No-Regret Learning in General Games

by

Maxwell K. Fishelson

B.S., MASSACHUSETTS INSTITUTE OF TECHNOLOGY, 2020

SUBMITTED TO THE DEPARTMENT OF ELECTRICAL ENGINEERING AND COMPUTER SCIENCE

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREES OF

MASTER OF SCIENCE

AT THE

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

FEBRUARY 2023

Authored by:	Maxwell K. Fishelson Department of Electrical Engineering and Computer Science January 23, 2023
Certified by:	Constantinos Daskalakis Professor of Electrical Engineering and Computer Science Thesis Supervisor
Approved by:	Leslie A. Kolodziejski Professor of Electrical Engineering and Computer Science Chair, Department Committee on Graduate Students

No-Regret Learning in General Games

by
Maxwell K. Fishelson

Submitted to the Department of Electrical Engineering and Computer Science on January 23, 2023
in Partial Fulfillment of the Requirements for the Degree of Master of Science

Abstract

This thesis investigates the regret performance of no-regret learning algorithms in the competitive, though not fully-adversarial, environment of games. We establish exponential improvements on previously best-known external and internal regret bounds for these settings.

We show that Optimistic Hedge – a common variant of multiplicative-weights-updates with recency bias – attains $\text{poly}(\log T)$ regret in multi-player general-sum games. In particular, when every player of the game uses Optimistic Hedge to iteratively update her strategy in response to the history of play so far, then after T rounds of interaction, *each player* experiences total regret that is $\text{poly}(\log T)$. Our bound improves, exponentially, the $O(T^{1/2})$ regret attainable by standard no-regret learners in games, the $O(T^{1/4})$ regret attainable by no-regret learners with recency bias [Syr+15], and the $O(T^{1/6})$ bound that was recently shown for Optimistic Hedge in the special case of two-player games [CP20]. A corollary of our bound is that Optimistic Hedge converges to coarse correlated equilibrium in general games at a rate of $\tilde{O}\left(\frac{1}{T}\right)$.

We then extend this result from external regret to internal and swap regret, thereby establishing uncoupled learning dynamics that converge to an approximate correlated equilibrium at the rate of $\widetilde{O}\left(T^{-1}\right)$. This substantially improves over the prior best rate of convergence for correlated equilibria of $O(T^{-3/4})$ due to Chen and Peng (NeurIPS '20), and it is optimal up to polylogarithmic factors in T.

The results presented here originate from my works [DFG21] and [Ana+22].

Thesis supervisor: Constantinos Daskalakis

Title: Professor of Electrical Engineering and Computer Science

1 Introduction

Online learning has a long history that is intimately related to the development of game theory, convex optimization, and machine learning. One of its earliest instantiations can be traced to Brown's proposal Brown [Bro49] of fictitious play as a method to solve two-player zero-sum games. Indeed, as shown by Robinson [Rob51], when the players of (zero-sum) matrix game use fictitious play to iteratively update their actions in response to each other's history of play, the resulting dynamics converge in the following sense: the product of the empirical distributions of strategies for each player converges to the set of Nash equilibria in the game, though the rate of convergence is now known to be exponentially slow [DP14]. Moreover, such convergence to Nash equilibria fails in non-zero-sum games [Sha64].

The slow convergence of fictitious play to Nash equilibria in zero-sum matrix games and non-convergence in general-sum games can be mitigated by appealing to the pioneering works Blackwell [Bla54] and Hannan [Han57] and the ensuing literature on no-regret learning [CL06a]. It is known that if both players of a zero-sum matrix game experience regret that is at most $\varepsilon(T)$, the product of the players' empirical distributions of strategies is an $O(\varepsilon(T)/T)$-approximate Nash equilibrium. More generally, if each player of a general-sum, multi-player game experiences regret that is at most $\varepsilon(T)$, the empirical distribution of joint strategies converges to a coarse correlated equilibrium[1] of the game, at a rate of $O(\varepsilon(T)/T)$. Importantly, a multitude of online learning algorithms, such as the celebrated Hedge and Follow-The-Perturbed-Leader algorithms, guarantee adversarial regret $O(\sqrt{T})$ [CL06a]. Thus, when such algorithms are employed by all players in a game, their $O(\sqrt{T})$ regret implies convergence to coarse correlated equilibria (and Nash equilibria of matrix games) at a rate of $O(1/\sqrt{T})$.

While standard no-regret learners guarantee $O(\sqrt{T})$ regret for each player in a game, the players can do better by employing specialized no-regret learning procedures. Indeed, it was established by Daskalakis, Deckelbaum, and Kim [DDK11] that there exists a somewhat complex no-regret learner based on Nesterov's excessive gap technique [Nes05], which guarantees $O(\log T)$ regret to each player of a two-player zero-sum game. This represents an exponential improvement over the regret guaranteed by standard no-regret learners. More generally, Syrgkanis, Agarwal, Luo, and Schapire [Syr+15] established that if players of a multi-player, general-sum game use any algorithm from the family of Optimistic Mirror Descent (MD) or Optimistic Follow-the-Regularized-Leader (FTRL) algorithms (which are analogoues of the MD and FTRL algorithms, respectively, with recency bias), each player enjoys regret that is $O(T^{1/4})$. This was recently improved by Chen and Peng [CP20] to $O(T^{1/6})$ in the special case of two-player games in which the players use OMWU, a particularly simple representative from both the Optimistic MD and Optimistic FTRL families.

The above results for general-sum games represent significant improvements over the $O(\sqrt{T})$ regret attainable by standard no-regret learners, but are not as dramatic as the logarithmic regret that has been shown attainable by no-regret learners, albeit more complex ones, in 2-player zero-sum games (e.g., Daskalakis, Deckelbaum, and Kim [DDK11]). Indeed, despite extensive work on no-regret learning, understanding the optimal regret that can be guaranteed by no-regret learning algorithms in general-sum games has remained elusive. This question is especially intruiging in light of experiments suggesting that polylogarithmic regret should be attainable [Syr+15; HAM21]. In this paper we settle this question by showing that no-regret learners can guarantee polylogarithmic

[1] In general-sum games, it is typical to focus on proving convergence rates for weaker types of equilibrium than Nash, such as coarse correlated equilibria, since finding Nash equilibria is PPAD-complete [DGP06; CDT09].

Table 1: Overview of prior work on fast rates for learning in games. m denotes the number of players, and n denotes the number of actions per player (assumed to be the same for all players). For OMWU, the adversarial regret bounds in the right-hand column are obtained via a choice of adaptive step-sizes. The $\tilde{O}(\cdot)$ notation hides factors that are polynomial in $\log T$.

Algorithm	Setting	Regret in games	Adversarial regret
Hedge (& many other algs.)	multi-player, general-sum	$O(\sqrt{T \log n})$ [CL06a]	$O(\sqrt{T \log n})$ [CL06a]
Excessive Gap Technique	2-player, 0-sum	$O(\log n(\log T + \log^{3/2} n))$ [DDK11]	$O(\sqrt{T \log n})$ [DDK11]
DS-OptMD, OptDA	2-player, 0-sum	$\log^{O(1)}(n)$ [HAM21]	$\sqrt{T \log^{O(1)}(n)}$ [HAM21]
OMWU	multi-player, general-sum	$O(\log n \cdot \sqrt{m} \cdot T^{1/4})$ [RS13b; Syr+15]	$\tilde{O}(\sqrt{T \log n})$ [RS13b; Syr+15]
OMWU	2-player, general-sum	$O(\log^{5/6} n \cdot T^{1/6})$ [CP20]	$\tilde{O}(\sqrt{T \log n})$
OMWU	multi-player, general-sum	$O(\log n \cdot m \cdot \log^4 T)$ **(Theorem 3.1)**	$\tilde{O}(\sqrt{T \log n})$ **(Corollary D.1)**

regret to each player in general-sum multi-player games. Moreover, this regret is attainable by a particularly simple algorithm – Optimistic Hedge:

Theorem 1.1 (Abbreviated version of Theorem 3.1)**.** *Suppose that m players play a general-sum multi-player game, with a finite set of n strategies per player, over T rounds. Suppose also that each player uses Optimistic Hedge to update her strategy in every round, as a function of the history of play so far. Then each player experiences $O(m \cdot \log n \cdot \log^4 T)$ regret.*

An immediate corollary of Theorem 1.1 is that the empirical distribution of play is a $O\left(\frac{m \log n \log^4 T}{T}\right)$-approximate coarse correlated equilibrium (CCE) of the game. We remark that Theorem 1.1 bounds the total regret experienced by *each player* of the multi-player game, which is the most standard regret objective for no-regret learning in games, and which is essential to achieve convergence to CCE. For the looser objective of the *average* of all players' regrets, [RS13b] established a $O(\log n)$ bound for OMWU in two-player zero-sum games, and [Syr+15] generalized this bound, to $O(m \log n)$ in m-player general-sum games. Note that since some players may experience *negative regret* [HAM21], the average of the players' regrets cannot be used in general to bound the maximum regret experienced by any individual player. Finally, we remark that several results in the literature posit no-regret learning as a model of agents' rational behavior; for instance, [Rou15; ST13; RST17] show that no-regret learners in smooth games enjoy strong Price-of-Anarchy bounds. By showing that *each agent* can obtain very small regret in games by playing OMWU, Theorem 1.1 strengthens the plausability of the common assumption made in this literature that each agent will choose to use such a no-regret algorithm.

1.1 Correlated Equilibria

We also investigate the more refined solution concept of multi-player general-sum games known as *correlated equilibria* (*CE*) [Aum74]. Like CCE, it is known that CE can be computed through uncoupled learning dynamics. The main follow-up question to Theorem 1.1 is:

Are there learning dynamics that, if followed by all players in a multi-player general-sum game, guarantee convergence with rate $\widetilde{O}\left(T^{-1}\right)$ to a correlated equilibrium? (♠)

We answer this question in the affirmative. Unlike in the case of CCE, typical no-external-regret dynamics such as MWU are known not to guarantee convergence to CE. Instead, specialized *no-internal-regret* or *no-swap-regret* algorithms have to be employed to converge to CE [CL06b]. Compared to no-external-regret dynamics, these learning dynamics are considerably more complex in that all known algorithms require the computation of the stationary distribution of a certain Markov chain at every iteration.

This work presents a refined analysis of the no-internal-regret algorithm of Stoltz and Lugosi [SL05], as well as the no-swap-regret algorithm of Blum and Mansour [BM07], both instantiated with Optimistic Multiplicative Weights Update (OMWU). Going forward, we will refer to those learning dynamics as SL-OMWU and BM-OMWU, respectively. Our primary contribution is to show that both of these algorithms exhibit a near-optimal convergence rate of $\widetilde{O}(T^{-1})$, settling Question (♠) in the affirmative. More precisely, for SL-OMWU our main theorem is summarized as follows.

Theorem 1.2. *Consider a general-sum multi-player game with m players, with each player $i \in [\![m]\!]$ having n_i actions. There exists a universal constant $C > 0$ such that, when all players select strategies according to algorithm* SL-OMWU *with step size $\eta = 1/(C \cdot m \log^4 T)$, the internal regret of every player $i \in [\![m]\!]$ is bounded by $O\big(m \log n_i \log^4 T\big)$. As a result, the average product distribution of play is an $O\big((m \log n \log^4 T)/T\big)$-approximate correlated equilibrium.*

This matches, up to constant factors, the rate of convergence for coarse correlated equilibria of Theorem 1.1, and it is optimal, within the no-regret framework,[2] up to polylogarithmic factors [DDK11]. This also substantially improves upon the $O(T^{-3/4})$ rate of convergence for correlated equilibria recently shown by Chen and Peng [CP20]. Moreover, since swap regret on an n-simplex is trivially at most n times larger than internal regret (*e.g.*, see Blum and Mansour [BM07, pp. 1311]), Theorem 1.2 directly gives a bound in terms of swap regret as well, stated as follows.

Corollary 1.3. *If all players select strategies according to algorithm* SL-OMWU*, the swap regret of every player $i \in [\![m]\!]$ is bounded by $O\big(m\, n_i \log n_i \log^4 T\big)$.*

For the algorithm BM-OMWU, we demonstrate the following.

Theorem 1.4. *Consider a general-sum multi-player game with m players, with each player $i \in [\![m]\!]$ having n_i actions. There exists a universal constant $C > 0$ such that, when all players select strategies according to algorithm* BM-OMWU *with step size $\eta = 1/(C \cdot m\, n_i^3 \log^4(T))$, the swap regret of every player $i \in [\![m]\!]$ is bounded by $O(m\, n_i^4 \log n_i \log^4(T))$. As a result, the average product distribution of play is an $O\big((m\, n^4 \log n \log^4 T)/T\big)$-approximate correlated equilibrium.*

[2] Finding a correlated equilibrium can be phrased as a linear programming problem, and thus ϵ-approximate correlated equilibria can be found in time $\mathrm{poly}(m, n, \log(1/\epsilon))$, where $n = \max_{i \in [\![m]\!]}\{n_i\}$, for succinct multi-player games [PR08]. However, the procedure for doing so, *ellipsoid against hope*, cannot be phrased as uncoupled dynamics and is unlikely to be run by players competing in a repeated game.

Finally, we remark that SL-OMWU and BM-OMWU instantiated with the learning rates described in Theorems 1.2 and 1.4 guarantee near-optimal swap regret (in T) when all players use the same dynamics, but might not against general, adversarial losses. To guarantee near-optimal swap regret in both the adversarial and the non-adversarial regime, an adaptive choice of learning rate similar to that in Appendix D can be used (see Corollary 4.5).

1.2 Related Work

Table 1 summarizes the prior works that aim to establish optimal regret bounds for no-regret learners in games. We remark that [CP20] shows that the regret of Hedge is $\Omega(\sqrt{T})$ even in 2-player games where each player has 2 actions, meaning that optimism is necessary to obtain fast rates. The table also includes a recent result of [HAM21] showing that when the players in a 2-player zero-sum game with n actions per player use a variant of OMWU with adaptive step size (a special case of their algorithms DS-OptMD and OptDA), each player has $\log^{O(1)} n$ regret. The techniques of [HAM21] differ substantially from ours: the result in [HAM21] is based on showing that the joint strategies $x^{(t)}$ rapidly converge, pointwise, to a Nash equilibrium $x^\star$. Such a result seems very unlikely to extend to our setting of general-sum games, since finding an approximate Nash equilibrium even in 2-player games is PPAD-complete [CDT09]. We also remark that the earlier work [Kan+18] shows that each player's regret is at most $O(\log T \cdot \log n)$ when they use a certain algorithm based on Optimistic MD in 2-player zero-sum games; their technique is heavily tailored to 2-player zero-sum games, relying on the notion of duality in such a setting.

[Fos+16] shows that one can obtain fast rates in games for a broader class of algorithms (e.g., including Hedge) if one adopts a relaxed (approximate) notion of optimality. [WL18] uses optimism to obtain adaptive regret bounds for bandit problems. Many recent papers (e.g., [DP19; GPD20; Lei+21; HAM21; Wei+21; Azi+21]) have studied the *last-iterate* convergence of algorithms from the Optimistic Mirror Descent family, which includes OMWU. Finally, a long line of papers (e.g., [HM03; Das+10; KLP11; BCM12; PP16; BP18; MPP18; BP19; CP19; Vla+20]) has studied the dynamics of learning algorithms in games. Essentially all of these papers do not use optimism, and many of them show *non-convergence* (e.g., divergence or recurrence) of the iterates of various learning algorithms such as FTRL and Mirror Descent when used in games.

No-internal-regret algorithms that require black-box access to a *single* no-external-regret minimizer are known in the literature [SL05; CL06b]. This is in contrast with the construction of Blum and Mansour [BM07] for the stronger notion of swap regret, which requires n independent no-external-regret minimizers—one per each action of the player. Nevertheless, both classes of algorithms involve computing the stationary distribution of a certain Markov chain at every iteration. The intrinsic complexity associated with the computation of a stationary distribution was arguably the main factor limiting our ability to give accelerated convergence guarantees for either class of algorithms. Indeed, while learning dynamics guaranteeing external regret bounded by $O(T^{1/4})$ have been known for several years [Syr+15], a matching bound for swap regret was only recently shown by Chen and Peng [CP20].

The setting studied in our paper (learning dynamics for correlated equilibrium) is substantially more challenging than the problem of giving accelerated learning dynamics for Nash equilibria in two-player zero-sum games, as well learning dynamics for *smooth games* [Rou15]. Indeed, while in our setting the convergence to the equilibrium is driven by the *maximum* internal (or swap) regret cumulated by the players, in the latter two settings the quality metric is driven by the *sum* of the external regrets. As shown by Syrgkanis, Agarwal, Luo, and Schapire [Syr+15], it is

possible to guarantee a *constant sum* of external regrets under a *broad* class of *predictive* no-regret algorithms which includes optimistic OMD and optimistic FTRL under very general distance-generating functions. This is in contrast with the case of no-regret dynamics for CCE and CE, where it remains an open question to give broad classes of algorithms that can achieve near-optimal convergence.

Finally, we point out that optimistic variants of FTRL such as OMWU have been shown to also converge in the *last-iterate sense*, and the convergence is known to be linear[3] [Das+18; DP18; DP19; Wei+21], but this holds only for restricted classes of games, such as two-player zero-sum games.

2 Preliminaries

2.1 General Preliminaries

Consider a finite *normal-form* game Γ consisting of a set of m players $[\![m]\!] \coloneqq \{1, 2, \dots, m\}$ such that every player i has an *action space* $\mathcal{A}_i \coloneqq [\![n_i]\!]$, for some $n_i \in \mathbb{N}$. The *joint* action space will be represented with $\mathcal{A} \coloneqq \mathcal{A}_1 \times \cdots \times \mathcal{A}_m$. For a given action profile $\boldsymbol{a} = (a_1, \dots, a_m) \in \mathcal{A}$, the *loss function* $\Lambda_i : \mathcal{A} \to [0, 1]$ specifies the loss of player i under the action profile $\boldsymbol{a}$; note that the normalization of the losses comes without any loss of generality. A *mixed strategy* $\boldsymbol{x}_i \in \Delta(\mathcal{A}_i)$ for a player $i \in [\![m]\!]$ is a probability distribution over i's action space $\mathcal{A}_i$, so that the coordinate $\boldsymbol{x}_i[j]$ indicates the probability that player i will select action $j \in [\![n_i]\!]$. A *deterministic strategy* refers to a mixed strategy supported on a single coordinate. Given a *joint* vector of mixed strategies $\boldsymbol{x} = (\boldsymbol{x}_1, \dots, \boldsymbol{x}_m)$, the *expected* loss $\boldsymbol{\ell}_i$ of player i is such that $\boldsymbol{\ell}_i[j] \coloneqq \mathbb{E}_{\boldsymbol{a}_{-i} \sim \boldsymbol{x}_{-i}}[\Lambda_i(j, \boldsymbol{a}_{-i})]$, for all $j \in [\![n_i]\!]$, where we used the notation $\boldsymbol{a}_{-i}$ to denote the vector $\boldsymbol{a}$ excluding the coordinate corresponding to i's action; i.e. $\boldsymbol{a}_{-i} \coloneqq (a_1, \dots, a_{i-1}, a_{i+1}, \dots, a_m)$.

Hindsight Rationality and Regret Minimization. A standard quality metric in the theory of learning in games is *hindsight rationality*. Hindsight rationality encodes the idea that a player has "learned" to play the game when, looking back at the history of play, there is no transformation of their strategies that—applied to the whole history of play—would have led to strictly better utility for that player. This notion is operationalized through *Phi-regret*. Formally, the Φ_i-regret incurred by the sequence of strategies $\boldsymbol{x}_i^{(1)}, \dots, \boldsymbol{x}_i^{(T)} \in \Delta^{n_i}$ selected by player $i \in [\![m]\!]$ is defined as

$$\mathrm{Reg}_{\Phi_i}^T \coloneqq \sum_{t=1}^{T} \langle \boldsymbol{x}_i^{(t)}, \boldsymbol{\ell}_i^{(t)} \rangle - \min_{\phi^* \in \Phi_i} \sum_{t=1}^{T} \langle \phi^*(\boldsymbol{x}_i^{(t)}), \boldsymbol{\ell}_i^{(t)} \rangle. \tag{1}$$

Notable special cases are identified based on the particular choice of transformations Φ_i, as follows.

(i) *External* regret (or simply *regret*) corresponds to the case where Φ_i is the set of all *constant* functions $\Phi_i^{\mathrm{const}} \coloneqq \{\phi : \Delta^{n_i} \to \Delta^{n_i}, \phi(\boldsymbol{x}) = \phi(\boldsymbol{x}') \ \forall \boldsymbol{x}, \boldsymbol{x}' \in \Delta^{n_i}\}$.

(ii) *Internal* regret corresponds to the set of linear transformation $\Phi_i = \Phi_i^{\mathrm{int}}$ that transport probability mass from some action j to some other action k. Formally, Φ_i^{int} is the convex hull $\Phi_i^{\mathrm{int}} \coloneqq \mathrm{co}\{\phi_{j\to k}\}_{j,k \in \mathcal{A}_i, j \neq k}$ of the functions $\phi_{j\to k}$ defined as

$$\phi_{j\to k} : \Delta^{n_i} \to \Delta^{n_i}, \quad \boldsymbol{x} \mapsto \boldsymbol{x} + (\boldsymbol{e}_k - \boldsymbol{e}_j)\, \boldsymbol{x}[j] = \left(\mathbf{I} + (\boldsymbol{e}_k - \boldsymbol{e}_j)\boldsymbol{e}_j^\top\right) \boldsymbol{x} =: \mathbf{E}_{j\to k}\, \boldsymbol{x}. \tag{2}$$

(iii) *Swap* regret corresponds to the set Φ_i^{swap} of all linear transformations $\Delta^{n_i} \ni \boldsymbol{x} \mapsto \mathbf{Q}^\top \boldsymbol{x}$ where $\mathbf{Q}$ is a *row-stochastic matrix*, that is, a non-negative matrix whose rows all sum to 1.

[3] However, the convergence rate of the last-iterate may *not* depend polynomially in the size of the game [Wei+21].

Optimistic Multiplicative Weights Update. A celebrated result in the theory of online learning is the existence of natural learning algorithms that guarantee that the *average regret* asymptotically vanishes as $T \to \infty$, even under *adversarially selected* loss functions [CL06b]. The *optimistic* variant of MWU, called *Optimistic Multiplicative Weights Update (OMWU)*[4] is a particular instantiation of the more general *optimistic FTRL* algorithm, and will be the focus of this work. At time $t = 1$, OMWU outputs the uniform distribution $\boldsymbol{x}^{(1)} = \frac{1}{n}\mathbf{1} \in \Delta^n$. Then, at each time $t \geq 1$, it computes the next iterate $\boldsymbol{x}^{(t+1)}$ according to the update rule

$$\boldsymbol{x}^{(t+1)}[j] = \frac{\exp\{-\eta(2\boldsymbol{\ell}^{(t)}[j] - \boldsymbol{\ell}^{(t-1)}[j])\}}{\sum_{k\in[\![n]\!]} \exp\{-\eta(2\boldsymbol{\ell}^{(t)}[k] - \boldsymbol{\ell}^{(t-1)}[k])\}\, \boldsymbol{x}^{(t)}[k]}\, \boldsymbol{x}^{(t)}[j], \tag{OMWU}$$

where $\eta > 0$ is a *learning rate* parameter and $\boldsymbol{\ell}^{(t)}$ is the feedback loss vector observed at time t (we conventionally let $\boldsymbol{\ell}^{(0)} = \mathbf{0}$).

Connections with Solution Concepts. There exist connections between the different notions of hindsight rationality described above and game-theoretic solution concepts. The two solution concepts of coarse correlated equilibria and correlated equilibria are the focus of this paper. We recall these definitions next [FV97; HM00].

Definition 2.1 (Coarse Correlated Equilibrium)**.** A probability distribution $\boldsymbol{\mu}$ over $\Pi_1 \times \cdots \times \Pi_m$ is said to be *an ϵ-correlated equilibrium*, where $\epsilon > 0$, if for every player $i \in [\![m]\!]$ and every $\phi \in \Phi_i^{\text{const}}$,

$$\mathbb{E}_{(\boldsymbol{\pi}_1,\ldots,\boldsymbol{\pi}_m)\sim\boldsymbol{\mu}}\big[\langle \boldsymbol{\ell}_i, \boldsymbol{\pi}_i\rangle - \langle \boldsymbol{\ell}_i, \phi(\boldsymbol{\pi}_i)\rangle\big] \leq \epsilon. \tag{3}$$

Definition 2.2 (Correlated Equilibrium)**.** A probability distribution $\boldsymbol{\mu}$ over $\Pi_1 \times \cdots \times \Pi_m$ is said to be *an ϵ-correlated equilibrium*, where $\epsilon > 0$, if for every player $i \in [\![m]\!]$ and every $\phi \in \Phi_i^{\text{int}}$,

$$\mathbb{E}_{(\boldsymbol{\pi}_1,\ldots,\boldsymbol{\pi}_m)\sim\boldsymbol{\mu}}\big[\langle \boldsymbol{\ell}_i, \boldsymbol{\pi}_i\rangle - \langle \boldsymbol{\ell}_i, \phi(\boldsymbol{\pi}_i)\rangle\big] \leq \epsilon. \tag{4}$$

A folklore result states that the average product distribution of play of no-external-regret players converges to an approximate *coarse correlated equilibrium* (*CCE*) and the average product distribution of play of no-internal-regret players converges to an approximate *correlated equilibrium* (*CE*). More precisely, the following holds.

Theorem 2.3. *Consider T repetitions of play in a game where every player $i \in [\![m]\!]$ employs an algorithm that produces strategies $\boldsymbol{x}_i^{(1)}, \ldots, \boldsymbol{x}_i^{(T)} \in \Delta^{n_i}$ with external regret Reg_i^T. Then, the average product distribution of play $\bar{\boldsymbol{\mu}} \coloneqq \frac{1}{T}\sum_{t=1}^T \boldsymbol{x}_1^{(t)} \otimes \cdots \otimes \boldsymbol{x}_m^{(t)}$ is an $O(\max_i \mathrm{Reg}_i^T/T)$-CCE.*

Theorem 2.4. *Consider T repetitions of play in a game where every player $i \in [\![m]\!]$ employs an algorithm that produces strategies $\boldsymbol{x}_i^{(1)}, \ldots, \boldsymbol{x}_i^{(T)} \in \Delta^{n_i}$ with internal regret IntReg_i^T. Then, the average product distribution of play $\bar{\boldsymbol{\mu}} \coloneqq \frac{1}{T}\sum_{t=1}^T \boldsymbol{x}_1^{(t)} \otimes \cdots \otimes \boldsymbol{x}_m^{(t)}$ is an $O(\max_i \mathrm{IntReg}_i^T/T)$-CE.*

[4]The OMWU algorithm is sometimes referred to as *Optimistic Hedge* in the literature.

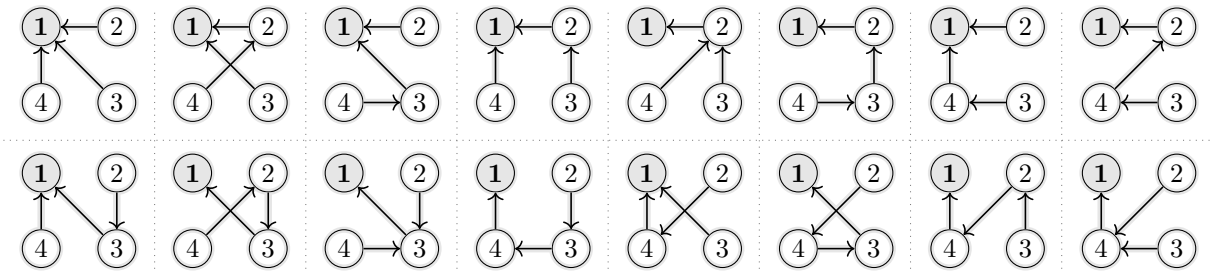

Figure 1: Set $\mathbb{T}_1^4$ of all directed trees rooted at node 1 in a graph with 4 nodes.

Distributions & divergences. For distributions $\boldsymbol{P}, \boldsymbol{Q}$ on a finite domain $[n]$, the *KL divergence* between $\boldsymbol{P}, \boldsymbol{Q}$ is $\mathrm{KL}(\boldsymbol{P}; \boldsymbol{Q}) = \sum_{j=1}^n \boldsymbol{P}(j) \cdot \log\left(\frac{\boldsymbol{P}(j)}{\boldsymbol{Q}(j)}\right)$. The *chi-squared divergence* between $\boldsymbol{P}, \boldsymbol{Q}$ is $\chi^2(\boldsymbol{P}; \boldsymbol{Q}) = \sum_{j=1}^n \boldsymbol{Q}(j) \cdot \left(\frac{\boldsymbol{P}(j)}{\boldsymbol{Q}(j)}\right)^2 - 1 = \sum_{j=1}^n \frac{(\boldsymbol{P}(j) - \boldsymbol{Q}(j))^2}{\boldsymbol{Q}(j)}$. For a distribution $\boldsymbol{P}$ on $[n]$ and a vector $\boldsymbol{v} \in \mathbb{R}^n$, we write $\mathrm{Var}_{\boldsymbol{P}}(\boldsymbol{v}) := \sum_{j=1}^n \boldsymbol{P}(j) \cdot (\boldsymbol{v}(j) - \sum_{k=1}^n \boldsymbol{P}(k)\boldsymbol{v}(k))^2$. Also define $\|\boldsymbol{v}\|_{\boldsymbol{P}} := \sqrt{\sum_{j=1}^n \boldsymbol{P}(j) \cdot \boldsymbol{v}(j)^2}$. If further $\boldsymbol{P}$ has full support, then define $\|\boldsymbol{v}\|_{\boldsymbol{P}}^\star = \sqrt{\sum_{j=1}^n \frac{\boldsymbol{v}(j)^2}{\boldsymbol{P}(j)}}$. The above notations will often be used when $\boldsymbol{P}$ is the mixed strategy profile $\boldsymbol{x}_i$ for some player i and $\boldsymbol{v}$ is a loss vector $\boldsymbol{\ell}_i$; in such a case the norms $\|\boldsymbol{v}\|_{\boldsymbol{P}}$ and $\|\boldsymbol{v}\|_{\boldsymbol{P}}^\star$ are often called *local norms*.

2.2 The Markov Chain Tree Theorem

Given an n-state ergodic (*i.e.*, aperiodic and irreducible) Markov chain with (row-stochastic) transition matrix $\mathbf{Q}$, the classic Markov chain tree theorem provides a closed-form solution for its *stationary distribution*, that is, the unique distribution $\boldsymbol{\pi} \in \Delta^n$ such that $\boldsymbol{\pi}^\top \mathbf{Q} = \boldsymbol{\pi}^\top$. The result requires the notion of a *directed rooted tree* (a.k.a. *arborescence*), recalled next.

Definition 2.5. Let $\mathcal{T} = (V, E)$, with $V = [\![n]\!]$, be an n-node directed graph. $\mathcal{T}$ is a *directed tree rooted at* $j \in V$ if (i) it contains no cycle (including self-loops); (ii) every node $V \setminus \{j\}$ has *exactly one* outgoing edge; and (iii) node j has *no* outgoing edges. We denote the set of all n-node directed trees rooted at $j \in [\![n]\!]$ with the symbol $\mathbb{T}_j^n$. Furthermore, we let $\mathbb{T}^n := \mathbb{T}_1^n \cup \cdots \cup \mathbb{T}_n^n$.

It is well-known that $|\mathbb{T}_j^n| = n^{n-2}$ for all $j \in [\![n]\!]$, and consequently $|\mathbb{T}^n| = n^{n-1}$ [Cay89]. An example for $n = 4$ is given in Figure 1. In this context, the Markov chain tree theorem asserts that the vector $(\Sigma_1, \ldots, \Sigma_n)$ of quantities

$$\Sigma_j := \sum_{\mathcal{T} \in \mathbb{T}_j} \prod_{(a,b) \in E(\mathcal{T})} \mathbf{Q}[a, b] \qquad (j \in [\![n]\!]) \tag{5}$$

is proportional to the stationary distribution. Specifically, we have the following.

Theorem 2.6 (Markov chain tree theorem)**.** *The stationary distribution $\boldsymbol{\pi}$ of an n-state ergodic Markov chain satisfies $\boldsymbol{\pi}[j] = \Sigma_j / \Sigma$ for all $j \in [\![n]\!]$, where $\Sigma := \Sigma_1 + \cdots + \Sigma_n$.*

For a proof of the Markov chain tree theorem, we refer the interested reader to the works of Anantharam and Tsoucas [AT89] and Kruckman, Greenwald, and Wicks [KGW10].

3 Near-Optimal No-External-Regret Dynamics via OMWU

Below we state our main theorem for external regret, which shows that when all players in a game play according to OMWU with appropriate step size, they all experience polylogarithmic individual regrets.

Theorem 3.1 (Formal version of Theorem 1.1)**.** *There are constants $C, C' > 1$ so that the following holds. Suppose a time horizon $T \in \mathbb{N}$ and a game G with m players and n_i actions for each player $i \in [m]$ is given. Suppose all players play according to OMWU with any positive step size $\eta \le \frac{1}{C \cdot m \log^4 T}$. Then for any $i \in [m]$, the regret of player i satisfies*

$$\mathrm{Reg}_i^T \le \frac{\log n_i}{\eta} + C' \cdot \log T. \tag{6}$$

In particular, if the players' step size is chosen as $\eta = \frac{1}{C \cdot m \log^4 T}$, then the regret of player i satisfies

$$\mathrm{Reg}_i^T \le O\left(m \cdot \log n_i \cdot \log^4 T\right). \tag{7}$$

A common goal in the literature on learning in games is to obtain an algorithm that achieves fast rates whan played by all players, and so that each player i still obtains the optimal rate of $O(\sqrt{T})$ in the adversarial setting (i.e., when i receives an arbitrary sequence of losses $\ell_i^{(1)}, \dots, \ell_i^{(T)}$). We show in Corollary D.1 (in the appendix) that by running OMWU with an adaptive step size, this is possible. Table 1 compares our regret bounds discussed in this section to those of prior work.

We overview the proof of Theorem 3.1; the full proof may be found in the appendix.

3.1 New adversarial regret bound

The first step in the proof of Theorem 3.1 is to prove a new regret bound (Lemma 3.2 below) for OMWU that holds for an adversarial sequence of losses. We will show in later sections that when *all* players play according to OMWU, the right-hand side of the regret bound (8) is bounded by a quantity that grows only poly-logarithmically in T.

Lemma 3.2. *There is a constant $C > 0$ so that the following holds. Suppose any player $i \in [m]$ follows the OMWU updates (OMWU) with step size $\eta < 1/C$, for an arbitrary sequence of losses $\ell_i^{(1)}, \dots, \ell_i^{(T)} \in [0,1]^{n_i}$. Then*

$$\mathrm{Reg}_i^T \le \frac{\log n_i}{\eta} + \sum_{t=1}^{T} \left(\frac{\eta}{2} + C\eta^2\right) \mathrm{Var}_{x_i^{(t)}}\left(\ell_i^{(t)} - \ell_i^{(t-1)}\right) - \sum_{t=1}^{T} \frac{(1 - C\eta)\eta}{2} \cdot \mathrm{Var}_{x_i^{(t)}}\left(\ell_i^{(t-1)}\right). \tag{8}$$

The detailed proof of Lemma 3.2 can be found in Section A, but we sketch the main steps here. The starting point is a refinement of [RS13a, Lemma 3] (stated as Lemma A.5), which gives an upper bound for Reg_i^T in terms of local norms corresponding to each of the iterates $x_i^{(t)}$ of OMWU. The bound involves the difference between the OMWU iterates $x_i^{(t)}$ and iterates $\tilde{x}_i^{(t)}$ defined by $\tilde{x}_i^{(t)} = \frac{x_i^{(t)}(j) \cdot \exp(-\eta \cdot (\ell_i^{(t)}(j) - \ell_i^{(t-1)}(j)))}{\sum_{k \in [n_i]} x_i^{(t)}(k) \cdot \exp(-\eta \cdot (\ell_i^{(t)}(k) - \ell_i^{(t-1)}(k)))}$:

$$\mathrm{Reg}_i^T \le \frac{\log n_i}{\eta} + \sum_{t=1}^{T} \left\| x_i^{(t)} - \tilde{x}_i^{(t)} \right\|_{x_i^{(t)}}^{\star} \sqrt{\mathrm{Var}_{x_i^{(t)}}\left(\ell_i^{(t)} - \ell_i^{(t-1)}\right)} - \frac{1}{\eta} \sum_{t=1}^{T} \mathrm{KL}(\tilde{x}_i^{(t)}; x_i^{(t)}) - \frac{1}{\eta} \sum_{t=1}^{T} \mathrm{KL}(x_i^{(t)}; \tilde{x}_i^{(t-1)}). \tag{9}$$

We next show (in Lemma A.2) that $\mathrm{KL}(\tilde{x}_i^{(t)}; x_i^{(t)})$ and $\mathrm{KL}(x_i^{(t)}; \tilde{x}_i^{(t-1)})$ may be lower bounded by $(1/2 - O(\eta)) \cdot \chi^2(\tilde{x}_i^{(t)}; x_i^{(t)})$ and $(1/2 - O(\eta)) \cdot \chi^2(x_i^{(t)}; \tilde{x}_i^{(t-1)})$, respectively. Note it is a standard fact that the KL divergence between two distributions is upper bounded by the chi-squared distribution between them; by contrast, Lemma A.2 can exploit that $x_i^{(t)}$, $\tilde{x}_i^{(t)}$ and $\tilde{x}_i^{(t-1)}$ are close to each other to show a reverse inequality. Finally, exploiting the exponential weights-style functional relationship between $x_i^{(t)}$ and $\tilde{x}_i^{(t-1)}$, we show (in Lemma A.3) that the χ^2-divergence $\chi^2(x_i^{(t)}; \tilde{x}_i^{(t-1)})$ may be lower bounded by $(1 - O(\eta)) \cdot \eta^2 \cdot \mathrm{Var}_{x_i^{(t)}}\left(\ell_i^{(t-1)}\right)$, leading to the term $\frac{(1-C\eta)\eta}{2} \mathrm{Var}_{x_i^{(t)}}\left(\ell_i^{(t-1)}\right)$ being subtracted in (8). The χ^2-divergence $\chi^2(\tilde{x}_i^{(t)}; x_i^{(t)})$, as well as the term $\left\|x_i^{(t)} - \tilde{x}_i^{(t)}\right\|^{\star}_{x_i^{(t)}}$ in (9) are bounded in a similar manner to obtain (8).

3.2 Finite differences

Given Lemma 3.2, in order to establish Theorem 3.1, it suffices to show Lemma 3.3 below. Indeed, (10) below implies that the right-hand side of (8) is bounded above by $\frac{\log n_i}{\eta} + \eta \cdot O(\log^5 T)$, which is bounded above by $O(m \log n_i \log^4 T)$ for the choice $\eta = \Theta\left(\frac{1}{m \cdot \log^4 T}\right)$ of Theorem 3.1.[5]

Lemma 3.3 (Abbreviated; detailed version in Section C.3)**.** *Suppose all players play according to OMWU with step size η satifying $1/T \le \eta \le \frac{1}{Cm \cdot \log^4 T}$ for a sufficiently large constant C. Then for any $i \in [m]$, the losses $\ell_i^{(1)}, \ldots, \ell_i^{(T)} \in \mathbb{R}^{n_i}$ for player i satisfy:*

$$\sum_{t=1}^{T} \mathrm{Var}_{x_i^{(t)}}\left(\ell_i^{(t)} - \ell_i^{(t-1)}\right) \le \frac{1}{2} \cdot \sum_{t=1}^{T} \mathrm{Var}_{x_i^{(t)}}\left(\ell_i^{(t-1)}\right) + O\left(\log^5 T\right). \tag{10}$$

The definition below allows us to streamline our notation when proving Lemma 3.3.

Definition 3.4 (Finite differences)**.** Suppose $L = (L^{(1)}, \ldots, L^{(T)})$ is a sequence of vectors $L^{(t)} \in \mathbb{R}^n$. For integers $h \ge 0$, the *order-h finite difference sequence* for the sequence L, denoted by $\mathrm{D}_h L$, is the sequence $\mathrm{D}_h L := \left((\mathrm{D}_h L)^{(1)}, \ldots, (\mathrm{D}_h L)^{(T-h)}\right)$ defined recursively as: $(\mathrm{D}_0 L)^{(t)} := L^{(t)}$ for all $1 \le t \le T$, and

$$(\mathrm{D}_h L)^{(t)} := (\mathrm{D}_{h-1} L)^{(t+1)} - (\mathrm{D}_{h-1} L)^{(t)} \tag{11}$$

for all $h \ge 1$, $1 \le t \le T - h$.[6]

Remark 3.5. *Notice that another way of writing (11) is: $\mathrm{D}_h L = \mathrm{D}_1 \mathrm{D}_{h-1} L$. We also remark for later use that*

$$(\mathrm{D}_h L)^{(t)} = \sum_{s=0}^{h} \binom{h}{s} (-1)^{h-s} L^{(t+s)}. \tag{12}$$

Let $H = \log T$, where T denotes the fixed time horizon from Theorem 3.1 (and thus Lemma 3.3). In the proof of Lemma 3.3, we will bound the finite differences of order $h \le H$ for certain sequences. The bound (10) of Lemma 3.3 may be rephased as upper bounding $\sum_{t=1}^{T} \mathrm{Var}_{x_i^{(t)}}\left((\mathrm{D}_1 \ell_i)^{(t-1)}\right)$, by $\frac{1}{2}\sum_{t=1}^{T} \mathrm{Var}_{x_i}\left(\ell_i^{(t-1)}\right)$; to prove this, we proceed in two steps:

[5] Notice that the factor $\frac{1}{2}$ in (10) is not important for this argument – any constant less than 1 would suffice.

[6] We remark that while Definition 3.4 is stated for a 1-indexed sequence $L^{(1)}, L^{(2)}, \ldots$, we will also occasionally consider 0-indexed sequences $L^{(0)}, L^{(1)}, \ldots$, in which case the same recursive definition (11) holds for the finite differences $(\mathrm{D}_h L)^{(t)}$, $t \ge 0$.

1. (*Upwards induction step*) First, in Lemma 3.6 below, we find an upper bound on $\left\|(\mathrm{D}_h \ell_i)^{(t)}\right\|_\infty$ for all $t \in [T]$, $h \geq 0$, which decays exponentially in h for $h \leq H$. This is done via *upwards induction* on h, i.e., first proving the base case $h = 0$ using boundedness of the losses $\ell_i^{(t)}$ and then $h = 1, 2, \ldots$ inductively. The main technical tool we develop for the inductive step is a weak form of the chain rule for finite differences, Lemma 3.7. The inductive step uses the fact that all players are following OMWU to relate the hth order finite differences of player i's loss sequence $\ell_i^{(t)}$ to the hth order finite differences of the strategy sequences $x_{i'}^{(t)}$ for players $i' \neq i$; then we use the exponential-weights style updates of OMWU and Lemma 3.7 to relate the hth order finite differences of the strategies $x_{i'}^{(t)}$ to the $(h-1)$th order finite differences of the losses $\ell_{i'}^{(t)}$.

2. (*Downwards induction step*) We next show that for all $0 \leq h \leq H$, $\sum_{t=1}^{T} \mathrm{Var}_{x_i^{(t)}}\left((\mathrm{D}_{h+1} \ell_i)^{(t-1)}\right)$ is bounded above by $c_h \cdot \sum_{t=1}^{T} \mathrm{Var}_{x_i^{(t)}}\left((\mathrm{D}_h \ell_i)^{(t-1)}\right) + \mu_h$, for some $c_h < 1/2$ and $\mu_h < O(\log^5 T)$. This shown via *downwards induction* on h, namely first establishing the base case $h = H$ by using the result of item 1 for $h = H$ and then treating the cases $h = H-1, H-2, \ldots, 0$. The inductive step makes use of the discrete Fourier transform (DFT) to relate the finite differences of different orders (see Lemmas 3.9 and 3.11). In particular, Parseval's equality together with a standard relationship between the DFT of the finite differences of a sequence to the DFT of that sequence allow us to first prove the inductive step in the frequency domain and then transport it back to the original (time) domain.

In the following subsections we explain in further detail how the two steps above are completed.

3.3 Upwards induction proof overview

Addressing item 1 in the previous subsection, the lemma below gives a bound on the supremum norm of the h-th order finite differences of each player's loss vector, when all players play according to OMWU and experience losses according to their loss functions $\mathcal{L}_1, \ldots, \mathcal{L}_m : \mathcal{A} \to [0, 1]$.

Lemma 3.6 (Abbreviated)**.** *Fix a step size $\eta > 0$ satisfying $\eta \leq o\left(\frac{1}{m \log T}\right)$. If all players follow OMWU updates with step size η, then for any player $i \in [m]$, integer h satisfying $0 \leq h \leq H$, and time step $t \in [T-h]$, it holds that $\| (\mathrm{D}_h \ell_i)^{(t)} \|_\infty \leq O(m\eta)^h \cdot h^{O(h)}$.*

A detailed version of Lemma 3.6, together with its full proof, may be found in Section B.4. We next give a proof overview of Lemma 3.6 for the case of 2 players, i.e., $m = 2$; we show in Section B.4 how to generalize this computation to general m. Below we introduce the main technical tool in the proof, a "boundedness chain rule," and then outline how it is used to prove Lemma 3.6.

Main technical tool for Lemma 3.6: boundedness chain rule. We say that a function $\phi : \mathbb{R}^n \to \mathbb{R}$ is a *softmax-type* function if there are real numbers $\xi_1, \ldots, \xi_n$ and some $j \in [n]$ so that for all $(z_1, \ldots, z_n) \in \mathbb{R}^n$, $\phi((z_1, \ldots, z_n)) = \frac{\exp(z_j)}{\sum_{k=1}^{n} \xi_k \cdot \exp(z_k)}$. Lemma 3.7 below may be interpreted as a "boundedness chain rule" for finite differences. To explain the context for this lemma, recall that given an infinitely differentiable vector-valued function $L : \mathbb{R} \to \mathbb{R}^n$ and an infinitely differentiable function $\phi : \mathbb{R}^n \to \mathbb{R}$, the higher order derivatives of the function $\phi(L(t))$ may be computed in terms of those of L and ϕ using the chain rule. Lemma 3.7 considers an analogous setting where the input variable t to L is discrete-valued, taking values in $[T]$ (and so we identify the function L

with the sequence $L^{(1)}, \ldots, L^{(T)}$). In this case, the *higher order finite differences* of the sequence $L^{(1)}, \ldots, L^{(T)}$ (Definition 3.4) take the place of the higher order derivatives of L with respect to t. Though there is no generic chain rule for finite differences, Lemma 3.7 states that, at least when ϕ is a softmax-type function, we may *bound* the higher order finite differences of the sequence $\phi(L^{(1)}), \ldots, \phi(L^{(T)})$. In the lemma's statement we let $\phi \circ L$ denote the sequence $\phi(L^{(1)}), \ldots, \phi(L^{(T)})$.

Lemma 3.7 ("Boundedness chain rule" for finite differences; abbreviated)**.** *Suppose that $h, n \in \mathbb{N}$, $\phi : \mathbb{R}^n \to \mathbb{R}$ is a softmax-type function, and $L = (L^{(1)}, \ldots, L^{(T)})$ is a sequence of vectors in $\mathbb{R}^n$ satisfying $\|L^{(t)}\|_\infty \leq 1$ for $t \in [T]$. Suppose for some $\alpha \in (0, 1)$, for each $0 \leq h' \leq h$ and $t \in [T - h']$, it holds that $\|\mathrm{D}_{h'} L^{(t)}\|_\infty \leq O(\alpha^{h'}) \cdot (h')^{O(h')}$. Then for all $t \in [T - h]$,*

$$|\left(\mathrm{D}_h \left(\phi \circ L\right)\right)^{(t)}| \leq O(\alpha^h) \cdot h^{O(h)}.$$

A detailed version of Lemma 3.7 may be found in Section B.3. While Lemma 3.7 requires ϕ to be a softmax-type function for simplicity (and this is the only type of function ϕ we will need to consider for the case $m = 2$) we remark that the detailed version of Lemma 3.7 allows ϕ to be from a more general family of analytic functions whose higher order derivatives are appropriately bounded. The proof of Lemma 3.6 for all $m \geq 2$ requires that more general form of Lemma 3.7.

The proof of Lemma 3.7 proceeds by considering the Taylor expansion $P_\phi(\cdot)$ of the function ϕ at the origin, which we write as follows: for $z = (z_1, \ldots, z_n) \in \mathbb{R}^n$, $P_\phi(z) := \sum_{k \geq 0, \gamma \in \mathbb{Z}^n_{\geq 0} :\ |\gamma| = k} a_\gamma z^\gamma$, where $a_\gamma \in \mathbb{R}$, $|\gamma|$ denotes the quantity $\gamma_1 + \cdots + \gamma_n$ and z^γ denotes $z_1^{\gamma_1} \cdots z_n^{\gamma_n}$. The fact that ϕ is a softmax-type function ensures that the radius of convergence of its Taylor series is at least 1, i.e., $\phi(z) = P_\phi(z)$ for any z satisfying $\|z\|_\infty \leq 1$. By the assumption that $\|L^{(t)}\|_\infty \leq 1$ for each t, we may therefore decompose $\left(\mathrm{D}_h \left(\phi \circ L\right)\right)^{(t)}$ as:

$$\left(\mathrm{D}_h \left(\phi \circ L\right)\right)^{(t)} = \sum_{k \geq 0, \gamma \in \mathbb{Z}^n_{\geq 0} :\ |\gamma| = k} a_\gamma \cdot \left(\mathrm{D}_h \, L^\gamma\right)^{(t)}, \tag{13}$$

where L^γ denotes the sequence of scalars $(L^\gamma)^{(t)} := (L^{(t)})^\gamma$ for all t. The fact that ϕ is a softmax-type function allows us to establish strong bounds on $|a_\gamma|$ for each γ in Lemma B.6. The proof of Lemma B.6 bounds the $|a_\gamma|$ by exploiting the simple form of the derivative of a softmax-type function to decompose each a_γ into a sum of $|\gamma|!$ terms. Then we establish a bijection between the terms of this decomposition and graph structures we refer to as *factorial trees*; that bijection together with the use of an appropriate generating function allow us to complete the proof of Lemma B.6.

Thus, to prove Lemma 3.7, it suffices to bound $\left|\left(\mathrm{D}_h \, L^\gamma\right)^{(t)}\right|$ for all γ. We do so by using Lemma 3.8.

Lemma 3.8 (Abbreviated; detailed vesion in Section B.2)**.** *Fix any $h \geq 0$, a multi-index $\gamma \in \mathbb{Z}^n_{\geq 0}$ and set $k = |\gamma|$. For each of the k^h functions $\pi : [h] \to [k]$, and for each $r \in [k]$, there are integers $h'_{\pi,r} \in \{0, 1, \ldots, h\}$, $t'_{\pi,r} \geq 0$, and $j'_{\pi,r} \in [n]$, so that the following holds. For any sequence $L^{(1)}, \ldots, L^{(T)} \in \mathbb{R}^n$ of vectors, it holds that, for each $t \in [T - h]$,*

$$\left(\mathrm{D}_h \, L^\gamma\right)^{(t)} = \sum_{\pi : [h] \to [k]} \prod_{r=1}^{k} \left(\mathrm{D}_{h'_{\pi,r}} \left(L(j'_{\pi,r})\right)\right)^{(t + t'_{\pi,r})}. \tag{14}$$

Lemma 3.8 expresses the hth order finite differences of the sequence L^γ as a sum of k^h terms, each of which is a product of k finite order differences of a sequence $L^{(t)}(j'_{\pi,r})$ (i.e., the $j'_{\pi,r}$th coordinate of the vectors $L^{(t)}$). Crucially, when using Lemma 3.8 to prove Lemma 3.7, the assumption of Lemma 3.7 gives that for each $j' \in [n]$, each $h' \in [h]$, and each $t' \in [T-h']$, we have the bound $\left|\left(\mathrm{D}_{h'} L(j')\right)^{(t')}\right| \leq O(\alpha^{h'}) \cdot (h')^{O(h')}$. These assumed bounds may be used to bound the right-hand side of (14), which together with Lemma 3.8 and (13) lets us complete the proof of Lemma 3.7.

Proving Lemma 3.6 using the boundedness chain rule. Next we discuss how Lemma 3.7 is used to prove Lemma 3.6, namely to bound $\| \left(\mathrm{D}_h \ell_i\right)^{(t)} \|_\infty$ for each $t \in [T-h]$, $i \in [m]$, and $0 \leq h \leq H$. Lemma 3.6 is proved using induction, with the base case $h = 0$ being a straightforward consequence of the fact that $\| \left(\mathrm{D}_0 \ell_i\right)^{(t)} \|_\infty = \|\ell_i^{(t)}\|_\infty \leq 1$ for all $i \in [m], t \in [T]$. For the rest of this section we focus on the inductive case, i.e., we pick some $h \in [H]$ and assume Lemma 3.6 holds for all $h' < h$.

The first step is to reduce the claim of Lemma 3.6 to the claim that the upper bound $\| \left(\mathrm{D}_h x_i\right)^{(t)} \|_1 \leq O\left(m\eta\right)^h \cdot h^{O(h)}$ holds for each $t \in [T-h], i \in [m]$. Recalling that we are only sketching here the case $m = 2$ for simplicity, this reduction proceeds as follows: for $i \in \{1, 2\}$, define the matrix $A_i \in \mathbb{R}^{n_1 \times n_2}$ by $(A_i)_{a_1 a_2} = \mathcal{L}_i(a_1, a_2)$, for $a_1 \in [n_1], a_2 \in [n_2]$. We have assumed that all players are using OMWU and thus $\ell_i^{(t)} = \mathbb{E}_{a_{i'} \sim x_{i'}^{(t)},\ \forall i' \neq i}[\mathcal{L}_i(a_1, \ldots, a_n)]$; for our case here ($m = 2$), this may be rewritten as $\ell_1^{(t)} = A_1 x_2^{(t)}$, $\ell_2^{(t)} = A_2^\top x_1^{(t)}$. Thus

$$\| \left(\mathrm{D}_h \ell_1\right)^{(t)} \|_\infty = \left\| A_1 \cdot \sum_{s=0}^{h} \binom{h}{s} (-1)^{h-s} x_2^{(t+s)} \right\|_\infty \leq \left\| \sum_{s=0}^{h} \binom{h}{s} (-1)^{h-s} x_2^{(t+s)} \right\|_1 = \| \left(\mathrm{D}_h x_2\right)^{(t)} \|_1,$$

where the first equality is from Remark 3.5 and the inequality follows since all entries of A_1 have absolute value ≤ 1. A similar computation allows us to show $\| \left(\mathrm{D}_h \ell_2\right)^{(t)} \|_\infty \leq \| \left(\mathrm{D}_h x_1\right)^{(t)} \|_1$.

To complete the inductive step it remains to upper bound the quantities $\| \left(\mathrm{D}_h x_i\right)^{(t)} \|_1$ for $i \in [m], t \in [T-h]$. To do so, we note that the definition of the OMWU updates (OMWU) implies that for any $i \in [m], t \in [T], j \in [n_i]$, and $t' \geq 1$, we have

$$x_i^{(t+t')}(j) = \frac{x_i^{(t)}(j) \cdot \exp\left(\eta \cdot \left(\ell_i^{(t-1)}(j) - \sum_{s=0}^{t'-1} \ell_i^{(t+s)}(j) - \ell_i^{(t+t'-1)}(j)\right)\right)}{\sum_{k=1}^{n_i} x_i^{(t)}(k) \cdot \exp\left(\eta \cdot \left(\ell_i^{(t-1)}(k) - \sum_{s=0}^{t'-1} \ell_i^{(t+s)}(k) - \ell_i^{(t+t'-1)}(k)\right)\right)}. \tag{15}$$

For $t \in [T]$, $t' \geq 0$, set $\bar{\ell}_{i,t}^{(t')} := \eta \cdot \left(\ell_i^{(t-1)} - \sum_{s=0}^{t'-1} \ell_i^{(t+s)} - \ell_i^{(t+t'-1)}\right)$. Also, for each $i \in [m]$, $j \in [n_i]$, $t \in [T]$, and any vector $z = (z(1), \ldots, z(n_i)) \in \mathbb{R}^{n_i}$ define $\phi_{t,i,j}(z) := \frac{x_i^{(t)}(j) \cdot \exp(z(j))}{\sum_{k=1}^{n_i} x_i^{(t)}(k) \cdot \exp(z(k))}$. Thus (15) gives that for $t' \geq 1$, $x_i^{(t+t')}(j) = \phi_{t,i,j}(\bar{\ell}_{i,t}^{(t')})$. Viewing t as a fixed parameter and letting t' vary, it follows that for $h \geq 0$ and $t' \geq 1$, $\left(\mathrm{D}_h\, x_i^{(t+\cdot)}(j)\right)^{(t')} = \left(\mathrm{D}_h \left(\phi_{t,i,j} \circ \bar{\ell}_{i,t}\right)\right)^{(t')}$.

Recalling that our goal is to bound $| \left(\mathrm{D}_h x_i(j)\right)^{(t+1)} |$ for each t, we can do so by using Lemma 3.7 with $\phi = \phi_{t,i,j}$ and $\alpha = O(m\eta)$, if we can show that its precondition is met, i.e. that $\| \left(\mathrm{D}_{h'} \bar{\ell}_{i,t}\right)^{(t')} \|_\infty \leq \frac{1}{B_1} \cdot \alpha^{h'} \cdot (h')^{B_0 h'}$ for all $h' \leq h$, the appropriate α and appropriate constants B_0, B_1. Helpfully, the definition of $\bar{\ell}_{i,t}^{(t')}$ as a partial sum allows us to relate the h'-th order finite

differences of the sequence $\bar{\ell}_{i,t}^{(t')}$ to the $(h'-1)$-th order finite differences of the sequence $\ell_i^{(t)}$ as follows:

$$\left(\mathrm{D}_{h'}\,\bar{\ell}_{i,t}\right)^{(t')} = \eta\cdot\left(\mathrm{D}_{h'-1}\,\ell_i\right)^{(t+t'-1)} - 2\eta\cdot\left(\mathrm{D}_{h'-1}\,\ell_i\right)^{(t+t')}. \tag{16}$$

Since $h'-1<h$ for $h'\le h$, the inductive assumption of Lemma 3.6 gives a bound on the ℓ_∞-norm of the terms on the right-hand side of (16), which are sufficient for us to apply Lemma 3.7. Note that the inductive assumption gives an upper bound on $\|\left(\mathrm{D}_{h'-1}\,\ell_i\right)^{(t)}\|_\infty$ that only scales with $\alpha^{h'-1}$, whereas Lemma 3.7 requires scaling of $\alpha^{h'}$. This discrepancy is corrected by the factor of η on the right-hand side of (16), which gives the desired scaling $\alpha^{h'}$ (since $\eta<\alpha$ for the choice $\alpha=O(m\eta)$).

3.4 Downwards induction proof overview

In this section we discuss in further detail item 2 in Section 3.2; in particular, we will show that there is a parameter $\mu=\tilde{\Theta}(\eta m)$ so that for all integers h satisfying $H-1\ge h\ge 0$,

$$\sum_{t=1}^{T-h-1}\mathrm{Var}_{x_i^{(t)}}\left(\left(\mathrm{D}_{h+1}\,\ell_i\right)^{(t)}\right)\le O(1/H)\cdot\sum_{t=1}^{T-h}\mathrm{Var}_{x_i^{(t)}}\left(\left(\mathrm{D}_{h}\,\ell_i\right)^{(t)}\right)+\tilde{O}\left(\mu^{2h}\right), \tag{17}$$

where $\tilde{O}$ hides factors polynomial in $\log T$. The validity of (17) for $h=0$ implies Lemma 3.3. On the other hand, as long we choose the value μ in (17) to satisfy $\mu\ge m\eta H^{\Omega(1)}$, then Lemma 3.6 implies that $\sum_{t=1}^{T-H}\mathrm{Var}_{x_i^{(t)}}\left(\left(\mathrm{D}_{H}\,\ell_i\right)^{(t)}\right)\le O(\mu^{2H})$. This gives that (17) holds for $h=H-1$. To show that (17) holds for all $H-1>h\ge 0$, we use downwards induction; fix any h, and assume that (17) has been shown for all h' satisfying $h<h'\le H-1$. Our main tool in the inductive step is to apply Lemma 3.9 below. To state it, for $\zeta>0$, $n\in\mathbb{N}$, we say that a sequence of distributions $P^{(1)},\dots,P^{(T)}\in\Delta^n$ is *ζ-consecutively close* if for each $1\le t<T$, it holds that $\max\left\{\left\|\frac{P^{(t)}}{P^{(t+1)}}\right\|_\infty,\left\|\frac{P^{(t+1)}}{P^{(t)}}\right\|_\infty\right\}\le 1+\zeta$.[7] Lemma 3.9 shows that given a sequence of vectors for which the variances of its second-order finite differences are bounded by the variances of its first-order finite differences, a similar relationship holds between its first- and zeroth-order finite differences.

Lemma 3.9. *There is a sufficiently large constant $C_0>1$ so that the following holds. For any $M,\zeta,\alpha>0$ and $n\in\mathbb{N}$, suppose that $P^{(1)},\dots,P^{(T)}\in\Delta^n$ and $Z^{(1)},\dots,Z^{(T)}\in[-M,M]^n$ satisfy the following conditions:*

1. *The sequence $P^{(1)},\dots,P^{(T)}$ is ζ-consecutively close for some $\zeta\in[1/(2T),\alpha^4/C_0]$.*
2. *It holds that $\sum_{t=1}^{T-2}\mathrm{Var}_{P^{(t)}}\left(\left(\mathrm{D}_2\,Z\right)^{(t)}\right)\le\alpha\cdot\sum_{t=1}^{T-1}\mathrm{Var}_{P^{(t)}}\left(\left(\mathrm{D}_1\,Z\right)^{(t)}\right)+\mu$.*

Then $\sum_{t=1}^{T-1}\mathrm{Var}_{P^{(t)}}\left(\left(\mathrm{D}_1\,Z\right)^{(t)}\right)\le\alpha\cdot(1+\alpha)\sum_{t=1}^{T}\mathrm{Var}_{P^{(t)}}\left(Z^{(t)}\right)+\frac{\mu}{\alpha}+\frac{C_0M^2}{\alpha^3}$.

Given Lemma 3.9, the inductive step for establishing (17) is straightforward: we apply Lemma 3.9 with $P^{(t)}=x_i^{(t)}$ and $Z^{(t)}=\left(\mathrm{D}_h\,\ell_i\right)^{(t)}$ for all t. The fact that $x_i^{(t)}$ are updated with OMWU may be used to establish that precondition 1 of Lemma 3.9 holds. Since $\left(\mathrm{D}_1\,Z\right)^{(t)}=\left(\mathrm{D}_{h+1}\,\ell_i\right)^{(t)}$ and $\left(\mathrm{D}_2\,Z\right)^{(t)}=\left(\mathrm{D}_{h+2}\,\ell_i\right)^{(t)}$, that the inductive hypothesis (17) holds for $h+1$ implies that precondition 2 of Lemma 3.9 holds for appropriate $\alpha,\mu>0$. Thus Lemma 3.9 implies that (17) holds for the value h, which completes the inductive step.

[7] Here, for distributions $P,Q\in\Delta^n$, $\frac{P}{Q}\in\mathbb{R}^n$ denotes the vector whose jth entry is $P(j)/Q(j)$.

On the proof of Lemma 3.9. Finally we discuss the proof of Lemma 3.9. One technical challenge is the fact that the vectors $P^{(t)}$ are not constant functions of t, but rather change slowly (as constrained by being ζ-consecutively close). The main tool for dealing with this difficulty is Lemma C.1, which shows that for a ζ-consecutively close sequence $P^{(t)}$, for any vector $Z^{(t)}$, $\frac{\mathrm{Var}_{P^{(t)}}\left(Z^{(t)}\right)}{\mathrm{Var}_{P^{(t+1)}}\left(Z^{(t)}\right)} \in [1-\zeta, 1+\zeta]$. This fact, together with some algebraic manipulations, lets us to reduce to the case that all $P^{(t)}$ are equal. It is also relatively straightforward to reduce to the case that $\langle P^{(t)}, Z^{(t)}\rangle = 0$ for all t, i.e., so that $\mathrm{Var}_{P^{(t)}}\left(Z^{(t)}\right) = \left\|Z^{(t)}\right\|_{P^{(t)}}^2$. We may further separate $\left\|Z^{(t)}\right\|_{P^{(t)}}^2 = \sum_{j=1}^n P^{(t)}(j)\cdot(Z^{(t)}(j))^2$ into its individual components $P^{(t)}(j)\cdot(Z^{(t)}(j))^2$, and treat each one separately, thus allowing us to reduce to a one-dimensional problem. Finally, we make one further reduction, which is to replace the finite differences $\mathrm{D}_h(\cdot)$ in Lemma 3.9 with *circular finite differences*, defined below:

Definition 3.10 (Circular finite difference)**.** Suppose $L = (L^{(0)}, \ldots, L^{(S-1)})$ is a sequence of vectors $L^{(t)} \in \mathbb{R}^n$. For integers $h \geq 0$, the *level-h circular finite difference sequence* for the sequence L, denoted by $\mathrm{D}_h^\circ L$, is the sequence defined recursively as: $\left(\mathrm{D}_0^\circ L\right)^{(t)} = L^{(t)}$ for all $0 \leq t < S$, and

$$\left(\mathrm{D}_h^\circ L\right)^{(t)} = \begin{cases} \left(\mathrm{D}_{h-1}^\circ L\right)^{(t+1)} - \left(\mathrm{D}_{h-1}^\circ L\right)^{(t)} & : 0 \leq t \leq S-2 \\ \left(\mathrm{D}_{h-1}^\circ L\right)^{(1)} - \left(\mathrm{D}_{h-1}^\circ L\right)^{(T)} & : t = S-1. \end{cases} \tag{18}$$

Circular finite differences for a sequence $L^{(0)}, \ldots, L^{(S-1)}$ are defined similarly to finite differences (Definition 3.4) except that unlike for finite differences, where $\left(\mathrm{D}_h L\right)^{(S-h)}, \ldots, \left(\mathrm{D}_h L\right)^{(S-1)}$ are not defined, $\left(\mathrm{D}_h^\circ L\right)^{(S-h)}, \ldots, \left(\mathrm{D}_h^\circ L\right)^{(S-1)}$ are defined by "wrapping around" back to the beginning of the sequence. The above-described reductions, which are worked out in detail in Section C.2, allow us to reduce proving Lemma 3.9 to proving the following simpler lemma:

Lemma 3.11. *Suppose $\mu \in \mathbb{R}$, $\alpha > 0$, and $W^{(0)}, \ldots, W^{(S-1)} \in \mathbb{R}$ is a sequence of reals satisfying*

$$\sum_{t=0}^{S-1}\left(\left(\mathrm{D}_2^\circ W\right)^{(t)}\right)^2 \leq \alpha \cdot \sum_{t=0}^{S-1}\left(\left(\mathrm{D}_1^\circ W\right)^{(t)}\right)^2 + \mu. \tag{19}$$

Then $\sum_{t=0}^{S-1}\left(\left(\mathrm{D}_1^\circ W\right)^{(t)}\right)^2 \leq \alpha \cdot \sum_{t=1}^{S-1}(W^{(t)})^2 + \mu/\alpha$.

To prove Lemma 3.11, we apply the discrete Fourier transform to both sides of (19) and use the Cauchy-Schwarz inequality in frequency domain. For a sequence $W^{(0)}, \ldots, W^{(S-1)} \in \mathbb{R}$, its (discrete) *Fourier transform* is the sequence $\widehat{W}^{(0)}, \ldots, \widehat{W}^{(S-1)}$ defined by $\widehat{W}^{(s)} = \sum_{t=0}^{S-1} W^{(t)} \cdot e^{-\frac{2\pi i s t}{S}}$. Below we prove Lemma 3.11 for the special case $\mu = 0$; we defer the general case to Section C.1.

Proof of Lemma 3.11 for special case $\mu = 0$. We have the following:

$$S \cdot \sum_{t=1}^{T}\left(\left(\mathrm{D}_1^\circ W\right)^{(t)}\right)^2 = \sum_{s=1}^{T}\left|\widehat{\mathrm{D}_1^\circ W}^{(s)}\right|^2 = \sum_{s=1}^{T}\left|\widehat{W}^{(s)}(e^{2\pi i s/T} - 1)\right|^2 \leq \sqrt{\sum_{s=1}^{T}\left|\widehat{W}^{(s)}\right|^2}\sqrt{\sum_{s=1}^{T}\left|\widehat{W}^{(s)}\right|^2\left|e^{2\pi i s/T}-1\right|^4},$$

where the first equality uses Parseval's equality, the second uses Fact C.3 (in the appendix) for $h = 1$, and the inequality uses Cauchy-Schwarz. By Parseval's inequality and Fact C.3 for $h = 2$, the right-hand side of the above equals $S \cdot \sqrt{\sum_{t=1}^{T}(W^{(t)})^2} \cdot \sqrt{\sum_{t=1}^{T}\left(\left(\mathrm{D}_2^\circ W\right)^{(t)}\right)^2}$, which, by assumption, is at most $S \cdot \sqrt{\sum_{t=1}^{T}(W^{(t)})^2} \cdot \sqrt{\alpha \cdot \sum_{t=1}^{T}\left(\left(\mathrm{D}_1^\circ W\right)^{(t)}\right)^2}$. Rearranging terms completes the proof. □

4 Near-Optimal No-Internal-Regret Dynamics via SL-OMWU

In this section we establish our main result regarding algorithm SL-OMWU, namely Theorem 1.2. In particular, we start in Section 4.1 by formalizing one of our key insights: SL-OMWU can be equivalently thought of as a certain linear transformation of the output of an external regret minimizer operating over the combinatorial space of directed trees; this equivalence is formalized in Theorem 4.1. Next, in Section 4.2 we leverage this connection to bound the internal regret of SL-OMWU using an extension of the techniques developed in Theorem 1.1.

Proof of Theorem 1.2: Analyzing SL-OMWU. We show that internal regret minimization on an n-simplex can be simulated by no-external-regret dynamics on the combinatorial space of all n-node *directed trees* (Theorem 4.1). Our equivalence result enables us to trade the computation of a stationary distribution of a polynomial-sized Markov chain for a (much more well-behaved) linear transformation on an exponential-sized set. To our knowledge, this is the first no-internal-to-no-external-regret reduction that sidesteps the computation of stationary distributions of Markov chains, and might have applications beyond the characterization of high-order smoothness of the dynamics. Based on our equivalence result, we then adapt and leverage the high-order smoothness result of Lemma 3.3 for no-external-regret dynamics. We stress that our analysis is eventually brought back to a "low-dimensional" regret minimizer, instead of solely operating over the space of directed trees; this step is crucial for obtaining the logarithmic dependence on the number of actions of each player for no-internal-regret dynamics (Theorem 1.2).

The equivalence result mentioned in the previous paragraph arises as a consequence of the classic Markov chain tree theorem, which provides a closed-form combinatorial formula for the stationary distribution of an ergodic Markov chain, and crucially relies on the multiplicative structure of the update rule of (O)MWU. Specifically, we prove that the stationary distributions of certain Markov chains whose transition probabilities are updated through (O)MWU are themselves linear transformations of iterates produced by (O)MWU. Our equivalence gives a direct way to argue about the high-order smoothness of stationary distributions of Markov chains, substantially extending the first-order smoothness observation of Chen and Peng [CP20]. Furthermore, we expect the equivalence to continue to hold beyond stationary distributions of Markov chains, to the more general problem of computing *fixed points* of linear transformations required in the framework of *Phi-regret* [SL07; GJ03; Far+21].

4.1 Equivalence Result

Before we proceed with the statement and proof of Theorem 4.1, we first summarize SL-OMWU in Algorithm 1; a more comprehensive overview of this construction is included in Appendix E.1. In addition, in Algorithm 2 we present an external regret minimizer over the space of all directed trees (a.k.a. arborescences). Both Algorithms 1 and 2 use the symbol $\boldsymbol{x}^{(t)}$ to denote the output strategies; this choice is justified by the following theorem (see also Figure 2).

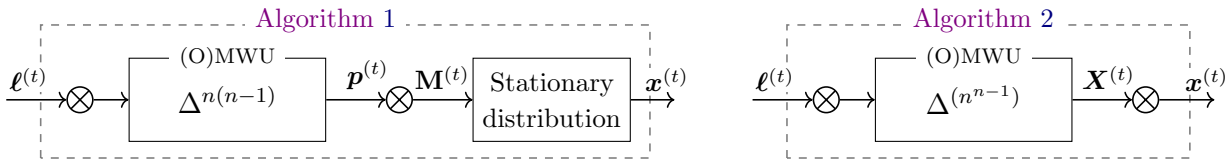

Figure 2: A schematic illustration of the equivalence result of Theorem 4.1; $\otimes$ in the figure represents a linear transformation.

Algorithm 1: Stoltz and Lugosi [SL05]

Data: $\mathcal{R}_\Delta$: OMWU algorithm for $\Delta^{n(n-1)}$ with learning rate $\eta > 0$

1 **function** NextStrategy()
2 $\quad \boldsymbol{p}^{(t)} \leftarrow \mathcal{R}_\Delta$.NextStrategy()
3 $\quad \mathbf{M}^{(t)} \leftarrow \sum_{j \neq k \in [\![n]\!]} \boldsymbol{p}^{(t)}[j \to k]\, \mathbf{E}_{j\to k}^\top$
4 $\quad$ **return** $\boldsymbol{x}^{(t)} \leftarrow$ StationDistrib($\mathbf{M}^{(t)}$) [8]
5 **function** ObserveUtility($\boldsymbol{\ell}^{(t)}$)
6 $\quad \boldsymbol{L}^{(t)} \leftarrow \mathbf{0} \in \mathbb{R}^{n(n-1)}$
7 $\quad$ **for** $j \neq k \in [\![n]\!]$ **do**
8 $\quad\quad \boldsymbol{L}^{(t)}[j \to k] \leftarrow \boldsymbol{x}^{(t)}[j](\boldsymbol{\ell}^{(t)}[k] - \boldsymbol{\ell}^{(t)}[j])$
9 $\quad \mathcal{R}_\Delta$.ObserveUtility($\boldsymbol{L}^{(t)}$)

Algorithm 2: Arborescence-based dynamics

Data: $\mathcal{R}_\Delta$: OMWU algorithm for $\Delta^{(n^{n-1})}$ with learning rate $\eta > 0$

1 **function** NextStrategy()
2 $\quad \boldsymbol{X}^{(t)} \leftarrow \mathcal{R}_\Delta$.NextStrategy()
3 $\quad$ **return** $\boldsymbol{x}^{(t)} \leftarrow \left(\sum_{\mathcal{T} \in \mathbb{T}_j^n} \boldsymbol{X}^{(t)}[\mathcal{T}]\right)_{j=1}^{n}$
4 **function** ObserveUtility($\boldsymbol{\ell}^{(t)}$)
5 $\quad \boldsymbol{\mathcal{L}}^{(t)} = \mathbf{0} \in \mathbb{R}^{|\mathbb{T}^n|} = \mathbb{R}^{(n^{n-1})}$
6 $\quad$ **for** $\mathcal{T} \in \mathbb{T}^n$ **do**
7 $\quad\quad \boldsymbol{\mathcal{L}}^{(t)}[\mathcal{T}] \leftarrow \sum_{(j,k) \in E(\mathcal{T})} \boldsymbol{x}^{(t)}[j](\boldsymbol{\ell}^{(t)}[k] - \boldsymbol{\ell}^{(t)}[j])$
8 $\quad \mathcal{R}_\Delta$.ObserveUtility($\boldsymbol{\mathcal{L}}^{(t)}$)

Theorem 4.1. *For any learning rate $\eta > 0$, Algorithms 1 and 2 produce the same strategies $\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(T)} \in \Delta^n$, assuming that they observe the same sequence of losses $\boldsymbol{\ell}^{(1)}, \dots, \boldsymbol{\ell}^{(T)} \in \mathbb{R}^n$.*

Proof. We will inductively show that the following property holds.

Property 4.2. *At all times $t \geq 1$, the following conditions hold:*

1. *there exists $N^{(t)} \in \mathbb{R}$ such that*

$$\prod_{(a,b) \in E(\mathcal{T})} \boldsymbol{p}^{(t)}[a \to b] = N^{(t)} \boldsymbol{X}^{(t)}[\mathcal{T}] \qquad \forall\, \mathcal{T} \in \mathbb{T}^n, \tag{20}$$

   where $\boldsymbol{p}^{(t)}$ and $\boldsymbol{X}^{(t)}$ are as defined in Algorithm 1 and Algorithm 2, respectively; and
2. *the strategies produced by Algorithms 1 and 2 are the same.*

Condition 2 of Property 4.2 immediately implies the statement. First, we establish the base case $t = 1$. The first iterate of OMWU is always the uniform distribution; so, $\boldsymbol{p}^{(1)} = 1/n(n-1)\, \mathbf{1} \in \Delta^{n(n-1)}$. Hence, the matrix $\mathbf{M}^{(1)}$ (Line 3 of Algorithm 1) has entries $\mathbf{M}[j,k] = 1/n(n-1)$ for all $j \neq k \in [\![n]\!]$. Correspondingly, the (unique) stationary distribution of $\mathbf{M}^{(1)}$ is the uniform distribution $\boldsymbol{x}^{(1)} = \frac{1}{n}\mathbf{1} \in \Delta^n$. We now verify that the same iterate is produced by Algorithm 2. Since $\boldsymbol{X}^{(1)}$ is computed

[8] The matrix $\mathbf{M}^{(t)}$ has strictly positive entries at all times t, since the iterates $\boldsymbol{p}^{(t)}$ produced by (OMWU) lie in the relative interior of the simplex. So, each $\mathbf{M}^{(t)}$ admits a *unique* fixed point (stationary distribution).

using OMWU, $\boldsymbol{X}^{(1)} \in \Delta^{n^{n-1}}$ is the uniform distribution $\boldsymbol{X}^{(1)}[\mathcal{T}] = 1/n^{n-1}$. So, using the fact that $|\mathbb{T}^n_j| = n^{n-2}$ for all $j \in [\![n]\!]$, each coordinate of the output strategy of Algorithm 2 is equal to $n^{n-2}/n^{n-1} = 1/n$, establishing the base case for Condition 2 of Property 4.2. Furthermore, since $\boldsymbol{p}^{(1)}$ and $\boldsymbol{X}^{(1)}$ are the uniform distributions over $\Delta^{n(n-1)}$ and $\Delta^{(n^{n-1})}$ respectively, Condition 1 follows directly from the fact that any directed tree $\mathcal{T} \in \mathbb{T}^n$ has exactly $n-1$ edges.

Next, we prove the inductive step. Assume that Property 4.2 holds at times $\tau = 1, \dots, t$, for some $t \geq 1$. We will prove that it will hold at time $t+1$ as well. Since $\boldsymbol{p}^{(t+1)} \in \Delta^{n(n-1)}$ in Algorithm 1 is updated using (OMWU), for all $j \neq k \in [\![n]\!]$

$$\boldsymbol{p}^{(t+1)}[j \to k] = \frac{\exp\{-\eta(2\boldsymbol{L}^{(t)}[j \to k] - \boldsymbol{L}^{(t-1)}[j \to k])\}}{\sum_{a \neq b} \exp\{-\eta(2\boldsymbol{L}^{(t)}[a \to b] - \boldsymbol{L}^{(t-1)}[a \to b])\}\boldsymbol{p}^{(t)}[a \to b]} \boldsymbol{p}^{(t)}[j \to k], \tag{21}$$

where the loss vector

$$\boldsymbol{L}^{(t)}[j \to k] \coloneqq \boldsymbol{x}^{(t)}[j](\boldsymbol{\ell}^{(t)}[k] - \boldsymbol{\ell}^{(t)}[j]) \tag{22}$$

is as defined on Line 8 of Algorithm 1. For convenience, we will denote the denominator in Equation (21) as $S^{(t+1)}$. Similarly, since the strategies $\boldsymbol{X}^{(t)} \in \Delta^{(n^{n-1})}$ of Algorithm 2 are updated using (OMWU) with the *same* learning rate $\eta > 0$, we have

$$\boldsymbol{X}^{(t+1)}[\mathcal{T}] = \frac{\exp\{-\eta(2\boldsymbol{\mathcal{L}}^{(t)}[\mathcal{T}] - \boldsymbol{\mathcal{L}}^{(t-1)}[\mathcal{T}])\}}{\sum_{\mathcal{T}' \in \mathbb{T}^n} \exp\{-\eta(2\boldsymbol{\mathcal{L}}^{(t)}[\mathcal{T}'] - \boldsymbol{\mathcal{L}}^{(t-1)}[\mathcal{T}']\})\, \boldsymbol{X}^{(t)}[\mathcal{T}']} \boldsymbol{X}^{(t)}[\mathcal{T}] \qquad \forall\, \mathcal{T} \in \mathbb{T}^n, \tag{23}$$

where $\boldsymbol{\mathcal{L}}^{(t)}$ is defined on Line 7 of Algorithm 2 as

$$\boldsymbol{\mathcal{L}}^{(t)}[\mathcal{T}] \coloneqq \sum_{(a,b) \in E(\mathcal{T})} \boldsymbol{x}^{(t)}[a](\boldsymbol{\ell}^{(t)}[b] - \boldsymbol{\ell}^{(t)}[a]) = \sum_{(a,b) \in E(\mathcal{T})} \boldsymbol{L}^{(t)}[a \to b] \qquad \forall\, \mathcal{T} \in \mathbb{T}^n. \tag{24}$$

(Note that in (22) and (24) we implicitly used the inductive hypothesis that the strategies of the two algorithms coincide at iterate t, as well as the assumption that the sequence of losses $\boldsymbol{\ell}^{(t)}$ observed by Algorithms 1 and 2 is the same.) To simplify the notation, we will use $\mathcal{S}^{(t+1)}$ to represent the denominator in (23). Next, we observe that for any $\mathcal{T} \in \mathbb{T}^n$

$$\begin{aligned}
\prod_{(a,b) \in E(\mathcal{T})} \boldsymbol{p}^{(t+1)}[a \to b] &= \left(\frac{1}{S^{(t+1)}}\right)^{n-1} \prod_{(a,b) \in E(\mathcal{T})} \exp\Big\{-\eta(2\boldsymbol{L}^{(t)}[a \to b] - \boldsymbol{L}^{(t-1)}[a \to b])\Big\}\boldsymbol{p}^{(t)}[a \to b] && (25)\\
&= \left(\frac{1}{S^{(t+1)}}\right)^{n-1} \exp\left\{-\eta \sum_{(a,b) \in E(\mathcal{T})} 2\boldsymbol{L}^{(t)}[a \to b] - \boldsymbol{L}^{(t-1)}[a \to b]\right\} \prod_{(a,b) \in E(\mathcal{T})} \boldsymbol{p}^{(t)}[a \to b] \\
&= \left(\frac{1}{S^{(t+1)}}\right)^{n-1} \exp\Big\{-\eta\Big(2\boldsymbol{\mathcal{L}}^{(t)}[\mathcal{T}] - \boldsymbol{\mathcal{L}}^{(t-1)}[\mathcal{T}]\Big)\Big\} \prod_{(a,b) \in E(\mathcal{T})} \boldsymbol{p}^{(t)}[a \to b], && (26)
\end{aligned}$$

where (25) uses (21) and the fact that every directed tree $\mathcal{T}$ has exactly $n-1$ edges, the second equality uses the multiplicative properties of the exponential, and (26) uses the definition of $\boldsymbol{\mathcal{L}}^{(t)}$ given in (24). From the inductive hypothesis, Condition 1 of Property 4.2 holds at time t; so,

continuing (26),

$$\prod_{(a,b)\in E(\mathcal{T})} \boldsymbol{p}^{(t+1)}[a\to b] = \left(\frac{1}{S^{(t+1)}}\right)^{n-1} \exp\Big\{-\eta\Big(2\boldsymbol{\mathcal{L}}^{(t)}[\mathcal{T}] - \boldsymbol{\mathcal{L}}^{(t-1)}[\mathcal{T}]\Big)\Big\} N^{(t)}\boldsymbol{X}^{(t)}[\mathcal{T}]$$
$$= \left(\frac{1}{S^{(t+1)}}\right)^{n-1} \mathcal{S}^{(t+1)} N^{(t)}\boldsymbol{X}^{(t+1)}[\mathcal{T}], \tag{27}$$

where (27) follows from (23) and the definition of $\mathcal{S}^{(t+1)}$. As a result, we have shown that Condition 1 of Property 4.2 holds at time $t+1$ for the parameter

$$N^{(t+1)} := \left(\frac{1}{S^{(t+1)}}\right)^{n-1} \mathcal{S}^{(t+1)} N^{(t)}. \tag{28}$$

We now show that Condition 2 of Property 4.2 holds at time $t+1$ as well. To do so, we analyze the (unique) stationary distribution of the matrix $\mathbf{M}^{(t+1)}$ defined on Line 3 of Algorithm 1. First, we start with the following simple observation.

Claim 4.3. *For any $j \neq k \in [\![n]\!]$, $\mathbf{M}^{(t+1)}[j,k] = \boldsymbol{p}^{(t+1)}[j\to k]$.*

Proof. For any $j \neq k \in [\![n]\!]$ the unique non-zero non-diagonal entry of the matrix $\mathbf{E}_{j\to k}$ appears as $\mathbf{E}_{j\to k}[k,j] = 1$ (recall (2)). Thus, the claim follows directly by the definition of the matrix $\mathbf{M}^{(t+1)}$ in Line 3 of Algorithm 1. □

In light of Claim 4.3, the j-th coordinate of the fixed point $\boldsymbol{x}^{(t+1)}$ of $\mathbf{M}^{(t+1)}$ can be expressed using the Markov chain tree theorem (Theorem 2.6) as

$$\boldsymbol{x}^{(t+1)}[j] = \frac{\sum_{\mathcal{T}\in\mathbb{T}_j^n} \prod_{(a,b)\in E(\mathcal{T})} \mathbf{M}^{(t+1)}[a,b]}{\sum_{j=1}^{n}\sum_{\mathcal{T}\in\mathbb{T}_j^n} \prod_{(a,b)\in E(\mathcal{T})} \mathbf{M}^{(t+1)}[a,b]} = \frac{\sum_{\mathcal{T}\in\mathbb{T}_j} \prod_{(a,b)\in E(\mathcal{T})} \boldsymbol{p}^{(t+1)}[a\to b]}{\sum_{j=1}^{n}\sum_{\mathcal{T}\in\mathbb{T}_j^n} \prod_{(a,b)\in E(\mathcal{T})} \boldsymbol{p}^{(t+1)}[a\to b]}. \tag{29}$$

Using (27) together with the fact that $\boldsymbol{X}^{(t+1)} \in \Delta^{(n^{n-1})}$, and therefore $\sum_{\mathcal{T}\in\mathbb{T}^n} \boldsymbol{X}^{(t+1)}[\mathcal{T}] = 1$, the denominator of (29) satisfies

$$\sum_{j=1}^{n}\sum_{\mathcal{T}\in\mathbb{T}_j^n} \prod_{(a,b)\in E(\mathcal{T})} \boldsymbol{p}^{(t+1)}[a\to b] = \sum_{\mathcal{T}\in\mathbb{T}^n} N^{(t+1)}\boldsymbol{X}^{(t+1)}[\mathcal{T}] = N^{(t+1)}. \tag{30}$$

Similarly, using Equation (27) together with (28), the numerator of (29) can be expressed as

$$\sum_{\mathcal{T}\in\mathbb{T}_j^n} \prod_{(a,b)\in E(\mathcal{T})} \boldsymbol{p}^{(t+1)}[a\to b] = N^{(t+1)} \sum_{\mathcal{T}\in\mathbb{T}_j^n} \boldsymbol{X}^{(t+1)}[\mathcal{T}]. \tag{31}$$

Plugging (30) and (31) into (29) we conclude that $\boldsymbol{x}^{(t+1)}[j] = \sum_{\mathcal{T}\in\mathbb{T}_j^n} \boldsymbol{X}^{(t+1)}[\mathcal{T}]$, which is exactly the j-th coordinate of the iterate produced by Algorithm 2 at time $t+1$ (Line 3). Thus, the strategies of Algorithms 1 and 2 at time $t+1$ are the same, completing the inductive proof. □

4.2 Bounding the Internal Regret

Here we explain how to leverage the high-order smoothnes techniques to bound the external regret of $\mathcal{R}_\Delta$ employed in Algorithm 1, which also bounds the internal regret of SL-OMWU (see Appendix E.1). To establish high-order smoothness, we use Theorem 4.1 to "lift" the analysis to the regret minimizer over the arborescences (Algorithm 2). Then, we leverage the particular structure of the losses in SL-OMWU to adapt Lemma 3.6, leading to the following guarantee.

Lemma 4.4. *Consider a parameter $\alpha \leq 1/(H+3)$. If all players employ* SL-OMWU *with learning rate $\eta \leq \frac{\alpha}{36e^5 m}$, then for any player $i \in [\![m]\!]$, $0 \leq h \leq H$ and $t \in [\![T-h]\!]$ it holds that*

$$\|(D_h \boldsymbol{L}_i)^{(t)}\|_\infty \leq \alpha^h h^{3h+1}.$$

Recall that $\boldsymbol{L}_i^{(t)}$ is the loss observed by algorithm $\mathcal{R}_\Delta$ at time t (see Algorithm 1). Armed with this crucial lemma, we are also able to extend the other technical ingredients used in the previous section, and we combine them to conclude the proof of Theorem 1.2, as we formally show in Appendix F.

Adversarial Bound. The learning algorithm can also be slightly modified in order to guarantee robustness when faced against adversarial losses. Indeed, the following corollary implies near-optimal internal regret in both regimes.

Corollary 4.5. *There exists a learning algorithm such that, if employed by all players, it guarantees that the internal regret of each player $i \in [\![m]\!]$ is bounded by $O(m \log n_i \log^4 T)$. Moreover, under adversarial losses the algorithm ensures internal regret bounded by $O(m \log n_i \log^4 T + \sqrt{\log n_i T})$.*

The idea is to use an adaptive choice of learning rate depending on whether the bound predicted in (158) has been violated in some repetition of the game.

5 Near-Optimal No-Snternal-Regret Dynamics via BM-OMWU

Proof of Theorem 1.4: Analyzing BM-OMWU. The techniques we described so far enable us to establish the near-optimal internal and swap regret bounds for SL-OMWU (Theorem 1.2 and Corollary 1.3), as well as the corresponding convergence to correlated equilibrium. However, different techniques are necessary to establish the regret bound of BM-OMWU (Theorem 1.4). At a high level, BM-OMWU runs n_i independent external regret minimizers for each player i, aggregates the outputs into a transition matrix of a Markov chain, and then computes its stationary distribution.

Thus, rather than arguing indirectly in terms of a supplementary external-regret minimizer, we *directly* analyze the high-order smoothness of a sequence of stationary distributions of Markov chains, at the cost of ultimately obtaining a worse dependence on the number of actions n_i in our swap regret bounds. Again using the high-order smoothness machinery (in particular, the boundedness chain rule of Lemma 3.7), doing so boils down to obtaining a bound on the Taylor series coefficients of the function that maps the entries of an ergodic matrix $\mathbf{Q}$ to $\mathbf{Q}$'s stationary distribution. Taken literally, such a bound is not quite possible, since the stationary distribution may have singularities around the non-ergodic matrices. However, we show that by using the Markov chain tree theorem together with the multi-dimensional version of Cauchy's integral formula, it is

possible to bound the Taylor series of the function mapping the *logarithms* of the entries of $\mathbf{Q}$ to its stationary distribution (see Lemma G.4). Leveraging the exponential-type structure of the OMWU updates, we then use this bound to obtain the desired guarantee on the high-order differences of the stationary distributions (Lemma 5.1).

In this section we give an overview of our analysis for the no-swap-regret algorithm of Blum and Mansour [BM07] (BM). Unlike the no-internal-regret algorithm of Stoltz and Lugosi [SL05], BM maintains n independent no-external-regret algorithms $\mathcal{R}_{\Delta,1}, \dots, \mathcal{R}_{\Delta,n}$, operating over Δ^n. For this reason, our previous approach appears to be no longer applicable. Instead, we develop more robust techniques that delve into the inner-workings of the high-order smoothness argument. In particular, we substantially generalize the approach by demonstrating how high-order smoothness bounds can be established even under the additional complexity of fixed point operations.

First, to keep the exposition reasonably self-contained, let us briefly recall the BM algorithm; a more detailed overview is included in Appendix E.2. At each time $t \geq 1$, every algorithm $\mathcal{R}_{\Delta,g}$ produces an iterate $\mathbf{Q}^{(t)}[g,\cdot] = (\mathbf{Q}^{(t)}[g,1], \dots, \mathbf{Q}^{(t)}[g,n]) \in \Delta^n$, for all $g \in [\![n]\!]$. Then, the algorithm computes a stationary distribution $\Delta^n \ni \boldsymbol{x}^{(t)} = (\mathbf{Q}^{(t)})^\top \boldsymbol{x}^{(t)}$ of the transition matrix $\mathbf{Q}^{(t)}$. Moreover, upon receiving a loss vector $\boldsymbol{\ell}^{(t)} \in \mathbb{R}^n$, the BM algorithm distributes the loss vector $\boldsymbol{x}^{(t)}[g]\boldsymbol{\ell}^{(t)}$ to each regret minimizer $\mathcal{R}_{\Delta,g}$ for $g \in [\![n]\!]$. In what follows, we will be concerned with the particular case where each regret minimizer is set to OMWU.

Our primary technical contribution in the analysis of BM-OMWU is to show that the losses observed by each individual regret minimizer exhibit high-order smoothness. Namely, we show the following lemma.

Lemma 5.1. *Fix a parameter $\alpha \in \left(0, \frac{1}{H+3}\right)$. There exists a sufficiently large universal constant C such that, if all players follow* BM-OMWU *with learning rate $\eta \leq \frac{1}{CH^2 n_i^3}$, then for any player $i \in [\![m]\!]$, integer h satisfying $0 \leq h \leq H$, time step $t \in [\![T-h]\!]$, and $g \in [\![n_i]\!]$, it holds that*

$$\| \left(\mathrm{D}_h \left(\boldsymbol{x}_i[g] \cdot \boldsymbol{\ell}_i\right)\right)^{(t)} \|_\infty \leq \alpha^h h^{3h+1}.$$

At a high level, the proof of this high-order smoothness lemma hinges on the cyclic relationship between the losses incurred by the players $\boldsymbol{\ell}_i^{(t)}$, the iterates produced by the copies of the OMWU algorithm $\mathbf{Q}_i^{(t)}$, and the final strategies $\boldsymbol{x}_i^{(t)}$ output by BM-OMWU. In particular, the iterates $\mathbf{Q}_i^{(t)}[g,\cdot]$ are determined based on the overall history of loss vectors $\sum_{t'<t} \boldsymbol{x}_i^{(t')}[g]\boldsymbol{\ell}_i^{(t')}$ weighted exponentially in terms of the *softmax* function; the strategies $\boldsymbol{x}_i^{(t)}$ are determined by applying the Markov Chain Tree Theorem (Theorem 2.6) to the stochastic matrix $\mathbf{Q}_i^{(t)}$ formed by the previous iterates; and, the losses $\boldsymbol{\ell}_i^{(t)}$ are determined as a function of the strategies of all the other players $\boldsymbol{x}_{-i}^{(t)}$. Therefore, using the "Boundedness Chain Rule for Finite Differences" (Lemma 3.7), we can demonstrate that bounds on the h-th order finite differences of $\boldsymbol{x}_i^{(t)}$ and $\boldsymbol{\ell}_i^{(t)}$ imply a bound on the $(h+1)$-th order finite differences of $\mathbf{Q}^{(t)}$.[9] This, in turn, implies a bound on the $(h+1)$-th order finite differences of $\boldsymbol{x}_i^{(t)}$, which then gives a bound on the $(h+1)$-th order finite differences of $\boldsymbol{\ell}_i^{(t)}$, as long as the Taylor coefficients of the softmax function and the Markov Chain Tree Theorem are sufficiently well-behaved. While this is not quite the case, we make a slight modification to this cycle of implication, bounding the finite differences of $\log \mathbf{Q}_i^{(t)}$ instead.

$$D_h\boldsymbol{x} \text{ bound}, D_h\boldsymbol{\ell} \text{ bound} \Rightarrow D_{h+1}\left(\log \mathbf{Q}\right) \text{ bound} \Rightarrow D_{h+1}\boldsymbol{x} \text{ bound} \Rightarrow D_{h+1}\boldsymbol{\ell} \text{ bound}.$$

[9] This is due to the fact that $D_{h+1}\mathbf{Q}^{(t)}[g,\cdot] = D_{h+1}\sum_{t'<t} \boldsymbol{x}_i^{(t')}[g]\boldsymbol{\ell}_i^{(t')} = D_h\boldsymbol{x}_i^{(t)}[g]\boldsymbol{\ell}_i^{(t)}$.

This cycle of implication enables an inductive argument that proves Lemma 5.1 as long as the log of the softmax function and an exponential version of the Markov Chain Tree Theorem exhibit bounded Taylor coefficients. These bounds are proved in Lemma G.7 and Lemma G.4 respectively. The proof of Lemma G.7 follows an explicit, combinatorial framework. On the other hand, Lemma G.4 introduces a novel technique for proving Taylor coefficient bounds, applying the Cauchy Integral Formula. This approach is far more general, and could help establish the necessary preconditions of Boundedness Chain Rule Lemma in much broader settings wherein combinatorial approaches are insufficient.

References

[Ana+22] Ioannis Anagnostides, Constantinos Daskalakis, Gabriele Farina, Maxwell Fishelson, Noah Golowich, and Tuomas Sandholm. "Near-optimal no-regret learning for correlated equilibria in multi-player general-sum games". In: *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*. 2022, pp. 736–749.

[AT89] V. Anantharam and P. Tsoucas. "A proof of the Markov chain tree theorem". In: *Statistics & Probability Letters* 8.2 (1989), pp. 189–192.

[Aum74] Robert Aumann. "Subjectivity and Correlation in Randomized Strategies". In: *Journal of Mathematical Economics* 1 (1974), pp. 67–96.

[Azi+21] Waïss Azizian, Franck Iutzeler, Jérome Malick, and Panayotis Mertikopoulos. "The Last-Iterate Convergence Rate of Optimistic Mirror Descent in Stochastic Variational Inequalities". In: *Conference on Learning Theory*. 2021, pp. 1–32.

[BCM12] Maria-Florina Balcan, Florin Constantin, and Ruta Mehta. "The Weighted Majority Algorithm does not Converge in Nearly Zero-sum Games". en. In: *ICML Workshop on Markets, Mechanisms, and Multi-Agent Models*. 2012.

[Bla54] David Blackwell. "Controlled Random Walks". In: *Proceedings of the International Congress of Mathematicians*. Vol. 3. 1954, pp. 336–338.

[BM07] Avrim Blum and Yishay Mansour. "From External to Internal Regret". In: *J. Mach. Learn. Res.* 8 (2007), pp. 1307–1324.

[BP18] James P. Bailey and Georgios Piliouras. "Multiplicative Weights Update in Zero-Sum Games". en. In: *Proceedings of the 2018 ACM Conference on Economics and Computation - EC '18*. Ithaca, NY, USA: ACM Press, 2018, pp. 321–338.

[BP19] James P. Bailey and Georgios Piliouras. "Fast and Furious Learning in Zero-Sum Games: Vanishing Regret with Non-Vanishing Step Sizes". In: *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, December 8-14, 2019, Vancouver, BC, Canada*. 2019, pp. 12977–12987.

[Bro49] George W Brown. *Some Notes on Computation of Games Solutions*. Tech. rep. RAND CORP SANTA MONICA CA, 1949.

[Cay89] Arthur Cayley. "A theorem on trees". In: *Quart. J. Math.* 23 (1889), pp. 376–378.

[CDT09] Xi Chen, Xiaotie Deng, and Shang-Hua Teng. "Settling the Complexity of Computing Two-Player Nash Equilibria". In: *Journal of the ACM* (2009).

[CL06a] Nicolo Cesa-Bianchi and Gábor Lugosi. *Prediction, learning, and games*. Cambridge university press, 2006.

[CL06b] Nicolo Cesa-Bianchi and Gabor Lugosi. *Prediction, learning, and games*. Cambridge University Press, 2006.

[CP19] Yun Kuen Cheung and Georgios Piliouras. "Vortices Instead of Equilibria in MinMax Optimization: Chaos and Butterfly Effects of Online Learning in Zero-Sum Games". In: *Proceedings of the Thirty-Second Conference on Learning Theory*. 2019, pp. 807–834.

[CP20] Xi Chen and Binghui Peng. "Hedging in games: Faster convergence of external and swap regrets". In: *Proceedings of the Annual Conference on Neural Information Processing Systems (NeurIPS)*. 2020.

[CS04] Imre Csiszár and Paul C. Shields. "Information Theory and Statistics: A Tutorial". In: *Commun. Inf. Theory* 1.4 (Dec. 2004), pp. 417–528.

[Das+10] Constantinos Daskalakis, Rafael Frongillo, Christos H. Papadimitriou, George Pierrakos, and Gregory Valiant. "On Learning Algorithms for Nash Equilibria". In: *Proceedings of the Third International Conference on Algorithmic Game Theory*. SAGT'10. Athens, Greece: Springer-Verlag, 2010, pp. 114–125.

[Das+18] Constantinos Daskalakis, Andrew Ilyas, Vasilis Syrgkanis, and Haoyang Zeng. "Training GANs with Optimism". In: *6th International Conference on Learning Representations, ICLR 2018*. OpenReview.net, 2018.

[DDK11] Constantinos Daskalakis, Alan Deckelbaum, and Anthony Kim. "Near-optimal no-regret algorithms for zero-sum games". In: *Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. 2011.

[DFG21] Constantinos Daskalakis, Maxwell Fishelson, and Noah Golowich. "Near-Optimal No-Regret Learning in General Games". In: *CoRR* abs/2108.06924 (2021).

[DGP06] Constantinos Daskalakis, Paul Goldberg, and Christos Papadimitriou. "The Complexity of Computing a Nash Equilibrium". In: *Proceedings of the Annual Symposium on Theory of Computing (STOC)*. 2006.

[DP14] Constantinos Daskalakis and Qinxuan Pan. "A counter-example to Karlin's strong conjecture for fictitious play". In: *Proceedings of the 55th Annual Symposium on Foundations of Computer Science (FOCS)*. 2014.

[DP18] Constantinos Daskalakis and Ioannis Panageas. "The Limit Points of (Optimistic) Gradient Descent in Min-Max Optimization". In: *NeurIPS 2018*. 2018, pp. 9256–9266.

[DP19] Constantinos Daskalakis and Ioannis Panageas. "Last-Iterate Convergence: Zero-Sum Games and Constrained Min-Max Optimization". In: *10th Innovations in Theoretical Computer Science Conference, ITCS 2019*. Ed. by Avrim Blum. Vol. 124. LIPIcs. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019, 27:1–27:18.

[Far+21] Gabriele Farina, Andrea Celli, Alberto Marchesi, and Nicola Gatti. "Simple Uncoupled No-Regret Learning Dynamics for Extensive-Form Correlated Equilibrium". In: *CoRR* abs/2104.01520 (2021).

[FKS19] Gabriele Farina, Christian Kroer, and Tuomas Sandholm. "Regret Circuits: Composability of Regret Minimizers". In: *International Conference on Machine Learning*. 2019, pp. 1863–1872.

[Fos+16] Dylan J Foster, Zhiyuan Li, Thodoris Lykouris, Karthik Sridharan, and Eva Tardos. "Learning in Games: Robustness of Fast Convergence". In: *Advances in Neural Information Processing Systems*. Vol. 29. Curran Associates, Inc., 2016.

[FV97] Dean Foster and Rakesh Vohra. "Calibrated Learning and Correlated Equilibrium". In: *Games and Economic Behavior* 21 (1997), pp. 40–55.

[GGM08] Geoffrey J Gordon, Amy Greenwald, and Casey Marks. "No-regret learning in convex games". In: *Proceedings of the 25th international conference on Machine learning*. ACM. 2008, pp. 360–367.

[GJ03] Amy Greenwald and Amir Jafari. "A General Class of No-Regret Learning Algorithms and Game-Theoretic Equilibria". In: *Conference on Learning Theory (COLT)*. Washington, D.C., 2003.

[GKP89] Ronald L. Graham, Donald E. Knuth, and Oren Patashnik. *Concrete Mathematics: A Foundation for Computer Science*. Reading: Addison-Wesley, 1989.

[GPD20] Noah Golowich, Sarath Pattathil, and Constantinos Daskalakis. "Tight last-iterate convergence rates for no-regret learning in multi-player games". In: *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, December 6-12, 2020, virtual*. 2020.

[HAM21] Yu-Guan Hsieh, Kimon Antonakopoulos, and Panayotis Mertikopoulos. "Adaptive Learning in Continuous Games: Optimal Regret Bounds and Convergence to Nash Equilibrium". In: *Conference on Learning Theory*. 2021.

[Han57] James Hannan. "Approximation to Bayes risk in repeated play". In: vol. III. Contributions to the Theory of Games. Princeton University Press, 1957, pp. 97–139.

[HM00] Sergiu Hart and Andreu Mas-Colell. "A Simple Adaptive Procedure Leading to Correlated Equilibrium". In: *Econometrica* 68 (2000), pp. 1127–1150.

[HM03] Sergiu Hart and Andreu Mas-Colell. "Uncoupled dynamics do not lead to Nash equilibrium". In: *American Economic Review* 93 (2003), pp. 1830–1836.

[Kan+18] Ehsan Asadi Kangarshahi, Ya-Ping Hsieh, Mehmet Fatih Sahin, and Volkan Cevher. "Let's be Honest: An Optimal No-Regret Framework for Zero-Sum Games". In: *Proceedings of the 35th International Conference on Machine Learning*. Vol. 80. Proceedings of Machine Learning Research. PMLR, Oct. 2018, pp. 2488–2496.

[KGW10] Alex Kruckman, Amy Greenwald, and John R. Wicks. *An elementary proof of the Markov chain tree theorem*. Tech. rep. 10-04. Brown University, 2010.

[KLP11] Robert Kleinberg, Katrina Ligett, and Georgios Piliouras. "Beyond the Nash equilibrium barrier". In: *In Innovations in Computer Science (ICS*. 2011, pp. 125–140.

[Lei+21] Qi Lei, Sai Ganesh Nagarajan, Ioannis Panageas, and xiao wang xiao. "Last iterate convergence in no-regret learning: constrained min-max optimization for convex-concave landscapes". In: *Proceedings of The 24th International Conference on Artificial Intelligence and Statistics*. Ed. by Arindam Banerjee and Kenji Fukumizu. Vol. 130. Proceedings of Machine Learning Research. PMLR, 13–15 Apr 2021, pp. 1441–1449.

[MPP18] Panayotis Mertikopoulos, Christos Papadimitriou, and Georgios Piliouras. "Cycles in Adversarial Regularized Learning". In: *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*. SODA '18. New Orleans, Louisiana: Society for Industrial and Applied Mathematics, 2018, pp. 2703–2717.

[Nes05] Yurii Nesterov. "Excessive Gap Technique in Nonsmooth Convex Minimization". In: *SIAM Journal of Optimization* 16.1 (2005).

[PP16] Christos Papadimitriou and Georgios Piliouras. "From Nash Equilibria to Chain Recurrent Sets: Solution Concepts and Topology". In: *Proceedings of the 2016 ACM Conference on Innovations in Theoretical Computer Science*. ITCS '16. Cambridge, Massachusetts, USA: Association for Computing Machinery, 2016, pp. 227–235.

[PR08] Christos H. Papadimitriou and Tim Roughgarden. "Computing correlated equilibria in multi-player games". In: *J. ACM* 55.3 (2008), 14:1–14:29.

[Rob51] Julia Robinson. "An iterative method of solving a game". In: *Annals of Mathematics* 54 (1951), pp. 296–301.

[Rou15] Tim Roughgarden. "Intrinsic Robustness of the Price of Anarchy". In: *J. ACM* 62.5 (2015), 32:1–32:42.

[RS13a] Alexander Rakhlin and Karthik Sridharan. "Online Learning with Predictable Sequences". In: *Conference on Learning Theory*. 2013, pp. 993–1019.

[RS13b] Alexander Rakhlin and Karthik Sridharan. "Optimization, learning, and games with predictable sequences". In: *Advances in Neural Information Processing Systems*. 2013, pp. 3066–3074.

[RST17] Tim Roughgarden, Vasilis Syrgkanis, and Éva Tardos. "The Price of Anarchy in Auctions". In: *J. Artif. Int. Res.* 59.1 (May 2017), pp. 59–101.

[Sha64] L. Shapley. "Some Topics in Two-Person Games". In: *Advances in Game theory* (1964).

[SL05] Gilles Stoltz and Gábor Lugosi. "Internal regret in on-line portfolio selection". In: *Machine Learning* 59.1-2 (2005), pp. 125–159.

[SL07] Gilles Stoltz and Gábor Lugosi. "Learning correlated equilibria in games with compact sets of strategies". In: *Games Econ. Behav.* 59.1 (2007), pp. 187–208.

[ST13] Vasilis Syrgkanis and Eva Tardos. "Composable and Efficient Mechanisms". In: *Proceedings of the Forty-Fifth Annual ACM Symposium on Theory of Computing*. STOC '13. Palo Alto, California, USA: Association for Computing Machinery, 2013, pp. 211–220.

[Syr+15] Vasilis Syrgkanis, Alekh Agarwal, Haipeng Luo, and Robert E Schapire. "Fast convergence of regularized learning in games". In: *Advances in Neural Information Processing Systems*. 2015, pp. 2989–2997.

[Vla+20] Emmanouil-Vasileios Vlatakis-Gkaragkounis, Lampros Flokas, Thanasis Lianeas, Panayotis Mertikopoulos, and Georgios Piliouras. "No-Regret Learning and Mixed Nash Equilibria: They Do Not Mix". In: *Advances in Neural Information Processing Systems*. Ed. by H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, and H. Lin. Vol. 33. Curran Associates, Inc., 2020, pp. 1380–1391.

[Wei+21] Chen-Yu Wei, Chung-Wei Lee, Mengxiao Zhang, and Haipeng Luo. "Linear Last-iterate Convergence in Constrained Saddle-point Optimization". In: 2021.

[WL18] Chen-Yu Wei and Haipeng Luo. "More Adaptive Algorithms for Adversarial Bandits". In: *Proceedings of the 31st Conference On Learning Theory*. 2018, pp. 1263–1291.

A Proofs for Section 3.1

In this section we prove Lemma 3.2. Throughout the section we use the notation of Lemma 3.2: in particular, we assume that any player $i \in [m]$ follows the OMWU updates (OMWU) with step size $\eta > 0$, for an arbitrary sequence of losses $\ell_i^{(1)}, \dots, \ell_i^{(T)}$.

A.1 Preliminary lemmas

The first few lemmas in this section pertain to vectors $P, Q \in \Delta^n$, for some $n \in \mathbb{N}$; note that such vectors P, Q may be viewed as distributions on $[n]$. Let $P/Q \in \mathbb{R}^n$ denote the Radon-Nikodym derivative, i.e., the vector whose jth component is $P(j)/Q(j)$.

Lemma A.1. *If $\|P/Q\|_\infty \le A$, then $\chi^2(P;Q) \le A \cdot \chi^2(Q;P)$.*

Proof. The lemma is immediate from the definition of the χ^2 divergence:

$$\chi^2(P;Q) = \sum_{j=1}^n \frac{(P(j)-Q(j))^2}{Q(j)} \le A \cdot \sum_{j=1}^n \frac{(P(j)-Q(j))^2}{P(j)} = A \cdot \chi^2(Q;P).$$

□

It is a standard fact (though one which we do not need in our proofs) that for all $P, Q \in \Delta^{n_i}$, $\mathrm{KL}(P;Q) \le \chi^2(P;Q)$. The below lemma shows an inequality in the opposite direction when $\|P/Q\|_\infty, \|Q/P\|_\infty$ are bounded:

Lemma A.2. *There is a constant C so that the following holds. Suppose that for $A \le \frac{3}{2}$ we have $\|P/Q\|_\infty \le A$ and $\|Q/P\|_\infty \le A$. Then $(1/2 - C(A-1)) \cdot \chi^2(P;Q) \le \mathrm{KL}(P;Q)$.*

Proof. There is a constant $C > 0$ so that for any $0 < \beta \le 1/2$, for all $|x| \le \beta$, we have

$$\log(1+x) \ge x - (1/2 + C\beta)x^2.$$

Set $a = A - 1$, so that $|P(j)/Q(j) - 1| \le a$ for all j by assumption. Then for $C' = C + 1/2$, we have

$$\begin{aligned}
\mathrm{KL}(P;Q) &= \sum_j P(j) \log \frac{P(j)}{Q(j)} \\
&\ge \sum_j P(j) \cdot \left(\left(\frac{P(j)}{Q(j)} - 1 \right) - (1/2 + Ca) \left(\frac{P(j)}{Q(j)} - 1 \right)^2 \right) \\
&\ge \chi^2(P;Q) - (1/2 + Ca) \sum_j P(j) \cdot \frac{(P(j)-Q(j))^2}{Q(j)^2} \\
&\ge \chi^2(P;Q) - \frac{1/2 + Ca}{A} \chi^2(P;Q) \\
&\ge \frac{1/2 - aC}{1+a} \cdot \chi^2(P;Q) \\
&\ge (1/2 - a \cdot (C + 1/2)) \cdot \chi^2(P;Q) \\
&= (1/2 - C'a) \cdot \chi^2(P;Q).
\end{aligned}$$

□

The next lemma considers two vectors $x, x' \in \Delta^n$ which are related by a multiplicative weights-style update with loss vector $w \in \mathbb{R}^n$; the lemma relates $\chi^2(x'; x)$ to $\|w\|_x^2$.

Lemma A.3. *There is a constant $C > 0$ so that the following holds. Suppose that $w \in \mathbb{R}^n$, $\alpha > 0$, $\|w\|_\infty \le \alpha/2 \le 1/C$, and $x, x' \in \Delta^n$ satisfy, for each $j \in [n]$,*

$$x'(j) = \frac{x(j) \cdot \exp(w(j))}{\sum_{k \in [n]} x(k) \cdot \exp(w(k))}. \tag{32}$$

Then

$$(1 - C\alpha) \cdot \mathrm{Var}_x(w) \le \chi^2(x'; x) \le (1 + C\alpha) \mathrm{Var}_x(w).$$

Proof. Let $w' = w - \langle x, w \rangle \mathbf{1}$, where $\mathbf{1}$ denotes the all-1s vector. Note that $\mathrm{Var}_x(w) = \mathrm{Var}_x(w')$, and that if we replace w with w', (32) remains true. Moreover, $\|w'\|_\infty \le 2\|w\|_\infty \le \alpha$. Thus, by replacing w with w', we may assume from here on that $\langle w, x \rangle = 0$ and that $\|w\| \le \alpha$.

Note that

$$\chi^2(x'; x) = -1 + \sum_{i=1}^n x(i) \cdot (x'(i)/x(i))^2 = -1 + \mathbb{E}\left(\frac{\exp(W)}{\mathbb{E}\exp(W)}\right)^2,$$

where W is a random variable that takes values $w(j)$ with probability $x(j)$. As long as C is a sufficiently large constant, we have that, for all z satisfying $|z| \le \alpha$,

$$1 + z + (1 - C\alpha)z^2/2 \le \exp(z) \le 1 + z + (1 + C\alpha)z^2/2. \tag{33}$$

Thus, for a sufficiently large constant C_0', we have, for all z satisfying $|z| \le \alpha$,

$$1 + 2z + (2 - C_0'\alpha)z^2 \le \exp(z)^2 \le 1 + 2z + (2 + C_0'\alpha)z^2. \tag{34}$$

Moreover, since $\mathbb{E}W = 0$, we have from (33) that $1 + (1 - C\alpha)\mathbb{E}W^2/2 \le \mathbb{E}\exp(W) \le 1 + (1 + C\alpha)\mathbb{E}W^2/2$. For a sufficiently large constant C_1' it follows that

$$1 + (1 - C_1'\alpha)\mathbb{E}W^2 \le (\mathbb{E}\exp(W))^2 \le 1 + (1 + C_1'\alpha)\mathbb{E}W^2. \tag{35}$$

Combining (34) and (35) and again using the fact that $\mathbb{E}W = 0$, we get, for some sufficiently large constant C'', as long as $\alpha < 1/C_1'$,

$$\begin{aligned}
(1 - C''\alpha)\mathbb{E}W^2 &\le -1 + \frac{1 + (2 - C_0'\alpha)\mathbb{E}W^2}{1 + (1 + C_1'\alpha)\mathbb{E}W^2} \\
&\le -1 + \frac{\mathbb{E}(\exp(W)^2)}{(\mathbb{E}\exp(W))^2} \\
&\le -1 + \frac{1 + (2 + C_0'\alpha)\mathbb{E}W^2}{1 + (1 - C_1'\alpha)\mathbb{E}W^2} \\
&\le (1 + C''\alpha)\mathbb{E}W^2.
\end{aligned}$$

By the assumption that $\langle w, x \rangle = 0$, we have $\mathbb{E}W^2 = \mathrm{Var}_x(w)$, and thus the above gives the desired result. □

We will need the following standard lemma:

Lemma A.4 ([RS13a], Eq. (26)). *For any $n \in \mathbb{N}$, $\ell \in \mathbb{R}^n$, $y \in \Delta^n$, if it holds that $x = \arg\min_{x' \in \Delta^n} \langle x', \ell \rangle + \mathrm{KL}(x'; y)$, then for any $z \in \Delta^n$,*

$$\langle x - z, \ell \rangle \leq \mathrm{KL}(z; y) - \mathrm{KL}(z; x) - \mathrm{KL}(x; y).$$

For $t \in [T]$, we define the vector $\tilde{x}_i^{(t)} \in \Delta^{n_i}$ by

$$\tilde{x}_i^{(t)}(j) := \frac{x_i^{(t)}(j) \cdot \exp(-\eta \cdot (\ell_i^{(t)}(j) - \ell_i^{(t-1)}(j)))}{\sum_{k \in [n_i]} x_i^{(t)}(k) \cdot \exp(-\eta \cdot (\ell_i^{(t)}(k) - \ell_i^{(t-1)}(k)))}. \tag{36}$$

Additionally define $\tilde{x}_i^{(0)} := (1/n_i, \ldots, 1/n_i)$ to be the uniform distribution over $[n_i]$.

The next lemma, Lemma A.5 is very similar to [RS13a, Lemma 3], and is indeed essentially shown in the course of the proof of that lemma. Note that no boundedness assumption is placed on the vectors $\ell_i^{(t)}$ in Lemma A.5. For completeness we provide a full proof of the lemma.

Lemma A.5 (Refinement of Lemma 3, [RS13a]). *Suppose that any player $i \in [m]$ follows the OMWU updates (OMWU) with step size $\eta > 0$, for an arbitrary sequence of losses $\ell_i^{(1)}, \ldots, \ell_i^{(T)} \in \mathbb{R}^{n_i}$. For any vector $x^\star \in \Delta^{n_i}$, it holds that*

$$\begin{aligned}\sum_{t=1}^{T} \langle x_i^{(t)} - x^\star, \ell_i^{(t)} \rangle \leq \frac{\log n_i}{\eta} + \sum_{t=1}^{T} \left\| x_i^{(t)} - \tilde{x}_i^{(t)} \right\|_{x_i^{(t)}}^\star \sqrt{\mathrm{Var}_{x_i^{(t)}} \left(\ell_i^{(t)} - \ell_i^{(t-1)} \right)} \\ - \frac{1}{\eta} \sum_{t=1}^{T} \mathrm{KL}(\tilde{x}_i^{(t)}; x_i^{(t)}) - \frac{1}{\eta} \sum_{t=1}^{T} \mathrm{KL}(x_i^{(t)}; \tilde{x}_i^{(t-1)}).\end{aligned} \tag{37}$$

Proof. For any $x^\star \in \Delta^{n_i}$, it holds that

$$\langle x_i^{(t)} - x^\star, \ell_i^{(t)} \rangle = \langle x_i^{(t)} - \tilde{x}_i^{(t)}, \ell_i^{(t)} - \ell_i^{(t-1)} \rangle + \langle x_i^{(t)} - \tilde{x}_i^{(t)}, \ell_i^{(t-1)} \rangle + \langle \tilde{x}_i^{(t)} - x^\star, \ell_i^{(t)} \rangle. \tag{38}$$

For $t \in [T]$, set $c^{(t)} = \langle x_i^{(t)}, \ell_i^{(t)} - \ell_i^{(t-1)} \rangle$. Using the definition of the dual norm and the fact $\langle x_i^{(t)} - \tilde{x}_i^{(t)}, \mathbf{1} \rangle = 0$, we have

$$\begin{aligned}\langle x_i^{(t)} - \tilde{x}_i^{(t)}, \ell_i^{(t)} - \ell_i^{(t-1)} \rangle &= \langle x_i^{(t)} - \tilde{x}_i^{(t)}, \ell_i^{(t)} - \ell_i^{(t-1)} - c^{(t)} \mathbf{1} \rangle \\ &\leq \left\| x_i^{(t)} - \tilde{x}_i^{(t)} \right\|_{x_i^{(t)}}^\star \cdot \left\| \ell_i^{(t)} - \ell_i^{(t-1)} - c^{(t)} \mathbf{1} \right\|_{x_i^{(t)}} \\ &\leq \left\| x_i^{(t)} - \tilde{x}_i^{(t)} \right\|_{x_i^{(t)}}^\star \cdot \sqrt{\mathrm{Var}_{x_i^{(t)}} \left(\ell_i^{(t)} - \ell_i^{(t-1)} \right)}.\end{aligned} \tag{39}$$

It is immediate from the definitions of $\tilde{x}_i^{(t)}$ (in (36)) and $x_i^{(t)}$ (in (OMWU)) that for $j \in [n_i]$,

$$x_i^{(t)}(j) = \frac{\tilde{x}_i^{(t-1)}(j) \cdot \exp(-\eta \cdot \ell_i^{(t-1)}(j))}{\sum_{k \in [n_i]} \tilde{x}_i^{(t-1)}(k) \cdot \exp(-\eta \cdot \ell_i^{(t-1)}(k))} = \arg\min_{x \in \Delta^{n_i}} \left\langle x, \eta \cdot \ell_i^{(t-1)} \right\rangle + \mathrm{KL}(x; \tilde{x}_i^{(t-1)}) \tag{40}$$

Using Lemma A.4 with $x = x_i^{(t)}, \ell = \eta \ell_i^{(t-1)}, y = \tilde{x}_i^{(t-1)}, z = \tilde{x}_i^{(t)}$, we obtain

$$\langle x_i^{(t)} - \tilde{x}_i^{(t)}, \ell_i^{(t-1)} \rangle \leq \frac{1}{\eta} \mathrm{KL}(\tilde{x}_i^{(t)}; \tilde{x}_i^{(t-1)}) - \frac{1}{\eta} \mathrm{KL}(\tilde{x}_i^{(t)}; x_i^{(t)}) - \frac{1}{\eta} \mathrm{KL}(x_i^{(t)}; \tilde{x}_i^{(t-1)}). \tag{41}$$

Next, we note that, again by (36) and (OMWU), for $j \in [n_i]$,

$$\tilde{x}_i^{(t)}(j) = \frac{\tilde{x}_i^{(t-1)}(j) \cdot \exp(-\eta \cdot \ell_i^{(t)}(j))}{\sum_{k\in[n_i]} \tilde{x}_i^{(t-1)}(k) \cdot \exp(-\eta \cdot \ell_i^{(t)}(k))} = \arg\min_{x\in\Delta^{n_i}} \left\langle x, \eta \cdot \ell_i^{(t)} \right\rangle + \mathrm{KL}(x; \tilde{x}_i^{(t-1)}).$$

Using Lemma A.4 with $x = \tilde{x}_i^{(t)}, \ell = \eta\ell_i^{(t)}, y = \tilde{x}_i^{(t-1)}, z = x^\star$, we obtain

$$\langle \tilde{x}_i^{(t)} - x^\star, \ell_i^{(t)} \rangle \le \frac{1}{\eta}\mathrm{KL}(x^\star; \tilde{x}_i^{(t-1)}) - \frac{1}{\eta}\mathrm{KL}(x^\star; \tilde{x}_i^{(t)}) - \frac{1}{\eta}\mathrm{KL}(\tilde{x}_i^{(t)}; \tilde{x}_i^{(t-1)}). \tag{42}$$

By (38), (39), (41), and (42), we have

$$\begin{aligned}
\langle x_i^{(t)} - x^\star, \ell_i^{(t)} \rangle &\le \left\| x_i^{(t)} - \tilde{x}_i^{(t)} \right\|^\star_{x_i^{(t)}} \cdot \sqrt{\mathrm{Var}_{x_i^{(t)}}\left(\ell_i^{(t)} - \ell_i^{(t-1)}\right)} \\
&\quad + \frac{1}{\eta}\mathrm{KL}(\tilde{x}_i^{(t)}; \tilde{x}_i^{(t-1)}) - \frac{1}{\eta}\mathrm{KL}(\tilde{x}_i^{(t)}; x_i^{(t)}) - \frac{1}{\eta}\mathrm{KL}(x_i^{(t)}; \tilde{x}_i^{(t-1)}) \\
&\quad + \frac{1}{\eta}\mathrm{KL}(x^\star; \tilde{x}_i^{(t-1)}) - \frac{1}{\eta}\mathrm{KL}(x^\star; \tilde{x}_i^{(t)}) - \frac{1}{\eta}\mathrm{KL}(\tilde{x}_i^{(t)}; \tilde{x}_i^{(t-1)}) \\
&= \left\| x_i^{(t)} - \tilde{x}_i^{(t)} \right\|^\star_{x_i^{(t)}} \cdot \sqrt{\mathrm{Var}_{x_i^{(t)}}\left(\ell_i^{(t)} - \ell_i^{(t-1)}\right)} + \frac{1}{\eta}\mathrm{KL}(x^\star; \tilde{x}_i^{(t-1)}) - \frac{1}{\eta}\mathrm{KL}(x^\star; \tilde{x}_i^{(t)}) \\
&\quad - \frac{1}{\eta}\mathrm{KL}(\tilde{x}_i^{(t)}; x_i^{(t)}) - \frac{1}{\eta}\mathrm{KL}(x_i^{(t)}; \tilde{x}_i^{(t-1)}).
\end{aligned} \tag{43}$$

The statement of the lemma follows by summing (43) over $t \in [T]$ and using the fact that for any choice of $x^\star$, $\mathrm{KL}(x^\star; x_i^{(0)}) \le \log n_i$. □

A.2 Proof of Lemma 3.2

Now we are ready to prove Lemma 3.2. For convenience we restate the lemma.

Lemma 3.2 [Restated] *There is a constant $C > 0$ so that the following holds. Suppose any player $i \in [m]$ follows the OMWU updates (OMWU) with step size $0 < \eta < 1/C$, for an arbitrary sequence of losses $\ell_i^{(1)}, \dots, \ell_i^{(T)} \in [0,1]^{n_i}$. Then for any vector $x^\star \in \Delta^{n_i}$, it holds that*

$$\sum_{t=1}^{T}\langle x_i^{(t)} - x^\star, \ell_i^{(t)} \rangle \le \frac{\log n_i}{\eta} + \sum_{t=1}^{T}\left(\frac{\eta}{2} + C\eta^2\right)\mathrm{Var}_{x_i^{(t)}}\left(\ell_i^{(t)} - \ell_i^{(t-1)}\right) - \sum_{t=1}^{T}\frac{(1-C\eta)\eta}{2}\cdot\mathrm{Var}_{x_i^{(t)}}\left(\ell_i^{(t-1)}\right). \tag{44}$$

Proof. Lemma A.5 gives that, for any $x^\star \in \Delta^{n_i}$,

$$\begin{aligned}
\sum_{t=1}^{T}\langle x_i^{(t)} - x^\star, \ell_i^{(t)} \rangle &\le \frac{\log n_i}{\eta} + \sum_{t=1}^{T}\left\| x_i^{(t)} - \tilde{x}_i^{(t)} \right\|^\star_{x_i^{(t)}} \sqrt{\mathrm{Var}_{x_i^{(t)}}\left(\ell_i^{(t)} - \ell_i^{(t-1)}\right)} \\
&\quad - \frac{1}{\eta}\sum_{t=1}^{T}\mathrm{KL}(\tilde{x}_i^{(t)}; x_i^{(t)}) - \frac{1}{\eta}\sum_{t=1}^{T}\mathrm{KL}(x_i^{(t)}; \tilde{x}_i^{(t-1)}).
\end{aligned} \tag{45}$$

Note that for any vectors $x, x' \in \Delta^{n_i}$, if there is a vector $\ell \in \mathbb{R}^{n_i}$ so that for all $j \in [n_i]$, $x'(j) = \frac{x(j)\cdot\exp(\eta\cdot\ell(j))}{\sum_k x(j)\cdot\exp(\eta\cdot\ell(k))}$, we have that

$$\exp(-2\eta\|\ell\|_\infty) \le \left\|\frac{x'}{x}\right\|_\infty \le \exp(2\eta\|\ell\|_\infty).$$

Therefore, by (36) and (40), respectively, we obtain that, for $\eta \le 1/4$,

$$\begin{aligned}\exp(-2\eta\|\ell_i^{(t)} - \ell_i^{(t-1)}\|_\infty) \le \left\|\frac{\tilde{x}_i^{(t)}}{x_i^{(t)}}\right\|_\infty \le \exp(2\eta\|\ell_i^{(t)} - \ell_i^{(t-1)}\|_\infty) \le \exp(4\eta) \le 1 + 8\eta \\ \exp(-2\eta\|\ell_i^{(t-1)}\|_\infty) \le \left\|\frac{x_i^{(t)}}{\tilde{x}_i^{(t-1)}}\right\|_\infty \le \exp(2\eta\|\ell_i^{(t-1)}\|_\infty) \le \exp(2\eta) \le 1 + 4\eta. \end{aligned} \tag{46}$$

(Above we have also used that $\|\ell_i^{(t)}\|_\infty \le 1$ for all t.) Thus, for $\eta \le \frac{1}{16}$, we can apply Lemma A.2 and show, for a sufficiently large constant C_0,

$$\mathrm{KL}(\tilde{x}_i^{(t)}; x_i^{(t)}) \ge \chi^2(\tilde{x}_i^{(t)}; x_i^{(t)}) \cdot (1/2 - C_0\eta) \tag{47}$$

$$\mathrm{KL}(x_i^{(t)}; \tilde{x}_i^{(t-1)}) \ge \chi^2(x_i^{(t)}; \tilde{x}_i^{(t-1)}) \cdot (1/2 - C_0\eta). \tag{48}$$

Note also that for vectors x, y we have that $\chi^2(x; y) = \left(\|x - y\|_y^\star\right)^2$. By Lemma A.3 and (36), we have that, for a sufficiently large constant C_1, as long as $\eta \le 1/C_1$,

$$\left(\left\|x_i^{(t)} - \tilde{x}_i^{(t)}\right\|_{x_i^{(t)}}^\star\right)^2 = \chi^2(\tilde{x}_i^{(t)}; x_i^{(t)}) \le (1 + C_1\eta)\eta^2 \cdot \mathrm{Var}_{x_i^{(t)}}\left(\ell_i^{(t)} - \ell_i^{(t-1)}\right) \tag{49}$$

and

$$\chi^2(\tilde{x}_i^{(t)}; x_i^{(t)}) \ge (1 - C_1\eta)\eta^2 \cdot \mathrm{Var}_{x_i^{(t)}}\left(\ell_i^{(t)} - \ell_i^{(t-1)}\right). \tag{50}$$

Next we lower bound $\chi^2(x_i^{(t)}; \tilde{x}_i^{(t-1)})$ as follows, where C_2 denotes a sufficiently large constant: as long as $\eta \le 1/C_2$,

$$\begin{aligned}\chi^2(x_i^{(t)}; \tilde{x}_i^{(t-1)}) &\ge \chi^2(\tilde{x}_i^{(t-1)}; x_i^{(t)}) \cdot \exp(-2\eta) &(51)\\ &\ge (1 - C_2\eta)\eta^2 \cdot \mathrm{Var}_{x_i^{(t)}}\left(\ell_i^{(t-1)}\right), &(52)\end{aligned}$$

where (51) follows from Lemma A.1 and (46), and (52) follows from Lemma A.3 and (40).

Combining (45), (47), (48), (49), (50), and (52) gives that for a sufficiently large constant C, as long as $\eta < 1/C$,

$$\sum_{t=1}^T \langle x_i^{(t)} - x^\star, \ell_i^{(t)}\rangle \le \frac{\log n_i}{\eta} + \sum_{t=1}^T (\eta/2 + C\eta^2) \cdot \mathrm{Var}_{x_i^{(t)}}\left(\ell_i^{(t)} - \ell_i^{(t-1)}\right) - \frac{(1 - C\eta)\eta}{2} \cdot \mathrm{Var}_{x_i^{(t)}}\left(\ell_i^{(t-1)}\right),$$

as desired. □

B Proofs for Section 3.3

In this section we give the full proof of Lemma 3.6. In Section B.1 we introduce some preliminaries. In Section B.2 we prove Lemma 3.7, the "boundedness chain rule" for finite differences. In Section B.4 we show how to use this lemma to prove Lemma 3.6.

B.1 Additional preliminaries

In this section we introduce some additional notations and basic combinatorial lemmas. Definition B.1 introduces the *shift operator* E_s, which like the finite difference operator D_h, maps one sequence to another sequence.

Definition B.1 (Shift operator)**.** Suppose $L = (L^{(1)}, \dots, L^{(T)})$ is a sequence of vectors $L^{(t)} \in \mathbb{R}^n$. For integers $s \geq 0$, the *s-shift sequence* for the sequence L, denoted by $\mathrm{E}_s L$, is the sequence $\mathrm{E}_s L = ((\mathrm{E}_s L)^{(1)}, \dots, (\mathrm{E}_s L)^{(T-s)})$, defined by $(\mathrm{E}_s L)^{(t)} = L^{(t+s)}$ for $1 \leq t \leq T - s$.

For sequences $L = (L^{(1)}, \dots, L^{(T)})$ and $K = (K^{(1)}, \dots, K^{(T)})$ of real numbers, we will denote the *product sequence* as $L \cdot K$ as the sequence of vectors $L \cdot K := (L^{(1)}K^{(1)}, \dots, L^{(T)}K^{(T)})$. Lemmas B.2 and B.3 below are standard analogues of the product rule for finite differences. The (straightforward) proofs are provided for completeness.

Lemma B.2 (Product rule; Eq. (2.55) of [GKP89])**.** *Suppose $L = (L^{(1)}, \dots, L^{(T)})$ and $K = (K^{(1)}, \dots, K^{(T)})$ are sequences of real numbers. Then the product sequence $L \cdot K$ satisfies*

$$\mathrm{D}_1 (L \cdot K) = L \cdot \mathrm{D}_1 K + \mathrm{D}_1 L \cdot \mathrm{E}_1 K.$$

Proof. We compute

$$\begin{aligned} \mathrm{D}_1 (L \cdot K)^{(t)} &= L^{(t+1)}K^{(t+1)} - L^{(t)}K^{(t)} \\ &= L^{(t+1)}K^{(t+1)} - L^{(t)}K^{(t+1)} + L^{(t)}K^{(t+1)} - L^{(t)}K^{(t)} \\ &= (L \cdot \mathrm{D}_1 K + \mathrm{D}_1 L \cdot \mathrm{E}_1 K)^{(t)}. \end{aligned}$$

□

Lemma B.3 (Multivariate product rule)**.** *Suppose that $m \in \mathbb{N}$ and for $1 \leq i \leq m$, $L_i = (L_i^{(1)}, \dots, L_i^{(T)})$ are sequences of real numbers. Then the product sequence $\prod_{i=1}^m L_i$ satisfies*

$$\mathrm{D}_1 \prod_{i=1}^m L_i = \sum_{i=1}^m \left(\prod_{i'<i} L_{i'} \right) \cdot \mathrm{D}_1 L_i \cdot \left(\prod_{i'>i} \mathrm{E}_1 L_{i'} \right).$$

Proof. We compute

$$\begin{aligned}\left(\mathrm{D}_1 \prod_{i=1}^m L_i\right)^{(t)} &= \prod_{i=1}^m L_i^{(t+1)} - \prod_{i=1}^m L_i^{(t)} \\ &= \sum_{i=1}^m \left(\prod_{i'\le i} L_{i'}^{(t+1)} \prod_{i'>i} L_{i'}^{(t)} - \prod_{i'<i} L_{i'}^{(t+1)} \prod_{i'\ge i} L_{i'}^{(t)}\right) \\ &= \sum_{i=1}^m \left(\prod_{i'<i} L_{i'}^{(t+1)} \cdot \prod_{i'>i} L_{i'}^{(t)} \cdot \left(L_i^{(t+1)} - L_i^{(t)}\right)\right) \\ &= \left(\sum_{i=1}^m \left(\prod_{i'<i} L_{i'}\right) \cdot \mathrm{D}_1 L_i \cdot \left(\prod_{i'>i} \mathrm{E}_1 L_{i'}\right)\right)^{(t)}.\end{aligned}$$

□

Lemma B.5 and Lemma B.4, which is used in the proof of the former, are used to bound certain sums with many terms in the proof of Lemma 3.7. To state Lemma B.4 we make one definition. For positive integers k, m and any $h, C > 0$, define

$$R_{h,m,k,C} = \sum_{0\le n_1,\cdots,n_k\le m} \left(\frac{\prod_{i=1}^k n_i^{n_i}}{h^{\sum_{i=1}^k n_i}}\right)^C,$$

where the sum is over integers $n_1, \ldots, n_k$ satisfying $0 \le n_i \le m$ for $i \in [k]$. In the definition of $R_{h,m,k,C}$, the quantity 0^0 (which arises when some $n_i = 0$) is interpreted as 1.

Lemma B.4. *For any positive integers k, m and any $h, C > 0$ so that $m \le h/2$, $C \ge 2$, and $h \ge 8$, then*

$$R_{h,m,k,C} \le \exp\left(\frac{2k}{h^C}\right).$$

Proof of Lemma B.4. We may rewrite $R_{h,m,k,C}$ and then upper bound it as follows:

$$\begin{aligned}R_{h,m,k,C} &= \left(\sum_{j=0}^m \left(\frac{j}{h}\right)^{Cj}\right)^k \\ &\le \left(1 + \left(\frac{1}{h}\right)^C + (m-1)\max\left(\left(\frac{2}{h}\right)^{2C}, \left(\frac{m}{h}\right)^{mC}\right)\right)^k \qquad (53) \\ &\le \left(1 + \left(\frac{1}{h}\right)^C + (h/2)\max\left(\left(\frac{2}{h}\right)^{2C}, \left(\frac{1}{2}\right)^{hC/2}\right)\right)^k\end{aligned}$$

where (53) follows since $\left(\frac{i}{h}\right)^{Ci}$ is convex in i for $i \ge 0$, and therefore, in the interval $[2, m] \subseteq [2, h/2]$, takes on maximal values at the endpoints. We see

$$(h/2)\left(\frac{2}{h}\right)^{2C} = \left(\frac{2}{h}\right)^{2C-1} \le \left(\frac{1}{h}\right)^C$$

for $h \geq 8$ when $C \geq 2$. Also,

$$(h/2)\left(\frac{1}{2}\right)^{hC/2} \leq \left(\frac{1}{h}\right)^{C}$$

for $h \geq 8$ when $C \geq 2$. (This inequality is easily seen to be equivalent to the fact that $(C+1)\log h - \frac{Ch}{2} \leq 1$, which follows from the fact that $\log h - h/2 \leq 0$ for $h \geq 8$ and $3\log h - h \leq 1$ for $h \geq 8$.) Therefore,

$$\begin{aligned} R_{h,m,k,C} &\leq \left(1 + \left(\frac{1}{h}\right)^{C} + (h/2)\max\left(\left(\frac{2}{h}\right)^{2C}, \left(\frac{1}{2}\right)^{hC/2}\right)\right)^{k} \\ &\leq \left(1 + 2\left(\frac{1}{h}\right)^{C}\right)^{k} \\ &\leq \exp\left(\frac{2k}{h^{C}}\right). \end{aligned}$$

□

Lemma B.5. *Fix integers $h \geq 0, k \geq 1$. For any function $\pi : [h] \to [k]$, define, for each $i \in [k]$, $h_i(\pi) = |\{q \in [h] | \pi(q) = i\}|$. Then, for any $C \geq 3$,*

$$\sum_{\pi:[h]\to[k]} \frac{\prod_{i=1}^{k} h_i(\pi)^{Ch_i(\pi)}}{h^{Ch}} \leq \max\left\{k^7, (hk+1)\cdot\exp\left(\frac{2k}{h^{C-1}}\right)\right\}. \tag{54}$$

Proof. In the case that $h \leq 7$, we simply use the fact that the number of functions $\pi : [h] \to [k]$ is $k^h \leq k^7$, and each term of the summation on the left-hand side of (54) is at most 1. In the remainder of the proof we may thus assume that $h \geq 8$.

For any tuple $(h_1, \cdots, h_k)$ of non-negative integers with $\sum_{i=1}^{k} h_i = h$, there are $\binom{h}{h_1, h_2, \cdots, h_k} \leq \frac{h^h}{\prod_i h_i^{h_i}}$ (see [CS04, Lemma 2.2] for a proof of this inequality) functions $\pi : [h] \to [k]$ such that $h_i(\pi) = h_i$ for all $i \in [k]$. Combining these like terms,

$$\begin{aligned} \sum_{\pi:[h]\to[k]} \frac{\prod_i h_i(\pi)^{Ch_i(\pi)}}{h^{Ch}} &\leq \sum_{\substack{h_1,\cdots,h_k \geq 0 \\ \sum h_i = h}} \frac{h^h}{\prod_i h_i^{h_i}} \cdot \left(\frac{\prod_i h_i^{h_i}}{h^h}\right)^{C} \\ &\leq \sum_{\substack{h_1,\cdots,h_k \geq 0 \\ \sum h_i = h}} \left(\frac{\prod_i h_i^{h_i}}{h^h}\right)^{C-1}. \end{aligned} \tag{55}$$

We evaluate this sum in 2 cases: whether or not $h_{\max} := \max_i\{h_i\}$ is greater than $h/2$. The

contribution to this sum coming from terms with $h_{\max} \le h/2$ is

$$\sum_{\substack{h_1,\cdots,h_k\ge 0\\ h_1,\cdots,h_k\le h/2\\ \sum h_i = h}} \left(\frac{\prod_i h_i^{h_i}}{h^h}\right)^{C-1} \le \sum_{\substack{h_1,\cdots,h_k\ge 0\\ h_1,\cdots,h_k\le h/2}} \left(\frac{\prod_i h_i^{h_i}}{h^{\sum h_i}}\right)^{C-1}$$

$$= R_{h,\lfloor h/2\rfloor,k,C-1}$$

$$\le \exp\left(\frac{2k}{h^{C-1}}\right), \tag{56}$$

by Lemma B.4.

We next consider the case where $h_{\max} > h/2$. For a specific term $(h_1,\cdots,h_k)$ with $\max_i\{h_i\} > h/2$, we know there is a unique $M \in [k]$ such that $h_M = \max_i\{h_i\}$ since $\sum_{i=1}^k h_i = h$. So, we can represent the contribution to the sum from this case as

$$\sum_{M=1}^{k} \sum_{\substack{h_1,\cdots,h_k\ge 0\\ h_M > h/2\\ \sum h_i = h}} \left(\frac{\prod_i h_i^{h_i}}{h^h}\right)^{C-1} = k \sum_{\substack{h_1,\cdots,h_k\ge 0\\ h_k > h/2\\ \sum h_i = h}} \left(\frac{\prod_i h_i^{h_i}}{h^h}\right)^{C-1} \tag{57}$$

$$\le k \sum_{d=0}^{\lfloor h/2\rfloor} \left(\frac{(h-d)^{h-d}}{h^{h-d}}\right)^{C-1} \sum_{\substack{h_1,\cdots,h_{k-1}\ge 0\\ \sum h_i = d}} \left(\frac{\prod_i h_i^{h_i}}{h^d}\right)^{C-1} \tag{58}$$

$$\le k \sum_{d=0}^{\lfloor h/2\rfloor} \sum_{\substack{h_1,\cdots,h_{k-1}\ge 0\\ h_1,\cdots,h_{k-1}\le d}} \left(\frac{\prod_i h_i^{h_i}}{h^{\sum h_i}}\right)^{C-1}$$

$$= k \sum_{d=0}^{\lfloor h/2\rfloor} R_{h,d,k-1,C-1}$$

$$\le kh \cdot \exp\left(\frac{2k}{h^{C-1}}\right), \tag{59}$$

where (57) follows by symmetry, (58) follows by factoring out the contribution of $\left(\frac{h_k^{h_k}}{h^{h_k}}\right)^C$ and letting $d = h - h_k$, and (59) follows by Lemma B.4.

The statement of the lemma follows from (55), (56), and (59). □

Lemma B.6. *For $n \in \mathbb{N}$, let $\xi_1,\ldots,\xi_n \ge 0$ such that $\xi_1 + \cdots + \xi_n = 1$. For each $j \in [n]$, define $\phi_j : \mathbb{R}^n \to \mathbb{R}$ to be the function*

$$\phi_j((z_1,\ldots,z_n)) = \frac{\xi_j \exp(z_j)}{\sum_{k=1}^n \xi_k \cdot \exp(z_k)}$$

and let $P_{\phi_j}(z) = \sum_{\gamma\in\mathbb{Z}^n_{\ge 0}} a_{j,\gamma}\cdot z^\gamma$ denote the Taylor series of ϕ_j. Then for any $j \in [n]$ and any integer $k \ge 1$,

$$\sum_{\gamma\in\mathbb{Z}^n_{\ge 0}:\ |\gamma|=k} |a_{j,\gamma}| \le \xi_j e^{k+1}.$$

Proof. Note that, for each $j \in [n]$,

$$a_{j,\gamma} = \frac{1}{\gamma_1!\gamma_2!\cdots\gamma_n!} \cdot \frac{\partial^k \phi_j(0)}{\partial z_1^{\gamma_1}\partial z_2^{\gamma_2}\cdots z_n^{\gamma_n}},$$

and so

$$\begin{aligned}
\sum_{\gamma\in\mathbb{Z}^n_{\geq 0}:\ |\gamma|=k} |a_{j,\gamma}| &= \sum_{\gamma\in\mathbb{Z}^n_{\geq 0}:\ |\gamma|=k} \frac{1}{\gamma_1!\gamma_2!\cdots\gamma_n!} \cdot |\frac{\partial^k \phi_j(0)}{\partial z_1^{\gamma_1}\partial z_2^{\gamma_2}\cdots z_n^{\gamma_n}}| \\
&= \frac{1}{k!} \sum_{\gamma\in\mathbb{Z}^n_{\geq 0}:\ |\gamma|=k} \frac{k!}{\gamma_1!\gamma_2!\cdots\gamma_n!} \cdot |\frac{\partial^k \phi_j(0)}{\partial z_1^{\gamma_1}\partial z_2^{\gamma_2}\cdots z_n^{\gamma_n}}| \\
&= \frac{1}{k!} \sum_{t\in[n]^k} |\frac{\partial^k \phi_j(0)}{\partial z_{t_1}\partial z_{t_2}\cdots\partial z_{t_k}}|.
\end{aligned}$$

It is straightforward to see that the following equalities hold for any $i \in [n]$, $i \neq j$:

$$\begin{aligned}
\frac{\partial \phi_j}{\partial z_j} &= \phi_j(1-\phi_j) \\
\frac{\partial \phi_j}{\partial z_i} &= -\phi_i\phi_j \\
\frac{\partial (1-\phi_j)}{\partial z_j} &= -\phi_j(1-\phi_j) \\
\frac{\partial (1-\phi_j)}{\partial z_i} &= \phi_i\phi_j
\end{aligned}$$

We claim that for any $(t_1, \ldots, t_k) \in [n]^k$, we can express $\frac{\partial^k \phi_j}{\partial z_{t_1}\cdots\partial z_{t_k}}$ as a polynomial in $\phi_1, \cdots, \phi_n, (1-\phi_1), \cdots, (1-\phi_n)$ comprised of $k!$ monomials each of degree $k+1$. We verify this by induction, first noting that after taking zero derivatives, the function ϕ_j is a degree-1 monomial. Assume that for some sequence $b_1, \ldots, b_{(\ell-1)!} \in \{0,1\}$, we can express

$$\frac{\partial^{\ell-1}\phi_j}{\partial z_{t_1}\cdots\partial z_{t_{\ell-1}}} = \sum_{f=1}^{(\ell-1)!} (-1)^{b_f} \prod_{d=0}^{\ell-1} m_{f,d}$$

where each $m_{f,d} \in \{\phi_1, \cdots, \phi_n, (1-\phi_1), \cdots, (1-\phi_n)\}$. We see that for each f, there is some sequence of bits $b_{f,0}, \ldots, b_{f,\ell-1} \in \{0,1\}$ so that

$$\frac{\partial}{\partial z_{t_\ell}} \prod_{d=0}^{\ell-1} m_{f,d} = \sum_{d=0}^{\ell-1} (-1)^{b_{f,d}} \cdot m_{f,0}\cdots m'_{f,d}\cdots m_{f,d,\ell} \tag{60}$$

where we define, for each $0 \leq d \leq \ell-1$,

$$m'_{f,d} \text{ and } m_{f,d,\ell} = \begin{cases}
m_{f,d} \text{ and } \phi_{t_\ell} & \text{if } m_{f,d} = \phi_i \text{ with } i \neq t_\ell \\
m_{f,d} \text{ and } (1-\phi_{t_\ell}) & \text{if } m_{f,d} = \phi_{t_\ell} \\
(1-m_{f,d}) \text{ and } \phi_{t_\ell} & \text{if } m_{f,d} = 1-\phi_i \text{ with } i \neq t_\ell \\
(1-m_{f,d}) \text{ and } (1-\phi_{t_\ell}) & \text{if } m_{f,d} = 1-\phi_{t_\ell}.
\end{cases}$$

Thus, $\frac{\partial^\ell \phi_j}{\partial z_{t_1} \cdots \partial z_{t_\ell}}$ can be expressed as a sum of $\ell!$ monomials of degree $(\ell+1)$, completing the inductive step.

This inductive argument also demonstrates a bijection between the $k!$ monomials of $\frac{\partial^k \phi_j}{\partial z_{t_1} \cdots \partial z_{t_k}}$ and a combinatorial structure that we call *factorial trees*. Formally, we define a factorial tree to be a directed graph on vertices $\{0, 1, \cdots, k\}$ such that each vertex $i \neq 0$ has a single incoming edge from one of the vertices in $[0, i-1]$. (For a non-negative integer i, we write $[0, i] := \{0, 1, \ldots, i\}$.) For a factorial tree f, let $p_f(\ell) \in [0, \ell - 1]$ denote the parent of a vertex ℓ. A particular factorial tree f represents the monomial that was generated by choosing the $p_f(\ell)^{\text{th}}$ term in (60) for derivation when taking the derivative $\frac{\partial}{\partial z_{t_\ell}}$, for each $\ell \in [k]$. (See Figure 3 for an example.)

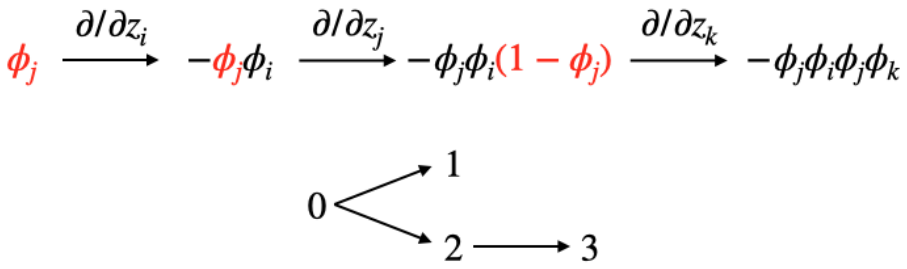

Figure 3: A monomial $-\phi_j\phi_i\phi_j\phi_k$ of $\frac{\partial^3 \phi_j}{\partial z_i \partial z_j \partial z_k}$ and its corresponding factorial tree

Each of the $k!$ monomials comprising $\frac{\partial^k \phi_j}{\partial z_{t_1} \cdots \partial z_{t_k}}$ is a product of $k+1$ terms corresponding to indices $j, t_1, \cdots, t_k$ (i.e., the first term in the product is either ϕ_j or $1 - \phi_j$, the second term is either ϕ_{t_1} or $1 - \phi_{t_1}$, and so on). We say that a term corresponding to index $i \in [n]$ is *perturbed* if it is $(1 - \phi_i)$ (as opposed to ϕ_i). From our construction, we see that the ℓ^{th} term is perturbed if $t_\ell = t_{p_f(\ell)}$ and there is no ℓ' such that $p_f(\ell') = \ell$. That is, ℓ is a leaf in the corresponding factorial tree f and the parent of ℓ corresponds to the same index as ℓ. One can think of $t_1, \cdots, t_k$ as a coloring of all the vertices of the factorial tree with n colors, except the root of the tree (vertex 0) which has fixed color j. Then, we can say the ℓ^{th} term is perturbed if and only if ℓ is a leaf with the same color as its parent. We call such a leaf a *petal*. For $t \in [n]$, we let $P_{f,t} \subseteq [k]$ be the set of petals on tree f with color t, $L_f \subseteq [k]$ be the set of leaves of tree f, and $B_f = [k] \setminus L_f$ be the set of all non-leaves other than the fixed-color root. Therefore,

$$\begin{aligned}\sum_{\gamma \in \mathbb{Z}^n_{\geq 0}:\ |\gamma| = k} |a_{j,\gamma}| &= \frac{1}{k!} \sum_{t \in [n]^k} \left| \frac{\partial^k \phi_j(0)}{\partial z_{t_1} \cdots \partial z_{t_k}} \right| \\ &\leq \frac{1}{k!} \sum_{t \in [n]^k} \sum_f \prod_{\ell=0}^{k} (\phi_{t_\ell}(0) \cdot \mathbb{1}[\ell \notin P_{f,t}] + (1 - \phi_{t_\ell}(0)) \cdot \mathbb{1}[\ell \in P_{f,t}])\end{aligned}$$

(where we let $t_0 = j$ for notational convenience)

$$= \frac{1}{k!} \sum_{t \in [n]^k} \sum_f \prod_{\ell=0}^{k} (\xi_{t_\ell} \cdot \mathbb{1}[\ell \notin P_{f,t}] + (1 - \xi_{t_\ell}) \cdot \mathbb{1}[\ell \in P_{f,t}])$$

$$= \frac{1}{k!} \sum_f \sum_{t_{B_f} \in [n]^{B_f}} \sum_{t_{L_f} \in [n]^{L_f}} \prod_{\ell=0}^{k} (\xi_{t_\ell} \cdot \mathbb{1}[\ell \notin P_{f,t}] + (1 - \xi_{t_\ell}) \cdot \mathbb{1}[\ell \in P_{f,t}]),$$

where in the last step we decompose, for each factorial tree f, $t \in [n]^k$ into the tuple of indices $t_{B_f} \in [n]^{B_f}$ corresponding to the non-leaves B_f, and the tuple of indices $t_{L_f} \in [n]^{L_f}$ corresponding to the leaves L_f.

We note that, fixing tree f and the colors of all non-leaves t_B,

$$\sum_{t_{L_f} \in [n]^{L_f}} \prod_{\ell \in L_f} (\xi_{t_\ell} \cdot \mathbb{1}[\ell \notin P_{f,t}] + (1 - \xi_{t_\ell}) \cdot \mathbb{1}[\ell \in P_{f,t}])$$

$$= \prod_{\ell \in L_f} \left(\sum_{t_\ell \in [n]} \xi_{t_\ell} \cdot \mathbb{1}[t_\ell \neq t_{p_f(\ell)}] + (1 - \xi_{t_\ell}) \cdot \mathbb{1}[t_\ell = t_{p_f(\ell)}] \right)$$

$$= \prod_{\ell \in L_f} \left(2 - 2\xi_{t_{p_f(\ell)}} \right)$$

$$\leq 2^{|L_f|}$$

And so,

$$\frac{1}{k!} \sum_f \sum_{t_{B_f} \in [n]^{B_f}} \sum_{t_{L_f} \in [n]^{L_f}} \prod_{\ell=0}^{k} (\xi_{t_\ell} \cdot \mathbb{1}[\ell \notin P_{f,t}] + (1 - \xi_{t_\ell}) \cdot \mathbb{1}[\ell \in P_{f,t}])$$

$$\leq \frac{1}{k!} \sum_f 2^{|L_f|} \sum_{t_{B_f} \in [n]^{B_f}} \prod_{\ell \in B_f \cup \{0\}} (\xi_{t_\ell} \cdot \mathbb{1}[\ell \notin P_{f,t}] + (1 - \xi_{t_\ell}) \cdot \mathbb{1}[\ell \in P_{f,t}])$$

$$= \frac{1}{k!} \sum_f 2^{|L_f|} \sum_{t_{B_f} \in [n]^{B_f}} \prod_{\ell \in B_f \cup \{0\}} \xi_{t_\ell}$$

(as no non-leaf can ever be a petal)

$$= \frac{\xi_j}{k!} \sum_f 2^{|L_f|} \prod_{\ell \in B_f} \left(\sum_{t_\ell \in [n]} \xi_{t_\ell} \right)$$

$$= \frac{\xi_j}{k!} \sum_f 2^{|L_f|} = \xi_j \mathbb{E}_{f \sim \mathcal{U}(\mathcal{F})} \left[2^{|L_f|} \right]$$

where $\mathcal{F}$ is the set of all factorial trees and $\mathcal{U}(\mathcal{F})$ is the uniform distribution over $\mathcal{F}$. For a specific vertex $\ell \in [0,k]$, we note that $\ell \in L_f$ if and only if it is not the parent of any vertex $\ell+1, \cdots, k$. So,

$$\Pr_{f\sim\mathcal{U}(\mathcal{F})}[\ell \in L_f] = \prod_{i=\ell+1}^{k} \frac{i-1}{i} = \frac{\ell}{k} \tag{61}$$

We will show via induction that, for any vertex set $S \subseteq [0,k]$

$$\Pr_{f\sim\mathcal{U}(\mathcal{F})}[S \subseteq L_f] \leq \prod_{\ell\in S} \frac{\ell}{k} \tag{62}$$

Having established the base case for every S with $|S| = 1$, we assume (62) holds for all S with $|S| < s$. For any set of s vertices V, consider an arbitrary partition of V into two sets $S \cup T = V$ with $|S|, |T| < s$. We see

$$\begin{aligned}
\Pr_{f\sim\mathcal{U}(\mathcal{F})}[V \subseteq L_f] &= \prod_{c=1}^{k} \Pr[p_f(c) \notin V] \\
&= \prod_{c=1}^{k} \Pr[p_f(c) \notin S] \Pr[p_f(c) \notin T | p_f(c) \notin S] \\
&\leq \prod_{c=1}^{k} \Pr[p_f(c) \notin S] \Pr[p_f(c) \notin T] \\
&= \Pr[S \subseteq L_f] \Pr[T \subseteq L_f] \\
&\leq \prod_{\ell\in V} \frac{\ell}{k}
\end{aligned}$$

by the inductive hypothesis, as desired. Thus, $\Pr[|L_f| \geq s] \leq \sum_{S:|S|=s} \Pr[S \subseteq L_f]$ is at most the s^{th} coefficient of the polynomial

$$R(x) = \prod_{\ell=0}^{k} \left(1 + \frac{\ell}{k} x\right)$$

and so

$$\begin{aligned}
\mathbb{E}_{f\sim\mathcal{U}(\mathcal{F})}\left[2^{|L_f|}\right] &\leq \sum_{s=0}^{k} 2^s \Pr[|L_f| \geq s] \\
&\leq R(2) \\
&\leq e^{\sum_{\ell=0}^{k} 2\ell/k} = e^{k+1}
\end{aligned}$$

and

$$\sum_{\gamma\in\mathbb{Z}^n_{\geq 0}:\ |\gamma|=k} |a_{j,\gamma}| \leq \xi_j e^{k+1},$$

as desired. □

Lemma B.7. *Let $\phi_1, \cdots, \phi_m$ be softmax-type functions. That is, for each ϕ_i, there is some $j_i \in [n]$ and indices $\xi_{i1}, \ldots, \xi_{in}$ such that*

$$\phi_i((z_1, \ldots, z_n)) = \frac{\exp(z_{j_i})}{\sum_{k=1}^n \xi_{ik} \cdot \exp(z_k)}$$

where $\xi_{i1} + \cdots + \xi_{in} = 1$ for all i. Let $P(z) = \sum_{\gamma \in \mathbb{Z}_{\geq 0}^n} a_\gamma z^\gamma$ denote the Taylor series of $\prod_i \phi_i$. Then for any integer k,

$$\sum_{\gamma \in \mathbb{Z}_{\geq 0}^n:\ |\gamma| = k} |a_\gamma| \leq (e^3 m)^k.$$

Proof. Letting $P_i(z) = \sum_{\gamma \in \mathbb{Z}_{\geq 0}^n} a_{i,\gamma} z^\gamma$ denote the Taylor series of ϕ_i for all i, we have $P(z) = \prod_i P_i(z)$ and therefore

$$\sum_{\gamma \in \mathbb{Z}_{\geq 0}^n:\ |\gamma| = k} |a_\gamma| \leq \sum_{\substack{k_1, \cdots, k_m \in \mathbb{Z}_{\geq 0} \\ \sum k_i = k}} \prod_i \sum_{\gamma \in \mathbb{Z}_{\geq 0}^n:\ |\gamma| = k_i} |a_{i,\gamma}|$$

We have that $\sum_{|\gamma| = k_i} |a_{i,\gamma}| \leq e^{2k_i}$ for all k_i since, for $k_i = 0$, $a_{i,0} = \phi_i(0) = 1$, and for $k_i \geq 1$,

$$\sum_{|\gamma| = k_i} |a_{i,\gamma}| \leq \frac{\xi_{ij}}{\xi_{ij}} e^{k_i + 1} \leq e^{2k_i} \tag{63}$$

from Lemma B.6. Note that the softmax-type functions discussed in Lemma B.6 have a ξ_{ij} term in the numerator, while those discussed here do not. This accounts for the extra ξ_{ij} term that appears in equation (63). Thus,

$$\begin{aligned}
\sum_{\substack{k_1, \cdots, k_m \in \mathbb{Z}_{\geq 0} \\ \sum k_i = k}} \prod_i \sum_{\gamma \in \mathbb{Z}_{\geq 0}^n:\ |\gamma| = k_i} |a_{i,\gamma}| &\leq \sum_{\substack{k_1, \cdots, k_m \in \mathbb{Z}_{\geq 0} \\ \sum k_i = k}} e^{2k} \\
&= e^{2k} \binom{m + k - 1}{k} \\
&\leq e^{2k} \left(\frac{e(m + k - 1)}{k} \right)^k \\
&= (e^3 m)^k
\end{aligned}$$

as desired. □

Lemma B.8. *Let $\phi((z_1, \ldots, z_n)) = \frac{\exp(z_j)}{\sum_{k=1}^n \xi_k \exp(z_k)}$ be any softmax-type function. Then the radius of convergence of the Taylor series of ϕ at the origin is at least 1.*

Proof. For a complex number z, write $\Re(z), \Im(z)$ to denote the real and imaginary parts, respectively, of z. Note that for any $\zeta_1, \ldots, \zeta_n \in \mathbb{C}$ with $|\zeta_k| \leq \pi/3$ for all $k \in [n]$, we have

$$\Re(\exp(\zeta_k)) \geq \cos(\pi/3) \cdot \exp(-\pi/3) > 1/10,$$

and thus $|\sum_{k=1}^{n} \xi_k \cdot \exp(\zeta_k)| \geq 1/10$. Moreover, for any such point $\zeta = (\zeta_1, \ldots, \zeta_n)$, it holds that $|\exp(\zeta_j)| \leq \exp(\pi/3) < 3$. It then follows that for such ζ we have $|\phi(z)| \leq 30$. In particular, ϕ is holomorphic on the region $\{\zeta : |\zeta_k| \leq \pi/3 \ \ \forall k \in [n]\}$.

Fix any $\gamma \in \mathbb{Z}^n_{\geq 0}$, and let $k = |\gamma|$. By the multivariate version of Cauchy's integral formula,

$$\begin{aligned}\left|\frac{d^\gamma}{dz^\gamma}\phi(z)\right| &= \left|\frac{\gamma!}{(2\pi i)^n}\int_{|\zeta_1 - z_1| = \pi/3} \cdots \int_{|\zeta_n - z_n| = \pi/3} \frac{\phi(\zeta_1, \ldots, \zeta_n)}{(\zeta_1 - z_1)^{\gamma_1+1} \cdots (\zeta_n - z_n)^{\gamma_n+1}} d\zeta_1 \cdots d\zeta_n\right| \\ &\leq \frac{30\gamma!}{(\pi/3)^{k+n}} \leq \frac{30\gamma!}{(\pi/3)^k}.\end{aligned}$$

The power series of ϕ at $\mathbf{0}$ is defined as $P_\phi(z) = \sum_{\gamma \in \mathbb{Z}^n_{\geq 0}} a_\gamma \cdot z^\gamma$, where $a_\gamma = \frac{1}{\gamma!}\frac{d^\gamma}{dz^\gamma}\phi(\mathbf{0})$. For any $\gamma \in \mathbb{Z}^n_{\geq 0}$ with $k = |\gamma|$, we have $|a_\gamma|^{1/k} \leq (30/(\pi/3)^k)^{1/k} = (30)^{1/k} \cdot 3/\pi$, which tends to $3/\pi < 1$ as $k \to \infty$. Thus, by the (multivariate version of the) Cauchy-Hadamard theorem, the radius of convergence of the power series of ϕ at $\mathbf{0}$ is at least $\pi/3 \geq 1$. □

B.2 Proof of Lemma 3.8

In this section prove Lemma 3.8, which, as explained in Section 3.3, is an important ingredient in the proof of Lemma 3.7. The detailed version of Lemma 3.8 is presented below; it includes several claims which are omitted for simplicity in the abbreviated version in Section 3.3.

Lemma 3.8 [Detailed] *Fix any integer $h \geq 0$, a multi-index $\gamma \in \mathbb{Z}^n_{\geq 0}$ and set $k = |\gamma|$. For each of the k^h functions $\pi : [h] \to [k]$, and for each $r \in [k]$, there are integers $h'_{\pi,r} \in \{0, 1, \ldots, h\}$, $t'_{\pi,r} \geq 0$, and $j'_{\pi,r} \in [n]$, so that the following holds. For any sequence $L^{(1)}, \ldots, L^{(T)} \in \mathbb{R}^n$ of vectors, it holds that*

$$\mathrm{D}_h\, L^\gamma = \sum_{\pi:[h]\to[k]} \prod_{r=1}^{k} \mathrm{E}_{t'_{\pi,r}}\, \mathrm{D}_{h'_{\pi,r}}\, (L(j'_{\pi,r})). \tag{64}$$

Moreover, the following properties hold:

1. *For each π and $r \in [k]$, $h'_{\pi,r} = |\{q \in [h] : \pi(q) = r\}|$. In particular, $\sum_{r=1}^{k} h'_{\pi,r} = h$.*
2. *For each π and $r \in [k]$, it holds that $0 \leq t'_{\pi,r} + h'_{\pi,r} \leq h$.*
3. *For each π, $r \in [k]$, and $j \in [n]$, $\gamma_j = |\{r \in [k] : j'_{\pi,r} = j\}|$.*

Proof of Lemma 3.8. We use induction on h. First note that in the case $h = 0$ and for any $k \geq 0$, we have that $(\mathrm{D}_h\, L^\gamma)^{(t)} = (L^{(t)})^\gamma$, and so for the unique function $\pi : \emptyset \to [k]$, for all $r \in [k]$, we may take $t'_{\pi,r} = 0$, $h'_{\pi,r} = 0$, and ensure that for each $j \in [n]$ there are γ_j values of r so that $j'_{\pi,r} = j$.

Now fix any integer $h > 0$, and suppose the statement of the claim holds for all $h' < h$. We have

that

$$
\begin{aligned}
&\mathrm{D}_h\, L^\gamma \\
&= \mathrm{D}_1\, \mathrm{D}_{h-1}\, L^\gamma \\
&= \mathrm{D}_1 \sum_{\pi:[h-1]\to[k]} \prod_{r=1}^{k} \mathrm{E}_{t'_{\pi,r}}\, \mathrm{D}_{h'_{\pi,r}}\, L(j'_{\pi,r}) \\
&= \sum_{\pi:[h-1]\to[k]} \sum_{r=1}^{k} \mathrm{D}_1\, \mathrm{E}_{t'_{\pi,r}}\, \mathrm{D}_{h'_{\pi,r}}\, L(j'_{\pi,r}) \cdot \prod_{r'=1}^{r-1} \mathrm{E}_{t'_{\pi,r'}}\, \mathrm{D}_{h'_{\pi,r'}}\, L(j'_{\pi,r'}) \cdot \prod_{r'=r+1}^{k} \mathrm{E}_{t'_{\pi,r'}+1}\, \mathrm{D}_{h'_{\pi,r'}}\, L(j'_{\pi,r'}) \quad (65) \\
&= \sum_{\pi:[h-1]\to[k]} \sum_{r=1}^{k} \mathrm{E}_{t'_{\pi,r}}\, \mathrm{D}_{h'_{\pi,r}+1}\, L(j'_{\pi,r}) \cdot \prod_{r'=1}^{r-1} \mathrm{E}_{t'_{\pi,r'}}\, \mathrm{D}_{h'_{\pi,r'}}\, L(j'_{\pi,r'}) \cdot \prod_{r'=r+1}^{k} \mathrm{E}_{t'_{\pi,r'}+1}\, \mathrm{D}_{h'_{\pi,r'}}\, L(j'_{\pi,r'}). \quad (66)
\end{aligned}
$$

where (65) uses Lemma B.3 and (66) uses the commutativity of $\mathrm{E}_{t'}$ and D_1. For each $\pi : [h-1] \to [k]$, we construct k functions $\pi_1, \ldots, \pi_k : [h] \to [k]$, defined by $\pi_r(q) = \pi(q)$ for $q < h$, and $\pi_r(h) = r$ for $r \in [k]$. Next, for $r, r' \in [k]$, we define the quantities $h'_{\pi_r,r'}, t'_{\pi_r,r'}, j'_{\pi_r,r'}$ as follows:

- Set $h'_{\pi_r,r'} = h'_{\pi,r'}$ if $r \neq r'$, and $h'_{\pi_r,r} = h'_{\pi,r} + 1$.
- Set $t'_{\pi_r,r'} = t'_{\pi,r'}$ if $r' \leq r$, and $t'_{\pi_r,r'} = t'_{\pi,r'} + 1$ if $r' > r$.
- Set $j'_{\pi_r,r'} = j'_{\pi,r'}$.

By (66) and the above definitions, we have

$$
\mathrm{D}_h\, L^\gamma = \sum_{\pi:[h]\to[k]} \prod_{r'=1}^{k} \mathrm{E}_{t'_{\pi,r}}\, \mathrm{D}_{h'_{\pi,r}}\, L(j'_{\pi,r}),
$$

thus verifying (14) for the value h.

Finally we verify that items 1 through 3 in the lemma statement hold. The definition of $h'_{\pi_r,r'}$ above together with the inductive hypothesis ensures that for all $r, r' \in [k]$, $h'_{\pi_r,r'} = |\{q \in [h] : \pi_r(q) = r'\}|$, thus verifying item 1 of the lemma statement. Since $h'_{\pi_r,r'} + t'_{\pi_r,r'} \leq h'_{\pi,r} + t'_{\pi,r} + 1$ for all r, r', it follows from the inductive hypothesis that $0 \leq h'_{\pi_r,r'} + t'_{\pi_r,r'} \leq h$; this verifies item 2. Finally, note that for any $j \in [n]$ and $r \in [k]$, $\{r' \in [k] : j'_{\pi,r'} = j\} = \{r' \in [k] : j'_{\pi_r,r'} = j\}$, and thus item 3 follows from the inductive hypothesis. □

B.3 Proof of Lemma 3.7

In this section we prove Lemma 3.7. To introduce the detailed version of the lemma we need the following definition. Suppose $\phi : \mathbb{R}^n \to \mathbb{R}$ is a real-valued function that is real-analytic in a neighborhood of the origin. For real numbers $Q, R > 0$, we say that ϕ is (Q, R)-*bounded* if the Taylor series of ϕ at $\mathbf{0}$, denoted $P_\phi(z_1, \ldots, z_n) = \sum_{\gamma \in \mathbb{Z}^n_{\geq 0}} a_\gamma z^\gamma$, satisfies, for each integer $k \geq 0$, $\sum_{\gamma \in \mathbb{Z}^n_{\geq 0} : |\gamma| = k} |a_\gamma| \leq Q \cdot R^k$. In the statement of Lemma 3.7 below, the quantity 0^0 is interpreted as 1 (in particular, $(h')^{B_0 h'} = 1$ for $h' = 0$).

Lemma 3.7 ["Boundedness chain rule" for finite differences; detailed] *Suppose that $h, n \in \mathbb{N}$, $\phi : \mathbb{R}^n \to \mathbb{R}$ is a (Q, R)-bounded function so that the radius of convergence of its power series at $\mathbf{0}$ is at least $\nu > 0$, and $L = (L^{(1)}, \ldots, L^{(T)}) \in \mathbb{R}^n$ is a sequence of vectors satisfying $\|L^{(t)}\|_\infty \leq \nu$ for $t \in [T]$. Suppose for some $\alpha \in (0, 1)$, for each $0 \leq h' \leq h$ and $t \in [T - h']$, it holds that $\|\mathrm{D}_{h'} L^{(t)}\|_\infty \leq \frac{1}{B_1} \cdot \alpha^{h'} \cdot (h')^{B_0 h'}$ for some $B_1 \geq 2e^2 R, B_0 \geq 3$. Then for all $t \in [T - h]$,*

$$|\left(\mathrm{D}_h\left(\phi \circ L\right)\right)^{(t)}| \leq \frac{12RQe^2}{B_1} \cdot \alpha^h \cdot h^{B_0 h + 1}.$$

Proof of Lemma 3.7. Note that the hth order finite differences of a constant sequence are identically 0 for $h \geq 1$, so by subtracting $\phi(\mathbf{0})$ from ϕ, we may assume without loss of generality that $\phi(\mathbf{0}) = 0$. (Here $\mathbf{0}$ denotes the all-zeros vector.)

By assumption, the radius of convergence of the power series of ϕ at the origin is at least ν, and so for each $\gamma \in \mathbb{Z}^n_{\geq 0}$, there is a real number a_γ so that for $z = (z_1, \ldots, z_n)$ with $|z_j| \leq \nu$ for each j,

$$\phi(z) = \sum_{k \in \mathbb{N}, \gamma \in \mathbb{Z}^n_{\geq 0} :\ |\gamma| = k} a_\gamma z^\gamma. \tag{67}$$

Let $A_k := \sum_{\gamma \in \mathbb{Z}^n_{\geq 0} : |\gamma| = k} |a_\gamma|$; by the assumption that ϕ is (Q, R)-bounded, we have that $A_k \leq Q \cdot R^k$ for all $k \in \mathbb{N}$.

For $\gamma \in \mathbb{Z}^n_{\geq 0}$, recall that L^γ denotes the sequence $((L^\gamma)^{(1)}, \ldots, (L^\gamma)^{(T)})$, defined by $(L^\gamma)^{(t)} = (L^{(t)}(1))^{\gamma_1} \cdots (L^{(t)}(n))^{\gamma_n}$. Then since $\|L^{(t)}\|_\infty \leq \nu$ for all $t \in [T]$, we have that, for $t \in [T - h]$, $\left(\mathrm{D}_h\left(\phi \circ L\right)\right)^{(t)} = \sum_{\gamma \in \mathbb{Z}^n_{\geq 0}} a_\gamma \cdot \left(\mathrm{D}_h\, L^\gamma\right)^{(t)}$.

We next upper bound the quantities $|\left(\mathrm{D}_h\, L^\gamma\right)^{(t)}|$. To do so, fix some $\gamma \in \mathbb{Z}^n_{\geq 0}$, and set $k = |\gamma|$. For each function $\pi : [h] \to [k]$ and $r \in [k]$, recall the integers $h'_{\pi,r} \in \{0, 1, \ldots, h\}$, $t'_{\pi,r} \geq 0$, $j'_{\pi,r} \in [n]$ defined in Lemma 3.8. By assumption it holds that for each $t \in [T - h]$, each $h' \leq h$, each $0 \leq t' \leq h$, $|\left(\mathrm{D}_{h'}\, L(j)\right)^{(t+t')}| \leq \frac{1}{B_1} \cdot \alpha^{h'} \cdot (h')^{B_0 h'}$. It follows that for each $t \in [T - h]$ and function $\pi : [h] \to [k]$,

$$\left|\prod_{r=1}^k \left(\mathrm{E}_{t'_{\pi,r}} \mathrm{D}_{h'_{\pi,r}} L(j'_{\pi,r})\right)^{(t)}\right| \leq \prod_{r=1}^k \frac{1}{B_1} \cdot \alpha^{h'_{\pi,r}} \cdot (h'_{\pi,r})^{B_0 h'_{\pi,r}} = \frac{\alpha^h}{B_1^k} \cdot \prod_{r=1}^k (h'_{\pi,r})^{B_0 h'_{\pi,r}},$$

where the last equality uses that $\sum_{r=1}^k h'_{\pi,r} = h$ (item 1 of Lemma 3.8). Then by Lemma 3.8, we have:

$$\begin{aligned}
\left|\left(\mathrm{D}_h\, L^\gamma\right)^{(t)}\right| &\leq \sum_{\pi : [h] \to [k]} \left|\prod_{r=1}^k \left(\mathrm{E}_{t'_{\pi,r}} \mathrm{D}_{h'_{\pi,r}} L(j'_{\pi,r})\right)^{(t)}\right| \\
&\leq \frac{\alpha^h}{B_1^k} \sum_{\pi : [h] \to [k]} \prod_{r=1}^k (h'_{\pi,r})^{B_0 h'_{\pi,r}} \\
&\leq \frac{\alpha^h}{B_1^k} \cdot h^{B_0 h} \max\left\{k^7, (hk + 1) \cdot \exp\left(\frac{2k}{h^{B_0 - 1}}\right)\right\},
\end{aligned} \tag{68}$$

where (68) follows from Lemma B.5, the fact that $B_0 \geq 3$, and that $h'_{\pi,r} = |\{q \in [h] : \pi(q) = r\}|$ (item 1 of Lemma 3.8).

We may now bound the order-h finite differences of the sequence $\phi \circ L$ as follows: for $t \in [T-h]$,

$$
\begin{aligned}
|\left(\mathrm{D}_h\left(\phi \circ L\right)\right)^{(t)}| \leq & \sum_{\gamma \in \mathbb{Z}^n_{\geq 0}} |a_\gamma| \cdot \left|\left(\mathrm{D}_h\, L^\gamma\right)^{(t)}\right| \\
\leq & \alpha^h \cdot h^{B_0 h+1} \sum_{\gamma \in \mathbb{Z}^n_{\geq 0}} |a_\gamma| \cdot B_1^{-|\gamma|} \cdot \max\left\{|\gamma|^7, (|\gamma|+1) \cdot \exp\left(\frac{2|\gamma|}{h^{B_0-1}}\right)\right\} \qquad (69) \\
\leq & \alpha^h \cdot h^{B_0 h+1} \cdot \sum_{k \in \mathbb{N}} A_k \cdot B_1^{-k} \cdot \left(k^7 + 2k \cdot \exp(2k/h^{B_0-1})\right) \\
\leq & \alpha^h \cdot h^{B_0 h+1} \cdot Q \cdot \left(\sum_{k \in \mathbb{N}} k^7 \cdot (R/B_1)^k + \sum_{k \in \mathbb{N}} 2k \cdot (R/B_1)^k \cdot e^{2k}\right) \qquad (70) \\
\leq & \frac{2RQe^2}{B_1} \cdot \alpha^h \cdot h^{B_0 h+1} \cdot \left(\sum_{k \in \mathbb{N}} k^7 \cdot (2e^2)^{-k} + 2\sum_{k \in \mathbb{N}} k \cdot 2^{-k}\right) \qquad (71) \\
= & \frac{12RQe^2}{B_1} \cdot \alpha^h \cdot h^{B_0 h+1}.
\end{aligned}
$$

where (69) uses (68), (70) uses the bound $A_k \leq QR^k$, and (71) uses the assumption $B_1 \geq 2e^2 R$. This gives the desired conclusion of the lemma. □

B.4 Proof of Lemma 3.6

In this section we prove Lemma 3.6. The detailed version of Lemma 3.6 is stated below.

Lemma 3.6 [Detailed] *Fix a parameter $\alpha \in \left(0, \frac{1}{H+3}\right)$. If all players follow OMWU updates with step size $\eta \leq \frac{\alpha}{36e^5 m}$, then for any player $i \in [m]$, integer h satisfying $0 \leq h \leq H$, time step $t \in [T-h]$, it holds that*

$$\|\left(\mathrm{D}_h\, \ell_i\right)^{(t)}\|_\infty \leq \alpha^h \cdot h^{3h+1}.$$

Proof. We have that for each agent $i \in [m]$, each $t \in [T]$, and each $a_i \in [n_i]$, $\ell_i^{(t)}(a_i) = \mathbb{E}_{a_{i'} \sim x_{i'}^{(t)}:\, i' \neq i}[\mathcal{L}_i(a_1, \ldots, a_m)]$. Thus, for $1 \leq t \leq T$,

$$
\begin{aligned}
\left|\left(\mathrm{D}_h\, \ell_i\right)^{(t)}(a_i)\right| &= \left|\sum_{s=0}^{h} \binom{h}{s} (-1)^{h-s} \ell_i^{(t+s)}(a_i)\right| \qquad (72) \\
&= \left|\sum_{a_{i'} \in [n_{i'}],\ \forall i' \neq i} \mathcal{L}_i(a_1, \ldots, a_m) \sum_{s=0}^{h} \binom{h}{s} (-1)^{h-s} \cdot \prod_{i' \neq i} x_{i'}^{(t+s)}(a_{i'})\right| \\
&\leq \sum_{a_{i'} \in [n_{i'}],\ \forall i' \neq i} \left|\sum_{s=0}^{h} \binom{h}{s} (-1)^{h-s} \cdot \prod_{i' \neq i} x_{i'}^{(t+s)}(a_{i'})\right| \\
&= \sum_{a_{i'} \in [n_{i'}],\ \forall i' \neq i} \left|\left(\mathrm{D}_h \left(\prod_{i' \neq i} x_{i'}(a_{i'})\right)\right)^{(t)}\right|, \qquad (73)
\end{aligned}
$$

where (72) and (73) use Remark 3.5 and in (73), $\prod_{i' \neq i} x_{i'}(a_{i'})$ refers to the sequence $\prod_{i' \neq i} x_{i'}^{(1)}(a_{i'})$, $\prod_{i' \neq i} x_{i'}^{(2)}(a_{i'}), \ldots, \prod_{i' \neq i} x_{i'}^{(T)}(a_{i'})$.

In the remainder of this lemma we will prepend to the loss sequence $\ell_i^{(1)}, \ldots, \ell_i^{(T)}$ the vectors $\ell_i^{(0)} = \ell_i^{(-1)} := \mathbf{0} \in \mathbb{R}^{n_i}$. We will also prepend $x_i^{(0)} := x_i^{(1)} = (1/n_i, \ldots, 1/n_i) \in \Delta^{n_i}$ to the strategy sequence $x_i^{(1)}, \ldots, x_i^{(T)}$. Next notice that for any agent $i \in [m]$, any $t_0 \in \{0, 1, \ldots, T\}$, and any $t \geq 0$, by the definition (OMWU) of the OMWU updates, it holds that, for each $j \in [n_i]$,

$$x_i^{(t_0+t+1)}(j) = \frac{x_i^{(t_0)}(j) \cdot \exp\left(\eta \cdot \left(\ell_i^{(t_0-1)}(j) - \sum_{s=0}^{t} \ell_i^{(t_0+s)}(j) - \ell_i^{(t_0+t)}(j)\right)\right)}{\sum_{k=1}^{n_i} x_i^{(t_0)}(k) \cdot \exp\left(\eta \cdot \left(\ell_i^{(t_0-1)}(k) - \sum_{s=0}^{t} \ell_i^{(t_0+s)}(k) - \ell_i^{(t_0+t)}(k)\right)\right)}.$$

Note in particular that our definitions of $\ell_i^{(0)}, \ell_i^{(-1)}, x_i^{(0)}$ ensure that the above equation holds even for $t_0 \in \{0, 1\}$. Now an integer t_0 satisfying $0 \leq t_0 \leq T$; for $t \geq 0$, let us write

$$\bar{\ell}_{i,t_0}^{(t)} := \ell_i^{(t_0-1)} - \sum_{s=0}^{t-1} \ell_i^{(t_0+s)} - \ell_i^{(t_0+t-1)}.$$

Also, for a vector $z = (z(1), \ldots, z(n_i)) \in \mathbb{R}^{n_i}$ and an index $j \in [n_i]$, define

$$\phi_{t_0,j}(z) := \frac{\exp(z(j))}{\sum_{k=1}^{n_i} x_i^{(t_0)}(k) \cdot \exp(z(k))}, \tag{74}$$

so that $x_i^{(t_0+t)}(j) = x_i^{(t_0)}(j) \cdot \phi_{t_0,j}(\eta \cdot \bar{\ell}_{i,t_0}^{(t)})$ for $t \geq 1$. In particular, for any $i \in [m]$, and any choices of $a_{i'} \in [n_{i'}]$ for all $i' \neq i$,

$$\prod_{i' \neq i} x_{i'}^{(t_0+t)}(a_{i'}) = \prod_{i' \neq i} x_{i'}^{(t_0)}(a_{i'}) \cdot \phi_{t_0,a_{i'}}(\eta \cdot \bar{\ell}_{i',t_0}^{(t)}). \tag{75}$$

Next, note that

$$\left(\mathrm{D}_1 \, \bar{\ell}_{i,t_0}\right)^{(t)} = \ell_i^{(t_0+t-1)} - 2\ell_i^{(t_0+t)} = \ell_i^{(t_0+t-1)} - 2\left(\mathrm{E}_1 \, \ell_i\right)^{(t_0+t-1)},$$

meaning that for any $h' \geq 1$,

$$\left(\mathrm{D}_{h'} \, \bar{\ell}_{i,t_0}\right)^{(t)} = \left(\mathrm{D}_{h'-1} \, \ell_i\right)^{(t_0+t-1)} - 2\left(\mathrm{E}_1 \, \mathrm{D}_{h'-1} \, \ell_i\right)^{(t_0+t-1)}. \tag{76}$$

We next establish the following claims which will allow us to prove Lemma 3.6 by induction.

Claim B.9. *For any $t_0 \in \{0, 1, \ldots, T\}$, $t \geq 0$, and $i \in [m]$, it holds that $\|\bar{\ell}_{i,t_0}^{(t)}\|_\infty \leq t + 2$.*

Proof of Claim B.9. The claim is immediate from the triangle inequality and the fact that $\|\ell_i^{(t)}\|_\infty \leq 1$ for all $t \in [T]$. □

Claim B.10. *Fix h so that $1 \leq h \leq H$. Suppose that for some $B_0 \geq 3$ and for all $0 \leq h' < h$, all $i \in [m]$, and all $t \leq T - h'$, it holds that $\| \left(\mathrm{D}_{h'} \, \ell_i\right)^{(t)} \|_\infty \leq \alpha^{h'} \cdot (h'+1)^{B_0(h'+1)}$. Suppose that the step size η satisfies $\eta \leq \min\left\{\frac{\alpha}{36e^5 m}, \frac{1}{12e^5(H+3)m}\right\}$. Then for all $i \in [m]$ and $1 \leq t \leq T - h$,*

$$\left\|\left(\mathrm{D}_h \, \ell_i\right)^{(t)}\right\|_\infty \leq \alpha^h \cdot h^{B_0 h + 1}. \tag{77}$$

Proof of Claim B.10. Set $B_1 := 12e^5 m$, so that the assumption of the claim gives $\eta \le \min\left\{\frac{\alpha}{3B_1}, \frac{1}{B_1(H+3)}\right\}$.

We first use Lemma 3.7 to bound, for each $0 \le t_0 \le T-h$, $i \in [m]$, and $a_{i'} \in [n_{i'}]$ for all $i' \ne i$, the quantity $\left|\left(\mathrm{D}_h\left(\prod_{i'\ne i} x_{i'}(a_{i'})\right)\right)^{(t_0+1)}\right|$. In particular, we will apply Lemma 3.7 with $n = \sum_{i'\ne i} n_{i'}$, $\nu = 1$, the value of h in the statement of Claim B.10, $T = h+1$, and the sequence $L^{(t)}$, for $1 \le t \le h+1$, defined as

$$L^{(t)} = \left(\eta \cdot \bar{\ell}^{(t)}_{1,t_0}, \ldots, \eta \cdot \bar{\ell}^{(t)}_{i-1,t_0}, \eta \cdot \bar{\ell}^{(t)}_{i+1,t_0}, \ldots, \eta \cdot \bar{\ell}^{(t)}_{m,t_0}\right),$$

namely the concatenation of the vectors $\eta \cdot \bar{\ell}^{(t)}_{1,t_0}, \ldots, \eta \cdot \bar{\ell}^{(t)}_{i-1,t_0}, \eta \cdot \bar{\ell}^{(t)}_{i+1,t_0}, \ldots, \eta \cdot \bar{\ell}^{(t)}_{m,t_0}$. The function ϕ in Lemma 3.7 is set to the function that takes as input the concatenation of $z_{i'} \in \mathbb{R}^{n_{i'}}$ for all $i' \ne i$ and outputs:

$$\phi_{t_0,a_{-i}}(z_1, \ldots, z_{i-1}, z_{i+1}, \ldots, z_m) := \prod_{i'\ne i} \phi_{t_0,a_{i'}}(z_{i'}), \tag{78}$$

where the function $\phi_{t_0,a_{i'}}$ are as defined in (74). We first verify the preconditions of Lemma 3.7. By Lemma B.7, $\phi_{t_0,a_{-i}}$ is a $(1, e^3 m)$-bounded function for some constant $C \ge 1$. By Lemma B.8, the radius of convergence of each function $\phi_{t_0,a_{i'}}$ at $\mathbf{0}$ is at least 1; thus the radius of convergence of $\phi_{t_0,a_{-i}}$ at $\mathbf{0}$ is at least $\nu = 1$. Claim B.9 gives that $\|\bar{\ell}^{(t)}_{i,t_0}\|_\infty \le t+2 \le h+3$ for all $t \le h+1$. Thus, since $\eta \le \frac{1}{B_1(H+3)}$,

$$\left\|\left(\mathrm{D}_0\left(\eta \cdot \bar{\ell}_{i,t_0}\right)\right)^{(t)}\right\|_\infty = \|\eta \cdot \bar{\ell}^{(t)}_{i,t_0}\|_\infty \le \eta \cdot (H+3) \le \frac{1}{B_1}$$

for $1 \le t \le h_0 + 1$. Next, for $1 \le h' \le h$ and $1 \le t \le h+1-h'$, we have

$$\begin{aligned}
\left\|\left(\mathrm{D}_{h'}\left(\eta \cdot \bar{\ell}_{i,t_0}\right)\right)^{(t)}\right\|_\infty \le& \eta \cdot \left\|\left(\mathrm{D}_{h'-1}\,\ell_i\right)^{(t_0+t-1)}\right\|_\infty + 2\eta \cdot \left\|\left(\mathrm{D}_{h'-1}\,\ell_i\right)^{(t_0+t)}\right\|_\infty && (79)\\
\le& 3\eta \cdot \alpha^{h'-1} \cdot (h')^{B_0(h')} && (80)\\
\le& \frac{1}{B_1} \cdot \alpha^{h'} \cdot (h')^{B_0(h')}, && (81)
\end{aligned}$$

where (79) follows from (76), (80) follows from the assumption in the statement of Claim B.10 and $t_0 + t + h' - 1 \le t_0 + h \le T$, and (81) follows from the fact that $3\eta \le \frac{\alpha}{B_1}$. It then follows from Lemma 3.7 and (75) that

$$\begin{aligned}
&\frac{1}{\prod_{i'\ne i} x^{(t_0)}_{i'}(a_{i'})} \cdot \left|\left(\mathrm{D}_h\left(\prod_{i'\ne i} x_{i'}(a_{i'})\right)\right)^{(t_0+1)}\right|\\
=& \left|\left(\mathrm{D}_h\left(\prod_{i'\ne i}\left(\phi_{t_0,a_{i'}} \circ \eta\bar{\ell}_{i',t_0}\right)\right)\right)^{(t_0+1)}\right| && (82)\\
=& \left|\left(\mathrm{D}_h\left(\phi_{t_0,a_{-i}} \circ (\eta\bar{\ell}_{1,t_0}, \ldots, \eta\bar{\ell}_{i-1,t_0}, \eta\bar{\ell}_{i+1,t_0}, \ldots, \eta\bar{\ell}_{m,t_0})\right)\right)^{(1)}\right| && (83)\\
\le& \frac{12e^5 m}{B_1} \cdot \alpha^h \cdot h^{B_0 h+1} = \alpha^h \cdot (h)^{B_0 h+1}. && (84)
\end{aligned}$$

(In particular, (82) uses (75), (83) uses the definition of $\phi_{t_0,a_{-i}}$ in (78), and (84) uses Lemma 3.7.)

Next we use (73), which gives that for each $i \in [m]$ and $t \geq 1$,

$$\begin{aligned}\left\|(\mathrm{D}_h\,\ell_i)^{(t)}\right\|_\infty \leq& \sum_{a_{i'}\in[n_{i'}],\ \forall i'\neq i} \left|\left(\mathrm{D}_h\left(\prod_{i'\neq i} x_{i'}(a_{i'})\right)\right)^{(t)}\right| \\ \leq& \sum_{a_{i'}\in[n_{i'}],\ \forall i'\neq i} \prod_{i'\neq i} x_{i'}^{(t_0)}(a_{i'})\cdot\alpha^h\cdot(h)^{B_0h+1} \qquad (85)\\ =&\alpha^h\cdot(h)^{B_0h+1},\end{aligned}$$

where (85) follows from (84) with $t = t_0 - 1$ (here we use that t_0 may be 0). This completes the proof of Claim B.10. □

It is immediate that for all $i \in [m], t \in [T]$, we have that $\|\,(\mathrm{D}_0\,\ell_i)^{(t)}\,\|_\infty \leq 1 = \alpha^0\cdot 1^{B_0\cdot 1}$. We now apply Claim B.10 inductively with $B_0 = 3$, for which it suffices to have $\eta \leq \frac{\alpha}{36e^5m}$ as long as $\alpha \leq 1/(H+3)$. This gives that for $0 \leq h \leq H$, $i \in [m]$, and $t \in [T-h]$, $\|\,(\mathrm{D}_h\,\ell_i)^{(t)}\,\|_\infty \leq \alpha^h\cdot h^{3h+1}$, completing the proof of Lemma 3.6. □

C Proofs for Section 3.4

The main goal of this section is to prove Lemma 3.9. First, in Section C.1 we prove some preliminary lemmas and then we prove Lemma 3.9 in Section C.2

C.1 Preliminary lemmas

Lemma C.1 shows that $\mathrm{Var}_P(W)$ and $\mathrm{Var}_{P'}(W)$ are close when the entries of P, P' are close; it will be applied with P, P' equal to the strategies $x_i^{(t)} \in \Delta^{n_i}$ played in the course of OMWU.

Lemma C.1. *Suppose $n \in \mathbb{N}$ and $M > 0$ are given, and $W \in \mathbb{R}^n$ is a vector. Suppose $P, P' \in \Delta^n$ are distributions with $\max\left\{\left\|\frac{P}{P'}\right\|_\infty, \left\|\frac{P'}{P}\right\|_\infty\right\} \leq 1+\alpha$ for some $\alpha > 0$. Then*

$$(1-\alpha)\,\mathrm{Var}_P(W) \leq \mathrm{Var}_{P'}(W) \leq (1+\alpha)\,\mathrm{Var}_P(W)\,. \qquad (86)$$

Proof. We first prove that $\mathrm{Var}_{P'}(W) \leq (1+\alpha)\,\mathrm{Var}_P(W)$. To do so, note that since adding a constant to every entry of W does not change $\mathrm{Var}_P(W)$ or $\mathrm{Var}_{P'}(W)$, by replacing W with $W - \langle P, W\rangle\cdot\mathbf{1}$, we may assume without loss of generality that $\langle P, W\rangle = 0$. Thus $\mathrm{Var}_P(W) = \sum_{j=1}^n P(j)W(j)^2$. Now we may compute:

$$\begin{aligned}\mathrm{Var}_{P'}(W) \leq& \sum_j P'(j)\cdot W(j)^2 \\ =& \sum_j P(j)\cdot W(j)^2 + \sum_j (P'(j)-P(j))\cdot W(j)^2 \\ =&(1+\alpha)\,\mathrm{Var}_P(W)\,, \qquad (87)\end{aligned}$$

where (87) uses the fact that $\left\|\frac{P'}{P}\right\|_\infty \leq 1+\alpha$.

By interchanging the roles of P, P', we obtain that

$$\mathrm{Var}_{P'}(W) \geq \frac{1}{1+\alpha}\,\mathrm{Var}_P(W) \geq (1-\alpha)\,\mathrm{Var}_P(W).$$

This completes the proof of the lemma. □

Next we prove Lemma 3.11 (recall that only the special case $\mu = 0$ was proved in Section 3.4). For convenience the lemma is repeated below.

Lemma 3.11 [Restated] *Suppose $\mu \in \mathbb{R}$, $\alpha > 0$, and $W^{(0)}, \ldots, W^{(S-1)} \in \mathbb{R}$ is a sequence of reals satisfying*

$$\sum_{t=0}^{S-1} \left((\mathrm{D}_2^\circ W)^{(t)}\right)^2 \leq \alpha \cdot \sum_{t=0}^{S-1} \left((\mathrm{D}_1^\circ W)^{(t)}\right)^2 + \mu. \tag{88}$$

Then

$$\sum_{t=0}^{S-1} \left((\mathrm{D}_1^\circ W)^{(t)}\right)^2 \leq \alpha \cdot \sum_{t=1}^{S-1} (W^{(t)})^2 + \mu/\alpha.$$

To prove Lemma 3.11 we need the following basic facts about the Fourier transform:

Fact C.2 (Parseval's equality)**.** *It holds that $\sum_{t=0}^{S-1} |W^{(t)}|^2 = \frac{1}{S}\sum_{s=0}^{S-1} |\widehat{W}^{(s)}|^2$.*

The second fact gives a formula for the Fourier transform of the circular finite differences; its simple form is the reason we work with *circular* finite differences in this section:

Fact C.3. *For $h \in \mathbb{Z}_{\geq 0}$, $\widehat{\mathrm{D}_h^\circ W}^{(s)} = \widehat{W}^{(s)} \cdot (e^{2\pi i s t/S} - 1)^h$.*

Proof of Lemma 3.11. Note that the discrete Fourier transform of $\mathrm{D}_1^\circ W$ satisfies $\widehat{\mathrm{D}_1^\circ W}^{(s)} = \widehat{W}^{(s)} \cdot (e^{2\pi i s/T} - 1)$, and similarly $\widehat{\mathrm{D}_2^\circ W}^{(s)} = \widehat{W}^{(s)} \cdot (e^{2\pi i s/T} - 1)^2$, for $0 \leq s \leq S-1$. By the Cauchy-Schwarz

inequality, Parseval's equality (Fact C.2), Fact C.3, and the assumption that (88) holds, we have

$$\begin{aligned}
\sum_{t=0}^{S-1}\left((\mathrm{D}_1^\circ W)^{(t)}\right)^2 =& \frac{1}{S}\sum_{s=0}^{S-1}\left|\widehat{\mathrm{D}_1^\circ W}^{(s)}\right|^2 \\
=& \frac{1}{S}\sum_{s=0}^{S-1}\left|\widehat{W}^{(s)}\cdot(e^{2\pi is/T-1})\right|^2 \\
=& \frac{1}{S}\sum_{s=0}^{S-1}\left|\widehat{W}^{(s)}\right|\cdot\left|\widehat{W}^{(s)}\right|\left|e^{2\pi is/T}-1\right|^2 \\
\le& \sqrt{\frac{1}{S}\sum_{s=0}^{S-1}\left|\widehat{W}^{(s)}\right|^2}\cdot\sqrt{\frac{1}{S}\sum_{s=0}^{S-1}\left|\widehat{W}^{(s)}\right|^2\cdot\left|e^{2\pi is/T}-1\right|^4} \\
=& \sqrt{\sum_{t=0}^{S-1}(W^{(t)})^2}\cdot\sqrt{\frac{1}{S}\sum_{s=0}^{S-1}\left|\widehat{\mathrm{D}_2^\circ W}^{(s)}\right|^2} \\
=& \sqrt{\sum_{t=0}^{S-1}(W^{(t)})^2}\cdot\sqrt{\sum_{t=0}^{S-1}\left((\mathrm{D}_2^\circ W)^{(t)}\right)^2} \\
\le& \sqrt{\sum_{t=0}^{S-1}(W^{(t)})^2}\cdot\sqrt{\alpha\cdot\sum_{t=0}^{S-1}\left((\mathrm{D}_1^\circ W)^{(t)}\right)^2+\mu}. \qquad (89)
\end{aligned}$$

Note that for real numbers $A>0$ and ϵ with $A+\epsilon>0$, it holds that

$$\frac{A^2}{A+\epsilon}=\frac{A}{1+\epsilon/A}\ge A\cdot(1-\epsilon/A)=A-\epsilon.$$

Taking $A=\sum_{t=0}^{S-1}\left((\mathrm{D}_1^\circ W)^{(t)}\right)^2$ and $\epsilon=\mu/\alpha$ (for which $A+\epsilon>0$ is immediate) and using (89) then gives

$$\sum_{t=0}^{S-1}\left((\mathrm{D}_1^\circ W)^{(t)}\right)^2-\mu/\alpha\le\frac{\left(\sum_{t=0}^{S-1}\left((\mathrm{D}_1^\circ W)^{(t)}\right)^2\right)^2}{\sum_{t=0}^{S-1}\left((\mathrm{D}_1^\circ W)^{(t)}\right)^2+\mu/\alpha}\le\alpha\cdot\sum_{t=0}^{S-1}\left(W^{(t)}\right)^2,$$

as desired. □

C.2 Proof of Lemma 3.9

Now we prove Lemma 3.9. For convenience we restate the lemma below with the exact value of the constant C_0 referred to in the version in Section 3.4.

Lemma 3.9 [Restated] *For any $M,\zeta,\alpha>0$ and $n\in\mathbb{N}$, suppose that $P^{(1)},\dots,P^{(T)}\in\Delta^n$ and $Z^{(1)},\dots,Z^{(T)}\in[-M,M]^n$ satisfy the following conditions:*

1. *The sequence $P^{(1)},\dots,P^{(T)}$ is ζ-consecutively close for some $\zeta\in[1/(2T),\alpha^4/8256]$.*

2. It holds that $\sum_{t=1}^{T-2} \mathrm{Var}_{P^{(t)}}\left((\mathrm{D}_2 Z)^{(t)}\right) \le \alpha \cdot \sum_{t=1}^{T-1} \mathrm{Var}_{P^{(t)}}\left((\mathrm{D}_1 Z)^{(t)}\right) + \mu$.

Then

$$\sum_{t=1}^{T-1} \mathrm{Var}_{P^{(t)}}\left((\mathrm{D}_1 Z)^{(t)}\right) \le \alpha \cdot (1+\alpha) \sum_{t=1}^{T} \mathrm{Var}_{P^{(t)}}\left(Z^{(t)}\right) + \frac{\mu}{\alpha} + \frac{1290M^2}{\alpha^3}. \tag{90}$$

Proof. Fix a positive integer $S < 1/(2\zeta) < T$, to be specified exactly below. For $1 \le t_0 \le T-S+1$, define $\mu_{t_0} \in \mathbb{R}$ by

$$\mu_{t_0} = \sum_{s=0}^{S-3} \mathrm{Var}_{P^{(t_0+s)}}\left((\mathrm{D}_2 Z)^{(t_0+s)}\right) - \alpha \cdot \sum_{s=0}^{S-3} \mathrm{Var}_{P^{(t_0+s)}}\left((\mathrm{D}_1 Z)^{(t_0+s)}\right). \tag{91}$$

Then

$$\begin{aligned}
&\sum_{t_0=1}^{T-S+1} \mu_{t_0} \\
=&\sum_{t=1}^{T-2} \mathrm{Var}_{P^{(t)}}\left((\mathrm{D}_2 Z)^{(t)}\right) \cdot \min\{S-2, t, T-t-1\} \\
&- \alpha \cdot \sum_{t=1}^{T-1} \mathrm{Var}_{P^{(t)}}\left((\mathrm{D}_1 Z)^{(t)}\right) \cdot \min\{S-2, t, T-t-1\} \\
\le&(S-2) \cdot \sum_{t=1}^{T-2} \mathrm{Var}_{P^{(t)}}\left((\mathrm{D}_2 Z)^{(t)}\right) - (S-2)\alpha \cdot \sum_{t=1}^{T-1} \mathrm{Var}_{P^{(t)}}\left((\mathrm{D}_1 Z)^{(t)}\right) + 8\alpha(S-2)^2M^2 \qquad (92)\\
\le&(S-2)\mu + 2\alpha(S-2)^2M^2, \qquad (93)
\end{aligned}$$

where (92) uses the fact that $\|Z^{(t)}\|_\infty \le M$ and so $\|\,(\mathrm{D}_1 Z)^{(t)}\,\|_\infty \le 2M$ for all $t \in [T]$, and the final inequality (93) follows from assumption 2 of the lemma statement.

By (91) and Lemma C.1 with $P = P^{(t_0)}$, we have, for some constant $C > 0$,

$$\begin{aligned}
\sum_{s=0}^{S-3} \mathrm{Var}_{P^{(t_0)}}\left((\mathrm{D}_2 Z)^{(t_0+s)}\right) \le&(1+2\zeta S) \cdot \sum_{s=0}^{S-3} \mathrm{Var}_{P^{(t_0+s)}}\left((\mathrm{D}_2 Z)^{(t_0+s)}\right) \\
=&(1+2\zeta S)\alpha \cdot \sum_{s=0}^{S-3} \mathrm{Var}_{P^{(t_0+s)}}\left((\mathrm{D}_1 Z)^{(t_0+s)}\right) + (1+2\zeta S)\mu_{t_0} \\
\le&(1+2\zeta S)^2\alpha \cdot \sum_{s=0}^{S-3} \mathrm{Var}_{P^{(t_0)}}\left((\mathrm{D}_1 Z)^{(t_0+s)}\right) + (1+2\zeta S)\mu_{t_0}. \qquad (94)
\end{aligned}$$

Here we have used that for $0 \le s \le S$, it holds that $\max\left\{\left\|\frac{P^{(t_0+s)}}{P^{(t_0)}}\right\|_\infty, \left\|\frac{P^{(t_0)}}{P^{(t_0+s)}}\right\|_\infty\right\} \le (1+\zeta)^S \le 1 + 2\zeta S$ since $\zeta S \le 1/2$.

For any integer $1 \le t_0 \le T-S+1$, we define the sequence $Z_{t_0}^{(s)} := Z^{(t_0+s)} - \langle P^{(t_0)}, Z^{(t_0+s)}\rangle \mathbf{1}$, for $0 \le s \le S-1$. Thus $\langle Z_{t_0}^{(s)}, P^{(t_0)}\rangle = 0$ for $0 \le s \le S-1$, which implies that for all $h \ge 0$,

$0 \le s \le S-1$, $\langle (\mathrm{D}_h^\circ Z_{t_0})^{(s)}, P^{(t_0)} \rangle = 0$, and thus

$$\mathrm{Var}_{P^{(t_0)}}\left((\mathrm{D}_h^\circ Z_{t_0})^{(s)}\right) = \sum_{j=1}^{n} P^{(t_0)}(j) \cdot (\mathrm{D}_h^\circ Z_{t_0})^{(s)}(j)^2. \tag{95}$$

By the definition of the sequence Z_{t_0}, for $0 \le s \le S-h-1$, we have

$$\mathrm{Var}_{P^{(t_0)}}\left((\mathrm{D}_h Z)^{(t_0+s)}\right) = \mathrm{Var}_{P^{(t_0)}}\left((\mathrm{D}_h Z_{t_0})^{(s)}\right) = \mathrm{Var}_{P^{(t_0)}}\left((\mathrm{D}_h^\circ Z_{t_0})^{(s)}\right). \tag{96}$$

For $1 \le t_0 \le T-S+1$, let us now define

$$\nu_{t_0,j} := \sum_{s=0}^{S-1} (\mathrm{D}_2^\circ Z_{t_0})^{(s)}(j)^2 - (1+2\zeta S)^2\alpha \cdot \sum_{s=0}^{S-1} (\mathrm{D}_1^\circ Z_{t_0})^{(s)}(j)^2, \tag{97}$$

so that, by (94), (95), and (96),

$$\begin{aligned}
&\sum_{j=1}^{n} P^{(t_0)}(j) \cdot \nu_{t_0,j} \\
=&\sum_{s=0}^{S-1} \mathrm{Var}_{P^{(t_0)}}\left((\mathrm{D}_2^\circ Z_{t_0})^{(s)}\right) - (1+2\zeta S)^2\alpha \cdot \sum_{s=0}^{S-1} \mathrm{Var}_{P^{(t_0)}}\left((\mathrm{D}_1^\circ Z_{t_0})^{(s)}\right) \\
\le&\left(\sum_{s=0}^{S-3} \mathrm{Var}_{P^{(t_0)}}\left((\mathrm{D}_2 Z)^{(t_0+s)}\right)\right) + \mathrm{Var}_{P^{(t_0+1)}}\left((\mathrm{D}_2^\circ Z)^{(t_0+S-2)}\right) + \mathrm{Var}_{P^{(t_0+1)}}\left((\mathrm{D}_2^\circ Z)^{(t_0+S-1)}\right) \\
&- (1+2\zeta S)^2\alpha \cdot \sum_{s=0}^{S-3} \mathrm{Var}_{P^{(t_0)}}\left((\mathrm{D}_1 Z)^{(t_0+s)}\right) \\
\le&(1+2\zeta S)\mu_{t_0} + \mathrm{Var}_{P^{(t_0)}}\left((\mathrm{D}_2^\circ Z)^{(t_0+S-2)}\right) + \mathrm{Var}_{P^{(t_0)}}\left((\mathrm{D}_2^\circ Z)^{(t_0+S-1)}\right).
\end{aligned} \tag{98}$$

By (97) and Lemma 3.11 applied to the sequence $Z_{t_0}^{(0)}, \dots, Z_{t_0}^{(S-1)}$, it holds that, for each $j \in [n]$,

$$\sum_{s=0}^{S-1} (\mathrm{D}_1^\circ Z_{t_0})^{(s)}(j)^2 \le (1+\zeta CS)^2\alpha \cdot \sum_{s=0}^{S-1} Z_{t_0}^{(s)}(j)^2 + \frac{\nu_{t_0,j}}{(1+\zeta CS)^2\alpha}. \tag{99}$$

Then we have:

$$\begin{aligned}
&\sum_{s=0}^{S-2} \mathrm{Var}_{P^{(t_0)}}\left((\mathrm{D}_1 Z)^{(t_0+s)}\right) \\
=&\sum_{s=0}^{S-2} \mathrm{Var}_{P^{(t_0)}}\left((\mathrm{D}_1^\circ Z_{t_0})^{(s)}\right) && (100) \\
\le&(1+2\zeta S)^2\alpha \cdot \sum_{s=0}^{S-1} \mathrm{Var}_{P^{(t_0)}}\left(Z_{t_0}^{(s)}\right) + \sum_{j=1}^{n} P^{(t_0)}(j) \cdot \frac{\nu_{t_0,j}}{(1+2\zeta S)^2\alpha} && (101) \\
\le&(1+2\zeta S)^2\alpha \sum_{s=0}^{S-1} \mathrm{Var}_{P^{(t_0)}}\left(Z_{t_0}^{(s)}\right) + \frac{\mu_{t_0}}{(1+2\zeta S)\alpha} + \frac{\mathrm{Var}_{P^{(t_0)}}\left((\mathrm{D}_2^\circ Z)^{(t_0+S-2)}\right) + \mathrm{Var}_{P^{(t_0)}}\left((\mathrm{D}_2^\circ Z)^{(t_0+S-1)}\right)}{(1+2\zeta S)^2\alpha}, && (102)
\end{aligned}$$

where (100) follows from (96), (101) follows from (99) and (95), and (102) follows from (98). Summing the above for $1 \le t_0 \le T-S+1$, we obtain, for some constant $C > 0$,

$$
\begin{aligned}
&(S-1)\cdot\sum_{t=1}^{T-1}\mathrm{Var}_{P^{(t)}}\left((\mathrm{D}_1 Z)^{(t)}\right)\\
\le&\sum_{t_0=1}^{T-S+1}\sum_{s=0}^{S-2}\mathrm{Var}_{P^{(t_0+s)}}\left((\mathrm{D}_1 Z)^{(t_0+s)}\right)+8(S-1)^2M^2 &(103)\\
\le&\sum_{t_0=1}^{T-S+1}(1+2\zeta S)\sum_{s=0}^{S-2}\mathrm{Var}_{P^{(t_0)}}\left((\mathrm{D}_1 Z)^{(t_0+s)}\right)+8(S-1)^2M^2 &(104)\\
\le&(1+2\zeta S)^3\alpha\sum_{t_0=1}^{T-S+1}\sum_{s=0}^{S-1}\mathrm{Var}_{P^{(t_0)}}\left(Z_{t_0}^{(s)}\right)+\sum_{t_0=1}^{T-S+1}\frac{\mu_{t_0}}{\alpha}+8(S-1)^2M^2\\
&+\sum_{t_0=1}^{T-S+1}\frac{\mathrm{Var}_{P^{(t_0)}}\left((\mathrm{D}_2^{\circ} Z)^{(t_0+S-2)}\right)+\mathrm{Var}_{P^{(t_0)}}\left((\mathrm{D}_2^{\circ} Z)^{(t_0+S-1)}\right)}{(1+2\zeta S)\alpha} &(105)\\
\le&(1+2\zeta S)^4\alpha\sum_{t_0=1}^{T-S+1}\sum_{s=0}^{S-1}\mathrm{Var}_{P^{(t_0+s)}}\left(Z^{(t_0+s)}\right)+\sum_{t_0=1}^{T-S+1}\frac{\mu_{t_0}}{\alpha}+8(S-1)^2M^2\\
&+\frac{4}{(1+2\zeta S)\alpha}\sum_{t_0=1}^{T-S+1}\Big[\mathrm{Var}_{P^{(t_0)}}\left(Z^{(t_0+S-2)}\right)+3\,\mathrm{Var}_{P^{(t_0)}}\left(Z^{(t_0+S-1)}\right)\\
&+3\,\mathrm{Var}_{P^{(t_0)}}\left(Z^{(t_0)}\right)+\mathrm{Var}_{P^{(t_0)}}\left(Z^{(t_0+1)}\right)\Big] &(106)\\
\le&(1+2\zeta S)^4\alpha S\sum_{t=1}^{T}\mathrm{Var}_{P^{(t)}}\left(Z^{(t)}\right)+\sum_{t_0=1}^{T-S+1}\frac{\mu_{t_0}}{\alpha}+8(S-1)^2M^2\\
&+\frac{32}{(1+2\zeta S)\alpha}\sum_{t=1}^{T}\mathrm{Var}_{P^{(t)}}\left(Z^{(t)}\right) &(107)\\
\le&(1+2\zeta S)^4\alpha S\cdot\left(1+\frac{32}{\alpha^2 S}\right)\sum_{t=1}^{T}\mathrm{Var}_{P^{(t)}}\left(Z^{(t)}\right)+\sum_{t_0=1}^{T-S+1}\frac{\mu_{t_0}}{\alpha}+8(S-1)^2M^2 &(108)\\
\le&(1+2\zeta S)^4\alpha S\cdot\left(1+\frac{32}{\alpha^2 S}\right)\sum_{t=1}^{T}\mathrm{Var}_{P^{(t)}}\left(Z^{(t)}\right)+\frac{(S-2)\mu}{\alpha}+10(S-1)^2M^2, &(109)
\end{aligned}
$$

where:

- (103) follows since $\|Z^{(t)}\|_\infty \le M$ and thus $\|\,(\mathrm{D}_1 Z)^{(t)}\,\|_\infty \le 2M$ for all t;
- (104) follows from Lemma C.1 and the fact that for $0 \le s \le S-2$, $\max\left\{\left\|\frac{P^{(t_0+s)}}{P^{(t_0)}}\right\|_\infty, \left\|\frac{P^{(t_0)}}{P^{(t_0+s)}}\right\|_\infty\right\} \le 1+2\zeta S$ as established above as a consequence of the fact that the distributions $P^{(t)}$ are ζ-consecutively close.
- (105) follows from (102);

- The first term in (106) is bounded using Lemma C.1 and the fact that the distributions $P^{(t)}$ are ζ-consecutively close, and the the final term in (106) is bounded using the fact that for any vectors $Z_1, \dots, Z_k \in \mathbb{R}^n$ and any $P \in \Delta^n$, we have $\mathrm{Var}_P(Z_1 + \cdots + Z_k) \le k \cdot (\mathrm{Var}_P(Z_1) + \cdots + \mathrm{Var}_P(Z_k))$;
- (107) and (108) by rearranging terms;
- (109) follows from (93).

Now choose $S = \lceil \frac{128}{\alpha^3} \rceil$, so that $\frac{32}{\alpha^2 S} \le \frac{\alpha}{4}$. Therefore, as long as $2\zeta S \le \frac{\alpha}{32}$, we have, since $\alpha \le 1/2$, that

$$(1 + 2\zeta S)^4 \alpha \cdot \frac{S}{S-1} \cdot \left(1 + \frac{32}{\alpha^2 S}\right) \le \alpha \cdot (1 + \alpha/4)^3 \le \alpha \cdot (1 + \alpha).$$

Then it follows from (109) that

$$\sum_{t=1}^{T-1} \mathrm{Var}_{P^{(t)}}\left((\mathrm{D}_1 Z)^{(t)}\right) \le \alpha(1+\alpha) \cdot \sum_{t=1}^{T} \mathrm{Var}_{P^{(t)}}\left(Z^{(t)}\right) + \frac{\mu}{\alpha} + 10SM^2, \tag{110}$$

Using that $S \le \frac{129}{\alpha^3}$, the inequality $2\zeta S \le \alpha/32$ can be satisfied by ensuring that $\zeta \le \frac{\alpha^4}{8256} = \frac{\alpha^4}{129 \cdot 42} \le \frac{\alpha}{64S}$. Note that our choice of S ensures that $\zeta S \le 1/2$, as was assumed earlier. Moreover, we have $10SM^2 \le \frac{1290M^2}{\alpha^3}$. Thus, (110) gives the desired result. □

C.3 Completing the proof of Theorem 3.1

Using the lemma developed in the previous sections we now can complete the proof of Theorem 3.1. We begin by proving Lemma 3.3. The lemma is restated formally below.

Lemma 3.3 [Detailed] *There are constants $C, C' > 1$ so that the following holds. Suppose a time horizon $T \ge 4$ is given, we set $H := \lceil \log T \rceil$, and all players play according to OMWU with step size η satisfying $1/T \le \eta \le \frac{1}{C \cdot mH^4}$. Then for any $i \in [m]$, the losses $\ell_i^{(1)}, \dots, \ell_i^{(T)} \in [0,1]^{n_i}$ for player i satisfy:*

$$\sum_{t=1}^{T} \mathrm{Var}_{x_i^{(t)}}\left(\ell_i^{(t)} - \ell_i^{(t-1)}\right) \le \frac{1}{2} \cdot \sum_{t=1}^{T} \mathrm{Var}_{x_i^{(t)}}\left(\ell_i^{(t-1)}\right) + C' H^5. \tag{111}$$

We state a generic version of this lemma that can be applied in more general settings.

Lemma C.4. *For any integers $n \ge 2$ and $T \ge 4$, we set $H := \lceil \log T \rceil$, $\alpha = 1/(4H)$, and $\alpha_0 = \frac{\sqrt{\alpha/8}}{H^3}$. Suppose that $Z^{(1)}, \dots, Z^{(T)} \in [0,1]^n$ and $P^{(1)}, \dots, P^{(T)} \in \Delta^n$ satisfy the following*

1. *For each $0 \le h \le H$ and $1 \le t \le T - h$, it holds that $\left\|(\mathrm{D}_h Z)^{(t)}\right\|_\infty \le H \cdot \left(\alpha_0 H^3\right)^h$*
2. *The sequence $P^{(1)}, \dots, P^{(T)}$ is ζ-consecutively close for some $\zeta \in [1/(2T), \alpha^4/8256]$.*

Then,

$$\sum_{t=1}^{T} \mathrm{Var}_{P^{(t)}}\left(Z^{(t)} - Z^{(t-1)}\right) \leq 2\alpha \sum_{t=1}^{T} \mathrm{Var}_{P^{(t)}}\left(Z^{(t-1)}\right) + 165120(1+\zeta)H^5 + 2 \tag{112}$$

where we say that a sequence $P^{(1)}, \ldots, P^{(T)}$ is ζ-consecutively close iff for all t

$$\max\left\{\left\|\frac{P^{(t)}}{P^{(t+1)}}\right\|_\infty, \left\|\frac{P^{(t+1)}}{P^{(t)}}\right\|_\infty\right\} \leq 1+\zeta$$

First, we demonstrate how Lemma C.4 implies Lemma 3.3.

Proof of Lemma 3.3. We hope to apply Lemma C.4 with $n = n_i$, $P^{(t)} = x_i^{(t)}$ and $Z^{(t)} = \ell_i^{(t)}$, as well as $\zeta = 7\eta$. To do so, we must verify that the preconditions of Lemma C.4 hold when our sequences arise from the dynamics of players playing OMWU with step size η satisfying $1/T \leq \eta \leq \frac{1}{C \cdot m H^4}$.

Set $C_1 = 8256$ (note that C_1 is the constant appearing in item 2 of Lemma C.4 and item 1 of Lemma 3.9 in Section C.2). Our assumption that $\eta \leq \frac{1}{C \cdot m H^4}$ implies that, as long as the constant C satisfies $C \geq 4^4 \cdot 7 \cdot C_1 = 14794752$,

$$\eta \leq \min\left\{\frac{\alpha^4}{7C_1}, \frac{\alpha_0}{36e^5 m}\right\}. \tag{113}$$

To verify precondition 1, we apply Lemma 3.6 with the parameter α in the lemma set to α_0: a valid selection as $\alpha_0 < \frac{1}{H+3}$. We conclude that, for each $i \in [m]$, $0 \leq h \leq H$ and $1 \leq t \leq T-h$, it holds that $\left\|\left(\mathrm{D}_h\, \ell_i\right)^{(t)}\right\|_\infty \leq H \cdot \left(\alpha_0 H^3\right)^h$ since $\eta \leq \frac{\alpha_0}{36e^5 m}$ as required by the lemma. To verify precondition 2, we first confirm that our selection of $\zeta = 7\eta$ places it in the desired interval $[1/(2T), \alpha^4/C_1]$ as $\eta \leq \frac{\alpha^4}{7C_1}$. By the definition of the OMWU updates, for all $i \in [m]$ and $1 \leq t \leq T$, we have $\max\left\{\left\|\frac{x_i^{(t)}}{x_i^{(t+1)}}\right\|_\infty, \left\|\frac{x_i^{(t+1)}}{x_i^{(t)}}\right\|_\infty\right\} \leq \exp(6\eta)$. Thus, the sequence $x_i^{(1)}, \ldots, x_i^{(T)}$ is (7η)-consecutively close (since $\exp(6\eta) \leq 1 + 7\eta$ for η satisfying (113)). Therefore, Lemma C.4 applies and we have

$$\begin{aligned}\sum_{t=1}^{T} \mathrm{Var}_{x_i^{(t)}}\left(\ell_i^{(t)} - \ell_i^{(t-1)}\right) &\leq 2\alpha \sum_{t=1}^{T} \mathrm{Var}_{x_i^{(t)}}\left(\ell_i^{(t-1)}\right) + 165120(1+7\eta)H^5 + 2 \\ &\leq \frac{1}{2} \cdot \sum_{t=1}^{T} \mathrm{Var}_{x_i^{(t)}}\left(\ell_i^{(t-1)}\right) + C' H^5\end{aligned}$$

for $C' = 2 + 165120(1 + 7/8256) = 165262$, as desired. □

So, it suffices to prove Lemma C.4.

Proof of Lemma C.4. Set $\mu = H \cdot \left(\alpha_0 H^3\right)^H$. Therefore, from item 1 we have,

$$\sum_{t=1}^{T-H} \mathrm{Var}_{P^{(t)}}\left(\left(\mathrm{D}_H\, Z\right)^{(t)}\right) \leq \alpha \cdot \sum_{t=1}^{T-H+1} \mathrm{Var}_{P^{(t)}}\left(\left(\mathrm{D}_{H-1}\, Z\right)^{(t)}\right) + \mu^2 T. \tag{114}$$

We will now prove, via reverse induction on h, that for all h satisfying $H-1 \geq h \geq 0$,

$$\sum_{t=1}^{T-h-1} \operatorname{Var}_{P^{(t)}}\left((\mathrm{D}_{h+1} Z)^{(t)}\right) \leq \alpha \cdot (1+2\alpha)^{H-h-1} \cdot \sum_{t=1}^{T-h} \operatorname{Var}_{P^{(t)}}\left((\mathrm{D}_h Z)^{(t)}\right) + \frac{2C_0 \cdot H^2 \cdot \left(2\alpha_0 H^3\right)^{2h}}{\alpha^3}. \tag{115}$$

with $C_0 = 1290$ (note that C_0 is the constant appearing in the inequality (90) of the statement of Lemma 3.9 in Section C.2). The base case $h = H-1$ is verified by (114) and the fact that $2^{2(H-1)} \geq 2^H \geq T$. Now suppose that (115) holds for some h satisfying $H-1 \geq h \geq 1$. We will now apply Lemma 3.9, with $P^{(t)} = P^{(t)}$ and $Z^{(t)} = (\mathrm{D}_{h-1} Z)^{(t)}$ for $1 \leq t \leq T-h+1$, as well as $M = H \cdot \left(2\alpha_0 H^3\right)^{h-1}$, $\zeta = \zeta$, $\mu = \frac{2C_0 \cdot H^2 \cdot \left(2\alpha_0 H^3\right)^{2h}}{\alpha^3}$, and the parameter α of Lemma 3.9 set to $\alpha \cdot (1+2\alpha)^{H-h-1}$. We verify that precondition 1 holds due to precondition 2 of Lemma C.4 and the fact that $\alpha \leq \alpha \cdot (1+2\alpha)^{H-h-1}$. Moreover, precondition 2 holds by our inductive hypothesis (115) and our choice of μ. Therefore, by Lemma 3.9 and the fact that $1 + \alpha \cdot (1+2\alpha)^H \leq 1+2\alpha$ for our choice of $\alpha = 1/(4H)$, it follows that

$$\begin{aligned}
\sum_{t=1}^{T-h} \operatorname{Var}_{P^{(t)}}\left((\mathrm{D}_h Z)^{(t)}\right) \leq & \alpha \cdot (1+2\alpha)^{H-h} \cdot \sum_{t=1}^{T-h+1} \operatorname{Var}_{P^{(t)}}\left((\mathrm{D}_{h-1} Z)^{(t)}\right) + \frac{2C_0 \cdot H^2 \cdot \left(2\alpha_0 H^3\right)^{2h}}{\alpha^4} \\
& + \frac{C_0 \cdot H^2 \cdot \left(2\alpha_0 H^3\right)^{2(h-1)}}{\alpha^3} \\
\leq & \alpha \cdot (1+2\alpha)^{H-h} \cdot \sum_{t=1}^{T-h+1} \operatorname{Var}_{P^{(t)}}\left((\mathrm{D}_{h-1} Z)^{(t)}\right) \\
& + \frac{C_0 \cdot H^2 \cdot (2\alpha_0 H^3)^{2(h-1)}}{\alpha^3} \cdot \left(1 + \frac{2(2\alpha_0 H^3)^2}{\alpha}\right) \\
\leq & \alpha \cdot (1+2\alpha)^{H-h} \cdot \sum_{t=1}^{T-h+1} \operatorname{Var}_{P^{(t)}}\left((\mathrm{D}_{h-1} Z)^{(t)}\right) + \frac{2C_0 \cdot H^2 \cdot (2\alpha_0 H^3)^{2(h-1)}}{\alpha^3},
\end{aligned}$$

where the final equality follows since α_0 is chosen so that $2(2\alpha_0 H^3)^2 = \alpha$. This completes the proof of the inductive step. Thus (115) holds for $h = 0$. Using again that the sequence $P^{(t)}$ is

(ζ)-exponentially close, we see that

$$\begin{aligned}
&\sum_{t=1}^{T} \mathrm{Var}_{P^{(t)}}\left(Z^{(t)} - Z^{(t-1)}\right) \\
\leq& 1 + \sum_{t=2}^{T} \mathrm{Var}_{P^{(t)}}\left(Z^{(t)} - Z^{(t-1)}\right) \\
\leq& 1 + (1+\zeta)\sum_{t=1}^{T-1} \mathrm{Var}_{P^{(t)}}\left((\mathrm{D}_1 Z)^{(t)}\right) && (116)\\
\leq& 1 + (1+\zeta)\cdot\left(\alpha(1+2\alpha)^{H-1}\sum_{t=1}^{T} \mathrm{Var}_{P^{(t)}}\left(Z^{(t)}\right) + \frac{2C_0H^2}{\alpha^3}\right) && (117)\\
\leq& 2 + (1+\zeta)\cdot\left(\alpha(1+2\alpha)^{H-1}(1+\zeta)\sum_{t=2}^{T} \mathrm{Var}_{P^{(t)}}\left(Z^{(t-1)}\right) + \frac{2C_0H^2}{\alpha^3}\right) && (118)\\
\leq& \alpha(1+2\alpha)^{H}\sum_{t=1}^{T} \mathrm{Var}_{P^{(t)}}\left(Z^{(t-1)}\right) + \frac{2(1+\zeta)C_0H^2}{\alpha^3} + 2 \\
\leq& 2\alpha\sum_{t=1}^{T} \mathrm{Var}_{P^{(t)}}\left(Z^{(t-1)}\right) + \frac{2(1+\zeta)C_0H^2}{\alpha^3} + 2, && (119)
\end{aligned}$$

where (116) and (118) follow from Lemma C.1, and (117) uses (115) for $h = 0$. Now, (119) verifies the statement of the lemma as $C_0 = 1290$ and $\alpha = \frac{1}{4H}$. □

We are finally ready to prove Theorem 3.1. For convenience the theorem is restated below.

Theorem 3.1 [Restated] *There are constants $C, C' > 1$ so that the following holds. Suppose a time horizon $T \in \mathbb{N}$ is given. Suppose all players play according to OMWU with any positive step size $\eta \leq \frac{1}{C \cdot m \log^4 T}$. Then for any $i \in [m]$, the regret of player i satisfies*

$$\mathrm{Reg}_i^T \leq \frac{\log n_i}{\eta} + C' \cdot \log T. \tag{120}$$

In particular, if the players' step size is chosen as $\eta = \frac{1}{C \cdot m \log^4 T}$, then the regret of player i satisfies

$$\mathrm{Reg}_i^T \leq O\left(m \cdot \log n_i \cdot \log^4 T\right). \tag{121}$$

Proof. The conclusion of the theorem is immediate if $T < 4$, so we may assume from here on that $T \geq 4$. Moreover, the conclusion of (120) is immediate if $\eta \leq 1/T$ (as $\mathrm{Reg}_i^T \leq T$ necessarily), so we may also assume that $\eta \geq 1/T$. Let C'' be the constant C of Lemma 3.2, let B be the constant called C in Lemma 3.3 and B' be the constant called C' in Lemma 3.3. As long as the constant C of Theorem 3.1 is chosen so that $C \geq B$ and $\eta \leq \frac{1}{C \cdot m \log^4 T}$ implies that $C''\eta \leq 1/6$, we have the

following:

$$\begin{aligned}
\mathrm{Reg}_i^T \leq & \frac{\log n_i}{\eta} + \sum_{t=1}^{T} \left(\frac{\eta}{2} + C\eta^2\right) \mathrm{Var}_{x_i^{(t)}}\left(\ell_i^{(t)} - \ell_i^{(t-1)}\right) - \sum_{t=1}^{T} \frac{(1-C\eta)\eta}{2} \mathrm{Var}_{x_i^{(t)}}\left(\ell_i^{(t-1)}\right) && (122)\\
\leq & \frac{\log n_i}{\eta} + \frac{2\eta}{3} \sum_{t=1}^{T} \mathrm{Var}_{x_i^{(t)}}\left(\ell_i^{(t)} - \ell_i^{(t-1)}\right) - \frac{\eta}{3} \sum_{t=1}^{T} \mathrm{Var}_{x_i^{(t)}}\left(\ell_i^{(t-1)}\right)\\
\leq & \frac{\log n_i}{\eta} + \frac{2\eta}{3} \cdot \left(B' \cdot (2\log T)^5\right) && (123)\\
\leq & \frac{\log n_i}{\eta} + 32B' \cdot \log T, && (124)
\end{aligned}$$

where (122) follows from Lemma 3.2, (123) follows from Lemma 3.3, and (124) follows from the upper bound $\eta \leq \frac{1}{Cm\log^4 T}$. We have thus established (120). The upper bound (121) follows immediately. □

D Adversarial regret bounds

In this section we discuss how OMWU can be modified to achieve an algorithm that obtains the fast rates of Theorem 3.1 when played by all players, and which still obtains the optimal rate of $O(\sqrt{T})$ in the adversarial setting. Such guarantees are common in the literature [DDK11; RS13b; Syr+15; Kan+18; HAM21]. The guarantees of this modification of OMWU are stated in the following corollary (of Lemmas 3.2 and 3.3):

Corollary D.1. *There is an algorithm $\mathcal{A}$ which, if played by all m players in a game, achieves the regret bound of $\mathrm{Reg}_i^T \leq O(m \cdot \log n_i \cdot \log^4 T)$ for each player i; moreover, when player i is faced with an* adversarial *sequence of losses, the algorithm $\mathcal{A}$'s regret bound is $\mathrm{Reg}_i^T \leq O(m \log n_i \cdot \log^4 T + \sqrt{T \log n_i})$.*

Proof. Let C be the constant called C in Theorem 3.1 and C' be the constant called C' in Lemma 3.3. The algorithm $\mathcal{A}$ of Corollary D.1 is obtained as follows:

1. Initially run OMWU, with the step-size $\eta = \frac{1}{Cm\log^4 T}$.
2. If, for some $T_0 \geq 4$, (111) first fails to hold at time T_0, i.e.,

$$\sum_{t=1}^{T_0} \mathrm{Var}_{x_i^{(t)}}\left(\ell_i^{(t)} - \ell_i^{(t-1)}\right) > \frac{1}{2} \cdot \sum_{t=1}^{T_0} \mathrm{Var}_{x_i^{(t)}}\left(\ell_i^{(t-1)}\right) + C' \lceil \log T \rceil^5, \tag{125}$$

   then set $\eta' = \sqrt{\frac{\log n_i}{T}}$ and continue on running OMWU with step size η'.

If there is no $T_0 \geq 4$ so that (125) holds (and by Lemma 3.3, this will be the case when $\mathcal{A}$ is played by all m players in a game), then the proof of Theorem 3.1 shows that the regret of each player i is bounded as $\mathrm{Reg}_i^T \leq O(m \log n_i \cdot \log^4 T)$. Otherwise, since T_0 is defined as the smallest integer at least 4 so that (125) holds, we have

$$\sum_{t=1}^{T_0} \mathrm{Var}_{x_i^{(t)}}\left(\ell_i^{(t)} - \ell_i^{(t-1)}\right) \leq \frac{1}{2} \cdot \sum_{t=1}^{T_0} \mathrm{Var}_{x_i^{(t)}}\left(\ell_i^{(t-1)}\right) + C' \lceil \log T \rceil^5 + 4,$$

and thus, by Lemma 3.2, for any $x^\star \in \Delta^{n_i}$,

$$\sum_{t=1}^{T_0} \langle \ell_i^{(t)}, x_i^{(t)} - x^\star \rangle \le \mathrm{Reg}_i^{T_0} \le O(m \log n_i \cdot \log^4 T_0). \tag{126}$$

Further, by the choice of step size $\eta' = \sqrt{\frac{\log n_i}{T}}$ for time steps $t > T_0$, we have, for any $x^\star \in \Delta^{n_i}$,

$$\sum_{t=T_0+1}^{T} \langle \ell_i^{(t)}, x_i^{(t)} - x^\star \rangle \le \frac{\log n_i}{\eta'} + \eta' \sum_{t=T_0+1}^{T} \|\ell_i^{(t)} - \ell_i^{(t-1)}\|_\infty^2 \tag{127}$$

$$\le \frac{\log n_i}{\eta'} + \eta' T \le O(\sqrt{T \log n_i}), \tag{128}$$

where (127) uses [Syr+15, Proposition 7]. Adding (126) and (128) completes the proof of the corollary. □

E Phi-Regret Minimization

In this section we give an overview of the no-internal-regret algorithm of Stoltz and Lugosi [SL05] and the no-swap-regret algorithm of Blum and Mansour [BM07]. To do so, we will use the template developed by Gordon, Greenwald, and Marks [GGM08] for performing Phi-regret minimization.

E.1 From Internal Regret to External Regret

In Section 2 we discussed the implications of having access to a no-internal-regret algorithm, namely Theorem 2.4, assuming that such objects indeed exist. In the sequel, this presumption will be justified by presenting an explicit construction that leads to such guarantees. In the sequel, we present the reduction from internal to external regret based on the template of Gordon, Greenwald, and Marks [GGM08] for performing Φ-regret minimization for a generic set Φ. We remark that the subsequent construction is equivalent to the one appearing in the book of Cesa-Bianchi and Lugosi [CL06b, pp. 84], which in turn is based on the work of Stoltz and Lugosi [SL05].

In this setting, the construction of Gordon, Greenwald, and Marks [GGM08] shows that Φ-regret minimization can be achieved starting from the following two ingredients:

1. An *external* regret minimizer for the set of transformations Φ;
2. Establishing an oracle that (efficiently) computes a fixed point for each $\phi \in \Phi$ (assuming that such a point exists).

In this subsection we are interested in the case $\Phi = \Phi_i^{\mathrm{int}}$ for each player $i \in [\![m]\!]$. Let us focus first on item 1. In particular, recalling the definition of the matrix $\mathbf{E}_{j \to k}$ associated with the linear transformation $\phi_{j \to k}$ given in (2), for all $j \neq k \in [\![n_i]\!]$, it is easy to see that one can construct a regret minimizer for the set Φ^{int} using a regret minimizer for the simplex $\Delta^{n_i(n_i-1)}$. This construction is summarized in Algorithm 3, and the formal guarantee is stated below:

Proposition E.1. *If the regret minimizer $\mathcal{R}_\Delta$ for the set $\Delta^{n_i(n_i-1)}$ incurs regret at most Reg_i^T, then the regret minimizer $\mathcal{R}_\Phi$ for the set Φ_i^{int} in Algorithm 3 has also regret bounded by Reg_i^T.*

Algorithm 3: Regret minimizer for the set Φ_i^{int}

Input: A regret minimizer $\mathcal{R}_\Delta$ for the simplex $\Delta^{n_i(n_i-1)}$

function NextStrategy()
  $\boldsymbol{p}^{(t)} \leftarrow \mathcal{R}_\Delta$.NextStrategy()
  return $\sum_{j \neq k} \boldsymbol{p}^{(t)}[j \to k]\, \mathbf{E}_{j \to k}$

function ObserveUtility($\mathbf{L}^{(t)}$)
  for $j \neq k \in [\![n]\!]$ **do**
    $\boldsymbol{L}^{(t)}[j \to k] \leftarrow \langle \mathbf{L}^{(t)}, \mathbf{E}_{j \to k} \rangle$
  $\mathcal{R}_\Delta$.ObserveUtility($\boldsymbol{L}^{(t)}$)

Algorithm 4: Internal Regret minimizer for the set Δ^{n_i}

Data: A regret minimizer $\mathcal{R}_\Phi$ for the set Φ_i^{int}, *e.g.*, via Algorithm 3

function NextStrategy()
  $\phi^{(t)} \leftarrow \mathcal{R}_\Phi$.NextStrategy()
  $\boldsymbol{x}^{(t)} \leftarrow$ StationDistrib($\phi^{(t)}$)
  return $\boldsymbol{x}^{(t)}$

function ObserveUtility($\boldsymbol{\ell}^{(t)}$)
  Construct the loss $\mathbf{L}^{(t)} \leftarrow \boldsymbol{\ell}^{(t)}(\boldsymbol{x}^{(t)})^\top$
  $\mathcal{R}_\Phi$.ObserveUtility($\mathbf{L}^{(t)}$)

Next, we turn our attention to item 2: establishing a fixed point oracle for each transformation $\phi \in \Phi^{\text{int}}$. To do so, we require the following key observation:

Claim E.2. *Each transformation $\phi \in \Phi_i^{\text{int}}$ represents the transition matrix of a Markov chain.*

Proof. Fix $j \neq k \in [\![n_i]\!]$. It is immediate to see that each column of the matrix $\mathbf{E}_{j \to k}$, defined in (2), corresponds to a point in the simplex Δ^{n_i}. Thus, the same holds for any convex combination of such matrices, concluding the proof. □

This observation ensures that each $\phi \in \operatorname{co} \Phi_i^{\text{int}}$ admits a fixed point (*e.g.*, as a corollary of the Perron-Frobenius theorem), associated with the stationary distribution of the corresponding Markov chain, and such a point can be computed in polynomial time. As a result, we have established both ingredients required in the template of Gordon, Greenwald, and Marks [GGM08], and the overall construction is presented in Algorithm 4. Formally, the guarantee of this reduction is stated in the following proposition:

Proposition E.3 (From External to Internal Regret)**.** *If the external regret of $\mathcal{R}_\Phi$ is bounded by $\operatorname{Reg}^T_{\Phi_i}$, then the internal regret of Algorithm 4 is also bounded by $\operatorname{Reg}^T_{\Phi_i}$.*

Proof. The conclusion follows directly from [GGM08, Theorem 1]. □

E.2 From Swap Regret to External Regret

Next, we give an overview of the algorithm of Blum and Mansour [BM07]. In particular, in this case $\Phi_i = \Phi_i^{\text{swap}}$ is the set of *all* linear transformations from Δ^{n_i} to Δ^{n_i}. Following the construction in Gordon, Greenwald, and Marks [GGM08], let us first focus on item 1: constructing an external regret minimizer for the set Φ_i^{swap}. We will use the following well-known characterization.

Fact E.4. *Any linear function $\Phi_i^{swap} \ni \phi : \Delta^{n_i} \to \Delta^{n_i}$ can be expressed in the form $\phi(\boldsymbol{x}) = \mathbf{Q}^\top \boldsymbol{x}$, where $\mathbf{Q}$ is a (row) stochastic matrix.*

Moreover, the set of all column stochastic matrices can be expressed as a concatenation of n_i independent columns, each of which is a vector in Δ^{n_i}. As a result, an external regret minimizer for the set Φ_i^{swap} can be constructed by simply applying a "regret circuit" for the Cartesian product

[FKS19]. Next, focusing on item 2, computing the fixed point of each $\phi \in \Phi_i^{\text{swap}}$ is also direct since ϕ is associated with a Markov chain (Fact E.4). With these ingredients at hand we are ready to present the no-swap-regret algorithm of Blum and Mansour [BM07] in Algorithm 5.

Algorithm 5: Blum and Mansour [BM07]

Input: Regret minimizers $\mathcal{R}_{\Delta,1}, \dots, \mathcal{R}_{\Delta,n_i}$ for the simplex Δ^{n_i}

function NextStrategy()
  $\mathbf{Q}^{(t)} \leftarrow \mathbf{0} \in \mathbb{R}^{n_i \times n_i}$
  for $g \in [\![n_i]\!]$ **do**
    $\mathbf{Q}^{(t)}[g, \cdot] \leftarrow \mathcal{R}_{\Delta,g}$.NextStrategy()
  $\boldsymbol{x}^{(t)} \leftarrow$ StationDistrib($\mathbf{Q}^{(t)}$)
  return $\boldsymbol{x}^{(t)}$

function ObserveUtility($\boldsymbol{\ell}^{(t)}$)
  for $g \in [\![n_i]\!]$ **do**
    $\mathcal{R}_{\Delta,g}$.ObserveUtility($\boldsymbol{x}^{(t)}[g]\, \boldsymbol{\ell}^{(t)}$)

F Analysis of the Stoltz-Lugosi Algorithm

In this section we provide all of the technical ingredients required for the analysis of SL-OMWU, and subsequently for the proof of Theorem 1.2. First, we prove Lemma 4.4, a lemma analogous to Lemma 3.6. For the convenience of the reader, the statement of the lemma is included below.

Lemma 4.4. *Consider a parameter $\alpha \leq 1/(H+3)$. If all players employ* SL-OMWU *with learning rate $\eta \leq \frac{\alpha}{36e^5 m}$, then for any player $i \in [\![m]\!]$, $0 \leq h \leq H$ and $t \in [\![T-h]\!]$ it holds that*

$$\|(D_h \boldsymbol{L}_i)^{(t)}\|_\infty \leq \alpha^h h^{3h+1}.$$

Proof of Lemma 4.4. First of all, we know that

$$\boldsymbol{\ell}_i^{(t)}[a_i] = \sum_{a_{i'} \in [\![n_{i'}]\!], i' \neq i} \Lambda(a_1, \dots, a_i, \dots, a_m) \prod_{i' \neq i} \boldsymbol{x}_{i'}^{(t)}[a_{i'}], \tag{129}$$

where recall that by assumption $\Lambda(\cdot) \in [0,1]$. In particular, given that $\boldsymbol{L}_i^{(t)}[j \to k] = \boldsymbol{x}_i^{(t)}[j](\boldsymbol{\ell}_i^{(t)}[k] - \boldsymbol{\ell}_i^{(t)}[j])$, we may conclude that

$$\boldsymbol{L}_i^{(t)}[j \to k] = \sum_{a_{i'} \in [\![n_{i'}]\!], i' \neq i} (\Lambda(a_1, \dots, k, \dots, a_m) - \Lambda(a_1, \dots, j, \dots, a_m)) \boldsymbol{x}_i^{(t)}[j] \prod_{i' \neq i} \boldsymbol{x}_{i'}^{(t)}[a_{i'}]. \tag{130}$$

However, we know from the equivalence shown in Theorem 4.1 that $\boldsymbol{x}_{i'}^{(t)}[a_{i'}] = \sum_{\mathcal{T}_{i'} \in \mathbb{T}_{i', a_{i'}}} \boldsymbol{X}_{i'}^{(t)}[\mathcal{T}_{i'}]$. Given that the set of directed trees with different roots are disjoint, (130) can be expressed as

$$\boldsymbol{L}_i^{(t)}[j \to k] = \sum_{\mathcal{T}_{i'} \in \mathbb{T}_{i'}} \Lambda'(\mathcal{T}_1, \dots, \mathcal{T}_m) \prod_{i' \in [\![m]\!]} \boldsymbol{X}_{i'}^{(t)}[\mathcal{T}_{i'}], \tag{131}$$

where $\Lambda' : \mathbb{T}_1 \times \cdots \times \mathbb{T}_m \to [-1, 1]$. In particular, this implies that

$$\left|(D_h \boldsymbol{L}_i)^{(t)}[j \to k]\right| = \left|\sum_{s=0}^{h} \binom{h}{s} (-1)^{h-s} \boldsymbol{L}_i^{(t+s)}[j \to k]\right| \tag{132}$$

$$= \left|\sum_{\mathcal{T}_{i'} \in \mathbb{T}_{i'}} \Lambda'(\mathcal{T}_1, \ldots, \mathcal{T}_m) \sum_{s=0}^{h} \binom{h}{s} (-1)^{h-s} \prod_{i' \in [\![m]\!]} \boldsymbol{X}_{i'}^{(t+s)}[\mathcal{T}_{i'}]\right| \tag{133}$$

$$\leq \sum_{\mathcal{T}_{i'} \in \mathbb{T}_{i'}} \left|\sum_{s=0}^{h} \binom{h}{s} (-1)^{h-s} \prod_{i' \in [\![m]\!]} \boldsymbol{X}_{i'}^{(t+s)}[\mathcal{T}_{i'}]\right| \tag{134}$$

$$= \sum_{\mathcal{T}_{i'} \in \mathbb{T}_{i'}} \left|\left(D_h \left(\prod_{i' \in [\![m]\!]} \boldsymbol{X}_{i'}[\mathcal{T}_{i'}]\right)\right)^{(t)}\right|, \tag{135}$$

where (132) uses the equivalent formulation (12) of h-order finite differences; (133) follows from (131); (134) follows from the triangle inequality and the fact that $|\Lambda'(\cdot)| \in [0, 1]$; and the final line uses again the equivalent formulation of finite differences of (12), with the convention that $\prod_{i' \in [\![m]\!]} \boldsymbol{X}_{i'}[\mathcal{T}_{i'}]$ refers to the sequence $\prod_{i' \in [\![m]\!]} \boldsymbol{X}_{i'}^{(1)}[\mathcal{T}_{i'}], \ldots, \prod_{i' \in [\![m]\!]} \boldsymbol{X}_{i'}^{(T)}[\mathcal{T}_{i'}]$. Next, it will be convenient to assume that $\boldsymbol{X}_i^{(0)}$ is the uniform distribution over the simplex $\Delta^{|\mathbb{T}_i|}$, as well as $\boldsymbol{\mathcal{L}}_i^{(0)} = \boldsymbol{\mathcal{L}}_i^{(-1)} = \mathbf{0}$. By construction, for every player $i \in [\![m]\!]$ the vector $\boldsymbol{X}_i^{(t)}$ is updated using (OMWU) under the sequence of losses $\boldsymbol{\mathcal{L}}^{(1)}, \ldots, \boldsymbol{\mathcal{L}}^{(T)}$, implying that for each $\mathcal{T} \in \mathbb{T}_i$,

$$\boldsymbol{X}_i^{(t_0+t+1)}[\mathcal{T}] = \frac{\exp\left\{\eta\left(\boldsymbol{\mathcal{L}}_i^{(t_0-1)}[\mathcal{T}] - \sum_{s=0}^{t} \boldsymbol{\mathcal{L}}_i^{(t_0+s)}[\mathcal{T}] - \boldsymbol{\mathcal{L}}_i^{(t_0+t)}[\mathcal{T}]\right)\right\} \boldsymbol{X}_i^{(t_0)}[\mathcal{T}]}{\sum_{\mathcal{T}' \in \mathbb{T}_i} \exp\left\{\eta\left(\boldsymbol{\mathcal{L}}_i^{(t_0-1)}[\mathcal{T}'] - \sum_{s=0}^{t} \boldsymbol{\mathcal{L}}_i^{(t_0+s)}[\mathcal{T}'] - \boldsymbol{\mathcal{L}}_i^{(t_0+t)}[\mathcal{T}']\right)\right\} \boldsymbol{X}_i^{(t_0)}[\mathcal{T}']}. \tag{136}$$

For notational convenience, we will let $\bar{\boldsymbol{\mathcal{L}}}_{i,t_0}^{(t)} = \boldsymbol{\mathcal{L}}_i^{(t_0-1)} - \sum_{s=0}^{t-1} \boldsymbol{\mathcal{L}}_i^{(t_0+s)} - \boldsymbol{\mathcal{L}}_i^{(t_0+t-1)}$, for $0 \leq t_0 \leq T$ and $t \geq 0$. Moreover, for a vector $\boldsymbol{z} \in \mathbb{R}^{|\mathbb{T}_i|}$ we define the following function:

$$\phi_{t_0,\mathcal{T}}(\boldsymbol{z}) = \frac{\exp\{\boldsymbol{z}[\mathcal{T}]\}}{\sum_{\mathcal{T}' \in \mathbb{T}_i'} \boldsymbol{X}_i^{(t_0)}[\mathcal{T}'] \exp\{\boldsymbol{z}[\mathcal{T}']\}}. \tag{137}$$

Equipped with this notation, we can equivalently write (136) as $\boldsymbol{X}_i^{(t_0+t)}[\mathcal{T}] = \boldsymbol{X}_i^{(t_0)}[\mathcal{T}] \phi_{t_0,\mathcal{T}}\left(\eta \bar{\boldsymbol{\mathcal{L}}}_{i,t_0}^{(t)}\right)$ for $t \geq 1$. In particular, this implies that for any $i' \in [\![m]\!]$ and $\mathcal{T}_{i'} \in \mathbb{T}_{i'}$,

$$\prod_{i' \in [\![m]\!]} \boldsymbol{X}_{i'}^{(t_0+t)}[\mathcal{T}_{i'}] = \prod_{i' \in [\![m]\!]} \boldsymbol{X}_{i'}^{(t_0)}[\mathcal{T}_{i'}] \phi_{t_0,\mathcal{T}_{i'}}\left(\eta \bar{\boldsymbol{\mathcal{L}}}_{i',t_0}^{(t)}\right). \tag{138}$$

Let us recall Definition B.1: the definition of the *shift operator*. For a sequence of vectors $\boldsymbol{z} = (\boldsymbol{z}^{(1)}, \ldots, \boldsymbol{z}^{(T)})$, the s-shifted sequence, denoted with $E_s \boldsymbol{z}$, is such that $(E_s \boldsymbol{z})^{(t)} = \boldsymbol{z}^{(t+s)}$, for $1 \leq t \leq T - s$. With this notation, we observe that

$$\left(D_1 \bar{\boldsymbol{\mathcal{L}}}_{i,t_0}\right)^{(t)} = \boldsymbol{\mathcal{L}}_i^{(t_0+t-1)} - 2\boldsymbol{\mathcal{L}}_i^{(t_0+t)} = \boldsymbol{\mathcal{L}}_i^{(t_0+t-1)} - 2(E_1 \boldsymbol{\mathcal{L}}_i)^{(t_0+t-1)},$$

which in particular implies that for any $h' \geq 1$,

$$\left(D_{h'}\bar{\boldsymbol{\mathcal{L}}}_{i,t_0}\right)^{(t)} = (D_{h'-1}\boldsymbol{\mathcal{L}}_i)^{(t_0+t-1)} - 2(E_1 D_{h'-1}\boldsymbol{\mathcal{L}}_i)^{(t_0+t-1)}. \tag{139}$$

Next we proceed with bounding $\bar{\boldsymbol{\mathcal{L}}}^{(t)}_{i,t_0}$, for any $0 \leq t_0 \leq T$ and $t \geq 0$. In particular, for a fixed $\mathcal{T} \in \mathbb{T}_i$, we know that $\boldsymbol{\mathcal{L}}^{(t)}_i[\mathcal{T}] = \sum_{(j,k)\in E(\mathcal{T})} \boldsymbol{x}^{(t)}_i[j](\boldsymbol{\ell}^{(t)}_i[k] - \boldsymbol{\ell}^{(t)}_i[j])$. Thus, given that $\boldsymbol{\ell}^{(t)}_i[j] \in [0,1]$, for any $j \in [\![n_i]\!]$ and $t \geq 0$, as follows from (129), we can conclude that there exists $\mathcal{T} \in \mathbb{T}_i$ such that

$$\|\boldsymbol{\mathcal{L}}^{(t)}_i\|_\infty = \left|\sum_{(j,k)\in E(\mathcal{T})} \boldsymbol{x}^{(t)}_i[j](\boldsymbol{\ell}^{(t)}_i[k] - \boldsymbol{\ell}^{(t)}_i[j])\right| \leq \sum_{(j,k)\in E(\mathcal{T})} \boldsymbol{x}^{(t)}_i[j]\left|\boldsymbol{\ell}^{(t)}_i[k] - \boldsymbol{\ell}^{(t)}_i[j]\right| \leq 1, \tag{140}$$

where we used that $\left|\boldsymbol{\ell}^{(t)}_i[k] - \boldsymbol{\ell}^{(t)}_i[j]\right| \leq 1$, as well as the fact that $\mathcal{T}$ is a directed tree, implying that every node has at most one outgoing (directed) edge (recall Definition 2.5). Along with the triangle inequality, this implies that

$$\|\bar{\boldsymbol{\mathcal{L}}}^{(t)}_{i,t_0}\|_\infty = \left\|\boldsymbol{\mathcal{L}}^{(t_0-1)}_i - \sum_{s=0}^{t-1} \boldsymbol{\mathcal{L}}^{(t_0+s)}_i - \boldsymbol{\mathcal{L}}^{(t_0+t-1)}_i\right\|_\infty \leq (t+2). \tag{141}$$

Before we proceed with the next claim, it will be useful to introduce the sequence $\mathfrak{L}^{(t)}$, defined as

$$\mathfrak{L}^{(t)} \coloneqq \left(\eta\bar{\boldsymbol{\mathcal{L}}}^{(t)}_{1,t_0}, \eta\bar{\boldsymbol{\mathcal{L}}}^{(t)}_{2,t_0}, \ldots, \eta\bar{\boldsymbol{\mathcal{L}}}^{(t)}_{m,t_0}\right).$$

We also restate the "boundedness chain rule" lemma for the convenience of the reader.

Lemma 3.7 ["Boundedness chain rule" for finite differences; detailed] *Suppose that $h, n \in \mathbb{N}$, $\phi : \mathbb{R}^n \to \mathbb{R}$ is a (Q,R)-bounded function so that the radius of convergence of its power series at $\mathbf{0}$ is at least $\nu > 0$, and $L = (L^{(1)}, \ldots, L^{(T)}) \in \mathbb{R}^n$ is a sequence of vectors satisfying $\|L^{(t)}\|_\infty \leq \nu$ for $t \in [T]$. Suppose for some $\alpha \in (0,1)$, for each $0 \leq h' \leq h$ and $t \in [T-h']$, it holds that $\|\,\mathrm{D}_{h'}\, L^{(t)}\|_\infty \leq \frac{1}{B_1} \cdot \alpha^{h'} \cdot (h')^{B_0 h'}$ for some $B_1 \geq 2e^2 R, B_0 \geq 3$. Then for all $t \in [T-h]$,*

$$|\,(\mathrm{D}_h\,(\phi \circ L))^{(t)}\,| \leq \frac{12RQe^2}{B_1} \cdot \alpha^h \cdot h^{B_0 h+1}.$$

Claim F.1. *Consider some parameter $\alpha \in (0,1)$ such that $\left\|\left(D_{h'}\left(\eta\bar{\boldsymbol{\mathcal{L}}}_{i,t_0}\right)\right)^{(t)}\right\|_\infty \leq \frac{1}{B_1}\alpha^{h'}(h')^{B_0 h'}$, for all $i \in [\![m]\!]$, $0 \leq h' \leq h$, and $t \in [\![h+1-h']\!]$, where $B_1 = 12e^5 m$ and $B_0 \geq 3$. Then, for any $0 \leq t_0 \in T-h-1$ and $\mathcal{T}_1 \in \mathbb{T}_1, \ldots, \mathcal{T}_m \in \mathbb{T}_m$, it holds that*

$$\left|\left(D_h\left(\prod_{i'\in[\![m]\!]} \boldsymbol{X}_{i'}[\mathcal{T}_{i'}]\right)\right)^{(t_0+1)}\right| \leq \alpha^h h^{B_0 h+1} \prod_{i'\in[\![m]\!]} \boldsymbol{X}^{(t_0)}_{i'}(\mathcal{T}_{i'}). \tag{142}$$

Proof. We will apply Lemma 3.7 with $n \coloneqq \sum_{i'\in[\![m]\!]} |\mathbb{T}_{i'}|$, time horizon $h+1$, the sequence $\mathfrak{L}^{(t)}$, and the function ϕ that maps the concatenation of $\boldsymbol{z}_{i'} \in \mathbb{R}^{|\mathbb{T}_{i'}|}$, for all $i' \in [\![m]\!]$, to the function $\prod_{i'} \phi_{t_0,\mathcal{T}_{i'}}(\boldsymbol{z}_{i'})$, where $\phi_{t_0,\mathcal{T}_{i'}}$ was defined in (137); that is,

$$\phi_{t_0}(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_m) \coloneqq \prod_{i'\in[\![m]\!]} \phi_{t_0,\mathcal{T}_{i'}}(\boldsymbol{z}_{i'}). \tag{143}$$

In this context, let us verify the conditions of Lemma 3.7. First, Lemma B.7 implies that the function ϕ_{t_0} is $(1, e^3 m)$-bounded (in the sense of Lemma 3.7), and $B_1 = 12e^5 m \geq 2e^2(e^3 m)$. Moreover, by Lemma B.8, we know that each function $\phi_{t_0,\mathcal{T}_{i'}}$ has radius of convergence—with respect to the origin $\mathbf{0}$—greater than 1, and hence, the radius of convergence of ϕ_{t_0} is also greater than 1. We also know, by assumption, that $\left\|(D_{h'}\mathfrak{L})^{(t)}\right\|_\infty \leq \frac{1}{B_1}\alpha^{h'}(h')^{B_0 h'}$, for all $0 \leq h' \leq h$ and $t \in [\![h+1-h']\!]$. As a result, Lemma 3.7 implies that

$$\left|(D_h(\phi \circ \mathfrak{L}))^{(1)}\right| \leq \frac{12e^5 m}{B_1}\alpha^h h^{B_0 h+1}. \tag{144}$$

Finally, we have that

$$\begin{aligned}
\frac{1}{\prod_{i'\in[\![m]\!]} \boldsymbol{X}_{i'}^{(t_0)}[\mathcal{T}_{i'}]}\left|\left(D_h\left(\prod_{i'\in[\![m]\!]} \boldsymbol{X}_{i'}[\mathcal{T}_{i'}]\right)\right)^{(t_0+1)}\right| &= \left|\left(D_h\left(\prod_{i'\in[\![m]\!]} \left(\phi_{t_0,\mathcal{T}_{i'}} \circ \left(\eta\bar{\boldsymbol{\mathcal{L}}}_{i',t_0}\right)\right)\right)\right)^{(1)}\right| && (145)\\
&= \left|\left(D_h\left(\phi_{t_0} \circ \left(\eta\bar{\boldsymbol{\mathcal{L}}}_{1,t_0}, \ldots, \eta\bar{\boldsymbol{\mathcal{L}}}_{m,t_0}\right)\right)\right)^{(1)}\right| && (146)\\
&= \left|D_h(\phi_{t_0} \circ \mathfrak{L})^{(1)}\right| \\
&\leq \frac{12e^5 m}{B_1}\alpha^h h^{B_0 h+1} = \alpha^h h^{B_0 h+1}, && (147)
\end{aligned}$$

where (145) follows from Equation (138); (146) simply uses the definition of ϕ_{t_0} in (143); and (147) follows from (144) (which in turn is a consequence of Lemma 3.7), as well as the fact that $B_1 = 12e^5 m$. This completes the proof. □

Claim F.2. *Fix some $1 \leq h \leq H$, a parameter $\alpha \in (0,1)$, and assume that the learning rate η is such that $\eta \leq \min\left\{\frac{\alpha}{36e^5 m}, \frac{1}{12e^5(H+3)m}\right\}$. Moreover, assume that for all $0 \leq h' < h$, $t \leq T - h'$, and $i \in [\![m]\!]$, it holds that $\|(D_{h'}\boldsymbol{\mathcal{L}}_i)^{(t)}\|_\infty \leq \alpha^{h'}(h'+1)^{B_0(h'+1)}$. Then, for all $t \in [\![T-h]\!]$ and $i \in [\![m]\!]$,*

$$\left\|(D_h\boldsymbol{\mathcal{L}}_i)^{(t)}\right\|_\infty \leq \alpha^h h^{B_0 h+1}. \tag{148}$$

Proof. Let us set $B_1 \coloneqq 12e^5 m$, so that $\eta \leq \min\left\{\frac{\alpha}{3B_1}, \frac{1}{B_1(H+3)}\right\}$. We know from (141) that for $t+2 \leq h+3$ it follows that

$$\left\|D_0\left(\eta\bar{\boldsymbol{\mathcal{L}}}_{i,t_0}\right)^{(t)}\right\|_\infty = \eta\left\|\bar{\boldsymbol{\mathcal{L}}}_{i,t_0}^{(t)}\right\| \leq \eta(t+2) \leq \frac{1}{B_1}, \tag{149}$$

where we used the fact that $\eta \leq 1/(B_1(H+3))$. Next, for $1 \leq h' \leq h$,

$$\begin{aligned}
\left\|\left(D_{h'}\left(\eta\bar{\boldsymbol{\mathcal{L}}}_{i,t_0}\right)\right)^{(t)}\right\|_\infty &\leq \eta\left\|(D_{h'-1}\boldsymbol{\mathcal{L}}_i)^{(t_0+t-1)}\right\|_\infty + 2\eta\left\|(D_{h'-1}\boldsymbol{\mathcal{L}}_i)^{(t_0+t)}\right\|_\infty && (150)\\
&\leq 3\eta\alpha^{h'-1}(h')^{B_0 h'} && (151)\\
&\leq \frac{1}{B_1}\alpha^{h'}(h')^{B_0 h'}, && (152)
\end{aligned}$$

where (150) follows from (139) and the triangle inequality; (151) is a consequence of the assumption in the claim; and (152) uses the fact that $\eta \leq \frac{\alpha}{3B_1}$. Next, given that $\boldsymbol{\mathcal{L}}_i^{(t)}[\mathcal{T}] = \sum_{(j,k)\in E(\mathcal{T})} \boldsymbol{x}_i^{(t)}[j](\boldsymbol{\ell}_i^{(t)}[k] - \boldsymbol{\ell}_i^{(t)}[j])$, where recall the expression of $\boldsymbol{\ell}_i^{(t)}$ in (129), we can infer that

$$\begin{aligned}\boldsymbol{\mathcal{L}}_i^{(t)} &= \sum_{a_{i'}\in[\![n_{i'}]\!]} \hat{\Lambda}(a_1,\ldots,a_m) \prod_{i'\in[\![m]\!]} \boldsymbol{x}_{i'}[a_{i'}] \\ &= \sum_{\mathcal{T}_{i'}\in\mathbb{T}_{i'}} \tilde{\Lambda}(\mathcal{T}_1,\ldots,\mathcal{T}_m) \prod_{i'\in[\![m]\!]} \boldsymbol{X}_{i'}[\mathcal{T}_{i'}], \end{aligned} \tag{153}$$

where $\hat{\Lambda}$ and $\tilde{\Lambda}$ are functions such that $\hat{\Lambda}(\cdot) \in [-1,1]$ and $\tilde{\Lambda}(\cdot) \in [-1,1]$. More precisely, (153) is obtained from Theorem 4.1. As a result, similarly to (135) we may conclude that

$$\left|(D_h\boldsymbol{\mathcal{L}}_i)^{(t)}[\mathcal{T}]\right| \leq \sum_{\mathcal{T}_{i'}\in\mathbb{T}_{i'}} \left|\left(D_h\left(\prod_{i'\in[\![m]\!]} \boldsymbol{X}_{i'}[\mathcal{T}_{i'}]\right)\right)^{(t)}\right|. \tag{154}$$

The next step is to invoke Claim F.1 in order to bound the induced term. Specifically, by (149) and (152) we see that its conditions are met, from which we can conclude that for $t \in [\![T-h]\!]$

$$\left\|(D_h\boldsymbol{\mathcal{L}}_i)^{(t)}\right\|_\infty \leq \sum_{\mathcal{T}_{i'}\in\mathbb{T}_{i'}} \left|\left(D_h\left(\prod_{i'\in[\![m]\!]} \boldsymbol{X}_{i'}[\mathcal{T}_{i'}]\right)\right)^{(t)}\right| \tag{155}$$

$$\leq \alpha^h h^{B_0 h+1} \sum_{\mathcal{T}_{i'}\in\mathbb{T}_{i'}} \prod_{i'\in[\![m]\!]} \boldsymbol{X}_{i'}^{(t-1)}[\mathcal{T}_{i'}] \tag{156}$$

$$= \alpha^h h^{B_0 h+1}, \tag{157}$$

where (155) follows from (153); (156) is an immediate application of Claim F.1; and (157) follows from the fact that each $\boldsymbol{X}_{i'}$ is a probability distribution over the space of directed trees $\mathbb{T}_{i'}$, for all $i' \in [\![m]\!]$, and as such the induced product distribution normalizes to 1. □

Finally, it follows from (140) that $\left\|(D_0\boldsymbol{\mathcal{L}}_i)^{(t)}\right\|_\infty \leq 1 = \alpha^0 1^{B_0 1}$, for all $i \in [\![m]\!]$. Thus, we can inductively invoke Claim F.2 for $B_0 = 3$ to infer that for all $0 \leq h \leq H$, $i \in [\![m]\!]$, and $t \in [\![T-h]\!]$ that $\left\|(D_h\boldsymbol{\mathcal{L}}_i)^{(t)}\right\|_\infty \leq \alpha^h h^{3h+1}$, as long as $\eta \leq \frac{\alpha}{36e^5 m}$ and $\alpha \leq 1/(H+3)$. Finally, following the argument given in the proof of Claim F.2, we obtain that for any $0 \leq h' \leq h$,

$$\left\|\left(D_{h'}\left(\eta\bar{\boldsymbol{\mathcal{L}}}_{i,t_0}\right)\right)^{(t)}\right\|_\infty \leq \frac{1}{B_1}\alpha^{h'}(h')^{3h'}.$$

Thus, we can invoke Claim F.1 to conclude that for any $0 \leq t_0 \leq T-h-1$ and $\mathcal{T}_1 \in \mathbb{T}_1, \ldots, \mathcal{T}_m \in \mathbb{T}_m$,

$$\left|\left(D_h\left(\prod_{i'\in[\![m]\!]} \boldsymbol{X}_{i'}[\mathcal{T}_{i'}]\right)\right)^{(t_0+1)}\right| \leq \alpha^h h^{3h+1} \prod_{i'\in[\![m]\!]} \boldsymbol{X}_{i'}^{(t_0)}(\mathcal{T}_{i'}).$$

As a result, plugging-in this bound into (135) we obtain that for $t \in [\![T-h]\!]$,

$$\begin{aligned}\left\|(D_h \boldsymbol{L}_i)^{(t)}\right\|_\infty &\le \sum_{\mathcal{T}_{i'} \in \mathbb{T}_{i'}} \left| \left(D_h \left(\prod_{i' \in [\![m]\!]} \boldsymbol{X}_{i'}[\mathcal{T}_{i'}] \right) \right)^{(t)} \right| \\ &\le \alpha^h h^{3h+1} \sum_{\mathcal{T}_{i'} \in \mathbb{T}_{i'}} \prod_{i' \in [\![m]\!]} \boldsymbol{X}_{i'}^{(t-1)}[\mathcal{T}_{i'}], \\ &= \alpha^h h^{3h+1},\end{aligned}$$

completing the proof. □

The final technical ingredient is the following lemma.

Lemma F.3. *There are universal constants $C, C' \ge 1$ so that for a time horizon $T \ge 4$ and $H := \lceil \log T \rceil$, if all players employ* SL-OMWU *with learning rate η such that $1/T \le \eta \le \frac{1}{CmH^4}$, then,*

$$\sum_{t=1}^{T} \mathrm{Var}_{\boldsymbol{p}_i^{(t)}}\left(\boldsymbol{L}_i^{(t)} - \boldsymbol{L}_i^{(t-1)}\right) \le \frac{1}{2}\mathrm{Var}_{\boldsymbol{p}_i^{(t)}}\left(\boldsymbol{L}_i^{(t-1)}\right) + C' H^5, \tag{158}$$

for any $i \in [\![m]\!]$.

Proof. Recall the following lemma from earlier.

Lemma C.4. *For any integers $n \ge 2$ and $T \ge 4$, we set $H := \lceil \log T \rceil$, $\alpha = 1/(4H)$, and $\alpha_0 = \frac{\sqrt{\alpha/8}}{H^3}$. Suppose that $Z^{(1)}, \dots, Z^{(T)} \in [0,1]^n$ and $P^{(1)}, \dots, P^{(T)} \in \Delta^n$ satisfy the following*

1. *For each $0 \le h \le H$ and $1 \le t \le T-h$, it holds that $\left\|(\mathrm{D}_h Z)^{(t)}\right\|_\infty \le H \cdot \left(\alpha_0 H^3\right)^h$*
2. *The sequence $P^{(1)}, \dots, P^{(T)}$ is ζ-consecutively close for some $\zeta \in [1/(2T), \alpha^4/8256]$.*

Then,

$$\sum_{t=1}^{T} \mathrm{Var}_{P^{(t)}}\left(Z^{(t)} - Z^{(t-1)}\right) \le 2\alpha \sum_{t=1}^{T} \mathrm{Var}_{P^{(t)}}\left(Z^{(t-1)}\right) + 165120(1+\zeta)H^5 + 2 \tag{112}$$

where we say that a sequence $P^{(1)}, \dots, P^{(T)}$ is ζ-consecutively close iff for all t

$$\max\left\{ \left\| \frac{P^{(t)}}{P^{(t+1)}} \right\|_\infty, \left\| \frac{P^{(t+1)}}{P^{(t)}} \right\|_\infty \right\} \le 1 + \zeta$$

We hope to apply Lemma C.4 with $n = n_i$, $P^{(t)} = \boldsymbol{p}_i^{(t)}$ and $Z^{(t)} = \boldsymbol{L}_i^{(t)}$, as well as $\zeta = 7\eta$. To do so, we must verify that the preconditions of the lemma hold. Set $C_1 = 8256$ (note that C_1 is the constant appearing in item 2). Our assumption that $\eta \le \frac{1}{C \cdot mH^4}$ implies that, as long as the constant C satisfies $C \ge 4^4 \cdot 7 \cdot C_1 = 14794752$,

$$\eta \le \min\left\{ \frac{\alpha^4}{7C_1}, \frac{\alpha_0}{36e^5 m} \right\}. \tag{159}$$

We verify precondition 1 from Lemma 4.4, which implies that for all $i \in [\![m]\!]$, $0 \le h \le H$, and $1 \le t \le T - h$, it holds that

$$\left\| (D_h \boldsymbol{L}_i)^{(t)} \right\|_\infty \le h\big(\alpha_0 h^3\big)^h \le H\big(\alpha_0 H^3\big)^h.$$

To verify precondition 2, we first confirm that our selection of $\zeta = 7\eta$ places it in the desired interval $[1/(2T), \alpha^4/C_1]$ as $\eta \le \frac{\alpha^4}{7C_1}$. We use the following claim.

Claim F.4 (Multiplicative Stability of (O)MWU)**.** *For every player $i \in [\![m]\!]$ the sequence $\boldsymbol{p}_i^{(1)}, \dots, \boldsymbol{p}_i^{(T)}$ is 7η-consecutively close.*

Proof. We know from the update rule of (OMWU) that

$$\boldsymbol{p}_i^{(t+1)}[j \to k] = \frac{\exp\{-\eta(2\boldsymbol{L}_i^{(t)}[j \to k] - \boldsymbol{L}_i^{(t-1)}[j \to k])\}}{\sum_{a \neq b} \exp\{-\eta(2\boldsymbol{L}_i^{(t)}[a \to b] - \boldsymbol{L}_i^{(t-1)}[a \to b])\}\boldsymbol{p}_i^{(t)}[a \to b]} \boldsymbol{p}_i^{(t)}[j \to k].$$

Thus, since $\|\boldsymbol{L}_i^{(t)}\|_\infty \le 1$, it follows that

$$\boldsymbol{p}_i^{(t+1)}[j \to k] \le \frac{\exp\{3\eta\}}{\sum_{a \neq b} \exp\{-3\eta\}\boldsymbol{p}_i^{(t)}[a \to b]} \boldsymbol{p}_i^{(t)}[j \to k] = \exp\{6\eta\}\boldsymbol{p}_i^{(t)}[j \to k],$$

and similar reasoning yields that $\boldsymbol{p}_i^{(t+1)}[j \to k] \ge \exp\{-6\eta\}\boldsymbol{p}_i^{(t)}[j \to k]$. Finally, the claim follows given that $\exp\{6\eta\} \le 1 + 7\eta$ for the η considered in the lemma. □

Therefore, Lemma C.4 applies and we have

$$\begin{aligned}\sum_{t=1}^{T} \mathrm{Var}_{\boldsymbol{p}_i^{(t)}}\left(\boldsymbol{L}_i^{(t)} - \boldsymbol{L}_i^{(t-1)}\right) &\le 2\alpha \sum_{t=1}^{T} \mathrm{Var}_{\boldsymbol{p}_i^{(t)}}\left(\boldsymbol{L}_i^{(t-1)}\right) + 165120(1+7\eta)H^5 + 2 \\ &\le \frac{1}{2} \cdot \sum_{t=1}^{T} \mathrm{Var}_{\boldsymbol{p}_i^{(t)}}\left(\boldsymbol{L}_i^{(t-1)}\right) + C'H^5\end{aligned}$$

for $C' = 2 + 165120(1 + 7/8256) = 165262$, as desired. This completes the proof of the lemma. □

Finally, we are ready to establish our main theorem, namely Theorem 1.2. Before with proceed with the proof, let us first restate the theorem for the convenience of the reader.

Theorem 1.2. *Consider a general-sum multi-player game with m players, with each player $i \in [\![m]\!]$ having n_i actions. There exists a universal constant $C > 0$ such that, when all players select strategies according to algorithm* SL-OMWU *with step size $\eta = 1/(C \cdot m \log^4 T)$, the internal regret of every player $i \in [\![m]\!]$ is bounded by $O\big(m \log n_i \log^4 T\big)$. As a result, the average product distribution of play is an $O\big((m \log n \log^4 T)/T\big)$-approximate correlated equilibrium.*

Proof. First of all, we know from Theorem 2.4 that it suffices to show that the internal regret of every player $i \in [\![m]\!]$ is bounded by $O(m \log n_i \log^4 T)$. In particular, by Proposition E.3 this reduces to showing that the external regret of the regret minimizer $\mathcal{R}_\Delta$ employed in Algorithm 1 operating over the pairs of actions is bounded by $O(m \log n_i \log^4 T)$. But the latter can be easily verified by Lemma 3.2 and Lemma F.3 (as in the proof of Theorem 3.1). □

G Analysis of the Blum-Mansour Algorithm

G.1 Near-Optimal Bounds for Swap Regret

Let us first recall our main theorem regarding algorithm BM-OMWU.

Theorem 1.4. *Consider a general-sum multi-player game with m players, with each player $i \in [\![m]\!]$ having n_i actions. There exists a universal constant $C > 0$ such that, when all players select strategies according to algorithm* BM-OMWU *with step size $\eta = 1/(C \cdot m\, n_i^3 \log^4(T))$, the swap regret of every player $i \in [\![m]\!]$ is bounded by $O(m\, n_i^4 \log n_i \log^4(T))$. As a result, the average product distribution of play is an $O\big((m\, n^4 \log n \log^4 T)/T\big)$-approximate correlated equilibrium.*

Proof. Recall our lemma bounding the external regret of OMWU in even the adversarial setting

Lemma 3.2 [Restated] *There is a constant $C > 0$ so that the following holds. Suppose we follow the OMWU updates with step size $0 < \eta < 1/C$, for an arbitrary sequence of losses $\boldsymbol{\ell}^{(1)}, \ldots, \boldsymbol{\ell}^{(T)} \in [0,1]^n$. Then for any vector $\boldsymbol{x}^\star \in \Delta^n$, it holds that*

$$\sum_{t=1}^{T} \langle \boldsymbol{x}^{(t)} - \boldsymbol{x}^\star, \boldsymbol{\ell}^{(t)} \rangle \le \frac{\log n}{\eta} + \sum_{t=1}^{T} \left(\frac{\eta}{2} + C\eta^2\right) \mathrm{Var}_{\boldsymbol{x}^{(t)}}\left(\boldsymbol{\ell}^{(t)} - \boldsymbol{\ell}^{(t-1)}\right) - \sum_{t=1}^{T} \frac{(1 - C\eta)\eta}{2} \cdot \mathrm{Var}_{\boldsymbol{x}^{(t)}}\left(\boldsymbol{\ell}^{(t-1)}\right). \tag{160}$$

We can use this bound to demonstrate that the Blum-Mansour (BM) algorithm achieves good swap-regret [BM07]. Recall that the BM algorithm maintains n copies of a no-regret algorithm $\mathcal{R}_{\Delta,1}, \ldots, \mathcal{R}_{\Delta,n}$, (we choose OMWU), where each algorithm $\mathcal{R}_{\Delta,g}$ produce iterates $\mathbf{Q}^{(t)}[g,\cdot] = (\mathbf{Q}^{(t)}[g,1], \ldots, \mathbf{Q}^{(t)}[g,n]) \in \Delta^n$ at time step t. The BM algorithm then chooses $\boldsymbol{x}^{(t)} \in \Delta^n$ satisfying $\boldsymbol{x}^{(t)} = \boldsymbol{x}^{(t)}\mathbf{Q}^{(t)}$ where $\mathbf{Q}^{(t)}$ is a matrix with g^{th} row $\mathbf{Q}^{(t)}[g,\cdot]$. Upon receiving a loss vector $\boldsymbol{\ell}^{(t)}$, the BM algorithm experiences loss of $\langle \boldsymbol{x}^{(t)}, \boldsymbol{\ell}^{(t)} \rangle$ and distributes the loss vector $\boldsymbol{x}^{(t)}[g] \cdot \boldsymbol{\ell}^{(t)}$ to $\mathcal{R}_{\Delta,g}$ for each $g \in [\![n]\!]$, which use them to update and produces iterates $\mathbf{Q}^{(t+1)}[g,\cdot]$. By Lemma 3.2, each algorithm $\mathcal{R}_{\Delta,g}$ experiences regret of

$$\begin{aligned}
&\sum_{t=1}^{T} \langle \mathbf{Q}^{(t)}[g,\cdot] - \boldsymbol{q}^\star, \boldsymbol{x}^{(t)}[g] \cdot \boldsymbol{\ell}^{(t)} \rangle \\
&\le \frac{\log n}{\eta} + \sum_{t=1}^{T} \left(\frac{\eta}{2} + C\eta^2\right) \mathrm{Var}_{\mathbf{Q}^{(t)}[g,\cdot]}\left(\boldsymbol{x}^{(t)}[g]\boldsymbol{\ell}^{(t)} - \boldsymbol{x}^{(t-1)}[g]\boldsymbol{\ell}^{(t-1)}\right) \\
&- \sum_{t=1}^{T} \frac{(1 - C\eta)\eta}{2} \cdot \mathrm{Var}_{\mathbf{Q}^{(t)}[g,\cdot]}\left(\boldsymbol{x}^{(t-1)}[g]\boldsymbol{\ell}^{(t-1)}\right)
\end{aligned}$$

Thus the overall regret of BM is bounded as follows: for any swap function $\phi : [\![n]\!] \to [\![n]\!]$, we have

$$\begin{aligned}
&\sum_{t=1}^{T} \langle \boldsymbol{x}^{(t)}, \boldsymbol{\ell}^{(t)} \rangle \\
=&\sum_{t=1}^{T} \langle \boldsymbol{x}^{(t)} \mathbf{Q}^{(t)}, \boldsymbol{\ell}^{(t)} \rangle \\
=&\sum_{t=1}^{T} \sum_{g=1}^{n} \langle \mathbf{Q}^{(t)}[g, \cdot], \boldsymbol{x}^{(t)}[g] \boldsymbol{\ell}^{(t)} \rangle \\
\leq&\sum_{g=1}^{n} \sum_{t=1}^{T} \left(\boldsymbol{x}^{(t)}[g] \cdot \boldsymbol{\ell}^{(t)}[\phi(g)] + \frac{\log n}{\eta} + \left(\frac{\eta}{2} + C\eta^2 \right) \mathrm{Var}_{\mathbf{Q}^{(t)}[g,\cdot]} \left(\boldsymbol{x}^{(t)}[g] \cdot \boldsymbol{\ell}^{(t)} - \boldsymbol{x}^{(t-1)}[g] \cdot \boldsymbol{\ell}^{(t-1)} \right) \right. \\
&\qquad \left. - \frac{(1 - C\eta)\eta}{2} \cdot \mathrm{Var}_{\mathbf{Q}^{(t)}[g,\cdot]} \left(\boldsymbol{x}^{(t-1)}[g] \cdot \boldsymbol{\ell}^{(t-1)} \right) \right), \qquad (161)
\end{aligned}$$

which implies that

$$\begin{aligned}
&\sum_{t=1}^{T} \langle \boldsymbol{x}^{(t)}, \boldsymbol{\ell}^{(t)} \rangle - \sum_{t=1}^{T} \sum_{g=1}^{n} \boldsymbol{x}^{(t)}[g] \cdot \boldsymbol{\ell}^{(t)}[\phi(g)] \\
\leq& \frac{n \log n}{\eta} + \sum_{t=1}^{T} \sum_{g=1}^{n} \left(\frac{\eta}{2} + C\eta^2 \right) \mathrm{Var}_{\mathbf{Q}^{(t)}[g,\cdot]} \left(\boldsymbol{x}^{(t)}[g] \cdot \boldsymbol{\ell}^{(t)} - \boldsymbol{x}^{(t-1)}[g] \cdot \boldsymbol{\ell}^{(t-1)} \right) \\
&\qquad - \sum_{t=1}^{T} \sum_{g=1}^{n} \frac{(1 - C\eta)\eta}{2} \cdot \mathrm{Var}_{\mathbf{Q}^{(t)}[g,\cdot]} \left(\boldsymbol{x}^{(t-1)}[g] \cdot \boldsymbol{\ell}^{(t-1)} \right).
\end{aligned}$$

Thus, we want to show the following lemma

Lemma G.1. *Suppose all players play according to BM-OMWU with step size η satifying $1/T \leq \eta \leq \frac{1}{C \cdot mn^3 \log^4(T)}$ for a sufficiently large constant C. Then for any player $i \in [\![m]\!]$ and any $g \in [\![n]\!]$, the overall losses for player i: $\boldsymbol{\ell}_i^{(1)}, \ldots, \boldsymbol{\ell}_i^{(T)} \in \mathbb{R}^n$, the weight player i places on strategy g: $\boldsymbol{x}^{(1)}[g], \ldots, \boldsymbol{x}^{(T)}[g] \in \mathbb{R}$, and the strategy vectors output by player i's g^{th} instance of OMWU $\mathcal{R}_{\Delta,i,g}$: $\mathbf{Q}_i[g, \cdot]^{(1)}, \ldots, \mathbf{Q}_i[g, \cdot]^{(T)} \in \mathbb{R}^n$ satisfy*

$$\sum_{t=1}^{T} \mathrm{Var}_{\mathbf{Q}^{(t)}[g,\cdot]} \left(\boldsymbol{x}^{(t)}[g] \cdot \boldsymbol{\ell}^{(t)} - \boldsymbol{x}^{(t-1)}[g] \cdot \boldsymbol{\ell}^{(t-1)} \right) \qquad (162)$$

$$\leq \frac{1}{2} \cdot \sum_{t=1}^{T} \mathrm{Var}_{\mathbf{Q}^{(t)}[g,\cdot]} \left(\boldsymbol{x}^{(t-1)}[g] \cdot \boldsymbol{\ell}^{(t-1)} \right) + O\left(\log^5(T) \right) \qquad (163)$$

Consequently, combining these two results gives, for $\eta \in \left[\frac{1}{T}, \frac{1}{C \cdot mn^3 \log^4(T)}\right]$,

$$\begin{aligned}
\mathrm{SwapReg}_i^T \leq & \frac{n \log n}{\eta} + \sum_{t=1}^{T} \sum_{g=1}^{n} \left(\frac{\eta}{2} + C\eta^2\right) \mathrm{Var}_{\mathbf{Q}^{(t)}[g,\cdot]} \left(\boldsymbol{x}^{(t)}[g] \cdot \boldsymbol{\ell}^{(t)} - \boldsymbol{x}^{(t-1)}[g] \cdot \boldsymbol{\ell}^{(t-1)}\right) \\
& - \sum_{t=1}^{T} \sum_{g=1}^{n} \frac{(1 - C\eta)\eta}{2} \cdot \mathrm{Var}_{\mathbf{Q}^{(t)}[g,\cdot]} \left(\boldsymbol{x}^{(t-1)}[g] \cdot \boldsymbol{\ell}^{(t-1)}\right) \\
\leq & \frac{n \log n}{\eta} + \sum_{g=1}^{n} \left(\frac{2\eta}{3} \sum_{t=1}^{T} \mathrm{Var}_{\mathbf{Q}^{(t)}[g,\cdot]} \left(\boldsymbol{x}^{(t)}[g] \cdot \boldsymbol{\ell}^{(t)} - \boldsymbol{x}^{(t-1)}[g] \cdot \boldsymbol{\ell}^{(t-1)}\right) \right. \\
& \left. - \frac{\eta}{3} \sum_{t=1}^{T} \mathrm{Var}_{\mathbf{Q}^{(t)}[g,\cdot]} \left(\boldsymbol{x}^{(t-1)}[g] \cdot \boldsymbol{\ell}^{(t-1)}\right) \right) \\
\leq & \frac{n \log n}{\eta} + \frac{2n\eta}{3} O\left(\log^5(T)\right). \qquad (164)
\end{aligned}$$

Hence, by setting $\eta = \frac{1}{C \cdot mn^3 \log^4(T)}$ we recover the bound stated in Theorem 1.4. □

The main technical tool for achieving this result is Lemma 5.1, establishing high order smoothness under the fixed point iterates of BM-OMWU.

G.2 Tools for BM-OMWU Swap Regret

Recall from earlier our definition of rooted trees and the Markov Chain Tree Theorem. We re-iterate those results here under a slightly different formulation, catering to the proof techniques for this section.

Definition G.2. Suppose Z is an $n \times n$ matrix and $G = (V, E)$ with $V = [\![n]\!]$ is the weighted directed graph associated with Z. We say a subgraph T of G is a directed tree rooted at $i \in [\![n]\!]$ if (1) T does not contain any cycles and (2) Node i has no outgoing edges, while every other node $j \in [\![n]\!]$ has exactly one outgoing edge. For each node $i \in [\![n]\!]$, we write $\mathbb{T}_i^n$ to denote the set of all directed trees rooted at node i.

We let $\mathbb{R}^{n \times n}$ (respectively, $\mathbb{C}^{n \times n}$) denote the space of real-valued (respectively, complex-valued) $n \times n$ matrices. For each $j \in [\![n]\!]$, we introduce the following functions $\Phi_{\mathrm{MCT},j} : \mathbb{C}^{n \times n} \to \mathbb{C}$:

$$\Phi_{\mathrm{MCT},j}(Z) := \frac{\sum_{T \in \mathbb{T}_j^n} \exp\left(\sum_{(a,b) \in T} Z_{ab}\right)}{\sum_{j=1}^{n} \sum_{T \in \mathbb{T}_j^n} \exp\left(\sum_{(a,b) \in T} Z_{ab}\right)}.$$

Note that for $Z \in \mathbb{R}^{n \times n}$, we have $\Phi_{\mathrm{MCT},j}(Z) \in \mathbb{R}$.

Theorem G.3 (Markov chain tree theorem)**.** *Let $\mathbf{Q} \in \mathbb{R}^{n \times n}$ be a Markov chain so that $\mathbf{Q}_{ij} > 0$ for all $i, j \in [\![n]\!]$. Then the stationary distribution of Q is given by the vector $(\Phi_{\mathrm{MCT},1}(\ln Q), \ldots, \Phi_{\mathrm{MCT},1}(\ln Q))$, where $\ln Q$ denotes the matrix whose (i, j) entry is $\ln \mathbf{Q}_{ij}$.*

Lemma G.4. *Fix any $Z^{(0)} \in \mathbb{R}^{n\times n}$. Consider any function of the form $\phi(Z) := \frac{\Phi_{\mathrm{MCT},j}(Z^{(0)}+Z)}{\Phi_{\mathrm{MCT},j}(Z^{(0)})}$. Then the sum of the Taylor series coefficients of order $k \geq 0$ of ϕ at $\mathbf{0}$ is bounded above by $30\cdot(2n^3)^k$, and has radius of convergence greater than $1/n$.*

Proof. Fix any multi-index $\gamma \in \mathbb{Z}_{\geq 0}^{n\times n}$; we will bound $\frac{d^\gamma}{dZ^\gamma}\phi(\mathbf{0})$. Fix any $Z \in \mathbb{C}^{n\times n}$ so that $\|Z\|_\infty \leq \pi/(3n)$. Set $\zeta = Z^{(0)} + Z$. Note that, for any $T \in \mathbb{T}_j^n$,

$$\begin{aligned}\left|\exp\left(\sum_{(a,b)\in T}\zeta_{ab}\right)\right| &= \left|\exp\left(\sum_{(a,b)\in T} Z^{(0)}_{ab} + \sum_{(a,b)\in T}(\zeta_{ab} - Z^{(0)}_{ab})\right)\right| \\ &\leq \exp(\pi/3)\cdot\exp\left(\sum_{(a,b)\in T} Z^{(0)}_{ab}\right),\end{aligned}$$

where the inequality uses the fact that $\left|\sum_{(a,b)\in T}(\zeta_{ab} - Z^{(0)}_{ab})\right| \leq \pi/3$ and that $Z^{(0)}$ is real-valued.

For any $a \in \mathbb{R}$ and $\zeta \in \mathbb{C}$ with $|\zeta| \leq \pi/3$, we have

$$\Re(\exp(a+\zeta)) = \exp(a)\cdot\Re(\zeta) \geq \exp(a)\cdot\cos(\pi/3)\cdot\exp(-\pi/3) > \exp(a)/10.$$

Thus, for any $j' \in [\![n]\!]$ and $T \in \mathbb{T}_{j'}^n$, it holds that

$$\begin{aligned}\Re\left(\exp\left(\sum_{(a,b)\in T}\zeta_{ab}\right)\right) &= \Re\left(\exp\left(\sum_{(a,b)\in T} Z^{(0)}_{ab} + \sum_{(a,b)\in T}(\zeta_{ab} - Z^{(0)}_{ab})\right)\right) \\ &\geq \frac{1}{10}\cdot\exp\left(\sum_{(a,b)\in T} Z^{(0)}_{ab}\right).\end{aligned}$$

Thus, since $Z^{(0)}$ is real-valued,

$$\begin{aligned}|\phi(Z)| &= \frac{|\Phi_{\mathrm{MCT},j}(\zeta)|}{\Phi_{\mathrm{MCT},j}(Z^{(0)})} \\ &\leq \frac{\sum_{j'=1}^n\sum_{T\in\mathbb{T}_{j'}^n}\exp\left(\sum_{(a,b)\in T} Z^{(0)}_{a,b}\right)}{\sum_{T\in\mathbb{T}_j^n}\exp\left(\sum_{(a,b)\in T} Z^{(0)}_{a,b}\right)}\cdot\frac{\sum_{T\in\mathbb{T}_j^n}\left|\exp\left(\sum_{(a,b)\in T} Z_{ab}\right)\right|}{\sum_{j'=1}^n\sum_{T\in\mathbb{T}_{j'}^n}\Re\left(\exp\left(\sum_{(a,b)\in T} Z_{ab}\right)\right)} \\ &\leq \frac{\sum_{j'=1}^n\sum_{T\in\mathbb{T}_{j'}^n}\exp\left(\sum_{(a,b)\in T} Z^{(0)}_{a,b}\right)}{\sum_{T\in\mathbb{T}_j^n}\exp\left(\sum_{(a,b)\in T} Z^{(0)}_{a,b}\right)}\cdot\frac{\pi/3\cdot\sum_{T\in\mathbb{T}_j^n}\exp\left(\sum_{(a,b)\in T} Z^{(0)}_{a,b}\right)}{1/10\cdot\sum_{j'=1}^n\sum_{T\in\mathbb{T}_{j'}^n}\exp\left(\sum_{(a,b)\in T} Z^{(0)}_{a,b}\right)} \\ &= \exp(\pi/3)\cdot 10 < 30.\end{aligned}$$

By the multivariate version of Cauchy's integral formula,

$$\begin{aligned}\left|\frac{d^\gamma}{dZ^\gamma}\phi(\mathbf{0})\right| &= \left|\frac{\gamma!}{(2\pi i)^{n^2}}\int_{|Z_{11}|=\pi/(3n)}\cdots\int_{|Z_{nn}|=\pi/(3n)}\frac{\phi(Z)}{\prod_{j_1,j_2\in[\![n]\!]}(Z_{j_1j_2})^{\gamma_{j_1j_2}+1}}d\zeta_{11}\cdots d\zeta_{nn}\right| \\ &\leq 30\cdot\frac{\gamma!}{(2\pi)^{n^2}}\cdot\frac{1}{(\pi/(3n))^{n^2+|\gamma|}}\cdot(2\pi\cdot\pi/(3n))^{n^2} = 30\cdot\gamma!\cdot(3n/\pi)^{|\gamma|}. \qquad (165)\end{aligned}$$

For any integer $k \geq 0$, the number of tuples $\gamma \in \mathbb{Z}_{\geq 0}^{n^2}$ with $|\gamma| = k$ is $\binom{k+n^2-1}{n^2-1} \leq (2n^2)^k$. Thus, if we write $\phi(Z) = \sum_{\gamma \in \mathbb{Z}_{\geq 0}^{n^2}} a_\gamma \cdot z^\gamma$, it follows that $a_\gamma = \frac{1}{\gamma!} \cdot \frac{d^\gamma}{dZ^\gamma}\phi(\mathbf{0})$, and so for each $k \geq 0$, $\sum_{\gamma:|\gamma|=k} |a_\gamma| \leq 30 \cdot (2n^2)^k \cdot (3n/\pi)^k \leq 30 \cdot (2n^3)^k$.

To see the lower bound on the radius of convergence, note that Equation (165) gives that for each $\gamma \in \mathbb{Z}_{\geq 0}^{n^2}$, letting $k := |\gamma|$, we have $|a_\gamma|^{1/k} \leq 30^{1/k} \cdot (3n/\pi)$, which tends to $3n/\pi < n$ as $k \to \infty$. Thus, by the multivariate version of the Cauchy-Hadamard theorem, the radius of convergence of the power series of ϕ at $\mathbf{0}$ is greater than $1/n$. □

In the statement of Lemma G.5 below, the quantity 0^0 is interpreted as 1 (in particular, $(h')^{B_0 h'} = 1$ for $h' = 0$).

Lemma G.5. *Fix any $B_1 \geq 4e^2 n^3$, $B_0 \geq 3$. Consider a sequence $\mathbf{Q}^{(0)}, \ldots, \mathbf{Q}^{(h)} \in \mathbb{R}^{n \times n}$ of ergodic Markov chains, so that $\| \log \mathbf{Q}^{(0)} - \log \mathbf{Q}^{(t)} \|_\infty \leq \frac{1}{B_1}$ for $0 \leq t \leq h$. Suppose that for some $\alpha \in (0,1)$, for each $1 \leq h' \leq h$ and $0 \leq t \leq h - h'$, it holds that $\| \left(\mathrm{D}_{h'} (\ln \mathbf{Q})\right)^{(t)} \|_\infty \leq \frac{1}{B_1} \cdot \alpha^{h'} \cdot (h')^{B_0 h'}$. Then, if $p^{(0)}, \ldots, p^{(h)} \in \Delta^n$ denotes the sequence of stationary distributions for $\mathbf{Q}^{(0)}, \ldots, \mathbf{Q}^{(h)}$, it holds that, for any $j \in [\![n]\!]$,*

$$\frac{\left|\left(\mathrm{D}_h\, p[j]\right)^{(t)}\right|}{p^{(0)}[j]} \leq \frac{720 n^3 e^2}{B_1} \cdot \alpha^h \cdot h^{B_0 h + 1}.$$

Proof. By the Markov chain tree theorem we have that, for each $j \in [\![n]\!]$, $p^{(t)}[j] = \Phi_{\mathrm{MCT},j}(\ln \mathbf{Q}^{(t)})$. We now apply Lemma 3.7 with $T = h+1$, $\mathbf{z}^{(t)} = \ln \mathbf{Q}^{(t-1)} - \ln \mathbf{Q}^{(0)}$ for $1 \leq t \leq h+1$, $R_1 = 1$, $R_2 = 2n^3$, and the values of B_0, B_1 given in the hypothesis of this lemma (Lemma G.5). Moreover, we set $\phi(Z) = \frac{\Phi_{\mathrm{MCT},j}((\ln \mathbf{Q}^{(0)}) + Z)}{\Phi_{\mathrm{MCT},j}(\ln \mathbf{Q}^{(0)})} = \frac{\Phi_{\mathrm{MCT},j}((\ln \mathbf{Q}^{(0)}) + Z)}{p^{(0)}[j]}$. Lemma G.4 gives that the function ϕ is $(30, 2n^3)$-bounded, and has radius of convergence greater than $1/n$ at the point $Z = \mathbf{0}$. Then the hypotheses of Lemma G.5 imply those of Lemma 3.7, and Lemma 3.7 gives that

$$\frac{\left|\left(\mathrm{D}_h\, p[j]\right)^{(t)}\right|}{p^{(0)}[j]} = \frac{\left|\left(\mathrm{D}_h \left(\phi \circ (\ln \mathbf{Q} - \ln \mathbf{Q}^{(0)})\right)\right)^{(0)}\right|}{p^{(0)}[j]} \leq \frac{720 n^3 e^2}{B_1} \cdot \alpha^h \cdot h^{B_0 h + 1}.$$

(Here we have used that $\phi(\ln \mathbf{Q}^{(t)} - \ln \mathbf{Q}^{(0)}) = \frac{\Phi_{\mathrm{MCT},j}(\ln \mathbf{Q}^{(t)})}{p^{(0)}[j]} = p^{(t)}[j]/p^{(0)}[j]$.) □

Lemma G.6. *For $n \in \mathbb{N}$, let $\xi_1, \ldots, \xi_n \geq 0$ so that $\xi_1 + \cdots + \xi_n = 1$. For any $j \in [\![n]\!]$, the function*

$$\phi_j((z_1, \ldots, z_n)) = \frac{\exp(z_j)}{\sum_{k=1}^n \xi_k \cdot \exp(z_k)}$$

satisfies, for any $z \in \mathbb{R}^n$ with $\|z\|_\infty \leq 1/4$,

$$|\log \phi_j(z)| \leq \|z\|_\infty \leq 6\|z\|_\infty.$$

Proof. For $0 \leq x \leq 1$, we have $1 + x \leq \exp(x) \leq 1 + 2x$. Then, for $\|z\|_\infty \leq 1/2$,

$$\phi_j(z) \leq \frac{1 + 2z_j}{\sum_{k=1}^n \xi_k \cdot (1 + z_k)} \leq \frac{1 + 2\|z\|_\infty}{1 - \|z\|_\infty} \leq (1 + 2\|z\|_\infty)^2$$

and

$$\phi_j(z) \geq \frac{1+z_j}{\sum_{k=1}^n \xi_k \cdot (1+2z_k)} \geq \frac{1-\|z\|_\infty}{1+2\|z\|_\infty} \geq (1-\|z\|_\infty)(1-2\|z\|_\infty).$$

Thus, for $\|z\|_\infty \leq 1/4$,

$$-6\|z\|_\infty \leq \log(1-\|z\|_\infty) + \log(1-2\|z\|_\infty) \leq \log \phi_j(z) \leq 2\log(1+2\|z\|_\infty) \leq 4\|z\|_\infty.$$

□

Lemma G.7. *For $n \in \mathbb{N}$, let $\xi_1, \ldots, \xi_n \geq 0$ such that $\xi_1 + \cdots + \xi_n = 1$. For each $j \in [\![n]\!]$, define $\phi_j : \mathbb{R}^n \to \mathbb{R}$ to be the function*

$$\phi_j((z_1, \ldots, z_n)) = \frac{\xi_j \exp(z_j)}{\sum_{k=1}^n \xi_k \cdot \exp(z_k)}$$

and let $P_{\log \phi_j}(z) = \sum_{\gamma \in \mathbb{Z}^n_{\geq 0}} a_{j,\gamma} \cdot z^\gamma$ denote the Taylor series of $\log \phi_j$. Then for any $j \in [\![n]\!]$ and any integer $k \geq 1$,

$$\sum_{\gamma \in \mathbb{Z}^n_{\geq 0} :\ |\gamma| = k} |a_{j,\gamma}| \leq e^k/k.$$

Proof. We have that

$$\frac{\partial \log \phi_j}{\partial z_t} = \frac{1}{\phi_j} \cdot \frac{\partial \phi_j}{\partial z_t} = \begin{cases} -\phi_t & \text{if } t \neq j \\ 1-\phi_j & \text{if } t = j \end{cases}$$

and so,

$$\begin{aligned}
\sum_{\gamma \in \mathbb{Z}^n_{\geq 0} :\ |\gamma| = k} |a_{j,\gamma}| &= \frac{1}{k!} \sum_{t \in [\![n]\!]^k} \left| \frac{\partial^k \log \phi_j(0)}{\partial z_{t_1} \partial z_{t_2} \cdots \partial z_{t_k}} \right| \\
&= \frac{1}{k!} \sum_{t_1 \in [\![n]\!]} \sum_{t_{-1} \in [\![n]\!]^{k-1}} \left| \frac{\partial^{k-1}}{\partial z_{t_2} \cdots \partial z_{t_k}} \frac{\partial \log \phi_j(0)}{\partial z_{t_1}} \right| \\
&= \frac{1}{k!} \sum_{t_1 \in [\![n]\!]} \sum_{t_{-1} \in [\![n]\!]^{k-1}} \left| \frac{\partial^{k-1}}{\partial z_{t_2} \cdots \partial z_{t_k}} \left(\phi_{t_1} - \mathbb{1}[t_1 = j]\right)(0) \right| \qquad (166)
\end{aligned}$$

For $k = 1$,

$$\frac{1}{k!} \sum_{t_1 \in [\![n]\!]} |\phi_{t_1}(0) - \mathbb{1}[t_1 = j]| \leq 1 + \sum_{t_1 \in [\![n]\!]} \xi_{t_1} = 2$$

For $k \geq 2$, the constant $\mathbb{1}[t_1 = j]$ will be eliminated by the derivative, and we will have

$$\begin{aligned}
\frac{1}{k!} \sum_{t_1 \in [\![n]\!]} \sum_{t_{-1} \in [\![n]\!]^{k-1}} \left| \frac{\partial^{k-1} \phi_{t_1}(0)}{\partial z_{t_2} \cdots \partial z_{t_k}} \right| &\leq \frac{1}{k!} \sum_{t_1 \in [\![n]\!]} \left((k-1)! \cdot \xi_{t_1} e^k \right) \qquad (167) \\
&= \frac{1}{k!} \cdot (k-1)! \cdot e^k \cdot \sum_{t_1 \in [\![n]\!]} \xi_{t_1} = e^k/k
\end{aligned}$$

where (167) follows from Lemma B.6.

□

G.3 Proof of BM High Order Smoothness

Lemma 5.1 [Detailed] *Fix a parameter $\alpha \in \left(0, \frac{1}{H+3}\right)$. If all players follow BM-OMWU updates with step size $\eta \le \frac{1}{311040e^7H^2n_i^3}$, then for any player $i \in [\![m]\!]$, integer h satisfying $0 \le h \le H$, time step $t \in [T-h]$, and $g \in [n_i]$, it holds that*

$$\| \left(\mathrm{D}_h \left(\boldsymbol{x}_i[g] \cdot \boldsymbol{\ell}_i\right)\right)^{(t)} \|_\infty \le \alpha^h \cdot h^{3h+1}.$$

Proof. We prove this inductively, showing that, for all $i \in [\![m]\!], 0 \le h \le H, t \in [T-h], g \in [n_i]$,

$$\begin{aligned}
\left\|\left(\mathrm{D}_h \boldsymbol{\ell}_i\right)^{(t)}\right\|_\infty &\le \alpha^h \cdot h^{3h+1} \\
\left\|\left(\mathrm{D}_h \boldsymbol{x}_i[g] \cdot \boldsymbol{\ell}_i\right)^{(t)}\right\|_\infty &\le \alpha^h \cdot h^{3h+1} \\
\left\|\left(\mathrm{D}_h \boldsymbol{x}_i\right)^{(t)}\right\|_\infty &\le \alpha^h \cdot h^{3h+1}.
\end{aligned}$$

The base case of $h = 0$ is evident from the fact that losses and strategy probabilities are in $[0,1]$, and therefore $\|\boldsymbol{\ell}_i^{(t)}\|_\infty, \|\boldsymbol{x}_i^{(t)}\|_\infty \le 1$. So, proving the following inductive statement is sufficient to prove the lemma.

Claim G.8. *Fix some $1 \le h \le H$. Suppose that for some $B_0 \ge 3$ and all integers h' satisfying $1 \le h' < h$, all $i \in [\![m]\!]$, all $t \le T - h'$, and all $g \in [n_i]$, it holds that*

$$\| \left(\mathrm{D}_{h'} \boldsymbol{\ell}_i\right)^{(t)} \|_\infty \le \alpha^{h'} \cdot (h')^{B_0 h'+1} \tag{168}$$

$$\left\|\left(\mathrm{D}_{h'} \boldsymbol{x}_i[g] \cdot \boldsymbol{\ell}_i\right)^{(t)}\right\|_\infty \le \alpha^{h'} \cdot (h')^{B_0 h'+1} \tag{169}$$

$$\| \left(\mathrm{D}_{h'} \boldsymbol{x}_i\right)^{(t)} \|_\infty \le \alpha^{h'} \cdot (h')^{B_0 h'+1}. \tag{170}$$

Suppose that the step size η satisfies $\eta \le \frac{1}{311040e^7H^2n_i^3}$. Then for all $i \in [\![m]\!]$ and $1 \le t \le T-h$,

$$\left\|\left(\mathrm{D}_h \boldsymbol{\ell}_i\right)^{(t)}\right\|_\infty \le \alpha^h \cdot h^{B_0 h+1} \tag{171}$$

$$\left\|\left(\mathrm{D}_h \boldsymbol{x}_i[g] \cdot \boldsymbol{\ell}_i\right)^{(t)}\right\|_\infty \le \alpha^h \cdot h^{B_0 h+1} \tag{172}$$

$$\left\|\left(\mathrm{D}_h \boldsymbol{x}_i\right)^{(t)}\right\|_\infty \le \alpha^h \cdot h^{B_0 h+1}. \tag{173}$$

Proof of Claim G.8. First, notice that for any agent $i \in [\![m]\!]$, any OMWU instance $\mathcal{R}_{\Delta,i,g}$ of agent i with $g \in [n_i]$, any $t_0 \in \{0, 1, \ldots, T\}$, and any $t \ge 0$, by the definition (OMWU) of the OMWU updates, it holds that, for each $j \in [n_i]$,

$$\mathbf{Q}_i^{(t_0+t+1)}[g,j] = \frac{\mathbf{Q}_i^{(t_0)}[g,j] \cdot \exp\left(\eta \cdot \left(\boldsymbol{x}_i^{(t_0-1)}[g]\boldsymbol{\ell}_i^{(t_0-1)}[j] - \sum_{s=0}^{t} \boldsymbol{x}_i^{(t_0+s)}[g]\boldsymbol{\ell}_i^{(t_0+s)}[j] - \boldsymbol{x}_i^{(t_0+t)}[g]\boldsymbol{\ell}_i^{(t_0+t)}[j]\right)\right)}{\sum_{k=1}^{n_i} \mathbf{Q}_i^{(t_0)}[g,k] \cdot \exp\left(\eta \cdot \left(\boldsymbol{x}_i^{(t_0-1)}[g]\boldsymbol{\ell}_i^{(t_0-1)}(k) - \sum_{s=0}^{t} \boldsymbol{x}_i^{(t_0+s)}[g]\boldsymbol{\ell}_i^{(t_0+s)}(k) - \boldsymbol{x}_i^{(t_0+t)}[g]\boldsymbol{\ell}_i^{(t_0+t)}(k)\right)\right)}$$

where $\mathbf{Q}_i^t[g,j]$ denotes the weight placed on action j by algorithm $\mathcal{R}_{\Delta,i,g}$ at time t. We can define edge values $\boldsymbol{\ell}_i^{(0)}, \boldsymbol{\ell}_i^{(-1)}, \mathbf{Q}_i^{(0)}$ to ensure that the above equation holds even for $t_0 \in \{0,1\}$. Now, for any $g,j \in [n_i]$, any integer t_0 satisfying $0 \le t_0 \le T$; and any integer $t \ge 0$, let us define

$$\mathbf{L}_{i,t_0}^{(t)}[g,j] := \boldsymbol{x}_i^{(t_0-1)}[g]\boldsymbol{\ell}_i^{(t_0-1)}[j] - \sum_{s=0}^{t-1} \boldsymbol{x}_i^{(t_0+s)}[g]\boldsymbol{\ell}_i^{(t_0+s)}[j] - \boldsymbol{x}_i^{(t_0+t-1)}[g]\boldsymbol{\ell}_i^{(t_0+t-1)}[j]$$

Also, for a vector $z = (z[1], \dots, z[n_i]) \in \mathbb{R}^{n_i}$ and indices $g,j \in [n_i]$, define

$$\phi_{t_0,g,j}(z) := \frac{\exp(z[j])}{\sum_{k=1}^{n_i} \mathbf{Q}_i^{(t_0)}[g,k] \cdot \exp(z(k))}, \tag{174}$$

so that

$$\mathbf{Q}_i^{(t_0+t)}[g,j] = \mathbf{Q}_i^{(t_0)}[g,j] \cdot \phi_{t_0,g,j}(\eta \cdot \mathbf{L}_{i,t_0}^{(t)}[g,\cdot]) \tag{175}$$

for $t \ge 1$, where $\mathbf{L}_{i,t_0}^{(t)}[g,\cdot]$ denotes the vector $(\mathbf{L}_{i,t_0}^{(t)}[g,1], \dots, \mathbf{L}_{i,t_0}^{(t)}[g,n_i])$.

Next, note that, for all $g,j \in [n_i]$,

$$(\mathrm{D}_1 \mathbf{L}_{i,t_0}[g,j])^{(t)} = \boldsymbol{x}_i^{(t_0+t-1)}[g]\boldsymbol{\ell}_i^{(t_0+t-1)}[j] - 2\boldsymbol{x}_i^{(t_0+t)}[g]\boldsymbol{\ell}_i^{(t_0+t)}[j]$$

and so, for any $1 \le h' \le h$,

$$\begin{aligned}\left|(\mathrm{D}_{h'}(\eta \cdot \mathbf{L}_{i,t_0}[g,j]))^{(t)}\right| &= \left|(\mathrm{D}_{h'-1}(\eta \cdot \boldsymbol{x}_i[g] \cdot \boldsymbol{\ell}_i[j]))^{(t_0+t-1)} - 2(\mathrm{D}_{h'-1}(\eta \cdot \boldsymbol{x}_i[g] \cdot \boldsymbol{\ell}_i[j]))^{(t_0+t)}\right| \\ &\le 3\eta \cdot \alpha^{h'-1} \cdot (h'-1)^{B_0 h'}\end{aligned} \tag{176}$$

where (176) follows from our inductive hypothesis (169).

Additionally, since $|\mathbf{L}_{i,t_0}[g,j]| \le t+2$ for all t_0, i, g, j, we have $|\eta \cdot \mathbf{L}_{i,t_0}[g,j]| \le \eta \cdot (t+2) \le 1$ for $\eta \le \frac{1}{t+2}$. By Lemma G.7, the function $z \mapsto \log \phi_{t_0,g,j}(z)$ is $(1,e)$-bounded, and so for each $1 \le h' \le h$ we may apply Lemma 3.7 with $h = h'$, $\boldsymbol{z}^{(t)} = \eta \cdot \mathbf{L}_{i,t_0}^{(t)}[g,\cdot]$ and $B_1 = \frac{1}{\eta} \cdot \min\left\{\frac{\alpha}{3}, \frac{1}{H+3}\right\} = \frac{\alpha}{3\eta}$. Thus, we can conclude, for all $1 \le h' \le h$, $t \le h+1$, and $g \in [n_i]$,

$$\left(\mathrm{D}_{h'} \log\left(\phi_{t_0,g,j}(\eta \cdot \mathbf{L}_{i,t_0}^{(t)}[g,\cdot])\right)\right)^{(t)} \le 12e^3 \cdot 3\eta \cdot \alpha^{h'} \cdot (h')^{B_0 h'+1} \le 36e^3 H\eta \cdot \alpha^{h'} \cdot (h')^{B_0 h'}. \tag{177}$$

Taking the logarithm of both sides of (175), we have

$$\log\left(\mathbf{Q}_i^{(t_0+t)}[g,j]\right) = \log \mathbf{Q}_i^{(t_0)}[g,j] + \log\left(\phi_{t_0,g,j}(\eta \cdot \mathbf{L}_{i,t_0}^{(t)}[g,\cdot])\right). \tag{178}$$

Let $\mathbf{Q}_i^{(t)} \in \mathbb{R}^{n_i \times n_i}$ be the matrix with entries $\mathbf{Q}_i^{(t)}[g,j]$ and $P_{i,t_0}^{(t)}$ be the matrix with entries $\phi_{t_0,g,j}\left(\eta \cdot \mathbf{L}_{i,t_0}^{(t)}[g,\cdot]\right)$, for $g,j \in [n_i]$. Then, for all $g \in [n_i]$ and all $t_0 \in [T-h]$, $t \in [\![h]\!]$, we have

$$\begin{aligned}\boldsymbol{x}_i^{(t_0+t)}[g] &= \Phi_{\mathrm{MCT},g}\left(\log \mathbf{Q}_i^{(t_0+t)}\right) \\ &= \Phi_{\mathrm{MCT},g}\left(\log \mathbf{Q}_i^{(t_0)} + \log\left(P_{i,t_0}^{(t)}\right)\right).\end{aligned}$$

Next note that, by Lemma G.6 and $t \le H+1$,

$$\begin{aligned}\left\|\log \mathbf{Q}_i^{(t_0+t)} - \log \mathbf{Q}_i^{(t_0)}\right\|_\infty \le& \max_{g,j\in[n_i]} \left|\log \phi_{t_0,g,j}(\eta \cdot \mathbf{L}_{i,t_0}^{(t)}[g,\cdot])\right| \\ \le& 6\eta\cdot(t+2) \le 36e^3 H\eta. \end{aligned} \tag{179}$$

We now apply Lemma G.5 with $\mathbf{Q}^{(t)} = P_{i,t_0}^{(t)}$, for $0 \le t \le h$, and $1/B_1 = 36e^3 H\eta$. The preconditions of the lemma hold from (177). Then Lemma G.5 gives that for all $g \in [n_i]$,

$$\left|(\mathrm{D}_h\, \boldsymbol{x}_i[g])^{(t_0+1)}\right| \le |\boldsymbol{x}_i^{(t_0)}[g]| \cdot 720 n_i^3 e^2 \cdot 36e^3 H\eta \cdot \alpha^{h'} \cdot (h)^{B_0 h+1}.$$

verifying the first of three desired inductive conclusions (173) as long as $\eta \le 1/(25920e^5 H n_i^3)$.

Next, we have that for each agent $i \in [\![m]\!]$, each $t \in [T]$, and each $a_i \in [n_i]$, $\boldsymbol{\ell}_i^{(t)}(a_i) = \mathbb{E}_{a_{i'}\sim \boldsymbol{x}_{i'}^{(t)}:\, i'\neq i}[\mathcal{L}_i(a_1,\dots,a_m)]$. Thus, for $1 \le t \le T$,

$$\begin{aligned}\left|(\mathrm{D}_h\, \boldsymbol{\ell}_i[a_i])^{(t)}\right| =& \left|\sum_{s=0}^{h}\binom{h}{s}(-1)^{h-s}\boldsymbol{\ell}_i^{(t+s)}[a_i]\right| \\ =& \left|\sum_{a_{i'}\in[n_{i'}],\ \forall i'\neq i} \mathcal{L}_i(a_1,\dots,a_m)\sum_{s=0}^{h}\binom{h}{s}(-1)^{h-s}\cdot\prod_{i'\neq i}\boldsymbol{x}_{i'}^{(t+s)}[a_{i'}]\right| \\ \le& \sum_{a_{i'}\in[n_{i'}],\ \forall i'\neq i}\left|\sum_{s=0}^{h}\binom{h}{s}(-1)^{h-s}\cdot\prod_{i'\neq i}\boldsymbol{x}_{i'}^{(t+s)}[a_{i'}]\right| \\ =& \sum_{a_{i'}\in[n_{i'}],\ \forall i'\neq i}\left|\left(\mathrm{D}_h\left(\prod_{i'\neq i}\boldsymbol{x}_{i'}[a_{i'}]\right)\right)^{(t)}\right|, \\ \le& 25920e^5 H n_i^3\eta\cdot\alpha^h\cdot(h)^{B_0h+1}\sum_{a_{i'}\in[n_{i'}],\ \forall i'\neq i}\left(\prod_{i'\neq i}\boldsymbol{x}_i^{(t)}[a_i']\right) \\ =& 25920e^5 H n_i^3\eta\cdot\alpha^h\cdot(h)^{B_0h+1}\prod_{i'\neq i}\left(\sum_{a_{i'}\in[n_{i'}]}\boldsymbol{x}_i^{(t)}[a_i']\right) \\ =& 25920e^5 H n_i^3\eta\cdot\alpha^h\cdot(h)^{B_0h+1}\end{aligned}$$

verifying the second of three desired inductive conclusions (171) as long as $\eta \le 1/(25920e^5 H n_i^3)$.

Lastly, for $\eta \le \frac{1}{12e^2H}\cdot\frac{1}{25920e^5Hn_i^3} = \frac{1}{311040e^7H^2n_i^3}$, we have now verified the inductive hypotheses

$$\begin{aligned}\|\,(\mathrm{D}_{h'}\,\boldsymbol{\ell}_i)^{(t)}\,\|_\infty &\le 25920e^5 H n_i^3\eta\cdot h'\alpha^{h'}\cdot(h')^{B_0h'} \le \frac{1}{2e^2}\alpha^{h'}\cdot(h')^{B_0h'} \\ \|\,(\mathrm{D}_{h'}\,\boldsymbol{x}_i)^{(t)}\,\|_\infty &\le 25920e^5 H n_i^3\eta\cdot h'\alpha^{h'}\cdot(h')^{B_0h'} \le \frac{1}{2e^2}\alpha^{h'}\cdot(h')^{B_0h'}\end{aligned}$$

for all h' up to and including h. Thus, we can apply Lemma 3.7 with $n = 2$, $\phi(a, b) = ab$ (which is $(1, 1)$-bounded), $B_1 = 12e^2$, and the sequence $\boldsymbol{z}^{(t)} = (\boldsymbol{\ell}_i^{(t)}[j], \boldsymbol{x}_i^{(t)}[g])$. Therefore, for the product sequence $\boldsymbol{x}_i[g] \cdot \boldsymbol{\ell}_i[j]$, we have, for all $t \in [T - h + 1]$

$$\left|(\mathrm{D}_h\, \boldsymbol{x}_i[g] \cdot \boldsymbol{\ell}_i[j])^{(t)}\right| \leq \alpha^h \cdot (h)^{B_0 \cdot h + 1} \tag{180}$$

verifying the final inductive conclusion (172), as desired.

□

□

G.4 Completing the proof of Theorem 1.4

Using the lemma developed in the previous sections we now can complete the proof of Lemma G.1, the last step towards the proof of Theorem 1.4. The lemma is restated formally below.

Lemma G.1 [Detailed] *Suppose a time horizon $T \geq 4$ is given, we set $H := \lceil \log T \rceil$, and all players play according to BM-OMWU with step size η satisfying $1/T \leq \eta \leq \frac{1}{C \cdot m H^4 n_i^3}$ with $C = 311040e^7$. Then for any $i \in [\![m]\!]$, $g \in [\![n_i]\!]$, the losses $\boldsymbol{\ell}_i^{(1)}, \ldots, \boldsymbol{\ell}_i^{(T)} \in [0, 1]^n$ for player i satisfy:*

$$\begin{aligned}
&\sum_{t=1}^{T} \mathrm{Var}_{\mathbf{Q}_i^{(t)}[g,\cdot]}\left(\boldsymbol{x}_i^{(t)}[g] \cdot \boldsymbol{\ell}_i^{(t)} - \boldsymbol{x}_i^{(t-1)}[g] \cdot \boldsymbol{\ell}_i^{(t-1)}\right) \\
&\leq \frac{1}{2} \cdot \sum_{t=1}^{T} \mathrm{Var}_{\mathbf{Q}_i^{(t)}[g,\cdot]}\left(\boldsymbol{x}_i^{(t-1)}[g] \cdot \boldsymbol{\ell}_i^{(t-1)}\right) + O\left(\log^5(T)\right)
\end{aligned}$$

Proof. Recall the following lemma from earlier.

Lemma C.4. *For any integers $n \geq 2$ and $T \geq 4$, we set $H := \lceil \log T \rceil$, $\alpha = 1/(4H)$, and $\alpha_0 = \frac{\sqrt{\alpha/8}}{H^3}$. Suppose that $Z^{(1)}, \ldots, Z^{(T)} \in [0, 1]^n$ and $P^{(1)}, \ldots, P^{(T)} \in \Delta^n$ satisfy the following*

1. *For each $0 \leq h \leq H$ and $1 \leq t \leq T - h$, it holds that $\left\|(\mathrm{D}_h\, Z)^{(t)}\right\|_\infty \leq H \cdot \left(\alpha_0 H^3\right)^h$*
2. *The sequence $P^{(1)}, \ldots, P^{(T)}$ is ζ-consecutively close for some $\zeta \in [1/(2T), \alpha^4/8256]$.*

Then,

$$\sum_{t=1}^{T} \mathrm{Var}_{P^{(t)}}\left(Z^{(t)} - Z^{(t-1)}\right) \leq 2\alpha \sum_{t=1}^{T} \mathrm{Var}_{P^{(t)}}\left(Z^{(t-1)}\right) + 165120(1 + \zeta)H^5 + 2 \tag{112}$$

where we say that a sequence $P^{(1)}, \ldots, P^{(T)}$ is ζ-consecutively close iff for all t

$$\max\left\{\left\|\frac{P^{(t)}}{P^{(t+1)}}\right\|_\infty, \left\|\frac{P^{(t+1)}}{P^{(t)}}\right\|_\infty\right\} \leq 1 + \zeta$$

We hope to apply Lemma C.4 with $n = n_i$, $P^{(t)} = \mathbf{Q}_i^{(t)}[g,\cdot]$ and $Z^{(t)} = \boldsymbol{x}_i^{(t)}[g] \cdot \boldsymbol{\ell}_i^{(t)}$, as well as $\zeta = 7\eta$. To do so, we must verify that the preconditions of the lemma hold. Set $C_1 = 8256$ (note that C_1 is the constant appearing in item 2). Our assumption that $\eta \le \frac{1}{C \cdot mH^4}$ implies that, as long as the constant C satisfies $C \ge 4^4 \cdot 7 \cdot C_1 = 14794752$,

$$\eta \le \min\left\{\frac{\alpha^4}{7C_1}, \frac{\alpha_0}{36e^5 m}\right\}. \tag{181}$$

To verify precondition 1, we apply Lemma 5.1 with the parameter α in the lemma set to α_0: a valid selection as $\alpha_0 < \frac{1}{H+3}$. We conclude that, for each $i \in [\![m]\!]$, $g \in [\![n_i]\!]$, $0 \le h \le H$ and $1 \le t \le T - h$, it holds that $\left\|\left(\mathrm{D}_h \boldsymbol{x}_i[g] \cdot \boldsymbol{\ell}_i\right)^{(t)}\right\|_\infty \le H \cdot \left(\alpha_0 H^3\right)^h$ since $\eta \le \frac{\alpha_0}{36e^5 m}$ as required by the lemma. To verify precondition 2, we first confirm that our selection of $\zeta = 7\eta$ places it in the desired interval $[1/(2T), \alpha^4/C_1]$ as $\eta \le \frac{\alpha^4}{7C_1}$. By the definition of the OMWU updates, for all $i \in [m]$ and $1 \le t \le T$, we have $\max\left\{\left\|\frac{\mathbf{Q}_i^{(t)}[g,\cdot]}{\mathbf{Q}_i^{(t+1)}[g,\cdot]}\right\|_\infty, \left\|\frac{\mathbf{Q}_i^{(t+1)}[g,\cdot]}{\mathbf{Q}_i^{(t)}[g,\cdot]}\right\|_\infty\right\} \le \exp(6\eta)$. Thus, the sequence $\mathbf{Q}_i^{(1)}[g,\cdot], \ldots, \mathbf{Q}_i^{(T)}[g,\cdot]$ is (7η)-consecutively close (since $\exp(6\eta) \le 1 + 7\eta$ for η satisfying (181)). Therefore, Lemma C.4 applies and we have

$$\begin{aligned}
&\sum_{t=1}^{T} \mathrm{Var}_{\mathbf{Q}_i^{(t)}[g,\cdot]}\left(\boldsymbol{x}_i^{(t)}[g] \cdot \boldsymbol{\ell}_i^{(t)} - \boldsymbol{x}_i^{(t-1)}[g] \cdot \boldsymbol{\ell}_i^{(t-1)}\right) \\
&\le 2\alpha \sum_{t=1}^{T} \mathrm{Var}_{\mathbf{Q}_i^{(t)}[g,\cdot]}\left(\boldsymbol{x}_i^{(t-1)}[g] \cdot \boldsymbol{\ell}_i^{(t-1)}\right) + 165120(1+7\eta)H^5 + 2 \\
&\le \frac{1}{2} \cdot \sum_{t=1}^{T} \mathrm{Var}_{\mathbf{Q}_i^{(t)}[g,\cdot]}\left(\boldsymbol{x}_i^{(t-1)}[g] \cdot \boldsymbol{\ell}_i^{(t-1)}\right) + C' H^5
\end{aligned}$$

for $C' = 2 + 165120(1 + 7/8256) = 165262$, as desired. □