Explicit solution of relative entropy weighted control

Joris Bierkens*, Hilbert J. Kappen†

May 1, 2014

Abstract

We consider the minimization over probability measures of the expected value of a random variable, regularized by relative entropy with respect to a given probability distribution. In the general setting we provide a complete characterization of the situations in which a finite optimal value exists and the situations in which a minimizing probability distribution exists. Specializing to the case where the underlying probability distribution is Wiener measure, we characterize finite relative entropy changes of measure in terms of square integrability of the corresponding change of drift. For the optimal change of measure for the relative entropy weighted optimization, an expression involving the Malliavin derivative of the cost random variable is derived. The theory is illustrated by its application to the case where the cost variable is the maximum of a standard Brownian motion over a finite time horizon.

Key words and phrases: stochastic optimal control, Itô calculus, Brownian motion, Malliavin calculus, diffusions, relative entropy, Kullback-Leibler divergence.

MSC2010 subject classifications: primary 93E20; secondary 60H07, 94A17.

The research leading to these results has received funding from the European Community’s Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 270327

1 Introduction

In certain situations in stochastic optimal control theory, the dynamic programming or Hamilton-Jacobi-Bellman equations may be transformed, through the Hopf-transform, into linear equations [Fle82], [FS09, Chapter VI]. In the past years, within the applied control and machine learning community, there has been a significant amount of interest in this class of problems (see e.g. [Kap05, Lod06, BH13]). This class of problems also occurs in risk sensitive control theory (see [FS09]) and the theory of large deviations (see [BD98]), and it occurs in modified form (constrained to equivalent martingale measures) in mathematical finance, in particular as the dual problem for a portfolio optimization problem [Mon13]. It is the goal of this paper to review and extend the mathematical underpinning of this optimization problem, as well as showcase some new results within this context.

The problem we consider is a minimization problem over probability measures that are absolutely continuous with respect to a given probability measure (referred to as the ‘uncontrolled measure’). The functional we wish to minimize is the sum of (i) the expectation of a given random variable with respect to any probability measure, and (ii) the relative entropy of that probability measure with respect to the uncontrolled measure. The density of the optimal probability measure with respect to the uncontrolled distribution is readily available through an explicit expression in terms of the cost random variable. The challenge is then to understand this probability measure within the context of the underlying problem.

* Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, The Netherlands, j.bierkens@science.ru.nl
† Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, The Netherlands, b.kappen@science.ru.nl
In particular, in the special case in which we are interested in this paper, the uncontrolled
distribution is Wiener measure on the space of continuous sample paths. Absolutely continuous
change of distribution then corresponds, by the Girsanov theorem, to a change of drift, which
we will interpret as control process. The regularizing relative entropy corresponds to squared
control cost, as we will discuss. The questions we wish to answer in this paper are: (i) under what
conditions does there exist an optimal change of drift corresponding to a given cost functional and
probability measure, and (ii) how can it be computed?

There is a closed relation to existing theory within the field of large deviations theory and
stochastic optimal control. For the reader who is familiar with this literature (in particular with
[BD98, DE97]), the new results of our paper consist of:

(i) A complete characterization of relative entropy weighted optimization problems in the gen-
eral setting (i.e. where we are only given a cost functional and a reference probability
measure). This extends, to our knowledge, the existing literature, where the cost random
variable is assumed to be bounded ([DE97, Proposition 1.4.2]), preventing it from being
applicable to a large class of problems in the field of stochastic control theory.

(ii) The characterization of changes of density with finite relative entropy in the case where the
underlying probability distribution is Wiener measure. The characterization is in terms of
the square integrability of the corresponding change of drift. The relation between square
integrability and finite relative entropy does appear in the literature (see [BD98, DE97, Sec-
tion 4.6.4]), but the implication that finite relative entropy implies finite square integrability
is, to our knowledge, novel.

(iii) The extension of the results of [BD98] to the important case (from applied point of view)
where the cost random variable is not bounded from above, along with the explicit consid-
eration of the existence of a minimizing control process, and of its properties.

(iv) The use of Malliavin calculus to compute the optimal control process. Malliavin calculus is
sometimes used in optimization problems (see e.g. [OK91, NOkP10]) but not before in the
context of relative entropy weighted control, even though it appears naturally.

(v) The solution of the problem where the cost random variable is the maximum of a standard
Brownian motion with controlled drift over a finite time horizon, as an application of (iv).

1.1 Outline

In Section 2, we consider the general relative entropy weighted optimization problem, and com-
pletely characterize the different situations that may arise. It turns out that the useful case is
the case of finite relative entropy, because only then there actually exists a probability measure
that minimizes the cost functional. Problems in which the optimal change of measure has finite
relative entropy are easily characterized in terms of conditions on the cost functional and the
probability measure. Then in Section 3 this case is further investigated within the context of a
Wiener process. It is shown that a change of measure with finite relative entropy corresponds to a
square integrable drift process, which in particular is the case for the optimal density under certain
assumptions. In Section 4 we show how the optimal control process may be computed in terms of
the maliavin derivative of the cost functional. To illustrate the usefulness of this approach, and
as an interesting result in its own right, we compute the optimal drift for the case where the cost
functional is the maximum of a one dimensional Wiener process with controlled drift, on a finite
time horizon in Section 5. Preliminary results on relative entropy, as well as some proofs of results
that are slightly more involved, are delegated to the Appendix.

1.2 Notation

As is common in probability theory, we will allow expectations of nonnegative random variables
to assume their values within the extended reals $[-\infty, \infty]$. Also, where $\log 0 = -\infty$, $\log \infty = \infty$,
$\exp(-\infty) = 0$ and $\exp(\infty) = \infty$.

2
The euclidean norm of $x \in \mathbb{R}^d$ is denoted by $|x|$. For an adapted process $\theta$ and a continuous local martingale $M$, both with values in $\mathbb{R}^d$, we write $\int_0^t (\theta_s, dM_s)$ to indicate $\sum_{i=1}^d \int_0^t \theta^i_s \, dM^i_s$.

If $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space we write $\mathbb{E}^\mathbb{P}$ for expectation with respect to the probability measure $\mathbb{P}$.

Lebesgue measure will be denoted by $\text{Leb}$.

Whenever $M$ is a continuous local martingale, we write $\mathcal{E}(M)$ for the exponential local martingale $\mathcal{E}(M)_t = \exp(M_t - \frac{1}{2}[M]_t)$.

## 2 Relative entropy weighted optimization

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. The probability measure $\mathbb{P}$ will be referred to as the *uncontrolled (probability) measure*. Let $C$ be a random variable in $[-\infty, \infty]$. The random variable $C$ indicates a cost we wish to minimize, as explained below.

Let $\mathcal{P}$ denote the set of probability measures that are measurable with respect to $\mathcal{F}$. We wish to find a probability measure $\mathbb{Q} \in \mathcal{P}$ such that

(i) is absolutely continuous with respect to $\mathbb{P}$ (denoted by $\mathbb{Q} \ll \mathbb{P}$),

(ii) reduces the expected cost $\mathbb{E}^\mathbb{Q} C$, but

(iii) has small deviation from $\mathbb{P}$. We take the relative entropy

$$\mathcal{H}(\mathbb{Q}; \mathbb{P}) = \int_\Omega \log \left( \frac{d\mathbb{Q}}{d\mathbb{P}} \right) \, d\mathbb{Q} = \mathbb{E}^\mathbb{Q} \left[ \log \left( \frac{d\mathbb{Q}}{d\mathbb{P}} \right) \right]$$

as a measure of this deviation (see Appendix A.1). Recall $\mathcal{H}(\mathbb{Q}; \mathbb{P}) \geq 0$ for any $\mathbb{Q}, \mathbb{P} \in \mathcal{P}$, and $\mathcal{H}(\mathbb{Q}; \mathbb{P}) = 0$ if and only if $\mathbb{Q} = \mathbb{P}$.

Note that (i) is a constraint and (ii) and (iii) are conflicting optimization targets.

Let $\mathcal{P}_A := \{ \mathbb{Q} \in \mathcal{P} : \mathbb{E}^\mathbb{P}[C] < \infty \}$ denote the set of admissible probability measures, and note that $\mathcal{P}_A$ is convex. Furthermore let $\mathcal{P}_0 := \{ \mathbb{Q} \in \mathcal{P}_A : \mathcal{H}(\mathbb{Q}; \mathbb{P}) < \infty \}$. Define the cost functional $J : \mathcal{P} \to \mathbb{R}$ by

$$J(\mathbb{Q}) := \begin{cases} \mathbb{E}^\mathbb{Q} C + \mathcal{H}(\mathbb{Q}; \mathbb{P}) = \mathbb{E}^\mathbb{Q} \left[ C + \log \left( \frac{d\mathbb{Q}}{d\mathbb{P}} \right) \right] & \text{if } \mathbb{Q} \in \mathcal{P}_A, \\ \infty & \text{otherwise.} \end{cases} \quad (1)$$

We arrive at the following problem:

**Problem 2.1** (Relative entropy weighted optimization). Compute $J^* = \inf_{\mathbb{Q} \in \mathcal{P}} J(\mathbb{Q})$, and if it exists, a minimizer $\mathbb{Q}^* \in \mathcal{P}_A$ such that $J(\mathbb{Q}^*) = J^*$.

The solution of this problem is well known for the case in which $\mathbb{P}(|C| < K) = 1$ for some $K > 0$, see e.g. [DE97] Proposition 1.4.2 or [BD98] Proposition 2.5. This case is too restrictive for practical purposes. The purpose of this section is to provide a complete characterization of the existence of solutions of Problem 2.1 in terms of conditions on $\mathbb{P}$ and $C$. To our knowledge this characterization has not appeared in the mathematical literature in this form.

To achieve this goal, we will consider the following further conditions on $C$ and $\mathbb{P}$.

**finite (relative) entropy:** $\mathbb{P}(C < \infty) > 0$ and $\mathbb{E}^\mathbb{P}[\exp(-C)|C|] < \infty$. (FE)

**integrability:** $0 < \mathbb{E}^\mathbb{P}[\exp(-C)] < \infty$. (I)

Note that

- (FE) implies (I) (since $\exp(-x) \mathbb{1}_{\{x \leq -1\}} \leq \exp(-x)|x|$, and $\exp(-x) \mathbb{1}_{\{x > -1\}} \leq \frac{1}{e}$);
- (I) implies $\mathbb{P}(C = -\infty) = 0$. 

3
If \( \mathcal{H} \) holds, then

\[
\frac{dQ^*}{dP} = Z^* := \frac{\exp(-C)}{\mathbb{E}[\exp(-C)]}.
\]  

(2)

defines a probability measure \( Q^* \) that is absolutely continuous with respect to \( P \).

**Lemma 2.2.** Suppose \( \mathcal{H} \) holds, and let \( Q^* \) and \( Z^* \) as defined by \( (2) \). Then

(i) For any \( Q \in \mathcal{P}_A \), we have that \( Q \ll Q^* \), and

\[
J(Q) = \mathcal{H}(Q; Q^*) - \log \mathbb{E}[\exp(-C)],
\]

(so that in particular \( J(Q) = \infty \) if \( \mathcal{H}(Q; Q^*) = \infty \));

(ii) \( J \) is strictly convex over \( \{ Q \in \mathcal{P}_A : \mathcal{H}(Q; Q^*) < \infty \} \) and is minimized at \( Q^* \), provided \( Q^* \in \mathcal{P}_A \).

**Proof.** (i) If \( Q \) is not absolutely continuous with respect to \( Q^* \), then there exists a set \( E \) of \( Q^* \)-measure zero, for which \( Q(E) > 0 \). We have \( Z^* = 0 \) on \( E \), so that \( C = \infty \) on \( E \). Therefore \( \mathbb{E}[Q|C] \geq \mathbb{E}[Q|E] = \infty \). So \( Q \notin \mathcal{P}_A \).

Now let \( Q \in \mathcal{P}_A \). We have just seen that \( Q \ll Q^* \), say with density \( Y = \frac{dQ}{dQ^*} \). We may choose a version of \( Y \) such that \( Y = 1 \) on \( \{ Z^* = 0 \} \). Then \( Q \ll P \) with density \( Z = \frac{dQ}{dP} = Y Z^* \). Write \( K = \mathbb{E}[\exp(-C)] \). Then

\[
\mathcal{H}(Q; Q^*) = \int_{\Omega} Z \log Y \, dP = \int_{\Omega \cap \{Z^* > 0\}} Z \log Y \, dP
\]

\[
= \int_{\Omega \cap \{Z^* > 0\}} Z (\log Z - \log Z^*) \, dP = \mathcal{H}(Q; P) - \int_{\Omega \cap \{Z^* > 0\}} Z \log Z^* \, dP
\]

\[
= \mathcal{H}(Q; P) + \int_{\Omega \cap \{Z^* > 0\}} Z (\log K + C) \, dP = \mathcal{H}(Q; P) + \log K + \mathbb{E}[^ZC].
\]

The term on the lefthand side, as well as the terms on the righthand side, are welldefined, with only the relative entropies possibly equal to \( +\infty \). In particular

\[
J(Q) = \mathbb{E}[ZC] + \mathcal{H}(Q; P) = \mathcal{H}(Q; Q^*) - \log K.
\]

(ii) This is an immediate consequence of Proposition \( \mathcal{A}.4 \) (ii), (iii).

The following proposition characterizes the situation in which a minimizer exists.

**Proposition 2.3.** Suppose Hypothesis \( \mathcal{H} \) holds. The following are equivalent.

(i) Hypothesis \( \mathcal{FE} \) holds;

(ii) \( \mathcal{P}_0 \) is non-empty, i.e. there exists an element \( Q \in \mathcal{P}_A \) such that \( \mathcal{H}(Q; P) < \infty \);

(iii) \( Q^* \), given by \( (2) \), is a minimizer for Problem 2.1 and \( J(Q^*) = -\log \mathbb{E}[\exp(-C)] \).

If any (hence all) of these cases hold, then the minimizer as mentioned under (iii) is unique.

**Proof.** (i) \( \Rightarrow \) (ii): For \( Q^* \) as defined by \( (2) \), \( Q^* \in \mathcal{P}_0 \).

(ii) \( \Rightarrow \) (iii): Let \( Q \in \mathcal{P}_0 \). In particular, \( J(Q) < \infty \). For \( Q^* \) as defined by \( (2) \), we have \( J(Q^*) \leq J(Q) < \infty \), by \( (3) \) and the non-negativity of relative entropy (Proposition \( \mathcal{A}.4 \)).

(iii) \( \Rightarrow \) (i): Since \( J(Q^*) < \infty \), in particular \( \mathcal{H}(Q^*; P) < \infty \). By Lemma \( \mathcal{A}.3 \) \( \mathbb{E}[\exp(-C)|C] < \infty \).

The stated uniqueness is an immediate consequence of Proposition \( \mathcal{A}.4 \).
The following proposition holds without restrictions on $\mathbb{P}$ and $C$. Since we are primarily interested in the case where Hypothesis $\text{FE}$ holds (so as to have the existence result of Proposition 2.3), the proof is delegated to the appendix. To our knowledge, Proposition 2.4 does not appear in the literature in this generality.

**Proposition 2.4.** We have

$$\inf_{Q \in \mathcal{P}} J(Q) = \inf_{Q \in \mathcal{P}} J(Q) = -\log \mathbb{E}^\mathbb{P} \exp(-C).$$

(4)

To conclude this section, we summarize by distinguishing the following cases, as immediate consequences of Proposition 2.3 and Proposition 2.4.

**Corollary 2.5.**

(i) If Hypothesis $\text{FE}$ holds, then there exists a minimizer for Problem 2.1;

(ii) If Hypothesis $\text{I}$ holds, but Hypothesis $\text{FE}$ does not hold, then Problem 2.1 has optimal value $-\infty < J^* < \infty$, but there does not exist a probability measure $Q \in \mathcal{P}_A$ at which this value is attained;

(iii) If Hypothesis $\text{I}$ does not hold, then Problem 2.1 has optimal value $J^* = \pm \infty$, with $J^* = +\infty$ if and only if $\mathbb{P}(C = \infty) = 1$.

2.1 Notes and remarks

**Remark 2.6.** One may wish to include a factor $\beta > 0$ in the problem formulation, to indicate the relative importance of minimizing $\mathbb{E}^Q C$ compared to minimizing $H(Q; \mathbb{P})$ to obtain the form $J(Q) = \mathbb{E}^Q C + \frac{1}{\beta}H(Q; \mathbb{P})$. In this case the optimal value becomes $J^* = -\frac{1}{\beta} \log \mathbb{E}^\mathbb{P} \exp(-\beta C)$, so that $\beta$ admits the interpretation of inverse temperature, in the context of statistical physics.

**Remark 2.7.** A sufficient condition for Hypothesis $\text{FE}$ to hold is that for some $\gamma > 1,\quad 0 < \mathbb{E}^\mathbb{P} \exp(-\gamma C) < \infty.$

(5)

Indeed, if this is the case, then, since $x \leq \frac{1}{\epsilon e} \exp(\epsilon x)$ for $x \geq 0$,

$$\mathbb{E}^\mathbb{P} \left[ \exp(-C)(-C)1_{C \leq 0} \right] \leq \frac{1}{(\gamma - 1)e} \mathbb{E}^\mathbb{P} \left[ \exp(-\gamma C)1_{C \leq 0} \right] < \infty.$$  

(And, as always, $\mathbb{E}^\mathbb{P} \left[ \exp(-C)|C|1_{C > 0} \right] \leq \frac{1}{\epsilon e}$. In turn $\text{I}$, and therefore $\text{FE}$, are implied by the conditions that $\mathbb{P}(C > K) = 1$ for some $K \in \mathbb{R}$, and $\mathbb{P}(C = \infty) < 1$.)

3 Relative entropy weighted optimization with respect to Wiener measure

In this section, we consider the important special situation where all randomness is generated by a $d$-dimensional Wiener process. Changes of measure (satisfying mild conditions) may in this case be expressed as a Girsanov type transformation. The stochastic process appearing in the exponent of the Girsanov density will constitute the ‘control process’. A crucial observation is that the relative entropy of such a transformation is given by the squared control costs (Proposition 3.3) (provided it is not infinite).

**Definition 3.1** (Canonical Wiener process). Let $\Omega = C([0,\infty); \mathbb{R}^d)$, i.e. the space of continuous functions mapping $[0,\infty)$ into $\mathbb{R}^d$, let $\mathcal{F}_t = \sigma(\{\omega(s) : 0 \leq s \leq t\})$, let $\mathbb{P}$ denote Wiener measure, let $\mathcal{F}_t$ be the right continuous completion of $\mathcal{F}_t$ with respect to $\mathbb{P}$ (see [RW94a, Section II.67]), and let $\mathcal{F} = \sigma(\cup_{t \geq 0} \mathcal{F}_t)$. Let $W_t(\omega) := \omega(t)$ for $\omega \in \Omega$ and $t \geq 0$, so that $W$ is a standard Brownian motion in $\mathbb{R}^d$. The collection $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P}, W)$ will be referred to as a canonical $d$-dimensional Wiener process.
Remark 3.2. We will often work with changes of density $Z = \frac{dQ}{dP}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ such that $Q$ and $P$ are equivalent probability measures. This is equivalent to the condition that $Z > 0$, $\mathbb{P}$-almost surely. Since $\mathcal{F}$ is assumed to be a complete filtration for $\mathbb{P}$, it is complete for $Q$. The fact that $(\mathcal{F}_t)_{t \geq 0}$ is right continuous is not affected by the change of measure, so for equivalent changes of measure the usual conditions remain to be satisfied. We will always work with continuous versions of the density process, that exist by virtue of the martingale representation theorem ([RW94b, Theorem IV.36.5]).

Let $\mathcal{U}$ denote the set of $\mathbb{R}^d$-valued progressively measurable stochastic processes $U$ such that the process $(Z^U_t)$ defined by
\begin{equation}
Z^U_t := \exp \left( \int_0^t \langle U_s, dW_s \rangle - \frac{1}{2} \int_0^t |U_s|^2 ds \right), \quad t \geq 0,
\end{equation}
is a martingale.

The set $\mathcal{U}$ will be called the *admissible set of controls* and $U \in \mathcal{U}$ will be called an *admissible control process*.

By Girsanov’s theorem [RW94b, Theorem IV.38.9], there exists a probability measure $Q^U$, such that for all $t \geq 0$,
\begin{equation}
\frac{dQ^U}{d\mathbb{P}} \bigg|_{\mathcal{F}_t} = Z^U_t,
\end{equation}
with respect to which the process
\begin{equation}
t \mapsto W^U_t := W_t - \int_0^t U_s \, ds, \quad t \geq 0,
\end{equation}
is a standard Brownian motion. Let $\mathbb{E}^U$ be a shorthand notation for $\mathbb{E}^{Q^U}$. Define
\begin{equation}
\mathcal{U}^2 := \left\{ U \in \mathcal{U} : \mathbb{E}^Q \int_0^\infty |U_t|^2 \, dt < \infty \right\}.
\end{equation}

The relative entropy of this change of measure is proportional to the expectation of the $L^2$-norm of $U$ as is shown in the following theorem. The direction (ii) $\Rightarrow$ (i) is observed in e.g. [DE97, Section 4.6.4] for the finite horizon case, but we have not encountered a proof of the less immediate implication (i) $\Rightarrow$ (ii) in the literature. The proof of this result may be found in the appendix.

**Proposition 3.3.** Suppose $Z^U > 0$, $\mathbb{P}$-almost surely. The following are equivalent.

(i) $\mathcal{H}(Q^U; \mathbb{P}) < \infty$;

(ii) $U \in \mathcal{U}^2$.

In any (hence both) of these cases, $\mathcal{H}(Q^U; \mathbb{P}) = \frac{1}{2} \mathbb{E}^Q \left[ \int_0^\infty |U_t|^2 \, dt \right]$. 

**Proof.** This is Proposition A.9. $\square$

Suppose again $C$ is a $\mathcal{F}$-measurable random variable. We will make use of the following hypothesis:

\begin{equation}
\text{finiteness: } \mathbb{P}(C < \infty) = 1. \quad (F)
\end{equation}

Define the *cost functional* $J(U)$ by
\begin{equation}
J(U) := \mathbb{E}^U C + \mathcal{H}(Q^U; \mathbb{P})
\end{equation}

Consider the following problem.

**Problem 3.4** (Dynamic relative entropy weighted optimization). Find the *optimal value* $J^\star$ defined by
\begin{equation}
J^\star := \inf_{U \in \mathcal{U}} J(U),
\end{equation}
and, provided it exists, a minimizer $U^\star = \arg \min_{U \in \mathcal{U}} J(U)$. 


Note the similarity to Problem 2.1. The main difference between the two problems is that in Problem 3.4 we restrict the possible probability measures to those parametrized by $U \in \mathcal{U}$, through their density given by (7). We have seen in Proposition 2.4 of Section 2 that if we would not impose this restriction, the optimal value would be given by (4). Therefore we immediately have

$$J^* \geq -\log \mathbb{E}^\mathbb{P}[\exp(-C)].$$

In fact, we have the following result.

**Theorem 3.5.** Suppose Hypotheses $\mathcal{F}$ and $\mathcal{FE}$ hold. Then Problem 3.4 admits a minimizer $U^* \in \mathcal{U}^2$, and $J(U^*) = -\log \mathbb{E}^\mathbb{P}[\exp(-C)]$. Furthermore $U^*$ is unique up to modification on $\mathbb{P} \otimes \text{Leb}$-null sets.

**Proof.** Let $Z$ be as in Proposition 2.3 (omitting $\ast$ in $Z^\ast$). The density process $(Z_t)_{t \geq 0}$, defined by

$$Z_t := \mathbb{E}^\mathbb{P}[Z \mid \mathcal{F}_t]$$

is a uniformly integrable martingale. By the martingale representation theorem (see [RW94], Theorem IV.36.5]), it has a continuous version, that we will work with. By the assumption that $C \leq \infty$, we have $Z > 0$, $\mathbb{P}$-almost surely and therefore $Z_t > 0$ for all $t \geq 0$, $\mathbb{P}$-almost surely, by Lemma A.8.

The process $M_t$ satisfying the SDE

$$dM_t = \frac{1}{Z_t} dZ_t, \quad t \geq 0, \quad \mathbb{P}$$

and $M_0 = 0$ is then a continuous local martingale, and $Z_t = \mathcal{E}(M)_t$ for $t \geq 0$. By Proposition A.9 $M$ is bounded in $\mathcal{L}^2(\mathbb{P})$. Let $M_\infty = \lim_{t \to \infty} M_t$. Again by the martingale representation theorem,

$$M_t = \int_0^t (U_s, dW_s),$$

and $U \in \mathcal{U}^2$. The uniqueness of $U$ is a consequence of the $\mathbb{P}$-almost sure uniqueness of $M$ (see [Kal02] Lemma 18.21]) and $\mathbb{P} \otimes$ Leb-uniqueness in the martingale representation theorem.

**Example 3.6.** Consider the case where $C := \inf\{t : |W_t| > 1\}$, where $W$ is a $d$-dimensional Wiener process. We have $\mathbb{P}(C = \infty) = 0$, and $C \geq 0$, so that Hypothesis $\mathcal{FE}$ is satisfied. Theorem 3.5 yields that the optimal control process $U$ is in $\mathcal{U}^2$. The optimal control can in this case be computed explicitly through the Hamilton-Jacobi-Bellman equation, e.g. $U_t = \sqrt{2 \tanh(\sqrt{2}W_t)}1_{\{|W_t|<1\}}$ in case $d = 1$.

Here we provide a formula that reduces the computation of $U$ to differentiation of a conditional expectation. This result applies to all $U \in \mathcal{U}^2$, and in particular to the solution $U^\ast$ of Problem 3.4 as provided by Theorem 3.5 with $Z^\ast = \exp(-C)/\mathbb{E}^\mathbb{P}\exp(-C)$.

**Proposition 3.7.** Suppose $U \in \mathcal{U}^2$, $Z^U > 0$, and for all $t \geq 0$, that $r \to \mathbb{E}^\mathbb{P}[W_r Z^U \mid \mathcal{F}_t]$, where $r \geq t$, is locally absolutely continuous with respect to Lebesgue measure, $\mathbb{P}$-almost surely. Then

$$U_t = \frac{d}{dr} \mathbb{E}^\mathbb{P}[W_r Z^U \mid \mathcal{F}_t]_{|r=t}, \quad \mathbb{P}$$

**Proof.** In the proof we fix $U \in \mathcal{U}$ and omit the superscripts $U$ in $Z^U$ etc. Let $t \geq 0$ and $r > t$. The process $(Z_s)$ is the stochastic exponential of $U_s$, so $dZ_s = Z_s \langle U_s, dW_s \rangle$. Then using Itô’s formula, for $0 \leq s \leq r$,

$$d(W_r Z_s) = Z_s(1 + W_s U_s) dW_s + U_s Z_s ds,$$

so that, for $t \leq s \leq r$, exchanging the integral over time and the conditional expectation,

$$\mathbb{E}^\mathbb{P}[Z W_r \mid \mathcal{F}_t] = \mathbb{E}^\mathbb{P}[Z W_r \mid \mathcal{F}_t] = Z_t W_t + \int_t^r \mathbb{E}^\mathbb{P}[U_s Z_s \mid \mathcal{F}_t] ds.$$
Taking the derivative of relation (12) with respect to $r$ and evaluating at $r = t$, gives
\[ \frac{d}{dr} E^p [ZW_r | \mathcal{F}_t] |_{r=t} = U_t Z_t. \]

3.1 Notes and remarks

Remark 3.8. In [BD98], it is proven that for $C$ bounded from above, and $\mathcal{F}_1$-measurable (i.e. in the finite time horizon case), we have
\[ \inf_{U \in \mathcal{U}} J(U) = -\log E^p \exp(-C). \]

This result is sufficient for application within the field of large deviations theory. Our Theorem 3.5 is perhaps more convenient for applications in control theory since (i) it is not restricted to finite time horizon problems; (ii) it establishes explicitly the existence of a minimizer; (iii) it does not impose the restriction on the cost functional to be bounded from above.

Remark 3.9. Proposition 3.7 provides a method to compute the optimal control process through Monte Carlo sampling, by employing the law of large numbers to approximate the (conditional) expectation.

Remark 3.10. Our set up, using Wiener measure on the time interval $[0, \infty)$ allows for control over an infinite time horizon. At the same time, Proposition 3.3 shows that conditions (FE) and (F) exclude e.g. the case where the control process is $\mathbb{Q}$-almost surely equal to a non-zero constant on an infinite time horizon (since then $U \notin \mathcal{U}^t$). However, conditions (FE) and (F) allow us to consider problems whose cost function is $\mathcal{F}_\tau$-measurable, with $\tau$ a stopping time that is unbounded but almost surely finite, as illustrated by Example 3.6.

Remark 3.11. It should be noted that finite time horizon problems fit well within the theory of this section. In this case one just considers a cost random variable $C$ that is $\mathcal{F}_T$-measurable for some deterministic $T > 0$. The optimal control process $U^*$ will be $\mathcal{F}_T$-measurable, by its construction.

4 Computation of the optimal control through Malliavin calculus

In this section we will apply the Clark-Ocone theorem of Malliavin calculus to obtain an explicit representation of the optimal control process in terms of the Malliavin derivative of the cost random variable. The results of this section have perhaps no important theoretical novelty (since the proofs are elementary and based upon well established results of Malliavin calculus) but are useful and should be noted within the context of this paper. Section 5 provides an illustration of the use of Corollary 4.4.

Throughout this section, let $(\Omega, (\mathcal{F}_t), \mathcal{F}, \mathbb{P}, W)$ denote a canonical $d$-dimensional Wiener process. Recall the following definitions and notations (see [Niu06]).

Let $H$ denote the Hilbert space $L^2([0, \infty); \mathbb{R}^d)$. With some abuse of notation, let $W$ also denote the mapping $W : H \rightarrow L^2(\Omega)$ given by $W(h) = \int_0^\infty \langle h(t), dW_t \rangle$. Let $C^\infty_p(\mathbb{R}^n)$ denote the set of all infinitely continuously differentiable functions $f : \mathbb{R}^n \rightarrow \mathbb{R}$ such that $f$ and all of its partial derivatives have polynomial growth. Let $\mathcal{S}$ denote the set of random variables $F$ of the form
\[ F = f(W(h_1), \ldots, W(h_n)), \]
where $f \in C^\infty_p(\mathbb{R}^n)$ and $h_1, \ldots, h_n \in H$ for some $n \in \mathbb{N}$. The Malliavin derivative of $F \in \mathcal{S}$ is the $H$-valued random variable given by
\[ DF = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(W(h_1), \ldots, W(h_n)) h_i. \]
The operator $D$ is closable as a linear operator $D : L^p((\Omega, \mathbb{P})) \to L^p((\Omega, \mathbb{P}); H)$ for any $p \geq 1$ (see [Nua06, Proposition 1.2.1]), and the domain of its closure is denoted by $D^{1,p}$, i.e. the closure of $S$ with respect to the norm

$$||F||_{1,p} = (\mathbb{E}^p|F|^p + \mathbb{E}^p||DF||_H^p)^{\frac{1}{p}}.$$  

If $F \in D^{1,2}$, so that $DF \in L^2(\Omega; H) = L^2(\Omega; L^2([0, \infty))) \cong L^2(\Omega \times [0, \infty))$, then we may identify $DF$ with a stochastic process, denoted $D_tF$, as usual:

$$(D_tF)(\omega) = DF(\omega, t), \quad \omega \in \Omega, \quad t \geq 0.$$  

Note that $(D_tF)_{t \geq 0}$ is not necessarily adapted, and the value of $D_tF(\omega)$ is defined $\mathbb{P} \otimes \text{Leb}$-almost everywhere.

**Lemma 4.1.** Suppose $F \in D^{1,2}$. Then $\mathbb{E}^p[F \mid \mathcal{F}_t] \in D^{1,2}$ for $t \geq 0$ and $D_t\mathbb{E}^q[F \mid \mathcal{F}_t] = \mathbb{E}^p[D_tF \mid \mathcal{F}_t]$.

*Proof.* This is an immediate consequence of [Nua06 Proposition 1.2.8].

**Lemma 4.2.** Suppose $Q$ is equivalent to $\mathbb{P}$ with Radon-Nikodým derivative $Z = \frac{dQ}{d\mathbb{P}} \in D^{1,2}$. Let $Z_t := \mathbb{E}^Q[Z \mid \mathcal{F}_t], t \geq 0$, denote the density process. Then $Z_t > 0$ and $Z_t \in D^{1,2}$ for all $t \geq 0$.

Define the $\mathbb{R}^d$-valued adapted process $V$ by

$$V_t := \frac{D_tZ_t}{Z_t}, \quad t \geq 0.$$  

(13)  

Then $Z_t = \exp \left(\int_0^t \langle V_s, dW_s \rangle - \frac{1}{2} \int_0^t |V_s|^2 \, ds\right)$, for $t \geq 0$.

*Proof.* Note that

$$D_sZ_s = D_s\mathbb{E}^Q[Z_t \mid \mathcal{F}_s] = \mathbb{E}^Q[D_sZ_t \mid \mathcal{F}_s] \quad \text{for} \quad 0 \leq s \leq t,$$

where the second equality is a consequence of Lemma 4.1. By the Clark-Ocone representation formula [Nua06 Proposition 1.3.14], therefore

$$Z_t = 1 + \int_0^t \langle D_sZ_s, dW_s \rangle = 1 + \int_0^t \langle D_sZ_s, dW_s \rangle = 1 + \int_0^t \langle V_sZ_s, dW_s \rangle$$

for $V_s = \frac{D_sZ_s}{Z_s}$.  

**Proposition 4.3.** Suppose $\exp(-C) \in D^{1,2}$ and Hypotheses FE and F hold. Then the stochastic process $(V_t)_{t \geq 0}$ defined by

$$V_t := \frac{\mathbb{E}^p[D_t\exp(-C) \mid \mathcal{F}_t]}{\mathbb{E}^p[\exp(-C) \mid \mathcal{F}_t]}, \quad t \geq 0,$$

(14)

is equal to the solution of Problem 3.1 (up to modification on a $\mathbb{P} \otimes \text{Leb}$-null set).

*Proof.* By Theorem 3.5 Problem 3.1 is solved by a density function $Z = \exp(-C)/\mathbb{E}^p \exp(-C)$ satisfying

$$Z_t = \exp \left(\int_0^t \langle U_s, dW_s \rangle - \frac{1}{2} \int_0^t |U_s|^2 \, ds\right), \quad t \geq 0,$$

for some $\mathbb{P} \otimes \text{Leb}$-almost everywhere unique process $U$. By Lemma 4.1 for $V$ by (14),

$$\frac{D_tZ_t}{Z_t} = \frac{\mathbb{E}^p[D_t\exp(-C) \mid \mathcal{F}_t]}{\mathbb{E}^p[\exp(-C) \mid \mathcal{F}_t]} = V_t.$$  

Therefore, applying Lemma 4.2 $Z_t = \exp \left(\int_0^t \langle V_s, dW_s \rangle - \frac{1}{2} \int_0^t |V_s|^2 \, ds\right)$, with $V$ as above. The result now follows from the stated uniqueness.
Corollary 4.4. Suppose \( P(0 \leq C < \infty) = 1 \), and \( C \in \mathbb{D}^{1,2} \). Then the stochastic process \((V_t)_{t \geq 0}\) defined by
\[
V_t := -\frac{\mathbb{E}^P[\exp(-C)D_tC \mid \mathcal{F}_t]}{\mathbb{E}^P[\exp(-C) \mid \mathcal{F}_t]}, \quad t \geq 0,
\]
is equal to the solution of Problem 3.4 (up to modification on a \( P \otimes \text{Leb} \)-null set).

Proof. Note that Hypothesis \( (\text{FE}) \) is satisfied. Also note that \( x \mapsto \exp(-x) \) has bounded derivative for \( x \) bounded from below. Therefore we may apply the chain rule of Malliavin calculus (see [Nua06, Proposition 1.2.3]), to conclude that \( \exp(-C) \in \mathbb{D}^{1,2} \) with \( D_t \exp(-C) = -\exp(-C) D_tC \). The result now follows from Proposition 4.3.

4.1 Notes and remarks

Remark 4.5. Expression (13) of Lemma 4.2 should be compared to Proposition 3.7, which states (phrased within the context of this section)
\[
V_t = \frac{d}{dt}\mathbb{E}^P[W_rZ \mid \mathcal{F}_t] \bigg|_{r=t}, \quad \mathbb{P}\text{-a.s. for } t \geq 0 \quad (15)
\]
As mentioned in Remark 3.9, Proposition 3.7 is perhaps best suited for Monte Carlo-sampling, whereas the results of this section may be applied for exact computations, as in Section 5. Of course, future research may find different uses for either of the two expression.

Remark 4.6. An alternative method for finding an expression for the optimal control process in the case of the control of diffusion processes is through the dynamic programming principle, i.e. the Hamilton-Jacobi-Bellman PDE. For the class of problems considered in this paper, the HJB may be transformed into a linear PDE through the Hopf- or logarithmic transform [FS09, Kap05].

This approach is however restricted to the case where \( C \) is of the form \( C = \int_0^T \varphi(s,X_s) \, ds + \psi(X_T) \) for some measurable functions \( \varphi : [0,T] \times \mathbb{R}^n \to \mathbb{R} \) and \( \psi : \mathbb{R}^n \to \mathbb{R} \) (along with possible variants). A detailed discussion of this approach requires the notion of viscosity solutions and is beyond the aim of this paper.

Remark 4.7. In [OK91], the Clark-Ocone formula is used in a different optimization context, namely for portfolio optimization in mathematical finance.

5 Minimization of the maximum of a Wiener process with drift over a finite time horizon

In this section we illustrate the theory by obtaining a new result on the minimization of the maximum of a Wiener process with drift over a finite time horizon. As it turns out, the value of optimal drift can be explicitly computed as a function of the difference between the running maximum and the current value of the Wiener process. This solution is obtained by applying the results of Section 4, illustrating the usefulness of the Malliavin calculus approach.

Let \((\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{F}, \mathbb{P}, W)\) denote a one-dimensional canonical Wiener process. Define \( M_t := \max_{0 \leq s \leq t} W_s \) and take \( C := M_T \) for some \( T > 0 \). Hypothesis \( (\text{FE}) \) is satisfied by Remark 2.7 and the observation that \( \mathbb{P}(M_T = \infty) = 0 \). For the distribution of \( M_t \) we have by virtue of the reflection principle [KS91, Section 2.8.A]
\[
\mathbb{P} (M_t \geq a) = \left( \frac{2}{\pi} \right)^{1/2} \int_0^a \frac{\exp(-\xi^2/2)}{\sqrt{\pi}} \, d\xi, \quad t \geq 0, a \geq 0.
\]

We will make use of the error function (erf) and complimentary error function (erfc), defined by
\[
erf(x) := \frac{2}{\sqrt{\pi}} \int_0^x \exp(-\eta^2) \, d\eta, \quad \text{erfc}(x) := 1 - \text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty \exp(-\eta^2) \, d\eta, \quad x \geq 0.
\]
We wish to compute the optimal control process corresponding to the minimization of $C$, using Corollary 4.4. This means we need to compute $\mathbb{E}^\mathbb{P}[\exp(-M_T)D_tM_T \mid \mathcal{F}_t]$ and $\mathbb{E}^\mathbb{P}[\exp(-M_T) \mid \mathcal{F}_t]$. We start with the latter. Conditional on $\mathcal{F}_t$, the event $M_T = M_t$ occurs when the maximum over $[t, T]$ does not exceed $y := M_t$. This has the same probability as the event that the maximum over $[0, T-t]$ does not exceed $y - W_t$, so

$$P(M_T = M_t \mid W_t = x, M_t = y) = P(M_{T-t} \leq y - x) = \left(\frac{2}{\pi}\right)^{1/2} \int_0^{\frac{M_t-W_t}{\sqrt{2(T-t)}}} \exp(-\xi^2/2) \, d\xi.$$  

For $0 \leq x \leq y < z$ we compute

$$P(M_T \geq z \mid W_t = x, M_t = y) = P(M_T = M_t W_t = x, M_t = y) + P(M_{T-t} \geq z - x)$$

$$= \left(\frac{2}{\pi}\right)^{1/2} \int_0^{\frac{M_t-W_t}{\sqrt{2(T-t)}}} \exp(-\xi^2/2) \, d\xi + \left(\frac{2}{\pi}\right)^{1/2} \int_{\frac{z-y}{\sqrt{2(T-t)}}}^\infty \exp(-\xi^2/2) \, d\xi.$$  

Therefore the density function of $M_T$ conditional on $\mathcal{F}_t$ is equal to

$$f_{M_T \mid \mathcal{F}_t}(\xi) = \left(\frac{2}{\pi(T-t)}\right)^{1/2} \exp\left(-\frac{(\xi - W_t)^2}{2(T-t)}\right), \quad \text{for } \xi > M_t \geq W_t.$$  

We compute

$$\mathbb{E}^\mathbb{P}[\exp(-M_T) \mid \mathcal{F}_t]$$

$$= \mathbb{E}^\mathbb{P}[\exp(-M_T)1_{M_T = M_t} \mid \mathcal{F}_t] + \mathbb{E}^\mathbb{P}[\exp(-M_T)1_{M_T > M_t} \mid \mathcal{F}_t]$$

$$= \exp(-M_t)P(M_T = M_t \mid \mathcal{F}_t) + \mathbb{E}^\mathbb{P}[\exp(-M_T)1_{M_T > M_t} \mid \mathcal{F}_t]$$

$$= \exp(-M_t)\left(\frac{2}{\pi}\right)^{1/2} \int_0^{\frac{M_t-W_t}{\sqrt{2(T-t)}}} \exp(-\xi^2/2) \, d\xi$$

$$+ \left(\frac{2}{\pi(T-t)}\right)^{1/2} \int_{\frac{M_t}{\sqrt{2(T-t)}}}^\infty \exp(-\xi) \exp\left(-\frac{(\xi - W_t)^2}{2(T-t)}\right) \, d\xi$$

$$= \exp(-M_t) \text{erf}\left(\frac{M_t - W_t}{\sqrt{2(T-t)}}\right) + \exp\left(-W_t + \frac{1}{2}(T-t)\right) \text{erfc}\left(\frac{M_t - W_t + (T-t)}{\sqrt{2(T-t)}}\right).$$  

An expression for the Malliavin derivative of $M_T$ is available: $D_t M_T = 1_{[0, \tau)}(t)$, where $\tau$ is the a.s. unique point where $W$ attains its maximum. See [Nua06, Exercise 1.2.11]. Note that $t < \tau$ if and only if $M_T > M_t$. Therefore

$$\phi_t := -\mathbb{E}^\mathbb{P}[\exp(-M_T)D_tM_T \mid \mathcal{F}_t] = -\mathbb{E}^\mathbb{P}[\exp(-M_T)1_{M_T > M_t} \mid \mathcal{F}_t]$$

$$= -\left(\frac{2}{\pi(T-t)}\right)^{1/2} \int_{M_t}^\infty \exp\left(-\xi - \frac{(\xi - W_t)^2}{2(T-t)}\right) \, d\xi$$

$$= -\exp\left(-W_t + \frac{1}{2}(T-t)\right) \text{erf}\left(\frac{M_t - W_t + (T-t)}{2(T-t)^{1/2}}\right).$$  

We finally compute

$$U^*_t := \mathbb{E}^\mathbb{P}[\exp(-M_T) \mid \mathcal{F}_t] = u(t, W_t, M_t),$$  

where, for $0 \leq t < T$, $w \in \mathbb{R}$, and $m \geq w$,

$$u(t, w, m) = -\exp\left(-w + \frac{1}{2}(T-t)\right) \text{erf}\left(\frac{m-w+T-t}{\sqrt{2(T-t)}}\right)$$

$$+ \exp(-m) \text{erfc}\left(\frac{m-w}{\sqrt{2(T-t)}}\right) + \exp\left(-w + \frac{1}{2}(T-t)\right) \text{erfc}\left(\frac{m-w+T-t}{\sqrt{2(T-t)}}\right).$$  

By Corollary 4.4, we have obtained the following result.
Consider the problem

\[
\text{minimize } J(U) := \mathbb{E}^U \left[ \max_{0 \leq t \leq T} \left( W_t^U + \int_0^t U_s \, ds \right) + \frac{1}{2} \int_0^T U_s^2 \right],
\]

with \( W^U \) a one dimensional standard Brownian motion under \( Q^U \), and where the minimum is over all adapted processes \((U_t)_{0 \leq t \leq T}\) satisfying \( \mathbb{E}^U \left[ \int_0^T U_s^2 \, ds \right] < \infty \). This problem is solved by \( U^* \) as given by (16), (17).

An illustration of this result is provided in Figure 1.

5.1 Notes and remarks

Remark 5.2. This example illustrates how the theory developed in this paper applies to non-Markovian processes, and therefore provides a method that applies where a dynamic programming (i.e. the HJB equation) cannot be used. Augmenting the state to \((W_t, M_t)\) would yield a Markov process, but solution through dynamic programming is then far from straightforward.

Remark 5.3. For a different optimization problem related to the maximum of a Wiener process see [HS91].

A Appendix

A.1 Relative entropy

In this section we define the notion of relative entropy and list some useful properties. Everything in this section is well established, see e.g. [DE97 Section 1.4].
In the following, \( Q \) and \( P \) denote probability measures over some measurable space \((\Omega, \mathcal{F})\). We write \( Q \ll P \) if \( Q \) is absolutely continuous with respect to \( P \). If \( Q \ll P \) we write \( Z = \frac{dQ}{dP} \) for the Radon-Nikodým derivative or density of \( Q \) with respect to \( P \), which is defined uniquely up to modification on \( P \)-null sets.

**Definition A.1.** The relative entropy (Kullback-Leibler divergence, information divergence) of \( Q \) with respect to \( P \) is defined as

\[
\mathcal{H}(Q;P) := \begin{cases} 
  \mathbb{E}^Q[\log Z], & \text{if } Q \ll P, \\
  \infty, & \text{otherwise.}
\end{cases}
\]  

(18)

**Remark A.2.** Relative entropy is defined without requiring that \( \log Z \in L^1(Q) \). This is not problematic since \( x \mapsto x \log x \) is bounded on \( x \leq 1 \). A slightly more careful definition of relative entropy would be

\[
\mathcal{H}(Q;P) := \begin{cases} 
  -\mathbb{E}^P[Z(-\log Z) \mathbb{1}_{\{Z \leq 1\}}] + \mathbb{E}^P[(Z \log Z) \mathbb{1}_{\{Z > 1\}}], & \text{if } Q \ll P, \\
  \infty, & \text{otherwise,}
\end{cases}
\]  

(19)

where both expectations are over nonnegative random variables, and the first expectation is finite.

From this observation we conclude:

**Lemma A.3.** The following are equivalent:

(i) \( \mathcal{H}(Q;P) < \infty \);

(ii) \( Q \ll P \) and \( Z \log Z \in L^1(P) \).

(iii) \( Q \ll P \) and \( \mathbb{E}^Q[(Z \log Z) \mathbb{1}_{\{Z > 1\}}] < \infty \).

The following proposition summarizes some useful properties of relative entropy. In particular, it indicates that relative entropy is a good indication of how similar two probability measures are.

**Proposition A.4.** Let \( P \) and \( Q \) be probability measures on \((\Omega, \mathcal{F})\). Then

(i) \( \mathcal{H}(Q;P) \in [0, \infty) \).

(ii) \( \mathcal{H}(Q;P) = 0 \) if and only if \( Q = P \) on \( \Omega \).

(iii) \( \mathcal{H}(\cdot;P) \) is strictly convex on \( \{ Q \ll P : \mathcal{H}(Q;P) < \infty \} \).

**Proof.** See [DE97, Section 1.4].

**A.2 Proof of Proposition 2.4**

Let \((\Omega, \mathcal{F}, P)\) be a probability space, and \( C \) be a \( \mathcal{F} \)-measurable random variable assuming its values within the extended reals. Write \( \mathcal{P} \) for the set of probability measures on \((\Omega, \mathcal{F})\), \( \mathcal{P}_A = \{ Q \in \mathcal{P} : \mathbb{E}^Q[C] < \infty \} \), and \( \mathcal{P}_0 = \{ Q \in \mathcal{P}_A : \mathcal{H}(Q;P) < \infty \} \). Let \( J : \mathcal{P} \to \mathbb{R} \) be given by \([1]\).

**Proposition A.5.** We have

\[
\inf_{Q \in \mathcal{P}} J(Q) = \inf_{Q \in \mathcal{P}_n} J(Q) = -\log \mathbb{E}^P \exp(-C).
\]  

(20)

The proof is similar to the proof of [BD98, Theorem 5.1] for the special case in which the underlying probability distribution is Wiener measure.

**Proof.** We distinguish three cases. Case (i): Suppose Hypothesis [1] holds. Define \( C_n := C \vee (-n) \). Since \( C_n \) is bounded from below, \( \mathbb{E}^P[\exp(-C_n)|C_n|] < \infty \). In particular, Proposition 2.3 may be applied to \( C_n \), to see that

\[
\inf_{Q \in \mathcal{P}_A(C_n)} \mathbb{E}^Q C_n + \mathcal{H}(Q;P) = -\log \mathbb{E}^P \exp(-C_n),
\]
where of course $\mathcal{P}_A(C_n) = \{Q \in \mathcal{P} : \mathbb{E}^Q|C_n| < \infty\}$. Note that $\mathbb{E}^P \exp(-C_n) \to \mathbb{E}^P \exp(-C)$ by dominated convergence (since $\exp(-C_n) \leq \exp(-C)$).

To establish the lower bound in (A.1), suppose $Q \in \mathcal{P}_A$. Since $\mathbb{E}^Q|C| < \infty$, also $\mathbb{E}^Q|C_n| < \infty$, so $Q \in \mathcal{P}_A(C_n)$. Let $N \in \mathbb{N}$ such that, for $n > N$, $\mathbb{E}^Q||C_n - C|| < \varepsilon$, which is possible by dominated convergence (since $|C_n - C| \leq |C|$). For any $n > N$,

\[
J(Q) \geq \mathbb{E}^Q[C_n] + \mathcal{H}(Q; \mathbb{P}) - \varepsilon \geq -\log \mathbb{E}^P(-C_n) - \varepsilon.
\]

Taking the limit of $n \to \infty$,

\[
J(Q) \geq \log \mathbb{E}^P(-C) - \varepsilon.
\]

Since $\varepsilon > 0$ is arbitrary, it follows that $J(Q) \geq -\log \mathbb{E}^P(-C)$ for any $Q \in \mathcal{P}_A$, and in particular $\inf_{Q \in \mathcal{P}_A} J(Q) \geq -\log \mathbb{E}^P(-C)$.

To establish the upper bound, suppose for $n \in \mathbb{N}$, that $Q_n \in \mathcal{P}_0$ is a minimizer of $Q \mapsto \mathbb{E}^Q C_n + \mathcal{H}(Q; \mathbb{P})$, so that $\mathbb{E}^Q_n C_n + \mathcal{H}(Q_n; \mathbb{P}) = -\log \mathbb{E}^P(-C_n)$. Now

\[
\inf_{Q \in \mathcal{P}_A} J(Q) \leq J(Q_n) = \mathbb{E}^Q_n C_n + \mathcal{H}(Q_n; \mathbb{P}) \leq \mathbb{E}^Q_n C_n + \mathcal{H}(Q_n; \mathbb{P}) = -\log \mathbb{E}^P(-C_n).
\]

As $n \to \infty$, we find

\[
\inf_{Q \in \mathcal{P}_A} J(Q) \leq -\log \mathbb{E}^P(-C).
\]

Case (ii): Suppose $\mathbb{E}^P(-C) = \infty$. Then $\lim_{n \to \infty} \mathbb{E}^P(-C_n) = \infty$, and analogous reasoning for the upper bound as before gives

\[
\inf_{Q \in \mathcal{P}_A} J(Q) \leq -\log \mathbb{E}^P(-C_n) \downarrow -\infty,
\]

as $n \to \infty$.

Case (iii): Suppose $\mathbb{E}^P(-C) = 0$. In this case, $\mathbb{P}(C = \infty) = 1$, so that $\mathcal{P}_0 = \emptyset$, and $\inf_{Q \in \mathcal{P}_0} J(Q) = \infty$, which gives (A.1) in this case.

To conclude the proof, note that in cases (i) and (ii), by the proofs given above, a minimizing sequence may be constructed in $\mathcal{P}_0$.

\[
\square
\]

### A.3 Square integrability and relative entropy

The goal of this section is to prove Theorem A.9 which characterizes Girsanov type changes of measure with finite relative entropy in the case of continuous sample paths. We are not aware of the existence of a proof of this result in the literature (in particular, of the implication (i) $\Rightarrow$ (ii)). An informal discussion of the (ii) $\Rightarrow$ (i) may be found in [DE97, Section 4.6.4].

Throughout this section, let $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P})$ denote a filtered probability space satisfying the usual conditions. Let $Q$ be a probability measure on $(\Omega, \mathcal{F})$ that is equivalent to $\mathbb{P}$ and let $Z := \frac{dQ}{d\mathbb{P}}$.

Since $\mathbb{P}$ and $Q$ are equivalent, we do not need to distinguish between $\mathbb{P}$- or $Q$-almost surely.

Let $\mathcal{M}_2^Q(\Omega)$ denote the space of continuous $Q$ martingales $M$ such that $\sup_{t \geq 0} \mathbb{E}^Q[M_t^2] \leq \infty$. As is well known, for $M \in \mathcal{M}_2^Q(\Omega)$, the limit $\lim_{t \to \infty} M_t$ exists almost surely and in $L^2(\mathbb{Q})$, making $\mathcal{M}_2^Q(\Omega)$ a Hilbert space equipped with norm $||M||_{\mathcal{M}_2^Q(\Omega)} = \sup_{t \geq 0} \mathbb{E}^Q[M_t^2] = \mathbb{E}^Q[M_\infty^2] = E^Q[M_\infty]$, (see e.g. [RW94b, Section IV.24]).

The density process $(Z_t)_{t \geq 0}$ is the uniformly integrable martingale defined as usual by

\[
Z_t := \mathbb{E}[Z | \mathcal{F}_t], \quad t \geq 0.
\]

Suppose, after a modification on $\mathbb{P}$-null sets, that $(Z_t)$ is a continuous martingale (which is always the case if, e.g., $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P})$ is the probability space of a canonical Wiener process by the martingale representation theorem [RW94b, Theorem IV.36.5]). By standard Girsanov change of measure arguments, there exists a continuous local $\mathbb{P}$-martingale $(M_t)$ such that $Z_t = \mathcal{E}(M_t) := \exp(M_t - \frac{1}{2}[M_t])$ for $t \geq 0$, almost surely, see e.g. [RW94b, Section IV.38]. Furthermore, $N := M - [M]$ is a $Q$-local martingale with quadratic variation process $[N] = [M]$. These processes $M$ and $N$ are fixed in this section.

The following lemma gives a useful way to understand the relative entropy of these Girsanov type changes of measure.

\[
\square
\]
Lemma A.6. Suppose that $M \in \mathcal{M}_2^1(\mathbb{Q})$, i.e.

$$E^\mathbb{Q}[|M|_{\infty}] < \infty.$$ \hspace{1cm} (21)

Then

$$\mathcal{H}(\mathbb{Q}; \mathbb{P}) = \frac{1}{2}E^\mathbb{Q}[|M|_{\infty}].$$

Proof. Since $[N] = [M]$ in fact $N$ is a $\mathbb{Q}$-martingale that is bounded in $L^2(\mathbb{Q})$. Note that

$$E^\mathbb{Q} |\log Z| = E^\mathbb{Q} |M_\infty - \frac{1}{2}[M]_{\infty}| = E^\mathbb{Q} [|N_\infty + \frac{1}{2}[M]_{\infty}|] \leq (E^\mathbb{Q} [N_\infty^2])^{1/2} + \frac{1}{2}E^\mathbb{Q}[|M|_{\infty}] < \infty.$$

In particular, $\log Z$ is $\mathbb{Q}$-integrable, so that $\mathcal{H}(\mathbb{Q}; \mathbb{P}) = E^\mathbb{Q} \log Z < \infty$. Furthermore,

$$\mathcal{H}(\mathbb{Q}; \mathbb{P}) - \frac{1}{2}E^\mathbb{Q}[|M|_{\infty}] = E^\mathbb{Q} [\log Z - \frac{1}{2}[M]_{\infty}] = E^\mathbb{Q} [N_\infty] = 0.$$

Lemma A.7. Suppose $Z[\log Z] \in L^1(\Omega, \mathcal{F}, \mathbb{P})$. Let $0 \leq \tau \leq \infty$ be an $(\mathcal{F}_t)_{t\geq0}$-stopping time. Then $Z_\tau[\log Z_\tau] \in L^1(\Omega, \mathcal{F}, \mathbb{P})$, and

$$E^\mathbb{P} [Z_\tau \log Z_\tau] \leq E^\mathbb{P} [Z \log Z].$$

Proof. Note that $(Z_t)$ is a uniformly integrable martingale so that the optional sampling theorem may be applied for any stopping time $0 \leq \tau \leq \infty$. On $\{\omega: Z_\tau(\omega) < 1\}$, using concavity of $x \mapsto \log x$ and the conditional Jensen inequality,

$$Z_\tau \log Z_\tau = Z_\tau(- \log Z_\tau) = Z_\tau(- \log E^\mathbb{P}[Z | \mathcal{F}_\tau]) \leq Z_\tau E^\mathbb{P}[- \log(Z) | \mathcal{F}_\tau] = E^\mathbb{P} [Z_\tau \log(Z) | \mathcal{F}_\tau].$$

On $\{\omega: Z_\tau(\omega) \geq 1\}$, using convexity of $x \mapsto f(x) := x \log x$ and the conditional Jensen inequality,

$$Z_\tau \log Z_\tau = f(E^\mathbb{P}[Z | \mathcal{F}_\tau]) \leq E^\mathbb{P} [f(Z) | \mathcal{F}_\tau] = E^\mathbb{P} [Z \log Z | \mathcal{F}_\tau].$$

Therefore

$$E^\mathbb{P} [Z_\tau \log Z_\tau] \leq E^\mathbb{P} [E^\mathbb{P}[Z_\tau \log(Z) | \mathcal{F}_\tau] 1_{\{Z_\tau < 1\}}] + E^\mathbb{P} [E^\mathbb{P}[Z \log(Z) | \mathcal{F}_\tau] 1_{\{Z_\tau \geq 1\}}]$$

$$= E^\mathbb{P} [Z \log Z].$$

Lemma A.8. For all $t \geq 0$,

$$\inf_{0 \leq s \leq t} Z_s > 0, \quad \text{almost surely.}$$

Proof. Let $t \geq 0$ and $A \in \mathcal{F}_t$ with $\mathbb{P}(A) > 0$. Then

$$E^\mathbb{P} [Z_t 1_A] = E^\mathbb{P} [Z 1_A] > 0,$$

so that $Z_t > 0$, almost surely. Therefore $(Z_t)$ has continuous, strictly positive trajectories. For such trajectories

$$\inf_{0 \leq s \leq t} Z_s = \min_{0 \leq s \leq t} Z_s > 0.$$

We arrive at the main result of this section, which characterizes Girsanov type change of measures with finite relative entropy in the case of a continuous density process.

Proposition A.9. The following are equivalent.
(i) $\mathcal{H}(Q;P) < \infty$;
(ii) $N = M - [M] \in M_2^c(Q)$.

If (i) and (ii) hold, then

$$\mathcal{H}(Q;P) = \frac{1}{2} \mathbb{E}^Q [[M]_\infty] < \infty.$$

**Proof.** Lemma A.6 establishes that (ii) implies (i). For the converse direction, suppose now that $Z|\log Z| \in L^1(P)$. Define for $n \in \mathbb{N}$ the stopping times

$$\tau_n^1 := \inf\{t \geq 0 : |Z_t| > n\} \quad (\uparrow \infty \text{ as } n \to \infty, \text{ almost surely}).$$

Then $Z^{\tau_n^1}$ is a bounded martingale for all $n$, and in particular $Z^{\tau_n^1}$ has bounded quadratic variation $[Z^{\tau_n^1}]_\infty$. By Lemma A.8, the trajectories of $Z^{-1}$ are locally bounded, so that we may define stopping times

$$\tau_n^2 := \inf\{t \geq 0 : Z^{-1}_t > n\} \quad (\uparrow \infty \text{ as } n \to \infty, \text{ almost surely}).$$

Let $\tau_n := \tau_n^1 \land \tau_n^2$. Then $Z^{\tau_n} = \mathcal{E}(M^{\tau_n})$. Since $dM_t = Z^{-1}_t dZ_t$,

$$[M^{\tau_n}]_\infty \leq \int_0^{\tau_n} Z^{-2}_s d[Z]_s \leq n^2 \int_0^{\tau_n} d[Z]_s \leq n^4,$$

almost surely. So, for all $n \in \mathbb{N}$, by Lemma A.6 and Lemma A.7,

$$\mathbb{E}^Q [[M^{\tau_n}]_\infty] = \mathbb{E}^P [Z_{\tau_n} \log Z_{\tau_n}] \leq \mathbb{E}^P [Z|\log Z]|.$$

By Fatou’s lemma,

$$\mathbb{E}^Q [[M]_\infty] \leq \liminf_{n \to \infty} \mathbb{E}^Q [[M^{\tau_n}]_\infty] \leq \mathbb{E}^P [Z|\log Z]|.$$

We may apply Lemma A.6 one more time to deduce that, if (i) and (ii) hold, then

$$\mathbb{E}^Q [[M]_\infty] = \mathcal{H}(Q;P).$$

\qed

**Acknowledgements**

We are grateful to dr. ir. O. van Gaans (Mathematical Institute, Leiden University) for his review of our work and his helpful suggestions that enabled us to improve the exposition. Also we wish to acknowledge the important suggestions of several anonymous reviewers. These comments have had significant impact upon this paper, for which we are very much indebted.

**References**

[BD98] M Boué and P Dupuis. A variational representation for certain functionals of Brownian motion. *The Annals of Probability*, pages 1641–1659, 1998.

[BH13] Ralf Banisch and Carsten Hartmann. Meshless discretization of LQ-type stochastic control problems. http://arxiv.org/abs/1309.7497, September 2013.

[DE97] Paul Dupuis and Richard S. Ellis. *A Weak Convergence Approach to the Theory of Large Deviations (Wiley Series in Probability and Statistics)*. Wiley-Interscience, 1997.

[Fle82] Wendell H Fleming. Logarithmic transformations and stochastic control. In *Advances in filtering and optimal stochastic control (Cocoyoc, 1982)*, volume 42 of *Lecture Notes in Control and Inform. Sci.*, pages 131–141. Springer, Berlin, 1982.
[FS09] Wendell H. Fleming and Halil Mete Soner. *Controlled Markov Processes and Viscosity Solutions (Stochastic Modelling and Applied Probability)*. Springer, 2009.

[HS91] AC Heinricher and RH Stockbridge. Optimal control of the running max. *SIAM journal on control and optimization*, 29(4):936–953, 1991.

[Kal02] Olav Kallenberg. *Foundations of Modern Probability (Probability and Its Applications)*. Springer, 2002.

[Kap05] Hilbert J. Kappen. Linear Theory for Control of Nonlinear Stochastic Systems. *Physical Review Letters*, 95(20):200201, November 2005.

[KS91] Ioannis Karatzas and Steven Shreve. *Brownian Motion and Stochastic Calculus (Graduate Texts in Mathematics)*. Springer, 1991.

[Mon13] Michael Monoyios. Malliavin calculus method for asymptotic expansion of dual control problems. *to appear in SIAM Journal on Financial Mathematics*, 2013.

[NOkP10] Giulia Di Nunno, Bernt Øksendal, and Frank Proske. *Malliavin Calculus for Lévy Processes with Applications to Finance (Universitext)*. Springer, 2010.

[Nua06] David Nualart. *The Malliavin Calculus and Related Topics*. 2006.

[OK91] Daniel L Ocone and Ioannis Karatzas. A generalized clark representation formula, with application to optimal portfolios. *Stochastics and Stochastic Reports*, 34(3-4):187–220, 1991.

[RW94a] L C G Rogers and David Williams. *Diffusions, Markov processes, and martingales. Vol. 1*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Ltd., Chichester, second edition, 1994.

[RW94b] L C G Rogers and David Williams. *Diffusions, Markov processes, and martingales. Vol. 2*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Ltd., New York, 2nd edition, 1994.

[Tod06] Emanuel Todorov. Linearly-solvable Markov decision problems. In B Schölkopf, J Platt, and T Hoffman, editors, *Advances in Neural Information Processing Systems 19*, number 1, pages 1369–1376. MIT Press, Cambridge, MA, 2006.