Research Report No. 1, 2025, *Probab. Statist. Group Manchester* (28 pp)

The Dubins Constants for Walsh's Spider Process

Goran Peskir

A long-standing open problem of L. E. Dubins seeks to determine the maximal expected range of Walsh's spider process on n edges per root of the expected stopping time. The solution is known for $n=1$ (1988) and $n=2$ (2009). In this paper we present the solution for $n \geq 3$.

1. Introduction

A long-standing open problem[1] of L. E. Dubins (cf. [6, p. 401]) seeks to determine the maximal expected range of Walsh's spider process on n edges per root of the expected stopping time. The solution is known for $n=1$ ([7], [14], [17]) and $n=2$ ([6]). The purpose of this paper is to present the solution for $n \geq 3$.

More specifically, let X be a symmetric Walsh spider process on edges $E_1, \ldots, E_n$ starting at 0 (the origin) where $n \geq 1$ is given and fixed (see Section 2 for fuller details). Let S^i denote the running maximum of X on the edge E_i for $1 \leq i \leq n$. The *Dubins problem* seeks to determine the *smallest* constant D_n such that

$$\mathsf{E}(S^1_\tau + \ldots + S^n_\tau) \leq D_n \sqrt{\mathsf{E}\tau} \tag{1.1}$$

for all stopping times τ of X . Note that the problem includes finding a (non-zero) stopping time at which equality in (1.1) is attained. An application of Lagrange multipliers shows that the Dubins problem (1.1) is equivalent to the optimal stopping problem

$$V = \sup_\tau \mathsf{E}(S^1_\tau + \ldots + S^n_\tau - c\tau) \tag{1.2}$$

where the supremum is taken over all stopping times τ of X such that $\mathsf{E}\tau < \infty$ and $c > 0$ is a given and fixed constant (Lagrange multiplier). Moreover, the Dubins constant D_n equals $2\sqrt{V}$ upon letting $c=1$ with equality in (1.1) being attained at the optimal stopping time in (1.2) (see Section 3 for fuller details).

When $n=1$ the process X is equal to $|B|$ where B is a standard Brownian motion. The solution to (1.2) with $D_1 = \sqrt{2}$ found in [7] was extended to Bessel processes in [8]. This

Mathematics Subject Classification 2020. Primary 60J65, 60G40, 60J60. Secondary 35R35, 49J40, 60E15.

Key words and phrases: Dubins constants, Walsh spider process, Brownian motion, optimal stopping, free-boundary problem, maximum process, smooth fit, the maximality principle, normal reflection, Lagrangian, Kirchhoff condition, exit probability, mean exit time, diffusion process on a graph, triple branching point, edge with (zero) leak, separation principle, pruning procedure, nonlinear hyperbolic integral/differential equation.

[1]Personal communication by L. A. Shepp in January 2006.

subsequently led to the identification of the *maximality principle* in [20] that applies to general one-dimensional diffusion processes (see [22, Section 13]). When $n = 2$ the process X is equal to B itself. The solution to (1.2) with $D_2 = \sqrt{3}$ was found in [6]. A different derivation based on the maximality principle was given in [21]. For $n = 3$ the problem (1.2) was studied in [9] and [10]. A modified version of the problem (1.2) where the sum of all running maxima is replaced by the sum of the two largest maxima was solved in [4].

In this paper we present a unifying approach that solves the problem (1.2) for all $n \geq 1$. Theorem 1 states the main results. Its proof consists of ten parts. There are also ten figures that help to visualise the rich underlying dynamics and geometry of the solution. The optimal stopping time in (2.1) is realised by means of the *pruning procedure* described in Theorem 1. This leads to the novel concepts of a *separation principle* and *edges with leaks*. The optimal stopping problem (1.2) is reformulated as a free-boundary problem where *Kirchhoff's condition* at the origin for $n \geq 2$ replaces *Neumann's condition* when $n = 1$. This reflects the subtle nature of the origin as a triple branching point for the state space of X when $n = 3$ (cf. [24]).

The free-boundary problem does not have a unique solution. We show that the maximality principle closes the system and makes the solution arising from the pruning procedure unique. This is achieved using *smooth fit* (at the optimal stopping boundary) and *normal reflection* (at diagonal points) combined with exit probabilities and mean exit times (calculated in Section 2) by showing that the optimal stopping boundary f used in the pruning procedure can be characterised as a unique continuous solution to a nonlinear integral equation that needs to be solved recursively over domains with curved boundaries determined (alongside the boundary values) by the solution f from the previous step. The nonlinear integral equation is equivalent to a nonlinear *hyperbolic* PDE and the appearance of hyperbolic equations in the probabilistic setting is both unusual and interesting. Combining the specified f with the optimal stopping boundary g on edges with leaks, which by the maximality principle happens to be a linear function, we obtain a full description of the pruning procedure that is sufficient to realise the optimal stopping. This, however, is still insufficient to calculate the value function at the origin as needed for finding the Dubins constants.

We resolve this issue in a novel way by adding an *extra edge* with *zero leak* to the state space of X. This yields the optimal stopping boundary $\bar{f}$ taking values in the added edge. We show that $\bar{f}$ can be characterised as the unique continuous solution to a linear (for its square) integral equation over the domain with a curved boundary determined (alongside the boundary values) by the final solution f to the nonlinear integral equation described above. We find that $V = c\bar{f}^2(0)$ so that $D_n = 2\sqrt{V} = 2\bar{f}(0)$ for $c = 1$. This settles the Dubins problem for all $n \geq 1$.

In Examples 1 and 2 we show how the general solution yields the known values $D_1 = \sqrt{2}$ and $D_2 = \sqrt{3}$ respectively. In Example 3 we briefly describe the *quadrature method* that yields the numerical approximation $D_3 = 1.91\ldots$ where the first two decimal places are exact. The absence of an explicit expression for D_3 is attributed to the fact that the origin is a triple branching point when $n = 3$ so that f admits no closed-form expression. The same conclusion extends to all $n \geq 3$. The quadrature method is applicable in any dimension although numerical approximations become increasingly computationally expensive (especially for $\bar{f}$) as the dimension n increases. Developing faster and more efficient methods/algorithms for solving the *nested* integral equations in higher dimensions to approximate the remaining Dubins constants $D_4, D_5, \ldots$ presents a challenging topic for further study.

2. The Walsh spider process

In this section we present a few basic facts about the Walsh spider process that are helpful in tackling the Dubins problem to be discussed in the next section.

1. *Definition and origin.* The *Walsh spider process* (often called *Walsh's Brownian motion*) is a diffusion process $X = (X_t)_{t \ge 0}$ on a star-shaped graph consisting of finitely or infinitely many half-lines $[0, \infty)$ (called edges, rays, ribs, legs) that are connected at a single vertex 0 (the origin). The motion of X on each edge coincides with a standard Brownian motion (Wiener process) when away from the origin. After hitting the origin from within an edge, the process X reflects instantaneously into another (possibly the same) edge, independently from the past motion, according to a given probability measure μ on $[0, 2\pi)$ representing the angle of the new edge in the plane to be taken by X. The motion of X then continues in exactly the same manner indefinitely (see Figure 1).

The Walsh spider process was introduced by Walsh in [25] as a generalisation of the Itô-McKean construction of a skew Brownian motion (when the angular measure μ is concentrated on two points). We refer to [2, Section 2] and the references therein for five different constructions of the Walsh spider process including the original one described by Walsh that we adopt. For constructions of more general Markov processes of this/similar kind see [11]. In addition to being *strong Markov* (and *continuous*), the Walsh spider process is known to be *Feller* in general, as well as *strong Feller* when the angular measure is concentrated on finitely many points (as in the Dubins problem). For a study of harmonic functions of the Walsh spider process see [12]. For applications of the Walsh spider process in queueing theory see [1].

2. *Radial and angular part.* In terms of polar coordinates in the plane, the Walsh spider process $X = (X_t)_{t \ge 0}$ can be written as

$$X_t = (R_t, \Phi_t) \tag{2.1}$$

for $t \ge 0$, where the *radial part* $R = (R_t)_{t \ge 0}$ is an instantaneously reflecting (at zero) standard Brownian motion in $[0, \infty)$, and the *angular part* $\Phi = (\Phi_t)_{t \ge 0}$ takes values in $[0, 2\pi)$ as described above. The local time of R at 0 is given by

$$\ell_t^0(R) = \lim_{\varepsilon \downarrow 0} \frac{1}{2\varepsilon} \int_0^t I(0 \le R_s \le \varepsilon)\, ds \tag{2.2}$$

for $t \ge 0$, and the process $B_t := R_t - R_0 - \ell_t^0(R)$ defines a standard Brownian motion for $t \ge 0$ with respect to the natural filtration $(\mathcal{F}_t^X)_{t \ge 0}$ of X (cf. [2, Lemma 2.2]). Hence we see that the radial part R of X solves

$$dR_t = dB_t + d\ell_t^0(R)\,. \tag{2.3}$$

It is known that $(\mathcal{F}_t^X)_{t \ge 0}$ is *not* a Brownian filtration unless the angular measure μ is concentrated on one or two points (cf. [24]) implying that extra randomness is needed essentially (in addition to B driving R) to create angles representing three or more edges to be occupied by X after R hits 0. More specifically, the angular part $\Phi = (\Phi_t)_{t \ge 0}$ is given by

$$\Phi_t = U_{n(t)} \tag{2.4}$$

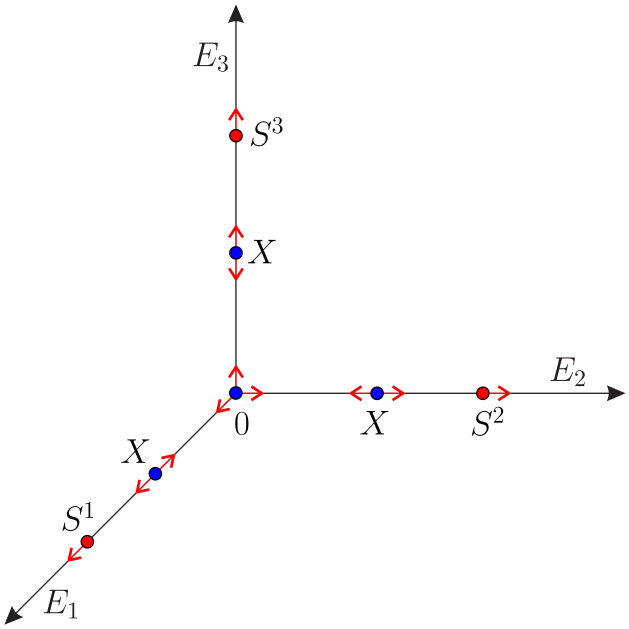

Figure 1. Motion of the Walsh spider process $X = (X_t)_{t \ge 0}$ on three edges: (i) X is a standard Brownian motion in each edge when away from the origin; (ii) X instantaneously reflects into another (or the same) edge when at the origin. The origin is a (triple) branching point of the one-dimensional (Euclidean) topological space $E := E_1 \sqcup E_2 \sqcup E_3$ where $E_i = [0, \infty)$ for $1 \le i \le 3$. The running maxima S^1, S^2, S^3 of X are being pushed infinitesimally upwards when (and only then) the process X reaches them respectively.

where $U_1, U_2, \ldots$ are independent (from B as well) μ-distributed random variables taking values in $[0, 2\pi)$ and $n(t)$ is the number of the excursion of B straddling t for $t \ge 0$. Note that Φ is flat when R is away from 0 and changes only when R returns to 0. Note also that the total time spent by R at 0 equals zero with probability one.

3. *Itô's formula.* The Walsh spider process $X = (X_t)_{t \ge 0}$ is a continuous semimartingale and the Itô formula reads

$$F(X_t) = F(X_0) + \int_0^t F'(X_s)\, I(X_s \ne 0)\, dB_s + \frac{1}{2} \int_0^t F''(X_s)\, I(X_s \ne 0)\, ds + \left(\int_0^{2\pi} \frac{\partial F}{\partial r}(0+, \varphi)\, \mu(d\varphi) \right) \ell_t^0(X) \tag{2.5}$$

for $t \ge 0$, where $F = F(r, \varphi)$ is a real-valued C^2 function for $r \in [0, \infty)$ and $\varphi \in [0, 2\pi)$, and the local time $\ell_t^0(X)$ of X at 0 (the origin) is defined in the same way (with X in place of R) as in (2.2) above (cf. [13], [16], [18]).

4. *Uniformly-atomic angular measure.* In the Dubins problem one assumes that the angular measure μ of the Walsh spider process $X = (X_t)_{t \ge 0}$ is uniformly concentrated on n points in $[0, 2\pi)$ for $n \ge 1$ given and fixed. In this case we refer to X as a *symmetric* Walsh spider process on n edges. The exact specification of the n points is irrelevant and one could take

them to be $(i-1)(2\pi/n)$ for $1 \le i \le n$. The state space of X then equals

$$E = \coprod_{i=1}^{n} E_i := \bigcup_{i=1}^{n} \{(x,i) \,|\, x \in E_i\} \tag{2.6}$$

where $E_i = [0,\infty)$ for $1 \le i \le n$. The graph E is viewed as a one-dimensional (Euclidean) topological space having a branching point at the origin. Points (x,i) in E will be denoted by x_i for $1 \le i \le n$. Thus x_i is a non-negative real number belonging to the edge E_i for $1 \le i \le n$. To indicate that X_t belongs to E_i we will write X_t^i for $t \ge 0$ with $1 \le i \le n$. The restriction of a function $F : E \to I\!R$ to E_i will be denoted by F_i meaning that the function $F_i : E_i \to I\!R$ is defined by $F_i(x_i) = F(x_i)$ for $x_i \in E_i$ and $1 \le i \le n$. In the rest of the paper we only consider a Walsh spider process X on n edges having the state space given by (2.6) above. For more general Walsh spider processes of this kind see [19]. Note that X has the same law as $|B|$ when $n=1$ and B when $n=2$ respectively.

5. *Infinitesimal characteristics.* If $X = (X_t)_{t \ge 0}$ is the Walsh spider process on n edges with angular probabilities $p_1, \ldots, p_n$, then its infinitesimal characteristics are specified by

$$I\!L_X = \frac{1}{2}\frac{\partial^2}{\partial x} \quad \text{for} \;\; x \neq 0 \tag{2.7}$$

$$\sum_{i=1}^{n} p_i \frac{\partial}{\partial x_i} = 0 \;\; \text{at} \;\; x = 0 \;\; \text{(Kirchhoff)}. \tag{2.8}$$

Note that the Kirchhoff condition (2.8) is consistent with the presence of the final term in Itô's formula (2.5) above when the angular measure is uniformly atomic. If $n=1$ then the Kirchhoff condition (2.8) reduces to the Neumann condition (of normal reflection). A precise meaning to (2.8) can be given in terms of the domain of the infinitesimal generator $I\!L_X$, however, we will not need this in the sequel and thus omit fuller details.

6. *Exit probabilities and mean exit times.* Given $a_1 \in E_1, \ldots, a_n \in E_n$ consider the first entry time of X to the set $A = \{a_1, \ldots, a_n\}$ given by

$$\tau_A = \inf\{\, t \ge 0 \mid X_t \in A \,\} \tag{2.9}$$

under the probability measure P_{x_i} such that $\mathsf{P}_{x_i}(X_0 = x_i) = 1$ for $0 \le x_i \le a_i$ with $1 \le i \le n$. Solving the boundary value problem based on (2.7)+(2.8) with $I\!L_X = 0$ for $x \neq 0$ one can verify that

$$\mathsf{P}_{x_i}(X_{\tau_A} \!=\! a_i) = \frac{1}{(a_i/p_i)\sum_{k=1}^{n} p_k/a_k}\Big(1 - \frac{x_i}{a_i}\Big) + \frac{x_i}{a_i} \quad (1 \le i \le n) \tag{2.10}$$

$$\mathsf{P}_{x_i}(X_{\tau_A} \!=\! a_j) = \frac{1}{(a_j/p_j)\sum_{k=1}^{n} p_k/a_k}\Big(1 - \frac{x_i}{a_i}\Big) \quad (1 \le i \neq j \le n) \tag{2.11}$$

Similarly, solving the boundary value problem based on (2.7)+(2.8) with $I\!L_X = -1$ for $x \neq 0$ one can verify that

$$\mathsf{E}_{x_i}(\tau_A) = \frac{\sum_{k=1}^{n} p_k a_k}{\sum_{k=1}^{n} p_k/a_k}\Big(1 - \frac{x_i}{a_i}\Big) + a_i x_i - x_i^2 \quad (1 \le i \le n)\,. \tag{2.12}$$

Denoting any of the functions from (2.10)-(2.12) by F , note that in these (and similar other) calculations in addition to the Kirchhoff condition (2.8) one also needs to employ the *continuity-at-the-origin* condition that reads

$$F_1(0+) = F_2(0+) = \ldots = F_n(0+) \tag{2.13}$$

where F_i denotes the restriction of F to E_i for $1 \le i \le n$.

3. The Dubins problem

In this section we present a few basic facts about the Dubins problem (cf. [6, p. 401]). These considerations will be continued in the next section.

1. *Initial formulation.* Let $X = (X_t)_{t \ge 0}$ be a *symmetric* Walsh spider process on n edges starting at 0 (the origin) where $n \ge 1$ is given and fixed. Consider the (running) maximum process $S^i = (S^i_t)_{t \ge 0}$ of X on the edge E_i defined by

$$S^i_t = \sup_{s \in [0,t] : X_s \in E_i} X^i_s \tag{3.1}$$

for $1 \le i \le n$. The *Dubins problem* seeks to determine the *smallest* constant D_n such that

$$\mathsf{E}(S^1_\tau + \ldots + S^n_\tau) \le D_n \sqrt{\mathsf{E}\tau} \tag{3.2}$$

for all stopping times τ of X . Note that the problem includes finding a (non-zero) stopping time at which equality in (3.2) is attained. The solution is known for $n = 1$ (cf. [7], [14], [17]) and $n = 2$ (cf. [6]) with the constants $D_1 = \sqrt{2}$ and $D_2 = \sqrt{3}$. Our main focus in this paper will therefore be on the case $n \ge 3$.

2. *Related formulations.* The Dubins problem can be reformulated by asking to solve

$$\sup_{\tau : \mathsf{E}\tau > 0} \frac{\mathsf{E}(S^1_\tau + \ldots + S^n_\tau)}{\sqrt{\mathsf{E}\tau}} . \tag{3.3}$$

Further two closely-related *constrained* problems are asking to solve

$$\sup_{\tau : \mathsf{E}\tau \le \alpha} \mathsf{E}(S^1_\tau + \ldots + S^n_\tau) \tag{3.4}$$

$$\inf_{\tau : \mathsf{E}(S^1_\tau + \ldots + S^n_\tau) \ge \beta} \mathsf{E}\tau \tag{3.5}$$

where $\alpha > 0$ and $\beta > 0$ are given and fixed constants. It turns out that the three problems (3.3)-(3.5) can be tackled simultaneously using *Lagrange multipliers* as follows. We will not discuss the constrained problems (3.4) and (3.5) in this paper and will only focus on (3.3).

3. *Lagrangian.* An application of Lagrange multipliers in either of the two constrained problems (3.4) and (3.5) implies that the Lagrange function (Lagrangian) is given by

$$V = \sup_\tau \mathsf{E}(S^1_\tau + \ldots + S^n_\tau - c\tau) \tag{3.6}$$

where the supremum is taken over all stopping times τ of X such that $\mathsf{E}\tau < \infty$ and $c > 0$ is a given and fixed constant (Lagrange multiplier). It is well known that solving (3.6) for all $c>0$ will also solve (3.4) and (3.5) as well. Moreover, due to the scaling property of X (inherited from B), it turns out that solving (3.6) for all $c>0$ (in fact $c=1$ is sufficient) will also solve (3.3). This can be seen as follows.

4. *Equivalence*. Denoting the value V in (3.6) by $V(0;c)$ to indicate its dependence on the initial point 0 and the Lagrange multiplier c, using the scaling property of X combined with the fact that the supremum in (3.6) is attained at the first entry time of the (Markov) process $(X,S^1,\dots,S^n)$ into a (closed) set (cf. [22, Corollary 2.9]), it is easily verified that

$$V(0;c)=\frac{1}{c}V(0;1) \tag{3.7}$$

for all $c>0$. From (3.6) and (3.7) we see that

$$\mathsf{E}(S^1_\tau+\ldots+S^n_\tau)\le V(0;c)+c\,\mathsf{E}\tau=\frac{1}{c}V(0;1)+c\,\mathsf{E}\tau \tag{3.8}$$

for every $c>0$ and every stopping time τ of X with $\mathsf{E}\tau<\infty$ given and fixed. Minimising the right-hand side in (3.8) over $c>0$, one finds that the unique minimum is attained at $c=\sqrt{V(0;1)/\mathsf{E}\tau}$. Evaluating the right-hand side in (3.8) at this c then gives

$$\mathsf{E}(S^1_\tau+\ldots+S^n_\tau)\le 2\sqrt{V(0;1)}\sqrt{\mathsf{E}\tau} \tag{3.9}$$

for all stopping times τ of X with $\mathsf{E}\tau<\infty$. Moreover, letting τ_1 denote the optimal stopping time in (3.6) with $c=1$, we claim that equality in (3.9) is attained at τ_1. For this, note that we have

$$\begin{aligned}\mathsf{E}(S^1_{\tau_1}+\ldots+S^n_{\tau_1})&=\mathsf{E}(S^1_{\tau_1}+\ldots+S^n_{\tau_1}-\tau_1)+\mathsf{E}\tau_1=V(0;1)+\mathsf{E}\tau_1\\&=\left(\frac{V(0;1)}{\sqrt{\mathsf{E}\tau_1}}+\sqrt{\mathsf{E}\tau_1}\right)\sqrt{\mathsf{E}\tau_1}\le 2\sqrt{V(0;1)}\sqrt{\mathsf{E}\tau_1}\end{aligned} \tag{3.10}$$

where the inequality follows from (3.9) above. From (3.10) we see that

$$\frac{V(0;1)}{\sqrt{\mathsf{E}\tau_1}}+\sqrt{\mathsf{E}\tau_1}\le 2\sqrt{V(0;1)} \tag{3.11}$$

or equivalently $(\sqrt{V(0;1)}-\sqrt{\mathsf{E}\tau_1})^2\le 0$ implying that $V(0;1)=\mathsf{E}\tau_1$. Making use of this identity in (3.10) above we obtain

$$\mathsf{E}(S^1_{\tau_1}+\ldots+S^n_{\tau_1})=V(0;1)+\mathsf{E}\tau_1=2V(0;1)=2\sqrt{V(0;1)}\sqrt{\mathsf{E}\tau_1} \tag{3.12}$$

showing that equality in (3.9) is attained at τ_1 as claimed. In turn this also shows that solving (3.6) for all $c>0$ (in fact $c=1$ is sufficient) will also solve (3.3) as claimed.

5. *Conclusion*. The previous considerations show that the Dubins problem is equivalent to the optimal stopping problem (3.6). Moreover, the Dubins constant D_n from (3.2) equals

$$D_n=\sup_{\tau:\,\mathsf{E}\tau>0}\frac{\mathsf{E}(S^1_\tau+\ldots+S^n_\tau)}{\sqrt{\mathsf{E}\tau}}=2\sqrt{V(0;1)} \tag{3.13}$$

with the supremum in (3.13), or equivalently equality in (3.2), being attained at the optimal stopping time in (3.6) with $c=1$. For this reason we focus on the optimal stopping problem (3.6) in the sequel.

4. Solution to the Dubins problem

In this section we present the solution to the Dubins problem discussed in the previous section. Recalling that the Dubins problem is equivalent to the optimal stopping problem (3.6) we concentrate on the latter problem.

Recalling (3.1) and setting $S = (S^1, \ldots, S^n)$ with $I = \{1, \ldots, n\}$, we will make use of the following notation $\bar{S}^{i_1} := (S^1, \ldots, S^{i_1-1}, S^{i_1+1}, \ldots, S^n)$ for $i_1 \in I$, $\bar{S}^{i_1,i_2} := (S^1, \ldots, S^{i_1-1}, S^{i_1+1}, \ldots, S^{i_2-1}, S^{i_2+1}, \ldots, S^n)$ with $\bar{S}^{i_2,i_1} := \bar{S}^{i_1,i_2}$ for $i_1 < i_2$ in I, and so on. Similarly, for $s = (s_1, \ldots, s_n) \in [0, \infty)^n$ we will write $\bar{s}_{i_1} := (s_1, \ldots, s_{i_1-1}, s_{i_1+1}, \ldots, s_n)$ for $i_1 \in I$, $\bar{s}_{i_1,i_2} := (s_1, \ldots, s_{i_1-1}, s_{i_1+1}, \ldots, s_{i_2-1}, s_{i_2+1}, \ldots, s_n)$ with $\bar{s}_{i_2,i_1} := \bar{s}_{i_1,i_2}$ for $i_1 < i_2$ in I, and so on.

The main results of the paper may be stated as follows.

Theorem 1. *Consider the optimal stopping problem* (3.6) *where $S^1, \ldots, S^n$ are the running maxima of a symmetric Walsh spider process X on n edges starting at 0 with $n \geq 1$ and $c > 0$ given and fixed.*

(I): The optimal stopping time τ_* in (3.6) *is obtained by the following pruning procedure where the optimal stopping boundaries f and g are specified in (III) below.*

Step 1: Consider

$$\tau_1 = \inf \{ t \geq 0 \mid f(\bar{S}_t^{i_1}) \leq g(S_t^{i_1}) \text{ for some } i_1 \in I \}. \tag{4.1}$$

Step 2: If $X_{\tau_1} \in E_{i_1}$ then

$$\tau_* = \inf \{ t \geq \tau_1 \mid X_t^{i_1} \leq g(S_t^{i_1}) \}. \tag{4.2}$$

If $X_{\tau_1} \notin E_{i_1}$ then consider

$$\tau_2 = \inf \{ t \geq \tau_1 \mid X_t^{i_1} \geq f(\bar{S}_t^{i_1}) \text{ or } f(\bar{S}_t^{i_1,i_2}) \leq g(S_t^{i_2}) \text{ for some } i_2 \in I \setminus \{i_1\} \}. \tag{4.3}$$

If $X_{\tau_2}^{i_1} \geq f(\bar{S}_{\tau_2}^{i_1})$ then $\tau_* = \tau_2$ otherwise proceed as after Step 1 above.

Step 3: If $X_{\tau_2} \in E_{i_2}$ then

$$\tau_* = \inf \{ t \geq \tau_2 \mid X_t^{i_2} \leq g(S_t^{i_2}) \}. \tag{4.4}$$

If $X_{\tau_2} \notin E_{i_2}$ then consider

$$\tau_3 = \inf \{ t \geq \tau_2 \mid X_t^{i_1} \vee X_t^{i_2} \geq f(\bar{S}_t^{i_1,i_2}) \text{ or } f(\bar{S}_t^{i_1,i_2,i_3}) \leq g(S_t^{i_3}) \text{ for some } i_3 \in I \setminus \{i_1, i_2\} \}. \tag{4.5}$$

If $X_{\tau_3}^{i_1} \vee X_{\tau_3}^{i_2} \geq f(\bar{S}_{\tau_3}^{i_1,i_2})$ then $\tau_* = \tau_3$ otherwise proceed as after Step 2 above.

Step 4: If $X_{\tau_3} \in E_{i_3}$ then

$$\tau_* = \inf \{ t \geq \tau_3 \mid X_t^{i_3} \leq g(S_t^{i_3}) \}. \tag{4.6}$$

If $X_{\tau_3} \notin E_{i_3}$ then consider

$$\tau_4 = \inf \{ t \geq \tau_3 \mid X_t^{i_1} \vee X_t^{i_2} \vee X_t^{i_3} \geq f(\bar{S}_t^{i_1,i_2,i_3}) \text{ or } f(\bar{S}_t^{i_1,i_2,i_3,i_4}) \leq g(S_t^{i_4}) \text{ for some } i_4 \in I \setminus \{i_1, i_2, i_3\} \}. \tag{4.7}$$

If $X_{\tau_4}^{i_1} \vee X_{\tau_4}^{i_2} \vee X_{\tau_4}^{i_3} \geq f(\bar{S}_{\tau_4}^{i_1,i_2,i_3})$ then $\tau_* = \tau_4$ otherwise proceed recursively as after Step 3 above until the final two edges are reached.

Step n: If $X_{\tau_{n-1}} \in E_{i_{n-1}}$ *then*

$$\tau_* = \inf\{\, t \ge \tau_{n-1} \mid X_t^{i_{n-1}} \le g(S_t^{i_{n-1}}) \,\}. \tag{4.8}$$

If $X_{\tau_{n-1}} \notin E_{i_{n-1}}$ *then*

$$\begin{aligned}\tau_* = \inf\{\, t \ge \tau_{n-1} \mid X_t^{i_1} \vee \ldots \vee X_t^{i_{n-1}} \ge f(\bar{S}_t^{i_1,\ldots,i_{n-1}}) \ \ or \\ X_t^{i_n} \le g(S_t^{i_n}) \,\}. \end{aligned} \tag{4.9}$$

[Note that $X_t^{i_1} \vee X_t^{i_2} \ge f(\bar{S}_t^{i_1,i_2})$ *in* (4.5) *is interpreted as either* $X_t^{i_1} \ge f(\bar{S}_t^{i_1,i_2})$ *holds (i.e.* $X_t \in E_{i_1}$ *and* $X_t \ge f(\bar{S}_t^{i_1,i_2})$ *) or* $X_t^{i_2} \ge f(\bar{S}_t^{i_1,i_2})$ *holds (i.e.* $X_t \in E_{i_2}$ *and* $X_t \ge f(\bar{S}_t^{i_1,i_2})$ *). A similar interpretation applies to* $X_t^{i_1} \vee X_t^{i_2} \vee X_t^{i_3}$ *in* (4.7) *and so on. Note also that* $X_{\tau_{n-1}} \notin E_{i_{n-1}}$ *is equivalent to* $X_{\tau_{n-1}} \in E_{i_n}$ *in Step n and that* $\bar{S}_t^{i_1,\ldots,i_{n-1}}$ *equals* $S_t^{i_n}$ *in* (4.9)*.]*

(II): The value function V *in* (3.6) *and the optimal stopping boundaries* f *and* g *solve the (parabolic) free-boundary problem* (4.17)-(4.24) *below, where the continuation set* C *is explicitly expressible in terms of* f *and* g *using the pruning procedure from (I) above as shown in* (4.25) *below. Explicit formulae for* V *on various parts of* C *(expressed in terms of* f *and* g *) alongside various properties of* f *itself (such as being permutationally invariant) are given in the proof below.*

(III): The optimal stopping boundary f *can be characterised as the unique continuous solution to the nonlinear (hyperbolic) integral equation*

$$\begin{aligned}&\frac{1}{4c}\Big(\frac{1}{s_1^*}+\ldots+\frac{1}{s_k^*}\Big) - c\big(s_1^*+\ldots+s_k^*\big) + \log\Big(\frac{s_1^*\cdot\ldots\cdot s_k^*}{s_1\cdot\ldots\cdot s_k}\Big)\\ &+ c\int_{s_1}^{s_1^*} \frac{f^2(s_1',s_2,\ldots,s_k)}{(s_1')^2}\,ds_1' + \ldots + c\int_{s_k}^{s_k^*} \frac{f^2(s_1,\ldots,s_{k-1},s_k')}{(s_k')^2}\,ds_k'\\ &= c\Big(\frac{1}{s_1}+\ldots+\frac{1}{s_k}\Big) f^2(s_1,\ldots,s_k) + 2c(n-k) f(s_1,\ldots,s_k)\end{aligned} \tag{4.10}$$

for $0 \le s_1 \le s_1^*, \ldots, 0 \le s_k \le s_k^*$ *solved recursively for* $k = 1,\ldots,n-1$ *, where* s_i^* *is the largest value of* s_i *which does not produce a (strict) leak either on the edge* E_i *or another edge, i.e. more specifically, setting* $s := (s_1,\ldots,s_k)$ *we either let*

$$s_i^* = f(\bar{s}_i) + \frac{1}{2c} \quad \text{if } f(\bar{s}_i) \in E_i \tag{4.11}$$

i.e. if $\bar{s}_i$ *belongs to the domain of* f *, or we set* $s_j := \max \bar{s}_i$ *and let* s_i^* *be the largest value of* s_i *in* $[0, 1/2c]$ *for which* $s_j \le f(\bar{s}_j)+1/2c$ *for* $1 \le i \le k$ *. [Note that the domain of* f *is changing as* k *runs from* 1 *to* $n-1$ *so that* f *solving* (4.10) *for* k *is being used to specify the domain boundary of* f *solving* (4.10) *for* k+1 *.] Moreover, we have* $f(s_1^*,\ldots,s_k^*) = f(s_1^*,\ldots,s_{i-1}^*,s_{i+1}^*,\ldots,s_k^*)$ *where* s_i^* *is the largest among* $s_1^*,\ldots,s_k^*$ *for* $1 \le k \le n-1$ *.*

The optimal stopping boundary g *is given explicitly by*

$$g(s) = s - \frac{1}{2c} \tag{4.12}$$

for $s \ge 1/2c$ *.*

(IV): Adding an extra edge with zero leak to the state space E of X yields the optimal stopping boundary $\bar{f}$ (taking values in the added edge) that can be characterised as the unique continuous solution to the linear (for its square) integral equation

$$\begin{aligned}
(4.13)\qquad & \frac{1}{4c}\Big(\frac{1}{s_1^*}+\ldots+\frac{1}{s_n^*}\Big)-c\big(s_1^*+\ldots+s_n^*\big)+\log\Big(\frac{s_1^*\cdot\ldots\cdot s_n^*}{s_1\cdot\ldots\cdot s_n}\Big)\\
&+c\int_{s_1}^{s_1^*}\frac{\bar{f}^2(s_1',s_2,\ldots,s_n)}{(s_1')^2}\,ds_1'+\ldots+c\int_{s_n}^{s_n^*}\frac{\bar{f}^2(s_1,\ldots,s_{n-1},s_n')}{(s_n')^2}\,ds_n'\\
&=c\Big(\frac{1}{s_1}+\ldots+\frac{1}{s_n}\Big)\bar{f}^2(s_1,\ldots,s_n)
\end{aligned}$$

for $0\le s_1\le s_1^*,\ldots,0\le s_n\le s_n^*$, where s_i^* is the largest value of s_i which does not produce a (strict) leak either on the edge E_i or another edge, i.e. more specifically, setting $s:=(s_1,\ldots,s_n)$ we either let

$$s_i^*=f(\bar{s}_i)+\frac{1}{2c}\quad \text{if } f(\bar{s}_i)\in E_i \tag{4.14}$$

i.e. if $\bar{s}_i$ belongs to the domain of f, or we set $s_j:=\max\bar{s}_i$ and let s_i^* be the largest value of s_i in $[0,1/2c]$ for which $s_j\le f(\bar{s}_j)+1/2c$ for $1\le i\le n$. [Note that f specifying the domain boundary in (4.14) *and afterwards is the final solution to* (4.10) *above obtained for $k=n-1$.] Moreover, we have $\bar{f}(s_1^*,\ldots,s_n^*)=f(s_1^*,\ldots,s_{i-1}^*,s_{i+1}^*,\ldots,s_n^*)$ where s_i^* is the largest among $s_1^*,\ldots,s_n^*$. [Note that the zero leak does not affect this conclusion.]*

(V): The Dubins constant from (3.2) *and* (3.13) *is given by*

$$D_n=2\bar{f}(0) \tag{4.15}$$

where $0:=(0,\ldots,0)\in[0,\infty)^n$ for $n\ge1$ and we set $c=1$.

Proof. The optimal stopping problem (3.6) is $(n+1)$-dimensional and the underlying Markov process equals $(X,S^1,\ldots,S^n)$. It is evident from the structure of the gain process in (3.6) that the *excursions* of X away from the running maxima $S^1,\ldots,S^n$ play a key role in the problem. Another important feature that underpins many arguments throughout is a complete (permutationally invariant) *symmetry* of the running maxima $S^1,\ldots,S^n$ in the problem. We will divide the proof into ten parts as follows.

1. *Optimal stopping problem.* To tackle the optimal stopping problem (3.6) we will enable the Markov process $(X,S^1,\ldots,S^n)$ to start at any point $(x,s_1,\ldots,s_n)$ in the state space $E\times[0,\infty)^n$ under the probability measure $\mathsf{P}_{x,s_1,\ldots,s_n}$. The optimal stopping problem (3.6) then extends as follows

$$V(x,s_1,\ldots,s_n)=\sup_\tau \mathsf{E}_{x,s_1,\ldots,s_n}(S_\tau^1+\ldots+S_\tau^n-c\tau) \tag{4.16}$$

for $(x,s_1,\ldots,s_n)\in E\times[0,\infty)^n$ where the supremum is taken over all stopping times τ of X such that $\mathsf{E}_{x,s_1,\ldots,s_n}(\tau)<\infty$ and $c>0$ is a given and fixed constant.

The motion of X and its excursions away from the running maxima $S^1,\ldots,S^n$ can be visualised as shown in Figure 1. An important initial observation is that the process $(X,S^1,\ldots,S^n)$ can never be optimally stopped when $X^i=S^i$ i.e. when X belongs to E_i and equals S^i for $1\le i\le n$. The same phenomenon is known to hold for optimal stopping of the maximum process associated with general one-dimensional diffusion processes (see [20, Proposition 2.1])

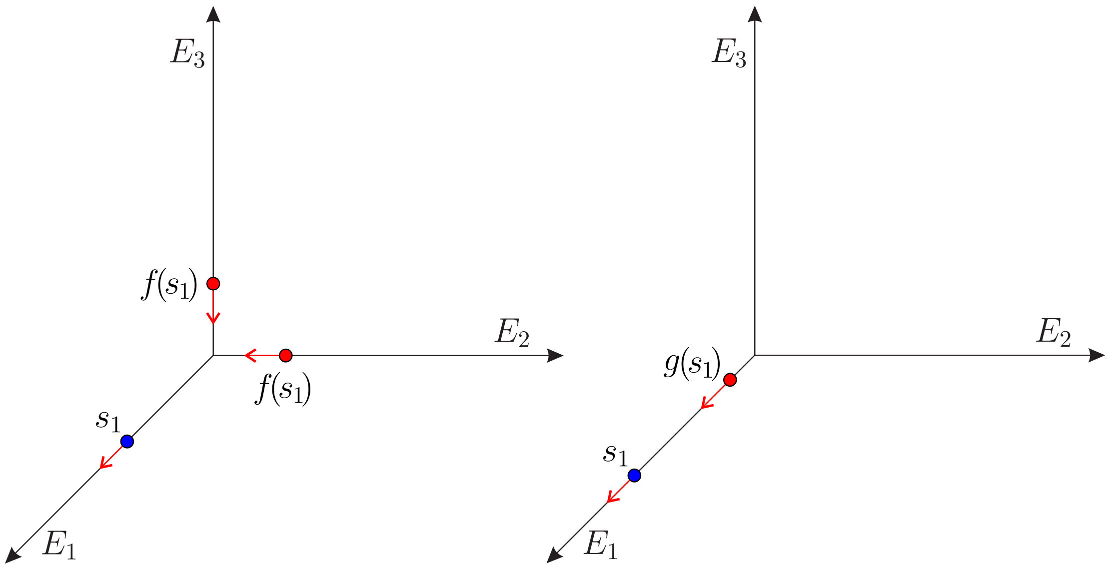

Figure 2. Optimal stopping boundaries f and g as functions of the running maximum s_1 attained at the edge E_1 (in the fictional absence of all other running maxima). At any point in E_2 or E_3 that is above $f(s_1)$, or at any point in E_1 that is below $g(s_1)$, there is no incentive to continue the motion (with a view to s_1 and its increase) because s_1 is too far away to offset the cost of getting there. When s_1 moves upwards, the optimal stopping value $f(s_1)$ moves downwards both within E_2 and E_3 until collapsing at the origin and turning into the optimal stopping value $g(s_1)$ which then moves upwards within E_1 ever afterwards. Exactly the same picture holds for the running maxima s_2 and s_3 attained at the edges E_2 and E_3 respectively. Understanding the interplay between the resulting interactions of the running maxima is of central importance for the Dubins problem (see Figures 4-6 below).

and the same arguments apply in the present setting without any changes. Before we formalise this in the next part below, let us recall that general theory of optimal stopping for Markov processes (see [22, Chapter 1]) implies that the *continuation set* in the problem (3.6) equals $C = \{ (x, s_1, \ldots, s_n) \in E \times [0, \infty)^n \mid V(x, s_1, \ldots, s_n) > s_1 + \ldots + s_n \}$ and the *stopping set* equals $D = \{ (x, s_1, \ldots, s_n) \in E \times [0, \infty)^n \mid V(x, s_1, \ldots, s_n) = s_1 + \ldots + s_n \}$, where we also know that the value function V is continuous so that C is open and D is closed. It means that the first entry time of $(X, S^1, \ldots, S^n)$ into D is optimal in the problem (3.6).

To determine the sets C and D we will begin by formalising the initial observation above. Any point $(x, s_1, \ldots, s_n)$ in the state space $E \times [0, \infty)^n$ such that $x \in E_i$ and $x = s_i$ for some $1 \leq i \leq n$ is referred to as a *diagonal* point.

2. *All diagonal points belong to* C . This is a simple consequence of the fact that if the process $(X, S^1, \ldots, S^n)$ starts at a diagonal point (meaning that $X_0 \in E_i$ and $X_0 = S^i$ for some $1 \leq i \leq n$) then the expected gain in the problem (3.6) is locally proportional to the root of time (because locally S^i is the running maximum of a standard Brownian motion) while the expected cost is linear in time so that it is not optimal to stop at once. We refer to the proof of Proposition 2.1 in [20] for fuller details.

Combining this conclusion with the same argument applied to the expected gain and cost for large times, which implies that it is not optimal to continue forever either, we see that a *sole in-*

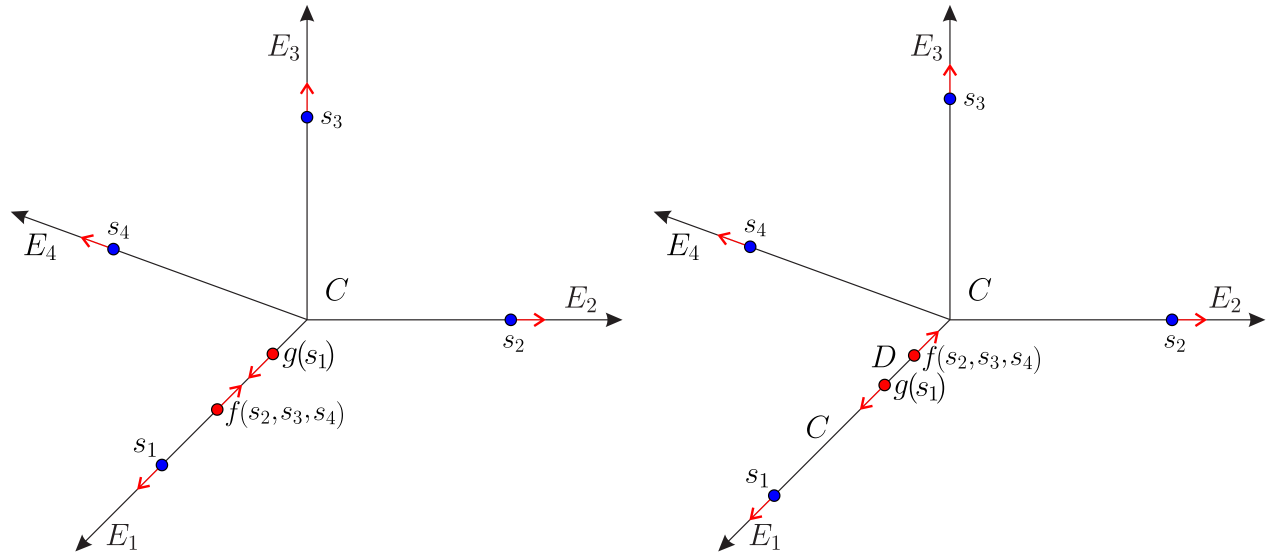

Figure 3. Optimal stopping boundaries f and g reflecting the closeness of $\{s_2, s_3, s_4\}$ and s_1 respectively. Initially all points in the state space E belong to the continuation set C (first graph). Then the increase of s_1, s_2, s_3, s_4 moves f and g across each other and creates points in the stopping set D (second graph). Removing S^1 from the gain process in the optimal stopping problem (3.6) creates a 'leak' on the edge E_1 meaning that the only incentive to continue the motion of X within E_1 below f is the closeness of X to the origin (to gain a potential entry to any of the running maxima on the edges E_2, E_3, E_4). The *separation principle* embodies the fact that adding S^1 to the gain process can only have two extreme effects: either (i) the optimal stopping value $f(s_2, s_3, s_4)$ is completely unaffected or (ii) the optimal stopping value $f(s_2, s_3, s_4)$ is no longer optimal (before at least one of the four running maxima is increased). This stands in sharp contrast with the usual behaviour of optimal stopping boundaries for diffusion processes on manifolds (non-graphs).

centive for continuation of the process $(X, S^1, \ldots, S^n)$ is the *closeness* of X to $\{S^1, \ldots, S^n\}$. Focusing on the single running maximum S^1 in the fictional absence of $S^2, \ldots, S^n$ as depicted in Figure 2, we see that this yields the existence of monotone optimal stopping boundaries f and g as functions of the running maximum s_1 attained at the edge E_1 (note that f transforms into g when passing through the origin). Exactly the same picture holds for any other running maximum (in isolation) among $S^2, \ldots, S^n$. Understanding the interplay between the resulting interactions of the running maxima is of central importance for the problem (3.6).

3. *Separation principle.* To address this interplay in fuller detail, let us assume that the process $(X, S^1, \ldots, S^n)$ starts at the origin (which clearly belongs to C as a diagonal point). The motion of X then triggers the interplay between f and g as depicted in Figure 3 for the edge E_1 and similar interplays take place simultaneously for all other edges too. Focusing on the edge E_1 alone for simplicity, the value $g(s_1)$ represents the optimal stopping value in the absence of $s_2, \ldots, s_n$, and similarly, the value $f(s_2, \ldots, s_n)$ represents the (combined) optimal stopping value in the absence of s_1 . Clearly the mapping $s_1 \mapsto g(s_1)$ is increasing and the mapping $s_i \mapsto f(s_2, \ldots, s_n)$ is decreasing for every $2 \le i \le n$. Hence after waiting sufficiently long the values $g(s_1)$ and $f(s_2, \ldots, s_n)$ will meet (see Figure 3). At this time something interesting happens. Namely, if the equality $g(s_1) = f(s_2, \ldots, s_n)$ is attained

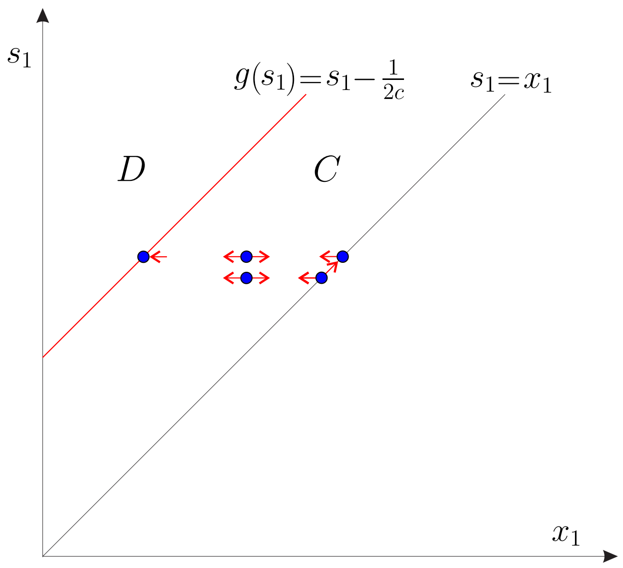

Figure 4. The optimal stopping boundary $s_1 \mapsto g(s_1)$ when a leak is created on the edge E_1 through the increase of S^1 . The resulting optimal stopping problem is two-dimensional so that the maximality principle [20] applies and the solution can be derived exactly (the optimal stopping boundary g is linear as shown in the graph above).

through the increase of s_1 , then the entire problem fully simplifies to two dimensions of the process (X^1, S^1) only, in which case the maximality principle [20] applies and the solution can be derived exactly (see Figure 4). On the other hand, if the equality $g(s_1) = f(s_2, \dots, s_n)$ is attained through the increase of s_i for some $2 \le i \le n$, then the problem reduces to the problem (3.6) where S^1 is removed from the gain process. The edge E_1 becomes an *edge with leak* in the latter case meaning that the only incentive to continue the motion of X within E_1 is its closeness to the origin (to gain a potential entry to any of the running maxima in the remaining edges). Thus, the problem is simplified to the edge E_1 alone in the former case, and the problem is reduced to the edges $E_2, \dots, E_n$ in the latter case (see the caption to Figure 3 for an equivalent/dual interpretation of this splitting). The *separation principle* just described stands in sharp contrast with the observed behaviour of optimal stopping boundaries for diffusion processes on manifolds (non-graphs).

4. *Edges with leaks*. Applying the separation principle inductively starting with (4.1) above leads to the *pruning procedure* described in Steps 1−n above. An edge with leak is created in each step and either the motion of $(X, S^1, \dots, S^n)$ is stopped at the boundary g if the leak is created through the increase of the running maximum in the leaked edge, or the motion of $(X, S^1, \dots, S^n)$ is continued and either stopped at the boundary f or a new leak is created (see Figure 5 for a leak created at E_1 in the first step followed by a leak created at E_2 in the second step). Proceeding recursively in this way until the final two edges are reached (given that it was not optimal to stop before) we arrive at Step n above where after a new leak is created the problem reduces to X being a skew Brownian motion (the first graph in Figure 6) and then a standard Brownian motion (the second graph in Figure 6) with a single running maximum in both cases, so that the maximality principle [20] applies as earlier and the solution

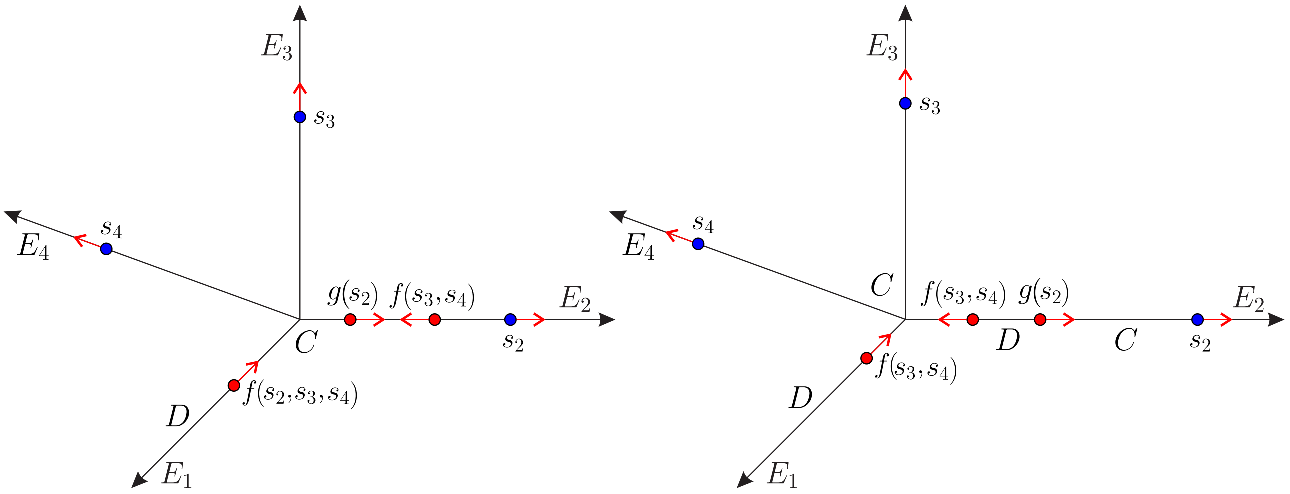

Figure 5. A leak created on E_1 in the first step (first graph) followed by a leak created on E_2 in the second step (second graph). The two leaked edges E_1 and E_2 having the angular probability $1/n$ each can be combined into a single (fictional) leaked edge having the angular probability $2/n$ (with $n=4$ in the graphs above).

can be derived exactly. Finally, the motion of $(X, S^1, \ldots, S^n)$ is optimally stopped for good at either the boundary f or the boundary g as stated in (4.9) above.

5. *Free-boundary problem.* To determine the value function V and the optimal stopping boundaries f and g, we are led to formulate the free-boundary problem

$$\mathbb{L}_X V(x, s_1, \ldots, s_n) = c \quad \text{for } (x, s_1, \ldots, s_n) \in C \tag{4.17}$$

$$V_1(0+, s_1, \ldots, s_n) = \ldots = V_n(0+, s_1, \ldots, s_n) = 0 \quad \text{(continuity)} \tag{4.18}$$

$$\textstyle\sum_{i=1}^n \frac{\partial V}{\partial x_i}(0+, s_1, \ldots, s_n) = 0 \quad \text{(Kirchhoff)} \tag{4.19}$$

$$\frac{\partial V}{\partial s_i}(x_i, s_1, \ldots, s_n)\Big|_{x_i = s_i -} \quad \text{(normal reflection)} \tag{4.20}$$

$$V(x_i, s_1, \ldots, s_n)\Big|_{x_i = g(s_i)+} = s_1 + \ldots + s_n \quad \text{(instantaneous stopping)} \tag{4.21}$$

$$\frac{\partial V}{\partial x_i}(x_i, s_1, \ldots, s_n)\Big|_{x_i = g(s_i)+} = 0 \quad \text{(smooth fit)} \tag{4.22}$$

$$V(x_i, s_1, \ldots, s_n)\Big|_{x_i = f(\bar{s}_{i_1, \ldots, i_k})-} = s_1 + \ldots + s_n \quad \text{(instantaneous stopping)} \tag{4.23}$$

$$\frac{\partial V}{\partial x_i}(x_i, s_1, \ldots, s_n)\Big|_{x_i = f(\bar{s}_{i_1, \ldots, i_k})-} = 0 \quad \text{(smooth fit)} \tag{4.24}$$

for $s_1, \ldots, s_n \in [0, \infty)$, all $1 \le i \le n$ in (4.20)-(4.22), and all $i \in \{i_1, \ldots, i_k\} \subseteq I$ with $1 \le k \le n-1$ in (4.23)-(4.24), where $\mathbb{L}_X$ is the infinitesimal generator of X given in (2.7) above and the continuation set C is given by

$$\begin{aligned} C = \Big\{ &(x, s_1, \ldots, s_n) \in E \times [0, \infty)^n \mid x \in E_i \text{ for some } i \in I \text{ and } \textit{either} \\ &(s_{i_1}, \ldots, s_{i_k - 1}) \text{ belongs to the domain of } f \text{ and } x < f(s_{i_1}, \ldots, s_{i_k - 1}) \\ &\text{with } k \in I \text{ satisfying } i_k = i \text{ for } s_{i_1} \le \ldots \le s_{i_n} \textit{ or } x > g(s_i) \Big\}. \end{aligned} \tag{4.25}$$

For the rationale and further details regarding free-boundary problems of this kind we refer to [22, Section 13] and the references therein (see also [5] for derivations of the smooth-fit conditions (4.22) and (4.24) in more general situations). These general arguments apply in the present setting without any changes. Note that (2.8) implies (4.19) because X is symmetric so that $p_1 = \ldots = p_n = 1/n$ and the representation (4.25) follows from the pruning procedure described in Steps 1−n above.

6. *Maximality principle.* The system (4.17)-(4.24) does not determine the value function V and the optimal stopping boundaries f and g uniquely. This is a consequence of the fact that when the optimal stopping problem (3.6) is reduced to a leaked edge E_i for some $1 \le i \le n$ as described in Steps 1−n above, then the system (4.17)-(4.24) collapses to a subsystem consisting of (4.17) combined with (4.20)-(4.22) above. Probabilistically this corresponds to a two-dimensional optimal stopping problem studied in [20] where the maximality principle implies that the optimal stopping boundary $s_i \mapsto g(s_i)$ is the maximal solution to the *nonlinear* differential equation

$$g'(s_i) = \frac{\sigma^2(g(s_i))\, L'(g(s_i))}{2c(g(s_i))\,(L(s_i) - L(g(s_i))} \tag{4.26}$$

staying strictly below the diagonal in the state space of (X^i, S^i), and the value function V is given explicitly by

$$V(x_i, s_1, \ldots, s_n) = s_1 + \ldots + s_n + \int_{g(s_i)}^{x_i} (L(x_i) - L(y))c(y)\, m(dy) \tag{4.27}$$

for $g(s_i) \le x_i \le s_i$ where $1 \le i \le n$ is the index of a leaked edge (note that every other solution g to (4.26) also makes V from (4.27) a solution to the subsystem consisting of (4.17) combined with (4.20)-(4.22) above). The equations (4.26) and (4.27) are expressed in a greater generality of one-dimensional diffusion processes and when specialised to the symmetric Walsh spider process X we know that the diffusion coefficient $\sigma(x) \equiv 1$, the cost function $c(x) \equiv c$, the scale function $L(x) = x$ and the speed measure $m(dx) = 2dx$. Inserting these formulae into (4.26) and (4.27) yields explicit expressions that we omit for brevity. The maximal solution to (4.26) is then given by

$$g(s_i) = s_i - \frac{1}{2c} \tag{4.28}$$

for $s_i \ge 1/2c$. This is the formula (4.12) above. For fuller details of these arguments we refer to Theorem 3.1 and Figure 1 in [20].

After the optimal stopping boundary g is determined uniquely on a leaked edge using the maximality principle, it turns out (as we will see below) that the optimal stopping boundary f is determined uniquely (and hence the value function V too). This is an important point to grasp for a deeper understanding of the solution in general. The first step in this direction corresponds to determining f which appears in (4.9) of the final Step n above (see the first graph in Figure 6). Clearly finding f is equivalent to solving a two-dimensional optimal stopping problem studied in [20] where X is a skew Brownian motion in $I\!R$ with the skewness parameter β for $[0, \infty)$ at 0 equal to $1/n$ (implying that the skewness parameter for $(-\infty, 0]$ at 0 equals $1-\beta = (n-1)/n$ as stated in Figure 6). The scale function of X is given by

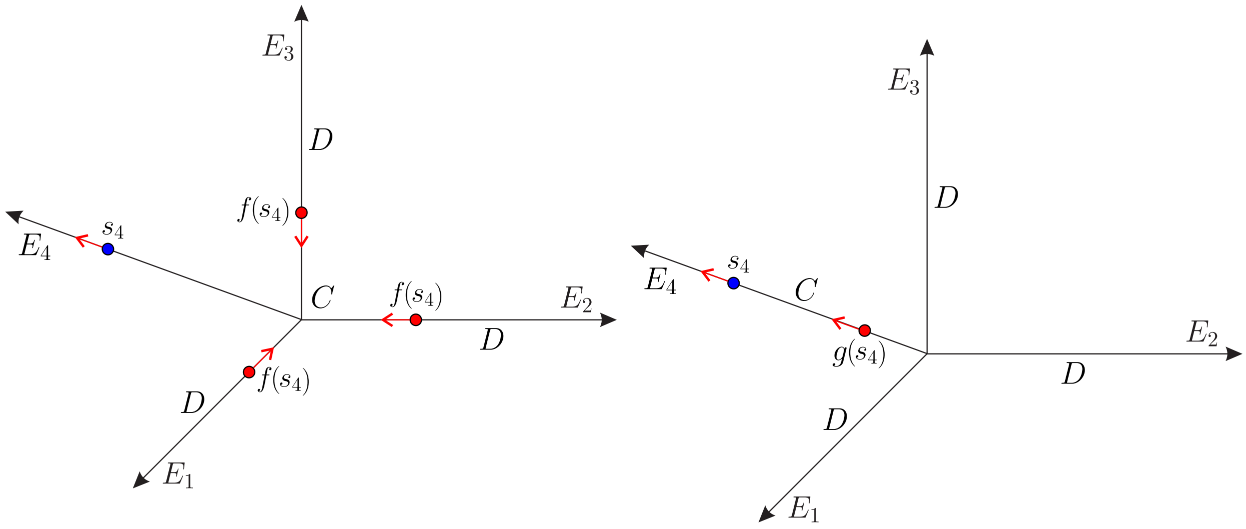

Figure 6. Proceeding recursively until the final edge is reached (given that it was not optimal to stop before and assuming that the final edge equals E_n without loss of generality) the problem reduces to X being (i) a skew Brownian motion with the angular probabilities $1/n$ and $(n-1)/n$ for E_n (with $n=4$ in the graphs above) and the fictional leaked edge respectively (first graph) and then (ii) a standard Brownian motion (second graph) with a single running maximum in both cases so that the maximality principle [20] applies and the solution can be derived exactly (the optimal stopping boundary f is curved and the optimal stopping boundary g is linear as shown in Figure 4 above).

$L(x) = (1-\beta)x$ if $x \geq 0$ and $L(x) = \beta x$ if $x \leq 0$, and the speed measure $m(dx) = 2dx$ is unchanged. The maximality principle implies that the optimal stopping boundary $s_i \mapsto g(s_i)$ is the maximal solution to (4.26) and the value function V is given explicitly by (4.27) above where i is the index of the remaining unleaked edge. Noting that the equation (4.26) is the same as in the non-skewed case when $g(s_i) \geq 0$ we see that the maximal solution to (4.26) is given by (4.28) above for $s_i \geq 1/2c$. The value $g(1/2c) = 0$ combined with continuity of g then determines the solution to (4.26) for $s_i \in [0, 1/2c]$ uniquely. Moreover, passing to an equivalent *linear* equation for the inverse function of g one can readily verify that

$$s_i(g) = \frac{1}{2c}\Big(\frac{\kappa^2-1}{\kappa^2}\Big)e^{2c\kappa g} + \frac{1}{\kappa}g + \frac{1}{2c\kappa^2} \tag{4.29}$$

for $g \in [g_0, 0]$ where $s_i(g_0) = 0$ and we set $\kappa = (1-\beta)/\beta$. Denoting the inverse function of s_i from (4.29) by g, we see that the sought f is given by

$$f(s_i) = -g(s_i) \tag{4.30}$$

for $s_i \in [0, 1/2c]$ (see Figure 9). This explains how f appearing in (4.9) of the final Step n above is determined. Before we show how this f can be used to determine the function f appearing in Step n−1 above uniquely, and proceed similarly by recursion until Step 1 is reached, we will first present a canonical result that underpins all the equations to be derived in this way. Note that the canonical result is obtained as a consequence of the strong Markov property combined with the condition of smooth fit (expressing a variational principle) and

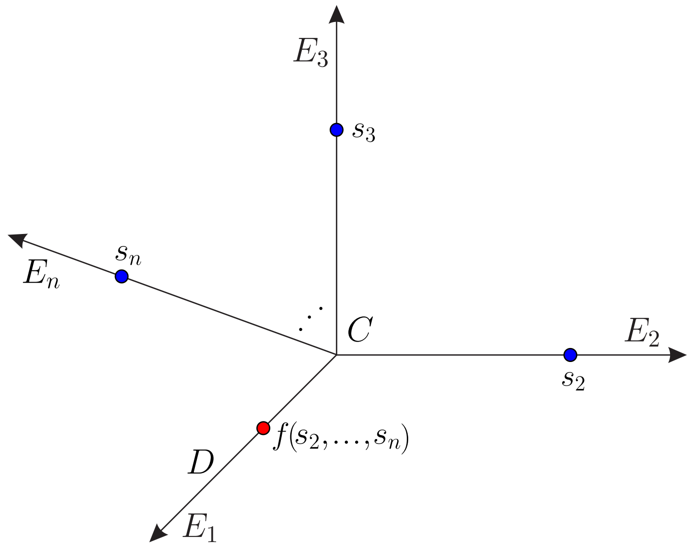

Figure 7. A canonical setting for derivation of all equations for f where E_1 is the (only) edge with a leak for a Walsh spider process X on n edges with angular probabilities $p_1, \ldots, p_n$ and $f(s_2, \ldots, s_n)$ is the optimal stopping point in E_1 for $s_2, \ldots, s_n$ given and fixed.

the condition of normal reflection (belonging to the infinitesimal characteristics of the process) in the free-boundary problem (4.17)-(4.24) above, and may be viewed as a realisation of the generally known fact that a harmonic function in the interior of a set is uniquely determined by its values at the boundary of the set.

7. *Smooth fit and normal reflection.* Consider the optimal stopping problem (3.6) for a Walsh spider process X on n edges with angular probabilities $p_1, \ldots, p_n$, and suppose that E_1 is the (only) edge with a leak (see Figure 7) where $f := f(s_2, \ldots, s_n)$ is the optimal stopping point in E_1 for the initial values $s_2, \ldots, s_n \in [0, \infty)$ of the running maxima $S^2, \ldots, S^n$ of X given and fixed. Because E_1 is the only edge with a leak, we know by the separation principle that the area in E surrounded by $f, s_2, \ldots, s_n$ is contained in the continuation set C. Focusing on that area we can therefore omit s_1 and S^1 from (4.16) as they play no role. Set

$$V(s_i) := V(x_i, s_2, \ldots, s_n)\big|_{x_i = s_i} \tag{4.31}$$

for $i = 2, \ldots, n$ and let $x_1 \in [0, f), x_2 \in [0, s_2), \ldots, x_n \in [0, s_n)$ be given and fixed. Applying the strong Markov property of $(X, S^2, \ldots, S^n)$ at the first entry time τ_A of X into the set $A := \{\, f, s_2, \ldots, s_n \,\}$ we find that

$$V(x_i, s_2, \ldots, s_n) = (s_2 + \ldots + s_n)\, \mathsf{P}_{x_i}(X_{\tau_A} = f) + \sum_{j=2}^{n} V(s_j)\, \mathsf{P}_{x_i}(X_{\tau_A} = s_j) - c\, \mathsf{E}_{x_i}(\tau_A) \tag{4.32}$$

for $1 \le i \le n$. Note that the identities (2.10)-(2.12) are applicable to the exit probabilities and mean exit times in (4.32) with $a_1 = f, a_2 = s_2, \ldots, a_n = s_n$.

Focusing on x_1 and rearranging the terms in (4.32) we see that

$$\frac{V(x_1, s_2, \ldots, s_n) - (s_2 + \ldots + s_n)}{1 - \mathsf{P}_{x_1}(X_{\tau_A} = f)} = -(s_2 + \ldots + s_n) + \sum_{j=2}^{n} V(s_j)\, \frac{\mathsf{P}_{x_1}(X_{\tau_A} = s_j)}{1 - \mathsf{P}_{x_1}(X_{\tau_A} = f)} \tag{4.33}$$

$$- c\,\frac{\mathsf{E}_{x_1}(\tau_A)}{1-\mathsf{P}_{x_1}(X_{\tau_A}=f)}\,.$$

Noting that $s_2+\ldots+s_n = V(f,s_2,\ldots,s_n)$ and making use of (2.10) above we find that

$$\frac{V(x_1,s_2,\ldots,s_n)-(s_2+\ldots+s_n)}{1-\mathsf{P}_{x_1}(X_{\tau_A}=f)} \tag{4.34}$$

$$= \frac{V(x_1,s_2,\ldots,s_n)-V(f,s_2,\ldots,s_n)}{f-x_1}\cdot\frac{f(p_1/f+\sum_{k=2}^n p_k/s_k)}{\sum_{k=2}^n p_k/s_k}$$

$$\longrightarrow \frac{\partial V}{\partial x}(f,s_2\ldots,s_n)\cdot\frac{f(p_1/f+\sum_{k=2}^n p_k/s_k)}{\sum_{k=2}^n p_k/s_k} = 0$$

as $x_1 \uparrow f$ where in the final step we use the smooth-fit condition (4.24) above.

Making use of (2.10) and (2.11) it is readily verified that

$$\frac{\mathsf{P}_{x_1}(X_{\tau_A}=s_j)}{1-\mathsf{P}_{x_1}(X_{\tau_A}=f)} = \frac{1}{(s_j/p_j)\sum_{k=2}^n p_k/s_k} \tag{4.35}$$

in (4.33) above. Similarly, making use of (2.10) and (2.12) it is readily verified that

$$\frac{\mathsf{E}_{x_1}(\tau_A)}{1-\mathsf{P}_{x_1}(X_{\tau_A}=f)} = \frac{p_1 f+\sum_{k=2}^n p_k s_k + (p_1+f\sum_{k=2}^n p_k/s_k)x_1}{\sum_{k=2}^n p_k/s_k} \tag{4.36}$$

in (4.33) above. Letting $x_1 \uparrow f$ in (4.33) and making use of (4.34)-(4.36) we obtain

$$\sum_{j=2}^n p_j\frac{V(s_j)}{s_j} = \Big(\sum_{k=2}^n s_k\Big)\Big(\sum_{k=2}^n p_k/s_k\Big) + c\Big(\sum_{k=2}^n p_k s_k + 2p_1 f + f^2\sum_{k=2}^n p_k/s_k\Big)\,. \tag{4.37}$$

Note that the identity (4.37) is derived by an essential use of the smooth-fit condition (4.24) (in addition to an implicit use of the infinitesimal-characteristics conditions (4.18) and (4.19) that underpin the identities (2.10)-(2.12) above).

Returning to (4.32) and making use of (2.10)-(2.12) combined with the identity (4.37) above, we find by a lengthy but straightforward calculation that

$$V(x_i,s_2,\ldots,s_n) = \Big(cf^2+\sum_{k=2}^n s_k\Big)\,\frac{s_i-x_i}{s_i} + V(s_i)\,\frac{x_i}{s_i} - cx_i(s_i-x_i) \tag{4.38}$$

for $2 \le i \le n$. Differentiating with respect to s_i in (4.38) and making use of the normal-reflection condition (4.20) above, we obtain

$$\frac{\partial}{\partial s_i}\Big(\frac{V(s_i)}{s_i}\Big) = c\big(1-f^2/s_i^2\big) - \sum_{k=2}^n s_k/s_i^2 \tag{4.39}$$

for $2 \le i \le n$. Note that the identity (4.39) is derived by an essential use of the smooth-fit condition (4.24) combined with the normal-reflection condition (4.20) (in addition to an implicit use of the infinitesimal-characteristics conditions (4.18) and (4.19) that underpin the identities

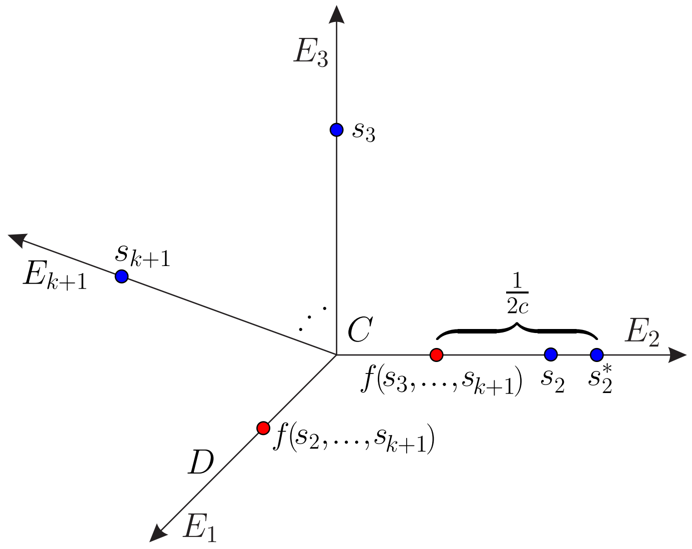

Figure 8. A realisation of the canonical setting from Figure 7 (with $k+1$ in place of n) where E_1 represents an aggregate of $n-k$ edges with leaks (having the angular probability $p_1 = (n-k)/n$) and the remaining k edges $E_2, \ldots, E_{k+1}$ (having the angular probabilities $p_2 = \ldots = p_{k+1} = 1/n$) contain no leak. With $f(s_3, \ldots, s_{k+1})$ determined in the previous step, and s_2^* being the largest value of s_2 which does not produce a leak on E_2, the aim is to find/determine the optimal stopping value $f(s_2, \ldots, s_{k+1})$ on each of the $n-k$ edges with leaks.

(2.10)-(2.12) above). We will use the identity (4.39) below to derive integral equations that characterise the optimal stopping boundaries f uniquely.

In addition to (4.38) above, returning to (4.32) and making use of (2.10)-(2.12) combined with the identity (4.37) above, we find by a lengthy but straightforward calculation that

$$V(x_1, s_2, \ldots, s_n) = \sum_{k=2}^{n} s_k + c(f - x_1)^2 . \tag{4.40}$$

Unlike in (4.37) note that the expressions on the right-hand side of (4.38)-(4.40) contain no explicit dependence on $p_1, \ldots, p_n$ (this dependence sits entirely in f instead). Note also that the function V given by (4.38) and (4.40) satisfies the Kirchhoff condition (2.8) (or equivalently (4.19) in the symmetric case) when the condition (4.37) holds.

8. *Integral equations.* Going backwards from Step n to Step 1 in the pruning procedure from (I) above, in each step we are faced with a particular realisation of the canonical setting from Figure 7 as shown in Figure 8. With k running from 1 to $n-1$ given and fixed, without loss of generality (using that X is symmetric) we may assume as in Figure 8 that the edge E_1 represents an aggregate of $n-k$ edges with leaks and the remaining k edges contain no leak. In the canonical setting of Figure 7 this corresponds to n being equal to $k+1$ with the angular probabilities $p_1 = (n-k)/n$ and $p_2 = \ldots = p_{k+1} = 1/n$. With $(s_3, \ldots, s_{k+1}) \mapsto f(s_3, \ldots, s_{k+1})$ determined in the previous step (note that the number of arguments of f equals $k-1$), we then aim to find/determine $(s_2, \ldots, s_{k+1}) \mapsto f(s_2, \ldots, s_{k+1})$ representing the optimal stopping value on each of the $n-k$ edges with leaks. If $f(s_3, \ldots, s_{k+1})$ belongs to E_2 , i.e. if $(s_3, \ldots, s_{k+1})$ belongs to the domain of f , then the largest value of s_2

which does not produce a (strict) leak on E_2 is given by $s_2^* = f(s_3,\dots,s_{k+1})+1/2c$ due to (4.12) above. Exactly the same argument (by symmetry) applies to any of the remaining edges $E_3,\dots,E_{k+1}$ with no leak and this leads to the values $s_3^*,\dots,s_{k+1}^*$ whenever the values of f are defined. On the other hand, if $(s_3,\dots,s_{k+1})$ does not belong to the domain of f, then using symmetry we set $s_j = \max\{s_3,\dots,s_{k+1}\}$ and let s_2^* be the largest value of s_2 in $[0,1/2c]$ for which $s_j \le f(s_2,\dots,s_{j-1},s_{j+1},\dots,s_{k+1})+1/2c$. In other words, we swap the roles of s_2 and s_j upon using that $(s_2,\dots,s_{j-1},s_{j+1},\dots,s_{k+1})$ must belong to the domain of f because the edges $E_2,\dots,E_{k+1}$ contain no leak (noting also that $s_2 \mapsto f(s_2,\dots,s_{j-1},s_{j+1},\dots,s_{k+1})$ is decreasing). Exactly the same argument (by symmetry) applies to any of the remaining edges $E_3,\dots,E_{k+1}$ with no leak and this leads to the values $s_3^*,\dots,s_{k+1}^*$ when the values of f are not defined either.

The key reason for singling out the points $s_2^*,\dots,s_{k+1}^*$ is that their extremity makes the value function from (4.31) with $k+1$ in place of n expressible explicitly at these points. Namely, calculating the integral in (4.27) explicitly as indicated above (with s_1 removed and $k+1$ in place of n), we see that

$$V(s_i^*) := V(s_i^*, s_2,\dots,s_{i-1},s_i^*,s_{i+1},\dots,s_{k+1}) = s_i^* + \sum_{j=2,j\neq i}^{k+1} s_j + \frac{1}{4c} \tag{4.41}$$

for $i = 2,\dots,k+1$. This expression can now be used to instantiate the identity (4.39) combined with the identity (4.37) above. For this, note that

$$\sum_{i=2}^{k+1} \frac{V(s_i^*)}{s_i^*} = \sum_{i=2}^{k+1} \frac{V(s_i)}{s_i} + \sum_{i=2}^{k+1} \int_{s_i}^{s_i^*} \frac{\partial}{\partial s_i'}\Big(\frac{V(s_i')}{s_i'}\Big)\, ds_i' \tag{4.42}$$

for $0 \le s_i \le s_i^*$ with $2 \le i \le k+1$. Making use of (4.41) in the first sum, (4.37) with $k+1$ in place of n and $p_1 = (n-k)/n$ alongside $p_2 = \ldots = p_{k+1} = 1/n$ in the second sum, and (4.39) in the third sum above, a lengthy but straightforward calculation shows that (4.42) reads

$$\begin{aligned}\sum_{i=2}^{k+1} \bigg[\frac{1}{4cs_i^*} - cs_i^* + \log\Big(\frac{s_i^*}{s_i}\Big) + c\int_{s_i}^{s_i^*} \frac{f^2(s_2,\dots,s_{i-1},s_i',s_{i+1},\dots,s_{k+1})}{(s_i')^2}\, ds_i' \bigg] \\ = c\Big(\sum_{i=2}^{k+1}\frac{1}{s_i}\Big) f^2(s_2,\dots,s_{k+1}) + 2c(n-k) f(s_2,\dots,s_{k+1})\end{aligned} \tag{4.43}$$

for $0 \le s_i \le s_i^*$ with $2 \le i \le k+1$. Note that (4.43) is exactly the equation (4.10) above stated with renamed variables.

9. *Uniqueness*. Probabilistic arguments presented above establish the existence of a non-negative solution $(s_1,\dots,s_k) \mapsto f(s_1,\dots,s_k)$ to the equation (4.10) with $1 \le k \le n-1$ given and fixed. From the closed-form expression (4.38) above we see that f is continuous (because V is so). To show that the solution is unique, suppose that g is another non-negative continuous solution to (4.10) and let $\varepsilon > 0$ be given and fixed. Setting $f := f(s_1,\dots,s_k)$ and $g := g(s_1,\dots,s_k)$ for $\varepsilon \le s_1 \le s_1^*,\dots,\varepsilon \le s_k \le s_k^*$ given and fixed, we see from (4.10) that

$$\Big[\Big(\frac{1}{s_1} + \ldots + \frac{1}{s_k}\Big)(f+g) + 2(n-k)\Big](f-g) \tag{4.44}$$

$$= \int_{s_1}^{s_1^*} \frac{(f^2-g^2)(s_1', s_2, \ldots, s_k)}{(s_1')^2}\, ds_1' + \ldots + \int_{s_k}^{s_k^*} \frac{(f^2-g^2)(s_1, s_2, \ldots, s_k')}{(s_k')^2}\, ds_k' \, .$$

Rearranging the terms and using the fact that both f and g are continuous and thus bounded on their compact domain, we find that

$$\begin{aligned}
(4.45) \qquad |f-g| &\le \frac{K_\varepsilon}{\varepsilon^2}\Big[\int_{s_1}^{s_1^*} |(f-g)(s_1', s_2, \ldots, s_k)|\, ds_1' + \ldots + \int_{s_k}^{s_k^*} |(f-g)(s_1, s_2, \ldots, s_k')|\, ds_k' \Big] \\
&= \frac{K_\varepsilon}{\varepsilon^2}\Big[\int_{s_1}^{s_1^*} |(f-g)(s_1', s_2, \ldots, s_k)|\, e^{\rho s_1'} e^{\rho s_2} \ldots e^{\rho s_k} e^{-\rho s_1'} e^{-\rho s_2} \ldots e^{-\rho s_k}\, ds_1' + \ldots \\
&\qquad + \int_{s_k}^{s_k^*} |(f-g)(s_1, s_2, \ldots, s_k')|\, e^{\rho s_1} \ldots e^{\rho s_{k-1}} e^{\rho s_k'} e^{-\rho s_1} \ldots e^{-\rho s_{k-1}} e^{-\rho s_k'}\, ds_k' \Big] \\
&\le \frac{K_\varepsilon}{\varepsilon^2}\, \|f-g\|_\varepsilon \Big[\frac{1}{\rho}\big(e^{-\rho s_1} - e^{-\rho s_1^*}\big) e^{-\rho s_2} \ldots e^{-\rho s_k} + \ldots + e^{-\rho s_1} \ldots e^{-\rho s_{k-1}} \frac{1}{\rho}\big(e^{-\rho s_k} - e^{-\rho s_k^*}\big)
\end{aligned}$$

for $K_\varepsilon > 0$ large enough and any $\rho > 0$ where we set

$$(4.46) \qquad \|f-g\|_\varepsilon := \sup_{\varepsilon \le s_i \le s_i^*, 1 \le i \le k} \Big(|(f-g)(s_1, s_2, \ldots, s_k)|\, e^{\rho s_1} \ldots e^{\rho s_k} \Big) .$$

Multiplying both sides in (4.45) by $e^{\rho s_1} \ldots e^{\rho s_k}$ and taking the supremum as in (4.46) we get

$$(4.47) \qquad \|f-g\|_\varepsilon \le \frac{K_\varepsilon}{\varepsilon^2} \frac{k}{\rho}\, \|f-g\|_\varepsilon \, .$$

Choosing ρ large enough so that $(K_\varepsilon k)/(\varepsilon^2 \rho) < 1$ we see that $\|f-g\|_\varepsilon$ must be zero. Letting $\varepsilon \downarrow 0$ we obtain $f = g$ as claimed.

10. *Edge with zero leak.* Having reached the final step $k = n-1$ in the system (4.10), and recalling from (3.13) that our aim is to determine the value $V(0, 0, \ldots, 0)$ in (4.16), we see from (4.40) that the latter value is still inaccessible because the optimal stopping boundary $(s_1, \ldots, s_n) \mapsto f(s_1, \ldots, s_{n-1})$ derived in the final step takes values in the edge E_1 having a leak so that $s_1 \ge 1/2c > 0$ (recall the paragraph containing (4.31) above) and we cannot let s_1 be or tend to 0. Potentially this could be a serious issue and the novel argument that we now present to resolve it is both insightful and beautiful.

The first step in this direction was already undertaken by introducing the general angular probabilities $p_1, \ldots, p_n$ in the paragraph containing (4.31) above. Moreover, it was shown above that these angular probabilities play an explicit role only in the smooth-fit relation (4.37) with p_1 standing out and leading to the appearance of the factor $2c(n-k)$ in the equation (4.10) making it nonlinear. Motivated by these observations and our aim recalled above we are thus naturally led to continue the recursion in the system (4.10) beyond $k = n-1$ by adding an extra edge E_0 to the state space $E = \coprod_{i=1}^n E_i$ of X having a zero leak. Probabilistically this means that X moves within E_0 as a standard Brownian motion, and after hitting the origin in $\bar{E} := \coprod_{i=0}^n E_i$ from within E_0, the process X reflects instantaneously into any of the edges $E_1, \ldots, E_n$ with probability $1/n$. The derivation of (4.37) and (4.39) presented above remains valid in this case with $p_1 = 0$ and $p_2 = \ldots = p_{n+1} = 1/n$ where n is replaced by $n+1$ throughout. The derivation of the integral equation (4.43) remains valid as well with the notable difference that $p_1 = 0$ kills the final term in (4.43). This leads to the *linear* integral equation (4.13) where the optimal stopping boundary $(s_1, \ldots, s_n) \mapsto \bar{f}(s_1, \ldots, s_n)$ takes values in E_0 and f specifying the domain boundary (alongside the boundary values) as stated in (4.14) above is the final solution to (4.10) corresponding to $k = n-1$.

Probabilistic arguments presented above establish the existence of a continuous non-negative solution $(s_1, \ldots, s_n) \mapsto \bar{f}(s_1, \ldots, s_n)$ to the equation (4.13). A minor modification of the uniqueness argument for the equation (4.10) given above shows that this solution is unique. The derivation of (4.40) remains valid as well and the resulting identity for the value function $\bar{V}$ on $\bar{E} \times [0, \infty)^n$ defined as in (4.16) reads as follows

$$\bar{V}(x_0, s_1, \ldots, s_n) = \sum_{k=1}^{n} s_k + c(\bar{f} - x_0)^2 \tag{4.48}$$

for $0 \leq x_0 \leq \bar{f} := \bar{f}(s_1, \ldots, s_n)$ with $0 \leq s_i \leq s_i^*$ for $1 \leq i \leq n$. The process X started at the origin in $\bar{E}$ coincides with a symmetric Walsh process on the edges $E_1, \ldots, E_n$. A formal verification of the fact that the solution V (with f and g) to the free boundary problem (4.17)-(4.24) constructed by means of τ_* equals the value function from (4.16) can be carried out in a standard way by applying Itô's formula to V composed with $(X, S^1, \ldots, S^n)$ (upon invoking a well-known justification [15, Remark 4.2] in these settings and using (2.5) above) combined with the optional sampling theorem and the fact that τ_* has a finite expectation (due to the maximality principle). Letting $x_0 = s_1 = \ldots = s_n = 0$ in (4.48) we see that the value function V from (4.16) satisfies

$$V(0) = c\bar{f}^2(0) \tag{4.49}$$

where $0 := (0, \ldots, 0) \in [0, \infty)^n$. Inserting this value with $c = 1$ in (3.13) we obtain (4.15) as claimed and the proof is complete. $\square$

Remark 1. As pointed out in the proof above, note that the pruning procedure of Steps 1−n is realised in a forward direction when observing and stopping the process X optimally, while the systems (4.10) and (4.13) need to be solved recursively in the opposite/backward direction when finding the optimal stopping boundary f used in the pruning procedure. Note also that a leak in the pruning procedure is always created on the edge having the largest running maximum among all edges with no leak.

Remark 2. In an attempt to specify the general solution to (4.17) within the free-boundary problem (4.17)-(4.24) we went on to exploit *symmetries* of the value function V (alongside its continuity) in the Kolmogorov-Arnold theorem and its refinement due to Sprecher [23]. Although this approach was eventually abandoned it may still be of interest for future studies.

Remark 3. We used the weighted sup norm (4.46) to establish uniqueness of the solution to the integral equations (4.10) and (4.13). Tricks of this kind are known in the literature on integral equations (see [26, pp 61-62]). Another way to derive the uniqueness could be based on multidimensional versions of the Wendroff inequality (see (5)-(6) in [3, p. 154]). This line of argument may also be of interest for further development.

Remark 4. Applying $\partial^2/\partial s_1 \partial s_2$ to both sides in (4.10) with $k = 2$ shows that the nonlinear integral equation (4.10) can be rewritten as the nonlinear PDE of *hyperbolic* type

$$\frac{\partial^2 f}{\partial s_1 \partial s_2} = -\frac{\frac{\partial f}{\partial s_1}\frac{\partial f}{\partial s_2}}{f + 1/(\frac{1}{s_1}+\frac{1}{s_2})} \tag{4.50}$$

where $f = f(s_1, s_2)$. The appearance of hyperbolic equations in the probabilistic setting is both unusual and interesting.

5. Special cases

In this section we specialise the result of Theorem 1 to dimensions $n = 1, 2, 3$. We show how the known Dubins constants D_1 and D_2 can be determined explicitly and how the long-sought Dubins constant D_3 can be determined numerically. The absence of an explicit expression for D_3 is attributed to the fact that the origin is a *triple branching point* in the state space of Walsh's spider process on 3 edges (cf. [24]) so that the optimal stopping boundary f given in (4.30) admits no closed-form expression (as is seen from (4.29) above). The same conclusion extends to all $n \geq 3$.

We assume throughout that X is a symmetric Walsh spider process on n edges and we consider the problem (4.16) and its solution presented in Theorem 1 above for $n \geq 1$.

Example 1. Consider the case $n = 1$. Adding an extra edge E_0 with zero leak to E_1 we see that the problem of finding $s_1 \mapsto \bar{f}(s_1)$ with values in E_0 reduces to solving a two-dimensional optimal stopping problem for a skew Brownian motion in $I\!R$ with the skewness parameter $\beta = 1$ for $[0, \infty)$ at 0. From the paragraph following (4.28) with $i = 1$ we know that the optimal stopping boundary $s_1 \mapsto g(s_1)$ in this problem is given by (4.28) for $s_1 \geq 1/2c$ and is expressed through its inverse (4.29) on $[0, 1/2c]$ where $\kappa = (1-\beta)/\beta = 0$ in the limit. Denoting the value in (4.29) by $s_1(g; \kappa)$ it is readily verified that

$$s_1(g) := \lim_{\kappa \downarrow 0} s_1(g; \kappa) = \frac{1}{2c} - cg^2 \tag{5.1}$$

for $g \in [-1/c\sqrt{2}, 0]$ where $s_1(-1/c\sqrt{2}) = 0$ (see the first graph in Figure 9). Recalling (4.30) we see that this yields $f(s_1) = (1/\sqrt{c})\sqrt{1/2c - s_1}$ for $s_1 \in [0, 1/2c]$ (see the second graph in Figure 9). In particular, we have $\bar{f}(0) = 1/c\sqrt{2}$ so that (4.15) with $c = 1$ implies that

$$D_1 = 2\bar{f}(0) = \sqrt{2}\,. \tag{5.2}$$

This is a known result of [7], [14], [17]. Note that one can also derive (5.2) more directly by the maximality principle without adding an extra edge with zero leak.

Example 2. Consider the case $n = 2$. Adding an extra edge E_0 with zero leak to $E_1 \coprod E_2$ we see that the problem of finding $(s_1, s_2) \mapsto \bar{f}(s_1, s_2)$ with values in E_0 reduces to solving the integral equation (4.13) that reads

$$\frac{1}{4c}\Big(\frac{1}{s_1^*}+\frac{1}{s_2^*}\Big) - c\big(s_1^*+s_2^*\big) + \log\Big(\frac{s_1^* s_2^*}{s_1 s_2}\Big) + c\int_{s_1}^{s_1^*} \frac{\bar{f}^2(s_1', s_2)}{(s_1')^2}\, ds_1' + c\int_{s_2}^{s_2^*} \frac{\bar{f}^2(s_1, s_2')}{(s_2')^2}\, ds_2' \tag{5.3}$$

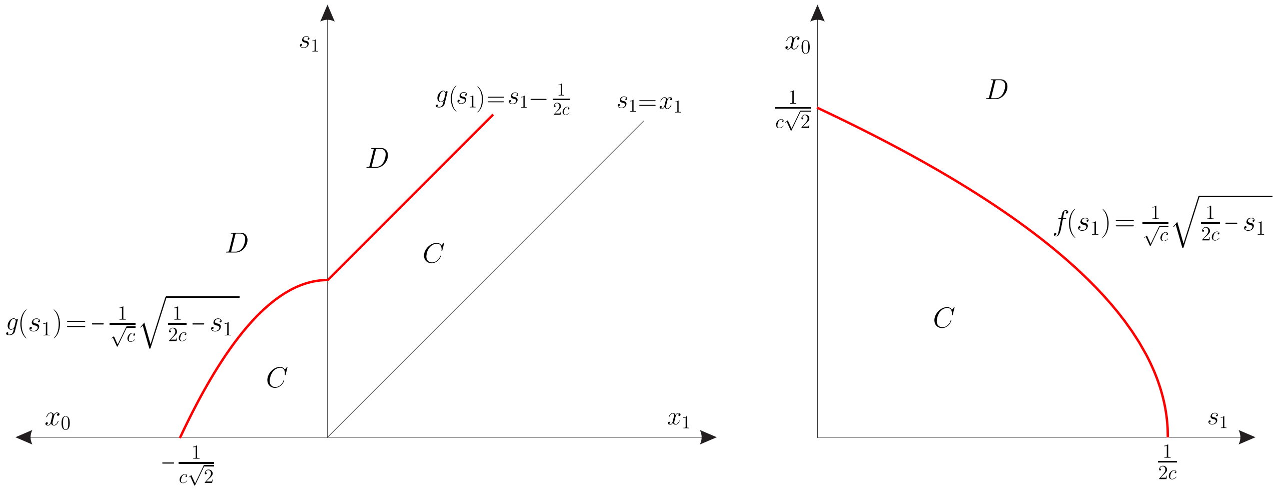

Figure 9. The optimal stopping boundary $s_1 \mapsto g(s_1)$ from (4.29) in the two-dimensional problem for a skew Brownian motion in $I\!R$ with a skewness parameter $\beta \uparrow 1$ for $[0,\infty)$ at 0 (left graph), and the optimal stopping boundary $s_1 \mapsto f(s_1)$ from (4.30) appearing in the final step (4.9) of the pruning procedure in Theorem 1 (right graph). Both graphs show the limiting case $\beta = 1$ only (see Example 1) while the graphs in all prelimiting cases look similar.

$$= c\Big(\frac{1}{s_1}+\frac{1}{s_2}\Big)\bar{f}^2(s_1,s_2)$$

for $0 \le s_1 \le s_1^*$ and $0 \le s_2 \le s_2^*$ where

$$f(s_1) = \frac{1}{2c} - s_1 \;\;\&\;\; f(s_2) = \frac{1}{2c} - s_2 \tag{5.4}$$

for $0 \le s_1 \le 1/2c$ and $0 \le s_2 \le 1/2c$ (because there is no skewness at 0 due to zero leak) so that $s_1^* = f(s_2) + 1/2c = 1/c - s_2$ for $0 \le s_2 \le 1/c$ and $s_2^* = f(s_1) + 1/2c = 1/c - s_1$ for $0 \le s_1 \le 1/c$. Note that the domain $D(\bar{f})$ of $\bar{f}$ equals the entire triangle $\{\,(s_1,s_2) \in [0,1/c]\times[0,1/c] \mid s_1+s_2 \le 1/c\,\}$ i.e. its (upper) boundary is not curved.

The equation (5.3) can be solved explicitly. For this, note that by applying $\partial^2/\partial s_1 \partial s_2$ to both sides in (5.3) we find that

$$\frac{\partial^2 \bar{f}^2}{\partial s_1 \partial s_2}(s_1,s_2) = 0 \tag{5.5}$$

(cf. (4.50) above) so that $\bar{f}^2(s_1,s_2) = h_1(s_1)+h_2(s_2)$ for $(s_1,s_2) \in D(\bar{f})$ with some functions h_1 and h_2. Since $\bar{f}^2$ is permutationally invariant (because f is so) and $\bar{f}^2(1/2c,1/2c) = 0$, we can conclude that $h_1 = h_2 =: h$. Inserting $\bar{f}^2(s_1,s_2) = h(s_1)+h(s_2)$ into (5.3) above, and separating variables, we find that

$$\frac{1}{4c}\Big(\frac{1}{s_j^*}\Big) - c s_j^* + \log\Big(\frac{s_j^*}{s_i}\Big) + c\int_{s_i}^{s_j^*} \frac{h(s')}{(s')^2}\,ds' - c\,\frac{h(s_i)}{s_j^*} = c\,\frac{h(s_i)}{s_i} \tag{5.6}$$

for $i \ne j$ in $\{1,2\}$ with $(s_1,s_2) \in D(\bar{f})$. Replacing either s_1 or s_2 by s and applying $\partial/\partial s$ to both sides in (5.6) we find that

$$h'(s) = s - \frac{1}{c} \tag{5.7}$$

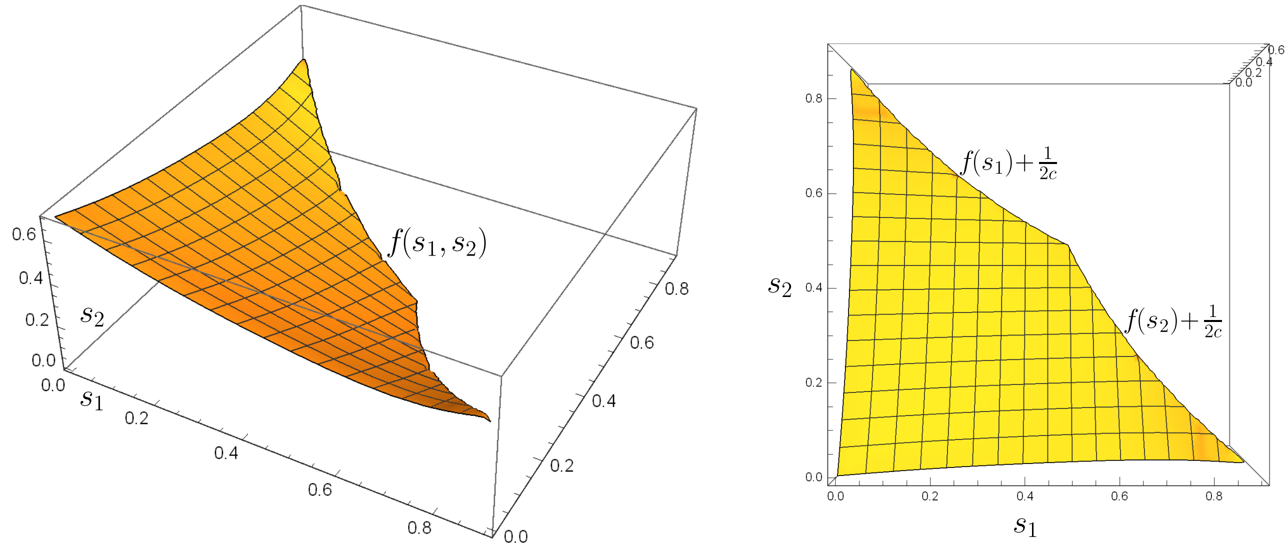

Figure 10. The optimal stopping boundary $(s_1, s_2) \mapsto f(s_1, s_2)$ solving the nonlinear integral equation (5.12) uniquely in the case $n = 3$ with $c = 1$ (see Example 3). The graph on the left shows a side view and the graph on the right shows a top view. The function $s \mapsto f(s)$ from Figure 9 determines the domain boundary of $(s_1, s_2) \mapsto f(s_1, s_2)$ and its boundary values (see the sentence containing (5.13) and (5.14) in Example 3).

for $0 \le s \le 1/2c$. The unique solution to (5.7) satisfying $h(1/2c) = 0$ is given by

$$h(s) = \frac{s^2}{2} - \frac{s}{c} + \frac{3}{8c^2} \tag{5.8}$$

for $0 \le s \le 1/2c$. Setting $c = 1$ this gives $\bar{f}(0, 0) = 2h(0) = 3/4$ and (4.15) implies that

$$D_2 = 2\bar{f}(0, 0) = \sqrt{3}\,. \tag{5.9}$$

This is a known result of [6] (see also [21] for a different derivation).

Example 3. Consider the case $n = 3$. Adding an extra edge E_0 with zero leak to $E_1 \coprod E_2 \coprod E_3$ we see that the problem of finding $(s_1, s_2, s_3) \mapsto \bar{f}(s_1, s_2, s_3)$ with values in E_0 reduces to solving the integral equation (4.13) that reads

$$\begin{aligned} &\frac{1}{4c}\Big(\frac{1}{s_1^*}+\frac{1}{s_2^*}+\frac{1}{s_3^*}\Big) - c\big(s_1^*+s_2^*+s_3^*\big) + \log\Big(\frac{s_1^* s_2^* s_3^*}{s_1 s_2 s_3}\Big) \\ &\quad + c\int_{s_1}^{s_1^*} \frac{\bar{f}^2(s_1', s_2, s_3)}{(s_1')^2}\, ds_1' + c\int_{s_2}^{s_2^*} \frac{\bar{f}^2(s_1, s_2', s_3)}{(s_2')^2}\, ds_2' + c\int_{s_3}^{s_3^*} \frac{\bar{f}^2(s_1, s_2, s_3')}{(s_3')^2}\, ds_3' \\ &\qquad = c\Big(\frac{1}{s_1}+\frac{1}{s_2}+\frac{1}{s_3}\Big)\bar{f}^2(s_1, s_2, s_3) \end{aligned} \tag{5.10}$$

for $0 \le s_1 \le s_1^*$, $0 \le s_2 \le s_2^*$ and $0 \le s_3 \le s_3^*$, where setting $s := (s_1, s_2, s_3)$ we either let

$$s_i^* = f(\bar{s}_i) + \frac{1}{2c} \quad \text{if } f(\bar{s}_i) \in E_i \tag{5.11}$$

i.e. if $\bar{s}_i$ belongs to the domain of f , or we set $s_j := \max \bar{s}_i$ and let s_i^* be the largest value of s_i in $[0,1/2c]$ for which $s_j \le f(\bar{s}_j)+1/2c$ for $1 \le i \le 3$ (recall the second paragraph in Section 4 above for the notation), and we have $f(s_1^*,s_2^*,s_3^*) = f(s_i^*,s_j^*)$ where s_k^* is the largest among s_1^*,s_2^*,s_3^* for different $i,j,k \in \{1,2,3\}$. The function $(s_1,s_2) \mapsto f(s_1,s_2)$ used in (5.11) and afterwards to specify the (upper) boundary of the domain of $\bar{f}$ (as well as the values of $\bar{f}$ at the boundary) solves (uniquely) the integral equation (4.10) that reads

$$\begin{aligned} \text{(5.12)} \qquad & \frac{1}{4c}\Big(\frac{1}{s_1^*}+\frac{1}{s_2^*}\Big) - c\big(s_1^*+s_2^*\big) + \log\Big(\frac{s_1^* s_2^*}{s_1 s_2}\Big) + c\int_{s_1}^{s_1^*} \frac{f^2(s_1',s_2)}{(s_1')^2}\, ds_1' + c\int_{s_2}^{s_2^*} \frac{f^2(s_1,s_2')}{(s_2')^2}\, ds_2' \\ & = c\Big(\frac{1}{s_1}+\frac{1}{s_2}\Big) f^2(s_1,s_2) + 2cf(s_1,s_2) \end{aligned}$$

for $0 \le s_1 \le s_1^*$ and $0 \le s_2 \le s_2^*$, where setting $s := (s_1,s_2)$ we either let

$$\text{(5.13)} \qquad s_1^* = f(s_2)+\frac{1}{2c} \;\;\&\;\; s_2^* = f(s_1)+\frac{1}{2c}$$

for $0 \le s_2 \le 1/2c$ and $0 \le s_1 \le 1/2c$, or we let

$$\text{(5.14)} \qquad s_1^* = f^{-1}\Big(s_2-\frac{1}{2c}\Big) \;\;\&\;\; s_2^* = f^{-1}\Big(s_1-\frac{1}{2c}\Big)$$

for $1/2c \le s_2 \le f(0)+1/2c$ and $1/2c \le s_1 \le f(0)+1/2c$, and we have $f(s_1^*,s_2^*) = f(s_i^*)$ where s_j^* is the larger of s_1^* and s_2^* for different $i,j \in \{1,2\}$. The function $s_1 \mapsto f(s_1)$ used in (5.13)+(5.14) and afterwards to specify the (upper) boundary of the domain of f (as well as the values of f at the boundary) is given by (4.30) above (where g is the inverse function of the function from (4.29) above). Note that the domain $D(f)$ of $(s_1,s_2) \mapsto f(s_1,s_2)$ equals $\{\, (s_1,s_2) \in [0,1/c]\times[0,1/c] \mid s_1 \in [0,1/2c]$ with $s_2 \le f(s_1)+1/2c$ or $s_2 \in [0,1/2c]$ with $s_1 \le f(s_2)+1/2c \,\}$, i.e. its (upper) boundary is curved (see the second graph in Figure 10). The curvature of the domain $D(\bar{f})$ of $(s_1,s_2,s_3) \mapsto \bar{f}(s_1,s_2,s_3)$ can be visualised using (5.11) (see the first graph in Figure 10). Recall that $(s_1,s_2) \mapsto f(s_1,s_2)$ and $(s_1,s_2,s_3) \mapsto \bar{f}(s_1,s_2,s_3)$ are permutationally invariant so that it is sufficient to determine the values of f and $\bar{f}$ on the generating sets consisting of $(s_1,s_2) \in D(f)$ with $s_1 \le s_2$ and $(s_1,s_2,s_3) \in D(\bar{f})$ with $s_1 \le s_2 \le s_3$ respectively.

The integral equations (5.10) and (5.12) can be solved numerically (in the reversed order) using a classic *quadrature method* that generally approximates each integral in the equation by a weighted sum of function values (a quadrature rule) to be determined by solving a system of algebraic equations. Using only a rudimentary knowledge of numerical analysis we have implemented such an algorithm using a simple endpoint quadrature rule for the equidistant mesh of the domain. The algorithm starts by making use of the known (curved) boundary values, and then continues by producing approximate values of f and $\bar{f}$ further away from the boundary and deeper inside the generating sets of their domains addressed above, until the values $f(0,0)$ and $\bar{f}(0,0,0)$ are finally reached. Only a single quadratic equation needs to be solved in each step for the (nonlinear) equation (5.12) and no algebraic equation needs to be solved for the (linear) equation (5.10). The equation (5.12) needs to be solved first. The numerical solution f found is then used as a boundary specification/condition to solve the equation (5.10) for $\bar{f}$. The latter task is more demanding because the equation (5.10) is three-dimensional while the equation (5.12) is two-dimensional. The quadrature method produces a stable and accurate algorithm in both cases while the speed of calculation is much slower for (5.10) than (5.12). Once $\bar{f}(0,0,0)$ is obtained with $c=1$ we know by (4.15) that

$$D_3 = 2\bar{f}(0,0,0) \tag{5.15}$$

approximately (the finer the mesh the more accurate the approximation). The following numerical approximations for $\bar{f}(0,0,0)$ were obtained using the quadrature algorithm described above for $c=1$ with the number of one-dimensional nodes in the equidistant mesh given in the brackets: 0.925505 (100); 0.952463 (500); 0.955386 (1,000); 0.956967 (1,500); 0.957801 (2,000). For 2,000 one-dimensional nodes using (5.15) this gives

$$D_3 = 1.91\ldots \tag{5.16}$$

where it took just over two minutes to solve the equation (5.12), and over seventy-seven hours to solve the equation (5.10), both using the quadrature algorithm described above (optimised by Wolfram Mathematica and run on a high-performance computing cluster). Developing faster and more efficient methods/algorithms for solving the *nested* integral equations (4.10) and (4.13) in higher dimensions to approximate the remaining Dubins constants $D_4, D_5, \ldots$ presents a challenging topic for further study.

References

[1] Atar, R. *and* Cohen, A. (2019). Serve the shortest queue and Walsh Brownian motion. *Ann. Appl. Probab.* 29 (613–651).

[2] Barlow, M. Pitman, J. *and* Yor, M. (1989). On Walsh's Brownian motions. *Sém. de Probab.* XXIII, *Lecture Notes in Math.* 1372, Springer (275–293).

[3] Beckenbach, E. F. *and* Bellman, R. (1961). *Inequalities.* Springer.

[4] Bednarz, E. Ernst, P. A. *and* Osękowski, A. (2024). On the diameter of the stopped spider process. *Math. Oper. Res.* 49 (346–365).

[5] De Angelis, T. *and* Peskir, G. (2020). Global C^1 regularity of the value function in optimal stopping problems. *Ann. Appl. Probab.* 30 (1007–1031).

[6] Dubins, L. E. Gilat, D. *and* Meilijson, I. (2009). On the expected diameter of an L_2-bounded martingale. *Ann. Probab.* 37 (393–402).

[7] Dubins, L. E. *and* Schwarz, G. (1988). A sharp inequality for sub-martingales and stopping times. *Astérisque* 157–158 (129–145).

[8] Dubins, L. E. Shepp, L. A. *and* Shiryaev, A. N. (1993). Optimal stopping rules and maximal inequalities for Bessel processes. *Theory Probab. Appl.* 38 (226–261).

[9] Ernst, P. (2016) Exercising control when confronted by a (Brownian) spider. *Oper. Res. Lett.* 44 (487–490).

[10] Ernst, P. (2025) Exercising control when confronted by a (Brownian) spider: part II. *Queueing Syst.* 109, Issue 2, Article 14 (20 pp).

[11] Evans, S. N. *and* Sowers, R. B. (2003). Pinching and twisting Markov processes. *Ann. Probab.* 31 (486–527).

[12] Fitzsimmons, P. J. *and* Kuter, K. E. (2014). Harmonic functions on Walsh's Brownian motion. *Stochastic Process. Appl.* 124 (2228–2248).

[13] Freidlin, M. *and* Sheu, S-J. (2000). Diffusion processes on graphs: stochastic differential equations, large deviation principle. *Probab. Theory Related Fields* 116 (181–220).

[14] Gilat, D. (1988). On the ratio of the expected maximum of a martingale and the L_p-norm of its last term. *Israel J. Math.* 63 (270–280).

[15] Graversen, S. E. *and* Peskir, G. (1998). Optimal stopping and maximal inequalities for geometric Brownian motion. *J. Appl. Probab.* 35 (856–872).

[16] Hajri, H. *and* Touhami, W. (2014). Itô's formula for Walsh's Brownian motion and applications. *Statist. Probab. Lett.* 87 (48–53).

[17] Jacka, S. D. (1991). Optimal stopping and best constants for Doob-like inequalities I: The case $p = 1$. *Ann. Probab.* 19 (1798–1821).

[18] Karatzas, I. *and* Yan, M. (2019). Semimartingales on rays, Walsh diffusions, and related problems of control and stopping. *Stochastic Process. Appl.* 129 (1921–1963).

[19] Lempa, J. Mordecki, E. *and* Salminen, P. (2024). Diffusion spiders: green kernel, excessive functions and optimal stopping. *Stochastic Process. Appl.* 167, 104229 (34 pp).

[20] Peskir, G. (1998). Optimal stopping of the maximum process: The maximality principle. *Ann. Probab.* 26 (1614–1640).

[21] Peskir, G. (2014). Quickest detection of a hidden target and extremal surfaces. *Ann. Appl. Probab.* 24 (2340–2370).

[22] Peskir, G. *and* Shiryaev, A. N. (2006). *Optimal Stopping and Free-Boundary Problems.* Lectures in Mathematics, ETH Zürich, Birkhäuser.

[23] Sprecher, D. A. (1965). On the structure of continuous functions of several variables. *Trans. Amer. Math. Soc.* 115 (340–355).

[24] Tsirelson, B. (1997). Triple points: from non-Brownian filtrations to harmonic measures. *Geom. Funct. Anal.* 7 (1096–1142).

[25] Walsh, J. B. (1978). A diffusion with discontinuous local time. *Astérisque* 52–53 (37–45).

[26] Walter, W. (1970). *Differential and Integral Inequalities.* Springer.

Goran Peskir
Department of Mathematics
The University of Manchester
Oxford Road
Manchester M13 9PL
United Kingdom
goran@maths.man.ac.uk