Coordinated Dynamic Bidding in Repeated Second-Price Auctions with Budgets

Yurong Chen [* 1] **Qian Wang** [* 1] **Zhijian Duan** [1] **Haoran Sun** [2] **Zhaohua Chen** [1] **Xiang Yan** [3] **Xiaotie Deng** [1 4]

Abstract

In online ad markets, a rising number of advertisers are employing bidding agencies to participate in ad auctions. These agencies are specialized in designing online algorithms and bidding on behalf of their clients. Typically, an agency usually has information on multiple advertisers, so she can potentially coordinate bids to help her clients achieve higher utilities than those under independent bidding. In this paper, we study coordinated online bidding algorithms in repeated second-price auctions with budgets. We propose algorithms that guarantee every client a higher utility than the best she can get under independent bidding. We show that these algorithms achieve maximal coalition welfare and discuss bidders' incentives to misreport their budgets, in symmetric cases. Our proofs combine the techniques of online learning and equilibrium analysis, overcoming the difficulty of competing with a multi-dimensional benchmark. The performance of our algorithms is further evaluated by experiments on both synthetic and real data. To the best of our knowledge, we are the first to consider bidder coordination in online repeated auctions with constraints.

1. Introduction

Online advertising has been the main source of revenue for important Internet companies (IAB, 2022). A transaction in the market usually goes as follows: upon the arrival of an ad query, an auction is held by the advertising platform among all advertisers, and the winner gets the opportunity to display her ad. Thus, advertisers engaging in such repeated

---

[*]Equal contribution [1]Center on Frontiers of Computing Studies, School of Computer Science, Peking University, Beijing, China [2]Peking University, Beijing, China [3]Huawei TCS Lab, Shanghai, China [4]Center for Multi-Agent Research, Institute for AI, Peking University, Beijing, China. Correspondence to: Xiaotie Deng <xiaotie@pku.edu.cn>.

auctions are confronted with an online decision problem. That is, they could not know the values of incoming queries, but have to decide their bids based on historical information to maximize their accumulated utilities.

Recently, more and more advertisers are delegating their bidding tasks to specialized market agencies (Decarolis et al., 2020). Advertisers only need to report their objectives and constraints within a given number of rounds. Typically, an advertiser aims to maximize her total utility subject to a budget constraint that limits her total payment. With the requests received, these agencies will design corresponding online bidding algorithms and bid on behalf of these advertisers. A bidding agency usually has information of all her clients, which leaves her a chance to coordinate bids to improve every client's utility.

However, most existing online bidding algorithms focus on bidding for one bidder in the campaign (Han et al., 2020; Golrezaei et al., 2021; Balseiro et al., 2022; Chen et al., 2022; Feng et al., 2023; Wang et al., 2023). The fundamental question of bidder coordination in online settings remains unanswered: can we design a coordinated online bidding strategy, such that each client advertiser enjoys a higher utility compared to the best they can get under independent bidding?

In this paper, we study bidder coordination in repeated second-price auctions with budgets. The above question can be modeled as an online multi-objective optimization: to show that an algorithm has the desired performance, one needs to compare each bidder's utility with her corresponding benchmark utility. The main difficulty, compared to the problem of designing individual online bidding algorithms, is how we deal with the interplay of bidders in dynamic repeated settings. Since auctions are multi-player games, a bidder's utility is influenced by others' bids.

Indeed, one way to overcome the problem is to assume that members who gain more than their respective benchmark utility are able to compensate other members with money. In this way, optimizing the welfare of the whole coalition is sufficient (Bachrach et al., 2011; Allouah & Besbes, 2017). However, we argue that monetary transfer is unrealistic in the ad auction scenario. In reality, advertisers' first goal is usually to enlarge the influence of their products. They participate in auctions to win more impressions for their ads,

rather than simply conducting financial investments.

Therefore, we focus on coordinated online bidding algorithms without monetary transfer. We propose an algorithm, **H**ybrid coordinated **P**acing (HP), under which each bidder's expected utility is theoretically guaranteed to outperform her corresponding benchmark utility asymptotically. In designing HP, an important observation we follow is that the problem of coordinated online bidding strategy design can be reduced to two design tasks: the selection rule inside the coalition and the budget management strategies. A simple application of this principle, **C**oordinated **P**acing (CP), is another algorithm proposed in the paper. It already has the desired utility guarantees when bidders in the coalition share the same value distribution and the same amount of budget (i.e., symmetric case).

In proving our theorems and overcoming the aforementioned difficulty, we combine the techniques of online learning and equilibrium analysis. We first show that these algorithms' performance converges. The problem then reduces to the comparison of utilities between the benchmark and the equilibrium these algorithms converge to. The analysis of equilibria, which are defined by non-linear complementarity problems and have no closed-form solutions, allows us to show that our algorithms maximize coalition welfare in the symmetric case. We believe all these techniques and results will be of generality and important use in future studies. Finally, we run experiments on both synthetic and real data to further illustrate how our algorithms improve every member's utility through coordination.

1.1. Related Work

In recent years, individual online bidding problems with budget constraints in repeated auctions have been widely discussed in literature, promoted by the surging of auto-bidding services in the industry (Aggarwal et al., 2019). Theoretically, this problem is closely related to contextual bandits with knapsacks (CBwK) (Badanidiyuru et al., 2014; Agrawal et al., 2016; Agrawal & Devanur, 2016; Sivakumar et al., 2022) as well as its specifications, e.g., online allocation problems (Li & Ye, 2021; Balseiro et al., 2022). For the dynamic bidding problem, Balseiro & Gur (2019) set the foundation of pacing in repeated second-price auctions, by which a bidder would shade values by a constant factor as her bids. They proposed an optimal online bidding algorithm that adaptively adjusts the pacing parameter. Subsequent works have extended the above results in multiple ways (Golrezaei et al., 2021; Gaitonde et al., 2023; Celli et al., 2022; Gaitonde et al., 2023). Our work considers the coordinated bidding scenario, in which bidders could cooperate to increase everyone's utility.

Graham & Marshall (1987) studied coordinated strategies in second-price auctions and first proposed to select only one member in the coalition as the representative to bid in the auction. Pesendorfer (2000); Leyton-Brown et al. (2002); Aoyagi (2003); Skrzypacz & Hopenhayn (2004); Romano et al. (2022) also consider models that do not allow monetary transfer in the coalition. Among these researches, Decarolis et al. (2020); Romano et al. (2022) studied the computation problem of optimizing each bidder's utility in GSP and VCG auctions. In comparison, we consider the coordination problem in online repeated second-price auctions with budgets, which is one of the most practical scenarios in current ad markets.

Technically, for the budget management strategy of each individual bidder, we adopt the adaptive pacing strategy introduced by Balseiro & Gur (2019), which is based on dual gradient descent. While algorithms and analysis based on duality are standard in online convex optimization (Nemirovski et al., 2009; Li & Ye, 2021; Golrezaei et al., 2021; Feng et al., 2023), due to the introduction of inner dual variables, our analysis of HP does not follow the standard dual analysis. The idea of selecting a representative to bid is from Graham & Marshall (1987) and confirmed to be applicable to the online setting in this paper. Moreover, it is important and nontrivial to design a fair adaptive selection rule, so as to make sure that everyone has fairly equal chances to bid: While CP can be regarded as a natural application of individual pacing to coordinated bidding, it can only guarantee performance in symmetric cases. Zhou et al. (2017) also studied the case when all players learn simultaneously in repeated games. They established the last-iterate convergence of mirror descent strategies when the one-shot game satisfies the property of variational stability. However, their results do not apply directly to the settings when players' actions across rounds satisfy certain budget constraints. Moreover, our analysis includes an equilibrium analysis that compares per-person utilities in two equilibria.

Other works utilize the multi-agent reinforcement learning framework to design algorithms (Jin et al., 2018; Guan et al., 2021; Wen et al., 2022; Tan et al., 2022). They usually need to explore beforehand to get enough samples. Compared to them, the algorithms we propose do not need samples. They consider accumulated utilities and achieve asymptotically better performance than the benchmark.

1.2. Paper Organization

In Section 2, we introduce our model and benchmark. In Section 3, we confirm the idea of selecting representatives to bid, and present the results of CP. Section 4 presents the results of HP. In Section 5, we show other properties our algorithms possess in symmetric cases. Section 6 show experiment results of our algorithms. Section 7 concludes.

2. Model and Benchmark

In this work, we consider the scenario in which N bidders participate in T rounds of repeated second-price auctions. In each round $t = 1, \ldots, T$, there is an available ad slot, auctioned by the advertising platform. There are K among N bidders who form a coalition $\mathcal{K} = \{1, \cdots, K\}$. The value bidder $k \in \mathcal{K}$ perceives for the ad slot in round t is denoted by $v_{k,t}$ and is assumed to be i.i.d. sampled from a distribution F_k. We assume that F_k has a bounded density function f_k, with a support over $[0, \bar{v}_k] \subseteq \mathbb{R}_+$. Each bidder $k \in \mathcal{K}$ has a budget constraint B_k, which limits the sum of her payments throughout the period, and we denote bidder k's target expenditure rate by $\rho_k := B_k/T \in (0, \bar{v}_k)$. As usual, we use bold symbols for vectors (or matrices), e.g., using $\boldsymbol{v}_t$ without subscript k to denote the vector $(v_{1,t}, \ldots, v_{K,t})$, and $\boldsymbol{v}$ to denote $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_T)$; the same goes for $\boldsymbol{\rho}$ as well as other variables to be defined.

For $k \in \mathcal{K}$, denote bidder k's bid at round t by $b_{k,t}$. For bidders outside the coalition, we assume their highest bid in round t, denoted by d_t^O, is sampled i.i.d. from a distribution H. This modeling follows the standard mean-field approximation (Iyer et al., 2014) when the number of bidders outside the coalition is large. Similar to F_k, H is also assumed to have a bounded density function h. Thus, the highest competing bid that bidder k faces is:

$$d_{k,t} := \max\left\{ \max_{i \in \mathcal{K}: i \neq k} b_{i,t},\ d_t^O \right\}.$$

Let $x_{k,t} := \mathbf{1}\{b_{k,t} \geq d_{k,t}\} \in \{0,1\}$ indicate whether bidder k wins the ad slot in round t. We denote by $u_{k,t} := x_{k,t}(v_{k,t} - d_{k,t})$ bidder k's utility in round t and by $z_{k,t} := x_{k,t} d_{k,t}$ her corresponding expenditure for a second-price auction.

Now we formally model the coordination among bidders in $\mathcal{K}$, which can also be viewed as a third-party agency bidding on behalf of all bidders in $\mathcal{K}$, given their budget and value information. We denote by $\mathcal{H}_{\mathcal{K}}^t$ the history available to the coalition before posting the bids in round t, defined as:

$$\mathcal{H}_{\mathcal{K}}^t := \left\{ \{\boldsymbol{v}_\tau, \boldsymbol{x}_\tau, \boldsymbol{z}_\tau\}_{\tau=1}^{t-1}, \boldsymbol{v}_t \right\}.$$

A coordinated strategy maps $\mathcal{H}_{\mathcal{K}}^t$ to a (possibly random) bid vector $\boldsymbol{b}_t$ for each t. We use $\mathcal{C}$ to denote the set of all coordinated strategies, and call a strategy $\mathrm{C} \in \mathcal{C}$ feasible if it guarantees that for each bidder, her cumulative expenditures never exceed her budget for any realizations of values and competing bids, i.e., $\forall \boldsymbol{v}, \boldsymbol{d}^O$,

$$\sum_{t=1}^T z_{k,t}^{\mathrm{C}} = \sum_{t=1}^T \mathbf{1}\left\{b_{k,t}^{\mathrm{C}} \geq d_{k,t}^{\mathrm{C}}\right\} d_{k,t}^{\mathrm{C}} \leq B_k, \forall k \in \mathcal{K}.$$

We denote by Π_k^{C} bidder k's expected payoff under coordi-

nated algorithm $\mathrm{C} \in \mathcal{C}$:

$$\Pi_k^{\mathrm{C}} := \mathbb{E}_{\boldsymbol{v}, \boldsymbol{d}^O}^{\mathrm{C}} \left[\sum_{t=1}^T u_{k,t}^{\mathrm{C}} \right]$$

$$= \mathbb{E}_{\boldsymbol{v}, \boldsymbol{d}^O}^{\mathrm{C}} \left[\sum_{t=1}^T \mathbf{1}\left\{b_{k,t}^{\mathrm{C}} \geq d_{k,t}^{\mathrm{C}}\right\} \left(v_{k,t} - d_{k,t}^{\mathrm{C}}\right) \right],$$

where the expectation is taken with respect to the randomness of algorithm C, the values of all coalition members $\boldsymbol{v}$, and the highest bids outside the coalition $\boldsymbol{d}^O$.

2.1. Optimal Individual Budget Management Algorithm and Benchmark

A desirable coordinated bidding algorithm should benefit every participant. That is, all participants should gain more utility than what they could obtain when they bid independently. This is also a requirement for stable coordination, as a bidder should have left the coalition if she gained less in it. A natural measurement is bidders' utilities when everyone uses optimal individual bidding algorithms independently.

As for individual bidding algorithms, it is known that adaptive pacing strategy (Balseiro & Gur, 2019) is asymptotically optimal in both stochastic and adversarial environments. The main idea of this strategy is to maintain a shading parameter λ, and bid $v/(1 + \lambda)$ when the true value is v. The algorithm updates λ according to each round's cost and tries to keep the bidder's average expenditure per auction close to the target expenditure rate. By doing so, the algorithm shares the risk as well as the opportunity among all rounds. For the completeness of our work, we state the individual adaptive pacing algorithm in Appendix A.

When all bidders follow adaptive pacing simultaneously, their shading parameters converge, assuming that the multi-valued function of their expected expenditures per round is strongly monotone (defined below). At the same time, their utilities also converge to the corresponding equilibrium utilities, which we regard as our benchmark.

Specifically, consider K bidders shade bids according to a profile of parameters $\boldsymbol{\lambda} \in \mathbb{R}_+^K$. Denote bidder k's expected expenditure per auction under $\boldsymbol{\lambda}$ by

$$G_k(\boldsymbol{\lambda}) := \mathbb{E}_{\boldsymbol{v}, d^O} \left[\mathbf{1}\{v_k \geq (1 + \lambda_k)d_k\} \cdot d_k \right], \quad (1)$$

and bidder k's expected utility over T rounds by

$$U_k(\boldsymbol{\lambda}) := T \cdot \mathbb{E}_{\boldsymbol{v}, d^O} \left[\mathbf{1}\{v_k \geq (1 + \lambda_k) d_k\} \cdot (v_k - d_k) \right],$$

where we omit the subscript t since the values and the highest bids are independent across rounds. We say function $\boldsymbol{F} : \mathbb{R}^K \to \mathbb{R}^K$ is γ-strongly monotone over a set $\mathcal{U} \subset \mathbb{R}^K$ if

$$(\boldsymbol{\lambda} - \boldsymbol{\lambda}')^\top (\boldsymbol{F}(\boldsymbol{\lambda}') - \boldsymbol{F}(\boldsymbol{\lambda})) \geq \gamma \|\boldsymbol{\lambda} - \boldsymbol{\lambda}'\|_2^2, \ \forall \boldsymbol{\lambda}, \boldsymbol{\lambda}' \in \mathcal{U}.$$

Assumption 2.1. There exists a constant $\gamma > 0$ such that $\boldsymbol{G} = (G_k)_{k=1}^K$ is a γ-strongly monotone function over $\prod_{k=1}^K [0, \bar{v}_k/\rho_k]$.

We call the special coordinated algorithm, the simultaneous adoption of individual adaptive pacing, IP for short. When Assumption 2.1 holds, parameters under IP converge to a fixed vector $\boldsymbol{\lambda}^*$, defined by the following complementarity conditions:

$$\lambda_k^* \geq 0 \perp G_k(\boldsymbol{\lambda}^*) \leq \rho_k, \quad \forall k \in \mathcal{K}, \tag{2}$$

where $\perp$ means that at least one condition holds with equality. The existence and uniqueness of $\boldsymbol{\lambda}^*$ are guaranteed by the result of Facchinei & Pang (2003). Moreover, its long-run average performance converges to the expected utilities achieved with vector $\boldsymbol{\lambda}^*$, i.e.,

$$\lim_{T \to \infty} \frac{1}{T} \left(\Pi_k^{\mathrm{IP}} - U_k(\boldsymbol{\lambda}^*) \right) = 0, \quad \forall k \in \mathcal{K}. \tag{3}$$

We consider $\boldsymbol{U}(\boldsymbol{\lambda}^*)$ as our benchmark. Notably, our goal is to design coordinated algorithms with better performance for every bidder inside the coalition, i.e., finding an algorithm C such that

$$\liminf_{T \to \infty} \frac{1}{T} \left(\Pi_k^{\mathrm{C}} - U_k(\boldsymbol{\lambda}^*) \right) \geq 0, \quad \forall k \in \mathcal{K}.$$

3. First Steps for Coordinated Strategy Design

In this section, we present two important elements used in our algorithms. We first show that we only need to consider a class of "bid rotation" strategies to design satisfying coordinated bidding algorithms. We then use a simple algorithm to illustrate the non-triviality of the problem and the importance of maintaining competition inside the coalition.

In a one-shot second-price auction where bidders have no constraint, as is revealed, the best collusive strategy is to select a representative to bid truthfully and let the other bidders inside the coalition bid zero. Such a strategy decreases the second-highest bid, alleviates the inner competition, and notably, improves the representative's utility without making others worse (Graham & Marshall, 1987). When it comes to the multi-round scenario with budget constraints, it is no longer optimal for bidders to bid truthfully. In fact, they tend to underbid to control their expenditures, making sure not to exceed their budgets (Aggarwal et al., 2019). However, we show that the idea of selecting representatives still works.

Definition 3.1. A coordinated bidding strategy is a *bid rotation*, if it lets at most one bidder bid non-zero in each round.

Lemma 3.2 indicates that there always exists a best coordinated bidding algorithm that is a bid rotation. Therefore, it suffices to consider bid rotation strategies only.

Algorithm 1 Coordinated Pacing (CP)
___
1: **Input:** $\epsilon = 1/\sqrt{T}$, $\bar{\xi}_k \geq \bar{v}_k/\rho_k$ for all $k \in \mathcal{K}$.
2: Select an initial multiplier $\boldsymbol{\xi}_0 \in [0, \bar{\xi}_k]^K$, and set the remaining budget of agent k to $\widetilde{B}_{k,1} = B_k = \rho_k T$.
3: **for** $t = 1$ **to** T **do**
4:     Observe the realization of $\boldsymbol{v}_t$.
5:     Select $k^* \in \mathrm{argmax}_{k \in \mathcal{K}} \min\{v_{k,t}/(1 + \xi_{k,t}), \tilde{B}_{k,t}\}$ (break ties arbitrarily).
6:     For each k, post a bid
7:       $b_{k,t} = \min \left\{ \mathbf{1}\{k = k^*\} v_{k,t}/(1 + \xi_{k,t}), \tilde{B}_{k,t} \right\}$.
8:     Observe the the expenditures $\boldsymbol{z}_t$.
9:     For each k, update the multiplier by
10:       $\xi_{k,t+1} = P_{[0,\bar{\xi}_k]} (\xi_{k,t} - \epsilon(\rho_k - z_{k,t}))$.
11:     For each k, update the remaining budget by
12:       $\widetilde{B}_{k,t+1} = \widetilde{B}_{k,t} - z_{k,t}$.
13: **end for**

Lemma 3.2. *For any feasible coordinated strategy* C*, there exists a feasible bid rotation strategy* C^{P} *such that every member's expected utility under algorithm* C^{P} *is at least as good as that under algorithm* C*.*

Hereafter, we divide the algorithm design into two parts: designing the budget control strategies in the real auction and the selection rule inside the coalition. We use adaptive pacing for the budget management part, and the main difficulty lies in designing the rule for selecting the representative bidder each round.

One natural idea is to choose the member with the highest bid to compete outside the coalition. Such an idea motivates coordinated pacing (CP, Algorithm 1). In CP, Line 5 shows the selection step inside the coalition: the one with the highest bid becomes the winner. In Line 6, the winner posts the same bid as in the coalition, while others post zero in the real auction. Finally, the pacing parameters are updated with the adaptive pacing technique. We note that the step size can be set in a more general way, with details in the appendix. Here, we use $1/\sqrt{T}$ to simplify the description.

Now we give our theoretical guarantees for the performance of CP. First, we make assumptions on the function of bidders' expected expenditures per round. In the case of CP, denote the expected expenditure of bidder k under CP to be G_k^{CP}, which is

$$G_k^{\mathrm{CP}}(\boldsymbol{\xi}) := \mathbb{E}_{\boldsymbol{v},d^O} \left[\mathbf{1}\{v_k \geq (1 + \xi_k)d_k\} d^O \right].$$

Assumption 3.3. $\boldsymbol{G}^{\mathrm{CP}}$ *is strongly monotone in* $[0, \bar{\xi}_k]^K$.

Denote the expected utility of bidder k under CP by Π_k^{CP}. In the symmetric case, we show that the performance of CP is better than the benchmark $U_k(\boldsymbol{\lambda}^*)$ asymptotically:

Theorem 3.4. *Suppose that Assumption 3.3 holds. When every bidder has the same value distribution and budget, we have*

$$\liminf_{T \to \infty} \frac{1}{T}(\Pi_k^{\text{CP}} - U_k(\boldsymbol{\lambda}^*)) > 0, \ \forall k \in \mathcal{K}. \quad (4)$$

The proof of the theorem is given in Appendix B. To prove Theorem 3.4, we first show that $\boldsymbol{\xi}$ converges to some equilibrium $\boldsymbol{\xi}^*$ and, therefore, the algorithm's performance converges to the corresponding expected utilities. The problem then reduces to the comparison of equilibrium utilities. Concretely, we show that $\boldsymbol{\xi}^* \le \boldsymbol{\lambda}^*$. This means that as an entirety, bidders in coalition $\mathcal{K}$ are more competitive than when they independently bid, and therefore the coalition wins more items. However, the competition inside the coalition also becomes fiercer.

Despite the good performance in the symmetric case, CP does not outperform the benchmark in all instances. In some cases, strong bidders may become stronger, and weak bidders become weaker. To see this, we consider a special case with two bidders forming a coalition. Bidder 1's value is always higher than the bidder 2's, and the outside competing bid d^O is always 0. For IP, which converges to our benchmark, bidder 1's expected expenditure per round is not zero. Therefore her shading parameter is non-zero, so long as her budget is sufficiently small. There is a chance that her bid is lower than bidder 2's, enabling bidder 2 to compete and win the item. In CP, however, bidder 1 keeps being selected as the representative, and bidder 2's expected utility will be lower than hers in the benchmark. A detailed experimental example is given in Section 6.

To avoid such cases and outperform the benchmark in all instances, one way is to independently maintain competition inside the coalition, giving each member a fair chance to become the representative. The ideas of bid rotation and maintaining the competition are important in designing HP, which we will show next.

4. Hybrid Coordinated Pacing

Inspired by CP (Algorithm 1), we further propose a hybrid coordinated pacing (HP) algorithm shown in Algorithm 2. HP ensures that every coalition member can achieve a higher utility than the benchmark.

More specifically, HP draws a clear line between the internal election and external competition. In addition to the multiplier μ_k used to compete with bidders outside the coalition, HP maintains for each bidder k a pseudo multiplier λ_k for the internal election. The coalition first holds an internal election to select a representative and recommends her to compete with other bidders outside the coalition. When the value is v_k (and the budget is still sufficient), bidder k bids

---

Algorithm 2 Hybrid Coordinated Pacing (HP)

1: **Input:** $\epsilon = 1/\sqrt{T}$, $\bar{\mu}_k \ge \bar{v}_k/\rho_k$ for all $k \in \mathcal{K}$.
2: Select an initial multiplier $\boldsymbol{\mu}_1 \in [0, \bar{\mu}_k]^K$, let the initial pseudo multiplier be $\boldsymbol{\lambda}_1 = \boldsymbol{\mu}_1$, and set the remaining budget of agent k to $\tilde{B}_{k,1} = B_k = \rho_k T$.
3: **for** $T = 1$ **to** t **do**
4:     Observe the realization of $\boldsymbol{v}_t$.
5:     Let $b_{k,t}^I = \min\left\{v_{k,t}/(1 + \lambda_{k,t}), \tilde{B}_{k,t}\right\}$ and $d_{k,t}^I = \max_{i \in \mathcal{K}: i \ne k} b_{i,t}^I$ for each k.
6:     Select $k^* \in \arg\max_k b_{k,t}^I$ (breaking ties arbitrarily).

7:     For each k, post a bid
8:     $b_{k,t}^O = \min\left\{\mathbf{1}\left\{k = k^*\right\} v_{k,t}/(1 + \mu_{k,t}), \tilde{B}_{k,t}\right\}$.
9:     Observe the allocations $\boldsymbol{x}_t$ and the expenditures $\boldsymbol{z}_t$.
10:    For each k, update the pseudo multiplier by
11:    $\lambda_{k,t+1} = P_{[0,\bar{\mu}_k]}\left(\lambda_{k,t} - \epsilon\left(\rho_k - z_{k,t}'\right)\right)$,
12:    where $z_{k,t}' = \mathbf{1}\left\{b_{k,t}^I \ge z_{k,t}\right\} \max\left(z_{k,t}, x_{k,t} d_{k,t}^I\right)$.

13:    For each k, update the multiplier by
14:    $\mu_{k,t+1} = P_{[0,\lambda_{k,t+1}]}\left(\mu_{k,t} - \epsilon\left(\rho_k - z_{k,t}\right)\right)$.
15:    For each k, update the remaining budget by
16:    $\tilde{B}_{k,t+1} = \tilde{B}_{k,t} - z_{k,t}$.
17: **end for**

$v_k/(1 + \lambda_k)$ in the internal election, and bids $v_k/(1 + \mu_k)$ in the real auction if she wins. To make a good distinction, we label the two ranking bids as b_k^I and b_k^O with superscripts.

By updating λ_k and μ_k separately, we are able to make sure that every member gets at least the same opportunities to bid outside the coalition and win the item, as they do in IP. To achieve this, we let the inner dual variables update exactly as the dual variables in IP. Therefore, in the updating rule of μ_k, we require that the inner variable always be greater than the outer variable, i.e., $\mu_{k,t} \le \lambda_{k,t}$, so that if an agent wins at a round in IP, she will also win at the same round in HP. In this way, HP can observe every necessary expenditure z_t that will be used to update the dual variables in IP.

We denote by $G_k^{\text{HP}}(\mu_k, \boldsymbol{\lambda})$ bidder k's expected expenditure per round when all coalition members shade bids according to $\boldsymbol{\mu}$ and $\boldsymbol{\lambda}$ (ignoring their budget constraints):

$$G_k^{\text{HP}}(\mu_k, \boldsymbol{\lambda}) = \mathbb{E}_{\boldsymbol{v},d^O}\left[\mathbf{1}\left\{\frac{v_k}{1 + \lambda_k} \ge d_k^I \wedge \frac{v_k}{1 + \mu_k} \ge d^O\right\} d^O\right], \quad (5)$$

and we denote by $U_k^{\text{HP}}(\mu_k, \boldsymbol{\lambda})$ her corresponding expected utility in this case:

$$U_k^{\text{HP}}(\mu_k, \boldsymbol{\lambda}) = T \cdot \\ \mathbb{E}_{\boldsymbol{v},d^O}\left[\mathbf{1}\left\{\frac{v_k}{1 + \lambda_k} \ge d_k^I \wedge \frac{v_k}{1 + \mu_k} \ge d^O\right\}(v_k - d^O)\right].$$

We now make the following assumptions.

Assumption 4.1. 1. There exists a constant $\gamma > 0$ such that $\boldsymbol{G} = (G_k)_{k=1}^K$ is a γ-strongly monotone function over $\prod_{k=1}^K [0, \bar{\mu}_k]$.

2. There exists $\underline{G}' > 0$ such that $\partial G_k^{\mathrm{HP}}/\partial \mu_k < -\underline{G}'$ over $\prod_{k=1}^{K+1}[0, \bar{\mu}_k]$ for every bidder $k \in \mathcal{K}$.

Assumption 4.1 regularizes the behavior of G_k^{HP} with two assumptions on $\boldsymbol{\lambda}$ and μ_k, respectively. The first one directly follows from Assumption 2.1, which guarantees the convergence of $\boldsymbol{\lambda}$. Interestingly, we only require $\boldsymbol{G}$ instead of $\boldsymbol{G}^{\mathrm{HP}}$ to be γ-strongly monotone with respect to $\boldsymbol{\lambda}$, which differs from that Theorem 3.4 requires the strong monotonicity of $\boldsymbol{G}^{\mathrm{CP}}$. The second one assures the strict monotonicity of G_k^{HP} and U_k^{HP} with respect to μ_k. Intuitively, for a fixed $\boldsymbol{\lambda}$, if both G_k^{HP} and U_k^{HP} strictly decrease in μ_k, bidder k tends to decrease μ_k as long as her budget constraint is not binding, so as to maximize her utility. Therefore, there exists a unique equilibrium μ_k^* such that either $\mu_k^* = 0$ or any μ_k lower than μ_k^* would break the constraint, and we will show that the dynamically updated sequence of $\mu_{k,t}$ converges to such an equilibrium under HP. We further provide a sufficient condition for the second assumption to hold in Lemma E.2.

Theorem 4.2. *Suppose that Assumption 4.1 holds. We have*

$$\liminf_{T \to \infty} \frac{1}{T}(\Pi_k^{\mathrm{HP}} - U_k(\boldsymbol{\lambda}^*)) \geq 0, \forall k \in \mathcal{K}, \qquad (6)$$

and the equality holds for at most one bidder.

Theorem 4.2 establishes that with HP, all but one bidders in the coalition can get higher utilities, compared to $U_k(\boldsymbol{\lambda}^*)$ in an asymptotic sense. Unlike CP, this performance guarantee holds in a general setting where bidders may have different budget constraints or value distributions. It is worth noticing that, both HP, CP and IP have the same convergence rates. Therefore, HP and CP are also comparable to IP, taking the convergence rates into consideration.

To prove Theorem 4.2, we first show the convergence of both the multipliers $\boldsymbol{\mu}$ and the pseudo multipliers $\boldsymbol{\lambda}$. Observe that the update of λ_k is equivalent to

$$\lambda_{k,t+1} = P_{[0,\bar{\lambda}_k]} \left(\lambda_{k,t} - \epsilon \left(\rho_k - \mathbf{1} \left\{ b_{k,t}^I \geq d_{k,t} \right\} d_{k,t} \right) \right), \qquad (7)$$

which coincides with the subgradient descent scheme of an individual adaptive pacing strategy. Note that the Euclidean projection operator in Line 14 assures that for all t,

$$\mu_{k,t} \leq \lambda_{k,t} \implies \frac{v_{k,t}}{1 + \lambda_{k,t}} \leq \frac{v_{k,t}}{1 + \mu_{k,t}}.$$

Thus by bidding $v_{k,t}/(1 + \mu_{k,t})$, we can always discern $\mathbf{1} \left\{ v_{k,t}/(1 + \lambda_{k,t}) \geq d_t^O \right\}$. This provides enough information to let the sequence of pseudo multipliers exactly behaves as if bidders are independently pacing and converges to the same $\boldsymbol{\lambda}^*$ defined by (2).

At the same time, we show that the sequence of multipliers $\boldsymbol{\mu}_t$ converges to a unique $\boldsymbol{\mu}^*$, defined by the complementarity conditions,

$$\mu_k^* \geq 0 \perp G_k^{\mathrm{HP}}(\mu_k^*, \boldsymbol{\lambda}^*) \leq \rho_k, \forall k \in \mathcal{K}. \qquad (8)$$

Consequently, we show that the expected utility of bidder k under HP converges to $U_k^{\mathrm{HP}}(\mu_k^*, \boldsymbol{\lambda}^*)$. Finally, the proof of Theorem 4.2 is concluded by a comparison between $U_k^{\mathrm{HP}}(\mu_k^*, \boldsymbol{\lambda}^*)$ and $U_k(\boldsymbol{\lambda}^*)$.

One may consider whether the inequality can strictly hold for everyone as in Theorem 3.4. Here we provide a concrete example in which a bidder gains no utility boost.

Example 4.3. *We consider two bidders with no budget constraints. Bidder 1's value v_1 is uniformly distributed over $[0, 2]$, while bidder 2' value v_2 is uniformly distributed over $[0, 1] \cup [4, 5]$. The highest competing bid of other bidders d^O follows a uniform distribution over $[1, 5]$. One can verify that this instance satisfies Assumption 4.1.*

When bidding independently, the optimal strategy for each bidder is to bid truthfully with respect to second-price auctions, i.e., $\lambda_1 = \lambda_2 = 0$. Then bidder 1 wins an auction only if $v_2 \in [0, 1]$ and $v_1 \geq d^O \geq 1$, which means her winning payment is always equal to d^O. Therefore, bidder 1 cannot benefit from forming a coalition with bidder 2.

5. Other Properties of Proposed Algorithms

In this section, we discuss the properties of CP and HP in the symmetric case. Recall that by showing the performance of the algorithms converges, we are able to study their utility properties by analyzing the equilibria they converge to. In what follows, we will assume the corresponding assumptions hold (i.e., Assumptions 2.1, 3.3 and 4.1), and all the results hold the in an asymptotic sense. We note that since these equilibria are defined as solutions to non-linear complementarity problems (NCP), and do not have closed forms, they can have various possibilities and are hard to analyze. Therefore, we leave the exploitation of their properties in asymmetric cases as an open question.

5.1. Coalition Welfare Maximization

We first show that CP and HP are the best coordinated algorithms in the sense of coalition welfare maximization and Pareto optimality. Specifically, we consider the hindsight maximal coalition welfare, given a realization of $(\boldsymbol{v}; \boldsymbol{d}^O)$, the values and highest bids outside the coalition:

$$\pi^{\mathrm{H}}\left(\boldsymbol{v}; \boldsymbol{d}^O\right) := \max_{\boldsymbol{x}_k \in \{0,1\}^{K \times T}} \sum_{t=1}^T \sum_{k \in \mathcal{K}} x_{k,t}(v_{k,t} - d_t^O),$$

$$\text{s.t.} \sum_{t=1}^T x_{k,t} d_t^O \leq T\rho_k, \forall k \in \mathcal{K}.$$

We further say a bidding algorithm C is Pareto optimal in the symmetric case, if there does not exist another coordinated bidding algorithm, such that in some symmetric instance, every bidder's expected utility is no worse than that under C, and at least one bidder's expected utility is strictly better.

Theorem 5.1. *When every bidder in the coalition has the same value distribution and budget, CP and HP maximize the total expected utility of the coalition asymptotically,*

$$\liminf_{T \to \infty} \frac{1}{T} \left(\sum_{k \in \mathcal{K}} \Pi^{\mathrm{CP/HP}} - \mathbb{E}[\pi^{\mathrm{H}}] \right) \geq 0.$$

Therefore, they are Pareto optimal.

A key of the proof is that the sum of expected utilities CP and HP converge to, $\sum_{k \in \mathcal{K}} U_k^{\mathrm{CP}}(\boldsymbol{\xi}^*)$ and $\sum_{k \in \mathcal{K}} U_k^{\mathrm{HP}}(\mu_k^*, \boldsymbol{\lambda}^*)$, have the same form as an upper bound of $\mathbb{E}[\pi^{\mathrm{H}}]$ in the symmetric case. The simplicity of CP allows us to use it as a surrogate to prove the properties of HP.

5.2. Truthfulness on Misreporting Budgets

In this section, we will study how a bidder's obtained value and utility at equilibrium change if she misreports her budget, with others truthfully behaving. We assume the values are publicly known and only consider the possibility that a bidder may under-report her budget, as over-reporting may let her total cost exceed her budget.

We directly consider $\boldsymbol{G}$ as a function of $\boldsymbol{\rho}$: $\boldsymbol{G}(\boldsymbol{\rho}) = \boldsymbol{G}(\boldsymbol{\lambda}^*)$ where $\boldsymbol{\lambda}^*$ is defined by (2) with respect to $\boldsymbol{\rho}$. $\boldsymbol{G}^{\mathrm{CP}}(\boldsymbol{\rho}), \boldsymbol{U}^{\mathrm{CP}}(\boldsymbol{\rho})$ and $\boldsymbol{V}(\boldsymbol{\rho})$ are defined similarly, where

$$V_k(\boldsymbol{\rho}) = V_k(\boldsymbol{\lambda}^*) = T \cdot \mathbb{E} \left[\mathbf{1} \{ v_k \geq d_k \} v_k \right]$$

is bidder k's expected obtained value under $\boldsymbol{\lambda}^*$ in IP. Theorem 5.2 shows that CP is incentive compatible.

Theorem 5.2. *Under CP, bidder k's expected utility U_k^{CP} will not increase, if she under-reports her budget and others report budgets truthfully.*

Contrary to Theorem 5.1, CP and HP have different forms when bidders deviate. Recall that IP can be regarded as a special coordinated bidding algorithm, and the pseudo parameters of HP converge to the same equilibrium as IP. Therefore, we study IP first.

Theorem 5.3. *Under IP, bidder k's expected obtained value V_k will not increase if she under-reports her budget and others report budgets truthfully.*

To prove Theorem 5.2 and 5.3, we compare the equilibrium the algorithm converges to when bidder k misreports with that when she truthfully behaves. Note that the equilibria the algorithms converge to are defined by NCPs, which are

generally difficult to study. We analyze all 8 possible solutions of the corresponding NCPs. By applying the definition of NCP and strong monotonicity, we show that bidder k's average expected obtained value at the equilibrium does not increase by misreporting.

Theorem 5.3 shows that IP has good incentive compatibility on obtained values. However, when it comes to utilities, there are chances that the decrease in cost outweighs that in value, and a bidder's utility may overall increase. An experimental example is shown in the appendix.

Moreover, as the analysis of IP is complicated enough, we make the incentive problem of HP an open problem. Compared with the intricacy of HP, CP has relative succinct structure and is easy to analyze.

6. Experiments

It is known that the expected performance of IP converges to the benchmark equilibrium (Balseiro & Gur, 2019). In this section, we run IP together with CP and HP to demonstrate their performance. We first use an example to provide useful insights. Afterward, we conduct experiments on real-world and synthetic data, as a supplement to our theoretical results.

6.1. An Asymmetric Example

Consider the example where there are only two bidders in the coalition. The valuation of bidder 1 is sampled from the uniform distribution in $[0, 1]$, $\mathcal{U}[0, 1]$, with probability p, and from $\mathcal{U}[1, 1 + \eta]$ otherwise. $p, \eta \in (0, 1)$ are two small constants. The valuation of bidder 2 is sampled from $U[0, 1]$. We consider the extreme case where d_t^O is always zero. As a result, the representative bidder in CP or HP has zero expenditure. For both CP and HP, We set the initial multipliers and their upper bounds as 0 and $\bar{\mu}_i = 3.0$, respectively. We also set $\rho_i = \bar{v}_i / \bar{\mu}_i$ for each bidder. The strong monotonicity assumption can be validated experimentally by computing the minimal marginal monotone parameter within small grids. When the grid width is 0.003, the value is 0.035. We use a step size $\epsilon = 0.1 T^{-0.5}$.

We demonstrate the empirical simulation of the example in Figure 1. The experiment is run $1,000$ times, each time with $20,000$ rounds. The average performance is presented. We can see from the figure that by using CP, bidder 2 receives lower utility than using IP. This result indicates that CP does not outperform the benchmark in this instance. Intuitively, in the extreme case where $p \to 0$, it always holds that $v_{1,t} > v_{2,t}$. Since bidder 1's expenditure per round is always zero, she can set her parameter as 0. Therefore, bidder 1 always wins the inner election, leaving bidder 2 no chance to win any item. In comparison, bidder 2 is more likely to win in IP and HP. The reason is that in IP, bidder 1 has a positive payment, so she has to lower her bid if her budget
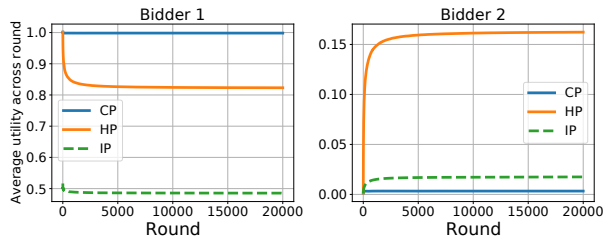
Figure 1. The utility convergence results of an asymmetric example. The y-axes are the average utilities across rounds (Π_k/T). CP does not outperform IP concerning bidder 2's utility.

is small. This enables bidder 2 to win when bidder 1's bid is low enough. In HP, the opportunities for competing outside the coalition are allocated the same way as in IP. Therefore, both bidders get chances to win, but their expenditures are lower than those under IP.

6.2. Experiments on Real Data and Synthetic Data

We present experiments on real-world data to evaluate the performance of our proposed algorithms, and justify its robustness through experiments on various synthetic data.

Experiments Set-Up. The real-world data we use, the iPinYou dataset [1] (Liao et al., 2014; Zhang et al., 2014), was provided for a real-time bidding competition organized by iPinYou. We use the testing data of Season 2, which contains $2,521,630$ records from 5 different bidders. Each piece of data contains the information of the user, the ad slot, and the bid, from which we mainly use the bidding (winning) price, paying price, and advertiser ID.

We first linearly scale all the bidding prices into $[0, 1]$. To make up K coordinated bidders, we randomly divide the historical bidding records of each of the 5 raw bidders into $K/5$ parts. Subsequently, we obtain K bidding records. We sample the valuation of each coalition member by drawing with replacement from her corresponding bidding record. Besides, we sample the highest bid outside the coalition, d_t^O, from $\mathcal{N}(0.5, 0.2)$. As for the budget, we first select $\bar{\rho}$, the mean of the target expenditure ratio, to be 0.01α with $\alpha \in \{1, \dots, 100\}$. Then we sample each coalition member's target expenditure ratio i.i.d. from $\mathcal{N}(\bar{\rho}, 0.1)$. We set the initial multipliers and their upper bounds as 1.0 and 20.0, respectively. We take step size $\epsilon = 0.1T^{-0.5}$ and $T = 20,000$. At last, we let $K \in \{5, 15\}$, and run 20 experiments for each setting with different random seeds.

We also manually design different distributions to verify the superiority of our algorithms. Due to page limits, we show only one typical example. We set $K = 5$ and compare

the coalition welfare and bidders' utility improvements in both symmetric and asymmetric cases, respectively. More results can be found in Appendix F. For the symmetric case, we sample each bidder's value i.i.d. from $\mathcal{N}(0.5, 0.2)$ and let their common target expenditure ratio be 0.01α for $\alpha \in \{1, \dots, 100\}$. For the asymmetric case, we sample each bidder's value from five different distributions, and each member's target expenditure ratio is selected the same as in real data experiments. For both cases, the highest bid outside the coalition is also sampled i.i.d. from $\mathcal{N}(0.5, 0.2)$. We note that the experiments of different distributions of the highest bid outside the coalition were conducted. The results are shown in Appendix F.

Experimental Results. Figure 2a and 2b show algorithms' performance in real data. We can see that Both CP and HP significantly improve coalition welfare compared with IP. The improvement increases with each member's budget. Such a result attributes to the reduction of inner competition in CP and HP. When bidder's budgets are large, the inner competition becomes fiercer under independent bidding. That is also why in IP, the coalition welfare tends to be 0 as people's budgets increase. Moreover, CP is slightly higher than HP in coalition welfare, whereas the method brings greater utility variance, and cannot guarantee that everyone gets better than IP. This coincides with our analysis in Section 3 that CP makes the entire coalition more competitive. A further observation is that the utility improvements of HP and CP decrease with the increase of K, the coalition size. The reason is that the number of rounds is fixed, and so does the number of items to be auctioned. Therefore, each person's share decreases as more people join the coalition.

The simulation results are shown in Figure 2c and 2d. In the symmetric case, we see that CP and HP are identical in both coalition welfare and utility improvement. This is consistent with the concrete forms of the algorithms. The asymmetric case reflects similar characteristics with real data. That is, HP and CP both improve total welfare significantly. In HP, all the participants receive higher utilities than in IP. In CP, the utilities of the vast majority can be improved. All the experimental results demonstrate the benefits of reducing the inner competition among members in the coalition.

Recall that the theoretic results predict that the averages of multipliers will converge. In the experiments, the variances of the last 100 rounds of the averages of multipliers are around 10^{-4}, indicating the convergence of multipliers. Further details show that they converge to the point predicted by our theorems, which are shown in Appendix F. Therefore, while we did not verify the assumptions made in previous sections, the experimental results further confirm the robustness of our algorithms.

(a) Coalition welfare in real-data experiments.

(b) Utility improvements in real-data experiments.

(c) Coalition welfare in simulation experiments.

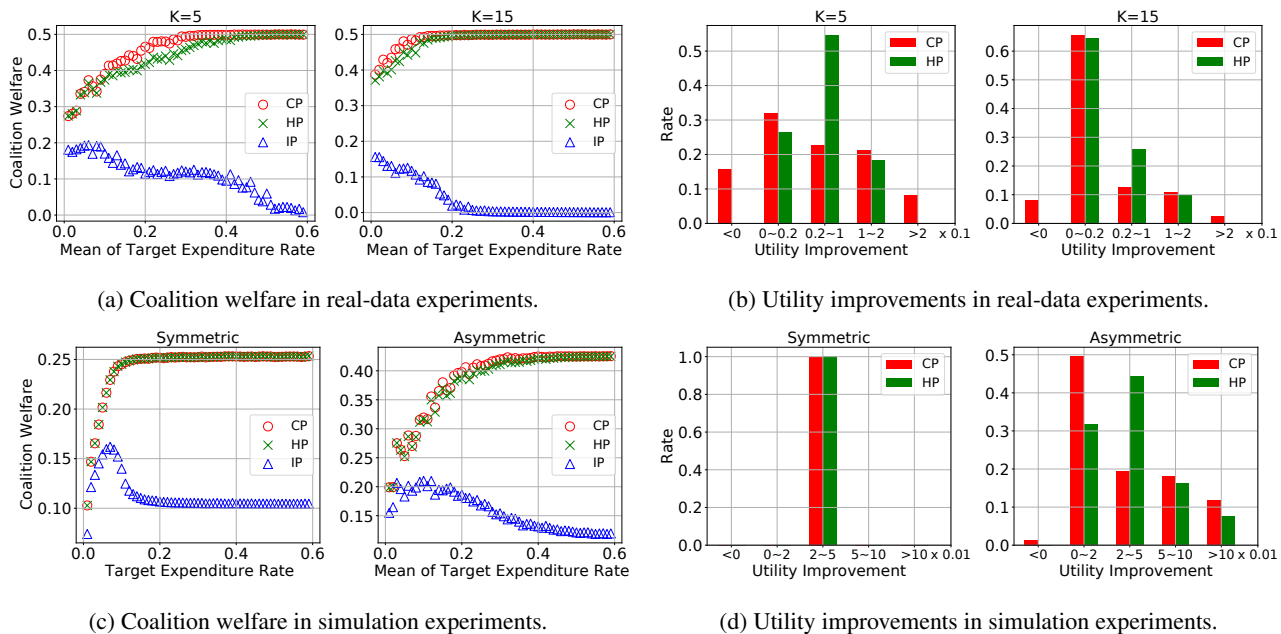(d) Utility improvements in simulation experiments.

Figure 2. Performance of coordinated pacing (CP) and hybrid coordinated pacing (HP) against individual pacing (IP) in real-data and simulated data: Figure 2a and 2c show the changes of coalition welfare ($\sum_{k \in \mathcal{K}} \Pi_k / T$) with target expenditure ratio varies. Figure 2a shows the result on iPinYou real data and Figure 2c on the simulated data. Each point is an average of 20 repeated experiments. Figure 2b and 2d the utility increase of coalition members increase after transferring to CP/HP from IP, captured by ($\Pi_k^{\mathrm{CP/HP}}/T - \Pi_k^{\mathrm{IP}}/T$) when the target expenditure rate ranges from 0.1 to 0.3. Figure 2b shows the results on real data, while Figure 2d on simulated data.

7. Conclusion

In this paper, we study designing online coordinated bidding algorithms in repeated second-price auctions with budgets. We propose algorithms that guarantee every coalition member a better utility than the best they can get under independent bidding. We analyze the game-theoretic properties of the algorithms like coalition welfare maximization and bidders' incentives to misreport their budget constraints. Experiment results on both synthetic data and real-world data show the effectiveness and robustness of our proposed algorithms.

There are quite a few open problems we can ask about. The first is whether we can relax the assumptions made in this paper, e.g., the strong monotonicity of bidders' expected expenditure functions. Second, in this paper, we assume that the highest bid outside the coalition is drawn from a stationary distribution. Another future work is to assume that agents outside the coalition also use IP to bid and study whether CP and HP can still improve every coalition member's utilities, compared to their independent pacing. As more bidders are taken into consideration, we can also ask about the fairness of the algorithms, before which an appropriate concept of fairness needs to be defined. One can also consider more complicated game-theoretic scenarios, e.g., taking the incentives of the bidding agency into account.

Acknowledgements

References

Aggarwal, G., Badanidiyuru, A., and Mehta, A. Autobidding with constraints. In *International Conference on Web and Internet Economics*, pp. 17–30. Springer, 2019.

Agrawal, S. and Devanur, N. Linear contextual bandits with knapsacks. *Advances in Neural Information Processing Systems*, 29, 2016.

Agrawal, S., Devanur, N. R., and Li, L. An efficient algorithm for contextual bandits with knapsacks, and an extension to concave objectives. In *Conference on Learning Theory*, pp. 4–18. PMLR, 2016.

Allouah, A. and Besbes, O. Auctions in the online display advertising chain: A case for independent campaign management. *Columbia Business School Research Paper*, (17-60), 2017.

Aoyagi, M. Bid rotation and collusion in repeated auctions. *Journal of economic Theory*, 112(1):79–105, 2003.

Bachrach, Y., Zadimoghaddam, M., and Key, P. A cooperative approach to collusion in auctions. *ACM SIGecom Exchanges*, 10(1):17–22, 2011.

Badanidiyuru, A., Langford, J., and Slivkins, A. Resourceful contextual bandits. In *Conference on Learning Theory*, pp. 1109–1134. PMLR, 2014.

Balseiro, S. R. and Gur, Y. Learning in repeated auctions with budgets: Regret minimization and equilibrium. *Management Science*, 65(9):3952–3968, 2019.

Balseiro, S. R., Lu, H., and Mirrokni, V. The best of many worlds: Dual mirror descent for online allocation problems. *Operations Research*, 2022.

Celli, A., Colini-Baldeschi, R., Kroer, C., and Sodomka, E. The parity ray regularizer for pacing in auction markets. In *Proceedings of the ACM Web Conference 2022*, pp. 162–172, 2022.

Chen, Z., Wang, C., Wang, Q., Pan, Y., Shi, Z., Tang, C., Cai, Z., Ren, Y., Zhu, Z., and Deng, X. Dynamic budget throttling in repeated second-price auctions. *arXiv preprint arXiv:2207.04690*, 2022.

Decarolis, F., Goldmanis, M., and Penta, A. Marketing agencies and collusive bidding in online ad auctions. *Management Science*, 66(10):4433–4454, 2020.

Facchinei, F. and Pang, J.-S. *Finite-dimensional variational inequalities and complementarity problems*. Springer, 2003.

Feng, Z., Padmanabhan, S., and Wang, D. Online bidding algorithms for return-on-spend constrained advertisers. In Ding, Y., Tang, J., Sequeda, J. F., Aroyo, L., Castillo, C., and Houben, G. (eds.), *Proceedings of the ACM Web Conference 2023, WWW 2023, Austin, TX, USA, 30 April 2023 - 4 May 2023*, pp. 3550–3560. ACM, 2023. doi: 10.1145/3543507.3583491. URL https://doi.org/10.1145/3543507.3583491.

Gaitonde, J., Li, Y., Light, B., Lucier, B., and Slivkins, A. Budget pacing in repeated auctions: Regret and efficiency without convergence. In Kalai, Y. T. (ed.), *14th Innovations in Theoretical Computer Science Conference, ITCS 2023, January 10-13, 2023, MIT, Cambridge, Massachusetts, USA*, volume 251 of *LIPIcs*, pp. 52:1–52:1. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2023. doi: 10.4230/LIPIcs.ITCS.2023.52. URL https://doi.org/10.4230/LIPIcs.ITCS.2023.52.

Golrezaei, N., Jaillet, P., Liang, J. C. N., and Mirrokni, V. Bidding and pricing in budget and ROI constrained markets. *arXiv preprint arXiv:2107.07725*, 2021.

Graham, D. A. and Marshall, R. C. Collusive bidder behavior at single-object second-price and English auctions. *Journal of Political economy*, 95(6):1217–1239, 1987.

Guan, Z., Wu, H., Cao, Q., Liu, H., Zhao, W., Li, S., Xu, C., Qiu, G., Xu, J., and Zheng, B. Multi-agent cooperative bidding games for multi-objective optimization in e-commercial sponsored search. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, pp. 2899–2909, 2021.

Han, Y., Zhou, Z., and Weissman, T. Optimal no-regret learning in repeated first-price auctions. *arXiv preprint arXiv:2003.09795*, 2020.

IAB, I. A. B. Internet advertising revenue report, full year 2021 results, 2022. URL https://www.iab.com/wp-content/uploads/2022/04/IAB_Internet_Advertising_Revenue_Report_Full_Year_2021.pdf.

Iyer, K., Johari, R., and Sundararajan, M. Mean field equilibria of dynamic auctions with learning. *Management Science*, 60(12):2949–2970, 2014.

Jin, J., Song, C., Li, H., Gai, K., Wang, J., and Zhang, W. Real-time bidding with multi-agent reinforcement learning in display advertising. In *Proceedings of the 27th ACM international conference on information and knowledge management*, pp. 2193–2201, 2018.

Leyton-Brown, K., Shoham, Y., and Tennenholtz, M. Bidding clubs in first-price auctions. In *AAAI/IAAI*, pp. 373–378, 2002.

Li, X. and Ye, Y. Online linear programming: Dual convergence, new algorithms, and regret bounds. *Operations Research*, 2021.

Liao, H., Peng, L., Liu, Z., and Shen, X. iPinYou global RTB bidding algorithm competition dataset. In *Proceedings of the Eighth International Workshop on Data Mining for Online Advertising*, pp. 1–6, 2014.

Nemirovski, A., Juditsky, A. B., Lan, G., and Shapiro, A. Robust stochastic approximation approach to stochastic programming. *SIAM J. Optim.*, 19(4):1574–1609, 2009. doi: 10.1137/070704277. URL https://doi.org/10.1137/070704277.

Pesendorfer, M. A study of collusion in first-price auctions. *The Review of Economic Studies*, 67(3):381–411, 2000.

Romano, G., Castiglioni, M., Marchesi, A., and Gatti, N. The power of media agencies in ad auctions: Improving utility through coordinated bidding. In *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence*, pp. 461–467, 2022.

Sivakumar, V., Zuo, S., and Banerjee, A. Smoothed adversarial linear contextual bandits with knapsacks. In *International Conference on Machine Learning*, pp. 20253–20277. PMLR, 2022.

Skrzypacz, A. and Hopenhayn, H. Tacit collusion in repeated auctions. *Journal of Economic Theory*, 114(1): 153–169, 2004.

Tan, J., Khalili, R., and Karl, H. Learning to bid long-term: Multi-agent reinforcement learning with long-term and sparse reward in repeated auction games. *arXiv preprint arXiv:2204.02268*, 2022.

Wang, Q., Yang, Z., Deng, X., and Kong, Y. Learning to bid in repeated first-price auctions with budgets. *arXiv preprint arXiv:2304.13477*, 2023.

Wen, C., Xu, M., Zhang, Z., Zheng, Z., Wang, Y., Liu, X., Rong, Y., Xie, D., Tan, X., Yu, C., et al. A cooperative-competitive multi-agent framework for auto-bidding in online advertising. In *Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining*, pp. 1129–1139, 2022.

Zhang, W., Yuan, S., Wang, J., and Shen, X. Real-time bidding benchmarking with iPinYou dataset. *arXiv preprint arXiv:1407.7073*, 2014.

Zhou, Z., Mertikopoulos, P., Moustakas, A. L., Bambos, N., and Glynn, P. W. Mirror descent learning in continuous games. In *56th IEEE Annual Conference on Decision and Control, CDC 2017, Melbourne, Australia, December 12-15, 2017*, pp. 5776–5783. IEEE, 2017. doi: 10.1109/CDC.2017.8264532. URL https://doi.org/10.1109/CDC.2017.8264532.

A. Individual Adaptive Pacing

For the completeness of our work, we present adaptive pacing (Balseiro & Gur, 2019), an optimal individual bidding algorithm, in Algorithm 3. We omit the subscript k to simplify the notation.

Algorithm 3 Individual Adaptive Pacing (IP)

1: **Input:** $\epsilon = 1/\sqrt{T}$, $\bar{\lambda} \geq \bar{v}/\rho$.
2: Select an initial multiplier $\lambda_1 \in [0, \bar{\lambda}]$, and set the remaining budget of agent k to $\widetilde{B}_1 = B = \rho T$.
3: **for** $t = 1$ **to** T **do**
4:     Observe the realization of v_t.
5:     Post a bid $b_t = \min\left\{ \frac{v_t}{1+\lambda_t}, \widetilde{B}_t \right\}$.
6:     Update the multiplier by $\lambda_{t+1} = P_{[0,\bar{\lambda}]}\left(\lambda_t - \epsilon(\rho - z_t) \right)$.
7:     Update the remaining budget by $\widetilde{B}_{t+1} = \widetilde{B}_t - z_t$.
8: **end for**

B. Missing Proofs in Section 3

B.1. Proof of Lemma 3.2

Lemma 3.2. *For any feasible coordinated strategy* C*, there exists a feasible bid rotation strategy* C^{P} *such that every member's expected utility under algorithm* C^{P} *is at least as good as that under algorithm* C*.*

Proof. Let the bidding profile produced by strategy C be $(\boldsymbol{b}_1, \boldsymbol{b}_2, \cdots, \boldsymbol{b}_T)$. Now randomly choose $k \in \arg\max \boldsymbol{b}_t$. Define new strategy C^P such that $b_{k,t}^{\mathrm{P}} = b_{k,t}$, and $b_{k',t}^{\mathrm{P}} = 0$ for all $k' \neq k$. This rule guarantees $\forall t \in [T]$, $\boldsymbol{b}_t^{\mathrm{P}}$ has at most one non-zero component, which means it is a bid rotation strategy.

For each round t, denote $I_t = \arg\max \boldsymbol{b}_t$. if $|I_t| = 1$, the bidder with the highest bid produced by C^{P} remains the same with that of C. The second-highest bid under C^{P} is no more than that under C. Therefore a coalition member will win the same item and gain the same value as in C, but pay no more than in C, leading to more utility. If $|I_t| > 1$, each bidder will pay at least the price as she does in C. This means C^{P} is feasible. Since the case $|I_t| > 1$ has measure zero, every member's expected utility under C^{P} is at least as good as that under C. $\square$

B.2. Details on the Choice of Step Sizes

In the algorithms presented in the main paper, we use $\epsilon = 1/\sqrt{T}$ as our step size. Generally, the results in Section 3 hold when ϵ satisfies the following conditions:

1. $\bar{v}\epsilon < 1$ and $2\gamma\epsilon < 1$;

2. $\lim_{T \to \infty} \epsilon = 0$ and $\lim_{T \to \infty} T\epsilon = \infty$.

One can check it when reading the following proofs.

B.3. Proof of Theorem 3.4

Theorem 3.4. *Suppose that Assumption 3.3 holds. When every bidder has the same value distribution and budget, we have*

$$\liminf_{T \to \infty} \frac{1}{T}(\Pi_k^{\mathrm{CP}} - U_k(\boldsymbol{\lambda}^*)) > 0, \quad \forall k \in \mathcal{K}. \tag{4}$$

To show that the performance of Algorithm 1 outperforms the benchmark in symmetric cases, we first prove that it converges to some equilibrium defined below. With the convergence results, we reduce the problem of comparing the asymptotic performance of our algorithms with players' benchmark utility, to that comparing the two equilibrium utilities.

Specifically, fix the pacing variables to be $\xi_1, \cdots, \xi_K$. Let $d_k := \max\left\{d_k^I, d^O\right\}$, where $d_k^I = \max_{i \in \mathcal{K}: i \neq k} \frac{v_i}{1+\xi_i}$. We can compute bidder k's expected expenditure per auction under Algorithm 1 to be

$$G_k^{\mathrm{CP}}(\boldsymbol{\xi}) := \mathbb{E}_{\boldsymbol{v}, d^O}\left[\mathbf{1}\left\{v_k \geq (1+\xi_k)d_k\right\} d^O\right],$$

and her corresponding expected utility function is

$$U_k^{\mathrm{CP}}(\boldsymbol{\xi}) := T\left(\mathbb{E}_{\boldsymbol{v}, d^O}\left[\mathbf{1}\left\{v_k \geq (1+\xi_k)d_k\right\}\left(v_k - d^O\right)\right]\right). \tag{9}$$

Since G_k^{CP} and U_k^{CP} are both decreasing in ξ_k, consider the equilibrium $\boldsymbol{\xi}^*$ defined by the following complementarity conditions:

$$\xi_k^* \geq 0 \perp G_k^{\mathrm{CP}}(\boldsymbol{\xi}^*) \leq \rho_k, \forall k \in \mathcal{K}. \tag{10}$$

Since $\boldsymbol{G}^{\mathrm{CP}}$ is strongly monotone, $\boldsymbol{\xi}^*$ exists and is unique.

First, with the help of the pacing strategy, we have the performance of Algorithm 1 converges to $U_k^{\mathrm{CP}}(\boldsymbol{\xi}^*)$, in arbitrary problem instances.

Lemma B.1. *Suppose that $\boldsymbol{G}^{\mathrm{CP}}$ is strongly monotone. We have*

$$\liminf_{T \to \infty} \frac{1}{T}(\Pi_k^{\mathrm{CP}} - U_k^{\mathrm{CP}}(\boldsymbol{\xi}^*)) \geq 0, \ \forall k \in \mathcal{K}. \tag{11}$$

Therefore, the problem reduces to the comparison of equilibrium utilities, $U_k(\boldsymbol{\lambda}^*)$ and $U_k^{\mathrm{CP}}(\boldsymbol{\xi}^*)$. In the symmetric cases, we first show that everyone's pacing parameter is lower compared to $\boldsymbol{\lambda}^*$. We can utilize this property, to show that the sets related to the indicator function in U_k^{CP} are supersets of that in U_k. By further showing that the utilities on those sets are no less in the former than the latter, we finished the comparison of two equilibria.

Specifically, to prove (4), by Lemma B.1, it suffices to show that

$$U_k^{\mathrm{CP}}(\boldsymbol{\xi}^*) > U_k(\boldsymbol{\lambda}^*). \tag{12}$$

By symmetry, for all $k \in \mathcal{K}$, we have

$$\lambda_k^* = \lambda_1^*, \text{ and } \xi_k^* = \xi_1^*,$$
$$G_k(\boldsymbol{\lambda}^*) = G_1(\boldsymbol{\lambda}^*) \text{ and } G_k^{\mathrm{CP}}(\boldsymbol{\xi}^*) = G_1^{\mathrm{CP}}(\boldsymbol{\xi}^*),$$
$$U_k(\boldsymbol{\lambda}^*) = U_1(\boldsymbol{\lambda}^*) \text{ and } U_k^{\mathrm{CP}}(\boldsymbol{\xi}^*) = U_1^{\mathrm{CP}}(\boldsymbol{\xi}^*).$$

Without loss of generality, we assume $T = 1$, then

$$U_k(\boldsymbol{\lambda}^*) = \mathbb{E}\left[\mathbf{1}\left\{(\forall i \in \mathcal{K} : i \neq k)\left(v_k \geq v_i\right) \wedge \left(v_k \geq (1+\lambda_1^*)d^O\right)\right\}\left(v_k - d_k\right)\right],$$
$$U_k^{\mathrm{CP}}(\boldsymbol{\xi}^*) = \mathbb{E}\left[\mathbf{1}\left\{(\forall i \in \mathcal{K} : i \neq k)\left(v_k \geq v_i\right) \wedge \left(v_k \geq (1+\xi_1^*)d^O\right)\right\}\left(v_k - d^O\right)\right].$$

We first show the following lemmas. Notice that in the symmetric setting, we can without loss of generality consider the case $\xi_k = \xi_1$ and write $G_k^{\mathrm{CP}}(\boldsymbol{\xi})$ as the function of ξ_1. The same holds for other functions as well.

Lemma B.2. *Suppose $\boldsymbol{G}^{\mathrm{CP}}(\boldsymbol{\xi})$ is strongly monotone. When every bidder has the same value distribution and budget, $G_k^{\mathrm{CP}}(\xi_1)$ is strictly decreasing on $[0, \bar{\xi}]$. Moreover, the bidders' expected value function $V_k^{\mathrm{CP}}(\xi_1)$ under Algorithm 1 is strictly decreasing on $[0, \bar{\xi}]$:*

$$V_k^{\mathrm{CP}}(\xi_1) := \mathbb{E}\left[\mathbf{1}\left\{(\forall i \in \mathcal{K} : i \neq k)\left(v_k \geq v_i\right) \wedge \left(v_k \geq (1+\xi_1)d^O\right)\right\} v_k\right].$$

Lemma B.3. $\xi_k^* < \bar{\xi}_k$ *for* $k \in \mathcal{K}$.

Lemma B.4. *For any* $\xi_1 \in [0, \bar{\xi})$, *we have*

$$G_k(\xi_1) > G_k^{\mathrm{CP}}(\xi_1).$$

13

Therefore when $\boldsymbol{\lambda}^* > 0$, we must have $\boldsymbol{\xi}^* < \boldsymbol{\lambda}^*$, and $G_k^{\mathrm{CP}}(\xi_1^*) \leq G_k(\lambda_1^*) = \rho_1$. Similarly, we can define bidder k's expected value function with respect to the benchmark, and it coincides with V_k^{CP}. Moreover, since G_k^{CP} and V_k^{CP} are strictly decreasing, we have

$$
\begin{aligned}
U_k(\lambda_1^*) &= V_k^{\mathrm{CP}}(\lambda_1^*) - G_k(\lambda_1^*) \\
&< V_k^{\mathrm{CP}}(\xi_1^*) - G_k^{\mathrm{CP}}(\xi_1^*) = U_k^{\mathrm{CP}}(\xi_1^*).
\end{aligned}
$$

When $\boldsymbol{\lambda}^* = 0$, we have $G_k^{\mathrm{CP}}(0) < G_k(0) \leq \rho_k$, which means $\boldsymbol{\xi}^* = 0$. Therefore

$$
U_k(\lambda_1^*) = U_k(0) = V_k^{\mathrm{CP}}(0) - G_k(0) < V_k^{\mathrm{CP}}(0) - G_k^{\mathrm{CP}}(0) = U_k(\xi_1^*).
$$

This completes the proof.

B.4. Missing Proofs of Lemmas in Theorem 3.4

In the following proofs of the lemmas, Denote the support of $\boldsymbol{F} := (F_1, \ldots, F_K, H)$ to be $\mathrm{Supp}(\boldsymbol{F})$. That is, for any $(\boldsymbol{v}, d^O) \in \mathrm{Supp}(\boldsymbol{F}) \subseteq \mathbb{R}^{K+1}$, $f_i(v_i) > 0$, for any $i \in \mathcal{K}$ and $h(d^O) > 0$. When we talk about sets of $(\boldsymbol{v}, d^O)$, we consider the subsets of $\mathrm{Supp}(\boldsymbol{F})$ without claiming explicitly. For any subsets of $\{v_i\}_{i=1}^K \cup \{d^O\}$, we denote their supports similarly. For example, $\mathrm{Supp}(F_i, F_j) := \{(\boldsymbol{v}, d^O) \in \mathbb{R}^{K+1} : f_i(v_i) > 0 \wedge f_j(v_j) > 0\}$.

Due to the convergence of the pacing strategies, the proof of Lemma B.1 is analogous to that in Balseiro & Gur (2019), which we shall omit here.

Lemma B.2. *Suppose $\boldsymbol{G}^{\mathrm{CP}}(\boldsymbol{\xi})$ is strongly monotone. When every bidder has the same value distribution and budget, $G_k^{\mathrm{CP}}(\xi_1)$ is strictly decreasing on $[0, \bar{\xi}]$. Moreover, the bidders' expected value function $V_k^{\mathrm{CP}}(\xi_1)$ under Algorithm 1 is strictly decreasing on $[0, \bar{\xi}]$:*

$$
V_k^{\mathrm{CP}}(\xi_1) := \mathbb{E}\left[\mathbf{1}\left\{(\forall i \in \mathcal{K} : i \neq k)\,(v_k \geq v_i) \wedge \left(v_k \geq (1+\xi_1)d^O\right)\right\} v_k\right].
$$

Proof. We first show that G_k^{CP} is strictly decreasing. For any $\xi_1 < \xi_1'$,

$$
\begin{aligned}
&(\xi_1 - \xi_1')\left(G_k^{\mathrm{CP}}(\xi_1) - G_k^{\mathrm{CP}}(\xi_1')\right) \\
&= (\xi_1 - \xi_1')\,\mathbb{E}\left[\mathbf{1}\left\{(\forall i \in \mathcal{K} : i \neq k)\,(v_k \geq v_i) \wedge \left((1+\xi_1')d^O > v_k \geq (1+\xi_1)d^O\right)\right\} d^O\right] \\
&= 1/K(\xi_1 \mathbf{1} - \xi_1' \mathbf{1})\left(\boldsymbol{G}^{\mathrm{CP}}(\xi_1 \mathbf{1}) - \boldsymbol{G}^{\mathrm{CP}}(\xi_1' \mathbf{1})\right) < 0.
\end{aligned}
$$

Moreover, we have

Claim B.5. For $\xi_1, \xi_1' \in [0, \bar{\xi}]$ such that $\xi_1 < \xi_1'$,

$$
\left\{(\forall i \in \mathcal{K} : i \neq k)\,(v_k \geq v_i) \wedge \left((1+\xi_1')d^O > v_k \geq (1+\xi_1)d^O\right)\right\}
$$

has non-zero measure.

Then

$$
V_k^{\mathrm{CP}}(\xi_1) - V_k^{\mathrm{CP}}(\xi_1') = \mathbb{E}\left[\mathbf{1}\left\{(\forall i \in \mathcal{K} : i \neq k)\,(v_k \geq v_i) \wedge \left((1+\xi_1')d^O > v_k \geq (1+\xi_1)d^O\right)\right\} v_k\right] > 0.
$$

This completes the proof. □

Lemma B.3. $\xi_k^* < \bar{\xi}_k$ *for* $k \in \mathcal{K}$.

Proof. $\xi_k^* < \bar{\xi}_k$ because

$$
G_k^{\mathrm{CP}}\left(\bar{\xi}_k, \boldsymbol{\xi}_{-k}^*\right) = \mathbb{E}\left[\mathbf{1}\left\{\frac{v_k}{1+\bar{\xi}_k} \geq d_k\right\} d^O\right] \overset{(a)}{\leq} \mathbb{E}\left[\mathbf{1}\left\{\frac{v_k}{1+\bar{v}_k/\rho_k} \geq d_k\right\} d^O\right] \leq \frac{v_k}{1+\bar{v}_k/\rho_k} < \rho_k,
$$

where (a) follows from $\bar{\xi}_k \geq \bar{v}_k/\rho_k$.

□

Lemma B.4. *For any $\xi_1 \in [0, \bar{\xi})$, we have*

$$G_k(\xi_1) > G_k^{\mathrm{CP}}(\xi_1).$$

Proof. Since

$$g(\xi_1) := G_k(\xi_1) - G_k^{\mathrm{CP}}(\xi_1) = \int_x \bar{F}(x) \int_0^{\frac{x}{1+\xi_1}} H(y) \, dy \, dL \geq 0,$$

where $\bar{F}(x) = 1 - F(x)$ and L is the cumulative probability function of $d_k^I := \max_{i \in \mathcal{K}: i \neq k} v_i$.

Since by Claim B.5 and the identical independence of bidders' distributions, we have, for any $\xi_1 \in [0, \bar{\xi})$,

$$\left\{ d_k^I \geq (1 + \xi_1) d^O \right\} \quad \text{and} \quad \left\{ v_k \geq d_k^I \right\}$$

have non-zero measure, and therefore $\mathbb{E}\left[\mathbf{1}\left\{ v_k \geq d_k^I \right\} \mathbf{1}\left\{ d_k^I \geq (1 + \xi_1) d^O \right\} d_k^I / (1 + \xi_1) \right] > 0$.

Note that for any $\xi_1 \in [0, \bar{\xi})$,

$$\frac{dg}{d\xi_1} = -\int_x \frac{x}{(1+\xi_1)^2} \bar{F}(x) H(\frac{x}{1+\xi_1}) \, dx \, dL = -\frac{1}{1+\xi_1} \mathbb{E}\left[\mathbf{1}\left\{ v_k \geq d_k^I \right\} \mathbf{1}\left\{ d_k^I \geq (1 + \xi_1) d^O \right\} d_k^I / (1 + \xi_1) \right] < 0.$$

We have $G_k(\xi_1) > G_k^{\mathrm{CP}}(\xi_1)$. This completes the proof.

$\square$

B.5. Other properties of CP in General Cases

In symmetric cases, $\boldsymbol{\xi}^* \leq \boldsymbol{\lambda}^*$. Here, we show this holds in general cases under mild conditions.

Lemma B.6. *If it holds that for any $k \in \mathcal{K}$,*

$$\left\{ \frac{v_k}{1 + \xi_k^*} \geq d^O \right\} \subsetneq \left\{ \frac{v_k}{1 + \lambda_k^*} \geq d^O \right\}.$$

then,

$$\boldsymbol{\lambda}^* \geq \boldsymbol{\xi}^*. \tag{13}$$

For example, the "if" condition holds when there exists $c > 0$, such that $H(d^O) \neq 0$, for $d^O \in (0, c)$. As H is the cdf of d^O and is an increasing function, this is equivalent to saying $H(d^O) > 0$ for all $d^O > 0$.

Proof. Writing down the specific formulation of G_k and G_k^{CP} respectively, we have

$$G_k(\boldsymbol{\lambda}) = \mathbb{E}\left[\mathbf{1}\left\{ \frac{v_k}{1 + \lambda_k} \geq \max\left\{ \max_{i \neq k} \frac{v_i}{1 + \lambda_i}, d^O \right\} \right\} \max\left\{ \max_{i \neq k} \frac{v_i}{1 + \lambda_i}, d^O \right\} \right]$$

$$= \mathbb{E}\left[\prod_{i \neq k} \mathbf{1}\left\{ \frac{v_k}{1 + \lambda_k} \geq \frac{v_i}{1 + \lambda_i} \right\} \mathbf{1}\left\{ \frac{v_k}{1 + \lambda_k} \geq d^O \right\} \max\left\{ \max_{i \neq k} \frac{v_i}{1 + \lambda_i}, d^O \right\} \right],$$

$$G_k^{\mathrm{CP}}(\boldsymbol{\xi}) = \mathbb{E}\left[\mathbf{1}\left\{ \frac{v_k}{1 + \xi_k} \geq \max\left\{ \max_{i \neq k} \frac{v_i}{1 + \xi_i}, d^O \right\} \right\} d^O \right]$$

$$= \mathbb{E}\left[\prod_{i \neq k} \mathbf{1}\left\{ \frac{v_k}{1 + \xi_k} \geq \frac{v_i}{1 + \xi_i} \right\} \mathbf{1}\left\{ \frac{v_k}{1 + \xi_k} \geq d^O \right\} d^O \right].$$

Now suppose for the sake of contradiction, $\xi_i^* > \lambda_i^*$ for $i \in \{1, \dots, s\}$, for some $s \leq K$, and $\xi_i^* \leq \lambda_i^*$ for $i > s$. Consider $k = 1$, and we have $G_1^{\mathrm{CP}}(\boldsymbol{\xi}^*) = \rho_1$. Since $\xi_i^* \leq \lambda_i^*$ for $i > s$, we have,

$$\left\{ \frac{v_1}{1 + \xi_1^*} \geq \frac{v_i}{1 + \xi_i^*} \right\} \subseteq \left\{ \frac{v_1}{1 + \lambda_1^*} \geq \frac{v_i}{1 + \lambda_i^*} \right\}, \forall i > s; \tag{14}$$

Since $\xi_1^* > \lambda_1^*$, we have,

$$\left\{ \frac{v_1}{1+\xi_1^*} \geq d^O \right\} \subsetneq \left\{ \frac{v_1}{1+\lambda_1^*} \geq d^O \right\}. \tag{15}$$

To make $G_1^{\mathrm{CP}}(\boldsymbol{\xi}_1^*) = \rho_1 \geq G_1(\boldsymbol{\lambda}_1^*)$, there must exist $\sigma(1) \in \{2, \ldots, s\}$, such that

$$\left\{ \frac{v_1}{1+\xi_1^*} \geq \frac{v_{\sigma(1)}}{1+\xi_{\sigma(1)}^*} \right\} \supseteq \left\{ \frac{v_1}{1+\lambda_1^*} \geq \frac{v_{\sigma(1)}}{1+\lambda_{\sigma(1)}^*} \right\},$$

since otherwise,

$$\left\{ \frac{v_1}{1+\xi_1^*} \geq \max\left\{ \max_{i\neq 1} \frac{v_i}{1+\xi_i^*}, d^O \right\} \right\} \subsetneq \left\{ \frac{v_1}{1+\lambda_1^*} \geq \max\left\{ \max_{i\neq 1} \frac{v_i}{1+\lambda_i^*}, d^O \right\} \right\}.$$

Since $\max\left\{ \max_{i\neq 1} \frac{v_i}{1+\lambda_i^*}, d^O \right\} \geq d^O$, we have $G_1(\boldsymbol{\lambda}^*) > G_1^{\mathrm{CP}}(\boldsymbol{\xi}^*)$, contradicting that $G_1^{\mathrm{CP}}(\boldsymbol{\xi}_1^*) \geq G_1(\boldsymbol{\lambda}_1^*)$.

This means that

$$\frac{1+\xi_1^*}{1+\xi_{\sigma(1)}^*} < \frac{1+\lambda_1^*}{1+\lambda_{\sigma(1)}^*}.$$

Similarly, for $\sigma(1)$, we have $\sigma^2(1) \in \{1, \ldots, s\}$, and moreover

$$\frac{1+\xi_{\sigma^{\ell-1}(1)}}{1+\xi_{\sigma^\ell(1)}} < \frac{1+\lambda_{\sigma^{\ell-1}(1)}}{1+\lambda_{\sigma^\ell(1)}}, \quad \text{for } \ell = 0, 1, \ldots \text{ and } \sigma^\ell(1) \in \{1, \ldots, s\}$$

Since $s < \infty$, there must exist ℓ_0, such that $1 = \sigma^{\ell_0}(1)$. Let us multiply the above inequalities from 0 to ℓ_0, we have

$$1 = \prod_{\ell=0}^{\ell_0} \frac{1+\xi_{\sigma^{\ell-1}(1)}}{1+\xi_{\sigma^\ell(1)}} < \prod_{\ell=0}^{\ell_0} \frac{1+\lambda_{\sigma^{\ell-1}(1)}}{1+\lambda_{\sigma^\ell(1)}} = 1,$$

leading to a contradiction.

$\square$

C. Omitted Proofs in Section 4

C.1. Details on the Choice of Step Sizes

In the algorithms presented in the main paper, we use $\epsilon = 1/\sqrt{T}$ as our step size. Generally, the results in Section 4 hold when ϵ satisfies the following conditions:

- $\bar{v}\epsilon < 1, 2\gamma\epsilon < 1$ and $\underline{G}'\epsilon < 1$;

- $\lim_{T\to\infty} \epsilon = 0$ and $\lim_{T\to\infty} T\epsilon^{3/2} = \infty$.

One can check it when reading the following proofs.

C.2. Proof of Theorem 4.2

Theorem 4.2. *Suppose that Assumption 4.1 holds. We have*

$$\liminf_{T\to\infty} \frac{1}{T}(\Pi_k^{\mathrm{HP}} - U_k(\boldsymbol{\lambda}^*)) \geq 0, \forall k \in \mathcal{K}, \tag{6}$$

and the equality holds for at most one bidder.

First, we show Line 11 in Algorithm 2 is equivalent to (7).

$$
\begin{aligned}
z'_{k,t} &= \mathbf{1}\left\{b^I_{k,t} \geq z_{k,t}\right\} \max\left(z_{k,t}, x_{k,t} d^I_{k,t}\right) \\
&= \mathbf{1}\left\{b^I_{k,t} \geq x_{k,t} d^O_t\right\} \max\left(x_{k,t} d^O_t, x_{k,t} d^I_{k,t}\right) \\
&= \mathbf{1}\left\{b^I_{k,t} \geq x_{k,t} d^O_t\right\} x_{k,t} d_{k,t}.
\end{aligned}
$$

When $x_{k,t} = 1$, as bidder k wins both the internal election and the real auction, we have $b^O_{k,t} \geq d^O_t$ and $b^I_{k,t} \geq d^I_{k,t}$. Hence,

$$
z'_{k,t} = \mathbf{1}\left\{b^I_{k,t} \geq d^O_t\right\} d_{k,t} = \mathbf{1}\left\{b^I_{k,t} \geq d_{k,t}\right\} d_{k,t}.
$$

When $x_{k,t} = 0$, either $b^O_{k,t} < d^O_t$ or $b^I_{k,t} < d^I_{k,t}$ holds. If $b^I_{k,t} \geq d^I_{k,t}$, we have $b^I_{k,t} \leq b^O_{k,t} < d^O_t \leq d_{k,t}$, so in either case

$$
z'_{k,t} = 0 = \mathbf{1}\left\{b^I_{k,t} \geq d_{k,t}\right\} d_{k,t}.
$$

As the update of λ_k coincides with the subgradient descent scheme of an individual adaptive pacing strategy, we can apply the results in Balseiro & Gur (2019) to characterize the convergence of the sequence of pseudo multipliers. In what follows we use the standard norm notation $\|\boldsymbol{y}\|_p := \left(\sum_{k=1}^K |y_k|^p\right)^{1/p}$ for a vector $\boldsymbol{y} \in \mathbb{R}^K$.

Lemma C.1 (Balseiro & Gur (2019)). *Suppose that Assumption 2.1 holds. There exists a unique $\boldsymbol{\lambda}^*$ that satisfies (2) and*

$$
\lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^T \mathbb{E}\left[\|\boldsymbol{\lambda}_t - \boldsymbol{\lambda}^*\|_1\right] = 0.
$$

The following Lemma C.2 establishes a similar convergence result for multipliers $\boldsymbol{\mu}$.

Lemma C.2. *For each k, there exists a unique $\mu^*_k \leq \lambda^*_k$ that satisfies (8) and for each $k \in \mathcal{K}$,*

$$
\lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^T \mathbb{E}\left[|\mu_{k,t} - \mu^*_k|\right] = 0.
$$

Lemma C.1 and Lemma C.2 ensure the existence and uniqueness of $(\boldsymbol{\mu}^*, \boldsymbol{\lambda}^*)$, so that we can bound the expected payoffs with a well-defined $U^{\mathrm{HP}}_k(\mu^*_k, \boldsymbol{\lambda}^*_k)$.

Lemma C.3.

$$
\liminf_{\substack{T \to \infty \\ B_k = \rho_k T}} \frac{1}{T}(\Pi^{\mathrm{HP}}_k - U^{\mathrm{HP}}_k(\mu^*_k, \boldsymbol{\lambda}^*)) \geq 0. \tag{16}
$$

Lemma C.3 establishes that the expected performance of Algorithm 2 can be asymptotically lower bounded by $U^{\mathrm{HP}}_k(\mu^*_k, \boldsymbol{\lambda}^*_k)$. The proof of Lemma C.3 consists of two parts. We first use the fact that budgets are not depleted too early as shown in the proof of Lemma C.2. Then we use that the expected utility functions are Lipschitz continuous as argued in Corollary E.3.

With Lemma C.3 in hand, it only remains to show that $U^{\mathrm{HP}}_k(\mu^*_k, \boldsymbol{\lambda}^*) \geq U_k(\boldsymbol{\lambda}^*)$. Since $\lambda^*_k \geq \mu^*_k \geq 0$ by Lemma C.2, we first have

$$
\begin{aligned}
U^{\mathrm{HP}}_k(\mu^*_k, \boldsymbol{\lambda}^*) &:= T \cdot \mathbb{E}_{\boldsymbol{v}, d^O}\left[\mathbf{1}\left\{\frac{v_k}{1+\lambda^*_k} \geq d^I_k \wedge \frac{v_k}{1+\mu^*_k} \geq d^O\right\}(v_k - d^O)\right] \\
&\geq T \cdot \mathbb{E}_{\boldsymbol{v}, d^O}\left[\mathbf{1}\left\{\frac{v_k}{1+\lambda^*_k} \geq d^I_k \wedge \frac{v_k}{1+\lambda^*_k} \geq d^O\right\}(v_k - d^O)\right] \\
&= U^{\mathrm{HP}}_k(\lambda^*_k, \boldsymbol{\lambda}^*).
\end{aligned}
$$

Then we decompose $U^{\mathrm{HP}}_k(\lambda^*_k, \boldsymbol{\lambda}^*)$ into two parts, $V^{\mathrm{HP}}_k(\lambda^*_k, \boldsymbol{\lambda}^*)$ and $G^{\mathrm{HP}}_k(\lambda^*_k, \boldsymbol{\lambda}^*)$, where $V^{\mathrm{HP}}_k(\mu_k, \boldsymbol{\lambda})$ is defined by

$$
V^{\mathrm{HP}}_k(\mu_k, \boldsymbol{\lambda}) = \mathbb{E}_{\boldsymbol{v}, d^O}\left[\mathbf{1}\left\{\frac{v_k}{1+\lambda_k} \geq d^I_k \wedge \frac{v_k}{1+\mu_k} \geq d^O\right\} v_k\right], \tag{17}
$$

and $G_k^{\mathrm{HP}}(\mu_k, \boldsymbol{\lambda})$ defined by (5). We also decompose the benchmark $U_k(\boldsymbol{\lambda}^*)$ into $V_k(\boldsymbol{\lambda}^*)$ and $G_k(\boldsymbol{\lambda}^*)$, where $V_k(\boldsymbol{\lambda})$ is defined by

$$V_k(\boldsymbol{\lambda}) = \mathbb{E}_{\boldsymbol{v},d^O} \left[\mathbf{1}\left\{ \frac{v_k}{1+\lambda_k} \geq d_k \right\} v_k \right], \tag{18}$$

and $G_k(\boldsymbol{\lambda})$ is defined by (1). Observe that $V_k^{\mathrm{HP}}(\lambda_k^*, \boldsymbol{\lambda}^*) = V_k(\boldsymbol{\lambda}^*)$.

Therefore, we have

$$
\begin{aligned}
U_k^{\mathrm{HP}}(\mu_k^*, \boldsymbol{\lambda}^*) \geq U_k^{\mathrm{HP}}(\lambda_k^*, \boldsymbol{\lambda}^*) &= T \cdot V_k^{\mathrm{HP}}(\lambda_k^*, \boldsymbol{\lambda}^*) - T \cdot G_k^{\mathrm{HP}}(\lambda_k^*, \boldsymbol{\lambda}^*) \\
&= T \cdot V_k(\boldsymbol{\lambda}^*) - T \cdot G_k^{\mathrm{HP}}(\lambda_k^*, \boldsymbol{\lambda}^*) \\
&\overset{(a)}{\geq} T \cdot V_k(\boldsymbol{\lambda}^*) - T \cdot G_k(\boldsymbol{\lambda}^*) = U_k(\boldsymbol{\lambda}^*),
\end{aligned}
$$

where (a) follows from Lemma E.4, which also indicates that the equality strictly holds with at most one bidder.

We defer the proofs of key lemmas to the next subsection. It is worth mentioning that the proof techniques differ from those of the similar results in previous work in the following ways: (1) First, the expected utility and expenditure functions are more complicated as they involve two multiplier vectors. (2) Second, Algorithm 2 imposes Euclidean projection with different intervals on μ_k in different rounds, i.e., $[0, \lambda^{k,t+1}]$ on $\mu_{k,t+1}$, so it is subtle to deal with the positive projection errors.

C.3. Proof of Lemmas in Theorem 4.2

Lemma C.2. *For each k, there exists a unique $\mu_k^* \leq \lambda_k^*$ that satisfies (8) and for each $k \in \mathcal{K}$,*

$$\lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}\left[|\mu_{k,t} - \mu_k^*| \right] = 0.$$

Proof. **Step 1: Existence and uniqueness of $\boldsymbol{\mu}^*$.** By Lemma C.1, $\boldsymbol{\lambda}^*$ defined by (2) exists and is unique. By Lemma E.4, we have

$$G_k^{\mathrm{HP}}(\lambda_k^*, \boldsymbol{\lambda}^*) \leq G_k(\boldsymbol{\lambda}^*). \tag{19}$$

As $G_k^{\mathrm{HP}}(\mu_k, \boldsymbol{\lambda})$ is continuous and strictly decreases in μ_k by Assumption 4.1 and $G_k(\boldsymbol{\lambda}^*) \leq \rho_k$ by the definition of $\boldsymbol{\lambda}^*$, we can find a unique $\mu_k^* \in [0, \lambda_k^*]$ that satisfies $\mu_k^* \geq 0 \perp G_k^{\mathrm{HP}}(\mu_k^*, \boldsymbol{\lambda}^*) \leq \rho_k$.

Step 2: bounding the ending time. Let $\tau_k = \sup\{t \leq T : B_{k,t} \geq \bar{v}_k\}$ be the last period in which bidder k's remaining budget is larger than her maximum value, and let $\tau = \min_k \tau_k$. We denote by $P_{k,t}$ the projection error introduced by Line 14 of Algorithm 2:

$$P_{k,t} := \mu_{k,t} - \epsilon(\rho_k - z_{k,t}) - P_{[0, \lambda_{k,t+1}]}\left(\mu_{k,t} - \epsilon(\rho_k - z_{k,t})\right),$$

and then $\mu_{k,t+1} = \mu_{k,t} - \epsilon(\rho_k - z_{k,t}) - P_{k,t}$. Summing up from $t = 1$ to τ_k, we have

$$\sum_{t=1}^{\tau_k} (z_{k,t} - \rho_k) = \frac{\mu_{k,\tau_k+1} - \mu_{k,1}}{\epsilon} + \frac{1}{\epsilon} \sum_{t=1}^{\tau_k} P_{k,t} \leq \frac{\mu_{k,\tau_k+1}}{\epsilon} + \frac{1}{\epsilon} \sum_{t=1}^{\tau_k} P_{k,t}^+.$$

When $\tau_k < T$, bidder k hits her budget constraint in round $\tau_k + 1$, so we should have $\sum_{t=1}^{\tau_k} z_{k,t} \geq \rho_k T - \bar{v}_k$. Together with $\mu_{k,t} \leq \lambda_{k,t} \leq \bar{\mu}_k$, we obtain

$$T - \tau_k \leq \frac{\bar{v}_k}{\rho_k} + \frac{\bar{\mu}_k}{\epsilon \rho_k} + \frac{1}{\epsilon \rho_k} \sum_{t=1}^{T} P_{k,t}^+. \tag{20}$$

Note that inequality (20) also trivially holds when $\tau_k = T$ since all the terms on the right hand side are positive.

Now we bound $P_{k,t}^+$ on the right hand side of (20). We first show that there is no positive projection error introduced by Line 11 of Algorithm 2 whenever $\epsilon \bar{v}_k \leq 1$.

$$\lambda_{k,t} - \epsilon \left(\rho - z'_{k,t} \right) \overset{(a)}{\leq} \lambda_{k,t} + \epsilon \left(b_{k,t}^I - \rho_k \right) \overset{(b)}{\leq} \lambda_{k,t} + \frac{\epsilon \bar{v}_k}{1 + \lambda_{k,t}} - \epsilon \rho_k$$

$$\overset{(c)}{\leq} \bar{\mu}_k + \frac{\epsilon \bar{v}_k}{1 + \bar{\mu}_k} - \epsilon \rho_k \overset{(d)}{\leq} \bar{\mu}_k,$$

where (a) follows from $z'_{k,t} = \mathbf{1}\left\{ b_{k,t}^I \geq d_{k,t} \right\} d_{k,t}$; (b) holds by $b_{k,t}^I \leq v_k / (1 + \lambda_{k,t}) \leq \bar{v}_k / (1 + \lambda_{k,t})$; (c) holds since $\lambda_{k,t} + \epsilon \bar{v}_k / (1 + \lambda_{k,t})$ increases in $\lambda_{k,t}$ whenever $\epsilon \bar{v}_k \leq 1$; and (d) holds since $\bar{\mu}_k \geq \bar{v}_k / \rho_k$. As a result, $\lambda_{k,t+1} \geq \lambda_{k,t} - \epsilon(\rho_k - z'_{k,t})$, and we have

$$P_{k,t}^+ = \left(\mu_{k,t} - \epsilon(\rho_k - z_{k,t}) - \lambda_{k,t+1} \right)^+$$

$$\leq \left(\mu_{k,t} - \lambda_{k,t} + \epsilon \left(z_{k,t} - z'_{k,t} \right) \right)^+ .$$

Since $\mu_{k,t} \leq \lambda_{k,t}$, $P_{k,t}^+ > 0$ is only possible when $d_t^O = z_{k,t} > z'_{k,t} = 0$. In this case, bidder k is the final winner with $b_{k,t}^O \geq d_t^O$ but she could not have won if using $b_{k,t}^I < d_t^O$ as her final bid. We have

$$P_{k,t}^+ \leq \left(\mu_{k,t} - \lambda_{k,t} + \epsilon d_t^O \right)^+ \leq \epsilon d_t^O \leq \epsilon b_{k,t}^O = \frac{\epsilon v_{k,t}}{1 + \mu_{k,t}} \leq \epsilon \bar{v}_k, \tag{21}$$

which also implies that a positive error requires $\lambda_{k,t} - \mu_{k,t} < \epsilon d_t^O$. Then we can bound the probability as follows,

$$\mathbb{P}\left[P_{k,t} > 0 \right] \leq \mathbb{P}\left[d_t^O \in \left(b_{k,t}^I, b_{k,t}^O \right] \right] \leq \mathbb{P}\left[d_t^O \in \left(\frac{v_{k,t}}{1 + \mu_{k,t} + \epsilon d_t^O}, \frac{v_{k,t}}{1 + \mu_{k,t}} \right] \right],$$

$$\leq \frac{\epsilon d_t^O v_{k,t}}{(1 + \mu_{k,t})(1 + \mu_{k,t} + \epsilon d_t^O)} \overline{h} \leq \epsilon \bar{v}_k^2 \overline{h}, \tag{22}$$

where we denote by $\overline{h}$ by the upper bound on the density funciton of d_t^O.

Combining (21) and (22), we have

$$\mathbb{E}_{d_t^O}\left[P_{k,t}^+ \mid \mathcal{H}_{k,t} \right] \leq \epsilon^2 \bar{v}_k^3 \overline{h}. \tag{23}$$

Taking expectations over (20) we have

$$\mathbb{E}\left[T - \tau_k \right] \leq \frac{\bar{v}_k}{\rho_k} + \frac{\bar{\mu}_k}{\epsilon \rho_k} + \frac{T \epsilon \bar{v}_k^3 \overline{h}}{\rho_k},$$

and thus

$$\mathbb{E}\left[T - \tau \right] \leq \frac{\bar{v}}{\underline{\rho}} + \frac{\bar{\mu}}{\epsilon \underline{\rho}} + \frac{T \epsilon \bar{v}^3 \overline{h}}{\underline{\rho}}, \tag{24}$$

where $\bar{v} = \max_k \bar{v}_k$, $\underline{\rho} = \min_k \rho_k$ and $\bar{\mu} = \max_k \bar{\mu}_k$.

Step 3: bounding the mean absolute errors. We consider the update formula of $\mu_{k,t}$ to prove that the sequence converges to μ_k^*.

$$(\mu_{k,t+1} - \mu_k^*)^2 \overset{(a)}{\leq} \left(\mu_{k,t} - \epsilon(\rho_k - z_{k,t}) - P_{k,t}^+ - \mu_k^* \right)^2$$

$$= (\mu_{k,t} - \mu_k^*)^2 - 2\epsilon(\rho_k - z_{k,t})(\mu_{k,t} - \mu_k^*) + \epsilon^2(\rho_k - z_{k,t})^2$$

$$+ (P_{k,t}^+)^2 - 2P_{k,t}^+ \left(\mu_{k,t} - \epsilon(\rho_k - z_{k,t}) - \mu_k^* \right)$$

$$\overset{(b)}{\leq} (\mu_{k,t} - \mu_k^*)^2 - 2\epsilon(\rho_k - z_{k,t})(\mu_{k,t} - \mu_k^*) + \epsilon^2 \bar{v}_k^2 + (P_{k,t}^+)^2 + 2P_{k,t}^+ \bar{\mu}_k, \tag{25}$$

where (a) follows from a standard contraction property of the Euclidean projection operator and (b) is due to $(\rho_k - z_{k,t})^2 \leq \bar{v}_k^2$ and $\mu_{k,t} - \epsilon(\rho_k - z_{k,t}) > \lambda_{k,t+1} \geq 0$ when $P_{k,t}^+ > 0$.

19

Define $\delta_{k,t} = \mathbb{E}\left[(\mu_{k,t} - \mu_k^*)^2 \mathbf{1}\{t \leq \tau\}\right]$. Taking expectations on (25) we obtain

$$\delta_{k,t+1} \overset{(a)}{\leq} \delta_{k,t} - 2\epsilon\mathbb{E}\left[(\rho_k - G_k^{\mathrm{HP}}(\mu_{k,t}, \boldsymbol{\lambda}_t))(\mu_{k,t} - \mu_k^*)\right] + \epsilon^2\bar{v}_k^2 + \mathbb{E}\left[(P_{k,t}^+)^2\right] + 2\bar{\mu}_k\mathbb{E}\left[P_{k,t}^+\right]$$

$$\overset{(b)}{\leq} \delta_{k,t} + 2\epsilon\mathbb{E}\left[(G_k^{\mathrm{HP}}(\mu_{k,t}, \boldsymbol{\lambda}_t) - \rho_k)(\mu_{k,t} - \mu_k^*)\right] + \epsilon^2\underbrace{\left(\bar{v}_k^2 + \epsilon\bar{v}_k^4\bar{h} + 2\bar{\mu}_k\bar{v}_k^3\bar{h}\right)}_{R} \tag{26}$$

where (a) holds since $\mathbf{1}\{t \leq \tau\}$ monotonically decreases with t and the expectation of $z_{k,t}$ is $G_k(\mu_{k,t}, \boldsymbol{\lambda}_t)$ conditioned on $\mu_{k,t}$ and $\boldsymbol{\lambda}_t$; (b) holds by inequalities (21) and (22). For the second term in (26), we have

$$\left(G_k^{\mathrm{HP}}(\mu_{k,t}, \boldsymbol{\lambda}_t) - \rho_k\right)(\mu_{k,t} - \mu_k^*) = \left(G_k^{\mathrm{HP}}(\mu_{k,t}, \boldsymbol{\lambda}_t) - G_k^{\mathrm{HP}}(\mu_k^*, \boldsymbol{\lambda}^*) + G_k^{\mathrm{HP}}(\mu_k^*, \boldsymbol{\lambda}^*) - \rho_k\right)(\mu_{k,t} - \mu_k^*)$$

$$\overset{(a)}{\leq} \left(G_k^{\mathrm{HP}}(\mu_{k,t}, \boldsymbol{\lambda}_t) - G_k^{\mathrm{HP}}(\mu_k^*, \boldsymbol{\lambda}^*)\right)(\mu_{k,t} - \mu_k^*)$$

$$\overset{(b)}{\leq} -\underline{G}'\left(\mu_{k,t} - \mu_k^*\right)^2 + \bar{v}_k^2\bar{f}|\mu_{k,t} - \mu_k^*|\|\boldsymbol{\lambda}_t - \boldsymbol{\lambda}^*\|_1$$

$$\overset{(c)}{\leq} -\underline{G}'\left(\mu_{k,t} - \mu_k^*\right)^2 + \frac{\underline{G}'}{2}\left(\mu_{k,t} - \mu_k^*\right)^2 + \frac{\bar{v}_k^4\bar{f}^2}{2\underline{G}'}\|\boldsymbol{\lambda}_t - \boldsymbol{\lambda}^*\|_2^2$$

$$= -\frac{\underline{G}'}{2}(\mu_{k,t} - \mu_k^*)^2 + \frac{\bar{v}_k^4\bar{f}^2}{2\underline{G}'}\|\boldsymbol{\lambda}_t - \boldsymbol{\lambda}^*\|_2^2, \tag{27}$$

where (a) holds by $\rho_k - G_k^{\mathrm{HP}}(\mu_k^*, \boldsymbol{\lambda}^*) \geq 0$, $\mu_{k,t} \geq 0$ and $\mu_k^*(\rho_k - G_k^{\mathrm{HP}}(\mu_k^*, \boldsymbol{\lambda}^*)) = 0$; (b) follows from Lemma E.2 and $\bar{f}$ denotes the upper bound on f_k; (c) holds by AM-GM inequality. Together with (26) one has

$$\delta_{k,t+1} \leq \left(1 - \epsilon\underline{G}'\right)\delta_{k,t} + \frac{\epsilon\bar{v}_k^4\bar{f}^2}{\underline{G}'}\eta_t + \epsilon^2 R,$$

where $\eta_t = \mathbb{E}\left[\|\boldsymbol{\lambda}_t - \boldsymbol{\lambda}^*\|_2^2\mathbf{1}\{t \leq \tau\}\right]$. The recursion gives

$$\delta_{k,t} \leq \left(1 - \epsilon\underline{G}'\right)^{t-1}\bar{\mu}_k^2 + \sum_{s=1}^{t-1}\left(1 - \epsilon\underline{G}'\right)^{t-1-s}\left(\frac{\epsilon\bar{v}_k^4\bar{f}^2}{\underline{G}'}\eta_s + \epsilon^2 R\right).$$

By the result of Balseiro & Gur (2019), we have

$$\eta_t \leq K\bar{\mu}_k^2\left(1 - 2\gamma\epsilon\right)^{t-1} + \frac{K\bar{v}_k^2}{2\gamma}\epsilon.$$

Thus we have

$$\delta_{k,t} \leq \left(1 - \epsilon\underline{G}'\right)^{t-1}\bar{\mu}_k^2 + \frac{\epsilon K\bar{\mu}_k^2\bar{v}_k^4\bar{f}^2}{\underline{G}'}\sum_{s=1}^{t-1}\left(1 - \epsilon\underline{G}'\right)^{t-1-s}\left(1 - 2\gamma\epsilon\right)^{s-1}$$

$$+ \left(\frac{K\bar{v}_k^6\bar{f}^2}{2\gamma\underline{G}'} + R\right)\epsilon^2\sum_{s=1}^{t-1}\left(1 - \epsilon\underline{G}'\right)^{t-1-s}$$

$$\leq \left(1 - \epsilon R_1\right)^{t-1}\bar{\mu}_k^2 + \epsilon R_2\sum_{s=1}^{t-1}\left(1 - \epsilon R_1\right)^{t-2} + R_3\epsilon^2\sum_{s=0}^{t-2}\left(1 - \epsilon R_1\right)^s$$

$$\leq \left(1 - \epsilon R_1\right)^{t-1}\bar{\mu}_k^2 + \epsilon R_2(t-1)\left(1 - \epsilon R_1\right)^{t-2} + \epsilon\frac{R_3}{R_1},$$

where we denote $\min\{\underline{G}', 2\gamma\}$ by R_1, $\frac{K\bar{\mu}_k^2\bar{v}_k^4\bar{f}^2}{\underline{G}'}$ by R_2 and $\left(\frac{K\bar{v}_k^6\bar{f}^2}{2\gamma\underline{G}'} + R\right)$ by R_3.

Consider the sum of square roots of $\delta_{k,t}$,

$$\sum_{t=1}^{T} \delta_{k,t}^{1/2} \overset{(a)}{\leq} \sum_{t=1}^{T} (1 - \epsilon R_1)^{(t-1)/2} \bar{\mu}_k + \sum_{t=1}^{T} \epsilon^{1/2} \sqrt{\frac{R_3}{R_1}} + \sum_{t=1}^{T} \epsilon^{1/2} \sqrt{R_2(t-1)} (1 - \epsilon R_1)^{(t-2)/2}$$

$$\overset{(b)}{\leq} \frac{2\bar{\mu}_k}{\epsilon R_1} + T\epsilon^{1/2} \sqrt{\frac{R_3}{R_1}} + \epsilon^{1/2} \sqrt{R_2} \sum_{t=1}^{T-1} t (1 - \epsilon R_1)^{(t-1)/2}$$

$$\overset{(c)}{\leq} \frac{2\bar{\mu}_k}{\epsilon R_1} + T\epsilon^{1/2} \sqrt{\frac{R_3}{R_1}} + \frac{4\sqrt{R_2}}{\epsilon^{3/2} R_1^2}, \tag{28}$$

where (a) holds since $\sqrt{x + y} \leq \sqrt{x} + \sqrt{y}$ for $x, y \geq 0$, (b) follows from $\sum_{t=1}^{T} (1 - \epsilon R_1)^{(t-1)/2} \leq \sum_{t=0}^{\infty} (1 - \epsilon R_1)^{t/2} = \frac{1}{1-\sqrt{1-\epsilon R_1}} \leq \frac{2}{\epsilon R_1}$, and (c) follows from $\sum_{t=1}^{T-1} t (1 - \epsilon R_1)^{(t-1)/2} \leq \frac{1}{(1-\sqrt{1-\epsilon R_1})^2} \leq \frac{4}{\epsilon^2 R_1^2}$.

Combining (24) and (28), we obtain

$$\sum_{t=1}^{T} \mathbb{E}\left[|\mu_{k,t} - \mu_k^*|\right] = \sum_{t=1}^{T} \mathbb{E}\left[|\mu_{k,t} - \mu_k^*|\mathbf{1}\{t \leq \tau\} + |\mu_{k,t} - \mu_k^*|\mathbf{1}\{t > \tau\}\right]$$

$$\overset{(a)}{\leq} \sum_{t=1}^{T} \delta_{k,t}^{1/2} + \bar{\mu}_k \mathbb{E}[T - \tau]$$

$$\leq \frac{2\bar{\mu}}{\epsilon R_1} + T\epsilon^{1/2} \sqrt{\frac{R_3}{R_1}} + \frac{4\sqrt{R_2}}{\epsilon^{3/2} R_1^2} + \frac{\bar{v}\bar{\mu}}{\rho} + \frac{\bar{\mu}^2}{\epsilon\rho} + \frac{T\epsilon\bar{v}^3\bar{\mu}\bar{h}}{\rho},$$

where (a) holds by $\mathbb{E}[|y|] \leq \sqrt{\mathbb{E}[y^2]}$. Dividing by T, the result then follows because each term goes to zero when $\epsilon \sim T^{-1/2}$. $\qquad \square$

Lemma C.3.

$$\liminf_{\substack{T \to \infty \\ B_k = \rho_k T}} \frac{1}{T} (\Pi_k^{\mathrm{HP}} - U_k^{\mathrm{HP}}(\mu_k^*, \boldsymbol{\lambda}^*)) \geq 0. \tag{16}$$

Proof. In the proof we consider an alternate framework in which bidders ignore their budgets, i.e., each bidder k can post $b_{k,t}^I = \frac{v_{k,t}}{1+\lambda_{k,t}}, b_{k,t}^O = \frac{v_{k,t}}{1+\mu_{k,t}}$ and make payments even after depleting her budget. Let $\tau = \sup\{t \leq T : \forall k, B_{k,t} \geq \bar{v}_k\}$ be the last period in which the remaining budget of every bidder is larger than her maximum value. The performance of both the original and the alternate frameworks coincide until time τ_k. Let $u_{k,t}$ be bidder k's payoff at round t in the alternate framework. We have

$$\Pi_k^{\mathrm{HP}} \overset{(a)}{\geq} \sum_{t=1}^{\tau} \mathbb{E}[u_{k,t}] \overset{(b)}{\geq} \sum_{t=1}^{T} \mathbb{E}[u_{k,t}] - \bar{v}_k \mathbb{E}[T - \tau], \tag{29}$$

where (a) holds by discarding all auctions after τ and (b) holds since $u_{k,t} \leq \bar{v}_k$.

We next give a lower bound on the expected utility per auction.

$$\mathbb{E}[u_{k,t}] \overset{(a)}{=} \mathbb{E}\left[\frac{1}{T} U_k^{\mathrm{HP}}(\mu_{k,t}, \boldsymbol{\lambda}_t)\right]$$

$$\overset{(b)}{\geq} \frac{1}{T} U_k^{\mathrm{HP}}(\mu_k^*, \boldsymbol{\lambda}^*) - K\bar{v}_k^2 \bar{f} \mathbb{E}[\|\boldsymbol{\lambda}_t - \boldsymbol{\lambda}^*\|_1] - \bar{v}_k^2 (\bar{h} + \bar{f}) \mathbb{E}[|\mu_{k,t} - \mu_k^*|],$$

where (a) holds by using the independence of $\boldsymbol{v}_t$ and d_t^O from the multipliers and taking conditional expectations; and (b) holds by the Lipschitz continuity of U_k^{HP}/T from Corollary E.3. Summing over $t = 1, \ldots, T$, we obtain

$$\sum_{t=1}^{T} \mathbb{E}[u_{k,t}] \geq U_k^{\mathrm{HP}}(\mu_k^*, \boldsymbol{\lambda}^*) - \underbrace{K\bar{v}_k^2 \bar{f}}_{C_1} \sum_{t=1}^{T} \mathbb{E}[\|\boldsymbol{\lambda}_t - \boldsymbol{\lambda}^*\|_1] - \underbrace{\bar{v}_k^2 (\bar{h} + \bar{f})}_{C_2} \sum_{t=1}^{T} \mathbb{E}[|\mu_{k,t} - \mu_k^*|], \tag{30}$$

By the result from Balseiro & Gur (2019) and Equation (28) in Lemma C.2, there exist positive constants C_3 through C_7 such that

$$\sum_{t=1}^{T} \mathbb{E}\left[\|\boldsymbol{\lambda}_t - \boldsymbol{\lambda}^*\|_1\right] \leq \frac{C_3}{\epsilon} + C_4\epsilon^{1/2}T. \tag{31}$$

$$\sum_{t=1}^{T} \mathbb{E}\left[|\mu_{k,t} - \mu_k^*|\right] \leq \frac{C_5}{\epsilon} + C_6\epsilon^{1/2}T + \frac{C_7}{\epsilon^{3/2}}. \tag{32}$$

Note that the sequences of multipliers in (31) and (32) are those in the alternate framework, so we do not involve $\mathbf{1}\{t < \tau\}$ as we did in Lemma C.2.

Putting (29), (30), (31), (32) and (24) together, we obtain

$$\Pi_k^{\mathrm{HP}} \geq U_k^{\mathrm{HP}}(\mu_k^*, \boldsymbol{\lambda}^*) - \frac{C_1C_3 + C_2C_5}{\epsilon} - (C_1C_4 + C_2C_6)\epsilon^{1/2}T - \frac{C_2C_7}{\epsilon^{3/2}} - \frac{\bar{v}^2}{\rho} - \frac{\bar{v}\bar{\mu}}{\epsilon\rho} - \frac{T\epsilon\bar{v}^4\bar{h}}{\rho},$$

which concludes the proof. In particular, a selection of step size of order $\epsilon \sim T^{-1/2}$ guarantees a long-run average convergence rate of order $T^{-1/4}$. □

D. Omitted Proofs in Section 5

D.1. Proof of Theorem 5.1

Theorem 5.1. *When every bidder in the coalition has the same value distribution and budget, CP and HP maximize the total expected utility of the coalition asymptotically,*

$$\liminf_{T\to\infty} \frac{1}{T}\left(\sum_{k\in\mathcal{K}} \Pi^{\mathrm{CP/HP}} - \mathbb{E}[\pi^{\mathrm{H}}]\right) \geq 0.$$

Therefore, they are Pareto optimal.

Proof. As it has been shown in Lemma B.1 that the performance of CP with respect to each bidder k, Π_k^{CP}, converges to $U_k^{\mathrm{CP}}(\boldsymbol{\xi}^*)$. It suffices for us to show $\sum_{k\in\mathcal{K}} U_k^{\mathrm{CP}}(\boldsymbol{\xi}^*)$ is an upper bound of $\mathbb{E}[\pi^{\mathrm{H}}]$.

Rewrite the form of π^{H} for a realization of values and the highest bids outside the coalition $(\boldsymbol{v}, \boldsymbol{d}^O)$:

$$\pi^{\mathrm{H}}\left(\boldsymbol{v}, \boldsymbol{d}^O\right) := \max_{\boldsymbol{x}\in\{0,1\}^{K\times T}} \sum_{t=1}^{T}\sum_{k\in\mathcal{K}} x_{k,t}(v_{k,t} - d_t^O),$$

$$\text{s.t. } \sum_{t=1}^{T} x_{k,t}d_t^O \leq T\rho_k, \forall k \in \mathcal{K}. \tag{33}$$

Introduce Lagrangian multipliers for each budget constraint. The Lagrangian of optimization problem (33) is

$$\begin{aligned}
L(\boldsymbol{v}, \boldsymbol{d}^O; \boldsymbol{\mu}) &= \sum_{t=1}^{T} \boldsymbol{x}_t^\top(\boldsymbol{v}_t - d_t^O\boldsymbol{e}) + \boldsymbol{\mu}^\top\sum_{t=1}^{T}\left(-d_t^O\boldsymbol{x}_t + T\boldsymbol{\rho}\right) \\
&= \sum_{t=1}^{T} \boldsymbol{x}_t^\top(\boldsymbol{v}_t - d_t^O(\boldsymbol{e} + \boldsymbol{\mu})) + \boldsymbol{\rho}^\top\boldsymbol{\mu}.
\end{aligned}$$

where $\boldsymbol{e}$ is all-one vector.

The dual problem is therefore given by

$$\inf_{\boldsymbol{\mu}\geq 0} \sup_{\boldsymbol{x}} L(\boldsymbol{\mu}, \boldsymbol{x}) = \inf_{\boldsymbol{\mu}\geq 0}\left\{\max_{k\in\mathcal{K}}\left(v_{k,t} - (1+\mu_k)d_t^O\right)^+ + \boldsymbol{\rho}^\top\boldsymbol{\mu}\right\} =: \inf_{\boldsymbol{\mu}\geq 0} \phi(\boldsymbol{\mu}),$$

which serves an upper bound of π^{H}. Take expectations on both sides,

$$\mathbb{E}\left[\pi^{\mathrm{H}}\right] \leq \mathbb{E}\left[\inf_{\boldsymbol{\mu} \geq 0} \phi(\boldsymbol{\mu})\right] \leq \inf_{\boldsymbol{\mu} \geq 0} \mathbb{E}\left[\phi(\boldsymbol{\mu})\right] = \inf_{\boldsymbol{\mu} \geq 0}\left\{\mathbb{E}\left[\max_{k \in \mathcal{K}}\left(v_{k,t} - (1+\mu_k)d_t^O\right)^+\right] + \boldsymbol{\rho}^\top \boldsymbol{\mu}\right\} =: \inf_{\boldsymbol{\mu} \geq 0} \Phi(\boldsymbol{\mu}).$$

Denote the expected expenditure per auction of each bidder k to be

$$G_k^{\mathrm{S}}(\boldsymbol{\mu}) = \mathbb{E}\left[\mathbf{1}\left\{k \in \mathrm{argmax}_{k' \in \mathcal{K}}\left(v_{k',t} - (1+\mu_{k'})d_t^O\right)\right\} \mathbf{1}\left\{v_{k,t} \geq (1+\mu_k)d_t^O\right\} d_t^O\right].$$

Then by the Karush-Kuhn-Tucker condition, the optimal dual solution $\boldsymbol{\mu}^{**}$ satisfies the complementarity condition, which is

$$G_k^{\mathrm{S}}(\boldsymbol{\mu}^{**}) \leq \rho_k \perp \mu_k^{**} \geq 0, \forall k \in \mathcal{K}$$

and the corresponding expected utility of bidder k is

$$U_k^{\mathrm{S}}(\boldsymbol{\mu}^{**}) = \mathbb{E}\left[\mathbf{1}\left\{k \in \mathrm{argmax}_{k' \in \mathcal{K}}\left(v_{k',t} - (1+\mu_{k'}^{**})d_t^O\right)\right\} \mathbf{1}\left\{v_{k',t} \geq (1+\mu_{k'}^{**})d_t^O\right\}\left(v_{k',t} - d_t^O\right)\right].$$

When every bidder has the same distribution and budget, we have $\mu_k^{**} = \mu_1^{**}$ for any $k \in \mathcal{K}$, and $\xi_k^* = \xi_1^*$, where $\boldsymbol{\xi}^*$ are the equilibria parameters defined by (10). Recall the forms of G_k^{CP} and U_k^{CP} in this case,

$$\begin{aligned}
G_k^{\mathrm{CP}}(\boldsymbol{\xi}^*) &= \mathbb{E}\left[\prod_{k' \in \mathcal{K}} \mathbf{1}\left\{v_{k,t} \geq v_{k',t}\right\} \mathbf{1}\left\{v_{k,t} \geq (1+\xi_k^*)d_t^O\right\} d_t^O\right], \\
U_k^{\mathrm{CP}}(\boldsymbol{\xi}^*) &= \mathbb{E}\left[\prod_{k' \in \mathcal{K}} \mathbf{1}\left\{v_{k,t} \geq v_{k',t}\right\} \mathbf{1}\left\{v_{k,t} \geq (1+\xi_k^*)d_t^O\right\}\left(v_{k,t} - d_t^O\right)\right].
\end{aligned} \tag{34}$$

We show $G_k^{\mathrm{S}}(\boldsymbol{\mu})$ and $U_k^{\mathrm{S}}(\boldsymbol{\mu})$ have the same form as (34), when bidders are homogeneous. Then by the uniqueness of $\boldsymbol{\xi}^*$, we have $\boldsymbol{\xi}^* = \boldsymbol{\mu}^{**}$. Moreover, we can conclude that CP maximizes the sum of bidders' utilities in expectation.

To show the two expected expenditures per auction and the total expected utilities have the same form, it suffices to show that the sets where the statements of the indicator functions hold are equal. Specifically, notice that the set $\{k \in \mathrm{argmax}_{k' \in \mathcal{K}}\left(v_{k',t} - (1+\mu_{k'})\right)\}$ can be rewritten as $\cap_{k' \in \mathcal{K}}\left\{v_{k,t} - (1+\mu_k)d_t^O \geq v_{k',t} - (1+\mu_{k'})d_t^O\right\}$. Moreover, when $\mu_k = \mu_1$,

$$v_{k,t} - (1+\mu_k)d_t^O \geq v_{k',t} - (1+\mu_{k'})d_t^O \Leftrightarrow v_{k,t} \geq v_{k',t}.$$

This completes the proof for CP. To show the results of HP, just notice that at equilibrium, the expected expenditures per auction and the total expected utilities of HP has the same form as those of CP. □

D.2. Proof of Theorem 5.2

Theorem 5.2. *Under CP, bidder k's expected utility U_k^{CP} will not increase, if she under-reports her budget and others report budgets truthfully.*

Proof. Let the expenditure rate of a member be ρ. By symmetry, all members' equilibrium parameters are the same: $\boldsymbol{\xi}^* = \xi^* \boldsymbol{e}$ for some $\xi^* \in [0, \bar{\xi}]$. When bidder k misreports a rate $\rho' < \rho$, denote her corresponding equilibrium parameter to be ξ_k^* and that of other member i to be ξ_1^*. By a slight abuse of notation, we let G_i^{CP} be the function of (ξ_1, ξ_k), the first parameter for other bidders, and the second for bidder k. Without loss of generality, let $k \neq 1$, and we only consider the truthful bidder 1 as other truthful bidders behave the same. We write down the form of G_i^{CP} and U_i^{CP}:

$$G_k^{\mathrm{CP}}(\xi^*, \xi^*) = \mathbb{E}\left[\mathbf{1}\left\{v_k \geq \max_{i \neq k} v_i\right\} \mathbf{1}\left\{v_k \geq (1+\xi^*)d^O\right\} d^O\right] = G_1^{\mathrm{CP}}(\xi^*, \xi^*) \leq \rho, \tag{35}$$

$$G_1^{\mathrm{CP}}(\xi_1^*, \xi_k^*) = \mathbb{E}\left[\Pi_{\substack{i \neq 1, \\ i \neq k}}\mathbf{1}\left\{v_1 \geq v_i\right\} \mathbf{1}\left\{v_1 \geq \frac{1+\xi_1^*}{1+\xi_k^*}v_k\right\} \mathbf{1}\left\{v_1 \geq (1+\xi_1^*)d^O\right\} d^O\right] \leq \rho, \tag{36}$$

$$G_k^{\mathrm{CP}}(\xi_1^*, \xi_k^*) = \mathbb{E}\left[\mathbf{1}\left\{v_k \geq \frac{1+\xi_k^*}{1+\xi_1^*}\max_{i \neq k} v_i\right\} \mathbf{1}\left\{v_k \geq (1+\xi_k^*)d^O\right\} d^O\right] \leq \rho' < \rho, \tag{37}$$

and,

$$U_k^{\mathrm{CP}}(\xi^*, \xi^*) = \mathbb{E}\left[\mathbf{1}\left\{ v_k \geq \max_{i \neq k} v_i \right\} \mathbf{1}\left\{ v_k \geq (1+\xi^*)d^O \right\} \left(v_k - d^O\right) \right] = U_1^{\mathrm{CP}}(\xi^*, \xi^*), \tag{38}$$

$$U_1^{\mathrm{CP}}(\xi_1^*, \xi_k^*) = \mathbb{E}\left[\Pi_{\substack{i \neq 1, \\ i \neq k}} \mathbf{1}\left\{ v_1 \geq v_i \right\} \mathbf{1}\left\{ v_1 \geq \frac{1+\xi_1^*}{1+\xi_k^*} v_k \right\} \mathbf{1}\left\{ v_1 \geq (1+\xi_1^*)d^O \right\} \left(v_1 - d^O\right) \right], \tag{39}$$

$$U_k^{\mathrm{CP}}(\xi_1^*, \xi_k^*) = \mathbb{E}\left[\mathbf{1}\left\{ v_k \geq \frac{1+\xi_k^*}{1+\xi_1^*} \max_{i \neq k} v_i \right\} \mathbf{1}\left\{ v_k \geq (1+\xi_k^*)d^O \right\} \left(v_k - d^O\right) \right]. \tag{40}$$

We consider all the possible cases the solutions of NCPs may be. We show that for the only reasonable cases that do not lead to contradiction with the definition of NCP and strong monotonicity, we have

$$\left\{ v_k \geq \frac{1+\xi_k^*}{1+\xi_1^*} \max_{i \neq k} v_i \right\} \subseteq \left\{ v_k \geq \max_{i \neq k} v_i \right\}, \tag{41}$$
$$\left\{ v_k \geq (1+\xi_k^*)d^O \right\} \subseteq \left\{ v_k \geq (1+\xi^*)d^O \right\},$$

therefore $U_k^{\mathrm{CP}}(\xi_1^*, \xi_k^*) \leq U_k^{\mathrm{CP}}(\xi^*, \xi^*)$.

When $\xi^* = \xi_1^* = \xi_k^* = 0$, all the members have sufficient budgets, after misreporting, bidder k's budget is still big enough. In this case, her obtained utility will not decrease.

When $\xi^* \neq 0$, then $G_k^{\mathrm{CP}}(\xi^*, \xi^*) = \rho$. First we have $\xi_1^* \neq \xi_k^*$:

- for $\xi_1^* = \xi_k^* > \xi^*$, we have $\rho = G_1^{\mathrm{CP}}(\xi_1^*, \xi_k^*) < G_k^{\mathrm{CP}}(\xi^*, \xi^*) = \rho$, leading to a contradiction;
- for $\xi_1^* = \xi_k^* < \xi^*$, we have $\rho' \geq G_k^{\mathrm{CP}}(\xi_1^*, \xi_k^*) = G_1^{\mathrm{CP}}(\xi_1^*, \xi_k^*) > G_k^{\mathrm{CP}}(\xi^*, \xi^*) = \rho$, leading to a contradiction.

Then we argue that $\xi_k^* \neq 0$, as otherwise it must be the case that $\xi_1^* \neq 0$ and

$$G_k^{\mathrm{CP}}(\xi^*, \xi^*) = G_1^{\mathrm{CP}}(\xi_1^*, \xi_k^*) = \rho,$$
$$G_k^{\mathrm{CP}}(\xi_1^*, \xi_k^*) \leq \rho' < \rho.$$

By the strong monotonicity of $\boldsymbol{G}^{\mathrm{CP}}$, we have

$$(\xi_k^* - \xi^*)(G_k^{\mathrm{CP}}(\xi^*, \xi^*) - G_k^{\mathrm{CP}}(\xi_1^*, \xi_k^*)) > 0,$$

and

$$\xi_k^* > \xi^*,$$

leading to a contradiction.

When $\xi_1^* = 0$, it must be the case that

$$\xi_k^* > \xi_1^* = 0, \text{ and } \xi^* \geq \xi_1^*.$$

and,

$$G_1^{\mathrm{CP}}(\xi_1^*, \xi_k^*) \leq G_k^{\mathrm{CP}}(\xi^*, \xi^*) = \rho,$$
$$G_k^{\mathrm{CP}}(\xi_1^*, \xi_k^*) = \rho' < \rho.$$

By the strong monotonicity of $\boldsymbol{G}^{\mathrm{CP}}$,

$$(\xi_k^* - \xi^*)(G_k^{\mathrm{CP}}(\xi^*, \xi^*) - G_k^{\mathrm{CP}}(\xi_1^*, \xi_k^*)) > (K-1)(\xi^* - \xi_1^*)(G_1^{\mathrm{CP}}(\xi^*, \xi^*) - G_1(\xi_1^*, \xi_k^*)) \geq 0,$$

which means $\xi_k^* > \xi^*$. In this case,

$$\left\{ v_k \geq \frac{1+\xi_k^*}{1+\xi_1^*} \max_{i \neq k} v_i \right\} \subseteq \left\{ v_k \geq \max_{i \neq k} v_i \right\},$$
$$\left\{ v_k \geq (1+\xi_k^*)d^O \right\} \subseteq \left\{ v_k \geq (1+\xi^*)d^O \right\}.$$

(41) holds.

Finally we discuss the case when $\xi^*, \xi_1^*, \xi_k^* > 0$, and

$$G_1^{\mathrm{CP}}(\xi_1^*, \xi_k^*) = G_k^{\mathrm{CP}}(\xi^*, \xi^*) = \rho, \tag{42}$$

$$G_k^{\mathrm{CP}}(\xi_1^*, \xi_k^*) = \rho' < \rho. \tag{43}$$

First by the strong monotonicity of $\boldsymbol{G}^{\mathrm{CP}}$, we have $\xi_k^* > \xi^*$. Then consider the following possible cases:

- $\xi_1^* \geq \xi_k^* > \xi^*$: $(1 + \xi_1^*)/(1 + \xi_k^*) > 1$, and

$$\begin{aligned}
\left\{ v_1 \geq \frac{1 + \xi_1^*}{1 + \xi_k^*} v_k \right\} &\subseteq \{v_1 \geq v_k\}, \\
\left\{ v_1 \geq (1 + \xi_1^*) d^O \right\} &\subseteq \left\{ v_1 \geq (1 + \xi^*) d^O \right\},
\end{aligned} \tag{44}$$

  which means that $G_1^{\mathrm{CP}}(\xi_1^*, \xi_k^*) \leq G_1^{\mathrm{CP}}(\xi^*, \xi^*)$. We state the following lemma:

  Lemma D.1. *When $\boldsymbol{G}^{\mathrm{CP}}$ is strongly monotone, then for any $\xi_1^* > \xi^*$ and $(1 + \xi_1^*)/(1 + \xi_k^*) > 1$, we have*

$$G_1^{\mathrm{CP}}(\xi_1^*, \xi_k^*) < G_1^{\mathrm{CP}}(\xi^*, \xi^*).$$

  By Lemma D.1, $G_1^{\mathrm{CP}}(\xi_1^*, \xi_k^*) < G_1^{\mathrm{CP}}(\xi^*, \xi^*)$, leading to a contradiction.

- $\xi_k^* > \xi_1^*$: $(1 + \xi_k^*)/(1 + \xi_1^*) > 1$, and

$$\begin{aligned}
\left\{ v_k \geq \frac{1 + \xi_k^*}{1 + \xi_1^*} \max_{i \neq k} v_i \right\} &\subseteq \left\{ v_k \geq \max_{i \neq k} v_i \right\}, \\
\left\{ v_k \geq (1 + \xi_k^*) d^O \right\} &\subseteq \left\{ v_k \geq (1 + \xi^*) d^O \right\}.
\end{aligned} \tag{45}$$

  (41) holds.

When $\xi^* = 0$, and one of ξ_1^*, ξ_k^* is non-zero, first we argue that $\xi_1^* \neq \xi_k^*$, as it contradicts that $G_1^{\mathrm{CP}}(\xi_1^*, \xi_k^*) \neq G_k^{\mathrm{CP}}(\xi_1^*, \xi_k^*)$.

If $\xi_k^* = 0$, then $G_1^{\mathrm{CP}}(\xi_1^*, \xi_k^*) = \rho \geq G_k^{\mathrm{CP}}(\xi^*, \xi^*)$, and $\xi_1^* > \xi^* = \xi_k^* = 0$, we have $(1 + \xi_1^*)/(1 + \xi_k^*) > 1$, and (44) holds. By Lemma D.1, $G_1^{\mathrm{CP}}(\xi_1^*, \xi_k^*) < G_1^{\mathrm{CP}}(\xi^*, \xi^*)$, leading to a contradiction.

If $\xi_1^* = 0$, then $G_k^{\mathrm{CP}}(\xi_1^*, \xi_k^*) = \rho'$, and $\xi_k^* > \xi_1^* = \xi^*$. In this case, $(1 + \xi_1^*)/(1 + \xi_k^*) < 1$, (45) and therefore (41) holds.

When $\xi_1^*, \xi_k^* \neq 0$, consider the two cases:

- if $\xi_1^* > \xi_k^* > \xi^* = 0$, then $(1 + \xi_1^*)/(1 + \xi_k^*) > 1$, (44) holds. By Lemma D.1, $G_1^{\mathrm{CP}}(\xi_1^*, \xi_k^*) < G_1^{\mathrm{CP}}(\xi^*, \xi^*) \leq \rho$, leading to a contradiction.

- if $\xi_k^* > \xi_1^* > \xi^* = 0$, then $(1 + \xi_1^*)/(1 + \xi_k^*) < 1$, (45) and therefore (41) holds.

$\square$

D.3. Proof of Theorem 5.3

Theorem 5.3. *Under IP, bidder k's expected obtained value V_k will not increase if she under-reports her budget and others report budgets truthfully.*

Proof. Let the expenditure rate of a member be ρ. By symmetry, all members' equilibrium parameters are the same: $\boldsymbol{\lambda}^* = \lambda^* \boldsymbol{e}$ for some $\lambda^* \in [0, \bar{\lambda}]$. When bidder k misreports a rate $\rho' < \rho$, denote her corresponding equilibrium parameter to be λ_k^* and that of another member i to be λ_1^*. By a slight abuse of notation, we let G_i^{CP} be the function of (λ_1, λ_k), the

first parameter for other bidders, and the second for bidder k. Without loss of generality, let $k \neq 1$, and we only consider the truthful bidder 1 as other truthful bidders behave the same. We write down the form of G_i and V_i:

$$G_k(\lambda^*, \lambda^*) = \mathbb{E}\left[\mathbf{1}\left\{v_k \geq \max_{i \neq k} v_i\right\}\mathbf{1}\left\{v_k \geq (1+\lambda^*)d^O\right\}\max\left\{\frac{\max_{i \neq k} v_i}{1+\lambda^*}, d^O\right\}\right] = G_1(\lambda^*, \lambda^*) \leq \rho,$$

$$G_1(\lambda_1^*, \lambda_k^*) = \mathbb{E}\left[\Pi_{\substack{i \neq 1 \\ i \neq k}}\mathbf{1}\left\{v_1 \geq v_i\right\}\mathbf{1}\left\{v_1 \geq \frac{1+\lambda_1^*}{1+\lambda_k^*}v_k\right\}\mathbf{1}\left\{v_1 \geq (1+\lambda_1^*)d^O\right\}\max\left\{\frac{\max_{\substack{i \neq 1 \\ i \neq k}} v_i}{1+\lambda_1^*}, \frac{v_k}{1+\lambda_k^*}, d^O\right\}\right] \leq \rho,$$

$$G_k(\lambda_1^*, \lambda_k^*) = \mathbb{E}\left[\mathbf{1}\left\{v_k \geq \frac{1+\lambda_k^*}{1+\lambda_1^*}\max_{i \neq k} v_i\right\}\mathbf{1}\left\{v_k \geq (1+\lambda_k^*)d^O\right\}\max\left\{\frac{\max_{i \neq k} v_i}{1+\lambda_1^*}, d^O\right\}\right] = \rho' < \rho,$$

and,

$$V_k(\lambda^*, \lambda^*) = \mathbb{E}\left[\mathbf{1}\left\{v_k \geq \max_{i \neq k} v_i\right\}\mathbf{1}\left\{v_k \geq (1+\lambda^*)d^O\right\}v_k\right] = V_1(\lambda^*, \lambda^*),$$

$$V_1(\lambda_1^*, \lambda_k^*) = \mathbb{E}\left[\Pi_{\substack{i \neq 1 \\ i \neq k}}\mathbf{1}\left\{v_1 \geq v_i\right\}\mathbf{1}\left\{v_1 \geq \frac{1+\lambda_1^*}{1+\lambda_k^*}v_k\right\}\mathbf{1}\left\{v_1 \geq (1+\lambda_1^*)d^O\right\}v_1\right],$$

$$V_k(\lambda_1^*, \lambda_k^*) = \mathbb{E}\left[\mathbf{1}\left\{v_k \geq \frac{1+\lambda_k^*}{1+\lambda_1^*}\max_{i \neq k} v_i\right\}\mathbf{1}\left\{v_k \geq (1+\lambda_k^*)d^O\right\}v_k\right].$$

We consider all the possible cases the solutions of NCPs may be. We show that for the only reasonable cases that do not contradict the definition of NCP and strong monotonicity, we have

$$\left\{v_k \geq \frac{1+\lambda_k^*}{1+\lambda_1^*}\max_{i \neq k} v_i\right\} \subseteq \left\{v_k \geq \max_{i \neq k} v_i\right\}, \tag{46}$$
$$\left\{v_k \geq (1+\lambda_k^*)d^O\right\} \subseteq \left\{v_k \geq (1+\lambda^*)d^O\right\},$$

therefore $V_k(\lambda_1^*, \lambda_k^*) \leq V_k(\lambda^*, \lambda^*)$.

When $\lambda^* = \lambda_1^* = \lambda_k^* = 0$, all the members have sufficient budgets, after misreporting, bidder k's budget is still big enough. In this case, her obtained value will not decrease.

When $\lambda^* \neq 0$, then $G_k(\lambda^*, \lambda^*) = \rho$. First we have $\lambda_1^* \neq \lambda_k^*$:

- for $\lambda_1^* = \lambda_k^* > \lambda^*$, we have $\rho = G_1(\lambda_1^*, \lambda_k^*) < G_k(\lambda^*, \lambda^*) = \rho$, leading to a contradiction;

- for $\lambda_1^* = \lambda_k^* < \lambda^*$, we have $\rho' \geq G_k(\lambda_1^*, \lambda_k^*) = G_1(\lambda_1^*, \lambda_k^*) > G_k(\lambda^*, \lambda^*) = \rho$, leading to a contradiction.

Then we argue that $\lambda_k^* \neq 0$, as otherwise it must be the case that $\lambda_1^* \neq 0$ and

$$G_k(\lambda^*, \lambda^*) = G_1(\lambda_1^*, \lambda_k^*) = \rho,$$
$$G_k(\lambda_1^*, \lambda_k^*) \leq \rho' < \rho.$$

By the strong monotonicity of G, we have

$$(\lambda_k^* - \lambda^*)(G_k(\lambda^*, \lambda^*) - G_k(\lambda_1^*, \lambda_k^*)) > 0,$$

and

$$\lambda_k^* > \lambda^*,$$

leading to a contradiction.

When $\lambda_1^* = 0$, it must be the case that

$$\lambda_k^* > \lambda_1^* = 0, \text{ and } \lambda^* \geq \lambda_1^*.$$

By the strong monotonicity of G,

$$(\lambda_k^* - \lambda^*)(G_k(\lambda^*, \lambda^*) - G_k(\lambda_1^*, \lambda_k^*)) > (\lambda^* - \lambda_1^*)(G_1(\lambda^*, \lambda^*) - G_1(\lambda_1^*, \lambda_k^*)) \geq 0,$$

which means $\lambda_k^* > \lambda^*$. In this case,

$$\left\{ v_k \geq \frac{1 + \lambda_k^*}{1 + \lambda_1^*} \max_{i \neq k} v_i \right\} \subseteq \left\{ v_k \geq \max_{i \neq k} v_i \right\},$$

$$\left\{ v_k \geq (1 + \lambda_k^*) d^O \right\} \subseteq \left\{ v_k \geq (1 + \lambda^*) d^O \right\}.$$

(46) holds.

Finally we discuss the case when $\lambda^*, \lambda_1^*, \lambda_k^* > 0$, and

$$G_1(\lambda_1^*, \lambda_k^*) = G_k(\lambda^*, \lambda^*) = \rho, \tag{47}$$

$$G_k(\lambda_1^*, \lambda_k^*) = \rho' < \rho. \tag{48}$$

First by strong monotonicity of G, we have $\lambda_k^* > \lambda^*$. Then consider the following possible cases:

- $\lambda_1^* \geq \lambda_k^* > \lambda^*$: $(1 + \lambda_1^*)/(1 + \lambda_k^*) > 1$, and

$$\left\{ v_1 \geq \frac{1 + \lambda_1^*}{1 + \lambda_k^*} v_k \right\} \subseteq \left\{ v_1 \geq v_k \right\},$$

$$\left\{ v_1 \geq (1 + \lambda_1^*) d^O \right\} \subseteq \left\{ v_1 \geq (1 + \lambda^*) d^O \right\},$$

$$\frac{\max_{\substack{i \neq 1 \\ i \neq k}} v_i}{1 + \lambda_1^*} < \frac{\max_{\substack{i \neq 1 \\ i \neq k}} v_i}{1 + \lambda^*}, \tag{49}$$

$$\frac{v_k}{1 + \lambda_k^*} < \frac{v_k}{1 + \lambda^*}.$$

$G_1(\lambda_1^*, \lambda_k^*) \leq G_1(\lambda^*, \lambda^*)$. We prove a similar lemma as Lemma D.1,

Lemma D.2. *When G is strongly monotone, then for any $\lambda_1^* > \lambda^*$ and $(1 + \lambda_1^*)/(1 + \lambda_k^*) > 1$, we have*

$$G_1(\lambda_1^*, \lambda_k^*) < G_1(\lambda^*, \lambda^*)$$

By Equation (49) and Lemma D.2, $G_1(\lambda_1^*, \lambda_k^*) < G_1(\lambda^*, \lambda^*)$, leading to a contradiction.

- $\lambda_k^* > \lambda_1^*$: $(1 + \lambda_k^*)/(1 + \lambda_1^*) > 1$, and

$$\left\{ v_k \geq \frac{1 + \lambda_k^*}{1 + \lambda_1^*} \max_{i \neq k} v_i \right\} \subseteq \left\{ v_k \geq \max_{i \neq k} v_i \right\},$$

$$\left\{ v_k \geq (1 + \lambda_k^*) d^O \right\} \subseteq \left\{ v_k \geq (1 + \lambda^*) d^O \right\}.$$

(46) holds.

When $\lambda^* = 0$, and one of λ_1^*, λ_k^* is non-zero, first we argue that $\lambda_1^* \neq \lambda_k^*$, as it contradicts that $G_1(\lambda_1^*, \lambda_k^*) \neq G_k(\lambda_1^*, \lambda_k^*)$.

If $\lambda_k^* = 0$, then $G_1(\lambda_1^*, \lambda_k^*) = \rho \geq G_1(\lambda^*, \lambda^*)$, and $\lambda_1^* > \lambda^* = \lambda_k^* = 0$, we have $(1 + \lambda_1^*)/(1 + \lambda_k^*) > 1$, and

$$\left\{ v_1 \geq \frac{1 + \lambda_1^*}{1 + \lambda_k^*} v_k \right\} \subseteq \left\{ v_1 \geq v_k \right\},$$

$$\left\{ v_1 \geq (1 + \lambda_1^*) d^O \right\} \subseteq \left\{ v_1 \geq (1 + \lambda^*) d^O \right\},$$

$$\frac{\max_{\substack{i \neq 1 \\ i \neq k}} v_i}{1 + \lambda_1^*} < \frac{\max_{\substack{i \neq 1 \\ i \neq k}} v_i}{1 + \lambda^*},$$

$$\frac{v_k}{1 + \lambda_k^*} \leq \frac{v_k}{1 + \lambda^*}.$$

By Lemma D.2, $G_1(\lambda_1^*, \lambda_k^*) < G_1(\lambda^*, \lambda^*)$, leading to a contradiction.

If $\lambda_1^* = 0$, then $G_k(\lambda_1^*, \lambda_k^*) = \rho'$, and $\lambda_k^* > \lambda_1^* = \lambda^*$. In this case, $(1 + \lambda_1^*)/(1 + \lambda_k^*) < 1$, and (46) holds.

When $\lambda_1^*, \lambda_k^* \neq 0$, consider the two cases:

- if $\lambda_1^* > \lambda_k^* > \lambda^* = 0$, then (49) holds, leading to a contradiction.

- if $\lambda_k^* > \lambda_1^* > \lambda^* = 0$, then $(1 + \lambda_1^*)/(1 + \lambda_k^*) < 1$, and (46) holds.

$\square$

D.4. An Experimental Example that IP is Not Truthful on Utilities

Consider a coalition of two bidders. Consider independent random varaibles X_1, X_2, distributed uniformly in $[0, 1]$, and Y_1, Y_2, distributed uniformly in $[1, 2]$, respectively. For $i \in 1, 2$, the random variable of bidder i's value is $Z_i = 1/3 X_i + 2/3 Y_i$. The competing bid outside the coalition is uniformly distributed in $[0, 1]$. We set the true expenditure ratio of each bidder to be 0.5, and let bidder 1 misreports a ratio of 0.49.

We run IP the same as we run the example in Section 6.1. And we repeated the experiment 100 times. For every experiment, all bidders have higher utilities when bidder 1's misreports: bidder 1's increases by about 0.003, and bidder 2's by about 0.006.

As we further check the changes of bidders' obtained values and the equilibrium parameters, we can find out that bidder 1's obtained value decreases by about 0.007. Denote the equilibrium parameter when bidder 1 truthfully reports by μ^*, the equilibrium parameter of bidder 1's by μ_1^* and of bidder 2's by μ_2^*, respectively, when bidder 1 misreports. It shows that $\mu_1^* > \mu_2^* > \mu^*$, which is consistent with the analysis above.

D.5. Proofs of Lemmas in Proving Theorem 5.2 and 5.3

We give the proof of Lemma D.1, as the proof of Lemma D.2 is similar.

Lemma D.1. *When $\boldsymbol{G}^{\mathrm{CP}}$ is strongly monotone, then for any $\xi_1^* > \xi^*$ and $(1 + \xi_1^*)/(1 + \xi_k^*) > 1$, we have*

$$G_1^{\mathrm{CP}}(\xi_1^*, \xi_k^*) < G_1^{\mathrm{CP}}(\xi^*, \xi^*).$$

Proof. We prove this by contradiction. Suppose $G_1^{\mathrm{CP}}(\xi_1^*, \xi_k^*) = G_1^{\mathrm{CP}}(\xi^*, \xi^*)$, then

$$
\begin{aligned}
0 =& G_1^{\mathrm{CP}}(\xi^*, \xi^*) - G_1^{\mathrm{CP}}(\xi_1^*, \xi_k^*) = \mathbb{E}\left[\prod_{\substack{i \in \mathcal{K}, \\ i \neq 1, i \neq k}} \mathbf{1}\left\{v_1 \geq v_i\right\} \right. \\
& \left. \cdot \mathbf{1}\left\{ \left((v_1 \geq v_k) \wedge \left((1 + \xi^*)d^O \leq v_1 < (1 + \xi_1^*)d^O\right)\right) \vee \left(\left(v_k \leq v_1 < \frac{1 + \xi_1^*}{1 + \xi_k^*} v_k\right) \wedge \left(v_1 \geq (1 + \xi^*)d^O\right)\right) \right\} \right],
\end{aligned}
$$

which means

$$
\bigcap_{\substack{i \in \mathcal{K}, \\ i \neq 1, i \neq k}} \{v_1 \geq v_i\} \bigcap \left(\left(\{v_1 \geq v_k\} \bigcap \{(1 + \xi^*)d^O \leq v_1 < (1 + \xi_1^*)d^O\}\right) \right.
$$

$$
\left. \bigcup \left(\left\{v_k \leq v_1 < \frac{1 + \xi_1^*}{1 + \xi_k^*} v_k\right\} \bigcap \{v_1 \geq (1 + \xi^*)d^O\}\right) \right) \tag{50}
$$

has zero measure, then for $\xi \in (\xi^*, \xi_1^*)$, the set

$$
\bigcap_{\substack{i \in \mathcal{K}, \\ i \neq 1, i \neq k}} \{v_1 \geq v_i\} \bigcap \left(\left(\{v_1 \geq v_k\} \bigcap \{(1 + \xi)d^O \leq v_1 < (1 + \xi_1^*)d^O\}\right) \right.
$$

$$
\left. \bigcup \left(\left\{v_k \leq v_1 < \frac{1 + \xi_1^*}{1 + \xi_k^*} v_k\right\} \bigcap \{v_1 \geq (1 + \xi)d^O\}\right) \right)
$$

is a subset of (50), and therefore has measure zero. This means that

$$G_1^{\mathrm{CP}}(\xi^*, \xi^*) - G_1^{\mathrm{CP}}(\xi_1^*, \xi_k^*) = 0 = G_1^{\mathrm{CP}}(\xi, \xi) - G_1^{\mathrm{CP}}(\xi_1^*, \xi_k^*),$$

equivalent to

$$G_1^{\mathrm{CP}}(\xi^*, \xi^*) = G_1^{\mathrm{CP}}(\xi, \xi).$$

This contradicts that $\boldsymbol{G}^{\mathrm{CP}}$ is strongly monotone. $\square$

E. Auxiliary Results

Lemma E.1. $G_k(\boldsymbol{\lambda})$, $V_k(\boldsymbol{\lambda})$, $G_k^{\mathrm{CP}}(\boldsymbol{\xi})$ *are Lipschitz continuous.*

Proof. We denote by L_k the cumulative distribution function of d_k^I:

$$L_k(x; \boldsymbol{\lambda}) = \prod_{i \in \mathcal{K}: i \neq k} F_i((1 + \lambda_i)x). \tag{51}$$

We omit $\boldsymbol{\lambda}$ when there is no confusion. In what follows, we denote by $\overline{f}$ the upper bound on f_k and by $\overline{h}$ the upper bound on h_k. We use $\bar{F}_k = 1 - F_k$ to denote the complementary cumulative distribution function of v_k.

1. We write down two expressions of $G_k(\boldsymbol{\lambda})$ and calculate its derivatives with respect to λ_k and λ_i for $i \neq k$.

$$G_k(\boldsymbol{\lambda}) = \int x \bar{F}_k((1 + \lambda_k)x) \, d\left(L_k(x)H(x)\right) \tag{52}$$

$$= \int L_k(x)H(x)\left((1 + \lambda_k)x f_k((1 + \lambda_k)x) - \bar{F}((1 + \lambda_k)x)\right) \, dx. \tag{53}$$

Using the first expression (52), we have

$$\frac{\partial G_k(\boldsymbol{\lambda})}{\partial \lambda_k} = -\int x^2 f((1 + \mu_k)x) \, d(L_k(x)H(x)),$$

and thus,

$$\left| \frac{\partial G_k(\boldsymbol{\lambda})}{\partial \lambda_k} \right| \leq \bar{v}\bar{f}.$$

Using the second expression (53), we have for $i \neq k$,

$$\frac{\partial G_k(\boldsymbol{\lambda})}{\partial \lambda_i} = \int \frac{t}{(1 + \lambda_k)^2} f_i\left(\frac{1 + \lambda_i}{1 + \lambda_k}t\right) \prod_{j \neq i,k} F_j\left(\frac{1 + \lambda_j}{1 + \lambda_k}t\right) H\left(\frac{t}{1 + \lambda_k}\right) (t f_k(t) - \bar{F}_k(t)) \, dt,$$

and thus,

$$\left| \frac{\partial G_k(\boldsymbol{\lambda})}{\partial \lambda_i} \right| \leq \bar{v}\bar{f}(\bar{v}\bar{f} + 1).$$

2. We write down the expression of $V_k(\boldsymbol{\lambda})$ and calculate its derivatives with respect to λ_k and λ_i for $i \neq k$.

$$V_k(\boldsymbol{\lambda}) = \int xH\left(\frac{x}{1 + \lambda_k}\right) L_k\left(\frac{x}{1 + \lambda_k}\right) \, dF_k(x).$$

Therefore

$$\frac{\partial V_k(\boldsymbol{\lambda})}{\partial \lambda_k} = -\int \left(\frac{x}{1 + \lambda_k}\right)^2 \left(h\left(\frac{x}{1 + \lambda_k}\right) L_k\left(\frac{x}{1 + \lambda_k}\right) + H\left(\frac{x}{1 + \lambda_k}\right) l_k\left(\frac{x}{1 + \lambda_k}\right)\right) \, dF_k(x),$$

where l_k is the density function of L_k. Since

$$l_k(x) = \sum_{i \neq k} (1 + \lambda_i) f_i((1 + \lambda_i)x) \prod_{j \neq i,k} F_j((1 + \lambda_i)x),$$

we have

$$\left| \frac{\partial V_k(\boldsymbol{\lambda})}{\partial \lambda_k} \right| \leq \bar{v}^2 \bar{h} + (K - 1)\bar{v}^2 \bar{f}.$$

And for $i \neq k$,

$$\frac{\partial V_k(\boldsymbol{\lambda})}{\partial \lambda_i} = \int \frac{x^2}{1 + \lambda_k} H\left(\frac{x}{1 + \lambda_k}\right) f_i\left(\frac{1 + \lambda_i}{1 + \lambda_k}x\right) \prod_{j \neq i,k} F_j\left(\frac{1 + \lambda_j}{1 + \lambda_k}x\right) dF_k(x),$$

$$\left| \frac{\partial V_k(\boldsymbol{\lambda})}{\partial \lambda_k} \right| \leq \bar{v}^2 \bar{f}.$$

3. We write down the expression of $G_k^{\mathrm{CP}}(\boldsymbol{\xi})$ and calculate its derivatives with respect to ξ_k and ξ_i for $i \neq k$.

$$G_k^{\mathrm{CP}}(\boldsymbol{\xi}) = \int \int_{x \geq y} \bar{F}_k((1 + \xi_k)x)y \, dL_k(x) \, dH(y) + \int \bar{F}_k((1 + \xi_k)y)L_k(y)y \, dH(y)$$

$$= \int y \int_{x \geq y} (1 + \xi_k)L_k(x)f_k((1 + \xi_k)x) \, dx \, dH(y),$$

Therefore, we have

$$\frac{\partial G_k^{\mathrm{CP}}(\boldsymbol{\xi})}{\partial \xi_k} = -\int \int_{x \geq y} x f_k((1 + \xi_k)x)y \, dL_k(x) \, dH(y) - \int y^2 f((1 + \xi_k)y) L_k(y) \, dH(y).$$

$$\left| \frac{\partial G_k^{\mathrm{CP}}(\boldsymbol{v})}{\partial \xi_k} \right| \leq 2\bar{v}^2 \bar{f}.$$

And for $i \neq k$,

$$\frac{\partial G_k^{\mathrm{CP}}(\boldsymbol{\xi})}{\partial \xi_i} = \int y \int_{t \geq (1 + \xi_k)y} t f_i\left(\frac{1 + \xi_i}{1 + \xi_k}t\right) f_k(t) \prod_{j \neq i,k} F_j\left(\frac{1 + \xi_j}{1 + \xi_k}t\right) dx \, dH(y),$$

$$\left| \frac{\partial G_k^{\mathrm{CP}}(\boldsymbol{\xi})}{\partial \xi_i} \right| \leq \bar{v}^2 \bar{f}^2.$$

$\square$

Lemma E.2. $G_k^{\mathrm{HP}}(\mu_k, \boldsymbol{\lambda})$ and $V_k^{\mathrm{HP}}(\mu_k, \boldsymbol{\lambda})$ are Lipschitz continuous. Moreover, if for all $k \in \mathcal{K}$, bidder k's density function has a lower bound $\underline{f} > 0$ over the interval $[0, \bar{v}_k]$, and density function h also has a lower bound $\underline{h} > 0$ over the interval $[0, \bar{v}]$, then there exists constants $\underline{G}' > 0$ such that

$$\frac{\partial G_k^{\mathrm{HP}}(\mu_k, \boldsymbol{\lambda})}{\partial \mu_k} \leq -\underline{G}' < 0.$$

Proof. In hybrid coordinated pacing algorithms, we still denote by L_k the cumulative distribution function of d_k^I:

$$L_k(x; \mu_k, \boldsymbol{\lambda}) = \prod_{i \in \mathcal{K}: i \neq k} F_i((1 + \lambda_i)x). \tag{54}$$

Note that L_k is independent of μ_k. We omit μ_k and $\boldsymbol{\lambda}$ when there is no confusion. In what follows, we denote by $\overline{f}$ the upper bound on f_k and by $\overline{h}$ the upper bound on h_k. We use $\bar{F}_k$ to denote the complementary cumulative distribution function of v_k.

We first consider bidder k's expected expenditure under Algorithm 2 and calculate its derivatives with respect to μ_k, λ_k and λ_i for $i \neq k$.

$$G_k^{\text{HP}}(\mu_k, \boldsymbol{\lambda}) = \int_y y \int_{x \geq \frac{1+\mu_k}{1+\lambda_k}y} \bar{F}_k((1+\lambda_k)x) \, dL_k(x) \, dH(y)$$

$$+ \int_y y \bar{F}_k((1+\mu_k)y) L_k \left(\frac{1+\mu_k}{1+\lambda_k}y \right) dH(y) \tag{55}$$

$$= \int_y y \int_{x \geq \frac{1+\mu_k}{1+\lambda_k}y} (1+\lambda_k) f_k((1+\lambda_k)x) L_k(x) \, dx \, dH(y). \tag{56}$$

1. Using the second expression (56) we obtain that

$$\frac{\partial G_k^{\text{HP}}(\mu_k, \boldsymbol{\lambda})}{\partial \mu_k} = -\int_y y^2 f_k \left((1+\mu_k)\, y \right) L_k \left(\frac{1+\mu_k}{1+\lambda_k}y \right) dH(y),$$

and thus,

$$\left| \frac{\partial G_k^{\text{HP}}(\mu_k, \boldsymbol{\lambda})}{\partial \mu_k} \right| \leq \bar{v}_k^2 \bar{f}.$$

Let $\underline{v} = \min_{i \in \mathcal{K}} \bar{v}_i$ and $\bar{\mu} = \max_{i \in \mathcal{K}} \bar{\mu}_i$. When $y \leq \underline{v}/(1+\bar{\mu})$ one has

$$\frac{(1+\lambda_i)(1+\mu_k)}{1+\lambda_k}y \leq (1+\lambda_i)y \leq \underline{v} \leq \bar{v}_i,$$

and consequently,

$$L_k \left(\frac{1+\mu_k}{1+\lambda_k}y \right) = \prod_{i \in \mathcal{K}: i \neq k} F_i \left(\frac{(1+\lambda_i)(1+\mu_k)}{1+\lambda_k}y \right)$$

$$\geq \left(\frac{(1+\lambda_i)(1+\mu_k)}{1+\lambda_k}y\underline{f} \right)^{K-1} \geq \left(\frac{y\underline{f}}{1+\bar{\mu}} \right)^{K-1}. \tag{57}$$

We can upper bound the derivative $\partial G_k^{\text{HP}}(\mu_k, \boldsymbol{\lambda})/\partial \mu_k$ as follows:

$$\frac{\partial G_k^{\text{HP}}(\mu_k, \boldsymbol{\lambda})}{\partial \mu_k} \overset{(a)}{\leq} -\underline{f}\underline{h} \int_0^{\bar{v}_k/(1+\mu_k)} y^2 L_k(\frac{1+\mu_k}{1+\lambda_k}y) \, dy$$

$$\overset{(b)}{\leq} -\underline{f}\underline{h} \int_0^{\underline{v}/(1+\bar{\mu})} y^2 \left(\frac{y\underline{f}}{1+\bar{\mu}} \right)^{K-1} dy$$

$$= -\frac{\underline{v}^{K+2}\underline{f}^K\underline{h}}{(K+2)(1+\bar{\mu})^{2K}},$$

where (a) follows $f_k \geq \underline{f}$ and $h_k \geq \underline{h}$ when $(1+\mu_k)y \leq \bar{v}_k$; (b) follows from (57) and $\bar{v}_k/(1+\mu_k) \geq \underline{v}/(1+\bar{\mu})$.

2. Using the first expression (55) we obtain that

$$\frac{\partial G_k^{\text{HP}}(\mu_k, \boldsymbol{\lambda})}{\partial \lambda_k} = -\int_y y \int_{x \geq \frac{1+\mu_k}{1+\lambda_k}y} x f_k \left((1+\lambda_x)x \right) dL_k(x) \, dH(y),$$

and thus,

$$\left| \frac{\partial G_k^{\text{HP}}(\mu_k, \boldsymbol{\lambda})}{\partial \lambda_k} \right| \leq \bar{v}_k^2 \bar{f}.$$

3. Using the second expression (56) we obtain that for $i \neq k$,

$$\frac{\partial G_k^{\text{HP}}(\mu_k, \boldsymbol{\lambda})}{\partial \lambda_i} = \int_y y \int_{x \geq \frac{1+\mu_k}{1+\lambda_k}y} (1+\lambda_k) f_k((1+\lambda_k)x) \frac{\partial L_k(x)}{\partial \lambda_i} \, dx \, dH(y),$$

where the derivative of $L_k(x)$ satisfies

$$\frac{\partial L_k(x; \mu_k, \boldsymbol{\lambda})}{\partial \lambda_i} = x f_i((1 + \lambda_i)x) \prod_{i \in \mathcal{K}: j \neq i, k} F_j((1 + \lambda_j)x) \leq \bar{v}_k \bar{f}.$$

Therefore, we have

$$\left| \frac{\partial G_k^{\mathrm{HP}}(\mu_k, \boldsymbol{\lambda})}{\partial \lambda_k} \right| \leq \bar{v}_k^2 \bar{f}.$$

We next consider $V_k^{\mathrm{HP}}(\mu_k, \boldsymbol{\lambda})$ and its derivatives.

$$V_k^{\mathrm{HP}}(\mu_k, \boldsymbol{\lambda}) = \int_x x L_k \left(\frac{x}{1 + \lambda_k} \right) H \left(\frac{x}{1 + \mu_k} \right) dF_k(x). \tag{58}$$

1. We have that

$$\frac{\partial V_k^{\mathrm{HP}}(\mu_k, \boldsymbol{\lambda})}{\partial \mu_k} = - \int_x \left(\frac{x}{1 + \mu_k} \right)^2 L_k \left(\frac{x}{1 + \lambda_k} \right) h \left(\frac{x}{1 + \mu_k} \right) dF_k(x),$$

and thus,

$$\left| \frac{\partial V_k^{\mathrm{HP}}(\mu_k, \boldsymbol{\lambda})}{\partial \mu_k} \right| \leq \bar{v}_k^2 \bar{h}.$$

2. We have

$$\frac{\partial V_k^{\mathrm{HP}}(\mu_k, \boldsymbol{\lambda})}{\partial \lambda_k} = - \int_x \left(\frac{x}{1 + \lambda_k} \right)^2 l_k \left(\frac{x}{1 + \lambda_k} \right) H \left(\frac{x}{1 + \mu_k} \right) dF_k(x),$$

and thus,

$$\left| \frac{\partial V_k^{\mathrm{HP}}(\mu_k, \boldsymbol{\lambda})}{\partial \lambda_k} \right| \leq (K - 1) \bar{v}_k^2 \bar{f}.$$

3. For $i \neq k$, the derivative with respect to λ_i is

$$\frac{\partial V_k^{\mathrm{HP}}(\mu_k, \boldsymbol{\lambda})}{\partial \lambda_i} = \int \frac{x^2}{1 + \lambda_k} f_i \left(\frac{1 + \lambda_i}{1 + \lambda_k} x \right) \prod_{j \neq i, k} F_j \left(\frac{1 + \lambda_j}{1 + \lambda_k} x \right) H \left(\frac{x}{1 + \mu_k} x \right) dF_k(x),$$

and thus,

$$\left| \frac{\partial V_k^{\mathrm{HP}}(\mu_k, \boldsymbol{\lambda})}{\partial \lambda_i} \right| \leq \bar{v}_k^2 \bar{f}.$$

$\square$

Corollary E.3. $U_k(\boldsymbol{\lambda})$, $U_k^{\mathrm{CP}}(\boldsymbol{\lambda})$ and $U_k^{\mathrm{HP}}(\mu_k, \boldsymbol{\lambda})$ are Lipschitz continuous.

Proof. The results directly follow from Lemma E.1 and Lemma E.2 since

$$\frac{1}{T} U_k(\boldsymbol{\lambda}) = V_k(\boldsymbol{\lambda}) - G_k(\boldsymbol{\lambda});$$

$$\frac{1}{T} U_k^{\mathrm{CP}}(\boldsymbol{\xi}) = V_k^{\mathrm{CP}}(\boldsymbol{\xi}) - G_k^{\mathrm{CP}}(\boldsymbol{\xi});$$

$$\frac{1}{T} U_k^{\mathrm{HP}}(\mu_k, \boldsymbol{\lambda}) = V_k^{\mathrm{HP}}(\mu_k, \boldsymbol{\lambda}) - G_k^{\mathrm{HP}}(\mu_k, \boldsymbol{\lambda}).$$

$\square$

Lemma E.4. *Suppose that Assumption 4.1 holds. We have*

$$G_k^{\mathrm{HP}}(\lambda_k^*, \boldsymbol{\lambda}^*) \le G_k(\boldsymbol{\lambda}^*), \forall k,$$

and the equality strictly holds for at most one bidder.

Proof. In what follows, we use $L_k(x)$ to denote the cumulative distribution function of d_k^I defined by (54) and $l_k(x)$ is the corresponding density function. We analyze how much $G_k^{\mathrm{HP}}(\lambda_k, \boldsymbol{\lambda})$ and $G_k(\boldsymbol{\lambda})$ differ by their expressions.

$$
\begin{aligned}
G_k^{\mathrm{HP}}(\lambda_k, \boldsymbol{\lambda}) &\overset{(a)}{=} \int_x (1+\lambda_k) f_k((1+\lambda_k)x) L_k(x) \left(\int_{y \le x} y \, dH(y) \right) dx \\
&\overset{(b)}{=} \int_x (1+\lambda_k) x f_k((1+\lambda_k)x) L_k(x) H(x) \, dx \\
&\quad - \int_x (1+\lambda_k) f_k((1+\lambda_k)x) L_k(x) \left(\int_{y \le x} H(y) \, dy \right) dx \\
&\overset{(c)}{=} \int_x \left((1+\lambda_k) x f_k((1+\lambda_k)x) - \bar{F}_k((1+\lambda_k)x) \right) L_k(x) H(x) \, dx \\
&\quad - \int_x \bar{F}_k((1+\lambda_k)x) l_k(x) \left(\int_{y \le x} H(y) \, dy \right) dx \\
&\overset{(d)}{=} G_k(\boldsymbol{\lambda}) - \underbrace{\int_x \bar{F}_k((1+\lambda_k)x) l_k(x) \left(\int_{y \le x} H(y) \, dy \right) dx}_{\Delta_k(\boldsymbol{\lambda})}.
\end{aligned}
$$

where (a) follows from (56) and changes the order of integration; (b) and (c) use integration by parts; and (d) follows from (53). Observe that $\Delta_k(\boldsymbol{\lambda})$ is always non-negative. Moreover, we claim that when Assumption 4.1 holds, $\Delta_k(\boldsymbol{\lambda}^*)$ must be strictly positive for all but one bidder, by which the proof can be completed.

To prove the claim, we first show that for every bidder $k \in \mathcal{K}$,

$$H\left(\frac{\hat{v}_k}{1+\lambda_k^*} \right) > 0, \tag{59}$$

where $\hat{v}_k := \sup\{v_k : \bar{F}_k(v_k) > 0\}$. Otherwise, $H(\hat{v}_k/(1+\lambda_k^*)) = 0$ implies that bidder k can, almost surely, never win the auction by bidding $v_k/(1+\lambda_k^*)$ or lower when bidders are independently bidding. Note that $\lambda_k^* < \bar{\mu}_k$ because

$$G_k(\bar{\mu}_k, \boldsymbol{\lambda}_{-k}^*) = \mathbb{E}\left[\mathbf{1}\left\{ \frac{v_k}{1+\bar{\mu}_k} \ge d_k \right\} d_k \right] \overset{(a)}{\le} \mathbb{E}\left[\mathbf{1}\left\{ \frac{v_k}{1+\bar{v}_k/\rho_k} \ge d_k \right\} d_k \right] \le \frac{v_k}{1+\bar{v}_k/\rho_k} < \rho_k,$$

where (a) follows from $\bar{\mu}_k \ge \bar{v}_k/\rho_k$. Thus, $G_k((\lambda_k, \boldsymbol{\lambda}_{-k}^*)) \equiv 0$ on $[\lambda_k^*, \bar{\mu}_k]$ contradicts the strong monotonicity assumption.

Let $k' = \arg\min_k \hat{v}_k/(1+\lambda_k^*)$ and let $\underline{d}^O := \inf\{d^O : H(d^O) > 0\}$. Note that by (59) and the continuity of H, we must have $\underline{d}^O < \hat{v}_{k'}/(1+\lambda_{k'}^*)$. For every $k \ne k'$, when $x \in (\underline{d}^O, \hat{v}_{k'}/(1+\lambda_{k'}^*))$, we have

1. $\bar{F}_k((1+\lambda_k^*)x) > 0$;

2. $\int_{y \le x} H(y) \, dy > 0$.

Suppose for contradiction that for $k \ne k'$, $\Delta_k(\boldsymbol{\lambda}^*) = 0$. Then $l_k(x) = 0$ holds almost everywhere on $[\underline{d}^O, \hat{v}_{k'}/(1+\lambda_{k'}^*)]$, i.e., the competing bid d_k^I lies in this interval with zero probability. However, $b_{k'} = v_{k'}/(1+\lambda_{k'}^*)$ belongs to $[\underline{d}^O, \hat{v}_{k'}/(1+\lambda_{k'}^*)]$ with positive probability by the definition of $\hat{v}_{k'}$. This implies there is almost always another bid higher than $b_{k'}$ so that $b_{k'}$ almost never becomes d_k^I, which further implies that bidder k' actually can, almost surely, never win in IP by bidding $v_{k'}/(1+\lambda_{k'}^*)$ or lower. Again by a similar argument, this contradicts the strong monotonicity assumption. Therefore, $\Delta_k(\boldsymbol{\lambda}^*) > 0$ holds at least for all bidders except bidder k'. $\square$
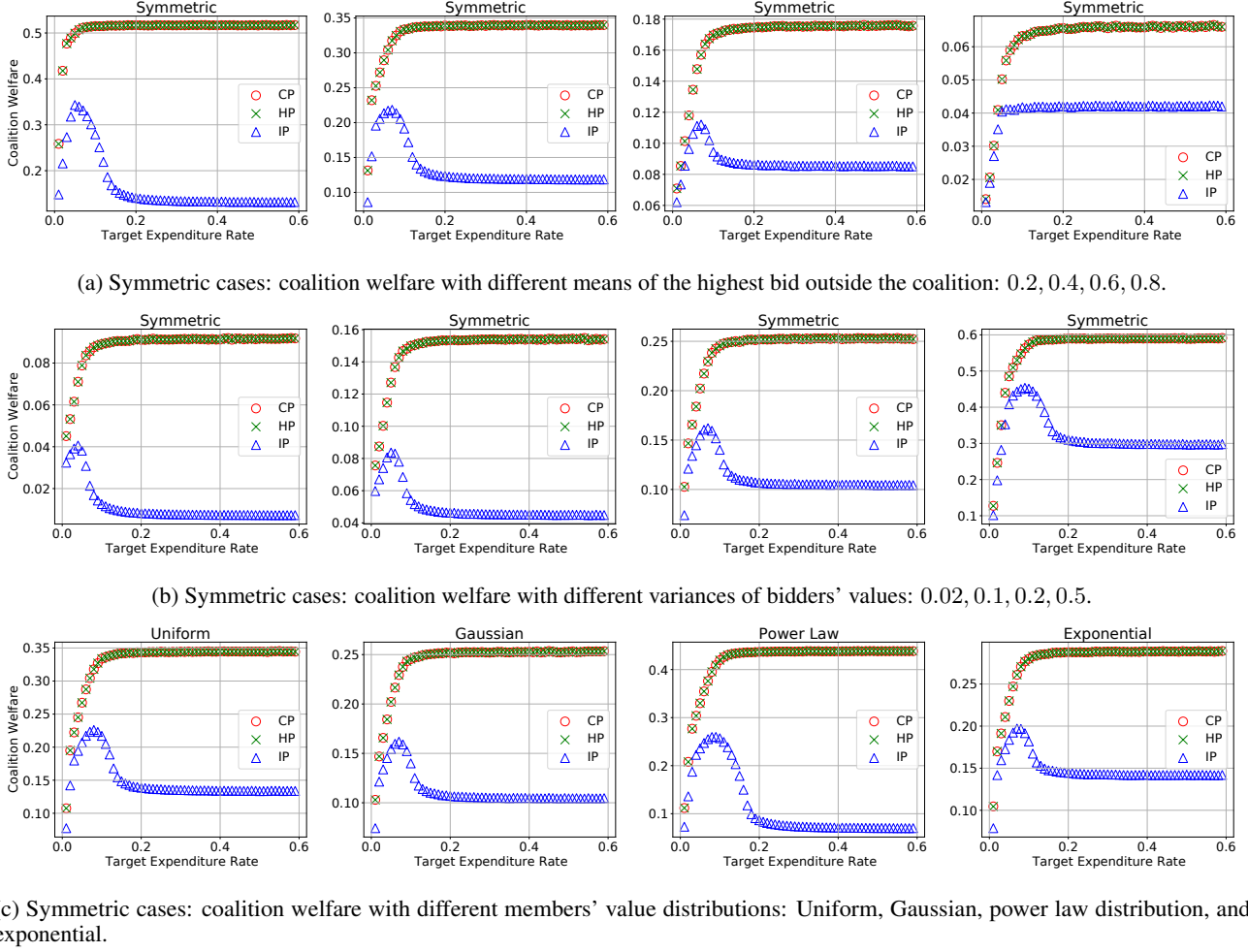
(a) Symmetric cases: coalition welfare with different means of the highest bid outside the coalition: $0.2, 0.4, 0.6, 0.8$.

(b) Symmetric cases: coalition welfare with different variances of bidders' values: $0.02, 0.1, 0.2, 0.5$.

(c) Symmetric cases: coalition welfare with different members' value distributions: Uniform, Gaussian, power law distribution, and exponential.

Figure 3. Experiment Results on Symmetric Synthetic Data

F. More Experiment Results on Synthetic Data

We present more experiment results on synthetic data here. All the following cases have the same number of coalition members: $K = 5$.
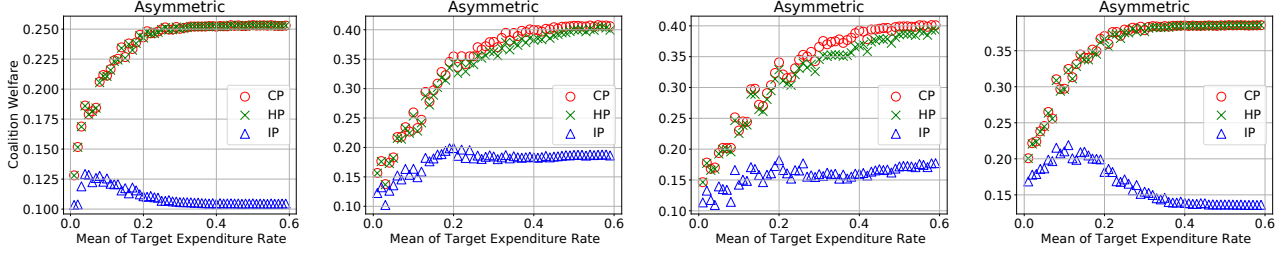
Figure 3 show the results on symmetric cases, where all 5 members have the same value distribution and the same budget. Specifically, Figure 3a shows the coalition welfare when the distribution of the highest bid outside the coalition d^O varies. From left to right, d^O is sampled from Gaussian distribution with μ equals to $0.2, 0.4, 0.6, 0.8$ respectively. With the mean increases, bidders' power outside the coalition becomes stronger.

In Figure 3b, the coalition members' value distributions are all Gaussian distributions with $\mu = 0.5$ but they have different standard deviations σ. From left to right, σ equals to $0.02, 0.1, 0.2, 0.5$ respectively.

The four subfigures in Figure 3c compare the coalition welfare when members' values are sampled from different distributions: uniform distribution, Gaussian distribution, power law distribution, and exponential distribution.

Figure 4 shows four asymmetric cases. Different from above, the target expenditure rates for coalition members have the same expectation rather than being completely the same. From left to right, coalition member's values are sampled from totally the same distributions, same distributions with different means, and completely different distributions, respectively.

In Figure 4a, each member's value is sampled from Gaussian distribution with $\mu = 0.5, \sigma = 0.2$. In Figure 4b, the five members' value are respectively sampled from Gaussian distributions with μ equals to $0.1, 0.3, 0.5, 0.7, 0.9$. In Figure 4c,

(a) Gaussian with same means. (b) Gaussian with different means. (c) Uniform with different means. (d) Different value distributions.

Figure 4. Experiment Results on Asymmetric Synthetic Data

the five members' value are respectively sampled from uniform distributions on $[0, 0.2]$, $[0.2, 0.4]$, $[0.4, 0.6]$, $[0.6, 0.8]$, $[0.8, 1]$. In Figure 4d, the value functions are more complicated. We choose one uniform distribution, one Gaussian distribution, two power law distributions and one exponential distribution. We adjust the parameters so that the value expectation of all members is 0.5, which is the same as the others in Figure 4.

To verify if the averages of the multipliers indeed converge to the profile defined by the non-linear complementarity conditions (2), (10) and (8), we check the variances of the averages and the empirical value of $(\sum_{t=1}^{T} \lambda_{k,t}/T)(\rho_k - G_k(\sum_{t=1}^{T} \boldsymbol{\lambda}_t/T))$ $((\sum_{t=1}^{T} \xi_{k,t}/T)(\rho_k - G_k^{\mathrm{CP}}(\sum_{t=1}^{T} \boldsymbol{\xi}_t/T))$ and $(\sum_{t=1}^{T} \mu_{k,t}/T)(\rho_k - G_k^{\mathrm{HP}}(\sum_{t=1}^{T} \boldsymbol{\lambda}_t/T, \sum_{t=1}^{T} \mu_{k,t}/T)))$. To approximate $G_k(\sum_{t=1}^{T} \boldsymbol{\lambda}_t)$, we use the empirical average expenditure $\sum_{t=1}^{T} z_{k,t}/T$. The same goes for $G_k(\sum_{t=1}^{T} \boldsymbol{\xi}_t)$ and $G_k(\sum_{t=1}^{T} \boldsymbol{\xi}_t)$. In the experiments, the variances of the last 100 rounds of $\sum \lambda_{k,t}/T$ ($\sum \xi_{k,t}/T$ and $\sum \mu_{k,t}/T$) are around 10^{-4}, indicating the convergence of multipliers. The values of the last 100 rounds of $(\sum \lambda_{k,t}/T)(\rho_k - \sum z_{k,t}/T)$ $((\sum \xi_{k,t}/T)(\rho_k - \sum z_{k,t}/T)$ and $(\sum \mu_{k,t}/T)(\rho_k - \sum z_{k,t}/T))$ are less than 10^{-3}, indicating the convergence to the point defined by (2), (10) and (8).