Bounded-Regret Sequential Learning using Prediction Markets[*] (Extended Abstract)

Sindhu Kutty
Department of Computer Science and Engineering
University of Michigan, Ann Arbor
skutty@umich.edu

Rahul Sami
School of Information
University of Michigan, Ann Arbor
rsami@umich.edu

October 15, 2011

Abstract

We demonstrate a relationship between prediction markets and online learning algorithms by using a prediction market metaphor to develop a new class of algorithms for learning exponential families with expert advice. The specific problem we consider is that of prediction when data is distributed according to a particular member of an exponential family. In such a case, cost function based prediction markets provide a convenient analytical tool for evaluating performance. Prediction markets also provide a natural technique for learning in an environment where expert advice is not available *simultaneously* but *sequentially*; and experts are either honest and informative, or dishonest and adversarial. As in traditional models, we combine advice using weights on experts. However, to exploit these particular features, we use a form of Kelly gambling to relate the weight of an expert to her budget. We give a formal description of this new model along with relevant definitions and show an equivalence between learning maximum likelihood estimates of the natural parameters of an exponential family and combining advice in prediction markets. We provide an abstract architecture for learning in this model that uses this equivalence to simulate a prediction market to update budgets of experts based on their individual loss. We apply this technique to construct a concrete algorithm that achieves bounded-regret.

1 Introduction

The problem of combining inputs from multiple experts and taking actions based on a combined forecast is one of the core areas of machine learning. There has been an equally rich, but mostly separate, literature in economics focused on the power of markets to aggregate information from multiple sources. This has led to the study and deployment of *prediction markets*, which are markets whose main purpose is to aggregate the information of the traders. One connection between online learning algorithms and

[*] This work was supported by the National Science Foundation awards IIS-0812042 and CCF-0728768

prediction markets has been noted recently by [Chen and Vaughan, 2010, Abernethy et al., 2011], who noted that optimization techniques from online learning can be used to develop tractable market makers. In this paper, we develop the reverse connection: we present a general technique to use prediction market metaphors to develop novel machine learning algorithms.

This approach is targeted at real-world machine learning domains that share many of the features of the environments in which markets operate: participation of agents (experts or traders) is incomplete and irregular; there are significant incentives for an attacker to disrupt the functioning of the system; and, information arrives sequentially over time. Markets appear to perform well even in these challenging environments. Additionally, markets also compose well with each other, in the sense that money made in one market is portable to other markets. One of our main goals in using algorithms based on a market metaphor is to develop "open" learning protocols that can be similarly composed across different domains.

In this paper, we describe a technique to develop machine learning algorithms for forecast aggregation systems based on a specific form of prediction market, the *market scoring rule* [Hanson, 2003]. In a market scoring rule, traders earn rewards proportional to the reduction in "loss" (measured using a proper scoring rule) caused by their trades; in other words, the difference in the loss of forecasting based on market price after their trade as compared to the market price before their trade. Our approach involves designing a learning algorithm by tracking a budget for each trader, and simulating a prediction market: Specific market form: Market scoring rule.

Algorithms based on prediction markets are attractive for the particular features of the domains we are interested in, because of the following reasons: First, traders' budgets allow us to control the total net impact of a single identity. By coupling traders' payoffs to the effects of their actions, and limiting their effect so that their budget is never negative, we can provide worst-case bounds against adversarial forecasters. Second, in a setting with honest agents but stochastic outcomes, a budget-proportional betting scheme (the Kelly criterion [Kelly, 1956]) leads to exponential growth in traders' budgets (in expectation), and thus the small initial budgets are not crippling to honest agents in the long run. Third, betting protocols have been used before in machine learning algorithms, for the reasons above (see, *e.g.*, [Shafer and Vovk, 2001]). Prediction markets are a natural extension of betting protocols to a sequential information setting. Traders' profits are based on the extent to which they *change* forecasts, thus ensuring that merely cloning previous information is not profitable.

We develop a modular framework for the construction of such algorithms. One module, the Influence Limiting and Scoring module, is domain independent; this implements the budget update and the determination of influence based on the current budget. The second, domain-specific, module will map updates of the learned forecast to trades in a prediction market.

We demonstrate the applicability of this framework by using it to construct an algorithm for a general problem: For a sequence of items, we need to learn parameters of exponential family models, using data samples received over time from a number of agents. The agents comprise of adversarial attackers and honest agents with stochastic samples of data. The ultimate objective is to minimize a natural notion of regret, over a sequence of inspected items.

The rest of this paper is structured as follows: In Section 1.1 we discuss related work. In Section 2 we provide a formal description of the model we are considering. We also provide performance measures for algorithms in this model. We show an equivalence between exponential families and prediction markets in Section 3. Then in Section 4 we provide an abstract construction for a learning algorithm in this new model and in Section 5 we provide an instantiation of this construction and provide a general technique to prove bounded regret in our model. Finally, in Section 6 we provide conclusions and potential directions for future work. For reviewing purposes, the proofs of the stated results have been deferred to Appendix A. For the readers' convenience, we have appended a table of notations in Appendix B.

1.1 Related Work

Our model has several distinguishing features from traditional online learning algorithms. First, we consider a hybrid model where experts are either honest and informative or dishonest and adversarial. Although hybrid models have been studied previously, these have focussed on distinguishing the labeling process from input generation. For instance, [Alessandro and Munos, 2009] consider a model where inputs are stochastically generated, while labels are adversarial. However, their model assumes that the labeling function is generated from a hypothesis class, rather than combining expert advice. Next, we assume sequential expert advice. Traditionally, expert advice is assumed to be available simultaneously, following Vovk's model [Vovk, 1995]. One consequence of having a sequential model is that it facilitates partial aggregation of expert advice. In other words, it is possible to have rounds where some experts abstain from making predictions. This is close to the 'sleeping experts' model that was originally studied in [Blum, 1997, Freund et al., 1997]; and further investigated under different benchmarks and modeling assumptions by [Kleinberg et al., 2008, Kanade et al., 2009]. However, in this model, the experts that do provide advice in a round are assumed to do so simultaneously thus precluding the need to analyze situations where an adversarial expert may imitate the advice of an honest one. To the best of our knowledge, the fact that later experts can have access to earlier advice has not been previously studied. This is a particular factor in our assessment of the value of each expert's prediction. Another point of distinction is that we consider a combination of expert advice as opposed to a comparison against a single best expert. For instance, the benchmark considered in [Kleinberg et al., 2008] is the best ordering of experts. Further, in an adversarial setting their algorithm requires a reduction where every possible ordering is mapped to a separate expert making it computationally inefficient. In [Freund et al., 1997] while they do consider combining experts, their bound gives higher weight to time instants where the best expert is awake and thus is significantly different.

[Azoury and Warmuth, 2001] consider learning exponential family distributions in a traditional online model (without experts) and provide worst case loss bounds relative to using an offline algorithm. [Dekel et al., 2008] are concerned with eliciting truthful advice from self-interested agents who each believe in different true distributions. Thus, rather than comparing against a particular true distribution, they use the average of all agent's beliefs as a benchmark. Yu et. al. [Yu et al., 2009] also present an algorithm to thwart sybil attacks in recommender systems. However, they do not consider a sequential ordering on the experts; further, unlike our model, their model assumes a strong similarity between the actual labels and the forecasts of some expert. Resnick and Sami [Resnick and Sami, 2007] consider an algorithm that uses a prediction market metaphor to make recommendations using influence limits. In prior work [Kutty and Sami, 2010] we have argued that a prediction market model is useful for learning in the presence of sequential information. Although [Kutty and Sami, 2010] proved information loss and damage bounds for a problem motivated by recommender systems, unlike the current paper, these did not lead to a regret bound. In this paper, we formalize the connection by giving a general technique for mapping learning problems to prediction markets. [Wagman and Conitzer, 2008] have previously posed a problem related to the sequential advice problem we examine here; they consider the selection of the true majority winner in a ballot. They provide rules to limit damage in cases where an agent may create multiple identities thus potentially affecting the outcome. [Chen and Vaughan, 2010] and [Abernethy et al., 2011] have previously explored the connection to learning algorithms to inform the design and understanding of prediction markets. In particular, [Chen and Vaughan, 2010] consider the correspondence between prediction markets with market scoring rules and the Follow the Regularized Leader algorithm proposed by [Kalai and Vempala, 2005] and thus provide insight into the aggregation mechanism of a prediction market. [Abernethy et al., 2011] use convex optimization techniques to design efficient markets. [Storkey, 2011] has considered machine learning algorithms, particularly aggregation methods, and has shown how to interpret these algorithms as prediction markets using appropriately defined utility functions. [Lay and Barbu, 2010] set up and simulate a prediction market to aggregate multiple classifiers and provide an interpretation for the resultant prices. Rather than providing a regret bound of the form we give here, they measure the performance of their algorithm experimentally. Independent of this work, we have learned that [Lahaie and Pennock, 2011] have noticed a connection between exponential family distributions and market scoring rules, similar to the connection we describe in section 3.

2 Model

In this section, we set out a new model of online learning and define a measure of algorithm performance in it. The class of problems we consider is that of fitting a member of an exponential family of probability distribution to each item in a sequence of items.

Formally, suppose that M items $\{1,2,\ldots,M\}$ are arriving into the system, one at a time. We assume that, at any point of time, only one item k is under consideration. Each item will have a realized value X (this may be an integer, real number, or vector). The realized value is drawn from an unknown distribution P_k.

The distribution P_k is assumed to be the member of a known *exponential family* $\mathcal{F}$ of distributions. An exponential family is a generalization of many commonly used statistical families e.g., binomial distributions with unknown probability of success, normal distributions with known variance, normal distributions with unknown mean and variance, etc. Formally, an exponential family $\mathcal{F}$ can be defined by specifying a vector of sufficient statistics $\boldsymbol{\phi}(X)$, and a set of possible parameters $\boldsymbol{\beta}$. Both $\boldsymbol{\phi}$ and $\boldsymbol{\beta}$ are of dimension t.

For exposition, we assume that all M items have distributions within the same family $\mathcal{F}$, although our algorithms and analysis do not require this. We also assume that the true parameter value $\boldsymbol{\beta}_k$ of the kth item is drawn from a known prior hyperdistribution (distribution over parameters) P_0^* and further that P_0^* is a member of the family of *Diaconis-Ylvisaker conjugate priors* for the exponential family $\mathcal{F}$ [Diaconis and Ylvisaker, 1979]. Thus the conjugate prior family $\mathcal{F}^*$ is itself an exponential family. Each member of $\mathcal{F}^*$ is parametrized by a natural parameter $\mathbf{b}$; given a distribution $P_{\mathbf{b}}^*$, the probability that the underlying distribution for item k has parameter $\boldsymbol{\beta}_k$ is given by:

$$P^*_{\mathbf{b}}(\boldsymbol{\beta}_k) = e^{\mathbf{b}.\boldsymbol{\beta}^*_k - \psi^*(\mathbf{b})}$$

where $\mathbf{b}$ has dimension $(t+1)$, and $\boldsymbol{\beta}^*_k = (\boldsymbol{\beta}_k, -\psi(\boldsymbol{\beta}_k))$. Here $\psi(\boldsymbol{\beta}_k)$, known as the log partition function, is the normalization coefficient of the original distribution P_k with parameters $\boldsymbol{\beta}_k$.

The task of the learning algorithm is, for each item k, to make a prediction $Q_k \in \mathcal{F}$ of the probability distribution governing item k. This prediction is made after receiving data from the agents (described below), and with knowledge of the prior P^*_0, but before observing the realized outcome X_k of item k. The true parameters $\boldsymbol{\beta}_k$ of the distribution governing item k are assumed to be independent for different values of k. As before, we assume for simplicity that all items have the same prior distribution P^*_0; this assumption is not necessary for our results.

Agents and data There are N agent identities $\{1, 2, \ldots N\}$ in the system. Of these, some subset $\mathcal{H}$ (the *honest set*) are assumed to be honest agents; the remaining set $\overline{\mathcal{H}}$ are identities controlled by an adversarial attacker.

On an item k, an arbitrary subset of honest agents receive data samples, in an arbitrary fixed order. Suppose that there are n agents who receive data on item k (the number, subset of agents, and order may be different for different k). Without loss of generality, let us denote the honest agents as $u_1, \cdots, u_n$ according to the order in which they make their reports on item k. Each honest agent u_i receives a datapoint x_i drawn from the true distribution P_k governing item k. Let $\mathbf{x}_k = (x_1, x_2, \cdots, x_n)$ denote the data reported by these honest agents on item k.

Based on these assumptions, the sequences $\mathbf{x} \stackrel{\text{def}}{=} (x_1, \cdots, x_n)$ are distributed according to a joint probability distribution. The set of all sequences $(x_1, \cdots, x_n)$ of the honest data sequence can be represented as a tree T_k, as follows: The root node corresponds to a the null sequence (no reports yet), each node at level 1 corresponds to a different value of x_1, each node at level 2 corresponds to a different value of the pair (x_1, x_2), and so on. We can index a node v of T_k at level j by the sequence $(x_1, x_2, \cdots x_j)$ of reports that lead to node v. Moreover, under the assumption that the prior is correct, the joint distribution will satisfy a consistency property: At each node v at level j, the posterior distribution of x_{j+1} is obtained by conditioning the prior on the sequence $(x_1, \cdots x_j)$ that indexes v. Further, the eventual outcome X_k is also jointly distributed with $\mathbf{x}$.

The tree T_k, together with the prior distribution, represents the underlying stochastic process governing the data on item k in the absence of attack. In order to simplify the notation, we will use $\mathbf{x} \sim T_k$ to indicate that $\mathbf{x}$ *and the eventual outcome X_K and true parameter* $\boldsymbol{\beta}_k$ are jointly distributed according to the hyperdistribution P^*_0.

Attack strategies Attackers can modify the observed data by adversarially injecting new data into the observed sequence, but the data they inject can only depend on honest data that has already been revealed.

Now, considering all the M items, the honest data can be described by a sequence $T_1, T_2, \cdots, T_M$ of trees, each with a corresponding joint distribution. By assumption, the data on different items is independent.

Definition 1 *An* **attack strategy** $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_M)$ *is a sequence of item attack strategies* $\mathcal{A}_k$*. Each item attack strategy* $\mathcal{A}_k$ *is a function mapping each node* v *of* T_k *to a series* $\mathcal{A}_k(v)$ *of datapoints to be submitted if and when node* v *is reached.* $\mathcal{A}_k(v)$ *may be empty, or may consist of a series of pairs of the form* $(a_i, \tilde{\mathbf{x}}_i)$*, where* $a_i \in \overline{\mathcal{H}}$ *is an attacker-controlled identity and* $\tilde{\mathbf{x}}_i$ *is a datapoint submitted by the attacker.*

The attack data $\mathcal{A}_k(v)$ may depend on the sequence of reports $(x_1, x_2, \cdots, x_j)$ leading to v, as well as the realized history on all items in $(1, 2, \ldots, k-1)$, the reputation of attack identities and honest identities at the start of the kth item, etc. The only relevant constraint is that $\mathcal{A}_k(v)$ must be independent of future data $(x_{j+1}, x_{j+2}, ..)$. For example, attack datapoints injected at a level-2 node cannot depend on the data that the honest agent u_3 will report in the future.

Apart from this constraint, the attack data are completely arbitrary, and the attacker is free to choose the timing of her data as well as their content. Thus, this notion of an attack strategy admits non-myopic attacks as well: the attacker may inject a datapoint $(a_1, \tilde{\mathbf{x}}_1)$ before anticipated honest data, not because of the immediate effect of the spurious data $\tilde{\mathbf{x}}_1$, but because of the eventual damage it will cause after future honest data. Indeed, the data may be injected to improve a_1's reputation on future items, whence a_1 can cause greater damage.

Now, consider a particular item k. Let $\mathbf{x}$ denote one realized sequence $(x_1, \ldots, x_n)$ of honest data. Let $\tilde{\mathbf{x}}_{\mathcal{A}_k}(\mathbf{x})$ denote the realized sequence of datapoints when the honest data is $\mathbf{x}$, and the attacker is following an attack strategy $\mathcal{A}_k$. We call $\tilde{\mathbf{x}}_{\mathcal{A}_k}(\mathbf{x})$ the **extended data sequence** corresponding to honest data $\mathbf{x}$ under attack $\mathcal{A}_k$. When $\mathcal{A}_k$ or $\mathbf{x}$ is implicit from the context, we abuse notation by dropping the respective symbol, to write $\tilde{\mathbf{x}}(\mathbf{x})$ or simply $\tilde{\mathbf{x}}$.

Learning algorithms A learning algorithm Z takes as input (for each k) an extended sequence $\tilde{\mathbf{x}}$ of data reports on item k, and produces a prediction $Q_{Z,k}(\tilde{\mathbf{x}})$. The algorithm may maintain state between items, so $Q_{Z,k}(\tilde{\mathbf{x}})$ may depend on the data and outcomes of each of the first $(k-1)$ items.

We consider two kinds of outcome information that may be revealed on each item:

1. In the **revealed data** model, only the realized outcome X_k of item k is revealed to the algorithm. In this case, $Q_{Z,k}(\tilde{\mathbf{x}}) \in \mathcal{F}$ is a member of the exponential family of distributions.
2. In the **revealed parameter** model, we make the stronger assumption that the true parameter $\boldsymbol{\beta}_k$ is revealed after each item k. In this case, $Q_{Z,k}(\tilde{\mathbf{x}}) \in \mathcal{F}^*$ is a forecast distribution over parameter values, which is a member of the conjugate prior family $\mathcal{F}^*$. Although we can prove damage bounds and information loss bounds in both models, we can also prove a regret bound in the revealed parameter model, and so we focus on this model in section 5.

We will compare the performance of our learning algorithms to an omniscient algorithm Z_O that knows the honest set $\mathcal{H}$. The algorithm Z_O can filter out any attack datapoints, and hence can compute the optimal posterior distribution given the data $\mathbf{x}$ on each item k.

Losses and incremental gains In this paper, the performance objective we consider is that of maximizing the log score (equivalently, minimizing the log loss). We consider two variants of the log-score, corresponding to the two different information models: In the revealed data model, the log score of a prediction Q_k when the true outcome turns out to be X is given by:

$$L(Q_k, X) \overset{\text{def}}{=} \log\left[Q_k(X)\right] = \hat{\boldsymbol{\beta}} \cdot \boldsymbol{\phi}(X) - \psi(\hat{\boldsymbol{\beta}})$$

where $\hat{\boldsymbol{\beta}}$ is the natural parameter of the forecast distribution Q_k.

In the revealed parameter model, the log score of a prediction Q_k when the true parameter value is $\boldsymbol{\beta}_k$ is given by:

$$L(Q_k, \boldsymbol{\beta}_k) \overset{\text{def}}{=} \log\left[Q_k(\boldsymbol{\beta}_k)\right] = \mathbf{b} \cdot \boldsymbol{\beta}_k^* - \psi^*(\mathbf{b})$$

where $\mathbf{b}$ is the natural parameter of the forecast hyperdistribution Q_k and $\boldsymbol{\beta}_k^*$ is as previously defined.

In each case, note that the loss can be evaluated after the true outcome is known. Now, for the appropriate loss function and a given learning algorithm Z and extended datapoint sequence $\tilde{\mathbf{x}}$, we can define the *incremental gain* for each datapoint as the change in log score that that datapoint induced:

Definition 2 *Consider agent i who reports the ith datapoint in the extended datapoint sequence $\tilde{\mathbf{x}} = (\tilde{x}_1, \ldots \tilde{x}_n)$ for item k. The incremental gain for an agent i in sequence $\tilde{\mathbf{x}}$, given an eventual true outcome X_k, is given by:*

$$G_Z(i, \tilde{\mathbf{x}}, X_k) = \left[L(Q_{Z,k}(\tilde{x}_1, \ldots, \tilde{x}_i), X_k) - L(Q_{Z,k}(\tilde{x}_1, \ldots, \tilde{x}_{i-1}), X_k)\right]$$

We reiterate that this may depend on the state of the algorithm Z at the start of the kth item. In the revealed parameter model, the incremental gain $G_Z(i, \tilde{\mathbf{x}}, \boldsymbol{\beta}_k)$ is defined accordingly.

We also define a notion of the informativeness (information content under optimal updating) of each honest agent u_i, on each item k, in terms of the expected gain under the omniscient algorithm Z_O. Recall that Z_O knows $\mathcal{H}$, and hence is not affected by attack data.

Definition 3 *The* **informativeness** *h_{ik} of agent u_i on item k is defined as:*

$$h_{ik} = E_{\mathbf{x} \sim T_k} G_{Z_O}(u_i, \mathbf{x}, X_k)$$

(In the revealed parameter model, the definition is analogous with $\boldsymbol{\beta}_k$ in place of X_k.)

The informativeness provides us with a benchmark for the incremental gain obtained from an agent i in practice; if the sum of expected incremental gains under an algorithm is equal to the sum of informativeness h_{ik} over all honest i, then the algorithm would be optimally using the received data.

Regret Now, consider a particular item k. Let $\mathbf{x}$ denote one realized sequence $(x_1, \ldots, x_n)$ of honest data. Let $\tilde{\mathbf{x}}_{\mathcal{A}}(\mathbf{x})$ denote the realized sequence of datapoints when the honest data is $\mathbf{x}$, and the attacker is following an attack strategy $\mathcal{A}$. With the definition of stochastic honest data and adversarial attack strategies, we can now define a hybrid stochastic-adversarial notion of *regret*:

Definition 4 *The* **regret** *of algorithm Z is defined as the maximum, over all possible honest sets $\mathcal{H}$, all possible attacks $\mathcal{A}$, and all prior distributions $T_i \in \mathcal{F}^*$, of the reduction in total log score over all items relative to the omniscient algorithm Z_O that knows $\mathcal{H}$.*

$$Reg(Z) = \max_{\mathcal{H}, \{T_i\}, \mathcal{A}} \left\{ \sum_{k=1}^{M} E_{(\mathbf{x}_1, \mathbf{x}_2, ..\mathbf{x}_k) \sim (T_1, T_2, ...T_k)} \left[L(Q_{Z,k}(\tilde{\mathbf{x}}_{\mathcal{A}}(\mathbf{x}_k)), X_k) - L(Q_{Z_O,k}(\mathbf{x}_k), X_k) \right] \right\}$$

Equivalently, the regret can be defined in terms of the incremental gains, as:

$$Reg(Z) = \max_{\mathcal{H},\{T_i\},\mathcal{A}} \left\{ \sum_{k=1}^{M} E_{(\mathbf{x}_1,\mathbf{x}_2,..\mathbf{x}_k)\sim(T_1,T_2,...T_k)} \left[\sum_{i\in\tilde{\mathbf{x}}_{\mathcal{A}}(\mathbf{x}_k)} G_Z(i,\tilde{\mathbf{x}}_{\mathcal{A}}(\mathbf{x}_k),X_k) - \sum_{j\in\mathcal{H}} h_{jk} \right] \right\}$$

Note that, in the definition of regret, the expectation was taken over all possible values of $(\mathbf{x}_1,\mathbf{x}_2,...\mathbf{x}_k)$, and not just over $\mathbf{x}_k$. The reason for this is that the state of the algorithm during item k (including, for example, budget, reputation, or influence values) may depend on the outcomes of earlier items.

In addition to regret, we will analyze weaker performance criteria of *damage* and *information loss*. We will defer their definition until section 4.3, where we define them within the context of our algorithm.

3 Exponential Families and Prediction Markets

In this section, we show that, for any exponential family, we can construct a prediction market (with a set of securities) such that optimal trade in the prediction market is equivalent to optimal updating in the exponential family. Here, we show this equivalence in the absence of attack; the constructed market has traders with unlimited budgets. We also point out an information-theoretic interpretation of this equivalence. This construction and equivalence forms the basis for our construction of a new learning algorithm in -5.

An exponential Family $\mathcal{F}$ with t sufficient statistics $\boldsymbol{\phi}(x)$ is defined as a collection of distributions over $x \in X$ of the form $P_{\boldsymbol{\beta}}(x) \propto \exp(\boldsymbol{\beta}^T\boldsymbol{\phi}(x) - \psi(\boldsymbol{\beta}))$ where $\boldsymbol{\beta}$, the vector of natural parameters, is of dimension t and $\psi(\boldsymbol{\beta}) \overset{\text{def}}{=} \log\int \exp\{\boldsymbol{\beta}^T\boldsymbol{\phi}(x)\}dx$ is the log partition function.

Corresponding to a member of an exponential family, we define a prediction market to be simulated by the learning algorithm. We claim that the maximum likelihood estimate (MLE) of the natural parameters of an exponential family distribution is exactly the same as an aggregation in a prediction market with log market scoring rule (LMSR) and infinite budget traders with relabelling. For a given exponential family distribution, the prediction market is defined as follows: For each sufficient statistic $l = 1,2,\dots,t$, we define a security s_l with payoff $\boldsymbol{\phi}^l(x)$. We define an additional security s_0 with payoff $\phi_0(x) := c - \sum_{l=1}^{t}\phi_l(x)$ where c is an appropriately chosen constant dependent on the range of $\boldsymbol{\phi}$ so that the payoff of s_0 is non-negative. Let $\mathbf{q}^* = (q_0^*, q_1^*, \dots, q_t^*)$ be the number of shares of each security held by the traders. We abuse notation slightly and define the cost function in this prediction market as $\psi(\mathbf{q}) := \log\int \exp\{\mathbf{q}^T\boldsymbol{\phi}(x)\}\ dx$. We define the interpretation function $I(\mathbf{q}) = (q_1 - q_0, q_2 - q_0, \dots, q_l - q_0) \overset{\text{def}}{=} (\beta_1, \beta_2, \dots, \beta_t) \overset{\text{def}}{=} \boldsymbol{\beta}$. This allows us to interpret the state of the market in terms of a prediction on the natural parameters of the distribution. We note that under the assumption of perfectly rational, risk-neutral traders with infinite budget and beliefs as indicated above, the gradient of the cost function at this point is $\left(\frac{\partial\psi(\mathbf{q})}{\partial q_l}\right)_{\mathbf{q}=\mathbf{q}^*} = E_{P_{\boldsymbol{\beta}}(x)}[\boldsymbol{\phi}^l(x)] \overset{\text{def}}{=} \mu_l$ where μ_l is believed to be the expected payoff of the l^{th} security by the last trader who traded in this market.

The relationship between the securities of the prediction market so defined and the MLE of the natural parameters is established by the following lemma:

Lemma 1 *The choice of parameters* $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_t) = (q_1 - q_0, q_2 - q_0, \dots, q_t - q_0)$ *satisfies* $\frac{\partial\psi(\boldsymbol{\beta})}{\partial\beta_l} = \mu_l$. *Further, the vector* $\boldsymbol{\beta}$ *is unique.*

If the exponential family is represented so that there is a unique parameter vector associated with each distribution, the representation is said to be minimal. The Bernoulli, Gaussian, and Poisson distributions all have minimal representations. Now, for an exponential family whose representation is minimal, the gradient mapping of the log partition function from the natural parameters to the mean parameter space is an injection [Wainwright and Jordan, 2008, p.64]. That is, there is a unique parameter vector $\boldsymbol{\beta}$ that satisfies $\frac{\partial \psi(\boldsymbol{\beta})}{\partial \beta_l} = \mu_l$, for each $l \in \{1, \ldots, t\}$. Thus, if the μ_l's correspond to the empirical means, since our choice of $\boldsymbol{\beta}$ satisfies the equality, it must also be the vector corresponding to the maximum likelihood estimate.

We observe a useful alternative view of the market scoring rule market for exponential family learning. We connect the cost, payoff and profit function to information-theoretic quantities associated with the exponential family.

The following result has been previously pointed out by Amari [Amari, 2001].

Lemma 2 (profit decomposition lemma)*: Consider an exponential family $\mathcal{F}$ of distributions over some set of statistics $\boldsymbol{\phi}(x)$, with natural parameters $\boldsymbol{\beta}$. Let $\pi, \rho \in \mathcal{F}$ be any two probability distributions in the family. We use $\boldsymbol{\beta}_\pi$ to denote the natural parameters of π, and likewise, we can define $\boldsymbol{\beta}_\rho, \boldsymbol{\mu}_\pi$, and $\boldsymbol{\mu}_\rho$. We abuse notation slightly and let $\psi(\rho)$ indicate the log partition function of ρ which technically depends on its natural parameters. Let $H(\pi)$ denote the entropy of the distribution π, and $K(\pi||\rho)$ denote the KL-divergence of ρ relative to π. Then, the following equality holds:*

$$K(\pi||\rho) + H(\pi) = \psi(\rho) - \boldsymbol{\beta}_\rho \cdot \boldsymbol{\mu}_\pi \tag{1}$$

Equation 1 gives us an alternative view of the market scoring rule construction. Assume that π is the true distribution, and consider two arbitrary distributions $\rho_1, \rho_2 \in \mathcal{F}$. Note that the first term in the RHS is independent of π, and the second term is linear in the probabilities $\pi(x)$. If we want to measure loss by the KL-divergence, we can do so (in expectation) by setting a cost function that captures the first term, and defining security quantities ($\boldsymbol{\beta}$) and payoffs ($\boldsymbol{\phi}$) to capture the second term. In particular, in a market with cost function ψ, if the market price initially implies a distribution ρ_1, and a trader moves the market to price than implies a distribution ρ_2, then the cost she incurs is $\psi(\rho_2) - \psi(\rho_1)$. The number of securities bought to make this trade is given by the vector ($\boldsymbol{\beta}_{\rho_2} - \boldsymbol{\beta}_{\rho_1}$), and the expected payoff of the securities are given by $\boldsymbol{\mu}_\pi$. Thus, by equation 1, the *net profit* of the trader is equal to $K(\pi||\rho_1) - K(\pi||\rho_2)$, *i.e.*, the reduction in KL-divergence with respect to the true distribution. We note one useful property of this construction: For a fixed vector $\boldsymbol{\beta}$ of purchased securities, the cost is independent of the outcome (and outcome distribution π), while the payoff is independent of the initial market state in which these securities were purchased.

4 Budget-limited Markets

In this section, we describe our solution for constructing learning algorithms based on simulated markets with limited trader budgets. The basic intuition, as discussed in section 1, is to use the budgets to bound the worst-case damage from attacker identities, while simultaneously using budget-proportional betting (Kelly gambling [Kelly, 1956]) to rapidly grow honest agents' budgets.

In this section, we describe a protocol for learning based on traders with limited budgets. The protocol described in this section is abstract, in the sense that it can be used with any prediction and loss measurement meeting certain requirements. We prove two theoretical results that show that this protocol can give us control over two metrics of robust learning: (1) We bound the expected loss of information from honest raters when the algorithm is used in the absence of attack. (2) We bound the worst-case damage that a set of attackers can cause, over any sequence of items and reports, assuming that the attack is carried out after all honest reports are in. In the following sections, we will describe multiple possible instantiations of this protocol, including one instantiation in which these bounds imply a regret bound.

4.1 Architecture

The architecture of the learning algorithms we describe is as follows. For each item k on which predictions are to be made, we receive a sequence of datapoints from users. The process of updating the forecast when a datapoint is received from an agent i is decomposed into two component modules:

- The *Influence Limiting and Scoring (ILS)* module uses the current budget of an agent i to compute an influence $y_{ik} \in [0,1]$, and passes this to the Weighted Trade Market (WTM) module (described below).
- The *Weighted Trade Market* module uses y_{ik} and i's datapoint x_i to update the forecast Q_k. It does not maintain any internal state across items.
- Subsequently, when the outcome on the item is realized, the Weighted Trade Market module uses this information to calculate the gain (profit) G_{ik} of trader i's trade, and passes this back to the ILS module. The ILS module updates the budgets

4.2 Weighted Trade Markets

Here, we describe the Weighted Trade Market module in terms of a set of properties that it must satisfy. Concrete implementations of this module, based on exponential family market models, will be described in section 5.

Consider an item k, and an arbitrary fixed sequence $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ of honest datapoints on this item. Let X_k denote the eventual outcome on this item. Let $\mathbf{y}_k = (y_{ik})$ denote the profile of agent influence values that is passed from the ILS module.

As the WTM module is internally stateless across items, the prediction Q_k depends only on $\mathbf{y}$ and the data sequence $\tilde{\mathbf{x}}$ (which may be $\mathbf{x}$ if there is no attack, or $\tilde{\mathbf{x}}_{\mathcal{A}}(\mathbf{x})$ if under attack). Thus, let us define a slightly modified notation for the gain that makes the role of the influence values explicit, as follows: Let $G(\mathbf{y}_k, i, \tilde{\mathbf{x}}, X_k)$ denote the incremental gain attributed to user i when the vector of influences is $\mathbf{y}_k$. As a further abuse of notation, when we are only considering the impact of one user i's influence y_{ik} while keeping the others fixed, we shall write $G(y_{ik}, i, \tilde{\mathbf{x}}, X_k)$.

With this notation, the WTM must satisfy the following properties:

- **(bounded gain)** Fixing all influence values other than i, we must always have $|G(y_{ik}, i, \tilde{\mathbf{x}}, X_k)| \leq y_{ik}$, i.e., the influence value effectively limits the range of feasible gains. Note that this implies that

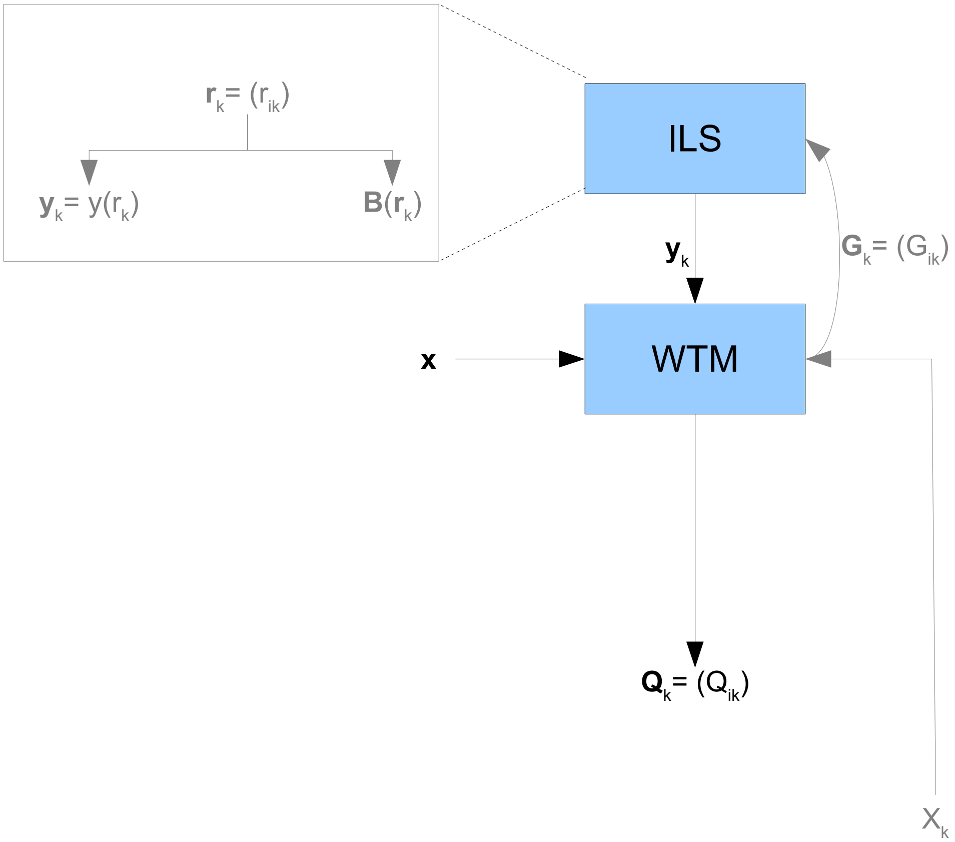
$\mathbf{r}_k = (r_{ik})$
$\mathbf{y}_k = y(r_k)$
$\mathbf{B}(\mathbf{r}_k)$
ILS
$\mathbf{y}_k$
$\mathbf{G}_k = (G_{ik})$
$\mathbf{x}$
WTM
$\mathbf{Q}_k = (Q_{ik})$
X_k

Figure 1: Architecture

$|G(y_{ik}, i, \tilde{\mathbf{x}}, X_k)| \leq 1$ for all y_{ik}, and $G(0, i, \tilde{\mathbf{x}}, X_k) = 0$.

- **(concave gain)** Fixing all influence values other than i, the gain function must satisfy the following concavity property:

  $$G(y_{ik}, i, \tilde{\mathbf{x}}, X_k) \geq y_{ik} G(1, i, \tilde{\mathbf{x}}, X_k)$$

  This property states that, in limiting influence to $y_{ik} < 1$, we get at least a proportional fraction of the gain that we would have got if there was no influence limiting.

- **(bounded variance)** Fix all influence values apart from i. For any honest agent u_i, and any datapoint sequence $\mathbf{x}$, there exists a universal positive constant c that bounds the variance of the gain in terms of its expectation:

  $$\mathrm{Var}[G(1, u_i, \mathbf{x}, X_k)] \leq cE[G(1, u_i, \mathbf{x}, X_k)]$$

- **(substitutes)** If there is no attack, *i.e.*, $\tilde{\mathbf{x}} = \mathbf{x}$, then, for any $\mathbf{y}_k$ such that $y_{ik} = 1$, the expected gain of i must be at least equal to her informativeness:

  $$E_{\mathbf{x} \sim T_k} G(\mathbf{y}, i, \mathbf{x}, X_k) \geq h_{ik}$$

  This is a ‘substitutes’ property in the following sense. In the absence of attack, the informativeness is equal to the expected gain if all agents j before i have full influence $y_{jk} = 1$. If $y_{jk} < 1$, we are making incomplete use of the information (data sample) from agent j. This property says that this should only increase (or keep constant) the value (in terms of expected gain) of the datapoint that agent i submits.

4.3 Influence Limits and Scoring

We now describe the ILS module. The main purpose of this module is to implement a form of Kelly gambling, by tracking a budget and limiting the influence of each agent based on her current budget. Rather than directly expressing the influence in terms of the budget, and the budget updates in terms of the gain of the limited trades, we define the connection indirectly: We express both the budget and the influence as functions of an underlying *reputation* value. This has the analytical advantage that the reputation values are additively updated, and hence easier to analyze for growth over a number of rounds. We then prove bounds on the damage by adversarial attackers, and the expected information loss from stochastic honest raters.

This module has one parameter, α, that is determined based on the smallest constant c for which the WTM bounded variance property is satisfied; setting $\alpha = \min \frac{1}{8}, \frac{1}{4c}$ is adequate. For any gain value G with bounded variance, define the variance-normalized gain $g = \alpha G$. We then have $E[g] = \alpha E[G]$, and $\mathrm{Var}[g] \leq \frac{E[g]}{4}$ when G is the gain of an honest rater.

Define the scoring function $S(g)$, a transformation of the gain, by:

$$S(g) = g - \frac{3}{4} g^2$$

In essence, this transformation makes the score *strictly* concave; we need control over the degree of concavity to prove the damage bound. If the WTM module in fact satisfies a strict concavity property, this transformation may not be necessary. However, as we seek to prove a general result that holds even for weakly concave gain, we work with the transformed gain.

Let $y(r)$ denote the logistic function: $y(r) = \frac{e^r}{1+e^r}$. For notational convenience, we define $\overline{y}(r) = 1 - y(r) = \frac{1}{1+e^r}$. The influence is calculated as follows: For a user i in round k, let $r_{ik} \in (-\infty, +\infty)$ denote the reputation of i at the start of the round. Then, the influence y_{ik} is set to $y(r_{ik})$. Note that $0 < y_{ik} < 1$.

Subsequently, the ILS receives a report of the gain $G_{ik} = G(y_{ik}, i, \tilde{\mathbf{x}}, X_k)$.

Let $\hat{G}_{ik} = (1/y_{ik})G_{ik}$; by the concavity property, $\hat{G}_{ik}$ is at least as high as the actual gain $G(1, i, \tilde{\mathbf{x}}, X_k)$ that would have occured if y_{ik} had been set to 1. The variance-normalized gain $g_{ik} = \alpha\hat{G}_{ik}$. Then, update the reputation by:

$$r_{i(k+1)} = r_{ik} + S(g_{ik})$$

We define the *budget function* $B(r) \overset{\text{def}}{=} \log(1 + e^r)$. Thus, we can speak of a trader with current reputation r_{ik} as having current budget $B(r_{ik})$. Note that $B(r_{ik}) > 0$. As we show below, the current budget serves as a bound on how much a trader can lose in the future.

4.4 Damage and Information Loss Bounds

In this section, we prove some basic properties about our learning algorithm. First, we begin by relating the budgets to the worst-case damage that an attacker identity can induce. Here, we measure damage by the negative of the incremental gain due to that identity.

Theorem 3 *Consider any sequence of items and sequences of reports on those items. For any user i, the net gain due to that user is bounded below in terms of i's initial budget:*

$$\sum_{k=1}^{M} G(y_{ik}, i, \tilde{\mathbf{x}}_k, X_k) \geq -\frac{1}{\alpha}B(r_{i1})$$

As stated, this theorem bounds the net gain attributed to each user. However, because the definition of gain is based on the change in loss, the theorem directly extends to the case of any number of successive reports by attackers. In this case, Theorem 3 can be interpreted as follows: consider an attack involving a number of attack users who insert data only after all honest users. Then, the total damage due to this attack, measured in terms of the gain, is bounded in terms of the total initial budget of all attack identities.

We now consider the expected impact of the ILS module on the use of data from a single honest user. For a user contributing informative data, the influence limits will restrict the contribution of that data, thereby reducing the overall improvement in expected performance. We will now prove a bound on this reduction in performance over all data from a single user.

For the next bound, we consider the case in which *no attack datapoints are inserted prior to the honest agents' data*. Consider the subsequence of items for which user i has submitted a datapoint; without loss of generality, let us assume that this set consists of items $\{1, 2, \ldots, M\}$. First, consider the sequence of forecasts in the absence of influence limits. For each $k \in \{1, 2, \ldots, M\}$, the measured gain $\hat{G}_{ik} = \frac{1}{y_{ik}}G(y_{ik}, i, \mathbf{x}_k, X_k)$ in round k would be a random variable. We use the shorthand notation $G_k = \hat{G}_{ik}$, and $g_k = \alpha\hat{G}_{ik}$.

By the concavity and substitutes properties of the gain function, the expected value of the gain G_k is at least i's informativeness h_{ik}. We assume that $h_{ik} > 0$; $h_{ik} \leq 1$ follows from the requirements on the gain. Finally, we use H_{ik} to denote the partial sum $H_{ik} = \sum_{t=1}^{k} h_{ik}$.

Now, consider the situation with influence limits. In this case, the realized gain in round k will be $G(y_{ik}(r_{ik}), i, \mathbf{x}_k, X_k)$. Here, the reputation r_{ik} is a random variable that is determined based on the realized gains in the previous rounds. We now seek to bound the expected information loss:

$$\mathrm{IL}_i = \sum_{k=1}^{M} h_{ik} - E\sum_{k=1}^{M}\left[G(y_{ik}(r_{ik}), i, \mathbf{x}_k, X_k)\right]$$

Using the concavity of gain, we obtain an alternate expression that is an upper bound on the information loss:

$$G(y_{ik}(r_{ik}), i, \mathbf{x}_k, X_k) = y(r_{ik})\hat{G}_{ik} \geq y(r_{ik})G(1, i, \mathbf{x}_k, X_k) \Rightarrow \mathrm{IL}_i \leq \sum_{k=1}^{M} E[(1-y(r_{ik}))h_{ik}] = \sum_{k=1}^{M} E[\bar{y}(r_{ik})]h_{ik} \tag{2}$$

Note that since r_{ik} is independent of the datapoints in this round $y(r_{ik})$ and $G(1, i, \mathbf{x}_k, X_k)$ are independent random variables.

The intuition behind the bound on information loss is as follows: Suppose that, in each round, the expected gain was some fixed quantity h. Suppose further that the realized score was exactly the expected gain, so that $r_{ik} = r_{i1} + (k-1)h$. Then, the expected information loss, over a very large number of items, would be approximately $\int_{r_{i1}}^{\infty} \bar{y}(x)dx = -\log(y(r_{i1})) = \log(1+e^{-r_{i1}})$ where we have introduced a change of variable with x denoting the range of values for r_{ik}. In other words, the logistic function approaches 1 at a fast enough rate that the total deficit is bounded.

For the actual bound, we need to handle several complications. Firstly, the score $S(g)$ is not a linear function of the gain $\hat{G}_{ik}$, and the expected score is lower than the expected gain. Second, the realized score is not the same as the expected score, and so we need to handle the full distribution of possible values of r_{ik}, and use concentration results to bound the loss. Finally, we need to take into account the fact that different rounds k have different expected gains h_{ik}.

Then, we can prove the information loss bound:

Theorem 4 *Let r_{i1} denote the initial reputation assigned to user i. Fix a sequence of items, and the datapoints submitted by users prior to i's report on each item. Then, the information lost from user i is bounded above by:*

$$IL_i \leq 2 + \frac{8}{\alpha} + \frac{8}{3\alpha}\log(1+e^{-r_{i1}})$$

Together, theorem 3 and theorem 4 imply a regret bound against one narrow class of attack: attacks in which all attack data is inserted after all honest data. In section 5, we will see that for specific instantiations of the WTM module, we can get a full regret bound.

5 Proving a Regret Bound

In this section, we use the fact that the family of conjugate priors is itself an exponential family to construct a market where the market state corresponds to the natural parameters of a *conjugate prior in $\mathcal{F}^*$*, and the securities correspond to the sufficient statistics of the family $\mathcal{F}^*$, which are given by $(\boldsymbol{\beta}, -\psi(\boldsymbol{\beta}))$. For now, we assume that we are operating in a revealed parameter environment, so that the true $\boldsymbol{\beta}$ is eventually known for each item.

In the following proofs, although we assume a certain finite range for $|\boldsymbol{\phi}(x)|$ and $|\boldsymbol{\beta}^*|$, the proofs are easily extended to the case where these ranges hold with high probability. Formally, suppose that we are given an exponential family $\mathcal{F}$ such that $|\boldsymbol{\phi}(x)|$ lies within $(0,1)$. Let $\boldsymbol{\beta}$ denote the natural parameter of the member of $\mathcal{F}$ corresponding to the true distribution of data samples, and let $\boldsymbol{\beta}^*$ denote the vector $(\boldsymbol{\beta}, -\psi(\boldsymbol{\beta}))$. By assumption, $\boldsymbol{\beta}$ is distributed according to the conjugate prior P_0^*. Let $D' > 0$ denote a sufficiently large range such that with $\boldsymbol{\beta}$ drawn from P_0^*, $|\boldsymbol{\beta}^*| < D'$. This implies from the Cauchy-Schwarz inequality that, for any $\boldsymbol{\beta}^*, \hat{\boldsymbol{\beta}^*}$ satisfying this property, and any x, $|(\boldsymbol{\phi}(x),1).(\boldsymbol{\beta} - \boldsymbol{\beta}')| \leq |(\boldsymbol{\phi}(x),1)|.|(\boldsymbol{\beta}^* - \hat{\boldsymbol{\beta}^*})| \leq \sqrt{2}(2D') = D$. This constant D will be the scaling parameter for the market payoffs.

The key property of conjugate prior distributions is that the parameter $\mathbf{b}$ can decomposed as $\mathbf{b} = (r\mathbf{z}, r)$, with $r > 0$, such that the prior $P_{\boldsymbol{\beta}^*}^*$ can be thought of as an observation of r datapoints with mean $\mathbf{z}$. In other words, given a prior distribution $P_{\mathbf{b}_0}^*$ and an observation $\boldsymbol{\phi}(x)$, the posterior distribution over $\boldsymbol{\beta}$ is the member $P_{\mathbf{b}}^*$ of $\mathcal{F}^*$ with parameters $\mathbf{b} = \mathbf{b}_0 + (\boldsymbol{\phi}(x), 1)$ [Diaconis and Ylvisaker, 1979].

The *conjugate prior* market is defined as follows: Consider a fixed item k. The initial state is the hyperparameter $\mathbf{b}_0$ corresponding to the prior P_0^*. At any point of time, the state of the market is determined by the current hyperparemeter $\mathbf{b}$. At this point, if a datapoint x_i is reported by an agent i with influence y_{ik}, the hyperparameter is updated to a value $\mathbf{b}_i$ determined by

$$\mathbf{b}_i = \mathbf{b}_{i-1} + y_{ik}(\boldsymbol{\phi}(x_i), 1)$$

This is treated as a purchase of a security vector $(y_{ik}\boldsymbol{\phi}(x_i), y_{ik})$, in a market with cost function $\psi^*(\mathbf{b})/D$. Eventually, when the true parameter $\boldsymbol{\beta}$ is revealed, the security payoffs are determined by $\boldsymbol{\beta}^* = (\boldsymbol{\beta}, \psi(\boldsymbol{\beta}))$. Thus, the net gain for agent i for item k is given by:

$$G_{ik} = \frac{(\mathbf{b}_i - \mathbf{b}_{i-1}) \cdot \boldsymbol{\beta}^* - [\psi^*(\mathbf{b}_i) - \psi^*(\mathbf{b}_{i-1})]}{D} = \frac{\log(P_{\mathbf{b}_i}^*(\boldsymbol{\beta})) - \log(P_{\mathbf{b}_{i-1}}^*(\boldsymbol{\beta}))}{D}$$

In other words, the gain as measured by this market corresponds to a scaled log score of predicting a hyperdistribution P^* for a true parameter $\boldsymbol{\beta}$.

We now show that this satisfies the properties of WTMs:

- **bounded gain**: Consider an agent i who reports datapoint x_i with influence y_{ik}, an arbitrary sequence of reported data $\tilde{\mathbf{x}}$, and an arbitrary outcome $\boldsymbol{\beta}$. Let $\mathbf{b}_y = \mathbf{b} + y(\boldsymbol{\phi}(x_i), 1)$.

  The gain of agent i is bounded by:

  $$G(y_{ik}, i, \tilde{\mathbf{x}}, \boldsymbol{\beta}) = \frac{\int_{y=0}^{y_{ik}} (\boldsymbol{\phi}(x_i), 1) \cdot (\boldsymbol{\beta}^* - \mathbf{m}(\mathbf{b}_y))dy}{D}$$

  where $\mathbf{m}(\mathbf{b}_y) \overset{\text{def}}{=} \mathbf{m}_{P_{\mathbf{b}_y}^*} = E_{P_{\mathbf{b}_y}^*}\boldsymbol{\beta}^*$ is the mean parameter of the hyperdistribution and corresponds to gradient of the cost function at $\mathbf{b}_y$. The integrand is bounded in absolute value by Ddy, and hence $|G(y_{ik}, i, \tilde{\mathbf{x}}, \boldsymbol{\beta}^*)| < y_{ik}$.

- **concave gain**: For this we will consider the terms constituting the gain separately: the (negative of the) cost function and the payoff function and show that these are separately concave in y_{ik}; and hence so is their sum. The cost function $\psi^*(\mathbf{b}_i)$ is convex in $\mathbf{b}_i$ because it is the cumulant generating function of an exponential family. Since $\mathbf{b}_i = \mathbf{b}_{i-1} + y_{ik}(\boldsymbol{\phi}(x_i), 1)$, it is convex in y_{ik}. The payoff of the purchased securities is linear in y_{ik}, because the quantity of each security purchased $\mathbf{b}_i - \mathbf{b}_{i-1}$ is proportional to y_{ik}. Hence, the net gain is concave in y_{ik}.

- **bounded variance**: Although we do not have a generic proof that the variance is bounded in terms of the mean, for a given family, and a bounded range D of $\boldsymbol{\beta}^*$, it should be possible to derive a constant ratio bound. In particular, for small quantities of securities, the variance is twice the expected score; we believe that this bound will hold in general.
- **substitutes**: The substitutes property holds for this market. This is a consequence of Lemma 6.

5.1 Proving a regret bound

In this section, we show that the combination of the ILS module and the conjugate-prior WTM satisfies a regret bound. The bound itself is simply a combination of the damage and information loss bounds presented in section 4.4. The main challenge is in proving that this bound applies even when attackers can insert datapoints before or between the honest datapoints. The information loss bound did not cover this case, because we could not rule out the possibility that an attacker's strategically chosen data could create more long-term impact than the immediate damage it causes.

The key to this proof is Theorem 5, which shows that under certain conditions, attackers would do better by attacking after an honest rating rather than before it. The relative order of ratings does not matter for the overall log score of our learning algorithm on a given item, because the final prediction Q_q is the same. However, it can affect the incremental gain attributed to different agents. Theorem 5 shows that, when the initial distribution P_b^* is accurate, an attacker would get higher measured gain (or lower measured damage) by moving an attack report to after an honest agent's report. In Theorem 7, we then use this to show that the total regret is bounded by the regret of an attack in which all attack data came after all honest data.

Damage Reduction Consider an exponential family $\mathcal{F}$ with natural parameter $\boldsymbol{\beta}$. Let $\mathcal{F}^*$ denote the Diaconis-Ylvisaker family of conjugate priors. Let $\mathbf{b}$ denote the natural parameters of $\mathcal{F}^*$, and let $\mathbf{m}$ denote its mean parameters. As before, for a distribution $P^* \in \mathcal{F}^*$, we use $\mathbf{b}_{P^*}$ and $\mathbf{m}_{P^*}$ to denote its natural and mean parameters, respectively.

Here we denote the posterior distribution after an observation of sufficient statistics, $\boldsymbol{\phi}(x)$ as P_x^* and the corresponding natural parameter as $\mathbf{b}_{P_x^*}$.

Theorem 5 (damage reduction property): *Let $P_0^* \in \mathcal{F}^*$ denote an initial distribution. Suppose that $\boldsymbol{\phi}(x)$ denotes an observation of sufficient statistics, distributed according to a distribution $\pi(x)$. Then, if P_x^* denotes the posterior hyperdistribution after conditioning on $\boldsymbol{\phi}(x)$; we must have $\mathbf{b}_{P_x^*} = \mathbf{b}_{P_0^*} + (\boldsymbol{\phi}(x), 1)$. The given P_0^* and π must be such that P_0^* is unbiased with respect to π: For any $\boldsymbol{\beta}$, we must have $P_0^*(\boldsymbol{\beta}) = E_\pi P_x^*(\boldsymbol{\beta})$. Consider a vector $\mathbf{a}$ of "attack" securities. The entries of $\mathbf{a}$ may be positive or negative. Let $\tilde{P}_0$ denote the hyperdistribution if the attack is carried out on the prior: $\mathbf{a}_0 = \mathbf{b}_0 + \mathbf{a}$. Likewise, let $\tilde{P}_1$ denote the hyperdistribution with natural parameter coordinates $\mathbf{a}_1 = \mathbf{b}_{P_x^*} + \mathbf{a}$.*

Then, the following condition must hold:

$$K(P_0^* || \tilde{P}_0) \geq E_\pi \left[K(P_1^* || \tilde{P}_1) \right] \tag{3}$$

In other words: the damage induced by a fixed vector of securities **a** *purchased at the initial distribution P_0^* is greater than the expected error of the same vector of securities after an additional informative observation x.*

Theorem 5 shows that the KL-divergence induced by a vector of attack securities is lower (in expectation) after additional honest datapoints have been received.

Regret bound We can now use the damage reduction property of the conjugate prior market to bound the total regret in a system with honest set $\mathcal{H}$ under any attack $\mathcal{A}$.

We start by considering a fixed item k. For now, let us treat the reputation values r_{ik} as fixed; consequently, the influence values $y_{ik} = y(r_{ik})$ are also fixed. Let $\bar{y}_{ik} = 1 - y_{ik}$. Consider a given attack strategy $\mathcal{A}$. There are two potential sources of regret: the effect of the attack data, and the information loss from the fractional use of honest data. Given fixed values of y_{ik}, we can unify these two by thinking of influence limiting as a form of "attack" on the system.

Consider an honest rating x_i, received from honest agent u_i when the market state has natural coordinates $\mathbf{b}_0$. The optimal use of $\boldsymbol{\phi}(x_i)$ would be to update the market state to $\mathbf{b} = \mathbf{b}_0 + (\boldsymbol{\phi}(x_i), 1)$. In actuality, because of influence limiting, the market state would be updated to $\mathbf{b}_j = \mathbf{b}_0 + y_{ik}(\boldsymbol{\phi}(x_i), 1)$. Noting that $\mathbf{b}_j = \mathbf{b}_{j-1} - \bar{y}_{ik}(\boldsymbol{\phi}(x_i), 1)$, we can think of the effect of influence limiting as involving an attack with *negatively weighted* data.

Thus, the effect of influence-limited update can be modeled as fully updating by the data x_i, followed by updating based on negatively-weighted data. Let $\bar{x}_i$ denote this latter half; in other words, $\bar{x}_i$ corresponds to updating the natural coordinates by adding $(-\bar{y}_{ik}\boldsymbol{\phi}(x_i), -\bar{y}_{ik})$. We can think of this negatively weighted datapoint as introduced by a phantom agent $\bar{u}_i$.

We will also use simplified notation to talk about attack. The proof only involves the aggregate effect of all attack identity, and the aggregate effect on the sum of all attackers' budgets. Thus, even though $\mathcal{A}$ may specify a sequence of attack datapoints at each point in T_k, we can lump them together into a single (weighted) datapoint. Let d_i denote the weighted sequence of ratings that the attacker a_i injects after x_i, with the understanding that the d_i may vary with $(x_1, ..., x_i)$. The weight can capture the effect of influence limiting of the attackers as well as the effect of multiple attack datapoints, and so we do not need to separately model a negative-weighted datapoint to capture the effect of an attack rating.

Now, for a given sequence $\mathbf{x} = (x_1, \ldots, x_n)$ of honest data, the combined effect of influence limiting and attack ratings can be modeled by an extended sequence $\tilde{\mathbf{x}}(\mathbf{x}, \mathcal{A}) = x_1\bar{x}_1 d_1 x_2\bar{x}_2 d_2 \cdots x_n\bar{x}_n d_n$.

For a given extended sequence $\tilde{\mathbf{x}}$, there is a fixed posterior distribution (determined by the honest data $\mathbf{x}$)) on the eventual outcome. Thus, the expected (over possible outcomes) gain of each datapoint in $\tilde{\mathbf{x}}$ is well defined. Denote by $G(u_i, \tilde{\mathbf{x}}, X_k)$, $G(\bar{u}_i, \tilde{\mathbf{x}}, X_k)$ and $G(a_i, \tilde{\mathbf{x}}, X_k)$ the expected gains due to the datapoints x_i, $\bar{x}_i$, d_i respectively .

As we consider influence limits as a special kind of attack datapoint $\bar{x}_i$, we can generalize the concept of an attack to denote any mapping from each $\mathbf{x}$ to the corresponding $\tilde{\mathbf{x}}$ such that d_i occurs after x_i and d_i depends only on $(x_1, \cdots, x_i)$. Starting from the given attack $\mathcal{A}$, let us define a new generalized attack $\mathcal{A}'$ by specifying its extended sequence mapping:

$$\tilde{\mathbf{x}}'(\mathbf{x}, \mathcal{A}') = x_1 x_2 x_3 ... x_n \bar{x}_1 d_1 \bar{x}_2 d_2 ... \bar{x}_n d_n$$

Note that $\mathcal{A}'$ does not necessarily correspond to any feasible combination of attack ratings and influence limits, because $\bar{x}_i$ has a negative weight whereas a real attack rating must have a positive weight. However, we are using $\mathcal{A}'$ purely for formal analysis of gain and regret, and so this does not matter.

Now, consider a given sequence $\mathbf{x}$, and let $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{x}}'$ denote the extended sequences corresponding to $\mathcal{A}$ and $\mathcal{A}'$ respectively. The sequence $\tilde{\mathbf{x}}'$ has the property that all the honest ratings are fully accounted for before the attack or influence-limit datapoints are received. For a given sequence $\mathbf{x}$, the regret of our algorithm given "attack" $\mathcal{A}'$ is therefore equal to the total error introduced by the attackers, which is the negative of the sum of gains of all attack and influence-limit datapoints. Then, the regret $\mathrm{Reg}_k(\mathcal{A}')$ of our algorithm on item k with respect to attack $\mathcal{A}'$ can be equivalently written as

$$\mathrm{Reg}_k(\mathcal{A}') = -E_{\{r_{ik}\}}E_{\mathbf{x}\sim T_k}\left[\sum_{i\in\mathcal{H}} G(\bar{u}_i, \tilde{\mathbf{x}}', X_k) + \sum_{i\notin\mathcal{H}} G(a_i, \tilde{\mathbf{x}}', X_k)\right]$$

We note that $\mathrm{Reg}_k(\mathcal{A})$ can not be analogously defined. This is due to the inherent complication of combining information loss and damage in extended sequences. Also note that, for any given $\mathbf{x}$, the final prediction of the algorithm is the same under both attacks $\mathcal{A}$ and $\mathcal{A}'$, because the order of the updates does not matter. It follows that, *for any given influence limits y_{ik}*, the regret $\mathrm{Reg}_k(\mathcal{A}')$ with respect to attack $\mathcal{A}'$ is the same as the regret with respect to attack $\mathcal{A}$ in round k.

The following lemma shows that each honest rater has a higher expected gain under $\mathcal{A}$ than under $\mathcal{A}'$. This implies that the constructed market satisfies the substitutes property.

Lemma 6 *Let $\Delta_i \overset{\text{def}}{=} G(u_i, \tilde{\mathbf{x}}, X_k) - G(u_i, \tilde{\mathbf{x}}', X_k)$. Then $\forall i\ \Delta_i \geq 0$*

The proof of this lemma involves using the error-reduction property (Theorem 5) inductively to reorder these datapoints. Using this lemma we can show that, in the actual attack $\mathcal{A}$, the total regret on any given item is no more than the sum of the damage that attackers are penalized for and the information losses of each honest expert. Then, we can bound the total regret by the sum of the damage and information loss bounds.

The following theorem states the final regret bound over all k items.

Theorem 7 *The regret of the conjugate-prior algorithm is bounded by:*

$$D\left\{\sum_{i\in\mathcal{H}}\left[2+\frac{8}{\alpha}+\frac{8}{3\alpha}\log(1+e^{-r_{i1}})\right]+\sum_{i\notin\mathcal{H}}\frac{1}{\alpha}B(r_{i1})\right\}$$

6 Conclusion and Future Work

Our aim has been to highlight some deep connections between prediction markets and online learning. To that end, we have demonstrated an equivalence between learning the estimating the natural parameters of a member of an exponential family distribution and aggregating trader beliefs in a prediction market. We used this equivalence to construct an online learning algorithm that makes predictions based on the advice of weighted experts. These weights are updated according to budget changes in the simulated prediction market; the performance of the experts in this algorithm is measured according to a particular loss function which satisfies certain properties. The algorithm so defined has certain requirements on the underlying loss function. We defined a particular instance of this abstract learning algorithm that uses the log-loss function. We prove information loss and damage bounds on this algorithm and provide a general technique to combine these to prove an overall regret bound.

This interesting equivalence leaves open a few questions. In our algorithm we have assumed that the true parameter is available to us as feedback. Relaxing this assumption would require scoring on the actual outcome rather than the distribution from which it is drawn. In this case, we could construct algorithms with damage and information loss bounds, but it is not clear whether a combined regret bound would be possible. We have assumed here that there is a bound the range of the natural parameter. How essential is this bound? Is it possible to generalize this technique so that even for an unknown family such parameter estimates are possible? As is usual in Bayesian updates, we have assumed that we have access to a true prior. It would be instructive to see the effect of a non-informative prior on the results.

References

[Abernethy et al., 2011] Abernethy, J., Chen, Y., and Wortman Vaughan, J. (2011). An optimization-based framework for automated market-making. In *Proceedings of the 12th ACM conference on Electronic commerce*, EC '11, pages 297–306, New York, NY, USA. ACM.

[Alessandro and Munos, 2009] Alessandro, L. and Munos, R. (2009). Hybrid stochastic-adversarial on-line learning. In *22nd Annual Conference on Learning Theory (COLT 2009)*.

[Amari, 2001] Amari, S.-I. (2001). Information geometry on hierarchy of probability distributions. *Information Theory, IEEE Transactions on*, 47(5):1701 –1711.

[Azoury and Warmuth, 2001] Azoury, K. S. and Warmuth, M. K. (2001). Relative loss bounds for on-line density estimation with the exponential family of distributions. *Machine Learning*, 43:211–246.

[Bennett, 1962] Bennett, G. (1962). Probability inequalities for the sum of independent random variables. *J. Amer. Statist. Soc.*, 57:33–45.

[Blum, 1997] Blum, A. (1997). Empirical support for winnow and weighted-majorityalgorithms: Results on a calendar scheduling domain. *Machine Learning*, 26:5–23.

[Chen and Vaughan, 2010] Chen, Y. and Vaughan, J. W. (2010). A new understanding of prediction markets via no-regret learning. In *Proceedings of the 11th ACM conference on Electronic commerce*, EC '10, pages 189–198, New York, NY, USA. ACM.

[Dekel et al., 2008] Dekel, O., Fischer, F., and Procaccia, A. D. (2008). Incentive compatible regression learning. In *Proceedings of the nineteenth annual ACM-SIAM symposium on Discrete algorithms*, SODA '08, pages 884–893, Philadelphia, PA, USA. Society for Industrial and Applied Mathematics.

[Diaconis and Ylvisaker, 1979] Diaconis, P. and Ylvisaker, D. (1979). Conjugate Priors for Exponential Families. *The Annals of Statistics*, 7(2):269–281.

[Freund et al., 1997] Freund, Y., Schapire, R. E., Singer, Y., and Warmuth, M. K. (1997). Using and combining predictors that specialize. In *Proceedings of the twenty-ninth annual ACM symposium on Theory of computing*, STOC '97, pages 334–343, New York, NY, USA. ACM.

[Gut, 1995] Gut, A. (1995). *An Intermediate Course in Probability*. Springer-Verlag.

[Hanson, 2003] Hanson, R. (2003). Combinatorial information market design. *Information Systems Frontiers*, 5(1):107–119.

[Kalai and Vempala, 2005] Kalai, A. and Vempala, S. (2005). Efficient algorithms for online decision problems. *J. Comput. Syst. Sci.*, 71:291–307.

[Kanade et al., 2009] Kanade, V., McMahan, H. B., and Bryan, B. (2009). Sleeping experts and bandits with stochastic action availability and adversarial rewards. In *Proceedings of the 12th International Conference on Artificial Intelligence and Statistics*.

[Kelly, 1956] Kelly, J. L. (1956). A new interpretation of information rate. *Bell System Technical Journal*, 35:917–926.

[Kleinberg et al., 2008] Kleinberg, R. D., Niculescu-Mizil, A., and Sharma, Y. (2008). Regret bounds for sleeping experts and bandits. In *21st Annual Conference on Learning Theory - COLT 2008*, pages 425–436.

[Kutty and Sami, 2010] Kutty, S. and Sami, R. (2010). A prediction market approach to learning with sequential advice. In *NIPS Workshop on Computational Social Science and the Wisdom of Crowds*, Vancouver, Canada.

[Lahaie and Pennock, 2011] Lahaie, S. and Pennock, D. (2011). Personal Communication.

[Lay and Barbu, 2010] Lay, N. and Barbu, A. (2010). Supervised Aggregation of Classifiers using Artificial Prediction Markets. In Fürnkranz, J. and Joachims, T., editors, *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*, pages 591–598, Haifa, Israel. Omnipress.

[Resnick and Sami, 2007] Resnick, P. and Sami, R. (2007). The influence limiter: Provably manipulation-resistant recommender systems. In *Proceedings of the ACM Recommender Systems Conference (RecSys07)*.

[Shafer and Vovk, 2001] Shafer, G. and Vovk, V. (2001). *Probability and Finance: It's Only a Game!* John Wiley and Sons.

[Storkey, 2011] Storkey, A. J. (2011). Machine learning markets. *Proceedings of AI and Statistics*.

[Tang and See, 2009] Tang, H.-K. and See, C.-T. (2009). Variance inequalities using first derivatives. *Statistics and Probability Letters*, 79(9):1277 – 1281.

[Vovk, 1995] Vovk, V. G. (1995). A game of prediction with expert advice. In *Proceedings of the eighth annual conference on Computational learning theory*, COLT '95, pages 51–60, New York, NY, USA. ACM.

[Wagman and Conitzer, 2008] Wagman, L. and Conitzer, V. (2008). Optimal false-name-proof voting rules with costly voting. In *AAAI'08: Proceedings of the 23rd national conference on Artificial intelligence*, pages 190–195. AAAI Press.

[Wainwright and Jordan, 2008] Wainwright, M. J. and Jordan, M. I. (2008). Graphical models, exponential families, and variational inference. *Found. Trends Mach. Learn.*, 1:1–305.

[Yu et al., 2009] Yu, H., Shi, C., Kaminsky, M., Gibbons, P. B., and Xiao, F. (2009). Dsybil: Optimal sybil-resistance for recommendation systems. In *Proceedings of the 2009 30th IEEE Symposium on Security and Privacy*, pages 283–298, Washington, DC, USA. IEEE Computer Society.

A Proofs of theorems

A.1 Proofs for Section 3

Lemma 1 *We define the interpretation function* $I(\mathbf{q}) = (q_1 - q_0, q_2 - q_0, \ldots, q_l - q_0) \overset{\text{def}}{=} (\beta_1, \beta_2, \ldots, \beta_t) \overset{\text{def}}{=} \boldsymbol{\beta}$*. This allows us to interpret the state of the market in terms of a prediction on the natural parameters of the distribution. This choice of parameters satisfies* $\frac{\partial\psi(\boldsymbol{\beta})}{\partial\beta_l} = \mu_l$*. Further, this vector is unique.*

Proof: We observe that for $l = 1, \ldots, t$,

$$\begin{aligned}
\mu_l = \left(\frac{\partial\psi(\mathbf{q})}{\partial q_i}\right)_{\mathbf{q}=\mathbf{q}^*} &= \left(\frac{\partial \log \int \exp(\mathbf{q}^T \boldsymbol{\phi}(x))dx}{\partial q_l}\right)_{\mathbf{q}=\mathbf{q}^*} \\
&= \left(\frac{\partial \log \int \exp\{q_0 \boldsymbol{\phi}^0(x) + \sum_{l=1}^t q_l \boldsymbol{\phi}^l(x)\}dx}{\partial q_l}\right)_{\mathbf{q}=\mathbf{q}^*} \\
&= \left(\frac{\partial \log \int \exp\{q_0(c - \sum_{l=1}^t \boldsymbol{\phi}^l(x)) + \sum_{l=1}^t q_l \boldsymbol{\phi}^l(x)\}dx}{\partial q_l}\right)_{\mathbf{q}=\mathbf{q}^*} \\
&= \left(\frac{\partial \log \int \exp\{q_0 c\} \exp\{\sum_{l=1}^t (q_l - q_0)\boldsymbol{\phi}^l(x)\}dx}{\partial q_l}\right)_{\mathbf{q}=\mathbf{q}^*} \\
&= \left(\frac{\partial(q_0 c)}{\partial q_i} \frac{\partial \log(\int \exp\{\sum_{l=1}^t (q_l - q_0)\boldsymbol{\phi}^l(x)\}dx)}{\partial q_l}\right)_{\mathbf{q}=\mathbf{q}^*} \\
&= \left(\frac{\partial\psi(\boldsymbol{\beta})}{\partial q_l}\right)_{\mathbf{q}=\mathbf{q}^*} \\
&= \left(\frac{\partial\psi(\boldsymbol{\beta})}{\partial\beta_l} \frac{\partial\beta_l}{\partial q_l}\right)_{\mathbf{q}=\mathbf{q}^*} \\
&= \left(\frac{\partial\psi(\boldsymbol{\beta})}{\partial\beta_l}\right)_{\boldsymbol{\beta}=\boldsymbol{\beta}^*}
\end{aligned}$$

■

The following result has been previously pointed out by Amari [Amari, 2001]. We reproduce the result and its proof as our notation is slightly different from that used by Amari.

Lemma 2 (profit decomposition lemma)*: Consider an exponential family* $\mathcal{F}$ *of distributions over some set of statistics* $\boldsymbol{\phi}(x)$*, with natural parameters* $\boldsymbol{\beta}$*. Let* $\pi, \rho \in \mathcal{F}$ *be any two probability distributions in the family. We use* $\boldsymbol{\beta}_\pi$ *to denote the natural parameters of* π*, and likewise, we can define* $\boldsymbol{\beta}_\rho$*,*$\boldsymbol{\mu}_\pi$*, and* $\boldsymbol{\mu}_\rho$*. We abuse notation slightly and let* $\psi(\rho)$ *indicate the log partition function of* ρ *which technically depends on its natural parameters. Let* $H(\pi)$ *denote the entropy of the distribution* π*, and* $K(\pi||\rho)$ *denote the KL-divergence of* ρ *relative to* π*. Then, the following equality holds:*

$$K(\pi||\rho) + H(\pi) = \psi(\rho) - \boldsymbol{\beta}_\rho \cdot \boldsymbol{\mu}_\pi \tag{4}$$

Proof:

$$\begin{aligned}
\text{LHS } = K(\pi||\rho) + H(\pi) &= \int_x \pi(x)\log\frac{\pi(x)}{\rho(x)}dx - \int_x \pi(x)\log\pi(x)dx \\
&= -\int_x \pi(x)\log\rho(x)dx \\
&= -\int_x \pi(x)\left[\boldsymbol{\beta}_\rho \cdot \boldsymbol{\phi}(x) - \psi(\rho)\right]dx \\
&= -\boldsymbol{\beta}_\rho \cdot \int_x \pi(x)\boldsymbol{\phi}(x)dx + \psi(\rho)\int_x \pi(x)dx \\
&= \text{RHS}
\end{aligned}$$

■

A.2 Proofs for Section 4

Theorem 3 *Consider any sequence of items and sequences of reports on those items. For any user i, the net gain due to that user is bounded below in terms of i's initial budget:*

$$\sum_{k=1}^{M} G(y_{ik}, i, \tilde{\mathbf{x}}_k, X_k) \geq -\frac{1}{\alpha}B(r_{i1})$$

Proof: Consider a fixed round k. For notational convenience, let $r_0 = r_{ik}$, and let $\hat{G}_{ik}$ denote the scaled gain $(1/y_{ik})G(y_{ik}, i, \tilde{\mathbf{x}}_k, X_k)$ in this round, let $g = \alpha\hat{G}_{ik}$, and let $r = r_0 + S(g)$ denote the final reputation. Recall that $B(r) = \log(1+e^r)$. Taking derivatives, we observe that $B'(r) = y(r)$, and $y'(r) = y(r)/(1+e^r)$.

Now, consider the function $B(r) = B(r_0 + S(g))$ as a function of g. We will show that this is a concave function of g. Differentiating twice with respect to g, we see that

$$\begin{aligned}
\frac{dB(r)}{dg} &= y(r)\frac{dS(g)}{dg} = y(r)[1-1.5g] \\
\frac{d^2B(r)}{dg^2} &= \left[(1-1.5g)\frac{y(r)}{1+e^r}(1-1.5g) - 1.5y(r)\right] \\
&\leq y(r)\left[(1-1.5g)^2 - 1.5\right]
\end{aligned}$$

We have $|g| \leq \frac{1}{8}$, and thus, $\frac{d^2B(r)}{dg^2} \leq 0$, and hence $B(r)$ is a concave function of g.

It follows from concavity that:

$$B(r) \leq B(r_0) + g\frac{dB}{dg}|_{g=0} \Rightarrow B(r) - B(r_0) \leq gy(r_0) \tag{5}$$

Now, we move to the theorem statement

$$\begin{aligned}
\alpha\sum_{k=1}^{M} G(y_{ik}, i, \tilde{\mathbf{x}}_k, X_k) &= \alpha\sum_k y_{ik}\hat{G}_{ik} \quad (\text{ by definition of } \hat{G}_{ik}) && (6) \\
&\geq \sum_k [B(r_{i(k+1)}) - B(r_{ik})] \quad (\text{by eqn. 5}) && (7) \\
&= [B(r_{iM}) - B(r_{i1})] \geq -B(r_{i1}) && (8)
\end{aligned}$$

Dividing both sides by α, we obtain the stated result. ■ In order to prove the information loss bound first, we show that the mean and variance of the score $S(g)$ are "well-behaved":

Lemma 8 *Suppose that G_k has mean h_{ik} and variance at most ch_{ik}. Then,*

$$E[S(g_k)] \geq \frac{3\alpha}{4} h_{ik}$$

and

$$Var[S(g_k)] \leq 0.5E[S(g_k)]$$

Proof: Consider two cases, based on the value of h_{ik}. For $h_{ik} \geq 0.5$, we have

$$E[G_k^2] = E[(\frac{1}{y_{ik}} G(y_{ik}, i, \mathbf{x}_k, X_k))^2] \leq 2h_{ik}$$

which follows by the bounded gain property. Thus,

$$E[S(g_k)] = \alpha E[G_k] - \frac{3\alpha^2}{4} E[G_k^2] \geq \alpha h_{ik}[1 - 1.5\alpha] \geq \alpha h_{ik}[1 - \frac{1.5}{8}] \geq \frac{3\alpha}{4} h_{ik}$$

Now, consider the case when $h_{ik} \leq 0.5$. We note that $E[G_k^2] = E[G_k]^2 + Var[G_k] \leq h_{ik}^2 + ch_{ik}$. Now,

$$\begin{aligned} E[S(g_k)] &= \alpha E[G_k] - \frac{3}{4}\alpha^2 E[G_k^2] \geq \alpha h_{ik} - \frac{3}{4}\alpha^2 [h_{ik}^2 + ch_{ik}] \\ &= \alpha h_{ik}[1 - \frac{3}{4}\alpha h_{ik} - \frac{3}{4}\alpha c] \geq \frac{3\alpha}{4} h_{ik} \quad (\text{ for } \alpha \leq 1/4c, \alpha h_{ik} < 1/16) \end{aligned}$$

Next, we consider the variance of $S(g_k)$. We bound it using a result due to Tang and See [Tang and See, 2009, Prop. 2] that states that if $|f'(x)| \leq a$, then $\mathrm{Var}(f(x)) \leq a^2 \mathrm{Var}(x)$. We see that $S'(g_k) = 1 - 1.5g_k$, and as $g_k \geq -\alpha$,

$$\begin{aligned} \mathrm{Var}[S(g_k)] &\leq (1 + 1.5\alpha)^2 \alpha^2 ch_{ik} \\ &\leq \frac{1}{4}(1 + 1.5\alpha)^2 \alpha h_{ik} \\ &\leq \frac{1}{4}(\frac{19}{16})^2 \alpha h_{ik} < 0.36\alpha h_{ik} \end{aligned}$$

Thus, $\mathrm{Var}[S(g_k)] < 0.36\alpha h_{ik} \leq 0.5E[S(g_k)]$. (The constants have not been optimized; tighter constants are possible, especially for particular values of c.) ■

The following lemma establishes the concentration result that we need.

Lemma 9 *Let r_{i1} denote the initial reputation. For any k , we must have:*

$$E[\overline{y}(r_{ik})] \leq e^{-\frac{\alpha H_{ik}}{8}} + \overline{y}(r_{i1} + \frac{3\alpha H_{ik}}{8})$$

Proof:

Recall that we denote $\sum_{t=1}^{k} h_{ik}$ as H_{ik}. First, note that the expected value of r_{ik} is at least $r_{i1} + \frac{3\alpha H_{ik}}{4}$ (by Lemma 8). We split the domain of possible values of r_{ik} into two components $[-\infty, r_{i1} + \frac{3\alpha H_{ik}}{8})$ and $[r_{i1} + \frac{3\alpha H_{ik}}{8}, \infty]$. Then, we have:

$$\begin{aligned} E[\overline{y}(r_{ik})] &= \int_{r_{ik}=-\infty}^{r_{i1}+\frac{3\alpha H_{ik}}{8}} \Pr(r_{ik})\overline{y}(r_{ik}) + \int_{r_{ik}=r_{i1}+\frac{3\alpha H_{ik}}{8}}^{\infty} \Pr(r_{ik})\overline{y}(r_{ik}) \\ &< \Pr(r_{ik} < r_{i1} + \frac{3\alpha H_{ik}}{8}) + \overline{y}(r_{i1} + \frac{3\alpha H_{ik}}{8}) \int_{r_{ik}=r_{i1}+\frac{3\alpha H_{ik}}{8}}^{\infty} \Pr(r_{ik}) \\ &\leq \Pr(r_{ik} < r_{i1} + \frac{3\alpha H_{ik}}{8}) + \overline{y}(r_{i1} + \frac{3\alpha H_{ik}}{8}) \end{aligned}$$

where in the first inequality we have used the fact that as $\bar{y}(r_{ik}) \leq 1$ in the first term and that $\bar{y}()$ is monotonically decreasing in the second.

Now, to bound the term $\Pr(r_{ik} < \frac{3\alpha H_{ik}}{8})$, we use Bennett's concentration inequality [Bennett, 1962]. Note that the gain in each round k is independent of previous and subsequent rounds. Thus, the change in reputation $r_{ik} - r_{i1}$ is the sum of independent random variables. By Lemma 8, the sum of variances of these random variables is at most $\frac{3\alpha H_{ik}}{8}$. We need to bound the probability that the sum of these variables is more than $\frac{3\alpha H_{ik}}{8}$ below its mean value. By Bennett's inequality, this is at most $e^{\frac{-3\alpha H_{ik}}{8}*0.38} \leq e^{-\frac{\alpha H_{ik}}{8}}$. ■

Theorem 4 *Let r_{i1} denote the initial reputation assigned to user i. Fix a sequence of items, and the datapoints submitted by users prior to i's report on each item. Then, the information lost from user i is bounded above by:*

$$IL_i \leq 2 + \frac{8}{\alpha} + \frac{8}{3\alpha}\log(1 + e^{-r_{i1}})$$

Proof: We begin with the upper bound on the information loss mentioned in Equation 2:

$$\mathrm{IL}_i \leq \sum_{k=1}^{M} E[\bar{y}(r_{ik})]h_{ik}$$

Let $\bar{k}$ be the lowest k such that $H_{ik} \geq 1$. Note that $H_{ik} \leq 2$ (since each user can have gain at most 1 per item), and $\bar{y}(r_{ik}) \leq 1$. Then, accounting for all information up to round $\bar{k}$ as lost, we have

$$\mathrm{IL}_i \leq 2 + \sum_{\bar{k}}^{M} E[\bar{y}(r_{ik})]h_{ik}$$

Now, consider a given $k \geq \bar{k}$. By lemma 9, this can be bounded by the sum:

$$\mathrm{IL}_i \leq 2 + \sum_{k=\bar{k}}^{M} e^{-\frac{\alpha H_{ik}}{8}} h_{ik} + \sum_{k=\bar{k}}^{M} \bar{y}(r_{i1} + \frac{3\alpha H_{ik}}{8})h_{ik}$$

In each of the two sums, the terms are monotonically decreasing. Moreover, the maximum h_{ik} is 1. Let $\overline{H}$ be the value of H_{ik} at $\bar{k}$. Thus, we can bound the sums by integrals, as:

$$\mathrm{IL}_i \leq 2 + \int_{H=\overline{H}}^{\infty} e^{-\frac{\alpha(H-1)}{8}} dH + \int_{H=\overline{H}}^{\infty} \bar{y}(r_{i1} + \frac{3\alpha(H-1)}{8})dH$$

The two integrals have closed-form solutions:

$$\int e^{-\frac{\alpha(H-1)}{8}} dH = \frac{-8}{\alpha} e^{-\frac{\alpha(H-1)}{8}}$$

and

$$\int \bar{y}(r_{i1} + \frac{3\alpha(H-1)}{8})dH = \frac{8}{3\alpha} \log\left[y(r_{i1} + \frac{3\alpha(H-1)}{8})\right]$$

Substituting these functions, and setting the range of the integrals to $(0, \infty)$, we get:

$$\mathrm{IL}_i \leq 2 + \frac{8}{\alpha} + \frac{8}{3\alpha}\log(1 + e^{-r_{i1}})$$

■

A.3 Proofs for Section 5

Before proving Theorem 5, we first prove the following straightforward but useful lemma:

Lemma 10 *Let $P_0^* \in \mathcal{F}^*$ denote an initial distribution. Suppose that $\boldsymbol{\phi}(x)$ denote an observation of sufficient statistics, distributed according to a distribution $\pi(x)$. Then, let P_1^* denote the posterior hyperdistribution after conditioning on x; we must have the corresponding natural parameter $\mathbf{b}_1 = \mathbf{b}_0 + (\boldsymbol{\phi}(x), 1)$. The given P_0^* and π must be such that P_0^* is unbiased with respect to π: For any $\boldsymbol{\beta}$, we must have $P_0^*(\boldsymbol{\beta}) = E_\pi P_1^*(\boldsymbol{\beta})$.*

For any vector $\mathbf{a}$ of the same dimension as $\boldsymbol{\beta}^*$, the variance of $\mathbf{a}\cdot\boldsymbol{\beta}^*$ at P_0^* is at least as high as the expected variance at P_x^*:

$$Var_{P_0^*}(\mathbf{a}\cdot\boldsymbol{\beta}^*) \geq E_x\left[Var_{P_x^*}(\mathbf{a}\cdot\boldsymbol{\beta}^*)\right]$$

Proof: The condition that P_0^* is unbiased implies that we can treat P_0^* as a joint distribution over $\boldsymbol{\phi}(x)$ and $\boldsymbol{\beta}^*$: $P_0^*(\boldsymbol{\phi}(x), \boldsymbol{\beta}^*) \overset{\text{def}}{=} \pi(x)P_x^*(\boldsymbol{\beta}^*)$ has marginal distribution $P_0^*(\boldsymbol{\beta}^*)$ on $\boldsymbol{\beta}^*$.

Now, treating $\mathbf{a}\cdot\boldsymbol{\beta}^*$ and $\boldsymbol{\phi}(x)$ as two random variables, we can use the standard result from probability theory (see, e.g., [Gut, 1995, p.39]) on the conditional variance:

$$\mathrm{Var}_{P_0^*}(\mathbf{a}\cdot\boldsymbol{\beta}^*) = E_x\mathrm{Var}(\mathbf{a}\cdot\boldsymbol{\beta}^*|\boldsymbol{\phi}(x)) + \mathrm{Var}_x[E(\mathbf{a}\cdot\boldsymbol{\beta}^*|\boldsymbol{\phi}(x))]$$

The second term on the right hand side is non-negative, so we get:

$$\mathrm{Var}_{P_0^*}(\mathbf{a}\cdot\boldsymbol{\beta}^*) \geq E_x\mathrm{Var}(\mathbf{a}\cdot\boldsymbol{\beta}^*|\boldsymbol{\phi}(x)) = E_x\left[\mathrm{Var}_{P_x^*}(\mathbf{a}\cdot\boldsymbol{\beta}^*)\right]$$

■

The next ingredient of the proof of Theorem 5 is to express the KL-divergence induced by $\mathbf{a}$ in terms of an integral over variances. This differential relationship is implicit in the literature on exponential families, but we include a self-contained proof for clarity and completeness:

Lemma 11 *Given any two distributions $P, \tilde{P} \in \mathcal{F}^*$ such that $\mathbf{b}_{\tilde{P}} = \mathbf{b}_P + \mathbf{a}$, the KL divergence can be expressed as follows:*

$$K(P||\tilde{P}) = \int_{t=0}^{1}\int_{u=0}^{t} Var_{P_u^*}[\mathbf{a}\cdot\boldsymbol{\beta}^*]du dt$$

(Here, P_u^* denotes the distribution with natural parameter coordinates $\mathbf{b}_{P_u^*} = \mathbf{b}_P + u\mathbf{a}$.)

Proof: We can prove this result by repeated differentiation:

$$\begin{aligned}
\frac{d}{du}K(P||P_u^*) &= -\frac{d}{du}\int_{\boldsymbol{\beta}^*} P(\boldsymbol{\beta}^*)\log P_u^*(\boldsymbol{\beta}^*)d\boldsymbol{\beta}^* \\
&= -\int_{\boldsymbol{\beta}^*} P(\boldsymbol{\beta}^*)\frac{d}{du}[\log P_u^*(\boldsymbol{\beta}^*)]d\boldsymbol{\beta}^* \\
&= -\int_{\boldsymbol{\beta}^*} P(\boldsymbol{\beta}^*)\frac{d}{du}[\mathbf{b}_{P_u^*}\cdot\boldsymbol{\beta}^* - \psi^*(P_u^*)] \\
&= -\int_{\boldsymbol{\beta}^*} P(\boldsymbol{\beta}^*)[\mathbf{a}\cdot\boldsymbol{\beta}^*]d\boldsymbol{\beta}^* + \frac{d}{du}\psi^*(P_u^*) \\
&= -\mathbf{a}\cdot\mathbf{m}_P + \mathbf{a}\cdot\mathbf{m}_{P_u^*}
\end{aligned}$$

In the last step, we used the definition of $\mathbf{m}_P$ and the well-known fact that the gradient of $\psi^*_{P_u^*}$ is $\mathbf{m}_{P_u^*}$.

Differentiating a second time, we get:

$$\begin{aligned}\frac{d^2}{du^2}K(P||P_u^*) &= \frac{d}{du}[-\mathbf{a}\cdot\mathbf{m}_P+\mathbf{a}\cdot\mathbf{m}_{P_u^*}] \\ &= \mathbf{a}\cdot\frac{d}{du}\mathbf{m}_{P_u^*}\end{aligned}$$

Now, we expand $\mathbf{m}_{P_u^*}$ by definition:

$$\begin{aligned}\mathbf{a}\cdot\frac{d}{du}\mathbf{m}_{P_u^*} &= \mathbf{a}\cdot\frac{d}{du}\int_{\boldsymbol{\beta}^*}\boldsymbol{\beta}^*P_u^*(\boldsymbol{\beta}^*)d\boldsymbol{\beta}^* \\ &= \mathbf{a}\cdot\int_{\boldsymbol{\beta}^*}\boldsymbol{\beta}^*\frac{d}{du}P_u^*(\boldsymbol{\beta}^*)d\boldsymbol{\beta}^*\end{aligned}$$

By definition of $P_u^*(\boldsymbol{\beta}^*)=\text{Exp}[\mathbf{b}_{P_u^*}.\boldsymbol{\beta}^*-\psi^*(P_u^*)]$, we have:

$$\frac{d}{du}P_u^*(\boldsymbol{\beta}^*)=P_u^*(\boldsymbol{\beta}^*).[\frac{d}{du}\mathbf{b}_{P_u^*}.\boldsymbol{\beta}^*-\frac{d}{du}\psi^*(P_u^*)]=P_u^*(\boldsymbol{\beta}^*)\cdot[\mathbf{a}\cdot\boldsymbol{\beta}^*-\mathbf{a}.\mathbf{m}_{P_u^*}]$$

Thus,

$$\begin{aligned}\mathbf{a}\cdot\frac{d}{du}\mathbf{m}_{P_u^*} &= \mathbf{a}\cdot\int_{\boldsymbol{\beta}^*}\boldsymbol{\beta}^*[\mathbf{a}\cdot\boldsymbol{\beta}^*-\mathbf{a}\cdot\mathbf{m}_{P_u^*}]P_u^*(\boldsymbol{\beta}^*)d\boldsymbol{\beta}^*) \\ &= \int_{\boldsymbol{\beta}^*}(\mathbf{a}\cdot\boldsymbol{\beta}^*)^2P_u^*(\boldsymbol{\beta}^*)d\boldsymbol{\beta}^*-(\mathbf{a}\cdot\mathbf{m}_{P_u^*})\int_{\boldsymbol{\beta}^*}\mathbf{a}\cdot\boldsymbol{\beta}^*d\boldsymbol{\beta}^* \\ &= \int_{\boldsymbol{\beta}^*}(\mathbf{a}\cdot\boldsymbol{\beta}^*)^2P_u^*(\boldsymbol{\beta}^*)d\boldsymbol{\beta}^*-(\mathbf{a}\cdot\mathbf{m}_{P_u^*})^2\end{aligned}$$

Finally, observing that $E_{P_u^*}(\mathbf{a}\cdot\boldsymbol{\beta}^*)=\mathbf{a}\cdot\mathbf{m}_{P_u^*}$, the RHS is observed to be, by definition, $\text{Var}_{P_u^*}(\mathbf{a}\cdot\boldsymbol{\beta}^*)$.

Integrating, and observing that the LHS is 0 when $t=0$, we have:

$$[-\mathbf{a}\cdot\mathbf{m}_P+\mathbf{a}\cdot\mathbf{m}_{P_t^*}]=\int_{u=0}^{t}\text{Var}_{P_u^*}(\mathbf{a}\cdot\boldsymbol{\beta}^*)du$$

Integrating a second time, and again observing that $K(P||P_t^*)=0$ when $t=0$, we have:

$$K(P||\tilde{P})=\int_{t=0}^{1}\int_{u=0}^{t}\text{Var}_{P_u^*}[\mathbf{a}\cdot\boldsymbol{\beta}^*]dudt$$

■

Now, we can return to the proof of Theorem 5:

Theorem 5 (damage reduction property): *Let $P_0^*\in\mathcal{F}^*$ denote an initial distribution. Suppose that $\boldsymbol{\phi}(x)$ denotes an observation of sufficient statistics, distributed according to a distribution $\pi(x)$. Then, if P_x^* denotes the posterior hyperdistribution after conditioning on $\boldsymbol{\phi}(x)$; we must have $\mathbf{b}_{P_x^*}=\mathbf{b}_{P_0^*}+(\boldsymbol{\phi}(x),1)$. The given P_0^* and π must be such that P_0^* is unbiased with respect to π: For any $\boldsymbol{\beta}$, we must have $P_0^*(\boldsymbol{\beta})=E_\pi P_x^*(\boldsymbol{\beta})$. Consider a vector $\mathbf{a}$ of "attack" securities. The entries of $\mathbf{a}$ may be positive or negative. Let $\tilde{P}_0$ denote the hyperdistribution if the attack is carried out on the prior: $\mathbf{a}_0=\mathbf{b}_0+\mathbf{a}$. Likewise, let $\tilde{P}_1$ denote the hyperdistribution with natural parameter coordinates $\mathbf{a}_1=\mathbf{b}_{P_x^*}+\mathbf{a}$.*

Then, the following condition must hold:

$$K(P_0^*||\tilde{P}_0)\geq E_\pi\left[K(P_1^*||\tilde{P}_1)\right] \tag{9}$$

Proof: The proof follows from a careful application of Lemma 10 to the decomposition given in Lemma 11. We seek to prove:

$$K(P^*_0||\tilde{P}_0) \geq E_\pi \left[K(P^*_x||\tilde{P}_x)\right]$$

By Lemma 11, this is equivalent to proving:

$$\int_{t=0}^{1}\int_{u=0}^{t} \mathrm{Var}_{P^*_u}[\mathbf{a}\cdot\boldsymbol{\beta}^*]dudt \geq E_\pi \int_{t=0}^{1}\int_{u=0}^{t} \mathrm{Var}_{P^*_{xu}}[\mathbf{a}\cdot\boldsymbol{\beta}^*]dudt$$

where P^*_{xu} is the distribution with $\mathbf{b}_{P^*_{xu}} = \mathbf{b}_{P^*_x} + u\mathbf{a}$.

It is therefore sufficient to prove that, for every $u \in (0,1)$,

$$\mathrm{Var}_{P^*_u}[\mathbf{a}\cdot\boldsymbol{\beta}^*] - E_\pi \mathrm{Var}_{P^*_{xu}}[\mathbf{a}\cdot\boldsymbol{\beta}^*] \geq 0 \tag{10}$$

Consider any fixed value of u. Based on the conjugate prior nature of $\mathcal{F}^*$, if we started with prior belief P^*_0 and observed a value $\mathbf{a}$ with weight u, the posterior distribution would be P^*_u. Then, conditioning P^*_u on a further observation of $\boldsymbol{\phi}(x)$ would yield the posterior distribution P^*_{xu}, because P^*_{xu} is the distribution obtained by conditioning P^*_0 on observing $\boldsymbol{\phi}x$ with weight j and $\mathbf{a}$ with weight u. Here, we have used the property of Bayesian updating that the order of observation does not affect the final posterior. We can also verify that

$$\begin{aligned} P^*_u(\boldsymbol{\beta}^*) &= P^*_0(\boldsymbol{\beta}^*|u\mathbf{a}\text{ observed}) = \int_x P^*_0(\boldsymbol{\beta}^*, \boldsymbol{\phi}x|u\mathbf{a}\text{ observed}) \\ &= \int_x \pi(x)P^*_0(\boldsymbol{\beta}^*|u\mathbf{a}\text{ observed}, \boldsymbol{\phi}_x observed) = \int_x \pi(x)P^*_{xu}(\boldsymbol{\beta}^*) \end{aligned}$$

Thus, the conditions of Lemma 10 are satisfied; using this result, we have that, for every value of u, equation 10 is satisfied. ■

In other words: the damage induced by a fixed vector of securities $\mathbf{a}$ purchased at the initial distribution P^*_0 is greater than the expected error of the same vector of securities after an additional informative observation x.

The following corollary shows that a similar result holds for the cost of purchasing the attack securities.

Corollary 12 *Under the same conditions of theorem 5, we must have:*

$$\psi^*(\tilde{P}_0) - \psi^*(P^*_0) \geq E_\pi \left[\psi^*(\tilde{P}_x) - \psi^*(P^*_x)\right]$$

Proof: By equation 1,

$$\psi^*(\tilde{P}_0) - \psi^*(P^*_0) = K(P^*_0||\tilde{P}_0) + (\mathbf{b}_{\tilde{P}_0} - \mathbf{b}_{P^*_0})\cdot\mathbf{m}_{P^*_0} = K(P^*_0||\tilde{P}_0) + \mathbf{a}\cdot\mathbf{m}_{P^*_0}$$

Likewise,

$$\psi^*(\tilde{P}_x) - \psi^*(P^*_x) = K(P^*_x||\tilde{P}_x) + \mathbf{a}\cdot\mathbf{m}_{P^*_x}$$

Thus, it is sufficient to prove that $\mathbf{a}\cdot\mathbf{m}_{P^*_0} = E_\pi[\mathbf{a}\cdot\mathbf{m}_{P^*_x}]$; the corollary will then follow from Theorem 5.

$$E_\pi \mathbf{a}\cdot\mathbf{m}_{P^*_x} = E_\pi E_{P^*_x}[\mathbf{a}\cdot\boldsymbol{\beta}^*]$$

By the condition on P^*_0, for any random variable Z, $P^*_0(Z=z) = \int_x \pi(x)P^*_x(Z=z)dx \Rightarrow E_{P^*_0}(Z) = E_\pi E_{P^*_x}(Z)$, and hence $E_{P^*_0}[\mathbf{a}\cdot\boldsymbol{\beta}^*] = E_\pi E_{P^*_x}[\mathbf{a}\cdot\boldsymbol{\beta}^*]$. ■

Let $\Delta_i \overset{\text{def}}{=} G(u_i, \tilde{\mathbf{x}}, X_k) - G(u_i, \tilde{\mathbf{x}}', X_k)$.

Lemma 6 $\forall i \Delta_i \geq 0$

Proof: The proof is a careful application of the error-reduction property (Theorem 5). Consider any i. Then, construct a new attack $\mathcal{A}_A$ from $\mathcal{A}$, defined as

$$\tilde{\mathbf{x}}_A(\mathbf{x}) = x_1 x_2 \cdots x_{i-1} \bar{x}_1 d_1 \bar{x}_2 d_2 .. \bar{x}_{i-1} d_{i-1} x_i \bar{x}_i d_i x_{i+1} \bar{x}_{i+1} d_{i+1} \cdots x_n \bar{x}_n d_n$$

Further, construct a new attack $\mathcal{A}_B$ from $\mathcal{A}$, defined as

$$\tilde{\mathbf{x}}_B(\mathbf{x}) = x_1 x_2 \cdots x_{i-1} x_i \bar{x}_1 d_1 \bar{x}_2 d_2 .. \bar{x}_{i-1} d_{i-1} \bar{x}_i d_i x_{i+1} \bar{x}_{i+1} d_{i+1} \cdots x_n \bar{x}_n d_n$$

As the gain $G(u_i, \tilde{\mathbf{x}}_A, X_k)$ depends only on the sum of weighted datapoints up to x_i, and this sequence is a reordering of the corresponding sequence in $\tilde{\mathbf{x}}$, we must have $G(u_i, \tilde{\mathbf{x}}_A, X_k) = G(u_i, \tilde{\mathbf{x}}, X_k)$.

Now, treat the sequence of ratings $\bar{x}_1 d_1 \bar{x}_2 d_2 ... \bar{x}_{i-1} d_{i-1}$ as a single entity, and let $\mathbf{a}$ denote the aggregate update implied by this series of datapoints. Note that $\tilde{\mathbf{x}}_A$ and $\tilde{\mathbf{x}}_B$ differ only in the relative position of the update $\mathbf{a}$ and the rating x_i. Further, the datapoints in $\mathbf{a}$ are independent of x_i, because they come before x_i in $\mathcal{A}$. Based on the assumption that the prior is accurate, and observing that the updates corresponding to $x_1, x_2, ..x_i$ represent accurate Bayesian updating, we find that the conditions of Theorem 5 are satisfied. Thus, we conclude that $G(u_i, \tilde{\mathbf{x}}_A, X_k) \geq G(u_i, \tilde{\mathbf{x}}_B, X_k)$.

As the gain $G(u_i, \tilde{\mathbf{x}}_B, X_k)$ depends only of the sum of weighted datapoints up to x_i, and this sequence is the same as the corresponding sequence in $\tilde{\mathbf{x}}'$, we must have $G(u_i, \tilde{\mathbf{x}}_B, X_k) = G(u_i, \tilde{\mathbf{x}}', X_k)$. Combining all the conditions gives $G(u_i, \tilde{\mathbf{x}}, X_k) \geq G(u_i, \tilde{\mathbf{x}}', X_k)$, and thus $\Delta_i \geq 0$. ■

Next, we need to tackle the gains of the influence limit and attack datapoints. We define the *unaccounted regret* $\mathrm{UR}_k(\mathcal{A})$ as the regret adjusted by the change in attackers' gains:

$$\mathrm{UR}_k(\mathcal{A}) = \mathrm{Reg}_k(\mathcal{A}) + E_{\{r_{ik}\}} E_{\mathbf{x} \sim T_k} \sum_{i \notin \mathcal{H}} G(a_i, \tilde{\mathbf{x}}, X_k)$$

$\mathrm{UR}_k(\mathcal{A}')$ is defined likewise. Intuitively, the unnacounted regret is of interest because the remainder of the regret – the attackers' gains – can be bounded in terms of the attackers' initial budgets, as in Theorem 3. We note that in general $\mathrm{UR}_k(\mathcal{A}') \neq \mathrm{UR}_k(\mathcal{A})$, because the gains attributed to each attack identity are different in sequence $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{x}}'$.

Lemma 13

$$UR_k(\mathcal{A}) = -E_{\{r_{ik}\}} E_{\mathbf{x} \sim T_k} \sum_{i \in \mathcal{H}} [G(\bar{u}_i, \tilde{\mathbf{x}}, X_k) + \Delta_i]$$

Proof: Fix a particular value of the reputations $\{r_{ik}\}$ and an honest data vector $\mathbf{x}$.

Note that, as $\mathrm{Reg}_k(\mathcal{A}) = \mathrm{Reg}_k(\mathcal{A}')$, and given the definition of $\mathrm{Reg}_k(\mathcal{A}')$, we have:

$$\mathrm{Reg}_k(\mathcal{A}) = \mathrm{Reg}_k(\mathcal{A}') \quad = \quad -E_{\{r_{ik}\}} E_{\mathbf{x} \sim T_k} \left[\sum_{i \in \mathcal{H}} G(\bar{u}_i, \tilde{\mathbf{x}}', X_k) + \sum_{i \notin \mathcal{H}} G(a_i, \tilde{\mathbf{x}}', X_k) \right]$$

From definition and the above,

$$\mathrm{UR}_k(\mathcal{A}) \quad = \quad -E_{\{r_{ik}\}} \quad E_{\mathbf{x} \sim T_k} \quad \left[\sum_{i \in \mathcal{H}} G(\bar{u}_i, \tilde{\mathbf{x}}', X_k) + \sum_{i \notin \mathcal{H}} G(a_i, \tilde{\mathbf{x}}', X_k) - \sum_{i \notin \mathcal{H}} G(a_i, \tilde{\mathbf{x}}, X_k) \right]$$

Further, from the fact that the total expected gain is the same under $\mathcal{A}$ and $\mathcal{A}'$, we have:

$$\begin{aligned} E_{\{r_{ik}\}} \quad & E_{\mathbf{x}\sim T_k}\left[\sum_{i\in\mathcal{H}} G(u_i,\tilde{\mathbf{x}}',X_k) + \sum_{i\in\mathcal{H}} G(\bar{u}_i,\tilde{\mathbf{x}}',X_k) + \sum_{i\notin\mathcal{H}} G(a_i,\tilde{\mathbf{x}}',X_k)\right] \\ = \quad & E_{\{r_{ik}\}} E_{\mathbf{x}\sim T_k}\left[\sum_{i\in\mathcal{H}} G(u_i,\tilde{\mathbf{x}},X_k) + \sum_{i\in\mathcal{H}} G(\bar{u}_i,\tilde{\mathbf{x}},X_k) + \sum_{i\notin\mathcal{H}} G(a_i,\tilde{\mathbf{x}},X_k)\right] \end{aligned}$$

Rearranging, we have

$$\begin{aligned} E_{\{r_{ik}\}} E_{\mathbf{x}\sim T_k} & \left[\sum_{i\in\mathcal{H}} G(\bar{u}_i,\tilde{\mathbf{x}}',X_k) + \sum_{i\notin\mathcal{H}} G(a_i,\tilde{\mathbf{x}}',X_k) - \sum_{i\notin\mathcal{H}} G(a_i,\tilde{\mathbf{x}},X_k)\right] \\ & = E_{\{r_{ik}\}} E_{\mathbf{x}\sim T_k} \sum_{i\in\mathcal{H}} \left[G(\bar{u}_i,\tilde{\mathbf{x}},X_k) + \Delta_i\right] \end{aligned}$$

The last equality follows from the definition of Δ_i. The LHS as we have shown above is $\mathrm{UR}_k(\mathcal{A})$. ■

Next, we bound the total gain of the influence limiting identities $\bar{x}_i$ by relating them to the gain of the honest entities x_i.

Lemma 14 *Fix reputation values, and hence influence values y_{ik}. Then, for each i,*

$$E_{\mathbf{x}\sim T_k} G(\bar{u}_i,\tilde{\mathbf{x}},X_k) \geq -\bar{y}_{ik}\left[E_{\mathbf{x}\sim T_k} G(u_i,\tilde{\mathbf{x}}',X_k) + \Delta_i\right]$$

Proof: Consider the sequence $\tilde{\mathbf{x}}$. Recall that the consecutive pair of "reports" $x_i\bar{x}_i$ is a model for the influence-limited report by agent i. By the concavity of the gain function, we have:

$$G(u_i,\tilde{\mathbf{x}},X_k) + G(\bar{u}_i,\tilde{\mathbf{x}},X_k) \geq y_{ik} G(u_i,\tilde{\mathbf{x}},X_k)$$

This implies that $G(\bar{u}_i,\tilde{\mathbf{x}},X_k) \geq -\bar{y}_{ik} G(u_i,\tilde{\mathbf{x}},X_k)$.

By definition of Δ_i, we have $E_{\mathbf{x}\sim T_k} G(u_i,\tilde{\mathbf{x}},X_k) = E_{\mathbf{x}\sim T_k} G(u_i,\tilde{\mathbf{x}}',X_k) + \Delta_i$, thus completing the proof. ■

Putting together Lemma 13 and Lemma 14, we get the following bound on $\mathrm{UR}_k(\mathcal{A})$:

Lemma 15

$$UR_k(\mathcal{A}) \leq E_{\{r_{ik}\}} \sum_{i\in\mathcal{H}} \bar{y}_{ik} E_{\mathbf{x}\sim T_k} G(u_i,\tilde{\mathbf{x}},X_k)$$

Proof: From Lemma 13 and Lemma 14, we get:

$$\mathrm{UR}_k(\mathcal{A}) \leq E_{\{r_{ik}\}} \sum_{i\in\mathcal{H}} \bar{y}_{ik} E_{\mathbf{x}\sim T_k} G(u_i,\tilde{\mathbf{x}},X_k) + \sum_{i\in\mathcal{H}} (\bar{y}_{ik}-1)\Delta_i$$

By Lemma 6, each $\Delta_i \geq 0$. Further, $\bar{y}_{ik}$ is always between 0 and 1. Thus, the last sum is always ≤ 0 and the lemma statement follows. ■ By definition, we have $\mathrm{Reg}_k(\mathcal{A}) = \mathrm{UR}_k(\mathcal{A}) - E_{\mathbf{x}\sim T_k} \sum_{i\notin\mathcal{H}} G(a_i,\tilde{\mathbf{x}},X_k)$. To prove the final regret bound, we sum this up over all k.

Theorem 7 *The regret of the conjugate-prior algorithm is bounded by:*

$$D\left\{\sum_{i\in\mathcal{H}}\left[2+\frac{8}{\alpha}+\frac{8}{3\alpha}\log(1+e^{-r_{i1}})\right]+\sum_{i\notin\mathcal{H}}\frac{1}{\alpha}B(r_{i1})\right\}$$

Proof: Consider any given attack $\mathcal{A}$. By lemma 15, the regret in round k is bounded by:

$$\mathrm{Reg}_k(\mathcal{A}) \leq E_{\{r_{ik}\}} \sum_{i\in\mathcal{H}} \bar{y}_{ik} E_{\mathbf{x}\sim T_k} G(u_i, \tilde{\mathbf{x}}, X_k) + E_{\mathbf{x}\sim T_k} \sum_{i\notin\mathcal{H}} (-G(a_i, \tilde{\mathbf{x}}, X_k))$$

Summing over k from 1 to M, we have:

$$\mathrm{Reg}(\mathcal{A}) \leq \sum_k E_{\{r_{ik}\}} \sum_{i\in\mathcal{H}} \bar{y}_{ik} E_{\mathbf{x}\sim T_k} G(u_i, \tilde{\mathbf{x}}, X_k) + \sum_k \sum_{i\notin\mathcal{H}} E_{\mathbf{x}\sim T_k} (-G(a_i, \tilde{\mathbf{x}}, X_k))$$

By Theorem 4, and using un-scaled log loss, the first sum is bounded by

$$\sum_{i\in\mathcal{H}} D\left[2 + \frac{8}{\alpha} + \frac{8}{3\alpha}\log(1+e^{-r_{i1}})\right]$$

By Theorem 3, and using un-scaled log loss, the second sum is bounded by $\sum_{i\notin\mathcal{H}} \frac{D}{\alpha} B(r_{i1})$. Thus, putting these two together, we have:

$$\mathrm{Reg}(\mathcal{A}) \leq D\left\{\sum_{i\in\mathcal{H}} \left[2 + \frac{8}{\alpha} + \frac{8}{3\alpha}\log(1+e^{-r_{i1}})\right] + \sum_{i\notin\mathcal{H}} \frac{1}{\alpha} B(r_{i1})\right\}$$

■

B Notation

For convenience, the following tables indicate notations for the key terms we use.

Symbol	Meaning
$\mathcal{H}$	set of honest agents
$\overline{\mathcal{H}}$	set of dishonest/adversarial agents
N	total number of agents: $\lvert\mathcal{H}\rvert + \lvert\overline{\mathcal{H}}\rvert = N$
n	number of honest agents: $\lvert\mathcal{H}\rvert = n$
M	number of items
X_k	realized value of each of the kth item
k	item index
P_k	distribution over X for item $k \in \{1, 2, \ldots, M\}$
$\mathcal{F}$	exponential family of distributions; $P_k \in \mathcal{F}$
$\mathcal{F}^*$	set of possible hyperdistributions; this is also assumed to be exponential family
$\boldsymbol{\phi}$	sufficient statistics that are a function of the input $x \in X$; $\lvert\boldsymbol{\phi}\rvert = t$
$\boldsymbol{\beta}$	natural parameters of the distribution on X; $\lvert\boldsymbol{\beta}\rvert = t$
P_0^*	prior distribution on $\boldsymbol{\beta}$ known as a hyperdistribution with parameter $\mathbf{b}_0$
$P_{\mathbf{b}}^*$	hyperdistribution with natural parameter $\mathbf{b}$; $P_{\mathbf{b}}^* \in \mathcal{F}^*$ and $P_{\mathbf{b}_j}^*$ may be simply written as P_j^*
$\mathbf{b}$	natural parameter of hyperdistribution $P_{\mathbf{b}}^*$
$\boldsymbol{\beta}^*$	sufficient statistics of the hyperdistribution over $\boldsymbol{\beta}$
ψ^*	log partition function of the hyperdistribution
ψ	log partition function of the distribution over X
$\mathbf{x}$	sequence of honest datapoints
T_k	set of all possible honest sequences represented as a tree
$\mathbf{x} \sim T_k$	shorthand for $\mathbf{x}$, the eventual outcome X_K, and true parameter $\boldsymbol{\beta}_k$ jointly distributed according to hyperdistribution P_0^*
$\mathcal{A}$ (and $\mathcal{A}_k$)	attack strategy (for item k)
$\hat{\boldsymbol{\beta}}$	natural parameter of the forecast distribution Q_k

Symbol	Meaning
i	agent index
u_i	honest agent; $u_i \in \mathcal{H}$
$\overline{u}_i$	the 'attack part' of the honest agent where $u_i \in \mathcal{H}$
a_i	dishonest agent/attacker; $a_i \in \overline{\mathcal{H}}$
$\tilde{\mathbf{x}}_i$	attack datapoints corresponding to the ith attacker
$\tilde{\mathbf{x}}$	the entire realized sequence of datapoints with honest data $\mathbf{x}$ and attack strategy $\mathcal{A}_k$; shorthand for $\tilde{\mathbf{x}}_{\mathcal{A}_k}(\mathbf{x})$
Q	prediction; a function of $\tilde{\mathbf{x}}$, typically parametrized by Z and k
Z and Z_O	learning algorithms; Z_O denotes an omniscient learning algorithm that is used a benchmark
$L(Q_k, \cdot)$	second parameter depends on whether revealed data/parameter
$G_Z(i, \tilde{\mathbf{x}}, X_k)$	incremental gain for an agent i in sequence $\tilde{\mathbf{x}}$, given an eventual true outcome X_k when it is obvious from context, this may be written as G_{ik}
y_{ik}	influence of agent i on item k
$\hat{G}_{ik} = (1/y_{ik})G_{ik}$	the scaled gain that is at least as high as the actual gain $G(1, i, \tilde{\mathbf{x}}, X_k)$
h_{ik}	informativeness of agent i on item k
$\mathbf{a}$	attack securities in the conjugate-prior market
$\tilde{P}_j$	posterior hyperdistribution conditioned on attack securities with prior P_j^* in the conjugate-prior market
$\mathrm{Reg}(Z)$	regret of algorithm Z

Family	Market State or natural parameters	Securities or sufficient statistics	Cost function
$\mathcal{F}$	$\boldsymbol{\beta}$	$\boldsymbol{\phi}(x)$	$\psi(x)$
$\mathcal{F}^*$	$\mathbf{b}$	$\boldsymbol{\beta}^* = (\boldsymbol{\beta}, -\psi(\boldsymbol{\beta}))$	$\psi^*(\boldsymbol{\beta})/D$