
Collaborative Prediction: Tractable Information Aggregation via Agreement

Anonymous Author(s)

Affiliation

Address

email

Abstract

1 We give efficient “collaboration protocols” through which two parties, who observe
2 different features about the same instances, can interact to arrive at predictions that
3 are more accurate than either could have obtained on their own. The parties only
4 need to iteratively share and update their own label predictions—without either
5 party ever having to share the actual features that they observe. Our protocols are
6 efficient reductions to the problem of learning on each party’s feature space alone,
7 and so can be used even in settings in which each party’s feature space is illegible
8 to the other—which arises in models of human/AI interaction and in multi-modal
9 learning. The communication requirements of our protocols are independent of
10 the dimensionality of the data. In an online adversarial setting we show how to
11 give regret bounds on the predictions that the parties arrive at with respect to a
12 class of benchmark policies defined on the joint feature space of the two parties,
13 despite the fact that neither party has access to this joint feature space. We also give
14 simpler algorithms for the same task in the “batch” setting in which we assume
15 that there is a fixed but unknown data distribution. We generalize our protocols
16 to a decision theoretic setting with high dimensional outcome spaces—the parties
17 in this setting do not need to communicate their (high dimensional) predictions
18 about the outcome, but can instead communicate only “best response actions” with
19 respect to a known utility function and their predicted outcome distribution.

20 Our theorems give a computationally and statistically tractable generalization of
21 past work on information aggregation amongst Bayesians who share a common and
22 correct prior, as part of a literature studying “agreement” in the style of Aumann’s
23 agreement theorem. Our results require no knowledge of (or even the existence of)
24 a prior distribution and are computationally efficient. Nevertheless we show how
25 to lift our theorems back to this classical Bayesian setting, and in doing so, give
26 new information aggregation theorems for Bayesian agreement. In particular we
27 give the first distribution-agnostic information aggregation theorems that do not
28 require making assumptions on the prior distribution, but instead are able to give
29 worst-case accuracy guarantees with respect to restricted classes of functions on
30 the parties’ joint feature spaces.

1 Introduction

31 Imagine that there are multiple parties who hold different kinds of information about the same
32 examples, and would like to collaboratively learn an optimal label predictor on their joint feature
33 space. The straightforward solution would be for them to pool their features and attempt to train
34 an optimal predictor on the data containing the pooled features. But there are settings in which this
35 approach is infeasible. For example, the features legible to one party might not be legible to another.
36 This is the case in e.g. models of human/AI interaction in which one of the parties is a human being
37

and another is a predictive model [Alur et al., 2024, Collina et al., 2025]. In this case the human being is in possession of qualitative features that are difficult to encode for a model (in a medical application, e.g. observations about patient demeanor, mood, smell, etc. that are difficult to formalize) and the model in turn has been trained on enormous amounts of data that cannot be easily used by the human being. In other settings, directly sharing the data might be impossible because of legal or contractual obligations, which is the case in regulated industries like healthcare. This setting motivates the field of “vertically federated learning” [Wei et al., 2022]. It may also be that the data itself is very high dimensional, and communication bandwidth constraints preclude sharing it in its entirety — which motivates the study of the communication complexity of distributed learning [Balcan et al., 2012]. Finally, the technical expertise to train on different kinds of features might be siloed: for example, the data held by different parties might be multi-modal; one party might hold image data (e.g. CT scans) while another might hold text data (e.g. physician notes). Each party might have the tooling necessary to learn on data in their own modality, but none may have the tooling to be easily able to learn on all of the data together.

Aumann’s agreement theorem suggests a tempting general (if stylized) solution to these problems: it states that two perfectly informed Bayesians with a common prior, different observations, and common knowledge of each other’s posteriors must share the same posterior [Aumann, 1976]. Subsequent work has given finite-time convergence protocols through which the different parties engage in a conversation about their beliefs about the outcome [Geanakoplos and Polemarchakis, 1982, Aaronson, 2005] without ever sharing their observations. In particular, Aaronson [2005] showed that for 1-dimensional real-valued outcomes, two Bayesians will reach approximate agreement quickly — in a number of rounds that depends only (polynomially) on the error parameters characterizing “approximate” agreement, independently of the dimensionality or complexity of the prior distribution. Agreement on its own does not in general solve the collaborative learning problem — it has been known since Geanakoplos and Polemarchakis [1982] that *agreement* does not imply *information aggregation*¹. In other words, although two Bayesians engaging in an agreement protocol can only improve the accuracy of their beliefs, they may *agree* on beliefs that are less accurate than those they would have arrived at had they instead shared their information and formed a posterior belief conditional on their joint observations. Nevertheless Kong and Schoenebeck [2023] and Frongillo et al. [2023] have studied conditions (on the prior distribution) under which agreement *does* imply full information aggregation — i.e. optimal learning on the joint feature space. Unfortunately, since it studies perfectly informed Bayesians, this literature makes implausible computational and epistemic assumptions (Why do the two parties share the same, perfect prior knowledge? How do they perform Bayes updates in complex settings?) which makes these approaches seemingly far from algorithmic solutions. Recently, Collina et al. [2025] showed how to recover and generalize quantitative agreement theorems without making any distributional assumptions (i.e. in online adversarial settings) using only computationally and statistically tractable calibration conditions that substantially relax Bayesian rationality. But the work of Collina et al. [2025] says nothing about information aggregation, and so does not provide a solution to the collaborative learning problem.

In this paper we generalize the connection between agreement and information aggregation and give computationally efficient protocols that provably result in information aggregation after only a small number of rounds of communication, that need not make any distributional assumptions at all. In our model, different parties hold different (possibly overlapping) features of the same examples, and may interact over a small number of rounds to share (only) their label predictions, computed from their own features, with each other. Because they hold different information from each other, different parties will likely initially make different predictions about the same example. Nevertheless, during the interaction, they may update their predictions in response to the predictions of their counterparty in the collaboration protocol. We would like them to converge on predictions that are more accurate than any single party could have obtained on their own — and ideally predictions that are *optimal* with respect to some benchmark class of predictors that are defined on their joint feature space, despite the fact that every participant in the protocol only has access to their own feature space. Because the parties only need to communicate their label predictions (to bounded precision) at each round, the communication complexity of our protocol is independent of the dimensionality of the data.

¹Consider a joint distribution on bits x_A, x_B , and y that are all marginally uniform, but such that $y = x_A + x_B \bmod 2$. If Alice is in possession of x_A and Bob is in possession of x_B they will agree that $\mathbb{P}[y = 1] = \frac{1}{2}$, even though they would know y with certainty if they shared their data with each other.

We give two variants of our protocol: one for prediction in the online adversarial setting, in which we make no distributional assumptions at all — and a simpler protocol for the batch setting, in which data is assumed to be drawn i.i.d. from a fixed but unknown and arbitrary distribution. In both cases we guarantee that the collaboration protocol is accuracy-improving, and give conditions under which the predictions are provably *optimal* with respect to a benchmark class of models defined on the pooled features. These conditions are frequentist “weak-learning” assumptions that substantially generalize the “information substitutes” condition on prior distributions used by Frongillo et al. [2023] in Bayesian settings. Moreover our protocols are computationally efficient to run in the sense that they are computationally efficient reductions from the problem of multi-party learning to the problem of single-party learning, and therefore efficient in the worst case whenever the single-party learning problem can be efficiently solved. Each party only needs to run their own learning algorithm, tailored to data from their own modality, on their own data a bounded number of times in order to engage in the protocol.

Finally, we show that all of our results “lift” back to the Bayesian setting of Aumann [1976], Aaronson [2005], Frongillo et al. [2023], resulting in new theorems about agreement and information aggregation in the classical setting in which examples are assumed to be drawn from a fixed and known prior, and a conversation between two perfect Bayesians occurs for a single draw from this prior. Among other things, we show that Bayesian agreement implies accuracy at least that of the best linear function on the joint feature space of the two parties, independently of any assumptions on the prior distribution — the first such distribution-independent information aggregation theorem we are aware of in the agreement literature.

1.1 Our Model and Results

We begin by describing our results in the online-adversarial setting, when our goal is to solve a one-dimensional regression problem; we then describe extensions to more complex prediction tasks and to simpler algorithms in the batch setting. Finally we describe how our results “lift” to one-shot interactions if both parties are perfect Bayesians and share a common and correct prior — the traditional setting for Aumann’s agreement theorem [Aumann, 1976].

There is a feature space $\mathcal{X} = \mathcal{X}_A \times \mathcal{X}_B$ partitioned into parts \mathcal{X}_A and \mathcal{X}_B which may each be arbitrary, as well as a label space \mathcal{Y} that initially we take to be $\mathcal{Y} = [0, 1]$. At each day t , an arbitrary adaptive adversarial process chooses an example $x^t = (x_A^t, x_B^t) \in \mathcal{X}$ and a label $y^t \in \mathcal{Y}$. Party 1 (Alice) receives x_A^t and Party 2 (Bob) receives x_B^t . Each day, Alice and Bob then engage in a “collaboration protocol” which is an interaction that takes place across K rounds. In odd rounds k , Alice produces a prediction $\hat{y}^{t,k} \in \mathcal{Y}$ that may be a function of x_A^t as well as all previously observed history (including Bob’s predictions at previous rounds on the same day). Similarly, in even rounds k , Bob produces a prediction $\hat{y}^{t,k} \in \mathcal{Y}$ that may be a function of x_B^t and all previously observed history. Crucially Alice and Bob never share their feature vectors with one another—only label predictions. At the final round K each day, they fix their final prediction $\hat{y}^t = \hat{y}^{t,K}$, at which point both Alice and Bob learn the true label y^t , and time proceeds to the next day.

Our goal in interaction is to arrive at a set of predictions $\hat{y}^1, \dots, \hat{y}^T$ that have squared error that is as low as the *best predictor in hindsight* in some class of models \mathcal{H}_J defined on the *joint* feature space of Alice and Bob: i.e. each function $h \in \mathcal{H}_J$ has the form $h : \mathcal{X} \rightarrow \mathbb{R}$ and produces a prediction $h(x_A, x_B)$ that is a function of the features available to both parties. For example, we might take \mathcal{H}_J to be the set of all norm-bounded linear functions on the joint feature space. In other words, we want to be able to guarantee, against an arbitrary adversarial sequence:

Definition 1.1 (Predictions have No (External) Regret to \mathcal{H}_J). *The final predictions $\hat{y}^1, \dots, \hat{y}^T$ have no (external) regret to \mathcal{H}_J if for every $h \in \mathcal{H}_J$:*

$$\sum_{t=1}^T (\hat{y}^t - y^t)^2 \leq \sum_{t=1}^T (h(x^t) - y^t)^2$$

The difficulty is that there may be no predictor defined over \mathcal{X}_A or \mathcal{X}_B individually that can obtain this — collaborating to use information from *both* parties is essential. What each party can try to do instead is to make predictions $\hat{y}^{t,k}$ during their own rounds of conversation that have no regret with respect to classes of functions \mathcal{H}_A and \mathcal{H}_B that are respectively defined only on their own feature spaces — i.e. each function $h_A \in \mathcal{H}_A$ is defined as $h_A : \mathcal{X}_A \rightarrow \mathbb{R}$ and each function $h_B \in \mathcal{H}_B$ is

defined as $h_B : \mathcal{X}_B \rightarrow \mathbb{R}$. For example, \mathcal{H}_A and \mathcal{H}_B might be the set of unit-norm linear functions defined only over \mathcal{X}_A and \mathcal{X}_B respectively. We take as an intermediate goal to produce a *single* sequence of predictions at some round $k \in \{1, \dots, K\}$ that has no *swap*-regret with respect to \mathcal{H}_A and \mathcal{H}_B simultaneously.

Definition 1.2 (Swap Regret (Informal Version of Definition A.5)). *A sequence of predictions $\hat{y}^{1,k}, \dots, \hat{y}^{T,k}$ has no swap regret with respect to a class \mathcal{H} if for every value $v \in \{\hat{y}^{1,k}, \dots, \hat{y}^{T,k}\}$ and for every $h \in \mathcal{H}$:*

$$\sum_{t=1}^T \mathbb{1}[\hat{y}^{t,k} = v](\hat{y}^{t,k} - y^t)^2 \leq \min_{h \in \mathcal{H}} \left(\sum_{t=1}^T \mathbb{1}[\hat{y}^{t,k} = v](h(x^t) - y^t)^2 \right)$$

Swap regret is a stronger condition than external regret — it requires that our predictions \hat{y}_k^t be as accurate as the best model $h \in \mathcal{H}$ not just marginally, but conditionally on the value of our own predictions. As an intermediate step towards our goal of obtaining predictions with no regret to \mathcal{H}_J , we will hope to produce a single sequence of predictions that has no swap regret with respect to classes \mathcal{H}_A and \mathcal{H}_B which are individually weaker than \mathcal{H}_J (as they are defined only on \mathcal{X}_A and \mathcal{X}_B respectively). To relate this condition to our ultimate goal, we define a weak learning condition (related to the weak learning condition used to characterize the relationship between multicalibration and boosting by Globus-Harris et al. [2023]) that relates \mathcal{H}_A , \mathcal{H}_B and \mathcal{H}_J .

Definition 1.3 (Weak Learning for Regression (Informal Version of Definition B.1)). *We say that \mathcal{H}_A and \mathcal{H}_B jointly satisfy the weak learning condition with respect to \mathcal{H}_J if for any joint distribution \mathcal{D} over $\mathcal{X} \times \mathcal{Y}$ we have that if:*

$$\min_{h_J \in \mathcal{H}_J} \mathbb{E}_{\mathcal{D}}[(h_J(x) - y)^2] < \mathbb{E}_{\mathcal{D}}[(\mu(\mathcal{D}) - y)^2]$$

Then there either exists $h_A \in \mathcal{H}_A$ such that:

$$\mathbb{E}_{\mathcal{D}}[(h_A(x_A) - y)^2] < \mathbb{E}_{\mathcal{D}}[(\mu(\mathcal{D}) - y)^2]$$

or there exists $h_B \in \mathcal{H}_B$ such that:

$$\mathbb{E}_{\mathcal{D}}[(h_B(x_B) - y)^2] < \mathbb{E}_{\mathcal{D}}[(\mu(\mathcal{D}) - y)^2]$$

where $\mu(\mathcal{D}) = \mathbb{E}_{\mathcal{D}}[y]$ is the label mean over the distribution.

In other words, the weak learning condition requires that if there is any model in the joint class \mathcal{H}_J that obtains predictions with lower error than a constant predictor, there must be some model in either Alice’s class \mathcal{H}_A or Bob’s class \mathcal{H}_B that also obtains lower error than a constant predictor. This is a weak learning condition because the model in the joint class might obtain *much* better error than a trivial constant predictor, but we only require that \mathcal{H}_A or \mathcal{H}_B obtain *slightly* better than trivial error.

In Section B we prove a “boosting” theorem: if \mathcal{H}_A and \mathcal{H}_B satisfy the weak learning condition with respect to \mathcal{H}_J , and a sequence of predictions $\{\hat{y}^{1,k}, \dots, \hat{y}^{T,k}\}$ has no swap regret with respect to \mathcal{H}_A and \mathcal{H}_B , then it must in fact have no external regret with respect to the joint class \mathcal{H}_J . In fact, we show that a slightly weaker condition than swap regret suffices. It is enough that the sequence of predictions has low *distance* to no-swap regret with respect to \mathcal{H}_A and \mathcal{H}_B — i.e. that it is possible to perturb the sequence of predictions by a small amount in the L_1 norm such that the perturbed predictions have no swap regret. This is related to recently studied notions of “distance to calibration” [Błasiok et al., 2023, Qiao and Zheng, 2024, Arunachaleswaran et al., 2025], and will be easier for us to satisfy. It then follows that any other sequence of predictions $\{\hat{y}^1, \dots, \hat{y}^T\}$ that has lower squared error than predictions $\{\hat{y}^{1,k}, \dots, \hat{y}^{T,k}\}$ must in turn have no (external) regret with respect to \mathcal{H}_J .

Given classes \mathcal{H}_A and \mathcal{H}_B that satisfy our weak learning condition with respect to a class of models \mathcal{H}_J defined on the joint feature space \mathcal{X} , our problem is thus reduced to giving a collaboration protocol that quickly converges to a sequence of predictions that simultaneously has no swap regret to both \mathcal{H}_A and \mathcal{H}_B . Towards this end, we ask that both Alice and Bob satisfy a condition that we call “conversation swap regret” relative to \mathcal{H}_A and \mathcal{H}_B respectively.

Definition 1.4 (Conversation Swap Regret (Definition A.7)). *We say that Bob’s predictions have no conversation swap regret with respect to \mathcal{H}_B if for every even round of conversation k and for every pair of values $v \in \{\hat{y}^{1,k}, \dots, \hat{y}^{T,k}\}$ and $v' \in \{\hat{y}^{1,k-1}, \dots, \hat{y}^{T,k-1}\}$:*

$$\sum_{t=1}^T \mathbb{1}[\hat{y}^{t,k-1} = v'] \mathbb{1}[\hat{y}^{t,k} = v](\hat{y}^{t,k} - y^t)^2 \leq \min_{h \in \mathcal{H}} \left(\sum_{t=1}^T \mathbb{1}[\hat{y}^{t,k-1} = v'] \mathbb{1}[\hat{y}^{t,k} = v](h(x^t) - y^t)^2 \right)$$

173 *If Alice satisfies a symmetric condition on odd rounds k with respect to \mathcal{H}_A we say that Alice has no*
 174 *conversation swap regret with respect to \mathcal{H}_A .*

175 In other words, conversation swap regret requires that Alice and Bob satisfy the no swap regret
 176 condition (with respect to their respective model classes \mathcal{H}_A and \mathcal{H}_B) not just marginally over the
 177 whole sequence, but on each subsequence defined by the other party’s prediction *at the previous*
 178 *round of interaction*. Whenever \mathcal{H}_A and \mathcal{H}_B contain all constant functions with range in $[0, 1]$, having
 179 no conversation swap regret implies satisfying the “conversation calibration” condition defined in
 180 Collina et al. [2025].

In Section C, we show that when both Alice and Bob make predictions with no conversation swap regret with respect to \mathcal{H}_A and \mathcal{H}_B respectively, then if the collaboration protocol runs for sufficiently many rounds K , there must exist some round $k \leq K$ at which the sequence of predictions $\{\hat{y}^{1,k}, \dots, \hat{y}^{T,k}\}$ has low distance to swap regret with respect to *both* \mathcal{H}_A and \mathcal{H}_B simultaneously. Although this round k may not be the final round K , we also show that the final set of predictions has only lower squared error than the predictions made at any previous round k :

$$\sum_{t=1}^T (\hat{y}^{t,K} - y^t)^2 \leq \sum_{t=1}^T (\hat{y}^{t,k} - y^t)^2$$

181 Thus, by applying our “boosting” theorem from Section B, we can conclude that if \mathcal{H}_A and \mathcal{H}_B satisfy
 182 the weak learning condition with respect to a joint class \mathcal{H}_J , then the final sequence of predictions
 183 $\{\hat{y}^1, \dots, \hat{y}^T\}$ has no (external) regret with respect to the joint class \mathcal{H}_J .

184 It remains to ask: for which classes \mathcal{H}_A and \mathcal{H}_B do there exist efficient algorithms for satisfying
 185 the no-conversation-swap regret condition, and are there examples of classes $(\mathcal{H}_A, \mathcal{H}_B, \mathcal{H}_J)$ that
 186 satisfy the weak learning condition? In Sections B.1 and C.1 we provide answers to these questions.
 187 Garg et al. [2024] gave an efficient reduction (in the style of Blum and Mansour [2007]) from the
 188 problem of obtaining no swap regret with respect to an arbitrary class of functions \mathcal{H} to the problem
 189 of obtaining no external regret with respect to \mathcal{H} . We in turn give an efficient reduction from the
 190 problem of obtaining no conversation swap regret with respect to an arbitrary class of functions \mathcal{H}
 191 to the problem of obtaining no swap regret with respect to \mathcal{H} . In combination, these results mean
 192 that there are computationally efficient algorithms for engaging in our collaboration protocol for
 193 any class of models $\mathcal{H}_A, \mathcal{H}_B$ that admit standard efficient online learning algorithms with regret
 194 guarantees — and “oracle efficient” algorithms for any class of models for which there are online
 195 learning algorithms with good (external) regret bounds, even if they are not computationally efficient
 196 in the worst case. Because there exist computationally efficient algorithms for online adversarial
 197 norm-bounded linear regression Azoury and Warmuth [2001], Vovk [2001] and related problems
 198 (e.g. squared error regression over reproducing kernel Hilbert spaces Vovk [2006]), this immediately
 199 implies efficient algorithms for obtaining no conversation swap regret with respect to classes $\mathcal{H}_A, \mathcal{H}_B$
 200 representing norm-bounded linear functions over \mathcal{X}_A and \mathcal{X}_B respectively. Moreover, we show
 201 in Section B.1 that norm-bounded linear functions over \mathcal{X}_A and \mathcal{X}_B respectively satisfy the weak
 202 learning condition with respect to norm-bounded linear functions on the joint feature space \mathcal{X} . In
 203 fact we show a more general theorem for any class of functions \mathcal{H}_J that can be represented as the
 204 Minkowski sum of classes \mathcal{H}_A and \mathcal{H}_B that are themselves bounded and star-shaped. Moreover we
 205 show (also in Section B.1) that the weak learning condition we prove is quantitatively tight for linear
 206 functions.

207 All together, this means that we have a computationally and statistically efficient collaboration
 208 protocol for learning predictors that are as accurate as the best linear function on the joint feature
 209 space (and more general classes of functions).

210 **Theorem 1.5** (Informal statement of Theorem C.6). *Fix any triple of hypothesis classes $\mathcal{H}_A, \mathcal{H}_B$,*
 211 *and \mathcal{H}_J . Suppose \mathcal{H}_A and \mathcal{H}_B consist of functions with bounded range and admit efficient online*
 212 *algorithms guaranteeing no external regret with respect to \mathcal{H}_A and \mathcal{H}_B respectively. If \mathcal{H}_A and \mathcal{H}_B*
 213 *satisfy the weak learning condition with respect to \mathcal{H}_J , and the conversation length K is sublinear in*
 214 *T (but not constant), then there is an efficient collaboration protocol such that:*

$$\sum_{t=1}^T (\hat{y}^t - y^t)^2 - \min_{h_J \in \mathcal{H}_J} \sum_{t=1}^T (h_J(x^t) - y^t)^2 \leq o(T)$$

215 *In particular, this is true for the classes of norm-bounded linear functions.*

216 1.1.1 Tightness of Our Approach

217 In Appendix D we give several lower bounds intended to illustrate the tightness of various aspects of
218 our approach, answering several questions:

219 **Is interaction necessary?** Perhaps for sufficiently simple classes of functions (e.g. linear functions)
220 that satisfy our weak learning condition, no interaction is necessary — maybe the optimal linear
221 predictor on \mathcal{X}_A and \mathcal{X}_B already contains enough information to compete with the best linear predictor
222 on the full feature space. We show that this is not the case, by exhibiting a lower bound instance
223 (Theorem D.1) such that the Bayes optimal predictors $h^*(x_A)$, $h^*(x_B)$, and $h^*(x)$ defined on \mathcal{X}_A ,
224 \mathcal{X}_B , and \mathcal{X} are all linear, and yet no function of $h^*(x_A)$ and $h^*(x_B)$ has squared error competitive
225 with $h^*(x)$.

226 **Is our weak learning condition necessary?** Can we relax our weak learning condition? We show
227 that the answer is no, at least for any similar approach. Our boosting theorem demonstrates that the
228 weak learning condition is sufficient for no swap regret with respect to \mathcal{H}_A and \mathcal{H}_B to imply no
229 external regret with respect to \mathcal{H}_J . We give a lower bound instance (Theorem D.2) showing that it is
230 also necessary: for any triple of function classes $\mathcal{H}_A, \mathcal{H}_B, \mathcal{H}_J$ that fail to satisfy the weak learning
231 condition, there is a distribution and a sequence of predictions such that the predictions have no swap
232 regret with respect to \mathcal{H}_A and \mathcal{H}_B and yet have positive external regret with respect to \mathcal{H}_J . We also
233 show that our weak learning condition is strictly weaker than the “information substitutes” condition
234 studied in Frongillo et al. [2023], and that indeed linear functions do *not* satisfy the information
235 substitutes condition on all distributions (Theorem D.4 and Theorem D.5).

236 **Is swap regret necessary?** Our collaboration protocol is designed to converge to a single sequence
237 of predictions that has low (distance to) swap regret with respect to \mathcal{H}_A and \mathcal{H}_B simultaneously —
238 despite the fact that our ultimate goal is simply to have no *external* regret with respect to \mathcal{H}_J . Might it
239 instead suffice to converge to a single sequence of predictions that has no *external* regret with respect
240 to \mathcal{H}_A and \mathcal{H}_B ? No. We give a lower bound instance (Theorem D.6) exhibiting that even for linear
241 functions \mathcal{H}_A and \mathcal{H}_B (which satisfy the weak learning condition relative to linear functions on the
242 joint feature space \mathcal{H}_J), predictions that have no external regret with respect to \mathcal{H}_A and \mathcal{H}_B can still
243 have positive external regret with respect to \mathcal{H}_J .

244 1.1.2 Lifting to the One-Shot Bayesian Setting

245 In Section E we show that the theorems we prove in the online adversarial section can be “lifted”
246 to the one-shot Bayesian setting in which agreement theorems have been traditionally studied
247 [Aumann, 1976, Geanakoplos and Polemarchakis, 1982, Aaronson, 2005, Kong and Schoenebeck,
248 2023, Frongillo et al., 2023]. This is, informally, because Bayesians with correct priors have beliefs
249 that are unbiased conditional on *any* event, and in particular their predictions are guaranteed in
250 expectation to have no conversation swap regret with respect to any fixed collection of benchmark
251 functions. For any class of benchmark functions for which empirical squared error converges
252 uniformly to expected squared error (e.g. any class of functions with bounded fat shattering dimension)
253 this means that they are guaranteed to satisfy the conditions of our boosting theorems on any
254 sufficiently long sequence of instances drawn from a known prior. Thus we can imagine that Bayesian
255 Alice and Bayesian Bob engage in an interaction for an arbitrarily long sequence of examples drawn
256 i.i.d. from their commonly shared prior, and apply our theorems to bound the accuracy of the
257 predictions that result along this imagined sequence. But when examples are drawn i.i.d. from a fixed
258 prior the final predictions at each day in this imagined sequence will also be i.i.d. Thus our theorems,
259 which generically apply to the average error of predictions over a sequence, actually in this case
260 apply to the expected squared error of the predictions that result from the collaboration protocol on
261 the *first* day of the sequence, and hence apply in the one-shot setting. The result is new information
262 aggregation theorems in the classical Bayesian setting.

263 1.1.3 Other Extensions

264 We give two additional extensions of our results:

265 **A Decision Theoretic Extension for Higher Dimensional Outcome Spaces.** In Appendix H we
266 give a decision theoretic extension of the online setting to high dimensional outcome spaces. Now

the outcome space $\mathcal{Y} \subseteq [0, 1]^d$ is d dimensional, and we model a decision maker with a finite action space \mathcal{A} and a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$ that maps an action and an outcome to a utility. The natural extension of our one-dimensional solution to a d -dimensional outcome space—by asking for swap regret with respect to outcome predictions themselves—would inherit exponential dependencies on d . We circumvent this difficulty by not communicating predictions of the outcome itself, but instead *actions* that are utility maximizing given agents’ predictions. Here, we appeal to *decision swap regret*, a coarser notion of swap regret given by Lu et al. [2025], to satisfy the no conversation swap regret condition; this allows us to invoke a fast agreement theorem from Collina et al. [2025]. Together with a decision version of the boosting theorem of Section B, we show that the sequence of *actions* that result from the collaboration protocol have no regret with respect to a collection of action policies defined on the joint feature space. Our regret bounds and our communication requirements are independent of the dimensionality of the data, depending instead only polynomially on the dimension of the outcome space and the cardinality of the action space.

Simpler Algorithms in the Batch Setting. In the bulk of this paper, we study the collaborative learning problem in the difficult online adversarial setting, in which examples are assumed to arrive adversarially. Of course the problem is still interesting in the more standard *batch* setting, in which examples (x, y) are assumed to be drawn i.i.d. from a fixed but unknown distribution. In Appendix I we give a simpler algorithm for this setting, which can be viewed as a two-party generalization of the “level-set boosting” algorithm given in Globus-Harris et al. [2023]. This algorithm is a reduction to the problem of squared error regression over the classes \mathcal{H}_A and \mathcal{H}_B respectively; we prove fast convergence and out-of-sample generalization theorems for it. Our algorithm in this setting uses *test-time* compute to make predictions on new instances: the two parties engage in a polynomial-length interaction exchanging and updating predictions about each test-time instance before agreeing on a final prediction.

1.2 Related Work

Agreement. Aumann’s classic “agreement theorem” [Aumann, 1976] states that two Bayesians with a common and correct prior, who have *common knowledge* of each other’s posterior expectation of any predicate must have the same posterior expectation of that predicate. “Common Knowledge” is the limit of an infinite exchange of information, but Geanakoplos and Polemarchakis [Geanakoplos and Polemarchakis, 1982] showed that whenever the underlying state space is finite, then agreement occurs after a finite number rounds in which the information exchanged in each round is the posterior expectation of each party. Aaronson [Aaronson, 2005] showed that for 1-dimensional expectations, ϵ -approximate agreement can be obtained (with probability $1 - \delta$ over the draw from the prior distribution) after the parties exchange only $O(1/\epsilon^2\delta)$ messages. Two papers [Kong and Schoenebeck, 2023, Frongillo et al., 2023] study conditions under which Aumannian agreement implies information aggregation — i.e. when “agreement” is reached at the same posterior belief that would have resulted had the two parties shared all of their information, rather than interacting within an agreement protocol. These papers all assume perfect Bayes updates based on a correctly specified and commonly known prior distribution, and so in general do not correspond to computationally tractable algorithms. Collina et al. [2025] generalizes Aaronson [2005] and proves agreement theorems without making any distributional assumptions (i.e. in an online adversarial setting as in this paper), and using tractable calibration conditions that relax Bayesian rationality — but says nothing about information aggregation. Our paper extends the work of Collina et al. [2025] to be able to give information-aggregation like statements in an online adversarial setting — in particular, regret bounds with respect to a class of models defined on the joint feature space across the two parties. When applied to the Bayes optimal predictors, our “weak learning” condition is strictly weaker than the “information substitutes” condition given by Frongillo et al. [2023], and our weak learning condition can be applied to any other class of models (not necessarily Bayes optimal). Our results can be lifted back to the Bayesian setting of [Aumann, 1976, Aaronson, 2005, Frongillo et al., 2023] to give new information aggregation theorems.

Calibration, Ensembling, and Boosting. Beyond Collina et al. [2025] which replaces the assumption of Bayesian rationality with tractable calibration conditions in the context of Aumann’s agreement theorem, several papers [Camara et al., 2020, Collina et al., 2024] have replaced traditional assumptions of Bayesian rationality (and common prior assumptions) with calibration assumptions

in *principal agent* problems arising e.g. in contract theory and Bayesian Persuasion. In particular, Collina et al. [2024] shows how to do this with tractable decision calibration conditions.

Our weak learning condition is a generalization of the weak learning condition given by Globus-Harris et al. [2023], which they showed characterizes when *multicalibration* [Hébert-Johnson et al., 2018] with respect to one class of functions implies error optimality with respect to another. An important step in our analysis is that agents with “conversation swap regret” converge quickly to predictions that agree on most days, which we obtain by showing that conversation swap regret implies conversation calibration as defined in Collina et al. [2025], which in turn implies fast agreement. The fact that swap regret with respect to squared loss implies low calibration error is a classical result originally due to Foster and Vohra [1999]. In the “action setting” in Appendix H, the conditions we require on each party are that they be decision calibrated and decision “cross-calibrated” with respect to a benchmark class of functions — conditions that were recently used in Lu et al. [2025]. These conditions are variants of decision calibration as studied by Zhao et al. [2021], Noarov et al. [2023] and “decision outcome indistinguishability” as studied by Gopalan et al. [2023]. We use the algorithm of Noarov et al. [2023] to constructively enforce these conditions. “Cross calibration” conditions have also been used to ensemble models in accuracy improving ways [Roth et al., 2023, Alur et al., 2024] — but with the exception of Globus-Harris et al. [2023] (which gives results in a single-party setting) these methods do not promise to compete with a benchmark class of models that is *strictly* more accurate than the initial models.

Other Related Work. The setting we study, in which different parties hold different features about the same example and want to coordinate on a single learning task resembles *co-training* as studied by Blum and Mitchell [1998], Balcan et al. [2004]. Models of co-training generally assume that the features each party hold are sufficient to learn a perfect model, but that labels are scarce: co-training protocols seek to use agreement with the other party as a regularization technique that allows them to learn with only small amounts of labeled data (together with larger amounts of unlabeled data). In contrast, our interest is in the setting in which each individual’s features are not sufficient to learn an accurate model, and the goal is to collaboratively learn a model that is more accurate than could be learned by any party alone, even with arbitrarily many samples. Blum et al. [2017] define *collaborative learning*, later studied by [Haghtalab et al., 2022, Donahue and Kleinberg, 2021, Blum et al., 2021, Zhang et al., 2024, Peng et al., 2024, Haghtalab et al., 2023]. In the collaborative learning setting, multiple parties have data from different distributions that are all labeled with the same function, and are interested in collaborating to learn their shared label function with fewer samples than it would take for each party to learn the function only from their own data. In contrast, in our setting, there is a single distribution (or no distribution, in the online adversarial setting), and it is the features that are distributed amongst parties. We defer a discussion on additional related work to Appendix G.

References

- Scott Aaronson. The complexity of agreement. In *Proceedings of the thirty-seventh annual ACM symposium on Theory of computing*, pages 634–643, 2005.
- Rohan Alur, Manish Raghavan, and Devavrat Shah. Human expertise in algorithmic prediction. *Advances in Neural Information Processing Systems*, 37:138088–138129, 2024.
- M. Anthony and P.L. Bartlett. *Neural Network Learning: Theoretical Foundations*. Cambridge University Press, 1999. ISBN 9780521573535. URL <https://books.google.com/books?id=UH6XR0EQ4h8C>.
- Eshwar Ram Arunachaleswaran, Natalie Collina, Aaron Roth, and Mirah Shi. An elementary predictor obtaining $2\sqrt{T} + 1$ distance to calibration. In *Proceedings of the 2025 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1366–1370. SIAM, 2025.
- Robert J. Aumann. Agreeing to Disagree. *The Annals of Statistics*, 4(6):1236 – 1239, 1976. doi: 10.1214/aos/1176343654. URL <https://doi.org/10.1214/aos/1176343654>.
- Katy S. Azoury and M. K. Warmuth. Relative loss bounds for on-line density estimation with the exponential family of distributions. In *Proceedings of the Fifteenth Conference on Uncertainty in*

372 *Artificial Intelligence*, UAI'99, page 31–40, San Francisco, CA, USA, 1999. Morgan Kaufmann
373 Publishers Inc. ISBN 1558606149.

374 Katy S Azoury and Manfred K Warmuth. Relative loss bounds for on-line density estimation with
375 the exponential family of distributions. *Machine learning*, 43:211–246, 2001.

376 Maria-Florina Balcan, Avrim Blum, and Ke Yang. Co-training and expansion: Towards bridging
377 theory and practice. *Advances in neural information processing systems*, 17, 2004.

378 Maria Florina Balcan, Avrim Blum, Shai Fine, and Yishay Mansour. Distributed learning, communi-
379 cation complexity and privacy. In *Conference on Learning Theory*, pages 26–1. JMLR Workshop
380 and Conference Proceedings, 2012.

381 Tadas Baltrušaitis, Chaitanya Ahuja, and Louis-Philippe Morency. Multimodal machine learning:
382 A survey and taxonomy. *IEEE transactions on pattern analysis and machine intelligence*, 41(2):
383 423–443, 2018.

384 Gagan Bansal, Tongshuang Wu, Joyce Zhou, Raymond Fok, Besmira Nushi, Ece Kamar, Marco Tulio
385 Ribeiro, and Daniel Weld. Does the whole exceed its parts? the effect of ai explanations on
386 complementary team performance. In *Proceedings of the 2021 CHI conference on human factors
387 in computing systems*, pages 1–16, 2021.

388 Jarosław Błasiok, Parikshit Gopalan, Lunjia Hu, and Preetum Nakkiran. A unifying theory of distance
389 from calibration. In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*,
390 pages 1727–1740, 2023.

391 Avrim Blum and Yishay Mansour. From external to internal regret. *Journal of Machine Learning
392 Research*, 8(6), 2007.

393 Avrim Blum and Tom Mitchell. Combining labeled and unlabeled data with co-training. In *Pro-
394 ceedings of the eleventh annual conference on Computational learning theory*, pages 92–100,
395 1998.

396 Avrim Blum, Nika Haghtalab, Ariel D Procaccia, and Mingda Qiao. Collaborative pac learning.
397 *Advances in Neural Information Processing Systems*, 30, 2017.

398 Avrim Blum, Nika Haghtalab, Richard Lanus Phillips, and Han Shao. One for one, or all for all:
399 Equilibria and optimality of collaboration in federated learning. In *International Conference on
400 Machine Learning*, pages 1005–1014. PMLR, 2021.

401 Modibo K Camara, Jason D Hartline, and Aleck Johnsen. Mechanisms for a no-regret agent: Beyond
402 the common prior. In *2020 IEEE 61st annual symposium on foundations of computer science (focs)*,
403 pages 259–270. IEEE, 2020.

404 Kewei Cheng, Tao Fan, Yilun Jin, Yang Liu, Tianjian Chen, Dimitrios Papadopoulos, and Qiang
405 Yang. Secureboost: A lossless federated learning framework. *IEEE intelligent systems*, 36(6):
406 87–98, 2021.

407 Natalie Collina, Aaron Roth, and Han Shao. Efficient prior-free mechanisms for no-regret agents.
408 *ACM Conference on Economics and Computation (EC)*, 2024.

409 Natalie Collina, Surbhi Goel, Varun Gupta, and Aaron Roth. Tractable agreement protocols. In
410 *Proceedings of the 57th Annual ACM Symposium on Theory of Computing (STOC)*, 2025.

411 Kate Donahue and Jon Kleinberg. Model-sharing games: Analyzing federated learning under volun-
412 tary participation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35,
413 pages 5303–5311, 2021.

414 Kate Donahue, Alexandra Chouldechova, and Krishnaram Kenthapadi. Human-algorithm collab-
415 oration: Achieving complementarity and avoiding unfairness. In *Proceedings of the 2022 ACM
416 Conference on Fairness, Accountability, and Transparency*, pages 1639–1656, 2022.

417 Dean P. Foster and Rakesh Vohra. Regret in the on-line decision problem. *Games and Economic
418 Behavior*, 29(1):7–35, 1999. ISSN 0899-8256. doi: <https://doi.org/10.1006/game.1999.0740>. URL
419 <https://www.sciencedirect.com/science/article/pii/S089982569907406>.

420 Rafael Frongillo, Eric Neyman, and Bo Waggoner. Agreement implies accuracy for substitutable
421 signals. In *Proceedings of the 24th ACM Conference on Economics and Computation*, pages
422 702–733, 2023.

423 Sumegha Garg, Christopher Jung, Omer Reingold, and Aaron Roth. Oracle efficient online multi-
424 calibration and omniprediction. In *Proceedings of the 2024 Annual ACM-SIAM Symposium on*
425 *Discrete Algorithms (SODA)*, pages 2725–2792. SIAM, 2024.

426 John D Geanakoplos and Heraklis M Polemarchakis. We can’t disagree forever. *Journal of Economic*
427 *theory*, 28(1):192–200, 1982.

428 Ira Globus-Harris, Declan Harrison, Michael Kearns, Aaron Roth, and Jessica Sorrell. Multicalibra-
429 tion as boosting for regression. In *Proceedings of the 40th International Conference on Machine*
430 *Learning*, ICML’23. JMLR.org, 2023.

431 Surbhi Goel, Steve Hanneke, Shay Moran, and Abhishek Shetty. Adversarial resilience in sequential
432 prediction via abstention. *Advances in Neural Information Processing Systems*, 36:8027–8047,
433 2023.

434 Ethan Goh, Robert Gallo, Jason Hom, Eric Strong, Yingjie Weng, Hannah Kerman, Joséphine A Cool,
435 Zahir Kanjee, Andrew S Parsons, Neera Ahuja, et al. Large language model influence on diagnostic
436 reasoning: a randomized clinical trial. *JAMA Network Open*, 7(10):e2440969–e2440969, 2024.

437 Catalina Gomez, Sue Min Cho, Shichang Ke, Chien-Ming Huang, and Mathias Unberath. Human-ai
438 collaboration is not very collaborative yet: A taxonomy of interaction patterns in ai-assisted
439 decision making from a systematic review. *Frontiers in Computer Science*, 6:1521066, 2025.

440 Parikshit Gopalan, Lunjia Hu, Michael P Kim, Omer Reingold, and Udi Wieder. Loss minimization
441 through the lens of outcome indistinguishability. In *14th Innovations in Theoretical Computer*
442 *Science Conference (ITCS 2023)*, 2023.

443 Ben Green and Yiling Chen. The principles and limits of algorithm-in-the-loop decision making.
444 *Proceedings of the ACM on human-computer interaction*, 3(CSCW):1–24, 2019.

445 Nika Haghtalab, Michael Jordan, and Eric Zhao. On-demand sampling: Learning optimally from
446 multiple distributions. *Advances in Neural Information Processing Systems*, 35:406–419, 2022.

447 Nika Haghtalab, Michael Jordan, and Eric Zhao. A unifying perspective on multi-calibration: Game
448 dynamics for multi-objective learning. *Advances in Neural Information Processing Systems*, 36:
449 72464–72506, 2023.

450 Nika Haghtalab, Tim Roughgarden, and Abhishek Shetty. Smoothed analysis with adaptive adver-
451 saries. *Journal of the ACM*, 71(3):1–34, 2024.

452 Stephen Hardy, Wilko Henecka, Hamish Ivey-Law, Richard Nock, Giorgio Patrini, Guillaume Smith,
453 and Brian Thorne. Private federated learning on vertically partitioned data via entity resolution
454 and additively homomorphic encryption. *arXiv preprint arXiv:1711.10677*, 2017.

455 Ursula Hébert-Johnson, Michael Kim, Omer Reingold, and Guy Rothblum. Multicalibration: Cali-
456 bration for the (computationally-identifiable) masses. In *International Conference on Machine*
457 *Learning*, pages 1939–1948. PMLR, 2018.

458 Yuqing Kong and Grant Schoenebeck. False consensus, information theory, and prediction markets. In
459 *14th Innovations in Theoretical Computer Science Conference (ITCS 2023)*, volume 251, page 81.
460 Schloss Dagstuhl–Leibniz-Zentrum f { " u } r Informatik, 2023.

461 Songtao Li and Hao Tang. Multimodal alignment and fusion: A survey. *arXiv preprint*
462 *arXiv:2411.17040*, 2024.

463 Jiuyao Lu, Aaron Roth, and Mirah Shi. Sample efficient omniprediction and downstream swap regret
464 for non-linear losses. *arXiv preprint arXiv:2502.12564*, 2025.

465 Balas K Natarajan. On learning sets and functions. *Machine Learning*, 4(1):67–97, 1989.

- 466 Georgy Noarov, Ramya Ramalingam, Aaron Roth, and Stephan Xie. High-dimensional prediction
467 for sequential decision making, 2023.
- 468 Gali Noti, Kate Donahue, Jon Kleinberg, and Sigal Oren. Ai-assisted decision making with human
469 learning. *arXiv preprint arXiv:2502.13062*, 2025.
- 470 Kenny Peng, Nikhil Garg, and Jon Kleinberg. A no free lunch theorem for human-ai collaboration.
471 *arXiv preprint arXiv:2411.15230*, 2024.
- 472 David Pollard. *Convergence of stochastic processes*. Springer Science & Business Media, 2012.
- 473 Mingda Qiao and Letian Zheng. On the distance from calibration in sequential prediction. In *The*
474 *Thirty Seventh Annual Conference on Learning Theory*, pages 4307–4357. PMLR, 2024.
- 475 Alexander Rakhlin, Karthik Sridharan, and Ambuj Tewari. Sequential complexities and uniform
476 martingale laws of large numbers. *Probability Theory and Related Fields*, 161, 02 2014. doi:
477 10.1007/s00440-013-0545-5.
- 478 Alexander Rakhlin, Karthik Sridharan, and Ambuj Tewari. Online learning via sequential complexi-
479 ties. *J. Mach. Learn. Res.*, 16(1):155–186, January 2015. ISSN 1532-4435.
- 480 Aaron Roth, Alexander Tolbert, and Scott Weinstein. Reconciling individual probability forecasts. In
481 *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*, pages
482 101–110, 2023.
- 483 Shai Shalev-Shwartz and Shai Ben-David. *Understanding machine learning: From theory to*
484 *algorithms*. Cambridge university press, 2014.
- 485 V.N. Vapnik and A. YA. Chervonenkis. On the uniform convergence of relative frequencies of events
486 to their probabilities, 1971.
- 487 Vladimir Vovk. On-line regression competitive with reproducing kernel hilbert spaces. In *Interna-*
488 *tional Conference on Theory and Applications of Models of Computation*, pages 452–463. Springer,
489 2006.
- 490 Volodya Vovk. Competitive on-line statistics. *International Statistical Review*, 69(2):213–248, 2001.
- 491 Kang Wei, Jun Li, Chuan Ma, Ming Ding, Sha Wei, Fan Wu, Guihai Chen, and Thilina Ran-
492 baduge. Vertical federated learning: Challenges, methodologies and experiments. *arXiv preprint*
493 *arXiv:2202.04309*, 2022.
- 494 Zihan Zhang, Wenhao Zhan, Yuxin Chen, Simon S Du, and Jason D Lee. Optimal multi-distribution
495 learning. In *The Thirty Seventh Annual Conference on Learning Theory*, pages 5220–5223. PMLR,
496 2024.
- 497 Shengjia Zhao, Michael Kim, Roshni Sahoo, Tengyu Ma, and Stefano Ermon. Calibrating predictions
498 to decisions: A novel approach to multi-class calibration. *Advances in Neural Information*
499 *Processing Systems*, 34:22313–22324, 2021.

500 A Preliminaries

501 We study a setting with two parties, Alice and Bob. Both parties are able to make predictions about a
502 label not only given their observed features, but as a function of an interaction that they have had with
503 their counterparty. With the exception of Section E and Appendix I, we consider the adaptive, online
504 setting where Alice and Bob interact to make label predictions over a sequence of days $t = 1, \dots, T$.
505 We let \mathcal{X}_A and \mathcal{X}_B denote feature spaces for Alice and Bob, respectively, and we let $\mathcal{X} = \mathcal{X}_A \times \mathcal{X}_B$
506 denote the joint feature space. We let \mathcal{Y} represent the outcome (label) space which we will take to be
507 $\mathcal{Y} = [0, 1]$ for much of the paper, generalizing it to higher dimensions in Appendix H.

508 On each day t , the parties converse for exactly K rounds about their predictions of that day’s outcome
509 y^t based on the features they each see: x_A^t and x_B^t , respectively. At each round k when they are
510 speaking, an agent makes a prediction of the label, denoted $\hat{y}_A^{t,k}$ and $\hat{y}_B^{t,k}$ respectively. This prediction

can be a function of everything the agent has observed so far — the features relevant to the instance, the predictions sent by the other party, and past outcomes on previous days.

They will alternate speaking, and we suppose that Alice (Party 1) acts in odd numbered rounds; Bob (Party 2) acts in even numbered rounds. In an odd round k , Alice sends her prediction $\hat{y}_A^{t,k}$, and then in the next round $k + 1$; Bob responds with a prediction $\hat{y}_B^{t,k+1}$. We use the subscript A and B for readability, so there is a clear distinction between Alice and Bob’s messages when possible. However, since which party is speaking is simply a function of the parity of the round k , we can also write $\hat{y}^{t,k}$ as shorthand for $\hat{y}_B^{t,k}$ or $\hat{y}_A^{t,k}$ when the round k is even or odd, respectively.

We formalize the interaction between the two agents in Protocol A—a generic “collaboration protocol.”

[ht] **Input** $(\mathcal{X}, \mathcal{Y}, K, T)$ each day $t = 1, \dots, T$ Receive $x^t = (x_A^t, x_B^t)$. Alice sees x_A^t and Bob sees x_B^t . each round $k = 1, 2, \dots, K$ k is odd Alice predicts $\hat{y}_A^{t,k} \in \mathcal{Y}$, and sends Bob $\hat{y}_A^{t,k}$. k is even Bob predicts $\hat{y}_B^{t,k}$, and sends Alice $\hat{y}_B^{t,k}$. Alice and Bob observe $y^t \in \mathcal{Y}$.

We informally refer to the history of interaction *within* any given day t as a “conversation.” This is, the sequence of predictions exchanged by Alice and Bob specifically about the currently unknown label y^t . We refer to the history of interaction *across* multiple days as a “conversation transcript.” It is an object that records the interactions between the agents and is visible to both, and which they can use to make their predictions.

Definition A.1 (Conversation Transcript $\pi^{1:T,1:K}$). A conversation transcript $\pi^{1:T,1:K} \in \{\mathcal{Y}^{K+1}\}^T$ is a sequence of tuples of predictions over rounds made by Alice and Bob (alternating across rounds), and the outcome, over T days:

$$\pi^{1:T,1:K} = \left\{ \left(\hat{y}_A^{1,1}, \hat{y}_B^{1,2}, \hat{y}_A^{1,3}, \dots, \hat{y}_A^{1,K}, y^1 \right), \dots, \left(\hat{y}_A^{T,1}, \hat{y}_B^{T,2}, \hat{y}_A^{T,3}, \dots, \hat{y}_A^{T,K}, y^T \right) \right\}.$$

We define $\pi^{1:T:k}$ to be the restriction to only round k of conversation across days as follows:

$$\pi^{1:T:k} = \begin{cases} \{(\hat{y}_A^{t,k}, y^t)\}_{t \in [T]} & \text{if } k \text{ is odd,} \\ \{(\hat{y}_B^{t,k}, y^t)\}_{t \in [T]} & \text{otherwise.} \end{cases}$$

We will use the notation $\pi^{1:T}$ to refer to a single sequence of predictions over T days, outside the context of a conversation.

Definition A.2 (Prediction Transcript $\pi^{1:T}$). A prediction transcript $\pi^{1:T} \in \{\mathcal{Y}^2\}^T$ is a sequence of tuples of predictions and outcomes over T days:

$$\pi^{1:T} = \{(\hat{y}^1, y^1), \dots, (\hat{y}^T, y^T)\}$$

A.1 Information Aggregation

Our focus is on giving algorithms in this collaborative learning setting that give strong information aggregation guarantees, in the sense that the parties, using only their own sets of features individually, converge on predictions that are optimal with respect to a benchmark class of predictors is defined with respect to *both* parties’ features.

In order to state such guarantees, we need to define a benchmark class. We first define the class of benchmark functions that map each of Alice and Bob’s features, individually, to predictions, and then the class of benchmark functions defined on their joint feature space.

Definition A.3 (Individual Hypothesis Classes $\mathcal{H}_A, \mathcal{H}_B$). Let $\mathcal{H}_A : \{h : \mathcal{X}_A \mapsto \mathbb{R}\}$ be a set of functions mapping from Alice’s feature set to \mathbb{R} . Analogously, let $\mathcal{H}_B : \{h : \mathcal{X}_B \mapsto \mathbb{R}\}$ be a set of functions mapping from Bob’s features to \mathbb{R} .

Definition A.4 (Joint Hypothesis Class \mathcal{H}_J). Let $\mathcal{H}_J : \{h : \mathcal{X} \mapsto \mathbb{R}\}$ be a set of functions mapping from the joint feature set $\mathcal{X} = \mathcal{X}_A \times \mathcal{X}_B$ to \mathbb{R} .

For simplicity, it will be convenient for us to assume that the hypothesis classes $\mathcal{H}_A, \mathcal{H}_B, \mathcal{H}_J$ contain constant functions (this is the case for most natural concept classes and is easy to enforce for any class for which it is not true originally)

Assumption 1. We assume that the hypothesis classes \mathcal{H} we work with contain the set of all constant functions $\{h(x) = v\}_{v \in [0,1]}$.

The goal of our collaboration protocol will be to guarantee that the sequence of predictions resulting from the interaction have error that is competitive with the best model in \mathcal{H}_J . In service of this, we will leverage the ability of Alice and Bob to make predictions that have low swap regret with respect to their individual hypothesis classes \mathcal{H}_A and \mathcal{H}_B respectively:

Definition A.5 ($((f, \mathcal{H})$ -Swap Regret). Fix an error function $f : \{1, \dots, T\} \rightarrow \mathbb{R}$ and a hypothesis class \mathcal{H} . A transcript $\pi^{1:T}$ has (f, \mathcal{H}) -swap regret if:

$$\sum_{t=1}^T (\hat{y}^t - y^t)^2 - \sum_v \min_{h \in \mathcal{H}} \left(\sum_{t=1}^T \mathbb{I}[\hat{y}^t = v] (h(x^t) - y^t)^2 \right) \leq f(T)$$

Here v ranges over values of the predictions: $v \in \{\hat{y}^1, \dots, \hat{y}^T\}$.

It will also be useful to have a notion of *distance* to swap regret. Distance to swap regret, which we define below, is analogous to the recently defined measure of *distance to calibration* Błasiok et al. [2023]. A sequence of predictions has low distance to swap regret, informally, if they are close (in ℓ_1 distance) to a sequence of predictions that itself has low swap regret.

Definition A.6 ($((q, f, \mathcal{H})$ -Distance to Swap Regret). Fix an error functions $f, q : \{1, \dots, T\} \rightarrow \mathbb{R}$ and a hypothesis class \mathcal{H} . Let $Q_{f, \mathcal{H}}$ be the set of prediction sequences $p^{1:T}$ that have (f, \mathcal{H}) -swap regret. A transcript $\pi^{1:T}$ has (q, f, \mathcal{H}) -distance to swap regret if:

$$\min_{p^{1:T} \in Q_{f, \mathcal{H}}} \|\hat{y}^{1:T} - p^{1:T}\|_1 \leq q(T)$$

A.2 Conversation Swap Regret

Our collaboration protocols involve “conversations” over k rounds. An important condition for us in our construction is called “conversation swap regret”, which informally requires that the predictions that Alice (resp. Bob) make at each round of conversation have no swap regret with respect to \mathcal{H}_A (resp. \mathcal{H}_B) not just marginally, but *conditionally* on the prediction that their counter-party made at the round before.

Definition A.7 ($((f, g, \mathcal{H})$ -Conversation Swap Regret). Fix an error function $f : \{1, \dots, T\} \rightarrow \mathbb{R}$, a bucketing function $g : \{1, \dots, T\} \rightarrow \mathbb{R}$ and a prediction class \mathcal{H}_B . Let v range over the values $\in \{\hat{y}_k^1, \dots, \hat{y}_k^T\}$. Given a conversation transcript $\pi^{1:T, 1:K}$ from an interaction in the Collaboration Protocol (Protocol A), Bob has (f, g, \mathcal{H}_B) -swap regret if for all even rounds k and buckets $i \in \{1, \dots, \frac{1}{g(T)}\}$:

$$\sum_{t \in T_A(k-1, i)} (\hat{y}^{t,k} - y^t)^2 - \sum_v \min_{h \in \mathcal{H}_B} \left(\sum_{t \in T_A(k-1, i)} \mathbb{I}[\hat{y}^{t,k} = v] (h(x^t) - y^t)^2 \right) \leq f(|T_A(k-1, i)|).$$

Where $T_A(k-1, i) = \{t : \hat{y}^{t,k-1} \in [(i-1)g(T), ig(T))\}$ is the subsequence of days where the predictions of Alice in round $k-1$ fall in bucket i .

If Alice satisfies a symmetric condition on odd rounds k with respect to \mathcal{H}_A , we say that Alice has (f, g, \mathcal{H}_A) -Conversation Swap Regret with respect to \mathcal{H}_A .

Assumption 2. We assume that all error functions $f(\cdot)$ are concave.

B Boosting for Collaboration

In this section we give a weak learning condition that characterizes when swap regret guarantees with respect to \mathcal{H}_A and \mathcal{H}_B on a single sequence of predictions imply regret guarantees with respect to a richer hypothesis class \mathcal{H}_J . We also show that linear functions (and substantial generalizations) over \mathcal{X}_A and \mathcal{X}_B indeed satisfy the weak learning condition with respect to \mathcal{H}_J , linear functions over the joint feature space \mathcal{X} . This justifies the algorithmic approach we pursue in Section C, giving

collaboration protocols whose aim is to arrive at a sequence of predictions that have no swap regret with respect to both \mathcal{H}_A and \mathcal{H}_B — the final accuracy guarantees in those sections will then follow from applying the boosting theorem we will prove here.

We first state our weak learning condition, which roughly speaking requires that on every distribution, if there is any model in \mathcal{H}_J that is able to obtain error lower than that of a constant predictor (by any margin γ), then there must also be a model in either \mathcal{H}_A or \mathcal{H}_B that can obtain error better than a constant predictor (by some smaller margin $w(\gamma)$). This is a generalization of a condition given in Globus-Harris et al. [2023] in the context of studying the boosting properties of multicalibration. Our definition below generalizes that of Globus-Harris et al. [2023] to multiple parties, and to a general margin function w (rather than just a linear function $w(\gamma) = \gamma$ as stated in Globus-Harris et al. [2023]). This generalization is important because as we will see, linear functions satisfy the weak learning condition only with the margin $w(\gamma) = \Theta(\gamma^2)$.

Definition B.1 ($w(\cdot)$ -Weak Learning Condition). *Let $\mathcal{H}_A = \{h_A : \mathcal{X}_A \rightarrow \mathcal{Y}\}$ and $\mathcal{H}_B = \{h_B : \mathcal{X}_B \rightarrow \mathcal{Y}\}$ be hypothesis classes over \mathcal{X}_A and \mathcal{X}_B respectively. Let \mathcal{H}_J be a hypothesis class of over the joint feature space $\mathcal{X} = \mathcal{X}_A \times \mathcal{X}_B$. Let $w : [0, 1] \rightarrow [0, 1]$ be a strictly increasing, continuous, convex function that satisfies $w(\gamma) \leq \gamma$. We say that \mathcal{H}_A and \mathcal{H}_B jointly satisfy the $w(\cdot)$ -weak learning condition with respect to \mathcal{H}_J if for any distribution \mathcal{D} over $\mathcal{X}_A \times \mathcal{X}_B \times \mathcal{Y}$, and any $\gamma \in [0, 1]$, if:*

$$\min_{c \in \mathbb{R}} \mathbb{E}_{\mathcal{D}}[(c - y)^2] - \min_{h_J \in \mathcal{H}_J} \mathbb{E}_{\mathcal{D}}[(h_J(x) - y)^2] \geq \gamma,$$

then there exists either $h_A \in \mathcal{H}_A$ or $h_B \in \mathcal{H}_B$ such that:

$$\min_{c \in \mathbb{R}} \mathbb{E}_{\mathcal{D}}[(c - y)^2] - \mathbb{E}_{\mathcal{D}}[(h_A(x_A) - y)^2] \geq w(\gamma)$$

or:

$$\min_{c \in \mathbb{R}} \mathbb{E}_{\mathcal{D}}[(c - y)^2] - \mathbb{E}_{\mathcal{D}}[(h_B(x_B) - y)^2] \geq w(\gamma)$$

Remark B.2. *We note that the conditions that w is convex and satisfies $w(\gamma) \leq \gamma$ is without loss. Indeed, if \mathcal{H}_A and \mathcal{H}_B jointly improve by a margin w' that is non-convex, there exists a convex function w such that $w(\gamma) \leq w'(\gamma)$ for all $\gamma \in [0, 1]$, and thus, \mathcal{H}_A and \mathcal{H}_B also jointly improve by the margin w . Similarly, if $w(\gamma) > \gamma$ for some γ — i.e. \mathcal{H}_A and \mathcal{H}_B jointly improve by more than γ — then they certainly improve by at least γ . We impose these conditions for technical reasons in the proof of Theorem B.3.*

We now state our “boosting” theorem. In fact, we will not need that our predictions have low swap regret — it will suffice that they have low *distance to swap regret*, which will be an easier condition to obtain. If we have a single sequence of predictions such that those predictions have low distance to swap regret with respect to \mathcal{H}_A and \mathcal{H}_B , and \mathcal{H}_A and \mathcal{H}_B satisfy our weak learning condition with respect to a stronger joint class of functions \mathcal{H}_J , then in fact the sequence of predictions has no regret with respect to \mathcal{H}_J .

Theorem B.3. *Let \mathcal{H}_J be a hypothesis class over the joint feature space \mathcal{X} . Let $\mathcal{H}_A = \{h_A : \mathcal{X}_A \rightarrow \mathcal{Y}\}$ and $\mathcal{H}_B = \{h_B : \mathcal{X}_B \rightarrow \mathcal{Y}\}$ be hypothesis classes over \mathcal{X}_A and \mathcal{X}_B respectively. Let $\mathcal{D} \in \Delta(\mathcal{X} \times \mathcal{Y})$ be the empirical distribution over a sequence $(x^t, y^t)_{t=1}^T$. If:*

- Predictions $\hat{y}^{1:T}$ have $(q, f, \mathcal{H}_A \cup \mathcal{H}_B)$ -distance to swap regret over \mathcal{D} , and
- \mathcal{H}_A and \mathcal{H}_B jointly satisfy the $w(\cdot)$ -weak learning condition with respect to \mathcal{H}_J

Then:

$$\mathbb{E}_{\mathcal{D}}[(\hat{y} - y)^2] - \min_{h_J \in \mathcal{H}_J} \mathbb{E}_{\mathcal{D}}[(h_J(x) - y)^2] \leq 2w^{-1} \left(\frac{f(T)}{T} \right) + 3 \frac{q(T)}{T}$$

whenever the inverse of w exists.

We first show that if our predictions have no distance to swap regret, then the weak learning condition implies low external regret with respect to \mathcal{H}_J . We will then argue that perturbing the predictions by a small amount cannot increase external regret by very much.

Lemma B.4. *Let \mathcal{H}_J be a hypothesis class over the joint feature space \mathcal{X} . Let $\mathcal{H}_A = \{h_A : \mathcal{X}_A \rightarrow \mathcal{Y}\}$ and $\mathcal{H}_B = \{h_B : \mathcal{X}_B \rightarrow \mathcal{Y}\}$ be hypothesis classes over \mathcal{X}_A and \mathcal{X}_B respectively. Let $\mathcal{D} \in \Delta(\mathcal{X} \times \mathcal{Y})$ be the empirical distribution over a sequence $(x^t, y^t)_{t=1}^T$. If:*

- Predictions $\hat{y}^{1:T}$ have $(f, \mathcal{H}_A \cup \mathcal{H}_B)$ -swap regret over \mathcal{D} , and
- \mathcal{H}_A and \mathcal{H}_B jointly satisfy the $w(\cdot)$ -weak learning condition with respect to \mathcal{H}_J

Then:

$$\mathbb{E}_{\mathcal{D}}[(\hat{y} - y)^2] - \min_{h_J \in \mathcal{H}_J} \mathbb{E}_{\mathcal{D}}[(h_J(x) - y)^2] \leq 2w^{-1} \left(\frac{f(T)}{T} \right)$$

whenever the inverse of w exists.

Proof. We show the contrapositive. Suppose there exists $h_J \in \mathcal{H}_J$ such that:

$$\frac{1}{T} \sum_v \sum_{t=1}^T \mathbb{1}[\hat{y}^t = v] (h_J(x^t) - y^t)^2 < \frac{1}{T} \sum_v \sum_{t=1}^T \mathbb{1}[\hat{y}^t = v] (v - y^t)^2 - 2w^{-1} \left(\frac{f(T)}{T} \right)$$

Since a swap benchmark is only stronger, there exists a collection $\{h_{J,v}\}_v \subseteq \mathcal{H}_J$ such that:

$$\frac{1}{T} \sum_v \sum_{t=1}^T \mathbb{1}[\hat{y}^t = v] (h_{J,v}(x^t) - y^t)^2 \leq \frac{1}{T} \sum_v \sum_{t=1}^T \mathbb{1}[\hat{y}^t = v] (h_J(x^t) - y^t)^2$$

and thus:

$$\frac{1}{T} \sum_v \sum_{t=1}^T \mathbb{1}[\hat{y}^t = v] (h_{J,v}(x^t) - y^t)^2 < \frac{1}{T} \sum_v \sum_{t=1}^T \mathbb{1}[\hat{y}^t = v] (v - y^t)^2 - 2w^{-1} \left(\frac{f(T)}{T} \right)$$

Let $S_v = \{t : \hat{y}^t = v\}$ be the level set corresponding to the subset of the domain that the prediction is v . Let $\bar{y}_v = \frac{1}{|S_v|} \sum_{t=1}^T \mathbb{1}[\hat{y}^t = v] y^t$ be the label mean of this subset. By Assumption 1, $\mathcal{H}_A \cup \mathcal{H}_B$ contains the set of all constant functions in $[0, 1]$. Let $\mathcal{H}_C \subset \mathcal{H}_A \cup \mathcal{H}_B$ denote the set of constant functions. Since, for every v , $h_c(x) = \bar{y}_v$ is the constant function that minimizes squared error, and $\hat{y}^{1:T}$ has (f, \mathcal{H}_C) -swap regret, we have that the average swap regret with respect to \mathcal{H}_C is bounded by:

$$\begin{aligned} & \frac{1}{T} \sum_v \sum_{t=1}^T \mathbb{1}[\hat{y}^t = v] (v - y^t)^2 - \frac{1}{T} \sum_v \sum_{t=1}^T \mathbb{1}[\hat{y}^t = v] (\bar{y}_v - y^t)^2 \\ &= \frac{1}{T} \sum_v \sum_{t=1}^T \mathbb{1}[\hat{y}^t = v] (v - y^t)^2 - \frac{1}{T} \sum_v \min_{h_c \in \mathcal{H}_C} \sum_{t=1}^T \mathbb{1}[\hat{y}^t = v] (h_c(x^t) - y^t)^2 \\ &\leq \frac{f(T)}{T} \\ &\leq w^{-1} \left(\frac{f(T)}{T} \right) \end{aligned}$$

In the last step, we use the fact that $w(\gamma) \leq \gamma$, and so $\gamma \leq w^{-1}(\gamma)$. Then, since the squared error of $\{h_{J,v}\}_v$ is less than the squared error of $\hat{y}^{1:T}$, and the squared error of $\hat{y}^{1:T}$ is close to the squared error of the label mean \bar{y}_v on each level set, we have that:

$$\begin{aligned} \frac{1}{T} \sum_v \sum_{t=1}^T \mathbb{1}[\hat{y}^t = v] (h_{J,v}(x^t) - y^t)^2 &< \frac{1}{T} \sum_v \sum_{t=1}^T \mathbb{1}[\hat{y}^t = v] (v - y^t)^2 - 2w^{-1} \left(\frac{f(T)}{T} \right) \\ &\leq \frac{1}{T} \sum_v \sum_{t=1}^T \mathbb{1}[\hat{y}^t = v] (\bar{y}_v - y^t)^2 + w^{-1} \left(\frac{f(T)}{T} \right) - 2w^{-1} \left(\frac{f(T)}{T} \right) \\ &= \frac{1}{T} \sum_v \sum_{t=1}^T \mathbb{1}[\hat{y}^t = v] (\bar{y}_v - y^t)^2 - w^{-1} \left(\frac{f(T)}{T} \right) \end{aligned}$$

Letting

$$\gamma_v = \frac{1}{|S_v|} \sum_{t=1}^T \mathbb{1}[\hat{y}^t = v] (\bar{y}_v - y^t)^2 - \frac{1}{|S_v|} \sum_{t=1}^T \mathbb{1}[\hat{y}^t = v] (h_{J,v}(x^t) - y^t)^2,$$

651 we can rewrite the expression above as:

$$\begin{aligned}
& \frac{1}{T} \sum_v \sum_{t=1}^T \mathbb{1}[\hat{y}^t = v](\bar{y}_v - y^t)^2 - \frac{1}{T} \sum_v \sum_{t=1}^T \mathbb{1}[\hat{y}^t = v](h_{J,v}(x^t) - y^t)^2 \\
&= \frac{1}{T} \sum_v |S_v| \cdot \frac{1}{|S_v|} \sum_{t=1}^T \mathbb{1}[\hat{y}^t = v](\bar{y}_v - y^t)^2 - \frac{1}{T} \sum_v |S_v| \cdot \frac{1}{|S_v|} \sum_{t=1}^T \mathbb{1}[\hat{y}^t = v](h_{J,v}(x^t) - y^t)^2 \\
&= \frac{1}{T} \sum_v |S_v| \gamma_v \\
&> w^{-1} \left(\frac{f(T)}{T} \right)
\end{aligned}$$

652 Observe that since \mathcal{H}_J contains the set of all constant functions (Assumption 1), there is always a
653 choice of $\{h_{J,v}\}_v$ such that γ_v is non-negative for all v .

654 Thus, by the $w(\cdot)$ -weak learning condition applied to the empirical distribution over the sequence
655 on which $\hat{y}^t = v$ for any level set v , if $h_{J,v}$ improves over the best constant prediction \bar{y}_v by γ_v ,
656 there is some $h_v \in \mathcal{H}_A \cup \mathcal{H}_B$ that improves over \bar{y}_v by $w(\gamma_v)$. That is, there exists a collection
657 $\{h_v\} \subseteq \mathcal{H}_A \cup \mathcal{H}_B$ such that:

$$\begin{aligned}
& \frac{1}{T} \sum_v \sum_{t=1}^T \mathbb{1}[\hat{y}^t = v](\bar{y}_v - y^t)^2 - \frac{1}{T} \sum_v \sum_{t=1}^T \mathbb{1}[\hat{y}^t = v](h_v(x^t) - y^t)^2 \\
&= \frac{1}{T} \sum_v |S_v| \cdot \frac{1}{|S_v|} \sum_{t=1}^T \mathbb{1}[\hat{y}^t = v](\bar{y}_v - y^t)^2 - \frac{1}{T} \sum_v |S_v| \cdot \frac{1}{|S_v|} \sum_{t=1}^T \mathbb{1}[\hat{y}^t = v](h_v(x^t) - y^t)^2 \\
&\geq \frac{1}{T} \sum_v |S_v| w(\gamma_v) \quad \text{(by the } w\text{-weak learning condition)} \\
&\geq w \left(\frac{1}{T} \sum_v |S_v| \gamma_v \right) \quad \text{(by convexity of } w \text{ and Jensen's inequality)} \\
&> w \left(w^{-1} \left(\frac{f(T)}{T} \right) \right) \quad \text{(by monotonicity of } w) \\
&= \frac{f(T)}{T}
\end{aligned}$$

658 In particular, this implies that:

$$\begin{aligned}
& \sum_v \min_{h_v^* \in \mathcal{H}_A \cup \mathcal{H}_B} \sum_{t=1}^T \mathbb{1}[\hat{y}^t = v](h_v^*(x^t) - y^t)^2 \leq \sum_v \sum_{t=1}^T \mathbb{1}[\hat{y}^t = v](h_v(x^t) - y^t)^2 \\
&< \sum_v \sum_{t=1}^T \mathbb{1}[\hat{y}^t = v](\bar{y}_v - y^t)^2 - f(T) \\
&\leq \sum_v \sum_{t=1}^T \mathbb{1}[\hat{y}^t = v](v - y^t)^2 - f(T) \\
&= \sum_{t=1}^T (\hat{y}^t - y^t)^2 - f(T)
\end{aligned}$$

659 Here, the third line follows from the fact that on level set v , the squared error of the constant prediction
660 v is at least the squared error of the best constant prediction \bar{y}_v . This violates the $(f, \mathcal{H}_A \cup \mathcal{H}_B)$ -swap
661 regret condition, which completes the proof.

662 □

663 We can now complete the proof by noting that squared error is Lipschitz in the predictions — so
664 perturbing predictions that have low swap regret to those that merely have low *distance* to swap regret
665 does not affect the final error bound by much:

666 *Proof of Theorem B.3.* By definition of distance to swap regret, there is a sequence $p^{1:T}$ with
 667 $(f, \mathcal{H}_A \cup \mathcal{H}_B)$ -swap regret such that $\|\hat{y}^{1:T} - p^{1:T}\|_1 \leq q(T)$. Furthermore, by Lemma B.4, $p^{1:T}$
 668 satisfies:

$$\frac{1}{T} \sum_{t=1}^T (p^t - y^t)^2 - \min_{h_J \in \mathcal{H}_J} \frac{1}{T} \sum_{t=1}^T (h_J(x^t) - y^t)^2 \leq 2w^{-1} \left(\frac{f(T)}{T} \right)$$

669 Applying Lemma K.14, we can conclude:

$$\begin{aligned} & \frac{1}{T} \sum_{t=1}^T (\hat{y}^t - y^t)^2 - \min_{h_J \in \mathcal{H}_J} \frac{1}{T} \sum_{t=1}^T (h_J(x^t) - y^t)^2 \\ & \leq \frac{1}{T} \sum_{t=1}^T (q^t - y^t)^2 - \min_{h_J \in \mathcal{H}_J} \frac{1}{T} \sum_{t=1}^T (h_J(x^t) - y^t)^2 + 3 \frac{q(T)}{T} \\ & \leq 2w^{-1} \left(\frac{f(T)}{T} \right) + 3 \frac{q(T)}{T} \end{aligned}$$

670 as desired. □

671 B.1 Function Classes Satisfying the Weak Learning Guarantee

672 Next, we show that a broad set of function classes satisfy our weak learning assumption. What we
 673 require is that \mathcal{H}_A and \mathcal{H}_B be “star shaped” (i.e. closed under downward scaling), bounded, and closed
 674 under additive shifts, and that \mathcal{H}_J be representable as the Minkowski sum of \mathcal{H}_A and \mathcal{H}_B — that is,
 675 for every $h_J \in \mathcal{H}_J$ there should be $h_A \in \mathcal{H}_A$ and $h_B \in \mathcal{H}_B$ such that $h_J(x) = h_A(x_A) + h_B(x_B)$.
 676 In particular, the class of linear functions over the feature spaces of Alice and Bob respectively satisfy
 677 our weak learning assumption relative to linear functions on their joint feature space.

678 In order to define Alice and Bob’s function classes, let us first define a few useful properties.

679 **Definition B.5** (Bounded and star-shaped function class). *For any class $\mathcal{F} = \{f : \mathcal{X} \rightarrow \mathbb{R}\}$ on*
 680 *domain \mathcal{X} , we say it is*

- 681 1. *C-bounded: if there exists $C > 0$ such that $\sup_{f \in \mathcal{F}, x \in \mathcal{X}} |f(x)| \leq C$*
- 682 2. *Star-shaped: if $f \in \mathcal{F}$ then $\alpha f \in \mathcal{F}$ for all $0 \leq \alpha \leq 1$.*

683 Note that the function class of linear functions with bounded norms $\mathcal{F} = \{x \mapsto \theta^\top x : \|\theta\|_2 \leq C\}$
 684 over bounded inputs $\mathcal{X} = \{x \in \mathbb{R}^d : \|x\|_2 \leq 1\}$ is C -bounded and star-shaped.

685 We now state our weak-learnability guarantees with respect to the Minkowski sum of our base function
 686 classes satisfying the above properties.

Theorem B.6. *Let $\mathcal{H}_A = \{f_A + b_A : f_A \in \mathcal{F}_A, b_A \in \mathbb{R}\}$ and $\mathcal{H}_B = \{f_B + b_B : f_B \in \mathcal{F}_B, b_B \in \mathbb{R}\}$
 where $\mathcal{F}_A = \{f_A : \mathcal{X}_A \rightarrow \mathbb{R}\}$ and $\mathcal{F}_B = \{f_B : \mathcal{X}_B \rightarrow \mathbb{R}\}$ are C -bounded and star-shaped. Let
 $\mathcal{H}_J = \{h_A + h_B : h_A \in \mathcal{H}_A, h_B \in \mathcal{H}_B\}$ be the Minkowski sum of \mathcal{H}_A and \mathcal{H}_B . If $C \geq 1/2$, then
 \mathcal{H}_A and \mathcal{H}_B jointly satisfy the $w(\cdot)$ -weak learning condition with respect to \mathcal{H}_J for:*

$$w(\gamma) = \frac{\gamma^2}{16C^2}$$

687 The key idea is to show that if a function in the joint class $h_J(x) = h_A(x_A) + h_B(x_B)$ improves
 688 over the constant predictor then this translates to at least one of the base functions $h_A(x_A)$ or $h_B(x_B)$
 689 having non-trivial correlation with the label y . Now appropriately choosing the scaling of the base
 690 function allows us to transfer this correlation to an improvement in squared loss over the constant
 691 predictor. This transfer is not exact and leads to the weaker γ^2 improvement, which we later show is
 692 actually tight!

693 *Proof of Theorem B.6.* Consider a distribution \mathcal{D} over $\mathcal{X}_A \times \mathcal{X}_B \times \mathcal{Y}$ with $\mu = \mathbb{E}_{\mathcal{D}}[y]$. Define
 694 $\bar{y} = y - \mu$ so that $\mathbb{E}_{\mathcal{D}}[\bar{y}] = 0$, and for any predictor $h(x)$, define the centered predictor $\bar{h}(x) =$
 695 $h(x) - \mu$. The best constant predictor for predicting \bar{y} is 0 with error $\mathbb{E}_{\mathcal{D}}[\bar{y}^2]$, and for any predictor,
 696 $\mathbb{E}_{\mathcal{D}}[(h(x) - y)^2] = \mathbb{E}_{\mathcal{D}}[(\bar{h}(x) - \bar{y})^2]$.

697 To prove the weak learnability condition, assume that there exists $h_J(x) = h_A(x_A) + h_B(x_B) =$
698 $f_A(x_A) + f_B(x_B) + b_A + b_B$ and its corresponding centered version $\bar{h}_J(x) = h_J(x) - \mu$ such that

$$\mathbb{E}_{\mathcal{D}}[(\bar{h}_J(x) - \bar{y})^2] \leq \mathbb{E}_{\mathcal{D}}[\bar{y}^2] - \gamma \implies -\mathbb{E}_{\mathcal{D}}[(\bar{h}_J(x))^2] + 2\mathbb{E}_{\mathcal{D}}[\bar{h}_J(x)\bar{y}] \geq \gamma.$$

699 Given that $\mathbb{E}_{\mathcal{D}}[(\bar{h}_J(x))^2] \geq 0$, we have:

$$\begin{aligned} \mathbb{E}_{\mathcal{D}}[\bar{h}_J(x)\bar{y}] &\geq \frac{\gamma}{2} \\ \implies \mathbb{E}_{\mathcal{D}}[(f_A(x_A) + f_B(x_B) + b_A + b_B - \mu)\bar{y}] &\geq \frac{\gamma}{2} \\ \implies \mathbb{E}_{\mathcal{D}}[f_A(x_A)\bar{y}] + \mathbb{E}_{\mathcal{D}}[f_B(x_B)\bar{y}] &\geq \frac{\gamma}{2}, \end{aligned}$$

700 where the last inequality follows from the fact that $\mathbb{E}_{\mathcal{D}}[\bar{y}] = 0$. This implies that either

$$\mathbb{E}_{\mathcal{D}}[f_A(x_A)\bar{y}] \geq \frac{\gamma}{4} \quad \text{or} \quad \mathbb{E}_{\mathcal{D}}[f_B(x_B)\bar{y}] \geq \frac{\gamma}{4}.$$

701 Without loss of generality, assume $\mathbb{E}_{\mathcal{D}}[f_A(x_A)\bar{y}] \geq \frac{\gamma}{4}$. Now let us construct $h_A(x_A) = \alpha f_A(x_A) + \mu$
702 and corresponding centered predictor $\bar{h}_1(x_1) = \alpha f_A(x_A)$ for $\alpha = \frac{\gamma}{4C^2}$. Note that $h_A \in \mathcal{H}_A$ since
703 $\alpha \leq 1$ (by assumption) and \mathcal{F}_A is star-shaped.

704 Now let us compute the error of $h_A(x_A)$.

$$\begin{aligned} \mathbb{E}_{\mathcal{D}}[\bar{y}^2] - \mathbb{E}_{\mathcal{D}}[(\bar{h}_A(x_A) - \bar{y})^2] &= -\mathbb{E}_{\mathcal{D}}[(\bar{h}_A(x_A))^2] + 2\mathbb{E}_{\mathcal{D}}[\bar{h}_A(x_A)\bar{y}] \\ &= -\mathbb{E}_{\mathcal{D}}[(\alpha f_A(x_A))^2] + 2\mathbb{E}_{\mathcal{D}}[\alpha f_A(x_A)\bar{y}] \\ &\geq -\alpha^2 C^2 + \frac{\alpha\gamma}{2} = \frac{\gamma^2}{16C^2}. \end{aligned}$$

705 where the inequality follows from C -boundedness of \mathcal{F}_A and $\mathbb{E}_{\mathcal{D}}[f_A(x_A)\bar{y}] \geq \frac{\gamma}{4}$. Removing the
706 centering gives us,

$$\mathbb{E}_{\mathcal{D}}[(\mu - y)^2] - \mathbb{E}_{\mathcal{D}}[(h_A(x_A) - y)^2] = \mathbb{E}_{\mathcal{D}}[\bar{y}^2] - \mathbb{E}_{\mathcal{D}}[(\bar{h}_A(x_A) - \bar{y})^2] \geq \frac{\gamma^2}{16C^2}.$$

707 □

708 We next establish the tightness of various aspects of our theorem, both qualitatively and quantitatively.
709 First, we have assumed that our function classes are bounded. This is necessary:

710 **Theorem B.7.** *There exists classes $\mathcal{F}_A = \{f_A : \mathcal{X}_A \rightarrow \mathbb{R}\}$ and $\mathcal{F}_B = \{f_B : \mathcal{X}_B \rightarrow \mathbb{R}\}$ that are star-*
711 *shaped but unbounded over some domain $\mathcal{X}_A, \mathcal{X}_B$ such that $\mathcal{H}_A = \{f_A + b_A : f_A \in \mathcal{F}_A, b_A \in \mathbb{R}\}$*
712 *and $\mathcal{H}_B = \{f_B + b_B : f_B \in \mathcal{F}_B, b_B \in \mathbb{R}\}$ do not jointly satisfy $w(\cdot)$ -weak learning with respect to*
713 *$\mathcal{H}_J = \{h_A + h_B : h_A \in \mathcal{H}_A, h_B \in \mathcal{H}_B\}$ for any strictly increasing w .*

714 To prove this theorem, we construct a simple distribution (\mathcal{X}_A and \mathcal{X}_B are one-dimensional and $\mathcal{H}_A,$
715 \mathcal{H}_B , and \mathcal{H}_J are linear functions) where both the features to the individual parties x_A and x_B have a
716 small signal to noise ratio and hence cannot predict the label y very accurately, but their difference
717 can exactly cancel out the noise to recover a scaled down version of the signal. Now scaling it up can
718 recover the label y exactly. The signal to noise ratio for the individual parties is inversely proportional
719 to the norm of the joint predictor, therefore we can make this arbitrarily small if the norms are allowed
720 to be unbounded.

721 Using the same construction, we establish that the quadratic dependence on the weak learning margin
722 $w(\gamma) = \Theta(\gamma^2)$ cannot be improved. In particular, despite bounding the norm of the predictors, the
723 perfect canceling of noise allows the joint predictor to do an order of magnitude better than any
724 individual predictor on noisy features.

725 **Theorem B.8.** *There exists classes $\mathcal{F}_A = \{f_A : \mathcal{X}_A \rightarrow \mathbb{R}\}$ and $\mathcal{F}_B = \{f_B : \mathcal{X}_B \rightarrow \mathbb{R}\}$ that are star-*
726 *shaped and 1-bounded over some domain $\mathcal{X}_A, \mathcal{X}_B$ such that $\mathcal{H}_A = \{f_A + b_A : f_A \in \mathcal{F}_A, b_A \in \mathbb{R}\}$*
727 *and $\mathcal{H}_B = \{f_B + b_B : f_B \in \mathcal{F}_B, b_B \in \mathbb{R}\}$ do not jointly satisfy $w(\cdot)$ -weak learning with respect to*
728 *$\mathcal{H}_J = \{h_A + h_B : h_A \in \mathcal{H}_A, h_B \in \mathcal{H}_B\}$ for any strictly increasing w such that $w(\gamma) = \omega(\gamma^2)$.*

729 C Collaboration in the Online Setting

730 In Section B, we established that if \mathcal{H}_A and \mathcal{H}_B satisfy our weak learning condition with respect to
 731 \mathcal{H}_J , a sequence of predictions that has low distance to swap regret with respect to $\mathcal{H}_A \cup \mathcal{H}_B$ has
 732 low external regret with respect to \mathcal{H}_J . In this section, we show how to arrive at such a prediction
 733 sequence via a collaboration protocol.

734 The high level idea of the proof is straightforward, but the details are surprisingly subtle. Collina
 735 et al. [2025] defines a notion called Conversation Calibration in settings (such as our collaboration
 736 protocol) in which two parties engage in conversations about predictions of a real valued outcome.
 737 This notion is formally defined in Appendix K.2. Informally speaking, conversation calibration
 738 requires that at each round k of the conversation, the sequence of predictions made over the T days is
 739 unbiased relative to the outcomes, conditional both on the prediction made at round k and on the
 740 prediction made by the other party at round $k - 1$. Collina et al. [2025] show that if both parties
 741 satisfy the conversation calibration condition across all rounds, then most conversations must quickly
 742 converge to approximate agreement. The conversation swap regret condition we require of our parties
 743 implies that the predictions also satisfy conversation calibration, and so the theorem of Collina et al.
 744 [2025] implies fast approximate agreement in our setting as well. The idea at a high level is that if
 745 Alice’s predictions have no swap regret with respect to \mathcal{H}_A at every round, and Bob’s predictions
 746 have no swap regret with respect to \mathcal{H}_B at every round, then when they agree, we will have a single
 747 sequence of predictions that has no swap regret with respect to both \mathcal{H}_A and \mathcal{H}_B simultaneously,
 748 exactly the condition that we need in order to invoke our boosting theorem.

749 However, several difficulties arise. First, the agreement theorem of Collina et al. [2025] states
 750 informally that conversations on most days must reach agreement quickly, but they might reach
 751 agreement at different rounds on different days. Just because the predictions at each round satisfy
 752 swap-regret guarantees does not mean that the sequence of final “agreed upon” predictions — stitched
 753 together from different rounds at different days — will have the same guarantee. To solve this
 754 problem, we use a different protocol than Collina et al. [2025]: rather than halting conversation at
 755 agreement, we continue each conversation for K rounds even if agreement is reached earlier. We
 756 generalize the agreement theorem of Collina et al. [2025] to show that (even if it is not the final
 757 round), for sufficiently large K there must *exist* a round k at which Alice’s predictions at round k are
 758 close to Bob’s predictions at round $k - 1$:

759 **Theorem C.1.** *If Alice has $(f_A, g_A, \mathcal{H}_A)$ -conversation swap regret and Bob has $(f_B, g_B, \mathcal{H}_B)$ -*
 760 *conversation swap regret and they engage in a Collaboration Protocol (Protocol A) for K rounds,*
 761 *then for any $\epsilon \in (0, 1)$, there is at least one round k such that*

$$\frac{1}{T} \sum_{t=1}^T \mathbb{I}[|\hat{y}^{t,k} - \hat{y}^{t,k-1}| \geq \epsilon] \leq \frac{1}{2K\epsilon^2} + \frac{\beta(T, f_A, f_B)}{2\epsilon^2}$$

That is, the fraction of predictions in round k that are ϵ -away from those in round $k - 1$ is at most

$$\frac{1}{2K\epsilon^2} + \frac{\beta(T, f_A, f_B)}{2\epsilon^2}$$

762 *Here, and for the other theorems following, we let $\beta(T, f_A, f_B) = \frac{f_A(g_A(T) \cdot T)}{T g_A(T)} + \frac{f_B(g_B(T) \cdot T)}{T g_B(T)} +$*
 763 *$g_A(T) + g_B(T)$, $f'_A(x) = \sqrt{x \cdot f_A(x)}$ and $f'_B(x) = \sqrt{x \cdot f_B(x)}$.*

764 The proof for this theorem (and all other theorems this section) can be found in Appendix K.

765
 766 If on Alice’s rounds, she has low swap regret with respect to \mathcal{H}_A and on Bob’s rounds, he has low
 767 swap regret with respect to \mathcal{H}_B , then if on a pair of adjacent rounds, they made *exactly* the same
 768 predictions, then on (both) of these rounds, the predictions would have no swap regret with respect
 769 to \mathcal{H}_A and \mathcal{H}_B simultaneously. Unfortunately Theorem C.1 does not guarantee a pair of rounds
 770 on which Alice and Bob’s predictions are exactly the same — it only guarantees a pair of adjacent
 771 rounds on which the predictions are *close* on *most days*. Naively, this gives us two sequences, one
 772 of which has low swap regret with respect to \mathcal{H}_A and low distance to swap regret with respect to
 773 \mathcal{H}_B , and the other of which has low swap regret with respect to \mathcal{H}_B and low distance to swap regret
 774 with respect to \mathcal{H}_A . But to apply our boosting theorem, we need a single sequence of predictions

that simultaneously has low distance to swap regret with respect to both \mathcal{H}_A and \mathcal{H}_B . The following theorem (Theorem C.2) shows that in fact the round k identified in Theorem C.1 has this property:

Theorem C.2. *If Alice has $(f_A, g_A, \mathcal{H}_A)$ -conversation swap regret and Bob has $(f_B, g_B, \mathcal{H}_B)$ -conversation swap regret, and they engage in a Collaboration Protocol (Protocol A) for K rounds, then there exists a round k of the protocol such that the transcript $\pi^{1:T,k}$ at round k has $(q, f, \mathcal{H}_A \cup \mathcal{H}_B)$ -distance to swap regret, where*

$$q = \frac{T}{2}(g_A(T) + g_B(T))$$

and

$$f = 8T \left(\frac{\beta(T, f'_A, f'_B) + 1/K}{2} \right)^{\frac{1}{3}} + \frac{11}{2}T\beta(T, f_A, f_B)$$

We have thus established that there must be a sequence of predictions corresponding to *some* round in the collaboration protocol which we can apply our boosting theorem to. However, this will not necessarily be the final round, and so the accuracy guarantees that we get from our boosting theorem will not necessarily apply to the final sequence of predictions. We show in the following theorem (Theorem C.3) that, while the final sequence of predictions do not necessarily have swap regret guarantees with respect to \mathcal{H}_A and \mathcal{H}_B , it nevertheless has external regret guarantees with respect to \mathcal{H}_J , the joint function class.

Theorem C.3. *Let \mathcal{H}_J be a hypothesis class over the joint feature space \mathcal{X} . Let $\mathcal{H}_A = \{h_A : \mathcal{X}_1 \rightarrow \mathcal{Y}\}$ and $\mathcal{H}_B = \{h_B : \mathcal{X}_2 \rightarrow \mathcal{Y}\}$ be hypothesis classes over \mathcal{X}_A and \mathcal{X}_B respectively. Consider some transcript $\pi^{1:T,1:K}$ generated via the Collaboration Protocol (Protocol A) between Alice and Bob over K rounds. If:*

- Alice has $(f_A, g_A, \mathcal{H}_A)$ -conversation swap regret
- Bob has $(f_B, g_B, \mathcal{H}_B)$ -conversation swap regret
- \mathcal{H}_A and \mathcal{H}_B jointly satisfy the $w(\cdot)$ -weak learning condition with respect to \mathcal{H}_J

The transcript $\pi^{1:T,K}$ on the last round K satisfies:

$$\sum_{t=1}^T (\hat{y}^{t,k} - y^t)^2 - \min_{h_J \in \mathcal{H}_J} \sum_{t=1}^T (h_J(x^t) - y^t)^2 \leq 2Tw^{-1} \left(8 \left(\frac{\beta(T, f'_A, f'_B) + 1/K}{2} \right)^{\frac{1}{3}} + \frac{11}{2}\beta(T, f_A, f_B) \right) + \frac{3}{2}T(g_A(T) + g_B(T)) + 3K\beta(T, f'_A, f'_B)$$

To prove this theorem, we apply our boosting theorem (Theorem B.3) to the round k identified in Theorem C.2, which establishes an external regret guarantee with respect to \mathcal{H}_J for the predictions made at round k . We then show that the swap regret conditions we assume of Alice and Bob also imply that the squared error cannot substantially increase at any subsequent round, which allows us to conclude that the error of our predictions at the final round K is not much larger than it is at the round k at which our boosting theorem applied. External regret (unlike swap regret) is monotone in the squared error of our predictions, which thus allows us to conclude that our final predictions satisfy the claimed external regret bound with respect to \mathcal{H}_J .

C.1 Reducing Conversation Swap Regret to External Regret

We have now established that two agents, engaging in our collaboration protocol, will arrive at predictions that have no external regret to \mathcal{H}_J if their predictions have no conversation swap regret with respect to classes \mathcal{H}_A and \mathcal{H}_B respectively. We now turn to reducing the algorithmic problem of engaging in our collaboration protocol with conversation swap regret guarantees with respect to a hypothesis class \mathcal{H} to the well studied problem of making predictions in an adversarial environment that simply have no *external* regret with respect to \mathcal{H} . Garg et al. [2024] give a generic reduction that efficiently transforms an algorithm guaranteeing no external regret with respect to \mathcal{H} into an algorithm that guarantees no *swap* regret with respect to \mathcal{H} . We in turn show how to transform any

algorithm guaranteeing no swap regret with respect to \mathcal{H} into one that can engage in a collaboration protocol and guarantee no *conversation* swap regret with respect to \mathcal{H} . Collina et al. [2025] use a similar reduction from conversation calibration to calibration. Together, this gives an efficient reduction from the problem of interacting with collaboration protocol A with no conversation swap regret guarantees (what is needed to invoke Theorem C.3) to the problem of making no (external) regret predictions. As we will see, whenever we start with an algorithm that guarantees sublinear external regret rates, we obtain an algorithm that guarantees sublinear conversation swap regret rates.

We begin by quoting the result of Garg et al. [2024].

Theorem C.4 (Theorem 3.1 of Garg et al. [2024]). *Fix a hypothesis class \mathcal{H} . If:*

- All $h \in \mathcal{H}$ satisfy $h(x)^2 \leq B$ for all $x \in \mathcal{X}$
- \mathcal{H} has finite sequential fat-shattering dimension (Definition K.11)
- There exists an efficient online algorithm producing predictions $\hat{y}^1, \dots, \hat{y}^T$ that achieve, for any sequence of outcomes y^1, \dots, y^T , external regret with respect to \mathcal{H} bounded by $r(T)$, i.e.:

$$\sum_{t=1}^T (\hat{y}^t - y^t)^2 - \min_{h \in \mathcal{H}} \sum_{t=1}^T (h(x^t) - y^t)^2 \leq r(T)$$

where $r(T)$ is a concave function.

Then, for any $m > 0$, there exists an efficient online algorithm which, with probability $1 - \rho$, guarantees (f, \mathcal{H}) -swap regret, where

$$f(T) \leq m \cdot r\left(\frac{T}{m}\right) + \frac{3T}{m} + m + \max(8B, 2\sqrt{B}) \cdot m \cdot C_{\mathcal{H}} \cdot \sqrt{T \log\left(\frac{4m}{\rho}\right)}$$

Here, $C_{\mathcal{H}}$ is a constant that depends on the sequential fat-shattering dimension of \mathcal{H} .

[ht] **Input** External regret algorithm M_0 , hypothesis class \mathcal{H} , bucketing function g

Let M be the swap regret algorithm given by Theorem C.4, when initiated with M_0 .

For every odd $k \in \{3, \dots, K\}$ and bucket $i \in \{1, \dots, 1/g(T)\}$, instantiate a copy of M , called $M_{k,i}$.
For the first round $k = 1$, instantiate a copy of M , called M_1 .

Let $\pi^{1:t,k|i}$ denote the transcript on round k up until day t , restricted to $\{t : \hat{y}^{t,k-1} \in [(i-1)g(T), ig(T))\}$, the subsequence where the previously communicated predictions falls into bucket i .

Let $M(\pi^{1:t,k|i}, \mathcal{H})$ denote the output of M given this transcript and hypothesis class \mathcal{H} .

each day $t = 1, \dots, T$ Receive x_A^t Make prediction $\hat{y}_A^{t,1} = M_1(\pi^{1:t-1,1}, \mathcal{H})$ Send to Bob $\hat{y}_A^{t,1}$ each odd round $k = 3, 5, \dots, K$ Observe Bob's prediction from the previous round $\hat{y}_B^{t,k-1}$ and let i be an integer such that $\hat{y}_B^{t,k-1} \in [(i-1)g(T), ig(T))$. Make prediction $\hat{y}_A^{t,k} = M_{k,i}(\pi^{1:t-1,k|i}, \mathcal{H})$ Send to Bob $\hat{y}_A^{t,k}$ Observe $y^t \in \mathcal{Y}$.

We formalize our reduction from conversation swap regret to external regret in Algorithm C.1 and prove its correctness in Theorem C.5. We state the algorithm from the perspective of Alice; Bob's is symmetric.

Theorem C.5. *Fix a hypothesis class \mathcal{H} . If:*

- All $h \in \mathcal{H}$ satisfy $h(x)^2 \leq B$ for all $x \in \mathcal{X}$
- \mathcal{H} has finite sequential fat-shattering dimension
- There exists an efficient online algorithm guaranteeing external regret with respect to \mathcal{H} bounded by $r(T)$ where $r(T)$ is a concave function.

Then, for any $m > 0$ and bucketing function g , Algorithm C.1 guarantees, with probability $1 - \rho$ (over the internal randomness of the algorithm), (f, g, \mathcal{H}) -conversation swap regret for:

$$f(|T(k-1, i)|) \leq m \cdot r\left(\frac{|T(k-1, i)|}{m}\right) + \frac{3|T(k-1, i)|}{m} + m + \max(8B, 2\sqrt{B}) \cdot m \cdot C_{\mathcal{H}} \cdot \sqrt{|T(k-1, i)| \log\left(\frac{2mK}{g(T)\rho}\right)}$$

where $T(k-1, i)$ is the subsequence of days where the predictions of round $k-1$ fall into bucket i and $C_{\mathcal{H}}$ is a constant that depends on the sequential fat-shattering dimension of \mathcal{H} .

C.2 End-to-End Results

Now we are able to state our end-to-end reduction which starts with algorithms with external regret guarantees to \mathcal{H}_A and \mathcal{H}_B respectively and instantiates a collaboration protocol with external regret guarantees to \mathcal{H}_J . In Theorem C.6, we show that as long as the external regret bounds we start with are sublinear in T and the number of rounds K that parameterize the collaboration protocol grows sublinearly with T (but is not constant), we obtain sublinear regret guarantees with respect to \mathcal{H}_J .

Theorem C.6. Fix any tuple of hypothesis classes $\mathcal{H}_A, \mathcal{H}_B$, and \mathcal{H}_J . If:

- All $h \in \mathcal{H}_A$ and $h \in \mathcal{H}_B$ satisfy $h(x)^2 \leq B$ for some constant B , for all $x \in \mathcal{X}$.
- \mathcal{H}_A and \mathcal{H}_B have finite sequential fat-shattering dimension
- There exists an efficient online algorithm guaranteeing external regret with respect to \mathcal{H}_A bounded by $r_A(T)$, and there exists an efficient online algorithm achieving external regret with respect to \mathcal{H}_B bounded by $r_B(T)$, where $r_A(T) \leq \tilde{O}(T^{\alpha_A})$ and $r_B(T) \leq \tilde{O}(T^{\alpha_B})$, $\alpha_1, \alpha_2 \in (0, 1)$, are sublinear in T
- \mathcal{H}_A and \mathcal{H}_B jointly satisfy the $w(\cdot)$ -weak learning condition with respect to \mathcal{H}_J

Then, there is an efficient online algorithm such that if Alice and Bob both use the algorithm to interact in the Collaboration Protocol (Protocol A), then the transcript $\pi^{1:T,K}$ at the last round K satisfies, with probability $1 - \rho$:

$$\begin{aligned} & \sum_{t=1}^T (\hat{y}^{t,K} - y^t)^2 - \min_{h_J \in \mathcal{H}_J} \sum_{t=1}^T (h_J(x^t) - y^t)^2 \\ & \leq 2Tw^{-1} \left(\tilde{O} \left(T^{-\alpha'} \sqrt{\log \left(\frac{K}{\rho} \right) + \frac{1}{K^{1/3}}} \right) \right) + \tilde{O} \left(KT^{1-\alpha''} \log^{1/4} \left(\frac{K}{\rho} \right) \right) + O(T^\alpha) \end{aligned}$$

for some constants $\alpha, \alpha', \alpha'' \in (0, 1)$.

Moreover, if $K = \omega(1)$ and $K = o(T^{\alpha''})$, then the transcript $\pi^{1:T,K}$ satisfies, with probability $1 - \rho$:

$$\sum_{t=1}^T (\hat{y}^{t,K} - y^t)^2 - \min_{h_J \in \mathcal{H}_J} \sum_{t=1}^T (h_J(x^t) - y^t)^2 \leq \tilde{O} \left(T^{\alpha'''} \log^{1/4} \left(\frac{1}{\rho} \right) \right) + o(T)$$

for some constant $\alpha''' \in (0, 1)$ and T sufficiently large (larger than a constant that depends on w, α_A, α_B , and ρ). Here, $o(T)$ is a sublinear term that depends on w, K, α_A , and α_B . That is, the transcript on the last round achieves sublinear regret.

Remark C.7. Observe that Theorem C.6 allows us to trade off K , the parameter controlling the length of the conversation at each day in our collaboration protocol, with the final regret bound. Increasing K can improve the regret bound, at the cost of increasing the amount of daily communication and computation. There is a range of choices of K , growing with T , that guarantee regret that grows only sublinearly with T . The algorithm itself is an efficient reduction to the external regret algorithms for \mathcal{H}_A and \mathcal{H}_B that we start with.

Finally, we derive concrete regret bounds when $\mathcal{H}_A, \mathcal{H}_B$, and \mathcal{H} are norm-bounded linear functions over the domains $\mathcal{X}_A, \mathcal{X}_B \subseteq \mathbb{R}^d$ and $\mathcal{X}_J \subseteq \mathbb{R}^{2d}$ respectively (recall that these classes satisfy the weak learning condition). First, for linear functions there indeed exists an efficient algorithm due to Azoury and Warmuth [1999] that achieves diminishing external regret — and thus conversation swap regret — and so we can apply our reductions to get worst-case polynomial-time algorithms to interact in our collaboration protocol.

Theorem C.8. [Azoury and Warmuth, 1999] There exists an efficient online algorithm producing predictions such that for $x^t \in \mathbb{R}^d, \|x^t\|_2 \leq 1$ and for all parameter vectors $\theta \in \mathbb{R}^d$:

$$\sum_{t=1}^T (\hat{y}^t - y^t)^2 - \sum_{t=1}^T (\langle \theta, x^t \rangle, y^t)^2 \leq 2d \ln(T+1) + \|\theta\|^2$$

Recall that norm-bounded linear functions satisfy the weak learning condition with margin $w(\gamma) = \Omega(\gamma^2)$ (Theorem B.6). Together with the conversation swap regret rates we have just derived, we can instantiate Theorem C.3 for norm-bounded linear functions on the joint feature space. Our result is Theorem C.9.

Theorem C.9. *Let $\mathcal{X}_A = \mathcal{X}_B = \{x \in \mathbb{R}^d : \|x\|_2 \leq 1\}$. Let $\mathcal{H}_A = \{x_A \mapsto \langle \theta, x_A \rangle : \|\theta\|_2 \leq C\}$ and $\mathcal{H}_B = \{x_B \mapsto \langle \theta, x_B \rangle : \|\theta\|_2 \leq C\}$ be the sets of all linear functions with bounded norm over \mathcal{X}_1 and \mathcal{X}_2 respectively, for $C \geq 1/2$. Let $\mathcal{H}_J = \{h_A + h_B : h_A \in \mathcal{H}_A, h_B \in \mathcal{H}_B\}$ be the Minowski sum of \mathcal{H}_A and \mathcal{H}_B . Consider some transcript $\pi^{1:T, 1:K}$ generated via the Collaboration Protocol between Alice and Bob over K rounds (Protocol A). There exists an online algorithm (Algorithm C.1, instantiated with the algorithm of Theorem C.8) such that the transcript $\pi^{1:T, K}$ at the last round K satisfies, with probability $1 - \rho$:*

$$\sum_{t=1}^T (\hat{y}^{t,K} - y^t)^2 - \min_{h_J \in \mathcal{H}_J} \sum_{t=1}^T (h_J(x^t) - y^t)^2 \leq \tilde{O} \left(T^{47/48} \sqrt{\max(C^2, C) d \log \left(\frac{KT^{1/8}}{\rho} \right)} + TK^{-\frac{1}{6}} + KT^{\frac{7}{8}} \sqrt{\max(C^2, C) d \log \left(\frac{KT^{1/8}}{\rho} \right)} \right)$$

Remark C.10. *By setting $K = T^{\frac{3}{28}}$, the external regret is sublinear:*

$$\begin{aligned} \tilde{O} \left(T^{47/48} \sqrt{\max(C^2, C) d \log \left(\frac{T^{15/56}}{\rho} \right)} + T^{55/56} + T^{55/56} \sqrt{\max(C^2, C) d \log \left(\frac{T^{15/56}}{\rho} \right)} \right) \\ = \tilde{O} \left(T^{55/56} \sqrt{\max(C^2, C) d \log \left(\frac{T^{15/56}}{\rho} \right)} \right) \end{aligned}$$

C.3 A Decision Theoretic Extension for Higher Dimensional Outcome Spaces

In Appendix H we extend our results in the online setting to high dimensional outcome spaces. Here we give an overview of our techniques. Now the outcome space $\mathcal{Y} \subseteq [0, 1]^d$ is d dimensional, and we model a decision maker with a finite action space \mathcal{A} and a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$ that maps an action and an outcome to a utility. The natural extension of our one-dimensional solution to a d -dimensional outcome space—by asking for swap regret with respect to outcome predictions themselves—would inherit exponential dependencies on d . We circumvent this difficulty by not communicating predictions \hat{y} of the outcome $y \in \mathcal{Y}$ itself. Instead, in each round k , parties produce predictions $\hat{y}^{t,k} \in \mathcal{Y}$ but communicate only actions $a^{t,k} \in \mathcal{A}$ that are utility maximizing given their predictions: $a^{t,k} = \text{BR}_u(\hat{y}^{t,k})$ where the best response function is defined as:

$$\text{BR}_u(\hat{y}) = \arg \max_{a \in \mathcal{A}} u(a, \hat{y})$$

We use a definition of decision calibration sufficient to guarantee swap regret of the best response actions first used by Noarov et al. [2023], generalizing the original definition given by Zhao et al. [2021] (the definition from Zhao et al. [2021] does imply swap regret bounds).

Definition C.11 (Decision Calibration (Definition H.7)). *Fix an action space \mathcal{A} and a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$. A sequence of outcome predictions $\{\hat{y}^{1,k}, \dots, \hat{y}^{T,k}\}$ is decision calibrated if for every action $a \in \mathcal{A}$:*

$$\left\| \sum_{t=1}^T \mathbb{1}[\text{BR}_u(\hat{y}^{t,k}) = a] (\hat{y}^{t,k} - y) \right\| = 0$$

We also use a definition of decision *cross-calibration* first used by Lu et al. [2025]:

Definition C.12 (Decision Cross Calibration (Definition H.8)). *Fix an action space \mathcal{A} , a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$, and a class of benchmark policies \mathcal{C} containing functions $c : \mathcal{X} \rightarrow \mathcal{A}$ mapping contexts to actions. A sequence of outcome predictions $\{\hat{y}^{1,k}, \dots, \hat{y}^{T,k}\}$ is decision cross-calibrated with respect to \mathcal{C} if for every pair of actions $a, a' \in \mathcal{A}$ and for every $c \in \mathcal{C}$:*

$$\left\| \sum_{t=1}^T \mathbb{1}[\text{BR}_u(\hat{y}^{t,k}) = a] \mathbb{1}[c(x^t) = a'] (\hat{y}^{t,k} - y) \right\| = 0$$

If a sequence of predictions $\{\hat{y}^{1,k}, \dots, \hat{y}^{T,k}\}$ is simultaneously decision calibrated and decision cross-calibrated with respect to \mathcal{C} , then the corresponding sequence of actions $a^{t,k} = \text{BR}_u(\hat{y}^{t,k})$ have no swap regret with respect to \mathcal{C} — i.e. for every $c \in \mathcal{C}$ and for every $a \in \mathcal{A}$:

$$\sum_{t=1}^T \mathbb{1}[a^{t,k} = a] u(a^{t,k}, y^t) \geq \max_{c \in \mathcal{C}} \sum_{t=1}^T \mathbb{1}[a^{t,k} = a] u(c(x^t), y^t)$$

We ask that both Alice and Bob are decision calibrated and decision cross calibrated conditional on the action that the other communicated at the previous round — which implies that both parties have no conversation swap regret with respect to \mathcal{C}_A and \mathcal{C}_B respectively on their own rounds. It also allows us to invoke a fast agreement theorem from Collina et al. [2025] which lets us establish fast convergence to a round of predicted actions that simultaneously has no swap regret to \mathcal{C}_A and \mathcal{C}_B . This lets us apply a similar boosting theorem to the one we develop in Section B to establish that the final sequence of actions $a^{1,K}, \dots, a^{T,K}$ that result from the collaboration protocol have no regret with respect to a collection \mathcal{C}_J of action policies defined on the joint feature space.

Theorem C.13 (Informal statement of Theorem H.26). *Fix any triple of policy classes $\mathcal{C}_A, \mathcal{C}_B$, and \mathcal{C}_J . If \mathcal{C}_A and \mathcal{C}_B satisfy the weak learning condition with respect to \mathcal{C}_J , and the conversation length K is sublinear in T (but not constant), then there is an efficient collaboration protocol such that:*

$$\max_{c_J \in \mathcal{C}_J} \sum_{t=1}^T u(c_J(x^t), y^t) - \sum_{t=1}^T u(a^{t,K}, y^t) \leq o(T)$$

D Lower Bounds: Necessity of Interaction, Weak Learning and Swap Regret

Next we provide qualitative lower bounds to motivate the design choices in our collaborative learning protocols. We demonstrate the necessity of interaction between parties, the necessity of a condition like our weak learning assumption for achieving information aggregation guarantees, and the necessity of using a stronger criterion than external regret (like swap regret) within the protocol.

Interaction is Necessary. One might wonder if interaction is necessary, especially when the underlying function classes are simple, like linear functions, which satisfy our weak learning condition (Theorem B.6). Perhaps some non-adaptive combination of the optimal *linear* predictors $h_A^*(x_A)$ and $h_B^*(x_B)$ is sufficient to achieve performance competitive with the optimal joint linear predictor $h_J^*(x)$. The following example, adapted from the proof of Theorem B.7, shows this is not the case. It demonstrates that even when the Bayes optimal predictors are themselves linear for Alice, Bob, and the joint feature space, the information required for optimal joint prediction might not be recoverable just by combining the optimal individual linear predictors.

Theorem D.1 (Interaction Necessity for Linear Functions). *There exists a joint distribution \mathcal{D} over $\mathcal{X}_A \times \mathcal{X}_B \times \mathcal{Y}$ and classes $\mathcal{H}_A, \mathcal{H}_B, \mathcal{H}_J$ corresponding to (bounded) linear functions over $\mathcal{X}_A, \mathcal{X}_B$, and $\mathcal{X} = \mathcal{X}_A \times \mathcal{X}_B$ respectively, such that for any $f : \mathcal{Y} \times \mathcal{Y} \rightarrow \mathcal{Y}$,*

$$\mathbb{E}_{\mathcal{D}} [(f(h_A^*(x_A), h_B^*(x_B)) - y)^2] > \min_{h_J \in \mathcal{H}_J} \mathbb{E}_{\mathcal{D}} [(h_J(x) - y)^2],$$

where h_A^*, h_B^* are the optimal linear predictors in $\mathcal{H}_A, \mathcal{H}_B$ respectively.

The proof uses a similar construction as Theorem B.7 where the label has no correlation with Alice's features, and weak correlation with Bob's feature. However, subtracting Alice's features from Bob's features gives the signal exactly. Since the optimal predictor on Alice's features is 0 (no correlation to the label), there is no way for any aggregation function to subtract Alice's features from Bob's to get the performance of the joint predictor.

The Weak Learning Condition is Necessary for Boosting. Our boosting result (Theorem B.3) shows that if \mathcal{H}_A and \mathcal{H}_B satisfy the weak learning condition with respect to \mathcal{H}_J , then achieving low swap regret with respect to $\mathcal{H}_A \cup \mathcal{H}_B$ implies low external regret with respect to \mathcal{H}_J . We now show this condition is necessary: if a triple $(\mathcal{H}_A, \mathcal{H}_B, \mathcal{H}_J)$ fails the weak learning condition, there exist distributions and prediction sequences with no swap regret to $\mathcal{H}_A \cup \mathcal{H}_B$ but positive external regret to \mathcal{H}_J .

Theorem D.2 (Necessity of Weak Learning for Boosting). *For any triple of function classes $(\mathcal{H}_A, \mathcal{H}_B, \mathcal{H}_J)$ that fails to satisfy the $w(\cdot)$ -weak learning condition (Definition B.1) for any strictly increasing function w , there exists a sequence of examples $(x_A^t, x_B^t, y^t)_{t=1}^T$ and predictions $\hat{y}^{1:T}$ such that, as $T \rightarrow \infty$, the sequence $\hat{y}^{1:T}$ has 0 swap regret with respect to \mathcal{H}_A and \mathcal{H}_B , but has positive external regret with respect to \mathcal{H}_J .*

The proof follows from observing that if the weak-learning condition is not satisfied for any $w(\cdot)$ then there is a distribution such that the joint predictor gets a non-zero gain over the constant predictor but both Alice and Bob do not improve over the constant predictor. Now predicting according to the best constant predictor guarantees no swap-regret to either Alice or Bob, but has non-zero external regret to the joint predictor, since there is a joint predictor better than the constant predictor on the distribution.

Weak Learning is Weaker than Information Substitutes. We show that our weak learning condition is strictly weaker than the “information substitutes” condition studied by Frongillo et al. [2023]. The concept of information substitutes, in the context of Bayesian agreement, fundamentally concerns the diminishing marginal value of information. When applied to predictors, it says that the improvement gained by adding Bob’s information (or signal) is smaller if the Alice’s information is already available, and vice-versa.

To translate this concept for comparing function classes $(\mathcal{H}_A, \mathcal{H}_B, \mathcal{H}_J)$, we need a measure of the “value” provided by each parties features when used by functions in one of these classes. In prediction tasks with squared error loss, a natural measure of value is the reduction in expected squared error compared to a baseline constant predictor. This gives us the following condition:

Definition D.3 (Information Substitutes for function classes). *Let $\mathcal{H}_A : \mathcal{X}_A \rightarrow \mathcal{Y}$ and $\mathcal{H}_B : \mathcal{X}_B \rightarrow \mathcal{Y}$ be hypothesis classes for Alice and Bob, respectively, and let \mathcal{H}_J be a hypothesis class over the joint features $\mathcal{H}_J : \mathcal{X}_A \times \mathcal{X}_B \rightarrow \mathcal{Y}$. We say model classes \mathcal{H}_A and \mathcal{H}_B satisfy information substitutes with respect to \mathcal{H}_J if, for all distributions \mathcal{D} ,*

$$\min_{h_A \in \mathcal{H}_A} \mathbb{E}[(h_A(x) - y)^2] - \min_{h_J \in \mathcal{H}_J} \mathbb{E}[(h_J(x) - y)^2] \leq \min_{c \in \mathbb{R}} \mathbb{E}[(c - y)^2] - \min_{h_B \in \mathcal{H}_B} \mathbb{E}[(h_B(x) - y)^2]$$

Information substitutes, as defined here, imposes a stronger, quantitative relationship on the magnitudes of the maximum achievable gains compared to weak-learning which asks if any positive gain with the joint features implies some positive gain for either individual class.

Lemma D.4. *If model classes \mathcal{H}_A and \mathcal{H}_B satisfy information substitutes with respect to \mathcal{H}_J , they also jointly satisfy the $w(\cdot)$ -weak learning condition with respect to \mathcal{H}_J for $w(\gamma) = \gamma/2$.*

Combining this with Theorem B.8 gives us that our weak-learning condition is significantly weaker than the information substitutes condition.

Corollary D.5. *There exists $\mathcal{H}_A, \mathcal{H}_B, \mathcal{H}_J$ that satisfy the $w(\cdot)$ -weak learnability condition for $w(\gamma) = \Theta(\gamma^2)$ but do not satisfy the information substitutes condition. In fact, the class of bounded linear functions over $\mathcal{X}_A = \mathcal{X}_B = [-1, 1]$ witnesses this gap.*

External Regret is Insufficient. Our protocol aims to produce predictions p that have low swap regret with respect to \mathcal{H}_A and \mathcal{H}_B . One might ask if the weaker condition of low external regret would suffice. That is, if p has low external regret to \mathcal{H}_A and low external regret to \mathcal{H}_B , does it follow (under the weak learning condition) that p has low external regret to \mathcal{H}_J ? The following example shows the answer is no, even for linear functions where the weak learning condition holds.

Theorem D.6. *There exists a joint distribution \mathcal{D} over $\mathcal{X}_A \times \mathcal{X}_B \times \mathcal{Y}$ and classes $\mathcal{H}_A, \mathcal{H}_B, \mathcal{H}_J$ corresponding to linear functions over $\mathcal{X}_A, \mathcal{X}_B$, and $\mathcal{X}_A \times \mathcal{X}_B$, respectively, such that there exists a sequence of examples $(x_A^t, x_B^t, y^t)_{t=1}^T$ and predictions $\hat{y}^{1:T}$ such that, as $T \rightarrow \infty$, the sequence $\hat{y}^{1:T}$ has no external regret to \mathcal{H}_A and \mathcal{H}_B , but has positive external regret with respect to \mathcal{H}_J .*

E Lifting to the One-Shot Bayesian Setting

Our paper primarily concerns itself with information aggregation in frequentist settings — both the online adversarial setting studied in Section C (and Appendix H) in which there is no distribution at all, and the batch setting studied in Appendix I in which there is a distribution, but the learners have

no prior knowledge of it except through a training sample. However, the theorems we prove can be lifted to the one-shot Bayesian setting studied by Aumann [1976], Aaronson [2005], Frongillo et al. [2023], which extends and generalizes the information aggregation result from Frongillo et al. [2023] in the original setting of Aumann’s agreement theorem. We generalize the information aggregation theorem of Frongillo et al. [2023] in two ways: first, our weak learning condition is strictly weaker than the information substitutes condition given by Frongillo et al. [2023] — for example, as we have shown, our weak learning condition is satisfied by *linear* functions, whereas the information substitutes condition is not (as we demonstrate in Section D). Second, our information aggregation theorems are agnostic in the sense that we can guarantee that *independently of the prior distribution*, Bayesians with a common prior must agree on predictions that are as accurate as the best model on their joint feature space in any hypothesis class with bounded fat shattering dimension, so long as the hypothesis class satisfies our weak learning assumption. In contrast Frongillo et al. [2023] apply their information substitutes condition only to the Bayes optimal predictors on \mathcal{X}_A , \mathcal{X}_B , and \mathcal{X} respectively.

Rather than the online adversarial setting we study in Sections C and H, we assume (as we do in Section I) that instances are drawn from \mathcal{D} : $(x_A, x_B, y) \sim \mathcal{D}$, where \mathcal{D} is a joint distribution over $\mathcal{X}_A \times \mathcal{X}_B \times \mathcal{Y}$. However, unlike in Section I we now assume that this distribution is known to both Alice and Bob as their (common) prior distribution. We now model Alice and Bob as perfect Bayesians, who at each round of conversation, form a posterior distribution conditional on all of their observations thus far (both the features visible to them and the transcript of the conversation so far) and communicate their posterior expectation of y . For simplicity, rather than communicating these expectations to arbitrary precision, Alice and Bob communicate expectations rounded to multiples of some discretization parameter $m \in \mathbb{N}$ (which guarantees among other things that the communication requires only a bounded number of bits). Let $\lceil \frac{1}{m} \rceil$ represent the discretization of the unit interval into m grid points: $\{0, \frac{1}{m}, \frac{2}{m}, \dots, 1\}$. We denote a prediction \hat{y} that is rounded to the nearest multiple of $\frac{1}{m}$ as \bar{y} .

Definition E.1 (Bayesian Learner). *Fix a joint distribution $\mathcal{D} \in \Delta(\mathcal{X}_A \times \mathcal{X}_B \times \mathcal{Y})$ over features observable to Alice, features observable to Bob, and labels. We say that Alice (resp., Bob) is a Bayesian Learner if for all $t, k > 0$, given observable features x_A^t , prediction transcript $\pi^{1:t-1}$, and conversation $C_{1:k-1}^t$, they make a prediction as*

$$\hat{y}_A^{t,k} = \mathbb{E}_{\mathcal{D}}[Y | x_A^t, \pi^{1:t-1}, C_{1:k-1}^t].$$

[ht] **Input** $(\mathcal{D}, \mathcal{Y}, K)$ each day $t = 1, \dots$ Receive $x^t = (x_A^t, x_B^t, y^t) \sim \mathcal{D}$. Alice sees x_A^t and Bob sees x_B^t . each round $k = 1, 2, \dots, K$ k is odd Alice predicts $\hat{y}_A^{t,k} \in \mathcal{Y}$, and sends Bob $\bar{y}_A^{t,k}$ (i.e. the rounded version of $\hat{y}_A^{t,k}$) k is even Bob predicts $\hat{y}_B^{t,k}$, and sends Alice $\bar{y}_B^{t,k}$. Alice and Bob observe $y^t \in \mathcal{Y}$.

Our argument will proceed as follows:

1. First, we observe that the predictions of a Bayesian are always unbiased at the time they are made. Among other things, this implies that a Bayesian always has no *expected* conversation swap regret with respect to *any* benchmark policy.
2. A consequence of this is that a Bayesian’s average realized conversation swap regret tends to zero as the number of days of interaction tends to infinity, for any benchmark class for which the realized squared error uniformly converges to the expected squared error with sufficiently many samples. This is the case for any benchmark class of policies with finite fat shattering dimension [Anthony and Bartlett, 1999].
3. Thus, if we imagine sampling T instances $(x_A, x_B, y) \sim \mathcal{D}$ from the prior distribution and two Bayesians collaborating on these instances, in the limit as $T \rightarrow \infty$, we can apply our information aggregation theorems with respect to any benchmark class that satisfies our weak learning condition and has bounded fat shattering dimension to bound the expected squared error of the final predictions.
4. Finally, we observe that since the examples are drawn i.i.d. and Bayesians will not condition on the history of past instances (as they are independent from the current instance), the distribution on the sequence of interactions is permutation invariant. Thus we can bound the expected squared error of the prediction arrived at for the *first* example, and hence our theorems apply even when $T = 1$.

The broad strokes of this proof strategy mirror how Collina et al. [2025] lifted their sequential agreement theorems to the one-shot Bayesian setting. Since we aim for the stronger goal of information aggregation, we must now reason about swap regret with respect to an infinite benchmark class (rather than simple calibration).

E.1 Bayesians and Conversation Swap Regret

We first want to establish that Bayesians will have low conversation swap regret (Definition A.7) when they participate in a sequential collaboration protocol (Protocol E). Then, in the following section, we can proceed by instantiating Theorem C.3. In fact, Bayesians always have zero *expected* swap regret with respect to any fixed class of benchmark functions. To bound their realized swap regret, we need to uniformly bound the loss with respect to its expectation across every function in the benchmark class \mathcal{H}_B , which is the step that causes us to require that \mathcal{H}_B has bounded fat shattering dimension.

Theorem E.2. Fix $\delta, \epsilon, m > 0$. Suppose the fat shattering dimension of \mathcal{H}_A is finite at any scale ϵ . Fix transcript $\pi^{1:T, 1:K}$. Let v range over values in $[\frac{1}{m}]$ and $g_B(T)$ be some bucketing. If Alice is a Bayesian learner with discretization m , with probability $1 - \delta$, Alice’s sequence of predictions resulting from Protocol E has low conversation swap regret with respect to bucketing $g_B(T)$ and function class \mathcal{H}_A : for all odd rounds k and buckets $i \in \{1, \dots, \frac{1}{g_B(T)}\}$ such that $\mathbb{P}[\bar{y}_B \in B_i] > 0$, if $|T_B(k-1, i)| > \frac{C_{\mathcal{H}}^{\epsilon/256} \ln(\frac{1}{\epsilon}) + \ln(\frac{1}{\delta})}{\epsilon^2}$:

$$\begin{aligned} & \sum_{t \in T_B(k-1, i)} (\bar{y}^{t,k} - y^t)^2 - \sum_v \min_{h \in \mathcal{H}_A} \sum_{t \in T_B(k-1, i)} \mathbb{I}[\bar{y}^{t,k} = v] (h(x^t) - y^t)^2 \\ & \leq 2\sqrt{2T \ln \frac{g_B(T)K}{\delta}} + \frac{T}{m^2} + mT\epsilon, \end{aligned}$$

where $T_B(k-1, i) = \{t \mid \bar{y}_B^{t,k-1} \in B_i(1/g_B(T))\}$ is the subsequence of days where the predictions of Bob at the previous round fall in bucket i , $C_{\mathcal{H}_A}^{\epsilon}$ is the fat shattering dimension of \mathcal{H}_A at scale ϵ , and K is the number of rounds on each day. A symmetric condition holds for Bob.

Before proceeding to the proof of Theorem E.2, we first formalize a simple observation: if we resample the label every day after the j^{th} round of conversation from the posterior distribution on the label conditional on the transcript of interaction so far, this does not change the distribution of transcripts. This allows us to conduct all of our subsequent analysis under this resampling thought experiment.

Lemma E.3 (Lemma 6.3 of Collina et al. [2025]). Let \mathcal{D} be a probability distribution over $\mathcal{X}_m \times \mathcal{X}_h \times \mathcal{Y}$ and fix a day $t \in [T]$. Fix a transcript through day $t-1$: $\pi^{1:t-1}$.

- Consider an interaction at day t under Protocol E. Let π^t be the transcript of day t from this interaction.
- Fix an arbitrary round j . Consider an interaction when (x_A, x_B, y^t) is sampled from \mathcal{D} at the beginning of day t and then Alice and Bob correspond according to Protocol E until round j . Then, in round j , the outcome is resampled from the posterior distribution conditional on the information observed so far: $y' \sim \mathcal{D}_{\mathcal{Y}}|x_A^t, \pi^{1:t-1}, C_{1:j-1}^t, \hat{y}_A^{t,j}$, where $\mathcal{D}_{\mathcal{Y}}$ is the marginal distribution on \mathcal{Y} . Let $\bar{\pi}_j^t$ be the transcript of day t from this interaction, with y^t replaced with y' .

For all rounds k ,

$$\mathbb{P}_{\mathcal{D}}[\pi^{t, 1:k}] = \mathbb{P}_{\mathcal{D}}[\bar{\pi}_j^{t, 1:k}].$$

Proofs in this section are deferred to Appendix M.

Now, we analyze the expected conversation swap regret of Alice and Bob. Recall that in the definition of swap regret (Definition A.7), we compare the squared error of Alice’s (or Bob’s) predictions to the predictions of the best comparator function in the benchmark class, separately for each level set of their prediction. Here, since predictions are restricted to the discretization $[\frac{1}{m}]$, we have m level sets.

1093 We first want to argue that for any possible swap function (i.e. selection of m functions from \mathcal{H}_A ,
 1094 one for each levelset), Alice’s expected swap regret is small.

1095 **Lemma E.4.** *Fix some bucketing function $g_B(\cdot)$. If Alice is a Bayesian as in Protocol E, she has*
 1096 *low expected conversation swap regret with respect to any fixed swap function $\{h_0, h_{\frac{1}{m}}, \dots, h_1\} \in$*
 1097 *$\{\mathcal{H}_A\}^m$, where h_v is the function she compares to her prediction $v \in [\frac{1}{m}]$. For all odd rounds k and*
 1098 *buckets $i \in \{1, \dots, \frac{1}{g_B(T)}\}$:*

$$\max_{\{h_0, h_{\frac{1}{m}}, \dots, h_1\} \in \{\mathcal{H}_A\}^m} \mathbb{E}_{\mathcal{D}} \left[(\bar{y}_A^{t,k} - y^t)^2 - \sum_v \mathbb{I}[\bar{y}_A^{t,k} = v] (h_v(x^t) - y^t)^2 \right] \leq \frac{1}{m^2}.$$

1099 Having established that Bayesians have low *expected* swap regret with respect to any fixed set of swap
 1100 functions, we now want to establish that they have low *realized* swap regret with high probability
 1101 over sufficiently long interactions, for large families of swap functions. We do this by applying two
 1102 concentration arguments. The first (which establishes that the realized squared error of each sequence
 1103 of predictions made by Alice and Bob are close to their expected squared error) is just an application
 1104 of Azuma’s inequality:

1105 **Lemma E.5.** *Fix $T, \delta > 0$ and bucketing $g_B(T)$. Let $\pi^{1:T, 1:K}$ be the transcript after running*
 1106 *Protocol E for T days. For all even rounds k and buckets $i \in \{1, \dots, \frac{1}{g_B(T)}\}$, with probability $1 - \delta$,*

$$\sum_{t=1}^T (\bar{y}_A^{t,k} - y^t)^2 - \mathbb{E}_{\mathcal{D}}[(\bar{y}_A^{t,k} - y^t)^2 | \pi^{1:t-1}] \leq 2\sqrt{2T \ln \frac{1}{\delta}}.$$

1107 To argue that Bayesians have low realized swap regret with respect to a (possibly infinite) benchmark
 1108 class of functions, we next need to argue that the squared error for every function in the benchmark
 1109 class (across each of the levelsets of our predictions) concentrates uniformly around its expectation.
 1110 To do this we recall the fat shattering dimension, which captures the capacity of real-valued function
 1111 classes [Anthony and Bartlett, 1999]. Full details are in Appendix M.

1112 **Lemma E.6.** *Fix $\varepsilon, \delta > 0$. Let $|T_B(k-1, i)| > \frac{C_{\mathcal{H}_A}^{\varepsilon/256} \ln(\frac{1}{\varepsilon}) + \ln(\frac{1}{\delta})}{\varepsilon^2}$, where $C_{\mathcal{H}_A}^{\varepsilon}$ is the fat shattering*
 1113 *dimension of \mathcal{H}_A at scale ε . Fix bucketing $g_B(T)$. Let $\pi^{1:T, 1:K}$ be the transcript after running*
 1114 *Protocol E for T days. For all even rounds k , buckets $i \in \{1, \dots, \frac{1}{g_B(T)}\}$ for Bob’s prediction in*
 1115 *round $k-1$, and level set $v \in [\frac{1}{m}]$ of Alice’s prediction in round k , with probability $1 - \delta$,*

$$\sup_{h \in \mathcal{H}_A} \left| \frac{1}{|T_B(k-1, i)|} \sum_{t \in T_B(k-1, i)} \mathbb{I}[\bar{y}_A^{t,k} = v] (h(x^t) - y^t)^2 - \mathbb{E}_{\mathcal{D}}[\mathbb{I}[\bar{y}_A(x) = v] (h(x) - y)^2 | \pi^{1:t-1}] \right| \leq \varepsilon.$$

1116 Finally, we can proceed to the proof of Theorem E.2, which gives a high probability bound on the
 1117 realized swap regret on the predictions made by Bayesians in Protocol E. The proof is deferred to
 1118 Appendix M, but follows directly from Lemmas E.4, E.5, and E.6.

1119 E.2 Online to One-Shot Reduction

1120 In this section, we show that if an instance is drawn from a common prior and both agents are
 1121 Bayesian, then our theorems which guarantee information aggregation with high probability on all
 1122 instances over an arbitrarily long sequence of length T hold in fact for a *single* conversation with
 1123 high probability.

1124 We can imagine an arbitrarily long sequence of conversations over many days. Each conversation on
 1125 any given day continues for exactly K rounds. We have shown that Bayesians satisfy our notion of
 1126 conversation swap regret with parameters growing sublinearly with T . In a Bayesian setting, since
 1127 instances are drawn i.i.d. from a fixed prior, Bayesians need not condition on information from prior
 1128 days. Thus, instances drawn each day (and subsequently, conversations each day) are distributed
 1129 identically. Therefore, the theorems we give which apply to the average cumulative regret over the
 1130 course of a subsequence of length T also holds in expectation over the draw from the prior, on any
 1131 single instance. Since we don’t actually need to run the protocol for T rounds to get predictions on
 1132 the first round, we can take $T \rightarrow \infty$ (as it is just a thought experiment).

1133 [ht] **Input** $\mathcal{D} \in \Delta(\mathcal{X}_h \times \mathcal{X}_m \times \mathcal{Y})$, instance for which you want information aggregation:
 1134 $(x_h^*, x_m^*, y^*) \sim \mathcal{D}$ **Parameter** number of samples: T Fix $(x_h^1, x_m^1, y^1) \sim \mathcal{D}$ For $t \in \{2, \dots, T\}$
 1135 draw $(x_h^t, x_m^t, y^t) \sim \mathcal{D}$ each day $t = 1, \dots, T$ Alice observes x_A^t and Bob observes x_B^t . each round
 1136 $k = 1, 2, \dots, L$ k is odd Alice predicts $\hat{y}_A^{t,k}$, and sends Bob $\hat{y}_A^{t,k}$ k is even Bob predicts $\hat{y}_B^{t,k} \in \mathcal{Y}$,
 1137 and sends Alice $\hat{y}_B^{t,k}$ Alice and Bob observe $y^t \in \mathcal{Y}$

1138 **Theorem E.7.** Let \mathcal{H}_J be a hypothesis class over the joint feature space \mathcal{X} . Let $\mathcal{H}_A = \{h_A : \mathcal{X}_A \rightarrow \mathcal{Y}\}$ and $\mathcal{H}_B = \{h_B : \mathcal{X}_B \rightarrow \mathcal{Y}\}$ be hypothesis classes over \mathcal{X}_A and \mathcal{X}_B . Consider instance
 1139 $(x_A, x_B, y) \sim \mathcal{D}$. If
 1140

- 1141 • Alice and Bob are both Bayesian learners
- 1142 • \mathcal{H}_A and \mathcal{H}_B have finite fat shattering dimension at every scale, and \mathcal{H}_A and \mathcal{H}_B jointly
- 1143 satisfy the $w(\cdot)$ -weak learning condition with respect to \mathcal{H}_J , for continuous $w(\cdot)$ such that
- 1144 $w(\gamma) > 0$ for all $\gamma > 0$,

1145 then, if they engage in K rounds of conversation on a single instance (x_A, x_B, y) , the prediction in
 1146 round K will have regret to the best function in \mathcal{H}_J bounded by:

$$\mathbb{E}[(\hat{y}^{1,K} - y)^2] - \min_{h_j \in \mathcal{H}_J} \mathbb{E}[(h_j(x) - y)^2] \leq O\left(w^{-1}\left(K^{-\frac{1}{3}}\right)\right).$$

1147 We can instantiate the above result for bounded norm linear functions, which satisfy our weak learning
 1148 guarantee (Theorem B.6).

1149 **Remark E.8.** If \mathcal{H}_A and \mathcal{H}_B are the classes of linear functions with bounded norm parameter
 1150 vectors: $\mathcal{H}_A = \{x_A \rightarrow \theta^T x : \|\theta\|_2 < C\}$ and $\mathcal{H}_B = \{x_B \rightarrow \theta^T x : \|\theta\|_2 < C\}$ and \mathcal{H}_J is the
 1151 Minkowski sum of \mathcal{H}_A and \mathcal{H}_B , then for an arbitrary prior distribution, when Bayesian learners
 1152 engage in a conversation of length K :

$$\mathbb{E}[(\hat{y}^{1,K} - y)^2] - \min_{h_j \in \mathcal{H}_J} \mathbb{E}[(h_j(x) - y)^2] \leq O(CK^{-\frac{1}{6}}).$$

1153 F Discussion and Future Work

1154 We present efficient protocols for collaborative information aggregation, enabling two parties with
 1155 distinct feature spaces—even if mutually illegible—to provably achieve the accuracy of joint feature
 1156 access without sharing their features. Our protocols require the two parties to operate only on their own
 1157 feature spaces and communication occurs solely through label predictions or best-response actions,
 1158 making our framework practical in modern AI systems, particularly human-AI interaction and multi-
 1159 modal settings, where challenges like privacy, data modality differences, and computational overheads
 1160 often render feature sharing impractical. Moreover, our protocols underscore the fundamental role
 1161 of interaction to achieve performance that surpasses that of the individual parties, or simple non-
 1162 interactive aggregation methods, opening up a new avenue of research in collaborative learning.

1163 Our work naturally leaves open several questions. Theoretically, extending the analysis of the weak
 1164 learning condition beyond the linear-like classes and Minkowski sum structure would broaden the
 1165 applicability of our framework to more complex function classes encountered in practice. Additionally,
 1166 our online guarantees hold against worst-case adversarial sequences, hence, exploring settings under
 1167 beyond-worst-case assumptions—for instance, leveraging models like smoothed analysis Haghtalab
 1168 et al. [2024] or incorporating mechanisms such as selective prediction Goel et al. [2023]—could
 1169 potentially yield tighter bounds and reduce the number of communication rounds.

1170 From a practical and safety perspective, the current protocols assume honest participation of both
 1171 parties. A crucial direction, particularly for human-AI collaboration, involves designing protocols
 1172 inherently robust to strategic manipulation, mitigating risks where a capable AI might deceptively
 1173 steer outcomes towards misaligned objectives. Ensuring trustworthiness in these interactions would
 1174 require designing strategy-proof protocols within our collaborative framework.

1175 Empirically evaluating the feasibility of our protocols is an important direction. While empirical
 1176 evaluations in realistic human-AI settings may be challenging, evaluations in the multi-modal setting
 1177 should be a good test ground for understanding the practical challenges of scalability, performance,
 1178 and communication efficiency, to guide further development of the framework.

G Additional Related Work

Vertically Federated Learning. Vertically federated learning (see e.g. Wei et al. [2022]) studies distributed learning problems in which features are distributed amongst parties, just as we do. The goal in this literature is to simulate learning on the shared feature space without sharing the data in the clear. Standard techniques in this literature involve running stochastic gradient descent over the full feature space over a cryptographic substrate — see e.g Hardy et al. [2017] who give an algorithm for solving logistic regression over the joint feature space using additively homomorphic encryption and Cheng et al. [2021] who give similar results for tree based models. In contrast to this line of work, our protocols require only learning on one’s own data and communicating only predictions. This is what allows us to lift our results to the Bayesian agreement setting (all of the learning conditions we need are satisfied by Bayesian reasoners), gives us protocols whose communication complexity is independent of the data dimension, and gives our protocols the form of direct reductions from multi-party learning to single-party learning, with no cryptographic overhead.

Human-AI Collaboration. The HCI literature on human-AI interaction has identified *complementarity* as a core goal — that a team consisting of a human and a model should perform measurably better than either of them could perform alone Bansal et al. [2021]. In particular, collaboration in the form of interaction is an explicit design goal Gomez et al. [2025], although one that has been hard to realize. Peng, Garg, and Kleinberg [Peng et al., 2024] prove a “no-free-lunch” theorem for human-AI collaboration, showing that for protocols that do not engage in an interaction (i.e. are just a post-processing of individual static predictors), non-trivial aggregation schemes (that do not always follow the prediction of a single model) must sometimes perform *worse* than the worst single predictor. Other empirical and theoretical studies of human-AI collaboration with the goal of improving over the best individual model include [Green and Chen, 2019, Donahue et al., 2022, Noti et al., 2025]. We give a protocol involving interaction (thus circumventing the barrier result proven by Peng et al. [2024]) that guarantees that a collaborative team can do strictly better than either alone. Additionally, common empirical approaches to human-AI collaboration often use insights into the model’s reasoning through ‘explanations’ as a form of communication. However, empirical studies show mixed results Bansal et al. [2021], Goh et al. [2024]; explanations can sometimes be ineffective or even misleading, potentially hindering human understanding or team performance, particularly if the explanations themselves are flawed. Our framework explores a different pathway for collaboration, that circumvents the need for explanations by replacing them with sharing only predictions.

Multi-modal Learning. Effectively integrating information across modalities like vision and language is a key challenge in multi-modal learning (see Baltrušaitis et al. [2018], Li and Tang [2024] and citations within). Standard techniques often involve either early fusion, combining representations before joint processing, or late fusion, typically averaging predictions from unimodal models. Early fusion may require complex joint models and careful feature alignment, while our theoretical results suggest simple late fusion can be suboptimal. In contrast, our protocols utilize iterative prediction or action exchange, requiring only learning on native data modalities. This mechanism avoids feature-level fusion entirely, enables communication complexity independent of data dimensionality, and represents a direct reduction to single-party learning, thus sidestepping the need for explicit feature alignment or joint model training overhead.

H Collaboration via Decisions

Thus far we have focused on real valued outcome spaces $\mathcal{Y} = [0, 1]$ in which we evaluate predictions by their squared error. Next we turn to an extension where the outcome space $\mathcal{Y} = [0, 1]^d$ is d -dimensional. The number of possible predictions (up to any reasonable discretization) now grows exponentially in d , and so the natural extension of our previous approach of asking the two parties to obtain no swap regret with respect to our predictions becomes infeasible — all known algorithms for obtaining this would have both run-time and regret bounds scaling exponentially with d or else regret bounds diminishing exponentially slowly with T . To circumvent this issue, we model Alice and Bob as decision makers who use predictions to inform downstream actions. More concretely, Alice and Bob have an action set \mathcal{A} and a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$ taking as input an action and outcome. As before, both parties will maintain predictions of the real-valued underlying

outcome. However, rather than communicating their estimates of the state directly, they will now simply communicate *actions* — specifically, the utility-maximizing action relative to their prediction.

1234 H.1 Decision Preliminaries

1235 **Definition H.1** (Best Response Action). *Fix a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$ and an*
 1236 *outcome/prediction $y \in \mathcal{Y}$. The best response to y according to u is the action $\text{BR}_u(y) =$*
 1237 *$\arg \max_{a \in \mathcal{A}} u(a, y)$.*

1238 Throughout this section, will assume that the utility function u is linear and Lipschitz in the outcome.

1239 **Assumption 3.** *We assume that the utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$ satisfies: for every action*
 1240 *$a \in \mathcal{A}$,*

- $u(a, \cdot)$ is linear in its second argument: for all $\alpha_1, \alpha_2 \in \mathbb{R}$, $y_1, y_2 \in [0, 1]^d$,

$$u(a, \alpha_1 y_1 + \alpha_2 y_2) = \alpha_1 u(a, y_1) + \alpha_2 u(a, y_2)$$

- $u(a, \cdot)$ is L -Lipschitz in its second argument in the L_∞ -norm: for all $y_1, y_2 \in [0, 1]^d$,

$$|u(a, y_1) - u(a, y_2)| \leq L \|y_1 - y_2\|_\infty.$$

1241 **Remark H.2.** *One natural special case is when y represents a probability distribution over d discrete*
 1242 *outcomes c_1, \dots, c_d , such that there is an arbitrary mapping $M(a, c)$ from action/outcome pairs to*
 1243 *utilities $[0, 1]$. In this case, $u(a, y)$ represents the expected utility of the action a over the outcome*
 1244 *distribution, which is linear in y by the linearity of expectation. The utility function is L -Lipschitz*
 1245 *in the L_∞ -norm, where $L = \max_{a, c_1, c_2} (M(a, c_1) - M(a, c_2)) \leq 1$. So our assumption is satisfied*
 1246 *by any risk neutral (expectation maximizing) decision maker with arbitrary utilities over d payoff*
 1247 *relevant states—and is only more general.*

1248 [ht] **Input** $\mathcal{X}, \mathcal{Y}, K, T$, action space \mathcal{A} , utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$ each day $t = 1, \dots, T$
 1249 Receive $x^t = (x_A^t, x_B^t)$. Alice sees x_A^t and Bob sees x_B^t . each round $k = 1, 2, \dots, K$ k is odd Alice
 1250 predicts $\hat{y}_A^{t,k} \in \mathcal{Y}$, and sends Bob $a_A^{t,k} = \text{BR}_u(\hat{y}_A^{t,k})$. k is even Bob predicts $\hat{y}_B^{t,k}$, and sends Alice
 1251 $a_B^{t,k} = \text{BR}_u(\hat{y}_B^{t,k})$. Alice and Bob observe $y^t \in \mathcal{Y}$.

1252 The interaction between Alice and Bob is formalized in Protocol H.1 (we will sometimes omit the
 1253 subscripts A and B when it is not important). The history of interaction is similarly captured by a
 1254 conversation transcript, which now additionally contains the actions communicated by both parties.

1255 **Definition H.3** (Conversation Transcript $\pi^{1:T, 1:K}$). *A conversation transcript $\pi^{1:T, 1:K} \in$*
 1256 *$\{\mathcal{Y}^{K+1} \times \mathcal{A}^K\}^T$ is a sequence of tuples of predictions made and actions chosen over rounds*
 1257 *by Alice and Bob (alternating across rounds), and the outcome, over T days:*

$$\pi^{1:T, 1:K} = \left\{ \left(\hat{y}_A^{1,1}, a_A^{1,1}, \hat{y}_B^{1,2}, a_B^{1,2}, \dots, \hat{y}_A^{1,K}, a_A^{1,K}, y^1 \right), \dots, \left(\hat{y}_A^{T,1}, a_A^{T,1}, \hat{y}_B^{T,2}, a_B^{T,2}, \dots, \hat{y}_A^{T,K}, a_A^{T,K}, y^T \right) \right\}.$$

1258 We define $\pi^{1:T:k}$ to be the restriction to only round k of conversation across days as follows:

$$\pi^{1:T:k} = \begin{cases} \{(\hat{y}_A^{t,k}, a_A^{t,k}, y^t)\}_{t \in [T]} & \text{if } k \text{ is odd,} \\ \{(\hat{y}_B^{t,k}, a_B^{t,k}, y^t)\}_{t \in [T]} & \text{otherwise.} \end{cases}$$

1259 Similarly, we will use the notation $\pi^{1:T}$ to refer to a single sequence of predictions and actions over
 1260 T days, outside the context of a conversation.

1261 **Definition H.4** (Prediction Transcript $\pi^{1:T}$). *A prediction transcript $\pi^{1:T} \in \{\mathcal{Y}^2 \times \mathcal{A}\}^T$ is a*
 1262 *sequence of tuples of predictions, actions, and outcomes over T days:*

$$\pi^{1:T} = \{(\hat{y}^1, a^1, y^1), \dots, (\hat{y}^T, a^T, y^T)\}$$

1263 Our goal is still to effectively aggregate information — in that the sequence of *actions* that results
 1264 from interaction between two parties only with access to their own features has *utility* comparable to
 1265 the best function mapping the parties *joint* feature space to actions in some benchmark policy class.
 1266 Below, we define benchmark classes for our setting as a collection of policies mapping contexts to
 1267 actions.

1268 **Definition H.5** (Individual Policy Classes $\mathcal{C}_A, \mathcal{C}_B$). Let $\mathcal{C}_A : \{\mathcal{X}_A \mapsto \mathcal{A}\}$ be a set of functions
 1269 mapping from Alice’s feature set to an action in \mathcal{A} . We analogously refer to \mathcal{C}_B for Bob.

1270 **Definition H.6** (Joint Policy Class \mathcal{C}_J). Let $\mathcal{C}_J : \{\mathcal{X} \mapsto \mathcal{A}\}$ be a set of functions mapping from the
 1271 entire feature set $\mathcal{X} = \mathcal{X}_A \times \mathcal{X}_B$ to an action in \mathcal{A} .

1272 **Assumption 4.** As before, we assume that all classes \mathcal{C} contain the set of all constant functions
 1273 $\{c(x) = a\}_{a \in \mathcal{A}}$.

1274 H.2 Decision Calibration and Regret

1275 We will appeal to a coarse notion of calibration suitable for high dimensional prediction problems
 1276 called “decision calibration” [Zhao et al., 2021, Noarov et al., 2023, Gopalan et al., 2023]. For a
 1277 single sequence of predictions, decision calibration asks that the predictions are unbiased conditional
 1278 not on the predictions themselves, but on the actions induced by best responding to the predictions.
 1279 The variant we use here is from Noarov et al. [2023].

1280 **Definition H.7** (f -Decision Calibration). Fix a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$. Fix an error
 1281 function $f : [T] \rightarrow \mathbb{R}$. We say that a transcript $\pi^{1:T}$ is f -decision calibrated with respect to u if for
 1282 all $a \in \mathcal{A}$:

$$\left\| \sum_{t=1}^T \mathbb{1}[a^t = a](\hat{y}_t - y_t) \right\|_{\infty} \leq f(|T(a)|)$$

1283 where $a^t = \text{BR}_u(\hat{y}^t)$ and $T(a) = \{t : a^t = a\}$ is the subsequence of days in which the best response
 1284 to \hat{y}^t according to u is a .

1285 When we are interested in competing with a benchmark class \mathcal{C} , another condition is also useful:
 1286 decision *cross-calibration* asks that predictions be unbiased conditional on the policy that best
 1287 responds to our predictions, *and* the decision made by each benchmark policy in \mathcal{C} :

1288 **Definition H.8** ((f, \mathcal{C}) -Decision Cross Calibration). Fix a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$ and a
 1289 policy class $\mathcal{C} : \{c : \mathcal{X} \rightarrow \mathcal{A}\}$. Fix an error function $f : [T] \rightarrow \mathbb{R}$. We say that a transcript $\pi^{1:T}$ is
 1290 (f, \mathcal{C}) -decision cross calibrated with respect to u if for all $a, a' \in \mathcal{A}$ and all $c \in \mathcal{C}$:

$$\left\| \sum_{t=1}^T \mathbb{1}[a^t = a, c(x^t) = a'](\hat{y}_t - y_t) \right\|_{\infty} \leq f(|T(a, a')|)$$

1291 where $a^t = \text{BR}_u(\hat{y}^t)$ and $T(a, a') = \{t : a^t = a, c(x^t) = a'\}$ is the subsequence of days in which
 1292 the best response to \hat{y}^t according to u is a and the action suggested by policy c is a' .

1293 We can also define an analogous notion of swap regret with respect to a policy class \mathcal{C} , which we will
 1294 call *decision swap regret*. Decision swap regret compares the utility of best response actions induced
 1295 by predictions \hat{y}^t to the counterfactual utility of actions suggested by policies in \mathcal{C} .

1296 **Definition H.9** ((f^S, \mathcal{C}) -Decision Swap Regret). Fix a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$ and
 1297 a policy class $\mathcal{C} : \{c : \mathcal{X} \rightarrow \mathcal{A}\}$. Fix an error function $f^S : [T] \rightarrow \mathbb{R}$. A transcript $\pi^{1:T}$ has
 1298 (f^S, \mathcal{C}) -decision swap regret if:

$$\sum_{a \in \mathcal{A}} \max_{c \in \mathcal{C}} \left(\sum_{t=1}^T \mathbb{1}[a^t = a] u(c(x^t), y^t) \right) - \sum_{t=1}^T u(a^t, y^t) \leq f^S(T)$$

1299 **Remark H.10.** This is the same as the notion of decision swap regret defined in Lu et al. [2025],
 1300 restricted to a single utility function (Lu et al. [2025] ask for this condition to hold over a class of
 1301 utility functions).

1302 Lu et al. [2025] relate decision calibration and decision swap calibration (conditions on *predictions*)
 1303 to decision swap regret on the sequence of actions that result from best-responding to the predictions:

1304 **Theorem H.11** (Theorem 1 of [Lu et al., 2025]). Fix a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$ and a
 1305 policy class $\mathcal{C} : \{c : \mathcal{X} \rightarrow \mathcal{A}\}$. If a transcript $\pi^{1:T}$ is f -decision calibrated and (f', \mathcal{C}) -decision cross
 1306 calibrated, and $a^t = \text{BR}_u(\hat{y}^t)$ for all $t \in [T]$, then $\pi^{1:T}$ has (f^S, \mathcal{C}) -decision swap regret, where:

$$f^S(T) \leq L|\mathcal{A}|f\left(\frac{T}{|\mathcal{A}|}\right) + L|\mathcal{A}|^2 f'\left(\frac{T}{|\mathcal{A}|^2}\right)$$

1307 **Remark H.12.** We remark that under the assumption that the class \mathcal{C} contains constant functions,
 1308 (f, \mathcal{C}) -decision cross calibration implies f -decision calibration (in fact, decision cross calibration
 1309 implies decision calibration even if \mathcal{C} does not contain constant functions, but with a loss of a factor of
 1310 $|\mathcal{A}|$ in decision calibration error). Thus, (f, \mathcal{C}) -decision cross calibration alone suffices to guarantee
 1311 diminishing decision swap regret.

1312 Moving to the collaboration setting, we define decision conversation calibration following Collina
 1313 et al. [2025]; this condition asks for decision calibration conditional on the previous action sent by
 1314 the other party. In other words, the predictions that each party makes should be unbiased conditional
 1315 on both their own best response action *and* the best response action communicated at the previous
 1316 round.

1317 **Definition H.13** (f -Decision Conversation Calibration). Fix a error function $f^S : [T] \rightarrow \mathbb{R}$. Given
 1318 a transcript $\pi^{1:T, 1:K}$ from an interaction in the Collaboration Protocol (Protocol H.1), Alice is
 1319 f -decision conversation calibrated if for all odd rounds k and all pairs of actions $a, a' \in \mathcal{A}$:

$$\left\| \sum_{t=1}^T \mathbb{I}[a_A^{t,k} = a, a_B^{t,k-1} = a'] (\hat{y}_A^{t,k} - y^t) \right\|_{\infty} \leq f(|T(k, a, a')|)$$

1320 where $T(k, a, a') = \{t : a_A^{t,k} = a \text{ and } a_B^{t,k-1} = a'\}$ is the subsequence of days in which Alice
 1321 communicates action a on round k and Bob communicates a' on round $k-1$.

1322 Symmetrically, Bob is f -decision conversation calibrated if for all even rounds k and all pairs of
 1323 actions $a, a' \in \mathcal{A}$:

$$\left\| \sum_{t=1}^T \mathbb{I}[a_B^{t,k} = a, a_A^{t,k-1} = a'] (\hat{y}_B^{t,k} - y^t) \right\|_{\infty} \leq f(|T(k, a, a')|)$$

1324 where $T(k, a, a') = \{t : a_B^{t,k} = a \text{ and } a_A^{t,k-1} = a'\}$ is the subsequence of days in which Bob
 1325 communicates action a on round k and Alice communicates a' on round $k-1$.

1326 Similarly, we extend conversation swap regret to *decision conversation swap regret*, which is the
 1327 decision swap regret conditional on the action chosen by the other party in the previous round.

1328 **Definition H.14** ((f^S, \mathcal{C}) -Decision Conversation Swap Regret). Fix a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow$
 1329 $[0, 1]$. Fix an error function $f^S : [T] \rightarrow \mathbb{R}$ and a policy class \mathcal{C}_A . Given a transcript $\pi^{1:T, 1:K}$ from an
 1330 interaction in the Collaboration Protocol (Protocol H.1), Alice has (f^S, \mathcal{C}_A) -decision conversation
 1331 swap regret if for all odd rounds k and all $a' \in \mathcal{A}$:

$$\sum_{a \in \mathcal{A}} \max_{c \in \mathcal{C}_A} \left(\sum_{t \in T_B(k-1, a')} \mathbb{I}[a_A^{t,k} = a] u(c(x_A^t), y^t) \right) - \sum_{t \in T_B(k-1, a')} u(a_A^{t,k}, y^t) \leq f^S(|T_B(k-1, a')|).$$

1332 where $a_A^{t,k} = \text{BR}_u(\hat{y}_A^{t,k})$ and $T_B(k-1, a') = \{t : \text{BR}_u(\hat{y}_B^{t,k-1}) = a'\}$ is the subsequence of days
 1333 where Bob's action in round $k-1$ is a' .

1334 If Bob satisfies a symmetric condition on even rounds k with respect to \mathcal{H}_B , we say that Bob has
 1335 (f, \mathcal{H}_B) -decision conversation swap regret.

1336 **Assumption 5.** As before, we assume that all error functions $f : [T] \rightarrow \mathbb{R}$ are concave.

1337 Our approach will be different compared to the one we took in Section C for real valued outcomes.
 1338 There, we argued that swap regret (with respect to the predictions) implied conversation calibration,
 1339 and hence fast agreement. In the action setting, decision swap regret does *not* necessarily imply
 1340 decision calibration, which is what is needed to invoke the fast agreement theorems of Collina et al.
 1341 [2025]. Instead we argue that decision calibration and decision cross calibration together imply both
 1342 decision conversation swap regret and decision conversation calibration.

1343 H.3 A Boosting Theorem for Decisions

1344 We now give a weak learning condition that parallels Definition B.1. Whereas Definition B.1 requires
 1345 that \mathcal{C}_A and \mathcal{C}_B jointly improve on the squared error of the best constant prediction whenever \mathcal{C}_J
 1346 does, the condition now requires that \mathcal{C}_A and \mathcal{C}_B jointly improve on the utility of the best constant
 1347 action whenever \mathcal{C}_J does.

1348 **Definition H.15** ($w(\cdot)$ -Weak Learning Condition for Decisions). Fix a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow$
1349 $[0, 1]$. Let \mathcal{C}_J be a policy class over the joint feature space \mathcal{X} . Let $\mathcal{C}_A = \{c_A : \mathcal{X}_A \rightarrow \mathcal{A}\}$ and
1350 $\mathcal{C}_B = \{c_B : \mathcal{X}_B \rightarrow \mathcal{A}\}$ be policy classes over \mathcal{X}_A and \mathcal{X}_B respectively. Let $w : [0, 1] \rightarrow [0, 1]$ be a
1351 strictly increasing, continuous, and convex function that satisfies $w(\gamma) \leq \gamma$. We say that \mathcal{C}_A and \mathcal{C}_B
1352 jointly satisfy the $w(\cdot)$ -weak learning condition with respect to \mathcal{C}_J if for any sequence of contexts
1353 $x^{1:T}$ and labels $y^{1:T}$, any $S \subseteq [T]$, and any $\gamma \in [0, 1]$, if:

$$\max_{c_J \in \mathcal{C}_J} \frac{1}{|S|} \sum_{t \in S} u(c_J(x^t), y^t) - \max_{a \in \mathcal{A}} \frac{1}{|S|} \sum_{t \in S} u(a, y^t) \geq \gamma,$$

1354 then there exists either $c_A \in \mathcal{C}_A$ or $c_B \in \mathcal{C}_B$ such that:

$$\frac{1}{|S|} \sum_{t \in S} u(c_A(x_A^t), y^t) - \max_{a \in \mathcal{A}} \frac{1}{|S|} \sum_{t \in S} u(a, y^t) \geq w(\gamma)$$

1355 or:

$$\frac{1}{|S|} \sum_{t \in S} u(c_B(x_B^t), y^t) - \max_{a \in \mathcal{A}} \frac{1}{|S|} \sum_{t \in S} u(a, y^t) \geq w(\gamma)$$

1356 Next we show that if \mathcal{C}_A and \mathcal{C}_B satisfy the weak learning condition with respect to \mathcal{C}_J , then low
1357 decision swap regret with respect to the classes \mathcal{C}_A and \mathcal{C}_B implies that the best response action
1358 obtains utility as high as any policy $c_J \in \mathcal{C}_J$ (up to regret terms). The proof mostly mirrors that of
1359 Theorem B.3.

1360 **Theorem H.16.** Fix a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$. Let \mathcal{C}_J be a policy class over the joint
1361 feature space \mathcal{X} . Let $\mathcal{C}_A = \{c_A : \mathcal{X}_A \rightarrow \mathcal{A}\}$ and $\mathcal{C}_B = \{c_B : \mathcal{X}_B \rightarrow \mathcal{A}\}$ be policy classes over \mathcal{X}_A
1362 and \mathcal{X}_B respectively. Fix a transcript $\pi^{1:T}$. If:

- 1363 • $\pi^{1:T}$ has $(f^S, \mathcal{C}_A \cup \mathcal{C}_B)$ -decision swap regret (Definition H.9)
- 1364 • \mathcal{C}_A and \mathcal{C}_B jointly satisfy the $w(\cdot)$ -weak learning condition with respect to \mathcal{C}_J (Definition
1365 H.15)

1366 Then, $\pi^{1:T}$ has $\left(2Tw^{-1}\left(\frac{f^S(T)}{T}\right), \mathcal{C}_J\right)$ -decision swap regret when choosing the best response action.
1367 That is:

$$\sum_{a \in \mathcal{A}} \max_{c_J \in \mathcal{C}_J} \sum_{t=1}^T \mathbb{1}[\text{BR}_u(\hat{y}^t) = a] u(c_J(x^t), y^t) - \sum_{t=1}^T u(\text{BR}_u(\hat{y}^t), y^t) \leq 2Tw^{-1}\left(\frac{f^S(T)}{T}\right)$$

1368 whenever the inverse of w exists.

1369 *Proof.* Let $a^t = \text{BR}(\hat{y}^t)$. We show the contrapositive. Suppose there exists a collection $\{c_{J,a}\}_{a \in \mathcal{A}} \subseteq$
1370 \mathcal{C}_J such that:

$$\sum_{a \in \mathcal{A}} \sum_{t=1}^T \mathbb{1}[a^t = a] u(c_{J,a}(x^t), y^t) > \sum_{t=1}^T u(a^t, y^t) + 2Tw^{-1}\left(\frac{f^S(T)}{T}\right)$$

1371 Equivalently,

$$\frac{1}{T} \sum_{a \in \mathcal{A}} \sum_{t=1}^T \mathbb{1}[a^t = a] u(c_{J,a}(x^t), y^t) > \frac{1}{T} \sum_{a \in \mathcal{A}} \sum_{t=1}^T \mathbb{1}[a^t = a] u(a, y^t) + 2w^{-1}\left(\frac{f^S(T)}{T}\right)$$

1372 Since $\pi^{1:T}$ has $(f^S, \mathcal{C}_A \cup \mathcal{C}_B)$ -decision swap regret, and \mathcal{C}_A and \mathcal{C}_B contain the set of all constant
1373 functions (Assumption 4), the decision swap regret with respect to the collection of best constant
1374 actions is:

$$\frac{1}{T} \sum_{a \in \mathcal{A}} \max_{a^* \in \mathcal{A}} \sum_{t=1}^T \mathbb{1}[a^t = a] u(a^*, y^t) - \frac{1}{T} \sum_{a \in \mathcal{A}} \sum_{t=1}^T \mathbb{1}[a^t = a] u(a, y^t) \leq \frac{f^S(T)}{T} \leq w^{-1}\left(\frac{f^S(T)}{T}\right)$$

1375 where the second inequality uses the fact that $w(\gamma) \leq \gamma$, and so $\gamma \leq w^{-1}(\gamma)$. Then, since the utility
 1376 of actions a^t is close to the utility of the collection of best constant actions, we have that:

$$\begin{aligned} \frac{1}{T} \sum_{a \in \mathcal{A}} \sum_{t=1}^T \mathbb{1}[a^t = a] u(c_{J,a}(x^t), y^t) &> \frac{1}{T} \sum_{a \in \mathcal{A}} \sum_{t=1}^T \mathbb{1}[a^t = a] u(a, y^t) + 2w^{-1} \left(\frac{f^S(T)}{T} \right) \\ &\geq \frac{1}{T} \sum_{a \in \mathcal{A}} \max_{a^* \in \mathcal{A}} \sum_{t=1}^T \mathbb{1}[a^t = a] u(a^*, y^t) - w^{-1} \left(\frac{f^S(T)}{T} \right) + 2w^{-1} \left(\frac{f^S(T)}{T} \right) \\ &= \frac{1}{T} \sum_{a \in \mathcal{A}} \max_{a^* \in \mathcal{A}} \sum_{t=1}^T \mathbb{1}[a^t = a] u(a^*, y^t) + w^{-1} \left(\frac{f^S(T)}{T} \right) \end{aligned}$$

Let $S_a = \{t : a^t = a\}$ and

$$\gamma_a = \frac{1}{|S_a|} \sum_{t=1}^T \mathbb{1}[a^t = a] u(c_{J,a}(x^t), y^t) - \max_{a^* \in \mathcal{A}} \frac{1}{|S_a|} \sum_{t=1}^T \mathbb{1}[a^t = a] u(a^*, y^t)$$

1377 Then, we can rewrite the expression above as:

$$\begin{aligned} &\frac{1}{T} \sum_{a \in \mathcal{A}} \sum_{t=1}^T \mathbb{1}[a^t = a] u(c_{J,a}(x^t), y^t) - \frac{1}{T} \sum_{a \in \mathcal{A}} \max_{a^* \in \mathcal{A}} \sum_{t=1}^T \mathbb{1}[a^t = a] u(a^*, y^t) \\ &= \frac{1}{T} \sum_{a \in \mathcal{A}} |S_a| \cdot \frac{1}{|S_a|} \sum_{t=1}^T \mathbb{1}[a^t = a] u(c_{J,a}(x^t), y^t) - \frac{1}{T} \sum_{a \in \mathcal{A}} |S_a| \max_{a^* \in \mathcal{A}} \frac{1}{|S_a|} \sum_{t=1}^T \mathbb{1}[a^t = a] u(a^*, y^t) \\ &= \frac{1}{T} \sum_{a \in \mathcal{A}} |S_a| \gamma_a \\ &> w^{-1} \left(\frac{f^S(T)}{T} \right) \end{aligned}$$

1378 Observe that since \mathcal{C}_J contains the set of all constant functions (Assumption 4), there is always a
 1379 choice of $\{c_{J,a}\}_{a \in \mathcal{A}}$ such that γ_a is non-negative for all a . Thus, we can invoke the weak learning
 1380 condition: on any subsequence S_a for which $c_{J,a}$ improves over the best constant action by γ_a , there
 1381 is some $c_a \in \mathcal{C}_A \cup \mathcal{C}_B$ that improves over the best constant action by $w(\gamma_a)$. Specifically, there exists
 1382 a collection $\{c_a\}_{a \in \mathcal{A}} \subseteq \mathcal{C}_A \cup \mathcal{C}_B$ such that:

$$\begin{aligned} &\frac{1}{T} \sum_{a \in \mathcal{A}} \sum_{t=1}^T \mathbb{1}[a^t = a] u(c_a(x^t), y^t) - \frac{1}{T} \sum_{a \in \mathcal{A}} \max_{a^* \in \mathcal{A}} \sum_{t=1}^T \mathbb{1}[a^t = a] u(a^*, y^t) \\ &= \frac{1}{T} \sum_{a \in \mathcal{A}} |S_a| \cdot \frac{1}{|S_a|} \sum_{t=1}^T \mathbb{1}[a^t = a] u(c_a(x^t), y^t) - \frac{1}{T} \sum_{a \in \mathcal{A}} |S_a| \max_{a^* \in \mathcal{A}} \frac{1}{|S_a|} \sum_{t=1}^T \mathbb{1}[a^t = a] u(a^*, y^t) \\ &\geq \frac{1}{T} \sum_{a \in \mathcal{A}} |S_a| w(\gamma_a) \quad \text{(by the } w\text{-weak learning condition)} \\ &\geq w \left(\frac{1}{T} \sum_{a \in \mathcal{A}} |S_a| \gamma_a \right) \quad \text{(by convexity of } w \text{ and Jensen's inequality)} \\ &> w \left(w^{-1} \left(\frac{f^S(T)}{T} \right) \right) \quad \text{(by monotonicity of } w) \\ &= \frac{f^S(T)}{T} \end{aligned}$$

1383 In particular, this implies that:

$$\begin{aligned}
\sum_{a \in \mathcal{A}} \max_{c_a^* \in \mathcal{C}_A \cup \mathcal{C}_B} \sum_{t=1}^T \mathbb{1}[a^t = a] u(c_a^*(x^t), y^t) &\geq \sum_{a \in \mathcal{A}} \sum_{t=1}^T \mathbb{1}[a^t = a] u(c_a(x^t), y^t) \\
&> \sum_{a \in \mathcal{A}} \max_{a^* \in \mathcal{A}} \sum_{t=1}^T \mathbb{1}[a^t = a] u(a^*, y^t) + f^S(T) \\
&\geq \sum_{a \in \mathcal{A}} \sum_{t=1}^T \mathbb{1}[a^t = a] u(a, y^t) + f^S(T) \\
&= \sum_{t=1}^T u(a^t, y^t) + f^S(T)
\end{aligned}$$

1384 which violates the $(f^S, \mathcal{C}_A \cup \mathcal{C}_B)$ -decision swap regret condition. This completes the proof.

1385

□

1386 H.4 Online Decision Collaboration

1387 We now extend our collaboration protocol to the action setting. We show that if both parties make
1388 predictions that have low decision conversation swap regret with respect to \mathcal{C}_A and \mathcal{C}_B respectively
1389 and are decision conversation calibrated, then they must quickly converge to a sequence of predictions
1390 at some round k (not necessarily the final round) at which they have low decision swap regret to both
1391 \mathcal{C}_A and \mathcal{C}_B simultaneously. At this round, if \mathcal{C}_A and \mathcal{C}_B satisfy the weak learning condition relative
1392 to a joint class \mathcal{C}_J , then we can argue that the predictions have utility as high as the best policy in the
1393 joint class. We then go on to show that the final sequence of predictions must have utility not much
1394 lower than the predictions at round k , and therefore also the best policy in \mathcal{C}_J .

1395 We begin by arguing that if Alice and Bob have low decision swap regret and are decision conversation
1396 calibrated with respect to their individual policy classes \mathcal{C}_A and \mathcal{C}_B , the best response actions at some
1397 round k will have low decision swap regret to both \mathcal{C}_A and \mathcal{C}_B . The argument will closely follow that
1398 of Theorem C.2. Since both Alice and Bob are decision conversation calibrated, there will exist some
1399 round k (assume for now that Alice communicates on round k) such that on most days, their actions
1400 ε -agree — that is, Alice’s action at round k is an ε -approximate best response for Bob at round $k + 1$,
1401 and vice versa (Lemma H.21). Our goal is to show that on round k , Alice has bounded decision
1402 swap regret simultaneously against \mathcal{C}_A and against \mathcal{C}_B . The first is simple: on round k , Alice has
1403 low decision conversation swap regret with respect to \mathcal{C}_A , and thus she has low decision swap regret
1404 with respect to \mathcal{C}_A (Lemma H.18). To argue the second: on round $k + 1$, Bob has bounded decision
1405 conversation swap regret with respect to \mathcal{C}_B . In particular, this means that conditioned on Alice’s
1406 action on round k , Bob’s actions are competitive against any policy in \mathcal{C}_B^2 . We will additionally show
1407 that since they agree, Alice’s actions at round k obtain similar utility to Bob’s actions at round $k + 1$
1408 (Lemma H.22). Thus, conditioned on Alice’s action on round k , Alice’s actions are also competitive
1409 against any policy in \mathcal{C}_B . Since this is true for any action that Alice chooses, Alice must also have
1410 low decision swap regret with respect to \mathcal{C}_B on this round.

1411 **Theorem H.17.** *Suppose Alice has (f_A^S, \mathcal{C}_A) -decision conversation swap regret and f_A -decision*
1412 *conversation calibration. Similarly, suppose Bob has (f_B^S, \mathcal{C}_B) -decision conversation swap regret*
1413 *and f_B -decision conversation calibration. If they engage in Protocol H.1 for T days, with K*
1414 *rounds each day, then there exists a round k of the protocol such that the transcript $\pi^{1:T,k}$ has*
1415 *$(\max\{\lambda_A, \lambda_B\}, \mathcal{C}_A \cup \mathcal{C}_B)$ -decision swap regret, where:*

$$\lambda_A \leq |\mathcal{A}| f_A^S \left(\frac{T}{|\mathcal{A}|} \right) + L |\mathcal{A}|^2 f_A \left(\frac{T}{|\mathcal{A}|^2} \right) + 2T \left(\frac{1}{(K-1)} + \beta(T) \right)^{1/2}$$

1416 and

$$\lambda_B \leq |\mathcal{A}| f_B^S \left(\frac{T}{|\mathcal{A}|} \right) + L |\mathcal{A}|^2 f_B \left(\frac{T}{|\mathcal{A}|^2} \right) + 2T \left(\frac{1}{(K-1)} + \beta(T) \right)^{1/2}$$

²Notice that the decision conversation swap regret condition is in fact stronger, since it guarantees that Bob’s actions are competitive conditioned on *both* Alice’s action on round k and Bob’s action on round $k + 1$. We will only use the weaker “external” regret guarantee at this step.

1417 Here, $\beta(T) = \frac{L|\mathcal{A}|^2}{T} \left(f_A \left(\frac{T}{|\mathcal{A}|^2} \right) + f_B \left(\frac{T}{|\mathcal{A}|^2} \right) \right)$.

1418 To prove Theorem H.17, we first introduce key lemmas we will use. In what follows, we denote
 1419 $a_A^{t,k} = \text{BR}_u(\hat{y}_A^{t,k})$ and $a_B^{t,k+1} = \text{BR}_u(\hat{y}_B^{t,k+1})$ for all $k \in [K]$ and $t \in [T]$. The first lemma shows
 1420 how to convert a decision conversation swap regret guarantee into a decision swap regret guarantee
 1421 for the sequence of predictions on any round k . Observe that decision conversation swap regret
 1422 stronger than decision swap regret, since it additionally conditions on the action chosen by the other
 1423 party in the previous round.

1424 **Lemma H.18.** *If Alice has (f_A^S, \mathcal{C}_A) -decision conversation swap regret, then for all odd $k \in [K]$,
 1425 the transcript $\pi^{1:T,k}$ satisfies (f'_A, \mathcal{C}_A) -decision swap regret, where:*

$$f'_A(T) \leq |\mathcal{A}| f_A^S \left(\frac{T}{|\mathcal{A}|} \right)$$

1426 A symmetric statement holds for Bob.

1427 *Proof.* We can compute the decision swap regret with respect to \mathcal{C}_A over round k :

$$\begin{aligned} & \sum_{a \in \mathcal{A}} \max_{c \in \mathcal{C}_A} \sum_{t=1}^T \mathbb{1}[a_A^{t,k} = a] u(c(x_A^t), y^t) - \sum_{t=1}^T u(a_A^{t,k}, y^t) \\ &= \sum_{a \in \mathcal{A}} \max_{c \in \mathcal{C}_A} \sum_{a' \in \mathcal{A}} \sum_{t \in T_B(k-1, a')} \mathbb{1}[a_A^{t,k} = a] u(c(x_A^t), y^t) - \sum_{a' \in \mathcal{A}} \sum_{t \in T_B(k-1, a')} u(a_A^{t,k}, y^t) \\ &\leq \sum_{a' \in \mathcal{A}} \left(\sum_{a \in \mathcal{A}} \max_{c \in \mathcal{C}_A} \left(\sum_{t \in T_B(k-1, a')} \mathbb{1}[a_A^{t,k} = a] u(c(x_A^t), y^t) \right) - \sum_{t \in T_B(k-1, a')} u(a_A^{t,k}, y^t) \right) \\ &\quad \text{(by the fact that moving the max inside the sum only strengthens the benchmark)} \\ &\leq \sum_{a' \in \mathcal{A}} f_A^S(|T_B(k-1, a')|) \quad \text{(by } (f_A^S, \mathcal{C}_A)\text{-decision conversation swap regret)} \\ &\leq |\mathcal{A}| f_A^S \left(\frac{T}{|\mathcal{A}|} \right) \quad \text{(by concavity of } f_A^S) \end{aligned}$$

1428 □

1429 We next argue that if Alice and Bob communicate for sufficiently many rounds, there will exist some
 1430 round where they ε -agree on a large fraction of days. To do this, we use a result from Collina et al.
 1431 [2025] showing that the utility must increase on any round they disagree.

1432 **Lemma H.19** (Lemma 5.4 of Collina et al. [2025]). *If Bob is f_B -decision conversation calibrated,
 1433 then after engaging in Protocol H.1 for T days, for all odd rounds $k \in [K]$, we have:*

$$\sum_{t=1}^T u(a_B^{t,k+1}, y^t) - \sum_{t=1}^T u(a_A^{t,k}, y^t) \geq \varepsilon |D(T^{k+1})| - 2L|\mathcal{A}|^2 f_B \left(\frac{T}{|\mathcal{A}|^2} \right)$$

1434 where $D(T^{k+1})$ is the subset of days over round $k+1$ such that Alice and Bob ε -disagree, i.e.:

$$\left| u(a_A^{t,k}, \hat{y}_A^{t,k}) - u(a_B^{t,k+1}, \hat{y}_A^{t,k}) \right| > \varepsilon$$

1435 or

$$\left| u(a_A^{t,k}, \hat{y}_B^{t,k+1}) - u(a_B^{t,k+1}, \hat{y}_B^{t,k+1}) \right| > \varepsilon$$

1436 Furthermore, if Alice is f_A -decision conversation calibrated, then after engaging in Protocol H.1 for
 1437 T days, for all even rounds $k \in [K]$, we have:

$$\sum_{t=1}^T u(a_A^{t,k+1}, y^t) - \sum_{t=1}^T u(a_B^{t,k}, y^t) \geq \varepsilon |D(T^{k+1})| - 2L|\mathcal{A}|^2 f_A \left(\frac{T}{|\mathcal{A}|^2} \right)$$

1438 **Remark H.20.** Lemma 5.4 of Collina et al. [2025] is stated for a slightly different setting where,
 1439 every day, the conversation protocol halts after both parties ε -agree (whereas our protocol runs for a
 1440 fixed number of rounds). There, the decrease in utility is a function of the number of days the protocol
 1441 advances to the next round. This is equivalent to the number of days Alice and Bob ε -disagree, and
 1442 so the result translates straightforwardly to our setting.

1443 **Lemma H.21.** After engaging in Protocol H.1 for T days, each with K rounds, there is at least one
 1444 round k (without loss, assume k odd) such that the fraction of days Alice and Bob ε -agree, i.e.:

$$\left| u(a_A^{t,k}, \hat{y}_A^{t,k}) - u(a_B^{t,k+1}, \hat{y}_A^{t,k}) \right| \leq \varepsilon$$

1445 and

$$\left| u(a_A^{t,k}, \hat{y}_B^{t,k+1}) - u(a_B^{t,k+1}, \hat{y}_B^{t,k+1}) \right| \leq \varepsilon,$$

1446 is at least $1 - \left(\frac{1}{(K-1)\varepsilon} + \frac{\beta(T)}{\varepsilon} \right)$, where $\beta(T) = \frac{L|\mathcal{A}|^2}{T} \left(f_A \left(\frac{T}{|\mathcal{A}|^2} \right) + f_B \left(\frac{T}{|\mathcal{A}|^2} \right) \right)$.

1447 *Proof.* Using Lemma H.19, we can calculate the difference in utility over two rounds:

$$\begin{aligned} & \sum_{t=1}^T u(a_A^{t,k+2}, y^t) - \sum_{t=1}^T u(a_A^{t,k}, y^t) \\ &= \sum_{t=1}^T u(a_A^{t,k+2}, y^t) - \sum_{t=1}^T u(a_B^{t,k+1}, y^t) + \sum_{t=1}^T u(a_B^{t,k+1}, y^t) - \sum_{t=1}^T u(a_A^{t,k}, y^t) \\ &\geq \varepsilon |D(T^{k+2})| - 2L|\mathcal{A}|^2 f_A \left(\frac{T}{|\mathcal{A}|^2} \right) + \varepsilon |D(T^{k+1})| - 2L|\mathcal{A}|^2 f_B \left(\frac{T}{|\mathcal{A}|^2} \right) \quad (\text{by Lemma H.19}) \\ &= \varepsilon (|D(T^{k+2})| + |D(T^{k+1})|) - 2T\beta(T) \quad (\text{by definition of } \beta(T)) \end{aligned}$$

1448 Now, to calculate the difference in utility over K rounds (we assume without loss that K is odd; we
 1449 obtain the same result if K is even), we iteratively apply the above $(K-1)/2$ times:

$$\begin{aligned} \sum_{t=1}^T u(a_A^{t,K}, y^t) - \sum_{t=1}^T u(a_A^{t,1}, y^t) &\geq \varepsilon \sum_{k=2}^K |D(T^k)| - \frac{K-1}{2} \cdot 2T\beta(T) \\ &= \varepsilon \sum_{k=2}^K |D(T^k)| - (K-1)T\beta(T) \end{aligned}$$

1450 Observe that since utilities are bounded between $[0, 1]$, the left hand side of this expression is at most
 1451 T . Thus, rearranging, we have that the total number of ε -disagreements is at most:

$$\sum_{k=2}^K |D(T^k)| \leq \frac{T + (K-1)T\beta(T)}{\varepsilon}$$

1452 Therefore, there must exist some round k^* with a number of ε -disagreements at most:

$$|D(T^{k^*})| \leq \frac{T + (K-1)T\beta(T)}{(K-1)\varepsilon} = \frac{T}{(K-1)\varepsilon} + \frac{T\beta(T)}{\varepsilon}$$

1453 That is, on round k^* , the fraction of ε -disagreements over T days is at most:

$$\frac{|D(T^{k^*})|}{T} \leq \frac{1}{(K-1)\varepsilon} + \frac{\beta(T)}{\varepsilon}$$

1454 which proves the claim. \square

1455 Finally, we show that on any round where Alice and Bob ε -agree, the utilities under their best
 1456 response actions do not differ by too much.

1457 **Lemma H.22.** Suppose that on some odd round $k \in [K]$, on at least $1 - \delta$ fraction of days $t \in [T]$,
 1458 we have:

$$\left| u(a_A^{t,k}, \hat{y}_B^{t,k+1}) - u(a_B^{t,k+1}, \hat{y}_B^{t,k+1}) \right| \leq \varepsilon$$

1459 If Bob is f_B -decision conversation calibrated, then:

$$\sum_{t=1}^T u(a_B^{t,k+1}, y^t) - \sum_{t=1}^T u(a_A^{t,k}, y^t) \leq (\varepsilon + \delta)T + L|\mathcal{A}|^2 f_B \left(\frac{T}{|\mathcal{A}|^2} \right)$$

1460 A symmetric statement holds for even round k and Alice.

1461 *Proof.* We can compute:

$$\begin{aligned} & \sum_{t=1}^T u(a_B^{t,k+1}, y^t) - \sum_{t=1}^T u(a_A^{t,k}, y^t) \\ & \leq \sum_{t=1}^T u(a_B^{t,k+1}, \hat{y}_B^{t,k+1}) - \sum_{t=1}^T u(a_A^{t,k}, y^t) \quad (\text{by definition of best response to } \hat{y}_B^{t,k+1}) \\ & = \sum_{t=1}^T u(a_B^{t,k+1}, \hat{y}_B^{t,k+1}) - \sum_{a \in \mathcal{A}} \sum_{a' \in \mathcal{A}} \sum_{t=1}^T \mathbb{1}[a_B^{t,k+1} = a, a_A^{t,k} = a'] u(a', y^t) \\ & = \sum_{t=1}^T u(a_B^{t,k+1}, \hat{y}_B^{t,k+1}) - \sum_{a \in \mathcal{A}} \sum_{a' \in \mathcal{A}} u \left(a', \sum_{t=1}^T \mathbb{1}[a_B^{t,k+1} = a, a_A^{t,k} = a'] y^t \right) \quad (\text{by linearity of } u) \\ & \leq \sum_{t=1}^T u(a_B^{t,k+1}, \hat{y}_B^{t,k+1}) - \sum_{a \in \mathcal{A}} \sum_{a' \in \mathcal{A}} u \left(a', \sum_{t=1}^T \mathbb{1}[a_B^{t,k+1} = a, a_A^{t,k} = a'] \hat{y}_B^{t,k+1} \right) + L|\mathcal{A}|^2 f_B \left(\frac{T}{|\mathcal{A}|^2} \right) \\ & = \sum_{t=1}^T u(a_B^{t,k+1}, \hat{y}_B^{t,k+1}) - \sum_{t=1}^T u(a_A^{t,k}, \hat{y}_B^{t,k+1}) + L|\mathcal{A}|^2 f_B \left(\frac{T}{|\mathcal{A}|^2} \right) \quad (\text{by linearity of } u) \\ & = \sum_{t=1}^T \mathbb{1} \left[\left| u(a_A^{t,k}, \hat{y}_B^{t,k+1}) - u(a_B^{t,k+1}, \hat{y}_B^{t,k+1}) \right| \leq \varepsilon \right] \left(u(a_B^{t,k+1}, \hat{y}_B^{t,k+1}) - u(a_A^{t,k}, \hat{y}_B^{t,k+1}) \right) \\ & \quad + \sum_{t=1}^T \mathbb{1} \left[\left| u(a_A^{t,k}, \hat{y}_B^{t,k+1}) - u(a_B^{t,k+1}, \hat{y}_B^{t,k+1}) \right| > \varepsilon \right] \left(u(a_B^{t,k+1}, \hat{y}_B^{t,k+1}) - u(a_A^{t,k}, \hat{y}_B^{t,k+1}) \right) \\ & \quad + L|\mathcal{A}|^2 f_B \left(\frac{T}{|\mathcal{A}|^2} \right) \\ & \leq \varepsilon(1 - \delta)T + \delta T + L|\mathcal{A}|^2 f_B \left(\frac{T}{|\mathcal{A}|^2} \right) \quad (\text{by assumption}) \\ & \leq (\varepsilon + \delta)T + L|\mathcal{A}|^2 f_B \left(\frac{T}{|\mathcal{A}|^2} \right) \quad (\text{since } \varepsilon, \delta \geq 0) \end{aligned}$$

1462 Here, the first inequality uses f_B -decision conversation calibration and the fact that u is L -Lipschitz;
 1463 we can see that for any $a' \in \mathcal{A}$:

$$\begin{aligned}
 & \left| \sum_{a \in \mathcal{A}} \sum_{a' \in \mathcal{A}} \left(u \left(a', \sum_{t=1}^T \mathbb{1}[a_B^{t,k+1} = a, a_A^{t,k} = a'] \hat{y}_B^{t,k+1} \right) - u \left(a', \sum_{t=1}^T \mathbb{1}[a_B^{t,k+1} = a, a_A^{t,k} = a'] y^t \right) \right) \right| \\
 & \leq \sum_{a \in \mathcal{A}} \sum_{a' \in \mathcal{A}} \left| u \left(a', \sum_{t=1}^T \mathbb{1}[a_B^{t,k+1} = a, a_A^{t,k} = a'] \hat{y}_B^{t,k+1} \right) - u \left(a', \sum_{t=1}^T \mathbb{1}[a_B^{t,k+1} = a, a_A^{t,k} = a'] y^t \right) \right| \\
 & \leq \sum_{a \in \mathcal{A}} \sum_{a' \in \mathcal{A}} L \left\| \sum_{t=1}^T \mathbb{1}[a_B^{t,k+1} = a, a_A^{t,k} = a'] (\hat{y}_B^{t,k+1} - y^t) \right\|_{\infty} \quad (\text{by } L\text{-Lipschitzness}) \\
 & \leq L \sum_{a \in \mathcal{A}} \sum_{a' \in \mathcal{A}} f_B(|T(k+1, a, a')|) \quad (\text{by } f_B\text{-decision conversation calibration}) \\
 & \leq L|\mathcal{A}|^2 f_B\left(\frac{T}{|\mathcal{A}|^2}\right) \quad (\text{by concavity of } f_B)
 \end{aligned}$$

1464 The second inequality follows from the fact that on at least $1 - \delta$ fraction of the days, the difference
 1465 in utility is at most ε . On the remaining days, the difference in utility is at most 1. \square

1466 Putting this all together, we can prove Theorem H.17.

1467 *Proof of Theorem H.17.* Let $\varepsilon = \left(\frac{1}{K-1} + \beta(T)\right)^{1/2}$. By Lemma H.21, there exists a round k such
 1468 that on $1 - \left(\frac{1}{(K-1)\varepsilon} + \frac{\beta(T)}{\varepsilon}\right)$ fraction of the days, Alice and Bob's actions are ε -approximate best
 1469 responses to each others' predictions. First, consider the case where k is odd, i.e. Alice communicates
 1470 on round k . We have that:

$$\left| u(a_A^{t,k}, \hat{y}_A^{t,k}) - u(a_B^{t,k+1}, \hat{y}_A^{t,k}) \right| \leq \varepsilon$$

1471 and

$$\left| u(a_A^{t,k}, \hat{y}_B^{t,k+1}) - u(a_B^{t,k+1}, \hat{y}_B^{t,k+1}) \right| \leq \varepsilon$$

1472 Since Alice has (f_A^S, \mathcal{C}_A) -decision conversation swap regret, by Lemma H.18, the transcript at round
 1473 k satisfies $\left(|\mathcal{A}| f_A^S\left(\frac{T}{|\mathcal{A}|}\right), \mathcal{C}_A\right)$ -decision swap regret. Next, we show that the transcript at round k
 1474 additionally has bounded decision swap regret with respect to \mathcal{C}_B .

1475 We can calculate the decision swap regret as:

$$\begin{aligned}
& \sum_{a \in \mathcal{A}} \max_{c \in \mathcal{C}_B} \sum_{t=1}^T \mathbb{1}[a_A^{t,k} = a] u(c(x_B^t), y^t) - \sum_{t=1}^T u(a_A^{t,k}, y^t) \\
& \leq \sum_{a \in \mathcal{A}} \max_{c \in \mathcal{C}_B} \sum_{t=1}^T \mathbb{1}[a_A^{t,k} = a] u(c(x_B^t), y^t) - \sum_{t=1}^T u(a_B^{t,k+1}, y^t) + \left(\varepsilon + \frac{1}{(K-1)\varepsilon} + \frac{\beta(T)}{\varepsilon} \right) T + L|\mathcal{A}|^2 f_B \left(\frac{T}{|\mathcal{A}|^2} \right) \\
& \quad \text{(by Lemma H.22)} \\
& = \sum_{a \in \mathcal{A}} \max_{c \in \mathcal{C}_B} \sum_{t=1}^T \mathbb{1}[a_A^{t,k} = a] u(c(x_B^t), y^t) - \sum_{t=1}^T u(a_B^{t,k+1}, y^t) + 2T \left(\frac{1}{(K-1)} + \beta(T) \right)^{1/2} + L|\mathcal{A}|^2 f_B \left(\frac{T}{|\mathcal{A}|^2} \right) \\
& \quad \text{(by our setting of } \varepsilon \text{)} \\
& = \sum_{a \in \mathcal{A}} \max_{c \in \mathcal{C}_B} \sum_{t=1}^T \mathbb{1}[a_A^{t,k} = a] \left(u(c(x_B^t), y^t) - u(a_B^{t,k+1}, y^t) \right) + 2T \left(\frac{1}{(K-1)} + \beta(T) \right)^{1/2} + L|\mathcal{A}|^2 f_B \left(\frac{T}{|\mathcal{A}|^2} \right) \\
& = \sum_{a \in \mathcal{A}} \max_{c \in \mathcal{C}_B} \sum_{a' \in \mathcal{A}} \sum_{t=1}^T \mathbb{1}[a_A^{t,k} = a, a_B^{t,k+1} = a'] \left(u(c(x_B^t), y^t) - u(a_B^{t,k+1}, y^t) \right) \\
& \quad + 2T \left(\frac{1}{(K-1)} + \beta(T) \right)^{1/2} + L|\mathcal{A}|^2 f_B \left(\frac{T}{|\mathcal{A}|^2} \right) \\
& \leq \sum_{a \in \mathcal{A}} \sum_{a' \in \mathcal{A}} \max_{c \in \mathcal{C}_B} \sum_{t=1}^T \mathbb{1}[a_A^{t,k} = a, a_B^{t,k+1} = a'] \left(u(c(x_B^t), y^t) - u(a_B^{t,k+1}, y^t) \right) \\
& \quad + 2T \left(\frac{1}{(K-1)} + \beta(T) \right)^{1/2} + L|\mathcal{A}|^2 f_B \left(\frac{T}{|\mathcal{A}|^2} \right) \\
& \leq \sum_{a \in \mathcal{A}} f_B^S(|T_A(k, a)|) + 2T \left(\frac{1}{(K-1)} + \beta(T) \right)^{1/2} + L|\mathcal{A}|^2 f_B \left(\frac{T}{|\mathcal{A}|^2} \right) \\
& \quad \text{(by } (f_B^S, \mathcal{C}_B)\text{-decision conversation swap regret)} \\
& \leq |\mathcal{A}| f_B^S \left(\frac{T}{|\mathcal{A}|} \right) + 2T \left(\frac{1}{(K-1)} + \beta(T) \right)^{1/2} + L|\mathcal{A}|^2 f_B \left(\frac{T}{|\mathcal{A}|^2} \right) \quad \text{(by concavity of } f_B^S \text{)}
\end{aligned}$$

1476 Here, the second inequality holds, since moving the max inside the sum can only make the quantity
1477 larger.

For brevity, let:

$$\lambda_A^{odd} := |\mathcal{A}| f_A^S \left(\frac{T}{|\mathcal{A}|} \right) \quad \text{and} \quad \lambda_B^{odd} := |\mathcal{A}| f_B^S \left(\frac{T}{|\mathcal{A}|} \right) + 2T \left(\frac{1}{(K-1)} + \beta(T) \right)^{1/2} + L|\mathcal{A}|^2 f_B \left(\frac{T}{|\mathcal{A}|^2} \right).$$

1478 Hence, the transcript at round k simultaneously has $(\lambda_A^{odd}, \mathcal{C}_A)$ -decision swap regret and $(\lambda_B^{odd}, \mathcal{C}_B)$ -
1479 decision swap regret. Therefore, it has $(\max\{\lambda_A^{odd}, \lambda_B^{odd}\}, \mathcal{C}_A \cup \mathcal{C}_B)$ -decision swap regret.

1480 Now, consider the case where k is even. Since all statements hold symmetrically, we have that, for:

$$\lambda_A^{even} := |\mathcal{A}| f_A^S \left(\frac{T}{|\mathcal{A}|} \right) + 2T \left(\frac{1}{(K-1)} + \beta(T) \right)^{1/2} + L|\mathcal{A}|^2 f_A \left(\frac{T}{|\mathcal{A}|^2} \right) \quad \text{and} \quad \lambda_B^{even} := |\mathcal{A}| f_B^S \left(\frac{T}{|\mathcal{A}|} \right),$$

1481 the transcript at round k simultaneously has $(\lambda_A^{even}, \mathcal{C}_A)$ -decision swap regret and $(\lambda_B^{even}, \mathcal{C}_B)$ -
1482 decision swap regret, and therefore $(\max\{\lambda_A^{even}, \lambda_B^{even}\}, \mathcal{C}_A \cup \mathcal{C}_B)$ -decision swap regret.

1483 Since $\lambda^{even} \geq \lambda^{odd}$ and $\lambda^{odd} \geq \lambda^{even}$, we can conclude that there exists a round k such that the
1484 transcript at round k has $(\max\{\lambda_A^{even}, \lambda_B^{odd}\}, \mathcal{C}_A \cup \mathcal{C}_B)$ -decision swap regret. \square

1485 Theorem H.17 shows that at some intermediate round, the transcript has bounded decision swap
1486 regret with respect to $\mathcal{C}_A \cup \mathcal{C}_B$. Our boosting result (Theorem H.16) states that if, additionally, \mathcal{C}_A
1487 and \mathcal{C}_B are weak learners for \mathcal{C}_J , then the transcript also has bounded decision swap regret with

respect to \mathcal{C}_J . Together, these results imply that at an intermediate round, the transcript has bounded decision swap regret with respect to \mathcal{C}_J .

One difficulty is that Alice and Bob will not know a priori which intermediate round will have these guarantees — and so it is not clear a priori which downstream action to take on any day. However, we will use a similar argument as we did in the proof of Theorem C.3 to argue that the transcript on the *last* round inherits an external regret guarantee. That is, as long as Alice and Bob act according to the last round, they are sure to achieve bounded external regret with respect to \mathcal{C}_J .

Theorem H.23. Fix a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$. Let \mathcal{C}_J be a policy class over the joint feature space \mathcal{X} . Let $\mathcal{C}_A = \{c_A : \mathcal{X}_A \rightarrow \mathcal{A}\}$ and $\mathcal{C}_B = \{c_B : \mathcal{X}_B \rightarrow \mathcal{A}\}$ be policy classes over \mathcal{X}_A and \mathcal{X}_B respectively. Fix a transcript $\pi^{1:T, 1:K}$ generated via Protocol H.1. If:

- Alice has (f_A^S, \mathcal{C}_A) -decision conversation swap regret and f_A -decision conversation calibration
- Bob has (f_B^S, \mathcal{C}_B) -decision conversation swap regret and f_B -decision conversation calibration
- \mathcal{C}_A and \mathcal{C}_B jointly satisfy the $w(\cdot)$ -weak learning condition with respect to \mathcal{C}_J

Then, there exists a round k of the protocol such that the transcript $\pi^{1:T, k}$ has $\left(2Tw^{-1}\left(\frac{\max\{\lambda_A, \lambda_B\}}{T}\right), \mathcal{C}_J\right)$ -decision swap regret, whenever the inverse of w exists. Moreover, on the last round K , the transcript $\pi^{1:T, K}$ satisfies:

$$\max_{c_J \in \mathcal{C}_J} \sum_{t=1}^T u(c_J(x^t), y^t) - \sum_{t=1}^T u(a^{t, K}, y^t) \leq 2Tw^{-1}\left(\frac{\max\{\lambda_A, \lambda_B\}}{T}\right) + (K-1)T\beta(T)$$

whenever the inverse of w exists. Here,

$$\lambda_A \leq |\mathcal{A}|f_A^S\left(\frac{T}{|\mathcal{A}|}\right) + L|\mathcal{A}|^2f_A\left(\frac{T}{|\mathcal{A}|^2}\right) + 2T\left(\frac{1}{(K-1)} + \beta(T)\right)^{1/2},$$

$$\lambda_B \leq |\mathcal{A}|f_B^S\left(\frac{T}{|\mathcal{A}|}\right) + L|\mathcal{A}|^2f_B\left(\frac{T}{|\mathcal{A}|^2}\right) + 2T\left(\frac{1}{(K-1)} + \beta(T)\right)^{1/2}$$

where $\beta(T) = \frac{L|\mathcal{A}|^2}{T}\left(f_A\left(\frac{T}{|\mathcal{A}|^2}\right) + f_B\left(\frac{T}{|\mathcal{A}|^2}\right)\right)$.

Proof. By Theorem H.17, there exists a round k^* of the protocol such that the transcript $\pi^{1:T, k^*}$ has $(\max\{\lambda_A, \lambda_B\}, \mathcal{C}_A \cup \mathcal{C}_B)$ -decision swap regret. Then, since \mathcal{C}_A and \mathcal{C}_B satisfy the $w(\cdot)$ -weak learning condition, Theorem H.16 gives us that $\pi^{1:T, k^*}$ has $\left(2Tw^{-1}\left(\frac{\max\{\lambda_A, \lambda_B\}}{T}\right), \mathcal{C}_J\right)$ -decision swap regret. This proves the first part of the theorem.

To prove the second part, we use Lemma H.19, which bounds the decrease in utility from every round k to $k+1$. We have that over two rounds, the change in utility is:

$$\begin{aligned} & \sum_{t=1}^T u(a_A^{t, k}, y^t) - \sum_{t=1}^T u(a_A^{t, k+2}, y^t) \\ &= \sum_{t=1}^T u(a_A^{t, k}, y^t) - \sum_{t=1}^T u(a_B^{t, k+1}, y^t) + \sum_{t=1}^T u(a_B^{t, k+1}, y^t) - \sum_{t=1}^T u(a_A^{t, k+2}, y^t) \\ &\leq 2L|\mathcal{A}|^2f_B\left(\frac{T}{|\mathcal{A}|^2}\right) - \varepsilon|D(T^{k+1})| + 2L|\mathcal{A}|^2f_A\left(\frac{T}{|\mathcal{A}|^2}\right) - \varepsilon|D(T^{k+2})| \quad (\text{by Lemma H.19}) \\ &\leq 2L|\mathcal{A}|^2f_B\left(\frac{T}{|\mathcal{A}|^2}\right) + 2L|\mathcal{A}|^2f_A\left(\frac{T}{|\mathcal{A}|^2}\right) \\ &= 2T\beta(T) \quad (\text{by definition of } \beta(T)) \end{aligned}$$

Thus, we can bound the decrease in utility by applying this expression iteratively from round k^* to the last round K . There are at most $K - 1$ rounds between k^* and K , and so applying this expression $(K - 1)/2$ times bounds the decrease in utility, i.e.:

$$\sum_{t=1}^T u(a^{t,k^*}, y^t) - \sum_{t=1}^T u(a^{t,K}, y^t) \leq \frac{K-1}{2} \cdot 2T\beta(T) = (K-1)T\beta(T)$$

Therefore, we can bound the external regret of the last round:

$$\begin{aligned} & \max_{c_J \in \mathcal{C}_J} \sum_{t=1}^T u(c_J(x^t), y^t) - \sum_{t=1}^T u(a^{t,K}, y^t) \\ & \leq \max_{c_J \in \mathcal{C}_J} \sum_{t=1}^T u(c_J(x^t), y^t) - \sum_{t=1}^T u(a^{t,k^*}, y^t) + (K-1)T\beta(T) \\ & \leq 2Tw^{-1} \left(\frac{\max\{\lambda_A, \lambda_B\}}{T} \right) + (K-1)T\beta(T) \end{aligned}$$

Here, the last line follows from the fact that external regret is upper bounded by decision swap regret, and we have previously bound the decision swap regret of the transcript at round k^* . This completes the proof. \square

H.5 Achieving Conversation Decision Cross Calibration Algorithmically

Finally, we turn attention to an algorithm that obtains low decision conversation swap regret and low decision conversation calibration; this will allow us to instantiate our results with concrete regret bounds. We use the algorithm of Lu et al. [2025], which guarantees diminishing decision calibration and decision cross calibration error and thus, by Theorem H.11, diminishing decision swap regret.

Theorem H.24 (Theorem 2 of Lu et al. [2025]). *Fix a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$. Fix a policy class \mathcal{C} . There is an algorithm that with probability $1 - \rho$, for any sequence of outcomes y^1, \dots, y^T , outputs predictions $\hat{y}^1, \dots, \hat{y}^T$ that are f -decision calibrated and (f, \mathcal{C}) -decision cross calibrated, where:*

$$f(\tau) \leq O \left(\ln(d|\mathcal{A}||\mathcal{C}|T) + \sqrt{T \ln \left(\frac{d|\mathcal{A}||\mathcal{C}|T}{\rho} \right)} \right)$$

for any $\tau \in [T]$.

To guarantee diminishing decision conversation swap regret and decision conversation calibration, we instantiate a copy of this algorithm for each pair of rounds k and actions a . On round k of day t , we call on the copy corresponding to that round and the action chosen in the previous round on that day. This gives us precisely what we want: diminishing decision swap regret and decision calibration, conditioned on every round *and* the most recently communicated action. This reduction is formalized in Algorithm H.5 (here, we take the perspective of Alice; Bob's is symmetric).

[ht] **Input** Algorithm M , policy class \mathcal{C}

For every odd $k \in [K]$ and $a \in \mathcal{A}$, instantiate a copy of M , called $M_{k,a}$. For the first round $k = 1$, instantiate a copy of M , called M_1 .

Let $\pi^{1:t,k|a}$ denote the transcript on round k up until day t , restricted to $\{t : a^{t,k-1} = a\}$, the subsequence where the previously communicated action was a .

Let $M(\pi^{1:t,k|a}, \mathcal{C})$ denote the output of M given this transcript.

each day $t = 1, \dots, T$ Receive x_A^t Make prediction $\hat{y}_A^{t,1} = M_1(\pi^{1:t-1,1}, \mathcal{C})$ Send to Bob $a_A^{t,1} = \text{BR}_u(\hat{y}_A^{t,1})$

each odd round $k = 3, 5, \dots, K$ Observe Bob's action from the previous round $a_B^{t,k-1}$ Make prediction $\hat{y}_A^{t,k} = M_{k,a_B^{t,k-1}}(\pi^{1:t-1,k|a_B^{t,k-1}}, \mathcal{C})$ Send to Bob $a_A^{t,k} = \text{BR}_u(\hat{y}_A^{t,k})$ Observe $y^t \in \mathcal{Y}$.

1548 **Theorem H.25.** Fix a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$. Fix a policy class \mathcal{C} . With probability $1 - \rho$,
 1549 Algorithm H.5, instantiated with the algorithm of Theorem H.24 and \mathcal{C} , obtains (f^S, \mathcal{C}) -decision
 1550 conversation swap regret and f -decision conversation calibration for:

$$f^S(\tau) \leq O \left(L|\mathcal{A}|^2 \ln(d|\mathcal{A}||\mathcal{C}|T) + L|\mathcal{A}| \sqrt{T \ln \left(\frac{dK|\mathcal{A}||\mathcal{C}|T}{\rho} \right)} \right)$$

1551 and

$$f(\tau) \leq O \left(\ln(d|\mathcal{A}||\mathcal{C}|T) + \sqrt{T \ln \left(\frac{dK|\mathcal{A}||\mathcal{C}|T}{\rho} \right)} \right)$$

1552 for any $\tau \in [T]$.

1553 *Proof.* Let M be the algorithm of Theorem H.24. Let $\rho' = \frac{2\rho}{K|\mathcal{A}|}$. By Theorem H.24, with probability
 1554 $1 - \rho'$, M produces predictions that are f -decision calibrated for:

$$f(\tau) \leq O \left(\ln(d|\mathcal{A}||\mathcal{C}|T) + \sqrt{T \ln \left(\frac{d|\mathcal{A}||\mathcal{C}|T}{\rho'} \right)} \right)$$

1555 Moreover, plugging the guarantees of M into Theorem H.11, we have that with probability $1 - \rho'$,
 1556 M obtains (f^S, \mathcal{C}) -decision swap regret for:

$$f^S(\tau) \leq O \left(L|\mathcal{A}|^2 \ln(d|\mathcal{A}||\mathcal{C}|T) + L|\mathcal{A}| \sqrt{T \ln \left(\frac{d|\mathcal{A}||\mathcal{C}|T}{\rho'} \right)} \right)$$

1557 By construction, on every odd round k , a separate copy $M_{k,a}$ is run for every subsequence on which
 1558 the action from the previous round $a^{t,k-1}$ is a . By a union bound, the probability that any one of
 1559 the copies fails is at most $\frac{K}{2}|\mathcal{A}|\rho' = \rho$. Therefore, since decision conversation calibration asks
 1560 for decision calibration on every such subsequence, with probability $1 - \rho$, Algorithm H.5 is also
 1561 (f, \mathcal{C}) -decision conversation calibrated. Likewise, since decision conversation swap regret measures
 1562 the decision swap regret on every such subsequence, with probability $1 - \rho$, Algorithm H.5 also
 1563 achieves (f^S, \mathcal{C}) -decision conversation swap regret. \square

1564 To end this section, we instantiate Theorem H.23 with the algorithmic bounds. As before, we face a
 1565 tradeoff in the choice of K , the length of the conversation. We show that for appropriately chosen K ,
 1566 we guarantee sublinear regret bounds with respect to \mathcal{C}_J .

1567 **Theorem H.26.** Fix a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$. Let \mathcal{C}_J be a policy class over the joint
 1568 feature space \mathcal{X} . Let $\mathcal{C}_A = \{c_A : \mathcal{X}_A \rightarrow \mathcal{A}\}$ and $\mathcal{C}_B = \{c_B : \mathcal{X}_B \rightarrow \mathcal{A}\}$ be policy classes over \mathcal{X}_A
 1569 and \mathcal{X}_B respectively. Suppose Alice and Bob interact via Protocol H.1. If:

- 1570 • Both Alice and Bob use Algorithm H.5, instantiated with the algorithm of Theorem H.24
 1571 and policy classes \mathcal{C}_A and \mathcal{C}_B respectively
- 1572 • \mathcal{C}_A and \mathcal{C}_B jointly satisfy the $w(\cdot)$ -weak learning condition with respect to \mathcal{C}_J

1573 Then, with probability $1 - \rho$, the transcript $\pi^{1:T,K}$ on the last round K satisfies:

$$\begin{aligned} \max_{c_J \in \mathcal{C}_J} \sum_{t=1}^T u(c_J(x^t), y^t) - \sum_{t=1}^T u(a^{t,K}, y^t) &\leq 2Tw^{-1} \left(O \left(\frac{L|\mathcal{A}|^3 \ln \left(\frac{dK|\mathcal{A}||\mathcal{C}_A||\mathcal{C}_B|T}{\rho} \right)}{T^{1/4}} + \frac{1}{\sqrt{K-1}} \right) \right) \\ &\quad + O \left((K-1)L|\mathcal{A}|^2 \ln \left(\frac{dK|\mathcal{A}||\mathcal{C}_A||\mathcal{C}_B|T}{\rho} \right) \sqrt{T} \right) \end{aligned}$$

1574 whenever the inverse of w exists.

Moreover, if $K = \omega(1)$ and $K = o(\sqrt{T})$, then the transcript $\pi^{1:T,K}$ satisfies, for some constant $\alpha \in (0, 1)$:

$$\begin{aligned} \max_{c_J \in \mathcal{C}_J} \sum_{t=1}^T u(c_J(x^t), y^t) - \sum_{t=1}^T u(a^{t,K}, y^t) &\leq 2Tw^{-1} \left(O \left(\frac{L|\mathcal{A}|^3 \ln \left(\frac{d|\mathcal{A}||\mathcal{C}_A||\mathcal{C}_B|T}{\rho} \right)}{T^{1/4}} + o(1) \right) \right) \\ &\quad + O \left(L|\mathcal{A}|^2 \ln \left(\frac{d|\mathcal{A}||\mathcal{C}_A||\mathcal{C}_B|T}{\rho} \right) T^\alpha \right) \\ &\leq o(T) \end{aligned}$$

That is, the transcript at the last round achieves sublinear external regret with respect to \mathcal{C}_J .

Proof. Let $\rho' = \rho/2$. By Theorem H.25, Algorithm H.5 achieves, with probability $1 - \rho'$, (f^S, \mathcal{C}_A) -decision conversation swap regret and f_A -decision conversation calibration for:

$$f_A^S(\tau) \leq O \left(L|\mathcal{A}|^2 \ln(d|\mathcal{A}||\mathcal{C}_A|T) + L|\mathcal{A}| \sqrt{T \ln \left(\frac{dK|\mathcal{A}||\mathcal{C}_A|T}{\rho} \right)} \right)$$

and

$$f_A(\tau) \leq O \left(\ln(d|\mathcal{A}||\mathcal{C}_A|T) + \sqrt{T \ln \left(\frac{dK|\mathcal{A}||\mathcal{C}_A|T}{\rho} \right)} \right)$$

for any $\tau \in [T]$. Likewise, Algorithm H.5 achieves, with probability $1 - \rho'$, (f_B^S, \mathcal{C}_B) -decision conversation swap regret and f_B -decision conversation calibration for:

$$f_B^S(\tau) \leq O \left(L|\mathcal{A}|^2 \ln(d|\mathcal{A}||\mathcal{C}_B|T) + L|\mathcal{A}| \sqrt{T \ln \left(\frac{dK|\mathcal{A}||\mathcal{C}_B|T}{\rho} \right)} \right)$$

and

$$f_B(\tau) \leq O \left(\ln(d|\mathcal{A}||\mathcal{C}_B|T) + \sqrt{T \ln \left(\frac{dK|\mathcal{A}||\mathcal{C}_B|T}{\rho} \right)} \right)$$

Thus, by a union bound, if Alice and Bob both use Algorithm H.5 to interact, then with probability $1 - 2\rho' = 1 - \rho$, Alice has (f_A^S, \mathcal{C}_A) -decision conversation swap regret and f_A -decision conversation calibration, and Bob has (f_B^S, \mathcal{C}_B) -decision conversation swap regret and f_B -decision conversation calibration.

Then, by Theorem H.23, the transcript $\pi^{1:T,K}$ on the last round satisfies:

$$\max_{c_J \in \mathcal{C}_J} \sum_{t=1}^T u(c_J(x^t), y^t) - \sum_{t=1}^T u(a^{t,K}, y^t) \leq 2Tw^{-1} \left(\frac{\max\{\lambda_A, \lambda_B\}}{T} \right) + (K-1)T\beta(T)$$

where:

$$\begin{aligned} \beta(T) &= \frac{L|\mathcal{A}|^2}{T} \left(f_A \left(\frac{T}{|\mathcal{A}|^2} \right) + f_B \left(\frac{T}{|\mathcal{A}|^2} \right) \right) \\ &\leq O \left(\frac{L|\mathcal{A}|^2 \ln(d|\mathcal{A}||\mathcal{C}_A|T)}{T} + L|\mathcal{A}|^2 \sqrt{\frac{\ln \left(\frac{dK|\mathcal{A}||\mathcal{C}_A|T}{\rho} \right)}{T}} + \frac{L|\mathcal{A}|^2 \ln(d|\mathcal{A}||\mathcal{C}_B|T)}{T} + L|\mathcal{A}|^2 \sqrt{\frac{\ln \left(\frac{dK|\mathcal{A}||\mathcal{C}_B|T}{\rho} \right)}{T}} \right) \\ &\leq O \left(\frac{L|\mathcal{A}|^2 \ln(d|\mathcal{A}||\mathcal{C}_A||\mathcal{C}_B|T)}{T} + L|\mathcal{A}|^2 \sqrt{\frac{\ln \left(\frac{dK|\mathcal{A}||\mathcal{C}_A||\mathcal{C}_B|T}{\rho} \right)}{T}} \right) \end{aligned}$$

(by Cauchy-Schwartz)

1590 and thus:

$$\begin{aligned}
\lambda_A &\leq |\mathcal{A}| f_A^S \left(\frac{T}{|\mathcal{A}|} \right) + L|\mathcal{A}|^2 f_A \left(\frac{T}{|\mathcal{A}|^2} \right) + 2T \left(\frac{1}{(K-1)} + \beta(T) \right)^{1/2} \\
&\leq O \left(L|\mathcal{A}|^3 \ln(d|\mathcal{A}||\mathcal{C}_A|T) + L|\mathcal{A}|^2 \sqrt{T \ln \left(\frac{dK|\mathcal{A}||\mathcal{C}_A|T}{\rho} \right)} \right. \\
&\quad \left. + \frac{T}{\sqrt{K-1}} + |\mathcal{A}| \sqrt{TL \ln(d|\mathcal{A}||\mathcal{C}_A||\mathcal{C}_B|T)} + |\mathcal{A}| \sqrt{L} \ln^{1/4} \left(\frac{dK|\mathcal{A}||\mathcal{C}_A||\mathcal{C}_B|T}{\rho} \right) T^{3/4} \right) \\
&\quad \text{(by concavity of the square root function)} \\
&\leq O \left(L|\mathcal{A}|^3 \ln(d|\mathcal{A}||\mathcal{C}_A||\mathcal{C}_B|T) + L|\mathcal{A}|^2 \sqrt{\ln \left(\frac{dK|\mathcal{A}||\mathcal{C}_A||\mathcal{C}_B|T}{\rho} \right) T^{3/4}} + \frac{T}{\sqrt{K-1}} \right)
\end{aligned}$$

1591 Since the expression for λ_B is symmetric, we have that:

$$\lambda_B \leq O \left(L|\mathcal{A}|^3 \ln(d|\mathcal{A}||\mathcal{C}_A||\mathcal{C}_B|T) + L|\mathcal{A}|^2 \sqrt{\ln \left(\frac{dK|\mathcal{A}||\mathcal{C}_A||\mathcal{C}_B|T}{\rho} \right) T^{3/4}} + \frac{T}{\sqrt{K-1}} \right)$$

1592 Hence, plugging this in, we can compute:

$$\begin{aligned}
&\max_{c_J \in \mathcal{C}_J} \sum_{t=1}^T u(c_J(x^t), y^t) - \sum_{t=1}^T u(a^{t,K}, y^t) \\
&\leq 2Tw^{-1} \left(\frac{\max\{\lambda_A, \lambda_B\}}{T} \right) + (K-1)T\beta(T) \\
&\leq 2Tw^{-1} \left(O \left(\frac{L|\mathcal{A}|^3 \ln(d|\mathcal{A}||\mathcal{C}_A||\mathcal{C}_B|T)}{T} + \frac{L|\mathcal{A}|^2 \sqrt{\ln \left(\frac{dK|\mathcal{A}||\mathcal{C}_A||\mathcal{C}_B|T}{\rho} \right)}}{T^{1/4}} + \frac{1}{\sqrt{K-1}} \right) \right) \\
&\quad + O \left((K-1)L|\mathcal{A}|^2 \ln(d|\mathcal{A}||\mathcal{C}_A||\mathcal{C}_B|T) + (K-1)L|\mathcal{A}|^2 \sqrt{T \ln \left(\frac{dK|\mathcal{A}||\mathcal{C}_A||\mathcal{C}_B|T}{\rho} \right)} \right) \\
&\leq 2Tw^{-1} \left(O \left(\frac{L|\mathcal{A}|^3 \ln \left(\frac{dK|\mathcal{A}||\mathcal{C}_A||\mathcal{C}_B|T}{\rho} \right)}{T^{1/4}} + \frac{1}{\sqrt{K-1}} \right) \right) \\
&\quad + O \left((K-1)L|\mathcal{A}|^2 \ln \left(\frac{dK|\mathcal{A}||\mathcal{C}_A||\mathcal{C}_B|T}{\rho} \right) \sqrt{T} \right)
\end{aligned}$$

1593 which proves the first part of the theorem.

1594 To prove the second part, suppose $K = \omega(1)$ and $K = o(\sqrt{T})$. Then, we can compute:

$$\begin{aligned}
&\max_{c_J \in \mathcal{C}_J} \sum_{t=1}^T u(c_J(x^t), y^t) - \sum_{t=1}^T u(a^{t,K}, y^t) \\
&\leq 2Tw^{-1} \left(O \left(\frac{L|\mathcal{A}|^3 \ln \left(\frac{d|\mathcal{A}||\mathcal{C}_A||\mathcal{C}_B|T}{\rho} \right)}{T^{1/4}} + o(1) \right) \right) + O \left(L|\mathcal{A}|^2 \ln \left(\frac{d|\mathcal{A}||\mathcal{C}_A||\mathcal{C}_B|T}{\rho} \right) T^\alpha \right)
\end{aligned}$$

1595 for some constant $\alpha \in (0, 1)$. Now, observe that any function $O \left(\frac{L|\mathcal{A}|^3 \ln \left(\frac{d|\mathcal{A}||\mathcal{C}_A||\mathcal{C}_B|T}{\rho} \right)}{T^{1/4}} + o(1) \right) \rightarrow$

1596 0 as $T \rightarrow \infty$. Hence, by Lemma K.18, $w^{-1} \left(O \left(\frac{L|\mathcal{A}|^3 \ln \left(\frac{d|\mathcal{A}||\mathcal{C}_A||\mathcal{C}_B|T}{\rho} \right)}{T^{1/4}} + o(1) \right) \right) \rightarrow 0$ as $T \rightarrow \infty$

1597 and thus,

$$Tw^{-1} \left(O \left(\frac{L|\mathcal{A}|^3 \ln \left(\frac{d|\mathcal{A}||\mathcal{C}_A||\mathcal{C}_B|T}{\rho} \right)}{T^{1/4}} + o(1) \right) \right) = o(T)$$

1598 Notice that since w is strictly increasing, w^{-1} exists for sufficiently large T (larger than a constant).
 1599 Therefore, for sufficiently large T , the regret is bounded by:

$$\max_{c_J \in \mathcal{C}_J} \sum_{t=1}^T u(c_J(x^t), y^t) - \sum_{t=1}^T u(a^{t,K}, y^t) \leq o(T) + O\left(L|\mathcal{A}|^2 \ln\left(\frac{d|\mathcal{A}||\mathcal{C}_A||\mathcal{C}_B|T}{\rho}\right) T^\alpha\right)$$

1600 which completes the proof. \square

1601 I Collaboration in the Batch Setting

1602 Thus far, we have studied the *online* setting, in which participants jointly predict the label on a new
 1603 adversarially chosen example every day. However, we can also study this form of collaborative
 1604 learning in the simpler *distributional* or *batch* setting, where Alice and Bob both receive different
 1605 features x_A and x_B drawn from a distribution. They will train on a sample of such data (paired
 1606 with labels) at training time, and then at test time (when labels are unavailable) will be evaluated on
 1607 examples drawn from the same distribution. This setting is strictly easier than the online adversarial
 1608 setting, and hence admits (morally if not notationally) simpler algorithms which we develop in this
 1609 section.

1610 At a high level, the algorithm here will proceed over R rounds that we index by r . In the training
 1611 phase, Alice and Bob will iteratively build their models as follows:

- 1612 • Bob will begin by generating an initial model and sending his model's initial predictions for
 1613 all of the points in the training set, P^0 , to Alice. These predictions will be discretized to a
 1614 finite range.
- 1615 • In the next round, Alice will refine her model according to Bob's predictions:
 - 1616 – First, she will bucket her data into level sets *according to Bob's predictions*. “Level set
 1617 v ” corresponds to all the points in the training set for which Bob predicted v .
 - 1618 – On each level set v in parallel, Alice will run an internal boosting procedure which
 1619 we call INTERNAL-BOOST with respect to her hypothesis class (defined only on her
 1620 own features), generating a model $\tilde{f}_A^{1,v}$. This internal boosting process is equivalent
 1621 to the LSBoost algorithm from Globus-Harris et al. [2023]. In essence, it repeatedly
 1622 performs squared error regression over \mathcal{H}_A on Alice's own level sets, until doing so no
 1623 longer substantially improves squared error. This procedure results in a (discretized)
 1624 ensemble of models from \mathcal{H}_A defined in parallel for each of the v level sets.
 - 1625 – For each level set v , Alice will look at the error of her resulting model on that level
 1626 set $\tilde{f}_A^{1,v}$, and compare it to the error of Bob's (constant) predictor v constrained to that
 1627 level set. Depending on whether her predictions improve substantially over Bob's, she
 1628 will either set $\tilde{f}_A^{1,v}$ to $\tilde{f}_A^{1,v}$ or to the constant predictor v (i.e. “agreeing” with Bob's
 1629 predictions on that levelset).
 - 1630 – She will define her final predictor at the end of round 1, f_A^1 , as an ensemble of these
 1631 models such that if a datapoint $x = (x_A, x_B)$ is given predicted label v by Bob's initial
 1632 predictor, $f_A^1(x_A) = \tilde{f}_A^{1,v}(x_A)$.
- 1633 • She will then evaluate f_A^1 on every point in the training sample and send the resulting
 1634 predictions P^1 to Bob.
- 1635 • In the next round, Bob will run a symmetric procedure using Alice's predictions P^1 . They
 1636 will continue in this manner in rounds until the predictions have converged to agreement.

1637 During this process, Alice and Bob will separately maintain transcripts of the models which they
 1638 have iteratively built across the rounds of communication. At test time, to make a prediction on a new
 1639 datapoint with features $x = (x_A, x_B)$ partitioned across Alice and Bob, they will again engage in an
 1640 interactive conversation, at each round making predictions according to the models recorded in the
 1641 transcript that was generated during training. This will proceed as follows:

- 1642 • Bob will look at his model transcript, extract his initial model, and evaluate it on x_B . He
 1643 will then send the prediction to Alice.

- Alice will extract from her transcript the model f^{1,v^*} corresponding to the value of Bob's prediction v^* , and send her prediction $f^{1,v^*}(x_A)$ to Bob.
- They will proceed in this manner across rounds until they have evaluated the final models stored in their transcripts, whose predictions they will output.

1.1 Preliminaries for the Batch Setting

Formally, as in the online setting, Alice and Bob have feature spaces \mathcal{X}_A and \mathcal{X}_B and there is a real-valued outcome space \mathcal{Y} . We now additionally assume that there is a joint distribution $\mathcal{D} \in \Delta(\mathcal{X}_A \times \mathcal{X}_B \times \mathcal{Y})$ from which examples are drawn. We will write \mathcal{D}_A to denote the marginal distribution over $(\mathcal{X}_A, \mathcal{Y})$ and \mathcal{D}_B to denote the marginal distribution over $(\mathcal{X}_B, \mathcal{Y})$.

1.1.1 Training Phase

In the training phase, a finite training set $S = \{(x_A^i, x_B^i, y_i)\}_{i \in [n]} \sim \mathcal{D}^n$ of size n is sampled i.i.d, where we write $[n]$ to denote $\{1, \dots, n\}$. Alice is given $S_A = \{(x_A^i, y^i)\}_{i \in [n]}$ and Bob is given $S_B = \{(x_B^i, y^i)\}_{i \in [n]}$. Importantly, i here indexes over the *same instances* whose features are split between parties: $x^i = (x_A^i, x_B^i)$. Over their rounds of communication, Alice and Bob's models will be generated by ensembling hypotheses h_A and h_B respectively in hypothesis classes \mathcal{H}_A and \mathcal{H}_B , where $h_A : \mathcal{X}_A \rightarrow \mathbb{R}$ and $h_B : \mathcal{X}_B \rightarrow \mathbb{R}$. In particular, we will assume that they generate these hypotheses via access to a *squared error regression oracle*:

Definition I.1. We say $\mathcal{O}_{\mathcal{H}} : \Delta(\mathcal{X} \times \mathcal{Y}) \rightarrow (\mathcal{X} \rightarrow \mathcal{Y})$ is a *squared error regression oracle* for a class of real-valued functions \mathcal{H} if for every distribution $\mathcal{D} \in \Delta(\mathcal{X} \times \mathcal{Y})$, $\mathcal{O}_{\mathcal{H}}$ outputs the squared-error minimizing function $h \in \mathcal{H}$ over the distribution. I.e., if $h = \mathcal{O}_{\mathcal{H}}(\mathcal{D})$ then

$$h \in \arg \min_{h' \in \mathcal{H}} \mathbb{E}_{(x,y) \sim \mathcal{D}} [(h'(x) - y)^2].$$

When we feed such an oracle a sample $S = (x_i, y_i)_{i \in [n]}$, we will interpret these expectations as over the sample.

Across their interactions, Alice and Bob will round their predictions to some discretization, defined by a discretization parameter $m \in \mathbb{Z}^+$. We will write $[1/m] := \{0, \frac{1}{m}, \dots, \frac{m-1}{m}, 1\}$ be a discretization of the range $[0, 1]$ into multiples of $1/m$. They will round their predictions as follows:

Definition I.2 ($\text{Round}(h; m)$). Let \mathcal{F} be the collection of all real valued functions from features $\mathcal{X} \rightarrow \mathbb{R}$. Then Round is a function $\text{Round} : \mathcal{F} \times \mathbb{Z}^+ \rightarrow \mathcal{F}$ where $\text{Round}(h; m)$ outputs a function \tilde{h} such that

$$\tilde{h}(x) = \min_{v \in [1/m]} |h(x) - v|.$$

During training, Alice and Bob will separately generate *model transcripts* of the models they have generated so far, which they will use to construct predictions of the model out of sample. In essence, these model transcripts are simply a collection of models in \mathcal{H}_A and \mathcal{H}_B respectively, with the exception that in some rounds, their algorithm will generate \perp instead of a model (indicating that they are deferring to their counter-party's prediction).

Definition I.3 (Transcript). Let \mathcal{H}_A be Alice's hypothesis class and let $m \in \mathbb{Z}^+$. Over her R rounds of interaction with Bob, she will within each round run an internal algorithm in parallel m times. This internal algorithm will either return \perp or run for at most $K \in \mathbb{Z}^+$ phases. Over the course of these interactions she will generate her model transcript, which is an object over both her interactions with Bob and her internal algorithm:

$$\Pi_A^R = \{\pi_A^0, \dots, \pi_A^R\} \in (\{\perp\} \cup \mathcal{H}_A^{Km})^{mR},$$

where for each round $r \in [R]$, we have

$$\pi_A^r = \{\pi_A^{r,v}\}_{v \in [1/m]}$$

1683 and each of these sub-transcripts $\pi_A^{r,v}$ describes the (at most) K phases of each of Alice's internal
 1684 algorithm.³

$$\pi_A^{r,v} \in \{\perp\} \cup \left\{ (\pi_A^{r,v,k})_{k \in [K]} \right\},$$

1685 and each

$$\pi_A^{r,v,k} = (h_A^{r,v,k,v'})_{v' \in [1/m]}, \quad \text{with } h_A^{r,v,k,v'} \in \mathcal{H}_A.$$

1686 Bob's transcript Π_B^R will be defined analogously.

1687 Alice and Bob act in alternating rounds, Alice in even rounds and Bob in odd ones. At the end of
 1688 each round r , the active player sends the other their current predictions on the training set (which are
 1689 all discretized to lie in $[1/m]$).

1690 **Definition I.4** (Prediction at round r). We will write $P^r \in [1/m]^n$ to be the n predictions generated
 1691 at round r for each $x^i \in S$. If r is odd, $P^r = P_A^r$ are Alice's predictions, and if r is even, $P^r = P_B^r$
 1692 are Bob's predictions. In our analyses, we will denote the i th prediction in the vector P^r as $P^{r,i}$.

1693 At the end of R rounds, Alice and Bob will know a collection of predictions

$$\mathcal{P}^R = (P^0, \dots, P^R) = \begin{cases} P_B^0, \dots, P_B^{R-1}, P_A^R & \text{if } R \text{ is even,} \\ P_B^0, \dots, P_A^{R-1}, P_B^R & \text{else.} \end{cases}$$

1694 **Remark I.5.** Note that the dimension of these predictions P^r is different than in the online setting.
 1695 There, only a single prediction $\hat{y}^{r,k}$ is communicated between the players in their conversation. Here,
 1696 we have a set of n predictions communicated in each round — one for each point in the training set.

1697 At round r , Alice will generate a model f_A^r . In the training algorithm defined in Section I.2.1, this
 1698 model will be only well-defined defined for the training sample; in Section I.2.2 we will discuss how
 1699 to generate predictions on new data using the training transcript.

1700 **Definition I.6** (Model at round r). At round r of training, Alice will generate a model $f_A^r : \mathcal{X}_A \times$
 1701 $[1/m] \rightarrow [1/m]$ which is based on her datapoint and Bob's prediction from the previous round. In
 1702 general this model will be invoked in contexts where Bob's prediction v is clearly defined so we will
 1703 write

$$f_A^r(x_A) = f_A^r(x_A, v).$$

1704 Bob's model f_B^r will be defined analogously.

1705 **Definition I.7.** At the final round R of our collaboration algorithm *COLLABORATE* (Algorithm
 1706 I.2.1), Alice and Bob will have two models f_A^R and f_B^R which will agree for all datapoints on both the
 1707 training sample and at test time, so we can equivalently consider them as represented by a single
 1708 model f^R . We will write \mathcal{F}^R to be the space of models which may be output by the collaboration
 1709 algorithm on samples of size n , i.e.,

$$\mathcal{F}^R = \{f^R | f^R \leftarrow \text{COLLABORATE}((S_A, S_B), \mathcal{O}_{\mathcal{H}_A}, \mathcal{O}_{\mathcal{H}_B}, m)\}_{(S_A, S_B) = \{(x_A^i, y^i), (x_B^i, y^i)\}_{i \in [n]}},$$

1710 where S_A and S_B have been generated from a joint sample $S \in (\mathcal{X}_A \times \mathcal{X}_B \times \mathcal{Y})^n$.

1711 I.1.2 Test Time Evaluation

1712 Once Alice and Bob have completed training, they will have models f_A^R and f_B^R and model transcripts
 1713 Π_A^R and Π_B^R . However, their final models will be recursively defined in terms of their predictions in
 1714 previous rounds. Thus, in order to evaluate f^R on a new sample (x_A, x_B) , they will have to again
 1715 interact over R rounds, sending each other their predictions \hat{y}^r at each round, which will be computed
 1716 based on their model transcripts Π_A^R and Π_B^R . Note that here, since the prediction is for a single
 1717 datapoint rather than a set of datapoints as it is in the training phase, we revert to the prediction
 1718 notation used in the rest of the paper (\hat{y}^r rather than P^r). This algorithm is formally described in
 1719 Section I.2.2.

³The internal algorithm will run for a variable number of phases across the rounds of the collaborative algorithm between Alice and Bob, but we can assume this variable number of phases is bounded by K . For the sake of notation, we can imagine instantiations with fewer phases to be padded with \perp to make them length K .

1720 I.2 Batch Collaboration Algorithm

1721 Our algorithm will make use of level sets of Alice and Bob’s model’s (discretized) predictions on
1722 their own data as well as the level sets of *each other’s* models.

1723 **Definition I.8** (Level Sets). *Let S_A be Alice’s sample. Let Alice’s predictions at round r for*
1724 *each point in her sample S_A be $P_A^r = \{P_A^{r,1}, \dots, P_A^{r,n}\}$ and Bob’s predictions at round r be*
1725 *$P_B^r = \{P_B^{r,1}, \dots, P_B^{r,n}\}$. Let $v \in \mathbb{R}$. We will say that*

$$\begin{aligned} LS(S_A, P_A^r, v) &= \{x_A^i | P_A^{r,i} = v\}_{i \in [n]} \\ &= \{x_A^i | f_A^r(x_A^i) = v\}_{i \in [n]} \end{aligned}$$

1726 *are Alice’s v th level set on her own predictions. Similarly, we will call Alice’s v th level set on Bob’s*
1727 *predictions*

$$\begin{aligned} LS(S_A, P_B^r, v) &= \{x_A^i | P_B^{r,i} = v\}_{i \in [n]} \\ &= \{x_A^i | f_B^r(x_B^i) = v\}_{i \in [n]}. \end{aligned}$$

1728 **Remark I.9.** *Note that the transcript at round r is directly computable based only on Alice and Bob’s*
1729 *knowledge of their and the other players’ predictions P_A^r and P_B^r —neither player has to recompute*
1730 *f_A^r or f_B^r , nor do they need access to the other players’ features.*

1731 In general, for subroutines we use a subscript \bullet to refer to either A or B , depending on whose inputs
1732 the subroutine was called on, and a subscript \circ to refer to the other player. With this notation in place,
1733 we can proceed to the algorithms.

1734 I.2.1 Training Algorithm

1735 While training, Alice and Bob will run Algorithm I.2.1, COLLABORATE, on their training samples
1736 (S_A, S_B) . This algorithm proceeds in rounds, with Alice and Bob alternating who sends whom their
1737 most current predictions. In each round, the current player will call a subroutine CROSS-BOOST
1738 (Algorithm I.2.1), in which that player boosts their predictions in parallel on each of their datasets’
1739 level sets as defined *by the other players’ predictions*. This “internal” boosting step which is done in
1740 parallel on each of these level sets is itself a boosting algorithm, which we call INTERNAL-BOOST
1741 (Algorithm I.2.1), and is equivalent to the level set boosting algorithm from [Globus-Harris et al.,
1742 2023]: we restate it here as our parametrization is slightly different and to make our notational
1743 choices clear for the sake of our later analysis. At the end of the process, Alice and Bob will have a
1744 collection of individual model transcripts, which they will later use to evaluate the final model on
1745 new datapoints.

1746 [H] 1.15 **Alice’s Input:** $\mathcal{O}_{\mathcal{H}_A}, S_A, m$ **Bob’s Input:** $\mathcal{O}_{\mathcal{H}_B}, S_B, m$ Let $h_B^0 \in \mathcal{O}_{\mathcal{H}_B}(S_B)$ and
1747 $f_B^0 = \text{Round}(h_B^0; m)$. Let $P^{-1} = \perp$ and $P_B^0 = \{f_B^0(x_B)\}_{(x_B, y) \in S_B}$. Bob sends $P^0 = P_B^0$ to Alice.
1748 Let $r = 0$, $\Pi_A^0 = \emptyset$, and $\Pi_B^0 = \{\pi_B^0\} = \{\{f_B^0\}\}$. $P^r \neq P^{r-1}$ r is even Alice plays, boosting her
1749 predictions on Bob’s predictor’s level sets: 0.8

$$f_A^{r+1}, \pi_A^{r+1} = \text{CROSS-BOOST}(S_A, \mathcal{O}_{\mathcal{H}_A}, P_B^r, m)$$

1750 Alice generates her predictions for this round, $P_A^{r+1} = \{f_A^{r+1}(x_A)\}_{(x_A, y) \in S_A}$ Alice sends her
1751 updated predictions $P^{r+1} = P_A^{r+1}$ to Bob. Alice updates her model transcript, setting $\Pi_A^{r+1} =$
1752 $\Pi_A^r \cup \{\pi_A^{r+1}\}$. Bob does nothing, and sets $f_B^{r+1} = f_B^r$ and $\Pi_B^{r+1} = \Pi_B^r$. Bob plays analogously,
1753 boosting his predictions on Alice’s predictor’s level sets: 0.8

$$f_B^{r+1}, \pi_B^{r+1} = \text{CROSS-BOOST}(S_B, \mathcal{O}_{\mathcal{H}_B}, P_A^r, m)$$

1754 Bob generates his predictions for this round, $P_B^{r+1} = \{f_B^{r+1}(x_B)\}_{(x_B, y) \in S_B}$ Bob sends his updated
1755 predictions $P^{r+1} = P_B^{r+1}$ to Alice. Bob updates his model transcript, setting $\Pi_B^{r+1} = \Pi_B^r \cup \{\pi_B^{r+1}\}$.
1756 Alice does nothing, and sets $f_A^{r+1} = f_A^r$. $r = r + 1$. **Alice’s Output:** f_A^r, Π_B^r **Bob’s Output:** f_B^r, Π_B^r
1757

1758 [H] 1.15 **Input :** $S_\bullet, \mathcal{O}_{\mathcal{H}_\bullet}, P_\circ^r, m$ each $v \in [1/m]$

1759 The player generates their v th level set on the other players’ predictions P_\circ^r , 0.7

$$S_\bullet^{r+1, v} = LS(S_\bullet, P_\circ^r, v)$$

1760 Using only their data constrained to this level set, they run the internal boosting algorithm, and
 1761 evaluate their updated model's performance: 0.7

$$\begin{aligned}\tilde{f}_{\bullet}^{r+1,v}, \tilde{\pi}^{r+1,v} &= \text{INTERNAL-BOOST}(S_{\bullet}^{r+1,v}, \mathcal{O}_{\mathcal{H}_{\bullet}}, m) \\ \widetilde{\text{err}}^{r+1,v} &= \mathbb{E}_{(x_{\bullet}, y) \in S_{\bullet}^{r+1,v}} \left[(\tilde{f}_{\bullet}^{r+1,v}(x_{\bullet}) - y)^2 \right]\end{aligned}$$

1762 They then compare their updated model's performance to their counter-party's constant predictor,
 1763 and determine which of the two to use as their final model: Let 0.4

$$\text{err}^v = \mathbb{E}_{(x_{\bullet}, y) \in S_{\bullet}^{r+1,v}} [(v - y)^2]$$

1764 $(\text{err}^v - \widetilde{\text{err}}^{r+1,v}) > 1/m^2$ $f_{\bullet}^{r+1,v}(x_{\bullet}) = \tilde{f}_{\bullet}^{r+1,v}$ $\pi^{r+1,v} = \tilde{\pi}^{r+1,v}$ $f_{\bullet}^{r+1,v}(x_{\bullet}) = v$ $\pi^{r+1,v} = \perp$ The
 1765 player then ensembles their models on each of the level sets of the others' predictions and updates
 1766 their transcript for the round:

$$f_{\bullet}^{r+1}(x_{\bullet}) = \sum_{v \in [1/m]} \mathbb{1}[x \in S_{\bullet}^{r+1,v}] \cdot f_{\bullet}^{r+1,v}(x_{\bullet}),$$

1767 $\pi^{r+1} = \{\pi^{r+1,v}\}_{v \in [1/m]}$ **Output:** $f_{\bullet}^{r+1}, \pi^{r+1}$

1768 [H] 1.2 **Input:** $S_{\bullet}, \mathcal{O}_{\mathcal{H}_{\bullet}}, m$ Let $k = 0$ Let $h_{\bullet}^{r,v,0} = \mathcal{O}_{\mathcal{H}_{\bullet}}(S_{\bullet})$ and $\pi^{r,v,0} = \{h_{\bullet}^{r,v,0}\}$ Let $f_{\bullet}^{r,v,k} =$
 1769 $\text{Round}(h_{\bullet}^{r,v,0}; m^2)$ Let $\text{err}_{-1} = \infty$ and $\text{err}_0 = \mathbb{E}_{(x_{\bullet}, y) \sim S_{\bullet}} [(h_{\bullet}^{r,v,0}(x_{\bullet}) - y)^2]$ $\text{err}_{k-1} - \text{err}_k \geq 1/m^2$
 1770 each $v' \in [1/m^2]$ $S_{\bullet}^{r,v,k+1,v'} = LS(S_{\bullet}, f_{\bullet}^{r,v,k}, v')$ Let $h_{\bullet}^{r,v,k+1,v'} = \mathcal{O}_{\mathcal{H}_{\bullet}}(S_{\bullet}^{r,v,k+1,v'})$. The player
 1771 ensembles their models: 0.8

$$\tilde{f}_{\bullet}^{r,v,k+1}(x_{\bullet}) = \sum_{v' \in [1/m]} \mathbb{1}[f_{\bullet}^{r,v,k}(x_{\bullet}) = v'] \cdot h_{\bullet}^{r,v,k+1,v'}(x_{\bullet})$$

1772

$$f_{\bullet}^{r,v,k+1}(x_{\bullet}) = \text{Round}(\tilde{f}_{\bullet}^{r,v,k+1}; m^2)$$

1773 Let $\text{err}_{k+1} = \mathbb{E}_{(x_{\bullet}, y) \sim S_{\bullet}} [(\tilde{f}_{\bullet}^{r,v,k+1}(x_{\bullet}) - y)^2]$ and $k = k + 1$. Let $\pi^{r,v,k+1} = \{h_{\bullet}^{r,v,k+1,v'}\}_{v' \in [1/m^2]}$.
 1774 Let $k = k + 1$. Let $\pi^{r,v} = (\pi^{r,v,0}, \dots, \pi^{r,v,k-1})$ **Output:** $f_{\bullet}^{r,v,k-1}, \pi^{r,v}$

1775 I.2.2 Test-time Evaluation of Collaborative Model

1776 Upon receiving a fresh datapoint (x_A, x_B) from the distribution, Alice and Bob will use their model
 1777 transcripts from training and a R -round interaction to evaluate f^R on the new datapoint. This is
 1778 described in detail in Algorithm I.2.2, which itself has two subroutines, CROSS-BOOST-EVAL
 1779 (Algorithm I.2.2) and INTERNAL-BOOST-EVAL (Algorithm I.2.2).

1780 [H] 1.2 **Alice's Input:** $x_A, \Pi_A^R = \{\pi_A^1, \dots, \pi_A^R\}, m$. **Bob's Input:** $x_B, \Pi_B^R = \{\pi_B^0, \dots, \pi_B^R\}, m$.
 1781 Bob extracts $f_B^0 = \text{Round}(h_B^0; m)$ from $\pi_B^0 = \{f_B^0\}$. Bob evaluates $\hat{y}^0 = f_B^0(x_B)$, and sends it to
 1782 Alice. Let $r = 0$. $r < R$ r is even Alice updates her prediction and sends it to Bob: She extracts
 1783 π_A^{r+1} from Π_A^R From her transcript from the round and Bob's predictions \hat{y}^r , she reconstructs f_A^{r+1}
 1784 and evaluates it on x_A , generating her prediction \hat{y}^{r+1} for this round: 0.7

$$\hat{y}^{r+1} = \text{CROSS-BOOST-EVAL}(x_A, \hat{y}^r, \pi_A^{r+1}, m)$$

1785 She sends her updated prediction \hat{y}^{r+1} to Bob. Bob does nothing. Bob updates his prediction
 1786 and sends it to Alice: He extracts π_B^{r+1} from Π_B^R He reconstructs f_B^{r+1} and evaluates it on x_B ,
 1787 generating his prediction \hat{y}^{r+1} for this round: 0.7

$$\hat{y}^{r+1} = \text{CROSS-BOOST-EVAL}(x_B, \hat{y}^r, \pi_B^{r+1}, m)$$

1788 He sends his updated prediction \hat{y}^{r+1} to Alice. Alice does nothing. $r = r + 1$ **Alice's Output:**
 1789 \hat{y}^R **Bob's Output:** \hat{y}^R

1790 [H] **Input:** $x_{\bullet}, \hat{y}^{r-1}, \pi^r = \{\pi^{r,v}\}_{v \in [1/m]}, m$. Let $v^* = \hat{y}^{r-1}$ be the value of the other player's
 1791 predictions on x_{\bullet} . The player extracts π^{r,v^*} from π^r . $\pi^{r,v^*} = \perp$ $\hat{y}^r = f_{\bullet}^r(x_{\bullet}) = v^*$. $\hat{y}^r =$
 1792 $\text{INTERNAL-BOOST-EVAL}(x_{\bullet}, \pi^{r,v^*}, m)$ **Output:** \hat{y}^r

1793 **[H] Input:** $x_\bullet, \pi^{r,v} = \{\pi^{r,v,0}, \dots, \pi^{r,v,K}\}, m$. The player extracts $\pi^{r,v,0} = \{h_\bullet^{r,v,0}\}$ from
 1794 $\pi^{r,v}$. Let $v_0^* = f_\bullet^{r,v,0}(x_\bullet) = \text{Round}(h_\bullet^{r,v,0}; m^2)(x_\bullet)$. Let $k = 0$. $k < K$ The player ex-
 1795 tracts $\pi^{r,v,k+1} = \{h_\bullet^{r,v,k+1,v'}\}_{v' \in [1/m]}$ from $\pi^{r,v}$. $\pi^{r,v,k+1} \neq \perp$ Let $v_{k+1}^* = f_\bullet^{r,v,k+1}(x_\bullet) =$
 1796 $\text{Round}(h_\bullet^{r,v,k+1,v_k^*}(x_\bullet); m^2)$. Let $k = k+1$ **Output:** $v_k^* = f_\bullet^{r,v,k}(x_\bullet)$. **Output:** $v_K^* = f_\bullet^{r,v,K}(x_\bullet)$.
 1797

1798 I.3 Algorithm Analysis

1799 We will first show that the COLLABORATE algorithm is guaranteed to converge in a small number
 1800 of rounds. We will then show that if Alice and Bob's model classes satisfy a joint weak learning
 1801 condition with respect to \mathcal{H}_J , then the output of the COLLABORATE algorithm will have low regret
 1802 with respect to \mathcal{H}_J , and finally will demonstrate that it generalizes out of sample.

1803 First, we state our convergence guarantee.

1804 **Theorem I.10.** *In training, the subprocess INTERNAL-BOOST converges after $K = m^2$ (sub)rounds,*
 1805 *and the COLLABORATE Algorithm I.2.1 converges after $R = m^2$ rounds on the training sample S .*

1806 *Proof.* To begin, assume that INTERNAL-BOOST always terminated after at most K rounds. At
 1807 round r , let err^r refer to the empirical squared error of the predictions P^r generated at round r :

$$\text{err}^r = \frac{1}{n} \sum_{i \in [n]} (P^{r,i} - y^i)^2.$$

1808 Consider what happens at round r of Algorithm I.2.1 when CROSS-BOOST is called. The CROSS-
 1809 BOOST algorithm has two kinds of updates that can occur on the level sets of the other players'
 1810 predictions: either the current player can choose to update their predictor to the output of INTERNAL-
 1811 BOOST or they can set their predictions on that level set to be equivalent to Bob's. Note that if, at any
 1812 round, they choose on all their level sets to use Bob's predictions, Algorithm I.2.1 will halt, because
 1813 $P^{r+1} = P^r$.

1814 Say that instead they choose to use the output of INTERNAL-BOOST on at least one level set v^* .
 1815 Then, on this level set their predictions will be equal to $\tilde{f}^{r+1,v^*}(x_\bullet)$. Note that the player's level sets
 1816 on the other players' predictions are disjoint, and that squared error is always non-negative. So,

$$\begin{aligned} \text{err}^{r+1} - \text{err}^r &= \text{err}^{r+1} - \sum_{v \in [1/m]} |S_\bullet^{r+1,v}| \cdot \text{err}^v \\ &= \frac{1}{n} \sum_{v \in [1/m]} |S_\bullet^{r+1,v}| \left(\sum_{x_\bullet^i \in S_\bullet^{r+1,v}} (P^{r+1,i} - y^i)^2 \right) - \sum_{v \in [1/m]} |S_\bullet^{r+1,v}| \cdot \text{err}^v \\ &= \frac{1}{n} \sum_{v \in [1/m]} |S_\bullet^{r+1,v}| \left(\sum_{x_\bullet^i \in S_\bullet^{r+1,v}} (P^{r+1,i} - y^i)^2 - \text{err}^v \right) \\ &\geq \frac{|S^{v^*}|}{n} \left(\sum_{x_\bullet^i \in S_\bullet^{r+1,v^*}} (P^{r+1,i} - y^i)^2 - \text{err}^{v^*} \right) \\ &= \frac{|S^{v^*}|}{n} \left(\sum_{x_\bullet^i \in S_\bullet^{r+1,v^*}} (\tilde{f}^{r+1,v^*}(x_\bullet^i) - y^i)^2 - \text{err}^{v^*} \right) \\ &= \widetilde{\text{err}}^{r+1,v^*} - \text{err}^{v^*} \\ &\geq 1/m^2 \end{aligned}$$

1817 Thus, at every round r in which they do not halt, they must improve the squared error of their
 1818 predictions by at least α . In the worst case, $\text{err}^0 = 1$, i.e. Bob's initial predictions are maximally

1819 incorrect. Squared error can never decrease below zero, so they must halt after at most $R = m^2$
 1820 rounds.

1821 It remains to show that INTERNAL-BOOST also terminates. This follows a similar potential
 1822 argument on the squared error as above. Modulo notational changes in our halting condition, the
 1823 complete proof is equivalent to that of the halting condition proved as part of Theorem 4.3 in
 1824 Globus-Harris et al. [2023]. \square

1825 We now prove an in-sample accuracy theorem for COLLABORATE. The proof of this statement
 1826 follows from our Boosting Lemma B.4 and a series of Lemmas.

- 1827 • Any time that INTERNAL-BOOST is invoked, the resulting model will have small swap
 1828 regret with respect to the players' own hypothesis class on the subset of data it was called
 1829 on. (Lemma I.12)
- 1830 • For any invocation of the CROSS-BOOST algorithm, either a model from INTERNAL-
 1831 BOOST will be used or a constant predictor from the other player will be. If a model from
 1832 INTERNAL-BOOST was used, it will have small swap regret on that subsample. And if
 1833 not, the regret of the constant predictor which is used instead cannot be too much bigger,
 1834 because the player only decided to use this constant predictor because the improvement from
 1835 using INTERNAL-BOOST instead was small. Summing over the players' level sets gives a
 1836 swap-regret guarantee on the entire model generated by CROSS-BOOST with respect to the
 1837 players' own hypothesis class and their sample. (Lemma I.13)
- 1838 • Because the final predictions by Alice and Bob generated by COLLABORATE always
 1839 agree, the final predictions have low swap regret on $\mathcal{H}_A \cup \mathcal{H}_B$. (Corollary I.14)
- 1840 • Hence, if \mathcal{H}_A and \mathcal{H}_B satisfy the weak learning condition with respect to \mathcal{H}_J , we can
 1841 directly apply the boosting result from Lemma B.4.

1842 **Theorem I.11.** *Let \mathcal{H}_J be a hypothesis class over the joint feature space \mathcal{X} , and let $\mathcal{H}_A = \{h_A : \mathcal{X}_A \rightarrow \mathcal{Y}\}$ and $\mathcal{H}_B = \{h_B : \mathcal{X}_B \rightarrow \mathcal{Y}\}$ be hypothesis classes over \mathcal{X}_A and \mathcal{X}_B respectively. Let f^R be the final model output by COLLABORATE. Then, if \mathcal{H}_A and \mathcal{H}_B jointly satisfy the $w(\cdot)$ -weak learning condition with respect to \mathcal{H}_J ,*

$$\mathbb{E}_S[(f^R(x) - y)^2] - \min_{h_J \in \mathcal{H}_J} \mathbb{E}_S[(h_J(x) - y)^2] \leq 2w^{-1} \left(\frac{3}{m} \right),$$

1846 *whenever the inverse of w exists.*

1847 We begin by proving a series of swap regret guarantees, first with respect to the individual runs of
 1848 INTERNAL-BOOST, then with respect to runs of CROSS-BOOST, and finally with respect to the
 1849 final model output by COLLABORATE.

1850 **Lemma I.12.** *Let $f^{r,v,K}$ be the model output by a run of INTERNAL-BOOST on a player's sample*
 1851 *S^\bullet , and let \mathcal{H}_\bullet be that player's own hypothesis class. Then, every time INTERNAL-BOOST is run by*
 1852 *a player, the final model has $(2/m^2, \mathcal{H}_\bullet)$ -swap regret on the sample S^\bullet it was run on:*

$$2/m^2 \geq \mathbb{E}_{(x_\bullet, y) \sim S_\bullet} \left[(f_\bullet^{r,v,K}(x_\bullet) - y)^2 \right] - \min_{h \in \mathcal{H}_\bullet} \mathbb{E}_{(x_\bullet, y) \sim S_\bullet} [\mathbb{1}[f_\bullet^{r,v,K}(x_\bullet) = v](h(x) - y)^2].$$

1853 *Proof.* Say that INTERNAL-BOOST is run on a sample S_\bullet and outputs the model from round K .
 1854 Recall that in INTERNAL-BOOST if the output model is the model from round K , then in fact the
 1855 algorithm ran for $K + 1$ rounds, and the stopping condition at the final round $K + 1$ is in terms of
 1856 the error of the *unrounded* predictors $\tilde{f}^{r,v,K}$ and $\tilde{f}^{r,v,K+1}$ which were generated at that round and

1857 the previous one. So, since the algorithm halted,

$$\begin{aligned}
1/m^2 &> \text{err}_K - \text{err}_{K+1} \\
&= \mathbb{E}_{(x_\bullet, y) \sim S_\bullet} \left[\left(\tilde{f}_\bullet^{r,v,K}(x_\bullet) - y \right)^2 \right] - \mathbb{E}_{(x_\bullet, y) \sim S_\bullet} \left[\left(\tilde{f}_\bullet^{r,v,K+1}(x_\bullet) - y \right)^2 \right] \\
&= \mathbb{E}_{(x_\bullet, y) \sim S_\bullet} \left[\left(\tilde{f}_\bullet^{r,v,K}(x_\bullet) - y \right)^2 \right] - \mathbb{E}_{(x_\bullet, y) \sim S_\bullet} \left[\left(\sum_{v' \in [1/m]} \mathbb{1}[f^{r,v,K}(x) = v'] \cdot h^{r,v,K+1,v'}(x) - y \right)^2 \right] \\
&\geq \mathbb{E}_{(x_\bullet, y) \sim S_\bullet} \left[\left(\tilde{f}_\bullet^{r,v,K}(x_\bullet) - y \right)^2 \right] - \sum_{v' \in [1/m]} \mathbb{E}_{(x_\bullet, y) \sim S_\bullet} \left[\mathbb{1}[f^{r,v,K}(x) = v'] \left(h^{r,v,K+1,v'}(x) - y \right)^2 \right] \\
&\quad \text{(by Cauchy-Schwartz)} \\
&= \mathbb{E}_{(x_\bullet, y) \sim S_\bullet} \left[\left(\tilde{f}_\bullet^{r,v,K}(x_\bullet) - y \right)^2 \right] - \min_{h \in \mathcal{H}_\bullet} \left(\mathbb{E}_{(x_\bullet, y) \sim S_\bullet} [\mathbb{1}[f^{r,v,K}(x_\bullet) = v](h(x) - y)^2] \right). \quad (\text{Eq. 1}) \\
&\quad \text{(by the definition of } h^{r,v,K+1,v'} \in \mathcal{O}_{\mathcal{H}_\bullet})
\end{aligned}$$

1858 This expression is nearly the swap regret statement we want, except we need to bound the swap regret
1859 of our *rounded* predictor $f^{r,v,K}$, rather than our unrounded $\tilde{f}^{r,v,K}$. However, note that pointwise,
1860 from the definition of Round, $|f^{r,v,K}(x) - \tilde{f}^{r,v,K}(x)| \leq 1/(2m^2)$. Hence,

$$\begin{aligned}
\mathbb{E}_{(x_\bullet, y) \sim S_\bullet} \left[\left(f_\bullet^{r,v,K}(x_\bullet) - y \right)^2 \right] &= \mathbb{E}_{(x_\bullet, y) \sim S_\bullet} \left[\left(\text{Round} \left(\tilde{f}_\bullet^{r,v,K}(x_\bullet); m^2 \right) - y \right)^2 \right] \\
&= \mathbb{E}_{(x_\bullet, y) \sim S_\bullet} \left[\left(\text{Round} \left(\tilde{f}_\bullet^{r,v,K}(x_\bullet); m^2 \right) \right)^2 \right] \\
&\quad - 2\mathbb{E}_{(x_\bullet, y) \sim S_\bullet} \left[\text{Round} \left(\tilde{f}_\bullet^{r,v,K}(x_\bullet); m^2 \right) \cdot y \right] + \mathbb{E}_{(x_\bullet, y) \sim S_\bullet} [y^2] \\
&\leq \mathbb{E}_{(x_\bullet, y) \sim S_\bullet} \left[\left(\tilde{f}_\bullet^{r,v,K}(x_\bullet) + \frac{1}{2m^2} \right)^2 \right] - 2 \left(\tilde{f}_\bullet^{r,v,K}(x_\bullet) - \frac{1}{2m^2} \right) y + y^2 \\
&\leq \mathbb{E}_{(x_\bullet, y) \sim S_\bullet} \left[\left(\tilde{f}_\bullet^{r,v,K}(x_\bullet) - y \right)^2 \right] + \frac{3}{4m^2}
\end{aligned}$$

1861 Combining this with the bound in Equation 1 gives us that

$$\begin{aligned}
1/m^2 &> \mathbb{E}_{(x_\bullet, y) \sim S_\bullet} \left[\left(\tilde{f}_\bullet^{r,v,K}(x_\bullet) - y \right)^2 \right] - \min_{h \in \mathcal{H}_\bullet} \left(\mathbb{E}_{(x_\bullet, y) \sim S_\bullet} [\mathbb{1}[f^{r,v,K}(x_\bullet) = v](h(x) - y)^2] \right) \\
&\geq \mathbb{E}_{(x_\bullet, y) \sim S_\bullet} \left[\left(f_\bullet^{r,v,K}(x_\bullet) - y \right)^2 \right] - \min_{h \in \mathcal{H}_\bullet} \left(\mathbb{E}_{(x_\bullet, y) \sim S_\bullet} [\mathbb{1}[f^{r,v,K}(x_\bullet) = v](h(x) - y)^2] \right) - \frac{3}{4m^2}
\end{aligned}$$

1862 And hence $f_\bullet^{r,v,K}(x_\bullet)$ has at most $2/m^2 > 1/m^2 + 3/4m^2$ swap-regret on S_\bullet with respect to
1863 \mathcal{H}_\bullet . \square

1864 **Lemma I.13.** Let f_\bullet^r be the model generated by CROSS-BOOST at round r on the player's sample
1865 S_\bullet . Then f_\bullet^r will have $(3/m, \mathcal{H}_\bullet)$ -swap regret on S_\bullet .

1866 *Proof.* Recall that in CROSS-BOOST, the player will bucket their sample into level sets based on the
1867 other players' predictions, which we call $S_\bullet^{r,v}$. Their final model f_\bullet^r will be an ensemble of models
1868 $f_\bullet^{r,v}$ generated on each level set v . On some of these level sets, their model will equal to a model
1869 $\tilde{f}_\bullet^{r,v}$ which is output by INTERNAL-BOOST. On these level sets, we can directly invoke the swap
1870 regret guarantee from Lemma I.12. However, if the INTERNAL-BOOST process did not sufficiently
1871 improve their squared error on $S_\bullet^{r,v}$, they will instead set $f_\bullet^{r,v}$ to always predict the other players'

1872 constant prediction v . We will first show that for any level set v where this happens, there is low swap
 1873 regret with respect to the sample $S_{\bullet}^{r,v}$.

1874 As in the statement of Algorithm I.2.1, let

$$\begin{aligned}\text{err}^v &= \mathbb{E}_{(x_{\bullet}, y) \sim S_{\bullet}^{r,v}} [(v - y)^2], \text{ and} \\ \widetilde{\text{err}}^{r,v} &= \mathbb{E}_{(x_{\bullet}, y) \sim S_{\bullet}^{r,v}} [(\tilde{f}_{\bullet}^{r,v}(x_{\bullet}) - y)^2].\end{aligned}$$

1875 Since the player chose to use Bob's predictor v , we know that $\text{err}^v - \widetilde{\text{err}}^{r,v} \leq 1/m^2$. But this means

$$\begin{aligned}1/m^2 &> \mathbb{E}_{(x_{\bullet}, y) \sim S_{\bullet}^{r,v}} [(v - y)^2] - \mathbb{E}_{(x_{\bullet}, y) \sim S_{\bullet}^{r,v}} [(\tilde{f}_{\bullet}^{r,v}(x_{\bullet}) - y)^2] \\ &= \left(\mathbb{E}_{(x_{\bullet}, y) \sim S_{\bullet}^{r,v}} [(v - y)^2] - \sum_{v' \in [1/m]} \min_{h \in \mathcal{H}_{\bullet}} \left(\mathbb{E}_{(x_{\bullet}, y) \sim S_{\bullet}^{r,v}} [\mathbb{1}[\tilde{f}_{\bullet}^{r,v}(x_{\bullet}) = v'](h(x) - y)^2] \right) \right) \\ &\quad - \left(\mathbb{E}_{(x_{\bullet}, y) \sim S_{\bullet}^{r,v}} [(\tilde{f}_{\bullet}^{r,v}(x_{\bullet}) - y)^2] - \sum_{v' \in [1/m]} \min_{h \in \mathcal{H}_{\bullet}} \left(\mathbb{E}_{(x_{\bullet}, y) \sim S_{\bullet}^{r,v}} [\mathbb{1}[\tilde{f}_{\bullet}^{r,v}(x_{\bullet}) = v'](h(x) - y)^2] \right) \right) \\ &> \mathbb{E}_{(x_{\bullet}, y) \sim S_{\bullet}^{r,v}} [(v - y)^2] - \sum_{v' \in [1/m]} \min_{h \in \mathcal{H}_{\bullet}} \left(\mathbb{E}_{(x_{\bullet}, y) \sim S_{\bullet}^{r,v}} [\mathbb{1}[\tilde{f}_{\bullet}^{r,v}(x_{\bullet}) = v'](h(x) - y)^2] \right) - 2/m^2\end{aligned}$$

(By Lemma I.12)

1876 Recall that low swap regret always implies low external regret. And for constant predictors, swap
 1877 regret and external regret are equivalent statements. So this inequality in turn implies that on the
 1878 subsample $S_{\bullet}^{r,v}$ where they used Bob's constant prediction v instead of $\tilde{f}_{\bullet}^{r,v}$,

$$\begin{aligned}3/m^2 &> \mathbb{E}_{(x_{\bullet}, y) \sim S_{\bullet}^{r,v}} [(v - y)^2] - \min_{h \in \mathcal{H}_{\bullet}} \left(\mathbb{E}_{(x_{\bullet}, y) \sim S_{\bullet}^{r,v}} [(h(x) - y)^2] \right), \\ &= \mathbb{E}_{(x_{\bullet}, y) \sim S_{\bullet}^{r,v}} [(\tilde{f}_{\bullet}^{r,v}(x_{\bullet}) - y)^2] - \sum_{v' \in [1/m]} \min_{h \in \mathcal{H}_{\bullet}} \left(\mathbb{E}_{(x_{\bullet}, y) \sim S_{\bullet}^{r,v}} [\mathbb{1}[\tilde{f}_{\bullet}^{r,v}(x_{\bullet}) = v'](h(x) - y)^2] \right).\end{aligned}$$

1879 In other words, the player will have at most $(3/m^2, \mathcal{H}_{\bullet})$ -swap regret with respect to the subsample
 1880 $S_{\bullet}^{r,v}$ on any subsample where they chose to follow the other players' prediction, which will be a
 1881 constant predictor on this subsample. We will now combine these marginal guarantees which are
 1882 with respect to the subsamples $S_{\bullet}^{r,v}$ into a swap regret guarantee on the entire sample S_{\bullet} .

1883 On any level set $S_{\bullet}^{r,v}$ where f_{\bullet}^r evaluates to $\tilde{f}_{\bullet}^{r,v}$, they will have $(2/m^2, \mathcal{H}_{\bullet})$ -swap regret with respect
 1884 to $S_{\bullet}^{r,v}$, and on any level set $S_{\bullet}^{r,v}$ where $f_{\bullet}^r = v$, they will have at most $(3/m^2, \mathcal{H}_{\bullet})$ -swap regret with
 1885 respect to $S_{\bullet}^{r,v}$. So in the worst case they will have swapped out to the other players' predictions on
 1886 each level set, and

$$\begin{aligned}\mathbb{E}_{(x_{\bullet}, y) \sim S_{\bullet}} [(f_{\bullet}^r(x_{\bullet}) - y)^2] &- \sum_{v \in [1/m]} \min_{h \in \mathcal{H}} \left(\mathbb{E}_{(x_{\bullet}, y) \sim S_{\bullet}} [\mathbb{1}[f_{\bullet}^r = v](h(x) - y)^2] \right) \\ &\leq \mathbb{E}_{(x_{\bullet}, y) \sim S_{\bullet}} \left[\sum_{v \in [1/m]} \mathbb{P}(x_{\bullet} \in S_{\bullet}^{r,v}) \left(\mathbb{E}_{(x_{\bullet}, y) \sim S_{\bullet}^{r,v}} [(f_{\bullet}^r(x_{\bullet}) - y)^2] \right. \right. \\ &\quad \left. \left. - \min_{h \in \mathcal{H}} \left(\mathbb{E}_{(x_{\bullet}, y) \sim S_{\bullet}^{r,v}} [\mathbb{1}[f_{\bullet}^r(x_{\bullet}) = v](h(x) - y)^2] \right) \right) \right], \\ &\leq m \left(\frac{3}{m^2} \right), \\ &= 3/m.\end{aligned}$$

1887

□

1888 **Corollary I.14.** Let f_A^R and f_B^R be the models output by Alice and Bob after running Algorithm I.2.1,
 1889 which halted after r rounds. Then the models will have $(3/m, \mathcal{H}_A \cup \mathcal{H}_B)$ -swap regret with respect to
 1890 the shared sample S .

1891 *Proof.* Note that at the final round, Alice and Bob's predictions will agree, because otherwise
 1892 Algorithm I.2.1 will not have terminated. We know from Lemma I.13 that their models on their
 1893 respective samples S_A and S_B will have $(3/m, \mathcal{H}_A)$ and $(3/m, \mathcal{H}_B)$ -swap regret respectively. So,
 1894 since they also agree at this round, it must be the case that they have swap regret bounded by $3/m$
 1895 with respect to $\mathcal{H}_A \cup \mathcal{H}_B$. \square

1896 This gives us all of the technical machinery needed for the proof of Theorem I.11, as stated in Section
 1897 I.3.

1898 *Proof of Theorem I.11.* We know from Corollary I.14 that the final models f_A^R and f_B^R output by
 1899 Algorithm I.2.1 have $(3/m, \mathcal{H}_A \cup \mathcal{H}_B)$ -swap regret on the sample S . By assumption, \mathcal{H}_A and
 1900 \mathcal{H}_B jointly satisfy the $w(\cdot)$ -weak learning condition with respect to \mathcal{H}_J . So, we can directly apply
 1901 boosting Lemma B.4, which will guarantee that

$$\mathbb{E}_S[(f^R(x) - y)^2] - \min_{h_J \in \mathcal{H}_J} \mathbb{E}_S[(h_J(x) - y)^2] \leq 2w^{-1} \left(\frac{3}{m} \right).$$

1902 \square

1903 This gives us in-sample accuracy guarantees.

1904 We will now state our generalization guarantee. As the models generated by the COLLABORATE
 1905 algorithm only include m possible values for the final predictor, we will leverage a multiclass uniform
 1906 convergence theorem which relies on the pseudodimension of \mathcal{H}_J . We will then in turn use a bound
 1907 on the Natarajan dimension of \mathcal{H}_J to bound its pseudodimension, applying a lemma that states that
 1908 if a model may be written as a decision rule over binary classifiers, then its Natarajan dimension is
 1909 bounded above by its pseudodimension. Writing our models as such decision rules will require a
 1910 small technical assumption that \mathcal{H}_J is "closed" with respect to \mathcal{H}_A and \mathcal{H}_B , i.e. that \mathcal{H}_J contains a
 1911 function equivalent to any function in \mathcal{H}_A or \mathcal{H}_B but defined over its input space $\mathcal{X}_A \times \mathcal{X}_B$.

1912 **Definition I.15** (Closure of \mathcal{H}_J with respect to \mathcal{H}_A and \mathcal{H}_B). We will say that \mathcal{H}_J is closed with re-
 1913 spect to \mathcal{H}_A and \mathcal{H}_B if for any $h_A \in \mathcal{H}_A$ there exists some $h \in \mathcal{H}_J$ such that $h(x) = h((x_A, x_B)) =$
 1914 $h_A(x_A)$ and for any $h_B \in \mathcal{H}_B$ there exists some $h \in \mathcal{H}_J$ such that $h(x) = h((x_A, x_B)) = h_B(x_B)$.

1915 We now state the generalization theorem.

1916 **Theorem I.16.** Let $\varepsilon, \delta > 0$ and let \mathcal{F} be the class of models output from Algorithm I.2.1 for any
 1917 joint input distribution \mathcal{D} . Let d be the pseudodimension of Alice and Bob's joint hypothesis class
 1918 \mathcal{H}_J , and assume that \mathcal{H}_J is closed with respect to \mathcal{H}_A and \mathcal{H}_B . Let $S = \{(x_A, x_B, y_i)\}_{i \in [n]} \sim \mathcal{D}^n$
 1919 be a sample of n iid points drawn from \mathcal{D} . Then, if

$$n \geq O \left(\frac{m^7 d \log(md) + \log(1/\delta)}{\varepsilon^2} \right),$$

1920

$$\mathbb{P} \left(\min_{f \in \mathcal{F}} |\mathbb{E}_{(x_A, x_B, y) \sim \mathcal{D}} [(y - f(x))^2] - \mathbb{E}_{(x_A, x_B, y) \sim S} [(y - f(x))^2]| \geq \varepsilon \right) \leq \delta.$$

1921 I.3.1 Definitions and Referenced Theorem Statements for Proof of Generalization

1922 In order to prove this statement, we will need to rely on a variety of different definitions and standard
 1923 results from the machine learning theory literature.

1924 **Definition I.17** (VC-dimension). Vapnik and Chervonenkis [1971] Let \mathcal{H} be a class of binary
 1925 classifiers $h : \mathcal{X} \rightarrow \{0, 1\}$. Let $S = \{x_1, \dots, x_n\}$ and let $\Pi_{\mathcal{H}}(S) = \{(h(x_1), \dots, h(x_n)) : h \in$
 1926 $\mathcal{H}\} \subseteq \{0, 1\}^n$. We say that S is shattered by \mathcal{H} if $\Pi_{\mathcal{H}}(S) = \{0, 1\}^n$. The Vapnik-Chervonenkis (VC)
 1927 dimension of \mathcal{H} , denoted $\text{VCdim}(\mathcal{H})$, is the cardinality of the largest set S shattered by \mathcal{H} .

Definition I.18. *Pseudodimension*[Pollard [2012]] Let \mathcal{H} be a class of functions from \mathcal{X} to \mathbb{R} . We say that a set $S = (x_1, \dots, x_m, y_1, \dots, y_m) \in \mathcal{X}^m \times \mathbb{R}^m$ is pseudo-shattered by \mathcal{H} if for any $(b_1, \dots, b_m) \in \{0, 1\}^m$ there exists $h \in \mathcal{H}$ such that $\forall i, h(x_i) > y_i \iff b_i = 1$. The pseudodimension of \mathcal{H} , denoted $\text{Pdim}(\mathcal{H})$ is the largest integer m for which \mathcal{H} pseudo-shatters some set S of cardinality m .

Definition I.19 (Shattering for multiclass functions). *Natarajan [1989], Shalev-Shwartz and Ben-David [2014]* A set $C \subseteq \mathcal{X}$ is shattered by \mathcal{H} if there exists two functions $f_0, f_1 : C \rightarrow [k]$ such that

1. For every $x \in C$, $f_0(x) \neq f_1(x)$.

2. For every $B \subseteq C$ there exists a function $h \in \mathcal{H}$ such that

$$\forall x \in B, h(x) = f_0(x) \text{ and } \forall x \in C \setminus B, h(x) = f_1(x).$$

Definition I.20 (Natarajan dimension). *Natarajan [1989], Shalev-Shwartz and Ben-David [2014]* The Natarajan dimension of \mathcal{H} , denoted $\text{Ndim}(\mathcal{H})$, is the maximal size of a shattered set $C \subseteq \mathcal{X}$.

Theorem I.21 (Multiclass uniform convergence). *Shalev-Shwartz and Ben-David [2014]* Let $\epsilon, \delta > 0$ and let \mathcal{H} be a class of functions $h : \mathcal{X} \rightarrow [1/k]$ such that the Natarajan dimension of \mathcal{H} is d . Let $\mathcal{D} \in \Delta(\mathcal{X} \times [0, 1])$ be an arbitrary distribution and let $D = \{(x_1, y_1), \dots, (x_n, y_n)\}_{(x_i, y_i) \sim \mathcal{D}}$ be a sample of n points from \mathcal{D} . Then for

$$n \geq O\left(\frac{d \log(k) + \log(1/\delta)}{\epsilon^2}\right),$$

$$\mathbb{P}\left[\max_{h \in \mathcal{H}} |\mathbb{E}_{(x,y) \sim \mathcal{D}}[(y - h(x))^2] - \mathbb{E}_{(x,y) \sim D}[(y - h(x))^2]| \geq \epsilon\right] \leq \delta.$$

Lemma I.22. *Shalev-Shwartz and Ben-David [2014]* Suppose we have ℓ binary classifiers from binary class \mathcal{H}_{bin} and a rule $r : \{0, 1\}^\ell \rightarrow [k]$ that determines a multiclass label according to the predictions of the ℓ binary classifiers. Define the hypothesis class corresponding to this rule as

$$\mathcal{H} = \{r(h_1(\cdot), \dots, h_\ell(\cdot)) : (h_1, \dots, h_\ell) \in (\mathcal{H}_{\text{bin}})^\ell\}.$$

Then, if $d = \text{VCdim}(\mathcal{H}_{\text{bin}})$,

$$\text{Ndim}(\mathcal{H}) \leq 3\ell d \log(\ell d).$$

I.3.2 Generalization Proof

First, we show that the models generated by the COLLABORATE algorithm may be written as decision rules over a polynomial number of binary predictors.

Lemma I.23. *Let K be an upper bound on the number of rounds that INTERNAL-BOOST ever runs for, and let r be a round of COLLABORATE. Then we can write the player's model f_\bullet^r at round r as a decision rule $\rho_\bullet^r : \{0, 1\}^\ell \rightarrow [1/m]$ over $\ell \leq m + rKm^3$ binary predictors. Assuming that \mathcal{H}_J is closed with respect to \mathcal{H}_A and \mathcal{H}_B , each of these binary predictor $g : \mathcal{X} \rightarrow \{0, 1\}$ will be a mapping from the full feature space $\mathcal{X} = (\mathcal{X}_A, \mathcal{X}_B)$ induced by a function $h \in \mathcal{H}_J$.*

Proof. We proceed by induction, first showing that f^r may be written as a decision rule over classifiers and then arguing that the number of total classifiers is bounded by $m + rKm^3$.

Base Case Consider the following m binary classifiers, $g^{0,v} : \mathcal{X} \rightarrow \{0, 1\}$ defined for each $v \in [1/m]$ and $x = (x_A, x_B) \in \mathcal{X}$:

$$\begin{aligned} g^{0,v}(x) &= \begin{cases} 1 & \text{if } f_B^0(x_B) = v, \\ 0 & \text{else} \end{cases} \\ &= \begin{cases} 1 & \text{if } \text{Round}(h_B^0; m)(x_B) = v, \\ 0 & \text{else} \end{cases} \end{aligned}$$

We can then write the following decision rule

$$\rho^0(\{g^{0,v}\}_{v \in [1/m]})(x) = \arg \max_{v \in [1/m]} v \cdot \mathbb{1}[g^{0,v}(x) = 1] = f_B^0(x_B),$$

which exactly reconstructs the starting model.

1963 **Induction step** Say that at round r Bob has played, and his model f_B^r may be written as a
 1964 decision rule ρ^r . We will now show that Alice's model f_A^{r+1} may be written as a decision rule
 1965 recursively defined in terms of ρ^r . First, we will fix v , and consider what happens internally to
 1966 INTERNAL-BOOST:

1967 **Base case** Consider the initial round of INTERNAL-BOOST, when $k = 0$. For
 1968 each $v' \in [1/m]$, let $g^{r+1,v,0,v'} : \mathcal{X} \rightarrow \{0, 1\}$ be a classifier

$$\begin{aligned} g^{r+1,v,0,v'}(x) &= \begin{cases} 1 & \text{if } f_A^{r+1,v,0}(x_A) = v', \\ 0 & \text{else,} \end{cases} \\ &= \begin{cases} 1 & \text{if Round}(h_A^{r+1,v,0}; m)(x_A) = v', \\ 0 & \text{else.} \end{cases} \end{aligned}$$

1969 As in our base case for the analysis for CROSS-BOOST, we can rewrite $f_A^{r+1,v,0}$
 1970 as a decision rule $\rho^{r,v,0}$ in terms of $g^{r,v,0,v'}$:

$$\rho^{r,v,0} \left(\{g^{r+1,v,0,v'}\}_{v' \in [1/m]} \right) (x) = \arg \max_{v' \in [1/m]} v' \cdot \mathbb{1}[g^{r+1,v,0,v'}(x) = 1] = f_A^{r+1,v,0}(x_A).$$

1971 **Induction step for INTERNAL-BOOST** Say that the claim holds at round k of
 1972 INTERNAL-BOOST, i.e. that there is a decision rule $\rho^{r+1,v,k}$ such that $\rho^{r+1,v,k} =$
 1973 $f_A^{r+1,v,k}$. Let $v', i \in [1/m]$ and define the following m^2 binary classifiers

$$g_{r+1,v,k+1}^{v',i}(x) = \begin{cases} 1 & \text{if Round}(h_A^{r+1,v,k+1,v'}(x_A); m) = i, \\ 0 & \text{else.} \end{cases}$$

1974 Then, we can write

$$\begin{aligned} \rho^{r+1,v,k+1}(\rho^{r+1,v,k}, \{g_{r+1,v,k+1}^{v',i}\}_{(v',i) \in [1/m]})(x) &= \sum_{(v',i) \in [1/m]} i \cdot \mathbb{1}[\rho^{r+1,v,k}(x) = v'] \mathbb{1}[g_{r+1,v,k+1}^{v',i}(x) = 1], \\ &= \sum_{v' \in [1/m]} \mathbb{1}[f_A^{r+1,v,k}(x_A) = v'] \cdot \sum_{i \in [1/m]} i \cdot \mathbb{1}[\text{Round}(h_A^{r+1,v,k+1,v'}(x_A); m) = i] \\ &= \sum_{v' \in [1/m]} \mathbb{1}[f_A^{r+1,v,k}(x_A) = v'] \cdot \text{Round}(h_A^{r+1,v,k+1,v'}(x_A); m), \\ &= f_A^{r+1,v,k+1}(x_A), \end{aligned}$$

1975 which concludes the induction internal to INTERNAL-BOOST.

1976 Following the induction argument in Globus-Harris et al. [2023], $\rho^{r+1,v,k+1}$ is a
 1977 decision rule over a total of $m + (k + 1)m^2$ classifiers.

1978 Now, we wish to show that f_A^{r+1} may be written as a decision rule ρ^{r+1} . Recall that in
 1979 CROSS-BOOST, on each level set of Bob's prediction, the updated model $f^{r,v}$ will *either* be
 1980 equivalent to Bob's predictions or a model output by INTERNAL-BOOST will be evaluated on the
 1981 point. Let $V_1 \subseteq [1/m]$ be the collection of level sets at round r where Alice's updated model was
 1982 equivalent to $\tilde{f}^{r+1,v}$ and let V_2 be the collection of level sets where her model used Bob's predictions.
 1983 I.e.,

$$\begin{aligned} V_1 &= \{v \in [1/m] : f_A^{r+1,v}(x_A) = \tilde{f}_A^{r+1,v}(x_A)\} \\ V_2 &= \{v \in [1/m] : f_A^{r+1,v}(x_A) = v\}. \end{aligned}$$

1984 Note that $[1/m] = V_1 \cup V_2$ and the two sets are disjoint. For $v \in V_1$, let K_v be the total number of
 1985 rounds that INTERNAL-BOOST ran for, and define

$$\rho^{r,v} = \begin{cases} \rho^{r,v,K_v} & \text{if } v \in V_1 \\ \rho^r & \text{if } v \in V_2. \end{cases}$$

1986 Then we can write

$$\begin{aligned} \rho^{r+1} \left(\rho^r, \{\rho^{r,v}\}_{v \in [1/m]} \right) (x) &= \sum_{v \in [1/m]} \mathbb{1}[\rho^r(x) = v] \cdot \rho^{r,v}(x) \\ &= \sum_{v \in V_1} \mathbb{1}[\rho^r(x) = v] \rho^{r,v,K_v}(x) + \sum_{v \in V_2} \mathbb{1}[\rho^r = v] \rho^r \\ &= \sum_{v \in V_1} \mathbb{1}[x \in S^{r+1,v}] \cdot \tilde{f}(x_A) + \sum_{v \in V_2} \mathbb{1}[x \in S^{r+1,v}] \cdot v \\ &= f_A^{r+1}. \end{aligned}$$

1987 In other words, Alice's model at round $r + 1$ may be written as a decision rule recursively defined in
 1988 terms of her decision rules from INTERNAL-BOOST on the level sets where these models are used
 1989 and on Bob's decision rule ρ^r .

1990 We now need to give an upper bound for the number of binary predictors which ρ^r is comprised of.
 1991 Let K be the maximum number of rounds that INTERNAL-BOOST ever runs for. Note ρ^0 is made
 1992 up of m classifiers, and say that ρ^r is made up of at most $m + r(m + Km^2)m$ classifiers. Note that
 1993 for any $v \in V_2$ no new classifiers will be invoked. So in the worst case, $V_1 = [1/m]$, i.e. for each of
 1994 Alice's level sets on Bob's predictor, INTERNAL-BOOST is invoked. Each of these runs will add at
 1995 most $m + Km^2$ classifiers to the decision rule, so in total there will be at most $m(m + Km^2)$ new
 1996 classifiers added to the decision rule. Hence, ρ^{r+1} will be comprised of at most

$$\begin{aligned} \ell &= m + r(m + Km^2) + (m + Km^2)m \\ &= m + (r + 1)m(m + Km^2) \\ &\leq m + (r + 1)(K + 1)m^3 \end{aligned}$$

1997 classifiers. □

1998 **Lemma I.24** (The VC dimension of $\mathcal{G}_{\mathcal{H}_J, \eta}$ is bounded by the pseudodimension of \mathcal{H}_J). *Let $\mathcal{G}_{\mathcal{H}_J, \eta}$*
 1999 *be the class of Boolean classifiers induced by $\text{Round}(h(x); m)$ for $h \in \mathcal{H}_J$. I.e., for any $g \in \mathcal{G}_{\mathcal{H}_J, \eta}$*
 2000 *there must be some $v \in [1/m]$ such that*

$$g(x) = \begin{cases} 1 & \text{if } \text{Round}(h(x); m) = v, \\ 0 & \text{else.} \end{cases}$$

2001 *Let d' be the VC dimension of $\mathcal{G}_{\mathcal{H}_J, \eta}$, and let d be the pseudodimension of \mathcal{H}_J . Then $d' < d$.*

2002 *Proof.* Let d' be the VC dimension of $\mathcal{G}_{\mathcal{H}_J, \eta}$, and let d be the pseudodimension of \mathcal{H}_J . First, consider
 2003 the richer hypothesis class of the set of linear thresholds induced by $\text{Round}(h(x); m)$. We will call
 2004 this class $\mathcal{G}_{\mathcal{H}_J, \eta}^{\leq}$; i.e., for any $g \in \mathcal{G}_{\mathcal{H}_J, \eta}^{\leq}$ there must be some $v \in [1/m]$ such that

$$g(x) = \begin{cases} 1 & \text{if } \text{Round}(h(x); m) \geq v, \\ 0 & \text{else.} \end{cases}$$

2005 Note that any function in $\mathcal{G}_{\mathcal{H}_J, m}^{\leq}$ can be written as an (infinite) disjunction over functions in $\mathcal{G}_{\mathcal{H}_J, m}$.
 2006 Hence, the VC dimension of $\mathcal{G}_{\mathcal{H}_J, m}^{\leq}$, which we will call d'' , must be greater than d' .

2007 We will now show that the pseudodimension of \mathcal{H}_J , d , bounds d'' . Say for contradiction that it
 2008 doesn't, and that $d < d''$. Since $d'' > d$, it must be the case that any $d + 1$ points in \mathcal{X} are shattered
 2009 by some $g \in \mathcal{G}_{\mathcal{H}_J, \eta}^{\leq}$. Say that the labels induced by g on these $d + 1$ points are (b_1, \dots, b_{d+1}) . By

2010 construction of $\mathcal{G}_{\mathcal{H}_J, \eta}^{\leq}$, there must be some $v \in [1/m]$ such that $\text{Round}(h(x_i); m) \rightarrow b_i = 1$. From
 2011 the definition of Round, this means there is some i such that $h(x_i) > i \Leftrightarrow b_i = 1$. But this is
 2012 the definition of pseudo-shattering, and hence \mathcal{H}_J must pseudo-shatter the $d + 1$ points. Hence by
 2013 contradiction $d'' < d$ and

$$d' < d'' < d.$$

2014

□

2015 **Lemma I.25** (Bound on Natarajan dimension of \mathcal{F}^R). *Let \mathcal{F}^R be the class of models that are output*
 2016 *by Algorithm I.2.1 after r rounds, and let d be the pseudodimension of \mathcal{H}_J . Then,*

$$\text{Ndim}(\mathcal{F}^R) \leq 3(m + m^7)d \log((m + m^7)d)$$

2017 *Proof.* Let r be the number of outer rounds that COLLABORATE runs for and let K be an upper
 2018 bound on any internal run of INTERNAL-BOOST. We combine the results of Lemmas I.23 and I.24:
 2019 In Lemma I.23, we showed that f^R may be written as a collection of decision rules over no more
 2020 than $\ell = m + RKm^3$ predictors in $\mathcal{G}_{\mathcal{H}_J, \eta}$. Let $d' = \text{VCdim}(\mathcal{G}_{\mathcal{H}_J, \eta})$. Plugging this in to Lemma
 2021 I.22 and using the bound from Lemma I.24,

$$\begin{aligned} \text{Ndim}(\mathcal{F}^R) &\leq 3(m + RKm^3)d' \log((2m + RKm^3)d') \\ &\leq 3(m + RKm^3)d \log((m + RKm^3)d) \quad (\text{By Lemma I.24}) \end{aligned}$$

2022 We know from Theorem I.10 that Algorithm I.2.1 will converge after no more than $R \leq m^2$ rounds
 2023 and the internal runs of INTERNAL-BOOST will run for no more than m^2 rounds. Plugging these in
 2024 as bounds on K and R , we get

$$\text{Ndim}(\mathcal{F}^R) \leq 3(m + m^7)d \log((m + m^7)d).$$

2025

□

2026 We now have all the components to prove our generalization theorem.

2027 *Proof of Theorem I.16.* This follows directly from Theorem I.21 and Lemma I.25 and suppressing the
 2028 smaller terms. □

2029 J Proofs of Tightness of Theorem B.6 from Section B

2030 We first give the formal proofs for the necessity of boundedness for weak-learning and the tightness
 2031 of quadratic guarantees. Then we show why some assumption on the joint class like the Minowski
 2032 sum one we make is necessary to get weak-learnability.

2033 J.1 Proof of Theorem B.7

2034 *Proof.* Let $\mathcal{X}_A = \mathcal{X}_B = [-1, 1]$ and $\mathcal{F}_A = \{x_A \mapsto w_A x_A : w_A \in \mathbb{R}\}$ and $\mathcal{F}_B = \{x_B \mapsto w_B x_B :$
 2035 $w_B \in \mathbb{R}\}$. Note that \mathcal{F}_A and \mathcal{F}_B are star-shaped since they are linear functions, but unbounded
 2036 since we have no bounds on the weights. For any strictly increasing function w , we will construct a
 2037 distribution such that the $w(\cdot)$ -weak-learnability condition does not hold for these function classes.

2038 Consider the following joint distribution \mathcal{D}_ρ over $\mathcal{X}_A \times \mathcal{X}_B \times \mathcal{Y}$ for any $\rho \geq 1$:

$$x_A = \frac{1}{2}\xi_A, \quad x_B = x_A + \frac{\xi_2}{2\rho} \quad \text{and} \quad y = \xi_B \quad \text{for} \quad \xi_A, \xi_B \sim_{\text{unif}} \{-1, +1\}.$$

2039 Observe that the optimal constant predictor $c^* = \mathbb{E}[Y] = 0$, giving $\min_{c \in \mathbb{R}} \mathbb{E}[(c - y)^2] =$
 2040 $\mathbb{E}[\xi_B^2] = 1$ and the optimal joint predictor is $h_J^*(x) = 2\rho x_B - 2\rho x_A = y$, yielding
 2041 $\min_{h_J \in \mathcal{H}_J} \mathbb{E}[(h_J(x) - y)^2] = 0$. This implies that

$$\min_{c \in \mathbb{R}} \mathbb{E}[(c - y)^2] - \min_{h_J \in \mathcal{H}_J} \mathbb{E}[(h_J(x) - y)^2] = 1.$$

2042 We will show that despite this, the improvement over the constant function for the optimal predictor
 2043 on either feature alone is much smaller. Observe that the the label y does not depend on x_A , hence the
 2044 optimal predictor over \mathcal{X}_A is $h_A^*(x) = 0$ which implies $\min_{h_A \in \mathcal{H}_A} \mathbb{E}[(h_A(x_A) - y)^2] = \mathbb{E}[y^2] = 1$.
 2045 This implies,

$$\min_{c \in \mathbb{R}} \mathbb{E}[(c - y)^2] - \min_{h_A \in \mathcal{H}_A} \mathbb{E}[(h_A(x_A) - y)^2] = 0 \leq w(0) < w(1).$$

2046 Here the last follows from $w(0) \in [0, 1]$ and w being strictly increasing.

2047 The label y does have correlation with x_B , and a simple calculation gives us that the optimal linear
 2048 predictor over \mathcal{X}_B has form $h_B^*(x_B) = w_B x_B$ where

$$w_B = \frac{\mathbb{E}[x_B y]}{\mathbb{E}[x_B^2]} = \frac{\frac{\mathbb{E}[\xi_A \xi_B]}{2} + \frac{\mathbb{E}[\xi_B^2]}{2\rho}}{\frac{\mathbb{E}[\xi_A^2]}{4} + \frac{\mathbb{E}[\xi_B^2]}{4\rho^2}} = \frac{2\rho}{\rho^2 + 1}.$$

2049 This gives us

$$\begin{aligned} \mathbb{E}[(h_B^*(x_B) - y)^2] &= \mathbb{E}\left[\left(\frac{2\rho}{\rho^2 + 1}x_B - \xi_B\right)^2\right] \\ &= \mathbb{E}\left[\left(\frac{\rho}{\rho^2 + 1}\xi_A - \frac{\rho^2}{\rho^2 + 1}\xi_B\right)^2\right] \\ &= \frac{\rho^2}{(\rho^2 + 1)^2} \mathbb{E}[\xi_A^2] + \frac{\rho^4}{(\rho^2 + 1)^2} \mathbb{E}[\xi_B^2] = \frac{\rho^2}{\rho^2 + 1} \end{aligned}$$

2050 This in turn implies:

$$\min_{c \in \mathbb{R}} \mathbb{E}[(c - y)^2] - \min_{h_B \in \mathcal{H}_B} \mathbb{E}[(h_B(x_B) - y)^2] = 1 - \frac{\rho^2}{\rho^2 + 1} = \frac{1}{\rho^2 + 1}.$$

2051 $w(\cdot)$ -weak learnability would require us to have $\frac{1}{\rho^2 + 1} \geq w(1)$. However, we can always choose
 2052 ρ large enough to make this not hold. In particular, any $\rho > \sqrt{\frac{1-w(1)}{w(1)}}$ will violate this condition.
 2053 Note that since w is strictly increasing, we will be guaranteed that $w(1) > w(0) \geq 0$, so such a ρ
 2054 exists. Therefore, for every fixed w , we can always construct a distribution that does not satisfy our
 2055 weak-learnability guarantee. \square

2056 J.2 Proof of Theorem B.8

2057 *Proof.* Let $\mathcal{X}_A = \mathcal{X}_B = [-1, 1]$ and $\mathcal{F}_A = \{x_A \mapsto w_A x_A : w_A \in \mathbb{R}, |w_A| \leq 1\}$ and $\mathcal{F}_B = \{x_B \mapsto$
 2058 $w_B x_B : w_B \in \mathbb{R}, |w_B| \leq 1\}$. Note that \mathcal{F}_A and \mathcal{F}_B are star-shaped since they are linear functions,
 2059 and 1-bounded since both the input and weights are bounded by 1. For any strictly increasing function
 2060 w , we will construct a distribution such that the $w(\cdot)$ -weak-learnability condition does not hold for
 2061 these function classes with respect to $\mathcal{H}_J = \{h_A + h_B : h_A \in \mathcal{H}_A, h_B \in \mathcal{H}_B\}$.

2062 We will consider the same joint distribution as in the proof of theorem B.7. We will further assume
 2063 that $\rho \geq 1$.

2064 Recall that the optimal joint predictor was $h_J(x) = 2\rho x_A - 2\rho x_B$ which required elements from
 2065 the base classes to have norm 2ρ which grows with increasing ρ . In our bounded class, however,
 2066 the optimal predictor is the scaled down version of this predictor to adhere to our norm constraints:
 2067 $h_J^*(x) = x_A - x_B = \frac{y}{2\rho}$. This gives us,

$$\mathbb{E}[(h_J^*(x) - y)^2] = \mathbb{E}\left[\left(\frac{y}{2\rho} - y\right)^2\right] = \left(1 - \frac{1}{2\rho}\right)^2 \mathbb{E}[y^2] = \frac{(2\rho - 1)^2}{4\rho^2}.$$

2068 Which in turn implies, that the gain of the joint predictor over the constant function is

$$\min_{c \in \mathbb{R}} \mathbb{E}[(c - y)^2] - \min_{h_J \in \mathcal{H}_J} \mathbb{E}[(h_J(x) - y)^2] = 1 - \frac{(2\rho - 1)^2}{4\rho^2} = \frac{4\rho - 1}{4\rho^2} \in \left[\frac{3}{4\rho}, \frac{1}{\rho}\right].$$

2069 Here the last follows from using the fact that $\rho \geq 1$.

2070 Recall that the optimal predictor over \mathcal{X}_A is $h_A^*(x_A) = 0$ which still belongs to our bounded class,
 2071 and its gain over the constant predictor was 0. The optimal predictor over \mathcal{X}_B in the unbounded case
 2072 is $h_B^*(x_B) = \frac{2\rho}{\rho^2+1}x_B$. Since $\rho^2 + 1 \geq 2\rho$ for all ρ , the norm of this predictor is actually bounded by
 2073 1. Therefore, for our bounded class, this remains an optimal predictor. The gain of this predictor over
 2074 the constant predictor is

$$\min_{c \in \mathbb{R}} \mathbb{E}[(c - y)^2] - \min_{h_B \in \mathcal{H}_B} \mathbb{E}[(h_B(x_B) - y)^2] = \frac{1}{\rho^2 + 1} \in \left[\frac{1}{2\rho^2}, \frac{1}{\rho^2} \right]$$

2075 Here the last follows from using the fact that $\rho \geq 1$.

2076 Therefore, for \mathcal{D}_ρ , the gain from the joint predictor over a constant is $\Theta(1/\rho)$ and from the best
 2077 individual predictor over constant is $\Theta(1/\rho^2)$ implying that there is no $w(\gamma) = \omega(\gamma^2)$ for this
 2078 distribution that satisfies weak-learnability. \square

2079 Finally, we establish that it is necessary to make some assumption on \mathcal{H}_J , such as the Minowski sum
 2080 structure we use—multiplicative rather than additive combinations would not work:

2081 **Theorem J.1.** *There exists classes $\mathcal{F}_A = \{f_A : \mathcal{X}_A \rightarrow \mathbb{R}\}$ and $\mathcal{F}_B = \{f_B : \mathcal{X}_B \rightarrow \mathbb{R}\}$ that are star-
 2082 shaped and 1-bounded over some domain $\mathcal{X}_A, \mathcal{X}_B$ such that $\mathcal{H}_A = \{f_A + b_A : f_A \in \mathcal{F}_A, b_A \in \mathbb{R}\}$
 2083 and $\mathcal{H}_B = \{f_B + b_B : f_B \in \mathcal{F}_B, b_B \in \mathbb{R}\}$ but do not jointly satisfy $w(\cdot)$ -weak learning with respect
 2084 to $\mathcal{H}_J = \{h_A \cdot h_B : h_A \in \mathcal{H}_A, h_B \in \mathcal{H}_B\}$ for any strictly increasing w .*

2085 *Proof.* We will consider the function classes as in the proof of Theorem B.8, that is, $\mathcal{X}_A = \mathcal{X}_B =$
 2086 $[-1, 1]$ and $\mathcal{F}_A = \{x_A \mapsto w_A x_A : w_A \in \mathbb{R}, |w_A| \leq 1\}$ and $\mathcal{F}_B = \{x_B \mapsto w_B x_B : w_B \in$
 2087 $\mathbb{R}, |w_B| \leq 1\}$. We know that this class is 1-bounded and star shaped.

2088 Now consider the following joint distribution over $\mathcal{X}_A \times \mathcal{X}_B \times \mathcal{Y}$:

$$x_A \sim_{\text{unif}} \{-1, +1\}, x_B \sim_{\text{unif}} \{-1, +1\} \text{ independent of } x_A, \text{ and } y = x_A x_B$$

2089 The best constant predictor on this is $\mathbb{E}[y] = 0$. This has loss $\mathbb{E}[y^2] = 1$. The best joint predictor for
 2090 this distribution is $h_J(x) = x_A x_B$ which can be constructed using $h_A(x_A) = x_A$ and $h_B(x_B) = x_B$.
 2091 Since this perfectly predicts the label, this has loss 0, therefore its gain over the constant predictor is 1.
 2092 However, the optimal predictor on either function alone is $h_A^*(x_A) = h_B^*(x_B) = 0$. This is because
 2093 the label is uniformly random given only information of either x_A or x_B . This implies that the gain
 2094 of the best predictor over the constant predictor is 0. This violates the weak-learning condition for
 2095 any strictly increasing w ($w(1) > w(0) \geq 0$). \square

2096 K Additional Material from Section C

2097 K.1 Calibration Preliminaries

2098 In this section we give the basic calibration definitions that we work with in our proofs.

2099 The standard measure of calibration of some sequence of predictions $\hat{y}^{1:T}$ to outcomes $y^{1:T}$ in a
 2100 sequential prediction setting is *expected calibration error*, defined as follows.

2101 **Definition K.1** (Expected Calibration Error). *Given a sequence of predictions $\hat{y}^{1:T}$ and outcomes*
 2102 *$y^{1:T}$, their expected calibration error is,*

$$\text{ECE}(\hat{y}^{1:T}, y^{1:T}) = \sum_{p \in [0,1]} \left| \sum_{t=1}^T \mathbb{1}[\hat{y}^t = p](\hat{y}^t - y^t) \right|$$

2103 *Here the outer sum is over the values p that appear in the sequence $\hat{y}^{1:T}$.*

2104 We will sometimes measure calibration error of a sequence instead using *distance to calibration*,
 2105 first defined by Błasiok et al. [2023] (we here use the definition given by Qiao and Zheng [2024]
 2106 in the sequential setting). Distance to calibration measures the ℓ_1 distance between a sequence of
 2107 predictions and the closest sequence of *perfectly calibrated* predictions.

2108 **Definition K.2** (Distance to Calibration). *Given a sequence of predictions $\hat{y}^{1:T}$ and outcomes $y^{1:T}$,*
 2109 *the distance to calibration is,*

$$\text{CalDist}(\hat{y}^{1:T}, y^{1:T}) = \min_{q^{1:T} \in \mathcal{C}(y^{1:T})} \|\hat{y}^{1:T} - q^{1:T}\|_1$$

2110 *where $\mathcal{C}(y^{1:T}) = \{q^{1:T} : \text{ECE}(q^{1:T}, y^{1:T}) = 0\}$ is the set of predictions that are perfectly calibrated*
 2111 *against outcomes $y^{1:T}$.*

2112 Calibration has a close relationship to squared error, which we will use as a potential function in
 2113 some of our analyses. Below we define the squared error of a sequence of predictions relative to a
 2114 sequence of outcomes:

2115 **Definition K.3** (Squared Error). *Given a sequence of predictions $\hat{y}^{1:T}$ and outcomes $y^{1:T}$, the*
 2116 *squared error between them is,*

$$\text{SQE}(\hat{y}^{1:T}, y^{1:T}) := \sum_{t \in [T]} (\hat{y}^t - y^t)^2.$$

2117 *We will overload this notation for the special case of constant sequences $\hat{y}^1 = \dots = \hat{y}^T = p$:*

$$\text{SQE}(p, y^{1:T}) := \sum_{t \in [T]} (p - y^t)^2.$$

2118 K.2 Conversation Calibration

2119 Here we formally define the notion of calibration introduced in Collina et al. [2025], called *conversa-*
 2120 *tion calibration*. This notion is defined over a transcript of days to $1 \dots T$ and varied-length rounds.
 2121 An agent is *conversation calibrated* if for every round k , the sequence of predictions (over days t)
 2122 that they make at round k of conversation is calibrated not just marginally, but *conditionally* on the
 2123 value of the prediction that the other agent made at round $k - 1$. We will condition on *bucketings* of
 2124 predictions.

2125 **Definition K.4** (Bucketing of the Prediction Space). *For bucket coarseness parameter n , let $B_n(i) =$*
 2126 *$[\frac{i-1}{n}, \frac{i}{n})$ and $B_n(n) = [\frac{n-1}{n}, 1]$ form a set \mathcal{B}_n of n buckets of width $1/n$ that partition the unit*
 2127 *interval.*

2128 **Definition K.5** (Conversation-Calibrated Predictions). *Fix an error function $f : \{1, \dots, T\} \rightarrow \mathbb{R}$*
 2129 *and bucketing function $g : \{1, \dots, T\} \rightarrow (0, 1]$. Given a prediction transcript $\pi^{1:T}$ resulting from an*
 2130 *interaction in the Collaboration Protocol, Bob is (f, g) -conversation-calibrated if for all even rounds*
 2131 *k and buckets $i \in \{1, \dots, 1/g(T)\}$:*

$$\text{CalDist}(\hat{y}_B^{T_A(k-1, i)}, y^{T_A(k-1, i)}) \leq f(|T_A(k-1, i)|),$$

2132 *where $T_A(k-1, i) = \{t \mid \hat{y}_A^{t, k-1} \in B_i(1/g(T))\}$ is the subsequence of days where the predictions*
 2133 *of Alice at the previous round fall in bucket i .*

2134 *Symmetrically, Alice is (f, g) -conversation-calibrated if for all odd rounds k and buckets $i \in$*
 2135 *$\{1, \dots, 1/g(T)\}$:*

$$\text{CalDist}(\hat{y}_A^{T_B(k-1, i)}, y^{T_B(k-1, i)}) \leq f(|T_B(k-1, i)|),$$

2136 *where $T_B(k-1, i) = \{t \mid \hat{y}_B^{t, k-1} \in B_i(1/g(T))\}$ is the subsequence of days where the predictions*
 2137 *of Bob at the previous round fall in bucket i .*

2138 We also introduce a function that checks whether, on a given day t and given even round k , the
 2139 prediction $\hat{y}^{t, k}$ is within ϵ of the prediction in the previous round $\hat{y}^{t, k-1}$. Formally, we define

Definition K.6 (Agreement Condition $A_{\pi^{1:T}}(t, k, \epsilon)$ and Disagreement Subsequence $D(T^k)$).

$$A_{\pi^{1:T}}(t, k) := \begin{cases} \mathbb{I}[|\hat{y}_A^{t, k} - \hat{y}_A^{t, k-1}| \leq \epsilon] & \text{if } \ell \text{ is odd,} \\ \mathbb{I}[|\hat{y}_B^{t, k} - \hat{y}_B^{t, k-1}| \leq \epsilon] & \text{if } \ell \text{ is even.} \end{cases}$$

2140 *Furthermore, let $D(T^k)$ be the subset of days t such that $A_{\pi^{1:T}}(t, k) = 0$.*

2141 We are now ready to discuss the relationship between conversation calibration and conversation swap
 2142 regret.

2143 **Theorem K.7.** *If \mathcal{H} contains all constant functions, then (f, g, \mathcal{H}) -Conversation Swap Regret implies*
 2144 *(f', g) -Conversation Calibration, where $f'(T) = \sqrt{T \cdot f(T)}$.*

2145 *Proof.* Assume that Bob satisfies (f, g, \mathcal{H}) -Conversation Swap Regret. Let $T_A(k-1, i)$ be the
 2146 subsequence of days where the predictions of Alice in round $k-1$ fall in bucket i . As \mathcal{H} contains all
 2147 constant functions, (f, g, \mathcal{H}) -Conversation Swap Regret directly implies that

$$\begin{aligned}
 & \sum_{t \in T_A(k-1, i)} (\hat{y}_k^t - y^t)^2 - \sum_v \min_{h \in \mathcal{H}_B} \left(\sum_{t \in T_A(k-1, i)} \mathbb{I}[\hat{y}_k^t = v] (h(x^t) - y^t)^2 \right) \leq f(|T_A(k-1, i)|) \\
 \Rightarrow & \sum_{t \in T_A(k-1, i)} (\hat{y}_k^t - y^t)^2 - \sum_v \min_{x^* \in [0, 1]} \left(\sum_{t \in T_A(k-1, i)} \mathbb{I}[\hat{y}_k^t = v] (x^* - y^t)^2 \right) \leq f(|T_A(k-1, i)|) \\
 \Rightarrow & \sum_v \left(\sum_{t \in T_A(k-1, i)} \mathbb{I}[\hat{y}_k^t = v] (\hat{y}_k^t - y^t)^2 - \min_{x_v^* \in [0, 1]} \sum_{t \in T_A(k-1, i)} \mathbb{I}[\hat{y}_k^t = v] (x_v^* - y^t)^2 \right) \leq f(|T_A(k-1, i)|) \\
 \Rightarrow & \sum_v \left(\sum_{t \in T_A(k-1, i)} \mathbb{I}[\hat{y}_k^t = v] (\hat{y}_k^t - y^t)^2 - \sum_{t \in T_A(k-1, i)} \mathbb{I}[\hat{y}_k^t = v] (x_v^a - y^t)^2 \right) \leq f(|T_A(k-1, i)|) \\
 & \quad \quad \quad \text{(Where } x_v^a \text{ is the average on the level set)} \\
 \Rightarrow & \sum_{t \in T_A(k-1, i)} \sum_v \mathbb{I}[\hat{y}_k^t = v] (\hat{y}_k^t - x_v^a)^2 \leq f(|T_A(k-1, i)|) \quad \quad \quad \text{(By Lemma K.12)}
 \end{aligned}$$

2148 Note that, by Cauchy-Schwartz, we have that $\sqrt{\sum_{t \in T_A(k-1, i)} \sum_v \mathbb{I}[\hat{y}_k^t = v] (\hat{y}_k^t - x_v^a)^2} \sqrt{|T_A(k-1, i)|} \geq$
 2149 $\sum_{t \in T_A(k-1, i)} \sum_v \mathbb{I}[\hat{y}_k^t = v] |\hat{y}_k^t - x_v^a|$, and therefore that $\sum_{t \in T_A(k-1, i)} \sum_v \mathbb{I}[\hat{y}_k^t = v] (\hat{y}_k^t - x_v^a)^2 \geq$
 2150 $\frac{(\sum_{t \in T_A(k-1, i)} \sum_v \mathbb{I}[\hat{y}_k^t = v] |\hat{y}_k^t - x_v^a|)^2}{|T_A(k-1, i)|}$. Thus, we can write

$$\begin{aligned}
 & \frac{(\sum_{t \in T_A(k-1, i)} \sum_v \mathbb{I}[\hat{y}_k^t = v] |\hat{y}_k^t - x_v^a|)^2}{|T_A(k-1, i)|} \leq f(|T_A(k-1, i)|) \\
 \Rightarrow & \frac{\sum_{t \in T_A(k-1, i)} \sum_v \mathbb{I}[\hat{y}_k^t = v] |\hat{y}_k^t - x_v^a|}{\sqrt{|T_A(k-1, i)|}} \leq \sqrt{f(|T_A(k-1, i)|)} \\
 & \quad \quad \quad \text{(Taking the square root of both sides)} \\
 \Rightarrow & \sum_{t \in T_A(k-1, i)} \sum_v \mathbb{I}[\hat{y}_k^t = v] |\hat{y}_k^t - x_v^a| \leq \sqrt{f(|T_A(k-1, i)|) \cdot |T_A(k-1, i)|} \\
 \Rightarrow & ECE(\hat{y}_k^{T_A(k-1, i)}, y^{T_A(k-1, i)}) \leq \sqrt{f(|T_A(k-1, i)|) \cdot |T_A(k-1, i)|} \\
 & \quad \quad \quad \text{(As the LHS is exactly ECE)} \\
 \Rightarrow & CalDist(\hat{y}_k^{T_A(k-1, i)}, y^{T_A(k-1, i)}) \leq \sqrt{f(|T_A(k-1, i)|) \cdot |T_A(k-1, i)|} \\
 & \quad \quad \quad \text{(As ECE upper bounds CalDist)}
 \end{aligned}$$

2151 As Conversation Swap Regret holds true for all rounds, this implies
 2152 $\sqrt{|T_A(k-1, i)| \cdot f(|T_A(k-1, i)|)}$ -conversation calibration. The proof holds symmetrically
 2153 for Alice. \square

2154 **Theorem K.8.** *If a sequence \hat{y}_k has (f, g, \mathcal{H}) -Conversation Swap Regret, then*

$$\sum_{t=1}^T (\hat{y}_k^t - y^t)^2 - \sum_v \min_{h \in \mathcal{H}} \left(\sum_{t=1}^T \mathbb{I}[\hat{y}_k^t = v] (h(x^t) - y^t)^2 \right) \leq \frac{f(g(T)T)}{g(T)}.$$

Proof.

$$\begin{aligned}
& \sum_{t=1}^T (\hat{y}_k^t - y^t)^2 - \sum_v \min_{h \in \mathcal{H}_B} \left(\sum_{t=1}^T \mathbb{I}[\hat{y}_k^t = v] (h(x^t) - y^t)^2 \right) = \\
& \sum_i \sum_{t \in T_A(k-1, i)} (\hat{y}_k^t - y^t)^2 - \sum_v \min_{h \in \mathcal{H}_B} \left(\sum_i \sum_{t \in T_A(k-1, i)} \mathbb{I}[\hat{y}_k^t = v] (h(x^t) - y^t)^2 \right) \\
& \leq \sum_i \sum_{t \in T_A(k-1, i)} (\hat{y}_k^t - y^t)^2 - \sum_i \sum_v \min_{h \in \mathcal{H}_B} \left(\sum_{t \in T_A(k-1, i)} \mathbb{I}[\hat{y}_k^t = v] (h(x^t) - y^t)^2 \right) \\
& \quad \text{(As by moving the sum over } i \text{ out of the min we are only strengthening the benchmark)} \\
& = \sum_i \left(\sum_{t \in T_A(k-1, i)} (\hat{y}_k^t - y^t)^2 - \sum_v \min_{h \in \mathcal{H}_B} \left(\sum_{t \in T_A(k-1, i)} \mathbb{I}[\hat{y}_k^t = v] (h(x^t) - y^t)^2 \right) \right) \\
& = \sum_i (f(|T_A(k-1, i)|)) \quad \text{(By the Conversation Swap Regret Condition)} \\
& \leq \frac{f(g(T)T)}{g(T)} \quad \text{(By the assumption that } f \text{ is concave)}
\end{aligned}$$

2155

□

2156 K.3 Additional Online Preliminaries

2157 **Definition K.9** (\mathcal{Z} -valued Tree). A \mathcal{Z} -valued tree \mathbf{z} of depth n is a rooted complete binary tree
2158 with nodes labeled by elements of \mathcal{Z} . We identify the tree \mathbf{z} with the sequence $(\mathbf{z}_1, \dots, \mathbf{z}_n)$ of labeling
2159 functions $\mathbf{z}_i : \{\pm 1\}^{i-1} \rightarrow \mathcal{Z}$ which provide the labels for each node. Here, $\mathbf{z}_1 \in \mathcal{Z}$ is the root of the
2160 tree, while $\mathbf{z}_i, i > 1$ is the label of the node obtained by following the path of length $i - 1$ from the
2161 root, with $+1$ indicating ‘right’ and -1 indicating ‘left.’

2162 **Definition K.10.** A \mathcal{Z} -valued tree \mathbf{z} of depth d is shattered by a function class $\mathcal{F} \subseteq \{\pm 1\}^{\mathcal{Z}}$ if

$$\forall \varepsilon \in \{\pm 1\}^d, \exists f \in \mathcal{F} \text{ s.t. } \forall t \in [d], f(\mathbf{z}_t(\varepsilon)) = \varepsilon_t.$$

2163 **Definition K.11** (Sequential Fat Shattering Dimension [Rakhlin et al., 2014]). A \mathcal{Z} -valued binary
2164 tree \mathbf{z} of depth d is α -shattered by a function class $\mathcal{F} \subseteq \mathbb{R}^{\mathcal{Z}}$ if there exists an \mathbb{R} -valued tree \mathbf{s} of
2165 depth d such that

$$\forall \varepsilon \in \{\pm 1\}^d, \exists f \in \mathcal{F} \text{ s.t. } \forall t \in [d], \varepsilon_t(f(\mathbf{z}_t(\varepsilon)) - \mathbf{s}_t(\varepsilon)) \geq \alpha/2.$$

2166 The sequential fat shattering dimension $\text{FAT}_\alpha(\mathcal{F}, \mathcal{Z})$ at scale α is the maximal d such that \mathcal{F}
2167 α -shatters a \mathcal{Z} -valued tree of depth d .

2168 K.4 Proof of Theorem C.1

2169 **Lemma K.12** (Lemma A.1 from Collina et al. [2025]). If $m = \frac{1}{T} \sum_{t=1}^T y^t$, then for any constant x ,

$$\text{SQE}(x, y^{1:T}) - \text{SQE}(m, y^{1:T}) = \sum_{t=1}^T (x - m)^2 \quad (1)$$

2170

2171 **Lemma K.13** (Lemma A.2 from Collina et al. [2025]). Let $T_k^{i, p_h} = \{t : \hat{y}_B^{t, k} = p_h \text{ and } \hat{y}_A^{t, k-1} \in$
2172 $B_i(\frac{1}{g(T)})\}$ be the subsequence of days such that the predicts p_h in round k and the model predicts in
2173 bucket $B_i(\frac{1}{g(T)})$ in round $k - 1$. If the human is $(\cdot, g_h(T))$ -conversation calibrated, then

$$\sum_{t \in T_k^{i, p_h}} (\hat{y}_A^{t, k-1} - y^t)^2 - \sum_{t \in T_k^{i, p_h}} (i \cdot g_h(T) - y^t)^2 \geq -g_h(T) \cdot |T_k^{i, p_h}| \quad (2)$$

2174

Lemma K.14 (Lemma A.3 from Collina et al. [2025]). *Consider any sequence of predictions and labels $p^{1:T}, y^{1:T}$ and some other sequence of predictions $q^{1:T}$ such that $\|p^{1:T} - q^{1:T}\| \leq \gamma$. Then,*

$$\sum_{t=1}^T (q^t - y^t)^2 - \sum_{t=1}^T (p^t - y^t)^2 \leq 3\gamma$$

2175

2176 **Lemma K.15.** *If Bob is $(0, g_B(T))$ -conversation-calibrated, then for any even k ,*

$$\text{SQE}(\hat{y}_B^{T,k}, y^{1:T}) \leq \text{SQE}(\hat{y}_A^{T,k-1}, y^{1:T}) - (\epsilon - g_B(T))^2 |D(T^k)| + g_B(T)T$$

2177 *And if Alice is $(0, g_A(T))$ -conversation-calibrated, for any odd k ,*

$$\text{SQE}(\hat{y}_A^{T,k}, y^{1:T}) \leq \text{SQE}(\hat{y}_B^{T,k-1}, y^{1:T}) - (\epsilon - g_A(T))^2 |D(T^k)| + g_A(T)T$$

2178 *Proof.* Let $T_k^{i,p_h} = \{t : t \in T^{\geq k} \text{ and } \hat{y}_B^{t,k} = p_h \text{ and } \hat{y}_A^{t,k-1} \in B_i(\frac{1}{g(T)})\}$ be the subsequence of
 2179 days such that Bob predicts p_h in round k and Alice predicts in bucket $B_i(\frac{1}{g(T)})$ in round $k-1$. Let

2180 $m_k^{i,p_h} = \frac{\sum_{t \in T_k^{i,p_h}} y^t}{|T_k^{i,p_h}|}$ be the true mean on this subsequence. The difference in squared errors over
 2181 this subsequence can be written as:

$$\begin{aligned} & \sum_{t \in T_k^{i,p_h}} (\hat{y}_A^{t,k-1} - y^t)^2 - \sum_{t \in T_k^{i,p_h}} (\hat{y}_B^{t,k} - y^t)^2 \\ &= \left[\sum_{t \in T_k^{i,p_h}} (\hat{y}_A^{t,k-1} - y^t)^2 - \sum_{t \in T_k^{i,p_h}} (m_k^{i,p_h} - y^t)^2 \right] - \left[\sum_{t \in T_k^{i,p_h}} (\hat{y}_B^{t,k} - y^t)^2 - \sum_{t \in T_k^{i,p_h}} (m_k^{i,p_h} - y^t)^2 \right] \\ & \quad \text{(Adding and subtracting } \sum_{t \in T_k^{i,p_h}} (m_k^{i,p_h} - y^t)^2) \\ &\geq \left[\sum_{t \in T_k^{i,p_h}} (i \cdot g_B(T) - y^t)^2 - |T_k^{i,p_h}| \cdot g_B(T) - \sum_{t \in T_k^{i,p_h}} (m_k^{i,p_h} - y^t)^2 \right] - \\ & \quad \left[\sum_{t \in T_k^{i,p_h}} (\hat{y}_B^{t,k} - y^t)^2 - \sum_{t \in T_k^{i,p_h}} (m_k^{i,p_h} - y^t)^2 \right] \quad \text{(By Lemma K.13)} \\ &= \left[\sum_{t \in T_k^{i,p_h}} (i \cdot g_B(T) - m_k^{i,p_h})^2 - |T_k^{i,p_h}| \cdot g_B(T) \right] - \left[\sum_{t \in T_k^{i,p_h}} (\hat{y}_B^{t,k} - y^t)^2 - \sum_{t \in T_k^{i,p_h}} (m_k^{i,p_h} - y^t)^2 \right] \\ & \quad \text{(By Lemma K.12)} \\ &= \left[\sum_{t \in T_k^{i,p_h}} (i \cdot g_B(T) - m_k^{i,p_h})^2 - |T_k^{i,p_h}| \cdot g_B(T) \right] - \left[\sum_{t \in T_k^{i,p_h}} (p_h - y^t)^2 - \sum_{t \in T_k^{i,p_h}} (m_k^{i,p_h} - y^t)^2 \right] \\ & \quad \text{(As by definition of } T_k^{i,p_h}, \hat{y}_B^{t,k} = p_h) \\ &\geq \left[\sum_{t \in T_k^{i,p_h}} (i \cdot g_B(T) - m_k^{i,p_h})^2 - |T_k^{i,p_h}| \cdot g_B(T) \right] - \left[\sum_{t \in T_k^{i,p_h}} (p_h - m_k^{i,p_h})^2 \right] \\ & \quad \text{(By Lemma K.12)} \\ &\geq -|T_k^{i,p_h}| \cdot g_B(T) + \sum_{t \in T_k^{i,p_h}} (i \cdot g_B(T) - p_h)^2 \\ & \quad \text{(As Bob is } (0, g_B(T))\text{-conversation calibrated, } p_h = m_k^{i,p_h}) \end{aligned}$$

2182 Using this analysis, we can write the difference in squared errors over the entire sequence $\hat{y}_B^{T,k}$ and
 2183 $\hat{y}_A^{T,k-1}$ as follows, where the first term comes from summing the above expression over all i, p_h :

$$\begin{aligned}
 & \text{SQE}(\bar{p}_A^{T,k-1}, y^{1:T}) - \text{SQE}(\bar{p}_B^{T,k}, y^{1:T}) \\
 &= \sum_{\forall i, p_h} \left(\sum_{t \in T_k^{i, p_h}} (\hat{y}_A^{t, k-1} - y^t)^2 - \sum_{t \in T_k^{i, p_h}} (\hat{y}_B^{t, k} - y^t)^2 \right) \\
 &= \sum_{\forall i, p_h} \left(-|T_k^{i, p_h}| \cdot g_B(T) + \sum_{t \in T_k^{i, p_h}} (i \cdot g_B(T) - p_h)^2 \right) \quad (\text{by the analysis above}) \\
 &= -g_B(T)T + \sum_{\forall i, p_h} \sum_{t \in T_k^{i, p_h}} (i \cdot g_B(T) - p_h)^2 \\
 &\quad \quad \quad (\text{As } g_B(T) \text{ is independent of } i \text{ and } p_h, \text{ and } \sum_{\forall i, p_h} |T_k^{i, p_h}| = T) \\
 &\geq -g_B(T)T + \sum_{\forall i, p_h} \sum_{t \in T_k^{i, p_h}} \mathbb{1}[|i \cdot g_B(T) - \hat{y}_B^{t, k}| \geq \epsilon - g_B(T)] (i \cdot g_B(T) - p_h)^2 \\
 &\geq -g_B(T)T + (\epsilon - g_B(T))^2 \sum_{\forall i, p_h} \sum_{t \in T_k^{i, p_h}} \mathbb{1}[|i \cdot g_B(T) - \hat{y}_B^{t, k}| \geq \epsilon - g_B(T)]
 \end{aligned}$$

2184 Note that, for all days in the subsequence T_k^{i, p_h} , in round $k-1$ Alice predicted in bucket $B_i(\frac{1}{g_B(T)}) =$
 2185 $i \cdot g_B(T)$, and therefore in each of these days, by the definition of our bucketing, $\hat{y}_A^{t, k-1} \geq (i-1) \cdot$
 2186 $g_B(T)$ and $\hat{y}_A^{t, k-1} \leq i \cdot g_B(T)$. So consider any round $t \in T_k^{i, p_h}$. If $|\hat{y}_B^{t, k} - \hat{y}_A^{t, k-1}| \geq \epsilon$, then we
 2187 have:

$$\begin{aligned}
 |\hat{y}_B^{t, k} - \hat{y}_A^{t, k-1}| &\leq |\hat{y}_B^{t, k} - i \cdot g_B(T)| + |i \cdot g_B(T) - \hat{y}_A^{t, k-1}| \\
 &= |\hat{y}_B^{t, k} - i \cdot g_B(T)| + i \cdot g_B(T) - \hat{y}_A^{t, k-1} \\
 &\leq |\hat{y}_B^{t, k} - i \cdot g_B(T)| + i \cdot g_B(T) - (i-1) \cdot g_B(T) \\
 &= |\hat{y}_B^{t, k} - i \cdot g_B(T)| + g_B(T), \\
 \implies |\hat{y}_B^{t, k} - i \cdot g_B(T)| &\geq |\hat{y}_B^{t, k} - \hat{y}_A^{t, k-1}| - g_B(T) \geq \epsilon - g_B(T).
 \end{aligned}$$

2188 Thus, if $|\hat{y}_B^{t, k} - \hat{y}_A^{t, k-1}| \geq \epsilon$, then $|i \cdot g_B(T) - \hat{y}_B^{t, k}| \geq \epsilon - g_B(T), \forall t \in T_k^{i, p_h}$. Therefore the set of
 2189 days for which the former condition holds is a subset of the latter condition, and we can write

$$\begin{aligned}
 & -g_B(T)T + (\epsilon - g_B(T))^2 \sum_{\forall i, p_h} \mathbb{1}[|i \cdot g_B(T) - p_h| \geq \epsilon - g_B(T)] \cdot |T_k^{i, p_h}| \\
 &\geq -g_B(T)T + (\epsilon - g_B(T))^2 \sum_{\forall i, p_h} \sum_{t \in T_k^{i, p_h}} \mathbb{1}[|\hat{y}_B^{t, k} - \hat{y}_A^{t, k-1}| \geq \epsilon] \\
 &= -g_B(T)T + (\epsilon - g_B(T))^2 |D(T^k)| \\
 &\quad (\text{As on every day and round where there is not agreement, Bob and Alice disagreed by at least } \epsilon)
 \end{aligned}$$

2190 As Bob and Alice are perfectly symmetrical, we also obtain the symmetrical result for Alice. \square

2191 **Theorem K.16.** *If Bob is $(f_B(\cdot), g_B(\cdot))$ -conversation-calibrated, then after engaging in the collabo-*
 2192 *ration protocol for T days:*

$$\text{SQE}(\hat{y}_B^{T, k}, y^{1:T}) \leq \text{SQE}(\hat{y}_A^{T, k-1}, y^{1:T}) - (\epsilon - g_B(T))^2 |D(T^k)| + g_B(T)T + 3 \frac{f_B(g_B(T) \cdot T)}{g_B(T)}$$

2193 And if Alice is $(f_A(\cdot), g_A(\cdot))$ -conversation-calibrated, then after engaging in the collaboration
 2194 protocol for T days:

$$\text{SQE}(\hat{y}_A^{T,k}, y^{1:T}) \leq \text{SQE}(\hat{y}_B^{T,k-1}, y^{1:T}) - (\epsilon - g_A(T))^2 |D(T^k)| + g_A(T)T + 3 \frac{f_A(g_A(T) \cdot T)}{g_A(T)}$$

2195 *Proof.* Let $T_m(k, i) = \{t : \hat{y}_A^{t,k-1} \in B_i(\frac{1}{g_B(T)})\}$ be the subsequence of days in which Alices
 2196 predicts in bucket $B_i(\frac{1}{g_B(T)})$ at round $k-1$.

2197 Note that Bob has distance to calibration of $f_B(|T_m(k, i)|)$ on every such subsequence defined this
 2198 way. Therefore, for predictions $p_h^{1:T,k}$ from Bob at round k :

$$\begin{aligned} \text{CalDist}(p_h^{T \geq k, k}, y^{1:T}) &= \min_{q^{1:T} \in C(y^{1:T})} \|p_h^{T \geq k, k} - q^{1:T}\|_1 \\ &\leq \sum_{i=1}^{\frac{1}{g_B(T)}} \min_{q^{1:T_m(k,i)} \in C^{T_m(k,i)}(y^{1:T})} \|p^{1:T} - q^{1:T}\|_1 \\ &\leq \sum_{i=1}^{\frac{1}{g_B(T)}} f_B(|T_m(k, i)|) \quad (\text{By the calibration distance of Bob}) \\ &\leq \frac{f_B(g_B(T) \cdot |T^{\geq k}|)}{g_B(T)} \quad (\text{By the assumption that } f_B \text{ is concave}) \\ &\leq \frac{f_B(g_B(T) \cdot T)}{g_B(T)} \end{aligned}$$

2199 Let q^k be the set of perfectly calibrated predictions that are $f_B(|T_m(k, i)|)$ -close to $p_h^{1:T,k}$. Then, we
 2200 have that

$$\begin{aligned} &\text{SQErr}(p_h^{T,k}, y^{1:T}) - \text{SQErr}(p_m^{T,k-1}, y^{1:T}) \\ &\leq \text{SQErr}(q^k, y^{1:T}) - \text{SQErr}(p_h^{T,k-1}, y^{1:T}) + 3 \frac{f_B(g_B(T) \cdot T)}{g_B(T)} \quad (\text{By Lemma K.14}) \\ &\leq -(\epsilon - g_B(T))^2 |D(T^k)| + g_B(T)T + 3 \frac{f_B(g_B(T) \cdot T)}{g_B(T)}. \quad (\text{By Lemma K.15}) \end{aligned}$$

2201 As Bob and Alice are symmetric, we also obtain the symmetric result for Alice. \square

2202 *Proof of Theorem C.1.* By composing the two results in Theorem K.16, we see that

$$\begin{aligned} &\text{SQErr}(\hat{y}_B^{T,k-2}, y^{1:T}) - \text{SQErr}(\hat{y}_B^{T,k}, y^{1:T}) \\ &\geq (\epsilon - g_B(T))^2 |D(T^k)| + (\epsilon - g_A(T))^2 |D(T^{k-1})| - g_A(T)T - 3 \frac{f_A(g_A(T) \cdot T)}{g_A(T)} - g_B(T)T - 3 \frac{f_B(g_B(T) \cdot T)}{g_B(T)} \\ &\geq (\epsilon - g_B(T))^2 |D(T^k)| + (\epsilon - g_A(T))^2 |D(T^{k-1})| - (g_A(T) + g_B(T))T - 3 \left(\frac{f_A(g_A(T) \cdot T)}{g_A(T)} + \frac{f_B(g_B(T) \cdot T)}{g_B(T)} \right). \end{aligned}$$

2203 Now we can apply the above expression recursively for k rounds in order to bound the total number
 2204 of days of disagreement:

$$\begin{aligned}
& \text{SQErr}(\hat{y}_B^{T,k}, y^{1:T}) \\
& \leq \text{SQErr}(\hat{y}_B^{T,2}, y^{1:T}) - (\epsilon - g_A(T))^2 \left(\sum_{k=1, k \text{ odd}}^k |D(T^k)| \right) - (\epsilon - g_B(T))^2 \left(\sum_{k=1, k \text{ even}}^k |D(T^k)| \right) \\
& \quad + (g_A(T) + g_B(T))rT + 3 \left(\frac{f_A(g_A(T) \cdot T)}{g_A(T)} + \frac{f_B(g_B(T) \cdot T)}{g_B(T)} \right) \left(\sum_{k=1, k \text{ even}}^k 1 \right) \\
& \leq \text{SQErr}(\hat{y}_B^{T,2}, y^{1:T}) - ((\epsilon - g_A(T))^2 + (\epsilon - g_B(T))^2) \left(\sum_{k=1}^k |D(T^k)| \right) \\
& \quad + (g_A(T) + g_B(T))rT + 3 \left(\frac{f_A(g_A(T) \cdot T)}{g_A(T)} + \frac{f_B(g_B(T) \cdot T)}{g_B(T)} \right) \frac{k}{2} \\
& \leq \text{SQErr}(\hat{y}_B^{T,2}, y^{1:T}) - 2\epsilon^2 \left(\sum_{k=1}^k |D(T^k)| \right) + 3k(g_A(T) + g_B(T))T + 3k \left(\frac{f_A(g_A(T) \cdot T)}{g_A(T)} + \frac{f_B(g_B(T) \cdot T)}{g_B(T)} \right) \\
& = \text{SQErr}(\hat{y}_B^{T,2}, y^{1:T}) - 2\epsilon^2 \left(\sum_{k=1}^k |D(T^k)| \right) + 3kT\beta(T, f_A, f_B)
\end{aligned}$$

2205 Finally we can compose this expression with one more instantiation of Theorem K.16:

$$\begin{aligned}
\text{SQE}(\hat{y}_B^{T,2}, y^{1:T}) & \leq \text{SQE}(\hat{y}_A^{T,1}, y^{1:T}) - (\epsilon - g_B(T))^2 |D(T^1)| + g_B(T)T + 3 \frac{f_B(g_B(T) \cdot T)}{g_B(T)} \\
& \leq \text{SQE}(\hat{y}_A^{T,1}, y^{1:T}) - \epsilon^2 |D(T^1)| + T\beta(T, f_A, f_B)
\end{aligned}$$

2206 and get a final expression of:

$$\text{SQErr}(\hat{y}_B^{T,k}, y^{1:T}) \leq \text{SQE}(\hat{y}_A^{T,1}, y^{1:T}) - 2\epsilon^2 \left(\sum_{k=1}^k |D(T^k)| \right) + 3kT\beta(T, f_A, f_B)$$

2207 Note also that $\text{SQE}(\hat{y}_A^{T,1}, y^{1:T}) \leq T$ and $\text{SQE}(\hat{y}_A^{T,k}, y^{1:T}) \geq 0$. Therefore, we have that

$$\begin{aligned}
0 & \leq T - 2\epsilon^2 \left(\sum_{k=1}^k |D(T^k)| \right) + rT\beta(T, f_A, f_B) \\
& \implies \sum_{k=1}^k |D(T^k)| \leq \frac{T + rT\beta(T, f_A, f_B)}{2\epsilon^2}
\end{aligned}$$

Thus, the round between 1 and k with the smallest number of disagreements has no more than $\frac{T + rT\beta(T, f_A, f_B)}{2r\epsilon^2}$ disagreements. Let k be the index of this round. As there are T predictions total in round k , the fraction of predictions in the round that are disagreements is

$$\frac{T + rT\beta(T, f_A, f_B)}{2rT\epsilon^2} = \frac{1}{2r\epsilon^2} + \frac{\beta(T, f_A, f_B)}{2\epsilon^2}$$

2208

□

2209 K.5 Proof of Theorem C.2

Lemma K.17. *If the sequence of real-valued predictions $a^{1:T}$ is (ϵ, δ) -close to the sequence $b^{1:T}$, and a, b, y are all bounded above by 1, then*

$$\sum_{t=1}^T (a^t - y^t)^2 - \sum_{t=1}^T (b^t - y^t)^2 \leq 4(\delta + \epsilon)T$$

2210

Proof.

$$\begin{aligned}
& \sum_{t=1}^T (a^t - y^t)^2 - \sum_{t=1}^T (b^t - y^t)^2 \\
&= \sum_{t=1}^T \mathbb{1}[|a^t - b^t| \geq \epsilon] ((a^t - y^t)^2 - (b^t - y^t)^2) + \sum_{t=1}^T \mathbb{1}[|a^t - b^t| < \epsilon] ((a^t - y^t)^2 - (b^t - y^t)^2) \\
&\leq \sum_{t=1}^T \mathbb{1}[|a^t - b^t| \geq \epsilon] (|a^t - b^t| \cdot |a^t + b^t| + 2|y^t| \cdot |a^t - b^t|) \\
&\quad + \sum_{t=1}^T \mathbb{1}[|a^t - b^t| < \epsilon] (|a^t - b^t| \cdot |a^t + b^t| + 2|y^t| \cdot |a^t - b^t|) \\
&\leq \sum_{t=1}^T \mathbb{1}[|a^t - b^t| \geq \epsilon] (|a^t + b^t| + 2|y^t|) + \sum_{t=1}^T \mathbb{1}[|a^t - b^t| < \epsilon] (\epsilon \cdot |a^t + b^t| + 2|y^t| \cdot \epsilon) \\
&\leq \sum_{t=1}^T \mathbb{1}[|a^t - b^t| \geq \epsilon] (4) + \sum_{t=1}^T \mathbb{1}[|a^t - b^t| < \epsilon] (4 \cdot \epsilon) \quad (\text{By the upper bounds on the values}) \\
&\leq 4\delta T + 4\epsilon(1 - \delta)T \leq 4T(\delta + \epsilon)
\end{aligned}$$

2211

□

2212 *Proof of Theorem C.2.* By Theorem K.7, Alice is (f'_A, g_A) -conversation calibrated and Bob is
2213 (f'_B, g_B) -conversation calibrated, where $f'_A(x) = \sqrt{x \cdot f_A(x)}$, and symmetrically for f'_B . Thus, by
2214 Theorem C.1, after the collaboration protocol is run for K rounds, there is at least one round $k+1 > 1$
2215 where the fraction of predictions that are ϵ -far from the previous round is at most $\frac{1}{2K\epsilon^2} + \frac{\beta(T, f'_A, f'_B)}{2\epsilon^2}$,
2216 where $\beta(T, f'_A, f'_B) = 3 \left(g_A(T) + g_B(T) + \frac{f'_A(g_A(T) \cdot T)}{g_A(T) \cdot T} + \frac{f'_B(g_B(T) \cdot T)}{g_B(T) \cdot T} \right)$. Consider the round be-
2217 fore round $k+1$, round k .

2218 First consider the case where k is an even round. Then, by definition, the predictions $\hat{y}_k^1, \dots, \hat{y}_k^T$ in
2219 this round have (f_B, g_B, H_B) -conversation swap regret. We will now define a sequence of predictions
2220 \bar{y} which is $g_B T$ -far in L_1 distance from $\hat{y}_k^1, \dots, \hat{y}_k^T$, and show that \bar{y} has low swap regret to $\mathcal{H}_A \cup \mathcal{H}_B$.
2221 This sequence is generated by combining level sets of $\hat{y}_k^1, \dots, \hat{y}_k^T$ such that each level set is mapped
2222 to the closest value in $\{\frac{1}{g_A(T)}, \dots, 1\}$. We will first compute the swap regret of \bar{y} with respect to \mathcal{H}_B :

$$\begin{aligned}
& \sum_{t=1}^T (\bar{y}^t - y^t)^2 - \sum_v \min_{h \in \mathcal{H}_B} \left(\sum_{t=1}^T \mathbb{1}[\bar{y}^t = v] (h(x^t) - y^t)^2 \right) \\
& \leq \sum_{t=1}^T (\bar{y}^t - y^t)^2 - \sum_v \min_{h \in \mathcal{H}_B} \left(\sum_{t=1}^T \mathbb{1}[\hat{y}_k^t = v] (h(x^t) - y^t)^2 \right) \\
& \quad \text{(As } \bar{y} \text{ has strictly coarser level sets than } \hat{y}_k, \text{ here we are only strengthening the benchmark)} \\
& = \sum_{t=1}^T (\hat{y}_k^t - y^t)^2 - \sum_v \min_{h \in \mathcal{H}_B} \left(\sum_{t=1}^T \mathbb{1}[\hat{y}_k^t = v] (h(x^t) - y^t)^2 \right) + \left(\sum_{t=1}^T (\bar{y}^t - y^t)^2 - \sum_{t=1}^T (\hat{y}_k^t - y^t)^2 \right) \\
& \leq \frac{f_B(g_B(T)T)}{g_B(T)} + \left(\sum_{t=1}^T (\bar{y}^t - y^t)^2 - \sum_{t=1}^T (\hat{y}_k^t - y^t)^2 \right) \quad \text{(By Theorem K.8)} \\
& = \frac{f_B(g_B(T)T)}{g_B(T)} + \sum_{t=1}^T ((\bar{y}^t)^2 - (\hat{y}_k^t)^2 + 2y^t(y_k^t - \bar{y}^t)) \\
& \leq \frac{f_B(g_B(T)T)}{g_B(T)} + \sum_{t=1}^T (|\bar{y}^t - \hat{y}_k^t| \cdot |\bar{y}^t + \hat{y}_k^t| + 2|y^t| \cdot |y_k^t - \bar{y}^t|) \\
& = \frac{f_B(g_B(T)T)}{g_B(T)} + \sum_{t=1}^T \left(\frac{g_A(T)}{2} \cdot |\bar{y}^t + \hat{y}_k^t| + 2|y^t| \cdot \frac{g_A(T)}{2} \right) \quad \text{(By construction of } \bar{y}) \\
& \leq \frac{f_B(g_B(T)T)}{g_B(T)} + \sum_{t=1}^T \left(\frac{3}{2} g_A(T) \right) = \frac{f_B(g_B(T) \cdot T)}{g_B(T)} + \frac{3g_A(T) \cdot T}{2} \quad \text{(As } y \leq 1)
\end{aligned}$$

2223 Next, we will compute the swap regret of \bar{y} with respect to \mathcal{H}_A . Here, we crucially use the fact that
 2224 the sequence \hat{y}_{k+1} has high agreement with \hat{y}_k , and furthermore that \hat{y}_{k+1} has low swap regret to
 2225 \mathcal{H}_A *exactly* on the level sets of \bar{y} . Let $T_B(k, i)$ be the subsequence of days on which Bob predicts in
 2226 bucket i in round k .

$$\begin{aligned}
& \sum_{t=1}^T (\bar{y}^t - y^t)^2 - \sum_v \min_{h \in \mathcal{H}_A} \left(\sum_{t=1}^T \mathbb{1}[\bar{y}^t = v] (h(x^t) - y^t)^2 \right) \\
&= \left(\sum_{t=1}^T (\bar{y}^t - y^t)^2 - \sum_{t=1}^T (\hat{y}_{k+1}^t - y^t)^2 \right) + \sum_{t=1}^T (\hat{y}_{k+1}^t - y^t)^2 - \sum_v \min_{h \in \mathcal{H}_A} \left(\sum_{t=1}^T \mathbb{1}[\bar{y}^t = v] (h(x^t) - y^t)^2 \right) \\
&= 4(\epsilon + g_A(T) + \frac{1}{2K\epsilon^2} + \frac{\beta(T, f'_A, f'_B)}{2\epsilon^2})T + \sum_{t=1}^T (\hat{y}_{k+1}^t - y^t)^2 - \sum_v \min_{h \in \mathcal{H}_A} \left(\sum_{t=1}^T \mathbb{1}[\bar{y}^t = v] (h(x^t) - y^t)^2 \right) \\
&\quad \text{(By K.17, and the fact that } \hat{y}_{k+1} \text{ is } (\epsilon + g_A(T), \frac{1}{2K\epsilon^2} + \frac{\beta(T, f'_A, f'_B)}{2\epsilon^2})\text{-close to } \bar{y}) \\
&= 4(\epsilon + g_A(T) + \frac{1}{2K\epsilon^2} + \frac{\beta(T, f'_A, f'_B)}{2\epsilon^2})T + \sum_i \sum_{t \in T_B(r, i)} (\hat{y}_{k+1}^t - y^t)^2 - \sum_v \min_{h \in \mathcal{H}_A} \left(\sum_{t=1}^T \mathbb{1}[\bar{y}^t = v] (h(x^t) - y^t)^2 \right) \\
&= 4(\epsilon + g_A(T) + \frac{1}{2K\epsilon^2} + \frac{\beta(T, f'_A, f'_B)}{2\epsilon^2})T + \sum_i \sum_{t \in T_B(r, i)} (\hat{y}_{k+1}^t - y^t)^2 - \sum_i \min_{h \in \mathcal{H}_A} \left(\sum_{t \in T_B(r, i)} (h(x^t) - y^t)^2 \right) \\
&\quad \text{(As } \bar{y} \text{ attains a particular value exactly when } t \in T_B(r, i); \text{ that is, when Bob predicts in bucket } i \text{ in round } k.) \\
&\leq 4(\epsilon + g_A(T) + \frac{1}{2K\epsilon^2} + \frac{\beta(T, f'_A, f'_B)}{2\epsilon^2})T + \sum_i \sum_{t \in T_B(r, i)} (\hat{y}_{k+1}^t - y^t)^2 \\
&\quad - \sum_i \sum_v \min_{h \in \mathcal{H}_A} \left(\sum_{t \in T_B(r, i)} \mathbb{1}[\hat{y}_{k+1}^t = v] (h(x^t) - y^t)^2 \right) \\
&\quad \text{(As we are only making the benchmark more powerful)} \\
&\leq 4(\epsilon + g_A(T) + \frac{1}{2K\epsilon^2} + \frac{\beta(T, f'_A, f'_B)}{2\epsilon^2})T \\
&\quad + \sum_i \left(\sum_{t \in T_B(r, i)} (\hat{y}_{k+1}^t - y^t)^2 - \sum_v \min_{h \in \mathcal{H}_A} \left(\sum_{t \in T_B(r, i)} \mathbb{1}[\hat{y}_{k+1}^t = v] (h(x^t) - y^t)^2 \right) \right) \\
&\leq 4(\epsilon + g_A(T) + \frac{1}{2K\epsilon^2} + \frac{\beta(T, f'_A, f'_B)}{2\epsilon^2})T + \sum_i (f_A(|T_B(r, i)|)) \\
&\quad \text{(By the conversation swap regret of Alice)} \\
&= 4(\epsilon + g_A(T) + \frac{1}{2K\epsilon^2} + \frac{\beta(T, f'_A, f'_B)}{2\epsilon^2})T + \frac{f_A(g_A(T) \cdot T)}{g_A(T)} \quad \text{(By the concavity of } f_A)
\end{aligned}$$

Thus, \bar{y} simultaneously has $(4(\epsilon + g_A(T) + \frac{1}{2K\epsilon^2} + \frac{\beta(T, f'_A, f'_B)}{2\epsilon^2})T + \frac{f_A(g_A(T) \cdot T)}{g_A(T)}, \mathcal{H}_A)$ -Swap Regret and $(\frac{3g_A(T) \cdot T}{2} + \frac{f_B(g_B(T) \cdot T)}{g_B(T)}, \mathcal{H}_B)$ -Swap Regret. Thus, it has at most

$$\left(4T(\epsilon + g_A(T) + \frac{1}{2K\epsilon^2} + \frac{\beta(T, f'_A, f'_B)}{2\epsilon^2}) + \frac{f_A(g_A(T) \cdot T)}{g_A(T)} + \frac{3g_A(T) \cdot T}{2} + \frac{f_B(g_B(T) \cdot T)}{g_B(T)}, \mathcal{H}_A \cup \mathcal{H}_B \right)$$

2227 -Swap Regret.

2228 Note that we can select the agreement parameter ϵ here however we like in order to minimize the swap

2229 regret. In particular, we would like to pick ϵ to minimize the expression $\epsilon + \frac{1}{2K\epsilon^2} + \frac{\beta(T, f'_A, f'_B)}{2\epsilon^2} =$

2230 $\epsilon + \frac{\beta(T, f'_A, f'_B) + 1/K}{2\epsilon^2}$. By setting $\epsilon = (\frac{\beta(T, f'_A, f'_B) + 1/K}{2})^{\frac{1}{3}}$, we get that

$$\begin{aligned}
\epsilon + \frac{\beta(T, f'_A, f'_B) + 1/K}{2\epsilon^2} &= \\
&= \left(\frac{\beta(T, f'_A, f'_B) + 1/K}{2} \right)^{\frac{1}{3}} + \frac{\beta(T, f'_A, f'_B) + 1/K}{2 \left(\frac{\beta(T, f'_A, f'_B) + 1/K}{2} \right)^{\frac{2}{3}}} \\
&= 2 \left(\frac{\beta(T, f'_A, f'_B) + 1/K}{2} \right)^{\frac{1}{3}}
\end{aligned}$$

2231 Plugging this back into the swap regret expression, we get that, if k is an even round, \bar{y}_k has at most

$$\left(8T \left(\frac{\beta(T, f'_A, f'_B) + 1/K}{2} \right)^{\frac{1}{3}} + \frac{11}{2} T g_A(T) + \frac{f_A(g_A(T) \cdot T)}{g_A(T)} + \frac{f_B(g_B(T) \cdot T)}{g_B(T)}, \mathcal{H}_A \cup \mathcal{H}_B \right)$$

2232 -swap regret.

2233 In the case where k is an odd round, by a symmetric argument in which we define \bar{y}_k by combining
2234 level sets of \hat{y}_k to map to the closest value in $g_B(T)$, \bar{y}_k has

$$\left(8T \left(\frac{\beta(T, f'_A, f'_B) + 1/K}{2} \right)^{\frac{1}{3}} + \frac{11}{2} T g_B(T) + \frac{f_B(g_B(T) \cdot T)}{g_B(T)} + \frac{f_A(g_A(T) \cdot T)}{g_A(T)}, \mathcal{H}_A \cup \mathcal{H}_B \right)$$

2235 -swap regret.

2236 Thus, in all cases, the swap regret of \bar{y}_k with respect to $\mathcal{H}_A \cup \mathcal{H}_B$ is at most

$$\begin{aligned}
&8T \left(\frac{\beta(T, f'_A, f'_B) + 1/K}{2} \right)^{\frac{1}{3}} + \frac{11}{2} T g_B(T) + \frac{11}{2} T g_A(T) + \frac{f_B(g_B(T) \cdot T)}{g_B(T)} + \frac{f_A(g_A(T) \cdot T)}{g_A(T)} \\
&= 8T \left(\frac{\beta(T, f'_A, f'_B) + 1/K}{2} \right)^{\frac{1}{3}} + \frac{11}{2} T \beta(T, f_A, f_B)
\end{aligned}$$

2237 Note that \hat{y}_k is close in L_1 distance to \bar{y} , as we have only modified each entry by either at most $\frac{g_A(T)}{2}$
2238 or at most $\frac{g_B(T)}{2}$, depending if it was an even or odd round. Therefore, \hat{y}_k has at most

$$\left(\frac{T}{2} (g_A(T) + g_B(T)), 8T \left(\frac{\beta(T, f'_A, f'_B) + 1/K}{2} \right)^{\frac{1}{3}} + \frac{11}{2} T \beta(T, f_A, f_B), \mathcal{H}_A \cup \mathcal{H}_B \right)$$

2239 -distance to swap regret.

2240 □

2241 K.6 Proof of Theorem C.3

2242 *Proof of Theorem C.3.* By Theorem C.2, if Alice has $(f_A, g_A, \mathcal{H}_A)$ -conversation swap regret and
2243 Bob has $(f_B, g_B, \mathcal{H}_B)$ -conversation swap regret, there exists a round k of the protocol that has
2244 $(\frac{T}{2}(g_A(T) + g_B(T)), 8T(\frac{\beta(T, f'_A, f'_B) + 1/K}{2})^{\frac{1}{3}} + \frac{11}{2}T\beta(T, f_A, f_B), \mathcal{H}_A \cup \mathcal{H}_B)$ -distance to swap regret,
2245 where $\beta(T, f_A, f_B) = \frac{f_A(g_A(T) \cdot T)}{T g_A(T)} + \frac{f_B(g_B(T) \cdot T)}{T g_B(T)} + g_A(T) + g_B(T)$, $f'_A(x) = \sqrt{x \cdot f_A(x)}$ and
2246 $f'_B(x) = \sqrt{x \cdot f_B(x)}$. Then by the fact that \mathcal{H}_A and \mathcal{H}_B jointly satisfy the $w(\cdot)$ -weak learning
2247 condition with respect to \mathcal{H}_J and via Theorem B.3, instantiating $f^S = 8T(\frac{\beta(T, f'_A, f'_B) + 1/K}{2})^{\frac{1}{3}} +$
2248 $\frac{11}{2}T\beta(T, f_A, f_B)$ and $f^D = \frac{T}{2}(g_A(T) + g_B(T))$, we have that for the predictions $\hat{y}^{k,t}$ in round k :

$$\begin{aligned}
&\sum_{t=1}^T (\hat{y}^{k,t} - y^t)^2 - \min_{h_J \in \mathcal{H}_J} \sum_{t=1}^T (h_J(x^t) - y^t)^2 \\
&\leq 2Tw^{-1} \left(\frac{8T(\frac{\beta(T, f'_A, f'_B) + 1/K}{2})^{\frac{1}{3}} + \frac{11}{2}T\beta(T, f_A, f_B)}{T} \right) + 3\frac{T}{2}(g_A(T) + g_B(T)) \\
&= 2Tw^{-1} \left(8 \left(\frac{\beta(T, f'_A, f'_B) + 1/K}{2} \right)^{\frac{1}{3}} + \frac{11}{2}\beta(T, f_A, f_B) \right) + 3\frac{T}{2}(g_A(T) + g_B(T))
\end{aligned}$$

2249 By Theorem K.16, we can upper bound the increase in squared error from round i to round $i + 2$ by
 2250 $3T\beta(T, f'_A, f'_B)$. The maximum number of rounds between k and K is K . Therefore, we have that

$$\sum_{t=1}^T (\hat{y}^{K,t} - y^t)^2 \leq \sum_{t=1}^T (\hat{y}^{k,t} - y^t)^2 + 3TK\beta(T, f'_A, f'_B)$$

2251 Combining the above results, we have that

$$\begin{aligned} & \sum_{t=1}^T (\hat{y}^{K,t} - y^t)^2 - \min_{h_J \in \mathcal{H}_J} \sum_{t=1}^T (h_J(x^t) - y^t)^2 \leq \\ & 2Tw^{-1} \left(8 \left(\frac{\beta(T, f'_A, f'_B) + 1/K}{2} \right)^{\frac{1}{3}} + \frac{11}{2} \beta(T, f_A, f_B) \right) + 3 \frac{T}{2} (g_A(T) + g_B(T)) + 3TK\beta(T, f'_A, f'_B) \end{aligned}$$

2252

□

2253 K.7 Proof of Theorem C.5

2254 *Proof of Theorem C.5.* Let $\rho' = \frac{2g(T)\rho}{K}$. Let M be the algorithm given by the reduction in Theorem
 2255 C.4, given an online algorithm M_0 that achieves external regret with respect to \mathcal{H} bounded by $r(\tau)$
 2256 for any $\tau \in [T]$. In particular, Theorem C.4 guarantees that with probability $1 - \rho'$, M achieves
 2257 (f, \mathcal{H}) -swap regret for:

$$f(\tau) \leq m \cdot r\left(\frac{\tau}{m}\right) + \frac{3\tau}{m} + m + \max(8B, 2\sqrt{B}) \cdot m \cdot C_{\mathcal{H}} \cdot \sqrt{\tau \log\left(\frac{2mK}{g(T)\rho}\right)}$$

2258 By construction, on every odd round k , a separate copy $M_{k,i}$ is run for every subsequence on which
 2259 the previous prediction falls into bucket i . By a union bound, the probability that any one of the
 2260 copies fails is $\frac{K}{2} \cdot \frac{1}{g(T)} \cdot \rho' = \rho$. Then, since conversation swap regret measures the swap regret
 2261 conditioned on subsequences on which the previous prediction falls into bucket i (as parameterized
 2262 by g), with probability $1 - \rho$, Algorithm C.1 also satisfies (f, g, \mathcal{H}) -conversation swap regret.

2263

□

2264 K.8 Proof of Theorem C.6

2265 **Lemma K.18.** *If w is continuous and strictly convex, $w(0) = 0$ and $\lim_{T \rightarrow \infty} s(T) = 0$, then*
 2266 $\lim_{T \rightarrow \infty} w^{-1}(s(T)) = 0$.

2267 *Proof.* Note that as w is strictly monotone, w^{-1} is defined everywhere in the range of $(0, c)$, where
 2268 $c = \lim_{x \rightarrow \infty} \inf(w(x))$ and $c > 0$. As $w(0) = 0$, it must be the case that $w^{-1}(0) = 0$. Furthermore,
 2269 as w is continuous, w^{-1} must be continuous. Now, we can proceed to reason about w^{-1} :

$$\begin{aligned} \lim_{T \rightarrow \infty} w^{-1}(x(T)) &= w^{-1} \lim_{T \rightarrow \infty} (x(T)) && \text{(By the continuity of } w^{-1}) \\ &= f(0) && \text{(By the fact that } \lim_{T \rightarrow \infty} s(T) = 0) \\ &= 0 && \text{(By the fact that } w^{-1}(0) = 0) \end{aligned}$$

2270

□

2271 *Proof of Theorem C.6.* Let $\rho' = \rho/2$. We set our parameters to be sublinear in T . Specifically, set
 2272 $m = T^{1/4}$ and $1/g_A(T) = 1/g_B(T) = T^{\alpha_g}$ for some constant $\alpha_g \in (0, 1)$. By Theorem C.5, there

2273 is an algorithm that achieves, with probability $1 - \rho'$, $(f_A, g_A, \mathcal{H}_A)$ -conversation swap regret for, for
 2274 any $\tau \in [T]$:

$$\begin{aligned}
 f_A(\tau) &\leq m \cdot r_A \left(\frac{\tau}{m} \right) + \frac{3\tau}{m} + m + \max(8B, 2\sqrt{B}) \cdot m \cdot C_{\mathcal{H}} \cdot \sqrt{\tau \log \left(\frac{2mK}{g_A(T)\rho'} \right)} \\
 &\quad \text{(by Theorem C.5)} \\
 &\leq m \cdot r_A \left(\frac{T}{m} \right) + \frac{3T}{m} + m + \max(8B, 2\sqrt{B}) \cdot m \cdot C_{\mathcal{H}} \cdot \sqrt{T \log \left(\frac{4mK}{g_A(T)\rho} \right)} \\
 &\leq T^{1/4} \cdot \tilde{O}((T^{3/4})^{\alpha_A}) + 3T^{3/4} + T^{1/4} + \max(8B, 2\sqrt{B}) \cdot C_{\mathcal{H}} \cdot T^{3/4} \sqrt{\log \left(\frac{4KT^{1/4+\alpha_g}}{\rho} \right)} \\
 &\leq \tilde{O} \left(T^{\alpha_1} \sqrt{\log \left(\frac{K}{\rho} \right)} \right)
 \end{aligned}$$

2275 for $\alpha_1 = \max\{1/4 + 3/4 \cdot \alpha_A, 3/4\} \in (0, 1)$. Since Bob's expression is symmetric, Theorem C.5
 2276 similarly implies that there is an algorithm that achieves, with probability $1 - \rho'$, $(f_B, g_B, \mathcal{H}_B)$ -
 2277 conversation swap regret for:

$$f_B(\tau) \leq \tilde{O} \left(T^{\alpha_2} \sqrt{\log \left(\frac{K}{\rho} \right)} \right)$$

2278 for $\alpha_2 = \max\{1/4 + 3/4 \cdot \alpha_B, 3/4\} \in (0, 1)$. Thus, by a union bound, with probability $1 - 2\rho' = 1 - \rho$,
 2279 Alice has $(f_A, g_A, \mathcal{H}_A)$ -conversation swap regret and Bob has $(f_B, g_B, \mathcal{H}_B)$ -conversation swap
 2280 regret.

2281 Now, by Theorem C.3, the transcript on the last round has regret bounded by:

$$\begin{aligned}
 &\sum_{t=1}^T (\hat{y}^{t,K} - y^t)^2 - \min_{h_J \in \mathcal{H}_J} \sum_{t=1}^T (h_J(x^t) - y^t)^2 \\
 &\leq 2Tw^{-1} \left(8 \left(\frac{\beta(T, f'_A, f'_B) + 1/K}{2} \right)^{\frac{1}{3}} + \frac{11}{2} \beta(T, f_A, f_B) \right) + 3 \frac{T}{2} (g_A(T) + g_B(T)) + 3TK\beta(T, f'_A, f'_B)
 \end{aligned}$$

2282 where for $\tau \leq T$:

$$\begin{aligned}
 f'_A(\tau) &= \sqrt{\tau \cdot f_A(\tau)} \leq \sqrt{T \cdot \tilde{O} \left(T^{\alpha_1} \sqrt{\log \left(\frac{K}{\rho} \right)} \right)} \leq \tilde{O} \left(T^{(1+\alpha_1)/2} \log^{1/4} \left(\frac{K}{\rho} \right) \right), \\
 f'_B(\tau) &= \sqrt{\tau \cdot f_B(\tau)} \leq \sqrt{T \cdot \tilde{O} \left(T^{\alpha_2} \sqrt{\log \left(\frac{K}{\rho} \right)} \right)} \leq \tilde{O} \left(T^{(1+\alpha_2)/2} \log^{1/4} \left(\frac{K}{\rho} \right) \right)
 \end{aligned}$$

2283 and thus:

$$\begin{aligned}
 \beta(T, f_A, f_B) &= \frac{f_A(g_A(T)T)}{Tg_A(T)} + \frac{f_B(g_B(T)T)}{Tg_B(T)} + g_A(T) + g_B(T) \\
 &\leq \tilde{O} \left((T^{\alpha_1+\alpha_g-1} + T^{\alpha_2+\alpha_g-1}) \sqrt{\log \left(\frac{K}{\rho} \right)} + T^{-\alpha_g} \right), \\
 \beta(T, f'_A, f'_B) &= \frac{f'_A(g_A(T)T)}{Tg_A(T)} + \frac{f'_B(g_B(T)T)}{Tg_B(T)} + g_A(T) + g_B(T) \\
 &\leq \tilde{O} \left((T^{\alpha_1/2+\alpha_g-1/2} + T^{\alpha_2/2+\alpha_g-1/2}) \log^{1/4} \left(\frac{K}{\rho} \right) + T^{-\alpha_g} \right)
 \end{aligned}$$

Suppose $\alpha_g < \min\{1/2 - \alpha_1/2, 1/2 - \alpha_2/2\}$. Then:

$$T^{\alpha_1+\alpha_g-1}, T^{\alpha_2+\alpha_g-1}, T^{\alpha_1/2+\alpha_g-1/2}, T^{\alpha_2/2+\alpha_g-1/2} \leq T^{-\alpha}$$

2284 for some constant $\alpha \in (0, 1)$. Hence, plugging in to the expression above, we have that:

$$\begin{aligned}
& \sum_{t=1}^T (\hat{y}^{t,K} - y^t)^2 - \min_{h_J \in \mathcal{H}_J} \sum_{t=1}^T (h_J(x^t) - y^t)^2 \\
& \leq 2Tw^{-1} \left(\tilde{O} \left(T^{-\alpha} \sqrt{\log \left(\frac{K}{\rho} \right)} + T^{-\alpha_g} + \frac{1}{K} \right)^{1/3} + \tilde{O} \left(T^{-\alpha} \sqrt{\log \left(\frac{K}{\rho} \right)} + T^{-\alpha_g} \right) \right) \\
& \quad + O(T^{1-\alpha_g}) + T \cdot \tilde{O} \left(T^{-\alpha} K \log^{1/4} \left(\frac{K}{\rho} \right) + KT^{-\alpha_g} \right) \\
& \leq 2Tw^{-1} \left(\tilde{O} \left(T^{-\alpha/3} \log^{1/6} \left(\frac{K}{\rho} \right) + T^{-\alpha_g/3} + \frac{1}{K^{1/3}} \right) + \tilde{O} \left(T^{-\alpha} \sqrt{\log \left(\frac{K}{\rho} \right)} + T^{-\alpha_g} \right) \right) \\
& \quad + O(T^{1-\alpha_g}) + T \cdot \tilde{O} \left(T^{-\alpha} K \log^{1/4} \left(\frac{K}{\rho} \right) + KT^{-\alpha_g} \right) \\
& \hspace{15em} \text{(by concavity of the cube root function)} \\
& \leq 2Tw^{-1} \left(\tilde{O} \left(T^{-\alpha/3} \sqrt{\log \left(\frac{K}{\rho} \right)} + T^{-\alpha_g/3} + \frac{1}{K^{1/3}} \right) \right) + O(T^{1-\alpha_g}) + \tilde{O} \left(KT^{1-\alpha} \log^{1/4} \left(\frac{K}{\rho} \right) + KT^{1-\alpha_g} \right) \\
& \leq 2Tw^{-1} \left(\tilde{O} \left(T^{-\alpha'} \sqrt{\log \left(\frac{K}{\rho} \right)} + \frac{1}{K^{1/3}} \right) \right) + O(T^{1-\alpha_g}) + \tilde{O} \left(KT^{1-\alpha''} \log^{1/4} \left(\frac{K}{\rho} \right) \right)
\end{aligned}$$

2285 where $\alpha' = \min\{\alpha/3, \alpha_g/3\} \in (0, 1)$ and $\alpha'' = \min\{\alpha, \alpha_g\} \in (0, 1)$. This proves the first part of
2286 the theorem.

2287 To argue the second part, suppose $K = \omega(1)$ and $K = O(T^{\alpha''-\varepsilon})$ for $\varepsilon > 0$. Then:

$$\begin{aligned}
& \sum_{t=1}^T (\hat{y}^{t,K} - y^t)^2 - \min_{h_J \in \mathcal{H}_J} \sum_{t=1}^T (h_J(x^t) - y^t)^2 \\
& \leq 2Tw^{-1} \left(\tilde{O} \left(T^{-\alpha'} \sqrt{\log \left(\frac{K}{\rho} \right)} + \frac{1}{K^{1/3}} \right) \right) + O(T^{1-\alpha_g}) + \tilde{O} \left(KT^{1-\alpha''} \log^{1/4} \left(\frac{K}{\rho} \right) \right) \\
& = 2Tw^{-1} \left(\tilde{O} \left(T^{-\alpha'} \sqrt{\log \left(\frac{T}{\rho} \right)} \right) + o(1) \right) + O(T^{1-\alpha_g}) + \tilde{O} \left(T^{1-\varepsilon} \log^{1/4} \left(\frac{T}{\rho} \right) \right)
\end{aligned}$$

Now, observe that any function $\tilde{O} \left(T^{-\alpha'} \sqrt{\log \left(\frac{T}{\rho} \right)} \right) + o(1) \rightarrow 0$ as $T \rightarrow \infty$. Thus, by Lemma

K.18, $w^{-1} \left(\tilde{O} \left(T^{-\alpha'} \sqrt{\log \left(\frac{T}{\rho} \right)} \right) + o(1) \right) \rightarrow 0$ as $T \rightarrow \infty$. In particular, this implies that

$$Tw^{-1} \left(\tilde{O} \left(T^{-\alpha'} \sqrt{\log \left(\frac{T}{\rho} \right)} + o(1) \right) \right) = o(T)$$

2288 i.e. is sublinear in T . Notice that since w is strictly increasing, w^{-1} exists for sufficiently large T
2289 (larger than a constant). Therefore, for sufficiently large T , the regret is bounded by:

$$\sum_{t=1}^T (\hat{y}^{t,K} - y^t)^2 - \min_{h_J \in \mathcal{H}_J} \sum_{t=1}^T (h_J(x^t) - y^t)^2 \leq o(T) + O(T^{1-\alpha_g}) + \tilde{O} \left(T^{1-\varepsilon} \log^{1/4} \left(\frac{T}{\rho} \right) \right)$$

2290 which completes the proof.

2291

□

2292 K.9 Proof of Theorem C.9

2293 **Theorem K.19.** [Rakhlin et al., 2015] Let $\mathcal{X} = \{x \in \mathbb{R}^d : \|x\|_2 \leq 1\}$ and $\mathcal{H} = \{x \mapsto \langle \theta, x \rangle : \|\theta\|_2 \leq C\}$ be the set of all linear functions with bounded norm. \mathcal{H} has finite sequential fat-shattering
2294 dimension.
2295

2296 **Corollary K.20.** Let $\mathcal{X} = \{x \in \mathbb{R}^d : \|x\|_2 \leq 1\}$ and $\mathcal{H} = \{x \mapsto \langle \theta, x \rangle : \|\theta\|_2 \leq C\}$ be the set of all
 2297 linear functions with bounded norm. Fix any bucketing function g . There exists an online algorithm
 2298 that, with probability $1 - \rho$, achieves $(\tilde{O}\left(\max(C^2, C)d \log\left(\frac{K}{g(T)\rho}\right) T^{3/4}\right), g, \mathcal{H})$ -conversation
 2299 swap regret.

2300 *Proof.* We have that $\langle \theta, x \rangle^2 \leq \|\theta\|_2^2 \|x\|_2^2 \leq C^2$. Therefore, by setting $m = T^{\frac{1}{4}}$ and instantiating
 2301 Theorem C.5 with the external regret algorithm of Theorem C.8, we have that, with probability $1 - \rho$,
 2302 Algorithm C.1 achieves (f, g, \mathcal{H}) -conversation swap regret for:

$$\begin{aligned}
 & f(|T(k-1), i|) \\
 & \leq T^{\frac{1}{4}} r \left(\frac{|T(k-1, i)|}{T^{\frac{1}{4}}} \right) + \frac{3|T(k-1, i)|}{T^{\frac{1}{4}}} + T^{\frac{1}{4}} + \max(8C^2, 2C)T^{\frac{1}{4}}C_{\mathcal{H}} \sqrt{|T(k-1, i)| \log\left(\frac{2KT^{\frac{1}{4}}}{g(T)\rho}\right)} \\
 & \hspace{15em} \text{(by Theorem C.5 and our setting of } m) \\
 & \leq T^{\frac{1}{4}} \left(2d \ln \left(\frac{|T(k-1, i)|}{T^{\frac{1}{4}}} + 1 \right) + C^2 \right) + \frac{3|T(k-1, i)|}{T^{\frac{1}{4}}} + T^{\frac{1}{4}} \\
 & \quad + \max(8C^2, 2C)T^{\frac{1}{4}}C_{\mathcal{H}} \sqrt{|T(k-1, i)| \log\left(\frac{2KT^{\frac{1}{4}}}{g(T)\rho}\right)} \hspace{5em} \text{(by Theorem C.8)} \\
 & \leq \tilde{O} \left(\max(C^2, C)d \log\left(\frac{K}{g(T)\rho}\right) T^{3/4} \right) \\
 & = \tilde{O} \left(\max(C^2, C)d \log\left(\frac{K}{g(T)\rho}\right) T^{3/4} \right)
 \end{aligned}$$

2303

□

2304 *Proof of Theorem C.9.* Let $\rho' = \rho/2$. By Corollary K.20, for any bucketing function g_A , there is an
 2305 algorithm that achieves, with probability $1 - \rho'$, $(f_A, g_A, \mathcal{H}_A)$ -conversation swap regret for:

$$f_A(|T_B(k-1, i)|) \leq \tilde{O} \left(\max(C^2, C)d \log\left(\frac{K}{g_A(T)\rho}\right) T^{3/4} \right)$$

2306 Likewise, for any bucketing function g_B , there is an algorithm that achieves, with probability $1 - \rho'$,
 2307 $(f_B, g_B, \mathcal{H}_B)$ -conversation swap regret for:

$$f_B(|T_A(k-1, i)|) \leq \tilde{O} \left(\max(C^2, C)d \log\left(\frac{K}{g_B(T)\rho}\right) T^{3/4} \right)$$

2308 Thus by a union bound, with probability $1 - 2\rho' = 1 - \rho$, Alice has $(f_A, g_A, \mathcal{H}_A)$ -conversation swap
 2309 regret and Bob has $(f_B, g_B, \mathcal{H}_B)$ -conversation swap regret. Let $g_A = g_B = T^{-\frac{1}{8}}$.

2310 Now, by Theorem B.6, \mathcal{H}_A and \mathcal{H}_B jointly satisfy the $w(\cdot)$ -weak learning condition with respect to
 2311 \mathcal{H}_J for $w(\gamma) = \frac{\gamma^2}{16C^2}$. In particular, we have that $w^{-1}(\gamma) = 4C\gamma^{1/2}$ for $\gamma \leq \frac{1}{16C^2}$. Therefore, by

2312 Theorem C.3, we have that the transcript $\pi^{1:T,K}$ at the last round satisfies:

$$\begin{aligned}
& \sum_{t=1}^T (\hat{y}^{t,K} - y^t)^2 - \min_{h_J \in \mathcal{H}_J} \sum_{t=1}^T (h_J(x^t) - y^t)^2 \\
& \leq 2Tw^{-1} \left(8 \left(\frac{\beta(T, f'_A, f'_B) + 1/K}{2} \right)^{\frac{1}{3}} + \frac{11}{2} \beta(T, f_A, f_B) \right) + 3 \frac{T}{2} (g_A(T) + g_B(T)) + 3TK\beta(T, f'_A, f'_B) \\
& = \tilde{O} \left(T \left((\beta(T, f'_A, f'_B) + 1/K)^{\frac{1}{3}} + \beta(T, f_A, f_B) \right)^{\frac{1}{2}} + T(g_A(T) + g_B(T)) + TK\beta(T, f'_A, f'_B) \right) \\
& \quad \text{(By Theorem B.6)} \\
& = \tilde{O} \left(T(\beta(T, f'_A, f'_B) + 1/K)^{\frac{1}{6}} + T\beta^{\frac{1}{2}}(T, f) + Tg_A(T) + Tg_B(T) + TK\beta(T, f'_A, f'_B) \right) \\
& = \tilde{O} \left(T\beta^{\frac{1}{6}}(T, f') + TK^{-\frac{1}{6}} + T\beta^{\frac{1}{2}}(T, f) + Tg_A(T) + Tg_B(T) + TK\beta(T, f'_A, f'_B) \right) \\
& = \tilde{O} \left(T\beta^{\frac{1}{6}}(T, f') + TK^{-\frac{1}{6}} + T\beta^{\frac{1}{2}}(T, f) + T^{\frac{7}{8}} + TK\beta(T, f'_A, f'_B) \right) \\
& \quad \text{(Instantiating } g_A \text{ and } g_B)
\end{aligned}$$

2313 Here, $\beta(T, f_A, f_B) = \frac{f(g_A(T) \cdot T)}{Tg_A(T)} + \frac{f(g_B(T) \cdot T)}{Tg_B(T)} + g_A(T) + g_B(T)$, and $f'(x) = \sqrt{x \cdot f(x)}$.

2314 Plugging in f_A and f_B , we have that:

$$\begin{aligned}
\beta(T, f_A, f_B) & \leq \tilde{O} \left(\frac{d \log \left(\frac{K}{g_A(T)\rho} \right)}{g_A(T)T^{1/4}} + \frac{d \log \left(\frac{K}{g_B(T)\rho} \right)}{g_B(T)T^{1/4}} + g_A(T) + g_B(T) \right) \\
& = \tilde{O} \left(d \log \left(\frac{KT^{\frac{1}{8}}}{\rho} \right) T^{-1/8} + T^{-1/8} \right) \\
& = \tilde{O} \left(d \log \left(\frac{KT^{\frac{1}{8}}}{\rho} \right) T^{-1/8} \right)
\end{aligned}$$

2315 Moreover:

$$\begin{aligned}
\beta(T, f'_A, f'_B) & \leq \tilde{O} \left(\frac{\sqrt{g_A(T) \cdot T \cdot f(g_A(T) \cdot T)}}{Tg_A(T)} + \frac{\sqrt{g_B(T) \cdot T \cdot f(g_B(T) \cdot T)}}{Tg_B(T)} + g_A(T) + g_B(T) \right) \\
& = \tilde{O} \left(\sqrt{\frac{f(g_A(T) \cdot T)}{Tg_A(T)}} + \sqrt{\frac{f(g_B(T) \cdot T)}{Tg_B(T)}} + g_A(T) + g_B(T) \right) \\
& = \tilde{O} \left(\sqrt{\frac{\max(C^2, C) d \log \left(\frac{K}{g_A(T)\rho} \right) T^{3/4}}{Tg_A(T)}} + \sqrt{\frac{\max(C^2, C) d \log \left(\frac{K}{g_B(T)\rho} \right) T^{3/4}}{Tg_B(T)}} + g_A(T) + g_B(T) \right) \\
& = \tilde{O} \left(T^{-\frac{1}{8}} g_A^{-\frac{1}{2}}(T) \sqrt{\max(C^2, C) d \log \left(\frac{K}{g_A(T)\rho} \right)} \right. \\
& \quad \left. + T^{-\frac{1}{8}} g_B^{-\frac{1}{2}}(T) \sqrt{\max(C^2, C) d \log \left(\frac{K}{g_B(T)\rho} \right)} + g_A(T) + g_B(T) \right) \\
& = \tilde{O} \left(T^{-\frac{1}{8}} T^{1/16} \sqrt{\max(C^2, C) d \log \left(\frac{KT^{1/8}}{\rho} \right)} + T^{-1/8} \right) \\
& = \tilde{O} \left(T^{-1/16} \sqrt{\max(C^2, C) d \log \left(\frac{KT^{1/8}}{\rho} \right)} \right)
\end{aligned}$$

2316 Plugging these expressions into the regret bound of the final round, we get that:

$$\begin{aligned}
& \tilde{O}\left(T\beta^{\frac{1}{6}}(T, f'_A, f'_B) + TK^{-\frac{1}{6}} + T\beta^{\frac{1}{2}}(T, f_A, f_B) + T^{\frac{7}{8}} + TK\beta(T, f'_A, f'_B)\right) \\
&= \tilde{O}\left(T(T^{-1/8}\sqrt{\max(C^2, C)d\log\left(\frac{KT^{1/8}}{\rho}\right)})^{1/6} + TK^{-\frac{1}{6}} + T(\max(C^2, C)d\log\left(\frac{KT^{1/8}}{\rho}\right)T^{-1/16})^{1/2}\right. \\
&\quad \left.+ T^{\frac{7}{8}} + KT^{\frac{7}{8}}\sqrt{\max(C^2, C)d\log\left(\frac{KT^{1/8}}{\rho}\right)}\right) \\
&= \tilde{O}\left(T^{47/48}(\max(C^2, C)d\log\left(\frac{KT^{1/8}}{\rho}\right))^{1/12} + TK^{-\frac{1}{6}} + T^{31/32}(\max(C^2, C)d\log\left(\frac{KT^{1/8}}{\rho}\right))^{1/2}\right. \\
&\quad \left.+ T^{\frac{7}{8}} + KT^{\frac{7}{8}}\sqrt{\max(C^2, C)d\log\left(\frac{KT^{1/8}}{\rho}\right)}\right) \\
&= \tilde{O}\left(T^{47/48}\sqrt{\max(C^2, C)d\log\left(\frac{KT^{1/8}}{\rho}\right)} + TK^{-\frac{1}{6}} + KT^{\frac{7}{8}}\sqrt{\max(C^2, C)d\log\left(\frac{KT^{1/8}}{\rho}\right)}\right)
\end{aligned}$$

2317

□

2318 L Proofs of Lower Bounds from Appendix D

2319 *Proof of Theorem D.1.* We adapt the construction from the proof of Theorem B.7. Define a joint
2320 distribution \mathcal{D} over $\mathcal{X}_A \times \mathcal{X}_B \times \mathcal{Y}$ where $\mathcal{X}_A, \mathcal{X}_B \subseteq \mathbb{R}$ as follows:

$$x_A = \xi_A, x_B = x_A + \xi_B = \xi_A + \xi_B, \text{ and } y = \xi_B = x_B - x_A,$$

2321 where ξ_A, ξ_B are independent random variables uniformly distributed in $\{-1, +1\}$.

2322 We consider $\mathcal{H}_A = \mathcal{H}_B = \{x \mapsto wx + b : |w| \leq 1, |b| \leq 1\}$ and $\mathcal{H}_J = \{(x_A, x_B) \mapsto w_A x_A +$
2323 $w_B x_B + b : |w_A| \leq 1, |w_B| \leq 1, |b| \leq 1\}$ to be the classes of bounded linear functions. Then we
2324 have the following:

2325 *Optimal Linear Predictor for Alice (h_A^*):* Since $y = \xi_B$ is independent of $x_A = \xi_A$, and $\mathbb{E}[y] =$
2326 $\mathbb{E}[\xi_B] = 0$, the optimal linear predictor is the constant predictor $h_A^*(x_A) = \mathbb{E}[y] = 0$. Its expected
2327 squared error is $\mathbb{E}[(0 - y)^2] = \mathbb{E}[\xi_B^2] = 1$.

2328 *Optimal Linear Predictor for Bob (h_B^*):* We seek $h_B^*(x_B) = w_B x_B + c_B$. Since $\mathbb{E}[y] = 0$ and
2329 $\mathbb{E}[x_B] = \mathbb{E}[x_A + \xi_B] = \mathbb{E}[x_A] + \mathbb{E}[\xi_B] = 0$, $c_B = 0$. The optimal $w_B = \frac{\mathbb{E}[x_B y]}{\mathbb{E}[x_B^2]}$. We have,

$$\begin{aligned}
\mathbb{E}[x_B y] &= \mathbb{E}[(\xi_A + \xi_B)\xi_B] = \mathbb{E}[\xi_A \xi_B] + \mathbb{E}[\xi_B^2] = 1. \\
\mathbb{E}[x_B^2] &= \mathbb{E}[(\xi_A + \xi_B)^2] = \mathbb{E}[\xi_A^2 + 2\xi_A \xi_B + \xi_B^2] = 2.
\end{aligned}$$

2330 Thus, $w_B = \frac{1}{2}$ and $h_B^*(x_B) = x_B/2$. Its expected squared error is $\mathbb{E}[(h_B^*(x_B) - y)^2] = \mathbb{E}[(x_B/2 -$
2331 $y)^2] = \mathbb{E}[(\xi_B/2)^2] = 1/4$.

2332 *Optimal Linear Predictor for Joint Features (h_J^*):* The conditional expectation $\mathbb{E}[y|x_A, x_B]$ is the
2333 optimal predictor overall. Here, $y = \xi_B = x_B - x_A$. Since this relationship is linear, the optimal
2334 linear predictor is $h_J^*(x) = x_B - x_A$. Its expected squared error is $\mathbb{E}[(h_J^*(x) - y)^2] = \mathbb{E}[(y - y)^2] = 0$.

2335 We have $h_A^*(x_A) = 0$ and $h_B^*(x_B) = x_B/2$. Any predictor $f(h_A^*(x_A), h_B^*(x_B))$ can only depend
2336 on x_B since $h_A^*(x_A) = 0$ is constant. The best predictor for y that is a function of x_B is in this case
2337 exactly the optimal linear predictor $h_B^*(x_B) = x_B/2$, which achieves an error of $1/4$. Thus, the
2338 minimum error achievable using only h_A^* and h_B^* by any function f is:

$$\mathbb{E}_{\mathcal{D}}[(f(h_A^*(x_A), h_B^*(x_B)) - y)^2] = \mathbb{E}_{\mathcal{D}}[(h_B^*(x_B) - y)^2] \geq 1/4 > 0 = \mathbb{E}_{\mathcal{D}}[(h_J^*(x) - y)^2].$$

2339

□

2340 *Proof of Theorem D.2.* Consider a triple $(\mathcal{H}_A, \mathcal{H}_B, \mathcal{H}_J)$ that fails the $w(\cdot)$ -weak learning condition
 2341 for any strictly increasing w . This implies there exists a distribution \mathcal{D} such that for some $\gamma > 0$:

$$\min_{c \in \mathbb{R}} \mathbb{E}_{\mathcal{D}}[(c - y)^2] - \min_{h_J \in \mathcal{H}_J} \mathbb{E}_{\mathcal{D}}[(h_J(x) - y)^2] \geq \gamma,$$

2342 but for all $h_A \in \mathcal{H}_A$ and $h_B \in \mathcal{H}_B$:

$$\min_{c \in \mathbb{R}} \mathbb{E}_{\mathcal{D}}[(c - y)^2] - \mathbb{E}_{\mathcal{D}}[(h_A(x_A) - y)^2] < w(\gamma)$$

2343

$$\min_{c \in \mathbb{R}} \mathbb{E}_{\mathcal{D}}[(c - y)^2] - \mathbb{E}_{\mathcal{D}}[(h_B(x_B) - y)^2] < w(\gamma).$$

2344 Since this must hold for any strictly increasing w (and $w(0) = 0$), it must be that the improvement
 2345 over the constant predictor for both \mathcal{H}_A and \mathcal{H}_B is zero. That is, $\min_{h_A \in \mathcal{H}_A} \mathbb{E}_{\mathcal{D}}[(h_A(x_A) - y)^2] =$
 2346 $\min_{h_A \in \mathcal{H}_A} \mathbb{E}_{\mathcal{D}}[(h_B(x_B) - y)^2] = \min_{c \in \mathbb{R}} \mathbb{E}_{\mathcal{D}}[(c - y)^2]$. Let $c^* = \arg \min_{c \in \mathbb{R}} \mathbb{E}_{\mathcal{D}}[(c - y)^2]$ be the
 2347 optimal constant predictor.

2348 Now consider the sequence of examples $(x_A^t, x_B^t, y^t)_{t=1}^T$ be drawn i.i.d. from the distribution \mathcal{D} and
 2349 the constant prediction sequence $\hat{y}^t = c^*$ for all $t = 1, \dots, T$. Since $\hat{y}^t = c^*$ for all t , the only
 2350 relevant level set is $v = c^*$, swap regret with respect to \mathcal{H}_A reduces to:

$$\frac{1}{T} \sum_{t=1}^T (\hat{y}^t - y^t)^2 - \min_{h_A \in \mathcal{H}_A} \frac{1}{T} \sum_{t=1}^T (h_A(x_A^t) - y^t)^2.$$

2351 As $T \rightarrow \infty$, by the law of large numbers, this reduces to

$$\mathbb{E}_{\mathcal{D}}[(c^* - y)^2] - \min_{h_A \in \mathcal{H}_A} \mathbb{E}_{\mathcal{D}}[(h_A(x_A) - y)^2] = 0.$$

2352 By the same argument, we get that the swap regret with respect to \mathcal{H}_B is also 0. However, the external
 2353 regret (as $T \rightarrow \infty$) with respect to \mathcal{H}_J ,

$$\mathbb{E}_{\mathcal{D}}[(c^* - y)^2] - \min_{h_J \in \mathcal{H}_J} \mathbb{E}_{\mathcal{D}}[(h_J(x) - y)^2] \geq \gamma > 0.$$

2354 Here the inequality follows from our assumption. This implies that the external regret with respect to
 2355 \mathcal{H}_J is positive, while swap regret with respect to both \mathcal{H}_A and \mathcal{H}_B is 0. \square

2356 *Proof of Theorem D.4.* In order to prove that \mathcal{H}_A and \mathcal{H}_B satisfy weak-learnability, let us assume
 2357 that for some distribution \mathcal{D} and $\gamma \in [0, 1]$

$$\min_{c \in \mathbb{R}} \mathbb{E}[(c - y)^2] - \min_{h_J \in \mathcal{H}_J} \mathbb{E}[(h_J(x) - y)^2] \geq \gamma.$$

2358 Now we will show that, either

$$\min_{c \in \mathbb{R}} \mathbb{E}[(c - y)^2] - \min_{h_A \in \mathcal{H}_A} \mathbb{E}[(h_A(x_A) - y)^2] \geq \gamma/2,$$

2359 or

$$\min_{c \in \mathbb{R}} \mathbb{E}[(c - y)^2] - \min_{h_B \in \mathcal{H}_B} \mathbb{E}[(h_B(x_B) - y)^2] \geq \gamma/2.$$

2360 Since \mathcal{H}_A and \mathcal{H}_B satisfy information substitutes with respect to \mathcal{H}_J , from the statement in Theo-
 2361 rem D.3, we have

$$\min_{h_A \in \mathcal{H}_A} \mathbb{E}[(h_A(x_A) - y)^2] + \min_{h_B \in \mathcal{H}_B} \mathbb{E}[(h_B(x_B) - y)^2] \leq \min_{c \in \mathbb{R}} \mathbb{E}[(c - y)^2] + \min_{h_J \in \mathcal{H}_J} \mathbb{E}[(h_J(x) - y)^2].$$

2362 Substituting the assumption on the joint feature improving over the constant function, we get

$$2 \min_{c \in \mathbb{R}} \mathbb{E}[(y - c)^2] - \min_{h_A \in \mathcal{H}_A} \mathbb{E}[(h_A(x_A) - y)^2] - \min_{h_B \in \mathcal{H}_B} \mathbb{E}[(h_B(x_B) - y)^2] \geq \gamma.$$

2363 This implies that either $\min_{c \in \mathbb{R}} \mathbb{E}[(y - c)^2] - \min_{h_A \in \mathcal{H}_A} \mathbb{E}[(h_A(x_A) - y)^2]$ or $\min_{c \in \mathbb{R}} \mathbb{E}[(y - c)^2] -$
 2364 $\min_{h_B \in \mathcal{H}_B} \mathbb{E}[(h_B(x_B) - y)^2]$ must be $\geq \gamma/2$. This gives us the desired weak-learning condition. \square

2365 *Proof of Theorem D.5.* Consider the class of bounded linear function over $\mathcal{X}_A = \mathcal{X}_B = [-1, 1]$
 2366 as defined in the proof of Theorem B.8. Suppose these classes satisfy the information substitutes
 2367 condition, then by Theorem D.4, we know that they must satisfy $w(\cdot)$ -weak learnability for $w(\gamma) =$
 2368 $\gamma/2$. However, from Theorem B.8, we know that these function classes can not satisfy $w(\cdot)$ -weak
 2369 learnability for $w(\gamma) = \gamma/2$ giving us a contradiction. Therefore, these classes could not have
 2370 satisfied the information substitutes condition. \square

2371 *Proof of Theorem D.6.* Consider the joint distribution \mathcal{D} over $\mathcal{X}_A \times \mathcal{X}_B \times \mathcal{Y}$ to be as follows:

$$x_A \sim_{\text{i.i.d.}} \{0, 1\}, x_B \sim_{\text{i.i.d.}} \{0, 1\}, y = x_A x_B.$$

2372 Let the class of functions be bounded linear functions which satisfy our weak-learning condition.

2373 Observe that the the best linear predictor in \mathcal{H}_A is $h_A^*(x_A) = \mathbb{E}[y|x_A] = x_A/2$ and the best linear
2374 predictor in \mathcal{H}_B is $h_B^*(x_B) = \mathbb{E}[y|x_B] = x_B/2$. Observe that,

$$\mathbb{E}[(h_A^*(x_A) - y)^2] = \mathbb{E}[(x_A/2 - x_A x_B)^2] = 1/8 = \mathbb{E}[(x_B/2 - x_A x_B)^2] = \mathbb{E}[(h_B^*(x_B) - y)^2].$$

2375 Now consider the sequence of examples $(x_A^t, x_B^t, y^t)_{t=1}^T$ be drawn i.i.d. from the distribution \mathcal{D} and
2376 the prediction sequence $\hat{y}^t = x_A^t/2$ for all $t = 1, \dots, T$. Observe that the external regret with respect
2377 to \mathcal{H}_A as $T \rightarrow \infty$ is

$$\mathbb{E}_{\mathcal{D}}[(x_A/2 - y)^2] - \min_{h_A \in \mathcal{H}_A} \mathbb{E}_{\mathcal{D}}[(h_A(x_A) - y)^2] = \mathbb{E}_{\mathcal{D}}[(x_A/2 - y)^2] - \mathbb{E}_{\mathcal{D}}[(x_A/2 - y)^2] = 0.$$

2378 Similarly the external regret with respect to \mathcal{H}_B as $T \rightarrow \infty$ is

$$\mathbb{E}_{\mathcal{D}}[(x_A/2 - y)^2] - \min_{h_B \in \mathcal{H}_B} \mathbb{E}_{\mathcal{D}}[(h_B(x_B) - y)^2] = \frac{1}{8} - \mathbb{E}[(h_B^*(x_B) - y)^2] = \frac{1}{8} - \frac{1}{8} = 0.$$

2379 Therefore the sequence of predictions has no external regret with respect to \mathcal{H}_A and \mathcal{H}_B .

2380 However, the best linear predictor defined on both \mathcal{X}_A and \mathcal{X}_B is $h_J^*(x) = (x_A + x_B)/2 - 1/4$. This
2381 has expected error over \mathcal{D} , $\mathbb{E}_{\mathcal{D}}[(h_J^*(x) - y)^2] = 1/16$. Thus, as $t \rightarrow \infty$, the predictions have external
2382 regret,

$$\mathbb{E}_{\mathcal{D}}[(x_A/2 - y)^2] - \mathbb{E}[(h_J^*(x) - y)^2] = \frac{1}{8} - \frac{1}{16} = \frac{1}{16} > 0.$$

2383 □

2384 M Proofs from Section E

2385 **Lemma M.1.** Let \mathcal{H} be a class of real-valued functions $h : \mathcal{X} \rightarrow \mathbb{R}$. Let $y : \mathcal{X} \rightarrow \mathcal{Y}$ be a fixed
2386 labeling function, and fix a label $v \in \mathcal{Y}$. Let \mathcal{H}_A^* be defined such that for each $h \in \mathcal{H}_A$, the
2387 corresponding function $h^* \in \mathcal{H}_A^*$ is given by:

$$h^*(x) = h(x) \cdot \mathbf{1}[y(x) = v].$$

2388 Then,

$$C_{\mathcal{H}_A^*}^\epsilon \leq C_{\mathcal{H}_A}^\epsilon$$

2389 In other words, for any scale ϵ , the fat-shattering dimension of \mathcal{H}_A^* is at most the fat-shattering
2390 dimension of \mathcal{H}_A .

2391 *Proof.* Let $S = \{x_1, \dots, x_n\} \subseteq \mathcal{X}$ be a set of size n that is ϵ -shattered by \mathcal{H}_A^* . That is, there exists
2392 a witness vector $\vec{r} = (r_1, \dots, r_n) \in \mathbb{R}^n$ such that for every binary vector $\vec{b} \in \{0, 1\}^n$, there exists a
2393 function $h^* \in \mathcal{H}_A^*$ satisfying:

$$\forall i \in [n], \quad \begin{cases} h^*(x_i) > r_i + \epsilon & \text{if } b_i = 1, \\ h^*(x_i) < r_i - \epsilon & \text{if } b_i = 0. \end{cases}$$

2394 But for any $h^* \in \mathcal{H}_A^*$, we have $h^*(x) = h(x) \cdot \mathbf{1}[y(x) = v]$ for some $h \in \mathcal{H}_A$. Therefore, $h^*(x_i) = 0$
2395 whenever $y(x_i) \neq v$. In particular, if x_i has $y(x_i) \neq v$, then the above inequalities cannot hold for
2396 any r_i with nonzero margin ϵ .

2397 Hence, only points x_i with $y(x_i) = v$ can be involved in the ϵ -shattering. Let $S_v = \{x_i \in S : y(x_i) = v\}$. Then the shattering must occur over S_v , and the effective shattering occurs only over this subset.

2400 Note that by construction, for each $h^* \in \mathcal{H}_A^*$, there is an $h \in \mathcal{H}_A$ such that $h^*(x_i) = h(x_i)$ for all
2401 $x_i \in S_v$. So the class \mathcal{H} , restricted to S_v , can realize the same shattering. Therefore:

$$C_{\mathcal{H}_A^*}^\epsilon \leq C_{\mathcal{H}_A|S_v}^\epsilon \leq C_{\mathcal{H}_A}^\epsilon$$

2402 □

2403 **M.1 Proof of Lemma E.4**

2404 *Proof of Lemma E.4.* Consider a modified interaction under Protocol E where, at each day in round
 2405 j (if the conversation reaches round j), the outcome is resampled according to the information seen
 2406 by Alice so far: $y' \sim \mathcal{D}_y|x_A^t, \pi^{t-1}, C_{j-1}^t, p_B^{t,j}$. Let $\hat{\pi}^j$ be the transcript from this interaction.

2407 First, we will show that $\mathbb{P}_{\mathcal{D}}[\pi] = \mathbb{P}_{\mathcal{D}}[\hat{\pi}^j]$, where π is the transcript under the unmodified Protocol E.

2408 Let $\hat{\pi}^{1:t,j}$ denote the transcript of this interaction up to day t . Note that this is distinct from $\bar{\pi}^{t,j}$,
 2409 which denotes the transcript of an interaction only on day t where the resampling only occurs in
 2410 round j . We will proceed via induction over days.

2411 • **Base Case:** $\mathbb{P}_{\mathcal{D}}[\pi^{1:1}] = \mathbb{P}_{\mathcal{D}}[\hat{\pi}^{1:1,j}]$.

2412 *Proof:* On day $t = 1$, we have $\mathbb{P}_{\mathcal{D}}[\pi^1] = \mathbb{P}_{\mathcal{D}}[\bar{\pi}^{1,j}]$, by Lemma E.3. Note that $\bar{\pi}^{1,j} = \bar{\pi}^{1:1,j} =$
 2413 $\hat{\pi}^{1:1,j}$, and therefore $\mathbb{P}_{\mathcal{D}}[\pi^{1:1}] = \mathbb{P}_{\mathcal{D}}[\hat{\pi}^{1:1,j}]$.

2414 • **Inductive Step:** If $\mathbb{P}_{\mathcal{D}}[\pi^{1:t}] = \mathbb{P}_{\mathcal{D}}[\hat{\pi}^{1:t,j}]$, then $\mathbb{P}_{\mathcal{D}}[\pi^{1:t+1}] = \mathbb{P}_{\mathcal{D}}[\hat{\pi}^{1:t+1,j}]$.

2415 *Proof:* Observe that the state of the model algorithm in any round $t + 1$ is a function only of
 2416 the algorithm M and the transcript until that round: $\pi^{1:t}$ or $\bar{\pi}^{1:t}$. By the Inductive Hypothesis,
 2417 $\mathbb{P}_{\mathcal{D}}[\pi^{1:t}] = \mathbb{P}_{\mathcal{D}}[\hat{\pi}^{1:t,j}]$ – and consequently, since the model algorithm M is the fixed between
 2418 both interactions, therefore, $\mathbb{P}_{\mathcal{D}}[\pi^{t+1,j}] = \mathbb{P}_{\mathcal{D}}[\bar{\pi}^{t+1,j}]$. By Lemma E.3, this is equal to $\mathbb{P}_{\mathcal{D}}[\pi^{t+1}]$.
 2419 As $\mathbb{P}_{\mathcal{D}}[\hat{\pi}^{1:t,j}] = \mathbb{P}_{\mathcal{D}}[\pi^{1:t}]$ and $\mathbb{P}_{\mathcal{D}}[\bar{\pi}^{t+1,j}] = \mathbb{P}_{\mathcal{D}}[\pi^{t+1}]$, we have that $\mathbb{P}_{\mathcal{D}}[\pi^{1:t+1}] = \mathbb{P}_{\mathcal{D}}[\hat{\pi}^{1:t+1,j}]$.

2420 Now, all that remains to show is that Alice’s sequence of predictions in $\hat{\pi}(j)$ has low expected regret
 2421 with respect to h . Recall that Alice is a Bayesian Learner (Definition E.1), which means that her
 2422 prediction in round k is deterministic after round $k - 1$, and is the posterior mean of the distribution
 2423 conditioned on the transcript up to day $t - 1$, their features on day t , and the conversation of day
 2424 t through round $k - 1$. Since squared error is a proper scoring rule, it follows that predicting the
 2425 mean of the sampling distribution, as Alice does, has lower squared error than predicting any other
 2426 post-processing of the information available to her, and in particular, the function $h \in \mathcal{H}_A$, which is
 2427 defined only on Alice’s features x_A , a subset of the information she has conditioned on. Therefore, it
 2428 follows that a perfect Bayesian will have 0 regret with respect to the swap function over her $[\frac{1}{m}]$ level
 2429 sets defined by the m fixed functions in \mathcal{H}_A , notated as $\{h_0, h_{\frac{1}{m}}, \dots, h_1\}$.

$$\mathbb{E}_{\mathcal{D}}[(\hat{y} - y)^2] \leq \mathbb{E}_{\mathcal{D}} \left[\sum_v \mathbb{I}[\hat{y} = v] (h_v(x) - y)^2 \right].$$

2430 However, since Alice and Bob are not *perfect* Bayesians in Protocol E, but instead round their
 2431 prediction to the nearest multiple of $\frac{1}{m}$, their expected regret with respect to h will depend on this
 2432 discretization.

$$\begin{aligned} \mathbb{E}_{\mathcal{D}}[(\bar{y} - y)^2] &= \mathbb{E}_{\mathcal{D}}[(\hat{y} - y + \bar{y} - \hat{y})^2] \\ &= \mathbb{E}_{\mathcal{D}}[(\hat{y} - y)^2 + (\bar{y} - \hat{y})^2 + 2(\hat{y} - y)(\bar{y} - \hat{y})]. \end{aligned}$$

2433 Since $|\hat{y} - y| < \frac{1}{m}$ and $\mathbb{E}_{\mathcal{D}}[2(\hat{y} - y)(\bar{y} - \hat{y})] = 0$, we have

$$\mathbb{E}_{\mathcal{D}}[(\bar{y} - y)^2] \leq \mathbb{E}_{\mathcal{D}}[(\hat{y} - y)^2] + \frac{1}{m^2}.$$

2434 We have shown the claim for an arbitrary set of m functions in $\mathcal{H}_A : \{h_0, h_{\frac{1}{m}}, \dots, h_1\}$, and can thus
 2435 conclude that it holds for any swap function with respect to \mathcal{H}_A . \square

2436 **Theorem M.2** (Azuma’s Inequality). *Let $\{X_0, X_1, \dots\}$ be a martingale sequence such that $|X_{i+1} -$*
 2437 *$X_i| < c$ for all i , then,*

$$\mathbb{P}[X_n - X_0 \geq \epsilon] \leq \exp \left(-\frac{\epsilon^2}{2c^2n} \right).$$

2438 *An immediate corollary of Theorem M.2 follows from appropriately setting parameters.*

2439 **Corollary M.3.** Letting $X_0 = 0, \varepsilon = c\sqrt{2n \ln \frac{1}{\delta}}$, then we have for any $\delta \in (0, 1)$, with probability
 2440 $1 - \delta$,

$$X_n \leq c\sqrt{2n \ln \frac{1}{\delta}}.$$

2441 **Lemma M.4.** Let $E : \Pi \rightarrow [0, 1]$ represent any conditioning event. Consider the random process
 2442 $\{\mathcal{Z}^t\}$ adapted to the sequence of random variables π^t for $t \geq 1$ and let

$$\mathcal{Z}^t := Z^{t-1} + E(\pi^{1:t-1}) \cdot (y^t(\pi^{1:t-1}) - \mathbb{E}_{y \sim \mathcal{D}}[y|\pi^{1:t-1}])$$

2443 Then,

$$\sum_{t=1}^T E(\pi^{1:t-1}) \cdot (y^t(\pi^{1:t-1}) - \mathbb{E}_{y \sim \mathcal{D}}[y|\pi^{1:t-1}]) \leq 2\sqrt{2T \ln \frac{1}{\delta}},$$

2444 with probability $1 - \delta$ over the randomness of \mathcal{D} and $\pi^{1:t-1}$.

2445 *Proof.* First, observe that the above sequence is a martingale as $\mathbb{E}_{\mathcal{D}}[E(\pi^{1:t-1}) \cdot (y^t(\pi^{1:t-1}) - \mathbb{E}_{y \sim \mathcal{D}}[y|\pi^{1:t-1}]) | \pi^{1:t-1}] = E(\pi^{1:t-1}) \cdot \mathbb{E}_{\mathcal{D}}[(y^t(\pi^{1:t-1}) - \mathbb{E}_{y \sim \mathcal{D}}[y|\pi^{1:t-1}]) | \pi^{1:t-1}] = 0$, since $E(\pi^{1:t-1})$ is a
 2446 constant at the start day t as it does not depend on the outcome y^t . Thus, $\mathbb{E}_{\mathcal{D}}[Z^{t+1}] = Z^t$. Next,
 2447 observe that since the outcomes $y \in [-1, 1]$, we have the bounded difference condition: $|Z^t - Z^{t-1}| < 2$ for all t . We can then instantiate Azuma's Inequality with $n = T$ and $c = 2$ to get the claim. \square
 2448
 2449

2450 M.2 Proof of Lemma E.5

2451 *Proof of Lemma E.5.* Fix bucket $i \in \{1, \dots, \frac{1}{g_B(T)}\}$ of Bob's prediction in round $k - 1$. Since
 2452 Alice's prediction is deterministic of round k is deterministic after round $k - 1$, we can instantiate
 2453 Lemma M.2 with the event $E(\pi^{1:T}) = \mathbb{I}[y_B^{t,k-1} \in i]$ and have, that with probability $1 - \delta$,

$$\left| \sum_{t=1}^T E(\pi^{1:t-1}) (\hat{y}_k^t - y^t(\pi^{1:t-1}))^2 - \mathbb{E}_{y \sim \mathcal{D}}[(y - y^t)^2 | \pi^{1:t-1}] \right| \leq 2\sqrt{2T \ln \frac{1}{\delta}}.$$

2454 \square

2455 **Definition M.5.** Let \mathcal{H} be a set of functions mapping from a domain \mathcal{X} to \mathbb{R} and suppose that
 2456 $S = \{x_1, \dots, x_m\} \subseteq \mathcal{X}$. Fix $\gamma > 0$. Then S is γ -shattered by \mathcal{H} if there are real numbers
 2457 r_1, \dots, r_m , such that for each $b \in \{0, 1\}^m$ there is a function h in \mathcal{H} satisfying, for all $i \in [m]$,

$$h(x_i) \geq r_i + \gamma \text{ if } b_i = 1$$

and

$$h(x_i) \leq r_i - \gamma \text{ if } b_i = 0.$$

2458 We say that $r = (r_1, \dots, r_m)$ witnesses the shattering.

2459 **Definition M.6** (Fat Shattering Dimension [Anthony and Bartlett, 1999]). Suppose that \mathcal{H} is a set of
 2460 functions from a domain \mathcal{X} to \mathbb{R} and that $\gamma > 0$. Then \mathcal{H} has γ -dimension d if d is the maximum
 2461 cardinality of subset S of \mathcal{X} that is γ -shattered by \mathcal{H} . If no such maximum exists, we say that
 2462 \mathcal{H} has infinite γ -dimension. The γ -dimension of \mathcal{H} is denoted $\text{FAT}_{\mathcal{H}}(\gamma)$. This defines a function
 2463 $\text{FAT}_{\mathcal{H}} : \mathbb{R}^+ \rightarrow \mathbb{N} \cup \{0, \infty\}$, which we call the fat shattering dimension of \mathcal{H} . We say that \mathcal{H} has finite
 2464 fat shattering dimension whenever it is the case that for all $\gamma > 0$, $\text{FAT}_{\mathcal{H}}(\gamma)$ is finite.

2465 **Theorem M.7** (Anthony and Bartlett [1999]). Let \mathcal{H} be a hypothesis space of real-valued functions
 2466 with finite fat-shattering dimension, then

$$\sup_{h \in \mathcal{H}} \left| \frac{1}{T} \sum_{i=1}^T (h(x^i) - y^i)^2 - \mathbb{E}_{\mathcal{D}}[(h(x^i) - y^i)^2] \right| \leq \varepsilon.$$

2467 for

$$M(\varepsilon, \delta) = O\left(\frac{C_{\mathcal{H}}^{\varepsilon/256} \ln(\frac{1}{\varepsilon}) + \ln(\frac{1}{\delta})}{\varepsilon^2}\right),$$

2468 where $M(\varepsilon, \delta)$ is the number of samples needed to reach ε uniform convergence with probability
 2469 $1 - \delta$.

2470 **M.3 Proof of Lemma E.6**

2471 *Proof of Lemma E.6.* Note that in a Bayesian setting, each (x^t, y^t) are sampled i.i.d. from \mathcal{D} every
 2472 day, which means that for a fixed round k , Bayesian predictions (and consequently the choice of
 2473 the benchmark function and thus value $h(x^t)$) are independent across days. Secondly, note that by
 2474 Lemma M.1, we have that for any scale $\varepsilon > 0$, $C_{\mathcal{H}_A^*}^\varepsilon \leq C_{\mathcal{H}_A}^\varepsilon$ where \mathcal{H}_A^* is the function class defined
 2475 as $\mathcal{H}_A^* = \{h(x) \cdot \mathbf{1}[y(x) = v] : \forall v \in \mathcal{Y}, h \in \mathcal{H}_A\}$.

2476 Thus, we can directly apply Theorem M.7.

$$\sup_{h \in \mathcal{H}_A} \left| \frac{1}{T} \sum_{i=1}^T \ell(h(x^i), y^i) - \mathbb{E}_{\mathcal{D}}[\ell(h(x^i), y^i)] \right| \leq \varepsilon.$$

2477 This means, that on the subsequence $T_B(k-1, i)$, for some level set v of Alice's prediction, we have:

$$\sup_{h \in \mathcal{H}_A} \left| \frac{1}{|T_B(k-1, i)|} \sum_{t \in T_B(k-1, i)} \mathbb{I}[\bar{y}_A^{t,k} = v](h(x^t) - y^t)^2 - \mathbb{E}_{\mathcal{D}}[\mathbb{I}[\bar{y}_A = v](h(x) - y)^2 | \pi^{1:t-1}] \right| \leq \varepsilon.$$

2478 □

2479 **M.4 Proof of Theorem E.2**

2480 *Proof of Theorem E.2.* With probability $1 - \delta$, we have that

$$\begin{aligned} & \sum_{t \in T_B(k-1, i)} (\bar{y}^{t,k} - y^t)^2 - \sum_v \min_{h \in \mathcal{H}_B} \sum_{t \in T_B(k-1, i)} \mathbb{I}[\bar{y}^{t,k} = v](h(x^t) - y^t)^2 \\ & \leq \sum_{t \in T_B(k-1, i)} \mathbb{E}[(\bar{y}^{t,k} - y^t)^2 | \pi^{1:t-1}] + 2\sqrt{2T \ln \frac{1}{\delta}} - \sum_v \min_{h \in \mathcal{H}_B} \sum_{t \in T_B(k-1, i)} \mathbb{I}[\bar{y}^{t,k} = v](h(x^t) - y^t)^2 \\ & \leq \sum_{t \in T_B(k-1, i)} \mathbb{E}[(\bar{y}^{t,k} - y^t)^2 | \pi^{1:t-1}] + 2\sqrt{2T \ln \frac{1}{\delta}} \\ & \quad - \sum_v \min_{h \in \mathcal{H}_B} \sum_{t \in T_B(k-1, i)} \mathbb{E}_{\mathcal{D}}[\mathbb{I}[\bar{y}^{t,k} = v](h(x^t) - y^t)^2 | \pi^{1:t-1}] + mT\varepsilon \\ & \leq 2\sqrt{2T \ln \frac{1}{\delta}} + \frac{T}{m^2} + mT\varepsilon, \end{aligned}$$

2481 where the first inequality comes from Lemma E.5, the second from applying Lemma E.6 to each
 2482 level set v of Alice's prediction, and the third from Lemma E.4. The final statement comes from
 2483 taking a union bound over all buckets $g_B(T)$ and rounds K . □

2484 **M.5 Proof of Theorem E.7**

2485 **Lemma M.8.** Let \mathcal{H}_J be a hypothesis class over the joint feature space \mathcal{X} . Let $\mathcal{H}_A = \{h_A : \mathcal{X}_A \rightarrow \mathcal{Y}\}$ and $\mathcal{H}_B = \{h_B : \mathcal{X}_B \rightarrow \mathcal{Y}\}$ be hypothesis classes over \mathcal{X}_A and \mathcal{X}_B . Consider instance
 2486 $(x_A, x_B, y) \sim \mathcal{D}$. If
 2487

- 2488 • Alice and Bob are both Bayesian learners, with discretization $m = T^{\alpha_g}$, for $\alpha \in [0, \frac{1}{4}]$
- 2489 • \mathcal{H}_A and \mathcal{H}_B jointly satisfy the $w(\cdot)$ -weak learning condition with respect to \mathcal{H}_J for any
 2490 continuous $w(\cdot)$ such that $w(\gamma) > 0$ for $\gamma > 0$,

2491 then under Protocol E.2 the prediction in round K will have low expected error with respect to \mathcal{H}_J
 2492 on day 1, with probability $1 - \delta$:

$$\mathbb{E}[(\hat{y} - y)^2] - \min_{f_j \in \mathcal{H}_J} \mathbb{E}[(f_j(x) - y)^2] \leq \frac{\tilde{O}(T^{\max(\frac{3}{4}, 1-\alpha_g)} \sqrt{\ln \frac{K}{\delta}})}{T}$$

2493 *Proof of Lemma M.8.* By Theorem E.2, we have that after T rounds, with probability $1 - \delta$, Alice
 2494 will have $(2\sqrt{2T \ln \frac{g_A(T)K}{\delta}} + \frac{T}{m^2} + m\sqrt{32T \ln \frac{4g_A(T)K}{\delta}}, g_A(T), \mathcal{H}_A)$ conversation swap regret
 2495 (symmetrically for Bob). We can instantiate this with parameters that are sublinear in T , specifically
 2496 $m = T^{\frac{1}{4}}$ and $g_A(T) = T^{-\alpha_g}$ for some constant $\alpha_g \in (0, 1)$. Then, we know that Alice, with
 2497 probability $1 - \delta'$, satisfy $(f_A, g_A, \mathcal{H}_A)$ -conversation swap regret, for:

$$\begin{aligned} f_A(T) &\leq 2\sqrt{2T \ln \frac{g_A(T)K}{\delta'}} + \frac{T}{m^2} + m\sqrt{32T \ln \frac{4g_A(T)K}{\delta'}} && \text{(by Theorem E.2)} \\ &\leq 2\sqrt{2T \ln \frac{T^{-\alpha_g}K}{\delta'}} + \sqrt{T} + T^{\frac{3}{4}}\sqrt{32 \ln \frac{4T^{-\alpha_g}K}{\delta'}} \\ &\leq \tilde{O}\left(T^{\frac{3}{4}}\sqrt{\ln\left(\frac{K}{\delta'}\right)}\right). \end{aligned}$$

2498 Since guarantees for Bob are symmetric, the same expression holds for him. Thus, by a union bound,
 2499 with probability $\delta' = \delta/2$, with probability $1 - \delta$, Alice and Bob simultaneously have $(f_A, g_A, \mathcal{H}_A)$ -
 2500 conversation swap regret and $(f_B, g_B, \mathcal{H}_B)$ -conversation swap regret, respectively. Protocol E.2 is
 2501 simply a special case of Protocol A, in which (x_A, x_B, y) are drawn from fixed distribution each day.
 2502 Therefore, the guarantees from Theorem C.3 hold, and we have that the predictions in round K have
 2503 low expected error with respect to \mathcal{H}_J :

$$\begin{aligned} \sum_{t=1}^T (p_K^t - y^t)^2 - \min_{h_J \in \mathcal{H}_J} \sum_{t=1}^T (h_J(x^t) - y^t)^2 &\leq \\ 2Tw^{-1} \left(8\left(\frac{\beta(T, f'_A, f'_B) + 1/K}{2}\right)^{\frac{1}{3}} + \frac{1}{2}\beta(T, f_A, f_B) \right) &+ 3\frac{T}{2}(g_A(T) + g_B(T)) + 3KT\beta(T, f'_A, f'_B), \end{aligned}$$

2504 where

$$f'_A(T) = f'_B(T) = \sqrt{T \cdot f_A(T)} = \sqrt{T \cdot \tilde{O}\left(T^{\frac{1}{2}}\sqrt{\ln \frac{K}{\delta}}\right)} = \tilde{O}\left(T^{\frac{3}{4}}\ln^{\frac{1}{4}}\left(\frac{K}{\delta}\right)\right)$$

2505 and thus:

$$\begin{aligned} \beta(T, f_A, f_B) &= \frac{f_A(g_A(T)T)}{Tg_A(T)} + \frac{f_B(g_B(T)T)}{Tg_B(T)} + g_A(T) + g_B(T) \\ &\leq \tilde{O}\left((T^{\alpha_g - \frac{1}{4}})\sqrt{\ln\left(\frac{K}{\delta}\right)} + T^{-\alpha_g}\right), \\ \beta(T, f'_A, f'_B) &= \frac{f'_A(g_A(T)T)}{Tg_A(T)} + \frac{f'_B(g_B(T)T)}{Tg_B(T)} + g_A(T) + g_B(T) \\ &\leq \tilde{O}\left(T^{-\frac{1}{4}}\ln^{\frac{1}{4}}\frac{K}{\delta} + T^{-\alpha_g}\right). \end{aligned}$$

2506 Returning to the expression from Theorem C.3, we see

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^T (p_K^t - y^t)^2 - \frac{1}{T} \min_{h_J \in \mathcal{H}_J} \sum_{t=1}^T (h_J(x^t) - y^t)^2 &\leq 2w^{-1} \left(8\left(\frac{\beta(T, f'_A, f'_B) + 1/K}{2}\right)^{\frac{1}{3}} + \frac{1}{2}\beta(T, f_A, f_B) \right) + 3\frac{1}{2}(g_A(T) + g_B(T)) + 3K\beta(T, f'_A, f'_B) \\ &\leq w^{-1} \left(\tilde{O}\left((T^{-\frac{1}{4}}\ln^{\frac{1}{4}}\frac{K}{\delta} + T^{-\alpha_g} + \frac{1}{K})^{\frac{1}{3}}\right) + \tilde{O}(T^{\alpha_g - \frac{1}{4}})\sqrt{\ln\left(\frac{K}{\delta}\right)} \right. \\ &\quad \left. + K\tilde{O}(T^{-\frac{1}{4}}\ln^{\frac{1}{4}}\frac{K}{\delta}) + T^{-\alpha_g} \right). \end{aligned}$$

2507

□

2508 *Proof of Theorem E.7.* By Lemma M.8 we have established that the cumulative regret grows as $o(T)$.
 2509 The claim we want to show is about the expected regret only on a single day, which pertains K rounds
 2510 of conversation about our instance of interest. In the Bayesian setting, since instances are drawn
 2511 i.i.d. and Bayesian agents make predictions independently across days, only as a function of the draw
 2512 from the prior at the beginning of that day - conversations are also identically and independently
 2513 distributed. Therefore, to argue about the expected error on the instance on any single day, it suffices
 2514 to reason about the average of the cumulative regret over all T days. We can consider what would
 2515 happen to the average expected regret in the limit as $T \rightarrow \infty$,

$$\lim_{T \rightarrow \infty} \frac{w^{-1} \left(\tilde{O}((T^{-\frac{1}{4}} \ln^{\frac{1}{4}} \frac{K}{\delta} + T^{-\alpha_g} + \frac{1}{K})^{\frac{1}{3}}) + \tilde{O}(T^{\alpha_g - \frac{1}{4}}) \sqrt{\ln(\frac{K}{\delta})} + K \tilde{O}(T^{-\frac{1}{4}} \ln^{\frac{-1}{4}}(\frac{K}{\delta}) + T^{-\alpha_g}) \right)}{T}$$

$$= w^{-1}(O(\frac{1}{K})^{\frac{1}{3}}).$$

2516 Thus, we have the claim. □