Continuous Prediction with Experts' Advice

Victor Sanches Portella VICTORSP@CS.UBC.CA
Christopher Liaw CVLIAW@GOOGLE.COM
Nicholas J. A. Harvey NICKHAR@CS.UBC.CA

Abstract

Prediction with experts' advice is one of the most fundamental problems in online learning and captures many of its technical challenges. A recent line of work has looked at online learning through the lens of differential equations and continuous-time analysis. This viewpoint has yielded optimal results for several problems in online learning.

In this paper, we employ continuous-time stochastic calculus in order to study the discrete-time experts' problem. We use these tools to design a continuous-time, parameter-free algorithm with improved guarantees for the *quantile regret*. We then develop an analogous discrete-time algorithm with a very similar analysis and identical quantile regret bounds. Finally, we design an *anytime* continuous-time algorithm with regret matching the optimal *fixed-time* rate when the gains are independent Brownian Motions; in many settings, this is the most difficult case. This gives some evidence that, even with adversarial gains, the optimal anytime and fixed-time regrets may coincide.

Keywords: experts, online learning, stochastic calculus, anytime, quantile regret

1. Introduction

One of the cornerstone online learning (OL) tasks is *prediction with experts' advice* or *experts' problem*. In this problem, at each round $t = 1, 2, ...$ a player picks a probability distribution p_t over n experts. Next, an adversary picks gains[1] $g_t \in [-1, 1]^n$ for each of the experts. At the end of round t the player receives the expected gains $p_t^{\mathsf{T}} g_t$ of the experts according to p_t. The performance of the player is given by the *regret*: the difference between the best expert's gains in hindsight and the players' gains. Albeit classical, the experts' problem already captures many of the key theoretical challenges in OL. Determining the minimax optimal regret has been a foundational research vein in OL. We focus on the analysis of optimal rates in two settings: *anytime regret* and *quantile regret*.

An intriguing question is to determine the optimal regret achievable by an *anytime* algorithm, that is, an algorithm that does not have access to the total number T of rounds. When the algorithm knows T beforehand, which we refer to as the *fixed-time setting*, the classical Multiplicative Weights Update (MWU) method (Vovk, 1990; Littlestone and Warmuth, 1994) suffers no more than $\sqrt{2T \ln n}$ regret, which is optimal in the worst-case (Cesa-Bianchi et al., 1997). For the time being, the best anytime regret guarantee known is $2\sqrt{T \ln n}$ using MWU with a time-varying step-size (Bubeck, 2011, Theorem 2.4). It is unknown, if for general n, there is an algorithm that guarantees regret smaller than $2\sqrt{T \ln n}$ or whether one can prove a lower bound strictly better than $\sqrt{2T \ln n}$.

The classical notion of regret may not always be ideal. For example, one might not mind if the player performs badly when compared to the single best expert if it performs well when compared to some ε-quantile of the top experts, denoted by *ε-quantile regret*. The first algorithms aimed provably good quantile regret were proposed by Chaudhuri et al. (2009). Currently, the best-known

1. In this paper we use gains in $[-1, 1]$ instead of costs in $[0, 1]$ due to parallels to random walks and Brownian motion.

ε-quantile regret guarantees are[2] $2\sqrt{3T\ln(1/\varepsilon)}$ in the fixed-time setting (Orabona and Pal, 2016) and $4\sqrt{t\ln(1/\varepsilon)}$ in the anytime setting (Chernov and Vovk, 2010). Recently, Negrea et al. (2021) showed a lower-bound of $\sqrt{2T\ln(1/\varepsilon)}$ on the ε-regret. Thus, the gap between the upper and lower bounds is $\sqrt{6}$ in the fixed-time setting and $2\sqrt{2}$ in the anytime setting.

Our contributions. We present a continuous-time variant of the experts problem and use it to study minimax optimal (quantile) regret rates. Our setting is a simpler variant of the framework proposed by Freund (2009), but we use it as a guide in algorithm design. Namely, working in continuous time allows us to utilize powerful analytical tools from stochastic calculus, which often allow for simpler analyses. In this paper, we use continuous-time techniques to obtain improved bounds on the minimax optimal ε-quantile regret and obtain intriguing results for the anytime continuous regret. Furthermore, we hope this to be a showcase of the potential of the impact of this continuous-time framework in research on the experts' problem. Our specific results are as follows.

- **Continuous MWU.** To demonstrate the parallels between the discrete and continuous time problems, we describe a continuous-time version of MWU. In Theorem 3.3, we show that we can easily obtain bounds on the continuous regret that match the best-known regret bounds for MWU: $\sqrt{2T\ln n}$ in the fixed time setting and $2\sqrt{T\ln n}$ in the anytime setting.
- **Continuous quantile regret.** Taking inspiration from Itô's formula from stochastic calculus, we propose a new algorithm for quantile regret in continuous time. This algorithm has anytime continuous quantile regret bounds whose leading constants are better than any known results for the discrete-case (Theorem 4.5). The bounds hold for all $\varepsilon \in (0, 1)$ and $T > 0$ simultaneously. The algorithm can be interpreted as "parameter-free", since it does not involve a learning rate.
- **Discretized quantile regret bound.** Next, we discretize the algorithm from the previous section while preserving the anytime quantile regret guarantees (Theorem 4.6). This algorithm is also parameter-free, and improves upon the best-known[3] quantile regret bounds in the literature. Furthermore, our analysis closely matches the continuous-time analysis.
- **Improved anytime continuous regret with independent experts.** We design an anytime continuous-time algorithm with $\sqrt{2T\ln n}$ regret (a.s. for all T), asymptotically in n, when the gains are independent Brownian Motion (Theorem 5.2). A simple argument shows that this is optimal (Proposition 5.3): for any algorithm and *fixed* time $T > 0$, the expected regret at time T exceeds $\sqrt{2T\ln n}(1 - o(1))$. Thus, against independent experts, the anytime setting is no harder than the fixed-time setting. This gives some evidence that $\sqrt{2T\ln n}$ anytime regret against all adversaries might be possible, matching the optimal fixed-time regret.

1.1 Related Work

Optimal regret in fixed and anytime settings. The most well-known algorithm for the experts setting is the multiplicative weights update (MWU) algorithm (Littlestone and Warmuth, 1994; Vovk, 1990). In the fixed-time setting (with gains in $[-1, 1]$), MWU achieves a regret bound of $\sqrt{2t\ln n}$ and this bound is tight (Cesa-Bianchi et al., 1997, Corollary 3.2.2). In the anytime setting,

2. Ignoring low order terms relative to $\sqrt{T\ln(1/\varepsilon)}$ and multiplying by 2 due to gains being in $[-1, 1]$ instead of $[0, 1]$.
3. In independent work, Zhang et al. (2022) developed an algorithm using coin-betting and a similar potential function to the one we use and which yields a similar quantile regret bound. We further discuss how to obtain the bound from their results in Section 4.1. In contrast, our analysis is quite self-contained, and avoids using the coin-betting framework.

MWU with a dynamic step size is known to achieve a regret bound of $2\sqrt{t \ln n}$ for all times $t \geq 0$ (e.g. Cesa-Bianchi and Lugosi, 2006, §14, Nesterov, 2009, Theorem 4, and Bubeck, 2011, §2.5). It is not known whether the constant 2 is tight; the best known lower bound for the anytime setting is the $\sqrt{2t \ln n}$ which is inherited from the fixed-time setting.

In the fixed-time setting, the minimax regret is known for $n = 2, 3, 4$ experts (Cover, 1967; Abbasi-Yadkori et al., 2017; Bayraktar et al., 2020a). In the anytime setting, we know an optimal algorithm only for $n = 2$ experts, where Harvey et al. (2020b) showed that the optimal regret is $\gamma\sqrt{t}$ where $\gamma \approx 1.3069$.

Another model for regret introduced by Gravin et al. (2016) is the geometric stopping time model in which the number of rounds is a geometric random variable. There is a growing body of work in exploring connections between PDEs and the expert problems in the pursuit of optimal algorithms (Andoni and Panigrahy, 2013; Bayraktar et al., 2020a,b; Drenska, 2017; Drenska and Kohn, 2020; Kobzar et al., 2020). Recently, Zhang et al. (2022) also used PDE techniques to obtain an optimal algorithm for unconstrained online linear optimization.

Quantile regret. Chaudhuri et al. (2009) introduced the notion of ε-regret where instead of comparing with the best expert, one compares with the $\lceil \varepsilon n \rceil$-th best expert (amongst n total experts). They devised the NormalHedge algorithm which they prove has an ε-quantile regret of $O(\sqrt{T \ln(1/\varepsilon)} + \ln^2 n)$. Moreover, the bound holds for all ε, T simultaneously. A somewhat different bound of $O(\sqrt{T(\ln \ln T + \ln 1/\varepsilon)})$ was proved by Luo and Schapire (2015) and Koolen and van Erven (2015). All of these works make use of a potential function to control the regret. Our work also makes use of a potential function which may be somewhat reminiscent of the potentials used by Chaudhuri et al. (2009) and Luo and Schapire (2015).

It is possible to improve upon the above bounds. Indeed, (Chernov and Vovk, 2010, Theorem 3), (Foster et al., 2015, Example 5.1), (Orabona and Pal, 2016, Corollary 6), and (Negrea et al., 2021, Corollary 2) show that it is possible to obtain an ε-quantile regret of $O(\sqrt{T \ln(1/\varepsilon)})$. This turns out to be tight up to constant factors (Negrea et al., 2021, Theorem 1). We note that Foster et al. (2015), Orabona and Pal (2016), and Negrea et al. (2021) derive regret bounds which depend on the KL divergence between a known prior and the player's probability distribution at a specific point of time; ε-quantile regret bounds can be recovered as a special case of such bounds. In this paper, we also recover the $O(\sqrt{T \ln(1/\varepsilon)})$ bound on the ε-quantile regret although, as we shall see, we obtain an improved constant in front of the $\sqrt{T \ln(1/\varepsilon)}$.

1.2 Basic Notation

We use $[n]$ to denote the set $\{1, ..., n\}$. For a predicate P, we write $[P]$ to be 1 if P is true and 0 otherwise. Moreover, if $[P]$ is multiplying an invalid expression (such as one with a division by 0) and P is false, we consider the whole expression to be 0. Set $[\alpha]_+ \coloneqq \max\{\alpha, 0\}$ for all $\alpha \in \mathbb{R}$. We use $\mathbb{1} \in \mathbb{R}^n$ to denote the all-ones vector and $e_i \in \mathbb{R}^n$ for $i \in [n]$ the indicator vector given by $e_i(j) \coloneqq [i = j]$ for all $j \in [n]$. We denote the $(n-1)$-dimensional probability simplex by $\Delta_n \coloneqq \{\, p \in [0,1]^n \,:\, \mathbb{1}^\mathsf{T} p = 1 \,\}$. For partial derivatives, we write $\partial_i \coloneqq \partial_{x_i}$ and $\partial_{ij} \coloneqq \partial_{x_i, x_j}$. Lastly, for $x \in \mathbb{R}^n$ and $\varepsilon \in (0,1)$, we write

$$\text{quantile}(\varepsilon, x) = x_{\pi(\lceil \varepsilon n \rceil)} \text{ where } \pi\colon [n] \to [n] \text{ is any permutation with } x_{\pi(1)} \geq \ldots \geq x_{\pi(n)}. \quad (1)$$

2. The Continuous Prediction Problem

As in the discrete experts' problem, we have a number $n \in \mathbb{N}$ of experts to choose from. In the continuous time setting, we model the cumulative gain of each expert as a mixture of n independent Brownian motions, as done by Freund (2009).

More formally, let $B_1, \ldots, B_n$ be n independent standard Brownian motion processes (or, equivalently, let B be a n-dimensional standard Brownian motion). The **cumulative gain process** $G_i(t)$ of expert $i \in [n]$ is given by the following stochastic differential equation (SDE)

$$\mathrm{d}G_i(t) \coloneqq \sum_{j=1}^{n} w_j^{(i)}(t)\,\mathrm{d}B_j(t) =: \langle w^{(i)}(t), \mathrm{d}B(t)\rangle, \qquad \forall t \geq 0, \forall i \in [n],$$

where $(w^{(i)}(t))_{t\geq 0}$ is any continuous stochastic process[4] in $\mathbb{R}^n$, not necessarily non-negative, such that $\left\|w^{(i)}(t)\right\|_2 = 1$ at all times $t \geq 0$. For example, if $w^{(i)}(t) = e_i$ for all $i \in [n]$ and $t \geq 0$, then $G^{(i)}(t)$ is an independent Brownian motion for each $i \in [n]$. An analogous situation in discrete time would be each expert receiving $\{\pm 1\}$ gain uniformly at random at each step, so each cumulative gain would be a standard random walk.[5] In our analysis the "instantaneous covariance matrix" $\Sigma(t)$ between the gain processes will be prominent. Formally, we define $\Sigma(t) \in \mathbb{R}^{n\times n}$ by

$$\Sigma_{ij}(t) \coloneqq \langle w^{(i)}(t), w^{(j)}(t)\rangle, \qquad \forall t \geq 0, \forall i, j \in [n].$$

From its definition, we have that $\Sigma(t)$ is a positive semi-definite matrix with ones along its diagonal.

Next, we define what a player strategy is in continuous-time and its corresponding regret. A **player (strategy)** is a left-continuous[6] process $(p(t))_{t\geq 0}$ on $\mathbb{R}^n$ such that $p(t) \in \Delta_n$ for all $t \geq 0$, where Δ_n is the $(n-1)$-dimensional simplex. The **player gain** process $(A(t))_{t\geq 0}$ is given by

$$\mathrm{d}A(t) \coloneqq \sum_{i=1}^{n} p_i(t)\,\mathrm{d}G_i(t) = \langle p(t), \mathrm{d}G(t)\rangle.$$

Moreover, the (continuous) **regret vector process** is given by

$$R_i(t) \coloneqq G_i(t) - A(t), \qquad \forall i \in [n], \forall t \geq 0.$$

That is, $R_i(t)$ is the regret—in the online learning sense—of the player with respect to expert i. Finally, the **continuous regret** (of the player strategy $(p(t))_{t\geq 0}$) is

$$\mathrm{ContRegret}(t) \coloneqq \max_{i\in[n]} R_i(t) = \max_{i\in[n]} G_i(t) - A(t)$$

Also, define the **continuous ε-quantile regret** to be $\mathrm{QuantRegret}(\varepsilon, t) \coloneqq \mathrm{quantile}(\varepsilon, R(t))$.

3. A Continuous Multiplicative Weights Update Method

In this section, we describe a continuous-time version of the classical Multiplicative Weights Update (MWU) method. This serves as a way to introduce some of our technical tools while avoiding the

4. Any stochastic process we mention in this paper is adapted to the filtration generated by $(B(t))_{t\geq 0}$
5. Intuitively, that is one of the reasons results in this setting mirrors the discrete-time case with costs in $[-1, 1]$.
6. One might loosen this to only assuming $(p(t))_{t\geq 0}$ is predictable. For a discussion, see Appendix E.

complexities we later introduce in the choice of potential function. Furthermore, we show bounds on its continuous regret that exactly match the bounds that the discrete algorithm enjoys, giving evidence of the parallels between the discrete and continuous time settings.

Analogous to the discrete version of MWU, we want the probability mass of an expert i at time t to be proportional to $\exp(\eta_t G_i(t))$, where η_t is some positive learning rate that is non-increasing in t. A familiar approach (see, e.g., Cesa-Bianchi and Lugosi, 2006, page 14 and Bubeck, 2011, §2.5) is to use the *LogSumExp* function given by

$$\Phi(t,x) \coloneqq \frac{[\eta_t > 0]}{\eta_t} \log\Big(\sum_{i=1}^{n} e^{\eta_t x_i}\Big) \quad \text{with } \eta_t \geq 0, \qquad \forall t \geq 0, \forall x \in \mathbb{R}^n. \tag{2}$$

In our case, the main property that we shall use from Φ is that $\nabla_x \Phi(t,\cdot)$ is the *softmax* function. That is, $\nabla_x \Phi(t,x) \in \Delta_n$ and $(\nabla_x \Phi(t,x))_i \propto \exp(\eta_t x_i)$, which is exactly the probability mass MWU places on expert i with cumulative gain x_i. Thus, we define the player strategy $(p(t))_{t\geq 0}$ by

$$p(t) \coloneqq \nabla_x \Phi(t, G(t)), \qquad \forall t \geq 0. \tag{3}$$

To analyze the regret of p, we need a way to handle $A(t) = \int_0^t \langle p(s), \mathrm{d}G(s)\rangle$. This is a stochastic integral, so we may use Itô's formula (Theorem C.1), which one can think of as the analogue of the fundamental theorem of calculus for stochastic integrals. Itô's formula coupled with the fact that $\Phi(t, G(t)) \geq \max_{i\in[n]} G_i(t)$ gives us the following lemma (whose proof we defer to Appendix D).

Lemma 3.1 *Let Φ be defined as in* (2) *and p be as in* (3)*. Then, almost surely,*

$$\mathrm{ContRegret}(T) \leq \Phi(0,0) + \int_0^T \Big(\partial_t \Phi(t, G(t)) + \frac{1}{2}\sum_{i,j\in[n]} \partial_{ij}\Phi(t, G(t))\Sigma_{ij}(t)\Big)\,\mathrm{d}t.$$

At this point, to bound the continuous regret of $(p(t))_{t\geq 0}$ it suffices to bound the partial derivatives of Φ. Lemma 3.2 bounds these partial derivatives in terms of a tunable learning rate η_t; minimizing the regret bound boils down to optimizing η_t. We defer the proof of the following lemma to Appendix D since it boils down to simple properties of Φ.

Lemma 3.2 *Let Φ be as in* (2)*. Let η_t be either constant in t or of the form $\frac{c}{\sqrt{t}}$, with $c > 0$. Then,*

$$\frac{1}{2}\sum_{i,j\in[n]} \partial_{ij}\Phi(t,x)\Sigma_{ij} \leq \frac{\eta_t}{2} \quad \textit{and} \quad \partial_t \Phi(t,x) \leq \frac{\log n}{2t\eta_t}, \qquad \forall t \geq 0, \forall x \in \mathbb{R}^n.$$

Theorem 3.3 summarizes the continuous regret bounds for MWU with properly chosen learning rates, both for the fixed-time and anytime settings. Crucially, these regret bounds match the best known regret bounds for the discrete-time MWU method (see Bubeck, 2011, Theorems 2.1 and 2.4).

Theorem 3.3 *Let Φ be as in* (2)*, and T be a positive number. If $\eta_t \coloneqq \sqrt{\ln n / 2T}$ for all $t \geq 0$, then $\mathrm{ContRegret}(T) \leq \sqrt{2T \ln n}$ almost surely. If $\eta_t \coloneqq [t > 0]\sqrt{\ln n / t}$ for all $t \geq 0$, then, almost surely, $\mathrm{ContRegret}(t) \leq 2\sqrt{t \ln n}$ for all $t \geq 0$.*

Proof Let us first consider the fixed-time case, that is, $\eta_t \coloneqq \sqrt{2 \ln n / T}$. In this case we have $\partial_t \Phi(t, x) = 0$ since $\Phi(\cdot, x)$ is constant for any $x \in \mathbb{R}^n$. Moreover, note that $\Phi(0,0) = \ln n/\eta_0 = \ln n/\eta_T$. Combining this with Lemmas 3.1 and 3.2, we have

$$\mathrm{ContRegret}(T) = \Phi(0,0) + \int_0^T \frac{1}{2} \sum_{i,j \in [n]} \partial_{ij}\Phi(t, G(t))\Sigma_{ij}\, \mathrm{d}t \leq \frac{\ln n}{\eta_T} + \frac{\eta_T T}{2} = \sqrt{2T \ln n}.$$

Let us now consider the anytime case, that is, when $\eta_t \coloneqq [t > 0]\sqrt{\ln n/t}$ for all $t \geq 0$. In this case we have $\Phi(0,0) = 0$, but $\partial_t \Phi(t,x)$ is not necessarily 0 anymore. By Lemma 3.2, we have

$$\partial_t \Phi(t, G(t)) + \frac{1}{2} \sum_{i,j \in [n]} \partial_{ij}\Phi(t, G(t))\Sigma_{ij} \leq \frac{\log n}{2t\eta_t} + \frac{\eta_t}{2} = \sqrt{\frac{\log n}{t}}, \qquad \forall t \geq 0.$$

Thus, for all $t \geq 0$, we have $\mathrm{ContRegret}(t) \leq \int_0^t \sqrt{\frac{\log n}{s}}\, \mathrm{d}s \leq 2\sqrt{t \log n}$. ■

It is intriguing that these bounds on the continuous regret differ by a factor of $\sqrt{2}$, exactly as in the discrete experts' problem. A natural question is whether there is an anytime algorithm that enjoys continuous regret bound smaller than $2\sqrt{t \log n}$. That is discussed in Section 5.

4. Quantile Regret Bounds with the Confluent Hypergeometric Potential

In this section, we design a different algorithm for the continuous prediction problem. We choose a potential function inspired by Itô's formula and obtain quantile regret bounds that are better than the ones known with a relatively simple proof. Furthermore, we show that this strategy has a simple discretization and obtain an algorithm with the same bounds for the discrete experts' problem. In Section 5 we shall see how a similar algorithm suggests an intriguing result for the anytime setting.

First of all, this time around we analyse a player strategy parameterized by a function of $R(t)$ instead of $G(t)$. That is, let $\Phi \colon \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ be a continuously differentiable function, which we refer to as a *potential function*. We consider the player strategy $(p(t))_{t \geq 0}$ given by[7]

$$p(t) \coloneqq \frac{1}{\mathbb{1}^{\mathsf{T}} \nabla_x \Phi(t, R(t))} \nabla_x \Phi(t, R(t)), \qquad \forall t \geq 0, \tag{4}$$

setting $p(t) \coloneqq \frac{1}{n}\mathbb{1}$ when $\mathbb{1}^{\mathsf{T}} \nabla_x \Phi(t, R(t)) = 0$. This class of player strategies mimics the potential-based strategies from the discrete experts' problems (Cesa-Bianchi and Lugosi, 2006, Chapter 2). As in the discrete case, if Φ is the LogSumExp potential from (2), we obtain the same player strategy of the last section. In the next lemma we use Itô's formula to get a useful expression for $\Phi(T, R(T))$ that, in turn, will guide us in the choice of Φ.

Lemma 4.1 *Let $\Phi \colon \mathbb{R}_{\geq 0} \times \mathbb{R}^n \to \mathbb{R}$ be one time continuously differentiable on its first argument and two-times continuously differentiable on its second argument. Let the player strategy $(p(t))_{t \geq 0}$*

7. Throughout this paper all entries of $\nabla_x \Phi(t, R(t))$ have the same sign, implying that $p(t) \in \Delta_n$. This $p(t)$ can be discontinuous when $\nabla_x \Phi(t, R(t)) = 0$, so we need to ensure it is predictable. This issue is discussed in Appendix E.

be as in (4)*. Then, almost surely for all* $T \geq 0$ *we have*

$$\Phi(T, R(T)) - \Phi(0,0) = \int_0^T \Big(\partial_t \Phi(t, R(t)) + \frac{1}{2} \sum_{i,j \in [n]} \partial_{ij} \Phi(t, R(t))(e_i - p(t))^\mathsf{T} \Sigma(t)(e_j - p(t))\Big)\, \mathrm{d}t. \tag{5}$$

In particular, if for all $t \geq 0$ *and* $x \in \mathbb{R}^n$ *we have* $\partial_{ij}\Phi(t,x) = 0$ *for all distinct* $i, j \in [n]$ *and* $\partial_{ii}\Phi(t,x) \leq 0$ *for each* $i \in [m]$*, then almost surely for all* $T \geq 0$ *we have*

$$\Phi(T, R(T)) - \Phi(0,0) \geq \int_0^T \Big(\partial_t \Phi(t, R(t)) + 2 \sum_{i=1}^n \partial_{ii} \Phi(t, R(t))\Big)\, \mathrm{d}t. \tag{6}$$

Proof Itô's formula gives us a useful formula to compute the evolution of the potential:

$$\Phi(T, R(T)) - \Phi(0,0) = \int_0^T \langle \nabla_x \Phi(t, R_t), \mathrm{d}R(t)\rangle + \int_0^T \partial_t \Phi(t, R(t))\, \mathrm{d}t + \frac{1}{2} \sum_{i,j \in [n]} \int_0^T \partial_{ij}\Phi(t, R(t))\, \mathrm{d}[R_i, R_j]_t.$$

For the first term above, note that $\langle \nabla_x \Phi(t, R(t)), \mathrm{d}R(t)\rangle = \mathbb{1}^\mathsf{T} \nabla_x \Phi(t, R(y)) \cdot \langle p(t), \mathrm{d}R(t)\rangle$ by the definition of $p(t)$ in (4). Furthermore, this is zero since

$$\langle p(t), \mathrm{d}R(t)\rangle = \langle p(t), \mathrm{d}G(t)\rangle - \langle p(t), \mathbb{1}\rangle\, \mathrm{d}A(t) = \mathrm{d}A(t) - \mathrm{d}A(t) = 0.$$

Finally, by Lemma F.1 we have $\mathrm{d}[R_i, R_j]_t = (e_i - p(t))^\mathsf{T}\Sigma(t)(e_j - p(t))\, \mathrm{d}t$ for all $i, j \in [n]$. This concludes the proof of (5).

Suppose that for all $t \geq 0$ and $x \in \mathbb{R}^n$ we have $\partial_{ij}\Phi(t,x) = 0$ for all distinct $i, j \in [n]$. Then,

$$\Phi(T, R_T) - \Phi(0,0) = \int_0^T \left(\partial_t \Phi(t, R(t)) + \frac{1}{2} \sum_{i=1}^n (e_i - p(t))^\mathsf{T} \Sigma(t)(e_i - p(t)) \partial_{ii} \Phi(t, R(t))\right) \mathrm{d}t.$$

Since $\Sigma(t)$ is a positive definite matrix with ones in its diagonal entries, we have $|\Sigma_{i,j}| \leq |\Sigma_{ii}| = 1$. Therefore, for any $v \in \mathbb{R}^n$, we have $v^\mathsf{T}\Sigma(t) v \leq \|v\|_1^2$. Thus, if $\partial_{ii}\Phi(t,x) \leq 0$ for all $i \in [n]$, then the second inequality stated in the lemma follows since $\|e_i - p(t)\|_1 \leq 2$ for all $i \in [n]$. ■

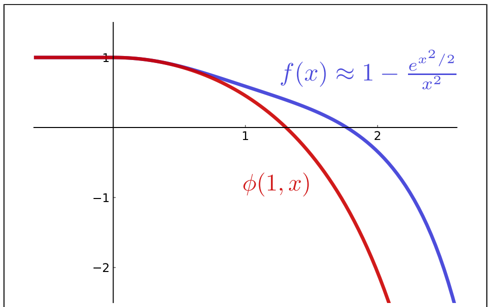

Figure 1: Plot of $\phi(1, x)$ in red and of the bound from Lemma A.3 in blue.

The second expression in Lemma 4.1 (Eq. (6)) hints at properties of potential functions Φ that may be particularly useful. More precisely, separable functions Φ that satisfy a diffusion constraint of the form $(\partial_t + 2\sum_{i=1}^n \partial_{ii})\,\Phi(t, \alpha) \geq 0$ would guarantee that $\Phi(t, R(t))$ is non-decreasing in t, which in turn may allow us to bound the continuous regret.

The player strategies in the rest of this paper involve the function M_0 defined as

$$M_0(\alpha) := e^{\alpha} - \sqrt{\pi\alpha}\, \mathrm{erfi}(\sqrt{\alpha}), \qquad \forall \alpha \in \mathbb{R}.$$

This is an example of a *confluent hypergeometric function* (of the first kind). We use M_0 in the form

$$\phi(t,\alpha) \coloneqq \sqrt{t} M_0\Big(\frac{[\alpha]_+^2}{2t}\Big), \quad \forall \alpha \in \mathbb{R}, \forall t > 0. \tag{7}$$

Similar functions have been used in the stochastic process literature (Breiman, 1967), (Davis, 1976), (Perkins, 1983), and in the online learning literature (Harvey et al., 2020b, eq. (2.6)) (Zhang et al., 2022, eq. (11)). Two particularly useful properties of ϕ are:

- $(\partial_t + \frac{1}{2}\partial_{xx})\phi(t,\alpha)$ is zero[8] for all $t > 0$ and $\alpha \geq 0$, and non-negative for all $\alpha < 0$. This is a PDE known as the *backwards heat equation* (BHE). Diffusion terms like these appear in Itô's formula, so functions satisfying the BHE are well-behaved under stochastic integration.
- $\phi(t,\alpha) \approx -\frac{1}{\alpha^2} e^{\alpha^2/2}$ (see Lemma A.3 and Figure 1), and so the potential resembles the normal distribution. Potentials of this form have been useful in the literature such as NormalHedge (Chaudhuri et al., 2009), AdaNormalHedge (Luo and Schapire, 2015). Moreover, Freund (2009) has used these normal-like potentials in continuous time.

The algorithm in this section uses the separable potential function Φ given by[9]

$$\Phi(t,x) \coloneqq \sum_{i=1}^{n} \phi\Big(t, \frac{x_i}{2}\Big) \qquad \forall t > 0, \forall x \in \mathbb{R}^n. \tag{8}$$

Lemma 4.2 *Let Φ and $(p(t))_{t\geq 0}$ be as in* (8) *and* (4)*, respectively. Then $\Phi(T, R(T)) \geq 0$.*

Note that if $\Phi(0,0) = 0$ then Lemma 4.2 would immediately follow from Lemma 4.1 (in particular, Eq. (6)) and our choice of Φ since Φ is concave and $\left(\partial_t + 2\sum_{i=1}^n \partial_{ii}\right)\Phi(t,x) \geq 0$. The only minor snag is that $\Phi(0,0)$ is not well-defined since Eq. (7) would involve a division by zero. Nonetheless, it is possible to resolve this issue; the details are relegated to Appendix F.1.

It remains to translate bounds on the value of the potential $\Phi(t, R(t))$ to bounds on the continuous regret. The following function is used in the regret bounds throughout the remainder of the paper.

Definition 4.3 *For $\alpha \in \mathbb{R}_{\geq 0}$, let $\lambda(\alpha) > 0$ be the unique positive solution to the equation*

$$\alpha = -M_0(\lambda(\alpha)^2/2) \equiv -\phi(1, \lambda(\alpha)). \tag{9}$$

We note that the function $M_0(x^2/2)$ is *strictly* decreasing on $\mathbb{R}_{\geq 0}$ and the image of M_0 on $\mathbb{R}_{\geq 0}$ is $(-\infty, 1]$ (see Appendix A). In particular, a solution to Eq. (9) exists and is unique so $\lambda(\alpha)$ is well-defined for all $\alpha \geq 0$. We also note that $\lambda(\alpha)$ is strictly increasing in α.

The next lemma show us how bounds in the value of $\Phi(t,x)$ can be translated into bounds on the quantiles of x, which were defined in (1).

8. Actually, ϕ is not doubly differentiable in its second argument because of the truncation in the definition of ϕ. Although this might seem like a problem to apply Itô's formula, we luckily have a single point of non-differentiability at each time $t \geq 0$, and the truncation only makes $(\partial_t + \frac{1}{2}\partial_{xx})\phi(t,x)$ no smaller everywhere else. Thus, standard smooth-truncation arguments can be made to apply Itô's formula. For an example, see Harvey et al. (2020a, Section 5.2.2). For the sake of simplicity, we set $\partial_{xx}\phi(t,0) \coloneqq \lim_{\varepsilon\to 0} \partial_{xx}\phi(t,\varepsilon)$
9. Ideally we would like to modify the potential by eliminating the denominator of 2. If this could be analyzed, it would yield an optimal quantile regret bound. At present, we have been unable to accomplish this.

Lemma 4.4 *Let $T > 0$ and $x \in \mathbb{R}^n$. Suppose that $\Phi(T, x) \geq 0$. Then, for any $\varepsilon \in (0, 1]$,*

$$\text{quantile}(\varepsilon, x) \;\leq\; 2\lambda\Big(\frac{1-\varepsilon}{\varepsilon}\Big)\sqrt{T} \;\leq\; 2\big(3 + \sqrt{2\ln(1/\varepsilon)}\big)\sqrt{T}.$$

Proof For simplicity, suppose $x_1 \geq x_2 \geq \ldots \geq x_n$. Since $\phi(T, \cdot)$ is decreasing, we have $\phi(t, x_{\varepsilon n}/2) \geq \phi(t, x_i/2)$ for every $i \leq \varepsilon n$. Summing this inequality for $i \in \{1, \cdots, \varepsilon n\}$, using the assumption $\sum_{i=1}^n \phi(t, x_i/2) = \Phi(T, x) \geq 0$ and since $-\phi(T, \alpha) \geq -\sqrt{T}$ for any $\alpha \in \mathbb{R}$, we have

$$\varepsilon n \phi(T, x_{\varepsilon n}/2) \geq \sum_{i=1}^{\varepsilon n} \phi(t, x_i/2) \geq -\sum_{i=\varepsilon n+1}^{n} \phi(t, x_i/2) \geq -(1-\varepsilon) n\sqrt{T}. \tag{10}$$

By the definition of ϕ, the above series of inequalities implies

$$\varepsilon n\sqrt{T} M_0\Big(\frac{([x_{\varepsilon n}]_+/2)^2}{2T}\Big) \geq -\big((1-\varepsilon)n\big)\sqrt{T} \implies M_0\Big(\frac{([x_{\varepsilon n}]_+/2)^2}{2T}\Big) \geq -\frac{1-\varepsilon}{\varepsilon} =: -\gamma.$$

Using the definition of λ and the fact that M_0 is a decreasing function, we have

$$M_0\Big(\frac{([x_{\varepsilon n}]_+/2)^2}{2T}\Big) \geq M_0\Big(\frac{\lambda(\gamma)^2}{2}\Big) \implies x_{\varepsilon n} \leq 2\lambda(\gamma)\sqrt{T}.$$

The second inequality in (10) follows from Lemma A.4. ■

Finally, we can combine Lemma 4.2 and Lemma 4.4 to prove a bound on the quantile regret in continuous-time. These quantile regret bounds improve upon the best known in the discrete case. In Section 4.1 we discretize this algorithm while preserving the same quantile regret bound.

Theorem 4.5 *Let Φ and $(p(t))_{t\geq 0}$ be as in* (8) *and* (4)*, respectively. Then*

$$\text{QuantRegret}(\varepsilon, T) \;\leq\; 2\lambda((1-\varepsilon)/\varepsilon)\sqrt{T} \;\leq\; 2\big(3 + \sqrt{2\ln(1/\varepsilon)}\big)\sqrt{T} \quad \textit{almost surely } \forall T \geq 0.$$

4.1 Discretization

In this section, we propose an algorithm for the original experts' problem based on the continuous-time solution of the previous section. As in the continuous setting, we have $n \in \mathbb{N}$ experts. At each round t, the player picks a probability vector $p_t \in \Delta_n$ and the adversary picks a gain vector $g_t \in [-1, 1]^n$. The **instantaneous regret vector** at round $t \geq 1$ is given by $r_t := g_t - \mathbb{1} \cdot p_t^{\mathsf{T}} g_t$. Moreover, define the **regret vector** at round t by $R_t := \sum_{s=1}^t r_s$.

To discretize the algorithm from the previous section, we shall make use of *discrete derivatives* in a way similar to Harvey et al. (2020a). For a bivariate function f, define its discrete derivatives as

$$\begin{aligned} f_t(t, x) &= f(t, x) - f(t-1, x), \\ f_x(t, x) &= \frac{f(t, x+1) - f(t, x-1)}{2}, \\ f_{xx}(t, x) &= (f(t, x+1) + f(t, x-1)) - 2f(t, x). \end{aligned} \tag{11}$$

Let Φ be defined as in Eq. (8). For $i \in [n]$, we define the discrete derivative of Φ as

$$\Phi_t(t, x) \;=\; \sum_{i=1}^n \phi_t\Big(t, \frac{x_i}{2}\Big); \quad \Phi_i(t, x) \;=\; \frac{1}{2}\phi_x\Big(t, \frac{x_i}{2}\Big); \quad \Phi_{ii}(t, x) \;=\; \frac{1}{4}\phi_{xx}\Big(t, \frac{x_i}{2}\Big).$$

For notation convenience, we also define the discrete gradient $\widetilde{\nabla}\Phi(t, x) = (\Phi_1(t, x), ..., \Phi_n(t, x))$.

Algorithm. The algorithm we use for the discrete setting is the natural analogue of the algorithm for the continuous setting as defined in Eq. (4). Specifically, for $t \in \mathbb{N}_{\geq 1}$, we set

$$p_t := \begin{cases} \frac{1}{n}\mathbb{1} & \text{if } \widetilde{\nabla}\Phi(t, R_{t-1}) = 0 \\ \frac{1}{\mathbb{1}^\mathsf{T}\widetilde{\nabla}\Phi(t,R_{t-1})}\widetilde{\nabla}\Phi(t, R_{t-1}) & \text{otherwise.} \end{cases} \tag{12}$$

Note that $\phi(t, x)$ is non-increasing in x so $\widetilde{\nabla}\Phi(t, x) \leq 0$ (component-wise), so $p_t \in \Delta_n$. Let us now analyze the performance of this algorithm. We summarize our results in the next theorem.

Theorem 4.6 *We have* $\mathrm{quantile}(\varepsilon, R_t) \leq 2\lambda((1-\varepsilon)/\varepsilon)\sqrt{t} \leq 2\sqrt{t} + \sqrt{8\ln(1/\varepsilon) + 20} \cdot \sqrt{t}$.

Let c be the optimal constant multiplying $\sqrt{t\ln(1/\varepsilon)}$ in the minimax optimal ε-quantile regret for anytime algorithms. Theorem 4.6 shows $c \leq 2\sqrt{2}$. Previously, Chernov and Vovk (2010, Theorem 3) and Negrea et al. (2021, Example 3) both proved[10] $c \leq 4$. On the lower bound side, Negrea et al. (2021, Theorem 1) proved that $c \geq \sqrt{2}$. So there remains a gap of 2 for the constant c. Finally, we note that if T is known beforehand (the fixed-time setting), Orabona and Pal (2016, Corollary 6) showed $c \leq 2\sqrt{3}$. Theorem 4.6 improves this to $c \leq 2\sqrt{2}$.

Interestingly, Zhang et al. (2022) proposed independently of us an algorithm using coin-betting with a potential similar to (8) also inspired by the work of Harvey et al. (2020b). Although their work is mostly on unconstrained online learning, one can obtain an algorithm for the experts' problem from their coin-betting algorithm whose ε-quantile regret is similar to our bound from Theorem 4.6, that is, roughly no more than $2\sqrt{2t\ln(1/\varepsilon)}$. More precisely, Orabona and Pal (2016, Theorem 4) show how to obtain an algorithm for the experts' problem with costs in $[0, 1]$ from a coin-betting algorithm together with regret guarantees against any comparison point $u \in \Delta_n$. We may obtain bounds on the ε-quantile regret by noting that it is the same as the regret against all points $u \in \Delta_n$ with $\lceil \varepsilon n \rceil$ non-zero equal entries. Combining this with Theorem 1 and Lemma B.2 of Zhang et al. (2022) yields the desired bound. In contrast, our analysis in this section is tailored specifically for the experts problem, and does not make use of the coin-betting framework.

To prove Theorem 4.6, we make use of the following lemma which is essentially a corollary of Harvey et al. (2020a, Lemma 3.13). It is similar to the discrete Itô's formula (Harvey et al., 2020a, Lemma 3.7) except that the equality is now an inequality. A proof can be found in Appendix G.1.

Lemma 4.7 *Let $x_1, x_2, \cdots$ be a sequence of real numbers such that $|x_t - x_{t-1}| \leq 1$. Then for any function f that is* concave *in its second argument and any integer $T \geq 2$, we have*

$$f(T, x_T) - f(1, x_1) \;\geq\; \sum_{t=2}^{T} f_x(t, x_{t-1}) \cdot (x_t - x_{t-1}) \;+\; \sum_{t=2}^{T} \left(\frac{1}{2} f_{xx}(t, x_{t-1}) + f_t(t, x_{t-1}) \right).$$

We now prove a discrete analogue of the second assertion of Lemma 4.1. Note that for technical reasons we start at $t = 1$ instead of $t = 0$. One can directly deal with the case $t = 0$ by customizing Eq. (7). However, this cumbersome approach does not yield improved bounds.

10. Note that these papers consider costs in $[0, 1]$ so their results must be multiplied by 2 to coincide with our setting.

Lemma 4.8 *Fix any $T \geq 2$. Then*

$$\begin{aligned}\Phi(T, R_T) - \Phi(1, R_1) \;&\geq\; \sum_{t=2}^{T}\left(\Phi_t(t, R_{t-1}) + 2\sum_{i=1}^{n}\Phi_{ii}(t, R_{t-1})\right) \\ &=\; \sum_{t=2}^{T}\sum_{i=1}^{n}\left(\phi_t(t, R_{t-1,i}/2) + \frac{1}{2}\phi_{xx}(t, R_{t-1,i}/2)\right).\end{aligned}$$

Proof Note that the equality follows from the definition of Φ_{ii} and Φ_t. Here, we prove that the term on the left is an upper bound on the final term on the right.

Recall that $\Phi(t, R_t) = \sum_{i=1}^{n}\phi(t, R_{t,i}/2)$. Since the gains are in $[-1, 1]$, it follows that $|(R_{t,i} - R_{t-1,i})/2| \leq 1$. From Lemma 4.7, since ϕ is concave in its second argument, we have

$$\begin{aligned}\sum_{i=1}^{n}\big(\phi(T, R_{t,i}/2) - \phi(1, R_{1,i}/2)\big) \;&\geq\; \sum_{t=2}^{T}\sum_{i=1}^{n}\phi_x(t, R_{t-1,i}/2)\cdot(R_{t,i} - R_{t-1,i}) \\ &\quad+ \sum_{t=2}^{T}\sum_{i=1}^{n}\left(\frac{1}{2}\phi_{xx}(t, R_{t-1,i}/2) + \phi_t(t, R_{t-1,i}/2)\right).\end{aligned} \tag{13}$$

To prove the lemma, it suffices to show that the first sum on the RHS of Eq. (13) is exactly 0. To that end, fix $t \in [T]$ such that $t \neq 1$. We have $R_{t,i} - R_{t-1,i} = g_{t,i} - p_t^{\mathsf{T}} g_t$. Hence,

$$\sum_{i=1}^{n}\phi_x(t, R_{t-1,i}/2)\cdot(R_{t,i} - R_{t-1,i}) \;=\; \sum_{i=1}^{n}\phi_x(t, R_{t,i}/2)\cdot(g_{t,i} - p_t^{\mathsf{T}} g_t). \tag{14}$$

If $\phi_x(t, R_{t-1,i}/2) = 0\ \forall i \in [n]$ then the RHS of (14) is 0. Otherwise, with the p_t as defined in (12),

$$p_t^{\mathsf{T}} g_t = \frac{\sum_{i=1}^{n}\phi_x(t, R_{t-1,i}/2)\cdot g_{t,i}}{\sum_{i=1}^{n}\phi_x(t, R_{t-1,i}/2)}. \tag{15}$$

Plugging (15) into (14) gives $\sum_{i=1}^{n}\phi_x(t, R_{t-1,i}/2)\cdot(R_{t,i} - R_{t-1,i}) = 0$, as required. ∎

In the continuous setting, we used the key fact that for any $t \in \mathbb{R}_{\geq 0}$ and $x \in \mathbb{R}$, we have $\left(\partial_t + \frac{1}{2}\partial_{xx}\right)\phi(t, x) \geq 0$. Luckily, the same fact holds in the discrete setting.

Lemma 4.9 *For any $t > 1$ and $x \in \mathbb{R}$, we have $\phi_t(t, x) + \frac{1}{2}\phi_{xx}(t, x) \geq 0$.*

The proof of Lemma 4.9 appears in Appendix G.2. We are now ready to prove Theorem 4.6.

Proof of Theorem 4.6 Lemma 4.8 and Lemma 4.9 imply that $\Phi(T, R_T) \geq \Phi(1, R_1)$ for all $T \geq 1$. Note that $\Phi(1, R_1) = \sum_{i=1}^{n}\phi(1, R_{1,i}/2) \geq n \cdot M_0(1/2) > n \cdot M_0(\gamma^2/2) = 0$. The bound on the quantile regret now follows from Lemma 4.4. ∎

5. Minimax Optimal Continuous Regret with Independent Experts

We have showed that an algorithm based on the hypergeometric potential (8) suffers no more than $2\sqrt{2t \ln n}$ at any round t (Theorem 4.5). Similarly, we saw that the *anytime* continuous MWU suffers at most $2\sqrt{t \ln n}$ regret at any round t (Theorem 3.3). A natural question is whether there are anytime algorithms in the continuous setting that enjoy regret better than $2\sqrt{t \ln n}$.

In the discrete version of the problem, we know that no algorithm can guarantee regret smaller than $\sqrt{2t \ln n}$. This lower-bound comes from the fact that the expected regret against an adversary that assigns ± 1 gains uniformly at random—or *uniformly random adversary*, for short—on the experts gets arbitrarily close to $\sqrt{2t \ln n}$ as n grows, regardless of the player's strategy. However, no tighter lower-bounds are known for *anytime* algorithms, that is, algorithms that do not know the length of the game. Furthermore, for two experts ($n = 2$) we know that *there is* a separation between the minimax optimal regret in the fixed-time and anytime settings, and both lower bounds arise from the expected regret against the uniformly random adversary! Namely, for fixed-time algorithms the best possible regret is $\sqrt{\frac{2T}{\pi}}$ (Cover, 1967) while for anytime algorithms the best possible regret is $\lambda(0)\sqrt{T}$ (where $\lambda(0) \approx 1.3069 > \sqrt{2/\pi} \approx 0.798$), both being the expected regret against the uniformly random adversary (the latter with a suitably chosen stopping-time).

Intriguingly, we show that in continuous time with *independent experts*—that is, when each $(G_i(t))_{t \geq 0}$ is an independent Brownian motion—there is an anytime algorithm whose regret is at most $\lambda(3n)\sqrt{t} \approx \sqrt{2t \ln n}$ for all $t \geq 0$. Furthermore, in Proposition 5.3 we show a matching lower-bound, just as in the discrete-time case. We conjecture that this algorithm can be discretized, and $\sqrt{2t \ln n}$ anytime regret against independent experts is possible in discrete time.

Algorithm. The player strategy we use in this section is similar from the one of the last section, with a crucial difference on Φ: we do not divide the second argument by 2. Namely,

$$\Phi(t, x) \coloneqq \sum_{i=1}^{n} \phi(t, x_i) \qquad \forall t \geq 0, \forall x \in \mathbb{R}^n, \tag{16}$$

and we set $(p(t))_{t \geq 0}$ as in (4) but using the above potential. We can still use Lemma 4.1 to get a formula for $\Phi(t, R(t))$. However, the lower-bound given in (6) is not of much use anymore since we do not have $(\partial_t + 2\sum_{i=1}^{n} \partial_{ii})\Phi(t, x) \geq 0$ anymore. Thus, we need to directly analyze the formula in (5). It turns out that when the instantaneous correlation matrix $\Sigma(t)$ is the identity matrix for all $t \geq 0$, we can show that this term does not become too negative. This term will be denoted

$$\mathrm{sBHT}(t) \coloneqq \partial_t \Phi(t, R(t)) + \frac{1}{2}\sum_{i=1}^{n} \partial_{ii}\Phi(t, R(t))(e_i - p(t))^{\mathsf{T}}(e_i - p(t))$$

That is, the sBHT is the integrand the appears in Lemma 4.1. Intuitively, since Φ satisfies the backwards-heat inequality, we should expect $\mathrm{sBHT}(t)$ to not be too negative. That is exactly what we show in the next lemma. A complete proof is in Appendix H, but we give a sketch here.

Lemma 5.1 *Let Φ be as in* (16). *Then* $\mathrm{sBHT}(t) \geq (2 - n)/\sqrt{t}$ *for all* $t > 0$.

Proof sketch For simplicity, fix $t > 0$ and assume $R_1(t) \geq R_2(t) \geq \cdots \geq R_n(t)$. Moreover, explicitly write the dependency of the sBHT on the regret vector by writing $\mathrm{sBHT}(t, R(t))$. The first step is to show that forcefully setting $R_n(t)$ to zero, denote such a vector vector by $\tilde{R}(t)$, can only decrease the sBHT. That is, $\mathrm{sBHT}(t, R(t)) \geq \mathrm{sBHT}(t, \tilde{R}(t))$. Then, we show that $\mathrm{sBHT}(t, \tilde{R}(t)) + 1/\sqrt{t}$ is equal to the sBHT restricted to the experts $1, \ldots, n-1$. Then, by induction we get that $\mathrm{sBHT}(t, R(t)) + (n-1)/\sqrt{t}$ is greater than the sBHT for a single expert, which one can verify that is at least $1/\sqrt{t}$, concluding the proof. ■

Finally, the above lemma together with the expression for $\Phi(t, R(t))$ given by (5) yields the desired regret bound when the instantaneous covariance matrix $\Sigma(t)$ is always the identity matrix.

Theorem 5.2 *Let Φ be defined as in* (16) *and let* $(p(t))_{t\geq 0}$ *be as in* (4). *Suppose* $\Sigma(t) = I\ \forall t \geq 0$. *Then, almost surely for all* $t \geq 0$ *we have* $\mathrm{ContRegret}(t) \leq \lambda(3n-1)\sqrt{t} \leq \sqrt{2t\ln n} + 6\sqrt{t}$.

Proof Let $T \in \mathbb{N}$. By Lemma 4.1 we have[11]

$$\Phi(T, R(T)) = \int_0^T \Big(\partial_t \Phi(t, R(t)) + \frac{1}{2}\sum_{i,j\in[n]} \partial_{ij}\Phi(t, R(t))(e_i - p(t))^{\mathsf{T}}\Sigma(t)(e_j - p(t))\Big)\,\mathrm{d}t. \quad (17)$$

Since $\Sigma(t) = I$ for all $t \geq 0$, the above integrand is exactly $\mathrm{sBHT}(t)$. Therefore, by Lemma 5.1,

$$(17) = \int_0^T \mathrm{sBHT}(t)\,\mathrm{d}t \geq \int_0^T \frac{2-n}{\sqrt{t}}\,\mathrm{d}t = 2(2-n)\sqrt{T} \geq -2n\sqrt{T}.$$

Finally, we translate this lower bound $\Phi(T, R(T))$ to an upper bound on $\mathrm{ContRegret}(T)$. Using Lemma 4.4, we have $\mathrm{ContRegret}(T) \leq \lambda(3n-1)\sqrt{t} \leq \sqrt{2t\ln(3n)} + 4\sqrt{t} \leq \sqrt{2t\ln n} + 6\sqrt{t}$.

A matching lower-bound. The next proposition shows that, against independent experts, the *expected* regret is always roughly $\sqrt{2t\ln n}$. This matches the discrete-time lower bound and shows that Theorem 5.2 is tight. The proof, given in Appendix H.1, is a straightforward modification of the analogous discrete-time result (Cesa-Bianchi and Lugosi, 2006, Theorem 3.7). It is important to note that Theorem 5.2 is considerably stronger than Proposition 5.3: the latter bounds the expectation separately for each t, whereas the former gives a bound that holds almost surely for all t.

Proposition 5.3 *Assume* $w^{(i)} = e_i$ *for each* $i \in [n]$. *Then, for any player strategy,*

$$\sqrt{2t\ln n}(1 - o(1)) \ \leq\ \mathrm{E}\left[\,\mathrm{ContRegret}(t)\,\right] \ \leq\ \sqrt{2t\ln n} \qquad \forall t > 0.$$

Acknowledgments

We would like to thank some anonymous reviewers for their feedback and for pointing out the connection with the work of Zhang et al. (2022).

References

Yasin Abbasi-Yadkori, Peter L. Bartlett, and Victor Gabillon. Near minimax optimal players for the finite-time 3-expert prediction problem. In *Annual Conference on Neural Information Processing Systems (NIPS)*, pages 3033–3042, 2017.

Alexandr Andoni and Rina Panigrahy. A differential equations approach to optimizing regret trade-offs. May 2013. URL `https://arxiv.org/abs/1305.1359`.

Erhan Bayraktar, Ibrahim Ekren, and Xin Zhang. Finite-time 4-expert prediction problem. *Communications in Partial Differential Equations*, 45(7):714–757, 2020a.

11. In fact, we should be careful with $\Phi(0,0)$ as we were in the proof of Theorem 4.5 in Appendix F. Since the exact same trick works in this case as well, we take $\Phi(0,0) := 0$ for the sake of simplicity.

Erhan Bayraktar, Ibrahim Ekren, and Yili Zhang. On the asymptotic optimality of the comb strategy for prediction with expert advice. *The Annals of Applied Probability*, 30(6):2517–2546, 2020b.

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.

Sébastien Bubeck. Introduction to online optimization, December 2011. URL `http://www.cse.iitd.ac.in/~naveen/courses/CSL866/BubeckLectureNotes.pdf`.

Nicolò Cesa-Bianchi and Gábor Lugosi. *Prediction, learning, and games*. Cambridge University Press, 2006. ISBN 978-0-521-84108-5.

Nicolò 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*, 44(3):427–485, 1997.

Kamalika Chaudhuri, Yoav Freund, and Daniel J. Hsu. A parameter-free hedging algorithm. In *Annual Conference on Neural Information Processing Systems (NIPS)*, pages 297–305, 2009.

Alexey V. Chernov and Vladimir Vovk. Prediction with advice of unknown number of experts. In *Proceedings of the Twenty-Sixth Conference on Uncertainty in Artificial Intelligence (UAI)*, pages 117–125, 2010.

Samuel N. Cohen and Robert J. Elliott. *Stochastic calculus and applications*. Probability and its Applications. Springer, Cham, second edition, 2015. ISBN 978-1-4939-2866-8; 978-1-4939-2867-5.

Thomas M. Cover. Behavior of sequential predictors of binary sequences. In *Trans. Fourth Prague Conference on Information Theory, Statistical Decision Functions, Random Processes*, pages 263–272. 1967.

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

Nadejda Drenska. *A PDE approach to a prediction problem involving randomized strategies*. PhD thesis, New York University, 2017.

Nadejda Drenska and Robert V Kohn. Prediction with expert advice: A PDE perspective. *Journal of Nonlinear Science*, 30(1):137–173, 2020.

Dylan J Foster, Alexander Rakhlin, and Karthik Sridharan. Adaptive online learning. *Advances in Neural Information Processing Systems (NIPS)*, 28:3375–3383, 2015.

Yoav Freund. A method for hedging in continuous time. October 2009. URL `http://arxiv.org/abs/0904.3356`.

Nick Gravin, Yuval Peres, and Balasubramanian Sivan. Towards optimal algorithms for prediction with expert advice. In *Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 528–547, 2016.

Nicholas J. A. Harvey, Christopher Liaw, Edwin Perkins, and Sikander Randhawa. Optimal anytime regret with two experts. February 2020a. URL `https://arxiv.org/abs/2002.08994v2`.

Nicholas J. A. Harvey, Christopher Liaw, Edwin A. Perkins, and Sikander Randhawa. Optimal anytime regret for two experts. In *61st IEEE Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1404–1415, 2020b.

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

Vladimir A. Kobzar, Robert V. Kohn, and Zhilei Wang. New potential-based bounds for prediction with expert advice. In *Conference on Learning Theory, (COLT)*, volume 125, pages 2370–2405, 2020.

Wouter M Koolen and Tim van Erven. Second-order quantile methods for experts and combinatorial games. In *Conference on Learning Theory (COLT)*, volume 40, pages 1155–1175, 2015.

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

Haipeng Luo and Robert E. Schapire. Achieving all with no parameters: AdaNormalHedge. In *Conference on Learning Theory (COLT)*, volume 40, pages 1286–1304, 2015.

Peter Mörters and Yuval Peres. *Brownian motion*, volume 30 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 2010.

Jeffrey Negrea, Blair Bilodeau, Nicolò Campolongo, Francesco Orabona, and Dan Roy. Minimax optimal quantile and semi-adversarial regret via root-logarithmic regularizers. *Advances in Neural Information Processing Systems (NeurIPS)*, 34, 2021.

Yurii Nesterov. Primal-dual subgradient methods for convex problems. *Math. Program.*, 120(1, Ser. B):221–259, 2009.

Frank W. J. Olver, Daniel W. Lozier, Ronald F. Boisvert, and Charles W. Clark, editors. *NIST handbook of mathematical functions*. Cambridge University Press, Cambridge, 2010.

Francesco Orabona and Dávid Pál. Optimal non-asymptotic lower bound on the minimax regret of learning with expert advice. November 2015. URL `http://arxiv.org/abs/1511.02176`.

Francesco Orabona and David Pal. Coin betting and parameter-free online learning. *Advances in Neural Information Processing Systems (NeurIPS)*, 29:577–585, 2016.

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

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

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

Martin J Wainwright. *High-dimensional statistics: A non-asymptotic viewpoint*, volume 48. Cambridge University Press, 2019.

Zhiyu Zhang, Ashok Cutkosky, and Ioannis Paschalidis. PDE-based optimal strategy for unconstrained online learning. January 2022. URL `https://arxiv.org/abs/2201.07877`.

A. Properties of the Confluent Hypergeometric Function

In this section we outline some of the main properties of the confluent hypergeometric function M_0 that we use throughout the paper. Many of the properties we use in this section come from Harvey et al. (2020a, Section 2.6).

Fact A.1 (Harvey et al., 2020a, Facts 2.4, 2.5, and 2.6) *For any $x \in \mathbb{R}$ we have*

(i) $M_0'(x) = -\frac{\pi}{2\sqrt{x}} \operatorname{erfi}(\sqrt{x})$

(ii) $M_0(x)$ *is strictly decreasing and concave on* $[0, \infty)$

From the above facts together with $M_0(0) = 1$ shows us that M_0 is strictly decreasing on $[0, +\infty)$ and its image over this domain is $(-\infty, 1]$ (since its derivative is negative and strictly decreasing). Furthermore, the above properties of M_0 allow us to derive many properties about the function $\phi(t, x) = \sqrt{t} \cdot M_0([x]_+^2/2t)$, as we show in the next lemma.

Lemma A.2 *Let $t \in \mathbb{R}_{\geq 0}$ and $x \in \mathbb{R}$. Then*

(i) $\phi(t, x)$ *is concave and non-increasing in* x*;*

(ii) $\partial_x \phi(t, x) = -\sqrt{\frac{\pi}{2}} \operatorname{erfi}(x/\sqrt{2t})$ *for* $x > 0$*;*

(iii) $\partial_{xx} \phi(t, x) = -\frac{1}{\sqrt{t}} \exp(x^2/2t)$ *for* $x > 0$*;*

(iv) $\partial_t \phi(t, x) = \frac{1}{2\sqrt{t}} \exp([x]_+^2/2t)$*;*

Proof Properties (i) and (ii) follow directly from Fact A.1 and the chain rule. Property (iii) follows from the fundamental theorem of calculus together with the chain rule since $\sqrt{\frac{\pi}{2}} \operatorname{erfi}(x) = \sqrt{2} \int_0^x e^{z^2} \, \mathrm{d}z$. For (iv), assume for notational simplicity only that $x > 0$. Then,

$$\begin{aligned}
\partial_t \left(\sqrt{t} M_0\left(\frac{x^2}{2t} \right) \right) &= \frac{1}{2\sqrt{t}} M_0\left(\frac{x^2}{2t} \right) - \frac{x^2}{2t^{3/2}} M_0'\left(\frac{x^2}{2t} \right) \\
&= \frac{1}{2\sqrt{t}} \left(\exp\left(\frac{x^2}{2t} \right) - \sqrt{\pi} \frac{x}{\sqrt{2t}} \operatorname{erfi}\left(\frac{x}{\sqrt{2t}} \right) - \frac{x^2}{t} \frac{\sqrt{\pi 2t}}{2x} \operatorname{erfi}\left(\frac{x}{\sqrt{2t}} \right) \right) \\
&= \frac{1}{2\sqrt{t}} \exp\left(\frac{x^2}{2t} \right).
\end{aligned}$$
∎

The next lemma gives an upper-bound to $M_0(x^2/2)$ for $x \geq 0$. Beyond other uses, this will be useful to upper-bound the regret bounds we derive with better-known functions.

Lemma A.3 *For every $x \geq 0$,*

$$1 - M_0(x^2/2) \;\geq\; \frac{\exp(x^2/2)}{x^2 + 1 + 2/x^2} \qquad \forall x \geq 0.$$

Proof Define $f(x) = 1 - M_0(x^2/2)$ and $g(x) = \exp(x^2/2)/(x^2 + 1 + 2/x^2)$. The derivatives are

$$f'(x) = \sqrt{\pi/2}\,\mathrm{erfi}(x/\sqrt{2}) \qquad \text{and} \qquad g'(x) = \exp(x^2/2) \cdot \frac{x^7 - x^5 + 2x^3 + 4x}{(x^4 + x^2 + 2)^2}.$$

The second derivatives are

$$f''(x) = \exp(x^2/2) \qquad \text{and} \qquad g''(x) = \exp(x^2/2) \cdot \frac{x^{12} - x^{10} + 9x^8 + 7x^6 - 32x^4 + 8x^2 + 8}{(x^4 + x^2 + 2)^3}.$$

We will show that $f''(x) \geq g''(x)$. By rearranging, this amounts to showing that

$$x^{12} - x^{10} + 9x^8 + 7x^6 - 32x^4 + 8x^2 + 8 \;\leq\; (x^4 + x^2 + 2)^3.$$

Expanding the right-hand side, we get

$$= \; x^{12} + 3x^{10} + 9x^8 + 13x^6 + 18x^4 + 12x^2 + 8.$$

The right-hand side coefficients are no smaller, which shows that $f''(x) \geq g''(x)$ for $x \geq 0$.

By integrating and using that $f'(0) = g'(0) = 0$, we obtain that $f'(x) \geq g'(x)$ for all $x \geq 0$. Finally, by integrating and using that $f(0) = g(0) = 0$, we obtain $f(x) \geq g(x)$ for all $x \geq 0$. ■

Recall from Definition 4.3 that λ is the non-negative inverse of $x \mapsto -M_0(x^2/2)$. The bound on M_0 from Lemma A.3 allows us to upper-bound λ as follows.

Lemma A.4 *Let $n \in \mathbb{R}$ be positive. Then,*

$$\lambda(n) \;\leq\; 3 + \sqrt{2\ln(n+1)}.$$

Consequently,

$$\lim_{n\to\infty} \frac{\lambda(n)}{\sqrt{2\ln(n)}} \;\leq\; 1.$$

Proof Define $\ell = 3 + \sqrt{2\ln(n+1)}$. Note that $\ell^2 = 2\ln(n+1) + 6\ell - 9$. So, by Lemma A.3,

$$1 - M_0(\ell^2/2) \;\geq\; \frac{\exp(\ell^2/2)}{\ell^2 + 1 + 2/\ell^2} \;=\; \frac{\exp(\ln(n+1) + 3\ell - 9/2)}{\ell^2 + 1 + 2/\ell^2}. \tag{18}$$

Since $\ell \geq 3$ we have $3\ell - 9/2 \geq \ell$, and also

$$e^{\ell} \;\geq\; \ell^2 + 1 + 2/\ell. \tag{19}$$

(This may be seen by a direct calculation for $\ell = 3$, then observing that the second derivative of the left-hand side exceeds the second derivative of the right-hand side for $\ell \geq 3$.) Combining (18) and (19) we obtain

$$1 - M_0(\ell^2/2) \;\geq\; \exp(\ln(n+1)) \;=\; n + 1.$$

Since $-M_0$ is monotonically increasing, it follows that $\lambda(n) \leq \ell$. ■

B. Additional Properties of the Continuous Expert's problem

From the definitions of Section 2, we can already obtain some useful properties. For example, for any $t \geq 0$ we have

$$\langle p(t), \mathrm{d}R(t)\rangle = \langle p(t), \mathrm{d}G(t)\rangle - \langle p(t), \mathbb{1}\rangle\, \mathrm{d}A(t) = \mathrm{d}A(t) - \mathrm{d}A(t) = 0. \tag{20}$$

At some points in ours proofs, it will be useful to write the weights of the different Brownian motions on the experts' gain processes in matrix form. Namely, define the continuous matrix-valued process $(W(t))_{t\geq 0}$ by

$$W(t) \cdot e_i \coloneqq w^{(i)}(t), \qquad \forall i \in [n], \qquad \forall t \geq 0.$$

That is, the i-th colum of $W(t)$ is $w^{(i)}(t)$. In particular, since $\Sigma_{i,j}(t) = \langle w^{(i)}(t), w^{(j)}(t)\rangle$, we have

$$\Sigma(t) = W(t)^{\mathsf{T}} W(t), \qquad \forall t \geq 0. \tag{21}$$

We use this definition in the following lemma, in which we directly write the regret process as stochastic integral with respect to the n-dimensional Brownian $(B(t))_{t\geq 0}$.

Lemma B.1 *For each $i \in [n]$ and all $t \geq 0$, we have $\mathrm{d}R_i(t) = \langle W(t)(e_i - p(t)), \mathrm{d}B(t)\rangle$.*

Proof Note that

$$\mathrm{d}A(t) = \langle p(t), \mathrm{d}G(t)\rangle = \sum_{i=1}^{n} p_i(t)\, \mathrm{d}G_i(t) = \sum_{i=1}^{n} p_i(t) \langle w^{(i)}, \mathrm{d}B(t)\rangle = \langle W(t)p(t), \mathrm{d}B(t)\rangle.$$

Thus,

$$\begin{aligned} \mathrm{d}R_i(t) = \mathrm{d}G_i(t) - \mathrm{d}A(t) &= \langle w^{(i)}(t), \mathrm{d}B(t)\rangle - \langle W(t)p(t), \mathrm{d}G(t)\rangle \\ &= \langle W(t)e_i, \mathrm{d}B(t)\rangle - \langle p(t), W(t)\, \mathrm{d}B(t)\rangle = \langle W(t)(e_i - p(t)), \mathrm{d}B(t)\rangle. \end{aligned}$$

■

C. Itô's Formula

In the following theorem we state Itô's formula as given by Revuz and Yor (1999). After the theorem statement we discuss some of the notation and applications the Itô's formula.

Theorem C.1 (Itô's Formula, Revuz and Yor, 1999, Theorem IV.3.3) *Let $F \colon \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable on its first argument and twice continuously differentiable on its second argument and let $(X_t)_{t\geq 0}$ be a continuous semimartingale in $\mathbb{R}^n$. Then, for any $T \geq 0$*

$$\begin{aligned} F(T, X(T)) - F(0, X(0)) = &\int_0^T \langle \nabla_x F(t, X(t)), \mathrm{d}X(t)\rangle + \int_0^T \partial_t F(t, X(t))\, \mathrm{d}t \\ &+ \frac{1}{2} \int_0^T \sum_{i,j\in[n]} \partial_{x_i, x_j} F(t, X(t))\, \mathrm{d}[X_i, X_j]_t \end{aligned}$$

In the third derivative above we use the bracket notation: for two continuous (local) martingales M and N, the process $[M,N]$, denoted as the **bracket of** M **and** N, is the unique adapted continuous process such that $MN - [M,N]$ is a local martingale (Revuz and Yor, 1999, Theorem IV.1.9). Although precise, this definition is not of much use for us since it does not gives us a way to compute this process.

Luckily, all the the process we deal with are defined as stochastic integrals with respect to other continuous martingales. For example, $A(t)$ is defined as (a sum of) stochastic integrals of a left-continuous and bounded function $p(t)$ with respect to the process $G(t)$. The latter is also a continuous martingale since it is a stochastic integral of a continuous and bounded function $w^{(i)}(t)$ with respect to the Brownian motion $B(t)$, which is also a martingale. This is specially useful to compute the bracket of two of these process. More specifically, using that

$$[B_i, B_j]_t = \begin{cases} 0 & \text{if } i \neq j, \\ t & \text{if } i = j, \end{cases} \tag{22}$$

we can compute the bracket of martingales by use of "box calculus" (Cohen and Elliott, 2015, Remark 14.2.7). Thus, for two continuous martingales M and N, we have

$$\mathrm{d}[M,N]_t = \mathrm{d}M(t) \cdot \mathrm{d}N(t),$$

and for the right-hand side above we usually can expand according to our definitions. In our case, we can always expand these expressions until they are written only in terms of the differentials of Brownian motions, and such expressions can be simplified using (22).

D. Missing Proofs of Section 3

Proof of Lemma 3.1 let $T > 0$. First of all, based on the remarks in Section C we have, for all $t \geq 0$,

$$\begin{aligned} \mathrm{d}[G_i, G_j]_t &= \mathrm{d}G_i(t)\,\mathrm{d}G_j(t) = \langle w^i(t), \mathrm{d}B(t)\rangle \cdot \langle w^j(t), \mathrm{d}B(t)\rangle \\ &= \sum_{k=1}^n w_k^{(i)}(t) \cdot w_k^{(j)}(t)\,\mathrm{d}t = \Sigma_{ij}(t)\,\mathrm{d}t, \end{aligned}$$

where in the last equation we used that $\Sigma_{ij} = \langle w^i(t), w^{(j)}(t)\rangle$.

Let $T > 0$. Itô's formula (Theorem C.1) allows us to express $A(T)$ as

$$\begin{aligned} A(T) &= \int_0^T \langle \nabla_x \Phi(t, G(t)), \mathrm{d}G(t)\rangle \\ &= \Phi(T, G(T)) - \Phi(0,0) + \int_0^T \Big(\partial_t \Phi(t, G(t)) + \frac{1}{2}\sum_{i,j\in[n]} \partial_{ij}\Phi(t, G(t))\Sigma_{ij}(t)\Big)\,\mathrm{d}t, \end{aligned}$$

Thus,

$$\begin{aligned} \mathrm{ContRegret}(T) = &\max_{i\in[n]} G_i(T) - \Phi(T, G(T)) + \Phi(0,0) \\ &+ \int_0^T \Big(\partial_t \Phi(t, G(t)) + \frac{1}{2}\sum_{i,j\in[n]} \partial_{ij}\Phi(t, G(t))\Sigma_{ij}(t)\Big)\,\mathrm{d}t. \end{aligned}$$

Finally, recall that the LogSumExp function smoothly approximates the maximum function, that is,

$$\max_{i\in[n]} x_i \leq \Phi(T,x) \leq \max_{i\in[n]} x_i + \frac{\log n}{\eta_T}, \qquad \forall x \in \mathbb{R}^n.$$

This implies that $\max_{i\in[n]} G_i(T) - \Phi(T, G(T)) \leq 0$. ■

Lemma D.1 *Let Φ be as in* (2) *and $\eta_t \geq 0$ for all $t \geq 0$. Then,*

$$\frac{1}{2}\sum_{i,j\in[n]} \partial_{ij}\Phi(t,x)\Sigma_{ij} \leq \frac{\eta_t}{2}, \qquad \forall t \geq 0, x \in \mathbb{R}^n.$$

Proof Define $\Theta \coloneqq (\sum_{i=1}^n \exp(\eta_t x_i))^2$. First of all, one may verify that

$$\partial_{ii}\Phi(t,x) = \frac{1}{\Theta}\eta_t e^{\eta_t x_i}\sum_{j\neq i} e^{\eta_t x_j} = \frac{1}{\Theta}\eta_t\Big(\sum_{j=1}^n e^{\eta_t(x_i+x_j)} - e^{2\eta_t x_i}\Big)$$

and that, for $i \neq j$,

$$\partial_{ij}\Phi(t,x) = -\frac{1}{\Theta}\eta_t e^{\eta_t(x_i+x_j)}.$$

Therefore, using that $\Sigma_{ii} = 1$ for any $i \in [n]$ and defining $v_i \coloneqq e^{\eta_t x_i}$ for each $i \in [n]$

$$\sum_{i,j\in[n]} \partial_{ij}\Phi(t,x)\Sigma_{ij} = \frac{\eta_t}{\Theta}\underbrace{\sum_{i,j\in[n]} e^{\eta_t(x_i+x_j)}}_{=\Theta} - \frac{\eta_t}{\Theta}\underbrace{\sum_{i,j} e^{\eta_t(x_i+x_j)}\Sigma_{ij}}_{=v^T\Sigma v\geq 0} \leq \eta_t.$$ ■

Lemma D.2 *Let Φ be as in* (2) *and η_t be either constant in t or of the form $[t > 0]c/\sqrt{t}$ for some $c > 0$. Then,*

$$\partial_t\Phi(t,x) \leq \frac{\log n}{2t\eta_t}, \qquad \forall t \geq 0, x \in \mathbb{R}^n.$$

Proof If η_t is constant as a function of t, then $\partial_t\Phi(t,x) = 0$. Otherwise, one may verify that (using the fact $\eta_t = \text{cst.}/\sqrt{t}$)

$$\begin{aligned}\partial_t\Phi(t,x) &= \frac{1}{2t\eta_t}\log\Big(\sum_{i=1}^n e^{\eta_t x_i}\Big) - \frac{1}{2\eta_t}\sum_{i=1}^n \frac{\eta_t}{t} x_i \frac{e^{\eta_t x_i}}{\sum_{j=1}^n e^{\eta_t x_j}} \\ &= \frac{1}{2t}\Bigg(\underbrace{\frac{1}{\eta_t}\log\Big(\sum_{i=1}^n e^{\eta_t x_i}\Big)}_{=\Phi(t,x)} - \underbrace{\sum_{i=1}^n x_i\frac{e^{\eta_t x_i}}{\sum_{j=1}^n e^{\eta_t C_j}}}_{=\langle C,p\rangle}\Bigg).\end{aligned}$$

Let us now show that

$$\frac{1}{\eta_t}\log\Big(\sum_{i=1}^n e^{\eta_t x_i}\Big) \leq \frac{\log n}{\eta_t} + \langle x, p\rangle, \tag{23}$$

which completes the proof of the lemma. We have

$$\frac{1}{\eta_t}\log\Big(\sum_{i=1}^n e^{\eta_t x_i}\Big) = \frac{\log n}{\eta_t} + \frac{1}{\eta_t}\log\Big(\sum_{i=1}^n \frac{1}{n}e^{\eta_t x_i}\Big).$$

Define $z_i \coloneqq \eta_t x_i$ for every $i \in [n]$. Then,

$$\begin{aligned}
&\frac{1}{\eta_t}\log\Big(\sum_{i=1}^n \frac{1}{n}e^{z_i}\Big) \leq \frac{1}{\eta_t}\sum_{i=1}^n \frac{e^{z_i}}{\sum_{j=1}^n e^{z_j}} z_i \\
\iff &\Big(\sum_{j=1}^n e^{z_j}\Big)\log\Big(\sum_{i=1}^n \frac{1}{n}e^{z_i}\Big) \leq \sum_{i=1}^n e^{z_i} z_i \\
\iff &\Big(\sum_{j=1}^n \frac{1}{n}e^{z_j}\Big)\log\Big(\sum_{i=1}^n \frac{1}{n}e^{z_i}\Big) \leq \sum_{i=1}^n \frac{1}{n}e^{z_i}\log(e^{z_i}),
\end{aligned}$$

and this last inequality is true by the convexity of $\alpha \in \mathbb{R}_{\geq 0} \mapsto \alpha \log \alpha$. This concludes the proof of (23) and, thus, of the lemma. ■

E. Ensuring Predictability of the Player Strategy

In Section 2, we required player strategies to be left-continuous in time. In fact, one could possibly loosen this assumption to only require $(p(t))_{t\geq 0}$ to be *predictable* (with respect to the filtration generated by $(B(t))_{t\geq 0}$) as defined in Revuz and Yor (1999, Definition IV.5.2). Yet, adapted left-continuous process are predictable (Mörters and Peres, 2010, Lemma 7.2) and are easier to reason about directly.

One should note, however, that the player strategies defined in (4) may not be left-continuous. This happens due to the discontinuity when the gradient entries sum to 0. For simplicity, we will discuss here how to modify player strategies generated by the potentials in (8) and (16) to ensure left-continuity, but similar techniques should for players generated by other potentials. In particular, the discontinuity problems may happen only when the gradient is 0. Finally, we note that the exact same predictability issue arises in the continuous NormalHedge algorithm due to Freund (2009).

Let us look at an example to see when p can be discontinuous and that it is not clear how to avoid the discontinuity in out case. Suppose $R(s) < 0$ for $s \in (t-\varepsilon, t]$ for some $t, \varepsilon > 0$, and that $R_1(s) > 0$ while $R_i(s) \leq 0$ for $s \in (t, t+\varepsilon)$. Then we for all $s \in (t-\varepsilon, t+\varepsilon)$ we have $p(s) = (1/n)\mathbb{1}$ if $s \leq t$ and $p(s) = e_1$ for $s > t$. Ideally, we would like a smooth transition between the uniform distribution and the point mass in the first expert, but there is no clear way to enforce that. In the case of the LogSumExp potential from Section 3, the gradient always lives in the *(relative) interior* of the simplex, so we never place 0 probability on any of the experts.

To ensure left-continuity of the player strategy, we can simply define it to be left-continuous. In our case, this will only modify the player strategy at the points of discontinuity we shall not affect our calculations. More precisely, let $(\hat{p}(t))_{t\geq 0}$ be the player strategy as described in (4) and define $(p(t))_{t\geq 0}$ by

$$p(t) \coloneqq \lim_{s \uparrow t} \hat{p}(s),$$

where $(R(t))_{t\geq 0}$ is still defined in terms of $(p(t))_{t\geq 0}$. One might worry that this definition becomes circular, but note that to define $p(t)$ we only need the values of $R(s)$ for $s < t$, and for $t = 0$ we have $R(t) = 0$. This ensures that $p(t)$ and $R(t)$ are well defined. Furthermore, since $(G(t))_{t\geq 0}$ is continuous, we also have that $(A(t))_{t\geq 0}$, and thus $(R(t))_{t\geq 0}$, are continuous, even though $(p(t))_{t\geq 0}$ may not be continuous (Cohen and Elliott, 2015, Remark 12.1.12).

Now *by definition* we have that p is left-continuous. Moreover, the points t of discontinuity for p are the points such that $R(t)$ enters or leaves the non-positive orthant $\mathcal{S} \coloneqq \{\, x \in \mathbb{R}^n : x \leq 0 \,\}$. Thus, we need only to ensure that any claims that explicitly use the form of p given by (4) also hold in the discontinuity points. In such points it is clear we have

$$\int_0^t \langle \nabla\Phi(s, R(s)), \mathrm{d}R(s)\rangle = 0,$$

as required by Lemma 4.1. This also should not affect the calculations in the proof of Lemma 5.1 since we can avoid the points of non-discontinuity my small perturbations without changing the value of the sBHT by much.

F. Missing Proofs for Section 4

Lemma F.1 *Let* $i, j \in [n]$. *We have* $\mathrm{d}[R_i, R_j]_t = (e_i - p(t))^T \Sigma(t)(e_j - p(t))\,\mathrm{d}t$.

Proof Using Lemma B.1 and the remarks in Section C we have

$$\begin{aligned}
\mathrm{d}[R_i, R_j]_t &= \mathrm{d}R_i(t)\cdot \mathrm{d}R_j(t) = \langle W(t)(e_i - p(t)), \mathrm{d}B(t)\rangle\langle W(t)(e_i - p(t)), \mathrm{d}B(t)\rangle \\
&= (e_i - p_t)^{\mathsf{T}} W^{\mathsf{T}} W(e_i - p_t) \qquad (\text{Since } \mathrm{d}B_i(t)\,\mathrm{d}B_j(t) = [i = j]\,\mathrm{d}t) \\
&= (e_i - p_t)^T \Sigma(e_i - p_t)\,\mathrm{d}t. \qquad (\text{By (21)})
\end{aligned}$$

■

F.1 Proof of Lemma 4.2

Proof of Lemma 4.2 Let $T \geq 0$. Intuitively, we want to show that $\Phi(T, R(T)) \geq 0$ by using Lemma 4.1. However, it is not clear what should be the value of $\Phi(0, 0)$. To handle this issue, let $\delta > 0$ and define $\Phi_\delta(t, x) \coloneqq \Phi(t + \delta, x)$ for all $t \geq 0$ and $x \in \mathbb{R}^n$, let $p^{(\delta)}$ be define as in (4) but replacing Φ by Φ_δ, and let R^δ be the continuous regret vector of p^δ. Our goal now is to show that

$$\Phi_\delta(T, R^\delta(T)) \geq \Phi_\delta(0, 0) = \sqrt{\delta} \text{ almost surely.} \tag{24}$$

Then, by taking the limit with δ tending to 0 we have $\Phi(T, R(T)) \geq 0$ almost surely. Furthermore, by a union bound we have $\Phi(t, R(t)) \geq 0$ for all rational $t \geq 0$, and since both Φ and R are continuous in t, this implies that $\Phi(t, R(t)) \geq 0$ for all $t \geq 0$ almost surely. Note, however, that there is a subtlety in the step in which we take the limit with $\delta \to 0$, since we are implicitly assuming that

$$\lim_{\delta\to 0} R^\delta(T) = R(T) \quad \text{almost surely.} \tag{25}$$

Let us prove that (25) indeed holds. Since $R_i^\delta(T) = G_i(T) - \int_0^T \langle p^{(\delta)}(t), \mathrm{d}G(t)\rangle$ and $G_i(T)$ is independent of δ for each $i \in [n]$, we only need to show that

$$\lim_{\delta\to 0} \int_0^T \langle p^{(\delta)}(t), \mathrm{d}G(t)\rangle = \int_0^T \langle p(t), \mathrm{d}G(t)\rangle \quad \text{almost surely.}$$

Since $p^{(\delta)}(t)$ is bounded and predictable (since it is left-continuous, see Appendix E) and G_i is a continuous martingale (since it is given by a stochastic integral of continuous functions with respect to a Brownian motion) for each $i \in [n]$, the above holds by the Dominated Convergence Theorem for stochastic integrals (Revuz and Yor, 1999, Theorem IV.2.12). It is also worth mentioning that we indeed have $\lim_{\delta \downarrow 0} p^\delta = p$ point-wise since taking δ to 0 would not make $x^2/2(t+\delta)$ cross the negative orthant where the points of discontinuity of p may be. This completes the proof of (25). We now proceed with the proof of (24).

Since Φ_δ is separable and $\phi(t+\delta, \cdot)$ is concave for any $t \geq 0$, by Lemma 4.1 we have

$$\Phi_\delta(T, R^\delta(T)) - \Phi_\delta(0,0) \geq \int_0^T \Big(\partial_t \Phi_\delta(t, R^\delta(t)) + 2\sum_{i=1}^n \partial_{ii} \Phi_\delta(t, R^\delta(t)) \Big)\, \mathrm{d}t.$$

Note that, for any $x \in \mathbb{R}^n$, we have $\partial_t \Phi_\delta(t,x) = \sum_{i=1}^n \partial_t \phi(t+\delta, x/2)$ and for all $i \in [n]$ we have $\partial_{ii}\Phi(t,x) = (1/4)\partial_{xx}\phi(t+\delta, x_i)$. Therefore,

$$\partial_t \Phi_\delta(t, R^\delta(t)) + 2\sum_{i=1}^n \partial_{ii}\Phi_\delta(t, R^\delta(t)) = \sum_{i=1}^n \Big(\partial_t \phi(t+\delta, R_i^\delta(t)) + \frac{1}{2}\partial_{xx}\phi(t+\delta, R_i^\delta(t)) \Big) \geq 0,$$

where the last equation holds since $\partial_t \phi(t,\alpha) + (1/2)\partial_{xx}\phi(t,\alpha) \geq 0$ for any $t > 0$ and $\alpha \in \mathbb{R}$. This implies that $\Phi_\delta(T, R^\delta(T)) \geq \Phi_\delta(0,0) = \sqrt{\delta}$ and $\Phi(T, R(T)) \geq 0$ by taking the limit $\delta \to 0$. ■

G. Missing Proofs for Section 4.1

G.1 Proof of Lemma 4.7

First, we require the following lemma. See, e.g., Klenke (2008, Example 10.9), Harvey et al. (2020a, Lemma 3.13).

Lemma G.1 (Discrete Itô's Formula) *Let $x_1, \cdots$ be a sequence of real numbers. Then for any function f and any fixed time $T \geq 2$, we have*

$$\begin{aligned} f(T, x_T) - f(1, x_1) = \sum_{t=2}^T f(t, x_t) &- \frac{f(t, x_{t-1}+1) + f(t, x_{t-1}-1)}{2} \\ &+ \sum_{t=2}^T \left(\frac{1}{2} f_{xx}(t, x_{t-1}) + f_t(t, x_{t-1}) \right). \end{aligned} \tag{26}$$

Proof of Lemma 4.7 We prove the following statement. Let f be a bivariate function that is concave in its second argument. Then for all $t, x \in \mathbb{R}$ and $y \in [-1,1]$ we have

$$f(t, x+y) - \frac{f(t,x+1) - f(t,x-1)}{2} \geq f_x(t,x) \cdot y.$$

Equality holds for $y \in \{-1, 1\}$. Since the LHS is concave in y and the RHS is linear in y, the inequality holds for all $y \in [-1,1]$. The lemma now follows by combining with Lemma G.1 with $x = x_{t-1}$ and $y = x_t - x_{t-1}$. ■

G.2 Proof of Lemma 4.9

We require the following lemma from Harvey et al. (2020a).

Lemma G.2 (Harvey et al., 2020a, Lemma 3.10) *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 of Lemma 4.9 Recalling the definition of the discrete derivatives (see Eq. (11)), we have that

$$\phi_t(t,x) + \frac{1}{2}\phi_{xx}(t,x) \;=\; \frac{\phi(t,x+1)+\phi(t,x-1)}{2} - \phi(t-1,x).$$

Hence, it suffices to prove that

$$\phi(t,x+1)+\phi(t,x-1) \geq 2\phi(t-1,x). \tag{27}$$

We consider several cases depending on the value of x.

Case 1: $x \geq 1$. In this case, Eq. (27) is equivalent to

$$\sqrt{t}\cdot M_0\left(\frac{(x+1)^2}{2t}\right) \;+\; \sqrt{t}\cdot M_0\left(\frac{(x-1)^2}{2t}\right) \;\geq\; 2\sqrt{t-1}\cdot M_0\left(\frac{x^2}{2(t-1)}\right). \tag{28}$$

Note that $t > 1$ so all terms are well-defined. Rearranging, this is equivalent to

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

which follows from Lemma G.2 by setting x and z in Lemma G.2 with $x/\sqrt{t}$ and $1/\sqrt{t}$, respectively.

Case 2: $x \in [0,1]$. Note that Eq. (28) holds for all $x \in \mathbb{R}$ (because Lemma G.2 holds for all $x \in \mathbb{R}$). Next, observe that $\phi(t,x) \geq M_0(x^2/2t)$ with equality whenever $x \geq 0$ (this is because $M_0(x^2/2t)$ is increasing on the interval $(-\infty, 0]$ while $\phi(t,x) = \phi(t,0)$ due to the truncation defined in Eq. (7)). So the LHS of Eq. (28) is upper bounded by the LHS of Eq. (27) and the RHS of Eq. (28) is equal to the RHS of Eq. (27). So Eq. (28) holds in this case as well.

Case 3: $x \in [-1,0]$. Note that $\phi(t,x+1)$ is the only term in Eq. (27) that is not constant on the interval $[-1,0]$. Further, note that Eq. (27) holds for $x=0$ by the previous case and it holds for $x=-1$ because the LHS is $2\sqrt{t}$ and the RHS is $2\sqrt{t-1}$. So the inequality holds since ϕ is concave in its second argument (Lemma A.2).

Case 4: $x \leq -1$. Due to the truncation in Eq. (7), Eq. (27) becomes $2\sqrt{t} \geq 2\sqrt{t-1}$. ■

H. Proof of Lemma 5.1

For our lower-bound, we do not need the sBHT to be evaluated at $R(t)$ to hold. Thus, we shall define some notation to analyse the sBHT evaluated at arbitrary points. Namely, let $x \in \mathbb{R}^n$ and $t > 0$. Define

$$p(t,x) \coloneqq \frac{1}{\mathbb{1}^\mathsf{T} \nabla_x \Phi(t,x)} \nabla_x \Phi(t,x), \qquad \forall t \geq 0,$$

setting $p(t,x) = (1/n)\mathbb{1}$ if $\mathbb{1}^\mathsf{T} \nabla_x \Phi(t,x) = 0$. Moreover, define

$$\mathrm{sBHT}(t,x) \coloneqq \partial_t \Phi(t,x) + \frac{1}{2} \sum_{i=1}^{n} \partial_{ii} \Phi(t,x) (e_i - p(t,x))^\mathsf{T} (e_i - p(t,x))$$

setting $p(t,x) \coloneqq \frac{1}{n}\mathbb{1}$ when $\mathbb{1}^\mathsf{T} \nabla_x \Phi(t,x) = 0$. Note now that we may assume $x \geq 0$. Indeed, assume $x_i < 0$ for some x_i. Then, due to the truncation in the definition of ϕ, we have $\partial_{ii}\Phi(t,x) = \partial_{xx}\phi(t,x_i) = 0$. Since $\partial_t \phi(t,x_i) = \frac{1}{\sqrt{2t}} \exp(\frac{x_i^2}{2t}) \geq 0$, we conclude that $\mathrm{sBHT}(t,x) \geq \mathrm{sBHT}(t, x - e_i x_i)$, that is, setting the i-th entry of x to zero can only decrease the value of the sBHT.

Furthermore, for the sake of simplicity assume $x_1 \geq x_2 \geq \cdots \geq x_n$ and set $p \coloneqq p(t,x)$. For each $i \in [n]$ define

$$q_i \coloneqq e^{x_i^2/2t}, \qquad Q_i \coloneqq \operatorname{erfi}\left(\frac{x_i}{\sqrt{2t}}\right),$$

$$\Theta \coloneqq \sum_{j=1}^{n} Q_j, \qquad p_i \coloneqq \frac{Q_j}{\Theta}.$$

Then, evaluating the derivatives according to Lemma A.2 and using that $p_i \propto \partial_x \phi(t, x_i) \propto Q_i$ we have

$$\begin{aligned} \mathrm{sBHT}(T,x) &= \frac{1}{2\sqrt{T}} \sum_{i=1}^{n} e^{x_i^2/2t} \left(1 - (e_i - p)^T (e_i - p)\right) \\ &= \frac{1}{2\sqrt{T}} \sum_{i=1}^{n} e^{x_i^2/2t} \left(2p_i - p^T p\right) \\ &= \frac{1}{2\Theta^2\sqrt{T}} \sum_{i=1}^{n} q_i \left(2\Theta Q_i - Q^T Q\right). \end{aligned}$$

Now, it suffices to show that

$$\partial_{x_n}(2\Theta^2 \sqrt{T}\, \mathrm{sBHT}(T,x)) = 2\sqrt{T} \partial_{x_n}(\Theta^2\, \mathrm{sBHT}(T,x)) \geq 0. \tag{29}$$

to prove the desired claim by induction. Indeed, note that if $\partial_{x_n}(\Theta^2\, \mathrm{sBHT}(T,x)) \geq 0$, then decreasing x_n all the way to zero only decreases the value of $\Theta^2\, \mathrm{sBHT}(T,x)$. More specifically, let $x' \coloneqq x - x_n e_n$ and define Θ', Q', and q' accordingly (that is, substituting x by x' in the original definition of these terms). In this case, we have $Q'_n = 0$ and $q'_n = 1$, yielding

$$\begin{aligned}
\Theta^2 \,\mathrm{sBHT}(T, x) \geq (\Theta')^2 \,\mathrm{sBHT}(T, x') &= \sum_{i=1}^{n} q'_i (2\Theta Q'_i - \langle Q', Q' \rangle) \\
&= \sum_{i<n} q'_i (2\Theta Q'_i - \langle Q', Q' \rangle) - \langle Q', Q' \rangle \\
&\geq \sum_{i<n} q'_i (2\Theta Q'_i - \langle Q', Q' \rangle) - \langle Q', Q' \rangle.
\end{aligned}$$

Dividing everything by $(\Theta')^2$, we get

$$\begin{aligned}
\left(\frac{\Theta}{\Theta'}\right)^2 \mathrm{sBHT}(T, x) &\geq \frac{1}{(\Theta')^2} \sum_{i<n} q'_i (2\Theta Q'_i - \langle Q', Q' \rangle) - \langle p', p' \rangle \\
&\geq \frac{1}{(\Theta')^2} \sum_{i<n} q'_i (2\Theta Q'_i - \langle Q', Q' \rangle) - 1.
\end{aligned}$$

If $\mathrm{sBHT}(T, x) \geq 0$, then the lower-bound we are trying to prove in the statement of Lemma 5.1 holds trivially. Otherwise, we have $\mathrm{sBHT}(T, x) \geq \left(\frac{\Theta}{\Theta'}\right)^2 \mathrm{sBHT}(T, x)$ since $\Theta \geq \Theta'$. Finally, the last summation in the above calculation is exactly the sBHT with $n - 1$ experts, and one can easily check that the sBHT for 1 expert is always at least 1. So by induction we have the claim of the theorem. Thus, let us finally proceed with the proof of (29).

First, let us summarize the properties on the partial derivatives:

$$\begin{aligned}
\partial_{x_n} Q_n &= \sqrt{\frac{2}{\pi}} q_n, & \partial_{x_n} q_n &= x_n q_n, \\
\partial_{x_n} \Theta = \partial_{x_n} Q_n &= \sqrt{\frac{2}{\pi}} q_n, & \partial_{x_n} (Q^T Q) &= 2\sqrt{\frac{2}{\pi}} Q_n q_n.
\end{aligned}$$

Then,

$$\begin{aligned}
&\partial x_n (\Theta^2 \,\mathrm{sBHT}(T, x)) \\
&= \sum_{i=1}^{n} \partial_{x_n} (q_i (2 Q_i \Theta - Q^T Q)) \\
&= \sum_{i<n} q_i (2 Q_i \partial_{x_n} \Theta - \partial_{x_n} (Q^T Q)) + (\partial_{x_n} q_n)(q_n (2 Q_n \Theta - Q^T Q)) + q_n \partial_{x_n} (2 Q_n \Theta - Q^T Q) \\
&= \sum_{i<n} 2\sqrt{\frac{2}{\pi}} q_i q_n (Q_i - Q_n) + x_n q_n (2 Q_n \Theta - Q^T Q) + q_n (2\Theta \partial_{x_n} (Q_n) + 2 Q_n \partial_{x_n} (\Theta) - \partial_{x_n} (Q^T Q)) \\
&= \sum_{i<n} 2\sqrt{\frac{2}{\pi}} q_i q_n (Q_i - Q_n) + x_n q_n (2 Q_n \Theta - Q^T Q) + \underbrace{2\sqrt{\frac{2}{\pi}} q_n^2 \Theta}_{\geq 0}. \qquad (30)
\end{aligned}$$

For the second term, since $Q_i \geq Q_n$ for any $i \in [n]$ we have

$$\begin{aligned} x_n q_n(2Q_n\Theta - Q^T Q) &= x_n q_n \sum_{i=1}^{n} (2Q_nQ_i - Q_i^2) \\ &= x_n q_n \sum_{i=1}^{n} Q_i(2Q_n - Q_i) \\ &\geq x_n q_n \sum_{i<n} Q_i(2Q_n - Q_i). \end{aligned}$$

Thus,

$$(30) \geq \sum_{i<n} \left(2\sqrt{\frac{2}{\pi}} q_i q_n (Q_i - Q_n) + x_n q_n Q_i (2Q_n - Q_i) \right).$$

Since $Q_i - Q_n \geq 0$, if $2Q_n - Q_i \geq 0$ we are done. Assume otherwise. The next lemma, which relies in a classical bound on erfi (Olver et al., 2010, Section 7.8), will be crucial for the rest of the proof.

Lemma H.1 *For any $z > 0$, we have*

$$\frac{\sqrt{\pi}}{2} \operatorname{erfi}(z) = \int_0^z e^{t^2} dt < \frac{e^{z^2} - 1}{x}. \tag{31}$$

In particular,

$$Q_i < 2\sqrt{\frac{2}{\pi}} \cdot \frac{q_i - 1}{x_i}, \qquad \forall i \in [n].$$

Proof The inequality from (31) can be found in Olver et al. (2010, Section 7.8). For the second inequality, note that

$$Q_i = \operatorname{erfi}\left(\frac{x_i}{\sqrt{2}}\right) \overset{(31)}{<} \frac{2}{\sqrt{\pi}} \cdot \frac{e^{x_i^2/2} - 1}{x_i/\sqrt{2}} = 2\sqrt{\frac{2}{\pi}} \cdot \frac{q_i - 1}{x_i}. \qquad \blacksquare$$

The above lemma together with $2Q_n - Q_i \leq 0$ implies, for each $i \in [n]$,

$$\begin{aligned} x_n q_n Q_i (2Q_n - Q_i) &\geq \underbrace{\frac{x_n}{x_i}}_{\leq 1} q_n 2\sqrt{\frac{2}{\pi}} \underbrace{(q_i - 1)}_{\leq q_i} (2Q_n - Q_i) \\ &\geq q_n 2\sqrt{\frac{2}{\pi}} q_i (2Q_n - Q_i). \end{aligned}$$

Thus,

$$\begin{aligned} &\sum_{i<n} \left(2\sqrt{\frac{2}{\pi}} q_i q_n (Q_i - Q_n) + x_n q_n Q_i (2Q_n - Q_i) \right) \\ \geq &\left(2\sqrt{\frac{2}{\pi}} q_i q_n (Q_i - Q_n) + q_n 2\sqrt{\frac{2}{\pi}} q_i (2Q_n - Q_i) \right) \\ = &2\sqrt{\frac{2}{\pi}} q_n q_i Q_n \geq 0. \end{aligned}$$

This completes the proof of (29) and, thus, of the lemma.

H.1 Proof of Proposition 5.3

Proof of Proposition 5.3 Since the functions in the stochastic differential equations in Section 2 are at least left-continuous and bounded, the stochastic integrals are well defined, vanish at time 0, and are martingales (Revuz and Yor, 1999, Def. IV.2.1 and IV.2.3). In particular, $\mathrm{E}\,[\,A(t)\,] = 0$ for all $t \geq 0$. For each $i \in [n]$ we have $G_i(t) = B_i(t)$ since $w^{(i)}(t) = e_i$. Thus each gain process is an independent Brownian motion. Thus,

$$\mathrm{E}\,[\,\mathrm{ContRegret}(t)\,] = \mathrm{E}\left[\max_{i \in [n]} B_i(t)\right] = \sqrt{2T \ln n}(1 - o(1)),$$

where in the last equation we used the well-known asymptotics for the maximum of n Gaussian random variables (e.g. Wainwright, 2019, Exercise 2.11 or Orabona and Pál, 2015, Theorem 3) and the fact that $B_i(t)$ follow a Gaussian distribution with mean zero and variance t. ■