JMLR: Workshop and Conference Proceedings vol 35:1–34, 2014

Higher-Order Regret Bounds with Switching Costs

Eyal Gofer EYALGOFE@CS.HUJI.AC.IL
The Rachel and Selim Benin School of Computer Science and Engineering
The Hebrew University
Givat Ram, Jerusalem 91904, Israel

Abstract

This work examines online linear optimization with full information and switching costs (SCs) and focuses on regret bounds that depend on properties of the loss sequences. The SCs considered are bounded functions of a pair of decisions, and regret is augmented with the total SC.

We show under general conditions that for any normed SC, $\sigma(\mathbf{x}, \mathbf{x}') = \|\mathbf{x} - \mathbf{x}'\|$, regret *cannot be bounded* given only a bound Q on the quadratic variation of losses. With an additional bound Λ on the total length of losses, we prove $O(\sqrt{Q + \Lambda})$ regret for Regularized Follow the Leader (RFTL). Furthermore, an $O(\sqrt{Q})$ bound holds for RFTL given a cost $\|\mathbf{x} - \mathbf{x}'\|^2$. By generalizing the Shrinking Dartboard algorithm, we also show an expected regret bound for the best expert setting with any SC, given bounds on the total loss of the best expert and the quadratic variation of any expert. As SCs vanish, all our bounds depend purely on quadratic variation.

We apply our results to pricing options in an arbitrage-free market with proportional transaction costs. In particular, we upper bound the price of "at the money" call options, assuming bounds on the quadratic variation of a stock price and the minimum of summed gains and summed losses.

Keywords: Online Learning, Regret Minimization, Switching Costs, Online Linear Optimization, Option Pricing

1. Introduction

Online linear optimization (OLO) models a wide range of sequential decision making problems in a possibly adversarial environment. In this setting, at each time step $t = 1, \ldots, T$ an online learning algorithm A (the *learner*) chooses a weight vector $\mathbf{x}_t$ taken from a non-empty compact and convex *decision (*or *action) set* $\mathcal{K} \subset \mathbb{R}^N$. Simultaneously, an adversary selects a loss vector $\mathbf{l}_t = (l_{1,t}, \ldots, l_{N,t}) \in \mathbb{R}^N$, and the algorithm experiences a loss $l_{A,t} = \mathbf{x}_t \cdot \mathbf{l}_t$. We denote $L_{A,t} = \sum_{\tau=1}^{t} l_{A,\tau}$ for the cumulative loss of A at time t and also $\mathbf{L}_t = \sum_{\tau=1}^{t} \mathbf{l}_\tau$. The aim of the learner is to achieve small *regret* w.r.t. the best fixed action with hindsight, regardless of the sequence of loss vectors chosen by the adversary. The regret is formally defined as $R_{A,T} = L_{A,T} - \min_{\mathbf{u}\in\mathcal{K}}\{\mathbf{u}\cdot\mathbf{L}_T\}$.[1] For randomized learners, the aim is to obtain small *expected* regret (alternatively, small with high probability). Importantly, it is assumed that $\mathbf{l}_t$ is also revealed to the learner once the loss $l_{A,t}$ is incurred, namely, a *full information* feedback mode. The alternative *bandit* feedback mode, which assumes that the learner is privy only to its own loss $l_{A,t}$, will not be considered in this paper.

An important special case of OLO is the classic *best expert* (BE) setting. This setting corresponds to picking the N-dimensional probability simplex Δ_N as the decision set $\mathcal{K}$. Thus, learner A's decision at time t is a probability vector $\mathbf{p}_t$, and the regret becomes $R_{A,T} = L_{A,T} - \min_j\{L_{j,T}\}$.

1. Thus, OLO is a special case of online convex optimization (OCO, Zinkevich, 2003), in which the learner incurs a loss of $f_t(\mathbf{x}_t)$, where f_t is a convex function chosen by the adversary, and the regret is $\sum_t f_t(\mathbf{x}_t) - \min_{\mathbf{u}\in\mathcal{K}}\{\sum_t f_t(\mathbf{u})\}$.

In this setting $\mathbf{l}_t$ may be thought of as the losses of N different experts, and the regret is measured w.r.t. the cumulative loss of the best expert.

In the above representation, $\mathbf{p}_t$ is a deterministically chosen point in the simplex, and may be seen as a fractional bet placed on *all* experts. This representation will be termed the *fractional view* of the BE setting. It is natural to also consider $\mathbf{p}_t$ as a random choice of a *single* expert at time t. Indeed, given a fractional learner, one may define another learner, which at time t picks vertex i of the simplex with probability $p_{i,t}$, for every i and t (a *randomized view* of the BE setting). Clearly, the expected regret of these two learners is the same. Nevertheless, they differ dramatically in their sequence of actions. For illustration, if $\mathbf{p}_t = (1/N, \ldots, 1/N)$ for every t, then the fractional learner stays put, while the randomized learner switches between vertices erratically.

In many real-world scenarios the amount of switching between decisions is material, and discouraging excessive switching is naturally motivated. In a physical system, changing configurations wastes energy. In an algorithmic system, changing algorithms implies overhead, delay, and possible user discomfort. Of particular interest for this work is the financial task of managing a portfolio, where switching between assets incurs *transaction costs* due to commissions, spreads, etc. This additional requirement may be incorporated in the OLO setting by defining a bounded SC function $\sigma : \mathcal{K} \times \mathcal{K} \to [0, B]$, where $B \geq 0$ and $\sigma(\mathbf{x}, \mathbf{x}) = 0$ for every $\mathbf{x} \in \mathcal{K}$, and aiming at minimizing the (expected) *augmented regret* $\widetilde{R}_{A,T} = R_{A,T} + \sum_{t=1}^{T-1} \sigma(\mathbf{x}_t, \mathbf{x}_{t+1})$ rather than the standard regret. This goal may be tackled in two conceptually different ways. One is to consider learners that are also *lazy*, that is, switch decisions a minimal number of times; indeed, such algorithms may accommodate any bounded σ. Another is to consider learners that move *smoothly*, namely, that keep $\sum_{t=1}^{T-1} \sigma(\mathbf{x}_t, \mathbf{x}_{t+1})$ small by making (possibly many) switches that are not costly given a specific σ. Both these approaches will be considered in this paper.

1.1. Existing Algorithms and Types of Regret Bounds

In the absence of SC, the most notable algorithm for the BE setting is the Hedge or Randomized Weighted Majority algorithm (Vovk, 1990; Littlestone and Warmuth, 1994; Freund and Schapire, 1997). Hedge takes as parameters a *learning rate* $\eta > 0$ and initial weights $w_{i,1} > 0$, $1 \leq i \leq N$. For each round $t = 1, \ldots, T$, the expert probabilities are defined by $p_{i,t} = w_{i,t}/W_t$, where $W_t = \sum_{i=1}^{N} w_{i,t}$, and after $\mathbf{l}_t$ is revealed, the update $w_{i,t+1} = w_{i,t} \exp(-\eta l_{i,t})$ is applied to the weights.

Given a bounded range for the losses $l_{i,t}$, Hedge may be shown to have vanishing average regret. If η is tuned solely as a function of time, we achieve the so-called zero-order regret bounds, which have the form $O(\sqrt{T \ln N})$. More refined bounds, called first-order bounds, are obtained when η is allowed to depend on L_T^*, the cumulative loss of the best expert. These bounds take the form $O(\sqrt{L_T^* \ln N})$. More recent second-order bounds obtained for various algorithms including Hedge (Cesa-Bianchi et al., 2007; Hazan and Kale, 2010; Chiang et al., 2012) replace the dependence on the time horizon T with quantities that measure the variability of the sequence of loss vectors $\mathbf{l}_t$. Such bounds generally offer an improvement over first-order bounds and may be significantly better in cases of low variability. Similarly, first-order bounds offer an improvement over zero-order bounds given an expert with small total loss. It should be noted that these various types of bounds are in general optimal.

One specific algorithm that achieves second-order bounds is the Polynomial Weights (PW) algorithm (Cesa-Bianchi et al., 2007), which replaces the exponential multiplicative update of Hedge with its first-order approximation $w_{i,t+1} = w_{i,t}(1 - \eta l_{i,t})$, for every i and t. The regret bound for

PW replaces the time horizon by a known bound on the quadratic variation of the losses of any expert, where the quadratic variation of an expert is the sum of its squared single-period losses.

When SCs are added, the randomized views of BE algorithms may generally incur $\Theta(T)$ total SCs.[2] This prompted the introduction of a lazy version of Hedge called the Shrinking Dartboard (SD) algorithm (Geulen et al., 2010), which was originally set in a context of online buffering. SD picks a single expert randomly at every round t and rarely switches it, but nevertheless has the same probability as Hedge to hold expert i at time t. It achieves the zero-order bound $O(\sqrt{T \ln N})$ on the expected augmented regret against an *oblivious* adversary (namely, one that fixes the losses in advance), and their proof implies a similar first-order bound as well.

For the more general OLO setting, in the absence of SCs, optimal regret bounds have been shown for the RFTL algorithm (Shalev-Shwartz and Singer, 2007; Hazan, 2011). This algorithm proceeds by "following the leader", namely, choosing the decision that would minimize the loss so far, but tempers its greediness by adding a regularization factor to the minimized expression. Formally, RFTL takes as parameters a learning rate $\eta > 0$ and a strongly convex regularizer function $\mathcal{R} : \mathcal{K} \to \mathbb{R}$, and on each round sets $\mathbf{x}_t = \arg\min_{\mathbf{x} \in \mathcal{K}} \{\mathbf{x} \cdot \mathbf{L}_{t-1} + (1/\eta)\mathcal{R}(\mathbf{x})\}$, where $\mathbf{L}_0 = \mathbf{0}$. RFTL generalizes both Hedge and the Lazy Projection variant of the Online Gradient Descent algorithm (OGD), defined by the rule $\mathbf{x}_{t+1} = \arg\min_{\mathbf{x} \in \mathcal{K}} \{\|\mathbf{x} + \eta \mathbf{L}_t\|_2\}$, via proper choices of $\mathcal{R}$. For a continuously twice-differentiable regularizer $\mathcal{R}$, RFTL guarantees, for a proper choice of η, an $O(\sqrt{T})$ regret bound. The work of Hazan and Kale (2010) showed second-order bounds for OGD in the OLO setting, which were later improved in Chiang et al. (2012) using a different algorithm.

When SCs are introduced, the "Follow the Leader" approach produces both lazy and smooth learners. The original Follow the Perturbed Leader algorithm and its lazy version, Follow the Lazy Leader (FLL) (Hannan, 1957; Kalai and Vempala, 2005), achieve $O(\sqrt{T})$ expected regret by using a random perturbation rather than regularization. For FLL a single large perturbation ensures $O(\sqrt{T})$ expected switches, and therefore $O(\sqrt{T})$ expected augmented regret against an oblivious adversary. OGD, on the other hand, has a smooth behavior and $O(\sqrt{T})$ augmented regret w.r.t. any normed SC $\sigma(\mathbf{x}, \mathbf{x}') = \|\mathbf{x} - \mathbf{x}'\|$ (see Andrew et al., 2013).

1.2. Contributions in This Paper

This work considers the problem of obtaining higher-order augmented regret bounds. As in the absence of SCs, the motivation is to achieve better bounds for sequences that are "easier" in some sense. Second-order bounds are of particular value because second-order quantities are linked to the standard deviation of random variables in general, and in finance, to the key volatility parameter.

Our first contribution is an infeasibility result. We show that for any normed SC there can be no augmented regret bound given only some bound on the quadratic variation $\sum_t \|\mathbf{l}_t\|_2^2$ (or other popular notions of variability). In particular, an adaptive (non-oblivious) adversary that knows the expectation of a learner's next action may force unbounded expected augmented regret. (The same is true for an oblivious adversary vs. a deterministic learner.) This result holds for general classes of the OLO setting, including the BE setting and any decision set that contains a ball around the origin (like the setting for OGD). It also holds for SCs like $\sigma(\mathbf{x}, \mathbf{x}') = 1_{\{\mathbf{x} \neq \mathbf{x}'\}}$ that are lower bounded by a normed SC. The proof technique is a novel application of the results of Gofer and Mansour (2012). Specifically, for every learner we construct a difficult sequence based on the behavior of the

2. If SCs are positive for each pair of vertices in Δ_N, then the only way for expected SCs to be $o(T)$ is for $\mathbf{p}_t$ to converge to a vertex of Δ_N. This does not hold in general for every loss sequence, certainly not for Hedge and PW.

learner and a second learner from the RFTL family. The properties of those algorithms ensure that the learner incurs either high standard regret or high SCs and non-negative standard regret.

Our second contribution consists of augmented regret bounds that circumvent our infeasibility result. For RFTL with a normed SC in the OLO setting we show an $O(\sqrt{Q+\Lambda})$ bound, where Q bounds the quadratic variation and Λ bounds the total length of the losses, $\sum_t \|\mathbf{l}_t\|_2$. We also give an $O(\sqrt{Q})$ bound for RFTL with the SC $\sigma(\mathbf{x}, \mathbf{x}') = \|\mathbf{x} - \mathbf{x}'\|^2$ for any norm. For the BE setting in particular, we derive expected augmented regret bounds that hold for any SC. This is done by generalizing the Shrinking Dartboard methodology to any algorithm that maintains strictly positive weights. Applied to PW and Hedge, this method yields bounds whose main terms are $O(\sqrt{(\mathcal{Q} + B(\mathcal{Q}+\mathcal{L}^*)) \ln N})$ for PW and $O(\sqrt{(q + B\mathcal{L}^*) \ln N})$ for Hedge. In these bounds, N is the number of experts, $\mathcal{Q}$ bounds $\sum_t l_{i,t}^2$ for any expert i, B bounds σ values, $\mathcal{L}^*$ bounds L_T^*, and q bounds the relative quadratic variation $\sum_t (\max_i\{l_{i,t}\} - \min_i\{l_{i,t}\})^2$. Importantly, the dependence of the above bounds on Λ and $\mathcal{L}^*$ disappears in the absence of SCs.

We apply our results to the financial problem of option pricing and specifically to pricing *"at the money" call options*.[3] In the financial setting, the fractional view of BE algorithms provides a recipe for allocating wealth among a set of assets. For call options, the assets are stock and cash. More importantly, regret bounds for such algorithms imply bounds on option prices in an arbitrage-free market (DeMarzo et al., 2006; Gofer and Mansour, 2012; Gofer, 2013).[4] We extend analysis given in Gofer and Mansour (2012) to a scenario where trading incurs proportional transaction costs. For "at the money" call options the augmented regret results obtained for SD yield a price upper bound of $\exp(2c \ln 2 + \sqrt{(q/2 + 8c\lambda^*) \ln 2}) - 1$. In this bound, q bounds the quadratic variation of the stock's log price ratios, c is the transaction cost rate for the stock, and λ^* bounds the minimum of summed gains and summed losses (in log price ratio terms). When $c = 0$, this bound has the same asymptotic behavior as the Black and Scholes bound for small values of q, namely, $\Theta(\sqrt{q})$.

We conclude by adapting our infeasibility result to show that similar methods can yield only trivial price bounds for these options, given only a bound on the quadratic variation. This provides an alternative regret-based proof for this well-known finance result in our model.

1.3. Related Work

The lazy approach to learning with SCs has been considered as a special case of learning against adversaries with bounded memory (Merhav et al., 2002). Specifically, it takes memory of depth 1 of the learner's actions for the adversary's loss values to simply incorporate the SC. Their $O(T^{2/3})$ regret bound for the full information case holds if a cost of 1 is incurred for any switching, but it is suboptimal. For the bandit setting, recent works have proved a $\widetilde{\Theta}(N^{1/3}T^{2/3})$ bound for N arms (experts) which is optimal in T (Arora et al., 2012; Cesa-Bianchi et al., 2013; Dekel et al., 2013), highlighting a fundamental difference between the two feedback models. Arora et al. (2012) have also shown $o(T)$ regret bounds for the OCO setting.

The FLL algorithm was modified in the work of Devroye et al. (2013) to work with perturbations that are independent symmetric random walks, retaining its performance guarantees. The work of György and Neu (2011) modified SD to work with an unknown time horizon, using variable learning rates, and employed it to solve the limited-delay universal lossy source coding problem.

3. A call option is a security that pays its holder at time T the sum of $\max\{S_T - K, 0\}$, where S_t is the price of a given stock at time t, and K is a set price, called the *strike price*. The option is "at the money" if $K = S_0$.

4. In an arbitrage-free market, no algorithm trading in financial assets can guarantee profit without any risk of losing money. For a randomized algorithm, we assume that even expected profit may not be guaranteed.

The smoothed approach in the OCO framework was considered by Andrew et al. (2013) for normed (and even seminormed) SCs. They concluded that OGD-type algorithms given by Zinkevich (2003) and Hazan et al. (2007) have the same order of augmented and standard regret bounds, namely, $O(\sqrt{T})$ or $O(\log T)$ (the latter holding for exp-concave losses).

The smoothed view was also considered by the online algorithms community in works on the metrical task system problem (Borodin et al., 1992). Here the decision space is finite, the losses are not necessarily even convex, and the SC function is a metric. Importantly, the learner knows the next loss and the goal is to minimize the competitive ratio rather than the (standard) regret. For works that consider interpolating between these two different goals, see (Blum and Burch, 2000; Buchbinder et al., 2012; Bera et al., 2013; Andrew et al., 2013). Interestingly, Andrew et al. (2013) showed that there are OCO problems with normed SC for which it is impossible to simultaneously obtain sublinear augmented regret and a finite competitive ratio.

Arbitrage-free option pricing with proportional transaction costs has been considered in the finance literature for the continuous-time, stochastically-based Black-Scholes-Merton (BSM) model as well as its discrete counterpart, the binomial model (see Musiela and Rutkowski, 1997 for more details). In the BSM model the cheapest way to almost surely super-replicate a call option (that is, to dominate its payoff) is to buy and hold the stock. This holds for any positive volatility, transaction cost rates, and strike price, and implies a trivial price bound. The same holds for the binomial model as the trading frequency goes to infinity.

In the learning literature, adversarial derivative pricing based on second-order regret bounds was pioneered by DeMarzo et al. (2006) for call options and extended to exotic (non-standard) derivatives by Gofer and Mansour (2011a,b). Optimal asymptotic lower bounds for the price of "at the money" call options were given in Gofer and Mansour (2012). These works show that the adversarial price of these call options behaves like $\sqrt{Q}$, up to multiplicative factors, if Q is the assumed quadratic variation of the stock's log price ratios, for small values Q. This asymptotic behavior matches the BSM pricing. Other works priced exotic options by super-replication (Dawid et al., 2011b,a; Koolen and Vovk, 2012), but did not consider second-order quantities.

Adversarial call option pricing may also be derived by considering the exact value of the game between the adversarial market and the trader (DeMarzo et al., 2006). Recently, Abernethy et al. (2012, 2013) have shown that given bounds on the quadratic variation of stock returns and the magnitude of price jumps, the strongest adversary is BSM's stochastic price process. Their works apply to call options and to more general payoffs. For "at the money" call options, the bound of Abernethy et al. (2013) behaves like $O(Q^{1/8})$ for an asymptotically small Q. These works as well as those based on regret minimization assumed the absence of transaction costs.

Numerous learning works deal with the problem of portfolio selection, where the aim is to maximize returns. Of those, some provide no adversarial performance guarantees (e.g., Li and Hoi, 2012), while others, beginning with Cover (1991), provide them w.r.t. rich classes of investment strategies. The pricing problems we consider require adversarial guarantees w.r.t. a small set of assets, making both types of results generally unhelpful. We briefly comment that for some of the rich strategy classes considered, there are provable regret bounds even in the presence of transaction costs. One class is *constantly rebalanced portfolios* (CRPs), namely, strategies that always keep a constant fraction of funds in each asset. For CRPs, the Universal Portfolio algorithm of Cover (1991) has been analyzed with transaction costs by Blum and Kalai (1999). Another benchmark includes all switching strategies (Singer, 1998).

Outline Section 2 provides some additional definitions. In Section 3 we give the infeasibility result for pure second-order bounds. Section 4 presents augmented regret bounds for RFTL. Bounds derived by generalizing the Shrinking Dartboard method are given in Section 5. Applications to option pricing are given in Section 6. We conclude the paper in Section 7. Note that background and technical results pertaining to the financial applications are deferred to Appendix A due to space constraints.

2. Additional Notation and Preliminaries

We next cover some additional facts, conventions and notation that will be needed later on.

It will be assumed that for a switching cost function, defined by $\sigma : \mathcal{K}\times\mathcal{K} \to [0, B]$, the value B is known to the learner. We note that all normed SCs are equivalent in the sense that for every norm $\|\cdot\|$ on $\mathbb{R}^N$ there are $c_1, c_2 > 0$ s.t. $c_1\|\mathbf{x}\|_1 \le \|\mathbf{x}\| \le c_2\|\mathbf{x}\|_1$ for every $\mathbf{x} \in \mathbb{R}^N$; as a result, all such norms are bounded on a compact $\mathcal{K}$.[5] In particular, $\|\mathbf{x}\|_\infty \le \|\mathbf{x}\|_p \le \|\mathbf{x}\|_1 \le N\|\mathbf{x}\|_\infty$ for every $\mathbf{x} \in \mathbb{R}^N$ and $p \in [1,\infty]$. That said, SCs do not have to be even metrics, as with the cost $\|\mathbf{x} - \mathbf{x}'\|^2$, which we consider. On the other hand, not every metric is bounded on a compact $\mathcal{K}$. While allowing unbounded costs might be considered, that would clearly invalidate the lazy approach.

We will sometimes rewrite the total SC as $\sum_{t=1}^{T} \sigma(\mathbf{x}_{t-1}, \mathbf{x}_t)$, defining $\mathbf{x}_0 = \mathbf{x}_1$ s.t. $\sigma(\mathbf{x}_0, \mathbf{x}_1) = 0$. Note that even if $\mathbf{x}_0$ were taken to be an arbitrary initial state, its effect would be bounded by B. We denote by K_t the number of changes to the decision vector in the first t rounds, namely, $K_t = \sum_{\tau=1}^{t} 1_{\{\mathbf{x}_{\tau-1}\neq\mathbf{x}_\tau\}}$. Thus, performing at most K_T switches means a total SC of at most BK_T.

We will consider primarily the quadratic variation of an entire loss sequence $\mathbf{l}_1, \ldots, \mathbf{l}_T$, defined as $Q_T = \sum_{t=1}^{T} \|\mathbf{l}_t\|_2^2$. For the BE setting in particular it is useful to consider the slightly different notion of *relative quadratic variation*, defined as $q_T = \sum_{t=1}^{T} \delta(\mathbf{l}_t)^2$, where $\delta(\mathbf{v}) = \max_i\{v_i\} - \min_i\{v_i\}$ for any $\mathbf{v} \in \mathbb{R}^N$. We assume $Q_T \le Q$ and $q_T \le q$ and that these bounds are known to the learner. Note that $q_T \le 2Q_T$, but it may be that $q_T = 0$ while Q_T is arbitrarily large.

In the BE setting we will sometimes use the notation $m(t) = \arg\min_i\{L_{i,t}\}$ for the index of the best expert, where the smallest such index is taken in case of a tie.

Our results make regular use of the properties of convex functions defined on $\mathbb{R}^N$. For a thorough coverage, see Rockafellar (1970), Boyd and Vandenberghe (2004), and Nesterov (2004), among others. We will denote $[\mathbf{x}, \mathbf{y}]$ for the line segment between $\mathbf{x}$ and $\mathbf{y}$, namely, $\{a\mathbf{x}+(1-a)\mathbf{y} : 0 \le a \le 1\}$. In addition, the convex conjugate of a function f will be denoted by f^*.

3. Infeasibility of Pure Second-Order Regret Bounds for Normed SCs

In this section we prove an impossibility result for obtaining second-order bounds for important classes of OLO problems given a normed SC. Specifically, we will show that any learner may suffer arbitrarily high augmented regret for some loss sequences with arbitrarily small quadratic variation. Note that this claim extends to the notions of variation considered by Hazan and Kale (2010) and Chiang et al. (2012) and also to relative quadratic variation, since those are all dominated, up to multiplicative factors, by the quadratic variation. We will consider a deterministic learner for simplicity, but the claim holds for randomized learners as well, as will be shown.

We first give the proof idea. For any learner A, we show the existence of a loss sequence for which the learner incurs either high standard regret or high SCs along with non-negative standard

5. The same holds for any seminorm, except that c_1 is non-negative.

regret. This is portrayed for convenience as constructing the loss $\mathbf{l}_t$ adaptively assuming knowledge of the learner's next decision vector $\mathbf{x}_t$. The loss vectors are all collinear with a pre-fixed vector $\mathbf{v}$ and have pre-fixed absolute sizes. However, the sign, or direction, of each loss vector is learner-dependent. Thus, for every t, $\mathbf{l}_t = d_t a_t \mathbf{v}$, where $a_t > 0$ and $d_t \in \{-1, 1\}$. Importantly, one may choose a sequence $\{a_t\}_{t=1}^{\infty}$ s.t. the cumulative path length $\|\mathbf{v}\|_2 \sum_{t=1}^{\infty} a_t$ is infinite and yet the quadratic variation $\|\mathbf{v}\|_2^2 \sum_{t=1}^{\infty} a_t^2$ is arbitrarily small.

The key element of the proof is that the direction d_t is decided by observing the next decision $\mathbf{x}'_t$ of a second online algorithm, denoted A', as well as the next decision $\mathbf{x}_t$ of the learner A. More concretely, *before* deciding on the direction of $\mathbf{l}_t$, the losses of A and A' in the upcoming round are compared, assuming we chose $d_t = d_{t-1}$ (where $d_0 = 1$); if A stands to do better than A', that is, $\mathbf{x}_t \cdot d_{t-1} a_t \mathbf{v} < \mathbf{x}'_t \cdot d_{t-1} a_t \mathbf{v}$, then the direction is reversed, otherwise, it is left the same. This construction ensures that the cumulative loss and regret of A are never better than those of A'.

We use a deterministic algorithm A' that calculates its decisions $\mathbf{x}'_t$ as the gradient of a concave potential function of the cumulative losses. Namely, $\mathbf{x}'_t = \nabla\Phi(\eta \mathbf{L}_{t-1})$ for a concave potential $\Phi : \mathbb{R}^N \to \mathbb{R}$ with a learning rate $\eta > 0$. We draw heavily on the special properties of these algorithms, and in particular, their non-negative or even strictly positive regret for any loss sequence (Gofer and Mansour, 2012). Another property of the second-order remainders of potentials in that family enables us to lower bound the SC of A' and A by $\Omega(\sum_{t=1}^{T} a_t)$.

We point out that all the details of the above construction may be known to the learner in advance, making the losses entirely predictable. Nevertheless, for any interval $[-\alpha\mathbf{v}, \alpha\mathbf{v}]$, the values $\mathbf{L}_t$ inevitably either oscillate indefinitely within the interval or at some time T depart it. In the former case, we show that SCs rise arbitrarily with $\sum_t a_t$, while the regret of A, lower bounded by the regret of A', is non-negative. In the latter case we show that $R_{A,T} \geq R_{A',T} = \Omega(1/\eta)$, which can be made arbitrarily high by picking an arbitrarily small η. It follows that in both cases the learner's augmented regret may reach any level we desire, whereupon the game may be stopped.

We now proceed to prove the result. We will assume the existence of a vector $\mathbf{v} \in \mathbb{R}^N$, a continuously twice-differentiable concave function $\Phi : \mathbb{R}^N \to \mathbb{R}$, and a scalar $\lambda > 0$, with the following properties:

- For every $\mathbf{L} \in \mathbb{R}^N$ it holds that $\nabla\Phi(\mathbf{L}) \in \mathcal{K}$.
- It holds that $\xi_1 = \inf_{|s|\geq\lambda}\{\Phi(s\mathbf{v}) - \Phi(\mathbf{0}) - \min_{\mathbf{x}\in\mathcal{K}}\{\mathbf{x}\cdot s\mathbf{v}\}\} > 0$.
- It holds that $\xi_2 = \inf_{|s|\leq\lambda}\{-\mathbf{v}^\top\nabla^2\Phi(s\mathbf{v})\mathbf{v}\} > 0$.

Such a triplet $(\mathbf{v}, \Phi, \lambda)$ will be termed *admissible*, and its existence will be justified later. We point out that the last requirement implies in particular that $\mathbf{v} \neq \mathbf{0}$.

The above vector $\mathbf{v}$ will be used when defining $\mathbf{l}_t = d_t a_t \mathbf{v}$. In addition, note that for any $\epsilon > 0$, one may set $a_t = \sqrt{6\epsilon}/(\pi t)$ for every $t = 1, 2, \ldots$ and satisfy both $\sum_{t=1}^{\infty} a_t = \infty$ and $\sum_{t=1}^{\infty} a_t^2 = \epsilon$. Now, for any $\eta > 0$ we denote $\Phi_\eta(\mathbf{L}) = (1/\eta)\Phi(\eta\mathbf{L})$ and may define A' by $\mathbf{x}'_t = \nabla\Phi_\eta(\mathbf{L}_{t-1}) = \nabla\Phi(\eta\mathbf{L}_{t-1})$, where $\mathbf{L}_0 = \mathbf{0}$. It then follows that for every T,

$$R_{A',T} \geq \Phi_\eta(\mathbf{L}_T) - \Phi_\eta(\mathbf{0}) - \min_{\mathbf{x}\in\mathcal{K}}\{\mathbf{x}\cdot\mathbf{L}_T\} \geq 0 \tag{1}$$

(Gofer and Mansour, 2012, Corollary 1). These considerations lead to the following lemma:

Lemma 1 *Let $(\mathbf{v}, \Phi, \lambda)$ be admissible, let $\eta > 0$, and define A' by $\mathbf{x}'_t = \nabla\Phi_\eta(\mathbf{L}_{t-1})$. If for some T it holds that $s = \sum_{t=1}^{T} d_t a_t$ satisfies $|s| \geq \lambda/\eta$, then $R_{A,T} \geq R_{A',T} \geq \xi_1/\eta$.*

Proof By Equation 1, $R_{A',T} \geq \frac{1}{\eta}\left(\Phi(\eta \mathbf{L}_T) - \Phi(\mathbf{0}) - \min_{\mathbf{x}\in\mathcal{K}}\{\mathbf{x}\cdot\eta\mathbf{L}_T\}\right)$. Since $\eta\mathbf{L}_T = (s\eta)\mathbf{v}$, $|s\eta| \geq \lambda$, and $\xi_1 = \inf_{|s'|\geq\lambda}\{\Phi(s'\mathbf{v}) - \Phi(\mathbf{0}) - \min_{\mathbf{x}\in\mathcal{K}}\{\mathbf{x}\cdot s'\mathbf{v}\}\} > 0$, we obtain $R_{A',T} \geq \xi_1/\eta$. As argued before, $R_{A,T} \geq R_{A',T}$ and the result follows. ■

Next, consider a case in which the values $\mathbf{L}_t$ all remain in the interval $[-\lambda\mathbf{v}/\eta, \lambda\mathbf{v}/\eta]$. This implies an unbounded number of direction reversals, and the SCs of A may be lower bounded as follows:

Lemma 2 *If for every $t \leq T$ it holds that $\mathbf{L}_t \in [-\lambda\mathbf{v}/\eta, \lambda\mathbf{v}/\eta]$, and $d_T \neq d_{T-1}$, then*

$$\sum_{t=1}^{T-1}\|\mathbf{x}_{t+1} - \mathbf{x}_t\|_1 > \frac{\eta\xi_2}{\|\mathbf{v}\|_\infty}\sum_{t=1}^{T-1} a_t .$$

The proof is in the appendix. Lemma 2 shows that to avoid arbitrarily large SCs, the vector $\mathbf{L}_t$ must leave the interval $[-\lambda\mathbf{v}/\eta, \lambda\mathbf{v}/\eta]$ at some point. Otherwise, the learner incurs arbitrarily large augmented regret, since $R_{A,t} \geq R_{A',t} \geq 0$ for any t (see Equation 1). However, if the cumulative loss does leave the interval, then by Lemma 1 a regret of at least ξ_1/η is incurred. Since $\eta > 0$ is arbitrarily small, the augmented regret in either case can be made arbitrarily large. We can now prove the following:

Theorem 3 *For any OLO problem with an admissible triplet, any normed SC, and any $Q > 0$, no learner may guarantee bounded augmented regret for every loss sequence with quadratic variation smaller than Q. In addition, no deterministic learner may guarantee such a bound against an oblivious adversary that can simulate the learner, and no learner may guarantee such a bound in expectation against an adaptive adversary that knows the expectation of the learner's next decision.*

We comment that Theorem 3 clearly holds even if the SC is only lower bounded by a normed SC.

We conclude by proving the existence of admissible triplets for two general classes of OLO problems. One is the BE setting, and the other includes all cases for which $\mathcal{K} \supseteq B(\mathbf{0}, a)$, where $B(\mathbf{0}, a)$ is the closed ball with radius a centered at $\mathbf{0}$, for some $a > 0$. For these settings the existence of entire classes of admissible triplets is implied by the results of Gofer and Mansour (2012), but note that for our purposes we require only a single representative per setting.

For the BE setting, we set $\Phi(\mathbf{L}) = -\ln((1/N)\sum_{i=1}^{N} e^{-L_i})$, which is the potential function of the Hedge algorithm, along with $\mathbf{v} = (1, 0, \ldots, 0)$ and any $\lambda > 0$. For the case $\mathcal{K} \supseteq B(\mathbf{0}, a)$, we set $\Phi(\mathbf{L}) = \min_{\mathbf{x}\in\mathcal{K}}\{\mathbf{x}\cdot\mathbf{L} + \frac{1}{2}\|\mathbf{x}\|_2^2\}$, which is the potential function of OGD with lazy projection, along with $\mathbf{v} = (1, 0, \ldots, 0)$, again, and $\lambda = a$. It is thus implied that the reference algorithm A', which uses the gradient of Φ_η, would in fact be Hedge in the former setting and OGD in the latter. More details may be found in the proof of the following corollary, given in the appendix:

Corollary 4 *Let $\mathcal{K}$ be the decision set of an OLO problem. If $\mathcal{K} = \Delta_N$ (the BE setting) or $\mathcal{K} \supseteq B(\mathbf{0}, a)$ for some $a > 0$, then for any learner A and any $Q > 0$, there exist loss sequences with quadratic variation smaller than Q for which A incurs arbitrarily large augmented regret.*

4. Higher-Order Augmented Regret Bounds for RFTL

The result of the previous section may be circumvented in two ways: One is to impose additional restrictions on the losses, and the other is to assume SCs other than normed ones. In this section we pursue both these routes and provide two types of bounds on the augmented regret of RFTL in the OLO setting. One bound is for normed SCs given an additional bound on the *length* of the cumulative loss path, $\sum_{t=1}^{T} \|\mathbf{l}_t\|_2$. The other depends purely on the quadratic variation, given that $\sigma(\mathbf{x}, \mathbf{x}') \leq c\|\mathbf{x} - \mathbf{x}'\|_2^2$. In what follows, $RFTL(\eta, \mathcal{R})$ stands for RFTL with learning rate η and a regularizer $\mathcal{R}$. Some useful properties of RFTL are given in the next theorem.

Theorem 5 *If $\eta > 0$ and $\mathcal{R} : \mathcal{K} \to \mathbb{R}$ is continuous and strongly convex with parameter α, then $\Phi(\mathbf{L}) = (-1/\eta)\mathcal{R}^*(-\eta\mathbf{L})$ is concave and continuously differentiable on $\mathbb{R}^N$, and for every $\mathbf{L} \in \mathbb{R}^N$, it holds that $\nabla\Phi(\mathbf{L}) = \arg\min_{\mathbf{x}\in\mathcal{K}}\{\mathbf{x} \cdot \mathbf{L} + \mathcal{R}(\mathbf{x})/\eta\}$ and $\Phi(\mathbf{L}) = \min_{\mathbf{x}\in\mathcal{K}}\{\mathbf{x} \cdot \mathbf{L} + \mathcal{R}(\mathbf{x})/\eta\}$. Furthermore, $\nabla\Phi$ is Lipschitz continuous with parameter η/α, namely, for any $\mathbf{L}, \mathbf{L}' \in \mathbb{R}^N$, $\|\nabla\Phi(\mathbf{L}') - \nabla\Phi(\mathbf{L})\|_2 \leq (\eta/\alpha)\|\mathbf{L} - \mathbf{L}'\|_2$.*

The Lipschitz continuity of $\nabla\Phi$ is proven in Appendix C. The rest of the above claims are found in Gofer and Mansour (2012) and their proofs are therefore omitted. We will continue to refer to the function $\Phi(\mathbf{L}) = \min_{\mathbf{x}\in\mathcal{K}}\{\mathbf{x} \cdot \mathbf{L} + \mathcal{R}(\mathbf{x})/\eta\}$ in what follows, and will also use the notation $D = \max_{\mathbf{u},\mathbf{v}\in\mathcal{K}}\{\mathcal{R}(\mathbf{u}) - \mathcal{R}(\mathbf{v})\}$. Next, the above theorem is applied in proving a general second-order regret bound for RFTL.

Theorem 6 *If $\eta > 0$ and $\mathcal{R} : \mathcal{K} \to \mathbb{R}$ is continuous and strongly convex with parameter α, then $R_{RFTL(\eta,\mathcal{R}),T} \leq D/\eta + \eta Q_T/\alpha$, and for $\eta = \sqrt{D\alpha/Q}$, it holds that $R_{RFTL(\eta,\mathcal{R}),T} \leq 2\sqrt{DQ/\alpha}$.*

We next consider SCs, where the Lipschitz continuity of $\nabla\Phi$ is key. It holds for every t that

$$\|\mathbf{x}_{t+1} - \mathbf{x}_t\|_2 = \|\nabla\Phi(\mathbf{L}_t) - \nabla\Phi(\mathbf{L}_{t-1})\|_2 \leq (\eta/\alpha)\|\mathbf{L}_t - \mathbf{L}_{t-1}\|_2 = (\eta/\alpha)\|\mathbf{l}_t\|_2 .$$

Thus, for a SC that satisfies $\sigma(\mathbf{x}, \mathbf{x}') \leq c\|\mathbf{x} - \mathbf{x}'\|_2$ we have that

$$\sum_{t=1}^{T-1} \sigma(\mathbf{x}_t, \mathbf{x}_{t+1}) \leq (\eta c/\alpha) \sum_{t=1}^{T-1} \|\mathbf{l}_t\|_2 ,$$

and for a SC satisfying $\sigma(\mathbf{x}, \mathbf{x}') \leq c\|\mathbf{x} - \mathbf{x}'\|_2^2$ we obtain

$$\sum_{t=1}^{T-1} \sigma(\mathbf{x}_t, \mathbf{x}_{t+1}) \leq (\eta^2 c/\alpha^2) \sum_{t=1}^{T-1} \|\mathbf{l}_t\|_2^2 \leq (\eta^2 c/\alpha^2) Q_T .$$

Together with Theorem 6 these observations lead to the following theorem:

Theorem 7 *(i) Let Λ be a known upper bound on $\sum_{t=1}^{T} \|\mathbf{l}_t\|_2$, the cumulative loss path length. If $\sigma(\mathbf{x}, \mathbf{x}') \leq c\|\mathbf{x} - \mathbf{x}'\|_2$, then $\widetilde{R}_{RFTL(\eta,\mathcal{R}),T} \leq D/\eta + \eta Q_T/\alpha + (\eta c/\alpha) \sum_{t=1}^{T} \|\mathbf{l}_t\|_2$, and for $\eta = \sqrt{\frac{D\alpha}{Q+c\Lambda}}$, it holds that $\widetilde{R}_{RFTL(\eta,\mathcal{R}),T} \leq 2\sqrt{(D/\alpha)(Q + c\Lambda)}$.*

(ii) If $\sigma(\mathbf{x}, \mathbf{x}') \leq c\|\mathbf{x} - \mathbf{x}'\|_2^2$, then $\widetilde{R}_{RFTL(\eta,\mathcal{R}),T} \leq D/\eta + \frac{\eta Q_T}{\alpha}\left(1 + \frac{\eta c}{\alpha}\right)$, and if we set $\eta = \sqrt{D\alpha/(2Q)}$ for $Q \geq Dc^2/(4\alpha)$ and $\eta = \sqrt[3]{D\alpha^2/(4Qc)}$ otherwise, we obtain $\widetilde{R}_{RFTL(\eta,\mathcal{R}),T} \leq 3.2 \cdot \sqrt{QD/\alpha}$ in the former case and $\widetilde{R}_{RFTL(\eta,\mathcal{R}),T} \leq 2.4 \cdot \sqrt[3]{QcD^2/\alpha^2}$ in the latter.

The proof of the last part of the second claim, which is purely technical, is given in the appendix.

The above theorem applies in particular to Hedge and OGD: Hedge corresponds to RFTL with $\mathcal{R}(\mathbf{x}) = \sum_{i=1}^{N} x_i \ln x_i$ defined on $\mathcal{K} = \Delta_N$, where $D = \ln N$ and $\alpha = 1$. OGD corresponds to $\mathcal{R}(\mathbf{x}) = (1/2)\|\mathbf{x}\|_2^2$ defined on, say, $\mathcal{K} = B(\mathbf{0}, 1)$, for which $D = 1/2$ and $\alpha = 1$.

We observe that while Theorem 7 applies to the BE setting, it is preferable to obtain bounds in terms that are optimized for a uniform translation of the losses in each round. Such a translation does not affect the regret, but does affect $\sum_t \|\mathbf{l}_t\|_2$ and $\sum_t \|\mathbf{l}_t\|_2^2$, which feature in the regret bounds. To optimize $\sum_i (l_{i,t}-\gamma_t)^2$ we take $\gamma_t = (1/N)\sum_{i=1}^N l_{i,t}$ and obtain $\frac{1}{N}\sum_i (l_{i,t}-\gamma_t)^2 \leq \frac{1}{4}(\max_i\{l_{i,t}\} - \min_i\{l_{i,t}\})^2$ by Popoviciu's inequality (Lemma 21 in the appendix). Thus, we may restate Theorem 7, replacing $\sum_{t=1}^T \|\mathbf{l}_t\|_2$ with $(\sqrt{N}/2)\sum_{t=1}^T (\max_i\{l_{i,t}\} - \min_i\{l_{i,t}\})$ and $Q_T = \sum_{t=1}^T \|\mathbf{l}_t\|_2^2$ with $q_T N/4$, and assuming known bounds on these new quantities instead of Q and Λ.

5. Mixed-Order Bounds for the Best Expert Setting

This section presents bounds on the expected augmented regret for the BE setting with any SC. These bounds combine first and second order terms and are based on an adaptation of the Shrinking Dartboard scheme. The SD algorithm modifies Hedge in a way that upper bounds $\mathbb{E}[K_T]$ while achieving the same regret as Hedge in expectation. We observe that the SD scheme is easily generalized and applied as a meta-algorithm to any BE algorithm A that deterministically assigns only *positive* weights to experts. This results in a modified algorithm denoted $SD(A)$. We next describe this construction and prove its properties.

Let $\mathbf{p}_t$ denote the decisions of A for $t \geq 1$. We recursively define a quantity Z_t by $Z_{t+1} = Z_t \cdot \min_i\{p_{i,t}/p_{i,t+1}\}$, where $Z_1 > 0$ is arbitrary. This definition is valid since $p_{i,t} > 0$ for every i and t. Observe that $p_{i,t}Z_t$ is positive and non-increasing in t for every i, and that Z_t may be computed at time t. The algorithm $SD(A)$ selects a single expert e_t at each time t as follows. It starts with the same probability vector $\mathbf{p}_1$ used by A. At time $t > 1$, $SD(A)$ flips a biased coin F_t with probability of success $f_t = \frac{p_{e_{t-1},t}Z_t}{p_{e_{t-1},t-1}Z_{t-1}}$. If $F_t = 1$, then $SD(A)$ sets $e_t = e_{t-1}$, and otherwise, it uses $\mathbf{p}_t$ to randomly choose e_t. Note that $f_t \in (0, 1]$, making this definition valid.

The next characterization of $SD(A)$ is an adaptation of claims given in Geulen et al. (2010) and is proved similarly. The proofs may be found in the appendix.

Lemma 8 *The algorithm $SD(A)$ satisfies that for every $1 \leq i \leq N$ and $1 \leq t \leq T$, $\mathbb{P}(e_t = i) = p_{i,t}$, and that $\mathbb{E}[K_T] \leq \ln(Z_1/Z_T)$.*

Using these properties we can bound the expected augmented regret of $SD(A)$ as follows:

Lemma 9 *For any switching cost σ upper bounded by B it holds that*

$$\mathbb{E}[\widetilde{R}_{SD(A),T}] \leq R_{A,T} + B\sum_{t=1}^{T-1} \ln \max_i\{p_{i,t+1}/p_{i,t}\} .$$

We now proceed to derive augmented regret bounds that involve second-order characteristics. Such bounds are available for PW and Hedge, and are given for completeness, along with additional required facts, in Appendix B (Theorems 22 and 23). We will denote $PW(\mathbf{p}_0, \eta)$ and $Hed(\mathbf{p}_0, \eta)$ for PW and Hedge, respectively, run with learning rate $\eta > 0$ and initial weights given by a probability vector $\mathbf{p}_0$ with non-zero entries.

Theorem 10 *Set $\mathbf{p}_0 = (1/N, \ldots, 1/N)$, and w.l.o.g., let $\min_i\{l_{i,t}\} = 0$ for every t.*

(i) Assume $\max_i\{l_{i,t}\} \le \mathcal{M}$ *for every* t, $\sum_{t=1}^{T} l_{i,t}^2 \le \mathcal{Q}$ *for every* i, *and* $L_T^* \le \mathcal{L}^*$, *where* $\mathcal{M} > 0$, $\mathcal{Q}$, *and* $\mathcal{L}^*$ *are known. If* $0 < \eta \le 1/(2\mathcal{M})$, *then*

$$\mathbb{E}[\widetilde{R}_{SD(PW(\mathbf{p}_0,\eta)),T}] \le (B + 1/\eta)\ln N + (\eta + B\eta^2)\mathcal{Q} + B\eta\mathcal{L}^* ,$$

and for $\eta = \min\left\{\frac{1}{2\mathcal{M}}, \sqrt{\frac{\ln N}{(1+B/(2\mathcal{M}))\mathcal{Q}+B\mathcal{L}^*}}\right\}$ *it holds that*

$$\mathbb{E}[\widetilde{R}_{SD(PW(\mathbf{p}_0,\eta)),T}] \le B\ln N + \max\left\{4\mathcal{M}\ln N, 2\sqrt{((1 + B/(2\mathcal{M}))\mathcal{Q} + B\mathcal{L}^*)\ln N}\right\} .$$

(ii) Assume $q_T \le q$ *and* $L_T^* \le \mathcal{L}^*$, *where* q *and* $\mathcal{L}^*$ *are known. For every* $\eta > 0$ *it holds that*

$$\mathbb{E}[\widetilde{R}_{SD(Hed(\mathbf{p}_0,\eta)),T}] \le (B + 1/\eta)\ln N + (\eta/8)\cdot q + B\eta\mathcal{L}^* ,$$

and for $\eta = \sqrt{\frac{8\ln N}{q+8B\mathcal{L}^*}}$ *we have* $\mathbb{E}[\widetilde{R}_{SD(Hed(\mathbf{p}_0,\eta)),T}] \le B\ln N + \sqrt{(q/2 + 4B\mathcal{L}^*)\ln N}$.

6. Application to Option Pricing with Transaction Costs

In this section we incorporate proportional transaction costs in the trading model examined in Gofer and Mansour (2012). We apply our augmented regret results to obtain new option price bounds based on a generalization of their analysis, given in detail in Appendix A.

We consider a discrete-time finite-horizon trading model with tradable assets $\mathbf{X}_1, \dots, \mathbf{X}_N$. The price of asset $\mathbf{X}_i$ at time $t \in \{0, 1, \dots, T\}$ is denoted by $X_{i,t}$, and we assume $X_{i,t} > 0$ for every i and t. We assume a zero risk-free interest rate, and that any real quantity of any asset may be bought or sold. Thus, for every $1 \le i \le N$ we may define the *fractional* assets $X_{i,0}^{-1}\mathbf{X}_i$, whose initial value is 1. For every asset $\mathbf{X}_i$ we denote by $r_{i,t}$ the single-period return between $t-1$ and t, so $X_{i,t} = X_{i,t-1}(1 + r_{i,t})$. A realization of the values $r_{i,1}, \dots, r_{i,T}$ is a *price path* for $\mathbf{X}_i$. A realization of the values $r_{i,t}$ for every i and t is simply called *a price path*.

We assume that trading incurs *proportional transaction costs*. Namely, buying or selling an amount worth x of $\mathbf{X}_i$ incurs a cost $c_i x$, where $0 \le c_i < 1$. We will denote $c_M = \max_i\{c_i\}$ and $c_m = \min_i\{c_i\}$. Note that if an asset is simply cash, its rate may reasonably be taken to be zero.[6]

The trading protocol involves a *trading algorithm* A, which is simply an algorithm for the BE setting with N experts. This algorithm starts with wealth U_0 (w.l.o.g. $U_0 = 1$). At every time period $t \ge 1$, A picks a probability vector $\mathbf{p}_t$ and divides its wealth U_{t-1} among the (fractional) assets according to this vector. This operation incurs transaction costs and leaves A with a total wealth V_{t-1}, of which $p_{i,t}V_{t-1}$ is placed with $\mathbf{X}_i$ for $1 \le i \le N$. Following that, the new asset prices $X_{1,t}, \dots, X_{N,t}$ become known, the wealth of the algorithm is updated to $U_t = \sum_{i=1}^{N} V_{t-1}p_{i,t}(1 + r_{i,t}) = V_{t-1}(1 + \sum_{i=1}^{N} p_{i,t}r_{i,t})$, and time period $t+1$ begins. We assume $V_0 = U_0$ (no setup cost) and also $V_T = U_T$, since there is no reason to change the distribution in the last round. Observe that the wealth U_t is divided according to a probability $\widehat{\mathbf{p}}_t$ defined by $\widehat{p}_{i,t} \propto p_{i,t}(1 + r_{i,t})$ for every i.

We assume that transaction costs are funded by the sale of assets, and point out that the procedure for reproportioning wealth among assets is not unique. However, the task of reproportioning wealth with minimal transaction costs is efficiently solvable (Blum and Kalai, 1999), and w.l.o.g. it may be assumed that an optimal procedure is employed by any trading algorithm.

6. Proportional transaction costs (Davis and Norman, 1990) have several different variants in the literature (see Musiela and Rutkowski, 1997). We note that like some, we do not differentiate between buying and selling rates.

We will price an *option* $\mathbf{\Psi}(\mathbf{X}_1, \ldots, \mathbf{X}_N, T)$, defined as a security that pays $\max_{1\le i\le N}\{X_{i,T}\}$ at time T. Specifically, we will upper bound its price at $t = 0$, denoted by $\Psi(\mathbf{X}_1, \ldots, \mathbf{X}_N, T)$. This may be achieved by devising a trading algorithm that *super-replicates* (or *dominates*) a fraction of the option's payoff. Namely, the algorithm guarantees $V_T \ge \beta \max_i\{X_{i,T}\}$ for some $\beta > 0$, for every price path in some set Π of allowed price paths. Investing $1/\beta$ with the algorithm and selling short the option at time 0 allows a guaranteed profit, or *arbitrage*, unless $\Psi(\mathbf{X}_1, \ldots, \mathbf{X}_N, T) \le 1/\beta$. Thus, a price bound is implied, assuming the market is *arbitrage-free*.[7] A randomized algorithm will require the *expected arbitrage-free assumption*, namely, that no *expected* profit may be guaranteed.[8]

The bound derived will be used to price a *European call option*. This security, denoted $\mathbf{C}(K, T)$, pays $\max\{S_T - K, 0\}$ at time T, where $K \ge 0$ is the *strike price* and S_T is the value of some asset $\mathbf{S}$ at time T. The option is "at the money" if $K = S_0$ (w.l.o.g., $S_0 = 1$). We denote $C(K, T)$ for the price of $\mathbf{C}(K, T)$ at time 0, and observe that $C(K, T) = \Psi(\mathbf{S}, \mathbf{K}, T) - K$, where $\mathbf{K}$ is K in cash.

6.1. Bounds on Option Prices

We apply bounds on augmented regret to option pricing based on an interpretation of a trading setup as a BE problem. The single-period losses of the experts are defined simply as $l_{i,t} = -\ln(1 + r_{i,t})$, for every $1 \le i \le N$, $1 \le t \le T$, implying that $L_{i,t} = -\ln(X_{i,t}/X_{i,0})$. In contrast, relating $L_{A,T}$, the cumulative loss of an algorithm A, and its final wealth V_T is more elaborate (see Appendix A, and especially Theorem 19, for the details.) These relations allow one to infer that $V_T \ge \beta \max_i\{X_{i,T}\}$ for every allowed price path, where β is derived from an upper bound on the regret of A. That in turn implies the price bound $\Psi(\mathbf{X}_1, \ldots, \mathbf{X}_N, T) \le 1/\beta$ via the arbitrage-free assumption (see Lemma 14 in Appendix A). This conclusion may be stated more generally:

Theorem 11 *Let A be a trading algorithm, and let $c_M \le 0.2$. It holds that for a known $\alpha_M = \alpha_M(c_m, c_M) \ge 0$, if A guarantees $L_{A,T} - L_{i,T} + \alpha_M \sum_{t=1}^{T} \|\widehat{\mathbf{p}}_t - \mathbf{p}_{t+1}\|_1 + \ln X_{i,0} \le \gamma$ for some γ, for every i and any valid price path, then $\Psi(\mathbf{X}_1, \ldots, \mathbf{X}_N, T) \le \exp(\gamma)$.*

Note that in reality $c_M \ll 0.2$. If $X_{i,0} = 1$ for every i, then γ in the above theorem becomes a bound on $R_{A,T} + \alpha_M \sum_{t=1}^{T} \|\widehat{\mathbf{p}}_t - \mathbf{p}_{t+1}\|_1$. This expression closely resembles regret augmented with a normed SC, but importantly, the role of $\mathbf{p}_t$ is taken by $\widehat{\mathbf{p}}_t$. For an algorithm that holds a single asset at each time t, we have that $\widehat{\mathbf{p}}_t = \mathbf{p}_t$ and the problem is solved. Otherwise, additional ad hoc arguments are necessary. Importantly, if A is probabilistic, then a variant of the above result may be applied. The guarantee of A may hold in expectation, and the result follows by invoking the *expected* arbitrage-free assumption and the concavity of $\ln x$. This variant may be used to derive concrete bounds on $\Psi(\mathbf{X}_1, \ldots, \mathbf{X}_N, T)$ by plugging in the bounds obtained for SD, either with Hedge or with PW. The next theorem will employ Hedge specifically to bound $C(1, T)$.

To the end of this section we will consider two assets: a stock $\mathbf{S}$ with $S_0 = 1$ whose price path is denoted by $(r_1, \ldots, r_T)$, and a unit of cash $\mathbf{1}$. It follows that $Q_T = q_T = \sum_{t=1}^{T} \ln^2(1 + r_t)$. We will also denote $l_t = -\ln(1 + r_t)$, $l_t^+ = \max\{l_t, 0\}$, and $l_t^- = \max\{-l_t, 0\}$, and assume that the transaction cost rate for $\mathbf{S}$ is $c = c_M \le 0.2$.

7. Short selling the option at time 0 and receiving its payoff at time T are assumed to incur no transaction costs. This, and the assumption on the costless setup of the trading portfolio are adopted from Musiela and Rutkowski (1997).

8. By *expected* arbitrage we mean w.r.t. the internal randomizations of a trading algorithm. This is different from the standard term *statistical arbitrage*, which assumes a statistical model of prices.

Theorem 12 *Assume all valid price paths satisfy* $\min\{\sum_{t=1}^{T} l_t^+, \sum_{t=1}^{T} l_t^-\} \leq \lambda^*$ *and* $q_T \leq q$, *where* q *and* λ^* *are given. Then it holds that* $C(1,T) \leq \exp(2c\ln 2 + \sqrt{(q/2 + 8c\lambda^*)\ln 2}) - 1$.

We emphasize that SD requires an oblivious adversary, so Theorem 12 must further assume that market prices are unaffected by the trading algorithm's actions. It is also easy to see how λ^* may grow indefinitely with T, even given $q_T \leq q$, trivializing the bound, unless prices move almost entirely in one direction. Nevertheless, for $c = 0$, the bound becomes $\exp(\sqrt{(q/2)\ln 2}) - 1$, and an optimal $\Theta(\sqrt{q})$ for a small q, matching a result in Gofer (2013) that assumes no transaction costs.

We end this section with an adaptation of the infeasibility result of Section 3 to option pricing. The option $\mathbf{\Psi}(\mathbf{S}, \mathbf{1}, T)$ is trivially dominated by buying and holding the stock and the cash, yielding $\Psi(\mathbf{S}, \mathbf{1}, T) \leq 2$, and thus $C(1,T) \leq 1$. The following theorem shows that given only assumptions on the quadratic variation, this bound may not be improved using our methods. This is expected, since similar results hold even for a stochastic price process (Musiela and Rutkowski, 1997).

Theorem 13 *Let* $c > 0$ *and let* Q *be a known upper bound on the quadratic variation* $\sum_{t=1}^{T} \ln^2(1+r_t)$. *For any* $Q > 0$ *and for any trading algorithm* A *there is a loss sequence with quadratic variation smaller than* Q *for which* $V_T/\max\{1, S_T\}$ *is arbitrarily close to* $1/2$. *As a result, Lemma 14 cannot provide a non-trivial price bound for an "at the money" call option on* $\mathbf{S}$.

7. Conclusion

This work considered regret bounds in the OLO setting with full information, where regret is augmented with SCs. We gave an infeasibility result for obtaining pure second-order bounds with normed SCs given only a bound on the quadratic variation. We also gave augmented regret upper bounds for RFTL and for variants of the Shrinking Dartboard scheme. Those bounds mostly feature an additional constraint on the loss sequence, such as a bound on the total length of losses or the cumulative loss of the best expert. In the absence of SCs, however, they become pure second-order bounds. Both positive and negative results were applied to the problem of option pricing with transaction costs.

Future work It would be interesting to consider upper bounds that involve the second-order quantities examined by Hazan and Kale (2010) and Chiang et al. (2012). In addition, one might consider alternatives to our constraint on the total length of losses. In this context, Lemma 9 in Hazan and Kale (2010) is of interest (although we point out that it requires an additional bound on the time horizon). Another interesting direction would be to consider second-order augmented regret bounds in the more general setting of OCO.

Acknowledgments

The author would like to thank Alon Cohen, Alon Gonen, Shai Shalev-Shwartz, and the anonymous reviewers for their helpful comments. This research was supported by The Israeli Centers of Research Excellence (I-CORE) program (Center No. 4/11) and by the Intel Collaborative Research Institute for Computational Intelligence (ICRI-CI).

References

Jacob Abernethy, Rafael M. Frongillo, and Andre Wibisono. Minimax option pricing meets black-scholes in the limit. In *Proceedings of the 44th Symposium on Theory of Computing*, pages 1029–1040, 2012.

Jacob Abernethy, Peter Bartlett, Rafael Frongillo, and Andre Wibisono. How to hedge an option against an adversary: Black-scholes pricing is minimax optimal. In *Advances in Neural Information Processing Systems*, pages 2346–2354, 2013.

Lachlan L. H. Andrew, Siddharth Barman, Katrina Ligett, Minghong Lin, Adam Meyerson, Alan Roytman, and Adam Wierman. A tale of two metrics: Simultaneous bounds on competitiveness and regret. In *COLT*, pages 741–763, 2013.

Raman Arora, Ofer Dekel, and Ambuj Tewari. Online bandit learning against an adaptive adversary: From regret to policy regret. *arXiv preprint arXiv:1206.6400*, 2012.

Suman K Bera, Anamitra R Choudhury, Syamantak Das, Sambuddha Roy, and Jayram S Thatchachar. Fenchel duals for drifting adversaries. *arXiv preprint arXiv:1309.5904*, 2013.

Avrim Blum and Carl Burch. On-line learning and the metrical task system problem. *Machine Learning*, 39(1):35–58, 2000.

Avrim Blum and Adam Kalai. Universal portfolios with and without transaction costs. *Machine Learning*, 35(3):193–205, 1999.

Allan Borodin, Nathan Linial, and Michael E Saks. An optimal on-line algorithm for metrical task system. *Journal of the ACM (JACM)*, 39(4):745–763, 1992.

Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

Niv Buchbinder, Shahar Chen, Joseph Naor, and Ohad Shamir. Unified algorithms for online learning and competitive analysis. *Journal of Machine Learning Research - Proceedings Track*, 23: 5.1–5.18, 2012.

Nicolò Cesa-Bianchi, Yishay Mansour, and Gilles Stoltz. Improved second-order bounds for prediction with expert advice. *Machine Learning*, 66(2-3):321–352, 2007.

Nicolò Cesa-Bianchi, Ofer Dekel, and Ohad Shamir. Online learning with switching costs and other adaptive adversaries. *arXiv preprint arXiv:1302.4387*, 2013.

Chao-Kai Chiang, Tianbao Yang, Chia-Jung Lee, Mehrdad Mahdavi, Chi-Jen Lu, Rong Jin, and Shenghuo Zhu. Online optimization with gradual variations. *Journal of Machine Learning Research - Proceedings Track*, 23:6.1–6.20, 2012.

T. M. Cover. Universal portfolios. *Mathematical Finance*, 1(1):1–29, 1991.

Mark HA Davis and Andrew R Norman. Portfolio selection with transaction costs. *Mathematics of Operations Research*, 15(4):676–713, 1990.

A Philip Dawid, Steven de Rooij, Peter Grunwald, Wouter M Koolen, Glenn Shafer, Alexander Shen, Nikolai Vereshchagin, and Vladimir Vovk. Probability-free pricing of adjusted american lookbacks. *arXiv preprint arXiv:1108.4113*, 2011a.

A. Philip Dawid, Steven de Rooij, Glenn Shafer, Alexander Shen, Nikolai Vereshchagin, and Vladimir Vovk. Insuring against loss of evidence in game-theoretic probability. *Statistics and Probability Letters*, 81(1):157–162, 2011b.

Ofer Dekel, Jian Ding, Tomer Koren, and Yuval Peres. Bandits with switching costs: T^{2/3} regret. *arXiv preprint arXiv:1310.2997*, 2013.

Peter DeMarzo, Ilan Kremer, and Yishay Mansour. Online trading algorithms and robust option pricing. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing*, pages 477–486, 2006.

Luc Devroye, Gábor Lugosi, and Gergely Neu. Prediction by random-walk perturbation. *arXiv preprint arXiv:1302.5797*, 2013.

Yoav Freund and Robert E Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55(1):119–139, 1997.

Sascha Geulen, Berthold Vöcking, and Melanie Winkler. Regret minimization for online buffering problems using the weighted majority algorithm. In *COLT*, pages 132–143, 2010.

Eyal Gofer. *Machine Learning Algorithms with Applications in Finance*. PhD thesis, Tel Aviv University, 2013. Submitted.

Eyal Gofer and Yishay Mansour. Regret minimization algorithms for pricing lookback options. In *Algorithmic Learning Theory*, pages 234–248, 2011a.

Eyal Gofer and Yishay Mansour. Pricing exotic derivatives using regret minimization. In *Algorithmic Game Theory*, pages 266–277, 2011b.

Eyal Gofer and Yishay Mansour. Lower bounds on individual sequence regret. In *Algorithmic Learning Theory*, pages 275–289, 2012.

András György and Gergely Neu. Near-optimal rates for limited-delay universal lossy source coding. In *Proceedings of the IEEE International Symposium on Information Theory*, pages 2218–2222, 2011.

James Hannan. Approximation to Bayes risk in repeated play. In M. Dresher, A. Tucker, and P. Wolfe, editors, *Contributions to the Theory of Games*, volume 3, pages 97–139. Princeton University Press, 1957.

Elad Hazan. The convex optimization approach to regret minimization. In Suvrit Sra, Sebastian Nowozin, and Stephen J. Wright, editors, *Optimization for Machine Learning*. MIT Press, 2011.

Elad Hazan and Satyen Kale. Extracting certainty from uncertainty: regret bounded by variation in costs. *Machine Learning*, 80(2-3):165–188, 2010.

Elad Hazan, Amit Agarwal, and Satyen Kale. Logarithmic regret algorithms for online convex optimization. *Machine Learning*, 69(2-3):169–192, 2007.

Adam Kalai and Santosh Vempala. Efficient algorithms for online decision problems. *Journal of Computer and System Sciences*, 71(3):291–307, 2005.

Wouter M Koolen and Vladimir Vovk. Buy low, sell high. In *Algorithmic Learning Theory*, pages 335–349, 2012.

Bin Li and Steven CH Hoi. On-line portfolio selection with moving average reversion. *arXiv preprint arXiv:1206.4626*, 2012.

N. Littlestone and M.K. Warmuth. The weighted majority algorithm. *Information and Computation*, 108:212–261, 1994.

Neri Merhav, Erik Ordentlich, Gadiel Seroussi, and Marcelo J Weinberger. On sequential strategies for loss functions with memory. *IEEE Transactions on Information Theory*, 48(7):1947–1958, 2002.

Marek Musiela and Marek Rutkowski. *Martingale methods in financial modelling*, volume 36. Springer, 1997.

Yurii Nesterov. *Introductory Lectures on Convex Optimization: A Basic Course*. Kluwer Academic Publishers, 2004.

R. Tyrrell Rockafellar. *Convex Analysis*. Princeton University Press, 1970.

Shai Shalev-Shwartz and Yoram Singer. A primal-dual perspective of online learning algorithms. *Machine Learning*, 69(2-3):115–142, 2007.

Yoram Singer. Switching portfolios. In *Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence*, pages 488–495, 1998.

Lieven Vandenberghe. Optimization methods for large-scale systems. http://www.seas.ucla.edu/~vandenbe/236C/lectures/smoothing.pdf.

Volodimir G. Vovk. Aggregating strategies. In *Proceedings of the 3rd Annual Workshop on Computational Learning Theory*, pages 371–383, 1990.

Martin Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *ICML*, pages 928–936, 2003.

Appendix A. Return, Regret, and Pricing with Transaction Costs

This appendix explains the relations between option pricing, the performance of trading algorithms, and the performance of algorithms for the BE setting. Those relations were developed previously assuming no transaction costs (DeMarzo et al., 2006; Gofer and Mansour, 2011b, 2012). The results given here extend this analysis to a trading model with proportional transaction costs.

Relating Option Pricing to Trading Algorithms

We start by linking the performance of trading algorithms to the pricing of options with the following simple lemma.

Lemma 14 (DeMarzo et al., 2006; Gofer and Mansour, 2011b) *If there exists an algorithm A trading in $\mathbf{X}_1, \ldots, \mathbf{X}_N$ and $\beta > 0$ s.t. for all possible price paths and every $1 \leq i \leq N$, $V_T \geq \beta X_{i,T}$, then*

$$\Psi(\mathbf{X}_1, \ldots, \mathbf{X}_N, T) \leq 1/\beta \ .$$

Proof It holds that $1/\beta$ units of cash invested with the algorithm will always dominate the payoff of $\mathbf{\Psi}(\mathbf{X}_1, \ldots, \mathbf{X}_N, T)$, implying an upper bound on the price of the option by the arbitrage-free assumption. ■

Importantly, if the algorithm is randomized, the condition must hold for the *expectation* of V_T instead of V_T itself, and the *expected* arbitrage-free assumption is invoked.

Relating Trading Performance and Regret

We next link the performance of a trading algorithm A to its performance as a BE algorithm. This result incorporates transaction costs in the analysis given in Gofer and Mansour (2012).

Recall that the single-period losses of the experts are defined as $l_{i,t} = -\ln(1 + r_{i,t})$, for every $1 \leq i \leq N$, $1 \leq t \leq T$. Thus, the cumulative losses of the experts are exactly minus the logarithms of the values of their respective fractional assets, that is,

$$L_{i,t} = \sum_{\tau=1}^{t} l_{i,\tau} = -\sum_{\tau=1}^{t} \ln(X_{i,\tau}/X_{i,\tau-1}) = -\ln(X_{i,t}/X_{i,0}) \ .$$

Such a simple transformation does not hold w.r.t. $L_{A,T}$ and V_T, but a useful link may nevertheless be established between these two quantities. We will require the next lemma.

Lemma 15 (Gofer and Mansour, 2012) *Let $\sum_{i=1}^{N} p_i = 1$, where $0 \leq p_i \leq 1$ for every $1 \leq i \leq N$, and let $z_i \in (-1, \infty)$ for every $1 \leq i \leq N$. Then*

$$\ln\left(1 + \sum_{i=1}^{N} p_i z_i\right) - (1/8)\ln^2\left(\frac{1 + \max_i\{z_i\}}{1 + \min_i\{z_i\}}\right) \leq \sum_{i=1}^{N} p_i \ln(1 + z_i) \ .$$

Before proceeding, we note that in the present context, the relative quadratic variation of $\mathbf{l}_1, \ldots, \mathbf{l}_T$ satisfies

$$q_T = \sum_{t=1}^{T} (\max_i\{l_{i,t}\} - \min_i\{l_{i,t}\})^2 = \sum_{t=1}^{T} \ln^2\left(\frac{1 + \max_i\{r_{i,t}\}}{1 + \min_i\{r_{i,t}\}}\right) \ .$$

We may now establish the following important relation:

Lemma 16 *It holds that*

$$0 \leq L_{A,T} + \sum_{t=1}^{T} \ln(U_t/V_t) + \ln V_T \leq q_T/8 \ ,$$

and as a result,

$$0 \le R_{A,T} + \sum_{t=1}^{T} \ln(U_t/V_t) + \ln \frac{V_T}{\max_i\{X_{i,0}^{-1} X_{i,T}\}} \le q_T/8 \ .$$

Proof We have that

$$\ln V_T - \sum_{t=1}^{T} \ln(V_t/U_t) = \sum_{t=1}^{T} \ln(U_t/V_{t-1}) = \sum_{t=1}^{T} \ln\left(1 + \sum_{i=1}^{N} p_{i,t} r_{i,t}\right) \ .$$

By the concavity of $\ln(1+z)$,

$$\sum_{t=1}^{T} \ln\left(1 + \sum_{i=1}^{N} p_{i,t} r_{i,t}\right) \ge \sum_{t=1}^{T}\sum_{i=1}^{N} p_{i,t} \ln(1 + r_{i,t}) = -\sum_{t=1}^{T}\sum_{i=1}^{N} p_{i,t} l_{i,t} = -L_{A,T} \ ,$$

and thus, $0 \le L_{A,T} + \ln V_T - \sum_{t=1}^{T} \ln(V_t/U_t)$, as required. For the other side, we have by Lemma 15 that

$$\begin{aligned} -L_{A,T} &= \sum_{t=1}^{T}\sum_{i=1}^{N} p_{i,t} \ln(1 + r_{i,t}) \\ &\ge \sum_{t=1}^{T} \left[\ln\left(1 + \sum_{i=1}^{N} p_{i,t} r_{i,t}\right) - (1/8)\ln^2\left(\frac{1+\max_i\{r_{i,t}\}}{1+\min_i\{r_{i,t}\}}\right)\right] \\ &= \ln V_T - \sum_{t=1}^{T} \ln(V_t/U_t) - q_T/8 \ , \end{aligned}$$

as needed. Since $\min_i\{L_{i,T}\} = -\ln \max_i\{X_{i,0}^{-1} X_{i,T}\}$, we have that

$$0 \le L_{A,T} - \min_i\{L_{i,T}\} - \ln \max_i\{X_{i,0}^{-1} X_{i,T}\} + \ln V_T - \sum_{t=1}^{T} \ln(V_t/U_t) \le q_T/8 \ ,$$

or equivalently,

$$0 \le R_{A,T} - \sum_{t=1}^{T} \ln(V_t/U_t) + \ln \frac{V_T}{\max_i\{X_{i,0}^{-1} X_{i,T}\}} \le q_T/8 \ ,$$

completing the proof. ■

Explicitly Bounding Transaction Costs

The relation given in Lemma 16 accounts for transaction costs only implicitly. The expressions $\ln(U_t/V_t)$ hide both the exact procedure for rearranging wealth among assets as well as the transaction cost rates $c_1, \ldots, c_N$ and relevant probabilities and losses. We turn next to derive explicit expressions.

Consider a single wealth reproportioning operation. Suppose that wealth $V > 0$, which is distributed among assets $\mathbf{X}_1, \ldots, \mathbf{X}_N$ according to distribution $\mathbf{p}$, is redivided according to distribution $\mathbf{p}'$. The total remaining wealth after transaction costs is denoted V'. The next two lemmas will show that transaction costs may be bounded from above and below in terms of $\|\mathbf{p} - \mathbf{p}'\|_1$. As mentioned in Section 6, we will assume that wealth reproportioning operations are optimized to minimize transaction costs.

Lemma 17 *It holds that*

$$\frac{\sum_i c_i|p'_i - p_i|}{1 + \sum_i c_i p'_i} \le 1 - \frac{V'}{V} \le \frac{\sum_i c_i|p'_i - p_i|}{1 - \sum_i c_i p'_i} ,$$

and as a result,

$$\frac{c_m}{1 + c_M} \cdot \|\mathbf{p} - \mathbf{p}'\|_1 \le 1 - \frac{V'}{V} \le \frac{c_M}{1 - c_M} \cdot \|\mathbf{p} - \mathbf{p}'\|_1 .$$

For the special case $N = 2$, one may obtain the improved bounds

$$\frac{c_m + c_M}{2(1 + c_M)} \cdot \|\mathbf{p} - \mathbf{p}'\|_1 \le 1 - \frac{V'}{V} \le \frac{c_m + c_M}{2(1 - c_M)} \cdot \|\mathbf{p} - \mathbf{p}'\|_1 .$$

Proof The optimal redistribution algorithm either buys or sells a certain quantity of each asset. Suppose $z_i \in \mathbb{R}$ is the amount of money spent on $\mathbf{X}_i$, where $z_i > 0$ stands for buying more of the asset, and $z_i < 0$ stands for selling. For every i, the new wealth in asset i is $p'_i V' = p_i V + z_i$, and the transaction cost incurred is $c_i|z_i|$. Summing over the assets we have $V' = V + \sum_i z_i$. In addition, the transaction costs account for the difference in total wealths, so $V - V' = \sum_i c_i|z_i|$. We may therefore write

$$V - V' = \sum_i c_i|z_i| = \sum_i c_i|(p'_i - p_i)V - p'_i(V - V')| .$$

We may now derive the result using the triangle inequality. On the one hand,

$$\begin{aligned}\sum_i c_i|(p'_i - p_i)V - p'_i(V - V')| &\le \sum_i c_i(|p'_i - p_i|V + p'_i|V - V'|) \\ &= V \cdot \sum_i c_i|p'_i - p_i| + (V - V') \cdot \sum_i c_i p'_i .\end{aligned}$$

On the other hand,

$$\begin{aligned}\sum_i c_i|(p'_i - p_i)V - p'_i(V - V')| &\ge \sum_i c_i(|p'_i - p_i|V - p'_i|V - V'|) \\ &= V \cdot \sum_i c_i|p'_i - p_i| - (V - V') \cdot \sum_i c_i p'_i .\end{aligned}$$

Together, we get the two-sided inequality

$$V \cdot \sum_i c_i|p'_i - p_i| - (V - V') \cdot \sum_i c_i p'_i \le V - V' \le V \cdot \sum_i c_i|p'_i - p_i| + (V - V') \cdot \sum_i c_i p'_i ,$$

and rearranging, we obtain

$$\frac{V \cdot \sum_i c_i |p'_i - p_i|}{1 + \sum_i c_i p'_i} \leq V - V' \leq \frac{V \cdot \sum_i c_i |p'_i - p_i|}{1 - \sum_i c_i p'_i} ,$$

and the first claim of the lemma follows, with the second claim of the lemma as an immediate result.

For $N = 2$, it holds that $|p'_2 - p_2| = |p'_1 - p_1| = (1/2)\|\mathbf{p} - \mathbf{p}'\|_1$ and the first claim of the lemma therefore gives

$$\frac{(c_1 + c_2)\|\mathbf{p} - \mathbf{p}'\|_1}{2(1 + \sum_i c_i p'_i)} \leq 1 - \frac{V'}{V} \leq \frac{(c_1 + c_2)\|\mathbf{p} - \mathbf{p}'\|_1}{2(1 - \sum_i c_i p'_i)} .$$

Since $\sum_i c_i p'_i \leq c_M$ and $c_1 + c_2 = c_m + c_M$, it follows that

$$\frac{c_m + c_M}{2(1 + c_M)} \cdot \|\mathbf{p} - \mathbf{p}'\|_1 \leq 1 - \frac{V'}{V} \leq \frac{c_m + c_M}{2(1 - c_M)} \cdot \|\mathbf{p} - \mathbf{p}'\|_1 ,$$

concluding the proof. ■

Using the previous lemma, we may derive bounds on $\ln(V'/V)$ that will be more useful for our purposes. We use the fact that for any $x \leq 1/2$ it holds that

$$-x - x^2 \leq \ln(1 - x) \leq -x . \tag{2}$$

By the second part of Lemma 17,

$$1 - \frac{c_M}{1 - c_M} \cdot \|\mathbf{p} - \mathbf{p}'\|_1 \leq \frac{V'}{V} \leq 1 - \frac{c_m}{1 + c_M} \cdot \|\mathbf{p} - \mathbf{p}'\|_1 .$$

Note that both the leftmost and rightmost expressions have the form $1 - x$, and we would like to show that $x \leq 1/2$ in both cases so that Equation 2 could be applied. To achieve that, since $\|\mathbf{p} - \mathbf{p}'\|_1 \leq 2$, it would suffice to require that $\frac{c_M}{1-c_M} \leq \frac{1}{4}$, or equivalently, that $c_M \leq 0.2$. We then have that

$$\ln \frac{V'}{V} \leq \ln \left(1 - \frac{c_m}{1 + c_M} \cdot \|\mathbf{p} - \mathbf{p}'\|_1\right) \leq -\frac{c_m}{1 + c_M} \cdot \|\mathbf{p} - \mathbf{p}'\|_1$$

and also that

$$\begin{aligned}
\ln \frac{V'}{V} &\geq \ln \left(1 - \frac{c_M}{1 - c_M} \cdot \|\mathbf{p} - \mathbf{p}'\|_1\right) \\
&\geq -\frac{c_M}{1 - c_M}\|\mathbf{p} - \mathbf{p}'\|_1 - \frac{c_M^2}{(1 - c_M)^2}\|\mathbf{p} - \mathbf{p}'\|_1^2 \\
&\geq -\frac{c_M}{1 - c_M}\|\mathbf{p} - \mathbf{p}'\|_1 - \frac{2c_M^2}{(1 - c_M)^2}\|\mathbf{p} - \mathbf{p}'\|_1 \\
&= -\frac{c_M(1 + c_M)}{(1 - c_M)^2}\|\mathbf{p} - \mathbf{p}'\|_1 .
\end{aligned}$$

Together, and since $c_M \leq 0.2$, we have

$$-2c_M\|\mathbf{p} - \mathbf{p}'\|_1 \leq \ln \frac{V'}{V} \leq -0.8c_m\|\mathbf{p} - \mathbf{p}'\|_1 .$$

The same considerations may be applied in the case $N = 2$, for which Lemma 17 yields

$$1 - \frac{c_m + c_M}{2(1 - c_M)} \cdot \|\mathbf{p} - \mathbf{p}'\|_1 \leq \frac{V'}{V} \leq 1 - \frac{c_m + c_M}{2(1 + c_M)} \cdot \|\mathbf{p} - \mathbf{p}'\|_1 \ .$$

It is easy to verify that if $c_M \leq 0.2$, we may apply Equation 2 and obtain that

$$\ln \frac{V'}{V} \leq -\frac{c_m + c_M}{2(1 + c_M)} \cdot \|\mathbf{p} - \mathbf{p}'\|_1$$

and also that

$$\begin{aligned} \ln \frac{V'}{V} &\geq -\frac{c_m + c_M}{2(1 - c_M)} \cdot \|\mathbf{p} - \mathbf{p}'\|_1 - \frac{(c_m + c_M)^2}{4(1 - c_M)^2} \cdot \|\mathbf{p} - \mathbf{p}'\|_1^2 \\ &\geq -\frac{c_m + c_M}{2(1 - c_M)} \cdot \|\mathbf{p} - \mathbf{p}'\|_1 - \frac{(c_m + c_M)^2}{2(1 - c_M)^2} \cdot \|\mathbf{p} - \mathbf{p}'\|_1 \\ &= -\frac{(c_m + c_M)(1 + c_m)}{2(1 - c_M)^2} \|\mathbf{p} - \mathbf{p}'\|_1 \ . \end{aligned}$$

We may thus obtain

$$-(c_m + c_M) \cdot \|\mathbf{p} - \mathbf{p}'\|_1 \leq \ln \frac{V'}{V} \leq -0.4(c_m + c_M)\|\mathbf{p} - \mathbf{p}'\|_1 \ ,$$

where we further used the assumption that $c_M \leq 0.2$. These results are summarized in the following lemma.

Lemma 18 *Assuming that $c_M \leq 0.2$, it holds that*

$$\ln \frac{V'}{V} = -\alpha \|\mathbf{p} - \mathbf{p}'\|_1 \ ,$$

where $\alpha \in [0.8c_m, 2c_M]$, and for the special case $N = 2$, $\alpha \in [0.4(c_m + c_M), c_m + c_M]$.

We may now give an explicit relation between loss, return, and transaction costs for a complete trading process.

Theorem 19 *Let A be an algorithm trading in N assets, and let $c_M \leq 0.2$. It holds that*

$$0 \leq L_{A,T} + \ln V_T + \alpha \sum_{t=1}^{T} \|\widehat{\mathbf{p}}_t - \mathbf{p}_{t+1}\|_1 \leq q_T/8 \ ,$$

and

$$0 \leq R_{A,T} + \alpha \sum_{t=1}^{T} \|\widehat{\mathbf{p}}_t - \mathbf{p}_{t+1}\|_1 + \ln \frac{V_T}{\max_i\{X_{i,0}^{-1} X_{i,T}\}} \leq q_T/8 \ ,$$

where $\widehat{\mathbf{p}}_t$ is a probability vector defined by $\widehat{p}_{i,t} \propto p_{i,t} \exp(-l_{i,t})$ for every i and t and $\alpha \in [0.8c_m, 2c_M]$, and for $N = 2$ in particular, $\alpha \in [0.4(c_m + c_M), c_m + c_M]$.

Proof By Lemma 18,

$$\sum_{t=1}^{T} \ln(V_t/U_t) = -\alpha \sum_{t=1}^{T} \|\widehat{\mathbf{p}}_t - \mathbf{p}_{t+1}\|_1 ,$$

where the distribution $\widehat{\mathbf{p}}_t$ is defined by

$$\widehat{p}_{i,t} = \frac{p_{i,t}(1+r_{i,t})}{\sum_{j=1}^{N} p_{j,t}(1+r_{j,t})} = \frac{p_{i,t}\exp(-l_{i,t})}{\sum_{j=1}^{N} p_{j,t}\exp(-l_{j,t})} .$$

The result is now immediate from Lemma 16. ■

Appendix B. Additional Claims

Lemma 20 (See, e.g., Vandenberghe) *Let $f : C \to \mathbb{R}$ be strongly convex with parameter α, where $C \subseteq \mathbb{R}^N$. Then f has at most one minimizer, and for such a minimizer $\mathbf{x}$, it holds that*

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \frac{1}{2}\alpha\|\mathbf{x} - \mathbf{y}\|_2^2$$

for every $\mathbf{y} \in C$.

Lemma 21 (Popoviciu's inequality) *If X is a bounded random variable with values in $[m, M]$, then $Var(X) \leq (M-m)^2/4$, with equality iff $\mathbb{P}(X = M) = \mathbb{P}(X = m) = 1/2$.*

Theorem 22 (Cesa-Bianchi et al., 2007) *Let A stand for $PW(\mathbf{p}_0, \eta)$. Assume that $l_{i,t} \leq \mathcal{M}$ for every $t = 1, \ldots, T$ and $i = 1, \ldots, N$ for some $\mathcal{M} > 0$. Then for any sequence of losses, expert i, $0 < \eta \leq 1/(2\mathcal{M})$, and $T \geq 1$, it holds that*

$$L_{A,T} \leq \frac{1}{\eta}\ln\frac{W_1}{W_{T+1}} \leq L_{i,T} + \frac{1}{\eta}\ln\frac{1}{p_{i,0}} + \eta\sum_{t=1}^{T} l_{i,t}^2 .$$

Theorem 23 *The algorithm $Hed(\mathbf{p}_0, \eta)$ satisfies that for every expert i,*

$$\frac{1}{\eta}\ln\frac{W_1}{W_{T+1}} \leq L_{i,T} + \frac{1}{\eta}\ln\frac{1}{p_{i,0}}$$

and

$$L_{Hed(\mathbf{p}_0,\eta),T} - L_{i,T} \leq \frac{1}{\eta}\ln\frac{1}{p_{i,0}} + \frac{\eta}{8}\cdot q_T .$$

If q is a known upper bound on q_T, then setting $\eta = \sqrt{(8/q)\ln N}$ and $p_{i,0} = 1/N$ for every i implies that $R_{Hed(\mathbf{p}_0,\eta),T} \leq \sqrt{(q/2)\ln N}$.[9]

9. This result from Gofer (2013) improves the constants of a result in the same spirit given in Cesa-Bianchi et al. (2007).

Proof Let A stand for $Hed(\mathbf{p}_0, \eta)$. Hedge may be defined as the gradient of the concave potential

$$\Phi_\eta(\mathbf{L}) = -\frac{1}{\eta} \ln \left(\sum_{j=1}^{N} p_{j,0} e^{-\eta L_j} \right) .$$

It holds that

$$R_{A,T} = \Phi_\eta(\mathbf{L}_T) - \min_j \{L_{j,T}\} - \frac{1}{2} \sum_{t=1}^{T} \mathbf{l}_t^\top \nabla^2 \Phi_\eta(\mathbf{z}_t) \mathbf{l}_t ,$$

where $\mathbf{z}_t \in [\mathbf{L}_{t-1}, \mathbf{L}_t]$ for $t = 1, \ldots, T$ (see, e.g., Theorem 2 in Gofer and Mansour, 2012). Equivalently,

$$L_{A,T} = \Phi_\eta(\mathbf{L}_T) - \frac{1}{2} \sum_{t=1}^{T} \mathbf{l}_t^\top \nabla^2 \Phi_\eta(\mathbf{z}_t) \mathbf{l}_t . \tag{3}$$

By definition of Φ_η it holds for every i that

$$\begin{aligned} \Phi_\eta(\mathbf{L}_T) - L_{i,T} &= -\frac{1}{\eta} \ln \left(\sum_{j=1}^{N} p_{j,0} e^{-\eta L_{j,T}} \right) + \frac{1}{\eta} \ln e^{-\eta L_{i,T}} \\ &= \frac{1}{\eta} \ln \left(\frac{\exp(-\eta L_{i,T})}{\sum_{j=1}^{N} p_{j,0} e^{-\eta L_{j,T}}} \right) = \frac{1}{\eta} \ln \frac{p_{i,T+1}}{p_{i,0}} \\ &\le \frac{1}{\eta} \ln \frac{1}{p_{i,0}} . \end{aligned} \tag{4}$$

Note that since $(1/\eta) \ln(W_1/W_{T+1}) = \Phi_\eta(\mathbf{L}_T)$, this yields the first of the required claims. Combining Equations 3 and 4, we have that for every i,

$$L_{A,T} - L_{i,T} \le \frac{1}{\eta} \ln \frac{1}{p_{i,0}} - \frac{1}{2} \sum_{t=1}^{T} \mathbf{l}_t^\top \nabla^2 \Phi_\eta(\mathbf{z}_t) \mathbf{l}_t . \tag{5}$$

Now we bound the sum on the right hand side. For every t, it holds that $-\mathbf{l}_t^\top \nabla^2 \Phi_\eta(\mathbf{z}_t) \mathbf{l}_t = \eta Var(Y_t)$, where Y_t is some discrete random variable that may attain only the values $l_{1,t}, \ldots, l_{N,t}$ (see, e.g., Lemma 6 in Gofer and Mansour, 2012). Thus, by Popoviciu's inequality (Lemma 21),

$$-\mathbf{l}_t^\top \nabla^2 \Phi(\mathbf{z}_t) \mathbf{l}_t \le \frac{\eta}{4} \cdot (\max_j \{l_{j,t}\} - \min_j \{l_{j,t}\})^2$$

for every t, and Equation 5 yields

$$L_{A,T} - L_{i,T} \le \frac{1}{\eta} \ln \frac{1}{p_{i,0}} + \frac{\eta}{8} \cdot q_T ,$$

as needed. Setting $p_{i,0} = 1/N$ for every i and $\eta = \sqrt{(8/q) \ln N}$ then yields

$$L_{A,T} - L_{i,T} \le \frac{1}{\eta} \ln N + \frac{\eta}{8} \cdot q = \sqrt{(q/2) \ln N} ,$$

completing the proof. ■

Appendix C. Missing Proofs

Proof of Lemma 2: Denote $t_1 < t_2 < \ldots$ for all times of true direction reversal, namely, times $t > 1$ when $d_t = -d_{t-1}$. Now consider a time interval $[\tau, \tau']$ between two reversals, with $\tau = t_i$ and $\tau' = t_{i+1}$. By our construction, $\mathbf{x}_\tau \cdot d_\tau a_\tau \mathbf{v} \geq \mathbf{x}'_\tau \cdot d_\tau a_\tau \mathbf{v}$ and $\mathbf{x}_{\tau'} \cdot d_\tau a_{\tau'} \mathbf{v} < \mathbf{x}'_{\tau'} \cdot d_\tau a_{\tau'} \mathbf{v}$, or equivalently, $\mathbf{x}_\tau \cdot d_\tau \mathbf{v} \geq \mathbf{x}'_\tau \cdot d_\tau \mathbf{v}$ and $\mathbf{x}_{\tau'} \cdot d_\tau \mathbf{v} < \mathbf{x}'_{\tau'} \cdot d_\tau \mathbf{v}$. These two inequalities combine to give

$$(\mathbf{x}_\tau - \mathbf{x}_{\tau'}) \cdot d_\tau \mathbf{v} > (\mathbf{x}'_\tau - \mathbf{x}'_{\tau'}) \cdot d_\tau \mathbf{v} \ . \tag{6}$$

Now,

$$\begin{aligned}(\mathbf{x}'_\tau - \mathbf{x}'_{\tau'}) \cdot d_\tau \mathbf{v} &= (\nabla\Phi_\eta(\mathbf{L}_{\tau-1}) - \nabla\Phi_\eta(\mathbf{L}_{\tau'-1})) \cdot \frac{d_\tau(\mathbf{L}_{\tau'-1} - \mathbf{L}_{\tau-1})}{d_\tau \sum_{t=\tau}^{\tau'-1} a_t} \\ &= (\nabla\Phi_\eta(\mathbf{L}_{\tau-1}) - \nabla\Phi_\eta(\mathbf{L}_{\tau'-1})) \cdot \frac{\mathbf{L}_{\tau'-1} - \mathbf{L}_{\tau-1}}{\sum_{t=\tau}^{\tau'-1} a_t} \ . \end{aligned} \tag{7}$$

By Taylor's expansion we have for any $\mathbf{L}, \mathbf{L}' \in \mathbb{R}^N$ that

$$\Phi_\eta(\mathbf{L}) - \Phi_\eta(\mathbf{L}') = \nabla\Phi_\eta(\mathbf{L}') \cdot (\mathbf{L} - \mathbf{L}') + \frac{1}{2}(\mathbf{L} - \mathbf{L}')^\top \nabla^2\Phi_\eta(\mathbf{z}_1)(\mathbf{L} - \mathbf{L}')$$

and

$$\Phi_\eta(\mathbf{L}') - \Phi_\eta(\mathbf{L}) = \nabla\Phi_\eta(\mathbf{L}) \cdot (\mathbf{L}' - \mathbf{L}) + \frac{1}{2}(\mathbf{L}' - \mathbf{L})^\top \nabla^2\Phi_\eta(\mathbf{z}_2)(\mathbf{L}' - \mathbf{L})$$

where $\mathbf{z}_1, \mathbf{z}_2 \in [\mathbf{L}, \mathbf{L}']$. Summing these two equations and rearranging then yields

$$\begin{aligned}(\nabla\Phi_\eta(\mathbf{L}') - \nabla\Phi_\eta(\mathbf{L})) \cdot (\mathbf{L} - \mathbf{L}') &= -\frac{1}{2}(\mathbf{L}' - \mathbf{L})^\top (\nabla^2\Phi_\eta(\mathbf{z}_1) + \nabla^2\Phi_\eta(\mathbf{z}_2))(\mathbf{L}' - \mathbf{L}) \\ &= -\frac{\eta}{2}(\mathbf{L}' - \mathbf{L})^\top (\nabla^2\Phi(\eta\mathbf{z}_1) + \nabla^2\Phi(\eta\mathbf{z}_2))(\mathbf{L}' - \mathbf{L}) \ .\end{aligned}$$

For $\mathbf{L} = \mathbf{L}_{\tau-1}$ and $\mathbf{L}' = \mathbf{L}_{\tau'-1}$ it holds that $\mathbf{z}_1 = s_1\mathbf{v}$ and $\mathbf{z}_2 = s_2\mathbf{v}$ with $|s_1|, |s_2| \leq \lambda/\eta$, and by admissibility,

$$(\nabla\Phi_\eta(\mathbf{L}_{\tau'-1}) - \nabla\Phi_\eta(\mathbf{L}_{\tau-1})) \cdot (\mathbf{L}_{\tau-1} - \mathbf{L}_{\tau'-1}) \geq \eta\xi_2 \left(\sum_{t=\tau}^{\tau'-1} a_t\right)^2 \ . \tag{8}$$

Therefore, together with Equation 7 we obtain

$$(\mathbf{x}'_\tau - \mathbf{x}'_{\tau'}) \cdot d_\tau \mathbf{v} \geq \eta\xi_2 \sum_{t=\tau}^{\tau'-1} a_t \ , \tag{9}$$

and with Equation 6 we now have

$$(\mathbf{x}_\tau - \mathbf{x}_{\tau'}) \cdot d_\tau \mathbf{v} > \eta\xi_2 \sum_{t=\tau}^{\tau'-1} a_t \ , \tag{10}$$

and since by Hölder's inequality, $(\mathbf{x}_\tau - \mathbf{x}_{\tau'}) \cdot d_\tau \mathbf{v} \leq \|\mathbf{v}\|_\infty \|\mathbf{x}_\tau - \mathbf{x}_{\tau'}\|_1$ we obtain that

$$\sum_{t=\tau}^{\tau'-1} \|\mathbf{x}_{t+1} - \mathbf{x}_t\|_1 \geq \|\mathbf{x}_{\tau'} - \mathbf{x}_\tau\|_1 > \frac{\eta\xi_2}{\|\mathbf{v}\|_\infty} \sum_{t=\tau}^{\tau'-1} a_t \ .$$

It may be verified that all the above arguments hold also if $\tau = 1$ and $\tau' = t_1$. Summing up over intervals including $[1, t_1]$ yields that for every n,

$$\sum_{t=1}^{t_n-1} \|\mathbf{x}_{t+1} - \mathbf{x}_t\|_1 > \frac{\eta \xi_2}{\|\mathbf{v}\|_\infty} \sum_{t=1}^{t_n-1} a_t \ ,$$

and the proof is complete. ■

Proof of Theorem 3: First, by the equivalence of norms, our results so far prove the first claim for any deterministic A given any normed SC. This claim holds also for randomized algorithms, since we never actually used the fact that A is deterministic.

Our construction requires knowledge of the learner's next decision. If A is deterministic, then an oblivious adversary that can simulate its run may clearly construct the sequence in advance.

Suppose then that the learner is randomized, and that for each t, just before A plays $\mathbf{x}_t$ the adversary may know $\mathbf{y}_t = \mathbb{E}_t[\mathbf{x}_t]$, where $\mathbb{E}_t$ is the conditional expectation at time t. The adversary may run the construction against a mock learner A_1 that plays $\mathbf{y}_1, \ldots, \mathbf{y}_T$. Note that the losses are completely determined by the learner's randomizations, which are the only source of randomness in the game between the learner and the adversary.

Note also that even though the length of the game is a random variable, it may be treated as a known constant, due to the nature of the construction. The reason is that for any target high value for the augmented regret, it is possible to find a bound T for the time it may take to reach it inside the interval, or earlier, by leaving the interval. Thus, it is always possible to pad the loss sequence with extra $\mathbf{0}$ values to ensure length T. On such values, the augmented regret of any learner cannot decrease.

By our construction, given some arbitrarily large R_0, the adversary may guarantee

$$R_0 \le \widetilde{R}_{A_1,T} = \sum_{t=1}^{T} \mathbf{y}_t \cdot \mathbf{l}_t - \min_{\mathbf{u} \in \mathcal{K}} \{\mathbf{u} \cdot \mathbf{L}_T\} + \sum_{t=1}^{T-1} \|\mathbf{y}_{t+1} - \mathbf{y}_t\|$$

regardless of the learner's randomizations. Now, for every t it holds that

$$\mathbb{E}[l_{A,t}] = \mathbb{E}[\mathbf{x}_t \cdot \mathbf{l}_t] = \mathbb{E}[\mathbb{E}_t[\mathbf{x}_t \cdot \mathbf{l}_t]] = \mathbb{E}[\mathbb{E}_t[\mathbf{x}_t] \cdot \mathbf{l}_t] = \mathbb{E}[\mathbf{y}_t \cdot \mathbf{l}_t] = \mathbb{E}[l_{A_1,t}] \ ,$$

implying that $\mathbb{E}[L_{A,T}] = \mathbb{E}[L_{A_1,T}]$ and thus, $\mathbb{E}[R_{A,T}] = \mathbb{E}[R_{A_1,T}]$. In addition,

$$\mathbb{E}[\|\mathbf{x}_{t+1} - \mathbf{x}_t\|] = \mathbb{E}[\mathbb{E}_{t+1}[\|\mathbf{x}_{t+1} - \mathbf{x}_t\|]] \ge \mathbb{E}[\|\mathbb{E}_{t+1}[\mathbf{x}_{t+1} - \mathbf{x}_t]\|] = \mathbb{E}[\|\mathbf{y}_{t+1} - \mathbf{x}_t\|] \ ,$$

where we used Jensen's inequality and the fact that all norms are convex. We therefore have that

$$\begin{aligned}\mathbb{E}[\|\mathbf{y}_{t+1} - \mathbf{y}_t\|] &\le \mathbb{E}[\|\mathbf{y}_{t+1} - \mathbf{x}_t\|] + \mathbb{E}[\|\mathbf{x}_t - \mathbf{x}_{t-1}\|] + \mathbb{E}[\|\mathbf{x}_{t-1} - \mathbf{y}_t\|] \\ &\le \mathbb{E}[\|\mathbf{x}_{t+1} - \mathbf{x}_t\|] + \mathbb{E}[\|\mathbf{x}_t - \mathbf{x}_{t-1}\|] + \mathbb{E}[\|\mathbf{x}_t - \mathbf{x}_{t-1}\|]\end{aligned}$$

for every $t \ge 2$. Consequently, denoting $D_0 = \max_{\mathbf{x},\mathbf{x}' \in \mathcal{K}} \{\|\mathbf{x} - \mathbf{x}'\|\}$ we obtain that

$$\sum_{t=1}^{T-1} \mathbb{E}[\|\mathbf{y}_{t+1} - \mathbf{y}_t\|] - D_0 \le \sum_{t=2}^{T-1} \mathbb{E}[\|\mathbf{y}_{t+1} - \mathbf{y}_t\|] \le 3 \sum_{t=1}^{T-1} \mathbb{E}[\|\mathbf{x}_{t+1} - \mathbf{x}_t\|] \ .$$

It follows that

$$\begin{aligned}\mathbb{E}[\widetilde{R}_{A,T}] &= \mathbb{E}\left[R_{A,T} + \sum_{t=1}^{T-1}\|\mathbf{x}_{t+1}-\mathbf{x}_t\|\right] = \mathbb{E}[R_{A_1,T}] + \sum_{t=1}^{T-1}\mathbb{E}[\|\mathbf{x}_{t+1}-\mathbf{x}_t\|] \\ &\geq \frac{1}{3}\left(\mathbb{E}[R_{A_1,T}] - D_0 + \sum_{t=1}^{T-1}\mathbb{E}[\|\mathbf{y}_{t+1}-\mathbf{y}_t\|]\right) = \frac{1}{3}\left(\mathbb{E}[\widetilde{R}_{A_1,T}] - D_0\right) \\ &\geq \frac{1}{3}\left(R_0 - D_0\right) \ ,\end{aligned}$$

where the first inequality uses the fact that our construction ensures that $R_{A_1,T} \geq 0$. Since R_0 is arbitrarily large, the proof is complete. ■

Proof of Corollary 4: It holds that $\Phi(\mathbf{L}) = -\ln((1/N)\sum_{i=1}^{N} e^{-L_i})$ is concave and continuously twice-differentiable with a gradient in Δ_N. In addition, for every $a > 0$ we have

$$\begin{aligned}\rho_1(a) &= \inf_{\delta(\mathbf{L})\geq a}\{\Phi(\mathbf{L}) - \Phi(\mathbf{0}) - \min_{\mathbf{u}\in\mathcal{K}}\{\mathbf{u}\cdot\mathbf{L}\}\} > 0 \\ \rho_2(a) &= \inf_{\delta(\mathbf{L})\leq a, \delta(\mathbf{l})=1}\{-\mathbf{l}^\top\nabla^2\Phi(\mathbf{L})\mathbf{l}\} > 0 \ .\end{aligned}$$

(See Gofer and Mansour, 2012, Subsection 5.2). It is easily verified that for $\mathbf{v} = (1, 0, \ldots, 0)$ and any $\lambda > 0$, $(\mathbf{v}, \Phi, \lambda)$ is admissible since $\xi_1 \geq \rho_1(\lambda)$ and $\xi_2 \geq \rho_2(\lambda)$.

For the case $\mathcal{K} \supseteq B(\mathbf{0}, a)$, it may be shown that $\Phi(\mathbf{L}) = \min_{\mathbf{x}\in\mathcal{K}}\{\mathbf{x}\cdot\mathbf{L} + \frac{1}{2}\|\mathbf{x}\|_2^2\}$ is concave and continuously twice-differentiable with a gradient in $\mathcal{K}$. In addition, we have

$$\begin{aligned}\rho_1(a) &= \inf_{\|\mathbf{L}\|_2\geq a}\{\Phi(\mathbf{L}) - \Phi(\mathbf{0}) - \min_{\mathbf{u}\in\mathcal{K}}\{\mathbf{u}\cdot\mathbf{L}\}\} > 0 \\ \rho_2(a) &= \inf_{\|\mathbf{L}\|_2\leq a, \|\mathbf{l}\|_2=1}\{-\mathbf{l}^\top\nabla^2\Phi(\mathbf{L})\mathbf{l}\} > 0 \ .\end{aligned}$$

(For details, see Gofer and Mansour, 2012, especially Subsection 5.1.) Again, it is easy to verify that for $\mathbf{v} = (1, 0, \ldots, 0)$ and $\lambda = a$, $(\mathbf{v}, \Phi, \lambda)$ is admissible since $\xi_1 \geq \rho_1(\lambda)$ and $\xi_2 \geq \rho_2(\lambda)$. ■

Proof of the Lipschitz continuity part of Theorem 5: Let $\mathbf{L}, \mathbf{L}' \in \mathbb{R}^N$ and denote $\mathbf{x} = \nabla\Phi(\mathbf{L})$ and $\mathbf{x}' = \nabla\Phi(\mathbf{L}')$. We may assume that $\mathbf{x} \neq \mathbf{x}'$, since otherwise the claim is trivial. The function $f : \mathcal{K} \to \mathbb{R}$ defined by $f(\mathbf{u}) = \eta\mathbf{u}\cdot\mathbf{L} + \mathcal{R}(\mathbf{u})$ is strongly convex with parameter α and is minimized by $\mathbf{x}$. Now, By Lemma 20, we have that

$$\eta\mathbf{x}'\cdot\mathbf{L} + \mathcal{R}(\mathbf{x}') \geq \eta\mathbf{x}\cdot\mathbf{L} + \mathcal{R}(\mathbf{x}) + \frac{1}{2}\alpha\|\mathbf{x} - \mathbf{x}'\|_2^2 \ .$$

The same argument also yields that

$$\eta\mathbf{x}\cdot\mathbf{L}' + \mathcal{R}(\mathbf{x}) \geq \eta\mathbf{x}'\cdot\mathbf{L}' + \mathcal{R}(\mathbf{x}') + \frac{1}{2}\alpha\|\mathbf{x}' - \mathbf{x}\|_2^2 \ ,$$

and adding the two equations, we have that

$$\eta\mathbf{x}'\cdot\mathbf{L} + \eta\mathbf{x}\cdot\mathbf{L}' \geq \eta\mathbf{x}\cdot\mathbf{L} + \eta\mathbf{x}'\cdot\mathbf{L}' + \alpha\|\mathbf{x}' - \mathbf{x}\|_2^2$$

or equivalently,

$$(\mathbf{x}' - \mathbf{x})\cdot(\mathbf{L} - \mathbf{L}') \geq (\alpha/\eta)\|\mathbf{x}' - \mathbf{x}\|_2^2 \ .$$

Therefore,

$$\begin{aligned}\|\nabla\Phi(\mathbf{L}') - \nabla\Phi(\mathbf{L})\|_2\|\mathbf{x}' - \mathbf{x}\|_2 &= \|\mathbf{x}' - \mathbf{x}\|_2^2 \\ &\leq (\eta/\alpha)(\mathbf{x}' - \mathbf{x}) \cdot (\mathbf{L} - \mathbf{L}') \\ &\leq (\eta/\alpha)\|\mathbf{x}' - \mathbf{x}\|_2\|\mathbf{L} - \mathbf{L}'\|_2 \ ,\end{aligned}$$

where the last step is by the Cauchy-Schwarz inequality. Dividing both sides by $\|\mathbf{x}' - \mathbf{x}\|_2$ yields the desired result. ■

Proof of Theorem 6: For every $1 \leq t \leq T$ it holds that $\Phi(\mathbf{L}_t) - \Phi(\mathbf{L}_{t-1}) \geq \nabla\Phi(\mathbf{L}_t) \cdot (\mathbf{L}_t - \mathbf{L}_{t-1})$ since Φ is concave (setting $\mathbf{L}_0 = \mathbf{0}$), and thus

$$\Phi(\mathbf{L}_t) - \Phi(\mathbf{L}_{t-1}) - (\nabla\Phi(\mathbf{L}_t) - \nabla\Phi(\mathbf{L}_{t-1})) \cdot (\mathbf{L}_t - \mathbf{L}_{t-1}) \geq \nabla\Phi(\mathbf{L}_{t-1}) \cdot (\mathbf{L}_t - \mathbf{L}_{t-1}) \ .$$

By the Cauchy-Schwarz inequality and the Lipschitz continuity of $\nabla\Phi$, we get

$$\begin{aligned}|(\nabla\Phi(\mathbf{L}_t) - \nabla\Phi(\mathbf{L}_{t-1})) \cdot (\mathbf{L}_t - \mathbf{L}_{t-1})| &\leq \|\nabla\Phi(\mathbf{L}_t) - \nabla\Phi(\mathbf{L}_{t-1})\|_2\|\mathbf{L}_t - \mathbf{L}_{t-1}\|_2 \\ &\leq (\eta/\alpha)\|\mathbf{L}_t - \mathbf{L}_{t-1}\|_2^2 \ ,\end{aligned}$$

and thus,

$$\Phi(\mathbf{L}_t) - \Phi(\mathbf{L}_{t-1}) + (\eta/\alpha)\|\mathbf{L}_t - \mathbf{L}_{t-1}\|_2^2 \geq \nabla\Phi(\mathbf{L}_{t-1}) \cdot (\mathbf{L}_t - \mathbf{L}_{t-1}) \ .$$

Summing up over $1 \leq t \leq T$, we have that

$$\Phi(\mathbf{L}_T) - \Phi(\mathbf{L}_0) + (\eta/\alpha)\sum_{t=1}^{T}\|\mathbf{L}_t - \mathbf{L}_{t-1}\|_2^2 \geq \sum_{t=1}^{T}\nabla\Phi(\mathbf{L}_{t-1}) \cdot (\mathbf{L}_t - \mathbf{L}_{t-1}) = L_{RFTL(\eta,\mathcal{R}),T} \ .$$

Therefore,

$$R_{RFTL(\eta,\mathcal{R}),T} \leq \Phi(\mathbf{L}_T) - \Phi(\mathbf{L}_0) + \eta Q_T/\alpha - \min_{\mathbf{x}\in\mathcal{K}}\{\mathbf{x} \cdot \mathbf{L}_T\} \ . \tag{11}$$

Now, let $\mathbf{x}_0 \in \mathcal{K}$ be a minimizer for $\mathbf{x} \cdot \mathbf{L}_T$, and let $\mathbf{x}_1 = \nabla\Phi(\mathbf{L}_0)$. By Theorem 5

$$\Phi(\mathbf{L}_0) = \mathbf{x}_1 \cdot \mathbf{L}_0 + \mathcal{R}(\mathbf{x}_1)/\eta = \mathcal{R}(\mathbf{x}_1)/\eta$$

and

$$\Phi(\mathbf{L}_T) \leq \mathbf{x}_0 \cdot \mathbf{L}_T + \mathcal{R}(\mathbf{x}_0)/\eta \ ,$$

and therefore

$$\begin{aligned}\Phi(\mathbf{L}_T) - \Phi(\mathbf{L}_0) - \min_{\mathbf{x}\in\mathcal{K}}\{\mathbf{x} \cdot \mathbf{L}_T\} &\leq \mathbf{x}_0 \cdot \mathbf{L}_T + \mathcal{R}(\mathbf{x}_0)/\eta - \mathcal{R}(\mathbf{x}_1)/\eta - \mathbf{x}_0 \cdot \mathbf{L}_T \\ &= \frac{1}{\eta}(\mathcal{R}(\mathbf{x}_0) - \mathcal{R}(\mathbf{x}_1)) \leq \frac{D}{\eta} \ .\end{aligned}$$

Plugging the above inequality into Equation 11, we obtain the first part of the theorem, and the second part follows immediately. ■

Proof of Theorem 7: If $Q \geq Dc^2/(4\alpha)$, then it holds that $\eta c/\alpha = \sqrt{Dc^2/(2\alpha Q)} \leq \sqrt{2}$, and thus

$$\begin{aligned}\widetilde{R}_{RFTL(\eta,\mathcal{R}),T} &\leq D/\eta + \frac{\eta Q}{\alpha}\left(1+\sqrt{2}\right) = \sqrt{2QD/\alpha} + \sqrt{QD/(2\alpha)}\left(1+\sqrt{2}\right) \\ &= \left(\sqrt{2}+1/\sqrt{2}+1\right)\sqrt{QD/\alpha} \\ &\leq 3.2\sqrt{QD/\alpha}\ .\end{aligned}$$

Otherwise, $Q < Dc^2/(4\alpha)$, and it holds that $\eta c/\alpha = \sqrt[3]{Dc^2/(4Q\alpha)} > 1$ and

$$\begin{aligned}\widetilde{R}_{RFTL(\eta,\mathcal{R}),T} &\leq D/\eta + 2\eta^2 cQ/\alpha^2 = \sqrt[3]{4QcD^2/\alpha^2} + \sqrt[3]{QcD^2/(2\alpha^2)} \\ &\leq 2.4\sqrt[3]{QcD^2/\alpha^2}\ .\end{aligned}$$

■

Proof of Lemma 8: The proof of the first claim proceeds by induction on t. For $t = 1$ the assertion is obvious, and we assume it is true for t and prove it for $t + 1$. It holds that

$$\mathbb{P}(e_{t+1} = i, F_{t+1} = 1) = \mathbb{P}(e_t = i, F_{t+1} = 1) = p_{i,t} \cdot \frac{p_{i,t+1}Z_{t+1}}{p_{i,t}Z_t} = \frac{p_{i,t+1}Z_{t+1}}{Z_t}\ ,$$

where the second equality used the induction assumption. In addition,

$$\begin{aligned}\mathbb{P}(e_{t+1} = i, F_{t+1} = 0) &= \mathbb{P}(e_{t+1} = i|F_{t+1} = 0)\mathbb{P}(F_{t+1} = 0) \\ &= p_{i,t+1}\sum_{j=1}^{N}\mathbb{P}(F_{t+1} = 0|e_t = j)\mathbb{P}(e_t = j) \\ &= p_{i,t+1}\sum_{j=1}^{N}\left(1 - \frac{p_{j,t+1}Z_{t+1}}{p_{j,t}Z_t}\right)\cdot p_{j,t} \\ &= p_{i,t+1}\sum_{j=1}^{N}\left(p_{j,t} - \frac{p_{j,t+1}Z_{t+1}}{Z_t}\right) \\ &= p_{i,t+1}\left(1 - \frac{Z_{t+1}}{Z_t}\right)\ ,\end{aligned}$$

where the third equality again used the induction assumption. Thus,

$$\mathbb{P}(e_{t+1} = i) = \frac{p_{i,t+1}Z_{t+1}}{Z_t} + p_{i,t+1}\left(1 - \frac{Z_{t+1}}{Z_t}\right) = p_{i,t+1}\ ,$$

as required.

For the second claim, note first that for every i and t,

$$\mathbb{P}(e_{t+1} \neq e_t|e_t = i) \leq 1 - \frac{p_{i,t+1}Z_{t+1}}{p_{i,t}Z_t}\ .$$

Denoting $\alpha_t = \mathbb{P}(e_{t+1} \neq e_t)$, we have by the first part that

$$\alpha_t = \sum_{i=1}^{N}\mathbb{P}(e_{t+1} \neq e_t|e_t = i)\mathbb{P}(e_t = i) \leq \sum_{i=1}^{N}\left(p_{i,t} - \frac{p_{i,t+1}Z_{t+1}}{Z_t}\right) = 1 - \frac{Z_{t+1}}{Z_t}\ .$$

Since $-\alpha_t \geq \ln(1-\alpha_t) \geq \ln(Z_{t+1}/Z_t)$, it follows that

$$\sum_{t=1}^{T-1} \alpha_t \leq -\sum_{t=1}^{T-1} \ln \frac{Z_{t+1}}{Z_t} = -\ln \frac{Z_T}{Z_1} \ .$$

Finally,

$$\mathbb{E}[K_T] = \mathbb{E}\left[\sum_{t=1}^{T-1} 1_{\{e_{t+1} \neq e_t\}}\right] = \sum_{t=1}^{T-1} \alpha_t \leq \ln \frac{Z_1}{Z_T} \ ,$$

concluding the proof. ■

Proof of Lemma 9: First, note that the expected standard regret of $SD(A)$ is identical to that of A since

$$\mathbb{E}[L_{SD(A),T}] = \mathbb{E}\left[\sum_{t=1}^{T}\sum_{i=1}^{N} \mathbb{P}(e_t = i) l_{i,t}\right] = \sum_{t=1}^{T}\sum_{i=1}^{N} p_{i,t} l_{i,t} = L_{A,T} \ ,$$

and consequently $\mathbb{E}[R_{SD(A),T}] = R_{A,T}$. Next, we have

$$\mathbb{E}[K_T] \leq \ln(Z_1/Z_T) = \sum_{t=1}^{T-1} \ln(Z_t/Z_{t+1}) = \sum_{t=1}^{T-1} \ln \max_i \{p_{i,t+1}/p_{i,t}\} \ .$$

The result now follows because

$$\mathbb{E}[\widetilde{R}_{SD(A),T}] \leq \mathbb{E}[R_{SD(A),T}] + B\mathbb{E}[K_T] \ .$$

■

Proof of Theorem 10: Note first that assuming $\min_i\{l_{i,t}\} = 0$ for every t is w.l.o.g. since we may bound the augmented regret given the translated losses $l_{i,t} - \min_j\{l_{j,t}\}$ instead of the original ones. The implication both for PW and Hedge is that for every t, $w_{i,t} \geq w_{i,t+1}$ for every i with at least one index j for which $w_{j,t} = w_{j,t+1}$. Therefore,

$$\max_i \{p_{i,t+1}/p_{i,t}\} = (W_t/W_{t+1}) \max_i \{w_{i,t+1}/w_{i,t}\} = W_t/W_{t+1}$$

for every t, and consequently,

$$\sum_{t=1}^{T-1} \ln \max_i \{p_{i,t+1}/p_{i,t}\} = \sum_{t=1}^{T-1} \ln(W_t/W_{t+1}) = \ln(W_1/W_T) \leq \ln(W_1/W_{T+1}) \ .$$

Thus, by Lemma 9

$$\mathbb{E}[\widetilde{R}_{SD(A),T}] \leq R_{A,T} + B\ln(W_1/W_{T+1}) \ , \tag{12}$$

where A stands for either algorithm.

(i) Let A stand for $PW(\mathbf{p}_0, \eta)$. It follows from Theorem 22 that for every expert i,

$$\begin{aligned} L_{A,T} - L_{i,T} + B\ln(W_1/W_{T+1}) &\leq \frac{1}{\eta}\ln\frac{1}{p_{i,0}} + \eta\sum_{t=1}^{T} l_{i,t}^2 + B\left(\eta L_{i,T} + \ln\frac{1}{p_{i,0}} + \eta^2 \sum_{t=1}^{T} l_{i,t}^2\right) \\ &= (B + 1/\eta)\ln\frac{1}{p_{i,0}} + (\eta + B\eta^2)\sum_{t=1}^{T} l_{i,t}^2 + B\eta L_{i,T} \end{aligned} \tag{13}$$

Setting $i = m(T)$ we obtain in conjunction with Equation 12 that

$$\mathbb{E}[\widetilde{R}_{SD(A),T}] \leq (B + 1/\eta)\ln N + (\eta + B\eta^2)\mathcal{Q} + B\eta\mathcal{L}^* \\ \leq B\ln N + (1/\eta)\ln N + \eta((1 + B/(2\mathcal{M}))\mathcal{Q} + B\mathcal{L}^*)\ .$$

This bound is minimized by $\eta = \min\left\{\frac{1}{2\mathcal{M}}, \sqrt{\frac{\ln N}{(1+B/(2\mathcal{M}))\mathcal{Q}+B\mathcal{L}^*}}\right\}$, yielding that

$$\mathbb{E}[\widetilde{R}_{SD(A),T}] \leq B\ln N + \max\left\{4\mathcal{M}\ln N, 2\sqrt{((1 + B/(2\mathcal{M}))\mathcal{Q} + B\mathcal{L}^*)\ln N}\right\}\ .$$

(ii) Let A stand for $Hed(\mathbf{p}_0, \eta)$. It follows from Theorem 23 that for every expert i,

$$L_{A,T} - L_{i,T} + B\ln(W_1/W_{T+1}) \leq \frac{1}{\eta}\ln\frac{1}{p_{i,0}} + \frac{\eta}{8}\cdot q_T + B\left(\eta L_{i,T} + \ln\frac{1}{p_{i,0}}\right) \\ = (B + 1/\eta)\ln\frac{1}{p_{i,0}} + \frac{\eta}{8}\cdot q_T + B\eta L_{i,T} \tag{14}$$

For $i = m(T)$ Equations 12 and 14 yield

$$\mathbb{E}[\widetilde{R}_{SD(A),T}] \leq (B + 1/\eta)\ln N + \frac{\eta}{8}\cdot q + B\eta\mathcal{L}^*\ ,$$

immediately implying the bound for the specific value of η. ■

Proof of Theorem 11: Since $L_{i,T} = -\ln(X_{i,T}/X_{i,0})$, we have by Theorem 19 that

$$0 \leq L_{A,T} - L_{i,T} - \ln(X_{i,T}/X_{i,0}) + \alpha_M\sum_{t=1}^{T}\|\widehat{\mathbf{p}}_t - \mathbf{p}_{t+1}\|_1 + \ln V_T\ ,$$

for a suitable α_M, and therefore

$$-\gamma \leq -(L_{A,T} - L_{i,T} + \ln X_{i,0} + \alpha_M\sum_{t=1}^{T}\|\widehat{\mathbf{p}}_t - \mathbf{p}_{t+1}\|_1) \leq \ln V_T - \ln X_{i,T}\ .$$

Thus, we have that $V_T/X_{i,T} \geq e^{-\gamma}$ for every i, and Lemma 14 yields the result. ■

Proof of Theorem 12: Observe that $SD(A)$ places all weight on a single asset, and therefore $\widehat{\mathbf{p}}_t = \mathbf{p}_t$. As a result, assuming the SC $\sigma(\mathbf{x}, \mathbf{x}') = \alpha_M\|\mathbf{x} - \mathbf{x}'\|_1$, one has for every i that

$$\widetilde{R}_{SD(A),T} + \ln\max_j\{X_{j,0}\} \geq L_{SD(A),T} + \alpha_M\sum_{t=1}^{T}\|\widehat{\mathbf{p}}_t - \mathbf{p}_{t+1}\|_1 - L_{i,T} + \ln X_{i,0} \\ \geq -(\ln V_T - \ln X_{i,T})\ ,$$

where the last inequality used Theorem 19 (which also provides a proper value for α_M). Thus, an upper bound U on $\mathbb{E}[\widetilde{R}_{SD(A),T}]$ implies that

$$U + \ln\max_j\{X_{j,0}\} \geq -\mathbb{E}[\ln V_T - \ln X_{i,T}] = \ln X_{i,T} + \mathbb{E}[-\ln V_T] \geq \ln X_{i,T} - \ln\mathbb{E}[V_T]\ ,$$

where the last inequality is by Jensen's inequality. Therefore,

$$\frac{\mathbb{E}[V_T]}{\max_j\{X_{j,T}\}} \geq \exp(-U - \ln \max_j\{X_{j,0}\}) ,$$

and by the expected version of Lemma 14 we have $\Psi(\mathbf{X}_1, \ldots, \mathbf{X}_N, T) \leq \max_j\{X_{j,0}\}\exp(U)$. In particular, a concrete bound for $C(1,T) = \Psi(\mathbf{S}, \mathbf{1}, T) - 1$ may be derived using Theorem 10. We will use the simpler bound for Hedge, and the valid bound $B = 2\alpha_M$. By Theorem 19 we have $\alpha_M = c$, and therefore,

$$\begin{aligned} C(1,T) + 1 &\leq \exp\left(B \ln 2 + \sqrt{(q/2 + 4B\mathcal{L}^*) \ln 2}\right) \\ &= \exp\left(2c \ln 2 + \sqrt{(q/2 + 8c\mathcal{L}^*) \ln 2}\right) . \end{aligned}$$

Crucially, however, Theorem 10 requires that the losses be transformed. Namely, if $\mathbf{l}_t = (l_t, 0)$, we must subtract $\min\{l_t, 0\}$ from both entries for every t. This transformation does not affect q_T or q, but means that $\mathcal{L}^*$ is actually an upper bound on $\min\{\sum_{t=1}^T l_t^+, \sum_{t=1}^T l_t^-\}$, and it is therefore replaced with λ^*. ■

Proof of Theorem 13: We employ the same construction used to prove the infeasibility result of Section 3, in the specific context of the BE setting with $N = 2$. The first expert will be the stock and the second will be the cash. Before proceeding, we note that the construction works regardless of any randomization on the side of the algorithm. For every time t denote $\mathbf{p}_t$ for the decision of A and $\boldsymbol{\pi}_t$ for the decision of A'. We will also denote $\widehat{\mathbf{p}}_t$ and $\widehat{\boldsymbol{\pi}}_t$ for the probabilities that satisfy $\widehat{p}_{i,t} \propto p_{i,t}\exp(-l_{i,t})$ and $\widehat{\pi}_{i,t} \propto \pi_{i,t}\exp(-l_{i,t})$, respectively, for every i and t. Note that the construction involves arbitrarily small quadratic variation, and is thus appropriate for any bound Q. Since proving the claim for smaller transaction cost rates is harder, we may assume w.l.o.g. that $c = c_M \leq 0.2$. Since $c_M > 0$, Theorem 19 implies that for $\alpha_m = 0.4c_M > 0$,

$$\ln \frac{V_T}{\max\{1, S_T\}} \leq q_T/8 - R_{A,T} - \alpha_m \sum_{t=1}^T \|\widehat{\mathbf{p}}_t - \mathbf{p}_{t+1}\|_1 ,$$

and since q_T is arbitrarily small it suffices to show that for an arbitrarily small $\epsilon > 0$

$$-R_{A,T} - \alpha_m \sum_{t=1}^T \|\widehat{\mathbf{p}}_t - \mathbf{p}_{t+1}\|_1 \ \leq \ln(1/2) + \epsilon ,$$

or equivalently, that

$$R_{A,T} + \alpha_m \sum_{t=1}^T \|\widehat{\mathbf{p}}_t - \mathbf{p}_{t+1}\|_1 \ \geq \ln 2 - \epsilon . \tag{15}$$

As already argued, we may obtain an admissible triplet $(\mathbf{v}, \Phi, \lambda)$ by setting $\mathbf{v} = (1, 0)$, $\Phi(\mathbf{L}) = -\ln((1/N)\sum_{i=1}^N e^{-L_i})$, and choosing some $\lambda > 0$, which may be arbitrarily large. The learning rate will be set as $\eta = 1 + \delta$ for some $\delta > 0$, which may be arbitrarily small. Thus, the reference algorithm A' is $Hed(\mathbf{p}_0, \eta)$, where $\mathbf{p}_0$ is uniform.

As before, the cumulative loss either leaves $[-(\lambda/\eta)\mathbf{v}, (\lambda/\eta)\mathbf{v}]$ for some T or remains inside it. For the first case, Lemma 1 guarantees that $R_{A,T} \geq \xi_1/\eta$, and it holds that $\xi_1 \geq \rho_1(\lambda) =$

$\ln \frac{N}{N-1+\exp(-\lambda)}$ (see Gofer and Mansour, 2012, Lemma 7). Thus, by choosing a large enough λ and a small enough δ we may obtain that $R_{A,T} \geq \xi_1/\eta \geq \ln 2 - \epsilon$ for an arbitrarily small $\epsilon > 0$, satisfying Equation 15, as required.

It now remains to show that Equation 15 is satisfied if $\mathbf{L}_t$ never leaves the interval. To do that, it suffices to show that $R_{A,T} - R_{A',T} + \alpha_m \sum_{t=1}^{T} \|\widehat{\mathbf{p}}_t - \mathbf{p}_{t+1}\|_1$ grows arbitrarily large, since A' has non-negative regret. Now, since $R_{A,T} - R_{A',T} = \sum_{t=1}^{T} \left(l_{A,t} - l_{A',t}\right)$, and $\|\widehat{\mathbf{p}}_t - \mathbf{p}_{t+1}\|_1 = 2|\widehat{p}_{1,t} - p_{1,t+1}|$ for every t, this expression equals $\sum_{t=1}^{T} \left(l_{A,t} - l_{A',t} + 2\alpha_m |\widehat{p}_{1,t} - p_{1,t+1}|\right)$. Furthermore, our construction guarantees $l_{A,t} \geq l_{A',t}$ for every t, so it would be sufficient even to show that

$$\lim_{T\to\infty} \sum_{t\in\mathcal{T}\cap[1..T]} \left(l_{A,t} - l_{A',t} + \kappa|\widehat{p}_{1,t} - p_{1,t+1}|\right) = \infty \ , \tag{16}$$

where $\mathcal{T}$ is some set of times and $\kappa = 2\alpha_m > 0$. We now proceed to show that such a set exists.

In what follows we will refer to $\mathbf{p}_t = (p_t, 1 - p_t)$ rather than $\mathbf{p}_t = (p_{1,t}, p_{2,t})$ for short, and similarly for all other two-dimensional probability vectors. As before, denote $t_1 < t_2 < \dots$ for times of true direction reversals, namely, times $t > 1$ when $d_t = -d_{t-1}$. Consider a single time interval $[\tau_1, \tau_2]$ between two reversals, with $\tau_1 = t_i$ and $\tau_2 = t_{i+1}$, where $d_{\tau_1} = -1$. We assume that i is large enough s.t. $\kappa > a_t$ for every $t \geq t_i$. Recall that $\mathbf{l}_t = d_t a_t \mathbf{v} = (d_t a_t, 0)$ for every t, so $l_{A,t} = p_t d_t a_t$ and $l_{A',t} = \pi_t d_t a_t$. Thus, the fact that $l_{A,t} \geq l_{A',t}$ necessitates that $p_t \leq \pi_t$ for $t \in [\tau_1, \tau_2)$ and $p_{\tau_2} > \pi_{\tau_2}$. We will prove that for any probability values $p_{\tau_1}, \dots, p_{\tau_2}$ that satisfy these conditions, it holds that

$$\sum_{t=\tau_1}^{\tau_2-1} \left(l_{A,t} - l_{A',t} + \kappa|\widehat{p}_t - p_{t+1}|\right) \geq \sum_{t=\tau_1}^{\tau_2-1} \kappa|\widehat{\pi}_t - \pi_{t+1}| \geq b \sum_{t=\tau_1}^{\tau_2-1} a_t \tag{17}$$

for some constant $b > 0$ that does not depend on the specific interval $[\tau_1, \tau_2]$. Now, note that since $\mathbf{L}_t$ remains inside a finite interval, the sums of steps in either direction ($d_t = 1$ or $d_t = -1$) must diverge. Thus, if we take $\mathcal{T} = \{t : d_t = -1\}$, then proving the claim in Equation 17 would satisfy Equation 16 and complete the proof.

We now proceed to prove the claim in Equation 17. Defining $f : [0,1] \times \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ by

$$f(x, y, z) = \frac{xe^{-yz}}{xe^{-yz} + 1 - x} = \frac{x}{x + (1-x)e^{yz}} \ ,$$

we have for every t that $\pi_{t+1} = f(\pi_t, \eta, d_t a_t)$, $\widehat{\pi}_t = f(\pi_t, 1, d_t a_t)$ and $\widehat{p}_t = f(p_t, 1, d_t a_t)$. Now, for any fixed values $y' > y$,

$$\begin{aligned} f(x, y', z) - f(x, y, z) &= \frac{x}{x + (1-x)e^{y'z}} - \frac{x}{x + (1-x)e^{yz}} \\ &= \frac{x(1-x)(e^{yz} - e^{y'z})}{(x + (1-x)e^{yz})(x + (1-x)e^{y'z})} \\ &= \frac{x(1-x)(y'-y)\exp(y''z)}{(x + (1-x)e^{yz})(x + (1-x)e^{y'z})} \cdot (-z) \ , \end{aligned} \tag{18}$$

for some $y'' \in [y, y']$. Clearly, given some $b_1, b_2 > 0$, then for any $x \in (b_1, 1-b_1)$ and $|z| < b_2$, there is some $b_3 > 0$ s.t. $|f(x, y', z) - f(x, y, z)| \geq b_3|z|$. Thus,

$$\sum_{t=\tau_1}^{\tau_2-1} |\widehat{\pi}_t - \pi_{t+1}| = \sum_{t=\tau_1}^{\tau_2-1} |f(\pi_t, 1, d_t a_t) - f(\pi_t, 1+\delta, d_t a_t)| \geq \sum_{t=\tau_1}^{\tau_2-1} b_4 a_t$$

for some $b_4 > 0$, since the values a_t are bounded and the probabilities π_t given by A' (namely, Hedge) are bounded away from 0 and 1 inside the interval $[-(\lambda/\eta)\mathbf{v}, (\lambda/\eta)\mathbf{v}]$. This proves the right inequality of (17), and we now turn to the left inequality. Note first that for every $t \in [\tau_1, \tau_2)$,

$$\begin{aligned}
\widehat{p}_t - \pi_{t+1} &= \widehat{p}_t - \widehat{\pi}_t + \widehat{\pi}_t - \pi_{t+1} \\
&= f(p_t, 1, -a_t) - f(\pi_t, 1, -a_t) + f(\pi_t, 1, -a_t) - f(\pi_t, \eta, -a_t) \\
&\leq f(\pi_t, 1, -a_t) - f(\pi_t, \eta, -a_t) < 0 \ .
\end{aligned} \tag{19}$$

(The first inequality holds since $p_t \leq \pi_t$ and $f(x, y, z)$ is increasing in x for every fixed y and z; the second inequality follows from Equation 18 and the fact that $\pi_t \in (0, 1)$.) Now, observe that the left inequality of (17) amounts to saying that if p_t were replaced by π_t for every $t \in [\tau_1, \tau_2]$, then the sum

$$\sum_{t=\tau_1}^{\tau_2-1} (p_t l_t - \pi_t l_t + \kappa|\widehat{p}_t - p_{t+1}|) = \sum_{t=\tau_1}^{\tau_2-1} \left(l_{A,t} - l_{A',t} + \kappa|\widehat{p}_t - p_{t+1}|\right)$$

may not increase for any valid values p_t. We will use induction to prove a stronger claim, namely, that the same is true if the change is applied only to a suffix of the indices, $[\tau_2 - t + 1, \tau_2]$. If $t = 1$, then the term $\kappa|\widehat{p}_{\tau_2-1} - p_{\tau_2}|$ is replaced by $\kappa|\widehat{p}_{\tau_2-1} - \pi_{\tau_2}|$. Using (19) and the fact that $p_{\tau_2} > \pi_{\tau_2}$ we have that $\widehat{p}_{\tau_2-1} - p_{\tau_2} < \widehat{p}_{\tau_2-1} - \pi_{\tau_2} < 0$, so the replacement does not increase the sum. the inductive step amounts to showing that the sum is not increased by replacing

$$p_{\tau_2-t} l_{\tau_2-t} + \kappa|\widehat{p}_{\tau_2-t-1} - p_{\tau_2-t}| + \kappa|\widehat{p}_{\tau_2-t} - \pi_{\tau_2-t+1}|$$

with

$$\pi_{\tau_2-t} l_{\tau_2-t} + \kappa|\widehat{p}_{\tau_2-t-1} - \pi_{\tau_2-t}| + \kappa|\widehat{\pi}_{\tau_2-t} - \pi_{\tau_2-t+1}|$$

if $\tau_2 - t > \tau_1$, and

$$p_{\tau_2-t} l_{\tau_2-t} + \kappa|\widehat{p}_{\tau_2-t} - \pi_{\tau_2-t+1}|$$

with

$$\pi_{\tau_2-t} l_{\tau_2-t} + \kappa|\widehat{\pi}_{\tau_2-t} - \pi_{\tau_2-t+1}|$$

if $\tau_2 - t = \tau_1$. It holds that $p_{\tau_2-t} \leq \pi_{\tau_2-t}$, and thus $p_{\tau_2-t} l_{\tau_2-t} \geq \pi_{\tau_2-t} l_{\tau_2-t}$. In addition, $f(x, y, z)$ is increasing in x for any given y and z, and therefore $\widehat{p}_{\tau_2-t} \leq \widehat{\pi}_{\tau_2-t}$. Furthermore, by Equation 18, $\widehat{\pi}_{\tau_2-t} < \pi_{\tau_2-t+1}$, and therefore, $|\widehat{p}_{\tau_2-t} - \pi_{\tau_2-t+1}| \geq |\widehat{\pi}_{\tau_2-t} - \pi_{\tau_2-t+1}|$. Combining these facts yields

$$p_{\tau_2-t} l_{\tau_2-t} + \kappa|\widehat{p}_{\tau_2-t} - \pi_{\tau_2-t+1}| \geq \pi_{\tau_2-t} l_{\tau_2-t} + \kappa|\widehat{\pi}_{\tau_2-t} - \pi_{\tau_2-t+1}| \ ,$$

proving the claim for the case $\tau_2 - t = \tau_1$. If $\tau_2 - t > \tau_1$, then since $p_{\tau_2-t} \leq \pi_{\tau_2-t}$, it suffices to show that the expression

$$p_{\tau_2-t} l_{\tau_2-t} + \kappa|\widehat{p}_{\tau_2-t-1} - p_{\tau_2-t}| + \kappa|\widehat{p}_{\tau_2-t} - \pi_{\tau_2-t+1}|$$

does not increase for $p_{\tau_2-t} \in [0, \pi_{\tau_2-t}]$. Observe that $\widehat{p}_{\tau_2-t} \leq \widehat{\pi}_{\tau_2-t} < \pi_{\tau_2-t+1}$ implies $\kappa|\widehat{p}_{\tau_2-t} - \pi_{\tau_2-t+1}| = -\kappa(\widehat{p}_{\tau_2-t} - \pi_{\tau_2-t+1})$, and therefore we may examine instead the expression

$$p_{\tau_2-t}l_{\tau_2-t} + \kappa|\widehat{p}_{\tau_2-t-1} - p_{\tau_2-t}| - \kappa\widehat{p}_{\tau_2-t} \ .$$

If $p_{\tau_2-t} \in [0, \widehat{p}_{\tau_2-t-1}]$, then we obtain $p_{\tau_2-t}l_{\tau_2-t} + \kappa\widehat{p}_{\tau_2-t-1} - \kappa\widehat{p}_{\tau_2-t} - \kappa p_{\tau_2-t}$, which clearly decreases in p_{τ_2-t}. If $p_{\tau_2-t} \in (\widehat{p}_{\tau_2-t-1}, \pi_{\tau_2-t}]$, then we need consider the expression $p_{\tau_2-t}l_{\tau_2-t} + \kappa p_{\tau_2-t} - \kappa\widehat{p}_{\tau_2-t}$, so it remains to show that $g(x) = xl + \kappa x - \kappa f(x,1,l)$ is non-increasing in x, for a negative l. Since $l < 0$ it holds that

$$\frac{\partial f(x,1,l)}{\partial x} = \frac{e^l}{(x + (1-x)e^l)^2} \geq e^l \geq 1 + l \ ,$$

and therefore

$$g'(x) \leq \kappa + l - \kappa(1 + l) = l(1 - \kappa) < 0 \ ,$$

completing the proof. ■