ORIGINAL ARTICLE

Relative arbitrage: Sharp time horizons and motion by curvature

Martin Larsson[1] | **Johannes Ruf**[2] (ORCID)

[1] Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA

[2] Department of Mathematics, London School of Economics and Political Science, London, UK

Correspondence
Martin Larsson, Department of Mathematical Sciences, Carnegie Mellon University, Wean Hall, 5000 Forbes Ave, Pittsburgh, PA 15213, USA.
Email: martinl@andrew.cmu.edu;
Johannes Ruf, Department of Mathematics, London School of Economics and Political Science, Columbia House, London, WC2A 2AE, UK.
Email: j.ruf@lse.ac.uk

Abstract

We characterize the minimal time horizon over which any equity market with $d \geq 2$ stocks and sufficient intrinsic volatility admits relative arbitrage with respect to the market portfolio. If $d \in \{2, 3\}$, the minimal time horizon can be computed explicitly, its value being zero if $d = 2$ and $\sqrt{3}/(2\pi)$ if $d = 3$. If $d \geq 4$, the minimal time horizon can be characterized via the arrival time function of a geometric flow of the unit simplex in $\mathbb{R}^d$ that we call the minimum curvature flow.

KEYWORDS
arbitrage, geometric flow, stochastic control, stochastic portfolio theory

JEL CLASSIFICATION
35J60, 49L25, 60G44, 91G10

1 INTRODUCTION

Consider a market with $d \geq 2$ stocks, modeled by a d-dimensional vector $\mu = (\mu_1, \ldots, \mu_d)^\top$ of continuous nonnegative semimartingales summing to one. Each component of μ corresponds to the relative market weight of one stock; thus μ_i is the capitalization of the i-th stock divided by the total market capitalization.

Fernholz and Karatzas (2005) observed that if all market weights are strictly positive and satisfy

$$\sum_{i=1}^{d} \int_0^t \frac{1}{\mu_i} \mathrm{d}[\mu_i, \mu_i] \geq t \quad \text{for all } t \geq 0, \tag{1}$$

then there exists some time $\overline{T}$ such that relative arbitrage is possible with respect to the market over any time horizon $[0, T]$ with $T > \overline{T}$ (precise definitions are provided below). However, this left open the question whether such markets allow for relative arbitrage over arbitrary time horizons, including very short ones. About ten years later, Fernholz et al. (2018) showed that this is not always possible. They provided an example of a market μ that satisfies (1) but does not allow for short-term relative arbitrage. Still, it remains open how to characterize the precise time $T_* \in (0, \overline{T}]$ such that a market satisfying (1) always allows for relative arbitrage over $[0, T]$ when $T > T_*$, but not always when $T < T_*$.

In this paper we take a step toward answering this question. To do so, we work with a strengthening of (1), only considering markets that satisfy the trace condition

$$\mathrm{tr}[\mu, \mu](t) = \sum_{i=1}^{d} [\mu_i, \mu_i](t) \geq t \quad \text{for all } t \geq 0. \tag{2}$$

Our goal is to determine the smallest time T_* such that every market satisfying (2) allows for relative arbitrage over any time horizon $[0, T]$ with $T > T_*$.

To achieve this goal, we formulate an optimal control problem. Its value function takes an argument μ_0 in the unit simplex and returns a nonnegative number $T_*(\mu_0)$. This number is the infimum over those T such that there exists a market μ, that satisfies $\mu(0) = \mu_0$ and (2), and, moreover, is free of relative arbitrage over $[0, T]$. Thanks to an application of the fundamental theorem of asset pricing, it actually suffices to consider markets μ that are martingales until they reach the boundary of the simplex.

In the case of $d = 3$ stocks, the Hamilton–Jacobi–Bellman equation of this control problem (see (14) below) turns out to be a well studied partial differential equation arising in geometric analysis. Indeed, in this case the solution is the arrival time function of the so called mean curvature flow. Using known properties of this flow, we are able to explicitly compute the worst-case time $T_* = \max_{\mu_0} T_*(\mu_0)$, where the maximum is taken over the unit simplex in $\mathbb{R}^3$. Its value is $T_* = \sqrt{3}/(2\pi)$; see Theorem 4.6. In higher dimensions, the Hamilton–Jacobi–Bellman equation still has a geometric meaning, now as a description of what we call the minimum curvature flow. This flow is less well studied than the mean curvature flow, though it is closely related to the codimension flow of Ambrosio and Soner (1996). The partial differential equation for the arrival time of the minimum curvature flow is the subject of our companion paper Larsson and Ruf (2020).

Here is the outline of this paper. Section 2 introduces some notation and the relevant financial concepts, such as trading strategies and relative arbitrage. Section 3 provides a representation of $T_* = \max_{\mu_0} T_*(\mu_0)$ in terms of an optimization problem that maximizes the essential infimum of the exit time of martingales with sufficient volatility. Section 4 discusses the case $d = 3$, where an explicit numerical value for T_* is obtained. Next, Section 5 provides the results for the general case $d \geq 4$ by connecting the value function of the control problem to the arrival time of the so called minimum curvature flow. Section 6 concludes by posing several open problems.

We end this introduction by providing some additional pointers to the literature discussing the existence of short-term relative arbitrage in the presence of sufficient intrinsic volatility, where intrinsic volatility has several but related meanings in the literature. Fernholz (2002) observes that there exists relative arbitrage over sufficiently long time horizons, provided that the market is diverse, that is, $\max_{1 \leq i \leq d} \mu_i \leq 1 - \delta$ for some $\delta > 0$ and the instantaneous covariance matrix of the stock returns is uniformly elliptic. Under the same conditions, Fernholz et al. (2005) later prove

that there exists relative arbitrage indeed over any time horizon. It is difficult to check statistically whether the instantaneous covariance matrix has eigenvalues bounded away from zero. Thus, Fernholz and Karatzas (2005) consider instead a scaled version of (1) which is easier to verify empirically. They prove that this suffices to guarantee relative arbitrage over long time horizons. Banner and Fernholz (2008) prove that relative arbitrage over any time horizon exists if no stock defaults and the smallest stock is sufficiently volatile. Fernholz and Karatzas (2010) prove that in a uniformly elliptic Markovian model satisfying (1), relative arbitrage over any time horizon exists. Pal (2019) shows the existence of asymptotic short-term arbitrage opportunities as the number of stocks tends to infinity and an appropriate notion of sufficient volatility is assumed to hold. For some general introduction to stochastic portfolio theory and relative arbitrage we refer to Fernholz and Karatzas (2009) and Vervuurt (2015).

2 | MARKET MODELS, ARBITRAGE, AND THE SMALLEST HORIZON

Let $\Delta^d = \{x \in [0,1]^d : x_1 + \cdots + x_d = 1\}$ for $d \geq 2$ denote the unit simplex in $\mathbb{R}^d$. A *market* or a *market weight process* is a Δ^d-valued continuous semimartingale $\mu = (\mu_1, \ldots, \mu_d)^\top$ defined on some stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathsf{P})$. In this paper, (in)equalities are understood P-almost surely. As mentioned in the introduction, each component of μ represents the relative capitalization of a stock. Note that μ is allowed to take values in all of Δ^d, including the boundary.

A *trading strategy* is an $\mathbb{R}^d$-valued μ-integrable predictable process θ, and its *relative value process* is defined by $V^\theta = \theta^\top \mu$. It is called *self-financing* if

$$V^\theta = V^\theta(0) + \int_0^\cdot \theta^\top \mathrm{d}\mu.$$

There is no bank account in this model. All self-financing trading strategies are fully invested in the stock market at all times. Moreover, wealth is measured in units of the total market capitalization, or *relative to the market*. Put differently, the market portfolio is the numeraire, and the market weights μ_i are the stocks' market capitalizations expressed in this numeraire. This explains the definition of the relative value process, the self-financing condition, as well as the usage of the word *relative*.

We now define the relevant arbitrage concept for this paper. This turns out to coincide with the classical no-arbitrage condition (NA), expressed in the numeraire currently in use. The terminology is chosen to emphasize this choice of numeraire.

Definition 2.1. Given a constant $T > 0$, a trading strategy θ is called a *relative arbitrage over (time horizon)* $[0, T]$ if $V^\theta \geq 0$, $V^\theta(T) \geq V^\theta(0)$, and $\mathsf{P}(V^\theta(T) > V^\theta(0)) > 0$.

Markets that satisfy (2) admit relative arbitrage over $[0, T]$, provided $T > 1 - |\mu(0)|^2$. This can be proved using (additively) functionally generated strategies, which are constructed as follows. Let $G : \mathbb{R}^d \to \mathbb{R}$ be a concave and C^2 function. The functionally generated trading strategy associated with G is then given by

$$\theta = \nabla G(\mu) + \big(G(\mu) + \Gamma^G - \nabla G(\mu)^\top \mu \big) \mathbf{1},$$

where **1** denotes the d-dimensional column vector of ones, and

$$\Gamma^G = -\frac{1}{2} \sum_{i,j=1}^{d} \int_0^{\cdot} \partial_{ij}^2 G(\mu) \mathrm{d}[\mu_i, \mu_j] \tag{3}$$

is a nondecreasing finite-variation process. An application of Itô's formula yields that θ is indeed self-financing, with relative value process

$$V^G = G(\mu) + \Gamma^G.$$

For further details on this class of trading strategies, we refer to Karatzas and Ruf (2017), where additively functionally generated trading strategies are introduced and studied.

Consider now the function

$$Q(x) = 1 - |x|^2 = 1 - \sum_{i=1}^{d} x_i^2, \qquad x \in \Delta^d. \tag{4}$$

We then get $\Gamma^Q = \mathrm{tr}[\mu, \mu]$ for the process in (3), and the functionally generated trading strategy associated with Q has relative value process $V^Q = Q(\mu) + \Gamma^Q \geq \Gamma^Q$. Since $\Gamma^Q(t) \geq t$ for all $t \geq 0$ if μ satisfies (2), we obtain the following proposition, well known in stochastic portfolio theory.

Proposition 2.2. *Any market that satisfies (2) allows for relative arbitrage over $[0, T]$, provided $T > V^Q(0) = 1 - |\mu(0)|^2$. Moreover, this relative arbitrage can be implemented by the self-financing trading strategy*

$$\theta = -2\mu + \left(1 + |\mu|^2 + \Gamma^Q\right)\mathbf{1}.$$

The condition (2) is a statement about volatility, and is a crucial property for this paper. Giving it a name will help to make the statements below clear.

Definition 2.3. A market μ is called *sufficiently volatile* if (2) holds.

If one used another concave function G in place of Q, one would obtain another finite-variation process Γ^G. For example, (2) relies on Γ^Q, while (1) relies on the corresponding finite-variation process of the so called entropy function (see also Definition 6.1 below and the discussion afterwards).

Remark 2.4. Usually in the literature one would find (2) (or, more specifically, (1)) with the right-hand side multiplied by a strictly positive (but small) constant η. Here we assume $\eta = 1$, which amounts to scaling time by a constant.

How quickly can one obtain relative arbitrage in a sufficiently volatile market? In general, the answer depends on the market. We are interested in the worst case:

> What is the smallest time horizon T_* beyond which relative arbitrage is possible in any sufficiently volatile market?

Without the qualification "sufficiently volatile" the answer is clearly $T_* = \infty$, since a constant market weight process rules out relative arbitrage over any time horizon. Slightly more formally, we write the time T_* as follows:

$$T_* = \inf \big\{T \geq 0 : \quad \text{every sufficiently volatile market admits relative arbitrage over } [0, T]\big\}.$$

Thus for $T > T_*$, any sufficiently volatile market admits relative arbitrage over $[0, T]$. For $T < T_*$, there exists at least one sufficiently volatile market that does not admit relative arbitrage over $[0, T]$.

The purpose of this paper is to characterize T_*. If $d = 2$, it is known that $T_* = 0$. Indeed, Fernholz et al. (2018, Proposition 5.13) show that every sufficiently volatile market with two assets allows for relative arbitrage over arbitrary time horizons. In the case $d \geq 3$, we know from Fernholz et al. (2018, Theorem 6.8) that

$$T_* \geq \max_{\mu_0 \in \Delta^3} Q(\mu_0) - \frac{1}{2} = \frac{1}{6},$$

where Q is the quadratic function introduced in (4); see also Subsection 4.2 below. Moreover, it is clear from Proposition 2.2 that

$$T_* \leq \max_{\mu_0 \in \Delta^d} Q(\mu_0) = 1 - \frac{1}{d}.$$

In the following sections we will prove that $T_* = \sqrt{3}/(2\pi)$, showing that both bounds are not tight.

3 | A REPRESENTATION OF T_* IN TERMS OF MARTINGALES

In this section we use a version of the fundamental theorem of asset pricing to prove that it suffices to only consider martingales when computing T_*. To this end, for any $\mathbb{R}^d$-valued process ν, define

$$\zeta_\nu = \inf \big\{t \geq 0 : \nu(t) \notin \Delta^d\big\}.$$

Recall that the *essential infimum* of a random variable X, denoted $\mathrm{ess\,inf}\, X$, is the largest deterministic lower bound on X. That is, $\mathrm{ess\,inf}\, X = \sup\{c \in \mathbb{R} : \mathrm{P}(X \geq c) = 1\}$.

Theorem 3.1. *We have the representation*

$$T_* = \sup \left\{ \mathrm{ess\,inf}\, \zeta_\nu : \begin{array}{l} \nu \text{ is an } \mathbb{R}^d - valued \;\; continuous \;\; martingale \\ with \; \mathrm{tr}[\nu, \nu](t) \geq t \, for \, all \, t \geq 0 \end{array} \right\}. \tag{5}$$

Proof. Denote the right-hand side of (5) by T_{**}. For any $T < T_{**}$ there exists a continuous martingale ν with $\mathrm{tr}[\nu, \nu](t) \geq t$ for all $t \geq 0$ that remains in Δ^d on $[0, T]$. Using this martingale, one can easily construct a sufficiently volatile market μ that is a martingale on $[0, T]$. The martingale property implies that no relative arbitrage on $[0, T]$ can be constructed. This yields $T \leq T_*$ and we deduce that $T_{**} \leq T_*$.

Suppose now for contradiction that $T_{**} < T_*$, and choose $T \in (T_{**}, T_*)$. Then there exists a sufficiently volatile market μ, defined on a stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathrm{P})$, that does not admit

relative arbitrage over $[0, T]$. We will show below that there exists a probability measure $Q \ll P$ such that μ is a Q-martingale on $[0, T]$. Then it is easy to construct an $\mathbb{R}^d$-valued Q-martingale ν, possibly on an extension of the stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, Q)$, such that $\nu = \mu$ on $[0, T]$ and $\mathrm{tr}[\nu, \nu](t) \geq t$ for all $t \geq 0$. In particular, since μ takes values in Δ^d, we have $\zeta_\nu \geq T$. This means that $T \leq T_{**}$, a contradiction, showing that $T_{**} = T_*$ as claimed.

It remains to argue the existence of the probability measure Q. Thanks to Delbaen and Schachermayer (1995, Theorem 1.4), we only need to show that μ satisfies the no-arbitrage condition (NA) on $[0, T]$. Namely, that there exists no trading strategy θ such that $\theta^\top(0)\mu(0) \geq 0$, $\int_0^{\cdot \wedge T} \theta^\top \mathrm{d}\mu \geq -\kappa$ for some $\kappa \geq 0$, $\int_0^T \theta^\top \mathrm{d}\mu \geq 0$, and $P(\int_0^T \theta^\top \mathrm{d}\mu > 0) > 0$. Assume for contradiction that such θ exists. Define the predictable process

$$\bar{\theta} = \theta + \left(\kappa + \int_0^{\cdot} \theta^\top \mathrm{d}\mu - \theta^\top \mu \right) \mathbf{1}.$$

Since $\int_0^{\cdot} \bar{\theta}^\top \mathrm{d}\mu = \int_0^{\cdot} \theta^\top \mathrm{d}\mu$, one sees that $\bar{\theta}$ is self-financing and a relative arbitrage in the sense of Definition 2.1. This contradicts the fact that μ does not admit relative arbitrage over $[0, T]$. We deduce that μ satisfies the no-arbitrage condition (NA) on $[0, T]$. $\square$

On the right-hand side of (5), the inequality in $\mathrm{tr}[\nu, \nu](t) \geq t$ can be replaced by an equality without changing the value of the supremum. To see this, consider $\tilde{\nu}(t) = \nu(A(t))$, where A is the right-continuous inverse of $\mathrm{tr}[\nu, \nu]$. Then $\tilde{\nu}$ is an $\mathbb{R}^d$-valued continuous martingale with $\mathrm{tr}[\tilde{\nu}, \tilde{\nu}](t) = t$ for all $t \geq 0$, and $\zeta_{\tilde{\nu}} \geq \zeta_\nu$. Motivated by this, we define the function $u : \mathbb{R}^d \to [0, \infty)$ by

$$u(y) = \sup \left\{ \begin{array}{ll} \mathrm{ess\,inf}\, \zeta_\nu : & \nu \text{ is an } \mathbb{R}^d\text{-valued continuous martingale with} \\ & \nu(0) = y \text{ and } \mathrm{tr}[\nu, \nu](t) = t \text{ for all } t \geq 0 \end{array} \right\}. \tag{6}$$

Thanks to Theorem 5 we then have

$$T_* = \sup_{y \in \mathbb{R}^d} u(y) = \sup_{y \in \Delta^d} u(y).$$

4 | THE CASE $d = 3$ AND THE APPEARANCE OF MEAN CURVATURE FLOW

In this section we focus on the case $d = 3$. We proceed in several steps. First, in Subsection 4.1 we map the hyperplane containing Δ^3 to $\mathbb{R}^2$. Then, in Subsection 4.2 we recall how Fernholz et al. (2018) obtained a lower (but not sharp) bound on T_*. Motivated by their approach, in Subsection 4.3 we introduce and discuss a boundary value problem whose solution will be used to characterize T_*. Subsection 4.4 discusses the existence of a weak solution to a related stochastic differential equation. Finally, Subsection 4.5 provides a computation of T_*.

4.1 | Mapping the market μ to a two-dimensional process

A market μ with $d = 3$ assets evolves in the unit simplex Δ^3, which is a two-dimensional subset of $\mathbb{R}^3$. Using a suitable projection $U : \mathbb{R}^3 \to \mathbb{R}^2$, one can express μ in terms of a two-dimensional

FIGURE 1 The mapping from the unit simplex Δ^3 to K [Color figure can be viewed at wileyonlinelibrary.com]

$$\mu(t) = (\mu_1(t), \mu_2(t), \mu_3(t))^\top \qquad X(t) = (X_1(t), X_2(t))^\top = U\mu(t)$$
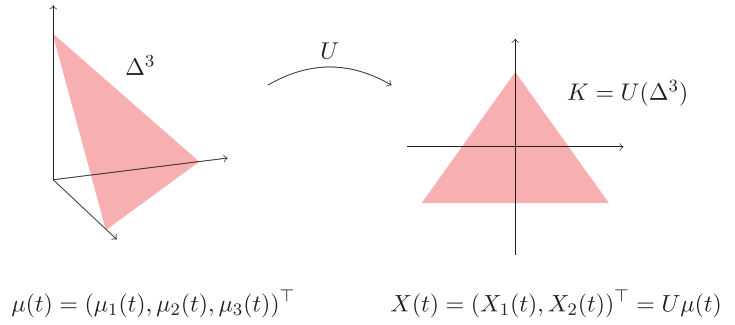
process X. This is illustrated in Figure 1. Formally, let $U \in \mathbb{R}^{2\times 3}$ be a matrix with orthonormal rows and $U\mathbf{1} = 0$. Equivalently, the rows of U form an orthonormal basis for $\mathbb{H} - \frac{1}{3}\mathbf{1}$, where

$$\mathbb{H} = \left\{ x \in \mathbb{R}^3 : x^\top \mathbf{1} = 0 \right\} + \frac{1}{3}\mathbf{1}$$

is the two-dimensional plane containing Δ^3. For example, we may choose

$$U = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} \end{pmatrix}.$$

The map $y \mapsto Uy$ from $\mathbb{H}$ to $\mathbb{R}^2$ is a bijection with inverse $x \mapsto U^\top x + \frac{1}{3}\mathbf{1}$. Since the rows of U are orthonormal, this map is an isometry. Thus ν is an $\mathbb{H}$-valued martingale if and only if $X = U\nu$ is an $\mathbb{R}^2$-valued martingale, and in this case $U^\top[X,X]U = [\nu,\nu]$. In particular, the quadratic variations satisfy $\operatorname{tr}[\nu,\nu] = \operatorname{tr}[X,X]$. The simplex Δ^3 is mapped to the compact domain

$$K = U(\Delta^3) \subset \mathbb{R}^2.$$

An advantage of this transformation is that K has an interior (in $\mathbb{R}^2$), while Δ^3 does not (in $\mathbb{R}^3$).

For any $\mathbb{R}^2$-valued semimartingale X we define

$$\tau_X = \inf \{t \geq 0 : X(t) \notin K\}. \tag{7}$$

We also define the value function $v : \mathbb{R}^2 \to [0, \infty)$ by

$$v(x) = \sup \left\{ \operatorname{ess\,inf} \tau_X : \begin{array}{l} X \text{ is an } \mathbb{R}^2 - \text{valued continuous martingale with} \\ X(0) = x \text{ and } \operatorname{tr}[X,X](t) = t \text{ for all } t \geq 0 \end{array} \right\}. \tag{8}$$

The isometry property of the map $y \mapsto Uy$ implies that the function v in (8) is related to u in (6) by the identity
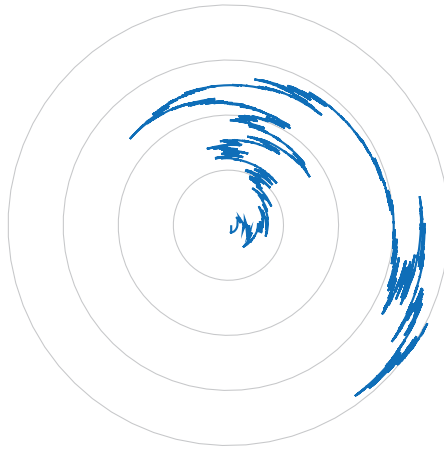
$$u(y) = v(Uy), \qquad y \in \Delta^3.$$

FIGURE 2 A sample path of the solution of (9) [Color figure can be viewed at wileyonlinelibrary.com]

As a consequence, we also get

$$T_* = \sup_{x\in\mathbb{R}^2} \upsilon(x) = \sup_{x\in K} \upsilon(x).$$

4.2 | The lower bound of Fernholz et al. (2018)

The above representation of T_* indicates that we should look for $\mathbb{R}^2$-valued martingales that do not slow down (in the sense that the trace of the quadratic variation remains equal to t), yet remain in K for a deterministic amount of time. To make headway, let us recall how Fernholz et al. (2018) derive the lower bound $T_* \geq 1/6$, mentioned at the end of Section 2. They consider the stochastic differential equation

$$\mathrm{d}X(t) = \mathrm{d}\begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} = \frac{1}{\sqrt{X_1^2(t) + X_2^2(t)}} \begin{pmatrix} X_2(t) \\ -X_1(t) \end{pmatrix} \mathrm{d}W(t), \qquad t \geq 0, \tag{9}$$

where W denotes a one-dimensional Brownian motion. It can be shown that this stochastic differential equation always admits a weak solution X satisfying $\mathrm{tr}[X, X](t) = t$ for all $t \geq 0$, even when $X(0) = 0$; see Theorem 4.3 and Remark 4.4 below. An application of Itô's formula yields

$$|X(t)|^2 = X_1^2(t) + X_2^2(t) = t, \qquad t \geq 0. \tag{10}$$

Thus at any time t, $X(t)$ lies on a centered circle with radius $\sqrt{t}$; see Figure 2 for a simulated sample path of X. It is clear that with τ_X as in (7) we have ess inf $\tau_X = r^2$, where r is the radius of the largest circle contained in K. Basic Euclidean geometry yields $r = 1/\sqrt{6}$, and hence ess inf $\tau_X = 1/6$. This gives the bound $T_* \geq 1/6$.

Before proceeding, we observe that (9) can be written

$$dX(t) = \frac{\nabla^\perp \overline{Q}(X(t))}{|\nabla \overline{Q}(X(t))|} dW(t), \qquad t \geq 0, \tag{11}$$

where $\overline{Q} : \mathbb{R}^2 \ni x \mapsto 1 - x_1^2 - x_2^2$, and $\nabla^\perp \overline{Q} = \begin{pmatrix} -\partial_2 \overline{Q} \\ \partial_1 \overline{Q} \end{pmatrix}$ is the so-called skew-gradient of $\overline{Q}$.

4.3 | The mean curvature equation

The concentric circles in Figure 2 are poorly adapted to the triangular geometry of K. For this reason $1/6$ is not a sharp lower bound for T_*. The goal is now to replace Q in (11) by some other function that better reflects the geometry of K. To illustrate the idea, consider a function $w \in C^2$ and assume, for the moment, that the stochastic differential equation

$$dX(t) = \frac{\nabla^\perp w(X(t))}{|\nabla w(X(t))|} dW(t), \qquad t \geq 0, \tag{12}$$

has a weak solution starting from any point in K. Using Itô's formula we then obtain

$$t + w(X(t)) = w(X(0)) + \int_0^t \left(1 + \frac{\nabla^\perp w^\top \nabla^2 w \nabla^\perp w}{2|\nabla w|^2}(X(s)) \right) ds, \qquad t \geq 0. \tag{13}$$

Let us additionally assume that the function w can be chosen so that the integrand is zero and $w|_{\partial K} = 0$. That is, we assume w is a solution of the boundary value problem

$$\begin{cases} 1 + \dfrac{\nabla^\perp w^\top \nabla^2 w \nabla^\perp w}{2|\nabla w|^2} = 0 & \text{in } K^\circ \\ \\ \qquad\qquad w = 0 & \text{on } \partial K. \end{cases} \tag{14}$$

With all these assumptions in place we set $t = \tau_X$ in (13) to get $\tau_X = w(X(0))$. This yields $v(X(0)) \geq$ ess inf $\tau_X = w(X(0))$, where v is the value function defined in (8). Hence we get the lower bound $T_* \geq \sup_{x \in K} w(x)$.

To turn these ideas into a rigorous argument, two main issues need to be resolved: solving the boundary value problem (14) and finding a solution of the stochastic differential equation (12). We deal with the latter in the next subsection. Here, we focus on (14).

It turns out that (14) describes the arrival time of the so-called *mean curvature flow of ∂K*, and we now discuss the physical phenomenon that this represents. The mean curvature (or curve shortening) flow is a construction that gradually deforms a given initial contour, in our case ∂K. Each point on the contour moves in the normal direction at a speed equal to the curvature at that point (in our case, half the curvature). Figure 3 illustrates this. The *arrival time function w* maps each point x to the time $w(x)$ it takes for the evolving contour to reach x. If the initial contour is convex (i.e., encloses a convex region), then $w(x)$ is well-defined and finite for every point in the enclosed region. The contour at a positive time $t > 0$ is then $\{x \in K^\circ : w(x) = t\}$. The contour eventually shrinks to a point and disappears. This happens at a finite extinction time $t^* < \infty$.
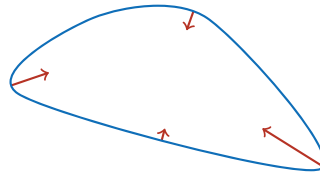
FIGURE 3   An illustration of the mean curvature flow in $\mathbb{R}^2$. The parts of the contour with high curvature move inwards fast, indicated by long arrows. Flat pieces move slowly, indicated by short arrows. The flow continues until the extinction time t^* [Color figure can be viewed at wileyonlinelibrary.com]
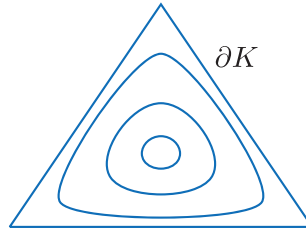
FIGURE 4   Mean curvature flow with initial contour ∂K. The flow deforms ∂K to resemble small circles (see Gage, 1984), before disappearing at the center of the triangle K [Color figure can be viewed at wileyonlinelibrary.com]

In our case, the initial contour ∂K is not smooth, and some care is needed to define the flow. The contour is however convex, and the mean curvature flow can be shown to exist and have a well-defined arrival time function. Its level curves are drawn schematically in Figure 4. The arrival time function w corresponding to ∂K turns out to be the solution of (14). This becomes clearer when the equation by simple algebra is written in the equivalent form

$$\frac{1}{|\nabla w|} = -\frac{1}{2} \operatorname{div}\left(\frac{\nabla w}{|\nabla w|} \right), \tag{15}$$

where $\operatorname{div}(f) = \partial_1 f_1 + \partial_2 f_2$ denotes the divergence of a function $f = (f_1, f_2) : \mathbb{R}^2 \to \mathbb{R}^2$. Evaluated at a point $\bar{x}$ on a contour $\{x \in K : w(x) = t\}$, the right hand side of (15) gives half the curvature of the contour at that point. The size of the gradient, $|\nabla w(\bar{x})|$, is expressed in units of time per distance, and gives the change in arrival time per unit displacement orthogonally to the level curve at $\bar{x}$. Thus $1/|\nabla w(\bar{x})|$, the left-hand side of (15), is the speed by which the contour moves inwards.

We are interested in classical solutions of (14). Some care is needed, because the left-hand side of (14) is undefined at critical points of w (i.e., at points where the gradient vanishes). With this in mind, we call a function $w \in C(K) \cap C^2(K^\circ)$ a *classical solution* of (14) if it satisfies the equation at every point $x \in K$ where $\nabla w(x) \neq 0$, and if it satisfies the boundary condition $w|_{\partial K} = 0$. Here one immediately runs into problems of uniqueness: under the above definition, the zero solution $w = 0$ is a classical solution, simply because *every* point in K° is a critical point. Therefore, we will not have uniqueness among *all* classical solutions, but only among those that have a finite number of critical points. Note that this definition of classical solution is stated slightly differently from the one in Colding and Minicozzi (2016, Section 4).

Theorem 4.1. *The boundary value problem (14) has a unique classical solution w with a single critical point $\bar{x} \in K^\circ$. The solution w is strictly positive, continuous on K, and C^3 in K°, and the Hessian at the critical point is $\nabla^2 w(\bar{x}) = -2I$.*

Proof. Evans and Spruck (1991, Section 5) (see also Chen et al., 1991) establish the existence of a so-called *generalized mean curvature flow* $(\Gamma_t)_{t \geq 0}$, where $\Gamma_0 = \partial K$ is the boundary of the convex domain K and each Γ_t is a closed curve. Furthermore, by Evans and Spruck (1992, Theorem 5.5), there exists some $t^* > 0$ such that Γ_t is in fact a smooth convex curve for all $t \in (0, t^*)$ and empty for all $t > t^*$. Evans and Spruck (1991, Section 6) then assert that $(\Gamma_t)_{0 < t < t^*}$ evolves by mean curvature in the classical differential geometric sense up to the extinction time t^* (see e.g., Gage & Hamilton, 1986, in particular Theorem 1.1). Thanks to the remark on page 75 in Gage and Hamilton (1986), Γ_t is also strictly convex, that is, does not contain any line segments, for all $t \in (0, t^*)$.

We next argue that

$$\bigcup_{t>0} \Gamma_t = K^\circ. \tag{16}$$

First of all, we have $\bigcup_{t>0} \Gamma_t \subset K^\circ$ since $(\Gamma_t)_{0<t<t^*}$ evolves by mean curvature in the classical differential geometric sense. Assume for contradiction that the other set inclusion does not hold, and let $\overline{\Gamma}_0$ denote the boundary of $\bigcup_{t>0} \Gamma_t$. Let $\bar{t}^*$ denote the extinction time of the generalized mean curvature flow $(\overline{\Gamma}_t)_{t \geq 0}$ with initial contour $\overline{\Gamma}_0$. Thanks to the comparison principle of Evans and Spruck (1991, Theorem 7.2) we have $\bar{t}^* = t^*$. Let now $\delta > 0$ denote the strictly positive distance between Γ_0 and $\overline{\Gamma}_0$ and choose $\varepsilon \in (0, t^*)$ small enough so that the diameter of $\Gamma_{t^*-\varepsilon}$ is strictly less than δ. Then the distance between $\Gamma_{t^*-\varepsilon}$ and $\overline{\Gamma}_{t^*-\varepsilon}$ is strictly less than δ, leading to a contradiction with Evans and Spruck (1991, Theorem 7.3). This proves (16).

Let $\overline{w} : K \to [0, t^*]$ denote the arrival time function: for each $x \in K$, $\overline{w}(x) = t$ if $x \in \Gamma_t$. This is well defined and positive on $\bigcup_{t>0} \Gamma_t = K^\circ$ thanks to the arguments in Evans and Spruck (1991, Section 7), and zero on ∂K. We claim that $\overline{w}$ is a C^3 classical solution of the partial differential equation

$$1 + \frac{\nabla^\perp \overline{w}^\top \nabla^2 \overline{w} \nabla^\perp \overline{w}}{|\nabla \overline{w}|^2} = 0 \tag{17}$$

in K° with a single critical point, at which the Hessian is $-I$. To show this, fix any $t_0 \in (0, t^*)$, and let $\widetilde{w}(x) = \overline{w}(x) - t_0$ be the arrival time of the time-shifted flow $\widetilde{\Gamma}_t = \Gamma_{t_0+t}$. The initial contour of this flow is the smooth and strictly convex curve Γ_{t_0}. By Kohn and Serfaty (2006, Lemma 3.1) (see also Huisken, 1993 for an earlier C^2 regularity result), $\widetilde{w}$ is a C^3 classical solution of (17) in the domain enclosed by Γ_{t_0}, has a single critical point $\bar{x}$ in this domain, and satisfies $\nabla^2 \widetilde{w}(\bar{x}) = -I$. Since t_0 was arbitrary, the properties carry over to $\overline{w}$ in the whole domain K°. It is clear that the critical point $\bar{x}$ does not depend on t_0.

Next, we argue that $\overline{w}$ is continuous up to the boundary. Indeed, let $(x_n)_{n \in \mathbb{N}} \subset K^\circ$ converge to some $x \in \partial K$ and suppose for contradiction that $t_n = \overline{w}(x_n)$ converges to a strictly positive limit. Then $\bar{t} = \inf_{n \in \mathbb{N}} t_n > 0$, and $x_n \in \bigcup_{t \geq \bar{t}} \Gamma_t$ for all $n \in \mathbb{N}$. However, then $(x_n)_{n \in \mathbb{N}}$ remains bounded away from ∂K, a contradiction.

Finally, defining $w = 2\overline{w}$, we obtain a classical solution of (14) with the required properties.

It remains to argue uniqueness of a classical solution for (14) with one critical point. Thanks to the comparison principle for viscosity solutions in Kohn and Serfaty (2006, Theorem 4) it suffices to argue that each classical solution for (14) with one critical point is also a viscosity solution in the sense of Kohn and Serfaty (2006, Definition 3). We will not give the details of this rather standard argument, but only indicate some key points. Viscosity solutions for (14) can be defined using the symbol

$$F(p, M) = -\frac{M_{11}p_2^2 - 2p_1p_2M_{12} + M_{22}p_2^2}{2|p|^2}, \qquad p \in \mathbb{R}^2 \setminus \{0\}, M \in \mathbb{S}^2,$$

where $\mathbb{S}^2$ denotes the set of 2×2 symmetric matrices. If u is a classical solution with a single critical point $\bar{x}$, then (14) states that $F(\nabla u(x), \nabla^2 u(x)) = 1$ for all $x \in K \setminus \{\bar{x}\}$. Letting F^* and F_* denote the upper and lower semicontinuous envelopes of F, it follows that $F_*(\nabla u(\bar{x}), \nabla^2 u(\bar{x})) \leq 1 \leq F^*(\nabla u(\bar{x}), \nabla^2 u(\bar{x}))$. Using (degenerate) ellipticity of F, it is now straightforward to verify that u is a viscosity solution of (14), as required. This completes the proof of the theorem. □

Remark 4.2. Although we have a triangle shaped domain K in mind, Theorem 4.1 holds true for any compact convex subset of $\mathbb{R}^2$. If K is indeed the triangle, then by symmetry the critical point $\bar{x}$ must be the center point of the triangle. Specifically, with K as in Subsection 4.1, we have $\bar{x} = 0$.

Let us mention that the mean curvature flow is more commonly studied via a parabolic equation satisfied by $v(t, x) = w(x) - t$; see for instance Evans and Spruck (1991) and Chen et al. (1991). The elliptic equation in (14) for the arrival time is however more natural in our context. This equation was first studied by Evans and Spruck (1991, Subsection 7.3), and subsequently by a number of other authors. In particular, Kohn and Serfaty (2006) discuss a deterministic game interpretation of (14) that is related to the stochastic representation that we use here. Other stochastic representations of the mean curvature flow have been obtained by Buckdahn et al. (2001) and Soner and Touzi (2002a, 2002b, 2003). These closely related methods have been very useful in solving a variety of stochastic control problems.

4.4 | Solving the stochastic differential equation in (12)

We assumed above that the stochastic differential equation in (12) allows for a weak solution whenever $X(0) \in K^\circ$. We now argue that that such a weak solution exists, at least up to a deterministic time.

Theorem 4.3. *Let w be the classical solution of the boundary value problem (14) from Theorem 4.1 and fix $x_0 \in K$. Then there exists an $\mathbb{R}^2$-valued martingale X with $X(0) = x_0$ and $\mathrm{tr}[X, X](t) = t$ for all $t \geq 0$ that satisfies the stochastic differential equation in (12) on $[0, \tau_X]$. Furthermore, $w(X(t)) = w(x_0) - t$ for all $t \in [0, w(x_0)]$, and $\tau_X = w(x_0)$*

Proof. The proof is similar to the proof of Larsson and Ruf (2020, Lemma 3.2). If $x_0 \in \partial K$ we let X be an $\mathbb{R}^2$-valued Brownian motion starting in x_0, scaled by $1/\sqrt{2}$ to ensure that $\mathrm{tr}[X, X](t) = t$ for all $t \geq 0$.

Suppose that $x_0 \in K^\circ \setminus \{\bar{x}\}$, where $\bar{x}$ is the unique critical point and global maximum of w. Since the stochastic differential equation has locally Lipschitz coefficients on the set

$$\{x \in K^\circ : \nabla w(x) \neq 0\} = K^\circ \setminus \{\bar{x}\},$$

there is a local solution X of (12) on $[0, \zeta)$, where

$$\zeta = \lim_{n \to \infty} \inf \left\{ t \geq 0 : \text{dist}(X(t), \partial K \cup \{\bar{x}\}) \leq \frac{1}{n} \right\}.$$

Since X is a bounded martingale on $[0, \zeta)$, the martingale convergence theorem implies that $X(\zeta) = \lim_{t \to \zeta} X(t)$ exists. After time ζ, we let X continue like a scaled $\mathbb{R}^2$-valued Brownian motion. The process X is now an $\mathbb{R}^2$-valued martingale with $X(0) = x_0$ and $\text{tr}[X, X](t) = t$ for all $t \geq 0$ that satisfies the stochastic differential equation in (12) on $[0, \zeta]$.

We now check the remaining properties. Since $\nabla^\perp w^\top \nabla w = 0$ on K°, Itô's formula yields

$$w(X(t)) = w(x_0) + \int_0^t \frac{\nabla^\perp w(X(s))^\top \nabla^2 w(X(s)) \nabla^\perp w(X(s))}{2|\nabla w(X(s))|^2} ds, \quad t < \zeta.$$

Since w satisfies (14) on $K^\circ \setminus \{\bar{x}\}$ and is continuous on K, we get

$$w(X(t)) = w(x_0) - t, \qquad t \leq \zeta. \tag{18}$$

Thus $w(X(\zeta)) \leq w(x_0) < w(\bar{x})$, hence ζ is the first hitting time of the boundary ∂K by X. Since X follows a Brownian motion after ζ, we actually have $\zeta = \tau_X$. Evaluating (18) at $t = \zeta = \tau_X$ yields $0 = w(x_0) - \tau_X$. This proves the theorem for $x_0 \neq \bar{x}$.

Suppose now that $x_0 = \bar{x}$, and select a sequence $(x_n)_{n \in \mathbb{N}} \subset K$ with $\lim_{n \to \infty} x_n = \bar{x}$. For each $n \in \mathbb{N}$, let X^n be the $\mathbb{R}^2$-valued martingale constructed above with $X^n(0) = x_n$. Fix $s \geq 0$ and $n \in \mathbb{N}$ and define

$$M(t) = |X^n(t) - X^n(s)|^2 - t + s, \qquad t \geq s.$$

Then M is a local martingale, hence a supermartingale, on $[s, \infty)$ with $[M, M] \leq 4 \int_s^\cdot |X(u) - X(s)|^2 du$. Thus for all $t \geq s$,

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

so that

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

Thus by Kolmogorov's continuity criterion (see Revuz & Yor, 1999, Theorem I.2.1 and its proof) we get, for any fixed $\alpha \in (0, \frac{1}{4})$,

$$E\left[\left(\sup_{0\le s<t\le w(\bar{x})}\frac{|X^n(t)-X^n(s)|}{|t-s|^\alpha}\right)^4\right]\le c$$

for some constant c that does not depend on $n \in \mathbb{N}$.

Since Hölder balls are relatively compact in $C([0, w(\bar{x})], \mathbb{R}^2)$ by the Arzelà–Ascoli theorem, it follows that the sequence $(X^n|_{[0,w(\bar{x})]})_{n\in\mathbb{N}}$ is tight in $C([0, w(\bar{x})], \mathbb{R}^2)$. Therefore, by Prokhorov's theorem there exists a continuous process Y on $[0, w(\bar{x})]$ such that a subsequence of $(X^n|_{[0,w(\bar{x})]})_{n\in\mathbb{N}}$ converges weakly to Y in $C([0, w(\bar{x})], \mathbb{R}^2)$. Since $X^n(t) \in K$ and $w(X^n(t)) = w(x_n) - t$ for all $t \in [0, w(x_n)]$ and all $n \in \mathbb{N}$, the continuous mapping theorem and the continuity of w give $Y(t) \in K$ and

$$w(Y(t)) = w(\bar{x}) - t \tag{19}$$

for all $t \in [0, w(\bar{x})]$. We now define the process X to be equal to Y on $[0, w(\bar{x})]$ and then continue like a scaled $\mathbb{R}^2$-valued Brownian motion. This yields directly $w(\bar{x}) = \tau_X$.

It now suffices to show that X on $[0, w(\bar{x})]$ is a weak solution to (12). By Karatzas and Shreve (1991, Proposition 5.4.6) it is enough to argue that the process

$$M^f(t) = f(X(t)) - \frac{1}{2}\int_0^t \frac{\nabla^\perp w(X(s))^\top \nabla^2 f(X(s))\nabla^\perp w(X(s))}{|\nabla w(X(s))|^2}ds, \qquad t \in [0, w(\bar{x})],$$

is a martingale for every $f \in C^\infty(\mathbb{R}^2)$. The continuous mapping theorem implies that M^f is a martingale whenever f vanishes in a neighborhood of $\bar{x}$. Now fix $s \in (0, w(\bar{x})]$ and a general $f \in C^\infty(\mathbb{R}^2)$. Let $g \in C^\infty(\mathbb{R}^2)$ coincide with f outside the set $\{x \in K : w(x) \ge w(\bar{x}) - s\}$ and vanish on $\{x \in K : w(x) \ge w(\bar{x}) - s/2\}$, which is a neighborhood of $\bar{x}$. In particular, M^g is a martingale. Moreover, (19) implies that X does not visit the set $\{x \in K : w(x) > w(\bar{x}) - s\}$ on $[s, w(\bar{x})]$. As a result, $M^f(t) - M^f(s) = M^g(t) - M^g(s)$ for all $t \in [s, w(\bar{x})]$, so that the process $(M^f(t))_{t\in[s,w(\bar{x})]}$ is a martingale. Since $s \in (0, w(\bar{x})]$ was arbitrary, this shows that $(M^f(t))_{t\in(0,w(\bar{x})]}$ is a martingale. The bounded convergence theorem finally implies that $(M^f(t))_{t\in[0,w(\bar{x})]}$ is a martingale as well. This completes the proof. $\square$

Remark 4.4. We can again repeat the observation of Remark 4.2, now concerning Theorem 4.3. This result can be applied for any closed convex subset of $\mathbb{R}^2$, provided that we use the appropriate function w. In particular, assume, for the moment only, that K denotes the unit disk in $\mathbb{R}^2$. Then it is easy to check that $w(x) = 1 - |x|^2$ for all $x \in K$ solves the corresponding boundary value problem in (12). The corresponding stochastic differential equation is (9), for which Theorem 4.3 now yields the existence of a weak solution.

4.5 | Worst-case time horizon

Let us summarize where we stand. In the case $d = 3$, we have used the solution of the arrival time formulation of mean curvature flow to obtain an $\mathbb{R}^2$-valued martingale X with $\text{tr}[X, X](t) = t$ for all $t \ge 0$, whose first exit time from K is deterministic (see Theorem 4.3). This yields the bound

$v(x) \geq w(x)$ for all $x \in K$. We will now argue that this bound is optimal, which also shows that X is optimal for the optimization problem on the right hand side of (8).

Theorem 4.5. *Let w be the classical solution of the boundary value problem (14) from Theorem 4.1. Then $w = v$ on K with v given in (8); that is, one has the stochastic representation*

$$w(x) = \sup_{X} \operatorname{ess\,inf} \tau_X, \qquad x \in K,$$

where the supremum extends over all $\mathbb{R}^2$-valued continuous martingales X with $X(0) = x$ and $\operatorname{tr}[X, X](t) = t$ for all $t \geq 0$, defined on arbitrary stochastic bases.

Proof. We use a verification argument. Theorem 4.3 exhibits, for every $x_0 \in K$, a valid martingale X with $X_0 = x_0$ and $\operatorname{ess\,inf} \tau_X = w(x_0)$. We now need to argue that this is the best one can do, so that this martingale is in fact optimal; that is, it remains to show that $\operatorname{ess\,inf} \tau_X \leq w(x_0)$ for any valid martingale X with $X_0 = x_0$.

To this end, fix $x_0 \in K$ and let X be an $\mathbb{R}^2$-valued continuous martingale with $X(0) = x_0$ and $\operatorname{tr}[X, X](t) = t$ for all $t \geq 0$, defined on a stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P)$. We need to prove that $\operatorname{ess\,inf} \tau_X \leq w(x_0)$. Since the quadratic variation actually has absolutely continuous trajectories almost surely, we have $[X, X] = \int_0^{\cdot} a(t) dt$ for some predictable $\mathbb{S}_+^2$-valued process a with $\operatorname{tr}(a) = 1$. Here $\mathbb{S}_+^2$ denotes the set of 2×2 symmetric positive semidefinite matrices.

Fix now $\lambda > 1$ and consider the function $w_\lambda : \lambda K \to [0, \infty)$, $x \mapsto \lambda^2 w(x/\lambda)$. One easily checks that w_λ satisfies the boundary value problem in (14) with K replaced by λK. Thanks to the continuity of w and since $\lambda > 1$ was chosen arbitrarily it now suffices to argue that $\operatorname{ess\,inf} \tau_X \leq w_\lambda(x_0)$. We shall use repeatedly that $\sup_{x \in K} |\nabla w_\lambda(x)| < \infty$ (which follows from the fact that ∇w_λ is continuous on λK).

Let now $k > 0$ be a constant to be determined later, and define the process

$$\widetilde{X} = X + k \int_0^{\cdot \wedge \tau_X} a(s) \nabla w_\lambda(X(s)) ds.$$

An application of Itô's formula gives

$$t \leq t + w_\lambda(X(t)) = w_\lambda(x_0) + \int_0^t \nabla w_\lambda(X)^\top d\widetilde{X} + \int_0^t \left(1 + \frac{1}{2} \operatorname{tr} \left(a(s) \nabla^2 w_\lambda(X(s))\right)\right.$$

$$\left. - k \nabla w_\lambda(X(s))^\top a(s) \nabla w_\lambda(X(s)) \right) ds \tag{20}$$

for all $t \leq \tau_X$.

We now fix $\varepsilon \in (0, 1)$ and claim that there exists $k > 0$ such that the integrand in the ds-integral on the right hand side of (20) is bounded from above by ε. To this end it suffices to show the existence of $k > 0$ such that

$$1 + \frac{1}{2} \operatorname{tr} \left(A \nabla^2 w_\lambda(x)\right) \leq \varepsilon + k \nabla w_\lambda(x)^\top A \nabla w_\lambda(x) \tag{21}$$

for all $(x, A) \in K \times S_+^2$ with $\text{tr}(A) = 1$. Let us argue by contradiction and assume that no such k exists. Then, since $\nabla^2 w_\lambda$ is bounded on K, there exists a sequence $(x_n, A_n)_{n \in \mathbb{N}} \subset K \times \mathbb{S}_+^2$ such that $\text{tr}(A_n) = 1$, $\lim_{n \to \infty} \nabla w_\lambda(x_n)^\top A_n \nabla w_\lambda(x_n) = 0$, and $1 + (1/2) \text{tr}(A_n \nabla^2 w_\lambda(x_n)) > \varepsilon$. After passing to a subsequence, we have $\lim_{n \to \infty} x_n = \widetilde{x}$ and $\lim_{n \to \infty} A_n = A$ for some $(\widetilde{x}, A) \in K \times \mathbb{S}_+^2$ with

$$\text{tr}(A) = 1, \qquad \nabla w_\lambda(\widetilde{x})^\top A \nabla w_\lambda(\widetilde{x}) = 0, \qquad \text{and} \qquad 1 + \frac{1}{2} \text{tr}(A \nabla^2 w_\lambda(\widetilde{x})) \geq \varepsilon. \qquad (22)$$

Let us first argue that $\nabla w_\lambda(\widetilde{x}) \neq 0$. If $\nabla w_\lambda(\widetilde{x}) = 0$ then $\nabla^2 w_\lambda(\widetilde{x}) = -2I$ by Theorem 4.1 (recall also Remark 4.2), contradicting the first and last equality in (22). Hence we have $|\nabla w_\lambda(\widetilde{x})| \neq 0$. This forces $\text{rank}(A) = 1$ and then $A = \nabla^\perp w_\lambda(\widetilde{x}) \nabla^\perp w_\lambda(\widetilde{x})^\top / |\nabla w_\lambda(\widetilde{x})|^2$. Thus, the last equality of (22) becomes

$$1 + \frac{\nabla^\perp w_\lambda(\widetilde{x})^\top \nabla^2 w_\lambda(\widetilde{x}) \nabla^\perp w_\lambda(\widetilde{x})}{|\nabla w_\lambda(\widetilde{x})|^2} \geq \varepsilon.$$

This contradicts the fact that w_λ solves the boundary value problem in (14) (with K replaced by K_λ).

We now argue that $(1 - \varepsilon) \text{ess inf } \tau_X \leq w_\lambda(x_0)$. Since $\varepsilon > 0$ was arbitrary, this will complete the proof. With k as determined in the previous paragraph, (20) and (21) yield

$$(1 - \varepsilon)t - w_\lambda(x_0) \leq \int_0^t \nabla w_\lambda(X)^\top d\widetilde{X}, \qquad t \leq \tau_X.$$

Consider now the strictly positive stochastic exponential Z, given by the stochastic differential equation

$$Z = 1 - k \int_0^{\cdot \wedge \tau_X} Z \nabla w_\lambda(X)^\top dX.$$

Since $\text{ess inf } \tau_X < \infty$ and ∇w_λ is bounded on K, Novikov's condition shows that Z is a martingale on $[0, \text{ess inf } \tau_X]$. Let Q be the probability measure on $\mathcal{F}_{\text{ess inf } \tau_X}$ induced by $Z_{\text{ess inf } \tau_X}$. Under Q, $\int_0^\cdot \nabla w_\lambda(X)^\top d\widetilde{X}$ is a local martingale on $[0, \text{ess inf } \tau_X]$ bounded from below, hence a supermartingale. Consequently,

$$(1 - \varepsilon) \text{ess inf } \tau_X - w_\lambda(x_0) \leq \mathsf{E}_Q \left[\int_0^{\text{ess inf } \tau_X} \nabla w_\lambda(X)^\top d\widetilde{X} \right] \leq 0,$$

yielding the claim. This completes the proof. $\qquad \square$

We have now argued that $v(x) = w(x)$ for all $x \in K$, where v is given in (8) and w denotes the arrival time function of the mean curvature flow provided by Theorem 4.1. Thanks to the observations in Subsection 4.1 we thus have $T_* = \sup_{x \in K} w(x)$. While an explicit expression for w is not available, its maximal value can be computed using Gage and Hamilton (1986, Lemma 3.1.7). It is observed there that the area A enclosed by a smooth simple closed curve that flows by mean

curvature satisfies

$$\frac{\mathrm{d}A(t)}{\mathrm{d}t} = -2\pi.$$

The extinction time of the flow is therefore $A(0)/2\pi$, which is the time it takes until the area becomes zero. In our case, the initial contour is not smooth. However, the arguments in the proof of Theorem 4.1 show that it immediately becomes smooth under the mean curvature flow, and it follows that the formula for the extinction time is still valid. Since w is twice the arrival time of the mean curvature flow we obtain

$$w(0) = \max_{x \in K} w(x) = 2\frac{\mathrm{area}(K)}{2\pi} = \frac{\mathrm{area}(\Delta^3)}{\pi} = \frac{\sqrt{3}}{2\pi} \approx 0.28,$$

where the first equality holds due to Remark 4.2, and the two areas coincide due to the isometry property of $U : \Delta^d \to K$ (see Subsection 4.1). We arrive at the following result.

Theorem 4.6. *If $d = 3$, then the smallest time horizon beyond which any sufficiently volatile market admits relative arbitrage is*

$$T_* = \frac{\sqrt{3}}{2\pi} \approx 0.28.$$

Note in particular that the true value of T_* lies strictly between the previously best known lower and upper bounds $1/6$ and $2/3$; see the end of Section 2.

5 | THE GENERAL CASE $d \geq 3$ AND MINIMUM CURVATURE FLOW

We now turn to the general case $d \geq 3$. The representation of T_* in Theorem 3.1 is still valid, and we approach it via the value function u in (6). Again, u can be characterized as the solution of a partial differential equation, but one that is more complicated than in the case $d = 3$. There are two essential differences.

First, it turns out that u no longer vanishes on the (relative) boundary of Δ^d. To see why, imagine a market μ that hits the interior of a $(d-1)$-dimensional boundary face of Δ^d. If $d \geq 4$, the construction in Subsection 4.2 shows that an optimal market will not immediately exit Δ^d, but instead spend some deterministic amount of time on this boundary face. This affects how we deal with boundary conditions.

Second, the equation itself no longer describes the arrival time of the mean curvature flow. Instead, it corresponds to another flow that we call the *minimum curvature flow*, which is more degenerate than the mean curvature flow. This flow is closely related to the *codimension-$(d-1)$ mean curvature flow* of Ambrosio and Soner (1996).

As in Subsection 4.1, we use an affine isometry $U : \mathbb{R}^d \to \mathbb{R}^{d-1}$ to identify Δ^d with a polytope $K \subset \mathbb{R}^{d-1}$ with nonempty interior $K^\circ \neq \emptyset$. Then $u(y) = v(Uy)$ for all $y \in \Delta^d$, where

$$v(x) = \sup \left\{ \mathrm{ess\,inf}\, \tau_X : \begin{array}{l} X \text{ is an } \mathbb{R}^{d-1} - \text{ valued continuous martingale with} \\ X(0) = x \text{ and } \mathrm{tr}[X, X](t) = t \text{ for all } t \geq 0 \end{array} \right\}; \quad (23)$$

here τ_X denotes the first exit time of X from K. Let us now describe heuristically how to obtain the partial differential equation satisfied by v.

Consider an $\mathbb{R}^{d-1}$-valued continuous martingale X with $\mathrm{tr}[X,X](t) = t$ for all $t \geq 0$. Its quadratic variation can be written

$$[X,X] = \int_0^{\cdot} a(s)\mathrm{d}s$$

for some $\mathbb{S}_+^{d-1}$-valued process $(a(t))_{t\geq 0}$ with $\mathrm{tr}\, a(t) = 1$ for all $t \geq 0$. For the sake of discussion, suppose v is C^2. Suppose also that v is zero on ∂K, so that $v(X(\tau_X)) = 0$. (In reality this only holds for strictly convex domains. Nevertheless, it still produces correct heuristics, basically because $v(X(\tau_X)) = 0$ still holds for the optimal choice of X.) Itô's formula then yields

$$\tau_X = v(X(0)) + \int_0^{\tau_X} \nabla v(X)^\top \mathrm{d}X + \int_0^{\tau_X} \left(1 + \frac{1}{2} \mathrm{tr}\left(a(t)\nabla^2 v(X(t))\right) \right) \mathrm{d}t.$$

We look for choices of X that lead to deterministic lower bounds on τ_X. To do so, we focus on those X for which the stochastic integral above vanishes, that is, $\nabla v(X(t))^\top \mathrm{d}X(t) = 0$ for all $t \in [0, \tau_X]$. That is, we require that $a\nabla v(X) = 0$. We then obtain

$$\tau_X = v(X(0)) + \int_0^{\tau_X} \left(1 + \frac{1}{2} \mathrm{tr}\left(a(t)\nabla^2 v(X(t))\right) \right) \mathrm{d}t$$

$$\leq v(X(0)) + \int_0^{\tau_X} \left(1 - F\left(\nabla v(X(t)), \nabla^2 v(X(t))\right) \right) \mathrm{d}t,$$

where for any $p \in \mathbb{R}^{d-1}$ and $M \in \mathbb{S}^{d-1}$ we define

$$F(p,M) = \inf \left\{ -\frac{1}{2}\mathrm{tr}(aM) : a \in \mathbb{S}_+^{d-1},\ \mathrm{tr}(a) = 1,\ ap = 0 \right\}. \tag{24}$$

This suggests that v should satisfy the partial differential equation

$$F(\nabla u, \nabla^2 u) = 1. \tag{25}$$

Indeed, in this case we obtain ess $\inf \tau_X \leq \tau_X \leq v(X(0))$, and if additionally $a(t)$ achieves the infimum in (24) with $p = \nabla v(X(t))$ and $M = \nabla^2 v(X(t))$, we obtain $\tau_X = v(X(0))$. This heuristic reasoning makes the following theorem plausible. The rigorous proof is more involved, and can be found in the companion paper Larsson and Ruf (2020).

Theorem 5.1. *For $d \geq 3$, the value function v in (23) is the unique continuous (on K) viscosity solution of (25) in $K°$ with zero boundary condition (in the viscosity sense). Moreover, v is quasi-concave, vanishes on the vertices and boundary lines of K, and is strictly positive elsewhere in K.*

Proof. Existence and uniqueness in the class of upper semicontinuous viscosity solutions follows from Larsson and Ruf (2020, Theorem 1.2). Continuity on K follows from Larsson and Ruf (2020,

Theorem 1.4 and Corollary 5.9). The remaining statements follow from Larsson and Ruf (2020, Theorem 1.3). $\qquad\square$

We will not delve into the technical aspects of this theorem here, nor its connection to the minimum curvature flow; details are available in Larsson and Ruf (2020). For completeness we nonetheless give the relevant definition of viscosity solution. A bounded function $u : K \to \mathbb{R}$ is called a viscosity subsolution of (25) in K° if

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

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

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

The function u is said to be a viscosity supersolution in K° with zero boundary condition if the above properties hold with u^*, F_*, max, $\leq$ replaced by u_*, F^*, min, $\geq$. It is a viscosity solution in K° with zero boundary condition if it is both a viscosity sub- and supersolution in K° with zero boundary condition.

Remark 5.2. Consider the case $d = 3$. For any $p \in \mathbb{R}^2 \setminus \{0\}$, if $a \in \mathbb{S}_+^2$ satisfies $ap = 0$, then rank$(a) \leq 1$ and we must have $a = qq^\top$ for some $q \in \mathbb{R}^2$ orthogonal to p. If in addition tr$(a) = 1$ then $|q| = 1$. Thus, in view of (24), provided $\nabla u \neq 0$, we obtain

$$F(\nabla u, \nabla^2 u) = -\frac{\nabla^\perp u^\top \nabla^2 u \nabla^\perp u}{2|\nabla u|^2},$$

and (25) reduces to the partial differential equation in (14).

Remark 5.3. Recall that $K = U(\Delta^d)$ for an affine isometry $\mathbb{R}^d \to \mathbb{R}^{d-1}$. Let us assume that U maps $(1/d)\mathbf{1}$ to the origin, so that K is $(d-1)$-dimensional polytope centered at the origin. It is worth noting that the function $Q(x) = \frac{d-1}{d} - |x|^2$ is nonnegative on K and is the solution of the partial differential equation

$$\inf\left\{-\frac{1}{2}\text{tr}(a\nabla^2 u) : a \in \mathbb{S}_+^{d-1}, \text{ tr}(a) = 1\right\} = 1$$

in K° with zero boundary condition (in the viscosity sense). This equation is similar but not identical to (25). It can be shown that it is the dynamic programming equation corresponding to the control problem

$$\sup\left\{\mathsf{E}[\tau_X] : \begin{array}{l} X \text{ is an } \mathbb{R}^{d-1} - \text{valued continuous martingale with} \\ X(0) = x \text{ and tr}[X, X](t) = t \text{ for all } t \geq 0 \end{array}\right\},$$

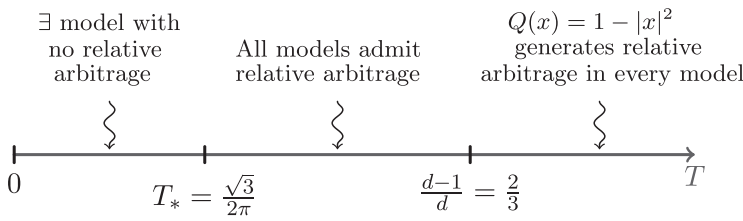where the expected exit time is maximized, rather than the essential infimum.

FIGURE 5 The various regimes for existence of relative arbitrage in the case $d = 3$

6 | CONCLUSION AND OPEN PROBLEMS

Let us summarize our findings. For sufficiently volatile markets with $d = 3$ assets, we have shown that relative arbitrage always exists beyond the critical time horizon $T_* = \frac{\sqrt{3}}{2\pi}$, but not always before T_*. The value of T_* is determined by analyzing the arrival time function of mean curvature flow in the plane. In general for $d \geq 3$, the critical time horizon T_* is determined through a partial differential equation connected to minimum curvature flow, which happens to coincide with the mean curvature flow for $d = 3$.

We conclude with four open problems.

- We have seen that in any sufficiently volatile market with $d = 3$ assets, the portfolio generating function $Q(x) = 1 - |x|^2$ generates relative arbitrage over $[0, T]$ for any $T > \frac{2}{3}$. However, relative arbitrage exists already for $T \in (\frac{\sqrt{3}}{2\pi}, \frac{2}{3}]$ (see Figure 5). What does the arbitrage strategy look like for these intermediate time horizons? Is there a portfolio generating function that works in every sufficiently volatile market? More generally, does there exist a single path-dependent strategy $\theta = (\theta(t, (\mu(s))_{s \leq t}))_{t \in [0,T]}$ that generates arbitrage over the time horizon $[0, T]$ with $T \in (\frac{\sqrt{3}}{2\pi}, \frac{2}{3}]$, in any sufficiently volatile market? Or is the arbitrage strategy inherently model dependent? We remark that the arrival time function w of the mean curvature flow will not yield such a portfolio generating function. This is because it fails to be concave near the vertices of the triangle.
- We have not provided any numerical values for T_* in dimension $d > 3$. Can one use numerics to produce quantitative bounds? In particular, lower (upper) bounds can be obtained by searching for explicit subsolutions (supersolutions) of (25).
- In Definition 2.3 we defined a sufficiently volatile market in terms of $\text{tr}[\mu, \mu]$. This is neither the only possible definition, nor the most natural. It can be generalized as follows.

Definition 6.1. Given a portfolio generating function G, a market μ is called G-*sufficiently volatile* if Γ^G in (3) is well defined and satisfies $\Gamma^G(t) \geq t$ for all $t \geq 0$.

- The condition (2) arises by taking $G(x) = Q(x) = 1 - |x|^2$. But there are other choices. For instance, one could take the *entropy function* $G_1(x) = -\sum_{i=1}^{d} x_i \log x_i$, or the *geometric average* $G_2(x) = \prod_{i=1}^{d} x_i^{1/d}$. For a given choice of G, one can then ask: what is the smallest time horizon beyond which relative arbitrage is possible in any G-sufficiently volatile market? If $G = G_2$, then $G|_{\partial \Delta^d} = 0$ and the answer seems to be that the critical time horizon is $\max_{y \in \Delta^d} G_2(y) = 1/d$; see also Fernholz et al. (2018, Remark 6.18). If $G = G_1$, then G-sufficiently volatile markets are

those that satisfy (1). It remains an open problem how to determine the critical time horizon for relative arbitrage across G_1–sufficiently volatile markets (or other choices of G other than Q and G_2). We conjecture that the arguments in this paper can be adapted to answer this open problem, and might lead to new geometric flows. In this context, let us also mention Pal and Wong (2018), who consider the geometries corresponding to different portfolio generating functions.

- Expanding on the previous question, one might ask how to use additional statistical knowledge about the market to further bound the smallest time beyond which relative arbitrage is possible in any such market. For example, assume $d = 3$ and restrict focus to sufficiently volatile markets that are in addition diverse, that is, $\max_{i=1,2,3} \mu_i \leq 1 - \delta$ for some $\delta \in (0, 1/2)$. What is the minimum time horizon now? Clearly, it is smaller than $\sqrt{3}/(2\pi)$. For this setup, when $d = 3$ and the markets are diverse, the answer is easy thanks to Remark 4.2 and the observations before Theorem 4.6. One only needs to compute the area of the region $\widetilde{K} = \{x \in \Delta^3 : \max_i x_i \leq 1 - \delta\}$, which is $\sqrt{3}(1 - 3\delta^2)/2$. This yields the minimum time

$$\frac{1}{\pi}\text{area}\big(\widetilde{K}\big) = \frac{\sqrt{3}}{2\pi}\big(1 - 3\delta^2\big).$$

ORCID
Johannes Ruf https://orcid.org/0000-0003-3616-2194

REFERENCES
Ambrosio, L., & Soner, H. M. (1996). Level set approach to mean curvature flow in arbitrary codimension. *Journal of Differential Geometry*, *43*(4), 693–737. ISSN 0022-040X. http://projecteuclid.org/euclid.jdg/1214458529.

Banner, A., & Fernholz, D. (2008). Short-term relative arbitrage in volatility-stabilized markets. *Annals of Finance*, *4*(4), 445–454.

Buckdahn, R., Cardaliaguet, P., & Quincampoix, M. (2001). A representation formula for the mean curvature motion. *Siam Journal on Mathematical Analysis*, *33*(4), 827–846. ISSN 0036-1410. https://doi.org/10.1137/S0036141000380334.

Chen, Y. G., Giga, Y., & Goto, S. (1991). Uniqueness and existence of viscosity solutions of generalized mean curvature flow equations. *Journal of Differential Geometry*, *33*(3), 749–786. ISSN 0022-040X. http://projecteuclid.org/euclid.jdg/1214446564.

Colding, T. H., & Minicozzi, II, W. P. (2016). Differentiability of the arrival time. *Communications on Pure and Applied Mathematics*, *69*(12), 2349–2363. ISSN 0010-3640. https://doi.org/10.1002/cpa.21635.

Delbaen, F., & Schachermayer, W. (1995). The existence of absolutely continuous local martingale measures. *The Annals of Applied Probability*, *5*(4), 926–945. ISSN 1050-5164. http://links.jstor.org/sici?sici=1050-5164(199511)5:4%3C926:TEOACL%3E2.0.CO;2-5&origin=MSN.

Evans, L. C., & Spruck, J. (1991). Motion of level sets by mean curvature. I. *Journal of Differential Geometry*, *33*(3), 635–681. ISSN 0022-040X. http://projecteuclid.org/euclid.jdg/1214446559.

Evans, L. C., & Spruck, J. (1992). Motion of level sets by mean curvature. III. *The Journal of Geometric Analysis*, *2*(2), 121–150. ISSN 1050-6926. https://doi.org/10.1007/BF02921385.

Fernholz, E. R. (2002). *Stochastic Portfolio Theory*, volume 48 of *Applications of Mathematics (New York)*. New York: Springer-Verlag. ISBN 0-387-95405-8. https://doi.org/10.1007/978-1-4757-3699-1. Stochastic Modelling and Applied Probability.

Fernholz, R., & Karatzas, I. (2005). Relative arbitrage in volatility-stabilized markets. *Annals of Finance*, *1*(2), 149–177.

Fernholz, D., & Karatzas, I. (2010). On optimal arbitrage. *The Annals of Applied Probability*, *20*(4), 1179–1204. ISSN 1050-5164. https://doi.org/10.1214/09-AAP642.

Fernholz, E. R., Karatzas, I., & Ruf, J. (2018). Volatility and arbitrage. *The Annals of Applied Probability*, *28*(1), 378–417. ISSN 1050-5164. https://doi.org/10.1214/17-AAP1308.

Fernholz, R., & Karatzas, I. (2009). Stochastic Portfolio Theory: an overview. In A. Bensoussan (Ed.), *Handbook of Numerical Analysis*, volume Mathematical Modeling and Numerical Methods in Finance. Elsevier.

Fernholz, R., Karatzas, I., & Kardaras, C. (2005). Diversity and relative arbitrage in equity markets. *Finance and Stochastics*, *9*(1), 1–27.

Gage, M. E. (1984). Curve shortening makes convex curves circular. *Inventiones Mathmaticae*, *76*(2), 357–364. ISSN 0020-9910. https://doi.org/10.1007/BF01388602.

Gage, M., & Hamilton, R. S. (1986). The heat equation shrinking convex plane curves. *Journal of Differential Geometry*, *23*(1), 69–96. ISSN 0022-040X. http://projecteuclid.org/euclid.jdg/1214439902.

Huisken, G. (1993). Local and global behaviour of hypersurfaces moving by mean curvature. In *Differential geometry: Partial differential equations on manifolds (Los Angeles, CA, 1990)*, volume 54 of *Proc. Sympos. Pure Math.* (pp. 175–191). Providence, RI: Amer. Math. Soc. https://doi.org/10.1090/pspum/054.1/1216584.

Karatzas, I., & Ruf, J. (2017). Trading strategies generated by Lyapunov functions. *Finance and Stochastics*, *21*(3), 753–787.

Karatzas, I., & Shreve, S. E. (1991). *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics* (2nd ed.). New York: Springer-Verlag. ISBN 0-387-97655-8. https://doi.org/10.1007/978-1-4612-0949-2.

Kohn, R. V., & Serfaty, S. (2006). A deterministic-control-based approach to motion by curvature. *Communications on Pure and Applied Mathematics*, *59*(3), 344–407. ISSN 0010-3640. https://doi.org/10.1002/cpa.20101.

Larsson, M., & Ruf, J. (2020). Minimum, curvature flow and martingale exit times. Preprint, arXiv:2003.13611.

Pal, S. (2019). Exponentially concave functions and high dimensional stochastic portfolio theory. *Stochastic Processes and Their Applications*, *129*(9), 3116–3128. ISSN 0304-4149. https://doi.org/10.1016/j.spa.2018.09.004.

Pal, S., & Wong, T.-K. L. (2018). Exponentially concave functions and a new information geometry. *The Annals of Probability*, *46*(2), 1070–1113. ISSN 0091-1798. https://doi.org/10.1214/17-AOP1201.

Revuz, D., & Yor, M. (1999). *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* (3rd ed.). Berlin: Springer-Verlag. ISBN 3-540-64325-7. https://doi.org/10.1007/978-3-662-06400-9.

Soner, H. M., & Touzi, N. (2002a). Dynamic programming for stochastic target problems and geometric flows. *Journal of the European Mathematical Society*, *4*(3), 201–236. ISSN 1435-9855. https://doi.org/10.1007/s100970100039.

Soner, H. M., & Touzi, N. (2002b). A stochastic representation for the level set equations. *Communications in Partial Differential Equations*, *27*(9-10), 2031–2053. ISSN 0360-5302. https://doi.org/10.1081/PDE-120016135.

Soner, H. M., & Touzi, N. (2003). A stochastic representation for mean curvature type geometric flows. *The Annals of Probability*, *31*(3), 1145–1165. ISSN 0091-1798. https://doi.org/10.1214/aop/1055425773.

Vervuurt, A. (2015). Topics in stochastic portfolio theory. Preprint, arXiv::1504.02988.