Self-similarity of Brownian motion and a large deviation principle for random fields on a binary tree

Nina Gantert
ETH Mathematikdepartement, CH-8092 ETH Zürich, Switzerland

Received September 17, 1991; in revised form June 30, 1993

Summary. Using self-similarity of Brownian motion and its representation as a product measure on a binary tree, we construct a random sequence of probability measures which converges to the distribution of the Brownian bridge. We establish a large deviation principle for random fields on a binary tree. This leads to a class of probability measures with a certain self-similarity property. The same construction can be carried out for $C[0,1]$-valued processes and we can describe, for instance, a $C[0,1]$-valued Ornstein–Uhlenbeck process as a large deviation of Brownian sheet.

Mathematics Subject Classifications (1991): 60F10, 60J65

Introduction

Self-similarity of Brownian motion induces a certain ergodic behaviour of the Brownian bridge. We investigate large deviations of this ergodic behaviour. Let $C[0,1]_{0,0}$ be the space of all functions $X$ in $C[0,1]$ with $X(0) = X(1) = 0$ and $P$ the distribution of the Brownian bridge. We define mappings $T_0, T_1$ of $C[0,1]_{0,0}$ on itself which describe rescalings of the left and the right half of the function $X$: $(T_0 X)_t = \sqrt{2(X_{t/2} - tX_{1/2})}$, $(T_1 X)_t = \sqrt{2(X_{t+1/2} - (1-t)X_{1/2})}$, respectively. Due to the self-similarity of Brownian motion, $P$ is invariant and, in fact, even ergodic under $T_0$ and $T_1$.

For each function $\omega \in C[0,1]_{0,0}$ and each $\theta \in \{0,1\}^N$, we now construct a sequence of probability distributions $R_{n,\theta}(\omega)$ on $C[0,1]_{0,0}: R_{n,\theta}(\omega) = (1/n) \sum_{k=0}^{n-1} \delta_{T_\theta \ldots T_\theta \omega}$, where $\delta_{\omega}$ denotes Dirac measure on $\omega$. This means that in each step we choose, according to $\theta$, the left or the right half of the function $\omega$ and rescale it. $R_{n,\theta}(\omega)$ is the empirical distribution corresponding to this sequence of functions. Let $\theta_1, \theta_2, \ldots$ be independent coin tossings under $\lambda$. We can show that $R_{n,\theta}(\omega)$ converges to $P$ for $P$-a.e. $\omega$ and $\lambda$-a.e. all $\theta$. This ergodic behaviour of the Brownian bridge $P$ says that we can reconstruct $P$ with an “infinitesimal” piece of a single “typical” trajectory around a “typical” point of the unit interval $[0,1]$, if we identify $\lambda$ with Lebesgue measure.
We get another description of $T_0$ and $T_1$ using the Lévy–Ciesielski construction of the Brownian bridge: each function in $C[0,1]_0,0$ can be written as a superposition of the Schauder functions $e_{n,k}$, $k = 1, 2, \ldots, 2^{n-1}, n = 1, 2, \ldots$:

$$X(t) = \sum_{n=1}^{\infty} \sum_{k=1}^{2^{n-1}} Y_{n,k}(X) e_{n,k}(t), 0 \leq t \leq 1.$$ 

This defines a mapping of $C[0,1]_0,0$ into $\mathbb{R}^I$, where $I := \{(n,k) | k = 1, 2, \ldots, 2^{n-1}, n = 1, 2, \ldots \}$ has the structure of a binary tree. $T_0$ and $T_1$, interpreted as mappings of $\mathbb{R}^I$ into $\mathbb{R}^I$, correspond to shifts of the tree to the left and to the right, respectively. $P$ corresponds to the product measure on $\mathbb{R}^I$ with marginal distribution $N(0,1)$, i.e. the random variables $Y_{n,k}$, $k = 1, 2, \ldots, 2^{n-1}, n = 1, 2, \ldots$ are independent with distribution $N(0,1)$ under $P$. Similar representations of stochastic processes as tree-indexed random variables have been investigated recently in the context of wavelet transforms, see [2].

We now look at large deviations of the convergence of $R_{n,0}$ to $P$. Note that $R_{n,0}$ would not correspond, on the lattice $\mathbb{Z}^d$, to the usual empirical field, but rather to the empirical distribution of a sequence of configurations we get shifting along the path of a (transient) random walk on the lattice. The same model has been considered independently by Ben Arous and Tamura. They get, for each fixed $\theta$, a large deviation principle for the distributions of $R_{n,\theta}$, where the rate function depends on $\theta$. We are interested in “uniform” bounds; in particular, we want to look at probability measures on the tree which are invariant under all shifts, not only for a fixed $\theta$. We prove that the finite-dimensional marginals of $R_{n,\theta}$ satisfy a large deviation principle and characterize the rate function as a “mean entropy” (Theorem 4.1). Minimizing this rate function leads to the class of self-similar probability measures, defined by invariance under $T_0$ and $T_1$. Such a self-similar probability distribution on $\mathbb{R}^I$ can be identified, under certain conditions, with a probability distribution on $C[0,1]_0,0$. We investigate some properties of the corresponding “self-similar” stochastic processes.

More generally, we may replace $\mathbb{R}$ with a Polish space $S$ and $P$ with a product measure on $S^I$. If we set $S = C[0,1]$ (see Sect. 5), the product measure on $S^I$ with Wiener measure as one-dimensional marginal can be identified with the distribution of Brownian sheet. We can describe then, for instance, a $C[0,1]$-valued Ornstein–Uhlenbeck process as a large deviation of Brownian sheet.

1 Lévy representation of functions in $C[0,1]_0$ as elements of $\mathbb{R}^I_0$

We consider the following representation of functions in $C[0,1]_0$ which is the space of all functions $X$ in $C[0,1]$ with $X(0) = 0$. Let the Haar functions $\varphi_0$, $\varphi_{n,k}$, $k = 1, 2, \ldots, 2^{n-1}$, $n = 1, 2, \ldots$ be defined as

$$\varphi_0(t) = 1 \quad 0 \leq t \leq 1$$

and

$$\varphi_{n,k}(t) = \begin{cases} 2^{(n-1)/2} \quad & (k - 1)/2^{n-1} \leq t < (k - 1/2)/2^{n-1} \\ -2^{(n-1)/2} \quad & (k - 1/2)/2^{n-1} \leq t < k/2^{n-1} \\ 0 \quad & \text{else.} \end{cases}$$

Then $\{\varphi_0, \varphi_{n,k}, k = 1, 2, \ldots, 2^n - 1, n = 1, 2, \ldots \}$ forms a complete, orthonormal system in $L^2[0,1]$. It is the oldest example of an orthonormal wavelet basis with
Self-similarity of Brownian motion

"mother wavelet" \( \varphi_{1,1} \) (see, for instance, [4]). We won’t represent \( X \) as a superposition of \( \varphi_0, \varphi_{n,k}, k = 1, 2, \ldots, 2^n-1, n = 1, 2, \ldots \) (wavelet transform), but as a superposition of the related \textit{Schauder functions} \( e_0, e_{n,k}, k = 1, 2, \ldots, 2^n-1, n = 1, 2, \ldots \), defined as follows

\[
e_0(t) = t, \quad e_{n,k}(t) = \int_0^t \varphi_{n,k}(s) \, ds, \quad 0 \leq t \leq 1, \quad k = 1, 2, \ldots, 2^n-1, n = 1, 2, \ldots
\]

For \( X \in \mathcal{C}[0,1] \), we set

\[
h_{n,k}(X) := X \left( (k - 1/2)/2^n - \varepsilon \right) - X \left( (k - 1)/2^n \right) + X \left( (k - 1)/2^n \right)
\]

and

\[
Y_0(X) = X(1), \quad Y_{n,k}(X) := 2^{(n+1)/2} \cdot h_{n,k}(X), \quad k = 1, 2, \ldots, 2^n-1, n = 1, 2, \ldots
\]

Let \( X^N(t) = Y_0(X) \cdot t + \sum_{n=1}^{\infty} \sum_{k=1}^{2^n-1} Y_{n,k}(X)e_{n,k}(t) \).

We then have

**Lemma 1.1** (i) \( X^N \) is the linear interpolation of \( X \) on the \( N \)-th dyadic partition of \([0,1]\). This implies

\[
\lim_{N \to \infty} \sup_{t \in [0,1]} |X^N(t) - X(t)| = 0.
\]

(ii) Let \( \langle X \rangle_1^N := \sum_{k=1}^{2^n} (X(k/2^n) - X((k - 1)/2^n))^2 \) be the quadratic variation of \( X \) on the \( N \)-th dyadic partition. We then have

\[
\langle X \rangle_1^N = \frac{1}{2^n} \left( Y_0^2 + \sum_{n=1}^{\infty} \sum_{k=1}^{2^n-1} (Y_{n,k})^2 \right).
\]

(iii) \( X \) is absolutely continuous with \( X' \in L^2([0,1]) \), i.e. \( X \) is in the Cameron–Martin space \( H \), if and only if \( Y_0^2 + \sum_{n=1}^{\infty} \sum_{k=1}^{2^n-1} (Y_{n,k})^2 < \infty \). In this case, we have

\[
\|X\|_H = \|X'\|_{L^2([0,1])} = Y_0^2 + \sum_{n=1}^{\infty} \sum_{k=1}^{2^n-1} (Y_{n,k})^2.
\]

**Sketch of a proof.** (i) follows by induction on \( N \);
(ii) follows by induction on \( N \);
(iii) follows by the representation of \( X' \) as a Fourier series with respect to the complete orthonormal system of Haar functions.

Probability did not enter until now. Let now \( P \) be Wiener measure on \( \mathcal{C}[0,1] \).

**Theorem 1.2** \( Y_0, Y_{n,k}, k = 1, 2, \ldots, 2^n-1, n = 1, 2, \ldots \) are iid random variables with distribution \( N(0,1) \) under \( P \).

**Sketch of a proof.** It is enough to show that \( Y_0, Y_{n,k}, k = 1, 2, \ldots, 2^n-1, n = 1, 2, \ldots \) are pairwise uncorrelated with distribution \( N(0,1) \) under \( P \).

**Remark 1.3** Theorem 1.2 leads to an algorithm for the simulation of the Brownian path. The law of large numbers and Lemma 1.1(ii) imply the following:
Corollary 1.4 $\langle X \rangle^N_t$ converges P-a.s. and in $L^2(P)$ to 1 if $N \to \infty$.

Remark 1.5 This construction of Brownian motion goes back to Lévy and Ciesielski (see [13] for references), it can be carried out for every complete orthonormal system in $L^2[0,1]$. Let $\{\varphi_k|k \geq 1\}$ be a complete orthonormal system in $L^2[0,1]$.

$$e_k(t) := \int_0^t \varphi_k(s)ds, \quad 0 \leq t \leq 1, \quad k = 1, 2, \ldots$$

Let $Y_k$, $k = 1, 2, \ldots$ be iid with distribution $N(0, 1)$ and set

$$X_t := \sum_{k=1}^{\infty} Y_k e_k(t) \quad 0 \leq t \leq 1.$$

Then $(X_t)_{0 \leq t \leq 1}$ is a Brownian motion (see Itô and Nisio [11] for a proof).

Let $I_0 := \{0, (n, k), k = 1, 2, \ldots, 2^{n-1}, n = 1, 2, \ldots\}$ be a binary tree:

![Fig. 1](image-url)

To each function in $C[0,1]_0$ corresponds a set of coefficients $Y_{0,k}$, $k = 1, 2, \ldots, 2^{n-1}$, $n = 1, 2, \ldots$ according to the mapping from $C[0,1]_0$ to $\mathbb{R}^{I_0}$ defined in (1.1). The converse, however, is not true: not each element of $\mathbb{R}^{I_0}$ is a function in $C[0,1]_0$, hence not each probability distribution on $\mathbb{R}^{I_0}$ is a probability distribution on $C[0,1]_0$.

Lemma 1.6 Let $Q$ be a probability distribution on $\mathbb{R}^{I_0}$, $M_n := \max_{k=1,2,\ldots,2^{n-1}} |Y_{n,k}|$ and $X^N$ as in Lemma 1.1. If

$$\sum_{n=1}^{\infty} Q[M_n > 2^n] < \infty \text{ for an } \alpha < 1/2 \quad (1.2)$$

then $Q[X^m \text{ converges uniformly}] = 1$, i.e. $Q$ is a probability distribution on $C[0,1]_0$. 
Proof. We equip $C[0, 1]_0$ with the supremum norm $\|X\| = \sup_{t \in [0, 1]} |X(t)|$. It is enough to show that
\[
\sum_{n=1}^{\infty} Q[\|X^n - X^{n-1}\| > a_n] < \infty
\]  
for a sequence $(a_n)_{n=1,2,\ldots} \subseteq \mathbb{R}$ with $\sum_{n=1}^{\infty} a_n < \infty$.

The Borel–Cantelli lemma then implies that $(X^m)_{m=1,2,\ldots}$ forms $Q$-a.s. a Cauchy sequence in $C[0,1]_0$. To show (1.3), we note that
\[
\|X^n - X^{n-1}\| \leq M_n \cdot 2^{-(n+1)/2},
\]  
which implies
\[
Q[\|X^n - X^{n-1}\| > a_n] \leq Q[M_n > 2^{(n+1)/2} a_n] .
\]  
If we set
\[ a_n = 2^{-\beta n - (1/2)} \text{ with } \beta = (1/2) - \alpha > 0 \]
and apply (1.2), the claim follows. $\square$

2 Construction of a random sequence of probability distributions which converges to the distribution of the Brownian bridge

Let $P$ be the distribution of the Brownian bridge. Then $Y_0 = 0$ $P$-a.s. and $Y_{n,k}, k = 1, 2, \ldots, 2^{n-1}, n = 1, 2, \ldots$ are iid random variables under $P$ with distribution $N(0,1)$. We set
\[ I := \{(n,k)|k = 1, 2, \ldots, 2^{n-1}, n = 1, 2, \ldots\} \]
and
\[ \Omega = C[0,1]_0, o := \{X \in C[0,1]|X(1) = X(0) = 0\}, \ X_t(\omega) := \omega(t) . \]
We then have $\Omega = C[0,1]_0, 0 \subseteq \mathbb{R}$. We denote the set of all probability distributions on $\Omega$ by $\mathcal{M}_1(\Omega)$.

We consider the mapping $T_0: \Omega \rightarrow \Omega$, defined as
\[ (T_0X)_t := \sqrt{2(X_{t/2} - tX_{1/2})} \quad 0 \leq t \leq 1 . \]

$T_0$ corresponds to a shift to the left of the tree, i.e. $Y_{n,k}(T_0\omega) = Y_{n+1,k}(\omega)$. In the same way, we define $T_1: \Omega \rightarrow \Omega$ as
\[ (T_1X)_t := \sqrt{2(X_{(t+1)/2} - (1-t)X_{1/2})} \quad 0 \leq t \leq 1 . \]

$T_1$ corresponds to a shift to the right of the tree, i.e. $Y_{n,k}(T_1\omega) = Y_{n+1,2^{n-1}+k}(\omega)$. $P$ is invariant under $T_0$ and under $T_1$.

We now consider the bigger space $\widetilde{\Omega}$, defined as
\[ \widetilde{\Omega} = \Omega \times \{0,1\}^N \]
\[ \widetilde{\omega} = (\omega, \theta) \quad \theta \in \{0,1\}^N \]
\[ \widetilde{P} := P \times \lambda , \]
where \( \lambda \) denotes product measure on \( \{0,1\}^\mathbb{N} \) with \( \lambda[\theta_i = 0] = \lambda[\theta_i = 1] = 1/2 \). Let the shift \( \bar{T} \) on \( \bar{\Omega} \) be defined as
\[
\bar{T} : \bar{\Omega} \to \bar{\Omega}
\]
\[
(\omega, (\theta_1, \theta_2, \ldots)) \to (T_{\theta_1} \omega, (\theta_2, \theta_3, \ldots)).
\]
\( \bar{P} \) is invariant under \( \bar{T} \). We can even show:

**Theorem 2.1** \( \bar{P} \) is ergodic with respect to \( \bar{T} \).

**Proof.** Let \( \mathcal{F}^* := \bigcap_n \sigma(\{T^m, m > n\}) \) be the tail-field on \( \bar{\Omega} \). \( \bar{P} \) is a product measure on \( \bar{\Omega} \), hence Kolmogorov's 0-1-law is satisfied:
\[
\bar{P}[A] = 0 \text{ or } \bar{P}[A] = 1 \text{ if } A \in \mathcal{F}^*.
\]
The \( \sigma \)-field \( \mathcal{S} := \{A | T^{-1}A = A\} \), generated by the shift-invariant sets, is contained in \( \mathcal{F}^* \), hence we have \( \bar{P}[A] = 0 \) or \( \bar{P}[A] = 1 \), if \( A \in \mathcal{S} \), i.e. \( \bar{P} \) is ergodic under \( \bar{T} \). \( \square \)

Let \( \delta_{T_k \omega} \) denote Dirac measure on \( T^k \omega \) and define the probability distribution \( \bar{R}_n(\bar{\omega}) \) by \( \bar{R}_n(\bar{\omega}) := (1/n) \sum_{k=1}^n \delta_{T_k \omega} \) (where \( T^0 \) denotes the identity). Let \( R_n,\theta(\omega) \) denote the marginal distribution of \( \bar{R}_n(\bar{\omega}) \) on \( \Omega \) for fixed \( \theta \):
\[
R_n,\theta(\omega) = \frac{1}{n} \sum_{k=1}^n \delta_{(T_\theta)^k \omega}
\]
where
\[
(T_\theta)^k \omega = T_{\theta_1} \circ \cdots \circ T_{\theta_k} \omega, \quad k \geq 1,
\]
\[
(T_\theta)^0 \omega = \omega.
\]
For each \( \theta \in \{0,1\}^\mathbb{N} \), \( R_n,\theta \) is a random variable with values in \( M_1(\Omega) \).

**Theorem 2.2** For \( \lambda \)-a.e. \( \theta \) and \( P \)-a.e. \( \omega \), \( R_n,\theta(\omega) \) converges weakly to \( P \).

**Proof.** Let \( f : \bar{\Omega} \to \mathbb{R} \) be measurable and bounded. With Birkhoff's ergodic theorem we get from Theorem 2.1:
\[
\frac{1}{n} \sum_{k=0}^{n-1} f \circ T_k \xrightarrow{n \to \infty} \int f \, d\bar{P} \quad \bar{P} \text{-a.s.},
\]
i.e.
\[
\int f \, dR_n,\theta(\omega) \to \int f \, d\bar{P} \quad \text{for } \lambda \text{-a.e. } \theta \text{ and } P \text{-a.e. } \omega.
\]
Since the set of all bounded, continuous functions on \( \bar{\Omega} \) is countably generated, this implies
\[
R_n,\theta \xrightarrow{w} P \quad \text{for } \lambda \text{-a.e. } \theta \text{ and } P \text{-a.e. } \omega. \quad \square
\]

**Remark 2.3** In fact, \( (R_n,\theta) \) converges weakly to \( P \) for each \( \theta \in \{0,1\}^\mathbb{N} \) (this was shown by Ben Arous and Tamura in an unpublished paper).

**Remark 2.4** Theorem 2.2 says that we can reconstruct Wiener measure with an "infinitesimal piece" of a single "typical" path around a "typical" point of the unit
interval. This property characterizes fractals: the information about a fractal object is contained in an arbitrary small part of the object. Of course, we are dealing with random fractals: the invariance of \( \tilde{P} \) under \( \tilde{T} \) corresponds to a "self-similarity in distribution":
\[
(\sqrt{2}(X_{t/2} - tX_{1/2}))_{0 \leq t \leq 1} \quad \text{and} \quad (\sqrt{2}(X_{(t+1)/2} - (1 - t)X_{1/2}))_{0 \leq t \leq 1}
\]
have the same distribution under \( P \) as \( (X_t)_{0 \leq t \leq 1} \). We get a deterministic fractal, if we set, for instance, all the coefficients \( Y_{n,k}, k = 1, 2, \ldots, 2^{n-1}, n = 1, 2, \ldots \) to the value 1, see Sect. 5, Example 5.3.

Remark 2.5 Let \( Q \in \mathcal{M}_1(\Omega) \) and let \( \tilde{Q} := Q \times \lambda \). Then \( Q \ll P \) implies \( \tilde{Q} \ll \tilde{P} \), hence \( R_{n,\theta}(\omega) \rightarrow P \tilde{Q} \)-a.s., i.e. for \( \lambda \)-a.e. \( \theta \), \( Q \)-a.e. \( \omega \). This means \( Q \) cannot be reconstructed in this way, but we get \( P \) from a path which is "typical" for \( Q \). The intuition behind this is the following: the drift of \( Q \) with respect to \( P \) is lost because of the iterated rescaling, see also Lemma 1.1(iii).

3 Generalization to probability distributions on \( S' \)

Let \( S \) be a Polish space. We may now replace \( \mathbb{R} \) with \( S \) and consider \( \Omega = S' \). Let \( \mu \) be a probability distribution on \( S \) and \( P = \prod \mu \) be the corresponding product measure on \( S' \).

We define \( T_0, T_1, \tilde{Q}, P, \tilde{T}, \tilde{R}, \) and \( R_{n,\theta} \) as in Sect. 2. We equip \( S' \) with the product topology and \( \mathcal{M}_1(S') \), the set of probability distributions on \( S' \), with the topology of weak convergence. We then get, in the same way as Theorem 2.2:

Theorem 3.1 For \( \lambda \)-a.e. \( \theta \) and \( P \)-a.e. \( \omega \), \( R_{n,\theta}(\omega) \) converges to \( P \) (in \( \mathcal{M}_1(S') \)).

Definition 3.2 We call \( Q \) stationary or self-similar, if \( Q \) is invariant under \( T_0 \) and under \( T_1 \). Let \( \mathcal{M}_1^S \) denote the set of all stationary probability distributions on \( \Omega \). We call \( Q \in \mathcal{M}_1^S \) ergodic, if \( \tilde{Q} := Q \times \lambda \) is ergodic on \( \tilde{Q} \) with respect to \( \tilde{T} \).

Remark 3.3 If \( Q \) is ergodic, \( R_{n,\theta}(\omega) \) converges to \( Q \) for \( Q \)-a.e. \( \omega \) and \( \lambda \)-a.e. \( \theta \).

4 Large deviations

In the following we investigate large deviations of the convergence of \( R_{n,\theta} \) to \( P \) in Theorem 3.1. Note that \( R_{n,\theta} \) would not correspond, on the lattice \( \mathbb{Z}^d \), to the usual empirical field, but rather to the empirical distribution of a sequence of configurations we get shifting along the path of a (transient) random walk on the lattice (see also Remark 4.8). Ben Arous and Tamura got, for each fixed \( \theta \), a large deviation principle for the distributions of \( R_{n,\theta} \), where the rate function depends heavily on \( \theta \) (see [3]). We are interested in estimates which do not depend on \( \theta \). In particular, our rate function will be finite only on probability distributions which are stationary, i.e. invariant under all shifts \( T_\theta \). There is, however, no large deviation principle holding "uniformly" in \( \theta \); we can get the bounds in Theorem 4.1 only for \( \lambda \)-a.e. \( \theta \). To understand why, look at the following example:

Let \( I = \mathbb{R} \) and \( A := \{ Q \mid Q[ Y_{2,2} > Y_{1,1} ] = 1 \} \). If we shift only to the left, i.e. take \( \theta = (0,0,0,\ldots) \), we have
\[
P[R_{n,\theta} \in A] = P[ Y_{2,2} > Y_{1,1} ]^n,
\]
hence \( (1/n) \log P[R_{n,\theta} \in A] \rightarrow -\infty \).

If we shift only to the right, i.e. take \( \theta = (1,1,1,\ldots) \), we get
\[
P[R_{n,\theta} \in A] = P[ Y_{1,1} < Y_{2,2} < Y_{3,4} < \ldots < Y_{n,2^{n-1}} ],
\]
hence \( (1/n) \log P[R_{n,\theta} \in A] \rightarrow -\infty \).
Also, there is no “global” rate function concentrated on \( \mathcal{M}_1^* (\Omega) \). This will follow from Theorem 4.1. We give an illustrative example:

Let \( I = \mathbb{R} \) and \( P = \prod_{n \in \mathbb{N}} N(0, 1) \) as in Sect. 2. Consider \( A_m : \{ Q | E_Q [Y_{m, 1}] = E_Q [Y_{m, 2}] = \ldots = E_Q [Y_{m, 2^{m-1}}] \geq b \} \) \( (m = 1, 2, \ldots) \) where \( b > 0 \). Since \( P \) is a product measure, we have, for all \( \theta \),

\[
P[R_{n, \theta} \in A_m] = P[\bar{Y}_n \geq b]^{2^{m-1}},
\]

where \( \bar{Y}_n \) is the arithmetic mean of \( Y_{1,1}, Y_{2,1}, Y_{3,1}, \ldots, Y_{n,1} \). So we get

\[
\lim_n \frac{1}{n} \log P[R_{n, \theta} \in A_m] = - 2^{m-1} \lambda(b)
\]

for all \( \theta \), where \( \lambda(x) = x^2/2 \) is the rate function for the large deviations of the arithmetic mean of iid random variables with distribution \( N(0, 1) \). On the other hand, \( A_m \cap \mathcal{M}_1^* (\Omega) = A_1 \cap \mathcal{M}_1^* (\Omega) \) for each \( m \).

We establish a principle of large deviations for the finite-dimensional marginal distributions of \( R_{n, \theta} \) on \( S^I (J \subseteq I, J \text{ finite}) \) under \( P \). Recall \( P \) is, in this section, a product measure on \( S^I \). Let us begin with some notation: we denote the mapping of the index set \( I \) on itself, which corresponds to \( T_0 \), again by \( T_0 \), i.e. \( T_0(n, k) = (n + 1, k) \) and, in the same way, \( T_1(n, k) = (n + 1, 2^n + k) \) \( (k = 1, 2, \ldots, 2^n - 1, n = 1, 2, \ldots) \). Let \( J \subseteq I \) and let \( T_{\theta_0} \) denote the identity. Let \( F_k(\theta, J) := \bigcup_{j=0}^{k-1} T_{\theta_j} \circ \cdots \circ T_{\theta_0} J \) be the set of coordinates, generated by \( J \) after \( k - 1 \) shifts according to \( \theta \), \( (k \geq 1) \), \( F_0(J) := J \). Let \( \mathcal{F}_k(\theta, J) := \sigma(\{ Y_i \in F_k(\theta, J) \}) \) be the corresponding \( \sigma \)-field. We write \( (T_0)^k \) instead of \( T_{\theta_0} \circ \cdots \circ T_{\theta_0} \), hence \( F_k(\theta, J) = \bigcup_{j=0}^{k-1} (T_0)^j J \), \( (k \geq 1) \). Let \( \mathcal{M}_1(S^I) \) denote the set of all probability distributions on \( S^I \) as before. We now consider subsets \( A_J \) of \( \mathcal{M}_1(S^I) \), which are characterized in the following way: let \( J \subseteq I, J \text{ finite}, B_J \subseteq \mathcal{M}_1(S^I) \) measurable and

\[
A_J := \{ Q | Q|_{\mathcal{F}_0(J)} \in B_J \}
\]

i.e. if \( Q \) is in \( A_J \) or not depends only on the finite-dimensional marginal of \( Q \) on \( \mathcal{F}_0(J) \). Then we have: \( A_J \) is open in \( \mathcal{M}_1(\Omega) \) \( \Leftrightarrow \) \( B_J \) is open in \( \mathcal{M}_1(S^I) \).

Let the relative entropy \( H(Q|P)|\mathcal{F}_0 \) of \( Q \) with respect to \( P \) on the \( \sigma \)-field \( \mathcal{F}_0 \) be defined as \( E_Q \left[ \log \frac{dQ}{dP} \right]_{\mathcal{F}_0} \), if \( Q \ll P \) on \( \mathcal{F}_0 \), and \( = + \infty \) else. Now we can state the following large deviation principle:

**Theorem 4.1** Let \( J \subseteq I, J \text{ finite}. Then there is a function \( I_J : \mathcal{M}_1(\Omega) \rightarrow [0, \infty] \), such that the following holds for all \( A_J \) of the form in (4.1):

\[
A_J \text{ open } \Rightarrow \lim_{n \to \infty} \frac{1}{n} \log P[R_{n, \theta} \in A_J] \geq - \inf_{Q \in A_J} I_J(Q)
\]

\[
A_J \text{ closed } \Rightarrow \lim_{n \to \infty} \frac{1}{n} \log P[R_{n, \theta} \in A_J] \leq - \inf_{Q \in A_J} I_J(Q)
\]

for \( \lambda \)-a.e. \( \theta \). Further, \( I_J : \mathcal{M}_1(\Omega) \rightarrow [0, \infty] \) is lower semicontinuous, \( I_J(Q) = + \infty \) if \( Q \notin \mathcal{M}_1^* \) and we have for \( Q \in \mathcal{M}_1^* \):

\[
I_J(Q) = \lim_{n \to \infty} \frac{1}{n} \int H(Q|P)|_{\mathcal{F}_0(\theta, J)} \lambda(d\theta).
\]
Proof. We fix $J$ and write $\mathcal{F}_0$, $\mathcal{F}_n(\theta)$, $A$ instead of $\mathcal{F}_0(\theta, J)$, $\mathcal{F}_n(\theta, J)$, $A_J$. The basic idea of the proof is to represent $R_{n, \theta} \mid \mathcal{F}_0$ as the empirical distribution of a Markov chain of order $m$. We will give an explicit proof of the lower bound and refer to a general theorem in the proof of the upper bound.

First we investigate the behaviour of the Radon–Nikodym derivatives of $Q$ with respect to $P$ on the $\sigma$-fields $\mathcal{F}_n(\theta)$: in fact, we need to know only the growths of the relative entropies $H(\mathcal{Q} \mid P) \mid \mathcal{F}_n(\theta)$ of $Q$ with respect to $P$ on the $\sigma$-fields $\mathcal{F}_n(\theta)$.

**Theorem 4.2** For each $Q \in \mathcal{M}_\lambda$, $(1/n) H(\mathcal{Q} \mid P) \mid \mathcal{F}_n(\theta)$ converges to $I_J(Q)$ for $\lambda$-a.e. $\theta$, where

$$ I_J(Q) = \lim_{n \to \infty} \frac{1}{n} \int H(\mathcal{Q} \mid P) \mid \mathcal{F}_n(\theta) \lambda(d\theta) \in [0, \infty] $$

Further, $I_J(\cdot)$ is lower semicontinuous and affine.

**Remark 4.3** If $Q$ is ergodic, we have for $\lambda$-a.e. $\theta$: $(1/n) \log \left| \frac{dQ}{dP} \right| \mid \mathcal{F}_n(\theta)$ converges $Q$-a.s. to $I_J(Q)$.

**Proof of Theorem 4.2** Let $\bar{Q} = Q \times \lambda$, $\bar{P} = P \times \lambda$. We make use of a theorem of A. Barron.

**Theorem 4.4** Let $(X_n)_{n=1,2,\ldots}$ be a stationary process with values in a Standard Borel space $E$. Let $Q$ be the distribution of $(X_n)$, and $P \in \mathcal{M}_1(E^\mathbb{N})$ be a “reference measure”: $P$ is stationary and Markov of order $m$ (i.e. $P[A \mid X_{n-1}, \ldots, X_1] = P[A \mid X_{n-1}, \ldots, X_{n-m}]$ for all $\sigma(X_{n},X_{n+1},\ldots)$ measurable sets $A$ and $n > m$). Define the $\sigma$-fields $A_n := \sigma(X_1, X_2, \ldots, X_n)$. Then the specific relative entropy $h(Q \mid P)$ of $Q$ with respect to $P$ exists:

$$ h(Q \mid P) = \lim_{n \to \infty} \frac{1}{n} H(Q \mid P) \mid \mathcal{F}_n \in [0, \infty] $$

See Barron [1, Theorem 1] for the proof.

Note that $H(Q \mid P) \mid \mathcal{F}_n$ is increasing, so $h(Q \mid P) = \infty$ if there is an $n$ such that $H(Q \mid P) \mid \mathcal{F}_n = \infty$.

We apply Theorem 4.4 with $\bar{Q}, \bar{P}, E := \{(Y_1, \ldots, Y_\ell) \mid Y_i \in \mathcal{S}\} \times \{0,1\}$, where $(1,2,\ldots,\ell)$ is an enumeration of $J$.

$\bar{P} \mid \sigma(J_1, J_2, J_3, \ldots) \times \mathcal{S}$ can be identified with a stationary measure, Markov of order $m$ on $E^\mathbb{N}$, where $X_n$ consists of the $\ell$-tuple of random variables $Y_i, i \in T^{n-1}_\theta J$ and $\theta_n$.

$$ m := \max \{n \mid \exists k \text{ with } (n,k) \in J \} - 1. \quad (4.2) $$

(Since the sets $T^{n}_\theta J, n = 1,2,\ldots$ are not disjoint in general, $\bar{P} \mid \sigma(J_1, J_2, J_3, \ldots) \times \mathcal{S}$ is in general not a product measure on $E^\mathbb{N}$). In the same way, we can identify a stationary and ergodic $\bar{Q}$ with a stationary and ergodic probability distribution on $E^\mathbb{N}$, respectively. Hence we get from Theorem 4.4

$$ \exists \bar{I}_J(\bar{Q}) = \lim_{n \to \infty} \frac{1}{n} H(\bar{Q} \mid \bar{P}) \mid \sigma(x_1, \ldots, x_n). $$
If we set $I_J(Q) := \overline{I_J(Q)}$, we get

$$I_J(Q) = \lim_{n \to \infty} \frac{1}{n} \int H(Q|P)|_{\mathscr{F}^n(\theta)} \lambda(d\theta).$$

(4.3)

It remains to show that $I_J$ is lower semicontinuous and affine; for this, we refer to [9].

Remark 4.5 $Q \in \mathcal{M}_1^*$ and $I_J(Q) = 0$ imply $Q = P$.

The next step in the proof of the lower bound is to show that the set of ergodic probability distributions is dense in $\mathcal{M}_1^*$.

Lemma 4.6 Let $Q \in \mathcal{M}_1^*$. Then there is a sequence of ergodic probability distributions $(Q_n)_{n=1,2,\ldots}$ such that $Q_n \Rightarrow Q$, and $I_J(Q_n) \to I_J(Q)$ for $n \to \infty$.

The proof is similar to the proof of Lemma 4.8 in Föllmer [7].

Note that Theorem 4.2, Lemma 4.6 and Remark 3.3 can be used to prove the lower bound with a standard argument (see [5, p. 76]).

For each $\theta$, $R_{n,\theta}|_{\mathscr{F}_\theta}$ is the $n$-th empirical distribution of a Markov chain of order $m$. This chain is, in the terminology of Deuschel et al. [6] R-mixing with $M = 1$ if $R \geq m$ (see (4.2)), since $\sigma(\{ T_{\theta \omega}, 0 \leq k \leq r \})$ and $\sigma(\{ T_{\theta \omega}^k, k \geq r + m \})$ are independent. A general theorem about uniform large deviations (see [6, p. 91]) implies, together with the contraction principle, that for each $\theta$, the distributions of $R_{n,\theta}|_{\mathscr{F}_\theta}$, $n \geq 1$, satisfy a large deviation principle with rate function $h_{\theta, J}: \mathcal{M}_1^*(\mathscr{F}_\theta) \to [0, \infty]$ and $h_{\theta, J}(v) = \inf\{ I_{\theta, J}(Q) \mid Q \in \mathcal{M}_1^*(\theta), Q|_{\mathscr{F}_\theta} = v \}$ with $I_{\theta, J}(Q) = \lim_n (1/n) H(Q|P)|_{\mathscr{F}_\theta^o}(\theta)$. (Here, $\mathcal{M}_1^*(\theta)$ denotes the set of probability distributions which are invariant under $T_\theta$.) It remains to identify this rate function with the rate function in Theorem 4.1. Theorem 4.2 implies that for $\lambda$-a.e. $\theta$, $\inf\{ I_{\theta, J}(Q) \mid Q \in \mathcal{M}_1^*(\theta), Q|_{\mathscr{F}_\theta} = v \} = \inf\{ I_{\theta, J}(Q) \mid Q \in \mathcal{M}_1^*(\theta), Q|_{\mathscr{F}_\theta} = v \}$ and $
lim_n (1/n) H(Q|P)|_{\mathscr{F}_\theta^o}(\theta) = I_{\theta, J}(Q)$, so $I_{\theta, J}(Q) = I_J(Q)$ for $\lambda$-a.e. $\theta$.

Note that for each $L \geq 0$, $\{ Q|_{\mathscr{F}_\theta} \mid I_J(Q) \leq L \}$ is a compact subset of $\mathcal{M}_1^*(S^1, \mathcal{F}_\theta)$, but $\{ Q \mid I_J(Q) \leq L \}$ is in general not compact in $\mathcal{M}_1(\Omega)$.

Of course, the arguments in [6] are much more general than our situation requires. For a direct proof of Theorem 4.1, we refer to [9].

Theorem 4.4 says that we have to minimize the rate functional $I_J$ over the set of probability distributions $\mathcal{M}_1^*(S^1)$. It is therefore natural to ask about the properties of probability distributions in $\mathcal{M}_1^*(S^1)$. We omitted $\omega$, but we can extend any stationary measure on $S^1$ to a stationary measure on $S^1_\omega$. Let $\mu = \prod_{t \in I_{\theta \omega}} \mu \in \mathcal{M}_1^*(S^1_\omega)$ denote a product measure and consider $\mathcal{M}_1^* = \mathcal{M}_1^*(S^1_\omega)$. Probability distributions in $\mathcal{M}_1^*$ are typically singular with respect to $P$. More precisely, $Q \in \mathcal{M}_1^*$, $Q < P \Rightarrow Q = P$. In particular, $Q \in \mathcal{M}_1^*$ has infinite relative entropy with respect to $P$, if $Q \neq P$. We can show, though, that a specific relative entropy with respect to $P$ exists for each $Q \in \mathcal{M}_1^*$. Consider the $\sigma$-fields $\mathcal{F}_2^* = \sigma(\{ Y_0, Y_{m,k} \mid m \leq n \})$. 


Lemma 4.7 Every $Q \in \mathcal{M}_1^*$ has a specific relative entropy $h(Q|P)$ with respect to $P$:

$$h(Q|P) = \lim_{n \to \infty} \frac{1}{2^n} H(Q|P)_{\mathcal{F}_{2^n}}$$

$$= \sup_{n} \frac{1}{2^n} H(Q|P)_{\mathcal{F}_{2^n}} \in [0, \infty].$$

In particular, $Q \in \mathcal{M}_1^*$, $h(Q|P) = 0 \Rightarrow Q = P$.

Further, $h(\cdot|P)$ is affine on $\mathcal{M}_1^* (\Omega)$.

Sketch of a proof. Let $\mathcal{V}$ be the set containing all subsets of $I_0$. Then the function $f: \mathcal{V} \to \mathbb{R}$, $f(V) = H(Q|P)_{\sigma_{(y_{i\in V})}}$ is a superadditive set function, i.e. $f(V \cup W) \geq f(V) + f(W)$ for disjoint sets $V, W \in \mathcal{V}$. Here we made use of the product structure of $P$. The rest of the proof is left to the reader (see also Georgii, [10, Chap. 15, Sect. 2], for a general argument).

Remark 4.8 Look at

$$R_\alpha (\omega) := \sum_{k=1}^{\infty} \delta(T_\alpha \omega).$$

$R_\alpha (\omega)$ is the analogon to the usual empirical field on the tree. Then the same arguments as in the proof of Theorem 2.2 show that $R_\alpha (\omega) \to P$ for $P$-a.e. $\omega$. The distributions of $R_\alpha, n \geq 1$, satisfy a large deviation principle with good rate function $I$ where $I(Q) = h(Q|P)$ for $Q \in \mathcal{M}_1^*, I(Q) = + \infty$ else. We don’t give a proof here, since it consists merely in carrying over arguments in [6] or [8] from the lattice to the tree structure. (The arguments in [6] or [8] can here be simplified of course, since we treat the particularly nice case of a product measure $P$.)

5 Examples

5.1 Stationary probability distributions on $\mathbb{R}^{I_0}$

Let $P = \prod_{t \in I_0} N(0, 1) \in \mathcal{M}_1^*(\mathbb{R}^{I_0})$ denote Wiener measure. Can we identify a self-similar probability distribution on $\mathbb{R}^{I_0}$ with a probability distribution on $C[0, 1]_0$? Notice this is not clear a priori. In Lemma 5.1 below we give a condition on $Q$ which guarantees that the support of $Q$ is contained in $C[0, 1]_0$. Our conjecture is, however, that this holds true for every $Q \in \mathcal{M}_1^*(\mathbb{R}^{I_0})$. Because we deal with stationary probability distributions, condition (5.1) below involves only the one-dimensional marginal distribution of $Q$.

Lemma 5.1 Let $Q \in \mathcal{M}_1^* = \mathcal{M}_1^*(\mathbb{R}^{I_0})$ and

$$\sum_{n=1}^{\infty} 2^{n-1} Q[|Y_0| \geq 2^n] < \infty \quad \text{for an } \alpha < 1/2.$$  \tag{5.1}

Then $Q$ is a probability distribution on $C[0, 1]_0$.

Proof. We can replace (1.2) in Lemma 1.6 with (5.1):

$$Q[M_n \geq 2^n] \leq Q \left[ \max_{k=1,2,\ldots,2^{n-2}} |Y_{n,k}| \geq 2^n \right] + Q \left[ \max_{k=2^n+1,\ldots,2^{n-1}} |Y_{n,k}| \geq 2^n \right]$$

$$= 2Q[M_{n-1} \geq 2^n],$$
since \(Q\) is stationary. By iteration, we conclude
\[
Q[M_n \geq 2^n] \leq 2^{n-1} Q[|Y_0| \geq 2^n].
\]

Let \(v\) be the distribution of \(Y_0\) under \(Q\). Sufficient for (5.1) to hold is, for instance,
\[
\int |x|^{2+\varepsilon} \, dv < \infty \quad \text{for an } \varepsilon > 0 \quad \text{or} \quad \sup_{t \in \mathbb{R}} \left| \frac{dv}{d\mu}(t) \right| \leq C, \quad \text{where } \mu = N(0,1).
\]

If the support of \(Q\) is contained in \(C[0,1]_o\), we get the coordinate process \((X_t)_{0 \leq t \leq 1}\) via
\[
X_t(\omega) = Y_0(\omega) \cdot t + \sum_{n=1}^{\infty} \sum_{k=1}^{2^n-1} Y_{n,k}(\omega) e_{n,k}(t) \quad (0 \leq t \leq 1)
\]
as in Sect. 1. In this case, we can write the \(\sigma\)-field \(\mathcal{F}_{2^n}\) in Lemma 4.7 as
\[
\mathcal{F}_{2^n} = \sigma(\{X_{k-2^n} \mid k = 0,1,\ldots,2^n\}),
\]
and we have an interpretation of the specific relative entropy \(h(Q \mid P)\) as a limit of entropies on the dyadic partitions of the unit interval.

The simplest case of a stationary probability distribution is, of course, a product measure \(Q = \prod_{i \in \mathbb{Z}_0} v\). What can we then say about \((X_t)_{0 \leq t \leq 1}\)?

**Lemma 5.2** Let \(Q = \prod_{i \in \mathbb{Z}_0} v\) and assume \(Q\) satisfies (5.1).

(i) Let \(\int x^2 \, dv < \infty\). Then \(E_Q[X_tX_s] = \int x^2 \, dv \cdot (t \wedge s)\). Further, \((X_t)_{0 \leq t \leq 1}\) has quadratic variation \(\langle X_t \rangle_{0 \leq t \leq 1}\) (along the sequence of dyadic partitions of \([0,1]\)) and
\[
\langle X \rangle_t = \left( \int x^2 \, dv \right) \cdot t \quad (0 \leq t \leq 1), \quad Q\text{-a.s.}
\]

(ii) Let \(\int x^4 \, dv < \infty\). Then \((X_t)_{0 \leq t \leq 1}\) is locally Hölder-continuous with exponent \(\gamma\), for each \(\gamma < 1/2\).

For the proof, we refer to [9].

In general, \((X_t)_{0 \leq t \leq 1}\) is not a Markov process under \(Q\). We can state the following "weakened Markov property": the conditional distribution of \(\{X_t \mid t \in [(k-1)/2^n, k/2^n]\}\), given \(\{X_t \mid t \in [0,(k-1)/2^n) \cup [k/2^n,1]\}\), depends only on \(X_{(k-1)/2^n}\) and \(X_{k/2^n}\), \(k = 1,2,\ldots,2^n\), \(n = 1,2,\ldots\).

**Example 5.3** Let \(v = N(a,1)\) and \(Q = \prod_{i \in \mathbb{Z}_0} v\). Of course, condition (5.1) in Lemma 5.1 is satisfied. \(Q\) is the distribution of \((B_t + a \cdot g(t))_{0 \leq t \leq 1}\), where \((B_t)_{0 \leq t \leq 1}\) is

\[g(t)\]

Fig. 2
a Brownian bridge and \( g \) the self-similar function
\[
g(t) = \sum_{n=1}^{\infty} \sum_{k=1}^{2^{n-1}} e_{n,k}(t) \quad (0 \leq t \leq 1)
\]
g is the profile of the fractal "mount Takagi" (see [14]).

5.2 Brownian sheet and other examples with \( S = C[0,1] \)

Except Brownian motion (multiplied with constants), real-valued diffusions resp.
their bridges are not self-similar in our sense, i.e. invariant under \( T_0 \) and under \( T_z \).
But if we allow the coefficients \( Y_0, Y_{n,k}, k = 1, 2, \ldots, 2^{n-1}, n = 1, 2, \ldots \) to have
values in a function space, we can describe well-known, "smooth" objects like the
\( C[0,1] \)-valued Ornstein–Uhlenbeck process. Take \( S = C[0,1] \) and \( P^\infty = \prod_{t \in S} P \in \mathcal{M}_5(S^0) \), where \( P \) denotes Wiener measure on \( C[0,1] \).
Set
\[
X(t,\tau) = Y_0(\tau) \cdot t + \sum_{n=1}^{\infty} \sum_{k=1}^{2^{n-1}} Y_{n,k}(\tau) e_{n,k}(t) \quad (0 \leq t, \tau \leq 1).
\]

\((X(t,\cdot))_{0 \leq t \leq 1}\) is then a \( C[0,1] \)-valued Brownian motion under \( P^\infty \) for \( t_1, t_2, t_3, t_4 \)
with \( 0 \leq t_1 < t_2 \leq t_3 \leq t_4 \leq 1 \) the increments \( X(t_2,\cdot) - X(t_1,\cdot) \), \( X(t_4,\cdot) - X(t_3,\cdot) \) are independent and
\( 1/(\sqrt{t_2 - t_1})(X(t_2,\cdot) - X(t_1,\cdot)) \) has distribution \( P \).
In the same way, \((X(\cdot,\tau))_{0 \leq \tau \leq 1}\) is a \( C[0,1] \)-valued Brownian motion under \( P^\infty \).
We call \( P^\infty \) infinite-dimensional Wiener measure or the distribution of "Brownian
sheet". Let us replace \( P^\infty \) with another Gaussian product measure
\( Q^\infty = \prod_{t \in S} Q \in \mathcal{M}_5(S^0) \).

Lemma 5.4 Let \( Q \) be a Gaussian probability distribution on \( C[0,1] \) with
\( E_Q[Y_0(\tau)] = 0 \) \((0 \leq \tau \leq 1)\) and set
\[
X(t,\tau) = Y_0(\tau) \cdot t + \sum_{n=1}^{\infty} \sum_{k=1}^{2^{n-1}} Y_{n,k}(\tau) e_{n,k}(t) \quad (0 \leq t, \tau \leq 1)
\]
where \( Y_0, Y_{n,k}, k = 1, 2, \ldots, 2^{n-1}, n = 1, 2, \ldots \) are independent with distribution \( Q \).
We then have (i) For each \( \tau \in [0,1] \), \((X(t,\tau))_{0 \leq t \leq 1}\) is a Brownian motion with variance
\( E_Q[Y_0(\tau)^2] \).

(ii) The \( C[0,1] \)-valued process \((X(t,\cdot))_{0 \leq t \leq 1}\) has independent increments, and
\( 1/(\sqrt{t_2 - t_1})(X(t_2,\cdot) - X(t_1,\cdot)) \) has distribution \( Q \) \((0 \leq t_1 < t_2 \leq 1)\).

The proof is easy: we refer to [9].

Example 5.5 Let \( Q \in \mathcal{M}_4(C[0,1]) \) be the distribution of an Ornstein–Uhlenbeck
process starting in 0, i.e. the distribution of \((Z_t)_{0 \leq t \leq 1}\), where \((Z_t)\) solves the
stochastic differential equation
\[
dZ_t = dW_t - Z_t dt
\]
and \((W_t)_{0 \leq t \leq 1}\) is a Brownian motion under \( Q \). Then the \( C[0,1] \)-valued process
\((X(t,\cdot))_{0 \leq t \leq 1}\) is a \( C[0,1] \)-valued Ornstein–Uhlenbeck process under \( Q^\infty \), i.e.
\( X_\tau := X(\cdot,\tau) \) solves the ("infinite-dimensional") stochastic differential equation
\[
dX_\tau = dW_\tau - X_\tau d\tau
\]
where \( W_t := W(\cdot, \cdot, t) \) is a \([0, 1]\)-valued Brownian motion under \( Q^\infty \).
We can describe \( Q^\infty \) with Theorem 4.1 as a "large deviation" of \( P^\infty \), i.e. as the solution of a variational problem where we have to minimize the rate function in Theorem 4.1 over a certain subset of \( \mathcal{M}_1(C[0, 1]^{I_0}) \). More precisely, set
\[
A := \{ \overline{R} \in \mathcal{M}_1(C[0, 1]^{I_0}) | \overline{R}|_{\mathcal{F}_0} \in B \}
\]
where
\[
B := \{ R | \int X_t^2 dR \leq 1 - e^{-t}, 0 \leq t \leq 1 \},
\]
\( B \subseteq \mathcal{M}_1(C[0, 1]) \). Here \( A \) is of the form in (4.1), i.e. \( \overline{R} \in A \) iff the one-dimensional marginal distribution of \( \overline{R} \) is in \( B \). The rate function \( I_f(\cdot) \) is here the specific relative entropy \( h(\cdot | P^\infty) \). Since \( Q \) minimizes the relative entropy \( H(\cdot | P) \) over \( B \) (this is shown in [9]), \( Q^\infty \) minimizes \( h(\cdot | P^\infty) \) over \( A \). In this way, we may see the \([0, 1]\)-valued Ornstein-Uhlenbeck process as a large deviation of Brownian sheet.

Acknowledgement. This work is part of a Ph.D. Thesis written at Bonn University. I would like to thank my adviser, Hans Föllmer, and also Michael Röckner for many helpful discussions.

References