MARTINGALE OPTIMAL TRANSPORT WITH STOPPING*

ERHAN BAYRAKTAR†, ALEXANDER M. G. COX‡, AND YAVOR STOEV†

Abstract. We solve the martingale optimal transport problem for cost functionals represented by optimal stopping problems. The measure-valued martingale approach developed in [A. M. G. Cox and S. Källblad, *SIAM J. Control Optim.*, 55 (2017), pp. 3409–3436] allows us to obtain an equivalent infinite dimensional controller-stopper problem. We use the stochastic Perron's method and characterize the finite dimensional approximation as a viscosity solution to the corresponding HJB equation. It turns out that this solution is the concave envelope of the cost function with respect to the atoms of the terminal law. We demonstrate the results by finding explicit solutions for a class of cost functions.

Key words. martingale optimal transport, dynamic programming, optimal stopping, stochastic Perron's method, viscosity solutions, concave envelope, distribution constraints

AMS subject classifications. 60G40, 93E20, 91A10, 91A60, 60G07

DOI. 10.1137/17M1114065

1. Introduction. The aim of this paper is to solve a class of martingale optimal transport problems for which the cost functional can be represented as an optimal stopping problem of the underlying cost function. Specifically, given a continuous and bounded cost function $f : \mathbb{R} \to \mathbb{R}$, we are interested in solving the martingale optimal transport problem

$$(1.1) \qquad \sup_{P_\mu} P^{\mathbb{P}}(f) \quad \text{with} \quad P^{\mathbb{P}}(f) = \sup_{\tau \in \mathcal{T}_0} \mathbb{E}[f(M_\tau)].$$

The outer supremum is taken over P_μ—the set of all pairs of filtered probability spaces $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ and continuous martingales $M = (M_t)_{t\geq 0}$ on them such that the filtration $(\mathcal{F}_t)_{t\geq 0}$ is generated by a Brownian motion and the terminal law is $M_T \sim \mu$ under $\mathbb{P}$. The inner stopping problem is over $\mathcal{T}_s$—the set of all $(\mathcal{F}_t)$-stopping times taking values in $[s, T]$ for $s \in [0, T]$ and some fixed terminal time $T > 0$.

The duality between martingale optimal transport and robust pricing problems was studied in a related setting in Dolinsky and Soner [8] for general path-dependent European-type cost functionals (i.e., payoffs) and continuous models. Recently Bayraktar and Miller [1] and Beiglböck et al. [5] obtained solutions to distribution-constrained optimal stopping problems by using dynamic programming and martingale transport methods, respectively. In contrast to our setting, however, the constraints in [1] and [5] are on the distribution of the stopping times and not on the marginal distribution at the terminal time. By using the concept of measure-valued martingales, Cox and Källblad [6] studied the robust pricing of Asian-type options subject to a marginal distribution constraint. The authors cast the original problem into a control theoretic framework and obtained a viscosity characterization of the solution.

†Department of Mathematics, University of Michigan, Ann Arbor, MI 48109 (erhan@umich.edu, ystoev@umich.edu).

‡Department of Mathematical Sciences, University of Bath, Bath BA2 7AY, UK (a.m.g. cox@bath.ac.uk).

Here we employ the control theoretic approach of [6] and [1] to analyze optimal martingale transport problems with cost functionals which are of American type. The difficulty in our setting is that we have an additional optimal stopping component. However, the fact that we optimize over continuous models allows us to prove that the resulting value function is time-independent up to the terminal time. Since the original problem is infinite dimensional, we use the continuity with respect to the terminal law to restrict it only to measures with finitely many atoms. Working in a Brownian filtration allows us to recast this finite dimensional approximation as a recursive sequence of controller-stopper problems with exit-time components. We prove that the value functions of these problems are viscosity solutions to the corresponding sequence of elliptic obstacle problems satisfying exact Dirichlet boundary conditions. We achieve this by applying the stochastic Perron's approach in the spirit of Bayraktar and Sirbu [4], where the obstacle problems are associated with Dynkin games, and Rokhlin [14], where an elliptic Dirichlet boundary problem arose from exit-time stochastic control. We circumvent the potential difficulty of proving a strong comparison result for viscosity sub/supersolutions satisfying *generalized* boundary conditions (see [14]) by using the recursive structure of the problem to show the exact attainment of these boundary conditions.

The main result in this paper, Theorem 3.1, is the characterization of the value function of the finite dimensional martingale transport problem as the concave envelope of the payoff with respect to the probability weights of the terminal law's atoms. In this final step we use a recent result of Oberman and Ruan [11] on characterizing convex envelopes as unique viscosity solutions to obstacle problems with appropriate Dirichlet boundary conditions. One possible application of our results is the robust pricing of American options. Indeed, the martingales over which we optimize can be seen as different models for the stock price with a given marginal distribution at the terminal time.

The rest of this paper is organized as follows: In section 2, we formulate the finite dimensional approximation of the martingale optimal transport problem; see (2.12). In section 3, we employ the stochastic Perron's method to characterize the value function as the unique viscosity solution of the corresponding Dirichlet obstacle problem and to show its concave envelope form in an appropriate phase space. Section 4 illustrates how our results can be achieved in a probabilistic framework and provides concrete examples.

2. Problem formulation. We define the set of measures $\mathcal{P}$ as

$$\mathcal{P} := \{\mu \in \mathcal{B}(\mathbb{R}_+) : \mu(\mathbb{R}_+) = 1 \text{ and } \int |x|\mu(dx) < \infty\},$$

and suppose that the terminal law μ of the martingales in the optimal transport problem (1.1) satisfies $\mu \in \mathcal{P}$. In the usual optimal transport framework we can regard the probability measures $\mathbb{P}$ contained in P_μ as transporting the initial Dirac measure δ_{M_0} (i.e., the law of M_0) to the terminal law μ under the cost functional $P^{\mathbb{P}}$—both of these laws are known at time $t = 0$. On the other hand, notice that the continuous martingale M satisfies

$$(2.1) \qquad M_t = \mathbb{E}[M_T|\mathcal{F}_t] = \int x\,\xi_t(dx) \quad \text{for} \quad t \in [0, T],$$

where ξ_t is the conditional law of M_T given $\mathcal{F}_t$ under the measure $\mathbb{P}$. In particular, we have that $\xi_0 = \mu$ and $\xi_T = \delta_{M_T}$. Therefore, similarly to the method proposed in

[6], we can rewrite (1.1) in its measure-valued martingale formulation as

$$(2.2) \qquad \sup_{(\xi_t)\in\Xi} \sup_{\tau\in\mathcal{T}_0} \mathbb{E}[f(M_\tau)] \quad \text{subject to} \quad \xi_0 = \mu,$$

where Ξ is the set of all terminating measure-valued (i.e., $\mathcal{P}$-valued) martingales (see Definition 2.7 in [6]) such that $(\int x\,\xi_t(dx))_{t\geq 0}$ is a continuous process a.s. with respect to the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ for all $(\xi_t)_{t\geq 0} \in \Xi$, where $(\mathcal{F}_t)_{t\geq 0}$ is a Brownian filtration. Moreover, as in [6], we fix the probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ which does not materially change our conclusions.

Let us write (2.2) in the Markovian form

$$(2.3) \qquad U(t,\xi) = \sup_{(\xi_r)\in\Xi} \sup_{\tau\in\mathcal{T}_t} \mathbb{E}[f(M_\tau)|\xi_t = \xi],$$

and note that we have the following variant of Lemma 3.1 in [6], the proof of which can be found in the appendix.

LEMMA 2.1. *If f is nonnegative and Lipschitz, then the function U is continuous in ξ (in the Wasserstein-1 topology) and independent of t for $t \in [0,T)$.*

The continuity in ξ allows us to apply the finite dimensional reduction from section 3.2 in [6]. In particular, we introduce the set $\mathbb{X}_N = \{x_0, \ldots, x_N\}$, where $0 \leq x_0 < x_1 < \cdots < x_N$, and let $\mathcal{P}^N = \mathcal{P} \cap \mathcal{M}(\mathbb{X}_N)$ and $\mathcal{P}(\mathbb{X}_\alpha) = \mathcal{P} \cap \mathcal{M}(\mathbb{X}_\alpha)$ for any $\alpha \subseteq \{0, 1, \ldots, N\}$, where $\mathcal{M}(\mathbb{X}_N)$, resp., $\mathcal{M}(\mathbb{X}_\alpha)$, denote the sets of all measures on $\mathbb{X}_N$, resp., $\mathbb{X}_\alpha := \{x_i : i \in \alpha\}$. We assume from now on that the terminal law ξ (i.e., also μ) is an atomic measure and satisfies $\xi \in \mathcal{P}^N$. Since we work in a Brownian filtration, by martingale representation for any terminating $\mathcal{P}^N$-valued martingale $(\xi_t)_{t\geq 0}$ it is true that the (nonnegative) martingales $\xi_t^n := \xi_t(\{x_n\})$ solve an SDE of the form

$$(2.4) \qquad d\xi_t^n = w_t^n dW_t$$

for $t \geq 0$ and $n = 0, \ldots, N$, where the vector of weights $\mathbf{w}_t = (w_t^0, \ldots, w_t^N)$ satisfies $\sum_{n=0}^N w_t^n = 0$, and $\xi_t^n \in \{0, 1\}$ implies that $w_t^n = 0$. The following result, by analogy to Corollary 3.6 in [6], follows directly from Lemma 3.4 in [6] and allows us to work with a bounded set of controls.

LEMMA 2.2. *Under the above assumption that $\mu \in \mathcal{P}^N$, the value function in (2.3) for $t \in [0, T)$ reduces to the value function*

$$(2.5) \qquad V(\xi) = \sup_{\mathbf{w}\in\mathcal{A}} \sup_{\tau\in\mathcal{T}_0} \mathbb{E}\left[f\left(\sum_{j=0}^N x_j\, \xi_{T_\tau^{-1}}^j\right)|\xi_0 = \xi\right],$$

where the admissible control set $\mathcal{A}$ is defined as

$$\mathcal{A} := \{(\mathbf{w}_r)_{r\geq 0} \text{ prog. meas.} : \mathbf{w}_r \in \mathrm{cl}(\mathbb{D}^{N+1}), \xi_r^n \in \{0,1\} \text{ implies } w_r^n = 0\},$$

with the disk $\mathbb{D}^{k+1}$ being the intersection of the open unit ball with the hyperplane $z_1 + \cdots + z_{k+1} = 0$ in $\mathbb{R}^{k+1}$, and T_r^{-1} is the continuous inverse of

$$(2.6) \qquad T_r := \int_0^r \lambda_s ds \quad \text{for} \quad r \geq 0,$$

where the strictly positive time change rate process $\lambda = (\lambda_r)_{r\geq 0}$ satisfies

$$(2.7) \qquad \|\mathbf{w}_r\|^2 + \lambda_r = 1 - I_{\{\xi_r = \delta_{x_i}\}} I_{\{T_r = T\}}.$$

The role of the time change in (2.6) is to stretch/compress the original time scale so as to bound the volatility of the state process (i.e., the control process $\mathbf{w}$). Thus we avoid technical difficulties arising from unbounded control sets later when proving the viscosity characterization of the value function.

Now notice that the value function $V(\xi)$ can be identified with $\tilde{V}_N(\xi)$ where for $k = 1, \ldots, N$, and $\xi \in \mathcal{P}(\mathbb{X}_\alpha)$, with $|\alpha| = k+1$, we introduce the sequence of problems

$$(2.8) \quad \tilde{V}_k(\xi) = \sup_{\mathbf{w} \in \mathcal{A}^\alpha} \sup_{\tau \in \mathcal{T}_0} \mathbb{E}\left[\tilde{V}_{k-1}(\xi_\sigma)I_{\{T_\sigma \leq \tau\}} + f\left(\sum_{j=0}^{N} x_j\, \xi^j_{T_\tau^{-1}}\right) I_{\{T_\sigma > \tau\}} \Big| \xi_0 = \xi\right],$$

with

$$(2.9) \qquad \mathcal{A}^\alpha := \{(\mathbf{w}_r)_{r \geq 0} \text{ prog. meas.} : \mathbf{w}_r \in \mathrm{cl}(\mathbb{D}^{N+1}),$$
$$w^i \equiv 0 \text{ for any } i \in \{0, 1, \ldots, N\} \setminus \alpha\},$$
$$(2.10) \qquad \sigma := \inf\{s \geq 0 : \xi_s \in \mathcal{P}(\mathbb{X}_{\alpha'}) \text{ for some } \alpha' \text{ with } |\alpha'| \leq k \text{ or } T_s = T\},$$

and $\tilde{V}_0(\xi) = f(x_i)$ for $\xi = \delta_{x_i}$. From now on we will denote the time changed filtration as $(\mathcal{G}_t)_{t \geq 0} := (\mathcal{F}_{T_t})_{t \geq 0}$ and suppress its dependence on λ for notational purposes. The following lemma shows that we can ignore controls which are small enough and that we can work with stopping times in the time changed filtration.

LEMMA 2.3. *The value function* $\tilde{V}_k(\xi)$ *can be written as*

$$(2.11) \quad \tilde{V}_k(\xi) = \sup_{\mathbf{w} \in \mathrm{int}(\mathcal{A}^\alpha_\varepsilon)} \sup_{\tau \in \mathcal{T}} \mathbb{E}\left[\tilde{V}_{k-1}(\xi_\sigma)I_{\{\sigma \leq \tau\}} + f\left(\sum_{j=0}^{N} x_j\, \xi^j_\tau\right) I_{\{\sigma > \tau\}} \Big| \xi_0 = \xi\right],$$

where $\mathrm{int}(\mathcal{A}^\alpha_\varepsilon) := \{(\mathbf{w}_r)_{r \geq 0} \in \mathcal{A}^\alpha : \mathbf{w}_r \in \mathbb{D}^{N+1}, \xi_r \neq \delta_{x_i} \text{ implies } \|\mathbf{w}_r\| \geq \varepsilon\}$ *for any* $\varepsilon \in [0, 1)$ *and* $\mathcal{T}$ *is the set of all* $(\mathcal{G}_t)$-*stopping times for an appropriately time changed filtration* $(\mathcal{G}_t)_{t \geq 0}$.

Proof. For any time change rate λ we have $\lambda_u > 0$ for $u \geq 0$, and from (2.7) it follows that $\|\mathbf{w}_u\| < 1$. Moreover, since λ is strictly positive, we have that T_r and T_t^{-1} are strictly increasing. It follows immediately that if $\tau \in [0, T]$ is an $(\mathcal{F}_t)$-stopping time, then $T_\tau^{-1} \geq 0$ is a $(\mathcal{G}_t)$-stopping time and, conversely, if $\tau \geq 0$ is a $(\mathcal{G}_t)$-stopping time, then $T_\tau \in [0, T]$ is an $(\mathcal{F}_t)$-stopping time. Therefore, in (2.8) we can substitute $\mathcal{T}_0$ with $\mathcal{T}$ and τ with T_τ.

What is left is to prove that we can take the outer supremum in (2.8) over $\mathrm{int}(\mathcal{A}^\alpha_\varepsilon) \subset \mathrm{int}(\mathcal{A}^\alpha)$. For $0 < \varepsilon < 1$ and any $\mathbf{w} \in \mathrm{int}(\mathcal{A}^\alpha) \setminus \mathrm{int}(\mathcal{A}^\alpha_\varepsilon)$ we can choose $\tilde{\mathbf{w}} \in \mathrm{int}(\mathcal{A}^\alpha_\varepsilon)$ defined as $\tilde{\mathbf{w}}_s^n := \sqrt{\bar{\varepsilon}_s}\mathbf{w}^n_{\phi(s)}$ where

$$\phi(s) = \int_0^s \bar{\varepsilon}_u du \quad \text{with} \quad \bar{\varepsilon}_s = \frac{\varepsilon^2}{\|\mathbf{w}_{\phi(s)}\|^2},$$

and $\phi(s)$ is the right-continuous inverse of the (nonstrictly) increasing continuous function $\phi^{-1}(s)$ given by

$$\phi^{-1}(s) = \int_0^s \frac{\|\mathbf{w}_u\|^2}{\varepsilon^2} du.$$

From (2.4) we see that ξ_r^n (corresponding to the control $\mathbf{w}$) has the same distribution as $\tilde{\xi}^n_{\phi^{-1}(r)}$ (corresponding to the control $\tilde{\mathbf{w}}$). Hence, for any $(\mathcal{G}_t)$-stopping time τ we have that $\tilde{\tau} = \phi^{-1}(\tau)$ is a $(\mathcal{G}_{\phi(t)})$-stopping time such that ξ_τ^n has the same law as $\tilde{\xi}^n_{\tilde{\tau}}$. We conclude from (2.8). $\square$

Before going further, we introduce some additional notation. Let $\alpha(\xi)$ be the subset of elements in $\mathbb{X}_N$ to which the atomic measure $\xi \in \mathcal{P}^N$ prescribes nonzero probability, and notice that we have the consistency conditions

$$\tilde{V}_k(\xi) = \tilde{V}_{|\alpha(\xi)|-1}(\xi) \quad \text{for} \quad k \geq |\alpha(\xi)|.$$

For every $\xi \in \mathcal{P}^N$ with $|\alpha(\xi)| = k+1$ it is true that $\xi = \sum_{j=0}^{k} \xi^{i_j} \delta_{x_{i_j}}$, where $\alpha(\xi) = \{x_{i_0}, \dots, x_{i_k}\} \subseteq \mathbb{X}_N$. Hence, we can identify every $\xi \in \mathcal{P}^N$ with the vector $\xi^{\alpha} := (\xi^{i_0}, \xi^{i_1}, \dots, \xi^{i_k}) \in \text{int}(\Delta^{k+1})$, where $\alpha = \{i_0, \dots, i_k\}$ and $\Delta^{k+1} := \{\mathbf{z} \in \mathbb{R}_{\geq 0}^{k+1} : \sum z_i = 1\}$. We let

$$(2.12) \qquad V_{\alpha}(\xi^{\alpha}) = \tilde{V}_{|\alpha(\xi)|-1}(\xi), \quad \bar{f}(\xi^{\alpha}) = f(\mathbf{x}^{\alpha} \cdot \xi^{\alpha}),$$

where $\mathbf{x}^{\alpha} := (x_{i_0}, \dots, x_{i_k})$. For any $r \geq 0$ and $\mathbf{w} = (w^0, \dots, w^N) \in \text{int}(\mathcal{A}^{\alpha})$ we also let $\boldsymbol{\xi}_u^{\mathbf{w},r,\xi^{\alpha}} := (\xi_u^{i_0,w^{i_0},r}, \xi_u^{i_1,w^{i_1},r}, \dots, \xi_u^{i_k,w^{i_k},r})$, where $\xi_u^{i_j,w^{i_j},r}$ is the unique strong solution to (2.4) with control w^{i_j} and initial condition $\xi_u^{i_j,w^{i_j},r} = \xi^{i_j}$ for $u \leq r$. Denote by $\xi^{\mathbf{w},r,\xi^{\alpha}}$ the $\mathcal{P}^N$-valued martingale corresponding to $\boldsymbol{\xi}^{\mathbf{w},r,\xi^{\alpha}}$, i.e., $\xi_u^{\mathbf{w},r,\xi^{\alpha}} := \sum_{j=0}^{k} \xi_u^{i_j,w^{i_j},r} \delta_{x_{i_j}}$. For short we let $\boldsymbol{\xi}^{\mathbf{w},\xi^{\alpha}} := \boldsymbol{\xi}^{\mathbf{w},0,\xi^{\alpha}}$ and $\xi^{\mathbf{w},\xi^{\alpha}} := \xi^{\mathbf{w},0,\xi^{\alpha}}$.

3. Viscosity characterization of the value function using stochastic Perron's method. We want to obtain the viscosity characterization of the value function V_{α}. Fix $0 < c < 1$ and $\alpha \subseteq \{0, \dots, N\}$ with $|\alpha| = k+1 \geq 2$ for some integer $k \geq 1$. Using (2.12) rewrite the value function from (2.11) as

$$(3.1) \qquad V_{\alpha}(\xi^{\alpha}) = \sup_{\mathbf{w} \in \text{int}(\mathcal{A}_c^{\alpha})} \sup_{\tau \in \mathcal{T}} \mathbb{E}\left[\tilde{V}_{k-1}(\xi_{\sigma}^{\mathbf{w},r,\xi^{\alpha}}) I_{\{\sigma \leq \tau\}} + \bar{f}(\boldsymbol{\xi}_{\tau}^{\mathbf{w},r,\xi^{\alpha}}) I_{\{\sigma > \tau\}} \right],$$

where $\xi^{\alpha} \in \Delta^{k+1}$. Our aim is to show that V_{α} is the unique viscosity solution (see, e.g., Definition 7.4 in [7]) to the associated Dirichlet obstacle problem given by

$$(3.2) \quad \min\left\{ -\sup_{\mathbf{w} \in \mathbb{D}_c^{k+1}} \frac{1}{2} \text{tr}(\mathbf{w}\mathbf{w}' D_{\boldsymbol{\xi}}^2 V_{\alpha}), V_{\alpha} - \bar{f} \right\} = 0 \quad \text{on} \quad \text{int}(\Delta^{k+1}),$$

$$(3.3) \qquad V_{\alpha}(\xi^{\alpha}) = g(\xi^{\alpha}) := V_{\alpha'}(\xi^{\alpha'}) \quad \text{on} \quad \partial\Delta^{k+1},$$

where $\xi^{\alpha'}$ and α' correspond to the nonzero components of ξ^{α} and α, and $\mathbb{D}_c^{k+1} := \{\mathbf{w} \in \mathbb{D}^{k+1} : \|\mathbf{w}\| > c\}$. The derivative $D_{\boldsymbol{\xi}}^2$ is to be understood in the directional sense—i.e., we restrict ourselves to second directional derivatives $\text{tr}(\mathbf{w}\mathbf{w}' D_{\boldsymbol{\xi}}^2)$ with respect to directions lying in the set $\mathbb{D}_c^{k+1}$.

We are now ready to state the main result of this paper—its proof relies on the stochastic Perron's method, and we present it in the next section.

THEOREM 3.1. *The function $V_{\alpha} : \Delta^{k+1} \to \mathbb{R}$ defined in (3.1) is the unique continuous viscosity solution of the obstacle problem (3.2) satisfying the Dirichlet boundary condition (3.3). Moreover, V_{α} is the concave envelope of $\bar{f}$ on Δ^{k+1}—i.e., denoting the projection of Δ^{k+1} onto $\mathbb{R}_{\geq 0}^k$ by $\tilde{\Delta}^k$ and the projected functions $\tilde{V}_{\alpha}, \tilde{f} : \tilde{\Delta}^k \to \mathbb{R}$ as*

$$(3.4) \qquad \tilde{V}_{\alpha}(z_0, \dots, z_{k-1}) := V_{\alpha}\left(z_0, \dots, z_{k-1}, 1 - \sum_{i=1}^{k-1} z_i \right),$$

$$(3.5) \qquad \tilde{f}(z_0, \dots, z_{k-1}) := \bar{f}\left(z_0, \dots, z_{k-1}, 1 - \sum_{i=1}^{k-1} z_i \right),$$

the function $\tilde{V}_\alpha$ is the concave envelope of $\tilde{f}$.

3.1. Proof of Theorem 3.1. We begin by introducing the notions of stochastic sub- and supersolutions.

DEFINITION 3.2. *The set of stochastic subsolutions to the PDE* (3.2) *with the boundary condition* (3.3), *denoted by* $\mathcal{V}^-$, *is the set of functions* $v : \Delta^{k+1} \to \mathbb{R}$ *that have the following properties:*
  (i) *They are continuous and bounded, and satisfy the boundary condition*

$$(3.6) \qquad\qquad v(\xi^\alpha) \le g(\xi^\alpha) \quad on \quad \partial\Delta^{k+1}.$$

  (ii) *For each* $\tau \in \mathcal{T}$ *and* $\boldsymbol{\xi} \in \mathcal{G}_\tau$ *with* $\mathbb{P}(\boldsymbol{\xi} \in \Delta^{k+1}) = 1$ *there exists a control* $\mathbf{w} \in \mathrm{int}(\mathcal{A}^\alpha)$ *such that for any* $\rho \in \mathcal{T}$ *with* $\rho \in [\tau, \sigma(\tau, \boldsymbol{\xi}, \mathbf{w})]$ *we have a.s. that*

$$(3.7) \qquad\qquad v(\boldsymbol{\xi}) \le \mathbb{E}[v(\boldsymbol{\xi}^{\mathbf{w},\tau,\boldsymbol{\xi}}_{\rho\wedge\tau_*(v)})|\mathcal{G}_\tau],$$

  where the $(\mathcal{G}_t)$-*stopping times* $\sigma(\tau, \boldsymbol{\xi}, \mathbf{w})$ *and* $\tau_*(v)$ *are defined as*

$$(3.8) \qquad \sigma(\tau, \boldsymbol{\xi}, \mathbf{w}) := \inf\{s \ge \tau : \boldsymbol{\xi}^{\mathbf{w},\tau,\boldsymbol{\xi}}_s \notin \mathrm{int}(\Delta^{k+1})\},$$
$$(3.9) \qquad \tau_*(v) \equiv \tau_*(v; \tau, \boldsymbol{\xi}, \mathbf{w}) := \inf\{s \ge \tau : v(\boldsymbol{\xi}^{\mathbf{w},\tau,\boldsymbol{\xi}}_s) \le \bar{f}(\boldsymbol{\xi}^{\mathbf{w},\tau,\boldsymbol{\xi}}_s)\}.$$

DEFINITION 3.3. *The set of stochastic supersolutions to the PDE* (3.2) *with the boundary condition* (3.3), *denoted by* $\mathcal{V}^+$, *is the set of functions* $v : \Delta^{k+1} \to \mathbb{R}$ *that have the following properties:*
  (i) *They are continuous and bounded, and satisfy the boundary condition*

$$(3.10) \qquad\qquad v(\xi^\alpha) \ge g(\xi^\alpha) \quad on \quad \partial\Delta^{k+1}.$$

  (ii) *For each* $\tau \in \mathcal{T}$ *and* $\boldsymbol{\xi} \in \mathcal{G}_\tau$ *with* $\mathbb{P}(\boldsymbol{\xi} \in \Delta^{k+1}) = 1$, *for any control* $\mathbf{w} \in \mathrm{int}(\mathcal{A}^\alpha_c)$ *and any* $\rho \in \mathcal{T}$ *with* $\rho \in [\tau, \sigma(\tau, \boldsymbol{\xi}, \mathbf{w})]$ *we have a.s. that*

$$(3.11) \qquad\qquad v(\boldsymbol{\xi}) \ge \mathbb{E}[v(\boldsymbol{\xi}^{\mathbf{w},\tau,\boldsymbol{\xi}}_\rho)|\mathcal{G}_\tau],$$

  where $\sigma(\tau, \boldsymbol{\xi}, \mathbf{w})$ *is defined as in* (3.8).

Clearly, $\mathcal{V}^-$ (resp., $\mathcal{V}^+$) is nonempty since $\bar{f}$ is bounded from below (resp., above) and any constant which is small (large) enough belongs to $\mathcal{V}^-$ (resp., $\mathcal{V}^+$). Actually, we can easily verify that $\bar{f} \in \mathcal{V}^-$. The following lemma proves an important property of the sets $\mathcal{V}^-$ and $\mathcal{V}^+$.

LEMMA 3.4. *For any two* $v^1, v^2 \in \mathcal{V}^-$ *we have that* $v^1 \vee v^2 \in \mathcal{V}^-$. *For any two* $v^1, v^2 \in \mathcal{V}^+$ *we have that* $v^1 \wedge v^2 \in \mathcal{V}^+$.

Proof. We will only prove the first part of the lemma—the second part follows in a similar way. Denote $v = v^1 \vee v^2$ and notice that item (i) in Definition 3.2 is clearly satisfied by v. Now fix $\tau \in \mathcal{T}$ and $\boldsymbol{\xi} \in \mathcal{G}_\tau$ as in item (ii) of Definition 3.2 and introduce the sequence of stopping time, control, and state process triples $(\gamma_n, \mathbf{w}^n, \boldsymbol{\xi}^n)_{n\ge-1}$ defined recursively as follows:

$$(\gamma_{-1}, \mathbf{w}^{-1}, \boldsymbol{\xi}^{-1}) \equiv (\gamma_0, \mathbf{w}^0, \boldsymbol{\xi}^0) := (\tau, \mathbf{1}_{\{v^1(\boldsymbol{\xi})\ge v^2(\boldsymbol{\xi})\}}\mathbf{w}^{0,1} + \mathbf{1}_{\{v^1(\boldsymbol{\xi})<v^2(\boldsymbol{\xi})\}}\mathbf{w}^{0,2}, \boldsymbol{\xi}^{\mathbf{w}^0,\tau,\boldsymbol{\xi}}),$$

where $\mathbf{w}^{0,1}, \mathbf{w}^{0,2}$ are the controls corresponding to the stochastic subsolutions v^1, v^2

starting at the pair $(\tau, \boldsymbol{\xi})$, and for $n = 0, 1, 2, \ldots$, the following hold:

(i) If $v(\boldsymbol{\xi}_{\gamma_n}^n) \le \bar{f}(\boldsymbol{\xi}_{\gamma_n}^n)$, then we set

$$(\gamma_{n+1}, \mathbf{w}^{n+1}, \boldsymbol{\xi}^{n+1}) := (\gamma_n, \mathbf{w}^n, \boldsymbol{\xi}^n).$$

(ii) If $v(\boldsymbol{\xi}_{\gamma_n}^n) = v^i(\boldsymbol{\xi}_{\gamma_n}^n) > \bar{f}(\boldsymbol{\xi}_{\gamma_n}^n)$ for $i \in \{1, 2\}$, then we set

$$\gamma_{n+1} := \sigma(\gamma_n, \boldsymbol{\xi}_{\gamma_n}^n, \mathbf{w}^n) \wedge \tau_*(v^i; \gamma_n, \boldsymbol{\xi}_{\gamma_n}^n, \mathbf{w}^n),$$

$$\mathbf{w}^{n+1} := \mathbf{w}^{n+1,i}, \quad \boldsymbol{\xi}^{n+1} := \boldsymbol{\xi}^{\mathbf{w}^{n+1,i}, \gamma_{n+1}, \boldsymbol{\xi}_{\gamma_{n+1}}^n},$$

where $\mathbf{w}^{n+1,i}$ is the control process corresponding to the stochastic subsolution v^i starting at the pair $(\gamma_{n+1}, \boldsymbol{\xi}_{\gamma_{n+1}}^n)$, and $\tau_*(v^i; \gamma_n, \boldsymbol{\xi}_{\gamma_n}^n, \mathbf{w}^n)$ is defined as in (3.9).

Define the control $\mathbf{w}$ by

$$\mathbf{w}_s := \sum_{n=1}^{\infty} \mathbf{1}_{\{s \in [\gamma_n, \gamma_{n+1})\}} \mathbf{w}_s^n$$

and notice that by construction, $\boldsymbol{\xi}_s^n = \boldsymbol{\xi}_s^{\mathbf{w}, \tau, \boldsymbol{\xi}}$ for $s \in [\gamma_n, \gamma_{n+1}]$ and any $n \ge 0$. For any stopping time $\rho \in [\tau, \sigma(\tau, \boldsymbol{\xi}, \mathbf{w})]$ denote $\rho \wedge \gamma_n = \rho_n$. By the definition of the sequence $(\gamma_n, \mathbf{w}^n, \boldsymbol{\xi}^n)$ we get that

$$\begin{aligned}
v(\boldsymbol{\xi}_{\rho_n}^n) &= (\mathbf{1}_{\{v^1 \ge v^2\}} v^1 + \mathbf{1}_{\{v^1 < v^2\}} v^2)(\boldsymbol{\xi}_{\rho_n}^n) \\
&\le \mathbb{E}[(\mathbf{1}_{\{v^1(\boldsymbol{\xi}_{\rho_n}^n) \ge v^2(\boldsymbol{\xi}_{\rho_n}^n)\}} v^1 + \mathbf{1}_{\{v^1(\boldsymbol{\xi}_{\rho_n}^n) < v^2(\boldsymbol{\xi}_{\rho_n}^n)\}} v^2)(\boldsymbol{\xi}_{\rho_{n+1}}^{n+1}) | \mathcal{G}_{\rho_n}] \\
&\le \mathbb{E}[v(\boldsymbol{\xi}_{\rho_{n+1}}^{n+1}) | \mathcal{G}_{\rho_n}],
\end{aligned}$$

and by iterating the above we conclude that

$$(3.12) \qquad v(\xi) \le \mathbb{E}[v(\boldsymbol{\xi}_{\rho_{n+1}}^{n+1}) | \mathcal{G}_\tau] = \mathbb{E}[v(\boldsymbol{\xi}_{\rho_{n+1}}^{\mathbf{w}, \tau, \boldsymbol{\xi}}) | \mathcal{G}_\tau]$$

for any $n \ge 0$. Now we apply the same reasoning as in the proof of Lemma 2.3 in [4] to conclude that

$$\lim_{n \to \infty} \gamma_n = \sigma(\tau, \boldsymbol{\xi}, \mathbf{w}) \wedge \tau_*(v; \tau, \boldsymbol{\xi}, \mathbf{w}) \quad \text{a.s.}$$

By taking $n \to \infty$ in (3.12) and using the bounded convergence theorem we finally obtain that v satisfies (3.7) and, hence, is a stochastic subsolution. $\qquad \square$

We introduce the following assumption.

Assumption 3.5. The boundary function g is continuous on $\partial \Delta^{k+1}$.

PROPOSITION 3.6. *Under Assumption* 3.5 *the lower stochastic envelope* $v^- :=$ $\sup_{v \in \mathcal{V}^-} v \le V_\alpha$ *is a viscosity supersolution and the upper stochastic envelope* $v^+ :=$ $\inf_{v \in \mathcal{V}^+} v \ge V_\alpha$ *is a viscosity subsolution of* (3.2) *and* (3.3).

Proof. The proof uses ideas from Theorem 3.1 (and Theorem 4.1) in [3] and Theorem 2 in [14]. We repeat the key steps for the lower stochastic envelope v^-.

Denote for short $V \equiv V_\alpha$. It is clear that $v^- \le V$ since in item (ii) of Definition 3.2 we can choose $\tau = 0$, a constant $\boldsymbol{\xi} \in \Delta^{k+1}$, and $\rho = \sigma(\tau, \boldsymbol{\xi}, \mathbf{w})$ for some control $\mathbf{w} \in \text{int}(\mathcal{A}^\alpha)$, and use the conditions (3.6) and (3.9).

We will prove the viscosity supersolution property of v^- by contradiction. Take a C^2 test function $\varphi : \Delta^{k+1} \to \mathbb{R}$ such that $v^- - \varphi$ achieves a strict local minimum

equal to 0 at some boundary point $\xi_0 \in \partial\Delta^{k+1}$ (the case when $\xi_0 \in \text{int}(\Delta^{k+1})$ is simpler). Assume that v^- is not a viscosity supersolution and hence

$$\max\left\{\left(-\sup_{\mathbf{w}\in\mathbb{D}_c^{k+1}} L^{\mathbf{w}}\varphi\right)(\xi_0), (\varphi - g)(\xi_0)\right\} < 0,$$

where

$$(L^{\mathbf{w}}\varphi)(\xi) := \frac{1}{2}\,\text{tr}(\mathbf{w}\mathbf{w}'D_{\boldsymbol{\xi}}^2\varphi(\xi)).$$

It follows that there exists $\tilde{\mathbf{w}} \in \mathbb{D}_c^{k+1}$ such that

(3.13) $(-L^{\tilde{\mathbf{w}}}\varphi)(\xi_0) < 0.$

By the continuity of φ, g and the lower semicontinuity of v^- we can find a small enough open ball $B(\xi_0, \varepsilon)$ and a small enough $\delta > 0$ such that

$$(-L^{\tilde{\mathbf{w}}}\varphi)(\xi) < 0, \quad \xi \in B(\xi_0, \varepsilon) \cap \Delta^{k+1},$$
$$\varphi < g \quad \text{on} \quad B(\xi_0, \varepsilon) \cap \partial\Delta^{k+1},$$
$$\varphi(\xi) < v^-(\xi), \quad \xi \in B(\xi_0, \varepsilon) \cap \Delta^{k+1} \setminus \{\xi_0\},$$
$$v^- - \delta \geq \varphi \quad \text{on} \quad (\overline{B(\xi_0, \varepsilon)} \setminus B(\xi_0, \varepsilon/2)) \cap \Delta^{k+1}.$$

Using Proposition 4.1 in [2] together with Lemma 3.4 above, we obtain an increasing sequence of stochastic subsolutions $v_n \in \mathcal{V}^-$ with $v_n \nearrow v^-$. In particular, since φ and the v_n's are continuous we can use an argument identical to the one in Lemma 2.4 in [4] to obtain for any fixed $\delta' \in (0, \delta)$ a corresponding $v = v_n \in \mathcal{V}^-$ such that

$$v - \delta' \geq \varphi \quad \text{on} \quad (\overline{B(\xi_0, \varepsilon)} \setminus B(\xi_0, \varepsilon/2)) \cap \Delta^{k+1}.$$

Now we can choose $\eta \in (0, \delta')$ small enough such that $\varphi^\eta := \varphi + \eta$ satisfies

$$(-L^{\tilde{\mathbf{w}}}\varphi^\eta)(\xi) < 0, \quad \xi \in B(\xi_0, \varepsilon) \cap \Delta^{k+1},$$
$$\varphi^\eta < g \quad \text{on} \quad B(\xi_0, \varepsilon) \cap \partial\Delta^{k+1},$$
$$\varphi^\eta < v \quad \text{on} \quad (\overline{B(\xi_0, \varepsilon)} \setminus B(\xi_0, \varepsilon/2)) \cap \Delta^{k+1}.$$

We define

$$v^\eta = \begin{cases} v \vee \varphi^\eta & \text{on } \overline{B(\xi_0, \varepsilon)} \cap \Delta^{k+1}, \\ v & \text{otherwise} \end{cases}$$

and notice that v^η is continuous and $v^\eta(\xi_0) = v^-(\xi_0) + \eta > v^-(\xi_0)$. Since condition (3.6) clearly also holds, we see that v^η satisfies item (i) of Definition 3.2. What is left is to check item (ii) in Definition 3.2 and obtain $v^\eta \in \mathcal{V}^-$ which will lead to a contradiction since $v^\eta(\xi_0) > v^-(\xi_0)$.

Choose $\tau \in \mathcal{T}$ and $\boldsymbol{\xi} \in \mathcal{G}_\tau$ with $\mathbb{P}(\boldsymbol{\xi} \in \Delta^{k+1}) = 1$, and, similarly to the proof of Lemma 3.4 above, introduce the sequence of stopping time, control, and state process triples $(\gamma_n, \mathbf{w}^n, \boldsymbol{\xi}^n)_{n \geq -1}$ defined recursively as follows:

$$(\gamma_{-1}, \mathbf{w}^{-1}, \boldsymbol{\xi}^{-1}) \equiv (\gamma_0, \mathbf{w}^0, \boldsymbol{\xi}^0) := (\tau, \tilde{\mathbf{w}}\mathbf{1}_A + \bar{\mathbf{w}}^0\mathbf{1}_{A^c}, \boldsymbol{\xi}^{\mathbf{w}^0, \tau, \boldsymbol{\xi}}),$$

where $\bar{\mathbf{w}}^0$ is the control corresponding to the stochastic subsolution v starting at the pair $(\tau, \boldsymbol{\xi})$, the event A is given by

$$A = A(\boldsymbol{\xi}) := \{\boldsymbol{\xi} \in B(\xi_0, \varepsilon/2) \cap \Delta^{k+1} \text{ and } \varphi^\eta(\boldsymbol{\xi}) > v(\boldsymbol{\xi})\},$$

and for $n = 0, 1, 2, \ldots$, the following hold:

(i) If $v^\eta(\boldsymbol{\xi}_{\gamma_n}^n) \leq \bar{f}(\boldsymbol{\xi}_{\gamma_n}^n)$, then we set

$$(\gamma_{n+1}, \mathbf{w}^{n+1}, \boldsymbol{\xi}^{n+1}) := (\gamma_n, \mathbf{w}^n, \boldsymbol{\xi}^n).$$

(ii) If $A(\boldsymbol{\xi}_{\gamma_n}^n)$ holds, then we set

$$\gamma_{n+1} := \sigma(\gamma_n, \boldsymbol{\xi}_{\gamma_n}^n, \mathbf{w}^n) \wedge \tau_1(\gamma_n, \boldsymbol{\xi}_{\gamma_n}^n, \mathbf{w}^n) \wedge \tau_*(\varphi^\eta; \gamma_n, \boldsymbol{\xi}_{\gamma_n}^n, \mathbf{w}^n),$$
$$\mathbf{w}^{n+1} := \tilde{\mathbf{w}}, \quad \boldsymbol{\xi}^{n+1} := \boldsymbol{\xi}^{\tilde{\mathbf{w}}, \gamma_{n+1}, \boldsymbol{\xi}_{\gamma_{n+1}}^n},$$

where the $(\mathcal{G}_t)$-stopping time τ_1 is defined by

$$\tau_1(\tau, \boldsymbol{\xi}, \mathbf{w}) := \inf\{s \geq \tau : \boldsymbol{\xi}_s^{\mathbf{w}, \tau, \boldsymbol{\xi}} \in \partial B(\xi_0, \varepsilon/2)\},$$

and τ_* is defined as in (3.9).

(iii) Otherwise we set

$$\gamma_{n+1} := \sigma(\gamma_n, \boldsymbol{\xi}_{\gamma_n}^n, \mathbf{w}^n) \wedge \tau_*(v; \gamma_n, \boldsymbol{\xi}_{\gamma_n}^n, \mathbf{w}^n),$$
$$\boldsymbol{\xi}^{n+1} := \boldsymbol{\xi}^{\mathbf{w}^{n+1}, \gamma_{n+1}, \boldsymbol{\xi}_{\gamma_{n+1}}^n},$$

where $\mathbf{w}^{n+1}$ is the control process corresponding to the stochastic subsolution v starting at the pair $(\gamma_{n+1}, \boldsymbol{\xi}_{\gamma_{n+1}}^n)$.

By construction we have that $\gamma_n \leq \tau_*(v^\eta; \tau, \boldsymbol{\xi}, \mathbf{w})$ where the control $\mathbf{w} \in \text{int}(\mathcal{A}_c^\alpha)$ is defined as

$$\mathbf{w}_s := \sum_{n=1}^{\infty} \mathbf{1}_{\{s \in [\gamma_n, \gamma_{n+1})\}} \mathbf{w}_s^n.$$

Introduce the event

$$B := \{\gamma_n < \tau_*(v^\eta; \tau, \boldsymbol{\xi}, \mathbf{w}) \wedge \sigma(\tau, \boldsymbol{\xi}, \mathbf{w}) \quad \text{for all } n \in \mathbb{N}\},$$

and notice that for each $\omega \in B$ there exists $n_0(\omega)$ such that

$$\varphi^\eta(\boldsymbol{\xi}_{\gamma_{n_0+2l+1}}^{n_0+2l+1}) \leq \bar{f}(\boldsymbol{\xi}_{\gamma_{n_0+2l+1}}^{n_0+2l+1}) \tag{3.14}$$
$$\text{if } \tau_*(\varphi^\eta; \gamma_{n_0+2l}, \boldsymbol{\xi}_{\gamma_{n_0+2l}}^{n_0+2l}, \mathbf{w}^{n_0+2l}) < \tau_1(\gamma_{n_0+2l}, \boldsymbol{\xi}_{\gamma_{n_0+2l}}^{n_0+2l}, \mathbf{w}^{n_0+2l}),$$

$$v^\eta(\boldsymbol{\xi}_{\gamma_{n_0+2l+1}}^{n_0+2l+1}) = v(\boldsymbol{\xi}_{\gamma_{n_0+2l+1}}^{n_0+2l+1}) \tag{3.15}$$
$$\text{if } \tau_*(\varphi^\eta; \gamma_{n_0+2l}, \boldsymbol{\xi}_{\gamma_{n_0+2l}}^{n_0+2l}, \mathbf{w}^{n_0+2l}) \geq \tau_1(\gamma_{n_0+2l}, \boldsymbol{\xi}_{\gamma_{n_0+2l}}^{n_0+2l}, \mathbf{w}^{n_0+2l}),$$

$$v(\boldsymbol{\xi}_{\gamma_{n_0+2l+1}}^{n_0+2l+1}) \leq \bar{f}(\boldsymbol{\xi}_{\gamma_{n_0+2l+1}}^{n_0+2l+1}) \tag{3.16}$$

for $l \geq 0$. Denoting $\gamma_\infty := \lim_n \gamma_n$ and noticing that $\boldsymbol{\xi}_s^{\mathbf{w}, \tau, \boldsymbol{\xi}} = \boldsymbol{\xi}_s^n$ for $s \in [\gamma_n, \gamma_{n+1})$ we take the limit in (3.16) to obtain

$$v(\boldsymbol{\xi}_{\gamma_\infty}^{\mathbf{w}, \tau, \boldsymbol{\xi}}) \leq \bar{f}(\boldsymbol{\xi}_{\gamma_\infty}^{\mathbf{w}, \tau, \boldsymbol{\xi}}). \tag{3.17}$$

Now assume there exists $C \subseteq B$ such that for each $\omega \in C$ we have

$$\varphi^\eta(\boldsymbol{\xi}^{\mathbf{w},\tau,\boldsymbol{\xi}}_{\gamma_\infty}) > \bar{f}(\boldsymbol{\xi}^{\mathbf{w},\tau,\boldsymbol{\xi}}_{\gamma_\infty}),$$

and conclude from (3.14)–(3.15) that there exists large enough positive integer $M(\omega)$ such that for all $n \geq M$ we have

$$v^\eta(\boldsymbol{\xi}^n_{\gamma_n}) = v(\boldsymbol{\xi}^n_{\gamma_n}).$$

By taking $n \to \infty$ above we get $v^\eta(\boldsymbol{\xi}^{\mathbf{w},\tau,\boldsymbol{\xi}}_{\gamma_\infty}) = v(\boldsymbol{\xi}^{\mathbf{w},\tau,\boldsymbol{\xi}}_{\gamma_\infty})$ on C. Hence, by using (3.17) we see that

$$v^\eta(\boldsymbol{\xi}^{\mathbf{w},\tau,\boldsymbol{\xi}}_{\gamma_\infty}) \leq \bar{f}(\boldsymbol{\xi}^{\mathbf{w},\tau,\boldsymbol{\xi}}_{\gamma_\infty})$$

on C. On the other hand, on $B \setminus C$ we have

$$\varphi^\eta(\boldsymbol{\xi}^{\mathbf{w},\tau,\boldsymbol{\xi}}_{\gamma_\infty}) \leq \bar{f}(\boldsymbol{\xi}^{\mathbf{w},\tau,\boldsymbol{\xi}}_{\gamma_\infty})$$

and again from (3.17) we get

$$v^\eta(\boldsymbol{\xi}^{\mathbf{w},\tau,\boldsymbol{\xi}}_{\gamma_\infty}) \leq \bar{f}(\boldsymbol{\xi}^{\mathbf{w},\tau,\boldsymbol{\xi}}_{\gamma_\infty})$$

on $B \setminus C$. It follows that $\gamma_\infty \geq \tau_*(v^\eta; \tau, \boldsymbol{\xi}, \mathbf{w})$ on B, and from the definition of B we conclude that $\gamma_\infty = \tau_*(v^\eta; \tau, \boldsymbol{\xi}, \mathbf{w}) \wedge \sigma(\tau, \boldsymbol{\xi}, \mathbf{w})$.

Now take any $\rho \in \mathcal{T}$ with $\rho \in [\tau, \sigma(\tau, \boldsymbol{\xi}, \mathbf{w})]$, let $\rho \wedge \gamma_n = \rho_n$, and notice that, by Itô's formula applied to φ^η and the subsolution property of v, we have

$$\begin{aligned}
v^\eta(\boldsymbol{\xi}^n_{\rho_n}) &= (\mathbf{1}_A \varphi^\eta + \mathbf{1}_{A^c} v)(\boldsymbol{\xi}^n_{\rho_n}) \\
&\leq \mathbb{E}[(\mathbf{1}_{A(\boldsymbol{\xi}^n_{\rho_n})} \varphi^\eta + \mathbf{1}_{A(\boldsymbol{\xi}^n_{\rho_n})^c} v)(\boldsymbol{\xi}^{n+1}_{\rho_{n+1}}) | \mathcal{G}_{\rho_n}] \leq \mathbb{E}[v^\eta(\boldsymbol{\xi}^{n+1}_{\rho_{n+1}}) | \mathcal{G}_{\rho_n}],
\end{aligned}$$

and by iterating the above we conclude that

$$(3.18) \qquad v(\boldsymbol{\xi}) \leq \mathbb{E}[v(\boldsymbol{\xi}^{n+1}_{\rho_{n+1}}) | \mathcal{G}_\tau] = \mathbb{E}[\boldsymbol{\xi}^{\mathbf{w},\tau,\boldsymbol{\xi}}_{\rho_{n+1}}) | \mathcal{G}_\tau].$$

By taking $n \to \infty$ in (3.18) and using the bounded convergence theorem, we obtain that v^η satisfies item (ii) in Definition 3.2. Hence $v^\eta \in \mathcal{V}^-$ and we obtain a contradiction and, consequently, the supersolution property of v^-. $\qquad \square$

Assumption 3.7. The boundary function g is the concave envelope of $\bar{f}$ on the simplex faces $\{\mathbf{z} \in \Delta^{k+1} : z_j = 0\}$ for all $j = 0, \ldots, k+1$.

PROPOSITION 3.8. *Under Assumption* 3.7 *we have that* $v^- = v^+ = g$ *on* $\partial\Delta^{k+1}$.

Proof. Let $\bar{v}$ be the concave envelope of $\bar{f}$ on the whole of Δ^{k+1}. From Assumption 3.7 it follows that $\bar{v} = g$ on $\partial\Delta^{k+1}$ and $\bar{v}$ satisfies item (i) of Definition 3.3. Now take any $\tau \in \mathcal{T}$, $\boldsymbol{\xi} \in \mathcal{G}_\tau$ with $\mathbb{P}(\boldsymbol{\xi} \in \Delta^{k+1}) = 1$, $\mathbf{w} \in \text{int}(\mathcal{A}^\alpha_c)$, and $\rho \in \mathcal{T}$ with $\rho \in [\tau, \sigma(\tau, \boldsymbol{\xi}, \mathbf{w})]$, and notice that, by the Itô–Tanaka formula (see, e.g., Theorem VI.1.5 in [13]) applied to the concave function $\bar{v}$ we have

$$\mathbb{E}[\bar{v}(\boldsymbol{\xi}^{\mathbf{w},\tau,\boldsymbol{\xi}}_\rho) | \mathcal{G}_\tau] = \mathbb{E}\left[\bar{v}(\boldsymbol{\xi}) + \int_\tau^\rho \bar{v}'(\boldsymbol{\xi}^{\mathbf{w},\tau,\boldsymbol{\xi}}_s) d\boldsymbol{\xi}^{\mathbf{w},\tau,\boldsymbol{\xi}}_s + \int_{\Delta^{k+1}} L^a_\rho \bar{v}''(da) | \mathcal{G}_\tau\right] \leq \bar{v}(\boldsymbol{\xi}),$$

where $\bar{v}'$ is the left derivative, the second derivative $\bar{v}''$ is understood in the sense of a negative measure, and L^a is the local time at a of the process $\boldsymbol{\xi}^{\mathbf{w},\tau,\boldsymbol{\xi}}$. Hence, item (ii)

of Definition 3.3 is also satisfied and $\overline{v}$ is a stochastic supersolution. Since v^+ satisfies (3.10) and $v^+ \leq \overline{v}$ it follows that $v^+ = g$ on $\partial\Delta^{k+1}$.

Fix a constant control $\mathbf{w} \in \text{int}(\mathcal{A}_c^\alpha)$ and define the function $\underline{v} : \Delta^{k+1} \to \mathbb{R}$ by

$$(3.19) \qquad \underline{v}(\xi^\alpha) = \sup_{\overline{\tau} \in \mathcal{T}} \mathbb{E}\left[\tilde{V}_{k-1}(\xi_\sigma^{\mathbf{w},\xi^\alpha}) I_{\{\sigma \leq \overline{\tau}\}} + \bar{f}(\boldsymbol{\xi}_{\overline{\tau}}^{\mathbf{w},\xi^\alpha}) I_{\{\sigma > \overline{\tau}\}} \right].$$

The continuity of $\underline{v}(\xi^\alpha)$ follows from the boundedness of the control $\mathbf{w}$ and standard results on optimal stopping problems (see, e.g., Theorem 3.1.5 in [9]). We have that $\underline{v}(\xi^\alpha) = V_{\alpha'}(\xi^{\alpha'}) = g(\xi^\alpha)$ for $\xi^\alpha \in \partial\Delta^{k+1}$, and we obtain that item (i) of Definition 3.2 is satisfied. Moreover, the optimal stopping time in (3.19) exists and is equal to $\tau^* = \sigma \wedge \tau_*(\underline{v}; 0, \xi^\alpha, \mathbf{w})$, and it follows that $\underline{v}(\boldsymbol{\xi}_{t \wedge \tau^*}^{\mathbf{w},\xi^\alpha})$ is a martingale (see, e.g., Theorems I.2.4 and I.2.7 in [12]). This means that (3.7) is satisfied with equality and $\underline{v}$ is a stochastic subsolution. By definition we know that $v^- \leq g$ on $\partial\Delta^{k+1}$ and $\underline{v} \leq v^-$. Hence, we conclude that $v^- = g$ on $\partial\Delta^{k+1}$. $\qquad\square$

Proof of Theorem 3.1. It is clear that if $|\alpha| = 1$, then $V_\alpha(\xi^\alpha) = \bar{f}(\xi^\alpha)$, where $\xi = \delta_{x_i}$ for some i and $\xi^\alpha = 1$. We continue by induction and assume that we have proven the statement for all $k' < k$. By the induction hypothesis, $V_{\alpha'}(\xi^{\alpha'})$ is the concave envelope of $\bar{f}$ on the corresponding to α' simplex face and hence Assumption 3.7 is satisfied. Moreover, value functions coincide on the intersection of their corresponding simplex faces, and therefore, Assumpton 3.5 is also satisfied. Define the Hamiltonian H as

$$H(A) := - \sup_{\mathbf{w} \in \mathbb{D}_c^{k+1}} \frac{1}{2} \text{tr}(\mathbf{w}\mathbf{w}'A) \quad \text{for} \quad A \in \mathbb{R}^{(k+1)\times(k+1)},$$

and notice that for small enough c the set $\mathbb{D}_c^{k+1}$ contains all directions in $\mathbb{R}^k$. On the other hand, V_α is a viscosity solution to (3.2) on $\text{int}(\Delta^{k+1})$ if and only if the projected function $\tilde{V}_\alpha$ defined in (3.4) is a viscosity solution of

$$(3.20) \qquad \min\left\{ - \sup_{\mathbf{w} \in \tilde{\mathbb{D}}_c^k} \frac{1}{2} \text{tr}(\mathbf{w}\mathbf{w}' D_{\boldsymbol{\xi}}^2 \tilde{V}_\alpha), \tilde{V}_\alpha - \tilde{f} \right\} = 0$$

on $\text{int}(\tilde{\Delta}^k)$, where $\tilde{\mathbb{D}}_c^k$ is the projection of $\mathbb{D}_c^{k+1}$ onto $\mathbb{R}^k$. Hence, the function V_α is a viscosity solution to $H(D_{\boldsymbol{\xi}}^2 V_\alpha) \geq 0$ if and only if $\tilde{V}_\alpha$ is a viscosity solution to $-\lambda_k[\tilde{V}_\alpha] \geq 0$, where $\lambda_k[\tilde{V}_\alpha]$ is the largest eigenvalue of the Hessian $D_{\boldsymbol{\xi}}^2 \tilde{V}_\alpha$. Therefore, we can apply Theorem 1 in [10] to obtain that any continuous viscosity solution to (3.20) is concave. Moreover, uniqueness of the solution to (3.20), together with the projected boundary condition

$$(3.21) \qquad \tilde{V}_\alpha(\xi^\alpha) = \tilde{V}_{\alpha'}(\xi^{\alpha'}),$$

follows from the comparison principle for Dirichlet problems stated in Theorem 2.10 of [11]. This leads to uniqueness and the comparison principle for our original problem (3.2)–(3.3). In particular, by Propositions 3.6 and 3.8 we have that $v^+ \leq v^-$ on $\text{int}(\Delta^{k+1})$. On the other hand, by Proposition 3.6 we also have $v^- \leq V_\alpha \leq v^+$ on Δ^{k+1}. Therefore, we can conclude that $v^- = V_\alpha = v^+$ on Δ^{k+1} and that V_α is the unique viscosity solution of (3.2) with the boundary condition (3.3), and the same is true for the projected versions.

Finally, from Theorem 2 in [10] we have that the concave envelope of the projected cost function $\tilde{f}$ solves (3.20), and since it also clearly satisfies (3.21), we conclude from the uniqueness that $\tilde{V}_\alpha$ is the concave envelope of $\tilde{f}$. $\qquad\square$

Remark 3.9. The value function V_α can be regarded as the concave envelope on the simplex Δ^{k+1} of the modified cost function $\bar{f}$. Indeed, we can ignore one direction in the state space vector $\boldsymbol{\xi}$ due to the fact that Δ^{k+1} is a k-dimensional simplex and any concave function on a k-dimensional simplex in $\mathbb{R}^{k+1}$ is concave in any k of its variables (and vice versa). Note that the optimal control weight vector $\mathbf{w}^*$ may not be unique. It is determined by the direction on the simplex Δ^{k+1} for which the second directional derivative of the value function V_α is zero—if the value function is linear at a point, then clearly many directions satisfy this condition.

Remark 3.10. When applying the stochastic Perron's method to controlled exit-time problems, one needs a comparison result for the corresponding PDE in order to characterize the value function as a viscosity solution (see, e.g., Definition 2 and Remark 1 in [14]). These comparison results are of a slightly different nature than the standard ones of, e.g., Theorems 7.9 and 8.2 in [7]—the latter requires an a priori knowledge of the behavior of the stochastic semisolutions at the boundary. We were able to exploit the specific structure of our exit-time problem in Proposition 3.8 to obtain the behavior at the boundary of the stochastic semisolutions. This allowed the application of the comparison result in [11].

4. Examples. Let us first provide some intuition behind the choice of optimal controls and stopping times. We will consider a general class of cost functions—namely all bounded, nonnegative Lipschitz continuous functions $f : \mathbb{R} \to \mathbb{R}$. This is the class for which Theorem 3.1 holds. We will use our concave envelope characterization to choose the optimal controls and verify that Brownian exit times are optimal.

We abuse notation and regard $\bar{f}$ as a function on the projected set of probability vectors $\tilde{\Delta}^N := \{\mathbf{z} \in \mathbb{R}_{\geq 0}^N : \sum z_i \leq 1\}$. Denote by $\mathrm{conc}(\bar{f})$ the concave envelope of $\bar{f}$ on $\tilde{\Delta}^N$. For any initial probability vector $z \in \tilde{\Delta}^N$ corresponding to some terminal law μ, e.g.,

$$\mu = \sum_{i=1}^N z_i \delta_{x_i} + \left(1 - \sum_{i=1}^N z_i\right)\delta_{x_0},$$

we will find a candidate optimal control weight process $(\mathbf{w}_r)_{r \geq 0}$ taking values in the projected admissible set $\hat{\mathbb{D}}_c^N$ (i.e., the projection of $\mathbb{D}_c^{N+1}$ onto $\mathbb{R}^N$) and a candidate optimal stopping time τ_* such that the resulting value function will be $\mathrm{conc}(\bar{f})$.

The usual characterization of optimal stopping times leads us to choose the candidate τ_* as

(4.1) $$\tau_* := \inf\{r \geq 0 : \mathrm{conc}(\bar{f})(\boldsymbol{\xi}_r^{\mathbf{w},z}) = \bar{f}(\boldsymbol{\xi}_r^{\mathbf{w},z})\}.$$

In particular, if the initial probability vector z is such that $\mathrm{conc}(\bar{f})(z) = \bar{f}(z)$, we can simply set $\tau_* = 0$. Assume now that $\mathrm{conc}(\bar{f})(z) > \bar{f}(z)$ and note that the point $(z, \mathrm{conc}(\bar{f})(z))$ belongs to a planar region of the graph of $\mathrm{conc}(\bar{f})(z)$ that contains a point $(z^{(1)}, \mathrm{conc}(\bar{f})(z^{(1)}))$ such that $\mathrm{conc}(\bar{f})(z^{(1)}) = \bar{f}(z^{(1)})$. In other words, all points on the line between $(z, \mathrm{conc}(\bar{f})(z))$ and $(z^{(1)}, \mathrm{conc}(\bar{f})(z^{(1)}))$ are also part of the graph of $\mathrm{conc}(\bar{f})$. We choose the control weight process as a constant vector in the direction of $z - z^{(1)}$, i.e., $\mathbf{w}_r \equiv c_1(z - z^{(1)})$, where the constant c_1 is such that $\mathbf{w}$ is admissible. Therefore, the probability vector process $(\boldsymbol{\xi}_r^{\mathbf{w},z})_{r \geq 0}$ evolves along the direction $z - z^{(1)}$ and either hits the point $z^{(1)}$ or hits the boundary of $\tilde{\Delta}^N$ at some point $z^{(2)}$. The point $z^{(2)}$ can be regarded as belonging to a lower dimensional projected set $\tilde{\Delta}^{\bar{N}} := \{\mathbf{z} \in \mathbb{R}_{\geq 0}^{\bar{N}} : \sum z_i \leq 1\}$ where $\bar{N} < N$. If $\mathrm{conc}(\bar{f})(z^{(2)}) > \bar{f}(z^{(2)})$,

we repeat the same procedure when choosing a control on this lower dimensional set—clearly this can happen at most N times.

For simplicity's sake assume that $\mathrm{conc}(\bar{f})(z^{(2)}) = \bar{f}(z^{(2)})$. In other words, by looking at (2.4) and (4.1), we get that τ_* is the first exit time of a Brownian motion from the interval with endpoints $v_1 = \frac{z_0^{(1)} - z_0}{c_1(z_0 - z_0')}$ and $v_2 = \frac{z_0^{(2)} - z_0}{c_1(z_0 - z_0')}$. Using the formula for the Brownian exit times from an interval, we obtain that the projected value function as defined in (3.4) satisfies

$$\tilde{V}_\alpha(z) = \frac{v_2}{v_2 - v_1}\bar{f}(z^{(1)}) + \frac{-v_1}{v_2 - v_1}\bar{f}(z^{(2)}),$$

and the point $(z, \tilde{V}_\alpha(z))$ lies on the line going through $(z, \mathrm{conc}(\bar{f})(z))$ and $(z', \mathrm{conc}(\bar{f})(z'))$; hence $\tilde{V}_\alpha(z) = \mathrm{conc}(\bar{f})(z)$. Similar calculation is valid for the case $\mathrm{conc}(\bar{f})(z^{(2)}) > \bar{f}(z^{(2)})$.

Finally, by application of the Itô–Tanaka formula as in the proof of Proposition 3.8, we conclude that $\mathrm{conc}(\bar{f})$ bounds the value function from above, and therefore, the two coincide.

Remark 4.1 (generalized put options). In fact, if the cost function is of the form

$$f(s) = (g(s))^+,$$

for some concave function g, by direct calculation we can check that the candidate control and stopping time described above are optimal among those controls that follow a fixed direction and those stopping times that are Brownian exit times from an interval. By applying Theorem 3.1, we see that optimization over this class is sufficient.

In what follows, using the observations above, we will construct the optimal controls and stopping times explicitly for a piecewise linear cost function which can be thought of as a call option spread.

4.1. Call option spread. We let f take the form

$$f(s) = (s - K_1)^+ - (s - K_2)^+$$

for $K_1 \in (-1, 1)$, $K_2 \in (0, 1)$, and $K_1 < K_2$, which can be seen as a bull call spread. Set $N = 2$, $\mathbb{X}_N = \{-1, 0, 1\}$, and assume that the law of M_T is given by

$$\mu = (1 - \gamma - \beta)\delta_{-1} + \beta\delta_0 + \gamma\delta_1$$

for $0 < \gamma, \beta < 1$ such that $0 < \gamma + \beta < 1$. Therefore, the initial probability vector is

$$\xi^\alpha \equiv (\xi_0^0, \xi_0^1, \xi_0^2) = (1 - \gamma - \beta, \beta, \gamma) \in \mathrm{int}(\Delta^3),$$

where $\alpha = \{0, 1, 2\}$. From the definition of the process M in (2.1) it follows that

(4.2) $\quad M_t = \gamma_{T_t^{-1}} - (1 - \gamma_{T_t^{-1}} - \beta_{T_t^{-1}}) = 2\gamma_{T_t^{-1}} + \beta_{T_t^{-1}} - 1 \quad \text{for} \quad t \in [0, T],$

where $\beta_r = \xi_r^1$ and $\gamma_r = \xi_r^2$ for $r \geq 0$. We introduce the constants $s^{-101} = 2\gamma + \beta - 1$, $s^{01} = \frac{\gamma}{\gamma+\beta}$, $s^1 = 1$, and $s^0 = 0$ corresponding to the value of M_0 taking various atoms of $\mathbb{X}_N$ into account. We use the notation $V_\alpha(\beta, \gamma) := V_\alpha(\xi^\alpha)$ and $\bar{f}(\beta, \gamma) := \bar{f}(\xi^\alpha)$.

We will now describe how to obtain a guess for the value function which, as expected, will turn out to be the concave envelope of the modified cost function

$\bar{f}$. Notice that f is nondecreasing and achieves its maximum for any $s \geq K_2$ and its minimum for any $s \leq K_1$. Therefore, for the martingale state process $\boldsymbol{\xi}^{\mathbf{w},\xi^\alpha}$ (or, equivalently, the law process $\xi^{\mathbf{w},\xi^\alpha}$), we want to offset any decrease of probability mass on the interval (K_2, ∞) with a corresponding decrease on the interval $(-\infty, K_1)$. We consider the following cases:

1. Assume $M_0 \equiv s^{-101} \geq K_2$. Then it is optimal to stop immediately, i.e., choose an optimal stopping time $\tau_* = 0$ and obtain $V_\alpha(\beta, \gamma) = K_2 - K_1$.

2. Assume $s^{01} \geq K_2 > s^{-101}$, and let the constant $\eta \in [0, 1-\gamma-\beta)$ be such that $\frac{\gamma-\eta}{\gamma+\beta+\eta} = K_2$. Then it is optimal to choose a stopping time τ_* and a control process $\mathbf{w}_r \equiv (w_r^0, w_r^1, w_r^2) = (-c_1 - \frac{\beta}{\gamma}c_1, \frac{\beta}{\gamma}c_1, c_1)$ for any $r \in [0, \tau_*]$, where the constant $c_1 > 0$ is such that $\mathbf{w}$ is an admissible control and the optimal stopping time τ_* is the first exit time of γ_r from the interval $(0, \frac{\gamma}{\gamma+\beta+\eta})$. Note that this choice of $\mathbf{w}$ is not unique.

   Equivalently, by using (4.2), we see that τ_* is the first exit time of M_{T_r} from the interval $(-1, K_2)$. This corresponds to letting the law $\xi^{\mathbf{w},\xi^\alpha}$ evolve until the stopping time τ_* when it separates into two measures of the form

$$\xi_{\tau_*}^{\mathbf{w},\xi^\alpha} = \begin{cases} \frac{\gamma\delta_1 + \beta\delta_0 + \eta\delta_{-1}}{\gamma+\beta+\eta} & \text{with probability } \gamma + \beta + \eta, \\ \delta_{-1} & \text{with probability } 1 - (\gamma + \beta + \eta). \end{cases}$$

   By the definition of η we have that $\gamma+\beta+\eta = \frac{2\gamma+\beta}{K_2+1}$, and therefore, $V_\alpha(\beta, \gamma) = \frac{2\gamma+\beta}{K_2+1}(K_2 - K_1)$.

3. Assume $K_2 > s^{01}$, and let the constant $\eta \in (0, \beta)$ be such that $\frac{\gamma}{\gamma+\eta} = K_2$. Then we choose a stopping time R_1 and a control process $\mathbf{w}_r \equiv (w_r^0, w_r^1, w_r^2) = (-c_1 - \frac{\eta-\beta(\gamma+\eta)}{\gamma-\gamma(\gamma+\eta)}c_1, \frac{\eta-\beta(\gamma+\eta)}{\gamma-\gamma(\gamma+\eta)}c_1, c_1)$ for any $r \in [0, R_1]$, where the constant $c_1 > 0$ is such that $\mathbf{w}$ is an admissible control and the stopping time R_1 is the first exit time of γ_r from the interval $(0, \frac{\gamma}{\gamma+\eta})$. Equivalently, by using (4.2), we see that R_1 is the first exit time of M_{T_r} from the interval $\left(-\frac{1-\gamma-\beta}{1-\gamma-\eta}, K_2\right)$. This corresponds to letting the law $\xi^{\mathbf{w},\xi^\alpha}$ evolve until time R_1 when it separates into two measures of the form

$$\xi_{R_1}^{\mathbf{w},\xi^\alpha} = \begin{cases} \frac{\gamma\delta_1 + \eta\delta_0}{\gamma+\eta} & \text{with probability } \gamma + \eta, \\ \frac{(\beta-\eta)\delta_0 + (1-\beta-\gamma)\delta_{-1}}{1-\gamma-\eta} & \text{with probability } 1 - (\gamma + \eta). \end{cases}$$

   In addition, if $s^0 \leq K_1$, we choose the optimal stopping time as $\tau_* \equiv R_1$ and we have $V_\alpha(\beta, \gamma) = \frac{\gamma}{K_2}(K_2 - K_1)$. This is due to the fact that if $\gamma_{R_1} = 0$ (i.e., the atom $\{1\}$ dies), it is not worth evolving the law $\xi^{\mathbf{w},\xi^\alpha}$ further because the cost function f will be 0 under any combination of the atoms $\{0, -1\}$. In other words we gain nothing from transferring probability mass between the atoms 0 and -1.

   On the other hand, if we also have that $s^0 > K_1$, on the event $A := \{\gamma_{R_1} = 0\}$ we let the control process be $\mathbf{w}_r = (-w_{R_1}^1, w_{R_1}^1, 0)$ for $r \in (R_1, R_2]$ and set the optimal stopping time

$$\tau_* = R_1 \mathbf{1}_{A^c} + R_2 \mathbf{1}_A,$$

   where the stopping time R_2 is the first exit time of β_u from the interval $(0, 1)$ for $u > R_1$. Equivalently, by using (4.2), we see that R_2 is the first exit time of M_{T_r} from the interval $(-1, 0)$ for $r > R_1$. This corresponds to further

evolving the law $\xi^{\mathbf{w},\xi^\alpha}$ until at the stopping time $R_2 > R_1$ it splits into three measures of the form

$$\xi_{R_2}^{\mathbf{w},\xi^\alpha} = \begin{cases} \frac{\gamma\delta_1 + \eta\delta_0}{\gamma+\eta} & \text{with probability } \gamma + \eta, \\ \delta_0 & \text{with probability } \beta - \eta, \\ \delta_{-1} & \text{with probability } 1 - \beta - \gamma. \end{cases}$$

Therefore we have

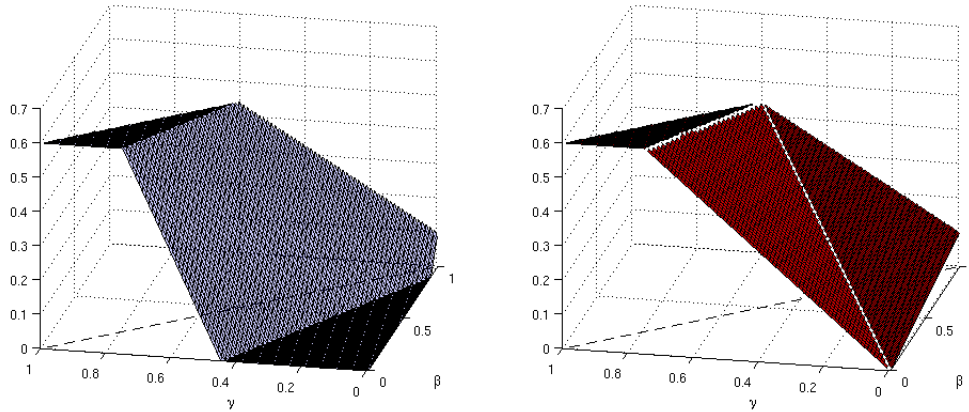$$V_\alpha(\beta,\gamma) = \frac{\gamma}{K_2}(K_2 - K_1) + (\beta - \eta)(-K_1) = \gamma(1 - K_1) - \beta K_1.$$

FIG. 1. *The modified cost function $\bar{f}(\beta,\gamma)$ on the left plotted together with the projected value function $V_\alpha(\beta,\gamma)$ on the right for $K_1 = -0.1$ and $K_2 = 0.5$. The three triangular planar regions correspond to the three cases above. It is evident that $V_\alpha(\beta,\gamma)$ is the concave envelope of $\bar{f}(\beta,\gamma)$.*

The candidate value function $V_\alpha(\beta,\gamma)$ is given by

$$V_\alpha(\beta,\gamma) = \begin{cases} K_2 - K_1 & \text{(i)} & s^{-101} \geq K_2, \\ \frac{2\gamma+\beta}{K_2+1}(K_2 - K_1) & \text{(ii)} & s^{01} \geq K_2 > s^{-101}, \\ \frac{\gamma}{K_2}(K_2 - K_1) & \text{(iii)} & K_2 > s^{01}, s^0 \leq K_1, \\ \gamma(1 - K_1) - \beta K_1 & \text{(iv)} & K_2 > s^{01}, s^0 > K_1, \end{cases}$$

and it is the concave envelope of $\bar{f}(\beta,\gamma)$ (see Figure 1).[1]

Appendix A. Proof of Lemma 2.1.

Proof. In order to prove the independence in the t variable we choose $0 \leq t_1 < t_2 < T$ and notice that $U(t_1,\xi) \geq U(t_2,\xi)$. Indeed, the supremum in (2.3) corresponding to $U(t_1,\xi)$ is taken over a larger set of stopping times than the one corresponding

---

[1]It turns out that the value function in this example is the same as in the Asian option setting of [6]; see the example in section 4.2 therein. This is because under their optimal model the stock price is a fixed random variable which is given by the average of our measure-valued martingale at τ^* using (2.1).

to $U(t_2, \xi)$. Conversely, for any $\xi \in \Xi$ and $\tau \in \mathcal{T}_{t_1}$ we can choose $\tilde{\xi} \in \Xi$ and $\tilde{\tau} \in \tilde{\mathcal{T}}_{t_2}$ such that

$$\tilde{\tau} = a\tau + b, \qquad \tilde{\xi}_{at+b} = \xi_t,$$

with $a = \frac{T-t_2}{T-t_1}$ and $b = \frac{T(t_2 - t_1)}{T-t_1}$. This choice leads to

$$\int x\, \xi_\tau(dx) = \int x\, \tilde{\xi}_{\tilde{\tau}}(dx),$$

which allows us to conclude that $U(t_2, \xi) \geq U(t_1, \xi)$ and hence $U(t_2, \xi) = U(t_1, \xi)$, and we have independence in t for $t \in [0, T)$.

To prove the continuity in ξ we first observe (e.g., see Lemma 3.1 in [6]) that if $(\xi_r)_{r \geq 0} \in \Xi$ with $\xi_t = \xi$ and $d_{\mathcal{W}_1}(\xi_t, \xi') < \varepsilon$ (here $d_{\mathcal{W}_1}$ is the Wasserstein-1 metric), then there is $(\xi'_r)_{r \geq 0} \in \Xi$ with $\xi'_t = \xi'$ such that $\mathbb{E}[|\int x\, \xi_\tau(dx) - \int x\, \xi'_\tau(dx)||\mathcal{F}_t] < \varepsilon$ for all $\tau \in \mathcal{T}_t$ with some fixed $\lambda \in \Lambda$. Indeed, we know that $\xi_s = \mathbb{E}[\xi_T | \mathcal{F}_s]$, and we can define

$$\xi'_s(dy) = \mathbb{E}\left[\int \xi_T(dx) m(x, dy) | \mathcal{F}_s\right], \quad s \geq t,$$

where the Borel family of probability measures $m(x, dy)$ is obtained by the disintegration of the transport plan $\Gamma(dx, dy) = \xi_t(dx) m(x, dy)$ such that $\Gamma(\mathbb{R}_+, dy) = \xi'(dy)$, $\Gamma(dx, \mathbb{R}_+) = \xi_t(dx)$, and $\int \int |x - y| \Gamma(dx, dy) < \varepsilon$. By optional stopping we get

$$\left|\int x\, \xi_\tau(dx) - \int x\, \xi'_\tau(dx)\right| \leq \mathbb{E}\left[\int \int |x - y| \xi_T(dx) m(x, dy) | \mathcal{F}_\tau\right],$$

and hence

$$\mathbb{E}\left[\left|\int x\, \xi_\tau(dx) - \int x\, \xi'_\tau(dx)\right| \Big| \mathcal{F}_t\right] \leq \int \int |x - y| \Gamma(dx, dy) < \varepsilon.$$

Denote by M^ξ the process corresponding to the measure-valued martingale $(\xi_r)_{r \geq 0}$ from (2.1). By the Lipschitz property of f and the above inequality we get

$$\mathbb{E}\left[\left|f(M_\tau^{\xi'}) - f(M_\tau^\xi)\right| \Big| \mathcal{F}_t\right] < \varepsilon.$$

Now fix $\varepsilon' > 0$ and consider $\xi, \xi' \in \mathcal{P}$ such that $d_{\mathcal{W}_1}(\xi, \xi') < \varepsilon'/2$. From the reasoning above, we can choose $(\xi_r)_{r \geq 0}, (\xi'_r)_{r \geq 0} \in \Xi$ with $\xi_t = \xi$ and $\xi'_t = \xi'$ such that $U(t, \xi) \leq \sup_{\tau \in \mathcal{T}_t} \mathbb{E}[f(M_\tau^\xi) | \mathcal{F}_t] + \varepsilon'/2$ and $\mathbb{E}\left[\left|f(M_\tau^{\xi'}) - f(M_\tau^\xi)\right| \Big| \mathcal{F}_t\right] < \varepsilon'/2$. Therefore, we obtain

$$U(t, \xi) \leq \sup_{\tau \in \mathcal{T}_t} \mathbb{E}[f(M_\tau^\xi) | \mathcal{F}_t] + \varepsilon'/2 \leq \sup_{\tau \in \mathcal{T}_t} \mathbb{E}[f(M_\tau^{\xi'}) | \mathcal{F}_t] + \varepsilon' \leq U(t, \xi') + \varepsilon',$$

and by symmetry we get $|U(t, \xi) - U(t, \xi')| \leq \varepsilon'$ and continuity follows.   □

REFERENCES

[1] E. Bayraktar and C. W. Miller, *Distribution-constrained optimal stopping*, Math. Finance, to appear, https://doi.org/10.1111/mafi.12171.

[2] E. Bayraktar and M. Sirbu, *Stochastic Perron's method and verification without smoothness using viscosity comparison: The linear case*, Proc. Amer. Math. Soc., 140 (2012), pp. 3645–3654.

[3] E. Bayraktar and M. Sirbu, *Stochastic Perron's method for Hamilton–Jacobi–Bellman equations*, SIAM J. Control Optim., 51 (2013), pp. 4274–4294, https://doi.org/10.1137/12090352X.

[4] E. Bayraktar and M. Sirbu, *Stochastic Perron's method and verification without smoothness using viscosity comparison: Obstacle problems and Dynkin games*, Proc. Amer. Math. Soc., 142 (2014), pp. 1399–1412.

[5] M. Beiglböck, M. Eder, C. Elgert, and U. Schmock, *Geometry of Distribution-Constrained Optimal Stopping Problems*, preprint, https://arxiv.org/abs/1612.01488, 2016.

[6] A. M. G. Cox and S. Källblad, *Model-independent bounds for Asian options: A dynamic programming approach*, SIAM J. Control Optim., 55 (2017), pp. 3409–3436, https://doi.org/10.1137/16M1087527.

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

[8] Y. Dolinsky and H. M. Soner, *Martingale optimal transport and robust hedging in continuous time*, Probab. Theory Related Fields, 160 (2014), pp. 391–427.

[9] N. Krylov, *Controlled Diffusion Processes*, Stoch. Model. Appl. Probab. 14, Springer-Verlag, Berlin, 2009.

[10] A. M. Oberman, *The convex envelope is the solution of a nonlinear obstacle problem*, Proc. Amer. Math. Soc., 135 (2007), pp. 1689–1694.

[11] A. M. Oberman and Y. Ruan, *A partial differential equation for the rank one convex envelope*, Arch. Ration. Mech. Anal., 224 (2017), pp. 955–984.

[12] G. Peskir and A. N. Shiryaev, *Optimal Stopping and Free-Boundary Problems*, Birkhäuser Verlag, Basel, 2006.

[13] D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion*, Springer-Verlag, Berlin, 1999.

[14] D. B. Rokhlin, *Verification by stochastic Perron's method in stochastic exit time control problems*, J. Math. Anal. Appl., 419 (2014), pp. 433–446.