Minimum curvature flow and martingale exit times*

Martin Larsson†   Johannes Ruf‡

Abstract

We study the following question: What is the largest deterministic amount of time T_* that a suitably normalized martingale X can be kept inside a convex body K in $\mathbb{R}^d$? We show, in a viscosity framework, that T_* equals the time it takes for the relative boundary of K to reach $X(0)$ as it undergoes a geometric flow that we call (positive) minimum curvature flow. This result has close links to the literature on stochastic and game representations of geometric flows. Moreover, the minimum curvature flow can be viewed as an arrival time version of the Ambrosio–Soner codimension-$(d-1)$ mean curvature flow of the 1-skeleton of K. Our results are obtained by a mix of probabilistic and analytic methods.

1 Introduction and main results

Let $d \geq 2$ and let $K \subset \mathbb{R}^d$ be a convex body, i.e. a nonempty compact convex set. If $X = (X_1, \ldots, X_d)$ is a d-dimensional continuous martingale that starts inside K and whose quadratic variation satisfies $\mathrm{tr}\langle X \rangle(t) = \langle X_1 \rangle(t) + \cdots + \langle X_d \rangle(t) \equiv t$, then X eventually leaves K. What is the maximal *deterministic* lower bound T_* on the exit time, across all such martingales X? The answer is linked to the evolution of the (relative) boundary of K as it undergoes a geometric flow that we refer to as *minimum curvature flow:* T_* is equal to the lifetime of this flow. The minimum curvature flow resembles the well-known mean curvature flow, in particular its version in codimension $d-1$ introduced by [2]. Our goal is to develop the connection between the exit time problem and the minimum curvature flow in detail.

†Department of Mathematical Sciences, Carnegie Mellon University, Wean Hall, 5000 Forbes Ave, Pittsburgh, Pennsylvania 15213, USA. E-mail: martinl@andrew.cmu.edu
‡Department of Mathematics, London School of Economics and Political Science, Columbia House, London, WC2A 2AE, UK. E-mail: j.ruf@lse.ac.uk

Our original motivation comes from a long-standing problem in mathematical finance, namely to characterize the worst-case time horizon for so-called *relative arbitrage*. In a suitably normalized setup, the answer turns out to be precisely T_*, with K being the standard d-simplex. We do not discuss this connection further here; instead we provide full details in the companion paper [26]. Let us however emphasize that this application motivates us to consider convex bodies K with nonsmooth boundary.

To give a precise description of our main results, let X denote the coordinate process on the Polish space $\Omega = C(\mathbb{R}_+, \mathbb{R}^d)$ of all continuous trajectories in $\mathbb{R}^d$ with the locally uniform topology. Thus $X(t, \omega) = \omega(t)$ for all $\omega \in \Omega$ and $t \in \mathbb{R}_+$. Write $\mathcal{P}(\Omega)$ for the set of all probability measures on Ω with the topology of weak convergence. For each $x \in \mathbb{R}^d$, define

$$\mathcal{P}_x = \{\mathsf{P} \in \mathcal{P}(\Omega) \colon X \text{ is a P-martingale and } \mathsf{P}(X(0) = x) = \mathsf{P}(\mathrm{tr}\langle X\rangle(t) \equiv t) = 1\},$$

where the martingale property is understood with respect to the (raw) filtration generated by X. We always take $K \subset \mathbb{R}^d$ to be compact, but not necessarily convex unless explicitly stated. The first exit time from K is

$$\tau_K = \inf\{t \geq 0 \colon X(t) \notin K\}, \tag{1.1}$$

and we are interested in computing the value function

$$v(x) = \sup_{\mathsf{P} \in \mathcal{P}_x} \mathsf{P}\text{-}\mathrm{ess\,inf}\, \tau_K. \tag{1.2}$$

This is the largest deterministic almost sure lower bound on the exit time τ_K across all martingale laws $\mathsf{P} \in \mathcal{P}_x$. At first glance one may suspect that v is identically zero. This is however not the case for $d \geq 2$; see for instance Example 1.8.

Our first result states that the value function solves a PDE with (degenerate) elliptic nonlinearity

$$F(p, M) = \inf\left\{-\frac{1}{2}\,\mathrm{tr}(aM) \colon a \succeq 0,\ \mathrm{tr}(a) = 1,\ ap = 0\right\}, \tag{1.3}$$

where a ranges through the set of all symmetric matrices of appropriate size, and $a \succeq 0$ refers to the positive semidefinite order. We write $\mathbb{S}^d$ for the set of all $d \times d$ symmetric matrices and $\mathbb{S}^d_+$ for the subset of positive semidefinite matrices. The theorem uses the notions of viscosity solution and boundary condition in the viscosity sense. This is because in general, the boundary condition is not satisfied pointwise, see e.g. Example 5.4, and it is not even known whether in the interior of K the solution is continuous. These concepts are reviewed in Section 3, where also the proof of the theorem is given.

Theorem 1.1. *Let $d \geq 2$ and suppose K is compact, but not necessarily convex. The value function v is an upper semicontinuous viscosity solution to the nonlinear equation*

$$F(\nabla u, \nabla^2 u) = 1 \tag{1.4}$$

in $\mathrm{int}(K)$ with zero boundary condition (in the viscosity sense).

The value function is always an *upper semicontinuous* viscosity solution. As our next result shows, it is actually the unique viscosity solution in this class, provided that K satisfies a certain additional condition. This condition holds for all strictly star-shaped compact sets, in particular for all convex bodies with nonempty interior. Our condition is however more general than that; see Example 4.2. We also show that uniqueness may fail for star-shaped but not strictly star-shaped domains; see Example 4.3. This answers a question of [24, Section 1.8]. The proof of the following uniqueness theorem is given in Section 4, and follows from a comparison principle proved there, Theorem 4.1.

Theorem 1.2. *Let $d \geq 2$ and suppose K is compact. Assume there exist invertible affine maps T_λ on $\mathbb{R}^d$, parameterized by $\lambda \in (0,1)$, such that $T_\lambda(K) \subset \operatorname{int}(K)$ and $\lim_{\lambda \to 1} T_\lambda = I$ (the identity). Then the value function v is the unique upper semicontinuous viscosity solution to (1.4) in $\operatorname{int}(K)$ with zero boundary condition (in the viscosity sense).*

Remark 1.3. We point out that this uniqueness result is designed to handle the non-smooth convex domains that arise in the financial applications of interest. There are however other natural domains that are not covered by this result, such a various non-convex domains with smooth boundary. Proving comparison theorems (and hence uniqueness results) for such domains is an interesting problem which we do not consider here; see however [36, 37, 5, 4].

Theorem 1.2 characterizes the value function even in cases where it is not continuous. In fact, we will give examples showing that the value function may be discontinuous even when K is a convex body.

Before describing this and related results, we briefly discuss links to the existing literature and the connection to geometric flows.

Our results tie in with a well established literature on stochastic representations of geometric PDEs, initiated by [11] and [33, 34, 35]. In particular, Soner and Touzi introduced the notion of *stochastic target problem* and based their analysis on an associated dynamic programming principle; see also [9]. Part of our analysis can be cast in the language of stochastic target problems, and this connection is described further in Remark 2.6.

The control problem (1.2) is formulated over an infinite time horizon. As a result, our PDE is elliptic rather than parabolic, and, as explained next, the solution acquires the interpretation of *arrival time* of an evolving surface. This is reminiscent of the two-person deterministic game introduced by [39] and linked to the positive curvature flow by [24]. In a similar spirit there are also the works of [29] on the tug-of-war game and infinity Laplacian, and more recently [18] and [12].

The geometric meaning of (1.4) is most clearly conveyed by reasoning as in [24, Section 1.2]. This is standard in the literature on geometric flows and paraphrased here for convenience. Let K be strictly convex with smooth boundary ∂K. Suppose we are given a family $\{\Gamma_t : t \geq 0\}$ of smooth convex surfaces with $\Gamma_0 = \partial K$, that evolve with normal velocity equal to (half) the smallest principal curvature at each point $x \in \Gamma_t$. It is natural to call this *minimum curvature flow*, by analogy with mean curvature flow whose normal velocity is the average curvature.

Let u be the *arrival time function:* for each $x \in K$, $u(x)$ is the time it takes the evolving front to reach x (we assume the front passes through each point in K exactly once.) Thus $\Gamma_t = \{x : u(x) = t\}$ is a level surface of u, and the gradient $\nabla u(x)$ is a normal vector at x. If $\nabla u(x) \neq 0$, the minimal principal curvature of Γ_t at x is the smallest value of

$$-\frac{y^\top \nabla^2 u(x) y}{|\nabla u(x)|}$$

as y ranges over all tangent unit vectors: $|y| = 1$ and $y^\top \nabla u(x) = 0$.[1] On the other hand, since $u(x)$ is the arrival time, the speed of normal displacement at x is $1/|\nabla u(x)|$. We therefore expect u to satisfy

$$\inf\left\{-\frac{1}{2} y^\top \nabla^2 u(x) y : |y| = 1,\ y^\top \nabla u(x) = 0\right\} = 1, \tag{1.5}$$

---

[1]Indeed, if $\gamma \colon \mathbb{R} \to \Gamma_t$ is a smooth geodesic curve with unit speed such that $\gamma(0) = x$ and $\gamma'(0) = y$, then $\nabla u(x)^\top \gamma''(0) + y^\top \nabla^2 u(x) y = 0$ and $\gamma''(0) = k \frac{\nabla u(x)}{|\nabla u(x)|}$, where k is the curvature of γ at 0.

at least at points where $\nabla u \neq 0$. It is not hard to check that this is precisely (1.4) thanks to the identity

$$\sup\{\operatorname{tr}(aM)\colon a \succeq 0, \operatorname{tr}(a) = 1, ap = 0\} = \sup\{y^\top My\colon |y| = 1, y^\top p = 0\}.$$

To verify this identity it suffices to show that the supremum on the left-hand side is attained by a rank-one matrix. For any candidate a the spectral theorem yields $a = \lambda_1 y_1 y_1^\top + \cdots + \lambda_d y_d y_d^\top$ for some orthonormal basis $y_1, \ldots, y_d$ and convex weights $\lambda_1, \ldots, \lambda_d$ such that $\lambda_i y_i^\top p = 0$ for all $i \in \{1, \ldots, d\}$. Since $\operatorname{tr}(aM) = \lambda_1 y_1^\top My_1 + \cdots + \lambda_d y_d^\top My_d$, we may improve by replacing a with the rank-one matrix $y_i y_i^\top$ for the index i such that $y_i^\top My_i$ is maximal.

In the planar case $d = 2$, Γ_t has only one principal curvature direction, and (1.5) reduces to the well-known arrival time PDE for the mean curvature flow,

$$\frac{1}{|\nabla u|} = -\frac{1}{2}\operatorname{div}\left(\frac{\nabla u}{|\nabla u|}\right).$$

Remark 1.4. Let us outline how the minimum curvature flow can be constructed rigorously using the *level set method* of [27, 13, 19] and then linked to (1.2) and (1.4). Fix a time horizon $T > \max_{x \in K} v(x)$ and consider the geometric parabolic equation

$$\partial_t U + F(\nabla U, \nabla^2 U) = 0 \text{ in } (0, T] \times \mathbb{R}^d$$

with an initial condition $U_0(x)$ that is positive on $\operatorname{int}(K)$, negative on K^c, and constant, say equal to -1, outside some large compact set. [13, Theorems 6.7–6.8] yields existence and uniqueness of a continuous solution $U(t, x)$ of the initial value problem. One now defines the evolving front of the minimum curvature flow at time t to be the boundary of the superlevel set, $\partial\{x\colon U(t, x) > 0\}$. The time $u(x) = \inf\{t\colon U(t, x) < 0\}$ at which the front passes through $x \in K$ can then, under suitable conditions, be shown to be an upper semicontinuous viscosity solution of the elliptic equation (1.4). If uniqueness holds for this equation, for instance if Theorem 1.2 is applicable, it follows that u actually coincides with the value function v in (1.2).

The link to the control problem (1.2) can be understood as follows. Proceeding informally, we assume a C^2 solution u of (1.4) with $u = 0$ on ∂K is given. By Itô's formula,

$$0 = u(X(\tau_K)) = u(x) + \int_0^{\tau_K} \nabla u(X(t))^\top dX(t) + \frac{1}{2}\int_0^{\tau_K} \operatorname{tr}(a(t)\nabla^2 u(X(t)))dt \qquad (1.6)$$

under any law $\mathsf{P} \in \mathcal{P}_x$, where $a(t)$ is the derivative of the quadratic variation of X and satisfies $\operatorname{tr}(a(t)) \equiv 1$. The discussion of minimum curvature flow suggests that optimally, X should fluctuate tangentially to the level surfaces of u, that is, $a(t)\nabla u(X(t)) \equiv 0$. Then, due to the definition (1.3) of F and since u solves (1.4),

$$\frac{1}{2}\operatorname{tr}(a(t)\nabla^2 u(X(t))) \leq -F(\nabla u(X(t)), \nabla^2 u(X(t))) = -1. \qquad (1.7)$$

Combining (1.6) and (1.7) leads to

$$0 = u(x) + \frac{1}{2}\int_0^{\tau_K} \operatorname{tr}(a(t)\nabla^2 u(X(t)))dt \leq u(x) - \tau_K,$$

showing that $\tau_K \leq u(x)$. If $a(t)$ maximizes the left-hand side of (1.7), we have equality and expect that u coincides with the value function. Still heuristically, this happens when X fluctuates only along the minimal principal curvature directions of the level surfaces

of u. This minimizes the speed at which X moves "outwards" toward ∂K, and maximizes the amount of time X spends in K.

This discussion suggests that optimally, X lies on the evolving front of the time-reversed minimum curvature flow. More precisely, before exiting K, one expects that X satisfies $v(X(t)) = v(x) - t$ under some optimal law $\mathsf{P} \in \mathcal{P}_x$, $x \in K$. Theorem 1.7 below shows that this is true if K is a polytope and v sufficiently regular. It is however false in general, even if v is smooth; see Example 2.3.

In the case where K is not convex, we get a somewhat different flow. Similarly to the positive curvature flow of [24], it is now the positive part of the minimum principal curvature that determines the speed of the flow.

We now return to our main results, and focus on the case where K is a convex body. Theorems 1.1 and 1.2 yield upper semicontinuity of the value function v and characterize it as a viscosity solution of (1.4) with zero boundary condition (in the viscosity sense). If K has empty interior we simply apply these results in the affine hull of K. The following result is a combination of Proposition 5.1 and Lemma 5.3 in Section 5.

Theorem 1.5. *Let $d \geq 2$ and suppose K is a convex body. Then the value function v is quasi-concave, vanishes on all faces of K of dimension zero and one, and is strictly positive elsewhere in K.*

In particular, if K is strictly convex, then all its boundary faces have dimension zero, and v vanishes everywhere on ∂K. Because of upper semicontinuity, this implies that it is continuous at ∂K. In fact, Theorem 1.6 below shows that v is continuous everywhere in this case.

However, many convex bodies K have boundary faces of higher dimension. In this case v does *not* vanish everywhere on ∂K. This includes the standard d-simplex appearing in our motivating financial application. Additionally, and more subtly, there are convex bodies for which the value function is actually *discontinuous*. This is because in dimension $d \geq 4$, there are convex bodies that admit boundary points x_n, all contained in 1-dimensional boundary faces, whose limit $\bar{x} = \lim_n x_n$ lies in the relative interior of a 2-dimensional boundary face; see Example 5.4. For such points, $v(x_n) = 0$ but $v(\bar{x}) > 0$, so continuity fails. This is in sharp contrast to the more familiar case of mean curvature flow, where the arrival time function is continuous for any convex initial surface; see [19, Theorem 7.4] and [20, Theorem 5.5].

We prove continuity under the following regularity condition on the geometry of K. We require that the *k-skeletons*, defined by

$$\mathcal{F}_k = \text{union of all faces of } K \text{ of dimension at most } k, \qquad (1.8)$$

be closed for $k = 1, \ldots, d$ (but not for $k = 0$, thus the set of extreme points need not be closed.) This condition is a weakening of a notion from convex geometry called *stability*, which is equivalent to all the k-skeletons being closed, including the 0-skeleton; see e.g. [28] and [32, p. 78]. Actually the d-, $(d-1)$- and $(d-2)$-skeletons of a convex body are always closed, so this does not have to be assumed separately; see Lemma 5.7.

The upshot is the following result, which is applicable in a number of interesting situations. In particular, it covers all convex bodies in $\mathbb{R}^3$, all polytopes in arbitrary dimension, and all convex bodies whose boundary faces all have dimension zero or one. It is a rewording of Theorem 5.8 in Section 5, and is proved using probabilistic arguments based on the control formulation (1.2).

Theorem 1.6. *Let $d \geq 2$ and suppose K is a convex body with $\mathcal{F}_k$ closed for $1 \leq k \leq d-3$. Then the value function v is continuous on K.*

The fact that v vanishes only at the 1-skeleton $\mathcal{F}_1$ (the extreme points and lines), but not elsewhere in K, suggests that (1.4) describes a geometric flow also of $\mathcal{F}_1$, not only of

∂K. This flow of $\mathcal{F}_1$ is the *codimension-$(d-1)$ mean curvature flow* of [2], although here the initial set $\mathcal{F}_1$ need not be a one-dimensional curve.

To spell this out, for any symmetric matrix A and eigenvector p of A, let $\lambda_{\min}(A, p)$ denote the smallest eigenvalue of A corresponding to an eigenvector orthogonal to p. Then (1.5) states that

$$\lambda_{\min}\left(-\frac{1}{2}P_{\nabla u(x)}\nabla^2 u(x)P_{\nabla u(x)}, \nabla u(x)\right) = 1, \tag{1.9}$$

where

$$P_{\nabla u} = I - \frac{\nabla u \nabla u^\top}{|\nabla u|^2}.$$

Modulo sign conventions and the factor $1/2$, the left-hand side of (1.9) is precisely the operator used by [2]. In fact, the function $V(t, x) = t - v(x)$, where v is the value function in (1.2), solves their parabolic equation on K with initial condition $V(0, x) = -v(x)$, whose zero set (in K) is the 1-skeleton $\mathcal{F}_1$. This suggests interpreting the minimum curvature flow of ∂K as a codimension-$(d-1)$ mean curvature flow of $\mathcal{F}_1$.

This perspective is particularly compelling when K is a polytope: $\mathcal{F}_1$ is then a finite union of closed line segments and thus one-dimensional, albeit with "branching". In this case, the one-dimensional initial contour instantly develops higher-dimensional features as it evolves under the flow, and eventually becomes a closed hypersurface. This is illustrated schematically in Figure 1, where K is the standard 3-simplex.
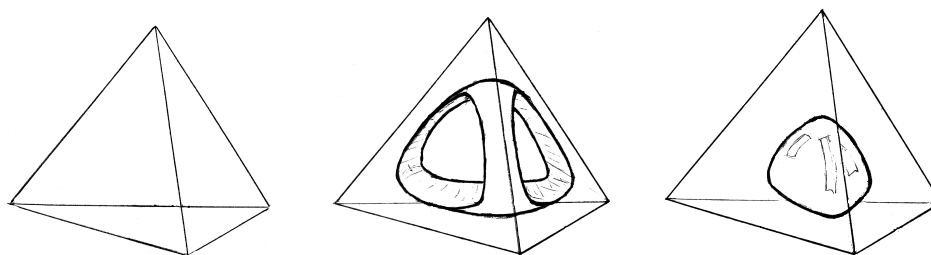

Figure 1: Schematic illustration of the minimum curvature flow of the 3-simplex, regarded as codimension-2 mean curvature flow of its 1-skeleton as initial contour. In the second and third panel, the 1-skeleton is still shown for reference.

Returning to the minimum curvature flow as a flow of surfaces starting from ∂K, we see that points inside two- and higher dimensional faces remain stationary for some period of time. This behavior is analogous to the behavior of mean curvature flow of *non-convex* contours; see [24, Figure 4] for an illustration. We thank R. Kohn for pointing this out to us. A similar phenomenon occurs for the Gauss curvature flow; see [22, 14, 16].

We do not have much information about the regularity of the value function v in general, beyond the continuity assertion in Theorem 1.6 and the counterexample in Example 5.4. An exception is the planar case $d = 2$, where we recover the standard mean curvature flow. In this case, for K strongly convex with smooth boundary, [24] proved that v is C^3 (see also [23] for an earlier proof that v is C^2). In general, let us *assume* that v is C^2 inside each face of K, with just one critical point. If in addition K has at most countably many faces, it is then possible to construct optimal solutions of (1.2) where the intuitive notion that X should fluctuate tangentially to, and remain on, the level surfaces of v becomes rigorous.

Theorem 1.7. *Let $d \geq 2$ and let K be a convex body with at most countably many faces. Assume the value function v lies in $C^2(K)$. Assume also that in each face F of*

dimension at least two, either v has no critical point, or v has one single critical point which additionally is a maximum. Then for every $\bar{x} \in K$ there is an optimal solution $\mathsf{P} \in \mathcal{P}_{\bar{x}}$ under which $v(X(t)) = v(\bar{x}) - t$ for all $t < \tau_K$. In particular,

$$. \int_0^t \nabla v(X(s))^\top dX(s) = 0, \quad t < \tau_K, \tag{1.10}$$

and X lies on the evolving front of the time-reversed minimum curvature flow in the sense that $X(t) \in \Gamma_{v(\bar{x})-t}$, where $\Gamma_t = \{x \colon v(x) = t\}$, until it leaves K.

The meaning of $C^2(K)$ and the notion of a critical point is explained in Section 6, where also the proof is given. The basic idea is to observe that v satisfies (1.9) classically at non-critical points. In particular, by definition of eigenvalue, the matrix

$$H(x) = \frac{1}{2} P_{\nabla v(x)} \nabla^2 v(x) P_{\nabla v(x)} + I$$

is singular at all such points, so is of rank at most $d - 1$. This can be used to construct a martingale law $\mathsf{P} \in \mathcal{P}_{\bar{x}}$ under which $H(X(t))d\langle X\rangle(t) = 0$. This turns out to imply $\nabla v(X(t))^\top dX(t) = 0$ and then $dv(X(t)) = -dt$. This is essentially the desired conclusion. Some effort is needed to construct P, basically because the Moore–Penrose inverse $H(x)^+$ of $H(x)$ is no longer continuous in x. Moreover, X is obtained by constructing martingales on each face of K separately and then "gluing" these martingales together. This introduces some technical hurdles, and explains why the proof is somewhat lengthy.

As an illustration, and for later use, we give a simple example where the value function v is known explicitly and happens to be smooth on K; see also [40] and [21].

Example 1.8. Let $d \geq 2$ and let $K = \{x \in \mathbb{R}^d \colon |x| \leq r\}$ be the centered closed ball of radius $r > 0$. In this case, $v(x) = r^2 - |x|^2$ for all $x \in K$. To see this, choose any $x \in K$ and $\mathsf{P} \in \mathcal{P}_x$. We have

$$|X(t)|^2 = |x|^2 + 2\int_0^t X(s)^\top dX(s) + t, \quad t \geq 0.$$

Evaluating at $t = \tau_K \wedge n$, taking expectations, and letting $n \to \infty$, one obtains $\mathsf{E}[\tau_K] = r^2 - |x|^2$. In particular, this shows that X escapes from any bounded set in finite time, P-a.s. Moreover, since of course $\mathsf{P}\text{-}\operatorname{ess\,inf} \tau_K \leq \mathsf{E}[\tau_K]$, we get $v(x) \leq |x|^2 - r^2$. In fact, we have equality. Indeed, let $\mathsf{P} \in \mathcal{P}_x$ be the law under which $X_3, \ldots, X_d$ are constant and $(X_1, X_2)^\top$ satisfies

$$d \begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} = \frac{1}{\sqrt{X_1(t)^2 + X_2(t)^2}} \begin{pmatrix} X_2(t) \\ -X_1(t) \end{pmatrix} dW(t),$$

where W denotes a one-dimensional Brownian motion. Such a probability measure P always exists, even if $x = 0$; see Lemma 3.4. An application of Itô's formula now yields $\tau_K = r^2 - |x|^2$, P_x-a.s. We deduce that $v(x) = r^2 - |x|^2$ for all $x \in K$. Furthermore, it is straightforward to verify that v satisfies (1.4) with boundary condition $v = 0$ on ∂K.

The reasoning in Example 1.8 directly yields the following upper bound on v.

Lemma 1.9. *If K is compact, $x \in K$, and $\mathsf{P} \in \mathcal{P}_x$, then $\mathsf{E}[\tau_K] \leq r^2$, where r is the radius of the smallest ball containing K. In particular, $\tau_K < \infty$, P-a.s., and the value function defined in (1.2) satisfies $v(x) \leq r^2$ for all x.*

The rest of the paper is organized as follows. Section 2 develops a number of general properties of the value function, as well as illustrative examples. In particular, a dynamic programming principle is proved. In Section 3 we prove Theorem 1.1 that the value function is a viscosity solution. In Section 4 we prove Theorem 4.1, a comparison

principle for viscosity solutions of (1.4), and use it to deduce Theorem 1.2. In Section 5 we focus on the case where K is a convex body, and establish in particular Theorem 5.8 on continuity of the value function. In Section 6 we prove Theorem 1.7.

We end with a technical remark regarding filtrations and stopping times. Whenever X is said to be a martingale, this is understood with respect to its own filtration $\mathbb{F}^X = (\mathcal{F}_t^X)_{t \geq 0}$ where $\mathcal{F}_t^X = \sigma(X_s, s \leq t)$. In this case, X is also a martingale for the right-continuous filtration $\mathbb{F}_+^X$ consisting of the σ-algebras $\bigcap_{u>t} \mathcal{F}_u^X$, and similarly for the filtrations obtained by augmenting $\mathbb{F}^X$ and $\mathbb{F}_+^X$ with nullsets. In particular, results such as the stopping theorem are applicable with τ_K in (1.1), which is an $\mathbb{F}_+^X$-stopping time but not an $\mathbb{F}^X$-stopping time.

2 The value function and dynamic programming

The purpose of this section is to establish a number of properties of the value function, in particular a dynamic programming principle. Throughout this section, K is compact but not necessarily convex.

Lemma 2.1. *The maps $\omega \mapsto \tau_K(\omega)$ from Ω to $[0, \infty]$ and $\mathsf{P} \mapsto \mathsf{P}\text{-}\mathrm{ess\,inf}\,\tau_K$ from $\mathcal{P}(\Omega)$ to $[0, \infty]$ are upper semicontinuous, where τ_K is the first exit time of K, given in (1.1).*

Proof. We claim that $\omega \mapsto \tau_K(\omega)$ is upper semicontinuous on Ω. To see this, let ω_n, ω satisfy $\tau_K(\omega) < \infty$ and $\omega_n \to \omega$ locally uniformly. Consider $\varepsilon > 0$ such that $\omega(\tau_K(\omega) + \varepsilon) \notin K$. Then for all large n, we have $\omega_n(\tau_K(\omega) + \varepsilon) \notin K$, and hence $\tau_K(\omega_n) \leq \tau_K(\omega) + \varepsilon$. Thus $\limsup_n \tau_K(\omega_n) \leq \tau_K(\omega) + \varepsilon$. This proves upper semicontinuity of τ_K since $\varepsilon > 0$ can be chosen arbitrarily small.

Next, for every $\lambda > 0$ the Portmanteau theorem yields that the map

$$\mathsf{P} \mapsto f_\lambda(\mathsf{P}) = -\frac{1}{\lambda} \log \mathsf{E}_\mathsf{P}[e^{-\lambda \tau_K}]$$

from $\mathcal{P}(\Omega)$ to $[0, \infty]$ is upper semicontinuous, and then so is $\mathsf{P} \mapsto \inf_{\lambda > 0} f_\lambda(\mathsf{P})$. We now show that $\inf_{\lambda > 0} f_\lambda(\mathsf{P}) = \mathsf{P}\text{-}\mathrm{ess\,inf}\,\tau_K$. First, $f_\lambda(\mathsf{P}) \geq \mathsf{P}\text{-}\mathrm{ess\,inf}\,\tau_K$ for all $\lambda > 0$. Next, if $\mathsf{P}\text{-}\mathrm{ess\,inf}\,\tau_K = \infty$, the opposite inequality is immediate. If $\mathsf{P}\text{-}\mathrm{ess\,inf}\,\tau_K < \infty$, pick any $c > \mathsf{P}\text{-}\mathrm{ess\,inf}\,\tau_K$ and note that

$$f_\lambda(\mathsf{P}) \leq -\frac{1}{\lambda} \log \mathsf{E}_\mathsf{P}[e^{-\lambda \tau_K} \mathbf{1}_{\tau_K \leq c}] \leq -\frac{1}{\lambda} \log \mathsf{E}_\mathsf{P}[e^{-\lambda c} \mathbf{1}_{\tau_K \leq c}] = c - \frac{1}{\lambda} \log \mathsf{P}(\tau_K \leq c).$$

We now send λ to infinity and then let c decrease to $\mathsf{P}\text{-}\mathrm{ess\,inf}\,\tau_K$. We conclude that $\mathsf{P} \mapsto \mathsf{P}\text{-}\mathrm{ess\,inf}\,\tau_K$ is upper semicontinuous, as claimed. $\qquad \square$

The proof of the following result uses the notion of conditional essential infimum. For a random variable Y and a sub-σ-algebra $\mathcal{G} \subset \mathcal{F}$, the conditional essential infimum of Y given $\mathcal{G}$ is defined as the largest $\mathcal{G}$-measurable random variable P-a.s. dominated by Y, denoted by $\mathsf{P}\text{-}\mathrm{ess\,inf}\{Y \mid \mathcal{G}\}$. Moreover, if $\{F_\omega\}_{\omega \in \Omega}$ is a regular conditional distribution of Y given $\mathcal{G}$, we have $\mathsf{P}\text{-}\mathrm{ess\,inf}\{Y \mid \mathcal{G}\}(\omega) = \mathrm{ess\,inf}\,F_\omega$ for P-a.e. ω, where we set $\mathrm{ess\,inf}\,F_\omega = \sup\{c \in \mathbb{R} \colon F_\omega([c, \infty)) = 1\}$. For further details, see [6, 25].

Proposition 2.2.   (i) *$\mathcal{P}_x$ is weakly compact for every $x \in \mathbb{R}^d$;*

  (ii) *v, given in (1.2), is upper semicontinuous and there is a measurable map $x \mapsto \mathsf{P}_x$ from $\mathbb{R}^d$ into $\mathcal{P}(\Omega)$ such that P_x lies in $\mathcal{P}_x$ and is optimal for all $x \in \mathbb{R}^d$;*

  (iii) *v satisfies the following dynamic programming principle: for every $x \in \mathbb{R}^d$ and every $\mathbb{F}^X$-stopping time θ,*

$$v(x) = \sup_{\mathsf{P} \in \mathcal{P}_x} \mathsf{P}\text{-}\mathrm{ess\,inf}\{\theta \wedge \tau_K + v(X(\theta)) \mathbf{1}_{\theta \leq \tau_K}\}.$$

(Note that $\tau_K < \infty$, P-a.s. for every $\mathsf{P} \in \mathcal{P}_x$ due to Lemma 1.9.) Moreover, the supremum is attained by any optimal $\mathsf{P} \in \mathcal{P}_x$.

Proof. (i): Consider any $\mathsf{P} \in \mathcal{P}_x$. Fix $s \geq 0$ and define $M(t) = |X(t) - X(s)|^2 - t + s$ for $t \geq s$. Then M is a P-martingale on $[s, \infty)$ and, letting $a(t)$ be the derivative of the quadratic variation of X with $\mathrm{tr}(a(t)) \equiv 1$, we have

$$\langle M \rangle(t) = 4 \int_s^t (X(u) - X(s))^\top a(u)(X(u) - X(s)) du \leq 4 \int_s^t |X(u) - X(s)|^2 du.$$

Thus

$$\mathsf{E}_\mathsf{P}[\langle M \rangle(t)] \leq 4 \int_s^t \mathsf{E}_\mathsf{P}[|X(u) - X(s)|^2] du = 4 \int_s^t (u - s) du = 2(t - s)^2,$$

so that

$$\mathsf{E}_\mathsf{P}[|X(t) - X(s)|^4] = \mathsf{E}_\mathsf{P}[(M(t) + t - s)^2] \leq 2\mathsf{E}_\mathsf{P}[\langle M \rangle(t)] + 2(t - s)^2 \leq 6(t - s)^2.$$

Kolmogorov's continuity criterion (see [30], Theorem I.2.1 and its proof) then gives, for any fixed $T > 0$ and $\alpha \in (0, \frac{1}{4})$,

$$\mathsf{E}_\mathsf{P}\left[\left(\sup_{0 \leq s < t \leq T} \frac{|X(t) - X(s)|}{|t - s|^\alpha} \right)^4 \right] \leq c$$

for some constant $c = c(T, \alpha)$ that does not depend on $\mathsf{P} \in \mathcal{P}_x$. Since Hölder balls are relatively compact in $C([0, T], \mathbb{R}^d)$ by the Arzelà–Ascoli theorem, it follows that $\mathcal{P}_x$ is tight and hence relatively compact by Prokhorov's theorem. To see that $\mathcal{P}_x$ is closed, note that the martingale property of both X and $|X|^2 - t$ (and hence the property $\mathrm{tr}\langle X \rangle(t) \equiv t$) carries over to weak limits of sequences in $\mathcal{P}_x$.

(ii): First observe that $\mathcal{P}_x$ consists of the pushforwards $(x + \cdot)_* \mathsf{P}$ with $\mathsf{P} \in \mathcal{P}_0$. Thus $v(x) = \sup_{\mathsf{P} \in \mathcal{P}_0} f(x, \mathsf{P})$, where $f(x, \mathsf{P}) = g((x + \cdot)_* \mathsf{P})$ and $g(\mathsf{P}) = \mathsf{P}\text{-ess inf } \tau_K$. By Lemma 2.1, the function g is upper semicontinuous. Since f is the composition of g with the continuous function $(x, \mathsf{P}) \mapsto (x + \cdot)_* \mathsf{P}$ from $\mathbb{R}^d \times \mathcal{P}(\Omega)$ to $\mathcal{P}(\Omega)$, it is also upper semicontinuous. Moreover, $\mathcal{P}_0$ is compact by (i). A suitable selection theorem, see e.g. [7, Proposition 7.33], yields upper semicontinuity of v as well as a measurable map $x \mapsto \mathsf{Q}_x$ from $\mathbb{R}^d$ into $\mathcal{P}_0$ such that $v(x) = f(x, \mathsf{Q}_x)$ for all $x \in \mathbb{R}^d$. Setting $\mathsf{P}_x = (x + \cdot)_* \mathsf{Q}_x$ gives the required map.

(iii): Fix $x \in \mathbb{R}^d$ and an $\mathbb{F}^X$-stopping time θ. We first first fix $\mathsf{P} \in \mathcal{P}_x$ and prove that

$$v(x) \geq \mathsf{P}\text{-ess inf}\{\theta \wedge \tau_K + v(X(\theta))\mathbf{1}_{\theta \leq \tau_K}\}. \tag{2.1}$$

To this end, consider the extended space $\Omega \times \Omega$ with coordinate process $(X, Y)(t, \omega, \tilde{\omega}) = (\omega(t), \tilde{\omega}(t))$ and define a law P' on $(\Omega \times \Omega, \mathcal{F} \otimes \mathcal{F})$ by $\mathsf{P}'(d\omega, d\tilde{\omega}) = \mathsf{P}_{X(\theta(\omega), \omega)}(d\tilde{\omega})\mathsf{P}(d\omega)$, where we use the measurable map $\mathbb{R}^d \ni y \mapsto \mathsf{P}_y \in \mathcal{P}$ from (ii). We now consider the process $X'(t) = X(t)\mathbf{1}_{t \leq \theta} + Y(t - \theta)\mathbf{1}_{t > \theta}$ and let Q denote the law of X'. Define next $\theta'(\omega, \tilde{\omega}) = \theta(X'(\omega, \tilde{\omega}))$; thus θ' depends on the trajectory of X' like θ depends on the trajectory of X. Since θ is an $\mathbb{F}^X$-stopping time, and since $X'(t)$ and $X(t)$ coincide for all $t \leq \theta$, it follows by Galmarino's test that $\theta'(\omega, \tilde{\omega}) = \theta(\omega)$ for all $(\omega, \tilde{\omega})$; see [17, Theorem IV.100(a)]. Consequently, for all bounded measurable maps $F, G: \Omega \to \mathbb{R}$, we have

$$\begin{aligned} \mathsf{E}_\mathsf{Q}[F(X(\cdot \wedge \theta))G(X(\theta + \cdot))] &= \mathsf{E}_{\mathsf{P}'}[F(X'(\cdot \wedge \theta'))G(X'(\theta' + \cdot))] \\ &= \mathsf{E}_{\mathsf{P}'}[F(X(\cdot \wedge \theta))G(Y)] \\ &= \mathsf{E}_\mathsf{P}[F(X(\cdot \wedge \theta))\mathsf{E}_{\mathsf{P}_{X(\theta)}}[G(X)]]. \end{aligned}$$

Thanks to the definition of Q we have

$$Q \in \mathcal{P}_x, \qquad Q|_{\mathcal{F}_\theta^X} = P|_{\mathcal{F}_\theta^X}. \tag{2.2}$$

Furthermore, with the notation $\tau_K(X(\theta + \cdot)) = \inf\{t \geq 0\colon X(\theta + t) \notin K\}$, one derives the identity

$$\tau_K = \theta \wedge \tau_K + \tau_K(X(\theta + \cdot))\mathbf{1}_{\theta \leq \tau_K}. \tag{2.3}$$

Finally, the Q-conditional distribution of $X(\theta + \cdot)$ given $\mathcal{F}_\theta^X$ equals the $\mathsf{P}_{X(\theta)}$-distribution of $X(\cdot)$. Since also P_y is optimal for every y, we get

$$\tau_K(X(\theta + \cdot)) \geq v(X(\theta)), \quad \text{Q-a.s.} \tag{2.4}$$

Combining the definition of $v(x)$, (2.2), (2.3), and (2.4), we get

$$
\begin{aligned}
v(x) \geq \mathsf{Q}\text{-}\operatorname{ess\,inf} \tau_K &= \mathsf{Q}\text{-}\operatorname{ess\,inf}\{\theta \wedge \tau_K + \tau_K(X(\theta + \cdot))\mathbf{1}_{\theta \leq \tau_K}\} \\
&\geq \mathsf{Q}\text{-}\operatorname{ess\,inf}\{\theta \wedge \tau_K + v(X(\theta))\mathbf{1}_{\theta \leq \tau_K}\} \\
&= \mathsf{P}\text{-}\operatorname{ess\,inf}\{\theta \wedge \tau_K + v(X(\theta))\mathbf{1}_{\theta \leq \tau_K}\}.
\end{aligned}
$$

In the last step we used that $\theta \wedge \tau_K$ and $\mathbf{1}_{\theta \leq \tau_K}$ are $\mathcal{F}_\theta^X$-measurable (even though τ_K is only an $\mathbb{F}_+^X$-stopping time) and hence have the same law under P as under Q due to (2.2). This proves (2.1).

It remains to prove that

$$v(x) = \mathsf{P}\text{-}\operatorname{ess\,inf}\{\theta \wedge \tau_K + v(X(\theta))\mathbf{1}_{\theta \leq \tau_K}\} \tag{2.5}$$

for any optimal $\mathsf{P} \in \mathcal{P}_x$. We thus fix any optimal $\mathsf{P} \in \mathcal{P}_x$. Using (2.3), we get

$$v(x) \leq \tau_K = \theta \wedge \tau_K + \tau_K(X(\theta + \cdot))\mathbf{1}_{\theta \leq \tau_K}, \quad \text{P-a.s.} \tag{2.6}$$

Next, let $\{Q_\omega\}_{\omega \in \Omega}$ be a regular conditional distribution of $X(\theta + \cdot)$ given $\mathcal{F}_\theta^X$; see [41, Theorem 1.3.4]. In particular, $\{F_\omega\}_{\omega \in \Omega}$ with $F_\omega = Q_\omega(\tau_K \in \cdot)$ is then a regular conditional distribution of $\tau_K(X(\theta + \cdot))$ given $\mathcal{F}_\theta^X$. We will now make use of conditional essential infima, which were reviewed before the statement of the proposition. Specifically, we take the $\mathcal{F}_\theta^X$-conditional essential infimum in (2.6). Since $\theta \wedge \tau_K$ and $\mathbf{1}_{\theta \leq \tau_K}$ are $\mathcal{F}_\theta^X$-measurable we get

$$
\begin{aligned}
v(x) &\leq \theta \wedge \tau_K + \mathbf{1}_{\theta \leq \tau_K}\mathsf{P}\text{-}\operatorname{ess\,inf}\{\tau_K(X(\theta + \cdot)) \mid \mathcal{F}_\theta^X\} \\
&= \theta \wedge \tau_K + \mathbf{1}_{\theta \leq \tau_K}\operatorname{ess\,inf} F_\omega \\
&= \theta \wedge \tau_K + \mathbf{1}_{\theta \leq \tau_K}Q_\omega\text{-}\operatorname{ess\,inf} \tau_K, \quad \text{P-a.s.}
\end{aligned}
$$

One readily verifies that $Q_\omega \in \mathcal{P}_{X(\theta,\omega)}$ for P-a.e. ω. Hence $Q_\omega\text{-}\operatorname{ess\,inf} \tau_K \leq v(X(\theta,\omega))$ for P-a.e. ω, and we deduce that $v(x) \leq \theta \wedge \tau_K + v(X(\theta))\mathbf{1}_{\theta \leq \tau_K}$, P-a.s. This yields (2.5), and completes the proof of the proposition. □

It is not true in general that, under an optimal law, $X(t)$ is located on the t-level surface of the value function, even if the value function is smooth. The following example illustrates this.

Example 2.3. Let $K \subset \mathbb{R}^3$ be the union of the line segment $L = (-1,1) \times \{(0,0)\}$ and the shifted unit discs $(1,0,0) + D$ and $(-1,0,0) + D$ with $D = \{(0,y,z)\colon y^2 + z^2 \leq 1\}$. Thanks to Example 1.8, at points $(\pm 1, y, z)$ in the shifted discs, the value function is $v(\pm 1, y, z) = 1 - y^2 - z^2$. At points $\bar{x} = (x,0,0) \in L$, the value function is $v(\bar{x}) = 1$. Indeed, X evolves as a Brownian motion along L until it hits $(\pm 1, 0, 0)$. This happens arbitrarily quickly, and at either point the value function is 1. Thus everywhere in K, $v(x,y,z) = 1 - y^2 - z^2$. We see that for $\bar{x} \in L$, under any optimal $\mathsf{P} \in \mathcal{P}_{\bar{x}}$ one has $v(X(t)) > v(\bar{x}) - t$ for all $t > 0$. Note that in this example, v is very smooth: on K it coincides with a polynomial.

Proposition 2.5 below can be viewed as an assertion about *propagation of continuity*: if the value function is continuous on a certain set, then it is also continuous on a larger set. Upper semicontinuity, which holds in general due to Proposition 2.2(ii), plays an important role. Furthermore, a refined version of this result is crucial in Section 5, where K will be a convex body. These results depend crucially on the following elementary lemma, which allows us to exploit the fact that the value function is always upper semicontinuous.

Lemma 2.4. *Let $C \subset \mathbb{R}^d$ be a compact set, and let $f \colon \mathbb{R}^d \to \mathbb{R}$ be a function that is upper semicontinuous at every point in C. If the restriction $f|_C$ is continuous, then there exists a modulus ω such that*

$$f(x) \leq f(y) + \omega(|x - y|) \text{ for all } x \in \mathbb{R}^d \text{ and } y \in C.$$

Proof. It suffices to pick any $\varepsilon > 0$ and exhibit $\delta > 0$ such that $f(x) \leq f(y) + \varepsilon$ holds whenever $x \in \mathbb{R}^d$, $y \in C$, and $|x - y| < \delta$. Since $f|_C$ is continuous and f is upper semicontinuous at C, for every $y \in C$ there exists $\delta_y > 0$ such that $|f(y) - f(y')| < \varepsilon/2$ and $f(x) < f(y) + \varepsilon/2$ whenever $y' \in C$, $|y - y'| < \delta_y$, $x \in \mathbb{R}^d$, $|y - x| < \delta_y$. The balls $B(y, \delta_y/2)$, $y \in C$, cover C. By compactness, there is a finite subcover $B(y_i, r_i)$, $i = 1, \ldots, n$, where $r_i = \delta_{y_i}/2$. Define $\delta = \min\{r_1, \ldots, r_n\}$. Suppose $x \in \mathbb{R}^d$, $y \in C$, and $|x - y| < \delta$. Then $y \in B(y_i, r_i)$ for some $i \in \{1, \ldots, n\}$, and hence $|x - y_i| \leq |x - y| + |y - y_i| < 2r_i \leq \delta_{y_i}$ and $|y - y_i| < \delta_{y_i}$. Therefore

$$f(x) < f(y_i) + \frac{\varepsilon}{2} \leq f(y) + |f(y_i) - f(y)| + \frac{\varepsilon}{2} < f(y) + \varepsilon,$$

as required. $\qquad\square$

Proposition 2.5. *Let K be compact, and assume $v|_{\partial K}$ is continuous. Then $v|_K$ is continuous.*

Proof. Since v is upper semicontinuous by Proposition 2.2(ii), since ∂K is compact, and since $v|_{\partial K}$ is continuous by assumption, Lemma 2.4 below gives a modulus ω such that

$$v(x) \leq v(y) + \omega(|x - y|) \text{ for all } x \in \mathbb{R}^d \text{ and } y \in \partial K. \tag{2.7}$$

Fix $\bar{x}, \bar{y} \in K$ and an optimal law $\mathsf{P} \in \mathcal{P}_{\bar{x}}$. Define the process $Y = X - \bar{x} + \bar{y}$ and the $\mathbb{F}^X$-stopping time $\theta = \inf\{t \geq 0 \colon Y(t) \notin \operatorname{int}(K)\}$. Note that $\mathsf{P}(\theta < \infty) = 1$ by Example 1.8. Since $Y(\theta) \in \partial K$, we have from (2.7) that

$$v(X(\theta)) \leq v(Y(\theta)) + \omega(|\bar{x} - \bar{y}|), \quad \mathsf{P}\text{-a.s.}$$

We now combine this with two applications of the dynamic programming principle of Proposition 2.2(iii). We get

$$\begin{aligned}
v(\bar{x}) &= \mathsf{P}\text{-}\operatorname{ess\,inf}\{\theta \wedge \tau_K + v(X(\theta))\mathbf{1}_{\theta \leq \tau_K}\} \\
&\leq \mathsf{P}\text{-}\operatorname{ess\,inf}\{\theta + v(Y(\theta))\} + \omega(|\bar{x} - \bar{y}|) \\
&\leq v(\bar{y}) + \omega(|\bar{x} - \bar{y}|).
\end{aligned}$$

In the last inequality, the application of the dynamic programming principle uses that the law of Y lies in $\mathcal{P}_{\bar{y}}$, that $\mathbb{F}^Y = \mathbb{F}^X$, and that $\theta \leq \inf\{t \geq 0 \colon Y(t) \notin K\}$, P-a.s. Since $\bar{x}, \bar{y} \in K$ were arbitrary, we deduce that $v|_K$ is uniformly continuous with modulus ω. $\quad\square$

As mentioned in Section 1, some of the analysis in this paper can be cast in the language of stochastic target problems. We end this section with a remark detailing this connection. Since this is not used in the analysis to come, we do not give proofs.

Remark 2.6. For any $t \in [0, \infty)$, the *target reachability set* when the target is K and the controlled state dynamics is described by $\mathcal{P}_x$ is defined by

$$V(t) = \{x \in \mathbb{R}^d \colon \exists \mathsf{P} \in \mathcal{P}_x \text{ such that } X(t) \in K, \text{ P-a.s.}\}.$$

This is a "time-to-maturity" version, in a weak formulation, of the definition in [33, Section 2.4]. Clearly $V(0) = K$, and one can show that $V(t) = \emptyset$ for all $t > \operatorname{diam}(K)^2/4$. One expects the following representation of the value function v in (1.2) in terms of the target reachability set:

$$v(x) = \sup\{t \geq 0 \colon x \in V(t)\}, \quad x \in K.$$

This equality can be shown to hold if K is convex, but there are non-convex examples where it fails. In such cases, one can work with the *obstacle version* of the stochastic target problem, where the reachability set is defined by

$$W(t) = \{x \in \mathbb{R}^d \colon \exists \mathsf{P} \in \mathcal{P}_x \text{ such that } X_s \in K \text{ for all } s \in [0, t], \text{ P-a.s.}\}.$$

This problem is discussed briefly in Section 7 of [33] and further in [10] (where the terminology "obstacle version" is introduced). It is straightforward to show that

$$v(x) = \sup\{T \geq 0 \colon x \in W(T)\}, \quad x \in K,$$

regardless of the geometry of K. A suitable weak-formulation version of the geometric dynamic programming principle in Theorem 7.1 of [33] or Theorem 2.1 of [10] could then be used to derive characterizations of $W(t)$, and hence $v(x)$, in terms of PDEs.

3 The value function is a viscosity solution

In this section we prove Theorem 1.1, the viscosity solution property, assuming that $d \geq 2$ and that K is compact but not necessarily convex. (We already know from Proposition 2.2(ii) that v is upper semicontinuous.) A bounded function $u \colon K \to \mathbb{R}$ is called a *viscosity subsolution* of $F(\nabla u, \nabla^2 u) = 1$ in $\operatorname{int}(K)$ if

$$\left.\begin{array}{l} (\bar{x}, \varphi) \in \operatorname{int}(K) \times C^2(\mathbb{R}^d) \text{ and} \\ (u^* - \varphi)(\bar{x}) = \max_K(u^* - \varphi) \end{array}\right\} \implies F_*(\nabla\varphi(\bar{x}), \nabla^2\varphi(\bar{x})) \leq 1,$$

where an upper (lower) star denotes upper (lower) semicontinuous envelope (restricting the function to K). We say that u has *zero boundary condition* (in the viscosity sense) if

$$\left.\begin{array}{l} (\bar{x}, \varphi) \in \partial K \times C^2(\mathbb{R}^d) \text{ and} \\ (u^* - \varphi)(\bar{x}) = \max_K(u^* - \varphi) \end{array}\right\} \implies F_*(\nabla\varphi(\bar{x}), \nabla^2\varphi(\bar{x})) \leq 1 \text{ or } u^*(\bar{x}) \leq 0.$$

The function u is said to be a *viscosity supersolution* in $\operatorname{int}(K)$ with *zero boundary condition* if the same conditions hold with u^*, F_*, max, $\leq$ replaced by u_*, F^*, min, $\geq$. It is a *viscosity solution* in $\operatorname{int}(K)$ with *zero boundary condition* if it is both a viscosity sub- and supersolution in $\operatorname{int}(K)$ with zero boundary condition.

To prove Theorem 1.1, we must establish the sub- and supersolution properties. We carry out these tasks separately in the following two subsections. To do so, the following description of the semicontinuous envelopes of F will be needed.

Lemma 3.1. *The nonlinearity* (1.3) *satisfies* $F_* = F$, *as well as* $F^*(p, M) = F(p, M)$ *for* $p \neq 0$, *and* $F^*(0, M) = -\lambda_2(M)/2$. *Here* $\lambda_1(M) \geq \lambda_2(M) \geq \cdots \geq \lambda_d(M)$ *are the eigenvalues of* $M \in \mathbb{S}^d$. *In particular,* F *is continuous on the set* $(\mathbb{R}^d \setminus \{0\}) \times \mathbb{S}^d$.

Proof. From the representation (1.5) we have $F(p, M) = -\frac{1}{2}\sup\{y^\top My\colon |y| = 1, \, y^\top p = 0\}$. One checks that this is continuous on the set $(\mathbb{R}^d \setminus \{0\}) \times \mathbb{S}^d$, and in particular equal to F_* and F^* there. Next, we claim that

$$-\frac{1}{2}\lambda_1(M) \leq F(p, M) \leq -\frac{1}{2}\lambda_2(M) \tag{3.1}$$

for all (p, M). The first inequality follows because $\sup\{y^\top My\colon |y| = 1\} = \lambda_1(M)$. For the second inequality, use the spectral theorem to write $M = \lambda_1(M)w_1w_1^\top + \cdots + \lambda_d(M)w_dw_d^\top$ for an orthonormal basis $w_1, \ldots, w_d$ of eigenvectors of M. Express p and y in this basis, say $p = \pi_1 w_1 + \cdots + \pi_d w_d$ and $y = \eta_1 w_1 + \cdots + \eta_d w_d$, to get

$$F(p, M) = -\frac{1}{2}\sup\left\{\sum_{i=1}^d \eta_i^2 \lambda_i(M)\colon \sum_{i=1}^d \eta_i^2 = 1, \sum_{i=1}^d \eta_i \pi_i = 0\right\}.$$

If $\pi_1 = 0$ one can take $\eta_1 = 1$ and $\eta_i = 0$ for $i \geq 2$ to get $F(p, M) \leq -\lambda_1(M)/2$. Otherwise one can take $\eta_2 = (1 + (\pi_2/\pi_1)^2)^{-1/2}$ and $\eta_1 = -\eta_2\pi_2/\pi_1$ to get $F(p, M) \leq -(\eta_1^2\lambda_1(M) + \eta_2^2\lambda_2(M))/2 \leq -\lambda_2(M)/2$. In either case, the second inequality of (3.1) holds.

For any fixed M, there is a sequence $(p_n, M_n) \to (0, M)$ with $F_*(0, M) = \lim_n F(p_n, M_n)$. Thus by (3.1) and since $\lambda_1(M)$ is continuous in M, we get

$$F(0, M) \geq F_*(0, M) = \lim_n F(p_n, M_n) \geq -\frac{1}{2}\lim_n \lambda_1(M_n) = -\frac{1}{2}\lambda_1(M) = F(0, M).$$

This shows that $F_*(0, M) = F(0, M)$. On the other hand, with w_1 an eigenvector of M with eigenvalue $\lambda_1(M)$, we have $F(n^{-1}w_1, M) = -\lambda_2(M)/2$. Sending $n \to \infty$ shows that $F^*(0, M) \geq -\lambda_2(M)/2$ and thus, by (3.1) and the continuity of $\lambda_2(M)$ in M, that $F^*(0, M) = -\lambda_2(M)/2$. $\qquad\square$

For later use, let us also record the following observations. We let $|A|_{\mathrm{op}} = \sup_{|y|=1} |Ay|$ denote the operator norm of a matrix A.

Lemma 3.2. *If $p \in \mathbb{R}^d$, $M \in \mathbb{S}^d$, $F^*(p, M) > 0$, and B is an $d \times d$ invertible matrix then*

$$F^*(p, M) \leq |(BB^\top)^{-1}|_{\mathrm{op}} F^*(B^\top p, B^\top MB).$$

Proof. Assume first $p \neq 0$, so that $F^*(p, M) = F(p, M)$ by Lemma 3.1. Consider any $\bar{a} \in \mathbb{S}_+^d$ with $\bar{a}B^\top p = 0$ and $\mathrm{tr}(\bar{a}) = 1$. Then $\mathrm{tr}(B\bar{a}B^\top) > 0$. Define now $a = (B\bar{a}B^\top)/\mathrm{tr}(B\bar{a}B^\top)$. Then $a \in \mathbb{S}_+^d$, $ap = 0$, and $\mathrm{tr}(a) = 1$. Thus from the definition of F,

$$F^*(p, M) = F(p, M) \leq -\frac{1}{2}\mathrm{tr}(aM) = -\frac{1}{2}\mathrm{tr}(\bar{a}B^\top MB)\frac{1}{\mathrm{tr}(B\bar{a}B^\top)}. \tag{3.2}$$

Since $F^*(p, M) \geq 0$ we have $\mathrm{tr}(\bar{a}B^\top MB) \leq 0$. Moreover, by expanding $B\bar{a}B^\top$ using the spectral theorem and applying the definition of the operator norm, we obtain $1 = \mathrm{tr}(B\bar{a}B^\top(BB^\top)^{-1}) \leq \mathrm{tr}(B\bar{a}B^\top)|(BB^\top)^{-1}|_{\mathrm{op}}$. Thus (3.2) yields

$$F^*(p, M) \leq -\frac{1}{2}\mathrm{tr}(\bar{a}B^\top MB)|(BB^\top)^{-1}|_{\mathrm{op}},$$

and taking infimum on the right-hand side gives the assertion, still for $p \neq 0$.

Consider now the case $p = 0$ and consider a sequence (p_n, M_n) converging to (p, M) with $p_n \neq 0$ such that $\lim_n F(p_n, M_n) = F^*(p, M)$. Since for sufficiently large n we have $F(p_n, M_n) > 0$, we get from the case just established that

$$F^*(p, M) = \lim_n F(p_n, M_n) \leq |(BB^\top)^{-1}|_{\mathrm{op}} \limsup_n F^*(B^\top p_n, B^\top M_n B)$$

$$\leq |(BB^\top)^{-1}|_{\mathrm{op}} F^*(B^\top p, B^\top MB),$$

as desired. $\qquad\square$

Corollary 3.3. *Let B be a $d \times d$ invertible matrix, viewed as a linear map. Define $K' = B^{-1}(K)$ and $\bar{w} = |(BB^\top)^{-1}|_{\text{op}}\, w \circ B$. If, on K, w is a lower semicontinuous viscosity supersolution of (1.4) with zero boundary condition, then so is $\bar{w}$ on K'.*

Proof. The statement follows from the definition of viscosity supersolution, in conjunction with Lemma 3.2. $\qquad\square$

3.1 Subsolution property

We now prove the subsolution property claimed in Theorem 1.1. Since v is upper semicontinuous and F is lower semicontinuous, we may drop the stars in the definition of subsolution.

Proof of the subsolution property. Fix $\bar{x} \in K$. If $\bar{x} \in \text{int}(K)$ then $v(\bar{x}) > 0$ by Example 1.8. If $\bar{x} \in \partial K$ and $v(\bar{x}) = 0$ then the subsolution property holds for this point. Hence, without loss of generality, we may assume that $v(\bar{x}) > 0$.

Fix now $\varphi \in C^2(\mathbb{R}^d)$ with $\varphi(\bar{x}) = v(\bar{x})$ and $\varphi(x) \geq v(x)$ for all $x \neq \bar{x}$. We assume that $F(\nabla\varphi(\bar{x}), \nabla^2\varphi(\bar{x})) > 1$ and work towards a contradiction. Without loss of generality, we may assume $\varphi(x) > v(x)$ for all $x \neq \bar{x}$.

We claim that there exists $\varepsilon \in (0, \sqrt{v(\bar{x})}/2)$ such that

$$
\begin{aligned}
&\text{for all } (x,a) \in (K \cap B_\varepsilon(\bar{x})) \times \mathbb{S}^d_+ \text{ with } \text{tr}(a) = 1, \text{ we have} \\
&1 + \frac{1}{2}\text{tr}(a\nabla^2\varphi(x)) > 0 \quad \text{implies} \quad \nabla\varphi(x)^\top a \nabla\varphi(x) \geq \varepsilon.
\end{aligned}
\tag{3.3}
$$

Indeed, if not, there exist $\varepsilon_n \to 0$ and $(x_n, a_n) \in (K \cap B_{\varepsilon_n}(\bar{x})) \times \mathbb{S}^d_+$ such that $\text{tr}(a_n) = 1$ and $\nabla\varphi(x_n)^\top a_n \nabla\varphi(x_n) \leq \varepsilon_n$, but $1 + \frac{1}{2}\text{tr}(a_n\nabla^2\varphi(x_n)) > 0$. In particular, $x_n \to \bar{x}$ and, after passing to a subsequence, we also have $a_n \to a$ for some $a \in \mathbb{S}^d_+$. Passing to the limit yields $\text{tr}(a) = 1$, $a\nabla\varphi(\bar{x}) = 0$, and $1 + \frac{1}{2}\text{tr}(a\nabla^2\varphi(\bar{x})) \geq 0$. This contradicts the assumption that $1 - F(\nabla\varphi(\bar{x}), \nabla^2\varphi(\bar{x})) < 0$, and proves the claim.

Note also that there exists some $c > 0$ such that for all $(x,a) \in (K \cap B_\varepsilon(\bar{x})) \times \mathbb{S}^d_+$ with $\text{tr}(a) = 1$ we have

$$
1 + \frac{1}{2}\text{tr}(a\nabla^2\varphi(x)) \leq 1 + \frac{1}{2}\lambda_1(\nabla^2\varphi(x)) \leq c < \infty,
\tag{3.4}
$$

where $\lambda_1(M)$ denotes the largest eigenvalue of a symmetric matrix M. The boundedness comes from the continuity of λ_1. Furthermore, we have

$$
\delta = \min_{K \cap \partial B_\varepsilon(\bar{x})} (\varphi - v) > 0.
\tag{3.5}
$$

Fix any optimal $\mathsf{P} \in \mathcal{P}_{\bar{x}}$. We then have a predictable $\mathbb{S}^d_+$-valued process $(a(s))_{s \geq 0}$ such that

$$
\langle X \rangle(t) = \int_0^t a(s)ds \quad \text{and} \quad \text{tr}(a(t)) = 1,\ dt \otimes d\mathsf{P}\text{-a.e.}
$$

Define the stopping time

$$
\theta = \inf\{t \geq 0 \colon X(t) \notin B_\varepsilon(\bar{x})\} \wedge v(\bar{x}).
$$

Clearly $\theta \leq \tau_K$ by definition of $v(\bar{x})$ and $\mathsf{P}[X(\theta) \in K \cap \partial B_\varepsilon(\bar{x})] > 0$ since $\varepsilon < \sqrt{v(\bar{x})}/2$ (recall Example 1.8).

We can now define the predictable set

$$
J = \{s \in [0,\theta) \colon 1 + \frac{1}{2}\text{tr}(a(s)\nabla^2\varphi(X(s))) > 0\}.
$$

Next, the dynamic programming principle of Proposition 2.2(iii) with $t \wedge \theta$ in place of θ yields

$$v(\bar{x}) \le t \wedge \theta + v(X(t \wedge \theta)), \quad \text{P–a.s.} \tag{3.6}$$

Using (3.6) and then (3.5), we get

$$\varphi(\bar{x}) = v(\bar{x}) \le t \wedge \theta + v(X(t \wedge \theta)) \le t \wedge \theta - \delta \mathbf{1}_{[\theta,\infty)}(t) \mathbf{1}_{\{X(\theta) \in K \cap \partial B_\varepsilon(\bar{x})\}} + \varphi(X(t \wedge \theta)).$$

Combining this with Itô's formula, the definition of J, and (3.4), we get

$$\begin{aligned}
\delta \mathbf{1}_{[\theta,\infty)}(t) \mathbf{1}_{\{X(\theta) \in K \cap \partial B_\varepsilon(\bar{x})\}} &\le t \wedge \theta + \varphi(X(t \wedge \theta)) - \varphi(\bar{x}) \\
&= \int_0^{t \wedge \theta} \nabla\varphi(X(s))^\top dX(s) + \int_0^{t \wedge \theta} (1 + \frac{1}{2}\operatorname{tr}(a(s)\nabla^2\varphi(X(s))))ds \\
&\le \int_0^{t \wedge \theta} \nabla\varphi(X(s))^\top dX(s) + c\int_0^{t \wedge \theta} \mathbf{1}_J(s)ds.
\end{aligned}$$

Now, define the process

$$\widetilde{X}(t) = X(t) + \frac{c}{\varepsilon}\int_0^t a(s)\nabla\varphi(X(s))\mathbf{1}_J(s)ds.$$

Due to (3.3) and the definition of J, we then have

$$\begin{aligned}
\delta \mathbf{1}_{[\theta,\infty)}(t) \mathbf{1}_{\{X(\theta) \in K \cap \partial B_\varepsilon(\bar{x})\}} &\le \int_0^{t \wedge \theta} \nabla\varphi(X(s))^\top d\widetilde{X}(s) \\
&\quad + \int_0^{t \wedge \theta} (c - \frac{c}{\varepsilon}\nabla\varphi(X(s))^\top a(X(s))\nabla\varphi(X(s)))\mathbf{1}_J(s)ds \\
&\le \int_0^{t \wedge \theta} \nabla\varphi(X(s))^\top d\widetilde{X}(s). \tag{3.7}
\end{aligned}$$

Consider now the exponential local martingale Z given by

$$\frac{dZ(t)}{Z(t)} = -\frac{c}{\varepsilon}\mathbf{1}_J(t)\nabla\varphi(X(t))^\top dX(t), \quad Z_0 = 1.$$

This is well-defined since $\nabla\varphi$ is bounded on the closure of $B_\varepsilon(\bar{x})$, which contains $X(t)$ for $t \in J$. An application of Itô's formula shows that multiplying (3.7) by $Z(t)$ gives a local martingale, and hence a supermartingale since it is nonnegative. Therefore,

$$0 < \delta \, \mathsf{E}[\mathbf{1}_{\{X(\theta) \in K \cap \partial B_\varepsilon(\bar{x})\}} Z(\theta)] \le \mathsf{E}\left[Z(\theta)\int_0^\theta \nabla\varphi(X(s))^\top d\widetilde{X}(s)\right] \le 0,$$

using that $\theta < \infty$, P-a.s., and $\mathsf{P}[X(\theta) \in K \cap \partial B_\varepsilon(\bar{x})] > 0$ for the first inequality. This contradiction completes the proof of the subsolution property. $\qquad\square$

3.2 Supersolution property

The following result is used in the proof.

Lemma 3.4. *Let $m \in \mathbb{N}$ with $m \ge 2$, and let S be a nonzero skew-symmetric $m \times m$ matrix and let $x, \bar{x} \in \mathbb{R}^m$. Then there exists a weak solution to the SDE*

$$dY(t) = \frac{S(Y(t) - \bar{x})}{|S(Y(t) - \bar{x})|}dW(t), \quad Y(0) = x,$$

that satisfies $|S(Y(t) - \bar{x})|^2 = |S(x - \bar{x})|^2 + t$ for all $t \ge 0$. Here W denotes a one-dimensional Brownian motion.

Proof. Suppose first that $S(x - \bar{x}) \neq 0$. Since the SDE has locally Lipschitz coefficients on the set $\{y \colon S(y - \bar{x}) \neq 0\}$, there is a local solution Y on $[0, \zeta)$, where $\zeta = \inf\{t \geq 0 \colon S(Y(t) - \bar{x}) = 0\}$. Itô's formula and the skew-symmetry of S give

$$d|S(Y(t) - \bar{x})|^2 = 2\frac{(Y(t) - \bar{x})^\top S^\top S^2 (Y(t) - \bar{x})}{|S(Y(t) - \bar{x})|} dW(t) + dt = dt, \quad t < \zeta,$$

so $\zeta = \infty$. Thus Y is actually a global solution, and $|S(Y(t) - \bar{x})|^2 = |S(x - \bar{x})|^2 + t$ for all $t \geq 0$. This proves the case where $S(x - \bar{x}) \neq 0$.

Suppose now that $S(x - \bar{x}) = 0$, and select points $x_n \in \mathbb{R}^m$ with $x_n \to x$ and $S(x_n - \bar{x}) \neq 0$. For each n, let Y_n be a solution to the SDE with $Y_n(0) = x_n$. Since $\mathrm{tr}\langle Y_n \rangle(t) \equiv t$, the law of $Y_n - x_n$ lies in $\mathcal{P}_0$, which is compact by Proposition 2.2(i). Thus after passing to a subsequence, we have $Y_n - x_n \Rightarrow Y - x$ for some limiting process Y with $Y(0) = x$. Since the set $C = \{\omega \colon |S(\omega(t) - \bar{x})|^2 = |S(\omega(0) - \bar{x})|^2 + t \text{ for all } t \geq 0\}$ is closed, and since Y_n lies in C almost surely for all n, the Portmanteau lemma implies that Y does as well. In particular, we have $|S(Y(t) - \bar{x})|^2 = t$ for all $t \geq 0$.

Now, for every $f \in C_c^\infty(\mathbb{R}^m)$, $k \in \mathbb{N}$, $0 \leq s_1 \leq \cdots \leq s_k < s < t$, and $g \in C_b((\mathbb{R}^m)^k)$, we have

$$\mathsf{E}\left[\left(f(Y_n(t)) - f(Y_n(s)) - \int_s^t Lf(Y_n(u))du \right) g(Y_n(s_1), \ldots, Y_n(s_k)) \right] = 0, \qquad (3.8)$$

where $Lf(y) = \frac{1}{2}(y - \bar{x})^\top S^\top \nabla^2 f(y) S(y - \bar{x})/|S(y - \bar{x})|^2$ is the operator associated to the given SDE. Since $|S(Y_n(u) - \bar{x})|^2 = |S(x_n - \bar{x})|^2 + u$, the integral in (3.8) can be written as

$$\int_s^t Lf(Y_n(u))du = \int_s^t \frac{(Y_n(u) - \bar{x})^\top S^\top \nabla^2 f(Y_n(u)) S(Y_n(u) - \bar{x})}{2(|S(x_n - \bar{x})|^2 + u)} du. \qquad (3.9)$$

The right-hand side of (3.9) is a bounded continuous function G_n of the trajectory of Y_n. Since $S(x_n - \bar{x}) \to S(x - \bar{x}) = 0$, the functions G_n converge uniformly to the function G given by $G(\omega) = \int_s^t (\omega(u) - \bar{x})^\top S^\top \nabla^2 f(\omega(u)) S(\omega(u) - \bar{x})/(2u)du$. Since the limiting process Y again satisfies $|S(Y(t) - \bar{x})|^2 = t$, we have $G(Y) = \int_s^t Lf(Y(u))du$. We may now pass to the limit in (3.8) and deduce that the corresponding equality holds for Y in place of Y_n. Since the set of all random variables $g(Y(s_1), \ldots, Y(s_k))$ with $k, s_1, \ldots, s_k, g$ as above generate $\sigma(Y(u) \colon u \leq s)$, we deduce that the process $M^f(t) = f(Y(t)) - f(Y(0)) - \int_0^t Lf(u, Y(u))du$, $t \geq 0$, is a martingale with respect to the filtration generated by Y. Since $f \in C_c^\infty(\mathbb{R}^m)$ was arbitrary, Y solves the martingale problem associated with the given SDE. Equivalently, Y is a weak solution, as desired. $\qquad\square$

We now turn to the supersolution property claimed in Theorem 1.1.

Proof of the supersolution property. Fix $\bar{x} \in K$. If $\bar{x} \in \partial K$ then there is nothing to prove since v is nonnegative. Hence, we may assume throughout the proof that $\bar{x} \in \mathcal{O}$, where we write $\mathcal{O} = \mathrm{int}(K)$.

Fix now $\varphi \in C^2(\mathbb{R}^d)$ with $\varphi \leq v_*$ and $\varphi(\bar{x}) = v_*(\bar{x})$. A standard perturbation argument relying on test functions $\varphi(x) - \varepsilon|x - \bar{x}|^2$ lets us suppose that $\varphi(x) < v_*(x)$ for all $x \neq \bar{x}$, and that the Hessian $\nabla^2 \varphi(\bar{x})$ is nonsingular. We consider three cases, depending on the properties of $\nabla \varphi(\bar{x})$ and $\nabla^2 \varphi(\bar{x})$.

Case 1: Suppose $\nabla \varphi(\bar{x}) \neq 0$. Assume for contradiction that $F^*(\nabla \varphi(\bar{x}), \nabla^2 \varphi(\bar{x})) < 1$. Since F^* equals F at this point, it follows from the representation (1.5) that there exists $\bar{\sigma} \in \mathbb{R}^d$ such that

$$|\bar{\sigma}| = 1, \ \bar{\sigma}^\top \nabla \varphi(\bar{x}) = 0, \text{ and } 1 + \frac{1}{2}\bar{\sigma}^\top \nabla^2 \varphi(\bar{x})\bar{\sigma} > 0.$$

In particular, there exists a skew-symmetric $d \times d$ matrix S such that $\bar{\sigma} = S\nabla\varphi(\bar{x})$; for instance,

$$S = \frac{1}{|\nabla\varphi(\bar{x})|^2}(\bar{\sigma}\nabla\varphi(\bar{x})^\top - \nabla\varphi(\bar{x})\bar{\sigma}^\top).$$

Furthermore, we can select $\varepsilon > 0$ such that the closure of $B_\varepsilon(\bar{x})$ is contained in $\mathcal{O}$ and

$$|S\nabla\varphi| \geq \frac{1}{2} \text{ and } |S\nabla\varphi|^2 + \frac{1}{2}\nabla\varphi^\top S^\top \nabla^2\varphi\, S\nabla\varphi \geq 0 \text{ on } B_\varepsilon(\bar{x}). \tag{3.10}$$

Fix any $x \in B_\varepsilon(\bar{x})$. Define

$$\theta = \inf\{t \geq 0 \colon X(t) \notin B_\varepsilon(\bar{x})\}, \tag{3.11}$$

and let P be the law under which X satisfies

$$dX(t) = \left(\frac{S\nabla\varphi(X(t))}{|S\nabla\varphi(X(t))|}\mathbf{1}_{[0,\theta)}(t) + e_1\mathbf{1}_{[\theta,\infty)}(t)\right)dW(t), \quad X_0 = x, \tag{3.12}$$

where W is a one-dimensional Brownian motion and e_1 is the first canonical unit vector (any other unit vector would also do). Note that $P \in \mathcal{P}_x$ and $\theta \leq \tau_K$, and thus $\theta < \infty$, P-a.s. by Lemma 1.9. Define

$$\delta = \min_{\partial B_\varepsilon(\bar{x})}(v_* - \varphi) > 0.$$

Using first that $v \geq v_* \geq \varphi + \delta$ on $\partial B_\varepsilon(\bar{x})$; then Itô's formula; and finally (3.10) along with the fact that $\nabla\varphi^\top S\,\nabla\varphi = 0$ by skew-symmetry of S, we get

$$\theta + v(X(\theta)) \geq \delta + \theta + \varphi(X(\theta))$$
$$= \delta + \varphi(x) + \int_0^\theta \frac{\nabla\varphi^\top S\,\nabla\varphi}{|S\nabla\varphi|}(X(s))dW(s) + \int_0^\theta (1 + \frac{\nabla\varphi^\top S^\top \nabla^2\varphi\, S\nabla\varphi}{2|S\nabla\varphi|^2}(X(s)))ds$$
$$\geq \delta + \varphi(x), \quad \text{P-a.s.}$$

Combining this with the dynamic programming principle of Proposition 2.2(iii) yields

$$v(x) \geq \text{P-ess inf}\{\theta + v(X(\theta))\} \geq \delta + \varphi(x).$$

Since $x \in B_\varepsilon(\bar{x})$ was arbitrary, we may send $x \to \bar{x}$ such that $v(x) \to v_*(\bar{x}) = \varphi(\bar{x})$, and deduce $0 \geq \delta$. This contradiction proves the supersolution property when $\nabla\varphi(\bar{x}) \neq 0$.

Case 2: Suppose now that $\nabla\varphi(\bar{x}) = 0$ and $\nabla^2\varphi(\bar{x})$ is negative definite. Assume for contradiction that $F^*(0, \nabla^2\varphi(\bar{x})) < 1$, meaning that $1 + \lambda_2(\nabla^2\varphi(\bar{x}))/2 > 0$ thanks to Lemma 3.1. We will replace φ by a simpler test function $\widetilde{\varphi}$. To this end, define $\gamma_i = \lambda_i(\nabla^2\varphi(\bar{x})) - \eta$ for $i = 1, \ldots, d$, where $\eta > 0$ is small enough so that $1 + \gamma_2/2 \geq 0$. Let $w_1, \ldots, w_d$ be an orthonormal basis of eigenvectors of $\nabla^2\varphi(\bar{x})$ corresponding to its ordered eigenvalues. Define

$$M = \gamma_2(w_1 w_1^\top + w_2 w_2^\top) + \gamma_3 w_3 w_3^\top + \cdots + \gamma_d w_d w_d^\top$$

and

$$\widetilde{\varphi}(x) = v_*(\bar{x}) + \frac{1}{2}(x - \bar{x})^\top M(x - \bar{x}).$$

Then $\widetilde{\varphi}(\bar{x}) = \varphi(\bar{x}) = v_*(\bar{x})$, $\nabla\widetilde{\varphi}(\bar{x}) = \nabla\varphi(\bar{x}) = 0$, and $\nabla^2\widetilde{\varphi}(\bar{x}) = M \prec \nabla^2\varphi(\bar{x})$. Thus $\widetilde{\varphi} \leq \varphi$ on some ball $B_\varepsilon(\bar{x})$ with positive radius $\varepsilon > 0$, whose closure is contained in $\mathcal{O}$. Define the skew-symmetric matrix

$$S = w_1 w_2^\top - w_2 w_1^\top.$$

Fix any $x \in B_\varepsilon(\bar{x})$, and let P be a law under which X satisfies

$$dX(t) = \frac{S(X(t) - \bar{x})}{|S(X(t) - \bar{x})|}dW(t), \quad X(0) = x,$$

and $|S(X(t) - \bar{x})|^2 = |S(x - \bar{x})|^2 + t$, where W is a one-dimensional Brownian motion. Such P exists by Lemma 3.4, and it is clear that $\mathsf{P} \in \mathcal{P}_x$. Itô's formula, the identity $MS = \gamma_2 S$, and the skew-symmetry of S give

$$
\begin{aligned}
\widetilde{\varphi}(X_t) &= \widetilde{\varphi}(x) + \int_0^t \frac{(X(s) - \bar{x})^\top MS(X(s) - \bar{x})}{|S(X(s) - \bar{x})|} dW(s) \\
&\quad + \frac{1}{2} \int_0^t \frac{(X(s) - \bar{x})^\top S^\top MS(X(s) - \bar{x})}{|S(X(s) - \bar{x})|^2} ds \\
&= \widetilde{\varphi}(x) + \frac{\gamma_2}{2} t, \quad t \geq 0, \quad \text{P-a.s.}
\end{aligned}
$$

As in Case 1, let θ be given by (3.11) and define $\delta = \min_{\partial B_\varepsilon(\bar{x})} (v_* - \widetilde{\varphi}) > 0$. We then get

$$
\theta + v(X(\theta)) \geq \delta + \theta + \widetilde{\varphi}(X(\theta)) \geq \delta + \widetilde{\varphi}(x) + \left(1 + \frac{\gamma_2}{2}\right) \theta \geq \delta + \widetilde{\varphi}(x), \quad \text{P-a.s.},
$$

using that $1 + \gamma_2/2 \geq 0$. The contradiction $v_*(\bar{x}) \geq \delta + v_*(\bar{x})$ is now obtained as in Case 1 using the dynamic programming principle and a limiting argument.

Case 3: Suppose finally that $\nabla\varphi(\bar{x}) = 0$ and $\nabla^2\varphi(\bar{x})$ has at least one strictly positive eigenvalue with eigenvector $\hat{e}$, say. Fix $\varepsilon_0 > 0$ such that the closure of $B_{\varepsilon_0}(\bar{x})$ is contained in $\mathcal{O}$, and define

$$
\delta = \min_{\partial B_{\varepsilon_0}(\bar{x})} (v_* - \varphi) > 0. \tag{3.13}
$$

Following [33] (specifically, Steps 6–7 in the proof of Theorem 4.1, see Section 8.2 in their paper), we define perturbed test functions

$$
\varphi_\varepsilon(x) = \varphi(x) + \varepsilon \hat{e}^\top (x - \bar{x}).
$$

The minimum of $v_* - \varphi_\varepsilon$ over the closure of $B_{\varepsilon_0}(\bar{x})$ is at most $v_*(\bar{x}) - \varphi_\varepsilon(\bar{x}) = 0$. Because of (3.13), for every sufficiently small $\varepsilon > 0$, the minimum cannot be attained on the boundary $\partial B_{\varepsilon_0}(\bar{x})$, so must be attained at some $x_\varepsilon \in B_{\varepsilon_0}(\bar{x})$. Moreover, since $\bar{x}$ is a strict minimizer of $v_* - \varphi$, we have $x_\varepsilon \to \bar{x}$ as $\varepsilon \to 0$. The argument in Step 7 of the proof of Theorem 4.1 in [33], which makes use of the fact that $\hat{e}$ is an eigenvector with strictly positive eigenvalue, yields that

$$
\nabla\varphi_\varepsilon(x_\varepsilon) \neq 0 \quad \text{for all sufficiently small } \varepsilon.
$$

Therefore, the result proved in Case 1 above implies that $F^*(\nabla\varphi_\varepsilon(x_\varepsilon), \nabla^2\varphi_\varepsilon(x_\varepsilon)) \geq 1$. Sending $\varepsilon \to 0$ gives $F^*(\nabla\varphi(\bar{x}), \nabla^2\varphi(\bar{x})) \geq 1$, which completes the proof of the supersolution property. $\square$

4 Comparison and uniqueness

The main result of this section is the following comparison principle, which is used to prove Theorem 1.2.

Theorem 4.1. *Let $d \geq 2$ and suppose K is compact. Assume there exist invertible affine maps T_λ on $\mathbb{R}^d$, parameterized by $\lambda \in (0,1)$, such that $T_\lambda(K) \subset \text{int}(K)$ and $\lim_{\lambda \to 1} T_\lambda = I$. Let u (w) be an upper (lower) semicontinuous viscosity subsolution (supersolution) of* (1.4), *both u and w with zero boundary condition (in the viscosity sense). Then $u \leq w^*$.*

Before giving the proof, let us show how this implies Theorem 1.2. Let u and v be two upper semicontinuous viscosity solutions of (1.4) with zero boundary condition. Applying the comparison principle with $w = v_*$ yields $u \leq (v_*)^* \leq v^* = v$. Letting u and v switch places yields $v \leq u$, and hence $u = v$.

Example 4.2. If K is strictly star-shaped about the origin, meaning that $\lambda K \subset \mathrm{int}(K)$ for all $\lambda \in (0,1)$, then it clearly satisfies the assumption of Theorem 4.1. In particular, this is the case if K is convex with $0 \in \mathrm{int}(K)$. Here is an example of a body that is not star-shaped but satisfies the assumptions of Theorem 4.1:

$$K = [-1,1]^2 \cap \{(x,y) \colon |y| \le 0.01 + x^2\}.$$

Indeed, one can use the linear maps $T_\lambda(x,y) = (\lambda x, \lambda^2 y)$. It is easily verified that K is not star-shaped.

If K is star-shaped but not strictly star-shaped, then uniqueness among upper semi-continuous viscosity solutions may fail, as the following example shows; see also [38, Example 8.2].

Example 4.3. Let D be the centered unit disk in $\mathbb{R}^2$, and set $K = (D+(1,0)) \cup (D-(1,0))$. Then K is star-shaped because $\lambda K \subset K$ for all $\lambda \in (0,1)$, but not strictly star-shaped because $\mathrm{int}(K)$ is not connected. The value function is upper semicontinuous and satisfies $v(0,0) \ge 1$, which can be seen by the argument in the proof of Proposition 5.1 below. However, it is easy to verify that the function $\bar{v}(x,y) = 1 - (|x|-1)^2 - y^2$ for $(x,y) \in K$ is a (continuous) viscosity solution of (1.4) with zero boundary condition. Since $\bar{v}(0,0) = 0 \ne v(0,0)$, this shows non-uniqueness.

We now give the proof of the comparison principle; see also [38, Section 9], [3, Section 2 and Theorem 4.3], and [24, Theorem 4] for related uniqueness statements.

Proof of Theorem 4.1. We assume for simplicity that the T_λ are linear, not just affine; we may then identify T_λ with its $d \times d$ matrix. Recall that $|\cdot|_{\mathrm{op}}$ denotes the operator norm. By Corollary 3.3, the function $w_\lambda = |(T_\lambda T_\lambda^\top)^{-1}|_{\mathrm{op}} \, w \circ T_\lambda$ is a lower semicontinuous viscosity supersolution of (1.4) with zero boundary condition on $K' = T_\lambda^{-1}(K)$. By the properties of T_λ, we have $K \subset \mathrm{int}(K')$. Theorem A.1 then yields $u \le w_\lambda$ on K for all $\lambda \in (0,1)$. We thus obtain $u \le \limsup_{\lambda \to 1} w_\lambda \le w^*$ on K as desired. $\qquad \square$

5 Convex bodies

Our next goal is to prove continuity of the value function v when $K \subset \mathbb{R}^d$ ($d \ge 2$) is a convex body, i.e. a nonempty compact convex set, satisfying an additional assumption. We first record the following simple property of the value function.

Proposition 5.1. *Let K be a convex body. Then v is quasi-concave.*

Proof. We must prove that v has convex super-level sets. Pick two distinct points $x, y \in K$, and let L be the line passing through x and y. Fix any point $z \in L$, and let P be the law under which X is a standard Brownian motion along L starting at z. Then $\mathsf{P} \in \mathcal{P}_z$, and with $\theta = \inf\{t \ge 0 \colon X(t) \in \{x,y\}\}$ the dynamic programming principle yields $v(z) \ge \mathsf{P}\text{-}\operatorname{ess\,inf}\{\theta + v(X(\theta))\} \ge v(x) \wedge v(y)$. This proves quasi-concavity. $\qquad \square$

Recall the following notions from convex geometry; see [31, 32] for more details. Let G be any subset of $\mathbb{R}^d$. The affine hull of G is denoted by $\mathrm{aff}(G)$, with dimension $\dim(G)$. The relative interior $\mathrm{ri}(G)$ is the interior of G in $\mathrm{aff}(G)$, and the relative boundary is $\mathrm{rbd}(G) = G \setminus \mathrm{ri}(G)$. A face of a convex set K is a convex subset $F \subset K$ such that every (closed) line segment $L \subset K$ with $\mathrm{ri}(L) \cap F \ne \emptyset$ satisfies $L \subset F$. In particular, every face is closed. A face is called a boundary face if it is nonempty and not all of K. The relative boundary $\mathrm{rbd}(K)$ is the union of all boundary faces. For every $x \in K$, there is a unique face of K whose relative interior contains x. We call this face F_x. For each $k = 0, \dots, d$, the k-skeleton is defined as in (1.8), namely

$$\mathcal{F}_k = \text{union of all faces } F \text{ of } K \text{ with } \dim(F) \le k.$$

In particular, $\mathcal{F}_0$ consists of all extreme points, $\mathcal{F}_1$ consists of all extreme points and line segments, $\mathcal{F}_{d-1}$ is the boundary of K, and $\mathcal{F}_d$ is K itself. For convenience we introduce the notation

$$\tau_F = \inf\{t \geq 0 \colon X(t) \notin F\} \tag{5.1}$$

for the first exit time of X from a set F. This notation is consistent with (1.1).

Lemma 5.2. *Let K be a convex body and consider a point $\bar{x} \in K$. For every $\mathsf{P} \in \mathcal{P}_{\bar{x}}$, we have $\tau_K = \tau_{F_{\bar{x}}}$, P-a.s.*

Proof. We assume without loss of generality that $\dim(K) = d$. If $\bar{x} \in \mathcal{F}_d \setminus \mathcal{F}_{d-1}$, then $F_{\bar{x}} = K$, and the statement is obvious. Otherwise, there exists a supporting halfspace $H_1 = \{x \in \mathbb{R}^d \colon a_1^\top x \geq b_1\}$ with $(a_1, b_1) \in \mathbb{R}^d \times \mathbb{R}$ such that $K \subset H_1$ and $\bar{x} \in \partial H_1$. Set $K_1 = K \cap \partial H_1$ and note that $\dim(K_1) < \dim(K) = d$ and $F_{\bar{x}} \subset K_1$. The scalar process $a_1^\top X^{\tau_K} - b_1$ is a nonnegative P-martingale starting at zero, hence is identically zero. Therefore $\tau_{K_1} = \tau_K$. If $K_1 = F_{\bar{x}}$, we are done. If not, we iterate the procedure and fix another halfspace $H_2 = \{x \in \mathbb{R}^d \colon a_2^\top x \geq b_2\} \neq H_1$ with $(a_2, b_2) \in \mathbb{R}^d \times \mathbb{R}$ such that $K_1 \subset H_2$ and $\bar{x} \in \partial H_2$. Setting $K_2 = K_1 \cap \partial H_2$ yields $\dim(K_2) < \dim(K_1)$ and $F_{\bar{x}} \subset K_2$. As above, we again obtain $\tau_{K_2} = \tau_{K_1} = \tau_K$. We proceed in the same way but, thanks to the reduction in dimension, at most d times until $K_k = F_{\bar{x}}$ for some $k \in \{1, \ldots, d\}$. We then have $\tau_{F_{\bar{x}}} = \tau_{K_k} = \ldots = \tau_{K_1} = \tau_K$, which proves the statement. $\square$

Lemma 5.3. *Let K be a convex body and consider a point $x \in K$. Then $v(x) = 0$ if and only if $\dim(F_x) \leq 1$.*

Proof. Suppose $\dim(F_x) = 0$, so that $F_x = \{x\}$ is a singleton. Then X leaves F_x immediately under any $\mathsf{P} \in \mathcal{P}_x$, that is, $\tau_{F_x} = 0$. Suppose instead that $\dim(F_x) = 1$, so that F_x is a line segment. Then under any $\mathsf{P} \in \mathcal{P}_x$, X evolves like a one-dimensional Brownian motion along the line segment F_x, at least until τ_{F_x}. Thus $\mathsf{P}\text{-}\operatorname{ess\,inf} \tau_{F_x} = 0$, since X reaches the endpoints of F_x arbitrarily quickly with positive probability. By Lemma 5.2, we have $\mathsf{P}\text{-}\operatorname{ess\,inf} \tau_K = 0$. Therefore, in either case, we deduce that $v(x) = 0$. For the converse direction, assume that $\dim(F_x) > 1$. Then there exists a $\dim(F_x)$-dimensional closed ball $B \subset F_x$ with radius $r > 0$. Since $\tau_B \leq \tau_K$, Example 1.8 yields $v(x) \geq r^2 > 0$. $\square$

We now discuss continuity of the value function v. It was shown in Proposition 2.2(ii) that the value function v is upper semicontinuous. Therefore, if $v(x) = 0$ at a point $x \in K$, then v must be continuous at x. Of course, many convex bodies K have boundary faces of dimension two or higher, in which case Lemma 5.3 shows that v will not be zero everywhere on the boundary. Still, even in such cases, one might hope that v remains continuous. Unfortunately, this is not true in general, as the following example shows.

Example 5.4. Let $C = \{(x, x(1-x), 0) \colon x \in (0, 1]\}$ denote a half-open arc in the xy-plane. Next let $K_0 \subset \mathbb{R}^3$ be the closed convex hull of $\{(0, 0, 1), (0, 0, -1)\} \cup C$. Then every point $\bar{x}_0 \in C$ is an extreme point of K_0, but the origin $(0, 0, 0) \in \{(0, 0)\} \times [-1, 1] \subset K_0$ is not, despite being a limit point of C. Now, define $K = K_0 \times [-1, 1] \subset \mathbb{R}^4$, which is compact and convex. If $\bar{x} = (\bar{x}_0, 0) \in C \times \{0\}$, then $\dim(F_{\bar{x}}) = 1$, so $v(\bar{x}) = 0$ by Lemma 5.3. On the other hand, the boundary face containing the origin is the square $F_0 = \{(0, 0)\} \times (-1, 1)^2$ with $\dim(F_0) = 2$, so that $v(0) > 0$. Since the origin is a limit point of $C \times \{0\}$, we conclude that v is not continuous on K, despite K being convex.

In Example 5.4, continuity of v fails because $\mathcal{F}_1$ is not closed. One might therefore hope that continuity can be proved if $\mathcal{F}_1, \ldots, \mathcal{F}_d$ are closed. (Requiring $\mathcal{F}_0$ closed should be, and is, unnecessary because v is zero on all of $\mathcal{F}_1$.) This condition indeed turns out to imply continuity. The proof iterates over the k-skeletons, in each step making use of the following refined version of the argument in Proposition 2.5. The argument is probabilistic and rests on the dynamic programming principle.

Lemma 5.5. *Let K be a convex body, fix $k \in \{1, \ldots, d\}$, and assume $v|_{\mathrm{cl}(\mathcal{F}_{k-1})}$ is continuous. Then there is a modulus ω such that the following holds. If $\bar{x}, \bar{y} \in \mathcal{F}_k$, $\dim(F_{\bar{x}}) \leq \dim(F_{\bar{y}})$, A is an affine subspace containing $F_{\bar{x}}$, and Q is an orthogonal $d \times d$ matrix such that the map $x \mapsto Q(x - \bar{x}) + \bar{y}$ maps A to $\mathrm{aff}(F_{\bar{y}})$, then*

$$v(\bar{x}) \leq v(\bar{y}) + \omega(c|Q - I| + |\bar{x} - \bar{y}|),$$

where $c = \mathrm{diam}(K)$ is the diameter of K.

Proof. Since v is upper semicontinuous by Proposition 2.2(ii), since $\mathrm{cl}(\mathcal{F}_{k-1})$ is compact, and since $v|_{\mathrm{cl}(\mathcal{F}_{k-1})}$ is continuous by assumption, Lemma 2.4 gives a modulus ω such that

$$v(x) \leq v(y) + \omega(|x - y|) \text{ for all } x \in \mathbb{R}^d \text{ and } y \in \mathrm{cl}(\mathcal{F}_{k-1}). \tag{5.2}$$

We now show that ω satisfies the claimed property. To this end, let $\bar{x}$, $\bar{y}$, A, and Q be as in the statement of the lemma, and select an optimal law $\mathsf{P} \in \mathcal{P}_{\bar{x}}$. Lemma 5.2 asserts that $X(t) \in F_{\bar{x}}$ for all $t \leq \tau_K$, P-a.s. By modifying the behavior after τ_K, which does not affect the optimality of P, we may therefore assume that

$$X(t) \in A \text{ for all } t \geq 0, \text{ P-a.s.} \tag{5.3}$$

Consider the affine isometry $\Phi \colon A \to \mathrm{aff}(F_{\bar{y}})$ given by $\Phi(x) = Q(x - \bar{x}) + \bar{y}$. Using this isometry, define

$$Y = \Phi(X) \text{ and } \theta = \inf\{t \geq 0 \colon Y(t) \notin \mathrm{ri}(F_{\bar{y}})\}.$$

Note that $\mathsf{P}(\theta < \infty) = 1$ by Example 1.8. Due to (5.3), Y takes values in $\mathrm{aff}(F_{\bar{y}})$, and hence $Y(\theta) \in \mathrm{rbd}(F_{\bar{y}}) \subset \mathcal{F}_{k-1}$, P-a.s. Thus by (5.2) and monotonicity of ω we have, P-a.s.,

$$\begin{aligned}
v(X(\theta)) &\leq v(Y(\theta)) + \omega(|X(\theta) - Y(\theta)|) \\
&= v(Y(\theta)) + \omega(|(I - Q)(X(\theta) - \bar{x}) + \bar{x} - \bar{y}|) \\
&\leq v(Y(\theta)) + \omega(c|I - Q| + |\bar{x} - \bar{y}|),
\end{aligned}$$

where $c = \mathrm{diam}(K)$. We now combine this with two applications of the dynamic programming principle of Proposition 2.2(iii). This is permissible because θ is P-a.s. equal to an $\mathbb{F}^X$-stopping time, despite not being an $\mathbb{F}^X$-stopping time itself in general. We get

$$\begin{aligned}
v(\bar{x}) &= \mathsf{P}\text{-}\operatorname{ess\,inf}\{\theta \wedge \tau_K + v(X(\theta))\mathbf{1}_{\theta \leq \tau_K}\} \\
&\leq \mathsf{P}\text{-}\operatorname{ess\,inf}\{\theta + v(Y(\theta))\} + \omega(c|I - Q| + |\bar{x} - \bar{y}|) \\
&\leq v(\bar{y}) + \omega(c|I - Q| + |\bar{x} - \bar{y}|).
\end{aligned}$$

In the last inequality, the application of the dynamic programming principle uses that the law of Y lies in $\mathcal{P}_{\bar{y}}$ due to the isometry property of Φ, that $\mathbb{F}^Y = \mathbb{F}^X$, and that $\theta \leq \inf\{t \geq 0 \colon Y(t) \notin K\}$, P-a.s. This completes the proof. $\square$

We now state the key propagation of continuity result, analogous to Proposition 2.5. Part of the proof is convenient to phrase in terms of convergence of affine subspaces. For affine subspaces A_n and A of $\mathbb{R}^d$, we say that $A_n \to A$ if $\dim(A_n) = \dim(A)$ for all large n, there are points $x_n \in A_n$ and $x \in A$ such that $x_n \to x$, and $A_n - x_n$ converges to $A - x$ as elements of the Grassmannian $\mathrm{Gr}(\dim(A), \mathbb{R}^d)$ of $\dim(A)$-dimensional linear subspaces of $\mathbb{R}^d$. In this case, there exist orthogonal $d \times d$ matrices Q_n such that $Q_n \to I$ and the map $y \mapsto Q_n(y - x) + x_n$ maps A to A_n for n sufficiently large. The Grassmannian is known to be compact. Therefore, whenever the affine subspaces A_n contain points x_n that converge to some limit, it is possible to select a convergent subsequence of the A_n.

Lemma 5.6. *Let K be a convex body, fix $k \in \{1, \dots, d\}$, and assume $v|_{\mathrm{cl}(\mathcal{F}_{k-1})}$ is continuous. Then $v|_{\mathcal{F}_k}$ is also continuous.*

Proof. Since v is upper semicontinuous by Proposition 2.2(ii), it suffices to show that $v|_{\mathcal{F}_k}$ is lower semicontinuous. Since $v|_{\mathrm{cl}(\mathcal{F}_{k-1})}$ is continuous by assumption, this amounts to showing that

$$\bar{x} \in \mathcal{F}_k, \; x_n \in \mathcal{F}_k \setminus \mathrm{cl}(\mathcal{F}_{k-1}), \; x_n \to \bar{x}, \; v(x_n) \to \alpha \in \mathbb{R} \quad \Longrightarrow \quad v(\bar{x}) \leq \alpha. \tag{5.4}$$

Let therefore $\bar{x}, x_n, \alpha$ be as in (5.4). Define $r_n = \mathrm{dist}(x_n, \mathrm{rbd}(F_{x_n}))$. This is the radius of the largest k-dimensional ball centered at x_n and contained in F_{x_n}. We consider two separate cases.

Case 1: Suppose $\liminf_{n \to \infty} r_n = 0$. After passing to a subsequence, we have $r_n \to 0$. Then there exist points $y_n \in \mathrm{rbd}(F_{x_n})$ such that $|x_n - y_n| \to 0$. Thus $y_n \in \mathcal{F}_{k-1}$ and $y_n \to \bar{x}$, so that $\bar{x} \in \mathrm{cl}(\mathcal{F}_{k-1})$ and $v(y_n) \to v(\bar{x})$. Moreover, applying Lemma 5.5 with $\bar{x} = y_n$, $\bar{y} = x_n$, $A = \mathrm{aff}(F_{x_n})$, and $Q = I$ then gives

$$v(x_n) = v(y_n) + v(x_n) - v(y_n) \geq v(y_n) - \omega(|x_n - y_n|) \to v(\bar{x}).$$

Thus $v(\bar{x}) \leq \alpha$, proving (5.4) in this case.

Case 2: Suppose instead there exists $r > 0$ such that $r_n \geq r$ for all n. Then each F_{x_n} contains a k-dimensional ball B_n of radius r centered at x_n. After passing to a subsequence, we have $\mathrm{aff}(F_{x_n}) \to A$ for some k-dimensional affine subspace A. Thus there exist orthogonal $d \times d$ matrices Q_n such that $Q_n \to I$ and the affine isometry $\Phi_n \colon x \mapsto Q_n(x - \bar{x}) + x_n$ maps A to $\mathrm{aff}(F_{x_n})$ for each n. Now, let $B \subset A$ be the k-dimensional ball of radius r centered at $\bar{x}$. There is only one such ball, and we have $B_n = \Phi_n(B)$ for all n. For any $x \in B$ we thus have $\Phi_n(x) \in F_{x_n} \subset K$ and $\Phi_n(x) \to x$. Since K is closed, it follows that $B \subset K$. Thanks to [32, Theorem 2.1.2], B is contained in the relative interior of a unique face of K. Because $\bar{x}$ is in B, this face must be $F_{\bar{x}}$. That is, we have $B \subset F_{\bar{x}}$, so that $A = \mathrm{aff}(B) \subset \mathrm{aff}(F_{\bar{x}})$. On the other hand, $\dim(A) = k \geq \dim(F_{\bar{x}})$, so in fact $A = \mathrm{aff}(F_{\bar{x}})$. We now apply Lemma 5.5 with $\bar{x}$, $\bar{y} = x_n$, $A = \mathrm{aff}(F_{\bar{x}})$, and $Q = Q_n$ to get

$$v(\bar{x}) \leq v(x_n) + \omega(c|Q_n - I| + |\bar{x} - x_n|)$$

with $c = \mathrm{diam}(K)$. Sending $n \to \infty$ yields $v(\bar{x}) \leq \alpha$ and proves (5.4). $\qquad\square$

In view of Lemma 5.6, it is of interest to know whether the k-skeletons of a given convex body are closed. For some values of k, closedness is automatic.

Lemma 5.7. *Let K be a convex body. Then $\mathcal{F}_d$, $\mathcal{F}_{d-1}$, and $\mathcal{F}_{d-2}$ are closed.*

Proof. Both $\mathcal{F}_d = K$ and $\mathcal{F}_{d-1} = \partial K$ are closed. To see that $\mathcal{F}_{d-2}$ is closed, assume for contradiction that there is a point $\bar{x} \in \mathrm{cl}(\mathcal{F}_{d-2}) \setminus \mathcal{F}_{d-2}$. Then $\bar{x}$ lies in ∂K but not in $\mathcal{F}_{d-2}$, so must lie in the relative interior of a $(d-1)$-dimensional boundary face F. But then $\bar{x}$ admits an open neighborhood contained in $\mathrm{ri}(F) \cup \mathrm{int}(K) \cup K^c$, and therefore cannot lie in the closure of $\mathcal{F}_{d-2}$. This contradiction finishes the proof. $\qquad\square$

Here is the main result of this section.

Theorem 5.8. *Let K be a convex body with $\mathcal{F}_k$ closed for $1 \leq k \leq d - 3$. Then $v|_K$ is continuous.*

Proof. By Lemma 5.3, v vanishes on $\mathcal{F}_1$ and is therefore continuous there. Continuity on K now follows by repeated application of Lemma 5.6, making use of the closedness hypothesis on the $(k-1)$-skeletons for $1 \leq k-1 \leq d-3$, and Lemma 5.7 for $k-1 \geq d-2$. $\qquad\square$

As an immediate corollary, several interesting cases are covered.

Corollary 5.9. *Each of the following conditions implies that $v|_K$ is continuous.*

(i) *All boundary faces of K have dimension zero or one.*

(ii) $\dim(K) \in \{0, 1, 2, 3\}$.

(iii) *K is a (convex) polytope.*

Proof. (i) and (ii) are immediate from Theorem 5.8. As for (iii), if K is a polytope, then $\mathcal{F}_k$ is closed for all $k = 0, \dots, d$; see e.g. [28, Theorem 2.3] and the subsequent discussion. Now apply Theorem 5.8. □

Thanks to Lemma 5.3, if $v|_K$ is continuous then $\mathcal{F}_1$ is necessarily closed. Are the other k-skeletons also closed in this case? If $d = 4$ the answer is yes thanks to Lemma 5.7. In general the answer is no, as shown in Example 5.10 below. Continuity of the value function therefore cannot be used to characterize closedness of the k-skeletons.

Example 5.10. Recall the set $K_0 \subset \mathbb{R}^3$ from Example 5.4. Let $K' = K_0 \times \mathbb{R}^2 \subset \mathbb{R}^5$ and set

$$K = K' \cap \{(x, y, z, u, w) \in \mathbb{R}^5 : z^2 + u^2 + w^2 \le 1\}.$$

Then K is a convex body, and one can check that

$$\mathcal{F}_1 = \{(x, y, z, u, w) \in \mathbb{R}^5 : (x, y, z) \in \partial K_0, z^2 + u^2 + w^2 = 1\};$$
$$\mathcal{F}_2 = \{(x, y, z, u, w) \in \mathbb{R}^5 : (x, y, z) \in K_0, z^2 + u^2 + w^2 = 1\}$$
$$\cup \{(x, y, 0, u, w) \in \mathbb{R}^5 : x \in (0, 1], y = x(1 - x), u^2 + w^2 \le 1\};$$
$$\mathrm{cl}(\mathcal{F}_1) = \mathcal{F}_1;$$
$$\mathrm{cl}(\mathcal{F}_2) = \mathcal{F}_2 \cup \{(0, 0, 0, u, w) : u^2 + w^2 \le 1\}.$$

Since $\mathcal{F}_1$ is closed, Lemmas 5.3 and 5.6 imply that $v|_{\mathcal{F}_2}$ is continuous. We claim that $v|_{\mathrm{cl}(\mathcal{F}_2)}$ is also continuous, but since $\mathcal{F}_2 \ne \mathrm{cl}(\mathcal{F}_2)$ we must argue this directly at the remaining points of $\mathrm{cl}(\mathcal{F}_2)$. Consider therefore such a point $(0, 0, 0, u, w) \in \mathrm{cl}(\mathcal{F}_2)$ with $u^2 + w^2 \le 1$ and an approximating sequence of points $(x_n, y_n, z_n, u_n, w_n) \in \mathrm{cl}(\mathcal{F}_2)$. Then we know that $v(x_n, y_n, z_n, u_n, w_n) \ge 1 - z_n^2 - u_n^2 - w_n^2$ for all $n \in \mathbb{N}$ thanks to Example 1.8. Thus by upper semicontinuity, we have

$$1 - u^2 - w^2 = v(0, 0, 0, u, w) \ge \limsup_n v(x_n, y_n, z_n, u_n, w_n)$$
$$\ge \liminf_n v(x_n, y_n, z_n, u_n, w_n) \ge 1 - u^2 - w^2.$$

This yields continuity of $v|_{\mathrm{cl}(\mathcal{F}_2)}$. Hence by Lemma 5.6, $v|_{\mathcal{F}_3}$ is continuous. Lemma 5.7 yields that $\mathcal{F}_3$, $\mathcal{F}_4$, $\mathcal{F}_5 = K$ are closed. Repeating the previous argument shows that $v|_K$ is continuous, even though $\mathcal{F}_2$ is not closed.

6 Smooth value functions

The goal of this section is to prove Theorem 1.7. Let us first introduce some terminology. Let K be a convex body. We say that a function f lies in $C^2(K)$ if $f|_K$ is continuous, and the restriction $f|_{\mathrm{ri}(F)}$ to the relative interior of any face F of K lies in $C^2(\mathrm{ri}(F))$, understood in the usual sense of twice continuous differentiability on the $\dim(F)$-dimensional open set $\mathrm{ri}(F) \subset \mathrm{aff}(F)$. The gradient and Hessian computed relative to this set are then denoted $\nabla_K f(x) = \nabla(f|_{\mathrm{ri}(F)})(x) \in \mathrm{aff}(F - x)$ and $\nabla_K^2 f(x) = \nabla^2(f|_{\mathrm{ri}(F)})(x)$ for any $x \in \mathrm{ri}(F)$. Thus $\nabla_K f(x)$ and $\nabla_K^2 f(x)$ are the projections of $\nabla f(x)$ and $\nabla^2 f(x)$ onto $\mathrm{aff}(F - x)$, whenever the latter exist. A critical point of f in F is a point $x \in \mathrm{ri}(F)$ where $\nabla_K f(x) = 0$.

To prove Theorem 1.7, it is enough to prove the following.

Theorem 6.1. *Let $d \geq 2$ and let K be a convex body with at most countably many faces. Assume the value function v lies in $C^2(K)$. Assume also that in each face F of dimension at least two, either v has no critical point, or v has one single critical point which additionally is a maximum. Then for every $\bar{x} \in K$ there is an optimal solution $\mathsf{P} \in \mathcal{P}_{\bar{x}}$ under which $v(X(t)) = v(\bar{x}) - t$ for all $t < \tau_K$.*

Before proving Theorem 6.1 we show how it implies Theorem 1.7.

Proof of Theorem 1.7. The only part of Theorem 1.7 that does not follow directly from Theorem 6.1 is (1.10). Let $F_1 = F_{\bar{x}}$ be the face that contains $\bar{x}$ in its relative interior, and consider a nested sequence of faces $F_1 \supset F_2 \supset \cdots \supset F_m$ such that $\dim(F_i) > \dim(F_{i+1})$ for $i = 1, \ldots, m-1$. Let $\theta_0 = 0$ and define $\theta_i = \inf\{t \geq \theta_{i-1} \colon X(t) \notin \mathrm{ri}(F_i)\}$ for $i = 1, \ldots, m$. Consider now the event $A = \{X(\theta_i) \in \mathrm{ri}(F_{i+1})$ for $i = 1, \ldots, m-1$ and $\theta_m = \tau_K\}$. If A has positive probability we may condition on A and get from Itô's formula that

$$v(X(t)) - v(\bar{x}) = \sum_{i=1}^m \left(v(X(\theta_i \wedge t)) - v(X(\theta_{i-1} \wedge t))\right) = \sum_{i=1}^m \int_{\theta_{i-1} \wedge t}^{\theta_i \wedge t} \nabla v(X(s))^\top dX(s) + V(t)$$

for $t \leq \tau_K$ on the event A, where V is a finite variation process. That is, we get $v(X(t)) = v(\bar{x}) + \int_0^t \nabla v(X(s))^\top dX(s) + V(t)$ for $t \leq \tau_K$ on A. Since K has at most countably many faces and X cannot move to a higher-dimensional without also leaving K in view of (the proof of) Lemma 5.2, the events of the form A are at most countably many and, up to a nullset, form a partition of Ω. We deduce that $v(X(t)) = v(\bar{x}) + \int_0^t \nabla v(X(s))^\top dX(s) + V(t)$ for $t \leq \tau_K$ holds almost surely. Since $v(X(t)) = v(\bar{x}) - t$ for all $t < \tau_K$ by Theorem 6.1, we see that the left-hand side of (1.10) is a martingale of finite variation and hence zero, as claimed. $\square$

Observe that the assumption that v lies in $C^2(K)$ immediately implies that $v|_{\mathrm{aff}(F)}$ is a classical solution of (1.4) in $\mathrm{ri}(F)$ away from the critical point, for every face F of dimension at least two; just use v itself as test function in the definition of viscosity sub- and supersolutions.

The proof of Theorem 6.1 proceeds by first constructing solution laws P under which X behaves in the desired manner while inside any given face F of K. Then these laws are pasted together as X reaches ever lower-dimensional faces, until it leaves K. To implement this idea, for any face F of K with $\dim(F) \geq 2$ and any point $x \in \mathrm{ri}(F)$, we define

$$\mathcal{P}_x^* = \{\mathsf{P} \in \mathcal{P}_x \colon v(X(t)) = v(x) - t \text{ for all } t < \tau_{\mathrm{ri}(F)} \text{ and } X(\tau_{\mathrm{ri}(F)}) \in \mathrm{rbd}(F)\},$$

where $\tau_{\mathrm{ri}(F)} = \inf\{t \geq 0 \colon X(t) \notin \mathrm{ri}(F)\}$ is the first time X leaves the relative interior of F. For points $x \in K^c \cup \mathcal{F}_1$, we somewhat arbitrarily set $\mathcal{P}_x^* = \mathcal{P}_x$.

Let now the hypotheses of Theorem 6.1 be in force. Our first goal is to prove that $\mathcal{P}_x^*$ is nonempty for every $x \in \mathbb{R}^d$. This rests on the following construction of a martingale with increments in the kernel of a given location-dependent matrix.

Lemma 6.2. *Let $\mathcal{O} \subset \mathbb{R}^d$ be open and let $H \colon \mathbb{R}^d \to \mathbb{S}^d$ be a locally bounded measurable map such that $H|_{\mathcal{O}}$ is continuous and $\mathrm{rank}\, H(x) \leq d-1$ for all $x \in \mathcal{O}$. For every $\bar{x} \in \mathcal{O}$ there exists a continuous martingale Y with $Y(0) = \bar{x}$ and $\mathrm{tr}\langle Y \rangle(t) \equiv t$ such that*

$$\int_0^t H(Y(s)) d\langle Y \rangle(s) = 0, \quad t < \tau_{\mathcal{O}}, \tag{6.1}$$

where $\tau_{\mathcal{O}} = \inf\{t \geq 0 \colon Y(t) \notin \mathcal{O}\}$.

Proof. Define
$$a(x) = \begin{cases} \dfrac{I - H(x)^+ H(x)}{d - \operatorname{rank} H(x)}, & x \in \mathcal{O} \\ d^{-1} I, & x \notin \mathcal{O} \end{cases}$$

where $H(x)^+$ is the Moore–Penrose generalized inverse of $H(x)$. Thus if the spectral decomposition of $H(x)$ is $H(x) = Q \operatorname{diag}(\lambda_1, \dots, \lambda_r, 0, \dots, 0) Q^\top$ with $\lambda_i \neq 0$ for $i = 1, \dots, r$, then $H(x)^+ H(x) = Q \operatorname{diag}(1, \dots, 1, 0, \dots, 0) Q^\top$, where the diagonal matrix contains r ones. It follows that

$$a(x) \succeq 0 \text{ and } \operatorname{tr}(a(x)) = 1 \text{ for all } x \in \mathbb{R}^d, \text{ and } H(x)a(x) = 0 \text{ for all } x \in \mathcal{O}.$$

Unless H has constant rank on $\mathcal{O}$, a is not continuous on $\mathcal{O}$. Consider therefore mollifications
$$a_n(x) = \int_{\mathbb{R}^d} \varphi_n(x - y) a(y) dy,$$

where $\varphi_n(x) = n^d \varphi(nx)$ for a positive mollifier φ supported on the centered unit ball. Then a_n is continuous, positive semidefinite, and has unit trace. Thus there exist weak solutions Y_n of the SDEs

$$dY_n(t) = a_n(Y_n(t))^{1/2} dW(t), \quad Y_n(0) = \bar{x},$$

where the positive semidefinite square root is understood, and W is d-dimensional Brownian motion. The law of Y_n lies in $\mathcal{P}_{\bar{x}}$ for each n, so Proposition 2.2(i) shows that after passing to a subsequence, $Y_n \Rightarrow Y$ for some limiting process Y whose law again lies in $\mathcal{P}_{\bar{x}}$. Since $\langle Y_n \rangle(t) = \int_0^t a_n(Y_n(s)) ds$ and the a_n are uniformly bounded, after passing to a further subsequence we actually have $(Y_n, \langle Y_n \rangle) \Rightarrow (Y, Q)$ in the space $C(\mathbb{R}_+, \mathbb{R}^d \times \mathbb{S}^d)$ for some process Q. Since $Y_n Y_n^\top - \langle Y_n \rangle$ is a martingale for each n, and using the uniform bound on the quadratic variations, we may pass to the limit to deduce that $YY^\top - Q$ is also martingale, and hence $Q = \langle Y \rangle$. Furthermore, by Skorohod's representation theorem (see [8, Theorem 6.7]), we may assume that the $(Y_n, \langle Y_n \rangle)$ and $(Y, \langle Y \rangle)$ are defined on a common probability space $(\Omega', \mathcal{F}', \mathsf{P}')$ and that, almost surely, $(Y_n, \langle Y_n \rangle) \to (Y, \langle Y \rangle)$ in $C(\mathbb{R}_+, \mathbb{R}^d \times \mathbb{S}^d)$, that is, locally uniformly.

We now verify (6.1). We first claim that

$$\text{if } x \in \mathcal{O} \text{ and } x_n \to x \text{ then } H(x_n)a_n(x_n) \to 0. \tag{6.2}$$

To prove this, note that

$$H(x_n)a_n(x_n) = \int_{\mathbb{R}^d} \varphi_n(x_n - y) H(x_n) a(y) dy = \int_{\mathbb{R}^d} \varphi_n(x_n - y)(H(x_n) - H(y)) a(y) dy.$$

Since a is bounded and the restriction $H|_{\mathcal{O}}$ is continuous, arguing component by component, we see that the right-hand side converges to zero. This proves (6.2). Now pick $t < \tau_{\mathcal{O}}$. Then $Y(s) \in \mathcal{O}$ for all $s \leq t$. Since $(Y_n, \langle Y_n \rangle) \to (Y, \langle Y \rangle)$ locally uniformly, the bounded convergence theorem and (6.2) yield that

$$\int_0^t H(Y_n(s)) d\langle Y_n \rangle(s) = \int_0^t H(Y_n(s)) a_n(Y_n(s)) ds \to 0.$$

On the other hand, the left-hand side converges to $\int_0^t H(Y(s)) d\langle Y \rangle(s)$. This yields (6.1) and completes the proof of the lemma. $\qquad\square$

Remark 6.3. An examination of the proof of Lemma 6.2 shows that the process Y is of the form $dY_t = \sigma_t dW_t'$ for some Brownian motion W', where $\sigma_t = a(Y_t)^{1/2}$ for all t such

that a is continuous at Y_t. By properties of the Moore–Penrose inverse, a is continuous except on the boundaries of the sets $\{x \in \mathcal{O}\colon \operatorname{rank} H(x) = r\}$, $r = 0, \dots, d-1$ and on $\partial \mathcal{O}$. Thus if Y can be shown to spend zero time in these sets, it is a bona fide weak solution of $dY_t = a(Y_t)^{1/2} dW_t'$.

Proposition 6.4. *Continue to assume $v \in C^2(K)$. Then $\mathcal{P}_{\bar{x}}^*$ is nonempty for every $\bar{x} \in \mathbb{R}^d$.*

Proof. If $\bar{x} \in K^c \cup \mathcal{F}_1$, then $\mathcal{P}_{\bar{x}}^* = \mathcal{P}_{\bar{x}}$ and the statement is obvious. Below we prove the statement for $\bar{x} \in \operatorname{int}(K)$; the case $\bar{x} \in \operatorname{ri}(F)$ for a face F with $\dim(F) \geq 2$ is identical since all considerations are then restricted to $\operatorname{aff}(F)$. So suppose $\bar{x} \in \operatorname{int}(K)$ and, initially, also that $\bar{x}$ is not the maximizer of v over K; in particular $\bar{x}$ is not a critical point in K.

Since v lies in $C^2(K)$, it is a classical solution of (1.4) in $\operatorname{int}(K)$ away from the critical point. As explained in Section 1, an alternative form of this equation at non-critical points is (1.9). That is,

$$\lambda_{\min}\left(-\frac{1}{2} P_{\nabla v(x)} \nabla^2 v(x) P_{\nabla v(x)}, \nabla v(x)\right) = 1, \tag{6.3}$$

where $\lambda_{\min}(A, p)$ denotes the smallest eigenvalue of A corresponding to an eigenvector orthogonal to p, and

$$P_{\nabla v(x)} = I - \frac{\nabla v(x) \nabla v(x)^\top}{|\nabla v(x)|^2}.$$

Let $\mathcal{O} = \{x \in \operatorname{int}(K)\colon \nabla v(x) \neq 0\}$ be the set of non-critical points in $\operatorname{int}(K)$. Define

$$H(x) = \frac{1}{2} P_{\nabla v(x)} \nabla^2 v(x) P_{\nabla v(x)} + I, \quad x \in \mathcal{O},$$

and arbitrarily set $H(x) = 0$ for $x \notin \mathcal{O}$. It is clear that H is locally bounded measurable and that $H|_{\mathcal{O}}$ is continuous. Moreover, the equation (6.3) satisfied by v implies that $H(x)$ is singular, i.e. $\operatorname{rank} H(x) \leq d-1$, for all $x \in \mathcal{O}$. We may thus apply Lemma 6.2 to obtain a martingale Y as in the statement of that lemma whose law we denote by P. Clearly $\mathsf{P} \in \mathcal{P}_{\bar{x}}$, and due to (6.1) we have

$$H(X(t))a(t) = 0 \text{ on } [0, \tau_\mathcal{O}), \text{ } \mathsf{P}\text{-a.s.},$$

where $a(t)$ satisfies $\langle X \rangle = \int_0^{\cdot} a(s)ds$ and $\operatorname{tr}(a(t)) = 1$, and $\tau_\mathcal{O} = \inf\{t \geq 0\colon X(t) \notin \mathcal{O}\}$. As a consequence, omitting the argument $X(t)$ for readability, we have for $t < \tau_\mathcal{O}$ that

$$0 = \nabla v^\top H a(t) = \frac{1}{2} \nabla v^\top P_{\nabla v} \nabla^2 v P_{\nabla v} a(t) + \nabla v^\top a(t) = \nabla v^\top a(t).$$

We thus have $\nabla v^\top a(t) = 0$, which yields $P_{\nabla v} a(t) P_{\nabla v} = a(t)$. Consequently,

$$0 = \operatorname{tr}(H a(t)) = 1 + \operatorname{tr}\left(\frac{1}{2} P_{\nabla v} \nabla^2 v P_{\nabla v} a(t)\right) = 1 + \frac{1}{2} \operatorname{tr}(a(t) \nabla^2 v).$$

An application of Itô's formula now gives

$$dv(X(t)) = \nabla v(X(t))^\top dX(t) + \frac{1}{2} \operatorname{tr}(a(t) \nabla^2 v(X(t)))dt = -dt.$$

These computations are valid for $t < \tau_\mathcal{O}$, so we deduce that $v(X(t)) = v(\bar{x}) - t$ for $t < \tau_\mathcal{O}$. In particular, $X(t)$ will not attain a critical point before $\tau_\mathcal{O}$, so in fact $\tau_\mathcal{O} = \tau_{\operatorname{int}(K)}$, the first exit time from $\operatorname{int}(K)$. Moreover, at the exit time, we have $X(\tau_{\operatorname{int}(K)}) \in \partial K$. This shows that $\mathsf{P} \in \mathcal{P}_{\bar{x}}^*$, as desired.

The case where $\bar{x}$ is a critical point still remains. In this case, we select points $x_n \in \operatorname{int}(K) \setminus \{\bar{x}\}$ with $x_n \to \bar{x}$, and let $\mathsf{P}_n \in \mathcal{P}_{x_n}^*$. In particular, the laws $\mathsf{Q}_n = (\cdot - x_n)_* \mathsf{P}_{x_n}$ lie in

$\mathcal{P}_0$, which is compact by Proposition 2.2(i). The Q_n are thus subsequentially convergent toward some $\mathsf{Q} \in \mathcal{P}_0$. Along this subsequence, the P_n converge to $\mathsf{P} = (\cdot + \bar{x})_* \mathsf{Q} \in \mathcal{P}_{\bar{x}}$. Lemma 6.5 below shows that the properties $v(X(t)) = v(X(0)) - t$ for all $t < \tau_{\text{int}(K)}$ and $X(\tau_{\text{int}(K)}) \in \partial K$ if $\tau_{\text{int}(K)} < \infty$ carry over to weak limits. This shows that $\mathsf{P} \in \mathcal{P}_x^*$, and completes the proof of the proposition. $\qquad\square$

We now turn to the task of pasting solutions together as X reaches ever lower-dimensional faces of K. This uses a measurable selection of laws from $\mathcal{P}_x^*$, which in turn requires suitable closedness properties of these sets. The following closedness result was already used in the proof of Proposition 6.4.

Lemma 6.5. *Let F be a face of K and write $\sigma = \tau_{\text{ri}(F)}$ for brevity. Then the set*

$$ C_F = \{\omega \in \Omega \colon v(\omega(t)) = v(\omega(0)) - t \ \forall\, t < \sigma(\omega),\ \text{and}\ \omega(\sigma(\omega)) \in \text{rbd}(F)\ \text{if}\ \sigma(\omega) < \infty\} $$

is closed in Ω. As a consequence, $\{\mathsf{P} \in \mathcal{P}(\Omega) \colon \mathsf{P}(C_F) = 1\}$ is closed in $\mathcal{P}(\Omega)$. The same conclusion holds if v is only known to be continuous, not necessarily $C^2(K)$.

Proof. It suffices to prove that C_F is closed, as the second statement then follows from the Portmanteau lemma. Pick $\omega_n \in C_F$ with $\omega_n \to \omega$ in Ω. Define $T = \liminf_n \sigma(\omega_n) \in [0, \infty]$, and pass to a subsequence to get $T = \lim_n \sigma(\omega_n)$. Then $v(\omega_n(t)) = v(\omega_n(0)) - t$ and $\omega_n(t) \in F$ if n is sufficiently large, for all $t < T$. Since v is continuous and F closed, we get $v(\omega(t)) = v(\omega(0)) - t$ and $\omega(t) \in F$ for all $t < T$. Provided $\sigma(\omega) \leq T$, this implies $\omega \in C_F$ and proves closedness. If $T = \infty$ then of course $\sigma(\omega) \leq T$. If $T < \infty$, then by definition of T we have $\sigma(\omega_n) \leq T + \varepsilon$ for any $\varepsilon > 0$ and all large n. Thus $\min_{t \leq T+\varepsilon} \text{dist}(\omega_n(t), \text{rbd}(F)) = 0$ for all large n. By continuity we get $\min_{t \leq T+\varepsilon} \text{dist}(\omega(t), \text{rbd}(F)) = 0$, and hence $\sigma(\omega) \leq T + \varepsilon$. Since $\varepsilon > 0$ was arbitrary, this yields $\sigma(\omega) \leq T$ as required. $\qquad\square$

The following lemma produces the required measurable selection. This is actually the only step that uses that K has countably many faces. If the lemma could be established without assuming this, the assumption could be dropped from Theorem 6.1 (and Theorem 1.7). In fact, the current proof works for the more general situation where K has countably many faces of dimension two and higher, and arbitrarily many faces of dimension zero and one.

Lemma 6.6. *Assume K has countably many faces, and continue to assume $v \in C^2(K)$. Them there is a measurable map $x \mapsto \mathsf{P}_x$ from $\mathbb{R}^d$ to $\mathcal{P}(\Omega)$ such that $\mathsf{P}_x \in \mathcal{P}_x^*$ for all x.*

Proof. We apply the selection theorem of Kuratowski and Ryll-Nardzewski; see [1, Theorem 18.13]. This requires that the set-valued map $x \mapsto \mathcal{P}_x^*$ be weakly measurable with nonempty closed values. By Proposition 6.4, $\mathcal{P}_x^*$ is nonempty for all x. For $x \in K^c \cup \mathcal{F}_1$, $\mathcal{P}_x^* = \mathcal{P}_x$ is closed (even compact) by Proposition 2.2(i). If F is a face of K with $\dim(F) \geq 2$ and $x \in \text{ri}(F)$, then

$$ \mathcal{P}_x^* = \mathcal{P}_x \cap \{\mathsf{P} \in \mathcal{P}(\Omega) \colon \mathsf{P}(C_F) = 1\}, $$

which is closed by Lemma 6.5. So $\mathcal{P}_x^*$ is closed for all x.

We now argue weak measurability, initially for the map $x \mapsto \mathcal{P}_x$. We must show that for every open subset $U \subset \mathcal{P}(\Omega)$, the set $\{x \in \mathbb{R}^d \colon \mathcal{P}_x \cap U \neq \emptyset\}$ is measurable; see [1, Definition 18.1]. But since $\mathcal{P}_x = (\cdot + x)_* \mathcal{P}_0$, the condition $\mathcal{P}_x \cap U \neq \emptyset$ means that there exists $\mathsf{P} \in \mathcal{P}_0$ such that $(\cdot + x)_* \mathsf{P} \in U$. If this holds for some $x \in \mathbb{R}^d$, then it also holds for all y in a neighborhood of x since U is open and $x \mapsto (\cdot + x)_* \mathsf{P}$ is continuous. Thus $\{x \in \mathbb{R}^d \colon \mathcal{P}_x \cap U \neq \emptyset\}$ is actually open, and in particular measurable. So $x \mapsto \mathcal{P}_x$ is weakly measurable.

Furthermore, the set-valued map $x \mapsto \varphi(x)$ specified by $\varphi(x) = \mathcal{P}(\Omega)$ for $x \in K^c \cup \mathcal{F}_1$ and $\varphi(x) = \{\mathsf{P} \in \mathcal{P}(\Omega) \colon \mathsf{P}(C_F) = 1\}$ for $x \in \text{ri}(F)$ is constant on K^c and on each face of

K. Since K has countably many faces, we deduce that $x \mapsto \varphi(x)$ is weakly measurable. By [1, Lemma 18.4(3)], it now follows that $x \mapsto \mathcal{P}_x^* = \mathcal{P}_x \cap \varphi(x)$ is weakly measurable, as required. $\qquad\square$

Proof of Theorem 6.1. For $k = 2, \ldots, d$, define $U_k = \mathcal{F}_k \setminus \mathcal{F}_{k-1}$. Equivalently, U_k is the (possibly empty) union of the relative interiors of all k-dimensional faces of K. We work on the d-fold product $\Omega^d = C(\mathbb{R}_+, \mathbb{R}^d)^d$ of the canonical path space, and let $(W, Y^2, \ldots, Y^d)$ be the $(\mathbb{R}^d)^d$-valued coordinate process. Let $x \mapsto \mathsf{P}_x \in \mathcal{P}_x^*$ be the measurable map given by Lemma 6.6; we will use it to specify the law of $Y^2, \ldots, Y^d$. Define random times $\tau_{k-1} = \inf\{t \geq 0 : Y^k(t) \notin U_k\}$. Given $\bar{x} \in K$, let Y^d have law $\mathsf{P}_{\bar{x}}$. Next, if the law of $(Y^d, \ldots, Y^k)$ has been specified for $k \geq 3$, then specify the law of Y^{k-1} to be conditionally independent of $(Y^d, \ldots, Y^k)$ given $Y^k(\tau_{k-1})$, which is finite almost surely, with law $Y^{k-1} \sim \mathsf{P}_{Y^k(\tau_{k-1})}$. That is, the regular conditional distribution of Y^{k-1} given $Y^k(\tau_{k-1}) = y$ is P_y. This procedure specifies the law of $Y^2, \ldots, Y^d$. Finally, let W have the law of an independent standard d-dimensional Brownian motion. Now, set $\tau_d = 0$ and define a process Y by

$$Y(t) = Y^k(t - (\tau_k + \cdots + \tau_d)), \quad t \in [\tau_k + \cdots + \tau_d, \tau_{k-1} + \cdots + \tau_d),$$

for $k = 2, \ldots, d$, and

$$Y(t) = Y^2(\tau_1) + d^{-1/2}W(t - (\tau_1 + \cdots + \tau_d)), \quad t \geq \tau_1 + \cdots + \tau_d.$$

Thus Y first follows the dynamics of Y^d while in the interior of K (a possibly empty time interval); then Y follows the dynamics of Y^{d-1} while inside the relative interior of a $(d-1)$-dimensional face, and so on, until it reaches a face of dimension zero or one. From that point onwards, it follows a Brownian motion, scaled so that the quadratic variation has unit trace. Since the law of each Y^k is chosen from the sets $\mathcal{P}_x^*$, it is straightforward but somewhat tedious to make this intuitive description rigorous. One also finds that Y is a continuous martingale, starting at $Y(0) = \bar{x}$ and with $\operatorname{tr}\langle Y\rangle(t) \equiv t$, and (using that $v|_K$ is continuous) such that $v(Y(t)) = v(\bar{x}) - t$ for all $t < \tau_{K \setminus \mathcal{F}_1} = \inf\{t \geq 0 : Y(t) \notin K \setminus \mathcal{F}_1\}$. Moreover, Y does not leave K before reaching $\mathcal{F}_1$, but then leaves K immediately since its dynamics switches to that of a scaled standard Brownian motion in $d \geq 2$ dimensions. In particular, $\tau_{K \setminus \mathcal{F}_1} = \tau_K$, and we have $\tau_K = v(\bar{x}) - v(Y(\tau_K)) = v(\bar{x})$, using also that $v = 0$ on $\mathcal{F}_1$. The law P of Y is therefore the required optimal law. $\qquad\square$

A A useful maximum principle

The proof of Theorem 4.1 relies on the following maximum principle, which holds for arbitrary compact sets K. The boundary conditions in its statement should be understood in the viscosity sense.

Theorem A.1. *Let $d \geq 2$ and suppose K is compact. Let u (w) be an upper (lower) semicontinuous viscosity subsolution (supersolution) of (1.4). Then there exists $\bar{x} \in \partial K$ that achieves $\max_K(u - w)$. Moreover, if u in addition satisfies the zero boundary condition, and if w is a lower semicontinuous viscosity supersolution of (1.4) with zero boundary condition on some compact set K' such that $K \subset \operatorname{int}(K')$, then $u \leq w$ on K.*

Proof. We proceed in several steps.

1. It is enough to prove the two assertions with u replaced by δu, for each $\delta \in (0, 1)$. Indeed, if $\delta u \leq w$ for all $\delta \in (0, 1)$ then also $u \leq w$, yielding the second assertion. For the first assertion, assume we have $\bar{x}_\delta \in \partial K$ that achieves $\max_K(\delta u - w)$, for each $\delta \in (0, 1)$. Then there exists a sequence (δ_n) such that $\lim_n \delta_n = 1$ and $\lim_n \bar{x}_{\delta_n} = \bar{x}$ for some $\bar{x} \in \partial K$.

Then for all $x \in K$ we have

$$\delta_n u(x) - w(x) \le \delta_n u(\bar{x}_{\delta_n}) - w(\bar{x}_{\delta_n}) = u(\bar{x}_{\delta_n}) - w(\bar{x}_{\delta_n}) - (1 - \delta_n) u(\bar{x}_{\delta_n}).$$

Sending n to infinity and using upper semicontinuity of $u - w$ then shows that $\bar{x}$ achieves $\max_K(u - w)$ as required. Now, δu is a subsolution of the equation

$$F(\nabla u, \nabla^2 u) = \delta, \tag{A.1}$$

and if u satisfies the zero boundary condition, then so does δu. Thus, by writing u instead of δu, we may and do assume throughout the proof that u itself is a subsolution of (A.1), where $\delta \in (0, 1)$ is arbitrary but fixed.

2. For the first assertion, for every $\varepsilon > 0$, define

$$\Phi_\varepsilon(x, y) = u(x) - w(y) - \frac{1}{\varepsilon^4} |x - y|^4$$

for $(x, y) \in K \times K$, and let $(x_\varepsilon, y_\varepsilon)$ maximize Φ_ε over $K \times K$. Then we have

$$\Phi_\varepsilon(x_\varepsilon, y_\varepsilon) \ge \max_{x \in K} \Phi_\varepsilon(x, x) = \max_K(u - w). \tag{A.2}$$

By compactness, $(x_\varepsilon, y_\varepsilon)$ converges to some $(\bar{x}, \bar{y}) \in K \times K$ as $\varepsilon \to 0$ along a suitable subsequence; in the following, ε is always understood to belong to this subsequence. Since $\varepsilon^{-4} |x_\varepsilon - y_\varepsilon|^4 \le \max_K u - \min_K w - \max_K(u - w)$, we actually have $\bar{x} = \bar{y}$. Moreover, (A.2) yields

$$\max_K(u - w) \le \limsup_{\varepsilon \to 0} \Phi_\varepsilon(x_\varepsilon, y_\varepsilon) \le \limsup_{\varepsilon \to 0} (u(x_\varepsilon) - w(y_\varepsilon)) \le u(\bar{x}) - w(\bar{x}),$$

by upper semicontinuity of u and of $-w$. Hence $\bar{x}$ maximizes $u - w$ over K. Thus, to show the first assertion it suffices to argue that $(x_\varepsilon, y_\varepsilon) \in \text{int}(K) \times \text{int}(K)$ is not possible. This forces $\bar{x} \in \partial K$ as desired.

For the second assertion, we define Φ_ε as above, but now on the larger set $K \times K'$. Let $(x_\varepsilon, y_\varepsilon)$ again denote the corresponding maximizers, which converge along a subsequence to some $(\bar{x}, \bar{y}) \in K \times K'$. We again obtain $\bar{x} = \bar{y}$, thus $\bar{y} \in K \subset \text{int}(K')$, and therefore $y_\varepsilon \in \text{int}(K')$ for all sufficiently small ε. We will use this to argue that u cannot satisfy the viscosity inequality at x_ε. This forces $x_\varepsilon \in \partial K$ and, due to the boundary condition, $u(x_\varepsilon) \le 0$. Together with nonnegativity of w (see Lemma A.2 below) this yields $\Phi_\varepsilon(x_\varepsilon, y_\varepsilon) \le 0$ and thus, thanks to (A.2), $\max_K(u - w) \le 0$. This is the second assertion.

3. Both assertions can now be argued by contradiction in the same manner: for any fixed small $\varepsilon > 0$, we assume that both u and w simultaneously satisfy the viscosity inequalities at x_ε and y_ε, respectively, and use this to derive a contradiction. (Indeed, to prove the first assertion we had to exclude that $(x_\varepsilon, y_\varepsilon) \in \text{int}(K) \times \text{int}(K)$, while for the second assertion we had to exclude that u satisfies the viscosity inequality at x_ε.)

4. Let us work under the assumptions of Step 3. Define

$$\zeta(x, y) = \frac{1}{\varepsilon^4} |x - y|^4.$$

To simplify notation, write

$$p = \nabla_x \zeta(x_\varepsilon, y_\varepsilon) = 2\varepsilon^{-4} |x_\varepsilon - y_\varepsilon|^2 (x_\varepsilon - y_\varepsilon),$$
$$H = \nabla_{xx}^2 \zeta(x_\varepsilon, y_\varepsilon) = 2\varepsilon^{-4} |x_\varepsilon - y_\varepsilon|^2 I + 4\varepsilon^{-4} (x_\varepsilon - y_\varepsilon)(x_\varepsilon - y_\varepsilon)^\top.$$

Then $\nabla_y \zeta(x_\varepsilon, y_\varepsilon) = -p$, $\nabla_{xy}^2 \zeta(x_\varepsilon, y_\varepsilon) = -H$, and $\nabla_{yy}^2 \zeta(x_\varepsilon, y_\varepsilon) = H$. We also define

$$A = \nabla^2 \zeta(x_\varepsilon, y_\varepsilon) = \begin{pmatrix} H & -H \\ -H & H \end{pmatrix}. \tag{A.3}$$

We now claim that $p \neq 0$. Suppose for contradiction that $p = 0$. Then $x_\varepsilon = y_\varepsilon$, $\nabla_y \zeta(x_\varepsilon, y_\varepsilon) = 0$, and $\nabla_{yy}\zeta(x_\varepsilon, y_\varepsilon) = 0$. Since y_ε minimizes $y \mapsto w(y) + \zeta(x_\varepsilon, y)$ over K (respectively, over K'), the supersolution inequality states that $0 = F^*(0,0) \geq 1$. This contradiction confirms that $p \neq 0$.

Ishii's lemma, see [15, Theorem 3.2], now gives $M, N \in \mathbb{S}^d$ such that

$$(p, M) \in \overline{J}_K^{2,+} u(x_\varepsilon), \qquad (-p, N) \in \overline{J}_K^{2,-} w(y_\varepsilon),$$

and

$$\begin{pmatrix} M & 0 \\ 0 & -N \end{pmatrix} \preceq A + A^2. \tag{A.4}$$

Pre- and post-multiplying (A.4) by vectors of the form (z, z) and using (A.3) shows that $M \preceq N$. Now we use the fact that (p, M) lies in the limiting superjet of the subsolution u at x_ε, the ellipticity of F, Lemma 3.1, the fact that $p \neq 0$, and finally that $(-p, N)$ lies in the limiting subjet of the supersolution w at y_ε to get

$$\delta \geq F(p, M) \geq F(p, N) = F(-p, N) = F^*(-p, N) \geq 1.$$

This is the required contradiction, which concludes the proof. $\qquad \square$

We used the following observation in the previous proof. The boundary condition in its statement should be understood in the viscosity sense.

Lemma A.2. *If w is a lower semicontinuous viscosity supersolution of* (1.4) *with zero boundary condition on some compact $K \subset \mathbb{R}^d$ with $d \geq 2$ then $w \geq 0$.*

Proof. The constant test function $\varphi \equiv \min_K w$ certifies that $w \geq 0$. Indeed, if $\bar{x}$ minimizes w over K and $w(\bar{x}) < 0$, then the supersolution inequality holds regardless of whether $\bar{x}$ lies in the interior or on the boundary. Thus $0 = F^*(\nabla\varphi(\bar{x}), \nabla^2\varphi(\bar{x})) = F^*(0,0) \geq 1$, a contradiction. So $w(\bar{x}) \geq 0$. $\qquad \square$

References

[1] Charalambos D. Aliprantis and Kim C. Border, *Infinite dimensional analysis*, third ed., Springer, Berlin, 2006, A Hitchhiker's Guide. MR2378491

[2] Luigi Ambrosio and H. Mete Soner, *Level set approach to mean curvature flow in arbitrary codimension*, J. Differential Geom. **43** (1996), no. 4, 693–737. MR1412682

[3] Guy Barles, H. Mete Soner, and Panagiotis E. Souganidis, *Front propagation and phase field theory*, SIAM J. Control Optim. **31** (1993), no. 2, 439–469. MR1205984

[4] Guy Barles and Francesca Da Lio, *On the generalized Dirichlet problem for viscous Hamilton-Jacobi equations*, J. Math. Pures Appl. (9) **83** (2004), no. 1, 53–75. MR2023054

[5] Guy Barles, Elisabeth Rouy, and Panagiotis E. Souganidis, *Remarks on the Dirichlet problem for quasilinear elliptic and parabolic equations*, in: Stochastic Analysis, Control, Optimization and Applications, Systems Control Found. Appl., Birkhäuser Boston, Boston, MA, 1999, pp. 209–222. MR1702961

[6] Emmanuel N. Barron, Pierre Cardaliaguet, and Robert Jensen, *Conditional essential suprema with applications*, Appl. Math. Optim. **48** (2003), no. 3, 229–253. MR2004284

[7] Dimitri P. Bertsekas and Steven E. Shreve, *Stochastic optimal control*, Mathematics in Science and Engineering, vol. 139, Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London, 1978, The Discrete Time Case. MR511544

[8] Patrick Billingsley, *Convergence of probability measures*, second ed., Wiley Series in Probability and Statistics: Probability and Statistics, John Wiley & Sons, Inc., New York, 1999, A Wiley-Interscience Publication. MR1700749

[9] Bruno Bouchard and Thanh Nam Vu, *The obstacle version of the geometric dynamic programming principle: application to the pricing of American options under constraints*, Appl. Math. Optim. **61** (2010), no. 2, 235–265. MR2585143

[10] Bruno Bouchard and Thanh Nam Vu, *The obstacle version of the geometric dynamic programming principle: application to the pricing of American options under constraints*, Appl. Math. Optim. **61** (2010), no. 2, 235–265. MR2585143

[11] Rainer Buckdahn, Pierre Cardaliaguet, and Marc Quincampoix, *A representation formula for the mean curvature motion*, SIAM J. Math. Anal. **33** (2001), no. 4, 827–846. MR1884724

[12] Jeff Calder and Charles K. Smart, *The limit shape of convex hull peeling*, Duke Math. J. **169** (2020), no. 11, 2079–2124. MR4132581

[13] Yun Gang Chen, Yoshikazu Giga, and Shun'ichi Goto, *Uniqueness and existence of viscosity solutions of generalized mean curvature flow equations*, J. Differential Geom. **33** (1991), no. 3, 749–786. MR1100211

[14] David Chopp, Lawrence C. Evans, and Hitosh Ishii, *Waiting time effects for Gauss curvature flows*, Indiana Univ. Math. J. **48** (1999), no. 1, 311–334. MR1722203

[15] Michael G. Crandall, Hitoshi Ishii, and Pierre-Louis Lions, *User's guide to viscosity solutions of second order partial differential equations*, Bull. Amer. Math. Soc. (N.S.) **27** (1992), no. 1, 1–67. MR1118699

[16] Panagiota Daskalopoulos and Ki-Ahm Lee, *Worn stones with flat sides all time regularity of the interface*, Invent. Math. **156** (2004), no. 3, 445–493. MR2061326

[17] Claude Dellacherie and Paul-André Meyer, *Probabilities and Potential*, volume 29 of *North-Holland Mathematics Studies*. North-Holland Publishing Co., Amsterdam-New York, 1978. MR0521810

[18] Nadejda Drenska and Robert V. Kohn, *Prediction with expert advice: a PDE perspective*, J. Nonlinear Sci. **30** (2020), no. 1, 137–173. MR4053484

[19] Lawrence C. Evans and Joel Spruck, *Motion of level sets by mean curvature. I*, J. Differential Geom. **33** (1991), no. 3, 635–681. MR1100206

[20] Lawrence C. Evans and Joel Spruck, *Motion of level sets by mean curvature. III*, J. Geom. Anal. **2** (1992), no. 2, 121–150. MR1151756

[21] E. Robert Fernholz, Ioannis Karatzas, and Johannes Ruf, *Volatility and arbitrage*, Ann. Appl. Probab. **28** (2018), no. 1, 378–417. MR3770880

[22] Richard S. Hamilton, *Worn stones with flat sides*, in: A Tribute to Ilya Bakelman (College Station, TX, 1993), Discourses Math. Appl., vol. 3, Texas A&M Univ., College Station, TX, 1994, pp. 69–78. MR1423370

[23] Gerhard Huisken, *Local and global behaviour of hypersurfaces moving by mean curvature*, Differential geometry: partial differential equations on manifolds (Los Angeles, CA, 1990), Proc. Sympos. Pure Math., vol. 54, Amer. Math. Soc., Providence, RI, 1993, pp. 175–191. MR1216584

[24] Robert V. Kohn and Sylvia Serfaty, *A deterministic-control-based approach to motion by curvature*, Comm. Pure Appl. Math. **59** (2006), no. 3, 344–407. MR2200259

[25] Martin Larsson, *Conditional infimum and recovery of monotone processes*, arXiv:1802.08628 (2018).

[26] Martin Larsson and Johannes Ruf, *Relative arbitrage: sharp time horizons and motion by curvature*, Math. Finance **31** (2021), no. 3, 885–906. MR4273539

[27] Stanley Osher and James A. Sethian, *Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations*, J. Comput. Phys. **79** (1988), no. 1, 12–49. MR965860

[28] Susanna Papadopoulou, *On the geometry of stable compact convex sets*, Math. Ann. **229** (1977), no. 3, 193–200. MR0450938

[29] Yuval Peres, Oded Schramm, Scott Sheffield, and David B. Wilson, *Tug-of-war and the infinity Laplacian*, J. Amer. Math. Soc. **22** (2009), no. 1, 167–210. MR2449057

[30] Daniel Revuz and Marc Yor, *Continuous martingales and Brownian motion*, third ed., Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 293, Springer-Verlag, Berlin, 1999. MR1725357

[31] R. Tyrrell Rockafellar, *Convex analysis*, Princeton Mathematical Series, No. 28, Princeton University Press, Princeton, N.J., 1970. MR0274683

[32] Rolf Schneider, *Convex bodies: the Brunn-Minkowski theory*, expanded ed., Encyclopedia of Mathematics and its Applications, vol. 151, Cambridge University Press, Cambridge, 2014. MR3155183

[33] H. Mete Soner and Nizar Touzi, *Dynamic programming for stochastic target problems and geometric flows*, J. Eur. Math. Soc. (JEMS) **4** (2002), no. 3, 201–236. MR1924400

[34] H. Mete Soner and Nizar Touzi, *A stochastic representation for the level set equations*, Comm. Partial Differential Equations **27** (2002), no. 9-10, 2031–2053. MR1941665

[35] H. Mete Soner and Nizar Touzi, *A stochastic representation for mean curvature type geometric flows*, Ann. Probab. **31** (2003), no. 3, 1145–1165. MR1988466

[36] H. Mete Soner, *Optimal control with state-space constraint. I*, SIAM J. Control Optim. **24** (1986), no. 3, 552–561. MR838056

[37] H. Mete Soner, *Optimal control with state-space constraint. II*, SIAM J. Control Optim. **24** (1986), no. 6, 1110–1122. MR861089

[38] H. Mete Soner, *Motion of a set by the curvature of its boundary*, J. Differential Equations **101** (1993), no. 2, 313–372. MR1204331

[39] Joel Spencer, *Balancing games*, J. Combinatorial Theory Ser. B **23** (1977), no. 1, 68–74. MR526057

[40] Daniel W. Stroock, *On the growth of stochastic integrals*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **18** (1971), 340–344. MR287622

[41] Daniel W. Stroock and S. R. Srinivasa Varadhan, *Multidimensional diffusion processes*, Classics in Mathematics, Springer-Verlag, Berlin, 2006, Reprint of the 1997 edition. MR2190038