Parallel Search for Information*

T. Tony Ke
Massachusetts Institute of Technology
kete@mit.edu

Wenpin Tang
University of California, Los Angeles
wenpintang@math.ucla.edu

J. Miguel Villas-Boas
University of California, Berkeley
villas@haas.berkeley.edu

Yuming Zhang
University of California, Los Angeles
yzhangpaul@math.ucla.edu

May 2019

*We thank Andrej Zlatos for the helpful discussions regarding the proof of Proposition 6. Very preliminary. Comments welcome.

Parallel Search for Information

Abstract

We consider an optimal stopping problem of a d-dimensional Brownian motion, where the payoff at stopping is the maximum component of the Brownian motion, and there is a running cost before stopping. Applications include choosing one among several alternatives while learning simultaneously about all the alternatives (parallel search), and exercising an option based on several assets. We present necessary and sufficient conditions for the solution, establishing existence and uniqueness. We show that the free boundary is star-shaped, and present asymptotic characterization of the value function and the free boundary. We also show properties of how the distance between the free boundary and the diagonal varies with the number of alternatives.

Keywords: *Optimal Stopping, Free Boundary Problem, Information, Search Theory, Brownian Motion*

1. Introduction

In several situations a decision-maker (DM) has to decide how long to gain information on several alternatives simultaneously at a cost before stopping to make an adoption decision. An important aspect considered here, is that the DM gains information on all alternatives at the same time and cannot choose which alternative to gain information on—which we call *parallel search*. This can be, for example, the case of a consumer trying to decide among several products in a product category and passively learning about the product category, or browsing through a web site that compares several products side by side. Another interesting application is a financial option based on several assets, where at the time of exercising the option, the investor decides which asset to take.

Let $B^x = (B_1^{x_1}(t), \dots, B_d^{x_d}(t))_{t\geq 0}$ be a d-dimensional Brownian motion starting at $x = (x_1, \dots, x_d)$. Each component of this Brownian motion could be the value of the alternative if the process is stopped. In the consumer learning application, this would be the expected value of that product at the time when the consumer makes the purchase decision. In the financial option application, this would be the value of the asset when the option is exercised. Let $\mathcal{T}$ be a suitable set of stopping times with respect to the natural filtration of B^x. We aim to determine the following value function,

$$u(x) := \sup_{\tau\in\mathcal{T}} \mathbb{E}\left[\max\left\{B_1^{x_1}(\tau), \dots, B_d^{x_d}(\tau), 0\right\} - c\tau\right], \tag{1}$$

where $c > 0$ is the cost per unit time.[1]

We first characterize the the value function u defined by (1). We give necessary conditions in Section 2: u is a viscosity solution to some partial differential equation (PDE) with at most linear growth. We then prove in Section 3 that the condition derived is also sufficient by establishing the existence and uniqueness of the solution to the PDE.

One important ingredient of the problem considered is that there is a free boundary where it is optimal to stop, and this boundary is determined by the solution to the PDE. In Section 4, we prove a geometric property of the free boundary: it is star-shaped with respect to the origin. In particular, how much is required from the best alternative in order to stop the process is increasing in the values of the other alternatives. We also compare this boundary with the boundary that results from the problem where alternatives can only be learnt sequentially—one alternative at each instance of time, and illustrate this comparison with numerical simulations. We consider also what happens if the DM can choose, at different costs, to gain information sequentially on one alternative at a time, or to gain information on all alternatives simultaneously.

[1] The problem could also be considered with time discounting. The case of the cost per unit of time could be seen as the costs of processing information when learning about different alternatives.

Although it is not possible to derive closed-form expressions for the value function or the free boundary, we can study the asymptotics of the value function as well as the free boundary as $x_1 = \ldots = x_d \to \infty$, which is presented in Section 5. We provide fine estimates of the distance from the free boundary to $\{x_1 = x_2\}$ at infinity for $d = 2$, while for general $d \geq 3$ we prove this distance is increasing in d, and is at most linear in d. To the best of our knowledge, this is one of the few results concerning the asymptotic geometry of the optimal stopping problem in dimension $d \geq 2$. See Peskir and Shiryaev (2006), Guo and Zervos (2010), Assing et al. (2014) for studies of optimal stopping problems for $d = 2$. The main difficulty in our problem is lack of closed-form expressions for the value function. Here we rely heavily on the PDE machinery.

There is some literature on gradual learning when a single alternative is considered or information is gathered to uncover a single uncertain value (e.g., Roberts and Weitzman 1981, Moscarini and Smith 2001, Branco et al. 2012, Fudenberg et al. 2018),[2] and the choice there is between adopting the alternative or not. In the face of more than one uncertain alternative (as is the case considered in this paper) the problem becomes more complicated. This is because opting for one alternative in a choice set means giving up potential high payoffs from other alternatives about which the decision maker has yet to learn more information. This paper can then be seen as extending this literature to allow for more than one alternative, which requires the solution to a partial differential equation. Another possibility, considered in Ke et al. (2016),[3] is that the DM can choose to search for information on one alternative at a time (with alternatives having independent values). That simplifies the analysis because in each region in which one alternative is searched, the value function satisfies an ordinary differential equation on the state of that alternative keeping the states of the other alternatives fixed. Here, the value function does not satisfy that property as the states of all alternatives move simultaneously. Consequently, the value function is determined by a partial differential equation (with free boundaries). We compare the solution in this case with the solution when the DM can choose to search for information on only one alternative at a time. We also consider what happens when the DM can choose to search for information on only one alternative, or search on all alternatives simultaneously at a higher cost, with economies of scale on the number of alternatives searched.

The literature on financial options based on multiple assets (rainbow options) is also related to this paper (see, for example, Stulz 1982, Johnson 1987, Rubinstein 1991, Broadie and Detemple

[2]The case with a single alternative can be traced back to the discrete costly sequential sampling in Wald (1945). The continuous time treatment of the single alternative case was also presented in Dvoretzky et al. (1953), Mikhalevich (1958), and Shiryaev (1967).

[3]Che and Mierendorff (2016) consider which type of information to collect in a Poisson-type model, when the decision maker has to choose between two alternatives, with one and only one alternative having a high payoff. See also Hébert and Woodford (2017) for a rational inattention formulation.

1997). In relation to that literature, we present a different specification related to consumer search for information and show existence of a unique solution, and compare the solution with the case in which the decision maker can choose to learn information about only one alternative at a time.

2. Analysis

We start with the general framework of the optimal stopping problem (1).[4] Let $\Omega \subset \mathbb{R}^d$ be a smooth domain. Consider the following stochastic differential equation (SDE):

$$dX(t) = b(X(t))dt + \sigma(X(t))dB(t), X(0) = x \in \Omega, \tag{2}$$

where $(B(t); t \geq 0)$ is a d-dimensional Brownian motion starting at $x \in \Omega$, $b : \mathbb{R}^n \to \mathbb{R}^n$ and $\sigma : \mathbb{R}^n \to \mathbb{R}^{n\times n}$ satisfy

- *Lipschitz condition*: there exists $C > 0$ such that

$$|b(x) - b(y)| + |\sigma(x) - \sigma(y)| \leq C|x - y|.$$

- *Linear growth condition*: there exists $K > 0$ such that

$$|b(x)| + |\sigma(x)| \leq K|x|.$$

It is well known that under these conditions, the SDE (2) has a strong solution which is pathwise unique. See, for example, Karatzas and Shreve (1991), Section 5.2, for background on strong solutions to SDEs. The vector $X(t)$ has as each element i the expected utility obtained if the DM were to decide to stop the search process at time t and choose alternative i.

Let

$$J_x(\tau) := \mathbb{E}\left[\int_0^\tau f(X(s))ds + g(X(\tau))\right], \tag{3}$$

where τ is a stopping time, and f, g are two smooth functions. We are interested in the value function

$$u(x) = \sup_{\tau\in\mathcal{T}} J_x(\tau), \tag{4}$$

[4]We present two examples of applications in the Appendix.

where $\mathcal{T}$ is a suitable set of stopping times. Let $\mathcal{L}$ be the infinitesimal generator of the SDE (2). That is,

$$\mathcal{L}h = \sum_{i=1}^{d} b_i \frac{\partial h}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^{d} (\sigma\sigma^T)_{ij} \frac{\partial^2 h}{\partial x_i \partial x_j},$$

for any suitably smooth test function $h : \mathbb{R}^n \to \mathbb{R}$.

A standard dynamic programming argument shows that u is a viscosity solution to the following partial differential equation (PDE):

$$\min(-\mathcal{L}u - f, u - g) = 0. \tag{5}$$

The notion of viscosity solution will be made precise later. Equation (5) is known as an *obstacle problem*, or a *variational inequality* (see Frehse 1972, Kinderlehrer and Stampacchia 1980). It exhibits two regimes:

- $-\mathcal{L}u = f$ when $u > g$,
- $-\mathcal{L}u \geq f$ when $u = g$,

which are separated by *free boundaries* $\Gamma(u)$. The set $\{u = g\}$ is called the *contact set*, or *coincidence set*. In general, a solution u to (5) is of class $\mathcal{C}^1$ but not $\mathcal{C}^2$, and the regularity depends on those of f, g. We refer to Caffarelli (1998) for details.

Now we consider the optimal stopping problem (4) with

$$f(x_1, \ldots, x_d) = -c \quad \text{and} \quad g(x_1, \ldots, x_d) = \max\{x_1, \ldots, x_d, 0\}. \tag{6}$$

The following lemma characterizes the value function u.

LEMMA 1: *Let u be the value function defined by (4), with $\mathcal{T} := \{\tau \text{ is a stopping time} : \mathbb{E}\tau < \infty\}$. Then u is a viscosity solution to*

$$\min\{-\mathcal{L}u + c, u - g\} = 0. \tag{7}$$

Moreover, if there exist $K_1 < c$ and $K_2 > 0$ such that

$$\sum_{i=1}^{d} \sup_{x \in \mathbb{R}} |b_i(x)| < K_1 \quad \textit{and} \quad \sup_{i,j} \sup_{x \in \mathbb{R}} |\sigma_{ij}(x)| < K_2,$$

then we have for some $C > 0$,

$$g \le u \le \sum_{i=1}^{d} |x_i| + C. \tag{8}$$

PROOF: The fact that u is a viscosity solution to (7) follows from the dynamic programming principle (5). By taking $\tau = 0$, we get $u \ge g(x_1, \ldots, x_d)$. It is easy to see that $g(x_1, \ldots, x_d) \le |x|$. So

$$\max\left\{X_1^{x_1}(\tau), \ldots, X_d^{x_d}(\tau), 0\right\} \le \sum_{i=1}^{d} |x_i| + g(X_1^0(\tau), \ldots, X_d^0(\tau)) \le \sum_{i=1}^{d} |x_i| + |X_\tau^0|.$$

Note that

$$|X_\tau^0| \le \sum_{i=1}^{d} \int_0^\tau |b_i(X_s^0)| ds + \sum_{i=1}^{d}\sum_{j=1}^{d} \left| \int_0^\tau \sigma_{ij}(X_s^0) dB_j(s) \right|,$$

which implies that

$$\begin{aligned}
\mathbb{E}|X_\tau^0| &\le K_1 \mathbb{E}\tau + \sum_{i=1}^{d}\sum_{j=1}^{d} \mathbb{E}\left| \int_0^\tau \sigma_{ij}(X_s^0) dB_j(s) \right| \\
&\le K_1 \mathbb{E}\tau + L \sum_{i=1}^{d}\sum_{j=1}^{d} \mathbb{E}\left[\left(\int_0^\tau \sigma_{ij}^2(X_s^0) ds \right)^{\frac{1}{2}} \right] \\
&\le K_1 \mathbb{E}\tau + L d^2 K_2 \sqrt{\mathbb{E}\tau},
\end{aligned}$$

where the second inequality is due to the *Burkholder-Davis-Gundy inequality* (see Revuz and Yor 1999, Chapter IV). Consequently,

$$\begin{aligned}
u &\le \sum_{i=1}^{d} |x_i| + \sup_{\tau \in \mathcal{T}} \left\{ (K_1 - c)\mathbb{E}\tau + L d^2 K_2 \sqrt{\mathbb{E}\tau} \right\} \\
&\le \sum_{i=1}^{d} |x_i| + \frac{L^2 d^4 K_2^2}{4(c - K_1)},
\end{aligned}$$

which yields (8). ■

Specializing to the optimal stopping problem (1), which is the focus of the analysis in the next sections, we get the following corollary.[5]

[5]In terms of the SDE (2) this is the case when $b = 0$, and $\sigma = I$ where I is the identity matrix. The application examples described in the Appendix are consistent with this case. Several of the results in the next section can also be obtained for the general SDE (2) under some conditions. This is a standard technical issue that is not central to the results presented here, and therefore not considered for ease of presentation.

Corollary 1: *Let u be the value function defined by (1), with $\mathcal{T} := \{\tau$ is a stopping time $: \mathbb{E}\tau < \infty\}$. Then u is a viscosity solution to*

$$\min\left\{-\frac{1}{2}\Delta u + c, u - g\right\} = 0, \tag{9}$$

where Δ is the Laplacian operator, $\sum_{i=1}^{d} \partial^2/\partial x_i^2$. Moreover, we have for some $C > 0$,

$$g \le u \le \sum_{i=1}^{d} |x_i| + C. \tag{10}$$

Corollary 1 asserts that the value function u satisfies the PDE (9), with at most linear growth. We will show in the synthesis part that such a solution is unique. Once the value function u is determined, then we construct an optimal strategy τ^* by

$$J_x(\tau^*) = u(x). \tag{11}$$

More precisely, starting at a position $x \in \{u > g\}$, the search will continue until it enters the contact set:

$$\tau^* = \inf\{t > 0 : B^x \in \{u = g\}\}. \tag{12}$$

3. Synthesis

In this section, we prove that there exists a unique viscosity solution to the PDE (9). We consider the notion of viscosity solution as in Crandall and Lions (1983), Ishii (1987, 1989) and Crandall et al. (1992). To avoid technical details, we state the definition of viscosity solution in our scenario.

Definition 1: *Let u be a continuous function.*

1. *We say that $-\frac{1}{2}\Delta u + c \le 0$ at x^0 in the viscosity sense if for any $\varphi \in \mathcal{C}^2$ which touches u at x^0 from above, we have $-\frac{1}{2}\Delta\varphi(x^0) + c \le 0$. We call u a subsolution to (9) if $-\frac{1}{2}\Delta u + c \le 0$ in the viscosity sense at all points where $u - g > 0$.*

2. *We say that $-\frac{1}{2}\Delta u + c \ge 0$ at x^0 in the viscosity sense if for any $\varphi \in \mathcal{C}^2$ which touches u at x^0 from below, we have $-\frac{1}{2}\Delta\varphi(x^0) + c \ge 0$. We call u a supersolution to (9) if $u - g \ge 0$ and $-\frac{1}{2}\Delta u + c \ge 0$ in the viscosity sense in $\mathbb{R}^d$.*

3. *We call u a viscosity solution to (9) if and only if u is both a subsolution and a supersolution to (9).*

For example, to have a supersolution, we only need to construct $f \in \mathcal{C}^1(\mathbb{R}^d)$ such that $f \geq g$, f is $\mathcal{C}^2$ in some open subsets $U \subset \mathbb{R}^d$ and $V = \mathbb{R}^d \backslash \bar{U}$, and

$$\Delta f|_U \leq 2c, \quad \Delta f|_V \leq 2c \text{ in the classical sense,}$$

where $\bar{U} = U \cup \partial U$, and ∂U is the boundary of U. Then automatically $\Delta f \leq 2c$ in $\mathbb{R}^d$ in the viscosity sense.

Moreover, if f_1, f_2 are two supersolutions, $\min\{f_1, f_2\}$ is also a supersolution. And the same is true for subsolutions if we change "min" to "max".

In the sequel, let $\mathcal{B}_R$ be the ball of radius $R > 0$, and $\partial\mathcal{B}_R$ its boundary. We first prove a comparison principle in bounded domains.

LEMMA 2 *(Comparison principle in $\mathcal{B}_R$): Assume that u_1 is a supersolution to (9), and u_2 is a subsolution to (9). If $u_1 \geq u_2$ on $\partial\mathcal{B}_R$ for some $R > 0$, then $u_1 \geq u_2$ in $\mathcal{B}_R$.*

PROOF: Assume by contradiction that

$$\sup(u_2 - u_1) = (u_2 - u_1)(x^0) > 0,$$

for some $x^0 \in \mathcal{B}_R$. Now we perturb u_1 a little bit and suppose that $u_2 - u_1 - \epsilon(x - x^0)^2$ obtains its one positive local maximum at x^0_ϵ near x^0 for some $\epsilon > 0$ small enough. Hence

$$\Delta(u_2 - u_1)(x^0_\epsilon) \leq 2\epsilon < 0. \tag{13}$$

Since u_1 is a supersolution, $u_1(x) \geq g(x)$ for all $x \in \mathbb{R}^d$. Consequently,

$$u_2(x^0_\epsilon) > u_1(x^0_\epsilon) \geq g(x^0_\epsilon),$$

and then by the fact that u_2 is a subsolution, we have

$$-\frac{1}{2}\Delta u_2(x^0_\epsilon) + c \leq 0.$$

Since u_1 is a supersolution, $-\frac{1}{2}\Delta u_1 + c \geq 0$ and therefore $\Delta(u_2 - u_1)(x^0_\epsilon) \geq 0$ which contradicts with (13), which completes the proof. ∎

Existence: Now we construct a viscosity solution to the PDE (9). Consider the following variational inequality

$$\begin{cases} \min\left\{-\frac{1}{2}\Delta u + c, u - g\right\} = 0 \text{ for } x \in \mathcal{B}_R, \\ u = g \text{ for } x \in \partial\mathcal{B}_R. \end{cases} \tag{14}$$

We say that u is a supersolution (resp. subsolution) to (14), if the corresponding inequalities in Definition 1 hold inside $\mathcal{B}_R$ and on $\partial\mathcal{B}_R$. A solution to (14) is both a supersolution and a subsolution to (14). For each R, the existence and uniqueness of the solution to (14) can be proved by the standard Perron's method (Crandall et al. 1992, Ishii 1989) and we denote the unique solution as u_R. In fact,

$$u_R(x) = \sup_{\tau\in\mathcal{T}} \mathbb{E}\left[\max\left\{B_1^{x_1}(\tau\wedge\tau_R), \ldots, B_d^{x_d}(\tau\wedge\tau_R), 0\right\} - c\tau\wedge\tau_R\right],$$

where $\tau_R := \inf\{t > 0 : B^x(t) \in \partial\mathcal{B}_R\}$.

It is easily checked that g is a subsolution to (14). It follows from Lemma 2 that $u_R \geq g$. Next we want to get an uniform upper bound of u_R. To do so, we define

$$\psi_\theta(x_1) = \begin{cases} 0 & \text{for } x_1 \leq -\frac{1}{4\theta} \\ \theta\left(x_1 + \dfrac{1}{4\theta}\right)^2 & \text{for } x_1 \in (-\frac{1}{4\theta}, \frac{1}{4\theta}) \\ x_1 & \text{for } x_1 \geq \frac{1}{4\theta}, \end{cases} \tag{15}$$

for some $\theta > 0$. This function is constructed as a modification of $\max\{x_1, 0\}$, where we replace the cusp by a parabola. Also

$$\psi_\theta \in \mathcal{C}^1(\mathbb{R}) \text{ and } \Delta\psi_\theta \leq 2\theta.$$

Therefore,

$$\Psi(x) := \sum_{i=1}^{d} \psi_{c/d}(x_i) \geq g(x)$$

satisfying $\Delta\Psi \leq 2c$ and $\Psi \in \mathcal{C}^1$, which indicates that Ψ is a supersolution to (14). By Lemma 2, $u_R \leq \Psi$ for all $R > 0$.

Note that for any $R_1 \geq R_2$, we have $u_{R_1} \geq u_{R_2}$ on $\partial\mathcal{B}_{R_2}$. Hence by comparison,

$$u_{R_1} \geq u_{R_2} \text{ in } \mathcal{B}_{R_2}.$$

Thus by the monotonicity, we can take the pointwise limit and set

$$u := \lim_{R\to\infty} u_R. \tag{16}$$

Since $u_R \geq g$, the limit is non-trivial. From Perron's method again we know that u is indeed a solution to (9). In addition to the existence result, since $\psi_{c/d} \leq \max\{x_i, 0\} + \frac{d}{16c}$, the function u constructed also satisfies the growth condition (10):

$$g \leq u \leq \Psi \leq \sum_{i=1}^{d} \max\{x_i, 0\} + \frac{d^2}{16c}. \tag{17}$$

Uniqueness: We now consider uniqueness and show that among continuous functions that have less than quadratic growth at infinity, the solution u obtained is unique.

LEMMA 3 *(Comparison principle): Let u_1, u_2 be respectively a subsolution and a supersolution to (9) in $\Omega \subseteq \mathbb{R}^d$, and suppose that*

$$\lim_{\sum_{i=1}^d x_i^2 \to \infty, (x_1,...,x_d) \in \Omega} \frac{\max\{u_1(x_1, ..., x_d), 0\} + \max\{-u_2(x_1, ..., x_d), 0\}}{\sum_{i=1}^d x_i^2} = 0, \tag{18}$$

and $u_2 \geq u_1$ on $\partial\Omega$ (this condition can removed if $\Omega = \mathbb{R}^d$).

Then we have $u_2 \geq u_1$ in Ω.

We prove this Lemma in the Appendix. With this Lemma, we are able to compare sub and supersolutions in $\mathbb{R}^d$ as long as condition (18) is satisfied. Combined with Lemma 1, we get a complete characterization of the value function u.

PROPOSITION 1: *Let u be the value function defined by (1), with $\mathcal{T} := \{\tau$ is a stopping time $: \mathbb{E}\tau < \infty\}$. Then u is the unique viscosity solution to (9) with at most linear growth.*

PROOF: By Corollary 1, u is a solution to (1) with linear growth at infinity. By the comparison principle we know this u is the unique viscosity solution to (9) among all continuous functions satisfying

$$\lim_{\sum_{i=1}^d x_i^2 \to \infty} \frac{|u(x_1, ..., x_d)|}{\sum_{i=1}^d x_i^2} = 0.$$

■

4. Star-shapedness of the Free Boundary

Let u be a solution of (9). Recall that the *free boundary* of u is the interface of the sets $\{u > g\}$ and $\{u = g\}$ which we denote by $\Gamma(u)$. Several regularity results of $\Gamma(u)$ can be found in Caffarelli

(1998). In this paper, we are interested in the global geometric property of $\Gamma(u)$. In this section we prove the star-shapedness.

First let us define “star-shapedness”. We say a hyperplane $S \subset \mathbb{R}^d$ is *star-shaped with respect to the origin* if for every $0 \neq x^0 \in S$,

$$\{tx^0, t > 1\} \cap S = \emptyset.$$

To prove the free boundary is star-shaped, we only need to prove the following result.

PROPOSITION 2: *Let u be a solution to (9). If $u(x) = g(x)$ for some x, then for any $t \geq 1$ we have $u(tx) = g(tx)$.*

PROOF: Let $v(x) := \frac{1}{t}u(tx)$. We first show that v is a subsolution to (9). In fact, for any $x \in \mathbb{R}^d$, if $v(x) > g(x)$ then

$$u(tx) > tg(x) = t \max\{x_1, ...x_d, 0\} = g(tx).$$

Thus,

$$-\frac{1}{2}\Delta u(tx) \leq c,$$

which implies that,

$$-\frac{1}{2}\Delta v(x) = -\frac{t}{2}\Delta u(tx) \leq tc \leq c.$$

So we conclude that v is a subsolution. Now take $x^0 \in \mathbb{R}^d$ such that $u(x^0) = g(x^0)$. From the order of u and v, we get

$$u(tx^0) \leq tu(x^0) = tg(x^0) = g(tx^0).$$

On the other hand, $u(tx^0) \geq g(tx^0)$ by definition, so we have that $u(tx^0) = g(tx^0)$. ∎

Figure 1 shows the continuation and stopping regions, as well as the free boundary separating them for the case of $d = 2$. The figure illustrates the star-shapedness of the free boundaries.

As shown by Figure 1, the optimal search strategy is quite intuitive—roughly speaking, the DM should stop searching and adopt alternative i if and only if x_i is relatively high compared with x_j and the outside option of 0, and she should stop searching and adopt the outside option when both x_1 and x_2 are relatively low. When x_j is relatively low, the DM will continue to search on the two alternatives if and only if x_i is near 0, so as to make a clear distinction between alternative i and the outside option. When both x_1 and x_2 are relatively high, the DM will continue to search if and only if x_1 and x_2 are close to each other, so as to to make a clear distinction between the two alternatives 1 and 2.

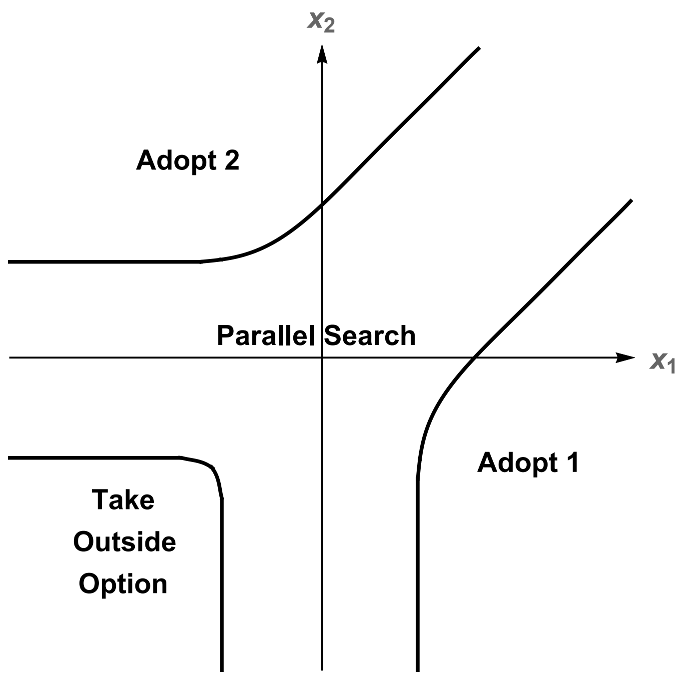

Figure 1: Optimal parallel search strategy in two dimensions.

5. Asymptotics

In this section, we study the free boundary of the solution near $x_1 = \ldots = x_d \to \infty$. We provide a detailed analysis for the case with $d = 2$, and compare it with the case in which the DM can only search sequentially, learning one alternative at a time. We also provide lower and upper bounds for the general case with $d \geq 2$.

5.1. Dimension of $d = 2$

The PDE (9) specializes to

$$\min\left\{-\frac{1}{2}\Delta u + c, u - \max\{x_1, x_2, 0\}\right\} = 0. \tag{19}$$

The PDE (19) does not have an explicit solution for the case of $d = 2$, so it is natural to ask about the properties of the solution, in particular those of free boundaries. There are three interesting regimes of asymptotic behavior:

1. $x_1 \to 0$ and $x_2 \to -\infty$,

2. $x_1 \to -\infty$ and $x_2 \to 0$,

3. $x_1 = x_2 \to \infty$.

The cases 1 and 2 boil down to the search problem of one alternative, since the other alternative has large negative value and thus loses the competition to its counterpart. A classical smooth-pasting technique shows that the distance of the free boundaries to x-axis (resp. y-axis) at $-\infty$ is $\frac{1}{4c}$, as illustrated in Figure 1. The case (3) is subtle, since the values of two products are close so there is a competitive search. One interesting question is to determine the distance from the free boundary to the line $x_1 = x_2$ at infinity.

We start with the following change of coordinates: $t = \frac{x_1+x_2}{\sqrt{2}}$ and $s = \frac{x_1-x_2}{\sqrt{2}}$. Consider the domain $t \geq 0$, and the PDE (19) becomes

$$\min\left\{-\frac{1}{2}\Delta\tilde{u} + c, \tilde{u} - \tilde{g}\right\} = 0 \text{ for all } (t,s) \in \mathbb{R}^2 \tag{20}$$

where

$$\tilde{u}(t,s) := u\left(\frac{t+s}{\sqrt{2}}, \frac{t-s}{\sqrt{2}}\right) \quad \text{and} \quad \tilde{g}(t,s) := \max\left\{\frac{t+|s|}{\sqrt{2}}, 0\right\}.$$

We first prove a lower bound on the free boundary $\Gamma(u)$ for $t \geq 0$ by the following lemma.

Lemma 4 *(Lower bound of the free boundary): For $\theta > 0$, let*

$$\eta_\theta(t,s) := \begin{cases} \dfrac{t}{\sqrt{2}} + \theta s^2 + \dfrac{1}{8\theta} & \text{for } |s| \leq \frac{1}{2\sqrt{2}\theta} \\ \dfrac{t}{\sqrt{2}} + \dfrac{|s|}{\sqrt{2}} & \text{for } |s| > \frac{1}{2\sqrt{2}\theta}. \end{cases}, \tag{21}$$

which is a $\mathcal{C}^1$ function. Then we have,

$$\tilde{u}(t,s) \geq \eta_c(t,s) \quad \text{in } \mathbb{R}^2.$$

Moreover, for $t \geq 0$, the free boundary $\Gamma(u)$ lies inside $\{|x_1 - x_2| = \sqrt{2}|s| \geq \frac{1}{2c}\}$.

Proof: Note that η_θ is an approximation of $\tilde{g}$ for $t \geq 0$. Moreover, it is not hard to check when $\theta = c$,

$$\min\left\{-\frac{1}{2}\Delta\eta_c + c, \eta_c - \rho\right\} = 0 \quad \text{for } (t,s) \in \mathbb{R}^2$$

where

$$\rho := \frac{t+|s|}{\sqrt{2}} \leq \tilde{g}.$$

We know that $\eta_c(\frac{x_1+x_2}{\sqrt{2}}, \frac{x_1-x_2}{\sqrt{2}})$ is actually a subsolution to (19) and the comparison principle yields $\tilde{u} \geq \eta_c$. Observe that $\eta_c = \rho = \tilde{g}$ for $|s| \geq \frac{1}{2\sqrt{2}c}$ and $t \geq 0$. Therefore, in the half plane $t \geq 0$, the free boundary $\Gamma(u)$ lies inside $\{|s| \geq \frac{1}{2\sqrt{2}c}\}$. ■

The result that $\Gamma(u)$ lies inside $\{|x_1 - x_2| \geq \frac{1}{2c}\}$ can be viewed as a "lower bound" of the free boundary.

Now we turn to look for an "upper" bound of the free boundary. We need the following result.

LEMMA 5: *For $\epsilon \in (0, c]$, let*

$$\bar{\varphi}_\epsilon(t,s) := \frac{1}{4c}h(\alpha t) + \eta_{c-\epsilon}(t,s)$$

where $h(t) := \max\{1-t, 0\}^2$ and $\alpha := 2\sqrt{c\epsilon}$. Then we have for all $t \geq 0$,

$$\tilde{u}(t,s) \leq \bar{\varphi}_\epsilon(t,s).$$

PROOF: It follows from (17) that

$$\tilde{g}(t,s) \leq \tilde{u}(t,s) \leq \max\left\{\frac{t+s}{\sqrt{2}}, 0\right\} + \max\left\{\frac{t-s}{\sqrt{2}}, 0\right\} + \frac{1}{4c}.$$

We get

$$\frac{|s|}{\sqrt{2}} = \tilde{g}(0,s) \leq \tilde{u}(0,s) \leq \frac{|s|}{\sqrt{2}} + \frac{1}{4c}. \tag{22}$$

Now we want to compare $\bar{\varphi}_\epsilon$ with u in the half plane $t > 0$. On the boundary of $t = 0$,

$$\begin{aligned}
\bar{\varphi}_\epsilon(0,s) &= \frac{1}{4c} + \eta_{c-\epsilon}(0,s) \\
&\geq \frac{1}{4c} + \tilde{g}(0,s) && \text{(by definition of } \eta_{c-\epsilon}) \\
&\geq \tilde{u}(0,s) && \text{(by (22))}.
\end{aligned}$$

Also it is not hard to check that $\bar{\varphi}_\epsilon \in \mathcal{C}^1$ and $\bar{\varphi}_\epsilon(t,s) \geq \tilde{g}(t,s)$ for all $t \geq 0, s \in \mathbb{R}$. Moreover when $|s| \leq \frac{1}{2\sqrt{2}(c-\epsilon)}$, we have

$$\Delta\bar{\varphi}_\epsilon = \frac{\alpha^2}{2c} + 2(c-\epsilon) \leq 2c \text{ if } \alpha \leq 2\sqrt{c\epsilon}.$$

When $|s| \geq \frac{1}{2\sqrt{2}(c-\epsilon)}$, there is $\Delta\bar{\varphi}_\epsilon \leq 2\epsilon \leq 2c$. Finally note that both $\tilde{u}$ and $\bar{\varphi}_\epsilon$ have linear growth at infinity. The comparison principle (Lemma 3 with $\Omega = \{t > 0\}$) yields $\tilde{u} \leq \bar{\varphi}_\epsilon$ for $t > 0$. ■

Based on Lemmas 4 and 5, we can obtain the asymptotic behavior of solutions u close to

$x_1 = x_2 \to +\infty$. To provide a quantitative description about the convergence of the free boundary of u to the one of η_c as $t = x_1 + x_2 \to \infty$, we define the distance function

$$d_{FB}(T) := \text{ distance between } \Gamma(u)|_{\frac{x_1+x_2}{\sqrt{2}} \geq T} \text{ and } \left\{|x_1 - x_2| = \frac{1}{2c}\right\}.$$

Here $\{|x_1 - x_2| = \frac{1}{2c}\}$ is the free boundary of η_c. By symmetry of $\Gamma(u)$ with respect to the line of $x_1 - x_2 = 0$, we only need to consider the situation when $x_1 - x_2 \geq 0$.

The following proposition characterizes the asymptotic behaviors of both the value function and the free boundary close to $x_1 = x_2 \to +\infty$, with the proof provided in Appendix.

Proposition 3 *(Upper bound of the free boundary): For (x_1, x_2) in the neighborhood of $x_1 = x_2 \to +\infty$,*

$$u(x_1, x_2) \to \eta_c\left(\frac{x_1 + x_2}{\sqrt{2}}, \frac{x_1 - x_2}{\sqrt{2}}\right) = \begin{cases} \dfrac{x_1 + x_2}{2} + \dfrac{c|x_1 - x_2|^2}{2} + \dfrac{1}{8c} & \text{for } |x_1 - x_2| \leq \frac{1}{2c} \\ \dfrac{x_1 + x_2}{2} + \dfrac{|x_1 - x_2|}{2} & \text{for } |x_1 - x_2| > \frac{1}{2c}. \end{cases},$$

$$\Gamma(u) \to \left\{|x_1 - x_2| = \frac{1}{2c}\right\}.$$

Moreover, for all $T \geq \frac{1}{2c}$, then

$$d_{FB}(T) \leq \frac{1}{8\sqrt{2}c^3T^2} + O\left(\frac{1}{c^5T^4}\right).$$

As for the limit $\eta_c(x_1 + x_2, x_1 - x_2)$, the distance of the free boundary to the line of $x_1 = x_2$ is always $\frac{1}{2^{3/2}c}$. Note that $\frac{1}{2^{3/2}c} > \frac{1}{4c}$ which is the distance of the free boundaries to x or y-axis at $-\infty$. This means that the search region is larger in case of competition. In other words, people have larger tolerance for search if two products are as good as each other.

5.2. Comparison with Sequential Search

One could consider a different technology for information search, as the one considered in Ke et al. (2016), where the DM searches costly and sequentially over multiple alternatives, learning only one alternative at a time. Let the sequential search cost be c'.

Suppose $c' = c/2$. That is, it costs twice as much to search two alternatives in parallel as to search one alternative at a time. Note that in the sequential search case, the DM could replicate any parallel search strategy considered above by alternating infinitely fast between the two alternatives.

Therefore, we have that the region in x_1-x_2 space where it is optimal to continue to search (i.e. $\{u > g\}$) is larger for the case of sequential search compared with that for the case of parallel search. In other words, the contact set is further away from the origin for the case of sequential search.

Figure 2 illustrates the sequential and parallel search strategies for the case with $d = 2$. The black solid lines represent the free boundaries for the case of parallel search, the same as Figure 1; while the gray solid lines represent the free boundaries for the case of sequential search. The gray dashed line represents $x_1 = x_2$. For the case of sequential search, when it is optimal for the DM to continue to search, the DM optimally searches alternative i if and only if $x_i \geq x_j$ (Ke et al. 2016). The figure illustrates that the gray lines are further away from the origin than the black lines.

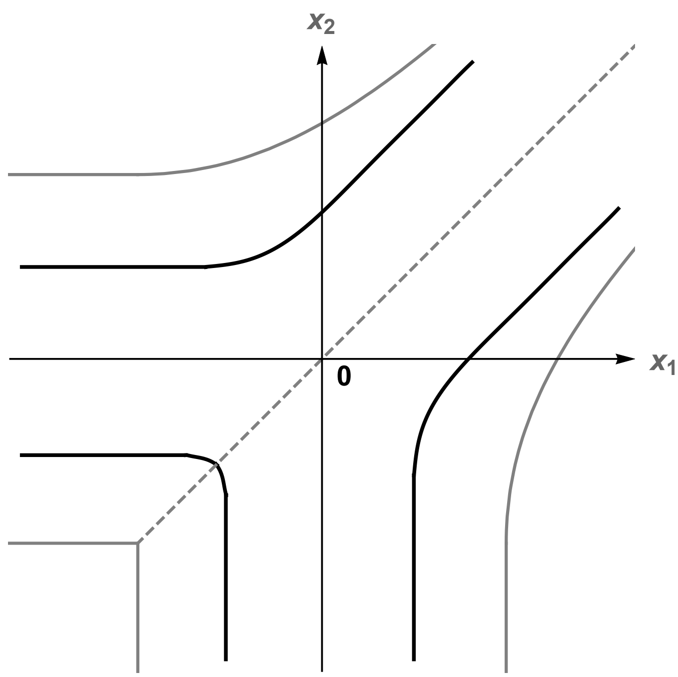

Figure 2: Comparison of the optimal parallel search strategy under search cost c with the optimal sequential search strategy under search cost $c' = c/2$.

One could also wonder how the asymptotic behavior of the free boundary compares between sequential and parallel search. On can obtain that when the DM searches sequentially, the distance of the free boundary to the line $x_1 = x_2$ when $x_1 = x_2 \to +\infty$ converges to $\frac{1}{4\sqrt{2c'}} = \frac{1}{2\sqrt{2c}}$ (Ke et al. 2016) which is the same as the distance in the case of parallel search. That is, the gray and black lines in Figure 2 will converge to $|x_1 - x_2| = \frac{1}{2c}$ as x_1 and x_2 go to positive infinity.

It is also interesting to consider the case in which the DM has the option to either search only one alternative at cost c' or search both alternatives at cost c with economies of scale on the number of alternatives searched, that is, $c' \in (c/2, c)$. Although the full-scale analysis of this problem is beyond the scope of this paper, one could expect that in such a setting when it is optimal to continue to search for information, the DM will choose to search for information on both alternatives simultaneously when the expected valuations of the two alternatives are relatively close, and choose to search for information on only one alternative otherwise.

It is interesting, however, that one can obtain a general result in that setting that for the state close to $x_1 = x_2 \to \infty$ it is always optimal to choose the search technology where both alternatives are being searched simultaneously.

Proposition 4: *Consider a DM, who can search either in parallel at cost c or sequentially at cost $c' \in (c/2, c)$. For x_1 and x_2 sufficiently high and close to each other, it is optimal for the DM to search in parallel.*

Here we provide an intuitive sketch of proof for the proposition. A formal proof can be obtained by applying Lemma 7 below and invoking the dynamic programming principle, and is omitted.

When x_1 and x_2 are high, the DM is most likely to choose one of the alternatives rather than the outside option, and just does not know which alternative to choose. The DM is then mostly concentrated on the difference $x_1 - x_2$ to see when this difference is high enough so that the DM makes a decision on which alternative to pick and stop the search process. As shown above, at the limit, when $|x_1 - x_2| \geq \frac{1}{2c}$, the DM prefers to stop and choose one alternative than to continue to search either sequentially or in parallel. On the other hand, at the limit, when $|x_1 - x_2| < \frac{1}{2c}$, the DM will choose to continue to search. By searching the two alternatives in parallel in an infinitesimal time dt, the DM pays a search cost of cdt and gets an update on $x_1 - x_2$ as $dx_1 - dx_2$, the variance of which is $2dt$; on the other hand, by searching one alternative (say, alternative 1) sequentially in an infinitesimal time dt, the DM pays a search cost of $c'dt$ and gets an update on $x_1 - x_2$ as dx_1, the variance of which is dt. Therefore, the parallel search yields variance per search cost $2/c$, which is greater than $1/c'$, the variance per search cost in the case of sequential search. To summarize, for $c' \in (c/2, c)$, it is less expensive to obtain a certain variation when searching two alternatives simultaneously, than just searching sequentially on one alternative. This implies that it is more cost-effective for the DM to search in parallel.

Figure 3 presents an example of the DM's optimal search strategy in this context of economics of scale of search costs, and illustrates that for x_1 and x_2 sufficiently high and close to each other, it is optimal to search the two alternatives in parallel.

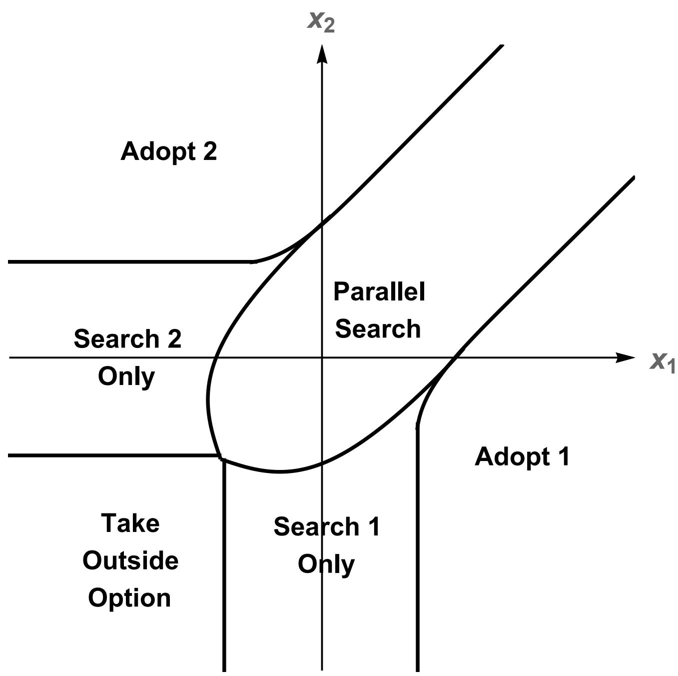

Figure 3: The DM's optimal search strategy when he can search either sequentially or in parallel, with $c' = 2c/3$.

5.3. General dimension

Now we study the quantitative properties of the free boundary in the general dimension case. We provide an "upper" and "lower" bound of the free boundary.

First for the upper bound, we will show that in the positive regime $x_i \geq 0$ for all i, the free boundary can not be too far away from the set $\{x_i = x_j \text{ for some } i \neq j\}$, and the distance grows at most linearly in the dimension d.

For any $\gamma > 0$, define

$$A(\gamma) = \{(x_1, ..., x_d) |\, x_i \geq 0,\, |x_i - x_j| \geq \gamma \text{ for all } i \neq j\}\,. \tag{23}$$

The following proposition presents the first main result in this section, with the proof in the Appendix.

PROPOSITION 5: *Let u_d be the solution to (9) in $\mathbb{R}^d$. There exists $\gamma > 1$ independent of d, c such that $\Gamma(u_d)$ lies inside the complement of $A(\frac{\gamma d}{c})$, i.e., $u_d(x) = g(x)$ for $x \in A(\frac{\gamma d}{c})$.*

Note that in the case of parallel search here, for fixed c, this result does not yield that the distance between the free boundary and $\{x_1 = \ldots = x_d\}$ is bounded when $d \to \infty$. We show that this is indeed the case when we investigate a "lower bound" of the free boundary in Proposition 6. In contrast, in the case of sequential search with cost c', when $d \to \infty$, that distance converges to $\frac{1}{\sqrt{2c'}}$ (see Ke et al. (2016), p. 3591). However, if we set $c' = c/d$ for a fixed c, then as $d \to \infty$, we would get $c' \to 0$, and correspondingly that distance for sequential search becomes unbounded. On the other hand, if we fix c' and let $c = dc'$ grow linearly with d, then by Proposition 5, we would have the distance between the free boundary and $\{x_1 = \ldots = x_d\}$ for parallel search to be bounded.

This can also be seen, by a similar argument used above, that as we can replicate parallel search by alternating among alternatives in sequential search, it must be that the "search region" (i.e., $\{u > g\}$) in the case of parallel search with cost c is a subset of that in the case of sequential search with cost $c' = c/d$. As the distance between the free boundary and $\{x_1 = \ldots = x_d\}$ is bounded for sequential search for a fixed c', it must be that it is also bounded for parallel search for $c = dc'$. Note also that even though the free boundary is unbounded when $d \to \infty$ for fixed c, the search process ends in finite time with probability one as the state moves, over time, away from $x_1 = ... = x_d$. The question of whether the distance for parallel search increases in d for fixed $c' = c/d$ remains open.

Next we study the "lower bound" of the free boundary. Let us consider the following auxiliary problems: for each $d \geq 1$ consider

$$\min\left\{-\frac{1}{2}\Delta w_d + c, w_d - \rho\right\} = 0 \text{ in } \mathbb{R}^d, \tag{24}$$

where $\rho = \max\{x_1, x_2, ..., x_d\}$. The free boundary of w_d, $(\Gamma(w_d))$ is defined as the boundary of the set $\{w_d = \rho\}$.

When $d = 1, 2$, by direct computation, we have that,

$$w_1 = \psi_c, \text{ and } w_2 = \eta_c,$$

where ψ_c and η_c are given by (15) and (21), respectively. Since $\rho \leq g$, w_d is a subsolution to the original PDE (9) and by comparison $u(= u_d) \geq w_d$. We will show that w_d provides the full information of the behavior of u near $x_1 = ... = x_d \to \infty$.

Let us introduce some notation. We write the positive $x_1, ..., x_d$ directions as $e_1, ..., e_d$ respectively and

$$\tau_d := \frac{\sum_{i=1}^d e_i}{\sqrt{d}}, \; H_{\tau_d} := \{v| \, v \cdot \tau_d = 0\}, \text{ and } t = \frac{\sum_{j=1}^d x_j}{\sqrt{d}}. \tag{25}$$

The following lemma shows that we can reduce the study of w_d to H_{τ_d}, where the proof is provided in Appendix.

Lemma 6: *The expression $w_d - \sum_{j=1}^d x_j/d$ is a constant function in the τ_d direction. The free boundary of w_d is the surface of one infinitely long columnar with τ_d as its longitudinal axis.*

In the following lemma, we show that the free boundary of w_d can be arbitrarily close to the one of u if $\sum_{j=1}^d x_j$ is large. Since we are only interested in the region near $x_1 = ... = x_d$, let us define the following open neighborhood:

$$A_1(R) := \left\{ x \in \mathbb{R}^d \,|\, dist(x, \{s\tau_d, s \in \mathbb{R}\}) \leq R \right\}.$$

Lemma 7: *For any $\epsilon \in (0,1)$ and $R \geq 1$, the distance between the free boundaries of u and w_d is bounded by $R\epsilon$ in the set*

$$A_1(R) \cap \left\{ \sum_{j=1}^d x_j \geq \frac{d}{c}\sqrt{\frac{\gamma}{\epsilon}} \right\}$$

where γ is a universal constant given in Proposition 5.

We provide the proof to Lemma 7 in the Appendix. Though more complicated, the idea of the proof follows from the one of Lemma 5.

We are interested in the most competitive region $x_1 = ... = x_d \to \infty$ where d products are close. From Proposition 5, we know that $\Gamma(u)$ can not be too far away from the axis $x_1 = ... = x_d$. Now we try to answer the question that how close this distance can be. By Lemma 7, we can identify $\Gamma(u)$ asymptotically with $\Gamma(w_d)$, and by Lemma 6, we only need to study $\Gamma(w_d) \cap H_{\tau_d}$.

We make the following definition: for each $d \geq 1$, define r_d to be the smallest number such that there exists $x \in H_{r_d}$ satisfying

$$|x| = r_d \text{ and } w_d(x) = \rho(x).$$

From the definition whenever $x \in H_{r_d}$ and $w_d(x) = \rho(x) = g(x)$, then $|x| \geq r_d$. For example, when $d = 1, 2$, by the definition of ψ_c and η_c, we have that $r_1 = \frac{1}{4c}$ and $r_2 = \frac{1}{2\sqrt{2}c}$.

Before the proposition, we need one technical lemma which compares w_d and $w_{d'}$ for $d \neq d'$.

Lemma 8: *For any $k > j \geq 1$, let $\{i_1, i_2, ..., i_k\}$ be a permutation of $\{1, ..., k\}$. Consider two solutions $w_k(x_1, ..., x_k)$ and $w_j(x_{i_1}, ..., x_{i_j})$. We can view w_j as a function in $\mathbb{R}^k$ by trivial extension:*

$$\tilde{w}_j(x_{i_1}, ..., x_{i_k}) := w_j(x_{i_1}, ..., x_{i_j}).$$

Then $w_k \geq \tilde{w}_j$ in $\mathbb{R}^k$.

The lemma is a direct result of the comparison principle. With a slight abuse of notation, we still write w_j instead of $\tilde{w}_j$.

Now we prove the second main result of this section, which provides a lower bound on the free boundary.

Proposition 6: *Let $u = u_d$ be the solution to (9) in dimension d and r_d be given as the above. In the half plane $\left\{t \geq \frac{1}{c}\sqrt{\frac{\gamma d}{\epsilon}}\right\}$, the distance from the free boundary of u to the ray $\{s\tau_d,\, s \geq 0\}$ lies in the interval*

$$[\, r_d \,,\, r_d + \frac{\gamma d}{c}\epsilon \,],$$

where τ_d is defined in (25), and γ is a universal constant given in Proposition 5. Furthermore, for each $d \geq 3$,

$$\left(d - 2 - \frac{1}{d}\right) r_d^2 \geq (d-2) r_{d-1}^2.$$

In particular, r_d is increasing in d. Furthermore, $r_d \to \infty$ as $d \to \infty$.

Proof: Due to Proposition 5, $r_d \leq \frac{\gamma d}{c}$. We apply Lemma 7 with $R = \frac{\gamma d}{c}$ and then the first part of the result follows from the definition of r_d.

For the second part, take $k \geq 3$ and $x \in H_{\tau_k}$. Without loss of generality we assume

$$\rho_k(x) := \rho(x_1, ..., x_k) = \max\{x_1, ..., x_k\} = x_1 > 0.$$

The inequality holds due to $x \in H_{\tau_k}$. Suppose $w_k(x) = x_1 = \rho_k(x)$. Take any $k-2$ different numbers

$$\{i_2, i_3, ..., i_{k-1}\} \subset \{2, ..., k\}.$$

If

$$x_1^2 + x_{i_2}^2 + ... + x_{i_{k-1}}^2 < r_{k-1}^2,$$

by Lemma 8 it follows that

$$w_k(x_1, ..., x_k) \geq w_{k-1}(x_1, x_{i_2}, ..., x_{i_{k-1}}) > \max\{x_1, x_{i_2}, ..., x_{i_{k-1}}\} = x_1 = \rho_k(x),$$

which cannot happen due to our assumption $w_k = \rho_k$ at x. Thus we must have

$$x_1^2 + x_{i_2}^2 + ... + x_{i_{k-1}}^2 \geq r_{k-1}^2.$$

We can vary the subscripts and add up all the inequalities with respect to different combinations of $\{i_2, ..., i_{k-1}\}$. It ends up with

$$(k-2)x_1^2 + (k-3)(\sum_{j=2}^{k} x_j^2) \geq (k-2)r_{k-1}^2. \tag{26}$$

Due to the facts that $\sum_{j=1}^{k} x_j = 0$ and $x_1 = \max\{x_1, ..., x_k\}$, we can show

$$x_1^2 \leq \frac{k-1}{k}|x|^2,$$

and equality can be obtained when,

$$x_1 = \sqrt{\frac{k-1}{k}}|x|,\ x_2 = x_3 = ... = x_k = -\frac{1}{\sqrt{k(k-1)}}|x|^2.$$

Therefore (26) leads to

$$\left(k - 3 + \frac{k-1}{k}\right)|x|^2 \geq (k-2)r_{k-1}^2.$$

According to the assumption $w_k = \rho_k$ and the definition of r_k,

$$r_k^2 \geq |x|^2 \geq \frac{(k-2)}{(k-2-\frac{1}{k})}r_{k-1}^2.$$

To prove that $r_d \to \infty$ as $d \to \infty$, suppose by contradiction that r_d is bounded as $d \to \infty$. Then for each $d \geq 2$, there exists $(x_1^d, \dots, x_d^d) \in \{u = g\}$ such that $\sup_{1 \leq i \leq d} x_i^d \leq K$ for some K independent of d. It is well known that for $Z_1, \dots, Z_d$ i.i.d. $\mathcal{N}(0,1)$, $\mathbb{E}(\max_{1 \leq i \leq d} Z_i) \sim \sqrt{2 \log d}$ as $d \to \infty$. This implies that given any $c > 0$,

$$\mathbb{E}\left(\max\left(B_1^{x_1^d}(1), \dots, B_d^{x_d^d}(1), 0\right)\right) > \max(x_1^d, \dots, x_d^d, 0) + c.$$

for d sufficiently large. Therefore,

$$u(x_1^d, \dots, x_d^d) \geq \mathbb{E}\left(\max\left(B_1^{x_1^d}(1), \dots, B_d^{x_d^d}(1), 0\right)\right) - c > g(x_1^d, \dots, x_d^d).$$

This contradicts the fact that $(x_1^d, \dots, x_d^d) \in \{u = g\}$. This concludes the proof. $\blacksquare$

APPENDIX

Examples of Applications

Parallel Search on Product Attributes

Consider a consumer, whose utility of product i, U_i is the sum of the utility derived from each attribute of the product. $U_i = x_i + \sum_{t=1}^{T} a_{it}$, where x_i is the consumer's initial expected utility, and a_{it} is the utility of attribute t of product i, which is uncertain to the consumer before search. It is also assumed that a_{it} is i.i.d. across t and i, and without loss of generality, $\mathbb{E}[a_{it}] = 0$. There is an outside option of zero.

Each time by paying a search cost c, the consumer checks one attribute a_{it} for all products $i = 1, \ldots, d$. The consumer decides when to stop searching and upon stopping which product to buy so as to maximize the expected utility. After checking t attributes, the consumer's conditional expected utility of the product i is,

$$X_i(t) = \mathbb{E}_t[U_i] = x_i + \sum_{s=1}^{t} a_{is}.$$

Therefore, $X_i(t)$ is a random walk, which converges to the Brownian motion $B_i^{x_i}(t)$, when we scale a_{is} and the search cost c proportionally to infinitesimally small and take T to infinity. In the continuous-attribute analog, the consumer's parallel search problem is formulated as the optimal stopping problem in (1).

Bayesian Learning with Evolving State

Suppose that $dX(t) = \sigma\ dB(t)$ where σ is a diagonal matrix, with general element σ_{ii} in the diagonal, and that the signal of $X(t), S(t)$, a d-dimension vector follows $dS(t) = X(t)\ dt + y\ d\widetilde{B}(t)$, with $\widetilde{B}(t)$ being a d-dimensional Brownian montion independent of $B(t)$, y is a diagonal matrix, with general element in the diagonal y_{ii}. Suppose also that the prior of $X(0)$ is a normal with mean $\widehat{X}(0)$ and variance-covariance $\widehat{\rho}(0)^2$, with $\widehat{\rho}(0)$ being a diagonal matrix, with general element in the diagonal $\widehat{\rho}_{ii}(0)$. Then, the posterior mean of $X(t)$, $\widehat{X}(t)$, follows $d\widehat{X}_i(t) = \frac{\widehat{\rho}_{ii}(t)}{y_{ii}^2}\ d\overline{B}_i(t)$, for all i, with $\overline{B}(t)$ being a d-dimensional Brownian montion, and $\frac{d\widehat{\rho}_{ii}(t)}{dt} = \sigma_{ii}^2 - \frac{\widehat{\rho}(t)^2}{y_{ii}^2}$ for all i. So, we have $\widehat{\rho}(t) \to \sigma y$ as $t \to \infty$. Then, if $\widehat{\rho}(0) = \sigma y$, we have that $\widehat{X}(t)$ is stationary, $d\widehat{X}_{ii}(t) = \frac{\sigma_{ii}}{y_{ii}}\ d\overline{B}_i(t)$ for all i.

Proof of Lemma 3:

Proof: Fix $r \geq 1$. For any $R > r+1$ and $\epsilon \in (0, 1/d)$, let

$$u_2^\epsilon := (1-d\epsilon)u_2 + c\epsilon\left(\sum_{i=1}^d x_i^2 + \frac{d^3}{4c^2}\right).$$

We claim that u_2^ϵ is a supersolution to (14). Since $\Delta u_2 \leq 2c$, we have

$$\Delta u_2^\epsilon = (1-d\epsilon)\Delta u_2 + 2cd\epsilon \leq 2c.$$

Also because $u_2 \geq g$, we get

$$\begin{aligned} u_2^\epsilon &= (1-d\epsilon)u_2 + \epsilon\sum_{i=1}^d \left(cx^2 + \frac{d^2}{4c}\right) \\ &\geq (1-d\epsilon)\max\{x_1, ..., x_d, 0\} + d\epsilon\sum_{i=1}^d |x_i| \\ &\geq \max\{x_1, ..., x_d, 0\} = g. \end{aligned}$$

Next for any small $\epsilon > 0$, if we pick $R(\epsilon)$ large enough and then by the condition (18),

$$u_1 \leq c\epsilon R^2 \leq u_2^\epsilon \text{ on } \partial\mathcal{B}_R \cup \partial\Omega.$$

By comparison, $u_2^\epsilon \geq u_1$ in $\mathcal{B}_R \cap \Omega$ and in particular in $\mathcal{B}_r \cap \Omega$. Consequently,

$$(1-d\epsilon)u_2 + c\epsilon\left(r^2 + \frac{d^3}{4c^2}\right) \geq u_1.$$

Since we can choose ϵ to be arbitrarily small and then r to be large, we conclude that $u_1 \leq u_2$ in Ω. ■

Proof of Proposition 3:

Proof: Consider the line $x_1 + x_2 = \sqrt{2}t$ with fixed $t \geq 1$. By Lemma 5,

$$u\left(\frac{t+s}{\sqrt{2}}, \frac{t-s}{\sqrt{2}}\right) = \tilde{u}(t,s) \leq \bar{\varphi}_\epsilon(t,s).$$

By definition, using the notation in Lemma 5, when

$$\alpha t \geq 1 \text{ i.e. } \epsilon \geq 1/(4ct^2), \tag{i}$$

we have $\tilde{u} \leq \bar{\varphi}_\epsilon = \eta_{c-\epsilon}$. This, combining with the fact that $\tilde{u} \geq \eta_{c+\epsilon}$, implies,

$$\begin{aligned}
&\tilde{u}(t,s) = u(x_1,x_2) \geq \max\{x_1,x_2,0\} = g(x_1,x_2) && \text{if } |x_1-x_2| \geq \frac{1}{2c};\\
&\tilde{u}(t,s) = u(x_1,x_2) \geq \frac{c|x_1-x_2|^2}{2} + \frac{1}{8c} + \frac{x_1+x_2}{2} > g(x_1,x_2) && \text{if } |x_1-x_2| < \frac{1}{2c};\\
&\tilde{u}(t,s) = u(x_1,x_2) \leq g(x_1,x_2) && \text{if } |x_1-x_2| \geq \frac{1}{2(c-\epsilon)}.
\end{aligned}$$

Thus,

$$\begin{aligned}
u(x_1,x_2) &> g(x_1,x_2) \text{ if } |x_1-x_2| < \frac{1}{2c},\\
u(x_1,x_2) &= g(x_1,x_2) \text{ if } |x_1-x_2| \geq \frac{1}{2(c-\epsilon)}.
\end{aligned}$$

We see that free boundary is between $|x_1-x_2| \in (\frac{1}{2c}, \frac{1}{2(c-\epsilon)})$ once $t = \frac{x_1+x_2}{\sqrt{2}}$ satisfying (i). Now take $\epsilon = \frac{1}{4ct^2}$ and to have $\epsilon < c$, we require $t \geq \frac{1}{2c}$. Finally, we conclude that,

$$\begin{aligned}
d_{FB}(t) &\leq \left(\frac{1}{2(c-\epsilon)} - \frac{1}{2c}\right)/\sqrt{2}\\
&= \frac{\epsilon}{2\sqrt{2}c(c-\epsilon)}\\
&= \frac{\epsilon}{2\sqrt{2}c^2} + O\left(\frac{\epsilon^2}{c^3}\right) = \frac{1}{8\sqrt{2}c^3t^2} + O\left(\frac{1}{c^5t^4}\right).
\end{aligned}$$

■

Proof of Proposition 5:

Proof: We first prove the following technical lemma.

Lemma A1: *There exists a universal constant C such that for all $d \geq 2$*

$$\int_0^1 \left|\frac{d}{dR}e^{-1/(1-R^2)}\right| R^{d-1}dR \leq C(d-1)\int_0^1 e^{-1/(1-R^2)}R^{d-1}dR.$$

Proof: Denote

$$J_d := \int_0^1 e^{-1/(1-R^2)}R^d dR. \tag{ii}$$

Integration by parts gives

$$\int_0^1 \left| \frac{d}{dR} e^{-1/(1-R^2)} \right| R^{d-1} dR = \int_0^1 \left(-\frac{d}{dR} e^{-1/(1-R^2)} \right) R^{d-1} dR$$
$$= (d-1) \int_0^1 e^{-1/(1-R^2)} R^{d-2} dR = (d-1) J_{d-2}.$$

By the Cauchy-Schwarz inequality, we have

$$J_{d-2}^2 \leq J_{d-1} J_{d-3}.$$

Thus,

$$\int_0^1 \left| \frac{d}{dR} e^{-1/(1-R^2)} \right| R^{d-1} dR / \left((d-1) \int_0^1 e^{-1/(1-R^2)} R^{d-1} dR \right)$$
$$= J_{d-2}/J_{d-1} \leq J_{d-3}/J_{d-2} \leq ... \leq J_2/J_1 = C.$$

■

Now, we prove the main proposition. From previous arguments, we know that g is a subsolution and $u \geq g$. We are going to construct a supersolution through g and it leads to an estimate of $\Gamma(u)$ from above.

Consider a symmetric modifier

$$\varphi(x) = \mu(|x|)/I_d$$

such that

$$\mu(R) = \begin{cases} e^{-1/(1-R^2)} & \text{if } R \leq 1, \\ 0 & \text{if } R > 1. \end{cases}$$

The numerical constant I_d ensures normalization, i.e.,

$$I_d = \int_{\mathbb{R}^d} \varphi(x) dx = A_d J_{d-1}$$

where A_d is the surface area of a unit d-dimensional ball and J_{d-1} is given in (ii).

Set $\varphi_r := r^d \varphi(rx)$ and then

$$
\begin{aligned}
& supp\{\varphi_r\} \subset \{|x| \le 1/r\}, \\
& \int_{\mathbb{R}^d} |\nabla \varphi_r| dx = r \int_{\mathcal{B}_1} |\nabla \varphi| dx \\
& = \frac{r}{I_d} \iint_{\mathcal{B}_1} |\nabla \mu(R)| R^{d-1} dR d\omega = \frac{r A_d}{I_d} \int_0^1 |\mu'(R)| R^{d-1} dR,
\end{aligned}
\tag{iii}
$$

where ∇ is the gradient operator. According to Lemma A1,

$$\int_{\mathbb{R}^d} |\nabla \varphi_r| dx \le C(d-1) r A_d J_{d-1} / I_d = C(d-1) r.$$

We claim that

$$\Phi_r := \varphi_r * g = \int_{\mathbb{R}^d} \varphi_r(x-y) g(y) dy$$

is a supersolution for some r small enough. Let us check the following two conditions,

$$\Delta \Phi_r \le 2c, \text{and} \Phi_r \ge g.$$

Since (by symmetry) $\varphi_r * x_i = x_i$ and $g = \max\{x_1, ..., x_d, 0\}$, we have

$$\Phi_r = \varphi_r * g \ge g.$$

Next we compute

$$
\begin{aligned}
|\Delta \Phi_r| &= |\nabla_x \int_{\mathbb{R}^d} (\nabla \varphi_r)(x-y) g(y) dy| \\
&= \left| \nabla_x \int_{\mathbb{R}^d} (\nabla \varphi_r)(y) g(x-y) dy \right| \\
&\le \int_{\mathbb{R}^d} |\nabla \varphi_r|(y) |\nabla g|(x-y) dy.
\end{aligned}
$$

By the fact $|\nabla g| \le 1$ and (iii), we obtain

$$|\Delta \Phi_r| \le C(d-1) r.$$

Thus for some universal $\gamma > 1$, we have $|\Delta \Phi_r| \le 2c$ if $r \le c/(\gamma d)$. In all we conclude that with this choice of r, Φ_r is a supersolution and $u \le \Phi_r$.

Fix any $x^0 \in A(\gamma d/c)$. By definition,

$$g(x) = x_k \text{ for some } k \text{ for all } x = (x_1, ..., x_d) \in \mathcal{B}_{\gamma d/c}(x^0)$$

and therefore $\Phi_r = g * \phi_r = x_k$. Hence in $A(\gamma d/c)$, we have $u \le \Phi_r = g$. Since $u \ge g$, we conclude that $u = g$ for $x \in A(\gamma d/c)$. ∎

Proof of Lemma 6:

Proof: Let us sketch the proof below. We are going to use the following cylindrical coordinates: for each $x \in \mathbb{R}^d$, write

$$x = t\tau_d + \sum_{j=1}^{d} s_j e_j,$$

where $\sum_{j=1}^{d} s_j e_j \in H_{\tau_d}$. Then $\omega := w_d - \frac{t}{\sqrt{d}}$ solves

$$\min\left\{-\frac{1}{2}\Delta\omega + c, \omega - \left(\rho - \frac{t}{\sqrt{d}}\right)\right\} = 0 \text{ in } \mathbb{R}^d. \tag{iv}$$

Notice that shifts in the τ_d direction preserve the value of $(\rho - \frac{t}{\sqrt{d}})$. Therefore by uniqueness of solutions to (iv), the shifts also preserve ω i.e. $\omega(x) = \omega(x + s\tau_d)$ for all $s \in \mathbb{R}$.

Now we consider the free boundary property of w_d. Again, for any $x \in \mathbb{R}^d$, write $x = t\tau_k + y$ with $y \in H_{\tau_d}$ and $t(x) = \sum_{j=1}^{d} x_j/\sqrt{d}$. From the above, $w_d(y) = \rho(y)$ if and only if

$$w_d(y) + \frac{t}{\sqrt{d}} = \rho(y) + \frac{t}{\sqrt{d}},$$

if and only if

$$w(x) = \max\{y_1, y_2, ..., y_d\} + \frac{t}{\sqrt{d}} = \max\{x_1, x_2, ..., x_d\} = \rho(x).$$

We used $x = y + t\tau_d$ in the second equality. Therefore $\Gamma(w_d)$ equals $\{\Gamma(w_d) \cap H_{\tau_d}\} \times \mathbb{R}\tau_d$. ∎

Proof of Lemma 7:

Proof: First, we want to give an upper bound of $u - g$ on H_{τ_d}. From the proof of Proposition 5, $\Phi_{1/r} = \varphi_{1/r} * g \ge u$ where $\varphi_{1/r}$ is a modifier supported in $\mathcal{B}_r$ with $r = \frac{\gamma d}{c}$. Because $|\nabla g| \le 1$ and $\varphi_{1/r} * g(x)$ can be viewed as a weighted average of g in $\mathcal{B}_r(x)$, we have

$$|\Phi_{1/r} - g| \le r.$$

In all, we find for $x \in H_{\tau_d}$

$$u - g \leq \Phi_{1/r} - g \leq r. \tag{v}$$

Second, let us construct a supersolution to (9). For $\epsilon \ll \min\{c, 1/d\}$, set

$$w_d^\epsilon := \frac{1}{1-\epsilon} w_d((1-\epsilon)x),$$

which then solves

$$\min\left(-\frac{1}{2}\Delta w_d^\epsilon + (1-\epsilon)c, w_d^\epsilon - \rho\right) = 0.$$

Next define a $\mathcal{C}^1$ function

$$\bar{\varphi} := rh(\alpha t) + w_d^\epsilon,$$

where $h(t) = (\max\{1-t, 0\})^2$ and $\alpha = \alpha(\epsilon)$ are to be determined.

In the third step, we want to show that $\bar{\varphi}$ is indeed a supersolution in the half hyperplane $\mathcal{D} := \left\{\frac{\sum_{j=1}^d x_j}{\sqrt{d}} = t > 0\right\}$. Since $\rho = g$ in $\mathcal{D}$, we have $\bar{\varphi} \geq g$ in $\mathcal{D}$. On the boundary $\partial\mathcal{D} = H_{\tau_d}$, it follows from (v) that

$$\bar{\varphi} = rh(0) + w_d^\epsilon \geq r + g \geq u.$$

Also by direct computation,

$$\Delta\bar{\varphi} = \Delta(rh) + \Delta w_d^\epsilon \leq 2r\alpha^2 + 2(1-\epsilon)c.$$

To make $\bar{\varphi}$ a subsolution, we only need $r\alpha^2 \leq c\epsilon$ which is equivalent to $\alpha \leq c\sqrt{\frac{\epsilon}{\gamma d}}$. Finally we can conclude that by comparison, $\bar{\varphi} \geq u$ in $\mathcal{D}$.

When $t \geq \frac{1}{\alpha} = \frac{1}{c}\sqrt{\frac{\gamma d}{\epsilon}}$, we have $\bar{\varphi} = w_d^\epsilon \geq u$. Hence we know $w_d^\epsilon \geq u \geq w_d$. Since

$$w_d^\epsilon(x) = \frac{1}{1-\epsilon} w_d((1-\epsilon)x),$$

then $w_d(x) = g(x)$ implies $w_d^\epsilon(x) = g(x)$. Therefore the free boundary of u ($\Gamma(u)$) lies between $\Gamma(w_d)$ and $\Gamma(w_d^\epsilon)$ when t is large. By Lemma 6, it is sufficient to compare $\Gamma(w_d)$ and $\Gamma(w_d^\epsilon)$ on H_{τ_d}.

We consider a R-neighbourhood of the origin in H_{τ_d} (seeing from Proposition 5, we may pick $R = \frac{\gamma d}{c}$). Again by definition of w_d^ϵ, inside $\mathcal{B}_R(t\tau_d) \cap H_{\tau_d}$, the distance between $\Gamma(w_d)$ and $\Gamma(w_d^\epsilon)$ is bounded by $R\epsilon$. Till here, we conclude that the distance between $\Gamma(u)$ and $\Gamma(w_d)$ is bounded by $R\epsilon$ in the set $A_1(R) \cap \{t \geq \frac{1}{c}\sqrt{\frac{\gamma d}{\epsilon}}\}$. ■

References

Sigurd Assing, Saul Jacka, and Adriana Ocejo. Monotonicity of the value function for a two-dimensional optimal stopping problem. *The Annals of Applied Probability*, 24(4):1554–1584, 2014.

Fernando Branco, Monic Sun, and J. Miguel Villas-Boas. Optimal search for product information. *Management Science*, 58(11):2037–2056, 2012.

M. Broadie and J. Detemple. The valuation of American options on multiple assets. *Mathematical Finance*, 7(3):241–286, 1997.

Luis Caffarelli. The obstacle problem revisited. *Journal of Fourier Analysis and Applications*, 4 (4-5):383–402, 1998.

Yeon-Koo Che and Konrad Mierendorff. Optimal sequential decision with limited attention. Working paper, Columbia University, 2016.

Michael Crandall and Pierre-Louis Lions. Viscosity solutions of hamilton-jacobi equations. *Transactions of the American mathematical society*, 277(1):1–42, 1983.

Michael Crandall, Hitoshi Ishii, and Pierre-Louis Lions. User's guide to viscosity solutions of second order partial differential equations. *Bulletin of the American Mathematical Society*, 27(1):1–67, 1992.

A. Dvoretzky, F. Kiefer, and J. Wolfowitz. Sequential decision problems for processes with continuous time parameter. *Annals of Mathematical Statistics*, 24(2):254–264, 1953.

Jens Frehse. On the regularity of the solution of a second order variational inequality. *Boll. Un. Mat. Ital.(4)*, 6:312–315, 1972.

Drew Fudenberg, Philipp Strack, and Tomasz Strzalecki. Speed, accuracy, and the optimal timing of choices. *American Economic Review*, 108:3651–3684, 2018.

Xin Guo and Mihail Zervos. π options. *Stochastic Processes and their Applications*, 120(7):1033–1059, 2010.

Benjamin Hébert and Michael Woodford. Rational inattention with sequential information sampling. Working paper, Stanford University and Columbia University, 2017.

Hitoshi Ishii. Perron's method for Hamilton-Jacobi equations. *Duke Mathematical Journal*, 55(2): 369–384, 1987.

Hitoshi Ishii. On uniqueness and existence of viscosity solutions of fully nonlinear second-order elliptic pde's. *Communications on Pure and Applied Mathematics*, 42(1):15–45, 1989.

H. Johnson. Options on the maximum or the minimum of several assets. *Journal of Financial Quantitative Analysis*, 22(3):277–283, 1987.

Ioannis Karatzas and Steven Shreve. *Brownian motion and stochastic calculus*, volume 113. Springer-Verlag, New York, second edition, 1991.

T.Tony Ke, Zuo-Jun Max Shen, and J.Miguel Villas-Boas. Search for information on multiple products. *Management Science*, 62(12):3576–3603, 2016.

David Kinderlehrer and Guido Stampacchia. *An introduction to variational inequalities and their applications*, volume 31. SIAM, 1980.

A.S. Mikhalevich. A Bayes test of two hypotheses concerning the mean of a normal process. *Visnyk Kyivskogo Universytetu*, 1:101–104, 1958.

Giuseppe Moscarini and Lones Smith. The optimal level of experimentation. *Econometrica*, 69(6): 1629–1644, 2001.

Goran Peskir and Albert Shiryaev. *Optimal stopping and free-boundary problems*. Springer, 2006.

Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion*, volume 293. Springer-Verlag, Berlin, third edition, 1999.

Kevin Roberts and Martin L. Weitzman. Funding criteria for research, development, and exploration projects. *Econometrica*, 49(5):1261–1288, 1981.

M. Rubinstein. Somewhere over the rainbow. *Risk*, 4:63–66, 1991.

A.N. Shiryaev. Two problems of sequential analysis. *Cybernetics*, 3:63–69, 1967.

R.M. Stulz. Options on the minimum or the maximum of two risky assets: analysis and applications. *Journal of Financial Economics*, 10(2):161–185, 1982.

Abraham Wald. Sequential tests of statistical hypotheses. *Annals of Mathematical Statistics*, 16 (2):117–186, 1945.