Safety in Numbers: Asymptotic Analysis of a Monitoring Problem

Reka Inovan, Emre Telatar
LTHI, EPFL
Lausanne, Switzerland

***Abstract*—In this work, we introduce a setup where a monitoring entity attempts to distinguish a cheating player among a group of regular players where all players behave in order to maximize their reward. We assume that the cheating player has an "information advantage" compared to the regular players. However, greedily exploiting this advantage will lead to the cheating player being easily distinguishable from its peers. Hence there is a tension between exploitation of the said advantage and the probability of being caught. We characterize this trade-off showing that the cheating player can obtain a higher reward as the number of regular players grows. We also show that, under a certain regime, a monitoring strategy based on the empirical divergence function attains the same normalized reward as the minimax reward.**

***Index Terms*—Monitoring Problem, Empirical Divergence, Minimax Reward, Method of Type.**

I. Monitoring Problem

Consider a scenario where the state of an environment X is observed by $M+1$ players. Among these players, there are M regular players and a cheating player. The index of the cheating player will be denoted by H which is distributed uniformly in $\{0,\dots,M\}$.

Let us denote the players' observations as $Z[0],\dots,Z[M]$. For the i-th player, conditioned on X and H, the observation $Z[i]$ is independent of $\{Z[j]\}_{j\neq i}$. These observations are noisy. For the regular players, i.e., $i\neq H$, these observations are obtained through identical channels characterized by $V_{Z^{(r)}|X}$. We assume that the cheating player benefits from a certain "information advantage", hence its observation is obtained through a different channel $V_{Z^{(c)}|X}$. We denote the distribution of X as V_X. We use superscripts (c) and (r) to differentiate the random variables of the cheating player and the regular player.

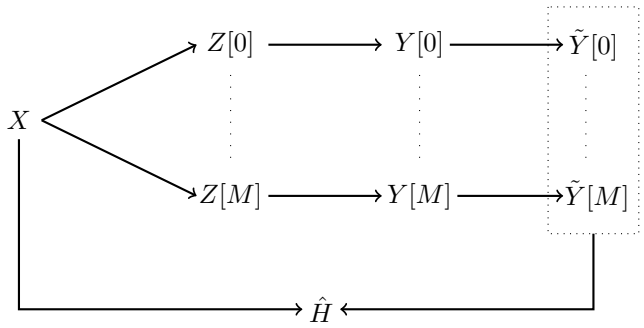
Each player uses its observation $Z[i]$ as a basis to take an action $Y[i]$. In this model, the action of the i-th player can only depend on $Z[i]$. This allows us to characterize the players' policies as (possibly probabilistic) mappings f_i for each player, so that $Y[i]=f_i(Z[i])$.

A player obtains a reward from its action depending on the state of the environment. The reward that the players obtain is determined by a bounded reward function $R(X,Y)$. All players, regular and cheating, have the same reward function. We assume that $\sup_{f^{(r)}}\mathbb{E}[R(X,Y^{(r)})]\leq\sup_{f^{(c)}}\mathbb{E}[R(X,Y^{(c)})]$, i.e., the cheating player can obtain a greater expected reward if it is allowed to act without any constraint.

The cheating player is constrained by a monitoring entity who gets to observe X and noisy observations of the players' actions. Let us denote the monitor's observation of the i-th player action as $\tilde{Y}[i]$. We also assume that conditioned on $Y[i]$, the observation $\tilde{Y}[i]$ is independent of $\{Z[j]\}_{j=0}^{M},\{Y[j]\}_{j\neq i}^{M}$ and $\{\tilde{Y}[j]\}_{j\neq i}^{M}$. Conditioned on the i-th player action $Y[i]$, the monitor's observation is distributed according to $V_{\tilde{Y}[i]|Y[i]}$. Based on $(X,\{\tilde{Y}[i]\}_{i=0}^{M})$, the monitor gives a guess of the cheating player's index. Let us denote this guess as $\hat{H}$.

We assume that the regular players disregard the possibility of being falsely accused. Hence the reward of a regular player is independent of $\hat{H}$ and is given by $\mathbb{E}[R(X,Y^{(r)})]$. For the cheating player, it can only reap its reward if it is not detected. So its reward is given by $\mathbb{E}[\mathbb{1}\{\hat{H}\neq H\}R(X,Y[H])]$. The cheating player wants to maximize its reward; this typically requires a small probability of being caught. On the other hand, the monitoring entity aims to minimize the reward of the cheating player.

The probabilistic model can be illustrated as follows:



Let us assume that all variables have finite alphabets.

Assuming that the monitor and the cheater are rational and that they both know about the reward structure, then the expected reward will be a minimax value.

Definition 1. *We define the minimax reward of the cheating player, $R^{(c)}$, as,*

$$\min_{\hat{H}\in\hat{\mathcal{H}}}\max_{f^{(c)}}\mathbb{E}\big[\mathbb{1}\{\hat{H}\neq H\}R\big(X,Y[H]\big)\big]. \tag{1}$$

where $\hat{\mathcal{H}}$ is a set of random variables which only depends on $\{Y[i]\}_{i=0}^{M}$.

At first glance, one would think that the formulation of reward is specific to the case where the decision rule is taken

before the cheating player decides on its policy. But this is not the case. This is a consequence of Von Neumann's minimax theorem (section 17.6 in [1]).

Proposition 1. *We have*

$$\min_{\hat{H}\in\hat{\mathcal{H}}}\max_{f^{(c)}}\mathbb{E}\big[\mathbb{1}\{\hat{H}\neq H\}R\big(X,Y[H]\big)\big] = \max_{f^{(c)}}\min_{\hat{H}\in\hat{\mathcal{H}}}\mathbb{E}[\mathbb{1}\{\hat{H}\neq H\}R(X,Y[H])] \quad (2)$$

Proof. See appendix A. The fact that we allow the policy $f^{(c)}$ to be probabilistic is essential to the proof of proposition 1. □

This setup models cases where a monitoring entity wants to ensure a level playing field for number of players which are motivated to maximize their rewards. Several settings where this might arise:

- In financial markets, where a regulator wants to prevent players from benefiting from non-public information.
- In cyber security, where an intruder wants to blend in with regular users.

The main contributions of this work are as follows:

- We introduce a new framework for a monitoring problem based on a zero-sum game formulation.
- When the game is played in a parallel fashion (see section IV), we characterize the trade-off between the cheating player linear growth of reward and the exponential growth of the number of regular players. This trade-off is expressed in a single-letter form in Theorem 1.
- The technical challenges en route to Theorem 1 are mainly resolved in Propositions 6, 7, and 9. Proposition 6 reduces the space of strategies in the case of parallel play, Proposition 7 shows a threshold phenomenon on the detection probability, Proposition 9 identifies an asymptotically optimal detection rule.

II. Related Works

The minimax characterization of hypothesis testing problem has been studied since the time of Hoeffding [2]. There has been several information theoretic studies of hypothesis testing problem with an adversary [3]–[5]. The main difference between the setting of our work and the literature on universal hypothesis testing lies on our assumption that there is an reward function which the cheating player wants to maximize.

There is also a recent interest on a related problem in the context of adversarial classification, e.g., [6], [7], which shares a similar concern of an adversary behaving strategically to maximize a reward function. Our work is different in that we are mainly interested in the asymptotic behaviour of the sequences of decision strategies instead of characterizing the equilibrium for a one-shot instance.

III. One Shot Case

Before we consider the asymptotic setting, it is useful to examine the one-shot version of the problem and establish several properties of the set of decision rules and the set of cheating policies. It is easy to see that we can characterize the behaviour of a regular player as follows.

Proposition 2. *For the optimal $f^{(r)}$, we have, $f^{(r)}(z)\in\arg\max_y\mathbb{E}\big[R(X,y)\,|\,Z^{(r)}=z\big]$. where the expectation is taken with respect to the regular player probability model.*

Furthermore, even if the set $\hat{\mathcal{H}}$ is uncountable due to randomization, one can characterize a subset which attains the minimum in (1). We will refer to this subset as the metric-based decision rules.

Definition 2. *We say that $\hat{H}$ is a* metric-based *decision rule if $\hat{H}$ is deterministic and there exists a $g:(\mathcal{X}\times\tilde{\mathcal{Y}})\to\mathbb{R}_{\geq 0}$ such that $\hat{H}\big((x,(\tilde{y}[i])_{i=0}^{M})\big)\in\arg\max_i g(x,\tilde{y}[i])$. In this case, we say that $\hat{H}$ is induced by g. Let $\hat{\mathcal{H}}_m$ be the set of all* metric-based *decision rules.*

Proposition 3. *Given a deterministic policy $f^{(c)}$, there exists a $\hat{H}\in\hat{\mathcal{H}}_m$ which attains the minimum of optimization problem, $\min_{\hat{H}\in\hat{\mathcal{H}}}\mathbb{E}\big[\mathbb{1}\{\hat{H}\neq H\}R\big(X,Y[H]\big)\big]$.*

Proof. See appendix B. □

Due to proposition 3, we can define an equivalence class on the set of all possible mappings g from $\mathcal{X},\tilde{\mathcal{Y}}$ to $\mathbb{R}_{\geq 0}$, where we say that g_1 and g_2 are equivalent if they induce the same $\hat{H}$. We can also see that although the space of g is uncountable, but the set of $\mathcal{H}_m$ is of finite size.

Proposition 4. *Given a decision rule $\hat{H}$, consider a deterministic policy, $f^{(c)}$ such that*

$$f^{(c)}(z)\in\arg\max_y\mathbb{E}[\mathbb{1}\{\hat{H}\neq H\}R(X,y)|Z^{(c)}=z] \quad (3)$$

where the expectation is taken under the cheater's probabilistic model. This policy attains the maximum of optimization problem,

$$\max_{f^{(c)}}\mathbb{E}[\mathbb{1}\{\hat{H}\neq H\}R(X,Y[H])]. \quad (4)$$

I.e., one can restrict the feasible set of cheating player's policies to a set of deterministic cheating policies.

As a consequence of the previous propositions, the class of randomized metric-based decision rules and the class of deterministic cheating policy is sufficient to characterize the minimax value. We states this formally below by using $P(\hat{\mathcal{H}}_m)$ and $\mathcal{F}_d$ to denote the class of randomized metric-based decision rule and the set of deterministic mapping from Z to Y.

Proposition 5. *We have*

$$\max_{f^{(c)}}\min_{\hat{H}\in\hat{\mathcal{H}}}\mathbb{E}[\mathbb{1}\{\hat{H}\neq H\}R(X,Y[H])] = \min_{\hat{H}\in P(\hat{\mathcal{H}}_m)}\max_{f^{(c)}\in\mathcal{F}_d}\mathbb{E}[\mathbb{1}\{\hat{H}\neq H\}R(X,Y[H])]. \quad (5)$$

Proof. See appendix C. □

Remark: However, note that we cannot exchange the minimum and the maximum after we constrain the admissible sets.

This is due to the fact that the set of all possible deterministic policies is not a convex set. One must be careful to note that constraining the inner maximization only on the deterministic policy is merely a calculation device (à la theorem 17:A in [1]). In general, the optimal cheater strategy will not correspond to a deterministic policy, but it is the case that the reward of the optimal (randomized) policy is equal to the reward of a certain deterministic policy.

By only considering metric-based decision rules, we transform the problem of optimizing a decision rule $\hat{H}$ over all possible realizations of $(X, \{\tilde{Y}_i\}_{i=0}^{M})$ into an easier problem of whether the cheating player metric is larger than the maximum of regular players metric. We present a restatement of proposition 4 which will be useful in the next section.

Corollary 1. *Given a decision rule in $P(\mathcal{H}_m)$, which is composed of $(\hat{H}_1, \dots, \hat{H}_k)$ with a probability distribution $(p_1, \dots, p_k)$, the deterministic policy for $f^{(c)}$ is given by,*

$$f^{(c)}(z) \in \arg\max_{y \in \mathcal{Y}} \sum_{x} V_{X|Z^{(c)}}(x|z) R(x,y) \sum_{i=1}^{k} p_i \Pr(\tilde{g}_i(x) \le g_i(x, \tilde{Y}^{(c)}) \,|\, Y^{(c)} = y) \quad (6)$$

where g_i is a metric which induces $\hat{H}_i$, and $\tilde{g}_i(x)$ is the maximum value of the metric g_i among the regular players given the realization of X.

IV. Memoryless Channels and Product Distributions

Consider a version of the monitoring problem where $(X, Z, Y, \tilde{Y})$ are replaced by n-vectors $\underline{X}, \underline{Z}, \underline{Y}, \underline{\tilde{Y}}$. We study a special case where,

- $(\underline{X}, \underline{Z})$ is distributed according to $V_{Z|X}^{\otimes n}$ and $V_X^{\otimes n}$.
- Probabilistic mapping between $\underline{Y}$ and $\underline{\tilde{Y}}$ is given by multiple uses of memoryless channel $V_{\tilde{Y}|Y}$, i.e., it is described by $V_{\tilde{Y}|Y}^{\otimes n}$.
- Both the regular players and the cheating player get to observe the whole realization of their respective $\underline{Z}$ before forming their action.

Furthermore, we will take the number of regular players M to scale as $M = \lfloor e^{nK} \rfloor$. We also assume that the reward function is additive, i.e., $R(\underline{X}, \underline{Y}) = \sum_{i=1}^{n} R(X_i, Y_i)$.

We are interested in studying the normalized reward of the cheating player,

$$\limsup_{n\to\infty} \frac{1}{n} \max_{f^{(c)}} \min_{\hat{H} \in \hat{\mathcal{H}}} \mathbb{E}\big[\mathbb{1}\{\hat{H} \neq H\} R\big(\underline{X}, \underline{Y}[H]\big)\big]. \quad (7)$$

We need to emphasize that this setting only captures the linear growth rate of an increasingly larger single-stage game where the monitor is given more observation, as opposed to a repeated game in the game theory setting.

We will refer to the setting which fulfills these requirements as the *product*-regime. Under this regime, the regular players' random variables fulfill a permutation symmetry. Hence, it is not very surprising that we can simplify the problem by only considering a class of "permutation-invariant" $\hat{H}$ and $f^{(c)}$.

To simplify the notation, it is useful to introduce the notion of type, see [8]. For a given n and a sequence $\underline{x}$ of length n, the type $P_{\underline{x}}(\underline{x})$ is defined as the empirical distribution of the sequence $\underline{x}$. We denote the set of possible types of sequences of length n as $\mathcal{P}_n(X)$. Furthermore, for a given $P_X \in \mathcal{P}_n(X)$. We define the set of sequences with type P as T_P^n. Given a joint type $P_{X,Y}$ we define $T^n_{P_{Y|X}}(\underline{x})$ as the set of $\underline{y}$ such that $(\underline{x}, \underline{y}) \in T^n_{P_{X,Y}}$. We will omit the length if it is clear from the context.

Definition 3. *We say that $\hat{H}$ is a* permutation-invariant *metric decision rule if there exists a metric $g(\underline{X}, \underline{\tilde{Y}})$ which induces $\hat{H}$ and this metric depends on $\underline{X}, \underline{\tilde{Y}}$ only through its joint type i.e., $g(\underline{X}, \underline{\tilde{Y}}) = g(P_{X,\tilde{Y}})$ where $P_{X,\tilde{Y}}$ is the empirical probability of realization $(\underline{X}, \underline{\tilde{Y}})$. Let us denote the set of* permutation-invariant *metric decision rules as $\hat{\mathcal{H}}_T$.*

Definition 4. *For a given n, we say that $f^{(c)}$, is a* permutation-invariant *policy if for every $\underline{z}$ and type $P_{Y,Z} \in \mathcal{P}(Y \times Z)$ we have $\Pr_{\underline{Y}|\underline{Z}}(.|\underline{z})$ is uniform on $T_{P_{Y|Z}}(\underline{z})$. Let us denote the set of* permutation-invariant *policies as $\mathcal{F}_T$.*

The analogue of proposition 5 for the product-regime is given by the following proposition.

Proposition 6. *Under the* product-*regime, we have*

$$\max_{f^{(c)}} \min_{\hat{H} \in \hat{\mathcal{H}}} \mathbb{E}\big[\mathbb{1}\{\hat{H} \neq H\} R\big(\underline{X}, \underline{Y}[H]\big)\big] = \max_{f^{(c)} \in \mathcal{F}_T} \min_{\hat{H} \in \hat{\mathcal{H}}_T} \mathbb{E}[\mathbb{1}\{\hat{H} \neq H\} R(\underline{X}, \underline{Y}[H])] \quad (8)$$

Proof. See appendix D. □

Given an $\underline{x}$ each permutation-invariant metric decision rule essentially ranks each possible type $P_{X,\tilde{Y}} \in \mathcal{P}(X \times \tilde{Y})$ which is compatible with $P_{\underline{x}}$. Therefore, the probability that the monitor makes an error is equivalent to the probability that one of the realizations of $\{\underline{\tilde{Y}}[i]\}_{i \neq H}$ is such that the joint type of $(\underline{X}, \underline{\tilde{Y}}[i])$ is ranked higher than the joint type of $(X, \underline{\tilde{Y}}[H])$.

For asymptotic analysis, there is a technical subtlety that the decision metrics that the monitor chooses can depend on n. It is reasonable to worry whether the ordering of types induced by the metrics will be preserved across the values of n. Hence, we need to introduce a formulation that take this concern into account.

Consider a sequence of metrics $g_1, g_2, \dots$ such that the monitoring entity uses g_n to form its decision rule for length n.

Let us define the sets,

$$\mathcal{P}_\infty(X) = \bigcup_{i=1}^{\infty} \mathcal{P}_i(X) \qquad \mathcal{P}_\infty(X \times \tilde{Y}) = \bigcup_{i=1}^{\infty} \mathcal{P}_i(X \times \tilde{Y}). \quad (9)$$

Given any joint type $P_{X,\tilde{Y}} \in \mathcal{P}_\infty(X \times \tilde{Y})$, we can define a set

$$\mathcal{U}_n(P_{X,\tilde{Y}}) = \{P'_{X,\tilde{Y}} \in \mathcal{P}_n(X \times \tilde{Y}) \,| \, P_X = P'_X, g_n(P'_{X,\tilde{Y}}) > g_n(P_{X,\tilde{Y}})\} \quad (10)$$

if $P_{X,\tilde{Y}} \in \mathcal{P}_n(X \times \tilde{Y})$, otherwise $\mathcal{U}_n(P_X, P_{X,\tilde{Y}}) = \varnothing$. Let us define

$$\mathcal{U}(P_{X,\tilde{Y}}) = \limsup_{n\to\infty} \mathcal{U}_n(P_{X,\tilde{Y}}). \tag{11}$$

An intuitive explanation of $\mathcal{U}(P_{X,\tilde{Y}})$ is that it corresponds to "level sets" on the probability simplex in which infinitely many g_n's agrees that the distribution on this set is ranked higher than $P_{X,\tilde{Y}}$.

We also define

$$G(P_{X,\tilde{Y}}) = \inf_{P'_{X,\tilde{Y}} \in \mathcal{U}(P_{X,\tilde{Y}})} D(P'_{X,\tilde{Y}} || V_{X,\tilde{Y}^{(r)}}) \tag{12}$$

i.e., $G(P_{X,\tilde{Y}})$ is the distance between the regular player joint distribution $V_{X,\tilde{Y}^{(r)}}$ and the set $\mathcal{U}(P_{X,\tilde{Y}})$ in KL divergence.

Proposition 7. *For a sequence of metrics $g_1, g_2, \ldots$ and any distribution $P_{X,\tilde{Y}} \in \mathcal{P}_\infty(X \times \tilde{Y})$,we have*

- *if $K > G(P_{X,\tilde{Y}})$ then*

$$\limsup_{n\to\infty} \Pr\left(\hat{H} \neq H \,\middle|\, (\underline{X}, \underline{\tilde{Y}}[H]) \in T_{P_{X,\tilde{Y}^{(c)}}}\right) = 1,$$

- *if $K < G(P_{X,\tilde{Y}})$ then*

$$\limsup_{n\to\infty} -\frac{1}{n} \log \Pr\left(\hat{H} \neq H \,\middle|\, (\underline{X}, \underline{\tilde{Y}}[H]) \in T_{P_{X,\tilde{Y}^{(c)}}}\right) > 0.$$

Proof. See appendix E. □

Proposition 7 is the main technical result of this work. This proposition asserts that there is a qualitative change depending on whether $G(P_{X,\tilde{Y}})$ is smaller or larger than K. Hence, given K we can determine the set of joint types of cheater's realization for which the monitoring entity will make an error with high probability. Let us define for $P_X \in \mathcal{P}_\infty(X)$

$$G^{-1}_{(g_n)_{i=1}^\infty}(P_X, K) = \limsup_{i\to\infty} G_n^{-1} \tag{13}$$

where

$$G_n^{-1} = \{P_{X,\tilde{Y}} \in \mathcal{P}_n(X, \tilde{Y}) : P_X = P'_X \quad g_n(P'_{X,\tilde{Y}}) \leq \max_{\substack{P'_{X,\tilde{Y}}: \\ D(P'_{X,\tilde{Y}} || V_{X,\tilde{Y}^{(r)}}) < K}} g_n(P'_{X,\tilde{Y}})\}, \tag{14}$$

and

$$G^{-1}_{(g_n)_{n=1}^\infty}(K) = \bigcap_{\epsilon>0} \bigcup_{\substack{P_X \in T_X^\infty \\ |V_X - P_X| \leq \epsilon}} G^{-1}_{(g_n)_{n=1}^\infty}(P_X, K). \tag{15}$$

Using proposition 7, we can give upper and lower bound on the expected linear growth of cheating player's reward.

Proposition 8. *Given a sequence of type-invariant metrics $(g_n)_{n=1}^\infty$, we have,*

$$\limsup_{n\to\infty} \frac{1}{n} \max_{f^{(c)} \in \mathcal{F}_T} \mathbb{E}\left[\mathbb{1}\{\hat{H} \neq H\} R\left(\underline{X}, \underline{Y}[H]\right)\right] \geq \sup_{\substack{P_{Y^{(c)}|Z^{(c)}}: \\ V_{X,\tilde{Y}^{(c)}} \in \mathit{IntCl}(G^{-1}_{(g_n)_{n=1}^\infty}(K))}} \mathbb{E}[R(X, Y^{(c)})], \tag{16}$$

where $\mathit{IntCl}(G^{-1}_{(g_n)_{n=1}^\infty}(K))$ is the interior of the closure of $G^{-1}_{(g_n)_{n=1}^\infty}(K)$ under the total variation metric. We also have,

$$\limsup_{n\to\infty} \frac{1}{n} \max_{f^{(c)} \in \mathcal{F}_T} \mathbb{E}\left[\mathbb{1}\{\hat{H} \neq H\} R\left(\underline{X}, \underline{Y}[H]\right)\right] \leq \max_{\substack{P_{Y^{(c)}|Z^{(c)}}: \\ V_{X,\tilde{Y}^{(c)}} \in \mathit{Cl}(G^{-1}_{(g_n)_{n=1}^\infty}(K))}} \mathbb{E}[R(X, Y^{(c)})], \tag{17}$$

where $\mathit{Cl}(G^{-1}_{(g_n)_{n=1}^\infty}(K))$ is the closure of $G^{-1}_{(g_n)_{n=1}^\infty}(K)$ and $V_{X,\tilde{Y}^{(c)}}$ is the induced distributed on the monitoring problem given $P_{Y^{(c)}|Z^{(c)}}$.

Proof. See appendix F. □

Remark: We have to split the proposition into two cases to take into account the effect of isolated points in $\mathrm{Cl}\big(G^{-1}_{(g_n)_{n=1}^\infty}(K)\big)$. Our achievability method requires the existence of a non-empty neighborhood around the sequence which approximates the optimal point. Note that this kind of separation also exists in several large deviation principle theorems given in [9], hence it might be non-trivial to remove the influence of isolated points. However, as we discuss later on, the sequence that induces the optimal strategy will not have isolated points, hence obviating the need for this consideration.

The interpretation of $\mathcal{U}$ as level sets is useful when we try to compare the asymptotic behaviour of two sequences of decision rules. In fact this notion allows us to argue that the empirical divergence decision rule is as good as any other sequence of decision rules.

To formalize this claim, let us denote the normalized reward of the cheating player under the sequence of decision rules $(g_n)_{n=1}^\infty$ as $R((g_n)_{n=1}^\infty)$. Let us also define the sequence of empirical divergence metrics as, $g_n^{ed}(\underline{X}, \underline{\tilde{Y}}) = D(P_{\underline{X},\underline{\tilde{Y}}} || V_{X,\tilde{Y}^{(r)}})$ where $P_{\underline{X},\underline{\tilde{Y}}}$ is the type of the realization $(\underline{X}, \underline{\tilde{Y}})$.

The following proposition guarantees the asymptotic optimality of the empirical divergence decision rule.

Proposition 9. *For any sequence of metric $(g_n)_{i=1}^\infty$,*

$$R((g_n^{ed})_{n=1}^\infty) \leq R((g_n)_{n=1}^\infty). \tag{18}$$

Proof. See appendix G. □

Corollary 2. *We have*

$$R((g_n^{ed})_{n=1}^\infty) = \max_{\substack{V_{Y^{(c)}|Z^{(c)}}: \\ D(V_{X,\tilde{Y}^{(c)}} || V_{X,\tilde{Y}^{(r)}}) \leq K}} \mathbb{E}[R(X, Y^{(c)})]. \tag{19}$$

Proof. Due to the continuity of divergence function and $D(.||V_{X,\tilde{Y}^{(r)}})$ is a convex function, the closure of $G^{-1}_{(g_n^{ed})_{n=1}^\infty}(K)$ does not have isolated points. Hence the lower bound and the upper bound in proposition 8 coincide. Finally, one obtains the RHS by substituting the empirical divergence metric to the definition of $G^{-1}_{(g_n^{ed})_{n=1}^\infty}(K)$. □

Proposition 9 not only gives us the minimizing metric. It also shows that this minimizing metric is independent of the

policy of the cheating player. For the purpose of summarizing the results, we state the following theorem.

Theorem 1. *For the* product-*regime, we have,*

$$\limsup_{n\to\infty} \frac{1}{n} \max_{f^{(c)}} \min_{\hat{H}\in\hat{\mathcal{H}}} \mathbb{E}[\mathbb{1}\{\hat{H} \neq H\} R(\underline{X}, \underline{Y}[H])] = \max_{\substack{V_{Y^{(c)}|Z^{(c)}}: \\ D(V_{X,\tilde{Y}^{(c)}} || V_{X,\tilde{Y}^{(r)}}) \leq K}} \mathbb{E}[R(X, Y^{(c)})]. \quad (20)$$

This result does not necessarily imply that the empirical divergence metric induces the monitor decision strategy at Nash equilibrium for every finite value of n. But one can deduce that the difference between the normalized reward of the Nash equilibrium strategy and the empirical divergence strategy is sub-linear in n.

V. A Coin Guessing Game

As an illustration, let us consider a specific instance of the problem. In this instance, $\underline{X}$ is a sequence of i.i.d. binary $Bernoulli(1/2)$ random variables. These are observed by the cheating player and the regular players through different memoryless channels. The regular player channel $V_{Z^{(r)}|X}$ is $BSC(p_r)$. The cheating player has a stochastically upgraded channel $BSC(p_c)$ with $p_c < p_r \leq 1/2$. The monitoring channel is $BSC(p_m)$, $p_m \leq 1/2$. The goal of the players is to guess the values of X_i's, so we have $R(\underline{X}, \underline{Y}) = \sum_{i=1}^{n} \mathbb{1}\{X_i = Y_i\}$.

The optimization problem in Theorem 1 is a convex optimization problem, as the objective function is a linear function of $P_{Y^{(c)}|Z^{(c)}}$. The constraint is also convex as $V_{\tilde{Y}^{(c)}|X}$ is linear w.r.t. to $P_{Y^{(c)}|Z^{(c)}}$, while $D(.||V_{\tilde{Y}^{(r)}|X}|V_X)$ is also convex.

For the regular player, the optimal strategy is to assign $Y^{(r)} = Z^{(r)}$. This strategy leads to the expected reward of $\mathbb{E}[R(X, Y^{(r)})] = 1 - p_r$. We can see that this problem is symmetric with respect to X, so the optimal $P_{Y^{(c)}|Z^{(c)}}$ is also BSC with certain flip probability. For this problem, we can express the optimal reward of cheating player in Theorem 1 as

$$\max_{\substack{p\in[0,1]: \\ d_2\left(p_c * p * p_m || p_r * p_m\right) \leq K}} 1 - (p * p_c) \quad (21)$$

where $p * p' := p(1-p') + p'(1-p)$ and $d_2(.||.)$ is the binary KL divergence.

A numerical example of the trade-off for a specific value of p_c and p_r is given in figure 1. We can observe several properties,

- Even if $p_m = 0$ (i.e., the monitor has perfect knowledge of all players' actions), the cheating player can still improve its normalized reward by exploiting its information advantage if the number of regular players is large enough.
- There is a cut-off K^*, after which the cheating player does not get any further advantage from larger K. This K^* corresponds to the conditional KL divergence between the optimal distribution of cheating player and the regular player.

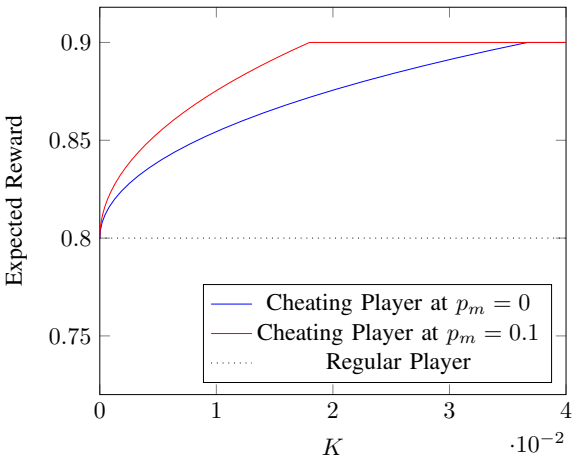


Fig. 1. Trade-off between the normalized reward and the growth rate K of the number of players. The plots are for $p_c = 0.1$ and $p_r = 0.2$.

VI. Conclusion

In this work, we introduce a model of monitoring problem for several players, where a cheating player has an information advantage compared to the regular players. We quantify the trade-off between the linear growth rate of cheating player's reward and the exponential growth of the number of regular players.

We show that there exists a detection strategy that the monitor can utilize to minimize the cheating player's normalized reward, regardless of the sequence of randomized strategies of the cheating player. It is surprising that this strategy only depends on the knowledge of the probabilistic model of the regular players.

There are several future directions that we have in mind. Currently the model assumes that the cheating player has access to very large amount of randomness and it is allowed to observe the realization of its whole private information before performing an action. An interesting extension is to study what the cheating player can achieve under a more restrictive model where it has limited amount of randomness and has to use an online policy. Intuitively, this problem requires the cheating player to simulate its chosen distribution under a more restrictive model, hence a connection with resolvability problems.

Another interesting extension is to study the optimal strategy under the sub-exponential growth of the number of regular players. This brings a connection with recent literature on finite length information theory.

Acknowledgment

We would like to thank the reviewers for their insightful suggestions and comments.

REFERENCES

[1] J. von Neumann, O. Morgenstern, and A. Rubinstein, *Theory of Games and Economic Behavior (60th Anniversary Commemorative Edition)*. Princeton University Press, 1944.

[2] W. Hoeffding, "Asymptotically Optimal Tests for Multinomial Distributions," *The Annals of Mathematical Statistics*, vol. 36, no. 2, pp. 369 – 401, 1965.

[3] M. Barni and B. Tondi, "The source identification game: An information-theoretic perspective," *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 3, pp. 450–463, 2013.

[4] M. Feder and N. Merhav, "Universal composite hypothesis testing: a competitive minimax approach," *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1504–1517, 2002.

[5] B. Tondi, M. Barni, and N. Merhav, "Detection games with a fully active attacker," in *2015 IEEE International Workshop on Information Forensics and Security (WIFS)*, 2015, pp. 1–6.

[6] S. Yasodharan and P. Loiseau, "Nonzero-sum Adversarial Hypothesis Testing Games," in *NeurIPS 2019 - Thirty-third Conference on Neural Information Processing Systems*, Vancouver, Canada, Dec. 2019, pp. 1–23.

[7] L. Dritsoula, P. Loiseau, and J. Musacchio, "A game-theoretic analysis of adversarial classification," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 12, pp. 3094–3109, 2017.

[8] I. Csiszar, "The method of types," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2505–2523, 1998.

[9] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*. Springer, 1998.

APPENDIX

A. Proof of Proposition 1

This is a consequence of the Von Neumann's minimax theorem (see section 17.6 in [1]). Due to our assumption that the alphabet of the random variables are finite, then the set of deterministic policies of the cheater and the set of deterministic decision rules of the monitor are finite. By the linearity of expectation, one can express the reward function as a bilinear form. Hence this problem fulfils the hypothesis of Von Neumann's minimax theorem.

B. Proof of Proposition 3

If for a certain i, $(x, \tilde{y}[i])$ has 0 under the regular player probabilistic model, then the monitor can choose that player and incur no error probability. Hence, such $(x, \tilde{y}[i])$ pairs do not contribute to the reward calculation.

Hence, let us only consider the case where the probability of all pairs $(x, \tilde{y}[i])$ are non-zero under the regular player probabilistic model. Note that we can write

$$\begin{aligned}&\mathbb{E}[\mathbb{1}\{\hat{H} \neq H\}R(X, Y[H])]\\ &= \mathbb{E}[R(X, Y[H])] - \mathbb{E}[\mathbb{1}\{\hat{H} = H\}R(X, Y[H])]. \end{aligned} \tag{22}$$

The goal of the monitor decision rule is equivalent to maximizing the second term in the RHS.

$$\begin{aligned}&\mathbb{E}[\mathbb{1}\{\hat{H} = H\}R(X, Y[H])] \\ &= \sum_{x, (\tilde{y}[k])_{k=0}^{M}} \frac{V_X(x)}{M+1} \prod_{j=0}^{M} V_{\tilde{Y}^{(r)}|X}(\tilde{y}[j]|x) \\ &\quad \mathbb{E}\left[\frac{V_{\tilde{Y}^{(c)}|X}(\tilde{y}[\hat{H}]|x)}{V_{\tilde{Y}^{(r)}|X}(\tilde{y}[\hat{H}]|x)} \mathbb{E}[R(x, Y^{(c)})|X = x, \tilde{Y}^{(c)} = \tilde{y}[\hat{H}]]\right]. \end{aligned} \tag{23, 24}$$

This quantity is maximized if $\hat{H}$ is chosen to maximize the following metric,

$$g(x, \tilde{y}) = \frac{V_{\tilde{Y}^{(c)}|X}(\tilde{y}|x)}{V_{\tilde{Y}^{(r)}|X}(\tilde{y}|x)} \mathbb{E}[R(x, Y^{(c)})|X = x, \tilde{Y}^{(c)} = \tilde{y}] \tag{25}$$

for $(x, \tilde{y})$ with non-zero probability in $V_{\tilde{Y}^{(r)}|X}$, and ∞ otherwise.

C. Proof of Proposition 5

Consider,

$$\begin{aligned}&\max_{f^{(c)}} \min_{\hat{H}} \mathbb{E}[\mathbb{1}\{\hat{H} \neq H\}R(X, Y[H])] \\ &= \max_{f^{(c)}} \min_{\hat{H} \in \hat{\mathcal{H}}_m} \mathbb{E}[\mathbb{1}\{\hat{H} \neq H\}R(X, Y[H])] \\ &= \max_{f^{(c)}} \min_{\hat{H} \in P(\hat{\mathcal{H}}_m)} \mathbb{E}[\mathbb{1}\{\hat{H} \neq H\}R(X, Y[H])] \\ &= \min_{\hat{H} \in P(\hat{\mathcal{H}}_m)} \max_{f^{(c)}} \mathbb{E}[\mathbb{1}\{\hat{H} \neq H\}R(X, Y[H])] \\ &= \min_{\hat{H} \in P(\hat{\mathcal{H}}_m)} \max_{f^{(c)} \in \mathcal{F}_d} \mathbb{E}[\mathbb{1}\{\hat{H} \neq H\}R(X, Y[H])] \end{aligned} \tag{26}$$

where the extension to randomized strategies on the second line is necessary so we can use the minimax theorem.

D. Proof of Proposition 6

First we show that,

$$\begin{aligned}&\max_{f^{(c)}} \min_{\hat{H}} \mathbb{E}[\mathbb{1}\{\hat{H} \neq H\}R(\underline{X}, \underline{Y}[H])] \\ &= \max_{f^{(c)} \in \mathcal{F}_T} \min_{\hat{H}} \mathbb{E}[\mathbb{1}\{\hat{H} \neq H\}R(\underline{X}, \underline{Y}[H])]. \end{aligned} \tag{27}$$

Given a $f^{(c)}$, let us consider the *permutation-invariant* version which takes a permutation π from the set of permutation Π uniformly at random, such that $\tilde{f}^{(c)}(\underline{z}) = \pi^{-1}(f^{(c)}(\pi(\underline{z})))$. Let us denote the distribution induced by this *permutation-invariant* version as $\tilde{P}$.

We have

$$\begin{aligned}&\min_{\hat{H}} \mathbb{E}_{\tilde{P}}[\mathbb{1}\{\hat{H} \neq H\}R(\underline{X}, \underline{Y}[H])] \\ &= \min_{\hat{H}} \sum_{\pi} \frac{1}{|\Pi|} \mathbb{E}_{\tilde{P}}[\mathbb{1}\{\hat{H} \neq H\}R(\underline{X}, \underline{Y}[H])|\pi] \\ &\geq \sum_{\pi} \frac{1}{|\Pi|} \min_{\hat{H}} \mathbb{E}_{\tilde{P}}[\mathbb{1}\{\hat{H} \neq H\}R(\underline{X}, \underline{Y}[H])|\pi] \end{aligned} \tag{28}$$

$$\begin{aligned}&\overset{(1)}{=} \sum_{\pi} \frac{1}{|\Pi|} \min_{\hat{H}} \mathbb{E}[\mathbb{1}\{\hat{H} \neq H\}R(\pi^{-1}(\underline{X}), \pi^{-1}(\underline{Y}[H]))] \\ &\overset{(2)}{=} \min_{\hat{H}} \mathbb{E}[\mathbb{1}\{\hat{H} \neq H\}R(\underline{X}, \underline{Y}[H])]. \end{aligned} \tag{29}$$

We require the facts that the channel is memoryless in (1), and that the reward is additive in (2). This inequality implies that there exists $\tilde{f}^{(c)} \in \mathcal{F}_t$ which is as good as any $f^{(c)} \notin \mathcal{F}_t$.

Finally, we show that,

$$\max_{f^{(c)}\in\mathcal{F}_T}\min_{\hat{H}}\mathbb{E}[\mathbb{1}\{\hat{H}\neq H\}R(\underline{X},\underline{Y}[H])] = \max_{f^{(c)}\in\mathcal{F}_T}\min_{\hat{H}\in\hat{\mathcal{H}}_T}\mathbb{E}[\mathbb{1}\{\hat{H}\neq H\}R(\underline{X},\underline{Y}[H])]. \quad (30)$$

We can show this by showing that the optimal metric that we established in eq. (25) is a *permutation invariant* decision rule under the assumption that $f^{(c)}$ is *permutation-invariant*. More formally, consider two tuples $(\underline{x},\underline{\tilde{y}})$ and $(\underline{x}',\underline{\tilde{y}}')$ both of the same joint type. This implies that there exists a permutation π such that $(\underline{x}',\underline{\tilde{y}}')=(\pi(\underline{x}),\pi(\underline{\tilde{y}}))$. Notice that for any $\pi\in\Pi$, then $\Pi=\cup_{\pi'\in\Pi}\{\pi\circ\pi'\}$. We have,

$$\begin{aligned}
&V_{\underline{\tilde{Y}}^{(c)}|\underline{X}}(\underline{\tilde{y}}',\underline{x}')\\
&=\sum_{\underline{y}'}V_{\underline{Y},\underline{\tilde{Y}}^{(c)}|\underline{X}}(\underline{y},\underline{\tilde{y}}',\underline{x}')\\
&=\sum_{\underline{y}'}V_{\underline{\tilde{Y}}^{(c)}|\underline{Y}^{(c)}}(\underline{\tilde{y}}',\underline{y}')V_{\underline{Y}^{(c)}|\underline{X}}(\underline{y}',\underline{x}')\\
&\overset{(1)}{=}\sum_{\underline{y}'}V_{\underline{\tilde{Y}}^{(c)}|\underline{Y}^{(c)}}(\underline{\tilde{y}}',\underline{y}')\sum_{\pi'}\frac{V_{\underline{Y}^{(c)}|\underline{X}}(\pi'(\underline{y}'),\pi'(\underline{x}'))}{|\Pi|}\\
&\overset{(2)}{=}\sum_{\underline{y}'}V_{\underline{\tilde{Y}}^{(c)}|\underline{Y}^{(c)}}(\pi(\underline{\tilde{y}}),\pi(\underline{y}))\sum_{\pi'}\frac{V_{\underline{Y}^{(c)}|\underline{X}}(\pi'\circ\pi(\underline{y}),\pi'\circ\pi(\underline{x}))}{|\Pi|}\\
&=\sum_{\underline{y}'}V_{\underline{\tilde{Y}}^{(c)}|\underline{Y}^{(c)}}(\underline{\tilde{y}},\underline{y})\sum_{\pi'}\frac{V_{\underline{Y}^{(c)}|\underline{X}}(\pi'(\underline{y}),\pi'(\underline{x}))}{|\Pi|}\\
&=V_{\underline{\tilde{Y}}^{(c)}|\underline{X}}(\underline{\tilde{y}},\underline{x})
\end{aligned}$$

We used the fact that $f^{(c)}$ is *permutation-invariant* in (1), and the fact that we are on the *product*-regime in (2). We can use the same argument to show that $\mathbb{E}[R(\underline{x},\underline{Y}^{(c)})|\underline{X}=\underline{x},\underline{\tilde{Y}}^{(c)}=\underline{\tilde{y}}]$ also has the same property. As every term of the optimal metric only depends on the joint type, the optimal metric also only depends only on the joint-type.

E. Proof of Proposition 7

For $P'_{\tilde{Y},X}$ such that $P'_X\neq P_X$, the conditioning event is a null event, hence we can ignore this event as we are working with discrete variables. So we will only focus on n such that $P'_X=P_X$. We have,

$$\begin{aligned}
&\Pr\left(\hat{H}=H\,\middle|\,(\underline{X},\underline{\tilde{Y}}[H])\in T_{P_{X,\tilde{Y}}}\right)\\
&=\left(1-\Pr(g_n(X,\tilde{Y}^{(r)})>g_n(P_{X,\tilde{Y}})|(\underline{X},\underline{\tilde{Y}}[H])\in T_{P_{X,\tilde{Y}}}\right)^{e^{nK}}\\
&\overset{(a)}{\leq}\left(1-e^{-n\min_{P'_{X,\tilde{Y}}\in U_n(P_{X,\tilde{Y}})}D(P'_{X,\tilde{Y}}||V_{X,\tilde{Y}^{(r)}})+o(n)}\right)^{e^{nK}}\\
&\overset{(b)}{\leq}\exp\left(-e^{n(K-\min_{P'_{X,\tilde{Y}}\in U_n(P_{X,\tilde{Y}})}D(P'_{X,\tilde{Y}}||V_{X,\tilde{Y}^{(r)}})+o(n)}\right)\\
&\overset{(c)}{\leq}\exp\left(-e^{n(K-D(P'_{X,\tilde{Y}}||V_{X,\tilde{Y}^{(r)}})+o(n)}\right) \quad (31)
\end{aligned}$$

where the last inequality holds for any $P'_{X,\tilde{Y}}\in\mathcal{U}_n(P_{X,\tilde{Y}})$. For inequality (a), we use the method of type estimate of the probability. For inequality (b), we use $(1-x)\leq\exp(-x)$. For inequality (c), we possibly choose a non-minimizer of the upper bound. Note that if a type $P'_{X,\tilde{Y}}$ is included in $\mathcal{U}(P_{X,\tilde{Y}})$, it is included in $\mathcal{U}_n(P_{X,\tilde{Y}})$ for infinitely many n. Hence, for any K and $P'_{X,\tilde{Y}}\in\mathcal{U}(P_{X,\tilde{Y}})$, we have

$$\begin{aligned}
&\liminf_{n\to\infty}\Pr\left(\hat{H}=H\,\middle|\,(\underline{X},\underline{\tilde{Y}}[H])\in T_{P_{X,\tilde{Y}^{(c)}}}\right)\\
&\leq\liminf_{n\to\infty}\exp\left(-e^{n(K-D(P'_{X,\tilde{Y}}||V_{X,\tilde{Y}^{(r)}})+o(n)}\right), \quad (32)
\end{aligned}$$

and thus,

$$\begin{aligned}
&\liminf_{n\to\infty}\Pr\left(\hat{H}=H\,\middle|\,(\underline{X},\underline{\tilde{Y}}[H])\in T_{P_{X,\tilde{Y}^{(c)}}}\right)\\
&\leq\liminf_{n\to\infty}\exp\left(-e^{n(K-G(P_{X,\tilde{Y}^{(c)}}))+o(n)}\right), \quad (33)
\end{aligned}$$

by the definition of $G(P_{X,\tilde{Y}^{(c)}})$. Where $K>G(P_{X,\tilde{Y}})$, eq. (33) proves the first part of the proposition.

For the second part, it is sufficient to use the union bound. We have,

$$\begin{aligned}
&\Pr\left(\hat{H}\neq H\,\middle|\,(\underline{X},\underline{\tilde{Y}}[H])\in T_{P_{X,\tilde{Y}}}\right)\\
&\leq e^{nK}\Pr(g_n(\underline{X},\underline{\tilde{Y}}^{(r)})>g_n(P_{X,\tilde{Y}})|(\underline{X},\underline{\tilde{Y}}[H])\in T_{P_{X,\tilde{Y}}})\\
&\leq e^{n(K-\min_{P'_{X,\tilde{Y}}\in U_n(P_{X,\tilde{Y}})}D(P'_{X,\tilde{Y}}||V_{X,\tilde{Y}^{(r)}})+o(n)} \quad (34)\\
&\leq e^{n(K-\min_{P'_{X,\tilde{Y}}\in\cup_{i\geq n}U_i(P_{X,\tilde{Y}})}D(P'_{X,\tilde{Y}}||V_{X,\tilde{Y}^{(r)}})+o(n)} \quad (35)
\end{aligned}$$

This upper bound is valid for all n. Since $\cup_{i\geq n}U_i(P_{X,\tilde{Y}})$ is a decreasing sequence of sets in n with limit $\mathcal{U}(P_{X,\tilde{Y}})$, for all $\epsilon>0$ there exists n^* such that for all $n>n^*$

$$\min_{P'_{X,\tilde{Y}}\in\cup_{i\geq n}U_i(P_{X,\tilde{Y}})}D(P'_{X,\tilde{Y}}||V_{X,\tilde{Y}})\geq G(P_{X,\tilde{Y}})-\epsilon \quad (36)$$

due to the definition of $G(P_{X,\tilde{Y}})$ and the continuity of the KL divergence. Hence,

$$\limsup_{n\to\infty}-\frac{1}{n}\log\Pr\left(\hat{H}\neq H\,\middle|\,(\underline{X},\underline{\tilde{Y}}[H])\in T_{P_{X,\tilde{Y}^{(c)}}}\right)>0 \quad (37)$$

if $K<G(P_{X,\tilde{Y}})$.

F. Proof of Proposition 8

To show the first part namely eq. (16), fix $\epsilon>0$ and consider a $P_{Y^{(c)}|Z^{(c)}}$ such that $V_{X,\tilde{Y}^{(c)}}\in\text{IntCl}(G^{-1}_{(g_n)_{n=1}^{\infty}}(K))$ and $\mathbb{E}_{P_{X,Y^{(c)}}}[R(X,Y^{(c)})]$ is ϵ-close to the supremum on the RHS of eq. (16).

Now fix n. Let us construct our $f^{(c)}$ by sampling from the distribution $P_{Y^{(c)}|Z}$ given the realization of $\underline{Z}$. From standard results in method of types we know that for every $\delta>0$, $P_X(\underline{X}\in T_{[P_X]_\delta})\geq 1-Ae^{-n\delta^2}$ and $V_{\tilde{Y}|X}(\underline{\tilde{Y}}\in T_{[V_{\tilde{Y}|X}]_{2\delta}}(\underline{X})|\underline{X}\in T_{[P_X]_\delta})\geq 1-A'e^{-n\delta^2}$ (A and A' only depend on the sizes of the alphabets of X and $\tilde{Y}$), where $T_{[V_{X,\tilde{Y}^{(c)}}]_\delta}$ is the set of all types with total variation distance at most δ from $V_{X,\tilde{Y}^{(c)}}$. As the $V_{X,\tilde{Y}^{(c)}}$ is in the $\text{IntCl}(G^{-1}_{(g_n)_{n=1}^{\infty}}(K))$, then for small enough δ we have,

$$T_{[V_{X,\tilde{Y}^{(c)}}]_{2\delta}}\subseteq\text{IntCl}(G^{-1}_{(g_n)_{n=1}^{\infty}}(K)) \quad (38)$$

Let us define the event Q_n as $\{X \in T_{[P_X]_\delta}, \underline{\tilde{Y}}[H] \in T_{[V_{\tilde{Y}^{(c)}|X}]_{2\delta}}(\underline{X})\}$. We note that $\lim_{n\to\infty} \Pr(Q_n) = 1$. So we have under $V_{X,\tilde{Y}^{(c)}}$,

$$\begin{aligned}
&\mathbb{E}[R(\underline{X}, \underline{Y}[H])|Q_n] \Pr(Q_n) \\
&= \mathbb{E}[R(\underline{X}, \underline{Y}[H])] - \mathbb{E}[R(\underline{X}, \underline{Y}[H])|Q_n^C](1 - \Pr(Q_n)) \\
&\geq \mathbb{E}[R(\underline{X}, \underline{Y}[H])] - O(ne^{-n\delta^2}).
\end{aligned} \tag{39}$$

We also have,

$$\begin{aligned}
&\frac{1}{n}\mathbb{E}[\mathbb{1}\{\hat{H} \neq H\}R(\underline{X}, \underline{Y}[H])] \\
&\geq \frac{1}{n}\mathbb{E}[\mathbb{1}\{\hat{H} \neq H\}R(\underline{X}, \underline{Y}[H])|Q_n] \Pr(Q_n)
\end{aligned} \tag{40}$$

Conditioned on Q_n, we have,

$$\frac{1}{n}R(\underline{X}, \underline{Y[H]}) \geq \mathbb{E}_V[R(\underline{X}, \underline{Y}^{(c)})](1 - O(\delta)). \tag{41}$$

Therefore

$$\begin{aligned}
&\frac{1}{n}\mathbb{E}[\mathbb{1}\{\hat{H} \neq H\}R(\underline{X}, \underline{Y}[H])] \\
&\geq \mathbb{E}_V[R(\underline{X}, \underline{Y}^{(c)})](1 - O(\delta)) \Pr(\hat{H} \neq H|Q_n) \Pr(Q_n).
\end{aligned} \tag{42}$$

Note that

$$\Pr(\hat{H} \neq H|Q_n) \geq b_n \tag{43}$$

where

$$b_n = \min_{P'_{\tilde{Y}|X} \in T_{[V_{\tilde{Y}|X}]_{2\delta}}} \Pr(\hat{H} \neq H|\underline{X} \in T_{[P_X]_\delta}, \underline{\tilde{Y}} \in T_{P'_{\tilde{Y}|X}}(\underline{X})).$$

Combining these terms give us,

$$\begin{aligned}
&\limsup_{n\to\infty} \frac{1}{n}\mathbb{E}[\mathbb{1}\{\hat{H} \neq H\}R(\underline{X}, \underline{Y}[H])] \\
&\geq (\mathbb{E}_V[R(\underline{X}, \underline{Y}^{(c)})](1 - O(\delta)) \limsup_{n\to\infty} \Pr(Q_n) b_n \\
&= (\mathbb{E}_V[R(\underline{X}, \underline{Y}^{(c)})](1 - O(\delta)) \limsup_{n\to\infty} b_n \\
&\overset{(*)}{=} \mathbb{E}_V[R(\underline{X}, \underline{Y}^{(c)})](1 - O(\delta))
\end{aligned} \tag{44}$$

for equality $(*)$ we used the fact that $T_{[V_{\tilde{Y}^{(c)}|X}]_{2\delta}} \subseteq \mathrm{IntCl}(G^{-1}_{(g_n)_{n=1}^\infty}(K))$, therefore $\limsup_{n\to\infty} b_n = 1$ due to the first part of proposition 7. Since $\mathbb{E}_V[R(\underline{X}, \underline{Y}^{(c)})]$ is ϵ-close to the supremum and δ is arbitrary, eq. (16) follows.

Now we will show eq. (17). As in the previous part, for any $\delta > 0$, let us define, $Q_n = \{(\underline{X}, \underline{Z}) \in T_{[V_{X,Z^{(c)}}]_\delta}\}$. So we have

$$\begin{aligned}
&\frac{1}{n}\mathbb{E}[\mathbb{1}\{\hat{H} \neq H\}R(\underline{X}, \underline{Y}[H])] \\
&\leq \frac{1}{n}\mathbb{E}[\mathbb{1}\{\hat{H} \neq H\}R(\underline{X}, \underline{Y}[H])|Q_n] \Pr(Q_n) + O(\exp(-n\delta^2))
\end{aligned} \tag{45}$$

where we used the fact that $\Pr\left((\underline{X}, \underline{Z}) \notin T_{[V_{X,Z^{(c)}}]_\delta}\right) \leq \exp(-n\delta^2)$. Note that $f^{(c)} \in \mathcal{F}_T$ can be parametrized by set of tuples $\{(P_Z, P_{Y|Z}) : P_Z \in \mathcal{P}(Z)\}$ i.e., the policy chooses a single conditional type $P_{Y|Z}$ for each P_Z that it observes. One can use this parametrization to further upper bound the non-vanishing part of eq. (45) as

$$\begin{aligned}
&\frac{1}{n}\mathbb{E}[\mathbb{1}\{\hat{H} \neq H\}R(\underline{X}, \underline{Y}[H])|Q_n] \Pr(Q_n) \\
&\leq \sum_{P_{X,Z} \in T_{[V_{X,Z^{(c)}}]_\delta}} \max_{P_{Y|Z}} \frac{1}{n}\mathbb{E}[\mathbb{1}\{\hat{H} \neq H\}R(\underline{X}, \underline{Y}[H])|(\underline{X}, \underline{Z}) \in T_{P_{X,Z}}] \\
&\qquad \Pr((\underline{X}, \underline{Z}) \in T_{P_{X,Z}}).
\end{aligned} \tag{46}$$

We will use the fact that $P_{X,Z} \in T_{[V_{X,Z^{(c)}}]_\delta}$ to note that $(\underline{X}, \underline{Z}, \underline{Y}) \in T_{[V_{X,Z^{(c)}} \circ P_{Y|Z}]_{2\delta}}$ with probability of 1, as given $\underline{Z}$ the policy takes $\underline{Y} \in T_{P_{Y|Z}}(\underline{Z})$. Finally, $\underline{\tilde{Y}}$ fulfils $(\underline{X}, \underline{Z}, \underline{Y}, \underline{\tilde{Y}}) \in T_{[V_{X,Z^{(c)}} \circ P_{Y|Z} \circ V_{\tilde{Y}|Y}]_{A\delta}}$ (A is a constant that depends on the size of the alphabet of Y) with high probability since the channel $V_{\tilde{Y}|Y}$ is memoryless. Note that $V_{X,Z^{(c)}} \circ P_{Y|Z} \circ V_{\tilde{Y}|Y}$ is the definition of $V_{X,\tilde{Y}^{(c)}}$, i.e., the distribution of $X, \tilde{Y}^{(c)}$ under the cheating player probabilistic model. Hence we have,

$$\begin{aligned}
&\frac{1}{n}\mathbb{E}[\mathbb{1}\{\hat{H} \neq H\}R(\underline{X}, \underline{Y}[H])|Q_n] \Pr(Q_n) \\
&\leq \max_{P_{Y|Z}} (1 + O(\delta))E_V[R(X, Y^{(c)})]c_{n,\delta}(V_{X,\tilde{Y}^{(c)}}) \\
&\qquad + O(-nA^2\delta^2)
\end{aligned} \tag{47}$$

where,

$$c_{n,\delta}(V_{X,\tilde{Y}^{(c)}}) = \sum_{P_{X,\tilde{Y}} \in T_{[V_{X,\tilde{Y}^{(c)}}]_{A'\delta}}} \Pr(\hat{H} \neq H, (\underline{X}, \underline{\tilde{Y}}) \in T_{P_{X,\tilde{Y}}}) \tag{48}$$

with A' is a constant which only depends on the size of the alphabets.

This allows us to give an upper bound,

$$\begin{aligned}
&\limsup_{n\to\infty} \frac{1}{n} \max_{f^{(c)} \in \mathcal{F}_T} \mathbb{E}[\mathbb{1}\{\hat{H} \neq H\}R(\underline{X}, \underline{Y}[H])] \\
&\leq \limsup_{n\to\infty} \max_{P_{Y|Z}} (1 + O(\delta))E_V[R(X, Y^{(c)})]c_{n,\delta}(V_{X,\tilde{Y}^{(c)}}).
\end{aligned}$$

We can decompose the region of the maximization such that,

$$\begin{aligned}
&\limsup_{n\to\infty} \frac{1}{n} \max_{f^{(c)} \in \mathcal{F}_T} \mathbb{E}[\mathbb{1}\{\hat{H} \neq H\}R(\underline{X}, \underline{Y}[H])] \\
&\leq \limsup_{n\to\infty} \max\{u_n, v_n\} \\
&= \max\{\limsup_{n\to\infty} u_n, \limsup_{n\to\infty} v_n\}
\end{aligned} \tag{49}$$

with,

$$\begin{aligned}
u_n &= \max_{\substack{P_{Y|Z}: \\ V_{X,Y^{(c)}} \in D_\delta}} (1 + O(\delta))E_V[R(X, Y^{(c)})]c_{n,\delta}(V_{X,\tilde{Y}^{(c)}}) \\
v_n &= \max_{\substack{P_{Y|Z}: \\ V_{X,Y^{(c)}} \notin D_\delta}} (1 + O(\delta))E_V[R(X, Y^{(c)})]c_{n,\delta}(V_{X,\tilde{Y}^{(c)}})
\end{aligned} \tag{50}$$

where

$$D_\delta = \{V_{X,Y^{(c)}} : |V_{X,Y^{(c)}} - G^{-1}_{(g_n)_{n=1}^\infty}(K)| \leq 2A'\delta\}, \tag{51}$$

i.e., D_δ is the $2A'\delta$ neighborhood of $G^{-1}_{(g_n)_{n=1}^\infty}(K)$ in total variation distance.

By the second part of preposition 7, we have that

$$\lim_{n\to\infty} c_{n,\delta}(V_{X,\tilde{Y}^{(c)}}) = 0 \tag{52}$$

if $V_{X,\tilde{Y}^{(c)}} \notin \text{Int}(D_\delta)$. Hence, v_n vanishes as $n \to \infty$. For the u_n, we can upper bound $c_{n,\delta}(.)$ by 1. This gives us,

$$\begin{aligned}&\limsup_{n\to\infty} \frac{1}{n} \max_{f^{(c)}\in\mathcal{F}_T} \mathbb{E}[\mathbb{1}\{\hat{H} \neq H\} R(\underline{X}, \underline{Y}[H])] \\ &\leq \max_{\substack{P_{Y|Z}: \\ V_{X,Y^{(c)}} \in D_\delta}} (1 + O(\delta)) E_V[R(X, Y^{(c)})].\end{aligned} \tag{53}$$

Finally, note that this upper bound holds for arbitrary δ and we have

$$\text{Cl}(G^{-1}_{(g_n)_{n=1}^\infty}(K)) = \bigcap_{\delta>0} D_\delta. \tag{54}$$

Combined with the fact that expectation is continuous under total variation distance, this upper bound establishes the second part of proposition 8.

G. Proof of Preposition 9

From proposition 8, one way to prove the statement is to show that,

$$\text{Cl}(G^{-1}_{(g_n^{ed})_{n=1}^\infty}(K)) \subseteq \text{ClInCl}(G^{-1}_{(g_n)_{n=1}^\infty}(K)). \tag{55}$$

The ClInCl(A) is the closure of the interior of closure of set A. If the inclusion holds, then the upper bound part of proposition 8 applied on $R((g_n^{ed})_{n=1}^\infty)$ is smaller than the lower bound part of proposition 8 applied on $R((g_n)_{n=1}^\infty)$. This is the approach that we will take.

First, observe that for every n and P_X we have,

$$\begin{aligned}&\{P'_{X,\tilde{Y}} : P'_X = P_X, g_n^{ed}(P'_{X,\tilde{Y}}) \leq \sup_{\substack{P^*_{X,\tilde{Y}}: \\ D(P^*_{X,\tilde{Y}}||V_{X,\tilde{Y}^{(r)}})<K}} g_n^{ed}(P^*_{X,\tilde{Y}})\} \\ &\overset{(*)}{=} \{P'_{X,\tilde{Y}} : P'_X = P_X, D(P'_{X,\tilde{Y}}||V_{X,\tilde{Y}^{(r)}}) < K\} \\ &\subseteq \{P'_{X,\tilde{Y}} : P'_X = P_X, g_n(P'_{X,\tilde{Y}}) \leq \sup_{\substack{P^*_{X,\tilde{Y}}: \\ D(P^*_{X,\tilde{Y}}||V_{X,\tilde{Y}^{(r)}})<K}} g_n(P^*_{X,\tilde{Y}})\}.\end{aligned} \tag{56}$$

Which imply that for every P_X,

$$G^{-1}_{(g_n^{ed})_{n=1}^\infty}(P_X, K) \subseteq G^{-1}_{(g_n)_{n=1}^\infty}(P_X, K) \tag{57}$$

and taking the appropriate $\lim$ gives us,

$$\begin{aligned}&G^{-1}_{(g_n^{ed})_{n=1}^\infty}(K) \\ &= \lim_{\epsilon\to 0} \bigcup_{\substack{P_X: \\ |V_X - P_X|<\epsilon}} G^{-1}_{(g_n^{ed})_{n=1}^\infty}(P_X, K) \\ &\subseteq \lim_{\epsilon\to 0} \bigcup_{\substack{P_X: \\ |V_X - P_X|<\epsilon}} G^{-1}_{(g_n)_{n=1}^\infty}(P_X, K) = G^{-1}_{(g_n)_{n=1}^\infty}(K)\end{aligned} \tag{58}$$

Furthermore, taking the ClIntCl in both sides gives us,

$$\text{ClIntCl}(G^{-1}_{(g_n^{ed})_{n=1}^\infty}(K)) \subseteq \text{ClIntCl}(G^{-1}_{(g_n)_{n=1}^\infty}(K)). \tag{59}$$

So we only need to show that,

$$\text{ClIntCl}(G^{-1}_{(g_n^{ed})_{n=1}^\infty}(K)) = \text{Cl}(G^{-1}_{(g_n^{ed})_{n=1}^\infty}(K)) \tag{60}$$

which is equivalent to showing that $\text{Cl}(G^{-1}_{(g_n^{ed})_{n=1}^\infty}(K))$ does not contain any isolated points. Let

$$\mathcal{A} = \{P'_{X,\tilde{Y}} : D(P'_{X,\tilde{Y}}||V_{X,\tilde{Y}^{(r)}}) < K\}.$$

So we have $G^{-1}_{(g_n^{ed})_{n=1}^\infty}(K) = \mathcal{P}_\infty \bigcap \mathcal{A}$. As $\mathcal{P}_\infty$ is dense, then $\text{Cl}(G^{-1}_{(g_n^{ed})_{n=1}^\infty}(K)) = \text{Cl}(\mathcal{A}) = \{P'_{X,\tilde{Y}} : D(P'_{X,\tilde{Y}}||V_{X,\tilde{Y}^{(r)}}) \leq K\}$ which is a closed convex set with no isolated points.