

# GEOMETRY OF DISTRIBUTION-CONSTRAINED OPTIMAL STOPPING PROBLEMS

MATHIAS BEIGLBÖCK, MANU EDER, CHRISTIANE ELGERT, AND UWE SCHMOCK

**ABSTRACT.** We adapt ideas and concepts developed in optimal transport (and its martingale variant) to give a geometric description of optimal stopping times  $\tau$  of Brownian motion subject to the constraint that the distribution of  $\tau$  is a given probability  $\mu$ . The methods work for a large class of cost processes. (At a minimum we need the cost process to be measurable and  $(\mathcal{F}_t^0)_{t \geq 0}$ -adapted. Continuity assumptions can be used to guarantee existence of solutions.) We find that for many of the cost processes one can come up with, the solution is given by the first hitting time of a barrier in a suitable phase space. As a by-product we recover classical solutions of the inverse first passage time problem / Shiryaev's problem.

**Keywords** distribution-constrained optimal stopping, optimal transport, inverse first passage problem, Shiryaev's problem

**MSC (2010)** Primary 60G42, 60G44; Secondary 91G20

## 1. APPETIZER

To whet the reader's appetite and to give some idea of the kind of problems that can be solved with the methods presented in this paper we would like to start with two corollaries to our main results. In Section 3 we will present these main results and in Section 4 we will use them to prove Corollary 1.1 from them.

Both Corollary 1.1 and Corollary 1.2 assert that the solutions of certain optimal stopping problems can be described by a barrier in an appropriate phase space.

In this section, let  $(B_t)_{t \geq 0}$  be a Brownian motion started<sup>1</sup> in 0 on some filtered probability space  $(\Omega, \mathcal{G}, (\mathcal{G}_t)_{t \geq 0}, \mathbb{P})$  satisfying the usual conditions and let  $\mu$  be a measure on  $(0, \infty)$ . First we consider optimal stopping problems of the following form.

**Problem** ( $\text{OPTSTOP}^{\psi(B_t, t)}$ ). Among all stopping times  $\tau \sim \mu$  on  $(\Omega, \mathcal{G}, (\mathcal{G}_t)_{t \geq 0}, \mathbb{P})$  find the maximizer of

$$\tau \mapsto \mathbb{E}[Z_\tau] ,$$

where the process  $Z$  is of the form  $Z_t = \psi(B_t, t)$ .

---

*Date:* September 13, 2017.

Corresponding author: Mathias Beiglböck, mathias.beiglboeck@tuwien.ac.at

The first and the second author were supported by FWF-grant Y00782, the third author by Jubiläumsfonds 16549, the fourth author by Jubiläumsfonds 16549 and FWF-grant P25216.

---

All authors

TU Vienna, Faculty of Mathematics, Wiedner Hauptstraße 8-10, 1040 Vienna.

---

<sup>1</sup>We note that the results presented in this section remain valid for Brownian motions started according to a general law  $\lambda$  at the cost of slightly more tedious moment conditions in the formulation of Corollaries 1.1 and 1.2.

**Corollary 1.1.** *Assume that  $\mu$  has finite first moment. There is an upper semicontinuous function  $\beta : \mathbb{R}_+ \rightarrow [-\infty, \infty]$  such that the stopping time*

$$\tau := \inf\{t > 0 : B_t \leq \beta(t)\} \quad (1.1)$$

*has distribution  $\mu$ .*

$\tau$  has the following uniqueness properties: *On the one hand it is the a.s. unique stopping time which has distribution  $\mu$  and which is of the form (1.1) (we will later say that such a stopping time is the hitting time of a downwards barrier).*

*On the other hand  $\tau$  is also the a.s. unique solution of  $(\text{OPTSTOP}^{\psi(B_t, t)})$  for a number of different  $\psi$ . Namely:*

- *Let  $p \geq 0$ , assume  $\mu$  has finite moment of order  $\frac{1}{2} + p + \varepsilon$  for some  $\varepsilon > 0$  and let  $A : \mathbb{R}_+ \rightarrow \mathbb{R}$  be strictly increasing and  $|A(t)| \leq K(1 + t^p)$  for some constant  $K$ .<sup>2</sup> Then we may choose*

$$\psi(B_t, t) = B_t A(t) .$$

- *Let  $p \geq 2$ , assume  $\mu$  has finite moment of order  $\frac{p}{2} + \varepsilon$  for some  $\varepsilon > 0$  and let  $\phi : \mathbb{R} \rightarrow \mathbb{R}$  satisfy  $\phi''' > 0$  as well as  $|\phi(y)| \leq K(1 + |y|^p)$  for some constant  $K$ . Then we may choose*

$$\psi(B_t, t) = \phi(B_t) .$$

To give an example of a slightly more complicated functional amenable to analysis with our tools consider

**Problem  $(\text{OPTSTOP}^{B_t^*})$ .** Among all stopping times  $\tau \sim \mu$  on  $(\Omega, \mathcal{G}, (\mathcal{G}_t)_{t \geq 0}, \mathbb{P})$  find the maximizer of

$$\tau \mapsto \mathbb{E}[B_\tau^*] ,$$

where  $B_t^* = \sup_{s \leq t} B(s)$ .

**Corollary 1.2.** *Assume that  $\mu$  has finite moment of order  $\frac{3}{2}$ . Then  $(\text{OPTSTOP}^{B_t^*})$  has a solution  $\tau$  given by*

$$\tau = \inf\{t > 0 : B_t - B_t^* \leq \beta(t)\}$$

*for some upper semicontinuous function  $\beta : \mathbb{R}_+ \rightarrow [-\infty, 0]$ .*

We emphasize that the solutions to the constrained optimal stopping problems provided in Corollaries 1.1 and 1.2 represent particular applications of the abstract results obtained below. Figure 1 presents graphical depictions of stopping rules of several further solutions of constrained optimal stopping problems (together with the respective optimality properties). These stopping rules can be derived – under suitable moment conditions – using arguments very similar to those required for Corollaries 1.1 and 1.2 (see also the comments in Remark 7.1 at the end of the paper).

## 2. BACKGROUND - MARTINGALE OPTIMAL TRANSPORT AND SHIRYAEV'S PROBLEM

In this article we consider distribution-constrained stopping problems from a mass transport perspective. Specifically we find that problems of the form exemplified in  $(\text{OPTSTOP}^{\psi(B_t, t)})$  and  $(\text{OPTSTOP}^{B_t^*})$  are amenable to techniques originally developed for the martingale version of the classical mass transport problem. This martingale optimal transport problem arises naturally in robust finance; papers

<sup>2</sup>One may of course choose  $0 \leq p < \frac{1}{2}$ ,  $\varepsilon := \frac{1}{2} - p$  and e.g.  $A(t) := t^p$  so that no moment conditions beyond those at the very beginning of this theorem are imposed on  $\mu$ .

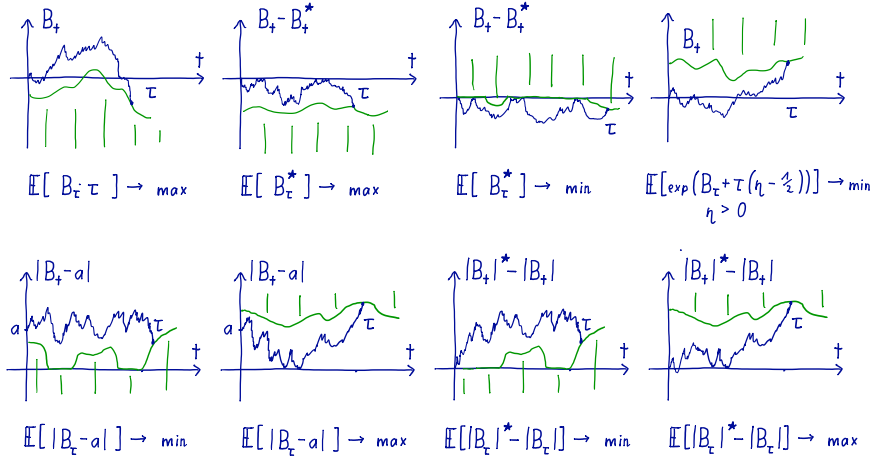


FIGURE 1. Solutions to constrained optimal stopping problems.

to investigate such problems include [25, 8, 18, 16, 12, 20, 31]. In mathematical finance, transport techniques complement the Skorokhod embedding approach (see [32, 24] for an overview) to model-independent/robust finance.

A fundamental idea in optimal transport is that the optimality of a transport plan is reflected by the geometry of its support set which can be characterized using the notion of *c-cyclical monotonicity*. The relevance of this concept for the theory of optimal transport has been fully recognized by Gangbo and McCann [19], based on earlier work of Knott and Smith [28] and Ruschendorf [36, 37] among others. Inspired by these ideas, the literature on martingale optimal transport has developed a ‘monotonicity principle’ which allows to characterize martingale transport plans through geometric properties of their support sets, cf. [9, 39, 7, 6, 22, 10].

The main contribution of this article is to establish a monotonicity principle which is applicable to distribution-constrained optimal stopping problems. This transport approach turns out to be remarkably powerful, in particular we will find that questions as raised in Problems  $(\text{OPTSTOP}^{\psi(B_t, t)})$  and  $(\text{OPTSTOP}^{B_t^*})$  can be addressed using a relatively intuitive set of arguments.

The distribution-constrained optimal stopping problem  $(\text{OPTSTOP})$  (and specifically  $(\text{OPTSTOP}^{B_t^*})$ ) arises naturally in financial and actuarial mathematics. We refer the reader to [23] which describes various examples (unit-linked life insurances, stochastic modelling for health insurances, the liquidation of an investment portfolio, the valuation of swing options).

Bayraktar and Miller [5] consider the same optimization problem that we treat here. However their setup and methods are rather distinct from the ones used here: they assume that the target distribution is given by finitely many atoms and that the target functional depends solely on the terminal value of Brownian motion. Following the measure valued martingale approach of Cox and Kallblad [14], [5] address the constrained optimal stopping problem using a Bellman perspective.

The problem to construct a stopping time  $\tau$  of Brownian motion such that the law of  $\tau$  matches a given distribution on the real line was proposed by Shiryaev in his Banach Center lectures in the 1970’s, it has since been called Shiryaev’s problem or inverse first passage problem. Dudley and Gutmann [17] provide an abstract measure-theoretic construction. An early barrier-type solution to the inverse first

passage problem was given by Anulova [3]. She constructs a symmetric two-sided barrier (corresponding to the case  $a = 0$  in the sixth picture of Figure 1). Anulova discretises the measure  $\mu$  and concludes through approximation arguments. The solution to the inverse first passage problem given in Corollary 1.1 was derived by Chen, Cheng, and Chadam, and Saunders [13] based on a variational inequality which describes the corresponding barrier. Notably, this is predated by a (formal) PDE description of such barriers given by Avellaneda and Zhu [4] in the context of credit risk modeling. Ekström and Janson [13] relate this solution to an optimal stopping problem and provide an integral equation for the barrier. Analytic solutions to the inverse first passage problem are known only in a few cases ([11, 29, 38, 33, 1, 2]). An interesting connection between the inverse first passage problem and Skorokhod's problem is provided by Jaimungal, Kreinin, and Valov [26].

### 3. STATEMENT OF MAIN RESULTS

**Assumption 1.** Throughout we will assume that  $(\Omega, \mathcal{G}, (\mathcal{G}_t)_{t \geq 0}, \mathbb{P})$  is a filtered probability space and that  $(B_t)_{t \geq 0}$  is an adapted process which has continuous paths on  $(\Omega, \mathcal{G}, (\mathcal{G}_t)_{t \geq 0})$ , such that  $B$  can be regarded as a measurable map from  $\Omega$  to  $C(\mathbb{R}_+)$ , the space of continuous functions from  $\mathbb{R}_+$  to  $\mathbb{R}$ . The cost function  $c$  will always be a measurable map  $C(\mathbb{R}_+) \times \mathbb{R}_+ \rightarrow \mathbb{R}$ .  $\mu$  will denote a probability measure on  $\mathbb{R}_+$ .

Then the problem we consider can be stated as follows.

**Problem (OPTSTOP).** Among all stopping times  $\tau \sim \mu$  find the minimizer of

$$\tau \mapsto \mathbb{E}[c(B, \tau)] .$$

Here we formulate our main optimization problem in terms of minimization, following the usual convention in the optimal transport literature (which is also used in the closely related paper [6]). Clearly, a sign change transforms this into a maximization problem and in our applications we will in fact turn to this latter version when resulting formulations appear more natural. We trust that this will not cause confusion.

Throughout we will also make the following assumptions without further mention:

**Assumption 2.**

- (1)  $c$  is measurable,  $(\mathcal{F}_t^0)_{t \geq 0}$ -adapted, where  $(\mathcal{F}_t^0)_{t \geq 0}$  is the filtration on  $C(\mathbb{R}_+)$  generated by the canonical process  $(\omega \mapsto \omega(t))_{t \in \mathbb{R}_+}$ .
- (2) There is a  $\mathcal{G}_0$ -measurable random variable  $U$  which is uniformly distributed on  $[0, 1]$  and independent of the process  $(B_t)_{t \geq 0}$ .
- (3) There is a probability measure  $\lambda$  s.t.  $(B_t)_{t \geq 0}$  is a Brownian motion with initial law  $\lambda$ , i.e.  $B_0 \sim \lambda$ .
- (4) The problem is well-posed in the sense that  $\mathbb{E}[c(B, \tau)]$  is defined and  $> -\infty$  for all stopping times  $\tau \sim \mu$  and that  $\mathbb{E}[c(B, \tau)] < \infty$  for at least one such stopping time.
- (5)  $\int t^{p_0} d\mu(t) < \infty$ , where  $p_0 \geq 0$  is some constant that we fix here and that can be chosen when applying the results from this section.

A note on language: The adjective ‘‘adapted’’ is usually applied to processes whose time argument is written in subscript form. For any filtered measurable space  $\tilde{\Omega}$  and any function  $f : \tilde{\Omega} \times \mathbb{R}_+ \rightarrow \mathbb{R}$  (or possibly  $f : \tilde{\Omega} \times \mathbb{R}_+ \rightarrow [-\infty, \infty]$ ) we will interchangeably think of  $f$  simply as a function or as the process  $Y_t(\omega) := f(\omega, t)$ . And so  $f$  being adapted means the same thing as  $(Y_t)_{t \in \mathbb{R}_+}$  being adapted. Similarly

for a subset  $\Gamma$  of  $\tilde{\Omega} \times \mathbb{R}_+$  we may also think of  $\Gamma$  as its indicator function or as the process  $Y_t(\omega) := 1_\Gamma(\omega, t)$  and will also say that the set  $\Gamma$  is adapted.

With that in mind, Assumption 2.1 should seem like an obvious thing to ask for from the cost function. Also, knowing about the existence of optional projections, it should be clear no later than Lemma 5.3 that Assumption 2.1 does not pose a real restriction on the class of problems we are treating.

The role of Assumption 2.2 should become clearer soon. We would like to note at this point though that often enough our results put together will imply that the solution of Problem (OPTSTOP) for a space  $(\Omega, \mathcal{G}, (\mathcal{G}_t)_{t \geq 0}, \mathbb{P})$  which satisfies Assumption 2.2 is essentially the same as the solution of the Problem for a space which may not satisfy said assumption, and we will find that we can describe this solution in detail. This can be seen executed in the proofs of the corollaries stated in the Appetizer.

The methods in this paper work not just for Brownian motion but for a class of processes which is conceptually bigger, but then turns out to not include much beyond Brownian motion – namely for any space-homogeneous but possibly time-inhomogeneous Markov process with continuous paths which has the strong Markov property. (Here space-homogeneous means that starting the process at location  $x$  and then moving its paths to start at location  $y$  results in a version of the process started at  $y$ .) If the reader so wishes, she may think of  $B$  as a process from this slightly larger class of processes. Care was taken not to reference any properties of Brownian motion beyond those stated here. In particular our results apply to multi-dimensional Brownian motion.

Assumption 2.4 is mostly just there to ensure that we are actually talking about an optimization problem in a meaningful sense. For the problems presented in the Appetizer, the moment conditions on  $\mu$  which are given in the statement of Corollary 1.1 and Corollary 1.2 ensure that Assumption 2.4 is satisfied (as we will see in the proofs of these corollaries).

The constant  $p_0$  in Assumption 2.5 will (implicitly) appear in the statement of Theorem 3.6, one of the main results. Its role is to ensure that  $\mathbb{E}[\varphi(B, \tau)]$  will be finite for some (class of) function(s)  $\varphi$  and any solution  $\tau$  of (OPTSTOP). (The choice  $\varphi(B, \tau) = \tau^{p_0}$  is somewhat arbitrary here.)

The main results are Theorem 3.1 and Theorem 3.6.

We give two versions of Theorem 3.1. Version A is easier to state and may feel more natural, but we will need Version B (which is more general and has essentially the same proof as Version A) in the proof of the corollaries in the Appetizer.

**Theorem 3.1.**

***Version A.** Assume that the cost function  $c$  is bounded from below and lower semicontinuous when we equip  $C(\mathbb{R}_+)$  with the topology of uniform convergence on compacts. Then the Problem (OPTSTOP) has a solution.*

***Version B.** Assume that the cost function  $c$  is lower semicontinuous when we equip  $C(\mathbb{R}_+) \times \mathbb{R}_+$  with the product topology of two Polish topologies which generate the right sigma-algebras on  $C(\mathbb{R}_+)$  and  $\mathbb{R}_+$  respectively and assume that the set  $\{c_-(B, \tau) : \tau \sim \mu, \tau \text{ is a stopping time}\}$  is uniformly integrable, where  $c_- := -c \vee 0$  denotes the negative part of  $c$ . Then the Problem (OPTSTOP) has a solution.*

To state Theorem 3.6 we need a few more definitions.

*Remark 3.2.* We will find it convenient to talk about processes that don't start at time 0 but instead at some time  $t > 0$ . Similarly we will consider stopping times taking values in  $[t, \infty)$ . These will be defined on the space  $C([t, \infty))$  equipped with the filtration  $(\mathcal{F}_s^t)_{s \geq t}$ , again generated by the canonical process  $(\omega \mapsto \omega(s))_{s \geq t}$ . We refer to the distribution of Brownian motion started at time  $t$  and location  $x$  by

$\mathbb{W}_x^t$ . This is a measure on  $C([t, \infty))$ . For a probability measure  $\kappa$  on  $\mathbb{R}$  we write  $\mathbb{W}_\kappa^t$  for the distribution of Brownian motion started at time  $t$  with initial law  $\kappa$ .

**Definition 3.3** (Concatenation). For every  $t \in \mathbb{R}_+$  we have an operation  $\odot$  of concatenation, which is a map into  $C([t, \infty))$  and is defined for  $(\omega, s) \in C([t, \infty)) \times [t, \infty)$  and  $\theta \in C([s, \infty))$  with  $\theta(s) = 0$  by

$$((\omega, s) \odot \theta)(r) = \begin{cases} \omega(r) & t \leq r \leq s \\ \omega(s) + \theta(r) & r > s \end{cases}. \quad (3.1)$$

**Definition 3.4** (Stop-Go pairs). The set of Stop-Go pairs  $\mathbf{SG} \subseteq (C(\mathbb{R}_+) \times \mathbb{R}_+) \times (C(\mathbb{R}_+) \times \mathbb{R}_+)$  is defined as the set of all pairs  $((\omega, t), (\eta, t))$  (note that the time components have to match) such that

$$c(\omega, t) + \int c((\eta, t) \odot \theta, \sigma(\theta)) d\mathbb{W}_0^t(\theta) < c(\eta, t) + \int c((\omega, t) \odot \theta, \sigma(\theta)) d\mathbb{W}_0^t(\theta) \quad (3.2)$$

for all  $(\mathcal{F}_s^t)_{s \geq t}$ -stopping times  $\sigma$  for which  $\mathbb{W}_0^t(\sigma = t) < 1$ ,  $\mathbb{W}_0^t(\sigma = \infty) = 0$ ,  $\int \sigma^{p_0} d\mathbb{W}_0^t < \infty$  and for which both sides in (3.2) are defined and finite.

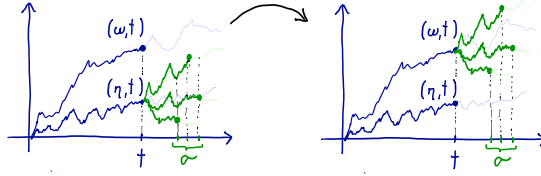


FIGURE 2. The left hand side of (3.2) corresponds to averaging the function  $c$  over the stopped paths on the left picture, the right hand side to averaging the function  $c$  over the stopped paths on the right picture.

A hopefully intuitive way of putting the definition of Stop-Go pairs into words is the following:  $((\omega, s), (\eta, t))$  form a Stop-Go pair iff, irrespective of how we might stop after time  $t$  (i.e. which stopping rule  $\sigma$  we might use after time  $t$ ), *Stopping*  $\omega$  at time  $t$  and letting  $\eta$  *Go* on is better – i.e. has lower cost – than stopping  $\eta$  and letting  $\omega$  go on.

As hinted at earlier, the definition of Stop-Go pairs depends on the parameter  $p_0$  from Assumption 2.5. A larger  $p_0$  means that we are asking for more in Assumption 2.5 and implies that we get a larger set  $\mathbf{SG}$ , as we are quantifying over fewer stopping times  $\sigma$  in the definition of  $\mathbf{SG}$ . This in turn implies that the conclusion of Theorem 3.6 below will be stronger.

**Definition 3.5** (Initial Segments). For a set  $\Gamma \subseteq C(\mathbb{R}_+) \times \mathbb{R}_+$  define the set  $\Gamma^< \subseteq C(\mathbb{R}_+) \times \mathbb{R}_+$  by

$$\Gamma^< = \{(\omega, s) : (\omega, t) \in \Gamma \text{ for some } t > s\}. \quad (3.3)$$

**Theorem 3.6** (Monotonicity Principle). *Assume that  $\tau$  solves (OPTSTOP). Then there is a measurable,  $(\mathcal{F}_t^0)_{t \geq 0}$ -adapted set  $\Gamma \subseteq C(\mathbb{R}_+) \times \mathbb{R}_+$  such that*

$$\mathbb{P}[(B_t)_{t \geq 0}, \tau] \in \Gamma = 1$$

and

$$\mathbf{SG} \cap (\Gamma^< \times \Gamma) = \emptyset. \quad (3.4)$$

The following lemma should give a first hint about how the Monotonicity Principle can be applied.

**Lemma 3.7.** *Let  $\tau$  be a solution of (OPTSTOP) and assume that the cost function  $c$  is such that there exists a measurable,  $(\mathcal{F}_t^0)_{t \geq 0}$ -adapted process  $(Y_t)_{t \geq 0}$  such that*

$$Y_t(\omega) < Y_t(\eta) \implies ((\omega, t), (\eta, t)) \in \text{SG} . \quad (3.5)$$

Define the barriers  $\mathcal{R}, \hat{\mathcal{R}} \subseteq \mathbb{R} \times \mathbb{R}_+$  by

$$\begin{aligned} \mathcal{R} &= \bigcup_{(\omega, t) \in \Gamma} (-\infty, Y_t(\omega)] \times \{t\} \\ \hat{\mathcal{R}} &= \bigcup_{(\omega, t) \in \Gamma} (-\infty, Y_t(\omega)) \times \{t\} , \end{aligned}$$

where  $\Gamma$  is a set with the properties in Theorem 3.6. Define the functions  $\tau$  and  $\hat{\tau}$  on  $C(\mathbb{R}_+)$  by

$$\begin{aligned} \tau(\tilde{\omega}) &= \inf \left\{ t \in \mathbb{R}_+ : (Y_t(B(\tilde{\omega})), t) \in \mathcal{R} \right\} \\ \hat{\tau}(\tilde{\omega}) &= \inf \left\{ t \in \mathbb{R}_+ : (Y_t(B(\tilde{\omega})), t) \in \hat{\mathcal{R}} \right\} . \end{aligned}$$

Then

$$\tau \leq \tau \leq \hat{\tau} \quad \mathbb{P}\text{-a.s.} \quad (3.6)$$

When applying this Lemma to show that some optimal stopping problem has a barrier-type solution as symbolized for example by the pictures in Figure 1 the process  $Y_t(B)$  is of course what we are labelling the vertical axes in the pictures with. So for the first picture  $Y_t(\omega) = \omega(t)$ , for the second one  $Y_t(\omega) = \omega(t) - \sup_{s \leq t} \omega(s)$ , for the third  $Y_t(\omega) = -(\omega(t) - \sup_{s \leq t} \omega(s))$  (the sign is flipped relative to the labelling in the picture because in this picture the barrier is drawn “up” instead of “down”), etc.

Notice that, contrary to customs, when we draw the barriers  $\mathcal{R}/\hat{\mathcal{R}}$  in the pictures in Figure 1 the first coordinate is the vertical axis and the second coordinate is the horizontal axis. This is because, to make cross-referencing and comparison with [6] easier, we follow their convention of always having time as the second coordinate but still in the pictures it seems more natural to put the independent variable on the horizontal axis.

Note that a priori  $\tau$  and  $\hat{\tau}$  need not be stopping times or even measurable, as we don’t know much about the sets  $\mathcal{R}$  and  $\hat{\mathcal{R}}$ .

Using the properties of a concrete process  $(Y_t)_{t \geq 0}$  we will in the proofs of Corollaries 1.1 and 1.2 be able to show that  $\tau = \hat{\tau}$  a.s. (this should not be surprising as for each time  $t$  the barriers  $\mathcal{R}$  and  $\hat{\mathcal{R}}$  differ by at most a single point) and therefore that the optimizer  $\tau$  is the hitting time of a barrier.

*Proof of Lemma 3.7.* Let  $\tilde{\omega} \in \Omega$  s.t.  $(B(\tilde{\omega}), \tau(\tilde{\omega})) \in \Gamma$ . By assumption this holds for  $\mathbb{P}$ -a.a.  $\tilde{\omega}$ . Then  $(Y_{\tau(\tilde{\omega})}(B(\tilde{\omega})), \tau(\tilde{\omega})) \in \mathcal{R}$  and therefore  $\tau(\tilde{\omega}) \leq \tau(\tilde{\omega})$ .

Next we show that  $\hat{\tau}(\tilde{\omega}) \geq \tau(\tilde{\omega})$ . Assume that  $(Y_t(B(\tilde{\omega})), t) \in \hat{\mathcal{R}}$ . We want to show that  $t \geq \tau(\tilde{\omega})$ . By the definition of  $\hat{\mathcal{R}}$  we find that there is  $\eta \in C(\mathbb{R}_+)$  with  $(\eta, t) \in \Gamma$  and  $Y_t(B(\tilde{\omega})) < Y_t(\eta)$ , so by (3.5) we know  $((B(\tilde{\omega}), t), (\eta, t)) \in \text{SG}$ . Assuming, if possible,  $t < \tau(\tilde{\omega})$  we get according to Definition 3.5 that  $(B(\tilde{\omega}), t) \in \Gamma^<$ . Therefore we have that  $((B(\tilde{\omega}), t), (\eta, t)) \in \text{SG} \cap (\Gamma^< \times \Gamma)$ , but this is a contradiction to  $\text{SG} \cap (\Gamma^< \times \Gamma) = \emptyset$ , so we must have  $t \geq \tau(\tilde{\omega})$ .  $\square$

*Remark 3.8 (Duality).* Problem (OPTSTOP) is an infinite-dimensional linear programming problem and one would hence expect that a corresponding dual problem

can be formulated. Indeed, assuming that  $c$  is lower semicontinuous and bounded from below, the value of the optimization problem equals

$$\sup_{M, \psi} \mathbb{E}[M_0] + \int \psi d\mu,$$

where the supremum is taken over bounded  $(\mathcal{G}_t)_{t \geq 0}$ -martingales  $M = (M_t)_{t \geq 0}$  and bounded continuous functions  $\psi : \mathbb{R}_+ \rightarrow \mathbb{R}$  satisfying (up to evanescence)

$$M_t + \psi(t) \leq c(B, t) .$$

This can be established in complete analogy to the duality result derived in [6, Theorem 1.2 / Section 4.2] and we do not elaborate.

#### 4. DIGESTING THE APPETIZER

We will now demonstrate how to use the Monotonicity Principle of Theorem 3.6 to derive Corollary 1.1. The proof of Corollary 1.2 is very similar but relies on understanding a technical detail which does not add much to the story at this point, so we leave it for the end of the paper.

Both of the sets  $\mathcal{R}$  and  $\hat{\mathcal{R}}$  in Lemma 3.7 have the property that (writing  $\mathcal{R}$  for the set in question)  $(y, t) \in \mathcal{R}$  and  $y' \leq y$  implies  $(y', t) \in \mathcal{R}$ . We call such sets (downwards) barriers. More specifically, for technical reasons in what follows it is slightly more convenient to talk about subsets of  $[-\infty, \infty] \times \mathbb{R}_+$  instead of subsets of  $\mathbb{R} \times \mathbb{R}_+$ , giving the following definition.

**Definition 4.1.** Let  $X$  be a topological space. A *downwards barrier* is a set  $\mathcal{R} \subseteq [-\infty, \infty] \times X$  such that  $\{-\infty\} \times X \subseteq \mathcal{R}$  and

$$(y, t) \in \mathcal{R} \text{ and } y' \leq y \text{ implies } (y', t) \in \mathcal{R}$$

Clearly, in Lemma 3.7, instead of talking about  $\mathcal{R} \subseteq \mathbb{R} \times \mathbb{R}_+$ , we could have talked about  $\mathcal{R} \cup (\{-\infty\} \times \mathbb{R}_+) \subseteq [-\infty, \infty] \times \mathbb{R}_+$  without anything really changing, and likewise for  $\hat{\mathcal{R}}$ .

The reader will easily verify the following lemma.

**Lemma 4.2.** *Let  $X$  be a topological space. There is a bijection between the set of all upper semicontinuous functions  $\beta : X \rightarrow [-\infty, \infty]$  and the set of all closed downwards barriers  $\mathcal{R} \subseteq [-\infty, \infty] \times X$  (where closure is to be understood in the product topology). This bijection maps any upper semicontinuous function  $\beta$  to the barrier  $\mathcal{R}$  which is the hypograph of  $\beta$*

$$\mathcal{R} := \{(y, x) : y \leq \beta(x)\} ,$$

while the inverse maps a barrier  $\mathcal{R}$  to the function  $\beta$  given by

$$\beta(x) := \sup\{y : (y, x) \in \mathcal{R}\} .$$

What we will show now, on the way to proving Corollary 1.1 is that the first hitting time after 0 of any downwards barrier by Brownian motion is a.s. equal to the first hitting time after 0 of the closure of that barrier. This serves to both resolve the question whether the times in Lemma 3.7 are stopping times and to show that  $\tau = \hat{\tau}$  a.s.

Let us assume for the rest of this section that  $B$  is actually a Brownian motion started in 0.

**Lemma 4.3.** *Let  $\mathcal{R}$  be a downwards barrier in  $[-\infty, \infty] \times \mathbb{R}_+$ . Let  $\bar{\mathcal{R}}$  be the closure of  $\mathcal{R}$  (in the product topology of the usual topologies on  $[-\infty, \infty]$  and  $\mathbb{R}_+$ ). Define*

$$\tau(\omega) := \inf\{t > 0 : (B_t(\omega), t) \in \mathcal{R}\}$$

$$\bar{\tau}(\omega) := \inf\{t > 0 : (B_t(\omega), t) \in \bar{\mathcal{R}}\} .$$



Then  $\tau = \bar{\tau}$  a.s.

*Proof.* As  $\bar{\mathcal{R}} \supseteq \mathcal{R}$  we clearly have  $\bar{\tau}(\omega) \leq \tau(\omega)$  for all  $\omega \in \Omega$ . Define

$$\bar{\tau}_\varepsilon(\omega) := \inf\{t > 0 : (B_t(\omega) + \varepsilon \cdot A(t), t) \in \bar{\mathcal{R}}\},$$

where  $A(t) := \frac{t}{1+t}$  is a bounded, strictly increasing function. Using just that  $\bar{\mathcal{R}}$  is the closure of  $\mathcal{R}$  one proves by elementary methods that  $\tau(\omega) \leq \bar{\tau}_\varepsilon(\omega)$  for all  $\omega \in \Omega$  and any  $\varepsilon > 0$ . Because  $A(t) = \int_0^t (1+s)^{-2} ds$  is the integral from 0 to  $t$  of a square integrable function we can apply Girsanov's theorem (see e.g. [34, Theorem 38.5]) to see that  $\bar{\tau}_{1/n}$  converges to  $\bar{\tau}$  in distribution as  $n \rightarrow \infty$ .

As  $(\bar{\tau}_{1/n})_n$  is a decreasing sequence bounded below by  $\bar{\tau}$  we get that convergence holds almost surely.  $\square$

The following is a particular case of [21, Corollary 2.3] (which in turn relies on arguments given in [35, 30]). Note that this lemma is purely a statement about barrier-type stopping times and is not directly connected to the optimization problem under consideration.

**Lemma 4.4** (Uniqueness of Barrier-type solutions). *Assume that  $(Y_t)_{t \geq 0}$  is a measurable,  $(\mathcal{F}_t^0)_{t \geq 0}$ -adapted process and that the process  $Z$  defined through  $Z_t := Y_t(B)$  has a.s. continuous paths. Let  $\mathcal{R}_1, \mathcal{R}_2 \subseteq [-\infty, \infty] \times \mathbb{R}_+$  be closed downwards barriers such that for*

$$\tau_i(\omega) := \inf\{t > 0 : (Z_t(\omega), t) \in \mathcal{R}_i\}$$

we have  $\tau_1 \sim \tau_2$ . Then  $\tau_1 = \tau_2$  a.s.

*Proof.* Is to be found in [21, Corollary 2.3].  $\square$

We now have the necessary prerequisites to use our main results in showing that the first optimization problem in the Appetizer admits a (unique) barrier-type solution.

*Proof of Corollary 1.1.* The strategy is as follows: We choose a cost function and leverage Theorem 3.1 to show that an optimizer exists, the Monotonicity Principle in the form of Theorem 3.6 and Lemma 3.7 will – with some help from Lemma 4.3 – show that any optimizer must be the hitting time of a barrier. Lemma 4.4 shows that any two barrier-type solutions must be equal.

We now provide the details. Start with a cost function  $c(\omega, t) := -\omega(t)A(t)$  for a strictly monotone function  $A : \mathbb{R}_+ \rightarrow \mathbb{R}$  which satisfies  $|A(t)| \leq K(1+t^p)$  and assume that  $\mu$  has moment of order  $\frac{1}{2}+p+\varepsilon$  for some  $\varepsilon > 0$ . To prove that a barrier-type solution exists when  $\mu$  has first moment, choose a bounded strictly increasing  $A$  and  $p = 0$ ,  $\varepsilon = \frac{1}{2}$  in this step. (These assumptions guarantee in particular that the optimization problems considered below have a finite value.) Clearly the problem (OPTSTOP) for  $c$  corresponds to (OPTSTOP) $^{\psi(B_t, t)}$  for  $\psi(B_t, t) = B_t A(t)$  (i.e.  $\psi$  takes the role of  $-c$  such that the minimal/maximal values agree up to a change of sign). We will deal with the case where  $\psi(B_t, t) = \phi(B_t)$  at the end of this proof.

We now check that the conditions in Version B of Theorem 3.1 are satisfied. We also need to check that Assumption 2 holds. Here we need the assumption that  $\mu$  has moment of order  $\frac{1}{2}+p+\varepsilon$ , as well as the Hölder and Burkholder-Davis-Gundy inequalities. The latter specialized to Brownian motion state that for all  $q > 0$  there are positive constants  $K_0$  and  $K_1$  such that for any stopping time  $\tau$  we have  $K_0 \mathbb{E}[\tau^{q/2}] \leq \mathbb{E}[|B|_\tau^*]^q \leq K_1 \mathbb{E}[\tau^{q/2}]$  (where  $|B|_\tau^* = \sup_{s \leq \tau} |B_s|$ ). With these in hand a straightforward calculation allows us to bound  $B_\tau A(\tau)$  in the  $L^{1+\delta}$ -norm for some  $\delta > 0$ , independently of the stopping time  $\tau \sim \mu$ .

This shows both that the uniform integrability condition in Version B of Theorem 3.1 is satisfied and that Assumption 2.4 is satisfied.

On  $C(\mathbb{R}_+)$  we may choose the (Polish) topology of uniform convergence on compacts. For the topology on  $\mathbb{R}_+$  we start with the usual topology and turn  $A$  into a continuous function (if it wasn't), by making use of the fact that any measurable function from a Polish space to a second countable space may be turned into a continuous function by passing to a larger Polish topology (with the same Borel sets) on the domain. (This can be found for example in [27, Theorem 13.11].)

In the statement of Corollary 1.1 we did not require that the probability space  $(\Omega, \mathcal{G}, (\mathcal{G}_t)_{t \geq 0}, \mathbb{P})$  satisfy Assumption 2.2. To remedy this we can enlarge the probability space by setting  $\tilde{\Omega} := \Omega \times [0, 1]$ ,  $\tilde{\mathcal{G}}_t := \mathcal{G}_t \otimes \mathcal{B}([0, 1])$  and  $\tilde{\mathbb{P}} := \mathbb{P} \otimes \mathcal{L}$ , where  $\mathcal{L}$  is Lebesgue measure on  $[0, 1]$ . On this space we consider the Brownian motion  $\tilde{B}_t(\omega, x) := B_t(\omega)$ . Theorem 3.1 now gives us an optimal stopping time  $\tilde{\tau}$  on the enlarged probability space. If we can show that this stopping time is in fact the hitting time of a barrier, then it follows that  $\tilde{\tau} = \tau \circ ((\omega, x) \mapsto \omega)$  for a stopping time  $\tau$  which is defined as the hitting time of the Brownian motion  $B$  of the same barrier. As there are *more* stopping times on  $(\tilde{\Omega}, \tilde{\mathcal{G}}, (\tilde{\mathcal{G}}_t)_{t \geq 0})$  than on  $(\Omega, \mathcal{G}, (\mathcal{G}_t)_{t \geq 0})$  in the sense that any stopping time  $\tau'$  on  $(\Omega, \mathcal{G}, (\mathcal{G}_t)_{t \geq 0})$  induces a stopping time  $\tilde{\tau}' := \tau' \circ ((\omega, x) \mapsto \omega)$  on  $(\tilde{\Omega}, \tilde{\mathcal{G}}, (\tilde{\mathcal{G}}_t)_{t \geq 0})$  we conclude that  $\tau$  must also be optimal among the stopping times on  $(\Omega, \mathcal{G}, (\mathcal{G}_t)_{t \geq 0})$ . With this out of the way, let us refer to our Brownian motion by  $B$ , to the optimal stopping time by  $\tau$  and to our filtered probability space by  $(\Omega, \mathcal{G}, (\mathcal{G}_t)_{t \geq 0}, \mathbb{P})$  irrespective of whether this is the original process and space we started with, or an enlarged one.

Choosing  $p_0 := \frac{1}{2} + p + \varepsilon$  in Assumption 2.5 we apply Theorem 3.6 to obtain a set  $\Gamma$  on which  $(B, \tau)$  is concentrated under  $\mathbb{P}$  and for which (3.4) holds. As  $\mu$  is concentrated on  $(0, \infty)$ , we may assume that  $\Gamma \cap (C(\mathbb{R}_+) \times \{0\}) = \emptyset$ . Next we want to show that Lemma 3.7 applies with  $Y_t(\omega) = \omega(t)$ .

Translating (3.5) to our situation, we want to prove that  $\omega(t) < \eta(t)$  implies

$$-\omega(t)A(t) - \mathbb{E}[(\eta(t) + \tilde{B}_\sigma)A(\sigma)] < -\eta(t)A(t) - \mathbb{E}[(\omega(t) + \tilde{B}_\sigma)A(\sigma)] , \quad (4.1)$$

where  $\tilde{B}$  is Brownian motion started in 0 at time  $t$  on  $C([t, \infty))$  and  $\sigma$  is any stopping time thereon with  $\mathbb{W}_0^t(\sigma = t) < 1$ ,  $\mathbb{W}_0^t(\sigma = \infty) = 0$ ,  $\int \sigma^{p_0} d\mathbb{W}_0^t < \infty$ . Again the Burkholder-Davis-Gundy inequality shows that  $\mathbb{E}[\tilde{B}_\sigma A(\sigma)] < \infty$ . So (4.1) turns into

$$\omega(t) \mathbb{E}[A(\sigma) - A(t)] < \eta(t) \mathbb{E}[A(\sigma) - A(t)]$$

which clearly follows from the assumptions. So we know that Lemma 3.7 holds, i.e. using the names from said lemma we have  $\tau \leq \tau \leq \hat{\tau}$   $\mathbb{P}$ -a.s.

$\Gamma \cap (C(\mathbb{R}_+) \times \{0\}) = \emptyset$  implies  $\mathcal{R} \cap (\mathbb{R} \times \{0\}) = \emptyset$  and therefore  $\tau(\omega) = \inf\{t > 0 : (B_t(\omega), t) \in \mathcal{R}\}$ , and likewise for  $\hat{\mathcal{R}}$  and  $\hat{\tau}$ . As  $\bar{\mathcal{R}} = \hat{\mathcal{R}} =: \bar{\mathcal{R}}$  it follows from Lemma 4.3 that  $\tau = \hat{\tau}$  a.s. and that  $\tau$  is of the form claimed in (1.1) with  $\beta(t) := \sup\{y \in \mathbb{R} : (y, t) \in \bar{\mathcal{R}}\}$ . The uniqueness claims follow from Lemma 4.4 and what we have already proven.

We now treat the case where  $\psi(B_t, t) = \phi(B_t)$  with  $\phi''' > 0$ ,  $|\phi(y)| \leq K(1 + |y|^p)$  and  $\mu$  has finite moment of order  $\frac{p}{2} + \varepsilon$  for some  $\varepsilon > 0$ . Most of the proof remains unchanged. Setting  $c(\omega, t) = -\phi(\omega(t))$  we may again use the Burkholder-Davis-Gundy inequalities to show that  $c(B_\tau, \tau)$  is bounded in  $L^{1+\delta}$ -norm, independently of the stopping time  $\tau \sim \mu$ , thereby showing both that Assumption 2.4 is satisfied and that the uniform-integrability condition in Version B of Theorem 3.1 is satisfied.

It remains to show that  $\omega(t) < \eta(t)$  implies  $((\omega, t), (\eta, t)) \in \text{SG}$ .  $\phi''' > 0$  implies that the map  $y \mapsto \phi(\eta(t) + y) - \phi(\omega(t) + y)$  is strictly convex. By the strict Jensen inequality  $\mathbb{E}[\phi(\eta(t) + \tilde{B}_\sigma) - \phi(\omega(t) + \tilde{B}_\sigma)] > \phi(\eta(t)) - \phi(\omega(t))$  for any stopping

time  $\sigma$  on  $C([t, \infty))$  which is almost surely finite, satisfies optional stopping and is not almost surely equal to  $t$ . As we may choose  $p_0 := \frac{t}{2} + \varepsilon$ , which is greater than 1, we may assume that the  $\sigma$  in the definition of **SG** has finite first moment, which is enough to guarantee that it satisfies optional stopping. Rearranging the last inequality gives (3.2).  $\square$

## 5. EXISTENCE OF AN OPTIMIZER

The proof of existence of solutions to the Problem (**OPTSTOP**) crucially depends on thinking of stopping times as the joint distribution of the process to be stopped and the stopping time. We introduce some concepts to make this precise and give a proof of Theorem 3.1 at the end of this section.

**Lemma 5.1.** *Let  $G : C([t, \infty)) \rightarrow \mathbb{R}$ , and  $s \geq t$ . The function*

$$\omega \mapsto \int G((\omega, s) \odot \theta) d\mathbb{W}_0^s(\theta)$$

*is a version of the conditional expectation  $\mathbb{E}_{\mathbb{W}_\lambda^t}[G|\mathcal{F}_s^t]$  (for any initial distribution  $\lambda$ ). Henceforth, by  $\mathbb{E}[G|\mathcal{F}_s^t]$  we will mean this function.*

*If  $G \in C_b(C([t, \infty)))$ , then  $\mathbb{E}[G|\mathcal{F}_s^t] \in C_b(C([t, \infty)))$ .*

*Proof.* Obvious.  $\square$

Here we use  $C_b(X)$  to denote the set of continuous bounded functions from a topological space  $X$  to  $\mathbb{R}$ . The last sentence of the lemma is of course true for any topology on  $C([t, \infty))$  for which the map  $\omega \mapsto \omega \odot \theta$  is continuous for all  $\theta$ , but we will only need it for the topology of uniform convergence on compacts.<sup>3</sup>

Given spaces  $X$  and  $Y$  we will denote the projection from  $X \times Y$  to  $X$  by  $\text{proj}_X$  (and similarly for  $Y$ ). For a measurable map  $F : X \rightarrow Y$  between measure spaces and a measure  $\nu$  on  $X$  we denote the pushforward of  $\nu$  under  $F$  by  $F_*(\nu) := D \mapsto \nu(F^{-1}[D])$ .

**Definition 5.2 (RST).** The set  $\text{RST}_\kappa^t$  of *randomized stopping times* (of Brownian motion started at time  $t$  with initial distribution  $\kappa$ ) is defined as the set of all subprobability measures  $\xi$  on  $C([t, \infty)) \times [t, \infty)$  such that  $(\text{proj}_{C([t, \infty))})_*(\xi) \leq \mathbb{W}_\kappa^t$  and that

$$\int F(r)(G(\omega) - \mathbb{E}[G|\mathcal{F}_s^t](\omega)) d\xi(\omega, r) = 0 \quad (5.1)$$

for all  $s > t$ , all  $G \in C_b(C([t, \infty)))$  and all  $F \in C_b([t, \infty))$  supported on  $[t, s]$ .

In this definition the topology on  $C([t, \infty))$  is that of uniform convergence on compacts and the topology on  $[t, \infty)$  is the usual topology.

Given a distribution  $\nu$  on  $C([t, \infty))$  we write

$$\text{RST}_\kappa^t(\nu) := \left\{ \xi \in \text{RST}_\kappa^t : (\text{proj}_{[t, \infty)})_*(\xi) = \nu \right\}.$$

We write  $\text{RST}_\kappa^t(\mathcal{P})$  for the set of all  $\xi \in \text{RST}_\kappa^t$  with mass 1 and call these the *finite randomized stopping times*.

In any of these, if we drop the superscript  $t$  then we will mean time  $t = 0$ , while, if we drop the subscript  $\kappa$ , then we mean that the initial distribution  $\kappa = \delta_0$ , i.e. the Brownian motion to be stopped is started deterministically in 0.

To explain the qualifier *finite* it may help to imagine that for a non-finite randomized stopping time of mass  $\alpha < 1$ , the mass  $1 - \alpha$  which is missing is placed along  $C([t, \infty)) \times \{\infty\}$ .

The following Lemma 5.3 from [6] shows that the problem (**OPTSTOP**) is equivalent to the following optimization problem (**OPTSTOP'**) in the sense that a solution

<sup>3</sup>And that choice is rather arbitrary itself, as close reading will reveal.

of one can be translated into a solution of the other and vice versa. This of course also implies that the values of the two problems are equal, thereby showing that the concrete space  $(\Omega, \mathcal{G}, (\mathcal{G}_t)_{t \geq 0}, \mathbb{P})$  has no bearing on this value, as long as Assumptions 1 and 2 are satisfied.

The definition we have given for a randomized stopping time is only the most convenient (for our purposes) of a number of possible equivalent definitions. Although Lemma 5.3 below should provide some intuition on what a randomized stopping time *is*, the reader may still wish to refer to [6, Theorem 3.8] for the other possible ways of defining randomized stopping times. The first step in connecting condition (5.1), which is one of the equivalent conditions listed in said theorem, to the others, is to notice that (5.1) can be rewritten as

$$\int \left( \int F(r) d\xi_\omega(r) \right) (G(\omega) - \mathbb{E}[G|\mathcal{F}_s^t](\omega)) d\mathbb{W}_\kappa^t(\omega) = 0,$$

where  $\xi_\omega$  is a disintegration of  $\xi$  with respect to  $\mathbb{W}_\kappa^t$ . This says that the function  $\omega \mapsto \int F(r) d\xi_\omega(r)$  is orthogonal to  $G - \mathbb{E}[G|\mathcal{F}_s^t]$  for all bounded continuous  $G$ , i.e. that it is a.s.  $\mathcal{F}_s^t$ -measurable whenever  $F$  is supported on  $[t, s]$ . A limit argument then shows that  $\omega \mapsto \xi_\omega([t, s])$  is a.s.  $\mathcal{F}_s^t$ -measurable. Again, we refer the reader to [6] for a more detailed exposition.

**Problem (OPTSTOP').** Among all randomized stopping times  $\xi \in \text{RST}_\lambda(\mu)$  find the minimizer of

$$\xi' \mapsto \int c d\xi'.$$

**Lemma 5.3** ([6, Lemma 3.11]). *Let  $\tau$  be a  $(\mathcal{G}_t)_{t \geq 0}$ -stopping time and consider*

$$\begin{aligned} \Phi : \Omega &\rightarrow C(\mathbb{R}_+) \times [0, \infty] \\ \Phi(\omega) &:= ((B_t(\omega))_{t \geq 0}, \tau(\omega)). \end{aligned}$$

*Then  $\xi := \Phi_*(\mathbb{P})|_{C(\mathbb{R}_+) \times \mathbb{R}_+}$  is a randomized stopping time, i.e.  $\xi \in \text{RST}_\lambda$ , and for any non-negative measurable process  $F : C(\mathbb{R}_+) \times \mathbb{R}_+ \rightarrow \mathbb{R}$  we have*

$$\int F d\xi = \mathbb{E}[(F \cdot \mathbf{1}_{C(\mathbb{R}_+) \times \mathbb{R}_+}) \circ \Phi] = \mathbb{E}[F(B, \tau) \cdot \mathbf{1}_{\mathbb{R}_+}(\tau)]. \quad (5.2)$$

*For any  $\xi \in \text{RST}_\lambda$ , we can find a  $(\mathcal{G}_t)_{t \geq 0}$ -stopping time  $\tau$  such that  $\xi = \Phi_*(\mathbb{P})$  and (5.2) holds.*

*$\xi$  is a finite randomized stopping time iff  $\tau$  is a.s. finite.*

*Proof of Theorem 3.1.* We prove Version B of the theorem. Version A is a special case. We show that Problem (OPTSTOP') has a solution. To this end we show that the set  $\text{RST}_\lambda(\mu)$  is compact (in the weak topology). From the fact that  $c$  is lower semicontinuous and bounded from below in an appropriate sense we then deduce by the Portmanteau theorem that the map

$$\begin{aligned} \hat{c} : \text{RST}_\lambda(\mu) &\rightarrow (-\infty, \infty] \\ \hat{c}(\zeta) &:= \int c d\zeta \end{aligned}$$

is lower semicontinuous and therefore that the infimum  $\inf_{\zeta \in \text{RST}_\lambda(\mu)} \hat{c}(\zeta)$  is attained.

Now for the details. On each of the spaces  $C(\mathbb{R}_+)$  and  $\mathbb{R}_+$  we are dealing with two topologies, one coming from the Definition 5.2 of randomized stopping times (to wit, the topology of uniform convergence on compacts on the space  $C(\mathbb{R}_+)$  and the usual topology on  $\mathbb{R}_+$ ) and one coming from the assumptions in the statement of this theorem. We can equip each of these spaces with the smallest topology which contains the two topologies in question. These are again Polish topologies and they still generate the standard sigma-algebras on the respective spaces. For the remainder of this proof all topological notions are to be understood relative to these topologies. So the topology on  $C(\mathbb{R}_+) \times \mathbb{R}_+$  is the product topology of

these two topologies, and the weak topology on the space of measures on  $C(\mathbb{R}_+) \times \mathbb{R}_+$  is to be understood relative to this product topology. The cost function  $c$  of course remains lower semicontinuous and by Lemma 5.1 the functions  $(\omega, r) \mapsto F(r)(G(\omega) - \mathbb{E}[G|\mathcal{F}_s^0])$  appearing in Definition 5.2 are continuous.

Note that for  $\xi \in \text{RST}_\lambda(\mu)$  as  $\mu$  has mass 1, so must  $\xi$  and  $(\text{proj}_{C(\mathbb{R}_+)}^*)_*(\xi)$ , which together with  $(\text{proj}_{C(\mathbb{R}_+)}^*)_*(\xi) \leq \mathbb{W}_\lambda^0$  implies  $(\text{proj}_{C(\mathbb{R}_+)}^*)_*(\xi) = \mathbb{W}_\lambda^0$ . So we deduce

$$\text{RST}_\lambda(\mu) = \left\{ \xi \in \Pi : \int F(s)(G - \mathbb{E}[G|\mathcal{F}_t^0])(\omega) d\xi(\omega, s) = 0 \quad \forall (t, F, G) \in \star \right\}$$

where

$$\begin{aligned} \xi \in \Pi &\iff (\text{proj}_{C(\mathbb{R}_+)}^*)_*(\xi) = \mathbb{W}_\lambda^0 \text{ and } (\text{proj}_{\mathbb{R}_+}^*)_*(\xi) = \mu \\ (t, F, G) \in \star &\iff t > 0, F: \mathbb{R}_+ \rightarrow \mathbb{R} \text{ is bounded and continuous in the} \\ &\text{usual topologies, and } 0 \text{ outside } [0, t], G: C(\mathbb{R}_+) \rightarrow \mathbb{R} \text{ is} \\ &\text{bounded and continuous as a function from the topology} \\ &\text{of uniform convergence on compacts.} \end{aligned}$$

The set  $\Pi$  is compact by Prokhorov's Theorem and the fact that pushforwards are continuous maps between measure spaces. It remains to show that  $\text{RST}_\lambda(\mu)$  is a nonempty closed subset. It is nonempty because the product measure  $\mathbb{W}_\lambda^0 \otimes \mu \in \text{RST}_\lambda(\mu)$ . It is closed because, as noted,  $(\omega, s) \mapsto F(s)(G - \mathbb{E}[G|\mathcal{F}_t^0])(\omega)$  is continuous for all  $(t, F, G) \in \star$ .

Now we show that  $\hat{c}$  is lower semicontinuous. The functions  $c^N := c \vee -N$  are each bounded from below and lower semicontinuous. By the Portmanteau theorem the maps  $\hat{c}^N := \zeta \mapsto \int c^N d\zeta$  are lower semicontinuous. On  $\text{RST}_\lambda(\mu)$  they converge uniformly to  $\hat{c}$  because

$$\sup_\zeta |\hat{c}(\zeta) - \hat{c}^N(\zeta)| \leq \sup_\zeta \int |c - c^N| d\zeta \leq \sup_{\zeta \in \text{RST}_\lambda(\mu)} \int c_- \cdot 1_{c_- \geq N} d\zeta,$$

which converges to 0 as  $N$  goes to  $\infty$  by the uniform integrability assumption. As a uniform limit of lower semicontinuous functions is again lower semicontinuous we see that  $\hat{c}$  is lower semicontinuous.  $\square$

## 6. GEOMETRY OF THE OPTIMIZER

This section is devoted to the proof of Theorem 3.6. The proof closely mimicks that of Theorem 1.3/Theorem 5.7 in [6]. For the benefit of those readers already familiar with said paper we will first describe the changes required to the proofs there to make them work in our situation and then – for the sake of a more self-contained presentation – indulge in reiterating the main arguments and only citing results from [6] that we can use verbatim.

*Sketch of differences in the proof of Theorem 3.6 relative to [6, Theorem 5.7].* Again the strategy is to show that for a larger set  $\widehat{\text{SG}}^\xi \supseteq \text{SG}$  we can find a set  $\Gamma \subseteq C(\mathbb{R}_+) \times \mathbb{R}_+$  such that  $\widehat{\text{SG}}^\xi \cap (\Gamma^< \times \Gamma) = \emptyset$ . The definition of  $\widehat{\text{SG}}^\xi$  must of course be adapted analogously to the changes required to the definition of  $\text{SG}$ .

Apart from that the only real changes are to [6, Theorem 5.8]. Whereas previously it was essential that the randomized stopping time  $\xi^{r(\omega, s)}$  is also a valid randomized stopping time of the Markov process in question when started at a different time but the same location  $\omega(s)$ , we now need that  $\xi^{r(\omega, s)}$  will also be a randomized stopping time of our Markov process when started at the same time  $s$  but in a different place. Of course, when we are talking about Brownian motion both are true, but this difference is the reason why in the case of the Skorokhod embedding the right class of processes to generalize the argument to is that of Feller processes while in our setup we don't need our processes to be time-homogeneous but we do need them to be space-homogeneous. That we are able to plant this

“bush”  $\xi^{r(\omega,s)}$  in another location is what guarantees that the measure  $\xi_1^\pi$  defined in the proof of Theorem 5.8 of [6] is again a randomized stopping time.

Whereas in the Skorokhod case the task is to show that the new better randomized stopping time  $\xi^\pi$  *embeds* the same distribution as  $\xi$  we now have to show that the randomized stopping time we construct *has* the same distribution as  $\xi$ . The argument works along the same lines though – instead of using that  $((\omega, s), (\eta, t)) \in \widehat{\text{SG}}^\xi$  implies  $\omega(s) = \eta(t)$  we now use that  $((\omega, s), (\eta, t)) \in \widehat{\text{SG}}^\xi$  implies  $s = t$ .  $\square$

We now present the argument in more detail.

As may be clear by now, what we will show is that if  $\xi \in \text{RST}_\lambda(\mu)$  is a solution of (OPTSTOP'), then there is a measurable,  $(\mathcal{F}_t^0)_{t \geq 0}$ -adapted set  $\Gamma \subseteq C(\mathbb{R}_+) \times \mathbb{R}_+$  such that  $\text{SG} \cap (\Gamma^< \times \Gamma) = \emptyset$ . Using Lemma 5.3 this implies Theorem 3.6.

We need to make some preparations. To align the notation with [6] and to make some technical steps easier it is useful to have another characterization of measurable,  $(\mathcal{F}_t^0)_{t \geq 0}$ -adapted processes and sets. To this end define

**Definition 6.1.**

$$\begin{aligned} S &:= \bigcup_{t \in \mathbb{R}_+} C([0, t]) \times \{t\} \\ r &: C(\mathbb{R}_+) \times \mathbb{R}_+ \rightarrow S \\ r(\omega, t) &:= (\omega_{|[0, t]}, t) \end{aligned}$$

$r$  has many right inverses. A simple one is

$$\begin{aligned} r' &: S \rightarrow C(\mathbb{R}_+) \times \mathbb{R}_+ \\ r'(f, s) &:= \left( t \mapsto \begin{cases} f(t) & \text{for } t \leq s \\ f(s) & \text{for } t > s \end{cases}, s \right). \end{aligned}$$

We endow  $S$  with the sigma algebra generated by  $r'$ .

[6, Theorem 3.2], which is a direct consequence of [15, Theorem IV. 97], asserts that a process  $X$  is measurable,  $(\mathcal{F}_t^0)_{t \geq 0}$ -adapted iff  $X$  factors as  $X = X' \circ r$  for a measurable function  $X' : S \rightarrow \mathbb{R}$ . So a set  $D \subseteq C(\mathbb{R}_+) \times \mathbb{R}_+$  is measurable,  $(\mathcal{F}_t^0)_{t \geq 0}$ -adapted iff  $D = r^{-1}[D']$  for some measurable  $D' \subseteq S$ .

Note that  $r(\omega, t) = r(\omega', t')$  implies  $(\omega, t) \odot \theta = (\omega', t') \odot \theta$  and therefore

$$\text{SG} = (r \times r)^{-1}[\text{SG}']$$

for a set  $\text{SG}' \subseteq S \times S$  which is described by an expression almost identical to that in Definition 3.4. Namely we can overload  $\odot$  to also be the name for the operation whose first operand is an element of  $S$ , such that  $(\omega, t) \odot \theta = r(\omega, t) \odot \theta$  and note that as  $c$  is measurable,  $(\mathcal{F}_t^0)_{t \geq 0}$ -adapted we can write  $c = c' \circ r$  and thus get a cost function  $c'$  which is defined on  $S$ .

Given an optimal  $\xi \in \text{RST}_\lambda(\mu)$  we may therefore rephrase our task as having to find a measurable set  $\Gamma \subseteq S$  such that  $r_*(\xi)$  is concentrated on  $\Gamma$  and that  $\text{SG}' \cap (\Gamma^< \times \Gamma) = \emptyset$ , where  $\Gamma^< := \{(g_{|[0, s]}, s) : (g, t) \in \Gamma, s < t\}$ .

Note that for  $\Gamma \subseteq S$  although  $(r^{-1}[\Gamma])^<$  is not equal to  $r^{-1}[\Gamma^<]$  we still have  $\text{SG} \cap (r^{-1}[\Gamma^<] \times r^{-1}[\Gamma]) = \emptyset$  iff  $\text{SG} \cap ((r^{-1}[\Gamma])^< \times r^{-1}[\Gamma]) = \emptyset$ .

One of the main ingredients of the proof of [6, Theorem 1.3] and of our Theorem 3.6 is a procedure whereby we accumulate many infinitesimal changes to a given randomized stopping time  $\xi$  to build a new stopping time  $\xi^\pi$ . The guiding intuition for the authors is to picture these changes as replacing certain “branches” of the stopping time  $\xi$  by different branches. Some of these branches will actually enter the statement of a somewhat stronger theorem (Theorem 6.8 below), so we

begin by describing these. Our way to get a handle on “branches” – i.e. infinitesimal parts of a randomized stopping time – is to describe them through a disintegration (wrt  $\mathbb{W}_\lambda^0$ ) of the randomized stopping time. We need the following statement from [6] which should also serve to provide more intuition on the nature of randomized stopping times.

**Lemma 6.2.** [6, Theorem 3.8] *Let  $\xi$  be a measure on  $C(\mathbb{R}_+) \times \mathbb{R}_+$ . Then  $\xi \in \text{RST}_\lambda$  iff there is a disintegration  $(\xi_\omega)_{\omega \in C(\mathbb{R}_+)}$  of  $\xi$  wrt  $\mathbb{W}_\lambda^0$  such that  $(\omega, t) \mapsto \xi_\omega([0, t])$  is measurable,  $(\mathcal{F}_t^0)_{t \geq 0}$ -adapted and maps into  $[0, 1]$ .*

Using Lemma 6.2 above let us fix for the rest of this section both  $\xi \in \text{RST}_\lambda(\mu)$  and a disintegration  $(\xi_\omega)_{\omega \in C(\mathbb{R}_+)}$  with the properties above. Both Definition 6.3 below and Theorem 6.8 implicitly depend on this particular disintegration and we emphasize that whenever we write  $\xi_\omega$  in the following we are always referring to the same fixed disintegration with the properties given in Lemma 6.2. Note that the measurability properties of  $(\xi_\omega)_{\omega \in C(\mathbb{R}_+)}$  imply that for any  $I \subseteq [0, s]$  we can determine  $\xi_\omega(I)$  from  $\omega|_{[0, s]}$  alone. For  $(f, s) \in S$  we will again overload notation and use  $\xi_{(f, s)}$  to refer to the measure on  $[0, s]$  which is equal to  $(\xi_\omega)|_{[0, s]}$  for any  $\omega \in C(\mathbb{R}_+)$  such that  $r(\omega, s) = (f, s)$ .

**Definition 6.3** (conditional randomized stopping time). Let  $(f, s) \in S$ . We define a new randomized stopping time  $\xi^{(f, s)} \in \text{RST}^s$  by setting

$$\xi_\omega^{(f, s)} := \begin{cases} \frac{1}{1 - \xi_{(f, s)}([0, s])} (\xi_{(f, s) \odot \omega})|_{(s, \infty)} & \text{for } \xi_{(f, s)}([0, s]) < 1 \\ \delta_s & \text{for } \xi_{(f, s)}([0, s]) = 1 \end{cases} \quad (6.1)$$

$$\int F d\xi^{(f, s)} := \int \int F(\omega, t) d\xi_\omega^{(f, s)}(t) d\mathbb{W}_0^s(\omega)$$

for all bounded measurable  $F : C([s, \infty)) \times [s, \infty) \rightarrow \mathbb{R}$ , i.e.  $(\xi_\omega^{(f, s)})_{\omega \in C([s, \infty))}$  is the disintegration of  $\xi^{(f, s)}$  wrt  $\mathbb{W}_0^s$ .

Here  $\delta_s$  is the Dirac measure concentrated at  $s$ . Really, the definition in the case where  $\xi_{(f, s)}([0, s]) = 1$  is somewhat arbitrary – it’s more a convenience to avoid partially defined functions. What we will use is that  $(1 - \xi_{(f, s)}([0, s]))\xi_\omega^{(f, s)} = (\xi_{(f, s) \odot \omega})|_{(s, \infty)}$ .

**Definition 6.4** (relative Stop-Go pairs). The set  $\text{SG}^\xi$  consists of all  $((f, t), (g, t)) \in S \times S$  (again the times have to match) such that either

$$c'(f, t) + \int c((g, t) \odot \theta, u) d\xi^{(f, t)}(\theta, u) < c'(g, t) + \int c((f, t) \odot \theta, u) d\xi^{(f, t)}(\theta, u) \quad (6.2)$$

or any one of

- (1)  $\xi^{(f, t)}(C(\mathbb{R}_+) \times \mathbb{R}_+) < 1$  or  $\int s^{p_0} d\xi^{(f, t)}(\theta, s) = \infty$
- (2) the integral on the right hand side equals  $\infty$
- (3) either of the integrals is not defined

holds. We also define

$$\widehat{\text{SG}}^\xi := \text{SG}^\xi \cup \{(f, s) \in S : \xi_{(f, s)}([0, s]) = 1\} \times S \quad (6.3)$$

Lemma 6.6 below says that the numbered cases above are exceptional in an appropriate sense and one may consider them a technical detail. Note that when we say  $((f, t), (g, t)) \in \text{SG}^\xi$  we are implicitly saying that  $\xi_{(f, t)}([0, t]) < 1$ .

Note that the sets  $\text{SG}^\xi$  and  $\widehat{\text{SG}}^\xi$  are measurable (in contrast to  $\text{SG}$ , which may be more complicated).

**Definition 6.5.** We call a measurable set  $F \subseteq S$  evanescent if  $r^{-1}[F]$  is evanescent, that is, if  $\mathbb{W}_\lambda^0\left(\text{proj}_{C(\mathbb{R}_+)}[r^{-1}[F]]\right) = 0$ .

**Lemma 6.6.** [6, Lemma 5.2] *Let  $F : C(\mathbb{R}_+) \times \mathbb{R}_+ \rightarrow \mathbb{R}$  be some measurable function for which  $\int F d\xi \in \mathbb{R}$ . Then the following sets are evanescent.*

- $\{(f, s) \in S : \xi^{(f,s)}(C(\mathbb{R}_+) \times \mathbb{R}_+) < 1\}$
- $\{(f, s) \in S : \int F((f, s) \odot \theta, u) d\xi^{(f,s)}(\theta, u) \notin \mathbb{R}\}$

*Proof.* See [6]. □

**Lemma 6.7** ([6, Lemma 5.4]).

$$\text{SG}' \subseteq \widehat{\text{SG}}^\xi$$

*Proof.* Can be found in [6]. Note that they fix  $p_0 = 1$ . □

**Theorem 6.8.** *Assume that  $\xi$  is a solution of (OPTSTOP'). Then there is a measurable set  $\Gamma \subseteq S$  such that  $r_*(\xi)(\Gamma) = 1$  and*

$$\widehat{\text{SG}}^\xi \cap (\Gamma^c \times \Gamma) = \emptyset. \quad (6.4)$$

Our argument follows [6, Theorem 5.7]. We also need the following two auxilliary propositions, which in turn require some definitions.

**Definition 6.9.** Let  $\nu$  be a probability measure on some measure space  $Y$ . The set  $\text{JOIN}_\lambda(\nu)$  is the set of all subprobability measures  $\pi$  on  $(C(\mathbb{R}_+) \times \mathbb{R}_+) \times Y$  such that

$$\begin{aligned} (\text{proj}_Y)_*(\pi) &\leq \nu \quad \text{and} \\ (\text{proj}_{C(\mathbb{R}_+) \times \mathbb{R}_+})_*(\pi|_{C(\mathbb{R}_+) \times \mathbb{R}_+ \times D}) &\in \text{RST}_\lambda \quad \text{for all measurable } D \subseteq Y. \end{aligned}$$

**Proposition 6.10.** *Let  $\xi$  be a solution of (OPTSTOP'). Then  $(r \times \text{Id})_*(\pi)(\text{SG}^\xi) = 0$  for all  $\pi \in \text{JOIN}_\lambda(r_*(\xi))$ .*

Here we use  $\times$  to denote the Cartesian product map, i.e. for sets  $X_i, Y_i$  and functions  $F_i : X_i \rightarrow Y_i$  where  $i \in \{1, 2\}$  the map  $F_1 \times F_2 : X_1 \times X_2 \rightarrow Y_1 \times Y_2$  is given by  $(F_1 \times F_2)(x_1, x_2) = (F_1(x_1), F_2(x_2))$ . Proposition 6.10 is an analogue of [6, Proposition 5.8] and it is where the material changes compared to [6] take place. We will give the proof at the end of this section.

**Proposition 6.11.** [6, Proposition 5.9] *Let  $(Y, \nu)$  be a Polish probability space and let  $E \subseteq S \times Y$  be a measurable set. Then the following are equivalent*

- (1)  $(r \times \text{Id})_*(\pi)(E) = 0$  for all  $\pi \in \text{JOIN}_\lambda(\nu)$
- (2)  $E \subseteq (F \times Y) \cup (S \times N)$  for some evanescent set  $F \subseteq S$  and a measurable set  $N \subseteq Y$  which satisfies  $\nu(N) = 0$ .

Proposition 6.11 is proved in [6] and we will not repeat the proof here.

*Proof of Theorem 6.8.* Using Proposition 6.10 we see that  $(r \times \text{Id})_*(\pi)(\text{SG}^\xi) = 0$  for all  $\pi \in \text{JOIN}_\lambda(r_*(\xi))$ . Plugging this into Proposition 6.11 we find an evanescent set  $F_1 \subseteq S$  and a set  $N \subseteq S$  such that  $r_*(\xi)(N) = 0$  and  $\text{SG}^\xi \subseteq (F_1 \times S) \cup (S \times N)$ . Defining for any Borel set  $E \subseteq S$  the analytic set

$$E^> := \{(g, t) \in S : \exists s < t, (g|_{[0,s]}, s) \in E\}$$

we observe that  $((E^>)^c)^< \subseteq E^c$  and find  $r_*(\xi)(F_1^>) = 0$ .

Setting  $F_2 := \{(f, s) \in S : \xi_{(f,s)}([0, s]) = 1\}$  and arguing on the disintegration  $(\xi_\omega)_{\omega \in C(\mathbb{R}_+)}$  we see that  $r_*(\xi)(F_2^>) = 0$ , so  $r_*(\xi)(F^>) = 0$  for  $F := F_1 \cup F_2$ .



This shows that  $S \setminus (N \cup F^>)$  has full  $r_*(\xi)$ -measure. Let  $\Gamma$  be a Borel subset of that set which also has full  $r_*(\xi)$ -measure.

Then

$$\begin{aligned} \Gamma^< \times \Gamma &\subseteq ((F^>)^c)^< \times N^c \subseteq F^c \times N^c \text{ and} \\ \widehat{S\mathcal{G}}^\xi &\subseteq (F \times S) \cup (S \times N) \end{aligned}$$

which shows  $\widehat{S\mathcal{G}}^\xi \cap (\Gamma^< \times \Gamma) = \emptyset$ .  $\square$

**Lemma 6.12.** *If  $\alpha \in \text{RST}_\lambda$  and  $G : C(\mathbb{R}_+) \times \mathbb{R}_+ \rightarrow [0, 1]$  is measurable,  $(\mathcal{F}_t^0)_{t \geq 0}$ -adapted, then the measure defined by*

$$F \mapsto \int F(\omega, t) G(\omega, t) d\alpha(\omega, t) \quad (6.5)$$

is still in  $\text{RST}_\lambda$ .

*Proof.* We use the criterion in Lemma 6.2. Let  $(\alpha_\omega)_{\omega \in C(\mathbb{R}_+)}$  be a disintegration of  $\alpha$  wrt  $\mathbb{W}_\lambda^0$  for which  $(\omega, t) \mapsto \alpha_\omega([0, t])$  is measurable,  $(\mathcal{F}_t^0)_{t \geq 0}$ -adapted and maps into  $[0, 1]$ . Then  $(\hat{\alpha}_\omega)_\omega$  defined by  $\hat{\alpha}_\omega := F \mapsto \int F(t) G(\omega, t) d\alpha_\omega(t)$  is a disintegration of the measure in (6.5) for which  $(\omega, t) \mapsto \hat{\alpha}_\omega([0, t])$  is measurable,  $(\mathcal{F}_t^0)_{t \geq 0}$ -adapted and maps into  $[0, 1]$ .  $\square$

**Lemma 6.13** (Strong Markov property for RSTs). *Let  $\alpha \in \text{RST}_\lambda$ . Then*

$$\int F(\omega, t) d\alpha(\omega, t) = \int \int F((\omega, t) \odot \tilde{\omega}, t) d\mathbb{W}_0^t(\tilde{\omega}) d\alpha(\omega, t)$$

for all bounded measurable  $F : C(\mathbb{R}_+) \times \mathbb{R}_+ \rightarrow \mathbb{R}$ .

*Proof.* Using integral notation instead of the more conventional  $\mathbb{E}$ , we may write the classical form of the strong markov property as

$$\begin{aligned} \int G(\Theta_{\tau(\omega)}(\omega)) H(\omega) \cdot 1_{\mathbb{R}_+}(\tau(\omega)) d\mathbb{W}_\lambda^0(\omega) = \\ \int \int G(\tilde{\omega}) H(\omega) \cdot 1_{\mathbb{R}_+}(\tau(\omega)) d\mathbb{W}_{\omega(\tau(\omega))}^{\tau(\omega)}(\tilde{\omega}) d\mathbb{W}_\lambda^0(\omega) \end{aligned}$$

for all bounded measurable  $G : C(\mathbb{R}_+) \rightarrow \mathbb{R}$  and all bounded  $\mathcal{F}_\tau^0$ -measurable  $H : C(\mathbb{R}_+) \rightarrow \mathbb{R}$ . Here  $\Theta_t$  is the function which cuts off the initial segment of a path up to time  $t$ . From this a simple monotone class argument shows that

$$\begin{aligned} \int K(\Theta_{\tau(\omega)}(\omega), \omega) \cdot 1_{\mathbb{R}_+}(\tau(\omega)) d\mathbb{W}_\lambda^0(\omega) = \\ \int \int K(\tilde{\omega}, \omega) \cdot 1_{\mathbb{R}_+}(\tau(\omega)) d\mathbb{W}_{\omega(\tau(\omega))}^{\tau(\omega)}(\tilde{\omega}) d\mathbb{W}_\lambda^0(\omega) \end{aligned}$$

for all bounded  $\mathcal{F}_\infty^0 \otimes \mathcal{F}_\tau^0$ -measurable  $K : C(\mathbb{R}_+) \times C(\mathbb{R}_+) \rightarrow \mathbb{R}$ .

We may then choose for  $K(\tilde{\omega}, \omega)$  the function  $F(\eta, \tau(\omega))$  where the path  $\eta$  is created by cutting off the tail of  $\omega$  after time  $\tau(\omega)$  and attaching  $\tilde{\omega}$  in its place. Noting the relationship between  $\mathbb{W}_x^{\tau(\omega)}$  and  $\mathbb{W}_0^{\tau(\omega)}$  we then get

$$\begin{aligned} \int F(\omega, \tau(\omega)) \cdot 1_{\mathbb{R}_+}(\tau(\omega)) d\mathbb{W}_\lambda^0(\omega) = \\ \int \int F((\omega, \tau(\omega)) \odot \tilde{\omega}, \tau(\omega)) \cdot 1_{\mathbb{R}_+}(\tau(\omega)) d\mathbb{W}_0^{\tau(\omega)}(\tilde{\omega}) d\mathbb{W}_\lambda^0(\omega) . \end{aligned}$$

Using Lemma 5.3 with  $\Omega = [0, 1] \times C(\mathbb{R}_+)$  and  $\mathcal{G}_t = \mathcal{B}([0, 1]) \otimes \mathcal{F}_t$  we find a  $(\mathcal{G}_t)_{t \geq 0}$ -stopping time  $\tau$  s.t. we may write  $\alpha$  as

$$\alpha = ((y, \omega) \mapsto (\omega, \tau(y, \omega)))_* (\mathcal{L} \otimes \mathbb{W}_\lambda^0) \upharpoonright_{C(\mathbb{R}_+) \times \mathbb{R}_+}$$

(where  $\mathcal{L}$  is Lebesgue measure on  $[0, 1]$ ). For a fixed  $y \in [0, 1]$ ,  $\omega \mapsto \tau(y, \omega)$  is an  $(\mathcal{F}_t^0)_{t \geq 0}$ -stopping time, so we may apply the previous equation to these stopping

times and integrate over  $y \in [0, 1]$  to get

$$\begin{aligned} \int F(\omega, \tau(y, \omega)) \cdot 1_{\mathbb{R}_+}(\tau(y, \omega)) d(\mathcal{L} \otimes \mathbb{W}_\lambda^0)(y, \omega) = \\ \iint F((\omega, \tau(y, \omega)) \odot \tilde{\omega}, \tau(y, \omega)) \cdot 1_{\mathbb{R}_+}(\tau(y, \omega)) d\mathbb{W}_0^\tau(y, \omega)(\tilde{\omega}) d(\mathcal{L} \otimes \mathbb{W}_\lambda^0)(y, \omega) . \end{aligned}$$

Using the equation for  $\alpha$  we see that this is what we wanted to prove.  $\square$

**Lemma 6.14** (Gardener's Lemma). *Assume that we have  $\xi \in \text{RST}_\lambda(\mathcal{P})$ , a measure  $\alpha$  on  $C(\mathbb{R}_+) \times \mathbb{R}_+$  and two families  $\beta^{(\omega, t)}$ ,  $\gamma^{(\omega, t)}$ , where  $(\omega, t) \in C(\mathbb{R}_+) \times \mathbb{R}_+$ , with  $\beta^{(\omega, t)}, \gamma^{(\omega, t)} \in \text{RST}^t(\mathcal{P})$  such that both maps*

$$\begin{aligned} (\omega, t) &\mapsto \int 1_D((\omega, t) \odot \tilde{\omega}, s) d\beta^{(\omega, t)}(\tilde{\omega}, s) \quad \text{and} \\ (\omega, t) &\mapsto \int 1_D((\omega, t) \odot \tilde{\omega}, s) d\gamma^{(\omega, t)}(\tilde{\omega}, s) \end{aligned}$$

are measurable for all Borel  $D \subseteq C(\mathbb{R}_+) \times \mathbb{R}_+$  and that

$$\xi(D) - \iint 1_D((\omega, t) \odot \tilde{\omega}, s) d\beta^{(\omega, t)}(\tilde{\omega}, s) d\alpha(\omega, t) \geq 0 \quad (6.6)$$

for all Borel  $D \subseteq C(\mathbb{R}_+) \times \mathbb{R}_+$ . Then for  $\hat{\xi}$  defined by

$$\begin{aligned} \int F d\hat{\xi} := \int F d\xi - \iint F((\omega, t) \odot \tilde{\omega}, s) d\beta^{(\omega, t)}(\tilde{\omega}, s) d\alpha(\omega, t) \\ + \iint F((\omega, t) \odot \tilde{\omega}, s) d\gamma^{(\omega, t)}(\tilde{\omega}, s) d\alpha(\omega, t) \end{aligned}$$

for all bounded measurable  $F$  we have  $\hat{\xi} \in \text{RST}_\lambda(\mathcal{P})$ .

*Remark 6.15.* The intuition behind the Gardener's Lemma is that we are replacing certain branches  $\beta^{(\omega, t)}$  of the randomized stopping time  $\xi$  by other branches  $\gamma^{(\omega, t)}$  to obtain a new stopping time  $\hat{\xi}$ . This process happens *along* the measure  $\alpha$ . Note that (6.6) implies that  $\int 1_D((\omega, t) \odot \tilde{\omega}) d\mathbb{W}_0^t(\tilde{\omega}) d\alpha(\omega, t) \leq \mathbb{W}_\lambda^0(D)$  for all Borel  $D \subseteq C(\mathbb{R}_+)$ . The authors like to think of  $\alpha$  as a stopping time and of the maps  $(\omega, t) \mapsto \beta^{(\omega, t)}$  and  $(\omega, t) \mapsto \gamma^{(\omega, t)}$  as adapted (in some sense that would need to be made precise). As these assumptions aren't necessary for the proof of the Gardener's Lemma, they were left out, but it might help the reader's intuition to keep them in mind.

*Proof of Lemma 6.14.* We need to check that the  $\hat{\xi}$  we define is indeed a measure, that  $(\text{proj}_{C(\mathbb{R}_+)})_*(\hat{\xi}) = \mathbb{W}_\lambda^0$  and that (5.1) holds for  $\hat{\xi}$ .

Checking that  $\hat{\xi}$  is a measure is routine – we just note that (6.6) guarantees that  $\hat{\xi}(D) \geq 0$  for all Borel  $D$ .

Let  $G : C(\mathbb{R}_+) \rightarrow \mathbb{R}$  be a bounded measurable function.

$$\begin{aligned} \int G(\omega) d\hat{\xi}(\omega, t) &= \int G(\omega) d\xi(\omega, t) - \iint G((\omega, t) \odot \tilde{\omega}) d\beta^{(\omega, t)}(\tilde{\omega}, s) d\alpha(\omega, t) \\ &\quad + \iint G((\omega, t) \odot \tilde{\omega}) d\gamma^{(\omega, t)}(\tilde{\omega}, s) d\alpha(\omega, t) \\ &= \int G d\mathbb{W}_\lambda^0 - \iint G((\omega, t) \odot \tilde{\omega}) d\mathbb{W}_0^t d\alpha(\omega, t) \\ &\quad + \iint G((\omega, t) \odot \tilde{\omega}) d\mathbb{W}_0^t d\alpha(\omega, t) \\ &= \int G d\mathbb{W}_\lambda^0 \end{aligned}$$

Now let  $F : \mathbb{R}_+ \rightarrow \mathbb{R}$  and  $G : C(\mathbb{R}_+) \rightarrow \mathbb{R}$  be bounded continuous functions, with  $F$  supported on  $[0, r]$ .

$$\begin{aligned} \int F(t)(G - \mathbb{E}[G|\mathcal{F}_r^0])(\omega) d\hat{\xi}(\omega, t) &= \int F(t)(G - \mathbb{E}[G|\mathcal{F}_r^0])(\omega) d\xi(\omega, t) \\ &\quad - \iint F(s)(G - \mathbb{E}[G|\mathcal{F}_r^0])((\omega, t) \odot \tilde{\omega}) d\beta^{(\omega, t)}(\tilde{\omega}, s) d\alpha(\omega, t) \\ &\quad - \iint F(s)(G - \mathbb{E}[G|\mathcal{F}_r^0])((\omega, t) \odot \tilde{\omega}) d\gamma^{(\omega, t)}(\tilde{\omega}, s) d\alpha(\omega, t) \quad (6.7) \end{aligned}$$

The first summand is 0 because  $\xi \in \text{RST}_\lambda(\mathcal{P})$ . Looking at the second summand we expand the definition of  $\mathbb{E}[G|\mathcal{F}_r^0]$ .

$$\begin{aligned} \mathbb{E}[G|\mathcal{F}_r^0](\omega, t) \odot \tilde{\omega} &= \int G((\omega, t) \odot \tilde{\omega}, r) \odot \theta \, d\mathbb{W}_0^r(\theta) \\ &= \int G((\omega, t) \odot ((\tilde{\omega}, r) \odot \theta)) \, d\mathbb{W}_0^r(\theta) \end{aligned}$$

whenever  $t \leq r$ , which is the case for those  $t$  which are relevant in the integrand above, because  $F(s) \neq 0$  implies  $s \leq r$  and moreover  $\beta^{(\omega, t)}$  is concentrated on  $(\tilde{\omega}, s)$  for which  $t \leq s$ .

Setting  $\hat{G}^{(\omega, t)}(\tilde{\omega}) := G((\omega, t) \odot \tilde{\omega})$  and  $\hat{F}^{(\omega, t)} := F_{|\cdot|t, \infty}$  we can write

$$\begin{aligned} \iint F(s)(G - \mathbb{E}[G|\mathcal{F}_r^0])(\omega, t) \odot \tilde{\omega} \, d\beta^{(\omega, t)}(\tilde{\omega}, s) \, d\alpha(\omega, t) = \\ \int \mathbb{1}_{[0, r]}(t) \int \hat{F}^{(\omega, t)}(s)(\hat{G}^{(\omega, t)} - \mathbb{E}[\hat{G}^{(\omega, t)}|\mathcal{F}_r^t])(\tilde{\omega}) \, d\beta^{(\omega, t)}(\tilde{\omega}, s) \, d\alpha(\omega, t) \end{aligned}$$

which is 0 because  $\beta^{(\omega, t)} \in \text{RST}^t(\mathcal{P})$  and therefore

$$\int \hat{F}^{(\omega, t)}(s)(\hat{G}^{(\omega, t)} - \mathbb{E}[\hat{G}^{(\omega, t)}|\mathcal{F}_r^t])(\tilde{\omega}) \, d\beta^{(\omega, t)}(\tilde{\omega}, s) = 0$$

for all  $(\omega, t)$  and  $r \geq t$ . The same argument works for the third summand in (6.7).  $\square$

*Proof of Proposition 6.10.* We prove the contrapositive. Assuming that there exists a  $\pi' \in \text{JOIN}_\lambda(r_*(\xi))$  with  $(r \times \text{Id})_*(\pi')(\text{SG}^\xi) > 0$ , we construct a  $\xi^\pi \in \text{RST}_\lambda(\mu)$  such that  $\int c \, d\xi^\pi < \int c \, d\xi$ .

If  $\pi' \in \text{JOIN}_\lambda(r_*(\xi))$ , then for any two measurable sets  $D_1, D_2 \subseteq S$ , because  $\pi'_{|(C(\mathbb{R}_+) \times \mathbb{R}_+) \times D_2} \in \text{RST}_\lambda$  and by making use of Lemma 6.12 we can deduce that  $(\text{proj}_{C(\mathbb{R}_+) \times \mathbb{R}_+})_* (\pi'_{|(r \times \text{Id})^{-1}[D_1 \times D_2]}) \in \text{RST}_\lambda$ . Using the monotone class theorem this extends to any measurable subset of  $S \times S$  in place of  $D_1 \times D_2$ . So we can set  $\pi := \pi'_{|(r \times \text{Id})^{-1}[\text{SG}^\xi]}$  and know that  $(\text{proj}_{C(\mathbb{R}_+) \times \mathbb{R}_+})_*(\pi) \in \text{RST}_\lambda$  and that  $\pi$  is concentrated on  $\text{SG}^\xi$ .

We will be using a disintegration of  $\pi$  wrt  $r(\xi)$ , which we call  $(\pi_{(g, t)})_{(g, t) \in S}$  and for which we assume that  $\pi_{(g, t)}$  is a subprobability measure for all  $(g, t) \in S$ . It will also be useful to assume that  $\pi_{(g, t)}$  is concentrated on the set  $\{(\omega, s) \in C(\mathbb{R}_+) \times \mathbb{R}_+ : s = t\}$  not just for  $r(\xi)$ -almost all  $(g, t)$  but for all  $(g, t)$ . Again this is no restriction of generality. We will also push  $\pi$  onto  $(C(\mathbb{R}_+) \times \mathbb{R}_+) \times (C(\mathbb{R}_+) \times \mathbb{R}_+)$ , defining a measure  $\bar{\pi}$  via

$$\int F \, d\bar{\pi} := \iint F((\omega, s), ((g, t) \odot \tilde{\eta}, t)) \, d\mathbb{W}_0^t(\tilde{\eta}) \, d\pi((\omega, s), (g, t))$$

for all bounded measurable  $F$ . Observe that by Lemma 6.13 the pushforward of  $\pi$  under projection onto the second coordinate (pair) is  $\xi$  and that a disintegration of  $\bar{\pi}$  wrt to  $\xi$  (again in the second coordinate) is given by  $(\pi_{r(\eta, t)})_{(\eta, t) \in C(\mathbb{R}_+) \times \mathbb{R}_+}$ . Let us name  $(\text{proj}_{C(\mathbb{R}_+) \times \mathbb{R}_+})_*(\pi) =: \zeta \in \text{RST}_\lambda$ . We will now use the Gardener's Lemma to define two modifications  $\xi_0^\pi, \xi_1^\pi$  of  $\xi$  such that  $\xi^\pi := \frac{1}{2}(\xi_0^\pi + \xi_1^\pi)$  is our improved randomized stopping time.

For all bounded measurable  $F : C(\mathbb{R}_+) \times \mathbb{R}_+ \rightarrow \mathbb{R}$  define

$$\begin{aligned} \int F \, d\xi_0^\pi := \int F \, d\xi + \int (1 - \xi_\omega([0, s])) \left( - \int F((\omega, s) \odot \tilde{\omega}, u) \, d\xi^{r(\omega, s)}(\tilde{\omega}, u) \right. \\ \left. + F(\omega, s) \right) \, d\zeta(\omega, s) \end{aligned}$$

$$\begin{aligned} \int F \, d\xi_1^\pi := \int F \, d\xi + \int (1 - \xi_\omega([0, s])) \left( - F(\eta, t) \right. \\ \left. + \int F((\eta, t) \odot \tilde{\omega}, u) \, d\xi^{r(\omega, s)}(\tilde{\omega}, u) \right) \, d\bar{\pi}((\omega, s), (\eta, t)) . \end{aligned}$$

The concatenation on the last line is well-defined  $\bar{\pi}$ -almost everywhere because  $\bar{\pi}$  is concentrated on  $(r \times r)^{-1}[\text{SG}^\xi]$  and so in the integrand above  $s = t$  on a set of full measure.

We need to check that the Gardener's Lemma applies in both cases. First of all observe that the product measure  $\mathbb{W}_0^t \otimes \delta_t$  is in  $\text{RST}^t(\mathcal{P})$  and that Lemma 6.13 implies

$$\int F(\omega, t) d\alpha(\omega, t) = \int \int F((\omega, t) \odot \tilde{\omega}, s) d(\mathbb{W}_0^t \otimes \delta_t)(\tilde{\omega}, s) d\alpha(\omega, t) .$$

for any randomized stopping time  $\alpha$ . So for  $\xi_0^\pi$  the measures  $\gamma^{(\omega, t)}$  are given by  $\mathbb{W}_0^t \otimes \delta_t$  and for  $\xi_1^\pi$  the measures  $\beta^{(\omega, t)}$  are given by  $\mathbb{W}_0^t \otimes \delta_t$ .

For  $\xi_0^\pi$  the measure along which we are replacing branches is given by

$$F \mapsto \int F(\omega, s)(1 - \xi_\omega([0, s])) d\zeta(\omega, s) .$$

The branches  $\beta^{(\omega, s)}$  we remove are  $\xi^{r(\omega, s)}$ . We need to check that

$$\int F d\xi - \int (1 - \xi_\omega([0, s])) \int F((\omega, s) \odot \tilde{\omega}, u) d\xi^{r(\omega, s)}(\tilde{\omega}, u) d\zeta(\omega, s) \geq 0$$

for all positive, bounded, measurable  $F : C(\mathbb{R}_+) \times \mathbb{R}_+ \rightarrow \mathbb{R}$ . Let us calculate.

$$\begin{aligned} \int (1 - \xi_\omega([0, s])) \int F((\omega, s) \odot \tilde{\omega}, u) d\xi^{r(\omega, s)}(\tilde{\omega}, u) d\zeta(\omega, s) &= \\ \iint F((\omega, s) \odot \tilde{\omega}, u) d((\xi_{(\omega, s) \odot \tilde{\omega}})_{\uparrow(s, \infty)})(u) d\mathbb{W}_0^s(\tilde{\omega}) d\zeta(\omega, s) &= \\ \iint F(\omega, u) d((\xi_\omega)_{\uparrow(s, \infty)})(u) d\zeta(\omega, s) \leq \iint F(\omega, u) d(\xi_\omega)(u) d\zeta(\omega, s) &\leq \\ \int \int F(\omega, u) d(\xi_\omega)(u) d\mathbb{W}_\lambda^0(\omega) = \int F(\omega, u) d\xi(\omega, u) \end{aligned}$$

Here we first used the definition of  $\xi^{r(\omega, s)}$  and then Lemma 6.13 and finally that  $(\text{proj}_{C(\mathbb{R}_+)})_*(\zeta) \leq \mathbb{W}_\lambda^0$ .

For  $\xi_1^\pi$  we replace branches along

$$\begin{aligned} F \mapsto \int F(\eta, t)(1 - \xi_\omega([0, s])) d\bar{\pi}((\omega, s), (\eta, t)) \\ = \int F(\eta, t) \int (1 - \xi_\omega([0, s])) d\pi_{r(\eta, t)}(\omega, s) d\xi(\eta, t) . \end{aligned}$$

The calculation above shows that

$$\int F d\xi - \int (1 - \xi_\omega([0, s])) F(\eta, t) d\bar{\pi}((\omega, s), (\eta, t)) \geq 0$$

for all positive, bounded, measurable  $F : C(\mathbb{R}_+) \times \mathbb{R}_+ \rightarrow \mathbb{R}$ . For  $\xi_1^\pi$  the branches  $\gamma^{(\eta, t)}$  that we add are given by

$$F \mapsto \frac{\int (1 - \xi_\omega([0, s])) \int F(\tilde{\omega}, u) d\xi^{r(\omega, s)}(\tilde{\omega}, u) d\pi_{r(\eta, t)}(\omega, s)}{\int (1 - \xi_\omega([0, s])) d\pi_{r(\eta, t)}(\omega, s)}$$

when  $\int (1 - \xi_\omega([0, s])) d\pi_{r(\eta, t)}(\omega, s) > 0$  and  $\delta_t$  otherwise (again, the latter is arbitrary). In the more interesting case  $\gamma^{(\eta, t)}$  is an average over elements of  $\text{RST}^t(\mathcal{P})$  and therefore itself in  $\text{RST}^t(\mathcal{P})$ . Here it is again crucial that for  $\pi_{r(\eta, t)}$ -almost all  $(\omega, s)$  we have  $s = t$ , otherwise we would be averaging randomized stopping times of our process started at unrelated times.

Putting this together we see that  $\xi^\pi := \frac{1}{2}(\xi_0^\pi + \xi_1^\pi)$  is a randomized stopping time and that

$$\begin{aligned} 2 \int F d(\xi^\pi - \xi) &= \int (1 - \xi_\omega([0, s])) \left( F(\omega, s) - \int F((\omega, s) \odot \tilde{\omega}, u) d\xi^{r(\omega, s)}(\tilde{\omega}, u) \right. \\ &\quad \left. - F(\eta, t) + \int F((\eta, t) \odot \tilde{\omega}, u) d\xi^{r(\omega, s)}(\tilde{\omega}, u) \right) d\bar{\pi}((\omega, s), (\eta, t)) \quad (6.8) \end{aligned}$$

for all bounded measurable  $F : C(\mathbb{R}_+) \times \mathbb{R}_+ \rightarrow \mathbb{R}$ . Specializing to  $F(\omega, s) = G(s)$  for  $G : \mathbb{R}_+ \rightarrow \mathbb{R}$  bounded measurable we find that

$$\int G(s) d(\xi - \xi^\pi)(\omega, s) = 0 ,$$

again because for  $\bar{\pi}$ -almost all  $((\omega, s), (\eta, t))$  we have  $s = t$ . This shows that  $\xi^\pi \in \text{RST}_\lambda(\mu)$ .

We now want to extend (6.8) to  $c$ . We first show that (6.8) also holds for  $F : C(\mathbb{R}_+) \times \mathbb{R}_+ \rightarrow \mathbb{R}$  which are measurable and positive and for which  $\int F d\xi < \infty$ . To see this, approximate such an  $F$  from below by bounded measurable functions (for which (6.8) holds) and note that by previous calculations both

$$\begin{aligned} \int (1 - \xi_\omega([0, s])) \int F((\omega, s) \odot \tilde{\omega}, u) d\xi^{r(\omega, s)}(\tilde{\omega}, u) d\bar{\pi}((\omega, s), (\eta, t)) &\leq \int F d\xi < \infty \\ \text{and} \quad \int (1 - \xi_\omega([0, s])) F(\eta, t) d\bar{\pi}((\omega, s), (\eta, t)) &\leq \int F d\xi < \infty . \end{aligned}$$

Looking at positive and negative parts of  $c$  and using Assumption 2.4 to see that  $\int c_- d(\xi^\pi - \xi) \in \mathbb{R}$  we get that indeed (6.8) holds for  $F = c$ .

Now we will argue that the integrand in the right hand side of (6.8) is negative  $\bar{\pi}$ -almost everywhere. This will conclude the proof.

By inserting an  $r$  in appropriate places we can read off from Definition 6.4 what it means that  $\bar{\pi}$  is concentrated on  $(r \times r)^{-1}[\text{SG}^\xi]$ . In the course of verifying that (6.8) applies to  $c$  we already saw that cases 2 and 3 in Definition 6.4 can only occur on a set of  $\bar{\pi}$ -measure 0. Lemma 6.6 excludes case 1  $\bar{\pi}$ -almost everywhere. This means that (6.2) holds  $\bar{\pi}$ -almost everywhere – or more correctly, that for  $\bar{\pi}$ -a.a.  $((\omega, s), (\eta, t))$  we have  $s = t$  and

$$\begin{aligned} c(\omega, s) - \int c((\omega, s) \odot \tilde{\omega}, u) d\xi^{r(\omega, s)}(\tilde{\omega}, u) \\ - c(\eta, t) + \int c((\eta, t) \odot \tilde{\omega}, u) d\xi^{r(\omega, s)}(\tilde{\omega}, u) < 0 , \end{aligned} \quad (6.9)$$

completing the proof.  $\square$

## 7. VARIATIONS ON THE THEME

We proceed to prove Corollary 1.2. This is closely modelled on the treatment of the Azema-Yor embedding in [6, Theorem 6.5]. As is the case there we run into a technical obstacle, though one which can be overcome by combining the ideas we have already seen in slightly new ways.

To demonstrate the problem let us begin an attempt to prove Corollary 1.2. Again, we read off  $c(\omega, t) = -\omega^*(t)$ , with  $\omega^*(t) = \sup_{s \leq t} \omega(s)$ . We may use Theorem 3.1 to find a solution  $\tau$  of the problem  $(\text{OPTSTOP}^{B_t^*})$  and we use Theorem 3.6 to find a set  $\Gamma \subseteq C(\mathbb{R}_+) \times \mathbb{R}_+$  for which  $\mathbb{P}[(B, \tau) \in \Gamma] = 1$  and  $\text{SG} \cap (\Gamma^< \times \Gamma) = \emptyset$ . Now we would like to apply Lemma 3.7 with  $Y_t(\omega) = \omega(t) - \omega^*(t)$ , as proposed by Corollary 1.2, so we want to prove that  $\omega(t) - \omega^*(t) < \eta(t) - \eta^*(t)$  implies  $((\omega, t), (\eta, t)) \in \text{SG}$ . Let us do the calculations. We start with an  $(\mathcal{F}_s^t)_{s \geq t}$ -stopping time  $\sigma$ , for which  $\mathbb{W}_0^t(\sigma = t) < 1$ ,  $\mathbb{W}_0^t(\sigma = \infty) = 0$  and for which both sides in (3.2) are defined and finite. To reduce clutter, let us name  $(\omega \mapsto (\omega, \sigma(\omega)))_*(\mathbb{W}_0^t) =: \alpha$ , so that (3.2), which we want to prove, reads

$$-\omega^*(t) + \int ((\omega, t) \odot \theta)^*(s) d\alpha(\theta, s) < -\eta^*(t) + \int ((\eta, t) \odot \theta)^*(s) d\alpha(\theta, s) \quad (7.1)$$

We may rewrite the left hand side as

$$\begin{aligned} \int (\omega^*(t) \vee (\omega(t) + \theta^*(s))) - \omega^*(t) d\alpha(\theta, s) = \\ \int 0 \vee (\omega(t) - \omega^*(t) + \theta^*(s)) d\alpha(\theta, s) . \end{aligned}$$

For the right hand side we get the same expression with  $\omega$  replaced by  $\eta$ . Looking at the integrands we see that if

$$0 < \eta(t) - \eta^*(t) + \theta^*(s) \tag{7.2}$$

then

$$0 \vee (\omega(t) - \omega^*(t) + \theta^*(s)) < 0 \vee (\eta(t) - \eta^*(t) + \theta^*(s)) ,$$

but in the other case

$$0 \vee (\omega(t) - \omega^*(t) + \theta^*(s)) = 0 = 0 \vee (\eta(t) - \eta^*(t) + \theta^*(s)) .$$

So if (7.2) holds for  $(\theta, s)$  from a set of positive  $\alpha$ -measure, then we proved what we wanted to prove. But if  $\theta^*(s) \leq \eta^*(t) - \eta(t)$  for  $\alpha$ -a.a.  $(\theta, s)$  then in (3.2) we have equality instead of strict inequality.

As in [6, Theorem 6.5], one way of getting around this is to introduce a secondary optimization criterion. One way to explain the idea of secondary optimization is to think about what happens if, instead of considering a cost function  $c : C(\mathbb{R}_+) \times \mathbb{R}_+ \rightarrow \mathbb{R}$  we consider a cost function  $c : C(\mathbb{R}_+) \times \mathbb{R}_+ \rightarrow \mathbb{R}^n$ . Of course, to be able to talk about optimization, we will then want to have an order on  $\mathbb{R}^n$ . For reasons that should become clear soon, we decide on the lexicographical order. For the case  $n = 2$  that we are actually interested in for Corollary 1.2 this means that

$$(x_1, x_2) \leq (y_1, y_2) \iff x_1 < y_1 \text{ or } (x_1 = y_1 \text{ and } x_2 \leq y_2) .$$

We claim that Theorem 3.6 is still true if we replace  $c : C(\mathbb{R}_+) \times \mathbb{R}_+ \rightarrow \mathbb{R}$  by  $c : C(\mathbb{R}_+) \times \mathbb{R}_+ \rightarrow \mathbb{R}^n$  and read any symbol  $\leq$  which appears between vectors in  $\mathbb{R}^n$  as the lexicographic order on  $\mathbb{R}^n$  (and of course likewise for all the derived symbols and notions  $<$ ,  $\geq$ ,  $>$ ,  $\inf$ , etc.). Moreover, the arguments are exactly the same. Indeed the crucial part that may deserve some mention is at the end of the proof of Proposition 6.10, where we use the assumption that (6.9) holds on a set of positive measure, i.e. that the integrand is  $< 0$  on a set of positive measure, and that the integrand is 0 outside that set, to conclude that the integral itself must be  $< 0$ . This implication is also true for the lexicographical order on  $\mathbb{R}^n$ . One more detail to be aware of is that integrating functions which map into  $\mathbb{R}^2$  may give results of the form  $(\infty, x)$ ,  $(x, -\infty)$ , etc. In the case of a one-dimensional cost function we excluded such problems by making Assumption 2.4. What we really want in the proof of Proposition 6.10 is that  $\int c d\xi$  and  $\int c d\xi^\pi$  should be finite. Clearly a sufficient condition to guarantee this is to replace Assumption 2.4 by

$$(4') \mathbb{E}[c(B, \tau)] \in \mathbb{R}^n \text{ for all stopping times } \tau \sim \mu .$$

This is not the most general version possible but it will suffice for our purposes.

To get an existence result we may assume that  $c = (c_1, c_2)$  is component-wise lower semicontinuous and that both  $c_1$  and  $c_2$  are bounded below (in either of the ways described in the two versions of Theorem 3.1). Note that – because we are talking about the lexicographic order –  $\xi \in \text{RST}_\lambda(\mu)$  is a solution of (OPTSTOP') for  $c$  iff  $\xi$  is a solution of (OPTSTOP') for  $c_1$  and among all such solutions  $\xi^t$ ,  $\xi$  minimizes  $\int c_2 d\xi^t$ . By Theorem 3.1 in the form that we have already proved the set of solutions of (OPTSTOP') for  $c_1$  is non-empty. It is also a closed subset of a compact set and therefore itself compact. This allows us to reiterate the argument that we used in the proof of Theorem 3.1 to find inside this set a minimizer of  $\xi^t \mapsto \int c_2 d\xi^t$ . This minimizer is the solution of (OPTSTOP') for  $c$ .

With this in hand we may pick up our

*Proof of Corollary 1.2.* The same arguments as in the proof of Corollary 1.1 apply, so we may assume that our probability space satisfies Assumption 2.2. We start with a cost function  $c(\omega, t) := (c_1(\omega, t), c_2(\omega, t)) := (-\omega^*(t), (\omega^*(t) - \omega(t))^3)$ .

$\|c_1(B, \tau)\|_{L^3} \leq \| |B|_\tau^* \|_{L^3} \leq K_1 \|\tau\|_{L^{3/2}}^{1/2}$ , by the Burkholder-Davis-Gundy inequalities, so  $(c_1)_-$  satisfies the uniform integrability condition and  $\mathbb{E}[c(B, \tau)]$  is finite for all stopping times  $\tau \sim \mu$ .  $c_2 \geq 0$  and by the Burkholder-Davis-Gundy inequalities  $\mathbb{E}[c_2(B, \tau)] \leq \mathbb{E}[(B^*(\tau))^3] \leq K_1 \mathbb{E}[\tau^{3/2}] = K_1 \int t^{3/2} d\mu(t)$  for some constant  $K_1$ . The last term is finite by assumption.

By our discussion in the preceding paragraphs we find a solution  $\tau$  of (OPTSTOP) for  $c$  and a measurable,  $(\mathcal{F}_t^0)_{t \geq 0}$ -adapted set  $\Gamma \subseteq C(\mathbb{R}_+) \times \mathbb{R}_+$ , for which  $\mathbb{P}[(B, \tau) \in \Gamma] = 1$  and  $\text{SG} \cap (\Gamma^c \times \Gamma) = \emptyset$ , where now  $((\omega, t), (\eta, t)) \in \text{SG}$  iff for all  $(\mathcal{F}_s^t)_{s \geq t}$ -stopping times  $\sigma$  for which  $\mathbb{W}_0^t(\sigma = t) < 1$ ,  $\mathbb{W}_0^t(\sigma = \infty) = 0$ ,  $\int \sigma^{3/2} d\mathbb{W}_0^t < \infty$ , setting  $\alpha := (\omega \mapsto (\omega, \sigma(\omega)))_*(\mathbb{W}_0^t)$  we have that either equation (7.1) holds or

$$-\omega^*(t) + \int ((\omega, t) \odot \theta)^*(s) d\alpha(\theta, s) = -\eta^*(t) + \int ((\eta, t) \odot \theta)^*(s) d\alpha(\theta, s) \quad (7.3)$$

and

$$c_2(\omega, t) - \int c_2((\omega, t) \odot \theta, s) d\alpha(\theta, s) < c_2(\eta, t) - \int c_2((\eta, t) \odot \theta, s) d\alpha(\theta, s). \quad (7.4)$$

Now we want to apply Lemma 3.7, so we want to show that  $\omega(t) - \omega^*(t) < \eta(t) - \eta^*(t)$  implies  $((\omega, t), (\eta, t)) \in \text{SG}$ . We already dealt with the case where  $\alpha$  is such that (7.2) holds on a set of positive  $\alpha$ -measure. We now deal with the other case, so we have

$$\theta^*(s) \leq \eta^*(t) - \eta(t) < \omega^*(t) - \omega(t) \quad (7.5)$$

for  $\alpha$ -a.a.  $(\theta, s)$  and we know that (7.3) holds. We show that (7.4) holds. Because of (7.5),  $((\omega, t) \odot \theta)^*(s) = \omega^*(t)$ , and so  $c_2((\omega, t) \odot \theta, s) = (\omega^*(t) - \omega(t) - \theta(s))^3$ . We calculate the left hand side of (7.4).

$$\begin{aligned} & \int (\omega^*(t) - \omega(t))^3 - (\omega^*(t) - \omega(t) - \theta(s))^3 d\alpha(\theta, s) = \\ & \int 3(\omega^*(t) - \omega(t))^2 \theta(s) - 3(\omega^*(t) - \omega(t))(\theta(s))^2 + (\theta(s))^3 d\alpha(\theta, s) = \\ & (\omega(t) - \omega^*(t))3 \int (\theta(s))^2 d\alpha(\theta, s) + \int (\theta(s))^3 d\alpha(\theta, s) \end{aligned}$$

Here the Burkholder-Davis-Gundy inequalities show that both  $\int (\theta(s))^3 d\alpha(\theta, s)$  and  $\int (\theta(s))^2 d\alpha(\theta, s)$  are finite so that we may split the integral and they also show that  $\{\tilde{B}_{\sigma \wedge T} : T \geq t\}$  is uniformly integrable so that by the optional stopping theorem  $\int \theta(s) d\alpha(\theta, s) = 0$ . ( $\tilde{B}$  is again Brownian motion started in 0 at time  $t$  on  $C([t, \infty))$ .)

For the right hand side of (7.4) we get the same expression with  $\omega$  replaced by  $\eta$ . This concludes the proof that  $\omega(t) - \omega^*(t) < \eta(t) - \eta^*(t)$  implies  $((\omega, t), (\eta, t)) \in \text{SG}$  and Lemma 3.7 gives us barriers  $\bar{\mathcal{R}}, \hat{\mathcal{R}}$  such that for their hitting times  $\bar{\tau}, \hat{\tau}$  by  $B_t - B_t^*$  we have  $\bar{\tau} \leq \tau \leq \hat{\tau}$  a.s.

Again we want to show that  $\bar{\tau} = \hat{\tau}$  a.s. and that they are actually stopping times. Again we do so by showing that they are both a.s. equal to the hitting time of the closure of the respective barrier. If  $\bar{\mathcal{R}} \cap (\{0\} \times \mathbb{R}_+) = \emptyset$  then this works in exactly the same way as in Lemma 4.3. (This time we define  $\bar{\tau}_\varepsilon := \inf\{t > 0 : (B_t^\varepsilon(\omega) - (B^\varepsilon)_t^*(\omega), t) \in \bar{\mathcal{R}}\}$  where  $B_t^\varepsilon(\omega) := B_t(\omega) + A(t)\varepsilon$ .) If  $\bar{\mathcal{R}} \cap (\{0\} \times \mathbb{R}_+) \neq \emptyset$  then  $(B_t^\varepsilon(\omega) - (B^\varepsilon)_t^*(\omega), t) \in \bar{\mathcal{R}}$  and  $t > 0$  need not imply  $B_t(\omega) - B_t^*(\omega) < B_t^\varepsilon(\omega) - (B^\varepsilon)_t^*(\omega)$ , which is essential for the topological argument showing that the hitting time of  $\bar{\mathcal{R}}$  is less than or equal  $\bar{\tau}_\varepsilon$ . But if  $\hat{\mathcal{R}} \cap (\{0\} \times \mathbb{R}_+) = \bar{\mathcal{R}} \cap (\{0\} \times \mathbb{R}_+) \neq \emptyset$ , then  $\bar{\tau}$  and  $\hat{\tau}$  are both almost surely  $\leq T$  where  $T := \inf\{t > 0 : (0, t) \in \hat{\mathcal{R}}\}$ , so in the step where we show that the hitting time of  $\bar{\mathcal{R}}$  is less than  $\bar{\tau}_\varepsilon$  we can argue under the assumption that  $\bar{\tau}_\varepsilon(\omega) < T$ . In this case we do have that  $(B_t^\varepsilon(\omega) - (B^\varepsilon)_t^*(\omega), t) \in \bar{\mathcal{R}}$  and  $t > 0$  implies  $B_t(\omega) - B_t^*(\omega) < B_t^\varepsilon(\omega) - (B^\varepsilon)_t^*(\omega)$ .  $\square$

*Remark 7.1.* We hope that the proofs of Corollary 1.1 and Corollary 1.2 have given the reader some idea of how to apply the main results of this paper to arrive at barrier-type solutions of constrained optimal stopping problems, as depicted in Figure 1.

We would like to conclude by giving a couple of pointers to the interested reader who may want to work through the proofs corresponding to the remaining pictures in Figure 1.

For the problem of minimizing  $\mathbb{E}[B_\tau^*]$ , it may actually happen that the times  $\tau, \hat{\tau}$  from Lemma 3.7 do not coincide. Specifically one has to expect this to happen on a non-negligible set when  $\mathcal{R}$  contains parts of the time axis which  $\hat{\mathcal{R}}$  does not contain. Under these circumstances an optimizer may turn out to be a true randomized stopping time, with a proportion of a path hitting the time axis at a certain point needing to be stopped while the rest continues. In this situation the picture alone does not completely describe the optimal stopping time.

For the problems involving absolute values one needs to make a minor modification in the proof of Proposition 6.10. Specifically one can allow “mirroring” the paths which are “transplanted” using the Gardener’s Lemma. This leads to a slightly different definition of Stop-Go pairs, which is perhaps most easily described by saying that in Figure 2 the green paths which are stopped by  $\sigma$  may be flipped upside-down on either side.

#### REFERENCES

- [1] Alili, L., Patie, P.: On the first crossing times of a Brownian motion and a family of continuous curves. *C. R. Math. Acad. Sci. Paris* **340**(3), 225–228 (2005). DOI 10.1016/j.crma.2004.11.008. URL <http://dx.doi.org/10.1016/j.crma.2004.11.008>
- [2] Alili, L., Patie, P.: Boundary crossing identities for Brownian motion and some nonlinear ODE’s. *Proc. Amer. Math. Soc.* **142**(11), 3811–3824 (2014). DOI 10.1090/S0002-9939-2014-12194-0. URL <http://dx.doi.org/10.1090/S0002-9939-2014-12194-0>
- [3] Anulova, S.V.: On Markov stopping times with a given distribution for a Wiener process. *Theory of Probability & Its Applications* **25**(2), 362–366 (1981). DOI 10.1137/1125045. URL <http://dx.doi.org/10.1137/1125045>
- [4] Avellaneda, M., Zhu, J.: Distance to default. *Risk* **12**(14), 125–129 (2001)
- [5] Bayraktar, E., Miller, C.W.: Distribution-constrained optimal stopping. *ArXiv e-prints* (2016)
- [6] Beiglböck, M., Cox, A.M.G., Huesmann, M.: Optimal transport and Skorokhod embedding. *Invent. Math.* **208**(2), 327–400 (2017). DOI 10.1007/s00222-016-0692-2. URL <http://dx.doi.org/10.1007/s00222-016-0692-2>
- [7] Beiglböck, M., Griessler, C.: An optimality principle with applications in optimal transport. *ArXiv e-prints* (2014)
- [8] Beiglböck, M., Henry-Labordère, P., Penkner, F.: Model-independent bounds for option prices: A mass transport approach. *Finance and Stochastics* **17**(3), 477–501 (2013)
- [9] Beiglböck, M., Juillet, N.: On a problem of optimal transport under marginal martingale constraints. *Ann. Probab.* **44**(1), 42–106 (2016). DOI 10.1214/14-AOP966. URL <http://dx.doi.org/10.1214/14-AOP966>
- [10] Beiglböck, M., Nutz, M., Touzi, N.: Complete duality for martingale optimal transport on the line. *Ann. Probab.*, to appear (2016)
- [11] Breiman, L.: First exit times from a square root boundary. In: *Proc. Fifth Berkeley Sympos. Math. Statist. and Probability* (Berkeley, Calif., 1965/66), Vol. II: Contributions to Probability Theory, Part 2, pp. 9–16. Univ. California Press, Berkeley, Calif. (1967)
- [12] Campi, L., Laachir, I., Martini, C.: Change of numeraire in the two-marginals martingale transport problem. *Finance Stoch.* **21**(2), 471–486 (2017). DOI 10.1007/s00780-016-0322-2. URL <http://dx.doi.org/10.1007/s00780-016-0322-2>
- [13] Chen, X., Cheng, L., Chadam, J., Saunders, D.: Existence and uniqueness of solutions to the inverse boundary crossing problem for diffusions. *Ann. Appl. Probab.* **21**(5), 1663–1693 (2011). DOI 10.1214/10-AAP714. URL <http://dx.doi.org/10.1214/10-AAP714>
- [14] Cox, A.M.G., Källblad, S.: Model-independent bounds for Asian options: A dynamic programming approach. *ArXiv e-prints* (2015)



- [15] Dellacherie, C., Meyer, P.A.: Probabilities and Potential, A, *North-Holland Mathematics Studies*, vol. 29. North-Holland Publishing Co., Amsterdam (1978)
- [16] Dolinsky, Y., Soner, H.M.: Martingale optimal transport and robust hedging in continuous time. *Probab. Theory Relat. Fields* **160**(1-2), 391–427 (2014). DOI 10.1007/s00440-013-0531-y. URL <http://dx.doi.org/10.1007/s00440-013-0531-y>
- [17] Dudley, R.M., Gutmann, S.: Stopping times with given laws pp. 51–58. *Lecture Notes in Math.*, Vol. 581 (1977)
- [18] Galichon, A., Henry-Labordère, P., Touzi, N.: A stochastic control approach to no-arbitrage bounds given marginals, with an application to lookback options. *Ann. Appl. Probab.* **24**(1), 312–336 (2014)
- [19] Gangbo, W., McCann, R.: The geometry of optimal transportation. *Acta Math.* **177**(2), 113–161 (1996)
- [20] Ghoussoub, N., Kim, Y.H., Lim, T.: Structure of optimal martingale transport plans in general dimensions. *ArXiv e-prints* (2016)
- [21] Grass, A.: Uniqueness and stability properties of barrier type Skorokhod embeddings. master thesis, Vienna, available online at <http://mstoch.tuwien.ac.at/grass/> (2016)
- [22] Guo, G., Tan, X., Touzi, N.: On the monotonicity principle of optimal Skorokhod embedding problem. *SIAM J. Control Optim.* **54**(5), 2478–2489 (2016). DOI 10.1137/15M1025268. URL <http://dx.doi.org/10.1137/15M1025268>
- [23] Hirhager, K.: Adapted dependence with applications to financial and actuarial risk management. PhD thesis, TU Vienna (2013)
- [24] Hobson, D.: The Skorokhod embedding problem and model-independent bounds for option prices. In: Paris-Princeton Lectures on Mathematical Finance 2010, *Lecture Notes in Math.*, vol. 2003, pp. 267–318. Springer, Berlin (2011). DOI 10.1007/978-3-642-14660-2\_4. URL [http://dx.doi.org/10.1007/978-3-642-14660-2\\_4](http://dx.doi.org/10.1007/978-3-642-14660-2_4)
- [25] Hobson, D., Neuberger, A.: Robust bounds for forward start options. *Mathematical Finance* **22**(1), 31–56 (2012)
- [26] Jaimungal, S., Kreinin, A., Valov, A.: The generalized Shiryaev problem and Skorokhod embedding. *Theory Probab. Appl.* **58**(3), 493–502 (2014). DOI 10.1137/S0040585X97986734. URL <http://dx.doi.org/10.1137/S0040585X97986734>
- [27] Kechris, A.: Classical Descriptive Set Theory. Graduate Texts in Mathematics. Springer New York (1995)
- [28] Knott, M., Smith, C.S.: On the optimal mapping of distributions. *J. Optim. Theory Appl.* **43**(1), 39–49 (1984). DOI 10.1007/BF00934745. URL <http://dx.doi.org/10.1007/BF00934745>
- [29] Lerche, H.R.: Boundary crossing of Brownian motion, *Lecture Notes in Statistics*, vol. 40. Springer-Verlag, Berlin (1986). DOI 10.1007/978-1-4615-6569-7. URL <http://dx.doi.org/10.1007/978-1-4615-6569-7>. Its relation to the law of the iterated logarithm and to sequential analysis
- [30] Loynes, R.M.: Stopping times on Brownian motion: Some properties of Root’s construction. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **16**, 211–218 (1970)
- [31] Nutz, M., Stebegg, F.: Canonical supermartingale couplings. *ArXiv e-prints* (2016)
- [32] Oblój, J.: The Skorokhod embedding problem and its offspring. *Probab. Surv.* **1**, 321–390 (2004). DOI 10.1214/154957804100000060. URL <http://dx.doi.org/10.1214/154957804100000060>
- [33] Peskir, G.: On integral equations arising in the first-passage problem for Brownian motion. *J. Integral Equations Appl.* **14**(4), 397–423 (2002). DOI 10.1216/jiea/1181074930. URL <http://dx.doi.org/10.1216/jiea/1181074930>
- [34] Rogers, L.C.G., Williams, D.: Diffusions, Markov Processes, and Martingales. Vol. 2. Cambridge Mathematical Library. Cambridge University Press, Cambridge (2000). DOI 10.1017/CBO9781107590120. URL <http://dx.doi.org/10.1017/CBO9781107590120>. Itô calculus, Reprint of the second (1994) edition
- [35] Root, D.H.: The existence of certain stopping times on Brownian motion. *Ann. Math. Statist.* **40**, 715–718 (1969)
- [36] Rüschendorf, L.: Fréchet-bounds and their applications. In: Advances in probability distributions with given marginals (Rome, 1990), *Math. Appl.*, vol. 67, pp. 151–187. Kluwer Acad. Publ., Dordrecht (1991)
- [37] Rüschendorf, L.: Optimal solutions of multivariate coupling problems. *Appl. Math. (Warsaw)* **23**(3), 325–338 (1995)
- [38] Salminen, P.: On the first hitting time and the last exit time for a Brownian motion to/from a moving boundary. *Adv. in Appl. Probab.* **20**(2), 411–426 (1988). DOI 10.2307/1427397. URL <http://dx.doi.org/10.2307/1427397>

- [39] Zaev, D.: On the Monge-Kantorovich problem with additional linear constraints. *Mat. Zametki* **98**(5), 664–683 (2015). DOI 10.4213/mzm10896. URL <http://dx.doi.org/10.4213/mzm10896>