2024 IEEE 63rd Conference on Decision and Control (CDC)
December 16-19, 2024. MiCo, Milan, Italy

Information Compression in Dynamic Information Disclosure Games

Dengwang Tang and Vijay G. Subramanian

2024 IEEE 63rd Conference on Decision and Control (CDC) | 979-8-3503-1633-9/24/$31.00 ©2024 IEEE | DOI: 10.1109/CDC56724.2024.10886413

***Abstract*— We consider a two-player dynamic information design problem for a game played between a principal and a receiver on top of a Markovian system controlled by the receiver's actions. The principal strategically obtains and shares some information about the underlying system with the receiver in order to influence their actions, and agents' instantaneous rewards depend only on the system state and receiver actions. In our game, both players have long-term objectives, and the principal sequentially commits to their strategies instead of at the beginning—at every turn the principal can choose randomized experiments to observe the system partially. The principal can share details about the experiments to the receiver. In our analysis the *truthful disclosure* rule is assumed—the principal is required to truthfully announce each experiment detail and result to the receiver immediately after the result is revealed. Based on the received information, when its their turn the receiver takes an action which influences the state of the underlying system. Using a constructive backward inductive procedure, we show that there exists a Perfect Bayesian Equilibrium in this game where both agents play Canonical Belief Based (CBB) strategies using a compressed version of their information, rather than their full information, to choose experiments (for the principal) or actions (for the receiver).**

I. INTRODUCTION

In many modern engineering and socioeconomic problems and systems, such as cyber-security, transportation networks, and e-commerce, information asymmetry is an inevitable aspect that crucially impacts decision making. In these systems, agents need to decide on their actions under limited information about the system and each other. In many situations, agents can overcome (some of) the information asymmetry by communicating with each other. However, agents can be unwilling to share information when agents' goals are not aligned with each other, since having some information that another agent does not know can be an advantage. In general, communication between agents with diverging incentives cannot be naturally established without collectively agreed upon rules/protocols, and all agents suffer due to the breakdown of the information exchange. For example, drug companies are required by regulations to disclose their trial results truthfully. The public can then trust the results and benefit from the drug. In turn, the drug companies can make a profit. Without government regulations, drug companies and the public will both suffer due to mistrust. In many real-world dynamic systems, information exchange and decision making can happen repeatedly as the system/environment changes over time—for example, public companies disclose information periodically which impacts stockholders' decisions; (COVID-19) vaccine producers conduct their trials and release results sequentially which impacts the government's purchasing decisions; during an epidemic, health authorities update their recommendations on the use of face masks over time according to changing levels of infection; etc. Therefore, in the face of information asymmetry, it is important to establish rules/protocols to facilitate repeated information exchange among agents in multi-agent dynamic systems.

In the economics literature, there are two main approaches to the above problem—mechanism design [2] and information design [3]. In mechanism design, less informed agents can extract information from more informed agents by committing beforehand to how they will use the collected information. Whereas in information design, more informed agents can (partially) disclose information to less informed agents. The more informed agents commit on the manner in which they (partially) disclose their information. In both approaches, all agents can benefit from the information exchange. The literature for both falls into two groups: (i) static settings, where both information disclosure and decision making take place only once; and (ii) dynamic settings, where agents repeatedly disclose information and take actions over time on top of an ever changing environment/physical system. In both cases dynamic settings are more challenging than static settings as agents need to anticipate future information disclosure when taking an action. Dynamic mechanism design has been studied extensively—see [4]–[7]. Most of the works in dynamic information design assume myopic receivers [8]–[14], which greatly simplifies receivers' decision making. There have been a few papers on information design problems where all agents in the system have long-term goals [15]–[25]. These papers typically assume that *the principal commits to their strategy for the whole game at the beginning*, a'la Stackelberg [26]. The bulk of this literature also assumes that the principal observes the underlying state perfectly. However, these assumptions can be inappropriate for many applications. If the protocol gives more informed agents the power to commit to a strategy for the whole time horizon at the beginning, then the more informed agents can implement punishment strategies by threatening to withhold information if the less informed agents do not obey their "instructions"—see Example 1. Thus, the more informed agents could abuse their com-

Full paper [1] at `https://arxiv.org/pdf/2403.12204.pdf`.

This work is supported by NSF Grant No. ECCS 1750041, ECCS 2038416, ECCS 1608361, CCF 2008130, ARO Award No. W911NF-17-1-0232, MIDAS Sponsorship Funds by General Dynamics, and Okawa Foundation Research Grant.

D. Tang is with Ming Hsieh Department of Electrical and Computer Engineering, University of Southern California, Los Angeles, CA 90089, USA dengwang@usc.edu

V. G. Subramanian is with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109, USA vgsubram@umich.edu

mitment power to implement otherwise non-credible threats instead of using it for efficient information disclosure. This is not a desirable outcome: for example, online map services should not threaten to withhold service if a driver refuses to take the recommended route; and similarly, public health authorities may want to use persuasion instead of threats to encourage mask wearing during an epidemic. Again, focusing on the public health setting, during an epidemic the authorities may not know the full extent of the disease spread, but only have an estimate of it (using testing or other methods such as media and social network feeds). In this context, transparency to the public on the part of the authorities—in disclosing measurement methods and data—is important for persuasion based schemes to be effective.

In this work, we focus on the dynamic information design problem that results in a dynamic game between a principal and a receiver on top of a Markovian system. Both the principal and the receiver have long-term objectives. The principal cannot directly (and perfectly) observe [1] the system state, but can choose (randomized) experiments to observe the system partially. The principal is allowed to choose any experiment, but they must announce the experimental setup and results truthfully to the receiver before the receiver takes their action. Both these aspects of our model are motivated by the public health setting described earlier. The receiver takes action on each turn based on the information received to date, which then influences the underlying system and also determines the instantaneous rewards for both agents.

Contributions: In the class of dynamic information disclosure games between a principal and a receiver discussed above, assuming the truthful disclosure rule, we identify equilibria where both players use compressed information based strategies, called Canonical Belief Based (CBB) strategies. Here both agents use strategies based on a compressed version of their full information—distilling it into beliefs—to choose their actions (experiments for the principal, and actions for the receiver). We develop a backward inductive sequential decomposition procedure for an equilibrium in such strategies, and show existence by proving that a solution always exists. Finally, we investigate examples of such games to provide insight into CBB-strategy-based equilibria.

Notation: We use superscripts to indicate agents, and subscripts to indicate time. We use $t_1 : t_2$ to indicate the set of timestamps $(t_1, t_1 + 1, \dots, t_2)$. For random variables or vectors, we use the corresponding script capital letters to denote the space of values these random vectors can take—for example, $\mathcal{H}_t^i$ denotes the space of values the random vector H_t^i can take. We use $\mathbb{P}(\cdot)$ and $\mathbb{E}[\cdot]$ to denote probabilities and expectations, respectively. We use $\Delta(\Omega)$ to denote the set of probability distributions on a set Ω.

The paper is organized as follows. We start with a motivating example in Section I-A that discusses issues with threat strategies based equilibria. The problem is formulated in Section II, and necessary discrete geometrical results are presented in Section III. Our main results are discussed in Section IV with outlines of the proofs. The proof details can be found in [1]. We present some numerical examples in Section V, discuss potential generalizations in Section VI, and conclude in Section VII with a discussion of future work.

A. A Motivating Example

The following is an example where a principal with the power to commit to a strategy for the whole game at the onset (a'la Stackelberg), can use otherwise non-credible threats.

Example 1. Consider a two-stage game of two players: the principal A, and the receiver B. The state of the system at time t is X_t. The states are uncontrolled, and X_1, X_2 are *i.i.d.* uniform random variables taking values in $\{0, 1\}$. The principal can observe X_t at time t while the receiver cannot. At stage t, the principal transmits message M_t to the receiver and the receiver takes an action $U_t \in \{a, b, c, d\}$. The instantaneous payoff for both players are given by

$$r_1^A(0,a) = 1, r_1^A(0,b) = 1.01, r_1^A(0,c) = r_1^A(0,d) = -1000$$
$$r_1^A(1,c) = 1, r_1^A(1,d) = 1.01, r_1^A(1,a) = r_1^A(1,b) = -1000$$
$$r_1^B(0,a) = 500, r_1^B(0,b) = 1, r_1^B(0,c) = r_1^B(0,d) = -1000$$
$$r_1^B(1,c) = 500, r_1^B(1,d) = 1, r_1^B(1,a) = r_1^B(1,b) = -1000$$

and $r_2^A(\cdot,\cdot) = r_2^B(\cdot,\cdot) = r_1^A(\cdot,\cdot)$.

Suppose that the principal has the power to commit to a strategy (g_1, g_2) at the beginning of the game. Then, (given the Stackelberg setting) an optimal strategy for the principal is the following: fully reveal the state at $t = 1$ (i.e. $M_1 = X_1$); if the receiver plays a or c at $t = 1$, then transmit no information at $t = 2$; and if the receiver plays b or d at $t = 1$, then fully reveal the state at $t = 2$. Then, the receiver's best response to the principal's strategy is the following: at $t = 1$: play b if $M_1 = 0$, and play d if $M_1 = 1$; and at time $t = 2$: play a if $M_2 = 0$, and play c if $M_2 = 1$.

In the resulting equilibrium, the principal effectively *threatens* the receiver to comply to their interest at time $t = 1$ by not giving information at time $t = 2$, even though the interests of both parties are aligned at $t = 2$. In fact, without posing a threat to the receiver at time 2, the principal cannot convince them to play b or d at time 1.

II. PROBLEM FORMULATION

We consider a finite-horizon two-player dynamic game between the principal A and the receiver B. The game consists of T stages, where the principal moves before the receiver in each stage. The game features an underlying dynamic system with state X_t. At each time $t \in [T]$, the receiver chooses an action U_t. Then, the system transits to the next state $X_{t+1} \sim P_t(X_t, U_t)$, where $P_t : \mathcal{X}_t \times \mathcal{U}_t \mapsto \Delta(\mathcal{X}_{t+1})$ is the transition kernel. The initial state X_1 has prior distribution $\hat{\pi} \in \Delta(\mathcal{X}_1)$. The initial distribution $\hat{\pi}$ and transition kernels $P = (P_t)_{t=1}^T$ are common knowledge to both players. We assume that neither player can observe the state X_t directly. However, at each time t, the principal can

[1]Generalizing to direct and perfect observations of the (dynamic) system state by the sender (resulting in information asymmetry) is technically challenging as existence of Nash equilibria and sequential refinements thereof, with or without compressing information, for infinite state or action dynamic games [27] is non-trivial or may not hold in great generality.

conduct an *experiment*[2], i.e. choosing an observation kernel, to learn about X_t. We impose the rule that the experiments are required to be public—both agents know the settings (the probabilities in the observation kernel), and the outcome (the observation itself) of the experiment. Specifically, at each time t, the principal chooses an observation kernel $\sigma_t : \mathcal{X}_t \mapsto \Delta(\mathcal{M}_t)$, and announces σ_t to the receiver. The experiment outcome M_t is then realized, and observed by both agents. If the horizon $T = 1$, then we have the classical information design problem [3], so we focus on $T \geq 2$.

Assumption 1. $\mathcal{X}_t, \mathcal{U}_t, \mathcal{M}_t$ are finite sets with $|\mathcal{M}_t|$ sufficiently large.[3]

The order of events happening at time t is given as the following: (1) The principal commits to an experiment σ_t, and announces it to the receiver; (2) The measurement result M_t is revealed to both the principal and receiver; (3) The receiver takes action U_t; and (4) X_t transits to the next state.

Let $\mathcal{S}_t$ be the space of experiments. The principal uses a (pure) strategy to choose their experiment $g_t^A : \mathcal{S}_{1:t-1} \times \mathcal{M}_{1:t-1} \times \mathcal{U}_{1:t-1} \mapsto \mathcal{S}_t$. For convenience, define $\mathcal{H}_t^A = \mathcal{S}_{1:t-1} \times \mathcal{M}_{1:t-1} \times \mathcal{U}_{1:t-1}$. The receiver uses a (pure) strategy $g_t^B : \mathcal{S}_{1:t} \times \mathcal{M}_{1:t} \times \mathcal{U}_{1:t-1} \mapsto \mathcal{U}_t$. For convenience, define $\mathcal{H}_t^B = \mathcal{S}_{1:t} \times \mathcal{M}_{1:t} \times \mathcal{U}_{1:t-1}$. The principal's goal is to maximize $J^A(g) = \mathbb{E}^g \left[\sum_{t=1}^T r_t^A(X_t, U_t) \right]$. The receiver's goal is to maximize $J^B(g) = \mathbb{E}^g \left[\sum_{t=1}^T r_t^B(X_t, U_t) \right]$. The instantaneous reward functions $(r_t^A, r_t^B)_{t=1}^T$ are common knowledge to both agents.

The belief of the principal at time t is a function $\mu_t^A : \mathcal{M}_{1:t-1} \times \mathcal{S}_{1:t-1} \times \mathcal{U}_{1:t-1} \mapsto \Delta(\mathcal{X}_{1:t})$. The belief of the receiver at time t (after knowing σ_t and observing M_t) is a function $\mu_t^B : \mathcal{M}_{1:t} \times \mathcal{S}_{1:t} \times \mathcal{U}_{1:t-1} \mapsto \Delta(\mathcal{X}_{1:t})$.

Inspired by the "mechanism picking game" defined in [30], we call the above game a *signal picking game*, and we will study Perfect Bayesian Equilibria [31] for our game.

Definition 1 (PBE)**.** A Perfect Bayesian Equilibrium [31] is a pair (g, μ), where (i) g is sequentially rational given $\mu = \left(\mu_{1:T}^A, \mu_{1:T}^B\right)$, and (ii) μ is consistent with g, i.e., Bayes law applies for updates if the denominator is non-zero.

III. BACKGROUND: DISCRETE GEOMETRY

In this section, we introduce some notations and results of discrete geometry that are necessary for our main results.

Definition 2. Let f be a real-valued function on a polytope[4] Ω. Then, f is called a (continuous) piecewise linear function if there exist polytopes $C_1, \cdots, C_k$ such that (i) f is linear on each C_j for $j = 1, \cdots, k$; and (ii) $C_1 \cup \cdots \cup C_k = \Omega$.

Lemma 1. *Let Ω_1, Ω_2 be polytopes. Let $\ell : \Omega_1 \mapsto \Omega_2$ be an affine function and $f : \Omega_2 \mapsto \mathbb{R}$ be a piecewise linear function. Then the composite function $f \circ \ell : \Omega_1 \mapsto \mathbb{R}$ is piecewise linear.*

[2]Information design literature [28], [29] deems such experiments *signals*.

[3]We assume a sufficiently large message space to rule out the complicating effect of limited communication bandwidth.

[4]A polytope is the convex hull of a finite set in $\mathbb{R}^d$ where $d < +\infty$.

Proof. See Appendix A of full version [1]. □

Next, we introduce the notion of a triangulation.

Definition 3. [32] Let Ω be a finite dimensional polytope. A *triangulation* γ of Ω is a finite collection of simplices (i.e. convex hulls of affinely independent set of points) such that (1) If a simplex $C \in \gamma$, then all faces of C are in γ; (2) For any two simplices $C_1, C_2 \in \gamma$, $C_1 \cap C_2$ is a (possibly empty) face of C_1; and (3) The union of γ equals Ω.

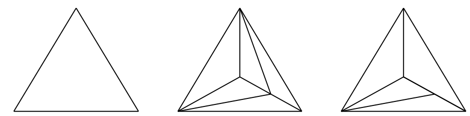

Fig. 1. Left: 2-D Polytope Ω; Center: A triangulation of Ω; Right: NOT a triangulation of Ω.

For a function $f : \Omega \mapsto \mathbb{R}$ and a triangulation γ, let $\mathbb{I}(f, \gamma)$ denote the linear interpolation of f based on the triangulation γ, i.e. $\mathbb{I}(f, \gamma)(\omega) := \alpha_1 f(\omega_1) + \cdots + \alpha_k f(\omega_k)$ if $\omega \in C$, where $C \in \gamma$ is a simplex with vertices $\omega_1, \cdots, \omega_k$, and $\omega = \alpha_1 \omega_1 + \cdots + \alpha_k \omega_k$ for some $\alpha_1, \cdots, \alpha_k \geq 0$ such that $\alpha_1 + \cdots + \alpha_k = 1$.

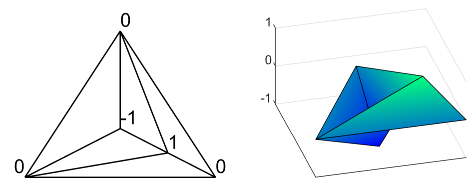

Fig. 2. Left: A triangulation γ labeled with the values of a function f on the vertices. Right: 3-D plot of $\mathbb{I}(f, \gamma)$.

Lemma 2. *For any real-valued function f on a polytope Ω, $\mathbb{I}(f, \gamma)$ is a well-defined, continuous piecewise linear function.*

Proof. See Appendix A of full version [1]. □

For each $\omega \in \Omega$ and triangulation γ, we have shown that there exists a unique way to represent ω as a convex combination of the vertices of one simplex from γ. One can treat this convex combination as a finite measure. Denote this finite measure by $\mathbb{C}(\omega, \gamma)$. Then we have $\mathbb{I}(f, \gamma)(\omega) = \int f(\cdot) \mathrm{d}\mathbb{C}(\omega, \gamma)$.

Definition 4. Let $f : \Omega \mapsto \mathbb{R}$. Its concave closure $\mathrm{cav}(f)$ is defined as a function ρ such that $\rho(\omega) := \sup\{z : (\omega, z) \in \mathrm{cvxg}(f)\}$ for all $\omega \in \Omega$, with $\mathrm{cvxg}(f) \subset \Omega \times \mathbb{R}$ being the convex hull of the graph of f.

For certain functions f, their concave closures can be represented as a triangulation based interpolation of the original function. Define the set of all such triangulations as $\arg\mathrm{cav}(f)$, i.e.

$$\arg\mathrm{cav}(f) := \{\gamma \text{ is a triangulation of } \Omega : \mathbb{I}(f, \gamma) = \mathrm{cav}(f)\}.$$

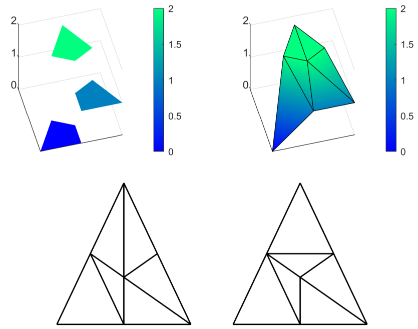

Fig. 3. Top-left: 3-D plot of a function f—an upper semi-continuous piecewise constant function taking values in $\{0, 1, 2\}$. Top-right: Concave closure of f. Bottom-left and bottom-right: 2-D visualization of two different triangulations in $\arg\mathrm{cav}(f)$.

The following lemma identifies a class of functions with the above property.

Lemma 3. *Let $f_1, \cdots, f_k, \rho_1, \cdots, \rho_k$ be continuous piecewise linear functions on a polytope Ω. For $\omega \in \Omega$, define*

$$\Upsilon(\omega) = \underset{j=1,\cdots,k}{\arg\max} f_j(\omega), \text{ and } \Psi(\omega) = \max_{j \in \Upsilon(\omega)} \rho_j(\omega).$$

Then $\arg\mathrm{cav}(\Psi)$ is non-empty, i.e. there exists a triangulation γ of Ω such that the concave closure of Ψ is equal to $\mathbb{I}(\Psi, \gamma)$.

Proof. See Appendix A of full version [1]. □

IV. MAIN RESULTS

In this section, we introduce our main result—Theorem 1—which provides a dynamic programming characterization of a subset of PBE of the signal picking game in Section II.

Note that due to the assumption of public experiments, the signal picking game is a game with symmetric information [5] after each experiment is conducted. The principal's advantages lies in the fact that they have the power to determine the choice of experiments. Thus, standard results on strategy-independence of beliefs (e.g. [33]) imply that the beliefs of both players in this game are strategy-independent, i.e. there is a canonical belief system. Similar strategy-independent belief systems are also constructed and used in [34]. We describe this belief system as follows.

Definition 5. Define the Bayesian update function $\xi_t : \Delta(\mathcal{X}_t) \times \mathcal{S}_t \times \mathcal{M}_t \mapsto \Delta(\mathcal{X}_t)$ by setting for each $x_t \in \mathcal{X}_t$

$$\xi_t(x_t|\pi_t, \sigma_t, m_t) := \frac{\pi_t(x_t)\sigma_t(m_t|x_t)}{\sum_{\tilde{x}_t} \pi_t(\tilde{x}_t)\sigma_t(m_t|\tilde{x}_t)},$$

for all (π_t, σ_t, m_t) such that the denominator is non-zero. When the denominator is zero, $\xi_t(\pi_t, \sigma_t, m_t)$ is defined to be the uniform distribution.

Definition 6. The canonical belief system is a collection of functions $(\kappa_t^A, \kappa_t^B)_{t\in\mathcal{T}}, \kappa_t^i : \mathcal{H}_t^i \mapsto \Delta(\mathcal{X}_t), i \in \{A, B\}$ defined recursively using new information via the following:

[5] As mentioned earlier—in Footnote 1—, there are significant technical challenges in generalizing to asymmetric information settings.

- $\kappa_1^A(h_1^A) := \hat{\pi}$, the prior distribution of X_1;
- $\kappa_t^B(h_t^B) := \xi_t(\kappa_t^A(h_t^A), \sigma_t, m_t)$ where $h_t^B = (h_t^A, \sigma_t, m_t)$;
- $\kappa_{t+1}^A(h_{t+1}^A) := \ell_t(\kappa_t^B(h_t^B), u_t)$, where $h_{t+1}^A = (h_t^B, u_t)$, and $\ell_t : \Delta(\mathcal{X}_t) \times \mathcal{U}_t \mapsto \Delta(\mathcal{X}_{t+1})$ is defined for each $x_{t+1} \in \mathcal{X}_{t+1}$ by
$$\ell_t(\pi_t, u_t)(x_{t+1}) := \sum_{\tilde{x}_t \in \mathcal{X}_t} \pi_t(\tilde{x}_t) P_t(x_{t+1}|\tilde{x}_t, u_t),$$
for all $\pi_t \in \Delta(\mathcal{X}_t), u_t \in \mathcal{U}_t$.

We consider a subclass of strategies for both the principal and the receiver, namely the CBB strategies [6], wherein agent $i \in \{A, B\}$ chooses their experiment or action, respectively, at time t based solely on beliefs $\Pi_t^i = \kappa_t^i(H_t^i)$ instead of H_t^i.

Let $\lambda_t^A : \Delta(\mathcal{X}_t) \mapsto \mathcal{S}_t$ be the CBB strategy of the principal, and $\lambda_t^B : \Delta(\mathcal{X}_t) \mapsto \mathcal{U}_t$ be the CBB strategy of the receiver. Then, saying that player i is using CBB strategy λ_t^i is equivalent to saying that they are using the strategy

$$g_t^i(h_t^i) = \lambda_t^i(\kappa_t^i(h_t^i)), \qquad \forall h_t^i \in \mathcal{H}_t^i.$$

Given an experiment and a distribution on the state, the posterior belief of the receiver is a random variable (a function of the random outcome). In an information disclosure game, it is helpful to consider the following sub-problem: *how to design an experiment such that the receiver's belief, as a random variable, follows a certain distribution?* The next definition formalizes this concept, which is used in classical one-shot information design setting [3] as well.

Definition 7. [3] An experiment $\sigma_t \in \mathcal{S}_t$ is said to *induce a distribution* $\eta \in \Delta_f(\Delta(\mathcal{X}_t))$—that is, η is a distribution with finite support on the set of distributions $\Delta(\mathcal{X}_t)$—*from* $\pi_t \in \Delta(\mathcal{X}_t)$ [3] if for all $\tilde{\pi}_t \in \Delta(\mathcal{X}_t)$,

$$\eta(\tilde{\pi}_t) = \sum_{\tilde{m}_t} \mathbf{1}_{\{\tilde{\pi}_t = \xi_t(\pi_t, \sigma_t, \tilde{m}_t)\}} \sum_{\tilde{x}_t} \sigma_t(\tilde{m}_t|\tilde{x}_t)\pi_t(\tilde{x}_t).$$

A distribution η is said to be *inducible from* π_t if there exists some experiment σ_t that induces η from π_t.

Remark 1. In [3], the authors showed that a distribution is $\eta \in \Delta_f(\Delta(\mathcal{X}_t))$ is inducible from π_t if and only if π_t is the center of mass of η, i.e. $\pi_t = \sum_{\tilde{\pi}_t \in \mathrm{supp}(\eta)} \eta(\tilde{\pi}_t) \cdot \tilde{\pi}_t$.

We now introduce our main result, which describes a backward induction procedure to find a PBE where both players use CBB strategies.

Theorem 1. *Let $V_{T+1}^A(\cdot) = V_{T+1}^B(\cdot) := 0$. For $t = T, T-1, \cdots, 1$ and $\pi_t \in \Delta(\mathcal{X}_t)$, define*

$$\hat{q}_t^i(\pi_t, u_t) := \sum_{\tilde{x}_t} r_t^i(\tilde{x}_t, u_t)\pi_t(\tilde{x}_t) + V_{t+1}^i(\ell_t(\pi_t, u_t)) \quad \forall i \in \{A, B\}; \tag{1a}$$

$$\Upsilon_t(\pi_t) := \arg\max_{u_t} \hat{q}_t^B(\pi_t, u_t); \tag{1b}$$

[6] Neither agent needs to retain the full information—H_t^i for agent i at time t—to obtain the beliefs, and a recursive procedure that uses only past values of the beliefs and new information suffices. Also, the agents cannot recreate the full information from the beliefs (so, they compress information).

$$\hat{v}_t^A(\pi_t) := \max_{u_t \in \Upsilon(\pi_t)} \hat{q}_t^A(\pi_t, u_t); \tag{1c}$$

$$\hat{v}_t^B(\pi_t) := \max_{u_t} \hat{q}_t^B(\pi_t, u_t); \tag{1d}$$

$$\gamma_t \in \arg \operatorname{cav}(\hat{v}_t^A); \tag{1e}$$

$$V_t^i(\pi_t) := \mathbb{I}(\hat{v}_t^i, \gamma_t) \quad \forall i \in \{A, B\}. \tag{1f}$$

Let $\lambda_t^{*B}(\pi_t)$ be any $u_t \in \mathcal{U}_t$ that attains the maximum in (1c). *Let $\lambda_t^{*A}(\pi_t)$ be any experiment that induces the finite measure $\mathbb{C}(\pi_t, \gamma_t)$ from π_t. Then, the CBB strategies $(\lambda^{*A}, \lambda^{*B})$ form (the strategy part of) a PBE, and $V_1^A(\hat{\pi})$ and $V_1^B(\hat{\pi})$ are the equilibrium payoffs for the principal and the receiver respectively in this PBE.*

Proof Outline. We first construct a belief system μ^* that is consistent with any strategy profile—details are in Appendix B of the full version [1]. Hence, we only need to show sequential rationality of λ^*.

To show the receiver's sequential rationality, we prove the following: Fixing the principal's strategy to be λ^{*A}, the receiver is facing an MDP with state Π_t^B and action U_t. The proof follows via standard stochastic control arguments [7].

To show the principal's sequential rationality, we prove the following: Fixing the receiver's strategy to be λ^{*B}, the principal is facing an MDP with state Π_t^A and action Σ_t. This proof follows using standard stochastic control arguments coupled with information design results [3].

The details of the proof are presented in Appendix B of the full version [1]. □

The following proposition states that the sequential decomposition procedure described in Theorem 1 is well defined and always has a solution.

Proposition 1. *There always exists a CBB strategy profile $(\lambda^{*A}, \lambda^{*B})$ that satisfies Eqs.* (1) *in Theorem 1.*

Proof. Proof with induction on time t.

Induction Invariant: V_t^A, V_t^B are well-defined continuous piecewise linear functions.

Induction Base: The induction variant is clearly true for $t = T + 1$ since V_{T+1}^A, V_{T+1}^B are constant functions.

Induction Step: Suppose that the induction invariant holds for $t + 1$, the result for t can be established as follows:

- Step 1: For each $u_t \in \mathcal{U}_t$, using the fact that $\ell_t(\pi_t, u_t)$ is affine in π_t, by Lemma 1, q_t^A, q_t^B are continuous piecewise linear functions in π_t.
- Step 2: By Lemma 3, γ_t is well-defined.
- Step 3: By Lemma 2, V_t^A, V_t^B are continuous piecewise linear functions.

This completes the proof. □

A. Extension

In many real-world settings, receivers have the option to quit the game at any time. Our model and results can be extended to finite horizon games where the receiver can decide to terminate the game at any time before time T.

[7]These are: i) Markov strategies are dominant; and ii) the Bellman recursion can be used to find optimal Markov strategies.

Proposition 2. *Let $\overline{\mathcal{U}}_t \subset \mathcal{U}_t$ be the set of actions that terminates the game at time t. If we define $V_t^i, q_t^i, \lambda_t^{*i}$ for each $i \in \{A, B\}, t \in \mathcal{T}$ as in* (1), *except* (1a) *is changed to*

$$\hat{q}_t^i(\pi_t, u_t) := \sum_{\tilde{x}_t} r_t^i(\tilde{x}_t, u_t)\pi_t(\tilde{x}_t) + \begin{cases} V_{t+1}^i(\ell_t(\pi_t, u_t)) & \text{if } u_t \notin \overline{\mathcal{U}}_t \\ 0 & \text{if } u_t \in \overline{\mathcal{U}}_t \end{cases}$$

for $i \in \{A, B\}$. Then, the CBB strategies $(\lambda^{*A}, \lambda^{*B})$ form (the strategy part of) a PBE, and $V_1^A(\hat{\pi})$ and $V_1^B(\hat{\pi})$ are the equilibrium payoff for the principal and the receiver, respectively, in this PBE.

Proof. Similar to Theorem 1. □

V. EXAMPLES

We implemented the sequential decomposition algorithm of Proposition 2 in MATLAB for binary state spaces (i.e. $|\mathcal{X}_t| = 2$). We ran the algorithm on the following examples of the signal picking game.

Example 2. Consider the quickest detection game defined in [25]. In this game, the underlying state X_t is binary and uncontrolled, with $\mathcal{X}_t = \{1, 2\}$. State 2 is an absorbing state, i.e. $\mathbb{P}(X_{t+1} = 2 \mid X_t = 2) = 1$, whereas the system can jump from state 1 to state 2 at any time with probability p, i.e. $\mathbb{P}(X_{t+1} = 2 \mid X_t = 1) = p$ where $p \in (0, 1)$.

The receiver would like to detect (the epoch of) the jump from state 1 to state 2 as accurately as possible. At each time the receiver has two options: $U_t = j$ stands for declaring state j for $j = 1, 2$. The instantaneous reward of the receiver is given by

$$r_t^B(X_t, U_t) = \begin{cases} -1 & \text{if } X_t = 1, U_t = 2 \\ -c & \text{if } X_t = 2, U_t = 1 \\ 0 & \text{otherwise} \end{cases},$$

where $c \in (0, 1)$. Once the receiver declares state 2, the game ends immediately.

The principal would like the receiver to stay in the system as long as possible. The instantaneous reward for the principal is

$$r_t^A(X_t, U_t) = \begin{cases} 1 & \text{if } U_t = 1 \\ 0 & \text{otherwise} \end{cases}.$$

Setting $p = 0.2$ and $c = 0.1$, we obtain the q_t^B and V_t^A functions specified in Proposition 2 in Figure 4. The horizontal axis represents $\pi_t(1)$. In the figures for V_t^A functions, the vertices of the triangulation γ_t are labeled. The vertices represent the set of beliefs that the principal could induce, and they completely describe the principal's CBB strategy. If the vertex is labeled with red circles, the receiver will take action $U_t = 1$ at this posterior belief. If, instead, the vertex is labeled with blue triangles, the receiver will take action $U_t = 2$ at this posterior belief.

From the figures, one can see that at any stage, there is only one possible belief that the principal would induce which leads to the receiver quitting the game (i.e. select

$U_t = 2$). This is consistent with the principal's objective of keeping the receiver in the system. Just like in static information design problems [3], [35], when it is better for the receiver to declare a change, i.e., quit, under the current belief, the principal promises to tell the receiver that the state is 2 with some probability $\tilde{p}$ when the state is indeed 2, and tell the receiver nothing otherwise. In doing so, the receiver would believe that the state is 1 with a higher probability when the principal does not tell the receiver anything. The principal chooses $\tilde{p}$ to be precisely the value for which the receiver is willing to stay in the system (as in [3]).

When t is close to T, the end of the game, the principal would only prefer to declare state 2 if they believe that $\pi_t(1)$ is very small. This is due to the fact that "false alarms" are costlier than delayed detection in this game. When t is further away from T, the threshold of $\pi_t(1)$ for the principal to declare state 2 becomes larger. This holds because when the game is close to end, the receiver has the "safe" option to declare state 1 (at a small cost) until the end to avoid false alarms (which are costly). However, this option is less preferable when the gap between t and T is large.

When t is further away from T, the principal's value function seems to converge. This is due to the fact that the receiver has the option to quit the game and staying in the game is costly in general.

Example 3. Consider a game between a principal and a detector. In this game, the underlying state X_t is binary and uncontrolled with $\mathcal{X}_t = \{-1, 1\}$. At any time, the system can jump to the other state with probability $p \in (0, 1)$, i.e.

$$\mathbb{P}(X_{t+1} = -j \mid X_t = j) = p, \qquad \forall j \in \{-1, 1\}.$$

The receiver has three actions: $U_t = j$ stands for declaring state j for $j = -1, 1$. Both $U_t = -1$ and $U_t = +1$ terminate the game. In addition, the receiver can choose to wait at a cost with action $U_t = 0$. The instantaneous reward of the receiver for $c \in (0, 1)$ is given by

$$r_t^B(X_t, U_t) = \begin{cases} 1 & \text{if } X_t = U_t \\ -c & \text{if } U_t = 0 \\ 0 & \text{otherwise} \end{cases}.$$

The principal would like the receiver to stay in the system as long as possible. The instantaneous reward for the principal is

$$r_t^A(X_t, U_t) = \begin{cases} 1 & \text{if } U_t = 0 \\ 0 & \text{otherwise} \end{cases}.$$

Setting $p = 0.2$ and $c = 0.15$, we obtain the q_t^B and V_t^A functions specified in Proposition 2—see Figure 5. The horizontal axis represents $\pi_t(-1)$. The figures follows the same interpretation as the figures in Example 2. (The markers for actions are different from previous figures, but they are self-explanatory.)

Different from Example 2, the value functions and CBB strategies at equilibrium oscillate with a period of 4 (given $p = 0.2, c = 0.15$) instead of converging as t gets further away from the horizon T.

VI. DISCUSSION

Naturally, one may consider extending the above result to two settings: (a) when a public noisy observation of the state is available in addition to the principal's experiment; and (b) when there are multiple receivers. However, our result is immediately extendable to neither setting. This is since the techniques we use in this paper depend heavily on the piecewise linear structure of $\hat{q}$ and V-functions in (1), as well as the preservation of this piecewise linear structure under backward induction. Specifically, when the functions $\hat{q}_t^A, \hat{q}_t^B$ are piecewise linear, the concave closure of $\hat{v}_t^A$ can be expressed as a triangulation based interpolation (through Lemma 3), which in turn allows us to apply the same triangulation to $\hat{v}_t^B$, and thus ensuring the continuity and piecewise linearity of V_t^B. However, this structure does not appear in general in the extensions.

We describe an attempt to extend Theorem 1 to settings (a) and (b) in the most straightforward way. In the case of setting (a), one needs to change the belief update in (1a) from $\ell_t(\pi_t, u_t)$ to some other update function that incorporates the public observation. However, unlike $\ell_t(\pi_t, u_t)$, the new update function may not be linear in π_t. Therefore this procedure cannot preserve piecewise linear properties.

In the case of setting (b), u_t will represent a vector of actions of all receivers, and one needs to change the definition of $\Upsilon_t(\pi_t)$ in (1b) to be the set of mixed strategy Nash equilibrium (or alternatively correlated equilibrium) action profiles of the following stage game: Receiver i chooses an action in $\mathcal{U}_t^i$, and receives payoff $\hat{q}_t^i(\pi_t, u_t)$. In this setting, $\Upsilon_t(\pi_t)$ is a set of probability measures on the product set $\mathcal{U}_t$. The new $\hat{v}_t^A$ function is then be given by

$$\hat{v}_t^A(\pi_t) = \operatorname*{arg\,max}_{\eta_t \in \Upsilon_t(\pi_t)} \sum_{\tilde{u}_t} q_t^A(\pi_t, \tilde{u}_t)\eta_t(\tilde{u}_t).$$

However, in this case, continuity and piecewise linearity of $\hat{q}_t$ are not enough to ensure that the value function V_t^A possesses the same property. To see this, consider the following example with two receivers B and C. Let $\mathcal{U}_t^B = \mathcal{U}_t^C = \mathcal{X}_t = \{1, 2\}$. Let $p = \pi_t(1)$. Then all functions of π_t can be expressed as a function of p. Suppose that

$$q_t^B(\pi_t, u_t) = \begin{cases} 1 & \text{if } u_t^B = u_t^C \\ 0 & \text{otherwise} \end{cases},$$

$$q_t^C(\pi_t, u_t) = \begin{cases} p + 1 & \text{if } u_t^B = 1, u_t^C = 2 \\ 1 & \text{if } u_t^B = 2, u_t^C = 1 \\ 0 & \text{otherwise} \end{cases}.$$

It can be verified that, under either the concept of Nash equilibrium or correlated equilibrium, $\Upsilon_t(\pi_t)$ contains only one element: player B plays action 1 with probability $\frac{1}{2+p}$, and player C plays their two actions with equal probability independent of player B's action. Now suppose that

$$q_t^A(\pi_t, u_t) = \begin{cases} p & u_t^B = 1 \\ 0 & \text{otherwise} \end{cases}.$$

Then we have $\hat{v}_t^A(\pi_t) = \frac{p}{2+p}$ for $p \in [0,1]$. Observe that $\hat{v}_t^A$ is a strictly concave function. Hence the concave closure of $\hat{v}_t^A$ is just $\hat{v}_t^A$ itself, which is not piecewise linear.

VII. Conclusion and Future Work

In this work, we formulated a dynamic information disclosure game, called the signal picking game, where the principal sequentially commits to a signal/experiment to communicate with the receiver. We showed that there exist equilibria where both the principal and the receiver make decisions based on canonical beliefs instead of their respective full information. We also provided a sequential decomposition procedure to find such equilibria.

Unlike the CIB-belief-based sequential decomposition procedures of [36]–[39], the sequential decomposition procedure of Theorem 1 always has a solution. The main reason is that the CIB belief in the signal picking game is strategy-independent, just as in [34].

There are a few future research directions arising from this work. The first is to extend the result to infinite horizon games. The second is to extend it to multiple senders settings. Another direction would be to explore different methodology for the extensions discussed in Section VI. Finally, it would be useful to apply the tools developed in [27] to study asymmetric information settings of our game.

References

[1] D. Tang and V. G. Subramanian, "Information compression in dynamic information disclosure games," *arXiv preprint arXiv:2403.12204*, 2024.

[2] R. B. Myerson, "Mechanism design," in *Allocation, Information and Markets*. Springer, 1989, pp. 191–206.

[3] E. Kamenica and M. Gentzkow, "Bayesian persuasion," *American Economic Review*, vol. 101, no. 6, pp. 2590–2615, 2011.

[4] D. Bergemann and J. Välimäki, "The dynamic pivot mechanism," *Econometrica*, vol. 78, no. 2, pp. 771–789, 2010.

[5] S. Athey and I. Segal, "An efficient dynamic mechanism," *Econometrica*, vol. 81, no. 6, pp. 2463–2485, 2013.

[6] A. Pavan, I. Segal, and J. Toikka, "Dynamic mechanism design: A Myersonian approach," *Econometrica*, vol. 82, no. 2, pp. 601–653, 2014.

[7] D. Bergemann and J. Välimäki, "Dynamic mechanism design: An introduction," *Journal of Economic Literature*, vol. 57, no. 2, pp. 235–74, 2019.

[8] D. Lingenbrink and K. Iyer, "Optimal signaling mechanisms in unobservable queues with strategic customers," in *Proceedings of ACM EC*, 2017, pp. 347–347.

[9] J. C. Ely, "Beeps," *American Economic Review*, vol. 107, no. 1, pp. 31–53, 2017.

[10] F. Farokhi, A. M. Teixeira, and C. Langbort, "Estimation with strategic sensors," *IEEE Transactions on Automatic Control*, vol. 62, no. 2, pp. 724–739, 2016.

[11] M. O. Sayin, E. Akyol, and T. Başar, "Strategic control of a tracking system," in *IEEE CDC*, 2016, pp. 6147–6153.

[12] J. Renault, E. Solan, and N. Vieille, "Optimal dynamic information provision," *Games and Economic Behavior*, vol. 104, pp. 329–349, 2017.

[13] J. Best and D. Quigley, "Honestly dishonest: A solution to the commitment problem in Bayesian persuasion," Economics Group, Nuffield College, University of Oxford, Tech. Rep., 2016.

[14] ——, "Persuasion for the long-run," Economics Group, Nuffield College, University of Oxford, Tech. Rep., 2016.

[15] P. B. Luh, S.-C. Chang, and T.-S. Chang, "Solutions and properties of multi-stage Stackelberg games," *Automatica*, vol. 20, no. 2, pp. 251–256, 1984.

[16] B. Tolwinski, "A Stackelberg equilibrium for continuous-time differential games," in *IEEE CDC*, 1983, pp. 675–681.

[17] ——, "Closed-loop Stackelberg solution to a multistage linear-quadratic game," *Journal of Optimization Theory and Applications*, vol. 34, no. 4, pp. 485–501, 1981.

[18] F. Farhadi, D. Teneketzis, and S. J. Golestani, "Static and dynamic informational incentive mechanisms for security enhancement," in *IEEE ECC*, 2018, pp. 1048–1055.

[19] L. Li, O. Massicot, and C. Langbort, "Sequential public signaling in routing games with feedback information," in *IEEE CDC*, 2018, pp. 2735–2740.

[20] M. O. Sayin and T. Başar, "Dynamic information disclosure for deception," in *IEEE CDC*, 2018, pp. 1110–1117.

[21] M. O. Sayin, E. Akyol, and T. Başar, "Hierarchical multistage Gaussian signaling games in noncooperative communication and control systems," *Automatica*, vol. 107, pp. 9–20, 2019.

[22] M. O. Sayin and T. Başar, "On the optimality of linear signaling to deceive Kalman filters over finite/infinite horizons," in *International Conference on Decision and Game Theory for Security*. Springer, 2019, pp. 459–478.

[23] E. Meigs, F. Parise, A. Ozdaglar, and D. Acemoglu, "Optimal dynamic information provision in traffic routing," *arXiv preprint arXiv:2001.03232*, 2020.

[24] H. Tavafoghi and D. Teneketzis, "Informational incentives for congestion games," in *2017 Allerton Conference*, 2017, pp. 1285–1292.

[25] F. Farhadi and D. Teneketzis, "Dynamic information design: A simple problem on optimal sequential information disclosure," *Dynamic Games and Applications*, vol. 12, no. 2, pp. 443–484, 2022.

[26] H. V. Stackelberg, *Market Structure and Equilibrium*. Springer, 2011.

[27] R. B. Myerson and P. J. Reny, "Perfect conditional ε-equilibria of multi-stage games with infinite sets of signals and actions," *Econometrica*, vol. 88, no. 2, pp. 495–531, 2020.

[28] M. Gentzkow and E. Kamenica, "Bayesian persuasion with multiple senders and rich signal spaces," *Games and Economic Behavior*, vol. 104, pp. 411–429, 2017.

[29] ——, "Disclosure of endogenous information," *Economic Theory Bulletin*, vol. 5, no. 1, pp. 47–56, 2017.

[30] L. Doval and V. Skreta, "Mechanism design with limited commitment," *arXiv preprint arXiv:1811.03579*, 2018.

[31] D. Fudenberg and J. Tirole, "Perfect Bayesian equilibrium and sequential equilibrium," *Journal of Economic Theory*, vol. 53, no. 2, pp. 236–260, 1991.

[32] J. De Loera, J. Rambau, and F. Santos, *Triangulations: Structures for algorithms and applications*. Springer Science & Business Media, 2010, vol. 25.

[33] P. R. Kumar and P. Varaiya, *Stochastic systems: Estimation, identification and adaptive control*. Prentice-Hall, Inc., 1986.

[34] A. Nayyar, A. Gupta, C. Langbort, and T. Başar, "Common information based Markov perfect equilibria for stochastic games with asymmetric information: Finite games," *IEEE Transactions on Automatic Control*, vol. 59, no. 3, pp. 555–570, 2013.

[35] D. Bergemann and S. Morris, "Information design: A unified perspective," *Journal of Economic Literature*, vol. 57, no. 1, pp. 44–95, 2019.

[36] Y. Ouyang, H. Tavafoghi, and D. Teneketzis, "Dynamic games with asymmetric information: Common information based perfect Bayesian equilibria and sequential decomposition," *IEEE Transactions on Automatic Control*, vol. 62, no. 1, pp. 222–237, 2016.

[37] H. Tavafoghi, Y. Ouyang, and D. Teneketzis, "On stochastic dynamic games with delayed sharing information structure," in *IEEE CDC*, 2016, pp. 7002–7009.

[38] D. Vasal, A. Sinha, and A. Anastasopoulos, "A systematic process for evaluating structured perfect Bayesian equilibria in dynamic games with asymmetric information," *IEEE Transactions on Automatic Control*, vol. 64, no. 1, pp. 81–96, 2019.

[39] D. Tang, H. Tavafoghi, V. Subramanian, A. Nayyar, and D. Teneketzis, "Dynamic games among teams with delayed intra-team information sharing," *Dynamic Games and Applications*, vol. 13, no. 1, pp. 353–411, 2023.

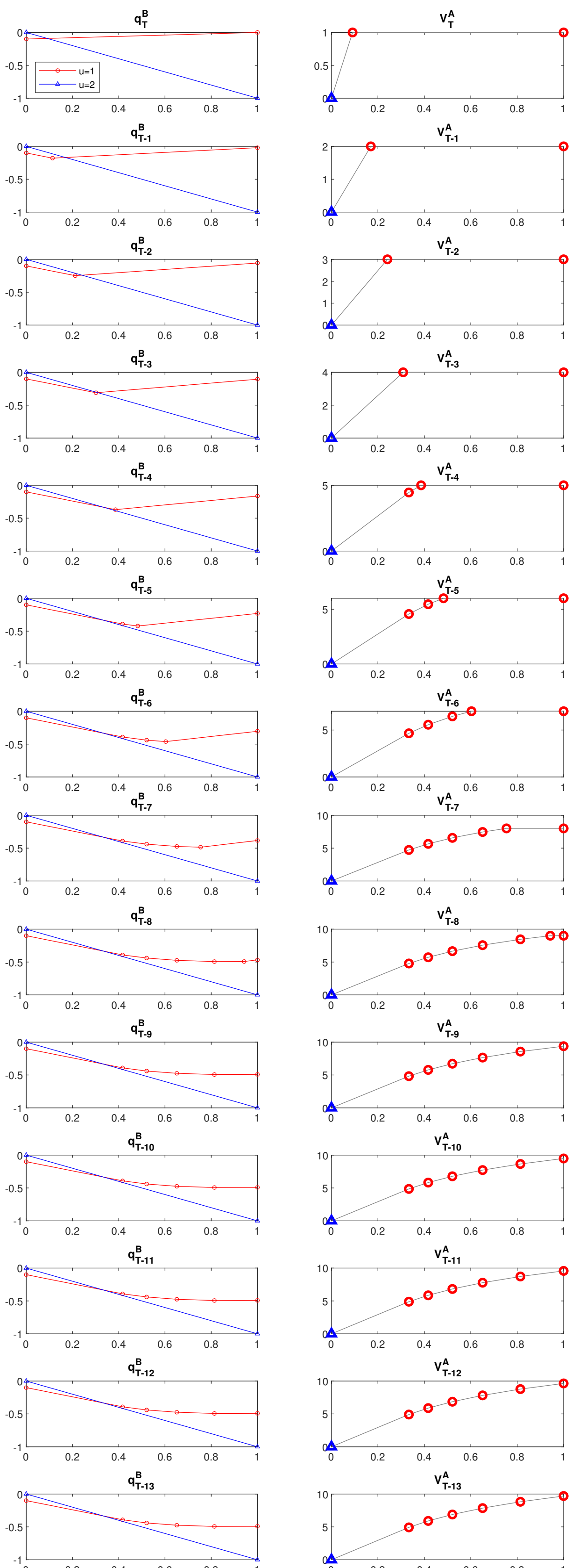

Fig. 4. The q_t^B and V_t^A functions for Example 2 with $p = 0.2, c = 0.1$ at times $t = T : T - 13$.

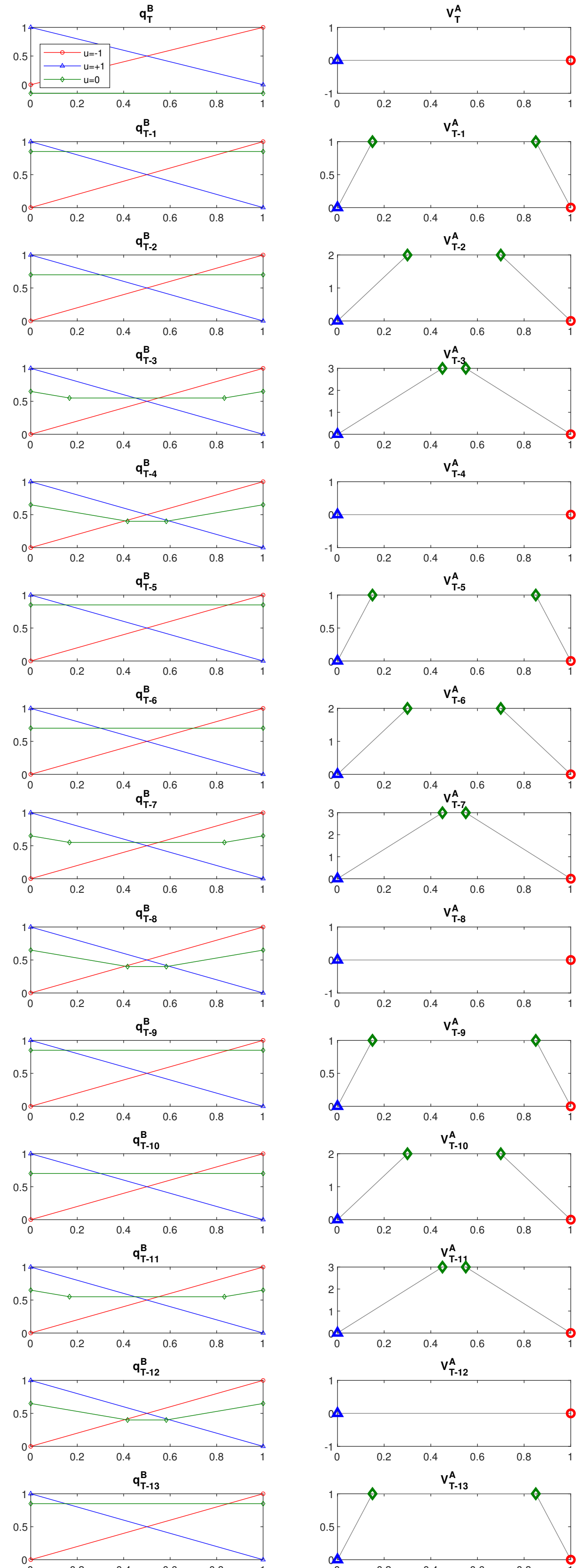

Fig. 5. The q_t^B and V_t^A functions for Example 3 with $p = 0.2, c = 0.15$ at times $t = T : T - 13$.