Skorokhod Embeddings via Stochastic Flows on the Space of Measures

Ronen Eldan

February 6, 2014

Abstract

We present a new construction of a Skorohod embedding, namely, given a probability measure \( \mu \) with zero expectation and finite variance, we construct an integrable stopping time \( T \) adapted to a filtration \( \mathcal{F}_t \), such that \( W_T \) has the law \( \mu \), where \( W_t \) is a standard Wiener process adapted to the same filtration. We find several sufficient conditions for the stopping time \( T \) to be bounded or to have a sub-exponential tail. In particular, our embedding seems rather natural for the case that \( \mu \) is a log-concave measure and the tail behaviour of \( T \) admits some tight bounds in that case. Our embedding admits the property that the stochastic measure-valued process \( \{ \mu_t \}_{t \geq 0} \), where \( \mu_t \) is the law of \( W_T \) conditioned on \( \mathcal{F}_t \), is a Markov process.

1 Introduction

The Skorokhod embedding problem, first presented by Skorokhod in [S] (1961), was originally formulated as follows: Given a prescribed centered probability measure \( \mu \) whose second moment is finite and a standard Wiener process \( W_t \) adapted to a filtration \( \mathcal{F}_t \), can one find an integrable stopping time \( T \), such that \( W_T \) has the law \( \mu \)?

This problem has encouraged rather extensive research in the past 50 years (e.g., by Azema, Bass, Dubins, Monroe, Obloj, Root, Rost, Yor and many others), some of which is devoted to constructing new solutions, some to formulating and proving more general cases of this problem, and some to establishing certain properties of the existing solutions. Some examples of properties one would be interested to establish about a solution are certain bounds on moments of the stopping time \( T \) given some information about the measure \( \mu \), monotonicity of \( T \) with respect to some quantities related to this measure and bounds related to the set \( \{ W_t, \ 0 \leq t \leq T \} \). For an extensive review of many of these results, the reader is referred to [OE]. See also [AHI] for a more recent construction.
In this note we present a solution to the Skorokhod embedding problem which, to our knowledge, does not appear in the literature. Furthermore, we present several properties that this solution admits, mainly concerning bounds the stopping time $T$ given some assumptions on $\mu$. One of the advantages of this solution is the existence of a certain formula with which the behaviour of $T$ can be analysed in many cases. In particular, the solution seems rather natural for log-concave measures.

The construction also has another property that may be notable: consider the measures $\mu_t$ defined to be the law of $W_T$ conditioned on $\mathcal{F}_t$. The process $\{\mu_t\}_{t \geq 0}$ is, in some sense, a Markov process whose "transition kernel" does not depend on the initial measure $\mu$.

Our construction has some similarities with the localization described in [E]. The main mechanism behind this construction is a certain flow on the space of densities on $\mathbb{R}$, defined by a system of stochastic differential equations. The central ideas behind it are described briefly in the beginning of section 2.

Let us formulate our theorems. Throughout this note, $\mu$ will denote some fixed Borel probability measure on $\mathbb{R}$. The only assumptions we will need for the construction of the Skorokhod embedding are that $\mu$ has expectation zero and a finite second moment:

$$\int_{\mathbb{R}} x^2 d\mu < \infty \text{ and } \int_{\mathbb{R}} x d\mu = 0. \tag{1}$$

Let $W_t$ be a standard Wiener process adapted to a filtration $\mathcal{F}_t$. Our first goal will be to construct a stopping time $T_\mu$. The main properties of this stopping time are described in the following theorem, whose point is that $T_\mu$ induces a Skorokhod embedding of $\mu$ into the probability space of $W_t$.

The actual definition of $T_\mu$ is postponed to the next section.

**Theorem 1.1** Let $\mu$ be a measure on $\mathbb{R}$ satisfying (1). Then the stopping time $T_\mu$ satisfies the following properties:

(i) The event $\{T_\mu \leq t\}$ is measurable with respect to the $\sigma$-algebra generated by $\{W_s\}_{s \leq t}$.

(ii) The random variable $W_{T_\mu}$ is distributed according to the law $\mu$.

(iii) One has $\mathbb{E}[T_\mu] = \text{Var}[\mu]$.

The next theorems in this note establish certain bounds on the distribution of $T_\mu$ given that $\mu$ satisfies some extra assumptions. Several estimates resembling some of our bounds have been established in [AS] for a Skorokhod embedding based on the solution of a backwards stochastic differential equation.

A measure $\mu$ is said to be log-concave if it is either a Dirac $\delta$-measure or has a density $f(x)$ with respect to the Lebesgue measure of the form $f(x) = e^{-\Phi(x)}$ where $\Phi : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ is a convex function. When the measure $\mu$ is log-concave, the stopping time $T_\mu$ admits a sub-exponential tail behaviour, namely we have the following.
Theorem 1.2 There exist universal constants $c, C > 0$ such that, if $\mu$ is a log-concave measure with $\mathbb{E}[\mu] = 0$ and $\text{Var}[\mu] = 1$, then

$$P(T_\mu > t) < Ce^{-ct}.$$ 

Remark 1.3 The above result is tight up to the constants $c, C$. To see this, let $\mu$ be the measure whose density is $\frac{1}{2\sqrt{\pi}} e^{-|x|/\sqrt{2}}$, and let $X$ have the law $\mu$. Then one has

$$P(|X| > x) > 0.5e^{-c_1x}$$

for some $c_1 > 0$. For a fixed $t > 0$, define the events,

$$A = \{T > t\}, \quad B = \{\exists s < t \text{ such that } |W_s| > t\}.$$ 

Then by well-known properties of Wiener processes,

$$P(B) < C_2e^{-c_2t}$$

for some $C, c$. Using a union bound,

$$P(|X| > t) \leq P(A) + P(B) \Rightarrow P(A) \geq 0.5e^{-c_1x} - C_2e^{-c_2t}.$$ 

Now choose $t$ such that $0.5e^{-c_1x} - C_2e^{-c_2t} = 0.4e^{-c_1x}$. Clearly $t > C_3x$ for some $C_3 > 0$, and we have

$$P(T > C_3x) \geq 0.4e^{-c_1x}.$$ 

For log-concave measures whose modulus of log-concavity is bounded from below, the stopping time will be bounded according to the following tight estimate:

Theorem 1.4 Let $\sigma > 0$. Let $\mu$ be a centered probability density such that,

$$\frac{d\mu}{dx}(x) = e^{-\frac{x^2}{2\sigma^2}} \Phi(x)$$

where $\Phi : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ is a convex function. Then one has almost surely,

$$T_\mu \leq \sigma^2.$$ 

Remark 1.5 The above bound is tight, as demonstrated by the case that $\mu$ is a gaussian measure whose variance is $\sigma^2$. 

For a Borel measure $\mu$ on $\mathbb{R}$, we denote by $\text{Supp}(\mu)$ the support of $\mu$ which is the minimal closed set of full measure. Our next task is to address measures whose support is a compact set. For such measures we can give deterministic bounds on the stopping time if the measure is either log-concave or absolutely continuous with respect to the Lebesgue measure with density bounded between two constants in an interval. This is summarized in the next two theorems:  

3
Theorem 1.6 Let \( \mu \) measure on \( \mathbb{R} \) which satisfies:

(i) \( \text{Supp}(\mu) \) is an interval contained in \([-L, L]\) for some \( L > 0 \).

(ii) \( \mu \) is absolutely continuous with respect to the Lebesgue measure and \( \alpha \leq \frac{d\mu}{dx}(x) \leq \beta \) for all \( x \in \text{Supp}(\mu) \). Then,

\[
T_\mu \leq 2L^2 \frac{\beta}{\alpha}
\]

almost surely.

Theorem 1.7 Let \( \mu \) a log concave measure on \( \mathbb{R} \) with \( \text{Supp}(\mu) \subseteq [-L, L] \) for some \( L > 0 \). Then,

\[
T_\mu \leq 2L^2.
\]

almost surely.

Remark 1.8 By inspecting the example of the uniform measure over an interval of length \( 2L \), it can be shown that the bound of the above theorem is tight up to the constant \( 2 \).

The structure of the remainder of this note is the following: In section 2 we construct the stopping time \( T_\mu \) and prove theorem 1.1. Section 3 deals with log-concave measures, in this section we prove theorems 1.2 and 1.4. In section 4 we prove theorems 1.6 and 1.7. In section 5, we define the “Markov property” satisfied by our construction and discuss some other possible constructions of Skorokhod embeddings with this property. In the appendix, we fill in some missing details left open in the construction of \( T_\mu \).

Throughout this note, we use the following notation: for an Itô process \( X_t \), we denote by \( dX_t \) the differential of \( X_t \), and by \( [X]_t \) the quadratic variation of \( X_t \). For a pair of continuous time stochastic processes \( X_t, Y_t \), the quadratic covariation will be denoted by \( [X, Y]_t \). For a measure \( \mu \) on \( \mathbb{R} \), we denote by \( E[\mu] \) and \( \text{Var}[\mu] \) its expectation and variance respectively. By slight abuse of notation, when we write \( X \sim \mu \) we mean that the random variable \( X \) is distributed according to the law of the measure \( \mu \).

Acknowledgements I would like to thank Ofer Zeitouni and Joseph Lehec for fruitful discussions.

2 Construction of the embedding

The goal of this section is to construct the stopping time \( T_\mu \) and to establish some of its basic properties.

Let us briefly describe the idea behind our construction. Given a measure \( \mu \), we will construct a random one-parameter family of measures, \( \{\mu_t\}_{t \geq 0} \)
such that \( \mu_0 = \mu \), and for which there exists some random time \( T > 0 \) such that the density \( \frac{d\mu}{dt} \) is a Gaussian density for all \( 0 \leq t < T \) and \( \mu_t \) is a Dirac \( \delta \)-measure for \( t \geq T \). Moreover,

(i) For any \( x \in \mathbb{R} \), the process \( \frac{d\mu}{dt}(x) \) is a martingale in \( 0 \leq t < T \),

(ii) The process \( \{\int_\mathbb{R} xd\mu_t(x)\}_{t \geq 0} \) is a Brownian motion for \( 0 \leq t < T \),

(iii) \( \mu_t \) converges (in \( L^2 \)) to a Dirac \( \delta \)-measure \( \mu_T \) as \( t \to T \) and \( \mathbb{E}[T] < \infty \),

(iv) The process \( \{\mu_t\} \) is an adapted process with respect to the filtration generated by the Brownian motion.

Recall that \( T \) is the time in which \( \mu_t \) becomes a Dirac measure. Properties (i) and (iii) will ensure us that \( \int_\mathbb{R} xd\mu_T(x) \) will have the law \( \mu \), property (ii) will ensure us that this quantity is, in fact, a brownian motion taken at time \( T \), and property (iv) will ensure us that no ”extra randomness” is used. The construction of the densities \( F_t(x) = \frac{d\mu_t}{d\mu}(x) \) is best described in formula (5) below, and the time \( T = T_\mu \) will be defined as the time in which the solution ”explodes” (hence, the solution ceases to exist).

We begin with some definitions. Let \( \mu \) be a probability measure on \( \mathbb{R} \), satisfying (1). For \( c \in \mathbb{R} \) and \( b \geq 0 \), we write

\[
V_\mu(b, c) = \int_\mathbb{R} e^{cx} - \frac{1}{2}bx^2 d\mu(x)
\]

and define two functions,

\[
a_\mu(b, c) = V_\mu^{-1}(b, c) \int_\mathbb{R} xe^{cx} - \frac{1}{2}bx^2 d\mu(x),
\]

and

\[
A_\mu(b, c) = V_\mu^{-1}(b, c) \int_\mathbb{R} (x - a_\mu(b, c))^2 e^{cx} - \frac{1}{2}bx^2 d\mu(x).
\]

It is easy to verify that under the assumption (1), the functions \( A_\mu \) and \( a_\mu \) are smooth functions in the domain \((b, c) \in (0, \infty) \times \mathbb{R}\).

Let \( W_t \) be a standard Wiener process and consider the following system of stochastic differential equations:

\[
c_0 = 0, \quad dc_t = A_\mu^{-1}(b_t, c_t)dW_t + A_\mu^{-2}(b_t, c_t)a_\mu(b_t, c_t)dt,
\]

\[
b_0 = 0, \quad db_t = A_\mu^{-2}(b_t, c_t)dt.
\]

First, we will explain why the solution exists under a stronger assumption, namely, that \( \mu \) has some finite exponential moment. The proof for the more general case, assuming only that the second moment exists, is left for the appendix. Assume that there exists a constant \( \delta > 0 \) such that

\[
\int_\mathbb{R} \left( e^{\delta x} + e^{-\delta x} \right) d\mu(x) < \infty.
\]

Under this assumption, it is not hard to check that the functions \( A_\mu^{-2}(b, c) \) and \( a_\mu(b, c) \) are smooth functions on the set

\[
(b, c) \in ((0, \infty) \times \mathbb{R}) \cup \{0\} \times [-\delta/2, \delta/2]\).
\]
Since $\frac{dB_t}{B_t}|_{t=0} > 0$, these functions can be modified so that they are smooth on the set $[0, \infty) \times \mathbb{R}$ without affecting the solution. In this case, we can use a standard existence and uniqueness theorem (see e.g., [Ok], section 5.2) to ensure the existence and uniqueness of a solution on some interval $[0, t_0)$ where $t_0$ is an almost-surely positive random variable.  

**Remark 2.1** Note that the fact that $\delta > 0$ is crucial for this argument, and the existence would not generally be true if $\delta = 0$. In general, the functions $A_\mu, a_\mu$ may not be defined in any neighbourhood of $(0,0)$ of the form $[0, \varepsilon] \times (-\varepsilon, \varepsilon)$, and are only bounded in parabolic sets the form $\{(b, c); \varepsilon c^2 < b\}$. A-priori, in order to ensure the existence of the solution, one has to prove that $(b_t, c_t)$ remain in such a set. This will be done in the appendix in an indirect fashion.

We are now ready to define our stopping time $T_\mu$: it will be defined as the supremum over the set of times in which the solution to (2) exists (and define $T_\mu = \infty$ if the solution exists for all $t \geq 0$). It is not hard to verify that the functions $A_\mu^{-1}(b_t, c_t)$ and $a_\mu(b_t, c_t)$ are smooth functions any set in which $b$ is bounded away from zero. Since $b_t$ is increasing and since $\frac{dB_t}{B_t}|_{t=0} > 0$, it follows that for any point $t$ in which $A_\mu^{-1}(b_t, c_t) > 0$ there exists some $\varepsilon > 0$ such that the solution may be continued in the interval $[t, t + \varepsilon]$. Consequently, if we define

$$T_\mu := \sup\{t \geq 0; A_t > 0\};$$

then almost surely, the solution of (2) exists exactly in the interval $[0, T_\mu)$. Our next main goal is to show that $W_{T_\mu}$ has the law $\mu$. We abbreviate $T = T_\mu$.

We begin with the construction a 1-parameter family of measures $\mu_t$ by writing

$$F_t(x) = V_\mu^{-1}(b_t, c_t)e^{c_t x - \frac{1}{2}b_t x^2}$$

and defining the measure $\mu_t$ by,

$$\frac{d\mu_t}{d\mu}(x) = F_t(x)$$

for all $0 \leq t < T$. Also, abbreviate

$$a_t = a_\mu(b_t, c_t), \quad A_t = A_\mu(b_t, c_t), \quad V_t = V_\mu(b_t, c_t),$$

so that $a_t$ and $A_t$ are respectively the expectation and the variance of the measure $\mu_t$.

The following lemma may shed some light on this construction.

**Lemma 2.2** For all $t \in [0, T)$ and for all $x \in \mathbb{R}$, the process $F_t(x)$ satisfies the following set of equations:

$$F_0(x) = 1, \quad dF_t(x) = (x - a_t)A_t^{-1}F_t(x)dW_t,$$

$$a_t = \int_{\mathbb{R}} xF_t(x)d\mu_t(x), \quad A_t = \int_{\mathbb{R}} (x - a_t)^2F_t(x)d\mu_t(x).$$


Proof:
Fix $x \in \mathbb{R}$. We will show that $dF_t(x) = (x - a_t)A_t^{-1}F_t(x)dw_t$. The correctness of the other equations is obvious. Define,

$$G_t(x) = V_tF_t(x) = e^{\frac{1}{2}bt^2x^2}.$$  

Equation (2) clearly implies that $|b|t = 0$. Moreover,

$$dc_t = A_t^{-1}dw_t + A_t^{-2}a_tdといった$$

It follows that,

$$d[c]_t = A_t^{-2}dt.$$  

Using Itô’s formula, we calculate

$$dG_t(x) = \left( cd_t - \frac{1}{2}d_t^2x^2 + \frac{1}{2}x^2d[c]_t \right) G_t(x) = (xA_t^{-1}dw_t + xA_t^{-2}a_td_ft - \frac{1}{2}A_t^{-2}x^2dt + \frac{1}{2}A_t^{-1}x^2dt) G_t(x) = x(A_t^{-1}dw_t + A_t^{-2}a_td_t) G_t(x).$$  

Next, we calculate,

$$dV_t = d\int e^{\frac{1}{2}bt^2x^2}d\mu(x) = \int dG_t(x)d\mu(x) = \int x(A_t^{-1}dw_t + A_t^{-2}a_tdw_t) G_t(x)d\mu(x) = V_tA_t^{-1}dw_t + V_tA_t^{-2}a_tdw_t.$$  

So, using Itô’s formula again,

$$dV_t^{-1} = -\frac{dV_t}{V_t^2} = -\frac{d[V]_t}{V_t^2} = -V_t^{-1}a_t(A_t^{-1}dw_t + A_t^{-2}a_tdw_t) + V_t^{-1}A_t^{-2}a_tdw_t.$$  

Applying Itô’s formula one last time yields,

$$dF_t(x) = d(V_t^{-1}G_t(x)) = G_t(x)dV_t^{-1} + V_t^{-1}dG_t(x) + d[V^{-1}, G(x)]_t = -V_t^{-1}a_t(A_t^{-1}dw_t + A_t^{-2}a_tdw_t) G_t(x) + V_t^{-1}A_t^{-2}a_t^2G_t(x)dt + V_t^{-1}x(A_t^{-1}dw_t + A_t^{-2}a_tdw_t) G_t(x) - A_t^{-2}a_txV_t^{-1}G_t(x)dt = A_t^{-1}dw_t(x - a_t)F_t(x).$$

This finishes the proof. 

An immediate consequence of this theorem is the following

**Lemma 2.3** One has,

(i) Almost surely, for every $0 \leq t < T$, $\mu_t$ is a probability measure.

(ii) For every measurable $E \subseteq \mathbb{R}$, the process $\mu_tE$ is a martingale for $0 \leq t < T$. 


Proof:
Part (ii) of the lemma follows immediately from formula (5). To prove part (i), we write
\[
\int \mu_t(x) = \int F_t(x) \mu(x) = V_t^{-1} \int e^{c_t x - \frac{b_t x^2}{2}} \mu(x) = 1.
\]

The next lemma is a simple calculation that extracts one of the main points of the construction.

**Lemma 2.4** For all \(0 \leq t < T\), one has \(a_t = W_t\).

**Proof:**
Using formula (5), we calculate
\[
da_t = d \int x F_t(x) \mu(x) = \int x d F_t(x) \mu(x) = \left( \int x(\mu_t(x) - \int x d \mu_t(x)) \right) dW_t.
\]

Now, by the definition of \(a_t\), one has
\[
\int a_t (x - a_t) A_t^{-1} F_t(x) \mu(x) = A_t^{-1} (a_t \mu_t(R) - \int x d \mu_t(x)) = 0
\]

Joining the two previous equations together gives
\[
da_t = \left( \int (x - a_t)^2 A_t^{-1} F_t(x) \mu(x) \right) dW_t = A_t^{-1} \left( \int (x - a_t)^2 F_t(x) \mu(x) \right) dW_t = dW_t.
\]

Next, we prove

**Lemma 2.5** One has for all \(0 \leq t < T\),
\[
dA_t = \left( \int (x - W_t)^3 \mu(x) \right) A_t^{-1} dW_t - dt.
\]

Moreover, \(T\) is almost-surely finite.

**Proof:**
We begin by calculating, using Itô’s formula:
\[
da_t = d \left( \int x^2 F_t(x) \mu(x) - a_t^2 \right) = d \left( \int x^2 F_t(x) \mu(x) - W_t^2 \right)
\]
\[
\left( \int x^2 A_t^{-1} (x - W_t) d \mu(x) \right) dW_t - 2W_t dW_t - dt =
\]

\[
8
\]
\[
\left( \int_{\mathbb{R}} \left( x^2 - 2W_t x + W_t^2 \right) A_t^{-1} (x - W_t) d\mu_t(x) \right) dW_t - dt = \\
\left( \int_{\mathbb{R}} (x - W_t)^3 d\mu_t(x) \right) A_t^{-1} dW_t - dt
\]
which settles (7).

To see that \( T \) is almost surely finite, write \( X_t = A_t + t \). The above equation suggests that \( X_t \) is a martingale up to time \( T \). Suppose by contradiction that with positive probability, a solution exists for all \( t > 0 \). This implies that \( A_t \) exists and is positive for all \( t > 0 \). By the martingale convergence theorem, we have
\[
P \left( \lim_{t \to \infty} X_t \text{ exists} \mid \forall t > 0, X_t \geq 0 \right) = 1,
\]
but observe that when \( \lim_{t \to \infty} X_t \) exists then \( \lim_{t \to \infty} A_t = -\infty \) which is clearly impossible. The lemma is complete.

We are now ready to establish the central property of \( T_{\mu} \) being a Skorokhod embedding, namely to show that \( W_T \) has the law \( \mu \).

**Proof of theorem 1.1**

Part (i) of the theorem is obvious from the definition of \( T \). To prove part (ii), let \( \varphi(x) \) be a smooth, compactly supported function. We have for all \( 0 \leq t < T \),
\[
\left| \int_{\mathbb{R}} \varphi(x) d\mu_t(x) - \varphi(W_t) \right| = \left| \int_{\mathbb{R}} (\varphi(x) - \varphi(a_t)) d\mu_t(x) \right| \leq \\
\sup_{x \in \mathbb{R}} |\varphi'(x)| \int_{\mathbb{R}} |x - a_t| d\mu_t(x) \leq \sup_{x \in \mathbb{R}} |\varphi'(x)| A_t^{1/2}.
\]
where in the first passage we used lemma 2.4 and in the last passage we used the Cauchy-Schwartz inequality and the fact that \( \text{Var}[\mu_t] = A_t \). Since \( \lim_{t \to T^-} A_t = 0 \), it follows that
\[
\lim_{t \to T^-} \int_{\mathbb{R}} \varphi(x) d\mu_t = \int_{\mathbb{R}} \varphi(x) d\mu_T. \tag{8}
\]
where we define the measure \( \mu_T \) to be a Dirac probability measure with an atom at \( W_T \). Recall that for a measurable set \( E \), \( \mu_t(E) \) is a martingale for \( t < T \). It follows that \( \int \varphi(x) d\mu_t(x) \) is also a martingale. By the optional stopping theorem,
\[
E \left[ \int_{\mathbb{R}} \varphi(x) d\mu_{T \wedge t}(x) \mid \mathcal{F}_s \right] = \int_{\mathbb{R}} \varphi(x) d\mu_{T \wedge s}(x), \quad \forall t > s \geq 0. \tag{9}
\]
Now, since \( \varphi \) is bounded and since \( \mu_{t \wedge T} \) is a probability measure, one has
\[
\int_{\mathbb{R}} |\varphi(x)| d\mu_{T \wedge t}(x) \leq \sup_{x \in \mathbb{R}} |\varphi(x)|, \quad \forall t \geq 0
\]
and the dominated convergence theorem implies that
\[
\lim_{t \to \infty} \mathbb{E} \left[ \int_R \varphi(x) d\mu_T(x) \bigg| F_s \right] = \mathbb{E} \left[ \lim_{t \to \infty} \int_R \varphi(x) d\mu_t(x) \bigg| F_s \right] = \\
\mathbb{E} \left[ \int_R \varphi(x) d\mu_T(x) \bigg| F_s \right]
\]
for all \( s \geq 0 \) where we used the fact that \( T \) is almost surely finite, proven in lemma 2.5, with formula (9). Combining the last equality with (9) gives,
\[
\mathbb{E} \left[ \int_R \varphi(x) d\mu_T(x) \bigg| F_s \right] = \int_R \varphi(x) d\mu_{T \land s}(x)
\]
for all \( s \geq 0 \). Taking \( s = 0 \) proves part (ii) of the theorem.

To prove part (iii) of the theorem, observe that it follows from the optional stopping theorem that
\[
\mathbb{E}[W_{T \land t}^2] = \mathbb{E}[T \land t], \quad \forall t > 0.
\]
By taking the limit \( t \to \infty \) on both sides, we see that it suffices to show that
\[
\lim_{t \to \infty} \mathbb{E}[W_{T \land t}^2] = \mathbb{E}[W_T^2] = \text{Var}[\mu].
\]
To that end, for all \( t > 0 \) define \( X_t = W_T - W_{t \land T} \). Equation (10) implies that for any compactly supported continuous test function \( \varphi \), one has
\[
\mathbb{E}[\varphi(X_s) | F_s] = \int_R \varphi(x - W_{s \land T}) d\mu_{s \land T}(x), \quad \forall s \geq 0.
\]
where we used the fact that \( W_{s \land T} \) is \( F_s \)-measurable. Consequently,
\[
\text{Var}[X_s | F_s] = A_{s \land T}, \quad \forall s \geq 0
\]
(where we define \( A_T = 0 \)). Moreover, since \( \mathbb{E}[\mu_{s \land T}] = W_{s \land T} \), it follows that \( \text{Cov}(X_s, W_{s \land T}) = 0 \) for all \( s \geq 0 \), which gives
\[
\text{Var}[X_s] + \text{Var}[W_{s \land T}] = \text{Var}[\mu]
\]
for all \( s \geq 0 \). Consequently,
\[
\text{Var}[W_{s \land T}] \leq \text{Var}[\mu], \quad \forall t \geq 0
\]
which implies that
\[
\lim_{t \to \infty} \mathbb{E}[W_{T \land t}^2] \leq \text{Var}[\mu]
\]
where the fact that the limit exists follows from the fact that the limit of the right hand side of formula (11) exists. We conclude that
\[
\mathbb{E}[T] = \lim_{t \to \infty} \mathbb{E}[T \land t] \leq \text{Var}[\mu].
\]
Now, since \( T \) has a finite first moment, we can use the optional stopping theorem once again with the martingale \( W_T^2 - t \) to get
\[
\mathbb{E}[W_T^2] = \mathbb{E}[T]
\]
and the theorem is complete. \( \square \)
3 Log concave measures

We begin by recalling a few basic things about log-concave measures. A log concave measure \( \mu \) on \( \mathbb{R} \) is either a Dirac measure or is absolutely continuous with respect to the Lebesgue measure.

A central tool we will use will be the following well-known estimate, proven via integration by parts:

**Theorem 3.1** Let \( V : \mathbb{R} \to \mathbb{R} \) be a strictly convex function, such that \( \int_{\mathbb{R}} e^{-V(x)} dx = 1 \). Let \( \mu \) be a probability measure on \( \mathbb{R} \) defined by \( \frac{d\mu}{dx} = e^{-V(x)} \). Then for every smooth function \( f : \mathbb{R} \to \mathbb{R} \),

\[
\int_{\mathbb{R}} \left( f(x) - \int_{\mathbb{R}} f(x) d\mu(x) \right)^2 d\mu(x) \leq \int_{\mathbb{R}} (V''(x))^{-1} (f'(x))^2 d\mu(x).
\]

**Remark 3.2** The above theorem is merely the one-dimensional version of a theorem of Brascamp-Lieb from [BL].

An application of this theorem with the function \( f(x) = x^2 \) gives,

**Proposition 3.3** Let \( \phi : \mathbb{R} \to \mathbb{R} \cup \{\infty\} \) be a convex function and let \( \sigma > 0 \). Suppose that,

\[
d\mu(x) = Z e^{-\phi(x) - \frac{1}{2\sigma^2} |x|^2} dx
\]

Then one has,

\[
\text{Var}[\mu] \leq \sigma^2.
\]

As a corollary, we have the following:

**Corollary 3.4** If \( \mu \) is a log-concave measure then

\[
A_t \leq b_t^{-1}
\]

for all \( 0 \leq t < T \).

Proof: Recall the formula (4), and apply proposition 3.3 with \( \sigma^2 = \frac{1}{b_t^2} \).

Recall that \( dB_t = A_t^{-2} dt \). In light of this equation, and with the help of the above corollary, we have the following bound for \( T \):

**Lemma 3.5** There exists a universal constant \( C > 0 \) such that the following holds whenever \( \mu \) is a log-concave measure: define the stopping time

\[
\tau = \min(\inf\{t; A_t \geq 2\}, 1).
\]

One has almost surely,

\[
T \leq C + \frac{C}{\tau}.
\]
Proof: If $T \leq 1$ then we’re done. Otherwise, we note that by the definition of $\tau$ and $b_t$ one has,
\[ db_t > \frac{1}{4} dt, \quad \forall 0 \leq t \leq \tau \]
which implies,
\[ b_{\tau} \geq \frac{\tau}{4}. \]
Combine the equation defining $b_t$, $db_t = A_t^{-2} dt$ with equation (14) to get,
\[ \frac{db_t}{dt} = A_t^{-2} \geq b_t^2. \]
Let $g(t)$ be the function solving the equations
\[ g(0) = \frac{\tau}{4}, \quad g'(t) = g(t)^2. \]
The reader may verify that the function $g(t) = \frac{1}{\tau - t}$ solves these equations.
Note that $g$ explodes for $t = \frac{\tau}{4}$. The fact that $b_{\tau} \geq g(0)$ and an application of a standard comparison theorem thus gives,
\[ T \leq \tau + 4 \leq 1 + \frac{4}{\tau}. \]
The proof is complete.

We are now ready to prove that when $\mu$ is log-concave, $T_\mu$ has a sub-exponential tail.

Proof of theorem: Recall equation (7),
\[ dA_t = \left( \int_\mathbb{R} (x - W_t)^3 d\mu_t(x) \right) A_t^{-1} dW_t - dt \]  
Define $S_t = \left( \int_\mathbb{R} (x - W_t)^3 d\mu_t(x) \right) A_t^{-1}$. A well-known fact about isotropic log-concave measures (see for example [LV, Lemma 5.7]) is that for every $p \geq 2$ there exists a constant $c(p)$ such that for every log-concave measure $\nu$ on $\mathbb{R},$
\[ \text{Var}[\nu]^{p/2} \leq \int |x - \mathbb{E}[\nu]|^p d\nu(x) \leq c(p) \text{Var}[\nu]^{p/2}. \]
Using the above with the measure $\mu_t$ and with $p = 3$ gives $|S_t| \leq C_1 A_t^{1/2}$ for some universal constant $C_1 > 0$ and for all $t > 0$. With the definition of $\tau$, this gives
\[ S_t < 2C_1, \quad \forall 0 \leq t < \min(\tau, T). \]  
Next, define $Y_t = A_t + t - 1$. By (7), we learn that $Y_t$ is a semi-martingale. By the Dambis / Dubins-Schwartz theorem, there exists a monotone time change $\Theta(t)$ such that $Y_{\Theta(t)} \sim \tilde{W}_t$ where $\tilde{W}_t$ is a standard Wiener process defined in the interval $[0, \Theta^{-1}(T)]$. Moreover,
\[ \Theta'(t) = \frac{d}{dt} [Y_t]^{-1} = S_t^{-2}. \]
Equation (16) implies,
\[ \Theta(t) \geq c_2 t, \quad \forall 0 \leq t \leq \min(\tau, T). \] (17)
for some universal constant \( c_2 > 0 \). An application of the so-called reflection principle now gives,
\[ P(\max_{t \in [0, p]} \tilde{W}_t \geq 2) = 2P(\tilde{W}_p \geq 2) < C_3 e^{-1/p}, \]
for some universal constant \( C_3 > 0 \), which implies that
\[ P(1/\tau > s) = P(\max_{0 \leq t \leq s-1} Y_t > 2) \leq P(\max_{0 \leq t \leq s-1} Y_{\Theta(t)} > 2) < C_3 e^{-c_2 s}. \]
Combining the last equation with lemma (3.5) finishes the proof.

We move on to proving theorem 1.4 which states that if the density of \( \mu \) with respect to some Gaussian measure is a log-concave function, then \( T \) is bounded by the variance of this Gaussian measure.

**Proof of theorem 1.4**
Thanks to the assumption of the theorem and to equation (4), we know that for all \( 0 \leq t < T \), \( \mu_t \) has the form
\[ \frac{d\mu_t}{dx}(x) = e^{-\left(\frac{1}{2}b_t x^2 + \frac{1}{2}\Phi_t(x)\right)} \]
for some \( \Phi_t : \mathbb{R} \to \mathbb{R} \cup \{\infty\} \) convex. Along with proposition 3.3 this gives
\[ A_t^{-1} \geq \sigma^{-2} + b_t, \quad \forall 0 \leq t < T. \]
Define \( e_t = \sigma^{-2} + b_t \). Combine this with the equation defining \( b_t, \) \( db_t = A_t^{-2} dt \) and with equation (14) to get,
\[ \frac{d}{dt} e_t = A_t^{-2} \geq e_t^2. \]
Note that the function \( g(x) = 1/(\sigma^2 - x) \) solves the equations
\[ g(0) = \frac{1}{\sigma^2}, \quad g'(x) = g^2(x). \]
A standard comparison theorem gives \( e_t \geq g(t) \) for all \( 0 \leq t \leq \sigma^2 \), which implies that there exists some \( t_0 \leq \sigma^2 \) such that
\[ \lim_{t \to t_0^+} b_t = +\infty \Rightarrow \lim_{t \to t_0^+} A_t = 0. \]
Thus,
\[ T \leq \sigma^2. \]
4 Measures with bounded support

Let \( \mu \) be a measure supported in the interval \([-L, L]\). Our main mean of using this fact will be the obvious observation that

\[
A_t \leq L^2, \quad \forall 0 \leq t < T.
\]  

(18)

The next lemma will be the main ingredient allowing us to take advantage of the fact that a measure has a density bounded between two constants on its support:

**Lemma 4.1** Suppose that \( \mu \) is absolutely continuous with respect to the Lebesgue measure on \( \mathbb{R} \) supported on some interval \( I \) and suppose that \( f(x) = \frac{d\mu}{dx}(x) \) satisfies

\[
0 < \alpha \leq f(x) \leq \beta, \quad \forall x \in I
\]

Let \( a \in \mathbb{R} \) and \( b > 0 \) and let \( \nu \) be a probability measure defined by the equation

\[
\frac{d\nu}{d\mu}(x) = Z^{-1} e^{-\frac{b}{2}(x-a)^2}
\]

where \( Z > 0 \) is a normalizing constant. Then

\[
\text{Var}[\nu] \leq \frac{\beta}{\alpha b}.
\]

**Proof:**

Define

\[
x_0 = \frac{\int_I x e^{-\frac{b}{2}(x-a)^2} dx}{\int_I e^{-\frac{b}{2}(x-a)^2} dx}.
\]

An application of proposition (3.3) with the function \( 1_{\{x \in I\}} e^{-\frac{b}{2}(x-a)^2} \) gives

\[
\frac{\int_I (x - x_0)^2 e^{-\frac{b}{2}(x-a)^2} dx}{\int_I e^{-\frac{b}{2}(x-a)^2} dx} \leq \frac{1}{b}.
\]

Now, we have

\[
\text{Var}[\nu] \leq \int_I (x - x_0)^2 d\nu(x) = \frac{\int_I (x - x_0)^2 e^{-\frac{b}{2}(x-a)^2} d\mu(x)}{\int_I e^{-\frac{b}{2}(x-a)^2} d\mu(x)} \leq \frac{\beta}{\alpha b} \int_I (x - x_0)^2 e^{-\frac{b}{2}(x-a)^2} dx \leq \frac{\beta}{\alpha b}.
\]

\( \square \)

By combining the above lemma with equation (18), we establish the bound for measures supported on an interval whose density is bounded between two constants:
Proof of theorem 1.6:
We conclude from the previous lemma that for all $0 \leq t < T$,
\[ A_t \leq \frac{\beta}{\alpha} b_t^{-1}. \]  
(19)
Using this estimate and the estimate (18) with the equation (2) for $db_t$
\[ \frac{d}{dt} b_t = A_t^2 \geq \max \left( \frac{\alpha^2}{\beta^2} b_t^2, \frac{1}{L^4} \right). \]  
(20)
Let $g(x)$ be a function satisfying
\[ g'(0) = \frac{1}{L^4}, \quad g'(x) = \frac{\alpha^2}{\beta^2} g(x)^2. \]
Then,
\[ g(x) = \frac{1}{L^2 \alpha} - \frac{\alpha^2}{\beta^2} x. \]
Define $t_0 = \frac{L^2 \alpha}{\beta^2}$ and note that, by (20),
\[ b_{t_0} \geq \frac{\beta}{\alpha L^2} = g(0). \]
By a standard comparison theorem
\[ b_{t+t_0} \geq g(t), \quad \forall 0 \leq t < T, \]
which implies that for some $t_1 \leq 2L^2 \frac{\beta}{\alpha}$, one has
\[ \lim_{t \to t_1} b_t = +\infty \]
In light of formula (19), this implies that
\[ T \leq 2L^2 \frac{\beta}{\alpha} \]
and the proof is complete.

The proof of theorem 1.7 follows the same lines, only lemma 4.1 is replaced by proposition 3.3:

Proof of theorem 1.7:
Using corollary 3.4 and equation (18) gives,
\[ A_t^{-2} \geq \max \left( b_t^2, \frac{1}{L^4} \right). \]
Plugging this into the formula for $db_t$, equation (2), gives
\[ \frac{d}{dt} b_t \geq \max \left( b_t^2, \frac{1}{L^4} \right). \]
now follow the proof of theorem 1.6 noting that equation (20) holds with $\frac{\alpha}{\beta} = 1$.  

5 Embeddings with a Markov property

The embedding constructed in section 2 admits a certain property that, as far as the author is informed, other known constructions of Skorokhod embeddings don’t seem to attain. To formulate this property, we begin with some definitions.

Let $W_t$ be a standard Wiener process with a corresponding filtration $\mathcal{F}_t$. Let $\mathcal{M}$ be the space of Borel probability measures $\mu$ on $\mathbb{R}$ such that $\text{Var}[\mu] < \infty$ and let $\mathcal{M}' \subset \mathcal{M}$ be the subsets of measures whose expectation is zero. Let $\mathcal{T}$ be the space of $\mathcal{F}$-stopping times. We define a Skorokhod embedding scheme as a function $T : \mathcal{M}' \to \mathcal{T}$ taking $\mu$ to a stopping time $T(\mu)$ such that $W_{T(\mu)}$ the law $\mu$.

For a measure $\mu$ and a Skorokhod embedding scheme $T$, define a random measure $\mu_t$ by,

$$\mu_t(E) = \mathbb{P}(W_{T(\mu)} \in E \mid \mathcal{F}_t)$$

for all measurable $E \subset \mathbb{R}$. It is easy to verify that $T$ is uniquely determined by the measure $\mu_t$ by the formula

$$T = \sup\{t \mid \text{Var}[\mu_t] > 0\}. \quad (22)$$

We denote by $\mathcal{SM}$ the space of signed measures on $\mathbb{R}$. Our main definition will be:

**Definition 5.1** (Markov property) A Skorokhod embedding scheme $T$ has a Markov property if there exists a function $F : \mathcal{M} \to \mathcal{SM}$ such that

$$d\mu_t = F(\mu_t)dW_t, \quad (23)$$

for all $\mu \in \mathcal{M}$ and for all $t > 0$, where $\mu_t$ is defined as in (21).

It is obvious from the definition of the stopping time $T_\mu$ constructed in section 2, that it has a Markov property.

**Remark 5.2** The above definition may be natural in a financial context: when one chooses a market strategy that maximizes the expectation of a certain quantity, in many cases the optimal strategy need not take the past into account, since the market, $W_t$, is a Markov process.

**Remark 5.3** Observe that if $T$ has a Markov property, then one has in particular,

$$T(\mu) | \mathcal{F}_t \sim T(\mu_t) + t$$

for every $\{W_t\}_t$ and every $t$ such that $t < T(\mu)$.

This definition gives rise to an entire family of Skorokhod embeddings, as demonstrated by the following proposition. We will omit its proof, as it basically follows the same lines as the ones described in section 2.
Proposition 5.4 Suppose that for a function $F : \mathcal{M} \to \mathcal{S}\mathcal{M}$, equation (23) has a unique solution given $\mu_0 = \mu$. In addition, suppose that the following conditions hold:

(i) $F(\mu)(\mathbb{R}) = 0$ for every $\mu \in \mathcal{M}$,

(ii) $\int_{\mathbb{R}} (x - E[\mu])dF(\mu)(x) = 1$ for every $\mu \in \mathcal{M}$, and

(iii) $\mathbb{E}[\sup\{t : \text{Var}[\mu_t] > 0\}] < \infty$.

Then the stopping time induced by equations (22) and (23) is a Skorokhod embedding.

It may be natural to consider the following construction: Denote the median of a measure $\nu$ by $\text{med}(\nu)$. Consider the equation $F_0 = 1, dF_t = (x - \text{med}(\mu_t))A_t^{-1}dW_t$,

$$\frac{d\mu_t}{d\mu} = F_t, A_t = \int_{\{x > \text{med}(\mu_t)\}} xd\mu_t(x) - \int_{\{x < \text{med}(\mu_t)\}} xd\mu_t(x)$$

note that if the solution of the above equations exists for some initial measure $\mu_0 = \mu$, then conditions (i) and (ii) in the last proposition hold if the measure $\mu$ has no atoms.

Question 5.5 Does the above construction induce a Skorokhod embedding for measures with no atoms?

6 Appendix

In this appendix we prove that the equation (2) has a unique solution whenever the second moment of $\mu$ is finite.

We begin with observing that it is enough to prove that almost surely, there exists some $t_0 > 0$ such that the equation has a solution in the interval $[0, t_0]$. Indeed, $A_t(b, c)$ and $a_t(b, c)$ are smooth functions any set in which $b$ is bounded away from zero. Since $A_t$ is continuous with respect to $t$ and since $A_0 > 0$, there will necessarily exist some $b' > 0$ such that $b_t \geq b'$ for all $t \geq t_0$ (in other words, the only "problematic" point is $t = 0$, since for any $t > 0$ the function $\mu_t$, surely has finite exponential moments).

We argue that there exists a function $c_\mu(\cdot, \cdot)$ satisfying,

$$a_\mu(b, c_\mu(a, b)) = a.$$  \hspace{1cm} (24)

for all $b > 0$ and $a \in \text{Conv}(\text{Supp}(\mu))$ (the convex hull of the support of $\mu$). Indeed, if we denote that partial derivatives of $a_\mu$ by $a_1(\cdot, \cdot)$ and $a_2(\cdot, \cdot)$ a straightforward calculation gives

$$a_2(b, c) = V_\mu(b, c)^{-1}\int_{\mathbb{R}} (x - a_\mu(b, c))e^{-\frac{1}{2}b^2x^2}d\mu(x) = A_\mu(b, c) > 0.$$  \hspace{1cm} (25)

for all $b > 0$. Moreover, it is easy to check that for all $b > 0$,

$$\lim_{c \to -\infty} a_\mu(b, c) = \min \text{Supp}(\mu) \text{ and } \lim_{c \to \infty} a_\mu(b, c) = \max \text{Supp}(\mu).$$
Observe also that $a_\mu$ is continuous on the domain $b > 0$. It follows from inverse function theorem that the function $c_\mu$ exists and is unique. Fix a realization of $W_t$ and consider the ODE,

$$\frac{d}{dt} b_t = A_\mu(b_t, c_\mu(W_t, b_t))^{-2}, \quad b_0 = 0.$$  

It is not hard to verify that $G(b, w) = A_\mu(b, c_\mu(w, b))^{-2}$ and $\frac{d}{dt} G(b, w)$ are continuous functions in the domain $(b, w) \in (0, \infty) \times \mathbb{R} \cup (0, 0)$. It follows from the standard theory of ordinary differential equations that there exists a unique solution $b_t$ on some interval $t \in [0, t_0]$.

Next define $c_t = c_\mu(W_t, b_t)$. Our main goal is to show that $b_t, c_t$ satisfy the equation (2). To that end, we use Itô’s formula to calculate

$$dc_t = c_1(W_t, b_t) dW_t + \frac{1}{2} c_{11}(W_t, b_t) dt + c_2(W_t, b_t) \frac{d}{dt} b_t dt,$$

where $c_{11}((\cdot, \cdot))$ is the second derivative of $c_\mu$ with respect to its first variable. According to the inverse function theorem and using equation (25),

$$c_1(a, b) = A_\mu(b, c_\mu(a, b))^{-1}.$$

Next, we have

$$a_1(b, c) = V^{-1}(b, c_\mu(a, b)) \left( \int R e^{c_\mu(a, b)x-\frac{1}{2} b^2} dx - \int R e^{c_\mu(a, b)x-\frac{1}{2} b^2} dx \right).$$

Differentiate equation (24) to get

$$a_2(b, c_\mu(a, b)) c_2(a, b) = 0$$

So,

$$c_2(a, b) = a_1(b, c_\mu(a, b)) a_2(b, c_\mu(a, b))^{-1} = A_\mu(b, c_\mu(a, b))^{-1} V^{-1}(b, c_\mu(a, b)) \left( \int R e^{c_\mu(a, b)x-\frac{1}{2} b^2} dx - \int R e^{c_\mu(a, b)x-\frac{1}{2} b^2} dx \right).$$

Lastly, we need the second derivative of $c$ with respect to the first variable. One has

$$c_{11}(a, b) = -a_2(b, c_\mu(a, b)) / a_2(b, c_\mu(a, b))^3 = A_\mu(b, c_\mu(a, b))^{-3} \left( V^{-1}(b, c_\mu(a, b))^{-1} \int R (x^3 - ax^2) e^{c_\mu(a, b)x-\frac{1}{2} b^2} dx + 2 A_\mu \right).$$

We finally get

$$dc_t = c_1(W_t, b_t) dW_t + \frac{1}{2} c_{11}(W_t, b_t) dt + c_2(W_t, b_t) \frac{d}{dt} b_t dt =$$

$$A_\mu(b_t, c_\mu(W_t, b_t))^{-1} dW_t + \frac{1}{2} A_\mu(b_t, c_t)^{-3} \left( V^{-1}(b_t, c_t)^{-1} \int R (x^3 - W_t x^2) e^{c_\mu(W_t, b_t)x-\frac{1}{2} b_t^2} dx - 2 A_\mu(b_t, c_t) \right) dt +$$

$$V^{-1}(b_t, c_t) \left( \int R e^{c_\mu(W_t, b_t)x-\frac{1}{2} b_t^2} dx W_t - \int R e^{c_\mu(W_t, b_t)x-\frac{1}{2} b_t^2} dx \right) A_\mu(b_t, c_\mu(W_t, b_t))^{-3} dt =$$

$$A_\mu(b_t, c_t)^{-2} dW_t + W_t A_\mu(b_t, c_t)^{-2} dt.$$  

In view of (6), we conclude that $b_t$ and $c_t$ satisfy (2) in some interval $[0, t_0)$, which proves the existence and uniqueness of the solution.
References

[AHI] S. Ankirchner, G. Heyne and P. Imkeller A BSDE approach to the Skorokhod embedding problem for the Brownian motion with drift. Stoch. Dyn. 08, 35 DOI: 10.1142/S0219493708002160 (2008)

[AS] S. Ankirchner and P. Strack Skorokhod embeddings in bounded time Stoch. Dyn. 11, No. 2-3, 215-226 ISSN 0219-4937 (2011).

[BL] H. J. Brascamp and E. H. Lieb. On extensions of the Brunn-Minkowski and Prékopa Leindler theorems, including inequalities for log concave functions, and with an application to the diffusion equation. J. Functional Analysis, 22(4):366-389, 1976

[Dur] R Durrett, Stochastic Calculus: A Practical Introduction Cambridge university press, 2003.

[E] R. Eldan Thin shell implies spectral gap up to polylog via a stochastic localization scheme. Geometric Aspects of Functional Analysis (to appear) (2012).

[LV] L. Lovász and S. Vempala, The geometry of logconcave functions and sampling algorithms. Random Structures & Algorithms, Vol. 30, no. 3, (2007), 307–358.

[Ob] J.K Obłój, The Skorokhod embedding problem and its offspring, Probability Surveys, 1, 321-392 (2004).

[Ok] B. Oksendal Stochastic Differential Equations: An Introduction with Applications. Berlin: Springer. ISBN 3-540-04758-1, (2003).

[RY] D. Revuz and M. Yor. Continuous martingales and Brownian motion, volume 293 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, Berlin, third edition, (1999)

[S] A. V. Skorohod. Studies in the theory of random processes. (Russian) Izdat. Kiev. Univ., Kiev MR185619 (1961).

e-mail address: ronenedan@gmail.com

February 6, 2014