



**HAL**  
open science

## Stopping with expectation constraints: 3 points suffice

Stefan Ankirchner, Nabil Kazi-Tani, Maike Klein, Thomas Kruse

► **To cite this version:**

Stefan Ankirchner, Nabil Kazi-Tani, Maike Klein, Thomas Kruse. Stopping with expectation constraints: 3 points suffice. 2017. hal-01525439

**HAL Id: hal-01525439**

**<https://hal.science/hal-01525439>**

Preprint submitted on 20 May 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Stopping with expectation constraints: 3 points suffice

Stefan Ankirchner\*    Nabil Kazi-Tani    Maike Klein  
Thomas Kruse

May 20, 2017

We consider the problem of optimally stopping a one-dimensional continuous-time Markov process with a stopping time satisfying an expectation constraint. We show that it is sufficient to consider only stopping times such that the law of the process at the stopping time is a weighted sum of 3 Dirac measures. The proof uses recent results on Skorokhod embeddings in order to reduce the stopping problem to a linear optimization problem over a convex set of probability measures.

## Introduction

Let  $(Y_t)_{t \in \mathbb{R}_{\geq 0}}$  be a one-dimensional strong Markov process with respect to a right-continuous filtration  $(\mathcal{F}_t)$ . Let  $f : \mathbb{R} \rightarrow \mathbb{R}$  be measurable and denote by  $\mathcal{T}(T)$  the set of  $(\mathcal{F}_t)$ -stopping times such that  $E[\tau] \leq T \in \mathbb{R}_{\geq 0}$ . In the following we consider the optimal stopping problem

$$\text{maximize } E[f(Y_\tau)] \quad \text{subject to } \tau \in \mathcal{T}(T). \quad (0.1)$$

The problem (0.1) arises whenever an average time constraint applies for any stopping rule. If a process has to be stopped repeatedly and independently of the previous stopping times, then it is reasonable to impose an average time constraint instead of a sharp constraint of the type  $\tau \leq T$ , a.s. For example think of the question of when to stop searching for a parking space. If you face this question every morning when driving to your work, it is more likely

---

\*Parts of the research results presented in this article have been obtained while S.A. and N. K.-T. were visiting the Vietnam Institute for Advanced Study in Mathematics (VIASM) in August and September 2016. Financial support from the VIASM is gratefully acknowledged.

that you impose an average constraint on your searching time than just a sharp upper bound.

Notice that there is no deterministic dependence of the constraint on time. For solving the stopping problem (0.1) one needs to turn the expectation constraint into a scenario-dependent constraint.

In this article we show that for the stopping problem (0.1) it is sufficient to consider only stopping times  $\tau$  such that the law of  $Y_\tau$  is a weighted sum of at most 3 Dirac measures. Any such stopping time can be interpreted as a composition of exit times from intervals.

We also show that in general a reduction to weighted sums of 2 Dirac measures is not possible. In particular, one can not split the state space into a deterministic stopping and continuation region. This is in contrast to stopping problems with a sharp bound on the stopping time and to stopping problems with infinite time horizon and discounting.

Our idea for proving a reduction to 3 Dirac measures is to rewrite the stopping problem (0.1) as a linear optimization problem over a set of probability measures. To this end we use recent results on the Skorokhod embedding problem characterizing the set  $\mathcal{A}(T)$  of probability distributions that can be embedded into  $Y$  (see [1] and [9]) with stopping times having expectation smaller than or equal to  $T$ . As for standard linear problems the maximal value of the optimization is attained by extreme points. The extreme points of  $\mathcal{A}(T)$  turn out to be contained in the set of probability measures that can be written as a weighted sums of at most 3 Dirac measures.

To the best of our knowledge, the idea of using Skorokhod embeddings to analyze optimal stopping problems first appeared in [19], where the authors solve an optimal stopping problem for the geometric Brownian motion, under the Choquet integral, and where the only condition imposed on the stopping times is that they are almost surely finite. When it comes to optimal stopping problems with constraints on the stopping time distribution, the literature is rather scarce: Kennedy [10] solves an optimal stopping problem with an expectation constraint for a discrete time process. The author uses Lagrangian techniques to write the problem as an unconstrained one. Within a continuous time setting, the article [2] formulates a dynamic programming principle for stopping problems with expectation constraints and derives a verification theorem. Different constraints have been recently studied: Bayraktar and Miller [3] consider the problem of optimally stopping the Brownian motion with a stopping time whose distribution is atomic with finitely many points of mass. Miller [11] analyzes stopping problems with time inconsistent constraints. In [4], the authors use optimal transport techniques to treat the problem of optimally stopping the Brownian motion with a stopping time having a fixed specified distribution. Further stopping problems with an expectation constraint on the stopping time have been solved by Urusov [17].

Let  $\theta \in [0, 1]$  be the moment at which a standard Brownian motion attains its maximal value on  $[0, 1]$  and let  $\alpha \geq 0$ . Then Urusov [17] characterizes the stopping time that minimizes  $E[(\tau - \theta)^+]$  over all stopping times  $\tau$  satisfying the expectation constraint  $E[(\tau - \theta)^-] \leq \alpha$ .

The article is organized as follows. In Section 1 we describe the precise setting of stopping problems considered. In Section 2 we show how one can reduce the stopping problem to an optimization over the set of probability measures that are weighted sums of 3 Dirac measures. In Section 3 we provide sufficient conditions for the existence of an optimal stopping time. Throughout we assume that the process to stop is in natural scale - this is, as explained in Section 4, not a restriction.

## 1. Stopping after consecutive exit times

In this section we rigorously set the framework for the optimal stopping problem. The process to stop is assumed to be a one-dimensional continuous strong Markov process. Let the state space  $J \subseteq \mathbb{R}$  be an open, half-open or closed interval and denote by  $I := (l, r)$  the interior of  $J$ , where  $-\infty \leq l < r \leq \infty$ . Let  $\Omega = \mathcal{C}([0, \infty), J)$  be the space of all continuous  $J$ -valued functions and  $(Y_t)_{t \geq 0}$  be the coordinate process, i.e.  $Y_t(\omega) = \omega(t)$ ,  $t \geq 0$ ,  $\omega \in \Omega$ . Let  $(\mathcal{F}_t^0)$  be the  $\sigma$ -algebra generated by  $(Y_t)_{t \geq 0}$  and  $\mathcal{F}^0 := \mathcal{F}_\infty^0 := \bigvee_{t \geq 0} \mathcal{F}_t^0$ . Denote by  $(\theta_t)_{t \geq 0}$  the family of shift operators on  $\Omega$  defined by  $(\theta_t \omega)(s) = \omega(t + s)$ ,  $s \geq 0$ . Let  $(P^x)_{x \in J}$  be a family of probability measures on  $(\Omega, \mathcal{F}^0)$  that is a regular diffusion in the sense of [16, Chapter V.45]. In particular, we have  $P^x[Y_0 = x] = 1$  for all  $x \in J$ . Regularity means that for every  $y \in I$  and  $x \in J$  we have that  $P^y[\tau_x < \infty] > 0$ , where  $\tau_x = \inf\{t \geq 0 : Y_t = x\}$ .

For a probability measure  $\mu$  on  $(J, \mathcal{B}(J))$  let

$$P^\mu(A) := \int P^x(A) \mu(dx), \quad A \in \mathcal{F}^0.$$

Let  $\mathcal{F}^\mu$  be completion of  $\mathcal{F}^0$  with respect to  $\mu$  and set  $\mathcal{F}_t^\mu = \sigma(\mathcal{F}_t^0, \mathcal{N})$ ,  $t \geq 0$ , where  $\mathcal{N}$  denotes the collection of  $P^\mu$ -null sets in  $\mathcal{F}^\mu$ . One can show that  $(\Omega, \mathcal{F}^\mu, (\mathcal{F}_t^\mu), P^\mu)$  satisfies the usual conditions. We set  $\mathcal{F}_t = \bigcap_\mu \mathcal{F}_t^\mu$  and  $\mathcal{F} = \bigcap_\mu \mathcal{F}^\mu$ . Observe that  $(\mathcal{F}_t)$  is right-continuous, but that in general  $(\Omega, \mathcal{F}, (\mathcal{F}_t), P^\mu)$  does not satisfy the usual conditions. The process  $(Y_t)_{t \geq 0}$  fulfills the strong Markov property (cf. Theorem 9.4, Chapter III, in [15]): For any bounded  $\mathcal{F}$ -measurable mapping  $\eta$  and any finite  $(\mathcal{F}_t)$ -stopping time  $\tau$  we have

$$E^\mu[\eta \circ \theta_\tau | \mathcal{F}_\tau] = E^{Y_\tau}[\eta], \quad P^\mu - a.s.$$

Let  $m$  be the speed measure of the diffusion  $(P^x)_{x \in J}$  on  $J$ . Since  $Y$  is regular we have for all  $a < b \in I$

$$0 < m([a, b]) < \infty.$$

For simplicity we assume that the diffusion  $Y$  is in natural scale. If  $Y$  is not in natural scale, then there exists a function  $s : J \rightarrow \mathbb{R}$ , the so-called scale function, such that  $s(Y_t)$ ,  $t \geq 0$ , is in natural scale. In Section 4 below we show how to reduce the general stopping problem to the case where the process to stop is in natural scale.

In addition, we assume that if an endpoint is accessible, then it is absorbing. This implies  $Y$  is a local martingale (see Corollary 46.15 in [16]).

For  $y \in I$  and  $x \in \bar{J}$  we define  $q_y : \bar{J} \rightarrow [0, \infty]$ ,

$$q_y(x) = \frac{1}{2}m(\{y\})|x - y| + \int_y^x m((y, u)) du \quad (1.1)$$

with the convention that  $m((y, u)) = -m((u, y))$  whenever  $u < y$ . Moreover, we set  $q_y(r) = \lim_{x \uparrow r} q_y(x)$  if  $r = \infty$  and  $q_y(l) = \lim_{x \downarrow l} q_y(x)$  if  $l = -\infty$ .

One can show that  $q_y(Y_t) - t$ ,  $t \in \mathbb{R}_{\geq 0}$ , is a local martingale with respect to  $P^y$  and  $(\mathcal{F}_t)$  (this follows e.g. from Theorem 3.6 in Chapter VII, [14]). Moreover, the behavior of  $q_y$  at  $l$  and  $r$  determines whether the process attains the boundary points with a positive probability or not.

**Lemma 1.1.** *We have  $q_y(r) < \infty$  if and only if  $r \in J$ . Similarly,  $q_y(l) < \infty$  if and only if  $l \in J$ .*

*Proof.* Since we have not found a reference with a proof, we provide a sketch of the proof in the appendix.  $\square$

Let  $f : J \rightarrow \mathbb{R}$  be a Borel-measurable function determining the payoff of the stopping problem. Throughout we make the following assumption on  $f$ :

**Assumption (A).** For every  $y \in J$  there exists a  $C(y) \in \mathbb{R}_{\geq 0}$  such that

$$f(x) \geq -C(y)(1 + q_y(x)), \quad x \in J. \quad (1.2)$$

For any  $T \in \mathbb{R}_{\geq 0}$ , let  $\mathcal{T}(T, y)$  be the set of all  $(\mathcal{F}_t)$ -stopping times  $\tau$  with  $E^y[\tau] \leq T$ .

**Remark 1.2.** Assumption (A) ensures that the expectation  $E^y[f(Y_\tau)]$  exists for all  $y \in J$ ,  $T \geq 0$  and  $\tau \in \mathcal{T}(T, y)$ . Indeed, for an appropriately chosen localizing sequence of stopping times  $(\tau_n)$ , it holds that

$$\begin{aligned} E^y[\{f(Y_\tau)\}^-] &\leq E^y[C(y)(1 + q_y(Y_\tau))] = C(y)(1 + E^y[\liminf_{n \rightarrow \infty} q_y(Y_{\tau \wedge \tau_n})]) \\ &\leq C(y)(1 + \liminf_{n \rightarrow \infty} E^y[q_y(Y_{\tau \wedge \tau_n})]) = C(y)(1 + \liminf_{n \rightarrow \infty} E^y[\tau \wedge \tau_n]) \\ &\leq C(y)(1 + T). \end{aligned}$$

We consider the problem of finding the stopping time in  $\mathcal{T}(T, y)$  that maximizes the expected payoff  $E^y[f(Y_\tau)]$ . The value function is defined by

$$v(T, y) = \sup_{\tau \in \mathcal{T}(T, y)} E^y[f(Y_\tau)], \quad (1.3)$$

for all  $T \geq 0$  and  $y \in J$ .

**Remark 1.3.** If Assumption (A) is replaced by the stronger assumption that for every  $y \in J$  there exists a  $C(y) \in \mathbb{R}_{\geq 0}$  such that

$$|f(x)| \leq C(y)(1 + q_y(x)), \quad x \in J, \quad (1.4)$$

then the value function is finite. Indeed, it follows by using similar arguments as in Remark 1.2 that  $\sup_{\tau \in \mathcal{T}(T, y)} E^y[f(Y_\tau)] \leq C(y)(1 + T)$  in this case. The following example shows that in general one can not dispense with condition (1.4) if we want to guarantee that  $v$  is finite. For a Brownian motion  $W$  we have  $q_0(x) = x^2$ . Let  $f(x) = |x|^{2+\varepsilon}$ ,  $\varepsilon > 0$ , be the payoff function. For every  $T > 0$  the first time  $\rho(a, T)$ ,  $a > 0$ , when  $W$  hits  $a$  or  $-T/a$  has expectation  $T$  under  $P^0$ . Hence,

$$v(T, 0) \geq \sup_{a > 0} E^0[f(W_{\rho(a, T)})] = \sup_{a > 0} \left\{ a^{2+\varepsilon} \frac{T}{a^2 + T} + \frac{a^2}{a^{2+\varepsilon}} \frac{T^{2+\varepsilon}}{a^2 + T} \right\} = \infty.$$

For stopping problems *without* an expectation constraint an optimal stopping time is given by the exit time of the continuation region (see Corollary 2.9, Chapter I in [13]). In particular, for solving unconstrained stopping problems it is enough to consider exit times from intervals. For constrained stopping problems a reduction to simple exit times is not possible. We show, however, that it is enough to consider at most three consecutive exit times.

To give a precise statement, we denote for  $a, b \in \mathbb{R}$  with  $a \leq b$  the first hitting time of  $a$  by  $\tau_a = \inf\{t \geq 0 : Y_t = a\}$  and the first exit time from the interval  $(a, b)$  after time  $r \geq 0$  by  $\tau_{a,b}(r) = \inf\{t \geq r : Y_t \notin (a, b)\}$ . Moreover, we write  $\mathcal{T}_3(T, y)$  for the collection of stopping times  $\tau \in \mathcal{T}(T, y)$  for which there exist  $p_1, p_2, p_3 \in [0, 1]$  with  $p_1 + p_2 + p_3 = 1$  and  $a, c, d \in \mathbb{R}$  with  $a \leq c \leq d$  such that

$$\tau = \tau_{a,b}(\tau_{\bar{\mu}}) + \mathbf{1}_{\{Y_{\tau_{a,b}(\tau_{\bar{\mu}})} = b\}} \inf\{t \geq 0 : Y_{t+\tau_{a,b}(\tau_{\bar{\mu}})} \in \{c, d\}\},$$

where  $\bar{\mu} = p_1 a + p_2 c + p_3 d$  and  $b = \frac{p_2 c + p_3 d}{1 - p_1}$ . Notice that  $b \in (c, d)$ .

One of our main results is that the stopping problem (1.3) can be simplified to the set  $\mathcal{T}_3(T, y)$ .

**Theorem 1.4.** *We have*

$$v(T, y) = \sup_{\tau \in \mathcal{T}_3(T, y)} E^y[f(Y_\tau)]. \quad (1.5)$$

We prove Theorem 1.4 in the following section. We do so by reducing problem (1.3) to an optimization over a set of probability measures.

Theorem 1.4 brings up the question whether the supremum is attained in  $\mathcal{T}_3(T, y)$ . In Section 3 below we provide sufficient conditions guaranteeing the existence of an optimal stopping time in  $\mathcal{T}_3(T, y)$ .

## 2. Optimal stopping as a measure optimization

In this section we first explain how one can reduce the stopping problem (1.3) to a linear optimization problem over a set of probability measures satisfying some integrability constraints. The linear nature of the measure optimization allows us then to conclude that the maximum values are attained by extreme points, which here are weighted sums of three Dirac measures.

We denote by  $\mathcal{M} = \mathcal{M}(J)$  the set of all probability measures on  $\mathbb{R}$  with support in  $J$  and by  $\mathcal{M}^1$  the set of all measures in  $\mathcal{M}$  with finite first moment  $\bar{\mu} = \int x \mu(dx)$ . Let  $\mathcal{A}(T, y)$  be the set of measures  $\mu \in \mathcal{M}^1$  satisfying the following properties:

1. a) If  $l > -\infty$ , then  $\bar{\mu} \leq y$ .  
b) If  $r < \infty$ , then  $\bar{\mu} \geq y$ .
2.  $\mu$  integrates  $q_y$  such that

$$\int q_y(x) \mu(dx) \leq T - H(y, \bar{\mu}),$$

where

$$H(y, \bar{\mu}) = \begin{cases} (y - \bar{\mu})(m((y, \infty)) + \frac{1}{2}m(\{y\})), & \bar{\mu} < y, \\ 0, & \bar{\mu} = y, \\ (\bar{\mu} - y)(m((-\infty, y)) + \frac{1}{2}m(\{y\})), & \bar{\mu} > y. \end{cases}$$

Results from [9] on the Skorokhod embedding problem for diffusions (and from [1] for processes described in terms of SDEs) imply that  $\mathcal{A}(T, y)$  coincides with the set of probability measures that can be embedded into  $Y$  under  $P^y$  with stopping times  $\tau$  satisfying  $E^y(\tau) \leq T$ . More precisely, we have the following:

**Proposition 2.1.** *Let  $\mu \in \mathcal{M}$ . There exists a stopping time  $\tau \in \mathcal{T}(T, y)$  with  $Y_\tau \sim \mu$  under  $P^y$  if and only if  $\mu \in \mathcal{A}(T, y)$ .*

*Proof.* Let  $\tau \in \mathcal{T}(T, y)$  be an embedding of  $\mu$  in  $Y$  under  $P^y$ , i.e. let  $Y_\tau$  have the distribution  $\mu$  under  $P^y$ . Then [7] and [12] imply that

- if  $l > -\infty$ , then  $\bar{\mu} \leq y$ ,

- if  $r < \infty$ , then  $\bar{\mu} \geq y$ .

Thus,  $\mu \in \mathcal{M}^1$  whenever  $r$  or  $l$  is finite. Section 3.5 in [9] shows that if  $I = (-\infty, \infty)$  and  $\tau$  is an integrable embedding for  $\mu$ , then  $\mu \in \mathcal{M}^1$ . If  $\bar{\mu} = y$ , then it follows from Theorem 2.4. in [9] that

$$\int q_y(x)\mu(dx) \leq E^y[\tau] \leq T = T - H(y, \mu).$$

If  $\bar{\mu} < y$ , then we conclude from Theorem 3.6 in [9] that

$$\int q_y(x)\mu(dx) + (y - \bar{\mu}) \left( m((y, \infty)) + \frac{1}{2}m(\{y\}) \right) \leq T,$$

which implies that 2. holds true. If  $\bar{\mu} > y$ , we again apply Theorem 3.6 in [9] to obtain  $\int q_y(x)\mu(dx) \leq T - H(y, \bar{\mu})$ .

For the reverse direction let  $\mu \in \mathcal{A}(T, y)$  and assume first that  $\mu$  is centered around  $y$ . Then  $\mu$  can be embedded in  $Y$  under  $P^y$  for  $-\infty \leq l < r \leq \infty$  by [7] and [12]. It follows from Theorem 3.4 in [9] that there exists a minimal stopping time  $\tau$  with  $Y_\tau \sim \mu$  under  $P^y$  and

$$E^y[\tau] = \int q_y(x)\mu(dx) \leq T.$$

Hence,  $\tau \in \mathcal{T}(T, y)$ .

Now let  $\mu \in \mathcal{A}(T, y)$  with  $\bar{\mu} < y$ . Then we have  $r = \infty$ . Theorem 3.6 in [9] shows the existence of a minimal embedding  $\tau$  of  $\mu$  in  $Y$  under  $P^y$  with

$$E^y[\tau] = \int q_y(x)\mu(dx) + (y - \bar{\mu}) \left( m((y, \infty)) + \frac{1}{2}m(\{y\}) \right) \leq T,$$

where the last inequality follows from the second property of  $\mu$ . Hence,  $\tau \in \mathcal{T}(T, y)$ .

Finally, for  $\mu \in \mathcal{A}(T, y)$  with  $\bar{\mu} > y$ , using similar arguments, one can show that there exists a stopping time  $\tau$  with  $Y_\tau \sim \mu$  under  $P^y$  and  $E^y[\tau] \leq T$ .  $\square$

**Remark 2.2.** The function  $q_y$  appearing in the definition of the set of measures  $\mathcal{A}(T, y)$  plays for the Markov process  $Y$  the same role than the function  $x \mapsto x^2$  plays for the Brownian motion. Indeed, we know that when  $Y$  is a Brownian motion starting in  $y = 0$ , we can find an embedding of  $\mu$  with an integrable stopping time if and only if  $\mu$  is centered and in  $L^2$ . The papers [9] and [1] identify the function  $q_y$  as the counterpart of the second-order moment when  $Y$  is a general diffusion.

**Remark 2.3.** When  $\mu$  is not centered around  $y$ , there is a non zero function  $H$  in the constraint 2. of  $\mathcal{A}(T, y)$ . To understand this condition, notice that  $H(y, \bar{\mu}) = E^y[\inf\{t \geq 0 : Y_t = \bar{\mu}\}]$ . This tells that to embed a measure  $\mu$  such that  $\bar{\mu} \neq y$ , we first wait until the first hitting time of  $\bar{\mu}$  and then we embed  $\mu$  in  $Y$ , started at  $\bar{\mu}$ .



Proposition 2.1 allows to reformulate the stopping problem (1.3) as a linear problem on  $\mathcal{M}$ .

**Corollary 2.4.** *We have*

$$v(T, y) = \sup_{\mu \in \mathcal{A}(T, y)} \int f(x) \mu(dx) \quad (2.1)$$

and for any optimal  $\mu \in \mathcal{A}(T, y)$  there exists an optimal stopping time  $\tau \in \mathcal{T}(T, y)$  in (1.3) with  $Y_\tau \sim \mu$  under  $P^y$ .

Notice that the functional  $\mu \mapsto \int f(x) \mu(dx)$  is linear on  $\mathcal{A}(T, y)$ . We have thus obtained a linear problem over a set of probability measures  $\mu$  with some integrability constraints. Recall that for standard linear problems the maximum value is attained by extreme points. We have a similar result for an optimization problem  $\int g d\mu$  over measures  $\mu \in \mathcal{M}$  satisfying moment constraints of the form  $\int f_i d\mu \leq c_i$ ,  $g$  and  $f_i$  measurable,  $c_i \in \mathbb{R}$ ,  $1 \leq i \leq n$ . The maximum value of  $\int g d\mu$  is also attained in the set of extreme points, see [18]. Furthermore, the extreme points are contained in the set of all weighted Dirac measures with at most  $n + 1$  mass points satisfying the moment constraints.

In the following we denote the extreme points of a convex set  $A \subseteq \mathcal{M}$  by  $\mathcal{E}(A)$  and for any  $M \subseteq \mathcal{M}$  we denote by  $M_3$  the set of all measures in  $M$  which are a weighted sum of at most 3 Dirac measures.

Now we reduce the optimization problem (2.1) to an optimization problem over weighted sums of Dirac measures.

**Theorem 2.5.** *We have*

$$v(T, y) = \sup_{\mu \in \mathcal{A}_3(T, y)} \int f(x) \mu(dx). \quad (2.2)$$

*Proof.* We consider two cases. In the first case we assume that all measures  $\mu$  in  $\mathcal{A}(T, y)$  are centered around  $y$ , i.e.  $\bar{\mu} = y$ . Observe that  $\bar{\mu} = y$  for all  $\mu \in \mathcal{A}(T, y)$  if and only if one of the following four cases is satisfied: 1.  $I$  is bounded, 2.  $l > -\infty$ ,  $r = \infty$  and  $m((y, \infty)) = \infty$ , 3.  $l = -\infty$ ,  $r < \infty$  and  $m((-\infty, y)) = \infty$  and 4.  $I = \mathbb{R}$ ,  $m((y, \infty)) = \infty$  and  $m((-\infty, y)) = \infty$ . The optimization problem (2.1) can be rewritten as

$$v(T, y) = \sup_{t \in [0, T]} \sup_{\mu \in \mathcal{D}(t, y)} \int f(x) \mu(dx),$$

where  $\mathcal{D}(t, y) = \{\mu \in \mathcal{M}^1 : \bar{\mu} = y \text{ and } \int q_y(x) \mu(dx) = t\}$ ,  $0 \leq t \leq T$ . Theorem 2.1(b) and Theorem 3.2 in [18] imply that

$$\sup_{\mu \in \mathcal{D}(t, y)} \int f(x) \mu(dx) = \sup_{\mu \in \mathcal{E}(\mathcal{D}(t, y))} \int f(x) \mu(dx) = \sup_{\mu \in \mathcal{D}_3(t, y)} \int f(x) \mu(dx)$$

because  $\mathcal{D}_3(t, y)$  coincides with  $\mathcal{E}(\mathcal{D}(t, y))$ . For  $t \in [0, T]$  we have  $\mathcal{D}_3(t, y) \subseteq \mathcal{A}_3(T, y)$ . Therefore,

$$v(T, y) = \sup_{t \in [0, T]} \sup_{\mu \in \mathcal{D}_3(t, y)} \int f(x) \mu(dx) = \sup_{\mu \in \mathcal{A}_3(T, y)} \int f(x) \mu(dx).$$

In the second case the set  $\mathcal{A}(T, y)$  also contains uncentered measures. We define

$$\mathcal{A}^+(T, y) = \begin{cases} \{\mu \in \mathcal{A}(T, y) : \bar{\mu} \geq y\}, & \text{if } \exists \mu \in \mathcal{A}(T, y) \text{ with } \bar{\mu} > y \\ \emptyset, & \text{if } \bar{\mu} \leq y \text{ for all } \mu \in \mathcal{A}(T, y), \end{cases}$$

$$\mathcal{A}^-(T, y) = \begin{cases} \{\mu \in \mathcal{A}(T, y) : \bar{\mu} \leq y\}, & \text{if } \exists \mu \in \mathcal{A}(T, y) \text{ with } \bar{\mu} < y \\ \emptyset, & \text{if } \bar{\mu} \geq y \text{ for all } \mu \in \mathcal{A}(T, y), \end{cases}$$

Observe that at least one of the sets  $\mathcal{A}^+(T, y)$  or  $\mathcal{A}^-(T, y)$  is nonempty and that (2.1) can be reduced to the two optimization problems  $\sup_{\mu \in \mathcal{A}^+(T, y)} \int f(x) \mu(dx)$  and  $\sup_{\mu \in \mathcal{A}^-(T, y)} \int f(x) \mu(dx)$ , where we follow the convention that the supremum over the empty set is equal to  $-\infty$ . If  $\mathcal{A}^+(T, y)$  is nonempty, then

$$\mathcal{A}^+(T, y) = \left\{ \mu \in \mathcal{M}^1 : \bar{\mu} \geq y, \int q_y(x) \mu(dx) \leq T - H(y, \bar{\mu}) \right\}$$

$$= \left\{ \mu \in \mathcal{M}^1 : \int -x \mu(dx) \leq -y, \int (q_y(x) + Cx) \mu(dx) \leq T + Cy \right\},$$

where  $C = m((-\infty, y)) + \frac{1}{2}m(\{y\}) < \infty$ . Therefore, Theorem 3.2 in [18] implies that

$$\sup_{\mu \in \mathcal{A}^+(T, y)} \int f(x) \mu(dx) = \sup_{\mu \in \mathcal{E}(A^+(T, y))} \int f(x) \mu(dx).$$

By Theorem 2.1(a) in [18] we have  $\mathcal{E}(A^+(T, y)) \subseteq \mathcal{A}_3^+(T, y)$ . Thus,

$$\sup_{\mu \in \mathcal{A}^+(T, y)} \int f(x) \mu(dx) = \sup_{\mu \in \mathcal{A}_3^+(T, y)} \int f(x) \mu(dx). \quad (2.3)$$

If  $\mathcal{A}^-(T, y)$  is nonempty, similar arguments show that (2.3) holds with  $\mathcal{A}^+(T, y)$  and  $\mathcal{A}_3^+(T, y)$  replaced by  $\mathcal{A}^-(T, y)$  and  $\mathcal{A}_3^-(T, y)$ , respectively. Since  $\mathcal{A}_3(T, y) = \mathcal{A}_3^-(T, y) \cup \mathcal{A}_3^+(T, y)$  we conclude that

$$v(T, y) = \sup_{\mu \in \mathcal{A}_3(T, y)} \int f(x) \mu(dx).$$

□

With Theorem 2.5 we can prove Theorem 1.4.

*Proof of Theorem 1.4.* Let  $\mu \in \mathcal{A}_3(T, y)$  with exactly three mass points  $a < c < d$ . First observe that we can assume that  $\mu$  is centered around  $y$ . Otherwise the first hitting time of  $\bar{\mu}$  is integrable wrt.  $P^y$  (Theorem 2.4 in [9]) and we wait until  $Y$  hits  $\bar{\mu}$  and then continue as in the centered case.

Extending the Balayage method developed by Chacon and Walsh in [6], we construct consecutive exit times as follows:

$$\begin{aligned}\tau_1 &= \inf\{t \geq 0 : Y_t \notin (a, b)\} \\ \tau_2 &= \tau_1 + \inf\{t \geq 0 : Y_t \notin (c, d)\} \circ \theta_{\tau_1},\end{aligned}$$

where  $b = (\mu(\{c\})c + \mu(\{d\})d)/(1 - \mu(\{a\}))$ . Notice that  $b \in (c, d)$ . The stopping time  $\tau_2$  is an embedding of  $\mu$  into  $Y$  under  $P^y$ . By using that  $q_y(Y_t) - t$  is a local martingale, one can show  $E^y[\tau_2] = E^y[q_y(Y_{\tau_2})] \leq T$ ; hence  $\tau_2 \in \mathcal{T}_3(T, y)$ .

If  $\mu$  has two mass points  $a < c$ , then  $\tau = \inf\{t \geq 0 : Y_t \notin (a, c)\} \in \mathcal{T}_3(T, y)$ . And similar, if  $\mu = \delta_a$ , then  $\tau = \inf\{t \geq 0 : Y_t \notin J \setminus \{a\}\} \in \mathcal{T}_3(T, y)$ .  $\square$

The following example shows that in general a reduction to  $\mathcal{A}_2(T, y)$ , the set of probability measures in  $\mathcal{A}(T, y)$  that are weighted sums of at most 2 Dirac measures, is not possible.

**Example 2.6.** Let  $(Y_t)_{t \geq 0}$  be a Brownian motion starting in 0 and let  $f(x) = 1_{\{|x| \geq 1\}}$ ,  $x \in \mathbb{R}$ , be the payoff function. We claim that

$$v = \sup_{\mu \in \mathcal{A}(T, 0)} \int f(x) \mu(dx) = T \wedge 1,$$

and

$$\sup_{\mu \in \mathcal{A}_2} \int f(x) \mu(dx) = \begin{cases} \frac{T}{1+T}, & T < 1, \\ 1, & T \geq 1. \end{cases}$$

Now we derive the value of the optimization problems over all measure  $\mu \in \mathcal{A}(T, 0)$  resp.  $\mu \in \mathcal{A}_2(T, 0)$ . For  $T \geq 1$  an optimal measure  $\mu^*$  is given by

$$\mu^* = \frac{1}{2} \delta_{-1} + \frac{1}{2} \delta_1.$$

Now let  $T < 1$  and observe that for every measure  $\mu \in \mathcal{A}_2(T, 0)$  at least one atom is contained in  $(-1, 1)$ . Due to the symmetry of the problem and the form of  $f$ , we can restrict ourselves to measures of the form

$$\mu^S = \frac{1}{1+S} \delta_{-S} + \frac{S}{1+S} \delta_1,$$

where  $S \in (0, T]$ . Then we obtain

$$\sup_{\mu \in \mathcal{A}_2(T, 0)} \int f(x) \mu(dx) = \sup_{S \in (0, T]} \int f(x) \mu^S(dx) = \frac{T}{1+T}.$$

Next we turn to measures with three atoms. Observe that all measures in  $\mathcal{A}_3(T, y)$  are centered. Theorem 1.4 and Theorem 2.5 imply that it is sufficient to consider measures  $\mu$  of the form

$$\mu = \frac{b}{1+b} \delta_{-1} + \frac{1-b}{(1+b)(1-c)} \delta_c + \frac{b-c}{(1+b)(1-c)} \delta_1,$$

where  $0 < b < 1$  and  $c < b < 1$ . The measure  $\mu$  corresponds to the distribution of  $Y_\tau$  under  $P^y$ , where

$$\tau = \tau_{-1,b}(0) + \mathbb{1}_{\{Y_{\tau_{-1,b}(0)}=b\}} \inf \{t \geq 0 \mid Y_{t+\tau_{-1,b}(0)} \in \{c, 1\}\}.$$

$\mu$  is a centered probability measure and

$$\int q_0(x) \mu(dx) = \int x^2 \mu(dx) = \frac{2b + bc - c}{1+b}.$$

Let  $S \in (0, T]$ ,  $b < S$  and set  $c = S - \frac{2(b-S)}{1-b}$ . Then we obtain  $c \in (-1, b)$  and  $\int q_0(x) \mu(dx) = S$ . The payoff of the measure  $\mu$  is given by

$$\int f(x) \mu(dx) = 1 - \frac{(1-b)^2}{(1+b)(1-3b+S+bS)},$$

which is maximized for  $b^S = \frac{S}{2-S} \in (0, S)$  with corresponding optimal  $c^S = 0$ . Hence, we obtain

$$\sup_{\mu \in \mathcal{A}(T,0)} \int f(x) \mu(dx) = \sup_{\mu \in \mathcal{A}_3(T,0)} \int f(x) \mu(dx) = \int f(x) \mu^*(dx) = T$$

with

$$\mu^* = \frac{T}{2} \delta_{-1} + (1-T) \delta_0 + \frac{T}{2} \delta_1.$$

### 3. Existence of an optimizer

The next example shows that the supremum in (2.2) is not always attained.

**Example 3.1.** Let  $f_1(x) = x^2 \frac{|x|}{1+|x|}$ ,  $x \in \mathbb{R}$ , and  $Y$  be a Brownian motion starting in 0 under  $P^y$ . In this case there does not exist an optimal stopping time. To prove this let  $v_1 := \sup_{\tau \in \mathcal{T}(T,0)} E^0[f_1(Y_\tau)]$ . Moreover, consider the second payoff function  $f_2(x) = x^2$ . Note that for any integrable stopping time  $\tau$  we have  $E^0[Y_\tau^2] = E^0[\tau]$ . Therefore,  $v_2 := \sup_{\tau \in \mathcal{T}(T,0)} E^0[f_2(Y_\tau)] = T$ .

One can show that  $v_1 = v_2$ . Indeed, on the one hand it must hold that  $v_1 \leq v_2$  since  $f_1 \leq f_2$ . On the other hand, for the stopping times  $\tau_n = \tau_{-1/n, nT}$  we have  $E^y[\tau_n] = T$  and

$$E^0[f_1(Y_{\tau_n})] = \frac{nT}{1/n + nT} \frac{1}{n^2} \frac{1/n}{1 + 1/n} + \frac{1/n}{1/n + nT} n^2 T^2 \frac{nT}{1 + nT} \longrightarrow T,$$

as  $n \rightarrow \infty$ , and hence  $v_1 \geq v_2$ .

From  $v_1 = v_2$  we can deduce that the supremum can not be attained in  $v_1$ , because for any stopping time  $\tau \neq 0$  with  $E^0[\tau] < \infty$  we have  $P^0[f_1(Y_\tau) < f_2(Y_\tau)] > 0$ .

In the case where  $I$  is bounded we have the following existence result.

**Theorem 3.2.** *Assume that  $I$  is a finite interval and  $f$  is upper-semicontinuous. Then there exists an optimal stopping time in (1.3).*

*Proof.* Let  $(\mu_n)$  be a sequence in  $\mathcal{A}(T, y)$  such that  $\lim_{n \rightarrow \infty} \int f d\mu_n = v$ . Since  $\bar{I}$  is compact,  $(\mu_n)$  is tight. By Prokhorov's Theorem  $(\mu_n)$  converges weakly along a subsequence to a probability measure  $\mu$  on  $J$ . We show that  $\mu \in \mathcal{A}(T, y)$ . Observe that  $y = \bar{\mu}_n := \int x \mu_n(dx) \rightarrow \int x \mu(dx) = \bar{\mu}$  as  $n \rightarrow \infty$ . Thus,  $\bar{\mu} = y$ . Next, lower-semicontinuity of  $q$  on  $J$  implies  $\int q d\mu \leq \liminf_{n \rightarrow \infty} \int q d\mu_n \leq T$ . Since  $f$  is upper-semicontinuous we have  $v = \lim_{n \rightarrow \infty} \int f d\mu_n \leq \int f d\mu \leq v$ . Hence  $\mu$  is optimal in (2.1) and by Corollary 2.4 there exists an optimal stopping in (1.3).  $\square$

If the state space equals  $\mathbb{R}$ , there exists an optimizer under conditions on the payoff function  $f$  and the speed measure  $m$ .

**Theorem 3.3.** *Assume that  $J = \mathbb{R}$ ,  $m((-\infty, y)) = m((y, \infty)) = \infty$ ,  $\limsup_{x \rightarrow \infty} \frac{f(x)}{q_y(x)} \leq 0$ ,  $\limsup_{x \rightarrow -\infty} \frac{f(x)}{q_y(x)} \leq 0$  and that  $f$  is upper semi-continuous. Then there exists an optimal stopping time in  $\mathcal{T}_3(T, y)$  for (1.5).*

*Proof.* First observe that the assumption  $m((-\infty, y)) = m((y, \infty)) = \infty$  ensures that  $\int x \nu(dx) = y$  for all  $\nu \in \mathcal{A}(T, y)$ . Now, let  $\mu_n = \sum_{j=1}^3 p_n^j \delta_{x_n^j} \in \mathcal{A}(T, y)$ ,  $n \in \mathbb{N}$ , be a sequence of measures such that

$$\lim_{n \rightarrow \infty} \int f d\mu_n = v(T, y).$$

If the sequence  $(x_n^1)_{n \in \mathbb{N}}$  is unbounded, choose a subsequence, also denoted by  $(x_n^1)_{n \in \mathbb{N}}$ , such that either  $\lim_{n \rightarrow \infty} x_n^1 = -\infty =: x^1$  or  $\lim_{n \rightarrow \infty} x_n^1 = \infty =: x^1$ . If  $(x_n^1)_{n \in \mathbb{N}}$  is bounded, extract a subsequence such that  $\lim_{n \rightarrow \infty} x_n^1 = x^1 \in \mathbb{R}$ . By extracting further subsequences, proceed in the same way with  $(x_n^2)_{n \in \mathbb{N}}$  and  $(x_n^3)_{n \in \mathbb{N}}$ . Then, refine once again the sequence to obtain that  $(p_n^1, p_n^2, p_n^3) \rightarrow (p^1, p^2, p^3) \in [0, 1]^3$  as  $n \rightarrow \infty$ . Overall we obtain for  $n \rightarrow \infty$  that

$$(x_n^1, x_n^2, x_n^3, p_n^1, p_n^2, p_n^3) \rightarrow (x^1, x^2, x^3, p^1, p^2, p^3) \in (\mathbb{R} \cup \{-\infty\} \cup \{\infty\})^3 \times [0, 1]^3.$$

Let  $K = \{j \in \{1, 2, 3\} | x^j \in \mathbb{R}\}$  and  $\mu = \sum_{j \in K} p_j \delta_{x^j}$ . Let  $i \in \{1, 2, 3\} \setminus K$ . From the fact that

$$p_n^i q_y(x_n^i) \leq \sum_{j=1}^3 p_n^j q_y(x_n^j) = \int q_y(x) \mu_n(dx) \leq T \quad (3.1)$$

it follows that

$$\limsup_{n \rightarrow \infty} p_n^i = \limsup_{n \rightarrow \infty} \frac{1}{q_y(x_n^i)} p_n^i q_y(x_n^i) = 0$$

and hence

$$\lim_{n \rightarrow \infty} p_n^i = 0. \quad (3.2)$$

Next observe that the assumption  $m((-\infty, y)) = m((y, \infty)) = \infty$  implies that  $\lim_{x \rightarrow \infty} \frac{q_y(x)}{|x|} = \lim_{x \rightarrow -\infty} \frac{q_y(x)}{|x|} = \infty$ . Combining this with (3.1) yields that

$$\limsup_{n \rightarrow \infty} |x_n^i| p_n^i = \limsup_{n \rightarrow \infty} \frac{|x_n^i|}{q_y(x_n^i)} p_n^i q_y(x_n^i) = 0$$

and hence

$$\lim_{n \rightarrow \infty} x_n^i p_n^i = 0. \quad (3.3)$$

Combining (3.1) and the assumptions  $\limsup_{x \rightarrow \infty} \frac{f(x)}{q_y(x)} \leq 0$  and  $\limsup_{x \rightarrow -\infty} \frac{f(x)}{q_y(x)} \leq 0$  yields that

$$\limsup_{n \rightarrow \infty} f(x_n^i) p_n^i = \limsup_{n \rightarrow \infty} \frac{f(x_n^i)}{q_y(x_n^i)} p_n^i q_y(x_n^i) \leq 0. \quad (3.4)$$

Observe that it follows from (3.2) that

$$\mu(\mathbb{R}) = \sum_{j \in K} p^j = \lim_{n \rightarrow \infty} \sum_{j \in K} p_n^j = \lim_{n \rightarrow \infty} \sum_{i=1}^3 p_n^i = \lim_{n \rightarrow \infty} \mu_n(\mathbb{R}) = 1.$$

Moreover, (3.3) proves that

$$\bar{\mu} = \int x \mu(dx) = \sum_{j \in K} x^j p^j = \lim_{n \rightarrow \infty} \sum_{j \in K} p_n^j x_n^j = \lim_{n \rightarrow \infty} \sum_{i=1}^3 p_n^i x_n^i = \lim_{n \rightarrow \infty} \int x \mu_n(dx) = y.$$

This implies that

$$\begin{aligned} \int q_y(x) \mu(dx) &= \sum_{j \in K} p^j q_y(x^j) = \lim_{n \rightarrow \infty} \sum_{j \in K} p_n^j q_y(x_n^j) \leq \limsup_{n \rightarrow \infty} \sum_{i=1}^3 p_n^i q_y(x_n^i) \\ &= \limsup_{n \rightarrow \infty} \int q_y(x) \mu_n(dx) \leq T = T - H(y, \bar{\mu}). \end{aligned}$$

Consequently, we have that  $\mu \in \mathcal{A}_3(T, y)$  and in particular  $v(T, y) \geq \int f d\mu$ . Finally, it follows from (3.4) and upper semi-continuity of  $f$  that

$$\begin{aligned} \int f d\mu &= \sum_{j \in K} p^j f(x^j) \geq \limsup_{n \rightarrow \infty} \sum_{j \in K} p_n^j f(x_n^j) \\ &\geq \limsup_{n \rightarrow \infty} \sum_{j \in K} p_n^j f(x_n^j) + \limsup_{n \rightarrow \infty} \sum_{j \notin K} p_n^j f(x_n^j) \\ &\geq \limsup_{n \rightarrow \infty} \sum_{i=1}^3 p_n^i f(x_n^i) = \lim_{n \rightarrow \infty} \int f d\mu_n = v(T, y), \end{aligned}$$

which implies that  $v(T, y) = \int f d\mu$ . □

**Remark 3.4.** Example 3.1 shows that the condition that  $\limsup_{x \rightarrow \infty} \frac{f(x)}{q_y(x)} = \limsup_{x \rightarrow -\infty} \frac{f(x)}{q_y(x)} = 0$  in Theorem 3.3 cannot be weakened in general.

## 4. General Diffusions

In this section we show how to deal with the optimal stopping problem (1.3) if  $Y$  is not in natural scale.

Let  $s$  be the scale function of  $Y$  and  $m$  its speed measure. Define  $Z_t = s(Y_t)$ ,  $t \geq 0$ . Then  $Z$  (or more precisely  $(Q^z)_{z \in s(J)}$ , where  $Q^z$  is the  $P^{s^{-1}(z)}$  distribution of  $(s(Y_t))_{t \geq 0}$ ) is a diffusion in natural scale on  $J^Z := s(J)$  by Theorem 46.12, in Chapter V, [16], respectively Theorem 2.1 in [5]. The speed measure  $m^Z$  of  $Z$  is given by  $m^Z = m \circ s^{-1}$ , i.e.  $m^Z((a, b)) = m((s^{-1}(a), s^{-1}(b)))$  for all  $[a, b] \subseteq J^Z$ . Denote by  $I^Z := (l^Z, r^Z) := (s(l), s(r))$  the interior of  $J^Z$  and for  $z \in I^Z$  let  $q_z^Z$  be the function on  $\overline{J^Z}$  defined in (1.1) with  $m = m^Z$ . In this section we impose the following assumption on  $f : J \rightarrow \mathbb{R}$ .

**Assumption:** For every  $y \in J$  there exists  $C(y) \in \mathbb{R}_{\geq 0}$  such that

$$f(w) \geq -C(y) \left( 1 + \int_{s(y)}^{s(w)} m((y, s^{-1}(u))) du \right), \quad w \in J.$$

The optimal stopping problem (1.3) can be rewritten as

$$v(T, y) = \sup_{\tau \in \mathcal{T}(T, y)} E^{P^y}[f(Y_\tau)] = \sup_{\tau \in \mathcal{T}^Z(T, s(y))} E^{Q^{s(y)}}[(f \circ s^{-1})(Z_\tau)],$$

where  $\mathcal{T}^Z(T, z)$  denotes the set of all  $(\mathcal{F}_t)$ -stopping times with  $E^{Q^z}[\tau] \leq T$ . Hence we can convert the optimal stopping problem with reward function  $f$  for the process  $Y$  under  $P^y$  into an optimal stopping problem with reward function  $f \circ s^{-1}$  for  $Z$  under  $Q^{s(y)}$ . Observe that  $v$  is well-defined, because we have for all  $z \in J^Z$

$$\begin{aligned} f(s^{-1}(z)) &\geq -C(y) \left( 1 + \int_{s(y)}^z m((y, s^{-1}(u))) du \right) \\ &= -C(y) \left( 1 + \int_{s(y)}^z m^Z(s(y), u) du \right) \\ &= -C(y) (1 + q_{s(y)}^Z(z)). \end{aligned}$$

Therefore, Assumption (A) is satisfied. Let  $\mathcal{A}^Z(T, z)$  be the set of all probability measures  $\nu$  on  $\mathbb{R}$  with support in  $J^Z$  and finite first moment satisfying

Properties 1. and 2. of Section 2, where  $l = l^Z$ ,  $r = r^Z$ ,  $q_z = q_z^Z$  and  $m = m^Z$ . Denote by  $\mathcal{A}_3^Z(T, z)$  the set of all probability measures in  $\mathcal{A}^Z(T, z)$  which can be written as a weighted sum of at most 3 Dirac measures. Theorem 2.5 implies that

$$v(T, y) = \sup_{\nu \in \mathcal{A}_3^Z(T, s(y))} \int f(s^{-1}(x)) \nu(dx). \quad (4.1)$$

In the reduced optimization problem (4.1) the set  $\mathcal{A}_3^Z(T, s(y))$  depends on the process  $Z$ , its state space  $J^Z$  and speed measure  $m^Z$ . Next we aim to characterize the set of measure  $\mathcal{A}_3^Z(T, s(y))$  in terms of the primal process  $Y$ , its state space  $J$  and speed measure  $m$ . First observe that an  $(\mathcal{F}_t)$ -stopping time  $\tau$  embeds  $\mu$  in  $Y$  under  $P^y$  if and only if  $\tau$  embeds  $\nu := \mu \circ s^{-1}$  in  $Z$  under  $Q^{s(y)}$ . In order to transfer the properties of  $\nu \in \mathcal{A}_3^Z(T, s(y))$  to  $\mu$  we introduce the set  $\mathcal{A}^s(T, y)$  of probability measures  $\rho$  on  $\mathbb{R}$  with support in  $J$  which satisfy the following properties:

1.  $\int_J |s(x)| \rho(dx) < \infty$ .
2. a) If  $s(l) > -\infty$ , then  $\int_J s(x) \rho(dx) \leq s(y)$ .  
b) If  $s(r) < \infty$ , then  $\int_J s(x) \rho(dx) \geq s(y)$ .
3.  $\rho$  integrates  $q_y$  such that

$$\begin{aligned} \frac{1}{2} m(\{y\}) \int_J |s(x) - s(y)| \rho(dx) + \int_J \int_{s(y)}^{s(x)} m((y, s^{-1}(u))) du \rho(dx) \\ \leq T - H\left(s(y), \int_J s(x) \rho(dx)\right), \end{aligned}$$

where

$$H(x, w) = \begin{cases} (x - w)(m((s^{-1}(x), r)) + \frac{1}{2}m(\{s^{-1}(x)\})), & w < x, \\ 0, & w = x, \\ (w - x)(m((l, s^{-1}(x))) + \frac{1}{2}m(\{s^{-1}(x)\})), & w > x. \end{cases}$$

Let  $\mathcal{A}_3^s(T, y)$  be the measures in  $\mathcal{A}^s(T, y)$  which can be written as a weighted sum of at most 3 Dirac measures.

Let  $y \in I$ . Then the mapping  $\mu \mapsto \nu := \mu \circ s^{-1}$  is a bijection from  $\mathcal{A}^s(T, y)$  to  $\mathcal{A}^Z(T, s(y))$ , because

$$\begin{aligned} \int_{J^Z} |x| \nu(dx) &= \int_J |s(x)| \mu(dx), \\ \int_{J^Z} x \nu(dx) &= \int_J s(x) \mu(dx), \\ \int_{J^Z} q_{s(y)}^Z(x) \nu(dx) &= \int_J q_{s(y)}^Z(s(x)) \mu(dx) \\ &= \frac{1}{2} m(\{y\}) \int_J |s(x) - s(y)| \mu(dx) + \int_J \int_{s(y)}^{s(x)} m((y, s^{-1}(u))) du \mu(dx). \end{aligned}$$



Furthermore, the number of mass points of  $\mu$  and  $\nu$  coincide. Thus, we have proven the following theorem.

**Theorem 4.1.** *For a diffusion  $Y$  with scale function  $s$  we have*

$$v(T, y) = \sup_{\nu \in \mathcal{A}_3^Z(T, s(y))} \int f(s^{-1}(x)) \nu(dx) = \sup_{\mu \in \mathcal{A}_3^S(T, y)} \int f(x) \mu(dx).$$

## A. Proof of auxiliary results

We prove in this appendix Lemma 1.1 that characterizes in terms of  $q_y$  whether a boundary point is attained with a positive probability or not.

*Proof of Lemma 1.1.* We prove the statement only for  $r$ .

If  $r = \infty$ , then  $q_y(r) = \infty$ . Next suppose that  $r \in J$  (in particular  $r < \infty$ ). Fix  $a \in (l, y)$  and let  $\tau_{a,r}$  be the first exit time of the interval  $(a, r)$ . Since  $q_y(Y_t) - t$  is a local martingale, we have  $E^y[\tau_{a,r}] = E^y[q_y(Y_{\tau_{a,r}})]$ . This further entails

$$E^y[\tau_{a,r}] = \frac{r-y}{r-a} q_y(a) + \frac{y-a}{r-a} q_y(r) \geq \frac{y-a}{r-a} q_y(r).$$

Proposition 3.1 in Chapter VII, [14], implies that  $\tau_{a,r}$  is integrable and, hence,  $q_y(r) < \infty$ .

Now consider the case when  $r \notin J$  and  $r < \infty$ . Then the first hitting time  $\tau_r$  of the point  $r$  satisfies  $\tau_r = \infty$ ,  $P^y$ -almost surely. Assume that  $q_y(r) < \infty$  for some  $y \in I$ . Let  $l < a < y < b < r$  and denote by  $\tau_{a,b}$  the first exit time of the interval  $(a, b)$ . Similar to the first case one can show

$$E^y[\tau_{a,b}] = E^y[q_y(Y_{\tau_{a,b}})] = \frac{b-y}{b-a} q_y(a) + \frac{y-a}{b-a} q_y(b).$$

Hence,

$$\begin{aligned} E^y[\tau_a] = E^y[\tau_{a,r}] &= \lim_{b \uparrow r} E^y[\tau_{a,b}] = \lim_{b \uparrow r} \left\{ \frac{b-y}{b-a} q_y(a) + \frac{y-a}{b-a} q_y(b) \right\} \\ &= \frac{r-y}{r-a} q_y(a) + \frac{r-a}{b-a} q_y(r). \end{aligned}$$

In particular,  $\tau_a < \infty$ ,  $P^y$ -almost surely, and  $\tau_a$  embeds the Dirac measure  $\delta_a$  into  $Y^y$ . But if  $r < \infty$ , a measure  $\mu$  can be embedded into  $Y$  under  $P^y$  if and only if  $\bar{\mu} = \int x \mu(dx) \geq y$  by [12] and [8]. Therefore,  $q_y(r) = \infty$ .  $\square$

## References

- [1] S. Ankirchner, D. Hobson, and P. Strack. Finite, integrable and bounded time embeddings for diffusions. *Bernoulli*, 21(2):1067–1088, 2015.
- [2] S. Ankirchner, M. Klein, and T. Kruse. A verification theorem for optimal stopping problems with expectation constraints. To appear in *Appl. Math. Opt.*, 2017.
- [3] E. Bayraktar and C. W. Miller. Distribution-constrained optimal stopping. 2016.
- [4] M. Beiglböck, M. Eder, C. Elgert, and U. Schmock. Geometry of distribution-constrained optimal stopping problems. preprint 2016.
- [5] J. K. Brooks. One-dimensional diffusions and their convergence in distribution. In *Handbook of measure theory, Vol. I, II*, pages 295–343. North-Holland, Amsterdam, 2002.
- [6] R. V. Chacon and J. B. Walsh. One-dimensional potential embedding. In *Séminaire de Probabilités, X (Première partie, Univ. Strasbourg, Strasbourg, année universitaire 1974/1975)*, pages 19–23. Lecture Notes in Math., Vol. 511. Springer, Berlin, 1976.
- [7] A. Cox and D. Hobson. An optimal embedding for diffusions. *Stochastic Processes and Applications*, 111:17–39, 2004.
- [8] A. Cox and D. Hobson. Skorokhod embeddings, minimality and non-centred target distributions. *Probability theory and Related Fields*, 135:395–414, 2005.
- [9] D. Hobson. Integrability of solutions of the Skorokhod embedding problem for diffusions. *Electron. J. Probab.*, 20:1–26, 2015.
- [10] D. P. Kennedy. On a constrained optimal stopping problem. *J. Appl. Probab.*, 19(3):631–641, 1982.
- [11] C. W. Miller. Non-linear PDE Approach to Time-Inconsistent Optimal Stopping. *arXiv preprint arXiv:1510.05766*, 2015.
- [12] J. L. Pedersen and G. Peskir. The Azéma-Yor embedding in non-singular diffusions. *Stochastic Process. Appl.*, 96(2):305–312, 2001.
- [13] G. Peskir and A. Shiryaev. *Optimal stopping and free-boundary problems*. Springer, 2006.
- [14] 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.
- [15] L. Rogers and D. Williams. *Diffusions, Markov processes, and Martingales. Vol. 1: Foundations*. Cambridge Mathematical Library. Cambridge University Press, second edition, 2000.

- [16] L. Rogers and D. Williams. *Diffusions, Markov processes, and Martingales. Vol. 2: Itô Calculus*. Cambridge Mathematical Library. Cambridge University Press, second edition, 2000.
- [17] M. A. Urusov. On a property of the moment at which Brownian motion attains its maximum and some optimal stopping problems. *Theory of Probability & Its Applications*, 49(1):169–176, 2005.
- [18] G. Winkler. Extreme points of moment sets. *Math. Oper. Res.*, 13(4):581–587, 1988.
- [19] Z. Q. Xu and X. Y. Zhou. Optimal stopping under probability distortion. *Ann. Appl. Probab.*, 23(1):251–282, 02 2013.