2020 54th Annual Conference on Information Sciences and Systems (CISS)

Sequential Hypothesis Testing Game

Guanze Peng
Dept of Electrical and Computer Engineering
New York University
Brooklyn, USA
guanze.peng@nyu.edu

Quanyan Zhu
Dept of Electrical and Computer Engineering
New York University
Brooklyn, USA
quanyan.zhu@nyu.edu

2020 54th Annual Conference on Information Sciences and Systems (CISS) 978-1-7281-4085-8/20/$31.00 ©2020 IEEE 10.1109/CISS48834.2020.1570617162

***Abstract*—In this work, we study a stopping time game problem in sequential hypothesis testing, where both of the two players perform hypothesis testing with distinct hypotheses. The payoff of the players depends on the order of stopping times. Therefore, apart from designing the decision function concerning the hypotheses, the players also determine the optimal stopping timings. We investigate the cases where the time horizon is finite or infinite and provide sufficient conditions of finding the equilibrium point. Moreover, we fully characterize the structural properties of the equilibrium strategies.**

***Index Terms*—Sequential hypothesis testing, stopping time game, randomized stopping time, Nash equilibrium**

I. Introduction

Over the past decades, hypothesis testing has been a fundamental problem in numerous areas, such as signal processing, statistics, and economics [1]–[3]. Many variants of hypothesis testing have been studied extensively, such as robust formulation, M-ary hypothesis testing, and sequential probability ratio test (SPRT) [4], [5].

In this paper, we explore the hypothesis testing problem in a multiagent setting where each player makes decisions sequentially to choose the stopping time and the hypothesis. The payoff of a player depends on the order of the stopping times of all the players and the risk associated with his decision. The agent's performance in terms of the hypothesis testing improves as the number of observation increases. However, if he stops later than the other agent, a large cost incurs. Therefore, there exists trade-off between stopping or continuing obtaining the observations. We formulate this problem as a stopping time game and the players aim to find an equilibrium strategies.

Our work is motivated by cybersecurity applications. Consider a computer network, in which each node has two types, *normal node* and *honeypot node* [6]. These two kinds of nodes provide different services to the user of the network. A user also has two types, *normal user* and *attacker*, each of which behaves differently in the network. Both the owner of the network and the user aim to determine each other's type by performing sequential hypothesis testing based on their observations. The owner can choose to kick the user out of the network or allow him to stay. The user can choose to leave the network or stay. Therefore, due to the sequential nature, it is natural to formulate this problem as a stopping time game in sequential hypothesis testing.

The natural solution concept associated with the sequential hypothesis testing game (SHTG) is the Nash equilibrium in which one player can benefit from unilateral deviation in their stopping time policy as well as decision function. We study the SHTG in two scenarios. One is the finite-horizon problem and the other one is the infinite-horizon one. We provide sufficient conditions to characterize the equilibria solution in both cases and obtain structural properties of the equilibrium strategies. We show that finding the equilibrium is equivalent to solving two interdependent dynamic programming problems. We show that when the costs induced by the order of stopping are sufficiently large, there exist fully randomized strategies. In this fully-randomized case, the randomized strategy of one player entirely depends on the other player's costs.

Our work is closely related to Dynkin's stopping time game [7], [8]. The formulation of SHTG is a variant of Dynkin's game in which players determine both the stopping times and their hypothesis. The formulation of sequential hypothesis testing game is naturally nonzero-sum [9] and it is an extension of one-player SPRT problem [5]. In our work, we do not incorporate cost of taking sequential observations and do not assume that the players have the same hypotheses as studied in most multi-agent hypothesis testing problems [10].

This paper is organized as the following. In Section II, we review the two classical hypothesis testing models and some fundamental results. We also point out the equivalence between these two models. In Section III, we formulate the stopping time game in sequential hypothesis testing. We first study the finite horizon case, provide sufficient condition for find the equilibria, and then extend it to infinite horizon case. In Section IV, we conclude this work and give possible directions of the future work.

II. Sequential Hypothesis Testing

In this section, we review the classical hypothesis testing formulation and the important results, which we will be revisited in subsequent sections. One may refer to [2], [5] for further references. The objective of hypothesis testing is to determine between two hypotheses H_0 and H_1 based on the observation $\{x_n\}_{n\in\mathbb{N}^+}$, which is the realization of independent and identically distributed (i.i.d.) random variables $\{X_n\}_{n\in\mathbb{N}^+}$. Let $\mathcal{X}$ be the space where the random variables X_n take values in. Under the different hypothesis, for $n \in \mathbb{N}^+$, the random variable has the following distribution.

$$H_0 \ : \ X_n \ \sim \ f(x_n|H_0) := f_0(x_n),$$
$$H_1 \ : \ X_n \ \sim \ f(x_n|H_1) := f_1(x_n).$$

Given observations $\mathbf{x}_1^N = \{x_1, x_2, ..., x_N\}$, a decision function $\delta(\mathbf{x}_1^N)$ is to decide the hypothesis. Formally,

$$\delta \ : \ \mathcal{X}^N \ \rightarrow \ \{0, 1\}.$$

If $\delta(\mathbf{x}_1^N) = 0$, then based on observations $\mathbf{x}_1^N$, we decide H_0 is true; and if $\delta(\mathbf{x}_1^N) = 1$, we decide H_1 is true.

We assume that there exist positive costs C_{01} and C_{10}. The cost C_{01} represents the cost of deciding that H_0 is true when H_1 holds, and C_{01} represents the cost of deciding that H_1 is true when H_0 holds. As a consequence, we can formulate the hypothesis testing problem as an optimization problem. The objective function, which is termed as *Bayes risk*, is given by the following

$$B_N(\delta) = \sum_{\substack{i,j=0 \\ i \neq j}}^{1} C_{ij} \mathbb{P}\left[\delta(\mathbf{X}_1^N) = i | H_j\right] \pi_j, \tag{1}$$

where π_0 and π_1 are the prior probabilities of H_0 and H_1 being true, respectively. Let $L(\mathbf{x}_1^N)$ be the likelihood ratio, i.e.,

$$L(\mathbf{x}_1^N) = \prod_{n=1}^{N} \frac{f_1(x_n)}{f_0(x_n)}.$$

Proposition 1. *The optimal decision function is given by*

$$\delta^*(\mathbf{x}_1^N) \ = \ \begin{cases} 0 & \text{if } L(\mathbf{x}_1^N) < T, \\ 1 & \text{if } L(\mathbf{x}_1^N) \geq T, \end{cases} \tag{2}$$

where the threshold is $T = C_{10}\pi_0/(C_{01}\pi_1)$.

In sequential hypothesis testing, one forms belief of the hypothesis, which is denoted by b_n. The belief b_n is a posterior probability,

$$b_n \ = \ \mathbb{P}\left[H_0 | \mathbf{x}_1^n\right], \quad \text{and} \quad b_0 = \pi_0.$$

The statistic b_n is updated according to the Bayesian rule as

$$\begin{aligned} b_{n+1} &= \frac{f_0(x_{n+1}) b_n}{f_0(x_{n+1}) b_n + f_1(x_{n+1})(1 - b_n)} \\ &= \frac{b_n}{b_n + L(x_{n+1})(1 - b_n)}. \end{aligned}$$

This posterior probability indicates the probability of H_0 being the true hypothesis given observation $\mathbf{x}_1^n$ at time n.

Corollary 1. *The Bayesian belief* b_N *is the sufficient statistic in optimizing* (1).

Proof. See Appendix A. □

Remark 1. *The corollary implies that,* b_N *provides the same information as* $\mathbf{x}_1^N$ *in optimizing the Bayes risk. As a result, one can use* b_N *to find the optimal decision function, instead of using* $\mathbf{x}_1^N$*. Moreover,* ***Corollary 1*** *points out the relation between hypothesis testing with repeated observations and the sequential hypothesis testing. For sequential hypothesis testing with a fixed length of observations, it is equivalent to a hypothesis testing with repeated observations.*

The optimal decision function with b_N as the argument is of the form

$$\tilde{\delta}_N^*(b_N) = \begin{cases} 0 & \text{if } b_N \geq C_{01}/(C_{10} + C_{01}), \\ 1 & \text{if } b_N < C_{10}/(C_{10} + C_{01}). \end{cases} \tag{3}$$

As **Remark 1** indicates, the equivalence between (2) and (3) holds for every $N \in \mathbb{N}^+$. Thus, we can replace this fixed length N with arbitrary $n \in \mathbb{N}^+$.

The advantage of using b_n is that it enables the sequential extension of hypothesis testing in following sense. At time n, the instantaneous Bayes risk induced by b_n is

$$R_n\left(\tilde{\delta}_n\right) = C_{01} b_n + C_{10}(1 - b_n). \tag{4}$$

By optimizing (8), we obtain exactly (3).

It is worth noting that we can consider b_n as a random variable. Indeed, b_n is a function of $\mathbf{X}_1^n$, thus its randomness is induced by $\mathbf{X}_1^n$. As a function of b_n, R_n also can be considered as a random variable. By observation, we have two fundamental lemmas.

Lemma 1. *The sequences* $\{b_n\}_{n \in \mathbb{N}^+}$ *is martingale.*

This lemma can be proved by computing $\mathbb{E}[b_{n+1}|b_n]$. Let $R_n = \min_{\tilde{\delta}_n} R_n\left(\tilde{\delta}_n\right)$.

Lemma 2. *The Bayes risk satisfies the following:*

1) R_n *is a concave function in* b_n *for every* $n \in \mathbb{N}^+$*;*
2) *The sequences* $\{R_n\}_{n \in \mathbb{N}^+}$ *is supmartingale, i.e.,*

$$\mathbb{E}\left[R_{n+1}|b_n\right] \leq R_n.$$

Proof. See Appendix. B. □

These two lemmas above are indispensable for the analysis of the game defined in the following sections.

III. Stopping Time Game

In this section, we discuss the stopping time game in hypothesis testing. Assume that there are two players, P1 and P2, both of which execute hypothesis testing. The hypotheses of P1 are H_0^1 and H_1^1, and the hypotheses of P1 are H_0^2 and H_1^2. P1 receives observations $\{x_n\}_{n \in \mathbb{N}^+}$, which is distributed according to

$$H_0^1 \ : \ X_n \ \sim \ f(x_n|H_0^1) := f_0(x_n),$$
$$H_1^1 \ : \ X_n \ \sim \ f(x_n|H_1^1) := f_1(x_n).$$

Likewise, P1 receives observations $\{y_n\}_{n \in \mathbb{N}^+}$, which is distributed according to

$$H_0^2 \ : \ Y_n \ \sim \ f(y_n|H_0^2) := g_0(y_n),$$
$$H_1^2 \ : \ Y_n \ \sim \ f(y_n|H_1^2) := g_1(y_n).$$

Moreover, we assume that the hypotheses of P1 and P2 are distinct.

Apart from devising the decision function, they also design stopping rules. Their objective functions are composed of two parts: the Bayes risk and the cost induced by the order stopping. Formally, their cost functions are given by

$$J^1(\tilde{\delta}^1_{\tau^1}, \tau^1, \tau^2) = R^1_{\tau^1}\left(\tilde{\delta}^1_{\tau^1}\right)\mathbb{1}_{\{\tau^1 \le \tau^2\}} + G^1\mathbb{1}_{\{\tau^1 > \tau^2\}}, \quad (5)$$

and

$$J^2(\tilde{\delta}^2_{\tau^2}, \tau^1, \tau^2) = R^2_{\tau^2}\left(\tilde{\delta}^2_{\tau^2}\right)\mathbb{1}_{\{\tau^2 \le \tau^1\}} + G^2\mathbb{1}_{\{\tau^2 > \tau^1\}}, \quad (6)$$

respectively. The costs G_1 and G_2 capture the cost incurred in the case where the other player stops first. The costs R^1_n and R^2_n are the instantaneous Bayes risk defined in Section II:

$$R^1_n\left(\tilde{\delta}^1_n\right) = C^1_{01}b^1_n + C^1_{10}(1 - b^1_n), \quad (7)$$

and

$$R^2_n\left(\tilde{\delta}^2_n\right) = C^2_{01}b^2_n + C^2_{10}(1 - b^2_n). \quad (8)$$

Here, b^1_n and b^2_n are the Bayesian beliefs of P1 and P2, which are determined by $\mathbf{x}^n_1$ and $\mathbf{x}^n_2$, respectively.

Based on the analysis in Section II, we note that for the sequences of observations of different lengths, the optimal decision function $\tilde{\delta}^{1*}_n$ (or $\tilde{\delta}^{2*}_n$) stays unchanged. With a slight abuse of notation, define

$$R^1_n := \min_{\tilde{\delta}^1_n} R^1_n\left(\tilde{\delta}^1_n\right), \quad \text{and} \quad R^2_n := \min_{\tilde{\delta}^1_n} R^2_n\left(\tilde{\delta}^2_n\right).$$

Similarly, define

$$\tilde{J}^i(\tau^1, \tau^2) = \min_{\tilde{\delta}^i_{\tau^i}} J^i(\tilde{\delta}^i_{\tau^i}, \tau^1, \tau^2), \quad i = 1, 2.$$

Definition 1. *(Randomized Stopping Times) Randomized stopping times for a strategy $\mathbf{p} = \{p_n\}_{n\in\mathbb{N}^+} \in \mathcal{P}$ and for a strategy $\mathbf{q} = \{q_n\}_{n\in\mathbb{N}^+} \in \mathcal{Q}$ are defined as*

$$\tilde{\tau}^1(\mathbf{p}) = \inf\{n \ge 1 : A_n \le p_n\},$$

and

$$\tilde{\tau}^2(\mathbf{q}) = \inf\{n \ge 1 : B_n \le q_n\},$$

respectively. Here $\{A_n\}_{n\in\mathbb{N}^+}$ and $\{B_n\}_{n\in\mathbb{N}^+}$ are i.i.d. random variables taking values in $[0, 1]$.

With a slight abuse of notation, let

$$\tilde{J}^1(\mathbf{p}, \mathbf{q}) := \tilde{J}^1(\tilde{\tau}^1(\mathbf{p}), \tau^2(\mathbf{q})),$$

and

$$\tilde{J}^2(\mathbf{p}, \mathbf{q}) := \tilde{J}^2(\tilde{\tau}^2(\mathbf{p}), \tau^2(\mathbf{q})).$$

A. Dynkin's Game

Before we point out the relation of our formulation with Dynkin's game, we first define the filtration at time n as $\mathcal{F}_n$. The filtration $\mathcal{F}_n$ contains all the information up to time n. More specifically,

$$\mathcal{F}_n = \sigma(\pi^1_0, x_1, ..., x_n, \pi^2_0, y_1, ..., y_n),$$

where $\sigma(\cdot)$ is the σ-algebra generator. π^1_0 and π^2_0 are the prior probabilities of H^1_0 and H^2_0, respectively. We first define the truncated feasible sets of the randomized stopping times:

$$\mathcal{P}^N_n = \left\{ \mathbf{p}^N_n \ : \ \begin{array}{l} p_m \in [0,1] \text{ and is adapted to } \mathcal{F}_m, \\ \forall\ n \le m \le N-1, \\ p_N = 1 \end{array} \right\},$$

and

$$\mathcal{Q}^N_n = \left\{ \mathbf{q}^N_n \ : \ \begin{array}{l} q_m \in [0,1] \text{ and is adapted to } \mathcal{F}_m, \\ \forall\ n \le m \le N-1, \\ q_N = 1 \end{array} \right\}.$$

As we can see from the definition of these feasible sets, $q_N = 1$ and $p_N = 1$ force the players to stop at time N. Denote $\mathcal{P}_n = \lim_{N\to\infty} \mathcal{P}^N_n$ and $\mathcal{Q}_n = \lim_{N\to\infty} \mathcal{Q}^N_n$. We are ready to give the formal definition of the equilibrium in this game.

Definition 2. *(Nash Equilibrium) A pair of randomized stopping time strategies $(\mathbf{p}^*, \mathbf{q}^*)$ is said to be a Nash equilibrium point if the following are satisfied:*

$$\mathbb{E}\left[\tilde{J}^1(\mathbf{p}^*, \mathbf{q}^*)\right] = \sup_{\mathbf{p}\in\mathcal{P}_0} \mathbb{E}\left[\tilde{J}^1(\mathbf{p}, \mathbf{q}^*)\right],$$

and

$$\mathbb{E}\left[\tilde{J}^2(\mathbf{p}^*, \mathbf{q}^*)\right] = \sup_{\mathbf{q}\in\mathcal{Q}_0} \mathbb{E}\left[\tilde{J}^2(\mathbf{p}^*, \mathbf{q})\right].$$

Definition 3. *(Equilibrium Point) A pair of stopping times (τ^{1*}, τ^{2*}) is said to be a Nash equilibrium point if the following are satisfied:*

$$\mathbb{E}\left[\tilde{J}^1(\mathbf{p}^*, \mathbf{q}^*)|\mathcal{F}_n\right] = \operatorname*{ess\,inf}_{\mathbf{p}\in\mathcal{P}_n} \mathbb{E}\left[\tilde{J}^1(\mathbf{p}, \mathbf{q}^*)|\mathcal{F}_n\right],$$

and

$$\mathbb{E}\left[\tilde{J}^2(\mathbf{p}^*, \mathbf{q}^*)|\mathcal{F}_n\right] = \operatorname*{ess\,inf}_{\mathbf{q}\in\mathcal{Q}_n} \mathbb{E}\left[\tilde{J}^2(\mathbf{p}^*, \mathbf{q})|\mathcal{F}_n\right].$$

Before we proceed to the main results, we explain the relation between the game considered in this work and Dynkin's game. Dynkin's game is a zero-sum stopping game, in which the players' costs (rewards) are modeled as three stochastic processes, each one of which corresponds to an order of stopping. We extend the classical Dynkin's game to a nonzero-sum setting [10]. In this extended Dynkin's game, the costs functions for P1 and P2 are given by

$$\bar{J}^1(\mathbf{p}, \mathbf{q}) = X^1_{\tilde{\tau}^1(\mathbf{p})}\mathbb{1}_{\{\tilde{\tau}^1(\mathbf{p}) < \tilde{\tau}^2(\mathbf{q})\}} + W^1_{\tilde{\tau}^1(\mathbf{p})}\mathbb{1}_{\{\tilde{\tau}^1(\mathbf{p}) = \tilde{\tau}^2(\mathbf{q})\}} + Y^1_{\tilde{\tau}^2(\mathbf{p})}\mathbb{1}_{\{\tilde{\tau}^1(\mathbf{p}) > \tilde{\tau}^2(\mathbf{q})\}},$$

and

$$\bar{J}^2(\mathbf{p}, \mathbf{q}) = X^2_{\tilde{\tau}^2(\mathbf{p})}\mathbb{1}_{\{\tilde{\tau}^2(\mathbf{p}) < \tilde{\tau}^1(\mathbf{q})\}} + W^2_{\tilde{\tau}^2(\mathbf{p})}\mathbb{1}_{\{\tilde{\tau}^2(\mathbf{p}) = \tilde{\tau}^1(\mathbf{q})\}} + Y^2_{\tilde{\tau}^1(\mathbf{p})}\mathbb{1}_{\{\tilde{\tau}^2(\mathbf{p}) > \tilde{\tau}^1(\mathbf{q})\}},$$

respectively. Moreover, we assume that each of the six stochastic processes is integrable[1]. It is clear that the game is a variant of Dynkin's game. In the following subsection, we give a generalized theorem to provide sufficient conditions to characterize equilibrium points in this game.

[1]**Integrability**: A random variable X is said to be integrable if the following is satisfied

$$\mathbb{E}[|X|] < \infty.$$

B. Finite-Horizon Case

We first consider the finite-horizon case, where both players are forced to stopping at the terminal time N. Define the bisequence $\left\{\left(\alpha_n^N,\beta_n^N\right)\right\}_{n=1}^N$, as the following. For $n=N$,

$$\left(\alpha_n^N,\beta_n^N\right)=(W_N^1,W_N^2),$$

and for $n=1,2,...,N-1$,

$$\left(\alpha_n^N,\beta_n^N\right)=\mathrm{VAL}\begin{bmatrix}(W_n^1,W_n^2) & (X_n^1,Y_n^2)\\ (Y_n^1,X_n^2) & \left(\mathbb{E}[\alpha_{n+1}^N|\mathcal{F}_n],\mathbb{E}[\beta_{n+1}^N|\mathcal{F}_n]\right)\end{bmatrix}.$$

The operator $\mathrm{VAL}(\cdot)$ operators on a bimatrix and it generates the pair of values of a bimatrix game defined by this bimatrix.

Now we present the main result of this section. The following proof is by extending the proof in [12].

Theorem 1. *For $n=1,2,...,N$, the bisequence (p_n,q_n) which constitutes (α_n^N,β_n^N) is an equilibrium point for the finite-horizon stopping time game.*

Proof. We commence by showing that for given $\mathbf{q}^*$, the following is satisfied:

$$\mathbb{E}\left[\operatorname*{ess\,inf}_{\mathbf{p}\in\mathcal{P}_n^N}\mathbb{E}\left[\bar{J}^1(\mathbf{p},\mathbf{q}^*)|\mathcal{F}_n\right]\middle|\mathcal{F}_{n-1}\right]=\operatorname*{ess\,inf}_{\mathbf{p}\in\mathcal{P}_n^N}\mathbb{E}\left[\bar{J}^1(\mathbf{p},\mathbf{q}^*)|\mathcal{F}_{n-1}\right] \tag{9}$$

In order to prove (9), we first note that $\forall\ \mathbf{p}\in\mathcal{P}_n^N,\forall\ \mathbf{q}\in\mathcal{Q}_n^N,\forall\ n\in\{1,2,...,N\}$,

$$\bar{J}^i(\mathbf{p},\mathbf{q})\le\max_{1\le n\le N}\left\{X_n^i+Y_n^i+W_n^i\right\},\quad i=1,2.$$

With the assumption of X_n^i, Y_n^i, and W_n^i, $i=1,2$, $n\in\{1,2,...,N\}$, all being integrable, $\bar{J}^i(\mathbf{p},\mathbf{q})$ is also integrable. Construct a new random variable as follows: for $\tilde{\mathbf{p}},\bar{\mathbf{p}}\in\mathcal{P}_n^N$,

$$\bar{\tau}^1(\mathbf{p})=\begin{cases}\tilde{\tau}^1(\tilde{\mathbf{p}}) & \text{if } \mathbb{E}\left[\bar{J}^1(\tilde{\mathbf{p}},\mathbf{q}^*)|\mathcal{F}_n\right]\ge\mathbb{E}\left[\bar{J}^1(\bar{\mathbf{p}},\mathbf{q}^*)|\mathcal{F}_n\right],\\ \tilde{\tau}^1(\bar{\mathbf{p}}) & \text{if } \mathbb{E}\left[\bar{J}^1(\tilde{\mathbf{p}},\mathbf{q}^*)|\mathcal{F}_n\right]<\mathbb{E}\left[\bar{J}^1(\bar{\mathbf{p}},\mathbf{q}^*)|\mathcal{F}_n\right].\end{cases}$$

Then clearly, $\mathbb{E}\left[\bar{J}^1(\mathbf{p},\mathbf{q}^*)\right]$ is integrable. So, $\mathbf{p}\in\mathcal{P}_n^N$. Therefore, for every $\tilde{\mathbf{p}},\bar{\mathbf{p}}\in\mathcal{P}_n^N$, there exists $\mathbf{p}\in\mathcal{P}_n^N$, such that

$$\mathbb{E}[\bar{J}^1(\mathbf{p},\mathbf{q}^*)|\mathcal{F}_n]=\min\left\{\mathbb{E}\left[\bar{J}^1(\tilde{\mathbf{p}},\mathbf{q}^*)|\mathcal{F}_n\right],\mathbb{E}\left[\bar{J}^1(\bar{\mathbf{p}},\mathbf{q}^*)|\mathcal{F}_n\right]\right\}.$$

Then, by definition, $\mathbb{E}\left[\bar{J}^1(\mathbf{p},\mathbf{q}^*)|\mathcal{F}_n\right]$ is an upwards directed set[2]. By **Proposition VI-1-1** in [13], there exists an increasing sequence $\{f_m\}_{m\in\mathbb{N}^+}$ such that

$$\operatorname*{ess\,inf}_{\mathbf{p}\in\mathcal{P}_n^N}\mathbb{E}\left[\bar{J}^1(\mathbf{p},\mathbf{q}^*)|\mathcal{F}_n\right]=\lim_{m\to\infty}f_m,\quad \text{a.s.},$$

where

$$f_m=\mathbb{E}\left[\bar{J}^1(\mathbf{p}_m,\mathbf{q}^*)|\mathcal{F}_n\right].$$

[2]**Upwards Directed Set**: A set F is said to be upwards directed if for all the $f_1,f_2\in F$, there exists $f_3\in F$ such that

$$f_3\ge f_1\quad\text{and}\quad f_3\ge f_2,\quad \text{a.s.}$$

And by dominated convergence theorem, (9) is prove.

Subsequently, we want to show that $\alpha_n^1=\mathbb{E}\left[\bar{J}^1(\mathbf{p}^*,\mathbf{q}^*)|\mathcal{F}_n\right],\ \forall\ n\in\{1,2,...,N\}$. By mathematical backwards induction, at time N,

$$\alpha_N^N=W_N^1=\mathbb{E}\left[\bar{J}^1(p_N^*,q_N^*)|\mathcal{F}_N\right].$$

Assume that at time $n+1$, the following holds:

$$\begin{aligned}\alpha_{n+1}^N&=\mathbb{E}\left[\bar{J}^1(\mathbf{p}^*,\mathbf{q}^*)|\mathcal{F}_{n+1}\right]\\&=\operatorname*{ess\,inf}_{\mathbf{p}\in\mathcal{P}_{n+1}^N}\mathbb{E}\left[\bar{J}^1(\mathbf{p},\mathbf{q}^*)|\mathcal{F}_{n+1}\right],\\ \beta_{n+1}^N&=\mathbb{E}\left[\bar{J}^2(\mathbf{p}^*,\mathbf{q}^*)|\mathcal{F}_{n+1}\right]\\&=\operatorname*{ess\,inf}_{\mathbf{q}\in\mathcal{Q}_{n+1}^N}\mathbb{E}\left[\bar{J}^2(\mathbf{p}^*,\mathbf{q})|\mathcal{F}_{n+1}\right].\end{aligned}$$

Define $(\mathrm{VAL_I}(\cdot),\mathrm{VAL_{II}}(\cdot))=\mathrm{VAL}(\cdot)$; i.e., given a bimatrix game, the operator $\mathrm{VAL_I}(\cdot)$ generates the value of the first player and $\mathrm{VAL_{II}}(\cdot)$ generates the value of the second player. At time n,

$$\begin{aligned}&\alpha_n^N\\&=\mathrm{VAL}_1\begin{bmatrix}W_n^1 & X_n^1\\ Y_n^1 & \mathbb{E}[\alpha_{n+1}^N|\mathcal{F}_n]\end{bmatrix}\\&=\mathrm{VAL}_1\begin{bmatrix}W_n^1 & X_n^1\\ Y_n^1 & \mathbb{E}\left[\operatorname{ess\,inf}_{\mathbf{p}\in\mathcal{P}_{n+1}^N}\mathbb{E}\left[R^1(\mathbf{p},\mathbf{q}^*)|\mathcal{F}_{n+1}\right]|\mathcal{F}_n\right]\end{bmatrix}\\&=\mathrm{VAL}_1\begin{bmatrix}W_n^1 & X_n^1\\ Y_n^1 & \operatorname{ess\,inf}_{\mathbf{p}\in\mathcal{P}_{n+1}^N}\mathbb{E}\left[R^1(\mathbf{p},\mathbf{q}^*)|\mathcal{F}_n\right]\end{bmatrix}\\&=\operatorname*{ess\,inf}_{\mathbf{p}\in\mathcal{P}_n^N}\mathbb{E}\left[\bar{J}^1(\mathbf{p},\mathbf{q}^*)|\mathcal{F}_n\right].\end{aligned}$$

The first equality is by the construction of α_n^N. The second equality is by the assumption of the induction. The third equality is justified by (9). The induction is complete.

By similar arguments, we have

$$\beta_n^N=\operatorname*{ess\,inf}_{\mathbf{q}\in\mathcal{Q}_n^N}\mathbb{E}\left[R^2(\mathbf{p}^*,\mathbf{q})|\mathcal{F}_n\right].$$

Then the theorem follows. □

The following corollary gives the exact value of (α_n^N,β_n^N) in the hypothesis testing game when the strategies are fully randomized case.

Corollary 2. *If $G^1>\frac{C_{01}^1C_{10}^1}{C_{01}^1+C_{10}^1}$ and $G^2>G^1>\frac{C_{01}^1C_{10}^1}{C_{01}^1+C_{10}^1}$, then the bimatrix game admits full randomized strategies at each time instant. Besides,*

$$\left(\alpha_n^N,\beta_n^N\right)=\left(R_n^1,R_n^2\right),\quad n=1,2,...,N.$$

Proof. We prove this corollary by mathematical induction. At time $N-1$, by the indifference principle [1], the pair of equilibrium point is given by

$$q_{N-1}^*=\frac{R_{N-1}^1-\mathbb{E}[R_N^1|\mathcal{F}_{N-1}]}{R_{N-1}^1-\mathbb{E}[R_N^1|\mathcal{F}_{N-1}]+G^1-R_{N-1}^1}$$

By the assumption of this corollary,

$$G^1>\frac{C_{01}^1C_{10}^1}{C_{01}^1+C_{10}^1}\ge R_n^1,\quad n=1,2,...,N.$$

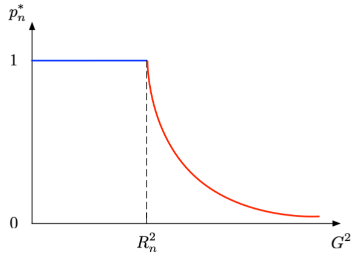

Fig. 1. p_n^* versus G^2 when $G^1 > R_n^1$: the red solid line stands for the fully randomized case and the blue solid line stands for the fully pure case .

Moreover, by **Lemma 2**, we can show that $0 \leq q_{N-1}^* \leq 1$. The similar arguments hold for p_{N-1}^*, where

$$p_{N-1}^* = \frac{R_{N-1}^2 - \mathbb{E}[R_N^2|\mathcal{F}_{N-1}]}{R_{N-1}^2 - \mathbb{E}[R_N^2|\mathcal{F}_{N-1}] + G^2 - R_{N-1}^2}$$

The value of the game for the first player at time $N-1$ is

$$\begin{aligned}\alpha_{N-1}^N = \ & p_{N-1}^* q_{N-1}^* R_{N-1}^1 + p_{N-1}^*(1-q_{N-1}^*)G^1 \\ & + (1-p_{N-1}^*)(1-q_{N-1}^*)\mathbb{E}\left[R_N|\mathcal{F}_{N-1}\right] \\ & + (1-p_{N-1}^*)q_{N-1}^* R_{N-1}^1 \\ = \ & R_{N-1}^1.\end{aligned}$$

Similarly,

$$\beta_{N-1}^N = R_{N-1}^2.$$

Now, we assume that at time $n+1$

$$\left(\alpha_{n+1}^N, \beta_{n+1}^N\right) = \left(R_{n+1}^1, R_{n+1}^2\right).$$

Then, at time n,

$$q_n^* = \frac{R_n^1 - \mathbb{E}[R_{n+1}^1|\mathcal{F}_n]}{R_n^1 - \mathbb{E}[R_{n+1}^1|\mathcal{F}_n] + G^1 - R_n^1}.$$

Directly following similar procedures at time $N-1$, we have

$$\left(\alpha_n^N, \beta_n^N\right) = \left(R_n^1, R_n^2\right).$$

And

$$0 \leq p_n^* \leq 1, \quad 0 \leq q_n^* \leq 1.$$

Hence the corollary follows. □

We give a full characterization of the equilibrium strategies.

1) When $G^1 \leq R_n^1$:
This game admits pure equilibria, $(p_n^*, q_n^*) = (1,1)$, as there exists dominant strategies.
2) When $G^1 > R_n^1$: We present the figure of p_n^* versus G_2. We see that, when $G^2 > R_n^2$, the strategies of P1 are based on the costs of P2 entirely. **what does it mean? This implies that P1 behaves according to the P2's attitude towards stopping**, which is quantified by G^2. Mathematically, at time n, P1 chooses to stop the game with probability p_n^*, which is given by

$$p_n^* = \frac{R_n^2 - \mathbb{E}[R_{n+1}^2|\mathcal{F}_n]}{G^2 - \mathbb{E}[R_{n+1}^2|\mathcal{F}_n]}.$$

As G^2 increases, P1 chooses to stop less likely. Similar arguments hold for P2.

C. Infinite-Horizon Case

In this subsection, we discuss the case where the time horizon is infinite.

Theorem 2. *As n goes to infinity, R_n^1 converges almost surely; i.e.,*

$$R_\infty^1 = \lim_{n\to\infty} R_n^1, \quad a.s.$$

Moreover,

$$R_\infty^1 = \limsup_{n\to\infty} R_n^1 = \frac{C_{01}C_{10}}{C_{01}+C_{10}}.$$

Proof. As $0 \leq R_n^1 \leq \max\{C_{01}, C_{10}\} < \infty$, for $n \in \mathbb{N}$,

$$\sup_{n\in\mathbb{N}^+} \mathbb{E}\left[|R_n^1|\right] < \infty.$$

The rest of the proof follows *Doob's Martingale Convergence Theorem.* □

By leveraging the theorem above, we can rewrite the first term in $\tilde{J}^1(\mathbf{p},\mathbf{q})$ as

$$\begin{aligned}R_{\tau^1}^1 \mathbb{1}_{\{\tau^1\leq\tau^2\}} &= R_{\tau^1}^1 \mathbb{1}_{\{\tau^1\leq\tau^2<\infty\}} + R_\infty^1 \mathbb{1}_{\{\tau^1=\tau^2=\infty\}} \\ &= R_{\tau^1}^1 \mathbb{1}_{\{\tau^1\leq\tau^2<\infty\}} + \frac{C_{01}C_{10}}{C_{01}+C_{10}} \mathbb{1}_{\{\tau^1=\tau^2=\infty\}}.\end{aligned}$$

Also, $G^1 > R_\infty^1$.

Define

$$(\alpha_n, \beta_n) = \left(\mathbb{E}\left[\tilde{J}^1(\mathbf{p}^*,\mathbf{q}^*)|\mathcal{F}_n\right], \mathbb{E}\left[\tilde{J}^2(\mathbf{p}^*,\mathbf{q}^*)|\mathcal{F}_n\right]\right).$$

Then (α_n, β_n) can be found by solving

$$(\alpha_n, \beta_n) = \text{VAL}\begin{bmatrix}(R_n^1, R_n^2) & (R_n^1, G^2) \\ (G^1, R_n^2) & (\mathbb{E}[\alpha_{n+1}|\mathcal{F}_n], \mathbb{E}[\beta_{n+1}|\mathcal{F}_n])\end{bmatrix}.$$

The main result is presented in the following theorem.

Theorem 3. *The bi-sequence $\{(p_n, q_n)\}_{n\in\mathbb{N}^+}$ which constitutes $\{(\alpha_n, \beta_n)\}_{n\in\mathbb{N}^+}$ is an equilibrium point for the infinite horizon stopping time game.*

Proof. Let (q_n^*, p_n^*) be the associated strategy. $\tau^* = \min\{\tau^1(p^*), \tau^2(q^*)\}$ that constitutes (α_n, β_n). Then, for $n \leq m \leq \tau^*$,

$$\begin{aligned}&\mathbb{E}\left[J^1(\mathbf{p}^*,\mathbf{q}^*)\right|\mathcal{F}_n] - \alpha_n \\ &= \mathbb{E}\left[\left(\prod_{k=n}^{m}(1-p_k^*)(1-q_k^*)\right)\right. \\ &\qquad \left.\cdot \left(\mathbb{E}\left[\tilde{J}^1(\mathbf{p}^*,\mathbf{q}^*)|\mathcal{F}_{m+1}\right] - \alpha_{m+1}\right)\middle|\mathcal{F}_n\right].\end{aligned}$$

By letting $m \to \infty$, we obtain

$$\lim_{m\to\infty} \mathbb{E}\left[\tilde{J}^1(\mathbf{p}^*,\mathbf{q}^*)|\mathcal{F}_{m+1}\right] = R_\infty^1,$$

and

$$\begin{aligned}&\lim_{m\to\infty} \alpha_m \\ &= \lim_{m\to\infty} \text{VAL}_\text{I}\begin{bmatrix}(R_m^1, R_m^2) & (R_m^1, G^2) \\ (G^1, R_m^2) & (\mathbb{E}[\alpha_{m+1}|\mathcal{F}_m], \mathbb{E}[\beta_{m+1}|\mathcal{F}_m])\end{bmatrix} \\ &= \text{VAL}_\text{I}\begin{bmatrix}(R_\infty^1, R_\infty^2) & (R_\infty^1, G^2) \\ (G^1, R_\infty^2) & (\lim_{m\to\infty}\alpha_m, \lim_{m\to\infty}\beta_m)\end{bmatrix},\end{aligned}$$

where the second equality is justified by **Theorem 4**. Define

$$\alpha_\infty = \lim_{m\to\infty} \alpha_m, \quad \text{and} \quad \beta_\infty = \lim_{m\to\infty} \beta_m.$$

Then

$$\alpha_\infty = \mathrm{VAL_I} \begin{bmatrix} (R^1_\infty, R^2_\infty) & (R^1_\infty, G^2) \\ (G^1, R^2_\infty) & (\alpha_\infty, \beta_\infty) \end{bmatrix}.$$

By solving this fixed point equation, we obtain that

$$(\alpha_\infty, \beta_\infty) = \left(R^1_\infty, R^2_\infty\right).$$

Thus,

$$\mathbb{E}\left[J^1(\mathbf{p}^*, \mathbf{q}^*)\right] |\mathcal{F}_n] - \alpha_n = 0.$$

Then we can prove that it coincides the definition of equilibrium point. □

IV. Conclusions and Future Work

A. Conclusions

In this work, we have studied the stopping time game in sequential hypothesis testing using an extended Dynkin's formulation. We have showed that the equilibrium points can be found backwards via two intertwined dynamic programming equations. For finite-horizon game, we have provided sufficient conditions of finding the equilibrium points. Besides, we have fully characterized the equilibrium strategy structure. And we have provided the condition under which the game admits fully randomized equilibrium strategies. In the infinite-horizon case, we have presented the theorem which also gives sufficient condition of finding the equilibrium points. We have also showed that in the fully-randomized case, the randomized strategy of one player entirely depends on the other player's costs.

B. Future Work

1) Information-Asymmetric Game: In an information-asymmetric game, the information is divided into two parts, *common information* and *private information*. In the game defined in this paper, there only exists common information as both players share the same sequence of filtration $\{\mathcal{F}_n\}_{n\in\mathbb{N}^+}$. We would investigate the case where the private information may be induced by delay of obtaining observation. For example, at time n, one player has the access to contemporary information $\mathcal{F}_n$ and the other player experiences one-step-delay and only has access to $\mathcal{F}_{n-1}$. A new concept of equilibrium is indispensable as the standard *Nash equilibrium* concept is defined under the assumption that both players only have common information.

2) Stopping Game with Self-Stopping: Another possible work can be explored in the case where the random processes defined in Dynkin's game do NOT satisfy $\sup_{n\in\mathbb{N}^+} \mathbb{E}\left[|\cdot|\right] < \infty$. Particularly, in hypothesis testing, we are interested in the case where there exists cost of taking observations. For example, the cost function of P1 is

$$J^1(\tau^1, \tau^2) = (R^1_{\tau^1} + \tau^1)\mathbb{1}_{\{\tau^1 \le \tau^2\}} + G^1 \mathbb{1}_{\{\tau^1 > \tau^2\}}.$$

Obviously, $\sup_{n\in\mathbb{N}^+} \mathbb{E}\left[|R^1_n + n|\right]$ is not bounded anymore. In this case, without the presence of P2, P1 himself will stop, as when P2 is absent (set $\tau^2 = \infty$), the problem reduces to SPRT.

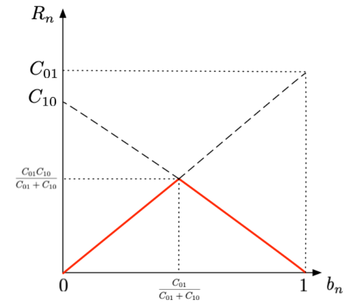

Fig. 2. R_n versus b_n: the red dotted line stands for the value of R_n.

References

[1] Peng G, Lu I.-T. Comparison of four super-resolution techniques for complex spectra or points estimation. In 2017 IEEE Long Island Systems, Applications and Technology Conference (LISAT) 2017 May 5 (pp. 1-5).

[2] Moulin P, Veeravalli V. *Statistical Inference for Engineers and Data Scientists*. Cambridge University Press, 2018 Oct 31.

[3] Tao Z, Zhu Q. On Incentive Compatibility in Dynamic Mechanism Design With Exit Option in a Markovian Environment. arXiv:1909.13720

[4] Gao R, Xie L, Xie Y, Xu H. Robust hypothesis testing using wasserstein uncertainty sets. In Advances in Neural Information Processing Systems 2018 (pp. 7902-7912).

[5] Levy BC. Principles of signal detection and parameter estimation. Springer Science & Business Media; 2008 Jul 7.

[6] Pawlick J, Colbert E, Zhu Q. A game-theoretic taxonomy and survey of defensive deception for cybersecurity and privacy. ACM Computing Surveys (CSUR). 2019 Aug 30; 52(4):82.

[7] Hamadene S, Zhang J. The continuous time nonzero-sum Dynkin game problem and application in game options. SIAM Journal on Control and Optimization. 2010 Mar 5; 48(5):3659-69.

[8] Cvitanic J, Karatzas I. Backward stochastic differential equations with reflection and Dynkin games. The Annals of Probability. 1996; 24(4):2024-56.

[9] Ohtsubo Y. A nonzero-sum extension of Dynkin's stopping problem. Mathematics of Operations Research. 1987 May; 12(2):277-96.

[10] Kim K, Shamma J. Multi-agent sequential hypothesis testing. In 53rd IEEE Conference on Decision and Control 2014 Dec 15 (pp. 1942-1947).

[11] Maschler M, Solan E, and Zamir E. Game Theory. Cambridge University Press, 2013.

[12] Yasuda M. On a randomized strategy in Neveu's stopping problem. Stochastic Processes and their Applications. 1985 Dec. 1; 21(1):159-66.

[13] Neveu J. *Discrete-parameter martingales*. Elsevier; 1975.

Appendix A
Proof of Corollary 1

By observation, we have $b_n = b_0/(b_0 + (1-b_0)L(\mathbf{x}_1^n))$. Thus, there exits one-to-one correspondence between b_N and $L(\mathbf{x}_1^N)$ for $N \in \mathbb{N}^+$.

Appendix B
Proof of Lemma 2

For the purpose of proof, with a slight abuse of notation, let $R_n(b_n) = R_n$.

1) The first statement is illustrated as shown in Fig. 2.
2) As R_k is a concave function in b_n, using Jensen's inequality, we obtain

$$\mathbb{E}\left[R_{n+1}(b_{n+1})|b_n\right] \le R_{n+1}\left(\mathbb{E}\left[b_n|b_n\right]\right) = R_n(b_n).$$