Incentivize Contribution and Learn Parameters Too: Federated Learning with Strategic Data Owners

Drashthi Doshi Avishek Ghosh Swaprava Nath
Indian Institute of Technology Bombay
drashthi@iitb.ac.in, avishek_ghosh@iitb.ac.in, swaprava@iitb.ac.in

Abstract

In federated learning (FL), a center trains a shared model from updates that clients compute on their private data. Collecting and training on data is costly, so a strategic client may contribute little data and free-ride on the model trained by the others. We study mechanisms that incentivize data contribution and learn the model at the same time. Our first mechanism, `UPBReD`, updates the clients' contributions and the model jointly, and converges at a geometric rate to a Nash equilibrium of the contribution game together with the welfare-maximizing model for those contributions. This equilibrium can involve under-contribution, so our second mechanism, `2P-UPBReD`, adds budget-balanced transfers that make full contribution a dominant strategy, after which training converges to the optimal model; we also give conditions under which every client prefers to participate. Both mechanisms trust what clients report about the data they use and the gradients they compute. We show that no payment rule can reward unverifiable contributions unless something the center observes depends on the data actually used. Building on this, `T-2P-UPBReD` checks each client's reported contribution against the other clients' realized valuations of the trained model, and prices gradient reports with a budget-balanced charge. It elicits truthful contribution reports, makes gradient reports approximately truthful, still reaches full contribution and the optimal model, and never runs a deficit. Experiments on CIFAR-10, FEMNIST, and Twitter show that `2P-UPBReD` matches the welfare of FedAvg trained on all data, while strategic clients without incentives lower the welfare of the learned model.

1 Introduction

Accurate machine learning models need large amounts of training data, but in many applications, such as language, vision, and disease modeling, the data are spread across many owners and cannot be pooled. Federated learning (FL) [22, 28, 35, 57] trains a shared model without moving the data: a center, the parameter server, maintains the model, and each data owner, which we call an agent, computes updates on its local data and sends only these updates to the center. Since an agent typically holds too little data to learn a good model alone, every agent benefits from the data of the others.

Classical FL assumes that agents contribute all their data voluntarily. In practice, collecting, cleaning, and training on data is costly, while the benefit of a better model is shared by everyone. A rational agent may therefore contribute little data and *free-ride* on the model trained by the others [27]. Consider hospitals in a region that jointly train a disease model on patient records. Sampling and annotating records is expensive, and hospitals differ in how much data they can collect, so a hospital may prefer to contribute little and still receive the shared model. If many hospitals reason this way, the model is trained on little data, and every hospital is worse off. This raises the question:

"Can a mechanism incentivize every agent to contribute all of its data and, at the same time, learn the optimal model, even though the center cannot verify how much data an agent actually uses?"

Recent work on incentives in FL [2, 27, 38] abstracts training away: the value of a contribution profile is given by a function of the profile, and the model is not learned inside the mechanism. In FL, however, contributions and the model evolve together, round by round, and the center sees only what the agents report, namely how much data they used and the gradients they computed. Neither report can be verified, and an agent may misreport either one to its advantage. This paper treats three questions

	Eq.	Learn.	Transfers	Reports	Truthful	Contrib.	Private
Procaccia et al. [42]	E	✗	BB	D	DSIC	P	✗
Bornstein et al. [3]	E	✗	✓	C	DSIC	F	✓
Blum et al. [2]	E	✗	✗	S	✗	P	✓
Murhekar et al. [38]	C	✗	BB	S,C	✗	P	✓
Karimireddy et al. [27]	E	✗	✓	S,C	✗	F	✓
`UPBReD` (Thm. 1)	C	J	✗	S,G	✗	P	✓
`T-UPBReD` (Thm. 4)	C	J	BB*	S,G	EPIC (S)†, ε (G)	P	✓
`2P-UPBReD` (Thm. 2, Prop. 1)	C	OPT	BB, IR	S,G	✗	F	✓
`T-2P-UPBReD` (Thm. 3)	C	OPT	BB*	S,G	EPIC (S)†, ε (G)	F	✓

Table 1: ***Comparison with incentive mechanisms for FL; our mechanisms are in the last four rows.*** *Eq.: existence (E) of, or convergence (C) to, a Nash equilibrium. Learn.: joint convergence of the model and the contributions (J), or convergence to the optimal parameter $w^{\texttt{OPT}}$ (OPT). Transfers: budget balanced (BB), budget balanced on the truthful path and never in deficit (BB*), individually rational (IR). Reports: data distribution (D), cost (C), contribution (S), gradient (G). Truthful: guarantee for the reported quantity; ε denotes ε-truthfulness with a vanishing gain from misreporting, and † marks guarantees that rely on agents observing realized valuations (Assumption 5); Proposition 2 shows that evidence of this kind is necessary. Contrib.: partial (P) or full (F) contribution at equilibrium. Private: data are not shared with the center.*

together: how contributions and the model evolve under joint dynamics, how to remove the inefficiency of the resulting equilibrium with transfers that sum to zero, and how to make the agents' reports truthful.

1.1 Our contributions

- **Joint dynamics of contributions and learning.** We model FL as a game in which each agent's valuation depends on the learned model and on all agents' contributions. In `UPBReD` (Algorithm 1), agents adjust their contributions by projected gradient steps while the center updates the model from their gradients. Under smoothness and concavity conditions, `UPBReD` converges at a geometric rate to a point at which the contributions form a Nash equilibrium and the model maximizes welfare for those contributions (Theorem 1). The best-response dynamics of Murhekar et al. [38] converge in the contributions alone, with no model being trained.
- **Free-riding at equilibrium, and a budget-balanced remedy.** The equilibrium of `UPBReD` can involve under-contribution: in Example 1, both agents contribute 7 of their 9 data points and the social welfare falls from 3 to 2.75. `2P-UPBReD` (Algorithm 2) adds a linear transfer that pays each agent in proportion to how much its contribution exceeds the average contribution of the other agents, and charges it when the difference is negative. The transfers sum to zero and make full contribution a dominant strategy, and training then converges to $w^{\texttt{OPT}}$ at a linear rate (Theorem 2). We give an explicit range of the transfer rate under which every agent prefers participating to training alone, and zero-sum rebates that restore participation otherwise (Proposition 1).
- **Truthful contribution requires evidence.** `2P-UPBReD` pays for reported contributions, which the center cannot verify. We show that no transfer rule can reward contributions unless something the center observes depends on the data actually used (Proposition 2), a reporting analogue of the informativeness principle of Holmström [20]. In particular, checking an agent's reports only against each other cannot make the subsidy of `2P-UPBReD` truthful.
- **A truthful mechanism that never runs a deficit.** `T-2P-UPBReD` (Algorithm 3) obtains this evidence from the other agents: after training, it checks each agent's reported contribution against the other agents' realized valuations of the trained model, and it prices gradient reports with a budget-balanced charge derived from VCG on a surrogate welfare. Truthful contribution and valuation reports are ex-post incentive compatible, the gain from misreporting a gradient vanishes as training converges,

full contribution remains a dominant strategy, the model converges to $w^{\texttt{OPT}}$ whether agents report gradients truthfully or play the equilibrium of each round, and the transfers sum to zero on the truthful path and never to a deficit (Theorem 3). Applied in every round, the same transfers make the reports of `UPBReD` truthful as well (`T-UPBReD`, Theorem 4).

- **Experiments.** On CIFAR-10, FEMNIST, and Twitter, with deep networks that lie beyond the scope of our theory, `2P-UPBReD` matches the welfare of `FedAvg` trained on all data, already for small transfer rates, while `UPBReD`, a strategic variant of FedAvg (`FedAvgStrategic`), and `FedBR-BG` [38] reach lower welfare (Section 6).

Table 1 compares these guarantees with those of prior incentive mechanisms for FL. Proofs, additional remarks, and further experiments are in the appendix.

1.2 Related work

Federated learning [28] has received wide attention over the last decade, largely owing to the FedAvg algorithm [35], which reduces communication by taking several local steps between aggregations. Subsequent work addressed data heterogeneity [14, 26], Byzantine robustness [16, 25, 55], communication overhead [15, 24, 45], and privacy [30, 50, 51].

Incentives through the model. Several works reward participation through the model itself: by weighting client updates so that the global model outperforms local ones [8], by valuing contributions with the Shapley value [49], by contract-theoretic accuracy shaping that elicits full contribution [27], by excluding malicious clients and rewarding helpful ones [12], and by differentiated model rewards during training [53]. Procaccia et al. [42] maximize utilitarian welfare in a heterogeneous setting under PAC learning constraints.

Monetary incentives. Yu et al. [56] allocate a dynamic budget to maximize utility and reduce inequality. Blockchain-based methods reward high-quality contributions under budget constraints [41] or distribute profits by the Shapley value [32], and client rewards have been shown to improve the final model [54]. Georgoulaki and Kollias [13] show a price of anarchy of two and a price of stability of one for utility sharing with budget-balanced payments. In auction-based designs, agents adjust their bids by reinforcement learning [47], and the center optimizes its budget to improve utility and reduce delays [48].

Coalitions and free-riding. Donahue and Kleinberg [9] model FL as a coalitional game, Ray Chaudhury et al. [43] ensure core stability with CoreFed, later extended to ordinal utilities [6], and Murhekar et al. [39] use the Shapley value for reciprocal fairness. Meng et al. [36] study how free-riding among competing firms affects the incentive to form FL coalitions, Chen et al. [7] and Tan et al. [46] propose conflict-aware collaboration among competitors, and Mao et al. [33] show that subgame perfect equilibria of repeated FL games can be inefficient and design budget-balanced mechanisms that restore efficiency and individual rationality.

Truthful mechanisms and contribution games. Bornstein et al. [3] elicit truthful cost reports, and Chakarov et al. [5] give a budget-balanced, Bayesian incentive compatible payment rule for heterogeneous agents. Closest to our work, Blum et al. [2] and Murhekar et al. [38] study equilibria of contribution games, the latter with best-response dynamics and budget-balanced payments. These works do not train the model inside the mechanism, and they treat contributions as observable. We analyze contributions and training jointly, and we make both the unverifiable contribution reports and the gradient reports incentive compatible, building on two-stage mechanisms for interdependent valuations [37, 40].

2 Preliminaries

Consider an FL setup where a set of data contributors, given by $N = \{1, 2, \ldots, n\}$, is interacting with a center. Each data contributor (agent) i has access to a private labeled dataset D_i sampled from distribution $\mathcal{D}_i$, with the size of the dataset given by $s_i^{\max} = |D_i|$. The agents are interested in learning a common parameter $w \in \mathbb{R}^m$ from their data for a common learning task (e.g., to perform a classification

task). As the agents have limited individual data, the model is trained on the consolidated data of all the n users. However, sampling data is costly, and agent i incurs a cost, given by a function $c_i(s_i)$, when it trains the model locally on its chosen dataset of size $s_i \in S_i := [s_i^{\min}, s_i^{\max}]$, where $s_i^{\min}$ is the smallest amount of data that the agent can contribute. Along with these costs, there may be some other arbitrary costs that an agent incurs based on the mechanism and other agents' contributions. Under this setup, each agent i gets a utility based on how much data s has been chosen by the agents to sample and train on. Hence, the utility is given by

$$u_i(w, s_i, s_{-i}) = v_i(w, s_i, s_{-i}) - z_i(s_i, s_{-i}), \tag{1}$$

where s_{-i} is the data chosen by the agents other than i in this FL process. The function $v_i(w, s_i, s_{-i})$ is the *valuation function*, which denotes the *benefit* to agent i if the parameter learned by the center is w and the agents contribute by running the FL algorithm on their dataset sizes given by the vector $s = (s_i, s_{-i})$. The function $z_i(s_i, s_{-i})$ is the *effective cost* to agent i, which may depend on the data contributions of all agents.

FL is an iterative process: the center updates a parameter w^t at round t, and agents use this to choose their dataset size s_i^{t+1} and to compute and share gradients $d_i^{t+1} \in \mathbb{R}^m$ with the center. The center aggregates gradients to update the parameter to w^{t+1}. We use the superscript t when referring to these terms in a specific iteration of the training round. Agents are myopic: in each round, agent i updates its contribution by a gradient step on its current-round utility given the broadcast (w^t, s^t), and does not anticipate the effect on future models. Myopic adjustment is a standard model of learning in games [11].

Remark (Arguments of the u_i function). Note that we have explicitly assumed that the utility function u_i depends on the parameter learned and the data contributions (w, s_i, s_{-i}). We now motivate this dependence. In FL problems, we typically run iterative algorithms to optimize the accuracy (or loss) function. Note that, with this, the weight w^t at time t depends on all the s_i^ℓ, $\ell < t, \forall i \in N$. However, the utility u_i at time t depends on (s_i^t, s_{-i}^t) as well, which is not captured through w^t. Hence, we require the said explicit dependencies. We represent this dependency via two functions: the valuation function v_i gives the benefit the agent gets from the parameter learned and the data contributions, while the effective cost is only dependent on all players' data contributions.

Remark (Realizing the v_i function). $v_i(w, s)$ denotes the benefit agent i derives from the learned parameters at a given contribution vector. It can be realized by an agent as the negative of the loss computed on its data for the given parameter w. For experiments in Section 6, we use the cross-entropy loss.

Since the agents are strategic, every agent i's aim is to maximize its utility by appropriately choosing its strategy s_i given the strategies s_{-i} of the other players and the parameter w chosen by the center. The center, on the other hand, is interested in learning the optimal parameter w that maximizes the sum of the valuations as follows, when all agents contribute their maximum data sizes, i.e., $s_i^{\max}, i \in N$.

$$w^{\texttt{OPT}} \in \operatorname*{argmax}_w \sum_{i \in N} v_i(w, s^{\max}) \tag{2}$$

As in prior work on data sharing in federated learning [27], the center's goal is the quality of the learned model, that is, the aggregate valuation $\sum_{i \in N} v_i$ obtained when all available data is used; the agents' data-collection costs are private and are not part of this objective. We therefore use *social welfare* to mean this learning welfare, not the net utility. We are interested in the question of whether we can design an FL algorithm that can make $(s_i^{\max}, s_{-i}^{\max})$ a Nash equilibrium of the underlying game, and the center can learn $w^{\texttt{OPT}}$. In this context, the Nash equilibrium is defined as follows.

Definition 1 (Nash equilibrium)**.** *A* pure strategy Nash equilibrium *(PSNE) for a given parameter w is a strategy profile (s_i^*, s_{-i}^*) of the agents such that $u_i(w, s_i^*, s_{-i}^*) \geqslant u_i(w, s_i, s_{-i}^*),\ \forall s_i \in S_i, \forall i \in N$.*

This definition is a modification of the standard definition of Nash equilibrium since the equilibrium profile (s_i^*, s_{-i}^*) depends on the parameter w (we do not write it explicitly for notational cleanliness). Consequently, we have the limit pair (w^*, s^*) as a fixed point where s^* is a PSNE of the game parameterized by w^*, and $w^* \in \arg\max_w \sum_{i \in N} v_i(w, s^*)$. We will refer to the term PSNE as Nash equilibrium (NE) in the rest of the paper.

In order to ensure convergence towards the optimal parameter, we introduce monetary transfers. We also use monetary transfers to satisfy two design objectives, defined below: incentive compatibility, so that truthful reporting is optimal for the agents, and budget balance, so that the transfers only redistribute money among the agents.

In every round, agent i takes an *action*, the contribution $s_i^t \in S_i$ that it actually uses, and sends *reports* to the center: a contribution report $\hat{s}_i^t$, a gradient report $\hat{d}_i^t \in \mathbb{R}^m$, and, in the truthful mechanisms, a report $\hat{v}_i$ of its valuation of a model it has observed. The center computes the model from the reported gradients, $w^{t+1} = w^t + \frac{\eta}{n} \sum_{j \in N} \hat{d}_j^t$, as in Equation (9), and computes any transfers p_i from the reports. Reporting is *truthful* if $\hat{s}_i^t = s_i^t$, if $\hat{d}_i^t = \nabla_w v_i(w^t, s^t)$ is the gradient computed on the data actually used, and if $\hat{v}_i = v_i(w, s)$ is the realized valuation of the evaluated model w at the contributions s actually used. Until Section 5.3 we assume truthful reporting. When reports may differ from the truth, agent i's utility is

$$u_i = v_i(w, s) - c_i(s_i) + p_i(\hat{s}, \hat{d}, \hat{v}), \tag{3}$$

where w is the model computed from the reports and s is the profile of contributions actually used. The valuation and the cost are realized by the agent and are evaluated at the actual contributions, whereas the transfers are computed by the center and are evaluated at the reports. Under truthful reporting, Equation (3) reduces to Equation (1) with $z_i = c_i - p_i$.

Truthfulness notions. Valuations here are *interdependent*: agent i's realized valuation depends on the model, which is computed from all agents' reports, and on the contributions actually used by all agents. With interdependent valuations, dominant-strategy implementation of efficient outcomes is generally impossible [21], so we target ex-post incentive compatibility, the standard goal in this setting. Following the two-stage construction of Mezzetti [37], our truthful mechanisms first collect contribution and gradient reports, compute the model, and then elicit the agents' realized valuations, which enter the transfers. A *reporting stage* is a step of a mechanism at which agents send reports; as elsewhere in the paper, agents are myopic, so the utility relevant to a report is the one realized at the stage in which it is sent.

Definition 2 (Ex-post incentive compatibility)**.** *A reporting stage of a mechanism is* ex-post incentive compatible *(EPIC) if, for every agent $i \in N$, every history of the mechanism up to that stage, and every report of agent i at that stage, agent i's utility Equation (3) when all agents report truthfully is at least its utility when it sends that report and all other agents report truthfully.*

Definition 3 (Approximate truthfulness)**.** *A reporting stage is* ε-truthful *for agent i if, for every history up to that stage, whenever all other agents report truthfully, no report of agent i raises its utility Equation (3) by more than ε over truthful reporting.*

In `T-2P-UPBReD`, the elicitation stage is EPIC for contribution and valuation reports, and every round of training is ε_i^t-truthful for gradient reports (Theorem 3). We use Equation (3) only when establishing truthfulness, in Section 5.3 and Section 5.4; elsewhere reports are truthful and we use the simpler form Equation (1).

Besides incentive compatibility, we want the transfers to only redistribute money among the agents. Transfers are computed from the reports, not from the actions the agents actually take, so we require this for every profile of admissible reports (for example, $\hat{s}_i \in S_i$), not only for truthful ones.

Definition 4 (Budget balance)**.** *Let r denote the profile of reports sent during a run of the mechanism (contribution reports $\hat{s}$, gradient reports $\hat{d}$, and valuation reports $\hat{v}$, as applicable), and let $p_i(r)$ be the total transfer paid* to *agent i over the run. The mechanism is*

- budget balanced *(BB) if $\sum_{i \in N} p_i(r) = 0$ for every report profile r;*
- weakly budget balanced *(W-BB) if $\sum_{i \in N} p_i(r) \leqslant 0$ for every report profile r, that is, the center never runs a deficit.*

Budget balanced ensures that the net monetary inflow or outflow is zero and the mechanism only allows monetary redistribution among the agents while weak budget balanced ensures that the center is not running a deficit. In the following section we discuss some realistic forms of the utility function along with some properties that satisfy the assumptions we make in future sections.

3 Examples of valuation functions

We now give a concrete form of the utilities that fits our framework and satisfies our assumptions. All agents learn the same model, parameterized by $w \in \mathbb{R}^m$, and each agent i contributes $s_i \in S_i = [1, s_i^{\max}]$, that is, $s_i^{\min} = 1$. A natural form of agent i's utility is

$$u_i(w, s) = \underbrace{r_i - L_i(w, s)}_{\text{valuation}} - \underbrace{c_i \cdot s_i}_{\text{cost}} + \underbrace{p_i(s)}_{\text{payments}}, \tag{4}$$

where $L_i(w, s)$ is agent i's loss when the model is w and the agents contribute s, $c_i > 0$ is a per-sample cost, and r_i is a constant large enough to keep utilities non-negative over a bounded set containing the iterates of our algorithms. A lower loss thus means a higher valuation $v_i(w, s) = r_i - L_i(w, s)$. Motivated by the additive scaling-law form of Hoffmann et al. [19] and Kaplan et al. [23], we model

$$L_i(w, s) = L_i(w) + \frac{1}{n} \sum_{j \in N} s_j^{-\alpha_i}, \qquad \alpha_i > 0, \tag{5}$$

where the first term is the loss achievable at the parameter w and the second is a finite-sample correction that decays as each agent contributes more data. Since the center aggregates the agents' updates with uniform weights $\frac{1}{n}$ (Equation (9)), the finite-sample error decomposes across data sources: specializing the result of Donahue and Kleinberg [9] to uniform weights yields a term proportional to $\sum_{j \in N} s_j^{-1}$, which is the case $\alpha_i = 1$ of Equation (5).

The first term $L_i(w)$ can be any standard loss, for example, the regularized cross-entropy and mean squared error losses of linear models, or the quadratic distance to the parameter that best fits agent i's own data. These functions satisfy the properties that the convergence results of Sections 4 and 5 require of the learning problem; the losses and the resulting constants are given below. The quadratic distance also makes concrete the knowledge of the valuations that the truthful mechanism of Section 5.3 requires of the center: agent i's valuation is determined by A_i, w_i^* and α_i, and truthful gradients identify A_i and w_i^* (see below). In `2P-UPBReD` the payments are $p_i = p_i^1$ of Equation (12); the transfers of the truthful mechanisms are defined in Section 5.3 and Section 5.4.

This decomposition, with linear models, serves only to show that our assumptions are satisfiable under a realistic functional form. Our experiments (Section 6) use deep networks and compute L_i directly as the empirical loss on each agent's data, with the cross-entropy loss, without imposing this form.

We now describe the losses $L_i(w)$ referred to above. For each, we record the two properties that our assumptions use, namely a Hessian that is bounded for every $w \in \mathbb{R}^m$ and strong convexity; the resulting constants are derived in the verification remarks of Assumption 1 and Assumption 3.

Cross-entropy loss. For classification tasks with K classes, the regularized cross-entropy loss for agent i is

$$L_i(w) = \mathbb{E}_{(x,y) \sim \mathcal{D}_i} \big[-\log p(y \mid x; w) \big] + \frac{\mu}{2} \|w\|_2^2, \tag{6}$$

where $p(y \mid x; w)$ is the softmax output of a linear model and $\mu > 0$ is the regularization weight. Under the standard assumption that inputs are bounded, $\|x\|_2 \leqslant B_x$, the Hessian of L_i is bounded for every $w \in \mathbb{R}^m$, and the regularization term makes L_i μ-strongly convex. Both properties are standard in the federated learning literature [26].

Mean squared error loss. For regression tasks with a linear predictor $\hat{y}(w, x) = w^\top x$, the regularized MSE loss for agent i is

$$L_i(w) = \mathbb{E}_{(x,y) \sim \mathcal{D}_i} \big[\big(y - \hat{y}(w, x) \big)^2 \big] + \frac{\mu}{2} \|w\|_2^2, \tag{7}$$

whose Hessian $2\,\mathbb{E}[xx^\top] + \mu \mathbb{I}$ is bounded under $\|x\|_2 \leqslant B_x$, and whose regularization term again makes it μ-strongly convex.

Distance to the local optimum. Let w_i^* denote the parameter that minimizes agent i's loss on its own data, and let $A_i \succ 0$ be a positive definite matrix. The loss

$$L_i(w) = (w - w_i^*)^\top A_i \, (w - w_i^*) \tag{8}$$

measures how far the learned parameter is from the one that best fits agent i's data. The MSE loss of Equation (7) is of this form up to an additive constant, with $A_i = \mathbb{E}_{x\sim\mathcal{D}_i}[xx^\top] + \frac{\mu}{2}\mathbb{I}$ and w_i^* the regularized least-squares solution on $\mathcal{D}_i$; for other losses, such as Equation (6), it is the local quadratic approximation of L_i around w_i^*. Since this loss carries no regularization term, Assumption 3 then holds with $\nu = 2\lambda_{\min}(\bar{A})$ and $M = 2\lambda_{\max}(\bar{A})$, where $\bar{A} := \frac{1}{n}\sum_{i\in N} A_i$, rather than with the regularization weight μ as for the losses Equations (6) and (7).

Evaluating the valuations. The truthful mechanisms require the center to evaluate each agent's valuation at the models it computes (Assumptions 5 and 7); under Equation (5) this amounts to knowing L_i at those models and the exponent α_i. The loss Equation (8) shows one way in which such knowledge can arise without access to the agents' data. Its gradient $\nabla L_i(w) = 2A_i(w - w_i^*)$ is affine in w, so gradients at $m+1$ affinely independent parameters determine A_i and w_i^*; when $A_i = \mathbb{I}$ a single gradient suffices, giving $w_i^* = w - \frac{1}{2}\nabla L_i(w)$. This reconstruction requires the gradients to be exact and reported truthfully, as in `UPBReD` and `2P-UPBReD`. In `T-2P-UPBReD` gradient reports are strategic, so Theorem 3 uses Assumption 5 as stated, under which the center's evaluation does not depend on the gradient reports. A gradient-based estimate would be biased: with $A_i = \mathbb{I}$, an agent playing the round equilibrium of part 4 of Theorem 3 reports $\nabla_w v_i(w^{t+1}, s^{\max}) = -2(w^{t+1} - w_i^*)$, so the estimate $w^t + \frac{1}{2}\hat{d}_i^t = w_i^* + (w^t - w^{t+1})$ is off by one step. Since the steps vanish as training converges, such an estimate, like an estimate from mini-batch gradients, is compatible with the tolerance version of the tests (remark after Theorem 3), but not with the exact tests of Equation (24), which would penalize truthful agents. Finally, knowing (A_i, w_i^*) for every agent is more than the mechanisms need, since it determines $w^{\texttt{OPT}} = \bar{A}^{-1}\frac{1}{n}\sum_{i\in N} A_i w_i^*$ in closed form; the mechanisms require only the losses at the models they compute.

4 Learning with arbitrary effective costs

In this section, we consider the effective cost of agent i, i.e., $z_i(s_i, s_{-i})$, to be an arbitrary function of $s = (s_i, s_{-i})$. We know that an alternative interpretation of NE (Definition 1) is a strategy profile (s_i^*, s_{-i}^*) where every agent's *best response* to the strategies of the other players is its own strategy in that profile [see, e.g., 34], i.e., $s_i^* \in \operatorname{argmax}_{s_i\in S_i} u_i(w, s_i, s_{-i}^*)$, $\forall i \in N$. Hence, an algorithm that simultaneously updates all agents' strategies with the best responses to the current strategies of the other players is called a *best response dynamics* of a strategic form game [see, e.g., 10, for a detailed description]. In our problem, this approach cannot be directly employed since the center also needs to learn and update the model parameters w as the agents choose their data contributions. We, therefore, propose a mechanism for FL that simultaneously updates both the agents' strategies and the center's choice of w. FL protocols iteratively use local gradients to update a global model; the mechanism we describe below captures strategic behavior at every iteration by allowing agents to update their strategy (data set sizes used for gradient computation) at every realization of the global model based on Equation (1).

In this mechanism, each agent i starts with some initial choices of s_i^0 and shares that with the center.[1] The center initializes a w^0 and broadcasts that and the entire initial data-contribution vector s^0 to all the agents. In every subsequent iteration t, each agent i locally computes two quantities: (i) its updated contribution s_i^{t+1} by taking one gradient ascent step w.r.t. its own contribution s_i, and (ii) the local gradient of agent i's valuation component of the social welfare w.r.t. w. Both are evaluated at the current values of w^t and s^t and sent back to the center. The center calculates an updated w^{t+1} that averages (in the spirit of the FedAvg algorithm [35]) all the local gradients sent by the agents. The center then shares w^{t+1} and s^{t+1} with all the agents. Formally, the updates are given as follows.

$$s^{t+1} = s^t + \gamma g(w^t, s^t, \mu^t), \qquad w^{t+1} = w^t + \eta\tilde{g}(w^t, s^t), \tag{9}$$

where the function $\tilde{g}(w^t, s^t) = \frac{1}{n}\sum_{i\in N}\nabla_w v_i(w^t, s^t)$ and $[g(w^t, s^t, \mu^t)]_i = \frac{\partial}{\partial s_i}u_i(w^t, s^t) + \mu_i^t$, with μ_i^t

[1] Note that only the number of data points, s_i, is shared with the center and not the data, which is consistent with the principle of FL.

defined below. Let $s_i' := s_i^t + \gamma \frac{\partial}{\partial s_i} u_i(w^t, s^t)$ denote the unconstrained gradient step. Then

$$\mu_i^t = \begin{cases} \big(s_i^{\max} - s_i'\big)/\gamma & \text{if } s_i' > s_i^{\max}, \\ \big(s_i^{\min} - s_i'\big)/\gamma & \text{if } s_i' < s_i^{\min}, \\ 0 & \text{otherwise}, \end{cases} \tag{10}$$

so that $s_i^{t+1} = s_i' + \gamma\mu_i^t = \Pi_{S_i}(s_i')$, where $S_i = [s_i^{\min}, s_i^{\max}]$ and $\Pi_{S_i}(x) := \min\{s_i^{\max}, \max\{s_i^{\min}, x\}\}$. Equivalently, $[g(w^t, s^t, \mu^t)]_i = \big(\Pi_{S_i}(s_i') - s_i^t\big)/\gamma$.

The mechanism is detailed in Algorithm 1. To guarantee the existence of and convergence to a Nash equilibrium, we require standard curvature and smoothness properties on the agents' utility functions. Intuitively, each agent's utility is strongly concave in its own contribution, the average valuation is strongly concave in w, and the cross-derivatives between w and s are bounded relative to these curvatures formalized in Assumption 1. Under these conditions, our main result shows that Algorithm 1 always converges to a Nash equilibrium.

We need a few matrices, defined as follows for every w and s. For $i, j \in N$ and $k, \ell \in \{1, \dots, m\}, G(w,s)_{ij} = \frac{\partial^2}{\partial s_j \partial s_i} u_i(w,s); \tilde{G}(w,s)_{k\ell} = \frac{1}{n}\sum_{i\in N} \frac{\partial^2}{\partial w_\ell \partial w_k} v_i(w,s); H(w,s)_{ik} = \frac{\partial^2}{\partial w_k \partial s_i} u_i(w,s); \tilde{H}(w,s)_{kj} = \frac{1}{n}\sum_{i\in N} \frac{\partial^2}{\partial s_j \partial w_k} v_i(w,s)$.

Assumption 1. *Consider the utility functions given by Equation (1) where functions $v_i, z_i, i \in N$ are such that the following properties hold for every $w \in \mathbb{R}^m$ and $s \in \prod_{i\in N} S_i$, $\forall i, j \in N, \forall k, \ell \in \{1, \dots, m\}$ (the matrices below are as defined above).*

1. *The matrices $G(w,s) + \lambda\mathbb{I}$ and $\tilde{G}(w,s) + \tilde{\lambda}\mathbb{I}$ are negative semi-definite, i.e., $v^\top \big(G(w,s) + \lambda\mathbb{I}\big)v \leqslant 0$ for all $v \in \mathbb{R}^n$ and $v'^\top \big(\tilde{G}(w,s) + \tilde{\lambda}\mathbb{I}\big)v' \leqslant 0$ for all $v' \in \mathbb{R}^m$.*
2. *We assume the following bounds: (i) $|G(w,s)_{ij}| \leqslant L$, (ii) $|\tilde{G}(w,s)_{k\ell}| \leqslant \tilde{L}$, (iii) $\|H(w,s)\|_{op} \leqslant P$, where $\|A\|_{op} := \inf\{c \geqslant 0 : \|Av\| \leqslant c\|v\|, \forall v\}$, (iv) $\|\tilde{H}(w,s)\|_{op} \leqslant \tilde{P}$.*

Remark. While the conditions on G align with assumptions on contribution dynamics of Murhekar et al. [38], the joint conditions on (w,s) via $\tilde{G}$ are novel to our work. These additional joint assumptions are used to establish a stronger guarantee: the simultaneous convergence of both the strategic contributions and the model parameter w.

Remark (Verification for Equation (4)). We verify Assumption 1 for the utility form of Equation (4) with the decomposition Equation (5), in the setting of `UPBReD`, where no payments are made. The valuation is additively separable in w and s: $L_i(w)$ does not depend on s, and the finite-sample correction does not depend on w. Every mixed derivative $\frac{\partial^2}{\partial w_k \partial s_j}$ therefore vanishes, so $H \equiv 0$ and $\tilde{H} \equiv 0$, conditions 2(iii) and (iv) hold with $P = \tilde{P} = 0$, and the requirements $\lambda > \tilde{P}$, $\tilde{\lambda} > P$ of Theorem 1 reduce to $\lambda, \tilde{\lambda} > 0$.

For the matrix G, differentiating Equation (4) with respect to s_i, the cost contributes the constant $-c_i$, so that

$$\frac{\partial}{\partial s_i} u_i(w,s) = \frac{\alpha_i}{n}\, s_i^{-\alpha_i - 1} - c_i.$$

This depends on s_i alone, because the correction in Equation (5) decomposes across agents. Differentiating once more, $G(w,s)_{ij} = 0$ for $j \neq i$ and

$$G(w,s)_{ii} = -\frac{\alpha_i(\alpha_i + 1)}{n}\, s_i^{-\alpha_i - 2},$$

so G is diagonal, and in particular symmetric. On $S_i = [1, s_i^{\max}]$ the diagonal entries lie between $-\frac{\alpha_i(\alpha_i+1)}{n}$ and $-\frac{\alpha_i(\alpha_i+1)}{n}(s_i^{\max})^{-\alpha_i-2}$. Hence condition 1 holds with $\lambda = \min_{i\in N} \frac{\alpha_i(\alpha_i+1)}{n}(s_i^{\max})^{-\alpha_i-2} > 0$, and condition 2(i) holds with $L = \max_{i\in N} \frac{\alpha_i(\alpha_i+1)}{n}$. The lower endpoint 1 of S_i is what keeps L finite.

For the matrix $\tilde{G} = -\frac{1}{n}\sum_{i\in N} \nabla^2 L_i(w)$, the losses Equations (6) and (7) satisfy $\nabla^2 L_i \succeq \mu\mathbb{I}$, so condition 1 holds with $\tilde{\lambda} = \mu$. For the cross-entropy loss, every entry of the softmax Hessian $\mathrm{diag}(p) - pp^\top$ has absolute value at most $\frac{1}{4}$, so the entries of $\nabla^2 L_i$ are bounded by $\tilde{L} = \frac{1}{4}B_x^2 + \mu$; for the MSE loss, by $\tilde{L} = 2B_x^2 + \mu$. These bounds hold for every $w \in \mathbb{R}^m$, so condition 2(ii) requires no bound on the weights.

Algorithm 1: Updated Parameter Best Response Dynamics (`UPBReD`)

Input: Step size γ, η, initialization w^0, s_i^0 for $i \in N$, number of iterations T
Output: w^T

for $t = 0$ ***to*** $T-1$ **do**
  Center broadcasts w^t, s^t
  foreach ***agent*** $i \in N$ *in parallel* **do**
    $s_i^{t+1} = s_i^t + \gamma[g(w^t, s^t, \mu^t)]_i$
    Compute local gradient: $d_i^{t+1} = \nabla_w v_i(w^t, s_i^t, s_{-i}^t)$
    Send s_i^{t+1}, d_i^{t+1} to the center
  Center updates: $w^{t+1} = w^t + \frac{\eta}{n}\sum_{i\in N} d_i^{t+1} = w^t + \eta\tilde{g}(w^t, s^t)$
return w^T

$$
\begin{aligned}
W_1 &= \sqrt{1+\gamma^2 n^2 L^2 - 2\gamma\lambda} + \sqrt{\tilde{P}^2\gamma^2},\\
W_2 &= \sqrt{1+\eta^2 m^2 \tilde{L}^2 - 2\eta\tilde{\lambda}} + \sqrt{P^2\eta^2},\\
E &= \|g(w^0, s^0, \mu^0)\|_2 + \|\tilde{g}(w^0, s^0)\|_2,\\
T_0(w^0, s^0) &= \left(\ln \tfrac{E}{\epsilon}\right) / \left(\ln \tfrac{1}{W}\right),
\end{aligned}
\tag{11}
$$

where $w^0 \in \mathbb{R}^m, s^0 \in \prod_{i\in N} S_i$ and $W = \max\{W_1, W_2\}$.

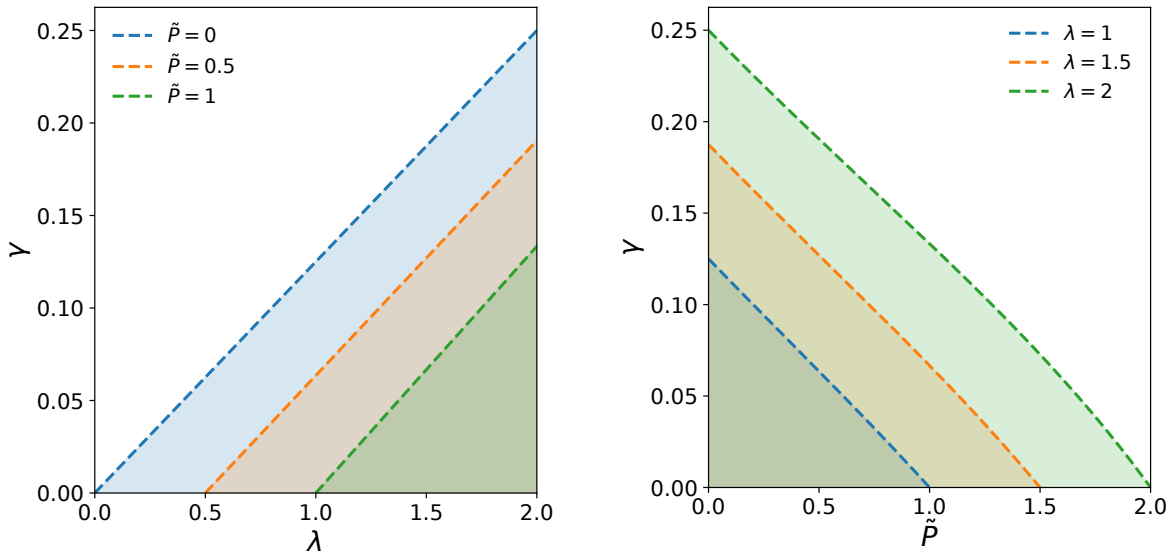
Theorem 1 (`UPBReD` convergence to NE)**.** *Suppose Assumption 1 holds with $\lambda > \tilde{P}$ and $\tilde{\lambda} > P$, and the step sizes satisfy $\gamma < \frac{2(\lambda-\tilde{P})}{n^2L^2-\tilde{P}^2}$ and $\eta < \frac{2(\tilde{\lambda}-P)}{m^2\tilde{L}^2-P^2}$. Then for every (w^0, s^0) the iterates of Algorithm 1 converge to a limit (w^*, s^*) such that s^* is a PSNE at w^* and w^* is the unique maximizer of $\frac{1}{n}\sum_{i\in N} v_i(\cdot, s^*)$. Moreover, for every $\epsilon > 0$ and every $T > T_0(w^0, s^0)$, we have $\|g(w^T, s^T, \mu^T)\|_2 < \epsilon$, $\|\tilde{g}(w^T, s^T)\|_2 < \epsilon$, and $\|s^T - s^\star(w^T, s^T)\|_2 < \frac{1+\gamma L}{\lambda}\epsilon$, where $s^\star(w, s)$ is the profile of best responses.*

Proof sketch. The proof tracks the gradient mappings g^t and $\tilde{g}^t$. Since the projection onto $S = \prod_{i\in N} S_i$ is non-expansive, the integral form of the mean value theorem bounds $\|g^{t+1}\|_2$ and $\|\tilde{g}^{t+1}\|_2$ through averages of G, H, $\tilde{G}$, $\tilde{H}$, which inherit Assumption 1; strong concavity and the bounds on the cross terms give $\|g^{t+1}\|_2 + \|\tilde{g}^{t+1}\|_2 \leqslant W(\|g^t\|_2 + \|\tilde{g}^t\|_2)$, and the step-size conditions are exactly those that make $W < 1$. The increments are then summable and the iterates converge, since $\mathbb{R}^m \times \mathbb{R}^n$ is complete. A zero of the projected step is a Nash equilibrium because each u_i is concave in s_i, a stationary point of the strongly concave average valuation is its unique maximizer, and an error bound for the projected step gives the distance to the best responses. The complete proof is in Appendix A.1. □

The theorem provides a guarantee of convergence of Algorithm 1 for any arbitrary initial condition (w^0, s^0). It needs to run for a minimum number of iterations $T_0(w^0, s^0)$, with appropriate parameters of γ and η that govern the gradient ascent rates of the agents' and center's objectives, respectively. Note that for a better convergence, i.e., a smaller ϵ, the algorithm needs to run for more iterations. The parameter W determines the contraction factor of the recurrence of $\|g(w^t, s^t)\|_2 + \|\tilde{g}(w^t, s^t)\|_2$ and is chosen to be smaller than unity by choosing γ and η appropriately. Figure 1 shows an example for such a feasible regions of γ for certain $n, L, \lambda, \tilde{P}$.

Remark. While the step sizes γ and η required by Theorem 1 scale inversely with m and n, these restrictions arise as sufficient conditions within our convergence analysis. In our experimental evaluations, we observe robust convergence even with small constant values of γ and η (see Section 6). Furthermore, when monetary transfers are permitted (Section 5), `2P-UPBReD` leverages its two-phase structure to establish convergence guarantees with step sizes that are naturally independent of m and n.

While agents strategically choose s_i to maximize their utility, we assume truthful reporting of these choices. In `UPBReD`, agents report s_i^t and d_i^t in each round. As Example 3 shows, misreporting can improve an agent's utility. In Section 5.4 we introduce payments that make the reports of s_i^t and d_i^t truthful.



(a) $\tilde{P} = 1$ **(b)** $\lambda = 1.1$

Figure 1: *Shaded regions show the feasible choices of γ in Theorem 1 for $n = 2$ and $L = 2$: (a) as a function of λ for the values of $\tilde{P}$ shown, and (b) as a function of $\tilde{P}$ for the values of λ shown. Dashed lines mark the boundary $\gamma = 2(\lambda - \tilde{P})/(n^2L^2 - \tilde{P}^2)$. The region is empty when $\lambda \leqslant \tilde{P}$. The feasible choices of η are obtained by replacing $(n, L, \lambda, \tilde{P})$ with $(m, \tilde{L}, \tilde{\lambda}, P)$.*

Application of Theorem 1: effective costs are only personal

One special but practical case of the above setup is when the effective cost is only borne by the agent. Mathematically, it is represented as $z_i(s_i, s_{-i}) = c_i(s_i), \forall s_i \in S_i, s_{-i} \in S_{-i}, \forall i \in N$.

We assume convex cost functions $c_i, \forall i \in N$, which is a commonly used assumption in the literature [31]. This setup is consistent with the assumptions of Assumption 1, and according to Theorem 1, Algorithm 1 converges to a Nash equilibrium, which we denote as (w^*, s^*). However, it is possible that either of the following happens: (i) $s^* \neq s^{\max}$, i.e., the agents do not contribute their entire data for the federated learning process, leading to a suboptimal learning, or (ii) $w^* \neq w^{\texttt{OPT}}$ (see Equation (2) for the definition), which is neither the objective of the center nor of the agents. The following example shows such an instance.

Example 1 (NE different from socially optimal)**.** *Consider two agents with linear costs who learn a model with two parameters, where agent 1's data informs the first parameter and agent 2's data informs the second. Both agents have the valuation*

$$v_i(w, s) = 2 - (1 - w_1)^2 - (2 - w_2)^2 - \frac{w_1^2}{s_1} - \frac{w_2^2}{s_2}, \qquad i = 1, 2,$$

in which the term w_k^2/s_k acts as a regularizer on the parameter informed by agent k, whose weight shrinks as agent k contributes more data. The costs are $c_1(s_1) = \frac{1}{64}s_1$ and $c_2(s_2) = \frac{1}{16}s_2$, and the strategy sets are $S_i = [1, 9]$.

For a fixed contribution profile s, maximizing v_i over w gives $w_1 = s_1/(s_1+1)$ and $w_2 = 2s_2/(s_2+1)$, so the learned parameter approaches the target $(1, 2)$ only as the contributions grow. At $s = s^{\max} = (9, 9)$ this yields $w^{\texttt{OPT}} = (0.9, 1.8)$ and the optimal social welfare $\sum_i v_i(w^{\texttt{OPT}}, s^{\max}) = 3$. This profile is not an NE: agent i's marginal utility at $(w^{\texttt{OPT}}, s^{\max})$ is $w_i^2/(s_i^{\max})^2 - c_i'$, which equals $0.01 - \frac{1}{64} < 0$ for agent 1 and $0.04 - \frac{1}{16} < 0$ for agent 2, so each agent prefers to reduce its contribution.

Starting from $(w^{\texttt{OPT}}, s^{\max})$ with $\gamma = \eta = 0.25$, Algorithm 1 converges to $w^* = (0.875, 1.75)$ and $s^* = (7, 7)$, where the social welfare is $2.75 < 3$. The agents contribute less than their full data, and as a consequence the center learns a parameter different from $w^{\texttt{OPT}}$. Full contribution is efficient here because each agent's data benefits both agents, whereas each agent weighs only its own benefit against its cost. □

The example shows that the Nash equilibrium reached by `UPBReD` need not be socially optimal. The same effect appears on real data: in Figure 2, the social welfare of `UPBReD` lies below that of `FedAvg`, in which every agent trains on its entire dataset. This motivates the payments of the next section, which make full contribution optimal for every agent.

5 Welfare maximization at Nash equilibrium

In this section, we consider the scenario where the center can make a payment $p_i(s_i, s_{-i})$ *to* agent i to alter its utility function. Therefore, the effective cost of agent i becomes $z_i(s_i, s_{-i}) = c_i(s_i) - p_i(s_i, s_{-i})$. Note that the payment can be either positive or negative, which determines whether the agent is *subsidized* or *taxed* respectively. It is reasonable to expect that the *data contributors* of the FL process are subsidized for their data contribution while the *data consumers* are charged a payment for obtaining the learned parameter. We assume that the derivatives of the valuation v_i and cost function c_i are bounded. We choose the following payment function:

$$p_i^1(s) = \beta(s_i - \tfrac{1}{n-1}\textstyle\sum_{j\neq i} s_j), \tag{12}$$

where β is a parameter of choice. Note that this payment mechanism is budget balanced by design, since $\sum_{i\in N} p_i^1(s_i, s_{-i}) = 0, \forall s_i \in S_i, s_{-i} \in S_{-i}$. With payment, the utility of agent i becomes $u_i(w, s) = v_i(w, s) - c_i(s_i) + p_i^1(s)$. We now discuss how the payment can be used to maximize the social welfare at a NE.

5.1 Convergence to the NE $(w^{\mathrm{OPT}}, s^{\max})$

In this setup, we first prove that the utility of every agent i can be made strictly increasing in its own contribution s_i in the following manner.

Assumption 2. *The partial derivatives of $v_i(w, s)$ are bounded below, i.e., $\frac{\partial}{\partial s_j} v_i(w, s) \geqslant -\tau$ for some $0 \leqslant \tau < \infty$, for all $i, j \in N$, all w and all s. Moreover, the cost functions $c_i(\cdot)$ have bounded derivatives, $c_i'(s_i) \leqslant \zeta$ for all s_i and i.*

Note that the above assumptions are mild. Previous works make stronger assumptions (like convexity) on c_i and v_i with respect to s_i [see 38]. Considering the utility form as described in Equation (4), Assumption 2 holds with $\tau = 0$ as $\partial v_i / \partial s_j = \frac{\alpha_i}{n} s_j^{-\alpha_i - 1} > 0$. We show that the above bounds on the derivatives of v_i and c_i are sufficient to claim our result.

Lemma 1 (Increasing utility)**.** *Suppose Assumption 2 holds and the payment function is given by Equation (12) with $\beta > \zeta + \tau$. Then for every $i \in N$, every w, and every $s \in \prod_j S_j$*

$$\frac{\partial}{\partial s_i} u_i(w, s_i, s_{-i}) \;>\; 0.$$

Proof sketch. By Equation (12), $\frac{\partial}{\partial s_i} u_i = \frac{\partial}{\partial s_i} v_i - c_i' + \beta \geqslant \beta - \tau - \zeta > 0$ by Assumption 2. The complete proof is in Appendix B.1. □

Remark (Knowledge of β). The knowledge of β is required for theoretical tractability only. In the experiments, we do not assume any knowledge of β; rather, we tune this hyperparameter. We observe that uniformly across all datasets, if β is chosen larger than a threshold, we obtain the best performance of our proposed algorithms, which also validates the above theoretical requirement.

Remark ($s^{\max}$ is a dominant strategy). The result of Lemma 1 is stronger than the Nash property invoked in Theorem 2, since Assumption 2 is already quantified over all w and all s. It is worth recording because it removes any equilibrium selection question from Phase 1: the agents need not coordinate, need not hold correct beliefs about one another, and need not observe the parameter, since increasing one's own contribution is optimal for any choice of the above.

This lemma implies that even in Algorithm 1 if we apply the above payment, s will converge to the maximum value $s^{\max}$. However, unlike Algorithm 1, in this section we provide an algorithm which gives the step sizes of the gradient ascents in a more concrete manner and is, therefore, a superior one. In

Algorithm 2: `2P-UPBReD`

Input: Step sizes γ, η; initialization w^0, $s_i^0 \in S_i$ for $i \in N$; iteration counts κ, T_0 as given by Theorem 2

Output: w^{T_0}

`// Phase 1: contributions, with the model fixed at` w^0

Center broadcasts w^0, s^0; set $k = 0$

while $k < \kappa$ **do**

  Center broadcasts s^k

  foreach ***agent*** $i \in N$ *in parallel* **do**

    $s_i^{k+1} = s_i^k + \gamma\,[g(w^0, s^k, \mu^k)]_i$; send s_i^{k+1} to center

  $k \leftarrow k + 1$

`// Under Equation (12),` $s^\kappa = s^{\max}$

`// Phase 2: model learning at` $s^{\max}$

for $t = 0$ **to** $T_0 - 1$ **do**

  Center broadcasts w^t

  foreach ***agent*** $i \in N$ *in parallel* **do**

    Compute $d_i^{t+1} = \nabla_w v_i(w^t, s^{\max})$; send d_i^{t+1} to center

  Center updates $w^{t+1} = w^t + \frac{\eta}{n} \sum_{i \in N} d_i^{t+1} = w^t + \eta\, \tilde{g}(w^t, s^{\max})$

return w^{T_0}

order to keep the convergence rate the same as Theorem 1, we assume that the negative[2] social welfare function at $s^{\max}$, given by $f(w, s^{\max}) = -\frac{1}{n} \sum_{i \in N} v_i(w, s^{\max})$, is M-smooth and ν-strongly convex in w. These properties are defined as follows.

Definition 5 (M-smoothness)**.** *A function $f : \mathbb{R}^m \to \mathbb{R}$ is M-smooth if for all $x, x' \in \mathbb{R}^m$, we have $f(x') \leqslant f(x) + \langle \nabla_x f(x), x' - x \rangle + \frac{M}{2} \|x - x'\|_2^2$.*

Definition 6 (ν-strongly convex)**.** *A function $f : \mathbb{R}^m \to \mathbb{R}$ is ν-strongly convex if for all $x, x' \in \mathbb{R}^m$, we have $f(x') \geqslant f(x) + \langle \nabla_x f(x), x' - x \rangle + \frac{\nu}{2} \|x - x'\|_2^2$.*

We propose a two-phase algorithm given by Algorithm 2. The algorithm, in the first phase, incentivizes the agents to contribute $s^{\max}$, and in the second phase, converges to $w^{\texttt{OPT}}$. Note that in the first phase of Algorithm 2, every agent runs a gradient ascent step. Thanks to the increasing utility property (Lemma 1), we have an increasing sequence of s_i^t for all $i \in N$. Moreover, from the definition of $[g(w^0, s^t, \mu^t)]_i$ it is ensured that $s_i^{\max}$ is the fixed point of this gradient ascent update. Hence after a finite number of iterations, every agent reaches $s_i^{\max}$. Note that in this phase the model parameter w^0 remains unchanged.

In the second phase of the algorithm, we update the model parameter w^t. Note that since all the agents contribute the maximum amount of data they own, this phase is simply *pure* FL (without incentive design). As such, each agent now computes the local gradient d_i^{t+1} and sends it to the center in each round.

We now provide the convergence guarantees of Algorithm 2. Before that, let us discuss the necessary assumptions.

Assumption 3. *The negative social welfare function at $s^{\max}$, $f(w, s^{\max})$, is M-smooth and ν-strongly convex in w, with $M > \nu$.*

Remark (Strong convexity and smoothness). The smoothness and strong convexity assumptions have featured in several previous papers on FL [26, 55]. Considering the utility function described in Equation (4), Assumption 3 holds under L2 regularization, consistent with standard convergence analyses of FedAvg and SCAFFOLD [26], and more recent analyses of more complex FL frameworks such as split FL [18]. We show in the following theorem that the convergence rate under Assumption 3 is $\mathcal{O}(\ln 1/\epsilon)$ which is the same as that of Algorithm 1. However, Algorithm 2 is cleaner in terms of the phases of

[2] We negate the welfare so that we can consider functions as convex to apply the results of convex analysis easily.

convergence and the choices of the step sizes γ, η. These structural assumptions of strong convexity and smoothness may be relaxed at the expense of weaker rates. For example, $O(1/\epsilon)$ iterations for convex smooth f, and $O(1/\epsilon^2)$ iterations to reach an ϵ-stationary point for non-convex smooth f. Since the second phase of 2P-UPBReD is pure FL, any convergence result in the FL literature with the corresponding assumptions on $f(w, s^{\max})$ can be used in place of the second phase of this algorithm. The main result of this section is as follows.

Theorem 2 (2P-UPBReD convergence to the optimal NE)**.** *Suppose Assumptions 2 and 3 hold and we consider the utility function with payment scheme given in Equation (12) with $\beta > \zeta + \tau$. Then under truthful reports, for every $w^0 \in \mathbb{R}^m$ and $s^0 \in \prod_{i\in N} S_i$, 2P-UPBReD (Algorithm 2) converges to the Nash equilibrium $(w^{\texttt{OPT}}, s^{\max})$ and is budget balanced. In particular, for any given $\epsilon > 0$, we have $\frac{1}{n}\sum_{i\in N} v_i(w^{\texttt{OPT}}, s^{\max}) - \frac{1}{n}\sum_{i\in N} v_i(w^T, s^{\max}) < \epsilon$, and $s^T = s^{\max}$ in $T = \kappa + T_0$ iterations provided we choose $\gamma = c > 0$ (a universal constant) and $\eta = 1/M$ with*

$$\begin{aligned} \kappa &\geqslant \max_i \left\{ \tfrac{(s_i^{\max} - s_i^0)}{c\,\Delta} \right\}, \\ T_0 &> \left(\ln \tfrac{f(w^0, s^{\max}) - f(w^{\texttt{OPT}}, s^{\max})}{\epsilon} \right) \Big/ \left(\ln \tfrac{1}{1 - \frac{\nu}{M}} \right) \end{aligned} \tag{13}$$

where $f(w, s) := -\frac{1}{n}\sum_{i\in N} v_i(w, s)$ and $\Delta = \beta - \tau - \zeta$.

Proof sketch. By Lemma 1, $\frac{\partial}{\partial s_i} u_i(w^0, s^k) \geqslant \Delta$, so each projected step raises s_i by at least $\gamma\Delta$ until it reaches $s_i^{\max}$, which is a fixed point of the step; hence $s^\kappa = s^{\max}$, which is each agent's unique best response at every w because its utility is strictly increasing in its own contribution. Phase 1 uses no concavity, only the lower bound Δ on the marginal utility. Budget balance holds because $\sum_{i\in N} p_i^1(s) = 0$ for every s. Phase 2 is gradient descent with $\eta = 1/M$ on the M-smooth, ν-strongly convex $f(\cdot, s^{\max})$, which reduces $f(w^t) - f(w^{\texttt{OPT}})$ by the factor $1 - \nu/M$ in every round [52], and strong convexity gives $\|w^{T_0} - w^{\texttt{OPT}}\|_2^2 \leqslant \frac{2}{\nu}\big(f(w^{T_0}) - f(w^{\texttt{OPT}})\big)$, so the iterates converge to $w^{\texttt{OPT}}$. The complete proof is in Appendix B.2. □

Remark (γ and rationality). Phase 1 of Algorithm 2 achieves the optimal contribution $s^{\max}$ without updating the learning parameter w. The speed of convergence of this phase depends on the parameter $\gamma = c$, which can be any positive constant. A larger c ensures faster convergence, however the algorithm converges for any positive $c > 0$. In Phase 2 of the algorithm, the agents stop updating their contributions and only update the gradients d_i and the center accumulates them to update w^t.

Thanks to the smoothness and strong convexity properties of $f(\cdot, s^{\max})$, we can show that running Algorithm 2 guarantees that the model iterate w^T converges to $w^{\texttt{OPT}}$ at an exponential speed as well.

Corollary 1 (Iterate convergence)**.** *With the same setup as above and $\gamma = c$ (a universal constant), $\eta = \frac{2}{M+\nu}$, we obtain $\|w^T - w^{\texttt{OPT}}\|_2 < \epsilon$, and $s^T = s^{\max}$, where $T = \kappa + \tilde{T}_0$, with κ as in Equation (13) and $\tilde{T}_0 > \left(\ln \frac{\|w^0 - w^{\texttt{OPT}}\|_2}{\epsilon} \right) / \left(\ln \frac{1 + \frac{\nu}{M}}{1 - \frac{\nu}{M}} \right)$.*

The proof follows that of Phase 2 of Theorem 2 with the step size $\eta = \frac{2}{M+\nu}$, for which gradient descent contracts the distance to $w^{\texttt{OPT}}$ by the factor $\frac{1-\nu/M}{1+\nu/M}$ in every round; it is in Appendix B.3.

5.2 Individual rationality

Participation in federated learning is voluntary. Theorem 2 shows that, once the agents have joined, the payment p^1 of Equation (12) makes full contribution optimal for every agent. This guarantee is only useful if the agents join in the first place: an agent that expects to be worse off inside the mechanism than outside it will not participate, and free-riding then reappears as non-participation rather than as under-contribution. We therefore ask whether every agent prefers taking part in 2P-UPBReD to training a model on its own data alone. This property is called *individual rationality*.

Two features of 2P-UPBReD make the question non-trivial. First, the transfers are budget balanced, so they redistribute money among the agents rather than inject it from the center. Since p_i^1 rewards contributing more than the average of the other agents, the redistribution runs from agents with small datasets to agents with large ones: an agent with below-average capacity pays, and it participates only if the shared model is worth more to it than both its payment and its data cost, compared with training

alone. Second, the shared model need not benefit every agent equally. An agent whose data distribution differs from the others' may value the federated model less than a model trained on its own data; if this agent has a large dataset, it must be compensated through its subsidy.

The parameter β governs both effects. Lemma 1 requires $\beta > \zeta + \tau$ so that contributing more is always profitable, and compensating large contributors whose gain from the shared model is small may require β to be larger still. On the other hand, the payment of each small contributor grows linearly in β, so a large β can drive small contributors out. Individual rationality therefore holds only for β in a window, whose endpoints depend on how much each agent gains from collaboration and on how far its capacity is from the average.

The rest of this subsection makes this precise. We first fix the cost accounting, the individual learning outcome and the notion of individual rationality. Proposition 1 then gives each agent's exact gain from participating at the outcome of `2P-UPBReD` and characterizes the window of β for which every agent participates, and Corollary 2 shows that the guarantee holds after finitely many rounds. Finally, when the window is empty, Corollary 3 shows that zero-sum lump-sum rebates restore individual rationality, without affecting incentives or budget balance, exactly when the agents gain from collaboration in aggregate.

Individual learning outcome. An agent that does not participate pays and receives no transfers and has no access to the federated model; it can only train on its own data. For $x \in S_i$, let $v_i^{\mathrm{loc}}(x)$ denote the valuation that agent i derives from the best model it can train alone on x of its own data points. Agent i's *individual learning outcome* is the best stand-alone utility,

$$\bar{u}_i := \sup_{x \in S_i} \left[v_i^{\mathrm{loc}}(x) - c_i(x) \right]. \tag{14}$$

We use the stand-alone benchmark rather than the value 0 because the valuation $v_i = r_i - L_i$ contains the arbitrary normalizing constant r_i, so comparing utilities with 0 would carry no meaning.

Outcome and individual rationality. By Theorem 2, `2P-UPBReD` reaches the full-contribution profile $s^{\max}$ and learns $w^{\mathtt{OPT}}$. The utility of agent i at this *limit outcome* is

$$U_i := v_i(w^{\mathtt{OPT}}, s^{\max}) - c_i(s_i^{\max}) + p_i^1(s^{\max}), \tag{15}$$

and its utility at the outcome after T rounds is $U_i^T := v_i(w^T, s^T) - c_i(s_i^T) + p_i^1(s^T)$.

Definition 7 (Individual rationality)**.** *The limit outcome of* `2P-UPBReD` *is* individually rational *(IR) if* $U_i \geqslant \bar{u}_i$ *for every* $i \in N$*. The outcome after* T *rounds is individually rational if* $U_i^T \geqslant \bar{u}_i$ *for every* $i \in N$*.*

Gain from collaboration and average capacity. We write

$$\pi_i := v_i(w^{\mathtt{OPT}}, s^{\max}) - c_i(s_i^{\max}) - \bar{u}_i \tag{16}$$

for agent i's *gain from collaboration*: the value of the federated model at full contribution, net of the cost of contributing all of its data, in excess of what the agent achieves alone. The gain π_i may be negative, for instance for an agent with high per-sample cost. We also write

$$\bar{s}^{\max} := \frac{1}{n} \sum_{j \in N} s_j^{\max} \tag{17}$$

for the average capacity, and we split the agents into those with below-average, average and above-average capacity,

$$N_- := \{i : s_i^{\max} < \bar{s}^{\max}\}, \qquad N_0 := \{i : s_i^{\max} = \bar{s}^{\max}\}, \qquad N_+ := \{i : s_i^{\max} > \bar{s}^{\max}\}.$$

Finally, we define the two thresholds

$$\beta_{\mathrm{low}} := \max\left\{0,\ \max_{i \in N_+} \frac{-(n-1)\,\pi_i}{n\,(s_i^{\max} - \bar{s}^{\max})}\right\}, \qquad \beta_{\mathrm{high}} := \min_{i \in N_-} \frac{(n-1)\,\pi_i}{n\,(\bar{s}^{\max} - s_i^{\max})}, \tag{18}$$

with the conventions that a maximum over the empty set is 0 and a minimum over the empty set is $+\infty$.

Proposition 1 (Individual rationality of 2P-UPBReD)**.** *Suppose Assumptions 2 and 3 hold and the payment is p^1 of Equation (12) with $\beta > \zeta + \tau$. Then the following hold.*

1. *At the limit outcome, agent i's gain from participating is*

$$U_i - \bar{u}_i \;=\; \pi_i + \frac{n\beta}{n-1}\big(s_i^{\max} - \bar{s}^{\max}\big). \tag{19}$$

2. *The limit outcome is individually rational if and only if $\pi_i \geqslant 0$ for every $i \in N_0$ and $\beta_{\text{low}} \leqslant \beta \leqslant \beta_{\text{high}}$.*
3. *Consequently, there exists a β for which Theorem 2 applies and the limit outcome is individually rational if and only if $\pi_i \geqslant 0$ for every $i \in N_0$, $\beta_{\text{low}} \leqslant \beta_{\text{high}}$, and $\zeta + \tau < \beta_{\text{high}}$. In particular, if $\pi_i \geqslant 0$ for every $i \in N$, such a β exists if and only if $\zeta + \tau < \beta_{\text{high}}$.*

Proof sketch. Evaluating Equation (12) at $s^{\max}$ gives $p_i^1(s^{\max}) = \frac{n\beta}{n-1}(s_i^{\max} - \bar{s}^{\max})$, which yields Part 1. The condition of Part 1 is then linear in β: it is an upper bound on β for agents in N_-, a lower bound for agents in N_+, and independent of β for agents in N_0, which gives Parts 2 and 3. The complete proof is in Appendix B.4. □

Remark (Interpretation and choice of β). Proposition 1 shows that individual rationality can fail only for agents with below-average capacity whose gain from collaboration is small: the tax such an agent pays at the limit outcome, $\frac{n\beta}{n-1}(\bar{s}^{\max} - s_i^{\max})$, grows linearly in β. Since Theorem 2 only needs $\beta > \zeta + \tau$, the center should choose β just above this threshold, which makes the window of Proposition 1 as easy to satisfy as possible.

The proposition concerns the limit outcome. The next corollary shows that the conclusion carries over to every sufficiently late round whenever the gains are strict.

Corollary 2 (Individual rationality after finitely many rounds)**.** *In addition to the hypotheses of Proposition 1, suppose that each valuation $v_i(\cdot, s^{\max})$ is continuous at w^{OPT}, that 2P-UPBReD is run with $\eta = 1/M$, and that $U_i > \bar{u}_i$ for every $i \in N$. Then there exists $\bar{T}$ such that the outcome after T rounds is individually rational for every $T \geqslant \bar{T}$.*

Proof sketch. After Phase 1, the contributions and the transfers no longer change, and $w^T \to w^{\text{OPT}}$ by strong convexity, so $U_i^T \to U_i$ by continuity; the strict inequality $U_i > \bar{u}_i$ then holds after finitely many rounds. The complete proof is in Appendix B.4. □

Lump-sum rebates. When the window of Proposition 1 is empty, the center can restore individual rationality without affecting incentives or budget balance. The reason is that Lemma 1 depends only on the slope $\frac{\partial}{\partial s_i} p_i^1 = \beta$ of the payment, not on its level. For constants $\rho_1, \dots, \rho_n \in \mathbb{R}$, consider the payment

$$p_i^\rho(s) := p_i^1(s) - \rho_i, \qquad \text{where} \qquad \sum_{i \in N} \rho_i = 0 \quad \text{and each } \rho_i \text{ is fixed before Phase 1.} \tag{20}$$

Requiring ρ_i to be fixed in advance, and in particular independent of agent i's own reports, is essential: if ρ_i were computed from agent i's reported final contribution, for instance as $\rho_i = p_i^1(\hat{s})$, it would cancel p_i^1 exactly and remove the incentive to contribute.

Corollary 3 (Individual rationality with lump-sum rebates)**.** *Suppose the hypotheses of Theorem 2 hold, and the payment p^1 is replaced by p^ρ of Equation (20). Then the following hold.*

1. *2P-UPBReD remains budget balanced, Lemma 1 still holds, and all conclusions of Theorem 2 remain valid.*
2. *There exists a choice of rebates for which the limit outcome is individually rational if and only if $\sum_{i \in N} \pi_i \geqslant 0$.*
3. *If the capacities $s_1^{\max}, \dots, s_n^{\max}$ are known to the center before Phase 1, the choice $\rho_i := p_i^1(s^{\max})$ satisfies Equation (20), makes every transfer at the limit outcome zero, and makes the limit outcome individually rational if and only if $\pi_i \geqslant 0$ for every $i \in N$.*

Proof sketch. The rebates are constants that sum to zero, so they change neither the slope of the payment, on which Lemma 1 and Theorem 2 rely, nor its budget balance. Summing the gains $\pi_i + p_i^1(s^{\max}) - \rho_i$ over the agents gives $\sum_i \pi_i$, which proves the necessity in Part 2, and the rebates that split $\sum_i \pi_i$ equally prove its sufficiency. The complete proof is in Appendix B.4. □

Remark (Information required by the rebates). Part 2 of Corollary 3 is a full-information statement: to set the rebates, the center must know each agent's gain π_i, which depends on private costs and stand-alone values. Part 3 needs only the capacities, which agents disclose when registering; since an agent could under-declare its capacity, this part assumes the declared capacities are correct.

5.3 Truthfulness and welfare maximization at NE

Why a truthful version of `2P-UPBReD`. Both `UPBReD` and `2P-UPBReD` rely on what the agents report about the data they use and the gradients they compute. We make `2P-UPBReD` truthful because it is the mechanism that reaches full contribution and learns $w^{\texttt{OPT}}$; truthful reports in `UPBReD` would only certify an equilibrium that may involve under-contribution, as in Example 1. The idea developed below, a cross-check of each agent's contribution report against the other agents' realized valuations together with a budget-balanced charge on gradient reports, applies to `UPBReD` as well: used in every round of `UPBReD`, it gives `T-UPBReD`, whose exposition and formal guarantees (Theorem 4) are in Section 5.4.

`2P-UPBReD` reaches full contribution and learns $w^{\texttt{OPT}}$, but it trusts two kinds of reports: the contributions $\hat{s}_i$, on which the subsidy p^1 is computed, and the gradients $\hat{d}_i^t$, from which the model is computed. In this subsection we ask whether both kinds of reports can be made incentive compatible (truthful) without giving up budget balance. The answer has three parts. First, Proposition 2 shows that no payment rule can reward contributions unless something the center observes depends on the data an agent actually used. Second, we introduce two assumptions under which the other agents' realized valuations provide such evidence, and a payment rule that exploits it through a cross-check, together with another payment that ensures the gradient reports are approximately truthful. Third, Theorem 3 shows that the resulting mechanism, `T-2P-UPBReD`, elicits contributions truthfully, reaches full contribution, learns $w^{\texttt{OPT}}$ under truthful as well as equilibrium gradient reporting, and never runs a deficit.

Why truthful contributions need evidence. Theorem 2 assumes that the contributions $\hat{s}_i$ on which p^1 is computed are the contributions actually used in training. The next result shows that this assumption cannot be removed by a cleverer payment rule alone.

Mechanisms with the structure of `2P-UPBReD`. We consider mechanisms that run in two phases, as `2P-UPBReD` does. Each agent i takes one *action*, its contribution $s_i \in S_i = [s_i^{\min}, s_i^{\max}]$, which is the amount of data it actually uses in training, and it incurs the one-time cost $c_i(s_i)$, in line with the cost accounting of Section 5.2. Throughout the run, agent i also sends *reports*: contribution reports in Phase 1, gradient reports in Phase 2, and possibly further reports, such as valuations, at the end. A *reporting strategy* m_i specifies all of these reports as functions of the information broadcast by the center so far, and M_i denotes the set of reporting strategies available to agent i. Reporting strategies do not take agent i's own contribution, or observations determined by it, as inputs; since agent i chooses the pair (s_i, m_i) jointly, it can still coordinate its reports with its contribution, so this is without loss of generality. The *record* R is everything the center observes during the run: all agents' reports and any additional observations the center may make. The center computes the final parameter $\bar{w} = W(R)$ from the record, for instance by aggregating the reported gradients as in Algorithm 2, and pays each agent a transfer $p_i(R)$. The transfer rule p is arbitrary: it need not be budget balanced, and it may depend on any part of the record. The record may be random, for instance through mini-batch sampling, in which case utilities are understood in expectation.

We fix the strategies of all agents other than i, both their actions and their reporting strategies, and we study agent i's choice of a pair (s_i, m_i). Agent i's utility is

$$u_i = v_i(\bar{w}, s) - c_i(s_i) + p_i(R), \tag{21}$$

where $s = (s_i, s_{-i})$ is the profile of contributions actually used.

If agents also value intermediate models, as in the myopic Phase 2 utility of `T-2P-UPBReD`, the term $\varphi_i(W(R))$ below is replaced by a sum of terms $\varphi_i(w^t)$, each of which is computed from the record; the proof is unchanged.

Definition 8 (Separable valuations)**.** *The valuations are* separable *if, for every agent i, there are functions φ_i and ψ_{ij}, $j \in N$, such that*

$$v_i(w, s) = \varphi_i(w) + \sum_{j \in N} \psi_{ij}(s_j) \qquad \text{for all } w \in \mathbb{R}^m \text{ and } s \in \textstyle\prod_j S_j.$$

The valuations of Section 3 are separable: with the loss Equation (5), $\varphi_i(w) = r_i - L_i(w)$ and $\psi_{ij}(s_j) = -\frac{1}{n} s_j^{-\alpha_i}$, and this holds for each of the losses Equations (6) to (8). Separability has a consequence that we use repeatedly: the gradient $\nabla_w v_i(w, s) = \nabla \varphi_i(w)$ does not depend on the contributions, so a gradient report carries no information about how much data was used, whether it is truthful or not.

Definition 9 (No evidence)**.** *Agent i's contribution leaves* no evidence *if*
(E1) the set M_i of reporting strategies does not depend on s_i; and
(E2) for every reporting strategy $m_i \in M_i$, the joint distribution of the record R and the other agents' contributions s_{-i} under (s_i, m_i) does not depend on s_i.

Condition (E1) says that agent i can send the same reports whatever amount of data it actually used, and condition (E2) says that neither the center nor the other agents receive any information about agent i's contribution except through agent i's own reports.

Remark (When the conditions hold and when they fail). In `2P-UPBReD` with separable valuations, both conditions hold when no agent observes another agent's contribution. Condition (E1) holds because the reports that matter are numbers that agent i can choose freely, and because, by separability, even a truthful gradient $\nabla \varphi_i(w^t)$ is the same for every contribution. Condition (E2) holds because the center's record consists of reports only, and because the other agents neither observe s_i nor, by separability, have their own best contributions depend on it. Each of the following violates a condition and therefore escapes the proposition:

- Assumption 5, under which the other agents report valuations that depend on agent i's actual contribution, violates (E2);
- another agent that observes s_i and reports it violates (E2);
- if gradients are truthful by assumption and the valuations are not separable, so that $\nabla_w v_i(w, s)$ depends on s_i, then the truthful gradient differs across contributions and (E1) fails; such gradients act as evidence in the sense of Green and Laffont [17].

By contrast, a public upper bound on each agent's capacity does *not* violate (E1): it restricts the reports in the same way whatever the agent actually contributes, so it rules out claims beyond $s_i^{\max}$ but not using less data than claimed.

We state the proposition below.

Proposition 2 (No evidence, no incentive)**.** *Suppose the valuations are separable, and agent i's contribution leaves no evidence. For a reporting strategy m_i, let*

$$\Pi_i(m_i) := \mathbb{E}\Big[\varphi_i\big(W(R)\big) + \sum_{j \neq i} \psi_{ij}(s_j) + p_i(R)\Big], \qquad A_i(s_i) := \psi_{ii}(s_i) - c_i(s_i), \tag{22}$$

where the expectation is taken under the pair (s_i, m_i), and assume that these expectations are finite. Then, for every transfer rule p, the following hold.

1. (Transfers cannot depend on the contribution.) *$\Pi_i(m_i)$ does not depend on s_i; in particular, neither does the expected transfer $\mathbb{E}[p_i(R)]$.*
2. (Contributions are as without transfers.) *Agent i's expected utility under (s_i, m_i) equals $A_i(s_i) + \Pi_i(m_i)$. A pair (s_i, m_i) is a best response if and only if s_i maximizes A_i over S_i and m_i maximizes Π_i over M_i. Since A_i does not involve p, agent i's best contributions are the same under every transfer rule, including the rule $p \equiv 0$.*
3. (Full contribution.) *Some transfer rule makes $s_i^{\max}$ part of a best response if and only if every transfer rule does, if and only if $s_i^{\max}$ maximizes A_i over S_i.*
4. (Approximate truthfulness.) *For each $x \in S_i$, let $m_i^T(x) \in M_i$ be a reporting strategy that reports the contribution x. If reporting truthfully is ε-optimal, that is, $\Pi_i\big(m_i^T(s_i)\big) \geqslant \sup_{m \in M_i} \Pi_i(m) - \varepsilon$ for every $s_i \in S_i$, then $\big|\Pi_i\big(m_i^T(x)\big) - \Pi_i\big(m_i^T(y)\big)\big| \leqslant \varepsilon$ for all $x, y \in S_i$.*

Proof sketch. By separability, agent i's realized utility splits into $A_i(s_i)$, which involves neither the transfers nor the reports, and a part that is a function of the record R and the other agents' contributions s_{-i} only. Condition (E2) makes the joint distribution of (R, s_{-i}), and hence the expectation of the second part, independent of s_i, which gives Part 1, and condition (E1) lets agent i choose s_i and m_i independently, so its expected utility $A_i(s_i)+\Pi_i(m_i)$ is maximized separately in each, which gives Parts 2 and 3. Part 4 follows because all values $\Pi_i(m_i^T(x))$ lie in an interval of length ε. The complete proof is in Appendix B.5. □

We now derive two consequences for `2P-UPBReD`.

Corollary 4 (Consequences for `2P-UPBReD`)**.** *In the setting of Proposition 2, consider* `2P-UPBReD` *with truthful gradient reports and the payment p^1 of Equation (12) computed on the reported contributions.*

1. *If, in addition, the valuations are those of Section 3 and the costs are linear, $c_i(s_i) = c_i s_i$, then agent i contributes $s_i^{\mathrm{nt}} := \min\left\{s_i^{\max}, \max\{s_i^{\min}, (\alpha_i/(nc_i))^{1/(\alpha_i+1)}\}\right\}$, whatever the value of β; so full contribution fails whenever $(\alpha_i/(nc_i))^{1/(\alpha_i+1)} < s_i^{\max}$.*
2. *Truthful contribution reporting under p^1 is ε-optimal only if $\varepsilon \geqslant \beta(s_i^{\max} - s_i^{\min})$; with $\beta > \zeta + \tau$, as Theorem 2 requires, this exceeds $(\zeta+\tau)(s_i^{\max} - s_i^{\min})$.*

Proof sketch. Part 1 maximizes $A_i(x) = -\frac{1}{n}x^{-\alpha_i} - c_i x$, which is strictly concave, over S_i. For Part 2, under separability the reported contribution changes neither the model nor the other agents' contributions, so two reporting strategies that differ only in the reported contribution differ in Π_i by exactly $\beta(x-y)$, and the comparison of $s_i^{\min}$ with $s_i^{\max}$ gives the bound. The complete proof is in Appendix B.5. □

Remark (Relation to the literature). Proposition 2 is closely related to the informativeness principle of Holmström [20], which states that an optimal incentive contract for an unobserved action should depend on a signal if and only if the signal is informative about that action. Our statement differs in three respects: it applies to every transfer rule rather than to optimal contracts; the center's information consists largely of reports that the agents choose strategically, rather than of exogenous signals; and it includes an approximate version (Part 4). Any truthful version of `2P-UPBReD` must therefore rely on some observation that depends on the actual contributions.

A natural design follows the two-stage mechanisms for interdependent valuations of Mezzetti [37] and Nath and Zoeter [40]: pay each agent the other agents' reported valuations, and check only its own valuation report against the reported contributions. Here this fails in two ways. An agent that over-reports its contribution and reports the matching valuation passes its own check and keeps the extra subsidy, as Proposition 2 predicts; and paying the others' valuations aligns each agent's utility with the welfare, so an agent gains by bending its reported gradient to move the model towards a better outcome. `T-2P-UPBReD` therefore checks each agent's contribution against the *other* agents' valuations, and controls the gradients with a separate payment.

The payments. `T-2P-UPBReD` uses three transfers. The first is the budget-balanced subsidy of `2P-UPBReD`,

$$p_i^1(\hat{s}) = \beta\Big(\hat{s}_i - \frac{1}{n-1}\sum_{j\neq i}\hat{s}_j\Big). \tag{23}$$

The second is computed once, after training, at the final model $\bar{w} := w^{T_0}$. Each agent reports the contribution $\hat{s}_i$ it used and its realized valuation $\hat{v}_i$ of $\bar{w}$, and pays

$$p_i^2(\hat{s},\hat{v}) = -\big(\hat{v}_i - v_i(\bar{w},\hat{s})\big)^2 - \chi_i(\hat{s},\hat{v})\,\big(\beta\hat{s}_i + K\big), \qquad \chi_i(\hat{s},\hat{v}) = \begin{cases} 1 & \text{if } \hat{v}_j \neq v_j(\bar{w},\hat{s}) \text{ for some } j\neq i,\\ 0 & \text{otherwise,}\end{cases} \tag{24}$$

where $K > 0$ is a constant. The first term is the own-report test of Nath and Zoeter [40], which makes truthful valuation reports strictly optimal. The second term is a *cross-check*: agent i is penalized whenever the *other* agents' valuation reports are inconsistent with the reported contributions. The penalty $\beta\hat{s}_i$ cancels the subsidy that a false contribution report would earn, and K makes the deviation strictly unprofitable. The third transfer is charged in every round t of Phase 2 on the reported gradients,

$$p_i^3(\hat{d}^t) = -\frac{\eta}{2n}\,\|\hat{d}_i^t\|_2^2 + \frac{\eta}{2n(n-1)}\sum_{j\neq i}\|\hat{d}_j^t\|_2^2. \tag{25}$$

The gradient step is the exact maximizer of a surrogate welfare built from the reports,

$$w^{t+1} = w^t + \frac{\eta}{n}\sum_{j\in N}\hat{d}_j^t = \operatorname*{argmax}_{w\in\mathbb{R}^m}\Big\{\sum_{j\in N}(\hat{d}_j^t)^\top(w-w^t) - \frac{n}{2\eta}\,\|w-w^t\|_2^2\Big\}, \tag{26}$$

and the first term of Equation (25) is the Clarke payment of VCG for this surrogate. The second term depends only on the other agents' reports, so it does not affect incentives, and it makes $\sum_i p_i^3 = 0$ for every report profile. Note that p^3 uses only the reported gradients. On the truthful path both penalties in Equation (24) vanish, so agent i's total transfer is $p_i^1(s) + \sum_t p_i^3(d^t)$, and agent i's utility is

$$v_i(\bar{w}, s) - c_i(s_i) + p_i^1(\hat{s}) + p_i^2(\hat{s}, \hat{v}) + \sum_{t<T_0} p_i^3(\hat{d}^t), \tag{27}$$

where the valuation and the cost are evaluated at the true contributions and the transfers at the reports.

The cross-check supplies the evidence that Proposition 2 requires, provided two properties hold: another agent's valuation must react to agent i's contribution, and agents must experience that reaction.

Assumption 4 (Strict interdependence)**.** *For every agent $i \in N$, every parameter w, every s_{-i} and every pair $\hat{s}_i \neq s_i$ in S_i, there exists an agent $j \neq i$ such that $v_j(w, \hat{s}_i, s_{-i}) \neq v_j(w, s_i, s_{-i})$.*

Assumption 5 (Realized payoffs)**.** *After training, each agent j observes its realized valuation $v_j(\bar{w}, s)$ at the contribution profile s that was actually used in training, whatever the agents reported, and the center can evaluate $v_j(\bar{w}, \hat{s})$ at any reported profile $\hat{s}$.*

Remark (Black-box valuations). We view each valuation as a black box: the world produces agent j's realized valuation from the model and the data actually used to train it, whatever the agents report. Agents observe only these realized outputs, while the center knows how the box operates and can evaluate it at any input, in particular at the reported profile, reflecting that it observes outcomes across many agents and deployments. This is the standard primitive of interdependent-value mechanisms [37, 40], and it is why valuations are elicited after training, when the effect of the data actually used is experienced. In our setup, the deterministic $\bar{w}$ of the analysis stands for a model trained on stochastic gradients computed from the samples actually used, and the finite-sample term represents the expected excess loss of that realized model [9]. In the terms of Proposition 2, the black box supplies the required evidence: by Assumption 4, the other agents' realized valuations depend on agent i's actual contribution, so the cross-check can tie the subsidy to the data actually used. Since realized valuations are noisy in practice, the tests in Equation (24) should use a tolerance (see the remark after Theorem 3).

Remark (On strict interdependence). Assumption 4 requires $n \geqslant 2$. It holds for the utility form of Equation (4) with the loss Equation (5): for every $j \neq i$, $\frac{\partial}{\partial s_i} v_j(w, s) = \frac{\alpha_j}{n} s_i^{-\alpha_j - 1} > 0$, so v_j is strictly increasing in s_i.

`T-2P-UPBReD` runs in three stages, given in Algorithm 3.

1. *Phase 1 (contributions), rounds $k = 0, \ldots, \kappa - 1$.* The model stays at w^0. The center broadcasts the current contribution reports $\hat{s}^k$; each agent i updates its actual contribution by a projected gradient step on $v_i - c_i + p_i^1$, evaluated at $(w^0, \hat{s}^k)$, and reports $\hat{s}_i^{k+1}$. No transfers are made in this phase.
2. *Phase 2 (training), rounds $t = 0, \ldots, T_0 - 1$.* The center broadcasts w^t; each agent sends a gradient report $\hat{d}_i^t$, whose truthful value is the gradient computed on the data it actually uses ($s^{\max}$ on the equilibrium path); the center updates $w^{t+1} = w^t + \frac{\eta}{n}\sum_{j\in N}\hat{d}_j^t$ and charges $p^3(\hat{d}^t)$ of Equation (25).
3. *Elicitation, once, at $\bar{w} = w^{T_0}$.* The center broadcasts $\bar{w}$; each agent observes its realized valuation $v_i(\bar{w}, s)$, and all agents send their reports $(\hat{s}_i, \hat{v}_i)$ simultaneously, without seeing each other's reports. The center then pays $p^1(\hat{s}) + p^2(\hat{s}, \hat{v})$ of Equations (23) and (24).

This order has three consequences, which the proof of Theorem 3 uses repeatedly: (T1) contribution reports sent in Phase 1 enter no transfer and do not affect the model; (T2) a gradient report affects only the next model and the charge of its own round; and (T3) the elicitation reports are sent after $\bar{w}$ is computed and all Phase 2 charges are paid, so they affect only p^1 and p^2. Because the elicitation reports are simultaneous, no agent can tailor its valuation report to another agent's contribution report. Under truthful reporting, Phases 1 and 2 coincide with those of `2P-UPBReD`.

The gradient guarantees need one more assumption on the individual valuations.

Assumption 6. *Each $v_i(\cdot, s^{\max})$ is differentiable with M_i-Lipschitz gradient, and $\eta M_i < n$ for all $i \in N$.*

Algorithm 3: Truthful Two Phase Updated Parameter Best Response Dynamics (`T-2P-UPBReD`)

Input: Step sizes γ, η; constant $K > 0$; initialization w^0, $s_i^0 \in S_i$ for $i \in N$; iteration counts κ, T_0 as per Theorem 3

Output: w^{T_0}

`// Phase 1: contributions, with the model fixed at` w^0

Center broadcasts $w^0, \hat{s}^0$; set $k = 0$

while $k < \kappa$ **do**

  Center broadcasts $\hat{s}^k$

  foreach ***agent*** $i \in N$ *in parallel* **do**

    $s_i^{k+1} = s_i^k + \gamma\,[g(w^0, \hat{s}^k, \mu^k)]_i$; send $\hat{s}_i^{k+1}$ to center

  $k \leftarrow k + 1$

`// Phase 2: training, with gradient charges`

for $t = 0$ **to** $T_0 - 1$ **do**

  Center broadcasts w^t

  foreach ***agent*** $i \in N$ *in parallel* **do**

    Send gradient report $\hat{d}_i^t$ (truthful: $\nabla_w v_i(w^t, s^{\max})$)

  Center updates $w^{t+1} = w^t + \frac{\eta}{n} \sum_{i \in N} \hat{d}_i^t$ and pays $p_i^3(\hat{d}^t)$ of Equation (25)

`// Elicitation, at the final model` $\bar{w} = w^{T_0}$

foreach ***agent*** $i \in N$ *in parallel* **do**

  Observe $v_i(\bar{w}, s)$; send $\hat{s}_i$ and $\hat{v}_i$ to center

Center pays $p_i^1(\hat{s}) + p_i^2(\hat{s}, \hat{v})$ of Equations (23) and (24)

return w^{T_0}

In Phase 2 agents are myopic, as elsewhere in the paper: in round t, agent i's utility is $v_i(w^{t+1}, s^{\max}) + p_i^3(\hat{d}^t)$. Write $f(w) := -\frac{1}{n} \sum_{i \in N} v_i(w, s^{\max})$, $\tilde{g}^t := \frac{1}{n} \sum_{j \in N} \nabla_w v_j(w^t, s^{\max})$, and $\Delta := \beta - \tau - \zeta$.

Theorem 3 (`T-2P-UPBReD`: truthfulness, full contribution, and learning)**.** *Suppose Assumptions 2 to 6 hold, $\beta > \zeta + \tau$ and $\eta = 1/M$. Then* `T-2P-UPBReD` *(Algorithm 3) satisfies the following.*

1. (Truthful contribution and valuation reports.) *The elicitation stage is EPIC: if the other agents report truthfully, then for* every *final model $\bar{w}$, and hence for every history of gradient reports, truthful reporting of $(\hat{s}_i, \hat{v}_i)$ is optimal. In particular, an agent that trains on fewer samples than it reports is penalized.*
2. (Full contribution.) *With truthful elicitation, for every w, contributing $s_i^{\max}$ maximizes $v_i(w, s) - c_i(s_i) + p_i^1(s)$ whatever the other agents do, and Phase 1 reaches $s^\kappa = s^{\max}$ for every $w^0 \in \mathbb{R}^m$ and $s^0 \in \prod_{i \in N} S_i$, where*
$$\kappa = \max_{i \in N} \left\lceil \frac{s_i^{\max} - s_i^0}{\gamma\,\Delta} \right\rceil. \tag{28}$$
3. (Approximately truthful gradients.) *In every round t of Phase 2, if the other agents report truthfully, no gradient report gains more than $\varepsilon_i^t := \frac{\eta^3 M_i^2}{2(n - \eta M_i)} \|\tilde{g}^t\|_2^2$ over truthful reporting. Along the truthful path, $\varepsilon_i^t \leqslant C_i \big(1 - \nu/M\big)^t$ with $C_i := \frac{M_i^2 (f(w^0) - f(w^{\texttt{OPT}}))}{M(nM - M_i)}$, so the sum of these one-shot bounds over the rounds of Phase 2 is at most $\frac{M_i^2 (f(w^0) - f(w^{\texttt{OPT}}))}{\nu(nM - M_i)}$.*
4. (Optimal parameter.) *If agents report gradients truthfully, $f(w^{T_0}) - f(w^{\texttt{OPT}}) < \epsilon$ after T_0 rounds, with T_0 as in Equation (13). If instead, in every round, each agent reports the gradient that maximizes its round utility given the other agents' reports, such a profile of reports exists and is unique, and $\|w^t - w^{\texttt{OPT}}\|_2 \leqslant (1 + \eta\nu)^{-t} \|w^0 - w^{\texttt{OPT}}\|_2$.*
5. (Budget.) *On the truthful path, the transfers sum to zero. For every report profile, $\sum_i p_i^3 = 0$ in every round and $\sum_i (p_i^1 + p_i^2) \leqslant 0$, so the center never runs a deficit.*

Proof sketch. Part 1: gradient reports do not depend on $\hat{s}_i$, so agent i's contribution report changes only $p^1 + p^2$. If $\hat{s}_i \neq s_i$, the other agents' realized valuations differ from $v_j(\bar{w}, \hat{s})$ for some j (Assumptions 4 and 5), so $\chi_i = 1$ and the extra subsidy $\beta(\hat{s}_i - s_i)$ is outweighed by the penalty $\beta \hat{s}_i + K$.

Part 2 is Lemma 1, which applies with $\beta > \zeta + \tau$ because no term $\sum_{j \neq i} v_j$ is paid.

Part 3: agent i's round utility is μ-strongly concave in its report with $\mu = \frac{\eta}{n}(1 - \frac{\eta M_i}{n})$, and its gradient at the truthful report is at most $\frac{\eta^2}{n} M_i \|\tilde{g}^t\|_2$.

Part 4: at the round equilibrium every agent reports $\nabla_w v_i(w^{t+1}, s^{\max})$, which makes the update a proximal point step on f.

Part 5: p^1 and p^3 are budget balanced, and both penalties in p^2 are non-negative and vanish on the truthful path. The full proof is in Appendix B.6. □

Example 2 illustrates two guarantees of `T-2P-UPBReD`: checking an agent's valuation report only against its own contribution report does not prevent misreporting, and the charge Equation (25) reduces the gain from misreporting a gradient.

Example 2. Contributions. *Take $n = 3$, $\beta = 2$, $K = 1$, $S_i = [1, 15]$ and the utility form of Equation (4), and let agent i train on $s_i = 5$ samples. Under a rule that checks only agent i's own valuation report, agent i could report $\hat{s}_i = 12$ with the matching valuation, pass its own test, and gain $\beta(12 - 5) = 14$. Under Equation (24), the other two agents' realized valuations are inconsistent with the reported profile, so $\chi_i = 1$ and the same deviation loses at least $\beta s_i + K = 11$. Likewise, an agent that reports 5 but trains on only 3 samples is caught, since the other agents' realized valuations reflect the data actually used, and it loses at least $\beta \cdot 3 + K = 7$ compared with reporting truthfully.*

Gradients. *True contributions are fixed at $s^{\max}$ in Phase 2, so the costs and the finite-sample term of the valuations are constant and do not affect gradient reports; we therefore specify only the parameter-dependent part of the valuations. Take $n = 4$, $m = 2$ and the quadratic valuations of Equation (8), $v_i(w) = -\frac{1}{2}(w - w_i^*)^\top A_i (w - w_i^*)$, with $A_1 = A_2 = \mathrm{diag}(4, \frac{1}{4})$, $A_3 = \mathrm{diag}(\frac{1}{2}, 1)$, $A_4 = \mathrm{diag}(1, \frac{1}{2})$, own optima $w_1^* = (0, 0)$, $w_2^* = (2, 0)$, $w_3^* = (0, 2)$, $w_4^* = (2, 2)$, $w^0 = (-1, -1)$ and $\eta = 1/M$. Here $\nu = 0.5$, $M = 2.375$ and $\eta M_i \leqslant 1.69 < n$. Without p^3, an agent's best gradient misreport steers the model towards its own optimum w_i^*; it gains between 1.82 and 3.34 in round 0, and the gain does not vanish: it is still between 0.41 and 2.49 in round 20, since at $w^{\mathtt{OPT}}$ every agent still prefers its own optimum. With p^3, the best gain is at most 2.50 in round 0, when the update is large, but falls to at most 8.2×10^{-3} in round 1, 1.2×10^{-4} in round 10 and 1.0×10^{-6} in round 20, always within the bound ε_i^t of part 3 of Theorem 3. The charges sum to zero in every round, and both truthful and equilibrium reporting converge to $w^{\mathtt{OPT}} = (1.053, 1.5)$.* □

Remark (Scope and limitations). (i) The guarantees are per round and assume myopic agents; deviations planned across rounds, and rounds in which some agents report gradients truthfully while others best-respond, are not covered. (ii) Realized valuations are noisy in practice, so the tests in Equation (24) need a tolerance; if every misreport larger than ε moves some other agent's valuation by more than the tolerance, such misreports remain unprofitable, while smaller ones gain at most $\beta\varepsilon$. (iii) A truthful agent can be penalized when another agent misreports; this does not affect EPIC, which concerns unilateral deviations. (iv) The Phase 2 charges sum to zero across agents but need not vanish for an individual agent: in each round, agent i's net Phase 2 transfer lies in $\left[-\frac{\eta}{2n} \max_j \|\hat{d}_j^t\|_2^2, \frac{\eta}{2n} \max_j \|\hat{d}_j^t\|_2^2\right]$, so its total over Phase 2 is at most $B := \frac{\eta T_0}{2n} \max_{t,j} \|\hat{d}_j^t\|_2^2$ in absolute value. Hence the limit outcome of `T-2P-UPBReD` is individually rational whenever the gain of Proposition 1 exceeds this amount, that is, whenever $\pi_i + \frac{n\beta}{n-1}\left(s_i^{\max} - \bar{s}^{\max}\right) \geqslant B$ for every $i \in N$.

Remark (Robustness to adversarial gradients). Theorem 3 handles *strategic* gradient reports, which aim to improve the reporting agent's own utility. *Byzantine* agents, which may send arbitrary gradients, are a different threat model; for them, our framework integrates with adversary-robust aggregation rules [16, 25, 55]. We describe one such method, trimmed mean [55], in Appendix C and implement `2P-UPBReD` with trimmed mean in Appendix D.6.

5.4 `T-UPBReD`: truthful reports in the joint dynamics

We now show that an agent can profit from misreporting its contribution in `UPBReD`, even when gradients are reported truthfully and the center aggregates them with uniform weights. Since the allocation is then unaffected by the reports, the manipulation works through the other agents' strategy updates, which use the broadcast contribution profile.

Example 3 (Profitable misreporting in UPBReD)**.** *Consider two agents who learn a model with two parameters and target* $w^\circ = (1, 2)$*. Both agents have the valuation*

$$v_i(w, s) = -\|w - w^\circ\|_2^2 + 10\,(s_1 + s_2) - \tfrac{1}{2}(s_1 + s_2)^2, \qquad i = 1, 2,$$

the costs $c_i(s_i) = s_i + \frac{1}{2}s_i^2$*, and the strategy sets* $S_i = [0, 5]$*. The valuation increases with the total contribution with diminishing returns on the whole domain, because the total never exceeds* 10*; the agents' data are therefore substitutes. The given valuation form satisfies Assumption 1.*

Truthful reporting. *Agent* i*'s marginal utility is* $\frac{\partial}{\partial s_i}u_i = 10 - (s_1 + s_2) - 1 - s_i$*. Setting it to zero for both agents gives the symmetric NE* $s^* = (3, 3)$*, and the parameter converges to* $w^* = w^\circ$*. Each agent's utility is* $10 \cdot 6 - \frac{1}{2} \cdot 6^2 - 3 - \frac{1}{2} \cdot 3^2 = 34.5$*.*

Misreporting. *Suppose agent* 1 *reports* $\hat{s}_1 = 0$ *in every round, while it updates its true contribution using its true value. Agent* 2*'s gradient step is evaluated at the broadcast profile, so its marginal utility becomes* $10 - (0 + s_2) - 1 - s_2 = 9 - 2s_2$*, and agent* 2 *converges to* $s_2 = 4.5$*. Agent* 1*'s marginal utility, evaluated at its true contribution and agent* 2*'s truthful report, is* $9 - s_2 - 2s_1$*, so agent* 1 *converges to* $s_1 = (9 - 4.5)/2 = 2.25$*. The total contribution is* 6.75*, and agent* 1*'s utility is*

$$10 \cdot 6.75 - \tfrac{1}{2} \cdot 6.75^2 - 2.25 - \tfrac{1}{2} \cdot 2.25^2 \;=\; 39.94 \;>\; 34.5,$$

while agent 2*'s utility falls to* 30.09*. By under-reporting, agent* 1 *makes agent* 2 *believe that the total contribution is lower than it is; this raises agent* 2*'s perceived marginal benefit, so agent* 2 *contributes more and agent* 1 *free-rides on the additional data while contributing less itself.* □

UPBReD relies on two kinds of reports in every round: the contributions $\hat{s}_i^t$, which are broadcast and enter the other agents' contribution updates, and the gradients $\hat{d}_i^t$, from which the model is computed. Example 3 shows that an agent can gain by misreporting its contribution, even when gradients are truthful, because the broadcast report changes how the other agents contribute. In this subsection we construct T-UPBReD, which adds to UPBReD per-round transfers that are analogous to those of T-2P-UPBReD (Section 5.3): a cross-check on the reported contributions and a budget-balanced charge on the reported gradients. Unlike T-2P-UPBReD, T-UPBReD pays no subsidy for contributions, so it does not change the equilibrium contributions of UPBReD; its purpose is to make the reports on which UPBReD relies reliable. By Proposition 2, this requires some observation that depends on the contributions actually used, and we use the per-round analogue of Assumption 5.

Timing of a round. T-UPBReD keeps the update rules of UPBReD (Equation (9)) and inserts reports and transfers into every round. Round $t \in \{0, \dots, T-1\}$ proceeds as follows.

1. The center broadcasts the current model w^t and the contribution reports $\hat{s}^t$ collected in the previous round (for $t = 0$, the initial reports $\hat{s}^0$).
2. Each agent i updates its actual contribution by a projected gradient step, $s_i^{t+1} = \Pi_{S_i}\big(s_i^t + \gamma\,\frac{\partial}{\partial s_i}u_i(w^t, s_i^t, \hat{s}_{-i}^t)\big)$, using its own actual contribution and the others' reports. It then sends a gradient report $\hat{d}_i^{t+1}$, whose truthful value is $\nabla_w v_i(w^t, s^t)$, together with a report $\hat{s}_i^{t+1}$ of its new contribution.
3. The center updates the model, $w^{t+1} = w^t + \frac{\eta}{n}\sum_{j \in N}\hat{d}_j^{t+1}$, charges the gradient transfer $p^{3,t}$ of Equation (30) below, and broadcasts w^{t+1}. It does *not* yet broadcast $\hat{s}^{t+1}$.
4. Each agent j observes its realized valuation $v_j(w^{t+1}, s^{t+1})$ and reports $\hat{v}_j^{t+1}$. The center then charges the transfer $p^{2,t}$ of Equation (29) below.

Under truthful reporting, steps 2 and 3 are exactly the updates of Algorithm 1, so T-UPBReD induces the same dynamics as UPBReD. The contribution reports $\hat{s}^{t+1}$ are revealed only in step 1 of the next round, after all valuation reports of round t have been collected, so no agent can tailor its valuation report to another agent's contribution report.

The payments. T-UPBReD uses two transfers in every round. The first is computed on the contribution and valuation reports of the round,

$$p_i^{2,t}\big(\hat{s}^{t+1}, \hat{v}^{t+1}\big) = -\big(\hat{v}_i^{t+1} - v_i(w^{t+1}, \hat{s}^{t+1})\big)^2 - K\,\chi_i^t, \qquad \chi_i^t = \begin{cases} 1 & \text{if } \hat{v}_j^{t+1} \neq v_j(w^{t+1}, \hat{s}^{t+1}) \text{ for some } j \neq i, \\ 0 & \text{otherwise,} \end{cases} \tag{29}$$

Algorithm 4: Truthful Updated Parameter Best Response Dynamics (`T-UPBReD`)

Input: Step sizes γ, η; constant $K > 0$; initialization w^0, $s_i^0 \in S_i$ for $i \in N$; number of rounds T
Output: w^T

Each agent i sends an initial report $\hat{s}_i^0$
for $t = 0$ **to** $T - 1$ **do**
  Center broadcasts w^t and $\hat{s}^t$
  foreach ***agent*** $i \in N$ *in parallel* **do**
    $s_i^{t+1} = \Pi_{S_i}\left(s_i^t + \gamma \frac{\partial}{\partial s_i} u_i(w^t, s_i^t, \hat{s}_{-i}^t)\right)$
    Send gradient report $\hat{d}_i^{t+1}$ (truthful: $\nabla_w v_i(w^t, s^t)$) and contribution report $\hat{s}_i^{t+1}$
  Center updates $w^{t+1} = w^t + \frac{\eta}{n} \sum_{i \in N} \hat{d}_i^{t+1}$, pays $p_i^{3,t}(\hat{d}^{t+1})$ of Equation (30), and broadcasts w^{t+1}
  foreach ***agent*** $i \in N$ *in parallel* **do**
    Observe $v_i(w^{t+1}, s^{t+1})$; send $\hat{v}_i^{t+1}$
  Center pays $p_i^{2,t}(\hat{s}^{t+1}, \hat{v}^{t+1})$ of Equation (29)
return w^T

where $K > 0$ is a constant. As in Equation (24), the first term is the own-report test of Nath and Zoeter [40], which makes truthful valuation reports strictly optimal, and the second term is the cross-check, which penalizes agent i whenever the other agents' valuation reports are inconsistent with the reported contributions. Since `T-UPBReD` pays no subsidy, the penalty consists of the constant K alone, instead of $\beta \hat{s}_i + K$ in `T-2P-UPBReD`. The second transfer is the gradient charge of Equation (25), applied in every round,

$$p_i^{3,t}\left(\hat{d}^{t+1}\right) = -\frac{\eta}{2n} \|\hat{d}_i^{t+1}\|_2^2 + \frac{\eta}{2n(n-1)} \sum_{j \neq i} \|\hat{d}_j^{t+1}\|_2^2. \tag{30}$$

As explained after Equation (26), the model update is the exact maximizer of a surrogate welfare built from the reported gradients, the first term of Equation (30) is the Clarke payment of VCG for this surrogate, and the second term depends only on the other agents' reports and makes the charges sum to zero.

Utilities. As in `UPBReD`, agents are myopic. In round t, agent i's utility is

$$u_i^t = v_i(w^{t+1}, s^{t+1}) - c_i(s_i^{t+1}) + p_i^{2,t}\left(\hat{s}^{t+1}, \hat{v}^{t+1}\right) + p_i^{3,t}\left(\hat{d}^{t+1}\right), \tag{31}$$

where the valuation and the cost are evaluated at the actual contributions and the transfers at the reports. On the truthful path both penalties in Equation (29) vanish, so the only transfer is the gradient charge.

Assumptions. The contribution guarantee uses Assumption 4 and the following per-round version of Assumption 5; the black-box interpretation of Assumption 5 applies unchanged.

Assumption 7 (Realized payoffs in every round)**.** *In every round t, each agent j observes its realized valuation $v_j(w^{t+1}, s^{t+1})$ at the contribution profile s^{t+1} actually chosen, whatever the agents reported, and the center can evaluate $v_j(w^{t+1}, \hat{s})$ at any reported profile $\hat{s}$.*

The gradient guarantee uses the following analogue of Assumption 6. Since contributions change from round to round in `T-UPBReD`, it also bounds how agent i's gradient responds to the contributions.

Assumption 8. *For every agent $i \in N$ there are constants $M_i, Q_i \geqslant 0$ with $\eta M_i < n$ such that, for all $w, w' \in \mathbb{R}^m$ and $s, s' \in \prod_{j \in N} S_j$,*

$$\left\|\nabla_w v_i(w', s) - \nabla_w v_i(w, s)\right\|_2 \leqslant M_i \|w' - w\|_2, \qquad \left\|\nabla_w v_i(w, s') - \nabla_w v_i(w, s)\right\|_2 \leqslant Q_i \|s' - s\|_2.$$

For separable valuations (Definition 8), and in particular for the valuations of Section 3, the gradient $\nabla_w v_i$ does not depend on the contributions, so $Q_i = 0$ and only the first condition remains.

We write $g^t := g(w^t, s^t, \mu^t)$ and $\tilde{g}^t := \tilde{g}(w^t, s^t)$ for the quantities of Equation (9) along the truthful path, and E, W for the constants of Equation (11).

Theorem 4 (T-UPBReD: truthfulness and learning)**.** *Suppose Assumptions 1, 4, 7 and 8 hold, and the step sizes satisfy the conditions of Theorem 1. Then* T-UPBReD *(Algorithm 4) satisfies the following.*

1. (Truthful contribution and valuation reports.) *In every round t, if the other agents report their contributions and valuations truthfully, then for every history and every profile of gradient reports, truthful reporting of $(\hat{s}_i^{t+1}, \hat{v}_i^{t+1})$ maximizes agent i's round utility, and every report $\hat{s}_i^{t+1} \neq s_i^{t+1}$ loses at least K.*
2. (Approximately truthful gradients.) *In every round t, if the other agents report their gradients truthfully, no gradient report gains more than*

$$\varepsilon_i^t := \frac{\eta \left(\eta M_i \|\tilde{g}^t\|_2 + Q_i \|s^{t+1} - s^t\|_2\right)^2}{2(n - \eta M_i)}$$

   over truthful reporting. Along the truthful path, $\varepsilon_i^t \leqslant \frac{\eta \max\{\eta M_i, \gamma Q_i\}^2 E^2}{2(n-\eta M_i)} W^{2t}$, so the sum of these one-shot bounds over all rounds is at most $\frac{\eta \max\{\eta M_i, \gamma Q_i\}^2 E^2}{2(n-\eta M_i)(1-W^2)}$.
3. (Learning.) *If all reports are truthful, the iterates of* T-UPBReD *coincide with those of* UPBReD*, and all conclusions of Theorem 1 hold. If instead the agents report contributions truthfully and gradients according to the Nash equilibrium of every round, this equilibrium exists and is unique; if, moreover, the valuations are separable, then s^t converges to the same PSNE as under truthful reporting, and $\|w^t - w^*\|_2 \leqslant (1+\eta\tilde{\lambda})^{-t} \|w^0 - w^*\|_2$, where w^* is the unique maximizer of $\frac{1}{n}\sum_{i \in N} v_i(\cdot, s)$, which does not depend on s.*
4. (Budget.) *On the truthful path, the transfers sum to zero in every round. For every report profile, $\sum_i p_i^{3,t} = 0$ and $\sum_i p_i^{2,t} \leqslant 0$ in every round, so the center never runs a deficit.*

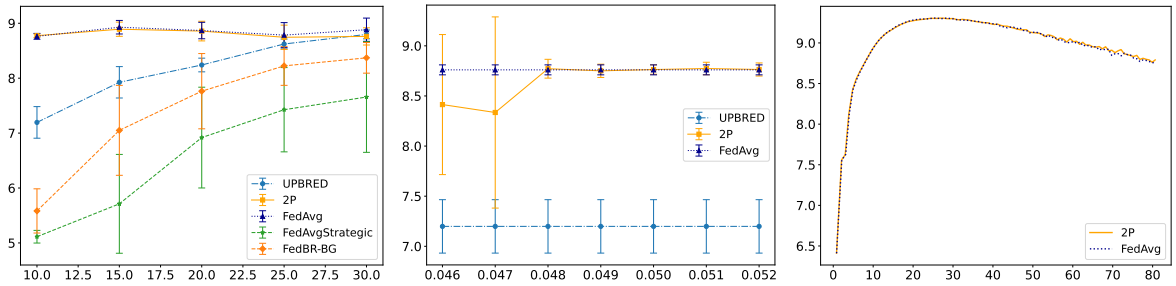
Proof sketch. Part 1 repeats Part 1 of Theorem 3 in every round, with the penalty K alone. In Part 2, both the parameter and the contributions change between the two gradients that determine the gain, which adds the term with Q_i, and the contraction in the proof of Theorem 1 gives the decay W^{2t}. In Part 3, under separability the contribution updates depend neither on the model nor on the other agents' contributions, so they are those of UPBReD, and the round equilibrium of gradient reports is again a proximal point step. Part 4 follows as in Theorem 3. The complete proof is in Appendix A.2. □

Remark (Scope of the guarantees). (i) The guarantees are per round and assume myopic agents. A contribution report also enters the other agents' contribution updates in the next round, which a forward-looking agent could exploit, as in Example 3. Under separable valuations this channel does not exist, since each agent's contribution update does not depend on the others' contributions, so Part 1 rules out every manipulation of contribution reports. Otherwise the penalty K must exceed the per-round gain from such a manipulation; in Example 3, a sustained misreport gains about 5.44 per round once the dynamics have settled, so any $K > 5.44$ makes it unprofitable under Assumption 7. (ii) As in T-2P-UPBReD, realized valuations are noisy in practice, so the tests in Equation (29) should use a tolerance, which turns Part 1 into an approximate guarantee. (iii) T-UPBReD pays no subsidy, so it leaves the equilibrium contributions of UPBReD unchanged; when these are below $s^{\max}$, as in Example 1, the subsidy of T-2P-UPBReD is needed. (iv) Under separable valuations the gradient charges $p^{3,t}$ do not depend on the contributions, so they do not affect the contribution incentives either; for non-separable valuations they can, and the contribution dynamics of T-UPBReD then follow Algorithm 4 as prescribed rather than a best response to the full round utility Equation (31).

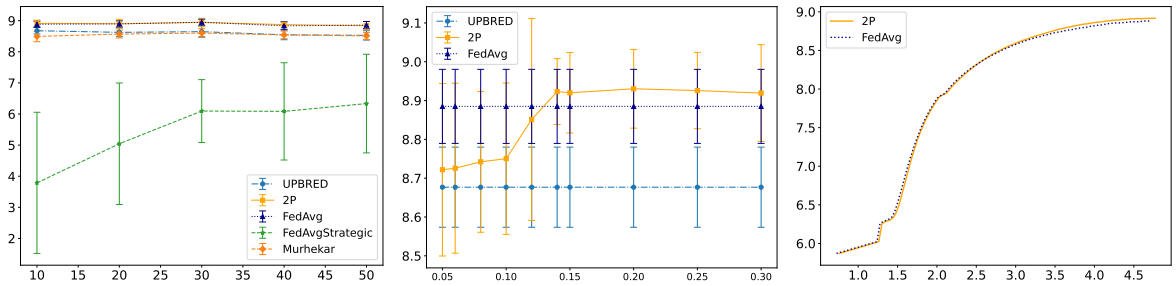
6 Experiments

Our results in Sections 4 and 5 guarantee convergence of Algorithms 1 and 2 to a pure Nash equilibrium under the given assumptions. However, it is important to ask how these FL algorithms behave in practice with real datasets under more general setups. We capture results in these realistic settings in our experiments. We train models on the CIFAR-10 [29], FEMNIST [4], and Twitter [1] datasets and compare the social welfare of UPBReD, 2P-UPBReD, and FedAvg and the FedBR-BG algorithm of Murhekar et al. [38].

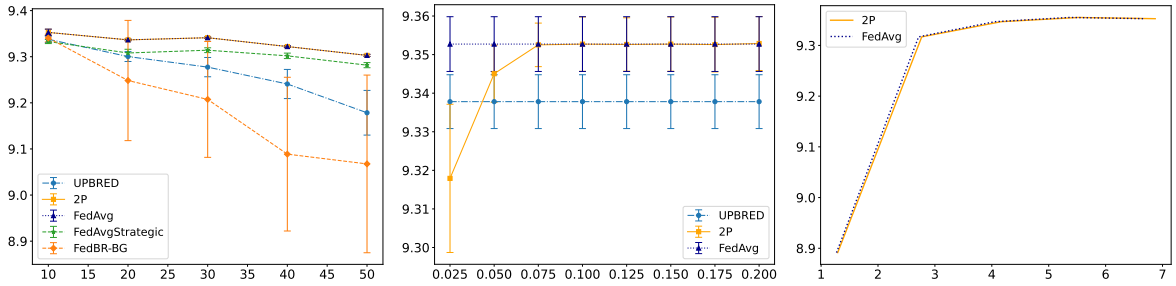
We also consider another algorithm, FedAvgStrategic, as a baseline. In this algorithm, the agents choose to contribute their data strategically as in UPBReD, and stick to this contribution in the entire training phase of FedAvg (details in Appendix D.2).



(a) *CIFAR-10*



(b) *FEMNIST*



(c) *Twitter*

Figure 2: *The three rows in this figure correspond to the experiments with the datasets CIFAR-10, FEMNIST, and Twitter, respectively. The* `y-axis` *in every plot is the social welfare (computed using* $10 - L_i(w)$*, where* $L_i(w)$ *is the test loss for agent* i*). The* `x-axis` *corresponds to the number of agents in the first plot, the choice of* β *in the second, and the time taken (in minutes) in the third plot, respectively, in every row.*

The cost function used is linear $c_i(s_i) = c_i \cdot s_i$, where $c_i = c$ for all agents is sampled from $U[0.01, 0.05]$ (sampled once and shared for all the agents for each run). We choose our valuation function to be $v_i(w^k, s^k) = r_i - L_i(w^k, s^k)$ where the $L_i(w^k, s^k)$ is the loss evaluated on agent i's test dataset at the k-th round of training and r_i is a sufficiently large constant such that u_i remains positive. The social welfare is given by $\sum_{i \in N} v_i(w^k, s^k)$. We perform strategy updates using the numerical derivative as follows: $s_i^{k+1} = s_i^k + \gamma \left(- \frac{L_i(w_i^k, s^k) - L_i(w^{k-1}, s^{k-1})}{s_i^k - s_i^{k-1}} - c_i + \beta \right)$ when $s_i^k \neq s_i^{k-1}$ and $s_i^{k+1} = s_i^k + \gamma \left(-c_i + \beta \right)$ otherwise. $\beta = 0$ for all simulations of `UPBReD` and `FedAvgStrategic`.

The model parameters obtained by agent i after performing one training step locally on the model initialized with w^{k-1} is denoted by w_i^k. We use the cross-entropy loss function for L_i on agent i's dataset. All the experiments reported are performed on *NVIDIA RTX A5000*, *NVIDIA RTX A6000*, and *NVIDIA RTX 5000* GPUs with 24GB, 48GB, and 32GB of memory, respectively. For comparing social welfare against β and number of agents, we randomize over 10 runs. For each run we initialize a seed, which is used to sample costs, initial strategies, model initialization and dataset splits. We report the average running time across 10 runs. We use the Flower framework to create IID splits of the datasets mentioned below.

Dataset specifics. For CIFAR-10 we train a ResNet-18 model with 11,173,962 parameters for 100 rounds. Hyperparameters: $\gamma = 0.01$, $\beta = 2$, $\eta = 0.003$, optimizer=Adam.

For FEMNIST we train a CNN with 6,603,710 parameters for 100 rounds. Hyperparameters: $\gamma =$

0.01, $\beta = 2$, $\eta = 0.001$, optimizer=Adam.

For Twitter we fine-tune all layers of the pretrained `distilbert-base-uncased` model [44] with a sequence-classification head for the three sentiment classes of TweetEval [1] (66,955,779 parameters), truncating inputs to 128 tokens, and train for 5 rounds. Hyperparameters: $\gamma = 0.1$, $\beta = 2$, $\eta = 0.00002$, optimizer=Adam. Unlike the CNNs used in the other datasets, which are trained from scratch over 50–100 rounds, the pre-trained DistilBERT model allows us to achieve full convergence within just 5 communication rounds.

All results are summarized in Figure 2, where the agents' valuations, and hence the social welfare, are computed with respect to the test loss. We provide the social welfare computed using test accuracy in Figure 3 (Appendix D.1). We provide more details of the model architecture for all models in Appendix D.5.

We observe that `2P-UPBReD` and `FedAvg` have the same social welfare performance w.r.t. the number of agents (first column), and running times are nearly the same for all the datasets (third column). `2P-UPBReD` outperforms `UPBReD` in nearly all the runs with respect to social welfare. Additionally, we present a study on how the convergence of `2P-UPBReD` is affected across different values of β and observe that low values of β are sufficient to ensure performance that matches `FedAvg`. `FedBR-BG` performs suboptimally in terms of social welfare compared to `FedAvg` and `2P-UPBReD`. We further comment on results regarding the utilities and payments for agents in Appendix D.4.

7 Conclusions and future work

In this paper, we proposed `UPBReD`, which ensures convergence to a Nash equilibrium with model parameters w, and its modification `T-UPBReD`, which enforces truthful reporting. Since w may deviate from the optimal $w^{\texttt{OPT}}$, we introduced a two-phase mechanism, `2P-UPBReD`, that guarantees full agent contribution and convergence to $w^{\texttt{OPT}}$ via internal agent transfers that are BB, along with its truthful version `T-2P-UPBReD`, which elicits truthful contribution reports, makes gradient reports approximately truthful, and still ensures full contribution, with transfers that are budget balanced on the truthful path. Its guarantee for contribution reports relies on agents observing their realized valuations after training; Proposition 2 shows that some evidence of this kind cannot be avoided. Looking ahead, our framework opens several promising avenues for future research. One natural extension is incorporating agent budget constraints to preserve approximate welfare guarantees, as well as adapting the mechanism to clustered FL architectures. Furthermore, relaxing current trust assumptions, for instance, by designing verification protocols for gradient computation or incorporating center-side learning techniques for unknown agent valuation functions v_i, presents a compelling frontier for building fully decentralized and robust FL mechanisms.

References

[1] Francesco Barbieri, Jose Camacho-Collados, Leonardo Neves, and Luis Espinosa-Anke. Tweeteval: Unified benchmark and comparative evaluation for tweet classification, 2020. URL https://arxiv.org/abs/2010.12421. (Cited on pages 24 and 26)

[2] Avrim Blum, Nika Haghtalab, Richard Lanas Phillips, and Han Shao. One for one, or all for all: Equilibria and optimality of collaboration in federated learning. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pages 1005–1014. PMLR, 18–24 Jul 2021. (Cited on pages 1, 2, and 3)

[3] Marco Bornstein, Amrit Singh Bedi, Abdirisak Mohamed, and Furong Huang. Fact or fiction: Can truthful mechanisms eliminate federated free riding? *Advances in Neural Information Processing Systems*, 37:69206–69229, 2024. (Cited on pages 2 and 3)

[4] Sebastian Caldas, Sai Meher Karthik Duddu, Peter Wu, Tian Li, Jakub Konečný, H. Brendan McMahan, Virginia Smith, and Ameet Talwalkar. Leaf: A benchmark for federated settings, 2019. (Cited on page 24)

[5] Dimitar Chakarov, Nikita Tsoy, Kristian Minchev, and Nikola Konstantinov. Incentivizing truthful collaboration in heterogeneous federated learning, 2025. (Cited on page 3)

[6] Bhaskar Ray Chaudhury, Aniket Murhekar, Zhuowen Yuan, Bo Li, Ruta Mehta, and Ariel D. Procaccia. Fair federated learning via the proportional veto core. In *Forty-first International Conference on Machine Learning*, 2024. (Cited on page 3)

[7] Mengmeng Chen, Xiaohu Wu, Xiaoli Tang, Tiantian He, Yew-Soon Ong, Qiqi Liu, Qicheng Lao, and Han Yu. Free-rider and conflict aware collaboration formation for cross-silo federated learning. *Advances in Neural Information Processing Systems*, 37:54974–55004, 2024. (Cited on page 3)

[8] Yae Jee Cho, Divyansh Jhunjhunwala, Tian Li, Virginia Smith, and Gauri Joshi. To federate or not to federate: Incentivizing client participation in federated learning. In *Workshop on Federated Learning: Recent Advances and New Challenges (in Conjunction with NeurIPS 2022)*, 2022. (Cited on page 3)

[9] Kate Donahue and Jon Kleinberg. Model-sharing games: Analyzing federated learning under voluntary participation. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35:5303–5311, May 2021. (Cited on pages 3, 6, and 19)

[10] Drew Fudenberg. *Game theory*. MIT press, 1991. (Cited on page 7)

[11] Drew Fudenberg and David K Levine. *The theory of learning in games*, volume 2. MIT press, 1998. (Cited on page 4)

[12] Liang Gao, Li Li, Yingwen Chen, Wenli Zheng, ChengZhong Xu, and Ming Xu. Fifl: A fair incentive mechanism for federated learning. In *Proceedings of the 50th International Conference on Parallel Processing*, ICPP '21. Association for Computing Machinery, 2021. ISBN 9781450390682. (Cited on page 3)

[13] Eirini Georgoulaki and Kostas Kollias. Arbitrary profit sharing in federated learning utility games. In *Algorithmic Game Theory*, pages 58–70. Springer Nature Switzerland, 2023. ISBN 978-3-031-43254-5. (Cited on page 3)

[14] Avishek Ghosh, Jichan Chung, Dong Yin, and Kannan Ramchandran. An efficient framework for clustered federated learning. *Advances in Neural Information Processing Systems*, 33:19586–19597, 2020. (Cited on page 3)

[15] Avishek Ghosh, Raj Kumar Maity, Swanand Kadhe, Arya Mazumdar, and Kannan Ramachandran. Communication efficient and byzantine tolerant distributed learning. In *2020 IEEE International Symposium on Information Theory (ISIT)*, pages 2545–2550. IEEE, 2020. (Cited on page 3)

[16] Avishek Ghosh, Raj Kumar Maity, Swanand Kadhe, Arya Mazumdar, and Kannan Ramchandran. Communication-efficient and byzantine-robust distributed learning with error feedback. *IEEE Journal on Selected Areas in Information Theory*, 2(3):942–953, 2021. (Cited on pages 3 and 21)

[17] Jerry R Green and Jean-Jacques Laffont. Partially verifiable information and mechanism design. *The Review of Economic Studies*, 53(3):447–456, 1986. (Cited on page 17)

[18] Pengchao Han, Chao Huang, Geng Tian, Ming Tang, and Xin Liu. Convergence analysis of split federated learning on heterogeneous data. *Advances in Neural Information Processing Systems*, 37: 103476–103544, 2024. (Cited on page 12)

[19] Jordan Hoffmann, Sebastian Borgeaud, Arthur Mensch, Elena Buchatskaya, Trevor Cai, Eliza Rutherford, Diego de Las Casas, Lisa Anne Hendricks, Johannes Welbl, Aidan Clark, Tom Hennigan, Eric Noland, Katie Millican, George van den Driessche, Bogdan Damoc, Aurelia Guy, Simon Osindero, Karen Simonyan, Erich Elsen, Jack W. Rae, Oriol Vinyals, and Laurent Sifre. Training compute-optimal large language models, 2022. URL https://arxiv.org/abs/2203.15556. (Cited on page 6)

[20] Bengt Holmström. Moral hazard and observability. *The Bell journal of economics*, pages 74–91, 1979. (Cited on pages 2 and 18)

[21] Philippe Jehiel and Benny Moldovanu. Efficient design with interdependent valuations. *Econometrica*, 69(5):1237–1259, 2001. doi: https://doi.org/10.1111/1468-0262.00240. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/1468-0262.00240. (Cited on page 5)

[22] Peter Kairouz, H Brendan McMahan, Brendan Avent, Aurélien Bellet, Mehdi Bennis, Arjun Nitin Bhagoji, Kallista Bonawitz, Zachary Charles, Graham Cormode, Rachel Cummings, et al. Advances and open problems in federated learning. *Foundations and trends in machine learning*, 14(1–2):1–210, 2021. (Cited on page 1)

[23] Jared Kaplan, Sam McCandlish, Tom Henighan, Tom B. Brown, Benjamin Chess, Rewon Child, Scott Gray, Alec Radford, Jeffrey Wu, and Dario Amodei. Scaling laws for neural language models. *CoRR*, abs/2001.08361, 2020. (Cited on page 6)

[24] Sai Praneeth Karimireddy, Quentin Rebjock, Sebastian Stich, and Martin Jaggi. Error feedback fixes signsgd and other gradient compression schemes. In *International Conference on Machine Learning*, pages 3252–3261. PMLR, 2019. (Cited on page 3)

[25] Sai Praneeth Karimireddy, Lie He, and Martin Jaggi. Byzantine-robust learning on heterogeneous datasets via bucketing. *arXiv preprint arXiv:2006.09365*, 2020. (Cited on pages 3 and 21)

[26] Sai Praneeth Karimireddy, Satyen Kale, Mehryar Mohri, Sashank Reddi, Sebastian Stich, and Ananda Theertha Suresh. Scaffold: Stochastic controlled averaging for federated learning. In *International conference on machine learning*, pages 5132–5143. PMLR, 2020. (Cited on pages 3, 6, and 12)

[27] Sai Praneeth Karimireddy, Wenshuo Guo, and Michael I Jordan. Mechanisms that incentivize data sharing in federated learning. *arXiv preprint arXiv:2207.04557*, 2022. (Cited on pages 1, 2, 3, and 4)

[28] Jakub Konečnỳ. Federated learning: Strategies for improving communication efficiency. *arXiv preprint arXiv:1610.05492*, 2016. (Cited on pages 1 and 3)

[29] Alex Krizhevsky. Learning multiple layers of features from tiny images. 2009. URL https://api.semanticscholar.org/CorpusID:18268744. (Cited on page 24)

[30] K Naveen Kumar, Ranjeet Ranjan Jha, C Krishna Mohan, and Ravindra Babu Tallamraju. Fortifying federated learning towards trustworthiness via auditable data valuation and verifiable client contribution. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4999–5009, June 2025. (Cited on page 3)

[31] Xiao-Bai Li and Srinivasan Raghunathan. Pricing and disseminating customer data with privacy awareness. *Decision Support Systems*, 59:63–73, 2014. ISSN 0167-9236. (Cited on page 10)

[32] Yuan Liu, Zhengpeng Ai, Shuai Sun, Shuangfeng Zhang, Zelei Liu, and Han Yu. Fedcoin: A peer-to-peer payment system for federated learning. In *Federated learning: privacy and incentive*, pages 125–138. Springer, 2020. (Cited on page 3)

[33] Wuxing Mao, Qian Ma, Guocheng Liao, and Xu Chen. Game analysis and incentive mechanism design for differentially private cross-silo federated learning. *IEEE Transactions on Mobile Computing*, 23(10):9337–9351, 2024. (Cited on page 3)

[34] Michael Maschler, Shmuel Zamir, and Eilon Solan. *Game theory*. Cambridge University Press, 2020. (Cited on page 7)

[35] Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Aguera y Arcas. Communication-efficient learning of deep networks from decentralized data. In *Artificial intelligence and statistics*, pages 1273–1282. PMLR, 2017. (Cited on pages 1, 3, and 7)

[36] Jiajun Meng, Jing Chen, Dongfang Zhao, and Lin Liu. Federated learning and free-riding in a competitive market. *arXiv preprint arXiv:2410.12723*, 2024. (Cited on page 3)

[37] Claudio Mezzetti. Mechanism design with interdependent valuations: Efficiency. *Econometrica*, 72 (5):1617–1626, 2004. (Cited on pages 3, 5, 18, and 19)

[38] Aniket Murhekar, Zhuowen Yuan, Bhaskar Ray Chaudhury, Bo Li, and Ruta Mehta. Incentives in federated learning: Equilibria, dynamics, and mechanisms for welfare maximization. In *Thirty-seventh Conference on Neural Information Processing Systems*, 2023. (Cited on pages 1, 2, 3, 8, 11, and 24)

[39] Aniket Murhekar, Jiaxin Song, Parnian Shahkar, Bhaskar Ray Chaudhury, and Ruta Mehta. You get what you give: Reciprocally fair federated learning. In *Forty-second International Conference on Machine Learning*, 2025. URL https://openreview.net/forum?id=ZdmMDz33Io. (Cited on page 3)

[40] Swaprava Nath and Onno Zoeter. A strict ex-post incentive compatible mechanism for interdependent valuations. *Economics Letters*, 121(2):321–325, 2013. ISSN 0165-1765. doi: https://doi.org/10.1016/j.econlet.2013.08.033. URL https://www.sciencedirect.com/science/article/pii/S0165176513004060. (Cited on pages 3, 18, 19, and 23)

[41] Shashi Raj Pandey, Lam Nguyen, and Petar Popovski. Fedtoken: Tokenized incentives for data contribution in federated learning. In *Workshop on Federated Learning: Recent Advances and New Challenges (in Conjunction with NeurIPS 2022)*, 2022. (Cited on page 3)

[42] Ariel D Procaccia, Han Shao, and Itai Shapira. Incentives in federated learning with heterogeneous agents. *arXiv preprint arXiv:2509.21612*, 2025. (Cited on pages 2 and 3)

[43] Bhaskar Ray Chaudhury, Linyi Li, Mintong Kang, Bo Li, and Ruta Mehta. Fairness in federated learning via core-stability. In *Advances in Neural Information Processing Systems*, volume 35, pages 5738–5750. Curran Associates, Inc., 2022. (Cited on page 3)

[44] Victor Sanh, Lysandre Debut, Julien Chaumond, and Thomas Wolf. Distilbert, a distilled version of bert: smaller, faster, cheaper and lighter, 2020. URL https://arxiv.org/abs/1910.01108. (Cited on page 26)

[45] Sebastian U Stich, Jean-Baptiste Cordonnier, and Martin Jaggi. Sparsified sgd with memory. *Advances in neural information processing systems*, 31, 2018. (Cited on page 3)

[46] Shanli Tan, Hao Cheng, Xiaohu Wu, Han Yu, Tiantian He, Yew Soon Ong, Chongjun Wang, and Xiaofeng Tao. Fedcompetitors: Harmonious collaboration in federated learning with competing participants. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pages 15231–15239, 2024. (Cited on page 3)

[47] Xiaoli Tang and Han Yu. Competitive-cooperative multi-agent reinforcement learning for auction-based federated learning. In *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence, IJCAI-23*, 2023. (Cited on page 3)

[48] Xiaoli Tang and Han Yu. Multi-session multi-objective budget optimization for auction-based federated learning. In *2024 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8, 2024. (Cited on page 3)

[49] Nurbek Tastan, Samar Fares, Toluwani Aremu, Samuel Horváth, and Karthik Nandakumar. Redefining contributions: Shapley-driven federated learning. In *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence, IJCAI-24*. (Cited on page 3)

[50] Stacey Truex, Ling Liu, Ka-Ho Chow, Mehmet Emre Gursoy, and Wenqi Wei. Ldp-fed: Federated learning with local differential privacy. In *Proceedings of the third ACM international workshop on edge systems, analytics and networking*, pages 61–66, 2020. (Cited on page 3)

[51] Kang Wei, Jun Li, Ming Ding, Chuan Ma, Howard H Yang, Farhad Farokhi, Shi Jin, Tony QS Quek, and H Vincent Poor. Federated learning with differential privacy: Algorithms and performance analysis. *IEEE transactions on information forensics and security*, 15:3454–3469, 2020. (Cited on page 3)

[52] Stephen J. Wright and Benjamin Recht. *Optimization for Data Analysis*. Cambridge University Press, 2022. (Cited on pages 13, 39, and 40)

[53] Zhaoxuan Wu, Mohammad Mohammadi Amiri, Ramesh Raskar, and Bryan Kian Hsiang Low. Incentive-aware federated learning with training-time model rewards. In *International Conference on Learning Representations*, volume 2024, pages 38263–38298, 2024. (Cited on page 3)

[54] Yunchao Yang, Yipeng Zhou, Miao Hu, Di Wu, and Quan Z. Sheng. Bara: Efficient incentive mechanism with online reward budget allocation in cross-silo federated learning. In *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence, IJCAI-23*, 2023. (Cited on page 3)

[55] Dong Yin, Yudong Chen, Ramchandran Kannan, and Peter Bartlett. Byzantine-robust distributed learning: Towards optimal statistical rates. In *International conference on machine learning*, pages 5650–5659. Pmlr, 2018. (Cited on pages 3, 12, 21, and 46)

[56] Han Yu, Zelei Liu, Yang Liu, Tianjian Chen, Mingshu Cong, Xi Weng, Dusit Niyato, and Qiang Yang. A fairness-aware incentive scheme for federated learning. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, AIES '20, page 393–399. Association for Computing Machinery, 2020. ISBN 9781450371100. (Cited on page 3)

[57] Chen Zhang, Yu Xie, Hang Bai, Bin Yu, Weihong Li, and Yuan Gao. A survey on federated learning. *Knowledge-Based Systems*, 216:106775, 2021. (Cited on page 1)

A Proofs for Section 4

A.1 Proof of Theorem 1

Proof. Throughout this proof we write $\Delta s^t := s^{t+1} - s^t$ and $\Delta w^t := w^{t+1} - w^t$ for the two increments produced by one round of Algorithm 1, and we abbreviate $g^t := g(w^t, s^t, \mu^t)$ and $\tilde{g}^t := \tilde{g}(w^t, s^t)$. We also write $S := \prod_{i \in N} S_i$, $z^t := (w^t, s^t)$, and $\nabla_s u(z) \in \mathbb{R}^n$ for the pseudo-gradient with entries $[\nabla_s u(z)]_i := \frac{\partial}{\partial s_i} u_i(z)$. We make no assumption that the iterates avoid the boundary of S.

Under Equation (10), the strategy update of Algorithm 1 is the projected gradient step

$$s^{t+1} = \Pi_S\big(s^t + \gamma\, \nabla_s u(z^t)\big), \tag{32}$$

where Π_S acts coordinatewise as $[\Pi_S(x)]_i = \min\{s_i^{\max}, \max\{s_i^{\min}, x_i\}\}$. Two consequences of Equation (32) are used throughout. First, every iterate is feasible, $s^t \in S$ for all t, since each s^{t+1} is the image of a point under Π_S and the parameter w is unconstrained. Second, the update rules Equation (9) give

$$\Delta s^t = \gamma\, g^t, \qquad \Delta w^t = \eta\, \tilde{g}^t, \tag{33}$$

so that $g^t = (s^{t+1} - s^t)/\gamma$ is the displacement actually produced by the step, after clipping. We stress that g is continuous but is *not* differentiable at the faces of S, where the projection has kinks, so we never differentiate g itself. We differentiate only the pseudo-gradient $\nabla_s u$, which is continuously differentiable on $\mathbb{R}^m \times S$ because u_i is twice continuously differentiable by Assumption 1, and whose Jacobians with respect to s and w are, by definition, the matrices G and H.

We use three elementary facts about the coordinatewise projection. First, since S is a product of intervals, $\|x - y\|_2^2 = \sum_{i \in N} (x_i - y_i)^2$ with each y_i constrained only to S_i, so minimizing $\|x - y\|_2$ over $y \in S$ decouples across coordinates, and Π_S coincides with the Euclidean projection onto S. Second, each Π_{S_i} is 1-Lipschitz on $\mathbb{R}$, since clipping to an interval never increases the distance between two points; summing the squared coordinatewise bounds gives the non-expansiveness

$$\|\Pi_S(x) - \Pi_S(y)\|_2 \leqslant \|x - y\|_2 \qquad \text{for all } x, y \in \mathbb{R}^n. \tag{34}$$

Third, for an interval $C = [\ell, h]$, the point $y = \Pi_C(x)$ is characterized by

$$y \in C \quad \text{and} \quad (x - y)(c - y) \leqslant 0 \ \text{ for all } c \in C, \tag{35}$$

as one checks in three cases: if $x \in C$ then $y = x$ and the product vanishes; if $x > h$ then $y = h$, so $x - y > 0$ and $c - y \leqslant 0$; and if $x < \ell$ then $y = \ell$, so $x - y < 0$ and $c - y \geqslant 0$.

Observation 1: $\lambda \leqslant L$ and $\tilde{\lambda} \leqslant \tilde{L}$. To see this, apply condition 1 of Assumption 1 with the test vector $v = e_i$, the i-th standard basis vector of $\mathbb{R}^n$. This gives $e_i^\top \big(G(w,s) + \lambda \mathbb{I}\big) e_i \leqslant 0$, that is, $G(w,s)_{ii} + \lambda \leqslant 0$, and therefore $\lambda \leqslant -G(w,s)_{ii} \leqslant |G(w,s)_{ii}| \leqslant L$, the last step being condition 2(i). The argument for $\tilde{\lambda} \leqslant \tilde{L}$ is identical, using condition 2(ii) in place of (i). Combining $\lambda \leqslant L$ with the standing hypothesis $\lambda > \tilde{P}$ of the theorem gives the chain $\tilde{P} < \lambda \leqslant L \leqslant nL$, and hence

$$n^2 L^2 - \tilde{P}^2 > 0, \tag{36}$$

so that the step-size bound stated in the theorem is a positive quantity. The symmetric chain $P < \tilde{\lambda} \leqslant \tilde{L} \leqslant m\tilde{L}$ gives $m^2 \tilde{L}^2 - P^2 > 0$.

Observation 2: each u_i is λ-strongly concave in its own contribution, and the average valuation is $\tilde{\lambda}$-strongly concave in the parameter. The computation of Observation 1 gives $G(w,s)_{ii} = \frac{\partial^2}{\partial s_i \partial s_i} u_i(w,s) \leqslant -\lambda < 0$ for every (w,s), which is λ-strong concavity of u_i in s_i with w and s_{-i} held fixed. Since $\tilde{G}$ is the Hessian of $\frac{1}{n}\sum_{i \in N} v_i(\cdot, s)$, the second half of condition 1 states that this function is $\tilde{\lambda}$-strongly concave in w. Observation 2 is not used in the contraction argument of Steps 1–8; it is what gives a vanishing gradient mapping its equilibrium content in Step 9.

Step 1: absorbing the projection. We wish to relate the gradient mappings at round $t+1$ to their values at round t. Fix $t \geqslant 0$. By Equation (32) applied at rounds $t+1$ and t, both s^{t+2} and s^{t+1} are images under Π_S:

$$s^{t+2} = \Pi_S\big(s^{t+1} + \gamma \nabla_s u(z^{t+1})\big), \qquad s^{t+1} = \Pi_S\big(s^t + \gamma \nabla_s u(z^t)\big).$$

Subtracting the second identity from the first and applying the non-expansiveness Equation (34) to the two arguments, we obtain

$$\|\Delta s^{t+1}\|_2 \;\leqslant\; \Big\| \Delta s^t + \gamma\big(\nabla_s u(z^{t+1}) - \nabla_s u(z^t)\big) \Big\|_2. \tag{37}$$

The right-hand side is the increment of the smooth map $\nabla_s u$ rather than of the clipped map g. It is this step that allows the argument to proceed without any assumption that the iterates stay in the interior of S.

Step 2: an exact expansion of the smooth maps. Since both arguments of $\nabla_s u$ change between rounds t and $t+1$, we first decompose the joint increment into two increments in which only one argument varies at a time, by adding and subtracting the intermediate value $\nabla_s u(w^t, s^{t+1})$:

$$\nabla_s u(z^{t+1}) - \nabla_s u(z^t) = \underbrace{\Big[\nabla_s u(w^{t+1}, s^{t+1}) - \nabla_s u(w^t, s^{t+1})\Big]}_{\text{only } w \text{ varies}} + \underbrace{\Big[\nabla_s u(w^t, s^{t+1}) - \nabla_s u(w^t, s^t)\Big]}_{\text{only } s \text{ varies}}. \tag{38}$$

Each bracket is the increment of a continuously differentiable map along a straight segment, and both segments lie in the domain on which Assumption 1 is stated: in the first, s is held at the feasible point s^{t+1}; in the second, s ranges over the segment between the feasible points s^t and s^{t+1}, which lies in S because S is convex. This is where the feasibility of the iterates guaranteed by Equation (32) is used; were the iterates allowed to leave S, part of a segment could fall outside the domain of Assumption 1.

To each bracket we apply the integral form of the mean value theorem, which for a continuously differentiable map F with Jacobian DF reads

$$F(b) - F(a) = \left(\int_0^1 DF\big(a + \tau(b-a)\big)\, d\tau \right) (b - a),$$

and follows from the fundamental theorem of calculus applied to $\tau \mapsto F(a + \tau(b-a))$. We emphasize that we use the integral form rather than the more familiar pointwise form $F(b) - F(a) = DF(\xi)(b-a)$. The pointwise form is valid only for scalar-valued F: for a vector-valued map the mean value theorem must be applied coordinatewise, producing a different intermediate point for each coordinate, and the resulting matrix, whose rows are evaluated at different points, is not $G(w,s)$ at any single point (w,s). Since Assumption 1 is a statement about G at a point, it would not apply to such a matrix. The integral form avoids this, as Lemma 2 below makes precise.

Accordingly, define the averaged matrices

$$\begin{aligned}
\overline{G}^{\,t} &:= \int_0^1 G\big(w^t,\, s^t + \tau\Delta s^t\big)\, d\tau \in \mathbb{R}^{n\times n},\\
\overline{H}^{\,t} &:= \int_0^1 H\big(w^t + \tau\Delta w^t,\, s^{t+1}\big)\, d\tau \in \mathbb{R}^{n\times m},\\
\overline{\tilde{G}}^{\,t} &:= \int_0^1 \tilde{G}\big(w^t + \tau\Delta w^t,\, s^t\big)\, d\tau \in \mathbb{R}^{m\times m},\\
\overline{\tilde{H}}^{\,t} &:= \int_0^1 \tilde{H}\big(w^{t+1},\, s^t + \tau\Delta s^t\big)\, d\tau \in \mathbb{R}^{m\times n}.
\end{aligned} \tag{39}$$

Applying the integral form to each bracket of Equation (38), and to the analogous decomposition of $\tilde{g}$ obtained by adding and subtracting $\tilde{g}(w^{t+1}, s^t)$, we obtain the exact identities

$$\begin{aligned}
\nabla_s u(z^{t+1}) - \nabla_s u(z^t) &= \overline{G}^{\,t}\,\Delta s^t + \overline{H}^{\,t}\,\Delta w^t,\\
\tilde{g}^{t+1} - \tilde{g}^t &= \overline{\tilde{G}}^{\,t}\,\Delta w^t + \overline{\tilde{H}}^{\,t}\,\Delta s^t.
\end{aligned} \tag{40}$$

The segments for $\tilde{g}$ also have their s-coordinate in S, by the same convexity argument. A dimension check confirms the shapes: $H(w,s)_{ik} = \frac{\partial^2}{\partial w_k \partial s_i} u_i(w,s)$ is the derivative of the i-th entry of $\nabla_s u$ with respect to w_k, so H is $n\times m$ and $\overline{H}^{\,t}\Delta w^t \in \mathbb{R}^n$ as required; symmetrically $\tilde{H}$ is $m\times n$ and $\overline{\tilde{H}}^{\,t}\Delta s^t \in \mathbb{R}^m$.

The following lemma records that the averaged matrices inherit the properties that Assumption 1 grants to the pointwise ones.

Lemma 2. *Let $A(\tau)$, $\tau\in[0,1]$, be a continuous family of matrices and let $\overline{A} = \int_0^1 A(\tau)\,d\tau$. If $\frac{1}{2}\big(A(\tau) + A(\tau)^\top\big) + \lambda\mathbb{I}$ is negative semi-definite for every τ, then so is $\frac{1}{2}\big(\overline{A} + \overline{A}^{\,\top}\big) + \lambda\mathbb{I}$. If $|A(\tau)_{ij}| \leqslant L$ for every τ and all i,j, then $|\overline{A}_{ij}| \leqslant L$ for all i,j. If $\|A(\tau)\|_{op} \leqslant P$ for every τ, then $\|\overline{A}\|_{op} \leqslant P$.*

Proof. For the first claim, fix any v and observe that $v^\top\overline{A}v = \int_0^1 v^\top A(\tau)v\,d\tau \leqslant \int_0^1(-\lambda\|v\|_2^2)\,d\tau = -\lambda\|v\|_2^2$, where the inequality holds pointwise in τ by hypothesis and integration preserves the inequality. The second and third claims follow from the triangle inequality for integrals: $|\overline{A}_{ij}| = \big|\int_0^1 A(\tau)_{ij}d\tau\big| \leqslant \int_0^1 |A(\tau)_{ij}|d\tau \leqslant L$, and for any unit vector v, $\|\overline{A}v\|_2 \leqslant \int_0^1 \|A(\tau)v\|_2 d\tau \leqslant P$. □

The families in Equation (39) are continuous in τ because u_i and v_i are twice continuously differentiable. We read condition 1 of Assumption 1 as a statement about the quadratic form $v^\top(G+\lambda\mathbb{I})v$, equivalently about the symmetric part $\frac{1}{2}(G+G^\top)+\lambda\mathbb{I}$. This is the relevant reading because $G(w,s)_{ij} = \frac{\partial^2}{\partial s_j \partial s_i} u_i(w,s)$ assembles second derivatives of n *different* functions and is not symmetric in general, and the proof uses the condition only through quadratic forms, which see G only through its symmetric part. The matrix $\tilde{G}$ is the Hessian of $\frac{1}{n}\sum_{i\in N} v_i$ in w and is symmetric, so no such care is needed for it.

Step 3: substituting the algorithm's update rules. Substituting the first line of Equation (40) into Equation (37), and then using Equation (33) to write $\Delta s^t = \gamma g^t$ and $\Delta w^t = \eta\tilde{g}^t$, the vector inside the norm on the right-hand side of Equation (37) becomes

$$\gamma g^t + \gamma\big(\gamma\overline{G}^{\,t} g^t + \eta\overline{H}^{\,t}\tilde{g}^t\big) = \gamma\Big[\big(\mathbb{I}_{n\times n} + \gamma\overline{G}^{\,t}\big)g^t + \eta\,\overline{H}^{\,t}\tilde{g}^t\Big],$$

where we have factored out the common γ and grouped the terms in g^t. On the left-hand side, Equation (33) gives $\|\Delta s^{t+1}\|_2 = \gamma\|g^{t+1}\|_2$. Dividing both sides of Equation (37) by $\gamma > 0$ therefore gives

$$\|g^{t+1}\|_2 \leqslant \big\|\big(\mathbb{I}_{n\times n} + \gamma\overline{G}^{\,t}\big)g^t + \eta\,\overline{H}^{\,t}\tilde{g}^t\big\|_2.$$

For the parameter block there is no projection, so substituting $\Delta w^t = \eta\tilde{g}^t$ and $\Delta s^t = \gamma g^t$ into the second line of Equation (40) and grouping the terms in $\tilde{g}^t$ gives the exact identity

$$\tilde{g}^{t+1} = \big(\mathbb{I}_{m\times m} + \eta\overline{\tilde{G}}^{\,t}\big)\tilde{g}^t + \gamma\,\overline{\tilde{H}}^{\,t} g^t.$$

Taking norms in the latter and applying the triangle inequality to both, we arrive at

$$\begin{aligned}
\|g^{t+1}\|_2 &\leqslant \big\|\big(\mathbb{I}_{n\times n} + \gamma\overline{G}^{\,t}\big)g^t\big\|_2 + \eta\big\|\overline{H}^{\,t}\tilde{g}^t\big\|_2,\\
\|\tilde{g}^{t+1}\|_2 &\leqslant \big\|\big(\mathbb{I}_{m\times m} + \eta\overline{\tilde{G}}^{\,t}\big)\tilde{g}^t\big\|_2 + \gamma\big\|\overline{\tilde{H}}^{\,t} g^t\big\|_2.
\end{aligned} \tag{41}$$

The remainder of the argument consists in bounding these four terms.

Step 4: bounding the diagonal terms. Consider the square of the first term on the right-hand side of the first line of Equation (41). Expanding the squared norm of a sum in the usual way,

$$\left\|\left(\mathbb{I}+\gamma\overline{G}^{\,t}\right)g^t\right\|_2^2=\|g^t\|_2^2+\gamma^2\left\|\overline{G}^{\,t}g^t\right\|_2^2+2\gamma\,(g^t)^\top\overline{G}^{\,t}g^t. \tag{42}$$

We bound the second and third terms in turn. For the second term, we use the fact that the operator norm of a matrix is dominated by its Frobenius norm, so that $\|\overline{G}^{\,t}g^t\|_2^2\leqslant\|\overline{G}^{\,t}\|_F^2\|g^t\|_2^2$, where $\|A\|_F:=\sqrt{\sum_i\sum_j|A_{ij}|^2}$. By Lemma 2 applied to condition 2(i) of Assumption 1, every entry of $\overline{G}^{\,t}$ is bounded in absolute value by L, and since $\overline{G}^{\,t}$ has n^2 entries we conclude that $\|\overline{G}^{\,t}\|_F^2\leqslant n^2L^2$. For the third term, we use Lemma 2 applied to condition 1 of Assumption 1: the averaged matrix satisfies $\frac{1}{2}(\overline{G}^{\,t}+\overline{G}^{\,t^\top})+\lambda\mathbb{I}\preceq 0$, and testing this quadratic form against the particular vector $v=g^t$ gives

$$(g^t)^\top\overline{G}^{\,t}g^t\leqslant-\lambda\|g^t\|_2^2. \tag{43}$$

Substituting both bounds into Equation (42) and collecting the common factor $\|g^t\|_2^2$,

$$\left\|\left(\mathbb{I}+\gamma\overline{G}^{\,t}\right)g^t\right\|_2^2\leqslant\left(1+\gamma^2n^2L^2-2\gamma\lambda\right)\|g^t\|_2^2. \tag{44}$$

Write $a:=1+\gamma^2n^2L^2-2\gamma\lambda$ for the bracketed constant. Before taking square roots we must check that $a\geqslant 0$; this requires no additional hypothesis. Viewing a as a quadratic in γ with positive leading coefficient, its discriminant is $4\lambda^2-4n^2L^2$, which is non-positive because Observation 1 gives $\lambda\leqslant L\leqslant nL$. Hence $a\geqslant 0$ for every choice of γ, and we may write

$$\left\|\left(\mathbb{I}+\gamma\overline{G}^{\,t}\right)g^t\right\|_2\leqslant\sqrt{a}\,\|g^t\|_2. \tag{45}$$

The identical computation applied to the first term of the second line of Equation (41), using conditions 1 and 2(ii) of Assumption 1 in place of 2(i), yields

$$\left\|\left(\mathbb{I}+\eta\overline{\tilde{G}}^{\,t}\right)\tilde{g}^t\right\|_2\leqslant\sqrt{\tilde{a}}\,\|\tilde{g}^t\|_2,\qquad\tilde{a}:=1+\eta^2m^2\tilde{L}^2-2\eta\tilde{\lambda}, \tag{46}$$

with $\tilde{a}\geqslant 0$ by the same discriminant argument applied to $\tilde{\lambda}\leqslant\tilde{L}\leqslant m\tilde{L}$.

Step 5: bounding the cross terms. The two remaining terms of Equation (41) are controlled directly by the operator norm bounds. By the definition of the operator norm and Lemma 2 applied to condition 2(iii) of Assumption 1,

$$\eta\left\|\overline{H}^{\,t}\tilde{g}^t\right\|_2\leqslant\eta\left\|\overline{H}^{\,t}\right\|_{op}\|\tilde{g}^t\|_2\leqslant\eta P\|\tilde{g}^t\|_2=\sqrt{b}\,\|\tilde{g}^t\|_2,\qquad b:=P^2\eta^2,$$

and symmetrically, using condition 2(iv),

$$\gamma\left\|\overline{\tilde{H}}^{\,t}g^t\right\|_2\leqslant\gamma\tilde{P}\|g^t\|_2=\sqrt{\tilde{b}}\,\|g^t\|_2,\qquad\tilde{b}:=\tilde{P}^2\gamma^2.$$

Step 6: putting them together and establishing the contraction factor. Substituting the four bounds of Steps 4 and 5 into Equation (41), we obtain the coupled system

$$\begin{aligned}\|g^{t+1}\|_2&\leqslant\sqrt{a}\,\|g^t\|_2+\sqrt{b}\,\|\tilde{g}^t\|_2,\\ \|\tilde{g}^{t+1}\|_2&\leqslant\sqrt{\tilde{a}}\,\|\tilde{g}^t\|_2+\sqrt{\tilde{b}}\,\|g^t\|_2.\end{aligned} \tag{47}$$

Adding the two inequalities and grouping the right-hand side by which of $\|g^t\|_2$ and $\|\tilde{g}^t\|_2$ each term multiplies, we get

$$\|g^{t+1}\|_2+\|\tilde{g}^{t+1}\|_2\leqslant\left(\sqrt{a}+\sqrt{\tilde{b}}\right)\|g^t\|_2+\left(\sqrt{\tilde{a}}+\sqrt{b}\right)\|\tilde{g}^t\|_2,$$

and bounding each of the two coefficients by their maximum,

$$\|g^{t+1}\|_2+\|\tilde{g}^{t+1}\|_2\leqslant\max\left\{\sqrt{a}+\sqrt{\tilde{b}},\ \sqrt{\tilde{a}}+\sqrt{b}\right\}\Big(\|g^t\|_2+\|\tilde{g}^t\|_2\Big). \tag{48}$$

Recalling the definitions preceding Theorem 1, the two quantities inside the maximum are precisely $W_1=\sqrt{a}+\sqrt{\tilde{b}}$ and $W_2=\sqrt{\tilde{a}}+\sqrt{b}$, and $W=\max\{W_1,W_2\}$. It is worth noting that W_1 depends only on γ and W_2 only on η, so the two step sizes may be chosen independently of one another.

Step 7: ensuring the contraction factor $W < 1$. It remains to determine when $W_1 < 1$ and $W_2 < 1$. We treat W_1; the argument for W_2 is obtained by the substitutions $\gamma \mapsto \eta$, $n \mapsto m$, $L \mapsto \tilde{L}$, $\lambda \mapsto \tilde{\lambda}$, $\tilde{P} \mapsto P$.

Since $\gamma > 0$ and $\tilde{P} \geqslant 0$ we have $\sqrt{\tilde{b}} = \gamma\tilde{P}$, so the requirement $W_1 < 1$ reads

$$\sqrt{1 + \gamma^2 n^2 L^2 - 2\gamma\lambda} < 1 - \gamma\tilde{P}. \tag{49}$$

A necessary condition for Equation (49) is that the right-hand side be positive, which is the requirement $\gamma < 1/\tilde{P}$. Given that, both sides of Equation (49) are non-negative, so squaring is an equivalence rather than merely an implication, and Equation (49) holds if and only if

$$1 + \gamma^2 n^2 L^2 - 2\gamma\lambda < 1 - 2\gamma\tilde{P} + \gamma^2\tilde{P}^2.$$

Cancelling the constant 1 from both sides and then dividing through by $\gamma > 0$,

$$\gamma n^2 L^2 - 2\lambda < -2\tilde{P} + \gamma\tilde{P}^2.$$

Collecting the terms carrying γ on the left and the constants on the right,

$$\gamma\big(n^2 L^2 - \tilde{P}^2\big) < 2(\lambda - \tilde{P}),$$

and dividing by $n^2 L^2 - \tilde{P}^2$, which is strictly positive by Equation (36), we arrive at the exact characterization

$$W_1 < 1 \quad \Longleftrightarrow \quad \gamma < \frac{2(\lambda - \tilde{P})}{n^2 L^2 - \tilde{P}^2}. \tag{50}$$

Note that the hypothesis $\lambda > \tilde{P}$ of the theorem is exactly what makes the right-hand side of Equation (50) positive, so that a feasible γ exists at all.

The auxiliary condition $\gamma < 1/\tilde{P}$ identified above is implied by Equation (50) and therefore need not be imposed separately. Indeed, it suffices to check that $2\tilde{P}(\lambda - \tilde{P}) \leqslant n^2 L^2 - \tilde{P}^2$, that is, $2\tilde{P}\lambda - \tilde{P}^2 \leqslant n^2 L^2$. Completing the square on the left, $2\tilde{P}\lambda - \tilde{P}^2 = \lambda^2 - (\lambda - \tilde{P})^2 \leqslant \lambda^2$, and $\lambda^2 \leqslant L^2 \leqslant n^2 L^2$ by Observation 1, which gives the claim. By the symmetric computation, $\eta < 2(\tilde{\lambda} - P)/(m^2\tilde{L}^2 - P^2)$ likewise implies $\eta < 1/P$.

Consequently, choosing

$$\gamma < \frac{2(\lambda - \tilde{P})}{n^2 L^2 - \tilde{P}^2} \quad \text{and} \quad \eta < \frac{2(\tilde{\lambda} - P)}{m^2\tilde{L}^2 - P^2} \tag{51}$$

guarantees $W_1 < 1$ and $W_2 < 1$, hence $W = \max\{W_1, W_2\} < 1$, and Equation (48) becomes a genuine contraction:

$$\|g^{t+1}\|_2 + \|\tilde{g}^{t+1}\|_2 \leqslant W\Big(\|g^t\|_2 + \|\tilde{g}^t\|_2\Big), \qquad W < 1. \tag{52}$$

Step 8: iterating the contraction over t. Since Equation (52) holds for every $t \geqslant 0$, applying it repeatedly, starting from $t = 0$ and iterating T times, gives

$$\|g^T\|_2 + \|\tilde{g}^T\|_2 \leqslant W^T\Big(\|g^0\|_2 + \|\tilde{g}^0\|_2\Big) = W^T E, \tag{53}$$

where $E := \|g(w^0, s^0, \mu^0)\|_2 + \|\tilde{g}(w^0, s^0)\|_2$ as defined preceding Theorem 1. Here $g^0 = (s^1 - s^0)/\gamma$ is the clipped step at the initial point, so that E depends on γ; this is harmless, since E is finite for every initialization. To make the right-hand side of Equation (53) smaller than ϵ we require $W^T E < \epsilon$. Taking logarithms of both sides and using that $\ln W < 0$, so that dividing by it reverses the inequality, this is equivalent to

$$T \ln \tfrac{1}{W} > \ln \tfrac{E}{\epsilon}, \qquad \text{that is,} \qquad T > T_0(w^0, s^0) := \frac{\ln\left(E/\epsilon\right)}{\ln\left(1/W\right)}.$$

Hence for every $T > T_0(w^0, s^0)$ we have $\|g^T\|_2 + \|\tilde{g}^T\|_2 < \epsilon$, and since both summands are non-negative, each is individually smaller than ϵ:

$$\|g(w^T, s^T, \mu^T)\|_2 < \epsilon \quad \text{and} \quad \|\tilde{g}(w^T, s^T)\|_2 < \epsilon. \tag{54}$$

Step 9: from a small gradient mapping to a Nash equilibrium. Equation (54) states that the step actually taken by each agent, after clipping, is small. We now convert this into the equilibrium statements

of the theorem. Because $[g]_i$ is the clipped step rather than the derivative $\frac{\partial}{\partial s_i} u_i$, a small $[g]_i$ does not imply a small derivative: at a best response on the boundary of S_i the derivative need not vanish at all. The right certificate of equilibrium is therefore the gradient mapping itself, and the following two lemmas make this precise, first qualitatively and then quantitatively.

Lemma 3 (Fixed points of the projected step are Nash equilibria)**.** *Under Assumption 1, for every $\gamma > 0$, $w \in \mathbb{R}^m$ and $s \in S$, we have $g(w, s, \mu) = 0$ if and only if s is a Nash equilibrium at the parameter w in the sense of Definition 1.*

Proof. Since Π_S acts coordinatewise, $g(w, s, \mu) = 0$ holds if and only if $[g(w, s, \mu)]_i = 0$ for every $i \in N$. Fix an agent i and write $\varphi(x) := u_i(w, x, s_{-i})$ for its utility as a function of its own contribution alone, so that $\frac{\partial}{\partial s_i} u_i(w, s) = \varphi'(s_i)$. By Equation (10), $[g(w, s, \mu)]_i = \big(\Pi_{S_i}(s_i + \gamma\varphi'(s_i)) - s_i\big)/\gamma$ is the clipped step divided by γ, so

$$[g(w, s, \mu)]_i = 0 \quad \Longleftrightarrow \quad s_i = \Pi_{S_i}\big(s_i + \gamma\varphi'(s_i)\big),$$

that is, the fixed-point condition states precisely that s_i is the projection of $s_i + \gamma\varphi'(s_i)$ onto S_i. We now apply the characterization Equation (35), which is an equivalence, with $C = S_i$, $x = s_i + \gamma\varphi'(s_i)$ and $y = s_i$. Its membership requirement $y \in C$ holds because $s \in S$ by hypothesis, and $x - y = \gamma\varphi'(s_i)$, so the fixed-point condition is equivalent to $\gamma\varphi'(s_i)(c - s_i) \leqslant 0$ for all $c \in S_i$. Since $\gamma > 0$ the factor γ may be dropped:

$$\varphi'(s_i)\,(c - s_i) \leqslant 0 \quad \text{for all } c \in S_i. \tag{55}$$

It remains to show that Equation (55) characterizes the maximizers of φ over S_i, and this is where Observation 2 enters: φ is concave on S_i. For a concave function the first-order condition Equation (55) is sufficient for optimality over a convex set, by concavity of φ:

$$\begin{aligned} \varphi(c) &\leqslant \varphi(s_i) + \varphi'(s_i)(c - s_i) \\ &\leqslant \varphi(s_i) \qquad\qquad \text{using Equation (55)} \end{aligned}$$

for every $c \in S_i$. It is also necessary: if $\varphi'(s_i)(c - s_i) > 0$ for some $c \in S_i$, then a small move from s_i toward c, which stays in S_i by convexity, strictly increases φ. Hence Equation (55) holds if and only if $s_i \in \operatorname{argmax}_{x \in S_i} u_i(w, x, s_{-i})$, and as this holds for every i, the condition $g = 0$ is equivalent to every agent playing a best response, which is Definition 1. □

Lemma 4 (Error bound for the gradient mapping)**.** *Under Assumption 1, for every $\gamma > 0$, $w \in \mathbb{R}^m$, $s \in S$ and $i \in N$, let $s_i^\star(w, s) := \operatorname{argmax}_{x \in S_i} u_i(w, x, s_{-i})$ be agent i's best response and $s^\star(w, s)$ the profile of best responses. Then*

$$\big|s_i - s_i^\star(w, s)\big| \;\leqslant\; \frac{1 + \gamma L}{\lambda}\,\big|[g(w, s, \mu)]_i\big|, \tag{56}$$

and consequently $\|s - s^\star(w, s)\|_2 \leqslant \frac{1+\gamma L}{\lambda}\,\|g(w, s, \mu)\|_2$.

Proof. As in Lemma 3 write $\varphi(x) := u_i(w, x, s_{-i})$, set $F := -\varphi'$, and abbreviate $s_i^\star := s_i^\star(w, s)$. Two properties of F follow from Assumption 1. It is strongly monotone with modulus λ: by the mean value theorem, which is valid here since F is scalar-valued,

$$(F(x) - F(y))(x - y) = -\varphi''(\xi)(x - y)^2 \geqslant \lambda(x - y)^2$$

using $\varphi'' \leqslant -\lambda$ from Observation 2. And it is Lipschitz with constant L, since $|F'| = |\varphi''| = |G_{ii}| \leqslant L$ by condition 2(i).

Write $x := s_i$, let $x^+ := \Pi_{S_i}(x - \gamma F(x))$ be the projected step, and let $r := x - x^+$, so that $|r| = \gamma|[g]_i|$ by Equation (33). We combine two inequalities. First, the characterization Equation (35) of x^+, tested at the point $c = s_i^\star$, gives $\big(x - \gamma F(x) - x^+\big)(s_i^\star - x^+) \leqslant 0$, which rearranges to $\gamma F(x)(x^+ - s_i^\star) \leqslant r(x^+ - s_i^\star)$. Second, the optimality of $s_i^\star$, in the form Equation (55) tested at the point $c = x^+$, gives $\gamma F(s_i^\star)(x^+ - s_i^\star) \geqslant 0$. Subtracting the second from the first,

$$\gamma\big(F(x) - F(s_i^\star)\big)(x^+ - s_i^\star) \;\leqslant\; r\,(x^+ - s_i^\star).$$

We now substitute $x^+ - s_i^\star = (x - s_i^\star) - r$ on both sides and expand. The left-hand side becomes $\gamma(F(x) - F(s_i^\star))(x - s_i^\star) - \gamma(F(x) - F(s_i^\star))\, r$, where the first term satisfies $\gamma(F(x) - F(s_i^\star))(x - s_i^\star) \geq \gamma\lambda|x - s_i^\star|^2$ by strong monotonicity and whose second term satisfies $-\gamma(F(x) - F(s_i^\star))r \geq -\gamma L|x - s_i^\star|\,|r|$ by the Lipschitz bound. The right-hand side becomes $r(x - s_i^\star) - r^2 \leqslant |r|\,|x - s_i^\star|$. Combining,

$$\gamma\lambda|x - s_i^\star|^2 - \gamma L|x - s_i^\star|\,|r| \;\leqslant\; |r|\,|x - s_i^\star|.$$

If $x = s_i^\star$ there is nothing to prove. Otherwise, dividing by $\gamma\lambda|x - s_i^\star| > 0$ and collecting terms gives $|x - s_i^\star| \leqslant \frac{1+\gamma L}{\gamma\lambda}|r|$, and substituting $|r| = \gamma|[g]_i|$ yields Equation (56). The vector form follows by squaring Equation (56), summing over $i \in N$, and taking square roots. □

With these two lemmas in hand we complete the proof. For convergence of the iterates, Equation (33) and Equation (53) give $\|\Delta s^t\|_2 + \|\Delta w^t\|_2 \leqslant \max\{\gamma, \eta\} W^t E$. Since $W < 1$ the geometric series $\sum_{t\geqslant 0} W^t$ converges, so the increments of both sequences are absolutely summable, and $(s^t)_t$ and $(w^t)_t$ are Cauchy sequences in finite-dimensional spaces; since $\mathbb{R}^n$ and $\mathbb{R}^m$ are complete, they therefore converge, to limits s^* and w^*, and $s^* \in S$ because S is closed. The map $z \mapsto \big(\Pi_S(s + \gamma\nabla_s u(z)) - s\big)/\gamma$ is continuous, as a composition of the continuous maps Π_S and $\nabla_s u$, and so is $\tilde{g}$. Passing to the limit in Equation (53) therefore gives $g(w^*, s^*, \mu^*) = 0$ and $\tilde{g}(w^*, s^*) = 0$.

For the equilibrium property, Lemma 3 shows that s^* is a Nash equilibrium at the parameter w^*. For the second, Observation 2 states that $\frac{1}{n}\sum_{i\in N} v_i(\cdot, s^*)$ is $\tilde{\lambda}$-strongly concave, and a stationary point of a strongly concave function is its unique global maximizer; hence w^* is the unique maximizer of $\frac{1}{n}\sum_{i\in N} v_i(\cdot, s^*)$.

Finally, Lemma 4 quantifies the equilibrium property at any finite T. For every $T > T_0(w^0, s^0)$, combining Equation (56) with Equation (54) gives

$$\big\|s^T - s^\star(w^T, s^T)\big\|_2 \;\leqslant\; \frac{1+\gamma L}{\lambda}\,\|g^T\|_2 \;<\; \frac{1+\gamma L}{\lambda}\,\epsilon, \tag{57}$$

that is, every agent's contribution lies within $\frac{1+\gamma L}{\lambda}\epsilon$ of its best response to the others at the parameter w^T. This completes the proof. □

A.2 Proof of Theorem 4

Proof. Throughout, *truthful reporting* in round t means $\hat{s}_i^{t+1} = s_i^{t+1}$, $\hat{d}_i^{t+1} = \nabla_w v_i(w^t, s^t)$ and $\hat{v}_i^{t+1} = v_i(w^{t+1}, s^{t+1})$. The proof uses three consequences of the timing of a round, which play the role of (T1)–(T3) in the proof of Theorem 3.

(R1) Agent i's new contribution s_i^{t+1} is computed in step 2 from $(w^t, s_i^t, \hat{s}_{-i}^t)$, before any report of round t is sent, so no report of round t affects it.

(R2) The model w^{t+1} and the charge $p^{3,t}$ depend only on the gradient reports $\hat{d}^{t+1}$.

(R3) The contribution and valuation reports of round t affect the round-t utility only through $p^{2,t}$. They also enter the other agents' updates in round $t+1$, which a myopic agent does not take into account; see the remark after the proof.

Part 1 (truthful contribution and valuation reports). Fix a round t, an agent i, an arbitrary history and an arbitrary profile of gradient reports, so that w^{t+1} and the actual contributions s^{t+1} are fixed. Suppose the other agents report $\hat{s}_{-i}^{t+1} = s_{-i}^{t+1}$ and $\hat{v}_j^{t+1} = v_j(w^{t+1}, s^{t+1})$, which they can do by Assumption 7. Write $\hat{s} := (\hat{s}_i^{t+1}, s_{-i}^{t+1})$.

Step 1: only $p^{2,t}$ changes. By (R1), agent i's reports do not affect s^{t+1}; by (R2), they do not affect w^{t+1} or $p^{3,t}$. Hence its valuation $v_i(w^{t+1}, s^{t+1})$, its cost $c_i(s_i^{t+1})$ and its gradient charge are the same under every report $(\hat{s}_i^{t+1}, \hat{v}_i^{t+1})$, and by (R3) the difference between its round utility under such a report and under truthful reporting equals the difference in $p_i^{2,t}$.

Step 2: $p^{2,t}$ vanishes under truthful reporting. Under truthful reporting the reported profile is s^{t+1}. Agent i's own report $v_i(w^{t+1}, s^{t+1})$ coincides with the center's evaluation, so the quadratic term vanishes, and every other agent's report coincides with the center's evaluation as well, so $\chi_i^t = 0$. Combining with Step 1,

$$u_i^t - u_i^{t,\mathrm{truth}} = -\big(\hat{v}_i^{t+1} - v_i(w^{t+1}, \hat{s})\big)^2 - K\chi_i^t. \tag{58}$$

Case 1: $\hat{s}_i^{t+1} = s_i^{t+1}$. Then $\hat{s} = s^{t+1}$, each other agent's report equals the center's evaluation at $\hat{s}$, and so $\chi_i^t = 0$. The difference Equation (58) reduces to $-\big(\hat{v}_i^{t+1} - v_i(w^{t+1}, s^{t+1})\big)^2 \leqslant 0$, with equality if and only if agent i reports its true valuation.

Case 2: $\hat{s}_i^{t+1} \neq s_i^{t+1}$. By Assumption 4, applied at the parameter w^{t+1} with the other agents' contributions s_{-i}^{t+1}, there is an agent $j \neq i$ with $v_j(w^{t+1}, \hat{s}) \neq v_j(w^{t+1}, s^{t+1})$. Agent j reports the latter, so its report fails the center's check and $\chi_i^t = 1$, whatever agent i reports, since χ_i^t does not depend on $\hat{v}_i^{t+1}$. Substituting into Equation (58) and dropping the non-positive quadratic term, $u_i^t - u_i^{t,\mathrm{truth}} \leqslant -K < 0$.

In both cases truthful reporting is optimal, and every misreport of the contribution loses at least K. Since the history and the gradient reports were arbitrary, this holds in every round.

Part 2 (approximately truthful gradients). Fix a round t and an agent i, and suppose the other agents report $d_j^{t+1} := \nabla_w v_j(w^t, s^t)$. Write $w := w^t$, $b := \sum_{j \neq i} d_j^{t+1}$, $\hat{v}_i(\cdot) := v_i(\cdot, s^{t+1})$, and let $x \in \mathbb{R}^m$ be agent i's report.

Step 1: the round utility as a function of the report. By (R1) and (R2), the report x affects only $w^{t+1} = w + \frac{\eta}{n}(x + b)$ and the charge $p_i^{3,t}$; the contributions s^{t+1}, and therefore the cost, are already fixed. In Equation (30) the rebate depends only on the other agents' reports. When the contribution and valuation reports are truthful, as Part 1 ensures, $p_i^{2,t} = 0$ whatever the model w^{t+1} is, by Step 2 of Part 1. Dropping the terms that do not depend on x, agent i's round utility is, up to a constant,

$$U(x) = \hat{v}_i\Big(w + \frac{\eta}{n}(x + b)\Big) - \frac{\eta}{2n}\|x\|_2^2.$$

Step 2: strong concavity. By the chain rule, $\nabla U(x) = \frac{\eta}{n}\nabla\hat{v}_i\big(w + \frac{\eta}{n}(x + b)\big) - \frac{\eta}{n}x$. The first condition of Assumption 8, applied at the fixed profile s^{t+1}, states that $\nabla\hat{v}_i$ is M_i-Lipschitz. Exactly as in Part 3, Step 2, of the proof of Theorem 3, the Cauchy–Schwarz inequality then gives $\langle \nabla U(x) - \nabla U(y), x - y\rangle \leqslant -\mu\|x - y\|_2^2$ with $\mu := \frac{\eta}{n}\big(1 - \frac{\eta M_i}{n}\big)$, which is positive because $\eta M_i < n$. Hence U is μ-strongly concave, and, by the argument leading to Equation (65), $U(x) - U(y) \leqslant \|\nabla U(y)\|_2^2/(2\mu)$ for all x and y.

Step 3: the gradient at the truthful report. At the truthful report $x = d_i^{t+1} = \nabla_w v_i(w^t, s^t)$, the other agents' truthfulness gives $x + b = n\,\tilde{g}(w^t, s^t) = n\,\tilde{g}^t$, so the new model is $w^{t+1} = w^t + \eta\tilde{g}^t$. Substituting into the formula of Step 2,

$$\nabla U(d_i^{t+1}) = \frac{\eta}{n}\Big(\nabla_w v_i(w^{t+1}, s^{t+1}) - \nabla_w v_i(w^t, s^t)\Big).$$

This is where the joint dynamics differ from `T-2P-UPBReD`: between the two gradients both the parameter and the contributions change. Adding and subtracting $\nabla_w v_i(w^t, s^{t+1})$ and applying the triangle inequality, then the two conditions of Assumption 8,

$$\begin{aligned}\|\nabla U(d_i^{t+1})\|_2 &\leqslant \frac{\eta}{n}\Big(\big\|\nabla_w v_i(w^{t+1}, s^{t+1}) - \nabla_w v_i(w^t, s^{t+1})\big\|_2 + \big\|\nabla_w v_i(w^t, s^{t+1}) - \nabla_w v_i(w^t, s^t)\big\|_2\Big)\\ &\leqslant \frac{\eta}{n}\Big(M_i\,\eta\|\tilde{g}^t\|_2 + Q_i\|s^{t+1} - s^t\|_2\Big),\end{aligned}$$

where the first term uses $w^{t+1} - w^t = \eta\tilde{g}^t$.

Step 4: the per-round bound. Squaring the bound of Step 3 and dividing by $2\mu = \frac{2\eta(n - \eta M_i)}{n^2}$, the factors n^2 cancel and one factor of η cancels, so for every report x

$$U(x) - U(d_i^{t+1}) \;\leqslant\; \frac{\eta\big(\eta M_i\|\tilde{g}^t\|_2 + Q_i\|s^{t+1} - s^t\|_2\big)^2}{2(n - \eta M_i)} = \varepsilon_i^t.$$

Step 5: the bound along the truthful path. On the truthful path the iterates are those of `UPBReD` (Part 3), so Equation (33) gives $s^{t+1} - s^t = \gamma g^t$. Bounding each coefficient by the larger of the two,

$$\eta M_i\|\tilde{g}^t\|_2 + \gamma Q_i\|g^t\|_2 \leqslant \max\{\eta M_i, \gamma Q_i\}\big(\|\tilde{g}^t\|_2 + \|g^t\|_2\big) \leqslant \max\{\eta M_i, \gamma Q_i\}\,W^t E,$$

where the second inequality is the geometric decay Equation (53) established in the proof of Theorem 1. Substituting into ε_i^t gives the stated bound proportional to W^{2t}. Since $W < 1$ under the step-size conditions, the geometric series $\sum_{t \geqslant 0} W^{2t}$ equals $\frac{1}{1 - W^2}$, which gives the bound on the sum. As in Theorem 3, this sum bounds one-shot deviations along the truthful path, not deviations in several rounds.

Part 3 (learning). *Truthful reports.* Under truthful reporting, each agent's contribution update in step 2 uses $\hat{s}^t_{-i} = s^t_{-i}$, so it is the update of Algorithm 1, and the model update uses the gradients $\nabla_w v_i(w^t, s^t)$, so it is $w^{t+1} = w^t + \eta\tilde{g}^t$. The iterates of `T-UPBReD` therefore coincide with those of `UPBReD`, and Theorem 1 applies.

Equilibrium gradient reports: existence and uniqueness. Fix a round t, so that w^t and s^{t+1} are fixed, and let $F^t(u) := -\frac{1}{n}\sum_{j\in N} v_j(u, s^{t+1})$. By Step 2 of Part 2, which did not use the truthfulness of the other agents, agent i's round utility is strictly concave in its report for any reports of the others, so its best response is the unique zero of its gradient, namely $x = \nabla_w v_i(w^{t+1}, s^{t+1})$ with $w^{t+1} = w^t + \frac{\eta}{n}(x + \sum_{j\neq i}\hat{d}^{t+1}_j)$. Exactly as in Part 4 of the proof of Theorem 3, a report profile is an equilibrium if and only if w^{t+1} solves $u = w^t - \eta\nabla F^t(u)$, which is the optimality condition of $\min_u\{F^t(u) + \frac{1}{2\eta}\|u - w^t\|_2^2\}$. Condition 1 of Assumption 1 states that $\frac{1}{n}\sum_j v_j(\cdot, s)$ is $\tilde{\lambda}$-strongly concave for every s, so F^t is $\tilde{\lambda}$-strongly convex, the proximal objective has a unique minimizer, and the round has a unique equilibrium.

Equilibrium gradient reports: convergence under separable valuations. Suppose now that the valuations are separable, $v_j(w, s) = \varphi_j(w) + \sum_k \psi_{jk}(s_k)$. Two simplifications follow. First, $F^t(u) = -\frac{1}{n}\sum_j \varphi_j(u)$ up to a constant, so the proximal step minimizes the same function $F := -\frac{1}{n}\sum_j \varphi_j$ in every round, whose unique minimizer w^* is the maximizer of $\frac{1}{n}\sum_j v_j(\cdot, s)$ for every s. The rate argument of Part 4 of the proof of Theorem 3, with the strong-convexity constant $\tilde{\lambda}$ of F in place of ν, gives $\|w^{t+1} - w^*\|_2 \leqslant (1 + \eta\tilde{\lambda})^{-1}\|w^t - w^*\|_2$, and iterating gives the stated rate. Second, since ψ_{ik} depends only on s_k and φ_i does not depend on s, the marginal utility $\frac{\partial}{\partial s_i}u_i(w, s) = \psi'_{ii}(s_i) - c'_i(s_i)$ depends neither on the model nor on the other agents' contributions. With truthful contribution reports (Part 1), the contribution dynamics are therefore exactly those of `UPBReD`, whatever the gradient reports. Under separability every mixed derivative $\frac{\partial^2}{\partial w_k \partial s_i}u_i$ vanishes, so $H \equiv 0$, $P = 0$ and $b = 0$, and the first line of Equation (47) reads $\|g^{t+1}\|_2 \leqslant \sqrt{a}\,\|g^t\|_2$, whatever the model updates are. Iterating this inequality gives $\|g^t\|_2 \leqslant (\sqrt{a})^t\|g^0\|_2$, where $\sqrt{a} \leqslant W_1 < 1$ under the step-size conditions. Since $s^{t+1} - s^t = \gamma g^t$ by Equation (33), the increments are absolutely summable, so $(s^t)_t$ is a Cauchy sequence in $\mathbb{R}^n$. Because $\mathbb{R}^n$ is complete, it converges to some s^∞, and $s^\infty \in S$ because S is closed. Under separability the gradient mapping does not depend on w, and it is continuous in s as a composition of Π_S and $\nabla_s u$; passing to the limit therefore gives $g(s^\infty) = 0$. Finally, Assumption 1 makes each u_i concave in s_i (Observation 2), so Lemma 3 applies and shows that s^∞ is a PSNE. It is the same PSNE as under truthful reporting, because the contribution dynamics are the same.

Part 4 (budget). In every round, summing Equation (30) over i and exchanging the order of summation, exactly as in Part 5 of the proof of Theorem 3, gives $\sum_i p_i^{3,t} = 0$ for every report profile. Both terms of Equation (29) are non-positive because $K > 0$ and $\chi_i^t \geqslant 0$, so $\sum_i p_i^{2,t} \leqslant 0$ for every report profile, and both vanish on the truthful path by Step 2 of Part 1. The total transfer in every round is therefore zero on the truthful path and non-positive in general, so the center never runs a deficit. □

B Proofs for Section 5

B.1 Proof of Lemma 1

Proof. Using Equation (12), we get the derivative of the utility function for $s_i \in (s_i^{\min}, s_i^{\max})$ (note that this holds only for the interior of S_i); at the interior, the updates are given by $[g(w, s, \mu)]_i = \frac{\partial}{\partial s_i}u_i(w, s) = \frac{\partial}{\partial s_i}v_i(w, s) - c'_i(s_i) + \beta$, where $c'_i(s_i) = \frac{d}{ds_i}c_i(s_i)$. Using Assumption 2 and choosing in Equation (12) the parameter $\beta > \zeta + \tau$, we have

$$[g(w, s, \mu)]_i = \frac{\partial}{\partial s_i}v_i(w, s) - c'_i(s_i) + \beta \geqslant -\tau - \zeta + \beta > 0.$$

The last inequality holds since $\beta > \zeta + \tau$, $\forall i \in N$. □

B.2 Proof of Theorem 2

Proof. **Phase 1.** Fix $i \in N$ and write $\Delta := \beta - \zeta - \tau > 0$ and $l_i := s_i^{\max} - s_i^0 \geq 0$. Throughout Phase 1 the model is fixed at w^0.

Step 1: each update increases s_i by at least $\gamma\Delta$, up to the cap. Let $s_i' := s_i^k + \gamma \frac{\partial}{\partial s_i} u_i(w^0, s^k)$ be the unconstrained step. By Lemma 1, $\frac{\partial}{\partial s_i} u_i(w^0, s^k) \geq \Delta$, so

$$s_i' \;\geq\; s_i^k + \gamma\Delta \;>\; s_i^k \;\geq\; s_i^{\min}. \tag{59}$$

Hence s_i' never falls below $s_i^{\min}$, and the projection can only clip from above:

$$s_i^{k+1} = \Pi_{S_i}(s_i') = \min\{s_i^{\max},\, s_i'\} \;\geq\; \min\{s_i^{\max},\, s_i^k + \gamma\Delta\}. \tag{60}$$

Equivalently, $[g(w^0, s^k, \mu^k)]_i = \big(s_i^{k+1} - s_i^k\big)/\gamma \geq 0$, and this is at least Δ whenever the step is not clipped.

Step 2: a lower bound on s_i^k. We claim that for all $k \geq 0$,

$$s_i^k \;\geq\; \min\{s_i^{\max},\, s_i^0 + \gamma\Delta k\}. \tag{61}$$

For $k = 0$ this holds because $s_i^0 \leq s_i^{\max}$. Suppose it holds for k. Since $x \mapsto \min\{s_i^{\max}, x + \gamma\Delta\}$ is nondecreasing, (60) gives

$$\begin{aligned} s_i^{k+1} &\geq \min\left\{s_i^{\max},\, \min\{s_i^{\max},\, s_i^0 + \gamma\Delta k\} + \gamma\Delta\right\} \\ &= \min\left\{s_i^{\max},\, s_i^{\max} + \gamma\Delta,\, s_i^0 + \gamma\Delta(k+1)\right\} = \min\{s_i^{\max},\, s_i^0 + \gamma\Delta(k+1)\}, \end{aligned}$$

which proves (61).

Step 3: reaching $s^{\max}$. Let

$$\kappa := \max_{i \in N} \left\lceil \frac{l_i}{\gamma\Delta} \right\rceil.$$

For $k \geq \kappa$ we have $s_i^0 + \gamma\Delta k \geq s_i^0 + l_i = s_i^{\max}$, so (61) gives $s_i^k \geq s_i^{\max}$. Since $s_i^k \in S_i$ implies $s_i^k \leq s_i^{\max}$, we conclude $s_i^k = s_i^{\max}$ for all $i \in N$ and all $k \geq \kappa$. In particular, $s^\kappa = s^{\max}$.

Step 4: $s^{\max}$ is a fixed point. At $s = s^{\max}$, Lemma 1 gives $\frac{\partial}{\partial s_i} u_i > 0$, so $s_i' = s_i^{\max} + \gamma \frac{\partial}{\partial s_i} u_i > s_i^{\max}$. Then $\mu_i = (s_i^{\max} - s_i')/\gamma = -\frac{\partial}{\partial s_i} u_i$, so $[g(w^0, s^{\max}, \mu)]_i = \frac{\partial}{\partial s_i} u_i + \mu_i = 0$ and $\Pi_{S_i}(s_i') = s_i^{\max}$. Thus $s^{\max}$ is a fixed point of $s^{k+1} = s^k + \gamma\, g(w^0, s^k, \mu^k)$, and the iterates remain at $s^{\max}$ from round κ onward.

Step 5: $s^{\max}$ is a Nash equilibrium at every parameter. Step 4 shows that $s^{\max}$ is a fixed point of the projected update. A fixed point is a Nash equilibrium only if every agent is best responding, and here we cannot invoke Lemma 3, since Theorem 2 does not assume the concavity of Assumption 1. We use instead the monotonicity given by Lemma 1. Fix $w \in \mathbb{R}^m$, $i \in N$ and s_{-i}. By Lemma 1, the function $x \mapsto u_i(w, x, s_{-i})$ has a strictly positive derivative on S_i. For any $x \in S_i$ with $x < s_i^{\max}$, the mean value theorem gives a point $\xi \in (x, s_i^{\max})$ with

$$u_i(w, s_i^{\max}, s_{-i}) - u_i(w, x, s_{-i}) = \frac{\partial}{\partial s_i} u_i(w, \xi, s_{-i}) \left(s_i^{\max} - x\right) > 0.$$

Hence $s_i^{\max}$ is agent i's unique best response to every s_{-i} at every parameter w, and $s^{\max}$ satisfies Definition 1 at every w, in particular at $w^{\texttt{OPT}}$.

Phase 2: For $t > \kappa$, we can focus on the second phase of the algorithm. From the fixed point property of $s^{\max}$, we can now focus entirely on the function $f(w, s^{\max}) = -\frac{1}{n}\sum_{i \in N} v_i(w, s^{\max})$. From the algorithmic description, the second phase is just a simple gradient descent, run by the center, on $f(w, s^{\max})$. This is easy to see since at every iteration the center gets d_i^{t+1} from all the agents $i \in N$, aggregates them and constructs $\tilde{g}(w^t, s^{\max})$, which is the gradient of $f(w, s^{\max})$ computed at w^t. We denote $f(w, s^{\max})$ with $f(w)$ in this part for notational cleanliness.

Note that the gradient descent is run with initialization w^κ (which is the same as w^0 as given by the first phase of the algorithm). We exploit the strong convexity and smoothness of $f(w, s^{\max})$. Running the second phase for T_0 iterations and using the convergence result of Wright and Recht [52], we obtain

$$f(w^{T_0}) - f(w^{\texttt{OPT}}) \leqslant (1 - \tfrac{\nu}{M})^{T_0} [f(w^0) - f(w^{\texttt{OPT}})].$$

Hence, for $f(w^{T_0}) - f(w^{\texttt{OPT}}) < \epsilon$, we require

$$T_0 > \left(\ln \frac{f(w^0) - f(w^{\texttt{OPT}})}{\epsilon}\right) \Big/ \left(\ln \frac{1}{\left(1 - \frac{\nu}{M}\right)}\right),$$

which gives the bound on the welfare gap stated in the theorem.

It remains to show that the iterates themselves converge. Since f is ν-strongly convex (Assumption 3), it has a unique minimizer $w^{\mathtt{OPT}}$, at which $\nabla f(w^{\mathtt{OPT}}) = 0$. Applying the definition of strong convexity (Definition 6) at $w^{\mathtt{OPT}}$ gives $f(w) - f(w^{\mathtt{OPT}}) \geqslant \frac{\nu}{2}\|w - w^{\mathtt{OPT}}\|_2^2$ for every w, and therefore

$$\|w^{T_0} - w^{\mathtt{OPT}}\|_2 \leqslant \sqrt{2\big(f(w^{T_0}) - f(w^{\mathtt{OPT}})\big)/\nu},$$

which tends to zero as T_0 grows. Combined with Step 5, and with the fact that $w^{\mathtt{OPT}}$ maximizes $\sum_{i\in N} v_i(\cdot, s^{\max})$ by Equation (2), the limit pair $(w^{\mathtt{OPT}}, s^{\max})$ is the fixed point described after Definition 1: $s^{\max}$ is a PSNE at $w^{\mathtt{OPT}}$, and $w^{\mathtt{OPT}}$ maximizes the welfare at $s^{\max}$. This proves the theorem. □

B.3 Proof of Corollary 1

Proof. The proof follows along the same lines as that of Theorem 2. Using the choice of β, we first ensure that with κ steps, we obtain $s_i^{\kappa} = s_i^{\max}$ for all $i \in N$. Now, the framework is the same as minimization of a strongly convex and smooth function $f(w)$ with initialization $w^{\kappa} = w^0$. Using the result of Wright and Recht [52] with $\eta = \frac{2}{M+\nu}$, we obtain the iterate convergence, namely $\|w^{\tilde{T}_0} - w^{\mathtt{OPT}}\|_2 \leqslant \left(\frac{1-\frac{\nu}{M}}{1+\frac{\nu}{M}}\right)^{\tilde{T}_0} \|w^0 - w^{\mathtt{OPT}}\|_2$. Taking logarithms of both sides implies the result. □

B.4 Proofs of the individual rationality results

Proof of Proposition 1. Part 1. We first compute the transfer that agent i receives at the full-contribution profile. Substituting $s = s^{\max}$ into Equation (12), and using that the other agents' capacities sum to $\sum_{j\neq i} s_j^{\max} = n\bar{s}^{\max} - s_i^{\max}$ by the definition Equation (17), we obtain

$$p_i^1(s^{\max}) = \beta\Big(s_i^{\max} - \frac{n\bar{s}^{\max} - s_i^{\max}}{n-1}\Big) = \frac{\beta\big((n-1)s_i^{\max} - n\bar{s}^{\max} + s_i^{\max}\big)}{n-1} = \frac{n\beta}{n-1}\big(s_i^{\max} - \bar{s}^{\max}\big),$$

where the second equality brings both terms over the common denominator $n-1$ and the third collects the two terms in $s_i^{\max}$. Subtracting the individual learning outcome $\bar{u}_i$ from the utility Equation (15) and grouping the first two terms and $-\bar{u}_i$ into π_i, as defined in Equation (16), gives Equation (19).

Part 2. By Definition 7, the limit outcome is individually rational if and only if the right-hand side of Equation (19) is non-negative for every agent. We examine this condition separately for the three groups of agents, since the sign of $s_i^{\max} - \bar{s}^{\max}$ determines how the condition depends on β. For $i \in N_0$, the transfer vanishes, so the condition reads $\pi_i \geqslant 0$ and does not involve β. For $i \in N_-$, the factor $s_i^{\max} - \bar{s}^{\max}$ is negative, so dividing the condition by $\frac{n}{n-1}(\bar{s}^{\max} - s_i^{\max}) > 0$, which preserves the inequality, shows that it is equivalent to $\beta \leqslant (n-1)\pi_i \big/ \big(n(\bar{s}^{\max} - s_i^{\max})\big)$. For $i \in N_+$, the factor is positive, and dividing by $\frac{n}{n-1}(s_i^{\max} - \bar{s}^{\max}) > 0$ shows that the condition is equivalent to $\beta \geqslant -(n-1)\pi_i \big/ \big(n(s_i^{\max} - \bar{s}^{\max})\big)$. Taking the most restrictive of the upper bounds and of the lower bounds, and recalling that $\beta > 0$ in Equation (12), all conditions hold simultaneously exactly when $\pi_i \geqslant 0$ on N_0 and $\beta_{\text{low}} \leqslant \beta \leqslant \beta_{\text{high}}$, with the thresholds defined in Equation (18).

Part 3. Theorem 2 requires $\beta > \zeta + \tau$, and by Part 2 individual rationality requires $\beta \in [\beta_{\text{low}}, \beta_{\text{high}}]$, together with $\pi_i \geqslant 0$ on N_0. A β meeting both requirements exists if and only if the interval $[\beta_{\text{low}}, \beta_{\text{high}}]$ is non-empty and contains a point strictly larger than $\zeta + \tau$, which happens exactly when $\beta_{\text{low}} \leqslant \beta_{\text{high}}$ and $\zeta + \tau < \beta_{\text{high}}$. If $\pi_i \geqslant 0$ for every agent, then every lower bound in Equation (18) is non-positive, so $\beta_{\text{low}} = 0$, the conditions on N_0 hold automatically, and $\beta_{\text{high}} \geqslant 0 = \beta_{\text{low}}$; the only remaining condition is $\zeta + \tau < \beta_{\text{high}}$. □

Proof of Corollary 2. We show that U_i^T converges to U_i for each agent, and then use the strict inequality $U_i > \bar{u}_i$.

First, contributions and transfers stop changing after Phase 1. By Theorem 2, $s^T = s^{\max}$ for all $T \geqslant \kappa$, so the cost term $c_i(s_i^T)$ and the transfer $p_i^1(s^T)$ in U_i^T coincide with the corresponding terms of U_i from round κ onwards.

Second, the parameter converges to $w^{\mathtt{OPT}}$. Write $f(w) := f(w, s^{\max})$. By Theorem 2, $f(w^T) - f(w^{\mathtt{OPT}})$ tends to 0 as T grows. Since f is ν-strongly convex by Assumption 3 and $w^{\mathtt{OPT}}$ minimizes f, so that

$\nabla f(w^{\mathtt{OPT}}) = 0$, the definition of strong convexity (Definition 6) applied at the point $w^{\mathtt{OPT}}$ gives

$$f(w^T) - f(w^{\mathtt{OPT}}) \;\geqslant\; \langle \nabla f(w^{\mathtt{OPT}}), w^T - w^{\mathtt{OPT}} \rangle + \frac{\nu}{2}\|w^T - w^{\mathtt{OPT}}\|_2^2 \;=\; \frac{\nu}{2}\|w^T - w^{\mathtt{OPT}}\|_2^2.$$

Rearranging, $\|w^T - w^{\mathtt{OPT}}\|_2 \leqslant \sqrt{2(f(w^T) - f(w^{\mathtt{OPT}}))/\nu}$, and the right-hand side tends to 0; hence $w^T \to w^{\mathtt{OPT}}$.

Third, by the continuity of $v_i(\cdot, s^{\max})$ at $w^{\mathtt{OPT}}$, it follows that $v_i(w^T, s^{\max}) \to v_i(w^{\mathtt{OPT}}, s^{\max})$. Combining the three observations, $U_i^T \to U_i$ for every $i \in N$.

Finally, set $\delta := \min_{i \in N}(U_i - \bar{u}_i)$, which is strictly positive by hypothesis because N is finite. By the convergence just established, there is a $\bar{T} \geqslant \kappa$ such that $|U_i^T - U_i| < \delta$ for every $i \in N$ and every $T \geqslant \bar{T}$. For such T we obtain $U_i^T > U_i - \delta \geqslant \bar{u}_i$, which is individual rationality after T rounds. □

Proof of Corollary 3. Part 1. Summing Equation (20) over agents, the term $\sum_i p_i^1(s)$ vanishes because p^1 is budget balanced, and $\sum_i \rho_i = 0$ by construction; hence $\sum_i p_i^\rho(s) = 0$ for every s. Since each ρ_i is a constant, it does not depend on s_i, so $\frac{\partial}{\partial s_i} p_i^\rho = \frac{\partial}{\partial s_i} p_i^1 = \beta$. The proof of Lemma 1 uses the payment only through this derivative, so the lemma still holds. The proof of Theorem 2 in turn uses the payment only through Lemma 1 in Phase 1, and Phase 2 involves no payments, so all of its conclusions remain valid.

Part 2. With rebates, agent i's gain at the limit outcome becomes, by the same computation as in Part 1 of Proposition 1,

$$\pi_i + p_i^1(s^{\max}) - \rho_i.$$

Suppose first that some choice of rebates makes the limit outcome individually rational. Summing these gains over all agents, using $\sum_i p_i^1(s^{\max}) = 0$ by budget balance and $\sum_i \rho_i = 0$ by construction, gives $\sum_i \pi_i$, and this sum is non-negative because every summand is. Conversely, suppose $\sum_i \pi_i \geqslant 0$ and choose

$$\rho_i := \pi_i + p_i^1(s^{\max}) - \frac{1}{n}\sum_{j \in N} \pi_j.$$

These rebates sum to zero, again because $\sum_i p_i^1(s^{\max}) = 0$, so they satisfy Equation (20); and substituting them shows that every agent's gain equals $\frac{1}{n}\sum_j \pi_j \geqslant 0$, so the limit outcome is individually rational. This choice splits the total gain from collaboration equally among the agents.

Part 3. If the capacities are known before Phase 1, the constants $\rho_i := p_i^1(s^{\max})$ can be fixed in advance, and they sum to zero by the budget balance of p^1, so they satisfy Equation (20). At the limit outcome each agent's transfer is $p_i^1(s^{\max}) - \rho_i = 0$, so by Part 1 of Proposition 1 its gain is exactly π_i, and individual rationality holds if and only if $\pi_i \geqslant 0$ for every agent. □

B.5 Proofs of the no-evidence results

Proof of Proposition 2. Fix a transfer rule p, a contribution $s_i \in S_i$ and a reporting strategy $m_i \in M_i$. Condition (E1) guarantees that every such pair is available to agent i, so agent i chooses from the product set $S_i \times M_i$.

Step 1: decomposing the utility. By the separability of the valuations (Definition 8), applied at $\bar{w} = W(R)$ and at the contribution profile s actually used, the valuation in Equation (21) splits as

$$v_i\big(W(R), s\big) = \varphi_i\big(W(R)\big) + \psi_{ii}(s_i) + \sum_{j \neq i} \psi_{ij}(s_j),$$

where we have separated the term $j = i$ from the sum. Substituting into Equation (21) and grouping the terms that involve agent i's own contribution, the realized utility is

$$u_i = \underbrace{\psi_{ii}(s_i) - c_i(s_i)}_{=A_i(s_i)} + \underbrace{\varphi_i\big(W(R)\big) + \textstyle\sum_{j \neq i} \psi_{ij}(s_j) + p_i(R)}_{=:\, Z_i}. \tag{62}$$

The first group is deterministic once s_i is chosen. The second group, Z_i, is a function of the pair (R, s_{-i}) alone: the parameter $W(R)$ and the transfer $p_i(R)$ are computed from the record, and the remaining terms involve only the other agents' contributions.

Step 2: the second group does not depend on the contribution (Part 1). Taking expectations in Equation (62) under the pair (s_i, m_i), linearity of expectation gives $\mathbb{E}[u_i] = A_i(s_i) + \mathbb{E}[Z_i]$. Since Z_i is

a function of (R, s_{-i}), its expectation is determined by the joint distribution of (R, s_{-i}). This is where condition (E2) is used: for the fixed reporting strategy m_i, that joint distribution is the same for every s_i, so $\mathbb{E}[Z_i]$ depends on m_i but not on s_i, and it equals $\Pi_i(m_i)$ as defined in Equation (22). The same argument, applied to the function $p_i(R)$ alone, shows that the expected transfer $\mathbb{E}[p_i(R)]$ does not depend on s_i. This proves Part 1.

Step 3: separate maximization (Part 2). By Steps 1 and 2, agent i's expected utility is $A_i(s_i)+\Pi_i(m_i)$, the sum of a function of s_i alone and a function of m_i alone, maximized over the product set $S_i \times M_i$. For such a sum, $\sup_{(s_i,m_i)} \left[A_i(s_i) + \Pi_i(m_i)\right] = \sup_{s_i} A_i(s_i) + \sup_{m_i} \Pi_i(m_i)$, because the two variables can be chosen independently. Consequently, a pair attains the supremum if and only if each component attains its own supremum: if s_i maximizes A_i and m_i maximizes Π_i, the sum equals the sum of the two suprema; and if either component falls short of its supremum by some $\delta > 0$, the pair falls short of the joint supremum by at least δ. This proves the characterization of best responses. Finally, $A_i(s_i) = \psi_{ii}(s_i) - c_i(s_i)$ involves neither the transfer rule nor the reports, so the set of maximizers of A_i, which is the set of agent i's best contributions, is the same for every transfer rule; taking $p \equiv 0$ identifies it with the set of best contributions without transfers. This proves Part 2.

Step 4: full contribution (Part 3). By Part 2, $s_i^{\max}$ is part of a best response under a given transfer rule if and only if $s_i^{\max}$ maximizes A_i. Since this condition does not involve the transfer rule, it holds for one rule if and only if it holds for every rule. This proves Part 3.

Step 5: approximate truthfulness (Part 4). Write $\bar{\Pi} := \sup_{m \in M_i} \Pi_i(m)$, which is finite by assumption. The hypothesis states that $\bar{\Pi} - \varepsilon \leqslant \Pi_i\big(m_i^T(s_i)\big)$ for every s_i, and by the definition of the supremum $\Pi_i\big(m_i^T(s_i)\big) \leqslant \bar{\Pi}$. Hence every value $\Pi_i\big(m_i^T(x)\big)$, $x \in S_i$, lies in the interval $[\bar{\Pi} - \varepsilon, \bar{\Pi}]$ of length ε, and any two numbers in an interval of length ε differ by at most ε. Note that Part 1 is what makes this hypothesis meaningful: since Π_i does not depend on the contribution, the same function Π_i is compared across all contributions. This proves Part 4. □

Proof of Corollary 4. Part 1. With the valuations of Section 3 and linear costs, $A_i(x) = -\frac{1}{n}x^{-\alpha_i} - c_i x$. Its derivative $A_i'(x) = \frac{\alpha_i}{n}x^{-\alpha_i-1} - c_i$ is strictly decreasing in x, so A_i is strictly concave and its unconstrained maximizer is the unique root of $A_i'(x) = 0$, namely $x^* = (\alpha_i/(nc_i))^{1/(\alpha_i+1)}$. For a strictly concave function of one variable, the maximizer over the interval S_i is the projection of the unconstrained maximizer onto S_i, which is s_i^{nt}. By Part 2 of Proposition 2, this is agent i's contribution under every transfer rule, and in particular under p^1 for every β.

Part 2. We compare two reporting strategies $m_i^T(x)$ and $m_i^T(y)$ that report gradients truthfully and differ only in the contribution reported, x versus y. Three observations show that only the transfer changes. First, the final parameter is computed from the gradient reports with uniform weights, and by separability the gradient reports $\nabla\varphi_i(w^t)$ do not depend on the reported contribution, so $W(R)$ is the same under both strategies. Second, the other agents' contributions are the same: by separability, agent j's marginal utility in Phase 1, $\psi_{jj}'(s_j) - c_j'(s_j) + \beta$, does not involve agent i's reported contribution, so agent j's updates are unaffected by it. Third, under p^1 the transfer differs only through agent i's own report, and p_i^1 is linear in that report with slope β, by Equation (12). Therefore

$$\Pi_i\big(m_i^T(x)\big) - \Pi_i\big(m_i^T(y)\big) = \beta\,(x - y) \qquad \text{for all } x, y \in S_i.$$

Now let agent i's true contribution be $s_i^{\min}$. By Part 2 of Proposition 2, its expected utility is $A_i(s_i^{\min}) + \Pi_i(m_i)$, so reporting $s_i^{\max}$ instead of the truthful $s_i^{\min}$ raises its expected utility by $\Pi_i\big(m_i^T(s_i^{\max})\big) - \Pi_i\big(m_i^T(s_i^{\min})\big) = \beta(s_i^{\max} - s_i^{\min})$. Truthful reporting is therefore ε-optimal only if $\varepsilon \geqslant \beta(s_i^{\max} - s_i^{\min})$, and $\beta > \zeta + \tau$ gives the final inequality. □

B.6 Proof of Theorem 3

Notation and timing. Throughout, $\bar{w} := w^{T_0}$ denotes the final model, p^1, p^2, p^3 are the transfers of Equations (23) to (25), and $s = (s_i)_{i \in N}$ denotes the contributions *actually used* in training, as opposed to the reported contributions $\hat{s}$. In the elicitation stage, *truthful reporting* by agent i means $\hat{s}_i = s_i$ and $\hat{v}_i = v_i(\bar{w}, s)$, its realized valuation. We write $v_i(\cdot) := v_i(\cdot, s^{\max})$ whenever the contributions are fixed at $s^{\max}$, and $f(w) := -\frac{1}{n}\sum_{i \in N} v_i(w)$, as in the statement.

The proof repeatedly uses the order in which Algorithm 3 collects reports, so we record three consequences of it.

(T1) Contribution reports sent during Phase 1 are used only in the other agents' Phase 1 updates; they enter no transfer and do not affect the model.
(T2) Gradient reports sent in round t of Phase 2 affect only the next model w^{t+1} and the round-t charges $p^3(\hat{d}^t)$.
(T3) The elicitation reports $(\hat{s}, \hat{v})$ are sent after $\bar{w}$ has been computed and after all Phase 2 charges have been paid; they affect only p^1 and p^2.

We now prove Theorem 3, taking its five parts in turn.

Part 1 (truthful contribution and valuation reports). Fix an agent i and an arbitrary history of Phase 1 and Phase 2 reports, so that the final model $\bar{w}$ and the actual contributions s are fixed. Suppose the other agents report truthfully in the elicitation stage, that is, $\hat{s}_{-i} = s_{-i}$ and $\hat{v}_j = v_j(\bar{w}, s)$ for every $j \neq i$; that the latter is possible is the first part of Assumption 5, which states that each agent observes its realized valuation at the contributions actually used. We compare agent i's utility under an arbitrary report $(\hat{s}_i, \hat{v}_i)$, with $\hat{s}_i \in S_i$, with its utility under truthful reporting.

Step 1: only p^1 and p^2 change. By (T3), agent i's elicitation report does not affect $\bar{w}$, and therefore does not affect its valuation $v_i(\bar{w}, s)$, which depends only on $\bar{w}$ and the actual contributions. It does not affect its cost $c_i(s_i)$, which depends only on its actual contribution, nor any Phase 2 charge, which has already been paid. Consequently the difference between agent i's utility under the report $(\hat{s}_i, \hat{v}_i)$ and under truthful reporting equals the difference between the corresponding values of $p_i^1 + p_i^2$.

Step 2: the difference in p^1. The other agents' reports $\hat{s}_{-i} = s_{-i}$ are the same in both cases, so in Equation (23) only the first term changes, and

$$p_i^1(\hat{s}_i, s_{-i}) - p_i^1(s_i, s_{-i}) = \beta\,(\hat{s}_i - s_i).$$

Step 3: the value of p^2 under truthful reporting. Under truthful reporting the reported profile is s, so the center evaluates $v_j(\bar{w}, s)$ for every j. Agent i's own report $v_i(\bar{w}, s)$ coincides with this evaluation, so the quadratic term of Equation (24) vanishes; and every other agent's report $v_j(\bar{w}, s)$ coincides with it as well, so $\chi_i = 0$. Hence $p_i^2 = 0$ under truthful reporting.

Combining Steps 1–3, and writing $\hat{s} := (\hat{s}_i, s_{-i})$ for the reported profile, the utility difference is

$$u_i - u_i^{\text{truth}} = \beta\,(\hat{s}_i - s_i) - \big(\hat{v}_i - v_i(\bar{w}, \hat{s})\big)^2 - \chi_i\,\big(\beta \hat{s}_i + K\big). \tag{63}$$

We now distinguish two cases according to the contribution report.

Case 1: $\hat{s}_i = s_i$. Then $\hat{s} = s$, and each other agent's report $\hat{v}_j = v_j(\bar{w}, s)$ equals the center's evaluation $v_j(\bar{w}, \hat{s})$, so $\chi_i = 0$. The first term of Equation (63) also vanishes, and the difference reduces to $-\big(\hat{v}_i - v_i(\bar{w}, s)\big)^2 \leqslant 0$, with equality if and only if $\hat{v}_i = v_i(\bar{w}, s)$, that is, if and only if agent i reports its true valuation.

Case 2: $\hat{s}_i \neq s_i$. Here we use Assumption 4, applied at the parameter $\bar{w}$ with the other agents' contributions s_{-i} and the pair $\hat{s}_i \neq s_i$: there is an agent $j \neq i$ with $v_j(\bar{w}, \hat{s}) \neq v_j(\bar{w}, s)$. Since agent j reports $\hat{v}_j = v_j(\bar{w}, s)$, its report differs from the center's evaluation $v_j(\bar{w}, \hat{s})$, and by the definition of χ_i in Equation (24), $\chi_i = 1$. This conclusion does not involve agent i's valuation report, because χ_i depends only on the other agents' reports and on the center's evaluations at $\hat{s}$. Substituting $\chi_i = 1$ into Equation (63), the terms $\beta \hat{s}_i$ cancel, and

$$u_i - u_i^{\text{truth}} = -\beta s_i - K - \big(\hat{v}_i - v_i(\bar{w}, \hat{s})\big)^2 \leqslant -\beta s_i - K < 0,$$

where the first inequality drops a non-positive term and the second uses $\beta > 0$, $s_i \geqslant 0$ (since $S_i \subseteq [0, \infty)$) and $K > 0$.

In both cases truthful reporting is optimal, and every report with $\hat{s}_i \neq s_i$ is worse by at least $\beta s_i + K$. Since $\bar{w}$ and the history were arbitrary, the conclusion holds for every history of gradient reports. Finally, the profile s in this argument is the profile actually used in training. An agent that trains on s_i samples but reports $\hat{s}_i \neq s_i$, in particular one that trains on fewer samples than it reports, is therefore in Case 2 and is penalized.

Part 2 (full contribution). *Step 1: the payoff that drives contributions.* By Part 1, at the elicitation stage agent i reports its actual contribution, receives p_i^1 evaluated at the actual contributions, and pays

$p_i^2 = 0$. The contribution-phase utility to which the Phase 1 dynamics of Algorithm 3 apply is therefore

$$u_i(w, s) = v_i(w, s) - c_i(s_i) + p_i^1(s).$$

This utility contains no term $\sum_{j \neq i} v_j$, so it coincides with the contribution-phase utility of `2P-UPBReD`.

Step 2: a uniform lower bound on the marginal utility. Differentiating with respect to s_i, the three terms contribute $\frac{\partial}{\partial s_i} v_i(w, s)$, $-c_i'(s_i)$, and, by Equation (23), exactly β. By Assumption 2, the first two are bounded below by $-\tau$ and $-\zeta$ respectively; this is the computation of Lemma 1. Hence

$$\frac{\partial}{\partial s_i} u_i(w, s) \;\geqslant\; \beta - \tau - \zeta \;=\; \Delta \;>\; 0 \qquad \text{for every } w \in \mathbb{R}^m \text{ and every } s \in \textstyle\prod_j S_j,$$

where $\Delta > 0$ because $\beta > \zeta + \tau$ by hypothesis.

Step 3: dominance. Fix w and s_{-i}. By Step 2, the function $x \mapsto u_i(w, x, s_{-i})$ has a strictly positive derivative on S_i, so by the mean value theorem it is strictly increasing on S_i and attains its maximum only at $s_i^{\max}$. Since w and s_{-i} were arbitrary, contributing $s_i^{\max}$ maximizes u_i whatever the other agents do, which is the first claim of Part 2.

Step 4: Phase 1 reaches $s^{\max}$. Each Phase 1 update is a projected step $s_i^{k+1} = \Pi_{S_i}\big(s_i^k + \gamma\, G_i^k\big)$, where G_i^k is agent i's marginal utility evaluated at the broadcast information $(w^0, \hat{s}^k)$. By Step 2, $G_i^k \geqslant \Delta$, whatever the broadcast reports are, because the lower bound holds at every (w, s). Hence $s_i^k + \gamma G_i^k \geqslant s_i^k + \gamma\Delta > s_i^k \geqslant s_i^{\min}$, and projecting a point above $s_i^{\min}$ onto $[s_i^{\min}, s_i^{\max}]$ only caps it at $s_i^{\max}$. Since the projection is monotone,

$$s_i^{k+1} \;\geqslant\; \min\big\{s_i^{\max},\, s_i^k + \gamma\Delta\big\}.$$

By induction on k, $s_i^k \geqslant \min\{s_i^{\max},\, s_i^0 + k\gamma\Delta\}$, so $s_i^k = s_i^{\max}$ as soon as $k\gamma\Delta \geqslant s_i^{\max} - s_i^0$, that is, for all $k \geqslant \lceil (s_i^{\max} - s_i^0)/(\gamma\Delta) \rceil$. Once reached, $s_i^{\max}$ is kept, because the step from $s_i^{\max}$ is positive and the projection returns $s_i^{\max}$. Taking the maximum over $i \in N$ gives Equation (28).

Step 5: misreports in Phase 1 are harmless. Step 4 used only the lower bound $G_i^k \geqslant \Delta$, which holds at every broadcast profile. Hence, consistently with (T1), contribution reports sent during Phase 1, truthful or not, change neither any agent's best response nor the number of steps needed to reach $s^{\max}$.

Part 3 (approximately truthful gradients). Fix a round t of Phase 2 and an agent i, and suppose the other agents report their true gradients $d_j^t := \nabla v_j(w^t)$. Write $w := w^t$ and $b := \sum_{j \neq i} d_j^t$, and let $x \in \mathbb{R}^m$ be agent i's report.

Step 1: agent i's round utility as a function of its report. By (T2), agent i's report affects only the next model $w^{t+1} = w + \frac{\eta}{n}(x + b)$ and the round-t charge. Agents are myopic in Phase 2, so agent i's round payoff is its valuation of w^{t+1} plus $p_i^3(\hat{d}^t)$. In Equation (25), the rebate term $\frac{\eta}{2n(n-1)} \sum_{j \neq i} \|d_j^t\|_2^2$ depends only on the other agents' reports, and the cost $c_i(s_i^{\max})$ depends only on the contribution, which was fixed in Phase 1. Dropping these terms, which do not depend on x, changes neither the maximizer nor any utility difference, and leaves

$$U(x) = v_i\Big(w + \frac{\eta}{n}(x + b)\Big) - \frac{\eta}{2n}\|x\|_2^2.$$

Step 2: U is μ-strongly concave. By the chain rule, since the inner map $x \mapsto w + \frac{\eta}{n}(x + b)$ has Jacobian $\frac{\eta}{n}\mathbb{I}$,

$$\nabla U(x) = \frac{\eta}{n}\, \nabla v_i\Big(w + \frac{\eta}{n}(x + b)\Big) - \frac{\eta}{n}\, x. \tag{64}$$

For any x, y, write $z_x := w + \frac{\eta}{n}(x + b)$ and $z_y := w + \frac{\eta}{n}(y + b)$, so that $z_x - z_y = \frac{\eta}{n}(x - y)$. Then

$$\big\langle \nabla U(x) - \nabla U(y),\, x - y \big\rangle = \frac{\eta}{n}\big\langle \nabla v_i(z_x) - \nabla v_i(z_y),\, x - y \big\rangle - \frac{\eta}{n}\|x - y\|_2^2 \leqslant \Big(\frac{\eta^2 M_i}{n^2} - \frac{\eta}{n}\Big)\|x - y\|_2^2 = -\mu\, \|x - y\|_2^2,$$

where the inequality applies the Cauchy–Schwarz inequality and the M_i-Lipschitz continuity of ∇v_i from Assumption 6, which give $\langle \nabla v_i(z_x) - \nabla v_i(z_y), x - y \rangle \leqslant M_i \|z_x - z_y\|_2 \|x - y\|_2 = \frac{\eta M_i}{n}\|x - y\|_2^2$, and where $\mu := \frac{\eta}{n}\big(1 - \frac{\eta M_i}{n}\big)$. The condition $\eta M_i < n$ of Assumption 6 ensures $\mu > 0$. This strong monotonicity of $-\nabla U$ is equivalent to μ-strong concavity of U. (When v_i is twice differentiable, the same conclusion reads $\nabla^2 U = \frac{\eta^2}{n^2}\nabla^2 v_i - \frac{\eta}{n}\mathbb{I} \preceq -\mu\mathbb{I}$.)

Step 3: the gain from any report is controlled by the gradient at the truthful report. For a μ-strongly concave function, for all x and y, $U(x) \leqslant U(y) + \langle \nabla U(y), x - y\rangle - \frac{\mu}{2}\|x-y\|_2^2$. Maximizing the right-hand side over $x - y$, the maximum of $\langle a, z\rangle - \frac{\mu}{2}\|z\|^2$ being $\|a\|^2/(2\mu)$, attained at $z = a/\mu$, gives

$$U(x) - U(y) \;\leqslant\; \frac{\|\nabla U(y)\|_2^2}{2\mu} \qquad \text{for all } x \in \mathbb{R}^m. \tag{65}$$

We apply Equation (65) with $y = d_i^t$, agent i's truthful report.

Step 4: the gradient at the truthful report. Substitute $x = d_i^t = \nabla v_i(w)$ in Equation (64). Because the other agents are truthful, $d_i^t + b = \sum_{j\in N} \nabla v_j(w) = n\,\tilde{g}^t$, so the model argument becomes $w + \eta\tilde{g}^t = w + \Delta w$, where $\Delta w := \eta\tilde{g}^t$ is the update under truthful reporting. The second term of Equation (64) becomes $-\frac{\eta}{n}\nabla v_i(w)$. Hence

$$\nabla U(d_i^t) = \frac{\eta}{n}\big(\nabla v_i(w + \Delta w) - \nabla v_i(w)\big),$$

and, by the M_i-Lipschitz continuity of ∇v_i,

$$\|\nabla U(d_i^t)\|_2 \leqslant \frac{\eta}{n} M_i \|\Delta w\|_2 = \frac{\eta^2}{n} M_i \|\tilde{g}^t\|_2.$$

Step 5: the per-round bound. Squaring the bound of Step 4 and dividing by $2\mu = \frac{2\eta(n-\eta M_i)}{n^2}$, as Equation (65) prescribes, the factors n^2 cancel and one factor of η cancels, so for every report x

$$U(x) - U(d_i^t) \;\leqslant\; \frac{\eta^4 M_i^2 \|\tilde{g}^t\|_2^2/n^2}{2\eta(n-\eta M_i)/n^2} = \frac{\eta^3 M_i^2}{2(n-\eta M_i)}\|\tilde{g}^t\|_2^2 = \varepsilon_i^t.$$

Step 6: the bound along the truthful path. Along the truthful path, $\tilde{g}^t = \frac{1}{n}\sum_j \nabla v_j(w^t) = -\nabla f(w^t)$ by the definition of f. By the descent lemma for the M-smooth function f (Assumption 3), applied at w^t with step $1/M$, $f(w^{\mathtt{OPT}}) \leqslant f\big(w^t - \frac{1}{M}\nabla f(w^t)\big) \leqslant f(w^t) - \frac{1}{2M}\|\nabla f(w^t)\|_2^2$, where the first inequality holds because $w^{\mathtt{OPT}}$ minimizes f. Rearranging,

$$\|\tilde{g}^t\|_2^2 = \|\nabla f(w^t)\|_2^2 \leqslant 2M\big(f(w^t) - f(w^{\mathtt{OPT}})\big).$$

Under truthful reports Phase 2 is gradient descent on f with step $\eta = 1/M$, so the Phase 2 argument in the proof of Theorem 2 gives $f(w^t) - f(w^{\mathtt{OPT}}) \leqslant (1 - \nu/M)^t\big(f(w^0) - f(w^{\mathtt{OPT}})\big)$; Phase 1 does not move the model, so Phase 2 starts at w^0. Substituting $\eta = 1/M$ into ε_i^t gives $\varepsilon_i^t = \frac{M_i^2\|\tilde{g}^t\|_2^2}{2M^2(nM - M_i)}$, and combining the last two displays,

$$\varepsilon_i^t \;\leqslant\; \frac{M_i^2\big(f(w^0) - f(w^{\mathtt{OPT}})\big)}{M(nM - M_i)}\Big(1 - \frac{\nu}{M}\Big)^t = C_i\Big(1 - \frac{\nu}{M}\Big)^t.$$

Since $0 < \nu \leqslant M$, the ratio $1 - \nu/M$ lies in $[0, 1)$, and the geometric series satisfies $\sum_{t\geqslant 0}(1 - \nu/M)^t = M/\nu$. Summing the per-round bounds over the rounds of Phase 2 therefore gives $\sum_{t<T_0}\varepsilon_i^t \leqslant C_i M/\nu = \frac{M_i^2(f(w^0) - f(w^{\mathtt{OPT}}))}{\nu(nM - M_i)}$. This total is the sum of the gains from one-shot deviations along the truthful path; a deviation in round t changes the later iterates, so the sum does not bound the gain from deviating in several rounds.

Part 4 (optimal parameter). *Truthful reports.* If every agent reports $\nabla v_j(w^t)$, the update is $w^{t+1} = w^t + \frac{\eta}{n}\sum_j \nabla v_j(w^t) = w^t - \eta\nabla f(w^t)$, which is gradient descent on f with $\eta = 1/M$. The claim is then the Phase 2 part of Theorem 2.

Equilibrium reports: characterization. Fix a round t and arbitrary reports $\hat{d}_{-i}^t$ of the other agents. The argument of Part 3, Step 2, did not use the truthfulness of the other agents, since b only shifts the argument of v_i; hence agent i's round utility is μ-strongly concave in its report for any $\hat{d}_{-i}^t$. A strongly concave differentiable function has a unique maximizer, characterized by a zero gradient. By Equation (64), agent i's best response is therefore the unique x with

$$x = \nabla v_i(w^{t+1}), \qquad w^{t+1} = w^t + \frac{\eta}{n}\Big(x + \sum_{j\neq i}\hat{d}_j^t\Big). \tag{66}$$

A report profile $\hat{d}^t$ is a Nash equilibrium of the round if and only if Equation (66) holds for every agent, that is, $\hat{d}_i^t = \nabla v_i(w^{t+1})$ for all i, with $w^{t+1} = w^t + \frac{\eta}{n}\sum_j \hat{d}_j^t$.

Equilibrium reports: reduction to a proximal step. We show that equilibria correspond one-to-one to solutions u of the equation

$$u = w^t - \eta \nabla f(u). \tag{67}$$

If $\hat{d}^t$ is an equilibrium, substituting $\hat{d}_j^t = \nabla v_j(w^{t+1})$ into the update gives $w^{t+1} = w^t + \frac{\eta}{n}\sum_j \nabla v_j(w^{t+1}) = w^t - \eta \nabla f(w^{t+1})$, so $u = w^{t+1}$ solves Equation (67). Conversely, if u solves Equation (67), set $\hat{d}_j^t := \nabla v_j(u)$ for every j; then $w^t + \frac{\eta}{n}\sum_j \hat{d}_j^t = w^t - \eta\nabla f(u) = u$, so the update returns $w^{t+1} = u$, and every agent's report equals $\nabla v_j(w^{t+1})$, which is the equilibrium condition. Equation Equation (67) is the first-order optimality condition of

$$\min_{u \in \mathbb{R}^m} \ \Phi(u), \qquad \Phi(u) := f(u) + \frac{1}{2\eta}\|u - w^t\|_2^2,$$

since $\nabla\Phi(u) = \nabla f(u) + \frac{1}{\eta}(u - w^t)$. By Assumption 3, f is ν-strongly convex, so Φ is $(\nu + \frac{1}{\eta})$-strongly convex. A strongly convex differentiable function has exactly one minimizer, and for a convex differentiable function the stationary points are exactly the minimizers. Hence Equation (67) has exactly one solution, the round has exactly one equilibrium, and the equilibrium update is the proximal-point step $w^{t+1} = \arg\min_u \Phi(u)$.

Equilibrium reports: rate. Since $w^{\texttt{OPT}}$ minimizes f, $\nabla f(w^{\texttt{OPT}}) = 0$, and Equation (67) with $u = w^{t+1}$ can be written as

$$w^{t+1} - w^{\texttt{OPT}} = (w^t - w^{\texttt{OPT}}) - \eta\big(\nabla f(w^{t+1}) - \nabla f(w^{\texttt{OPT}})\big).$$

Taking the inner product of both sides with $e := w^{t+1} - w^{\texttt{OPT}}$ gives $\|e\|_2^2 = \langle w^t - w^{\texttt{OPT}}, e\rangle - \eta\langle \nabla f(w^{t+1}) - \nabla f(w^{\texttt{OPT}}), e\rangle$. By ν-strong convexity, the gradient of f is strongly monotone, $\langle \nabla f(u) - \nabla f(u'), u - u'\rangle \geqslant \nu\|u - u'\|_2^2$, so the last inner product is at least $\nu\|e\|_2^2$. Moving it to the left-hand side and applying the Cauchy–Schwarz inequality to the remaining inner product,

$$(1 + \eta\nu)\|e\|_2^2 \ \leqslant\ \langle w^t - w^{\texttt{OPT}}, e\rangle \ \leqslant\ \|w^t - w^{\texttt{OPT}}\|_2\,\|e\|_2.$$

If $e = 0$ there is nothing to prove; otherwise dividing by $\|e\|_2$ gives $\|w^{t+1} - w^{\texttt{OPT}}\|_2 \leqslant (1 + \eta\nu)^{-1}\|w^t - w^{\texttt{OPT}}\|_2$, and iterating from w^0 gives $\|w^t - w^{\texttt{OPT}}\|_2 \leqslant (1 + \eta\nu)^{-t}\|w^0 - w^{\texttt{OPT}}\|_2$.

Part 5 (budget). *p^1 is budget balanced for every report profile.* Summing Equation (23) over i,

$$\sum_{i\in N} p_i^1(\hat{s}) = \beta\sum_{i\in N}\hat{s}_i - \frac{\beta}{n-1}\sum_{i\in N}\sum_{j\neq i}\hat{s}_j = \beta\sum_{i\in N}\hat{s}_i - \frac{\beta}{n-1}(n-1)\sum_{j\in N}\hat{s}_j = 0,$$

where the middle equality exchanges the order of summation: each $\hat{s}_j$ appears in the inner sum of the $n-1$ agents $i \neq j$.

p^3 is budget balanced in every round. Write $Q := \sum_{j\in N}\|\hat{d}_j^t\|_2^2$. Summing Equation (25) over i, the first terms give $-\frac{\eta}{2n}Q$, and by the same exchange of summation the rebates give $\frac{\eta}{2n(n-1)}(n-1)Q = \frac{\eta}{2n}Q$. Hence $\sum_i p_i^3(\hat{d}^t) = 0$ for every report profile.

p^2 is non-positive, and zero on the truthful path. In Equation (24), the quadratic term is non-positive, and the cross-check term is non-positive because $\chi_i \geqslant 0$, $\hat{s}_i \geqslant 0$ for reports in S_i, $\beta > 0$ and $K > 0$. Hence $p_i^2 \leqslant 0$ for every agent and every report profile, and $\sum_i(p_i^1 + p_i^2) = \sum_i p_i^2 \leqslant 0$. On the truthful path both terms vanish, as shown in Step 3 of Part 1, so $p^2 \equiv 0$ there.

Conclusion. The total transfer to the agents is $\sum_i\big(p_i^1 + p_i^2\big) + \sum_{t<T_0}\sum_i p_i^3$. By the three facts above, it is zero on the truthful path and non-positive for every report profile, so the center never runs a deficit. □

C Adversarial agents

Our results for `2P-UPBReD` assume that agents are honest and report their gradients truthfully. However, the agents may be adversaries who misreport their gradients. Prior work [55] describes some simple ways to deal with this. When the ratio of adversarial agents in the system is low, we use `2P-UPBReD` along with the trimmed mean algorithm of Yin et al. [55].

Algorithm 5: Trimmed mean for `2P-UPBReD` (trim-`2P-UPBReD`)

Input: Step sizes γ, η; initialization w^0, $s_i^0 \in S_i$ for $i \in N$; iteration counts κ, T_0 as given by Theorem 2; trimming fraction β^{trim}

Output: w^{T_0}

Set $b = \lfloor n\beta^{trim} \rfloor$

`// Phase 1: contributions, with the model fixed at` w^0 `(identical to 2P-UPBReD)`

Center broadcasts w^0, s^0; set $k = 0$

while $k < \kappa$ **do**

  Center broadcasts s^k

  foreach ***agent*** $i \in N$ *in parallel* **do**

    $s_i^{k+1} = s_i^k + \gamma\,[g(w^0, s^k, \mu^k)]_i$; send s_i^{k+1} to center

  $k \leftarrow k + 1$

`// Under Equation (12),` $s^\kappa = s^{\max}$

`// Phase 2: model learning at` $s^{\max}$ `with trimmed-mean aggregation`

for $t = 0$ **to** $T_0 - 1$ **do**

  Center broadcasts w^t

  foreach ***agent*** $i \in N$ *in parallel* **do**

    Compute $d_i^{t+1} = \nabla_w v_i(w^t, s^{\max})$ (adversarial agents may send arbitrary d_i^{t+1}); send d_i^{t+1} to center

  foreach *coordinate* $j \in [m]$ **do**

    Sort $\{d_{ij}^{t+1}\}_{i \in N}$ in ascending order as $c_{(1)} \leq \cdots \leq c_{(n)}$

    $w_j^{t+1} = w_j^t + \frac{\eta}{n-2b} \sum_{r=b+1}^{n-b} c_{(r)}$

return w^{T_0}

Let α^{adv} be the fraction of adversarial agents in the system. Phase 1 of the algorithm remains the same with agents first converging to $s^{\max}$. In the second phase, which is a classic FedAvg algorithm, the agents only report gradients computed on their training data. Instead of aggregating all gradients, for each coordinate of the gradients, the values of the gradient corresponding to that coordinate are sorted. The lowest and highest α^{adv} fraction of these values are discarded and the remaining gradient coordinates are used to update the parameter. We detail the algorithm in Algorithm 5. In Appendix D.6 we show that this algorithm is robust to adversaries when the fraction of adversarial agents is small.

D Appendix to Section 6

D.1 Social welfare results using test accuracy

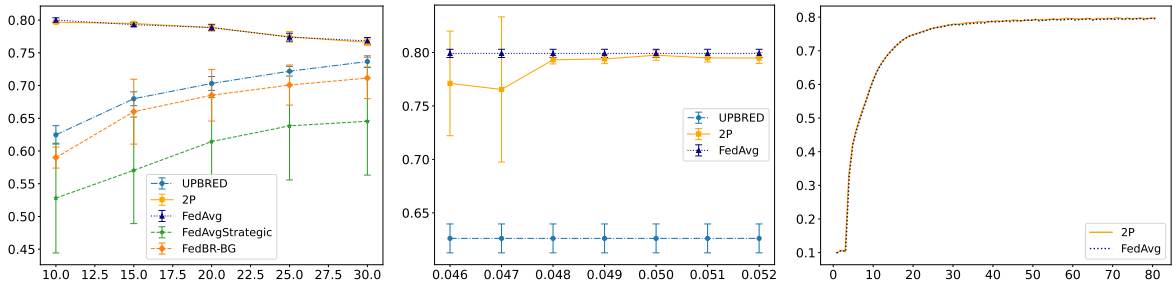
D.2 Algorithm for `FedAvgStrategic`

In the `FedAvgStrategic` algorithm, the agents choose to contribute their data strategically as in `UPBReD`, and stick to this contribution in the entire training phase of `FedAvg`. `FedAvgStrategic` is formally given by Algorithm 6.

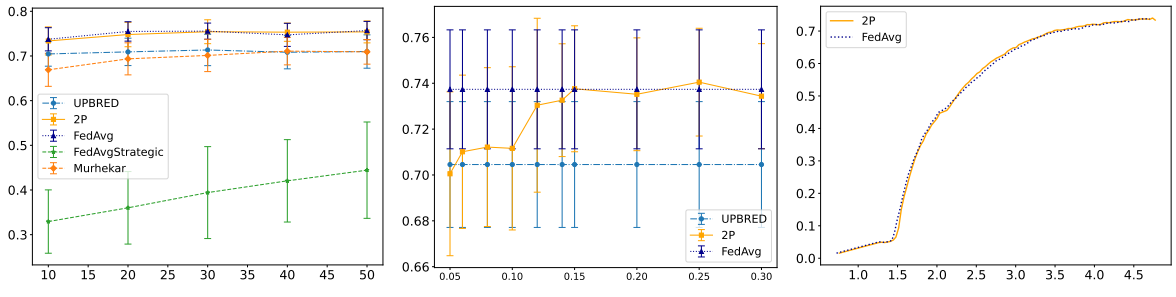
D.3 Additional experiments run using the Flower framework

Using the Flower framework, we partitioned the CIFAR-10 dataset into three regimes: IID, Dirichlet (concentration parameter $\alpha^D = 0.1$, self-balanced), and Pathological (3 classes per partition, random assignment). Distribution profiles for Pathological and Dirichlet are shown in Figure 5. Hyperparameter details: $\beta = 2, \gamma = 1, \eta = 0.0001$, optimizer=Adam. We compare the algorithms `UPBReD`, `2P-UPBReD`, `FedAvgStrategic` and `FedAvg`. The results are summarized in Figure 4, averaged across 5 runs.

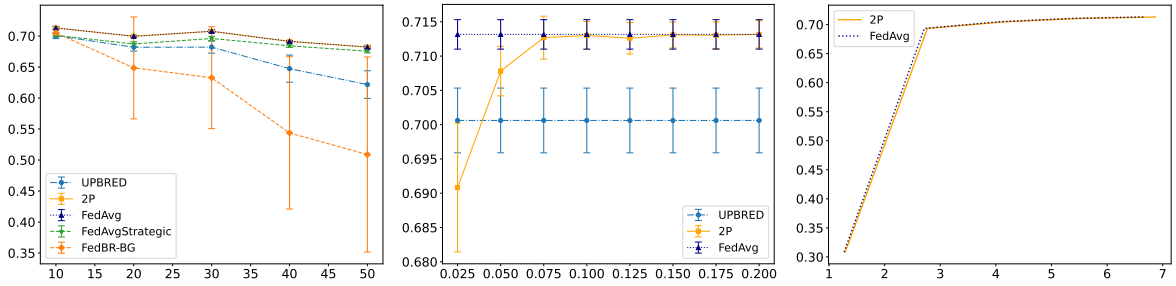
In Appendix D.6 we show that `2P-UPBReD` with trimmed mean (described in Appendix C) is robust to adversarial agents.



(a) *CIFAR-10*



(b) *FEMNIST*



(c) *Twitter*

Figure 3: *The three rows in this figure correspond to the experiments with the datasets CIFAR-10, FEMNIST, and Twitter, respectively. The* `y-axis` *in every plot is the social welfare (computed using test accuracy for each agent). The* `x-axis` *corresponds to the number of agents in the first plot, the choice of β in the second, and the time taken (in minutes) in the third plot, respectively, in every row.*

D.4 Utilities and payments for agents

We also note that the data contributors get paid and data consumers pay in `2P-UPBReD`, and their payments are monotone non-decreasing with data contribution (Figure 6). The utilities vary because the dataset on which the model is tested to obtain the accuracy varies for the agents.

D.5 Model architecture

We provide details of the model architecture used in our experiments in Tables 2 and 3.

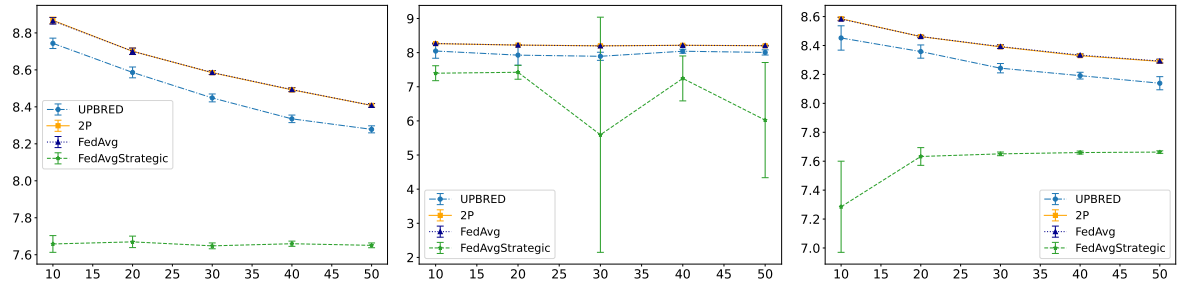
D.6 Experiments with adversarial agents

In Appendix C, we describe the trimmed mean algorithm for `2P-UPBReD` designed to be robust in the presence of adversarial agents. In this section, we demonstrate this robustness on the FEMNIST and CIFAR-10 datasets. We vary the adversarial fraction $\alpha^{adv} \in \{0.025, 0.05, 0.075, 0.1, 0.125, 0.15, 0.175, 0.2\}$, selecting the trimming parameter $\beta^{trim} = \alpha^{adv}$. Figure 7 shows the ratio of social welfare with and without adversaries for `UPBReD`, `2P-UPBReD` (both without trimming) and trimmed mean `2P-UPBReD`. While `UPBReD` and `2P-UPBReD` without trimming suffer welfare loss under adversaries, applying the trimmed mean algorithm (Algorithm 5) keeps the welfare ratio close to 1 for trimmed mean `2P-UPBReD`. Adversarial agents transmit random updates generated from a standard multivariate Gaussian distribution.

Algorithm 6: `FedAvgStrategic`

Input: Step size γ, η, initialization w^0, s_i^0 for $i \in N$, number of iterations T, and a tolerance ϵ
Output: w^T

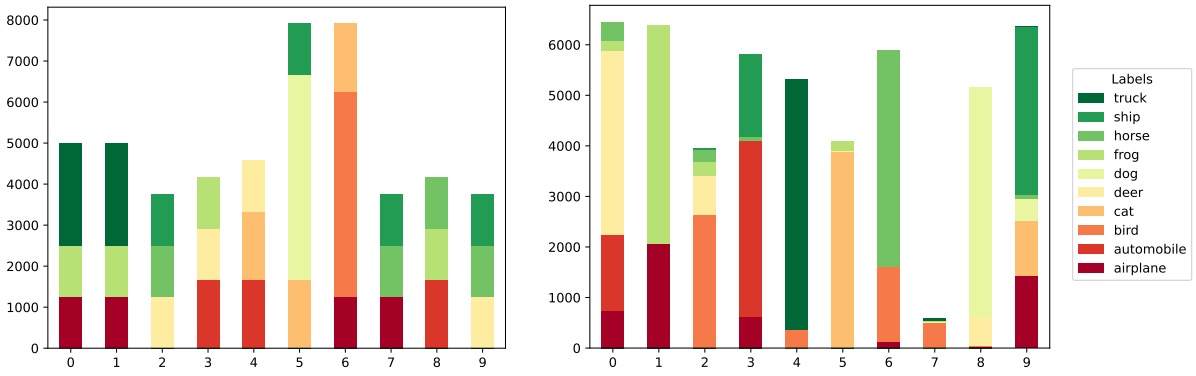
Center broadcasts w^0, s^0, set $t = 0$
while $\exists i \in N$ *such that* $\|[g(w^0, s^t, \mu^t)]_i\| \geq \epsilon$ **do**
  Center broadcasts s^t
  for ***agent*** $i \in N$ *in parallel* **do**
    $s_i^{t+1} = s_i^t + \gamma[g(w^0, s^t, \mu^t)]_i$
    Send s_i^{t+1} to the center
  $t = t + 1$
$s^* = s^t$
for $t = 0$ ***to*** T**do**
  Center broadcasts w^t
  for ***agent*** $i \in N$ *in parallel* **do**
    Compute local gradient: $d_i^{t+1} = \nabla_w v_i(w^t, s_i^*, s_{-i}^*)$
    Send d_i^{t+1} to the center
  Center updates: $w^{t+1} = w^t + \frac{\eta}{n} \sum_{i \in N} d_i^{t+1} = w^t + \eta \tilde{g}(w^t, s^*)$
return w^T



(a) *IID Split* **(b)** *Dirichlet Split* **(c)** *Pathological Split*

Figure 4: *Social welfare vs. number of clients for different CIFAR-10 dataset splits using the Flower framework.*

For a model with m parameters, the adversarial gradient vector is drawn as $g \sim \mathcal{N}(0, \mathbf{I}_m)$, where each coordinate is sampled independently from $\mathcal{N}(0, 1)$.



(a) *Pathological partitioner (3 classes per client).* **(b)** *Dirichlet partitioner ($\alpha = 0.1$).*

Figure 5: *Label distributions for different data partitioning strategies using Flower with* 10 *agents. The* `x-axis` *represents the client ID and the* `y-axis` *represents the number of data points of the given label.*

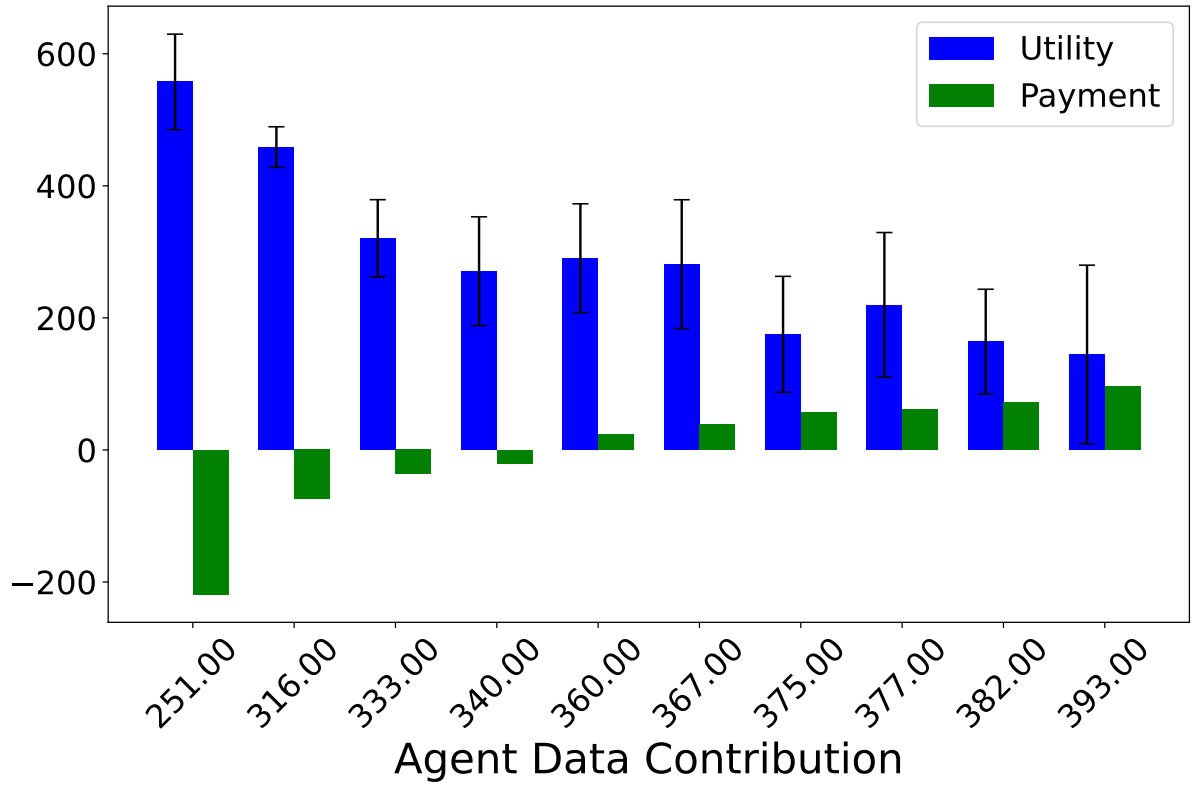


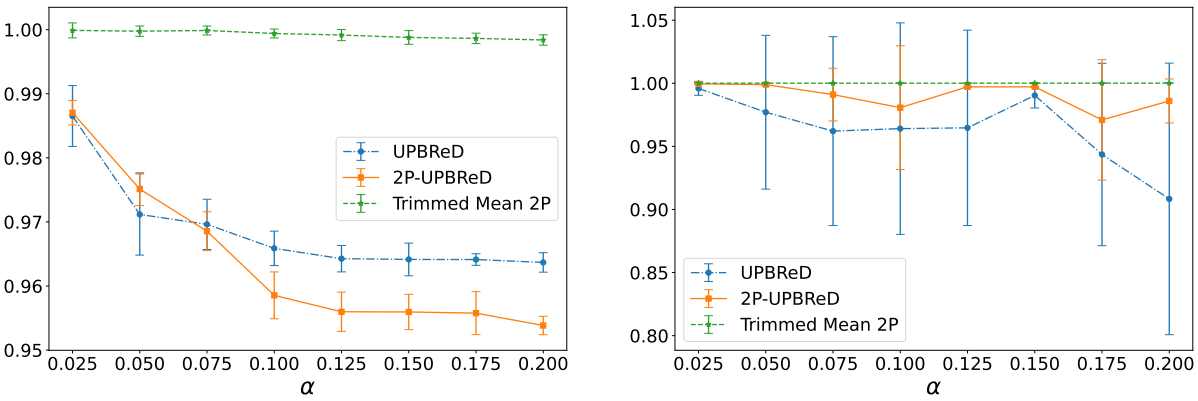
Figure 6: *Utilities and payments received by agents for* `2P-UPBReD` *based on their data contributions for FEMNIST (averaged over 10 runs).*

Layer (type:depth-idx)	Output Shape	Param #
Cifar10ResNet	[1, 10]	–
Conv2d: 1-1	[1, 64, 32, 32]	1,728
BatchNorm2d: 1-2	[1, 64, 32, 32]	128
ReLU: 1-3	[1, 64, 32, 32]	–
Identity: 1-4	[1, 64, 32, 32]	–
Sequential (layer1): 1-5	[1, 64, 32, 32]	147,968
Sequential (layer2): 1-6	[1, 128, 16, 16]	525,568
Sequential (layer3): 1-7	[1, 256, 8, 8]	2,099,712
Sequential (layer4): 1-8	[1, 512, 4, 4]	8,393,728
AdaptiveAvgPool2d: 1-9	[1, 512, 1, 1]	–
Linear: 1-10	[1, 10]	5,130
Total Params		**11,173,962**
Trainable Params		**11,173,962**
Non-trainable Params		**0**
Total Mult-Adds (GFLOPs)		**0.56**
Input Size (MB)		**0.01**
Forward/Backward Pass Size (MB)		**5.84**
Params Size (MB)		**42.63**
Estimated Total Size (MB)		**48.48**

Table 2: *Layer-wise Summary of Cifar10ResNet for CIFAR-10*

Layer (type:depth-idx)	Output Shape	Param #
FemnistCNN	[1, 62]	–
Conv2d: 1-1	[1, 32, 28, 28]	832
MaxPool2d: 1-2	[1, 32, 14, 14]	–
Conv2d: 1-3	[1, 64, 14, 14]	51,264
MaxPool2d: 1-4	[1, 64, 7, 7]	–
Linear: 1-5	[1, 2048]	6,424,576
Linear: 1-6	[1, 62]	127,038
Total Params		**6,603,710**
Trainable Params		**6,603,710**
Non-trainable Params		**0**
Total Mult-Adds (M)		**17.21**
Input Size (MB)		**0.00**
Forward/Backward Pass Size (MB)		**0.39**
Params Size (MB)		**26.41**
Estimated Total Size (MB)		**26.81**

Table 3: *Layer-wise Summary of the CNN for FEMNIST*



(a) *CIFAR-10* **(b)** *FEMNIST*

Figure 7: *Ratio of social welfare with adversaries and without adversaries.*