Appendix
Competing AI: How does competition feedback affect machine learning?

A Extended Discussion

Due to space constraints, we continue a supplementary discussion here (Appendix A). Experimental details and supplementary figures can be found in Appendix B and mathematical details can be found in Appendix C.

A.1 Economic & Multi-Agent Theory

Many of the concepts and quantities we study in this work have parallels in economic theory [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]. For example, quality of prediction is a notion of consumer welfare. The economics literature tends to focus on competition between firms or mathematical agents, rather than on specific ML predictors. It would be an interesting direction for future work to connect and extend our learning competition framework from the economic perspective.

A.2 Softmax Model

We include additional details and justifications for the model proposed in Section 2. Our proposed model of user choice satisfies Luce's choice axioms [11, 12, 13, 14] and emerges from the established information-theoretic notion of rational inattention in economics [15, 16, 17, 18]. The softmax form in particular can be derived from optimal decision making under information processing constraints [19, 20, 21].

A.3 Feedback Loops in ML Systems

Feedback loops, where two systems repeatedly influence and have access to only the decisions of the other, have been studied in supervised and reinforcement learning [22]. A related and interesting example includes the feedback in online reviews on digital media platforms [23, 24, 25]. Existing works examine feedback loops in the single-agent setting [26], with particular branches proposing metrics to recover counterfactuals of consumer preferences fixing consumer strategy [27, 28]. The effects of competition in ML holds significant implications to sociology [29, 30], economics [31, 32, 33], electoral systems [34, 35], and recommendation systems [36, 37, 38, 39]. [40] studies an interesting but substantially different model of bandits in matching markets. The model we study here is also distinct from standard settings for online learning or active learning, where typically a single algorithm gets to explore and select data. In our setup, each data selects one among several competing algorithms.

B Experimental Details

The implementation for the experiments may be found on GitHub in repository: `tginart/competing-ai`.

B.1 Supervised learning

We ran three datasets for our supervised learning simulations: `Postures` [41, 42], `Adult Income` [42], and `FashionMNIST` [43]. `Postures` has 5 classes. `Adult Income` has 2 classes. `FashionMNIST` has 10 classes. In the main text we reported results for NN and MLP. Our implementation is in Python [44] using Pytorch [45] and Numpy [46] frameworks.

B.1.1 Hyperparameters and Training Protocols

Seed samples For each data set we set a number of seed samples that was sufficient to train a model to perform slightly (a few percentage points) better than random guessing. For `Postures` and `Adult` we set a seed size of 3 samples. For `FashionMNIST` we set a seed size of 100.

Algorithm 1 Competing predictors in supervised learning

Input: Set of competing predictors $\{A^{(1)},...,A^{(k)}\}$, general population distribution $\mathcal{D}$, prediction quality function q.
 # To initialize, all predictors warm-start with $s \in \mathbb{N}$ i.i.d. seed data
for $i \in [k]$ **do**
  IID random sample $\mathcal{D}^{(i)}_{\text{seed}} := \{(x^{(i)}_{0,j}, y^{(i)}_{0,j})\}_{j=1}^{s} \sim \mathcal{D}$
  Train a predictor $A^{(i)}_1$ using $\mathcal{D}^{(i)}_{\text{seed}}$.
end for
 # The competition begins
for $1 \leq t < T$ **do**
  Random sample $(x_t, y_t) \sim \mathcal{D}$ *# User samples from the general population*
  For all $i \in [k]$, predictor $A^{(i)}_t$ proposes a prediction $\hat{y}^{(i)}_t$, *i.e.*, $\hat{y}^{(i)}_t = A^{(i)}_t(x_t)$
  User selects a winner $w_t \in [k]$ based on $\{q(y_t, \hat{y}^{(1)}_t), ..., q(y_t, \hat{y}^{(k)}_t)\}$ *# User decides which predictor to query based on the prediction quality*
  Winner updates the current model $A^{(w_t)}_t$ and defines $A^{(w_t)}_{t+1}$ by retraining with a new datum (x_t, y_t)
end for

NN For NN, we always use *one* nearest-neighbor to keep things simple. Because the method is non-parametric, training the model just consists of appending each new sample to the data matrix. Refer to [47, 48] for more details on the nearest-neighbor algorithm.

MLP For the MLP, we use the same architecture hyper-parameters for each α in order to ensure consistency. All predictors used the same hyper-parameters and training protocol as well. We selected these by doing a small amount of manual tuning. We use Pytorch's Adam optimizer [49] to train our MLPs (with default settings). For all datasets, we use 1 hidden layer. At the start of the competition, we used Pytorch's default initialization. After this, we never reinitialized the weights. Instead, we always fine-tune the weights from the previous training pass. Refer to [50] for more details about MLPs.

For `Postures` the input width is 16 and the hidden width is 16. We used a learning rate of 10^{-3}. After the initial training on the seed samples, we retrained after every 4th new data point was added to the set. When training, we used we batch size of (up to) 32. We randomly shuffled the data for each training instance. We trained until we had updated on 1,000 data points or reached 32 epochs (whichever was met first – this depended on the number of data samples the predictor had observed).

For `Adult Income` the input width is 50 and the hidden width is 64. We used a learning rate of 10^{-3}. After the initial training on the seed samples, we retrained after every 32nd new data point was added to the set. When training, we used we batch size of (up to) 32. We randomly shuffled the data for each training instance. We trained until we had updated on 1,000 data points or reached 32 epochs.

For `FashionMNIST` the input width is 784 and the hidden width is 400. We used a learning rate of 10^{-4}. After the initial training on the seed samples, we retrained after every 500th new data point was added to the set. When training, we used we batch size of (up to) 50. We randomly shuffled the data for each training instance. We trained until we had reached 30 epochs.

Number of Iterations The simulation takes time (roughly) quadratic in the number of iterations because of the frequent retraining that takes place. For `Postures` and `Adult`, we run the competition for 2,000 rounds when using the NN and we run the competition for 4,000 rounds when using the MLP. For `FashionMNIST` we run the competition for 10,000 rounds for both algorithms.

B.2 Collaborative Filtering

We describe the protocol for the collaborative filtering experiments in detail. Recall that at each round t, one of m users is sampled uniformly at random. The sampled user at time t, denoted U_t, must then query one of k recommenders. The selected recommender then recommends one of r items. We have an underlying preference matrix, $M \in [0,1]^{r \times m}$ such that M_{ij} is the j-th user's preference score for item . The goal of the users is to maximize the the preference score

Preference Matrix For our simulation we set $r \leftarrow 12$ and $m \leftarrow 64$. We sweep over k as seen in Fig. 4. We sample $W_{ij} \sim N(0,1)$ for $W \in \mathbb{R}^{m \times 3}$ and we sample $V_{ij} \sim N(0,1)$ for $V \in \mathbb{R}^{r \times 3}$. We compute $M' = VW^T$ and then affine scale M' onto $[0,1]$ to produce the final preference matrix M. The final rank of M is 4 (the rank increases by 1 due to the affine scaling).

Users Each user operates independently of the others without sharing data or otherwise communicating. Users treat each recommender as an arm in a k-arm bandit problem. Users operate as if the arms are stationary in time (although this is actually false since the recommenders can improve in time). The reward each user obtains from each recommendation is dictated the underlying preference matrix. Each user uses an ϵ-greedy strategy to **SELECT** a recommender. When user i is selecting $\epsilon \leftarrow \tau_i^{-0.3}$ where τ_i is the total number of rounds in which user i has participated thus far.

Recommenders Recommenders use online matrix factorization to reconstruct the underlying preference matrix. Each recommender uses an ϵ-greedy strategy to recommend items to users based on their current representation. Recommenders do not directly observe M_{ij} when they recommend item i to user j. Instead, they observe noisy feedback sampled from a Bernoulli trial with mean M_{ij}. Recommender a has their own copy of $\hat{W}^{(a)} \in \mathbb{R}^{m \times 4}$ and $\hat{V}^{(a)} \in \mathbb{R}^{r \times 4}$ which are initialized with independent uniform random entries. See algorithm below for the psuedo-code for this implementation. The dependence on a is implicit (again, each recommender solves their own instance of the problem with their own data).

Algorithm 2 Online Matrix Factorization Update for Collaborative Filtering

Input: User index i, item index j, $\tilde{Y} \sim \mathbf{Bern}(M_{ij})$, number of observations for pair (i,j), denoted ν_{ij}, learning rate γ and regularization penalty λ

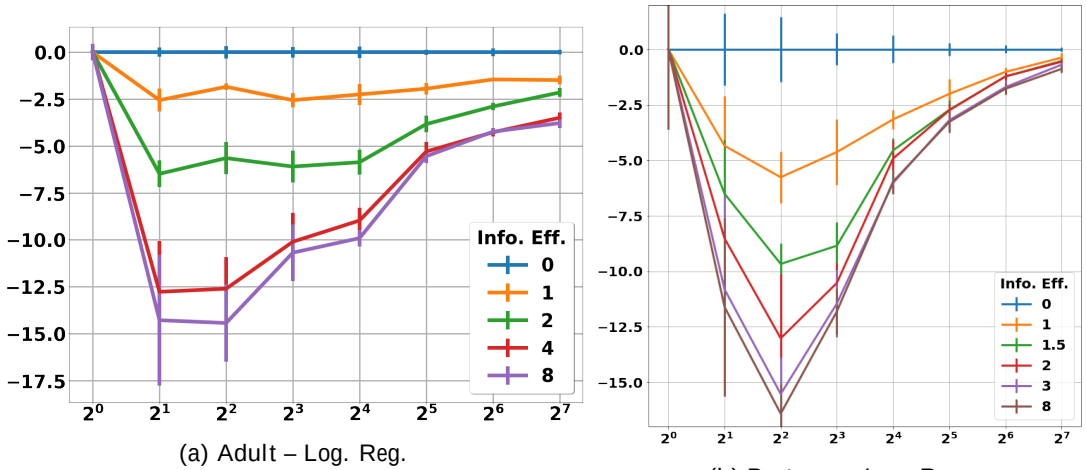
$\hat{M}_{ij} \leftarrow \frac{\nu_{ij}-1}{\nu_{ij}} \hat{M}_{ij} + \frac{1}{\nu_{ij}} \tilde{Y}$ # *Update running average pair* (i,j)
$\hat{V}_i \leftarrow \hat{V}_i - \gamma(\hat{M}_{ij} - V_i W_j^T)\hat{W}_j - \lambda \hat{V}_i$ # *Update row from user matrix*
$\hat{W}_j \leftarrow \hat{W}_j - \gamma(\hat{M}_{ij} - V_i W_j^T)\hat{V}_i^{(a)} - \lambda \hat{W}_j^{(a)}$ # *Update row from item matrix*

For our simulation we use $\lambda = 10^{-4}$ and $\gamma = 0.1$. When recommender a chooses an item to select for user i they select $\max_j \{\hat{V}_i^{(a)} \hat{W}_j^{(a)T}\}$ with probability $1-\epsilon$ and a uniformly random item with probability ϵ. For consistency with the users, we set $\epsilon \leftarrow \tau_a^{-0.3}$ where τ_a is the number of recommendations given out by the recommender a thus far.

B.3 Additional Experiments

We include some additional experiments and figures. In particular, we sweep wider ranges of α and we also include some additional simulations using logistic regression as the prediction algorithm.

In Fig. 6 we report the experiment in Fig. 3 using the logistic regression model. The trends are similar to what we saw with the MLP and NN models. In Fig. 7 we report the experiment in Fig. 4 but for a wider range of α. When α is large the prediction quality for users is monotone increasing and when α it is monotone decreasing.



(a) Adult – Log. Reg.

(b) Postures – Log. Reg.

Figure 6: Figure 3 experiment with Log. Reg.

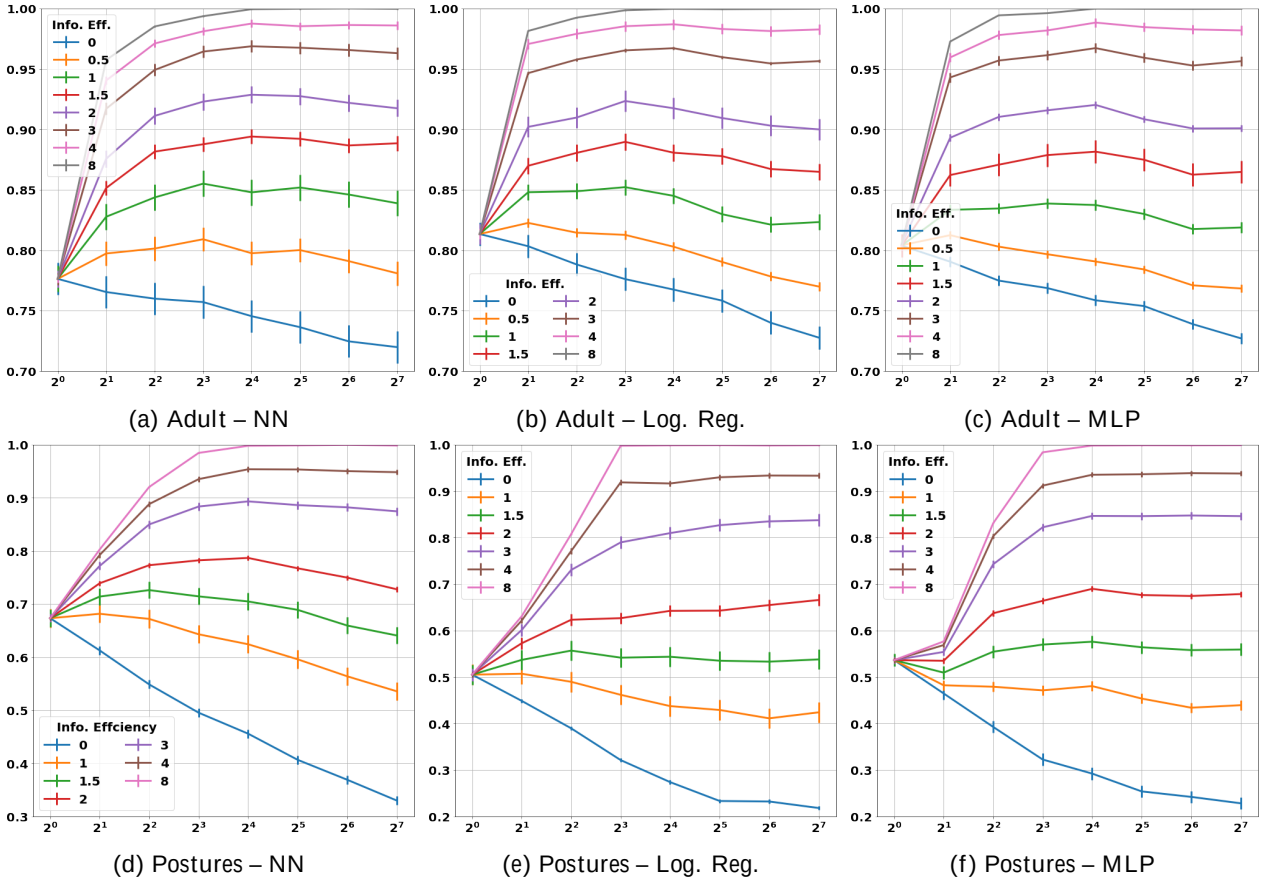


Figure 7: Figure 4 experiment with a larger sweep of α reported

C Mathematical Details

We include mathematical and theoretical details here. This includes formal definitions for the learning competition as well as proofs of the theorems in Sec. 4.

C.1 Definitions

We proceed by formally defining the details of the competition dynamics which the theorems in the main text use. This particular instance can be generalized to include a wider range of scenarios, including the collaborative filtering simulations.

Definition C.1. *(Supervised Prediction Competition)*

Prediction competition $\mathcal{G}=(\mathcal{D},\mathbf{A},\mathbf{SELECT},T)$ *where*

1. $\mathcal{D}$ *is a general population distribution over* $\mathcal{X}\times\mathcal{Y}\subseteq\mathbb{R}^d\times\{1,...,|\mathcal{Y}|\}$ *defining a supervised classification task*
2. $\mathbf{A}=\{A^{(1)},..,A^{(k)}\}$ *is a set of* k *competing predictors such that each predictor* $A^{(i)}$ *updates a learning algorithm over time as described in Sec. 2 of the main text.*
3. *A selection rule* $\mathbf{SELECT}:\mathbb{R}^k\to[k]$ *is a function over* $\mathbb{R}^k$*, i.e., a space of* k *prediction qualities, and outputs one of the* k *predictors.*
4. $T\in\mathbb{N}$ *is the number of rounds in the competition.*

When it is unambiguous we may identify a predictor $A^{(i)}$ with their index, $i\in[k]$. As mentioned in Sec. 2, we will use the softmax rule, parameterized by $0\leq\alpha\leq\infty$ as our user choice model at each round. We use 0-1 loss for ℓ.

C.2 Proofs

C.2.1 Proof of Theorem 4.1

We proceed to restate and prove Thm 4.1. Note that many non-parametric methods are C-approximations to the Bayes error rate. For example, the nearest neighbors method is a 2-approximation [51]. Here, we use the standard notion of a C-approximation ratio in optimization (see [52]).

Theorem C.2 (Theorem 4.1)**.** *Suppose users have perfect information ($\alpha=\infty$) and each predictor is trained using a non-parametric method that is asymptotically a C-approximation to the Bayes error rate. Let $s=|D_0|$ be the number of i.i.d. seed samples that each predictor starts with and assume $s\geq 5$. Then there exists $\mathcal{D}$ such that for any $k>1$, $\lim_{t\to\infty}\frac{\mathcal{R}_t^k}{\mathcal{R}_t^1}=\infty$.*

Proof. It will suffice to construct a distribution that can be easily analyzed. Let P denote this distribution. At a high-level, the strategy will be to construct a noise-free ground truth P that results in any particular predictor's expected error rate to be bounded away from 0. This suffices to complete the claim, since any non-zero expected risk is sufficient.

We proceed to define the distribution P. Let P_X be a distribution over $\mathbb{R}$.

$$P_X(x)=\begin{cases}\frac{1}{s} & x=1\\ 1-\frac{1}{s} & x=0\\ 0 & \textbf{else}\end{cases}$$

and the marginal $P_{Y|X}$ is defined by $Y=X$.

Fix an arbitrary predictor. Let $\mathcal{E}$ be the event that this fixed predictor's seed set will lack any points labeled 1. Then:

$$\mathbf{Pr}(\mathcal{E})=\left(1-\frac{1}{s}\right)^s\geq\frac{1}{4} \tag{1}$$

Where the inequality holds for any positive s. On the other hand, let $\mathcal{F}$ be the event that at least one other of the $k-1$ remaining predictors do sample a seed set that does contain a point labeled 1. Then:

$$\mathbf{Pr}(\mathcal{F})=1-\left(1-\frac{1}{s}\right)^{s(k-1)}\geq 1-e^{-k+1} \tag{2}$$

Since the events are independent, combining them yields:

$$\mathbf{Pr}(\mathcal{F}\cap\mathcal{E})\geq\frac{1}{4}(1-e^{-k+1}) \tag{3}$$

Finally, to complete the argument, we point out that in the event $\mathcal{F}\cap\mathcal{E}$, the fixed predictor will never obtain a sample labeled 1 because they cannot predict a 1 due to the fact that they lack any such points in their seed sets. Meanwhile, there exists another predictor who does have a seed point labeled 1 and will correctly predict all points labeled 1. Thus, the fixed predictor will never obtain a zero error rate. In particular we have that

$$\mathcal{R}_t^k\geq\mathbf{Pr}(\mathcal{F}\cap\mathcal{E}\cap\{Y=1\})=\frac{1}{4ks}(1-e^{-k+1})>0 \tag{4}$$

for all t.

Of course, it is unavoidable that the error rate not decay to 0 for large s, but the linear rate of decay in the seed size is also not particularly fast.

It is clear that because $P_{Y|X}$ is deterministic, a C-approximation should also converge to zero in large t, but we have shown that $\mathcal{R}_t^k$ is bounded away from 0 for all t when $k>1$. This completes the proof.

□

C.2.2 Proof of Theorem 4.2

It is worth clarifying that for the Thm 4.2, proved below, we will be assuming that the nearest-neighbor algorithm uses a majority vote tie-breaking procedure when the nearest-neighbor is non-unique.

Theorem C.3 (Theorem 4.2)**.** *Suppose $k=2$ and both predictors use the nearest-neighbor algorithm. Assume that $\mathbf{Pr}(Y=f(X))>1-\epsilon$ for some function f and $\epsilon<\frac{1}{3}$, and $\alpha>\log\frac{1-\epsilon}{\epsilon}$. Let $s=|D_0|$ be the number of i.i.d. seed samples that each predictor starts with and assume $s\geq 2$. Then there exists $\mathcal{D}$ such that*

$$\lim_{t\to\infty}\frac{\mathcal{R}_t^2}{\mathcal{R}_t^1}\geq 1+\frac{(4\epsilon(1-\epsilon))^{s/2}}{9\sqrt{2s}}\left(\frac{1}{2}-\epsilon\right)\left(1-2\frac{1-\epsilon}{1-\epsilon+\epsilon e^{\alpha}}\right)^2$$

A slightly looser version that removes ϵ dependence holds whenever $\alpha>\log(2)$:

$$\lim_{t\to\infty}\frac{\mathcal{R}_t^2}{\mathcal{R}_t^1}\geq 1+\frac{1}{54\sqrt{2s}}\left(\frac{8}{9\sqrt{s}}\right)^{s/2}\left(1-\frac{2}{2+e^{\alpha}}\right)^2$$

Proof. Suppose $X=0$. Let f be constant with $f(\cdot)=1$. Let $Y=1$ with probability $1-\epsilon$ and $Y=0$ with probability ϵ for some $\epsilon<\frac{1}{2}$. This joint distribution P_{XY} satisfies the assumptions. The fact that $Y|X$ is non-deterministic is essential. Without this, as long as α is finite, an interpolating non-parametric method would asymptotically ε-cover the input domain, which would result in no penalty.

We will use the following proof strategy with P as our example distribution. For some event $\mathcal{E}$:

$$\mathcal{R}_t^2\geq\mathbf{Pr}(\mathcal{E})\mathcal{R}_t^2|\mathcal{E}+(1-\mathbf{Pr}(\mathcal{E}))\mathcal{R}^*\tag{5}$$

which implies

$$\frac{\mathcal{R}_t^2}{\mathcal{R}^*}\geq 1+\frac{\mathbf{Pr}(\mathcal{E})(\mathcal{R}_t^2|\mathcal{E}-\mathcal{R}^*)}{\mathcal{R}^*}\tag{6}$$

where $\mathcal{R}_t^2|\mathcal{E}$ denotes the expected error rate conditioned on event $\mathcal{E}$ and $\mathcal{R}^*$ is the Bayes error rate. We will also show that under P:

$$\lim_{t\to\infty}\mathcal{R}_t^1\to\mathcal{R}^*=\epsilon\tag{7}$$

Combining (6) and (7) will then yield the claim. To show (7), we can simply note that the fraction of 1s in the predictor's data set will concentrate around $1-\epsilon$ due to the law of large numbers [53]. Since this will be the majority, the majority tie-breaker will be in effect, meaning that the predictor will always predict 1. This prediction is the Bayes optimal prediction. This implies (7) holds.

To finish the proof, we must give expressions or bounds for $\mathbf{Pr}(\mathcal{E})$ and $\mathcal{R}_t^2|\mathcal{E}$ under a suitably defined event $\mathcal{E}$.

Let $Q_0(t)$ and $Q_1(t)$ be the fraction of 1s in the data sets for each predictor, respectively, at round t. Thus, $Q_i(0)$ is the fraction of 1s in the seed set of agent i.

$$sQ_i(0)\sim\mathbf{Bin}(s,1-\epsilon)\tag{8}$$

We can use the following bounds on the deviations of a Binomial [54, 55, 56] to yield the following bounds:

$$\exp\left(-sD(\frac{1}{2}||1-\epsilon)\right)\geq\mathbf{Pr}(Q_i(0)<\frac{1}{2})\geq\frac{1}{\sqrt{2s}}\exp\left(-sD(\frac{1}{2}||1-\epsilon)\right)\tag{9}$$

Where $i\in\{0,1\}$ and $D(a||b)$ is the binary relative entropy, given by

$$D(a||b)=a\log(\frac{a}{b})+(1-a)\log(\frac{1-a}{1-b})\tag{10}$$

Using elementary logarithmic identities we can simplify:

$$\exp\left(-sD(\frac{1}{2}||1-\epsilon)\right)=(4\epsilon(1-\epsilon))^{s/2} \tag{11}$$

The tail bound inequalities simplify into:

$$\sqrt{2s}\mathfrak{B}\geq\mathbf{Pr}(Q_i(0)<\frac{1}{2})\geq\mathfrak{B} \tag{12}$$

Where, as a shorthand, we let $\mathfrak{B}=\frac{(4\epsilon(1-\epsilon))^{\frac{s}{2}}}{\sqrt{2s}}$.

We proceed to define a suitable $\mathcal{E}$. In order to do so, we define event $\mathcal{F}(\tau)$, parameterized by $0\leq\tau\leq\infty$, as follows:

$$\mathcal{F}(\tau)=\left(\bigcap_{0\leq l<\tau}\{Q_0(l)<\frac{1}{2}<Q_1(l)\}\right)\cup\left(\bigcap_{0\leq l<\tau}\{Q_1(l)<\frac{1}{2}<Q_0(l)\}\right) \tag{13}$$

Let $\iota=\text{argmin}_i\{Q_i(0)\}$ and $\bar{\iota}=\text{argmax}_i\{Q_i(0)\}$. Let $\mathcal{E}=\mathcal{F}(\infty)$. Recall that $\hat{Y}_t^{(j)}$ is agent j's prediction at time t. Observe that $\mathcal{F}(t)$ implies both $\hat{Y}_t^{(\iota)}=0$ and $\hat{Y}_t^{(\bar{\iota})}=1$ due to the majority rule. With $\mathcal{E}$ now defined, we proceed to concoct a bound for $\mathbf{Pr}(\mathcal{E})$.

In event $\mathcal{F}(0)$, it must be the case that predictor ι has at least one fewer 1s than 0s in the seed set and predictor $\bar{\iota}$ has at least one more 1 than 0. En route for our bound on $\mathbf{Pr}(\mathcal{E})$ we derive the lower bound for $\mathbf{Pr}(\mathcal{F}(0))$ as follows. Notice that $\mathcal{F}(0)$ can be easily expressed in terms of $\{Q_i(0)\leq\frac{1}{2})\}$ which makes $\mathfrak{B}$ an ideal expression for bounding $\mathbf{Pr}(\mathcal{F}(0))$.

$$\mathbf{Pr}(\mathcal{F}(0))\geq\mathfrak{B}(1-\sqrt{2s}\mathfrak{B}) \tag{14}$$

In turn, this yields us the following:

$$\mathbf{Pr}(\mathcal{E})\geq\mathbf{Pr}(\mathcal{E}|\mathcal{F}(0))\mathbf{Pr}(\mathcal{F}(0))\geq\mathbf{Pr}(\mathcal{E}|\mathcal{F}(0))\mathfrak{B}(1-\sqrt{2}\mathfrak{B}) \tag{15}$$

Thus, we are left with the task of bounding $\mathbf{Pr}(\mathcal{E}|\mathcal{F}(0))$ to find an expression to bound $\mathbf{Pr}(\mathcal{E})$. We proceed by using random walk theory ([57, 58] are suitable references for the uninitiated). We will study an integer-valued stochastic process over the integers. At time $\mathfrak{t}\in\mathbb{N}$, $\mathfrak{X}_\mathfrak{t}\in\mathbf{Z}$. Furthermore, $\mathfrak{X}_{t+1}\in\{\mathfrak{X}_t-1,\mathfrak{X}_t,\mathfrak{X}_t+1,\}$. The distribution over the increments is defined by $(\mathfrak{q},\mathfrak{p},\mathfrak{r})$ as follows:

$$\mathbf{Pr}(\mathfrak{X}_{\mathfrak{t}+1}=\mathfrak{X}_\mathfrak{t}-1)=\mathfrak{q}\text{ and }\mathbf{Pr}(\mathfrak{X}_{\mathfrak{t}+1}=\mathfrak{X}_\mathfrak{t}+1)=\mathfrak{p}\text{ and }\mathbf{Pr}(\mathfrak{X}_{\mathfrak{t}+1}=\mathfrak{X}_\mathfrak{t})=\mathfrak{r} \tag{16}$$

We denote the random walk distribution by $\mathbf{RW}(q,p,r)$ and write $\{\mathfrak{X}\}_{\mathfrak{t}=0}^{\infty}\sim\mathbf{RW}(\mathfrak{q},\mathfrak{p},\mathfrak{r})$ to associate the random variable to the distribution. We will also define and use an independent copy of the random walk, denoted by $\{\mathfrak{X}'\}_{\mathfrak{t}=0}^{\infty}\sim\mathbf{RW}(\mathfrak{q}',\mathfrak{p}',\mathfrak{r}')$. The analysis of random walks is a rich subject and many techniques are known for computing various probabilities for events of interest. Of particular utility here is the fact that if $\mathfrak{X}_0=1$ and $\frac{\mathfrak{q}}{\mathfrak{q}+\mathfrak{p}}<\frac{1}{2}$, then the probability that the walk never reaches the origin is given by

$$\mathbf{Pr}(\mathfrak{X}_\mathfrak{t}>0,\forall\mathfrak{t})=1-\frac{2\mathfrak{q}}{\mathfrak{q}+\mathfrak{p}} \tag{17}$$

This result that can be obtained by solving a recurrence relation, as in [57, 58]. Notice that the probability that a random walk with bias (q,p,r) ever reaches the origin is the same as a random walk with bias $(\frac{q}{q+p},\frac{p}{q+p},0)$. We will assume that the random walks start at 1: $\mathfrak{X}_0=\mathfrak{X}'_0=1$.

For convenience, we will define events $\mathcal{H}(\tau)$ as follows.

$$\mathcal{H}(\tau)=\bigcap_{0\leq l\leq\tau}(\{\mathfrak{X}_l>0\}\cap\{\mathfrak{X}'_l>0\}) \tag{18}$$

As we will see, our target $\mathbf{Pr}(\mathcal{E}|\mathcal{F}(0))$ can be bounded by the probability that the random walks in question never reach the origin: $\mathbf{Pr}(\mathcal{H}(\infty))=\mathbf{Pr}(\mathfrak{X}_\mathfrak{t}>0,\mathfrak{X}'_\mathfrak{t}>0,\forall\mathfrak{t})$. To see how this works, we begin by considering the following choices for $(\mathfrak{q},\mathfrak{p},\mathfrak{r})$ and $(\mathfrak{q},\mathfrak{p},\mathfrak{r})$:

$$(\mathfrak{q},\mathfrak{p},\mathfrak{r})=\left(\frac{\epsilon}{1+e^{\alpha}},\frac{(1-\epsilon)e^{\alpha}}{1+e^{\alpha}},\frac{1-\epsilon+\epsilon e^{\alpha}}{1+e^{\alpha}}\right) \tag{19}$$

$$(\mathfrak{q}',\mathfrak{p}',\mathfrak{r}')=\left(\frac{1-\epsilon}{1+e^{\alpha}},\frac{\epsilon e^{\alpha}}{1+e^{\alpha}},\frac{(1-\epsilon)e^{\alpha}+\epsilon}{1+e^{\alpha}}\right) \tag{20}$$

Also note that it is easy to give $\mathbf{Pr}(\mathcal{H}(\infty))$ in terms of $(\mathfrak{q},\ \mathfrak{p},\ \mathfrak{r})$ and $(\mathfrak{q}',\ \mathfrak{p}',\ \mathfrak{r}')$ by noting that $\mathbf{Pr}(\mathcal{H}(\infty))=\mathbf{Pr}(\mathfrak{X}_{\mathfrak{t}}>0,\forall \mathfrak{t})\mathbf{Pr}(\mathfrak{X}'_{\mathfrak{t}}>0,\forall \mathfrak{t})$.

$$\mathbf{Pr}(\mathcal{H}(\infty))=\left(1-\frac{2\mathfrak{q}}{\mathfrak{q}+\mathfrak{p}}\right)\left(1-\frac{2\mathfrak{q}'}{\mathfrak{q}'+\mathfrak{p}'}\right) \tag{21}$$

The reason for this choice will be made apparent later. By assumption we know that α is not too small and ϵ is not too large. Precisely, the conditions are: $\alpha>\log\frac{1-\epsilon}{\epsilon}$ and $\epsilon<\frac{1}{2}$. With elementary algebra one can verify these conditions are sufficient to imply that $\frac{\mathfrak{q}}{\mathfrak{q}+\mathfrak{p}}<\frac{1}{2}$ and $\frac{\mathfrak{q}'}{\mathfrak{q}'+\mathfrak{p}'}<\frac{1}{2}$.

We proceed to construct a relationship between $\mathfrak{X}_t$ and $Q_{\bar{\iota}}(t)$ and between $\mathfrak{X}'_t$ and $Q_{\iota}(t)$. We introduce $\mathbf{S}(t)=\mathbf{SELECT}(t)$ as a shorthand notation. Let $\mathbf{1}(\cdot)$ denote the indicator function. Consider W_t and W'_t defined as below:

$$W_t=\left(s+\sum_{l=1}^{t}\mathbf{1}(\mathbf{S}(l)=\bar{\iota})\right)(2Q_{\bar{\iota}}(t)-1) \tag{22}$$

$$W'_t=\left(s+\sum_{l=1}^{t}\mathbf{1}(\mathbf{S}(l)=\iota)\right)(1-2Q_{\iota}(t)) \tag{23}$$

Intuitively, W is renormalizing $Q_{\bar{\iota}}$ from $\frac{\#\text{ of 1s}}{\text{total }\#}$ to (# of 1s)$-$(# of 0s) (and analogously for W'_t and Q_{ι}). Under this transform, we can express $\mathcal{F}(\tau)$ in terms of W and W' as follows:

$$\mathcal{F}(\tau)=\bigcap_{0\leq l<\tau}(\{W_l>0\}\cap\{W'_l>0\}) \tag{24}$$

Because, of course, $W_t>0$ if and only if $Q_{\bar{\iota}}>\frac{1}{2}$ (and analogously for W'_t and Q_{ι}). Recalling that **SELECT** uses the softmax rule and with Bayes rule [57, 59] we can obtain the following expressions for the distribution over the increments to W and W'. Notice that when conditioning on $\mathcal{F}(t-1)$, we know that $\hat{Y}_t^{(\iota)}=0$ because learner ι's dataset must have a majority of 0s. Similarly, $\hat{Y}_t^{(\bar{\iota})}=1$.

$$\mathbf{Pr}(W_t=W_{t-1}-1|\mathcal{F}(t-1))=\mathbf{Pr}(\mathbf{S}(t)=\bar{\iota},Y_t=0|\mathcal{F}(t-1))=\frac{\epsilon}{1+e^{\alpha}}=\mathfrak{q} \tag{25}$$

$$\mathbf{Pr}(W_t=W_{t-1}+1|\mathcal{F}(t-1))=\mathbf{Pr}(\mathbf{S}(t)=\bar{\iota},Y_t=1|\mathcal{F}(t-1))=\frac{(1-\epsilon)e^{\alpha}}{1+e^{\alpha}}=\mathfrak{p} \tag{26}$$

$$\mathbf{Pr}(W_t=W_{t-1}|\mathcal{F}(t-1))=\mathbf{Pr}(\mathbf{S}(t)=\iota|\mathcal{F}(t-1))=\frac{1-\epsilon+\epsilon e^{\alpha}}{1+e^{\alpha}}=\mathfrak{r} \tag{27}$$

$$\mathbf{Pr}(W'_t=W'_{t-1}-1|\mathcal{F}(t-1))=\mathbf{Pr}(\mathbf{S}(t)=\iota,Y_t=1|\mathcal{F}(t-1))=\frac{1-\epsilon}{1+e^{\alpha}}=\mathfrak{q}' \tag{28}$$

$$\mathbf{Pr}(W'_t=W'_{t-1}+1|\mathcal{F}(t-1))=\mathbf{Pr}(\mathbf{S}(t)=\iota,Y_t=0|\mathcal{F}(t-1))=\frac{\epsilon e^{\alpha}}{1+e^{\alpha}}=\mathfrak{p}' \tag{29}$$

$$\mathbf{Pr}(W'_t=W'_{t-1}|\mathcal{F}(t-1))=\mathbf{Pr}(\mathbf{S}(t)=\bar{\iota}|\mathcal{F}(t-1))=\frac{(1-\epsilon)e^{\alpha}+\epsilon}{1+e^{\alpha}}=\mathfrak{r}' \tag{30}$$

This reveals the reasoning behind the choices for $(\mathfrak{q},\mathfrak{p},\mathfrak{r})$ and $(\mathfrak{q}',\mathfrak{p}',\mathfrak{r}')$. We proceed to bound the increments of W and W' by using $\mathfrak{X}$ and $\mathfrak{X}'$, but this requires a bit more work. Observe that $(W,W')_t$ are coupled whereas $(\mathfrak{X},\ \mathfrak{X}')_t$ are independent. The trick will be the couple $(W,W')_t$ and $(\mathfrak{X},\ \mathfrak{X}')_t$ in a prudent way. To this end, we will introduce a third random process, $(Z,Z')_t$. First, define $\bar{\rho}(t)$ and $\rho(t)$ as follows:

$$\bar{\rho}(t)=\min_{m\in S_t(\{\mathfrak{X}_j\})}m \tag{31}$$

$$\rho(t) = \min_{m \in S_t(\{\mathfrak{X'}_j\})} m \tag{32}$$

where S_t is a set-valued function defined over sequences $\{x_j\}_{j=0}^{\infty}$ as:

$$S_t(\{x_j\}) = \{i : i \in \mathbb{N}, \sum_{j=1}^{i} \mathbf{1}(x_j \neq x_{j-1}) \geq t\} \tag{33}$$

where $\mathbb{N}$ is the set of natural numbers (including zero). Intuitively, $\bar{\rho}(t)$ is computing the index of process $\mathfrak{X}$ that corresponds to t-th non-zero increment (and analogously for $\rho(t)$ and $\mathfrak{X}'$). As for Z_t and Z'_t:

$$Z_t = \mathfrak{X}_{\bar{\rho}(t)} \tag{34}$$

$$Z'_t = \mathfrak{X'}_{\rho(t)} \tag{35}$$

Intuitively Z_t corresponds to the sequence one would obtain from $\mathfrak{X}_t$ with the zero increments deleted (and analogously for Z'_t and $\mathfrak{X'}_t$). Let

$$\mathcal{Z}(\tau) = \bigcap_{0 \leq l < \tau} (\{Z_l > 0\} \cap \{Z'_l > 0\}) \tag{36}$$

Any $\mathfrak{X}_t$ sequence that reaches the origin in finite time will do so with the non-zero increments deleted, which implies that Z_t will also reach the origin (and analogously for Z'_t and $\mathfrak{X'}_t$). This implies the key fact that event $\mathcal{Z}(\infty)$ occurs if $\mathcal{H}(\infty)$ occurs. Also, $\mathcal{H}(\infty)$ occurs almost surely if $\mathcal{Z}(\infty)$ occurs. Events $\mathcal{Z}(\infty)$ and $\mathcal{H}(\infty)$ are equivalent up to a set of measure zero. Namely, the measure zero event that either $\mathfrak{X}_t$ or $\mathfrak{X'}_t$ produces a finite number of non-zero increments.

$$\mathbf{Pr}(\mathcal{Z}(\infty)) = \mathbf{Pr}(\mathcal{H}(\infty)) \tag{37}$$

It remains to couple W_t with Z_t and W'_t with Z'_t. We do so as follows by defining $\bar{\psi}(t)$ and $\psi(t)$

$$\bar{\psi}(t) \sim \mathbf{Bin}(t, \mathfrak{r}) \tag{38}$$

$$\psi(t) = t - \bar{\psi}(t) \tag{39}$$

$$W_t = Z_{\bar{\psi}(t)} + W_0 - 1 \tag{40}$$

$$W'_t = Z'_{\psi(t)} + W'_0 - 1 \tag{41}$$

Notice that because $\mathfrak{r} + \mathfrak{r}' = 1$, the marginals are preserved which makes this a valid coupling. Also notice that the event $\mathcal{Z}(\infty)$ implies $\mathcal{F}(\infty)|\mathcal{F}(0)$ outside of the measure zero event that either $\bar{\psi}(t)$ or $\psi(t)$ remain bounded as $t \to \infty$:

$$\mathbf{Pr}(\mathcal{F}(\infty)|\mathcal{F}(0)) \geq \mathbf{Pr}(\mathcal{Z}(\infty)) \tag{42}$$

$$\mathbf{Pr}(\mathcal{E}|\mathcal{F}(0)) = \mathbf{Pr}(\mathcal{F}(\infty)|\mathcal{F}(0)) \geq \mathbf{Pr}(\mathcal{Z}(\infty)) = \mathbf{Pr}(\mathcal{H}(\infty)) \tag{43}$$

We proceed to establish the bound on $\mathbf{Pr}(\mathcal{E}|\mathcal{F}(0))$. Consider the following:

$$\mathbf{Pr}(\mathcal{E}|\mathcal{F}(0)) = \mathbf{Pr}(\mathcal{F}(\infty)|\mathcal{F}(0)) \geq \mathbf{Pr}(\mathcal{H}(\infty)) = \left(1 - \frac{2\mathfrak{q}}{\mathfrak{q}+\mathfrak{p}}\right)\left(1 - \frac{2\mathfrak{q}'}{\mathfrak{q}'+\mathfrak{p}'}\right) \tag{44}$$

Thus, we can combine our bounds for $\mathbf{Pr}(\mathcal{F}(0))$ and $\mathbf{Pr}(\mathcal{E}|\mathcal{F}(0))$ to obtain a bound on $\mathbf{Pr}(\mathcal{E})$:

$$\mathbf{Pr}(\mathcal{E}) \geq \mathfrak{B}(1 - \sqrt{2s}\mathfrak{B})\left(1 - \frac{2\mathfrak{q}}{\mathfrak{q}+\mathfrak{p}}\right)\left(1 - \frac{2\mathfrak{q}'}{\mathfrak{q}'+\mathfrak{p}'}\right) \tag{45}$$

It now remains to get a bound for $\mathcal{R}_t^2|\mathcal{E}$. Of course, this is easy now that we have ascertained that in event $\mathcal{E}$, we know that ι ($\bar{\iota}$) always has a majority 0s (1s) thus always predicts 0 (1). From this it follows immediately by taking an average that

$$\mathcal{R}_t^2|\mathcal{E} = \frac{1}{2}\ \forall t \implies \lim_{t\to\infty} \mathcal{R}_t^2|\mathcal{E} = \frac{1}{2} \tag{46}$$

Combining the bounds gives:

$$\lim_{t\to\infty} \frac{\mathcal{R}_t^2}{\mathcal{R}_t^1} \geq 1 + \frac{\mathbf{Pr}(\mathcal{E})(\mathcal{R}_t^2|\mathcal{E} - \mathcal{R}^*)}{\mathcal{R}^*} \geq 1 + \frac{1}{\epsilon}\mathfrak{B}(1-\sqrt{2s}\mathfrak{B})\left(1-\frac{2\mathfrak{q}}{\mathfrak{q}+\mathfrak{p}}\right)\left(1-\frac{2\mathfrak{q}'}{\mathfrak{q}'+\mathfrak{p}'}\right)\left(\frac{1}{2}-\epsilon\right) \tag{47}$$

At the expense of a looser bound, we can simplify the bound by a series of additional approximations in order to obtain an interpretable expression:

$$\left(1-\frac{2\mathfrak{q}}{\mathfrak{q}+\mathfrak{p}}\right)\left(1-\frac{2\mathfrak{q}'}{\mathfrak{q}'+\mathfrak{p}'}\right) \geq \left(1-\frac{2\mathfrak{q}'}{\mathfrak{q}'+\mathfrak{p}'}\right)^2 = \left(1-2\frac{1-\epsilon}{1-\epsilon+\epsilon e^{\alpha}}\right)^2 \tag{48}$$

Which follows because $\frac{\mathfrak{q}'}{\mathfrak{q}'+\mathfrak{p}'} > \frac{\mathfrak{q}}{\mathfrak{q}+\mathfrak{p}}$.

We can also crudely simplify the $\mathfrak{B}$-term as follows. Recall:

$$\mathfrak{B} = \frac{(4\epsilon(1-\epsilon))^{s/2}}{\sqrt{2s}} \tag{49}$$

With the assumption that $\epsilon \leq 1/3$ we note that the following holds:

$$\max_{0\leq\epsilon\leq 1/3}\{4\epsilon(1-\epsilon)\} = 8/9 \tag{50}$$

As well as the following identity:

$$(4\epsilon(1-\epsilon)) \geq \epsilon \text{ for } 0 \leq \epsilon \leq 1/3 \tag{51}$$

Together, we can bound $\mathfrak{B}$ from both above and below:

$$\frac{(8/9)^{s/2}}{\sqrt{2s}} \geq \mathfrak{B} \geq \frac{\epsilon^{s/2}}{\sqrt{2s}} \tag{52}$$

Allowing us to conclude:

$$1-\sqrt{2s}\mathfrak{B} \geq 1/9 \tag{53}$$

And,

$$\mathfrak{B}(1-\sqrt{2s}\mathfrak{B}) \geq \frac{\mathfrak{B}}{9} \geq \frac{\epsilon^{s/2}}{9\sqrt{2s}} \tag{54}$$

Thus, we arrive at a neater expression in terms of ϵ, s, and α:

$$\lim_{t\to\infty} \frac{\mathcal{R}_t^2}{\mathcal{R}_t^1} \geq 1 + \frac{(4\epsilon(1-\epsilon))^{s/2}}{9\sqrt{2s}}\left(\frac{1}{2}-\epsilon\right)\left(1-2\frac{1-\epsilon}{1-\epsilon+\epsilon e^{\alpha}}\right)^2 \tag{55}$$

$$\geq 1 + \frac{\epsilon^{s/2}}{9\sqrt{2s}}\left(\frac{1}{2}-\epsilon\right)\left(1-2\frac{1-\epsilon}{1-\epsilon+\epsilon e^{\alpha}}\right)^2 \tag{56}$$

Finally, due to monotonicity of the expression in ϵ we can remove the dependence on ϵ by fixing $\epsilon = \frac{1}{3}$:

$$\lim_{t\to\infty}\frac{\mathcal{R}_t^2}{\mathcal{R}_t^1}\geq 1+\frac{1}{54\sqrt{2s}}\left(\frac{8}{9\sqrt{s}}\right)^{s/2}\left(1-\frac{2}{2+e^{\alpha}}\right)^2 \tag{57}$$

□

C.2.3 Proof of Theorem 4.3

It will be helpful to separately prove a lemma for use in Theorem 4.3's proof. This lemma upper bounds the variance of a symmetrical truncated Binomial with the variance of a usual Binomial with the same number of trials as support left in the truncated Binomial. To clarify notion in the proof of the lemma, note that we define $f(x)\propto g(x)$ to mean:

$$f(x)=Cg(x)\ \forall x\in\mathcal{X} \tag{58}$$

where C is some fixed constant independent of x and the equation holds over all choices of x in some set $\mathcal{X}$ which can be inferred from context.

Lemma C.4. *If $X\sim\mathbf{Bin}(2n,\frac{1}{2})$, then for any integer c such that $n>c>0$:*

$$\mathbf{Var}(X|n-c\leq X\leq n+c)\geq\mathbf{Var}(\mathbf{Bin}(2c,1/2))=\frac{c}{2}$$

Proof. As a notional shorthand, we will use $X_{\mathbf{tr}}$ to denote the truncated version of X:

$$X_{\mathbf{tr}}\sim X|\{n-c\leq X\leq n+c\} \tag{59}$$

By the definition of truncation, $\mathbf{Pr}(X_{\mathbf{tr}}=x)\propto\mathbf{Pr}(X=x)$ for all x in the truncated support. From this, it follows that $X_{\mathbf{tr}}$ inherits symmetry and unimodality from X. Furthermore, from the Binomial pmf it also directly follows that $\mathbf{Pr}(X=x)\propto\binom{n}{x}$ because $p=\frac{1}{2}$. Let $X_{\mathbf{sup}}$ denote the usual Binomial defined over the truncated support:

$$X_{\mathbf{sup}}\sim\mathbf{Bin}(2c,1/2)+n-c \tag{60}$$

The constant translation is just an aesthetic to keep the supports of $X_{\mathbf{sup}}$ and $X_{\mathbf{tr}}$ identical. Observe that for both $X_{\mathbf{sup}}$ and $X_{\mathbf{tr}}$, n is both the mean and the mode outcome (keep in mind the symmetry and unimodality of both). We will complete the proof by showing that:

$$1\geq\frac{\mathbf{Pr}(X_{\mathbf{tr}}=n+j)}{\mathbf{Pr}(X_{\mathbf{tr}}=n)}\geq\frac{\mathbf{Pr}(X_{\mathbf{sup}}=n+j)}{\mathbf{Pr}(X_{\mathbf{sup}}=n)}\text{ for all }j\text{ in the support} \tag{61}$$

Notice that this inequality would imply that $X_{\mathbf{tr}}$ is strictly less concentrated that $X_{\mathbf{sup}}$, and since they have the same support, are unimodal, and are symmetric, it follows that $X_{\mathbf{tr}}$ has higher variance. We proceed to demonstrate this key inequality.

Based on our established proportionality rules we can express the inequality in terms of factorials:

$$\frac{\mathbf{Pr}(X_{\mathbf{tr}}=n+j)}{\mathbf{Pr}(X_{\mathbf{tr}}=n)}=\frac{\binom{2n}{n+j}}{\binom{2n}{n}}=\frac{(n!)^2}{(n+j)!(n-j)!}=\prod_{i=1}^{j}\frac{n-i}{n+i} \tag{62}$$

$$\frac{\mathbf{Pr}(X_{\mathbf{sup}}=n+j)}{\mathbf{Pr}(X_{\mathbf{sup}}=n)}=\frac{\binom{2c}{c+j}}{\binom{2c}{c}}\frac{(c!)^2}{(c+j)!(c-j)!}=\prod_{i=1}^{j}\frac{c-i}{c+i} \tag{63}$$

In order to compare these quantities, consider the following:

$$\prod_{i=1}^{j}\frac{c+1-i}{c+1+i}>\prod_{i=1}^{j}\frac{c-i}{c+i} \tag{64}$$

In order to see the correctness of the above inequality, consider any particular term in the product:

$$\frac{c+1-i}{c+1+i} > \frac{c-i}{c+i} \tag{65}$$

$$(c+1-i)(c+i) > (c+1+i)(c-i) \tag{66}$$

$$2i > 0 \tag{67}$$

From which we can conclude that the inequality over the entire product: $\prod_{i=1}^{j} \frac{c+1-i}{c+1+i} > \prod_{i=1}^{j} \frac{c-i}{c+i}$ must hold since it holds over each term and all are positive. By inductively applying the product inequality, we can then conclude the key inequality because $n > c$:

$$\prod_{i=1}^{j} \frac{n-i}{n+i} \geq \prod_{i=1}^{j} \frac{c-i}{c+i} \tag{68}$$

□

Theorem C.5 (Theorem 4.3)**.** *Suppose the data is generated from a linear model $Y = XW + \epsilon$ with $\mathbf{E}(\epsilon|X) = 0$. Assume each predictor is uses an ordinary least-squares linear regression. Let $s \geq 1$ be the number of i.i.d. seed samples each predictor starts with. We have the following:*

(i) If $\alpha > 0$ and $k \geq 2$ then $\lim_{t\to\infty} \sup_{\mathcal{D}} \frac{\mathcal{R}_t^k}{\mathcal{R}_t^1} \geq 1 + \frac{1}{3567 s^{3/2}}$

(ii) If $\alpha = \infty$ then $\lim_{t\to\infty} \sup_{\mathcal{D}} \frac{\mathcal{R}_t^k}{\mathcal{R}_t^1} \geq \frac{2k}{k+1}$

Proof. For both parts it will be helpful to recall the seminal result of White which shows the consistency of the OLS estimate for linear models [60]. Thus, $\mathcal{R}_t^1$ converges to the minimum mean square error.

Proof of (ii) We first prove part (ii). It is sufficient to construct a P_{XY} satisfying the linear model assumptions that can be easily analyzed. In this proof, we construct P_{XY} such that with perfect information and OLS updates, the learning dynamics reduce into the *sequential K-means* dynamics studied in [61]. To be more precise, the resultant random process over the tuple of each predictor's OLS weight estimates is almost the same as that over the tuple of centroids in sequential K-means, in a sense that will be made formal in this proof. First we restate the notion of a sequential K-means process, originally defined in [61].

Definition C.6. *(Sequential K-means process)*

Let P be a non-atomic distribution over $\mathbb{R}$ with bounded first and second moments. Let $V(t) \in \mathbb{R}^K$ be a vector-valued random process. At time $t = 0$, define $V_\kappa(0) \sim Q$ under the product coupling (i.e. independently sampled) and $N(t) = 0 \in \mathbb{N}^K$

We define $V_\kappa(t)$ for $t > 0$ recursively as follows. $Z(t) \sim Q$ is an i.i.d sample from the distribution. Let $i(t) = \operatorname{argmin}_{\kappa\in\{1,\dots,k\}} |V_\kappa(t) - Z(t)|$. Then, $N(t) = N(t-1) + \mathbf{1}_{i(t)}$ where $\mathbf{1}_{i(t)}$ is the one-hot vector at index $i(t)$ and

$$V_j(t) = \begin{cases} \frac{1}{N_j(t)}(N_j(t-1)V_j(t-1) + Z(t)) & j = i(t) \\ V_j(t-1) & j \neq i(t) \end{cases}$$

For sequential K-means processes, the following holds [61]:

Lemma C.7. *(J. MacQueen, 1967)*

Let $V \in \mathbb{R}^d$ be a a sequential k-means process. Then V converges a.s. to an unbiased partition *jointly satisfying:*

$$\textit{(i)} V_j(\infty) = \operatorname*{argmin}_{v\in\mathcal{S}_j} \int_{z\in\mathcal{S}_j} (v-z)^2 dQ(z)$$

and

(ii) $\mathcal{S}_j = \{z : j = \underset{[K]}{\operatorname{argmin}} |V_j(\infty) - z|\} \subset \mathbb{R}$

Equivalently, these conditions state that V_j is the centroid of its Voronoi interval [62, 63] with respect to the vector V.

We proceed to give a reduction from the competitive learning market with OLS predictors with a particular choice of P_{XY}. In turn, this will enable us to concretely characterize the weight estimates asymptotically via MacQueen's lemma.

Let $W = 1$. Let P_X be a point mass at 1. Let $\epsilon_i \sim \mathbf{Unif}(-\delta, \delta)$ for $0 < \delta$. In this construction, $P_Y \sim \mathbf{Unif}(1-\delta, 1+\delta)$ since $Y = 1 + \epsilon$. In the scalar case, the OLS rule is $\hat{W} = \frac{X^n \bullet Y^n}{||X^n||^2}$. For the assumed distribution, the weight estimate is determined by the empirical mean:

$$\hat{W} = \frac{\sum_i y_i}{n} = 1 + \bar{\epsilon} = \hat{Y} \tag{69}$$

Let $\hat{W}_t(i)$ denote the i-th predictor's weight estimate at round t, with $\hat{W}_0(i)$ being the initial estimate based on the the seed samples. Let $\nu_t(a) = s + \sum_{j=0}^{t} \mathbf{1}(\mathbf{SELECT}((j) = a)$ denote the total number of observations made by agent a by round t (where $\mathbf{1}()$ is the indicator function). When there is no ambiguity, we will write $a_t = \mathbf{SELECT}(t)$ as shorthand.

At round t, the OLS update is given by:

$$\hat{W}_t(a_t) \leftarrow \frac{\nu_t(i) - 1}{\nu_t(i)} \hat{W}_{t-1}(a_t) + \frac{1}{\nu_t(i)} y_t \tag{70}$$

Under the assumption of perfect information and rational consumers ($\alpha = \infty$) we have that:

$$a_t = \underset{a \in \mathcal{A}}{\operatorname{argmin}} |\hat{W}_t(a) - y_t| \tag{71}$$

In prose, at time t, the sample Y_t is averaged into the closest of the OLS estimates at time t. This is precisely the update rule use in [61] to describe the sequential K-means. However, while the update rule is the same, there remains one blemish that we must smooth over before we can apply MacQueen's lemma. Namely, that the intializations between our case and the sequential K-means are not the same, since here the seed dataset can consist of more than one example. It turns out, however, that this issue is easily remedied by noting that the tail of any sequential K-means process is conditionally independent of $Y(t)$ given only $V(t)$ and $N(t)$. In other words, for $\tau > 0$:

$$V(t+\tau) \perp\!\!\!\perp (Y(0), .., Y(t)) | (V(t), N(t)) \tag{72}$$

Therefore, when we have s seed points per k agents, we may construct an equivalence relation between the our process at time 0 and the sequential K-means process at time $(s-1)k$ via:

$$\forall i : N_i((s-1)k) \leftarrow \nu_i(0) = s \tag{73}$$

and

$$V((s-1)k) \leftarrow \hat{W}(0) \tag{74}$$

Thus, we can treat our process $\hat{W}_t$ as a sequential K-means process conditioned on the sequence head as just stated. From the definition of the sequential K-means process, it is clear that conditioning on this measure zero event remains well-posed. Furthermore, we can conclude that it does not alter the convergence of the tail because for any $I \subset \mathbb{R}^K$ we have that

$$\mathbf{Pr}(\hat{W}(0) \in I) > 0 \tag{75}$$

if and only if

$$\mathbf{Pr}(\{V((s-1)k) \in I\} \cap \{N_i((s-1)k) = s, \forall i\}) > 0 \tag{76}$$

which establishes absolute continuity [64] of the corresponding probability measures over the possible initialization. From this follows the a.s. convergence of $\hat{W}$ [65], and we may dispense with the issue of the initialization and conclude that $\hat{Y}$ converge a.s. to an unbiased partition (recalling that $\hat{Y}=\hat{W}$ for our chosen P_{XY}).

It remains to make use of MacQueen's lemma to finish our claim. To do so, we point out the well-known result that there is only one unbiased partition of the uniform distribution over an interval $(1-\delta,1+\delta)$ [66], namely, the uniform quantization that sets:

$$\hat{Y}_i(\infty)=\frac{i2\delta}{k+1}-\delta+1 \tag{77}$$

or some permutation thereof. Also notice that this result is translation invariant. Translating the interval corresponds to a translation of the quanta.

It remains to compute the expected MSE over the $Y_i(\infty)$, noting that *a priori*, each predictor is uniformly likely to converge to any of the quanta.

$$\mathbf{E}[(Y-\hat{Y}(\infty))^2]=\sum_{i=1}^{n}\mathbf{Pr}[\hat{Y}(\infty)=\frac{i2\delta}{k+1}-\delta+1]\mathbf{E}[(Y-\frac{i2\delta}{k+1}+\delta-1)^2] \tag{78}$$

$$=\sum_{i=1}^{k}\mathbf{Pr}[\hat{Y}(\infty)=\frac{i2\delta}{k+1}-\delta]\int_{-\delta}^{\delta}\frac{1}{2\delta}(\delta-\frac{i2\delta}{k+1}+y)^2dy \tag{79}$$

$$=\frac{1}{k}\sum_{i=1}^{k}\frac{1}{2\delta}\int_{-\delta}^{\delta}(\delta-\frac{i2\delta}{k+1}+y)^2dy \tag{80}$$

$$=\frac{1}{k}\sum_{i=1}^{k}\frac{1}{2\delta}\left(\frac{8\delta^3(3i^2+3i(k+1)+(k+1)^2)}{3(k+1)^2}\right) \tag{81}$$

$$=\frac{2\delta^2k}{3(k+1)} \tag{82}$$

To complete the proof, we point out that $W^*=1$ and that the MMSE for predicting Y given X is given by:

$$\mathbf{MSE}^*(P_{Y|X})=\frac{\delta^2}{3} \tag{83}$$

And that we may lower bound the $\sup_P\frac{\mathcal{R}_t^k}{\mathcal{R}_t^1}$ with our particular choice of P_{XY}.

Thus, taking the limit:

$$\lim_{T\to\infty}\sup_P\frac{\mathcal{R}_T^k}{\mathcal{R}_T^1}\geq\lim_{T\to\infty}\frac{\mathbf{E}[(Y-\hat{Y}(T)^2]}{\mathcal{R}_T^1}=\frac{\mathbf{E}[(Y-\hat{Y}(\infty)^2]}{\mathbf{MSE}^*(P_{Y|X})}=\frac{2k}{(k+1)} \tag{84}$$

This completes the proof of part (ii).

Proof of (i) We now turn out attention to the proof of part (i). It will suffice to construct a different distribution that can be easily analyzed. Let $X=\frac{1}{2}\Delta$ and define ϵ as follows:

$$\epsilon=\begin{cases}\frac{1}{2}\Delta & \text{with prob.}\frac{1}{2}\\ -\frac{1}{2}\Delta & \text{with prob.}\frac{1}{2}\end{cases}$$

Thus, we have that $Y=0$ with prob. $1/2$ and $Y=\Delta$ with prob. $1/2$. The MMSE estimate $W^*=1$ and the corresponding MSE is $\frac{1}{4}\Delta^2$. Let $\mu_i=\frac{1}{\#\text{ samples for }i}\sum_j y_j$ denote the empirical mean of the samples observed by agent i. The OLS estimate will satisfy $\hat{Y}_i=\frac{1}{2}\Delta\hat{W}_i=\mu_i$. Due to the OLS-specified bijection between $\hat{Y}$ and $\hat{W}$ we will work with $\hat{Y}$ for convenience without loss of precision.

Recall we have n predictors. At round $t=0$, each has been seeded with s i.i.d samples. Let μ_i be the empirical mean of the seed samples for the i-th predictor. Recall that OLS estimate $\hat{Y}_i = \mu_i$.

Let $i^+ = \text{argmax}\{\hat{Y}_i\}$ and $i^- = \text{argmin}\{\hat{Y}_i\}$. For now, we will assume that these extreme predictors are unique, and later address the case in which they are not.

Let $\hat{a}_t = \textbf{SELECT}(t)$. In the limit as $\Delta \to \infty$ we have that *only* the two extreme predictors, a_{i^+} and a_{i^-} will ever win consumer queries: $\mathbf{Pr}(\hat{a}_t = a_{i^+} | Y = \Delta) \to 1$ and $\mathbf{Pr}(\hat{a}_t = a_{i^-} | Y = 0) \to 1$ as $\Delta \to \infty$. To see this, observe that:

$$\frac{\mathbf{Pr}(\hat{a}_t = a_{i^+} | Y = \Delta)}{\mathbf{Pr}(\hat{a}_t = a_j | Y = \Delta)} = \frac{\exp\big(\alpha(\Delta - \mu_{i^+})^2\big)}{\exp(\alpha(\Delta - \mu_j)^2)} = \exp\big(\alpha(\mu_{i^+}^2 - \mu_j^2 + 2\Delta(\mu_{i^+} - \mu_j))\big) \tag{85}$$

And so, when $j \neq i^+$ and $\alpha > 0$:

$$\lim_{\Delta \to \infty} \frac{\mathbf{Pr}(\hat{a}_t = a_{i^+} | Y = \Delta)}{\mathbf{Pr}(\hat{a}_t = a_j | Y = \Delta)} = \infty \tag{86}$$

because of the fact that $\mu_{i^+} \geq \mu_j$. In the case when the inequality is strict, we can immediately conclude that:

$$\lim_{\Delta \to \infty} \mathbf{Pr}(\hat{a}_t = a_{i^+} | Y = \Delta) = 1 \tag{87}$$

and more precisely, for some $\varphi > 0$

$$\mathbf{Pr}(\hat{a}_t = a_{i^+} | Y = \Delta) = 1 - \Theta(e^{-\varphi \Delta}) \tag{88}$$

Due to the fact that number outcomes (i.e. the consumer decision) is finite and a probability vector is normalized. If the argmax is not unique, then in the first round at which $Y = \Delta$, one of the maximizing predictors will be selected arbitrarily, which will break the equality. Along those lines, note that $\hat{W}_{i^+}$ is monotone increasing after observing more samples of $Y = \Delta$. A similar argument holds in the case when $Y = 0$.

Combining these two cases, we conclude that only the 2 extreme predictors will ever receive additional samples, whereas the other $n-2$ predictors will maintain the weight estimates based solely on their seed samples.

Later on, it will be useful to let Δ scale with the number of rounds T. This does not pose a difficulty when the scaling is chosen judiciously. With $\Delta = \sqrt{T}$, it is easy to verify that the aforementioned limits hold over all rounds. Let $\mathcal{E}(\tau)$ be the event that $\hat{a}_t = a_{i^+}$ for all $1 \leq t < \tau$ with $Y_t = \Delta$. Then:

$$\mathbf{Pr}(\mathcal{E}(T)) \geq \prod_{1 \leq t < T} \mathbf{Pr}(\hat{a}_t = a_{i^+} | Y_t = \Delta, \mathcal{E}(t-1)) \geq \prod_{1 \leq t < T} \mathbf{Pr}(\hat{a}_0 = a_{i^+} | Y_0 = \Delta) \tag{89}$$

because of the conditional independence of consumer decisions and monotone increasing trajectory of $\hat{W}_{i^+}$ in time. However, recalling the convergence rate in Δ previously established:

$$\mathbf{Pr}(\mathcal{E}) = \Theta(1 - e^{-\varphi\sqrt{T}})^T \tag{90}$$

And thus, with the superpolynomial convergence we still obtain:

$$\lim_{T \to \infty} \mathbf{Pr}(\mathcal{E}) = 1 \tag{91}$$

We will return to this analysis later to complete the proof. For now, we can proceed to lower bound the gap by taking a weighted average of the MSE in between the extremal and non-extremal predictors. For the two extreme predictors, we will have that $\hat{Y} \to \Delta$ and $\hat{Y} \to 0$. In this case the MSE is $\frac{1}{2}\Delta^2$ for both.

For the remaining non-extremal predictors, we can bound the expected MSE as follows. Since they do not obtain further samples beyond the seed set, we can directly analyze the expected MSE of the OLS estimate under s seed samples. However, we must still account for the fact that we are conditioning on the event that these are non-extremal estimates. This is rather cumbersone and is difficult to do exactly, but we can use the following 3 steps of approximations to obtain a lower bound.

(i) The conditional variance of a non-extremal estimate is most reduced when $n = 3$.

$$\hat{Y}_{i^+} = \mathbf{max}\{\hat{Y}_j\}_{j=1}^n \tag{92}$$

$$\hat{Y}_{i^-} = \mathbf{min}\{\hat{Y}_j\}_{j=1}^n \tag{93}$$

Thus, we have that $\hat{Y}_j \in [\hat{Y}_{i^-}, \hat{Y}_{i^+}]$ for all j. Recall we are seeking a lower bound on

$$\mathbf{Var}(\hat{Y}_j | j \neq i^+, j \neq i^-) = \mathbf{Var}(\hat{Y}_j | \hat{Y}_j \in [\hat{Y}_{i^-}, \hat{Y}_{i^+}]) \tag{94}$$

From this, we immediately obtain that for $n'' \geq n'$:

$$\mathbf{Var}(\hat{Y}_j | j \neq i^+, j \neq i^-, n = n') \leq \mathbf{Var}(\hat{Y}_j | j \neq i^+, j \neq i^-, n = n'') \tag{95}$$

because the cdf of $\hat{Y}_{i^+}$ ($\hat{Y}_{i^-}$) is monotone increasing (decreasing) in n, given that it is the maximum (minimum) of n i.i.d random variables. Therefore, lowering bounding the case when $n = 3$ is sufficient to lower bound all cases.

(ii) We may lower bound the probability of a lower tail deviations of $\hat{Y}_{i^-}$ by ignoring that it is the minimum – i.e. treating it as any generic i.i.d sample – and then by using the standard bound [55]:

$\mathbf{Pr}(\mathbf{Bi}(s, \frac{1}{2}) < k) \geq \frac{1}{15} \exp\left(\frac{-16}{k}(k/2 - k)^2\right)$

(Recalling that $\hat{Y} \sim \frac{\Delta}{2s}\mathbf{Bi}(s, \frac{1}{2})$).

Setting a deviation of $\frac{\sqrt{s}}{4}$ results in a lower bound of $\frac{1}{15e}$ on the lower tail. We can use symmetry to apply the argument to $\hat{Y}_{i^+}$ as a lower bound to the upper tail. Thus, we have that

$$\mathbf{Pr}(\frac{\Delta}{2} + \frac{\Delta\sqrt{s}}{4} \leq \hat{Y}_{i^-} \leq \frac{\Delta}{2} - \frac{\Delta\sqrt{s}}{4}) \geq \frac{1}{1764} \tag{96}$$

(iii) Once an interval has been established, we can treat the non-extremal estimate $\hat{W}_j$ as a truncated Binomial. By Lemma C.4, we can lower bound the variance of the truncated Binomial with a Binomial over the truncated support. If $Z \sim \mathbf{Bi}(s, \frac{1}{2})$ then

$$\mathbf{Var}\left(Z | \{\frac{\sqrt{s}}{4} \leq Z - \frac{s}{2} \leq \frac{\sqrt{s}}{4}\}\right) \geq \frac{\sqrt{s}}{8} \tag{97}$$

and

$$\mathbf{Var}\left(\frac{\Delta}{2s} Z | \{\frac{\sqrt{s}}{4} \leq Z - \frac{s}{2} \leq \frac{\sqrt{s}}{4}\}\right) \geq \frac{\Delta}{16 s^{3/2}} \tag{98}$$

Recalling that $\hat{Y}_j \sim \frac{\Delta}{2s}\mathbf{Bi}(s, \frac{1}{2})$, we have the variance conditioned on event $\mathcal{H} = \{j \neq i^+, j \neq i^-\}$ is lower bounded as follows:

$$\mathbf{Var}(Y_j | \mathcal{H}) = \mathbf{Var}(Y_j | \mathcal{H}, \mathcal{F})\mathbf{Pr}(\mathcal{F}) + \mathbf{Var}(Y_j | \mathcal{H}, \bar{\mathcal{F}})(1 - \mathbf{Pr}(\mathcal{F})) \geq \mathbf{Var}(Y_j | \mathcal{H}, \mathcal{F})\mathbf{Pr}(\mathcal{F}) \tag{99}$$

And since we know have a bound for $\mathbf{Pr}(\mathcal{F})$ and $\mathbf{Var}(Y_j | \mathcal{H}, \mathcal{F})$ we conclude:

$$\mathbf{Var}(Y_j | \mathcal{H}) \geq \mathbf{Var}(Y_j | \mathcal{H}, \mathcal{F})\mathbf{Pr}(\mathcal{F}) \geq \frac{\Delta^2}{(7056 \times 4) s^{3/2}} \tag{100}$$

Where $\mathcal{F} = \{\hat{Y}_{i^-} \leq \frac{\Delta}{2} - \frac{\Delta\sqrt{s}}{4}, \hat{Y}_{i^-} \geq \frac{\Delta}{2} + \frac{\Delta\sqrt{s}}{4}\}$

Note that $\hat{Y}_j$ is unbiased. From this, it follows that the excess MSE risk is given by its variance. Define $\delta = \hat{Y}_j - \mathbf{E}\hat{Y}_j = \hat{Y}_j - \frac{\Delta}{2}$.

$$\mathbf{MSE}(\hat{Y}_j) = \mathbf{E}[\frac{1}{2}(\frac{\Delta}{2} - \delta)^2 + \frac{1}{2}(\frac{\Delta}{2} + \delta)^2] = \frac{\Delta^2}{4} + \mathbf{E}\delta^2 = \mathbf{MSE}^* + \mathbf{Var}(\hat{Y}_j)$$

Note that we have established that for the extremal predictors, the asymptotic MSE is $\frac{\Delta^2}{2}$, and for the non-extremal predictors, the asymptotic MSE is at least $\frac{\Delta^2}{4}+\frac{\Delta^2}{(7056\times4)s^{3/2}}$. Combining these in-expectation gives: $\frac{k-2}{k}(\frac{\Delta^2}{4}+\frac{\Delta^2}{(7056\times4)s^{3/2}})+\frac{2}{k}\frac{\Delta^2}{2}$

Thus, to conclude the proof, let us scale $\Delta=\sqrt{T}$:

$$\lim_{T\to\infty}\sup_P\frac{\mathcal{R}_T^k}{\mathcal{R}_T^1}\geq\lim_{T\to\infty}\frac{\frac{k-2}{k}(\frac{\Delta^2}{4}+\frac{\Delta^2}{(7056\times4)s^{3/2}})+\frac{2}{k}\frac{\Delta^2}{2}}{\frac{\Delta^2}{4}}\geq1+\frac{1}{7056s^{3/2}}+\frac{2}{k}$$

.

which yields the inequality in part (ii).

□

C.2.4 Proof of Theorem 4.4.

As a clarification, when we refer to pairwise covariance ρ, we refer to:

$$\rho=\mathbf{Cov}(W_1,W_2)=\mathbf{E}(W_1W_2)-\mathbf{E}(W_1)\mathbf{E}(W_1) \tag{101}$$

where $W_1=\mathbf{1}(\hat{Y}^1=Y)$ and $W_2=\mathbf{1}(\hat{Y}^2=Y)$ for $\hat{Y}^1$ and $\hat{Y}^2$ denote the predictions from agent 1 and agent 2 respectively. Notice that $\mathbf{E}(W_1W_2)$ equals the probability that both predictions are correct. At any given time t, the quantity $\mathbf{E}(W_1W_2)$ itself is stochastic, given that it depends on the particular sample path taken by the random competition up until the given time. In our proof, we will use $P(t)$ to denote the measure over predictor correctness at a given time t

In order to proceed, we must first recall the definition of $\mathbb{A}_\tau$ from the main text: $\mathbb{A}_\tau=\mathbf{E}(\mathbf{1}\{\hat{Y}_\tau^{(w_\tau)}=Y_\tau\})$. Note that this expectation implicitly takes place over the entire randomness in the learning competition $\mathcal{G}$ (refer to Def. C.1). For the purposes of this Theorem, from a formal perspective, we will be comparing the the quantities $\mathbb{A}_\tau$ corresponding to different instances of $\{\mathcal{G}_\kappa\}_{\kappa=1}^\infty$ that vary only in the number of predictors. We write $\mathbb{A}_\tau^k$ to refer to the expectation over the randomness of learning competition $\mathcal{G}_k$ with k predictors:

$$\mathbb{A}_\tau^k=\mathbf{E}_{\mathcal{G}_k}(\mathbf{1}\{\hat{Y}_\tau^{(w_\tau)}=Y_\tau\})$$

Theorem C.8. *Assume a learning competition at round τ. Define $\mathcal{A}(A^{(i)};\mathcal{D})=1-\mathcal{R}(A^{(i)};\mathcal{D})$ and $\mathcal{A}_k(t)=1-\mathcal{R}_t^k$. When the parameter t is omitted, assume $t=\tau$: $\mathcal{A}_k=\mathcal{A}_k(\tau)$. Define $\delta=\mathcal{A}_1-\mathcal{A}_2$ and $\varepsilon=\mathcal{A}_k(0)-\frac{1}{2}$. Let ρ be the pairwise covariance between two predictors. Assume the following holds:*

1. *There is su cient data to train two predictors well: $0<\delta<\frac{1}{6}$*
2. *Predictors are weak with only seed data: $\varepsilon<1/14$*
3. *The predictors are not too correlated: $\rho<\mathcal{A}_k-\mathcal{A}_k^2-6\delta$*
4. *The expected accuracy of a predictor monotonically increases with dataset size: $\mathbf{E}_{\mathcal{G}}\left[\mathcal{A}(A_t^{(i)};\mathcal{D})|r=|D_t^{(i)}|\right]$*

Then there exists $0<c_1<c_2<\infty$ such that if $c_1<\alpha<c_2$ then $\mathbb{A}_\tau^k$ at round τ is maximized by some k^* number of predictors such that $1<k^*<\infty$. In particular, $c_1<\log\frac{\mathcal{A}_1-(\mathcal{A}_1-\delta)^2-\rho}{\mathcal{A}_1-(\mathcal{A}_1-\delta)^2-\rho-2\delta}$ and $c_2>\log\frac{(1-4\varepsilon)\mathcal{A}_1}{1-\mathcal{A}_1}$.

Proof. For notational convenience, we slightly modify the notation for expected accuracy $\mathcal{A}_k(t)$ compared with the notation $\mathcal{A}_t^k$ introduced in the main text. We have move the time index t from a subscript into a parenthetical function argument and moved the number of predictors k from a superscript to a subscript.

By assumption, we have that predictors at round 0 are weak predictors, meaning they are independently accurate with probability $\mathcal{A}_k(0)=\frac{1}{2}+\varepsilon$. Note we thus have $\varepsilon<1/14$ by the assumption. As we will see shortly, in the limit of infinite predictors, the algorithm's performance after obtaining additional samples is immaterial.

As defined in the main text, let denote $\mathbb{A}_\tau^k$ the expected prediction quality for users at time τ of $\mathcal{G}_k$. In other words, $\mathbb{A}_\tau^k$ is the expected prediction quality for users when the competition has k learners. Let $\hat{Y}_\tau^{(i)}$ be the i-th learner's prediction at time τ (or equivalently let $\hat{Y}_\tau^{(a)}$ be learner a's prediction). When it is unambiguous to do so, we will define $w_\tau=\mathbf{SELECT}(\mathbf{q}_\tau)$ as described in Section 2 of the main text. Because most of the variables in this proof are implicit at time τ, when the time is not explicitly stated or sub-scripted, assume that the variable refers to time τ, henceforth.

$$\mathbb{A}_\tau^k = \sum_{a \in \mathsf{A}} \mathbf{Pr}[a = w_\tau] \mathbf{Pr}[\hat{Y}_\tau^{(a)} = Y_\tau | a = w_\tau] \tag{102}$$

where we recall from the definition of **SELECT**:

$$\mathbf{Pr}[a = w_\tau] = \frac{\exp(\alpha \mathbf{1}\{\hat{Y}_\tau^{(a)} = Y_\tau\})}{Z_\tau} \tag{103}$$

where $Z_\tau = \sum_{a \in \mathcal{A}} \exp(\alpha \mathbf{1}\{\hat{Y}_\tau^{(a)} = Y\})$.

We proceed to give an expression for $\mathbf{Pr}[a = a_t]$ in the limit as $k \to \infty$. Let $\mathsf{B}_\tau \subset \mathsf{A}$ be the subset of predictors that have been queried at least once by some time τ (subscript on the B is omitted when it may be safely inferred). It follows that $\frac{|\mathsf{B}|}{|\mathsf{A}|} \leq \frac{\tau}{k}$ which implies $\frac{|\mathsf{B}|}{|\mathsf{A}|} \to 0$ as $k \to \infty$. From this, it follows that the $\mathbf{Pr}[a \in \mathsf{B}] \leq \frac{\tau e^\alpha}{k}$ which also vanishes as k gets large. With this, we can revise Eq. 102 as follows:

$$\mathbb{A}_\tau^k = \sum_{a \in \mathsf{B}} \mathbf{Pr}[a = w_\tau | a \in \mathsf{B}] \mathbf{Pr}(\hat{Y}^a = Y | a = w_\tau) + \sum_{a \notin \mathsf{B}} \mathbf{Pr}[a = w_\tau | a \notin \mathsf{B}] \mathbf{Pr}(\hat{Y}^a = Y | a \notin \mathsf{B}) \tag{104}$$

Here, notice that if $w_\tau \notin \mathsf{B}$, then $\mathbf{Pr}(\hat{Y}^a = Y | a \notin \mathsf{B}) = \mathbf{Pr}(\hat{Y}^a = Y | a = w_\tau)$ because all predictors that have not been selected are modeled as weak learners and are thus interchangeable.

We give the following lower and upper bounds for $\mathbb{A}$, which can be easily derived using the law of total probability [59] and the fact that $0 \leq \mathbb{A} \leq 1$:

$$\mathbf{Pr}(w_t \notin \mathsf{B}_t)\big(\mathbb{A}_t^k | \{w_t \notin \mathsf{B}_t\}\big) \leq \mathbb{A}_t^k \leq \mathbf{Pr}(w_t \notin \mathsf{B}_t)\big(\mathbb{A}_t^k | \{w_t \notin \mathsf{B}_t\}\big) + \mathbf{Pr}[w_t \in \mathsf{B}_t] \tag{105}$$

Where because $\mathbb{A}_t^k$ is already an expectation, we use the notation $\mathbb{A}_t^k | \{a_t \notin B_t\}$ to mean the conditional expectation.

Taking the limit in k vanishes $\mathbf{Pr}[a \in \mathsf{B}_t]$, yielding:

$$\mathbb{A}_t^\infty = \mathbb{A}_t^\infty | \{w_t \notin \mathsf{B}_t\} \tag{106}$$

We proceed to compute this quantity $\mathbb{A}^\infty | \{a_t \notin \mathsf{B}\}$. To do so, we make use of the weak predictor assumption. From that, may treat the aggregate predictions from predictors not in B_t as following a Binomial distribution with success probability of $\frac{1}{2} + \varepsilon$. Define $\kappa_t = |\mathsf{A} / \mathsf{B}_t| \in [k-t, k-1] = \Theta(k)$. Let V_t be the number of weak predictors with correct predictions at time t. Then $V \sim \mathbf{Bin}(\kappa_t, \frac{1}{2} + \varepsilon)$. Let μ denote the mean of V, $\mathbf{E}V = \mu$. Then, the conditional probability that the consumer at t selects any correct weak predictor is:

$$\mathbf{Pr}(\hat{Y}^{w_t} = Y | w_t \notin \mathsf{B}_t, V = \nu) = \frac{\nu e^\alpha}{\nu(e^\alpha - 1) + \kappa} \tag{107}$$

Let us rewrite V in terms of its deviation from its mean: $V = \mu + \Delta$ for implicitly defined random deviation Δ. By the central limit theorem [59, 64, 67, 68], we know that deviations of the Binomial with κ trials [59, 64, 67] are order $\Theta(\sqrt{k})$ implying $\mathbf{Pr}(|\Delta| \geq Ck^{0.6}) \to 0$ as $k \to \infty$ for any finite C. Thus, let us introduce the deviation into the above equation and normalize by κ:

$$\mathbf{Pr}(\hat{Y}^{w_t} = Y | w_t \notin \mathsf{B}_t, \Delta = \boldsymbol{\Delta}) = \frac{(\frac{\mu}{\kappa} + \frac{\boldsymbol{\Delta}}{\kappa}) e^\alpha}{(\frac{\mu}{\kappa} + \frac{\boldsymbol{\Delta}}{\kappa})(e^\alpha - 1) + 1} = \frac{(\frac{1}{2} + \varepsilon + \frac{\boldsymbol{\Delta}}{\kappa}) e^\alpha}{(\frac{1}{2} + \varepsilon + \frac{\boldsymbol{\Delta}}{\kappa})(e^\alpha - 1) + 1} \tag{108}$$

By the aforementioned line of reasoning, because the the risk is bounded, we may ignore any large deviations for Δ with respect the mean when taking a limit in k. Computing the limit is direct and yields:

$$\lim_{k \to \infty} \frac{(\frac{1}{2} + \varepsilon + \frac{\boldsymbol{\Delta}}{\kappa}) e^\alpha}{(\frac{1}{2} + \varepsilon + \frac{\boldsymbol{\Delta}}{\kappa})(e^\alpha - 1) + 1} = \frac{(\frac{1}{2} + \varepsilon) e^\alpha}{(\frac{1}{2} + \varepsilon)(e^\alpha - 1) + 1} = \frac{e^\alpha}{e^\alpha + 1 - \chi} \tag{109}$$

For $\chi = \frac{4\varepsilon}{2\varepsilon + 1}$. Notice that $\varepsilon < \frac{1}{6}$ implies $\chi < \frac{1}{2}$. From this we establish: $\mathbb{A}^\infty = \frac{e^\alpha}{e^\alpha + 1 - \chi}$. By definition, we know $\mathbb{A}^1 = \mathcal{A}_1$.

Given that the joint distribution for 2 predictors is determined by the marginals and the covariance, we can solve for $\mathbb{A}^2$. We introduce P, the probability measure over the correctness of each of the two predictors at a fixed

time t (the dependence from P on t is only stated here and omitted in notation). P_{11} denotes the probability that both predictors are correct and P_{00} denotes the probability that both are incorrect. P_{10} denotes the probability that the first predictor is correct and the second is incorrect and P_{01} denotes the probability of the opposite. There are the four possible outcomes under measure P for only two predictors ($k=2$). Of course, P itself is a randomized object, given that it depends on the randomness generated by the competition $\mathcal{G}$ until time t.

We find $\mathbb{A}_t^2$ in terms of P:

$$\mathbb{A}_t^2 = \mathbf{E}_{P\sim\mathcal{G}_2}(P_{11} + \frac{e^\alpha}{e^\alpha+1}(P_{10}+P_{01})) \tag{110}$$

where the above follows directly from the structure of the competition.

From the definition of covariance we know that:

$$\rho = P_{11} - (P_{11}+P_{10})(P_{11}+P_{01}) \tag{111}$$

Substituting for P_{11} yields:

$$\mathbb{A}_t^2 = \mathbf{E}_{P\sim\mathcal{G}_2}\left[\rho + (P_{11}+P_{10})(P_{11}+P_{01}) + \frac{e^\alpha}{e^\alpha+1}(P_{10}+P_{01})\right] \tag{112}$$

Furthermore (we drop the explicit expectation over $\mathcal{G}_2$):

$$\mathbb{A}_t^2 = \mathbf{E}[\rho] + \mathbf{E}[(P_{11}+P_{10})(P_{11}+P_{01})] + \frac{e^\alpha}{e^\alpha+1}\mathbf{E}(P_{10}+P_{01}) \tag{113}$$

And notice that $P_{11}+P_{10}$ is the marginal probability that the first predictor is correct. Thus, $P_{11}+P_{10} = \mathcal{A}(A^{(1)};\mathcal{D})$ and $P_{01}+P_{11} = \mathcal{A}(A^{(2)};\mathcal{D})$ by definition (for predictor i we may omit the parenthesis in the superscript, i.e. we write A^i instead of $A^{(i)}$).

$$\mathbb{A}_t^2 = \mathbf{E}[\rho] + \mathbf{E}[\mathcal{A}(A^1;\mathcal{D})\mathcal{A}(A^2;\mathcal{D})] + \frac{e^\alpha}{e^\alpha+1}\mathbf{E}(P_{10}+P_{01}) \tag{114}$$

where $\sim$ denotes equality of distribution.

Below, we break up the expectation of $(P_{10} + P_{01})$ into two terms and we add $P_{11} - P_{11}$ and $(P_{11}+P_{10})(P_{11}+P_{01}) - (P_{11}+P_{10})(P_{11}+P_{01})$ to each term for a net effect of zero.

$$\mathbf{E}(P_{10}+P_{01}) = \mathbf{E}[P_{10}+P_{11}-P_{11}+(P_{11}+P_{10})(P_{11}+P_{01})-(P_{11}+P_{10})(P_{11}+P_{01})] \tag{115}$$

$$+\mathbf{E}[P_{01}+P_{11}-P_{11}+(P_{11}+P_{10})(P_{11}+P_{01})-(P_{11}+P_{10})(P_{11}+P_{01})] \tag{116}$$

Regrouping the terms produces:

$$\mathbf{E}(P_{10}+P_{01}) = \mathbf{E}\left[\mathcal{A}(A^1;\mathcal{D}) - \rho - \mathcal{A}(A^1;\mathcal{D})\mathcal{A}(A^2;\mathcal{D})\right] + \mathbf{E}\left[\mathcal{A}(A^2;\mathcal{D}) - \rho - \mathcal{A}(A^1;\mathcal{D})\mathcal{A}(A^2;\mathcal{D})\right] \tag{117}$$

As a shorthand, let $\mathfrak{a} = \frac{e^\alpha}{e^\alpha+1}$.

Substituting it back into the full equation yields:

$$\mathbb{A}_t^2 = (1-2\mathfrak{a})\mathbf{E}[\rho] + \mathfrak{a}\mathbf{E}(\mathcal{A}(A^1;\mathcal{D})) + \mathfrak{a}\mathbf{E}(\mathcal{A}(A^2;\mathcal{D})) + (1-2\mathfrak{a})\mathbf{E}[\mathcal{A}(A^1;\mathcal{D})\mathcal{A}(A^2;\mathcal{D})] \tag{118}$$

Notice that we may assume $\mathcal{A}(A^2;\mathcal{D})$ and $\mathcal{A}(A^1;\mathcal{D})$ are ***negatively*** correlated. This follows from the assumption that $\mathbf{E}_{\mathcal{G}}\left[\mathcal{A}(A_t^{(i)};\mathcal{D})|r=|D_t^{(i)}|\right]$ is monotone increasing in r. Because $|D_t^{(1)}|+|D_t^{(2)}| = t-2s$, as one predictor gets more data, the other must get less (the $-2s$ term comes from accounting for the seed sets). Thus, by assumption:

$$\mathbf{E}[\mathcal{A}(A^1;D)\mathcal{A}(A^2;D)] \leq \mathbf{E}[\mathcal{A}(A^1;D)]\mathbf{E}[\mathcal{A}(A^2;D)] \tag{119}$$

Because $\mathfrak{a} \geq 1/2$ we know that $(1-2\mathfrak{a}) \leq 0$:

$$(1-2\mathfrak{a})\mathbf{E}[\mathcal{A}(A^1;D)\mathcal{A}(A^2;D)] \geq (1-2\mathfrak{a})\mathbf{E}[\mathcal{A}(A^1;D)]\mathbf{E}[\mathcal{A}(A^2;D)] \tag{120}$$

Implying,

$$\mathbb{A}_t^2 \geq (1-2\mathfrak{a})\mathbf{E}[\rho] + \mathfrak{a}\mathbf{E}(\mathcal{A}(A^1;\mathcal{D})) + \mathfrak{a}\mathbf{E}(\mathcal{A}(A^2;\mathcal{D})) + (1-2\mathfrak{a})\mathbf{E}[\mathcal{A}(A^1;\mathcal{D})]\mathbf{E}[\mathcal{A}(A^2;\mathcal{D})] \tag{121}$$

$$\mathbb{A}_t^2 = \mathbf{E}[\rho] + \mathbf{E}[\mathcal{A}(A_1;D)\mathcal{A}(A_2;D)] + \frac{e^\alpha}{e^\alpha+1}\mathbf{E}(P_{10}+P_{01}) \tag{122}$$

By the symmetry assumption (i.e. predictors 1 and 2 use the same algorithm), we have that:

$$\mathcal{A}(A_1;\mathcal{D}) \sim \mathcal{A}(A_2;\mathcal{D}) \tag{123}$$

Implying,

$$\mathbb{A}_t^2 \geq (1-2\mathfrak{a})\mathbf{E}[\rho] + 2\mathfrak{a}\mathbf{E}(\mathcal{A}(A^2;\mathcal{D})) + (1-2\mathfrak{a})\mathbf{E}[\mathcal{A}(A^2;\mathcal{D})]^2 \tag{124}$$

Since ρ is given by assumption, we have the simplification that $\mathbf{E}\rho = \rho$. Further, recalling the definition of $\mathcal{A}_2$ yields the simplification that $\mathbf{E}\mathcal{A}(A^2;\mathcal{D}) = \mathcal{A}_2(t)$. Simplifying and rearranging yields:

$$\mathbb{A}_t^2 \geq \rho + \mathcal{A}_2^2 + \frac{2e^\alpha(\mathcal{A}_2 - \rho - \mathcal{A}_2^2)}{e^\alpha+1} \tag{125}$$

This lower bound on $\mathbb{A}^2$ will suffice to get us through the rest of the argument. Note that if both $\mathbb{A}^2 > \mathbb{A}^1$ and $\mathbb{A}^2 > \mathbb{A}^\infty$ hold, then the theorem stands. For expediency, we will prove a sufficient condition instead. Namely, we will show that given the assumed initial conditions, we have that $\mathbb{A}^2 > \mathbb{A}^1 > \mathbb{A}^\infty$. The reason for choosing this particular order is not fundamental since the potential order $\mathbb{A}^2 > \mathbb{A}^\infty > \mathbb{A}^1$ would also be sufficient. However, the since the quantity $\mathbb{A}^1$ is trivial, it is simpler to tame the pairwise comparisons $\mathbb{A}^2 > \mathbb{A}^1$ and $\mathbb{A}^1 > \mathbb{A}^\infty$ than the comparison between $\mathbb{A}^2 > \mathbb{A}^\infty$.

Before proceeding, we briefly pause to provide the reader with a bit of intuition. In a sense, we can see that $\mathbb{A}^2$ is larger than $\mathbb{A}^1$ when the boost gained from having the user selection outweighs the penalty due to the competition between the predictors. Thus, we need a ***lower bound*** on α in order for $\mathbb{A}^2 > \mathbb{A}^1$. On the other hand, if α is too large, then we get too much of a boost from user selection in the $k \to \infty$ limit making $\mathbb{A}^\infty$ too large. If we want $\mathbb{A}^2 > \mathbb{A}^\infty$ we also need an ***upper bound*** on α. We proceed to quantify these bounds and show that under the assumed conditions, in between these bounds remains the ***sweet spot*** interval for α in which both $\mathbb{A}^2 > \mathbb{A}^\infty$ and $\mathbb{A}^2 > \mathbb{A}^1$ which implies the non-monotonic phenomena.

Let us examine the constraints placed on key quantities $\mathcal{A}_1$, ρ, δ, ε, and α based on the two inequalities. Recall $\mathcal{A}_2 = \mathcal{A}_1 - \delta$. Of course, there are an intrinsic set of constraints on these five quantities by assumption or definition:

$$1 > \mathcal{A}_1 > \frac{2}{3} \tag{126}$$

$$\frac{1}{6} > \delta > 0 \tag{127}$$

$$\alpha \geq 0 \tag{128}$$

$$1 \geq \rho \geq -1 \tag{129}$$

$$\frac{1}{4} > \chi > 0 \tag{130}$$

As discussed previously, there is a 1-to-1 mapping between χ to ε given by$\chi = \frac{4\varepsilon}{2\varepsilon+1}$. This mapping happens to be concave in ε to χ, so any tangent line to the curve serves as on overestimate from ε to χ. Noting that $(\chi,\varepsilon) = (\frac{1}{4},\frac{1}{14})$ lies on the curve lets us conclude that $\chi(\varepsilon) < 4\varepsilon$. This will later on helpful when transforming bounds into terms of ε.

The question remains what subset of this region remains after imposing that $\mathbb{A}^2 > \mathbb{A}^1 > \mathbb{A}^\infty$. In other words, how should we refine this subset in order to accomplish the desired relationships in $\mathbb{A}$? Let us begin by analyzing $\mathbb{A}^2 > \mathbb{A}^1$:

$$\rho + (\mathcal{A}_1 - \delta)^2 + \frac{2e^\alpha(\mathcal{A}_1 - \delta - \rho - (\mathcal{A}_1 - \delta)^2)}{e^\alpha + 1} > \mathcal{A}_1 \tag{131}$$

It is not difficult to check (perhaps by using a computer algebra system such as Mathematica[1]) that this inequality is satisfied if the additional constraints hold:

$$\alpha > \log\frac{\psi - \rho}{\psi - \rho - 2\delta} \text{ and } \psi - 2\delta > \rho \text{ where } \psi = \mathcal{A}_1 - (\mathcal{A}_1 - \delta)^2 \tag{132}$$

Loosely, these conditions say that ρ cannot be too large. This is natural since the users gain no benefit from two learners when they are perfectly correlated.

We can proceed to look at the second pairing: $\mathbb{A}^1 > \mathbb{A}^\infty$:

$$\mathcal{A}_1 > \frac{e^\alpha}{e^\alpha + 1 - \chi} \tag{133}$$

Elementary algebra yields:

$$\log\frac{(1-\chi)\mathcal{A}_1}{1 - \mathcal{A}_1} > \alpha \tag{134}$$

It remains to show that this interval $\log\frac{(1-\chi)\mathcal{A}_1}{1-\mathcal{A}_1} > \alpha > \log\frac{\psi-\rho}{\psi-\rho-2\delta}$ is guaranteed to exist.

To do this, note that since $0 < \chi < \frac{1}{6}$ and $1 > \mathcal{A}_1 > 2/3$ we have that:

$$\frac{(1-\chi)\mathcal{A}_1}{1-\mathcal{A}_1} > \min_{0<x<\frac{1}{6}, \frac{2}{3}<a<1} \frac{(1-x)a}{1-a} = \frac{5}{3} \tag{135}$$

Thus we are left to ponder the inequality: $5/3 > \frac{\psi-\rho}{\psi-\rho-2\delta}$. As before, with the aid of computer algebra[2] we can translate the above inequality into constraints on the quantities of interest. The following additional constraint on ρ is sufficient to imply $\mathbb{A}^2 > \mathbb{A}^\infty$:

$$\rho < \psi - 6\delta \tag{136}$$

To conclude, then there exists an non-empty interval for α given by:

$$\log\frac{(1-\chi)\mathcal{A}_1}{1-\mathcal{A}_1} > \alpha > \log\frac{\psi-\rho}{\psi-\rho-2\delta} \tag{137}$$

such that we can guarantee:

$$1 < k^* < \infty \tag{138}$$

if the following also hold:

$$\mathcal{A}_1 > \frac{2}{3} \tag{139}$$

$$\frac{1}{6} > \delta \tag{140}$$

[1] For example, using the following one line Mathematica command: `Reduce[{1 > rho > -1, expalpha > 1, 1 > A1 > 2/3, 1/5 > delta > 0, (-delta + A1)^2 + (2 expalpha (-delta + A1 - (-delta + A1)^2 - rho))/(1 + expalpha) + rho > A1}, {expalpha}]`

[2] For example, using the following one line Mathematica command: `Reduce[{A1-A1^2 -2delta + 2A1 delta - delta^2 > rho >-1 ,1/5>delta>0, 1 > A1 > 2/3, (A1 - A1^2 + 2A1 delta - delta^2 - rho)/(A1 - A1^2 - 2 delta +2 A1 delta - delta^2 - rho) <5/3},{rho}]`

$$\frac{1}{4} > \chi \tag{141}$$

$$\mathcal{A}_2 - \mathcal{A}_2^2 - 6\delta > \rho \tag{142}$$

Finally, we translate back to ε from χ by nothing that since $\chi < 4\varepsilon$ it follows: $\log \frac{(1-\chi)\mathcal{A}_1}{1-\mathcal{A}_1} > \log \frac{(1-4\varepsilon)\mathcal{A}_1}{1-\mathcal{A}_1}$. Transforming the constraints then yield:

$$\frac{1}{14} > \varepsilon \tag{143}$$

$$\log \frac{(1-4\varepsilon)\mathcal{A}_1}{1-\mathcal{A}_1} > \alpha > \log \frac{\psi - \rho}{\psi - \rho - 2\delta} \tag{144}$$

□

References

[1] Kotaro Suzumura. *Oligopolistic Competition and Economic Welfare The E ects of Ownership Structures*, pages 93–107. Springer US, Boston, MA, 1996.

[2] Arnold C. Harberger. Monopoly and resource allocation. *The American Economic Review*, 44(2):77–87, 1954.

[3] Yoon-Ho Alex Lee and Donald J. Brown. *Competition, Consumer Welfare, and the Social Cost of Monopoly*, pages 47–68. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008.

[4] Takashi Negishi. Monopolistic competition and general equilibrium. In *General Equilibrium*, pages 194–201. Springer, 1989.

[5] Steve Salop. Information and monopolistic competition. *The American Economic Review*, 66(2):240–245, 1976.

[6] Edward Chamberlain. They theory of monopolistic competition, 1946.

[7] Avinash K Dixit and Joseph E Stiglitz. Monopolistic competition and optimum product diversity. *The American economic review*, 67(3):297–308, 1977.

[8] Asher Wolinsky. True monopolistic competition as a result of imperfect information. *The Quarterly Journal of Economics*, 101(3):493–511, 1986.

[9] Andrea Prat and Tommaso M Valletti. Attention oligopoly. *Available at SSRN 3197930*, 2019.

[10] George K Yarrow. Welfare losses in oligopoly and monopolistic competition. *The Journal of Industrial Economics*, pages 515–529, 1985.

[11] Timothy J Pleskac. Decision and choice: Luce's choice axiom. *International encyclopedia of the social & behavioral sciences*, 5:895–900, 2015.

[12] R Duncan Luce. Individual choice behavior, a theoretical analysis. *Bull. Amer. Math. Soc*, 66(1960):259–260, 1960.

[13] R Duncan Luce. *Individual choice behavior: A theoretical analysis*. Courier Corporation, 2012.

[14] Byron J Morgan. On luce's choice axiom. *Journal of Mathematical Psychology*, 1974.

[15] Christopher A Sims. Rational inattention and monetary economics. In *Handbook of monetary economics*, volume 3, pages 155–181. Elsevier, 2010.

[16] Christopher A Sims. Rational inattention: Beyond the linear-quadratic case. *American Economic Review*, 96(2):158–163, 2006.

[17] Christopher A Sims. Implications of rational inattention. *Journal of monetary Economics*, 50(3):665–690, 2003.

[18] Bartosz Mackowiak and Mirko Wiederholt. Optimal sticky prices under rational inattention. *American Economic Review*, 99(3):769–803, 2009.

[19] Pedro A Ortega and Daniel A Braun. Thermodynamics as a theory of decision-making with information-processing costs. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 469(2153):20120683, 2013.

[20] Pedro Alejandro Ortega Jr. *A unified framework for resource-bounded autonomous agents interacting with unknown environments*. PhD thesis, University of Cambridge, 2011.

[21] Daniel Alexander Ortega and Pedro Alejandro Braun. Information, utility and bounded rationality. In *International Conference on Artificial General Intelligence*, pages 269–274. Springer, 2011.

[22] D. Sculley, Gary Holt, Daniel Golovin, Eugene Davydov, Todd Phillips, Dietmar Ebner, Vinay Chaudhary, Michael Young, Jean-François Crespo, and Dan Dennison. Hidden technical debt in machine learning systems. In C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, and R. Garnett, editors, *Advances in Neural Information Processing Systems 28*, pages 2503–2511. Curran Associates, Inc., 2015.

[23] Michael Luca and Georgios Zervas. Fake it till you make it: Reputation, competition, and yelp review fraud. *Management Science*, 62(12):3412–3427, 2016.

[24] Karen L Xie, Chihchien Chen, and Shinyi Wu. Online consumer review factors affecting offline hotel popularity: evidence from tripadvisor. *Journal of Travel & Tourism Marketing*, 33(2):211–223, 2016.

[25] Raffaele Filieri, Salma Alguezaui, and Fraser McLeay. Why do travelers trust tripadvisor? antecedents of trust towards consumer-generated media and its influence on recommendation adoption and word of mouth. *Tourism Management*, 51:174–185, 2015.

[26] Ray Jiang, Silvia Chiappa, Tor Lattimore, András György, and Pushmeet Kohli. Degenerate feedback loops in recommender systems. *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, Jan 2019.

[27] Léon Bottou, Jonas Peters, Joaquin Quiñonero Candela, Denis X. Charles, D. Max Chickering, Elon Portugaly, Dipankar Ray, Patrice Simard, and Ed Snelson. Counterfactual reasoning and learning systems: The example of computational advertising. *J. Mach. Learn. Res.*, 14(1):3207–3260, January 2013.

[28] Ayan Sinha, David F. Gleich, and Karthik Ramani. Deconvolving feedback loops in recommender systems. *CoRR*, abs/1703.01049, 2017.

[29] Richard A. Berk. An introduction to sample selection bias in sociological data. *American Sociological Review*, 48(3):386–398, 1983.

[30] John Baffoe-Bonnie. Black–white wage differentials in a multiple sample selection bias model. *Atlantic Economic Journal*, 37(1):1–16, September 2008.

[31] Lars Osberg, R.L. Mazany, Richard Apostle, and Don Clairmont. Job mobility, wage determination and market segmentation in the presence of sample selection bias. *The Canadian Journal of Economics / Revue canadienne d'Economique*, 19(2):319–346, 1986.

[32] Lung-fei Lee. Self-selection. *A Companion to Theoretical Econometrics*, pages 383–409, 2001.

[33] Seung-Hoon Yoo, Joo Suk Lee, and Seung-Jun Kwak. Estimation of residential electricity demand function in seoul by correction for sample selection bias. *Energy Policy*, 35(11):5702–5707, 2007.

[34] Joëlle Pianzola. Selection biases in voting advice application research. *Electoral Studies*, 36:272 – 280, 2014.

[35] Peter Kellner. Can online polls produce accurate findings? *International Journal of Market Research*, 46(1):3–22, 2004.

[36] Mark A Kramer. Self-selection bias in reputation systems. In *IFIP International Conference on Trust Management*, pages 255–268. Springer, 2007.

[37] Jeanette B Ruiz and Robert A Bell. Understanding vaccination resistance: vaccine search term selection bias and the valence of retrieved information. *Vaccine*, 32(44):5776–5780, 2014.

[38] Harald Steck. Item popularity and recommendation accuracy. In *Proceedings of the fifth ACM conference on Recommender systems*, pages 125–132, 2011.

[39] Harald Steck. Evaluation of recommendations: rating-prediction and ranking. In *Proceedings of the 7th ACM conference on Recommender systems*, pages 213–220, 2013.

[40] Lydia T. Liu, Horia Mania, and Michael I. Jordan. Competing bandits in matching markets. *CoRR*, abs/1906.05363, 2019.

[41] Andrew Gardner, Christian A Duncan, Jinko Kanno, and Rastko Selmic. 3d hand posture recognition from small unlabeled point sets. In *2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 164–169. IEEE, 2014.

[42] Dheeru Dua and Casey Graff. UCI machine learning repository, 2017.

[43] Han Xiao, Kashif Rasul, and Roland Vollgraf. Fashion-mnist: a novel image dataset for benchmarking machine learning algorithms, 2017.

[44] Guido Van Rossum et al. Python programming language. In *USENIX annual technical conference*, volume 41, page 36, 2007.

[45] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. Pytorch: An imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems*, pages 8024–8035, 2019.

[46] Stéfan van der Walt, S Chris Colbert, and Gael Varoquaux. The numpy array: a structure for efficient numerical computation. *Computing in Science & Engineering*, 13(2):22–30, 2011.

[47] Jerome Friedman, Trevor Hastie, and Robert Tibshirani. *The elements of statistical learning*, volume 1. Springer series in statistics New York, 2001.

[48] Naomi S Altman. An introduction to kernel and nearest-neighbor nonparametric regression. *The American Statistician*, 46(3):175–185, 1992.

[49] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[50] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep learning*. MIT press, 2016.

[51] Thomas Cover and Peter Hart. Nearest neighbor pattern classification. *IEEE transactions on information theory*, 13(1):21–27, 1967.

[52] Giorgio Ausiello, Pierluigi Crescenzi, Giorgio Gambosi, Viggo Kann, Alberto Marchetti-Spaccamela, and Marco Protasi. *Complexity and approximation: Combinatorial optimization problems and their approximability properties*. Springer Science & Business Media, 2012.

[53] Pao-Lu Hsu and Herbert Robbins. Complete convergence and the law of large numbers. *Proceedings of the National Academy of Sciences of the United States of America*, 33(2):25, 1947.

[54] Richard Arratia and Louis Gordon. Tutorial on large deviations for the binomial distribution. *Bulletin of mathematical biology*, 51(1):125–131, 1989.

[55] Jiří Matoušek and Jan Vondrák. The probabilistic method. *Lecture Notes, Department of Applied Mathematics, Charles University, Prague*, 2001.

[56] B Robert. Ash. information theory, 1990.

[57] Willliam Feller. *An introduction to probability theory and its applications*. John Wiley & Sons, 2008.

[58] Frank Spitzer. *Principles of random walk*, volume 34. Springer Science & Business Media, 2013.

[59] Geoffrey Grimmett and Dominic Welsh. *Probability: an introduction*. Oxford University Press, 2014.

[60] Halbert White. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica: journal of the Econometric Society*, pages 817–838, 1980.

[61] James MacQueen et al. Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, volume 1, pages 281–297. Oakland, CA, USA, 1967.

[62] Steven Fortune. Voronoi diagrams and delaunay triangulations. In *Computing in Euclidean geometry*, pages 193–233. World Scientific, 1992.

[63] Franz Aurenhammer and Rolf Klein. Voronoi diagrams. *Handbook of computational geometry*, 5(10):201–290, 2000.

[64] Patrick Billingsley. *Probability and measure*. John Wiley & Sons, 2008.

[65] Olav Kallenberg. *Random measures, theory and applications*. Springer, 2017.

[66] Stuart P. Lloyd. Least squares quantization in pcm. *IEEE Trans. Inf. Theory*, 28:129–136, 1982.

[67] Achim Klenke. *Probability theory: a comprehensive course*. Springer Science & Business Media, 2013.

[68] Murray Rosenblatt. A central limit theorem and a strong mixing condition. *Proceedings of the National Academy of Sciences of the United States of America*, 42(1):43, 1956.