution.

\[ \square \]
1.4 Central limit theorems

The remainder of this section is devoted to the proof of a central limit theorem for normalized, multivariate continuous local martingales. The first step is a lemma concerning nested, one-dimensional martingales. Again, it will be convenient to formulate the nesting condition in terms of filtrations. Following Feigin (1985), we call a sequence of filtrations \( \mathcal{F}_t^n \) on a probability space \((\Omega, \mathcal{F}, P)\) nested if there exists a sequence \( t_n \downarrow 0 \) such that

\[
\mathcal{F}_{t_n}^n \subseteq \mathcal{F}_{t_{n+1}}^{n+1}
\]

for all \( n \in \mathbb{N} \), and

\[
\bigvee_{n=1}^{\infty} \mathcal{F}_{t_n}^n = \bigvee_{n=1}^{\infty} \mathcal{F}_{t_n}^n.
\]

Any sequence \( t_n \downarrow 0 \) for which these conditions are satisfied is called an \( N \)-sequence. A sequence of adapted processes \( X^n = (X^n_t, \mathcal{F}_t^n : t \geq 0) \) on \((\Omega, \mathcal{F}, P)\) is called nested if the corresponding filtrations \( \mathcal{F}_t^n \) are nested.

**Lemma 1.4.2.** Let \( (M_t^n, \mathcal{F}_t^n : t \geq 0) \) be a nested sequence of continuous local martingales such that \( \langle M^n \rangle_\infty = \infty \) almost surely and assume that there exists an \( N \)-sequence \( t_n \) such that \( \langle M^n \rangle_{t_n} \xrightarrow{P} 0 \). Suppose that for a fixed \( t \geq 0 \) there exists a nonnegative random variable \( C \) such that

\[
\langle M^n \rangle_t \xrightarrow{P} C \tag{1.5}
\]

as \( n \to \infty \). Then for each random element \( X \) defined on the same probability space, with values in some Polish space \( \mathcal{X} \), we have \( (M^n_t, X) \Rightarrow (W_C, X) \) as \( n \to \infty \), where \( W \) is a Brownian motion that is independent of \((C, X)\).

**Proof.** Let \( (W^n_t, \mathcal{G}_t^n : t \geq 0) \) be the Brownian motion corresponding to \( (M^n_t, \mathcal{F}_t^n : t \geq 0) \) via theorem 1.1.1. Then the random variable \( \tau_n = \langle M^n \rangle_{t_n} \) is a \( (\mathcal{G}_t^n) \)-stopping time. By construction, all conditions of theorem 1.3.1 are satisfied. Using also (1.5) the theorem implies that

\[
(W^n, \langle M^n_t \rangle, X) \Rightarrow (W, C, X),
\]

where \( W \) is a Brownian motion that is independent of the pair \((C, X)\). Now write \( (M^n_t, X) = \xi(W^n, \langle M^n_t \rangle, X) \), with \( \xi : C[0, \infty) \times \mathbb{R}^+ \times \mathcal{X} \to \mathbb{R} \times \mathcal{X} \) defined by \( \xi(f, t, x) = (f(t), x) \). Then by the continuous mapping theorem we have \( (M^n_t, X) = \xi(W^n, \langle M^n_t \rangle, X) \Rightarrow \xi(W, C, X) = (W_C, X) \). \( \square \)
We will need a simple result from linear algebra. Euclidean norms of vectors are denoted by $|\cdot|$. If $A$ is an $n \times m$ matrix we write $\|A\| = \sup\{|Ax| : x \in \mathbb{R}^m, |x| = 1\}$ for the operator norm of $A$.

**Lemma 1.4.3.** Let $A_n$ be a sequence of symmetric, nonnegative definite $d \times d$ matrices and suppose that $A_n \rightarrow A$, where $A$ is symmetric and (strictly) positive definite. Then if $x_n$ is a sequence of vectors in $\mathbb{R}^d$ such that $|x_n| \rightarrow \infty$, it holds that $x_n^T A_n x_n \rightarrow \infty$.

**Proof.** Observe that $x_n^T A_n x_n \geq \lambda_n |x_n|^2$, where $\lambda_n$ is the smallest eigenvalue of $A_n$. It thus suffices to show that $\lambda_n \rightarrow \lambda$, where $\lambda > 0$ is the smallest eigenvalue of $A$. Since the determinant is a continuous function and the matrix $A$ is invertible, the matrices $A_n$ are also invertible for $n$ large enough. The matrix map $B \mapsto B^{-1}$ is continuous on the domain of all invertible $d \times d$ matrices. It follows that $A_n^{-1} \rightarrow A^{-1}$ and therefore also $\|A_n^{-1}\| \rightarrow \|A^{-1}\|$. It is easily seen that the operator norm of a symmetric, positive definite matrix is equal to its largest eigenvalue. So we find that the largest eigenvalue of $A_n^{-1}$ converges to the largest eigenvalue of $A^{-1}$. Since the eigenvalues of the inverse $B^{-1}$ of a matrix $B$ are the reciprocals of the eigenvalues of $B$, this implies that the smallest eigenvalue of $\lambda_n$ of $A_n$ converges to the smallest eigenvalue $\lambda$ of $A$. \hfill \box

We can now prove our multivariate central limit theorem for continuous local martingales, which was first presented in the paper Van Zanten (2000a). Recently, Küchler and Sørensen (1999) reported a result that is similar to ours. In their setup $M$ is a multi-dimensional square integrable martingale (not necessarily continuous) with covariance matrices $\Sigma_t = E(M_t M_t^T)$. In addition to our condition (1.6) and a Lindeberg-type condition (if $M$ has jumps), they needed the assumption that there exists a (strictly) positive definite limit of $K_t \Sigma_t K_t^T$ as $t \rightarrow \infty$. The latter condition can be very tedious to verify in applications. This is the case for instance in the asymptotic analysis of the maximum likelihood estimator for multi-dimensional, linear stochastic differential equations (see Küchler and Sørensen (1999), p. 489).

**Theorem 1.4.4.** Let $(M_t, \mathcal{F}_t : t \geq 0)$ be a $d$-dimensional continuous local martingale. Suppose that there exist invertible, deterministic $d \times d$ matrices
(1.4 Central limit theorems)

\((K_t)_{t \geq 0}\) such that \(\|K_t\| \to 0\) as \(t \to \infty\) and

\[
K_t \langle M \rangle_t K_t^T \overset{P}{\to} \eta \eta^T,
\]

(1.6)
where \(\eta\) is an invertible, random \(d \times d\) matrix. Then for each \(\mathbb{R}^k\)-valued random vector \(X\) defined on the same probability space as \(M\) we have \((K_t M_t, X) \Rightarrow (\eta Z, X)\) as \(t \to \infty\), where \(Z\) has a \(N_d(0, I)\)-distribution and \(Z\) is independent of \((\eta, X)\).

**Proof.** First observe that for all \(x \in \mathbb{R}^d\) and \(y \in \mathbb{R}^k\), we have

\[
E e^{ix^T \eta Z + iy^T X} = E e^{-\frac{1}{2} x^T \eta \eta^T x + iy^T X}.
\]

So in terms of characteristic functions we have to prove that for all \(x \in \mathbb{R}^d\) and \(y \in \mathbb{R}^k\)

\[
E e^{ix^T K_t M_t + iy^T X} \to E e^{-\frac{1}{2} x^T \eta \eta^T x + iy^T X}.
\]

That is, we need to prove that

\[
(x^T K_t M_t, Y) \Rightarrow (x^T \eta Z, Y)
\]

(1.7)
for all \(x \in \mathbb{R}^d\) and all real-valued random variables \(Y\), where \(Z\) has a \(N_d(0, I)\)-distribution and \(Z\) is independent of \((\eta, Y)\).

Let \(a_n\) be an arbitrary sequence such that \(a_n \to \infty\). We introduce the one-dimensional continuous processes \(M^n\) by putting

\[
M^n_t = x^T K_{a_n} M_{a_n t}.
\]

Observe that for every \(n \in \mathbb{N}\) the process \(M^n\) is a continuous local martingale with respect to the filtration \((\mathcal{F}_{a_n t})\) and that its quadratic variation is given by

\[
\langle M^n \rangle_t = x^T K_{a_n} \langle M \rangle_{a_n t} K_{a_n t}^T x.
\]

(1.8)
In this notation (1.7) reduces to

\[
(M^n_1, Y) \Rightarrow (x^T \eta Z, Y).
\]

(1.9)
In order to prove (1.9) we will show that every subsequence \(a_{i_n}\) of \(a_n\) has a further subsequence \(a_{k_n}\), such that

\[
(M^n_{k_1}, Y) \Rightarrow (x^T \eta Z, Y).
\]
Continuous martingales

We can choose a subsequence $a_{k_n}$ of $a_{t_n}$ and numbers $0 < t_n \downarrow 0$, so that

$$a_{k_n} t_n \uparrow \infty \quad \text{and} \quad \|K_{a_{k_n}} K_{a_{k_n} t_n}^{-1}\| \to 0. \quad (1.10)$$

Indeed, since $\|K_{a_{t_n}}\| \to 0$ and $1 = \|I\| \leq \|K_{a_{t_n}}\|\|K_{a_{t_n}}^{-1}\|$, we have $\|K_{a_{t_n}}^{-1}\| \to \infty$. So we can choose the subsequence $a_{k_n}$ in such a way that the following inequalities are satisfied:

$$\|K_{a_{k_n}}\| \leq \frac{1}{n\|K_{a_{t_n}}^{-1}\|} \quad \text{and} \quad a_{k_n} \geq n a_{t_n}. \quad (1.11)$$

Now put $t_n = a_{t_n}/a_{k_n}$. By the second of the inequalities we have $t_n \leq 1/n$, so $t_n \downarrow 0$. Moreover, $a_{k_n} t_n = a_{t_n}$, which tends to $\infty$ so that the first condition in (1.10) is satisfied. The second condition in (1.10) is satisfied as well since by the inequality in (1.11)

$$\|K_{a_{k_n}} K_{a_{k_n} t_n}^{-1}\| \leq \|K_{a_{k_n}}\|\|K_{a_{k_n} t_n}^{-1}\| \leq \frac{1}{n},$$

which means that the sequences $a_{k_n}$ and $t_n$ possess the desired properties.

We are going to apply lemma 1.4.2 to the local martingales $M^{k_n}$. We saw already that $M^{k_n}$ is a continuous local martingale with respect to the filtration $(\mathcal{F}_{a_{k_n} t})$, so it is clear that the $M^{k_n}$ are nested. By the first relation in (1.10), $t_n$ is an $N$-sequence. Moreover, by (1.8) we have

$$\| \langle M^{k_n} \rangle_{t_n} \| = \| x^T K_{a_{k_n}} \langle M \rangle_{a_{k_n} t_n} K_{a_{k_n} t_n}^T x \|$$

$$= \| x^T (K_{a_{k_n}} K_{a_{k_n} t_n}^{-1}) K_{a_{k_n} t_n} \langle M \rangle_{a_{k_n} t_n} K_{a_{k_n} t_n}^T (K_{a_{k_n}} K_{a_{k_n} t_n}^{-1})^T x \|$$

$$\leq \|K_{a_{k_n}} K_{a_{k_n} t_n}^{-1}\|^2 \|K_{a_{k_n} t_n} \langle M \rangle_{a_{k_n} t_n} K_{a_{k_n} t_n}^T \| \|x\|^2.$$ 

So it follows by the second relation in (1.10) and by the assumption on $\langle M \rangle$ that $\langle M^{k_n} \rangle_{t_n} \xrightarrow{P} 0$. Also observe that for fixed $n \in \mathbb{N}$

$$\langle M^{k_n} \rangle_{t} = y_t^T K_{a_{k_n} t} \langle M \rangle_{a_{k_n} t} K_{a_{k_n} t}^T y_t,$$

with

$$y_t = (K_{a_{k_n} t}^T)^{-1} K_{a_{k_n} t}^T x.$$ 

We have $|y_t| \geq \|K_{a_{k_n} t}^T x\|/\|K_{a_{k_n} t}\|$. So if $x \neq 0$, which we may assume without loss of generality, then $|y_t| \to \infty$ as $t \to \infty$. Lemma 1.4.3 thus implies that

$$\langle M^{k_n} \rangle_{t} \xrightarrow{P} \infty.$$
as \( t \to \infty \). It follows that \( \langle M^{k_n} \rangle_\infty = \infty \) almost surely.

The preceding paragraph shows that lemma 1.4.2 can be indeed by applied to the local martingales \( M^{k_n} \). By the assumption on the quadratic variation of \( M \) we have

\[
\langle M^{k_n} \rangle_1 = x^T K_{a_{k_n}} \langle M \rangle_{a_{k_n}} K_{a_{k_n}}^T x \xrightarrow{P} x^T \eta^T x.
\]

It thus follows from lemma 1.4.2 that

\[
(M_1^{k_n}, Y) \Rightarrow (W_{x^T \eta^T x}, Y),
\]

where \( W \) is a Brownian motion that is independent of \( (x^T \eta^T x, Y) \). Finally, use the independence of \( W \) and \( (x^T \eta^T x, Y) \) to see that \( (W_{x^T \eta^T x}, Y) \) has the same distribution as \( (x^T \eta Z, Y) \), where \( Z \) has a \( N_d(0, I) \)-distribution and \( Z \) is independent of \( (\eta, Y) \).

### 1.5 Entropy methods

Suppose that on some filtered probability space, we have a family \((M^\theta : \theta \in \Theta)\) of continuous local martingales, indexed by a countable set \( \Theta \). Moreover, suppose that we are given positive constants \((k_t)_{t>0}\). In this section we study certain properties of the random maps \( Z_t \) on \( \Theta \) defined by

\[
Z_t(\theta) = \frac{1}{\sqrt{k_t}} M^\theta_t.
\]  

Nishiyama (2000) introduced an important quantity related to the map \( Z_t \), the so-called quadratic modulus. For a given semimetric \( d \) on \( \Theta \), this is the process \( \|Z\|_d \) defined by

\[
\|Z\|_{d,t} = \sup_{d(\theta, \psi) > 0} \frac{\sqrt{\frac{1}{k_t} \langle M^\theta_t - M^\psi_t \rangle_t}}{d(\theta, \psi)}.
\]

Recall that the quadratic variation process of a continuous local martingale is unique up to indistinguishability. Hence, since \( \Theta \) is countable, the same holds for the process \( \|Z\|_d \).
Recall that a random map $Z$ on $\Theta$ is called sub-Gaussian with respect to a semimetric $\rho$ on $\Theta$ if

$$P(|Z(\theta) - Z(\psi)| > x) \leq 2e^{-\frac{1}{2} \frac{x^2}{\rho^2(\theta, \psi)}}$$

for all $\theta, \psi \in \Theta$ and $x > 0$. Using a maximal inequality for sub-Gaussian random maps (taken from Van der Vaart and Wellner (1996)) we can easily obtain a maximal inequality involving the expected modulus of $d$-continuity of the map (1.12). This bound is given in terms of so-called entropy integrals of the form

$$\int_0^\delta \sqrt{\log N(\varepsilon, \Theta, d)} \, d\varepsilon,$$

where the numbers $N(\varepsilon, \Theta, d)$ are the covering numbers of the semimetric space $(\Theta, d)$. So $N(\varepsilon, \Theta, d)$ is the minimal number of balls of $d$-radius $\varepsilon$ needed to cover $\Theta$.

Entropy methods for martingales were studied systematically by Nishiyama (1997, 1999, 2000). However, Nishiyama did not use the maximal inequality for sub-Gaussian random maps to prove his basic maximal inequality for continuous local martingales. The approach that we take here has the advantage that it works already if $d$ is only a semimetric on the parameter space $\Theta$. Nishiyama's argument requires that $d$ is a proper metric.

**Theorem 1.5.1.** Let $d$ be a semimetric on $\Theta$. There exists a universal constant $c > 0$ such that for all $t \geq 0$ and $\delta, L > 0$

$$E \sup_{d(\theta, \psi) \leq \delta} |Z_t(\theta) - Z_t(\psi)|1_{\|Z\|_{d,t} \leq L} \leq cL \int_0^\delta \sqrt{\log N(\varepsilon, \Theta, d)} \, d\varepsilon.$$

**Proof.** Let $L > 0$ and $t \geq 0$ be fixed and consider the random map $R$ on $\Theta$ defined by

$$R(\theta) = Z_t(\theta)1_{\|Z\|_{d,t} \leq L}.$$

Then for all $\theta, \psi \in \Theta$ and $z > 0$ we have

$$P(|R(\theta) - R(\psi)| > z)$$

$$= P(|Z_t(\theta) - Z_t(\psi)| > z, \|Z\|_{d,t} \leq L)$$

$$\leq P(|M^\theta_t - M^\psi_t| > z\sqrt{k_t}, \langle M^\theta - M^\psi \rangle_t \leq d^2(\theta, \psi)L^2k_t)$$

$$\leq 2e^{-\frac{1}{2} \frac{z^2}{L^2d^2(\theta, \psi)}}.$$
by theorem 1.2.1. In other words, the random map $R$ is sub-Gaussian with respect to the semimetric $Ld$. By corollary 2.2.8 of Van der Vaart and Wellner (1996) this implies that there exists a universal constant $c > 0$ such that for every $\delta > 0$

$$\mathbb{E} \sup_{Ld(\theta, \psi) \leq \delta} |R(\theta) - R(\psi)| \leq c \int_0^\delta \sqrt{\log N(\varepsilon, \Theta, Ld)} \, d\varepsilon.$$ 

It follows that

$$\mathbb{E} \sup_{d(\theta, \psi) \leq \delta} |Z_t(\theta) - Z_t(\psi)| 1_{\{\|Z\|_{d,t} \leq L\}}$$

$$= \mathbb{E} \sup_{d(\theta, \psi) \leq \delta} |R(\theta) - R(\psi)|$$

$$\leq c \int_0^{L\delta} \sqrt{\log N(\varepsilon, \Theta, Ld)} \, d\varepsilon$$

$$= cL \int_0^\delta \sqrt{\log N(L\varepsilon, \Theta, Ld)} \, d\varepsilon$$

$$= cL \int_0^\delta \sqrt{\log N(\varepsilon, \Theta, d)} \, d\varepsilon.$$ 

This completes the proof of the theorem. $\square$

The following result is the first application of theorem 1.5.1. It gives sufficient conditions for the uniform $d$-continuity of the random map (1.12). We follow the proof of theorem 2.4.3 of Nishiyama (2000).

**Theorem 1.5.2.** Let $d$ be a semimetric on $\Theta$. If $\|Z\|_{d,t} < \infty$ almost surely and

$$\int_0^1 \sqrt{\log N(\varepsilon, \Theta, d)} \, d\varepsilon < \infty, \quad (1.13)$$

then the random map $Z_t$ is almost surely uniformly $d$-continuous on $\Theta$.

**Proof.** By theorem 1.5.1 there exists for every $n$ a positive number $\delta_n$ such that for every $L > 0$

$$\mathbb{E} \sup_{d(\theta, \psi) \leq \delta_n} |Z_t(\theta) - Z_t(\psi)| 1_{\{\|Z\|_{d,t} \leq L\}} \leq cL4^{-n},$$
where \( c \) is the universal constant of theorem 1.5.1. For every \( n \), define the event

\[
A_n = \left\{ \sup_{d(\theta, \psi) \leq \delta_n} |Z_t(\theta) - Z_t(\psi)| > 2^{-n} \right\}.
\]

Then for every \( L > 0 \)

\[
\sum_n P(A_n \cap \{\|Z\|_{d,t} \leq L\}) \leq cL \sum_n 2^{-n} < \infty.
\]

So by the Borel-Cantelli lemma

\[
P(A_n \text{ i.o.} \cap \{\|Z\|_{d,t} \leq L\}) = 0
\]

for every \( L > 0 \). Since \( P(\|Z\|_{d,t} < \infty) = 1 \) by assumption, it follows that

\[
P(A_n \text{ i.o.}) = P(A_n \text{ i.o.} \cap \{\|Z\|_{d,t} < \infty\}) \leq \sum_{L=1}^{\infty} P(A_n \text{ i.o.} \cap \{\|Z\|_{d,t} \leq L\}) = 0.
\]

So we almost surely have

\[
\sup_{d(\theta, \psi) \leq \delta_n} |Z_t(\theta) - Z_t(\psi)| \leq 2^{-n}
\]

for all \( n \) large enough. This proves that with probability 1, the map \( Z_t \) is uniformly \( d \)-continuous on \( \Theta \). \( \square \)

Now suppose that with probability 1, the random maps \( Z_t \) are elements of the space \( \ell^\infty(\Theta) \) of bounded functions on \( \Theta \) and that we want to prove that they have a weak limit in this space. The first basic result in the general theory of weak convergence in \( \ell^\infty \)-spaces is that weak convergence in \( \ell^\infty(\Theta) \) is equivalent to finite-dimensional convergence and asymptotic tightness (see Van der Vaart and Wellner (1996), theorem 1.5.4). To show finite-dimensional convergence we can use the results of the preceding section. Theorem 1.5.3 below gives sufficient conditions for asymptotic tightness. It is well-known that asymptotic tightness in \( \ell^\infty(\Theta) \) is equivalent to tightness of the marginals and asymptotic equicontinuity (see Van der Vaart and Wellner (1996), theorem 1.5.7). Recall that if \( d \) is a semimetric on \( \Theta \), then a sequence \( R_n \) of random maps on \( \Theta \) is called asymptotically uniformly \( d \)-equicontinuous in probability if for every \( \varepsilon, \eta > 0 \) there exists a \( \delta > 0 \) such that

\[
\limsup_{n \to \infty} P \left( \sup_{d(\theta, \psi) \leq \delta} |R_n(\theta) - R_n(\psi)| > \varepsilon \right) < \eta.
\]
We have the following result.

**Theorem 1.5.3.** Let \( d \) be a semimetric on \( \Theta \). If \( \|Z\|_{d,t} = O_P(1) \) for \( t \to \infty \) and
\[
\int_0^1 \sqrt{\log N(\epsilon, \Theta, d)} \, d\epsilon < \infty,
\]
then the random maps \( Z_t \) are asymptotically uniformly \( d \)-equicontinuous in probability.

**Proof.** Observe that for every \( \epsilon, \delta, L > 0 \) we have
\[
P \left( \sup_{d(\theta, \psi) \leq \delta} |Z_t(\theta) - Z_t(\psi)| > \epsilon \right)
\leq P \left( \sup_{d(\theta, \psi) \leq \delta} |Z_t(\theta) - Z_t(\psi)|1_{\{\|Z\|_{d,t} \leq L\}} > \epsilon \right) + P(\|Z\|_{d,t} > L).
\]
So by Markov's inequality and theorem 1.5.1 there exists a universal constant \( c > 0 \) such that
\[
P \left( \sup_{d(\theta, \psi) \leq \delta} |Z_t(\theta) - Z_t(\psi)| > \epsilon \right)
\leq \frac{cL}{\epsilon} \int_0^\delta \sqrt{\log N(\nu, \Theta, d)} \, d\nu + P(\|Z\|_{d,t} > L)
\]
for all \( \epsilon, \delta, L > 0 \). To complete the proof, suppose that \( \epsilon, \eta > 0 \) are given. Then by the assumption on the quadratic modulus \( \|Z\|_d \) we can choose \( L \) so large that
\[
\limsup_{t \to \infty} P(\|Z\|_{d,t} > L) < \eta/2.
\]
By the entropy assumption we can now choose \( \delta > 0 \) so small that
\[
\frac{cL}{\epsilon} \int_0^\delta \sqrt{\log N(\nu, \Theta, d)} \, d\nu < \eta/2.
\]
With this choice for \( \delta \) we then have
\[
\limsup_{t \to \infty} P \left( \sup_{d(\theta, \psi) \leq \delta} |Z_t(\theta) - Z_t(\psi)| > \epsilon \right) < \eta,
\]
which is what we had to show. \( \square \)
Continuous martingales