The Root solution to the multi–marginal embedding problem: an optimal stopping and time–reversal approach[*]

Alexander M. G. Cox[†]   Jan Obłój[‡]   Nizar Touzi[§]

February 6, 2017

Abstract

We provide a complete characterisation of the Root solution to the Skorokhod embedding problem (SEP) by means of an optimal stopping formulation. Our methods are purely probabilistic and the analysis relies on a tailored time-reversal argument. This approach allows us to address the long-standing question of a multiple marginals extension of the Root solution of the SEP. Our main result establishes a complete solution to the n–marginal SEP using first hitting times of barrier sets by the time-space process. The barriers are characterised by means of a recursive sequence of optimal stopping problems. Moreover, we prove that our solution enjoys a global optimality property extending the one-marginal Root case. Our results hold for general, one-dimensional, martingale diffusions.

1 Introduction

The Skorokhod embedding problem (SEP) for Brownian motion $(B_t)_{t \geq 0}$ consists of specifying a stopping time σ such that B_σ is distributed according to a given probability measure μ on $\mathbb{R}$. It has been an active field of study in probability since the original paper by Skorokhod (1965), see Obłój (2004) for an account. One of the most natural ideas for a solution is to consider σ as the first hitting time of some *shape* in time–space. This was carried out in an elegant paper of Root (1969). Root showed that for any centred and square integrable distribution μ there exists a *barrier* $\mathcal{R}$, i.e. a subset of $\mathbb{R}_+ \times \mathbb{R}$ such that $(t, x) \in \mathcal{R}$ implies $(s, x) \in \mathcal{R}$ for all $s \geq t$, for which $B_{\sigma_{\mathcal{R}}} \sim \mu$, $\sigma_{\mathcal{R}} = \inf\{t : (t, B_t) \in \mathcal{R}\}$. The barrier is (essentially) unique, as argued by Loynes (1970).

Root's solution enjoys a fundamental optimality property, established by Rost (1976), that it minimises the variance of the stopping time among all solutions to the SEP. More generally, $\mathbb{E}f(\sigma_{\mathcal{R}}) \leq \mathbb{E}f(\sigma)$ for any convex function $f \geq 0$ and any stopping time σ with $B_\sigma \sim B_{\sigma_{\mathcal{R}}}$. This led to a recent revival of interest in this construction in the mathematical finance literature, where optimal solutions to SEP are linked to robust pricing and hedging of derivatives, see Hobson (1998, 2011). More precisely, optimality of the Root solution translates into lower bounds on prices of options written on the realised volatility.

In recent work Cox and Wang (2013b) show that the barrier $\mathcal{R}$ may be written as the unique solution to a Free Boundary Problem (FBP) or, more generally, to a Variational Inequality (VI). This yields directly its representation by means of an optimal stopping problem. This observation was the starting point for our study here. Subsequently, Gassiat et al. (2014) used analytic methods based on the theory of viscosity solutions to extend Root's existence result to the case of general, integrable starting and target measures satisfying the convex ordering condition. Using methods from optimal transport, Beiglböck et al. (2016) have also recently proved the existence and optimality of Root solutions for one-dimensional Feller processes and, under suitable assumptions on the target measure, for Brownian motion in higher dimensions.

The first contribution of our paper is to show that one can obtain the barrier $\mathcal{R}$ directly from the optimal stopping formulation, and to prove the embedding property using purely probabilistic methods. This also allows us to determine a number of interesting properties of $\mathcal{R}$ by means of a time-reversal technique. Our results will hold for a general one-dimensional diffusion.

Beyond the conceptual interest in deriving the Root solution from the optimal stopping formulation, the new perspective enables us to address the long–standing question of extending the Root solution of the Skorokhod embedding problem to the multiple-marginals case, i.e. given a non-decreasing (in convex order) family of n probability measures $(\mu_0, \ldots, \mu_n)$ on $\mathbb{R}$ with finite first moment, and a diffusion X started from the measure μ_0, find stopping times $\sigma_1 \leq \ldots \leq \sigma_n$ such that $X_{\sigma_i} \sim \mu_i$, and $X_{\cdot \wedge \sigma_n}$ is uniformly integrable. Our second contribution, and the main result of the paper, provides a complete characterisation of such a solution to the SEP which extends the Root solution in the sense that it enjoys the following two properties:

- first, the stopping times are defined as hitting times of a sequence of barriers, which are completely characterized by means of a recursive sequence of optimal stopping problems;
- second, similar to the one-marginal case, we prove that our solution of the multiple marginal SEP minimizes the expectation of any non-decreasing convex function of ρ_n among all families of stopping times $\rho_1 \leq \ldots \leq \rho_n$, such that $X_{\rho_i} \sim \mu_i$.

It is well known that solutions to the multiple marginal SEP exist if and only if the measures are in convex order, however finding optimal solutions to the multiple marginal SEP is more difficult. While many classical constructions of solutions to embedding problems can, in special cases, be ordered (see Madan and Yor (2002)), in general the ordering condition is not satisfied except under strong conditions on the measures. The first paper to produce optimal solutions to the multiple marginal SEP was Brown et al. (2001), who extended the single marginal construction of Azéma and Yor (1979) to the case where one intermediate marginal is specified. More recently, Obłój and Spoida (2016) and Henry-Labordère et al. (2016) extended these results to give an optimal construction for an arbitrary sequence of n marginals satisfying a mild technical condition.

There are also a number of papers which make explicit connections between optimal stopping problems and solutions to the SEP, including Peskir (1999), Obłój (2007) and Cox et al. (2008). In these papers, the key observation is that the optimal solution to the SEP can be closely connected to a particular optimal stopping problem; in all these papers, the *same* stopping time gives rise to both the optimal solution

to the SEP, and the optimal solution to a related optimal stopping problem. In this paper, we will see that the key connection is not that the same stopping time solves both the SEP and a related optimal stopping problem, but rather that there is a time-reversed optimal stopping problem which has the same stopping region as the SEP, and moreover, the value function of the optimal stopping problem has a natural interpretation in the SEP. The first paper we are aware of to exploit this connection is McConnell (1991), who works in the setting of the solution of Rost (1971) and Chacon (1985) to the SEP (see also Cox and Wang (2013a); Gassiat et al. (2014)), and uses analytic methods to show that Rost's solution to the SEP has a corresponding optimal stopping interpretation. More recently[1] De Angelis (2015) has provided a probabilistic approach to understanding McConnell's connection, using a careful analysis of the differentiability of the value function to deduce the embedding properties of the SEP; both the papers of McConnell and De Angelis also require some regularity assumptions on the underlying measures in order to establish their results. In contrast, we consider the Root solution to the SEP. As noted above, a purely analytic connection between Root's solutions to the SEP and a related (time-reversed) optimal stopping problem was observed in Cox and Wang (2013b). In this paper, we are not only able to establish the embedding problems based on properties of the related optimal stopping problem, but we are also able to use our methods to prove new results (in this case, the extension to multiple marginal solutions, and characterisation of the corresponding stopping regions), without requiring any assumptions on the measures which we embed (beyond the usual convex ordering condition).

The paper is organized as follows. Section 2 formulates the multiple marginals Skorokhod embedding problem, reviews the Root solution together with the corresponding variational formulation, and states our optimal stopping characterization of the Root barrier. In Section 3, we report the main characterisation of the multiple marginal solution of the SEP, and we derive the corresponding optimality property. The rest of the paper is devoted to the proof of the main results. In Section 4, we introduce some important definitions relating to potentials, state the main technical results, and use these to prove our main result regarding the embedding properties. The connection with optimal stopping is examined in Section 5. Given this preparation, we report the proof of the main result in Section 6 in the case of locally finitely supported measures. This is obtained by means of a time reversal argument. Finally, we complete the proof in the case of general measures in Section 7 by a delicate limiting procedure.

Notation and Standing Assumptions: In the following, we consider a regular, time-homogenous, martingale diffusion taking values on an interval $\mathcal{I}$, defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ satisfying the usual hypotheses. For $(t, x) \in \mathbb{R}_+ \times \mathbb{R}$, we write $\mathbb{E}^{t,x}$ for expectations under the measure for which the diffusion departs from x at time t. We also write $\mathbb{E}^x = \mathbb{E}^{0,x}$. We use both (X_t) and (Y_t) to denote the diffusion process. While X and Y denote the same object, the double notation allows us to distinguish between two interpretations: with a fixed reference time-space domain $\mathbb{R}_+ \times \mathbb{R}$, we think of (X_t) as starting in (t, x) and *running forward in time* and of (Y_t) as starting in (t, x) and *running backwards in time*. For a distribution ν on $\mathbb{R}$, we interpret $\mathbb{E}^\nu[.] = \int \mathbb{E}^x[.]\nu(dx)$.

We suppose that the diffusion coefficient is $\eta(x)$, so $d\langle X \rangle_t = \eta(X_t)^2 dt$, where η is locally Lipschitz, $|\eta(x)|^2 \le C_\eta(1 + |x|^2)$, for some constant C_η, and strictly positive on $\mathcal{I}^\circ$, where we write $\mathcal{I}^\circ = (a_\mathcal{I}, b_\mathcal{I})$, and without loss of generality, assume that $0 \in \mathcal{I}^\circ$; in addition, we use $\bar{\mathcal{I}}$ for the closure of $\mathcal{I}$, and $\partial \bar{\mathcal{I}}$ for the boundary, so $\partial \bar{\mathcal{I}} = \{a_\mathcal{I}, b_\mathcal{I}\}$. We assume that the corresponding endpoints are either absorbing (in which case they are in $\mathcal{I}$), or inaccessible (in which case, if for example $b_\mathcal{I}$ is inaccessible and finite, then $\mathbb{P}(X_t \to b_\mathcal{I} \text{ as } t \to \infty) > 0$). The measures we wish to embed will be assumed to be supported on $\bar{\mathcal{I}}$, and in the case where $\mathcal{I} \ne \bar{\mathcal{I}}$, it may be possible to embed mass at $\partial \bar{\mathcal{I}}$ by taking a stopping time which takes

---

[1]Indeed, we were made aware of this paper only in the final stages of completing this work.

the value ∞. We define $\mathcal{S} := [0, \infty] \times \bar{\mathcal{I}}$. We note also[2] that as a consequence of the assumption on η, we have $\mathbb{E}^x [X_t^2] < \infty$, and we further write $m_{\mu_0}(t) := \mathbb{E}^{\mu_0} |X_t|$ for suitable measures μ_0.

We will also frequently want to restart the space-time process, given some stopped distribution in both time and space, and we will write ξ for a general probability measure on $\mathcal{S}$, with typically $\xi \sim (\sigma, X_\sigma)$ for some stopping time σ. With this notation, we have, $\mathbb{E}^\xi [A] = \int \mathbb{E}^{t,x} [A] \, \xi(dt, dx)$ and we denote (T_ξ, X_{T_ξ}) the random starting point, which then has law ξ. Since ξ may put mass on $\bar{\mathcal{I}} \setminus \mathcal{I}$, we interpret the process started at such a point as the constant process. For each of these processes, L_t^x denotes the (semimartingale) local time at x corresponding to the process X_t, with the convention that $L_t^x = 0$ for $t \leq T_\xi$. In addition, given a barrier $\mathcal{R}$, we define the corresponding hitting time of $\mathcal{R}$ by X under $\mathbb{P}^\xi$ by:

$$\sigma_\mathcal{R} = \inf\{t \geq T_\xi : (t, X_t) \in \mathcal{R}\}.$$

Similarly, given a stopping time σ_0 we write

$$\sigma_\mathcal{R}(\sigma_0) = \inf\{t \geq \sigma_0 : (t, X_t) \in \mathcal{R}\}.$$

Finally, we observe that, as a consequence of the (local) Lipschitz property of η, we know there exists a continuous transition density, $p : (0, \infty) \times \mathcal{I}^\circ \times \mathcal{I}^\circ$, so that

$$\mathbb{E}^x [f(X_t)] = \int p(t, x, y) f(y) \, dy,$$

whenever f is supported in $\mathcal{I}$ (see e.g. (Rogers and Williams, 2000, Theorem V.50.11)). We observe that we then have the following useful identities for the local time (see e.g. (Karatzas and Shreve, 1991, Theorem 3.7.1)):

$$\int_0^t f(X_s) \eta^2(X_s) \, ds = \int f(a) L_t^a \, da$$

and

$$\mathbb{E}^y [L_t^x] = \eta(x)^2 \int_0^t p(s, y, x) \, ds. \tag{1.1}$$

2 The Root solution of the Skorokhod embedding problem

2.1 Definitions

Throughout this paper, we consider a sequence of centred probability measures $\boldsymbol{\mu}_n := (\mu_i)_{i=0,\dots,n}$ on $\bar{\mathcal{I}}$:

$$\int_\mathcal{I} |x| \mu_i(dx) < \infty, \quad \text{and} \quad \int_\mathcal{I} x \mu_i(dx) = 0, \quad i = 0, \dots, n. \tag{2.1}$$

We similarly denote $\boldsymbol{\mu}_k = (\mu_0, \mu_1, \dots, \mu_k)$ for all $k \leq n$. We say that $\boldsymbol{\mu}_k$ is in convex order, and we denote $\mu_0 \preceq_{\text{cx}} \dots \preceq_{\text{cx}} \mu_k$, if

$$\int_\mathbb{R} c(x) \mu_{i-1}(dx) \leq \int_\mathbb{R} c(x) \mu_i(dx), i = 1, \dots, k \quad \text{for all convex functions} \quad c. \tag{2.2}$$

The lower and the upper bounds of the support of μ_k *relative to* μ_{k-1} are denoted by

$$\ell_k := \inf \left\{ x : \mu_k \big[(-\infty, x) \big] \neq \mu_{k-1} \big[(-\infty, x) \big] \right\} \quad \text{and} \quad r_k := \sup \left\{ x : \mu_k \big[(x, \infty) \big] \neq \mu_{k-1} \big[(x, \infty) \big] \right\}. \tag{2.3}$$

We exclude the case where $\mu_k = \mu_{k-1}$ as a trivial special case, and so we always have $\ell_k < r_k$ for all $k = 1, \dots, n$, as a consequence of the convex ordering. The potential of a probability measure μ is defined by

$$U^\mu(x) := -\int_\mathbb{R} |x - y| \mu(dy); \quad x \in \mathbb{R}, \tag{2.4}$$

---

[2]See the proof of Lemma 5.1 below for a suitable argument.

see Chacon (1977). For centred measures $\boldsymbol{\mu}_n$ in convex order, we have

$$U^{\mu_k} \leq U^{\mu_{k-1}} \quad \text{and} \quad U^{\mu_k} = U^{\mu_{k-1}} \quad \text{on} \quad (\ell_k, r_k)^c, \quad \text{for all} \quad k = 1, \ldots, n. \tag{2.5}$$

Recall that $(X_t)_{t \in \mathbb{R}_+}$ is a martingale diffusion. A stopping time σ (which may take the value ∞ with positive probability) is said to be uniformly integrable (UI) if the process $(X_{t \wedge \sigma})_{t \geq 0}$ is uniformly integrable under $\mathbb{P}^{\mu_0}$. We denote by $\mathcal{T}$ the collection of all UI stopping times.

The classical Skorokhod embedding problem with starting measure μ_0 and target measure μ_1 is:

$$\text{SEP}(\boldsymbol{\mu}_1): \qquad \text{find } \sigma \in \mathcal{T} \text{ such that } X_\sigma \sim \mu_1 \text{ under } \mathbb{P}^{\mu_0}. \tag{2.6}$$

We consider the problem with multiple marginals:

$$\text{SEP}(\boldsymbol{\mu}_n): \qquad \text{find } 0 \leq \sigma_1 \ldots \leq \sigma_n \in \mathcal{T} \text{ such that } X_{\sigma_k} \sim \mu_k, \ k = 1, \ldots, n \text{ under } \mathbb{P}^{\mu_0}. \tag{2.7}$$

In this paper, our interest is in a generalisation of the Root (1969) solution of the Skorokhod embedding problem so that each stopping time σ_k is the first hitting time, after σ_{k-1}, by $(t, X_t)_{t \geq 0}$ of some subset $\mathcal{R}$ in $\mathcal{S}$. Further, and crucially, we require that $\mathcal{R}$ is a barrier in the following sense:

Definition 2.1. *A set $\mathcal{R} \subset \mathcal{S}$ is called a* barrier *if*

- *$\mathcal{R}$ is closed;*
- *if $(t, x) \in \mathcal{R}$ then $(s, x) \in \mathcal{R}$ for all $s \geq t$;*
- *if $x \in \{a_\mathcal{I}, b_\mathcal{I}\}$ is finite, $(0, x) \in \mathcal{R}$.*

Given a barrier $\mathcal{R}$, for $x \in \bar{\mathcal{I}}$, we define the corresponding barrier function:

$$\bar{t}_\mathcal{R}(x) \quad := \quad \inf\{t \geq 0 : (t, x) \in \mathcal{R}\} \in [0, \infty]. \tag{2.8}$$

Since $\mathcal{R}$ is closed it follows, as observed by Root (1969) and Loynes (1970), that $\bar{t}_\mathcal{R}(\cdot)$ is lower semi–continuous on $\mathcal{I}$. Also, from the second property, we see that a barrier is the epigraph of the corresponding barrier function in the (t, x)-plane:

$$\mathcal{R} \quad = \quad \big\{(t, x) \in \mathbb{R}_+ \times \mathcal{I} : \ t \geq \bar{t}_\mathcal{R}(x)\big\}.$$

Definition 2.2. (i) *We say that a barrier is* regular *if $\{x \in \mathcal{I}^\circ : \bar{t}_\mathcal{R}(x) > 0\}$ is an open interval containing zero.*

(ii) *For a probability measure $\xi = \xi(dt, dx)$ on $\mathcal{S}$, we say that a barrier is ξ-regular if*

$$\mathbb{P}^\xi\big[\sigma_\mathcal{R} = \sigma_{\mathcal{R}^{(t,x)}}\big] < 1 \quad \text{for all} \quad (t, x) \notin \mathcal{R}, \quad \text{where} \quad \mathcal{R}^{(t,x)} = \mathcal{R} \cup \big([t, \infty) \times \{x\}\big),$$

i.e. the barrier cannot be enlarged without altering the stopping distribution of the space-time diffusion started with law ξ and run to the hitting of $\mathcal{R}$.

Observe that a regular barrier is a $\delta_{(0,0)}$-regular barrier. We have the following characterisation:

Remark 2.3. *A barrier $\mathcal{R}$ is ξ-regular if and only if $\mathbb{E}^\xi\big[L_{t \wedge \sigma_\mathcal{R}}^x\big] < \mathbb{E}^\xi\big[L_{\sigma_\mathcal{R}}^x\big]$ for all $(t, x) \notin \mathcal{R}$.*

Lemma 2.4. *Let ξ be a probability measure on $\mathcal{S}$ and $\mathcal{R}$ a barrier such that $\inf_{x \in \mathcal{I}} \bar{t}_\mathcal{R}(x) < \infty$. Then $\sigma_\mathcal{R} < \infty$ or $\lim_{t \to \infty} X_t \in \{a_\mathcal{I}, b_\mathcal{I}\}$ $\mathbb{P}^\xi$-a.s. Further, if $\mathcal{R}$ is not ξ-regular then there exists a ξ-regular barrier $\tilde{\mathcal{R}} \supseteq \mathcal{R}$ such that $X_{\sigma_\mathcal{R}} \sim X_{\sigma_{\tilde{\mathcal{R}}}}$ $\mathbb{P}^\xi$-a.s.*

Proof. For some $x_0 \in \mathcal{I}$, we have $\bar{t}_\mathcal{R}(x_0) < \infty$ and $\{(t, x) : t \geq \bar{t}_\mathcal{R}(x_0)\} \subset \mathcal{R}$. If $\mathcal{I} = \mathbb{R}$ then $\limsup_t X_t = \infty$ and $\liminf X_t = -\infty$ and it is clear that $\sigma_\mathcal{R} < \infty$ $\mathbb{P}^\xi$-a.s. Otherwise $\lim_{t \to \infty} X_t \in \{a_\mathcal{I}, b_\mathcal{I}\}$. If $\mathcal{R}$ is not ξ-regular then by definition the set of all barriers $\tilde{\mathcal{R}}$ for which $X_{\sigma_\mathcal{R}} \sim X_{\sigma_{\tilde{\mathcal{R}}}}$ $\mathbb{P}^\xi$-a.s. is not a singleton. Then for any two such barriers $\tilde{\mathcal{R}}_1, \tilde{\mathcal{R}}_2$ their union is also such a barrier, as shown by Loynes (1970). It follows that there exists a minimal such barrier with respect to the inclusion which then necessarily has to be ξ-regular. $\square$

It follows that, without loss of generality, we may restrict our attention to ξ-regular barriers. Henceforth, whenever a barrier is given it is assumed that it is a ξ-regular barrier, where the measure ξ will be clear from the context.

2.2 Root's solution and its PDE characterisation

The main result of Root (1969) is the following.

Theorem 2.5 (Root (1969)). *Let $\mu_0 = \delta_0$, $\eta(x) \equiv 1$, and μ_1 be a centred probability measure on $\mathbb{R}$ with a finite second moment. Then there exists a barrier $\mathcal{R}^*$ such that $\sigma_{\mathcal{R}^*}$ is a solution of $\mathrm{SEP}(\boldsymbol{\mu}_1)$.*

The first significant generalisation of this result is due to Rost (1976) who showed that the result generalised to transient Markov processes under certain conditions. The condition that the probability measure μ_1 has finite second moment has only very recently been further relaxed to the more natural condition that the measure has a finite first moment. This was first achieved by Gassiat et al. (2014), who have extended Root's result to the case of one-dimensional (time-inhomogeneous) diffusions using PDE methods. The result was also obtained by Beiglböck et al. (2016) using methods from Optimal Transport theory.

Remark 2.6. *Loynes (1970) showed, as used above in Lemma 2.4, that in Theorem 2.5 the barrier can be taken to be regular and is then unique.*

We next recall the recent work of Cox and Wang (2013b) and Gassiat et al. (2014). For a function $u : (t,x) \in \mathbb{R}_+ \times \mathbb{R} \longmapsto u(t,x) \in \mathbb{R}$, we denote by $\partial_t u$ the $t-$derivative, Du, D^2u the first and second spacial derivatives, i.e. with respect to the x-variable, and we introduce the (heat) second order operator

$$\mathcal{L}u \quad := \quad -\partial_t u + \frac{1}{2}\eta^2 D^2 u. \tag{2.9}$$

Consider the variational inequality or obstacle problem:

$$\min\big\{ -\mathcal{L}u \,,\, u - U^{\mu_1} \big\} = 0 \quad \text{and} \quad u(0,\cdot) = U^{\mu_0}. \tag{2.10}$$

Then, based on the existence result of Root (1969), Cox and Wang (2013b) proved the following result.

Theorem 2.7 (Theorem 4.2, Cox and Wang (2013b); Theorem 2, Gassiat et al. (2014)). *Let $\boldsymbol{\mu}_1 = (\mu_0, \mu_1)$ be centred probability measures on $\mathbb{R}$ in convex order. Then, there is a unique solution u^1 of (2.10) which extends continuously to $[0,\infty] \times [-\infty,\infty]$, and the Root solution of the $\mathrm{SEP}(\boldsymbol{\mu}_1)$ is induced by the regular barrier*

$$\mathcal{R}^* \quad = \quad \big\{(t,x) \in [0,\infty] \times [-\infty,\infty] : u^1(t,x) = U^{\mu_1}(x)\big\}.$$

Moreover, we have the representation $u^1(t,x) = -\mathbb{E}\big|X_{t\wedge\sigma_{\mathcal{R}^*}} - x\big|$, for all $t \geq 0, x \in \mathbb{R}$.

In Cox and Wang (2013b), the solution to the variational inequality was determined as a solution in an appropriate Sobolev space, while Gassiat et al. (2014) show that the solution can be understood in the viscosity sense.

2.3 Optimal stopping characterisation

The objective of this paper is to provide a probabilistic version of the last result, and its generalisation to the multiple marginal problem. Our starting point is the classical probabilistic representation of the solution to (2.10) as an optimal stopping problem. Define now

$$u^1(t,x) := \sup_{\tau \in \mathcal{T}^t} J^1_{t,x}(\tau) \quad \text{with} \quad J^1_{t,x}(\tau) := \mathbb{E}^x\big[U^{\mu_0}(Y_\tau) + (U^{\mu_1} - U^{\mu_0})(Y_\tau)\mathbf{1}_{\{\tau<t\}}\big], \tag{2.11}$$

where $\mathcal{T}^t$ is the collection of all $(\mathcal{F}_t)$–stopping times $\tau \leq t$. Then, using classical results, see e.g. Bensoussan and Lions (1982), when properly understood, u_1 in (2.11) is a solution to (2.10). Uniqueness, in an appropriate sense, of solutions to (2.10), then allows to deduce that the characterisation of the Root barrier given in Theorem 2.7 corresponds to the stopping region of the optimal stopping problem (2.11)

$$\mathcal{R}^1 \quad := \quad \big\{(t,x) \in [0,\infty] \times [-\infty,\infty] : u^1(t,x) = U^{\mu_1}(x)\big\}. \tag{2.12}$$

The probabilistic approach we develop in this paper provides a self-contained construction of the Root solution, and does not rely on the existence result of Root (1969) or PDE results. Indeed, these follow from the following direct characterisation which is a special case of Theorem 3.1 below.

Theorem 2.8. *Let $\boldsymbol{\mu}_1 = (\mu_0, \mu_1)$ be centred probability measures on $\bar{\mathcal{I}}$ in convex order. Then, $\mathcal{R}^1$ defined by (2.11)-(2.12) is the regular barrier inducing the Root solution of the* $\mathrm{SEP}(\boldsymbol{\mu}_1)$. *Moreover,*

$$u^1(t,x) \quad = \quad -\mathbb{E}^{\mu_0}\big|X_{t \wedge \sigma_{\mathcal{R}^1}} - x\big|, \quad \textit{for all} \ \ t \geq 0, \ x \in \mathbb{R}.$$

3 Multiple Marginal Root Solution of the SEP: main results

3.1 Iterated optimal stopping and multiple marginal barriers

In order to extend the Root solution to the multiple marginals $\mathrm{SEP}(\boldsymbol{\mu}_n)$, we now introduce the following natural generalisation of the previous optimal stopping problem. Denote

$$\delta U^k(x) := U^{\mu_k}(x) - U^{\mu_{k-1}}(x), \quad \text{and} \quad u^0(t,x) := U^{\mu_0}(x), \quad t \in [0,\infty], x \in \bar{\mathcal{I}}.$$

The main ingredient for our construction is the following iterated sequence of optimal stopping problems:

$$u^k(t,x) := \sup_{\tau \in \mathcal{T}^t} J^k_{t,x}(\tau) \quad \text{where} \quad J^k_{t,x}(\tau) := \mathbb{E}^x\Big[u^{k-1}(t-\tau, Y_\tau) + \delta U^k(Y_\tau)\mathbf{1}_{\{\tau < t\}}\Big], \ 1 \leq k \leq n. \tag{3.1}$$

The stopping regions corresponding to the above sequence of optimal stopping problems are given by:

$$\mathcal{R}^k := \big\{(t,x) \in \mathcal{S} : \delta u^k(t,x) = \delta U^k(x)\big\} \quad \text{with} \quad \delta u^k := u^k - u^{k-1}, k = 1,\ldots,n, \tag{3.2}$$

and the optimal stopping time which solves (3.1) is the first entry to $\mathcal{R}^k$ by the time space process starting in (t,x) and running backwards in time: $\tau^t(k) := \inf\{s \geq 0 : (t-s, Y_s) \in \mathcal{R}^k\} \wedge t$.

Our main result shows that the same barriers used to stop the process running forward in time:

$$\sigma_0 = 0, \quad \sigma_k := \sigma_{\mathcal{R}^k}(\sigma_{k-1}) = \inf\big\{t \geq \sigma_{k-1} : (t, X_t) \in \mathcal{R}^k\big\}, \ k = 1,\ldots,n, \tag{3.3}$$

give the multiple marginals Root solution of $\mathrm{SEP}(\boldsymbol{\mu}_n)$. It is important to note that the barriers in (3.2) are not necessarily nested – both $\mathcal{R}^k$ and $\mathcal{R}^{k-1}$ may contain points which are not in the other barrier. An example of a possible sequence of stopping times is depicted in Figure 1. Since the barriers are not necessarily nested, in general σ_k will not be equal to the first entry time to the barrier, only the first entry time *after* σ_{k-1}. It may also be the case that $\sigma_{k-1} = \sigma_k$. Both cases are shown in Figure 1.

Finally, it will be useful to introduce the (time-space) measures on $\mathcal{S}$ defined for all Borel subsets A of $\mathcal{S}$ by:

$$\xi^k[A] := \mathbb{P}^{\mu_0}\big[(\sigma_k, X_{\sigma_k}) \in A\big], \quad k = 0,\ldots,n.$$

We are now ready to state our main result, which includes Theorem 2.8 as a special case.

Theorem 3.1. *Let $\boldsymbol{\mu}_n$ be a vector of centred probability measures on $\bar{\mathcal{I}}$ in convex order. Then $\mathcal{R}^k$ is a* ξ^{k-1}*-regular barrier for all $k = 1,\ldots,n$, and $(\sigma_1, \sigma_2, \ldots, \sigma_n)$ solves* $\mathrm{SEP}(\boldsymbol{\mu}_n)$. *Moreover, we have*

$$u^k(t,x) = -\mathbb{E}^{\mu_0}\big|X_{t \wedge \sigma_k} - x\big|, \quad \textit{for all} \quad t \geq 0, \ x \in \bar{\mathcal{I}}, \ k = 1,\ldots,n. \tag{3.4}$$

Our proof will proceed by induction. Its main ingredients will be summarised in Section 4.
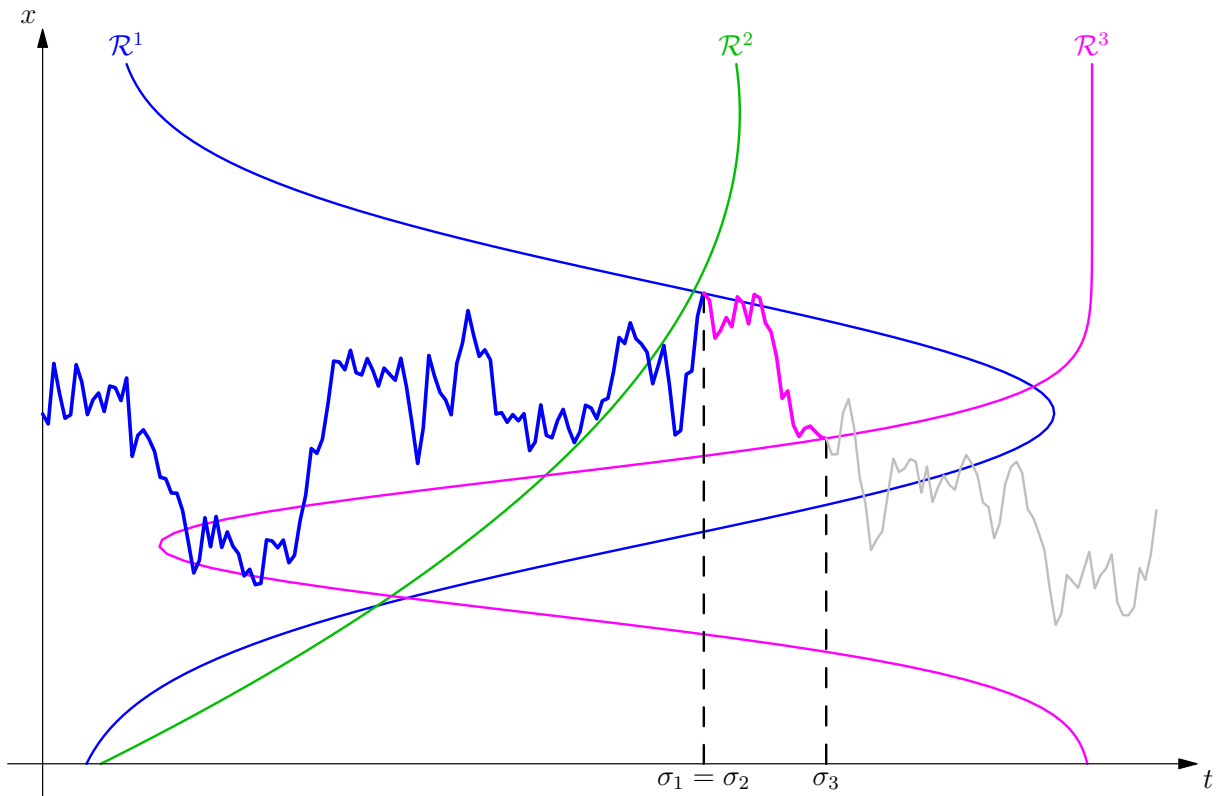
Figure 1: A realisation of a Root-type solution to the multiple marginal problem. Here we depict three barriers which are not ordered (in the sense that $\mathcal{R}^1 \supsetneq \mathcal{R}^2 \supsetneq \mathcal{R}^3$). As a result, the given realisation can enter the second and third barriers before the first stopping time. Note also that since the first stopping time, σ_1, happens at a point which is also inside the second barrier, we have here $\sigma_1 = \sigma_2$.

3.2   Optimality

In this section, we show optimality of the constructed n-fold Root solution of the multiple marginal Skorokhod embedding problem. We recall the main ingredients of our embedding defined in (3.1)–(3.3). We also denote $\bar{t}^k := \bar{t}_{\mathcal{R}^k}$. Define the set of all solutions to SEP($\boldsymbol{\mu}_n$) in (2.7):

$$\mathcal{T}(\boldsymbol{\mu}_n) \quad := \quad \big\{\rho = (\rho_1, \ldots, \rho_n) \in \mathcal{T}^n : \ \rho_1 \leq \ldots \leq \rho_n, \ \text{and} \ X_{\rho_i} \sim \mu_i, \ i = 1, \ldots, n \big\}.$$

For a given function $f : \mathbb{R} \longrightarrow \mathbb{R}_+$ we consider the optimal n-fold embedding problem:

$$\inf_{\rho \in \mathcal{T}(\boldsymbol{\mu}_n)} \mathbb{E}^{\mu_0}\Big[\int_0^{\rho_n} f(t) dt \Big]. \tag{3.5}$$

Theorem 3.2. *Let $\boldsymbol{\mu}_n$ be a vector of centred probability measures on $\bar{\mathcal{I}}$ in convex order and f a non-negative non-decreasing function. Then the n-tuple $\sigma = (\sigma_1, \ldots, \sigma_n)$ in (3.3) is a solution of (3.5):*

$$\sigma \in \mathcal{T}(\boldsymbol{\mu}_n) \quad and \quad \mathbb{E}^{\mu_0}\Big[\int_0^{\sigma_n} f(t) dt \Big] \ \leq \ \mathbb{E}^{\mu_0}\Big[\int_0^{\rho_n} f(t) dt \Big] \quad for \ all \quad \rho \in \mathcal{T}(\boldsymbol{\mu}_n).$$

The above remains true for any stopping times $\rho_1, \ldots, \rho_n$ which embed $\boldsymbol{\mu}$ since if ρ is not uniformly integrable then it is not minimal, see (Obłój, 2004, Section 8), and we can find smaller stopping times $\tilde{\rho} \in \mathcal{T}(\boldsymbol{\mu}_n)$ for which the above bound is already satisfied.

Similar to many proofs of optimality of particular solutions to SEP, see e.g. Hobson (1998); Cox et al. (2008); Henry-Labordère et al. (2016), at the heart of our argument lies identification of a suitable pathwise inequality. Interpreting (3.5) as an iterated Martingale Optimal Transport problem, the pathwise

inequality amounts to an explicit identification of the dual optimiser in the natural Kantorovich-type duality. Our inequality is inspired by the one developed by Cox and Wang (2013b).

For all $(t,x) \in \mathbb{R}_+ \times \bar{\mathcal{I}}$ and $k = n, \ldots, 0$, we introduce the functions

$$\varphi_{n+1}(t,x) := f(t), \qquad\qquad \varphi_k(t,x) := \mathbb{E}^{t,x}\big[\varphi_{k+1}\big(\sigma_{\mathcal{R}^k}, X_{\sigma_{\mathcal{R}^k}}\big)\big],$$

$$\phi_k(x) := \int_0^x \varphi_k(0,y)\eta(y)^{-2}dy, \qquad\qquad \psi(x) := 2\int_0^x \int_0^y \eta(z)^{-2}\, dz.$$

Our main result below involves the following functions:

$$h_k(t,x) := \int_0^t \varphi_k(s,x)ds - 2\int_0^x \phi_k(y)dy, \quad \text{and} \quad \lambda_k(x) := (h_{k+1} - h_k)\big(\bar{t}_k(x), x\big), \ (t,x) \in \mathbb{R}_+ \times \mathcal{I}. \quad (3.6)$$

Lemma 3.3. *Let f be a non-negative non-decreasing function. Then for all $(s_i, x_i)_{0 \le i \le n} \subset \mathbb{R}_+ \times \bar{\mathcal{I}}$, with $0 = s_0 \le s_1 \le \cdots \le s_n$, we have:*

$$\int_0^{s_n} f(t)dt \ \ge \ \sum_{i=1}^n \lambda_i(x_i) + h_1(s_0, x_0) + \sum_{i=1}^n \big[h_i(s_i, x_i) - h_i(s_{i-1}, x_{i-1})\big] - \psi(x_n)f(0), \quad (3.7)$$

and equality holds if $(s_i, x_i) \in \mathcal{R}^i$ for $i = 1, \ldots, n$.

The proof of the above inequality is entirely elementary, even if not immediate, and is reported in Appendix A. The optimality in Theorem 3.2 then essentially follows by evaluating the above on stopped paths (ρ_i, X_{ρ_i}) and taking expectations. Technicalities in the proof are mainly related to checking suitable integrability of various terms and the proof is also reported in Appendix A.

Finally, we note that the above pathwise inequality could be evaluated on paths of arbitrary martingale and, after taking expectations, would lead to a martingale inequality. The inequality would be sharp in the sense that we have equality for X stopped at σ in (3.3). This method of arriving at martingale inequalities is linked to the so-called Burkholder method, see e.g. Burkholder (1991), and has been recently exploited in number of works, see e.g. Acciaio et al. (2013); Beiglböck and Nutz (2014); Obłój et al. (2015).

4 The inductive step

In this section we outline the main ideas behind the proof of Theorem 3.1. The proof proceeds by induction. At the end of each step in the induction, we will determine a stopping time σ^ξ, and the time-space distribution ξ, which corresponds to the distribution of the stopped process $(\sigma^\xi, X_{\sigma^\xi})$ under the starting measure μ_0. This measure will be the key part of the subsequent definitions. Given this stopping time, and a new law β, we proceed to determine a new stopping time σ^{ξ^β}, and the corresponding time-space distribution ξ^β. This stopping time will embed the law β. This inductive step is summarised in Theorem 4.1 below.

This stopping time σ^{ξ^β} is constructed as the solution of an optimal stopping problem u^β, introduced below, with obstacle function appropriately defined by combining the potential function v^ξ of the stopped process $X_{\cdot \wedge \sigma^\xi}$ and the difference of potentials between the starting distribution – the spatial marginal of ξ denoted α^ξ – and the target distribution β. We will also show that the function u^β is equal to the potential function v^{ξ^β}, allowing us to iterate the procedure.

We now introduce the precise definitions. The measure μ_0 will be a fixed integrable measure throughout, and so we will typically not emphasise the dependence of many terms on this measure.

Let ξ be the $\mathbb{P}^{\mu_0}$–time-space distribution of $(\sigma^\xi, X_{\sigma^\xi})$ for some UI stopping time $\sigma^\xi \in \mathcal{T}$. The stopped potential v^ξ is defined as the $\mathbb{P}^{\mu_0}$–potential of $X_{t \wedge \sigma^\xi}$:

$$v^\xi(t,x) := -\mathbb{E}^{\mu_0}\big[|X_{t \wedge \sigma^\xi} - x|\big], \quad t \ge 0, \quad x \in \bar{\mathcal{I}}. \quad (4.1)$$

Motivated by the iterative optimal stopping problems (3.1), we also introduce, for any probability measure β on $\bar{\mathcal{I}}$, the difference of potentials

$$w^\beta := U^\beta - U^{\alpha^\xi} \quad \text{where} \quad \alpha^\xi(A) := \xi([0,\infty) \times A), A \in \mathcal{B}(\bar{\mathcal{I}}),$$

and $\alpha^\xi \preceq_{\mathrm{cx}} \beta$ is equivalent to $w^\beta \leq 0$. Moreover, since σ^ξ is UI, we have

$$\mu_0 \preceq_{\mathrm{cx}} \alpha^\xi, \quad v^\xi(0,.) = U^{\mu_0}, \quad \text{and} \quad v^\xi(t,.) \searrow v^\xi(\infty,.) := U^{\alpha^\xi} \text{ pointwise as } t \nearrow \infty. \tag{4.2}$$

The optimal stopping problem which will serve for our induction argument is:

$$u^\beta(t,x) \quad := \quad \sup_{\tau \in \mathcal{T}^t} \mathbb{E}^x \left[v^\xi(t - \tau, Y_\tau) + w^\beta(Y_\tau) \mathbf{1}_{\{\tau < t\}} \right] \quad t \geq 0, x \in \bar{\mathcal{I}}. \tag{4.3}$$

We also introduce the corresponding stopping region

$$\mathcal{R}^\beta \quad := \left\{ (t,x) : u^\beta(t,x) = v^\xi(t,x) + w^\beta(x) \right\}, \tag{4.4}$$

and we set

$$\sigma^{\xi^\beta} := \inf\{t \geq \sigma^\xi : (t, X_t) \in \mathcal{R}^\beta\}, \text{ and } \xi^\beta[A] := \mathbb{P}^\xi \left[(\sigma^{\xi^\beta}, X_{\sigma_{\mathcal{R}^\beta}}) \in A \right] \text{ for all } A \in \mathcal{B}(\mathcal{S}). \tag{4.5}$$

Theorem 4.1. *Let $\sigma^\xi \in \mathcal{T}$ with corresponding time-space distribution ξ, and β an integrable measure such that $\beta \succeq_{cx} \alpha^\xi$. Then σ^{ξ^β} is a UI stopping time embedding β and $u^\beta = v^{\xi^\beta}$. Moreover, $\mathcal{R}^\beta$ is a ξ^β-regular barrier.*

We now show that Theorem 3.1 follows from Theorem 4.1.

Proof of Theorem 3.1. Consider the first marginal. Let $\xi = \delta_0 \otimes \mu_0$ so that $\sigma^\xi = 0$, $\alpha^\xi = \mu_0$, and let $\beta = \mu_1$. Then $v^\xi(t,x) = U^{\mu_0}(x)$ and $u^\beta, \mathcal{R}^\beta$ in (4.3)–(4.4) are equal to, respectively, $u^1, \mathcal{R}^1$ in (3.1)-(3.2). It follows from Theorem 4.1 that the stopping time $\sigma_1 = \sigma^{\xi^\beta}$ induced by $\mathcal{R}^1 = \mathcal{R}^\beta$ is a UI stopping time solving SEP($\boldsymbol{\mu}_1$) and $u^1 = u^\beta = v^{\xi^\beta}$, as required. We next iterate the arguments. Given the UI stopping time σ_k from the kth step with its space-time measure ξ we know that $v^\xi = u^k$ so that, with $\beta = \mu_{k+1}$, we have $u^\beta = u^{k+1}$ and $\mathcal{R}^\beta = \mathcal{R}^{k+1}$. Applying Theorem 4.1 we get that σ_{k+1} embeds μ_{k+1}, is UI and $v^{\xi^\beta} = u^{k+1}$ as required. The proof finishes after n iterations. $\square$

The rest of this paper is dedicated to the proof of Theorem 4.1. The following result isolates the main steps needed for this.

Lemma 4.2. *Let $\sigma^\xi \in \mathcal{T}$ with corresponding time-space distribution ξ, and $\alpha^\xi \preceq_{\mathrm{cx}} \beta$. Assume further that $u^\beta(t,.) \longrightarrow U^\beta$, pointwise as $t \nearrow \infty$, and $u^\beta = v^{\xi^\beta}$. Then, σ^{ξ^β} is a UI stopping time embedding β.*

Proof. From the assumptions and the definition of v^{ξ^β} we obtain

$$-U^\beta(x) = -\lim_{t \to \infty} v^{\xi^\beta}(t,x) = \lim_{t \to \infty} \mathbb{E}^{\mu_0} \left[|X_{t \wedge \sigma^{\xi^\beta}} - x| \right] \geq \mathbb{E}^{\mu_0} \left[|X_{\sigma^{\xi^\beta}} - x| \right] = -U^{\alpha^{\xi^\beta}}(x),$$

where the inequality follows from Fatou's Lemma. This in particular implies that α^{ξ^β} is an integrable probability measure on $\bar{\mathcal{I}}$, $U^{\alpha^{\xi^\beta}}(x) > -\infty$ for all $x \in \bar{\mathcal{I}}$, and $U^{\xi^\beta}(x) - |x - m^{\alpha^{\xi^\beta}}| \longrightarrow 0$ as $x \to \partial\bar{\mathcal{I}}$, where $m^{\alpha^{\xi^\beta}} := \int x \alpha^{\xi^\beta}(dx)$. Since also $U^{\xi^\beta}(x) - |x - m^\beta| \longrightarrow 0$ as $x \to \partial\bar{\mathcal{I}}$, we deduce from the above inequality $U^\beta \leq U^{\alpha^{\xi^\beta}}$ that $m^{\alpha^{\xi^\beta}} = m^\beta$, and therefore $|U^{\alpha^{\xi^\beta}}(x) - U^\beta(x)| \to 0$ as $x \to \partial\bar{\mathcal{I}}$. Then for $x, y \in \bar{\mathcal{I}}$, it follows from the dominated convergence theorem that

$$\begin{aligned} U^\beta(x) - U^\beta(y) &= \lim_{t \to \infty} \left[v^{\xi^\beta}(t,x) - v^{\xi^\beta}(t,y) \right] \\ &= \lim_{t \to \infty} \mathbb{E}^{\mu_0} \left[|X_{t \wedge \sigma^{\xi^\beta}} - y| - |X_{t \wedge \sigma^{\xi^\beta}} - x| \right] \\ &= \mathbb{E}^{\mu_0} \left[|X_{\sigma^{\xi^\beta}} - y| - |X_{\sigma^{\xi^\beta}} - x| \right] \\ &= U^{\alpha^{\xi^\beta}}(x) - U^{\alpha^{\xi^\beta}}(y). \end{aligned}$$

10

In particular, $U^\beta(x) = U^{\alpha^{\xi^\beta}}(x) + c$ for some $c \in \mathbb{R}$, for all $x \in \bar{\mathcal{I}}$, and by the above, sending $x \to \partial\bar{\mathcal{I}}$, we see that $c = 0$. We conclude that $\alpha^{\xi^\beta} = \beta$, i.e. $X_{\sigma^{\xi^\beta}} \sim \beta$, which is the required embedding property. Moreover, it follows from the Tanaka formula together with the monotone convergence theorem that

$$U^\beta(x) \quad = \quad U^{\alpha^{\xi^\beta}}(x) \quad = \quad -\mathbb{E}^{\mu_0}\big[|X_{\sigma^{\xi^\beta}} - x|\big] \quad = \quad U^{\mu_0}(x) - \mathbb{E}^{\mu_0}\big[L^x_{\sigma^{\xi^\beta}}\big], \quad \text{for all} \quad x \in \bar{\mathcal{I}}.$$

The uniform integrability of the stopping time σ^{ξ^β} now follows from (Elworthy et al., 1999, Corollary 3.4). $\qquad\square$

The pointwise convergence of $u^\beta(t,.)$ towards U^β, as $t \to \infty$ will be stated in Lemma 5.5 (iii), while the equality $u^\beta = v^{\xi^\beta}$ is more involved and will be shown through a series of results, see Lemma 7.3.

Remark 4.3. *We have $u^\beta = v^{\xi^\beta}$ if and only if $(v^\xi - u^\beta)(t,x) = \mathbb{E}^\xi\big[L^x_{t \wedge \sigma_{\mathcal{R}^\beta}}\big]$, for all $t \geq 0, x \in \bar{\mathcal{I}}$. Indeed, by the Tanaka formula,*

$$v^{\xi^\beta}(t,x) \quad = \quad U^{\mu_0}(x) - \mathbb{E}^{\mu_0}\left[L^x_{t \wedge \sigma^{\xi^\beta}}\right] \quad = \quad v^\xi(t,x) - \mathbb{E}^{\mu_0}\left[L^x_{t \wedge \sigma^{\xi^\beta}} - L^x_{t \wedge \sigma^\xi}\right].$$

Recalling that, under $\mathbb{P}^\xi$, $\sigma_{\mathcal{R}^\beta} = \inf\{t > T_\xi : (t, X_t) \in \mathcal{R}^\beta\}$, and (under $\mathbb{P}^{\mu_0}$), $\sigma^{\xi^\beta} = \inf\{t > \sigma^\xi : (t, X_t) \in \mathcal{R}^\beta\}$. Recall that, under $\mathbb{P}^\xi$, the local time is set to $L^x_t = 0$ for $t \leq T_\xi$, by convention. Then from the strong Markov property, we have $\mathbb{E}^{\mu_0}\left[L^x_{t \wedge \sigma^{\xi^\beta}} - L^x_{t \wedge \sigma^\xi}\right] = \mathbb{E}^{(\sigma^\xi, X_{\sigma^\xi})}\left[L^x_{t \wedge \sigma_{\mathcal{R}^\beta}}\right] = \mathbb{E}^\xi\left[L^x_{t \wedge \sigma_{\mathcal{R}^\beta}}\right]$, and therefore:

$$v^{\xi^\beta}(t,x) \quad = \quad v^\xi(t,x) - \mathbb{E}^\xi\left[L^x_{t \wedge \sigma_{\mathcal{R}^\beta}}\right], \tag{4.6}$$

justifying the claimed equivalence.

Remark 4.4. *Observe that the regularity of the barrier can now be seen as an easy consequence of Lemma 4.2. Suppose (in the setting of Theorem 4.1), we have $u^\beta = v^{\xi^\beta}$ and $u^\beta(t,.) \to U^\beta$ pointwise as $t \to \infty$. From (4.6), (4.2) and applying monotone convergence to $\mathbb{E}^\xi\left[L^x_{t \wedge \sigma_{\mathcal{R}^\beta}}\right]$ as $t \to \infty$, we deduce that*

$$\mathbb{E}^\xi\left[L^x_{\sigma_{\mathcal{R}^\beta}}\right] = U^{\alpha^\xi}(x) - U^\beta(x) = -w^\beta(x).$$

Now suppose that $(t,x) \notin \mathcal{R}^\beta$. Then $\mathbb{E}^\xi\left[L^x_{\sigma_{\mathcal{R}^\beta}}\right] = -w^\beta(x) > (v^\xi - u^\beta)(t,x) = (v^\xi - v^{\xi^\beta})(t,x) = \mathbb{E}^\xi\left[L^x_{t \wedge \sigma_{\mathcal{R}^\beta}}\right]$, by (4.6). In view of Remark 2.3, this shows that $\mathcal{R}^\beta$ is $\xi-$regular.

5 Stopped potential and the optimal stopping problem

5.1 Properties of the stopped potential function

The following lemma provides some direct properties of the stopped potential. Recall the definition $m_\mu(t) := \mathbb{E}^\mu |X_t|$.

Lemma 5.1. *Let $\sigma^\xi \in \mathcal{T}$ with corresponding time-space distribution ξ. Then, v^ξ is concave and 1-Lipschitz-continuous in x, and non-increasing, and $v^\xi(t,x)$ is (uniformly in x) $\frac{1}{2}$-Hölder continuous on $[0,T]$ for all $T > 0$. In addition*

$$0 \quad \leq \quad U^{\mu_0}(x) - v^\xi(t,x) \quad = \quad \mathbb{E}^{\mu_0}\left[L^x_t\right] - \mathbb{E}^\xi\left[L^x_t\right] \quad \leq \quad \sqrt{2C_\eta t}\, m_{\mu_0}(t) \mathrm{e}^{C_\eta t},$$

and the following identity holds in the distribution sense:

$$\left(\mathcal{L}v^\xi\right)(t,dx) \quad = \quad -\int_0^t \eta(x)^2\, \xi(ds,dx); \quad t \geq 0, \; x \in \bar{\mathcal{I}},$$

11

by which we mean that, for any stopping time $\sigma \leq t$, we have

$$\mathbb{E}^x \left[v^\xi(t - \sigma, Y_\sigma) \right] - v^\xi(t, x) = -\mathbb{E}^x \left[\int_0^\sigma \eta(Y_s)^2 \, ds \int_0^s \delta_{Y_s}(y) \, \xi(dr, dy) \right] \tag{5.1}$$

$$= -\int_0^t \int_{\bar{\mathcal{I}}} q_\sigma(t - s, y) \eta(y)^2 \int_0^s \xi(dr, dy) \, ds$$

where q_σ is the space-time density of the process Y_s (started at t and running backwards in time) up to the stopping time σ.

Proof. The definition of $v^\xi(t, x)$ in (4.1) immediately shows that v^ξ is concave, $1-$Lipschitz in x, and non-increasing in t. As in Remark 4.3 above, using Tanaka's formula and the strong Markov property we obtain

$$v^\xi(t, x) = U^{\mu_0}(x) - \mathbb{E}^{\mu_0}[L^x_{t \wedge \sigma^\xi}] = U^{\mu_0}(x) - \mathbb{E}^{\mu_0}[L^x_t] + \mathbb{E}^{\mu_0}[(L^x_t - L^x_{\sigma^\xi}) \mathbf{1}_{\{\sigma^\xi \leq t\}}]$$

$$= U^{\mu_0}(x) - \mathbb{E}^{\mu_0}[L^x_t] + \mathbb{E}^\xi[L^x_t] \tag{5.2}$$

$$= U^{\mu_0}(x) - \mathbb{E}^{\mu_0}[L^x_t] + \int_{[0,t] \times \bar{\mathcal{I}}} \mathbb{E}^y \left[L^x_{t-s} \right] \xi(ds, dy).$$

We now consider continuity properties of $\mathbb{E}^y[L^x_t]$. First observe that, by the martingale property of X_t, we have

$$\mathbb{E}^y \left[(X_t - x)^2 \right] = \mathbb{E}^y \left[(X_t - y)^2 \right] + (x - y)^2.$$

Using the fact that $\eta(x)^2 \leq C_\eta(1 + |x|^2)$ and the martingale property of X, we deduce

$$\mathbb{E}^y \left[(X_t - y)^2 \right] \leq \mathbb{E}^y \left[\int_0^t \eta(X_s)^2 \, ds \right]$$

$$\leq C_\eta \mathbb{E}^y \left[\int_0^t \left(1 + |X_s|^2 \right) ds \right]$$

$$\leq C_\eta \left(t + 2\mathbb{E}^y \left[\int_0^t \left[(X_s - y)^2 + y^2 \right] ds \right] \right)$$

$$\leq 2C_\eta \left(t(1 + y^2) + \mathbb{E}^y \left[\int_0^t (X_s - y)^2 \, ds \right] \right),$$

where the first inequality follows via localisation and limiting argument using Fatou's lemma and monotone convergence. It now follows by Grönwall's lemma that

$$\mathbb{E}^y \left[(X_t - y)^2 \right] \leq 2C_\eta(1 + y^2) t \mathrm{e}^{2C_\eta t},$$

from which we deduce that

$$\mathbb{E}^y \left[L^x_t \right] = \mathbb{E}^y |X_t - x| - |x - y| \leq \sqrt{\mathbb{E}^y \left[(X_t - x)^2 \right]} - |x - y| \leq \sqrt{2C_\eta t} \, (1 + |y|) \mathrm{e}^{C_\eta t}. \tag{5.3}$$

Writing $\mathbb{E}^{\mu_0} \left[L^x_{t'} \right] - \mathbb{E}^{\mu_0} \left[L^x_t \right] = \mathbb{E}^{\mu_0} \left[\mathbb{E}^{X_t} \left[L^x_{t'-t} \right] \right]$ for $t < t' \leq T$, we see that

$$\mathbb{E}^{\mu_0} \left[L^x_{t'} \right] - \mathbb{E}^{\mu_0} \left[L^x_t \right] \leq \sqrt{2C_\eta(t' - t)} \, (1 + m_{\mu_0}(t)) \mathrm{e}^{C_\eta(t' - t)} \leq \sqrt{2C_\eta(t' - t)} \, (1 + m^{\mu_0}(T)) \mathrm{e}^{C_\eta(t' - t)}$$

and we deduce that $v^\xi(x, t)$ is $\frac{1}{2}$-Hölder continuous on $[0, T]$. Equation (5.3) also provides the inequality

$$v^\xi(t, x) \geq U^{\mu_0}(x) - \mathbb{E}^{\mu_0} \left[L^x_t \right] \geq U^{\mu_0}(x) - \sqrt{2C_\eta t} \, (1 + m_{\mu_0}(t)) \mathrm{e}^{C_\eta t}.$$

It remains to compute $\mathcal{L} v^\xi$. First, since v^ξ is non-increasing in t and concave in x, the partial derivatives $\partial_t v^\xi$ and $D^2 v$ are well-defined as distributions on $\bar{\mathcal{I}}$, so $\mathcal{L} v^\xi$ makes sense in terms of measures.

We first consider the case where η is suitably differentiable (say smooth). Note that by a monotone convergence argument, we can restrict to the case where Y remains in a compact subinterval of $\mathcal{I}^\circ$ up to σ, and hence is bounded. Let $p(t,x,y)$ be the transition density for the diffusion and recall that $\mathbb{E}^y\left[L_t^x\right] = \eta(x)^2 \int_0^t p(r,y,x)dr$. It follows that for an arbitrary starting measure ν, we have $\mathbb{E}^\nu\left[L_t^x\right] = \int \nu(dy)\eta(x)^2 \int_0^t p(r,y,x)dr$, and we directly compute (using Kolmogorov's Forward Equation, which holds due to the smoothness assumption on η) that

$$\begin{aligned}
\mathcal{L}\mathbb{E}^\nu\left[L_t^x\right] &= \int \nu(dy)\Big(-\eta(x)^2 p(t,y,x) + \eta(x)^2 \int_0^t \frac{1}{2}D^2\left(\eta(x)^2 p(r,y,x)\right)\,dr\Big) \\
&= \int \nu(dy)\eta(x)^2\Big(-p(t,y,x) + \int_0^t \partial_t p(r,y,x)\,dr\Big) \\
&= -\eta(x)^2 \int \nu(dy)p(0,y,x) = -\eta(x)^2\nu(dx) = \frac{1}{2}\eta(x)^2 D^2 U^\nu(dx).
\end{aligned}$$

Suppose in addition that ξ has a smooth density with respect to Lebesgue measure (which we also denote by ξ). We then compute from (5.2) and the equation above that

$$\begin{aligned}
\mathcal{L}v^\xi(t,dx) &= \mathcal{L}\int_0^t \int_{\bar{\mathcal{I}}} \mathbb{E}^{s,y}\left[L_t^x\right]\xi(s,y)\,ds\,dy = \mathcal{L}\int_0^t \int_{\bar{\mathcal{I}}} \mathbb{E}^y\left[L_{t-s}^x\right]\xi(s,y)\,ds\,dy \\
&= -\int_0^t \int_{\bar{\mathcal{I}}} \eta(x)^2 \delta_{\{y\}}(dx)\xi(s,y)\,ds\,dy = -\int_0^t \eta(x)^2\xi(s,x)\,ds\,dx.
\end{aligned}$$

Applying Itô's lemma, we see that

$$\mathbb{E}^x\left[v^\xi(t-\sigma,Y_\sigma)\right] = v^\xi(t,x) - \mathbb{E}^x\left[\int_0^\sigma \int_0^s \eta(Y_s)^2\xi(r,Y_s)\,dr\,ds\right] \tag{5.4}$$
$$= v^\xi(t,x) - \int_0^t \int_{\bar{\mathcal{I}}} q_\sigma(t-s,y)\eta(y)^2 \int_0^s \xi(dr,dy)\,ds.$$

We now argue that our results hold for an arbitrary, locally Lipschitz function η. Keeping ξ fixed as above, with a smooth density, let η_n be a sequence of Lipschitz functions obtained from η by mollification. Note that since we are on a compact interval, η and hence η_n are all bounded and from the mollification, we may assume that there exists K such that η, η_n are all K-Lipschitz; moreover ξ is bounded on the corresponding compact time-space set.

Write Y^n for the solution to the SDE $dY_t^n = \eta^n(Y_t^n)\,dW_t$, and note in particular, by standard results for SDEs (e.g. (Protter, 2005, Theorem V.4.15)) that $\sup_{r\in[0,t]}|Y_r^n - Y_r| \to 0$ almost surely (possibly after restricting to a subsequence), and in $\mathcal{L}^1$ as $n \to \infty$. Hence, by bounded convergence, we get convergence of the corresponding expectations on the right-hand side of (5.4), as $n \to \infty$. In addition, writing $v^{\xi,n}$ for the functions corresponding to the diffusions Y^n, we see from the first half of the proof that the functions $v^{\xi,n}, v^\xi$ are 1-Lipschitz in x, and uniformly Hölder continuous in t, for some common Hölder coefficient. It follows from the Arzelà-Ascoli theorem that $v^{\xi,n}$ converge uniformly (possibly down a subsequence) to v^ξ. We deduce that (5.1) holds for general η and smooth ξ.

Finally, approximating the measure ξ by smooth measures through a mollification argument, and observing that the local times for the diffusion are jointly continuous in x and t (by (1.1) and the discussion preceeding this equation) we conclude that we can pass to the limit on the right-hand side of (5.2), and hence on the left-hand side of (5.1). On the other hand, when q_σ is continuous, we can also pass to the limit on the right-hand side of (5.1). Moreover we can approximate σ by a sequence of stopping times $\sigma^n \searrow \sigma$ such that q_{σ^n} has a continuous density, and this gives us the required result after a monotone convergence argument. $\qquad\square$

For the next statement, we introduce the processes

$$V^t := \left\{V_s^t := v^\xi(t-s,Y_s), s \in [0,t]\right\}, \quad t \in [0,\infty], \tag{5.5}$$

13

where V^∞ is defined through $v^\xi(\infty, x) = U^{\alpha^\xi}(x)$ as in (4.2), i.e. $V_s^\infty = U^{\alpha^\xi}(Y_s)$.

Lemma 5.2. *Let $\sigma^\xi \in \mathcal{T}$ with corresponding time-space distribution ξ. Then the processes V^t and $V^{t'} - V^t$ are $\mathbb{P}^x$-supermartingales for all $t \le t' \le \infty$, and $x \in \bar{\mathcal{I}}$.*

Proof. In this proof we will want to take expectations with respect to both the X and Y processes at the same time; we will assume that these are defined on a product space, where the processes are independent. Then we will denote expectation with respect to the X process alone by $\mathbb{E}_X^\mu[A]$, etc, and the filtrations generated by the respective processes by $\mathcal{F}_s^X$ and $\mathcal{F}_s^Y$.

We first prove the supermartingale property for the process V^t. The case $t = \infty$ is an immediate consequence of the Jensen inequality. Next, fix $t \in [0, \infty)$, and recall that $v^\xi(t, x) = -\mathbb{E}_X^{\mu_0}|X_{t \wedge \sigma^\xi} - x|$ for $t \ge 0, x \in \bar{\mathcal{I}}$. Then we need to show, for $0 \le u \le s$,

$$-\mathbb{E}_Y^{\nu_0}\left[\mathbb{E}_X^{\mu_0}\left[|X_{(t-s)\wedge\sigma^\xi} - Y_s|\right]\big|\mathcal{F}_u^Y\right] \le -\mathbb{E}_X^{\mu_0}\left|X_{(t-u)\wedge\sigma^\xi} - Y_u\right|.$$

Using Hunt's switching identity we have

$$\mathbb{E}_Y^y|x - Y_{s-u}| = \mathbb{E}_X^x|X_{s-u} - y|.$$

Using the Strong Markov property, and $\tilde{Y}, \tilde{\mathbb{E}}$ to denote independent copies of Y etc., we deduce

$$\begin{aligned}
-\mathbb{E}_Y^{\nu_0}\left[\mathbb{E}_X^{\mu_0}\left[|X_{(t-s)\wedge\sigma^\xi} - Y_s|\right]\big|\mathcal{F}_u^Y\right] &= -\mathbb{E}_X^{\mu_0}\left[\tilde{\mathbb{E}}_Y^{Y_u}\left|X_{(t-s)\wedge\sigma^\xi} - \tilde{Y}_{s-u}\right|\right] \\
&= -\mathbb{E}_X^{\mu_0}\left[\tilde{\mathbb{E}}_X^{X_{(t-s)\wedge\sigma^\xi}}\left|\tilde{X}_{(s-u)} - Y_u\right|\right] \\
&= -\mathbb{E}_X^{\mu_0}\left[\left|X_{(t-s)\wedge\sigma^\xi + (s-u)} - Y_u\right|\right] \\
&\le -\mathbb{E}_X^{\mu_0}\left[\left|X_{(t-u)\wedge\sigma^\xi} - Y_u\right|\right]
\end{aligned}$$

where, in the final line, we used Jensen's inequality and the fact that $(t-s) \wedge \sigma^\xi + (s-u) \ge (t-u) \wedge \sigma^\xi$.

Now suppose $t' \ge t$, and consider $V_s^{t'} - V_s^t$ for $0 \le s \le t \le t'$. A similar calculation to that above shows that for $u \le s$,

$$\begin{aligned}
\mathbb{E}_Y^{\nu_0}\left[V_s^{t'} - V_s^t|\mathcal{F}_u^Y\right] &= \mathbb{E}_X^{\mu_0}\left[\left|X_{(t-s)\wedge\sigma^\xi+(s-u)} - Y_u\right| - \left|X_{(t'-s)\wedge\sigma^\xi+(s-u)} - Y_u\right|\right] \\
&= \mathbb{E}_X^{\mu_0}\left[\left|X_{(t-u)\wedge\tilde{\sigma}} - Y_u\right| - \left|X_{(t'-u)\wedge\tilde{\sigma}} - Y_u\right|\right],
\end{aligned}$$

where $\tilde{\sigma} := \sigma^\xi + (s-u)$. Note that for any $r' > r$, the process $|X_{r\wedge u} - y| - |X_{r'\wedge u} - y|$ is a supermartingale for $u \ge 0$. It follows that, since $\sigma^\xi \le \tilde{\sigma}$,

$$\begin{aligned}
\mathbb{E}_Y^{\nu_0}\left[V_s^{t'} - V_s^t|\mathcal{F}_u^Y\right] &= \mathbb{E}_X^{\mu_0}\left[\left|X_{(t-u)\wedge\tilde{\sigma}} - Y_u\right| - \left|X_{(t'-u)\wedge\tilde{\sigma}} - Y_u\right|\right], \\
&\le \mathbb{E}_X^{\mu_0}\left[\left|X_{(t-u)\wedge\sigma^\xi} - Y_u\right| - \left|X_{(t'-u)\wedge\sigma^\xi} - Y_u\right|\right] \\
&= V_u^{t'} - V_u^t.
\end{aligned}$$

$\square$

5.2 The optimal stopping problem

In this section we derive some useful properties of the function $u^\beta(t, x)$. We first state some standard facts from the theory of optimal stopping. Introduce

$$\tau^t := \inf\{s \ge 0 : (t - s, Y_s) \in \mathcal{R}^\beta\} \wedge t, \quad \text{for all} \quad t \ge 0. \tag{5.6}$$

14

Proposition 5.3. *Let $\sigma^\xi \in \mathcal{T}$ with corresponding time-space distribution ξ, and $\alpha^\xi \preceq_{\mathrm{cx}} \beta$. Then, for all $(t,x) \in \mathcal{S}$, $\tau^t \in \mathcal{T}^t$ is an optimal stopping rule for the problem u^β in (4.3):*

$$u^\beta(t,x) = \mathbb{E}^x \left[v^\xi(t - \tau^t, Y_{\tau^t}) + w^\beta(Y_{\tau^t})\mathbf{1}_{\{\tau^t < t\}} \right], \tag{5.7}$$

and the process $\left(u^\beta(t-s, Y_s) \right)$ is a $\mathbb{P}^x$-martingale for $s \in [0, \tau^t]$ and a $\mathbb{P}^x$-supermartingale for $s \in [0,t]$.

Proof. Recall that under $\mathbb{P}^{t,x}$ the diffusion Y_r, $r \geq t$ departs from x at time t, and when $t = 0$, we write $\mathbb{P}^{0,x} = \mathbb{P}^x$. Then we have for $0 \leq s \leq t$:

$$u^\beta(t-s, x) = u^t(s, x) := \sup_{s \leq \tau \leq t} \mathbb{E}^{s,x} \left[v^\xi(t - \tau, Y_\tau) + w^\beta(Y_\tau)\mathbf{1}_{\{\tau < t\}} \right]. \tag{5.8}$$

Notice that $u^t(s, x)$ is a classical optimal stopping problem with maturity t, and obstacle $Z_s := v^\xi(t - s, Y_s) + w^\beta(Y_s)\mathbf{1}_{\{s < t\}}$, $s \in [0, t]$, satisfying the condition of upper semicontinuity under expectation, i.e. $\limsup_{n \to \infty} \mathbb{E}^x[Z_{\theta_n}] \leq \mathbb{E}^x[Z_\theta]$ for any monotone sequence of stopping times θ_n converging to θ. Under this condition, it is proved in El Karoui (1981) that the standard theory of optimal stopping holds true. In particular, the process $\left(u^\beta(t-s, Y_s) \right)_{s \leq t}$ satisfies the announced martingale and supermartingale properties, and an optimal stopping time for the problem $u^t(0, x) = u^\beta(t, x)$ is

$$t \wedge \inf \left\{ s \geq 0 : u^t(s, Y_s) = v^\xi(t - s, Y_s) + w^\beta(Y_s) \right\},$$

which is exactly τ^t. $\qquad\square$

Remark 5.4. *Note that, taking $\tau = t$ in (4.3), $u^\beta(t,x) \geq \mathbb{E}^x[U^{\mu_0}(Y_t)] = U^{\mu_0}(x) + \mathbb{E}^x \left[\int U^{\mu_0}{}''(dy)L_t^y \right] := U^t(x)$. Suppose $([0,t] \times \mathcal{I}^\circ) \cap \mathcal{R}^\beta = \emptyset$ then, from (5.7), $u^\beta(t,x) = U^t(x) > U^\beta(x)$ for all $x \in \mathbb{R}$. We now consider the cases where $\mathcal{I} = \mathbb{R}$, $\bar{\mathcal{I}} = (-\infty, b_{\mathcal{I}}]$, $b_{\mathcal{I}} < \infty$ and $\bar{\mathcal{I}} = [a_{\mathcal{I}}, b_{\mathcal{I}}]$ a finite interval separately.*

In the case where $\mathcal{I} = \mathbb{R}$, we have $\mathbb{E}^x[L_t^y] \to \infty$ as $t \to \infty$, for any $x, y \in \mathcal{I}$. As $U^t(x) \to -\infty$ for all x as $t \to \infty$, it is impossible that $U^t(x) > U^\beta(x)$ for all $x \in \mathbb{R}$ and all $t \geq 0$. So there always exists $x \in \mathcal{I}$ with $\bar{t}^\beta(x) < \infty$ and hence $\mathcal{R}^\beta \neq \emptyset$.

Similarly consider the case where $\bar{\mathcal{I}} = (-\infty, b_{\mathcal{I}}]$. From the properties of the diffusion, we know that $X_t \to b_{\mathcal{I}}$ almost surely as $t \to \infty$. Moreover, since $\mathbb{E}^x[X_t] = x$ and $-|x| \geq U^{\mu_0}(x) \geq -|x| - c$, for some $c \in [0, \infty)$, we must have $U^t(x) \to -b_{\mathcal{I}} + (x - b_{\mathcal{I}})$ as $t \to \infty$ for $x \in \bar{\mathcal{I}}$. Since β is centred, $|U^\beta(x) + |x|| \to 0$ as $x \to -\infty$, and hence we cannot have $U^t(x) > U^\beta(x)$ for all $x \in \mathcal{I}^\circ$ and all $t \geq 0$. Hence there always exists $x \in \mathcal{I}^\circ$ with $\bar{t}^\beta(x) < \infty$ and hence $\mathcal{R}^\beta \neq \emptyset$.

Finally consider the case where $\bar{\mathcal{I}} = [a_{\mathcal{I}}, b_{\mathcal{I}}]$. Hence $\lim_{t \to \infty} X_t \in \{a_{\mathcal{I}}, b_{\mathcal{I}}\}$, and a similar argument to above gives $U^t(x) \to a_{\mathcal{I}} - \frac{x - a_{\mathcal{I}}}{b_{\mathcal{I}} - a_{\mathcal{I}}}(b_{\mathcal{I}} + a_{\mathcal{I}})$ as $t \to \infty$, for $x \in \bar{\mathcal{I}}$. This limit corresponds to $U^{\tilde{\nu}}(x)$, where $\tilde{\nu}$ is the centred measure supported on $\{a_{\mathcal{I}}, b_{\mathcal{I}}\}$, and it is easy to check that this potential is strictly smaller than the potential of any other centred measure supported on $\bar{\mathcal{I}}$, and so for any other measure, there always exists $x \in \mathcal{I}^\circ$ with $\bar{t}^\beta(x) < \infty$ and hence $\mathcal{R}^\beta \neq \emptyset$. The case of the measure $\tilde{\nu}$ is trivial, and we exclude this from subsequent arguments.

Lemma 5.5. *Let $\sigma^\xi \in \mathcal{T}$ with corresponding time-space distribution ξ, and $\alpha^\xi \preceq_{\mathrm{cx}} \beta$. Then:*

(i) *the function u^β is 1-Lipschitz-continuous in x, non-increasing and u^β is $\frac{1}{2}$-Hölder-continuous in t, and there is a constant C which is independent of β such that $|u^\beta(t,x) - u^\beta(t',x)| \leq C(1 + |x|)\sqrt{|t - t'|}$;*

(ii) *$u^\beta - v^\xi$ is non-increasing in t; in particular, u^β is non-increasing in t and concave in x;*

(iii) *$u^\beta(0, .) = U^{\mu_0}$, $U^\beta \leq v^\xi + w^\beta \leq u^\beta \leq v^\xi$, and $u^\beta(t, .) \searrow U^\beta$ pointwise as $t \nearrow \infty$.*

Proof. (i) The 1-Lipschitz-continuity of $u^\beta(t,x)$ in x follows directly from the Lipschitz continuity of v^ξ and w^β in x. Then, the $\frac{1}{2}$-Hölder continuity in t follows by standard arguments using the dynamic

programming principle (for example, as a simple modification of the proof of Proposition 2.7 in Touzi (2012)).

(ii) Let $t' > t$, fix $\varepsilon > 0$, and let $\tau' \in \mathcal{T}^{t'}$ be such that

$$u^\beta(t', x) - \varepsilon \leq \quad \mathbb{E}^x \left[v^\xi(t' - \tau', Y_{\tau'}) + w^\beta(Y_{\tau'}) \mathbf{1}_{\{\tau' < t'\}} \right] = \mathbb{E}^x \left[V_{\tau'}^{t'} + w^\beta(Y_{\tau'}) \mathbf{1}_{\{\tau' < t'\}} \right].$$

Recall from Lemma 5.2 the supermartingale properties of the process V^t introduced in (5.5). Then

$$\mathbb{E}^x \left[V_{\tau'}^{t'} \right] \ \leq \ \mathbb{E}^x \left[V_{t \wedge \tau'}^{t'} \right] \ = \ \mathbb{E}^x \left[V_{t \wedge \tau'}^{t'} - V_{t \wedge \tau'}^t \right] + \mathbb{E}^x \left[V_{t \wedge \tau'}^t \right] \ \leq \ V_0^{t'} - V_0^t + \mathbb{E}^x \left[V_{t \wedge \tau'}^t \right].$$

In addition, since $w^\beta \leq 0$, we have:

$$\mathbb{E}^x \left[w^\beta(Y_{\tau'}) \mathbf{1}_{\{\tau' < t'\}} \right] \ \leq \ \mathbb{E}^x \left[w^\beta(Y_{\tau'}) \mathbf{1}_{\{\tau' < t\}} \right] \ = \ \mathbb{E}^x \left[w^\beta(Y_{\tau' \wedge t}) \mathbf{1}_{\{\tau' < t\}} \right].$$

Putting these together, we conclude that

$$u^\beta(t', x) - v^\xi(t', x) - \varepsilon \ \leq \ \mathbb{E}^x \left[V_{t \wedge \tau'}^t + w^\beta(Y_{\tau' \wedge t}) \mathbf{1}_{\{\tau' < t\}} \right] - v^\xi(t, x) \ \leq \ u^\beta(t, x) - v^\xi(t, x).$$

By the arbitrariness of $\varepsilon > 0$, this shows $u^\beta - v^\xi$ is non-increasing in t, and implies that u^β inherits from v^ξ the non-increase in t. By the supermartingale property of the process $\left(u^\beta(t - s, Y_s) \right)_{s \in [0, t]}$ in Proposition 5.3, this in turns implies that u^β is concave in x.

(iii) By definition, $u^\beta(0, x) = v^\xi(0, x) = U^{\mu_0}(x)$. Since $v^\xi(t, x) \geq U^{\alpha^\xi}(x)$, we have $u^\beta(t, x) \geq v^\xi(t, x) + w^\beta(x) \geq U^\beta(x)$. On the other hand, since $w^\beta(x) \leq 0$, we have $u^\beta(t, x) \leq \sup_{\tau \leq t} \mathbb{E}^x \left[v^\xi(t - \tau, Y_\tau) \right] \leq v^\xi(t, x)$ by the supermartingale property of V^t established in the previous Lemma 5.2.

In the rest of this proof, we show that $u^\beta(t, x) \to U^\beta(x)$ as $t \to \infty$ for all $x \in \bar{\mathcal{I}}$. We consider three cases:

- Suppose $(t_0, x) \in \mathcal{R}^\beta$ for some $t_0 \geq 0$. Then, for any $t \geq t_0$, $\tau^t = 0$ and $u^\beta(t, x) = v^\xi(t, x) + w^\beta(x)$ which converges to $U^{\alpha^\xi}(x) + w^\beta(x) = U^\beta(x)$, as $t \to \infty$.

- Suppose that $(t_n, x_n) \in \mathcal{R}^\beta$ for some sequence $(t_n, x_n)_{n \geq 1}$ with $x_n \to x$. Then it follows from the previous case that $u^\beta(t, x_n) \to U^\beta(x_n)$, and therefore $u^\beta(t, x) \to U^\beta(x)$ by the Lipschitz-continuity of u^β.

- Otherwise, suppose that $[0, \infty] \times (x - \varepsilon, x + \varepsilon)$ does not intersect $\mathcal{R}^\beta$ for some $\varepsilon > 0$. Let $(a_x, b_x) := \cup (a, b)$ over all $a \leq x - \varepsilon < x + \varepsilon \leq b$ such that $[0, \infty] \times (a, b)$ does not intersect $\mathcal{R}^\beta$. By Remark 5.4, we may assume $\mathcal{R}^\beta$ is not empty and hence $(a_x, b_x) \neq \mathcal{I}^\circ$. In the subsequent argument, we assume that a_x is finite, the case where b_x is finite follows by the same line of argument. The optimal stopping time τ^t in (5.6) satisfies $\tau^t \geq H_{a_x, b_x} := \inf\{r \geq 0 : Y_t \notin (a_x, b_x)\}$ and $\tau^t \to H_{a_x, b_x}$, $\mathbb{P}^x$-almost surely. If both a_x and b_x are finite, we use the inequality $u^\beta(t, x) \geq U^\beta(x)$, together with Fatou's Lemma, Lemmas 5.1 and 5.2, and bounded convergence, to see that

$$
\begin{aligned}
U^\beta(x) \leq \lim_{t \to \infty} u^\beta(t, x) &= \lim_{t \to \infty} \mathbb{E}^x \left[v^\xi(t - \tau^t, Y_{\tau^t}) + w^\beta(Y_{\tau^t}) \right] \\
&\leq \lim_{t \to \infty} \mathbb{E}^x \left[v^\xi(t - H_{a_x, b_x}, Y_{H_{a_x, b_x}}) \right] + \mathbb{E}^x \left[w^\beta(Y_{H_{a_x, b_x}}) \right] \qquad (5.9) \\
&\leq \mathbb{E}^x \left[\lim_{t \to \infty} v^\xi(t - H_{a_x, b_x}, Y_{H_{a_x, b_x}}) + w^\beta(Y_{H_{a_x, b_x}}) \right] = \mathbb{E}^x \left[U^\beta(Y_{H_{a_x, b_x}}) \right] \leq U^\beta(x).
\end{aligned}
$$

Hence $\lim_{t \to \infty} u^\beta(t, x) = U^\beta(x)$, and U^β is linear on (a_x, b_x).

For the general case where b_x may be infinite, a more careful argument is needed. Since $w^\beta := (U^\beta - U^{\alpha^\xi})(x) \to 0$ as $|x| \to 0$, it follows that $\delta := \max(-w^\beta) < \infty$. Fix $\varepsilon > 0$ and choose c sufficiently large that $\delta/(c - a_x) < \varepsilon$. Let $H_c := \inf\{s \geq 0 : Y_s \geq c\}$ and note that $\tau^t \wedge H_c \to H_{a_x, c} = \inf\{t \geq 0 : Y_t \notin (a_x, c)\}$

16

as $t \to \infty$. Then by the martingale property of u^β on $t \le \tau^t$, and the fact that $u^\beta \le v^\xi$, we have

$$
\begin{aligned}
u^\beta(t,x) &= \mathbb{E}^x \left[u^\beta(t - \tau^t \wedge H_c, Y_{\tau^t \wedge H_c}) \right] \\
&\le \mathbb{E}^x \left[\mathbf{1}_{\{\tau^t \le H_c\}} (v^\xi + w^\beta)(t - \tau^t \wedge H_c, Y_{\tau^t \wedge H_c}) + \mathbf{1}_{\{\tau^t > H_c\}} v^\xi(t - \tau^t \wedge H_c, Y_{\tau^t \wedge H_c}) \right] \\
&\le \mathbb{E}^x \left[v^\xi(t - \tau^t \wedge H_c, Y_{\tau^t \wedge H_c}) + w^\beta(Y_{\tau^t \wedge H_c}) \mathbf{1}_{\{\tau^t \wedge H_c < t\}} \right] + \delta \mathbb{P}^x[\tau^t > H_c],
\end{aligned}
$$

where we wrote $w^\beta(t,x) = w^\beta(x)$. Taking limits as $t \to \infty$, and using Fatou as above, it follows from the definition of c that:

$$
U^\beta(x) \;\le\; \lim_{t \to \infty} u^\beta(t,x) \;\le\; \mathbb{E}^x \left[U^\beta(Y_{H_{a_x,c}}) \right] + \varepsilon \;=\; \frac{x - a_x}{c - a_x} U^\beta(c) + \frac{c - x}{c - a_x} U^\beta(a_x) + \varepsilon. \tag{5.10}
$$

Taking $\varepsilon \searrow 0$ and using concavity of U^β we get that $\lim_{t \to \infty} u^\beta(t,x) = U^\beta(x)$, and U^β is linear on (a_x, c). Letting $c \to \infty$ we conclude that U^β is linear on (a_x, ∞). $\qquad\square$

5.3 Existence and basic properties of the barrier

We denote the barrier function corresponding to the regular barrier $\mathcal{R}^\beta$ defined in (4.4) with $\bar{t}^\beta := \bar{t}_{\mathcal{R}^\beta}$. It will be used on many occasions in our proofs. Recall from (2.3) the definition of the support of a measure μ_k in terms of the measure μ_{k-1}. In what follows, we write ℓ^β, r^β for the bounds of the support of β in terms of the measure α^ξ.

Corollary 5.6. *Let $\sigma^\xi \in \mathcal{T}$ with corresponding time-space distribution ξ, and $\alpha^\xi \preceq_{\mathrm{cx}} \beta$. Then, the set $\mathcal{R}^\beta$ is a (closed) barrier, and moreover*
  (i) $\left([0,\infty] \times (\ell^\beta, r^\beta)^c\right) \subset \mathcal{R}^\beta$;
  (ii) $\mathcal{R}^\beta \cap ([0,\infty] \times (a,b)) = \emptyset$ *if and only if* $\beta[(a,b)] = 0$ *and* $w^\beta < 0$ *on* (a,b);
  (iii) $\bar{t}^\beta(x) = 0$ *if and only if* $w^\beta(x) = 0$.

Proof. For $(t,x) \in \mathcal{R}^\beta$, we have $u^\beta(t,x) = v^\xi(t,x) + w^\beta(x)$ and it is then immediate from (iii) and (ii) of Lemma 5.5 that $u^\beta(t',x) = v^\xi(t',x) + w^\beta(x)$ and so $(t',x) \in \mathcal{R}^\beta$, for all $t' > t$. By the continuity of v^ξ and u^β, established in Lemmas 5.1 and 5.5, we conclude that $\mathcal{R}^\beta$ is a closed barrier.

(i) For $x \notin (\ell^\beta, r^\beta)$, we have $U^{\alpha^\xi}(x) = U^\beta(x)$ and hence $w^\beta(x) = 0$. It follows from Lemma 5.5 (iii) that $u^\beta(t,x) = v^\xi(t,x)$ and hence $(t,x) \in \mathcal{R}^\beta$ for all $t \ge 0$ so that $[0,\infty] \times (\ell^\beta, r^\beta)^c \subset \mathcal{R}^\beta$.

(ii) In the proof of Lemma 5.5 (iii), it was shown that the condition $\mathcal{R}^\beta \cap ([0,\infty] \times (a,b)) = \emptyset$ implies that U^β is linear on (a,b), i.e. $\beta[(a,b)] = 0$, see (5.10). Moreover, the last argument in (i) above also implies that $w^\beta(x) < 0$ for all $x \in (a,b)$ whenever $\mathcal{R}^\beta \cap ([0,\infty] \times (a,b)) = \emptyset$. This provides the implication $\Longrightarrow$.

Suppose now that $\beta[(a,b)] = 0$ and $w^\beta < 0$ on (a,b). For fixed $x \in (a,b)$, we have:

$$
\begin{aligned}
u^\beta(t,x) &\ge \mathbb{E}^x \left[v^\xi(t - H_{a,b} \wedge t, Y_{H_{a,b} \wedge t}) + w^\beta(Y_{H_{a,b} \wedge t}) \mathbf{1}_{\{H_{a,b} < t\}} \right] \\
&> \mathbb{E}^x \left[v^\xi(t - H_{a,b} \wedge t, Y_{H_{a,b} \wedge t}) + w^\beta(Y_{H_{a,b} \wedge t}) \right] \\
&\ge v^\xi(t,x) - U^{\alpha^\xi}(x) + U^\beta(x) = v^\xi(t,x) + w^\beta(x).
\end{aligned}
$$

Here we have used the strict inequality $w^\beta(y) < 0$ for all $y \in (a,b)$ to get the second line. To get the final line, we use Lemma 5.2 to deduce that $\mathcal{L}v^\xi(t,dx) = -\eta(x)^2 \int_0^t \xi(ds,dx) \ge -\eta(x)^2 \alpha^\xi(dx) = \mathcal{L}U^{\alpha^\xi}(dx)$, and hence that $v^\xi(t-s, Y_s) + w^\beta(Y_s)$ is a submartingale up to $H_{a,b} \wedge t$, given that $U^\beta(x)$ is linear on (a,b).

This shows that $u^\beta(t,x) > v^\xi(t,x) + w^\beta(x)$, and hence $(t,x) \notin \mathcal{R}^\beta$, for all $t \ge 0$, and $x \in (a,b)$.

(iii) If $w^\beta(x) = 0$ then $u^\beta(t,x) = v^\xi(t,x)$ for all t, by (iii) of Lemma 5.5, and so $(t,x) \in \mathcal{R}^\beta$ for all $t \ge 0$. Recalling that $v^\xi(0,x) = u^\beta(0,x) = U^{\mu_0}(x)$, we conclude that $(0,x) \in \mathcal{R}^\beta$ only if $w^\beta(x) = 0$. $\qquad\square$

17

Remark 5.7 (On $\mathcal{R}^\beta$ having rays for arbitrary large $|x|$). *We can now deduce from the proof of the convergence $u^\beta \searrow U^\beta$, as $t \nearrow \infty$ in Lemma 5.5 (iii), that for any $N > 0$ there exist $x \le (-N) \vee a_\mathcal{I} < N \wedge b_\mathcal{I} \le y$ such that $\bar{t}^\beta(x) < \infty$ and $\bar{t}^\beta(y) < \infty$. In the proof, we show that for any point x such that $\bar{t}^\beta(x) = \infty$ either there exists points $a < x < b$ such that $\bar{t}^\beta(a), \bar{t}^\beta(b) < \infty$ or there exists an a less than x such that for any c large enough U^β is linear on (a,c). Letting $c \to \infty$, and using the fact that $U^\beta(c) + |c| \to 0$, we conclude that $U^\beta(y) = -|y|$ for all $y \ge a$. Then $U^\beta(y) \le U^{\alpha^\xi}(y) \le U^{\mu_0}(y) \le -|y| = U^\beta(y)$ implies $U^\beta(y) = U^{\alpha^\xi}(y)$. In particular, $w^\beta(x) = 0$, and by Corollary 5.6 we contradict the initial assumption that x is not in the barrier.*

Remark 5.8 (On the structure of the stopping region). *Let α^ξ, β be integrable measures in convex order. It follows from Corollary 5.6 that the barrier can be divided into at most countably many (possibly infinite) non-overlapping open intervals $J_1, J_2, J_3, \ldots$ such that $J_k = (a_k, b_k)$, for $a_k < b_k$, on which $\bar{t}^\beta(x) > 0$ for all $x \in (a_k, b_k)$ and $\left(\left(\bigcup_{k=1}^\infty J_k \right)^{\complement} \times [0, \infty] \right) \subseteq \mathcal{R}^\beta$.*

Observing that in both the embedding, and the optimal stopping perspectives, the process started from $x \in J_k$ never exits each interval J_k, it is sufficient to consider each interval separately, noting that in such a case, $u^\beta(t,x) = v^\xi(t,x)$ for all $t \ge 0$, and all $x \in \left(\bigcup_{k=1}^\infty J_k \right)^{\complement}$. In the subsequent argument, we will assume that we are on a single such interval J_k, which may then be finite, semi-infinite, or equal to $\mathcal{I}^\circ$. In addition, if the measures α^ξ, β are in convex order, then their restrictions to each J_k are also in convex order.

Remark 5.9 (On $\mathcal{R}^\beta$ for atomic measures). *Let α^ξ, β be integrable measures in convex order. Bearing in mind Remark 5.8, we suppose that β is a probability measure on $\bar{\mathcal{I}}$ such that for some integer $n' \ge 1$, and some ordered scalars $x_1 < \ldots < x_{n'}$, we have $\sum_{i=1}^{n'} \beta[\{x_i\}] = 1$ and $\beta[\{x_i\}] > 0$ for all $i = 1, \ldots, n'$. From the representation of the optimal stopping time τ^t, see Proposition 5.3 above, and the form of the set $\mathcal{R}^\beta$ implied by Corollary 5.6, it follows that*

$$u^\beta(t,x) \quad = \quad \sup_{\tau \in \mathcal{T}(x_1, \ldots, x_{n'})} \mathbb{E}^x \left[v^\xi(t - \tau, Y_\tau) + w^\beta(Y_t) \mathbf{1}_{\{\tau < t\}} \right], \tag{5.11}$$

where $\mathcal{T}(x_1, \ldots, x_{n'})$ is the set of stopping times τ such that $\tau \le H_{x_1, x_{n'}}$ and $Y_\tau \in \{x_1, \ldots x_{n'}\}$ a.s.

6 Locally finitely supported measures

A probability measure β is said to be α^ξ–locally finitely supported if its support intersects any compact subset of $\mathrm{supp}(\alpha^\xi, \beta) = \overline{\{x : U^{\alpha^\xi}(x) > U^\beta(x)\}}$ at a finite number of points. The measure β is α^ξ–finitely supported if its support intersects $\mathrm{supp}(\alpha^\xi, \beta)$ at a finite number of points. Throughout, α^ξ will be fixed, so we will typically only refer to (locally) finitely supported measures. Observe that an integrable, centred measure β can only be finitely supported if ℓ^β and r^β are both finite — indeed, in this case a locally finitely supported measure is finitely supported if and only if r^β and ℓ^β are both finite.

6.1 Preparation

We start with two preliminary results which play crucial roles in the next section where we establish the main result for finitely supported measures. The first result is the key behind the time-reversal methodology which underpins the main results, see Section 3.1. Here, we give a natural proof in the case where $X = B$ is a Brownian motion, when the proof has a simple intuition[3]. In Appendix B we give a

---

[3]Given its importance, we have discussed this result with many colleagues. Our first proof used an explicit formula for the density p_c in $\mathbb{P}_c^x[B_s \in \mathrm{d}y, s < H_{0,c}] = p_c(s, x, y)\mathrm{d}y$, see Proposition 2.8.10 p.98 in Karatzas and Shreve (1991). The current proof uses a clever coupling trick devised by Tigran Atoyan.

PDE proof which works in the more general diffusion setting.

To understand the importance of the result, it is helpful to think of the local time of X and of Y on the two sides of the announced equality. This result is then used to obtain the key equality $v^{\xi^\beta} = u^\beta$ in a "box" setting where the barrier is locally composed of two rays. The case of finitely supported measures is then obtained with an inductive argument in Section 6.2.

Lemma 6.1. *Let L be the local time of a Brownian motion B. For any $a < x < y < b$ and $t \geq 0$ we have $\mathbb{E}^x\big[L^y_{t \wedge H_{a,b}}\big] = \mathbb{E}^y\big[L^x_{t \wedge H_{a,b}}\big]$.*
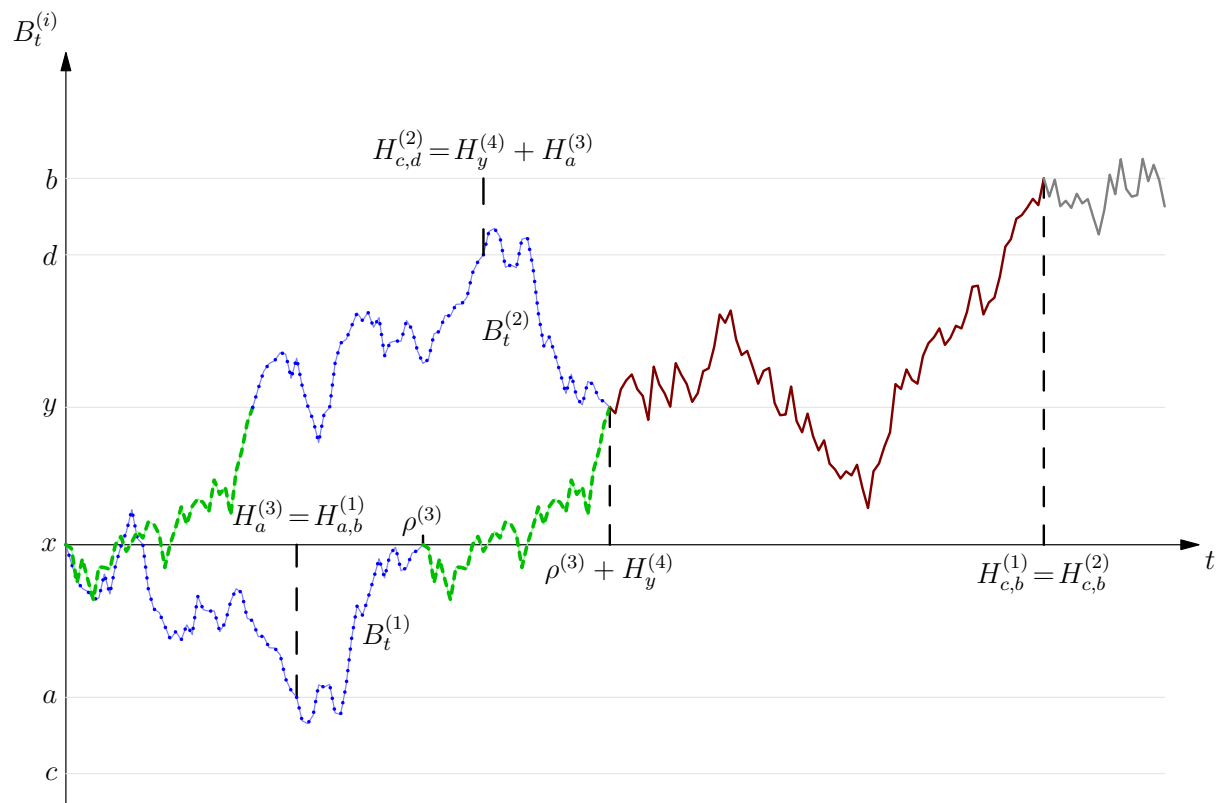

Figure 2: A depiction of the Brownian motions $B^{(1)}$ and $B^{(2)}$ constructed in the proof of Lemma 6.1. Observe that the blue and green sections in each process are mirror images, up to translation, while the magenta sections are equal, up to translation.

Proof. Without loss of generality we suppose $b - y > x - a$ and introduce two additional points $c = x - (b - y)$ and $d = y + (x - a)$ so that $c < a < x < y < d < b$ with $b - d = a - c$. Note that by translation invariance and symmetry of Brownian motion we have

$$\mathbb{E}^y\big[L^x_{t \wedge H_{a,b}}\big] = \mathbb{E}^x\big[L^y_{t \wedge H_{c,d}}\big].$$

Using this in the desired equality, and subtracting $\mathbb{E}^x\big[L^y_{t \wedge H_{c,b}}\big]$, we see that it suffices to show that

$$\mathbb{E}^x\big[L^y_{t \wedge H_{c,b}} - L^y_{t \wedge H_{a,b}}\big] = \mathbb{E}^x\big[L^y_{t \wedge H_{c,b}} - L^y_{t \wedge H_{c,d}}\big].$$

Finally, by shift invariance, we may suppose without loss of generality that $x = 0$. Consider three independent Brownian motions $B^{(3)}, B^{(4)}, B^{(5)}$ starting from 0 and denote $H^{(i)}$ the hitting times for

$B^{(i)}$. Further, let $\rho^{(3)} = \inf\{t \geq H_a^{(3)} : B_t^{(3)} = 0\}$. Define two new processes

$$B_t^{(1)} := B_{t \wedge \rho^{(3)}}^{(3)} + B_{(t-t \wedge \rho^{(3)}) \wedge H_y^{(4)}}^{(4)} + B_{t-t \wedge (\rho^{(3)}+H_y^{(4)})}^{(5)}$$

$$B_t^{(2)} := B_{t \wedge H_y^{(4)}}^{(4)} - B_{(t-t \wedge H_y^{(4)}) \wedge \rho^{(3)}}^{(3)} + B_{t-t \wedge (\rho^{(3)}+H_y^{(4)})}^{(5)}$$

(6.1)

and observe these are standard Brownian motions. This construction is depicted in Figure 2. We denote $L^{y,(i)}$ the local time of $B^{(i)}$ at level y.

Recall that $c < a < d < b$ and consider $L_{t \wedge H_{c,b}^{(1)}}^{y,(1)} - L_{t \wedge H_{a,b}^{(1)}}^{y,(1)}$. For this quantity to be non-zero the following have to happen prior to t: first $B^{(1)}$ has to hit a without reaching b, then it has to come back to $x = 0$ and continue to y without ever reaching c. This happens at time $\rho^{(3)} + H_y^{(4)}$ and from then onwards the local time $L^{y,(1)}$ is counted before time $t \wedge H_{c,b}^{(1)}$ and we see that it simply corresponds to $L^{0,(5)}$. With a similar reasoning for $L^{y,(2)}$, we see that our construction gives us the desired coupling:

$$L_{t \wedge H_{c,b}^{(1)}}^{y,(1)} - L_{t \wedge H_{a,b}^{(1)}}^{y,(1)} = L_{(t-t \wedge (\rho^{(3)}+H_y^{(4)})) \wedge H_{c-y,b-y}^{(5)}}^{0,(5)} = L_{t \wedge H_{c,b}^{(2)}}^{y,(2)} - L_{t \wedge H_{c,d}^{(2)}}^{y,(2)}$$

and taking expectations gives the required result. □

We now prove an important consequence of the above result, which will form the basis of an induction argument.

Lemma 6.2. *Let $\sigma^\xi \in \mathcal{T}$ with corresponding time-space distribution ξ, and $\alpha^\xi \preceq_{\mathrm{cx}} \beta$. Let $a < b$ and $t_0 > 0$ be such that $[t_0, \infty] \times \{a, b\} \subset \mathcal{R}^\beta$, $(0, \infty) \times (a, b) \cap \mathcal{R}^\beta = \emptyset$, and $(v^{\xi^\beta} - u^\beta)(t_0, \cdot) = 0$ on $[a, b]$. Then $v^{\xi^\beta} - u^\beta = 0$ on $[t_0, \infty) \times [a, b]$.*

Proof. In view of Remark 4.3, and the continuity of $v^{\xi^\beta} - u^\beta$, it is sufficient to show that

$$v^\xi(t, x) - u^\beta(t, x) + u^\beta(t_0, x) - v^\xi(t_0, x) = \mathbb{E}^\xi\left[L_{t \wedge \sigma_{\mathcal{R}^\beta}}^x\right] - \mathbb{E}^\xi\left[L_{t_0 \wedge \sigma_{\mathcal{R}^\beta}}^x\right] \text{ for } t \geq t_0, \ x \in (a, b). \quad (6.2)$$

We fix $x \in (a, b)$. Since $[t_0, \infty] \times \{a, b\} \subset \mathcal{R}^\beta$, $(0, \infty) \times (a, b) \cap \mathcal{R}^\beta = \emptyset$, we have the decomposition

$$\begin{aligned}
\mathbb{E}^\xi\left[L_{t \wedge \sigma_{\mathcal{R}^\beta}}^x\right] - \mathbb{E}^\xi\left[L_{t_0 \wedge \sigma_{\mathcal{R}^\beta}}^x\right] &= \mathbb{E}^\xi\left[\left(L_{t \wedge \sigma_{\mathcal{R}^\beta}}^x - L_{t_0 \wedge \sigma_{\mathcal{R}^\beta}}^x\right)\mathbf{1}_{\{T_\xi < t_0\}}\right] \\
&\quad + \mathbb{E}^\xi\left[\left(L_{t \wedge \sigma_{\mathcal{R}^\beta}}^x - L_{T_\xi \wedge t}^x\right)\mathbf{1}_{\{t_0 \leq T_\xi < t, X_{T_\xi} \in (a,b)\}}\right] \\
&= \mathbb{E}^\xi\left[\left(L_{t \wedge \sigma_{\mathcal{R}^\beta}}^x - L_{t_0 \wedge \sigma_{\mathcal{R}^\beta}}^x\right)\mathbf{1}_{\{T_\xi < t_0 < \sigma_{\mathcal{R}^\beta}\}}\right] \\
&\quad + \mathbb{E}^\xi\left[\left(L_{t \wedge H_{a,b}}^x - L_{T_\xi \wedge t}^x\right)\mathbf{1}_{\{t_0 \leq T_\xi < t, X_{T_\xi} \in (a,b)\}}\right] \\
&= \int_{(a,b)} \mathbb{E}^{(t_0, y)}\left[L_{t \wedge H_{a,b}}^x\right] m(dy) + \int_{[t_0, t]} \int_{(a,b)} \mathbb{E}^{(s,y)}\left[L_{t \wedge H_{a,b}}^x\right] \xi(ds, dy), (6.3)
\end{aligned}$$

where we introduced the measure $m(dy) := \mathbb{P}^\xi\left[X_{t_0} \in dy, T_\xi < t_0 < \sigma_{\mathcal{R}^\beta}\right]$, and used the fact that, conditional on starting in $\{t_0\} \times (a, b)$, the stopping times $\sigma_{\mathcal{R}^\beta}$ and $H_{a,b}$ are equal (and starting on $\{t_0\} \times (a, b)^{\complement}$, we never hit x before $\sigma_{\mathcal{R}^\beta}$). Observe that for $y \in (a, b)$, we have

$$\begin{aligned}
m(dy) + \xi(dy; s \geq t_0) &= \mathbb{P}^\xi\left[X_{t_0} \in dy, T_\xi < t_0 < \sigma_{\mathcal{R}^\beta}\right] + \mathbb{P}^\xi\left[X_{T_\xi} \in dy, T_\xi \geq t_0\right] \\
&= \mathbb{P}^\xi\left[X_{(t_0 \wedge \sigma_{\mathcal{R}^\beta}) \vee T_\xi} \in dy\right] =: \lambda(dy), \quad (6.4)
\end{aligned}$$

since $X_{\sigma_{\mathcal{R}^\beta}} \notin (a, b)$ by the assumptions on $\mathcal{R}^\beta$. Moreover, since σ^ξ is a UI embedding of α^ξ, it follows from the Tanaka formula that for $y \in (a, b)$, we have

$$U^\lambda(y) = U^{\alpha^\xi}(y) - \mathbb{E}^\xi\left[L_{t_0 \wedge \sigma_{\mathcal{R}^\beta}}^y\right] = U^{\alpha^\xi}(y) - (v^\xi - u^\beta)(t_0, y),$$

where the last equality follows from the assumption that $(v^{\xi^{\beta}} - u^{\beta})(t_0, .) = 0$ on $[a, b]$ together with Remark 4.3. Since $D^2 U^{\lambda}(dy) = \lambda(dy)$, this provides by substituting in (6.4) that for $y \in (a, b)$:

$$m(dy) = -\frac{1}{2} D^2 U^{\lambda}(y) dy - \xi(dy, s \geq t_0) = \frac{1}{2} D^2 \left(v^{\xi} - u^{\beta} \right)(t_0, dy) + \xi(dy, s < t_0).$$

Plugging this expression in (6.3), we get

$$\mathbb{E}^{\xi}\left[L^x_{t \wedge \sigma_{\mathcal{R}^{\beta}}} \right] - \mathbb{E}^{\xi}\left[L^x_{t_0 \wedge \sigma_{\mathcal{R}^{\beta}}} \right] = \frac{1}{2} \int_{(a,b)} \mathbb{E}^{(t_0, y)}\left[L^x_{t \wedge H_{a,b}} \right] D^2 (v^{\xi} - u^{\beta})(t_0, dy)$$
$$+ \int_{[0,t]} \int_{(a,b)} \mathbb{E}^{(s \vee t_0, y)}\left[L^x_{t \wedge H_{a,b}} \right] \xi(ds, dy).$$

The required result now follows from the following claims involving $\zeta := \inf\{s \geq 0 : (t - s, Y_s) \notin [0, t - t_0] \times (a, b)\}$:

$$\int_{(a,b)} \int_{[0,t]} \mathbb{E}^{(s \vee t_0, y)}\left[L^x_{t \wedge H_{a,b}} \right] \xi(ds, dy) = v^{\xi}(t, x) - \mathbb{E}^x\left[v^{\xi}(t - \zeta, Y_{\zeta}) \right], \tag{6.5}$$

$$\frac{1}{2} \int_{(a,b)} \mathbb{E}^{(t_0, y)}\left[L^x_{t \wedge H_{a,b}} \right] D^2 v^{\xi}(t_0, dy) = \mathbb{E}^x\left[v^{\xi}(t_0, Y_{\zeta}) \right] - v^{\xi}(t_0, x), \tag{6.6}$$

$$-\frac{1}{2} \int_{(a,b)} \mathbb{E}^{(t_0, y)}\left[L^x_{t \wedge H_{a,b}} \right] D^2 u^{\beta}(t_0, dy) = u^{\beta}(t_0, x) - u^{\beta}(t, x) + \mathbb{E}^x\left[v^{\xi}(t - \zeta, Y_{\zeta}) - v^{\xi}(t_0, Y_{\zeta}) \right], \tag{6.7}$$

which we now prove.

(i) To prove (6.5), we use Itô's formula (possibly after mollification) to get

$$v^{\xi}(t, x) = \mathbb{E}^x\left[v^{\xi}(t - \zeta, Y_{\zeta}) \right] + \mathbb{E}^x\left[\int_0^{\zeta} \mathcal{L} v^{\xi}(t - s, Y_s)\, ds \right].$$

Using Lemma 5.1 and writing $p_{\zeta}(r, x, y) dy := \mathbb{P}^x(Y_r \in dy, r < \zeta)$, this provides:

$$v^{\xi}(t, x) - \mathbb{E}^x\left[v^{\xi}(t - \zeta, Y_{\zeta}) \right] = \int_{y \in (a,b)} \int_0^{t - t_0} \eta(y)^2 p_{\zeta}(r, x, y)\, dr \left(-\int_0^{t - r} \xi(ds, dy) \right)$$
$$= \int_{y \in (a,b)} \int_{t_0}^t \eta(y)^2 p_{\zeta}(t - u, x, y)\, du \left(-\int_0^u \xi(ds, dy) \right)$$
$$= \int_{y \in (a,b)} \int_0^t \int_{t_0 \vee s}^t \eta(y)^2 p_{\zeta}(t - u, x, y)\, du\, \xi(ds, dy)$$
$$= \int_{y \in (a,b), s \in [0,t]} \mathbb{E}^{(s \vee t_0, y)}\left[L^x_{t \wedge H_{a,b}} \right] \xi(ds, dy).$$

(ii) We next prove (6.6). Since $v^{\xi}(t_0, .)$ is concave by Lemma 5.1, it follows from the Itô-Tanaka formula that:

$$\mathbb{E}^x\left[v^{\xi}(t_0, Y_{\zeta}) \right] - v^{\xi}(t_0, x) = \frac{1}{2} \int_{(a,b)} \mathbb{E}^x\left[L^y_{\zeta} \right] D^2 v^{\xi}(t_0, dy) = \frac{1}{2} \int_{(a,b)} \mathbb{E}^{(t_0, y)}\left[L^x_{t \wedge H_{a,b}} \right] D^2 v^{\xi}(t_0, dy),$$

where the last equality follows from Lemma 6.1 together with a coordinate shift.

(iii) Finally we turn to (6.7). Recall that $u^{\beta} = v^{\xi} + w^{\beta}$ on $[t_0, \infty) \times \{a, b\} \subset \mathcal{R}^{\beta}$. Then, since $Y_{\zeta} \in \{a, b\}$ on $\{\zeta < t - t_0\}$, we have:

$$u^{\beta}(t - \zeta, Y_{\zeta}) = u^{\beta}(t_0, Y_{\zeta}) \mathbf{1}_{\{\zeta = t - t_0\}} + \left(v^{\xi}(t - \zeta, Y_{\zeta}) + w^{\beta}(Y_{\zeta}) \right) \mathbf{1}_{\{\zeta < t - t_0\}}$$
$$= u^{\beta}(t_0, Y_{\zeta}) \mathbf{1}_{\{\zeta = t - t_0\}} + \left(v^{\xi}(t - \zeta, Y_{\zeta}) + w^{\beta}(Y_{\zeta}) \right) \mathbf{1}_{\{\zeta < t - t_0\}}$$
$$+ \left(v^{\xi}(t - \zeta, Y_{\zeta}) - v^{\xi}(t_0, Y_{\zeta}) \right) \mathbf{1}_{\{\zeta = t - t_0\}}$$
$$= u^{\beta}(t_0, Y_{\zeta}) \mathbf{1}_{\{\zeta = t - t_0\}} + v^{\xi}(t - \zeta, Y_{\zeta}) - v^{\xi}(t_0, Y_{\zeta})$$
$$+ \left(w^{\beta}(Y_{\zeta}) + v^{\xi}(t_0, Y_{\zeta}) \right) \mathbf{1}_{\{\zeta < t - t_0\}}$$
$$= u^{\beta}(t_0, Y_{\zeta}) + v^{\xi}(t - \zeta, Y_{\zeta}) - v^{\xi}(t_0, Y_{\zeta}).$$

21

We next use the fact that $[0,\infty] \times (a,b)$ does not intersect $\mathcal{R}^\beta$ to compute for $x \in (a,b)$ that

$$
\begin{aligned}
u^\beta(t,x) &= \mathbb{E}^x\left[u^\beta(t-\zeta,Y_\zeta)\right] \\
&= \mathbb{E}^x\left[u^\beta(t_0,Y_\zeta) + \left(v^\xi(t-\zeta,Y_\zeta) - v^\xi(t_0,Y_\zeta)\right)\right] \\
&= u^\beta(t_0,x) + \frac{1}{2}\mathbb{E}^x\left[\int_{(a,b)} L_\zeta^y D^2 u^\beta(t_0,dy)\right] + \mathbb{E}^x\left[v^\xi(t-\zeta,Y_\zeta) - v^\xi(t_0,Y_\zeta)\right],
\end{aligned}
$$

by application of the Itô-Tanaka formula, due to the concavity of the function $u^\beta(t,.)$, as established in Lemma 5.5. We finally conclude from Lemma 6.1/B.1 that

$$
u^\beta(t,x) = u^\beta(t_0,x) + \frac{1}{2}\int_{(a,b)} \mathbb{E}^{(t_0,y)}\left[L_{t\wedge H_{a,b}}^x\right] D^2 u^\beta(t_0,dy) + \mathbb{E}^x\left[v^\xi(t-\zeta,Y_\zeta) - v^\xi(t_0,Y_\zeta)\right].
$$

$\square$

6.2 The case of finitely supported measures

We now start the proof of Theorem 4.1 for a (relatively) finitely supported probability measure β. Recall from Lemma 4.2 and Lemma 5.5 (iii) that we need to prove that $u^\beta = v^{\xi^\beta}$. When there is no risk of confusion we write σ^β for $\sigma_{\mathcal{R}^\beta}$. In the sequel, we will say that β is α^ξ-supported on n points if the measure β restricted to (ℓ^β, r^β) is a discrete measure, supported on n points.

Proposition 6.3. *Let $\sigma^\xi \in \mathcal{T}$ with corresponding time-space distribution ξ, and β an α^ξ-finitely supported measure such that $\alpha^\xi \preceq_{cx} \beta$. Then $u^\beta = v^{\xi^\beta}$ and Theorem 4.1 holds for β.*

The proof proceeds by induction on the number of points in the support of $\beta|_{(\ell^\beta, r^\beta)}$. The case where $\alpha^\xi = \beta$ is trivial, since it follows immediately from (iii) of Corollary 5.6 that $\mathcal{R}^\beta = \mathcal{S}$. Hence we suppose that $\ell^\beta < r^\beta$. We start with the case where $\beta|_{(\ell^\beta, r^\beta)}$ contains no points, and therefore all mass starting in (ℓ^β, r^β) under ξ will be embedded at the two points ℓ^β, r^β.

Lemma 6.4. *Let $\sigma^\xi \in \mathcal{T}$ with corresponding time space distribution ξ, and $\alpha^\xi \preceq_{cx} \beta$ with $\beta((\ell^\beta, r^\beta)) = 0$. Then $v^{\xi^\beta} = u^\beta$ holds for all $(t,x) \in \mathcal{S}$.*

Proof. Note first that the convex ordering of β and α^ξ implies that $\alpha^\xi([\ell^\beta, r^\beta]) = \beta([\ell^\beta, r^\beta])$. Moreover, as we ruled out the case $\beta = \alpha^\xi$ and U^β is linear on (ℓ^β, r^β), we have $U^{\alpha^\xi}(x) > U^\beta(x)$ for all $x \in (\ell^\beta, r^\beta)$. It then follows from Corollary 5.6 that $\mathcal{R}^\beta = \mathbb{R}_+ \times (\bar{\mathcal{I}} \setminus (\ell^\beta, r^\beta))$ and $\sigma^\beta = \inf\{t \geq 0 : Y_t \notin (\ell^\beta, r^\beta)\}$ is the first hitting time of $(\bar{\mathcal{I}} \setminus (\ell^\beta, r^\beta))$. The result now follows from an application of Lemma 6.2. $\square$

The proof of Proposition 6.3 will be complete when we establish that the following induction step works.

Lemma 6.5. *Let $\sigma^\xi \in \mathcal{T}$ with time-space distribution ξ. Assume $v^{\xi^\beta} = u^\beta$ for any $\beta \succeq_{cx} \alpha^\xi$ which is α^ξ-supported on n points. Then, $v^{\xi^\beta} = u^\beta$ for any measure β which is α^ξ-supported on $n+1$ points.*

Proof. Let β be a centred probability measure α^ξ-supported on the $n+1$ ordered points $\mathbf{x} := \{x_1, \ldots, x_{n+1}\}$, with $\beta[\{x_i\}] > 0$ for all $i = 1, \ldots, n+1$. By Remark 5.9, the set $\mathcal{R}^\beta$ is of the form

$$
\mathcal{R}^\beta = \left([0,\infty] \times (\bar{\mathcal{I}} \setminus (\ell^\beta, r^\beta))\right) \bigcup_{i=1}^{n+1} \left([t_i, \infty) \times \{x_i\}\right) \quad \text{for some} \quad t_1, \ldots, t_{n+1} > 0.
$$

Let j be such that $t_j = \max_i t_i$, so that $[t_j, \infty) \times \{x_j\}$ is a horizontal ray in $\mathcal{R}^\beta$ starting farthest away from zero. Define a centred probability measure α^ξ-supported on $\mathbf{x}^{(-j)} := \mathbf{x} \setminus \{x_j\}$ by conveniently distributing the mass of β at x_j among the closest neighboring points:

$$
\beta^* = \beta + \beta[\{x_j\}]\left(-\delta_{\{x_j\}} + \frac{x_{j+1} - x_j}{x_{j+1} - x_{j-1}}\delta_{\{x_{j-1}\}} + \frac{x_j - x_{j-1}}{x_{j+1} - x_{j-1}}\delta_{\{x_{j+1}\}}\right).
$$

1. Let $I_j = (x_{j-1}, x_{j+1})$. We first prove that

$$u^\beta(t,x) \;=\; u^{\beta^*}(t,x), \quad (t,x) \in \big([0,\infty] \times \bar{\mathcal{I}} \setminus I_j\big) \cup \big([0,t_j] \times I_j\big). \tag{6.8}$$

By a direct calculation, we see that $U^{\beta^*}(x) = U^\beta(x)$ for $x \notin I_j$, and U^{β^*} is affine and strictly smaller than U^β on I_j. Consider first $x \notin I_j$. Recall (5.7) with the optimal stopping time τ^t being the minimum of t and the first entry to $\mathcal{R}^\beta$ for the diffusion X started in (t,x) and running backward in time. However since $\max\{t_{j-1}, t_{j+1}\} \le t_j$ it follows that $Y_{\tau^t} \ne x_j$ on $\tau^t < t$. In consequence, we can rewrite (5.11) as

$$u^\beta(t,x) \;=\; \sup_{\tau \in \mathcal{T}(\mathbf{x})} J_{t,x}^\beta(\tau) \;=\; \sup_{\tau \in \mathcal{T}(\mathbf{x}^{(-j)})} J_{t,x}^\beta(\tau) \;=\; \sup_{\tau \in \mathcal{T}(\mathbf{x}^{(-j)})} J_{t,x}^{\beta^*}(\tau) \;=\; u^{\beta^*}(t,x) \text{ for } t \ge 0, \ x \notin I_j.$$

An analogous argument shows $u^\beta(t,x) = u^{\beta^*}(t,x)$ for $x \in I_j \setminus \{x_j\}$ and $t \le t_j$ and for $x = x_j$ and $t < t_j$. By continuity of u^β we also have $u^\beta(x_j, t_j) = u^{\beta^*}(x_j, t_j)$.

2. We now prove that $u^\beta = v^{\xi^\beta}$ holds for all (t,x).

2.1. From the fact that $u^\beta(t,x) = u^{\beta^*}(t,x)$, for $x \notin I_j$, together with $\beta^*(I_j) = 0$, it follows that $\mathcal{R}^\beta = \mathcal{R}^{\beta^*} \cup \big([t_j, \infty) \times \{x_j\}\big)$. Consequently, for all $t \le t_j$ and all $s \ge 0$,

$$X_{t \wedge \sigma_{\mathcal{R}^{\beta^*}}} = X_{t \wedge \sigma_{\mathcal{R}^\beta}} \quad \text{and} \quad X_{s \wedge \sigma_{\mathcal{R}^{\beta^*}}} \mathbf{1}_{\bar{\mathcal{I}} \setminus I_j}(X_{s \wedge \sigma_{\mathcal{R}^{\beta^*}}}) = X_{s \wedge \sigma_{\mathcal{R}^\beta}} \mathbf{1}_{\bar{\mathcal{I}} \setminus I_j}(X_{s \wedge \sigma_{\mathcal{R}^\beta}}), \quad \text{a.s.}$$

It follows from the induction hypothesis that $u^\beta = v^{\xi^\beta}$ holds for all $x \in \mathbb{R}$, $t \le t_j$, and for all $x \notin I_j$.

2.2. It remains to consider $x \in (x_{j-1}, x_{j+1})$ and $t > t_j$. For $x \in (x_j, x_{j+1})$, we now know that $u^\beta = v^{\xi^\beta}$ holds at $t = t_j$, and $\mathcal{R}^\beta$ places no points in $[0, \infty) \times (x_j, x_{j+1})$. Then, it follows from Lemma 6.2 that $u^\beta = v^{\xi^\beta}$ on (x_j, x_{j+1}). The same argument applies for $x \in (x_{j-1}, x_j)$.

$\square$

6.3 The case of locally finitely supported measures

In this subsection, we consider the case of measures β which are α^ξ–finitely supported on any compact subset of $\mathcal{R}$, but could have an accumulation of atoms at $-\infty$ or ∞. We will establish Theorem 4.1 for such β by suitably approximating β with a sequence of measures with α^ξ–finite support. Recall that $\ell^\beta = \sup\{x : \alpha^\xi((-\infty, y]) = \beta((-\infty, y]) \ \forall y \le x\} = \sup\{x : U^{\alpha^\xi}(y) = U^\beta(y) \ \forall y \le x\}$, and similarly for r^β. The desired result has already been shown when $-\infty < \ell^\beta \le r^\beta < \infty$, see Proposition 6.3, so we consider the case where at least one of these is infinite. For simplicity, we suppose that both are infinite (and hence $\mathcal{I} = \mathbb{R}$), the case where only one is being similar. The approximation is depicted graphically in Figure 3.

For $N > 0$, we observe that we can define a new measure β^N, and constants $\ell^N < N, r^N > N$ such that $\beta^N([-N, N] \cap A) = \beta([-N, N]) \cap A$ for $A \in \mathcal{B}(\mathbb{R})$, $\beta^N([\ell^N, r^N]^\complement \cap A) = \alpha^\xi([\ell^N, r^N]^\complement \cap A)$, and $\beta^N((\ell^N, -N) \cup (N, r^N)) = 0$. In particular, to construct such a measure, we can set $U^{\beta^N}(x) = U^\beta(x)$ for $x \in [-N, N]$, and extend linearly to the right of N, with gradient $(U^\beta)'_+(N)$ until the function meets U^{α^ξ}, at the point r^N, from which point on, we take $U^{\beta^N}(x) = U^{\alpha^\xi}(x)$; a similar construction follows from $-N$. The existence of the point r^N follows from the fact that $U^\beta(x) - U^{\alpha^\xi}(x) \to 0$ as $x \to \infty$, which in turn is a consequence of the convex ordering property. This construction guarantees

$$U^{\beta^N}(x) \ge U^\beta(x) \text{ for all } x \in \mathbb{R},$$
$$U^{\beta^N} \text{ converges uniformly to } U^\beta \text{ and}$$
$$U^{\beta^N}(x) = U^{\alpha^\xi}(x) \text{ for } x \notin (\ell^N, r^N).$$

In particular, β^N is a sequence of atomic measures with α^ξ–finite support. Hence, by Proposition 6.3, Theorem 4.1 holds for these measures. Moreover, we can prove the following:
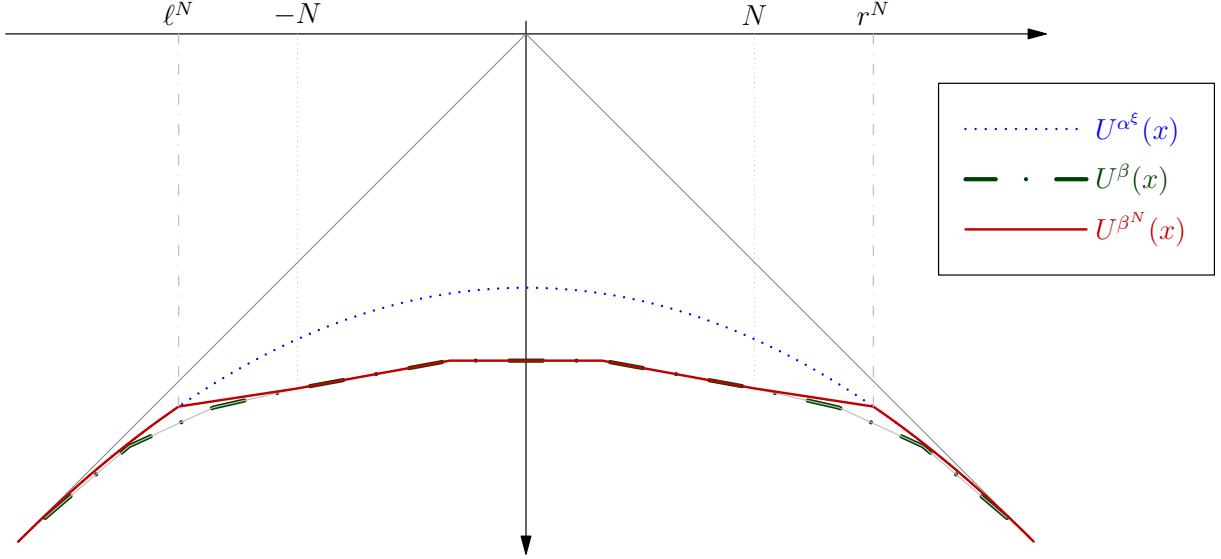
Figure 3: A graphical representation of the construction of the measure β^N in terms of the potential functions of the measures α^ξ and β.

Lemma 6.6. *Let $\sigma^\xi \in \mathcal{T}$ with corresponding time-space distribution ξ, and β a locally finitely supported measure such that $\alpha^\xi \preceq_{cx} \beta$. Let β^N be the sequence of measures constructed above. Then the sequence $\left(\mathcal{R}^{\beta^N} \cap ([0, \infty) \times [-N, N])\right)_{N \geq 1}$ is non-decreasing, and*

$$\mathcal{R}^\beta = \mathcal{R} := \overline{\bigcup_{N \geq 1} \left(\mathcal{R}^{\beta^N} \cap ([0, \infty) \times [-N, N])\right)}.$$

Proof. We proceed in two steps:

1. We first show that $\left(\mathcal{R}^{\beta^N} \cap ([0, \infty) \times [-N, N])\right)_{N \geq 1}$ is non-decreasing and $\mathcal{R}^\beta \supseteq \mathcal{R}$. Recall that $U^\beta(x) \leq U^{\beta^{N'}}(x) \leq U^{\beta^N}(x)$ for $N' \geq N$. Then, by definition of the optimal stopping problem, we see that $u^\beta(t, x) \leq u^{\beta^{N'}}(t, x) \leq u^{\beta^N}(t, x)$. However, we have $U^\beta(x) = U^{\beta^{N'}}(x) = U^{\beta^N}(x)$ for $x \in [-N, N]$ by construction, and so if it is optimal to stop for β^N, it is also optimal to stop for $\beta^{N'}$ and for β. It follows that, for $x \in [-N, N]$, $(t, x) \in \mathcal{R}^{\beta^N}$ implies $(t, x) \in \mathcal{R}^{\beta^N}$ and $(t, x) \in \mathcal{R}^\beta$. The desired monotonicity follows instantly and $\mathcal{R}^\beta \supseteq \mathcal{R}$ follows since $\mathcal{R}^\beta$ is closed.

2. It remains to show the reverse inclusion $\mathcal{R} \supseteq \mathcal{R}^\beta$.

First, observe that for the points where $\bar{t}_\mathcal{R}(x) = 0$ or $\bar{t}_\mathcal{R}(x) = \infty$ the inclusion holds. This is an immediate consequence of Corollary 5.6 together with the relation between the measures β and β^N.

The rest of the proof is devoted to showing that for a point x in the support of β with $0 < t' := \bar{t}_\mathcal{R}(x) < \infty$, we have $(t, x) \notin \mathcal{R}^\beta$ for all $t < t'$. We first carry our preparatory computations which follow two cases. Then we combine the two to give the final result.

2.1. Since Theorem 4.1 holds for β^N, we have $u^{\beta^N} = v^{\xi^{\beta^N}}$. It then follows from Remark 4.4 that $\mathbb{E}^\xi \left[L^x_{\sigma_{\beta^N}}\right] = \left(U^{\alpha^\xi} - U^{\beta^N}\right)(x) =: \varepsilon_0$ and $\varepsilon_0 > 0$ for our x. Denote $H_{-N_0, N_0} = \inf\{t \geq T_\xi : |X_t| \geq N_0\}$. Then, for sufficiently large N_0, we have $\mathbb{E}^\xi \left[L^x_{\sigma_{\mathcal{R}^{\beta^N}} \wedge H_{-N_0, N_0}}\right] > \varepsilon_0/2$ for all $N \geq N_0$. Note that $\lim_{N \to \infty} \mathcal{R}^{\beta^N} \cap ([0, \infty) \times [-N_0, N_0]) = \mathcal{R} \cap ([0, \infty) \times [-N_0, N_0])$. Letting $N \to \infty$, we conclude that

$$\mathbb{E}^\xi \left[L^x_{\sigma_\mathcal{R} \wedge H_{-N_0, N_0}}\right] \geq \varepsilon_0/2.$$

This means that, for all $t < t'$ with $t' - t$ sufficiently small there is a positive probability under $\mathbb{P}^\xi$ that the process reaches (t, x) before hitting $\mathcal{R}$ (and hence also $\mathcal{R}^{\beta^N}$) or exiting $[-N_0, N_0]$. In particular,

24

considering possible paths, we can reverse this: for any such $t < t'$, running backwards, there exists a positive probability that we will reach the support of ξ before hitting $\mathcal{R}$ or exiting a bounded interval. More specifically, writing $x_- = \sup\{y < x : (0, y) \in \mathcal{R}\}$, $x_+ = \inf\{y > x : (0, y) \in \mathcal{R}\}$, and $\varepsilon = t' - t$, for some ε sufficiently small at least one of the following two cases described below is true. We refer to Figure 4 for a graphical interpretation of the two cases, and a number of the important quantities described below.
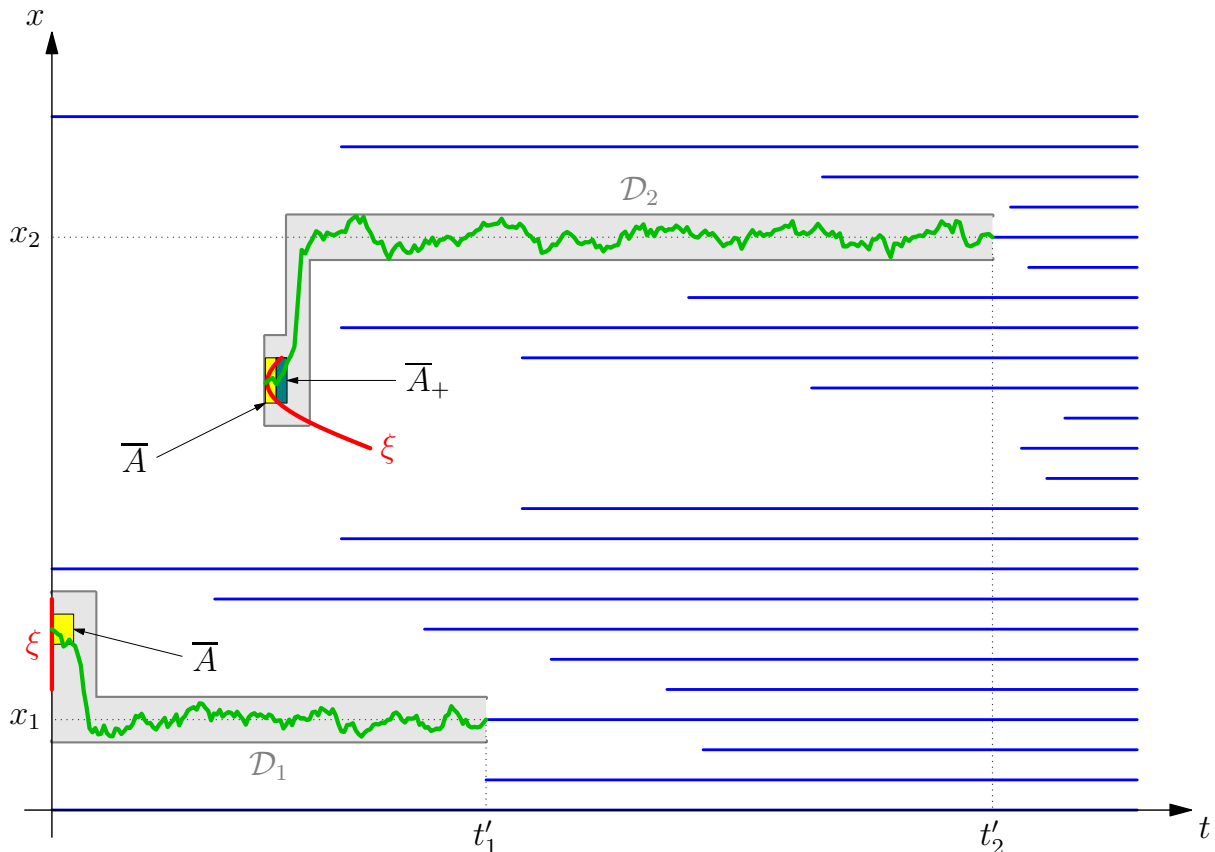

Figure 4: The possible cases considered in step 3.1. of the proof of Lemma 6.6. In the first case, shown in the bottom half of the diagram, paths starting at (t'_1, x_1) can only reach points in the support of ξ (denoted by the red line) which are at time 0. In this case, we are interested on the behaviour of the process on the set $\overline{A}$ shown, given that it does not leave the set $\mathcal{D}_1$. In the second case, the process starting at (t'_2, x_2) can reach points in the support of ξ which are not in the set $\{t = 0\}$. In this case, we are interested in the behaviour of the process on the sets $\overline{A}$ and $\overline{A}_+$ depicted, given that the process does not leave $\mathcal{D}_2$.

Case 1 The only points of the support of ξ which can be reached from (t', x) without exiting $\mathcal{R}$ are in $\{0\} \times (x_-, x_+)$. Let $A \subseteq (x_-, x_+)$ be a closed and bounded interval such that $\xi(\{0\} \times A) > 0$. Observe that the measures β^N are α^ξ-finitely supported, and hence $\mathcal{R}^{\beta^N} \cap (\mathbb{R}_+ \times ([x - \varepsilon, x + \varepsilon] \setminus \{x\})) = \emptyset$ for some $\varepsilon > 0$, and all N. Moreover, we may assume that ε is also sufficiently small that $[0, 2\varepsilon] \times [\inf A \wedge x - \varepsilon, \sup A \vee x + \varepsilon] \cap \mathcal{R} = \emptyset$.

For such an ε, write

$$\mathcal{D} := ([0, 2\varepsilon] \times [\inf A \wedge x - \varepsilon, \sup A \vee x + \varepsilon] \cup [0, t') \times [x - \varepsilon, x + \varepsilon])$$

and note that $\mathcal{R} \cap \mathcal{D} = \emptyset$.

Our aim is now to use the expression of $\mathcal{L}v^\xi$ in Lemma 5.1, to show that $V^{t'}$ is a strict supermartingale on $\overline{A} := [0, \varepsilon] \times A$. Recall that $t = t' - \varepsilon$ and define

$$\tau_N = \inf\{s > 0 : (t' - s, Y_s) \in \mathcal{R}^{\beta^N}\} \wedge t, \qquad \tau = \inf\{s > 0 : (t' - s, Y_s) \in \mathcal{R}\} \wedge t$$

$$\tau_N^\varepsilon = \inf\{s > 0 : (t' - s, Y_s) \in \mathcal{R}^{\beta^N}\} \wedge t', \qquad \tau^\varepsilon = \inf\{s > 0 : (t' - s, Y_s) \in \mathcal{R}\} \wedge t'$$

and

$$\tau^\mathcal{D} = \inf\{s > 0 : (t' - s, Y_s) \notin \mathcal{D}\} \wedge t'.$$

Recall the family of supermartingales V^t defined in (5.5). We want to show that $\mathbb{E}^x\left[V_{\tau_N}^{t'} - V_{\tau_N^\varepsilon}^{t'}\right] \geq \delta > 0$ for some constant δ which is independent of N. Since $\tau^\mathcal{D} \wedge t \leq \tau_N \leq \tau_N^\varepsilon$ for all N, the event $\{\tau^\mathcal{D} > t\}$ is $\mathcal{F}_{\tau_N}$-measurable. Hence it is sufficient to show that $\mathbb{E}^x\left[\left(V_{\tau_N}^{t'} - V_{\tau_N^\varepsilon}^{t'}\right)\mathbf{1}_{\{\tau^\mathcal{D} > t\}}\right] \geq \delta$. Using the supermartingale property of $V^{t'}$, we can further reduce this to showing that

$$\mathbb{E}^x\left[\left(V_{\tau_N}^{t'} - V_{\tau_N^\varepsilon \wedge \tau^\mathcal{D}}^{t'}\right)\mathbf{1}_{\{\tau^\mathcal{D} > t\}}\right] \geq \delta.$$

Note that on $\{\tau_D > t\}$ we have $\tau_N = t$ and $\tau_N^\epsilon \geq \tau_D$. We now write $q(t' - s, y)$ for the space-time density of the process $(t' - s, x + Y_s)$ killed when it leaves $\mathcal{D}$, i.e.

$$\mathbb{E}^x\left[f(Y_s); s < \tau_D\right] = \int q(t' - s, y) f(y)\, dy$$

for smooth functions f. Then from the form of $\mathcal{D}$, we know that q is bounded away from zero on $\overline{A}$, and applying Lemma 5.1 we have

$$\mathbb{E}^x\left[\left(V_{\tau_N}^{t'} - V_{\tau_N^\varepsilon \wedge \tau^\mathcal{D}}^{t'}\right)\mathbf{1}_{\{\tau^\mathcal{D} > t\}}\right] \geq -\int_{(t'-s, y) \in \overline{A}} q(t' - s, y) \mathcal{L}v^\xi(t' - s, dy)\, ds$$

$$\geq \int_{(t'-s, y) \in \overline{A}} \eta(y)^2 q(t' - s, y) \xi(0, dy)\, ds,$$

by the assumption on the support of ξ under consideration. By the assumption on ξ, and the fact that q is bounded below on $\overline{A}$, this final term is strictly positive, and independent of N, so:

$$\mathbb{E}^x\left[V_{\tau_N}^{t'} - V_{\tau_N^\varepsilon}^{t'}\right] \geq \delta \tag{6.9}$$

for some $\delta > 0$ independent of N.

Case 2   There exists a bounded rectangle $\overline{A} \subset (0, t') \times (x_-, x_+)$ such that $\xi(\overline{A}) > 0$, all points of $\overline{A}$ can be reached from (t', x) via a continuous path which does not enter $\mathcal{R}$, and the process spends a strictly positive time in $\overline{A}$. More specifically, for all sufficiently small $\varepsilon > 0$, we can choose a_ℓ, a_r, s_A such that $\overline{A} = [s_A, s_A + \varepsilon/2] \times [a_\ell, a_r]$, $\xi(\overline{A}) > 0$, $s_A + 3\varepsilon < t'$ and the set

$$\mathcal{D} := ([s_A, s_A + \varepsilon] \times [a_\ell - \varepsilon, a_r + \varepsilon]) \cup ([s_A + \varepsilon, s_A + 2\varepsilon] \times [a_\ell \wedge x - \varepsilon, a_r \vee x + \varepsilon])$$
$$\cup ([s_A + 2\varepsilon, t'] \times [x - \varepsilon, x + \varepsilon])$$

satisfies $\mathcal{D} \cap \mathcal{R} = \emptyset$. Further, recalling the definitions of $\tau^\mathcal{D}$ and τ_N above, we have $\tau^\mathcal{D} \leq \tau_N$ $\mathbb{P}^x$-a.s.. In a similar manner to above, we now write $\tilde{q}(t' - s, y)$ for the space-time density of the process $(t' - s, x + Y_s)$ killed when it leaves $\mathcal{D}$, and observe that $\tilde{q}$ is bounded away from zero on the set

$\overline{A}_+ := [s_A + \varepsilon/2, s_A + \varepsilon] \times [a_\ell, a_r]$. It follows from Lemmas 5.1 and 5.2 that:

$$\mathbb{E}^x \left[\int_0^{\tau_N} \left(\mathcal{L}v^\xi(t - s, Y_s) - \mathcal{L}v^\xi(t' - s, Y_s) \right) ds \right]$$

$$\geq \mathbb{E}^x \left[\int_0^{\tau^D} \left(\mathcal{L}v^\xi(t - s, Y_s) - \mathcal{L}v^\xi(t' - s, Y_s) \right) ds \right]$$

$$\geq \int_{(t'-s,y)\in\mathcal{D}} \tilde{q}(t' - s, y) \left(\mathcal{L}v^\xi(t - s, y) - \mathcal{L}v^\xi(t' - s, y) \right) ds \, dy$$

$$\geq \int_{(t'-s,y)\in\overline{A}_+} \eta(y)^2 \tilde{q}(t' - s, y) \xi([s_A, s_A + \varepsilon/2), dy) ds$$

where in the last line we applied Lemma 5.1 and the fact that for $(t' - s, y) \in \overline{A}_+$

$$\left(\mathcal{L}v^\xi(t - s, y) - \mathcal{L}v^\xi(t' - s, y) \right) dy = \eta(y)^2 \xi([t - s, t' - s), dy) \geq \eta(y)^2 \xi([s_A, s_A + \varepsilon/2), dy).$$

It follows that we can choose $\delta > 0$ independent of N such that

$$\mathbb{E}^x \left[\int_0^{\tau_N} \left(\mathcal{L}v^\xi(t - s, Y_s) - \mathcal{L}v^\xi(t' - s, Y_s) \right) ds \right] \geq \delta,$$

which, by an application of Itô's formula, implies that

$$\mathbb{E}^x \left[V_{\tau_N}^t - V_{\tau_N}^{t'} \right] \geq v^\xi(t, x) - v^\xi(t', x) + \delta. \tag{6.10}$$

Observe finally that, in view of the supermartingale properties of Lemma 5.2, we can combine (6.9) and (6.10) to get:

$$\mathbb{E}^x \left[V_{\tau_N}^t - V_{\tau_N}^{t'} \right] + \mathbb{E}^x \left[V_{\tau_N}^{t'} - V_{\tau_N^\varepsilon}^{t'} \right] \geq v^\xi(t, x) - v^\xi(t', x) + \delta \tag{6.11}$$

for some $\delta > 0$ independent of N, and for *any* ξ satisfying the conditions of the lemma.

2.2. We are now ready to exploit the above to establish that $(t, x) \notin \mathcal{R}^\beta$ for $t < t'$. Take the values of t, ε, δ determined above, and consider the following calculation:

$$\begin{aligned}
u^{\beta^N}(t, x) - v^\xi(t, x) &\geq \mathbb{E}^x \left[V_{\tau_N}^t + w^{\beta^N}(Y_{\tau_N}) \mathbf{1}_{\{\tau_N < t\}} \right] - v^\xi(t, x) \\
&\geq \mathbb{E}^x \left[V_{\tau_N}^t - V_{\tau_N}^{t'} \right] + \mathbb{E}^x \left[V_{\tau_N}^{t'} - V_{\tau_N^\varepsilon}^{t'} \right] \\
&\quad + \mathbb{E}^x \left[w^{\beta^N}(Y_{\tau_N}) \mathbf{1}_{\{\tau_N < t\}} - w^{\beta^N}(Y_{\tau_N^\varepsilon}) \mathbf{1}_{\{\tau_N^\varepsilon < t'\}} \right] \\
&\quad + \mathbb{E}^x \left[V_{\tau_N^\varepsilon}^{t'} + w^{\beta^N}(Y_{\tau_N^\varepsilon}) \mathbf{1}_{\{\tau_N^\varepsilon < t'\}} \right] - v^\xi(t, x) \\
&\geq \left(v^\xi(t, x) - v^\xi(t', x) \right) + \delta + u^{\beta^N}(t', x) - v^\xi(t, x).
\end{aligned}$$

Here we use (6.11) for the first two terms in the second inequality; the third term in the second inequality is at least 0 using the fact that $\tau_N < t$ implies that $\tau_N^\varepsilon = \tau_N < t$, and $w^{\beta_N}(\cdot) \leq 0$. It then follows, since v^ξ is non-increasing in t, that

$$u^{\beta^N}(t, x) - v^\xi(t, x) \geq u^{\beta^N}(t', x) - v^\xi(t', x) + \delta \geq w^{\beta^N}(x) + \delta \geq w^\beta(x) + \delta.$$

We now use the fact that $\delta > 0$ independently of N, and $u^{\beta^N}(t, x) \to u^\beta(t, x)$ as $N \to \infty$ to deduce that $u^\beta(t, x) - v^\xi(t, x) > w^\beta(x)$. In particular, it is not optimal to stop immediately for the u^β optimal stopping problem at (t, x) with $t < t'$, whenever $0 < \bar{t}_\mathcal{R}(x) < \infty$.

$\square$

Proposition 6.7. *Let $\sigma^\xi \in \mathcal{T}$ with corresponding time-space distribution ξ, and β a locally finitely supported measure such that $\alpha^\xi \preceq_{\mathrm{cx}} \beta$. Then $u^\beta = v^{\xi^\beta}$ and Theorem 4.1 holds for β.*

27

Proof. It follows from Lemma 6.6 that σ^{β^N} decreases to σ^β, and $X_{\sigma^{\beta^N}}$ converges to X_{σ^β} in probability, and therefore $X_{\sigma^\beta} \sim \beta$. Finally, if we write $H_{\pm N} = \inf\{t \geq T_\xi : |X_t| = N\}$, we also have

$$
\begin{aligned}
\mathbb{E}^\xi\left[L^x_{t \wedge \sigma^\beta}\right] &= \lim_{N \to \infty} \mathbb{E}^\xi\left[L^x_{t \wedge \sigma^\beta \wedge H_{\pm N}}\right] \\
&= \lim_{N \to \infty} \mathbb{E}^\xi\left[L^x_{t \wedge \sigma^{\beta^N} \wedge H_{\pm N}}\right] \\
&= \lim_{N \to \infty} \left[v^\xi(t, x) - u^{\beta^N}(t, x)\right] \\
&= v^\xi(t, x) - u^\beta(t, x),
\end{aligned}
$$

where we used (4.6) and monotone convergence. It follows from Remark 4.3 that $v^{\xi^\beta} = u^\beta$.

Since $X_{\sigma^\beta} \sim \beta$ and $X_0 \sim \alpha$, and $v^\xi(t, x) - u^\beta(t, x) \to -w^\beta(x)$ as $t \to \infty$, by monotone convergence, we have $\mathbb{E}^\xi\left[L^x_{t \wedge \sigma^\beta}\right] = -w^\beta(x)$, and hence by (Elworthy et al., 1999, Corollary 3.4), σ^β is a UI stopping time. Finally, we deduce that $\mathcal{R}^\beta$ is ξ-regular by observing from (4.4) and taking limits in the equation above that $(t, x) \in \mathcal{R}^\beta$ if and only if $\mathbb{E}^\xi\left[L^x_{t \wedge \sigma^\beta}\right] = w^\beta(x) = \mathbb{E}^\xi\left[L^x_{\sigma^\beta}\right]$. From Remark 2.3, it follows that $\mathcal{R}^\beta$ is ξ-regular. $\qquad \square$

7 The general case

In this section, we complete the proof of Theorem 4.1. We fix $\sigma^\xi \in \mathcal{T}$ with its corresponding time-space distribution ξ, and let β be an arbitrary integrable measure such that $\beta \succeq_{\mathrm{cx}} \alpha^\xi$. We start by approximating β with a sequence of locally finitely supported measures. Let

$$
I^k_m := [k2^{-m}, (k+1)2^{-m}] \cap \bar{\mathcal{I}}, \quad \text{and} \quad t^k_m := \min_{x \in I^k_m} \bar{t}^\beta(x) = \bar{t}^\beta(x^k_m) \text{ with } (t^k_m, x^k_m) \in \mathcal{R}^\beta, \; x^k_m \in I^k_m \tag{7.1}
$$

We set $t^k_m = \infty$ when there are no points of $\mathcal{R}^\beta$ in $[0, \infty) \times I^k_m$, see Corollary 5.6 for a characterisation. The existence of a minimizer x^k_m follows from the lower semicontinuity of the barrier function $\bar{t}^\beta$ which, in turn, is implied by the closedness property of the barrier $\mathcal{R}^\beta$. If there exists more than one minimiser, we choose the smallest: $x^k_m = \min\{x \in I^k_m : \bar{t}^\beta(x) = t^k_m\}$, so that if $(t, x) = (t^k_m, x^k_m)$, then $(t, x) = (t^{k'}_{m+1}, x^{k'}_{m+1})$ for some k'. Note that $0 \leq x^{k+1}_m - x^k_m \leq 2^{-m+1}$.

We now determine a sequence of approximating measures defined as follows: the measure β^m is defined through its potential function, $U^{\beta^m}(x)$, and we set $U^{\beta^m}(x)$ to be the smallest concave function such that $U^{\beta^m}(x^k_m) = U^\beta(x^k_m)$ for all k. In particular, we deduce that $U^{\beta^m}(x) \leq U^{\beta^{m+1}}(x) \leq U^\beta(x)$; moreover, β^m has the same mean as β, $\beta^m \succeq_{\mathrm{cx}} \beta^{m+1} \succeq_{\mathrm{cx}} \beta$ and $U^{\beta^m}(x) - U^\beta(x) \to 0$ as $x \to \partial\mathcal{I}$ for each m. This approximation is depicted in Figure 5.

Each β^m is locally finitely supported, and so we can apply Proposition 6.7 to each β^m. Write $\mathcal{R}^m := \mathcal{R}^{\beta^m}$ for the corresponding barrier. A typical sequence of barriers are depicted in Figure 6. Since the potentials of the measures are increasing, we have $u^{\beta^m}(t, x) \leq u^{\beta^{m+1}}(t, x)$; in addition, the function $U^{\beta^m}(x)$ is piecewise linear, and so $(t, x) \in \mathcal{R}^m$ implies $x = x^k_m = x^{k'}_{m+1}$, for some k, k', and $U^{\beta^m}(x) = U^{\beta^{m+1}}(x) = U^\beta(x)$. In consequence, for such an x we have

$$
v^\xi(t, x) + w^\beta(x) = v^\xi(t, x) + w^{\beta^{m+1}}(x) = v^\xi(t, x) + w^{\beta^m}(x) \leq u^{\beta^m}(t, x) \leq u^{\beta^{m+1}}(t, x) \leq u^\beta(t, x) \tag{7.2}
$$

and it follows from the optimal stopping formulation that $\bar{t}_{\mathcal{R}^m}(x) \leq \bar{t}_{\mathcal{R}^{m+1}}(x) \leq \bar{t}^\beta(x)$ — i.e. new spikes may appear, but existing spikes get smaller. Taking a sequence k_m such that $x = x^{k_m}_m$ for all $m \geq m_0$, for some m_0, we see that $\bar{t}_{\mathcal{R}^m}(x)$ increases to a limit. We now establish that this limit is equal to $\bar{t}^\beta(x)$.
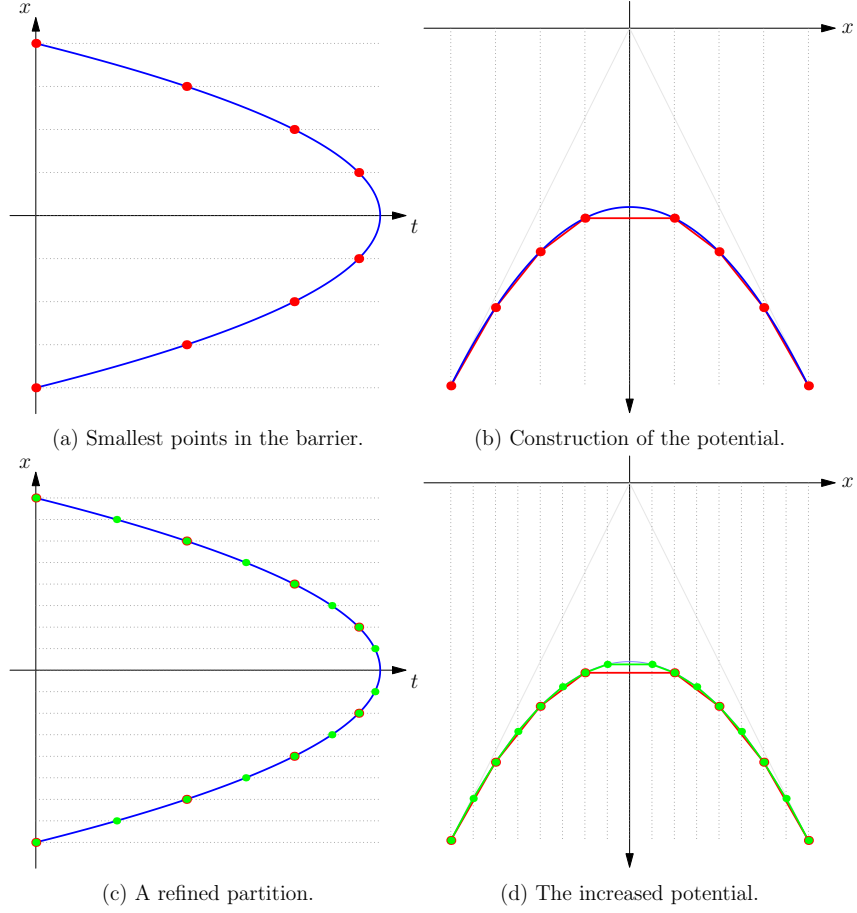
(a) Smallest points in the barrier.

(b) Construction of the potential.

(c) A refined partition.

(d) The increased potential.

Figure 5: The approximation sequence of a general measure β. In (a), the red points denote the smallest point in the barrier for the given subdivisions (marked in gray). In (b), the original potential (in blue) is interpolated at the corresponding x-values, to produce a smaller potential corresponding to a measure β^m. In (c), a finer set of intervals are used to produce additional approximating points. Note that the previous (red) points are all in the new set of approximating points. In (d), these points are used to produce the potential of a new measure β^{m+1}.

Lemma 7.1. *Let*

$$\mathcal{R} := \overline{\bigcap_{m \geq 0} \bigcup_{k \geq m} \mathcal{R}^k}. \tag{7.3}$$

Then $\mathcal{R} = \mathcal{R}^\beta$ and for any x of the form $x = x_m^{k_m}$, for some sequence of indices (k_m), $\bar{t}_{\mathcal{R}^m}(x) \nearrow \bar{t}^\beta(x)$.

Proof. We first show $\mathcal{R} \subseteq \mathcal{R}^\beta$. Let $(t, x) \in \bigcap_{m \geq 0} \bigcup_{k \geq m} \mathcal{R}^k$. Then, for all $m \geq 1$, there is $k_m \geq m$ such that $(t, x) \in \mathcal{R}^{k_m}$, i.e. $(u^{\beta^{k_m}} - v^\xi)(t, x) = w^{\beta^{k_m}}(x) = w^\beta(x)$. However $u^{\beta^{k_m}}(t, x) \to u^\beta(t, x)$ as $m \to \infty$, and so $(u^\beta - v^\xi)(t, x) = w^\beta(x)$, proving that $(t, x) \in \mathcal{R}^\beta$. This shows that $\bigcap_{m \geq 0} \bigcup_{k \geq m} \mathcal{R}^k \subset \mathcal{R}^\beta$, and therefore $\mathcal{R} \subset \mathcal{R}^\beta$ by the closeness of $\mathcal{R}^\beta$.

We now show the reverse inclusion, $\mathcal{R}^\beta \subseteq \mathcal{R}$. For $(t, x) \in \mathcal{R}^\beta$, and $\varepsilon > 0$, choose m_0 so that $2^{-m_0} < \varepsilon$. Then there exists x' such that $|x - x'| < \varepsilon$ and $(t', x') \in \mathcal{R}^{m_0}$ for some t' and $\bar{t}^\beta(x') \leq \bar{t}^\beta(x) \leq t$ by our choice of points x_m^k. Further, as argued above, $\bar{t}_{\mathcal{R}^{m_0}}(x') \leq \bar{t}_{\mathcal{R}^m}(x') \leq \bar{t}^\beta(x') \leq t$ so that $(t, x') \in \mathcal{R}^m$ for all $m \geq m_0$. It follows that $(t, x) \in \mathcal{R}$.

The above shows $\mathcal{R} = \mathcal{R}^\beta$, or equivalently $\bar{t}_{\mathcal{R}} = \bar{t}^\beta$. As observed above, for $x = x_m^{k_m}$, we have $\bar{t}_{\mathcal{R}^m}(x)$ is an increasing sequence in m and hence converges to some limit which we denote $t(x)$. By the barrier property of each $\mathcal{R}_m$ and the definition of $\mathcal{R}$ we see that $(t(x), x) \in \mathcal{R}$. It follows that $\bar{t}_{\mathcal{R}}(x) \leq t(x) \leq \bar{t}^\beta(x)$
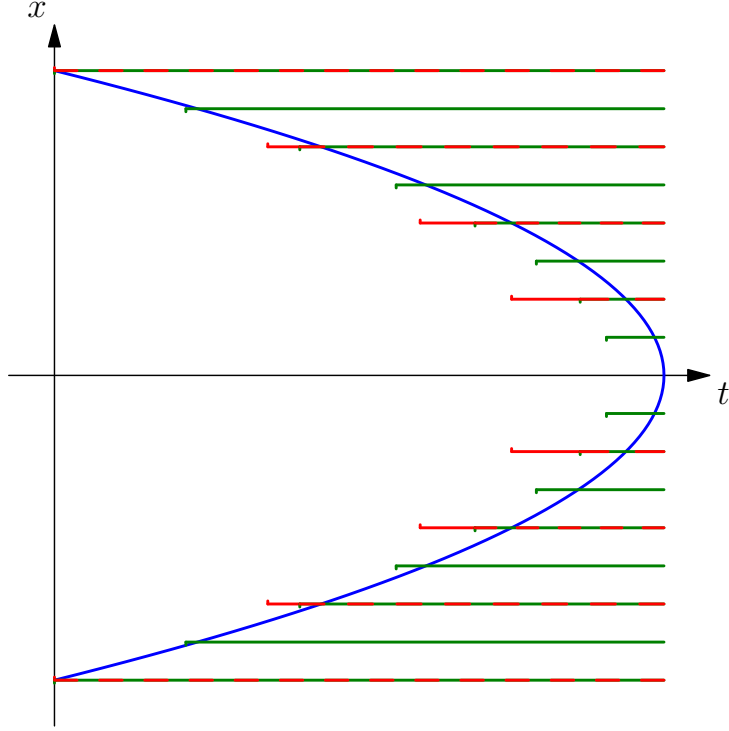
29

Figure 6: The sequence of barriers constructed by the approximation sequence. The red barrier corresponds to β^m, and the green barrier to β^{m+1}. Where the barriers have common atoms, the green barrier is to the right of the left barrier, however new 'spikes' appear for the green barrier. The blue line denotes the barrier $\mathcal{R}^\beta$.

and hence all three are equal. □

Proposition 7.2. *Consider the approximation sequence above and define $\sigma^m = \sigma_{\mathcal{R}^m} \wedge \sigma_{\mathcal{R}^\beta}$. Then:*

(i) *the process $(X_{t \wedge \sigma^m})_{t \geq T_\xi}$ is uniformly integrable under $\mathbb{P}^\xi$;*

(ii) *$\sigma^m \to \sigma_{\mathcal{R}^\beta}$;*

(iii) *$\mathbb{E}^\xi \left[L^x_{t \wedge \sigma_{\mathcal{R}^\beta}} \right] \leq v^\xi(t, x) - u^\beta(t, x)$.*

Proof. (i) By definition $\sigma^m \leq \sigma_{\mathcal{R}^m}$ and, from Proposition 6.7, the same process stopped at $\sigma_{\mathcal{R}^m}$ is uniformly integrable, which implies the result.

(ii) Suppose that σ^m does not converge a.s. to $\sigma_{\mathcal{R}^\beta}$. Take ω such that, possibly passing to a subsequence, we have that $\sigma^m(\omega) \to t_\infty$ for some $t_\infty < \sigma_{\mathcal{R}^\beta}(\omega)$. Then necessarily $(t_m, y_m) := (\sigma^m(\omega), X_{\sigma^m(\omega)}) \in \mathcal{R}^m$ for m large enough. This gives

$$v^\xi(t_m, y_m) + w^\beta(y_m) = v^\xi(t_m, y_m) + w^{\beta^m}(y_m) = u^{\beta^m}(t_m, y_m).$$

We take limits on both sides. The left-hand side converges to $v^\xi(t_\infty, y_\infty) + w^\beta(y_\infty)$ by continuity of w^β and joint continuity of v^ξ, see (5.2). For the right-hand side we use 1-Lipschitz continuity of each $u^{\beta^m}(t, \cdot)$ and 1/2-Hölder continuity in t as given in Lemma 5.5. This shows, with $X_{t_\infty} =: y_\infty$, that

$$|u^{\beta^m}(t_\infty, y_\infty) - u^{\beta^m}(t_m, y_m)| \leq C(1 + |y_\infty|)\sqrt{|t_\infty - t_m|} + |y_\infty - y_m|$$

for a constant C independent of m, and hence

$$\lim_{m \to \infty} u^{\beta^m}(t_m, y_m) = \lim_{m \to \infty} u^{\beta^m}(t_\infty, y_\infty) = u^\beta(t_\infty, y_\infty),$$

which then shows that $(t_\infty, y_\infty) \in \mathcal{R}^\beta$ and hence $\sigma_{\mathcal{R}^\beta}(\omega) \leq t_\infty$ which gives the desired contradiction.

(iii) Using the above, together with Proposition 6.7 and Remark 4.3, we deduce that

$$\mathbb{E}^\xi \left[L^x_{t \wedge \sigma_{\mathcal{R}^\beta}} \right] = \lim_{m \to \infty} \mathbb{E}^\xi \left[L^x_{t \wedge \sigma^m} \right] \leq \lim_{m \to \infty} \mathbb{E}^\xi \left[L^x_{t \wedge \sigma_{\mathcal{R}^m}} \right] = \lim_{m \to \infty} \left[v^\xi(t,x) - u^{\beta^m}(t,x) \right] = v^\xi(t,x) - u^\beta(t,x).$$

$\square$

Lemma 7.3. *We have $v^{\xi^\beta} = u^\beta$ and σ^{ξ^β} is a UI stopping time embedding β.*

Proof. It suffices to show the first equality as the rest follows from Lemma 5.5 (iii) and Lemma 4.2. Given (iii) of Proposition 7.2 and Remark 4.3, it remains only to show that $\mathbb{E}^\xi \left[L^x_{t \wedge \sigma_{\mathcal{R}^\beta}} \right] \geq v^\xi(t,x) - u^\beta(t,x)$. We consider the alternative approximating sequence: $\tilde{\mathcal{R}}^m := \mathcal{R}^m \cap \mathcal{R}^\beta$. Recall from above that if $(t,x) \in \mathcal{R}^m$ then $\bar{t}_{\mathcal{R}_m}(x) \leq \bar{t}_{\mathcal{R}_{m+1}}(x) \leq \bar{t}^\beta(x)$ from which it follows that $\tilde{\mathcal{R}}^m$ is an increasing sequence of barriers. Moreover, from the definition of the points x_m^k, we have $\sigma_{\tilde{\mathcal{R}}^m} \searrow \sigma_{\mathcal{R}^\beta}$, since when we hit $\mathcal{R}^\beta$, we are guaranteed to hit $\tilde{\mathcal{R}}^m$ as soon as we have travelled at least 2^{-m+1} in both directions. However $\sigma_{\tilde{\mathcal{R}}^m} \geq \sigma_{\mathcal{R}^m}$, and therefore:

$$\mathbb{E}^\xi \left[L^x_{t \wedge \sigma_{\mathcal{R}^m}} \right] \leq \mathbb{E}^\xi \left[L^x_{t \wedge \sigma_{\tilde{\mathcal{R}}^m}} \right] \to \mathbb{E}^\xi \left[L^x_{t \wedge \sigma_{\mathcal{R}^\beta}} \right].$$

But also $\mathbb{E}^\xi \left[L^x_{t \wedge \sigma_{\mathcal{R}^m}} \right] = v^\xi(t,x) - u^{\beta^m}(t,x) \to v^\xi(t,x) - u^\beta(t,x)$ and the result follows. $\square$

We note that ξ^β-regularity of $\mathcal{R}^\beta$ now follows from Remark 4.4. The proof of Theorem 4.1 is complete.

A   Proofs of the optimality results

We prove here the results announced in Section 3.2. We start with establishing the required pathwise inequality.

Proof of Lemma 3.3. We proceed in three steps.

1. We first observe that $\varphi_k \geq \varphi_{k+1}$ for all $k = 1, \ldots, n$, and $\varphi_n = \varphi_{n+1}$ on $\mathcal{R}^n$. Indeed, notice that $\varphi_k(t,x) = \mathbb{E}^{t,x}[f(\zeta^k)]$, where ζ^k is the first time we enter $\mathcal{R}^n$, having previously entered the barriers $\mathcal{R}^{n-1}, \mathcal{R}^{n-2}, \ldots, \mathcal{R}^k$ in sequence. Then $\zeta^k \geq \zeta^{k+1}$, $\mathbb{P}^{t,x}$-a.s. implying that $\varphi_k \geq \varphi_{k+1}$ by the non-decrease of f.

2. We next compute that:

$$(h_k - h_{k-1})(t,x) - \lambda_{k-1}(x) \quad = \quad \int_t^{\bar{t}^{k-1}(x)} (\varphi_{k-1} - \varphi_k)(s,x)ds. \tag{A.1}$$

Then, $h_k - h_{k-1} - \lambda_{k-1} \geq 0$ for $t \leq \bar{t}^{k-1}(x)$, by Step 1. Next, notice that $t \geq \bar{t}^{k-1}(x)$ if and only if $(s,x) \in \mathcal{R}^{k-1}$ for all $s \in [\bar{t}^{k-1}(x), t]$, and that in this case $\sigma_{\mathcal{R}^{k-1}} = s$, $\mathbb{P}^{s,x}$-a.s., implying that $\varphi_{k-1}(s,x) = \varphi_k(s,x)$. Hence:

$$h_k \geq h_{k-1} + \lambda_{k-1} \quad k = 2, \ldots, n, \text{ with equality on } \mathcal{R}^{k-1}. \tag{A.2}$$

3. By the previous steps, we have:

$$\sum_{i=1}^{n} \lambda_i(x_i) + \sum_{i=1}^{n} \big[h_i(s_i, x_i) - h_i(s_{i-1}, x_{i-1}) \big] + h_1(s_0, x_0)$$

$$= \sum_{i=1}^{n} \lambda_i(x_i) + \sum_{i=1}^{n-1} \big[h_i(s_i, x_i) - h_{i+1}(s_i, x_i) \big] + h_n(s_n, x_n)$$

$$\leq \lambda_n(x_n) + h_n(s_n, x_n), \quad \text{with } \text{``=''} \text{ if } (s_i, x_i) \in \mathcal{R}^i, \ i = 1, \ldots, n-1,$$

$$= \int_0^{s_n} f(t)dt - \psi(x_n)f(0) - \int_{s_n}^{\bar{t}^n(x_n)} (\varphi_n - \varphi_{n+1})(t, x_n)dt$$

$$\leq \int_0^{s_n} f(t)dt - \psi(x_n)f(0), \quad \text{with } \text{``=''} \text{ if } (s_n, x_n) \in \mathcal{R}^n,$$

where we used (A.1) and $\phi_{n+1}(x) = f(0)\int_0^x \eta(y)^{-2}\,dy$. $\qquad\square$

To be able to take expectations in the pathwise inequality when applied to the stopped diffusion, we need to establish suitable (sub)martingale properties. These are isolated in the following lemma.

Lemma A.1. *Let f be bounded non-negative and non-decreasing, and assume*

$$\int_0^{\cdot} \phi_k(X_s)dX_s \quad \text{is a } \mathbb{P}^{\mu_0}-\text{martingale for all} \quad k = 1, \ldots, n+1. \tag{A.3}$$

Then, for all $k = 1, \ldots, n$, the process $\{h_k(t, X_t) - h_k(0, X_0), t \geq 0\}$ is a $\mathbb{P}^{\mu_0}$-submartingale, and a $\mathbb{P}^{\mu_0}$-martingale on $[\sigma_{k-1}, \sigma_k]$.

Proof. First, applying the Itô-Tanaka formula to the second term in the definition of h_k, we have

$$h_k(t, X_t) = h_k(0, X_0) + \int_0^t \varphi_k(u, X_t)du - 2\int_0^t \phi_k(X_u)dX_u - \int_0^t \varphi_k(0, X_u)du, \quad t \geq 0.$$

Since $0 \leq \varphi(u, x) \leq \|f\|_\infty < \infty$, (A.3) shows that $h_k(t, X_t) - h_k(0, X_0)$ differs from a martingale by a bounded random variable and in particular is integrable. We now proceed in two steps.

1. For $0 \leq s \leq t$, using the above decomposition and (A.3), we have

$$\mathbb{E}_s^{\mu_0} \big[h_k(t, X_t) \big] - h_k(0, X_0) = \int_0^t \mathbb{E}_s^{\mu_0} \big[\varphi_k(u, X_t) \big] du - \int_0^t \mathbb{E}_s^{\mu_0} \big[\varphi_k(0, X_u) \big] du - 2\int_0^s \phi_k(X_u)dX_u$$

where $\mathbb{E}_s^{\mu_0} := \mathbb{E}^{\mu_0}[.|\mathcal{F}_s]$. We shall prove in Step 2 below that

$$\mathbb{E}_s^{\mu_0} \big[\varphi_k(u, X_t) \big] \geq \mathbb{E}_s^{\mu_0} \big[\varphi_k\big(u - (t-s), X_s\big) \big] \quad \text{for} \quad u \in [t-s, t], \tag{A.4}$$

$$\mathbb{E}_s^{\mu_0} \big[\varphi_k(u, X_t) \big] \geq \mathbb{E}_s^{\mu_0} \big[\varphi_k(0, X_{t-u}) \big] \quad \text{for} \quad u \in [0, t-s], \tag{A.5}$$

and

$$\text{equality holds in (A.4)–(A.5) if} \quad \sigma_{k-1} \leq s \leq t \leq \sigma_k. \tag{A.6}$$

Then,

$$\mathbb{E}_s^{\mu_0} \big[h_k(t, X_t) \big] - h_k(0, X_0) \geq \int_0^{t-s} \mathbb{E}_s^{\mu_0} \big[\varphi_k(0, X_{t-u}) \big] du + \int_{t-s}^t \mathbb{E}_s^{\mu_0} \big[\varphi_k(u - (t-s), X_s) \big] du$$

$$- \int_0^t \mathbb{E}_s^{\mu_0} \big[\varphi_k(0, X_u) \big] du - 2\int_0^s \phi_k(X_u)dX_u$$

$$= \int_s^t \mathbb{E}_s^{\mu_0} \big[\varphi_k(0, X_u) \big] du + \int_0^s \mathbb{E}_s^{\mu_0} \big[\varphi_k(u, X_s) \big] du$$

$$- \int_0^t \mathbb{E}_s^{\mu_0} \big[\varphi_k(0, X_u) \big] du - 2\int_0^s \phi_k(X_u)dX_u$$

$$= h_k(s, X_s) - h_k(0, X_0)$$

32

with equality if $\sigma_{k-1} \leq s \leq t \leq \sigma_k$.

2. (i) We first argue, for all $(s,x) \in \mathbb{R}_+ \times \mathbb{R}$, that

$$\left\{\varphi_k\big(t,X_t\big)\right\}_{t \geq s} \text{ is a submartingale on } [s,\infty), \text{ and a martingale on } [s,\sigma_{\mathcal{R}^k}], \ \mathbb{P}^{s,x}-\text{a.s.} \qquad (A.7)$$

The martingale property is immediate from the definition of φ_k. The submartingale property follows from the following induction. First, the claim is obvious for $k = n+1$ by the fact that f is non-decreasing. Next, suppose that the submartingale property in (A.7) holds for some $k+1$. Introduce the stopping times $\sigma^t_{\mathcal{R}^k} := \inf\{u \geq t : (u, X_u) \in \mathcal{R}^k\}$, and notice that $\sigma^t_{\mathcal{R}^k} \geq \sigma^r_{\mathcal{R}^k}$ for $s \leq r \leq t$. Then, denoting by $\tilde{X}, \tilde{\sigma}$ independent copies of the same objects, and using the induction hypothesis, we see that:

$$
\begin{aligned}
\mathbb{E}^{(s,x)}\left[\varphi_k(t,X_t)|\mathcal{F}_r\right] = \mathbb{E}^{(s,x)}\left[\mathbb{E}^{(t,X_t)}\left[\varphi_{k+1}(\tilde{\sigma}_{\mathcal{R}^k}, \tilde{X}_{\tilde{\sigma}_{\mathcal{R}^k}})\right]|\mathcal{F}_r\right] &= \mathbb{E}^{(s,x)}\left[\varphi_{k+1}(\sigma^t_{\mathcal{R}^k}, X_{\sigma^t_{\mathcal{R}^k}})|\mathcal{F}_r\right] \\
&\geq \mathbb{E}^{(s,x)}\left[\varphi_{k+1}(\sigma^r_{\mathcal{R}^k}, X_{\sigma^r_{\mathcal{R}^k}})|\mathcal{F}_r\right] \\
&= \varphi_k(r, X_r).
\end{aligned}
$$

(ii) We now prove (A.4). For $u \geq t - s$, it follows from (A.7) that

$$
\begin{aligned}
\mathbb{E}^{\mu_0}_s\left[\varphi_k(u, X_t)\right] = \mathbb{E}^{0,X_s}\left[\varphi_k(u, \tilde{X}_{t-s})\right] &= \mathbb{E}^{u-(t-s),X_s}\left[\varphi_k(u, \tilde{X}_u)\right] \\
&= \mathbb{E}^{u-(t-s),X_s}\left[\varphi_{k+1}(\tilde{\sigma}^u_{\mathcal{R}^k}, \tilde{X}_{\tilde{\sigma}^u_{\mathcal{R}^k}})\right] \\
&\geq \mathbb{E}^{u-(t-s),X_s}\left[\varphi_{k+1}(\tilde{\sigma}_{\mathcal{R}^k}, \tilde{X}_{\tilde{\sigma}_{\mathcal{R}^k}})\right] \\
&= \varphi_k\big(u-(t-s), X_s\big), \ \ \mathbb{P}^{\mu_0}-\text{a.s.}
\end{aligned}
\qquad (A.8)
$$

(iii) We next prove (A.5). For $u \leq t - s$, using again (A.7), we see that:

$$
\begin{aligned}
\mathbb{E}^{\mu_0}_s\left[\varphi_k(u, X_t)\right] &= \mathbb{E}^{\mu_0}_s\left[\mathbb{E}^{0,X_{t-u}}\left[\varphi_k(u, \tilde{X}_u)\right]\right] \\
&\geq \mathbb{E}^{\mu_0}_s\left[\mathbb{E}^{0,X_{t-u}}\left[\varphi_k(0, \tilde{X}_0)\right]\right] \\
&= \mathbb{E}^{\mu_0}_s\left[\varphi_k(0, X_{t-u})\right].
\end{aligned}
\qquad (A.9)
$$

(iv) Finally, to prove (A.6), we observe that the equality was lost in (A.4) and (A.5) only because of the inequalities in (A.8) and (A.9), which in turn become equalities provided that (u, X_u) does not enter $\mathcal{R}^k$ for $u \in [s,t)$. The condition that $\sigma_{k-1} \leq s \leq t \leq \sigma_k$ ensures this is true. $\qquad\square$

Proof of Theorem 3.2. Finally, we complete the proof of the main result in Section 3.2. First, by monotone convergence arguments and since ψ is convex, note that

$$\mathbb{E}^{\mu_0}\left[\rho_n\right] = \int \psi(x)(\mu_n - \mu_0)(dx) = \int (U^{\mu_0}(x) - U^{\mu_n}(x))dx \qquad (A.10)$$

is the same for all $\rho \in \mathcal{T}(\boldsymbol{\mu}_n)$ so that adding a constant to f does not change the problem. We shall normalise f by taking $f(0) = 0$ and exclude the trivial case $f \equiv 0$. If the quantities in (A.10) are equal to $+\infty$ then there is nothing to prove. We thus assume that (A.10) is finite. Note that this might be so even if $\int \psi(x)\mu_i(dx) = \infty$ for each $0 \leq i \leq n$. More generally, thanks to the convex ordering of measures, one can define the integral $\int g(x)(\mu_j - \mu_i)(dx)$ for a convex g and $0 \leq i \leq j \leq n$. This is done by considering $g_k \nearrow g$ which are convex, equal to g on a compact set and affine on the complement. Further, if $h = h - g + g$ with $(h-g)$ and g convex with finite integrals against $(\mu_j - \mu_i)$ then the integral $\int h(x)(\mu_j - \mu_i)(dx)$ is also well defined and finite, see Beiglböck et al. (2016) for details. We shall use this fact below repeatedly together with $\int \psi(x)(\mu_j - \mu_i)(dx) < \infty$ which follows from (A.10).

Without loss of generality, we may assume that f is bounded, the general case follows from a direct monotone convergence argument. Then $0 \leq \varphi_i \leq \|f\|_\infty$ for all i, and in particular, $|\phi_i(x)| \leq \|f\|_\infty |\psi'(x)|$.

We define $\kappa_i(x) := \int_0^x \phi_i(y)dy = -h_i(0,x)/2$, and observe that $\kappa_i(x)$ is then a non-negative, convex function with $0 \le \kappa_i''(x) \le \|f\|_\infty \eta(x)^{-2}$ so that $\|f\|_\infty \psi(x) - \kappa_i(x)$ is a non-negative convex function. We conclude that $\int \kappa_i(x)(\mu_j - \mu_k)(dx) < \infty$, $0 \le k \le j \le n$. Moreover, we have $\kappa_i(x) \ge \kappa_{i+1}(x)$ for all $x \in \bar{\mathcal{I}}$ since $\varphi_i \ge \varphi_{i+1}$, as argued above.

The aim is now to take expectations in (3.7) for $(s_i, x_i) = (\rho_i, X_{\rho_i})$, where $\rho \in \mathcal{T}(\boldsymbol{\mu}_n)$. To do this, we need to check that the expectations under $\mathbb{P}^{\mu_0}$ of individual terms on the right-hand side of (3.7) are well defined.

We can rewrite the first two terms on the right-hand side of (3.7) as:

$$\sum_{i=1}^n \lambda_i(x_i) + h_1(0, x_0) = \sum_{i=1}^n \int_0^{\bar{t}_i(x_i)} \left(\varphi_{i+1}(s, x_i) - \varphi_i(s, x_i)\right) ds + 2\sum_{i=1}^n \int_{x_{i-1}}^{x_i} \phi_i(y) dy,$$

where we used that $f(0) = 0$ so that $\kappa_{n+1} \equiv 0$ and $h_{n+1}(t, x) = \int_0^t f(u)du$. The expectation of the first two terms is then equal to

$$\sum_{i=1}^n \int \int_0^{\bar{t}_i(x)} (\varphi_{i+1}(s, x) - \varphi_i(s, x)) ds \, \mu_i(dx) + 2\sum_{i=1}^n \int \kappa_i(x)(\mu_i - \mu_{i-1})(dx).$$

The integrals in the second sum are well defined and finite by the discussion above. As for the first sum, observe that

$$0 \le \int \int_0^{\bar{t}_i(x)} (\varphi_i(s, x) - \varphi_{i+1}(s, x)) ds \, \mu_i(dx) \le \|f\|_\infty \mathbb{E}^{\mu_0} \left[\bar{t}_i(X_{\sigma_i})\right] \le \|f\|_\infty \mathbb{E}^{\mu_0}\left[\sigma_i\right] < \infty.$$

Using $|\varphi_i| \le \|f\|_\infty$, $\mathbb{E}^{\mu_0}\left[\rho_n\right] < \infty$ and integrability properties of κ_i we see that the local martingale

$$\int_0^t \phi_i(X_u) dX_u = \kappa_i(X_t) - \kappa_i(X_0) - \frac{1}{2}\int_0^t \varphi_i(0, X_u) du,$$

is a martingale on $[0, \rho_n]$. It then follows from Lemma A.1 that

$$\mathbb{E}^{\mu_0}\left[h_i(\rho_i, X_{\rho_i}) - h_i(\rho_{i-1}, X_{\rho_{i-1}})\right] \ge 0, \quad i = 1, \ldots, n,$$

with equality if $\rho_i = \sigma_i$. Taking expectations under $\mathbb{P}^{\mu_0}$ in (3.7), we deduce that

$$\mathbb{E}^{\mu_0}\left[\int_0^{\rho_n} f(t)\, dt\right] \ge \sum_{i=1}^n \int \int_0^{\bar{t}_i(x)} (\varphi_{i+1}(s, x) - \varphi_i(s, x)) ds \, \mu_i(dx) + 2\sum_{i=1}^n \int \kappa_i(x)(\mu_i - \mu_{i-1})(dx),$$

with equality when we replace ρ_n with σ_n. $\qquad\square$

B   Extension to continuous Markov local martingales

The following statement extends Lemma 6.1 to a class of continuous Markov local martingales.

Lemma B.1. *Let X be a local martingale with $d\langle X\rangle_t = \eta(X_t)^2 dt$, for some locally Lipschitz function η, and let $a < b$ be fixed points in $\mathcal{I}^\circ$, and $H_{a,b}$ the first exit time of X from the interval (a, b). Then*

$$\mathbb{E}^x\left|X_{t \wedge H_{a,b}} - y\right| = \mathbb{E}^y\left|X_{t \wedge H_{a,b}} - x\right| \quad \text{for all} \quad x, y \in [a, b].$$

Proof. Le $y \in (a, b)$ be fixed, and denote $X^H := X_{. \wedge H_{a,b}}$. We decompose the proof in three steps.

Step 1: By dominated convergence the function $u(t, x) := \mathbb{E}^x\left|X_t^H - y\right|$ is continuous, and it follows from classical argument using the tower property that u is a viscosity solution of the equation

$$\begin{aligned}
&\left(\partial_t u - \tfrac{1}{2}\eta^2 D^2 u\right)(t, x) = 0 \quad \text{for} \quad t \ge 0, \ x \in (a, b) \\
&u(x, a) = y - a, \quad u(x, b) = b - y, \ x \in (a, b).
\end{aligned} \tag{B.1}$$

Step 2: Similarly, the function $v(t,x) := \mathbb{E}^y |X_t^H - x|$ is a continuous function, and is in addition convex in the $x-$variable. Denote by $L(X^H)$ the local time of the continuous martingale X^H. Using the Itô-Tanaka formula, we see that:

$$v(t+h,x) - v(t,x) \quad = \quad \mathbb{E}^y \big[L_{t+h}^x(X^H) - L_t^x(X^H) \big].$$

By the density occupation formula, this provides for all Borel subset A of $[a,b]$:

$$\int_A \int_t^{t+h} \partial_t v(ds,x) dx \quad = \quad \int_A \big(v(t+h,x) - v(t,x) \big) dx \quad = \quad \int_A \eta^2(x) \int_t^{t+h} \mathbb{P}^{X_s^H}(dx) ds,$$

where $\mathbb{P}^{X_s^H}$ denotes the distribution function of X_s^H. Notice that $\mathbb{P}^{X_s^H} = \frac{1}{2} D^2 v(s,.)$. Then:

$$\int_A \int_t^{t+h} \partial_t v(ds,x) dx \quad = \quad \int_A \int_t^{t+h} \frac{1}{2} \eta^2(x) ds D^2 v(s,dx).$$

Let φ_ε be a $C^\infty-$molifier, and set $v_\varepsilon(t,x) = \int v(t-s,x-y)\varphi_\varepsilon(s,y)dsdy$. Then, v_ε is smooth, and it follows from the last equality that

$$\int_A \int_t^{t+h} \big(\partial_t v_\varepsilon - \frac{1}{2} \eta^2 D^2 v_\varepsilon - R_\varepsilon \big)(s,x) dsdx \quad = \quad 0,$$

where $R_\varepsilon(s,x) := \int \big(\eta^2(x) - \eta^2(x-y) \big) D^2 v(r-s,x-y)\varphi_\varepsilon(r,y)drdy$. Since η is Lipschitz on $[a,b]$, and v is bounded, we see that

$$
\begin{aligned}
\big| R_\varepsilon(s,x) \big| &\leq c \int D^2 \{ |x-y| \varphi_\varepsilon(r-s,x-y) \} drdy \\
&= c \int \big[D\{ |x-y| \varphi_\varepsilon(r-s,x-y) \} \big]_a^b dr =: r_\varepsilon \longrightarrow 0, \quad \text{as} \quad \varepsilon \to 0.
\end{aligned}
$$

By the arbitrariness of $h > 0$ and the Borel subset A of $[a,b]$, this shows that

$$\partial_t v_\varepsilon - \frac{1}{2} \eta^2 D^2 v_\varepsilon - r_\varepsilon \geq 0 \quad \text{and} \quad \partial_t v_\varepsilon - \frac{1}{2} \eta^2 D^2 v_\varepsilon + r_\varepsilon \leq 0 \quad \text{on} \quad \mathbb{R}_+ \times (a,b).$$

Since $v_\varepsilon \longrightarrow v$, locally uniformly, it follows from the stability result of viscosity solutions that v is a viscosity solution of $\partial_t v - \frac{1}{2} \eta^2 D^2 v = 0$ on $\mathbb{R}_+ \times (a,b)$. We also directly see that $v(t,a) = y - a$ and $v(t,b) = b - y$. Hence v is also a viscosity solution of (B.1).

Step 3: To conclude that $u = v$, we now use the fact that equation (B.1) has a unique $C^0(\mathbb{R}_+ \times [a,b])$ viscosity solution. Indeed the corresponding equation satisfied by $e^{\lambda t} u(t,x)$, for an arbitrary $\lambda > 0$, satisfies the conditions of Theorem 8.2 of Crandall et al. (1992). $\qquad \square$

References

Acciaio, B., Beiglböck, M., Penkner, F., Schachermayer, W., and Temme, J. (2013). A trajectorial interpretation of Doob's martingale inequalities. *The Annals of Applied Probability*, 23(4):1494–1505.

Azéma, J. and Yor, M. (1979). Une solution simple au problème de Skorokhod. In *Séminaire de Probabilités, XIII (Univ. Strasbourg, Strasbourg, 1977/78)*, volume 721 of *Lecture Notes in Math.*, pages 90–115. Springer, Berlin.

Beiglböck, M., Cox, A. M. G., and Huesmann, M. (2016). Optimal transport and Skorokhod embedding. *Inventiones mathematicae*. (online first).

Beiglböck, M. and Nutz, M. (2014). Martingale Inequalities and Deterministic Counterparts. *Electron. J. Probab.*, 19(95):1–15.

Beiglböck, M., Nutz, M., and Touzi, N. (2016). Complete duality for martingale optimal transport on the line. *Ann. Probab.* to appear.

Bensoussan, A. and Lions, J.-L. (1982). *Applications of variational inequalities in stochastic control*, volume 12 of *Studies in Mathematics and its Applications*. North-Holland Publishing Co., Amsterdam–New York.

Brown, H., Hobson, D., and Rogers, L. C. G. (2001). The maximum maximum of a martingale constrained by an intermediate law. *Probab. Theory Related Fields*, 119(4):558–578.

Burkholder, D. L. (1991). Explorations in martinagle theory and its applications. In *École d'Été de Probabilités de Saint Flour XIX – 1989*, volume 1464 of *Lecture Notes in Math.*, pages 1–66. Springer Berlin.

Chacon, P. (1985). *The filling scheme and barrier stopping times*. PhD thesis, Univ. Washington.

Chacon, R. V. (1977). Potential processes. *Trans. Amer. Math. Soc.*, 226:39–58.

Cox, A. M. G., Hobson, D. G., and Obłój, J. (2008). Pathwise inequalities for local time: applications to Skorokhod embeddings and optimal stopping. *Ann. Appl. Probab.*, 18(5):1870–1896.

Cox, A. M. G. and Wang, J. (2013a). Optimal robust bounds for variance options. arXiv: 1308.4363.

Cox, A. M. G. and Wang, J. (2013b). Root's barrier: Construction, optimality and applications to variance options. *Annals of Applied Probability*, 23(3):859–894.

Crandall, M., Ishii, H., and Lions, P.-L. (1992). User's guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc.*, 27(1):1–67.

De Angelis, T. (2015). From optimal stopping boundaries to Rost's reversed barriers and the Skorokhod embedding. arXiv: 1505.02724.

El Karoui, N. (1981). Les aspects probabilistes du contrôle stochastique. In Hennequin, P., editor, *Ecole d'été de Probabilités de Saint-Flour IX-1979*, volume 876 of *Lecture Notes in Mathematics*, pages 73–238. Springer Berlin Heidelberg.

Elworthy, K. D., Li, X.-M., and Yor, M. (1999). The importance of strictly local martingales; applications to radial Ornstein–Uhlenbeck processes. *Probability Theory and Related Fields*, 115(3):325–355.

Gassiat, P., Oberhauser, H., and dos Reis, G. (2014). Root's barrier, viscosity solutions of obstacle problems and reflected FBSDEs. arXiv:1301.3798v4.

Henry-Labordère, P., Obłój, J., Spoida, P., and Touzi, N. (2016). The maximum maximum of a martingale with given n marginals. *Ann. Appl. Prob*, 26(1):1–44.

Hobson, D. (2011). The Skorokhod Embedding Problem and Model-Independent Bounds for Option Prices. In Carmona, R., Çinlar, E. amd Ekeland, I., Jouini, E., Scheinkman, J., and Touzi, N., editors, *Paris-Princeton Lectures on Mathematical Finance 2010*, volume 2003 of *Lecture Notes in Math.*, pages 267–318. Springer.

Hobson, D. G. (1998). Robust hedging of the lookback option. *Finance Stoch.*, 2(4):329–347.

Karatzas, I. and Shreve, S. E. (1991). *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag.

Loynes, R. M. (1970). Stopping times on Brownian motion: Some properties of Root's construction. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 16:211–218.

Madan, D. and Yor, M. (2002). Making Markov martingales meet marginals: with explicit constructions. *Bernoulli*, 8(4):509–536.

McConnell, T. R. (1991). The two-sided Stefan problem with a spatially dependent latent heat. *Trans. Amer. Math. Soc.*, 326(2):669–699.

Obłój, J. (2004). The Skorokhod embedding problem and its offspring. *Probab. Surv.*, 1:321–390 (electronic).

Obłój, J. (2007). The Maximality Principle Revisited: On Certain Optimal Stopping Problems. In Donati-Martin, C., Émery, M., Rouault, A., and Stricker, C., editors, *Séminaire de Probabilités, XL*, number 1899 in Lecture Notes in Mathematics, pages 309–328. Springer Berlin Heidelberg.

Obłój, J. and Spoida, P. (2016). An iterated Azéma-Yor type embedding for finitely many marginals. *Ann. Probab.* to appear.

Obłój, J., Spoida, P., and Touzi, N. (2015). Martingale inequalities for the maximum via pathwise arguments. In Donati-Martin, C., Lejay, A., and Rouault, A., editors, *In Memoriam Marc Yor - Séminaire de Probabilités XLVII*, volume 2137 of *Lecture Notes in Mathematics*, pages 227–247. Springer Berlin Heidelberg.

Peskir, G. (1999). Designing options given the risk: the optimal Skorokhod-embedding problem. *Stochastic Process. Appl.*, 81(1):25–38.

Protter, P. E. (2005). *Stochastic Integration and Differential Equations.* Springer Science & Business Media, second edition edition. Google-Books-ID: mJkFuqwr5xgC.

Rogers, L. C. G. and Williams, D. (2000). *Diffusions, Markov processes, and martingales. Vol. 2.* Cambridge Mathematical Library. Cambridge University Press, Cambridge. Itô calculus, Reprint of the second (1994) edition.

Root, D. H. (1969). The existence of certain stopping times on Brownian motion. *Ann. Math. Statist.*, 40:715–718.

Rost, H. (1971). The stopping distributions of a Markov Process. *Invent. Math.*, 14:1–16.

Rost, H. (1976). Skorokhod stopping times of minimal variance. In *Séminaire de Probabilités, X (Première partie, Univ. Strasbourg, Strasbourg, année universitaire 1974/1975)*, pages 194–208. Lecture Notes in Math., Vol. 511. Springer, Berlin.

Skorokhod, A. V. (1965). *Studies in the theory of random processes.* Translated from the Russian by Scripta Technica, Inc. Addison-Wesley Publishing Co., Inc., Reading, Mass.

Touzi, N. (2012). *Optimal Stochastic Control, Stochastic Target Problems, and Backward SDE*, volume 29. Springer Science & Business Media.