Optimal anytime regret with two experts

Nicholas J. A. Harvey [*] Christopher Liaw [†] Edwin Perkins [‡] Sikander Randhawa [§]

Abstract

The multiplicative weights method is an algorithm for the problem of prediction with expert advice. It achieves the minimax regret asymptotically if the number of experts is large, and the time horizon is known in advance. Optimal algorithms are also known if there are exactly two or three experts, and the time horizon is known in advance.

In the anytime setting, where the time horizon is *not* known in advance, algorithms can be obtained by the "doubling trick", but they are not optimal, let alone practical. No minimax optimal algorithm was previously known in the anytime setting, regardless of the number of experts.

We design the first minimax optimal algorithm for minimizing regret in the anytime setting. We consider the case of two experts, and prove that the optimal regret is $\gamma\sqrt{t}/2$ at all time steps t, where γ is a natural constant that arose 35 years ago in studying fundamental properties of Brownian motion. The algorithm is designed by considering a continuous analogue, which is solved using ideas from stochastic calculus.

[*]Email: `nickhar@cs.ubc.ca`. University of British Columbia, Department of Computer Science.
[†]Email: `cvliaw@cs.ubc.ca`. University of British Columbia, Department of Computer Science.
[‡]Email: `perkins@math.ubc.ca`. University of British Columbia, Department of Mathematics.
[§]Email: `srand@cs.ubc.ca`. University of British Columbia, Department of Computer Science.

Contents

1 Introduction

We study the classical problem of prediction with expert advice, whose origin can be traced back as early as the 1950s [28]. The problem can be formulated as a sequential game between an adversary and an algorithm as follows. At each time t, an adversary chooses a cost for each of n possible experts. Without knowledge of the adversary's move, the algorithm must choose (perhaps randomly) one of the n experts to follow. The cost of each expert is then revealed to the algorithm, and the algorithm incurs the cost that its chosen expert incurred. The goal is to design an algorithm whose *regret* is small, i.e. the algorithm's expected total cost is small relative to the total cost of the best expert. In the theoretical computer science community, algorithms for this problem and its variants have been a key component in many results; we refer the reader to [3] and the references therein for a survey on some of these results.

The most well-known algorithm for the experts problem is the celebrated multiplicative weights update algorithm (MWU) which was introduced, independently, by Littlestone and Warmuth [31] and by Vovk [40]. The algorithm itself is very elegant and commonly taught in courses on algorithms, machine learning, and algorithmic game theory. An analysis of MWU shows that, in the fixed-time setting (where a time horizon T is known in advance), it achieves a regret of $\sqrt{(T/2)\ln n}$ at time T, where n is the number of experts [11, 10]. This bound on the regret of MWU is known to be tight whenever $n \geq 2$ is an even integer [25]. It is also known [11] that $\sqrt{(T/2)\ln n}$ is asymptotically optimal for any algorithm as $n, T \to \infty$. Hence, MWU is a minimax optimal[1] algorithm as $n, T \to \infty$. Interestingly, MWU is *not* optimal for small values of n. For $n = 2$, Cover [14] observed that a natural dynamic programming formulation of the problem leads to a simple analysis showing that the minimax optimal regret is $\sqrt{T/2\pi}$.

The assumption that the time horizon T is known in advance may be problematic in some scenarios; examples include any sort of online tasks (e.g., online learning), or tasks requiring convergence over time (e.g., convergence to equilibria). These scenarios may be better suited for the *anytime setting*, which has the stronger requirement that the regret be controlled at *all* points in time. Another interesting setting is the *geometric horizon setting*, introduced by Gravin, Peres, and Sivan [24], in which the time horizon is a geometric random variable of known distribution. In this setting, they gave the optimal algorithm for three experts. (The optimal algorithm for two experts is inherited from Cover's result).

The anytime setting is the focus of this work. There is a well-known "doubling trick" [11, §4.6] that can be used to convert algorithms for the fixed-time setting to algorithms for the anytime setting. Typically, the doubling trick involves restarting the fixed-time horizon algorithm every power-of-two steps with new parameters. If the fixed-time algorithm has regret $O(T^c)$ for some $c \in (0, 1)$ then the doubling trick yields an algorithm with regret $O(t^c)$ for every $t \geq 1$. On the one hand, this is a conceptually simple and generic reduction from the anytime setting to the fixed-time setting. On the other hand, this approach is inelegant, wasteful, and turns useful algorithms into algorithms of dubious practicality.

Instead of using the doubling trick, one can instead use variants of MWU with a dynamic step size; see, e.g., [12, §2.3], [34, Theorem 1], [6, §2.5]. This is a much more elegant and practical approach than the doubling trick (and is even simpler to implement). However, the analysis is somewhat different and more difficult than the standard MWU analysis, and is rarely taught. It is known that, with an appropriate choice of step sizes, MWU can guarantee[2] a regret of $\sqrt{t \ln n}$ for all $t \geq 1$ and all $n \geq 2$ (see [6, Theorem 2.4]). However, it is unknown whether $\sqrt{t \ln n}$ is the minimax optimal anytime regret, for any value of n.

A motivating conjecture. Our research agenda is motivated by the following question.

Are there settings in which the minimax anytime regret equals the minimax fixed-time regret?

In any such setting, the fixed-time algorithm would be making the unnecessary assumption that the time horizon is known in advance. One might as well use the anytime algorithm since it has fewer assumptions

[1] This means that the algorithm minimizes the maximum, over all adversaries, of the regret.

[2] It can be shown, by modifying arguments of [25], that this is the optimal anytime analysis for MWU with step sizes $c/\sqrt{t}$.

and achieves a stronger guarantee. At present the question has been answered negatively for the setting of two experts [32]: the anytime regret is strictly worse than the fixed-time regret, as discussed below.

Nevertheless, we conjecture that the question has a positive answer in the limit $n \to \infty$, which is the only setting in which MWU is known to be optimal. This would imply that there is an anytime algorithm with weaker hypotheses and stronger guarantees than MWU. More specifically, we conjecture that the optimal anytime regret for n experts is intimately related to the roots of a special function (a confluent hypergeometric function, to be discussed later). That leads to the following conjectured regret bound.

Conjecture 1.1. The optimal regret for n experts in the anytime setting is $\sqrt{\frac{t}{2}\ln n}$ for all t, as $n \to \infty$. This matches the guarantee of MWU for the fixed-time setting.

Results and techniques. As a first step in our research agenda, this work considers the anytime setting with $n = 2$ experts. We show that the optimal regret is $\frac{\gamma}{2}\sqrt{t}$, where $\gamma \approx 1.30693$ is a fundamental constant that arises in the study of Brownian motion [35]. A very concise algorithm achieving the minimax regret is presented in Algorithm 1 on page 4.

Our techniques to derive and analyze this algorithm are a significant departure from previous work on regret minimization. Firstly, our algorithm is designed by considering a continuous-time analogue of the discrete-time algorithmic process. This idea of viewing discrete algorithms in continuous time has recently been fruitful in unrelated lines of work, e.g., [7, 9, 13, 1, 20, 41, 30, 8, 17]. Secondly, we use tools of stochastic calculus to design and analyze algorithms for our continuous-time problem. Our work appears to be the first to solve a regret minimization problem using tools from continuous-time stochastic calculus.

Lastly, we use confluent hypergeometric functions to design and analyze the optimal continuous-time algorithm. These functions may seem exotic, but they turn out to be inherent to our problem since they also arise in the matching lower bound. The constant γ in the minimax regret may be defined as $\alpha(1/2)$, where α is a function giving the root of a confluent hypergeometric function with certain parameters (see Claim A.5 and [35]). While attempting to generalize our techniques to n experts (a task that is still underway) we observed that $\lim_{n\to\infty} \alpha(1/n)/\sqrt{\ln n} = \sqrt{1/2}$, which lends support towards Conjecture 1.1.

Applications. The prediction problem with two experts is closely related to the problem of *predicting binary sequences*; in fact, this was the problem originally considered by Cover [14]. A notable[3] paper by Feder et al. [21] pursued this problem further, defining the notion of universal s-state predictors, and showing connections to Lempel-Ziv compression. They derive [21, Theorem 1 and Eq. (14)] a universal online predictor whose expected performance converges to the performance of the best 1-state predictor at rate $1/\sqrt{t} + O(1/t)$ where t is the sequence length. We describe a different online predictor achieving the better convergence rate $\gamma/2\sqrt{t}$, and show that no other online predictor can improve the constant $\gamma/2$.

We also give a second application of our techniques to a problem purely in probability theory, and not involving regret at all. Let $(X_t)_{t\geq 0}$ be a standard random walk. Then $\mathrm{E}\,[\,|X_\tau|\,] \leq \gamma\,\mathrm{E}\,[\,\sqrt{\tau}\,]$ for every stopping time τ; moreover, the constant γ cannot be improved. This result is originally due to Davis [16, Eq. (3.8)], who proved it first for Brownian motion then derived the result for random walks via the Skorokhod embedding. We will prove it in Appendix B as a simple corollary of our main result.

2 Discussion of results and techniques

2.1 Formal problem statement

The problem may be stated formally as follows. For each integer $t \geq 1$, there is a prediction task, which is said to occur at time t. The task involves a deterministic algorithm $\mathcal{A}$, which must pick a vector $x_t \in [0,1]^n$, and an adversary $\mathcal{B}$, which knows $\mathcal{A}$ and picks a vector $\ell_t \in [0,1]^n$. The vector x_t must satisfy

[3] Their paper won the 1993 IEEE Information Theory Society Paper Award.

$\sum_{j=1}^{n} x_{t,j} = 1$ and may depend on $\ell_1, \ldots, \ell_{t-1}$ (and implicitly $x_1, \ldots, x_{t-1}$). The vector ℓ_t may depend on $\mathcal{A}$ and on $\ell_1, \ldots, \ell_{t-1}$ (and implicitly $x_1, \ldots, x_t$, since $\mathcal{A}$ is deterministic and known).

The dimension n denotes the number of experts. The coordinate $\ell_{t,j}$ denotes the cost of the jth expert at time t. The vector x_t may be viewed as a probability distribution, so the inner product $\langle\, x_t,\, \ell_t \,\rangle$ is the expected cost of the algorithm at time t. Thus, the total expected cost of the algorithm up to time t is $\sum_{i=1}^{t} \langle\, x_i,\, \ell_i \,\rangle$. For $j \in [n]$, the total cost of the jth expert up to time t is $L_{t,j} = \sum_{i=1}^{t} \ell_{i,j}$. The regret at time t of algorithm $\mathcal{A}$ against adversary $\mathcal{B}$ is the difference between the algorithm's total expected cost and the total cost of the best expert, i.e.,

$$\text{Regret}(n, t, \mathcal{A}, \mathcal{B}) \;=\; \sum_{i=1}^{t} \langle\, x_i,\, \ell_i \,\rangle \;-\; \min_{j \in [n]} L_{t,j}.$$

Anytime setting. This work focuses on the anytime setting where the algorithm's objective is to minimize, for *all* t, the regret normalized by $\sqrt{t}$. Specifically, the minimax optimal algorithm must solve

$$\text{AnytimeNormRegret}(n) \;:=\; \inf_{\mathcal{A}} \sup_{\mathcal{B}} \sup_{t \geq 1} \frac{\text{Regret}(n, t, \mathcal{A}, \mathcal{B})}{\sqrt{t}}. \tag{2.1}$$

As mentioned above, MWU with a time-varying step size achieves $\text{AnytimeNormRegret}(n) \leq \sqrt{\ln n}$ for all $n \geq 2$ [6, §2.5]. It is unknown whether this bound is tight, although as $n \to \infty$ it can be loose by at most a factor $\sqrt{2}$ due to the lower bound from the fixed time horizon setting [11]. The minimax optimal anytime regret is unknown even in the case of $n = 2$ experts. The best known bounds at present are

$$0.564 \;\approx\; \sqrt{1/\pi} \;\leq\; \text{AnytimeNormRegret}(2) \;\leq\; \sqrt{\ln 2} \;\approx\; 0.833. \tag{2.2}$$

The lower bound, due to [32], demonstrates a gap between the anytime setting and the fixed-time setting, where the optimal normalized regret is $\sqrt{1/2\pi}$ [14]. We will show that neither inequality in (2.2) is tight.

2.2 Statement of results

To state our main theorem and our algorithm, we must define two special functions.

$$\begin{aligned} \text{erfi}(x) \;&=\; \frac{2}{\sqrt{\pi}} \int_0^x e^{z^2}\, \mathrm{d}z \\ M_0(x) \;&=\; e^x - \sqrt{\pi x}\, \text{erfi}(\sqrt{x}) \end{aligned} \tag{2.3}$$

The first one is the well-known imaginary error function. The second one is a confluent hypergeometric function with certain parameters, as discussed in Appendix A. A key constant used throughout this paper is γ, which is the smallest[4] positive root of $M_0(x^2/2)$, i.e.,

$$\gamma \;:=\; \min \left\{\, x > 0 \;:\; M_0(x^2/2) = 0 \,\right\} \;\approx\; 1.3069\ldots \tag{2.4}$$

It is known that the constant γ relates to the slow points[5] of Brownian motion [33, §10.3].

Theorem 2.1 (Main result). In the anytime setting with two experts, the minimax optimal normalized regret (over deterministic algorithms $\mathcal{A}$ and adversaries $\mathcal{B}$) is

$$\text{AnytimeNormRegret}(2) \;=\; \inf_{\mathcal{A}} \sup_{\mathcal{B}} \sup_{t \geq 1} \frac{\text{Regret}(2, t, \mathcal{A}, \mathcal{B})}{\sqrt{t}} \;=\; \frac{\gamma}{2}.$$

[4] In fact, γ is the *unique* positive root. See Fact A.4.

[5] γ is the smallest value such that Brownian motion almost surely has a two-sided γ-slow point [35]. We will not use this fact.

The proof of this theorem has two parts: a lower bound, in Section 4, which exhibits an optimal randomized adversary, and an upper bound, in Section 3, which exhibits an optimal algorithm. The algorithm is very short, and it appears below in Algorithm 1. Remarkably, the quantity γ arises in both the lower bound and upper bound for seemingly unrelated reasons. In the lower bound γ is the maximizer in (4.3), and in the upper bound γ is the minimizer in (5.14).

Remark. Our lower bound can be strengthened to show that, for any algorithm $\mathcal{A}$,

$$\sup_{\mathcal{B}} \limsup_{t \geq 1} \frac{\mathrm{Regret}(2, t, \mathcal{A}, \mathcal{B})}{\sqrt{t}} \geq \frac{\gamma}{2}.$$

In particular, even if $\mathcal{A}$ is granted a "warm-up" period during which its regret is ignored, an adversary can still force it to incur large regret afterwards. A sketch of this is in Appendix E.1.

The algorithm's description and analysis heavily relies on a function $R\colon \mathbb{R}_{\geq 0} \times \mathbb{R} \to \mathbb{R}$ defined by

$$R(t, g) = \begin{cases} 0 & (t = 0) \\ \frac{g}{2} + \kappa\sqrt{t} \cdot M_0\left(g^2/2t\right) & (t > 0 \text{ and } g \leq \gamma\sqrt{t}) \\ \frac{\gamma\sqrt{t}}{2} & (t > 0 \text{ and } g \geq \gamma\sqrt{t}) \end{cases} \qquad \text{where} \qquad \kappa = \frac{1}{\sqrt{2\pi}\,\mathrm{erfi}(\gamma/\sqrt{2})} \tag{2.5}$$

and M_0 is defined in (2.3). The function R may seem mysterious at first, but in fact arises naturally from the solution to a stochastic calculus problem in Section 5. In our usage of this function, t will correspond to the time and g will correspond to the *gap* between (i.e., absolute difference of) the total loss for the two experts. One may verify that R is continuous on $\mathbb{R}_{>0} \times \mathbb{R}$ because the second and third cases agree on the curve $\left\{ (t, \gamma\sqrt{t}) : t > 0 \right\}$ since γ satisfies $M_0(\gamma^2/2) = 0$. We next define the function p to be

$$p(t, g) = \tfrac{1}{2}\big(R(t, g+1) - R(t, g-1)\big), \tag{2.6}$$

which is the discrete derivative of R at time t and gap g. It will be shown later that $p(t, g) \in [0, 1/2]$ whenever $t \geq 1$ and $g \geq 0$. The algorithm constructs its distribution x_t so that $p(t, g)$ is the probability mass assigned to the worst expert. We remark that $p(t, 0) = 1/2$ (Lemma 3.3) for all $t \geq 1$ so that when both experts are equally good, the algorithm places equal mass on both experts.

An interesting observation is that Algorithm 1 assigns *zero* probability mass to the worst expert when the gap is sufficiently large. Specifically, $p(t, g_{t-1}) = 0$ whenever $g_{t-1} \geq \gamma\sqrt{t} + 1$. Typical[6] regret minimization algorithms, such as MWU, do not have this property.

Algorithm 1 An algorithm achieving the minimax anytime regret for two experts. It is assumed that each cost vector $\ell_t \in [0, 1]^2$.

```
Initialize L_0 ← [0, 0].
for t = 1, 2, ... do
    If necessary, swap indices so that L_{t-1,1} ≥ L_{t-1,2}.
    The current gap is g_{t-1} ← L_{t-1,1} − L_{t-1,2}.
    Set x_t ← [ p(t, g_{t-1}), 1 − p(t, g_{t-1}) ], where p is the function defined by (2.6).
    ▷ Observe cost vector ℓ_t and incur expected cost ⟨ x_t, ℓ_t ⟩.
    L_t ← L_{t-1} + ℓ_t
end for
```

[6] We are aware of two other algorithms that do have this property: Feder et al. [21, Eq. (9)] and Foster and Vohra [23].

2.3 Techniques

Lower Bound. The common approach to prove lower bounds in the experts problem is to consider a random adversary that changes the gap by ± 1 at each step and to consider the regret at a fixed time T. Although we do consider a random adversary, looking at a fixed time T will not be able to yield a good lower bound. The first key idea is to replace the fixed time with a suitable *stopping time*. In particular, the stopping time we use is the first time that the gap process (which is evolving as a reflected random walk) crosses a $c\sqrt{t}$ boundary where $c > 0$ is a constant to be optimized.

To analyze this, we use an elementary identity known as Tanaka's formula for random walks that allows us to write the regret process as $\mathrm{Regret}(t) = Z_t + g_t/2$ where Z_t is a martingale with $Z_0 = 0$ and g_t is the current gap at time t. At this point, it might seem we are ready to apply the optional stopping theorem, which states that if we have a stopping time τ then $\mathrm{E}[\, Z_\tau \,] = Z_0 = 0$. In particular, by choosing τ as the first time that the gap g_t exceeds the $c\sqrt{t}$ boundary, one might expect that $\mathrm{E}[\, \mathrm{Regret}(\tau) \,] = \mathrm{E}[\, g_\tau \,]/2 \geq \mathrm{E}[\, c\sqrt{\tau} \,]/2$. Unfortunately, the argument cannot be so simple since the adversary is allowed to choose $c > 0$ and, by taking c sufficiently large, it would violate known upper bounds on the regret.

The issue lies in the fact that the optional stopping theorem requires certain conditions on the martingale and stopping time. It turns out that the conditions used in most textbooks are too weak for us to derive the optimal regret bound. Fortunately there is a strengthening of the optional stopping theorem that leads to optimal results in our setting. Namely, if Z_t is a martingale with bounded increments (i.e. $\sup_{t \geq 0} |Z_{t+1} - Z_t| \leq K$ for some $K > 0$) and τ is a stopping time satisfying $\mathrm{E}[\, \sqrt{\tau} \,] < \infty$ then $\mathrm{E}[\, Z_\tau \,] = 0$. (The crucial detail is the square root.) This result is stated formally in Theorem 4.2. The question is now to choose as large a boundary as possible such that the associated stopping time of hitting the boundary satisfies $\mathrm{E}[\, \sqrt{\tau} \,] < \infty$. Using classical results of Breiman [4] and Greenwood and Perkins [26], we will show that the optimal choice of c is γ.

Upper Bound. Our analysis of the upper bound uses a fairly standard, undergraduate-style potential function argument with the function R defined in (2.5) as the potential. Specifically, we show that the change in regret from time $t-1$ and gap g_{t-1} to time t and gap g_t is at most $R(t, g_t) - R(t-1, g_{t-1})$. This implies that $\max_g R(t, g)$ is an upper bound on the regret at time t. It is not difficult to see that $R(t, g) \leq \gamma\sqrt{t}/2$ for all $t \geq 0$, which establishes our main upper bound. One interesting twist is that our potential function is bivariate: it depends both on the state g of the algorithm and on time t. To capture how the potential's evolution depends on time, we use a simple identity known as the discrete Itô formula.

The function R and the use of discrete Itô do not come "out of thin air"; both of these ideas come from considering a continuous-time analogue of the problem. The reason for taking this continuous viewpoint is that it brings a wealth of analytical tools that may not exist (or are more cumbersome) in the discrete setting. In order to formulate the continuous-time problem, we will assume that the continuous adversary evolves the gap between the best and worst expert as a reflected Brownian motion. This assumption is motivated by the discrete-time lower bound, since Brownian motion is the continuous-time analogue of a random walk. Using this adversary, the continuous-time regret becomes a stochastic integral.

An important tool at our disposal is the (continuous) Itô formula (Theorem 5.3), which provides an insightful decomposition of the continuous-time regret. This decomposition suggests that the algorithm should satisfy an analytic condition known as the *backwards heat equation*. A key resulting idea is: if the algorithm satisfies the backward heat equation, then there is a natural potential function that upper bounds the regret of the algorithm. This affords us a systematic approach to obtain an explicit continuous-time algorithm and a potential function that bounds the continuous algorithm's regret. Remarkably, this same potential function can be used to analyze the discrete-time algorithm with exactly the same regret bound.

2.4 An application

Let us consider in more detail the problem of predicting a binary sequence, as discussed by Feder et al. [21]. At each time step $t \geq 1$, the algorithm must randomly predict whether the next bit is a 0 or a 1, and the adversary chooses the bit's true value b_t. For any finite sequence $b \in \{0,1\}^*$, let $\pi_s(b)$ be the smallest fraction of errors achieved by any s-state predictor (that may be chosen with knowledge of b). Let $\hat{\pi}(b)$ denote the expected fraction of errors achieved by some online algorithm (or "universal sequential predictor"), whose behavior is independent of $|b|$.

The main objective of Feder et al. is to study algorithms for which $\hat{\pi}(b)$ approximates $\pi_s(b)$. In particular, their Theorem 1 describes an algorithm for which $\hat{\pi}(b) - \pi_1(b) \leq 1/\sqrt{t} + 1/t$ for all $b \in \{0,1\}^*$, where $t = |b|$. They then build on this result to approximate any s-state predictor. They appear to have made efforts to optimize the constant multiplying $1/\sqrt{t}$; see remark 2 on page 1260 and the final paragraph of their Appendix A. We determine the optimal convergence rate for the problem considered by Feder et al.

Theorem 2.2. There is an algorithm achieving $\hat{\pi}(b) - \pi_1(b) \leq \gamma/2\sqrt{t}$ for all $b \in \{0,1\}^*$, where $t = |b|$. Moreover, no algorithm can achieve such a guarantee with a constant smaller than $\gamma/2$.

Proof sketch. The universal sequential prediction problem reduces easily to the problem of bounding anytime regret for prediction with two experts. Intuitively, one expert always predicts that the next bit is 0, whereas the other expert always predicts that it is 1. The adversary chooses a cost vector $[0,1]$ or $[1,0]$ to indicate which expert's prediction was correct. The quantity $t \cdot \pi_1(b)$ equals the cost of the best expert, and $t \cdot \hat{\pi}(b)$ equals the cost of the algorithm, so $t \cdot (\hat{\pi}(b) - \pi_1(b))$ equals the regret. If Algorithm 1 is used for the random prediction, then Theorem 2.1 implies the first statement of the theorem.

Conversely, for any sequential predictor, we may use our adversaries from the proof of Theorem 2.1 to generate the binary sequence (since they only use cost vectors $[0,1]$ or $[1,0]$). For any $\epsilon > 0$, there is an adversary that ensures that the regret is at least $(\gamma - \epsilon)\sqrt{t}/2$ at some time t. It follows that there exists $b \in \{0,1\}^*$ for which $\hat{\pi}(b) - \pi_1(b) \geq (\gamma - \epsilon)/2\sqrt{t}$. Taking $\epsilon \to 0$, the second statement follows. □

2.5 An expression for the regret involving the gap

In our two-expert prediction problem, the most important scenario restricts each cost vector ℓ_t to be either $[0,1]$ or $[1,0]$. This restricted scenario is equivalent to the condition $g_t - g_{t-1} \in \{\pm 1\}$ $\forall t \geq 1$, where $g_t := |L_{t,1} - L_{t,2}|$ is the gap at time t. To prove the optimal lower bound it suffices to consider this restricted scenario. The optimal upper bound will first be proven in the restricted scenario, then extended to general cost vectors in Appendix D. With the sole exception of Appendix D, we will assume the restricted scenario.

We now present an expression, valid for any algorithm, that emphasizes how the regret depends on the *change* in the gap. This expression will be useful in proving both the upper and lower bounds. Henceforth we will often write $\mathrm{Regret}(t) := \mathrm{Regret}(2, t, \mathcal{A}, \mathcal{B})$ where $\mathcal{A}$ and $\mathcal{B}$ are usually implicit from the context.

Proposition 2.3. Assume the restricted setting in which $g_t - g_{t-1} \in \{\pm 1\}$ for every $t \geq 1$. When $g_{t-1} \neq 0$, let p_t denote the probability mass assigned by the algorithm to the worst expert; this quantity may depend arbitrarily on $\ell_1, \ldots, \ell_{t-1}$. Then

$$\mathrm{Regret}(T) \;=\; \sum_{t=1}^{T} p_t \cdot (g_t - g_{t-1}) \cdot \mathbf{1}[g_{t-1} \neq 0] \;+\; \sum_{t=1}^{T} \langle\, x_t,\, \ell_t \,\rangle \cdot \mathbf{1}[g_{t-1} = 0]. \tag{2.7}$$

Remark. If the cost vectors are randomly chosen so that the gap process $(g_t)_{t \geq 0}$ is the absolute value of a standard random walk, then (2.7) is the Doob decomposition [29, Theorem 10.1] of the regret process $\big(\mathrm{Regret}(t)\big)_{t \geq 0}$, i.e., the first sum is a martingale and the second sum is an increasing predictable process.

Proof. Define $\Delta_{\mathrm{R}}(t) = \mathrm{Regret}(t) - \mathrm{Regret}(t-1)$. The total cost of the best expert at time t is $L_t^* := \min\{L_{t,1}, L_{t,2}\}$. The change in regret at time t is the cost incurred by the algorithm minus the change in the total cost of the best expert, so $\Delta_{\mathrm{R}}(t) = \langle x_t, \ell_t \rangle - (L_t^* - L_{t-1}^*)$.

Case 1: $g_{t-1} \neq 0$. In this case, the best expert at time $t-1$ remains a best expert at time t. If the worst expert incurs cost 1, then the algorithm incurs cost p_t and the best expert incurs cost 0, so $\Delta_{\mathrm{R}}(t) = p_t$ and $g_t - g_{t-1} = 1$. Otherwise, the best expert incurs cost 1 and the algorithm incurs cost $1 - p_t$, so $\Delta_{\mathrm{R}}(t) = -p_t$ and $g_t - g_{t-1} = -1$. In both cases, $\Delta_{\mathrm{R}}(t) = p_t \cdot (g_t - g_{t-1})$.

Case 2: $g_{t-1} = 0$. Both experts are best, but one incurs no cost, so $L_t^* = L_{t-1}^*$ and $\Delta_{\mathrm{R}}(t) = \langle x_t, \ell_t \rangle$. □

3 Upper bound

In this section, we prove the upper bound in Theorem 2.1 via a sequence of simple steps. We remind the reader that for simplicity, we will assume that the gap changes by ± 1 at each step, which corresponds to the loss vectors $\ell_t \in \{[0,1],[1,0]\}$. The analysis can be extended to general loss vectors in $[0,1]^2$ through the use of concavity arguments. The details of this extension are not particularly enlightening, so we relegate them to Appendix D.

The proof in this section uses the potential function R which, as explained in Subsection 2.3, is defined via continuous-time arguments in Section 5. Moreover, the structure of the proof is heavily inspired by the proof in the continuous setting. Finally, we remark that the analysis of this section uses the potential function in a modular way[7], and could conceivably be used to analyze other algorithms.

Moving forward, we will need a few observations about the functions R and p, which were defined in equations (2.5) and (2.6).

Lemma 3.1. For any $t > 0$, $R(t,g)$ is concave and non-decreasing in g.

The proof of Lemma 3.1 is a calculus exercise and appears in Appendix C.1. As a consequence, we can easily get the maximum value of $R(t,g)$ for any t.

Lemma 3.2. For any $t > 0$, we have $R(t,g) \leq \gamma\sqrt{t}/2$.

Proof. Lemma 3.1 shows that $R(t,g)$ is non-decreasing in g. By definition, $R(t,g)$ is constant for $g \geq \gamma\sqrt{t}$. It follows that $\max_g R(t,g) \leq R(t,\gamma\sqrt{t}) = \gamma\sqrt{t}/2$. □

In the definition of the prediction task, the algorithm must produce a probability vector x_t. Recalling the definition of x_t in Algorithm 1, it is not a priori clear whether x_t is indeed a probability vector. We now verify that it is, since Lemma 3.3 implies that $p(t,g) \in [0,1/2]$ for all t,g.

Lemma 3.3. Fix $t \geq 1$. Then
(1) $p(t,0) = 1/2$;
(2) $p(t,g)$ is non-increasing in g; and
(3) $p(t,g) \geq 0$.

Proof. For the first assertion, we have

$$p(t,0) = \frac{1}{2}(R(t,1) - R(t,-1)) = \frac{1}{2}\left(\frac{1}{2} + \kappa\sqrt{t}M_0(1/2t) + \frac{1}{2} - \kappa\sqrt{t}M_0(1/2t)\right) = \frac{1}{2}.$$

[7] Our analysis may also be viewed as an amortized analysis. With this viewpoint, the algorithm incurs amortized regret at most $\frac{\gamma}{2}(\sqrt{t} - \sqrt{t-1}) \approx \gamma/4\sqrt{t}$ at each time step t.

For the second equality, we used that $1 \le \gamma \le \gamma\sqrt{t}$ for all $t \ge 1$. The second assertion follows from concavity of R, which was shown in Lemma 3.1, and an elementary property of concave functions (Fact A.6). The final assertion holds because R is non-decreasing in g, which was also shown in Lemma 3.1. □

3.1 Analysis when gap increments are ± 1

In this subsection we prove the upper bound of Theorem 2.1 for a restricted class of adversaries (that nevertheless capture the core of the problem). The analysis is extended to all adversaries in Appendix D.

Theorem 3.4. Let $\mathcal{A}$ be the algorithm described in Algorithm 1. For any adversary $\mathcal{B}$ such that each cost vector ℓ_t is either $[0,1]$ or $[1,0]$, we have

$$\sup_{t \ge 1} \frac{\mathrm{Regret}(2, t, \mathcal{A}, \mathcal{B})}{\sqrt{t}} \;\le\; \frac{\gamma}{2}.$$

Our analysis will rely on an identity known as the discrete Itô formula, which is the discrete analogue of Itô's formula from stochastic analysis (see Theorem 5.3). To make this connection (in addition to future connections) more apparent, we define the discrete derivatives of a function f to be

$$\begin{aligned} f_g(t,g) &= \frac{f(t,g+1) - f(t,g-1)}{2}, \\ f_t(t,g) &= f(t,g) - f(t-1,g), \\ f_{gg}(t,g) &= \big(f(t,g+1) + f(t,g-1)\big) - 2f(t,g). \end{aligned}$$

It was remarked earlier that $p(t,g)$ is the discrete derivative of R, and this is because

$$p(t,g) \;=\; R_g(t,g). \tag{3.1}$$

Lemma 3.5 (Discrete Itô formula). Let $g_0, g_1, \ldots$ be a sequence of real numbers satisfying $|g_t - g_{t-1}| = 1$. Then for any function f and any fixed time $T \ge 1$, we have

$$f(T, g_T) - f(0, g_0) \;=\; \sum_{t=1}^{T} f_g(t, g_{t-1}) \cdot (g_t - g_{t-1}) \;+\; \sum_{t=1}^{T} \left(\frac{1}{2} f_{gg}(t, g_{t-1}) + f_t(t, g_{t-1}) \right). \tag{3.2}$$

This lemma is a small generalization of [29, Example 10.9] to accommodate a bivariate function f that depends on t. The proof is essentially identical, and appears in Appendix C.2 for completeness.

Now we show how the regret has a formula similar to (3.2). Recall that Lemma 3.3(1) guarantees $p(t,0) = 1/2$. Plugging this into (2.7) with $x_t = [1/2, 1/2]$ and $\ell_t \in \{[0,1],[1,0]\}$, the last summation in (2.7) becomes $\sum_t \mathbf{1}[g_{t-1} = 0]/2$. This equals $\sum_t p(t,0) \cdot (g_t - g_{t-1}) \mathbf{1}[g_{t-1} = 0]$, since $g_t = 1$ on the event $g_{t-1} = 0$. Collapsing the two sums in (2.7) into one, we derive that

$$\mathrm{Regret}(T) = \sum_{t=1}^{T} p(t, g_{t-1}) \cdot (g_t - g_{t-1}) = \sum_{t=1}^{T} R_g(t, g_{t-1}) \cdot (g_t - g_{t-1}), \tag{3.3}$$

where $g_0 = 0$ and $g_t \ge 0$ for all $t \ge 1$. Note that the difference between (3.3) and (3.2) is the quantity $\frac{1}{2} f_{gg}(t, g_{t-1}) + f_t(t, g_{t-1})$. In the continuous setting, we will see that a key idea is to try to obtain a solution satisfying $(\frac{1}{2}\partial_{gg} + \partial_t) f = 0$; this is the well-known backwards heat equation. In the discrete setting, we will show that $\frac{1}{2} f_{gg}(t, g_{t-1}) + f_t(t, g_{t-1}) \ge 0$ which suffices for our purposes.

Lemma 3.6 (Discrete backwards heat inequality). $\frac{1}{2} R_{gg}(t,g) + R_t(t,g) \ge 0$ for all $t \in \mathbb{R}_{\ge 1}$ and $g \in \mathbb{R}_{\ge 0}$.

This lemma is the most technical part of the discrete analysis. Its proof appears in Appendix C.3. We now have all the ingredients needed to prove our main theorem (in the present special case).

Proof (of Theorem 3.4). Apply Lemma 3.5 to the function R and the sequence $g_0, g_1, \ldots$ of (integer) gaps produced by the adversary $\mathcal{B}$. Then, for any time $T \geq 0$,

$$\begin{aligned}
&R(T, g_T) - R(0, g_0) \\
&= \sum_{t=1}^{T} R_g(t, g_{t-1}) \cdot (g_t - g_{t-1}) + \sum_{t=1}^{T} \Big(\frac{1}{2} R_{gg}(t, g_{t-1}) + R_t(t, g_{t-1}) \Big) && \text{(by Lemma 3.5)} \\
&\geq \sum_{t=1}^{T} p(t, g_{t-1}) \cdot (g_t - g_{t-1}) && \text{(by (3.1) and Lemma 3.6)} \\
&= \mathrm{Regret}(T) && \text{(by (3.3))}.
\end{aligned}$$

Since $g_0 = 0$ and $R(0, 0) = 0$, applying Lemma 3.2 shows that $\mathrm{Regret}(T) \leq R(T, g_T) \leq \gamma\sqrt{T}/2$. □

The reader at this point, may be wondering why γ is the right constant to appear in the analysis. In Section 5, we will define the function R specifically to obtain γ in the preceding analysis. In the next section, our matching lower bound will prove that γ is indeed the right constant.

4 Lower bound

The main result of this section is the following theorem, which implies the lower bound in Theorem 2.1.

Theorem 4.1. For any algorithm $\mathcal{A}$ and any $\epsilon > 0$, there exists an adversary $\mathcal{B}_\epsilon$ such that

$$\sup_{t \geq 1} \frac{\mathrm{Regret}(2, t, \mathcal{A}, \mathcal{B}_\epsilon)}{\sqrt{t}} \;\geq\; \frac{\gamma - \epsilon}{2}. \tag{4.1}$$

As remarked earlier, the sup can be replaced by a lim sup, see Appendix E.1.

It is common in the literature for regret lower bounds to be proven by random adversaries; see, e.g., [12, Theorem 3.7]. We will also consider a random adversary, but the novelty is the use of a non-trivial stopping time at which it can be shown that the regret is large.

A random adversary. Suppose an adversary produces a sequence of cost vectors $\ell_1, \ell_2, \ldots \in \{0, 1\}^2$ as follows. For all $t \geq 1$,

- If $g_{t-1} > 0$ then ℓ_t is randomly chosen to be one of the vectors $[1, 0]$ or $[0, 1]$, uniformly and independent of $\ell_1, \ldots, \ell_{t-1}$. Thus $g_t - g_{t-1}$ is uniform in $\{\pm 1\}$.
- If $g_{t-1} = 0$ then $\ell_t = [1, 0]$ if $x_{t,1} \geq 1/2$, and $\ell_t = [0, 1]$ if $x_{t,2} > 1/2$. In both cases $g_t = 1$.

As remarked above, the process $(g_t)_{t \geq 0}$ has the same distribution as the absolute value of a standard random walk (which is also known as a reflected random walk).

We now obtain from (2.7) a lower bound on the regret of any algorithm against this adversary. The adversary's behavior when $g_{t-1} = 0$ ensures that $\langle\, x_t,\, \ell_t \,\rangle \geq 1/2$, showing that

$$\mathrm{Regret}(T) \;\geq\; \underbrace{\sum_{t=1}^{T} p_t\,(g_t - g_{t-1}) \cdot \mathbf{1}[g_{t-1} \neq 0]}_{\text{martingale}} \;+\; \underbrace{\frac{1}{2} \sum_{t=1}^{T} \mathbf{1}[g_{t-1} = 0]}_{\text{local time}} \qquad \forall T \in \mathbb{N}.$$

(Equality holds if the algorithm sets $x_t = [1/2, 1/2]$ whenever $g_{t-1} = 0$.) The first sum is a martingale indexed by t. (This holds because $g_t - g_{t-1}$ has conditional expectation 0 when $g_{t-1} \neq 0$, and $\mathbf{1}[g_{t-1} \neq 0] =$

0 when $g_{t-1} = 0$.) The second sum is called the local time of the random walk. Using Tanaka's formula [29, Ex. 10.8], the local time can be written as $\sum_{t=1}^{T} \mathbf{1}[g_{t-1} = 0] = g_t - Z'_t$ where Z'_t is a martingale with uniformly bounded increments and $Z'_0 = 0$. Thus, combining the two martingales, we have

$$\mathrm{Regret}(t) \;\geq\; Z_t + \frac{g_t}{2} \qquad \forall t \in \mathbb{Z}_{\geq 0}, \tag{4.2}$$

where Z_t is a martingale with uniformly bounded increments and $Z_0 = 0$.

Intuition for a stopping time. Optional stopping theorems assert that, under some hypotheses, the expected value of a martingale at a stopping time equals the value at the start. Using such a theorem, at a stopping time τ it would hold that $\mathrm{E}\,[\,\mathrm{Regret}(\tau)\,] \geq \mathrm{E}\,[\,g_\tau\,]/2$ (under some hypotheses on τ and Z). Thus it is natural to design a stopping time τ that maximizes $\mathrm{E}\,[\,g_\tau\,]$ and satisfies the hypotheses. We know from (2.2) that the optimal anytime regret at time t is $\Theta(\sqrt{t})$, so one reasonable stopping time would be

$$\tau(c) \;:=\; \min\Big\{\; t > 0 \;:\; g_t \geq c\sqrt{t}\;\Big\}$$

for some constant c yet to be determined. If $\tau(c)$ and Z satisfy the hypotheses of the optional stopping theorem, then it will hold that $\mathrm{E}\,[\,\mathrm{Regret}(\tau(c))\,] \geq \frac{c}{2}\,\mathrm{E}[\,\sqrt{\tau(c)}\,]$. From this, it follows, fairly easily, that $\mathrm{AnytimeNormRegret}(2) \geq c/2$; this will be argued more carefully later.

An optional stopping theorem. The optional stopping theorems appearing in standard references require one of the following hypotheses: (i) τ is almost surely bounded, or (ii) $\mathrm{E}\,[\,\tau\,]$ is bounded and the martingale has bounded increments, or (iii) the martingale is almost surely bounded and τ is almost surely finite. See, e.g., [5, Theorem 5.33], [19, Theorem 4.8.5], [29, Theorem 10.11], [27, Theorem 12.5.1], [37, Theorem II.57.4], or [42, Theorem 10.10]. These will not suffice for our purposes, and we will require the following theorem, which has a weaker hypothesis (due to the square root). We are unable to find a reference for this theorem, although it is presumably folklore, so we provide a proof in Appendix E.

Theorem 4.2. Let Z_t be a martingale and $K > 0$ a constant such that $|Z_t - Z_{t-1}| \leq K$ almost surely for all t. Let τ be a stopping time. If $\mathrm{E}\,[\,\sqrt{\tau}\,] < \infty$ then $\mathrm{E}\,[\,Z_\tau\,] = \mathrm{E}\,[\,Z_0\,]$.

Optimizing the stopping time. Since the martingale Z_t defined above has bounded increments, Theorem 4.2 may be applied so long as $\mathrm{E}[\,\sqrt{\tau(c)}\,] < \infty$, in which case the preceding discussion yields $\mathrm{AnytimeNormRegret}(2) \geq c/2$. So it remains to determine

$$\sup\{\; c \geq 0 \;:\; \mathrm{E}[\,\sqrt{\tau(c)}\,] < \infty \;\}, \tag{4.3}$$

where $\tau(c)$ is the first time at which a standard random walk crosses the two-sided boundary $\pm c\sqrt{t}$. We will use the following result, in which M is the confluent hypergeometric function defined in Appendix A. Some discussion of our statement of this theorem appears in Appendix E.

Theorem 4.3 (Breiman [4], Theorem 2). Let $c > 1$ and $a < 0$ be such that c is the smallest positive root of the function $x \mapsto M(a, 1/2, x^2/2)$. Then there exists a constant K such that $\Pr\,[\,\tau(c) > u\,] \sim K u^a$.

Recall the definition of γ in (2.4). For intuition, let us apply Theorem 4.3 with $c = \gamma$, which is defined so that it is the root for $a = -1/2$ (see Eq. (A.2) and Fact A.2). It then follows that

$$\mathrm{E}\left[\,\sqrt{\tau(\gamma)}\,\right] \;=\; \int_0^\infty \Pr\left[\,\sqrt{\tau(\gamma)} > s\,\right]\,\mathrm{d}s \;=\; \int_0^\infty \Pr\left[\,\tau(\gamma) > s^2\,\right]\,\mathrm{d}s \;\sim\; K\int_0^\infty s^{-1}\,\mathrm{d}s,$$

by Theorem 4.3. This integral is infinite, so Theorem 4.2 cannot be applied to $\tau(\gamma)$. However, the integral is on the cusp of being finite. By slightly decreasing a below $-1/2$, and slightly modifying c to be the new root, we should obtain a finite integral, showing that $\mathrm{E}[\,\sqrt{\tau(c)}\,]$ is finite. The following proof uses analytic properties of M to show that this is possible.

Proof (of Theorem 4.1). Fix any $\epsilon > 0$ that is sufficiently small. Consider the random adversary and the stopping times $\tau(c)$ described above. By Claim A.5, there exists $a_\epsilon \in (-1, -1/2)$ and $c_\epsilon \geq \gamma - \epsilon$ such that c_ϵ is the unique positive root of $z \mapsto M(a_\epsilon, 1/2, z^2/2)$. As in the above calculations, Theorem 4.3 shows that

$$\mathrm{E}\left[\sqrt{\tau(c_\epsilon)}\right] \;=\; \int_0^\infty \Pr\left[\tau(c_\epsilon) > s^2\right] \mathrm{d}s \;\sim\; K \int_0^\infty s^{2a_\epsilon}\,\mathrm{d}s \;<\; \infty, \tag{4.4}$$

since $a_\epsilon < -1/2$. It follows that $\tau(c_\epsilon)$ is almost surely finite, and therefore $\mathrm{Regret}(\tau(c_\epsilon))$ and $g_{\tau(c_\epsilon)}$ are almost surely well defined. Applying Theorem 4.2 to the martingale Z_t appearing in (4.2), we obtain that

$$\mathrm{E}\left[\mathrm{Regret}(\tau(c_\epsilon))\right] \;\geq\; \frac{1}{2}\mathrm{E}\left[g_{\tau(c_\epsilon)}\right] \;=\; \frac{1}{2}\mathrm{E}\left[c_\epsilon\sqrt{\tau(c_\epsilon)}\right].$$

By the probabilistic method, there exists a finite sequence of cost vectors $\ell_1, \ldots, \ell_t$ (depending on $\mathcal{A}$ and ϵ) for which the regret of $\mathcal{A}$ at time t is at least $c_\epsilon\sqrt{t}/2$. The adversary $\mathcal{B}_\epsilon$ (which knows $\mathcal{A}$) provides this sequence of cost vectors to algorithm $\mathcal{A}$, thereby proving (4.1). □

5 Derivation of a continuous-time analogue of Algorithm 1

The purpose of this section is to show how the potential function R defined in (2.5) arises naturally as the solution of a stochastic calculus problem. The derivation of that function is accomplished by defining, then solving, an analogue of the regret minimization problem in continuous time. The main advantage of considering this continuous setting is the wealth of analytic methods available, such as stochastic calculus.

5.1 Defining the continuous regret problem

Continuous time regret problem. The continuous regret problem is inspired by (3.3). Notice that, when the adversary chooses costs in $\{[0,1],[1,0]\}$, the sequence of gaps $g_0, g_1, g_2, \ldots$ live in the support of a reflected random walk. The goal in the discrete case is to find an algorithm p that bounds the regret over all possible sample paths of a reflected random walk. In continuous time it is natural to consider a stochastic integral with respect to reflected Brownian motion instead. Our goal now is to find a continuous-time algorithm whose regret is small for almost all reflected Brownian motion paths.

Definition 5.1 (Continuous Regret). Let $p : \mathbb{R}_{>0} \times \mathbb{R}_{\geq 0} \to [0,1]$ be a continuous function that satisfies $p(t,0) = 1/2$ for every $t > 0$. Let B_t be a standard one-dimensional Brownian motion. Then, the *continuous regret* of p with respect to B is the stochastic integral

$$\mathrm{ContRegret}(T, p, B) \;=\; \int_0^T p(t, |B_t|)\,\mathrm{d}\,|B_t|\,. \tag{5.1}$$

Remark. The condition $p(t,0) = 1/2$ is due to (5.1) being inspired by (3.3), which requires this condition. It is also possible to show that assuming $p(t,0) = 1/2$ in lieu of $p(t,0) \geq 1/2$ is without loss of generality.

In this definition we may think of p as a continuous-time algorithm and B as a continuous-time adversary. The goal for the remainder of this section is to prove the following result.

Theorem 5.2. There exists a continuous-time algorithm p^* such that

$$\mathrm{ContRegret}(T, p^*, B) \;\leq\; \frac{\gamma\sqrt{T}}{2} \quad \forall T \in \mathbb{R}_{\geq 0},\ \text{almost surely}. \tag{5.2}$$

5.2 Connections to stochastic calculus and the backward heat equation

Since $\mathrm{ContRegret}(T)$ evolves as a stochastic integral with respect to a semi-martingale[8] (namely reflected Brownian motion), Itô's lemma provides an insightful decomposition. The following statement of Itô's lemma is a specialization of [36, Theorem IV.3.3] for the special case of reflected Brownian motion.[9]

Notation. Up to now, we have used the symbol g as the second parameter to the bivariate functions p and R. Henceforth, it will be more consistent with the usual notation in the literature to use x to denote g. We will also use the notation $C^{1,2}$ to denote the class of bivariate functions that are continuously differentiable in their first argument and twice continuously differentiable in their second argument.

Theorem 5.3 (Itô's formula). Let $f\colon \mathbb{R}_{\geq 0} \times \mathbb{R} \to \mathbb{R}$ be $C^{1,2}$. Then, almost surely,

$$f(T, |B_T|) - f(0, |B_0|) \;=\; \int_0^T \partial_x f(t, |B_t|)\,\mathrm{d}\,|B_t| + \int_0^T \Big[\underbrace{\partial_t f(t, |B_t|) + \tfrac{1}{2}\partial_{xx} f(t, |B_t|)}_{=:\overset{*}{\Delta} f(t,|B_t|)}\Big]\,\mathrm{d}t. \tag{5.3}$$

The integrand of the second integral is an important quantity arising in PDEs and stochastic processes (see, e.g., [18, pp. 263]). We will denote it by $\overset{*}{\Delta} f(t,x) := \partial_t f(t,x) + \frac{1}{2}\partial_{xx} f(t,x)$. Some discussion about the statement of Theorem 5.3 appears in Appendix F.7.

Applying Itô's formula to the continuous regret. Comparing (5.1) and (5.3), it is natural to assume that $p = \partial_x f$ for a function f that is $C^{1,2}$ with $f(0,0) = 0$, $\partial_x f \in [0,1]$, and $\partial_x f(t,0) = 1/2$; the latter two conditions are needed for Definition 5.1 to be applicable. Itô's formula then yields

$$\mathrm{ContRegret}(T, p = \partial_x f, B) \;=\; \int_0^T \partial_x f(t, |B_t|)\,\mathrm{d}\,|B_t| \;=\; f(T, |B_T|) - \int_0^T \overset{*}{\Delta} f(t, |B_t|)\,\mathrm{d}t. \tag{5.4}$$

Path independence and the backward heat equation. At this point a useful idea arises: as a thought experiment, suppose that $\overset{*}{\Delta} f = 0$. Then the second integral would vanish, and we would have the appealing expression $\mathrm{ContRegret}(T, p, B) = f(T, |B_T|)$. Moreover, since f is a deterministic function, the right-hand side depends only on $|B_T|$ rather than the entire Brownian path $B|_{[0,T]}$. Thus, the same must be true of the left-hand side: at time T, the continuous regret of the algorithm p depends only on T and $|B_T|$ (the gap). We say that say that such an algorithm has *path independent regret.* Our supposition that led to these attractive consequences was only that $\overset{*}{\Delta} f = 0$, which turns out to be a well studied condition.

Definition 5.4. Let $f\colon \mathbb{R}_{>0} \times \mathbb{R} \to \mathbb{R}$ be a $C^{1,2}$ function. If $\overset{*}{\Delta} f(t,x) = 0$ for all $(t,x) \in \mathbb{R}_{>0} \times \mathbb{R}$ then we say that f satisfies the *backward heat equation.* A synonymous statement is that f is *space-time harmonic.*

We may summarize the preceding discussion with the following proposition.

Proposition 5.5. Let $f : \mathbb{R}_{>0} \times \mathbb{R} \to \mathbb{R}$ be a $C^{1,2}$ function that satisfies $\overset{*}{\Delta} f = 0$ everywhere with $f(0,0) = 0$. Let $p = \partial_x f$. Then,

$$\int_0^T p(t, |B_t|)\,\mathrm{d}\,|B_t| \;=\; f(T, |B_t|). \tag{5.5}$$

Suppose that a function f satisfies the hypothesis of Proposition 5.5 *and in addition* $p = \partial_x f \in [0,1]$ with $p(t,0) = 1/2$. Then, we would have

$$\mathrm{ContRegret}(T, p, B) = f(T, |B_T|). \tag{5.6}$$

[8] A semi-martingale is a stochastic process that can written as the sum of a local martingale and a process of finite variation.

[9] Specifically, we are using the statement of Itô's formula that appears in Remark 1 after Theorem IV.3.3 in [36] with $X_t = |B_t|$ and $A_t = t$. Note that y in their notation is t in ours and $\langle |B|, |B| \rangle_t = t$.

We are unable to derive a function that satisfies the properties required for (5.6) to hold along with $\max_{x\geq 0} f(T, |B_T|) \leq \gamma\sqrt{T}/2$. Instead, we will begin by relaxing the constraint that $p(t,x) \in [0,1]$ and allow $p(t,x)$ to be negative. We will overload the notation $\mathrm{ContRegret}(\cdot)$ to include such functions. In the next section, we will derive a family of such functions that all achieve $\mathrm{ContRegret}(T, p, |B_T|) = f(T, |B_T|) = O(\sqrt{T})$. This is done by setting up and solving the backwards heat equation. Next, we use a smoothing argument to obtain a family of functions that all achieve $\mathrm{ContRegret}(T, p, |B_T|) = O(\sqrt{T})$, and that *do* satisfy $p(t,x) \in [0,1]$. Finally, we will optimize $\mathrm{ContRegret}(T, \cdot, |B_T|)$ over this family of functions to prove Theorem 5.2.

5.2.1 Satisfying the backward heat equation

The main result of this section is the derivation of a family of functions $\tilde{p} : \mathbb{R}_{>0} \times \mathbb{R} \to \mathbb{R}$ that satisfy $\tilde{p}(t,x) \leq 1$, $\tilde{p}(t,0) = 1/2$ and

$$\mathrm{ContRegret}(T, \tilde{p}, B) \;=\; f(T, |B_T|) \;=\; O(\sqrt{T}), \tag{5.7}$$

but do not necessarily satisfy $\tilde{p}(t,x) \geq 0$.

The first step is to find a function f which satisfies the partial differential equation $\overset{*}{\Delta} f = 0$. Since the boundary condition $\tilde{p}(t,0) = 1/2$ is a condition on $\tilde{p} = \partial_x f$, not on f itself, it will be convenient to solve a PDE for $\tilde{p}$ instead, and then derive f by integrating. However, some care is needed since not all antiderivates of $\tilde{p}$ (in x) will satisfy the backwards heat equation. Fortunately, we have a useful lemma showing that if $\tilde{p}$ satisfies the backward heat equation, then we can construct an f that also does. This is proven in Appendix F.1.

Lemma 5.6. Suppose that $h : \mathbb{R}_{>0} \times \mathbb{R} \to \mathbb{R}$ is a $C^{1,2}$ function. Define

$$f(t,x) := \int_0^x h(t,y)\,\mathrm{d}y - \frac{1}{2}\int_0^t \partial_x h(s,0)\,\mathrm{d}s.$$

Then,

(1) $f \in C^{1,2}$,
(2) If $\overset{*}{\Delta} h = 0$ over $\mathbb{R}_{>0} \times \mathbb{R}$ then $\overset{*}{\Delta} f = 0$ over $\mathbb{R}_{>0} \times \mathbb{R}$,
(3) $h = \partial_x f$.

Defining boundary conditions for p. Obtaining a particular solution to the backward heat equation requires sufficient boundary conditions in order to uniquely identify $\tilde{p}$. The boundary condition mentioned above is that $\tilde{p}(t,0) = 1/2$ for all t. This condition together with the backward heat equation clearly do not suffice to uniquely determine $\tilde{p}$. Therefore, we impose some reasonable boundary conditions on $\tilde{p}$.

What should the value be at the boundary? Intuitively, $x \mapsto \tilde{p}(t,x)$ should be a decreasing function because $\tilde{p}$ represents the weight placed on the worst expert. Therefore, it is natural to consider an "upper boundary" which specifies the point at which the difference in experts' total costs is so great that the algorithm places zero weight on the worst expert. The upper boundary can be specified by a curve, $\{\, (t, \phi(t)) \,:\, t > 0 \,\}$ for some continuous function $\phi : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$. We will incorporate this idea by requiring $\tilde{p}(t, \phi(t)) = 0$ for all $t > 0$.

Where should the boundary be? One reasonable choice for the boundary is to use $\phi_\alpha(t) = \alpha\sqrt{t}$ for some constant $\alpha > 0$, as this is similar to the boundary used by the random adversary in the lower bound of Section 4. These conditions are combined into the following partial differential equation:

$$\text{(backward heat equation)} \qquad \partial_t u(t,x) + \tfrac{1}{2}\partial_{xx} u(t,x) \;=\; 0 \qquad \text{for all } (t,x) \in \mathbb{R}_{>0} \times \mathbb{R} \tag{5.8}$$

$$\text{(upper boundary)} \qquad u(t, \alpha\sqrt{t}) \;=\; 0 \qquad \text{for all } t > 0 \tag{5.9}$$

$$\text{(lower boundary)} \qquad u(t,0) \;=\; \tfrac{1}{2} \qquad \text{for all } t > 0. \tag{5.10}$$

Next we show that the following function solves this PDE. Define $\tilde{p}_\alpha : \mathbb{R}_{>0} \times \mathbb{R} \to \mathbb{R}$ by

$$\tilde{p}_\alpha(t,x) \;:=\; \frac{1}{2}\left(1 - \frac{\operatorname{erfi}\left(x/\sqrt{2t}\right)}{\operatorname{erfi}\left(\alpha/\sqrt{2}\right)}\right). \tag{5.11}$$

Lemma 5.7. $\tilde{p}_\alpha$ satisfies the following properties:
(1) $\tilde{p}_\alpha$ is $C^{1,2}$ over $\mathbb{R}_{>0} \times \mathbb{R}$,
(2) $\tilde{p}_\alpha$ satisfies the constraints in (5.8), (5.9) and (5.10), and
(3) For all $t > 0$ and all $x \geq 0$, $\tilde{p}_\alpha(t,x) \leq 1/2$.

The proof of Lemma 5.7 appears in Appendix F.2. It shows that $\tilde{p}_\alpha(t,x)$ nearly defines a valid continuous time algorithm, in that it satisfies the conditions of Definition 5.1 except for non-negativity. Next, we will integrate $\tilde{p}_\alpha$ as described in Lemma 5.6. Define the function $\tilde{R}_\alpha \colon \mathbb{R}_{>0} \times \mathbb{R} \to \mathbb{R}$ as

$$\tilde{R}_\alpha(t,x) = \frac{x}{2} + \kappa_\alpha \sqrt{t} \cdot M_0\left(\frac{x^2}{2t}\right) \qquad \text{where} \qquad \kappa_\alpha = \frac{1}{\sqrt{2\pi}\operatorname{erfi}(\alpha/\sqrt{2})}.$$

Lemma 5.8. $\tilde{R}_\alpha(t,x) = \int_0^g \tilde{p}_\alpha(t,y)\ \mathrm{d}y - \frac{1}{2}\int_0^t \partial_g \tilde{p}_\alpha(s,0)\ \mathrm{d}s$.

The proof of Lemma 5.8 appears in Appendix F.3. By Lemma 5.7, the function $\tilde{p}_\alpha$ satisfies the hypothesis of the function h in Lemma 5.6. Hence, we can apply Lemma 5.6 with $h = \tilde{p}_\alpha$ and $f = \tilde{R}_\alpha$ to assert the following properties on $\tilde{R}_\alpha$.

Lemma 5.9. $\tilde{R}_\alpha$ satisfies the following properties:
(1) $\tilde{R}_\alpha$ is $C^{1,2}$,
(2) $\tilde{R}_\alpha$ satisfies $\overset{*}{\Delta} R_\alpha = 0$ over $\mathbb{R}_{>0} \times \mathbb{R}$,
(3) $\partial_x \tilde{R}_\alpha(t,x) = \tilde{p}_\alpha(t,x)$.

Since $\operatorname{erfi}(\cdot)$ is a strictly increasing function with $\operatorname{erfi}(0) = 0$, observe that $\tilde{p}_\alpha$ has exactly one root at $\alpha\sqrt{t}$. Therefore, for every T, we have

$$\operatorname{ContRegret}(T, \tilde{p}_\alpha, B) \;=\; \tilde{R}_\alpha(T, |B_T|) \;\leq\; \max_{x \geq 0} \tilde{R}_\alpha(T,x) \;\leq\; \left(\frac{\alpha}{2} + \kappa_\alpha M_0\left(\frac{\alpha^2}{2}\right)\right)\sqrt{T}.$$

This establishes (5.7), as desired.

5.2.2 Resolving the non-negativity issue

The only remaining step is to modify $\tilde{p}_\alpha$ so that it lies in the interval $[0, 1/2]$. We will modify $\tilde{p}_\alpha$ in the most natural way: by modifying all negative values to be zero. Specifically, we set

$$p_\alpha(t,x) \;:=\; \partial_x R_\alpha(t,x) = \begin{cases} 0 & t = 0 \\ \frac{1}{2}\left(1 - \frac{\operatorname{erfi}(x/\sqrt{2t})}{\operatorname{erfi}(\alpha/\sqrt{2})}\right)_+ & t > 0 \end{cases}. \tag{5.12}$$

Here, we use the notation $(x)_+ = \max\{0, x\}$. Note that $p_\alpha(t,0) = 1/2$ for all $t > 0$ and $p_\alpha \in [0, 1/2]$. So p_α defines a valid continuous-time algorithm. From (5.12), we obtain a truncated version of $\tilde{R}_\alpha$ as

$$R_\alpha(t,x) \;:=\; \begin{cases} 0 & t = 0 \\ \tilde{R}_\alpha(t,x) & t > 0 \wedge x \leq \alpha\sqrt{t} \\ \tilde{R}_\alpha(t, \alpha\sqrt{t}) & t > 0 \wedge x \geq \alpha\sqrt{t} \end{cases}. \tag{5.13}$$

If R_α were sufficiently smooth then we could immediately apply (5.6) (or Theorem 5.3) to obtain a formula for the regret of p_α. The only flaw is that $\partial_{xx} R_\alpha$ is not well-defined on the curve $\{ (t, \alpha\sqrt{t}) : t > 0 \}$ so R_α is not in $C^{1,2}$ and Theorem 5.3 cannot be applied directly. The reader who believes that this issue is unlikely to be problematic may wish to take Lemma 5.10 on faith and skip ahead to Subsection 5.3.

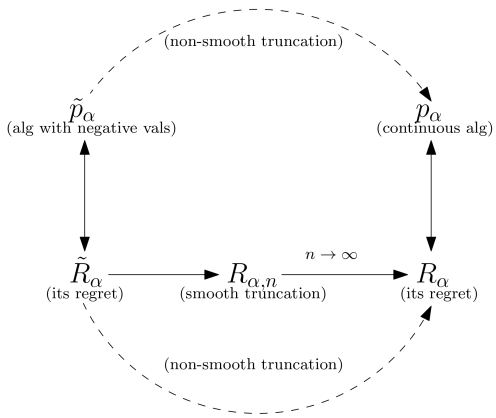

Figure 1: The relationships between $\tilde{p}_\alpha$, $\tilde{R}_\alpha$, $R_{\alpha,n}$, p_α, and R_α

Lemma 5.10. Fix $\alpha > 0$. Then, almost surely, for all $T \geq 0$, $\mathrm{ContRegret}(T, p_\alpha, B) \leq R_\alpha(T, |B_T|)$.

Here, we will present a high-level overview of the proof of this lemma; the details can be found in Appendix F.4. Let $\phi(x)$ be a smooth function satisfying $\phi(x) = 1$ for $x \leq 0$ and $\phi(x) = 0$ for $x \geq 1$. For $n \in \mathbb{N}$, define $\phi_n(x) = \phi(nx)$ and the approximations

$$R_{\alpha,n}(t, x) := \tilde{R}_\alpha(t, x)\phi_n(x - \alpha\sqrt{t}) + \tilde{R}_\alpha(t, \alpha\sqrt{t})(1 - \phi_n(x - \alpha\sqrt{t})).$$

It is relatively straightforward to check that $R_{\alpha,n}(t, x) \xrightarrow{n\to\infty} R_\alpha(t, x)$ pointwise and similarly for the derivatives. The important property is that $R_{\alpha,n}$ is smooth so Itô's formula may be applied. Lemma 5.10 is then proved by taking limits and controlling the error terms.

The remainder of this section proves Theorem 5.2 by setting $p^* = p_\alpha$ for the optimal α.

5.3 Optimizing the boundary to minimize the continuous regret problem

By Lemma 5.10, $\mathrm{ContRegret}(T, \partial_x R_\alpha, B) \leq R_\alpha(T, |B_T|) \leq R_\alpha(T, \alpha\sqrt{T})$, where the last inequality is because $\partial_g R_\alpha(t, g) = p_\alpha(t, g)$ is positive for $g \in [0, \alpha\sqrt{t})$ and 0 for $g \geq \alpha\sqrt{t}$. Define

$$h(\alpha) := R_\alpha(1, \alpha) = \frac{\alpha}{2} + \kappa_\alpha M_0(\alpha^2/2)$$

and note that $R_\alpha(T, \alpha\sqrt{T}) = \sqrt{T} \cdot h(\alpha)$. Thus, the only remaining task is now to solve the following optimization problem.

$$\min_{\alpha>0} h(\alpha) = \min_{\alpha>0} \left\{ \frac{\alpha}{2} + \kappa_\alpha \cdot M_0\left(\frac{\alpha^2}{2}\right) \right\} \tag{5.14}$$

The following lemma verifies that there exists some α for which $\mathrm{ContRegret}(T, \partial_x R_\alpha, B) \leq \frac{\gamma\sqrt{T}}{2}$, completing the proof of Theorem 5.2

Lemma 5.11. Fix $T > 0$. Then $\min_\alpha R_\alpha(T, \alpha\sqrt{T}) = R_\gamma(T, \gamma\sqrt{T}) = \frac{\gamma\sqrt{T}}{2}$.

Lemma 5.11 follows easily from the following claim whose proof appears in Appendix F.6.

Claim 5.12. $h'(\alpha) = -\frac{\exp(\alpha^2/2)}{\pi \operatorname{erfi}(\alpha/\sqrt{2})} \cdot M_0(\alpha^2/2)$. In particular, $h'(\alpha) < 0$ for $\alpha \in (0, \gamma)$, $h'(\gamma) = 0$, and $h'(\alpha) > 0$ for $\alpha \in (\gamma, \infty)$.

Proof of Lemma 5.11. Claim 5.12 implies that γ is the global minimizer for $h(\alpha)$. Therefore, for every $\alpha > 0$, we have $R_\alpha(T, \alpha\sqrt{T}) = \sqrt{T} \cdot h(\alpha) \geq \sqrt{T} \cdot h(\gamma) = R_\gamma(T, \gamma\sqrt{T})$. This proves the first equality. The second equality is because $M_0(\gamma^2/2) = 0$ by definition of γ. $\square$

A Standard facts

A.1 Basic facts about confluent hypergeometric functions

For any $a, b \in \mathbb{R}$ with $b \notin \mathbb{Z}_{\leq 0}$, the confluent hypergeometric function of the first kind is defined as

$$M(a, b, z) = \sum_{n=0}^{\infty} \frac{(a)_n z^n}{(b)_n n!}, \tag{A.1}$$

where $(x)_n \coloneqq \prod_{i=0}^{n-1}(x+i)$ is the Pochhammer symbol. See, e.g., Abramowitz and Stegun [2, Eq. (13.1.2)].

For notational convenience, for $i \in \{0, 1, 2, \ldots, \}$, we write

$$M_i(x) = M(i - 1/2, i + 1/2, x). \tag{A.2}$$

Fact A.1. If $b \notin \mathbb{Z}_{\leq 0}$ then $\frac{\mathrm{d}}{\mathrm{d}x} M(a, b, x) = \frac{a}{b} \cdot M(a+1, b+1, x)$. Consequently,

(1) $M_0'(x) = -M_1(x)$; and

(2) $M_1'(x) = \frac{1}{3} \cdot M_2(x)$.

Proof. See [2, Eq. (13.4.9)]. □

Fact A.2. The following identities hold:

(1) $M_0(x) = -\sqrt{\pi x}\,\mathrm{erfi}(\sqrt{x}) + e^x$.

(2) $M_1(x) = \frac{\sqrt{\pi}\,\mathrm{erfi}(\sqrt{x})}{2\sqrt{x}}$.

(3) $M_2(x) = \frac{3(2e^x\sqrt{x} - \sqrt{\pi}\,\mathrm{erfi}(\sqrt{x}))}{4x^{3/2}}$.

(4) $\frac{2}{3} \cdot M_2(x) \cdot x + M_1(x) = e^x$.

Proof.

(2): See [2], equations (7.1.21) or (13.6.19), and use that $\mathrm{erfi}(x) = -i\,\mathrm{erf}(ix)$, where $i = \sqrt{-1}$.

(1): Differentiating the right-hand side (using the definition of erfi in (2.3)) yields $-\frac{\sqrt{\pi}\,\mathrm{erfi}(\sqrt{x})}{2\sqrt{x}}$. So the right-hand side is an anti-derivative of $-M_1(x)$, by part (2). Thus, the identity (1) follows from Fact A.1(1) and the initial condition $M_0(0) = 1$.

(3): This follows directly by differentiating (2) and Fact A.1(2).

(4): Immediate from (2) and (3). □

Fact A.3. The function $M_0(x)$ is decreasing and concave on $[0, \infty)$.

Remark. In fact, $M_0(x)$ is decreasing and concave on $\mathbb{R}$ but we will not require this fact.

Proof. By Fact A.1, we have $M_0'(x) = -M_1(x)$ and $M_0''(x) = -\frac{1}{3} \cdot M_2(x)$. Note that the coefficients of $M_1(x), M_2(x)$ in their Taylor series are all non-negative. As $x \geq 0$, we have that $M_0'(x), M_0''(x) \leq 0$ as desired. □

Fact A.4. The function $x \mapsto M_0(x^2/2)$ has a unique positive root at $x = \gamma$. Moreover $M_0(x^2/2) > 0$ for $x \in (0, \gamma)$ and $M_0(x^2/2) < 0$ for $x \in (\gamma, \infty)$.

Proof. The Maclaurin expansion of $M_0(x^2/2)$ is given by

$$M_0\left(\frac{x^2}{2}\right) = 1 - \sum_{k=1}^{\infty} \frac{1}{(2k-1)k!} \frac{x^{2k}}{2^k}.$$

Note that $M_0(0) = 1$. It is clear, from the series expansion above (and Fact A.3), that $M_0(x^2/2)$ is strictly decreasing in x on $(0, \infty)$ and $\lim_{x\to\infty} M_0(x^2/2) = -\infty$. Hence, $M_0(x^2/2)$ contains a positive root γ and it is unique. Finally, it is clear that $M_0(x^2/2)$ is positive on $(0, \gamma)$ and negative on (γ, ∞). □

Claim A.5. For any $\epsilon > 0$, there exists $a_\epsilon \in (-1, -1/2)$ such that the smallest[10] positive root c_ϵ of $z \mapsto M(a_\epsilon, 1/2, z^2/2)$ satisfies $c_\epsilon \geq \gamma - \epsilon$.

Proof. Following Perkins' notation [35], let $\lambda_0(-c, c)$ be such that c is the smallest positive root of $x \mapsto M(-\lambda_0(-c, c), 1/2, x^2/2)$. By [35, Proposition 1], the map $c \mapsto \lambda_0(-c, c)$ is strictly decreasing and continuous on $\mathbb{R}_{>0}$, so it has a continuous inverse α. From (2.4) and Fact A.2(1), we see that $\lambda_0(-\gamma, \gamma) = 1/2$, hence $\alpha(1/2) = \gamma$. By continuity, for all $\epsilon > 0$, there exists $\delta \in (0, 1/2)$ such that $\alpha(1/2 + \delta) > \gamma - \epsilon$. Then we may take $a_\epsilon = -(1/2 + \delta)$ and $c_\epsilon = \alpha(1/2 + \delta)$. □

A.2 Other standard facts

Fact A.6. Suppose $f : \mathbb{R} \to \mathbb{R}$ is concave. Then for any $\alpha < \beta$, the function $g(t) = f(t + \beta) - f(t + \alpha)$ is non-increasing.

Fact A.7. Suppose that $f : \mathbb{R} \to \mathbb{R}$ is concave. Let $\alpha < \beta$. Then $f(x) \geq \min\{f(\alpha), f(\beta)\}$ for all $x \in [\alpha, \beta]$.

B Application to random walks

In this section we prove the following result, which was mentioned in Section 1.

Theorem B.1 (Davis [16]). Let $(X_t)_{t \geq 0}$ be a standard random walk. Then $\mathrm{E}\,[\,|X_\tau|\,] \leq \gamma\,\mathrm{E}\,[\,\sqrt{\tau}\,]$ for every stopping time τ; moreover, the constant γ cannot be improved.

Proof. We begin by showing that for every stopping time τ, $\mathrm{E}\,[\,|X_\tau|\,] \leq \gamma\,\mathrm{E}\,[\,\sqrt{\tau}\,]$. Assume that $\mathrm{E}\,[\,\sqrt{\tau}\,] < +\infty$, otherwise there is nothing to show. Let $\mathcal{B}$ be the adversary in the proof of Theorem 4.1 and let $\mathrm{Regret}(T)$ be the regret incurred by Algorithm 1. By (4.2), we may write $\mathrm{Regret}(T) = Z_T + \frac{|X_T|}{2}$ where Z is a martingale with uniformly bounded increments. Then by Theorem 3.4 we have

$$\frac{\gamma}{2}\sqrt{T} \;\geq\; \mathrm{Regret}(T) \;=\; Z_T + \frac{|X_T|}{2},$$

for all T. Therefore, the above inequality also holds at time τ. We obtain our desired result by taking expectation and applying Theorem 4.2 — which says that $\mathrm{E}\,[\,Z_\tau\,] = \mathrm{E}\,[\,Z_0\,]$ when Z is a martingale with bounded increments and $\mathrm{E}\,[\,\sqrt{\tau}\,] < \infty$.

Next, let $c < \gamma$. Then, define by the arguments in the proof of Theorem 4.1, c' such that $\gamma > c' > c$ and $\mathrm{E}\left[\,\sqrt{\tau(c')}\,\right] < \infty$ where $\tau(c') = \left\{\, t > 0 \,:\, |X_t| \geq c'\sqrt{t} \,\right\}$. Then,

$$\mathrm{E}\left[\,\left|X_{\tau(c')}\right|\,\right] \;=\; c'\,\mathrm{E}\left[\,\sqrt{\tau(c')}\,\right] \;>\; c\,\mathrm{E}\left[\,\sqrt{\tau(c')}\,\right]. \qquad \square$$

C Technical results from Section 3

C.1 Proof of Lemma 3.1

The following two lemmas are essentially special cases of Lemma F.1 since $\tilde{R}_\gamma = \tilde{R}$ and $R_\gamma = R$. We restate them here without the subscript for convenience.

Lemma C.1. Consider the function $\tilde{R}(t, g) = \frac{g}{2} + \kappa\sqrt{t}M_0\left(\frac{g^2}{2t}\right)$. Then $\frac{\partial}{\partial g}\tilde{R}(t, g) = \frac{1}{2}\left(1 - \frac{\mathrm{erfi}(g/\sqrt{2t})}{\mathrm{erfi}(\gamma/\sqrt{2})}\right)$.

[10] In fact, there is a unique positive root.

Lemma C.2. $\frac{\partial}{\partial g} R(t,g) = \frac{1}{2}\left(1 - \frac{\operatorname{erfi}(g/\sqrt{2t})}{\operatorname{erfi}(\gamma/\sqrt{2})}\right)_+$.

Proof (of Lemma 3.1). The fact that $R(t,g)$ is non-decreasing in g follows from Lemma C.2. The concavity of $R(t,g)$ (in g) follows from the fact that erfi is non-decreasing, so $\frac{\partial}{\partial g} R(t,g)$ is non-increasing in g. □

C.2 Proof of Lemma 3.5

Proof (of Lemma 3.5). By telescoping, $f(T, g_T) - f(0, g_0) = \sum_{t=1}^{T} f(t, g_t) - f(t-1, g_{t-1})$. Consider a fixed $t \in [T]$. We can write

$$\begin{aligned} f(t,g_t) - f(t,g_{t-1}) = &\left(f(t,g_t) - \frac{f(t,g_{t-1}+1) + f(t,g_{t-1}-1)}{2}\right) \\ &+ \left(\frac{f(t,g_{t-1}+1) + f(t,g_{t-1}-1)}{2} - f(t-1,g_{t-1})\right). \end{aligned} \tag{C.1}$$

For the first bracketed term, by considering the cases $g_t = g_{t-1}+1$ and $g_t = g_{t-1}-1$, we have

$$\begin{aligned} f(t,g_t) - \frac{f(t,g_{t-1}+1) + f(t,g_{t-1}-1)}{2} &= \frac{f(t,g_{t-1}+1) - f(t,g_{t-1}-1)}{2} \cdot (g_t - g_{t-1}) \\ &= f_g(t,g_{t-1}) \cdot (g_t - g_{t-1}). \end{aligned} \tag{C.2}$$

Note that the above step is the only place we used the assumption that $|g_t - g_{t-1}| = 1$. For the second bracketed term, we have

$$\begin{aligned} \frac{f(t,g_{t-1}+1) + f(t,g_{t-1}-1)}{2} - f(t-1,g_{t-1}) &= \frac{f(t,g_{t-1}+1) + f(t,g_{t-1}-1) - 2f(t,g_{t-1})}{2} \\ &\quad + (f(t,g_{t-1}) - f(t-1,g_{t-1})) \\ &= \frac{1}{2} f_{gg}(t,g_{t-1}) + f_t(t,g_{t-1}). \end{aligned}$$

This gives the desired formula. □

C.3 Proof of Lemma 3.6

Lemma C.3. For all $u \in [0, 1/2]$, we have $M_0(u) \geq \sqrt{1-2u}$.

Proof. The Maclaurin expansion of $M_0(u)$ is given by

$$M_0(u) = 1 - \sum_{k=1}^{\infty} \frac{1}{(2k-1)k!} u^k.$$

Note that $\frac{\mathrm{d}^k}{\mathrm{d}x^k}\sqrt{1-2x} = -\frac{(2k-3)!!}{(1-2x)^{(2k-1)/2}}$, where $(n)!!$ denotes the double factorial (note that $(-1)!! = 1$).[11] Hence, the Maclaurin expansion of $\sqrt{1-2u}$ is

$$\sqrt{1-2u} = 1 - \sum_{k=1}^{\infty} \frac{(2k-3)!!}{k!} u^k.$$

It is not hard to verify that $(2k-3)!! \geq \frac{1}{2k-1}$. This implies that $M_0(u) \geq \sqrt{1-2u}$. □

[11] If $n \in \mathbb{Z}_{\geq 0}$, we define $(n)!! = \prod_{k=0}^{\lceil n/2 \rceil - 1}(n-2k)$. If $n \in \mathbb{Z}_{<0}$, we define $(n)!!$ via the recursive relation $(n)!! = \frac{(n+2)!!}{n+2}$ so that $(-1)!! = \frac{(1)!!}{1} = 1$.

Lemma C.4. For all $z \in [0,1)$ and $x \in \mathbb{R}$, we have

$$M_0\left(\frac{(x+z)^2}{2}\right) + M_0\left(\frac{(x-z)^2}{2}\right) \geq 2\sqrt{1-z^2}M_0\left(\frac{x^2}{2(1-z^2)}\right).$$

Proof. Fix $z \in [0,1)$ and consider the function

$$h_z(x) = M_0\left(\frac{(x+z)^2}{2}\right) + M_0\left(\frac{(x-z)^2}{2}\right) - 2\sqrt{1-z^2}M_0\left(\frac{x^2}{2(1-z^2)}\right).$$

Note that $h_z(0) \geq 0$ by applying Lemma C.3 with $u = z^2/2$. We will show that $x = 0$ is the minimizer of h_z which implies the lemma.

Indeed, computing derivatives, we have

$$h_z'(x) = -M_1\left(\frac{(x+z)^2}{2}\right)\cdot(x+z) - M_1\left(\frac{(x-z)^2}{2}\right)\cdot(x-z) + 2M_1\left(\frac{x^2}{2(1-z^2)}\right)\cdot\frac{x}{\sqrt{1-z^2}}.$$

As $h_z'(0) = 0$, $x = 0$ is a critical point of h_z. We will now show that h_z is convex which certifies that $x = 0$ is indeed a minimizer.

To obtain h_z'', we differentiate term-by-term. Let $u = \frac{(x+z)^2}{2}$. Then

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x}M_1\left(\frac{(x+z)^2}{2}\right)\cdot(x+z) &= \frac{M_2\left(\frac{(x+z)^2}{2}\right)\cdot(x+z)^2}{3} + M_1\left(\frac{(x+z)^2}{2}\right) \\
&= \frac{2M_2(u)\cdot u}{3} + M_1(u) \\
&= \frac{2u(2e^u\sqrt{u} - \sqrt{\pi}\,\mathrm{erfi}(\sqrt{u}))}{4u^{3/2}} + \frac{\sqrt{\pi}\,\mathrm{erfi}(\sqrt{u})}{2\sqrt{u}} \\
&= e^u = \exp\left(\frac{(x+z)^2}{2}\right).
\end{aligned}$$

The first equality is by Fact A.1 and the third equality is by identities (2) and (3) in Fact A.2. We can similarly show that

$$\frac{\mathrm{d}}{\mathrm{d}x}M_1\left(\frac{(x-z)^2}{2}\right)\cdot(x-z) = \exp\left(\frac{(x-z)^2}{2}\right).$$

Finally, for the last term, we have

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x}M_1\left(\frac{x^2}{2(1-z^2)}\right)\cdot\frac{x}{\sqrt{1-z^2}} &= M_2\left(\frac{x^2}{2(1-z^2)}\right)\cdot\frac{x^2}{(1-z^2)^{3/2}} + M_1\left(\frac{x^2}{2(1-z^2)}\right)\cdot\frac{1}{\sqrt{1-z^2}} \\
&= \frac{1}{\sqrt{1-z^2}}\left(M_2\left(\frac{x^2}{2(1-z^2)}\right)\cdot\frac{x^2}{(1-z^2)} + M_1\left(\frac{x^2}{2(1-z^2)}\right)\right) \\
&= \frac{\exp\left(\frac{x^2}{2(1-z^2)}\right)}{\sqrt{1-z^2}},
\end{aligned}$$

where the first equality uses Fact A.1 and the last equality is by identity (4) in Fact A.2.

Hence, we have

$$h_z''(x) = \frac{2e^{x^2/2(1-z^2)} - (e^{(x+z)^2/2} + e^{(x-z)^2/2})\sqrt{1-z^2}}{\sqrt{1-z^2}}.$$

So to check that $h_z''(x) \geq 0$ for all $x \in \mathbb{R}$, it suffices to check that

$$\frac{(e^{(x+z)^2/2} + e^{(x-z)^2/2})\sqrt{1-z^2}}{2} \leq e^{x^2/2(1-z^2)}.$$

Indeed, we have

$$\begin{aligned}
\frac{(e^{(x+z)^2/2} + e^{(x-z)^2/2})\sqrt{1-z^2}}{2} &\leq \frac{(e^{(x+z)^2/2} + e^{(x-z)^2/2})e^{-z^2/2}}{2} \\
&= e^{x^2/2}\frac{(e^{xz} + e^{-xz})}{2} \\
&\leq e^{x^2/2}e^{x^2z^2/2} \\
&= e^{x^2(1+z^2)/2} \\
&\leq e^{x^2/2(1-z^2)},
\end{aligned}$$

where the first inequality is because $1 - a \leq e^{-a}$ for all $a \in \mathbb{R}$, the second inequality is because $(e^a + e^{-a})/2 = \cosh(a) \leq e^{a^2/2}$ for all $a \in \mathbb{R}$, and the last inequality is because $1 + a \leq 1/(1-a)$ for all $a < 1$. This proves that h_z is convex which concludes the proof that $x = 0$ is a minimizer for h_z and hence, completes the proof of the lemma. □

Proof (of Lemma 3.6). The inequality $R_t(t,g) + \frac{1}{2}R_{gg}(t,g) \geq 0$ is equivalent to

$$R(t, g+1) + R(t, g-1) \geq 2R(t-1, g). \tag{C.3}$$

We first prove the claim for $t = 1$. In this case, the RHS of (C.3) is identically 0. On the other hand, the LHS of (C.3) is non-decreasing in g by Lemma 3.1. Hence, it suffices to prove the inequality for $g = 0$. With $t = 1$ and $g = 0$, we have

$$R(1,1) + R(1,-1) = 2\kappa M_0(1/2).$$

As M_0 is decreasing (Fact A.3) and $1/2 \leq \gamma^2/2$, we have $M_0(1/2) \geq M_0(\gamma^2/2) = 0$. So (C.3) holds for $t = 1$ and $g \geq 0$.

For the remainder of the proof, we assume that $t > 1$. Observe that $\gamma\sqrt{t} - 1 \leq \gamma\sqrt{t-1} \leq \gamma\sqrt{t} + 1$ (since $t \geq 1$).[12] We will consider a few cases depending on the value of g.

Case 1: $g \leq \gamma\sqrt{t} - 1$. In this case, $g + 1 \leq \gamma\sqrt{t}$, $g \leq \gamma\sqrt{t-1}$, and $g - 1 \leq \gamma\sqrt{t}$. Hence,

$$\begin{aligned}
R(t, g+1) &= \frac{g+1}{2} + \kappa\sqrt{t} \cdot M_0\left(\frac{(g+1)^2}{2t}\right) \\
R(t, g-1) &= \frac{g-1}{2} + \kappa\sqrt{t} \cdot M_0\left(\frac{(g-1)^2}{2t}\right) \\
R(t-1, g) &= \frac{g}{2} + \kappa\sqrt{t} \cdot M_0\left(\frac{g^2}{2(t-1)}\right).
\end{aligned}$$

So (C.3) is equivalent to

$$\sqrt{t} \cdot M_0\left(\frac{(g+1)^2}{2t}\right) + \sqrt{t} \cdot M_0\left(\frac{(g-1)^2}{2t}\right) \geq 2\sqrt{t-1} \cdot M_0\left(\frac{g^2}{2(t-1)}\right), \tag{C.4}$$

[12] The inequality $\gamma\sqrt{t} - 1 \leq \gamma\sqrt{t-1}$ is equivalent to $\sqrt{t} - \sqrt{t-1} \leq 1/\gamma$. As $t \mapsto \sqrt{t}$ is concave and $t \geq 1$, the LHS is maximized at $t = 1$ (Fact A.6). Hence, the inequality is true provided $\sqrt{2} \leq 1 + 1/\gamma$. One can check numerically that this last inequality is true as $\gamma \leq 2$.

or rearranging, is equivalent to

$$M_0\left(\frac{(g+1)^2}{2t}\right)+M_0\left(\frac{(g-1)^2}{2t}\right)\geq 2\sqrt{1-1/t}\cdot M_0\left(\frac{g^2}{2(t-1)}\right).$$

The latter inequality is true by Lemma C.4 using $x=g/\sqrt{t}$ and $z=1/\sqrt{t}\in(0,1)$.

Case 2: $\gamma\sqrt{t}-1\leq g\leq\gamma\sqrt{t-1}$. Let $\tilde{R}$ be the function defined in Lemma C.1. In this case, we have

$$R(t,g+1)=\gamma\sqrt{t}=\tilde{R}(t,\gamma\sqrt{t})\geq\tilde{R}(t,g+1)=\frac{g+1}{2}+\kappa\sqrt{t}\cdot M_0\left(\frac{(g+1)^2}{2t}\right).$$

The inequality is by Lemma C.1 which implies that $\tilde{R}(t,g+1)$ is *non-increasing* for $g\in(\gamma\sqrt{t}-1,\infty)$. Using the lower bound on $R(t,g+1)$, (C.3) is again implied by (C.4) and we have already verified that (C.4) is true.

Case 3: $\gamma\sqrt{t-1}\leq g$. Note that for $g\geq\gamma\sqrt{t-1}$, the functions $R(t-1,g)$ and $R(t,g+1)$ are constant in g but $R(t,g-1)$ is non-decreasing in g. Hence, it suffices to check (C.3) for $g=\gamma\sqrt{t-1}$ which holds by case 2. □

D Analysis of Algorithm 1 for general cost vectors

In this section, we prove the upper bound of Theorem 2.1 in full generality.

Theorem D.1. Let $\mathcal{A}$ be the algorithm described in Algorithm 1. For any adversary $\mathcal{B}$ (allowing any cost vectors $\ell_t\in[0,1]^2$), we have

$$\sup_{t\geq 1}\frac{\mathrm{Regret}(2,t,\mathcal{A},\mathcal{B})}{\sqrt{t}}\leq\frac{\gamma}{2}.$$

In Subsection 3.1, since the gap was integer-valued, the identity of the best expert could only change when the gap is exactly 0 (at which time there are two best experts). In general, the gap can be real-valued, so the best expert can switch abruptly, which affects our formula for the regret. We will need to generalize Proposition 2.3 to deal with this possibility. Let $\Delta_{\mathrm{R}}(t)=\mathrm{Regret}(t)-\mathrm{Regret}(t-1)$.

Proposition D.2. Let g_{t-1} be the gap after time $t-1$ but before playing an action at time t. Let g_t be the gap after time t. Let $p(t,g_{t-1})$ denote the probability mass assigned to the worst expert at time t. Suppose that $p(t,0)=1/2$ for all $t\geq 1$.

1. If a best expert at time $t-1$ remains a best expert at time t then

$$\Delta_{\mathrm{R}}(t)=(g_t-g_{t-1})p(t,g_{t-1}).$$

2. If a best expert at time $t-1$ is no longer a best expert at time t then

$$\Delta_{\mathrm{R}}(t)=g_t-(g_t+g_{t-1})p(t,g_{t-1}).$$

   Moreover, $g_t+g_{t-1}\leq 1$.

The proof of this is very similar to that of Proposition 2.3 and appears in Appendix D.1

Remark. Note that, at any specific time, the set of best experts may have size either one or two so the choice of the best expert in Proposition D.2 may be ambiguous. However, note that if $g_{t-1}=0$ (i.e., there are two best experts at time $t-1$) then $p(t,g_{t-1})=1/2$ so both formulas give $\Delta_{\mathrm{R}}(t)=\frac{1}{2}g_t$. On the other hand, if $g_t=0$ (i.e., there are two best experts at time t) then both formulas give $\Delta_{\mathrm{R}}(t)=-g_{t-1}p(t,g_{t-1})$. Hence there is no issue with the ambiguity.

We will need the following identity which is essentially the same as Lemma 3.5 but without specializing to the case where $|g_t - g_{t-1}| = 1$.

Lemma D.3. Let $g_0, g_1, \ldots$ be a sequence of real numbers. Then for any function f and any fixed time $T \geq 1$, we have

$$\begin{aligned} f(T, g_T) - f(0, g_0) &= \sum_{t=1}^{T} f(t, g_t) - \frac{f(t, g_{t-1}+1) + f(t, g_{t-1}-1)}{2} \\ &\quad + \sum_{t=1}^{T} \left(\frac{1}{2} f_{gg}(t, g_{t-1}) + f_t(t, g_{t-1}) \right). \end{aligned} \tag{D.1}$$

Proof. The proof is identical to the proof of Lemma 3.5 except that we do not perform the simplification in (C.2). □

When we assumed the gaps were integer-valued, we had

$$\Delta_{\mathrm{R}}(t) = R(t, g_t) - \frac{R(t, g_{t-1}+1) + R(t, g_{t-1}-1)}{2}$$

because both sides were equal to $R_g(t, g_{t-1}) \cdot (g_t - g_{t-1})$. This does not hold in the general setting, but we will be able to prove the following inequality.

Lemma D.4. For all $t \geq 1$,

$$\Delta_{\mathrm{R}}(t) \leq R(t, g_t) - \frac{R(t, g_{t-1}+1) + R(t, g_{t-1}-1)}{2}.$$

The proof of Lemma D.4 appears in Appendix D.2. Given Lemma D.4, we can now prove our upper bound in general.

Proof (of Theorem D.1). Fix any $T \geq 1$. Then

$$\begin{aligned} R(T, g_T) - R(0, g_0) &= \sum_{t=1}^{T} R(t, g_t) - \frac{R(t, g_{t-1}+1) + R(t, g_{t-1}-1)}{2} \\ &\quad + \sum_{t=1}^{T} \left(\frac{1}{2} R_{gg}(t, g_{t-1}) + R_t(t, g_{t-1}) \right) && \text{(Lemma D.3)} \\ &\geq \sum_{t=1}^{T} \Delta_{\mathrm{R}}(t) && \text{(Lemma D.4 and Lemma 3.6)} \\ &= \mathrm{Regret}(T). \end{aligned}$$

As $g_0 = 0$ and $R(0, 0) = 0$, we have $\mathrm{Regret}(T) \leq R(T, g_T) \leq \gamma\sqrt{T}/2$, where the last inequality is by Lemma 3.2. □

D.1 Proof of Proposition D.2

Proof (of Proposition D.2). Fix t and for notational convenience, let $p = p(t, g_{t-1})$ throughout the proof. In addition, throughout the proof, we use expert 1 to refer to the worst expert at time $t-1$ (chosen arbitrarily if the choice of worst expert is not unique) and use expert 2 to refer to the other expert. Let

$\ell_{t,1}, \ell_{t,2} \in [0,1]$ be the respective losses at time t and $L_{t,1}, L_{t,2}$ be the respective *cumulative* losses up to time t. Note that $g_{t-1} = L_{t-1,1} - L_{t-1,2}$. Finally, we set $L_t^* = \min_{i\in[2]} L_{t,i}$. By assumption, $L_{t-1}^* = L_{t-1,2}$.

For the first assertion we have $L_t^* = L_{t,2}$ (because a best expert remains a best expert). Note that $\ell_{t,1} + \ell_{t,2} = (L_{t,1} - L_{t,2}) - (L_{t-1,1} - L_{t-1,2}) = g_t - g_{t-1}$. So the cost of the algorithm can be can be written as

$$p\ell_{t,1} + (1-p)\ell_{t,2} = p(g_t - g_{t-1}) + \ell_{t,2}.$$

On the other hand, $L_t^* - L_{t-1}^* = L_{t,2} - L_{t-1,2} = \ell_{t,2}$. Subtracting this from the above display equation gives $\Delta_{\mathrm{R}}(t) = (g_t - g_{t-1})p$.

In the second assertion, we have $L_t^* = L_{t,1}$. Again, the algorithm incurs cost $p\ell_{t,1} + (1-p)\ell_{t,2}$. This time, note that $\ell_{t,1} - \ell_{t,2} = (L_{t,1} - L_{t,2}) - (L_{t-1,1} - L_{t-1,2}) = -g_t - g_{t-1}$. So the algorithm incurs cost $-p(g_t + g_{t-1}) + \ell_{t,2}$. On the other hand,

$$L_t^* - L_{t-1}^* = L_{t,1} - L_{t-1,2} = L_{t,1} - L_{t-1,1} + L_{t-1,1} - L_{t-1,2} = \ell_{t,1} + g_{t-1} = \ell_{t,2} - g_{t-1},$$

where the last equality uses the identity $\ell_{t,1} - \ell_{t,2} = -g_t - g_{t-1}$. Subtracting this last quantity with the change in the algorithm's cost gives $\Delta_{\mathrm{R}}(t) = g_{t-1} - p(g_t + g_{t-1})$.

To complete the proof for the second assertion, it remains to check that $g_t + g_{t-1} \le 1$. From above, we have the identity, $g_t + g_{t-1} = \ell_{t,2} - \ell_{t,1} \le \ell_{t,2} \le 1$, as desired. □

D.2 Proof of Lemma D.4

Proof (of Lemma D.4). Fix $t \ge 1$. We will consider the two cases corresponding to the two cases in Proposition D.2.

Case 1: A best expert at time $t-1$ remains a best expert at time t. In this case, $\Delta_{\mathrm{R}}(t) = (g_t - g_{t-1})p(t, g_{t-1})$, so it suffices to check that

$$p(t, g_{t-1}) \cdot (g_t - g_{t-1}) \le R(t, g_t) - \frac{R(t, g_{t-1}+1) + R(t, g_{t-1}-1)}{2}. \tag{D.2}$$

Rearranging, the above inequality is equivalent to

$$R(t, g_t) - \frac{R(t, g_{t-1}+1) + R(t, g_{t-1}-1)}{2} - p(t, g_{t-1}) \cdot (g_t - g_{t-1}) \ge 0.$$

If g_{t-1} is fixed then notice that the LHS of the above expression is concave in g_t. To see this, Lemma 3.1 implies that $R(t, g_t)$ is concave in g_t, the second term is constant in g_t, and the last term is linear in g_t. Hence, it suffices to verify the inequality when $g_t = g_{t-1} \pm 1$ (Fact A.7). Indeed, if $|g_t - g_{t-1}| = 1$ then

$$\begin{aligned} R(t, g_t) - \frac{R(t, g_{t-1}+1) + R(t, g_{t-1}-1)}{2} &= \frac{R(t, g_{t-1}+1) - R(t, g_{t-1}-1)}{2} \cdot (g_t - g_{t-1}) \\ &= p(t, g_{t-1}) \cdot (g_t - g_{t-1}), \end{aligned}$$

where the second equality used the definition of p.

Case 2: A best at time $t-1$ is no longer a best expert at time t. This case is nearly identical to the previous case but in this case $\Delta_{\mathrm{R}}(t) = g_t - (g_t + g_{t-1})p(t, g_{t-1})$ with the promise that $g_t + g_{t-1} \le 1$. Hence, the inequality we need to verify is that

$$g_t - (g_t + g_{t-1})p(t, g_{t-1}) \le R(t, g_t) - \frac{R(t, g_{t-1}+1) + R(t, g_{t-1}-1)}{2}. \tag{D.3}$$

Once again, we do this via a concavity argument. Fix g_{t-1}. Since $g_t + g_{t-1} \le 1$, we have $g_t \in [0, 1 - g_{t-1}]$. Notice that LHS of (D.3) is linear in g_t and the RHS of (D.3) is concave in g_t (by Lemma 3.1). Hence, it

suffices to check the inequality assuming $g_t \in \{0, 1 - g_{t-1}\}$. Note that the case $g_t = 0$ is handled by case 1 since the LHS of (D.2) and (D.3) are identical (see also the remark after Proposition D.2).

Now assume that $g_t = 1 - g_{t-1}$. Then (D.3) becomes

$$1 - g_{t-1} - p(t, g_{t-1}) \leq R(t, 1 - g_{t-1}) - \frac{R(t, g_{t-1} + 1) + R(t, g_{t-1} - 1)}{2}$$

Recall that $p(t, g) = \frac{R(t,g+1) - R(t,g-1)}{2}$ so that the above inequality is equivalent to

$$1 - g_{t-1} - \frac{R(t, g_{t-1} + 1) + R(t, g_{t-1} - 1)}{2} \leq R(t, 1 - g_{t-1}) - \frac{R(t, g_{t-1} + 1) + R(t, g_{t-1} - 1)}{2}.$$

Rearranging the inequality becomes

$$1 \leq g_{t-1} + R(t, 1 - g_{t-1}) - R(t, g_{t-1} - 1).$$

Note that $g_{t-1} \leq 1 \leq \gamma\sqrt{t}$ (when $t \geq 1$). Hence, by definition of R, the RHS of the above inequality is

$$\begin{aligned} g_{t-1} + R(t, 1 - g_{t-1}) - R(t, g_{t-1} - 1) &= g_{t-1} + \frac{1 - g_{t-1}}{2} + \kappa\sqrt{t} M_0 \left(\frac{(1 - g_{t-1})^2}{2} \right) \\ &\quad - \frac{1 - g_{t-1}}{2} - \kappa\sqrt{t} M_0 \left(\frac{(g_{t-1} - 1)^2}{2} \right) \\ &= 1, \end{aligned}$$

and obviously, $1 \leq 1$. □

E Additional proofs for Section 4

Before proving Theorem 4.2, some preliminary definitions are required. For a martingale $(X_t)_{t \in \mathbb{N}}$, define its maximum process $X_t^* = \max_{0 \leq s \leq t} |X_s|$ and its quadratic variation process $[X]_t = \sum_{1 \leq s \leq t} (X_s - X_{s-1})^2$.

Theorem E.1 (Davis [15]). There exists a constant C such that for any martingale $(X_t)_{t \in \mathbb{N}}$ with $X_0 = 0$, $\mathrm{E}\left[X_\infty^* \right] \leq C \, \mathrm{E}\left[[X]_\infty^{1/2} \right]$.

We will prove a more general variant of Theorem 4.2. To recover Theorem 4.2, we apply the following theorem with $\sigma = 0$ and then take expectations to get that $\mathrm{E}\left[Z_\tau \right] = Z_0$.

Theorem E.2. Let $(Z_t)_{t \in \mathbb{Z}_{\geq 0}}$ be a martingale with respect to the filtration $\{\mathcal{F}_t\}$ and $K > 0$ a constant such that $|Z_t - Z_{t-1}| \leq K$ almost surely for all t. Let $\sigma \leq \tau$ be stopping times and suppose that $\mathrm{E}\left[\sqrt{\tau} \right] < \infty$. Then the random variables Z_σ, Z_τ are almost surely well-defined and $\mathrm{E}\left[Z_\tau \mid \mathcal{F}_\sigma \right] = Z_\sigma$.

Proof. Define the stopped process $Z_{t \wedge \tau}$, which is also a martingale [29, Theorem 10.15]. Since $\mathrm{E}\left[\sqrt{\tau} \right] < \infty$ we have $\Pr\left[\tau < \infty \right] = 1$. On the event $\{\tau < \infty\}$, $(Z_{t \wedge \tau})_{t \geq 0}$ has a well-defined limit, which is used as the almost sure definition of Z_τ. As $\{\tau < \infty\} \subseteq \{\sigma < \infty\}$, the same argument shows that $(Z_{t \wedge \sigma})_{\geq 0}$ has a well-defined limit, and we use this as the almost sure definition of Z_σ.

We claim that also $Z_{t \wedge \tau} \xrightarrow{L_1} Z_\tau \in L_1$ and $Z_{t \wedge \sigma} \xrightarrow{L_1} Z_\sigma \in L_1$, from which the theorem concludes as follows. By the definition of conditional expectation, we need to check that $\mathrm{E}\left[Z_\tau \mathbf{1}_A \right] = \mathrm{E}\left[Z_\sigma \mathbf{1}_A \right]$ for all $A \in \mathcal{F}_\sigma$. To that end, fix $A \in \mathcal{F}_\sigma$ and note that $A \cap \{\sigma \leq t\} \in \mathcal{F}_{\sigma \wedge t}$. For any fixed t, $t \wedge \sigma \leq t \wedge \tau \leq t$, so the optional sampling theorem [29, Theorem 10.11] applied to the stopped process yields $\mathrm{E}\left[Z_{t \wedge \tau} \mid \mathcal{F}_{t \wedge \sigma} \right] = Z_{t \wedge \sigma}$. Hence,

$$\mathrm{E}\left[Z_{\tau \wedge t} \mathbf{1}_A \mathbf{1}_{\{\sigma \leq t\}} \right] = \mathrm{E}\left[Z_{\sigma \wedge t} \mathbf{1}_A \mathbf{1}_{\{\sigma \leq t\}} \right]. \tag{E.1}$$

Since $Z_{\tau\wedge t} \xrightarrow{L_1} Z_\tau$, it follows that $Z_{\tau\wedge t}\mathbf{1}_A\mathbf{1}_{\{\sigma\le t\}} \xrightarrow{L_1} Z_\tau\mathbf{1}_A\mathbf{1}_{\{\sigma<\infty\}}$. This is because

$$\begin{aligned}\mathrm{E}\left[|Z_{\tau\wedge t}\mathbf{1}_A\mathbf{1}_{\sigma\le t} - Z_\tau\mathbf{1}_A\mathbf{1}_{\sigma<\infty}|\right] &\le \mathrm{E}\left[\,|Z_{\tau\wedge t}\mathbf{1}_A\mathbf{1}_{\sigma\le t} - Z_\tau\mathbf{1}_A\mathbf{1}_{\sigma\le t}|\,\right] + \mathrm{E}\left[\,|Z_\tau\mathbf{1}_A\mathbf{1}_{\sigma<\infty} - Z_\tau\mathbf{1}_A\mathbf{1}_{\sigma\le t}|\,\right] \\ &\le \mathrm{E}\left[\,|Z_{t\wedge\tau} - Z_\tau|\,\right] + \mathrm{E}\left[\,|Z_\tau|\mathbf{1}_{t<\sigma<\infty}\,\right].\end{aligned}$$

The quantity $\mathrm{E}\left[\,|Z_{t\wedge\tau} - Z_\tau|\,\right] \to 0$ because $Z_{t\wedge\tau} \xrightarrow{L_1} Z_\tau$. Next, $Z_\tau \in L_1$ and $\mathbf{1}_{t<\sigma<\infty} \to 0$ a.s. so $\mathrm{E}\left[\,|Z_\tau|\mathbf{1}_{t<\sigma<\infty}\,\right] \to 0$ by dominated convergence. Finally, note that $Z_\tau\mathbf{1}_A\mathbf{1}_{\sigma<\infty} = Z_\tau\mathbf{1}_A$ as $\mathbf{1}_{\sigma<\infty} = 1$ a.s. Hence,

$$\mathrm{E}\left[\,Z_{\tau\wedge t}\mathbf{1}_A\mathbf{1}_{\{\sigma\le t\}}\,\right] \xrightarrow{t\to\infty} \mathrm{E}\left[\,Z_\tau\mathbf{1}_A\,\right]. \tag{E.2}$$

Similarly,

$$\mathrm{E}\left[\,Z_{\sigma\wedge t}\mathbf{1}_A\mathbf{1}_{\{\sigma\le t\}}\,\right] \xrightarrow{t\to\infty} \mathrm{E}\left[\,Z_\sigma\mathbf{1}_A\,\right]. \tag{E.3}$$

Combining (E.1), (E.2), and (E.3) gives $\mathrm{E}\left[\,Z_\tau\mathbf{1}_A\,\right] = \mathrm{E}\left[\,Z_\sigma\mathbf{1}_A\,\right]$ as desired.

It remains to show that $Z_{\tau\wedge t} \xrightarrow{L_1} Z_\tau \in L_1$ and $Z_{\sigma\wedge t} \xrightarrow{L_1} Z_\sigma \in L_1$. We will only prove the convergence for $Z_{\tau\wedge t}$ as the two arguments are identical. The L_1 convergence is proven using the dominated convergence theorem [29, Corollary 6.26], which requires exhibiting a random variable that bounds $|Z_{t\wedge\tau}|$ for all t and has finite expectation. For notational convenience, let $X_t = Z_{t\wedge\tau}$. Clearly $|X_t| \le X_t^* \le X_\infty^*$, so it remains to show that $\mathrm{E}\left[\,X_\infty^*\,\right] < \infty$. Using Theorem E.1 and that Z has increments bounded by K,

$$\mathrm{E}\left[\,X_\infty^*\,\right] \;\le\; C\,\mathrm{E}\left[\,[X]_\infty^{1/2}\,\right] \;=\; C\,\mathrm{E}\left[\,\Big(\sum_{1\le s\le\tau}(Z_s - Z_{s-1})^2\Big)^{1/2}\,\right] \;\le\; CK\,\mathrm{E}\left[\,\tau^{1/2}\,\right] \;<\; \infty.$$

The dominated convergence theorem states that $Z_{t\wedge\tau} \xrightarrow{L_1} Z_\tau \in L_1$, as required. □

Remark (on Theorem 4.3). Breiman's result is not stated in exactly this form because he focused on the case $a \in \mathbb{Z}_{<0}$, in which case M degenerates to a polynomial. One can show by direct calculation that the function $\theta(a)$ in his equation (2.6) is identical to our function $M(a, 1/2, c^2/2)$ for all $a \in \mathbb{R}$.

An alternative approach is to use a result of Greenwood and Perkins [26, Theorem 5], which shows in a more general context that $\Pr\left[\,\tau(c) > u\,\right] = u^{-\lambda_0(-c,c)}\pi(u)$ where $-\lambda_0(-c,c)$ is the largest non-positive eigenvalue of a certain Sturm-Liouville equation and $\pi(u)$ is a "slowly-varying function". It is shown by Perkins [35, Proposition 1] that c is the smallest positive root of $x \mapsto M(-\lambda_0(-c,c), 1/2, x^2/2)$. A standard result [22, Lemma VIII.8.2] states that any slowly-varying function π satisfies $\pi(u) = O(u^\epsilon)$ for every $\epsilon > 0$. This alternative approach suffices to prove Theorem 4.1 since (4.4) is unaffected by the slowly-varying function.

E.1 Large regret infinitely often

In this subsection, we sketch the following theorem.

Theorem E.3. For any algorithm $\mathcal{A}$ and any $\epsilon > 0$, there exists an adversary $\mathcal{B}_\epsilon$ such that

$$\limsup_{t\ge1} \frac{\mathrm{Regret}(2, t, \mathcal{A}, \mathcal{B}_\epsilon)}{\sqrt{t}} \;\ge\; \frac{\gamma - \epsilon}{2}. \tag{E.4}$$

Sketch. We use the same adversary as in Theorem 4.1 so that

$$\mathrm{Regret}(t) \ge Z_t + \frac{g_t}{2},$$

where Z_t is a martingale with $Z_0 = 0$ and g_t evolves as a reflected random walk. Let $\mathcal{F}_t \coloneqq \sigma(g_0, \ldots, g_t)$ be the natural filtration. Finally, let $c_\epsilon \geq \gamma - \epsilon$ be as in the proof of Theorem 4.1.

Define the stopping times $\tau_0 \coloneqq 0$ and $\tau_i \coloneqq \inf\left\{ t > \tau_{i-1} : g_t \geq c_\epsilon \sqrt{t} \right\}$ for $i \geq 1$. Note that, by the strong Markov property, for each $i \geq 1$, the process $\{g_{\tau_{i-1}+t}\}_{t \geq 0}$ is a reflected random walk started at position $g_{\tau_{i-1}} > 0$. Moreover, observe that τ_i is similar to the stopping time used in Theorem 4.1 in that the asymptotics of the boundary are the same but the starting point is perturbed by a (random) additive constant. It is not hard to show (via [26, Theorem 5]) that $\mathrm{E}\left[\sqrt{\tau_i}\right] < \infty$.[13] Hence, we can apply Theorem E.2 to obtain that $\mathrm{E}\left[Z_{\tau_i} \mid \mathcal{F}_{\tau_{i-1}} \right] = Z_{\tau_{i-1}}$ for all $i \geq 1$.

We will now inductively construct a sequence of events which satisfy the conclusions of the theorem. To that end, define the events

$$A_i = \{\tau_i < \infty, Z_{\tau_i} \geq \ldots \geq Z_{\tau_1} \geq 0\}.$$

For the base case, we have $A_1 = \{\tau_1 < \infty, Z_{\tau_1} \geq 0\}$. In the proof of Theorem 4.1, we have already verified that $\Pr[A_1] > 0$ (this also follows from the previous paragraph). For the inductive step, suppose that $\Pr[A_{i-1}] > 0$. The condition that $\mathrm{E}\left[Z_{\tau_i} \mid \mathcal{F}_{\tau_{i-1}} \right] = Z_{\tau_{i-1}}$ implies that, for any $B \in \mathcal{F}_{\tau_{i-1}}$ with $\Pr[B] > 0$, the event $B \cap \left\{\tau_i < \infty, Z_{\tau_i} \geq Z_{\tau_{i-1}}\right\}$ has positive probability. Taking $B = A_{i-1}$ implies that $\Pr[A_i] > 0$.

To conclude, for any $n \geq 1$, the event A_n has positive probability. Hence, there exists a sequence of times $T_1, \ldots, T_n < \infty$ and loss vectors up to time T_n that guarantee $g_{T_i} \geq c_\epsilon \sqrt{T_i}$ for all $i \in [n]$ and $Z_{T_n} \geq \ldots \geq Z_{T_1} \geq 0$. In particular, for all $i \in [n]$,

$$\mathrm{Regret}(T_i) \geq Z_{T_i} + \frac{g_{T_i}}{2} \geq \frac{c_\epsilon}{2}\sqrt{T_i}.$$

As $n \geq 1$ was arbitrary, the theorem follows. □

F Additional proofs for Section 5

F.1 Proof of Lemma 5.6

Proof (of Lemma 5.6). First, we check that $f \in C^{1,2}$. Let $(t, x) \in \mathbb{R}_{>0} \times \mathbb{R}$. It is easy to check via standard applications of the Dominated Convergence Theorem (DCT) and the Fundamental Theorem of Calculus (FTC) that

(1) $\partial_t f(t, x) = \int_0^x \partial_t h(t, y)\, \mathrm{d}y - \frac{1}{2}\partial_x h(s, 0)$,
(2) $\partial_x f(t, x) = h(t, x)$, and
(3) $\partial_{xx} f(t, x) = \partial_x h(t, x)$.

All of the above partial derivatives are clearly continuous since h is $C^{1,2}$.

Next, we show that if $\overset{*}{\Delta} h(t, x) = 0$ for all $(t, x) \in \mathbb{R}_{>0} \times \mathbb{R}$, then $\overset{*}{\Delta} f(t, x) = 0$ for all $\mathbb{R}_{>0} \times \mathbb{R}$. By

[13] Verifying that $\mathrm{E}\left[\sqrt{\tau_i}\right] < \infty$ is the only non-rigorous portion of the proof.

DCT and FTC,

$$\begin{aligned}
\overset{*}{\Delta} f(t,x) &= \left(\partial_t + \frac{1}{2}\partial_{xx}\right)\left(\int_0^x h(t,y)\,\mathrm{d}y - \int_0^t \frac{1}{2}\partial_x h(s,0)\,\mathrm{d}s\right) \\
&= \int_0^x \partial_t h(t,y)\,\mathrm{d}y + \frac{1}{2}\partial_{xx}\int_0^x h(t,y)\,\mathrm{d}y - \left(\partial_t + \frac{1}{2}\partial_{xx}\right)\int_0^t \frac{1}{2}\partial_x h(s,0)\,\mathrm{d}s \qquad \text{(by DCT)} \\
&= \int_0^x \partial_t h(t,y)\,\mathrm{d}y + \frac{1}{2}\partial_x h(t,x) - \frac{1}{2}\partial_x h(t,0) \qquad \text{(by FTC)} \\
&= \int_0^x \underbrace{\left(\partial_t h(t,y) + \frac{1}{2}\partial_{xx} h(t,y)\right)}_{=0}\,\mathrm{d}y \qquad \text{(by FTC)} \\
&= 0.
\end{aligned}$$

An application of FTC shows that $\partial_x f(t,x) = h(t,x)$ for every (t,x) as $y \mapsto h(t,y)$ is continuous. □

F.2 Proof of Lemma 5.7

Proof (of Lemma 5.7). Let us assume that we can write $u(t,x) = v(x/\sqrt{t})$. Then, we have $\partial_t u(t,x) = -\frac{x}{2t^{3/2}} v'(x/\sqrt{t})$, and $\frac{1}{2}\partial_{xx} u(t,x) = \frac{1}{2t} v''(x/\sqrt{t})$. The backward heat equation enforces that $v''(x/\sqrt{t}) = \frac{x}{\sqrt{t}} v'(x/\sqrt{t})$. By a change of variables ($z = x/\sqrt{t}$), we obtain the following ordinary differential equation

$$v''(z) \;=\; z \cdot v'(z). \tag{F.1}$$

Hence, $v'(z) = C \cdot e^{\frac{z^2}{2}}$ for some constant C. We can then integrate to obtain $v(z) = \int_0^z Ce^{y^2/2}\,\mathrm{d}y + D = \int_0^{z/\sqrt{2}} \sqrt{2}Ce^{u^2}\,\mathrm{d}u + D$, for some constant D. For the last equality, we made the change of variables $u = y/\sqrt{2}$ in the integral. Therefore, by the definition of erfi (and replacing $C\sqrt{2}$ with C, we have $v(z) = C\,\mathrm{erfi}(z/\sqrt{2}) + D$. Hence, for some constants $C, D \in \mathbb{R}$, we have

$$u(t,x) \;=\; C\,\mathrm{erfi}(x/\sqrt{2t}) + D.$$

Plugging in the boundary condition at $x = 0$ and recalling that $\mathrm{erfi}(0) = 0$ we see that $D = 1/2$. Plugging in the boundary condition that $u(t, \alpha\sqrt{t}) = 0$ and using that $D = 1/2$ we see that $C = -\frac{1}{2\,\mathrm{erfi}(\alpha/\sqrt{2})}$. Therefore, we have that the following function

$$q(t,x) \;=\; \frac{1}{2}\left(1 - \frac{\mathrm{erfi}\left(x/\sqrt{2t}\right)}{\mathrm{erfi}\left(\alpha/\sqrt{2}\right)}\right)$$

satisfies the backwards heat equation and the boundary conditions. Moreover, $q \in C^{1,2}$ on $\mathbb{R}_{>0} \times \mathbb{R}$. □

F.3 Proof of Lemma 5.8

Recall that $\tilde{R}_\alpha = \frac{g}{2} + \kappa_\alpha\sqrt{t} \cdot M_0\left(x^2/2t\right)$ where $\kappa_\alpha = \frac{1}{\sqrt{2\pi}\,\mathrm{erfi}(\alpha/\sqrt{2})}$. First we need to compute some derivatives.

Lemma F.1. The following identities hold for every $\alpha > 0$.

1. $\partial_x \tilde{R}_\alpha(t,x) = \tilde{p}_\alpha(t,x) = \frac{1}{2}\left(1 - \frac{\mathrm{erfi}(x/\sqrt{2t})}{\mathrm{erfi}(\alpha/\sqrt{2t})}\right)$.
2. $\partial_{xx} \tilde{R}_\alpha(t,x) = \partial_x \tilde{p}_\alpha(t,x) = -\kappa_\alpha \cdot \frac{\exp(x^2/2t)}{\sqrt{2t}}$.

Proof. The proof is a straightforward calculation. We have

$$\begin{aligned}\partial_x \tilde{R}_\alpha(t,x) &= \frac{1}{2} - \kappa_\alpha \frac{x}{\sqrt{t}} \cdot M_1\left(\frac{x^2}{2t}\right) \\ &= \frac{1}{2} - \frac{1}{\sqrt{2\pi}\,\mathrm{erfi}(\alpha/\sqrt{2})} \cdot \frac{x}{\sqrt{t}} \cdot \frac{\sqrt{\pi}\,\mathrm{erfi}(x/\sqrt{2t})}{2 \cdot x/\sqrt{2t}} \\ &= \frac{1}{2}\left(1 - \frac{\mathrm{erfi}(x/\sqrt{2t})}{\mathrm{erfi}(\alpha/\sqrt{2})}\right),\end{aligned}$$

where the first equality uses Fact A.1 and the second equality uses the identity (2) in Fact A.2. This proves the first identity.

For the second identity, using the definition of $\mathrm{erfi}(\cdot)$, we have

$$\partial_{xx}\tilde{R}_\alpha = \partial_x \tilde{p}_\alpha(t,x) = -\frac{\exp(x^2/2t)}{\sqrt{2\pi}\,\mathrm{erfi}(\alpha/\sqrt{2})\sqrt{s}} = -\kappa_\alpha \cdot \frac{\exp(x^2/2t)}{\sqrt{2t}}. \qquad \square$$

Proof (of Lemma 5.8). By the first identity in Lemma F.1, we have

$$\int_0^x \tilde{p}_\alpha(t,y)\ \mathrm{d}y = \tilde{R}_\alpha(t,x) - \tilde{R}_\alpha(t,0) \tag{F.2}$$

Note that $\tilde{R}_\alpha(t,0) = \kappa_\alpha\sqrt{t}$. Next, the second identity of Lemma F.1 implies that $-\partial_x \tilde{p}_\alpha(s,0) = \frac{\kappa_\alpha}{2\sqrt{s}}$. Hence,

$$-\frac{1}{2}\int_0^t \partial_x \tilde{p}_\alpha(s,0)\ \mathrm{d}s = \kappa_\alpha\sqrt{t} = \tilde{R}_\alpha(t,0). \tag{F.3}$$

Summing (F.2) and (F.3) gives

$$\int_0^x \tilde{p}_\alpha(t,y)\ \mathrm{d}y - \frac{1}{2}\int_0^t \partial_x \tilde{p}_\alpha(s,0)\ \mathrm{d}s = \tilde{R}_\alpha(t,x) - \tilde{R}_\alpha(t,0) + \tilde{R}_\alpha(t,0) = \tilde{R}_\alpha(t,x). \qquad \square$$

F.4 Proof of Lemma 5.10

The main idea of the proof is that we will approximate R_α by a sequence of smooth functions (i.e. functions in $C^{2,2}$).

Fix $\alpha > 0$. Recall that $\tilde{R}_\alpha(t,x) = \frac{x}{2} + \kappa_\alpha\sqrt{t} \cdot M_0\left(\frac{x^2}{2t}\right)$ for $t > 0, x \in \mathbb{R}$, where $\kappa_\alpha = \frac{1}{\sqrt{2\pi}\,\mathrm{erfi}(\alpha/\sqrt{2})}$. (For $t = 0$, it suffices to define $\tilde{R}_\alpha(t,x) = 0$.) We also have the truncated version, R_α, defined as

$$R_\alpha(t,x) = \begin{cases} \tilde{R}_\alpha(t,g) & t > 0 \wedge x \le \alpha\sqrt{t} \\ \tilde{R}_\alpha(t,\alpha\sqrt{t}) & t > 0 \wedge x \ge \alpha\sqrt{t} \\ 0 & t = 0 \end{cases}.$$

Recall also that $p_\alpha = \partial_x R_\alpha$. For convenience, we restate the lemma.

Lemma 5.10. Fix $\alpha > 0$. Then, almost surely, for all $T \ge 0$, $\mathrm{ContRegret}(T, p_\alpha, B) \le R_\alpha(T, |B_T|)$.

For the remainder of this section, we will write $\tilde{f} = \tilde{R}_\alpha$ and $f = R_\alpha$. Let $\phi(x)$ be any non-increasing C^2 function satisfying $\phi(x) = 1$ for $x \le 0$ and $\phi(x) = 0$ for $x \ge 1$. For concreteness, we may take

$$\phi(x) = \begin{cases} 1 & x \le 0 \\ (1-x) + \frac{1}{2\pi}\sin(2\pi x) & x \in [0,1] \\ 0 & x \ge 1 \end{cases}. \tag{F.4}$$

We leave it as an easy calculus exercise to verify that ϕ is indeed a non-increasing C^2 function.

Next, define $\phi_n(x) = \phi(nx)$ and

$$f_n(t,x) = \tilde{f}(t,x) \cdot \phi_n(x - \alpha\sqrt{t}) + f(t, \alpha\sqrt{t}) \cdot \left(1 - \phi_n(x - \alpha\sqrt{t})\right).$$

Note that $f_n \in C^{2,2}$ on $\mathbb{R}_{>0} \times \mathbb{R}$ for all n. The function f_n is a smooth approximation to f and its limit is exactly f $(= R_\alpha)$.

Claim F.2. For every $t > 0, x \in \mathbb{R}$, $\lim_{n\to\infty} f_n(t,x) = f(t,x)$.

Proof. If $x \leq \alpha\sqrt{t}$ then $\phi_n(x - \alpha\sqrt{t}) = 1$ so $f_n(t,x) = \tilde{f}(t,x) = f(t,x)$. In particular, this also holds for the limit. Next, suppose that $a = x - \alpha\sqrt{t} > 0$. If $n > 1/a$ then $\phi_n(x - \alpha\sqrt{t}) = 0$ so $f_n(t,x) = \tilde{f}(t, \alpha\sqrt{t}) = f(t,x)$. □

Recall that our goal is to relate $f(T, |B_T|)$ and $\int_0^T \partial_x f(t, |B_t|)\ \mathrm{d}|B_t|$. However, one cannot apply Itô's formula to f directly as it is not in $C^{1,2}$. Instead, we will apply Itô's formula to the smoothed version of f, namely f_n, and then take limits. The remainder of this section does this limiting argument carefully.

For technical reasons (namely that $\tilde{f}(t,x)$ has a pole when $t \to 0$ and $x \neq 0$), we will not be able to start the stochastic integral at 0. Hence, we will fix $\epsilon > 0$ and, at the end of the proof, we will allow $\epsilon \to 0$.

The following lemma bounds the stochastic integral of f_n with respect to $|B_t|$.

Lemma F.3. Almost surely, for every $T \geq \epsilon$

$$\begin{aligned}\int_\epsilon^T \partial_x f_n(t, |B_t|)\ \mathrm{d}|B_t| \leq\ & f_n(T, |B_T|) - f_n(\epsilon, |B_\epsilon|) \\ & - \int_\epsilon^T \frac{\alpha}{2\sqrt{t}} \cdot \phi_n'(|B_t| - \alpha\sqrt{t}) \cdot \left(f(t, \alpha\sqrt{t}) - \tilde{f}(t, |B_t|)\right)\ \mathrm{d}t \\ & - \frac{1}{2}\int_\epsilon^T \phi_n''(|B_t| - \alpha\sqrt{t}) \cdot \left(f(t, \alpha\sqrt{t}) - \tilde{f}(t, |B_t|)\right)\ \mathrm{d}t.\end{aligned} \tag{F.5}$$

Proof. The proof is by Itô's formula (Theorem 5.3) applied to f_n. We have, for all $T \geq \epsilon$,

$$f_n(T, |B_T|) - f_n(\epsilon, |B_\epsilon|) = \int_\epsilon^T \partial_x f_n(t, |B_t|)\ \mathrm{d}|B_t| + \int_\epsilon^T \partial_t f_n(t, |B_t|) + \frac{1}{2}\partial_{xx} f_n(t, |B_t|)\ \mathrm{d}t. \tag{F.6}$$

Computing derivatives of f_n, we have

$$\begin{aligned}\partial_t f_n(t,x) =\ & (\partial_t \tilde{f}(t,x)) \cdot \phi_n(x - \alpha\sqrt{t}) - \frac{\alpha}{2\sqrt{t}} \tilde{f}(t,g)\phi_n'(x - \alpha\sqrt{t}) \\ & + \partial_t(f(t, \alpha\sqrt{t})) \cdot (1 - \phi_n(x - \alpha\sqrt{t})) + \frac{\alpha}{2\sqrt{t}} f(t, \alpha\sqrt{t}) \cdot \phi_n'(x - \alpha\sqrt{t})\end{aligned} \tag{F.7}$$

$$\partial_x f_n(t,x) = (\partial_x \tilde{f}(t,x)) \cdot \phi_n(x - \alpha\sqrt{t}) + \tilde{f}(t,x)\phi_n'(x - \alpha\sqrt{t}) - f(t, \alpha\sqrt{t})\phi_n'(x - \alpha\sqrt{t}) \tag{F.8}$$

$$\begin{aligned}\partial_{xx} f_n(t,x) =\ & (\partial_{xx} \tilde{f}(t,x)) \cdot \phi_n(x - \alpha\sqrt{t}) + 2(\partial_x \tilde{f}(t,x))\phi_n'(x - \alpha\sqrt{t}) \\ & + \left(\tilde{f}(t,x) - f(t, \alpha\sqrt{t})\right) \phi_n''(x - \alpha\sqrt{t}).\end{aligned} \tag{F.9}$$

Recalling the notation $\overset{*}{\Delta} = \partial_t + \frac{1}{2}\partial_{xx}$, we have

$$\begin{aligned}\overset{*}{\Delta} f_n(t,x) =\ & \left(\overset{*}{\Delta} \tilde{f}(t,x)\right) \cdot \phi_n(x - \alpha\sqrt{t}) + \partial_t(f(t, \alpha\sqrt{t})) \cdot (1 - \phi_n(x - \alpha\sqrt{t})) \\ & + (\partial_x \tilde{f}(t,x))\phi_n'(x - \alpha\sqrt{t}) \\ & + \frac{\alpha}{2\sqrt{t}} \cdot (f(t, \alpha\sqrt{t}) - \tilde{f}(t,x)) \cdot \phi_n'(x - \alpha\sqrt{t}) + \frac{1}{2}\left(\tilde{f}(t,x) - f(t, \alpha\sqrt{t})\right) \phi_n''(x - \alpha\sqrt{t}).\end{aligned} \tag{F.10}$$

By Lemma 5.9, $\overset{*}{\Delta}\tilde{f} = 0$. By Claim F.4 below, $\partial_t(f(t, \alpha\sqrt{t})) > 0$. Next, observe that $(\partial_x \tilde{f}(t,x)) \cdot \phi_n'(x - \alpha\sqrt{t}) \geq 0$. To see this, if $x \leq \alpha\sqrt{t}$ then $\phi_n'(x - \alpha\sqrt{t}) = 0$. On the other hand, if $x > \alpha\sqrt{t}$ then $\phi_n'(x - \alpha\sqrt{t}) \leq 0$ because ϕ_n is non-increasing and $\partial_x \tilde{f}(t,x) \leq 0$ by Lemma 5.9 and (5.11). Hence, we can lower bound (F.10) by

$$\overset{*}{\Delta} f_n(t,x) \geq \frac{\alpha}{2\sqrt{t}} \cdot (f(t, \alpha\sqrt{t}) - \tilde{f}(t,x)) \cdot \phi_n'(x - \alpha\sqrt{t}) + \frac{1}{2}\left(\tilde{f}(t,x) - f(t,\alpha\sqrt{t})\right) \phi_n''(x - \alpha\sqrt{t}). \tag{F.11}$$

Plugging (F.11) into (F.6) gives

$$\begin{aligned} f_n(T, |B_T|) - f_n(\epsilon, |B_\epsilon|) \geq & \int_\epsilon^T \partial_x f_n(t, |B_t|) \, \mathrm{d}|B_t| \\ & + \int_\epsilon^T \frac{\alpha}{2\sqrt{t}} \cdot \phi_n'(|B_t| - \alpha\sqrt{t}) \cdot \left(f(t,\alpha\sqrt{t}) - \tilde{f}(t,|B_t|)\right) \, \mathrm{d}t \\ & + \frac{1}{2}\int_\epsilon^T \phi_n''(|B_t| - \alpha\sqrt{t}) \cdot \left(f(t,\alpha\sqrt{t}) - \tilde{f}(t,|B_t|)\right) \, \mathrm{d}t. \end{aligned} \tag{F.12}$$

Rearranging (F.12) gives the lemma. □

Claim F.4. If $t > 0$ then $\partial_t(\tilde{f}(t, \alpha\sqrt{t})) > 0$.

Proof. Note that

$$\tilde{f}(t, \alpha\sqrt{t}) = \sqrt{t} \cdot \left(\frac{\alpha}{2} + \frac{M_0(\alpha^2/2)}{\sqrt{2\pi}\,\mathrm{erfi}(\alpha/\sqrt{2})}\right) = \sqrt{t} \cdot f(1, \alpha).$$

So it suffices to check that $\tilde{f}(1,\alpha) > 0$. To see this, note that $\tilde{f}(1,0) = \kappa_\alpha > 0$ and $\partial_x \tilde{f}(1,x) \geq 0$ as long as $x \leq \alpha$ (by the first identity of Lemma F.1). Hence, $\tilde{f}(1,\alpha) > 0$. □

At this point, we would like to take limits on both sides of (F.5). This is achieved by the following two lemmas.

Lemma F.5. Almost surely, for every $T \geq \epsilon$,

1. $\lim_{n\to\infty} \int_\epsilon^T \frac{\alpha}{2\sqrt{t}} \cdot \phi_n'(|B_t| - \alpha\sqrt{t}) \cdot \left(f(t,\alpha\sqrt{t}) - \tilde{f}(t,|B_t|)\right) \, \mathrm{d}t = 0$; and
2. $\lim_{n\to\infty} \int_\epsilon^T \phi_n''(|B_t| - \alpha\sqrt{t}) \cdot \left(f(t,\alpha\sqrt{t}) - \tilde{f}(t,|B_t|)\right) \, \mathrm{d}t = 0$.

Lemma F.6. For every $T \geq \epsilon$,

$$\int_\epsilon^T \partial_x f_n(t, |B_t|) \, \mathrm{d}|B_t| \xrightarrow{L^2} \int_\epsilon^T \partial_x f(t, |B_t) \, \mathrm{d}|B_t|$$

as $n \to \infty$.

Within this section, $X_n \xrightarrow{L^2} X$ means that $\mathrm{E}\left[(X_n - X)^2\right] \to 0$ as $n \to \infty$. We relegate the proofs of Lemma F.5 and Lemma F.6 to Appendix F.5. We now take limits on both sides of (F.5) to obtain the following lemma.

Lemma F.7. Almost surely, for every $T \geq \epsilon$,

$$\int_\epsilon^T \partial_x f(t, |B_t|) \, \mathrm{d}|B_t| \leq f(T, |B_T|) - f(\epsilon, |B_\epsilon|). \tag{F.13}$$

Proof. By Lemma F.6, for every $T \geq \epsilon$,

$$\int_\epsilon^T \partial_x f_n(t, |B_t|)\ \mathrm{d}|B_t| \xrightarrow{L^2} \int_\epsilon^T \partial_x f(t, |B_t)\ \mathrm{d}|B_t|.$$

Hence, there exists a subsequence n_k such that

$$\int_\epsilon^T \partial_x f_{n_k}(t, |B_t|)\ \mathrm{d}|B_t| \xrightarrow{\text{a.s.}} \int_\epsilon^T \partial_x f(t, |B_t)\ \mathrm{d}|B_t|.$$

Using Lemma F.3 to bound the left-hand-side and then Lemma F.5 to take limits gives that (F.13) holds for any fixed $T \geq \epsilon$. Hence, almost surely, (F.13) holds for all rational $T \geq \epsilon$. As both sides of (F.13) are continuous as a function of T, (F.13) holds for all $T \geq \epsilon$. □

The proof of this lemma appears in Appendix F.5.

Proof (of Lemma 5.10). We will work in the probability 1 set where Lemma F.7 holds and $t \mapsto B_t$ is continuous.

Fix $T > 0$. Note that $\mathrm{ContRegret}(T, \partial_x f, B)$ is defined because $\partial_x f \in [0, 1/2]$ and $\partial_x f(t, 0) = 1/2$ for all $t > 0$ (see (5.12)). Recalling Definition 5.1, we have, for $\epsilon \leq T$,

$$\begin{aligned}
\mathrm{ContRegret}(T, \partial_x f, B) &= \int_0^T \partial_x f(t, |B_t|)\ \mathrm{d}|B_t| \\
&= \int_\epsilon^T \partial_x f(t, |B_t|)\ \mathrm{d}|B_t| + \int_0^\epsilon \partial_x f(t, |B_t|)\ \mathrm{d}|B_t| \\
&\leq f(T, |B_T|) - f(\epsilon, |B_\epsilon|) + \int_0^\epsilon \partial_x f(t, |B_t|)\ \mathrm{d}|B_t| \qquad \text{(Lemma F.7)}.
\end{aligned}$$

The right-hand-side is continuous in ϵ so taking $\epsilon \to 0$ (and recalling that $f(0, 0) = 0$), gives

$$\mathrm{ContRegret}(T, \partial_x f, B) \leq f(T, |B_T|).$$ □

F.5 Additional proofs from Appendix F.4

Before we prove Lemma F.5, we will need one key observation.

Lemma F.8. Fix $\epsilon > 0$. Then there is a constant $C_\epsilon > 0$ (depending also on α) such that for $t > 0$ and x satisfying $|x - \alpha\sqrt{t}| \leq 1$,

1. $|\tilde{f}(t, x) - f(t, \alpha\sqrt{t})| \leq C_\epsilon \cdot (x - \alpha\sqrt{t})^2$; and
2. $|\partial_x \tilde{f}(t, x)| \leq C_\epsilon \cdot |x - \alpha\sqrt{t}|$.

Proof. The key observation is that $f(t, \alpha\sqrt{t})$ is already a first-order Taylor expansion of $\tilde{f}(t, x)$ (in x) about the point $\gamma\sqrt{t}$. Indeed, $\tilde{f}(t, \alpha\sqrt{t}) = f(t, \alpha\sqrt{t})$ and $(\partial_x \tilde{f})(t, \alpha, \sqrt{t}) = 0$. Hence, by Taylor's Theorem (see e.g. [39, Theorem 5.15])

$$|\tilde{f}(t, x) - f(t, \alpha\sqrt{t})| \leq \frac{1}{2} \cdot (x - \alpha\sqrt{t})^2 \cdot \sup_{t \geq \epsilon, |x - \alpha\sqrt{t}| \leq 1} |\partial_{xx} \tilde{f}(t, x)|$$

By the second identity in Lemma F.1, we have

$$|\partial_{xx} \tilde{f}(t, x)| = \frac{\kappa_\alpha \exp(x^2/2t)}{\sqrt{2t}}.$$

Since $t \geq \epsilon$ and $x \leq 1 + \alpha\sqrt{t}$, we have

$$\begin{aligned}
|\partial_{gg}\tilde{f}(t,x)| &\leq \frac{\kappa_\alpha \exp\big((1+\alpha\sqrt{t})^2/2t\big)}{\sqrt{2\epsilon}} \\
&= \frac{\kappa_\alpha \exp(\alpha^2/2 + \alpha/\sqrt{t} + 1/t)}{\sqrt{2\epsilon}} \\
&\leq \frac{\kappa_\alpha \exp(\alpha^2/2 + \alpha/\sqrt{\epsilon} + 1/\epsilon)}{\sqrt{2\epsilon}}.
\end{aligned}$$

So one can take $C_\epsilon = \frac{\kappa_\alpha \exp(\alpha^2/2+\alpha/\sqrt{\epsilon}+1/\epsilon)}{\sqrt{2\epsilon}}$. This gives the first assertion.

The second assertion is similar. Indeed, since $(\partial_x \tilde{f})(t, \alpha\sqrt{t}) = 0$, we have

$$\begin{aligned}
|(\partial_x \tilde{f})(t,x)| &= |(\partial_x \tilde{f})(t,x) - (\partial_x \tilde{f})(t,\alpha\sqrt{t})| \\
&\leq |x - \alpha\sqrt{t}| \cdot \sup_{t\geq\epsilon, |x-\alpha\sqrt{t}|\leq 1} |\partial_{xx}\tilde{f}(t,x)| \\
&\leq C_\epsilon \cdot |x - \alpha\sqrt{t}|.
\end{aligned}$$

□

We also need a simple claim which bounds the value of $|\phi_n'(x)|$ and $|\phi_n''(x)|$.

Claim F.9. There is an absolute constant $C > 0$ such that $|\phi_n'(x)| \leq Cn$ and $|\phi_n''(x)| \leq Cn^2$.

Proof. Note that $\phi_n'(x) = n \cdot \phi'(x)$ and $n^2 \cdot \phi''(x)$. It is easy to see, from differentiating (F.4) or by continuity and compact arguments, that there exists $C > 0$ such that $|\phi'(x)|, |\phi''(x)| \leq C$ for all $x \in \mathbb{R}$. □

Proof (of Lemma F.5). We start with the second assertion. The first assertion is similar but simpler. We claim that there exists a constant C' (depending on ϵ and α) such that

$$\left|\phi_n''(|B_t| - \alpha\sqrt{t}) \cdot \Big(f(t,\alpha\sqrt{t}) - \tilde{f}(t,|B_t|)\Big)\right| \leq C' \mathbf{1}[|B_t| - \alpha\sqrt{t} \in [0, 1/n]] \tag{F.14}$$

Indeed, if $|B_t| - \alpha\sqrt{t} \notin [0, 1/n]$ then $\phi_n''(|B_t| - \alpha\sqrt{t}) = 0$ so both sides of (F.14) are equal to 0. On the other hand, if $|B_t| - \alpha\sqrt{t} \in [0, 1/n]$ then Lemma F.8 shows that $|f(t,\alpha\sqrt{t}) - \tilde{f}(t,|B_t|)| \leq C_\epsilon/n^2$ where C_ϵ is the constant from Lemma F.8. Next, Claim F.9 gives $|\phi_n''(|B_t| - \alpha\sqrt{t})| \leq Cn^2$. So taking $C' = C_\epsilon \cdot C$ gives (F.14). Hence,

$$\begin{aligned}
\left|\int_\epsilon^T \phi_n''(|B_t| - \alpha\sqrt{t}) \cdot \Big(f(t,\alpha\sqrt{t}) - \tilde{f}(t,|B_t|)\Big)\ \mathrm{d}t\right| &\leq \int_\epsilon^T C' \cdot \mathbf{1}[|B_t| - \alpha\sqrt{t} \in [0,1/n]]\ \mathrm{d}t \\
&= C' \cdot m\left(\Big\{\ t \in [\epsilon, T]\ :\ |B_t| - \alpha\sqrt{t} \in [0,1/n]\ \Big\}\right),
\end{aligned}$$

where m denotes the Lebesgue measure. By continuity of measure, we have

$$\lim_n m\left(\Big\{\ t \in [\epsilon, T]\ :\ |B_t| - \alpha\sqrt{t} \in [0,1/n]\ \Big\}\right) = \int_\epsilon^T \mathbf{1}\Big[|B_t| = \alpha\sqrt{t}\Big]\ \mathrm{d}t = 0 \quad \text{a.s.}$$

This proves the second assertion.

For the first assertion, we can use the bound (from Lemma F.8 and Claim F.9)

$$\left|\phi_n'(x - \alpha\sqrt{t}) \cdot \Big(f(t,\alpha\sqrt{t}) - \tilde{f}(t,x)\Big)\right| \leq \frac{C'}{n}\mathbf{1}[x - \alpha\sqrt{t} \in [0, 1/n]] \leq \frac{C'}{n}. \tag{F.15}$$

Hence,

$$\begin{aligned}
\left|\int_\epsilon^T \frac{\alpha}{2\sqrt{t}} \cdot \phi_n'(|B_t| - \alpha\sqrt{t}) \cdot \Big(f(t,\alpha\sqrt{t}) - \tilde{f}(t,|B_t|)\Big)\ \mathrm{d}t\right| &\leq \int_\epsilon^T \frac{\alpha}{2\sqrt{t}}\frac{C'}{n}\ \mathrm{d}t \\
&\leq C'\alpha\sqrt{T}/n \to 0.
\end{aligned}$$

□

Proof (of Lemma F.6). By (F.8), we have

$$\begin{aligned}\partial_x f_n(t,x) - \partial_x f(t,x) &= \Big(\partial_x \tilde{f}(t,x)\phi_n(x-\alpha\sqrt{t}) - \partial_x f(t,x)\Big) \\ &\quad + \Big(\phi_n'(x-\alpha\sqrt{t}) \cdot \Big(\tilde{f}(t,x) - f(t,\alpha\sqrt{t})\Big)\Big).\end{aligned} \tag{F.16}$$

For the first bracketed term, since $\partial_x \tilde{f}(t,x) = \partial_x f(t,x)$ when $x \le \alpha\sqrt{t}$ and $\partial_x f(t,x) = 0$ when $x \ge \alpha\sqrt{t}$, we have

$$\begin{aligned}\Big|\partial_x \tilde{f}(t,x)\phi_n(x-\alpha\sqrt{t})\Big| &= \Big|\partial_x \tilde{f}(t,x)\phi_n(x-\alpha\sqrt{t})\Big| \, \mathbf{1}[x-\alpha\sqrt{t} \in [0,1/n]] \\ &\le \frac{C'}{n},\end{aligned}$$

where the final inequality is by the second assertion in Lemma F.8. The second bracketed term has been bounded in (F.15), and so we have proved

$$\Big|\partial_x f_n(t,x) - \partial_x f(t,x)\Big| \le \frac{C''}{n} \text{ for all } t \ge \varepsilon \text{ and all } x. \tag{F.17}$$

Tanaka's formula (see [38, Theorem IV.43.3]) states that

$$|B_t| = \int_0^t \operatorname{sign}(B_s)\, \mathrm{d}B_s + L_t =: W_t + L_t,$$

where L is the local time at zero of B and W is a Brownian motion. Recall that $t \mapsto L_t$ is a continuous non-decreasing random process which increases only on the set $\{\, t \,:\, B_t = 0 \,\}$. Therefore by the Itô isometry property, for any $T \ge \varepsilon$,

$$\begin{aligned}&\mathrm{E}\Big[\Big(\int_\varepsilon^T \partial_x f_n(t,|B_t|)\, \mathrm{d}|B|_t - \int_\varepsilon^T \partial_x f(t,|B_t|)\, \mathrm{d}|B|_t\Big)^2\Big] \\ &\quad \le 2\,\mathrm{E}\Big[\Big(\int_\varepsilon^T (\partial_x f_n - \partial_x f)(t,|B_t|)\, \mathrm{d}W_t\Big)^2\Big] + 2\,\mathrm{E}\Big[\Big(\int_\varepsilon^T (\partial_x f_n - \partial_x f)(t,|B_t|))\, \mathrm{d}L_t\Big)^2\Big] \\ &\quad = 2\,\mathrm{E}\Big[\int_\varepsilon^T (\partial_x f_n - \partial_x f)(t,|B_t|)^2\, \mathrm{d}t\Big] + 2\,\mathrm{E}\Big[\Big(\int_\varepsilon^T (\partial_x f_n - \partial_x f)(t,0)\, \mathrm{d}L_t\Big)^2\Big].\end{aligned}$$

Now use (F.17) to bound the right-hand side by

$$2(C''/n)^2 T + 2(C''/n)^2\, \mathrm{E}\,[\, L_T^2 \,] \le C''' n^{-2} T,$$

where the last inequality uses Tanaka's formula (and the fact that W_t is also a standard Brownian motion) to bound

$$\mathrm{E}\,[\, L_T^2 \,] = \mathrm{E}\,[\, (|B_T| - W_T)^2 \,] \le 2\,\mathrm{E}\,[\, |B_T|^2 \,] + 2\,\mathrm{E}\,[\, |W_T|^2 \,] = 4\,\mathrm{E}\,[\, |B_T|^2 \,] = O(T).$$

The result follows. □

F.6 Proof of Claim 5.12

Proof of Claim 5.12. Recall that $h(\alpha) = \frac{\alpha}{2} + \frac{M_0(\alpha^2/2)}{\sqrt{2\pi}\,\operatorname{erfi}(\alpha/\sqrt{2})}$. Hence,

$$\begin{aligned}h'(\alpha) &= \frac{1}{2} - \frac{\alpha \cdot M_1(\alpha^2/2)}{\sqrt{2\pi}\,\operatorname{erfi}(\alpha/\sqrt{2})} - \frac{\exp(\alpha^2/2)\cdot M_0(\alpha^2/2)}{\pi\,\operatorname{erfi}(\alpha/\sqrt{2})} && \text{(by Fact A.1)} \\ &= -\frac{\exp(\alpha^2/2)\cdot M_0(\alpha^2/2)}{\pi\,\operatorname{erfi}(\alpha/\sqrt{2})} && \text{(by Fact A.2(2)).}\end{aligned}$$

This proves the first assertion.

Next, observe that $\frac{\exp(\alpha^2/2)}{\mathrm{erfi}(\alpha/\sqrt{2})}$ is positive for all $\alpha \in \mathbb{R}$. Hence, by Fact A.4, we have that $h'(\alpha) < 0$ for $\alpha \in (0, \gamma)$, $h'(\gamma) = 0$, and $h'(\alpha) > 0$ for $\alpha \in (\gamma, \infty)$. □

F.7 Discussion on the statement of Theorem 5.3

In this paper, we use the version of Itô's formula that appears in Remark 1 after Theorem IV.3.3 in [36]. It states that if $f \in C^{1,2}$, X is a continuous semimartingale[14] and A is a process with bounded variation then

$$\begin{aligned} f(A_T, X_T) - f(A_0, X_0) = \int_0^T \partial_x f(A_t, X_t)\ \mathrm{d}X_t + \int_0^T \partial_t f(A_t, X_t)\ \mathrm{d}A_t \\ + \frac{1}{2}\int_0^T \partial_{xx} f(A_t, X_t)\ \mathrm{d}\langle\, X,\, X\,\rangle_t. \end{aligned} \tag{F.18}$$

In our setting, we take $X_t = |B_t|$ and $A_t = t$. We now explain the notation $\langle\, X,\, X\,\rangle$.

(1) For a continuous local martingale M, $\langle\, M,\, M\,\rangle$ is the unique increasing continuous process vanishing at 0 such that $M^2 - \langle\, M,\, M\,\rangle$ is a martingale [36, Theorem IV.1.8].

(2) If X is a continuous semimartingale with M being the (continuous) local martingale part then $\langle\, X,\, X\,\rangle = \langle\, M,\, M\,\rangle$ [36, Definition IV.1.20].

Tanaka's formula [38, Theorem IV.43.3] asserts that $|B_t| = W_t + L_t$ where W_t is a Brownian Motion and L_t is the local time of B_t at 0, which is an increasing, continuous, adapated process. Hence, $|B_t|$ is a semimartingale with $\langle\, |B|,\, |B|\,\rangle_t = \langle\, W,\, W\,\rangle_t = t$. Plugging these into (F.18) gives

$$f(T, |B_T|) - f(0, |B_0|) \;=\; \int_0^T \partial_x f(t, |B_t|)\, \mathrm{d}\,|B_t| + \int_0^T \Big[\partial_t f(t, |B_t|) + \tfrac{1}{2}\partial_{xx} f(t, |B_t|)\Big]\, \mathrm{d}t,$$

which is what appears in Theorem 5.3.

References

[1] Sepehr Abbasi-Zadeh, Nikhil Bansal, Guru Guruganesh, Aleksandar Nikolov, Roy Schwartz, and Mohit Singh. Sticky Brownian rounding and its applications to constraint satisfaction problems. *arXiv preprint arXiv:1812.07769*, 2018.

[2] Milton Abramowitz and Irene A. Stegun. *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*, volume 55. Courier Corporation, 1965.

[3] Sanjeev Arora, Elad Hazan, and Satyen Kale. The multiplicative weights update method: a meta-algorithm and applications. *Theory of Computing*, 8(1):121–164, 2012.

[4] Leo Breiman. First exit times for a square root boundary. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 2: Contributions to Probability Theory, Part 2*, pages 9–16. University of California Press, 1967.

[5] Leo Breiman. *Probability*. SIAM, 1992.

[6] Sébastien Bubeck. Introduction to online optimization, December 2011. unpublished.

[14] A continuous semimartingale X is a process that can be written as $X = M + N$ where M is a continuous local martingale and N is a continuous adapted process of finite variation.

[7] Sébastien Bubeck, Michael B. Cohen, Yin Tat Lee, James R. Lee, and Aleksander Mądry. k-server via multiscale entropic regularization. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, pages 3–16. ACM, 2018.

[8] Sébastien Bubeck, Ronen Eldan, and Joseph Lehec. Sampling from a log-concave distribution with projected langevin monte carlo. *Discrete Comput. Geom.*, 59(4):757–783, June 2018.

[9] Gruia Calinescu, Chandra Chekuri, Martin Pál, and Jan Vondrák. Maximizing a monotone submodular function subject to a matroid constraint. *SIAM Journal on Computing*, 40(6):1740–1766, 2011.

[10] Nicolo Cesa-Bianchi. Analysis of two gradient-based algorithms for on-line regression. *Journal of Computer and System Sciences*, 59(3):392–411, 1999.

[11] Nicolo Cesa-Bianchi, Yoav Freund, David Haussler, David P Helmbold, Robert E Schapire, and Manfred K Warmuth. How to use expert advice. *Journal of the ACM (JACM)*, 44(3):427–485, 1997.

[12] Nicolò Cesa-Bianchi and Gábor Lugosi. *Prediction, learning, and games*. Cambridge University Press, 2006.

[13] Chandra Chekuri, TS Jayram, and Jan Vondrák. On multiplicative weight updates for concave and submodular function maximization. In *Proceedings of the Conference on Innovations in Theoretical Computer Science (ITCS)*, pages 201–210, 2015.

[14] Thomas M. Cover. Behavior of sequential predictors of binary sequences. In *Proceedings of the 4th Prague Conference on Information Theory, Statistical Decision Functions, Random Processes*. Publishing House of the Czechoslovak Academy of Sciences, Prague, 1965.

[15] Burgess Davis. On the intergrability[15] of the martingale square function. *Israel Journal of Mathematics*, 8:187–190, 1970.

[16] Burgess Davis. On the L_p norms of stochastic integrals and other martingales. *Duke Math. J*, 43(4):697–704, 1976.

[17] Jelena Diakonikolas and Lorenzo Orecchia. The approximate duality gap technique: A unified theory of first-order methods. *SIAM Journal on Optimization*, 29(1):660–689, 2019.

[18] J. L. Doob. *Classical Potential Theory and Its Probabilistic Counterparts*. Springer-Verlag, 1984.

[19] Rick Durrett. *Probability: Theory and Examples*. Cambridge University Press, fifth edition, 2019.

[20] Ronen Eldan and Assaf Naor. Krivine diffusions attain the Goemans–Williamson approximation ratio. *arXiv preprint arXiv:1906.10615*, 2019.

[21] Meier Feder, Neri Merhav, and Michael Gutman. Universal prediction of individual sequences. *IEEE Transactions on Information Theory*, 38(4):1258–1270, 1992.

[22] William Feller. *An Introduction to Probability Theory and Its Applications*. John Wiley & Sons, second edition, 1971.

[23] Dean P. Foster and Rakesh V. Vohra. A randomization rule for selecting forecasts. *Operations Research*, 41(4):704–709, 1993.

[15] This appears to be a typographical error in the title of the paper.

[24] Nick Gravin, Yuval Peres, and Balasubramanian Sivan. Towards optimal algorithms for prediction with expert advice. In *Proceedings of the twenty-seventh annual ACM-SIAM symposium on Discrete algorithms*, pages 528–547. SIAM, 2016.

[25] Nick Gravin, Yuval Peres, and Balasubramanian Sivan. Tight Lower Bounds for Multiplicative Weights Algorithmic Families. In Ioannis Chatzigiannakis, Piotr Indyk, Fabian Kuhn, and Anca Muscholl, editors, *44th International Colloquium on Automata, Languages, and Programming (ICALP 2017)*, volume 80 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 48:1–48:14, Dagstuhl, Germany, 2017. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

[26] Priscilla Greenwood and Edwin Perkins. A conditioned limit theorem for random walk and brownian local time on square root boundaries. *Annals of Probability*, 11:227–261, 1983.

[27] Geoffrey Grimmett and David Stirzaker. *Probability and Random Processes.* Oxford University Press, third edition, 2001.

[28] James Hannan. Approximation to bayes risk in repeated play. *Contributions to the Theory of Games*, 3:97–139, 1957.

[29] Achim Klenke. *Probability Theory: A Comprehensive Course.* Springer, 2008.

[30] Yin Tat Lee and Santosh S. Vempala. Geodesic walks in polytopes. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, pages 927–940. ACM, 2017.

[31] Nick Littlestone and Manfred K. Warmuth. The weighted majority algorithm. *Information and computation*, 108(2):212–261, 1994.

[32] Haipeng Luo and Robert E. Schapire. Towards minimax online learning with unknown time horizon. In *Proceedings of ICML*, 2014.

[33] Peter Mörters and Yuval Peres. *Brownian Motion.* Cambridge University Press, 2010.

[34] Yurii Nesterov. Primal-dual subgradient methods for convex problems. *Mathematical Programming*, 120(1):221–259, 2009.

[35] Edwin Perkins. On the Hausdorff dimension of the brownian slow points. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 64:369–399, 1983.

[36] Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion*, volume 293. Springer Science & Business Media, 2013.

[37] L. C. G. Rogers and David Williams. *Diffusions, Markov Processes and Martingales. Volume 1: Foundations.* Cambridge University Press, second edition, 2000.

[38] L. C. G. Rogers and David Williams. *Diffusions, Markov Processes and Martingales. Volume 2: Itô Calculus*, volume 2. Cambridge University Press, second edition, 2000.

[39] Walter Rudin. *Principles of Mathematical Analysis.* John Wiley & Sons, third edition, 1976.

[40] Volodimir G. Vovk. Aggregating strategies. *Proc. of Computational Learning Theory, 1990*, 1990.

[41] Andre Wibisono, Ashia C. Wilson, and Michael I. Jordan. A variational perspective on accelerated methods in optimization. *Proceedings of the National Academy of Sciences*, 113(47):E7351–E7358, 2016.

[42] David Williams. *Probability with Martingales.* Cambridge University Press, 1991.