Artificial Intelligence 323 (2023) 103975

Contents lists available at ScienceDirect

Artificial Intelligence

journal homepage: www.elsevier.com/locate/artint

Budget-feasible mechanisms for proportionally selecting agents from groups ☆

Xiang Liu [a], Hau Chan [b], Minming Li [c], Weiwei Wu [a,*], Yingchao Zhao [d]

[a] Southeast University, Nanjing, China
[b] University of Nebraska-Lincoln, Lincoln, NE, USA
[c] City University of Hong Kong, Hong Kong Special Administrative Region of China
[d] Caritas Institute of Higher Education, Hong Kong Special Administrative Region of China

ARTICLE INFO

Article history:
Received 7 January 2022
Received in revised form 1 February 2023
Accepted 12 July 2023
Available online 1 August 2023

Keywords:
Proportional representation
Budget-feasibility
Mechanism design
Auction

ABSTRACT

In many social domains involving collective decision-making (*e.g.*, committee selection and survey sampling), it is often desirable to select individuals from different population groups to achieve proportional representation (*e.g.*, to represent the opinions of each group). For instance, in the selection of a committee (*e.g.*, to form a working group within a company), the planner would like to select agents from different groups to represent their respective groups proportionally. Typically, there are intrinsic private costs for agents to represent their groups, and the planner would like to compensate the selected agents via some form of payments, which is constrained by the planner's available budget. As the costs are unknown to the planner, the planner is required to design incentive mechanisms to elicit agents' real costs and provide payments (or monetary incentives) to the selected agents to ensure proportional representation and the total payments do not exceed the budget. Such a mechanism design setting falls into the budget-feasible mechanism design paradigm. However, existing budget-feasible mechanisms only consider all agents to be in the same group with non-proportional objectives. To study the above-mentioned setting, we consider the problem of designing budget-feasible mechanisms for selecting agents with private costs from various groups to ensure proportional representation, where the minimum proportion of the overall value of the selected agents from each group is maximized. We study this problem by first considering the setting with homogeneous agents who have identical values to the planner. Depending on agents' membership in the groups, we consider two models: a single group model where each agent belongs to only one group, and a multiple group model where each agent may belong to multiple groups. We propose novel budget-feasible proportion-representative mechanisms for these models that require different selection methods, *i.e.*, a novel greedy mechanism that considers all possible proportion ratios for the single group model and a novel mechanism that leverages the Max-Flow algorithm to evaluate the proportional representation for the multiple group model, to choose representative agents from each group. The proposed mechanisms guarantee theoretical properties of individual rationality, budget-feasibility, truthfulness, and approximation performance on maximizing the minimum proportional representation of each group. We also provide a matching lower bound for budget-feasible

☆ The previous version of this paper appeared at IJCAI 2021 [1] with title "Budget-feasible Mechanisms for Representing Groups of Agents Proportionally". The earlier paper focused on budget-feasible proportion-representative mechanisms on homogeneous agents who hold identical values. We non-trivially extend the previous mechanisms to the heterogeneous agent setting where agents have different values. We elaborate on our new contributions in Section 5.

* Corresponding author.
E-mail addresses: xiangliu@seu.edu.cn (X. Liu), hchan3@unl.edu (H. Chan), minming.li@cityu.edu.hk (M. Li), weiweiwu@seu.edu.cn (W. Wu), yczhao@cihe.edu.hk (Y. Zhao).

https://doi.org/10.1016/j.artint.2023.103975

proportion-representative mechanisms. Finally, we non-trivially extend these mechanisms to the settings of heterogeneous agents who can have different values to the planner under the two models.

1. Introduction

In many real-world scenarios such as conducting surveys [2], selecting representatives [3], public opinion prediction [4], voting [5] and school choice [6], it is important to select a proportional number of agents from each population group to best represent the overall population. For instance, when conducting surveys in crowdsourcing platforms [7–9], the organizer typically is required to gather information from individuals among different populations (e.g., geographic locations) via websites [10] or mobile smartphones [11]. For the survey data to be representative or useful for various applications (*e.g.*, for building generalizable prediction of public opinion [12] or forecasting models to reflect the underlying population distributions [13]), we need to select proportional numbers of individuals from different populations. Inadequate proportional representation can affect generalisability, *e.g.*, inaccurate poll predictions due to lack of representative samples [9]. While achieving proportional representation can be a straightforward process in crowdsourcing surveys, the challenge is to identify appropriate payments to the agents such that we can incentivize agents to participate in the surveys [14,15] and the total payment is no more than the allowable budget.[1] Indeed, different agents might require different desirable rewards to participate, and the reward information might not be known to the organizer publicly.

Another example is committee selection, where the planner needs to select members from a set of candidates described by various attributes (*e.g.*, gender, age, profession, and division) such that each attribute offers a certain representation [18]. Typically, the chosen member is required to undertake additional duties, *e.g.*, managing the company and organizing campaigns, which result in the consumption of time and attention. In exchange for their services, it is natural to reward chosen members via monetary payments [19] (*e.g.*, salaries, bonus, or funding for traveling), which should be constrained by a predefined budget (*e.g.*, funding for constructing a committee [20]). However, the required rewards for members can be different and privately known by themselves.

Because the agent desired rewards are often unknown to the social planner and the social planner has a budget in many domains (including the above-mentioned settings), we must consider selecting appropriate agents and determining their payment schemes in order to achieve proportional representation (*i.e.*, choosing proportional agents from each group). Without paying agents correctly, they can be unwilling to participate in the activity (*e.g.*, completing the surveys or serving on the committee), which can lead to low participation rates (*e.g.*, see studies regarding low survey response rates [21,22]). Therefore, in this paper, we aim to address the following question.

Can we design a mechanism that elicits private reward information from agents and selects agents to participate in the campaign in order to achieve proportional representation such that the total payments to the selected agents are no more than the budget?

The problem at hand can be cast naturally into a budget-feasible mechanism design setting [23][2] where the social planner seeks to design a computationally efficient mechanism that elicits true cost information from agents, selects agents to optimize some objective, and derives payments to the selected agents to ensure their total payment is no more than the budget. However, the design of budget-feasible mechanisms has not been considered in the considered proportional representation setting with different groups of agents.

1.1. Our contributions

We consider the problem of designing budget-feasible mechanisms for selecting agents proportionally from groups satisfying standard desirable properties (*i.e.*, individual rationality, budget-feasibility, and truthfulness [23]). In particular, there are n agents $S = \{s_1, s_2, ..., s_n\}$ and m groups $G = \{G_1, G_2, ..., G_m\}$, where group G_j is a non-empty subset of S. We use $\mathcal{G}(s_i)$ to denote the set of groups that agent s_i belongs to. Agent s_i has a value v_i to the planner. The planner has a budget $B \in \mathbb{R}_+$ and each agent s_i has a private cost $c_i \in \mathbb{R}_+$.

1.1.1. Homogeneous agents

We first consider homogeneous agents who have identical values to the planner, *i.e.*, $v_i = 1, \forall 1 \le i \le n$. As selecting a proportional number of agents from each population group to best represent the overall population has been well studied in many real-world scenarios such as conducting surveys [2], selecting representatives [3], public opinion prediction [4],

[1] Typically, the organizer has a limited budget for conducting a crowdsourcing task [16,17]. Trivially, if there is no budget, the organizer can pay agents a very high amount which might be unnatural for many domains.

[2] We refer to the mechanism that guarantees budget-feasibility as a budget-feasible mechanism.

voting [5] and school choice [6]. Thus, we are inspired to consider the proportionality among groups that we want to choose proportional agents from each group under the budget constraint. To ensure proportional representation, we carefully formulate the planner's objective as to select agents from different groups that maximizes the minimum proportional ratio of the selected agents among groups, *i.e.*, $\max\min_{1\le j\le m}\frac{Q_j}{\sum_{s_i\in G_j}v_i}$ where Q_j is the total value of chosen agents in group G_j. This objective is based on a well-studied notion of proportional representation (*e.g.*, the electoral system requires the minimum vote proportion which is the ratio between the number of voters and population in each group for representation [24,25], specifying an exact allocation of votes to different groups proportionally [26] and maximizing the minimum diversity fairness among different groups [27,28]).

We differentiate two general models that depend on whether each agent belongs to (1) one group (*i.e.*, $|\mathcal{G}(s_i)|=1$) or (2) multiple groups (*i.e.*, $|\mathcal{G}(s_i)|\ge 1$). Within the multiple group model, we use $x_{ij}\in\{0,1\}, s_i\in S_w$ to denote whether group G_j is counted once when s_i is chosen. Then, we further consider (2a) the single counting case in which a selected agent is counted exactly once in one of the groups she belongs to, *i.e.*, $\sum_{j\in\mathcal{G}(s_i)}x_{ij}=1$ if $s_i\in S_w$, and (2b) multiple counting case in which a selected agent is counted in all groups she belongs to, *i.e.*, $\forall j\in\mathcal{G}(s_i), x_{ij}=1$ if $s_i\in S_w$. For example, in survey collection, consider surveying individuals with different occupations (*e.g.*, doctors and teachers) or work experiences (*e.g.*, 1-2 years, ..., 10 years or more), where each category can be viewed as a group. The single group model can be appropriate for survey gathering when the data analysis focuses on a single attribute such as occupations or work experiences only (*e.g.*, to measure certain characteristics) or political polling (*e.g.*, to measure population opinions). Similarly, the multiple group model with multiple counting can be appropriate for survey gathering when the data analysis is conditional on a certain attribute (*e.g.*, conditional on an occupation to gather statistics with respect to work experiences). The multiple group model with single counting is particularly useful for data gathering dealing with A/B testing [29] or human subject research [30] with randomized trials where each subject is matched to a trial that is appropriate for their attribute (*e.g.*, either a particular occupation or a particular work experience to observe their proficiency or skill levels for given tasks). Because different trials often have overlapping components and subjects can learn from previous trials, subjects often participated in only one randomized trial to limit spillover effects.

For these models, under the goal of maximizing the minimum proportional ratio of selected agents among groups, we design budget-feasible proportion-representative mechanisms. The proposed mechanisms achieve desirable theoretical properties, including budget-feasibility, individual rationality, truthfulness, and approximation guarantee. We note that existing budget-feasible mechanism approaches do not apply directly (see Section 2 for more discussion). As a result, we design several novel mechanisms for our models. Table 1 shows the performance of our mechanisms (when the mechanisms can select at least one agent from each group). Note that, as we will show, other possible mechanisms can also fail to obtain good approximation ratios when our mechanisms cannot select more than one agent. See Section 4 and 5 for more detail.

In particular, for (1), we construct a novel greedy mechanism that selects agents proportionally from each group and pays chosen agents appropriately for budget-feasibility. The proposed mechanism achieves approximation performance that depends on the size of the largest and smallest groups. Moreover, we show a matching lower bound for any budget-feasible proportion-representative mechanism.

For the multiple group model (2a) or (2b), we construct a novel mechanism that leverages the Max-Flow algorithm [31] to evaluate the proportional representation under a given subset of agents, in which we can find the candidate agent set with the greatest proportional representation within the budget constraint. We then apply the minimum weight matching to identify the final selected agents from the candidate agent set, whereby the estimated maximum proportional representation can be obtained, and the corresponding payment can be determined. The designed mechanisms in this model achieve approximation performance that depends on the size of the groups.

1.1.2. Heterogeneous agents

In the above settings, each agent represents a single unit in their corresponding population groups (*e.g.*, contributing to a single survey or randomized trial). In some scenarios, an agent can actually provide more contributions beyond a single unit. For instance, consider a crowdsensing setting for monitoring the air qualities of different regions [32]. Each region consists of a set of agents corresponding to the region size. Each agent can measure the air qualities of a different number of (randomly sampled non-overlapping) locations within the regions predetermined by the planner [33,34]. Therefore, an agent's contribution is heterogeneous, and the social planner needs to select proportional numbers of agents in order to obtain good estimations of air qualities across different regions. The single group model and multiple group model with single counting can naturally capture the crowdsensing setting in which agents can only be assigned to measure the air qualities in a single region. The multiple group model with multiple counting can be used to capture situations where an agent measures the air qualities of locations in the boundary of different regions. As a result, an agent can belong to two different regions and be accounted for in both regions simultaneously.

As another example, when conducting surveys (*e.g.*, regarding habits or certain behaviors), each agent can gather or provide multiple units of responses from their friends, families, coworkers, or neighbors within their regions or groups [35,36]. Therefore, to obtain a representative of the different groups or regions of populations, the planner should select appropriate agents whose total number of provided surveys is proportional to the sum of surveys that can be collected by all agents in this group. The considered single group and multiple group models can be used to capture different situations of conducting surveys for data analysis (as discussed earlier).

Table 1
Our Approximation Results. The parameters m, α, σ, and η denote the number of groups, the ratio of the maximum and the minimum number of agents among groups, the ratio of the maximum and minimum value of agents, and the ratio of the maximum and minimum total value of agents among groups, respectively. UB and LB refer to upper bounds and lower bounds, respectively. This table shows the approximation performance when mechanisms can select at least one agent from each group. We also show that no budget-feasible proportion-representative mechanisms can do much better either when our mechanisms cannot select at least one agent in Section 4 and 5.

			UB	LB
Homogeneous	Single Group		$3+\alpha$ (Th. 4)	$\Omega(\alpha)$ (Th. 5)
	Multiple Group	Single Counting	$m\alpha(\alpha+2)+1$ (Th. 8)	
		Multiple Counting	$m\alpha(\alpha+2)+1$ (Th. 10)	
Heterogeneous	Single Group		$(m+1)(1+\sigma\eta)$ (Th. 13)	$\Omega(\sigma\eta)$ (Th. 14)
	Multiple Group	Single Counting	$m\eta(1+\sigma)(1+\sigma\eta)+1$ (Th. 17)	
		Multiple Counting	$m\eta(1+\sigma)(1+\sigma\eta)+1$	

The above two examples illustrate the importance of considering the heterogeneous values of agents, *i.e.*, agents' values v_i can be different.

The social planner now has the objective of maximizing the minimum proportional ratio of the total value of the selected agents from each group. Similarly, we consider the single group and multiple group models. We non-trivially extend the mechanisms for homogeneous agents to the heterogeneous setting. In particular, the main difference is that for the single group model (1), we design a non-trivial payment scheme to ensure agents' truthfulness. For the multiple group model (2a) and (2b), we carefully choose the fractional budget used to select agents to ensure budget-feasibility when agents hold different values.

1.2. Comparison with traditional budget-feasible mechanisms

Generally, the standard approach for designing budget-feasible mechanisms has leveraged a greedy approach to select agents based on their value-to-cost ratios [23]. However, this approach cannot be used to achieve proportional selections directly as they choose agents by their bids without considering group attribution. As our initial starting point, we investigate whether a similar greedy approach can work for our setting. It turns out that we can design appropriate selection and payment schemes for each group, *e.g.*, use a greedy approach to select agents in each group (along with other techniques such as integer program and max-min flow in different models/settings). Therefore, the proposed (new greedy) mechanisms can achieve the objective of proportional representations while guaranteeing truthfulness and budget-feasibility.

An earlier version of the paper appeared at IJCAI 2021 focusing only on the homogeneous setting with omitted proofs. In this journal version, we included all the proofs and extended our results to the heterogeneous setting.

2. Related work

Budget-feasible mechanisms Since the original seminal work of budget-feasible mechanisms [23], many research efforts have been invested in designing budget-feasible mechanisms for various valuation functions of the buyer. Chen et al. [37] further develop improved mechanisms with better approximation ratios for the submodular value function, while Amanatidis et al. [38] consider symmetric submodular valuations, a prominent class of non-monotone submodular functions. Anari et al. [17] design a constant-approximation budget-feasible mechanism for large markets where sellers' costs are far less than the buyer's budget and show that it is impossible to achieve bound approximation ratio without the large market assumption when sellers' items are divisible. Singer and Mittal [39] focus on designing pricing mechanisms with the objective of maximizing the number of finished tasks while guaranteeing budget-feasibility. However, these mechanisms in existing literature do not perform well for our settings directly as they do not consider agents' groups and ensure proportional representation. In particular, these mechanisms greedily select agents with the lowest cost-per-value ratios irrespective of the group memberships, which may lead to the selected agents belonging to one group only if a similar greedy manner is used in the group setting (*e.g.*, all the members in a single group have very low cost-per-value ratio). Thus, such a mechanism cannot ensure proportional representation.

Fairness/diversity in optimization perspective Note that proportional representation has been studied from the optimization or algorithmic perspective in various areas such as voting and electoral systems. For example, Procaccia et al. [5] focus on analyzing the complexity of achieving proportional representation. Buisseret and Prato [40] consider the voter preferences in proportional representation systems to understand the candidate selection and behavior. Flanigan et al. [3] pay attention to 'citizens' assemblies in which a panel of randomly selected constituents contributes to questions of policy and develop selection algorithms that can select representative people and in the spirit of democratic equality, individuals would ideally be selected to serve on this panel with equal probability. In addition, some works consider diversity fairness in matching/allocation [41–43], school choice [44,45] (*e.g.*, each school is endowed with a lower and an upper quota for each distinct type), housing allocation [46] (*e.g.*, every ethnic group must not own more than a certain percentage in a housing project). These works mainly view diversity as a constraint while optimizing the general objectives (*e.g.*, social welfare) or consider balancing diversity and efficiency simultaneously [42]. However, in our model, we want to maximize the minimum selection ratio among different groups to ensure proportional representation.

Another related research direction is the (allocation) apportionment problem [47–49] in which some public resources (*e.g.*, seats in the parliament) should be proportionally divided into different groups (*e.g.*, parties or states). However, they mainly focus on the complete information scenario. In our model, each agent has a privately known cost and the chosen agent should obtain a payment as the reward (may not be equal to her cost due to the truthfulness requirement) under the budget constraint.

Fairness/diversity in mechanism design There are also some works taking the group attributes and diversity fairness into account when designing (auction) mechanisms. Ilvento et al. [50] consider the problem of fairness in advertising and propose the inter-category and intra-category fairness desiderata. Kuo et al. [51] and Finocchiaro et al. [52] consider the mechanism design problem by using a machine learning method to address fairness. Chawla et al. [53] first express the fairness constraint as a kind of stability condition and then introduce a new class of allocation algorithms to achieve a near-optimal trade-off between fairness and social welfare. Some literature considers the design of fair mechanisms in job processing. They study the max-min problem, *e.g.*, makespan minimization [54,55], which is similar to our proportion-representative objective. Moreover, many works on school choice [56–58] consider the mechanism design problem with diversity constraints, that is, there is a specific minimum (maximum) quota of school for each student type, where agents may misreport their private preferences. However, these works do not consider the planner's budget constraint and cannot guarantee the budget-feasibility. In addition, in our model, we aim to optimize the proportionality of agents in different groups, *i.e.*, maximize the minimum selection ratio of agents among groups, which is much different from the required diversity constraints in previous works.

3. Preliminaries

In this section, we define the budget-feasible proportion-representative selection settings and the desirable properties of the mechanisms.

3.1. The model

There is a planner and a set of n agents $S = \{s_1, s_2, ..., s_n\}$. The agents have group attributes, specifying one or more groups the agent belongs to, *e.g.*, genders, ages, ethnicities, regions, and educational levels. There are m groups $G = \{G_1, G_2, ..., G_m\}$, where group G_j is a non-empty subset of S, *i.e.*, $\emptyset \neq G_j \subseteq S$ and $G_1 \cup G_2 \cup \cdots \cup G_m = S$. Let $\mathcal{G}(s_i)$ denote the set of groups that agent s_i belongs to. Let n_j be the number of agents in group G_j, *i.e.*, $|G_j| = n_j$. Denote by n_{min} and n_{max} the minimum and maximum number of agents among all the groups respectively, *i.e.*, $n_{min} = \min_{1 \le j \le m} n_j$, $n_{max} = \max_{1 \le j \le m} n_j$. The agents are to be selected by the planner for proportional representation.

The planner has a budget $B \in \mathbb{R}_+$ and each agent s_i has a private cost $c_i \in \mathbb{R}_+$ (*e.g.*, her required cost for time, privacy, or fees) when selected to represent her group(s). We use $\mathbf{c} = (c_i)_{i=1}^n$ to denote agents' costs. Let $\mathbf{c}_{-i}$ denote all costs except s_i's cost c_i. Moreover, agent s_i has a value v_i, and agents are homogeneous if values are identical, *i.e.*, $v_i = 1, \forall 1 \le i \le n$, otherwise, agents are heterogeneous. Let $\mathbf{v} = (v_i)_{i=1}^n$ denote agents' values. The agents may act strategically to maximize their own utilities by misreporting their costs. Each agent bids a cost b_i that may be different from her real cost c_i in order to maximize her utility (defined below). Let $\mathbf{b} = (b_i)_{i=1}^n$ denote agents' bid profile and $\mathbf{b}_{-i}$ denote all bids except s_i's bid b_i. We sometime use $(b_i, \mathbf{b}_{-i})$ to represent $\mathbf{b}$ to highlight s_i's bid. We use $P = (p_i)_{i=1}^n$ to denote agents' payments.

3.2. The mechanism

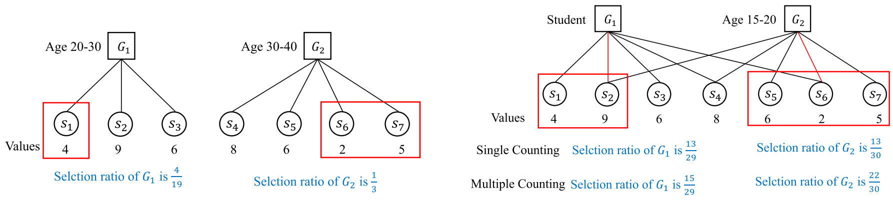
A mechanism $M = (X, P)$ consists of an allocation rule $X(\mathbf{b}) = (x_i)_{i=1}^n : \mathbb{R}_+^n \to \{0, 1\}^n$ deciding the selected agents (who are chosen by the planner) and a payment scheme $P(\mathbf{b}) : \mathbb{R}_+^n \to \mathbb{R}_+^n$ deciding the payment to each agent. Denote by S_w the selected agent set, *i.e.*, $S_w = \{s_i \mid x_i(\mathbf{b}) = 1\}$. Given a mechanism M, the utility of agent s_i is defined as the difference between the payment she receives and her true cost, *i.e.*,

$$u_i(b_i, \mathbf{b}_{-i}) = p_i(\mathbf{b}) - x_i(\mathbf{b}) \cdot c_i. \tag{1}$$

We consider both the single group model problem (SGP) where each agent belongs to only one group, and the multiple group model problem (MGP) each agent may belong to multiple groups. Let Q_j denote the total value of selected agents in group G_j. Next, we formally define two problems.

Single Group Model Problem (SGP): Since each agent belongs to only one group, we have $|\mathcal{G}(s_i)| = 1$. Then, we have $Q_j = \sum_{s_i \in G_j} x_i v_i$.

Multiple Group Model Problem (MGP): In this model, an individual agent may belong to multiple groups, *i.e.*, $1 \le |\mathcal{G}(s_i)| \le m$. Depending on whether each selected agent can be counted into all groups, we further consider two cases: the Single Counting (MGP-SC) case where a selected agent is counted just once in one of the groups she belongs to, and



(a) An example of single group setting. (b) An example of multiple group setting.

Fig. 1. Examples for the problem model.

the Multiple Counting (MGP-MC) case where a selected agent is counted in all groups she belongs to. For example, when forming a committee, the agent selected can only represent one of the groups to which she belongs, or when she is selected, all the groups to which the agent belongs are happy.

(1) MGP-SC: Each selected agent is only included in the selected agents of the group she is matched to. Let $x_{ij} = 1$ indicate that agent s_i is matched to group $G_j \in \mathcal{G}(s_i)$, otherwise, $x_{ij} = 0$. We also have $x_i = \sum_{j:G_j \in \mathcal{G}(s_i)} x_{ij} \leq 1$ where $x_{ij} \in \{0, 1\}, \forall 1 \leq i \leq n$. Thus, we have $Q_j = \sum_{s_i \in G_j} x_{ij} v_i$. Notice that the allocation rule of the mechanism is now defined over x_{ij} to additionally include matching a selected agent to one of their groups. We will use this allocation rule when the context is clear.

(2) MGP-MC: Each selected agent is counted in all groups she belongs to. Thus, we have $Q_j = \sum_{s_i \in G_j} x_i v_i$.

To obtain a proportion-representative selection of agents, we define the *selection ratio* of group G_j as $\frac{Q_j}{\sum_{s_i \in G_j} v_i}$, representing the ratio between the total value of selected agents and the total value of agents in group G_j.

Example 1. We now use the example in Fig. 1 to further explain our model. As shown in Fig. 1(a), we consider the SGP problem in which there are two groups: group G_1 contains students aged 20 to 30 and Group G_2 contains students aged 30 to 40. Specifically, we have $G_1 = \{s_1, s_2, s_3\}, G_2 = \{s_4, s_5, s_6, s_7\}$. We use rectangles and circles to denote groups and students, respectively. Numbers below circles are students' values, and there exists a line between a student and a group if such a student belongs to this group. Each student can only belong to one age group, *i.e.*, each student has one line connected to a group. Suppose we have chosen student s_1 from group G_1, *i.e.*, $x_1 = 1$ and the total value of the chosen student in group G_1 is $Q_1 = 4$, then the selection ratio of G_1 is $\frac{4}{4+9+6} = \frac{4}{19}$. If we select one more student s_2, *i.e.*, $x_1 = 1, x_2 = 1$ and $Q_1 = 4 + 9 = 13$, then the selection ratio of G_1 will be $\frac{13}{19}$. In Fig. 1(b), we construct an instance of the MGP problem with two groups: student group G_1 contains all students and age group G_2 contains students aged 15 to 20. Students may belong to multiple groups (*e.g.*, student s_2 aged 18 belongs to group G_1 and G_2 at the same time). Specifically, $G_1 = \{s_1, s_2, s_3, s_4, s_6\}, G_2 = \{s_2, s_4, s_5, s_6, s_7\}$. Suppose that we have chosen students s_1, s_2, s_5, s_6, s_7. In the single counting case, each chosen student can only represent one of the groups she belongs to, *e.g.*, in Fig. 1(b), the chosen students s_2, s_6 represent student group G_1 and G_2, *i.e.*, $x_{21} = 1, x_{22} = 0$ and $x_{61} = 0, x_{62} = 1$, respectively. Then, the selection ratio of G_1 and G_2 are $\frac{4+9}{4+9+6+8+2} = \frac{13}{29}$ and $\frac{6+2+5}{9+8+6+2+5} = \frac{13}{30}$, respectively. In the multiple counting case, each chosen student can represent all groups she belongs to. Thus, the selection ratio of G_1 and G_2 are $\frac{4+9+2}{4+9+6+8+2} = \frac{15}{29}$ and $\frac{6+2+5+9}{9+8+6+2+5} = \frac{22}{30}$.

3.3. Our proportion-representative objective and goal

Given the above models, we aim to maximize the minimum selection ratio of groups, *i.e.*,

$$\max \min_{1 \leq j \leq m} \frac{Q_j}{\sum_{s_i \in G_j} v_i},$$

when designing budget-feasible proportion-representative mechanisms.[3] This objective is based on a well-studied notion of proportional representation (*e.g.*, the electoral system requires the minimum vote proportion which is the ratio between the number of voters and population in each group for representation [24,25], specifying an exact allocation of votes to different groups proportionally [26] and maximizing the minimum diversity fairness among different groups [27,28]). Moreover, we want the designed proportion-representative mechanism M to satisfy the following properties:

- **Budget-feasibility.** The total payment of the planner does not exceed her budget B, *i.e.*, $\sum_{1 \leq i \leq n} p_i(\mathbf{b}) \leq B$.
- **Individual rationality.** The utility of each agent s_i is non-negative, *i.e.*, $u_i(b_i, \mathbf{b}_{-i}) \geq 0$ for any $\mathbf{b}$.
- **Truthfulness.** Each agent achieves the maximum utility by bidding her real cost, *i.e.*, $u_i(c_i, \mathbf{b}_{-i}) \geq u_i(b_i, \mathbf{b}_{-i})$ for any b_i.

[3] We refer to mechanisms that aim to maximize the minimum selection ratio of groups as proportion-representative mechanisms.

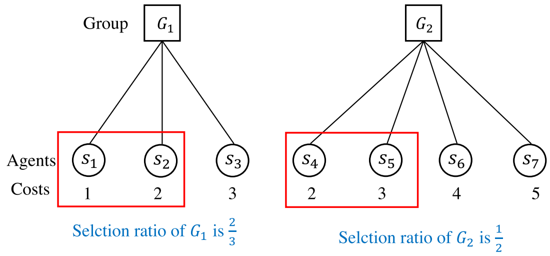


Fig. 2. An example where the optimal solution is manipulable.

- **Computational efficiency.** The output of the mechanism can be computed in polynomial time with respect to the number of agents and groups.
- **Approximation.** Let $ALG(I)$ be the minimum selection ratio among groups of the proposed mechanism M on input instance I. We compare the output of the mechanism with the minimum selection ratio of the optimal solution where agents' costs are known in advance. Formally, the optimal solution can be formulated as the solution of the following integer program,

$$\begin{aligned} &\max \min_{1\le j\le m} \frac{Q_j}{\sum_{s_i\in G_j} v_i} \\ &s.t.\ x_i \in \{0,1\}, \forall i \le n \\ &\sum_{i\le n} c_i \cdot x_i \le B \end{aligned} \tag{2}$$

For the MGP-SC case, the above integer program can be modified to include variables that are defined over group (*i.e.*, x_{ij}) and constraints that enforce single counting (*i.e.*, $\sum_{j:G_j\in\mathcal{G}(s_i)} x_{ij} \le 1$). The Q_j terms are defined accordingly as in the above subsections for different models. We say that a mechanism is β-approximate if $ALG(I) \ge \frac{1}{\beta} OPT(I)$ for any instance I.

Notice that, in the above properties, we are interested in mechanisms that approximately optimize the objective because an optimal solution cannot induce truthfulness in general. Below, we provide such an example. As shown in Fig. 2, we construct an instance where there are two groups $G_1 = \{s_1, s_2, s_3\}$ and $G_2 = \{s_4, s_5, s_6, s_7\}$ with costs $\{1, 2, 3\}$ and $\{2, 3, 4, 5\}$, respectively. The planner has a budget 8.5. The optimal solution will choose agent s_1, s_2, s_4, s_5 and pay the chosen agent reported cost. Then, if agent s_1 reported cost 1.5, then she will still be selected and achieve 1.5 payment which is higher than 1. Therefore, if we run the optimal solution, agents have incentives to report false costs to increase their utilities.

4. Homogeneous agent setting

We first consider homogeneous agents who have identical values, *i.e.*, $v_i = 1, \forall 1 \le i \le n$. Then the selection ratio in each group reduces to the ratio between the number of selected agents and the number of agents in this group. In Subsection 4.1 and Subsection 4.2, we develop mechanisms for the single group models and multiple group models (with single counting and multiple counting). When developing these mechanisms, we consider ideas from the standard budget-feasible mechanism literature. More specifically, the standard approach for designing budget-feasible mechanisms has leveraged a greedy approach to select agents based on their value-to-cost ratios. However, this approach cannot be used to achieve proportional representations directly as the approach does not consider groups. As our initial starting point, we investigate whether a similar greedy approach can work for our setting. It turns out that we can design new mechanisms with appropriate selection and payment schemes for each group using a greedy approach (along with other techniques such as integer program and max-min flow in different models/settings). The proposed (new greedy) mechanisms can (approximately) achieve the objective of proportional representations while guaranteeing truthfulness and budget-feasibility.

4.1. Mechanism for the single group model

In this section, we start with the single group model and introduce a **B**udget-feasible **P**roportion-representative mechanism for the **S**ingle **G**roup model (BPSG).

The main idea of Mechanism BPSG is as follows. We first generate a *virtual ratio set* which contains all possible selection ratios for each group when selecting a different number of agents from this group. In order to maximize the minimum selection ratio among all groups within the budget constraint, we find all feasible virtual ratios which ensure that the selection ratio for each group does not fall below that ratio and the current total payment (based on the ratio) does not exceed the budget. Specifically, the payments for the selected agents depend on the bids of agents after the last selected agent in each group, and thus the ratio which results in selecting all agents from some group will not be feasible. Among all these feasible virtual ratios, we use the maximum one as the final selection ratio for all groups.

The detail of Mechanism BPSG is shown in Algorithm 1. In order to distinguish the agents belonging to different groups, we use b_i^j to denote the i-th agent's (agent s_i^j) bid in group G_j (i.e., we sort all agents in the same group $G_{j:1\le j\le m}$ in the weakly increasing order of their bids, *i.e.*, $b_1^j \le b_2^j \le \cdots \le b_{n_j}^j$). Denote by p_i^j the payment for agent s_i^j. We then generate a *virtual ratio set* $\mathcal{R}$ which consists of possible selection ratios among all groups, *i.e.*,

$$\mathcal{R} = \cup_{0\le i\le n_j, 1\le j\le m}\left\{\frac{i}{n_j}\right\}, \tag{3}$$

and sort all ratios in the weakly increasing order of their values, where γ_l is the l-th element in $\mathcal{R}$, *i.e.*, $\gamma_1 < \gamma_2 < \cdots < \gamma_l < \cdots < \gamma_{|\mathcal{R}|}$. Denote by r_f (*final base selection ratio*) the minimum selection ratio among groups in the solution that the mechanism selects.

To find the final base selection ratio, we iteratively consider ratios in $\mathcal{R}$ starting with the first ratio γ_1.[4] Suppose that we are now considering ratio γ_l for $l > 1$. Let $I_j(\gamma_l)$ denote the minimum number of agents which ensures that the selection ratio in group G_j is at least γ_l, and we thus have $I_j(\gamma_l) = \lceil \gamma_l \cdot n_j \rceil$. Specifically, since Mechanism BPSG decides the payment for each chosen agent as the lowest bid of non-selected agents in each group, BPSG will select up to $n_j - 1$ agents from group G_j to ensure that we have at least one non-selected agent in this group. Thus, when trying ratio γ_l, BPSG will terminate and output γ_{l-1} as the final base selection ratio if there exists group G_j with $I_j(\gamma_l) = n_j$ (line 6). If we have $I_j(\gamma_l) < n_j, \forall 1 \le j \le m$, we compute the current payment for each of the first $I_j(\gamma_l)$ agents in group G_j as the bid of agent $s^j_{I_j(\gamma_l)+1}$, *i.e.*, $p_i^j = b^j_{I_j(\gamma_l)+1}$, $\forall 1 \le i \le I_j(\gamma_l)$. Thus, when all groups have a selection ratio of at least γ_l, the total payment, denoted by P_{γ_l}, is

$$P_{\gamma_l} = \sum_{1\le j\le m} I_j(\gamma_l) \cdot b^j_{I_j(\gamma_l)+1}. \tag{4}$$

It is easy to see that P_{γ_l} is increasing with γ_l. If $P_{\gamma_l} \le B$, we continue to try the next ratio γ_{l+1}. Otherwise, the *final base selection ratio* is $r_f = \gamma_{l-1}$.

Once we decide the final base selection ratio, we then determine the final selected agents and corresponding payments. Let k_j denote the number of selected agents in group G_j, *i.e.*, $k_j = I_j(r_f) = Q_j$. In each group G_j, the first k_j agents are selected, *i.e.*, $s_i^j \in S_w$, $\forall 1 \le i \le k_j$, and we have $k_j < n_j$. Then we have

$$p_i^j = \begin{cases} b^j_{k_j+1}, & if\ s_i^j \in S_w \\ 0, & otherwise. \end{cases} \tag{5}$$

Example 2 *(A running example of Mechanism BPSG).* Suppose there are nine agents who can be divided into two groups G_1 and G_2. Group G_1 has four agents $G_1 = \{s_1^1, s_2^1, s_3^1, s_4^1\}$ with bids $\{1, 1.5, 3, 4\}$ and group G_2 has five agents $G_2 = \{s_1^2, s_2^2, s_3^2, s_4^2, s_5^2\}$ with bids $\{0.5, 1.5, 2, 3, 5\}$. Thus, we have $n_1 = 4$ and $n_2 = 5$. The virtual ratio set is $\mathcal{R} = \{0, 0.2, 0.25, 0.4, 0.5, 0.6, 0.75, 0.8, 1\}$. The planner has a budget $B = 10$. We now try virtual ratios by starting from the first non-zero ratio 0.2:

(1) Try ratio 0.2: We have $I_1(0.2) = 1$ and $I_2(0.2) = 1$. Thus, we will select s_1^1 from G_1 and pay her 1.5, and select s_1^2 from G_2 and pay her 1.5. Then, the total payment is $3 < B = 10$, and we will try the next ratio of 0.25.

(2) Try ratio 0.25: We have $I_1(0.25) = 1$ and $I_2(0.25) = 2$. Thus, we will select s_1^1 from G_1 and pay her 1.5, and select s_1^2, s_2^2 from G_2 and pay each of them 2. Then, the total payment is $5.5 < 10$, and we will try the next ratio of 0.4.

(3) Try ratio 0.4: We have $I_1(0.4) = 2$ and $I_2(0.4) = 2$. Thus, we will select s_1^1, s_2^1 from G_1 and pay each of them 3, and select s_1^2, s_2^2 from G_2 and pay each of them 2. Then, the total payment is $10 = B$, and we will try the next ratio of 0.5.

(4) Try ratio 0.5: We have $I_1(0.5) = 2$ and $I_2(0.5) = 3$. Thus, we will select s_1^1, s_2^1 from G_1 and pay each of them 3, and select s_1^2, s_2^2, s_3^2 from G_2 and pay each of them 3. Then, the total payment is $15 > B = 10$.

Then, Mechanism BPSG terminates with the final base selection ratio 0.4. The selected agent set is $S_w = \{s_1^1, s_2^1, s_1^2, s_2^2\}$ with payments $p_1^1 = 3, p_2^1 = 3, p_1^2 = 2, p_2^2 = 2$, while the payments for the unselected agents are zero.

Next, we analyze the performance of Mechanism BPSG. We first provide a well-known Myerson's characterization for truthful mechanisms in the single parameter domain which we will rely on to show the truthfulness of the proposed mechanisms.

[4] The first virtual ratio in $\mathcal{R}$ is 0 and note that at this ratio, we select no agents and pay each agent zero.

Algorithm 1: Mechanism **BPSG**(B, b, S, G).

Input: B, b, S, G.
Output: P, S_w

$P \leftarrow 0, S_w \leftarrow \emptyset$;
Sort agents in $G_j (\forall 1 \le j \le m)$, in the weakly increasing order of their bids $b_1^j \le b_2^j \le \cdots \le b_{n_j}^j$ and generate the virtual ratio set $\mathcal{R}$ with value sorted and indexed by γ_l's;
// **Determine the final base selection ratio**;
for $1 \le l \le |\mathcal{R}|$ **do**
  Compute $I_j(\gamma_l) = \lceil \gamma_l \cdot n_j \rceil$ for any $1 \le j \le m$;
  if $I_j(\gamma_l) < n_j, \forall 1 \le j \le m$ **then**
    Compute the payment P_{γ_l} according to (4);
    if $P_{\gamma_l} \le B$ **then**
      $l \leftarrow l + 1$;
    else
      break;
    end
  else
    break;
  end
end
$r_f \leftarrow \gamma_{l-1}$;
// **Agent selection and payment scheme**;
Add agent $s_i^j (\forall 1 \le j \le m)$ with $1 \le i \le k_j = I_j(r_f)$ into the selected agent set S_w;
Decide the payments to agents according to (42);
return P, S_w

Theorem 1. *(**Monotone Theorem, [59]**) In the single parameter domains,*[5] *a mechanism $M = (X, P)$ guarantees sellers' truthfulness if and only if:*

*(1) X **is monotone:** $\forall s_i \in S$, if $b_i \le c_i$, then $x_i(c_i, \mathbf{c}_{-i}) = 1$ implies $x_i(b_i, \mathbf{c}_{-i}) = 1$ for every $\mathbf{c}_{-i}$;*
*(2) **winners are paid threshold payments**: the payment to each winning bidder is the critical value* $\inf\{c_i : x_i(c_i, \mathbf{c}_{-i}) = 0\}$.

The above theorem shows that truthful mechanisms satisfy monotonicity and agents are paid the threshold payments. Monotonicity means that when the selected agent reports a lower cost, she will still be selected. Threshold payments guarantee that if an agent reports her cost higher than the threshold payment, this agent will not be selected. We prove the truthfulness of Mechanism BPSG by leveraging the theorem above. Without loss of generality, we now assume that the final base selection ratio $r_f = \gamma_l$ is the l-th element in virtual set $\mathcal{R}$.

Theorem 2. *Mechanism BPSG guarantees truthfulness.*

Proof. Note that the virtual set $\mathcal{R}$ is generated by the number of agents in each group, which will not change when any agent bids any false cost. Recall that, in group G_j, agent s_i^j is the i-th agent and the last selected agent is $s_{k_j}^j$ where $k_j < n_j$. Let O_j denote the order of agents in group G_j, *i.e.*, $O_j = \langle s_1^j, s_2^j, \ldots, s_{n_j}^j \rangle$. In addition, the payment to each selected agent $s_i^j (\forall 1 \le i \le k_j)$ in group G_j is $b_{k_j+1}^j$ and all other agents will receive a payment of zero. Since the final base selection ratio is $r_f = \gamma_l$, we have

$$P_{\gamma_l} = \sum_{1 \le h \le m} k_h \cdot b_{k_h+1}^h \le B. \tag{6}$$

Let $\hat{k}_j$ denote the number of selected agents in group G_j when the selection ratio is at least γ_{l+1}, *i.e.*, $\hat{k}_j = I_j(\gamma_{l+1})$. Depending on the final base selection ratio determined in Mechanism BPSG, we consider the following two cases:

Case 1: There exists no group G_j in which all its agents are selected when considering γ_{l+1}, *i.e.*, $\hat{k}_j = I_j(\gamma_{l+1}) < n_j, \forall 1 \le j \le m$. As γ_{l+1} is not selected by the mechanism, we thus have

$$P_{\gamma_{l+1}} = \sum_{1 \le h \le m} \hat{k}_h \cdot b_{\hat{k}_h+1}^h > B \tag{7}$$

since the final base selection ratio is $r_f = \gamma_l$. Additionally, it is clear that

[5] The single parameter setting means that each agent only holds one type of private information. In our model, it is easy to see that each agent has only one kind of private information, that is, the private cost.

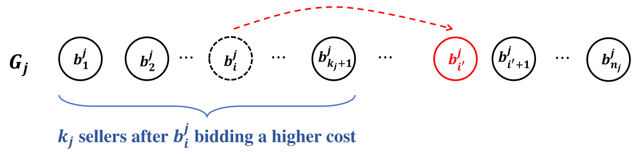


Fig. 3. An example of a selected agent s_i^j bidding a higher bid for the proof of Theorem 2. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

$$\gamma_l \le \frac{k_j}{n_j} < \frac{k_j+1}{n_j}, \forall 1 \le j \le m. \tag{8}$$

According to the generation of virtual set $\mathcal{R}$ in (3), we must have $\frac{k_j+1}{n_j} \in \mathcal{R}$ for each j, and then $\gamma_{l+1} \le \frac{k_j+1}{n_j}$ since $\gamma_{l+1} = \min_{l<h\le|\mathcal{R}|} \gamma_h$ is the next ratio after γ_l. Thus, we have

$$\hat{k}_j = I_j(\gamma_{l+1}) = \lceil \gamma_{l+1} \cdot n_j \rceil \le k_j + 1 \tag{9}$$

and

$$b^j_{\hat{k}_j+1} \le b^j_{\min\{n_j, k_j+2\}}. \tag{10}$$

We next introduce new notations after agent s_i^j reports a false cost when fixing the costs of the remaining agents. Let r'_f, k'_j and O'_j denote the new final base selection ratio, the number of selected agents in group G_j, and the new order of agents, respectively.

Monotonicity: For any selected agent s_i^j in group G_j, *i.e.*, $1 \le i \le k_j$, if she reports a lower cost $b^j_{i'}$ where $b^j_{i'} < b^j_i \le b^j_{k_j+1}$, it is easy to see that the final base selection ratio is still γ_l, and she will still be selected. Therefore, Mechanism BPSG is monotonic.

Threshold payments: We consider two cases: (1) If agent s_i^j with $1 \le i \le k_j$ reports a cost $b^j_{i'}$ which is $b^j_{i'} > b^j_{n_j} \ge b^j_{k_j+1}$, she is the last agent in the new order O'_j, and will never be selected since Mechanism BPSG never selects all agents in group G_j for each j. (2) As shown in Fig. 3, assume that agent s_i^j with $1 \le i \le k_j$ reports a cost $b^j_{i'}$ which is $b^j_{i'} \in (b^j_{k_j+1}, b^j_{n_j}]$, and thus changes her position to that of $s^j_{i'}$ marked by the red circle, *i.e.*, the new order of agents in group G_j is

$$O'_j = \langle s^j_1, s^j_2, \ldots, s^j_{i-1}, s^j_{i+1}, \ldots, s^j_{k_j+1}, \ldots, s^j_{i'}, \ldots, s^j_{n_j} \rangle \tag{11}$$

while the original order of agents is $O_j = \langle s^j_1, s^j_2, \ldots, s^j_i, s^j_{i+1}, \ldots, s^j_{k_j+1} \ldots, s^j_{n_j} \rangle$. Denote by $O'_j(i')$ the number of agents from s^j_1 to $s^j_{i'}$ in (11). If she is not selected after bidding such a false cost, her utility is zero. If she is still a selected agent, agents from s^j_1 to $s^j_{i'}$ in order (11) should be selected, and we have $O'_j(i') \ge k_j + 1$. Due to $\lceil r_f \cdot n_j \rceil = k_j$ when s_i^j reports her real cost, the new final base selection ratio r'_f must satisfy $\lceil r'_f \cdot n_j \rceil \ge O'_j(i') \ge k_j + 1 \ge \lceil \gamma_{l+1} \cdot n_j \rceil$ due to (9) which implies $r'_f \ge \gamma_{l+1}$. Thus, under ratio r'_f, the mechanism will select the first $k'_h = \lceil r'_f \cdot n_h \rceil \ge \lceil \gamma_{l+1} \cdot n_h \rceil = \hat{k}_h$ agents from group G_h for $h \le m$ and $h \ne j$, and at least $O'_j(i')$ agents from group G_j. Let $s^j_{i'+1}$ denote the agent after $s^j_{i'}$ in group G_j. Then, the payment under ratio r'_f is

$$\begin{aligned} P_{r'_f} &\ge O'_j(i') \cdot b^j_{i'+1} + \sum_{1\le h\le m, h\ne j} k'_h \cdot b^h_{k'_h+1} \\ &\ge (k_j+1) \cdot b^j_{\hat{k}_j+1} + \sum_{1\le h\le m, h\ne j} \hat{k}_h \cdot b^h_{\hat{k}_h+1} \\ &\ge \sum_{1\le h\le m} \hat{k}_h \cdot b^h_{\hat{k}_h+1} > B \end{aligned} \tag{12}$$

There are three reasons for the second inequality in (12): (i) $O'_j(i') \ge k_j + 1$, (ii) Since we have $b^j_{i'} > b^j_{k_j+1}$, we have $i' + 1 \ge k_j + 2$. Then we have $b^j_{i'+1} \ge b^j_{\hat{k}_j+1}$ due to $b^j_{i'+1} \ge b^j_{k_j+2}$ and Eq. (10), (iii) $k'_h \ge \hat{k}_h$. The third and the last equality in (12) is due to (10) and (7), respectively. Thus, the total payment is higher than budget B if s_i^j is selected after bidding false cost. Then, by the above contradiction, s_i^j will not be selected and have zero utility. Therefore, the value $b^j_{k_j+1}$ is the critical value and the selected agents are paid threshold payments.

Case 2: There exists a group G_j in which all its agents are selected when considering γ_{l+1}, *i.e.*, $\exists j, 1 \le j \le m, \hat{k}_j = I_j(\gamma_{l+1}) = n_j$**.** As $r_f = \gamma_l$ is the final base selection ratio, Mechanism BPSG will not select all agents from G_j with ratio r_f, *i.e.*, $\lceil r_f \cdot n_j \rceil \le n_j - 1$. If $\lceil r_f \cdot n_j \rceil < n_j - 1$, we have $\lceil \gamma_{l+1} \cdot n_j \rceil \le n_j - 1$ since γ_{l+1} is the minimum virtual ratio in $\mathcal{R}$, which implies $I_j(\gamma_{l+1}) \le n_j - 1$, contradicting to $I_j(\gamma_{l+1}) = n_j$. Thus, we have

$$r_f \cdot n_j = \gamma_l \cdot n_j = n_j - 1, \tag{13}$$

and the payment for the selected agent in group G_j is $b^j_{n_j}$.

Monotonicity: If the selected agent s^j_i in group G_j reports a lower cost $b^j_{i'} < b^j_i \le b^j_{n_j}$, it is obvious that the final base selection ratio is sill γ_l, and she will still be selected. Thus, Mechanism BPSG is monotonic.

Threshold payments: Note that the payment for the selected agent s^j_i in group G_j is $b^j_{n_j}$ due to (13). If she reports a cost higher than $b^j_{n_j}$, she will be the last agent and never be selected. In addition, for the selected agent s^h_i in group $G_{h:h \ne j}$, the payment is $b^h_{I_h(\gamma_l)+1}$. If she reports a cost higher than $b^h_{I_h(\gamma_l)+1}$, she will never be selected unless the new final base selection ratio is higher than $r_f = \gamma_l$, which results in that there will exist a group G_j in which all its agents are selected, leading to the contradiction.

Therefore, Mechanism BPSG guarantees truthfulness by Theorem 1. □

Then, we show that Mechanism BPSG guarantees individual rationality, budget-feasibility, and computational efficiency.

Theorem 3. *Mechanism BPSG guarantees individual rationality, budget feasibility, and computational efficiency.*

Proof. 1) Individual rationality: Since Mechanism BPSG is truthful, we have $b^j_i = c^j_i$ where c^j_i is the true cost of agent s^j_i. For each selected agent s^j_i, we have $b^j_i \le b^j_{k_j+1}$ where $i \le k_j$ in group G_j, and her payment is $b^j_{k_j+1}$ which implies that her utility is $b^j_{k_j+1} - c^j_i = b^j_{k_j+1} - b^j_i \ge 0$ which is non-negative. **2) Budget-feasibility:** After determining the selection ratio, it is easy to see that the total payment is $\sum_{1 \le j \le m} k_j \cdot b^j_{k_j+1} \le B$ which is no greater than the budget B. **3) Computational efficiency:** The running time of Mechanism BPSG is dominated by the sorting (line 2) and the loop in determining the final base selection ratio (line 4-15) as shown in Algorithm 1. Therefore, the total computational complexity is $O(n \log n)$. This completes the proof. □

Next, we introduce a useful property of Mechanism BPSG. Let $r_j(\gamma_l)$ denote the selection ratio of group G_j when the final base selection ratio is γ_l (after selecting k_j agents), *i.e.*, $r_j(\gamma_l) = \frac{k_j}{n_j}$. We use r_{max} and r_{min} to denote the maximum and minimum selection ratios among groups when the final base selection ratio is γ_l, *i.e.*, $r_{max} = \max_{1 \le j \le m}\{r_j(\gamma_l)\}$ and $r_{min} = \min_{1 \le j \le m}\{r_j(\gamma_l)\}$. Specifically, we have $r_{min} = \gamma_l = r_f$ since there must exist at least one group G_j whose selection ratio is $r_j(\gamma_l) = \gamma_l$ due to the generation of $\mathcal{R}$ in (3). Denote by α the ratio between n_{max} and n_{min}, *i.e.*, $\alpha = \frac{n_{max}}{n_{min}}$.

Lemma 1. *Mechanism BPSG has the following two properties:*

1. $\gamma_{h+1} - \gamma_h \le \frac{1}{n_{max}}, \forall 1 \le h \le |\mathcal{R}| - 1$.
2. $r_{max} - r_{min} < \frac{1}{n_{min}}$.

Proof. (1) Assume that group G_j has the maximum total number of agents, *i.e.*, $n_j = n_{max}$. There must exist an integer i, $0 \le i \le n_j - 1$, such that $\frac{i}{n_j} \le \gamma_h < \frac{i+1}{n_j}, \forall h, 1 \le h \le |\mathcal{R}|$ due to the generation of the virtual ratio set $\mathcal{R}$. Since the next virtual ratio γ_{h+1} is the minimum virtual ratio after γ_h, we thus have

$$\gamma_{h+1} - \gamma_h \le \frac{i+1}{n_j} - \frac{i}{n_j} = \frac{1}{n_j} = \frac{1}{n_{max}}. \tag{14}$$

(2) Assume that r_{max} and r_{min} are the selection ratios in group G_{j_1} and G_{j_2}, respectively. We have $r_{max} = \frac{k_{j_1}}{n_{j_1}} \ge \gamma_l$, and $\frac{k_{j_1}-1}{n_{j_1}} < \gamma_l = r_{min}$ since the number of selected agents in G_{j_1} is k_{j_1} rather than $k_{j_1} - 1$. Then, we have

$$r_{max} - r_{min} < \frac{k_{j_1}}{n_{j_1}} - \frac{k_{j_1}-1}{n_{j_1}} \le \frac{1}{n_{min}}. \quad \square \tag{15}$$

Given the above lemma, we consider the approximation guarantee of Mechanism BPSG. Let ALG and OPT denote the minimum selection ratio of BPSG and the optimal solution, respectively. Our analysis considers two separate cases. In the

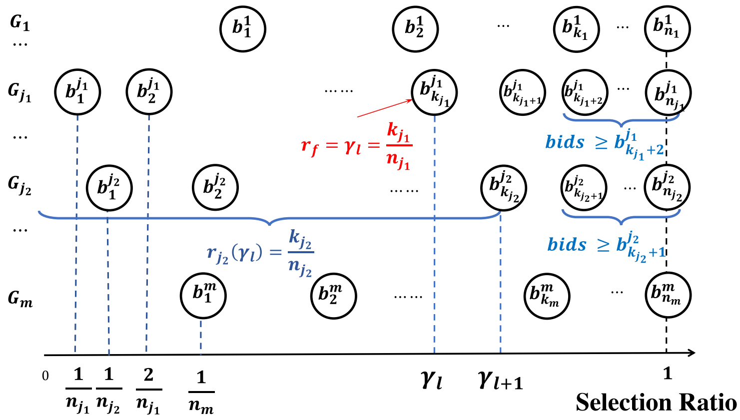


Fig. 4. An example for the proof of Theorem 4.

first case when BPSG is able to select at least one agent from each group, we show that BPSG obtains an approximation ratio that depends on the ratio between the maximum and minimum number of agents among groups. In fact, the ratio is asymptomatically tight as we will show later in Theorem 5. For the second case, where Mechanism BPSG cannot select at least one agent from each group, we show that no budget-feasible proportion-representative mechanisms can output a better solution than BPSG.

Theorem 4. *(1) Mechanism BPSG achieves $(3+\alpha)$-approximation ratio if BPSG selects at least one agent from each group (i.e., $ALG \geq \frac{1}{n_{max}}$) where $\alpha = \frac{n_{max}}{n_{min}}$.*

(2) No budget-feasible proportion-representative mechanism $\mathcal{M}$, that guarantees truthfulness and individual rationality, can achieve $ALG_{\mathcal{M}} \geq \theta$ for any $\theta > 0$ where $ALG_{\mathcal{M}}$ is the solution of $\mathcal{M}$ if Mechanism BPSG cannot select at least one agent from each group, i.e., $ALG = 0$.

Proof. (1) As shown in Fig. 4, we use circles to indicate the agents, and we sort agents in the weakly increasing order of their bids in each group, *e.g.*, $b_1^{j_1} \leq b_2^{j_1} \leq \cdots \leq b_{n_{j_1}}^{j_1}$ in group G_{j_1}. The distance between any two neighboring agents s_i^j, s_{i+1}^j in group G_j is the marginal gain of the selection ratio after adding s_{i+1}^j into the selected agent set where the first i agents have been selected. For example, the distance between $s_1^{j_1}$ and $s_2^{j_1}$ is $\frac{1}{n_{j_1}}$.

Suppose that the final base selection ratio is the l-th element in the virtual ratio set $\mathcal{R}$, and we have $r_f = \gamma_l = ALG$ and $k_j = I_j(\gamma_l)$ is the number of selected agents in group G_j. We have $n_j \geq 2, \forall 1 \leq j \leq m$ since BFSG can select at least one agent from each group by the design of the mechanism. Depending on the conditions at the termination of BPSG, we consider the following two cases:

Case 1: There exists a group G_j in which we select all agents when considering γ_{l+1}, i.e., $\exists j \leq m, I_j(\gamma_{l+1}) = n_j$. We have $\gamma_l \cdot n_j = n_j - 1$ according to (13). Thus, we have $ALG \geq \min_{1\leq j\leq m}\{\frac{n_j-1}{n_j}\}$ and $OPT \leq 1$, which implies $\frac{OPT}{ALG} \leq \frac{1}{\min_{1\leq j\leq m}\{\frac{n_j-1}{n_j}\}} \leq 2$.

Case 2: There exists no group G_j in which all its agents are selected when considering γ_{l+1}, i.e., $I_j(\gamma_{l+1}) < n_j, \forall 1 \leq j \leq m$. Let G' denote the set of all groups that can choose exactly $r_f \cdot n_j = k_j$ agents making the selection ratio in group G_j equal to r_f, $\forall G_j \in G'$. As shown in Fig. 4, G_{j_1} is one of the groups in G'. Thus, we have $r_j(r_f) > r_f, \forall G_j \notin G'$. In addition, the next virtual ratio γ_{l+1} must satisfy

$$\gamma_{l+1} = \min\left\{\min_{G_j\notin G'}\{r_j(r_f)\}, \min_{G_j\in G'}\left\{\frac{k_j+1}{n_j}\right\}\right\}. \tag{16}$$

According to the payment scheme of Mechanism BPSG, we have

$$P_{\gamma_l} = \sum_{1\leq j\leq m} k_j \cdot b^j_{k_j+1} \leq B. \tag{17}$$

Depending on whether the next ratio γ_{l+1} is generated from a group in G' or not, we consider the following two subcases:

Subcase 1: Assume that the virtual ratio γ_{l+1} is generated from group $G_{j_2} \notin G'$ where $j_1 \neq j_2$ as shown in Fig. 4, *i.e.*,

$$\gamma_{l+1} = r_{j_2}(r_f) = \min\left\{\min_{G_j\notin G'}\{r_j(r_f)\}, \min_{G_j\in G'}\left\{\frac{k_j+1}{n_j}\right\}\right\}. \tag{18}$$

Thus, we have $r_j(r_f) \geq r_{j_2}(r_f) = \gamma_{l+1}, \forall G_j \notin G'$, and

$$k_j = \lceil r_f \cdot n_j \rceil \leq \lceil \gamma_{l+1} \cdot n_j \rceil \leq \lceil r_j(r_f) \cdot n_j \rceil = k_j$$

which means the number of selected agents with ratio γ_{l+1} in group G_j where $G_j \notin G'$ does not change and should be k_j. For the group G_j in G', the number of selected agents will be $k_j + 1$ since $\frac{k_j}{n_j} = \gamma_l$ and $\frac{k_j+1}{n_j} \geq \gamma_{l+1}$ due to (18). Thus, we have $k_j + 1 = I_j(\gamma_{l+1}) \leq n_j - 1$. Since γ_l is the final base selection ratio, the total payment with ratio γ_{l+1} exceeds the budget, *i.e.*,

$$P_{\gamma_{l+1}} = \sum_{G_j \in G'} (k_j + 1) \cdot b^j_{k_j+2} + \sum_{G_j \notin G'} k_j \cdot b^j_{k_j+1} > B. \tag{19}$$

We divide agents in S into two parts: $\tilde{S} = \{s^j_i \mid i \leq k_j, \forall G_j \notin G'\} \cup \{s^j_i \mid i \leq k_j + 1, \forall G_j \in G'\}$ and $S \setminus \tilde{S}$ for further analysis to bound the minimum selection ratio among groups in each part.

For set $\tilde{S}$: The optimal solution can select all agents in $\tilde{S}$ with cost zero in the best case and spend the budget on the remaining agents in $S \setminus \tilde{S}$. Since OPT selects k_j agents from each group G_j not in G', and $k_j + 1$ agents from each group G_j in G', we have

$$\max_{G_j \notin G'} \frac{k_j}{n_j} = \max_{G_j \notin G'} r_j(r_f) = r_{max}$$

and

$$\max_{G_j \in G'} \frac{k_j + 1}{n_j} = r_f + \max_{G_j \in G'} \frac{1}{n_j} = r_{min} + \frac{1}{n_{min}}.$$

Thus, after choosing all agents in $\tilde{S}$, the maximum selection ratio among all groups is $\max\{r_{max}, r_{min} + \frac{1}{n_{min}}\} \leq r_{min} + \frac{1}{n_{min}}$ according to Lemma 1, and the minimum selection ratio is γ_{l+1} due to (18).

For set $S \setminus \tilde{S}$: For group G_j where $G_j \notin G'$, the bids of agents after s_{k_j} are at least $b^j_{k_j+1}$, while the bids of agents after $b^j_{k_j+1}$ are at least $b^j_{k_j+2}$ in group $G_j \in G'$. With budget B, the optimal solution can select k_j agents from the remaining agents after s_{k_j} with costs $b^j_{k_j+1}$ from group G_j for any $G_j \notin G'$, and $k_j + 1$ agents with costs $b^j_{k_j+2}$ from the group $G_j \in G'$ due to Eq. (19). According to the analysis for the set $\tilde{S}$, the minimum selection ratio among groups by selecting these numbers of agents is γ_{l+1} and the maximum selection ratio is $r_{min} + \frac{1}{n_{min}}$. Otherwise, OPT should select more agents to achieve a higher base selection ratio which will exceed the budget.

In set $\tilde{S}$ and $S \setminus \tilde{S}$, we know that the maximum and minimum selection ratios in these two parts for OPT are $r_{min} + \frac{1}{n_{min}}$ and γ_{l+1}, respectively. By combining these two parts, the minimum selection ratio among groups for OPT is no greater than the sum of $r_{min} + \frac{1}{n_{min}}$ and γ_{l+1}, *i.e.*, $OPT \leq r_{min} + \frac{1}{n_{min}} + \gamma_{l+1} \leq 2r_f + \frac{1}{n_{max}} + \frac{1}{n_{min}}$ where the last inequality is due to the first property in Lemma 1.

Subcase 2: Assume that the virtual ratio γ_{l+1} is generated from group $G_{j_2} \in G'$, *i.e.*,

$$\gamma_{l+1} = \frac{k_{j_2} + 1}{n_{j_2}} = \min\left\{\min_{G_j \notin G'}\{r_j(r_f)\}, \min_{G_j \in G'}\left\{\frac{k_j + 1}{n_j}\right\}\right\}.$$

Using similar arguments in the Subcase 1, we have $OPT \leq 2r_f + \frac{1}{n_{max}} + \frac{1}{n_{min}}$.

Thus, we have $OPT \leq 2ALG + \frac{1}{n_{min}} + \frac{1}{n_{max}}$. Specifically, when Mechanism BPSG can select at least one agent from each group, *i.e.*, $ALG \geq \frac{1}{n_{max}}$, we have $\frac{OPT}{ALG} \leq 3 + \alpha$ which means that Mechanism BPSG achieves $(3 + \alpha)$-approximation ratio.

(2) Assume that there are m groups, and we sort agents in weakly increasing order of their bids in each group. Suppose that the budget of the planner is B. According to Mechanism BPSG, there are two cases as follows due to $ALG = 0$:

Case i) The total payment is $\sum_{1 \leq j \leq m} b^j_2$ where b^j_2 denotes the second lowest bid in group G_j when trying the first non-zero virtual ratio in $\mathcal{R}$, that is, choosing the first agent from each group, and we have $\sum_{1 \leq j \leq m} b^j_2 > B$ due to $ALG = 0$ which implies that we cannot select any agent from each group. Assume for the purpose of contradiction that there exists a budget-feasible proportion-representative mechanism $\mathcal{M}$ with better performance $ALG_{\mathcal{M}} \geq \theta$. In this case, Mechanism $\mathcal{M}$ must select at least one agent from each group. Myerson's characterization [59] implies that the threshold payment for the first agent in group G_j should be at least b^j_2, and thus the total payment now is $\sum_{1 \leq j \leq m} b^j_2 > B$. By truthfulness and budget-feasibility, $\mathcal{M}$ cannot select one agent from each group using that payment which contradicts $ALG_{\mathcal{M}} \geq \theta$.

Case ii) There exists a group G_j which contains only one agent, and we will select all agents from G_j when trying the first non-zero virtual ratio which implies that we cannot select at least one agent from each group. Assume that we have $ALG_{\mathcal{M}} \geq \theta$. In this case, Mechanism $\mathcal{M}$ must always select at least one agent from each group. Thus, the payment for the

Algorithm 2: Mechanism **BPMG-S**(B, b, S, G).

Input: B, b, S, G.
Output: P, S_w

$P \leftarrow 0, S_w \leftarrow \emptyset$;
Assign each agent s_i a weight $w_i = 2^{z_i}$ where $z_i \in N_+$ is an arbitrary integer such that no two agents have the same weight, *i.e.*, $z_i \neq z_{i'}$ for any $i \neq i'$;
Sort all agents in weakly increasing order of their bids b_i, *i.e.*, $b_1 \leq b_2 \leq \cdots \leq b_n$;
// **Determine the candidate agent set**;
$k \leftarrow 0$;
for $1 \leq i \leq n$ **do**
  Compute $r(s_i)$ according to (20);
  $F(s_i) \leftarrow \sum_{1 \leq j \leq m} \lceil r(s_i) \cdot n_j \rceil$;
  if $b_i \cdot F(s_i) \leq B$ **then**
    $k \leftarrow k + 1$;
  else
    break;
  end
end
// **Agent selection and payment scheme**;
$X, S_w \leftarrow AgentSelect(S_k)$;
The payment for $s_i \in S_w$ is $p_i = \min\{\frac{B}{F(s_k)}, b_{k+1}\}$;

agent in G_j can be infinity by Myerson's characterization [59] which violates budget-feasibility. Therefore, $\mathcal{M}$ also cannot select at least one agent from each group using that payment which contradicts $ALG_{\mathcal{M}} \geq \theta$. □

Next, we provide a lower bound for all budget-feasible proportion-representative mechanisms.

Theorem 5. *No budget-feasible proportion-representative mechanism obtains an approximation ratio better than* $\Omega(\alpha)$.

Proof. Suppose that we have m groups. Group G_1 has n_1 agents with costs $0, B-\epsilon, \ldots, B-\epsilon$, while group G_m has n_m agents, each with cost ϵ. Each remaining group $G_j (\forall j, 2 \leq j \leq m-1)$ has n_j agents with cost zero. Specifically, assume that $n_1 \leq n_2 \leq \cdots \leq n_m$ and $B = \frac{1}{n_1} n_m \epsilon$. Thus, we have $n_{min} = n_1$ and $n_{max} = n_m$. By Myerson's characterization [59], the payment for the first agent in G_1 is at least $B - \epsilon$ and for the first agent in G_m is at least ϵ to achieve truthfulness. Thus, by truthfulness and budget-feasibility, the solution of any budget-feasible proportion-representative mechanism is no better than $\frac{1}{n_m}$ since they can only select at most one agent each from G_1 and G_m. However, the optimal solution is $\frac{1}{n_1}$. Thus, no budget-feasible proportion-representative mechanism can achieve an approximation ratio better than $\frac{n_m}{n_1} = \alpha$. This finishes the proof. □

In particular, it is worth stating that this lower bound also applies to the multiple group model which we will consider in the next section. Furthermore, based on the instance in Theorem 5, we can directly obtain the lower bound for the approximation ratio of Mechanism BPSG, that is also α.

4.2. Mechanisms for multiple group models

In this section, we consider the multiple group model where each agent s_i might belong to multiple groups, *i.e.*, $1 \leq |\mathcal{G}(s_i)| \leq m$. We distinguish two subcases according to whether the contribution of the agent is counted only once or not: single-counting case and multiple-counting case. In the single counting case, the selection ratio in one of the groups $\mathcal{G}(s_i)$, say G_j, would increase by $\frac{1}{n_j}$ where $n_j = |G_j|$ if s_i is selected and contributes to G_j; while in the multiple counting case, the selection ratios of all groups in $\mathcal{G}(s_i)$ would increase by $\frac{1}{n_j}$ for any $G_j \in \mathcal{G}(s_i)$, once s_i is selected.

4.2.1. Single counting case

In this section, we introduce a **B**udget-feasible **P**roportion-representative mechanism for the **M**ultiple **G**roup model in **S**ingle counting case, called BPMG-S.

Intuitively, the mechanism first considers the available agents under a given fixed payment. Given all agents who have bids lower than this payment, we can compute the solution that maximizes the minimum selection ratio among groups. This minimum ratio will increase as the given payment increases. We then try to find the payment which can maximize such a ratio and ensure budget-feasibility simultaneously.

In detail, we first sort all the agents in weakly increasing order of their bids $b_1 \leq b_2 \leq b_3 \leq \cdots \leq b_n$. Let $x_{ij} = 1$ indicate that agent s_i is matched to group G_j, otherwise, $x_{ij} = 0$. Agent s_i can only be matched to at most one group, *i.e.*, $\sum_{1 \leq j \leq m} x_{ij} \leq 1, \forall 1 \leq i \leq n$. Denote by $\mathcal{S}(s_h) = \{s_i \mid 1 \leq i \leq h\}$ the set containing agents before agent s_{h+1}.

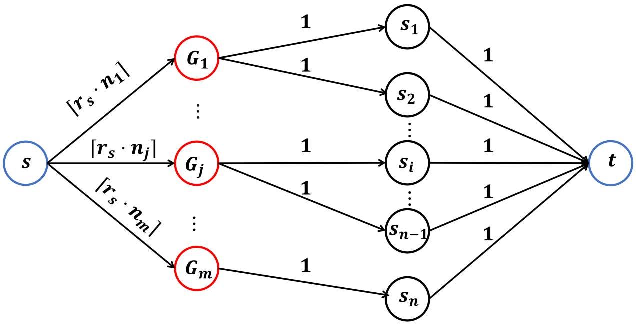


Fig. 5. An example of a flow network under ratio r_s.

Algorithm 3: Function **AgentSelect**(S, G, k).

Input: S, G, k.
Output: X, S_w

$X \leftarrow 0$, $S_w \leftarrow \emptyset$;
Compute the final allocation $X_w(s_k)$ according to (22);
$X \leftarrow X_w(s_k)$;
Add agent $s_i (\forall 1 \le i \le n)$ into S_w if $\sum_{1\le j\le m} x_{ij} = 1$;

Before introducing our mechanism, we first introduce an important component, an integer program formulation to compute a matching under a given agent set $\mathcal{S}(s_h)$, that maximizes the minimum selection ratio among groups while ignoring the costs and budget constraint, denoted by $ILP(s_h)$ as follows,

$$\begin{aligned} &\max \min_{1\le j\le m} \frac{\sum_{1\le i\le h} x_{ij}}{n_j} \\ s.t., &\sum_{1\le j\le m} x_{ij} \le 1, \forall 1 \le i \le h \\ &x_{ij} \in \{0,1\}, \forall 1 \le i \le h, G_j \in \mathcal{G}(S_i) \\ &x_{ij} = 0, \forall 1 \le i \le h, G_j \notin \mathcal{G}(S_i) \end{aligned} \tag{20}$$

where the three conditions indicate that agents in $\mathcal{S}(s_h)$ can only be matched to at most one of the groups they belong to. Notice that the optimal solution in (20) can be computed in polynomial time by constructing Max-Flow networks as follows. In detail, given an agent set $\mathcal{S}(s_h)$, all possible selection ratios for group G_j are in $\cup_{y\le|\mathcal{S}(s_h)\cap G_j|}\{\frac{y}{n_j}\}$. Thus, the optimal selection ratio in the solution of (20) must be one of the ratios in set $R = \cup_{y\le|\mathcal{S}(s_h)\cap G_j|, j\le m}\{\frac{y}{n_j}\}$. We say a ratio $r_s \in R$ is *feasible* if we can find a matching result ensuring that the selection ratio in each group is at least r_s and satisfying all conditions in (20). To find the optimal solution, we take $r_s \in R$ as input and construct a flow network based on such a ratio. As shown in Fig. 5, we use blue, red and black circles to represent source node s/terminal node t, groups and agents, respectively. Specifically, there exists a directed edge from source s to each group G_j, an edge from each G_j to s_i if $s_i \in G_j$, and from each agent s_i to terminal node t. Each edge from s to G_j is assigned with a capacity $\lceil r_s \cdot n_j \rceil$, while the capacity of each directed edge from G_j to s_j and from s_i to t are all 1. If the maximum flow on this network equals to $\sum_{1\le j\le m}\lceil r_s \cdot n_j \rceil$, then r_s is a feasible ratio. We can find the optimal solution in (20) by testing every possible input of r_s and validating its feasibility, *i.e.*, the maximum one among all feasible ratios is the desired optimal solution.

Let $X(s_h)$ denote the solution of $ILP(s_h)$, and $r(s_h)$ be the minimum selection ratio among groups under $X(s_h)$, *i.e.*, $r(s_h) = \min_{1\le j\le m} \frac{\sum_{1\le i\le h} x_{ij}}{n_j}, x_{ij} \in X(s_h)$. Denote by $F(s_h)$ the total number of agents under ratio $r(s_h)$, *i.e.*,

$$F(s_h) = \sum_{1\le j\le m} \lceil r(s_h) \cdot n_j \rceil. \tag{21}$$

It is not hard to see that $F(s_h) \le |\mathcal{S}(s_h)|$.

Candidate agent selection Before selecting agents, we first decide on a candidate agent set from which we finally select agents. We iteratively consider each agent's bid starting from the first agent's bid b_1. Suppose that we are now trying agent s_i and calculate the value of $F(s_i)$ by Eq. (21). If $b_i \cdot F(s_i) \le B$, we consider the next agent s_{i+1}, otherwise, we select agents from the previous $i-1$ agents $\mathcal{S}(s_{i-1})$. Assume that s_k is the last agent who satisfies $b_k \cdot F(s_k) \le B$ which implies $b_{k+1} \cdot F(s_{k+1}) > B$. After determining the value k, we define agent set $\mathcal{S}(s_k) = \{s_i : i \le k\}$ as the *candidate agent set*, where $|\mathcal{S}(s_k)| \ge F(s_k)$ follows.

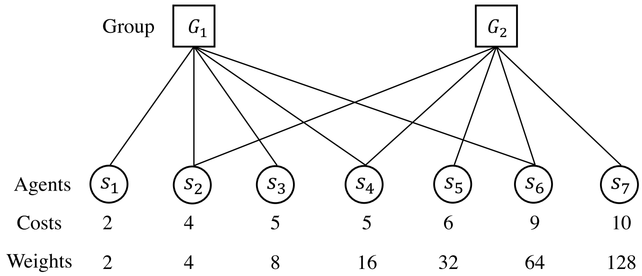


Fig. 6. Running example of Mechanism BPMG-S.

Agent selection and payment scheme Now we select agents from the candidate agent set $\mathcal{S}(s_k)$. At the beginning of the mechanism, we assign each agent s_i a weight $w_i = 2^{z_i}$ where $z_i \in N_+$ is an arbitrary integer such that no two agents have the same weight, *i.e.*, $z_i \neq z_{i'}$ for any $i \neq i'$. Then, we use the function $AgentSelect(S, G, k)$ to select $F(s_k)$ agents from $\mathcal{S}(s_k)$. The detail of $AgentSelect(S, G, k)$ is shown in Algorithm 3. We try to find a minimum weight matching between agents and groups that can minimize the total weight of matched agents satisfying that each group G_j has $\lceil r(s_k) \cdot n_j \rceil$ matched agents from set $\mathcal{S}(s_k)$ as follows,[6] denoted by $ILP_w(s_k)$,

$$\begin{aligned} &\min \sum_{1\le i\le k} \sum_{1\le j\le m} w_i x_{ij} \\ s.t., &\sum_{1\le j\le m} x_{ij} \le 1, \forall i \le k \\ &\sum_{1\le i\le k} x_{ij} = \lceil r(s_k)\cdot n_j \rceil, \forall 1 \le j \le m \\ &x_{ij} \in \{0,1\}, \forall i \le k, G_j \in \mathcal{G}(S_i) \\ &x_{ij} = 0, \forall i \le k, G_j \notin \mathcal{G}(S_i). \end{aligned} \tag{22}$$

Since we assign agents exponential weights, there is a unique matching result in (22). Let $X_w(s_k)$ denote the solution of $ILP_w(s_k)$. If agent $s_{i:i\le k}$ is matched to one of the groups, *i.e.*, $\sum_{1\le j\le m} x_{ij} = 1$, then she is selected, *i.e.*, $s_i \in S_w$. The payment for each selected agent is $p_i = \min\{\frac{B}{F(s_k)}, b_{k+1}\}$ while the payments for unselected agents are zero.[7]

Example 3. We now show a running example of Mechanism BPMG-S. As shown in Fig. 6, suppose there are seven agents and two groups G_1, G_2. Group G_1 has five agents $G_1 = \{s_1, s_2, s_3, s_4, s_6\}$ with bids $\{2, 4, 5, 5, 9\}$ and group G_2 has five agents $G_2 = \{s_2, s_4, s_5, s_6, s_7\}$ with bids $\{4, 5, 6, 9, 10\}$. Thus, we have $n_1 = 5$ and $n_2 = 5$. The planner has a budget $B = 30$.

(1) Given agent set $\mathcal{S}(s_2) = \{s_1, s_2\}$: According to the optimal solution of Eq. (20), the minimum selection ratio among groups under $\mathcal{S}(s_2)$ is 0.2, and we have $F(\mathcal{S}(s_2)) = 2$, where s_1 is matched to group G_1 and s_2 is matched to group G_2 by Eq. (22). And we have $F(\mathcal{S}(s_2)) \cdot b_3 = 10 < B = 30$. Thus, we will select s_1, s_2 and pay each of them $\min\{\frac{30}{2}, 5\} = 5$. Then, we will try agent set $\mathcal{S}(s_3) = \{s_1, s_2, s_3\}$.

(2) Given agent set $\mathcal{S}(s_3) = \{s_1, s_2, s_3\}$: According to the optimal solution of Eq. (20), the minimum selection ratio among groups under $\mathcal{S}(s_3)$ is still 0.2, and we have $F(\mathcal{S}(s_3)) = 2$, where s_1 is matched to group G_1 and s_2 is matched to group G_2 by Eq. (22). And we have $F(\mathcal{S}(s_3)) \cdot b_4 = 10 < B = 30$. Thus, we will select s_1, s_2 and pay each of them $\min\{\frac{30}{2}, 5\} = 5$. Then, we will try agent set $\mathcal{S}(s_4) = \{s_1, s_2, s_3, s_4\}$.

(3) Given agent set $\mathcal{S}(s_4) = \{s_1, s_2, s_3, s_4\}$: According to the optimal solution of Eq. (20), the minimum selection ratio among groups under $\mathcal{S}(s_4)$ is 0.4, and we have $F(\mathcal{S}(s_4)) = 4$, where s_1, s_3 is matched to group G_1 and s_2, s_4 is matched to group G_2 by Eq. (22). And we have $F(\mathcal{S}(s_4)) \cdot b_5 = 24 < B = 30$. Thus, we will select s_1, s_2, s_3, s_4 and pay each of them $\min\{\frac{30}{4}, 6\} = 6$. Then, we will try agent set $\mathcal{S}(s_5) = \{s_1, s_2, s_3, s_4, s_5\}$.

(4) Given agent set $\mathcal{S}(s_5) = \{s_1, s_2, s_3, s_4, s_5\}$: According to the optimal solution of Eq. (20), the minimum selection ratio among groups under $\mathcal{S}(s_4)$ is 0.4, and we have $F(\mathcal{S}(s_5)) = 4$, where s_1, s_3 is matched to group G_1 and s_2, s_4 is matched to group G_2 by Eq. (22). And we have $F(\mathcal{S}(s_5)) \cdot b_6 = 36 > B = 30$.

Then, Mechanism BPMG-S terminates with the final base selection ratio 0.4. The selected agent set is $S_w = \{s_1, s_2, s_3, s_4\}$ and each of them gets payment 6 while the payments for the unselected agents are zero.

Next, we analyze the performance of Mechanism BPMG-S.

[6] This problem can be similarly solved in polynomial time by constructing Max-Flow Min-Cost networks.

[7] The payment for each selected agent is $\frac{B}{F(s_k)}$ if $k = n$.

Theorem 6. *Mechanism BPMG-S guarantees individual rationality, budget-feasibility, and computational efficiency.*

Proof. The proof for the first two properties is straightforward, and we will not go into detail. For computational efficiency, the running time of Mechanism BPMG-S is dominated by the loop in determining the candidate agent set (line 4-12) in Algorithm 2, and the agent selection. The complexity of solving the problem in (20) is $O(mn^3(n+m))$ by using the Max-Flow method [60] where n and m are the number of agents and the number of groups, respectively. When computing the candidate agent set, we try n agents, and thus the total complexity is $O(mn^4(n+m))$. For the agent selection determination, the complexity of solving (22) is $O(n(n+m)\log n)$ by using the Min-Cost Max-Flow method [31]. Therefore, the total computational complexity is $O(mn^4(n+m))$. □

We prove the truthfulness of Mechanism BPMG-S by showing that it satisfies Theorem 1.

Theorem 7. *Mechanism BPMG-S guarantees truthfulness.*

Proof. According to the definition of $F(s_{k+1})$, it is obvious that $F(s_{k+1}) \geq F(s_k)$. Since s_k is the last agent among candidate agents, we have,

$$b_k \cdot F(s_k) \leq B, b_{k+1} \cdot F(s_{k+1}) > B \tag{23}$$

The payment for agent $s_i \in S_w$ is $p_i = \min\{\frac{B}{F(s_k)}, b_{k+1}\}$. Denote by $O = \langle s_1, s_2, ..., s_k, ..., s_n \rangle$ the order of agents in weakly increasing order of their bids.

Monotonicity: For any agent s_i with $i \leq k$, s_i is selected when bidding real cost c_i. If s_i decreases her bid to $b'_i < c_i$, agents before s_{k+1} will not change and the candidate agent set is still $\mathcal{S}(s_k)$ due to (23). Thus, s_i will still be selected since the unique matching under the same agent set in the agent selection scheme by (22). Therefore, Mechanism BPMG-S is monotonic.

Threshold payments: For any selected agent $s_i \in S_w$, her payment is $p_i = \min\{\frac{B}{F(s_k)}, b_{k+1}\}$. According to the relationship between $\frac{B}{F(s_k)}$ and b_{k+1}, we further consider two subcases:

- $\frac{B}{F(s_k)} \leq b_{k+1}$: For agent s_i with $i \leq k$ and $s_i \in S_w$, we have $p_i = \frac{B}{F(s_k)}$ and $b_i \leq \frac{B}{F(s_k)}$. If she bids a higher cost $b'_i > p_i$ (let $s_{i'}$ denote the agent after bidding a false cost), she will be after agent s_k, and the new order now is

$$O' = \langle s_1, ..., s_{i-1}, s_{i+1}, ..., s_k, ..., s_{i'}, ..., s_n \rangle \tag{24}$$

  Let $\mathcal{S}(s_h)'$ denote agents before s_h (including herself) in the new sequence, and $F(s_h)'$ denote the new total number of agents with ratio computed by (20) with input $\mathcal{S}(s_h)'$. We have $\mathcal{S}(s_k) \subseteq \mathcal{S}(s_{i'})'$ which means $F(s_{i'})' \geq F(s_k)$. Thus, we must have $b'_i \cdot F(s_{i'})' > p_i \cdot F(s_k) = B$ which implies that agent s_i will not be selected, and her utility is zero. Therefore, the value $\frac{B}{F(s_k)}$ is the critical value and the selected agents are paid threshold payments.
- $b_{k+1} < \frac{B}{F(s_k)}$: For agent s_i with $i \leq k$ and $s_i \in S_w$, we have $p_i = b_{k+1}$. If she bids a higher cost $b'_i > p_i$, it is easy to find that $\mathcal{S}(s_{k+1}) \subseteq \mathcal{S}(s_{i'})'$ and $F(s_{i'})' > F(s_{k+1})$. Thus, we have $b'_i \cdot F(s_{i'})' > b_{k+1} \cdot F(s_{k+1}) > B$ which implies that agent s_i will not be selected. Therefore, the value b_{k+1} is the critical value and the selected agents are paid threshold payments.

Therefore, Mechanism BPMG-S guarantees truthfulness. □

Let OPT_s and ALG_s denote the optimal solution and the solution of BPMG-S, respectively. We show that BPMG-S generally achieves an approximation ratio with respect to the size of groups when at least one agent is selected in each group, while there is a small gap between BPMG-S and any other budget-feasible proportion-representative mechanism when BPMG-S cannot select at least one agent from each group.

Theorem 8. *(1) Mechanism BPMG-S achieves $(m\alpha(\alpha+2)+1)$-approximation ratio if BPMG-S can select at least one agent from each group (i.e., $ALG_s \geq \frac{1}{n_{max}}$) where $\alpha = \frac{n_{max}}{n_{min}}$.*

(2) No budget-feasible proportion-representative mechanism $\mathcal{M}'$, that guarantees truthfulness and individual rationality, can obtain $ALG_{\mathcal{M}'} \geq \theta$ for any $\theta > \frac{1}{n_{min}}$ where $ALG_{\mathcal{M}'}$ is the solution of $\mathcal{M}'$ if Mechanism BPMG-S cannot select at least one agent from each group, i.e., $ALG_s = 0$.

Proof. **(1)** The bids of agents before s_{k+1} are no greater than b_{k+1} while the bids of agents after s_{k+1} are no lower than b_{k+1}. According to the Mechanism BPMG-S, we have $ALG_s = r(s_k)$. For the agents before s_{k+1}, the optimal solution OPT_s can select all these k agents with zero costs in the best case. Then, the minimum selection ratio among these agents is still $r(s_k)$.

For the agents after s_k, OPT_s can select at most $F(s_{k+1})$ agents due to (23). For set $\mathcal{S}(s_k)$, the maximum increase in the minimum selection ratio among all groups is $\max_{1\le j\le m}\frac{1}{n_j}=\frac{1}{n_{min}}$ after adding agent s_{k+1} into $\mathcal{S}(s_k)$, *i.e.*, $r(s_{k+1})\le r(s_k)+\frac{1}{n_{min}}$. Thus, we have

$$F(s_{k+1})=\sum_{1\le j\le m}\lceil r(s_{k+1})\cdot n_j\rceil \le \sum_{1\le j\le m}\left\lceil\left(r(s_k)+\frac{1}{n_{min}}\right)\cdot n_j\right\rceil. \tag{25}$$

In the best case, the maximum increase in the minimum selection ratio among all groups is $\frac{F(s_{k+1})}{n_{min}}$, that is, adding $F(s_{k+1})$ agents to $\mathcal{S}(s_k)$ and matching all $F(s_{k+1})$ agents to the group with the minimum number of agents. It follows that

$$\begin{aligned}\frac{F(s_{k+1})}{n_{min}} &\le \sum_{1\le j\le m}(\lceil r(s_k)\cdot n_j+\frac{n_j}{n_{min}}\rceil)/n_{min}\\ &\le \sum_{1\le j\le m}\frac{r(s_k)\cdot n_j+\frac{n_j}{n_{min}}+1}{n_{min}}\\ &\le m\cdot\left(r(s_k)\cdot\frac{n_{max}}{n_{min}}+\frac{1}{n_{min}}+\frac{n_{max}}{n_{min}^2}\right)\end{aligned}$$

where the first inequality and the last inequality are due to (25) and $n_j\le n_{max}$, respectively. Thus, we have $OPT_s\le r(s_k)+\frac{F(s_{k+1})}{n_{min}}\le(m\cdot\frac{n_{max}}{n_{min}}+1)\cdot ALG_s+m\cdot(\frac{1}{n_{min}}+\frac{n_{max}}{n_{min}^2})$.

Additionally, when Mechanism BPMG-S selects at least one agent from each group, *i.e.*, $ALG_s\ge\frac{1}{n_{max}}$, we have $\frac{OPT_s}{ALG_s}\le(m\cdot\frac{n_{max}}{n_{min}}+1)+\frac{m\cdot(\frac{1}{n_{min}}+\frac{n_{max}}{n_{min}^2})}{ALG_s}\le m\alpha(\alpha+2)+1$.

(2) Since s_k is the last agent among candidates, we have $F(s_k)=0$ when $ALG_s=0$ and consider the following two cases:

- $F(s_k)=0$ $(k=n)$: There is no feasible ratio in (20) with input $\mathcal{S}(s_n)$, *i.e.*, we cannot find a matching that allocates one agent to each group. Thus, it is obvious that no budget-feasible proportion-representative mechanism $\mathcal{M}'$ can achieve $ALG_{\mathcal{M}'}\ge\theta$ for any $\theta>0$.
- $F(s_k)=0$ $(k<n)$: We have $r(s_k)=0$, $b_k\cdot F(s_k)\le B$, and $b_{k+1}\cdot F(s_{k+1})>B$. For the set $\mathcal{S}(s_k)$, the maximum increase in the minimum selection ratio among all groups is $\max_{1\le j\le m}\frac{1}{n_j}=\frac{1}{n_{min}}$, *i.e.*, $r_{s_{k+1}}\le\frac{1}{n_{min}}$ after adding agent s_{k+1} into $\mathcal{S}(s_k)$. For the agents before s_{k+1}, Myerson's characterization [59] implies that the threshold payment for these agents should be at least b_{k+1}. For the agents after s_k, we have $b_i\ge b_{k+1},\forall i>k$. Thus, by truthfulness and budget-feasibility, mechanism $\mathcal{M}$ can select at most $F(s_{k+1})$ agents, which implies $ALG_{\mathcal{M}}=r(s_{k+1})\le\frac{1}{n_{min}}$.

This completes the proof. □

Next, we show that the lower bound of the approximation ratio for Mechanism BPMG-S is $\Omega(m\alpha)$ as follows. As shown in Fig. 7, there are m groups, where group G_j for $j\le m-1$ has n_j agents with costs 0, and group G_m has two agents with costs 0 while costs of the remaining agents are ϵ. Suppose that $n_1=n_2=\cdots=n_{m-1}>n_m$ and $\alpha=\frac{n_1}{n_m}$. Let $B=(\sum_{j\le m}\lceil\frac{2}{n_m}\cdot n_j\rceil-1)\epsilon$. When trying agent set $\mathcal{S}_k$, we have $F(\mathcal{S}_k)\cdot b_{k+1}=0$. Assume that $F(\mathcal{S}_{k+1})\cdot\epsilon=(\sum_{j\le m}\lceil\frac{2}{n_m}\cdot n_j\rceil)\cdot\epsilon>B$. Then, we have $ALG_s=\frac{1}{n_m}$. However, for the optimal solution, OPT_s can obtain the first $k+1$ agents with payment 0 and $(\sum_{j\le m}\lceil\frac{2}{n_m}\cdot n_j\rceil)-1$ agents under budget B. Thus, we have $OPT_s\ge\frac{2}{n_m}+\frac{\frac{2}{n_m}\cdot n_j\rceil)-1}{n_m}\ge(m-1)\frac{2n_1}{n_m^2}+3\frac{1}{n_m}$. Then, we have $\frac{OPT_s}{ALG_s}\ge(m-1)\alpha+3$.

4.2.2. Multiple counting case

In this section, we consider the multiple counting case where the selection ratio of each group G_j in $\mathcal{G}(s_i)$ can increase by $\frac{1}{n_j}$ for $G_j\in\mathcal{G}(s_i)$ when s_i is selected. We propose a modified version of Mechanism BPMG-S, called BPMG-M.

In general, Mechanism BPMG-M applies the framework in BPMG-S. We first sort all the agents in weakly increasing order of their bids $b_1\le b_2\le b_3\le\cdots\le b_n$. We then use an integer program formulation to compute a result given an agent set $\mathcal{S}(s_h)=\{s_i:i\le h\}$, that maximizes the minimum selection ratio among groups, denoted by $ILP^m(s_h)$ as follows,

$$\begin{aligned}&\max\min_{1\le j\le m}\frac{\sum_{s_i\in G_j}x_i}{n_j}\\ &s.t., x_i\in\{0,1\},\forall 1\le i\le h\\ &x_i=0,\forall i>h\end{aligned} \tag{26}$$

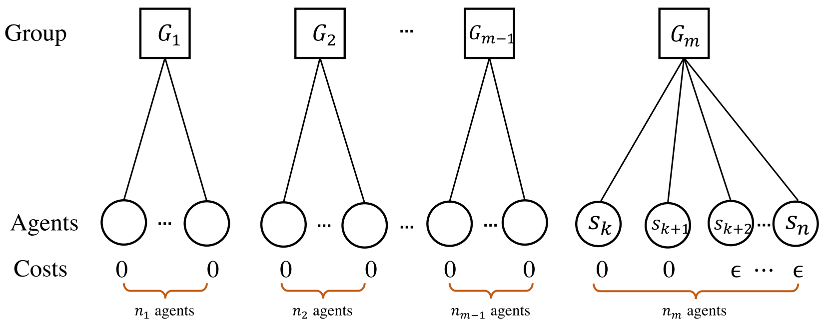


Fig. 7. Example for lower bound.

where the two conditions mean that we only consider agents in $\mathcal{S}(s_h)$. Similar to that of (20), we can solve this problem by using the Max-Flow method [60]. Let $X^m(s_h)$ denote the solution of $ILP^m(s_h)$, and $r_m(s_h)$ denote the minimum selection ratios under $X_m(s_h)$. We denote $F_m(s_h)$ as the total counted number of selected agents among groups by ratio $r_m(s_h)$, *i.e.*, $F_m(s_h) = \sum_{1 \le j \le m} \lceil r_m(s_h) \cdot n_j \rceil$.

Candidate agent selection We also decide on a set of candidate agents by iteratively testing each agent's bid starting from the first agent's bid b_1. That is, we find the last agent s_k who ensures $b_k \cdot F_m(s_k) \le B$ and $b_{k+1} \cdot F_m(s_{k+1}) > B$. Then, the agent set $\mathcal{S}(s_k)$ is the candidate agent set.

Agent selection and payment scheme Similar to the agent selection function $AgentSelect(S_k, G, B)$, at the beginning of the mechanism, we assign each agent s_i a weight $w_i = 2^{z_i}$ where $z_i \in N_+$ is an arbitrary integer such that no two agents have the same weight, *i.e.*, $z_i \ne z_{i'}$ for any $i \ne i'$. We try to select agents with the minimum total weight satisfying that at least $\lceil r_m(s_k) \cdot n_j \rceil$ agents in each group G_j are selected from set $\mathcal{S}(s_k)$ as follows,[8] denoted by $ILP^m_w(s_k)$,

$$\begin{aligned} &\min \sum_{i \le k} w_i x_i \\ s.t., &\sum_{i \in G_j} x_i \ge \lceil r_m(s_k) \cdot n_j \rceil, \forall 1 \le j \le m \\ &x_i \in \{0, 1\}, \forall i \le k \\ &x_i = 0, \forall i > k \end{aligned} \tag{27}$$

Let $X^m_w(s_k)$ denote the solution of $ILP^m_w(s_k)$. If $x_i = 1$, agent s_i is selected and her payment is $p_i = \min\{\frac{B}{F_m(s_k)}, b_{k+1}\}$, otherwise, $p_i = 0$.

Theorem 9. *Mechanism BPMG-M guarantees individual rationality, budget-feasibility, computational efficiency, and truthfulness.*

Proof. We refer to the earlier proofs in Section 4.2.1 for all the parts. □

Let ALG_m denote the solution of Mechanism BPMG-M.

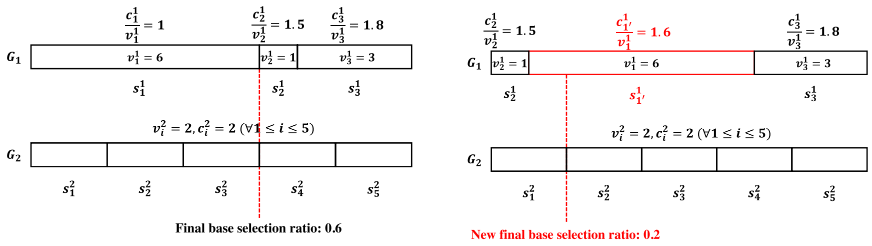
Theorem 10. *(1) Mechanism BPMG-M achieves $(m\alpha(\alpha+2)+1)$-approximation ratio if BPMG-M can obtain at least one agent from each group (i.e., $ALG_m \ge \frac{1}{n_{max}}$) where $\alpha = \frac{n_{max}}{n_{min}}$.*

(2) No budget-feasible proportion-representative mechanism $\mathcal{M}''$, that guarantees truthfulness and individual rationality, can obtain $ALG_{\mathcal{M}''} \ge \theta$ for any $\theta > \frac{2m}{n_{min}}$ where $ALG_{\mathcal{M}''}$ is the solution of $\mathcal{M}''$ if Mechanism BPMG-M cannot select at least one agent from each group, i.e., $ALG_m = 0$.

Proof. (1) The proof for the first property is similar to that in Theorem 8, and we will not repeat it here.

(2)Similar to the proof of the second property in Theorem 8, we still have $r_m(s_{k+1}) \le \frac{1}{n_{min}}$ after adding agent s_{k+1} into $\mathcal{S}(s_k)$ and $\mathcal{M}''$ can select at most $F_m(s_{k+1})$ agents by truthfulness and budget-feasibility. Thus, we have $F_m(s_{k+1}) \le \sum_{1 \le j \le m} \lceil \frac{1}{n_{min}} \cdot n_j \rceil$. Since each selected agent belongs to at most m groups, the minimum selection ratio is at most

[8] Similarly, this problem can be solved in polynomial time by constructing Min-Cost Max-Flow networks.



(a) Selected agents when s_1^1 reports true cost.

(b) Selected agents when s_1^1 misreports cost.

Fig. 8. An example shows that MPSG cannot be directly applied to heterogeneous agent scenarios.

$$\frac{F_m(s_{k+1})}{n_{max}} \le \frac{\sum_{1\le j\le m}(\frac{1}{n_{min}} \cdot n_j + 1)}{n_{max}} \le \frac{2m}{n_{min}}$$

when selecting $F_m(s_{k+1})$ agents. □

5. Heterogeneous agent setting

In the previous sections, we consider the proportional representation of groups where we want to obtain a proportional number of agents from each group. Such a setting corresponds to the case where the agents have identical values (i.e., $v_i = 1, \forall i \le n$). In this section, we consider the setting where agents have different values, *i.e.*, agent s_i has a value v_i which may be different from other agents. Based on the applicability of greedy approaches for designing mechanisms for the homogeneous setting, we naturally consider similar approaches for designing mechanisms for this setting. While our earlier mechanisms do not apply directly to the heterogeneous agent settings (see an example below), it turns out that we can extend our earlier mechanisms non-trivially to heterogeneous agents.

Let v_{max} and v_{min} denote the maximum and minimum value among agents, *i.e.*, $v_{max} = \max_{1\le i\le n} v_i$ and $v_{min} = \min_{1\le i\le n} v_i$. Moreover, let $\mathcal{V}_j = \sum_{s_i \in G_j} v_i$ denote the sum of values of agents in group G_j. Denote by $\mathcal{V}_{max}$ and $\mathcal{V}_{min}$ the maximum and minimum total value of agents among groups, *i.e.*, $\mathcal{V}_{max} = \max_{1\le j\le m} \mathcal{V}_j$ and $\mathcal{V}_{min} = \min_{1\le j\le m} \mathcal{V}_j$. We define the cost-per-value of agent s_i as $\frac{c_i}{v_i}$, which is useful to evaluate each agent's cost efficiency.

As agents have different values, our previous mechanisms cannot directly be applied to address heterogeneous agents due to the violation of truthfulness. For example, as shown in Fig. 8, we represent agents by rectangles while the lengths of rectangles indicate their values. We consider two groups $G_1 = \{s_1^1, s_2^1, s_3^1\}$ and $G_2 = \{s_1^2, s_2^2, s_3^2, s_4^2, s_5^2\}$. The cost-per-values and values of agents in group G_1 are $\{1, 1.5, 1.8\}$ and $\{6, 1, 3\}$, while the cost-per-value and value for each agent in group G_2 are 1 and 2 respectively. Assume that the budget of the planner is 15. By applying Mechanism BPSG, we select agent s_1^1 from group G_1 with payment 9, and select s_1^2, s_2^2, s_3^2 from group G_2 each with payment 2. Suppose that agent s_1^1 misreports her cost-per-value as 1.6 which changes her position to $s_{1'}^1$ as shown in Fig. 8(b). Then, agent $s_{1'}^1$ will be selected with payment $1.8 \times 6 = 10.8$ achieving more utility. Thus, it fails to ensure agents' truthfulness when agents have various values.

5.1. Mechanism for the single group model

To address the above challenges on ensuring agents' truthfulness, we introduce a **B**udget-feasible **P**roportion-representative mechanism for the **S**ingle **G**roup model with heterogeneous agents (BPSG-H) below.

5.1.1. Mechanism design

We mainly follow the framework of Mechanism BPSG for the homogeneous agent setting. However, to ensure agents' truthfulness, we design a novel payment scheme which finds threshold payments for the selected agents. The detail of Mechanism BPSG-H is shown in Algorithm 4.

We use b_i^j and v_i^j to denote the i-th agent's (agent s_i^j) bid and value in group G_j, respectively. That is, we sort all agents in the same group $G_{j:1\le j\le m}$ in weakly increasing order of their bid-per-values, *i.e.*, $\frac{b_1^j}{v_1^j} \le \frac{b_2^j}{v_2^j} \le \cdots \le \frac{b_{n_j}^j}{v_{n_j}^j}$. Denote by p_i^j the payment for agent s_i^j. We generate a *virtual ratio set* $\mathcal{R}_h$ according to the weakly increasing order of bid-per-values as follows,

$$\mathcal{R}_h = \cup_{0\le i\le n_j, 1\le j\le m} \left\{ \sum_{1\le h\le i} \frac{v_h^j}{\mathcal{V}_j} \right\} \tag{28}$$

We sort all ratios in weakly increasing order of their values where γ_l is the l-th element in $\mathcal{R}_h$.

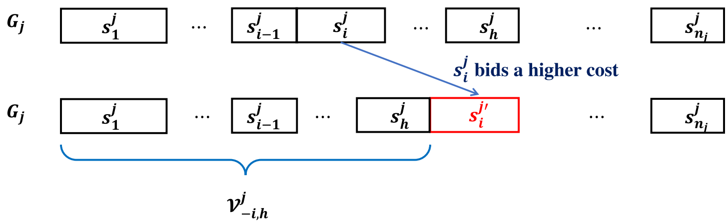


Fig. 9. An example of the position of agent s_i^j in group G_j after she bids a higher cost.

Agent selection We iteratively consider ratios in $\mathcal{R}_h$ starting with the first ratio γ_1. In the selection phase, we only use $\frac{B}{m}$ budget to select agents from all groups to ensure budget-feasibility. Suppose that we are now considering ratio γ_l. Recall that $I_j(\gamma_l)$ denotes the minimum number of agents which ensures that the selection ratio in group G_j is at least γ_l, *i.e.*,

$$I_j(\gamma_l) = \arg\min_{1\le i\le n_j} \{ \sum_{1\le h\le i} v_h \ge \gamma_l \cdot \mathcal{V}_j \}.$$

Specifically, Mechanism BPSG-H will select up to $n_j - 1$ agents from each group G_j (to ensure truthfulness). Thus, when trying ratio γ_l, BPSG-H will terminate and output γ_{l-1} as the final base selection ratio if there exists group G_j with $I_j(\gamma_l) = n_j$. When $I_j(\gamma_l) < n_j, \forall 1 \le j \le m$, we continue to try the next ratio γ_{l+1} if the following threshold holds,

$$P_{\gamma_l}^H = \sum_{1\le j\le m} \left(\frac{b^j_{I_j(\gamma_l)+1}}{v^j_{I_j(\gamma_l)+1}} \cdot \sum_{1\le i\le I_j(\gamma_l)} v_i^j \right) \le \frac{B}{m}. \tag{29}$$

It is not hard to see that $P_{\gamma_l}^H$ increases with γ_l. Otherwise, the *final base selection ratio* is $r_f = \gamma_{l-1}$. After deciding the final base selection ratio, we determine the final selected agents. Let k_j denote the number of selected agents in group G_j, *i.e.*, $k_j = I_j(r_f)$. In each group G_j, the first k_j agents are selected, *i.e.*, $s_i^j \in S_W$, $\forall i \le k_j$, and we have $k_j < n_j$.

Payment determination Next, we decide on the payment for each selected agent. For each selected agent $s_i^j (i \le k_j)$, we want to decide her responding threshold payment that is the maximum cost she can bid to be selected. As shown in Fig. 9, suppose s_i^j bids a higher cost $\frac{b_i^{j'}}{v_i^j} \in (\frac{b_h^j}{v_h^j}, \frac{b_{h+1}^j}{v_{h+1}^j}]$ where $k_j \le h < n_j$. We generate a new virtual ratio set $\mathcal{R}_h'$ under the new weakly increasing order of their bid-per-values. Let $\mathcal{V}_{-i,h}^j$ denote the sum of values of the first h agents in group G_j except value v_i^j. Agent s_i^j will still be selected if we can find a virtual ratio $r_{i,h} = \arg\min_{\gamma\in\mathcal{R}_h'} \{\gamma > \frac{\mathcal{V}_{-i,h}^j}{\mathcal{V}_j}\}$ under which the following threshold is still satisfied, *i.e.*,

$$P_{r_{i,h}}^H = (\mathcal{V}_{-i,h}^j + v_i^j) \frac{b_{h+1}^j}{v_{h+1}^j} + \sum_{u\ne j} \left(\frac{b^u_{I_u(r_{i,h})+1}}{v^u_{I_u(r_{i,h})+1}} \cdot \sum_{1\le i\le I_u(r_{i,h})} v_i^u \right) \le \frac{B}{m}.$$

Then, we find $h_i = \arg\max_{k_j\le h\le n_j-1} \{P_{r_{i,h}}^H \le \frac{B}{m}\}$, and the payment of $s_i^j \in S_W$ is

$$p_i^j = \frac{b_{h_i+1}^j}{v_{h_i+1}^j} \cdot v_i^j. \tag{30}$$

5.1.2. Performance of mechanism BPSG-H

Theorem 11. *Mechanism BPSG-H guarantees individual rationality and budget-feasibility.*

Proof. It is easy to show individual rationality. We next focus on budget-feasibility. Denote by S_W^j the selected agents in group $G_j, \forall 1 \le j \le m$, *i.e.*, $S_W^j = \{s_1^j, \ldots, s_{k_j}^j\}$. Assume that agent $s_i^j (i \le k_j)$ achieves the maximum payment-per-value among agents in S_W^j, *i.e.*, $\frac{p_i^j}{v_i^j} = \max_{1\le h\le k_j} \frac{p_h^j}{v_h^j}$. According to Eq. (29), we have $P_{r_{i,h_i}}^H \le \frac{B}{m}$ which implies $\sum_{1\le h\le k_j} p_h^j = \sum_{1\le h\le k_j} v_h^j \cdot \frac{p_h^j}{v_h^j} \le (\mathcal{V}_{-i,h_i}^j + v_i^j) \cdot \frac{p_i^j}{v_i^j} \le \frac{B}{m}$. Thus, the total payment of selected agents in group G_j will not exceed $\frac{B}{m}$, and the total payment of m groups will not exceed B, that is, budget-feasibility holds. □

Algorithm 4: Mechanism **BPSG-H**(B, b, S, G).

Input: B, b, S, G.
Output: P, S_w

$P \leftarrow 0, S_w \leftarrow \emptyset$;
Sort agents in $G_j (\forall 1 \le j \le m)$, in weakly increasing order of their bid-per-values and generate the virtual ratio set $\mathcal{R}_h$ with value sorted and indexed by γ_l's;
// **Determine the final base selection ratio**;
for $1 \le l \le |\mathcal{R}|$ **do**
  Compute $I_j(\gamma_l) = \arg\min_{1 \le i \le n_j} \{\sum_{1 \le h \le i} v_h \ge \gamma_l \cdot \mathcal{V}_j\}$ for any $1 \le j \le m$;
  if $I_j(\gamma_l) < n_j, \forall 1 \le j \le m$ **then**
    Compute the payment $P^H_{\gamma_l}$ according to (29);
    if $P^H_{\gamma_l} \le \frac{B}{m}$ **then**
      $l \leftarrow l + 1$;
    else
      break;
    end
  else
    break;
  end
end
$r_f \leftarrow \gamma_{l-1}$;
// **Agent selection and payment scheme**;
Add agent $s^j_i (\forall 1 \le j \le m)$ with $i \le k_j = I_j(r_f)$ into the selected agent set S_w;
Decide the payments to agents according to (30);
return P, S_w

Theorem 12. *Mechanism BPSG-H guarantees truthfulness.*

Proof. It is easy to prove that each selected agent would still be selected if she bids a lower cost. Moreover, the payment for each selected agent is the threshold payment according to the payment determination in Mechanism BPSG-H. Thus, Mechanism BPSG-H guarantees truthfulness by the characterization of truthfulness in [59]. □

Assume that γ_l is the final base selection ratio. Recall that $r_j(\gamma_l)$ is the selection ratio of group G_j after selecting the first k_j agents, *i.e.*, $r_j(\gamma_l) = \frac{\sum_{i \le k_j} v^j_i}{\mathcal{V}_j}$, and r_{max} and r_{min} are the maximum and minimum selection ratios among groups, *i.e.*, $r_{max} = \max_{1 \le j \le m}\{r_j(\gamma_l)\}$ and $r_{min} = \min_{1 \le j \le m}\{r_j(\gamma_l)\}$. Specifically, we have $r_{min} = \gamma_l = r_f$ since there must exist at least one group G_j whose selection ratio is $r_j(\gamma_l) = \gamma_l$ due to the generation of $\mathcal{R}_h$ in (28). Let ALG_H and OPT_H denote the minimum selection ratio of Mechanism BPSG-H and the optimal solution, respectively.[9]

Theorem 13. *(1) Mechanism BPSG-H achieves $(m+1)(1+\sigma\eta)$-approximation ratio where $\sigma = \frac{v_{max}}{v_{min}}$, $\eta = \frac{\mathcal{V}_{max}}{\mathcal{V}_{min}}$ when BPSG-H can select at least one agent from each group, i.e., $ALG_H \ge \frac{v_{min}}{\mathcal{V}_{max}}$.*

(2) No budget-feasible proportion-representative mechanism $\mathcal{M}_H$, that guarantees truthfulness and individual rationality, can achieve $ALG_{\mathcal{M}_H} \ge m \cdot \frac{v_{max}}{\mathcal{V}_{min}} + \theta$ for any $\theta > 0$ where $ALG_{\mathcal{M}_H}$ is the solution of $\mathcal{M}_H$ if Mechanism BPSG cannot select at least one agent from each group, i.e., $ALG_H = 0$.

Proof. **(1)** As shown in Fig. 10, we sort agents in weakly increasing order of their bids in each group, *e.g.*, $\frac{b^{j_1}_1}{v^{j_1}_1} \le \frac{b^{j_1}_2}{v^{j_1}_2} \le \cdots \le \frac{b^{j_1}_{n_{j_1}}}{v^{j_1}_{n_{j_1}}}$ in group G_{j_1}. The length of each rectangle s^j_i in group G_j is the marginal gain of the selection ratio after adding s^j_i into the selected agent set when the first $i-1$ agents have been selected. For example, the length of the rectangle representing $s^{j_1}_1$ is $\frac{v^{j_1}_1}{\mathcal{V}_{j_1}}$. Mechanism BPSG-H selects the first k_j agents from group G_j (green rectangles).

Suppose that the final base selection ratio is the l-th element in virtual ratio set $\mathcal{R}_h$, and we have $r_f = \gamma_l = ALG_H$ and $k_j = I_j(\gamma_l)$ is the number of selected agents in group G_j. We have $n_j \ge 2, \forall 1 \le j \le m$ since BFSG-H can select at least one agent from each group. Depending on the conditions at the termination of BPSG-H, we consider the following two cases:

Case i) Suppose that there exists group G_j from which we select all agents, *i.e.*, $\exists j \le m, I_j(\gamma_{l+1}) = n_j$, when trying the next virtual ratio γ_{l+1}. We have $\gamma_l \cdot \mathcal{V}_j = \sum_{i \le n_j - 1} v^j_i$ and $OPT_H \le 1$, which implies $\frac{OPT_H}{ALG_H} \le \frac{1}{\gamma_l} \le 1 + \frac{v_{max}}{v_{min}}$.

[9] Generally speaking, the framework of the approximation poofs in the heterogeneous setting are similar to that of homogeneous.

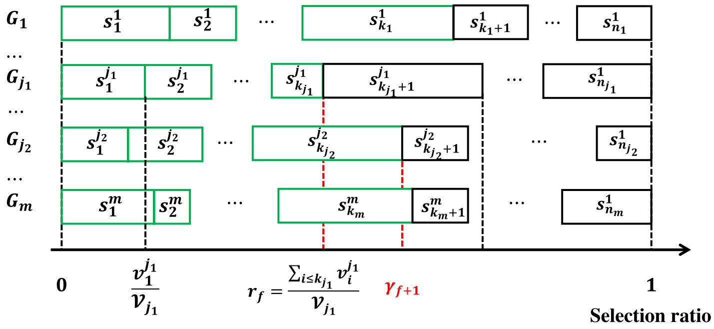


Fig. 10. An example for the proof of Theorem 13.

Case ii) Next, we focus on the case that not all agents are selected from group G_j when trying γ_{l+1}, *i.e.*, $I_j(\gamma_{l+1}) < n_j, \forall 1 \le j \le m$. Let G' denote the set of all groups that can choose the first k_j agents making the selection ratio in group G_j equal to r_f, *i.e.*, $r_f \cdot \mathcal{V}_j = \sum_{i\le k_j} v_i^j$, $\forall G_j \in G'$. As shown in Fig. 10, G_{j_1} is one of the groups in G'. Thus, we have $r_j(r_f) > r_f, \forall G_j \notin G'$. In addition, the next virtual ratio γ_{l+1} must be,

$$\gamma_{l+1} = \min\left\{\min_{G_j\notin G'}\{r_j(r_f)\}, \min_{G_j\in G'}\left\{\frac{\sum_{i\le k_j+1} v_i^j}{\mathcal{V}_j}\right\}\right\}.$$

According to the agent selection scheme of Mechanism BPSG-H, we have

$$P_{\gamma_l}^H = \sum_{1\le j\le m}\left(\frac{b_{k_j+1}^j}{v_{k_j+1}^j}\cdot\sum_{1\le i\le k_j} v_i^j\right) \le \frac{B}{m} \tag{31}$$

Depending on whether the next ratio γ_{l+1} is generated from a group in G' or not, we consider the following two subcases:

Subcase 1: Assume that the virtual ratio γ_{l+1} is generated from group $G_{j_2} \notin G'$ where $j_1 \ne j_2$ as shown in Fig. 4, *i.e.*,

$$\gamma_{l+1} = r_{j_2}(r_f) = \min\left\{\min_{G_j\notin G'}\{r_j(r_f)\}, \min_{G_j\in G'}\left\{\frac{\sum_{i\le k_j+1} v_i^j}{\mathcal{V}_j}\right\}\right\}. \tag{32}$$

Thus, we have $r_j(r_f) \ge r_{j_2}(r_f) = \gamma_{l+1}, \forall G_j \notin G'$, and the number of selected agents with ratio γ_{l+1} in group G_j where $G_j \notin G'$ should be k_j. For the group G_j in G', the number of selected agents will be $k_j + 1$. Since γ_l is the final base selection ratio, with ratio γ_{l+1}, we have

$$P_{\gamma_{l+1}}^H = \sum_{G_j\in G'}\sum_{i\le k_j+1} v_i^j\cdot\frac{b_{k_j+2}^j}{v_{k_j+2}^j} + \sum_{G_j\notin G'}\sum_{i\le k_j} v_i^j\cdot\frac{b_{k_j+1}^j}{v_{k_j+1}^j} > \frac{B}{m}. \tag{33}$$

We divide agents into two parts: $\tilde{S} = \{s_i^j | i \le k_j, \forall G_j \notin G'\} \cup \{s_i^j | i \le k_j + 1, \forall G_j \in G'\}$ and $S \setminus \tilde{S}$ for further analysis.

For set $\tilde{S}$: The optimal solution can select all agents in $\tilde{S}$ with cost zero in the best case and spend the budget on the remaining agents in $S \setminus \tilde{S}$. Since OPT_H selects k_j agents from each group G_j not in G', and $k_j + 1$ agents from each group G_j in G', we have

$$\max_{G_j\notin G'}\frac{\sum_{i\le k_j} v_i^j}{\mathcal{V}_j} = \max_{G_j\notin G'} r_j(r_f) \le r_{max}$$

and

$$\max_{G_j\in G'}\frac{\sum_{i\le k_j+1} v_i^j}{\mathcal{V}_j} \le r_f + \max_{G_j\in G'}\frac{v_{k_j+1}^j}{\mathcal{V}_j} \le r_{min} + \frac{v_{max}}{\mathcal{V}_{min}}.$$

Thus, after choosing all agents in $\tilde{S}$, the maximum selection ratio among all groups is $\max\{r_{max}, r_{min} + \frac{v_{max}}{\mathcal{V}_{min}}\} \le r_{min} + \frac{v_{max}}{\mathcal{V}_{min}}$, and the minimum selection ratio is γ_{l+1} due to (32). Thus, the optimal solution OPT_H can achieve selection ratio γ_{l+1} by choosing all agents in $\tilde{S}$.

For set $S \setminus \tilde{S}$: For group G_j where $G_j \notin G'$, the bid-per-values of agents after $s^j_{k_j}$ are at least $\frac{b^j_{k_j+1}}{v^j_{k_j+1}}$, while the bid-per-values of agents after $s^j_{k_j+1}$ are at least $\frac{b^j_{k_j+2}}{v^j_{k_j+2}}$ in group $G_j \in G'$. With budget B, the optimal solution can select at most $m\sum_{i\le k_j} v^j_i$ value from the remaining agents after s_{k_j} with cost-per-value $\frac{b^j_{k_j+1}}{v_{k_j+1}}$ from group G_j for any $G_j \notin G'$, and $m\sum_{i\le k_j+1} v^j_i$ value with cost-per-value $\frac{b^j_{k_j+2}}{v_{k_j+2}}$ from the group $G_j \in G'$ due to Eq. (33). According to the analysis for the set $\tilde{S}$, the minimum selection ratio among groups by selecting these numbers of agents is $m\gamma_{l+1}$ and the maximum selection ratio is $m(r_{min} + \frac{v_{max}}{\mathcal{V}_{min}})$.

Thus, by combining set $\tilde{S}$ and $S \setminus \tilde{S}$, we have $OPT_H \le r_{min} + \frac{v_{max}}{\mathcal{V}_{min}} + m(r_{min} + \frac{v_{max}}{\mathcal{V}_{min}}) \le (m+1)(r_f + \frac{v_{max}}{\mathcal{V}_{min}})$.

Subcase 2: Assume that the virtual ratio γ_{l+1} is generated from group $G_{j_2} \in G'$, *i.e.*,

$$\gamma_{l+1} = \frac{\sum_{i\le k_{j_2}+1} v^{j_2}_i}{\mathcal{V}_{j_2}} = \min\left\{\min_{G_j\notin G'}\{r_j(r_f)\}, \min_{G_j\in G'}\left\{\frac{\sum_{i\le k_j+1} v^j_i}{\mathcal{V}_j}\right\}\right\}.$$

Using similar arguments to the case above, we still have $OPT_H \le (m+1)(r_f + \frac{v_{max}}{\mathcal{V}_{min}})$.

Therefore, we have $OPT_H \le (m+1)(ALG + \frac{v_{max}}{\mathcal{V}_{min}})$. Specifically, if Mechanism BPSG can select at least one agent from each group, *i.e.*, $ALG_H \ge \frac{v_{min}}{\mathcal{V}_{max}}$, we have $\frac{OPT_H}{ALG_H} \le (m+1)(1+\sigma\eta)$ which means Mechanism BPSG achieves $(m+1)(1+\sigma\eta)$-approximation ratio.

(2) Assume that there are m groups, and we sort agents in weakly increasing order of their bids in each group. Suppose that the budget of the planner is B. According to Mechanism BPSG-H, there are two cases as follows due to $ALG_H = 0$:

i) When choosing the first agent from each group, we have $\sum_{1\le j\le m} v^j_1 \cdot \frac{b^j_2}{v^j_2} > \frac{B}{m}$ which implies that we cannot select at least one agent from each group. Moreover, the minimum selection ratio among groups is at most $\frac{v_{max}}{\mathcal{V}_{min}}$ when selecting the first agent in each group. Assume for purpose of contradiction that there exists a budget-feasible proportion-representative mechanism $\mathcal{M}$ with better performance $ALG_{\mathcal{M}_H} \ge m \cdot \frac{v_{max}}{\mathcal{V}_{min}} + \theta$. Myerson's characterization [59] implies that the threshold payment for the first agent in group G_j should be at least $\frac{v^j_1 \cdot b^j_2}{v^j_2}$, and thus the total payment now is $\sum_{1\le j\le m} \frac{v^j_1 b^j_2}{v^j_2} > \frac{B}{m}$. By truthfulness and budget-feasibility, $\mathcal{M}$ cannot select at most $m \cdot v^j_1$ value from group G_j implying that $\mathcal{M}$ can achieve at most $m \cdot \frac{v_{max}}{\mathcal{V}_{min}}$ which contradicts $ALG_{\mathcal{M}_H} \ge m \cdot \frac{v_{max}}{\mathcal{V}_{min}} + \theta$.

ii) There exists a group G_j which contains only one agent, and we will select this agent from G_j when trying the first virtual ratio which implies that we cannot select at least one agent from each group. Assume that we have $ALG_{\mathcal{M}} \ge \theta$. In this case, Mechanism $\mathcal{M}_H$ must always select at least one agent from each group. Thus, the payment for the agent in G_j can be infinity by Myerson's characterization [59], which violates budget-feasibility. Therefore, $\mathcal{M}_H$ also cannot select at least one agent from each group using that payment which contradicts $ALG_{\mathcal{M}_H} \ge \theta$. □

Theorem 14. *No budget-feasible proportion-representative mechanism obtains an approximation ratio better than* $\Omega(\sigma\eta)$ *where* $\sigma = \frac{v_{max}}{v_{min}}$ *and* $\eta = \frac{\mathcal{V}_{max}}{\mathcal{V}_{min}}$.

Proof. Suppose that we have m groups. Group G_1 has n_1 agents with cost-per-values $\{0, \frac{B-\epsilon}{v_{max}}, \ldots, \frac{B-\epsilon}{v_{max}}\}$ and values $\{v_{max}, \cdots, v_{max}\}$. While group G_m has n_m agents, each with cost-per-value $\frac{\epsilon}{v_{min}}$ and values $\{v_{min}, \cdots, v_{min}\}$. Each remaining group $G_j(\forall j, 2 \le j \le m-1)$ has n_j agents with cost zero. Specifically, assume that $\mathcal{V}_1 \le \mathcal{V}_2 \le \cdots \le \mathcal{V}_m$ and $B = v_m\epsilon > 2\epsilon$. Thus, we have $\mathcal{V}_{min} = \mathcal{V}_1, \mathcal{V}_{max} = \mathcal{V}_m$. By Myerson's characterization [59], the payment in G_1 is at least $B-\epsilon$ for the first agent and is at least ϵ for the first agent in G_m to achieve truthfulness. Thus, by truthfulness and budget-feasibility, the solution of any budget-feasible proportion-representative mechanism is no better than $\frac{v_{min}}{\mathcal{V}_{max}}$ since they can only select at most the first agent from G_1 and G_m. However, the optimal solution is $\frac{v_{max}}{\mathcal{V}_{min}}$. Thus, no budget-feasible proportion-representative mechanism can achieve an approximation ratio better than $\frac{v_{max}}{v_{min}} \cdot \frac{\mathcal{V}_{max}}{\mathcal{V}_{min}}$. This completes the proof. □

5.2. Mechanisms for multiple group models

Recall that in the multiple group model, each agent s_i might belong to multiple groups, *i.e.*, $1 \le |\mathcal{G}(s_i)| \le m$. We first consider the single counting case where the selection ratio in one of the groups $\mathcal{G}(s_i)$, say G_j, would increase by $\frac{v_i}{\mathcal{V}_j}$ if s_i is selected and contributes to G_j, while in the multiple counting case, the selection ratios of all groups in $\mathcal{G}(s_i)$ would increase by $\frac{v_i}{\mathcal{V}_j}$ for any $G_j \in \mathcal{G}(s_i)$, once s_i is selected.

Algorithm 5: Mechanism **BPMG-SH**(B, b, S, G).

Input: B, b, S, G.
Output: P, S_W
$P \leftarrow 0, S_W \leftarrow \emptyset$;
$\hat{B} \leftarrow \frac{v_{min}}{v_{max}+v_{min}} B$;
Sort all agents in the weakly increasing order of their bid-per-values, *i.e.*, $\frac{b_1}{v_1} \leq \frac{b_2}{v_2} \leq \cdots \leq \frac{b_n}{v_n}$;
// **Determine the candidate agent set**;
$k \leftarrow 0$;
for $1 \leq i \leq n$ **do**
  Compute $r^H(s_i)$ according to (34);
  $F(s_i) \leftarrow \sum_{1 \leq j \leq m} r^H(s_i) \cdot \mathcal{V}_j$;
  if $\frac{b_i}{v_i} \cdot F(s_i) \leq \hat{B}$ **then**
    $k \leftarrow k+1$;
  else
    break;
  end
end
$k \leftarrow i-1$;
// **Agent selection and payment scheme**;
Let $X^H(s_h)$ denote the solution of $ILP^H(s_h)$;
The payment for $s_i \in S_W$ is $p_i = v_i \min\{\frac{\hat{B}}{F(s_k)}, \frac{b_{k+1}}{v_{k+1}}\}$;

5.2.1. Single counting case

In this section, we introduce a **B**udget-feasible **P**roportion-representative mechanism for the **M**ultiple **G**roup setting in **S**ingle counting case with heterogeneous agents, called BPMG-SH.

Intuitively, we modify Mechanism BPMG-S to address heterogeneous agents. Given different values of agents, we use $\hat{B} = \frac{v_{min}}{v_{max}+v_{min}} B$ when selecting agents for budget-feasibility. In detail, we first sort all agents in weakly increasing order of their bid-per-values $\frac{b_1}{v_1} \leq \frac{b_2}{v_2} \leq \cdots \leq \frac{b_n}{v_n}$. Recall that $\mathcal{S}(s_h) = \{s_i | i \leq h\}$ is the set containing agents before agent b_{h+1}. We modify the integer program formulation (20) to compute a matching result under a given agent set $\mathcal{S}(s_h)$, that maximizes the minimum selection ratio among groups while ignoring the costs and budget constraint, denoted by $ILP^H(s_h)$ as follows,

$$\begin{aligned} & \max \min_{1 \leq j \leq m} \frac{\sum_{i \leq h} x_{ij} v_i}{\mathcal{V}_j} \\ s.t., & \sum_{1 \leq j \leq m} x_{ij} \leq 1, \forall 1 \leq i \leq h \\ & x_{ij} \in \{0, 1\}, \forall 1 \leq i \leq h, G_j \in \mathcal{G}(S_i) \\ & x_{ij} = 0, \forall 1 \leq i \leq h, G_j \notin \mathcal{G}(S_i) \end{aligned} \tag{34}$$

where the three conditions indicate that agents in $\mathcal{S}(s_h)$ can only be matched to at most one of the groups they belong to. Similar to earlier sections, optimal solution for (34) can be computed in polynomial time by constructing Min-cost Max-Flow networks. To guarantee truthfulness, we need to ensure the algorithm for solving Min-cost Max-Flow returns the same solution under the same agent set $\mathcal{S}(s_h)$. To do this, different from the homogeneous agent setting, we will set a unique index for each node of the flow network. Note that the network is the same when the input agent set does not change. After the construction of the networks, we then run the Ford-Fulkerson algorithm [61] and use Bellman-Ford algorithm to consider the path with the smallest lexicographical order of nodes' indexes when searching the augmenting path, which can always output the same solution under the fixed flow network.

Let $X^H(s_h)$ denote the solution of $ILP^H(s_h)$, and $r^H(s_h)$ be the minimum selection ratio among groups under $X^H(s_h)$. In the heterogeneous agent scenario, $F(s_h)$ denotes the sum of values that should be selected under $r^H(s_h)$ in each group, *i.e.*,

$$F(s_h) = \sum_{1 \leq j \leq m} r^H(s_h) \cdot \mathcal{V}_j. \tag{35}$$

It is obvious that $F(s_h) \leq \sum_{s_i \in \mathcal{S}(s_h)} v_i$.

Agent selection and payment scheme Now we are ready to select agents. We iteratively test each agent starting from the first agent s_1. Suppose that we are now trying agent s_i and calculating the value of $F(s_i)$ by Eq. (35). If $\frac{b_i}{v_i} \cdot F(s_i) \leq \hat{B}$, we consider the next agent s_{i+1}, otherwise, we select agents from the previous $i-1$ agents $\mathcal{S}(s_{i-1})$. Assume that s_k is the last agent who satisfies $\frac{b_k}{v_k} \cdot F(s_k) \leq \hat{B}$ which implies $\frac{b_{k+1}}{v_{k+1}} \cdot F(s_{k+1}) > \hat{B}$. After determining value k, we decide winners and corresponding payments. If agent $s_i (i \leq k)$ is matched to one of the groups in $X^H(s_k)$, *i.e.*, $\sum_{1 \leq j \leq m} x_{ij} = 1$, then she is selected, *i.e.*, $s_i \in S_W$. The payment for each selected agent is $p_i = v_i \cdot \min\{\frac{\hat{B}}{F(s_k)}, \frac{b_{k+1}}{v_{k+1}}\}$ while the payments for unselected agents are zero. Specifically, the payment for each selected agent is $\frac{\hat{B}}{F(s_k)}$ if $k = n$.

Next, we analyze the performance of Mechanism BPMG-SH.

Theorem 15. *Mechanism BPMG-SH guarantees individual rationality and budget-feasibility.*

Proof. It is easy to show individual rationality. Next, we consider budget-feasibility. Recall that $F(s_k) = \sum_{1 \le j \le m} r^H(s_k) \cdot \mathcal{V}_j$. In group G_j, the total value of the selected agents must satisfy $\sum_{1 \le i \le n} x_{ij} v_i \le r^H(s_k) \cdot \mathcal{V}_j + v_{max}$ due to (34). Thus, we have

$$\begin{aligned}\sum_{s_i \in S_W} p_i &= \sum_{1 \le j \le m} \sum_{1 \le i \le n} x_{ij} v_i \cdot \min\{\frac{\hat{B}}{F(s_k)}, \frac{b_{k+1}}{v_{k+1}}\} \\ &\le \min\{\frac{\hat{B}}{F(s_k)}, \frac{b_{k+1}}{v_{k+1}}\} \sum_{1 \le j \le m} (r^H(s_k) \cdot \mathcal{V}_j + v_{max}) \\ &\le \min\{\frac{\hat{B}}{F(s_k)}, \frac{b_{k+1}}{v_{k+1}}\} (F(s_k) + m v_{max}) \le B\end{aligned}$$

where the last inequality is because Mechanism BPMG-SH will select at least one agent from each group under $r^H(s_k)$, *i.e.*, $F(s_h) \ge m \cdot v_{min}$. □

Theorem 16. *Mechanism BPMG-SH guarantees truthfulness.*

Proof. According to the definition of $F(s_{k+1})$, it is obvious that $F(s_{k+1}) \ge F(s_k)$. Since s_k is the last agent among candidate agents, we have,

$$\frac{b_k}{v_k} \cdot F(s_k) \le \hat{B}, \frac{b_{k+1}}{v_{k+1}} \cdot F(s_{k+1}) > \hat{B} \tag{36}$$

The payment for agent $s_i \in S_W$ is $p_i = v_i \cdot \min\{\frac{\hat{B}}{F(s_k)}, \frac{b_{k+1}}{v_{k+1}}\}$. Denote by $O = \langle s_1, s_2, ..., s_k, ..., s_n \rangle$ the order of agents in weakly increasing order of their bids.

Monotonicity: For any agent s_i with $i \le k$ and $s_i \in S_W$, s_i is selected when bidding real cost c_i. If s_i decreases her bid to $b'_i \le c_i$, agents before s_{k+1} will not change and s_i will still be selected. Therefore, Mechanism BPMG-S is monotonic.

Threshold payments: For any agent s_i with $i \le k$ and $s_i \in S_W$, her payment is $p_i = v_i \cdot \min\{\frac{\hat{B}}{F(s_k)}, \frac{b_{k+1}}{v_{k+1}}\}$. According to the relationship between $\frac{\hat{B}}{F(s_k)}$ and $\frac{b_{k+1}}{v_{k+1}}$, we further consider two subcases:

- $\frac{\hat{B}}{F(s_k)} \le \frac{b_{k+1}}{v_{k+1}}$: For agent s_i with $i \le k$, we have $p_i = \frac{v_i \cdot \hat{B}}{F(s_k)}$ and $\frac{b_i}{v_i} \le \frac{\hat{B}}{F(s_k)}$. If she bids a higher cost $\frac{b'_i}{v_i} > \frac{\hat{B}}{F(s_k)}$ (let $s_{i'}$ denote the new position after bidding a false cost), she will be after agent s_k and the new order now is

  $$O' = \langle s_1, ..., s_{i-1}, s_{i+1}, ..., s_k, ..., s_{i'}, ..., s_n \rangle$$

  Let $\mathcal{S}(s_h)'$ denote agents before s_h (including herself) in the new sequence O', and $F(s_h)'$ denote the new value with ratio computed by (34) with input $\mathcal{S}(s_h)'$. We have $\mathcal{S}(s_k) \subseteq \mathcal{S}(s_{i'})'$ which means $F(s_{i'})' \ge F(s_k)$. Thus, we must have $\frac{b'_i}{v_i} \cdot F(s_{i'})' > \frac{\hat{B}}{F(s_k)} \cdot F(s_k) = \hat{B}$ which implies that agent s_i will not be selected, and her utility is zero. Therefore, the value $v_i \cdot \frac{B}{F(s_k)}$ is the critical value for the selected agent $s_i \in S_W$ and the selected agents are paid threshold payments.
- $\frac{\hat{B}}{F(s_k)} > \frac{b_{k+1}}{v_{k+1}}$: For agent s_i with $i \le k$, we have $p_i = v_i \cdot \frac{b_{k+1}}{v_{k+1}}$. If she bids a higher cost $\frac{b'_i}{v_i} > \frac{b_{k+1}}{v_{k+1}}$, it is easy to find that $\mathcal{S}(s_{k+1}) \subseteq \mathcal{S}(s_{i'})'$ and $F(s_{i'})' > F(s_{k+1})$. Thus, we have $\frac{b'_i}{v_i} \cdot F(s_{i'})' > \frac{b_{k+1}}{v_{k+1}} \cdot F(s_{k+1}) > \hat{B}$ which implies that agent s_i will not be selected. Therefore, the value $v_i \cdot \frac{b_{k+1}}{v_{k+1}}$ for the selected agent $s_i \in S_W$ is the critical value and the selected agents are paid threshold payments.

Therefore, Mechanism BPMG-S guarantees truthfulness. □

Let ALG^s_H and OPT^s_H denote the minimum selection ratio of Mechanism BPMG-SH and the optimal solution, respectively.

Theorem 17. *(1) Mechanism BPMG-SH achieves $[m\eta(1+\sigma)(1+\sigma\eta)+1]$-approximation ratio if BPMG-SH can select at least one agent from each group (i.e., $ALG^s_H \ge \frac{v_{min}}{\mathcal{V}_{max}}$) where $\sigma = \frac{v_{max}}{v_{min}}, \eta = \frac{\mathcal{V}_{max}}{\mathcal{V}_{min}}$.*

(2) No budget-feasible proportion-representative mechanism $\mathcal{M}'_H$, that guarantees truthfulness and individual rationality, can obtain $ALG_{\mathcal{M}'_H} \geq \theta$ for any $\theta > (1+\frac{v_{max}}{v_{min}})\frac{v_{max}}{\mathcal{V}_{min}}$ where $ALG_{\mathcal{M}'_H}$ is the solution of $\mathcal{M}'_H$ if Mechanism BPMG-S cannot select at least one agent from each group, i.e., $ALG^s_H = 0$.

Proof. **(1)** The bid-per-values of agents before s_{k+1} are at most $\frac{b_{k+1}}{v_{k+1}}$ while the bid-per-values of agents after s_{k+1} are at least $\frac{b_{k+1}}{v_{k+1}}$. According to Mechanism BPMG-SH, we have $ALG^s_H = r^H(s_k)$. For the agents before s_{k+1}, the optimal solution OPT^s_H can select all these k agents with zero costs in the best case. Then, the minimum selection ratio among these agents is still $r^H(s_k)$.

For the agents after s_k, OPT^s_H can select at most $F(s_{k+1}) \cdot \frac{v_{min}+v_{max}}{v_{min}}$ value due to (36) with budget B. For the set $\mathcal{S}(s_k)$, the maximum increase in the minimum selection ratio among all groups is $\max_{1\leq j\leq m} \frac{v_{k+1}}{\mathcal{V}_j} \leq \frac{v_{max}}{\mathcal{V}_{min}}$ after adding agent s_{k+1} into $\mathcal{S}(s_k)$, *i.e.*, $r(s_{k+1}) \leq r(s_k) + \frac{v_{max}}{\mathcal{V}_{min}}$. Thus, we have

$$F(s_{k+1}) = \sum_{1\leq j\leq m} r(s_{k+1}) \cdot \mathcal{V}_j \leq \sum_{1\leq j\leq m} \left(r(s_k)\mathcal{V}_j + \frac{v_{max}}{\mathcal{V}_{min}}\mathcal{V}_j \right).$$

In the best case, the maximum increase in the minimum selection ratio among all groups is $\frac{F(s_{k+1})\frac{v_{min}+v_{max}}{v_{min}}}{\mathcal{V}_{min}}$ by adding $\frac{v_{min}+v_{max}}{v_{min}}F(s_{k+1})$ value. Thus, we have

$$\begin{aligned} \frac{F(s_{k+1})\frac{v_{min}+v_{max}}{v_{min}}}{\mathcal{V}_{min}} &\leq \frac{v_{min}+v_{max}}{v_{min}} \sum_{1\leq j\leq m} \left(r(s_k)\mathcal{V}_j + \frac{v_{max}}{\mathcal{V}_{min}}\mathcal{V}_j \right) / \mathcal{V}_{min} \\ &\leq m\frac{v_{min}+v_{max}}{v_{min}} \cdot \left(r(s_k) \cdot \frac{\mathcal{V}_{max}}{\mathcal{V}_{min}} + \frac{v_{max}\mathcal{V}_{max}}{\mathcal{V}^2_{min}} \right). \end{aligned}$$

Thus, we have $OPT^s_H \leq r(s_k) + \frac{F(s_{k+1})}{\mathcal{V}_{min}} \leq (m \cdot (1+\frac{v_{max}}{v_{min}})\frac{\mathcal{V}_{max}}{\mathcal{V}_{min}} + 1) \cdot ALG^s_H + m \cdot (1+\frac{v_{max}}{v_{min}})\frac{v_{max}\mathcal{V}_{max}}{\mathcal{V}^2_{min}}$. Additionally, if Mechanism BPMG-SH can select at least one agent from each group, *i.e.*, $ALG^s_H \geq \frac{v_{min}}{\mathcal{V}_{max}}$, we have $\frac{OPT^s_H}{ALG^s_H} \leq (m \cdot (1+\frac{v_{max}}{v_{min}})\frac{\mathcal{V}_{max}}{\mathcal{V}_{min}} + 1) + m \cdot (1+\frac{v_{max}}{v_{min}})\frac{v_{max}\mathcal{V}^2_{max}}{v_{min}\mathcal{V}^2_{min}} \leq m\eta(1+\sigma)(1+\sigma\eta)+1$.

(2) Since s_k is the last agent among candidates, we have $F(s_k) = 0$ when $ALG^s_H = 0$ and consider the following two cases:

- $F(s_k) = 0 (k = n)$: There is no feasible ratio in (34) with input $\mathcal{S}(s_n)$, *i.e.*, we cannot find a matching that allocates non-zero value to each group. Thus, no budget-feasible proportion-representative mechanism $\mathcal{M}'_H$ can achieve $ALG_{\mathcal{M}'_H} \geq \theta$ for any $\theta > 0$.
- $F(s_k) = 0 (k < n)$: We have $r^H(s_k) = 0$ and $\frac{b_k}{v_k} \cdot F(s_k) \leq \hat{B}$, $\frac{b_{k+1}}{v_{k+1}} \cdot F(s_{k+1}) > \hat{B}$. For the set $\mathcal{S}(s_k)$, the maximum increase in the minimum selection ratio among all groups is $\frac{v_{max}}{\mathcal{V}_{min}}$, *i.e.*, $r^H(s_{k+1}) \leq \frac{v_{max}}{\mathcal{V}_{min}}$, after adding agent s_{k+1}. For the agent s_i before s_{k+1}, Myerson's characterization [59] implies that her threshold payment should be at least $v_i \cdot \frac{b_{k+1}}{v_{k+1}}$. For the agents after s_k, we have $\frac{b_i}{v_i} \geq \frac{b_{k+1}}{v_{k+1}}, \forall i > k$. Thus, by truthfulness and budget-feasibility, mechanism $\mathcal{M}$ can select at most $(1+\frac{v_{max}}{v_{min}})F(s_{k+1})$ value with budget B which implies $ALG_{\mathcal{M}} \leq r^H(s_{k+1}) \leq (1+\frac{v_{max}}{v_{min}})\frac{v_{max}}{\mathcal{V}_{min}}$.

This finishes the proof. □

5.2.2. Multiple counting case

We can easily modify the Mechanism BPMG-SH to address the multiple counting case. When computing a matching result under a given agent set $\mathcal{S}(s_h)$, we modify the integer program formulation (34) as follows,

$$\begin{aligned} &\max \min_{1\leq j\leq m} \frac{\sum_{i\leq h} x_i v_i}{\mathcal{V}_j} \\ &s.t.\, x_i \in \{0,1\}, \forall 1 \leq i \leq h, G_j \in \mathcal{G}(S_i) \\ &\quad x_i = 0, \forall 1 \leq i \leq h, G_j \notin \mathcal{G}(S_i) \end{aligned} \tag{37}$$

Let $X^{H,m}(s_h)$ denote the solution of the above programming, and $r^{H,m}(s_h)$ be the minimum selection ratio among groups under $X^{H,m}(s_h)$. Recall that $F(s_h)$ denotes the sum of values that should be selected under $r^{H,m}(s_h)$ in each group, *i.e.*,

$$F(s_h) = \sum_{1 \le j \le m} r^{H,m}(s_h) \cdot \mathcal{V}_j. \tag{38}$$

Agent selection and payment scheme Now we are ready to select agents. We find s_k such that $\frac{b_k}{v_k} \cdot F(s_k) \le \hat{B}$ and $\frac{b_{k+1}}{v_{k+1}} \cdot F(s_{k+1}) > \hat{B}$. After determining value k, we define agent set $\mathcal{S}(s_k) = \{s_i : i \le k\}$ as the candidate agent set and decide the final winners. If $x_i = 1$ in $X^{H,m}(s_h)$, then s_i is selected, *i.e.*, $s_i \in S_w$. The payment for each selected agent is $p_i = v_i \cdot \min\{\frac{\hat{B}}{F(s_k)}, \frac{b_{k+1}}{v_{k+1}}\}$ while the payments for unselected agents are zero. Specifically, the payment for each selected agent is $\frac{\hat{B}}{F(s_k)}$ if $k = n$.

The above mechanism can achieve the same approximation ratio as that in Section 5.2.1.

6. Extension to weighted proportional selection

In this section, we consider the weighted proportional selection setting where there is a weight φ_j assigned to group G_j for any $j \le m$ and the objective is to maximize the minimum weighted selection ratio among groups, *i.e.*, $\max \min_{j \le m} \frac{\varphi_j Q_j}{\sum_{s_i \in G_j} v_i}$. A higher φ_j results in the mechanism placing less emphasis on selecting agents from group G_j. Next, we modify the proposed mechanisms for the homogeneous setting to address the weighted proportional selection. A similar modification can be applied to the heterogeneous setting. Here, we only give mechanisms without approximation guarantees.

6.1. Single group model

We first modify Mechanism BPSG to address the weighted proportional selection for the homogeneous setting. The main difference is that the generated *virtual ratio set* $\mathcal{R}$ should consist of possible weighted selection ratios among all groups, *i.e.*,

$$\mathcal{R} = \cup_{0 \le i \le n_j, 1 \le j \le m} \left\{ \varphi_j \cdot \frac{i}{n_j} \right\}, \tag{39}$$

and sort all weighted ratios in the weakly increasing order of their values, where γ_l is the l-th element in $\mathcal{R}$, *i.e.*, $\gamma_1 < \gamma_2 < \cdots < \gamma_l < \cdots < \gamma_{|\mathcal{R}|}$.

To find the final base selection ratio, we iteratively consider weighted ratios in $\mathcal{R}$ starting with the first ratio γ_1. Suppose that we are now considering ratio γ_l for $l > 1$. Recall that $I_j(\gamma_l)$ denotes the minimum number of agents which ensures that the selection ratio in group G_j is at least γ_l, and we thus have $I_j(\gamma_l) = \lceil \frac{\gamma_l}{\varphi_j} \cdot n_j \rceil$, which should depend on the weight φ_j of group G_j. Then, similar to Mechanism BPSG, we find the final base selection ratio $r_f = \gamma_l$ which satisfies

$$P_{\gamma_l} = \sum_{1 \le j \le m} I_j(\gamma_l) \cdot b^j_{I_j(\gamma_l)+1} \le B \tag{40}$$

and

$$P_{\gamma_{l+1}} = \sum_{1 \le j \le m} I_j(\gamma_{l+1}) \cdot b^j_{I_j(\gamma_{l+1})+2} > B \tag{41}$$

Once deciding the final base selection ratio, we determine the final selected agents and corresponding payments. Let k_j denote the number of selected agents in group G_j, *i.e.*, $k_j = I_j(r_f) = Q_j$. In each group G_j, the first k_j agents are selected, *i.e.*, $s^j_i \in S_w$, $\forall 1 \le i \le k_j$, and we have $k_j < n_j$. Then we have

$$p^j_i = \begin{cases} b^j_{k_j+1}, & if\ s^j_i \in S_w \\ 0, & otherwise. \end{cases} \tag{42}$$

6.2. Multiple group model

We now modify Mechanism BPMG-S for the single counting case in multiple group model.[10] When using integer program formulation to compute a matching under a given agent set $\mathcal{S}(s_h)$, we should update the objective as follows,

[10] We can apply the similar modification for the multiple counting case.

$$\begin{aligned}
&\max \min_{1\le j\le m} \frac{\varphi_j \cdot \sum_{1\le i\le h} x_{ij}}{n_j}\\
s.t., &\sum_{1\le j\le m} x_{ij} \le 1, \forall 1 \le i \le h\\
&x_{ij} \in \{0,1\}, \forall 1 \le i \le h, G_j \in \mathcal{G}(S_i)\\
&x_{ij} = 0, \forall 1 \le i \le h, G_j \notin \mathcal{G}(S_i).
\end{aligned} \tag{43}$$

Note that the optimal solution of the above integer program can still be computed in polynomial time by constructing Max-Flow networks as follows. However, there are two difference from Mechanism BPMG-S: 1) Given an agent set $\mathcal{S}(s_h)$, all possible selection ratios for group G_j are in $\cup_{y\le|\mathcal{S}(s_h)\cap G_j|}\{\varphi_j \frac{y}{n_j}\}$; 2) When constructing a flow network based on such ratio r_s, each edge from source s to G_j is assigned with a capacity $\lceil \frac{r_s}{\varphi_j} \cdot n_j \rceil$, and if the maximum flow on such network equals to $\sum_{1\le j\le m} \lceil \frac{r_s}{\varphi_j} \cdot n_j \rceil$, then r_s is a feasible ratio.

Recall that $r(s_h)$ is the minimum selection ratio among groups in (43) and we can compute the total number of agents under ratio $r(s_h)$, *i.e.*,

$$F(s_h) = \sum_{1\le j\le m} \lceil \frac{r(s_h)}{\varphi_j} \cdot n_j \rceil. \tag{44}$$

Candidate agent selection Similar to Mechanism BPMG-S, we find s_k such that $b_k \cdot F(s_k) \le B$ and $b_{k+1} \cdot F(s_{k+1}) > B$. After determining the value k, we define agent set $\mathcal{S}(s_k) = \{s_i : i \le k\}$ as the *candidate agent set*.

Agent selection and payment scheme Now we select agents from the candidate agent set $\mathcal{S}(s_k)$. The main difference from Mechanism BPMG-S is the minimum weight matching between agents and groups that can minimize the total weight of matched agents satisfying that each group G_j has $\lceil \frac{r(s_k)}{\varphi_j} \cdot n_j \rceil$ matched agents from set $\mathcal{S}(s_k)$ as follows

$$\begin{aligned}
&\min \sum_{1\le i\le k} \sum_{1\le j\le m} w_i x_{ij}\\
s.t., &\sum_{1\le j\le m} x_{ij} \le 1, \forall i \le k\\
&\sum_{1\le i\le k} x_{ij} = \lceil \frac{r(s_k)}{\varphi_j} \cdot n_j \rceil, \forall 1 \le j \le m\\
&x_{ij} \in \{0,1\}, \forall i \le k, G_j \in \mathcal{G}(S_i)\\
&x_{ij} = 0, \forall i \le k, G_j \notin \mathcal{G}(S_i).
\end{aligned} \tag{45}$$

If agent $s_{i:i\le k}$ is matched to one of the groups, *i.e.*, $\sum_{1\le j\le m} x_{ij} = 1$, then she is selected, *i.e.*, $s_i \in S_w$. The payment for each selected agent is $p_i = \min\{\frac{B}{F(s_k)}, b_{k+1}\}$ while the payments for unselected agents are zero.

7. Conclusion

In this paper, we consider the proportion representation budget-feasible mechanism design problem where agents may have diverse group attributes and belong to different groups. We focus on designing budget-feasible mechanisms that can select appropriate proportions of agents from various groups satisfying individual rationality, budget-feasibility, and truthfulness under several settings. We start with homogeneous agent scenarios where agents have identical values. For the single group model, we propose Mechanism BPSG, which iteratively considers each virtual ratio generated by the distribution of agents across groups and finds the maximum one as the base selection ratio for all groups. For the multiple group model, we first consider the single counting case and propose Mechanism BPMG-S, which leverages the Max-Flow technique to measure the supply of agents under a fixed payment and identifies the payment which can maximize the minimum selection ratio among groups. We then extend BPMG-S to the multiple counting case to account for the situation in which agents can represent all of their respective groups. Finally, we consider heterogeneous agents where the agents can contribute different values under the single group and multiple group models. We carefully extend the designed mechanisms from the homogeneous agents to the heterogeneous setting. All the designed budget-feasible proportion-representative mechanisms can guarantee desirable properties like individual rationality, budget-feasibility, truthfulness, and approximation performance on proportional representation.

Future work on this problem may concentrate on the following directions: (i) It is natural to design a general mechanism or randomized mechanism for the heterogeneous setting that can simultaneously address the homogeneous setting efficiently. (ii) It is promising to consider different methods of computing the selection ratio for each group, e.g., considering decreasing marginal contributions of choosing additional agents. (iii) It is also potential to consider proportion-representative pricing mechanisms where the organizer posts prices to agents without requiring agents to bid their costs.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgements

This work was supported in part by the National Key Research and Development Program of China under grant No. 2019YFB2102200, Natural Science Foundation of China under Grant No. 62232004, 61972086, the Project 11771365 supported by NSFC, Key Laboratory of Computer Network and Information Integration (Southeast University), Ministry of Education, and the Research Grants Council of the Hong Kong Special Administrative Region, China, under Project UGC/FDS11/E03/21. Moreover, Chan was supported by the National Institute of General Medical Sciences of the National Institutes of Health [P20GM130461] and the Rural Drug Addiction Research Center at the University of Nebraska-Lincoln. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or the University of Nebraska.

References

[1] X. Liu, H. Chan, M. Li, W. Wu, Budget-feasible mechanisms for representing groups of agents proportionally, in: 30th International Joint Conference on Artificial Intelligence, IJCAI 2021, in: International Joint Conferences on Artificial Intelligence, 2021, pp. 313–320.
[2] M.A. Bargsted, O. Kedar, Coalition-targeted Duvergerian voting: how expectations affect voter choice under proportional representation, Am. J. Polit. Sci. 53 (2) (2009) 307–323.
[3] B. Flanigan, P. Gölz, A. Gupta, B. Hennig, A.D. Procaccia, Fair algorithms for selecting citizens' assemblies, Nature 596 (7873) (2021) 548–552.
[4] T.R. Marshall, Public opinion, representation, and the modern Supreme Court, Am. Polit. Q. 16 (3) (1988) 296–316.
[5] A.D. Procaccia, J.S. Rosenschein, A. Zohar, On the complexity of achieving proportional representation, Soc. Choice Welf. 30 (3) (2008) 353–362.
[6] H. Aziz, Z. Sun, School choice with flexible diversity goals and specialized seats, in: IJCAI, 2021, pp. 17–23.
[7] M.J. Taylor, J. McWilliam, H. Forsyth, S. Wade, Methodologies and website development: a survey of practice, Inf. Softw. Technol. 44 (6) (2002) 381–391.
[8] N.M. Bradburn, S. Sudman, B. Wansink, Asking Questions: the Definitive Guide to Questionnaire Design–for Market Research, Political Polls, and Social and Health Questionnaires, John Wiley & Sons, 2004.
[9] N. Jackson, M.S. Lewis-Beck, C. Tien, Pollster problems in the 2016 US presidential election: vote intention, vote prediction, Quad. Oss. Elet. 83 (1) (2020) 17–28.
[10] P. Whitla, Crowdsourcing and its application in marketing activities, Contemp. Manag. Res. 5 (1) (2009).
[11] N. Eagle, txteagle: mobile crowdsourcing, in: International Conference on Internationalization, Design and Global Development, Springer, 2009, pp. 447–456.
[12] L. Leemann, F. Wasserfallen, Extending the use and prediction precision of subnational public opinion estimation, Am. J. Polit. Sci. 61 (4) (2017) 1003–1022.
[13] N.J. Briscoe, J. Elith, R. Salguero-Gómez, J.J. Lahoz-Monfort, J.S. Camac, K.M. Giljohann, M.H. Holden, B.A. Hradsky, M.R. Kearney, S.M. McMahon, et al., Forecasting species range dynamics with process-explicit models: matching methods to applications, Ecol. Lett. 22 (11) (2019) 1940–1956.
[14] C. Harris, You're hired! An examination of crowdsourcing incentive models in human resource tasks, in: Proceedings of the Workshop on Crowdsourcing for Search and Data Mining (CSDM) at the Fourth ACM International Conference on Web Search and Data Mining (WSDM), Hong Kong, China, 2011, pp. 15–18.
[15] D. Yang, G. Xue, X. Fang, J. Tang, Crowdsourcing to smartphones: incentive mechanism design for mobile phone sensing, in: Proceedings of the 18th Annual International Conference on Mobile Computing and Networking, 2012, pp. 173–184.
[16] D. Zhao, X.-Y. Li, H. Ma, Budget-feasible online incentive mechanisms for crowdsourcing tasks truthfully, IEEE/ACM Trans. Netw. 24 (2) (2014) 647–661.
[17] N. Anari, G. Goel, A. Nikzad, Mechanism design for crowdsourcing: an optimal 1-1/e competitive budget-feasible mechanism for large markets, in: 2014 IEEE 55th Annual Symposium on Foundations of Computer Science, IEEE, 2014, pp. 266–275.
[18] J. Lang, P. Skowron, Multi-attribute proportional representation, Artif. Intell. 263 (2018) 74–106.
[19] Pay for committee chairs, https://www.theipsa.org.uk/publications/consultations/pay-for-committee-chairs.
[20] A. Knopf, Appropriations bill signed, funding many sud programs, Alcohol. Drug Abuse Week 30 (39) (2018) 5–6.
[21] E.L. Dey, Working with low survey response rates: the efficacy of weighting adjustments, Res. High. Educ. 38 (1997) 215–227.
[22] R.R. Rindfuss, M.K. Choe, N.O. Tsuya, L.L. Bumpass, E. Tamaki, Do low survey response rates bias results? Evidence from Japan, Demogr. Res. 32 (2015) 797–828.
[23] Y. Singer, Budget feasible mechanisms, in: 2010 IEEE 51st Annual Symposium on Foundations of Computer Science, IEEE, 2010, pp. 765–774.
[24] D.P. Baron, D. Diermeier, Elections, governments, and parliaments in proportional representation systems, Q. J. Econ. 116 (3) (2001) 933–967.
[25] P. Norris, Choosing electoral systems: proportional, majoritarian and mixed systems, Int. Polit. Sci. Rev. 18 (3) (1997) 297–312.
[26] D. Cassone, D. Ben-Arieh, Successive proportional additive numeration using fuzzy linguistic labels (Fuzzy Linguistic Span), Fuzzy Optim. Decis. Mak. 4 (3) (2005) 155.
[27] E. Diana, W. Gill, M. Kearns, K. Kenthapadi, A. Roth, Minimax group fairness: algorithms and experiments, in: Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society, 2021, pp. 66–76.
[28] R. Becker, G. D'angelo, S. Ghobadi, H. Gilbert, Fairness in influence maximization through randomization, J. Artif. Intell. Res. 73 (2022) 1251–1283.
[29] R. Kohavi, R. Longbotham, Online controlled experiments and A/B testing, Encycl. Mach. Learn. Data Min. 7 (8) (2017) 922–929.
[30] D.P. Schultz, The human subject in psychological research, Psychol. Bull. 72 (3) (1969) 214.

[31] J.B. Orlin, A polynomial time primal network simplex algorithm for minimum cost flows, Math. Program. 78 (2) (1997) 109–129.
[32] F. Corno, T. Montanaro, C. Migliore, P. Castrogiovanni, SmartBike: an IoT crowd sensing platform for monitoring city air pollution, Int. J. Electr. Comput. Eng. 7 (6) (2017) 3602.
[33] J. Dutta, C. Chowdhury, S. Roy, A.I. Middya, F. Gazi, Towards smart city: sensing air quality in city based on opportunistic crowd-sensing, in: Proceedings of the 18th International Conference on Distributed Computing and Networking, 2017, pp. 1–6.
[34] C. Feng, W. Wang, Y. Tian, X. Que, X. Gong, Estimate air quality based on mobile crowd sensing and big data, in: 2017 IEEE 18th International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM), IEEE, 2017, pp. 1–9.
[35] S.d.M.C. Grossmann, M.D.G. Moura, M.D.P. Matias, S.M. Paiva, R.A. Mesquita, The use of social networks in scientific research with questionnaires, Braz. J. Oral Sci. 17 (2018) e18162.
[36] H.N. Eke, N.J. Odoh, The use of social networking sites among the undergraduate students of University of Nigeria, Nsukka, Libr. Philos. Pract. (2014) 0_1.
[37] N. Chen, N. Gravin, P. Lu, On the approximability of budget feasible mechanisms, in: Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms, SIAM, 2011, pp. 685–699.
[38] G. Amanatidis, G. Birmpas, E. Markakis, On budget-feasible mechanism design for symmetric submodular objectives, in: International Conference on Web and Internet Economics, Springer, 2017, pp. 1–15.
[39] Y. Singer, M. Mittal, Pricing mechanisms for crowdsourcing markets, in: Proceedings of the 22nd International Conference on World Wide Web, 2013, pp. 1157–1166.
[40] P. Buisseret, C. Prato, Voting behavior under proportional representation, J. Theor. Polit. 32 (1) (2020) 96–111.
[41] N. Benabbou, M. Chakraborty, X.-V. Ho, J. Sliwinski, Y. Zick, The price of quota-based diversity in assignment problems, ACM Trans. Econ. Comput. 8 (3) (2020) 1–32.
[42] S. Ahmadi, F. Ahmed, J.P. Dickerson, M. Fuge, S. Khuller, An algorithm for multi-attribute diverse matching, in: Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, 2021, pp. 3–9.
[43] H. Aziz, S. Gaspers, Z. Sun, T. Walsh, From matching with diversity constraints to matching with regional quotas, in: Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems, 2019, pp. 377–385.
[44] L. Ehlers, I.E. Hafalir, M.B. Yenmez, M.A. Yildirim, School choice with controlled choice constraints: hard bounds versus soft bounds, J. Econ. Theory 153 (2014) 648–683.
[45] H. Aziz, S. Gaspers, Z. Sun, T. Walsh, From matching with diversity constraints to matching with regional quotas, arXiv preprint, arXiv:2002.06748.
[46] N. Benabbou, M. Chakraborty, X.-V. Ho, J. Sliwinski, Y. Zick, Diversity constraints in public housing allocation, in: 17th International Conference on Autonomous Agents and MultiAgent Systems, AAMAS 2018, 2018.
[47] M. Viana, T.A. Kuhlbusch, X. Querol, A. Alastuey, R.M. Harrison, P.K. Hopke, W. Winiwarter, M. Vallius, S. Szidat, A.S. Prévôt, et al., Source apportionment of particulate matter in Europe: a review of methods and results, J. Aerosol Sci. 39 (10) (2008) 827–849.
[48] P.K. Hopke, Review of receptor modeling methods for source apportionment, J. Air Waste Manage. Assoc. 66 (3) (2016) 237–259.
[49] H.F. Niemeyer, A.C. Niemeyer, Apportionment methods, Math. Soc. Sci. 56 (2) (2008) 240–253.
[50] C. Ilvento, M. Jagadeesan, S. Chawla, Multi-category fairness in sponsored search auctions, in: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, 2020, pp. 348–358.
[51] K. Kuo, A. Ostuni, E. Horishny, M.J. Curry, S. Dooley, P.-y. Chiang, T. Goldstein, J.P. Dickerson, ProportionNet: balancing fairness and revenue for auction design with deep learning, arXiv preprint, arXiv:2010.06398.
[52] J. Finocchiaro, R. Maio, F. Monachou, G.K. Patro, M. Raghavan, A.-A. Stoica, S. Tsirtsis, Bridging machine learning and mechanism design towards algorithmic fairness, arXiv preprint, arXiv:2010.05434.
[53] S. Chawla, M. Jagadeesan, Fairness in ad auctions through inverse proportionality, arXiv preprint, arXiv:2003.13966.
[54] C. Daskalakis, S.M. Weinberg, Bayesian truthful mechanisms for job scheduling from bi-criterion approximation algorithms, in: Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms, SIAM, 2014, pp. 1934–1952.
[55] P. Dhangwatnotai, S. Dobzinski, S. Dughmi, T. Roughgarden, Truthful approximation schemes for single-parameter agents, SIAM J. Comput. 40 (3) (2011) 915–933.
[56] A. Abdulkadiroğlu, T. Sönmez, School choice: a mechanism design approach, Am. Econ. Rev. 93 (3) (2003) 729–747.
[57] Y.A. Gonczarowski, N. Nisan, L. Kovalio, A. Romm, Matching for the Israeli “mechinot” gap-year programs: handling rich diversity requirements, in: Proceedings of the 2019 ACM Conference on Economics and Computation, 2019, p. 321.
[58] H. Aziz, S. Gaspers, Z. Sun, Mechanism design for school choice with soft diversity constraints, in: Proceedings of the Twenty-Ninth International Conference on International Joint Conferences on Artificial Intelligence, 2021, pp. 153–159.
[59] R.B. Myerson, Optimal auction design, Math. Oper. Res. 6 (1) (1981) 58–73.
[60] J.B. Orlin, Max flows in O (nm) time, or better, in: Proceedings of the Forty-Fifth Annual ACM Symposium on Theory of Computing, 2013, pp. 765–774.
[61] U. Zwick, The smallest networks on which the Ford-Fulkerson maximum flow procedure may fail to terminate, Theor. Comput. Sci. 148 (1) (1995) 165–170.