Achieving Efficient Collaboration in Decentralized Heterogeneous Teams using Common-Pool Resource Games

Piyush Gupta Shaunak D. Bopardikar Vaibhav Srivastava

Abstract— **We consider a team of heterogeneous agents that is collectively responsible for servicing and subsequently reviewing a stream of homogeneous tasks. Each agent (autonomous system or human operator) has an associated mean service time and mean review time for servicing and reviewing the tasks, respectively, which are based on their expertise and skill-sets. The team objective is to collaboratively maximize the number of "serviced and reviewed" tasks. To this end, we formulate a Common-Pool Resource (CPR) game and design utility functions to incentivize collaboration among team-members. We show the existence and uniqueness of the Pure Nash Equilibrium (PNE) for the CPR game. Additionally, we characterize the structure of the PNE and study the effect of heterogeneity among the agents at the PNE. We show that the formulated CPR game is a best response potential game for which both sequential best response dynamics and simultaneous best reply dynamics converge to the Nash equilibrium. Finally, we numerically illustrate the price of anarchy for the PNE.**

I. Introduction

Success of a project is often contingent upon effective and efficient collaboration among members of diverse, dynamic, digital and dispersed teams [1]. An effective collaboration requires each team-member to efficiently work on their tasks while backing up other team-members by monitoring and providing review and feedback. Such team backup behavior improves team performance by mitigating the lack of certain skills in some team-members. Lack of incentives to backup other members may result in team-members operating individually and a poor team performance. Therefore, for effective team performance, it is imperative to design appropriate incentives that facilitate collaboration among the agents while ensuring that their individual performance does not suffer.

In this paper, we study incentive design mechanisms to facilitate aforementioned team backup behavior among the heterogeneous agents. In particular, we pose this problem as a CPR game [2, 3] and design utilities that yield the desired behavior. CPR games is a class of resource sharing games in which the players jointly manage a common pool of resource and make strategic decisions to maximize their utilities.

Human-team-supervised autonomy is a class of motivating problems for our setup. Queueing theory has emerged as a popular paradigm to study these problems [4–7]. However, these works predominantly consider a single human operator. There have been limited studies on human-team-supervised autonomy. These include simulation based studies [8–10], ad hoc design [11], or non-interacting operators [12]. Here, we focus on a game-theoretic approach to study one of the key features of the human-team-supervised autonomy: the team backup behavior, which refers to the extent to which team-members help each other perform their roles [9].

We model team backup behavior in the following way. We consider an unlimited supply of tasks from which each team-member may admit tasks for servicing at a constant rate. We assume that each serviced task is stored in a common review pool for a second review. Each team-member can choose to spend a fraction of their time to review tasks from the common review pool and provide backup to improve the quality. But without any incentives, members may not choose to review the tasks as it may affect their individual performance. We focus on design of incentives, within the CPR game formalism, to facilitate the team backup behavior.

Our formulation has features similar to the CPR game studied in [3, 13, 14]. In these works, authors utilize prospect theory to capture the risk aversion behavior of the players investing into a fragile CPR that fails if there is excessive investment in it. In case of its failure, no player receives any return from the CPR. While our design of the common review pool is similar to the fragile CPR, our failure model incorporates the constraint that only serviced tasks can be reviewed. In contrast to the agent heterogeneity due to prospect-theoretic risk preferences in [3], heterogeneity in our model arises due to differences in agents' mean service and review times.

The major contributions of this work are fourfold. First, we present a novel formulation of team backup behavior and design incentives, within the CPR game formalism, to facilitate such behavior. Second, we show existence and uniqueness of the PNE for the proposed game. Third, we show that the proposed game is a best response potential game [15], for which both sequential best response dynamics [16] and simultaneous best reply dynamics [17] converge to the PNE. Thus, the policies of self-interested agents in a decentralized team will converge to the PNE. Finally, we numerically illustrate the inefficiency of the PNE using Price of Anarchy and show its variation as a function of heterogeneity.

The rest of the paper is structured in the following way. In Section II, we describe our problem, pose it as a CPR game, and design utilities that facilitate the team backup behavior. In Section III, we show the existence and uniqueness of the PNE and show that the proposed game is a best response potential game for which the best response dynamics converge to the Nash equilibrium. Numerical illustrations showing the effects of heterogeneity on the Price of Anarchy are discussed in Section IV. Finally, we conclude in Section V.

This work has been supported by NSF Award IIS-1734272.

Piyush Gupta (guptapi1@msu.edu), Shaunak D. Bopardikar (shaunak@egr.msu.edu) and Vaibhav Srivastava (vaibhav@egr.msu.edu) are with Department of Electrical and Computer Engineering, Michigan State University, East Lansing, Michigan, 48824, USA.

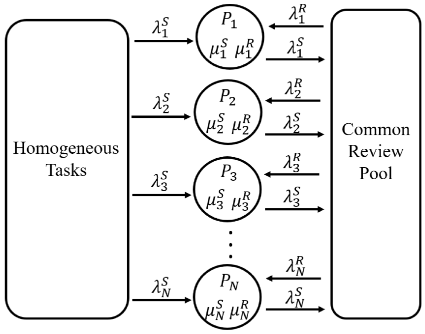

Fig. 1: Player i devotes her time to service homogeneous tasks (at a constant service admission rate λ_i^S) while reviewing the serviced tasks from the common review pool (at a constant review admission rate λ_i^R). The maximum admission rate for player i for servicing and reviewing the tasks is given by μ_i^S and μ_i^R, respectively.

II. Background and Problem Formulation

In this section, we describe the problem setup and formulate the problem using a game-theoretic framework. We also present some preliminaries that will be used in the paper.

A. Problem Description

We consider a heterogeneous team of $N \in \mathbb{N}$ agents tasked with servicing a stream of homogeneous tasks. These agents could be autonomous systems or human operators. Each task after getting serviced by a team-member gets stored in a common review pool for a second review. This second review is a feedback process in which any team-member can re-examine the serviced task from the common review pool for performance monitoring and quality assurance purposes. Each agent $i \in \mathcal{N} = \{1, \dots, N\}$ may choose to spend a portion of her time to review the tasks from the common pool while spending her remaining time to service the incoming tasks. We consider heterogeneity among the agents due to the difference in their level of expertise and skill-sets in servicing and reviewing the tasks. This heterogeneity is captured by the average service time $(\mu_i^S)^{-1} \in \mathbb{R}_{>0}$ and average review time $(\mu_i^R)^{-1} \in \mathbb{R}_{>0}$ spent by the agent $i \in \mathcal{N}$ on servicing and reviewing a task, respectively.

Let $\lambda_i^S \in [0,\ \mu_i^S]$ and $\lambda_i^R \in [0,\ \mu_i^R]$ be the deterministic service and review admission rates, i.e., the rates at which the agent i chooses to admit tasks for servicing and reviewing, respectively. We assume that each agent i can choose their admission rates independent of the other agents. The range of λ_i^S and λ_i^R have been chosen to satisfy the stability conditions (including marginal stability) for the service and review queues for each agent $i \in \mathcal{N}$ [18, Chapter 8].

Suppose the agent i selects λ_i^S and λ_i^R as their service and review admission rates, then

$$\frac{\lambda_i^S}{\mu_i^S} + \frac{\lambda_i^R}{\mu_i^R} \le 1,$$

where $\frac{\lambda_i^S}{\mu_i^S}$ (respectively, $\frac{\lambda_i^R}{\mu_i^R}$) is the average time the agent spends on servicing (respectively, reviewing) the tasks within a unit time. Thus, if the agent has selected a review admission rate λ_i^R, then the service admission rate satisfies

$$\lambda_i^S \le \mu_i^S - \left(\frac{\mu_i^S}{\mu_i^R}\right)\lambda_i^R. \tag{1}$$

We will assume that the agents operate at their maximum capacity and equality holds in (1). Fig. 1 shows the schematic of our problem setup. Notice that every task reviewed by the agents arrives from the common review pool of serviced tasks. Therefore, the total review admission rate, i.e., the rate at which tasks get reviewed from the common review pool, given by $\sum_{i=1}^N \lambda_i^R$, is upper bounded by the total arrival rate into the common review pool given by $\sum_{i=1}^N \lambda_i^S$. Therefore,

$$\sum_{i=1}^N \lambda_i^R \le \sum_{i=1}^N \lambda_i^S. \tag{2}$$

Eq. (2) captures the constraint that only serviced tasks are available for review. By substituting (1) in (2), we obtain,

$$\sum_{i=1}^N a_i \lambda_i^R \le \sum_{i=1}^N \mu_i^S, \tag{3}$$

where $a_i := \left(1 + \frac{\mu_i^S}{\mu_i^R}\right)$. Eq. (3) represents the system constraint on the review admission rate chosen by the agents.

We are interested in incentivizing the collaboration among the agents for better team performance. Towards this end, we propose a game-theoretic setup defined below.

B. A Common-Pool Resource Game Formulation

We now formulate our problem in a Common-Pool Resource (CPR) game setup. Henceforth, we would refer to each agent as a player. A maximum service admission rate μ_i^S and a maximum review admission rate μ_i^R is associated with each player i based on her skill-set and level of expertise. Without loss of generality, let the players be ordered in the increasing order of the ratio of their maximum service admission rate to the maximum review admission rate, i.e., $\frac{\mu_1^S}{\mu_1^R} \le \frac{\mu_2^S}{\mu_2^R} \le \dots \le \frac{\mu_N^S}{\mu_N^R}$.

Let $S_i := [0,\ \mu_i^R]$ be a non-empty convex and compact strategy set for each player i, from which the player chooses its admission rate λ_i^R for reviewing the tasks from the common review pool. Once the player i decides her review admission rate $\lambda_i^R \in S_i$, her service admission rate for servicing the tasks λ_i^S is given by the right hand side of (1) with equality. Let $S = \prod_{i \in \mathcal{N}} S_i$ be the joint strategy space of all the players, where $\prod$ denotes the Cartesian product. Furthermore, we define $S_{-i} = \prod_{j \in \mathcal{N}, j \ne i} S_j$ as the joint strategy space of all the players except player i.

For brevity of notation, we denote the total service admission rate and the total review admission rate by $\lambda_T^S = \sum_{i=1}^N \lambda_i^S$ and $\lambda_T^R = \sum_{i=1}^N \lambda_i^R$, respectively. Similarly, $\mu_T^S = \sum_{i=1}^N \mu_i^S$ and $\mu_T^R = \sum_{i=1}^N \mu_i^R$ denote the aggregated sum of the maximum service admission rates and the maximum review admission rates of all the players, respectively.

Each player i receives a constant reward $r^S \in \mathbb{R}_{>0}$ for servicing each task. Hence, the service utility $u_i^S : S_i \mapsto \mathbb{R}_{>0}$ for each player i servicing the tasks at λ_i^S is given by:

$$u_i^S = \lambda_i^S r^S. \tag{4}$$

To incentivize collaboration among the agents, we design the review utility $u_i^R : S \mapsto \mathbb{R}_{>0}$ received for reviewing the tasks from the common review pool using two functions: a rate of return, $r^R : S \mapsto \mathbb{R}_{>0}$ for each reviewed task and a probability of failure $p : S \mapsto [0,1]$.

Let $x \in \mathbb{R}$ defined by $x = \lambda_T^S - \lambda_T^R = \mu_T^S - \sum_{i=1}^N a_i \lambda_i^R$ be the slackness parameter for the system constraint (3). The constraint (3) is violated for the negative values of x. The slackness parameter characterizes the gap between the total service admission rate and the total review admission rate for all the players. In order to maximize the high quality team throughput, i.e., the number of tasks that are both serviced and reviewed, it is desired to incentivize the team to operate close to $x = 0$.

We assume that the rate of return r^R and the failure probability function p both depend on the strategy of all the players only through the slackness parameter x. Furthermore, we assume that r^R is strictly decreasing in x. Since for each $x \in [0,\ \mu_T^S]$, the system constraint (3) is satisfied, the rate of return is maximized at $x = 0$. Such choice may correspond to a scenario in which, e.g., an employer generates higher revenue based on the high quality throughput of her company, i.e., when the team efficiently reviews all the serviced tasks, which she redistributes among her employees as an incentive based on their contribution to the review process.

Since the system constraint (3) is a hard constraint that must be satisfied at all times, the failure probability function $p(x) = 1$ if the system constraint (3) gets violated, i.e., the slackness parameter $x < 0$. We assume that the failure probability p is non-increasing in x, and approaches 1 as x approaches 0. If the common review pool fails, then $u_i^R = 0$ for each player i. Therefore, we define the utility u_i^R by

$$u_i^R(\lambda_i^R, \lambda_{-i}^R) = \begin{cases} 0, & \text{with probability } p(\lambda_i^R, \lambda_{-i}^R), \\ \lambda_i^R r^R(\lambda_i^R, \lambda_{-i}^R), & \text{otherwise.} \end{cases} \tag{5}$$

Let $u_i(\lambda_i^R, \lambda_{-i}^R) = u_i^S + u_i^R$ be the total utility of the player $i \in \mathcal{N}$. Each player i tries to maximize her expected utility $\tilde{u}_i : S \mapsto \mathbb{R}$ defined by

$$\begin{aligned} \tilde{u}_i &= \mathbb{E}[u_i^S + u_i^R], \\ &= \lambda_i^S r^S + \lambda_i^R r^R(\lambda_i^R, \lambda_{-i}^R)(1 - p(\lambda_i^R, \lambda_{-i}^R)), \end{aligned} \tag{6}$$

where the expectation is computed over the failure event. Since r^R and p depend on the review admission rates of all players only through the slackness parameter x, with a slight abuse of notation, we express $r^R(\lambda_i^R, \lambda_{-i}^R)$ and $p(\lambda_i^R, \lambda_{-i}^R)$ by $r^R(x)$ and $p(x)$, respectively. Substituting (1) in (6), yields

$$\begin{aligned} \tilde{u}_i &= \mu_i^S r^S + \lambda_i^R \left[r^R(x)(1 - p(x)) - \left(\frac{\mu_i^S}{\mu_i^R} \right) r^S \right], \\ &=: \mu_i^S r^S + \lambda_i^R f_i(x), \end{aligned} \tag{7}$$

where $f_i : S \mapsto \mathbb{R}$ is defined by

$$f_i(\lambda_i^R, \lambda_{-i}^R) = f_i(x) = r^R(x)(1 - p(x)) - \left(\frac{\mu_i^S}{\mu_i^R} \right) r^S. \tag{8}$$

The function f_i is the incentive for the player i to review the tasks. Note that the player i will choose a non-zero λ_i^R if and only if she has a positive incentive to review the tasks, i.e., $f_i(x) > 0$. Otherwise, the player i drops out without reviewing any task from the common review pool ($\lambda_i^R = 0$) and focuses solely on servicing of tasks ($\lambda_i^S = \mu_i^S$), thereby maximizing her expected utility given by $\tilde{u}_i = \mu_i^S r^S$.

In the following, we will refer to the above CPR game by $\Gamma = (\mathcal{N}, \{S_i\}_{i \in \mathcal{N}}, \{\tilde{u}_i\}_{i \in \mathcal{N}})$. In this paper, we are interested in the equilibrium strategies for the players that are characterized by a PNE defined below.

*Definition 1 (**Pure Nash Equilibrium**):* A PNE is a strategy profile $\lambda^{R^*} = \{\lambda_i^{R^*}\}_{i \in \mathcal{N}} \in S$, such that for each player $i \in \mathcal{N}$, $\tilde{u}_i(\lambda_i^{R^*}, \lambda_{-i}^{R^*}) \geq \tilde{u}_i(\lambda_i^R, \lambda_{-i}^{R^*})$, for any $\lambda_i^R \in S_i$.

Let $b_i : S_{-i} \mapsto S_i$ defined by

$$b_i(\lambda_{-i}^R) \in \underset{\lambda_i^R \in S_i}{\operatorname{argmax}}\ \tilde{u}_i(\lambda_i^R,\ \lambda_{-i}^R),$$

be a *best response* of the player i to the review admission rates of the other players λ_{-i}^R. A PNE exists if and only if there exists an invariant strategy profile, $\lambda^{R^*} = \{\lambda_i^{R^*}\}_{i \in \mathcal{N}} \in S$, such that $\lambda_i^{R^*} = b_i(\lambda_{-i}^{R^*})$, for each $i \in \mathcal{N}$.

C. Social Welfare

Social welfare corresponds to an optimal (centralized) allocation by the players with respect to a social welfare function. We choose a typical social welfare function $\Psi(\lambda^R) : S \mapsto \mathbb{R}$ defined by the sum of expected utility of all players, i.e.,

$$\begin{aligned} \Psi = \sum_{i=1}^N \tilde{u}_i &= \sum_{i=1}^N [\mu_i^S r^S + \lambda_i^R f_i(x)] \\ &= (\lambda_T^R + x) r^S + \lambda_T^R r^R(x)(1 - p(x)). \end{aligned} \tag{9}$$

A *social welfare solution* is an optimal allocation that maximizes the social welfare function. It can be shown that the social welfare solution for $\sum_{i=1}^N a_i \lambda_i^R = c$, for any given $c \in \mathbb{R}_{\geq 0}$ can be analytically determined. Since $\sum_{i=1}^N a_i \lambda_i^R \in [0,\ \mu_T^S + \mu_T^R]$, we can employ a bisection algorithm to compute the optimal c and hence, the optimal social welfare solution.

*Lemma 1 (**Social welfare solution**):* For the CPR game Γ with the constraint $\sum_{i=1}^N a_i \lambda_i^R = c$ and the players ordered in the increasing order of $\frac{\mu_i^S}{\mu_i^R}$, the associated social welfare solution, $\lambda^R \in S$ is given by

$$\lambda^R = \left[\mu_1^R,\ \mu_2^R, \cdots,\ \mu_{k-1}^R,\ \frac{1}{a_k}(c - \sum_{i=1}^{k-1} a_i \mu_i^R),\ 0, \cdots,\ 0 \right].$$

For brevity of space, we skip the proof of Lemma 1. The details of the proof can be found in [19].

III. Existence and Uniqueness of PNE and Convergence to the PNE

In this section, we study the existence and uniqueness of the PNE for the CPR game Γ and show that the best response dynamics converge to the unique PNE. Each player $i \in \mathcal{N}$ chooses a review admission rate from her strategy set $S_i = [0,\ \mu_i^R]$ and receives an expected utility $\tilde{u}_i$ given by (7). For any given $\lambda_{-i}^R \in S_{-i}$, we obtain an upper-bound $\overline{\lambda}_i^R : S_{-i} \mapsto S_i$ on λ_i^R defined by

$$\overline{\lambda}_i^R = \begin{cases} 0, & \text{if } \frac{\mu_T^S - \sum_{j \in \mathcal{N}, j \neq i} a_j \lambda_j^R}{a_i} < 0, \\ \frac{\mu_T^S - \sum_{j \in \mathcal{N}, j \neq i} a_j \lambda_j^R}{a_i}, & \text{if } 0 \leq \frac{\mu_T^S - \sum_{j \in \mathcal{N}, j \neq i} a_j \lambda_j^R}{a_i} \leq \mu_i^R, \\ \mu_i^R, & \text{if } \frac{\mu_T^S - \sum_{j \in \mathcal{N}, j \neq i} a_j \lambda_j^R}{a_i} > \mu_i^R, \end{cases}$$

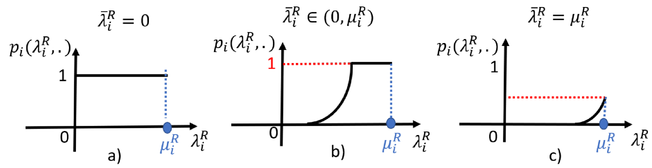

Fig. 2: Failure probability of the player $i \in \mathcal{N}$ as $\overline{\lambda}_i^R$ varies from 0 to μ_i^R. a) For $\overline{\lambda}_i^R = 0$, the failure probability $p_i(\lambda_i^R, \cdot) = 1,\ \forall \lambda_i^R \in S_i$, b) for $\overline{\lambda}_i^R \in (0,\ \mu_i^R)$, $p_i(\lambda_i^R, \cdot)$ is convex for $\lambda_i^R \in [0,\ \overline{\lambda}_i^R)$, with $p_i(\lambda_i^R, \cdot) \mapsto 1$ as $\lambda_i^R \mapsto \overline{\lambda}_i^R$, and $p_i(\lambda_i^R, \cdot) = 1,\ \forall \lambda_i^R \in [\overline{\lambda}_i^R,\ \mu_i^R]$, and c) for $\overline{\lambda}_i^R = \mu_i^R$, $p_i(\lambda_i^R, \cdot)$ is convex in λ_i^R and $p_i(\lambda_i^R, \cdot) < 1,\ \forall \lambda_i^R \in S_i$.

such that for $\lambda_i^R \in [0,\ \overline{\lambda}_i^R) \subset S_i$, the constraint (3) is automatically satisfied, and for $\lambda_i^R \in (\overline{\lambda}_i^R,\ \mu_i^R] \subset S_i$, the constraint (3) is violated. For $\lambda_i^R = \overline{\lambda}_i^R$, the constraint (3) is satisfied if $\overline{\lambda}_i^R \in (0,\ \mu_i^R]$, and is violated if $\overline{\lambda}_i^R = 0$.

We study the properties of the game Γ under the following assumptions:

(A1) For a given $\lambda_{-i}^R \in S_{-i},\ i \in \mathcal{N}$, we assume that the rate of return $r^R(\lambda_i^R, \cdot)$ for reviewing the tasks is continuously differentiable, strictly increasing and strictly concave for $\lambda_i^R \in S_i$, with $r^R(0,0) = 0$. Equivalently, $r^R(x)$ is continuously differentiable, strictly decreasing and strictly concave for $x \in [0,\ \mu_T^S]$, with $r^R(\mu_T^S) = 0$.

(A2) For a given $\lambda_{-i}^R \in S_{-i},\ i \in \mathcal{N}$, we assume that the failure probability $p(\lambda_i^R; \cdot)$ is (i) continuous on S_i; (ii) is continuously differentiable, non-decreasing and convex for $\lambda_i^R \in (0,\ \overline{\lambda}_i^R) \subset S_i$; and (iii) is equal to 1, for $\lambda_i^R \in (\overline{\lambda}_i^R,\ \mu_i^R]$. See Fig. 2 for an illustration. Equivalently, $p(x)$ is continuously differentiable, non-increasing and convex for $x \in (0,\ \mu_T^S]$, and $p(x) \to 1$, as $x \to 0$. Furthermore, $p = 1$, for all $x < 0$.

(A3) We assume $f_i(\mu_i^R, 0) = r^R(\mu_i^R, 0)(1 - p(\mu_i^R, 0)) - \frac{\mu_i^S r^S}{\mu_i^R} > 0$, for each $i \in \mathcal{N}$, i.e., if no other player reviews any task, then each player i has a positive incentive to review the tasks with their maximum admission rate μ_i^R.

*Theorem 1 (**Existence of PNE**):* The CPR game Γ, under the Assumptions (A1-A3), admits a PNE.

We prove Theorem 1 using Brouwer's fixed point theorem [20, Appendix C] applied to the best response mapping with the help of the following lemmas (Lemma 2-4). Recall that $b_i(\lambda_{-i}^R)$ is the best response of the player i to the review admission rates of the other players λ_{-i}^R. For brevity of notation, we will represent $r^R(\lambda_i^R, \lambda_{-i}^R)$, $p(\lambda_i^R, \lambda_{-i}^R)$, $f_i(\lambda_i^R, \lambda_{-i}^R)$, $\tilde{u}_i(\lambda_i^R, \lambda_{-i}^R)$ using r^R, p, f_i, $\tilde{u}_i$, respectively. Furthermore, let q' and q'', respectively, represent the first and the second partial derivatives of a generic function q with respect to λ_i^R.

*Lemma 2 (**Strict concavity of Incentive**):* For the CPR game Γ, under the Assumptions (A1-A2), the incentive function $f_i : S \mapsto \mathbb{R}$ is strictly concave in λ_i^R, for $\lambda_i^R \in [0,\ \overline{\lambda}_i^R]$ and any fixed λ_{-i}^R. Equivalently, $f_i(x)$ is strictly concave in x for $x \in [0,\ \mu_T^S - \sum_{j \in \mathcal{N}, j \neq i} a_j \lambda_j^R]$.

Proof: The first and the second partial derivative of the incentive function f_i with respect to λ_i^R in the interval $\lambda_i^R \in [0,\ \overline{\lambda}_i^R]$ is given by:

$$f_i' = r'^R(1 - p) - r^R p' = -a_i \frac{df_i}{dx}, \tag{10a}$$

$$f_i'' = r''^R(1 - p) - 2r'^R p' - r^R p'' = a_i^2 \frac{d^2 f_i}{dx^2}. \tag{10b}$$

From the Assumptions (A1) and (A2), we have $f_i'' < 0$ and $\frac{d^2 f_i}{dx^2} < 0$ in the interval where the derivative of f_i exists, thereby proving the strict concavity of f_i in λ_i^R and x. ■

*Lemma 3 (**Best response mapping**):* For the CPR game Γ, under the Assumptions (A1-A2), the best response mapping $b_i(\lambda_{-i}^R)$ is unique for $\lambda_{-i}^R \in S_{-i}$ and is given by:

$$b_i(\lambda_{-i}^R) = \begin{cases} 0, & \text{if } f_i(\lambda_i^R, \cdot) \le 0,\ \ \forall \lambda_i^R \in S_i, \\ \alpha_i, & \text{if } \exists \alpha_i \in S_i \text{ s.t. } \frac{\partial \tilde{u}_i}{\partial \lambda_i^R}(\alpha_i) = 0, \text{ and } f_i(\alpha_i, \cdot) > 0, \\ \mu_i^R, & \text{otherwise}. \end{cases}$$

Proof [Sketch]: The proof utilizes the strict concavity of f_i (Lemma 2) and considers the structure of $\tilde{u}_i$ with respect to f_i. Specifically, it can be shown that the best response b_i is unique and either satisfies $\frac{\partial \tilde{u}_i}{\partial \lambda_i^R} = 0$, or occurs at the boundary of S_i. The details of the proof can be found in [19]. ■

*Lemma 4 (**Continuity of best response mapping**):* For the CPR game Γ, under the Assumptions (A1-A3), the best response mapping $b_i(\lambda_{-i}^R)$ is continuous at every $\lambda_{-i}^R \in S_{-i}$.

Proof [Sketch]: The proof utilizes the continuity of maximum admissible review admission rate (say $h(\lambda_{-i}^R)$) above which the system constraint (3) is violated with respect to λ_{-i}^R. We establish continuity of the best response mapping b_i with respect to λ_{-i}^R by establishing its continuity with respect to the continuous mapping $h(\lambda_{-i}^R)$. The details of the proof can be found in [19]. ■

Proof of Theorem 1: To prove the existence of a PNE, define a mapping $M : S \mapsto S$ as follows:

$$M(\lambda_1^R,\ \lambda_2^R, \ldots,\ \lambda_N^R) = (b_1(\lambda_{-1}^R),\ b_2(\lambda_{-2}^R), \ldots,\ b_N(\lambda_{-N}^R)).$$

The mapping M is unique (Lemma 3) and continuous (Lemma 4), and maps the compact convex set S (S_i is convex and compact, $\forall i \in \mathcal{N}$) to itself. Hence, application of the Brouwer's fixed point theorem [20, Appendix C] yields that there exists a strategy profile $\lambda^R = \{\lambda_i^{R^*}\}_{i \in \mathcal{N}} \in S$ which is invariant under the best response mapping and therefore is a PNE of the game. □

*Proposition 1 (**Structure of PNE**):* For the CPR game Γ with the players ordered in the increasing order of $\frac{\mu_i^S}{\mu_i^R}$, let $\lambda^{R^*} = [\lambda_1^{R^*}, \lambda_2^{R^*}, \cdots, \lambda_N^{R^*}]$ be a PNE, then the following statements hold:

(i) If for any player k_1, $\lambda_{k_1}^{R^*} < \mu_{k_1}^R$, then $\lambda_{k_1}^{R^*} \ge \lambda_{k_2}^{R^*}$ for each $k_2 > k_1$; and

(ii) if $\lambda_l^{R^*} = 0$, for any $l \in \mathcal{N}$, then $\lambda_i^R = 0$, for each $i \in \{j \in \mathcal{N} \mid j \ge l\}$.

Proof: The proof leverages two intermediate results: (a) $\lambda_i^{R^*} = 0$, if and only if, $f_i(\lambda_i^{R^*}, \lambda_{-i}^{R^*}) \le 0$; and (b) the review admission rate for the player i at a PNE is non-zero and satisfies the implicit equation

$$\lambda_i^{R^*} = \min\Big\{-\frac{f_i(\lambda_i^{R^*},\lambda_{-i}^{R^*})}{f_i'(\lambda_i^{R^*},\lambda_{-i}^{R^*})},\ \mu_i^R\Big\},$$

if and only if, $f_i(\lambda_i^{R^*},\lambda_{-i}^{R^*}) > 0$. The proof of these results can be found in [19, Corollary 2].

We now establish the Proposition. Let $\lambda_{k_1}^{R^*}$ and $\lambda_{k_2}^{R^*}$ be the review admission rates at a PNE for the players k_1 and k_2, respectively, with $\frac{\mu_{k_1}^S}{\mu_{k_1}^R} \le \frac{\mu_{k_2}^S}{\mu_{k_2}^R}$. We assume $\lambda_{k_1}^{R^*} < \lambda_{k_2}^{R^*}$ and prove the first statement using a contradiction argument.

Case 1: $\lambda_{k_1}^{R^*} = 0$. From the statement (a) above, $f_{k_1}(\lambda_{k_1}^{R^*},\lambda_{-k_1}^{R^*}) \le 0$. From (8), the incentives f_{k_1} and f_{k_2} for the players k_1 and k_2 at a PNE satisfies:

$$f_{k_2} = f_{k_1} + \Big(\frac{\mu_{k_1}^S}{\mu_{k_1}^R} - \frac{\mu_{k_2}^S}{\mu_{k_2}^R}\Big)r^S = f_{k_1} + (a_{k_1} - a_{k_2})r^S < 0.$$

Thus, from the statement (a) above, $\lambda_{k_2}^{R^*} = 0$, which is a contradiction. This also proves the Proposition statement (ii).

Case 2: $\lambda_{k_1}^{R^*} > 0$. By assumption $\lambda_{k_1}^{R^*} < \lambda_{k_2}^{R^*}$, from the statement (b) above, $\lambda_{k_1}^{R^*}$ and $\lambda_{k_2}^{R^*}$ satisfies the implicit equations, $\lambda_{k_1}^{R^*} = \min\big\{-\frac{f_{k_1}(\lambda_{k_1}^{R^*},\ \lambda_{-k_1}^{R^*})}{f_{k_1}'(\lambda_{k_1}^{R^*},\ \lambda_{-k_1}^{R^*})},\ \mu_{k_1}^R\big\}$ and $\lambda_{k_2}^{R^*} = \min\big\{-\frac{f_{k_2}(\lambda_{k_2}^{R^*},\ \lambda_{-k_2}^{R^*})}{f_{k_2}'(\lambda_{k_2}^{R^*},\ \lambda_{-k_2}^{R^*})},\ \mu_{k_2}^R\big\}$, respectively. By the hypothesis $\lambda_{k_1}^{R^*} < \mu_{k_1}^R$, and therefore, $\lambda_{k_1}^{R^*} = -\frac{f_{k_1}}{f_{k_1}'}$. From (8) and (10a):

$$\begin{aligned}\lambda_{k_2}^{R^*} &= \min\Big\{-\frac{f_{k_2}}{f_{k_2}'},\ \mu_{k_2}^R\Big\}\\ &\le -\frac{f_{k_2}}{f_{k_2}'} = -\frac{f_{k_1} + (a_{k_1} - a_{k_2})r^S}{\frac{a_{k_2}}{a_{k_1}}f_{k_1}'} \le \lambda_{k_1}^{R^*},\end{aligned}$$

which is a contradiction. Hence, if $\lambda_{k_1}^{R^*} < \mu_{k_1}^R$, then $\lambda_{k_1}^{R^*} \ge \lambda_{k_2}^{R^*}$ for each $k_2 > k_1$. ∎

It follows from Proposition 1 that the review admission rate of each player i at a PNE is monotonically decreasing with the ratio $\frac{\mu_i^S}{\mu_i^R}$. Therefore, at a PNE, as the heterogeneity among the players become very large, players with small (respectively, large) ratio of $\frac{\mu_i^S}{\mu_i^R}$ review the tasks with high (respectively, zero) review admission rate. Thus, a PNE has characteristics similar to the social welfare solution obtained in Lemma 1. We illustrate this further in Section IV.

*Theorem 2 (**Uniqueness of PNE**):* The PNE admitted by the CPR game Γ, under the Assumptions (A1-A3), is unique.

Proof [Sketch]: Uniqueness of the PNE is established by considering multiple PNEs, and showing by contradiction arguments that the number of players with non-zero review admission rate and the slackness parameter x are the same at each PNE. This leads to uniqueness of the PNE. The details of the proof can be found in [19]. ∎

*Theorem 3 (**Convergence to the PNE**):* For the CPR game Γ, under the Assumptions (A1-A3), the best response dynamics converges to the unique PNE.

Proof [Sketch]: It can be verified that the proposed CPR game Γ under the Assumptions (A1-A3) belongs to a class of *Quasi Aggregative games* defined in [21] in which the expected utility function $\tilde{u}_i$ for each player i is a function of the player's own strategy λ_i^R, and an interaction function $\sigma_i(\lambda_{-i}^R) : S_{-i} \mapsto \mathbb{R}$ independent of λ_i^R. Applying [21, Theorem 1], we conclude that if the best response for all the players is decreasing in $\sigma_i(\lambda_{-i}^R) = \sum_{j=1,j\neq i}^N a_j\lambda_j^R$, then the CPR game Γ is a best response pseudo-potential game [22]. By realizing $x = \mu_T^S - a_i\lambda_i^R - \sigma_i(\lambda_{-i}^R)$ and differentiating b_i with respect to $\sigma_i(\lambda_{-i}^R)$, we formally show that the best response mapping $b_i(\lambda_{-i}^R)$ is monotonically decreasing in $\sigma_i(\lambda_{-i}^R)$, for each $i \in \mathcal{N}$ (see [19] for detailed proof).

Remark 1 in [16] states that a best response pseudo-potential game with a unique best response is an instance of best response potential game [15]. Therefore, the CPR game Γ, with its unique (Lemma 3) and monotonically decreasing best response b_i in $\sigma_i(\lambda_{-i}^R)$, is a best response potential game. Hence, simple best response dynamics such as sequential best response dynamics [16] and simultaneous best response dynamics [17] converge to its unique PNE. ∎

IV. Numerical Illustrations

In this section, we numerically illustrate the inefficiency of the PNE by comparing its structure with the social welfare solution as well as by studying the variation of Price of Anarchy (PoA) [23] with increasing heterogeneity among the players. PoA is defined as the ratio of the social welfare function Ψ, evaluated at the social welfare solution and the PNE, respectively. Therefore, $PoA = \frac{(\Psi)_{SW}}{(\Psi)_{PNE}} \ge 1$.

In our numerical illustrations, we obtain the PNE by simulating the sequential best response dynamics of players with randomized initialization of their strategy. We verify the uniqueness of the PNE for different choices of functions, $r^R(x)$ and $p(x)$ satisfying Assumptions (A1-A2), and by following sequential best response dynamics with multiple random initializations for the strategy of each player. Furthermore, in our numerical simulations, we relax the Assumption (A3) and still obtain the unique PNE.

A comparison of the social welfare solution (obtained using fmincon in MATLAB) and the PNE, for low and high heterogeneity among players, is shown in Fig. 3. For our numerical illustrations, we choose the number of players, $N = 6$, and choose the functions $r^R(x)$ and $p(x)$, satisfying the Assumptions (A1-A2) as following:

$$r^R(\lambda_i^R,\lambda_{-i}^R) = r^R(x) = 5[1 - \exp\{0.5(x - \mu_T^S)\}],$$

$$p_i^R(\lambda_i^R,\lambda_{-i}^R) = p(x) = \begin{cases}1, & \text{if} \quad x \le 0,\\ \exp(-0.5x), & \text{otherwise},\end{cases}$$

where $x = \mu_T^S - \sum_{i=1}^N a_i\lambda_i^R$ is the slackness parameter. To characterize the heterogeneity among the players, we do a random sampling of the players' maximum service admission rate and maximum review admission rate from normal distributions with fixed means, $M_{\mu_S} = 40$, and $M_{\mu_R} = 80$, and identical standard deviation, $\rho \in \mathbb{R}_{>0}$. Any non-positive realizations were discarded. We consider the standard deviation of the distributions as the measure of heterogeneity among the players.

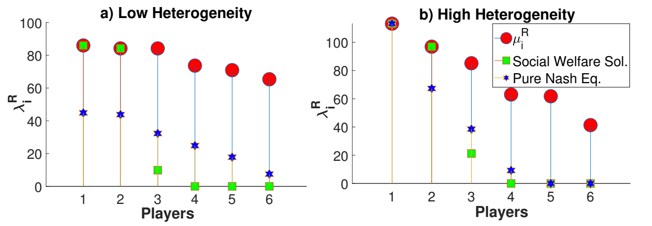

Fig. 3: Social welfare solution and the pure Nash equilibrium for a) low and b) high heterogeneity among the players, respectively. μ_i^R denotes the maximum review admission rate for each player i.

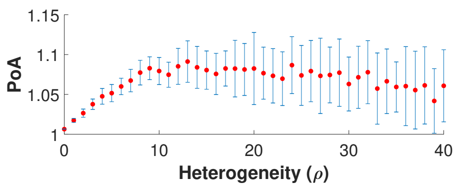

Fig. 4: Price of Anarchy with the increasing heterogeneity among the agents

Fig. 3 shows that in the social welfare solution, the players with low ratio of $\frac{\mu_i^S}{\mu_i^R}$ review the tasks at maximum review admission rate and the players with high ratio of $\frac{\mu_i^S}{\mu_i^R}$ drop out of the game. At PNE, the strategy profile of the players follow the characteristics described by Lemma 1. Lastly, with the increase in heterogeneity among the players, the PNE starts to approach the social welfare solution.

Fig. 4 shows the variation of the PoA with increasing heterogeneity among the players. In the case of homogeneous players, i.e., $\rho = 0$, any strategy profile with same λ_T^R, produces same value of the social welfare function, Ψ (see (9)), and hence results in PoA = 1. As we initially increase the heterogeneity among the players, the PNE starts to deviate from the social welfare solution, resulting in an increase in PoA. Finally, with a large increase in the heterogeneity among the players, the PNE starts to approach the social welfare solution, i.e. the players with small ratio of $\frac{\mu_i^S}{\mu_i^R}$ starts reviewing the tasks with high review admission rate, and the players with very large ratio of $\frac{\mu_i^S}{\mu_i^R}$ starts to drop out of the game (see Lemma 1). We note that the PoA ≤ 1.15, suggesting that the unique PNE is close to the optimal centralized social welfare solution.

V. Conclusions and Future Directions

We studied incentive design mechanisms to facilitate effective team collaboration among the agents servicing a stream of homogeneous tasks. In particular, we designed a Common-Pool Resource (CPR) game to incentivize team collaboration and showed the existence and uniqueness of PNE. We showed that the proposed CPR game is an instance of the best response potential game and by playing the sequential best response against each other, players converge to the unique PNE. At PNE, the review admission rate of the players decreases with the increasing ratio of $\frac{\mu_i^S}{\mu_i^R}$, i.e., the review admission rate is higher for the players that are "better" at reviewing the tasks than servicing the tasks (characterized by their average service and review time).

There are several possible avenues of future research. It is of interest to extend the results for broader class of games with less restrictive choice of utility functions, i.e. games that are not quasi-aggregative or commonly used games of weak strategic substitutes (WSTS) or complements (WSTC). An interesting open problem is to consider a team of agents processing stream of heterogeneous tasks. In such a setting, incentivizing team collaboration based on the task-dependent skill-set of the agents is also of interest.

*Bibliography

[1] M. Haas and M. Mortensen, "The secrets of great teamwork." *Harvard Business Review*, vol. 94, no. 6, pp. 70–6, 2016.

[2] C. Keser and R. Gardner, "Strategic behavior of experienced subjects in a common pool resource game," *International Journal of Game Theory*, vol. 28, no. 2, pp. 241–252, 1999.

[3] A. R. Hota, S. Garg, and S. Sundaram, "Fragility of the commons under prospect-theoretic risk attitudes," *Games and Economic Behavior*, vol. 98, pp. 135–164, 2016.

[4] J. Peters, V. Srivastava, G. Taylor, A. Surana, M. P. Eckstein, and F. Bullo, "Human supervisory control of robotic teams: Integrating cognitive modeling with engineering design," *IEEE Control System Magazine*, vol. 35, no. 6, pp. 57–80, 2015.

[5] K. Savla and E. Frazzoli, "A dynamical queue approach to intelligent task management for human operators," *Proceedings of the IEEE*, vol. 100, no. 3, pp. 672–686, 2012.

[6] V. Srivastava, R. Carli, C. Langbort, and F. Bullo, "Attention allocation for decision making queues," *Automatica*, vol. 50, no. 2, pp. 378–388, 2014.

[7] P. Gupta and V. Srivastava, "Optimal fidelity selection for human-in-the-loop queues using semi-Markov decision processes," in *American Control Conference*, Philadelphia, PA, Jul. 2019, pp. 5266–5271.

[8] F. Gao, M. L. Cummings, and E. T. Solovey, "Modeling teamwork in supervisory control of multiple robots," *IEEE Transactions on Human-Machine Systems*, vol. 44, no. 4, pp. 441–453, 2014.

[9] B. Mekdeci and M. Cummings, "Modeling multiple human operators in the supervisory control of heterogeneous unmanned vehicles," in *Proceedings of the 9th Workshop on Performance Metrics for Intelligent Systems*. ACM, 2009, pp. 1–8.

[10] Y. Guo, "Using agent-based models to understand multi-operator supervisory control," *International Journal of Precision Engineering and Manufacturing*, 2012.

[11] A. Hong, "Human-robot interactions for single robots and multi-robot teams," Ph.D. dissertation, University of Toronto, 2016.

[12] V. Srivastava and F. Bullo, "Knapsack problems with sigmoid utilities: Approximation algorithms via hybrid optimization," *European Journal of Operational Research*, vol. 236, no. 2, pp. 488–498, 2014.

[13] A. R. Hota, S. Garg, and S. Sundaram, "Resource sharing games with failures and heterogeneous risk attitudes," in *Annual Allerton Conference on Communication, Control, and Computing*, 2013, pp. 535–542.

[14] A. R. Hota and S. Sundaram, "Controlling human utilization of failure-prone systems via taxes," *arXiv preprint arXiv:1802.09490*, 2018.

[15] M. Voorneveld, "Best-response potential games," *Economics letters*, vol. 66, no. 3, pp. 289–295, 2000.

[16] P. Dubey, O. Haimanko, and A. Zapechelnyuk, "Strategic complements and substitutes, and potential games," *Games and Economic Behavior*, vol. 54, no. 1, pp. 77–94, 2006.

[17] M. K. Jensen, "Stability of pure strategy nash equilibrium in best-reply potential games," *University of Birmingham, Tech. Rep*, 2009.

[18] C. G. Cassandras and S. Lafortune, *Introduction to Discrete Event Systems*. Springer Science & Business Media, 2009.

[19] P. Gupta, S. D. Bopardikar, and V. Srivastava, "Incentivizing collaboration in heterogeneous teams via common-pool resource games," *arXiv preprint arXiv: 1908.03938*, Aug. 2019.

[20] T. Basar and G. J. Olsder, *Dynamic Noncooperative Game Theory*. SIAM, 1999, vol. 23.

[21] M. K. Jensen, "Aggregative games and best-reply potentials," *Economic Theory*, vol. 43, no. 1, pp. 45–66, 2010.

[22] B. Schipper, "Pseudo-potential games," University of Bonn, Germany, Tech. Rep., 2004, working paper. [Online]. Available: available at citeseer.ist.psu.edu/schipper04pseudopotential.html

[23] J. P. Hespanha, *Noncooperative Game Theory: An Introduction for Engineers and Computer Scientists*. Princeton University Press, 2017.