A PDE Approach to the Prediction of a Binary Sequence with Advice from Two History-Dependent Experts

NADEJDA DRENSKA
University of Minnesota
AND
ROBERT V. KOHN
Courant Institute, NYU

Abstract

The prediction of a binary sequence is a classic example of online machine learning. We like to call it the "stock prediction problem," viewing the sequence as the price history of a stock that goes up or down one unit at each time step. In this problem, an investor has access to the predictions of two or more "experts," and strives to minimize her final-time regret with respect to the best-performing expert. Probability plays no role; rather, the market is assumed to be adversarial. We consider the case when there are *two history-dependent* experts, whose predictions are determined by the d most recent stock moves. Focusing on an appropriate continuum limit and using methods from optimal control, graph theory, and partial differential equations, we discuss strategies for the investor and the adversarial market, and we determine associated upper and lower bounds for the investor's final-time regret. When $d \leq 4$ our upper and lower bounds coalesce, so the proposed strategies are asymptotically optimal. Compared to other recent applications of partial differential equations to prediction, ours has a new element: there are two timescales, since the recent history changes at every step whereas regret accumulates more slowly. © 2000 Wiley Periodicals, Inc.

1 Introduction

Prediction with expert advice is an area of online machine learning [8], in which an agent has access to several experts' predictions and uses them to make a prediction of her own. Probabilistic modeling is not involved; instead, the agent's goal is to "minimize regret" – i.e. to minimize her worst-case shortfall with respect to the (retrospectively) best-performing expert.

Much of the vast literature in this area explores easily-implemented prediction strategies, assessing their performance on broad classes of prediction problems. Our focus is different: we examine a special example – the prediction of a binary sequence using guidance from two history-dependent experts (described informally below, and more formally in Section 2) – looking for prediction strategies that take into account the character of the example. Our work is interesting because:

Communications on Pure and Applied Mathematics, Vol. 000, 0001–0055 (2000)

(a) we show how considering a scaled version of the problem facilitates the use of PDE methods;
(b) we show how the study of history-dependent experts leads naturally to the use of graph-theoretic methods; and
(c) we identify *asymptotically optimal* strategies when the experts use only the most recent $d \leq 4$ stock moves.

(We only consider two experts, and when $d \geq 5$ we do not show that our strategies are optimal. Subsequent to the work presented here, the first author and J. Calder have obtained further results. In [4] they identify asymptotically optimal strategies for any value of d and any number of history-dependent experts, using methods rather different from those of the present paper. Then in [5] they confirm the conjectures in Section 4 of the present paper (extended to any number of experts), and they provide an alternative (simpler) analysis of the limit $\varepsilon \to 0$. For additional information on these developments, see the end of Section 2.6.)

The idea that scaling should facilitate the use of PDE methods is not unprecedented. Our prediction problem can be viewed as a two-person game, and the PDE describes the value of the game in an appropriate scaling limit. Other two-person games leading similarly to nonlinear parabolic equations have been studied, for example, in [2, 3, 18, 19, 20, 21, 22]. Much of this literature uses viscosity-solution techniques, which are required when the solution of the relevant PDE is not smooth. In this paper we don't need viscosity solutions, since our PDE has a smooth solution; as a result, our arguments are rather elementary, using what a control theorist would call "verification arguments."

While the standard name for our problem is the "prediction of a binary sequence using expert advice," we like to call it the "stock prediction problem," viewing the sequence as the price history of a stock that goes up or down one unit at each time step. The case with two *static* experts (one who always predicts the stock will go up, the other who always predicts the stock will go down) has a long history, going back at least to Thomas Cover's 1965 paper [9]. A concise treatment can be found in Section 8.4 of [8]. The optimal strategies can be determined using dynamic programming, and the predictor's worst-case regret has an explicit formula, involving the expected value of a function of many Bernoulli random variables. When the number of steps is large, the explicit formula can be evaluated using the central limit theorem. An alternative PDE-based perspective on this continuum limit was developed by Kangping Zhu in [25]; roughly speaking, it takes the continuum limit of the dynamic programming principle rather than the limit of the explicit solution formula.

In this paper we consider two *history-dependent* experts, whose predictions depend on the most recent $d \geq 1$ stock movements. Our overall approach is similar to that of [25] – which is only natural, since when $d = 0$ the experts are static rather than history-dependent. When $d \geq 1$, however, the problem is fundamentally different from those considered by Cover [9] and Zhu [25]. Briefly: regret accumulates slowly, while the recent history changes quickly. In more detail: if the stock price moves by ± 1 at each time step then the worst-case regret should be of order $\sqrt{N}$ after N time steps, by arguments related to Blackwell's approachability theorem (see e.g. Chapter 2 of [8]). So the regret increases about $1/\sqrt{N}$ per time step, which is small (since we are interested in the asymptotic behavior as $N \to \infty$). The recent history, by contrast, is different at each time step. Thus for history-dependent experts the problem has two well-separated time scales. In our analysis, the fast timescale is handled using graph-theoretic methods, while the slow timescale is handled using methods analogous to those of [25].

Here is an informal description of our version of the stock prediction problem. Consider a stock whose price goes up or down one unit at each time step. An investor can invest (at each timestep) in $-1 \leq f \leq 1$ units of stock. There are two experts, whose investment choices are determined by two functions (call them q and r) of the most recent d stock moves. The clock stops after N time steps, and the investor's goal is to minimize her worst-case final-time shortfall (i.e. "regret") with respect to the (retrospectively) best-performing expert. Since the focus is on worst-case behavior, this is a two-person game, whose players are the investor and an adversarial market (which chooses the price history so as to maximize the investor's final-time regret). To be clear: the final time N and the experts' investment rules q and r are public information. At the kth timestep, the investor announces her investment f_k, then the market chooses whether the stock goes up ($b_k = 1$) or down ($b_k = -1$); these determine the investor's gain $b_k f_k$ and also the experts' gains $b_k q(m_k)$ and $b_k r(m_k)$; here m_k represents k days of history prior to timestep k. (For a more complete description see Section 2.1.)

The investor's worst-case final-time regret can in principle be determined by dynamic programming, using as state variables the investor's regret with respect to each expert and the past d stock price moves (see Section 2.3). But one can do better by taking advantage of the fact that regret accumulates slowly while the recent history changes at each time step. Indeed, as time proceeds the sequence of recent histories determines a walk on a certain directed graph (see Section 2.2). Simplifying somewhat, let us suppose that the market's choices involve well-selected cycles on this graph (a key task in our analysis of the investor's strategy will be to justify this simplification). Then the rapidly-changing recent history can be accounted for by considering, for each cycle, how regret accumulates if the market chooses that cycle. This gives an estimate of the worst-case final-time regret that's independent of the recent history – a function only of the investor's regret with respect to each expert and number of timesteps that remain. (The case when only $d = 1$ timestep of history is used is especially transparent; it is discussed in detail in Section 3.)

This game is discrete (in much the way that a random walk on a lattice is discrete). To connect the game with a PDE we consider a scaling limit (entirely analogous to the way that scaled random walks lead to Brownian motion, whose backward Kolmogorov equation is the linear heat equation). To explain heuristically, suppose $v(j, y_1, y_2)$ is the worst-case regret at the final time N, given that the investor's regret with respect to experts 1 and 2 are y_1 and y_2 at timestep j (and ignoring for now the importance of the recent history). As already noted above, we expect v to be of order $\sqrt{N}$ when $(y_1, y_2) = (0,0)$ and $j = 0$, by arguments related to Blackwell's theorem. Therefore it is natural to consider the rescaled regret $u = v/\sqrt{N}$ as a function of $x = y/\sqrt{N}$ and $t = j/N$. Writing $\varepsilon = 1/\sqrt{N}$, the rescaled dynamic programming principle for u involves spatial steps of order ε and timesteps of order ε^2. The relevant PDE for $u(x,t)$ is, roughly speaking, the one for which this rescaled dynamic programming principle is a convergent numerical scheme in the limit $\varepsilon \to 0$.

The preceding discussion of scaling neglects the importance of the recent stock price history. In fact, we identify a limiting PDE for the investor's worst-case regret only when the experts use up to $d = 4$ days of history. This is because we achieve a full understanding of how the stock price history determines the players' optimal choices only when $d \leq 4$. But we also have results for $d > 4$, involving upper (respectively lower) bounds for the worst-case regret, obtained by solving PDEs associated with specific strategies for the investor (respectively the market).

The PDE's that emerge are 2nd-order nonlinear parabolic equations, solved backward in time. Such PDE are commonly seen in stochastic control; here the second-order character comes not from stochasticity, but rather from the game's Hannan consistency (a situation analogous to that of [18]). Based on our account of the game, the final-time value should be $\varphi(x_1, x_2) = \max\{x_1, x_2\}$, since the predictor's goal was to minimize her worst-case regret with respect to the best-performing expert, i.e. $\max\{y_1, y_2\}$. (Since this function is homogeneous of degree one, the passage from regret to scaled regret does not change its form.) It is, however, both natural and convenient to consider the scaled dynamic programming principle with a more general final-time value $\varphi(x)$. Indeed, our methods rely on the smoothness of u, so they require φ to be sufficiently smooth. We adapt them to the classic case $\varphi(x) = \max\{x_1, x_2\}$ by approximating it (above or below) by a smooth function.

What are the PDE's? Actually, there is essentially just *one* PDE. It is convenient to change variables to $\xi = x_1 - x_2$ and $\eta = x_1 + x_2$. The PDE is then

$$u_t + Cu_\eta^{-2}\left(u_{\xi\xi}u_\eta^2 - 2u_{\xi\eta}u_\xi u_\eta + u_{\eta\eta}u_\xi^2\right) = 0 \quad \text{for } t < T \tag{1.1}$$

with the final-time condition $u(T, \xi, \eta) = \varphi(\xi, \eta)$. The constant C must of course be chosen properly; it reflects the rate at which regret accumulates (see Section 2.3 for an account of how such a PDE emerges from the dynamic programming principle). The PDE (1.1) looks very nonlinear, but it simplifies dramatically when the final-time condition has the form $\varphi(\xi, \eta) = c\eta + \overline{\varphi}(\xi)$. (Note that the classic choice $\max(x_1, x_2)$ has this form, since it equals $\frac{1}{2}(\eta + |\xi|)$.) Then it is natural to expect that $u(t, \xi, \eta) = c\eta + \overline{u}(t, \xi)$, and substitution of this ansatz into the PDE gives the linear heat equation $\overline{u}_t + C\overline{u}_{\xi\xi} = 0$ with the final-time condition $\overline{u} = \overline{\varphi}$ at $t = T$. Surprisingly, something similar can be done for a much broader class of final-time conditions $\varphi(\xi, \eta)$. Indeed, assuming certain structural conditions on φ, for each y the level set $u(t,\xi,\eta) = y$ has the form $\eta = h(t,\xi;y)$ where h solves a linear heat equation in ξ and t, with appropriate final-time data. This is explained in [25]; we offer a self-contained treatment as an Appendix.

How is the investor's strategy linked to the PDE? For two static experts (i.e. when the two experts make no use of the recent stock price history) the investor's optimal strategy resembles what is known in the literature as a potential-based strategy, using the solution of the PDE as a (time-dependent) potential [25]. In our history dependent setting, the situation is quite different: the investor's optimal strategies depend on the recent stock price history as well as on the solution of the PDE. Our analysis suggests that if the recent history is m, the investor should buy $f_m = f_m^* + \varepsilon f_m^\#$ units of stock. The leading-order term f_m^* is state-dependent, but it is nevertheless obtained by using the solution of the PDE as something like a potential; the correction $\varepsilon f_m^\#$ is obtained by an entirely different argument, involving the graph that represents the evolution of stock price histories (see Sections 2.4, 3.1, and 4.1 for further information in this direction).

What about the market's strategies? Briefly (and informally), knowing that the investor should choose a strategy of the form $f_m = f_m^* + \varepsilon f_m^\#$, the market identifies a particular choice of $f_m^\#$ that works to its advantage, and "forces" the investor to use this strategy by penalizing other choices (see Sections 2.4, 3.2 and 4.2 for further information in this direction).

To put our work in context with respect to the online machine learning literature, we note that there are numerous articles on the "prediction of binary sequences," of which [6, 7] are representative. In much of this literature, the focus is on identifying easily-implemented strategies for the predictor (involving "potential functions" for example) that work relatively well (assuring regret of order at most $\sqrt{N}$ after N timesteps). Our focus is different: we would like to understand an asymptotically *optimal* strategy for the predictor, and the optimal prefactor of $\sqrt{N}$ in the estimate of the worst-case regret. This requires identifying asymptotically optimal strategies for the market as well as for the predictor. Such a complete understanding has to date been achieved for only a few prediction problems. In the present paper we achieve such clarity when the history-dependence involves at most $d = 4$ recent stock movements. For $d > 4$, as noted above, we obtain upper and lower bounds for the worst-case regret, but we don't know whether they match; this question reduces, as we explain in Section 4, to one about the cycles in certain de Bruijn graphs.

The online machine learning literature is not restricted to the prediction of sequences. A different class of examples focuses directly on the experts' outcomes, letting those be determined directly (rather than through a time series) by the market. (A PDE-based discussion of one such problem can be found in [11], and a PDE perspective on potential-based strategies can be found in [16, 17, 23]; PDE methods have also been applied to another class of problems known as "drifting games" [13].) In the present setting the experts' outcomes are highly constrained, since (i) their progress must be consistent with a time series, and (ii) their predictions depend, at a given time, on the past d items in that time series. The graph-theoretic methods we use to deal with the fast time scale are essentially a means for understanding the effect of these constraints. Perhaps related methods might be useful for some other prediction problems with similar constraints on the experts' advice.

In a generic problem involving prediction of sequences, the goal might be to actually guess the next element of the sequence. In such a setting the "loss function" would measure the error; a typical example would be the "absolute value loss" $|f - b_k|$. In the stock prediction problem the situation is different – the investor and the experts focus instead on how much money they gain or lose. This seems at first glance very different from the absolute value loss; however in the context of the stock prediction problem (where each stock move b_k is limited to ± 1, and the focus is exclusively on cumulative regret) the use of absolute value loss rather than financial loss would not really change matters. (This fact is well-known and elementary; for completeness we nevertheless review it in Section 2.1.)

We briefly highlight a few recent papers with connections to our work. The stock prediction problem with two static experts, a discounted payoff, and no final time was studied in [1]. The focus there (like here) is on identification of the investor's and the market's optimal behavior. While the paper's methods seem somewhat different from ours, differential equations do play a fundamental role. A related though somewhat different discussion can be found in Section 1.5.2 of [25], where the problem is considered for a discounted payoff and a fixed final time.

The stock prediction problem with two static experts is a special case of a more general class of problems studied in [15]. The focus there is rather different from ours: rather than considering just the predictor's regret (i.e. her shortfall with respect to the best-performing expert), the paper [15] discusses algorithms that limit the worst-case loss as well as the worst-case regret. The analysis there permits any number of experts.

The rest of this paper is organized as follows: Section 2 sets the stage, by presenting our problem in full detail and linking it heuristically to our PDE; it closes, in Section 2.6, with a summary of our main results. Section 3 proves our upper and lower bounds in the special case when the experts use only $d = 1$ days of history; this case is treated separately because it is relatively transparent, while still capturing the main ideas used for general d. Section 4 introduces the graph-based methods used to determine the investor's and market's strategies. A good strategy for the investor leads to an upper bound for the worst-case final-time regret, while a good strategy for the market leads to a lower bound for the same quantity; our proofs of these bounds occupies Section 5. Our analysis relies on the PDE (1.1) having a sufficiently smooth solution (with uniformly bounded derivatives) when the final-time function φ satisfies some structural properties. The required estimates follow easily from basic facts about the linear heat equation when the final-time function has the form $\varphi(\xi,\eta) = c\eta + \overline{\varphi}(\xi)$. They are, however, valid for a broader class of final-time functions φ; this is proved in [25] but we provide a self-contained treatment here as Appendix A, for the convenience of the reader.

Our analysis requires a minor and apparently technical restriction on the experts (see (2.2)); Appendix B explains what changes if it is relaxed.

2 Getting started

This section describes our problem in detail and provides a heuristic discussion of our approach. It closes, in Section 2.6, with a summary of our main results.

2.1 The game

The stock prediction problem is a zero-sum, two-person, sequential game played by the investor and the market. We represent the stock movements by a binary data stream $b_1, b_2, \ldots$, each with value -1 or $+1$.

Since our experts make their predictions using d days of history, we need some notation for the most recent d stock movements (the current "state"). While the obvious representation is $(b_{k-d}, \ldots, b_{k-1})$ at time k, it is convenient to use binary notation instead, recording -1 as 0 and $+1$ as 1. With this convention, the state at time k is

$$m_k = \frac{1}{2}(b_{k-d}+1, \ldots, b_{k-1}+1) \in \{0,1\}^d,$$

which can be viewed as an integer ranging from 0 to $2^d - 1$.

At each time step the investor can buy $|f| \le 1$ units of stock. Her choice of f can be viewed as a prediction, since if she buys f_k shares at time k and the stock then moves by b_k, she has a profit of $b_k f_k$ (or a loss, if this is negative). Our two history-dependent experts' predictions are expressed similarly: the experts are characterized by two functions $q(m)$ and $r(m)$, defined on $\{0,1\}^d$, which represent their investment choices. The experts should of course not be identical,

$$q(m) \neq r(m) \text{ for at least one state } m, \tag{2.1}$$

and their guidance should be admissible: $|q(m)| \le 1$ and $|r(m)| \le 1$ for all m. For technical reasons, we need the preceding inequalities to be strict:

$$\text{we assume that } |q(m)| < 1 \text{ and } |r(m)| < 1 \text{ for all states } m. \tag{2.2}$$

(This condition assures that the investor's leading-order bid f_m^* satisfies $|f_m^*| < 1$, as we'll see in Section 2.5. It is natural to wonder whether the condition is really necessary. The answer is yes, as we explain in Appendix B.)

Following the convention of the prediction literature, we focus on the investor's and the experts' losses rather than their gains. In terms of the *financial loss* function

$$l(f, b_k) = -f b_k \tag{2.3}$$

the investor's and experts' losses due to stock movement k are evidently $l(f_k, b_k)$, $l(q_k, b_k)$, and $l(r_k, b_k)$, if the investor's prediction is f, the recent history is m_k, and the experts' bids are $q_k = q(m_k)$ and $r_k = r(m_k)$. So the investor's *regret* (or shortfall, or relative loss) with respect to expert q at step k is

$$l(f, b_k) - l(q_k, b_k) = -f b_k + q_k b_k = b_k(q_k - f), \tag{2.4}$$

and her regret with respect to expert r at step k is similarly

$$l(f,b_k) - l(r_k,b_k) = -fb_k + r_k b_k = b_k(r_k - f). \tag{2.5}$$

The investor's cumulative regret is a key quantity, since it captures her overall performance compared to the experts. We record the cumulative regret with respect to experts q and r by (x_1,x_2). To be explicit: if the investor's prediction at step k is f_k then after the Mth step,

$$x_1 = \sum_{k=1}^{M} l(f_k,b_k) - l(q_k,b_k) \quad \text{and} \quad x_2 = \sum_{k=1}^{M} l(f_k,b_k) - l(r_k,b_k). \tag{2.6}$$

The language of investment makes the financial loss function $l(f,b) = -fb$ seem very natural. But (as noted in the Introduction) the situation would be no different if we used the absolute value loss function $l'(f,b) = |f-b|$. Indeed, since b takes only the values ± 1, the associated regret with respect to expert q at time k would be

$$l'(f,b_k) - l'(q_k,b_k) = |f-b_k| - |q_k - b_k| = b_k(q_k - f)$$

and similarly the associated regret with respect to expert r would be $b_k(r_k - f)$. Comparing with (2.4)–(2.5), we see that when only the regret (rather than the total loss) is being considered, the financial loss and absolute value loss functions are equivalent.

When the game starts there is no history; we suppose the two experts have some rule for handling this (for example $q_k = r_k = 0$ if $1 \le k \le d$). The specific choice of this rule will have no effect on our analysis, since we are interested in what happens over many time steps.

2.2 The underlying graph

Since the state $m \in \{0,1\}^d$ records the recent history, it changes from one step to the next. For example, if $d = 2$ and the most recent moves were $b_{k-2} = -1$ and $b_{k-1} = 1$ then $m_k = (0,1)$ and the next state m_{k+1} can be either $(1,1)$ (if $b_k = 1$) or $(1,0)$ (if $b_k = -1$). Introducing some notation: if the current state is m, we define

$$\begin{aligned} m_+ &= \text{subsequent state if } b_k = 1, \text{ and} \\ m_- &= \text{subsequent state if } b_k = -1. \end{aligned}$$

Thus when

$$m = \frac{1}{2}(b_{k-d}+1,\ldots,b_{k-1}+1) \in \{0,1\}^d$$

we have

$$\begin{aligned} m_+ &= \frac{1}{2}(b_{k-d+1}+1,...,b_{k-1}+1,2), \text{ and} \\ m_- &= \frac{1}{2}(b_{k-d+1}+1,...,b_{k-1}+1,0). \end{aligned}$$

As the game proceeds, the sequence of states can be visualized using an appropriate directed graph. It has the states as vertices, and each vertex m is connected to

m_+ and m_-; the case $d = 2$ is shown in Figure 2.1. Evidently the graph has 2^d vertices; each vertex has exactly two outgoing edges; and each vertex has exactly two incoming edges. It is, in fact, a well-studied graph, known as the d-dimensional de Bruijn graph on 2 symbols (see e.g. page 61 of [24]).

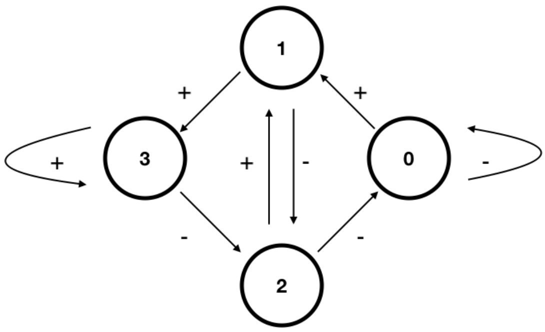

FIGURE 2.1. the underlying graph, when $d = 2$.

Our analysis uses the well-known fact that a closed walk can be decomposed as a union of simple cycles. To keep our treatment self-contained, we review the relevant definitions and give the proof of this result.

Definition 2.1. A closed walk is a list of vertices and edges $a_0, e_1, a_1, \ldots, e_k, a_k$ such that the initial and final vertices are the same ($a_k = a_0$), and for $1 \le i \le k$ the edge e_i has endpoints a_{i-1} and a_i.

Definition 2.2. A simple cycle is a walk without repeated vertices apart from the first one and the last one, which are equal.

Where it introduces no confusion, we will identify a cycle by listing its vertices in order. As an example, when $d = 2$ (the case shown in Figure 2.1) there are 6 simple cycles: two with length one (00 and 33), one with length two (121), two with length three (0120 and 1321), and one with length four (01320).

Lemma 2.3. *A closed walk on a directed graph is a union of simple cycles.*

Proof. The argument is constructive. Given a closed walk, consider the list of vertices it visits (in order): $a_0, a_1, a_2, \ldots, a_k$. (By hypothesis, $a_0 = a_k$). Now start traversing the walk, and consider the first time a vertex is repeated: $a_i = a_j$ with $j > i$. Evidently $a_i, a_{i+1}, \ldots, a_j$ is a simple cycle. Remove this simple cycle from the walk, keeping a_i. This gives a new (shorter) closed walk, to which the same argument can be applied. Continuing, we eventually get a walk that has no repetitions aside from its intial and final vertices. It is evidently a simple cycle, so the proof is complete. □

2.3 The dynamic programming principle before and after scaling

The investor's worst-case regret is determined by a dynamic programming principle. The goal of this subsection is to make this explicit. As already explained in the introduction, our main focus will be on a *scaled* version of the game, and we'll get to that presently. But first we discuss the dynamic programming problem for the game described by Section 2.2.

Recall that the investor's cumulative regret is recorded by (x_1, x_2). As already noted in the Introduction, it is convenient to work with $\xi = x_1 - x_2$ and $\eta = x_1 + x_2$. If the investor's choice at step k is f_k then the change in x associated with actions at step k is, according to (2.4)–(2.6), $\Delta x = [b_k(q_k - f_k), b_k(r_k - f_k)]$. So the changes in ξ and η are $\Delta\xi = b_k(q_k - r_k)$ and $\Delta\eta = b_k(q_k + r_k - 2f_k)$. We suppose the game ends at step N (i.e. the last prediction is made at $k = N-1$).

The dynamic programming principle determines the function

$$U(k,m,\xi,\eta) = \begin{cases} \text{the investor's worst-case minimum final-time regret at step } k, \\ \text{if the recent history is } m, \text{ the value of } x_1 - x_2 \text{ (based on times} \\ \text{through } k-1) \text{ is } \xi, \text{ and the value of } x_1 + x_2 \text{ is } \eta, \end{cases}$$

by working backward in time: for $k \leq N-1$

$$\begin{aligned} &U(k,m,\xi,\eta) = \\ &\min_{|f|\leq 1} \max_{b_k=\pm 1} U\big(k+1, m_{b_k}, \xi + b_k[q(m) - r(m)], \eta + b_k[q(m) + r(m) - 2f]\big), \end{aligned} \tag{2.7}$$

where m_{b_k} denotes the next state, which is determined from m by the value of b_k. Since the investor's regret with respect to the best-performing expert is $\max\{x_1, x_2\} = (\eta + |\xi|)/2$, the final-time condition is

$$U(N,m,\xi,\eta) = \frac{\eta + |\xi|}{2}. \tag{2.8}$$

Our hypotheses about the flow of information are reflected by the min-max in (2.7): the investor chooses f knowing that the adversarial market will see her choice and do whatever is worst for her.

We are interested in what happens over large numbers of steps. As discussed in the Introduction, we expect $U(1,m,0,0) \sim \sqrt{N}$. Therefore it is natural to scale time by N and regret by $\sqrt{N}$. Writing $\varepsilon = 1/\sqrt{N}$, the scaled version of U is

$$u^\varepsilon(t,m,\xi,\eta) = \varepsilon U\left(\frac{t}{\varepsilon^2}, m, \frac{\xi}{\varepsilon}, \frac{\eta}{\varepsilon}\right). \tag{2.9}$$

It is defined for times t that are multiples of $\varepsilon^2 = 1/N$ between 0 and 1. The dynamic programming principle (2.7) is equivalent to

(2.10)

$$\begin{aligned} &u^\varepsilon(t,m,\xi,\eta) = \\ &\min_{|f|\leq 1} \max_{b_t=\pm 1} u^\varepsilon\big(t+\varepsilon^2, m_{b_t}, \xi + \varepsilon b_t[q(m) - r(m)], \eta + \varepsilon b_t[q(m) + r(m) - 2f]\big) \end{aligned}$$

with the same final-time condition as before:

$$u^{\varepsilon}(1,m,\xi,\eta) = \frac{\eta + |\xi|}{2}. \tag{2.11}$$

Our goal is to study $u = \lim_{\varepsilon\to 0} u^{\varepsilon}$ by combining ideas from PDE and graph theory. In our study of the continuum limit, it is natural to consider any final time T (rather than just $T = 1$), and it is natural to consider other final-time conditions (not just $(\eta + |\xi|)/2$). Therefore we shall henceforth assume that u^{ε} satisfies (2.10) up to a fixed final time T, with a more general final-time condition

$$u^{\varepsilon}(T,m,\xi,\eta) = \varphi(\xi,\eta). \tag{2.12}$$

Our analysis requires some conditions on the final-time condition φ; these will be discussed in Section 2.5.

To be sure the definition is clear: u^{ε} is defined, in general, by the dynamic programming principle (2.10). It is defined discretely in time (at $t = T$, $T - \varepsilon^2$, $T - 2\varepsilon^2$, etc.) for all $(\xi,\eta) \in \mathbb{R}^2$ and all states m in the relevant graph. Assuming only that φ is continuous, one easily proves by induction backward in time that $u^{\varepsilon}(t,m,\xi,\eta)$ is well-defined and continuous with respect to ξ and η at each discrete time. If φ is not one-homogeneous, then u^{ε} does not have the form (2.9) for any solution U of our unscaled game.

2.4 A formal connection to PDE

To see a connection with PDE, we begin by ignoring the dependence of u^{ε} on ε and m, and combining the dynamic programming principle (2.10) with Taylor expansion. Replacing u^{ε} by u in (2.10) and assuming u is sufficiently smooth, this gives

(2.13)
$$\begin{aligned} &u(t,m,\xi,\eta) \\ &= \min_{|f_{t,m}|\le 1} \max_{b_t=\pm 1} u\big(t+\varepsilon^2, m_{b_t}, \xi + \varepsilon b_t[q(m)-r(m)], \eta + \varepsilon b_t[q(m)+r(m)-2f]\big) \\ &= \min_{|f_{t,m}|\le 1} \max_{b_t=\pm 1} \big\{u(t,m,\xi,\eta) + \varepsilon A + \varepsilon^2 B + O(\varepsilon^3)\big\} \end{aligned}$$

with

$$A = b_t\big([q(m)-r(m)]u_{\xi} + [q(m)+r(m)-2f_{t,m}]u_{\eta}\big) \tag{2.14}$$

and

(2.15)
$$B = u_t + \frac{\langle D^2u[q(m)-r(m), q(m)+r(m)-2f_{t,m}], [q(m)-r(m), q(m)+r(m)-2f_{t,m}]\rangle}{2}.$$

In the latter expression, D^2u is the Hessian of u with respect to its spatial variables ξ and η:

$$D^2u = \begin{pmatrix} u_{\xi\xi} & u_{\xi\eta} \\ u_{\xi\eta} & u_{\eta\eta} \end{pmatrix}. \tag{2.16}$$

The order ε^0 term in the min-max (2.13) is not interesting – it is independent of $f_{t,m}$ and b_t, and in fact it cancels with the u that's on the left. So the first interesting term in the min-max is εA. If it is nonzero, then the market can choose the sign of b_t to make this term positive – a bad outcome for the investor. It is tempting to conclude that the investor's choice of $f_{t,m}$ should make A vanish, but that's not quite right. Rather, the investor's choice of $f_{t,m}$ should make this term small enough that it interacts with the ε^2 term. This occurs when $A = O(\varepsilon)$, i.e.

$$[q(m) - r(m)]u_\xi + [q(m) + r(m) - 2f_{t,m}]u_\eta = O(\varepsilon),$$

or equivalently

$$f_{t,m} = \frac{[q(m) - r(m)]u_\xi + [q(m) + r(m)]u_\eta}{2u_\eta} + O(\varepsilon).$$

So the investor's choice should take the form

$$f_{t,m} = f^*_{t,m} + \varepsilon f^\#_{t,m} \tag{2.17}$$

with

$$f^*_{t,m} = \frac{[q(m) - r(m)]u_\xi + [q(m) + r(m)]u_\eta}{2u_\eta}. \tag{2.18}$$

(We shall explain later, in Section 2.5, why this choice is admissible – that is, why $|f_{t,m}| \le 1$ – under suitable hypotheses upon the final-time function φ.) When $f_{t,m}$ has the form (2.17), the "first-order term" from the Taylor expansion becomes

$$\begin{aligned} \varepsilon A &= \varepsilon b_t \big([q(m) - r(m)]u_\xi + [q(m) + r(m) - 2(f^*_{t,m} + \varepsilon f^\#_{t,m})]u_\eta\big) \\ &= -\varepsilon^2 2 b_t u_\eta f^\#_{t,m}. \end{aligned}$$

As for the "second-order term" $\varepsilon^2 B$: evaluating B at $f^*_{t,m}$ rather than $f_{t,m}$ introduces an error of order ε; this leads (after some algebra) to

$$\varepsilon^2 B = \varepsilon^2 \left(u_t + (q(m) - r(m))^2 \frac{1}{2} \langle D^2u \frac{\nabla^\perp u}{u_\eta}, \frac{\nabla^\perp u}{u_\eta} \rangle \right) + O(\varepsilon^3), \tag{2.19}$$

using the notation

$$\nabla^\perp u = (-u_\eta, u_\xi). \tag{2.20}$$

Summarizing the preceding calculation: when f has the form (2.17), (2.13) reduces (after division by ε^2) to

$$0 = \min_{f^\#_{t,m}} \max_{b_t = \pm 1} \left\{ u_t + (q(m) - r(m))^2 \frac{1}{2} \langle D^2u \cdot \frac{\nabla^\perp u}{u_\eta}, \frac{\nabla^\perp u}{u_\eta} \rangle - 2 b_t u_\eta f^\#_{t,m} \right\} \tag{2.21}$$

if we ignore the error associated with higher-order terms in the Taylor expansion.

We have thus far ignored the possible dependence of u on m. It seems from (2.21) that u should indeed depend on the state m as well as on t, ξ, and η. This would be a mess, since we would need to keep track of how the states evolve.

In fact, the function u we want is a little different from the one discussed thus far. The min-max should be over the players' (multistep) *strategies*, and the expression under the min-max should involve the *time-averaged value* of the expression under the min-max in (2.21). By proceeding this way, we shall obtain a function u that depends only on time and space (not on the current state).

To avoid unwieldy expressions, we introduce the notation

$$D_k := \frac{1}{2}\langle D^2 u \cdot \frac{\nabla^\perp u}{u_\eta}, \frac{\nabla^\perp u}{u_\eta}\rangle, \tag{2.22}$$

where the right hand side is evaluated at time t_k and location ξ_k, η_k. The expression to be time-averaged is thus

$$L(t_k, m, b, \xi_k, \eta_k, f^{\#}_{t_k,m}) := u_t + (q(m) - r(m))^2 D_k - 2bu_\eta f^{\#}_{t_k,m}. \tag{2.23}$$

To identify the optimal strategies and implement the time averaging, let us focus on the case $d = 1$, for which the graph is shown in Figure 2.2. It has three simple cycles: $0-0$, $1-1$, and $0-1-0$. The investor knows that the market could choose

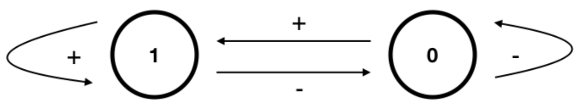

FIGURE 2.2. The directed graph when $d = 1$.

to simply repeat one of these cycles. For the cycle $0-0$, the value of L is

$$L(t_k, 0, -1, \xi_k, \eta_k, f^{\#}_{t_k,0}) = u_t + (q(0) - r(0))^2 D_k - 2(-1)u_\eta f^{\#}_{t_k,0}; \tag{2.24}$$

for the cycle $1-1$ it is

$$L(t_k, 1, 1, \xi_k, \eta_k, f^{\#}_{t_k,1}) = u_t + (q(1) - r(1))^2 D_k - 2(+1)u_\eta f^{\#}_{t_k,1}; \tag{2.25}$$

for the two-step cycle $0-1-0$ the time-averaged value is

$$\begin{aligned} &\frac{1}{2}\{L(t_k, 1, -1, \xi_k, \eta_k, f^{\#}_{t_k,1}) + L(t_k, 0, 1, \xi_k, \eta_k, f^{\#}_{t_k,0})\} = \\ &u_t + \frac{1}{2}\{(q(1) - r(1))^2 D_k - 2(-1)u_\eta f^{\#}_{t_k,1} + (q(0) - r(0))^2 D_k - 2(+1)u_\eta f^{\#}_{t_k,0}\}. \end{aligned} \tag{2.26}$$

Knowing that the market is adversarial, the investor assumes that the market will choose whichever cycle gives her the worst result, and chooses $f^{\#}_{t,m}$ to optimize the outcome associated with the worst-case cycle. This leads, for general d, to a linear program (the "investor's linear program" studied in Section 4.1). For $d = 1$, however, the solution of the linear program has a simple and intuitive character:

the best choice of $f^{\#}_{t,m}$ is the one that makes the investor *indifferent* with respect to the three simple cycles. (A similar phenomenon occurs for $d = 2,3,4$, as we will show in Section 4.4.) Writing $D = D_k$ (a harmless convention, since D_k changes only slightly at each timestep) and writing α_m rather than $f^{\#}_{t,m}$ (a significant choice, since it restricts the investor's strategy to depend on time only through the state m and the values of u_η and D), the condition for indifference is that α_0 and α_1 satisfy

$$
\begin{aligned}
(q(1)-r(1))^2 D - 2u_\eta \alpha_1 &= (q(0)-r(0))^2 D + 2u_\eta \alpha_0 \\
&= \frac{1}{2}\left[(q(0)-r(0))^2 + (q(1)-r(1))^2\right] D + u_\eta \alpha_1 - u_\eta \alpha_0.
\end{aligned}
$$

This amounts to two linear equations in the two unknowns α_0 and α_1; the unique solution is easily seen to be

$$\alpha_0 = \alpha_1 = \frac{(q(1)-r(1))^2 - (q(0)-r(0))^2}{4u_\eta} D$$

and the time-averaged value of L for each cycle is then

$$u_t + \frac{(q(0)-r(0))^2 + (q(1)-r(1))^2}{2} D. \tag{2.27}$$

Taking this choice of strategy into account, we obtain from (2.21) the desired PDE for the case $d = 1$:

$$u_t + \frac{1}{2} C^{\#}_1 \langle D^2 u \frac{\nabla^\perp u}{u_\eta}, \frac{\nabla^\perp u}{u_\eta} \rangle = 0 \tag{2.28}$$

with

$$C^{\#}_1 = \frac{(q(0)-r(0))^2 + (q(1)-r(1))^2}{2}. \tag{2.29}$$

It is to be solved for $t < T$, with final-time data $u = \varphi$ at time T.

The preceding derivation is of course quite heuristic. We will justify it with full mathematical rigor in Section 3. The ideas presented above provide a strategy for the investor, and the outline of a proof that by following it she can do at least as well as the solution of the PDE. The heuristic argument considered only simple cycles, while the rigorous one (presented in Section 3.1) must consider any possible actions the market might take. This is where Lemma 2.3 comes in: while the market's choices might not produce a walk that's restricted to a simple cycle, the walk they produce can be decomposed as a union of simple cycles – so a strategy that works well for every simple cycle actually works well for every walk.

How do we know there is no *better* strategy for the investor? This will be shown in Section 3.2 by considering a particular strategy for the market. Briefly, the strategy "forces" the investor to use the values of f^*_m and $f^{\#}_m$ identified by our heuristic argument, by penalizing other choices.

Our heuristic discussion has focused on the case $d = 1$. The cases $d = 2,3,4$ are similar, since in those cases too the parameters $f^{\#}_m$ can be chosen to make the investor indifferent with respect to the various simple cycles; the associated PDEs have the same form as (2.28), except that the "diffusion constant" is different. (The values of the constants $C_2^\#$, $C_3^\#$, and $C_4^\#$, corresponding to $d = 2,3,4$, are determined in Section 4.4.) For $d \geq 5$ we do not know whether indifference is achievable. As a result, for $d \geq 5$ our method only gives upper and lower bounds for the investor's worst-case final-time regret. Both bounds involve PDE's of the form (2.28) with suitable "diffusion constants." The arguments are largely parallel to those sketched in this section for $d = 1$. To find a good strategy for the investor, we choose the parameters $f_m^\#$ to minimize the rate at which regret accumulates for the most dangerous cycle. To find a good strategy for the market, we choose the parameters $f_m^\#$ to maximize the rate at which regret accumulates for the cycle most favorable to the investor. The optimal choices of $f_m^\#$ are determined by a pair of linear programs, presented in Section 4. One might have expected these linear programs to be dual, but they are not; Section 4.3 explains why not.

2.5 Properties of u required for our analysis

The heuristic arguments in Section 2.4 rely heavily on Taylor expansion, and so do our proofs. Therefore we need that our PDE has a solution $u(t,\xi,\eta)$ that is sufficiently smooth, in the sense that

(2.30)
$$\begin{gathered} u \text{ has continuous, uniformly bounded derivatives} \\ \text{in } (\xi,\eta) \text{ of order up to 4,} \\ u_t \text{ has continuous, uniformly bounded derivatives} \\ \text{in } (\xi,\eta) \text{ of order up to 2, and} \\ u_{tt} \text{ is continuous and uniformly bounded,} \end{gathered}$$

all bounds being uniform as $t \uparrow T$. Our setup also has some additional requirements: it requires that

(2.31)
$$u_\eta > c > 0$$

(with a lower bound c that is uniform in space and valid for all times between t and T), and

(2.32)
$$|u_\xi| \leq u_\eta.$$

In Section 5 we shall also need

(2.33)
$$u_t \leq 0.$$

The relevance of (2.31) is clear, since many of the expressions that emerged from our heuristic discussion had u_η in the denominator. To explain the relevance of (2.32), we recall that $\xi = x_1 - x_2$ and $\eta = x_1 + x_2$, where (x_1,x_2) are the investor's regrets with respect to the two experts. By chain rule, $u_{x_1} = u_\xi + u_\eta$ and $u_{x_2} = -u_\xi + u_\eta$; so (2.32) is equivalent to $u_{x_1} \geq 0$ and $u_{x_2} \geq 0$. These conditions are quite intuitive: since regret accumulates as the game proceeds, starting at time t with more regret should mean ending at time T with more regret. They are also necessary for our heuristic calculation to make sense. Indeed, recall from (2.17) that if the investor is in state m at time t, her choice takes the form $f_{t,m} = f_{t,m}^* +$

$\varepsilon f^{\#}_{t,m}$, where $f^*_{t,m}$ is given by (2.18). Rewriting the formula for $f^*_{t,m}$ in terms of u_{x_1} and u_{x_2} instead of u_ξ and u_η, one finds that

$$f^*_{t,m} = q(m)\frac{u_{x_1}}{u_{x_1}+u_{x_2}} + r(m)\frac{u_{x_2}}{u_{x_1}+u_{x_2}}. \tag{2.34}$$

Since u_{x_1} and u_{x_2} are nonnegative (and their sum is strictly positive, by (2.31)), $f^*_{t,m}$ is a weighted average of the experts' choices $q(m)$ and $r(m)$. The rules of the game require that $|f_{t,m}| \leq 1$. Since $f_{t,m} = f^*_{t,m} + \varepsilon f^{\#}_{t,m}$ and we don't know the sign of $f^{\#}_{t,m}$ in advance, we need $|f^*_{t,m}| < 1$. This is assured by (2.34), together with our hypothesis (2.2)) that $|q(m)| < 1$ and $|r(m)| < 1$ for every state m.

Condition (2.33) reflects again the idea that regret accumulates as the game proceeds. If we start the game a little later (with the same initial regrets, and the same final time T) then the total time that the game runs is smaller, so the final-time regret should be smaller. Thus $u(t,\xi,\eta)$ should be a decreasing function of t when ξ and η are held fixed.

These conditions on u are effectively conditions on the final-time data φ. As already noted in the Introduction, it is very natural to focus on final-time data of the form

$$\varphi(\xi,\eta) = c\eta + \overline{\varphi}(\xi). \tag{2.35}$$

Then the PDE

$$u_t + \frac{1}{2}C\langle D^2u\frac{\nabla^\perp u}{u_\eta}, \frac{\nabla^\perp u}{u_\eta}\rangle = 0 \tag{2.36}$$

with $u = \varphi$ at $t = T$ is solved by $u = c\eta + \overline{u}(t,\xi)$, where

$$\overline{u}_t + \frac{1}{2}C\overline{u}_{\xi\xi} = 0 \text{ for } t < T, \text{ with } \overline{u} = \overline{\varphi} \text{ at } t = T. \tag{2.37}$$

(We do not need to discuss in what class the solution of (2.36) is unique, since our verification arguments rely on the smoothness of u, not its uniqueness; however the results in Appendix A can be used to prove uniqueness within a suitable class.) By standard results about the linear heat equation, u has the required properties provided that $\overline{\varphi}$ is C^4 with uniformly bounded fourth derivatives, $|\overline{\varphi}_\xi| \leq c$ for all ξ, and $\overline{\varphi}_{\xi\xi} \geq 0$.

Interestingly, our PDE can be reduced to the linear heat equation for a much more general class of final-time data. We first learned this from Y. Giga. The reduction was studied in detail by Kangping Zhu in [25], and it is reviewed in Appendix A. In particular, we show there that our PDE (2.36) with final data φ has a solution satisfying (2.30)–(2.33) provided

$$\varphi \text{ is } C^4 \text{ with uniformly bounded derivatives of order up to 4,} \tag{2.38}$$

$$\varphi_\eta > c > 0 \text{ for some constant } c, \tag{2.39}$$

$$|\varphi_\xi| \le \varphi_\eta, \text{ and} \tag{2.40}$$

$$\varphi_{\xi\xi}\varphi_\eta^2 - 2\varphi_{\xi\eta}\varphi_\xi\varphi_\eta + \varphi_{\eta\eta}\varphi_\xi^2 \ge 0. \tag{2.41}$$

For machine learning, we are especially interested in the final-time data $\varphi(\xi,\eta) = \frac{1}{2}(\eta+|\xi|) = \max\{x_1,x_2\}$. This has the form (2.35) with $\overline{\varphi} = |\xi|/2$, so u is determined by solving the linear heat equation (2.37) with final-time data $|\xi|/2$. Evidently u is *not* uniformly C^4, since its second derivatives with respect to ξ blow up at $t = T$. Our method can still be applied, however, by bounding this φ above or below by a smooth function and considering the associated u. For our implementation of this idea, see Sections 3.3 and 5.3. (It makes a difference, of course, exactly how φ is smoothed; we use parabolic smoothing.)

2.6 Summary of our main results

In Section 2.4 we argued heuristically that our scaled value function u^ε is related, in the limit $\varepsilon \to 0$, to the solution of a PDE. The rest of this paper is devoted to proving rigorous results of this type. Careful statements will be found in the subsequent sections, but here is a brief summary:

- In our heuristic discussion, we eventually focused on the special case $d = 1$, when the experts' predictions depend only on the most recent market move. Our rigorous analysis also begins with this case, since it permits us to address the essential issues in a relatively uncluttered environment. Theorems 3.1 and 3.2 assume that the solution u of our heuristic PDE (2.28) satisfies (2.30)–(2.32). They provide upper and lower bounds on u^ε, which taken together show that $|u^\varepsilon - u| \le C[(T-t)+\varepsilon]\varepsilon$. While the classic case $\varphi(\xi,\eta) = \frac{1}{2}(\eta+|\xi|)$ is not covered by Theorems 3.1 and 3.2, the additional ideas needed to handle it were already present in [25]; Theorem 3.3 uses them to show that for this classic case (and still assuming $d = 1$) we have $|u^\varepsilon - u| \le C\varepsilon|\log\varepsilon|$.
- In our heuristic discussion, the predictor's strategy had the form $f_m = f_m^* + \varepsilon f_m^\#$ (see e.g. (2.17)). The value of f_m^* was immediately clear, but the value of $f_m^\#$ required more thought. In Section 2.4 we argued that the best choice of $f_m^\#$ was the one that made the investor indifferent between the simple cycles of the relevant graph. For general d we do not know whether there exists such a choice of $f_m^\#$. But every choice of $f_m^\#$ represents a candidate strategy for the investor, and there is a linear program that identifies the best. Our strategies for the market have a similar character: the market "forces" the investor to make a particular choice of $f_m^\#$ (by penalizing her if she doesn't), and there is a linear program that identifies the market's best choice. These linear programs are discussed in Section 4. For general d we

do not have explicit solutions of the linear programs, and we do not know whether their optimal values match. Our understanding is, however, more complete for $d = 2,3,4$; we shall show in Section 4.4 that these cases are like $d = 1$: there is a choice of $f_m^\#$ that achieves indifference. This choice is an explicit solution to both linear programs, and it demonstrates that their optimal values match.

- Theorems 5.2 and 5.5 show that any admissible choice for the investor's linear program determines a PDE-based upper bound for u^ε, and any admissible choice for the market's linear program determines a PDE-based lower bound for u^ε. When the two linear programs have the same optimal value (which happens at least for $d = 2,3,4$) we obtain an estimate of the form $|u^\varepsilon - u| \le C[(T-t)+\varepsilon]\varepsilon$, entirely analogous to the situation for $d = 1$. For the classic final-time data $\varphi = \frac{1}{2}(\eta + |\xi|)$ similar estimates hold, except that (as for $d = 1$) the error is of order $\varepsilon|\log\varepsilon|$. (This is the assertion of Theorem 5.8.)

Throughout this paper there are only *two* history-dependent experts. Subsequent to this work, the first author and J. Calder have obtained further results in [4, 5]. These papers consider any number of experts n and any finite d. The paper [4] (which came first) identifies $u = \lim_{\varepsilon\to 0} u^\varepsilon$ for any n and d using methods rather different from those of this paper; however it obtains a convergence rate that's slower than the one obtained here for $n = 2$ and $d \le 4$. (Our rate is ε for smooth data and $\varepsilon|\log\varepsilon|$ for the classic case, whereas the method of [4] gives only $\varepsilon^{1/3}$ and $\varepsilon^{1/3}|\log\varepsilon|$). The later paper [5] uses methods closer to ours, and therefore its convergence rates are the same as ours. Its contributions include a proof that our upper and lower bound linear programs have the same value for any n and d, and an alternative (simpler) analysis of the limit $\varepsilon \to 0$.

3 Upper and lower bounds for $d = 1$

This section provides a fully rigorous treatment of the special case $d = 1$, when the experts' advice depends only on the most recent stock move. The upper bound is proved by considering a good strategy for the investor – namely, the one developed in Section 2.4. The lower bound is proved by considering a good strategy for the market – namely, the one described at the beginning of Section 3.2.

3.1 The upper bound, when φ is regular

We assume throughout this section that $d = 1$. Let $u^\varepsilon(t,m,\xi,\eta)$ be defined by the dynamic programming principle (2.10) with final value $u^\varepsilon(T,m,\xi,\eta) = \varphi(\xi,\eta)$ (it is defined only for times t such that $(T-t)/\varepsilon^2$ is an integer). Let

$u(t,\xi,\eta)$ be the solution of the PDE

$$u_t + \frac{1}{2}C_1^{\#}\langle D^2u\frac{\nabla^{\perp}u}{u_\eta}, \frac{\nabla^{\perp}u}{u_\eta}\rangle = 0, \quad u(T,\xi,\eta) = \varphi(\xi,\eta), \tag{3.1}$$

where

$$C_1^{\#} = \frac{(q(1)-r(1))^2 + (q(0)-r(0))^2}{2}. \tag{3.2}$$

Our goal is to prove

Theorem 3.1. *Suppose $d = 1$ and assume the PDE solution u satisfies* (2.30)–(2.32). *Then there is a constant C (independent of ε, t, and T) such that*

$$u^{\varepsilon}(t,m,\xi,\eta) \le u(t,\xi,\eta) + C[(T-t)+\varepsilon]\varepsilon \tag{3.3}$$

for $t < T$, $\xi \in \mathbb{R}$, $\eta \in \mathbb{R}$, and $m \in \{0,1\}$, provided that ε is small enough and t is such that $N = (T-t)/\varepsilon^2$ is an integer.

Proof. An outline of the proof is as follows: to estimate $u^{\varepsilon}(t_0,m_0,\xi_0,\eta_0)$, we shall define a sequence (t_k,m_k,ξ_k,η_k) along which u^{ε} is monotone

$$u^{\varepsilon}(t_0,m_0,\xi_0,\eta_0) \le u^{\varepsilon}(t_1,m_1,\xi_1,\eta_1) \le \cdots \le u^{\varepsilon}(t_N,m_N,\xi_N,\eta_N)$$

with $t_N = T$, so that

$$u^{\varepsilon}(t_N,m_N,\xi_N,\eta_N) = \varphi(\xi_N,\eta_N) = u(t_N,\xi_N,\eta_N).$$

Then we'll show that u is nearly constant along the sequence:

$$|u(t_0,\xi_0,\eta_0) - u(t_N,\xi_N,\eta_N)| \le C[(T-t)+\varepsilon]\varepsilon. \tag{3.4}$$

These estimates lead immediately to (3.3).

To define the sequence (t_k,m_k,ξ_k,η_k), we need only explain how (t_1,m_1,ξ_1,η_1) is chosen (then the rest of the sequence is determined similarly, step by step). Recall the dynamic programming principle (2.10), which we can write more compactly by defining, in terms of the investor's choice f,

$$v = \begin{pmatrix} v^1 \\ v^2 \end{pmatrix} = \begin{pmatrix} q(m)-r(m) \\ q(m)+r(m)-2f \end{pmatrix}. \tag{3.5}$$

The dynamic programming principle then says

$$u^{\varepsilon}(t,m,\xi,\eta) = \min_{|f|\le 1}\max_{b_t=\pm 1} u^{\varepsilon}(t+\varepsilon^2, m_{b_t}, \xi+\varepsilon b_t v^1, \eta+\varepsilon b_t v^2). \tag{3.6}$$

It follows that for any choice of f,

$$u^{\varepsilon}(t,m,\xi,\eta) \le \max_{b_t=\pm 1} u^{\varepsilon}(t+\varepsilon^2, m_{b_t}, \xi+\varepsilon b_t v^1, \eta+\varepsilon b_t v^2). \tag{3.7}$$

Applying this with $(t,m,\xi,\eta) = (t_0,m_0,\xi_0,\eta_0)$ and taking b_0 to achieve the max on the RHS, we see that for $t_1 = t_0+\varepsilon^2$, $m_1 = (m_0)_{b_0}$, $\xi_1 = \xi_0+\varepsilon b_0 v^1$, $\eta_1 = \eta_0 + \varepsilon b_0 v^2$ we have the desired inequality $u^{\varepsilon}(t_0,m_0,\xi_0,\eta_0) \le u^{\varepsilon}(t_1,m_1,\xi_1,\eta_1)$. While the preceding argument works for any choice of f, we must make a special choice if we want the solution of the PDE to be nearly constant along the sequence. We therefore choose f as indicated by the heuristic discussion in Section 2.4:

$$\begin{aligned} f_{t_0,1} &= f^*_{t_0,1} + \varepsilon f^\#_{t_0,1} \\ f_{t_0,0} &= f^*_{t_0,0} + \varepsilon f^\#_{t_0,0} \\ f^*_{t_0,m} &= \frac{(q(m)-r(m))u_\xi + (q(m)+r(m))u_\eta}{2u_\eta} \\ f^\#_{t_0,1} &= f^\#_{t_0,0} = \frac{(q(1)-r(1))^2 - (q(0)-r(0))^2}{4u_\eta} D, \end{aligned} \tag{3.8}$$

with the familiar convention that

$$D = \frac{1}{2}\langle D^2 u \cdot \frac{\nabla^\perp u}{u_\eta}, \frac{\nabla^\perp u}{u_\eta}\rangle, \tag{3.9}$$

and the understanding that u_η, u_ξ, and D are evaluated at (t_0, ξ_0, η_0). As we explained in Section 2.5, this choice of f is admissible (i.e. $|f| \le 1$) if ε is sufficiently small. (This is our only smallness condition on ε.) In summary: given $(t_0, m_0, \xi_0, \eta_0)$, the point $(t_1, m_1, \xi_1, \eta_1)$ is the location that shows up on the RHS of the dynamic programming principle when f is chosen by (3.8) and b_0 achieves the max over $b_0 = \pm 1$. The rest of the sequence $(t_k, m_k, \xi_k, \eta_k)$ is determined similarly, step by step. (The values of D, f^*_m, b, and $f^\#_m$ at step k will be called D_k, $f^*_{k,m}$, b_k, and $f^\#_{k,m}$.)

Our remaining task is to prove the near-constancy of u along our sequence, (3.4). We begin by estimating the increments

$$U_k := u(t_{k+1}, \xi_{k+1}, \eta_{k+1}) - u(t_k, \xi_k, \eta_k).$$

There are four cases, depending on the values of m_k and b_k.

Case 1: $m_k = 0, b_k = -1$.
This case is relatively easy, since the market is effectively following one of the cycles in the $d = 1$ graph (namely, the cycle $0-0$). Taylor expanding u around (t_k, ξ_k, η_k) as we did in (2.13), the calculation in Section 2.4 shows that

$$\begin{aligned} U_k &= \varepsilon^2 \left[u_t + (q(0)-r(0))^2 D_k - 2(-1) u_\eta f^\#_{k,0} \right] + O(\varepsilon^3) \\ &= \varepsilon^2 L(t_k, 0, -1, \xi_k, \eta_k, f^\#_{k,0}) + O(\varepsilon^3) \end{aligned}$$

where for the second line we used the definition of L, (2.23) (which reduces in this case to (2.24)). Moreover, the value of $L(t_k, 0, -1, \xi_k, \eta_k, f^\#_{t_k,0})$ is exactly $u_t + C^\#_1 D_k$ by (2.27), which equals zero by the PDE (3.1). Thus, in case 1 we have

$$U_k = O(\varepsilon^3).$$

Case 2: $m_k = 1, b_k = 1$.
This case is similar to the first, since the market is effectively following another cycle in the $d = 1$ graph (namely, the cycle $1-1$). Arguing as in Case 1, we get

$$\begin{aligned} U_k &= \varepsilon^2\left[u_t + (q(1)-r(1))^2 D_k - 2(+1)u_\eta f^\#_{k,1}\right] + O(\varepsilon^3) \\ &= \varepsilon^2 L(t_k, 1, 1, \xi_k, \eta_k, f^\#_{k,1}) + O(\varepsilon^3) \\ &= O(\varepsilon^3). \end{aligned}$$

Case 3: $m_k = 0, b_k = 1$.
This case is different, since the market is doing just half of the two-step cycle $0-1-0$. Starting as in the previous cases, we have

$$\begin{aligned} U_k &= \varepsilon^2\left[u_t + (q(0)-r(0))^2 D_k - 2(+1)u_\eta f^\#_{k,0}\right] + O(\varepsilon^3) \\ &= \varepsilon^2 L(t_k, 0, 1, \xi_k, \eta_k, f^\#_{k,0}) + O(\varepsilon^3). \end{aligned}$$

However the value of L in this case is no longer 0. Rather, it is

$$\begin{aligned} &u_t + (q(0)-r(0))^2 D_k - 2u_\eta f^\#_{k,0} \\ &\qquad = u_t + (q(0)-r(0))^2 D_k - \frac{1}{2}\left[(q(1)-r(1))^2 - (q(0)-r(0))^2\right] D_k \\ &= u_t + \frac{1}{2}\left[(q(0)-r(0))^2 + (q(1)-r(1))^2\right] D_k + \left[(q(0)-r(0))^2 - (q(1)-r(1))^2\right] D_k. \end{aligned}$$

The sum of the first two terms on the right vanishes, as a consequence of the PDE; therefore

$$U_k = \varepsilon^2\left[(q(0)-r(0))^2 - (q(1)-r(1))^2\right] D_k + O(\varepsilon^3).$$

Case 4: $m_k = 1, b_k = -1$.
This case is similar to the last one, since the market is again doing just half of the two-step cycle $0-1-0$. We have

$$\begin{aligned} U_k &= \varepsilon^2(u_t + (q(1)-r(1))^2 D_k - 2(-1)u_\eta f^\#_{k,1} + O(\varepsilon^3) \\ &= \varepsilon^2 L(t_k, 1, -1, \xi_k, \eta_k, f^\#_{k,1}) + O(\varepsilon^3) \end{aligned}$$

and the value of L this time is

$$\begin{aligned} &u_t + (q(1)-r(1))^2 D_k + 2u_\eta f^\#_{k,1} \\ &\qquad = u_t + (q(1)-r(1))^2 D_k + \frac{1}{2}\left[(q(1)-r(1))^2 - (q(0)-r(0))^2\right] D_k \\ &= u_t + \frac{1}{2}\left[(q(0)-r(0))^2 + (q(1)-r(1))^2\right] D_k + \left[(q(1)-r(1))^2 - (q(0)-r(0))^2\right] D_k. \end{aligned}$$

The sum of the first two terms on the right vanishes as before, so

$$U_k = \varepsilon^2\left[(q(1)-r(1))^2 - (q(0)-r(0))^2\right] D_k + O(\varepsilon^3).$$

(Notice that the ε^2 terms from cases 3 and 4 sum to zero. This had to be so, since we know from Section 2.4 that the average value of L over the cycle $0-1-0$ is $u_t + C_1^{\#} D_k$, which equals 0.)

We want to show that $|U_0 + \ldots U_{N-1}| \le C[(T-t)+\varepsilon]\varepsilon$. The $O(\varepsilon^3)$ terms in the estimates for U_k are consistent with this: since there are $N = (T-t)/\varepsilon^2$ of them, they accumulate at worst to an error that's bounded by $C(T-t)\varepsilon$. So our task is to control the sum of the order-ε^2 terms from Cases 3 and 4. If the value of D_k didn't change with k this would be easy; but alas, it does change with k, since D_k is the value of (3.9) evaluated at (t_k, ξ_k, η_k).

Consider the walk on the $d = 1$ graph that's associated with our sequence. Suppose transitions from 1 to 0 happen at steps $i_1 < i_2 < \ldots$ and transitions from 0 to 1 happen at steps $j_1 < j_2 < \ldots$. These transitions must – by their essential character – be ordered. Making a choice about which comes first, we may assume (without loss of generality) that $i_1 < j_1$. Depending upon which type of transition comes last, the full list of transtions between 0 and 1 is either of the form

$$i_1 < j_1 < i_2 < j_2 < \ldots i_K < j_K \tag{3.10}$$

or

$$i_1 < j_1 < i_2 < j_2 < \ldots i_K < j_K < i_{K+1}. \tag{3.11}$$

Either way, we have

$$\sum_{n=1}^{K} (j_n - i_n) \le N. \tag{3.12}$$

If (3.10) holds, then the sum of the ε^2 terms (in our estimates for U_k) is precisely

$$\varepsilon^2 \left[(q(1) - r(1))^2 - (q(0) - r(0))^2 \right] \sum_{n=1}^{K} (D_{i_n} - D_{j_n}). \tag{3.13}$$

Now, D_{i_n} and D_{j_n} represent the same function (3.9) evaluated at different points in space-time, which differ in time by $(j_n - i_n)\varepsilon^2$ and in space by at most a constant times $(j_n - i_n)\varepsilon$. Our hypotheses on u assure that the expression being evaluated has uniformly bounded derivatives with respect to t, ξ, and η, so it is globally Lipschitz continuous. We conclude that

$$|D_{i_n} - D_{j_n}| \le C(j_n - i_n)\varepsilon. \tag{3.14}$$

Combining this with (3.12) gives

$$\varepsilon^2 \sum_{n=1}^{K} |D_{i_n} - D_{j_n}| \le CN\varepsilon^3 = C(T-t)\varepsilon.$$

So by (3.13), the ε^2 terms in our estimates for U_k sum to at most a constant times $(T-t)\varepsilon$.

If the situation is (3.11) rather than (3.10), the the same argument applies, but the ε^2 term in $U_{i_{K+1}}$ must be handled separately. In this case the ε^2 terms in our estimates for U_k sum to at most a constant times $(T-t)\varepsilon+\varepsilon^2$.

In either situation, these estimates establish (3.4) , completing the proof of the theorem. □

3.2 The lower bound, when φ is regular

Our lower bound shows that Theorem 3.1 is asymptotically sharp. Its proof is largely parallel to that of the upper bound, except that this time we use a good strategy for the market. To describe the strategy, recall that our heuristic discussion used Taylor expansion to estimate the increments of u. As we observed just after (2.13), the investor must make the "first-order term" εA nearly vanish, since otherwise this term dominates and the market can choose b to make it positive; this determined f_m^*, the leading-order term in the investor's strategy. A more subtle analysis led us to guess the optimal next-order term, and the resulting strategy $f_m = f_m^* + \varepsilon f_m^\#$ was at the heart of our upper bound.

Our strategy for the market reflects the two-step character of the heuristic discussion:

Case 1: If the investor's choice doesn't nearly zero out the "first-order term," then the market chooses b to make that term positive; more quantitatively,

(3.15) if the investor's choice f has $|f-f_m^*|\geq\gamma\varepsilon$ then the market chooses b so that $-b(f-f_m^*)\geq 0$.

Case 2: When case 1 doesn't apply, it is convenient to express the investor's choice f as $f=f_m^*+\varepsilon f_m^\#+\varepsilon X$ (this relation defines X). The investor is more optimistic than our conjectured optimal strategy if $X>0$, and more pessimistic if $X<0$. In the former case the market makes the stock go down, and in the latter case it makes the stock go up – in each case giving the investor an unwelcome surprise; more quantitatively:

(3.16) if the investor's choice f has $|f-f_m^*|<\gamma\varepsilon$ then the market chooses b so that $-bX\geq 0$.

Here γ is a constant, which cannot be too small; specifically, it must satisfy (3.24) and (3.25) below. This strategy lies at the heart of the following lower bound.

Theorem 3.2. *Suppose $d=1$, let u^ε be defined by the dynamic programming principle* (2.10) *with final-time condition φ, let u solve the PDE* (3.1) *with the same final-time condition, and assume u satisfies* (2.30)–(2.32). *Then there is a constant C (independent of ε, t, and T) such that*

$$u^\varepsilon(t,m,\xi,\eta)\geq u(t,\xi,\eta)-C[(T-t)+\varepsilon]\varepsilon \tag{3.17}$$

for $t<T$, $\xi\in\mathbb{R}$, $\eta\in\mathbb{R}$, and $m\in\{0,1\}$, provided that ε is small enough and t is such that $N=(T-t)/\varepsilon^2$ is an integer.

Proof. For the upper bound, we estimated $u^\varepsilon(t_0,m_0,\xi_0,\eta_0)$ by choosing a sequence (t_k,m_k,ξ_k,η_k) along which u^ε was increasing and u was nearly constant. For the lower bound, we shall use a different sequence along which u^ε is decreasing

$$u^\varepsilon(t_0,m_0,\xi_0,\eta_0) \geq u^\varepsilon(t_1,m_1,\xi_1,\eta_1) \geq \cdots \geq u^\varepsilon(t_N,m_N,\xi_N,\eta_N) \tag{3.18}$$

with $t_N = T$, so that

$$u^\varepsilon(t_N,m_N,\xi_N,\eta_N) = \varphi(\xi_N,\eta_N) = u(t_N,\xi_N,\eta_N). \tag{3.19}$$

The sequence will be chosen so that

$$u(t_N,\xi_N,\eta_N) - u(t_0,\xi_0,\eta_0) \geq -C[(T-t)+\varepsilon]\varepsilon. \tag{3.20}$$

These estimates lead immediately to (3.17).

We shall identify the sequence by explaining the choice of (t_1,m_1,ξ_1,η_1) (the rest of the sequence is then determined similarly, step by step). Recall our compact form of the dynamic programming principle, equation (3.6). Let f be the investor's optimal choice at (t_0,m_0,ξ_0,η_0); then the dynamic programming principle becomes

$$u^\varepsilon(t_0,m_0,\xi_0,\eta_0) = \max_{b=\pm1} u^\varepsilon(t_0+\varepsilon^2,(m_0)_b,\xi_0+\varepsilon b v^1,\eta_0+\varepsilon b v^2) \tag{3.21}$$

where v^1 and v^2 are defined by (3.5). Evidently, either choice $b=1$ or $b=-1$ gives a point $t_1 = t_0+\varepsilon^2$, $m_1 = (m_0)_b$, $\xi_1 = \xi_0+\varepsilon v^1$, $\eta_1 = \eta_0+\varepsilon v^2$ for which the desired inequality $u^\varepsilon(t_0,m_0,\xi_0,\eta_0) \geq u^\varepsilon(t_1,m_1,\xi_1,\eta_1)$ holds. We shall show that if b is chosen according to the proposed strategy (3.15)–(3.16) then we obtain the desired control on u.

We suppose henceforth that the sequence (t_k,m_k,ξ_k,η_k) has been chosen by applying the preceding argument inductively (using the proposed market strategy to determine the value of b_k at each step). Properties (3.18) and (3.19) are immediately clear. Our plan for demonstrating (3.20) is to show that the increments $U_k = u(t_{k+1},\xi_{k+1},\eta_{k+1}) - u(t_k,\xi_k,\eta_k)$ satisfy

$$U_k \geq \varepsilon^2\left[u_t + (q(m_k)-r(m_k))^2 D_k - 2b_k u_\eta f^{\#}_{t_k,m_k}\right] + O(\varepsilon^3). \tag{3.22}$$

Crucially: the RHS of (3.22) is the expression we found for the increment in our proof of the upper bound.

Given (3.22), the rest is easy: the desired control on u, (3.20), follows immediately from the argument we used for the upper bound (applied, of course, to the walk determined by our sequence (t_k,m_k,ξ_k,η_k)).

To prepare for the proof of (3.22), we state now the conditions we need for the constant γ (recall that Case 1 of the market's strategy applies when $|f-f^*_m| \leq \gamma\varepsilon$). Remember (from Cases 1–4 in the proof of the upper bound) that the ε^2 term on the RHS of (3.22) is either 0 or else

$$\pm\varepsilon^2\left[(q(1)-r(1))^2 - (q(0)-r(0))^2\right]D_k. \tag{3.23}$$

The expression in brackets is a fixed constant. Since we have a uniform bound on $|D_k|$ and a positive lower bound on u_η (by our hypotheses on u), by taking γ sufficiently large we can have

$$\gamma \geq \left|(q(1)-r(1))^2-(q(0)-r(0))^2\right| \max_{\xi,\eta\in\mathbb{R}^2,\ t<T} \frac{|D|}{u_\eta} \tag{3.24}$$

where D is defined as usual by (3.9). Next, recall from our heuristic discussion that when estimating the RHS of our dynamic programming principle (3.6) by Taylor expansion, the second-order-term is

$$\varepsilon^2\left[u_t+\frac{1}{2}\langle D^2u\cdot v_k, v_k\rangle\right]$$

where v_k is defined by (3.5). Since the investor must choose $|f|\leq 1$, we have a uniform bound $|v_k|\leq M$. Since we are assuming uniform bounds for u_t and D^2u (and recalling that u_η is bounded away from 0) by taking γ sufficiently large we can have

$$\gamma \geq \max_{\substack{\xi,\eta\in\mathbb{R}^2,\ t<T\\ |v|\leq M}} \frac{\left|u_t+\frac{1}{2}\langle D^2u\cdot v, v\rangle\right|}{u_\eta}. \tag{3.25}$$

These are the only conditions we place on γ.

We now prove (3.22) for the market's Case 1. At step k (when the current state is m_k) this means the investor's choice f_k has $|f_k-f^*_{k,m_k}|\geq\gamma\varepsilon$. (Here we write f_k and f^*_{k,m_k} rather than f_{t_k} and $f^*_{t_k,m_k}$ to simplify the notation.) Proceeding as we did in the heuristic discussion, the increment is

$$U_k=\varepsilon A+\varepsilon^2B+O(\varepsilon^3) \tag{3.26}$$

with

$$A=b_k(v^1u_\xi+v^2u_\eta)=b_k\big([q(m_k)-r(m_k)]u_\xi+[q(m_k)+r(m_k)-2f_k]u_\eta\big)$$

and

$$B=u_t+\frac{1}{2}\langle D^2u\cdot v_k, v_k\rangle.$$

Since A is an affine function of f_k and it vanishes when $f_k=f^*_{k,m_k}$, the expression for A simplifies to

$$A=-2b_ku_\eta(f_k-f^*_{k,m_k}). \tag{3.27}$$

Since we are in Case 1, the market chooses b_k so that A is positive, and we have

$$\varepsilon A\geq\varepsilon^2[2\gamma u_\eta].$$

Since $B\leq\gamma u_\eta$ by (3.25), we have

$$\varepsilon A+\varepsilon^2B\geq\varepsilon^2[\gamma u_\eta].$$

Making use now of (3.24), we conclude that

$$\varepsilon^2[\gamma u_\eta]\geq\varepsilon^2\left|(q(1)-r(1))^2-(q(0)-r(0))^2\right|D_k;$$

in view of (3.23), this implies the desired estimate (3.22).

Turning now to the market's Case 2, we start again with (3.26). The value of A is again given by (3.27). Writing $f_k - f^*_{k,m_k} = \varepsilon f^\#_{k,m_k} + \varepsilon X$ and choosing b_k by the market's strategy (3.16), we have

$$\varepsilon A \geq \varepsilon^2[-2b_k u_\eta f^\#_{k,m_k}].$$

Since $|f_k - f^*_{k,m_k}| \leq \gamma\varepsilon$, the "second-order term" $B = u_t + \frac{1}{2}\langle D^2 u \cdot v_k, v_k\rangle$ can be estimated as we did in the heuristic calculation: replacing f_k by f^*_{k,m_k} in the expression for v_k makes an error of order ε, and this leads to

$$\varepsilon^2 B = \varepsilon^2[u_t + (q(m) - r(m))^2 D_k] + O(\varepsilon^3).$$

Adding, we conclude that

$$\begin{aligned} U_k &= \varepsilon A + \varepsilon^2 B + O(\varepsilon^3) \\ &\geq \varepsilon^2\left[u_t + (q(m) - r(m))^2 D_k - 2b_k u_\eta f^\#_{k,m_k}\right] + O(\varepsilon^3), \end{aligned}$$

which is precisely (3.22). The proof of the theorem is now complete. □

3.3 The classic case, when $\varphi = (\eta + |\xi|)/2$

The classic goal of minimizing regret with respect to the best-performing expert corresponds, as we explained in Section 2.3, to using final-time data $\varphi = (\eta + |\xi|)/2$. Since this φ is not C^4, the solution of the associated PDE does not admit a uniform C^4 bound up to $t = T$, so Theorems 3.1 and 3.2 do not include this case. However we can obtain similar estimates (with an error estimate of order $\varepsilon|\log\varepsilon|$ rather than ε) by repeating the proofs of those theorems with proper attention to the error terms. A result of this type was proved by Zhu for two constant experts [25]; the proof of the following theorem uses essentially the same arguments.

Theorem 3.3. *Suppose $d = 1$ and consider the classic final-time data $\varphi = (\eta + |\xi|)/2$. We consider, as usual, the function $u^\varepsilon(t,m,\xi,\eta)$ defined by the dynamic programming principle* (2.10) *with final-time data φ, and the function $u(t,\xi,\eta)$ defined by the PDE* (3.1) *with final-time data φ. There is a constant C (independent of ε, t, and T) such that*

$$\left|u^\varepsilon(t,m,\xi,\eta) - u(t,\xi,\eta)\right| \leq C\varepsilon|\log\varepsilon| \tag{3.28}$$

for $t < T$, $\xi \in \mathbb{R}$, $\eta \in \mathbb{R}$, and $m \in \{0,1\}$, provided that ε is small enough and t is such that $N = (T-t)/\varepsilon^2$ is an integer.

Proof. As we observed in Section 2.5, the PDE simplifies dramatically in this case, and the solution is $u(t,\xi,\eta) = \frac{1}{2}\eta + \overline{u}(t,\xi)$ where $\overline{u}$ solves the linear heat equation $\overline{u}_t + \frac{1}{2}C_1^\#\overline{u}_{\xi\xi} = 0$ for $t < T$ with $\overline{u} = \frac{1}{2}|\xi|$ at $t = T$. We note that the "diffusion constant" $\frac{1}{2}C_1^\#$ (which is determined by (2.29)) is nonzero, since we assumed in (2.1) that the experts were distinct. Using the explicit solution formula for the linear heat equation we have

$$|\partial_\xi^k \partial_t^l \overline{u}| \leq C_{k,l}(T-t)^{-(k+2l-1)/2} \quad \text{for } k \geq 0 \text{ and } l \geq 0 \text{ with } k + 2l \geq 1. \tag{3.29}$$

Our overall strategy is to repeat the arguments used for Theorems 3.1 and 3.2, with $u(t,\xi,\eta)$ replaced by the smooth function $\tilde{u}(t,\xi,\eta) = u(t-\delta,\xi,\eta)$ (which satisfies the same PDE, with final-time data that's a smooth approximation to φ). We will choose the value of δ below to optimize the resulting estimate.

To see what emerges from this strategy, we need to revisit the various "error terms" that entered the proofs of the theorems, namely:

(i) those incurred by estimating the increments of u using Taylor expansion, i.e. the $O(\varepsilon^3)$ term in the estimate

$$u(t_{k+1},\xi_{k+1},\eta_{k+1}) - u(t_k,\xi_k,\eta_k) = \varepsilon A + \varepsilon^2 B + O(\varepsilon^3)$$

where A and B are given by (2.14) and (2.15) evaluated at step k;

(ii) those made by evaluating B at $f^*_{t_k,m_k}$ rather than f_{t_k,m_k}, i.e. the $O(\varepsilon^3)$ term in (2.19);

(iii) our estimate (3.14) for $|D_{i_n} - D_{j_n}|$; and

(iv) our estimate for the value of D_{K+1}, which was needed when analyzing a walk with a different number of transitions from 0 to 1 vs 1 to 0 (i.e. when (3.11) applies).

We also need to monitor errors associated with

(v) the difference $|u(t-\delta,\xi,\eta) - u(t,\xi,\eta)|$, evaluated both at the time t that enters the estimate, and at the final time T when $u(T,\xi,\eta) = \varphi(\xi,\eta)$.

We begin with some easy estimates, which are too crude to give (3.28) but already give a nontrivial result. For any $\delta > 0$ we have uniform bounds on the derivatives of $\tilde{u}(t,\xi,\eta) = u(t-\delta,\xi,\eta)$ for $t < T$, obtained by taking $(T-t) = \delta$ in (3.29). Applying this:

(i) The errors of type (i) are most easily estimated by writing the increment $u(t_{k+1},\xi_{k+1},\eta_{k+1}) - u(t_k,\xi_k,\eta_k)$ as

$$[u(t_{k+1},\xi_{k+1},\eta_{k+1}) - u(t_{k+1},\xi_k,\eta_k)] + [u(t_{k+1},\xi_k,\eta_k) - u(t_k,\xi_k,\eta_k)],$$

then estimating each term separately using Taylor's theorem. If $\delta^{-1}\varepsilon^2 \le 1/2$, the error due to estimating a single increment by Taylor expansion is of order $\delta^{-1}\varepsilon^3$. There are $N = (T-t)/\varepsilon^2$ such increments, so these accumulate to a term of order $\delta^{-1}(T-t)\varepsilon$.

(ii) Since the spatial second derivatives of $\tilde{u}$ are at most of order $\delta^{-1/2}$, the errors of type (ii) in the proof of the upper bound are at most $\delta^{-1/2}\varepsilon^3$ at each time step, accumulating to $\delta^{-1/2}(T-t)\varepsilon$ after summation over all steps. The situation is slightly worse in the proof of the lower bound, since in the market's Case 2 we only have $|f_{t,m} - f^*_{t,m}| \le \gamma\varepsilon$ and the size of γ is controlled by (3.24)–(3.25). These conditions can be met with $\gamma \sim \delta^{-1/2}$; this leads to a type-(ii) error of order $\delta^{-1}\varepsilon^3$ at each time step, accumulating to $\delta^{-1}(T-t)\varepsilon$ after summation over all steps.

(iii) If $\delta^{-1}\varepsilon^2 \le 1/2$ then the increment in the value of D_k from one step to the next is at most of order $\delta^{-1}\varepsilon$. This gives a substitute for (3.14), namely $|D_{i_n} - D_{j_n}| \le C\delta^{-1}\varepsilon(j_n - i_n)$.

(iv) Since each D_k is at most of order $\delta^{-1/2}$, a single term $\varepsilon^2 D_{K+1}$ is estimated by $\delta^{-1/2}\varepsilon^2$.

(v) Since $u(t-\delta,\xi,\eta) - u(t,\xi,\eta) = \int_{t-\delta}^{t} u_s(s,\xi,\eta)\,ds$, the value at $t = T$ is of order $\delta^{1/2}$, and the value at an earlier time t is not larger.

Arguing as for Theorems 3.1 and 3.2 and using these estimates, one finds that

$$\left|u^\varepsilon(t,m,\xi,\eta) - u(t,\xi,\eta)\right| \le C\big(\delta^{-1}(T-t)\varepsilon + \delta^{-1/2}\varepsilon^2 + \delta^{1/2}\big)$$

provided $\delta^{-1}\varepsilon^2 \le 1/2$. To make the first and last terms similar it is natural to take $\delta \sim (T-t)^{2/3}\varepsilon^{2/3}$. (This is consistent with $\delta^{-1}\varepsilon^2 \le 1$ since $(T-t) \ge \varepsilon^2$.) The middle term is then of order $(T-t)^{-1/3}\varepsilon^{5/3}$, which is dominated by the other two; thus we conclude that $|u^\varepsilon - u| \le C(T-t)^{1/3}\varepsilon^{1/3}$.

The preceding estimates are too crude, because they estimate the errors at *every* time step using uniform bounds for the derivatives of $\tilde{u}$, which are overly pessimistic when $t \ll T$. To do better, we should use a k-dependent estimate for the errors at time t_k. Since $\tilde{u}(t,\xi,\eta) = u(t-\delta,\xi,\eta)$, the derivatives of $\tilde{u}$ are estimated at time t_k by (3.29) with $T-t$ replaced by $T - t_k + \delta$. For errors that accumulate over many time steps we must sum the resulting series. This is easily done by comparison to a suitable integral: as t_k runs from t to T in increments of ε^2, $\tau_k = T - t_k + \delta$ runs from $T - t + \delta$ to δ, and a sum of the form $\sum_{k=0}^{N} \tau_k^{-a}\varepsilon^2$ is essentially a Riemann sum for $\int_\delta^{T-t+\delta} s^{-a}\,ds$; as a result we have

$$\sum_{k=0}^{N} \tau_k^{-a} \le \begin{cases} C_a\varepsilon^{-2}\delta^{1-a} & \text{when } a > 1, \text{ and} \\ C\varepsilon^{-2}|\log\delta| & \text{when } a = 1. \end{cases} \tag{3.30}$$

We now review the "error terms" from this perspective:

(i) As in the previous calculation, we write $\tilde{u}(t_{k+1},\xi_{k+1},\eta_{k+1}) - \tilde{u}(t_k,\xi_k,\eta_k)$ as

$$[\tilde{u}(t_{k+1},\xi_{k+1},\eta_{k+1}) - \tilde{u}(t_{k+1},\xi_k,\eta_k)] + [\tilde{u}(t_{k+1},\xi_k,\eta_k) - \tilde{u}(t_k,\xi_k,\eta_k)].$$

When we estimate the first term by 2nd order Taylor expansion in space and the second by 1st order Taylor expansion in time, we introduce an error of order

$$\tau_{k+1}^{-1}\varepsilon^3 + \tau_{k+1}^{-3/2}\varepsilon^4$$

by (3.29). Using (3.30) and assuming $\delta^{-1}\varepsilon^2 \le 1/2$, we see that these errors sum to at most a constant times $|\log\delta|\varepsilon + \varepsilon$.

(ii) The type-(ii) error in the upper-bound argument is controlled at step k by $|\tilde{u}_{\xi\xi}|\varepsilon^3 \le \tau_k^{-1/2}\varepsilon^3$. The situation is slightly worse in the lower bound argument, since the constant γ in the market's Case 2 scales like $\tau_k^{-1/2}$ at step

k, leading to a type-(ii) error of order $\tau_k^{-1}\varepsilon^3$. By (3.30), these errors sum to a term of order at most $|\log\delta|\varepsilon$.

(iii) The increment in the value of D_k from one step to the next is estimated by $\varepsilon^2|\partial_t D| + \varepsilon|\partial_\xi D|$. At step k this is at most $\varepsilon^2\tau_k^{-3/2} + \varepsilon\tau_k^{-1}$. Using (3.30) and assuming $\delta^{-1}\varepsilon^2 \leq 1/2$, we see that the sum of these errors is at most of order $|\log\delta|\varepsilon + \varepsilon$.

(iv) At time k, a single term $\varepsilon^2 D_k$ is estimated by $\tau_k^{-1/2}\varepsilon^2$. Since $\tau_k \geq \delta$, when $\delta^{-1}\varepsilon^2 \leq 1/2$ this term is of order ε.

(v) Our previous estimate of the type-(v) error remains adequate for our present purpose: it is at most $\delta^{1/2}$.

Thus: both the upper-bound and lower-bound arguments, applied using $\tilde{u}$ rather than u, give

$$|u^\varepsilon(t,m,\xi,\eta) - u(t,\xi,\eta)| \leq C\big(\varepsilon|\log\delta| + \varepsilon + \sqrt{\delta}\big)$$

provided $\delta^{-1}\varepsilon^2 \leq 1/2$. Choosing $\delta = 2\varepsilon^2$, we obtain the desired estimate (3.28). □

We note that for the classic final-time data our estimate (3.28) is independent of time, while in Theorems 3.1 and 3.2 it was proportional to $(T-t)$. This difference arises because for the classic final-time data, u_η is constant and all the ξ derivatives of u of order 2 or more decay to 0 as $T-t \to \infty$ (note that u_ξ solves the linear heat equation with final time data $\frac{1}{2}\operatorname{sgn}\xi$). In the more general setting of Theorems 3.1–3.2 the derivatives driving the error terms are controlled, but it is not clear that they tend to 0 as $T-t \to \infty$.

4 The linear programs

We have thus far focused on experts that use only the most recent market move (the case $d = 1$). We identified a strategy for the investor that makes her indifferent with respect to the cycles on the $d = 1$ graph (Section 2.4); then we used that strategy to prove upper and lower bounds that match at leading order (Theorems 3.1 and 3.2).

The situation is similar for experts that use up to four recent market moves (the cases $d = 2,3,4$): there is a strategy for the investor that achieves indifference with respect to the cycles on the relevant graph (see Section 4.4), and it leads to upper and lower bounds that match at leading order (see Remark 5.7 in Section 5.2).

For experts that use more history ($d \geq 5$) we do not know whether indifference is achievable. Our methods still lead to upper and lower bounds, but it is no longer clear that they match at leading order.[1] The upper bound is associated with a strategy for the investor that minimizes the maximum rate at which regret accumulates,

[1] As already mentioned in Sections 1 and 2.6, work subsequent to the present paper has shown that indifference is achievable for any d [5]. As a consequence, our upper and lower bounds do indeed match at leading order.

among all possible cycles on the graph. The lower bound is associated with a strategy for the market that maximizes the minimum rate at which regret accumulates, among all possible cycles on the graph. The identification of these strategies is the focus of this section.

In the proofs of Theorems 3.1 and 3.2, most of the work involved considering how the solution u of our PDE changed along a well-chosen path $(t_0, m_0, \xi_0, \eta_0)$, $(t_1, m_1, \xi_1, \eta_1), \ldots$ Within the special class of strategies described by (2.17) and (2.18), we showed in Section 2.4 that the increments are

$$u(t_{k+1}, \xi_{k+1}, \eta_{k+1}) - u(t_k, \xi_k, \eta_k) = \varepsilon^2 L(t_k, m_k, b_k, \xi_k, \eta_k, f^{\#}_{t_k, m_k}) + O(\varepsilon^3)$$

where

$$L(t, m, b, \xi, \eta, f^{\#}_{t,m}) = u_t + (q(m) - r(m))^2 D - 2bu_\eta f^{\#}_{t,m}$$

(see (2.13) – (2.23)); here we use our usual convention $D = \frac{1}{2}\langle D^2 u \cdot \frac{\nabla^\perp u}{u_\eta}, \frac{\nabla^\perp u}{u_\eta}\rangle$, and b_k is determined by the relation $(m_k)_{b_k} = m_{k+1}$. Since u represents our estimate of the investor's worst-case regret, we think of $\varepsilon^2 L$ as the *increment of regret*.

Our linear programs are concerned with the average rate at which regret accumulates, as the state m traverses various cycles. In formulating them, we will treat u_t, D, and u_η as constants, ignoring the fact that they are functions evaluated at different points in space-time. This seems reasonable, since the state changes at every step while the location in space-time changes slowly (by increments of order ε in space and ε^2 in time). Our rigorous bounds, presented in Section 5, will of course take into account the fact that u_t, D, and u_η are not really constant.

As an introduction to the linear programs, it is convenient to revisit the case $d = 1$. Its graph, shown in Figure 2.2, has three cycles: 00, 11, and 010. The condition that for each cycle, the average rate at which regret accumulates is at most $\varepsilon^2 R$ is therefore

$$\begin{aligned} u_t + (q(0) - r(0))^2 D + 2u_\eta f^{\#}_0 &\leq R \\ u_t + (q(1) - r(1))^2 D - 2u_\eta f^{\#}_1 &\leq R \\ 2u_t + [(q(0) - r(0))^2 + (q(1) - r(1))^2] D - 2u_\eta f^{\#}_0 + 2u_\eta f^{\#}_1 &\leq 2R. \end{aligned} \tag{4.1}$$

The investor wants to make R as small as possible. Since u_t is being treated as a constant, we can move it to the right hand side. Setting $M = (R - u_t)/D$ and

$$\beta_m = \frac{2u_\eta}{D} f^{\#}_m, \tag{4.2}$$

and assuming $D > 0$, we see that (4.1) is equivalent to

$$\begin{aligned} (q(0) - r(0))^2 + \beta_0 &\leq M \\ (q(1) - r(1))^2 - \beta_1 &\leq M \\ [(q(0) - r(0))^2 + (q(1) - r(1))^2] - \beta_0 + \beta_1 &\leq 2M. \end{aligned} \tag{4.3}$$

(In practice, the parameter D will come from the solution of our PDE, and it is nonnegative by (2.33). We need not be concerned about the exceptional case $D = 0$, which is handled in Section 5 by taking $f_m^\# = 0$.) The *investor's linear program* for $d = 1$ is thus to find β_0, β_1, and M that minimize M subject to (4.3).

A similar discussion applies for the market. The condition that for each simple cycle, the average rate at which regret accumulates is at least $\varepsilon^2 R$ is

$$\begin{aligned} u_t + (q(0)-r(0))^2 D + 2u_\eta f_0^\# &\geq R \\ u_t + (q(1)-r(1))^2 D - 2u_\eta f_1^\# &\geq R \\ 2u_t + [(q(0)-r(0))^2 + (q(1)-r(1))^2]D - 2u_\eta f_0^\# + 2u_\eta f_1^\# &\geq 2R. \end{aligned}$$

Changing variables as before, this is equivalent to

$$\begin{aligned} (q(0)-r(0))^2 + \beta_0 &\geq M \\ (q(1)-r(1))^2 - \beta_1 &\geq M \\ [(q(0)-r(0))^2 + (q(1)-r(1))^2] - \beta_0 + \beta_1 &\geq 2M. \end{aligned} \tag{4.4}$$

The *market's linear program* for $d = 1$ is thus to find β_0, β_1, and M that maximize M subject to (4.4).

The following properties of these linear programs are immediately evident:

(i) If β_0, β_1, and M are admissible for (4.3), then adding the constraints gives

$$M \geq \frac{(q(0)-r(0))^2 + (q(1)-r(1))^2}{2}.$$

(ii) If β_0, β_1, and M are admissible for (4.4), then adding the constraints gives

$$M \leq \frac{(q(0)-r(0))^2 + (q(1)-r(1))^2}{2}.$$

(iii) In view of (i) and (ii), the values of the two linear programs coincide when there is a choice of β_0 and β_1 that achieves indifference, in the sense that the LHS of each of the inequalities in (4.3) (or equivalently, each of the inequalities in (4.4)) takes the same value; moreover the common value is then $\frac{1}{2}[(q(0)-r(0))^2 + (q(1)-r(1))^2]$.

(iv) The values of the two linear programs do indeed coincide, since the choice $\beta_0 = \beta_1 = \frac{1}{2}[(q(1)-r(1))^2 - (q(0)-r(0))^2]$ achieves indifference.

We shall show in Sections 4.1 and 4.2 that analogues of (i)–(iii) hold for any d. We do not know whether the analogue of (iv) holds for any d, however we shall show in Section 4.4 that indifference is achievable when $d = 2$, 3, or 4.

4.1 The investor's linear program

The investor's linear program for $d = 1$, presented above as the minimization of M subject to (4.3), can be written as

$$\min M \quad \text{such that}$$
$$A \begin{pmatrix} -\beta_0 \\ -\beta_1 \\ M \end{pmatrix} \leq g \tag{4.5}$$

with

$$A = \begin{pmatrix} -1 & 0 & -1 \\ 0 & 1 & -1 \\ 1 & -1 & -2 \end{pmatrix} \quad \text{and} \quad g = -\begin{pmatrix} (q(0)-r(0))^2 \\ (q(1)-r(1))^2 \\ (q(0)-r(0))^2 + (q(1)-r(1))^2 \end{pmatrix}.$$

Notice that each row of A corresponds to a simple cycle; the final entry in a row is minus the length of the cycle, while the row's other elements reflect the details of the cycle. As for g: each row is minus the sum of $(q(m)-r(m))^2$, as m ranges over the vertices participating in the associated cycle.

The investor's linear program is similar for any value of d: it always has the form

$$\min M \quad \text{such that}$$
$$A \begin{pmatrix} -\beta_0 \\ \cdots \\ -\beta_{2^d-1} \\ M \end{pmatrix} \leq g \tag{4.6}$$

where each row of A corresponds to a simple cycle on the underlying graph (the d-dimensional de Bruijn graph on 2 symbols, see Section 2.2). If the graph has l simple cycles then A is an $l \times (2^d + 1)$ matrix. Its first 2^d columns are in correspondence with the vertices: for $0 \leq m \leq 2^d - 1$, the $(m+1)$th column reports, for each cycle, whether the cycle includes vertex m, and if so then whether the cycle leaves it by a $+$ edge or a $-$ edge. The last column of A is, up to sign, the number of edges in the cycle. The elements of g are, up to sign, the sum of $(q(m)-r(m))^2$ as m ranges over the vertices in the given cycle. More explicitly: the first 2^d columns of the matrix A are determined by

$$\begin{aligned} a_{i,m+1} &= 1 && \text{if cycle } i \text{ contains an edge from } m \text{ to } m_+; \\ a_{i,m+1} &= -1 && \text{if cycle } i \text{ contains an edge from } m \text{ to } m_-; \\ a_{i,m+1} &= 0 && \text{if cycle } i \text{ does not pass through vertex } m; \end{aligned} \tag{4.7}$$

the last column of A has entries

$$a_{i,2^d+1} = -|s_i| \text{ where } |s_i| \text{ is the number of edges in cycle } i; \tag{4.8}$$

and the elements of g are

$$g_i = -\sum_{\text{vertices}\, m\, \text{in}\, \text{cycle}\, i} (q(m)-r(m))^2. \tag{4.9}$$

While A happens to be square when $d=1$, it is usually rectangular; for example, when $d=2$ there are 6 simple cycles and 4 vertices, so A is 6×5.

The interpretation of the investor's linear program (4.6) is the same for any d as it was for $d=1$: if β_m ($0\le m\le 2^d-1$) and M are admissible then the choice $f^{\#}_m = \frac{D}{2u_\eta}\beta_m$ assures that the average rate at which regret accumulates over any simple cycle is at most $\varepsilon^2 R$ with $R = u_t + MD$.

Lemma 4.1. *The investor's linear program* (4.6) *has the following properties:*

(a) The feasible set is nonempty; moreover, every feasible point has the property that

$$M \ge \frac{\sum_{m=0}^{2^d-1}(q(m)-r(m))^2}{2^d}. \tag{4.10}$$

(b) The optimum is achieved.

(c) If there is a strategy that achieves indifference (i.e. if there is a feasible point for which the constraints all hold with equality) then the optimal value of the investor's linear program is $2^{-d}\sum_{m=0}^{2^d-1}(q(m)-r(m))^2$.

Proof. The existence of a feasible point is easy: if we take $\beta_m=0$ for all m, then the constraints are satisfied provided that for each cycle i,

$$M \ge \frac{1}{|s_i|}\sum_{\text{vertices}\, m\, \text{in}\, \text{cycle}\, i} (q(m)-r(m))^2$$

where $|s_i|$ is the number of edges in cycle i.

To complete the proof of (a) we must establish the lower bound (4.10). We use here the fact that our de Bruijn graph is *Eulerian*, i.e. there exists a closed walk on the graph that traverses each edge exactly once (see e.g. [24]). Consider a decomposition of this walk as a union of simple cycles. Clearly no cycle can appear more than once; and in the union of all the cycles that are used, each vertex m appears twice – once in connection with the edge from m to m_+, and once in connection with the edge from m to m_-. When we add the constraints corresponding the cycles that are used, the terms involving β_m cancel (since each β_m appears twice, with opposite signs – once due to the edge from m to m_+, and once due to the edge from m to m_-). The sum of these constraints thus reduces to

$$2\sum_{m=0}^{2^d-1}(q(m)-r(m))^2 \le 2^{d+1}M,$$

which is equivalent to (4.10).

Assertion (b) follows immediately from (a), using the general result from linear programming that if an LP of the form (4.6) is feasible and bounded below then the optimal value is achieved. Alternatively (and more constructively), one can see (b) by considering the simplex method (with a suitable scheme to prevent cycling). At each step the simplex method moves in a direction that improves the objective. A move to infinity cannot occur, since the objective is bounded below. Therefore the simplex method moves from vertex to vertex, terminating at one that is optimal.

Assertion (c) is clear from the proof of (4.10). Indeed, we proved that result by adding the inequalities associated with the cycles of an Eulerian circuit. A strategy that achieves indifference turns those inequalities into equalities, so equality must also hold in (4.10). □

4.2 The market's linear program

Comparing (4.3) and (4.4), we see that for $d = 1$ the market's linear program is obtained from the investor's by changing the direction of the inequalities and changing the objective from $\min M$ to $\max M$. The situation is the same for any d: the market's linear program is

$$\max M \quad \text{such that}$$

$$A \begin{pmatrix} -\beta_0 \\ \cdots \\ -\beta_{2^d-1} \\ M \end{pmatrix} \geq g \tag{4.11}$$

where A and g are still defined by (4.7)–(4.9). Its interpretation is analogous to the case $d = 1$: if β_m $(0 \leq m \leq 2^d - 1)$ and M are admissible for (4.11) then the choice $f_m^\# = \frac{D}{2u_\eta}\beta_m$ assures that the average rate at which regret accumulates over any simple cycle is at least $\varepsilon^2 R$ with $R = u_t + MD$.

The following analogue of Lemma 4.1 is proved using the same arguments.

Lemma 4.2. *The market's linear program* (4.11) *has the following properties:*

(a) The feasible set is nonempty; moreover, every feasible point has the property that

$$M \leq \frac{\sum_{m=0}^{2^d-1}(q(m) - r(m))^2}{2^d}. \tag{4.12}$$

(b) The optimum is achieved.

(c) If there is a strategy that achieves indifference (i.e. if there is a feasible point for which the constraints all hold with equality) then the optimal value of the market's linear program is $2^{-d}\sum_{m=0}^{2^d-1}(q(m) - r(m))^2$.

Corollary 4.3. *The optimal value of the market's linear program is less than or equal to that of the investor's linear program. If there is a strategy that achieves indifference then the optimal values are equal.*

Proof. The first assertion is clear by combining part (a) of Lemma 4.1 with part (a) of Lemma 4.2. The second assertion follows from part (c) of each Lemma. □

4.3 Why are the LP's not convex duals?

In finite dimensions, zero-sum two-person games using mixed strategies can be analyzed using the duality theory of linear programming, leading to a pair of dual linear programs – one for the "minimizing" player, the other for the "maximizing" player. We emphasize that this is *not* the relationship between our linear programs (4.6) and (4.11).

It is natural to ask whether duality might somehow be relevant to our problem. The answer is this:

(a) If the rules of the game *required* the investor to choose $f_m = f_m^* + \varepsilon f_m^\#$, as envisioned by the heuristic argument presented in Section 2.4, then duality would be relevant (at least formally). Indeed, as we shall show in a moment, the dual of the investor's linear program provides a convex combination of simple cycles that makes the market indifferent to the investor's choice of $f_m^\#$ and provides (at least formally) a lower bound that matches our upper bound.

(b) Alas, the rules of our game do not require the investor to behave this way (and we have not proved that the optimal choice has this form). In our lower bound for $d \geq 2$, presented in Section 5, the idea of the proof is similar to what we did in Section 3.2 for $d = 1$. The market considers a *particular* investor strategy of the form $f_m = f_m^* + \varepsilon f_m^\#$, and chooses the stock movement $b = \pm 1$ based on the sign of $\hat{f} - f_m$, where $\hat{f}$ is the investor's choice. Roughly speaking, the market uses its discretion over the stock price evolution to "force" the investor to use the particular strategy (much as we did for $d = 1$ in (3.15)–(3.16)). In doing so, the market loses all control over the cycle decomposition of the resulting walk; therefore its worst-case estimate involves the minimum rate at which regret accumulates (the minimization being over all simple cycles). Our market's linear program, which chooses $f_m^\#$ to maximize this, seems quite different from the dual of the investor's linear program.

The rest of this subsection explains point (a). The investor's linear program chooses $f_m^\#$ to

$$\min_{f_m^\#} \max_{\text{cycles}} \left(\text{average rate at which regret accumulates}\right).$$

Normalizing as we did in Section 4.1, and indexing the simple cycles by $i = 1, \dots, n$, this amounts to

$$\min_{\beta_m} \max_{1 \leq i \leq n} \frac{1}{|s_i|} \sum_{\substack{\text{vertices}\, m \\ \text{on cycle}\, i}} \left[(q(m) - r(m))^2 - b_{i,m}\beta_m\right] \tag{4.13}$$

where $|s_i|$ is the number of edges in cycle i, and $b_{i,m} = +1$ if cycle i leaves vertex m by the edge from m to m_+, while $b_{i,m} = -1$ if cycle i leaves vertex m by the edge from m to m_-. (Note that $b_{i,m} = a_{i,m+1}$, according to (4.7).) The inner max is not changed if we permit "mixed strategies," i.e. a probability distribution over the possible cycles. So the investor's linear program solves

$$\min_{\beta_m} \max_{\substack{p_1+\cdots+p_n=1 \\ p_i \geq 0}} \sum_{i=1}^{n} p_i \frac{1}{|s_i|} \sum_{\substack{\text{vertices}\, m \\ \text{on cycle}\, i}} [(q(m)-r(m))^2 - b_{i,m}\beta_m].$$

The dual is obtained by switching the min and the max, then evaluating the inner minimization. Since the variables β_m are unbounded,

$$\min_{\beta_m} \sum_{i=1}^{n} p_i \frac{1}{|s_i|} \sum_{\substack{\text{vertices}\, m \\ \text{on cycle}\, i}} [(q(m)-r(m))^2 - b_{i,m}\beta_m] = \tag{4.14}$$

$$\begin{cases} -\infty & \text{if the variables } \beta_m \text{ don't cancel out} \\ \sum_{i=1}^{n} p_i \frac{1}{|s_i|} \sum_{\substack{\text{vertices}\, m \\ \text{on cycle}\, i}} (q(m)-r(m))^2 & \text{if they do.} \end{cases}$$

Thus: the dual of the investor's linear program involves mixed strategies over the simple cycles that make the market insensitive to the investor's choice of β_m (or equivalently, $f_m^{\#}$); the optimal mixed strategy is the one that maximizes the value of (4.14). Since min-max equals max-min in this setting, the value achieved by the optimal mixed strategy is the same as that of the investor's linear program.

Does the market have access to such mixed strategies? We suppose so, since the market can choose any walk it likes.

As noted earlier, throughout this section we have ignored the fact that u_t, u_η, and $D = \frac{1}{2}\langle D^2 u \cdot \frac{\nabla^\perp u}{u_\eta}, \frac{\nabla^\perp u}{u_\eta}\rangle$ are not really constant. Our upper and lower bounds for $d \geq 2$, presented in Section 5, deal with this issue – but only for linear programs discussed in Sections 4.1 and 4.2. We have not attempted to address this issue for the dual of the investor's linear program.

4.4 Coalescence of the optimal values for $d \leq 4$

Recall from Corollary 4.3 that the investor's linear program and the market's linear program have the same optimal value if there is a choice of $\{\beta_m\}_{m=0}^{2^d-1}$ that achieves indifference. We have already shown the existence of such a choice when $d = 1$. This subsection shows the existence of such a choice when $d = 2$, 3, or 4. Alas, since our analysis is by brute force, it offers little insight about whether indifference is also achievable for $d \geq 5$.

THE CASE $d = 2$. The graph for $d = 2$ was shown in Figure 2.1. It is easy to see that there are six simple cycles, and we enumerated them in Section 2.2. Introducing the notation

$$\gamma_m = (q(m)-r(m))^2,$$

indifference requires that $\beta_0, \ldots, \beta_3$ satisfy

$$\begin{aligned}
\gamma_0 + \beta_0 &= M \\
\gamma_3 - \beta_3 &= M \\
\gamma_1 + \gamma_2 + \beta_1 - \beta_2 &= 2M \\
\gamma_0 + \gamma_1 + \gamma_2 - \beta_0 + \beta_1 + \beta_2 &= 3M \\
\gamma_1 + \gamma_2 + \gamma_3 - \beta_1 - \beta_2 + \beta_3 &= 3M \\
\gamma_0 + \gamma_1 + \gamma_2 + \gamma_3 - \beta_0 - \beta_1 + \beta_2 + \beta_3 &= 4M
\end{aligned} \tag{4.15}$$

and we know the value of M from Lemma 4.1(c). Elementary manipulation reveals that indifference is achieved when

$$\begin{aligned}
M &= \frac{\gamma_0 + \gamma_1 + \gamma_2 + \gamma_3}{4} \\
\beta_0 &= \beta_2 = M - \gamma_0 \\
\beta_1 &= \beta_3 = \gamma_3 - M.
\end{aligned}$$

THE CASE $d = 3$. It is straightforward to draw the $d = 3$ graph (see e.g. Figure 1 of [12]). It is possible (though laborious) to check by hand that there are 19 simple cycles, and to write down the $d = 3$ analogue of (4.15) (a system of 19 linear equations in the 9 unknowns $\gamma_0, \ldots, \gamma_7$ and M). Solving that system, one finds that indifference is achieved for $d = 3$ when

$$\begin{aligned}
M &= \frac{\sum_{m=0}^{7} \gamma_i}{8} \\
\beta_0 &= \beta_4 = M - \gamma_0 \\
\beta_1 &= \beta_5 = -M + \frac{-\gamma_2 + \gamma_3 + \gamma_6 + \gamma_7}{2} \\
\beta_2 &= \beta_6 = M - \frac{\gamma_0 + \gamma_1 + \gamma_4 - \gamma_5}{2} \\
\beta_3 &= \beta_7 = \gamma_7 - M.
\end{aligned}$$

THE CASE $d = 4$. The $d = 4$ graph is not planar, but it is still not difficult to visualize (see e.g. Figure 1 of [12]). We wrote a Matlab program to enumerate the simple cycles; it found 179 of them. Using the results, we looked numerically for linear combinations of $\{\gamma_m\}_{m=0}^{15}$ that achieve indifference. This led to the conclusion that indifference is achieved for $d = 4$ when

$$
\begin{aligned}
M &= \frac{\sum_{i=0}^{15} \gamma_i}{16} \\
\beta_0 = \beta_8 &= M - \gamma_0 \\
\beta_1 = \beta_9 &= -M + \frac{-2\gamma_2 + 2\gamma_3 - \gamma_4 - \gamma_5 + \gamma_6 + \gamma_7 + \gamma_{12} + \gamma_{13} + \gamma_{14} + \gamma_{15}}{4} \\
\beta_2 = \beta_{10} &= M - \frac{\gamma_0 + \gamma_1 + \gamma_2 + \gamma_3 + 2\gamma_4 - 2\gamma_5 + \gamma_8 + \gamma_9 - \gamma_{10} - \gamma_{11}}{4} \\
\beta_3 = \beta_{11} &= -M + \frac{-2\gamma_6 + 2\gamma_7 + 2\gamma_{14} + 2\gamma_{15}}{4} \\
\beta_4 = \beta_{12} &= M - \frac{2\gamma_0 + 2\gamma_1 + 2\gamma_8 - 2\gamma_9}{4} \\
\beta_5 = \beta_{13} &= -M + \frac{-\gamma_4 - \gamma_5 + \gamma_6 + \gamma_7 - 2\gamma_{10} + 2\gamma_{11} + \gamma_{12} + \gamma_{13} + \gamma_{14} + \gamma_{15}}{4} \\
\beta_6 = \beta_{14} &= M - \frac{\gamma_0 + \gamma_1 + \gamma_2 + \gamma_3 + \gamma_8 + \gamma_9 - \gamma_{10} - \gamma_{11} + 2\gamma_{12} - 2\gamma_{13}}{4} \\
\beta_7 = \beta_{15} &= -M + g_{15}.
\end{aligned}
$$

We wonder whether indifference might be achievable for any d. Alas, the brute force approach we have used for $d \le 4$ does not provide much guidance. (It does, however, provide some hints; in particular, the choices achieving indifference for $d = 2, 3, 4$ all have the symmetry that $\beta_m = \beta_{m+2^{d-1}}$ for $0 \le m \le 2^{d-1} - 1$. This means that β_m actually depends only on the most recent $d-1$ market moves.)

5 Upper and lower bounds for general *d*

This section provides our fully rigorous upper and lower bounds for the general case, when our two experts use d days of data. The arguments are in many ways parallel to those presented in Section 3 where we considered the case $d = 1$. The main differences are:

(i) For $d = 1$ there is a choice of β_m that achieves indifference (in the sense of Lemmas 4.1(c) and 4.2(c)), and the proofs of both the upper and lower bounds used the associated value of $f_m^\#$ (this was the logic behind (3.8)). For general d, our upper bound uses the investor's linear program to determine $f_m^\#$, whereas our lower bound uses the market's linear program.

(ii) In formulating the linear programs at the beginning of Section 4 we ignored the fact that u_t, u_η, and $D = \frac{1}{2}\langle D^2 u \cdot \frac{\nabla^\perp u}{u_\eta}, \frac{\nabla^\perp u}{u_\eta}\rangle$ are functions of x and t. For $d = 1$, the errors associated with their variability were relatively easy to control (see (3.10)–(3.14)). While that argument did not make explicit use of the graph, it basically relied on the simple cycle structure of the $d = 1$ graph. The corresponding arguments for general d are more complicated since the cycle structure of the graph is less controlled.

In connection with the second point, we shall need the following refinement of Lemma 2.3.

Lemma 5.1. *When a closed walk on a directed graph is decomposed into a union of simple cycles by the argument used to prove Lemma 2.3, the decomposition has the following property: for any simple cycle s that appears in the decomposition more than once, each instance of s is completed before the next begins (as one traverses the walk from beginning to end).*

Proof. Fixing the graph under consideration, we use complete induction on the length of the walk. The shortest possible length of a closed walk is the length of the shortest simple cycle; for walks of this length the result is trivial.

For the inductive step, we must show that if the result is true for closed walks of length up to $N-1$ then it is also true for closed walks of length N. So consider a closed walk of length N whose vertices (in order) are $a_0, a_1, \dots, a_N$ with $a_N = a_0$. Consider (as in the proof of Lemma 2.3) the beginning of the walk, up to the first time a vertex is repeated:

$$a_0 \dots a_i \dots a_j \dots a_N \quad \text{where } a_j = a_i \text{ is the first repetition.}$$

Then $s = a_i \dots a_j$ is a simple cycle, and it is the first cycle in the decomposition of the walk. The rest of the decomposition is obtained by considering the walk that remains after removal of this cycle, namely $a_0 \dots a_{i-1}, a_j, a_{j+1} \dots a_N$ and applying the same argument (repeatedly). The inductive hypothesis applies to this shortened walk.

For simple cycles other than s, the inductive hypotheses assures us that each instance is complete before the next begins, as one traverses the shortened walk. Therefore the same is true of the original walk.

As for the simple cycle s: since $a_i = a_j$ was the first repeated node, none of the vertices $a_0 \dots a_{i-1}$ participate in s. Therefore in the decomposition of the shortened walk, no instances of s are begun during the initial segment $a_0 \dots a_{i-1}$. So the first instance of s is completed before the others begin; and by the inductive hypothesis, each other instance of s is completed before the next begins. ☐

5.1 The upper bound, when φ is regular

Let $u^\varepsilon(t, m, \xi, \eta)$ be defined by the dynamic programming principle (2.10) with final value $u^\varepsilon(T, m, \xi, \eta) = \varphi(\xi, \eta)$ (it is defined only for times t such that $(T - t)/\varepsilon^2$ is an integer); and let $u(t, \xi, \eta)$ be the solution of the PDE

$$\begin{aligned} u_t + \frac{1}{2} C_d^{\#} \langle D^2 u \frac{\nabla^\perp u}{u_\eta}, \frac{\nabla^\perp u}{u_\eta} \rangle &= 0, \\ u(T, \xi, \eta) &= \varphi(\xi, \eta), \end{aligned} \tag{5.1}$$

where $C_d^{\#}$ is the optimal value of the investor's linear program (4.6). Our goal is to prove

Theorem 5.2. *Let $d \geq 1$ be an integer and assume the solution u of* (5.1) *satisfies* (2.30)–(2.33)*. Then there is a constant C (independent of ε, t, and T) such that*

$$u^{\varepsilon}(t,m,\xi,\eta) \leq u(t,\xi,\eta) + C[(T-t)+\varepsilon]\varepsilon \tag{5.2}$$

for $t < T$, $\xi \in \mathbb{R}$, $\eta \in \mathbb{R}$, and $m \in \{0,1\}^d$, provided that ε is small enough and t is such that $N = (T-t)/\varepsilon^2$ is an integer.

Proof. Our overall strategy is similar to the proof of Theorem 3.1: to estimate $u^{\varepsilon}(t_0,m_0,\xi_0,\eta_0)$, we shall define a sequence (t_k,m_k,ξ_k,η_k) along which u^{ε} is monotone

$$u^{\varepsilon}(t_0,m_0,\xi_0,\eta_0) \leq u^{\varepsilon}(t_1,m_1,\xi_1,\eta_1) \leq \cdots \leq u^{\varepsilon}(t_N,m_N,\xi_N,\eta_N) \tag{5.3}$$

with $t_N = T$, so that

$$u^{\varepsilon}(t_N,m_N,\xi_N,\eta_N) = \varphi(\xi_N,\eta_N) = u(t_N,\xi_N,\eta_N). \tag{5.4}$$

Then we'll show that

$$u(t_N,\xi_N,\eta_N) - u(t_0,\xi_0,\eta_0) \leq C[(T-t_0)+\varepsilon]\varepsilon. \tag{5.5}$$

These estimates lead immediately to (5.2).

Our choice of $\{(t_k,m_k,\xi_k,\eta_k)\}_{k=1}^{N}$ is entirely parallel to what was done for Theorem 3.1. It suffices to explain the choice of (t_1,m_1,ξ_1,η_1) (then the rest of the sequence is determined similarly, step by step). Recall that the dynamic programming principle was written compactly in equations (3.5)–(3.6), and it gave

$$u^{\varepsilon}(t_0,m_0,\xi_0,\eta_0) \leq \max_{b_0=\pm 1} u^{\varepsilon}(t_0+\varepsilon^2, m_{b_0}, \xi_0+\varepsilon b_0 v^1, \eta_0+\varepsilon b_0 v^2) \tag{5.6}$$

where (v^1,v^2) are determined by the investor's choice of f via (3.5). Our choice of $f = f^*_{t_0,m_0} + \varepsilon f^{\#}_{t_0,m_0}$ is guided by the heuristic discussion in Section 2.4 (which determines $f^*_{t_0,m_0}$) combined with the investor's linear program (which determines $f^{\#}_{t_0,m_0}$ via (4.2)):

$$\begin{aligned} f_{t_0,m_0} &= f^*_{t_0,m} + \varepsilon f^{\#}_{t_0,m_0}, \text{ with} \\ f^*_{t_0,m_0} &= \frac{(q(m_0)-r(m_0))u_{\xi} + (q(m_0)+r(m_0))u_{\eta}}{2u_{\eta}} \text{ and} \\ f^{\#}_{t_0,m_0} &= \frac{D}{2u_{\eta}}\beta_{m_0}. \end{aligned} \tag{5.7}$$

Here u_{ξ}, u_{η}, and $D = \frac{1}{2}\langle D^2u \cdot \frac{\nabla^{\perp}u}{u_{\eta}}, \frac{\nabla^{\perp}u}{u_{\eta}}\rangle$ are evaluated at (t_0,ξ_0,η_0), while β_{m_0} comes from the solution of the investor's linear program (4.6). (If the linear program has more than one solution, we choose one and use it throughout the proof.) The proper choice of (t_1,m_1,ξ_1,η_1) is now clear: taking b_0 to achieve the max on the RHS of (5.6), the choice $t_1 = t_0+\varepsilon^2$, $m_1 = (m_0)_{b_0}$, $\xi_1 = \xi_0+\varepsilon b_0 v^1$, $\eta_1 = \eta_0+\varepsilon b_0 v^2$ satisfies the desired inequality $u^{\varepsilon}(t_0,m_0,\xi_0,\eta_0) \leq u^{\varepsilon}(t_1,m_1,\xi_1,\eta_1)$. The rest of the sequence (t_k,m_k,ξ_k,η_k) is determined similarly, up to $k=N$, when $t_N = T$.

We now depart a bit from the proof of Theorem 3.1. The sequence $m_0, m_1, \ldots, m_N$ is a walk on our directed graph, but it is not in general closed (since we cannot expect that $m_N = m_0$). It can, however, be extended to a closed walk by adding some steps at the end. Indeed, as already noted in Section 4 our graph is Eulerian, i.e. there is a closed walk (an Eulerian circuit) that traverses each edge exactly once; the desired extension can be achieved by adding part of an Eulerian circuit, starting from m_N and stopping upon arrival at m_0. We let N' be the final index of the extended walk (so $m_{N'} = m_0$); note that the number of extra steps is bounded by the total number of edges in the graph:

$$N' - N \leq 2^{d+1}. \tag{5.8}$$

The expression $u^{\varepsilon}(t_k, m_k, \xi_k, \eta_k)$ is undefined for $k > N$. However it is convenient to introduce a suitable choice of (t_k, ξ_k, η_k) for $k > N$, so that $u(t_k, \xi_k, \eta_k)$ will make sense. We take the time increments to be trivial: $t_k = t_N = T$ for $k > N$. For the spatial increments we use the same definition as for $k < N$: defining $b_k = \pm 1$ for $k \geq N$ by $(m_k)_{b_k} = m_{k+1}$, we set

$$\xi_{k+1} = \xi_k + \varepsilon b_k v^1, \quad \eta_{k+1} = \eta_k + \varepsilon b_k v^2$$

for $k \geq N$, where (v^1, v^2) are determined by the investor's choice of f via (3.5), the choice of f still being given by (5.7) (with t_0 and m_0 replaced by t_k and m_k, and with u_ξ, u_η, and $D = \frac{1}{2}\langle D^2 u \cdot \frac{\nabla^\perp u}{u_\eta}, \frac{\nabla^\perp u}{u_\eta}\rangle$ evaluated at (t_k, ξ_k, η_k)). This choice has the key feature that when the increment

$$U_k = u(t_{k+1}, \xi_{k+1}, \eta_{k+1}) - u(t_k, \xi_k, \eta_k)$$

is estimated for $k \geq N$ by Taylor expansion, there is no u_t term (since $t_{k+1} = t_k$) and the first-order terms involving the increments of ξ and η vanish (see (2.13)–(2.19)); as a result, we have

$$|U_k| \leq C\varepsilon^2 \quad \text{for } k \geq N. \tag{5.9}$$

As a reminder, we also have a convenient estimate for the increments at $k < N$:

$$U_k = \varepsilon^2 L(t_k, m_k, b_k, \xi_k, \eta_k, f^{\#}_{t_k, m_k}) + O(\varepsilon^3) \tag{5.10}$$

where L is defined by (2.23).

We now start the proof of (5.5). In view of (5.8) and (5.9), it suffices to prove that

$$u(t_{N'}, \xi_{N'}, \eta_{N'}) - u(t_0, \xi_0, \eta_0) \leq C[(T - t_0) + \varepsilon]\varepsilon.$$

Since the walk $m_0, \ldots, m_{N'}$ is closed, it is a union of simple cycles by Lemma 2.3. Now,

$$u(t_{N'}, \xi_{N'}, \eta_{N'}) - u(t_0, \xi_0, \eta_0) = \sum_{k=0}^{N'-1} U_k, \tag{5.11}$$

and each increment is associated with a step of our closed walk, so RHS of (5.11) can be reorganized using the walk's cycle decomposition.

The total contribution of all cycles that involve the extended part of the walk (i.e., cycles that include a node m_k with $k > N$) is of order ε^2, since the number of such cycles is finite (with an estimate depending only on d) and each instance of any cycle makes a contribution of order ε^2. We shall refer to the cycles that don't involve the extended part of the walk as the *remaining cycles*.

The $O(\varepsilon^3)$ error terms in (5.10) don't bother us, since they accumulate to an error of at most $C(T-t)\varepsilon$. Thus to prove (5.5) we need only show that

(5.12) the terms $\varepsilon^2 L(t_k, m_k, b_k, \xi_k, \eta_k, f^{\#}_{t_k,m_k})$ coming from the remaining cycles sum to at most $C(T-t)\varepsilon$.

Let $\{s_i\}_{i=1}^n$ be a list of the simple cycles on the graph; we shall (as usual) write $|s_i|$ for the number of edges in s_i. Suppose that among the remaining cycles, s_i appears σ_i times. We sum $L(t_k, m_k, b_k, \xi_k, \eta_k, f^{\#}_{t_k,m_k})$ over the remaining cycles in stages: first over the edges of an instance α of a cycle s_i, then over the σ_i instances of this cycle, then over all the distinct cycles $s_1, \ldots, s_n$. Thus, the quantity to be estimated is

(5.13)
$$\varepsilon^2 \sum_{\substack{\text{remaining}\\ \text{cycles}}} L(t_k, m_k, b_k, \xi_k, \eta_k, f^{\#}_{t_k,m_k}) = \varepsilon^2 \sum_{i=1}^{n} \sum_{\alpha=1}^{\sigma_i} \sum_{j=1}^{|s_i|} L(t^{\alpha}_{i,j}, m^{\alpha}_{i,j}, b^{\alpha}_{i,j}, \xi^{\alpha}_{i,j}, \eta^{\alpha}_{i,j}, f^{\#}_{t^{\alpha}_{i,j},m^{\alpha}_{i,j}}).$$

Here the time steps have been relabeled using the cycle decomposition: $t^{\alpha}_{i,j}$ is the time when instance α of cycle s_i takes its jth step, $j = 1, \ldots, |s_i|$. Note that by Lemma 5.1, for each i and each $\alpha < \sigma_i$,

$$t^{\alpha}_{i,1} < t^{\alpha}_{i,2} < \ldots < t^{\alpha}_{i,|s_i|} < t^{\alpha+1}_{i,1} < t^{\alpha+1}_{i,2} < \ldots < t^{\alpha+1}_{i,|s_i|}. \tag{5.14}$$

In formulating our linear program we treated u_t, u_η, etc. as being constant, whereas in fact they vary with space and time. To deal with this, we need to compare $L(t^{\alpha}_{i,j}, m^{\alpha}_{i,j}, b^{\alpha}_{i,j}, \xi^{\alpha}_{i,j}, \eta^{\alpha}_{i,j}, f^{\#}_{t^{\alpha}_{i,j},m^{\alpha}_{i,j}})$ with $L(t^{\alpha}_{i,1}, m^{\alpha}_{i,j}, b^{\alpha}_{i,j}, \xi^{\alpha}_{i,1}, \eta^{\alpha}_{i,1}, f^{\#}_{t^{\alpha}_{i,1},m^{\alpha}_{i,j}})$, the latter being the analogue of the former obtained by freezing u_t, u_η, etc to their values at time $t^{\alpha}_{i,1}$ when the αth instance of the ith cycle begins. Combining the definition (2.23) of L with the choice (5.7) of $f^{\#}_{t,m}$, we have

$$L(t, m, b, \xi, \eta, f^{\#}_{t,m}) = u_t + (q(m) - r(m))^2 D - b\beta_m D,$$

in which u_t and $D = \frac{1}{2}\langle D^2 u \cdot \frac{\nabla^\perp u}{u_\eta}, \frac{\nabla^\perp u}{u_\eta}\rangle$ are evaluated at (t, ξ, η). Our hypotheses (2.30)–(2.31) assure that u_t and D are uniformly Lipschitz continuous functions of t, ξ, and η. Since the number of time steps from $t^{\alpha}_{i,1}$ to $t^{\alpha}_{i,j}$ is $(t^{\alpha}_{i,j} - t^{\alpha}_{i,1})/\varepsilon^2$ and the change in (t, ξ, η) is of order ε at each time step,

(5.15)
$$L(t^{\alpha}_{i,j}, m^{\alpha}_{i,j}, b^{\alpha}_{i,j}, \xi^{\alpha}_{i,j}, \eta^{\alpha}_{i,j}, f^{\#}_{t^{\alpha}_{i,j},m^{\alpha}_{i,j}}) = L(t^{\alpha}_{i,1}, m^{\alpha}_{i,j}, b^{\alpha}_{i,j}, \xi^{\alpha}_{i,1}, \eta^{\alpha}_{i,1}, f^{\#}_{t^{\alpha}_{i,1},m^{\alpha}_{i,j}}) + \frac{t^{\alpha}_{i,j} - t^{\alpha}_{i,1}}{\varepsilon^2} O(\varepsilon).$$

Focusing on the terms associated with a particular instance of a particular cycle (thus, fixing α and i), we conclude that

$$
\begin{aligned}
(5.16)\quad & \sum_{j=1}^{|s_i|} L(t_{i,j}^\alpha, m_{i,j}^\alpha, b_{i,j}^\alpha, \xi_{i,j}^\alpha, \eta_{i,j}^\alpha, f^\#_{t_{i,j}^\alpha, m_{i,j}^\alpha}) \\
&= \sum_{j=1}^{|s_i|} L(t_{i,1}^\alpha, m_{i,j}^\alpha, b_{i,j}^\alpha, \xi_{i,1}^\alpha, \eta_{i,1}^\alpha, f^\#_{t_{i,1}^\alpha, m_{i,j}^\alpha}) + \sum_{j=1}^{|s_i|} \varepsilon^{-2}(t_{i,j}^\alpha - t_{i,1}^\alpha) O(\varepsilon) \\
&= |s_i| \left[u_t + D_{i,1}^\alpha \sum_{j=1}^{|s_i|} \frac{1}{|s_i|} \left[(q(m_{i,j}^\alpha) - r(m_{i,j}^\alpha))^2 - b_{i,j}^\alpha \beta_{m_{i,j}^\alpha} \right] \right] + \sum_{j=1}^{|s_i|} (t_{i,j}^\alpha - t_{i,1}^\alpha) O(\varepsilon^{-1})
\end{aligned}
$$

where $D_{i,1}^\alpha$ is the value of $\frac{1}{2}\langle D^2 u \cdot \frac{\nabla^\perp u}{u_\eta}, \frac{\nabla^\perp u}{u_\eta}\rangle)$ evaluated at $(t_{i,1}^\alpha, \xi_{i,1}^\alpha, \eta_{i,1}^\alpha)$, and u_t is also evaluated at this location. We come now to the *key point*:

$$
(5.17)\qquad \sum_{j=1}^{|s_i|} \frac{1}{|s_i|} \left[(q(m_{i,j}^\alpha) - r(m_{i,j}^\alpha))^2 - b_{i,j}^\alpha \beta_{m_{i,j}^\alpha} \right] \le C_d^\#,
$$

since $(\beta^0, \dots, \beta_{2^d-1}, C_d^\#)$ *is a feasible point for the investor's linear program.* Our hypothesis that $u_t \le 0$ and the PDE (5.1) assure that $D_{i,1}^\alpha \ge 0$; so (5.16) is bounded above by

$$
(5.18)\qquad |s_i|[u_t + \frac{C_d^\#}{2}\langle D^2 u \frac{\nabla^\perp u}{u_\eta}, \frac{\nabla^\perp u}{u_\eta}\rangle] + \sum_{j=1}^{|s_i|} (t_{i,j}^\alpha - t_{i,1}^\alpha) O(\varepsilon^{-1}).
$$

Since the first term vanishes by (5.1), we have shown that

$$
(5.19)\qquad \sum_{j=1}^{|s_i|} L(t_{i,j}^\alpha, m_{i,j}^\alpha, b_{i,j}^\alpha, \xi_{i,j}^\alpha, \eta_{i,j}^\alpha, f^\#_{t_{i,j}^\alpha, m_{i,j}^\alpha}) \le \sum_{j=1}^{|s_i|} (t_{i,j}^\alpha - t_{i,1}^\alpha) O(\varepsilon^{-1})
$$

for every i and α.

To finish, we now sum over i and α and change the order of summation:

$$
\text{value of (5.13)} \le \varepsilon^2 \sum_{i=1}^{n} \sum_{\alpha=1}^{\sigma_i} \sum_{j=1}^{|s_i|} (t_{i,j}^\alpha - t_{i,1}^\alpha) O(\varepsilon^{-1}) = \sum_{i=1}^{n} \sum_{j=1}^{|s_i|} \sum_{\alpha=1}^{\sigma_i} (t_{i,j}^\alpha - t_{i,1}^\alpha) O(\varepsilon).
$$

By (5.14) this is at most

$$
\sum_{i=1}^{n} \sum_{j=1}^{|s_i|} [T - t_0] O(\varepsilon).
$$

Since n (the number of cycles) and $\max_i |s_i|$ (the maximum length of a cycle) are constants (depending only on d), we conclude as desired that

$$
\text{value of (5.13)} \le C(T - t_0)\varepsilon,
$$

where C depends only on d and the constants implicit in our hypotheses on u, (2.30)–(2.33). □

Remark 5.3. The only property of $C_d^{\#}$ that we used in the proof was the existence of $\{\beta_m\}$ such that $(\beta_0, \ldots, \beta_{2^d-1}, C_d^{\#})$ is a feasible point for the investor's linear program (see (5.17)). As a reminder, this means that for each cycle s on the graph,

$$\frac{1}{|s|} \sum_{\substack{\text{vertices } m \\ \text{on cycle } s}} [(q(m) - r(m))^2 - b_{i,m}\beta_m] \leq C_d^{\#}.$$

where $|s|$ is the number of edges in s and $b_{i,m}$ was defined after (4.13). By choosing $C_d^{\#}$ to be the optimal value of the investor's linear program we make $C_d^{\#}$ as small as possible, optimizing the resulting bound. But since our argument uses only feasibility (not optimality), *any* feasible point for the investor's linear program determines an upper bound on u^ε. For example, to get an upper bound that's worse than that of Theorem 5.2 but somewhat more explicit, one can take $\beta_m = 0$ for all m and replace $C_d^{\#}$ in (5.1) by

$$\max_{s \in \{\text{cycles}\}} \frac{1}{|s|} \sum_{\substack{\text{vertices } m \\ \text{on cycle } s}} (q(m) - r(m))^2.$$

Remark 5.4. The hypotheses of Theorem 5.2 include that $u_t \leq 0$, an assumption we didn't need when $d = 1$. It was needed for the passage from (5.17) to (5.18). If there is a strategy that achieves indifference (in the sense of Lemmas 4.1(c) and 4.2(c)) then the inequality in (5.17) becomes an equality and the sign of u_t becomes irrelevant. As a reminder: we showed in Section 4.4 that indifference is achievable for $d \leq 4$.

5.2 The lower bound, when φ is regular

For our lower bound on u^ε, the function u solves a PDE similar to that used for the lower bound, but with a different "diffusion constant": throughout this subsection, $u(t, \xi, \eta)$ is the solution of the PDE

$$\begin{aligned} & u_t + \frac{1}{2} C_d^* \langle D^2 u \frac{\nabla^\perp u}{u_\eta}, \frac{\nabla^\perp u}{u_\eta} \rangle = 0, \\ & u(T, \xi, \eta) = \varphi(\xi, \eta), \end{aligned} \tag{5.20}$$

where C_d^* is the optimal value of the market's linear program (4.11). Our goal is to prove

Theorem 5.5. *Let $d \geq 1$ be an integer, and assume the solution u of* (5.20) *satisfies* (2.30)–(2.33). *Then there is a constant C (independent of ε, t, and T) such that*

$$u^\varepsilon(t, m, \xi, \eta) \geq u(t, \xi, \eta) - C[(T - t) + \varepsilon]\varepsilon \tag{5.21}$$

for $t < T$, $\xi \in \mathbb{R}$, $\eta \in \mathbb{R}$, and $m \in \{0, 1\}$, provided that ε is small enough and t is such that $N = (T - t)/\varepsilon^2$ is an integer.

Proof. No new ideas are needed, beyond those introduced for the lower bound when $d = 1$ (Theorem 3.2) and the upper bound for general d (Theorem 5.2). Therefore we will be somewhat brief.

Our lower bound for $d = 1$ was associated with a specific strategy for the market, which was summarized at the beginning of Section 3.2. Our lower bound for general d uses an analogous strategy. Briefly: the market identifies a specific investor strategy of the form $f_m = f_m^* + \varepsilon f_m^\#$ by solving the market's linear program, and chooses the market's movements to penalize the investor if she doesn't make that choice. In more detail: the market's strategy is organized around the analogue of (5.7) obtained using the market's linear program rather than the investor's:

$$\begin{aligned} f_{t,m} &= f_{t,m}^* + \varepsilon f_{t,m}^\#, \text{ with} \\ f_{t,m}^* &= \frac{(q(m)-r(m))u_\xi + (q(m)+r(m))u_\eta}{2u_\eta} \text{ and} \\ f_{t,m}^\# &= \frac{D}{2u_\eta}\beta_m. \end{aligned} \tag{5.22}$$

where $\{\beta_m\}_{m=0}^{2^d-1}$ come from the solution of the market's linear program (4.6). There are two cases, identical to (3.15) and (3.16):

Case 1: If the investor's choice doesn't nearly zero out the "first-order term" in the Taylor-expansion-based estimate of the increment of u, then the market chooses b to make that term positive; quantitatively,

$$\begin{aligned} &\text{if the investor's choice } f \text{ has } |f - f_m^*| \geq \gamma\varepsilon \\ &\text{then the market chooses } b \text{ so that } -b(f - f_m^*) \geq 0. \end{aligned}$$

Case 2: When case 1 doesn't apply, it is convenient to express the investor's choice f as $f = f_m^* + \varepsilon f_m^\# + \varepsilon X$, where f_m^* and $f_m^\#$ are defined by (5.22) (this relation defines X). The investor is more optimistic than our conjectured optimal strategy if $X > 0$, and more pessimistic if $X < 0$. In the former case the market makes the stock go down, and in the latter case it makes the stock go up – in each case giving the investor an unwelcome surprise; quantitatively:

$$\begin{aligned} &\text{if the investor's choice } f \text{ has } |f - f_m^*| < \gamma\varepsilon \\ &\text{then the market chooses } b \text{ so that } -bX \geq 0. \end{aligned}$$

Under conditions on γ analogous to (3.24) and (3.25), an argument entirely parallel to that of Theorem 3.2 produces a sequence $(t_k, m_k, \xi_k, \eta_k)$, starting from any given $(t_0, m_0, \xi_0, \eta_0)$ and ending when $k = N = (T - t_0)/\varepsilon^2$ so that $t_N = T$, such that

$$u^\varepsilon(t_0, m_0, \xi_0, \eta_0) \geq u^\varepsilon(t_1, m_1, \xi_1, \eta_1) \geq \cdots \geq u^\varepsilon(t_N, m_N, \xi_N, \eta_N) \tag{5.23}$$

for which the increments of u satisfy the analogue of (3.22):

$$U_k = u(t_{k+1}, \xi_{k+1}, \eta_{k+1}) - u(t_k, \xi_k, \eta_k) \geq \varepsilon^2 L(t_k, m_k, b_k, \xi_k, \eta_k, f_{t_k, m_k}^\#) + O(\varepsilon^3) \tag{5.24}$$

Then arguments entirely parallel to the those used for Theorem 5.2 show that

$$u(t_N,\xi_N,\eta_N)-u(t_0,\xi_0,\eta_0)\geq -C[(T-t_0)+\varepsilon]\varepsilon. \tag{5.25}$$

The analogue of (5.17) in this setting is of course

$$\sum_{j=1}^{|s_i|}\frac{1}{|s_i|}\left[(q(m_{i,j}^{\alpha})-r(m_{i,j}^{\alpha}))^2-b_{i,j}^{\alpha}\beta_{m_{i,j}^{\alpha}}\right]\geq C_d^*, \tag{5.26}$$

which holds since $(\beta^0,\dots,\beta_{2^d-1},C_d^*)$ is a feasible point for the market's linear program. Combining (5.23)–(5.25) with the fact that $u^{\varepsilon}(t_N,m_N,\xi_N,\eta_N)=\varphi(t_N,\xi_N,\eta_N)=u(t_N,\xi_N,\eta_N)$ gives the desired lower bound

$$u^{\varepsilon}(t_0,m_0,\xi_0,\eta_0)\geq u(t_0,\xi_0,\eta_0)-C[(T-t)+\varepsilon]\varepsilon.$$

□

Remark 5.6. The observations in Remarks 5.3 and 5.4 apply here as well: since the proof of the lower bound uses only the *feasibility* of $(\beta_0,\dots,\beta_{2^d-1},C_d^*)$ for the market's linear program, the same argument can be applied using *any* feasible point for that linear program. For example, to get an upper bound that's worse than that of Theorem 5.5 but more explicit, one can take $\beta_m=0$ for all m and replace C_d^* in (5.20) by

$$\min_{s\in\{\text{cycles}\}}\frac{1}{|s|}\sum_{\substack{\text{vertices}\,m\\ \text{on cycle}\,s}}(q(m)-r(m))^2.$$

Remark 5.7. We know from Lemmas 4.1 and 4.2 that $C_d^*\leq C_d^{\#}$. Our upper and lower bounds match asymptotically in the limit $\varepsilon\to 0$ if and only if $C_d^*=C_d^{\#}$. By Lemmas 4.1(c) and 4.2(c), this relation holds when there is a strategy that achieves indifference.

5.3 The classic case, when $\varphi=\frac{1}{2}(\eta+|\xi|)$

The classic goal of minimizing regret with respect to the best-performing expert corresponds to using the non-smooth final-time data $\varphi=(\eta+|\xi|)/2$. In Section 3, which focused on the case $d=1$, we showed in Theorem 3.3 how our upper and lower bounds can be adapted to this case. The proof involved approximating φ by something a bit smoother, then examining the various error terms.

For the general case $d\geq 1$, the proofs of our upper and lower bounds (Theorems 3.1 and 3.2) were largely parallel to the case $d=1$. The main difference was in some sense bookkeeping, namely our use of the cycle decomposition of a closed walk on the relevant graph; for $d=1$ the cycle decomposition was so simple that we avoided discussing it explicitly, though it was implicit in our discussion of (3.10)–(3.14).

In view of the parallels between our bounds for $d=1$ and those for the general case $d\geq 1$, it is not surprising that the results in this section can be extended to the classic case $\varphi=(\eta+|\xi|)/2$.

Theorem 5.8. *Let $d \geq 1$ be an integer and consider the classic final-time data $\varphi = (\eta + |\xi|)/2$. We consider, as usual, the function $u^\varepsilon(t,m,\xi,\eta)$ defined by the dynamic programming principle* (2.10) *with final-time data φ. Let $u_+(t,\xi,\eta)$ solve our upper-bound PDE* (5.1) *and let $u_-(t,\xi,\eta)$ solve our lower-bound PDE* (5.20)*, with final-time data φ in both cases. Then there is a constant C (independent of ε, t, and T) such that*

$$u_-(t,\xi,\eta) - C\varepsilon|\log\varepsilon| \leq u^\varepsilon(t,m,\xi,\eta) \leq u_+(t,\xi,\eta) + C\varepsilon|\log\varepsilon| \tag{5.27}$$

for all $t < T$, $\xi \in \mathbb{R}$, $\eta \in \mathbb{R}$, and $m \in \{0,1\}^d$, provided that ε is small enough and t is such that $N = (T-t)/\varepsilon^2$ is an integer.

Proof. The proof requires no new ideas: one simply reviews the proofs of Theorems 5.2 and 5.5, estimating the magnitude of each error term as in the proof of Theorem 3.3. We leave the details to the reader. □

Appendix A: Solving our PDE using the linear heat equation

Our bounds involve solutions of the final-value problem

$$\begin{aligned} &u_t + \frac{1}{2}C\langle D^2u\frac{\nabla^\perp u}{u_\eta}, \frac{\nabla^\perp u}{u_\eta}\rangle = 0, \\ &u(T,\xi,\eta) = \varphi(\xi,\eta) \end{aligned} \tag{A.1}$$

with various choices of the "diffusion constant" C. We rely on the existence of a sufficiently smooth solution with certain qualitative properties, namely (2.30)–(2.33). We asserted in Section 2.5 the existence of such a solution under certain conditions upon φ, namely (2.38)–(2.41). The main goal of this appendix is to prove that assertion. We also briefly discuss a geometric interpretation of the PDE, and we give an explicit formula for u in the classic case $\varphi = \frac{1}{2}(\eta + |\xi|)$. All these results are already known [25], but we present a self-contained treatment here for the reader's convenience.

A.1 Motivation

We begin with the observation that if $\varphi(\xi,\eta)$ satisfies

$$\varphi_\eta \geq c > 0 \tag{A.2}$$

for some constant c, then for each $y \in \mathbb{R}$ there is a unique value $g(\xi;y)$ such that

$$\varphi(\xi, g(\xi;y)) = y. \tag{A.3}$$

Thus: for each y the level set $\{\varphi = y\}$ is the graph of a function $\xi \mapsto g(\xi;y)$. These graphs foliate the (ξ,η) plane and provide an alternative representation of the function φ. Implicit differentiation in ξ reveals that

$$\varphi_\xi + \varphi_\eta g_\xi = 0 \quad \text{and} \quad \varphi_{\xi\xi} + 2\varphi_{\xi\eta}g_\xi + \varphi_{\eta\eta}g_\xi^2 + \varphi_\eta g_{\xi\xi} = 0$$

when $\eta = g(\xi; y)$. In view of (A.2) we can solve for g_ξ and $g_{\xi\xi}$, obtaining

$$g_\xi = -\frac{1}{\varphi_\eta}\varphi_\xi \quad \text{and} \quad g_{\xi\xi} = -\frac{1}{\varphi_\eta^3}(\varphi_{\xi\xi}\varphi_\eta^2 - 2\varphi_{\xi\eta}\varphi_\xi\varphi_\eta + \varphi_{\eta\eta}\varphi_\xi^2). \tag{A.4}$$

There is also a simple formula for g_y: differentiation of (A.3) with respect to y gives

$$g_y = 1/\varphi_\eta. \tag{A.5}$$

If we accept that the PDE (A.1) has a solution $u(t,\xi,\eta)$ with $u_\eta > c > 0$, then the preceding discussion applies to it as well (treating the time t as a parameter). So there is a function $h(t,\xi;y)$ such that

$$u(t,\xi,h(t,\xi;y)) = y \tag{A.6}$$

for all ξ and y, and all $t < T$. Differentiation in time gives

$$u_t + u_\eta h_t = 0$$

and differentiation with respect to ξ gives the analogue of (A.4):

$$h_\xi = -\frac{1}{u_\eta}u_\xi \quad \text{and} \quad h_{\xi\xi} = -\frac{1}{u_\eta^3}(u_{\xi\xi}u_\eta^2 - 2u_{\xi\eta}u_\xi u_\eta + u_{\eta\eta}u_\xi^2) \tag{A.7}$$

when $\eta = h(t,\xi;y)$. Since

$$\langle D^2u\nabla^\perp u, \nabla^\perp u\rangle = u_{\xi\xi}u_\eta^2 - 2u_{\xi\eta}u_\xi u_\eta + u_{\eta\eta}u_\xi^2, \tag{A.8}$$

we see that (A.1) holds exactly if $h(t,\xi;y)$ solves the linear heat equation

$$h_t + \frac{1}{2}Ch_{\xi\xi} = 0 \tag{A.9}$$

for all y and ξ and all $t < T$, with the final-time condition

$$h(T,\xi;y) = g(\xi;y). \tag{A.10}$$

The hypothesis that u_η is positive is consistent with this PDE characterization of h: indeed, differentiating (A.6) with respect to y gives

$$u_\eta = 1/h_y, \tag{A.11}$$

and differentiation of (A.9)–(A.10) reveals that

$$\varphi_\eta \geq c \text{ implies } u_\eta \geq c \tag{A.12}$$

for any constant $c > 0$, using the maximum principle for the linear heat equation and the relations $u_\eta = 1/h_\eta$, $g_\eta = 1/\varphi_\eta$.

Summarizing: we have shown that if u exists and is sufficiently smooth, then for each y the level set $\{u = y\}$ is the graph of a function $h(t,\xi;y)$, where h solves the final-value problem (A.9)–(A.10) in t and ξ.

A.2 Existence and properties of u

Our construction of u reverses the preceding discussion: we solve the linear heat equation to get h, then use h to get u.

Theorem A.1. *Suppose $\varphi(\xi,\eta)$ is a C^4 function on $\mathbb{R}^2$, with*

$$\varphi_\eta \geq c > 0 \tag{A.13}$$

for some constant c. While φ can (indeed, must) have linear growth at infinity, we assume that $\varphi_\xi, \varphi_\eta$, and all higher derivatives of order up to 4 are bounded. Then for any positive constant C the final-value problem

$$\begin{aligned} u_t + \frac{1}{2}C\langle D^2u\frac{\nabla^\perp u}{u_\eta}, \frac{\nabla^\perp u}{u_\eta}\rangle = 0 \textit{ for } t < T \\ u(T,\xi,\eta) = \varphi(\xi,\eta) \end{aligned} \tag{A.14}$$

has a classical solution $u(t,\xi,\eta)$ such that

$$u_\eta \geq c, \tag{A.15}$$

(A.16) *u has continuous, uniformly bounded derivatives in (ξ,η) of order up to 4,*

(A.17) *u_t has continuous, uniformly bounded derivatives in (ξ,η) of order up to 2, and*

(A.18) *u_{tt} is continuous and uniformly bounded,*

all bounds being uniform as $t \uparrow T$. Moreover, the following structural properties of φ persist to u:

(A.19) *if $|\varphi_\xi| \leq \varphi_\eta$ for all ξ,η, then $|u_\xi| \leq u_\eta$ for all ξ,η and all $t \leq T$;*

(A.20) *if $\varphi_{\xi\xi}\varphi_\eta^2 - 2\varphi_{\xi\eta}\varphi_\xi\varphi_\eta + \varphi_{\eta\eta}\varphi_\xi^2 \geq 0$ for all ξ,η, then $u_t \leq 0$ for all ξ,η and all $t \leq T$;*

(A.21) *if φ is odd in ξ then so is u.*

Proof. We begin with the level-set representation of φ, i.e. with the function $g(\xi;y)$ defined by (A.3). To assess the smoothness of g we apply the inverse function theorem to the maps G and Φ defined by

$$(\xi,y) \xrightarrow{G} (\xi,g(\xi;y)) \quad \text{and} \quad (\xi,\eta) \xrightarrow{\Phi} (\xi,\varphi(\xi,\eta)). \tag{A.22}$$

The relation $\varphi(\xi,g(\xi;y)) = y$ says that $\Phi \circ G$ is the identity map, so G is the inverse of Φ. The Jacobian of Φ is uniformly bounded and positive, by (A.13). It follows, by the inverse function theorem (see e.g. [10]) that G has the same smoothness as Φ; in particular, since φ has uniformly bounded derivatives of order up to 4, so does g.

Now consider the solution $h(t,\xi;y)$ of the final-value problem (A.9)–(A.10). Differentiating the PDE in ξ and/or y and applying the maximum principle for the linear heat equation, we see that the derivatives of h in ξ, y of order up to 4 are uniformly bounded. Since g_y is uniformly positive as well as uniformly bounded (by (A.5)), the same is true of h_y. So for each t, the graphs $\eta = h(t,\xi;y)$ foliate the (ξ,η) plane as y varies, determining a unique function u such that (A.6) holds. The arguments in Section A.1 show that u is a classical solution of the PDE (A.14), and that $u_\eta \geq c$. To assess its regularity in ξ and η, we observe that for each fixed t the functions

$$(\xi, y) \to (\xi, h(t,\xi;y)) \quad \text{and} \quad (\xi,\eta) \to (\xi, u(t,\xi,\eta))$$

are inverse to one another; so the regularity of h implies, via the inverse function theorem, that u has uniformly bounded derivatives of order up to 4 in ξ and η. The regularity of u in t is best assessed using the PDE: since $\langle D^2 u \frac{\nabla^\perp u}{u_\eta}, \frac{\nabla^\perp u}{u_\eta}\rangle$ has bounded second spatial derivatives, so does u_t. It follows immediately (by differentiating the PDE in t) that u_{tt} is also uniformly bounded.

Turning now to qualitative properties, recall from (A.4) and (A.7) that $|\varphi_\xi| \leq \varphi_\eta$ is equivalent to $|g_\xi| \leq 1$, and $|u_\xi| \leq u_\eta$ is equivalent to $|h_\xi| \leq 1$. Since $|g_\xi| \leq 1$ implies $|h_\xi| \leq 1$ by the maximum principle for the linear heat equation, (A.19) is clear.

Next, recall from (A.4)–(A.8) that $\varphi_{\xi\xi}\varphi_\eta^2 - 2\varphi_{\xi\eta}\varphi_\xi\varphi_\eta + \varphi_{\eta\eta}\varphi_\xi^2 \geq 0$ is equivalent to $g_{\xi\xi} \leq 0$, while $u_{\xi\xi}u_\eta^2 - 2u_{\xi\eta}u_\xi u_\eta + u_{\eta\eta}u_\xi^2 \geq 0$ is equivalent to both $h_{\xi\xi} \leq 0$ and $u_t \leq 0$. Since $h_{\xi\xi}$ solves a linear heat equation with $g_{\xi\xi}$ as final-time data, (A.20) follows once again from the maximum principle for the linear heat equation.

Finally, recall from (A.3) and (A.6) that $\varphi(\xi,\eta) = \varphi(-\xi,\eta)$ is equivalent to $g(\xi) = g(-\xi)$, and $u(\xi,\eta) = u(-\xi,\eta)$ is equivalent to $h(t,\xi,\eta) = h(t,-\xi,\eta)$. So (A.21) follows from the fact that the solution of a linear heat equation is odd if the final-time data are odd. □

A.3 Geometric interpretation of u

We have shown that for each y, the evolution of the level set $u = y$ can be found without considering any other level sets (by solving a linear heat equation). Second-order parabolic PDE's with this property are called "geometric," because the normal velocity of each level set can be written in terms of its normal direction and curvature [14].

It is natural to ask what our PDE (A.1) looks like from this perspective (though the main part of our paper makes no use of this result). The answer is that the normal velocity of the level set $\{u = y\}$, viewed as a curve in the (ξ,η) plane, is related to the level set's unit normal $\vec{n}$ and curvature κ by

$$v_{nor} = -\frac{C}{2}\frac{|\nabla u|^2}{|u_\eta|^2}\kappa = -\frac{C}{2(\vec{n}\cdot(0,1))^2}\kappa. \tag{A.23}$$

Indeed, the curvature of a level set

$$\kappa = -\text{div}\left(\frac{\nabla u}{|\nabla u|}\right)$$

has the property that

$$\langle D^2 u \frac{\nabla^\perp u}{|\nabla u|}, \frac{\nabla^\perp u}{|\nabla u|}\rangle = -\kappa |\nabla u|,$$

so the PDE (A.1) says

$$u_t = \frac{1}{2} C\kappa \frac{|\nabla u|^3}{u_\eta^2}.$$

This leads directly to (A.23), since the velocity of the level set is (by definition) $v_{nor} = -u_t/|\nabla u|$ and the unit normal is $\vec{n} = (u_\xi, u_\eta)/|\nabla u|$.

A.4 Explicit solution for the classic final-time data

As we observed in the Introduction, in the classic case $\varphi = \frac{1}{2}(\eta + |\nabla \xi|)$ the solution of (A.1) has the form $u = \frac{1}{2}\eta + \overline{u}(t,\xi)$, where $\overline{u}$ solves the linear heat equation $\overline{u}_t + \frac{1}{2} C \overline{u}_{\xi\xi} = 0$ for $t < T$ with $\overline{u}(T,\xi) = \frac{1}{2}|\xi|$. It is amusing to observe that this function $\overline{u}$ has an explicit formula, namely

$$\overline{u}(t,\xi) = \sqrt{T-t}\, G\left(\xi/\sqrt{T-t}\right) \tag{A.24}$$

where

$$G(z) = \sqrt{\frac{C}{2\pi}} \exp\left(\frac{-z^2}{2C}\right) + \frac{z}{2} \operatorname{erf}\left(\frac{z}{\sqrt{2C}}\right) \tag{A.25}$$

(with the usual convention that $\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-s^2}\, ds$, so that $\operatorname{erf}(x) \to 1$ as $x \to \infty$). Indeed, a brief calculation reveals that a function of the form (A.24) solves the desired PDE exactly if G solves

$$G(z) - zG'(z) - CG''(z) = 0,$$

and it is easy to check that the proposed function G (given by (A.25)) meets this requirement. Turning to the final-time behavior: it is easy to check that our G satisfies

$$\lim_{z\to\infty} \frac{G(z)}{z} = \frac{1}{2} \quad \text{and} \quad \lim_{z\to-\infty} \frac{G(z)}{z} = -\frac{1}{2},$$

and it follows that

$$\lim_{t\to T} \overline{u}(t,\xi) = |\xi|/2$$

as expected.

Appendix B: Necessity of the hypothesis $|q(m)| < 1$ and $|r(m)| < 1$

We assumed in (2.2) that the experts' choices satisfy $|q(m)| < 1, |r(m)| < 1$ for all states m. We explained in Section 2.5 the utility of this hypothesis: it assures that the predictor's optimal choice of f_m^* satisfies $|f_m^*| < 1$, so that the optimal strategy $f_m = f_m^* + \varepsilon f_m^\#$ is admissible when ε is sufficiently small. It is natural to wonder, however, what would change if the experts' choices were constrained only by $|q(m)| \leq 1$ and $|r(m)| \leq 1$.

To capture the essential phenomenon, let us focus on what happens when the experts use just the most recent stock move ($d = 1$) and their choices satisfy

$$q(0) = r(0) = 1 \text{ and}$$
$$q(1) \neq r(1), \text{ each taking a value in } (-1,1).$$

(Recall that when $d = 1$ the relevant graph has just two states, $m = 0$ and $m = 1$, as shown in Figure 2.2.) Our heuristic discussion of the case $d = 1$ in Section 2.4 assumed that the market chose one of the graph's three cycles. Our rigorous treatment in Section 3 showed that the conclusions reached by such a discussion are in fact correct. Therefore it suffices to use the framework of Section 2.4 for the analysis of this example.

The investor's optimal choice of f_m^* and $f_m^\#$ were characterized by (3.8) (assuming $|q(m)| < 1$ and $|r(m)| < 1$). Recalling the alternative expression (2.34) for f_m^*, namely

$$f_m^* = q(m)\frac{u_{x_1}}{u_{x_1} + u_{x_2}} + r(m)\frac{u_{x_2}}{u_{x_1} + u_{x_2}},$$

we see that $f_0^* = 1$ while $|f_1^*| < 1$. Equation (3.8) would give

$$f_0^\# = f_1^\# = \frac{[q(1) - r(1)]^2 D}{4u_\eta} > 0 \tag{B.1}$$

(using our usual convention $D = \frac{1}{2}\langle D^2u \cdot \frac{\nabla^\perp u}{u_\eta}, \frac{\nabla^\perp u}{u_\eta}\rangle$, and remembering that D and u_η are both positive). Evidently $f_0^* + \varepsilon f_0^\# > 1$, which violates the condition that $|f_m| \leq 1$.

To identify the optimal strategy in our example, we briefly review the argument of Section 2.4, which involved the "rate at which regret accumulates" when the market chooses one of the three cycles. For cycle $0-0$ this rate is

$$[q(0) - r(0)]^2 D + 2u_\eta f_0^\# = 2u_\eta f_0^\# \tag{B.2}$$

by (2.24); similarly, for cycle $1-1$ it is

$$[q(1) - r(1)]^2 D - 2u_\eta f_1^\# \tag{B.3}$$

by (2.25) and for cycle $0-1-0$ it is

$$\tfrac{1}{2}[q(1) - r(1)]^2 D - u_\eta f_0^\# + u_\eta f_1^\# \tag{B.4}$$

by (2.26). The choice (B.1) makes these three expressions equal; their common value

$$\frac{1}{2}[q(1)-r(1)]^2 D \tag{B.5}$$

is the rate at which regret accumulates when the investor and market behave optimally (and $|q(m)|<1, |r(m)|<1$ for $m=0,1$). However, in the example under consideration here the investor must optimize the max of these three expression subject to the constraint $f_0^{\#} \le 0$. A brief calculation reveals that the optimal choice is

$$f_0^{\#}=0, \quad f_1^{\#}=\frac{[q(1)-r(1)]^2 D}{6u_\eta},$$

for which the value of (B.2) is 0 while the values of (B.3) and (B.4) are both

$$\frac{2}{3}[q(1)-r(1)]^2 D. \tag{B.6}$$

Evidently, the market can choose either cycle $1-1$ or cycle $0-1-0$, and the rate at which regret accumulates is given by (B.6), which is different from (B.5).

In short: our methods can be used even if the experts' strategies satisfy $|q(m)|=1$ and $|r(m)|=1$ for some states m; however if $|f_m^*|=1$ then the linear programs considered in Section 4 acquire an inequality constraint on the associated $f_m^{\#}$, and this can affect the optimal strategies and the rate at which regret accumulates. While our discussion has focused on the case $d=1$, this is in fact the situation for any value of d.

Acknowledgment. This research was partially supported by NSF grant DMS-1311833.

Bibliography

[1] Andoni, A.; Panigrahy, R. A differential equations approach to optimizing regret trade-offs. *arXiv:1305.1359* (2013). Available at: `http://arxiv.org/abs/1305.1359`

[2] Antunovic, T.; Peres, Y.; Sheffield, S.; Somersille, S. Tug-of-war and infinity Laplace equation with vanishing Neumann boundary condition. *Communications in Partial Differential Equations* **37** (2012), no. 10, 1839–1869. Available at: `https://doi.org/10.1080/03605302.2011.642450`

[3] Armstrong, S. N.; Smart, C. K. A finite difference approach to the infinity Laplace equation and tug-of-war games. *Transactions of the American Mathematical Society* **364** (2012), no. 2, 595–636. Available at: `http://dx.doi.org/10.1090/S0002-9947-2011-05289-X`

[4] Calder, J.; Drenska, N. Online prediction with history-dependent experts: the general case. *Communications on Pure and Applied Mathematics* (2020). In press; preprint available as arXiv:2008.00052.

[5] Calder, J.; Drenska, N. Asymptotically optimal strategies for online prediction with history-dependent experts. *Journal of Fourier Analysis and Applications* **27** (2021), no. 2, 1–20. Available at: `https://doi.org/10.1007/s00041-021-09815-4`

[6] Cesa-Bianchi, N.; Lugosi, G. On prediction of individual sequences. *Ann. Statist.* **27** (1999), no. 6, 1865–1895. Available at: `http://dx.doi.org/10.1214/aos/1017939242`

[7] Cesa-Bianchi, N.; Lugosi, G. Potential-based algorithms in on-line prediction and game theory. *Machine Learning* **51** (2003), 239–261. Available at: `https://doi.org/10.1023/A:1022901500417`

[8] Cesa-Bianchi, N.; Lugosi, G. *Prediction, Learning, and Games*, Cambridge University Press, New York, 2006.

[9] Cover, T.: Behaviour of sequential predictors of binary sequences, in *Proceedings of the Fourth Prague Conference on Information Theory, Statistical Decision Functions, and Random Processes*, Publishing House of the Czechoslovak Academy of Sciences, Prague, 1965 pp. 263–272.

[10] Dieudonné, J. *Foundations of Modern Analysis*, Academic Press, 1960.

[11] Drenska, N.; Kohn, R. V. Prediction with expert advice: A PDE perspective. *Journal of Nonlinear Science* **30** (2020), no. 1, 137–173. Available at: `https://doi.org/10.1007/s00332-019-09570-3`

[12] Fredricksen, H. A new look at the de Bruijn graph. *Discrete Applied Mathematics* **37/38** (1992), 193–203. Available at: `https://www.sciencedirect.com/science/article/pii/0166218X9290133U`

[13] Freund, Y.; Opper, M. Drifting games and Brownian motion. *Journal of Computer and System Sciences* **64** (2002), no. 1, 113–132. Available at: `https://www.sciencedirect.com/science/article/pii/S0022000001918021`

[14] Giga, Y. *Surface Evolution Equations*, Springer, 2006.

[15] Kapralov, M.; Panigrahy, R. Prediction strategies without loss. in *Advances in Neural Information Processing Systems 24*, edited by J. Shawe-Taylor; R. S. Zemel; P. L. Bartlett; F. Pereira; K. Q. Weinberger, pp. 828–836, Curran Associates, Inc., 2011. Available at: `http://papers.nips.cc/paper/4388-prediction-strategies-without-loss.pdf`

[16] Kobzar, V. A.; Kohn, R. V.; Wang, Z.: New potential-based bounds for prediction with expert advice, in *Proceedings of Thirty Third Conference on Learning Theory*, *Proceedings of Machine Learning Research*, vol. 125, edited by J. Abernethy; S. Agarwal, PMLR, 2020 pp. 2370–2405. Available at: `http://proceedings.mlr.press/v125/kobzar20a.html`

[17] Kobzar, V. A.; Kohn, R. V.; Wang, Z.: New potential-based bounds for the geometric-stopping version of prediction with expert advice, in *Proceedings of The First Mathematical and Scientific Machine Learning Conference*, *Proceedings of Machine Learning Research*, vol. 107, edited by J. Lu; R. Ward, PMLR, 2020 pp. 537–554. Available at: `http://proceedings.mlr.press/v107/kobzar20a.html`

[18] Kohn, R. V.; Serfaty, S. A deterministic-control-based approach motion by curvature. *Communications on Pure and Applied Mathematics* **59** (2006), no. 3, 344–407. Available at: `http://dx.doi.org/10.1002/cpa.20101`

[19] Kohn, R. V.; Serfaty, S. A deterministic-control-based approach to fully nonlinear parabolic and elliptic equations. *Communications on Pure and Applied Mathematics* **63** (2010), no. 10, 1298–1350. Available at: `http://dx.doi.org/10.1002/cpa.20336`

[20] Lewicka, M.; Manfredi, J. J. The obstacle problem for the p-Laplacian via optimal stopping of tug-of-war games. *Probability Theory and Related Fields* **167** (2017), 349–378. Available at: `http://dx.doi.org/10.1007/s00440-015-0684-y`

[21] Peres, Y.; Schramm, O.; Sheffield, S.; Wilson, D. B. Tug-of-war and the infinity Laplacian. *Journal of the American Mathematical Society* **22** (2009), no. 1, 167–210. Available at: `http://dx.doi.org/10.1090/S0894-0347-08-00606-1`

[22] Peres, Y.; Sheffield, S. Tug-of-war with noise: A game-theoretic view of the p-Laplacian. *Duke Mathematical Journal* **145** (2008), no. 1, 91–120. Available at: `http://dx.doi.org/10.1215/00127094-2008-048`

[23] Rokhlin, D. B. PDE approach to the problem of online prediction with expert advice: a construction of potential-based strategies. *International Journal of Pure and Applied Mathematics* **114** (2017), no. 4, 907–915. Available at: `http://dx.doi.org/10.12732/ijpam.v114i4.20`

[24] West, D. B. *Introduction to Graph Theory*, Prentice Hall, 2001, 2nd ed.

[25] Zhu, K.: Two problems in applications of pde, Ph.D. thesis, New York University. 2014. Available at: `http://pqdtopen.proquest.com/pubnum/3635320.html`

Received Month 200X.